UNIVERSITY OF CENTRAL OKLAHOMA

Edmond, Oklahoma

Jackson College of Graduate Studies

Cultivating Success in Forensic Molecular Biology Students by Developing Laboratory Skills

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements

For the degree of

MASTER OF SCIENCE IN FORENSIC SCIENCE

By

Claire Joyce

Edmond, Oklahoma

2019

Cultivating Success in

Forensic Molecular Biology Students

by Developing Laboratory Skills

By: Claire Joyce

University of Central Oklahoma

A THESIS

APPROVED FOR THE W. ROGER WEBB FORENSIC SCIENCE INSTITUTE

December 6, 2019

By _____

Dr. James Creecy Committee Chair

By _____

Dr. Mark McCoy Committee Member

By _____

Dr. Brandt Cassidy Committee Member

Acknowledgements

To my husband, Nick, for his love and support and for never letting me give up.

To Dr. James Creecy, for his patience, advice, and support. Thank you for your words of wisdom and for holding me to a higher standard than I had previously experienced in my academic career.

To Dr. Mark McCoy and Dr. Brandt Cassidy for joining my thesis committee and for providing advice.

To all the students and professors at the FSI who have impacted my time here. Not only did this program provide me with a solid education in forensic science, but also helped me to grow as a person and a citizen. For that, I am forever grateful to this school and the FSI.

To my family for all their support. Special thanks to my parents, Ricky and Jo, who have always supported me, and never expected anything in return., and to Brianna and Colin for providing advice, sympathy, and feedback.

"Knowing is not enough, we must apply. Willing is not enough, we must do." – Bruce Lee

Abstract

The University of Central Oklahoma's Forensic Science Institute offers a series of courses for students pursuing a career as a forensic DNA analyst, providing students with the knowledge and experience to prepare them for such a career. One of the courses in the Forensic Molecular Biology series is Advanced Forensic DNA Analysis. This project creates a laboratory component for the Advanced Forensic DNA Analysis course. Addition of a laboratory course will further prepare Forensic Molecular Biology students for the real-world application of methods learned in the course. The laboratory involves experiences working with DNA profiles that vary in composition and complexity to simulate real-life casework. The lab also implements the use of statistical software for DNA analysis of both single source and mixed profiles. The course allows students to develop a working knowledge of interpreting various DNA profiles with a range of complexity, assigning statistical weight to their interpretation, designing standard operating procedures for methodology, and the validation process. The need to implement a laboratory alongside the course was identified through discussion with students and professors, as well as by comparison to similar programs at other universities. The laboratory course is designed with research-based educational methods to maximize student learning and course effectiveness. Professors of the course are provided with a learning framework that is adaptable to continuous improvement of the course to accommodate developments in the field and differing student needs. This project provides students with a greater foundation for success in the field of Forensic DNA Analysis.

Table of Contents

List of Figures

Introduction

Knowledge is valuable, but that value is difficult to measure without the experience of application. Learning is a unique and individual process. Countless studies have determined that students have different learning styles, so utilizing various methods of learning during classroom time is more effective than using a single approach (O'Connor, 2015). In particular, the addition of a "hands-on" experiential approach to a traditional lecture often benefits students (S. Wurdinger & Allison, 2017). The application of learned knowledge to real world situations also gives students in higher education a foundation for what they may be doing in a future workplace. This type of hands-on learning is best presented in a laboratory component accompanying a lecture course and is seen in most Science, Technology, Engineering, and Mathematics (STEM) fields (Council, 2003). Forensic science, being an applied science, involves extensive amounts of hands-on work which prepares students for actual casework in forensic disciplines. Therefore, providing a laboratory setting in which students have the opportunity to develop and practice hands-on techniques is highly beneficial to student learning.

The University of Central Oklahoma (UCO) is home to the W. Roger Webb Forensic Science Institute (FSI), which maintains accreditation for Bachelor of Science in Forensic Science, Digital Forensics, Forensic Chemistry, and Forensic Molecular Biology, making it one of the most comprehensive forensic science programs in the country (Commission, 2019). The FSI trains students in various forensic science disciplines within the above degree categories, including digital evidence analysis, drug chemistry, impression evidence (fingerprint analysis), trace evidence, and forensic molecular biology (DNA analysis). Courses in the subject of forensic DNA analysis currently allow students to practice the process of evidence handling and

DNA extraction. However, forensic molecular biology students would further benefit from opportunities to apply the statistics or advanced computational methods used in a working forensic DNA laboratory. This project advocates for the addition of a laboratory course to UCO's Advanced Forensic DNA Analysis course and produces the teaching tools and materials necessary to administer the course. The new laboratory is specifically designed to guide students to develop the higher knowledge and skills that are relevant to a career as a forensic DNA analyst. The course is designed using research-based educational methods and is justified by an exploration of the body of literature regarding all STEM education but especially forensic educational practices.

Science education theory comprises a long-standing, constantly developing body of work which guides educators to current research-based methodologies. Research-based methods use evidence to establish best practices in science education which, when followed, ensure that students receive the best possible education. Educational technology and learning theory tell us that it is the responsibility of teachers to facilitate learning (Robinson, Molenda, & Rezabek, 2008). Learning is facilitated by providing students with the tools, resources, and environment to promote learning. Various educational philosophies have been developed and published as "educational theories" or "learning theories" which assist educators in developing effective curricula. Educational theories focus on the ways that students best learn certain types of material and may suggest concrete methods to facilitate learning. The most commonly encountered educational theories are behaviorism, cognitivism, and constructivism (Robinson et al., 2008). The complex nature of the material taught in the advanced forensic DNA analysis course would best adapt to the cognitivist and constructivist learning theory principles.

Cognitivism in educational theory defines a set of educational frameworks which focus on mental structures and processes as a foundation for learning (Januszewski & Molenda, 2008). Cognitivist frameworks are highly structured and easily adapt to use in scientific education. Constructivism as an educational theory indicates educational frameworks which allow students to "construct" their own knowledge (Driscoll, 2005). Students are presented with ideas and then given opportunities to explore the ideas independently or in groups to form knowledge. Constructivism is often used in courses where the students already have the foundations of knowledge and are working on more advanced concepts, such as in the Advanced Forensic DNA Analysis course discussed in this project. However, constructivism may give students too much freedom in a course where accuracy and standard procedure are of utmost importance. Therefore, in the proposed course constructivism is subordinate to cognitivism. The use of cognitivism and constructivism will best facilitate student learning in the Advanced Forensic DNA Analysis course. These educational theories are discussed in greater detail in a later section and applied throughout the Advanced Forensic DNA Analysis Laboratory.

One of the primary purposes of education is to give students the tools to be useful to society after graduation (Kivunja, 2014). Pragmatically, the principles learned in forensic DNA analysis are not useful until applied in a real or simulated case. Practical exercises may be developed and performed in the laboratory setting to provide students with an understanding of how to apply their knowledge. However, student success is ultimately dependent on graduating and applying their knowledge in the workplace as a forensic DNA analyst. Advancing a career in the field requires working knowledge and experience with techniques. Upon beginning a forensic DNA analyst job, laboratories require a minimum amount of on-the-job training to ensure competency of the analyst ("Quality Assurance Standards for Forensic DNA Testing

Laboratories," 2011). This training period may be longer or shorter depending on the readiness of the analyst. A higher level of prior knowledge or experience would help to decrease the length of this training to the minimum requirement, which is valuable to employers. Therefore, the addition of the proposed laboratory course gives students an advantage over potential competitors entering the forensic workforce. Facilitating the path to a successful career as a forensic scientist is the ultimate goal of the forensic science program for most of the students in the program. Continuous improvement of the program through projects such as this one will assist students in meeting their goals and becoming successful post-graduation.

Statement of Problem

The need for a laboratory for the Advanced Forensic DNA Analysis course is readily apparent to students in the FSI Forensic Molecular Biology program. Other courses leading up to the Advanced Forensic DNA Analysis course all have laboratory components, including the Forensic Molecular Biology and Forensic Serology courses. Students have expressed the desire for a laboratory through course evaluations. Professors who have taught the course in the past have also voiced concern over the lack of a laboratory component. Addition of a laboratory course brings this course up to the standard set by other courses within the program and meets the needs of students and professors in the program.

The UCO FSI Mission Statement states:

> *The Forensic Science Institute is devoted to academic excellence, through a unique multidisciplinary program, that provides outstanding educational, research, and professional training opportunities for practicing professionals and both undergraduate and graduate students. The Forensic Science Institute is a comprehensive training and*

> *research organization in all aspects of evidence collection, preservation, analysis, reporting and testimony ("Forensic Science Institute," 2019).*

For the Advanced Forensic DNA Analysis course to fulfill the requirement of a comprehensive training program, it must provide students with the opportunity for practical experience and application of concepts. Addition of a laboratory component that focuses on the professional training and practical application of knowledge both fulfills the stated requirements and benefits students' education and future work endeavors. By meeting the goals outlined in the mission statement, the revised course better represents the FSI program to both potential students and employers.

Project Statement and Objectives

Both graduate and undergraduate students interested in DNA analysis are provided with a laboratory course curriculum. The curriculum includes a syllabus, course schedule, and laboratory manual to guide learning. The laboratory course is designed using the modern educational theories of cognitivism and constructivism and implements research-based educational methods such as formative assessments and active learning. Professors are provided with a research-based framework for facilitating student learning of advanced DNA analysis techniques. The proposed course improvements allow students to apply the concepts learned in lecture to scenarios which may be encountered in a future workplace. Students will emerge from the Advanced Forensic DNA Analysis course with the skills and experience necessary for a successful career in forensic DNA analysis.

Scope of the Proposed Course

Students and professionals enrolled in courses at the Forensic Science Institute are seeking to gain the knowledge and skills which would be directly applicable to working in a forensic investigation setting. Students may customize the program to their specific interests by taking classes on a variety of forensic topics, including impression evidence, drug chemistry, trace evidence, and biological evidence. The programs offered by the Forensic Science Institute include a Forensic Science Education Programs Accreditation Commission (FEPAC) accredited Bachelor of Science and Master of Science in Forensic Science – Molecular Biology. Either of these programs require the completion of two sequential courses titled Forensic Molecular Biology and Advanced Forensic DNA Analysis. Forensic Molecular Biology, the first course in this series, comprises the basics of DNA evidence collection and chemical preparation of DNA samples. The preliminary course includes a laboratory component where students practice the techniques they learn about in lecture. The second course, Advanced Forensic DNA Analysis, comprises the actual analysis of a forensic DNA profile, from receiving an electropherogram to presenting a conclusion to a jury. Currently, this course does not involve a laboratory component, forcing instructors to instead use limited classroom time to teach, practice, and improve techniques.

The complexity of DNA analysis requires students to learn and understand many various logical and mathematical techniques. The Advanced DNA Analysis course teaches students the following general ideas, skills, and techniques: Interpretation of Single Source DNA profiles and Mixed Source DNA profiles, validation of methods used in a Forensic DNA lab, identification of artifacts associated with DNA analysis, use of thresholds in DNA analysis, familiarity with software used in DNA analysis including expert systems, statistical spreadsheets, and

probabilistic genotyping systems, performing statistical calculations to the determine the weight of evidence for mixed or single source profiles, interpreting and writing standard operating procedures for any of the techniques learned in the course, and presenting results in a court setting. The laboratory curriculum presented in this project will include practice in all the above concepts. In addition to building on knowledge gained in the pre-requisite Forensic Molecular Biology course, the skills students learn in the Advanced Forensic DNA Analysis laboratory are directly applicable to casework in a real forensic DNA laboratory.

Design-Based Research

A relatively recent innovation in educational research is development of the concept of design-based research. Design-based research is related to the innovation and implementation of new educational technology through collaboration between researchers and field practitioners to ensure effectiveness and continuous improvement of methods (Amiel & Reeves, 2008). In the design-based research framework, new methods of educational technology go through iterative cycles of refinement and improvement (see Figure 1). Design-based research in the classroom allows educators to try new technologies or educational techniques in courses to gauge effectiveness and continuously improve courses.

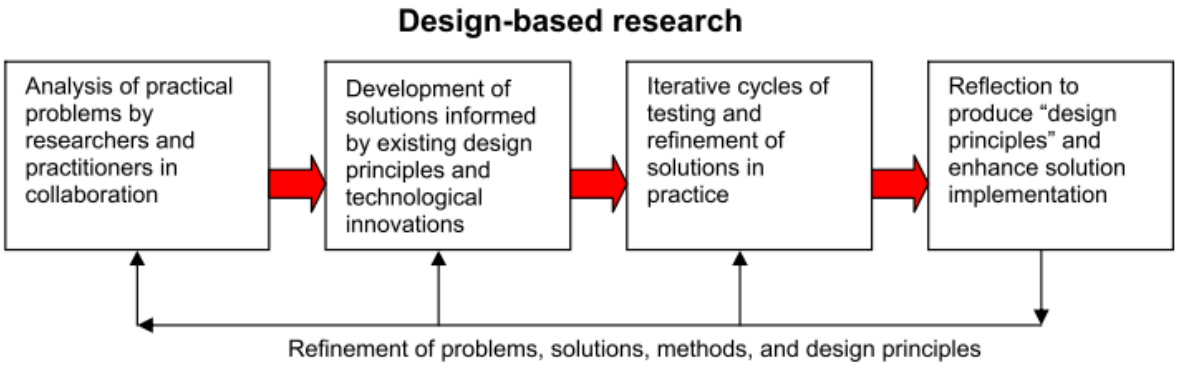


Figure 1: The Design-based research iterative process. Amiel & Reeves, 2008

While this project does not propose a new educational technology, the proposed course is itself innovative, and has the potential for refinement with each iteration of the course through student feedback, professor experience, and experts and innovations in the field. The design-based research framework allows professors to continuously improve their teaching methods each semester the class is taught. This continuous development is especially important in the context of applied science courses, in which the course must evolve to keep up with the industry standard techniques. Forensic DNA analysis is a field of constantly developing science and methodology, and therefore professors will find it necessary to utilize the design-based research framework to keep classroom experiences up to date.

Identifying the Need for Laboratory Experience

In the field of forensic DNA analysis, the need for laboratory experience is widely recognized. From students to experts in the field, the benefits of experience are known and desired. Within the FSI itself, student evaluations of the Advanced Forensic DNA Analysis course and the prerequisite Forensic Molecular Biology course reveal student desire for a laboratory course. The Forensic Molecular Biology course currently has a laboratory component where students may extract and produce DNA profiles. In student evaluations obtained through personal communication with the course instructor, inclusion of a laboratory was highly valued. In response to the question "What did you like about this course?" students remarked: "I loved actually working with the techniques and instruments I learned about in lecture," "The lab was typically on topic with the lecture which made it easier to understand," "Doing lab work hands-on instead of just theoretically," and "Lab was ran in a manner a lab in the field operates so gives experience in the use of molecular forensic techniques that are used for casework" . In contrast, evaluations of the Advanced Forensic DNA Analysis course revealed that students felt the need

for more application of concepts. In response to the question “What can the instructor do to improve this course?” students replied: “Please have the software programs mentioned in class to help students understand how to calculate stats and understand likelihood ratio,” “The lab component would be useful if it was more like deconvoluting mixtures and using the probabilistic software, that sort of stuff,” and “Labs for this class that I personally think would be useful include labs to work with actual deconvolution software after learning how to deconvolute profile by hand. Statistical analysis software also after learning statistical analysis by hand”. This feedback from students enrolled in the forensic molecular biology program provides strong evidence that addition of a laboratory to the Advanced Forensic DNA Analysis course would greatly facilitate students achieving their goals.

Research into STEM educational programs has examined the benefits of implementing laboratory courses with hands-on work. A 2013 study from the Loughborough University engineering department (Gibbins & Perkins, 2013) explored the state of laboratory courses in STEM programs. This included a survey of fifty engineering students who participated in laboratory courses. Surveyed students responded that they valued laboratory courses to 1) aid in understanding concepts learned in lecture, 2) learn skills through hands-on activities, and 3) apply theory to real-world problems. Student feedback such as this is an important part of identifying strengths and deficiencies in educational programs. Students reporting the benefits and advantages of laboratory courses is strong evidence of the need for hands-on experience and laboratory courses.

A UK-based study (Welsh & Hannis, 2011) which surveyed forensic science students and employers identified recent forensic science graduates as lacking technical skills and experiences. This led to employers implementing initial training for new hires starting with the

basics of forensic science, which is both costly and time-consuming. Preparing forensic science students for their intended job field requires hands-on learning and experience which can be gained in a laboratory setting. Forensic education programs implementing hands-on learning and experience in real techniques improve upon the deficiencies identified by this study. This project both addresses the identified deficiencies and provides a roadmap for other programs hoping to achieve the same goal.

The 2016 PCAST report ("Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods," 2016) evaluated practices in the field of forensic science for current methods and best practices. The report determined that methods of single-source and simple two-person mixture are valid and well-applied in the field. However, complex mixture analysis, which is a significant component of UCO's Advanced Forensic DNA Analysis course, is less understood in the community and requires improvement. The report recommends the use of probabilistic genotyping, which requires extensive training to be well understood. Analysts must be able to demonstrate their ability to reliably apply methodology, which is stronger if they are able to perform these methods in a classroom/laboratory setting prior to entering the workforce. In addition to probabilistic genotyping programs, students are afforded the opportunity to use an expert system and a calculation spreadsheet for DNA analysis, which further prepares them for future laboratory work. To satisfy the PCAST recommendations, any academic program which teaches forensic DNA analysis should utilize a laboratory component which would allow students to gain experience practicing the techniques they would use in a working lab.

A comparison of UCO's forensic molecular biology program to similar programs at other universities reveals that UCO is behind in giving students laboratory experience in advanced

DNA analysis methods. Virginia Commonwealth University, rated the #1 forensic science university in 2018 (helptostudy.com, 2018), offers an "Advanced Forensic DNA Analysis" course consisting of two hours lecture and three hours laboratory weekly. The course description states that "Students gain individualized, hands-on experience with DNA procedures and instrumentation in the laboratory exercises" (University, 2019). The Forensic Genetics program at the University of North Texas Health Science Center likewise demonstrates a focus on laboratory experiences, with a three-course sequence of "Methods in Forensic Molecular Biology" from Basic to Advanced level. The courses are characterized as "intensive laboratory courses". The Advanced Methods in Forensic/Molecular Genetics course, which most closely resembles the Advanced Forensic DNA Analysis course at UCO, is described as providing "in-depth experience and expertise in performing DNA analysis techniques" to students (Center, 2019). In order to be competitive with other forensic science education programs, the UCO Forensic Molecular Biology program needs to implement a focus on laboratory experiences.

Development of a laboratory course for the Advanced Forensic DNA Analysis course has overwhelming support from students, professors, and the FSI department. Because the need for laboratory experience is identified, educators have a responsibility to provide these critical experiences to students. Students and higher education programs both benefit from the use of research-based methods of laboratory instruction by increasing competitiveness and quality of graduates. Providing students with the tools and experiences they will need to work in a career in their desired discipline increases student success.

Literature Review

Industry Standard Techniques for DNA Analysis and Interpretation

Forensic DNA analysis is a field of continually advancing research and innovation which requires precise and accurate methods of analysis. While no standard methodology is set for the interpretation of DNA evidence, guidelines are provided by government or regulatory entities, and there exist industry standard expectations which are generally followed by laboratories. In order to ensure consistent and correct methodologies across laboratories, a formal education process is needed for DNA analysts. Analysis of forensic DNA evidence is subject to review by legal and forensic experts and may have life or death consequences for defendants involved in legal cases. Therefore, it is critical that DNA analysts are provided with an education that gives them the best possible foundation and experience in the discipline. The general process of DNA analysis outlined below is consistent with material presented in universally highly regarded textbooks such as John Butler's *Advanced Topics in Forensic DNA Typing: Methodology* and *Interpretation* and in professional laboratory procedures.

In the United States, the Federal Bureau of Investigation (FBI) sets a standard for DNA analysis to consist of the examination of 20 core genetic loci and a sex marker (Hares, 2015). While individuals within a population may share the same alleles (or markers) at some loci, the statistical likelihood of two people matching at all loci is exceptionally low. This principle forms the basis for forensic DNA analysis (Bar, Brinkmann, Lincoln, Mayr, & Rossi, 1994). When a DNA sample collected from a person or crime scene is analyzed, the resulting profile may be compared to another. If all loci match, the unknown sample may be probabilistically "identified" as belonging to the person with whom the unknown is compared. In samples which are not significantly degraded, and which come from only one person, this analysis is a relatively

straightforward process. The necessary steps include DNA extraction and quantification, PCR amplification, and capillary electrophoresis. This initial process generates an electropherogram with the alleles at each locus represented by “peaks” of varying height, which correspond to the amount of that particular allele present in the sample (see Figure 2). Loci may be heterozygous, with two different alleles, or homozygous, with two of the same alleles. Therefore, in a single-source sample, loci may contain one or two peaks. Any additional peaks are either artifacts generated in the analysis process, rare tri-allelic results, or from an additional contributor. While additional contributors may complicate the analysis process, they do not necessarily inhibit it. A mixed sample may be manually interpreted or “deconvoluted” if one contributor gave significantly more or less sample than other contributors. Otherwise, probabilistic genotyping software systems may be used to perform this process, with varying degrees of success.

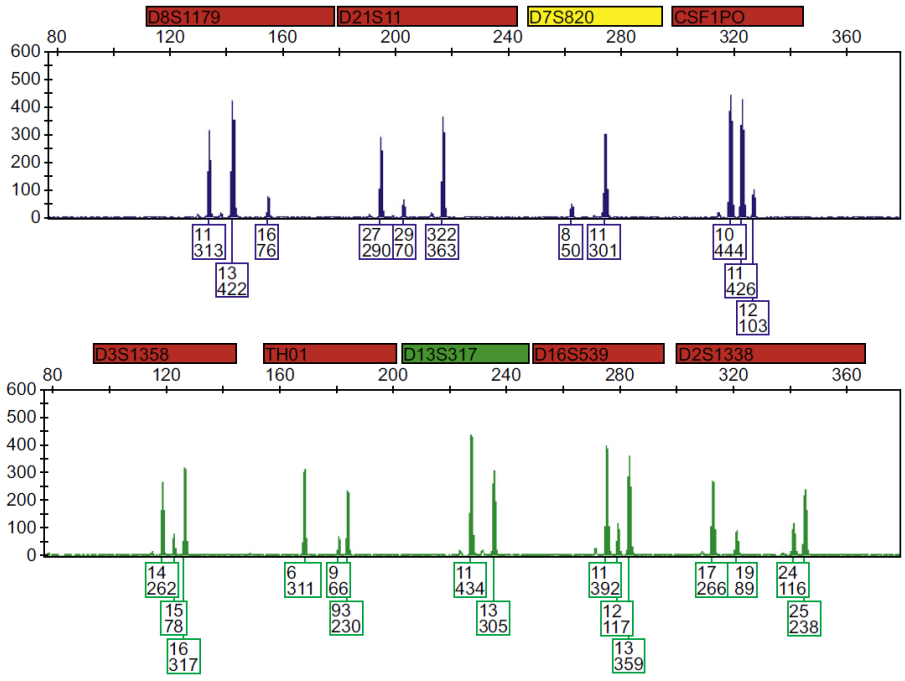


Figure 2: Example electropherogram data, demonstrating loci and allele peaks of a possible mixed-sample. Butler, 2015.

Mixture interpretation is a critical part of the Advanced Forensic DNA Analysis course and will therefore take a large amount of class and laboratory time. In order to deconvolute a mixed sample, the presence of more than one contributor must first be determined by the analyst. Mixed samples may be indicated by several factors, including: significant heterozygote imbalance, high stutter ratio (stutter peaks >15%), and three or more alleles per locus. Once the presence of a mixture is determined, each allele must be identified. The number of contributors, relative amount of each contribution, and each possible genotype combination is then determined. Finally, the sample may be compared to reference samples. In 2010, the Scientific Working Group on DNA Analysis Methods (SWGDAM) published interpretation guidelines (Methods, 2010) which specify that if a major and minor profile may be determined due to a significant difference in contribution to a sample, then one or both profiles may be determined. However, if multiple contributors give similar amounts of sample, then the profiles may be indistinguishable. Even an indistinguishable mixture may still allow for inclusion or exclusion of individuals. Additionally, if an unknown sample is taken directly from an individual (i.e. in cases of sexual assault), and an additional known sample is also taken from that individual, that person's profile may be differentiated during analysis, allowing for discernment of other profiles present in the mixture. Many students entering the course have no prior experience handling DNA mixture analysis, requiring extra time and focus to be spent on the topic. Students' lack of experience in critical topics such as DNA mixture interpretation demonstrate the need for a high-quality laboratory course.

DNA analysis is given real meaning only when a conclusion is formed and the weight of evidence is determined, so these steps are vital to students learning the process of DNA interpretation. After profiles are determined and compared, the DNA analyst must form a

conclusion to their analysis. This conclusion may be an inclusion, exclusion, or a statement that the evidence is not suitable for analysis. If the analysis results in an inclusion, then statistical analysis must be performed to indicate the weight or evidentiary value of the inclusion (Evett & Weir, 1998). There are several statistical methods for this, depending on the type of resultant profile. In general, single-source profiles use the Random Match Probability, indistinguishable mixed profiles use Combined Probability of Inclusion, and mixed profiles use the Likelihood Ratio. These designations should not be confused with requirements as the likelihood ratio calculation may be used on any type of profile, and the RMP may be modified for use with certain mixtures. Analysts should always be aware of their lab's SOP which gives guidance for which calculation to use in which situation. Students participating in the Advanced Forensic DNA Analysis laboratory will gain useful experience in not only forming conclusions to their analysis, but also in understanding and following laboratory SOPs.

Modern technologies used in the field of forensic DNA analysis are implemented in the laboratory curriculum to give students experiences with current techniques in the field. One such technology, probabilistic genotyping system (PGS) software, is a relatively recent addition to the forensic DNA analyst toolbox. These software programs consist of algorithms which determine the possible and most likely DNA profiles from a mixture and provide statistics to back up the data. Use of PGS software programs is recommended by the 2016 PCAST report ("Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods," 2016) as the most objective method of complex DNA mixture interpretation. Because PGS software is new to many laboratories, some professionals in the field may not yet be comfortable using this software. Early exposure to the use and functions of

PGS programs gives students a lead in becoming familiar and comfortable with this type of software.

Upon entering the Advanced Forensic DNA Analysis course, most students are already familiar with many of the above topics. The pre-requisite classes teach students how to handle DNA up to the interpretation steps, and possibly through interpretation of single-source samples; these lessons are review for many students. The more advanced techniques of interpreting complex DNA samples and using new technologies are more challenging for students. Use of modern educational theory to optimize student learning assists students in understanding the more advanced topics.

Application of Learning Theory

Construction of an effective laboratory course requires an exploration of modern educational theory and methodology. Educational and learning theory provides three main avenues to follow for course design and facilitation of learning: behaviorism, cognitivism, and constructivism. While the bodies of research on these three principles are extensive, this project is focused on the application of cognitivism and constructivism to laboratory course design. A blended approach of these two methods has been shown to be successful in higher science education (O'Connor, 2015). Behaviorism, while a valuable educational theory, is better suited to courses of lower complexity. The advanced nature of the Advanced Forensic DNA Analysis course would benefit from the more in-depth character of the cognitivist and constructivist theories (Robinson et al., 2008).

Behaviorism is defined as “the acquisition of new behavior based on environmental conditions” and is highly effective as a technique to train memorization and recall of skills. Behaviorism provides effective strategies for learning foundational knowledge because it

primarily requires memorization of facts. This learning theory uses methods such as classical conditioning to train behavior in students (Robinson et al., 2008). In an advanced level course, behaviorism is less effective because students are required to do more than just memorize and recall facts. In order to developing understanding and skills, students need more advanced techniques to facilitate learning.

Cognitivism is an educational theory largely based on the idea that learning is dependent on brain function. Therefore, sequence and order of learning become an important part of constructing curriculum. Cognitivist researchers have created an array of frameworks for learning. An example of such a framework is Gagne's Events of Instruction, which prescribes the following sequence for instructors: Gain attention by giving reason and purpose, give clear expectations and goals, review background knowledge as a foundation for new knowledge, demonstrate a new skill, guide learners through the content or skill, provide opportunities to practice, provide feedback, test mastery of the skill on problems, and finally develop skill by giving varied/advanced problems (Gagne & Medsker, 1996). In addition to providing guidance to overall course design, this sequence of learning provides a useful format for formation of a laboratory manual to guide student learning. Cognitivist theory provides a high level of structure to learning, making it well suited as a foundation for teaching advanced skills and concepts in a laboratory setting. Studies involving undergraduate chemistry and physics laboratories describe the need for cognitive structures to ensure learning and cognitive growth.

Simonson and Thompson provided cognitivist guidelines for important factors of computer-based learning: 1) the willingness to study is important to the process of learning, 2) the structure and types of knowledge being taught are critical to student understanding, 3) the sequence of presenting materials is important to students' processing information, 4) time and

place, or environment, determine the best teaching style, and 5) learning through exploration and discovery is an effective approach (Simonson & Thompson, 1997). These factors are important to consider when developing classes with a high amount of technical and computer-based learning. While the laboratory developed in this project is not entirely computer-based, it contains several lessons dealing with computer software programs. Therefore, the guidelines given by Simonson and Thompson are used throughout the design process.

Domin in 1999 analyzed undergraduate chemistry lab manuals to evaluate the use of cognitive skill development (Domin, 1999). Cognitive skills are the things which cognitivist learning schemes seek to develop in students and include both higher-order and lower-order skills. Domin's analysis of chemistry lab manuals specifically looked for higher-order skill development, which is appropriate for college-level coursework. Higher-order cognitive skills are developed when learning more complex concepts, and are demonstrated in behaviors like evaluating, inferring, or planning. By carefully analyzing the phrasing of questions and instructions in the manuals, Domin was able to determine when higher-order skills were being developed. In his analysis, Domin reports that the best manuals place more of the instructional burden on the students by making students generate their own procedure or work from a partial one. Additionally, higher-order cognitive learning was facilitated by the use of varied instructional strategies, rather than relying on a single approach. Domin's findings are useful in the formation of a laboratory manual to ensure higher-order cognitive skill development.

Wieman performed analysis of the cognitive tasks required for experimental physics research (Wieman, 2015). The requisite cognitive tasks are: establishing a research goal, defining criteria for suitable evidence, determining feasibility of the experiment, experimental design, construction and testing of instrumentation/apparatus, analyzing data, evaluating results,

determining implications of the results (novel, unexpected, confirmed, etc.), and presenting the work. The study found that most of the required cognitive tasks are performed by instructors when designing the laboratory, leaving little cognitive work for students. Wieman cites this disconnect between instructor and student work as a primary reason for students to be disinterested and frustrated with laboratory work. By increasing the cognitive workload for students enrolled in laboratory courses, student interest and expertise in the subject likewise increases.

Constructivism, like cognitivism, is well suited to advanced concepts which are expected in an advanced course such as Advanced Forensic DNA Analysis. Unlike cognitivism, however, constructivism is less structured, and is rather based on "self-directed learning, discovery learning, practical learning, co-operative learning in groups," (Terhart, 2003). The main application of constructivism in this course is demonstrated in the semester-spanning project of constructing a Standard Operating Procedure (SOP) document outlining the methodologies learned in the course. As a group effort, the project encourages collaboration and peer review. The peer review process, as well as instructor assessment and feedback, ensures that students are correct in their understanding of topics. While the "learner-centered" focus of constructivism is useful to develop student's critical thinking, it must be carefully managed in a field such as forensics, where students must learn methods and topics correctly and precisely. Therefore, the combination of constructivism with highly-structured cognitivism benefits students by ensuring both accurate knowledge and a high level of understanding of the topic.

Writing in 1999, Shiland utilized the theory of constructivism in conjunction with the National Science Education Standards to outline modifications to improve chemistry laboratory courses (Shiland, 1999). Shiland makes the following recommendations for improvement of

laboratory courses: increase cognitive activity of the learner, design labs to identify and correct misconceptions, challenge present knowledge, include group activities, and require students to demonstrate applications of their new knowledge. Shiland additionally provides examples of these improvements, such as students designing their own procedures or making predictions prior to performing the experiment. The described course improvements increase student learning by ensuring that students understand the underlying science to activities they are performing.

Another constructivist course redesign saw the implementation of a virtual chemistry laboratory as early as 2005 (Molphy & Pocknee, 2005). This course is designed following Jonassen's Constructivist Learning Environment Model, which prescribes: real world environments for the relevant learning, realistic approaches to real-world problems, instructor as coach and strategy analyzer, interrelated concepts and diverse perspectives of content, flexible instructional goals, evaluation as feedback and improvement tool, incorporating multimedia tools and environments, and allowing learners to control learning (Jonassen, Mayes, & McAleese, 1993). The course design placed the onus of responsibility for learning on the student, requiring students to manage their own time and learning. Students were provided with resources including video lectures, video tutorials, online discussion forums, tutoring, weekly feedback, and practice quizzes. Despite these resources, students reported difficulties with the lack of physical instructor, lack of participation from other students in the discussion forums, and confusion with the delivery of course material. In 2019, virtual learning environments are much more commonplace than in 2005, so some of these barriers may be mitigated today. However, a pure constructivist approach to a complex scientific subject may create more issues than benefits to students, due to the lack of structure and direction.

Another theory of learning which heavily overlaps with cognitivism and constructivism is contextualism. Contextualism as a learning theory concentrates on the relation of subject matter to real world problems and situations. By relating learning to the real world, students become more engaged with the material and have a greater understanding of the applications of their new knowledge (Hudson & Whisler, 2007). Contextualism is easily used in conjunction with other learning theories to optimize student learning. In an applied science such as forensic DNA analysis, understanding theory in its applied context is of utmost importance for student learning and success.

In practice, many educators use a combination of learning styles to teach their students. Bakar and Zaman developed a virtual laboratory for their chemistry class using the cognitivism-constructivism-contextual approach (Bakar & Zaman, 2008). This three-pronged approach to developing a virtual chemistry lab showed experimental success, with students performing significantly better after using the new lab than the traditional one. Bakar et al. directly measured the effectiveness of their new lab by comparing students before and after taking the chemistry lab, with a control group using conventional laboratory methods, while the experimental group used the new virtual laboratory platform. O'Connor, likewise, advocates a "blended approach" to learning in upper level chemistry classrooms to meet "diversity in learners needs" (O'Connor, 2015). O'Connor describes teaching strategies in higher chemistry education by identifying issues and determining a teaching method to address the issue (see Figure 3). Teachers at all educational levels may improve their teaching methods by using a similar diagnosing strategy to pair learning needs with relevant learning theories to find solutions. The proposed laboratory course utilizes learning theories to address teaching deficiencies, such as implementing cognitivist pre-lab work to ensure students are prepared for the laboratory material.

Issue in Chemistry Education	Learning and Teaching Strategy	Relevant Learning Theory
Student diversity	Creating a VLE with a blend of support material	Constructivism
Variety in Maths & IT ability	Sample answers with accompanying podcasts	Cognitivism
Variety in prior knowledge	Peer to peer teaching, group work	Constructivism/ Cognitivism/ Social constructivism
Lack of preparation time	Pre-lab and pre-lecture work	Cognitivism/Behaviourism
Inconsistency in tutorials	Tutorial workbook/ sample answers for tutors	Behaviourism
Surface learning in lab practicals and in class	Context/ project based labs, workshop style classes (problem based driven)	Constructivism/ Social Constructivism
Student engagement/ motivation	Industry visits, debating, context based learning	Social constructivism

Figure 3: O'Connor's strategy of pairing issues with solutions using relevant learning theory. O'Connor, 2015.

Rieger compares the differing learning theories in the context of adult education and concludes that each theory has a place in learning (Rieger, 2017). The study provides the requisite performance tasks and measures of success for the three theories, as in Figure 4. Rieger's definitions of success for each learning theory are useful for implementing the theories to a course design. Student's enrolled in the Advanced Forensic DNA Analysis laboratory will be able to define individual goals and put theory into practice through utilization of the constructivist and cognitivist learning theories.

Table 2. Practice combination between learning performance / learning success and scientific theories

	Learning Performance	Learning Success
Constructivism	Individual learning material; background, knowledge	Define individual goals
Cognitivism	Learner should find solutions	Put theory into practice
Behaviorism	Conditioning	Focus on improvements

Figure 4: Rieger's association of each learning theory to its measures of performance and success. Rieger, 2017.

The differing benefits of the learning theories show how a combination of the learning theories provides a more comprehensive educational framework. The above studies demonstrate that different learning theories may address different student or course needs and should be applied as they are needed.

Experiential Learning Theory is a model proposed by David Kolb which might be thought of as a synthesis of cognitivism, constructivism, and contextualism. In experiential learning, students learn a topic through the combination of abstract conceptualization and concrete experience, followed by reflective observation and active experimentation to transform the experience to knowledge (Kolb & Kolb, 2005). While lecture or discussion style courses provide opportunities for abstract conceptualization and reflective observation, they leave out the other two modes of learning. A laboratory course creates an environment where concrete experience and active experimentation of a subject are required. Kolb draws on the theories of John Dewey and Carl Jung to explore the variable learning ability and styles of students. Kolb describes the importance of experiential learning to higher education by helping students understand the connection between education, work, and personal development. The goal of experiential learning is teaching learners "how to think" rather than "what to think."

Kolb advocates the use of appropriate learning space to a course. Practical application of scientific theories is best applied in a laboratory setting. Kolb writes of the importance of "learning teams" to promote effective learning – satisfied by the use of "lab partners" or group activities. These activities promote collaboration and effective communication, which are especially important to forensic science students who will encounter technical reviews and peer reviews in future casework. Expertise is developed through the repeated practice of applied techniques to build knowledge. In studying art and management education programs, Kolb found

that arts education is built on a "demonstration – practice – critique" process where active expression and testing are continuously part of the learning process. This continuous process would be effective in the applied science of Forensic DNA analysis by providing students with opportunities to practice their new skills and receive prompt feedback. Experiential learning promotes "self-authorship" which describes the process of constructing one's own knowledge rather than passively receiving it from another. Allowing students to author their own laboratory SOP gives them the opportunity to create their own knowledge by expressing what they learned and sharing with their peers. Experiential learning, therefore, provides the course with an effective framework for providing students with a well-founded education.

Brooks and Simpson surveyed students on the impact of experiential learning to future employment and personal and professional growth. Specifically, this study analyzed the experiences of internships and study-abroad programs. However, experiential learning may also be attained in a traditional classroom or laboratory setting through the use of simulations of real-life scenarios. Through surveys, Brooks and Simpson found that students attained skills and development through experiential learning opportunities which were transferable to future employment and life in general. The authors further recommend the use of experiential learning to improve the employability of graduates (Brooks & Simpson, 2014) .

Experiential learning causes effective learning by "promoting hands-on learning, using a problem solving process, addressing real world problems, encouraging student interaction with each other and the content, engaging in direct experiences," (S. D. Wurdinger & Carlson, 2009). A survey of instructors on their perception of experiential learning found that although instructors are aware of the benefits of this type of learning, they could not overcome obstacles to implementing it (S. Wurdinger & Allison, 2017). The surveyed instructors provided multiple

reasons for the lack of experiential learning in their classrooms, including large class sizes, unaccommodating class structures, lack of time, lack of funding, and faculty or departmental resistance. Overcoming these obstacles is the first step to improving student learning in the university classroom.

In the Advanced Forensic DNA Analysis laboratory course, a blended learning approach is taken. Blending learning theories accommodate 1) a diverse student body, 2) differing learning needs, and 3) the complex subject matter. Cognitivism and experiential learning primarily guide the course design to provide structure and experience with "real-life" techniques and scenarios. Instructors are provided with the flexibility to constantly develop and improve based on course effectiveness and student feedback, while maintaining a research-based learning framework and keeping pace with new technologies.

Best Practices in Science Education

Forensic Science as a discipline fits squarely within the realm of STEM education, as it may involve all the requisite characteristics of science, technology, engineering, and mathematics. Forensic DNA analysis in particular is found at the intersection of molecular biology science, applied technology, and statistics. Education in STEM fields relies largely on consistency in methodology. From lab to lab, school to school, country to country, the scientific principles being taught remain the same, therefore they must be communicated to students in a way that ensures students learn the principles correctly. Best practices in science education have largely been established. The National Research Council (NRC) in 1996 published the "National Science Education Standards" (Council, 1996) for use in K-12 science education. In 2013 they revised and updated these standards with "Next Generation Science Standards." While such standards do not specifically exist for higher education, these existing standards may be extended

and applied to higher education courses. In addition, the NRC has multiple publications exploring success factors in undergraduate STEM education (Council, 2003, 2012, 2015; National Academies of Sciences, 2016, 2018). The importance of educational standards is evidenced by the National Institute of Standards and Technology (NIST) which offers research grants to universities to develop standards in education. In 2018, NIST provided a grant to Bowling Green State University to develop a standardized digital forensics curriculum, consisting of lecture slides, case studies, and hands-on laboratory activities (Roy, Wu, & LaVenia, 2019). Roy et al. found that students' computer science skills increased as a result of a standard laboratory curriculum. Because educational standards do not exist for the field of forensic DNA analysis, best practices in forensic DNA education must be compiled from various sources, including NRC publications, FBI Quality Assurance Standards, and SWGDAM guidelines.

Research into undergraduate STEM education has revealed a variety of methods to improve student learning. Current practices advocate the use of evidence-based STEM educational practices and programs, which are defined as:

> *"Educational practices meeting at least one of the following criteria: the preponderance of published literature suggests that it will be effective across settings or in the specific local setting, or the practice is built explicitly from accepted theories of teaching and learning and is faithful to best practices of implementation, or locally collected, valid, and reliable evidence, based on a sound methodological research approach, suggests that the practice is effective." (National Academies of Sciences, 2018)*

Evidence-based methods include entire course overhauls, such as Student-Centered Active Learning Environment with Upside-down Pedagogies (SCALE-UP) classrooms, or techniques

which may be incorporated into already existing class structures, such as formative assessment and active learning. The NRC outlines goals for undergraduate STEM education as: 1) Increased student mastery of concepts and skills using evidence-based educational practices and programs, 2) Increased equity, diversity, and inclusion in the field of study, and 3) Increased completion of STEM credential to ensure an adequate number of STEM professionals in the field (National Academies of Sciences, 2018). By using evidence-based educational techniques, these goals may be met in the field of forensic DNA analysis.

Some programs have found success with entire restructuring of courses to improve student learning. The SCALE-UP model redesigns the classroom to use student collaboration as the primary means of instruction (Beichner et al., 2007). In a SCALE-UP classroom, the lecture and lab components of the course would be combined. Students would use time outside of class to learn the basic course content through reading and homework. Class time would be used for classroom discussion, group work, hands-on activities, problems, simulations, and experiments to challenge students to think more deeply and critically. While the proposed course would maintain the traditional divide between lecture and lab, the design of this course would easily adapt to the SCALE-UP format in the future if professors found it more beneficial to students. In existing research, the SCALE-UP class format appears to be most beneficial to introductory level courses. In the case of the advanced forensic DNA analysis course, maintaining the separation between the course and lab ensures that professors have adequate class time to communicate concepts and answer questions effectively, which is beneficial for the advanced nature of the course.

Another classroom restructure is described in the Process Oriented Guided Inquiry Learning (POGIL) approach. In a POGIL classroom, students are presented with information paired with

leading questions that guide them to formulate their own conclusions, often in small groups (Moog & Spencer, 2008). Professors take a back seat in this type of classroom, observing and facilitating learning when students have questions. While group work and inquiry-based learning is implemented in the Advanced Forensic DNA Analysis course, the sensitive nature of DNA analysis work requires more structured teaching to ensure that students are learning information correctly.

Smaller changes to course structure are easier to achieve and implement to a pre-existing course. One method that is strongly represented in the literature is formative assessment. Formative assessment is a tool intended to provide rapid feedback to students and instructors on how well the material is being understood. These assessments are "used to diagnose where a student is relative to learning goals and to inform students and instructors of any actions needed to address learning gaps… administered in the course of instruction and have low or no stakes attached to them." (National Academies of Sciences, 2018). These assessments typically take the form of miniature exams or quizzes which allow professors to diagnose student learning and adjust the pace of the class based on the results. Instructors may implement strategies such as leading whole-class discussions, briefly discuss the answer and move on, or have students discuss with their peers based on the learning needs of the class. An example of the implementation of formative assessments are ConcepTests paired with Peer Instruction, as described by Mazur. In this method, students are presented with multiple choice conceptual questions and given time to think about their answer and then discuss with a group to reach a consensus. The questions are designed to focus on important concepts, include incorrect answers based on common misconceptions, and be challenging enough that only 35-70% of students reach the correct answer (Mazur, 1997).

The value of formative assessments are reinforced by another NRC report (Council, 2003) which posits six main principles to improving undergraduate STEM learning: 1) Formative assessment is critical to student learning, 2) Formative assessment benefits students and teachers in a feedback loop that allows professors to improve their teaching strategies and course design, 3) Information gained from regular formative assessment shows effectiveness of teaching methods, 4) Formative assessment improves departments and programs as a whole by improving faculty, 5) Effective evaluation practices can foster communities of teaching and learning, and 6) Effective evaluation and teaching practices allow institutions to demonstrate the value of teaching and courses. In summary, formative assessments should be a primary tool used by instructors to evaluate and improve their teaching.

Another method of formative assessment uses writing activities to gauge student learning (Council, 2015). These writing assessments include reading reflections, prompts to make students communicate their thought processes on a concept, or short essays in which students articulate the most important or most difficult concepts that they are learning. While large classrooms would have difficulty implementing writing as a tool due to the time constraints of reading and grading writing assignments, UCO's Advanced Forensic DNA Analysis course is fortunate to have small class sizes. Writing assignments in the form of an SOP writing project make students both articulate their thinking as well as regurgitate the steps of the methodologies they are learning. In an applied science such as forensic DNA analysis, understanding and being able to follow procedure are just as important as scientific thought processes.

In addition to formative assessment, the literature also advocates active learning as a technique to improve student learning. Active learning may be defined as "practices that cognitively engage students in building understanding." Implementation of active learning may

include experiments, group activities, problem-based learning, and peer instruction. Active learning has been shown to increase students' knowledge of STEM content, understanding of concepts, and problem-solving skills (National Academies of Sciences, 2018). Freeman defines active learning as engaging students in the process of learning through activities and/or discussion in class, as opposed to passively listening to an expert (Freeman et al., 2014). Active learning emphasizes higher-order thinking and often involves group work. Freeman found that active learning increased student success in examinations by almost half a standard deviation, and that traditional lecture courses increased failure rates by 55%, regardless of discipline, class size, or course level. Additionally, it was found that active learning is most effective in small classes. Due to the small class size and advanced subject matter, addition of a laboratory with active learning activities best facilitates learning.

Discipline-Based Education Research (DBER) advocates customized learning methods based on the discipline. Forensic Biology is largely unrepresented in the literature, however biology in general is discussed. Biology laboratories are typically traditional or inquiry-based—which are designed around the learning cycle, allowing students to ask questions, confront misconceptions, develop hypotheses, and design experiments to test them. An example of this type of lab would be students answering research questions using online data sets. Upper division courses such as the Advanced Forensic DNA Analysis course at UCO are also underrepresented in the literature as far as the effectiveness of various instructional approaches. Across the board, however, use of technology has been demonstrated to improve student learning, retention of knowledge, and attitudes about science learning when used appropriately. The role of laboratories in collegiate STEM courses is also unexamined in the literature, despite their traditional importance in STEM disciplines (Council, 2012).

In an effort to construct the best possible course, the new laboratory curriculum primarily follows the structure of a traditional lab, while integrating technology and other evidence-based methods recommended by the literature. By including research-based methods such as active learning and formative assessment, students will find greater success in learning. Active learning activities are designed to make students perform tasks that professional DNA analysts would do. Formative questions are designed using pre-existing class material, the textbook, and other sources in the literature to ensure students think and reflect more deeply on the material as they learn. In addition, formation of a curriculum framework using current research-based techniques provides other forensic science programs with a launching point for establishing their own laboratory courses.

Related Work in the Literature

The literature is ripe with examples of course reforms to align with new educational research and practices. Upper-division courses seem to be the slowest to adapt new research, but there are some examples to be found in the literature. More common are cases of redesigning introductory level courses to accommodate large class sizes or new educational standards. The growing field of forensic science education in particular has a significant representation in educational research. The current project joins a growing body of work by graduate students to improve courses which they have previously taken, as seen in Wang (2017) and Jeremias (2018).

In 2017, Casagrand and Semsar published their work on redesigning an upper level neurophysiology course to include student-centered learning methods over four years (Casagrand & Semsar, 2017). Student success was measured through evaluation of cognitive skill development on exam questions. The authors found that students were more successful in assessments post-redesign. Success was attributed to four factors: changing multiple course

components, use of formative evaluation, student acceptance of the changes, and instructional support. Changes were gradually implemented over four years by adding evidence-based learning methods, including homework assignments, peer collaboration opportunities, in-class formative assessment, learning goals, and exam revisions. Throughout the four years of development, both homework questions and assessment questions were continuously revised and updated to meet the needs of students, resulting in iterative improvement of the course.

Several forensic science programs have implemented labs or hands-on activities to their courses with great success. Coticone and van Houten designed and implemented a special topics course covering basic forensic DNA and drug identification topics at Florida Gulf Coast University in 2013-2014 (Coticone & Van Houten, 2015). The course met twice per week and consisted of 45 minutes of lecture and 30 minutes of "mini laboratory exercises." Student feedback after the class requested more labs and hands-on learning. Additional expectations in comments from the students included wanting to know "How DNA is used in the real world" and to "Learn how DNA testing and drug analysis are performed and used in a forensic scene." Therefore, a laboratory course which presents problems and techniques which would be used in a working laboratory would be valuable to students.

Arwood designed and taught a course that utilized forensic science concepts to teach cell biology to non-science undergraduates (Arwood, 2004). She reported that the hands-on learning provided in the course led to students reporting higher confidence in their scientific ability after the course. Students reported anecdotally that they "learn better by actually doing science." Indeed, 61.7% of the students who took the course made positive comments about the hands-on activities in course evaluations. In an applied science such as forensic DNA analysis, some students may not have strong science backgrounds. Although certain science courses are required

to become a DNA analyst, some students may be criminal justice or other majors taking DNA courses as electives. These students would especially benefit from the addition of a laboratory course which would improve their confidence and scientific ability.

Kanu et al. describe their success in hosting a forensic chemistry workshop to give students and instructors hands-on experience in laboratory techniques. This workshop consisted of a staged crime scene and required participants to be involved in not only laboratory examination, but also the crime scene analysis and evidence collection (Kanu et al., 2015). At UCO, this comprehensive experience is already provided to forensic science students, who are required to take classes such as crime scene investigation. The integration of hands-on laboratory experiences to more specialized classes, such as advanced DNA and chemistry classes, benefits the students who seek to enter the workplace in that field by providing them with meaningful experience in the industry standard techniques and methods.

Wang, in her 2017 Master's Thesis, developed an innovative mobile learning platform for an organic chemistry laboratory course (Wang, 2017). The learning framework utilized cognitivism in its formation and incorporated techniques such as formative assessment. The mobile platform went through several rounds of testing and improvements, following the design-based research process. Wang found that the mobile learning platform improved student performance on examination but noted that other modes of learning such as peer discussion would further benefit students. New technologies do appear to be useful to student learning, but no one tool can help all students achieve success. Therefore, a range of course improvements is more beneficial to students than a single new method.

In a similar project, Jeremias developed a casework simulation activity to be used in the Digital Forensics program at UCO (Jeremias, 2018). The simulation activity provided digital

forensics students with an opportunity to perform work in a realistic environment and gain hands-on experience in work as a digital evidence analyst. Also provided in the project are guidelines for generating such simulations. The project therefore benefited both students participating in UCO's digital forensics program and students and instructors at other universities studying digital forensics.

Measures of Student Success and Course Effectiveness

Student success may be measured at many points, but not all measures are useful. Success is often measured based on graduation rate, which is the primary goal of most students. Other metrics found in the literature include course completion, accumulation of credits, time to degree, retention and transfer rates, and diversity and learning outcomes (National Academies of Sciences, 2016). The Association of American Universities measures student success based on the improvement of instruction and culture. Schreiner et al. define student success as students "thriving" in multiple facets, including academic engagement and determination, interpersonal relationships, and psychological well-being (Schreiner, Pothoven, Nelson, & McIntosh, 2009). In some studies, such as Bakar and Zaman (2008), student success was measured by comparing test scores of students taking the original versus redesigned laboratory. Course effectiveness was measured through pre- and post-course assessments for students enrolled in both the original course format and in the redesigned virtual laboratory format. Figure 5 demonstrates the formation of the effectiveness construct for Bakar and Zaman's laboratory redesign. Such a test is not possible at this time for the Advanced Forensic DNA Analysis course due to lack of data for the newly designed laboratory course. However, professors may find such a model useful for determining course effectiveness and making future course improvements.

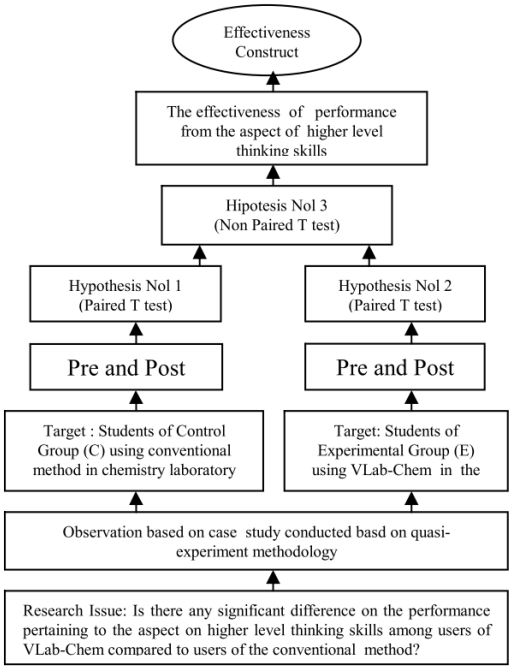


Figure 5: Effectiveness construct for course redesign. Bakar and Zaman, 2008.

Student success in this course is primarily measured through assessment, course completion, and student evaluation feedback. Assessments utilized include formative assessment to continuously gauge student learning, laboratory assignments to assess how well students are grasping and applying concepts, and periodic examinations. This project does not extend to directly measure and analyze student performance after completion of the course. Rather, professors are provided with the tools to make further adjustments to the course depending on its effectiveness. The literature has provided evidence that the techniques implemented in this course will maximize students' chances for success in the forensic DNA workplace.

Methodology

The methodology utilized in creating the proposed laboratory course primarily followed examples from the literature of successful use of cognitivist and experiential learning principles in forming science courses. The course design is supplemented with constructivist learning methods. Each laboratory lesson is designed to reflect real-world experiences of practicing forensic DNA analysts while utilizing learning theory to maximize student success. Each lesson is introduced with a review of the foundational knowledge necessary to learn the new skill, followed by activities to teach the new skill, and concluded with opportunities for students to practice their new skill and share their thoughts and understanding of the lesson with questions and a writing assignment. Each lesson is designed with a stated purpose, and a pathway for students to achieve the stated goal. This format provides students with a foundation to successfully achieve their educational and career goals in forensic DNA analysis.

General Course Design

The proposed course changes maintain a traditional format of alternating lecture and laboratory. The separation of the laboratory is consistent with other classes offered in the course of the forensic molecular biology degree. This consistency of format is important for student learning, because students already have the mental structures in place to learn in this way. Students will appreciate the familiarity of this course to the pre-requisite forensic molecular biology course. A separate laboratory and lecture also allows professors to spend adequate class time teaching theoretical concepts while also allowing students adequate time to apply and practice these concepts. This format applies readily to learning and teaching DNA analysis,

which requires both theoretical knowledge of underlying genetic science and the ability to apply knowledge to real situations.

The laboratory itself consists of approximately three-hour segments in which students review the concepts learned in class and then apply those concepts to simulated evidence material. Both the conceptual overview and the practice problems are contained in a comprehensive laboratory manual which accompanies this document. This provided material allows students to look ahead and look back at their work in order to easily follow the class structure and learning objectives.

Laboratories are structured to provide students the best opportunity for learning by following the sequence of: 1) Review information, 2) Give examples, 3) Allow practice and feedback (peer and professor), 4) Give more challenging problems, 5) Give assessments to make students think and reinforce. This sequence follows the general cognitivist framework found in Gagne's Events of Instruction or in Kolb's Experiential Learning model. Besides structuring the class to maximize student learning, the laboratory implements technology and assessments to assist in instruction. In-class activities are the primary mode of student learning by allowing students to practice the techniques they just learned and receive feedback from both peers and instructors.

Activities Design

The activities presented in each laboratory lesson are designed to simulate casework a student might encounter in a future workplace. As such, sources are used which provide a realistic simulation of a working laboratory design, including the Oklahoma State Bureau of Investigation Forensic Biology Unit Manual (Investigation, 2018). Activities are selected based

on 1) the realities of working as a forensic DNA analyst, 2) the limitations of the course and students, and 3) ease of implementation. In studies on STEM education, difficulty and cost are often cited as reasons for not implementing new educational technology (Ejiwale, 2013). This author is hopeful to demonstrate that implementing new types of learning into classrooms need not be difficult or costly, especially when added to an existing course.

The activities in the laboratory fall into three general categories, with some overlap: DNA profile interpretation (labs 1, 3, 4, 6), statistical calculations (labs 5, 6), and use of software programs (labs 2, 5, 7). Additional shorter exercises are included which involve method validations, paternity testing, report writing, and an online activity on wildlife forensics. The order of laboratories follows the format of the lecture portion of the class to gradually build students' knowledge of DNA analysis methods. Students learn the primary task of a DNA analyst: to analyze and interpret DNA profiles and perform the necessary statistical calculations. Integrating software programs gives students a foundation for technologies they will encounter in a working forensic laboratory.

An exemplar laboratory is Laboratory 5: Statistical Weight for Single Source Profiles, which includes a group activity, DNA profile interpretation, statistical calculations, and use of a software program. The laboratory manual is contained in Appendix A; the following example material is taken from Laboratory 5. The laboratory begins with describing the purpose and goals of the lab, which is prescribed by cognitivism to ensure students understand what they are learning. The purpose statement consists of a brief summary of the lab objectives which provides students with a task to perform, as in Figure 6.

Laboratory 5: Statistical Weight for Single Source Profiles

Objective

The purpose of this laboratory is to learn strategies to determine statistical weight of evidence for single source DNA profiles.

Figure 6: Example of Laboratory Objective. Advanced Forensic DNA Analysis Laboratory Manual, Appendix A.

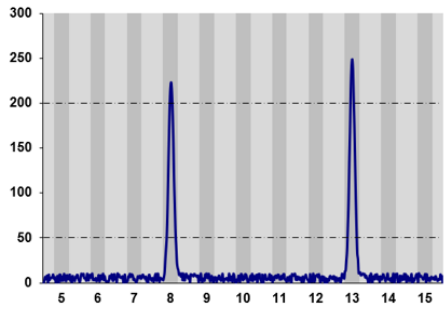
Following the purpose statement is a review of the material covered in the lecture portion of the course. The background material presents the knowledge students need to complete the laboratory. In most cases the review contains examples of the types of problems students will encounter in the laboratory procedure and assignment, as in Figure 7.

Calculating an RMP requires using the product rule to combine the frequency of each allele at each locus. Alleles are treated differently depending on if they are heterozygous or homozygous. Heterozygous alleles have two possible combinations, for example Allele A and Allele B may be A, B or B, A. Therefore, the probability of observing these alleles is 2xPxQ where P and Q represent the population allele frequencies. When interpreting homozygous alleles, on the other hand, order does not matter, so the probability of observing homozygous allele A would be P^2 where P represents the population allele frequency.

Finally, the frequency of each locus is multiplied together and inverted to reach the RMP.

Example:

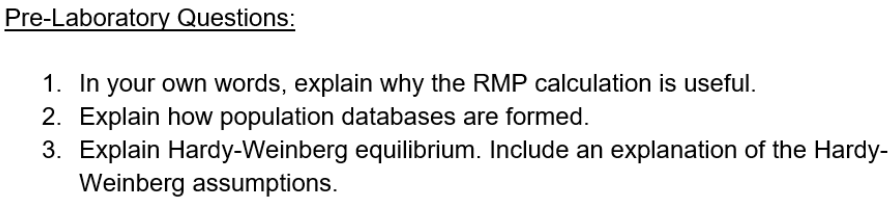
For the following locus, D16S539:



Since the alleles are heterozygous, genotype frequency = 2pq
We can utilize population data (Caucasian) to determine 2(0.0180) (0.163) = 0.00587
This calculation is repeated at each locus and the results multiplied together to determine the profile frequency. The result is divided by 1 to reach the RMP.

Figure 7: Example of laboratory background information and example calculation. Advanced Forensic DNA Analysis Laboratory Manual, Appendix A.

Students are required to read the material and answer pre-laboratory questions prior to beginning the laboratory activity. Professors may choose to go over the questions as a class at the beginning of the laboratory section to gauge student knowledge of the topic and clear any misconceptions. Pre-laboratory questions cover material presented in the laboratory background information or in the corresponding class lecture or textbook chapter. Questions are intended to make students think more deeply about the material, for example:



Pre-Laboratory Questions:

1. In your own words, explain why the RMP calculation is useful.
2. Explain how population databases are formed.
3. Explain Hardy-Weinberg equilibrium. Include an explanation of the Hardy-Weinberg assumptions.

Figure 8: Example of pre-laboratory questions. Advanced Forensic DNA Analysis Laboratory Manual, Appendix A.

The laboratory procedure section outlines the step-by-step process students follow to learn a new skill. Activities are chosen and designed to give students opportunities to actively learn techniques that they would apply in a working forensic DNA analysis laboratory. Most of the activities consist of interpreting DNA profiles and performing calculations related to profile analysis. Others involve the use of software programs. All activities require the student to actively participate in performing the technique. Procedures are lengthier for activities involving software programs to ensure that students do not get overwhelmed by the unfamiliar software environment. Other procedures are quite brief and require students to understand how to perform calculations presented in the course material. The different activity types are demonstrated in Figure 9 below, taken from Laboratory 5:

Procedure

Population Frequency:

Choose a prominent feature that you can determine from your classmates such as hair color, shirt color, type of shoe, etc. This feature will represent alleles of a gene.

1. Gather data on the occurrence of this feature in your class population.
2. Calculate the frequency of this feature in your class population.
3. Share your results with your class
4. Now, using your class data, determine the rarity of your "genotype".

Calculating the RMP:

For the following electropherogram:

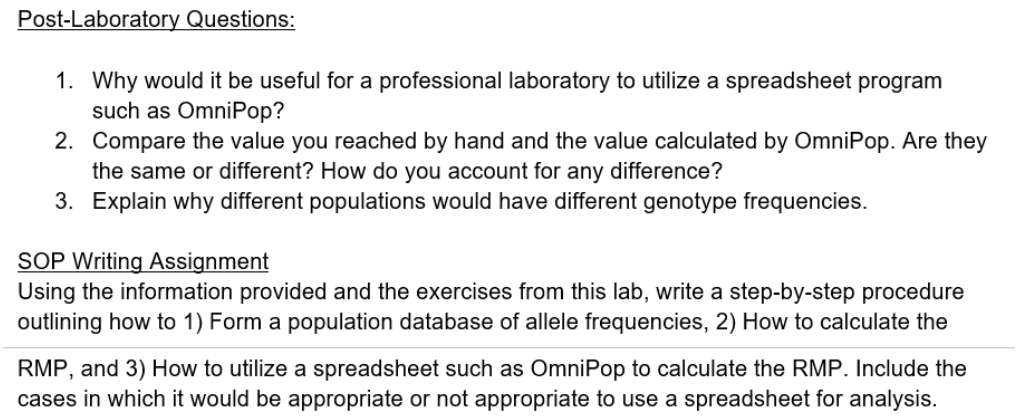
1. Determine the genetic profile for the sample.
2. Calculate the RMP using the allele population frequency from Appendix I. Show your work.

OmniPop:

1. Open OmniPop from your computer desktop. Look for an Excel Spreadsheet file called "OmniPop200.1.xls"
2. When the file opens, it will start on the "Main" sheet, indicated by the green tab at the bottom of the window. Note the other sheets next to this: "WorldMap", "Data", "Refs", and "Info". We will revisit these tabs later.
3. At the top of the "Main" sheet, you will see the name of the program and publisher. Below this is a list of DNA loci and alleles. Click on the first allele for locus D8S1179. A drop-down arrow will appear next to the cell. If you click on the arrow, it will show a list of possible alleles. By selecting different alleles from the list, you can see the graph to the right, and the values in the spreadsheet below change. This is indicating the program is recalculating the profile frequency.
4. Note the two grey rectangles below the allele list: "Show Frequencies" and "Show Map".
5. First, click on "Show Frequencies". This will bring you further down in the current sheet to show you the genotype frequencies for each population calculated by the database. Notice that there are many populations in this database. For the purposes of this class, we will only utilize the FBI databases for Caucasian, Hispanic, and African American. Click "Return to Top" or scroll top the top of the page.
6. Now, click "Show Map". This will take you to the "WorldMap" sheet. A world map is displayed which indicates the most likely origins of the genotype.
7. Next, click on the "Data" tab at the bottom of the window to open the Data sheet. This page lists the allele frequencies for each locus for each population. Scroll down to find loci D2S1338 and D19S433. The original downloadable version of this spreadsheet does not contain population frequency information in the FBI database columns for these loci. They have been updated in the classroom version of the spreadsheet using the new FBI data from 2015. If you download this program to use outside of class, you will need to fill in these sections using the new FBI data.
8. The "Refs" and "Info" tabs contain the program references and description respectively.
9. Now, using the profile provided in part 2, enter the alleles into the spreadsheet.
10. Use the spreadsheet to determine the genotype frequency for the three major US populations.

Figure 9: Example laboratory activities, including group activity, DNA profile interpretation, statistical calculations, and use of technology. Advanced Forensic DNA Analysis Laboratory Manual, Appendix A.

Following the activity, the laboratory concludes with assessment questions and the SOP writing project assignment. Students should work alone or with their professor on individual questions to gauge their own understanding and identify misconceptions. Professors may compare results from the pre- and post-laboratory questions to determine effectiveness of the laboratory activity. The SOP project involves group work to give students the opportunity for peer instruction and feedback. Both post-laboratory assessments make students think more deeply about the concepts covered in the laboratory, as in Figure 10. The laboratory questions and SOP assignment should be used by professors to diagnose student learning and adjust teaching to maximize student learning.



Post-Laboratory Questions:

1. Why would it be useful for a professional laboratory to utilize a spreadsheet program such as OmniPop?
2. Compare the value you reached by hand and the value calculated by OmniPop. Are they the same or different? How do you account for any difference?
3. Explain why different populations would have different genotype frequencies.

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to 1) Form a population database of allele frequencies, 2) How to calculate the RMP, and 3) How to utilize a spreadsheet such as OmniPop to calculate the RMP. Include the cases in which it would be appropriate or not appropriate to use a spreadsheet for analysis.

Figure 10: Example post-laboratory questions and SOP Writing Assignment. Advanced Forensic DNA Analysis Laboratory Manual, Appendix A.

The sequence of each laboratory follows the cognitive process to maximize student learning. Students will gain skills which they will be able to apply in future work as a DNA analyst. Professors are provided with ample opportunities to continuously improve the course to meet student needs and increase student success.

Selection of Software

Integrating technology into the laboratory has a two-fold purpose: to expose students to the modern technological methods which they will likely encounter in future workplaces, and to assist in learning. Three main technologies are integrated into this laboratory: Omnipop, a statistical spreadsheet, OSIRIS, an expert system, and Lab Retriever, a probabilistic expert system. At the time of writing, these programs fit well into the structure and needs of the course. Software programs are always at risk of becoming obsolete due to how fast technology develops. Therefore, the framework established by this project is designed to easily adapt to new and different technologies. Software should be selected based on the needs of students and professors, and the resources available to the program.

OmniPop is a freely available spreadsheet containing the data and formulas required to determine genotype profile frequency in various populations ("STRBase Population Data," 2015). It is utilized to calculate the RMP statistic. OmniPop was written by Brian Burritt of the San Diego Police Department and is published online by NIST's STRBase project. STRBase was created by John Butler, who also authored the textbook used in the Advanced Forensic DNA Analysis course. The STRBase website, strbase.nist.gov, contains abundant information and tools for DNA analysts. Many labs develop their own spreadsheets with similar functionality to OmniPop, so students will benefit from using the software in a classroom setting. OSIRIS, or "Open Source, Independent Review and Interpretation System", is a software program used to view and analyze STR data (Goor, Forman Neall, Hoffman, & Sherry, 2011). The software is used to analyze DNA profiles of single-source samples. OSIRIS may function as an expert system in the interpretation of single source samples. Expert systems are becoming more common in laboratories because of the speed and efficiency of their analysis. These programs

can be used to replace the technical review step for single-source DNA profiles, allowing labs to decrease backlogs. With the constant development and improvement of artificial intelligence algorithms, expert systems will likely become even more prominent in the future. Students hoping to enter the forensic DNA field in the next few years will benefit from early exposure to the use of these programs. Lab Retriever is a software program based on the work of Dr. David Balding. Dr. Balding published the likeLTD PGS program in 2013 as a package for the statistical computing program "R". While mathematicians and statisticians may be familiar with the usage of R, it would present a steep learning curve for most forensic scientists. Nonetheless, the program effectively allowed analysts to compute likelihood ratios for DNA profiles, including complex mixtures and dropout. In 2015, a team led by Kirk Lohmueller of the University of California, Los Angeles developed Lab Retriever by implementing the algorithms of likeLTD into a more user-friendly software program (Inman et al., 2015). By re-coding the program and redesigning some parts of Balding's algorithm, the Lab Retriever authors were able to increase the program speed. A graphical user interface was implemented to allow forensic scientists without a strong mathematics or technology background to easily use the software. Lab Retriever is utilized to determine the likelihood ratio for single- and mixed-source samples. As with the other programs used in this course, Lab Retriever introduces students to the use of PGS programs, which they will use in future forensic DNA analysis laboratories. The selection of these software programs for use in an upper level forensic DNA course gives students an advantage that will increase success in their careers in the field.

A review of the literature has demonstrated that the types of technologies described above are present in working DNA analysis laboratories in the United States and beyond. Haned reviews and compares currently available expert systems (Haned & Gill, 2016). In an earlier

review of statistical evaluation software, Steele and Balding undergo a more in-depth discussion of the use of these software programs (Steele & Balding, 2014). A comparison of the technologies discussed by the two papers reveals how quickly technology can develop, change, or become obsolete. This demonstrates the importance of understanding the principles behind the software, rather than just memorizing steps. Early exposure to the use of these methods provides students with a foundation for future success in the field of forensic DNA analysis by understanding how to learn and understand statistical software. The selected software programs are all open source, meaning they are freely available to the public. Open source programs are selected for use in this course because 1) this presents no additional cost to the program, which is a potential barrier for implementation of technology in STEM classrooms, 2) students (and professors) are able to install these programs on their own personal computing devices to get extra practice and learning outside of the classroom, 3) student with more advanced computing knowledge or curiosity into programming may examine the source code of these programs for their own knowledge. Providing students with easily accessible tools and resources for learning facilitates them taking control of constructing their own education.

Assessment Design

A final addition to the Advanced Forensic DNA Analysis course provided by the new laboratory curriculum is several methods of student assessment. Formative assessment is advocated in the literature as a primary mechanism for adjusting the pace of a course to fit student needs. Typically taking the form of short quizzes, questions, or writing assignments, formative assessments are a critical tool for gauging student understanding. Therefore, assessment questions written to accompany the provided laboratory format and class schedule benefit both students and professors by allowing them to readily adjust the class pacing.

Formative assessments are based in research and have been shown to increase student success. In this course, formative assessments come in three modes: pre- and post-laboratory questions, SOP writing assignments, and class examinations. All three of these methods of assessment measure student knowledge and understanding of the subject matter as well as increase high-order cognitive skills such as critical thinking and applying knowledge to new situations. Course examinations already exist for the lecture section of the Advanced Forensic DNA Analysis course. While the new laboratory curriculum does not add to these examinations, it does increase student readiness and confidence entering the examinations.

Laboratory questions serve a dual purpose of measuring student knowledge both before and after a laboratory activity. This helps professors to measure the effectiveness of laboratory activities. Questions are formed from sources including the course textbook, lecture slides, and regulatory bodies including SWGDAM and the FBI. The questions range in complexity from basic knowledge recall (lower-order cognitive skills) to critical thinking questions (higher-order cognitive skills). Besides ensuring that students are learning and understanding the skills they are being taught, laboratory questions provide students with opportunities to work with peers and to prepare for exam material.

Each laboratory is followed by an "SOP Writing Assignment" in which students are tasked with reiterating the laboratory method they used in formal procedure language. Every professional laboratory has a written standard operating procedure document which instructs employees how to perform tasks. The laboratory assignment gives students the experience of not only following a procedure, but also collaborating with their peers to write one. Students work in groups to ensure that the procedure is correct, comprehensive, and coherent. By regurgitating knowledge into writing in their own words, students reinforce what they have learned.

Additionally, any errors in students' thought process are caught either by peers or instructor upon review of this assignment. The SOP writing assignment is in line with constructivist and experiential learning principles in which students perform "real-life" tasks in order to construct their own knowledge. Likewise, working in groups with peers further enhances effective student learning.

Designing a laboratory based in experiential learning theory as well as cognitivist and constructivist principles results in optimal student learning. Students not only become more successful in this discipline but also develop learning skills that they will bring with them to other courses. Additionally, other courses at the university may benefit from the example set by this project.

Opportunities for Continuous Improvement

Besides the specific activities provided in the new laboratory curriculum, the learning framework being established allows professors to take opportunities to further improve the course. Design-based research shows that continuous feedback and improvement of course design results in a better course with each iteration. Professors may adjust the course pace based on the results of formative assessment of the class. For example, poor assessment performance on mixture interpretation would lead professors to give additional time and practice on the topic. If students perform poorly on the SOP writing project, professors may diagnose that more time should be spent on understanding the step-by-step procedure of a technique. Formative assessments not only allow professors to improve the course each semester, but also to adjust the course for specific class dynamics, which change with each semester. By separating the laboratory activities from the lecture, professors can effectively manage their time and may

potentially incorporate experiential learning opportunities such as visiting a working DNA laboratory or inviting an expert guest speaker. The questions and procedures provided in the lab manual would easily incorporate different technologies or activities. If one of the provided software programs becomes obsolete, then it may be replaced with a new one. By designing an adaptable and accommodating educational framework, professors can increase opportunities for student success with each subsequent semester.

Conclusion

The purpose of this project is to design a laboratory course to accompany the Advanced Forensic DNA Analysis course at the University of Central Oklahoma Forensic Science Institute which increases student success in the Forensic Molecular Biology program and beyond. A laboratory manual was created to provide the laboratory structure including practical exercises for students to perform. Current educational theory has been used in the construction of the proposed course to ensure maximum student success post-completion of the course and upon entry into the working field of forensic DNA analysis. Educational methods such as formative assessment, group work, and active learning are all integrated into the course design. Student success is measured by assessments, student feedback and evaluations, and post-graduation success. Professors of the course are provided with an education framework to continuously improve the course design based on student needs and developments in the field. The goal of this project is accomplished by providing the structure and tools necessary for students to increase their success in the field of forensic DNA analysis by developing laboratory skills.

Graduate Project Limitations

This project is limited in that it does not span the completion of the fully-integrated course, and therefore does not directly measure student success. In addition, students are limited to the use of open source technology for practical exercises, rather than the more expensive software programs which students are likely to encounter in future workplaces. This drawback is mitigated, however, by the fact that students benefit more from understanding the principles behind such software, rather than memorizing steps to run a program. Finally, the literature is lacking in research of whether students benefit more from separate lecture and lab, or from

integrated laboratory exercises. This author is hopeful that this work adds to the body of literature and furthers this research.

Significance of Findings

This work is beneficial to students because the new course format integrates concepts from learning theory to maximize student success. In addition, students have the opportunity to apply concepts and use technologies that they previously might have seen for the first time upon entry to the professional field of forensic DNA analysis. In a highly competitive field, students emerging from this program have an advantage over other applicants. Professors benefit from this project because the creation of a separate laboratory with pre-made laboratory exercises frees up classroom time for more advanced and detailed learning. It also allows professors to explore other avenues of teaching, such as integrating a mock court or taking a field trip to a working DNA laboratory. The FSI and the Forensic Molecular Biology program are improved by this project because the integration of modern technology and education theory make the program more competitive. Students seeking higher education in this field want to enter a program which gives them the best chance of success in their desired field of work. By improving student success, the program itself becomes more successful. Additionally, other courses within the program may benefit from this project's demonstration of implementing hands-on activities by integrating modern educational theory and technology. Finally, this work contributes to the body of literature on forensic DNA analysis by exploring new ideas in the forensic DNA classroom. Research involving higher education in STEM is limited, especially regarding Forensic DNA Analysis. The specific topic of this work provides future researchers with a launching point into improving their own programs in forensic DNA analysis education.

Suggestions for Future Research

Future work should examine the impact of separate laboratory and lecture versus integrated laboratory exercises. The effectiveness of laboratory courses and the key elements necessary for an effective laboratory course should be examined. Additionally, more research is needed on the direct effects of classroom laboratory experience on student success in the workplace. In the United States, forensic DNA laboratories have a mandatory training period for new hires, which cannot be bypassed by having student experience in the techniques. However, many DNA analysts entering the workplace take up to two or three times the mandatory training period. By having previous experience in applying the techniques used by working laboratories, students emerging from this program will hopefully be able to finish their training closer to the minimum amount of time. Future research, therefore, could survey working DNA analysts on their education and experience prior to entering the working field of DNA analysis. Finally, more research into student success factors would be beneficial to the literature. Direct evidence of student success beyond the classroom would give insight into which educational methods most benefit students in the long term.

References

Amiel, T., & Reeves, T. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Educational Technology and Society, 11*(4), 29-40.

Arwood, L. (2004). Teaching cell biology to nonscience majors through forensics, or how to design a killer course. *Cell Biol Educ, 3*(2), 131-138. doi:10.1187/cbe.03-12-0023

Bakar, H. N. B., & Zaman, H. B. (2008). *Virtual laboratory for chemistry based on the constructivism-cognitivism-contextual approach (VLab-Chem).* Paper presented at the 2008 International Symposium on Information Technology.

Bar, W., Brinkmann, B., Lincoln, P., Mayr, W. R., & Rossi, U. (1994). DNA recommendations--1994 report concerning further recommendations of the DNA Commission of the ISFH regarding PCR-based polymorphisms in STR (short tandem repeat) systems. *International Journal of Legal Medicine, 107*, 175-176.

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D. L., Allain, R. J., & Risley, J. S. (2007). The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project. *Research-Based Reform of University Physics, 1*(1), 2-39.

Brooks, L., & Simpson, L. (2014). Experiential Learning in Higher Education: Implications for Improving Employability of FCS Graduates. *Journal of Family and Consumer Sciences, 106*(2), 16-21.

Casagrand, J., & Semsar, K. (2017). Redesigning a course to help students achieve higher-order cognitive thinking skills: from goals and mechanics to student outcomes. *Adv Physiol Educ, 41*(2), 194-202. doi:10.1152/advan.00102.2016

Center, U. N. T. H. S. (2019). 2019-2020 Course Catalog. Retrieved from http://catalog.unthsc.edu/content.php?filter%5B27%5D=-1&filter%5B29%5D=&filter%5Bcourse_type%5D=-1&filter%5Bkeyword%5D=forensic+genetics&filter%5B32%5D=1&filter%5Bcpage%5D=1&cur_cat_oid=7&expand=&navoid=760&search_database=Filter#acalog_template_course_filter

Commission, F. S. E. P. A. (2019). Accredited Universities.

Coticone, S. R., & Van Houten, L. B. (2015). DNA, Drugs, and Detectives: An Interdisciplinary Special Topics Course for Undergraduate Students in Forensic Science. *Journal of College Science Teaching, 45*(2), 24-29.

Council, N. R. (1996). *National Science Education Standards*. Washington, DC: The National Academies Press.

Council, N. R. (2003). *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics*. Washington, DC: The National Academies Press.

Council, N. R. (2012). *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*. Washington, DC: The National Academies Press.

Council, N. R. (2015). *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. Washington, DC: The National Academies Press.

Domin, D. (1999). A Content Analysis of General Chemistry Laboratory Manuals for Evidence of Higher-Order Cognitive Tasks. *Journal of Chemical Education, 76*(1), 109-112.

Driscoll, M. P. (2005). *Psychology of learning for instruction* (3 ed.). Boston: Allyn & Bacon.

Ejiwale, J. (2013). Barriers to Successful Implementation of STEM Education. *Journal of Education and Learning, 7*(2), 63-74.

Evett, I. W., & Weir, B. S. (1998). *Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists*: Sinauer Associates Inc.

Forensic Science Institute. (2019). Retrieved from https://www.uco.edu/fsi/

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci U S A, 111*(23), 8410-8415. doi:10.1073/pnas.1319030111

Gagne, R. M., & Medsker, K. L. (1996). *The conditions of learning: Training applications.* Fort Worth, TX: Harcourt Brace College Publishers.

Gibbins, L., & Perkins, G. (2013). Laboratories for the 21st Century in STEM Education.

Goor, R. M., Forman Neall, L., Hoffman, D., & Sherry, S. T. (2011). A mathematical approach to the analysis of multiplex DNA profiles. *Bull Math Biol, 73*(8), 1909-1931. doi:10.1007/s11538-010-9598-0

Haned, H., & Gill, P. (2016). Expert Systems in DNA Interpretation. In A. Jamieson & S. Bader (Eds.), A Guide to Forensic DNA Profiling: John Wiley & Sons.

Hares, D. R. (2015). Selection and implementation of expanded CODIS core loci in the United States. *Forensic Sci Int Genet, 17*, 33-34. doi:10.1016/j.fsigen.2015.03.006

helptostudy.com. (2018). Best Forensic Science Colleges in the U.S. Retrieved from https://www.helptostudy.com/best-forensic-science-colleges-in-the-us/

Hudson, C. C., & Whisler, V. R. (2007). Contextual teaching and learning for practitioners. *Valdosta: Vadolsta State University*.

Inman, K., Rudin, N., Cheng, K., Robinson, C., Kirschner, A., Inman-Semerau, L., & Lohmueller, K. E. (2015). Lab Retriever: a software tool for calculating likelihood ratios incorporating a probability of drop-out for forensic DNA profiles. *BMC Bioinformatics, 16*, 298. doi:10.1186/s12859-015-0740-8

Investigation, O. S. B. (2018). Forensic Biology Unit Policy Manual. Oklahoma.

Januszewski, A., & Molenda, M. (Eds.). (2008). *Educational Technology* (2 ed.): Routledge.

Jeremias, M. S. (2018). Computer Crimes Case Simulation and Design Model. *University of Central Oklahoma Archives and Special Collections*.

Jonassen, D., Mayes, T., & McAleese, R. (1993). A manifesto for a constructivist approach to uses of technology in higher education. In *Designing environments for constructive learning* (pp. 231-247): Springer.

Kanu, A. B., Pajski, M., Hartman, M., Kimaru, I., Marine, S., & Kaplan, L. J. (2015). Exploring Perspectives and Identifying Potential Challenges Encountered with Crime Scene Investigations when Developing Chemistry Curricula. *Journal of Chemical Education, 92*(8), 1353-1358. doi:10.1021/ed500671x

Kivunja, C. (2014). Do You Want Your Students to Be Job-Ready with 21st Century Skills? Change Pedagogies: A Pedagogical Paradigm Shift from Vygotskyian Social Constructivism to Critical Thinking, Problem Solving and Siemens' Digital Connectivism. *International Journal of Higher Education, 3*(3). doi:10.5430/ijhe.v3n3p81

Kolb, A. Y., & Kolb, D. A. (2005). Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. *Academy of Management Learning & Education, 4*(2), 193-212.

Mazur, E. (1997). *Peer Instruction: A User's Manual*: Pearson.

Methods, S. W. G. o. D. A. (2010). SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories.

Molphy, M., & Pocknee, C. (2005). Chemistry online in a constructivist environment. *Proceedings of ASCILITE*.

Moog, R. S., & Spencer, J. N. (2008). POGIL: An Overview. In *Process Oriented Guided Inquiry Learning (POGIL)* (pp. 1-13).

National Academies of Sciences, E., and Medicine. (2016). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. Washington, DC: The National Academies Press.

National Academies of Sciences, E., and Medicine. (2018). *Indicators for Monitoring Undergraduate STEM Education*. Washington, DC: The National Academies Press.

O'Connor, C. (2015). A Practice-Led Approach to Aligning Learning Theories with Learning and Teaching Strategies in Third Level Chemistry Education. *Irish Journal of Academic Practice, 4*(1). doi:10.21427/D7B42F

Quality Assurance Standards for Forensic DNA Testing Laboratories, (2011).

Report to the President: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. (2016).

Rieger, G. (2017). Conceptualizing Success And Performance for Adult Learners: Merging the Contexts of Adult Education and Professional Training. *Journal of Pedagogy and Psychology "Signum Temporis", 9*(1), 57-62. doi:10.1515/sigtem-2017-0004

Robinson, R., Molenda, M., & Rezabek, L. (2008). Facilitating Learning. In A. Januszewski & M. Molenda (Eds.), *Educational Technology*. New York: Routledge.

Roy, S., Wu, Y., & LaVenia, K. (2019). Experience of Incorporating NIST Standards in a Digital Forensics Curricula. *IEEE*.

Schreiner, L. A., Pothoven, S., Nelson, D., & McIntosh, E. J. (2009). *College student thriving: Predictors of success and retention.* Paper presented at the Annual meeting of the Association for the Study of Higher Education, Vancouver, British Columbia.

Shiland, T. W. (1999). Constructivism: The Implications for Laboratory Work. *Journal of Chemical Education, 76*(1), 107-108.

Simonson, M. R., & Thompson, A. (1997). *Educational computing foundations*: Prentice-Hall (Merrill) Publishing Co.

Steele, C. D., & Balding, D. J. (2014). Statistical Evaluation of Forensic DNA Profile Evidence. *Annual Review of Statistics and Its Application, 1*(1), 361-384. doi:10.1146/annurev-statistics-022513-115602

STRBase Population Data. (2015). Retrieved from https://strbase.nist.gov/populationdata.htm

Terhart, E. (2003). Constructivism and teaching: a new paradigm in general didactics? *Journal of Curriculum Studies, 35*(1), 25-44.

University, V. C. (2019). Forensic Science Course Catalog. Retrieved from http://bulletin.vcu.edu/azcourses/frsc/

Wang, S. (2017). Learning Theory Guided Iterative Development of a Mobile Learning Paradigm Aimed at Enhancing Organic Chemistry Laboratory Instruction at Undergraduate Level. *UC San Diego Electronic Theses and Dissertations*.

Welsh, C., & Hannis, M. (2011). Are UK undergraduate Forensic Science degrees fit for purpose? *Sci Justice, 51*(3), 139-142. doi:10.1016/j.scijus.2011.03.003

Wieman, C. (2015). Comparative Cognitive Task Analyses of Experimental Science and Instructional Laboratory Courses. *The Physics Teacher*, *53*(6), 349-351. doi:10.1119/1.4928349

Wurdinger, S., & Allison, P. (2017). Faculty Perceptions and Use of Experiential Learning In Higher Education. *Journal of e-Learning and Knowledge Society, 13*(1), 15-26.

Wurdinger, S. D., & Carlson, J. A. (2009). *Teaching for experiential learning: Five approaches that work*: R&L Education.

Appendix A: Advanced Forensic DNA Analysis Laboratory Manual

Advanced Forensic DNA Analysis Laboratory Manual

Prepared by: Claire Joyce
University of Central Oklahoma
Forensic Science Institute

Contents

Introduction

The purpose of this laboratory course is to give the student a greater understanding and a working knowledge of the procedures used in working labs to interpret forensic DNA profiles. Upon completion of this course, the student will be able to not only interpret single and mixed source profiles, but also will have experience in using modern software technology to assist in analysis. The skills learned in this laboratory will be directly applicable to working in a forensic laboratory.

This manual should be utilized within the Advanced Forensic DNA Analysis course at the University of Central Oklahoma Forensic Science Institute. Students will follow this manual to learn the methods utilized for analysis of DNA evidence. When conducting laboratory experiments, students should first review the introduction material and answer the pre-laboratory questions. Then, students may proceed to the laboratory activity and post-laboratory assignment. When questions or discrepancies arise, students should refer to the referenced sources or to the course instructor.

Standards and Guidelines

This lab manual is meant to be used as a companion to J. M. Butler's *Advanced Topics in Forensic DNA Typing: Interpretation*. The methods and guidelines outlined in that textbook and in this manual are based on recommendations by the Scientific Working Group on DNA Analysis Methods (SWGDAM), the National Research Council (NRC), the Federal Bureau of Investigation (FBI) Quality Assurance Standards (QAS), and the International Society for Forensic Genetics (ISFG). Students are encouraged to read and reference these sources throughout this course.

Students seeking additional help beyond the above resources are recommended to look at the resources on www.bioforensics.com/videos/, as well as other resources provided throughout this manual.

Standard Operating Procedure - Semester Project

As each method is learned in lab, it will be written into a shared standard operating procedure (SOP) for the class to use in subsequent labs. Students may be divided into groups for this project.

In general, the steps of forensic DNA analysis are as follows:

1. Determine if the profile is useable
2. Determine if the profile is a mixture
3. Identify peaks vs. artifacts
4. Determine the number of contributors
5. Utilize necessary calculations

6. Make allele calls
7. Compare to reference samples
8. Calculate weight of evidence

Laboratory 1: Preparing Data for Interpretation

Objective

The purpose of this laboratory is to learn the steps to process data prior to interpretation. These steps include determining the usability of data, identifying mixtures vs. single source samples, differentiating artifacts from true alleles, and setting and using thresholds.

Introduction

When receiving a DNA profile for analysis, the first step is to determine the usability of the profile. If the profile has enough data present to perform analysis, it may be used. The usability of a profile does not determine if a conclusion may be formed. If the analyst determines that the profile is useable, then they must determine if the profile is single source or a mixture. A mixture may be identified by the presence of more than two alleles at multiple loci or by loci with only two alleles having extreme peak imbalance. If the profile is determined to be a mixture, then the steps of analysis will change to account for mixture interpretation. However, if the profile is determined to be single-source, then analysis may proceed to allele calls.

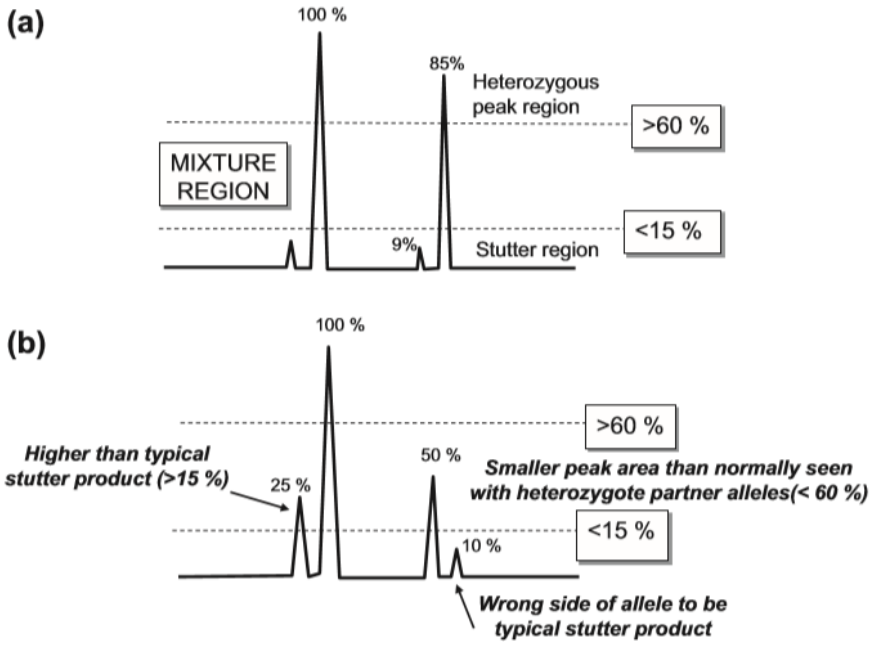
Alleles are designated by the number of repeat units present at a genetic locus. Generally, each person has two variants, one inherited from each parent. These two variants are described as heterozygous if they are different, or homozygous if they are the same.

Determining Usability

The development of increasingly sensitive DNA collection and analysis techniques in recent years has led to an increase in DNA samples which are unsuitable for analysis. Touch DNA often has such a small number of cells present that a partial DNA profile may be produced which does not have any significant evidentiary value. Unfortunately, DNA analysis laboratories lack outside guidance on when to cease their analysis and call the profile unusable, and must instead produce their own guidelines through internal validation. Some labs have solved this problem by introducing a “complexity threshold” which allows them to throw out profiles with high stochastic effects or drop-out. Others have simply advocated common sense as the best method to determine if a profile has valid evidentiary value or not.

Determining Number of Contributors

Another effect of improved DNA collection methodology is an increase in the number of mixed samples. Minor contributors to a DNA sample may appear in an electropherogram even if their contribution consisted of only a few cells. Analysts must be able to determine the presence of a mixture early in their analysis to ensure the correct application of interpretation methods. Mixed samples may be indicated by several factors, including: significant heterozygote imbalance, high stutter ratio (0.15), or the presence of three or more alleles in any locus.



Butler's Interpretation, Chapter 6: DNA Mixtures

Thresholds

Threshold levels are determined on a lab-by-lab basis due to variation in instrumentation, supplies, and environmental factors. There is no standard number at which thresholds are set, but there is a general range in which thresholds are expected to be. Labs use validation studies to determine their own threshold levels and incorporate those into their internal SOPs. The Oklahoma State Bureau of Investigation Forensic Biology Unit, for example, has set a stochastic threshold of 950 RFU using the GlobalFiler kit.

Two main types of thresholds are set and used for forensic DNA analysis:

- Analytical threshold is the RFU value below which observed peaks cannot be reliably distinguished from instrument noise. Therefore, peaks observed above this threshold indicate the presence of an allele (or artifact).
- Stochastic threshold is the RFU value which indicates that stochastic issues may be present with peaks falling below the threshold. A commonly encountered stochastic effect would be drop-out of a sister allele to a below-threshold peak. In other words, peaks observed below the stochastic threshold cannot be assumed to be homozygous, even if no other peaks are observed. Additionally, peak height ratios may be lower than typical when alleles fall below the stochastic threshold.

Threshold levels are set based on instrument signal ranges and limitations. These limitations are quantified in the Limit of Detection (LOD), Limit of Quantitation (LOQ), and Limit of Linearity (LOL). In turn, the limit values are calculated based on the average background noise level. The LOD indicates the point at which data may be reliably distinguished from background noise and is typically set at three standard deviations above the average noise variation. Likewise, the LOQ indicates the point at which data may be reliably quantified and is set at ten standard deviations above average noise variation. Either value may be used as the analytical threshold value, but analysts must be aware that quantitative information is only available above the LOQ. The LOL may be thought of as the upper limit of reliable data. Above the LOL, the instrument is unable to gather reliable data due to signal saturation. Data landing between the analytical threshold (whether using the LOD or LOQ) and the LOL is within the "Dynamic Range," indicating that data may be reliably used for analysis.

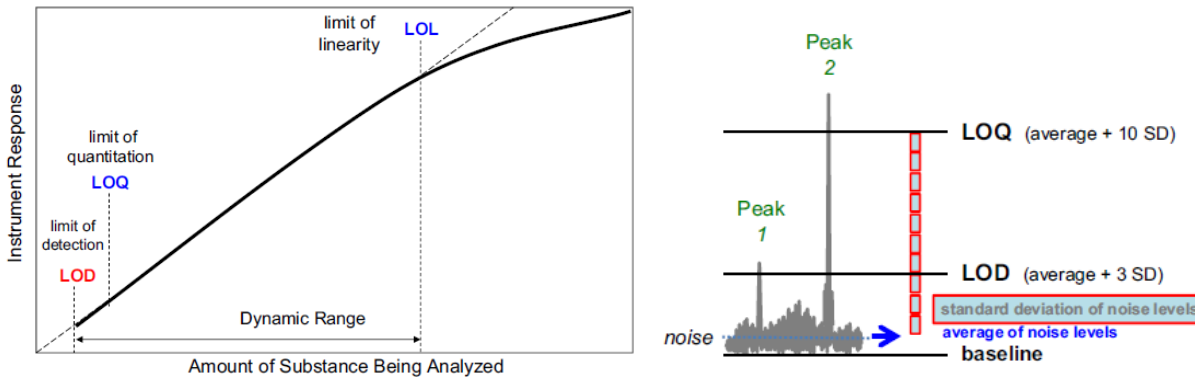


TABLE 2.2 Threshold Decisions and Potentially Useful Validation Data

Thresholds to Determine	Decisions to Make (lab- & kit-specific)	Useful Validation Data
Analytical = ___ RFU	Will a single overall value or color specific values be used?	Noise levels in negative controls or non-peak areas of positive controls.
Stochastic = ___ RFU	Will a minimum peak height value or alternative criteria such as a quantitation value be used? Or will a probabilitistic genotype approach be used?	Level where dropout occurs in low-level single-source heterozygous samples under conditions used (e.g. different injection times, post-PCR cleanup).
Stutter filter = ___%	Will profile-wide, locus-specific, or allele-specific stutter filters be applied?	Stutter in single-source samples (helpful if examined at multiple DNA quantities).
Peak Height Ratio = ___%	Will profile-wide, locus-specific, or signal height (quantity) specific PHRs be applied?	Heterozygote peak height ratios in single-source samples (helpful if examined at multiple DNA quantities).
Major/Minor Ratio = ___	At what mixture ratio will an attempt be made to separate components of a two-person mixture into major and minor contributors for profile deductions?	Defined mixture ratios (e.g. 1:1, 1:3, 1:9) with known samples to observe consistency across loci and to assess ability to deduce correct contributor profiles.

Butler's Interpretation, Chapter 2: Data, Models, Thresholds

Artifacts

In the course of DNA analysis, there are multiple types of "artifacts" which may be encountered among true data. These artifacts may be caused by several stages of the DNA analysis process including PCR, the instrument or kit itself, or from contaminants.

Types of commonly encountered artifacts include:

- Stutter is a type of artifact introduced during the PCR step. Stutter is observed as a smaller peak typically appearing one repeat unit more or less than the true peak which is below the stutter threshold (typically 15% or less of true peak height) due to polymerase slippage during PCR. Software programs for analysis of electropherograms are typically able to identify stutter peaks, but analysts must be aware of the possibility of mixtures before disregarding these peaks. Stutter thresholds are determined through internal validation and may be different for specific loci.
- Minus A is another artifact which is produced by PCR. During the PCR process, polymerase tends to add an adenosine to the end of PCR product. This is encouraged for consistency between DNA strands. When strands are produced without addition of the terminal adenosine, it results in creation of a peak at the -1bp position, which is referred to as a minus A artifact. Often this artifact is eliminated by extra extension time at the end of PCR, dependent on the lab's internal validation data.
- Spikes occur due to unpredictable events during capillary electrophoresis and are non-reproducible. They are identified by a high peak running through all dye channels at the same location.
- Pull-up occurs when a peak in one dye channel bleeds-through into one or more other channels. Typically, this is due to too much DNA template in the original sample and the instrument being unable to resolve the spectral overlap. If pull-up appears to interfere with true alleles or if it interferes with analysis, the sample may need to be re-amplified.
- Dye blobs are often consistent with a dye kit lot due to superfluous primer dyes being dissociated from PCR product. Unlike true alleles, they typically have a rounded appearance rather than a vertical peak. If a dye blob interferes with analysis, the sample should be re-injected.
- Additionally, kits may report specific artifacts found during their developmental validation. For example, GlobalFiler reports a 90 RFU amplicon in TH01 at n-10 to n-12. Kit documents and validations should be thoroughly reviewed prior to internal validation when implementing a new kit.

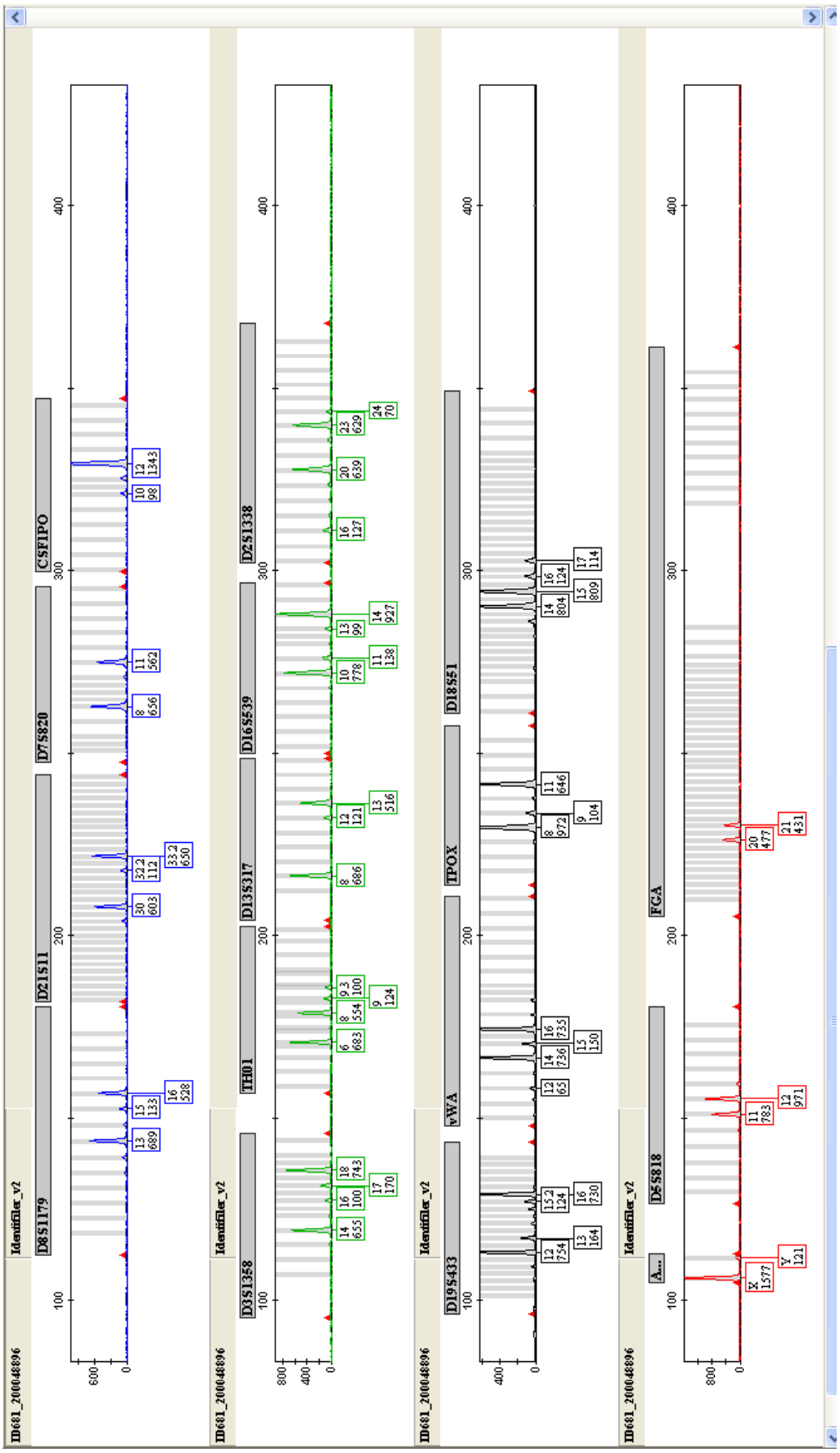
Pre-Laboratory Questions:

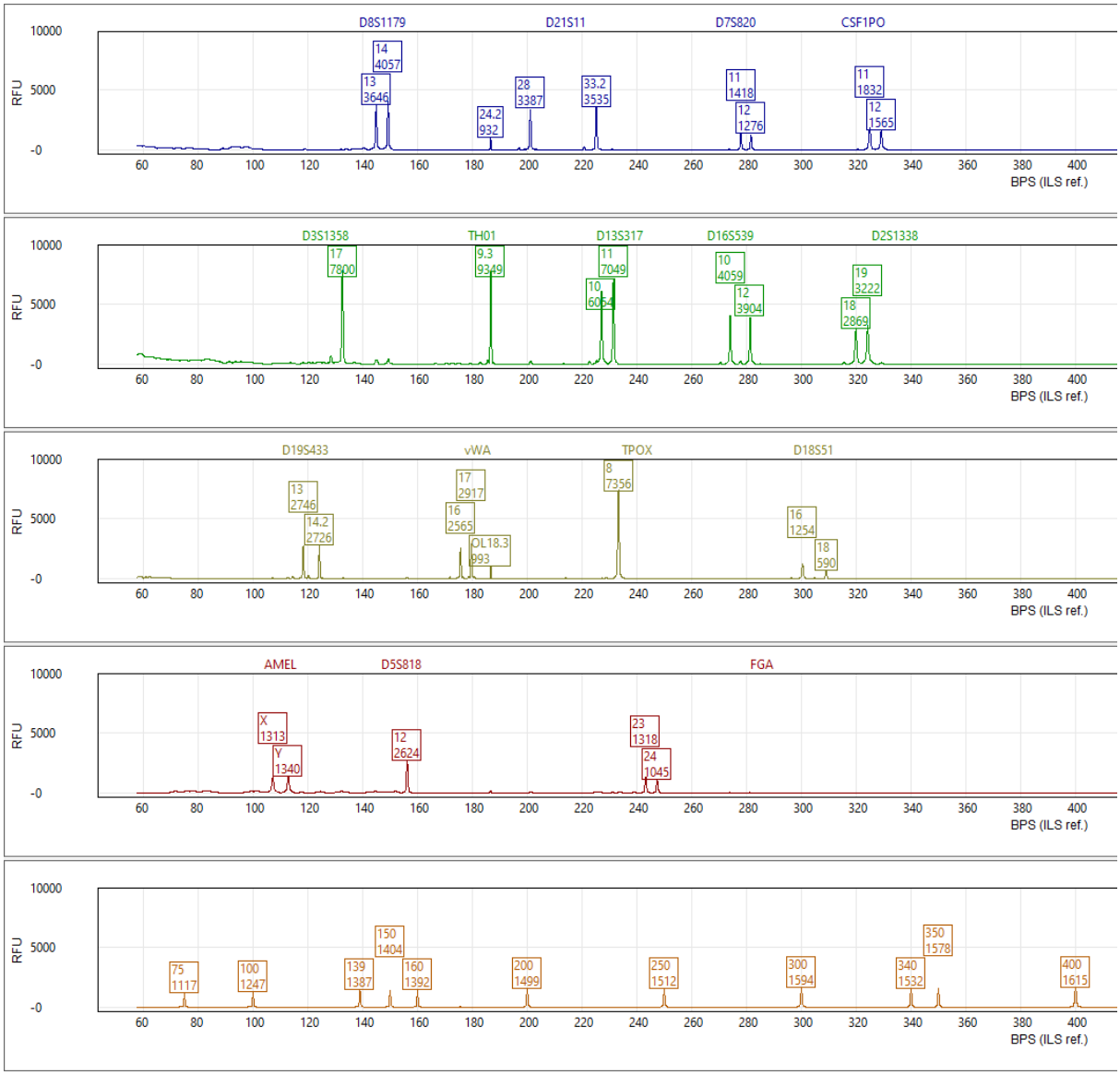
1. Explain how an electropherogram is obtained from a biological sample.
2. What are SOPs and what should they include?
3. Describe what a real STR peak looks like.

Procedure

For the following electropherograms, determine:

- If the profile is useable for interpretation and describe your reasoning.
- If the profile has one or more DNA contributor(s) and describe your reasoning.
- Explain how you would determine an appropriate analytical threshold and describe your reasoning.
- For any artifacts present, identify the type of artifact and describe your reasoning.

1. 
D681_20004896
Identifiler_v2
D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
D5S818
FGA
Amelogenin

2. 

Post-Laboratory Questions:

1. Which type of artifact do you think is most commonly encountered? Why?
2. Describe how you would differentiate a stutter peak from a true peak.
3. What would be the LOD for an electropherogram in which the average baseline noise was 15 RFU with standard deviation of 3 RFU? What is the LOQ?

SOP Writing Assignment:

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to 1. Determine usability, 2. determine the number of contributors, 3. set and use thresholds, and 4. identify and handle artifacts.

Laboratory 2: Single source profiles

Objective

The purpose of this laboratory is to review the interpretation of single source profiles and introduce tools to assist in analysis such as OSIRIS expert system software. Students will practice interpreting profiles by hand, and then use OSIRIS for analysis to compare results.

Introduction

Review: Single Source Profile Analysis

Once an electropherogram is determined to have only one contributor, the rest of the analysis process is relatively straightforward. For each locus examined, alleles may be called to form a profile. Each individual has two alleles per locus, one inherited maternally, the other paternally. If these alleles are the same, they are called homozygous; different alleles are called heterozygous. On an electropherograms, homozygous alleles cannot be visually differentiated from each other, instead forming one peak. Ideally, peak heights may be used to indicate homozygous alleles because the combined peak heights should form a peak twice as high as similarly-sized heterozygous peaks. Due to variations in the real world, however, this may not always be the case. Likewise, heterozygous peaks within a locus should be similarly sized to each other. The size similarity is quantified using the Peak Height Ratio (PHR), and a certain expected value is set through internal validation. 100% PHRs are never seen between real heterozygous alleles due to stochastic variation. PHRs are calculated one of two ways depending on laboratory SOP:

Method in Figure 4.1	Equation	Definition of terms	Reference
Method #1	$Hb' = h^{(2)} = \frac{\emptyset_{smaller}}{\emptyset_{larger}}$	*smaller* = smaller peak height *larger* = larger peak height	Leclair et al. 2004 Kelly et al. 2012
Method #2	$Hb = h = \frac{\emptyset_{HMW}}{\emptyset_{LMW}}$	h = heterozygote balance $\emptyset$ = peak height HMW = higher molecular weight allele LMW = lower molecular weight allele	Leclair et al. 2004 Kelly et al. 2012

Butler's Interpretation, Chapter 4: STR Genotypes

In this course, we will use method 1, in which we will divide the smaller peak height value (RFUs) by the larger peak height value, and multiply by 100 to obtain a percentage. The obtained

value indicates the differing levels of allele recovery. The expected PHR for sister alleles will vary by lab SOP based on internal validations.

Expert Systems

Interpretation of non-complex, single-source profiles has been largely automated through the use of expert system software. Programs such as FSS-i^3, TrueAllele, GeneMapper ID-X, and the open source OSIRIS allow analysts to automate all or part of the interpretation process. In practice, these systems are used on known offender and reference samples for the purpose of uploading to databases such as the National DNA Index System (NDIS). Expert systems may also be used as part of the review process for casework samples but are not used as the sole analyzer for unknown samples.

OSIRIS is the only open source expert system which is approved by the FBI for use on reference samples being entered into NDIS.

For the purposes of this exercise, OSIRIS will be available for students on laboratory-provided computers. Students wishing to explore OSIRIS outside of laboratory should follow the instructions for installation and use provided by the NCBI at: https://www.ncbi.nlm.nih.gov/osiris/help/#installing-and-upgrading-osiris

Upon launching the OSIRIS program, there will be three menus available in the toolbar at the top left. “File” gives options for starting a New Analysis or Opening a previous analysis. “Tools” allows users to select different options for analysis. “Help” contains the user manual with instructions for use under “Documentation”. Students are encouraged to consult the Documentation for questions about OSIRIS not covered in this Lab Manual. The procedure to be followed in this lab will generally follow the steps given in the OSIRIS documentation.

To begin, a brief discussion of file types is in order. OSIRIS accepts two file types for analysis: .fsa and .hid. Both types of file are fragment analysis data files created by DNA sequencers and analyzers. Different analysis programs will provide different file types. Other types of files students may encounter are .csv, .fasta, and .fastq. While these files may contain genetic information and be used by different analysis programs, they are not suitable for use with OSIRIS.

The OSIRIS software package comes with several sets of test or example data for which it gives step-by-step instructions for analysis. To familiarize students with the software, part 1 of this lab will guide students through one such example.

Pre-Laboratory Questions:

1. What are two situations where an electropherogram may seem to depict a mixture, but there is another valid explanation for the alleles present? How would you identify these situations and how would you explain them to a jury?
2. Describe three advantages of using an expert system in the forensic DNA analysis process.
3. How would you differentiate an off-ladder allele from an artifact?

Procedure

Part One: Example Profile in OSIRIS

Follow the steps below to utilize OSIRIS to analyze a DNA profile.

1. Navigate to the OSIRIS installation folder containing the example data sets on your computer. For most Windows computers, this is found under "This PC > User (C:) > Program Files (x64) > NCBI > Osiris > TestAnalysis"
2. Within this folder are six different data sets from different analysis software programs. For this example, we will use "Identifiler". Open this folder to view the contents.
3. Two more folders are visible, "Identifiler_Artifacts" and "STRbaseIF". The first folder contains various ladders and examples of artifacts found in Identifiler kits. For this exercise, we will be using the second folder, "STRbaseIF". Open this folder to view the contents.
4. Students will see six .fsa files within this folder. Notice that three of the files are unknown DNA profiles, labeled with "victim." The remaining three contain the ladder, negative control, and positive control. These files will be imported by the OSIRIS program for analysis.
5. Open OSIRIS on your computer. This may be done by opening the software through an icon on the desktop or toolbar, or through the installation folder "This PC > User (C:) > Program Files (x64) > NCBI > Osiris > OsirisAnalysis.exe"
6. Select "File > New Analysis" from the toolbar at the top left. This will open a pop-up window where users may specify the Input Directory and Output Directory, type of analysis procedure, ILS, RFU thresholds, and type of input data.
7. The Input Directory tells the program where to look for source files. In this example, the data is in the "Identifiler" folder identified in Step 2. Click Browse next to Input Directory and navigate to this folder. Output Directory may be specified by the user. In this example, you may create a folder on the desktop with your name and select that folder as the output directory.
8. Select the Operating Procedure Name for your analysis. In this example, we are using the Identifiler procedure, so select Identifiler from the dropdown list.

9. The program automatically selects data RFU thresholds based on the operating procedure you have selected. The Identifiler procedure sets all thresholds at 150. For this example we will leave the default values of 150, however students may practice with adjusting these values to see how the data changes.
10. We will leave the ILS and data type as the defaults of "ABI-LIZ450" and "Raw" respectively.
11. Press OK to proceed with analysis.
12. OSIRIS will analyze data from both folders present in the input directory: "Identifiler_Artifacts" and "STRbaseIF". For now we will only examine data from the STRbaseIF folder. Select this analysis from the list and click "View Selection".
13. The window that opens is by default on the "Preview" or table view. The top menu shows the files which were analyzed, from "Identifiler LADDER" to "MIX05case4_victim". Each row will display allele data gained from the file.
14. In the middle of the window is a graphical representation of the selected file. "Identifiler LADDER" should be selected by default, so the graph will show the peaks of the ladder.
15. The bottom section of the window will display any notices from the file or program. It should be empty for this file.
16. Select file 4 "MIX05case2_victim" from the top section.
17. Now, the graph and notices sections will change. Students will see a graphical representation of peaks from all the dye channels in the middle section. Below will be a series of notices about the file. Browse through the list of notices to get an idea of issues that may be present in the sample. Note that OSIRIS recommends reamplification. The program has the option for the analyst to accept each of these recommendations by clicking on the hyperlinks on the left and accepting the notices.
18. To view the separate dye channels, select the Graph option at the top left toolbar.
19. Now, the data is separated by dye channel. To view all dye channels easily, from the Graph button at the top left, select "Resizable Plots". From the Graph button again, hover over and select "Synchronize Axes" and then "Reset Axes".
20. The view may be changed by right clicking on any channel plot and selecting "Remove Plot" to hide plots, or "Remove other Plots" to focus to one channel. Selecting "Multiple Plots" will return to viewing all plots.
21. To zoom to peak data, from any plot, click and drag to form a box around the data you want to zoom too. This may be all visible alleles, one locus, or individual peaks. You may zoom back out by right clicking on any plot and selecting "Reset Axes". Be aware that scrolling with your mouse over any plot will cause the plot view to change. You can manually scroll back to your original position, or right clicking and selecting "Reset Axes" on any plot.
22. Peak information may be edited by clicking on the peak label. A pop-up window will open with options to disable the peak, mark it as an artifact, or to Edit Locus. Selecting

Edit Locus will allow you to view and change information about all peaks present at the locus.

23. The graphical view may be exported and printed using Graph > Export Graphic File. Save a copy of your graph, zoomed to contain data from all of the channels.

Part Two: Applying Knowledge

Review the example provided above. We will now use the second profile from the example data set, "MIX05case3_victim".

Utilize the OSIRIS software and your knowledge of DNA profiles to determine 1) the peak height ratios of each locus, 2) how many contributors are present in the profile, and 3) the genetic profile for the contributor(s) for the profile "MIX05case3_victim".

Part Three: More Practice

In order for students to gain practice using expert system software, students will use the example data sets provided with the OSIRIS software. Following the example procedure above, repeat Part 1 (begin at Step 6) and Part 2 for the profiles in folders "GlobalFilerHID" and "ProfilerPlus". Students may choose any one of the profiles included in these sets.

Post-Laboratory Questions:

1. After using the OSIRIS expert system software, do you feel more or less confident in your ability to interpret DNA profiles? Why?
2. What would cause a locus to only show one allele, if the individual is known to be heterozygous at that locus?
3. What would cause a locus to only show two alleles if every other locus indicates a mix of two individuals.

SOP Writing Assignment:

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to utilize expert systems such as OSIRIS to assist in DNA analysis interpretation. Include the cases in which it would be appropriate or not appropriate to use an expert system for analysis.

Laboratory 3: Mixed source profiles

Objective

The purpose of this laboratory is to learn how to identify and interpret simple mixed-source DNA profiles.

Introduction

DNA samples are said to be a "mixture" when they contain genetic material from more than one individual. Remember from Laboratory 1 that mixed samples may be indicated by: significant heterozygote imbalance, high stutter ratio (0.15), or three or more alleles per locus. If any one of these factors is seen, the analyst should be prompted to look more closely at the entire profile to determine the number of contributors. The presence of a mixture inevitably complicates the analysis process. Variation in the number of contributors, amounts they contributed, types of body source, and level of degradation can all be factors affecting analysis. This lesson will focus on simple mixtures, or major/minor mixtures, which contain genetic material from no more than two individuals, of which one individual contributed significantly more DNA than the other. In contrast, indistinguishable two-person mixtures are those in which two contributors may be identified, but sister alleles cannot be paired up through peak heights. Alternatively, certain types of mixtures may be separated prior to DNA extraction by performing a differential extraction. The most prominent example of this is sexual assault cases, in which the two individuals' DNA is contained by epithelial cells and sperm cells respectively. Differing characteristics of these two cell types allows them to be mechanically separated so that the DNA never becomes mixed.

Steps to Mixture Deconvolution: Clayton et al. (1998)

1. **Identify the presence of a mixture**
2. **Designate allele peaks**
3. **Identify the number of potential contributors**
4. **Estimate the relative ratio of the individuals contributing to the mixture**
5. **Consider all possible genotype combinations**
6. **Compare reference samples**
7. **Determine statistical weight of evidence**

We have already covered steps 1 and 2 in previous laboratories. Now we will tackle steps 3-5.

The number of potential contributors can be estimated by looking at the number of alleles present in each locus. We know that each individual contributor gave two alleles per locus, but we do not know if they are heterozygous or homozygous, and we do not know if there is allele overlap between contributors. Therefore, analysts must look across all loci to determine the number of contributors. If there are no more than four alleles present at any one locus, we can assume the presence of two contributors: 4 alleles/2 alleles/individual = 4 alleles x individuals / 2

alleles = 2 individuals. If an odd number of alleles are present, we must round up to account for homozygous alleles or overlap: 5 alleles x (individuals / 2 alleles) = 2.5 individuals = 3 individuals (at least). As the number of contributors increases, or if low levels of DNA are present, this estimation becomes much more difficult due to allele overlap and preferential amplification skewing the electropherogram results. Laboratory SOPs should address when analysts must proceed or discontinue analysis based on these complicating factors.

The relative ratio of contributions to a mixture are calculated using the mixture ratio (MR): this value compares the relative amounts of genetic material contributed to a sample by two individuals. Within a locus, MR= sum of major peak height values (RFU) / sum of minor peak height values (RFU). The MR is presented in ratio format such as 1:1 or 1:5. Analysts should start with loci that do not demonstrate overlap or homozygosity.

Major/minor deconvolution

Mixtures in which contributors gave significantly different amounts of genetic material may often be deconvoluted by designating major and minor contributors and differentiating the genotypes of those individuals. Because each chromosome pair is represented equally throughout all cells, alleles on an electropherogram should be represented at approximately the same intensity for a single individual. While various environmental, preferential, and stochastic factors do cause some variation in allele representation, it is not significant enough to prevent the formation of genotypes for individuals. Most often, this approach is only useful in two person mixtures, but an analyst should make the determination based on their laboratory SOP. The previously discussed PHR (As discussed in Laboratory 2. For two alleles within a locus, PHR=smaller peak height (RFU) / larger peak height (RFU) x 100) and MR calculations may be used to assist in this determination, as well as the Major/Minor Proportion. The major/minor proportion (MP) value quantifies the amount of genetic material an individual contributed to the whole sample. Within a locus, MP = sum of minor (or major) peaks / sum of all peaks. The opposite value is found by subtracting from one, for example, 1 -minor proportion=major proportion.

Restricted vs Unrestricted:

When analyzing a DNA mixture, all genotype combinations must be considered, unless the analyst is able to restrict combinations using peak height ratios and other quantitative information. When the analyst can restrict genotype combinations, often through the use of major/minor deconvolution, it is referred to as a restricted approach, whereas an unrestricted approach must consider all genotype combinations. Even combinations of alleles which have peak height ratios below the acceptable threshold would be considered in this approach.

Another method to separate mixed profiles is the use of backing out reference samples. In a case where DNA is taken from the body of a victim, such as skin found under fingernails, it is likely

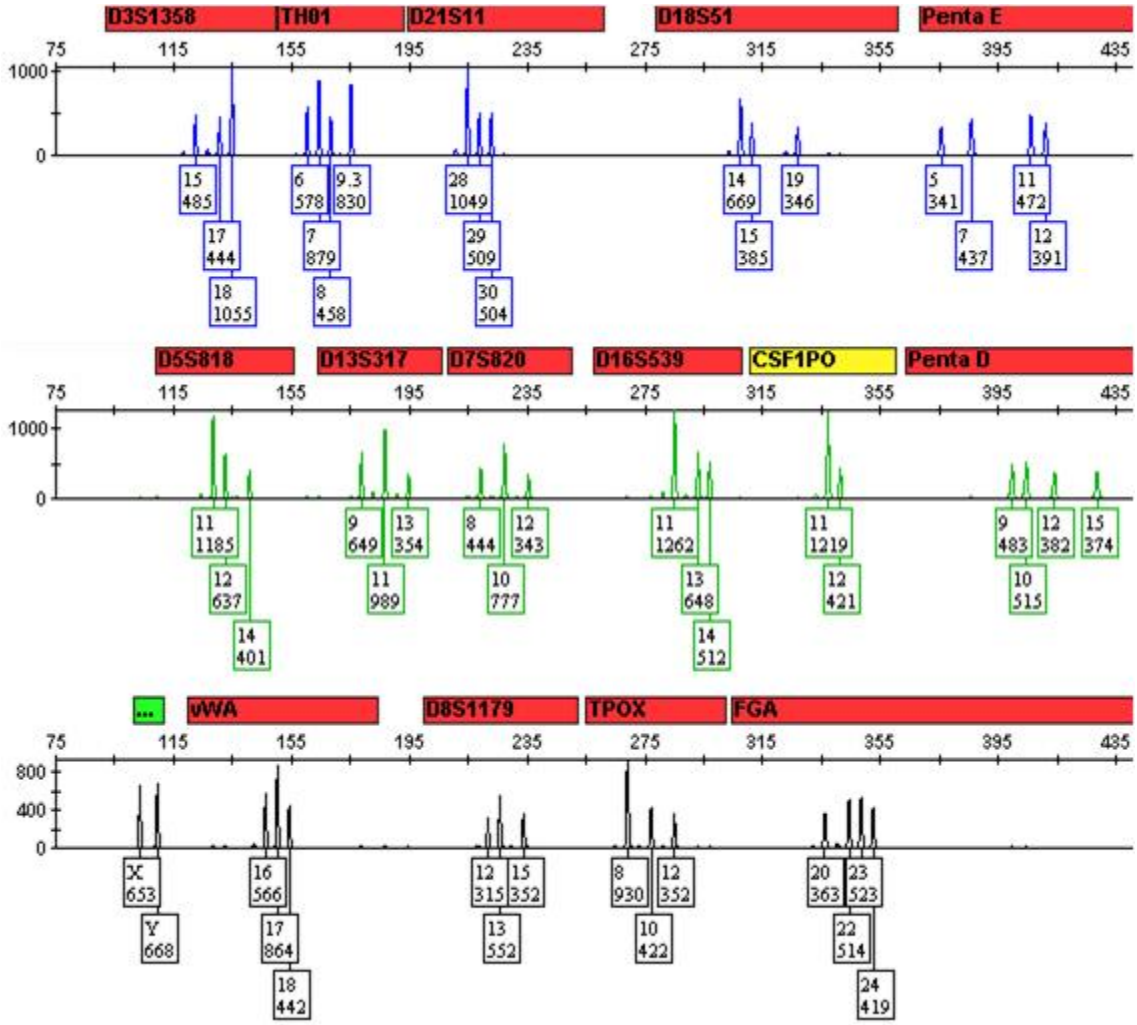
that the victim's DNA is present as well as the perpetrator. In these cases, a reference sample is taken from the victim for comparison. When the resulting DNA profile clearly contains the victim's genotype, those alleles may be "backed out" or excluded from analysis to allow for analysis of only the remaining DNA. Analysts must be careful to not completely exclude alleles which may be overlapped the other individual's alleles. After a genotype has been identified or called, it may then be compared to suspect reference samples.

Pre-Laboratory Questions:

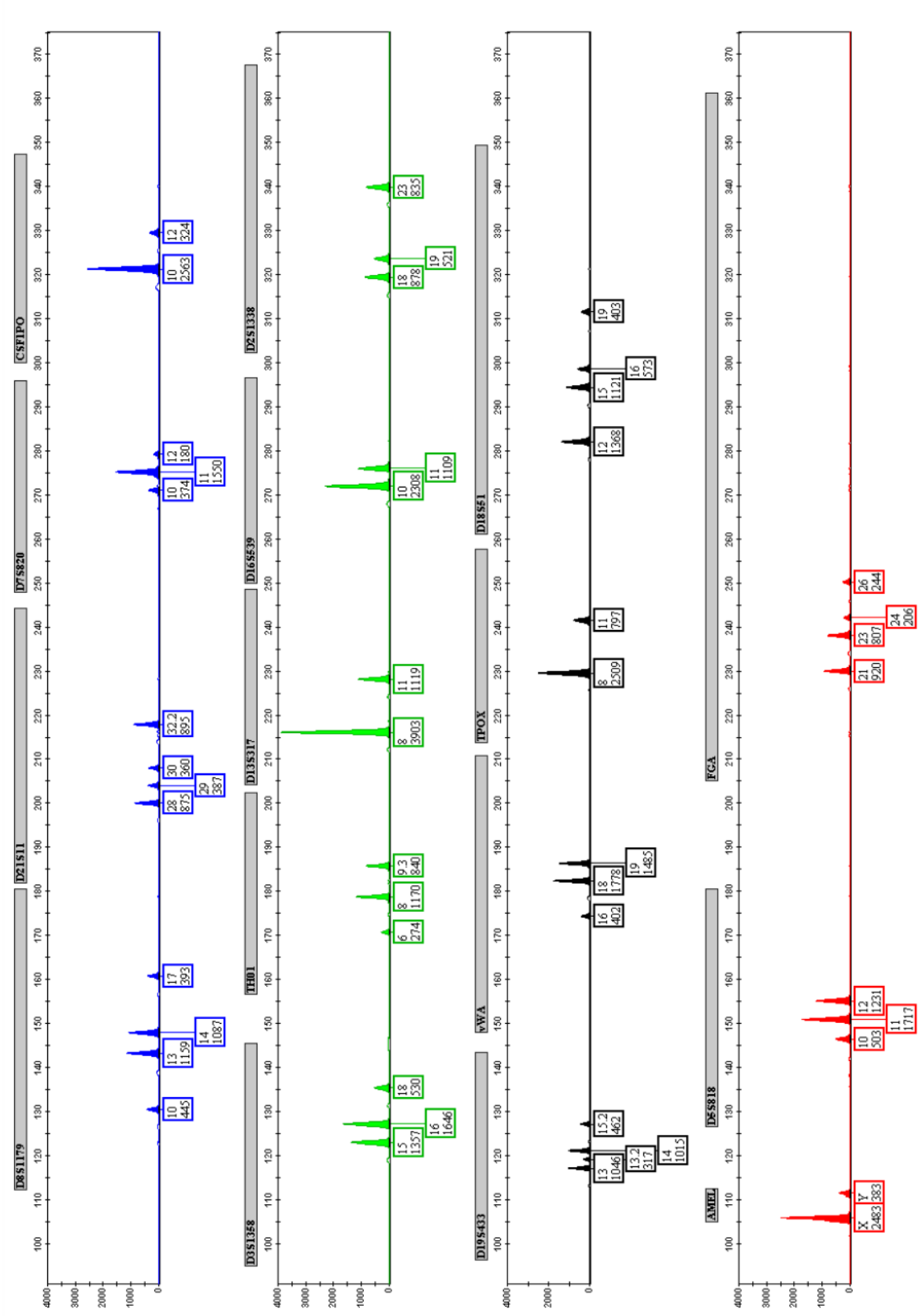
1. Explain why an analyst should never look at DNA profiles taken from known suspects before evaluating the evidentiary profile.
2. How many individuals could potentially be present in a mixture if a locus has 7 alleles visible?
3. Describe what differential extraction is and why it is valuable when handling mixtures.

Procedure

For the following electropherograms, determine the minimum number of contributors, the type of mixture (major/minor, indistinguishable, etc.) and describe a strategy for interpreting the profile.

1. 
D3S1358
TH01
D21S11
D18S51
Penta E
D5S818
D13S317
D7S820
D16S539
CSF1PO
Penta D
...
vWA
D8S1179
TPOX
FGA
75
115
155
195
235
275
315
355
395
435
1000
0
800
400
0
15 485
17 444
18 1055
6 578
9.3 830
7 879
8 458
28 1049
29 509
30 504
14 669
19 346
15 385
5 341
7 437
11 472
12 391
11 1185
12 637
14 401
9 649
13 354
11 989
8 444
12 343
10 777
11 1262
13 648
14 512
11 1219
12 421
9 483
12 382
15 374
10 515
X 653
Y 668
16 566
17 864
18 442
12 315
15 352
13 552
8 930
12 352
10 422
20 363
23 523
22 514
24 419

2.


D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
AMEL
D5S818
FGA

Post-Laboratory Questions:

1. Calculate the mixture ratios and major/minor proportions for the two electropherograms above.
2. Explain why major/minor deconvolution is possible.
3. How would you handle a mixed profile in which the minor profile appears to be missing alleles due to drop-out?

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to deconvolute a major/minor profile to determine the two genotypes.

Laboratory 4: Complex Mixtures

Objective

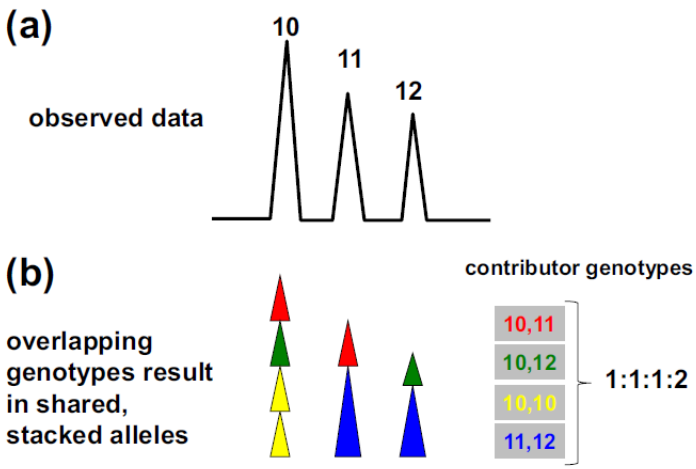
The purpose of this laboratory is to explore strategies DNA analysts may use in the interpretation of complex DNA mixtures.

Introduction

DNA samples containing genetic material from more than one individual may be simple or complex to interpret depending on the number of contributors and on the amounts contributed by the individuals. Samples containing DNA from more than two individuals is defined as a complex mixture. This laboratory will focus on these profiles, which are more difficult to analyze and interpret. In addition to the tools previously used such as peak height ratio, mixture ratios, and major or minor proportions, complex mixtures may require the use of additional tools such as consensus profiles or probabilistic genotyping.

SWGDAM guidelines recommend that labs perform validation studies and establish rules for interpretation of mixtures in which no major or minor may be discerned and further recommends the use of probabilistic genotyping, which will be covered in a future lab. Across the field, analysts are encouraged to avoid dealing with highly complex mixtures whenever possible. If there are multiple DNA samples from a crime scene, analysts should begin with those that are easier to interpret. In addition, analysts should only report the results they are confident in, and never attempt to guess.

Complex mixtures are challenging because of allele sharing and the potential presence of low-template DNA (LTDNA). Allele sharing occurs when two (or more) contributors both exhibit the same allele, making the two alleles indistinguishable on an electropherogram. The occurrence of allele sharing makes it difficult to differentiate the full genotypes. LTDNA occurs in mixtures because PCR is limited to 1ng of total DNA. When there are multiple contributors, they are each diluted by the others, creating the possibility of drop-out.



Butler's Interpretation, Chapter 7: Low-Level DNA and Complex Mixtures

Replicate PCR amplifications may be performed to create a consensus or composite profile to assist in analysis. This method uses more sample material but eliminates problems with drop-in and may assist with drop-out issues. If the original sample size is small, however, sub-dividing it into multiple PCR runs may increase drop-out. A consensus profile includes all alleles seen in the replicate runs, while a composite profile includes only repeated alleles to avoid drop-in. Both types of profile may be used to approximate the true profile(s).

Drop-out and low-quality samples make the use of a stochastic threshold very important. If allele peaks are visible below the stochastic threshold, then a dropped-out sister allele may exist. When allele stacking is a possibility, however, missing alleles cannot always be attributed to drop-out. Validation studies assist in this effort by testing replicate samples of known DNA genotypes to determine when drop-out can be expected. Analysts should be cautious, however, to avoid using expectations from validation studies when analyzing more complex samples that are not consistent with the validation study.

At any given locus, there are multiple explanations for the presence or absence of alleles depending on the possibility of drop-out, drop-in, homozygosity or heterozygosity. The probability of drop-out, Pr(D) and of drop-in Pr(C), is generally calculated during a lab's internal validation based on the amplification kit used. Pr(D) represents an estimation of how often alleles may be expected to not amplify under PCR conditions with LTDNA. The calculation of Pr(D) and Pr(C) is beyond the scope of this class, but students may read more about it on John Buckleton's blog: https://johnbuckleton.wordpress.com/john-butler/

Some laboratories set a threshold of complexity beyond which they will not interpret samples. The OSBI lab, for example, will analyze profiles with up to 6 alleles per locus, but not more. Any profile showing seven or more alleles in any locus will be deemed not suitable for interpretation. Additionally, the OSBI requires partial mixtures to have interpretable data at a minimum of six loci, and single-source minor components to have interpretable data at a minimum of four loci. Different laboratories will have different requirements for this based on their internal validation studies.

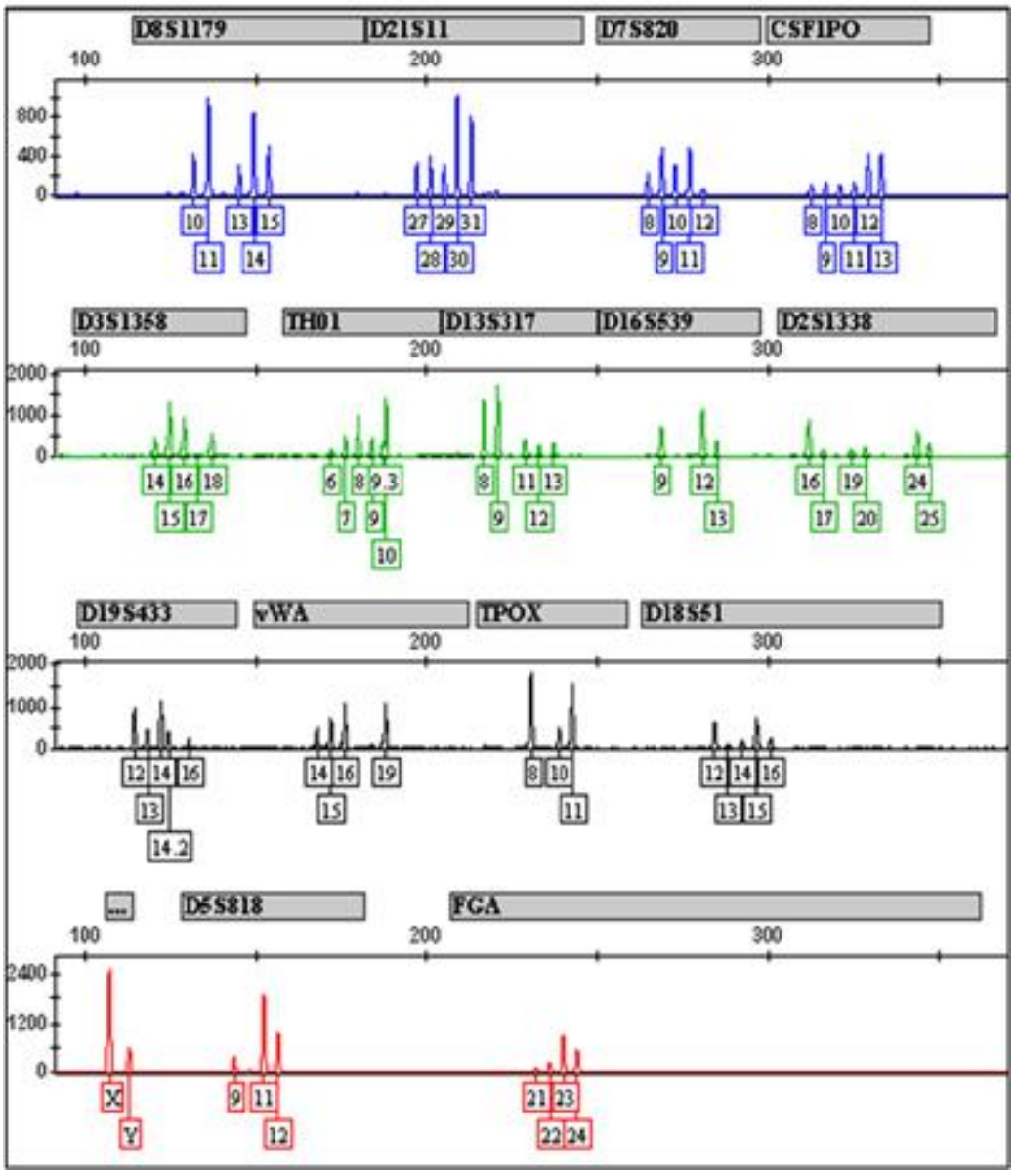
Need more help? Visit:
https://www.nist.gov/featured-stories/dna-mixtures-forensic-science-explainer

Pre-Laboratory Questions:

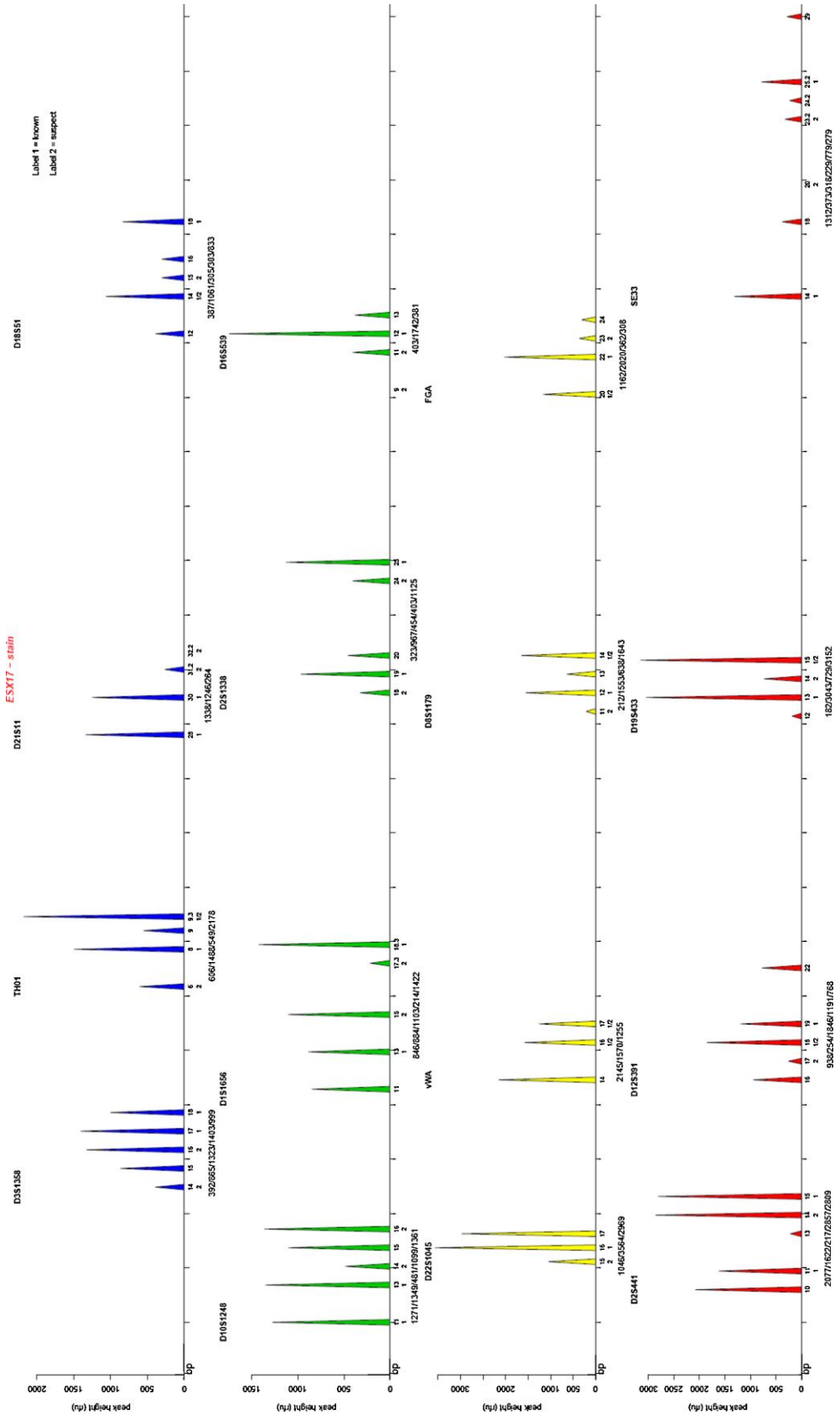
1. If you were setting a complexity threshold for your lab, what would your rule be?
2. How would you explain to a jury that you were unable to draw conclusions from a DNA sample because the mixture was too complex?
3. Describe how a complex mixture could still be useful as evidence, despite being unable to deconvolute it.

Procedure

For the following electropherograms, determine the minimum number of contributors, the type of mixture (major/minor, indistinguishable, etc.) and the possible genotype combinations.

1. 
D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
...
D5S818
FGA
100
200
300
800
400
0
2000
1000
2400
1200
X
Y

2.


Label 1 = known
Label 2 = suspect
ESX17 – stain
D3S1358
TH01
D21S11
D18S51
D1S1248
D22S1045
D10S1248
D16S539
D2S1338
D8S1179
FGA
D19S433
D12S391
D2S441
SE33
peak height (rfu)
bp

Post-Laboratory Questions:

1. How many genotype combinations are possible in a four-person mixture?
2. Imagine you are analyzing a mixed DNA profile that appears to contain genetic material from three individuals. The victim of the crime has given testimony that she was assaulted by only one individual and no other people were present. How do you attribute the third set of genetic material? How do you explain it to a jury?
3. Explain why using a composite or consensus profile would be useful. Which method would be more useful?

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to identify a complex mixture, and how to proceed with interpretation based on the different types of mixtures (major/minor, complex, indistinguishable, etc).

Laboratory 5: Statistical Weight for Single Source Profiles

Objective

The purpose of this laboratory is to learn strategies to determine statistical weight of evidence for single source DNA profiles.

Introduction

When forming a conclusion of their interpretation, analysts must be able to quantify the certainty of their answer. This certainty or "weight" of the evidence is only necessary to support a result of inclusion or any type of "cannot be excluded". Weight is provided through statistical calculations, typically in the form of a Random Match Probability, Combined Probability of Inclusion, or Likelihood Ratio. All of these methods generally evaluate the certainty that the evidentiary or unknown sample came from the reference source (or, inversely, came from a source other than the reference source). In order to determine these values, representative population databases have been established which provide known values for the rarity of alleles in major populations. The statistical combination of the alleles seen in a profile may be compared to the level of that genotype in the population to determine how likely it is that an unknown, unrelated individual contributed the sample. Cases in which involved individuals are related would not use the same calculations.

RMP

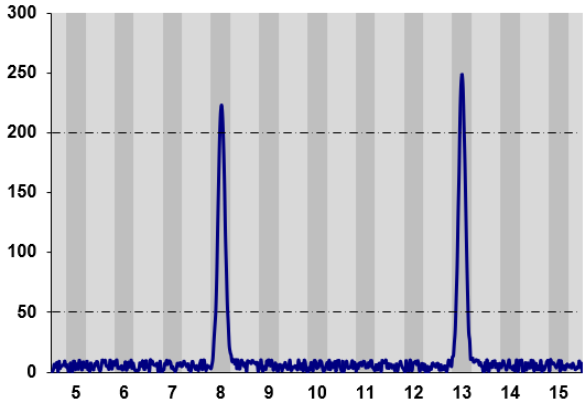
In cases of single-source samples, the calculation used to determine evidentiary weight is the Random Match Probability. In this calculation, analysts determine the frequency of the DNA profile in the major population. The underlying principle of DNA interpretation is that genotypes are rare to the point of uniqueness. Even "common" alleles in a population may form a rare profile through combination. The FBI policy of evaluating 20 loci ensures sufficient rareness in a population that analysts may determine an individualization of a person from the profile alone.

Calculating an RMP requires using the product rule to combine the frequency of each allele at each locus. Alleles are treated differently depending on if they are heterozygous or homozygous. Heterozygous alleles have two possible combinations, for example Allele A and Allele B may be A, B or B, A. Therefore, the probability of observing these alleles is 2xPxQ where P and Q represent the population allele frequencies. When interpreting homozygous alleles, on the other hand, order does not matter, so the probability of observing homozygous allele A would be P^2 where P represents the population allele frequency.

Finally, the frequency of each locus is multiplied together and inverted to reach the RMP.

Example:

For the following locus, D16S539:



Since the alleles are heterozygous, genotype frequency = 2pq
We can utilize population data (Caucasian) to determine 2(0.0180) (0.163) = 0.00587
This calculation is repeated at each locus and the results multiplied together to determine the profile frequency. The result is divided by 1 to reach the RMP.

Many laboratories will utilize a software program or a spreadsheet that incorporates the formulas and population frequency data needed to calculate the RMP. In this laboratory, we will use a freely available spreadsheet that performs this calculation, called OmniPop. OmniPop is published by NIST and is available for download from the STRBase webpage: https://strbase.nist.gov/populationdata.htm. Students are encouraged to download the program on their own devices for extra practice outside of class. OmniPop allows a user to input allele values at each locus, and the program automatically calculates the population frequency for each population group in its database.

Pre-Laboratory Questions:

1. In your own words, explain why the RMP calculation is useful.
2. Explain how population databases are formed.
3. Explain Hardy-Weinberg equilibrium. Include an explanation of the Hardy-Weinberg assumptions.

Procedure

Population Frequency:

Choose a prominent feature that you can determine from your classmates such as hair color, shirt color, type of shoe, etc. This feature will represent alleles of a gene.

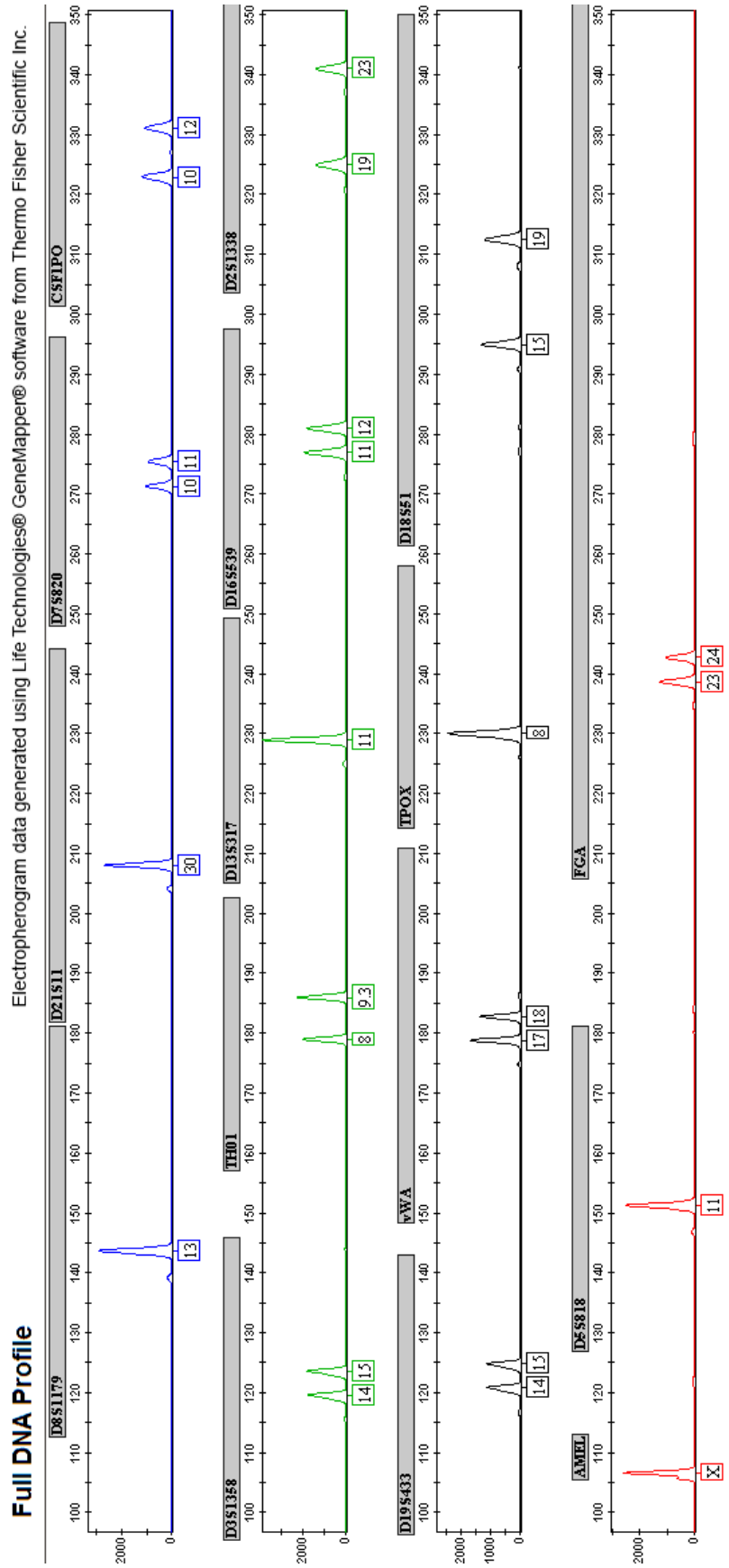
1. Gather data on the occurrence of this feature in your class population.
2. Calculate the frequency of this feature in your class population.

3. Share your results with your class
4. Now, using your class data, determine the rarity of your "genotype".

Calculating the RMP:

For the following electropherogram:

1. Determine the genetic profile for the sample.
2. Calculate the RMP using the allele population frequency from Appendix 1 of your textbook. Show your work.


Full DNA Profile
Electropherogram data generated using Life Technologies® GeneMapper® software from Thermo Fisher Scientific Inc.
D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
AMEL
D5S818
FGA

OmniPop:

1. Open OmniPop from your computer desktop. Look for an Excel Spreadsheet file called “OmniPop200.1.xls”
2. When the file opens, it will start on the “Main” sheet, indicated by the green tab at the bottom of the window. Note the other sheets next to this: “WorldMap”, “Data”, “Refs”, and “Info”. We will revisit these tabs later.
3. At the top of the “Main” sheet, you will see the name of the program and publisher. Below this is a list of DNA loci and alleles. Click on the first allele for locus D8S1179. A drop-down arrow will appear next to the cell. If you click on the arrow, it will show a list of possible alleles. By selecting different alleles from the list, you can see the graph to the right, and the values in the spreadsheet below change. This is indicating the program is recalculating the profile frequency.
4. Note the two grey rectangles below the allele list: “Show Frequencies” and “Show Map”.
5. First, click on “Show Frequencies”. This will bring you further down in the current sheet to show you the genotype frequencies for each population calculated by the database. Notice that there are many populations in this database. For the purposes of this class, we will only utilize the FBI databases for Caucasian, Hispanic, and African American. Click “Return to Top” or scroll top the top of the page.
6. Now, click “Show Map”. This will take you to the “WorldMap” sheet. A world map is displayed which indicates the most likely origins of the genotype.
7. Next, click on the “Data” tab at the bottom of the window to open the Data sheet. This page lists the allele frequencies for each locus for each population. Scroll down to find loci D2S1338 and D19S433. The original downloadable version of this spreadsheet does not contain population frequency information in the FBI database columns for these loci. They have been updated in the classroom version of the spreadsheet using the new FBI data from 2015. If you download this program to use outside of class, you will need to fill in these sections using the new FBI data.
8. The “Refs” and “Info” tabs contain the program references and description respectively.
9. Now, using the profile provided in part 2, enter the alleles into the spreadsheet.
10. Use the spreadsheet to determine the genotype frequency for the three major US populations.

Post-Laboratory Questions:

1. Why would it be useful for a professional laboratory to utilize a spreadsheet program such as OmniPop?
2. Compare the value you reached by hand and the value calculated by OmniPop. Are they the same or different? How do you account for any difference?
3. Explain why different populations would have different genotype frequencies.

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to 1) Form a population database of allele frequencies, 2) How to calculate the RMP, and 3) How to utilize a spreadsheet such as OmniPop to calculate the RMP. Include the cases in which it would be appropriate or not appropriate to use a spreadsheet for analysis.

Laboratory 6: Statistical Weight for Mixed Source Profiles

Objective

The purpose of this laboratory is to learn strategies to determine statistical weight of evidence for mixed source DNA profiles.

Introduction

mRMP

For samples in which a major and minor profile may be distinguished, the RMP calculation may be used on the individual profiles. The RMP calculation in this case is performed as on a single source profile. This approach is dependent, however, on the number of contributors and the ability to distinguish contributors.

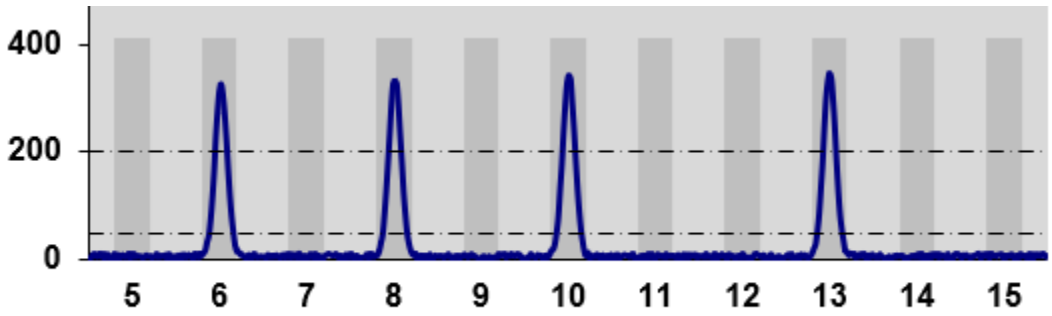
CPI

When interpreting indistinguishable mixtures, in which mixture proportions are similar, analysts may calculate exclusion or inclusion probabilities. In this random man not excluded (RMNE) approach, the combined probability of inclusion (CPI) is calculated. All genotype combinations are given equal weight in this approach. The observed alleles at a locus are summed and the result squared to reach the combination of all possible genotypes, then the individual locus values are multiplied to obtain a profile CPI. This value does not depend on any reference data, but only on the evidence. In addition, because all possible genotype combinations are considered, the calculation does not assume the number of contributors. The combined probability of exclusion (CPE) is 1-CPI. This approach is being phased out by labs and analysts because it wastes information and is generally inferior to other methods.

When using the CPI calculation, several assumptions are made that the analyst must keep in mind: no allele drop-out, all contributors are unrelated, and peak heights are irrelevant. If drop-out is suspected another approach must be taken because the CPI would not be a conservative measure for analysis.

Example CPI:

For the following locus D7S820:



CPI = 6,6 + 6,8 + 6,10 + 6,13 + 8,8 + 8,10 + 8,13 + 10,10 + 10,13 + 13,13
CPI = $(freq6 + freq8 + freq10 + freq13)^2$
CPI = $(0.00693 + 0.144 + 0.256 + 0.0346)^2$
CPI = 0.883 or 88.3% at this locus

This calculation must be repeated across all loci and multiplied to obtain the profile CPI.

LR

The most robust method used to assign weight to DNA evidence is the likelihood ratio. This method compares probabilities of two differing hypotheses for the evidence; hypotheses are exhaustive and mutually exclusive and reflect the positions of the prosecution and defense. The likelihood ratio may be thought of as a standard measure in information that summarizes in a single number the data support for a hypothesis. This approach may be difficult to explain for non-mathematicians and as a result is not yet prominent in the United States. Recent efforts to broaden its use have increased its usage in laboratories, especially with the advent of probabilistic genotyping software. In addition, for single source samples, the RMP calculation is somewhat easier to explain, even though in a single-source profile, the LR simply becomes the inverse of the genotype frequency. This method may be used for single or mixed source profiles, but for either it requires an assumption of the number of contributors. The likelihood ratio is generally more useful and conservative than other methods because it preserves more information; all available data is used, nothing is excluded from the calculation. When calculating this ratio, the result will be large (>=1) when defendant and perpetrator are same person and will be small (<1) if the defendant is not the perpetrator (Q≠K). Relative levels of LR give different weights, and value ranges are typically assigned verbal equivalencies depending on lab or jurisdiction. For example, an LR of one to ten may indicate limited support, ten to one hundred moderate support, one hundred to one thousand moderately strong support, one thousand to ten thousand strong support, and greater than ten thousand very strong support. Analysts must be wary when reporting the LR to avoid the “transposed conditional”. The transposed conditional arises when analysts incorrectly frame their conclusion as the likelihood that the suspect committed the crime, rather than the likelihood that the evidence supported the hypothesis or not. Experts recommend using alphabetical order to remember to always give E

before H when reporting, i.e. “the evidence is 50 times more likely if H_1 is correct than if H_2 is correct.”

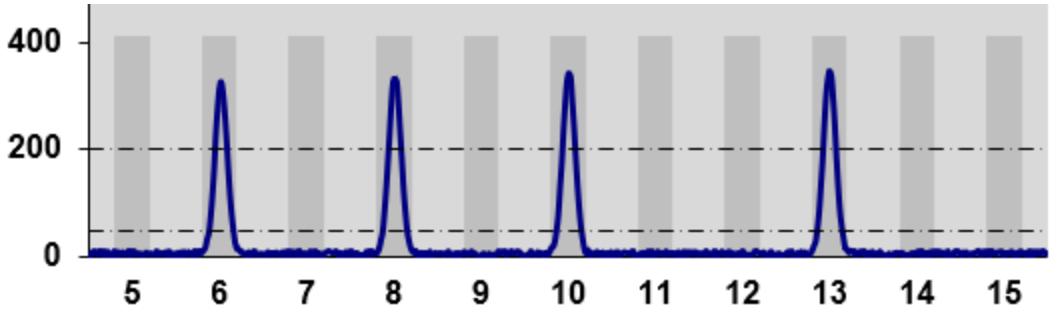
When setting up an LR calculation, analysts may understand the prosecution hypothesis (H_p) as the null hypothesis, where no difference exists between the evidence and the reference and the defense hypothesis (H_d) as the alternative hypothesis, where there is a difference between the reference and evidence.

The numerator and denominator of the LR calculation depend on the hypothesis of the prosecution or defense. If the profile has a single unknown contributor and the prosecution hypothesis posits that the defendant committed the crime, then H_p=1, meaning that the prosecution assumes 100% probability the defendant is the perpetrator. The defense in this case claims that the perpetrator is an unknown, unrelated individual, which is represented by the frequency of the given genotype in the population. The result of this calculation is the inverse of the RMP.

The LR calculation may be performed “restricted” or “unrestricted” based on the ability of the analyst to deconvolute the contributor genotypes. If one contributor (typically the victim) may be separated from the unknown profile, then the calculation may be performed only on the unknown. If there is more than one unknown contributor, however, the calculation must be unrestricted. The likelihood ratio calculation becomes more complicated when an analyst must account for more contributors to the sample, degraded DNA, drop-in, drop-out, or stochastic factors. SWGDAM interpretation guidelines provide several examples of these complex cases.

Example LR:

For the following locus D7S820:
Given victim profile 10,13



$$\frac{P(E|H_1)}{P(E|H_2)} = \frac{V+S}{V+U} = \frac{2(f_{10})(f_{13})+1}{2(f_{10})(f_{13})+2(f_6)(f_8)} = \frac{1}{2(f_6)(f_8)} = \frac{1}{2(0.00693)(0.144)} = 501$$

This calculation must be repeated across all loci and multiplied to obtain the profile LR.

SWGDAM Guidelines, Section 4B, begins p. 56, more examples:
https://docs.wixstatic.com/ugd/4344b0_50e2749756a242528e6285a5bb478f4c.pdf

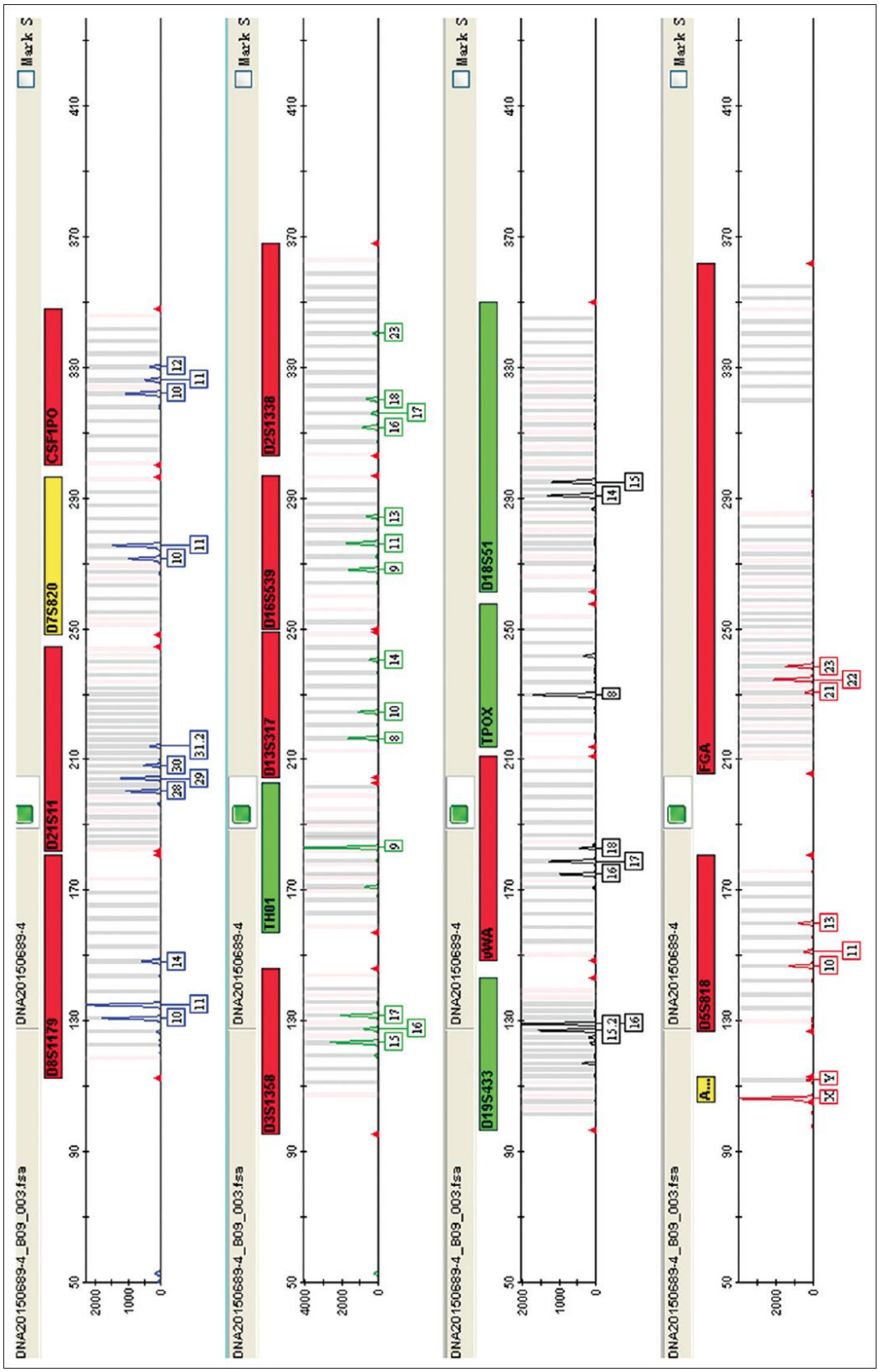
Need more help understanding likelihood ratios? Check out this presentation from Mark Perlin of Cybergenetics' TrueAllele, "Sherlock Holmes and the DNA Likelihood Ratio":
https://youtu.be/eZRRKKj_GPM

Pre-Laboratory Questions:

1. What would the LR calculation look like for a three-person mixture containing one known victim?
2. Why is the LR is becoming the standard calculation for determining weight of evidence?
3. Why should you not use the CPI if drop-out has occurred?

Procedure

For the following electropherogram, determine 1) the type of mixture, 2) the minimum number of contributors, 3) the possible genotype combinations, 4) and which of the above statistical calculations is the appropriate choice for the profile (there may be more than one correct answer, describe your reasoning). Then show how to set up both the CPI and LR calculation for the profile.


DNA20150689-4_B09_003.fsa
DNA20150689-4
Mark S
D8S1179
D21S11
D7S820
CSF1PO
D3S1358
TH01
D13S317
D16S539
D2S1338
D19S433
vWA
TPOX
D18S51
D5S818
FGA
A...
X
Y

Post-Laboratory Questions:

1. What assumptions must be made when using a LR?
2. In what scenario can you utilize a "restricted" LR? What is the advantage of restricting the analysis?
3. What does the LR actually tell you in regard to the DNA evidence?

SOP Writing Assignment:

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to determine statistical weight of evidence for mixed profiles. Include how you would go about determining the correct method to calculate statistical weight for different types of profiles.

Laboratory 7: Probabilistic Genotyping Systems

Objective

The purpose of this laboratory is to explore the use of probabilistic genotyping systems in the interpretation of forensic DNA.

Introduction

With the advent of more sensitive DNA detection methods came an increase in the issue of complex mixtures and allele drop-out. In these situations, binary interpretation models (those which use thresholds to make inclusion or exclusion decisions) fall short of providing a conservative interpretation result. Probabilistic genotyping systems (PGS) are the method recommended by SWGDAM and others as the best way to handle mixture interpretation and drop-out. There are two types of PGS approaches: Semi-continuous and fully continuous.

Semi-continuous

This model, also called discrete, focuses on the alleles present without accounting for peak heights. This approach is somewhat simplified, and does not make full use of the data present by examining peak height ratios, stutter percentages, etc. However, these limitations make these PGS programs much easier to program, easier to understand, and faster. Like fully continuous methods, the likelihood ratio is used to determine statistical weight of evidence.

One such semi-continuous PGS software program is called Lab Retriever, which will be used in this laboratory. Lab Retriever calculates likelihood ratios for samples with the ability to account for probability of drop-out and for assumed or suspected contributions. Therefore, the LR is easily calculated for different hypotheses of the case. Lab Retriever is freely available online, and students are encouraged to download it for practice outside of class from the website: https://scieg.org/lab-retriever/. The manual may be downloaded from the same page. Based on the likeLTD program published by Dr. David Balding, the Lab Retriever software streamlines the software and incorporates a graphical user interface, which makes it more accessible for DNA analysts who are not familiar with computer software programming.

Fully Continuous

The fully continuous model is advocated by sources in the literature as the best and most conservative method of PGS analysis. This approach utilizes all available data from a profile, including mixture ratios, peak heights, etc. In this method, the Markov-chain Monte Carlo algorithm is used to formulate simulations to fit the observed data to the best set of potential genotype combinations. The programming of fully continuous PGS software is much more involved and difficult for the uninitiated to understand. However, DNA analysts should be aware of the underlying DNA principles. Lack of understanding of these software systems can lead to erroneous interpretation of evidence.

An example of issues arising from lack of understanding of this software can be seen in the case of the Forensic Statistical Tool (FST). This proprietary software was created and used by the New York Office of the Chief Medical Examiner for use in analysis of DNA evidence. It was used in the analysis of evidence from approximately 1350 cases before concerns were raised that the software was not giving correct results. Eventually a judge ordered the release of the software code, and it was found to contain mathematical errors. At this time, every case which involved the software is being reexamined. Other software programs, such as TrueAllele, currently do not release the program code. Some sources in the literature advocate the release of proprietary source code for any software program used in this type of evidence analysis.

Pre-Laboratory Questions:

1. Why would it be good for the source code for forensic software programs to be available to the public?
2. Describe the possible advantages or disadvantages to using PGS software.
3. Lab Retriever is a semi-continuous PGS program. What does this type of program do? How is a fully continuous PGS different?

Procedure

1. Open Lab Retriever on your computer.
2. Notice the panel on the left side of the window which contains customizable parameters. P(DI) is 1 by default, however P(DO) is empty and must be filled.
3. Probability of Drop-out (P(DO) is calculated by using the average peak height of the profile and logistic regression of population data based on the kit being used. For the purposes of this class, P(DO) values will be provided with each profile.
4. Select the relevant hypothesis using the H1 and H2 drop-down lists.
5. Click “Load a file” near the top of the window. We will select the .csv files containing the profile we want to analyze. The file must contain the Q data and any available K data.
6. After loading the file, we will select the detected profile, assumed profile of any known contributors, and suspected profiles of any potential or suspected contributors.
7. Once all profiles have been selected, click “RUN!” above the profile table to calculate the LR.
8. Use Lab Retriever to determine the LR for the following profiles contained by .csv files:
    a. Single.csv
    b. 2-Person.csv
    c. 5-Person.csv
9. After you have run each file in Lab Retriever, save your results to a .csv file using the “Save” button. These files may be opened in Excel to print the LR results.

Post-Laboratory Questions:

1. How would you explain a resulting LR to a jury? Describe how the LR supports the evidence.
2. In your own words, how would you describe the function of a PGS program to a jury?
3. Would a PGS utilize a stochastic threshold? Why or why not?

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to utilize probabilistic genotyping systems such as Lab Retriever to assist in DNA analysis interpretation. Include the cases in which it would be appropriate or not appropriate to use a PGS for analysis.

Laboratory 8: Validating Methods

Objective

The purpose of this laboratory is to gain experience in the validation process used in many DNA laboratories, as well as other types of laboratory settings.

Introduction

Validation is defined by SWGDAM guidelines as "a process by which a procedure is evaluated to determine its efficacy and reliability to forensic casework and/or database analysis". Any method used in the process of evidence analysis requires validation. Techniques requiring validation include, but are not limited to, DNA extraction, PCR amplification, and probabilistic genotyping software. SWGDAM has published separate validation guidelines for the adoption of probabilistic genotyping software, due to the increased complexity of these programs. The validation process consists of two steps: developmental validation and internal validation. Developmental validation is performed on a new technique to determine the conditions and limitations of the method on different types of data (that is forensic knowns or unknowns, databases or casework samples). Additionally, developmental validation must utilize published, peer-reviewed scientific principles. Following developmental validation, laboratories may adopt techniques by performing their own internal validation studies to demonstrate that the new method performs as expected. Internal validations are useful for establishing analysis parameters such as thresholds, instrument sensitivity and precision, and mixture interpretation guidelines.

In the US, both developmental and internal validation of forensic DNA analysis techniques must follow the FBI Quality Assurance Standards and the SWGDAM Validation Guidelines for DNA Analysis Methods. These documents outline the process for performing validations and the requirements for doing so. Developmental validation must include characterization of genetic markers, species specificity, sensitivity studies, stability studies, precision and accuracy of the assay (including repeatability and reproducibility), case-type samples, population studies, mixture studies, and PCR-based studies. Similarly, internal validations should include known and mock evidence samples, sensitivity and stochastic studies, precision and accuracy, mixture studies, and contamination assessment.

Once a validation has been completed, the instrument or technique must only be used within the confines of the validation study. For instance, if a software program is adopted by a lab and the internal validation approves it for use with only single-source samples, then it may not be used for mixed-source samples, regardless of the software capabilities. Modification of a process, whether it consists of movement or routine maintenance of an instrument or adjustment of PCR parameters, requires additional validation of the adjustment.

<u>Pre-Laboratory Questions:</u>

1. What is the purpose of method validations?
2. Describe what is established in the internal validation process.
3. Imagine that the capillary electrophoresis instrument in your lab was knocked over and broken. It is replaced by a new instrument of the same model. Does your lab need to perform a new internal validation? Why or why not?

Procedure

Choose two labs from this manual and write a procedure to validate the method you would use. Be sure to include information on all the involved variables such as what types of samples and how many you would use.

Post-Laboratory Questions:

1. In a working DNA lab, list the techniques or instruments that might require validation.
2. Imagine you are working as a DNA analyst and you are tasked with interpreting an evidentiary DNA sample. The first electropherogram you generate shows several instances of drop-out. You think you can get a better outcome by re-amplifying the sample with more PCR cycles, but that is outside your lab's internal validation. How do you proceed?

SOP Writing Assignment

Using the information provided and the exercises from this lab, write a step-by-step procedure outlining how to perform an internal validation of a new technique being introduced to your lab.

Additional Laboratory Exercises

The following exercises consist of short lessons and exercises exploring other areas of forensic DNA analysis. These short labs may be used to supplement other lessons or to fill extra class time.

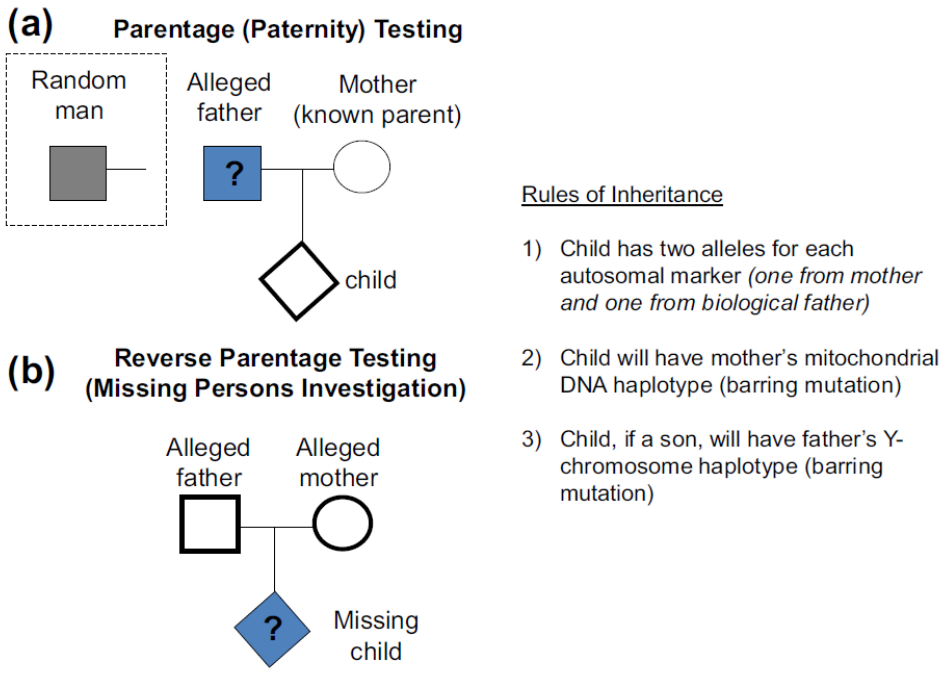
Familial Matching

Objective

In this exercise, we will learn how to determine relatedness and how relatedness may be used in an investigation.

Introduction

The analysis methods discussed so far in this course have all relied on the assumption that the individuals involved are unrelated. Several cases exist in which related individuals would be tested, including parentage testing, missing persons cases, mass disasters, immigration cases, and familial searching of DNA databases. In any type of case, analysts should be careful to 1) ask the correct question and 2) keep in mind the relationship between the individuals involved. In this laboratory, we will focus on cases of establishing parentage. Two cases are visualized below:



Butler's Interpretation, Chapter 14: Relationship Testing, Kinship Statistics

When handling kinship cases, analysts will use a concept of Identical-by-Descent (IBD) alleles. The probability of IBD alleles may be estimated by knowing the relationship of the individuals. An individual shares two IBD alleles with him/herself, one IBD allele with a parent/child, and zero with an unrelated person. Other relationships have different fractions of relatedness: full siblings, for example, have ¼ probability of sharing zero or two alleles, and ½ probability of sharing 1 allele. These known probability values may be used to establish a conditional probability between the two potential relatives and to perform a statistical calculation. The likelihood ratio is used to express the likelihood of the evidence given two differing hypotheses. The likelihood ratio in relationship cases is also called the relationship index. In paternity or maternity cases, the LR is referred to as the paternity or maternity index, respectively. A likelihood ratio may be determined by comparing the conditional probability that the person of interest is the father versus a random man from the same genetic population is the father.

In parentage-determination cases, analysts may have access to DNA from the child, mother, and alleged father. However, this is not always the case. “Deficient” paternity testing may occur when either the mother or father are unknown. Cases where all three family profiles are available—mother, potential father, and child—are ideal because they provide “obligate” alleles, which have a known source in one parent. Backing out the obligate alleles facilitates discernment of the other parent’s genotype. If the mother is unknown, however, it is difficult to determine obligate alleles, thereby complicating analysis and possibly leading to incorrect results (false inclusions or exclusions). The discrimination power of paternity testing is greatly reduced if the mother’s DNA is not available and should be avoided whenever possible according to the AABB (AABB, 2010).

There are 21 different explanations for a parentage profile. The table below outlines each scenario:

#	G_C	G_M	G_{TM}	Numerator (X)	Denominator (Y)	Paternity Index (PI)	PI if p=0.134, q=0.170	Random Man Not Excluded (RMNE)	RMNE if p=0.134, q=0.170	#
1	PP	PP	PP	1	p	1/p	7.46	p(2-p)	0.250	*1*
2			PQ	1/2	p	1/2p	3.73	p(2-p)	0.250	*2*
3			QR	0	p	0	0	0	0	*3*
4		PQ	PP	1/2	p/2	1/p	7.46	p(2-p)	0.250	*4*
5			PQ	1/4	p/2	1/2p	3.73	p(2-p)	0.250	*5*
6			PR	1/4	p/2	1/2p	3.73	p(2-p)	0.250	*6*
7			QR	0	p/2	0	0	0	0	*7*
8	PQ	PP	QQ	1	q	1/q	5.88	q(2-q)	0.311	*8*
9			PQ	1/2	q	1/2q	2.94	q(2-q)	0.311	*9*
10			QR	1/2	q	1/2q	2.94	q(2-q)	0.311	*10*
11			RS	0	q	0	0	0	0	*11*
12		PQ	PP	1/2	(p+q)/2	1/(p+q)	3.29	(p+q)(2−p−q)	0.516	*12*
13			PQ	1/2	(p+q)/2	1/(p+q)	3.29	(p+q)(2−p−q)	0.516	*13*
14			PR	1/4	(p+q)/2	1/[2(p+q)]	0.608	(p+q)(2−p−q)	0.516	*14*
15			QR	1/4	(p+q)/2	1/[2(p+q)]	0.608	(p+q)(2−p−q)	0.516	*15*
16			RS	0	(p+q)/2	0	0	0	0	*16*
17		QR	QQ	0	p/2	0	0	p(2−p)	0.250	*17*
18			PQ	1/4	p/2	1/2p	3.73	p(2−p)	0.250	*18*
19			QR	0	p/2	0	0	p(2−p)	0.250	*19*
20			QS	0	p/2	0	0	p(2−p)	0.250	*20*
21			RS	0	p/2	0	0	0	0	*21*

Butler's Interpretation, Chapter 14: Relationship Testing: Kinship Statistics

Example:

Given a child (C), mother (M), and alleged father (AF), with alleles at STR locus D18 as follows:

M: 14,15 C: 14,14 AF: 14,16

Since the child is homozygous, we know that both the mother and father must have that allele. The alleged father does in fact have allele 14, therefore that allele's frequency will be used to calculate PI at this locus.

PI = (1/4)/(p/2) = 1/2p

The PI will be calculated at each STR locus and the values multiplied to form the combined paternity index.

Procedure

Explain how you would set up and utilize a LR calculation in a case where relatedness was possible. Include in your explanation a discussion of IBD and different levels of relatedness that could be present.

Communicating Laboratory Results

Objective

The purpose of this laboratory is to gain experience communicating the results of DNA analysis and interpretation.

Introduction

Forensic DNA analysts must be able to effectively communicate the results of their analysis both on paper for review by peers and verbally in a courtroom setting. In the process of DNA analysis, the workflow consists of: analysis of evidentiary samples, comparison to known/reference samples, determine conclusion of analysis, formulate report, technical review of report (by peer or colleague), administrative review of report (by technical reviewer or supervisor), submission of report to law enforcement, review of report by prosecution and defense, presentation of conclusions in court. In this process, the report will be used and reviewed by at least six entities (two colleagues, prosecution/defense, and judge/jury). Therefore, it is important for the report to be clear, concise, and correct.

When writing reports, DNA analysts must keep in mind the question they are attempting to answer and the potential answers to the question. Analysts must also avoid common pitfalls when forming a conclusion. For instance, DNA analysis never determines guilt or innocence of a suspect. Instead, DNA analysis is forming a conclusion about a piece of evidence and its possible attribution. Very precise wording is therefore required to prevent analysts misrepresenting their data or conclusions. Statistical calculations are included in “inclusion” conclusions to provide weight to the evidence being presented.

A general analysis report follows the format of: 1) Case identifier, 2) description of the technology and amplification system used, 3) results and conclusions with a quantitative and qualitative interpretation statement, 4) date of issuance of report, 5) disposition of evidence, 6) signature of the report writer and reviewer(s). This formatting both meets FBI quality assurance standards and adequately represents the work performed and conclusions made by the analyst.

When forming conclusions, analysts should defer to their laboratory SOP for the specific wording to be used in a report. For the broad conclusion of “inclusion”, the evidence may be described by several phrases: “match” when there exists agreement at every locus between the Q and K samples, “consistent with” where a Q profile is partial, but agrees with K at the observed genotypes, or “cannot be excluded” where alleles from K are present in a Q mixture. Other conclusions that may be drawn are “exclusion” when the questioned profile does not match any potential reference samples, “inconclusive” when the profile cannot be clearly included or excluded from the reference profile, or “unsuitable for analysis” if the evidence sample is a complex mixture or too degraded to have evidentiary value.

Procedure

For the following profiles, determine what conclusion you would reach as a DNA analyst. Write a concluding statement that you might include in an analysis report for each profile.

1.

Locus:	Q	K
CSF1PO	11, 12	11, 12
FGA	19, 21	19, 21
THO1	7, 8	7, 8
TPOX	9, 12	9, 12
VWA	15, 17	15, 17
D3S1358	13, 13	13, 13
D5S818	9, 12	9, 12
D7S820	8, 8	8, 8
D8S1179	10, 13	10, 13
D13S317	11, 13	11, 13
D16S539	9, 11	9, 11
D18S51	11, 15	11, 15
D21S11	28, 30.2	28, 30.2
D1S1656	13, 16	13, 16
D2S441	11.3, 14	11.3, 14
D2S1338	17, 25	17, 25
D10S1248	11, 15	11, 15
D12S391	15, 16	15, 16
D19S433	12, 13.2	12, 13.2
D22S1045	12, 15	12, 15

2.

Locus:	Q	K
CSF1PO	11, 12	11, 12
FGA	19, 21	19, 21
THO1	7, 8	7, 8
TPOX	9, 12	9, 12

VWA	15, 17	15, 17
D3S1358	12, 13	13, 13
D5S818	9, 12	9, 14
D7S820	8, 8	8, 8
D8S1179	10, 13	10, 13
D13S317	11, 13	11, 13
D16S539	9, 11	10, 11
D18S51	11, 15	11, 15
D21S11	28, 30.2	28, 30.2
D1S1656	13, 16	13, 16
D2S441	11.3, 14	11, 14
D2S1338	17, 25	17, 25
D10S1248	11, 14	11, 15
D12S391	15, 16	15, 16
D19S433	12, 13.2	12, 13.2
D22S1045	12, 15	12, 15

3.

Locus:	Q	K
CSF1PO		11, 12
FGA	19, 21, 22	19, 21
THO1	8	7, 8
TPOX	9, 12	9, 12
VWA	15, 17	15, 17
D3S1358		13, 13
D5S818		9, 12
D7S820	8	8, 8
D8S1179	10, 13	10, 13
D13S317		11, 13
D16S539	9	9, 11
D18S51	11, 15	11, 15
D21S11		28, 30.2
D1S1656	13, 16	13, 16

D2S441	12, 13, 14	11.3, 14
D2S1338	17, 25	17, 25
D10S1248	11	11, 15
D12S391	15	15, 16
D19S433		12, 13.2
D22S1045	12, 15	12, 15

Wildlife Forensics

Objective

The purpose of this laboratory is to learn about the applications of forensic DNA analysis in a poaching investigation.

Introduction

A less known but growing field of forensic science is Wildlife Forensics. This discipline is related to the investigation of illegally obtained or traded animal and plant products. DNA analysis is used in this field to determine the geographic or biological source of seized goods, such as elephant ivory or tiger skins. Research is actively being performed to collect data on endangered or protected wildlife species so that species can easily be genetically identified.

This laboratory will use an online interactive activity from HHMI BioInteractive to learn about wildlife forensics. The activity consists of a series of videos and worksheets in which you will learn about how DNA analysis is applied in the field of wildlife forensics.

CSI Wildlife Interactive Activity:
https://www.biointeractive.org/classroom-resources/csi-wildlife

Procedure

1. If you are performing the activity in class, launch the BioInteractive CSI Wildlife program from your computer. (If you are using your own computer, follow the link above to download and install the Desktop App for Windows or Mac OS, and then launch the program.
2. Obtain Worksheet #1, Student Supplement, and Worksheet #2. Follow directions on the worksheets to guide you through the interactive program.

Sources

1. AABB. (2010). Annual Report Summary for Testing in 2010: Prepared by the Relationship Testing Program Unit. Available at http://www.aabb.org/sa/facilities/Documents/rtannrpt10.pdf.
2. Bieber, F. R., et al. (2016). "Evaluation of forensic DNA mixture evidence: protocol for evaluation, interpretation, and statistical calculations using the combined probability of inclusion." BMC Genet **17**(1): 125.
3. Butler, J. M. (2011). Advanced Topics in Forensic DNA Typing : Methodology, Elsevier.
4. Butler, J. M. (2014). Advanced Topics in Forensic DNA Typing: Interpretation, Elsevier.
5. Evett, I. W. and B. S. Weir (1998). Interpreting DNA Evidence: Statistical Genetics for Forensic Scientists, Sinauer Associates Inc.
6. Inman, K., et al. (2015). "Lab Retriever: a software tool for calculating likelihood ratios incorporating a probability of drop-out for forensic DNA profiles." BMC Bioinformatics **16**: 298.
7. Investigation, O. S. B. o. (2018). Forensic Biology Unit Policy Manual. Oklahoma.
8. Methods, S. W. G. o. D. A. (2010). SWGDAM Interpretation Guidelines for Autosomal STR Typing by Forensic DNA Testing Laboratories.
9. Methods, S. W. G. o. D. A. (2016). Validation Guidelines for DNA Analysis Methods. SWGDAM.
10. Perlin, M. (2010). Explaining the Likelihood Ratio in DNA Mixture Interpretation. Promega's Twenty First International Symposium on Human Identification, Pittsburgh, PA.
11. (2011). Quality Assurance Standards for Forensic DNA Testing Laboratories. D. A. Board.
12. (2015). "STRBase Population Data." 2019, from https://strbase.nist.gov/populationdata.htm.