Development and Implementation of Forensic Science Research and Training Programs at the University at Albany's Northeast Regional Forensic Institute (NERFI)

Authors: Donald D. Orokos, Ph.D. (Forensic DNA Training, and Leadership Assessment Workshops)
Address: University at Albany / Biology 112
Phone: 518-442-4308
Fax: 518-442-4767
E-mail: orokos@albany.edu

John Hicks, MPA (Forensic DNA Training and Leadership Assessment Workshops)
Address: University at Albany / NERFI: Biology 225
Phone: 518-437-3791
Fax: 518-442-4767
E-mail: jhicks@uamail.albany.edu

Igor Lednev, Ph.D. (Raman Method Validation)
Address: University at Albany / LSRB 1107
Phone: 518-591-8863
E-mail: lednev@albany.edu

Ron Stevens, (Forensic Science Impact in the Court Room)
Address: University at Albany / Biology 225
Phone: 518-860-6425
E-mail: ronaldstevens@mac.com

Tomek Strzalkowski, Ph.D. (ChatMinder: A Safe Internet Tool for Parent)
Address: University at Albany / SS 262B
Phone: 518-442-2608
E-mail: tomek@albany.edu

Sanjay Goel, Ph.D. (Capitol Region Cyber Crime Partnership)
Address: University at Albany / BA 310B
Phone: 518-442-4925
E-mail: goel@albany.edu

Submission Date: August 31, 2010

Award Number: 2008-DD-BX-K301

This project was supported by Grant No. 2008-DD-BX-K301 awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. Points of view in this document are those of the author and do not necessarily represent the official position or policies of the US Department of Justice.

TABLE OF CONTENTS

ABSTRACT

The Northeast Regional Forensic Institute (NERFI)-instructed ABI 3130xl™, ABI RTPCR™, and Leadership Assessment workshops were attended by forensic scientists from across the United States. Course evaluations were completed and indicate a favorable assessment of the workshops content, the instructors, mode of instruction, and NERFI facilities and support staff. Offering travel and per diem for all participants, linking NERFI's web site with NIJ's web site, and meeting with various directors and supervisors from the forensic community provided the impetus on educating 104 trainees in DNA Analysis and Leadership Assessment skills.

Near-infrared (NIR) Raman spectroscopy was used to measure spectra of dried human semen samples from multiple donors. The major chemical components that contributed to the Raman spectrum of semen were determined and used to identify the principal spectral components. Advanced statistical analysis of spectra obtained from multiple spots on dry samples showed that dry semen is heterogeneous and its Raman spectra could be presented as a linear combination of a fluorescent background and three spectral components. The relative contribution of each of the three components varies with donor, so no single spectrum could effectively represent an experimental Raman spectrum of dry semen in a quantitative way. The combination of the three spectral components could be considered to be a spectroscopic signature for semen.

The Judicial Awareness of Forensic Science project evaluated the use, impact, and effectiveness of forensic evidence in the courtroom. The preliminary results of this review indicate that forensic evidence definitely played a crucial role in some of the cases and investigations reviewed under this project. In addition, eighty first responders were trained on the proper collection, handling and storage of biological evidence. Furthermore, funding from this award was used to purchase equipment from Porter-Lee and update existing technology to maximize the ability of law enforcement to process crime scenes in the most efficient and effective manner and to permit real-time access to uniform evidence inventory and management files by the District Attorney as well as the police department..

The objective of the ChatMinder project was to conduct a study of the dialogues occurring in the on-line chat rooms. A secure chat-room at the ILS Institute was established and chat data collected from experiments from recruited subjects (SUNY students under an IRB protocol). Over 20 hours of chat involving groups of 3 to 6 people on topics ranging from movies to organic food to state of the economy was collected for this study. VCA technology represents an important advance in automated human-computer communication with potential applications in cross-cultural social modeling, influence operations, advertising, law enforcement, and national security.

Analyzing crime data to make defensible judgments has become increasingly difficult due to the disparate sources of data - each element of which paints only a part of the complete picture. In addition, data can be both structured (log files, spread sheets, and databases) and unstructured (plain text, web pages, word documents etc.). Amalgamating all the evidence into a cohesive stream of evidence requires analysis capability that not

only handles varied data sets but was able to find correlations among the data. Under the Capitol Region Cyber Crime Partnership section of this award, a suite of tools was employed to collect data from open sources, analyze text data, and correlate information from multiple sources including sexual predator behavior and detecting attacks on the network.

EXECUTIVE SUMMARY

As described in the topical headings below, the overall goal of the National Institute of Justice Congressionally-directed award (2008-DD-BX-K301) was to provide the University at Albany's Northeast Regional Forensic Institute (NERFI) funds for forensic science research, and to develop and implement various professional development training programs for the criminal justice community. Most of the research described in this report was funded at levels only to explore novel concepts in the forensic sciences. The majority of funds from this grant were used to support cost-free instruction and travel in state-of-the-art DNA Analysis Technology and Managerial Assessment workshops. In view of the diverse nature of the individual activities carried out under this project, each activity is characterized herein as a "Module".

Module: Forensic DNA Training, Leadership Assessment and Research

- A: ABI 3130xl™ Training (Appendix 1)
  - Three separate weeks of lecture plus hands-on
    - 6-8 trainees per session
- B: ABI RTPCR™ Training (Appendix 7)
  - 2 - 2 ½ day training sessions
    - At least 6 trainees per session
- C: Leadership Assessment Workshops for Managers (Appendix 13)
  - 2 - 2 day training sessions
    - 6-10 trainees per session
- D: Raman Method Validation Applied to Biological Evidence

Module: Judicial Awareness of Forensic Science

- Determine the impact of new forensic technologies in the court room
- Identify forensic evidence applications in Schenectady County criminal cases
- Train Schenectady law enforcement first responders in the recognition, collection, and preservation of forensic evidence (Appendix 17)
  - Approximately eighty (80) officers were trained in groups of 20
- Set up a system from Porter-Lee to bar-code evidence in the Schenectady, New York DA's Office and to permit real-time tracking for evidence analysis, inventory control, and disposition purposes in a shared system with the Schenectady PD.

Module: ChatMinder: A Safe Internet Tool for Parents

Graduate level assistance is needed to perform:

- Design and apply new advanced language and information technologies developed at the University at Albany for assisting automated intelligence gathering and analysis to detect and prevent of online crime against children as well as other persons or organizations.

Module: Capitol Region Cyber Crime Partnership

- Design and apply new Cyber Crime tools to identify and prevent sexual offenses against children on the internet.

Module: Forensic DNA Training, Leadership Assessment Workshops, and Research

The forensic science community needs structured training and research in an academic environment. It is important to note, that over 700,000 sworn police officers have available 632 police academies in the United States (Hickman, 2002). The academies are collectively expending over $725,000,000. The Alcohol Tobacco and Firearms (http://www.atf.treas.gov/training/arsonex.htm) funds (except travel and per diem) an eleven month firearms academy and numerous other short courses. The time is now to properly support forensic DNA training programs.

NERFI has a proven record of providing structured learning in an academic environment. NERFI was conceived to address a similar need for structured forensic laboratory training developed with graduate level academic standards delivered in an intensive 16-week program in lieu of inefficient, unstructured mentor-based training programs traditionally afforded newly hired forensic laboratory personnel in state and local crime laboratories. NERFI's five-year record of accomplishment for delivering graduate level forensic training for DNA casework and databank scientists is unique and well established for expansion. New York State DNA legislation (Article 49-B – NYS Executive Law) was enacted in 1994 and then expanded in 1999, 2004 and 2006 increasing designated offenses from serious felonies to all felonies and selected misdemeanors. This unprecedented expansion led to the signing of a Memorandum of Understanding in 2000 between the University at Albany (UA) and the New York State Police (NYSP) to collaborate on forensic science work force development and research programs. Accordingly, in 2002 the UA and NYSP developed a Graduate Program in Forensic Biology (FB) which now matriculates 6-12 students per year and received the full accreditation from the Forensic Education Program Accreditation Commission (FEPAC) in 2008.

NERFI's programs provide a long term investment in forensic laboratories' most valuable assets: forensic scientists. Structured workshops providing forensic scientists professional development is a necessity in the forensic science community. Most forensic laboratories have little or no training funds and view training as a non-reoccurring cost / expenditure instead of an *investment* in the most significant asset (human resources) in their organizations. Laboratories need training for new and experienced employees in all disciplines and all job titles. In addition, complimentary skills, such as presentation of results, team work, communication (verbal and written), and testimony are also the major focus of in-house training after hiring. Lack of access to efficient training programs and resources can also be seen as a contributor to delays in analyses or backlogs in DNA analyses of criminal cases. Local and federal funding must be utilized to stabilize these education and training programs for all disciplines.

Governmental agencies will need to invest in future criminal justice training to keep abreast of new technology issues related to DNA analyses and CODIS data base needs, and provide the necessary funds to advance basic and applied forensic science research.

In doing so, the following training and research programs were included in award 2008-DD-BX-K301: Module: Forensic DNA Training, Leadership Assessment Workshops, and Research.

A: Capillary Electrophoresis & Data Analysis Using Applied Biosystems' ABI™ 3130xl Genetic Analyzer & GeneMapper® ID Course (3 one-week laboratory workshops with 6-8 trainees per session):
This 5-day course was designed to provide the theoretical and practical background necessary to perform capillary electrophoresis and data analysis. Today's advanced technology has led to an exponential number of cases being submitted to the crime lab for DNA testing. For this technique to be successful, it is imperative that the biological evidence be processed and analyzed effectively. A total of twenty-three DNA Analysts attended three different ABI 3130xl™ workshops.

B: ABI™ RT-PCR training Real-Time PCR Using Applied Biosystems' 7500 Instrument and Chemistries (2 – 2 ½ day with at least 6 trainees per session):
This 2.5 day course was designed to provide the theoretical and practical background necessary to perform quantitative PCR and data analysis. This course also included a multiple-choice exam and issuance of completion certificates. NERFI originally proposed at least six trainees per training session. However, a total of sixty-two DNA Analysts attended the 2-2½ day training sessions on Real-Time PCR Using Applied Biosystems' 7500 Instrument & Chemistries.

C: Leadership Assessment Workshop for Managers, Supervisors and Directors:
NERFI provided two 2 - day workshops for managers, supervisors, and directors in human resource strategies to help increase efficiencies in selection, retention and promotion of forensic scientists and managers. The first group of forensic managers attending *Leadership Assessment: Developing the Next Generation of Leaders* in January, 2010, included three lab directors, an assistant director, three managers, and a lab supervisor, two technical managers, and a senior scientist from various forensic sections – crime scene, biology, latent prints, drug analysis, trace, and QA. The second group of forensic managers attending *Leadership Assessment: Developing the Next Generation of Leaders* in February, 2010, included one director, a forensic science coordinator, three supervisors, a senior scientist, two technical leaders, a quality manager, and a latent print examiner with no management experience. Professionals from the forensic science fields of DNA, drug chemistry, crime scene, firearms and toolmarks, QA and latent prints were represented in this session.

The Leadership Assessment Workshops focused on developing both strategic and behavioral skills for managing employees. These leadership assessment workshops provided participants tools to increase efficiency in selection and retention of new hires and internal promotion decisions, thereby reducing costly turnover. Turnover costs can exceed twice the annual salary plus benefits of the employee (Cascio, 2000).

The leadership assessment workshops also provided an introduction to behavioral observation techniques. This can be used by forensic managers, supervisors and directors

to enhance their skills in the effective assessment of employees. The workshop used managerial exercises with behavioral examples to familiarize participants with the assessment concept. There was an enthusiastic and strong positive response by all participants in these workshops and a high demand for more of these kinds of sessions (Appendix 16).

D: Raman Method Validation—The use of Raman Technology to Detect Human Biological Fluids: The identification and characterization of body fluids and stains discovered at a crime scene is a major part of forensic investigation today. The three most common fluids found are blood, semen, and saliva, and there are several methods used currently to distinguish one from another. Blood can be presumptively tested for using different color spot tests, but these tests are destructive to the sample and can also have false positives (Siegel, 2000). If only a small amount of sample is available, careful decisions must be made as to whether the presumptive test is necessary. There are also confirmatory tests for blood that conclusively prove blood is present, and some of these tests can distinguish between species. Semen is similar in that there are destructive presumptive tests as well as confirmatory tests. Saliva, however, has no confirmatory tests. So, an examiner can never be positive about the presence of saliva (Siegel, 2000). Most presumptive tests can be performed in the field, but some sample preparation such as extraction is often necessary. Most confirmatory tests must be done in the laboratory, so forensic experts responding at a crime scene will not know the confirmed identity of a fluid until much later on. The largest problem with these tests is the consumption of the sample. Sometimes a case can be broken with just the smallest amount of biological evidence, so it is crucial that these small quantities are examined as efficiently as possible and nondestructively at the crime scene. Another issue is the ambiguity of the tests. Current simple on-field screening tests cannot confirm the presence of a particular fluid, and saliva can never be confirmed. Finally, mixtures of fluids are frequently found, and this can make identification and subsequent DNA analysis even more difficult. *The forensic community is in need of a reliable and ultimately portable method that can exclusively distinguish between the common and uncommon body fluids, as well as not destroy the sample in the process.*

Long-term goal is to develop an intelligent and user-friendly methodology with in-the-field capabilities based on Raman spectroscopy for characterizing traces of body fluids at a crime scene. Our hypothesis is based on the fact that every body fluid has a unique composition and should have a unique Raman spectroscopic signature, which can be used for its identification. Recently, Virkler and Lednev reported that Raman spectroscopy can be potentially used to distinguish different body fluids (Virkler and Lednev, 2008) as well as provide non-destructive, confirmatory identification of body fluids at the scene of a crime (Virkler and Lednev, 2009). However, this analysis was carried out on only one sample of each body fluid and did not take into account any variations that might occur between different donors of the same fluid. Since each donor's sample is heterogeneous within itself due to many different chemical components, we would also like to investigate the effect these chemical components have on the Raman spectral components of a body fluid. The main purpose of this project was to understand the role of heterogeneity within a sample as well as among multiple donors for human semen.

This project was designed to determine the level of spectral heterogeneity of human semen based on principal components and to find out how much variation there is in the spectra from different donors. If there is very little change in the spectrum from one donor to another, then the technique of Raman spectroscopy can be considered to be reproducible in identifying a sample to be semen-based on the application of a calculated spectroscopic signature. This signature, which could be fitted to a semen sample collected from any donor, could be produced based on several spectral components found in semen that are present due to the heterogeneous distribution of the many chemical species in semen. A unique signature can ultimately be developed for other body fluids as well so that an unknown body fluid discovered at a crime scene could potentially be identified in a confirmatory manner.

A spectroscopic signature for human semen was developed based on the heterogeneous chemical composition of semen using NIR Raman Spectroscopy. A set of 50 semen samples were obtained from anonymous donors at an in-vitro fertilization clinic. No information about the donors was available. Statistical analysis found that the spectrum of a dried semen sample contained three major spectral components in addition to a fluorescent background component; a component matching tyrosine, a component containing albumin and choline, and a component matching spermine phosphate hexahydrate. This project also demonstrated qualitatively that there are no significant visual changes in the Raman spectra of dried semen acquired from multiple donors, and we showed that the spectrum of dried semen varies considerably when compared to the spectra of dried blood and saliva. We did not perform any quantitative statistical analysis to compare Raman spectra acquired for dry semen sample collected from different donors. Instead, a novel approach has been developed based on multi-dimensional spectroscopic signature to take into account both sample heterogeneity and possible variations with a donor. The combination of the three principal components can be used as a *unique* spectroscopic signature to identify the presence of semen and *possibly* distinguish it from other body fluids and substances of artificial nature found at a crime scene. The term “possibly” was used to indicate that several important steps need to be taken before the final conclusion about the differentiating power of the method could be made. These steps include (1) the development of multidimensional spectroscopic signatures for all body fluids of interest for forensic science, (2) validating the application of the developed method for mixtures of body fluids, (3) validating the application of the method for body fluid traces containing various contaminants, and (4) testing the interference of substrates. The signature’s specificity to semen is additionally reinforced by the determination that two of the three spectral components are dominated by choline and spermine, respectively, and these chemical components are unique to semen and have been used as forensic identification techniques for semen in the past. This spectroscopic signature can be fitted to all of the dried semen samples with high goodness-of-fit statistical results, and this outcome shows how the signature can be applied to any human semen sample to potentially identify it. This proof of concept experiment showed promising results, but many more samples with known demographic information should be investigated.

We envision the potential use of this method for nondestructive detection and confirmatory identification of semen at a crime scene, both in its pure form and even as part of a stain. In addition, the ability to not damage the sample while making these conclusions would be a valuable feature since it would allow the possibility of additional testing on the same sample. More experiments need to be performed involving semen stains on different materials such as clothing, paper, wood, etc., but the technique introduced in this paper shows the potential for the Raman spectroscopic signature of semen to be useful in identifying semen at crime scenes.

Continuing investigation of semen samples and other body fluids is currently taking place in our laboratory. Future work will focus on developing unique spectroscopic signatures for other body fluids to support the assumption that the different fluids can be distinguished from one another using Raman spectroscopy since they are composed of different chemical components. We are also experimenting with a more advanced statistical method which uses principal component analysis (PCA) to mathematically compare multiple spectra of different body fluids as well as spectra from different animal species of the same fluid.

Module: Forensic Science Impact in the Court Room

There is very little extant research studying the impact of forensic evidence and technologies in their impact on the outcomes of criminal investigations. Even more alarming is that we have limited documentation of the impact, in general, of the overall forensic science process to outcomes in violent criminal investigations. While we have studied the impact of new technology on forensic scientists (for example, see, Becker & Dale, 2007; Dale & Becker, 2005) the impact on the wider community remains undocumented.

The criminal justice community (prosecutors, detectives, medical examiners) benefits from a team-based structure. The forensic science community has benefited from advances in technology with concomitant increased capacity for evidence processing and analysis. Technological advances raise new questions related to the effectiveness of how agencies respond in concert to violent crime investigations. Most critically this involves how personnel in mission-critical areas function as a single enterprise, as in the coordination and execution during real-time response to a violent crime. These important questions are best addressed using the extant knowledge from social science research in teamwork that will help the forensic community adapt and apply this knowledge quickly.

The application to the criminal justice community is that it must measure and manage the impact of DNA and advanced technology to violent crime investigations. Combined DNA Index System (CODIS), National Integrated Ballistic Identification Network (NIBIN) and Automated Fingerprint Identification System (AFIS) represent technologies that have revolutionized forensic science, providing critical leads to help assure successful criminal investigations. The criminal justice community is a customer of the services of forensic science. However, the criminal justice community lacks explicit measures of the effectiveness of dollars spent on advanced technology. District attorneys and detectives were involved in the construction of the measures in this project. This is

important in that we recognize that those involved in solving high-profile crimes demand quick processing of evidence.

Increased synergy between forensic constituencies will help to stop criminals earlier in their criminal careers, in that offenders identified more quickly as a result of minor criminal activity are not able to advance to more serious crimes. Ninety-four percent of convicted offenders previously committed minor crimes (Haapanen, 1998). Felons whose most serious prior convictions were for forgery or passing bad checks had DNA matches in 12 rape cases, 8 homicides, 1 rape-homicide, an assault, a robbery, and a carjacking (Simon, 1997).

This pilot project analyzed the amount and type of evidence and forensic science technology as used by law enforcement agencies in Schenectady County, New York.

Plea bargains and verdicts were reviewed. The review assessed whether evidence could have been collected that may have improved the outcome of the case. An assessment was also made on whether the prosecutors could have further enhanced the case by seeking other evidence after arrest. Best practices were evaluated with respect to interaction with the lab throughout the criminal justice process.

This project also evaluated the use and effectiveness of forensic technology in the courtroom. Verdicts were examined and some jurors were asked, after trial, to voluntarily take a survey to assist in determining how important the presence or lack of, forensic science was in the verdict. In addition, training for eighty first responders from the Schenectady PD on proper evidence collection was done. Furthermore, funding was provided to purchase equipment from Porter-Lee to bar-code evidence in the Schenectady, New York DA's Office.

Module: ChatMinder: A Safe Internet Tool for Parents

Graduate level assistance was provided with funds from this award to perform literature research and interviews with professionals to identify the need for a new advanced language and information technologies developed at the University at Albany for assisting automated intelligence gathering and analysis to detect and prevent of online crime against children as well as other persons or organizations. Job analysis was performed to identify the necessary fundamental knowledge and skills needed by cyber crime investigators. New Chat Minder tools were then designed using the data from literature research, job analyses and interviews with criminal justice cyber crime professionals.

In recent years, there has been a proliferation in the use of Internet chat rooms – virtual online communities where users from all over the world can interact. Chat rooms attract people of all ages, from all walks of life, although their use is particularly prolific among teenagers. Because encounters are anonymous, the risk of being deceived is high. For example, many chat room sites offer areas for specific age groups, but there is little to prevent someone from creating a profile reflecting a user of that age, and entering the room. There is no guarantee that the users of children's chat areas are in fact all children.

While the vast numbers of interactions in such chat rooms are innocent, there are an increasing number of documented instances of pedophiles posing as children and 'grooming' target children for potential abuse. In response, a number of countries have adopted policies to target online crime against children. In the U.S., the bill is called the Children's Internet Protection Act (CIPA). In 2003, the Virtual Global Taskforce was created to address online crimes against children (http://www.virtualglobaltaskforce.com/). Despite these efforts, current technological capabilities to detect and prevent on-line crime are extremely limited.

The aim of this project was to leverage advanced language and information technologies developed for assisting automated intelligence gathering and analysis and apply them to detection and prevention of online crime against children as well as other persons or organizations.

Analysis of the collected data led to construction of preliminary models of social behavior in online discourse. Conversations were annotated for communicative links, dialogue acts, and topic and focus shifts, which created the basis for building computational models of conversational behavior. Some of these models, e.g., how to effectively change the topic of conversation, were subsequently implemented into an automated Virtual Chat Agent (VCA), a Chat Minder prototype. VCA has been demonstrated to perform effectively and convincingly in Internet conversation with human participants.

Module: Capital Region Cyber Crime Partnership

For this part of the project, the team continued to work with the district attorneys in the Capital Region, the New York State Division of Criminal Justice Services, and the New York State Computer Crime Unit to expedite the prosecution of computer crime and capital cases involving computer forensic evidence.

It is no longer sufficient to use a single data source or a single analytic technique while analyzing data. In addition, there is a need to analyze unstructured text data since it can provide valuable clues on criminal behavior and intentions. The three methods discussed are mutually complementary and address the needs for law enforcement in fighting crime: 1) open source data collection, 2) natural language processing, 3) identifying correlations between disparate data sources. The data collection robots automate the process of collecting online data making it efficient. Linguistic analysis can be used for behavior analysis and integration tools can be used for tagging and correlating data.

Law enforcement is saddled with a growing backlog of cases of online crime and traditional crime that rely on online evidence. Crime labs around the country have been increasing capacity to handle this growing backlog. In addition to increasing capacity, efficiency of analysis also needs to increase. The crime scene today is often not a physical location but the Internet. Being able to rapidly collect data from online sources will make it feasible for investigators to pursue more crimes. Being able to gather corroborating evidence from chats, instant messaging, and web sites can improve the rate

of conviction. A suite of tools for data analysis will be employed for a host of problems including, 1) psychological profiling of sexual predators and determining precursors to crime 2) identifying hacker motivations for committing crime.

Future studies will involve developing algorithms for data correlations and developing best practices for law enforcement to use. In addition, other linguistic characteristics beyond content analysis while examining text data will be developed and tested. Some of the techniques developed above will be useful for identifying recidivism in certain crimes.

MAIN BODY—Module: Forensic DNA Training, Leadership Assessment, and Research

A-C: DNA Training and Leadership Assessment Workshops

(Authors: Dr. Donald Orokos and Mr. John Hicks)

1. Introduction: A-C

The Northeast Regional Forensic Institute (NERFI) was created in 2003 as a more resourceful, efficient, and cost-saving approach to meet the critical training needs of new DNA personnel hired by public forensic laboratories to perform DNA analysis. Education and training programs are the core of NERFI's mission to support the forensic science and criminal justice community in staying abreast of the ever-evolving technological in the field. NERFI also strives to develop novel, state-of-the-art applications of new technologies through on-going research. Past and current interactions between the University and crime laboratories; other federal and state governmental agencies; NIJ's Office of Science and Technology; and DoD are essential to advancing education, training, and research. These cooperative relationships culminate to improve the speed, accuracy, and delivery of forensic DNA analysis.

Since 2005, NERFI has provided education and training in the form of DNA academies and professional development workshops. For example, the NERFI 16-week intensive DNA academy has successfully trained over 120 newly hired, full-time forensic scientists in the theory and practice of DNA analysis; with most trainees completing all of their training requirements, and beginning casework, in 6-8 months (as opposed to the 12 to 18 months typically required to prepare casework-ready DNA scientists through in-house mentoring programs routinely used by crime laboratories).

Studies have shown that the current population of forensic scientists in the United States is about 10,000 (Becker and Dale, 2004; Dale and Becker, 2003); however, an additional 10,000 new forensic scientists will have to enter the work-force in the next 10 years to alleviate the expanding casework backlog (Fisher, 2003; Long 2001). NERFI's education and training programs are designed to produce critical thinking, problem solving scientists - not technicians. NERFI's generic DNA training curricula are designed to embrace all forensic laboratory methods using fundamental scientific theory. The NERFI program strives to provide each trainee with the fundamental forensic biology knowledge and instrumentation necessary to perform the laboratory techniques properly and execute corrective actions when unacceptable results occur.

A major component of this project is to offer to the forensic community a set of workshops that include three separate weeks of AB 3130xl™, two 2.5-day sessions of AB RT-PCR™, and two 2-day training sessions on Leadership Assessment for Managers.

- A: AB 3130xl™ Training
  - Three separate weeks of lecture plus hands-on
    - 6-8 trainees per session
- B: AB RT-PCR™ Training
  - Two 2.5-day training sessions

        - At least 6 trainees per session
- C: Leadership Assessment Workshops
    - Two 2-day training sessions
        - 6-10 trainees per session

2. Methods: A-C

Instruction:

A key factor in providing continuing professional development to the forensic community was to employ a team of highly qualified scientists with years of practical experience, combined with a passion for educating and training the next generation of forensic DNA analysts and managers. In addition, NERFI only hires instructors and support staff willing and capable of being a team player. NERFI's ultimate goal is to serve the needs of the customer by providing cost-effective, high-quality training and education in forensic DNA and leadership assessment. Nearly all of the current NERFI instructors, support staff, and expert contractors have participated in all twelve DNA academies and countless workshops. For Modules A-C, NERFI trainers used various methods of instruction for the AB 3130xl™, AB RT-PCR™, and Leadership Assessment Workshops. For example, the AB 3130xl™ training used a combination of PowerPoint lectures (Appendices 2, 3, and 4) with hands-on instruction, whereas the AB RT-PCR™ employed mostly traditional lectures (Appendices 8 and 9). Both the AB 3130xl™ and AB RT-PCR™ instituted an interactive review session prior to the final examination. Instructors for the Leadership Assessment workshop predominantly used round-table discussions of case studies from journal articles; example-driven discussions on critical topics; and role-playing with modeling combined with traditional PowerPoint presentations (Appendix 14).

Assessment and Evaluation:

NERFI's professional development curriculum incorporates trainee performance measures and assessments that are substantially more rigorous than those typically found in traditional forensic training programs. For example, trainees in the AB 3130xl™ and AB RT-PCR™ workshops were required to earn a grade of 'B' or better on their final written examination (Appendices 5, 10 and 11) in order to receive a NERFI completion certificate. All workshop examinations were developed and reviewed by a team of NERFI instructors and staff, and each examination was graded for content by at least two instructors. NERFI has continuously improved the professional development program from feedback provided by: past trainees, NERFI and UAlbany faculty and staff, and members of the forensic science and criminal justice community.

Dissemination:

NERFI explored several avenues to inform the community about its grant-funded training opportunities. Approximately four thousand brochures were sent to the forensic community in February 2009; John Hicks and Dr. Becker spoke at sessions during the 37th Annual ASCLD Symposium on September 16, 2009; and encouraged word of mouth between members of the forensic community familiar with NERFI's quality programming.

Unfortunately, there was limited response, with many laboratory managers/supervisors/directors suggesting travel and per diem be included along with workshop attendance. Based on feedback from the forensic community and poor response generated from advertising, NERFI sent NIJ a revised budget for award 2008-DD-BX-K301 of which was approved on July 10, 2009. The revised budget included trainee travel and per diem for all workshops under this award. NERFI also updated their website in early May 2009 with information on Module #1 workshops and linked the newly updated website to NIJ's website (http://www.ojp.usdoj.gov/nij/training/welcome.htm) - many inquiries from all over the U.S. (and world-wide) were received.

NERFI also held a one-day conference on May 28, 2009, inviting laboratory directors, technical leaders, and supervisors from the northeast to discuss the upcoming AB 3130xl™, AB RT-PCR™, and managerial workshops - as well as other possible future training opportunities. Those in attendance were also asked to complete a brief survey regarding training. From the information gathered, NERFI was able to determine that these technologies (Applied Biosystems' capillary electrophoresis 3130xl and real-time PCR 7500) were priorities. Next, NERFI set about delivering programs to fulfill the needs expressed by the community. As such, the following dates were scheduled for the above mentioned workshops:

- A: AB 3130xl™ Training
  (Three separate weeks of lecture plus hands-on training for 6-8 trainees)
  - January 11-15, 2010, at Westchester County Forensic Lab
  - January 25-29, 2010, at NERFI
  - February 8-12, 2010, at Honolulu Police Dept.
- B: AB RT-PCR™ Training
  (Two 2.5-day training sessions for at least 6 trainees per session)
  - June 3-5, 2009, at Massachusetts State Police Crime Laboratory
  - January 5-7, 2010, at Westchester County Forensic Lab
- C: Leadership Assessment Workshops
  (Two 2-day training sessions for 6-10 trainees
  - January 14-15, 2010, at NERFI
  - February 18-19, 2010, at NERFI

3. Results: A-C

A: AB 3130xl™

(Three separate weeks of lecture plus hands-on training for 6-8 trainees)

December 7-11, 2009, at NERFI:

NERFI had to cancel the first AB 3130xl™ due to lack of attendees.

January 11-15, 2010, at Westchester County Forensic Lab:

Six trainees from the Westchester County Crime Laboratory (Table 1) were in attendance for this on-site workshop held Jan. 11-15, 2010. Three of the trainees had several years of experience and were acting as casework scientists. The other three were technicians, processing case samples in the lab. All of the trainees were well-versed in the use and maintenance of the 3130 instrument.

The workshop began with theoretical lectures on electrophoresis, the 3130xl instrument, and GeneMapper® ID (Appendices 2, 3, and 4). Because many of the trainees possessed years of hands-on experience with the 3130xl, and the group showed a heightened interest, NERFI staff chose to focus the workshop on the analysis and interpretation of electronic simulated case data, beginning with simple data followed by more challenging and complex.

On the final day of the workshop, NERFI staff held a review session and presented information on the recently released AB GeneMapper® ID-X software. In addition, each trainee completed a final exam (Appendix 5), an evaluation form, and received a completion certificate. Table 1 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 1: Participants and their Agency— January 11-15, 2010 at Westchester Lab

Last Name	First Name	Agency	Certificate
Vialotti	Angela	Westchester Co. Crime Lab	X
King	Nicole	Westchester Co. Crime Lab	X
Leung	Helen	Westchester Co. Crime Lab	X
Gonzalez	Lissette	Westchester Co. Crime Lab	X
Davis	Alexandra	Westchester Co. Crime Lab	X
Bradshaw	Niyrai	Westchester Co. Crime Lab	X

January 25-29, 2010, at NERFI:
Nine students from across the continental U.S. descended on the University at Albany / Northeast Regional Forensic Institute the week of January 25-29, 2010, for a 5-day workshop on Capillary Electrophoresis & Data Analysis.

There were several trainees attending this workshop (Table 2) that were currently responsible for instrumentation maintenance in their labs, as well as one individual who was slated to be the first analysts in a new DNA section. In addition, another trainee was working as a fingerprint examiner with no prior DNA bench experience, with the remaining students somewhere in between.

NERFI staff started the workshop with theoretical lectures on electrophoresis, the 3130xl instrumentation, and GeneMapper® ID v3.2.1 (Appendices 2, 3, and 4). Afterwards, each trainee did analysis and interpretation using GeneMapper® ID v3.2.1 from electronic simulated case data. All trainees were required to replace the 16-capillary array, perform both routine polymer delivery system maintenances, and both the spatial and spectral calibrations on the 3130xl. The workshop concluded with a presentation on GeneMapper® ID-X software and a review session. Each trainee completed the required exam and course evaluation.

Table 2 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 2: Participants and their Agency— January 25-29, 2010 at NERFI

Last Name	First Name	Agency	Certificate
Hou	Geoge	Los Angeles County Sheriff's Office	X
Couch	Amy	AZ DPS	X
Fejes	Ildiko	AZ DPS	X
Oliver	Dianne	TX Dept. of Public Safety	X
Bryan	Lauren	North LA Crime Lab	X
Punte	Dana	North LA Crime Lab	X
Ho	Ranee	St. Louis Metro PD	X
Smith	Juline	Trinity DNA Solutions	X
Lockhart	Brigid	Oakland Co. Sheriff's Office	X

February 8-12, 2010, at Honolulu, HI:
Eight scientists, including the DNA Technical Leader, from the Scientific Investigation Section of the Honolulu Police Department (Table 3), hosted and participated in the Capillary Electrophoresis & Data Analysis workshop on-site. The workshop consisted of several theoretical lectures (Appendices 2-4) and introductory 3130xl instrumentation hands-on work as well. Because many of the trainees already had years of hands-on experience with the 3130xl, NERFI staff spent the majority of the workshop on the analysis and interpretation of electronic simulated data sets and providing a lecture on real-time PCR.

On the final day of the workshop, the trainees participated a review session covering the instrumentation and GeneMapper® ID v3.2.1 software. Afterwards, each trainee completed the required exam (Appendix 5) and workshop evaluation form. Table 3 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 3: Participants and their Agency—February 8-12, 2010 at Honolulu, HI

Last Name	First Name	Agency	Certificate
Matsuoka	Cathy	Honolulu Police Dept.	X
Chua-Chiaco	Barrie	Honolulu Police Dept.	X
Kashimoto	Samantha	Honolulu Police Dept.	X
Tsang	Elizabeth	Honolulu Police Dept.	X
Crabbe	Sean	Honolulu Police Dept.	X
Fuller	Kim	Honolulu Police Dept.	X
Esaki	David	Honolulu Police Dept.	X
Young	Michael	Honolulu Police Dept.	X

B: Two 2.5-day training sessions on "Real-Time PCR Using Applied Biosystems' 7500 Instrument & Chemistries"

June 3-5, 2009, at Massachusetts State Police Crime Laboratory:
Thirty-six scientists from four New England Forensic Laboratories attended from June 3-5, at Massachusetts State Police Crime Lab (Table 4), the Advanced Lecture with Bruce McCord Ph.D., the Introductory Lecture with NERFI staff, or both. This 2.5-day workshop was divided into one day of Advanced Lecture (Appendix 9) followed by one and a half days of Introductory Lecture (Appendix 8).

The Advanced Lecture (Appendix 9), presented by Bruce McCord from Florida International University, consisted of five separate lectures covering a range of topics related to and affecting quantitation (via real-time PCR) of human DNA in forensic casework. Topics included: 1) DNA Quantitation by Real-Time PCR: Advanced Issues, 2) Investigation of the Effects of Sample Degradation and Inhibition in Forensic DNA Typing with Reference to qPCR, 3) A Comparison Between Plexor and Quantifiler Duo, 4) qPCR and Low Copy Template, and 5) Y STRs and qPCR. Dr. McCord lectured throughout the day and answered, as well as posed, insightful questions as he went. The attendees were engaged and focused.

The Introductory Lecture (Appendix 8) with NERFI staff began with an in-depth look at: why there is a need to quantitate, possible outcomes of not quantitating, history of quantitation, Polymerase Chain Reaction, Real-Time PCR, Fluorescent Detection (Stokes Shift & Emission Spectra), Filters, Multicomponent Analysis, AB Human Quantifiler kit, TaqMan Probe Technology, Definitions associated with Real-Time Analysis, Standard Curves and three criteria (slope, y-intercept, correlation coefficient), Troubleshooting the Standard Curve, Instrument Calibrations, and a Software Demo. The trainees were then divided into groups of two and issued laptops with SDS software and sample data. The trainee pairs were encouraged to explore the many functions of the software as the instructors worked with each pair individually.

The following day the students reviewed with NERFI staff all of the material covered in the past two days in a game of RT-PCR Jeopardy followed by a final examination (Appendices 10 and 11) and an evaluation of the workshop. The NERFI staff found this method of review to be very successful; the students were very engaged and interactive. Table 4 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 4: Participants and their Agency—June 3-5, 2009 at Massachusetts Lab

Last Name	First Name	Agency	Certificate
Ross	Joseph	City of Boston Crime Lab	X
Lynch	Julie	City of Boston Crime Lab	X
Webster	Rebecca	City of Boston Crime Lab	X
Pilla	Angela	City of Boston Crime Lab	X
Muniec	David	Maine State Police Lab	X
Sabean	Jennifer	Maine State Police Lab	X

Waterhouse	Christine	Maine State Police Lb	X
LaFountain	Marcia	Vermont Forensic Laboratory	X
Herrick	Rebekah	Vermont Forensic Laboratory	X
Drugan	Cailin	Massachusetts State Police	X
Farnam	Leanna	Massachusetts State Police	X
Harrington	Kim	Massachusetts State Police	X
Walsh	Sharon	Massachusetts State Police	X
Sgueglia	Joanne	Massachusetts State Police	X
Haddad	Sandra	Massachusetts State Police	X
Ruiz	Elisse	Massachusetts State Police	X
Tremblay	Kara	Massachusetts State Police	X
Jennings	Laurie	Massachusetts State Police	X
Ordyna	Chrissy	Massachusetts State Police	X
Simson	Crystal	Massachusetts State Police	X
Sullivan	Kristen	Massachusetts State Police	X
Collins	Sidney	Massachusetts State Police	X
Gould	Kathleen	Massachusetts State Police	X
Barber	Amy	Massachusetts State Police	X
Scott	Abbey	Massachusetts State Police	X
Brachold	Jaime	Massachusetts State Police	X
Lemire	Christine	Massachusetts State Police	X
Frederick	Alanna	Massachusetts State Police	X
Marengo	Denise	Massachusetts State Police	X
Lindauer	Kim	Massachusetts State Police	X
Wilcox	Kenton	Massachusetts State Police	X
Dindinger	Matt	Massachusetts State Police	X
O'Connor	Jessica	Massachusetts State Police	X
Schneeweis	Lynn	Massachusetts State Police	did not attend full

January 5-7, 2010, at Westchester County Forensic Lab:
Twenty-six scientists from three New York Forensic Laboratories attended the Advanced Lecture with Bruce McCord (Jan. 8, 2010), the Introductory Lecture with NERFI (Jan. 6-7, 2010), or both. The Advanced Lecture, presented by Dr. Bruce McCord covered the same topics and material presented during June 3-5, 2009. The Introductory Lecture with NERFI staff covered the same topics as described above during the June 3-5, 2009 RT-PCR. On the last day of the workshop, a review session was held using the Jeopardy-type format. Once again the NERFI staff found this method of review to be extremely successful. At the conclusion, each attendee completed the assessment examination (Appendices 10 and 11) as well as an evaluation form. Table 5 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 5: Participants and their Agency—January 5-7, 2010 at Westchester Lab

Last Name	First Name	Agency	Certificate
Ramprashad	Alanna	Westchester Co. Crime Lab	X
Davis	Alexandra	Westchester Co. Crime Lab	X
Vialotti	Angela	Westchester Co. Crime Lab	X
D'Amato	Chris	Westchester Co. Crime Lab	X
San Pietro	David	Westchester Co. Crime Lab	X
Schwartz	Elayne	Westchester Co. Crime Lab	X
Leung	Helen	Westchester Co. Crime Lab	X
O'Connor	Holly	Westchester Co. Crime Lab	X
Hoey	Jaime	Westchester Co. Crime Lab	X
Reilly	Jennifer	Westchester Co. Crime Lab	X
Chernjawski	Joselyn	Westchester Co. Crime Lab	X
MacLaren	Kevin	Westchester Co. Crime Lab	X
Gonzalez	Lissette	Westchester Co. Crime Lab	X
Stout	Lynn	Westchester Co. Crime Lab	X
Tsocanos	Maria	Westchester Co. Crime Lab	X
Eustace	Mary	Westchester Co. Crime Lab	X
King	Nicole	Westchester Co. Crime	X

		Lab	
Bradshaw	Niyrai-Daun	Westchester Co. Crime Lab	X
Walters	Samantha	Westchester Co. Crime Lab	X
Flaherty	Susan	Westchester Co. Crime Lab	X
Baumann	Bob	Suffolk Co. Crime Lab	X
Scioli	Ramona	Suffolk Co. Crime Lab	X
Wallman	Rebecca	Suffolk Co. Crime Lab	X
Gettig	Russell	New York State Police	X
Murray	Lola	New York State Police	X
Brown	Tory	New York State Police	X

C: Leadership Assessment Workshops
(Two 2-day training sessions for 6-10 trainees
January 14-15, 2010, at NERFI

The first group of forensic managers attending *Leadership Assessment: Developing the Next Generation of Leaders* in January, 2010 included three lab directors, an assistant director, three managers, and a lab supervisor, two lab managers, and a senior scientist from various forensic sections – crime scene, biology, latent prints, drug analysis, trace, and QA (Table 6). All of the nine participants had some management experience, from as little as 3 to as much as 21 years. As shown in Table 6, experience in forensic science ranged from 7 to 30 years.

Table 6: Participant Position and Experience—January 14-15, 2010 at NERFI

Participant Title	Years of Forensic experience	Years in Management	Years in Current Job
Director	30	21	2
Assistant Director	25.5	10	2
Director	22	13	10
Quality Manager	21	20	4
Lab Manager	18	3	1
Section Supervisor	17	11	11
Director	15	10	1
Lab Manager	14	10	10
Senior Scientist	7	3	3

The first day of the two-day workshop began with a look at the leadership challenges facing lab managers today, including forensic scientist turnover. This was followed with an overview of the need for responsive leadership in forensic labs as implicated in the NRC's *Strengthening Forensic Science in the United State: A Path Forward* (NAS Report, 2009) and of the importance of leadership training in such a technology-intensive field. After acknowledging the paucity of forensic lab management studies, Dr. Becker summarized the lessons learned from the few management studies that have been conducted in forensic lab situations. Dr. Pavur compared the roles of the forensic scientist and the supervising scientists, which centered on the nature of intellectual capital. Both instructors also discussed methods of assessing and managing intellectual capital by

focusing on various models of organizational behavior – creating Forensic Advisory Boards for stakeholder input, "people make the place," the **A**ttraction-**S**election-**S**ocialization-**A**ttrition (ASSA) framework, and STAR interviewing practices (**S**ituation or **T**ask, **A**ction, **R**esult).

The second day began with a tour of the New York State Police Forensic Investigation Center followed by instruction and discussion of management leadership theory, concentrating on the theories of Mary Parker Follett and the relationship between power, leadership and conflict. Dr. Becker introduced the concepts of empowerment and collaborative learning (Dr. Gary Yukl's articles) and offered specific strategies for effective leadership. Both instructors ended the workshop by leading a role play activity involving experimentation with eleven influencing behaviors.

Over the two days, the NERFI Leadership Assessment workshop engaged in co-instructional techniques of case studies from journal articles, example-driven discussions on critical topics, Power Point presentations, and role play with modeling. Table 7 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 7: Participants and their Agency—January 14-15, 2010 at NERFI

Last Name	First Name	Agency	Certificate
Adamo	Robert	Westchester Co. Crime Laboratory	X
Baral	Sanghamitra	Prince Geoge's Co. PD	X
Crenshaw	Karin	Palm Beach Co. Sheriff's Office	X
Grady	David	Worchester PD	X
Kamb	Valerie	Johnson Co. Crime Laboratory	X
Lakhkar	Bharat	Westchester Co. Crime Laboratory	X
Mayo	Nellie	Prince Geoge's Co. PD	X
Murga	Kim	Las Vegas Metro. PD	X
Eastman	Dr. Allison	Forensic Identity & Profiling	X

February 18-19, 2010

The second group of forensic managers attending *Leadership Assessment: Developing the Next Generation of Leaders* in February, 2010, included one director, a forensic science coordinator, three supervisors, a senior scientist, two technical leaders, a quality manager, and a latent print examiner with no management experience. Professionals from the forensic science fields of DNA, drug chemistry, crime scene, firearms and toolmark, QA and latent prints were represented in this session (Table 8). Nine of the ten participants had some management experience, from as little as 6 months to as much as 9 ½ years. Experience in forensic science ranged from 3 to 24 years. Table 8 below summarizes the position and experiences of the ten participants.

Table 8: Participant Position and Experience—February 18-19, 2010 at NERFI

Participant Title	Years of Forensic experience	Years in Management	Years in Current Job
Senior Scientist/QM Manager	24	3	2
Senior Scientist	20	0	0.3
Technical Leader, Supervisor	19	1.5	10
FS Coordinator	17	9.5	9.5
Supervisor	15	3	3
Director	14	4	4
Supervisor of Forensic Services	11	4	4
Technical Leader	11	2.5	2.5
Supervisor	5	0.58	0.58
Latent Print Examiner	3	0	0

The workshop content presented, and activities conducted, were nearly identical to the first workshop.

Table 9 below describes each attendee, their associated agency, and whether or not they received a completion certificate.

Table 9: Participants and their Agency—February 18-19, 2010 at NERFI

Last Name	First Name	Agency	Certificate
Werry	Brandon	Ohio State Highway Patrol	X
Saul	Doug	DuPage Co. Sheriff's Crime Lab	X
West	Sarah	MS Crime Lab	X
Duffy	Linda	Westchester Co. Crime Laboratory	X
Gombos	Jennifer	Montgomery Co. Crime Lab	X
Spannhake	William	FL Dept. of Law Enforcement	X
Anderson	Denise	Sorenson Forensics	X
Davis	Paul	Burlington Laboratories	X
Hurbanek	Nichole	New York State Police	X
Levine	Cathryn	NY State DCJS	X

4. Conclusions and Implications: A-C

A: AB 3130xl™

(Three separate weeks of lecture plus hands-on training for 6-8 trainees)

A total of twenty-three DNA Analysts attended three different AB 3130xl™ workshops. A fourth workshop was originally scheduled for the week of December 7, 2009, at NERFI—unfortunately, the workshop was cancelled due to a lack of interest. Feedback from the forensic community indicated that the lack of interest was due to the workshop's

scheduling too close to the holiday season, and that most DNA analysts had already satisfied their required 8 hours of continuing education for 2009.

Each workshop varied in content due to the level of expertise—however, the basic theory of capillary electrophoresis and GMID software, plus the data analysis and interpretation using simulated data was included for all three workshops. In addition, trainees also had hands-on training with the AB 3130xl™. All three workshops reviewed the material prior to taking the final examination. Attendees who were present for the entire workshop and who passed the assessment exam (letter grade of 'B' or better) were issued a completion certificate (Tables 1-5, 7, and 9). All attendees participating in the NERFI workshops passed with a 'B' or better. NERFI goes to great lengths to present the information in various formats: presentation, handouts, animation, and discussion - by more than one instructor, with interactive review sessions, to ensure the students in the group have the best possible chance of coming away with the intended knowledge. In addition, one of the necessary qualifications of a NERFI instructor is that they are knowledgeable, approachable and have years of forensic science training and education experience. These traits have proven to be invaluable when trying to engage students and encouraging them to ask questions when they are unsure. NERFI instructors are also well versed in recognizing students that may be apprehensive and approaching them after lectures, starting conversations, and enticing questions from these students that may not otherwise ask.

A program evaluation for all three AB 3130xl™ workshops provided feedback (*Poor* 1 to 5 *Excellent*) on five areas: Program Overall, Speakers, Quality of Audio-Visual, Quality of Handouts and Facilities (Appendix 6). Attendees rated the program very high, especially with regard to the Course Overall and Instructors, with comments indicating the topics covered were very *informative* and *thorough*. The instructors themselves were described as *clear and well-spoken* and *sufficiently answered all of the questions*. Participants enjoyed the small group interaction and commented that they would like to see similar offerings in the future.

B: Two 2.5-day training sessions on "Real-Time PCR Using Applied Biosystems' 7500 Instrument & Chemistries"

NERFI originally proposed at least six trainees per training session. However, a total of sixty-two DNA Analysts attended the two 2.5-day training sessions on Real-Time PCR Using Applied Biosystems' 7500 Instrument & Chemistries. The interest level for the two RT-PCR workshops was exceptional for two reasons. First, having two separate host-sites in the northeast allowed more trainees to attend per session. Second, many of the trainees appreciated the advanced lectures on real-time PCR presented by Dr. Bruce McCord from Florida International University. However, high marks from trainee feedback, for both workshops, were given to NERFI staff for their Introductory RT-PCR presentations (Appendix 12).

C: Leadership Assessment Workshops
(Two 2-day training sessions for 6-10 trainees
January 14-15, 2010, at NERFI

An evaluation requested scored feedback (*Poor* 1 to 5 *Excellent*) on five areas: Program Overall, Speakers, Quality of Audio-Visual, Quality of Handouts and Facilities (Appendix 16). Participants scored all aspects of the program favorably, with the highest marks given to the speakers themselves as *professional, knowledgeable, highly specialized, and effective.* Participants enjoyed the small group interaction and commented that they were leaving with management tools they could use in their current positions.

February 18-19, 2010, at NERFI
A program evaluation requested scored feedback (*Poor* 1 to 5 *Excellent*) on five areas: Program Overall, Speakers, Quality of Audio-Visual, Quality of Handouts and Facilities. Participants scored all aspects of the program favorably, with the highest marks given to the Program Overall and a comment indicating the topics covered were useful (Appendix 16). The instructors themselves were described as *very knowledgeable* and praised for prompting *very thought-provoking discussions*. Participants enjoyed sharing experiences with other forensic scientists. Small group interaction and commented that they would like to see similar offerings in the future.

The most interesting /useful topics included influencing tactics, the relationship between morale and retention, and the psychology of being a leader.

The instructors for the Leadership Assessment workshop revised the content and schedule for the second session based on the comments from the first session participants, the experience level of the second group, and questions from the participants on the first day. The January group indicated that the influencing tactics were quite useful. Therefore, the instructors moved that activity to the first day to provide more time to practice influencing behaviors. In addition, there was discussion on morale, engagement, and loyalty, based on questions from the January session. The February group (Table 8) had fewer long-tenured senior managers or directors than the January group (Table 6), so the second day included more team relationship and customer aspects, and less material on external liaisons for upper management. As the instructors make adjustments to meet the needs of the participants, some of the PowerPoint presentations are modified and do not follow the pre-printed handouts.

Lessons learned, future thoughts for improved workshops:
Over the years NERFI has relied heavily on feedback from the forensic community for all of its education and training programs. For example, after each training session NERFI instructors and staff review evaluations and revise the curricula accordingly. NERFI has future plans for all potential attendees to complete a brief questionnaire before attending a workshop. Some of the topics covered in the newly developed pre-workshop questionnaire will include: level of experience, current job title, and training and education needs. Armed with this information, NERFI plans to offer two workshops for each topic in the future, one introductory and one advanced. This approached was used for the real-time PCR workshop and the NERFI instructors found this method of instruction to be very successful. It was noted that this information would have been very useful for the 3130xl™ training workshop.

After reviewing the Leadership Assessment evaluation forms it was determined that the attendees would like to extend the workshop to include one more day. One experienced manager explained it quite well when she commented that she felt like she had the background after two days and was ready to use it if she had just had one more day. Part of the issue is that the sessions are mixed with new managers, people in line of succession to be managers, and very experienced managers. Because of this, the instructors covered all -- theory, research -- and yet still offered training in specific skill sets. New managers acquired a huge amount of what they needed but experienced managers, while appreciating the theory and research, are ready for the more seasoned business management approach which there was just not time for in these sessions.

Implications for the field of criminal justice and the study of forensic science
Because of the recent economic downturn and budgetary constraints, training budgets are the first items to be eliminated in most forensic science labs. However, the FBI Quality Assurance Standards require that all DNA Analysts receive at least 8 hours of continuous education each fiscal year. In Modules A-C, over 100 trainees from across the nation received their annual mandatory professional development by successfully completing the NERFI 3130XL, RT-PCR or Leadership Assessment for Managers workshops. More importantly, all trainees received this specialized and required training free-of-charge. All of the trainees that took the written examinations scored at least an "80" or better (no failures) with an overall average of 90.38.

Dissemination
NERFI is in the process of producing a peer-reviewed report on the various methods of training currently being done throughout the forensic science community. Information obtained from the 3130XL, RT-PCR or Leadership Assessment for Managers workshops will be included in this paper, and presented to scientists, policy makers, and practitioners at a criminal justice and scientific meetings.

D—Raman Method Validation--The use of Raman Technology to Detect Human Biological Fluids
(Author: Dr. Igor Lednev)

1. Introduction: D
In recent years, forensic analysis has become one of the most growing areas of bioanalytical chemistry (Brettell etal, 2007). The ability to identify traces of body fluids discovered at crime scenes is a very important aspect of forensic investigations (Li, 2008; Shaler, 2002; Jones, 2005). With DNA analysis being one of the most popular and informative forensic techniques, it is imperative that any potential body fluid sample is not destroyed during the initial identification process. Fluids such as blood, semen, saliva, and vaginal fluid can be very useful in identifying a victim or suspect (Best, 2007), and they can also help answer questions regarding the events of a crime. An analytical technique that could identify a particular body fluid rapidly, simply, and non-destructively at the scene of a crime would be a valuable tool for forensic investigators.

Virkler and Lednev have recently reported that Raman spectroscopy can be potentially used to distinguish different body fluids (Virkler and Lednev, 2008) as well as provide non-destructive, confirmatory identification of body fluids at the scene of a crime (Virkler and Lednev, 2009). However, this analysis was carried out on only one sample of each body fluid and did not take into account any variations that might occur between different donors of the same fluid. Since each donor's sample is heterogeneous within itself due to many different chemical components, we would also like to investigate the effect these chemical components have on the Raman spectral components of a body fluid. This paper investigates the role of heterogeneity within a sample as well as among multiple donors for human semen.

In addition to blood, semen is one of the most prevalent body fluids found during criminal investigations, especially in cases involving sexual assault (Shaler, 2002). There are currently several tests, both presumptive and confirmatory, that can be used to identify an unknown fluid found at a crime scene to be semen (Virkler and Lednev, 2009). Some of the popular presumptive tests include searching for stains using an alternate light source (ALS) and looking for the presence of seminal acid phosphatase (SAP) using chemical tests (Li, 2008; Greenfield and Sloan, 2003). Some commercial ALS instruments have been developed such as the Wood's Lamp (Santucci et al, 1999), Bluemaxx™ BM500 (Nelson and Santucci, 2002), Polilight® (Vandenberg and vanOorshot, 2006), and the Lumatec Superlight 400 (Fielder et al, 2008), but these are not exclusive to semen identification and can only be used as a screening technique. The tests for SAP (Li, 2008; Watson, 2004) are much more reliable, but they are destructive to the sample and there is still some potential for false positive results. The most popular methods for confirmatory identification of semen include the microscopic visualization of sperm cells using specific stains and immunological tests for prostate specific antigen (PSA) (Li, 2008; Greenfield and Sloan, 2003). The staining method will of course not be helpful if the donor is azoospermic, so this technique has limited applications including having to perform the test in a laboratory. Several commercial PSA test kits which can be used at a crime scene have been developed including Biosign® PSA (Maher et al, 2002), OneStep ABAcard® (Hochmeister et al, 1999), Chembio, Medpro, Onco-screen (Healy et al, 2007), PSA-check-1, Seratec® PSA Semiquant (Hochmeister et al, 1999), and SMITEST (Sato, etal 2002; Yokota et al, 2001). Like with the presumptive tests, these PSA test kits do show false positive results and are destructive to the sample.

The destructive nature of both the presumptive and confirmatory tests is the largest concern that needs to be addressed. Sometimes a very small amount of semen evidence can solve a case if examined properly, so it is crucial that the available evidence is processed efficiently and non-destructively so that further analysis, including DNA typing, can be performed (Budowle and vanDaal, 2009) . Another issue is the potential of false positive results (Denison et al, 2004). The current easy-to-use test kits do not absolutely confirm the presence of semen either in pure form or as part of a stain. The forensic community is in great need of a reproducible, non-destructive, and portable method that can exclusively identify the presence of semen at a crime scene and distinguish it from other body fluids.

Raman spectroscopy is a forensic technique that has increased in popularity over the last several years (Macleod and Matousek, 2008; Bartick, 2002), and it can be paired with infrared (IR) absorption spectroscopy to gain information about the structure and properties of materials based on their vibrational transitions (Nafie, 2001). Some applications being used today involve the identification of fibers (Thomas et al, 2005), drugs (Hodges and Akhavan, 1990), and lipsticks (Rodger and Broughton, 1998), as well as ink (Mazzella and Buzzini, 2005), paint (Suzuki and Carrabba, 2001), and condom lubricant (Coyle and Anwar, 2008) analysis. The theory behind Raman spectroscopy involves the inelastic scattering of a low-intensity, monochromatic, and nondestructive laser light by a solid, liquid or gas sample. There is little to no sample preparation, and no reagents are needed for analysis. Most importantly, the required amount of sample needed for Raman analysis can be as low as several picograms or femtoliters, and the sample will not be destroyed so that further analysis can still be performed. A typical Raman spectrum reveals a specific vibrational signature of the sample being measured based on the energy of the scattered light, and this feature is very useful in identifying an unknown substance. Raman spectroscopy is also very appropriate for the analysis of disordered and heterogeneous samples (Colomban and Gouadec, 2009) which are common properties of body fluids. Finally, Raman spectroscopy shows very little interference from water (Grasselli, 1981) which makes it a great technique for analyzing body fluids and their traces. Portable Raman spectrometers are available now (Eckenrode et al, 2001; Yan and Vo-Dinh, 2007), and these designs along with advanced software could be applied to the identification of semen at a crime scene.

Surprisingly, there have been no publications of any experiments involving the identification of semen using Raman spectroscopy. The objective of this study is to determine the heterogeneity of a dried semen sample from one donor as well as analyze the qualitative variation among samples from different donors using NIR Raman spectroscopy. Determining the specific principal components that contribute to the overall spectrum of semen is an important task since it is unlikely that a single library spectrum of semen will match an unknown semen sample. If careful statistical analysis is not performed, there could possibly be a false positive result for another body fluid that is similar in composition. The goal is to use automatic mapping to develop a spectroscopic signature specific to dried semen which will be generated by combining multiple spectra based on the different components that make up a basis sample due to heterogeneity. Our hypothesis was that advanced statistical analysis could be used to separate the basis semen spectrum into individual components, and that these components could potentially be characterized based on the known chemical composition of semen. The characterized principal components could then be used in a “multi-dimensional analysis” of dried semen as opposed to a “single-dimensional analysis” which only involves the comparison of a single average library spectrum. The principal components found from the basis semen sample will be fitted to the average dried semen spectrum obtained from multiple donors to illustrate the capability of a unique spectroscopic signature to be applied to all semen samples. herein this study, the NIR Raman dried semen component spectra found by analyzing a single semen sample as well as the spectra obtained from analyzing dried semen samples from many donors. Preliminary assignments of major Raman peaks and possible identities of the semen components were made based on literature data.

2. Methods: D

2.1 Samples

A set of 50 semen samples were obtained from anonymous donors at an in-vitro fertilization clinic. A small 10 µL drop of each sample was placed on a circular glass slide designed for use with an automatic mapping stage and allowed to dry completely. The samples were analyzed using automatic mapping that scanned a sample area of 75x75 µm and measured Raman spectra from 10 random points within the area with 6 ten-second accumulations at each point.

A single basis semen sample was prepared in the same way as the others, and automatic mapping was again used scanning 36 random points with 6 ten-second accumulations at each point. The spectra obtained from this sample were used to determine the number and identities of the principal components of semen and to develop the spectroscopic signature.

2.2 Raman Microscope

A Renishaw inVia confocal Raman spectrometer equipped with a research-grade Leica microscope, 20x long-range objective (numerical aperture of 0.35), and WiRE 2.0 software were used. For the automatic mapping, the lower plate of a Nanonics AFM MultiView 1000 system was set up under the microscope, and measurements were taken using Quartz II and QuartzSpec software. A 785-nm laser light was utilized for excitation. The laser power on the dried samples was about 115 mW.

2.2 Data Treatment

All of the spectra obtained from the automatic mapping of the dried semen samples were first treated using GRAMS/AI 7.01 software to remove any cosmic ray interference. The spectra were then imported into MATLAB 7.4.0 for statistical analysis and normalized to adjust for the varying amount of background interference in each spectrum. The number of principal components in the basis sample was determined using significant factor analysis (SFA), and the individual component spectra were extracted using the alternate least squares (ALS) function. The components found in the original basis sample were used to create a spectroscopic signature, and this signature was fitted to each average spectrum found from the remaining semen samples. The Curve Fitting Toolbox in MATLAB was used to perform residual analysis on the difference between the fitted and experimental spectra, and "goodness-of-fit" statistics were calculated based on how well the signature matched the experimental spectrum.

3. Results: D

3.1 Main approach

The main goals of this study were to determine the level of spectral heterogeneity of human semen based on principal components and to find out how much variation there is in the spectra from different donors. If there is very little change in the spectrum from one donor to another, then the technique of Raman spectroscopy can be considered to be reproducible in identifying a sample to be semen based on the application of a calculated spectroscopic signature. This signature, which could be fitted to a semen sample collected from any donor, could be produced based on several spectral components found

in semen that are present due to the heterogeneous distribution of the many chemical species in semen. A unique signature can ultimately be developed for other body fluids as well so that an unknown body fluid discovered at a crime scene could potentially be identified in a confirmatory manner.

3.2 Single Sample Heterogeneity

A single basis semen sample was used to develop the spectroscopic signature that would be applied to all samples. The basis sample spectra were imported into MATLAB, and SFA analysis was performed to determine the number of principal components that were present. The results of this analysis (data not shown) indicated 6 principal components. The ALS function was applied to extract the spectra of each of these components, and further examination of the results revealed that there were actually only 3 unique components that were spectral representations of the chemical species in the semen sample. The remaining components consisted of background fluorescence, a virtual duplicate of one of the 3 real components, and a nonsense component that appeared to just be noise. Figure 1 shows the average spectrum of the basis sample along with the 3 principal components.

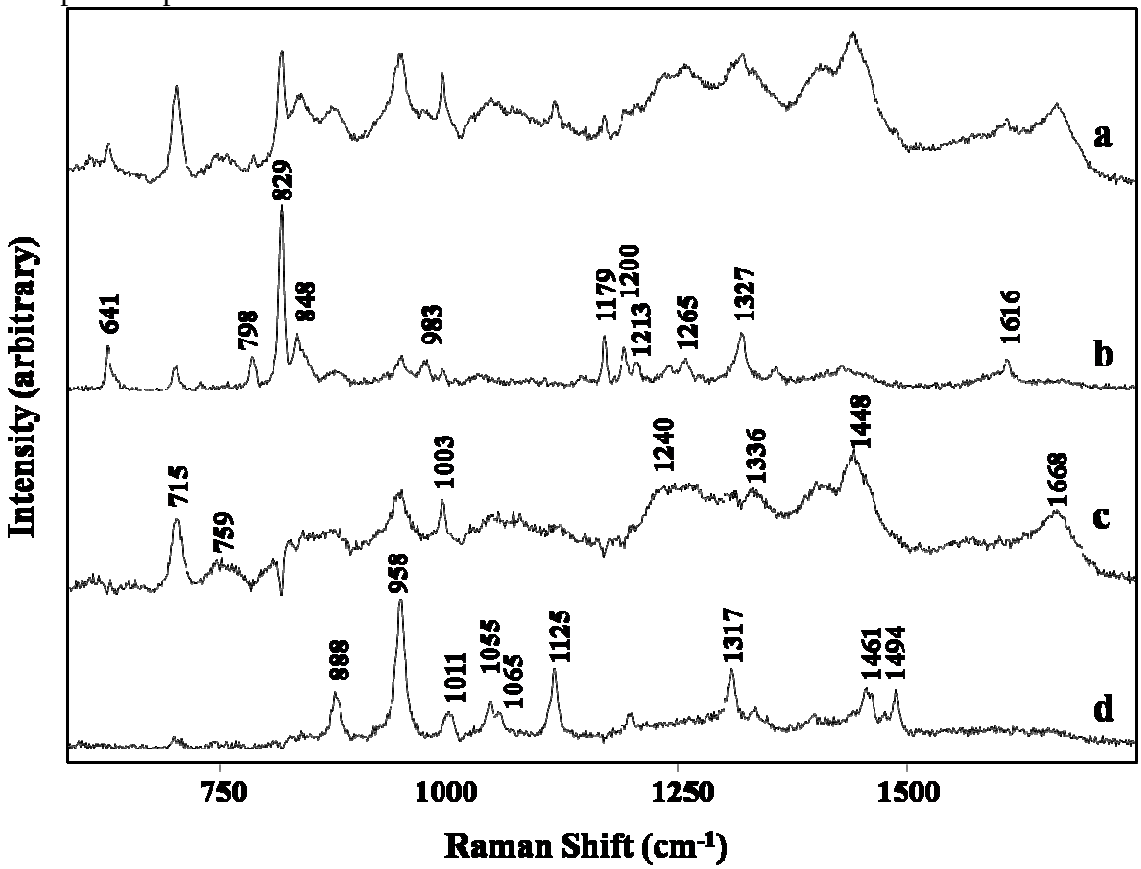


Fig. 1. The average Raman spectrum of the basis semen sample (a), and the Raman spectra of semen spectral components 1 (b), 2 (c), and 3 (d) with major peaks labeled.

The wave number range of 670-1750 cm^{-1} is shown in the figure and will be the used to create the spectroscopic signature since this is the region that contains most of the important characteristic peaks. The major Raman peaks that define each component are labeled and are listed in Table 1.

Table 1 – Raman peak assignments of dried semen

Raman shift (cm^{-1})	Spectral component	Vibrational mode	Chemical component
641	1	Ring deformation [58]	Tyrosine
715	2	CN stretching [46]	Choline
759	2	Ring vibrations (Trp) [59]	Albumin
798	1	CH_2 deformations in ring [58]	Tyrosine
829	1	Ring breathing [38, 58]	Tyrosine
848	1	Ring bending [58, 60]	Tyrosine
888	3	Phosphate mode [48]	SPH
958	3	PO_4^{3-} sym. stretching [49, 61]	SPH
983	1	CH_2 wagging [58]	Tyrosine
1003	2	Aromatic ring breathing (Phe) [43]	Albumin
1011	3	CC stretching [48, 58]	SPH
1055	3	CN sym. stretching [48]	SPH
1065	3	PO_4^{3-} asym. stretching [49, 61]	SPH
1125	3	CN asym. stretching [48]	SPH
1179	1	CH_2/NH_3 rocking [58]	Tyrosine
1200	1	CC stretching [60]	Tyrosine
1213	1	CH_2 twist and rock [58]	Tyrosine
1240	2	Amid III [44]	Albumin
1265	1	Sym. ring deformation [62]	Tyrosine
1317	3	CH vibration [58]	SPH
1327	1	Ring stretching [58]	Tyrosine
1336	2	CH bending (Trp) [39, 44]	Albumin
1448	2	CH_2, CH_3 bend (Trp) [32]	Albumin
1461	3	CH_2 bending [32]	SPH
1494	3	NH_3 sym. bending [49]	SPH
1616	1	CC stretching [60]	Tyrosine
1668	2	Amid I [44]	Albumin

Possible assignments and vibrational modes for each peak are also listed, and these assignments were made based on the known composition of semen and literature data.

The spectral components of semen are complex, and some have contribution from multiple chemical species. According to several literature sources (Altman, 1961; Owen and Katz, 2005; Mann, 1975), some of the chemical components of semen that are present in the highest concentrations are fructose, choline, spermine, citric acid, acid phosphatase, and albumin. Other chemical components of lower abundance are glucose, inositol, lactic acid, and urea. Despite this list of dominant chemical species, component 1 shown in Figure 1 appears to be a match to the amino acid tyrosine. The peaks at 641, 798, 829, 848, 983, 1179, 1200, 1213, 1265, 1327, and 1616 cm^{-1} are almost exactly the same as peaks depicted in literature sources on the Raman spectrum of tyrosine (De

Gelder et al, 2007; Johnson et al, 1986). There are also residual peaks from other chemical species, but these are more dominant in the other two components. It is surprising that a single amino acid would be found to be one of the major principal spectral components of semen instead of a more complex chemical compound, but there are free amino acids known to be in semen (Altman, 1961), and it is possible that tyrosine is abundant in albumin and acid phosphatase which are both large contributors to the composition of semen. Phosphorylated tyrosine residues have been reported to have a protecting effect on the membranes of sperm cells and help stabilize lipids (Sancho et al, 2006), so an abundance of free tyrosine in semen is practical. In addition, it has been found that tyrosine phosphorylation occurs when sperm undergo capacitation which is necessary before fertilization (Liu et al, 2006).

It is obvious at first glance that component 2 is dominated by a protein due to the presence of amid I and amid III (Carter and Edwards, 2001) peaks at 1668 cm^{-1} and 1240 cm^{-1}, respectively. It has been reported that the protein albumin makes up about one third of the protein content of semen (Owen and Katz, 2005), therefore this is a logical assignment as a contributor to component 2. Comparison of literature data on the Raman spectrum of albumin supports this conclusion (Liang et al, 2008; Ivanov et al, 1994), with matching peaks occurring around 759, 1003, 1336, and 1448 cm^{-1} in addition to the amid I and III peaks already mentioned. It is also possible that the enzyme acid phosphatase is contributing to this component due to its protein qualities and large abundance in semen, but there is not a lot of literature data available to compare with, so albumin will be considered the major contributor of component 2 for now. Finally, choline also appears to be present in component 2. The large peak at 715 cm^{-1} is very likely due to the C-N symmetric stretch found in choline which has been previously reported (Edsall, 1943; Koyama et al, 1977; Spiker and Levin, 1975). It is also likely that the CH_2 scissoring in choline is contributing to the peak at 1448 cm^{-1} (Koyama et al, 1977). This peak is very large in component 2 so it is probable that more than one chemical species is contributing to it. As with component 1, there are other peaks present that are much stronger in the other two components, so they are not considered to be dominating in component 2. Component 3 appears to oppose component 1 when comparing spectra from one donor collected from different spots. When the peaks present in component 1 are strong, the peaks for component 3 are weak and vice versa (Figure 2).

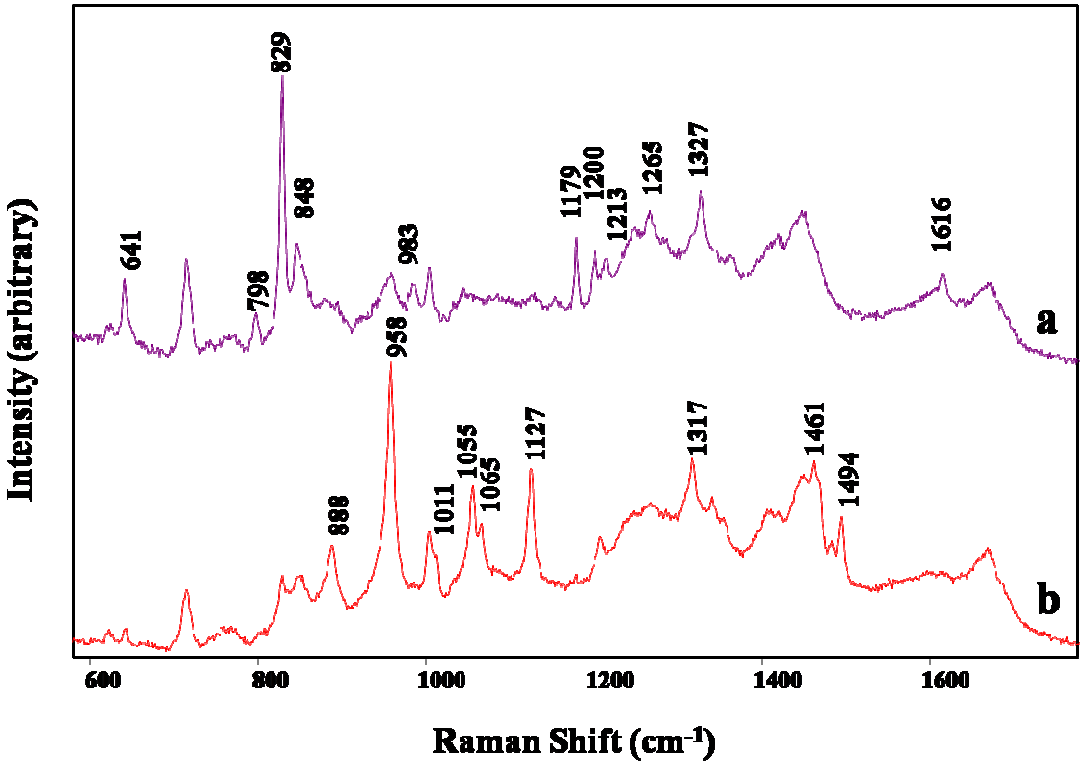


Fig. 2. A dry trace of semen is strongly heterogeneous. Raman spectra acquired from different spots in the same dried semen sample dominated by component 1 (a) and component 3 (b).

The chemical contributor to component 3 was fairly simple like with the case of component 1. The spectrum for component 3 is a match to spermine phosphate hexahydrate (SPH) that has previously been reported (Bertoluzza et al, 1983; Eapen and Joe, 1997). The peaks found in component 3 at 888, 958, 1011, 1055, 1065, 1125, 1317, 1461, and 1494 cm^{-1} all appear in the known spectrum of SPH and are listed in Table 1 with vibrational assignments. As previously mentioned, spermine is present in large concentrations in human semen, and it has even been the basis of forensic semen identification in the past (Gonmori et al, 1994; Tsutsumi, 1987; Suzuki et al, 1980; Sato et al, 1996). The basic nature of spermine causes it to interact with the phosphoric acid groups of nucleic acids and form strong bonds (Bertoluzza et al, 1983; Eapen and Joe, 1997). This binding leads to a precipitation of SPH in semen (Iitaka and Huse, 1965). These crystals were first observed by Leeuwenhoek in 1678 (van Leeuwenhoek, 1678), and they have since been observed in semen as the fluid begins to dry (Rosenheim, 1924). These properties regarding the presence of SPH in semen make it an understandable assignment of component 3. As with the other two principal components of semen, component 3 contains residual peaks from chemical species which are more dominant in components 1 and 2.

3.4 Multiple Donors

The second objective of this study was to determine the amount of spectral variation from one semen donor to another. The first step was qualitative in nature and involved the

visual comparison of the average spectrum from all of the different donors. All of the spectra appeared to be very similar and contained all of the same major peaks. There were changes in intensity of some peaks for different donors, but this is expected since the relative contribution of the chemical species in semen will likely change with each donor and can even change within the same donor (Mann, 1975). Figure 3 shows the average spectra of five semen samples (black lines) from different donors as an example of their similarities. To demonstrate the spectral differences between body fluids, the spectra of blood and saliva that we have previously reported (Virkler and Lednev, 2008) are also included in this figure. Unlike the slight intensity changes within the semen samples, blood and saliva have major peak differences which make them distinguishable. The other features of Figure 3 will be discussed in more detail in the following paragraphs.

After achieving visual confirmation that there is consistency among the spectra of semen from different donors, a more quantitative approach was developed. A spectroscopic signature was created that consisted of the 3 principal components found in the basis semen sample along with a horizontal line and a line with a slope equal to that of the fluorescence background. These five basis spectra were linearly fitted to the average basis semen spectrum, and the two spectra overlapped very well. The spectroscopic signature was also applied to the spectra from each of the remaining 49 semen samples to determine if it could universally be fitted to a sample from any donor. Figure 3 shows the fitting to only five representative semen samples, but all of the samples had very similar fits. The bottom of Figure 3 contains the results of fitting the spectroscopic signature to the spectra of blood and saliva, and it is visually obvious that they are very poor matches.

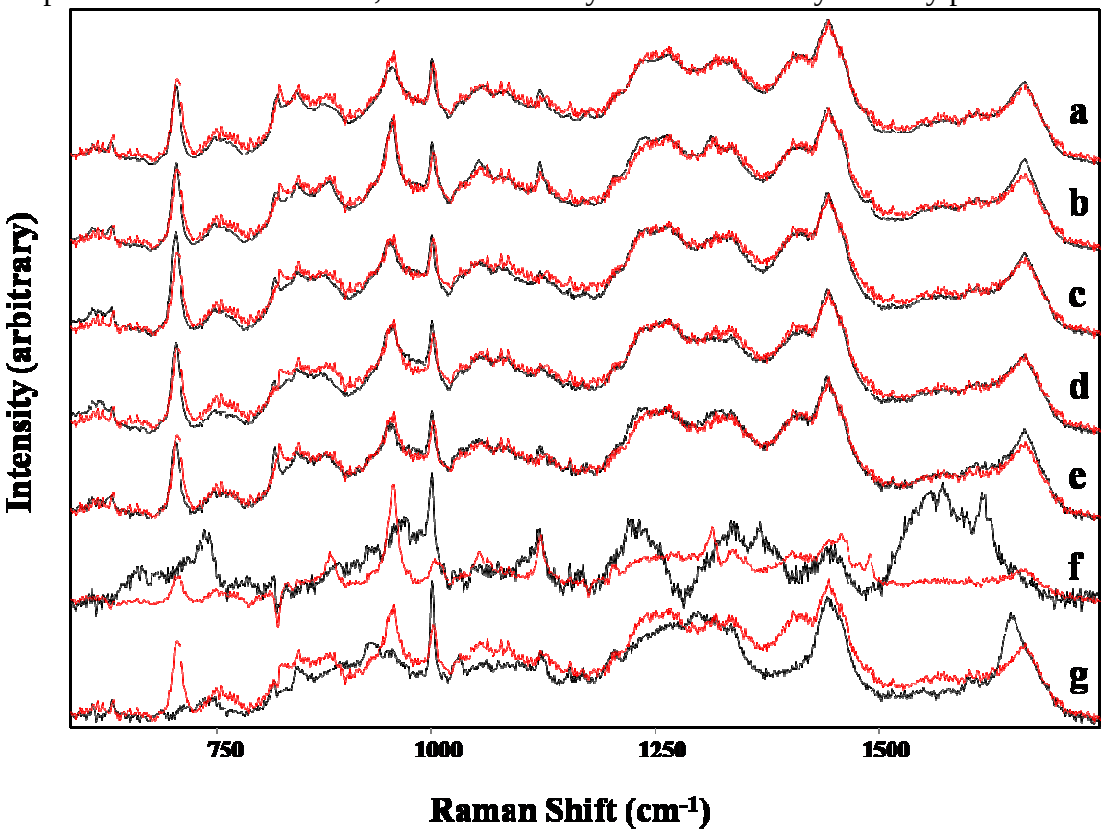


Fig. 3. The average Raman spectra of five semen samples (black) with the fitted spectroscopic signature (a-e), and the Raman spectra of blood (f) and saliva (g) with the fitted spectroscopic signature.

These results qualitatively show the specificity of this signature to semen and its potential ability to be used as an identification technique for forensic purposes.

A quantitative statistical analysis was performed to determine how well the spectroscopic signature fit the experimental spectra. Using the Curve Fitting Toolbox in MATLAB, the intensity values for the basis experimental spectrum and fitted spectrum were plotted on an axis as the x- and y- coordinates, respectively. All of the spectra were normalized to a maximum value of 1, so that is the highest value for both the x- and y- axis. Two identical spectra would yield a scatter plot matching a line with the equation of y=x, so this line was used for comparison and was fit to the plotted data points of the basis sample evaluation to determine how close of a match the experimental spectrum and signature were. The statistical result was three quantitative goodness-of-fit values which statistically confirmed the qualitative match of the experimental and fitted spectra (Table 2).

Table 2 - Goodness-of-fit statistical results for semen signature fitting

Sample	SSE	R-square	RMSE
Semen basis	0.0556	0.998	0.00630
1	0.299	0.990	0.0146
2	0.379	0.982	0.0164
3	0.450	0.979	0.0179
4	0.311	0.981	0.0149
5	0.319	0.979	0.0151
6	0.200	0.992	0.0119
7	0.324	0.987	0.0152
8	0.164	0.996	0.0108
9	0.161	0.992	0.0107
10	0.189	0.989	0.0116
11	0.107	0.996	0.00875
12	0.111	0.995	0.00888
13	0.122	0.995	0.00932
14	0.107	0.996	0.00875
15	0.281	0.987	0.0142
16	0.191	0.991	0.0117
17	0.460	0.984	0.0181
18	0.0645	0.999	0.00678
19	0.251	0.990	0.0134
20	0.0736	0.998	0.00725
21	0.118	0.996	0.00919
22	0.220	0.991	0.0125
23	0.0853	0.998	0.00780
24	0.174	0.993	0.0112
25	0.177	0.993	0.0112
Blood	1.38	0.967	0.0313
Saliva	2.25	0.822	0.0401

These values are the sum of squares due to error (SSE), R-square, and root mean squared error (RMSE). The SSE value measures the total deviation of the data points from the y=x line, and a value closer to 0 means there are fewer random errors (Wakefield, 2008). The R-square value indicates how well the y=x best-fit line explains variation in the data, and a value closer to 1 indicates that a higher proportion of the variance is accounted for by the line (Wakefield, 2008). A value closer to 1 also means that the fitted signature and experimental spectrum are a better match. Finally, the RMSE value estimates the standard deviation of the random data components. Again, a value closer to 0 indicates that the y=x line is a better fit (Wakefield, 2008), and that the signature better fits the experimental spectrum.

This same fitting procedure was performed for the remaining 49 samples. The graphical result of a typical fit is shown for sample 22 in Figure 4A as an example. The top half is a graph showing the fit of the line y=x to the comparison of the signature and sample, and the bottom half is a plot of the residuals which are found by subtracting the best fit line from the scatter plot. A residual plot with random points around zero that do not form a pattern indicates a good fit (Wakefield, 2008), as is the case here. The results for SSE, R-square, and RMSE for 25 of the 50 samples are shown in Table 2. As expected, the basis sample results indicated the best match since that sample was the template for the spectroscopic signature determination, but the statistical values for all of the samples are very close fall within a certain range that suggests a good fit. To put in perspective how well the signature fits the semen samples, it was also applied to the spectra of human blood and saliva which we have already reported (Virkler and Lednev, 2008). The goodness-of-fit statistics for those fits are listed at the bottom of Table 2, and it is easy to see how poorly the semen signature matches the spectra of the other two body fluids when all three statistics are taken into account. The R-square value for the semen signature fit to blood is not too much different than the average fit to semen, but the SSE and RMSE values definitely indicate a much worse fit. The visual results of the signature fit to the saliva sample are also shown in Figure 4B.

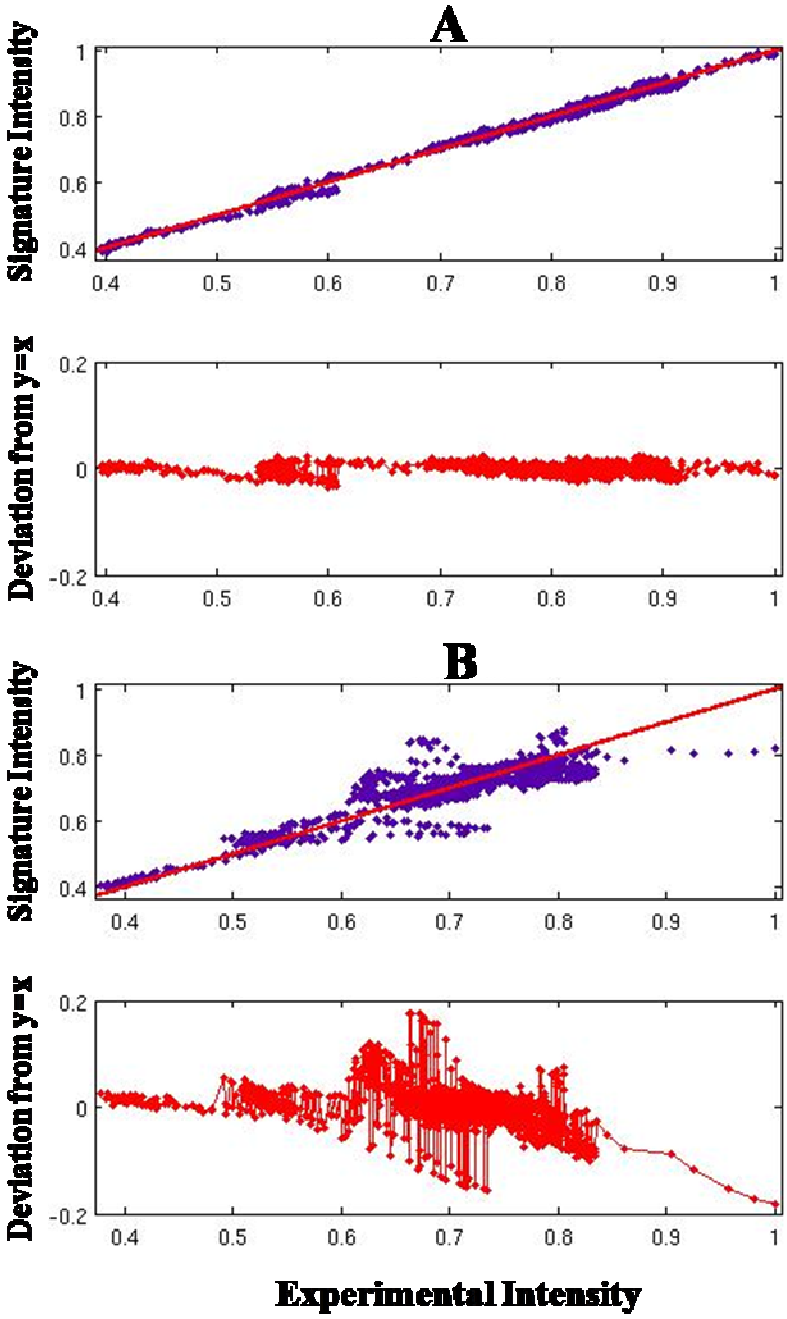


Fig. 4. Quantitative evaluation of the fitting quality. Comparison of the line y=x with the semen signature fit for a semen sample (4A, top) along with the residual plot (4A, bottom). The semen signature fit for a saliva sample (4B, top) along with the residual plot (4B, bottom).

There is a large amount of disagreement between the best fit line and scatter plot in the top graph, and there is an obvious pattern of digression away from zero in the residual plot.

As revealed in Figure 3, Figure 4, and Table 2, the semen signature closely matches the experimental semen spectra and clearly does not fit the blood or saliva spectra. This result shows that a spectroscopic signature created from one basis semen sample can be fitted to multiple other semen samples from different donors, and this technique can potentially be used to identify an unknown sample to be semen without yielding a false positive match with other body fluids. However, at this stage of the project, the identification could be made only for individual body fluid traces (not mixtures of body fluids) without any contamination or substrate interference as indicated above (steps for further research).

4. Conclusions and Implications: D

A spectroscopic signature for human semen was developed based on the heterogeneous chemical composition of semen using NIR Raman Spectroscopy. Statistical analysis found that the spectrum of a dried semen sample contained three major spectral components in addition to a fluorescent background component; a component matching tyrosine, a component containing albumin and choline, and a component matching spermine phosphate hexahydrate. The results demonstrated qualitatively that there are no significant visual changes in the Raman spectra of dried semen acquired from multiple donors, and that the spectrum of dried semen varies considerably when compared to the spectra of dried blood and saliva. The combination of the three principal components can be used as a unique spectroscopic signature to identify the presence of semen and possibly distinguish it from other body fluids and substances of artificial nature found at a crime scene. The signature's specificity to semen is additionally reinforced by the determination that two of the three spectral components are dominated by choline and spermine, respectively, and these chemical components are unique to semen and have been used as forensic identification techniques for semen in the past. This spectroscopic signature can be fitted to all of the dried semen samples with high goodness-of-fit statistical results, and this outcome shows how the signature can be applied to any human semen sample to potentially identify it. This proof of concept experiment showed promising results, but many more samples with known demographic information should be investigated.

We envision the use of this method for nondestructive detection and confirmatory identification of semen at a crime scene, both in its pure form and even as part of a stain. A forensic investigator would be able to determine the true identity of a suspected semen sample and whether it was pure and not contaminated. The ability to make these identifications and conclusions, especially at the scene of a crime, would be major progress in the area of forensic semen analysis. In addition, the ability to not damage the sample while making these conclusions would be a valuable feature since it would allow the possibility of additional testing on the same sample. More experiments need to be performed involving semen stains on different materials such as clothing, paper, wood, etc., but the technique introduced in this paper shows the potential for the Raman spectroscopic signature of semen to be useful in identifying semen at crime scenes.

Continuing investigation of semen samples and other body fluids is currently taking place in our laboratory. Future work will focus on developing unique spectroscopic signatures

for other body fluids to support the assumption that the different fluids can be distinguished from one another using Raman spectroscopy since they are composed of different chemical components. In addition, a more advanced statistical method which uses principal component analysis (PCA) to mathematically compare multiple spectra of different body fluids as well as spectra from different animal species of the same fluid will be tested in future work.

Implications for the field of criminal justice and the study of forensic science:
The needs and current status of various methods for detection and identification of body fluids for forensic purposes are briefly outlined above. In much greater details, we have disseminated this information in our recent review article "Analysis of Body Fluids for Forensic Purposes: From Laboratory Testing to Non-Destructive Rapid Confirmatory Identification at a Crime Scene" (*Forensic Sci. Int.* 2009, 188, 1-17). Great progress has been made in developing multi-dimensional Raman spectroscopic signatures for dry traces of various body fluids. In addition to the future research steps outlined above, practical application of our new method will require the understanding (i) how aging effect the spectral response from body fluid traces. In addition, our preliminary results clearly indicate that Raman spectroscopy combined with advanced statistical analysis is capable to (ii) differentiate species based on the blood traces. Another area of potential expansion of the method capabilities is (iii) genetic profiling, which is of great importance for forensic science. The overall project includes two major phases: Phase 1 targets the development and evaluation of the proposed novel methodology under controlled (laboratory) conditions. Phase 2 will involve building and certifying a portable easy-to-use automatic instrument based on the requirements identified in Phase 1.

MAIN BODY: Module—Forensic Science Impact in the Court Room

Author: Mr. Ronald Stevens

1. Introduction

This study will examine the extent to which forensic DNA technology is exploited in the range of criminal investigations within Schenectady County and to assess the impact of its use in the outcomes of the individual investigations and subsequent proceedings in court. Published studies suggest benefits to the expanded use of DNA technology (Cascio, 2000; Weedn and Hicks, 1998). Under this project, training will be developed and instructed by NERFI staff for first responding officers on the identification and collection of potential DNA evidence at "routine" crime scenes – a "routine" scene being one in which the department's crime scene/evidence collection unit would not typically be deployed.

To further facilitate the use of forensic services – especially DNA testing – an integration system initiative through Porter Lee will be set up to link the evidence management and tracking systems of the Schenectady DAs office with the Schenectady PD and with the NY State Police Forensic Investigations Center. When fully implemented, the Schenectady County initiative will serve as the pilot for a wider proposal (subject to available funding) linking ten Counties in the Capital Region as described herein.

Schenectady County is one of sixty-two counties in New York State, located in an area known as upstate New York or the Capital District area and one of ten Counties within the NY State Police Troop G territory. Schenectady County is 206 square miles in size with a population of more than one hundred and fifty one thousand. The District Attorney, Robert Carney, holds the highest office in the county judicial structure. Seven organized law enforcement agencies operate within the County, the largest of which is the Schenectady City Police Department. DA Carney's office is staffed with approximately thirty-seven employees. The Schenectady Police Department patrols the City and is the sixth largest law enforcement agency within the State of New York. All law enforcement forensic needs of the county are performed by the New York State Police Forensic Investigation Center located just a few miles away in Albany.

The New York State Division of Criminal Justice Services (DCJS) annually provides the following data in the *"Crime Index Crime Summary 2008"* (New York State Division of Criminal Justice Services, 2008) Schenectady County years 2007 vs. 2008 statistics indicate an increase in all categories except motor vehicle theft as illustrated:

				Forcible		Aggravated				MV
Year	**Total**	**Violent**	Murder	Rape	Robbery	Assault	**Property**	Burglary	Larceny	Theft
2007	**5,355**	**689**	5	36	288	360	**4,666**	1,026	3,282	358
2008	**5,634**	**740**	9	43	311	377	**4,894**	1,124	3,482	288
% change	**5.2%**	**7.4%**	80.0%	19.4%	8.0%	4.7%	**4.9%**	9.6%	6.1%	-19.6%

Law enforcement's success in solving crime depends on its ability to identify, collect and preserve physical evidence of probative value and then to present that evidence to a

forensic laboratory for analysis. Based on studies cited above, the recovery of DNA evidence in any criminal investigation is expected to improve the "solvability factor" significantly through the exclusion or inclusion of suspects or particular acts and has been demonstrated to significantly affect the course of the investigation and subsequent judicial process (Becker and Dale, 2007; Dale and Becker, 2005).

To determine the scope and extent of the use of forensic DNA evidence within Schenectady County, the primary service provider, the New York State Police Forensic Investigation Center was asked to provide statistics on cases presented to them from all law enforcement agencies within the County. The following graph displays a summary of the data provided:

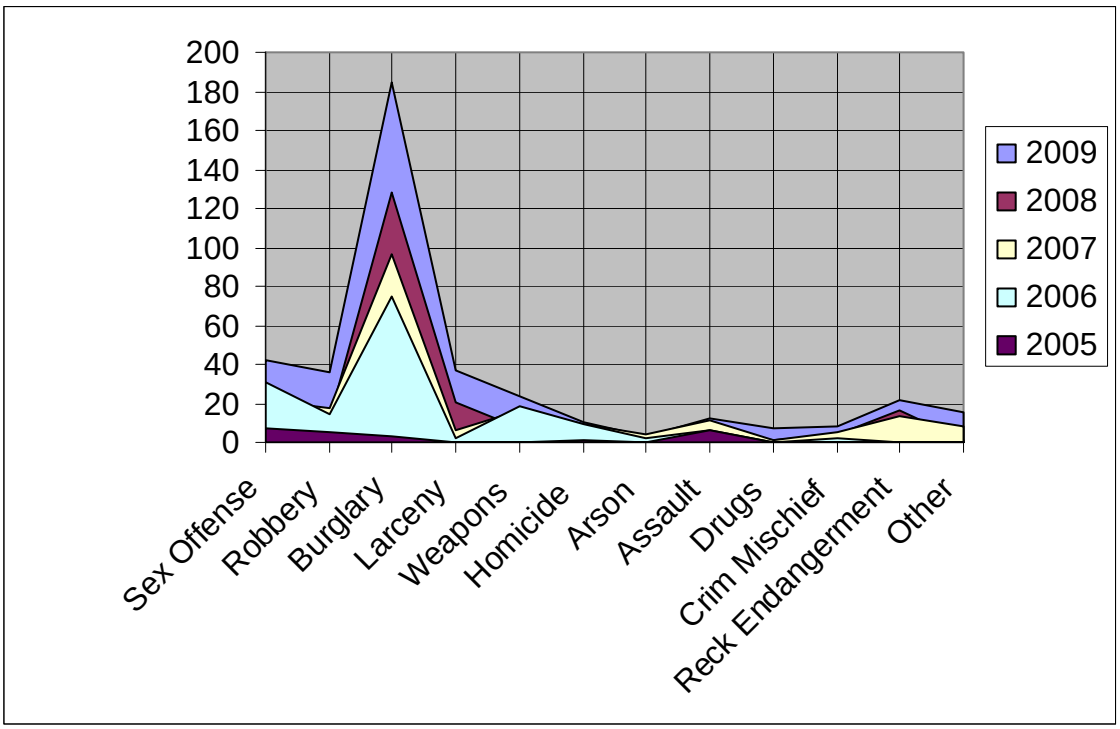


Figure 1. Comparing the results of the Index Crime Summary to the Laboratory's DNA statistics on the same four (4) categories for period 2007 vs. 2008 provided the following results.

Larceny
The Index Crime Summary 2007 vs. 2008 revealed an increase of 6.1% on reported Larceny investigations whereas the laboratory reports an increase of 55.6% in Larceny cases submitted with DNA evidence.

Burglary
The Crime Index 2007 vs. 2008 revealed an increase of 9.6% on reported Burglaries investigations whereas the laboratory reports an increase of 14.3% in Burglary cases submitted with DNA evidence.

Robbery
The Index Crime Summary 2007 vs. 2008 revealed an increase of 8% on reported Robbery investigations whereas the laboratory reports a decrease of 17.2% in Robbery cases submitted with DNA evidence.

Homicide
The Index Crime Summary 2007 vs. 2008 revealed an increase of 80% on reported Homicide investigations whereas the laboratory reports a decrease of 38.4% in Homicide cases submitted with DNA evidence.

During the period covered by these data, it is significant to note the dramatic increase in physical evidence submissions from law enforcement agencies within the County requesting DNA analysis – especially for "non-traditional" DNA crimes.

2009 New York State Index Crime statistics are not yet available but will prove to be extremely interesting as DNA submissions 2008 vs. 2009 increased for the reported categories by 23.8%.

2. Methods

For purposes of this project, DA Carney authorized a part-time employee (a representative of the Northeast Regional Forensic Institute) to be placed within his office to serve the following functions:

- Liaison between the police agencies, the District Attorney's Office, and the NYSP Forensic Investigation Center
- Facilitate training for first responding (patrol) officers of the Schenectady PD in the recognition, detection, and collection of forensic evidence - a total of eighty officers (Appendix 17).
- Coordinate the acquisition and installation within the District Attorney's Office of a fully automated evidence inventory and disposition tracking system (purchased from Porter-Lee, Schaumberg, Illinois) that is fully compatible with the system in use by the Schenectady PD.
- Identify and measure evidence collection data and, to the extent possible within the period of this study, to monitor case outcomes:
    - Types of crime reported
    - Amount and types of evidence collected
    - Identify opportunities to apply new forensic technologies
    - Determine impact of evidence and new forensic technologies on case outcomes

Plea bargains and verdicts were reviewed. If evidence could have been collected that may have improved the outcome of the case, police work was also reviewed. If the

prosecutors could have further enhanced the case by seeking other evidence after arrest, this was explored.

This project also attempted to evaluate the use and effectiveness of forensic technology in the courtroom. Verdicts were to be examined and jurors asked after trial to voluntarily take a survey to assist in determining how important the presence or lack of, forensic science was in the verdict; however, as noted below, timely access to the jurors and to individual cases before the courts was precluded by legal and policy considerations.

3. Results and Conclusions and Implications

Attempting to identify case outcomes became very problematic as the gathering of case information critical to this study was challenging. Through a cooperative arrangement with the State Police Forensic Investigation Center, a spread sheet was created from case information gleaned from the laboratory DNA case submission information for years 2006 through 2008 but significant problems were encountered which precluded the populating of this record as conceived in the original project proposal. Historical case records are all paper boxed, labeled and stored at an off-site location to the limited office space. The existing District Attorney's office electronic records did not contain detailed information on the nature of specific evidence items collected and forensic analysis outcomes. In addition, the length of time required for a criminal case to go from arrest and charging action to court disposition proved counterproductive to the strategies originally conceived for this study. A significant amount of time was spent in the Special Victim Unit which practices vertical prosecution, trying cases of domestic and sexual assault. During the grant period, some time was spent actually observing trial testimony of witnesses including testimony of forensic laboratory analysts in cases in which DNA evidence was introduced.

One noteworthy case resulted in a sentence of 75 yrs. to life. DNA evidence played a crucial role in the investigation of a sexual assault as the assailant was unknown and unidentifiable to the victim who was rendered unconscious due to strangulation and very severely beaten.

A review was conducted of sentences imposed in ten historical sexual assault cases. Five of the cases resulted in plea bargaining and five resulted in trial. In the cases in which plea bargains were negotiated, the sentences imposed averaged twelve years. The cases that went to trial resulted in an average sentence of forty-two years.

An anonymous and voluntary survey form was designed to be completed by jurors when available; however, upon review by the District Attorney and Counsel to the DA, it was determined that such a survey would prove problematic based on a recent court decision weighing post-trial interview and survey of jurors.

As an ancillary development in this project, the DA's office requested a review of two 1995 cold case homicides with DNA evidence that remain unsolved. After reviewing the investigations and analyzing the evidence secured, a spread sheet was created organizing and coordinating the evidence recovered at the scene. Several new items that potentially

contain DNA evidence were identified. Several meetings were held with the District Attorney's office, Schenectady Police Department and the New York State Police Forensic Investigation Center resulting in the submission of new evidence for testing and resubmission of older DNA evidence.

"Crime Scene Analysis Evidence Collection and Packaging" training was conducted for eighty one (81) officers (first responders) attending four 4-hour training sessions at the Schenectady Police Department on forensic evidence recognition and collection (Appendices 17 and 18) by NERFI staff covered under this grant. In addition, all eighty-one officers received four hours of further training by Schenectady Police Department staff (not covered by this grant) which included a redirection on policy and procedure with regard to "routine" crime scenes. The Schenectady Police Department is implementing a new policy where first response officers collect appropriate forensic evidence from misdemeanor offenses, primarily automobile break-ins, for submission to Forensic Investigation Center. The program included an overview of the New York DNA Database program and qualifying offenders who are required to provide DNA specimens to the Database (Appendix 18). Other lectures were given on the capabilities of a forensic laboratory, the potential for effective use of DNA evidence, and general contamination and collection issues. A Schenectady Police Department detective and supervisor of the Crime Scene Unit provided a four hour agency specific collection and processing training session to the officers, alternating with the forensic laboratory training lectures.

The training provided officers with an overview describing the benefit of forensic evidence to a criminal investigation, this included discussion of cases where forensics was critical to identifying and convicting the perpetrator of the crime.

This grant provided the ability to purchase an evidence management system ("The Beast" by Porter-Lee corp.) for the Schenectady County District Attorney's Office. A system from the Porter-Lee Corporation was purchased due to its compatibility with the Laboratory Information Management System (LIMS) developed by the same corporation currently in use by the Schenectady PD evidence management unit and the New York State Police Forensic Laboratory. Not only will this evidence management system be used to record and track the evidence in the District Attorney's office it will also be used to track the volumes of boxed case material needed for prosecution. These boxes are stored throughout the District Attorney's office until they are moved offsite. Additionally the Schenectady Police Department is a subscriber to the Porter-Lee system and the largest contributor to case load for prosecution by the District Attorney's office. The system has the potential of allowing the District Attorney's office to remotely view the evidence and case records of the Schenectady Police Department. This alone would benefit a prosecutor in case preparation and evidence transfer for trial along with assisting the two agencies in decisions needed for evidence retention and destruction. The installation of this system would have made this study much more successful. Unfortunately, full installation and operability could not be completed within the project cycle.

The case management system will be delivered and installed in the District Attorney's office by Porter-Lee during the month of March 2010. Though this grant will have expired, the project manager retained under the grant, Ron Stevens, has agreed to assist in the implementation and training to insure the successful implementation of this system.

When implemented, the Schenectady PD/Schenectady DA/State Police Forensic Investigation Center integration project will serve as a pilot for a wider, ten County, Capital Area Initiative linking these agencies with the Forensic Investigation Center. The concept and basis for the expanded initiative is described further below.

Bipartisan legislation now pending in Congress would require the federal government to collect data on the number of untested rape kits nationwide, and to prioritize the testing of this evidence. A pending House bill also would provide incentives to state and local law enforcement to eliminate the backlog of untested kits. The U.S. Department of Justice has declared its support for initiatives that help assure rape kit evidence is processed on a timely basis. When implemented, the integration project will help assure that criminal justice agencies can readily and effectively identify criminal cases in which physical evidence has been recovered and track the progress of laboratory analyses.

Other federal legislation has been proposed to establish DNA evidence retention requirements. The legislation is in response to expressed concerns that such evidence may be subjected to testing using technological advances unavailable at the time of the original crime and ensuing investigation that may lead to new information bearing upon the guilt or innocence of suspects or persons convicted of those crimes.

Criminal justice agencies face significant challenges in maintaining the integrity and control of the wide variety of physical evidence recovered during the course of criminal investigations. Automated systems are available from several sources that are designed to aid in inventory control, in documenting the chain of custody as required when introduced in subsequent court proceedings, and in tracking the progress of forensic processing of the evidence.

The NYSP Forensic Investigations Center uses a Laboratory Information Management System (LIMS) to track all cases in which physical evidence is submitted for laboratory analysis. The system employs bar-coding of physical evidence specimens and is a product developed by the Porter-Lee Corporation, Schaumburg, IL. The same system is used by many of the public crime laboratories operating in New York State. Porter-Lee also produces an evidence management system (EMS) which is used by many police departments across the state.

Implications for the field of criminal justice and the study of forensic science:
This project will prove helpful in identifying and addressing the technical specifications and operating rules necessary to (1) effectively track and document the movement of physical evidence from the point of collection to . . . (2) storage at the police agency to . . . (3) transfer to the laboratory for forensic analysis to . . . (4) return to the police agency to . . . (5) transfer to the District Attorney for introduction in a court proceeding and . . . (6) ultimate return for retention at the police evidence storage and control. Throughout this movement of the physical evidence, interested agencies (the PD, DA and Lab) will be able to independently inquire

electronically in real time on the location of the evidence, the status of pending forensic tests, and obtain reports of the results of forensic testing on the evidence. The documentation involved in the transfer of evidence will be accomplished without redundancy in data input as the two-dimensional bar-codes employed will automatically track and capture the specimen identification number with descriptive information, the relevant case information, and the submitting and receiving agencies identifying information and points of contact.

The acquisition of computer equipment and software for other District Attorneys offices in the capital region and for selected law enforcement agencies within those counties will proceed as resources are available. The police departments at Albany and Troy, the Rensselaer County Sheriff's department, and Albany District Attorneys office already use the Porter-Lee evidence management system. In discussions held with State Police administrators and IT managers as well as FIC managers, there is strong interest in expanding this system to include the major users of forensic testing services available through the FIC.

With the improved capacity to manage forensic evidence, training will be provided to departments with an emphasis on the effective use of DNA technology to resolve a wider variety of crimes. Traditionally, DNA technology has been applied primarily in the investigation of violent crimes. Crime data show that there are nearly twice as many burglaries in New York State as compared to the number of violent crimes – with most burglaries going unsolved. Recent advances in the technology, coupled with the expanded processing capacity of public forensic DNA laboratories (largely accomplished with federal funding assistance), provide opportunities for its effective use in property crimes. According to recent studies, DNA is more likely to be recovered from a crime scene than a fingerprint, and DNA Databank "hits" linking offenders to evidence from a violent crime showed that in 75% of the cases the basis for collecting the offender's DNA specimen was a conviction for a "lesser" crime such as a burglary or drug offense. The Capital Area Integration Initiative will assure the first responding officers to a burglary scene will be equipped and trained to recognize potential DNA evidence and properly collect it in a manner that meets the requirements of the forensic laboratories and, ultimately, the courts. In order to insure the quality and reliability of DNA analysis services provided to the law enforcement agencies in the capital region and to assure analysts are prepared with the latest and most efficient analytical techniques, highly specialized training will be provided to forensic DNA laboratory personnel in workshops and seminars.

Finally, the immediate success of Forensic Science Impact in the Courtroom can not be judged solely on the statistics and information provided from the results of this study. Instead, future studies may be warranted to determined the overall success from the advise, instruction and equipment provided under Module—Forensic Science Impact in the Court Room.

MAIN BODY: Module - ChatMinder—A Safe Internet Tool for Parents

Author: Dr. Tomek Strzalkowski

1. Introduction

The objective of ChatMinder project is to conduct and deliver a study of dialogues occurring in the on-line chat rooms. A secure chat-room at the ILS Institute will be used to collect chat data from experiments with recruited subjects (SUNY students under an IRB protocol). Approximately 20 hours of chat involving groups of 3 to 6 people on topics ranging from movies to organic food to state of the economy will be collected and analyzed for this study.

One hypothesis pursued is that an automatic agent could be developed to operate in a live chat room, monitoring conversational behavior. When this automatic agent detects behavior that is unsuitable, it can try to intervene. There is a range of actions such an agent could take, including reporting malicious behavior to some administrator. However, a more subtle intervention could be an effort at changing the topic of conversation, presumably away from potentially dangerous areas, to safer topics.

2. Methods

Methods used in this research include data collection and annotation, conversational modeling, and software development and testing.

1 Data

Chat data was collected through controlled exercises with participants in the secure chat-room located at ILS labs. While large volumes of data may be obtained from public chat-rooms, it was of limited value for the type of modeling tasks that were of interest in part because of the high-level of noise, lack of focus, and rapidly shifting, chaotic nature, which makes any longitudinal studies virtually impossible. Public chat-rooms may be excellent sources of data for studies involving on-line language usage (e.g., novel uses of vocabulary, syntax), general conversational etiquette, and related issues. However, for deriving more complex models of conversational behavior, this project required the interaction to be reasonably focused on a task and/or social objectives within a group. There are a number of available resources that demonstrate that data collected under monitored conditions is still effective for modeling language (cf. ICSI-MRDA corpus (Shriberg et al., 2004), Switchboard corpus (Jurafsky et al., 1998) and Map Task corpus (Anderson et al, 1991)). The purpose of conversational data analysis was two-fold: (1) understanding how certain social behaviors are reflected in language use, and (2) building an automated chat agent that could effectively use appropriate linguistic forms to achieve certain social objectives in the chat-room.

One specific social behavior that was focused on in this project included Topic Change, which was an attempt by a participant to change the flow of discussion from one topic to another. Modeling this behavior was of interest because it was directly applicable to a future Chat Minder prototype. To ensure that this social behavior was present in our chat

data, a multi-tiered collection process was devised in which the subjects started from simple, free-flowing conversations and progressed towards more complex and structured interactions. Approximately 20 hours of chat dialogue spread out over 14 sessions of 90 minutes each were collected during this study, amounting to a total of 7317 individual utterances.

2 Data Annotation

The data was annotated for presence of linguistic elements that correlate with social behavior on three different levels, as detailed in Sections 2.1 through 2.3 below. In Section 2.4, we describe the annotation tool used to facilitate annotation. Section 2.5 describes the annotated set and provides pertinent statistics.

2.1 Communication Links

Communication links capture associations between utterances. Utterance includes the sequence of words that are entered by a participant in a single turn in chat. In multi-party chat the relationships between utterances are often ambiguous; it is not readily apparent who is speaking to whom, particularly when there is no addressing information in the utterance. However, it is an integral part of understanding how social behavior is manifested in language, as we need to determine between which participants the conversation flows. These are situations where one participant's utterance responds or relates to a previous utterance by another participant. This also includes situations where one participant is addressing another participant or a group of participants who may subsequently respond to him or her.

At this level, we were interested in annotating utterances that were addressed to the entire group or to some specific participant, uttered in response to a specific previous utterance, or are continuations of a previous utterance by the same participant. Accordingly, there are 3 possible communication links, one of which was assigned to each utterance in our annotated data: Addressed-to (a specific participant), Response-to (a specific prior utterance by a different participant), and Continuation-of (a specific prior utterance by the same participant).

2.2 Dialogue Acts

The functional or dialogic aspect of an utterance has to do with its role or purpose in conversation. Statements, questions, answers, offers, acceptances and rejections, as well as expressions of thanks are all examples of such functions in a dialogue, which we call Dialogue Acts (DA). At this level our objective was to capture how an utterance functions in dialogue, which may or may not be directly related to its form. For example, the utterance "Can you close the window?" can function as a question or as a directive, depending upon the context in which it is used. It was therefore important to consider the context of the utterance to make a decision of which dialogue act label to assign. We created a hierarchy of 21 dialogue act labels, which we grouped into three top-level categories, namely – Statements and Responses, Questions and Directives and Conversational Norms.

By marking utterances at this level, we wanted to identify the pragmatic composition of an utterance. This is especially useful when trying to discover certain social behaviors, such as the change of conversation topic. The mechanism of changing the topic may vary by participant, the kind of topic change, and by the time at which it is attempted relative to the sequence of conversation. A participant may introduce a new topic by asking a question about that topic, thereby obligating the other participants to respond to them; or they may make a provocative statement and thus attempt to steer the conversation towards another topic. We wanted to determine whether there was any correlation between utterances marked as topic change utterances (explain in subsection 3.3 below) and the dialogue act assigned to them.

2.3 Topic and Focus of Conversation

A topic is the subject matter under discussion in some part of the dialogue. Focus is another semantic property of an utterance and it pertains to an entity or event that is the most salient in the utterance. The difference between topic and focus is often only the matter of degree: topics are larger subjects (such as technology, movies, or politics), while foci are more narrow subthemes within a topic (e.g., cell phones, or Tom Hanks, or Eiffel Tower). However, it is often impossible to define topics and focus *a priori* (e.g., Tom Hanks may be a focus of a discussion about movies; but it would be a topic if the discussion is about his career, while a specific movie he starred in may be the focus.)

In this exercise we primarily concentrated on topic and focus *changes*. In other words, we wanted to be able to tell where a topic/focus starts and when it ends in a dialogue. We postulate that annotating topic change and focus shifts will help us determine behavior, which induces a modification of the flow of conversation. We asked annotators to determine the topic and the focus of each utterance using labels of their own choosing (based on what is being said). When the subject of the dialogue changed in any significant way, we asked the annotators to change the topic or the focus label. That utterance would then denote the beginning of a new topic. Note that in chat room conversations, due to the asynchronous nature of dialogue and multiple participants, there may be different topics being discussed within a short period of time, often being interleaved with one another. The same may be the case for focus as well.

2.4 Chat Annotation Tool

An annotation tool (developed under another project) was used to facilitate the annotation process. This tool was implemented in Java and deployed as a web applet, so that annotators were able to work from remote locations. All annotators were trained on a sample set of data. Since our annotation paradigm is multi-tiered, we asked the annotators to annotate each level in separate passes through the data to reduce cognitive effort. This also had the effect that annotators become more familiar with the data with each pass. Figure 1 shows a screenshot of the Chat Annotation Tool.

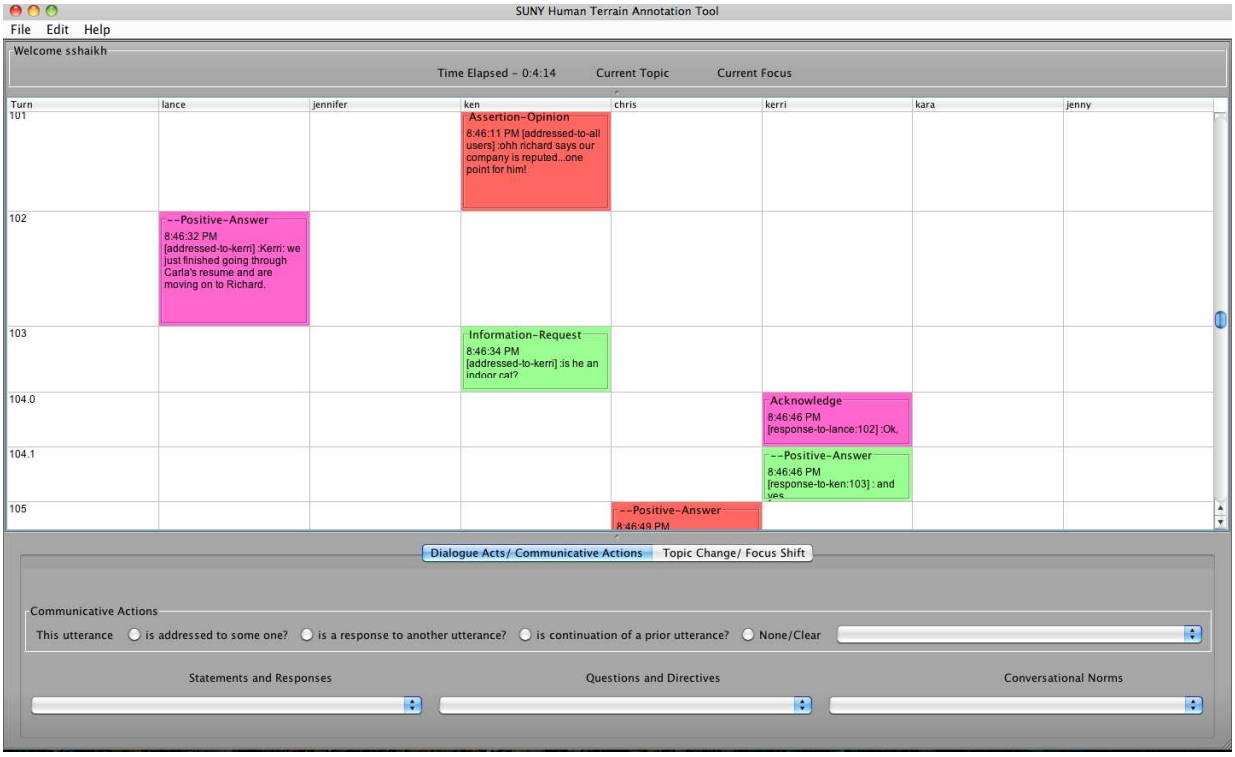


Figure 1. Screenshot of Chat Annotation Tool

In the screenshot above, note that every utterance by the participants was represented as a row and all participants were separated into columns. Annotators mark the communicative links and dialogue acts which show in the utterance cell – e.g turn 102 by participant Lance was marked as "addressed-to-Kerri" communicative link and a "Positive-Answer" dialogue act. Each different color represents a separate topic, so in this short snippet of conversation in the screenshot there were three, albeit interleaved, topics that were annotated.

2.5 Annotated Data Set

Of the 14 sessions we collected, we selected 10 for annotation, with at least 3 annotators for each session. In Table 2 some of the overall statistics computed from this set are shown. We computed inter-annotator agreement on all three levels of our annotation, i.e. Communication Links, Dialogue Acts and Topic/Focus Shifts. Topic and Focus shifts had the highest inter-annotator agreement scores on different measures such as Krippendorf's Alpha and Fliess' Kappa. In Figure 2, we show inter-annotator agreement measures on Topic/Focus shift annotation. With such high degree of agreement, we can reliably derive models of topic shift behavior from our annotated data.

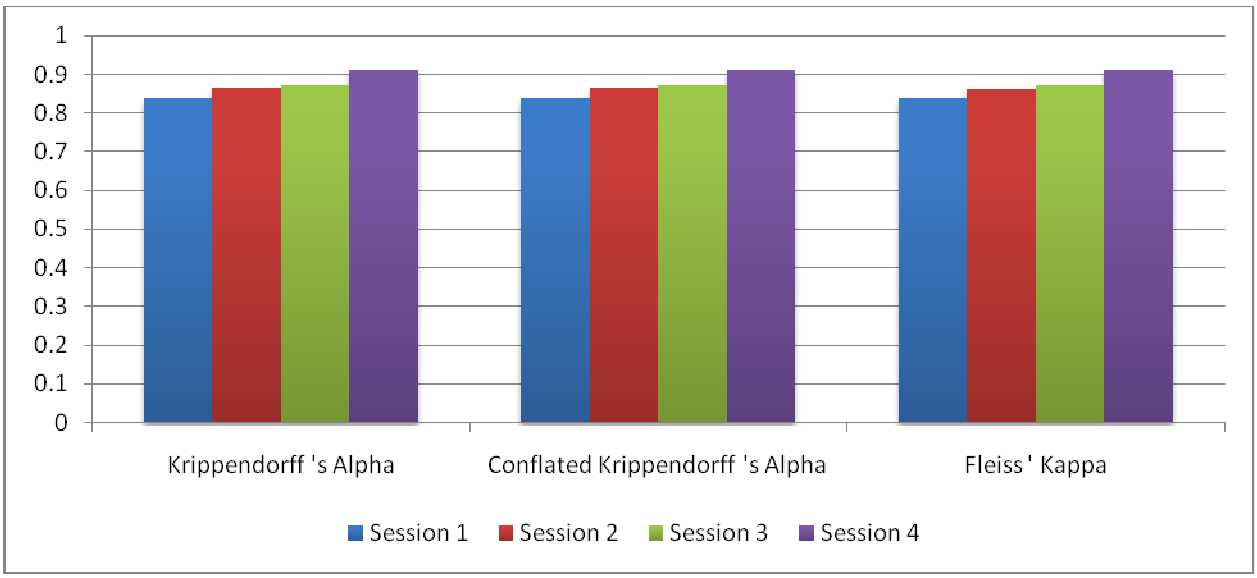


Figure 2. Inter-annotator agreement measures for Topic/Focus shifts

Total Number of Sessions Annotated	10
Number of annotators per file	3
Total Utterances Annotated	4640
Average number of utterances per session	~520
Total topics identified per session	174
Total topic shifts identified per session	344

Table 2. Selected statistics from annotated data set

3. Results

Analysis of the collected data led to construction of preliminary models of social behavior in online discourse. Conversations were annotated for communicative links, dialogue acts, and topic and focus shifts, which created the basis for building computational models of conversational behavior. Some of these models, e.g., how to effectively change the topic of conversation, were subsequently implemented into an automated Virtual Chat Agent (VCA), a Chat Minder prototype. VCA has been demonstrated to perform effectively and convincingly in Internet conversation with human participants.

3.1 Conversational Modeling for VCA design

A virtual chat agent is an automated program with the ability to respond to utterances in chat. Our VCA was distinctive in its ability to participate in multi-party chat and manage to steer the flow of conversation to a new topic. We exploit the dialogue mechanism underlying HITIQA (Small et al. 2009) to drive the dialogue in VCA.

The topic as defined by the information contained in the participant's utterance was used to mine outside data sources (e.g., the web) in order to locate additional information about that topic. The objective was to identify some of the salient concepts that appear

associated with the topic, but are not directly mentioned in any recent utterances. Such associations may be postulated because additional concepts are repeatedly found in many web pages near the concepts that are mentioned in chat.

Based on our annotated corpus, we also determined that a common method that participants employ to achieve a topic change in conversation is to introduce a new concept (or aspect) that is shared between the current topic and the new topic. This was schematically illustrated in Figure 2 below, where the current conversation topic (technology) was changed to a new topic T2 (music bands). Participant K introduces the topic of music bands by finding a common concept that forms a bridge between these two topics, which is "Lars Ulrich". By introducing Lars Ulrich K opens the window for the conversation to shift to music bands, which indeed happens as speaker N picks up on the association, as intended. Subsequently, K clinches the transition by adding another utterance on the new topic.

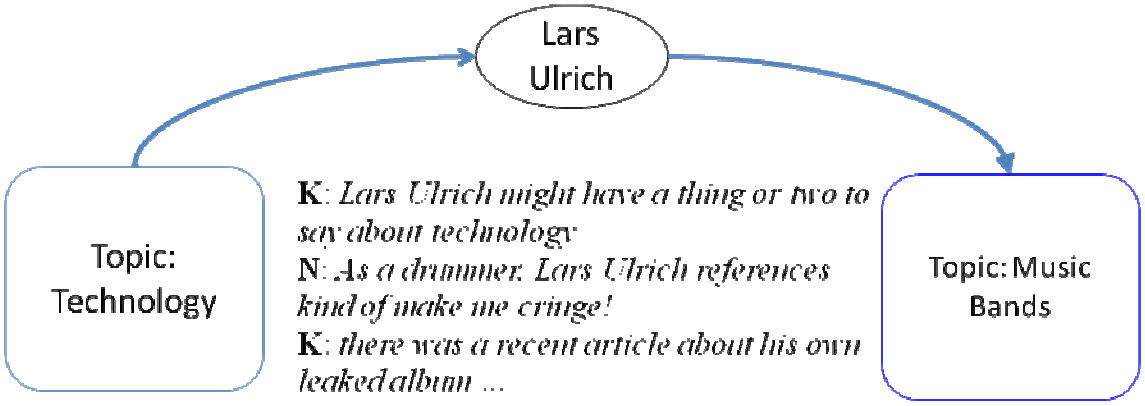


Figure 3. Example of topic change by participant K.

The effect of topic change was apparent when a subsequent utterance by another participant is about the same topic. This is a successful attempt at changing the topic. Below is an example of topic shift annotated in our chat data collection.

AA: *My kids tell me the band you're going to hear (dc for cutie) is great.*
(TOPIC: music bands, dc for cutie)

KA: *oh cool! Their lyrics are nice, I think. What kind of music do you guys listen to?*
(TOPIC: music bands, dc for cutie; BRIDGE: music)

KN: *I don't really have a favorite genre….you on youtube right now?*
(TOPIC: you tube)

Example 1. A topic change in dialogue

Note that in this example, the second part of utterance by participant KA – "what kind of music do you guys listen to?" is deployed to shift the topic from "music band, dc for cutie" to "music". This, in turn allows participant KN to shift the topic to You Tube.

Another example is participant KA's utterance in the transcript below inducing participant KI to change her topic from navigation to texting:

KI: *I actually went on a 3 hr drive yesterday, i just used a map instead of his GPS*
(TOPIC: navigation technology/GPS)
JR: *hahaha its still possible to exisit without all the technology but google and mapquest have def made my life more livable-*
(TOPIC: navigation technology/google, mapquest)
KA: *what about text messaging*
(TOPIC: technology, BRIDGE: text messaging)
KI: *I hate texting. I'd rather just talk on the phone*
(TOPIC: text messaging)

Example 2. Another topic change in dialogue

We found this model of topic change fairly consistently exhibited, where the participants would ask an open question, to get other participants to respond to them, thereby changing the course of conversation. We collected all utterances marked topic shifts and created a set of templates from them. These templates served as a model for the VCA to utilize when creating a response.

Another model of behavior that we found as a consequence of topic change was topic sustain. This is an instance where the utterance was marked to be on the same topic as the one currently being discussed, for example, JR's utterance in the Example 2 above. Topic-sustain utterances typically offer a new in-topic aspect, but provide no bridge to another topic as noted previously. Typical linguistic forms used were offers of support, an agreement with a previous utterance, or a question about any known in-topic aspect.

3.2 VCA Software Architecture

The chart in Figure 4 shows an overall architecture of the VCA.

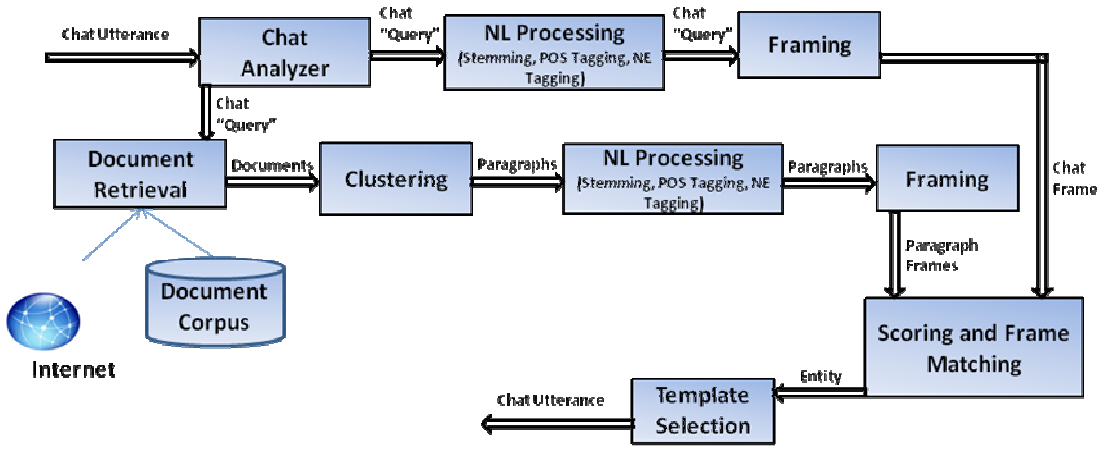


Figure 4. VCA Architecture and Components.

Chat Analyzer

Every utterance in chat was considered to be a candidate for response by the VCA. It was first analyzed by the Chat Analyzer component. This process removes stop words, emoticons and punctuation, as well as any participant nicknames from the utterance. We postulate that the remaining content bearing words in the utterance represent the topic of that utterance. We call this analyzed utterance our chat "query" which was sent in parallel to the Document Retrieval and NL Processing component.

Document Retrieval

The document retrieval process retrieves documents from either the web or a test document corpus. We use Google AJAX api for our web retrieval process and InQuery (Callan et al., 1992) retrieval engine for our offline mode of operation. The test document corpus was collected by mining the web for all utterances in our data collection, thus creating a stable document set for experimental purposes. In addition, this test corpus ensured that we had a collection of documents pertinent to the realistic topics that were discussed by the participants in chat sessions. Currently, the document corpus contains about 1Gb of text data. We retrieve, on average, 20 documents per chat query, for both online and offline modes of operation.

Clustering

Paragraphs in documents were retrieved using clustering method in Hardy et al. (Hardy et al. 2009). This process groups the paragraphs containing salient entities into sets of closely associated concepts. From each cluster, we choose the most representative paragraph, usually called the "seed" paragraph for further NL processing. Each seed paragraph and the chat query undergo the same further NL processing sequence.

NL Processing

We process each chat query by performing stemming, part-of-speech tagging and named-entity recognition on it. Each seed paragraph is also run through same three natural language processing tasks. We are using Stanford POS tagger for our part-of-speech tagging. For named entity recognition, we used BBN's IdentiFinder.

Framing

Frames were from the entities and attributes found in both the chat query and the paragraphs. A frame is a structured template that provides salient information about the underlying text. This work extends the concept of framing developed for HITIQA (Small et al, 2009) and COLLANE (Strzalkowski, 2009). The frame built from chat query was the Chat Frame and those frame built from paragraphs are Paragraph Frames. Framing provides an informative handle on text, which can be exploited to compare the underlying textual representations, as we explain in the next section.

Scoring and Frame Matching

Using the information in the frames built in the previous step, we compared the chat query frame to the frames created from the paragraphs. We assign a score for each paragraph frame based on how many attributes and their corresponding values match. Note that for a paragraph frame to be eligible for scoring it should contain all the attributes found in the chat query frame. Of all the paragraph frames we select the highest scoring frames and select the attribute-value pairs that are not part of the chat query frame. For example, as show in Figure 5 below, the chat utterance "Aruba might be nice!" created the following chat query frame.

[POS]
NNP, Aruba
JJ, nice
[ENT] PLACE

Figure 5. Example chat query frame

Correspondingly, we will select all PLACE type entities from the highest-ranking paragraph frames. These are shown in Figure 6 as Aruba Entity list. The entities "NASCAR", "Women Seeking Men" and "Mateo" are not of entity type – PLACE, we assign them a score of 0. Only the entities that match the type of entity in the chat query frame get a positive score. This score is the frequency of occurrence of that entity in the paragraph; in this example it is found to be 1. Assigning scores by frequency of occurrence ensures that the most commonly occurring concept around the one that is being discussed in the chat query utterance will be used to respond with.

Template Selection

Once we have chosen the entity to respond with, we select a template from the set of templates for that entity. These were templates that were created based on the models created from topic change utterances annotated in our data set. For a select group of entities, which were quite frequently encountered in our data collection such as PLACE, PERSON, ORGANIZATION etc., we have a set of templates specific to that entity type. We also have several generic templates that may be used if the entity type does not match the ones that we have selected. For example, a PLACE specific template is "Have you ever been to __?" and a PERSON specific template is "You heard about __?". Not all templates are formulated as questions. An example of a generic template is "__ rules!".

User1: ***Aruba might be nice!***
[POS]
NNP, Aruba
JJ, nice
[ENT] PLACE

Aruba Entity list:
VALUE = NASCAR AND TYPE = ORGANIZATION AND SCORE = 0
VALUE = Dallas AND TYPE = PLACE AND SCORE = 1
VALUE = Women Seeking Men AND TYPE = ORGANIZATION AND SCORE = 0
VALUE = Mateo AND TYPE = PERSON AND SCORE = 0
....
VCA: ***How about Dallas?***

Figure 6. Example of frame matching and scoring

3.3 VCA Experiments

Figures 7 and 8 represent examples of the VCA in action in a simulated environment. In both examples the VCA was the participant "renee". Figure 7 shows an example of a topic sustain behavior while Figure 8 shows an example of a topic change behavior.

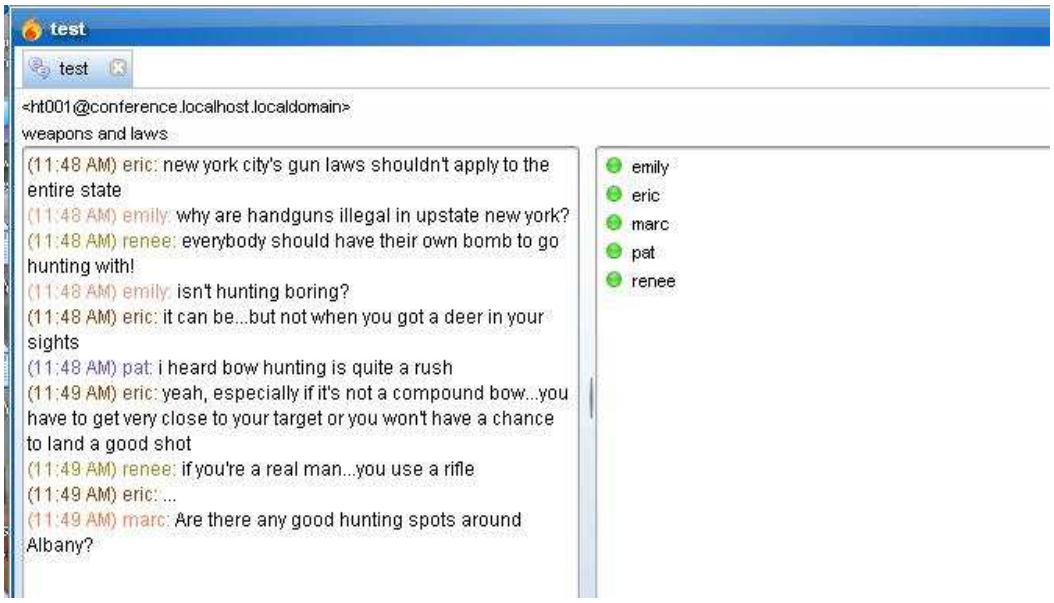


Figure 7. Example of VCA in action: topic change

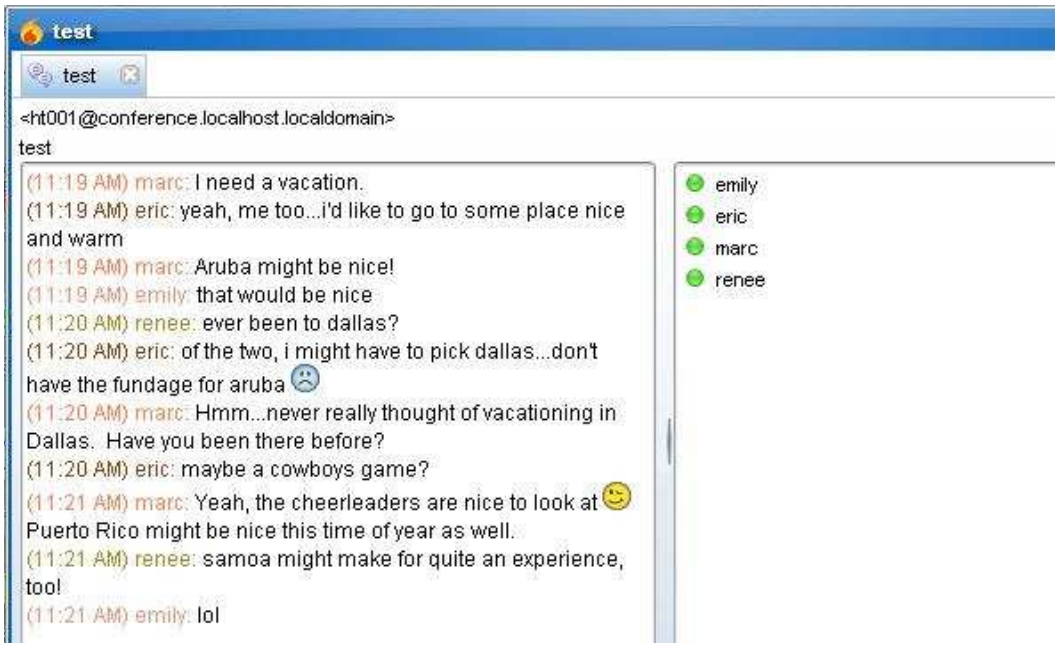


Figure 8. Another example of VCA in action: topic sustainment

3.4 Preliminary Evaluation

We have developed a two-stage evaluation protocol in order to test the effectiveness of the VCA prototype in a realistic setting. In the first stage, we tested the performance of the VCA in the laboratory chat room with human participants. At some point during the dialogue, one of the participants was tasked to withdraw from the chat and (silently) pass the control to this VCA for a period of time. This participant remains on-line but all conversation was now conducted by the VCA. After N minutes, where N is an experimental parameter, the control of conversation returns to the human user. In this test, it was important that the other participants were not aware ahead of time when the switch was supposed to occur.

The initial metric of VCA effectiveness was calculated as a proportion of utterances generated by the VCA during the number of utterances made by all participants during this period. These utterances were subsequently judged for appropriateness using the metric developed for the Companions Project (Webb, 2009). In general, utterances that were not on topic or out of place were judged as not appropriate. Such utterances may give off the VCA to other members of the group.

Since a VCA was tasked to perform a specific function in the chat room – i.e., to accomplish a topic change – the effectiveness with which this task was accomplished was also measured. One simple metric that was developed included an index of subsequent mentions of the new topic by others in the discourse. Initially, annotators track the topics manually over the log of the chat (we noted previously that there is a high level of inter-annotator agreement in this category). For the purpose of automating this metric, the topic may be equated with the named entity introduced by the VCA into the conversation for the first time, e.g., “How about Dallas?” introduces Dallas as a topic of conversation if it wasn’t a current topic, and has not been one before (or at least not recently). We count the number of utterances by other participants in the dialogue that contain references to the new topic (e.g., Dallas) in the immediately following dialogue. The topic change was considered successful if at least one subsequent mention of the topic occurs in the dialogue. The topic change was considered lasting if there are at least two subsequent mentions by two different participants. The topic change was considered permanent, if the dialogue does not return to the topic before the change occurred. Evaluation of topic sustain action was performed in a similar manner. In this case the VCA picks up an already introduced topic and generates an utterance meant to induce a response from other participants that would require a mention of the topic, e.g., “What one can do in Dallas?” Again we measure the number of subsequent mentions of this topic, which may be direct or via a pronoun. We also count the number of turns in between these to measure the extent of dialogue over which the topic continues and to what degree it engages the participants.

4. Conclusions and Implications

In the ChatMinder project, a study of language uses had been conducted from dialogues occurring in the on-line chat rooms. At least 20 hours of chat involving groups of 3 to 6 people on topics ranging from movies to organic food to state of the economy had been analyzed for this project. We subsequently performed initial studies with a prototype autonomous chat agent (VCA) that can effectively change the topic of conversation in a

chat room with human participants. VCA technology represents an important advance in automated human-computer communication with potential applications in cross-cultural social modeling, influence operations, advertising, law enforcement, and national security.

It was our continuing hypothesis that an automated agent such as this can be used to monitor live chat rooms where at risk constituents (such as children) can be participating in online conversation. There were a number of flags and filters that can be used to spot potentially troubling conversation. One method of dealing with these was to report transgressors to chat room administration. Another, more preventative measure could be to allow automated agents to change the topic of conversation when it believes the current topic was dangerous or inappropriate. This method could also be used in situations where people were posting inflammatory statements about some issue, and rather than terminate the discussion an automated agent could attempt to divert the course of the conversation, rather than allow it to continue. Of course, these are potential future applications of this technology, which require more extensive evaluation of the current prototype.

For a more realistic future evaluation, a live Internet chat room experiments need to be conducted, where the VCA is entered either autonomously or with a human "handler". Future work needs to focus on transferring VCA technology to the chat-rooms used by children. This may be accomplished in collaboration with researchers at UA School of Social Welfare, the Families Together of Albany County, and Prevent Child Abuse NY, and would include additional studies of suspicious on-line behavior when a change of topic intervention by a VCA may be warranted.

Implications for the field of criminal justice and the study of forensic science:
There are mechanisms for interaction specific to groups we may wish to target. For example, the language used, or social conventions that any automated agent some observe if communication is likely to be successful (e.g., the ECO project currently underway at ILS). The strength of our model is that it doesn't rely on any domain model or prior knowledge, which is costly to include and brittle to maintain. A weakness is that there may have to be adaptation to individual user groups, or some sort of maintained 'hot list' of topics that should be monitored. These will need to be acquired, either from monitoring the groups themselves, or more likely, by law enforcement or monitoring agencies, who are likely aware of the issues facing particular groups or communities.

Future research may thus focus on (a) adapting the findings of this study to other chat-room user populations, specifically to children; and (b) building and evaluating of an autonomous chat agent. The deliverable from this project is the information included in this report.

MAIN BODY: Module—Capital Region Cyber Crime Partnership

Author: Dr. Sanjay Goel

1. INTRODUCTION

Crime data analysis necessitates combining pieces of information from disparate sources to make meaningful deductions. The number of sources continues to increase and the data is complex – making it difficult for forensic analysts to collect and process data manually. Consequently, the process of crime analysis is often slow and labor-intensive. While, it is not possible to replicate human deductive ability tools can be used to process data and reduce the cognitive overload on the analysts. The objective of this investigation is to develop techniques that improving the efficiency and effectiveness of crime data analysis, including, collection and analysis. Crime data comes from several sources, such as: computer logs, past crime records, databanks, and reports. The data can be both unstructured such as reports, articles or structured such as databases and spreadsheets. The goal of this research is to develop crime analysis techniques using both structured and unstructured data. There are three components of this work. Our goal was to develop a set up an infrastructure that will assist in data collection and analysis. This includes:

1) Determination of data sources
2) Creation of data collection robots from online sources
3) Development of natural language processing capability from text data
4) Creation of techniques for tagging and correlating data

Our initial goal was to use the infrastructure to investigate the problem of recidivism in crime (especially sexual crime) based on crime records available as well as information from public sources. However, obtaining criminal data, especially DNA data, proved to be very difficult and we focused our efforts on public data sources and creating the infrastructure to efficiently analyze them.

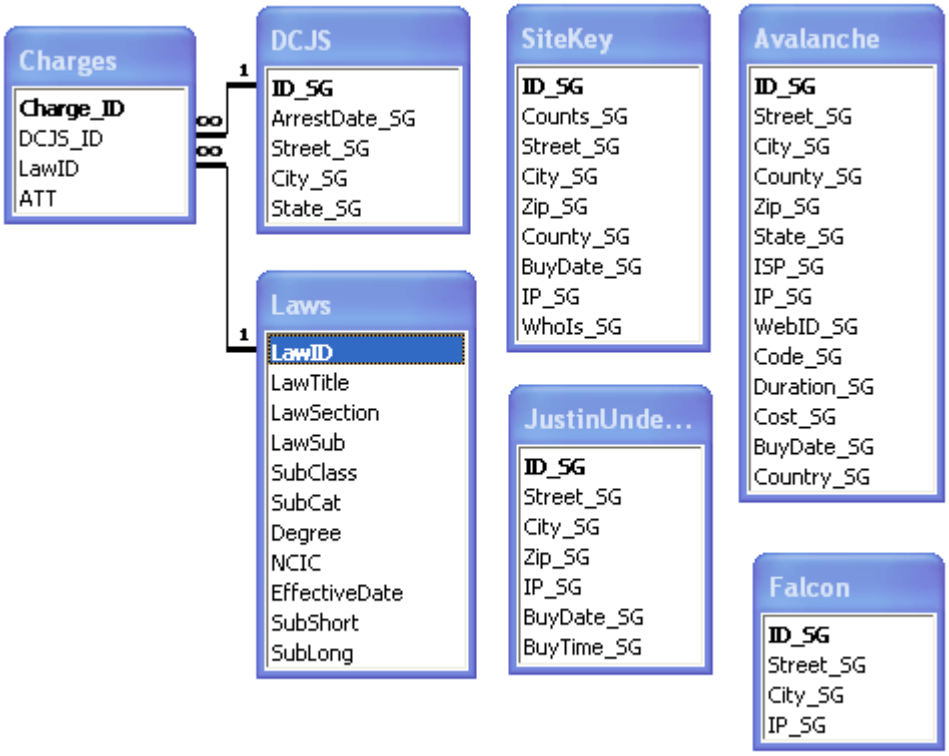
The focus of this work is to build the infrastructure that will allow for efficient data collection, and correlation and analysis of data to draw reliable conclusions. Infrastructure building involved in house code development as well as acquisition of third party tools. Law enforcement already has capabilities of searching through databases efficiently. They now also have standard forensic procedures for electronic data and new procedures are evolving as technology evolves. There are however missing elements of the infrastructure which could add to the repertoire of law enforcement agencies, including: 1) online data collection (e.g. robots), 2) text data analysis, and 3) correlations across disparate data sets.

2. DATA COLLECTION

Data collection for crime analysis has traditionally come from law enforcement agencies and other government sources. Increasingly, data for crime analysis is coming from public sources. Clues to motivations and behaviors of criminals can be buried in their online activities which can provide precursors to future criminal activity or help in

corroborating evidence from other sources. The data set can contain both text and numerical data which needs to be analyzed in conjunction.

The focus of this work has been on Internet Crimes against Children (ICAC) and we have identified several public and non-public sources to compare and analyze. The non-public sources were collected as a part of a previous investigation by the Capital Region Cyber Crime Partnership. The DCJS data includes level 1-3 sex offender data with arrest dates, address information, and charges of those individuals who used a computer to commit their offense prior to 2007. Site Key, Avalanche, Justin, Underscore, and Falcon are sting operations conducted by the New York State Police where data was collected from websites offering child pornography subscriptions which included address, IP, and purchase dates. A database schema including attributes is included below.



We have identified several sources of data including publically available sex offender registries, which can be used to solve important questions related to criminal justice. We are in the process of looking at Perverted Justice chat logs, Wikisposure, pedophile forums (including some which are defunct), PeeJ forums, Absolute Zero and AntiPaedo blogs, Corporate Sex Offender. In addition, business and school information is also being collected.

We hope to be able to develop techniques for identification of online sex offenders through analysis of these sources. Some cyber crime analysis methods can be used for this purpose. Website domain histories can be reviewed as well as IP addresses correlated to geographic locations for associating with known information about an individual.

Identification can also be done through text-analysis and authorship determination (linguistics and psychology).

3. TOOLS AND METHODS

Data Collection Robots

As a part of this work, we have developed data collection robots and models to collect data from publicly available sources using Kapow[1]. This software tool allows scanning of thousands of sites to collect data and insert it into structured databases. As an initial pilot, the robot was configured to collect information on sex offenders listed in the NYS Department of Criminal Justice Sex Offender Registry. This robot can collect information on personal information, location, crime, and conviction of the sex offenders. Robots were also developed to collect information from hacker forums to observe patterns of behavior.

Linguistic Analysis and Identification

We have also developed software that can do content analysis on unstructured data such as chat logs, web pages, online postings, and emails. Unstructured text can be used to profile criminals, identify their motives, and predict their inclination to coming crimes. Natural language processing allows us to analytically identify such markers in text. In addition, tools were developed for parsing unstructured text and identifying keywords. Frequency of usage of keywords in blogs, chat logs, and postings of internet users can be used to understand the psychological behavior of a user. Using data from known offenders and clean data from other blogs can be used to create a classifier. Keyword markers exist for several different attributes, such as anger, fear, sexual behavior, etc. We have programmed these keywords in the code so that we can conduct behavior/motive analysis based on the content. Beyond use of keywords, we also plan to investigate linguistic features in the text that can provide us clues to their behavior.

Correlating information obtained from the web and attempting to associate these with the existing data on sexual offenders can help in detection of potentially dangerous situations where chat room text analysis can expose precursors to sexual crime. Based on preliminary analysis we can identify keywords that predators may use to ask about clothing sizes, genitalia descriptions, and parental presence. Users using such language may be more likely to be sexual predators.

Correlating Disparate Sources of Data

If these predators can be linked to past crime history the likelihood of their committing another crime increases manifold. Another attribute is the frequency of such behavior, i.e. similar behavior with multiple potential victims. We have deployed the Palantir Government platform, which allows us to tag data from disparate sources and correlate the information creating a comprehensive data view for analysts who can then use their expert judgment while analyzing capabilities. Palantir Government is increasingly being used by intelligence and law enforcement agencies to assist analysts. By working with

[1] http://www.kapowtech.com

Palantir, we hope our work will be able to help generate or refine an ontology which can benefit law enforcement and justice entities.

Finally, capability has been developed where information from multiple sources can be exposed as XML (semi-structured) and correlations can be made across different sets of information. SOLR search engine has been deployed for data analysis; work is also under way for installing the Palantir Government analytic engine for more sophisticated data correlations. This infrastructure has multiple tools for data collection and analysis and will be leveraged for a variety of research projects mainly focused on cyber crime. Initial investigations planned using this infrastructure are: 1) to analyze behavior of sexual predators and develop psychological profiles that can be linked to their online behavior expressed through online postings; 2) develop identifiers from online chatter that suggest the onset of a cyber attack.

4. RESULTS

One of the most pernicious dangers of the Internet is its potential use for online sexual predation since it allows predator access to children in a relatively anonymous environment. Researchers believe that increased social activity online through chats and social networking sites (i.e. MySpace, Facebook) have increased the chances of contact between potential predators and their victims. A majority of research has focused on the behavior of offenders using data once a crime has been committed. The objective of this research is to identify the potential for a criminal to commit such activities (either for the first or subsequent times) by evaluation of their virtual identities.

Some of the initial analysis has revealed use of some of the techniques discussed can also benefit parole officers. For example, in just a preliminary analysis of the sex offender registry, it was found that some of the listed residences are hotels or motels (temporary housing). By correlating business data (in the hotel/motel category) we can flag residences which are questionable and should be checked out. Also, sometimes addresses listed as places of work refer to large plazas which can include many different places such as restaurants, grocery stores, and child care facilities in close proximity. Other times, places of business are just residential homes. Sometimes, these residential homes come up for sale on real estate sites and alerts relating to this activity are important. These are actual instances that we discovered on live sex offender registry members.

There are psychological markers that are exhibited in sexual predators that may be detected from their online behavior. Similarly there are psychological traits of potential victims that make them more vulnerable to advances of sexual predators. Evaluating either of the two in isolation may not accurately predict the risk of a potential sexual crime. Four components that contribute, in differing degrees and forms, to the development of a sexual predators behavior are 1) arousal, 2) emotional congruence, 3) blockage, and 4) disinhibition (Finkelhor, 1984). Emotional congruence relates to the pedophile’s emotional need to relate to children. It is usually expressed in terms of pedophiles having difficulty relating to other adults. Sometimes children meet pedophile emotional needs, which other adults cannot. A feeling of lacking dominance can be

attempted to be fulfilled with children who are inherently less dominant. Similarly, psychological and social immaturity, low self-esteem, sexual abuse in childhood, narcissism, self-centeredness, and social inadequacy can lead to similar manifestations of emotional need. This can result in blockage – where bad experiences with age appropriate adults, sexual dysfunction, limited social skills, marital disturbance, or social and religious pressures limit age-appropriate sexual opportunities. These experiences can include: difficulty in relating to adults of the opposite sex; deficient social skills; anxiety over sexual matters; unresolved oedipal dynamics; disturbances in adult sexual relationships; and repressive norms about sexual behavior. Disinhibition reflects the abusers lack of control through impulse control deficits, psychosis, alcohol, drugs, stress, or nonexistent family rules-coupled with sexual arousal conditioning. This disinhibition can be caused by impulsive disorders senility, mental retardation, alcohol or drug abuse, situation stress, and tolerance of incest within the culture. Lastly, arousal of an adult by child has frequently been cultural or familial conditioning to sexual activity with children: such as corporal punishment and/or sexual child abuse. Approximately 50% of all sex offenders were victims of sexual assault (Smith and Israel, 1987; Johnson, 1989; Gladwell, 2007), 70% of child sex offenders have between 1 and 9 victims; at least 20% have 10 to 40 victims, and serial child sex offenders may have as many as 400 victims in a lifetime (Elliot et al. 1995; Elliot, 2009). With today's electronic world, even with cultural intolerance, pedophiles can find support in large virtual communities validating their lifestyles. The objective of this research is to identify markers for these attributes in the writings and evaluate the potential for sexual crime. Data from sources such as chat logs from existing sexual offenders, train a classifier, and then use it to classify logs from current online activity in social networks.

4. CONCLUSIONS AND IMPLICATIONS

Discussion of findings: It is no longer sufficient to use a single data source or a single analytic technique while analyzing data. In addition, there is a need to analyze unstructured text data since it can provide valuable clues on criminal behavior and intentions. The three methods discussed are mutually complementary and address the needs for law enforcement in fighting crime: 1) open source data collection, 2) natural language processing, 3) identifying correlations between disparate data sources. The data collection robots automate the process of collecting online data making it efficient. Linguistic analysis can be used for behavior analysis and integration tools can be used for tagging and correlating data.

Implications for policy and practice: Law enforcement is saddled with a growing backlog of cases of online crime and traditional crime that rely on online evidence. Crime labs around the country have been increasing capacity to handle this growing backlog. In addition to increasing capacity, efficiency of analysis also needs to increase. The crime scene today is often not a physical location but the Internet. Being able to rapidly collect data from online sources will make it feasible for investigators to pursue more crimes. Being able to gather corroborating evidence from chats, instant messaging, and web sites can improve the rate of conviction. A suite of tools for data analysis will be employed for

a host of problems including, 1) psychological profiling of sexual predators and determining precursors to crime 2) identifying hacker motivations for committing crime.

Implications for further research: The current work builds the basic infrastructure for analyzing crime using multiple data sources including unstructured text data. This infrastructure can be used for facilitating other problems, i.e. sexual predators, hackers, and white collar criminals. Each of these problems requires an understanding of psychological behavior markers as well as data collection and analysis. Further research will involve developing algorithms for data correlations and developing best practices for law enforcement to use. We also need to add other linguistic characteristics beyond content analysis while examining text data. Some of the techniques developed above will be useful for identifying potential for recidivism in certain crimes.

Dissemination: The work will be disseminated through publications that will be written on specific problems that are addressed using this infrastructure. The New York State Police will work with UAlbany on specific analysis using these tools, including, sexual predator analysis, and white collar crime.

REFERENCES

Altman, P. L., *Blood and Other Body Fluids*. Federation of American Societies for Experimental Biology: Washington, D.C., 1961.

Anderson, A., M. Bader, E. Bard, E. Boyle, and G. Doherty. 1991. The HCRC Map Task Corpus. Language and Speech, 34:351–366.

Aubrey, K. L.; Thomas, G. J., Jr., Raman spectroscopy of filamentous bacteriophage Ff (fd, M13, f1) incorporating specifically-deuterated alanine and tryptophan side chains. Assignments and structural interpretation. *Biophys. J.* 1991, 60, (6), 1337-49.

Bartick, E., Forensic analysis by Raman spectroscopy: An emerging technique. In *16th Meeting of the International Association of Forensic Sciences*, Medimond, Inc.: Montpellier, France, 2002.

Becker, W. S. and Dale, W. M. A Case of Forensic Scientist Turnover, *Forensic Science Communications* [Online]. (2004) 6(3). Available: http://www.fbi.gov/hq/lab/fsc/backissu/july2004/research/2004_03_research04.htm.

Becker, W.S., and Dale, W.M. (2007). Critical human resource issues: Scientists under pressure. *Forensic Science Communications, 9, 2.*

Bertoluzza, A.; Fagnano, C.; Finelli, P.; Morelli, M. A.; Simoni, R.; Tosi, R., Raman and infrared spectra of spermidine and spermine and their hydrochlorides and phosphates as a basis for the study of the interactions between polyamines and nucleic acids. *J. Raman Spectrosc.* 1983, 14, (6), 385-394.

Best, M. A., Statistical presentation of forensic data. In *Molecular Forensics*, Rapley, R.; Whitehouse, D., Eds. John Wiley & Sons, Ltd.: West Sussex, England, 2007; pp 185-195.

Boba, Rachel (2005). *Crime Analysis and Crime Mapping*. Sage Publications.

Brettell, T. A.; Butler, J. M.; Almirall, J. R., Forensic science. *Anal Chem* 2007, 79, (12), 4365-84.

Budowle, B.; van Daal, A., Extracting evidence from forensic DNA analyses: future molecular biology directions. *Biotechniques* 2009, 46, (5), 339-50.

Callan, J. P., W. B. Croft, and S. M. Harding. 1992. *The INQUERY Retrieval System*, in Proceedings of the 3rd Inter- national Conference on Database and Expert Systems.

Carter, E. A.; Edwards, H. G. M., Biological applications of Raman spectoscopy. In *Infrared and Raman Spectroscopy of Biological Materials*, Gremlich, H. U.; Yan, B., Eds. Marcel Dekker, Inc.: New York, 2001.

Cascio, W. F. *Costing Human Resources: The Financial Impact of Behavior in Organizations*. 4th ed., South-Western College, Cincinnati, Ohio, 2000.

Colomban, P.; Gouadec, G., Raman scattering theory and elements of Raman instrumentation. In *Raman Spectroscopy for Soft Matter Applications*, Amer, M. S., Ed. John Wiley & Sons, Inc.: Hoboken, NJ, 2009.

Coyle, T.; Anwar, N., A novel approach to condom lubricant analysis: In-situ analysis of swabs by FT-Raman spectroscopy and its effects on DNA analysis. *Science and Justice* 2008, doi:10.1016/j.scijus.2008.04.003, 1-11.

Dale, W. M., & Becker, W. S. (2003). *Strategy for Staffing Forensic Scientists*, *Journal of Forensic Sciences*, 48(2), pg 465-466.

Dale, W.M., and Becker, W.S. (2005). Managing intellectual capital. *Forensic Science Communications, 7,* 4.

de Aza, P. N.; Santos, C.; Pazo, A., Vibrational properties of calcium phosphate compounds: Raman spectrum of b-tricalcium phosphate. *Chem. Mater.* 1997, 9, 912-915.

De Gelder, J.; De Gussem, K.; Vandenabeele, P.; Moens, L., Reference database of Raman spectra of biological molecules. *J. Raman Spectrosc.* 2007, 38, 1133-1147.

Denison, S. J.; Lopes, E. M.; D'costa, L.; Newman, J. C., Positive prostate-specific antigen (PSA) results in semen-free samples. *Can. Soc. Forensic Sci. J.* 2004, 37, (4), 197-206.

Eapen, A.; Joe, I. H., Vibrational and SERS spectra of spermine phosphate hexahydrate. *Spectrosc. Lett.* 1997, 30, (4), 751-770.

Eckenrode, A.; Bartick, E. G.; Harvey, S.; Vucelick, M. E.; Wright, B. W.; Huff, R. A., Portable Raman spectroscopy systems for field analysis. *Forensic Sci. Comm.* 2001, 3.

Edsall, J. T., Raman specrta of amino acids and related compounds VI: Sarcosine, ethanolamine, choline, betaine, and betaine derivatives *J. Am. Chem. Soc.* 1943, 65, (9), 1767-1770.

Elliot, I.A., Beech A.R., Mandeville-Norden, R., and Hayes, H., "Psychological Profiles of Internet Sexual Offenders", Sexual Abuse: A Journal of Research and Treatment, Vol. 21, No. 1, 76-92 (2009).

Elliot, M., Browne, K., & Kilcoyne, J. (1995). Child sexual abuse prevention: What offenders tell us. Child Abuse & Neglect, 5, 579-594.

Fiedler, A.; Rehdorf, J.; Hilbers, F.; Johrdan, L.; Stribl, C.; Benecke, M., Detection of Semen (human and boar) and saliva on fabrics by a very high powered UV-/VIS-light source. *Open Forensic Sci. J.* 2008, 1, 12-15.

Finkelhor, David. 1984 Child sexual abuse : New Theory and Research / David Finkelhor Free Press, New York.

Fisher, B. A. J., (2003), Field needs adequate funding, National Forensic Science Commission, *Forensic Focus: Advancing the Forensic Science and Criminal Justice Communities* [Online]. Available: http://forensicfocusmag.com/articles/3b1persp1.html.

Gladwell, M. (2007-11-12). "Dangerous Minds: Criminal profiling made easy."". The New Yorker (New York). http://www.newyorker.com/reporting/2007/11/12/071112fa_fact_gladwell.

Gonmori, K.; Yoshioka, N.; Chattopadhyay, P. K., Identification of human semen stains by thin-layer chromatography. *Act. Crim. Japon.* 1994, 60, (4), 149-154.

Grasselli, J., *Chemical applications of Raman spectroscopy*. John Wiley & Sons: New York, 1981.

Greenfield, A.; Sloan, M. A., Identification of Biological Fluids and Stains. In *Forensic science: an introduction to scientific and investigative techniques*, James, S. H.; Nordby, J. J., Eds. CRC Press: Boca Raton, 2003; pp 203-220.

Guicheteau, J.; Argue, L.; Hyre, A.; Jacobson, M.; Christensen, S. D., Raman and surface-enhanced Raman spectroscopy of amino acids and nucleotide bases for target biological vibrational mode identification. *Proc. of SPIE* 2006, 6281, (62810O), 1-11.

Hardy, H., Shimizu, N., Tomek Strzalkowski, Ting Liu, Bowden Wise and Xinyang Zhang. 2002. Cross-document summarization by concept classification. In *Proceedings of ACM SIGIR '02 Conference*, pages 121-128, Tampere, Finland.

Healy, D. A.; Hayes, C. J.; Leonard, P.; McKenna, L.; O'Kennedy, R., Biosensor developments: application to prostate-specific antigen detection. *Trends in Biotech.* 2007, 25, (3).

Hochmeister, M. N.; Budowle, B.; Rudin, O.; Gehrig, C.; Borer, U.; Thali, M.; Dirnhofer, R., Evaluation of prostate-specific antigen (PSA) membrane test assays for the forensic identification of seminal fluid. *J. Forensic Sci.* 1999, 44, 1057-1060.

Hodges, C. M., Akhavan, J., The use of Fourier transform Raman spectroscopy in the forensic identification of illicit drugs and explosives. *Molecular Spectroscopy* 1990, 46, (2), 303-307.

Hickman, M. J. and Peterson, J.L. 50 largest Crime Labs, Bureau of Justice Statistics, U.S. Department of Justice, 2002, NCJ: 20598.

Iitaka, Y.; Huse, Y., The Crystal Structure of Spermine Phosphate Hexahydrate. *Acta Crystallogr* 1965, 18, 110-21.

Ivanov, A. I.; Zhbankov, R. G.; Korolenko, E. A.; Korolik, E. V.; Meleshchenko, L. A.; Marchewka, M.; Ratajczak, H., Infrared and Raman spectroscopic studies of the structure of hman serum albumin under various ligand loads. *J. Appl. Spectrosc.* 1994, 60, (5-6), 305-309.

Johnson, C. R.; Ludwig, M.; Asher, S. A., Ultraviolet Resonance Raman Characterization of Photochemical Transients of Phenol, Tyrosine, and Tryptophan. *J. Am. Chem. Soc.* 1986, 108, (5), 905-912.

Johnson, T. C. (1989). Female child perpetrators: Children who molest other children. *Child Abuse and Neglect,* 13(4), 571-585

Jones Jr., E. L., The identification of semen and other body fluids. In *Forensic science handbook*, Saferstein, R., Ed. Prentice Hall: Upper Saddle River, NJ, 2005; Vol. 2, pp 329-382.

Jurafsky, D., R. Bates, N. Coccaro, R. Martin, M. Meteer, K. Ries, E. Shriberg, A. Stolcke, P. Taylor, and C. Van Ess-Dykema. 1998. Switchboard Discourse Language Modeling Project Final Report. Research Note 30, Center for Language and Speech Processing, Johns Hopkins University, Baltimore.

Koyama, Y.; Toda, S.; Kyogoku, Y., Raman spectra and conformation of the glycerophosphorylcholine headgroup. *Chem Phys Lipids* 1977, 19, (1), 74-92.

Li, R., *Forensic Biology*. CRC Press: Boca Raton, 2008.

Liang, J.; Cheng, Y.; Han, H., Study on the interaction of bovine serum albumin and CdTe quantum dots with spectroscopic techniques. *J. Mol. Struct.* 2008, 892, 116-120.

Liu, D. Y.; Clarke, G. N.; Baker, H. W., Tyrosine phosphorylation on capacitated human sperm tail detected by immunofluorescence correlates strongly with sperm-zona pellucida (ZP) binding but not with the ZP-induced acrosome reaction. *Hum Reprod* 2006, 21, (4), 1002-8.

Long, J., (2001), Employment outlook (2002), *Chemical and Engineering News* 79(46): pg 38-39.

Ludwig, M.; Asher, S. A., Ultraviolet resonance Raman excitation profiles of tyrosine: Dependence of Raman cross sections on excited-state intermediates. *J. Am. Chem. Soc.* 1988, 110, (4), 1005-1011.

Macleod, N. A.; Matousek, P., Emerging non-invasive Raman methods in process control and forensic applications. *Pharm Res* 2008, 25, (10), 2205-15.

Maher, J.; Vintiner, S.; Elliot, D.; Melia, L., Evaluation of the BioSign PSA membrane test for the identification of semen stains in forensic casework. *New Zeal. Med. J.* 2002, 115, (1147), 48-49.

Mann, T., Biochemistry of semen. In *Handbook of Physiology, Section 7; Endocrinology*, American Physiological Society: Washington, 1975; Vol. 5, pp 461-471.

Mazzella, W. D.; Buzzini, P., Raman spectroscopy of blue gel pen inks. *Forensic Sci Int* 2005, 152, (2-3), 241-7.

Nafie, L. A., Theory of Raman scattering. In *Handbook of Raman Spectroscopy*, Lewis, I. R.; Edwards, H. G. M., Eds. Marcel Dekker, Inc.: New York, 2001.

National Research Council of the National Academies, (2009), *Strengthening Forensic Science in the United States: A Path Forward*, The National Academies Press, Washington, D.C.: Chapter 8.

Nelson, D. G.; Santucci, K. A., An alternate light source to detect semen. *Acad. Emerg. Med.* 2002, 9, (10).

New York State Divison of Criminal Justice Services. http://www.criminaljustice.state.ny.us/pio/annualreport/crimeupdate2008.pdf .

Osborne, Deborah and Susan Wernicke (2003) *Introduction to Crime Analysis: Basic Resources for Criminal Justice Practice*. Haworth Press.

Owen, D. H.; Katz, D. F., A review of the physical and chemical properties of human semen and the formulation of a semen simulant. *J Androl* 2005, 26, (4), 459-69.

Rodger, C., Broughton, D., The in-situ analysis of lipsticks by surface enhanced resonance Raman scattering. *Analyst* 1998, 1823-1826.

Rosenheim, O., The Isolation of Spermine Phosphate from Semen and Testis. *Biochem J* 1924, 18, (6), 1253-1262 1.

Sancho, S.; Rodriguez-Gil, J. E.; Pinart, E.; Briz, M.; Garcia-Gil, N.; Badia, E.; Bassols, J.; Pruneda, A.; Bussalleu, E.; Yeste, M.; Casas, I.; Palomo, M. J.; Ramio, L.; Bonet, S., Effects of exposing boars to different artificial light regimens on semen plasma markers and "in vivo" fertilizing capacity. *Theriogenology* 2006, 65, (2), 317-31.

Santucci, K. A.; Nelson, D. G.; McQuillen, K. K.; Duffy, S. J.; Linakis, J. G., Wood's lamp utility in the identification of semen. *Pediatrics* 1999, 104, 1342-1344.

Sato, S.; Moriya, F.; Yamamoto, Y.; Ishizu, H. In *Assay of polyamine in semen and seminal stains by HPLC with fluorometric determination of o-phthalaldehyde-deivatives*, Current topics in forensic science: proceedings of the 14th meeting of the International Association of Forensic Sciences, Tokyo, Japan, 1996; Shunderson Communications: Tokyo, Japan, 1996; pp 40-43.

Sato, I.; Sagi, M.; Ishiwari, A.; Nishijima, H.; Ito, E.; Mukai, T., Use of the "SMITEST" PSA card to identify the presence of prostate-specific antigen in semen and male urine. *Forensic Sci. Int.* 2002, 127, 71-74.

Shaikh, S, and Strzalkowski T., Sarah Taylor and Jonathan Smith (2009) Comparing an Integrated QA system performance - A Preliminary Model. Proceedings of PACLING Conference, Sapporo, Japan.

Shaler, R. C., Modern forensic biology. In *Forensic science handbook*, Saferstein, R., Ed. Prentice Hall: Upper Saddle River, NJ, 2002; Vol. 1, pp 529-546.

Shriberg, E., R. Dhillon, S. Bhagat, J. Ang, and H. Carvey. 2004. The ICSI Meeting Recorder Dialog Act (MRDA) Corpus. In Special Interest Group on Discourse and Dialogue (SIGdial), Boston, USA.

Siegel, J. A., *Encyclopedia of Forensic Sciences.* Academic Press: San Diego, 2000; Vol. 3, p 5.

Simon, L. M. J. (1997). Do criminal offenders specialize in crime types? *Applied and Preventive Psychology, 6,* 3553.

Small, S and Strzalkowski, T. 2009. HITIQA: High-Quality Intelligence through Interactive Question Answering. *Journal of Natural Language Engineering*, Vol. 15 (1), pp. 31—54. Cambridge.

Smith, H., and Israel, E. (1987). Sibling incest: A study of the dynamics of 25 cases. *Child Abuse and Neglect,* 11(1), 101-108.

Spiker, R. C., Jr.; Levin, I. W., Raman spectra and vibrational assignments for dipalmitoyl phosphatidylcholine and structurally related molecules. *Biochim Biophys Acta* 1975, 388, (3), 361-73.

Stewart, S.; Fredericks, P. M., Surface-enhanced Raman spectroscopy of amino acids absorbed on an electrically prepared silver surface. *Spectrochim. Acta A* 1999, 55, 1641-1660.

Strzalkowski, T 2009. ECO: Effective Communication through Internet Social Media. Internal report. ILS.

Suzuki, E. M.; Carrabba, M., In situ identification and analysis of automotive paint pigments using line segment excitation Raman spectroscopy: I. Inorganic topcoat pigments. *J Forensic Sci* 2001, 46, (5), 1053-69.

Suzuki, O.; Oya, M.; Katsumata, Y.; Matsumoto, T.; Yada, S., A new enzymatic method for the demonstration of spermine in human seminal stains. *J. Forensic Sci.* 1980, 25, 99.

Thomas, J.; Buzzini, P.; Massonnet, G.; Reedy, B.; Roux, C., Raman spectroscopy and the forensic analysis of black/grey and blue cotton fibres Part 1: investigation of the effects of varying laser wavelength. *Forensic Sci Int* 2005, 152, (2-3), 189-97.

Tsutsumi, A., Assay of seminal spermine by isotachophoresis and its application to forensic science. *Jpn. J. Legal Med.* 1987, 41, (4), 294-307.

Vandenberg, N.; van Oorshot, R. A. H., The use of Polilight in the detection of seminal fluid, saliva, and bloodstains and comparison with conventional chemical-based screening tests. *J. Forensic Sci.* 2006, 51, (1).

van Leeuwenhoek, A., Observationes D. Anthonii Leeuwenhoek, de natis e semine genitali animalculis. *Philos.Trans. R. Soc.* 1678, 12, 1040-1043.

Virkler, K.; Lednev, I. K., Raman spectroscopy offers great potential for the nondestructive confirmatory identification of body fluids. *Forensic Sci. Int.* 2008, 181, e1-e5.

Virkler, K.; Lednev, I. K., Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene. *Forensic Sci Int* 2009, 188, (1-3), 1-17.

Wakefield, A., Curve Fitting Toolbox User's Guide. In The Mathworks, Inc.: Natick, MA, 2008.

Watson, N., The analysis of body fluids. In *Crime scene to court; the essentials of forensic science,* Royal Society of Chemist.: Cambridge, UK, 2004; pp 377-413.

Webb, Nick. 2009. Evaluating Human-Machine Conversation for Appropriateness. Online at http://ils.albany.edu/tech_reports.html.

Weedn and Hicks, The Unrealized Potential of DNA Testing,, USDOJ, June 1998.

Yan, F.; Vo-Dinh, T., Surface-enhanced Raman scattering detection of chemical and biological agents using a portable Raman integrated tunable sensor. *Sens. Actuators, B* 2007, 61-66.

Yokota, M.; Mitani, T.; Tsujita, H.; Kobayashi, T.; Higuchi, T.; Akane, A.; Nasu, M., Evaluation of prostate-specific antigen (PSA) membrane test for forensic examination of semen. *Legal Medicine* 2001, 3, 171-176.

APPENDICES

Appendix 1: 3130 XL and Genemapper ID General Information

NORTHEAST REGIONAL FORENSIC INSTITUTE
UNIVERSITY AT ALBANY
1400 WASHINGTON AVENUE
ALBANY, NEW YORK 12222
WWW.ALBANY.EDU/NERFI

CAPILLARY ELECTROPHORESIS & DATA ANALYSIS USING APPLIED BIOSYSTEMS' 3130XL GENETIC ANALYZER & GENEMAPPER® ID

Instructors: Lucy A. Davis & Jamie L. Belrose

COURSE DESCRIPTION

This five-day course is designed to provide the theoretical and practical background necessary to perform capillary electrophoresis and data analysis. Today's advanced technology has led to an exponential number of cases being submitted to the crime lab for DNA testing. For this technique to be successful, it is imperative that the biological evidence is processed and analyzed effectively.

The course enrollment is limited to six (6) individuals per session – due to the hands-on element.

There are four sessions scheduled:
- December 7 – 11, 2009
- January 11 – 15, 2010
- January 25 – 29, 2010
- February 8 – 12, 2010

Sessions will be held from 8:00am till 4:00pm (with a one-hour lunch) each day at the NERFI labs. We are also interested in traveling to host labs and providing this workshop. If your lab is interested in this option, please contact us for additional information. These workshops are being held via NIJ funding and you may attend free-of-charge. The cost to attend the course is covered, as well as travel and per diem (GSA: Albany, NY rate $160/day).

This course will cover a variety of topics (lecture & hands-on) including:

I. Electrophoretic Theory: Ohm's Law, Cathode, Anode and Molecular Separation

II. Genetic Analyzers:
- **A.** Sample Preparation with Formamide
- **B.** Electrokinetic Injection
- **C.** Sample Stacking
- **D.** The Capillaries used in Forensic STR Analysis
- **E.** Electric Osmotic Flow
- **F.** Performance Optimized Polymer
- **G.** Constant Charge-to-Mass-Ratio
- **H.** Stokes Shift
- **I.** Optics & Fluorescent Detection
- **J.** Charged-Couple Device Camera
- **K.** Spectral Calibration, How & Why
- **L.** Spatial Calibration, How & Why
- **M.** Instrument Components & Function
- **N.** Data Collection Software: instrument protocols, results groups, plate records, plate map editor, running the instrument, preparing samples, and maintenance calibrations (daily, weekly, monthly, annually).

III. Data Analysis

A. GeneMapper Theory:

1. Internal Lane Standard
2. Local Southern Method
3. Analysis Parameters
4. Baseline Smoothing
5. Baseline Window
6. Polynomial Degree
7. Peak Window Size and Slope Threshold
8. Advanced / Classic Algorithms
9. Marker Specific Stutter Ratios
10. Defining and Analysis Range

B. Use of GeneMapper"

1. Creating a New Project
2. Navigating Through the Software
3. Analyzing / Evaluating Data: Size Match Editor, Raw Data Review, 250bp peak Migration, Ladders (Panels & Bins), Negative Controls, Positive Controls, Unknown Samples, Artifacts and OL Calculations.

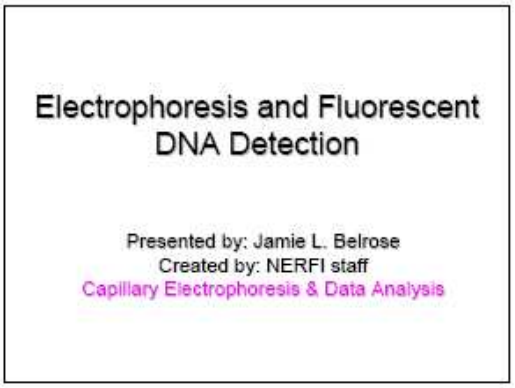
C. Demonstration of GMID-IDX

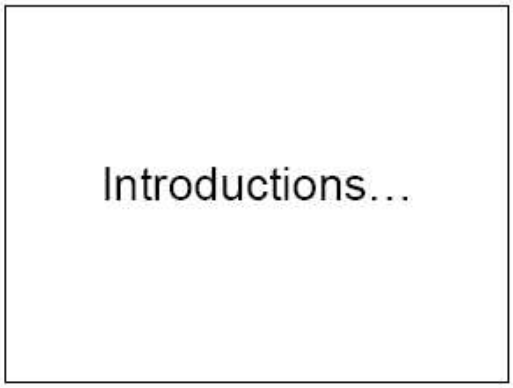
The course will conclude with a multiple-choice exam and issuance of course completion certificates.

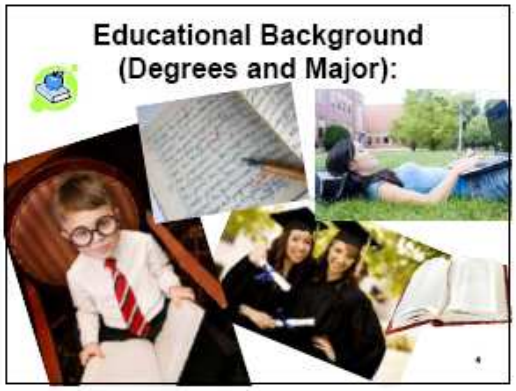
For more information, please contact:

Ms. Jamie L. Belrose, M.S.
Forensic Training Lecturer
Manager of Daily Operations
Northeast Regional Forensic Institute (NERFI)
University at Albany; Biology Rm 228
1400 Washington Avenue
Albany, New York 12222
ph: 518.442.4355
fax: 518.442.5401
jamie.belrose@gmail.com
www.albany.edu/NERFI

Appendix 2: Electrophoresis Presentation







Jamie L. Belrose

Educational Background

AS: Adirondack Community College – Math/Science

BS: University of New Haven – Forensic Science
~ Internship: New York State Police

MS: University at Albany – Forensic Molecular Bio
~ Internship: Federal Bureau of Investigation

9

Professional Background

Taconic Biotechnology: Assay Design Specialist

OCME NYC: Criminalist II

NERFI: Forensic Training Lecturer & Manager of Daily Operations

10

11

12

13

14

15

PIA

Short for –

Pain In the A……

16

Interesting Fact about me

I collect magnets…

I was named after Wonder Woman

My name in French translates to:

J'aime Belle Rose

" I love the beautiful rose."

17

What is **Electrophoresis?**

- A method of separating large molecules (such as DNA fragments or Proteins) from a mixture of similar molecules.
- An electric current is passed through a medium containing the sample, and each molecule travels through the medium at a different rate, depending on its electrical charge and size.
- Agarose and acrylamide gels are the media commonly used for electrophoresis of proteins and nucleic acids

Gel Electrophoresis

- The pH and other buffer conditions are arranged so that the molecules being separated carry a negative charge. This allows them to move within an electric field.
- As they move through the gel, the larger molecules will be held up as they try to pass through the pores of the gel, while the smaller molecules will be impeded less and move faster.
- This results in a separation by size, with the larger molecules closer to the starting point and the smaller molecules further away.

Electrophoresis of DNA

- The Phosphate groups on the backbone of the DNA molecule readily give up their H^+ ions, therefore nucleic acids are negatively charged in most buffered systems.
- DNA molecules will migrate away from the negative electrode (cathode), and migrate towards the positive electrode (anode).
- The higher the voltage, the greater the force felt by the DNA molecule, and thus the faster they will migrate in an electric field.

Electrophoresis of DNA

- The DNA molecule is made up of nucleotide subunits each consisting of a sugar, Phosphate and nitrogenous base.
- Each of these nucleotide subunits carries a negative charge (because of Phosphate).
- As the DNA molecule incorporates additional nucleotide subunits, it acquires more negative charges.

Electrophoresis of DNA

- Constant Charge to Mass Ratio:
  - During electrophoresis the DNA fragments are separated based solely on size. This is because each of the molecules, regardless of size, experiences the same amount of "pull" from the electric current.
  - The structure of DNA is comprised of nucleotide units. Each unit contains: 1 phosphate, 1 sugar, and 1 nitrogenous base. For every unit there is one negative charge from the phosphate.
  - As the DNA molecule grows there are more units added, each with a negative charge.
  - This means the charge is proportional for each DNA molecule. More units = More negative charges!

Topics to be Discussed

- Agarose Gel Electrophoresis
- Acrylamide Gel Electrophoresis
  - Native conditions
  - Denaturing conditions
- Capillary Electrophoresis
  - 3130xl

Gel Matrices Used for Electrophoresis of DNA

- **Agarose** Gels have fairly large pore sizes and are used for separating larger DNA molecules (ie RFLP fragments).
- **Polyacrylamide** Gels are used to obtain high resolution separations for smaller DNA molecules (STR and DNA sequence analysis).

Agarose Gel Supplies

- An electrophoresis chamber and power supply
- Gel casting trays, which are available in a variety of sizes and composed of UV-transparent plastic.
- Sample combs, around which molten agarose is poured to form sample wells in the gel.
- Electrophoresis buffer, usually Tris-acetate-EDTA (TAE) or Tris-borate-EDTA (TBE).

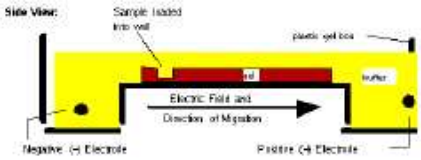
Electrophoresis Buffer

- Several different buffers have been recommended for electrophoresis of DNA.
- The most commonly used for duplex DNA are TAE (Tris-acetate-EDTA) and TBE (Tris-borate-EDTA). DNA fragments will migrate at somewhat different rates in these two buffers due to differences in ionic strength. Buffers establish a pH and provide ions to support conductivity.
  - TAE has buffering capacity and dsDNA runs faster in TAE vs. TBE

Agarose Gel Supplies

- Loading buffer (for loading samples) contain something dense (i.e. glycerol) to allow the sample to sink to the bottom of the sample wells, and one or two tracking dyes, which migrate in the gel (ahead of the DNA) and allow monitoring or how far the electrophoresis has proceeded.
- A fluorescent dye used for staining nucleic acids, such as Ethidium bromide or Sybr Green.
- Transilluminator or Fluorescent Gel Scanner for photo-documentation

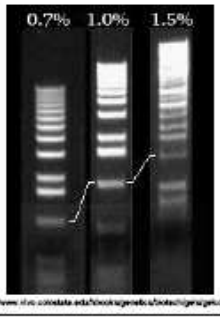
Agarose Gel Electrophoresis System



Agarose Concentration

- By using gels with different concentrations of agarose, one can resolve different sizes of DNA fragments. Higher concentrations of agarose facilitate separation of small DNAs, while low agarose concentrations allow resolution of larger DNAs.
- High concentration = small pores = good for small fragments.
- Low concentration = large pores = good for large fragments.

Agarose Concentration



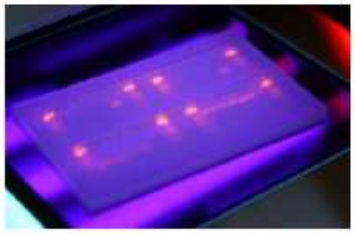
Ethidium Bromide

- This compound contains a planar group that **intercalates** between the stacked bases of double-stranded DNA.
- UV radiation at 254 nm is absorbed by the DNA and transmitted to the bound dye.
- The energy is then emitted at 590 nm in the **red-orange** region of the spectrum.

Agarose Gels – Ethidium Bromide

- Ethidium Bromide intercalates between the stacked bases of the DNA strands.
- Ethidium Bromide can be co-polymerized with the agarose or stained post-electrophoresis by incubating the gel in an Ethidium bromide solution.

After completion of the run, visualize the DNA under UV light.



Acrylamide Gel Electrophoresis

Polyacrylamide Gel Electrophoresis

- Polyacrylamide gels are poured between two glass plates held apart by spacers and sealed with tape.
- The length of the gel can vary between 10 cm and 100 cm depending on the separation required. Longer gels = more separation (sequencing)
- They are always run vertically with TBE as a the buffer.

Polyacrylamide Gel Electrophoresis

- Polyacrylamide gels have enough resolving power to separate fragments differing by one base pair in size, but their range is ~ 5 to 1000 bp.
- They are much more difficult to handle than agarose gels, fragile.

Non-denaturing gels (Native)

- Run at low voltages - 8V/cm - and 1X TBE to prevent denaturation of small fragments of DNA by the heat generated in the gel during electrophoresis.
- The rate of migration is approximately inversely proportional to log of their size. However, the base sequence composition can alter the electrophoretic mobility of DNAs such that two DNAs of the same size may show up to a 10% difference in electrophoretic mobility.

Double-stranded DNA

Denaturing Gels

- These gels are polymerized with a denaturant that suppresses base pairing between nitrogenous bases - this is usually Urea but can be Formamide.
- Denatured DNA migrates through the gel at a rate which is *almost* completely independent of its composition or sequence.

Single-stranded DNA

Capillary Electrophoresis

Heat Dissipation

- In conventional slab gel electrophoresis the heat associated with the generation of electric current during separation can cause problems of peak / band broadening.
- Heat causes the formation of convection currents within the gel.
- Heat generation restricts the operating voltages that can be used in slab gel electrophoresis which produces longer analysis times.

Heat Dissipation

- Performing electrophoresis in a capillary allows the heat to be effectively dissipated through the capillary walls which reduces any convection related peak broadening.
- This improved heat dissipation means that higher operating voltages can be used in CE which can produce significantly faster analysis times.

Capillary Electrophoresis

- Performing electrophoretic separations in capillaries was shown to offer the possibility of automated analytical equipment and fast analysis times.
- The capillary was inserted through the optical center of a detector which allowed real time capillary detection. You can view the data as it is detected by the instrument.

Capillary Electrophoresis

- Operation of a CE system involves application of a high voltage (typically 10-30kV) across a narrow bore (25-100μm) capillary.
- The capillary is filled with electrolytic solution which conducts current through the inside of the capillary.
- The ends of the capillary are dipped into reservoirs filled with electrolyte.

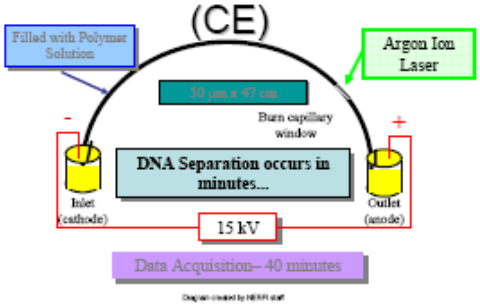
Capillary Electrophoresis

- Electrodes made of an inert material, such as Platinum, are also inserted into the electrolyte reservoirs to complete the electrical circuit.
- A small volume of sample is injected into one end of the capillary.
- The capillary passes through a detector at the opposite end.

Capillary Electrophoresis

- Application of a voltage causes movement of sample ions towards their appropriate electrode passing through the detector.
- The plot of detector response (RFU) versus time (data point) is generated, which is termed an electropherogram.

Capillary Electrophoresis (CE)



The Data

Chemistry Involved

- Injection
  - electrokinetic injection process
  - importance of sample preparation (formamide)
- Separation
  - capillary
  - POP-4 polymer
  - buffer
- Detection
  - fluorescent dyes with excitation and emission traits
  - virtual filters (hardware/software issues)

Injection

- **Electrokinetic injection:**
  Definition: A mechanism that forces a small amount of sample into the capillary as a result of an applied voltage.

  Process: A small amount of sample is forced into the capillary when the capillary and the electrode are submerged into the sample vial and voltage is applied. Since the sample is ionized, ions from the sample migrate into the capillary.

**Increasing the amount of time that the capillary and electrode are in contact with the sample should increase the amount of sample (DNA) drawn into the capillary*

Injection

- The efficiency of electrokinetic injection is dependent on the amount of ions present in the sample.
- Sample stacking describes the preferential sample injection between the DNA molecules and the ions.
- As the ion content of a sample increases, the sample conductivity also increases (ionic strength increases). The ions then compete with the DNA in the sample to be injected into the capillary. Thus, fewer DNA molecules are injected.

The importance of Formamide

- Formamide and single-stranded DNA
  - Samples to be analyzed are first diluted in a large amount (proportionally 1:25) of high quality deionized formamide (HI-DI). Formamide serves two purposes:
    - 1. By diluting the PCR product in a large volume of solution the salts associated with PCR are diluted. The salt dilution aids in the sample injection process.
    - 2. The formamide forms hydrogen bonds with the nucleotide bases preventing complimentary hybridization; forcing the DNA into a single-stranded formation. This single-stranded form is necessary for high resolution. When the DNA is single-stranded it is more flexible and is better able to interact with the sieving medium (POP-4), resulting in separation of closely sized molecules and thus increased resolution.
      - In addition, secondary structure is eliminated.

Injection

- The amount of DNA injected into the capillary is inversely proportional to the ionic strength of the sample.

Injection

- Electrokinetic Injection:
  - Since the introduction of DNA into the capillary is based on the application of an electric voltage, the salt ions present in a sample will also be affected. If the ionic strength of a sample is high, the ions will be competing with DNA molecules for injection into the capillary.
  - The ionic strength of the sample and the amount of DNA injected are inversely proportional to one another. Low ionic strength = Lots of DNA on cap!
  - Ionic strength is also used to "level the playing field".

Injection

Sample Stacking:

- When DNA is injected from a solution that is lower in ionic strength, than the buffer inside the capillary, sample stacking results.
- When the injection voltage is applied the DNA within the sample solution rushes into the capillary. As the DNA enters the capillary it encounters a buffer that is much higher in ionic strength, bringing the racing DNA molecules to a halt. Regardless of size/length, the DNA molecules stack on top of one another forming a sharp band.
- The large DNA molecules will be starting from the same point as the small. This "evens the race through the capillary".
- The sample is focused!

Injections

A few nanoliters of DNA enters the capillary:

- The capillaries used in STR analysis are made of fused silica (glass). They are 47cm in total length and 36cm from the inlet to the detection window. The inner diameter is 50µm.
- Some available lengths of capillary are: 47cm and 50cm. The longer the capillary, the longer the separation time, the better the resolution. The longer capillaries are traditionally used in sequencing where 1bp separation is required.

Injection

- The outside of the capillary is coated with a layer of plastic polyimide. This adds strength to the otherwise frail capillary. This coating however interferes with the fluorescent detection and must be removed. A small window is burned away to allow the laser to penetrate the capillary.
- The inside of capillary is uncoated from the manufacturer. It does become coated however, after it has been filled with POP-4.

POP 4

Performance Optimized Polymer – 4:

- The "4" refers to the 4% concentration of linear uncross-linked poly(dimethylacrylamide) or PDMA. The PDMA molecules coat the inner wall of capillary masking the charge of the silica. Allowing for non-retarded migration of the DNA.
- POP-4 also contains 8M UREA and 5% 2-pyrrolidinone to maintain the single-stranded formation of DNA during electrophoresis.
- The buffer that helps to maintain pH is 100mM *N*-tris(hydroxymethyl)methyl-3-aminopropane-sulfonic acid (TAPS). This buffer also aids in sample stacking.

Separation

- Size based separation due to interaction of DNA molecules with entangled polymer strands
- Polymers are not cross-linked (as in slab gels)
- "Gel" is not attached to the capillary wall
- "Pumpable" -- can be replaced after each run
- Polymer length and concentration determine the separation characteristics

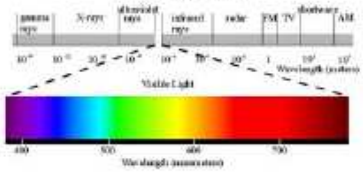
Separation Issues

- Run temperature -- 60 °C helps reduce secondary structure on DNA and improves precision
- Electrophoresis buffer -- Urea in Pop 4 helps keep DNA strands denatured
- Capillary wall coating -- dynamic (constantly changing) coating with polymer
- Polymer solution -- POP-4

Detection

- Fluorescent dyes
  - spectral emission overlap
  - relative levels on primers used to label PCR products
  - dye "blobs" (free dye)
- Virtual filters
  - hardware (CCD camera)
  - software (color matrix)

Detection



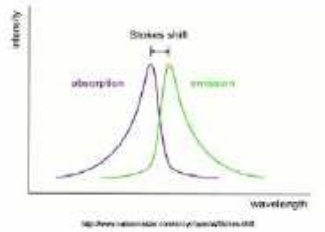
Do you remember the electromagnetic spectrum?

- Energy travels in waves
  - Shorter waves have more energy
- A photon is the elementary particle of electromagnetic radiation for a given wavelength (light).

Detection

- Molecules (fluorophores) exist at a "ground state" energy level.
- Energy (photons) is emitted from a light source (laser) to excite the electrons of a molecule.
  - Electrons must absorb or release energy when changing levels (absorb=↑;release=↓).
  - Electrons absorb energy and are elevated to a higher energy level. This excited state is only temporary, approx. 1-10 nanoseconds.
  - Interactions and conformational changes within the molecule cause electrons to transition to a lower energy level, causing the release of energy in the form of a photon (light).
  - The energy of the released photon is different than the energy that was absorbed from the laser - the difference between the two states is Stoke's shift.

Detection



Remember that EM waves with shorter wavelengths have more energy!

- Energy from the laser is of a shorter wavelength and higher energy, while the energy emitted from the photon is of a longer wavelength and less energy.

Detection

- Energy absorption and fluorescent emission properties are specific to individual dyes. Therefore one dye can be differentiated from another
- Utilizing dyes with different emission fluorescence allows for multiple DNA fragments to be run simultaneously.

Detection

Methods:

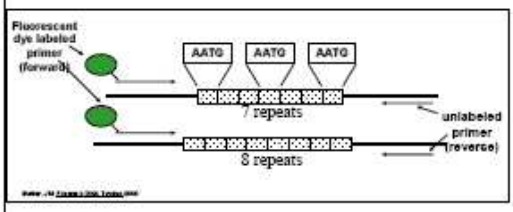
Intercalating fluorescent dyes – Binds to the DNA, only allows for analysis in a single color

Adding labeled deoxynucleotides (dNTPs) into the PCR product – Impacts mobility of the DNA fragments in an electric field; the ionic charge on the dye also changes the charge-to-mass ratio of the DNA-Dye conjugate

Detection

- Dye attached to the 5' end of the amplification primer – Only labels one strand of the amplified product for ease in analysis (second strand is not detected). Multiple amplicons can be labeled simultaneously.
- AB AmpF/STR kits: it is the forward primers that are labeled.

Primer - 5' end Dye Attachment



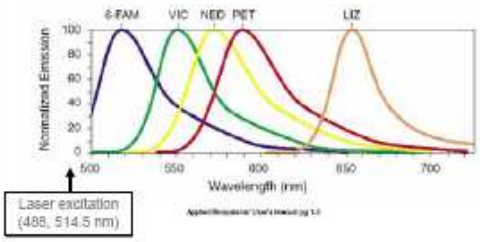
Fluorescent Dyes

- Each manufacturer of Fluorescent DNA kits/STR Genotyping kits employs a combination of dyes optimized for their multiplex system, the modes of detection, and the laser used.
- The manufacturer also optimizes the concentrations of the dyes to promote balanced electronic signals across each loci in a multiplex system.

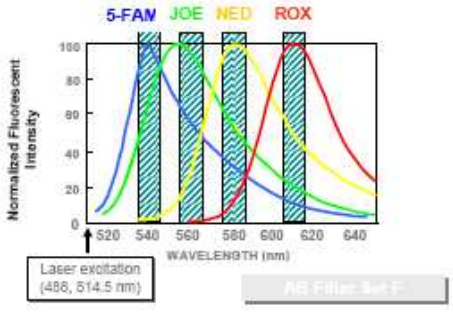
Detection

- Using a multiplex kit requires multi-component analysis to compensate for the spectral overlap.
- Multi-component analysis is mathematically performed by subtracting out the fluorescent contribution of overlapping dyes leaving only the dye of interest.

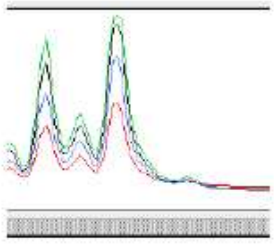
Detection Identifiler



Fluorescent Emission Spectra



Detection



A matrix, or spectral, is used to create a "virtual filter" to permit only the specific dye in a wavelength range to be visualized.

Detection

- Creating a matrix file involves running dye standards and capturing the fluorescence for each individual dye – no evidence STR data associated.
- The fluorescence is then tabulated and the amount of overlap between the dyes is calculated.
- This calculation is then applied to all subsequent runs of evidentiary samples.

Detection

POP4STRMODF

Reactions

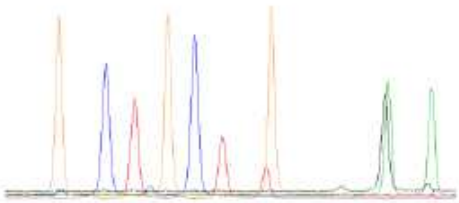
	B	G	Y	R
B	1.0000	0.8502	0.1390	0.0009
G	0.8300	1.0000	0.7622	0.0051
Y	0.6416	0.8524	1.0000	0.1102
R	0.4495	0.6484	0.7851	1.0000

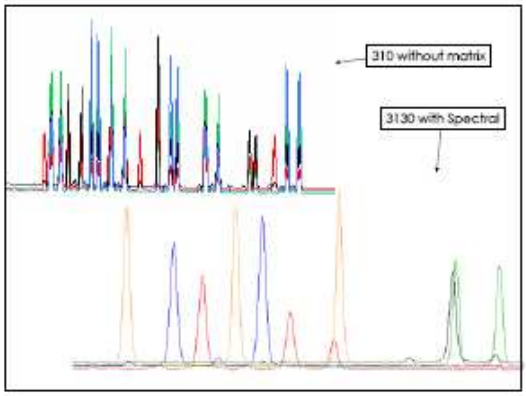
These values are used by the Analysis Software to separate the various dye colors from one another. The letters B, G, Y, and R represent the dye colors Blue, Green, Yellow, and Red, respectively.

Detection

- The Matrix is defined as a Spectral on the 3100 series instruments
- The spectral is applied to the data when it is collected by the camera, and is NOT a separate analysis. Thus, there is no true raw data associated with the 3100 series instruments.

Raw data from 3130 with Spectral applied





Detection

- Since the fluorescent dye is affected by its environment, varying conditions can influence the effectiveness of the matrix.
- Therefore matrix files should be generated as often as needed when run/detections conditions are altered.
- When are new matrix files needed?
  - When observable artifacts are seen during analysis, aka pull-up.

Appendix 3: 3130XL Presentation:

Applied Biosystems' 3130*xl* Genetic Analyzers: The Theory

Northeast Regional Forensic Institute
Capillary Electrophoresis & Data Analysis
Presented by: Jamie L. Belrose
Created by: NERFI Staff

Sample Prep with Formamide

Samples are first diluted in a large amount (proportionally 1:24) of High quality De-Ionized Formamide (Hi-Di).

Formamide serves two roles:

1. Diluting the PCR product in a large volume of solution also dilutes the salts associated with PCR (this is important in the sample injection process in that it decreases the ionic strength of the sample solution).
2. Keeps DNA single stranded
   - Forms hydrogen bonds with the nucleotide bases preventing complimentary hybridization, or re-annealing.
   - Why do we want single-stranded DNA?
     - Better resolution (dsDNA is bulkier)
     - Single-stranded DNA has increased flexibility allowing for easier movement through the POP-4, resulting in better separation of closely sized molecules and thus increased resolution.
   - In addition, secondary structure is eliminated.

Electrokinetic Injection

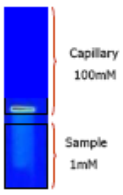
- The single-stranded (negatively charged) DNA is drawn into the capillary via the application of an electric voltage (15kV for 5s, *based on validation*).
- By using electricity, only a sample's charged species enter the capillary.
- Electrokinetic injection is highly dependent on the ionic strength of the sample solution and the buffer within the capillary.

Electrokinetic Injection

- Since the introduction of DNA into the capillary is based on the application of an electric voltage, the salt ions present in a sample will also be affected. If the ionic strength of a sample is high, the salt ions will be competing with DNA molecules for injection into the capillary.
- The ionic strength of the sample and the amount of DNA injected are inversely proportional to one another.
  Low ionic strength of sample = Lots of DNA on the capillary!
- Ionic strength is also used to "level the playing field".

Sample Stacking

- When DNA is injected from a sample solution that is lower in ionic strength, than the TAPS buffer inside the capillary, sample stacking results.
- When the injection voltage is applied the DNA within the sample solution rushes into the capillary. As the DNA enters the capillary it encounters a buffer that is much higher in ionic strength, bringing the racing DNA molecules to a halt. Regardless of size, the DNA molecules stack on top of one another forming a sharp focused band.
- This evens the race. The large DNA molecules will be starting from the same point as the small.
  The sample is focused!



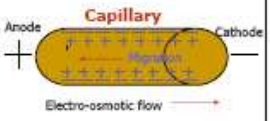
Capillary

- Above a pH of 6, silica (glass) is negatively charged.
- If the Inside of the capillary were to remain uncoated, the inside wall of the capillary would become covered in a layer of negative charges. The TAPS buffer within the capillary would create a layer of positive charges on top of this; creating what is known as the "double-layer".
- The negatively charged DNA molecules would be attracted to these positive charges.
- When the separation voltage is applied, the mobile cations (positive) would migrate toward the cathode (negative) and pull the (DNA) molecules in the same direction, *back to start*.

The dynamic internal coating of POP-4 is necessary for efficient migration. It also insures no cross-contamination between samples.

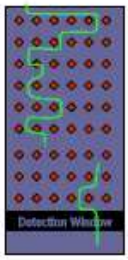
Electro-Osmotic Flow (EOF)

- In other words, the positive charges (cations from the buffer) that formed on top of the negative charges (anions from the glass) will migrate toward the negative electrode and pull the DNA back to where it started from.
  - Or, it retards migration.
- This phenomenon is known as electro-osmotic flow (EOF).
- However, the POP4 dynamically coats the inside of the capillary to prevent EOF.



Separation based on size

- When the second voltage is applied the DNA begins to migrate toward the positive anode. As it passes through the POP-4 it interacts with the entangled linear PDMA molecules.
- The shorter DNA fragments can more readily navigate through the sieving medium and reach the detection window sooner.
- The longer the fragment, the longer it takes to migrate the length of the capillary.



Constant Charge-to-Mass Ratio

- During electrophoresis the DNA fragments are separated based solely on size. This is because each of the molecules, regardless of size, experiences the same amount of "pull" from the electric current.
- The structure of DNA is comprised of nucleotide units. Each unit contains: 1 phosphate, 1 sugar, and 1 nucleotide base. For every unit there is one negative charge from the phosphate.
- As the DNA molecule grows there are more units added, each with a negative charge.
- This means the charge is proportional to molecule's size for each DNA fragment.
  - More nucleotide units = More negative charges!
  - More negative charges = More mass to pull...

Fluorescence and Excitation

- When the DNA fragments reach the detection window they encounter a laser. The fluorescent tag located at the 5' end of the forward primer is excited.
- L.A.S.E.R. = Light Amplification by Stimulated Emission of Radiation.
- The excitation energy from the laser causes an electron of the fluorophore molecule to ascend to a higher energy level. The electron then undergoes a conformational change and returns to the original ground state.
- As it is doing so, it emits a photon of light.

Stokes Shift

- The excitation energy from the laser is of a shorter wavelength and higher energy.
- The emission energy from the fluorescent tag is of a longer wavelength and lower energy. This enables the detection system to differentiate between the two wavelengths and only collect and record the emission.
- Each of the 5 dyes used has a different emission wavelength, allowing the instrument to determine which one it is "seeing".
- This phenomenon allows us to simultaneously use 5 dyes.

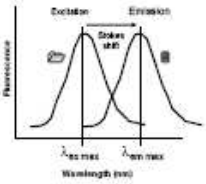


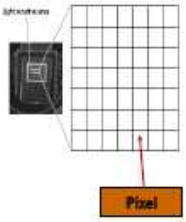
Figure 13.1, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Optics and the Detection System

- L.A.S.E.R:
  - The laser is the source of excitation energy needed to detect the fluorescent dye emissions.
    - 3130 has dual laser illumination to evenly excite all of the capillaries simultaneously.

Charge-Coupled Device (CCD) Camera

- Uses a small, rectangular piece of silicon rather than a piece of film to receive incoming light.
- This piece of silicon is divided into many light sensitive cells, each cell represents one pixel (picture element) of the entire picture. The surface of the camera is 256 pixels wide by 550 pixels long.



Charged-Coupled Device

How does a Charged-Coupled Device work?

- CCD can receive charge via the photoelectric effect and electronic images can be created.
  - The photoelectric effect is synonymous with Stokes Shift.
- Each one of the pixels on the surface of the camera accumulates (builds up) an electric charge (from the emitted photos) proportional to the light intensity at that ms in time.
- All at once (every ms) this stored electrical energy is sent as an electronic signal to the data collection software.
- This information is then converted (using mathematical algorithms) into the peaks that you see in an electropherogram.

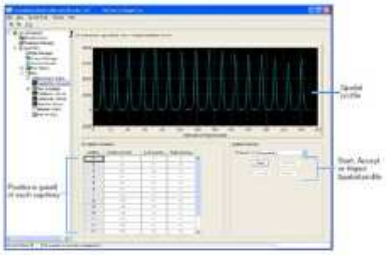
Optical Calibrations 3130

Optical Calibration

Spatial (3130):

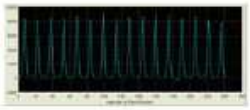
- A spatial calibration maps the position of each of the 16 capillaries on the surface of the CCD camera.
- A spatial is performed when a new array is installed, the detection window is opened, and routinely to ensure proper instrument function.
- The spatial calibration uses the Raman water signal from the polymer within the array to determine the precise location of each array.

Spatial Calibration

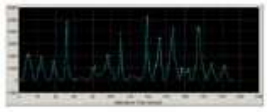


Spatial Calibration

Good ☺



Bad ☹

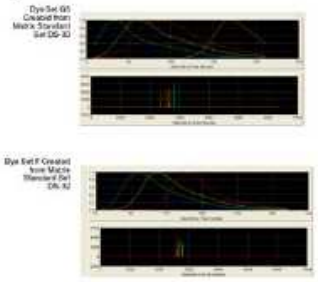


Spectral Calibrations

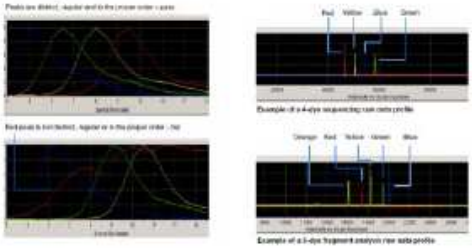
Spectral (3130):

- Equivalent to a Matrix on the 310
- A spectral calibration creates a mathematical algorithm to correct for the overlapping of fluorescence emission spectra of the dyes.
- Spectral calibrations are run if pull-up is seen often, after service calls involving the optics, and approximately once a year to ensure accurate results.

Optical Calibrations



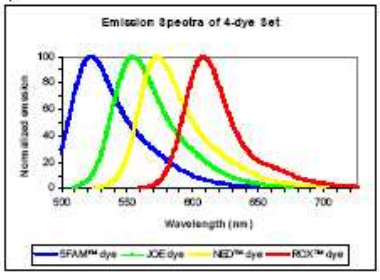
Optical Calibrations



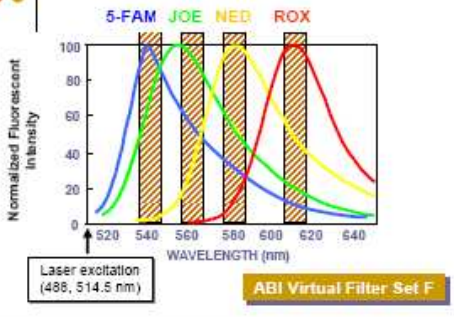
The Spectral

- Generated by a separate run / analysis of a dye set.
- Contains information about how much of the collected light falling on a virtual filter is due to the intended light (color) emission, and how much is from the overlapping *other* colors.
- Using a mathematical algorithm (we don't need to know) the overlapping colors are subtracted out, leaving only the peak / color of interest.

Dye Set F (Profiler Plus, COfiler)



Fluorescent Emission Spectra for AB Dyes



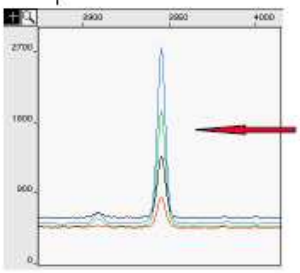
4-dye Matrix
(you can actually view this file on the 310)

The amount of spectral overlap observed for each dye in each virtual filter

Reactions

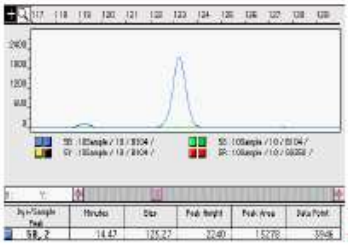
	B	G	Y	R
B	1.0000	0.7622	0.0583	0.0027
G	0.5998	1.0000	0.5824	0.0098
Y	0.4036	0.7003	1.0000	0.1583
R	0.1877	0.3812	0.5470	1.0000

Raw data (*only on the 310*)



No color subtraction

Analyzed data
(after the spectral is applied)



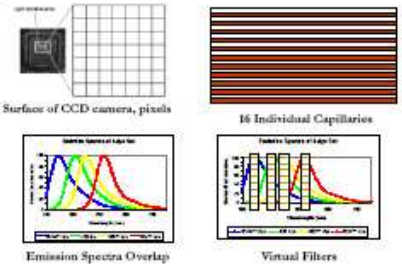
table

This is different for the 3130

- The matrix file is known as a Spectral Calibration.
- For the 3100 series instruments the spectral file is applied to the data as it is collected, thus no true raw data.
- How is the Spectral applied simultaneously as the data is collected, you ask?!?

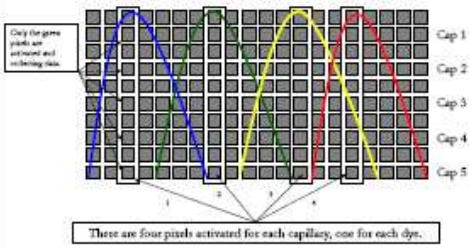
Excellent Question!

This is Complicated…Listen Closely!!!



All Overlaid On Top of Each Other!!!

Tah-dah!!!

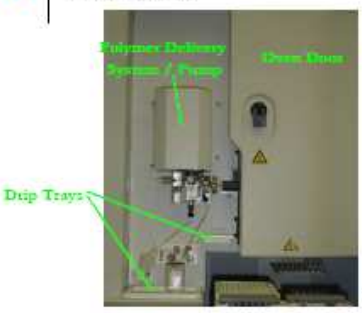


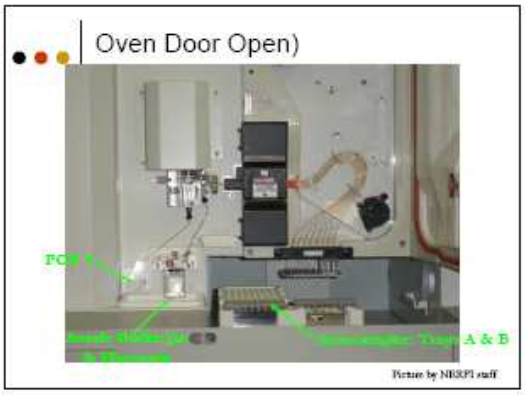
One intersection point

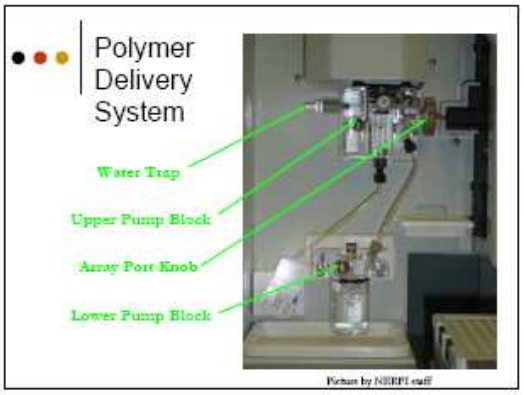
- But remember (from slide 18) that the surface of the CCD is 256 X 550 pixels = 140,800.
- …..and each one of the activated pixels on the previous slide is really 3 X 14 = 42 pixels.
- So, 42 * 4 = 168/capillary
- 168 * 16 = 2688 pixels activated.

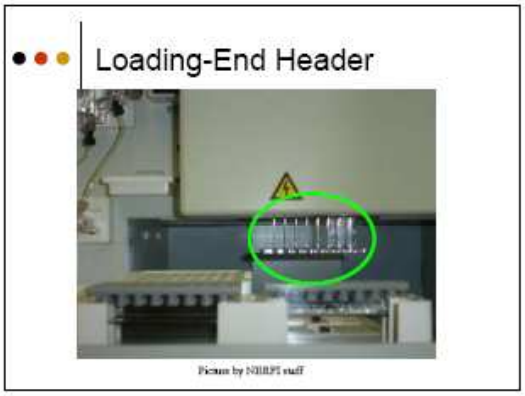
Nuts and Bolts of 3130

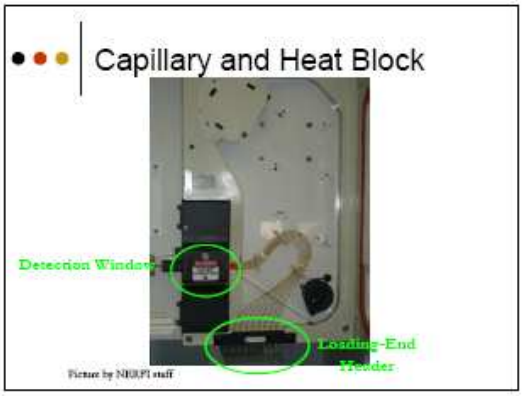
The Guts

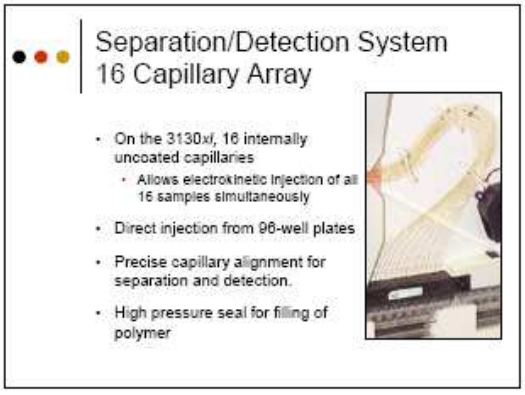


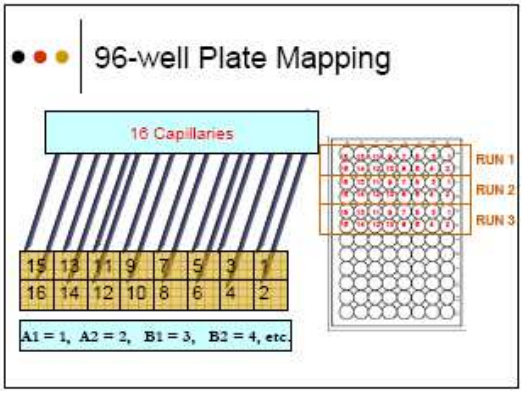












Workflow Diagram

Data Collection Software Features

- Collects data from electrophoresis runs
- Generation of sample files = .fsa files
- Allows operator to control and monitor activity of instrument
- Create/import/export plate records
- Edit modules to control instrument
- Perform and view instrument calibrations
- Wizard-based set up and maintenance
- Interface between operator and Oracle database

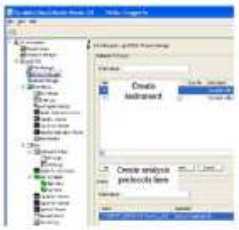
Navigator Pane

Creating an Instrument Protocol

In the tree pane of the Data Collection software, click GA Instruments > ga3130(f or ga3130 > Protocol Manager

Click NEW to open Protocol Editor

The Instrument protocol will hold all the settings necessary to run the instrument



Creating an Instrument Protocol

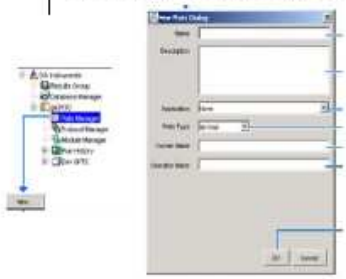
Name	HIDFragAnalysis36_F_POP4
Description	Optional
Type	Regular
Run Module	HIDFragmentAnalysis36_POP4_1
Dye Set	F

Results Group

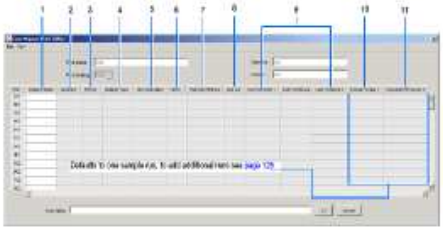
The Results Group organizes samples and user settings under a single name.

It will be used to name, sort, and deliver samples (to a storage file) that result from a run.

Creating a Plate Record



GeneMapper Plate Editor



GeneMapper Plate Editor

1. Sample name	Name of sample
2. Comment	List sequential three digit numbers (ie 001, 002, 003, etc.)
3. Priority	Default 100 (all samples run in order)
4. Sample Type	Sample, Positive Control, Negative Control, Allelic Ladder
5. Size Standard	CE_HID_GS600
6. Panel	DS-32/DS-33 (COfiler-Profiler Plus/Identifiler)
7. Analysis Method	Default (analyze in GeneMapperID)
8. SNP set	None
9. User defined columns	Comments (not currently used)
10. Results Group	Here will be CADEWORK
11. Instrument Protocol	Here will be HID_FragAnalysis36_GS_POP4_5s

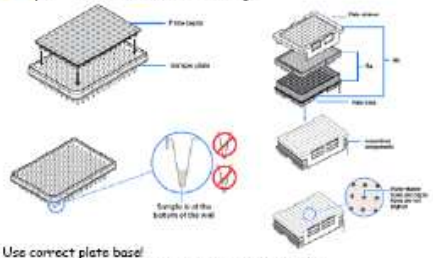
How to perform multiple injections of the same samples?

- Use the Add Sample Run tool under the Edit column
- Additional Results Group and Instrument Protocol columns are added.
- This is useful if you would like to inject a sample (from the same well) twice with different injection times.

Running the Instrument

1. Prepare Samples and Plate Assemblies
2. Place Plate onto Autosampler
3. Schedule a Run
4. Link the Plate
5. Run the Instrument

Plate Assembly



- Use correct plate base!
- Make sure reservoir septa are firmly seated and flat!
- Make sure holes in the plate retainer align with the holes in the septa!

The Autosampler

- Positions plates correctly for sample injection
- Holds 96 and 384 well plates
- Electric sensors detect presence/type of plate
- Holds cathode buffer vial, waste, and wash reservoirs
- Capillaries embedded in electrodes
- Autosampler calibration only as needed.

Linking the Plate

- Under the Run scheduler, select the desired plate record
- Click the plate position (A or B)
- Plate position indicator changes from yellow to green when linked
- The green *run* button becomes active

Running the Instrument

- Very Easy

Run Procedure

1. Capillary Fill –
   - Capillary moves from the buffer reservoir to waste reservoir
   - Capillary is filled with Pop-4 by the force applied by the cylinder.
2. Pre-electrophoresis
2. Water wash of the capillary
   - Capillary dipped twice into water reservoir
   - Removes buffer salts
3. Sample injection
   - The autosampler moves the sample tray so that the sample plate is below the capillary loading end header
   - The autosampler then raises the tray up to dip the capillaries into the samples.
   - 15 kV for 5 seconds to draw the DNA up into the capillary

Run Procedure (cont.)

5. Water wash of the capillary
   - Dipped several times in waste reservoir in order to remove any extra sample, etc. on the outside of the capillary.
6. Water dip
   - Capillary is dipped in clean water several times.
7. Electrophoresis
   - Capillary moved to buffer reservior
   - Another voltage is applied to allow the DNA to traverse the capillary
   - DNA begins to separate through the POP-4 polymer
   - Approximately 40 minutes
8. Detection
   - Data collection begins
   - Mathematical Algorithms established by the spectral are applied simultaneously
   - Can see the data in real-time

Things to be mindful of…

- LED lights located on the front of the instrument
  - Steady green = ready
  - Blinking green = running
  - Steady green and blinking amber = end of run
  - Blinking amber = paused state, door open, self test
  - Steady red = failure (check error log)
- Cathode
  - Made of platinum ($$$ ☺)
  - Check for crystals before each run
  - Very easy to bend!! Wipe gently with kimwipe and distilled water

Things to be mindful of…

- Pump Blocks
  - Polymer is changed every 7 days
    - Or sooner if you notice a decrease in resolution (broader peaks)
  - Be sure the polymer blocks are clean and dry

Things to be mindful of…

- Capillary
  - Change it when you notice poor resolution
  - Do not let the ends of the capillary be exposed to air for more than 30 min
  - Be sure the capillary window is over the laser detector window, no fingerprints!
  - The capillary is tough, yet fragile!
    - Don't be afraid to handle it, but avoid applying undue stress, such as twisting
    - Capillary window is fragile

3130*xl* Maintenance Tasks

The following should be performed **BEFORE EACH BATCH OF RUNS**:

- Reservoir septa are firmly seated and flat
- Holes in the plate retainer align with the holes in the septa
- Plate assembly sits snugly on the plate deck
- Clean instrument surfaces, checking for dried polymer
- Check for leaks around array knob and interconnections

3130*xl* Maintenance Tasks

The following should be performed **DAILY or BEFORE EACH BATCH OF RUNS:**

- Clean and replenish water and 1X GAB (Genetic Analyzer Buffer) reservoirs.
- Remove bubbles in the polymer system (via *Wizard*)
- Check the loading-end header to ensure the capillary tips are not crushed, damaged or have visible crystals.
- Check the level of polymer to ensure that there is sufficient volume for runs intended.
  - Apx. 60-80ul per set of 16 injections

The Wonders of Wizards

Change Polymer Wizard
- To change DIFFERENT polymer types, e.g. POP-6 to POP-4

Instrument Shutdown Wizard
- To prepare instrument for being idle more than one week

Replenish Polymer Wizard
- To replenish array with SAME polymer (same or different lot)

Bubble Remove Wizard
- To remover bubbles in PDP chamber, channels, and tubing

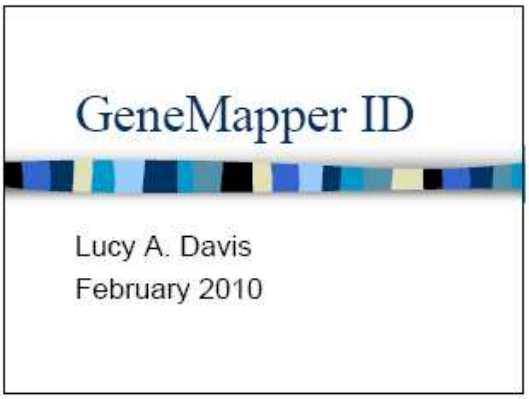
Water Wash Wizard
- A part of the recommended maintenance procedures

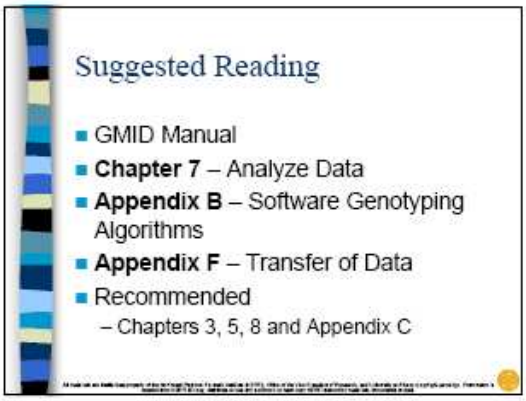
Install Array Wizard
- To install or replace a new array

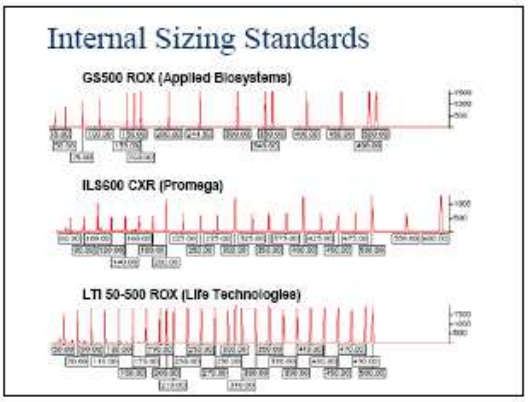
Electropherograms

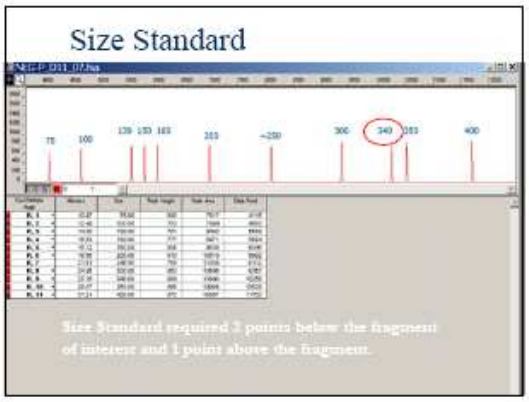
GMID lecture with Lucy ☺

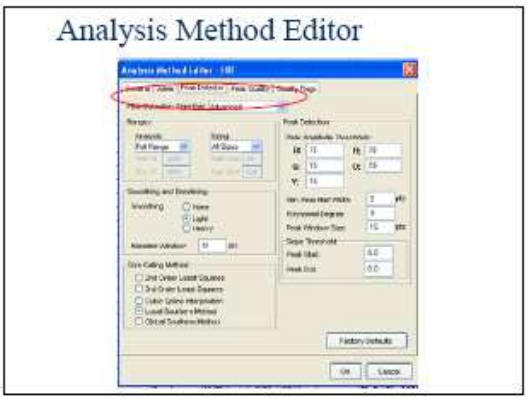
Appendix 4: Genemapper ID Presentation:

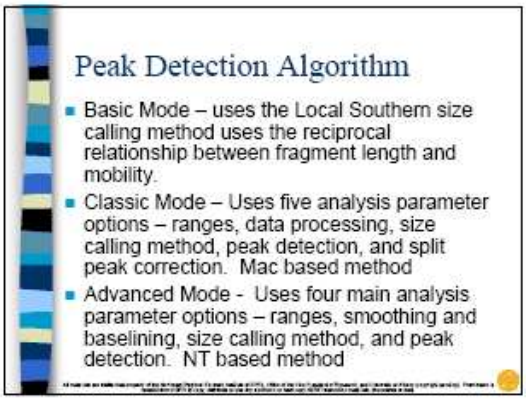


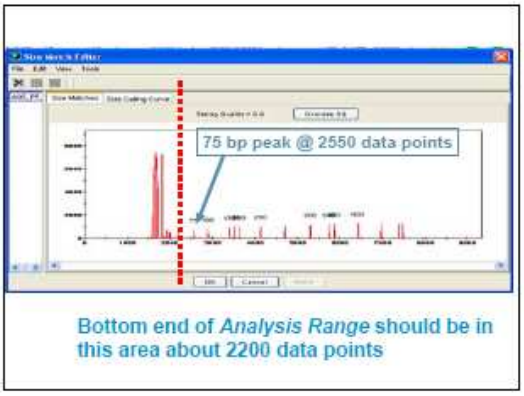


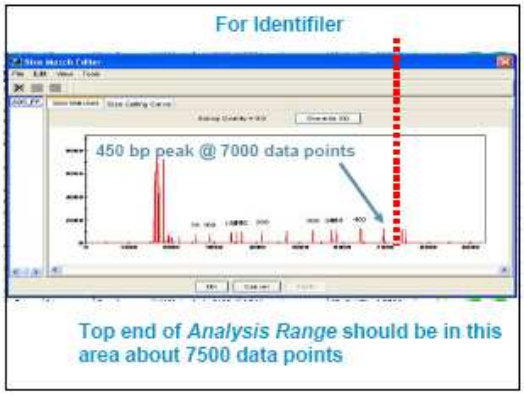


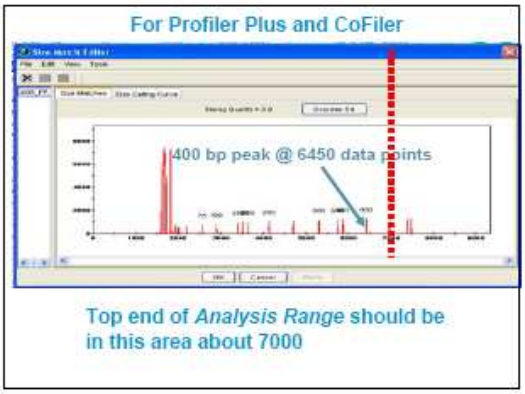


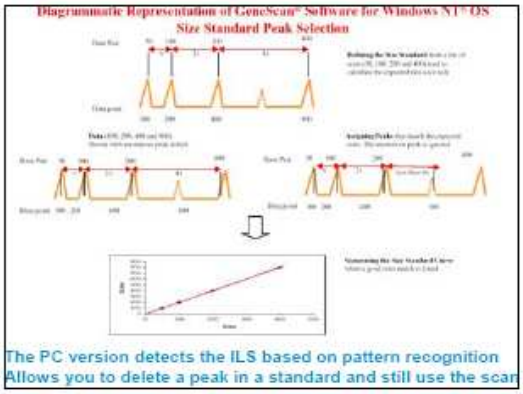


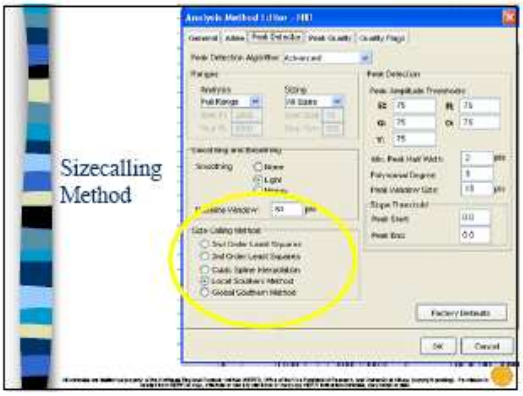








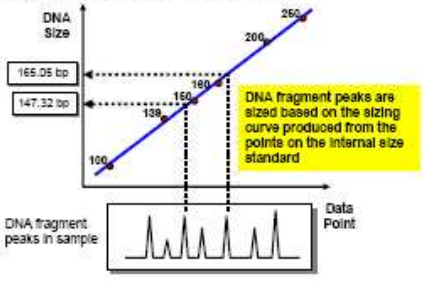




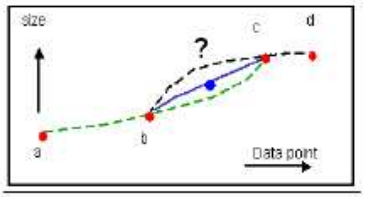
Sizecalling Methods

- **Local Southern** – Type used by forensics. Determines the sizes of fragments by using the reciprocal relationship between fragment length and mobility
- 2nd (or 3rd) Order Least Squares – Uses regression analysis to build best-fit curve
- Cubic Spline – Forces sizing curve through all known points
- Global Southern – Similar to the Least Squares but compensates for electrophoresis migration abnormalities, creates a best-fit line through all available points

Process of Sizing DNA Fragments Using an Internal Standard



Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4335755 Rev. C

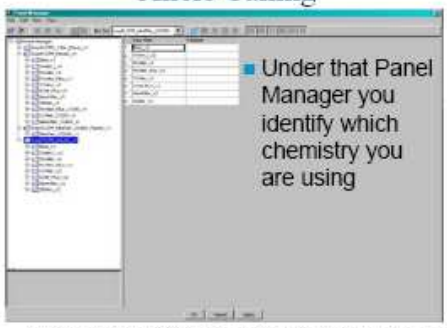


Blue point is the data point of the allele of interest

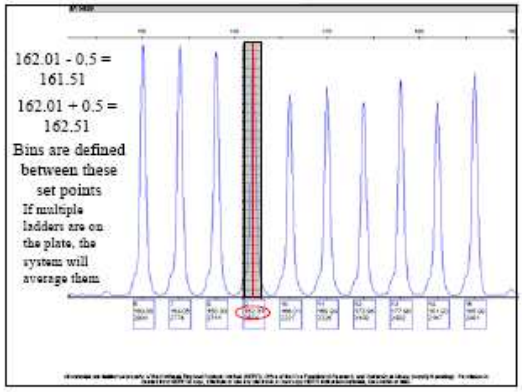
1. Fit a curve through points a, b, and c
2. Fit a curve through points b, c, and d
3. Average the calculated base pair size determined by steps 1 and 2, average those 2 numbers together to get the base pair sized of the fragment of interest

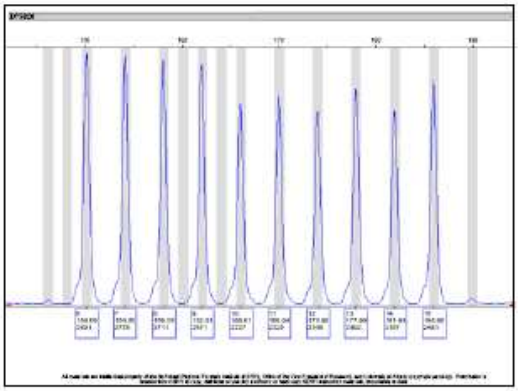
GeneMapper® ID-X Software Version 1.0 Reference Guide, Part Number 4375871 Rev. A, Page 20

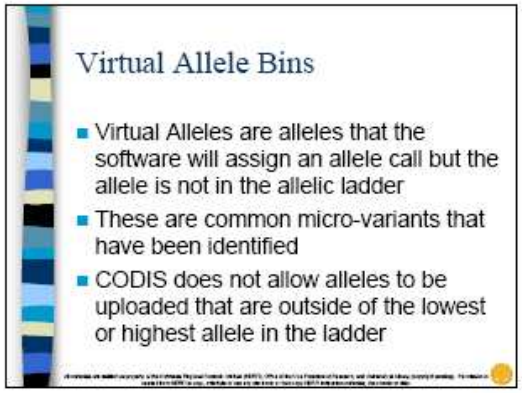
Allele Calling

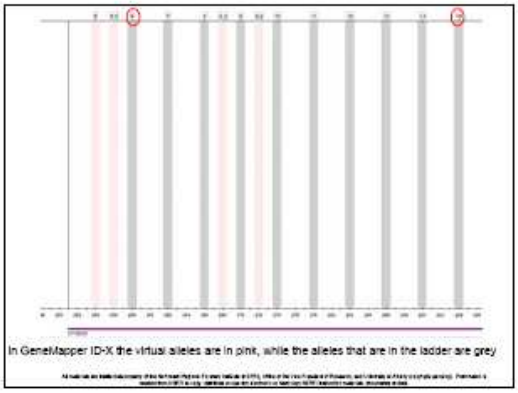


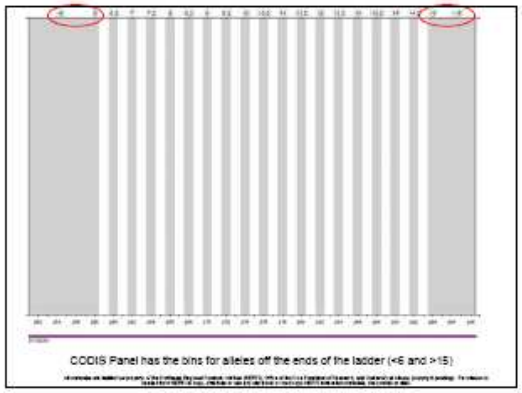
- Under that Panel Manager you identify which chemistry you are using

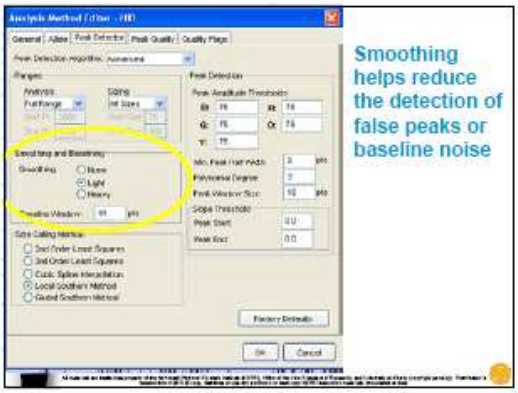






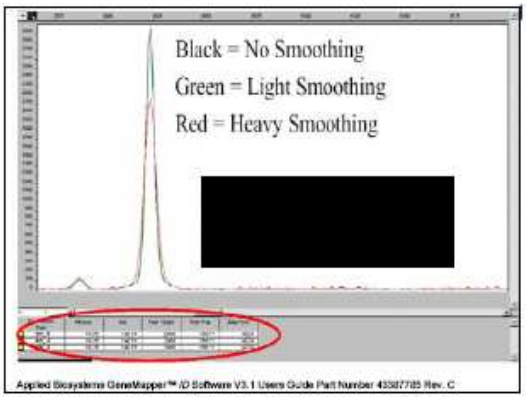






Smoothing Option

- Reduces the number of false peaks detected by the software.
  - None – if the data has very sharp, narrow peaks of interest
  - Light – provides best results for typical data
  - Heavy – apply to data from slower runs that has very broad peaks, or to avoid the detection of sharp edges. Might reduce peak size or eliminate narrow peaks



Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4335775 Rev. C

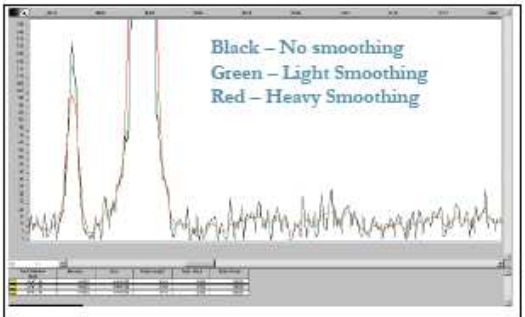


Figure 4 Electropherogram showing the effects of smoothing on the smaller peak and baseline by changing the y scale from Figure 3

Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4335775 Rev. C

Baseline

- GMID computes the baseline for each dye color independently
- GMID baselines an electropherogram by subtracting the baseline from the raw electropherogram
- The baseline value is the lowest GMID allows you to choose the baseline data point that sets where the software will set the lowest value on the electropherogram

Baseline

- A small (low) baseline window
  - Causes the baseline to creep into the peaks, resulting in shorter peaks in the analyzed data
- A large (high) baseline window
  - Causes the baseline to ride to low, resulting in elevated and possibly not baseline-resolved peaks

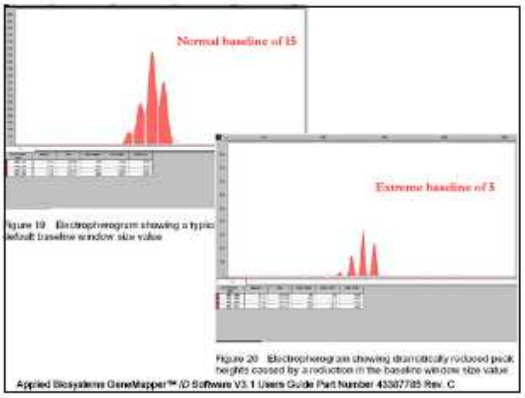


Figure 19 Electropherogram showing a typical default baseline window size value

Figure 20 Electropherogram showing dramatically reduced peak heights caused by a reduction in the baseline window size value

Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4335775 Rev. C

Baselining example

The red trace is a baseline of 2801

The blue trace is a baseline of 51

The black trace is a baseline of 5

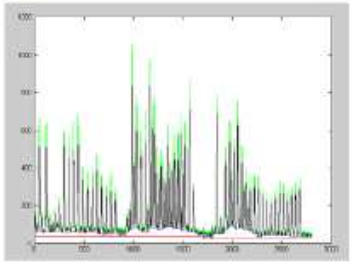
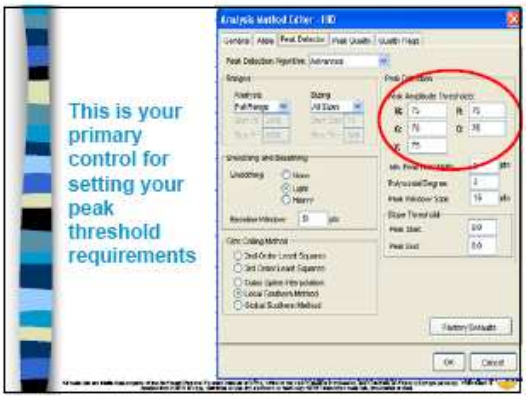
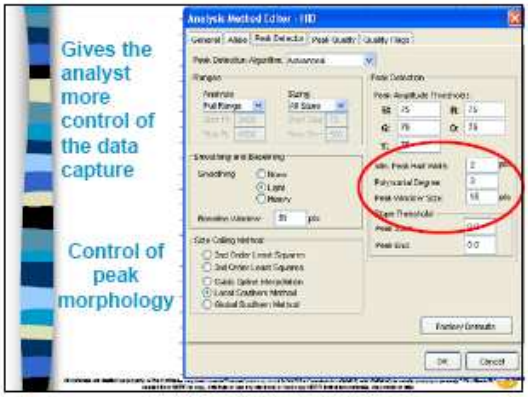


Figure 1 Baseline Window example

GeneMapper® ID-X Software Version 1.0 Reference Guide, Part Number 4375671 Rev. A, Page 7



This is your primary control for setting your peak threshold requirements



Gives the analyst more control of the data capture

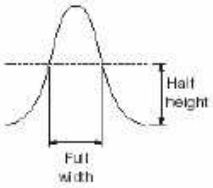
Control of peak morphology

Peak Detection

- The Minimal Peak Height Width, Polynomial Degree, and Peak Window Size settings defines what the software labels as a peak
- It sets the parameters for the software to evaluate
- How the software keeps from labeling a spike or dye blob as an allele

Min. Peak Half Width

- Specifies the smallest full width at half maximum height for peak detection. Can be used to ignore noise spikes.
- Defines what constitutes a peak. Software ignores peak half widths smaller than the specified value.
- If a peak meets the value, then the software will assign an allele call, if not the software ignores the peak



Range is 2-99

Full width of the peak measured at half its height

Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4336785 Rev. C

Polynomial Degree & Peak Window Size

- Affects the sensitivity of peak detection
- These can be adjusted to detect a single base pair difference while minimizing the detection of shoulder effects or noise
- The peak detector computes the first derivative of a polynomial curve to the data within a window that is centered on each data point in the analysis range

Polynomial Degree

- A polynomial is an expression of length constructed from mathematical equations (i.e. $x^2 - 5x + 2$)
- The degree is the number of terms defined in the equation (the above equation has a degree of 3)
- Curves with larger polynomial degrees (more variables in the equation that defines the line) will identify shaper increases in the slope of the line making the peak on the EPG

Polynomial Degree & Peak Window Size

- Using curves with larger polynomial degrees (more variables in the equation that defines the line) allows the curve to more closely approximate the signal and therefore the peak detector captures more peak structure in the electropherogram
- The peak window size sets the width (in data points) of the window to which the polynomial curve is fitted to data. Higher peak window size values smooth out the polynomial curve, which limits the structure being detected. Smaller window size values allow a curve to better fit the underlying data.

Constant window size 15

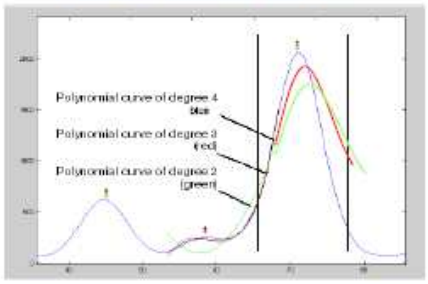


Figure 6 Electropherogram showing peaks detected with the same window size and three different polynomial degrees

Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4338775 Rev. C

Constant Polynomial curve 3

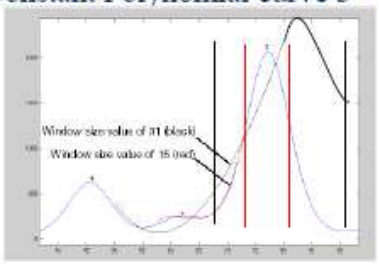


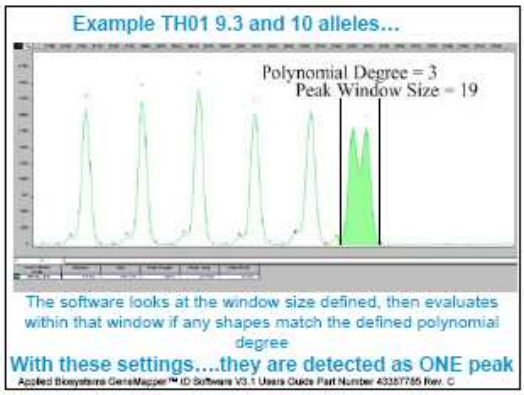
Figure 7 Electropherogram showing the same peaks as in Figure 6 after increasing the window size value while keeping the polynomial degree the same

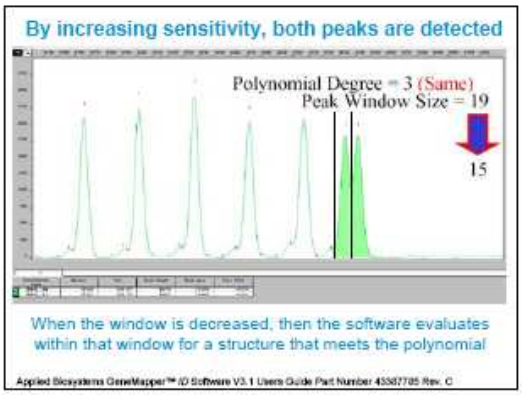
Applied Biosystems GeneMapper™ ID Software V3.1 Users Guide Part Number 4338775 Rev. C

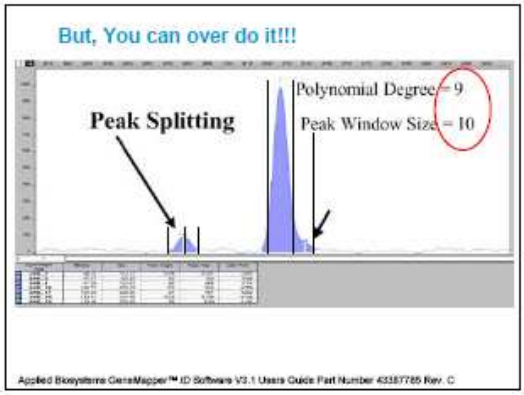
Peak detection Sensitivity

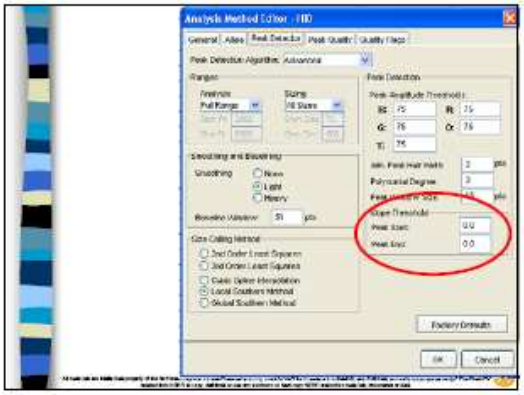
Controlled by:
Polynomial Degree & Peak Window Size

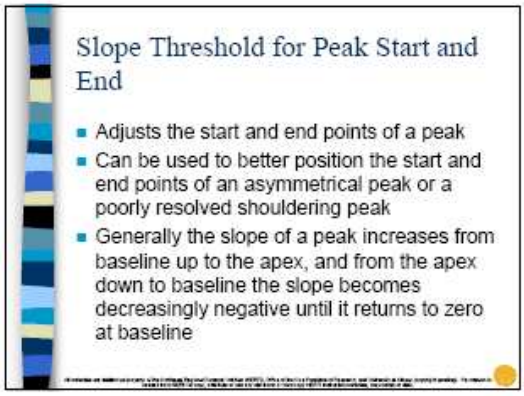
More Sensitive
(Increase) Polynomial Degree
(Decrease) Peak Window Size
Less Sensitive
Polynomial Degree
Peak Window Size

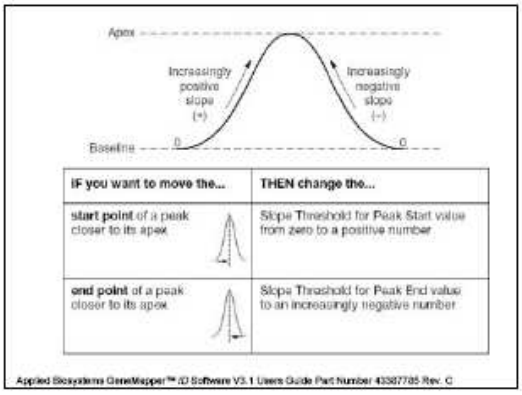


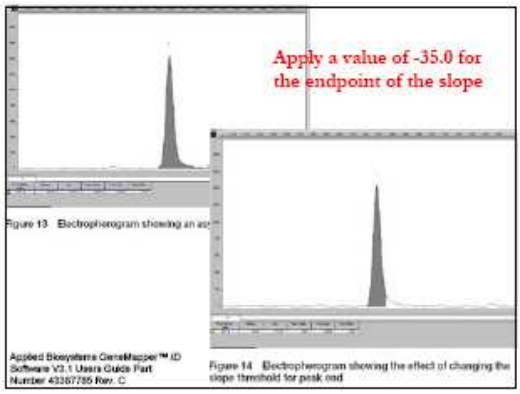












Data Analysis - NERFI Laptops GeneMapper ID v3.2.1

- Log on (printed on the left hand side of key pad
  - User name: NERFI
  - Password: rox!500!
- GeneMapper ID (GMID) Log on (printed on the right hand side of key pad
  - User name: gmid
  - Password: gm3130

Create Analysis Method – NERFI Workshop 100208

Most common mistake made while creating analysis methods in GMID is not identifying Bin Set

Importing Settings

- Import Analysis Parameters
- Copy the parameters to a specific location on your computer or USB drive
- GeneMapper Manager→Click on *Analysis Setting* tab→Import→Highlight appropriate parameters → Import

Importing Settings

- Import previous GMID projects
- This is where you will import projects created on other computers
- If you were using a 310 you would have to import matrices here

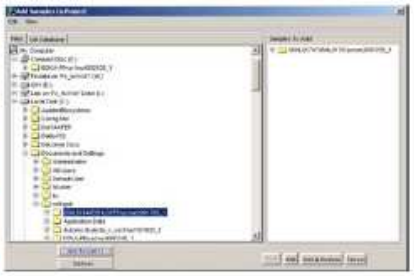
Creating a new project

- Once GMID is open, there will be a blank project on the screen
- From here, the sample files from the run folder will need to be added.
- To do this, go to File>Add Samples to Project, Ctrl K, or the icon

Navigating to your run folder

- Navigate to your run folder that you want to analyze
- Move over the entire run file by selecting add to list

-Batch run folders will already be created specific to each case off the 3130
-When you copy you data files make sure you put your sample files in a folder under you're "My Documents" file. Do NOT put them on your desktop and leave them there



Sorting with the Comment

- The Comment column will allow you to sort in the order of injection
- If you don't do this, your plots will be in alphabetical order, not injection order
- Make sure the Comment is selected in the samples table

-To check this, select the table icon or Ctrl T, and be sure the Comments is checked
-Next, go under Edit>Sort (Ctrl G), and sort by Comments

Now you can analyze your project

- Be sure the correct analysis parameters are selected

if you modify the run range, MAKE SURE you reset it to original settings

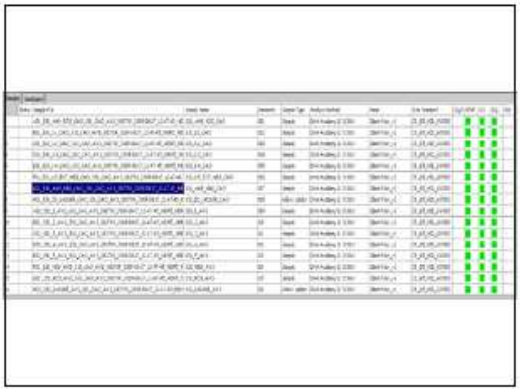
Be sure to check all size standards
-CE_G5_HID_GS500

- Verify sample type column is correct

-Allelic ladders are listed as 'Allelic Ladder'
-9947a is listed as 'Positive Control' (optional)
-Kit negative and reagent blanks are designated as 'Negative Control' (optional)
-Blanks may be deleted from project (optional)

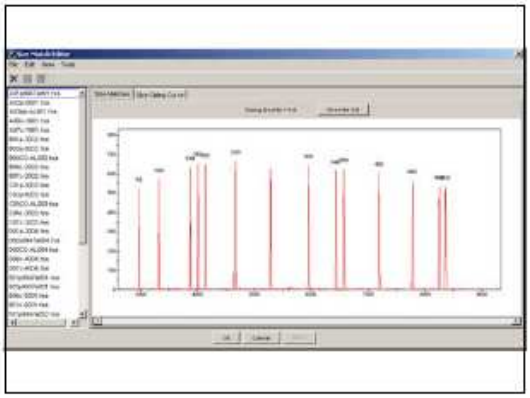
- Verify the panel information is correct

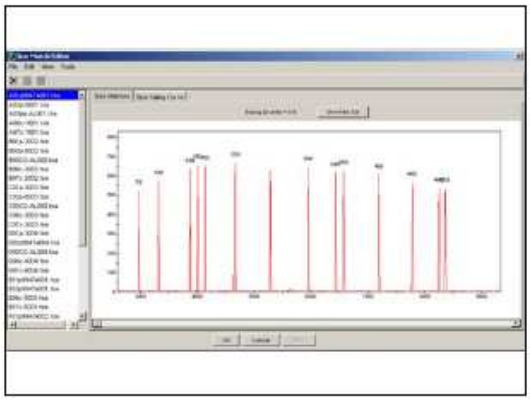
-You are using Identifiler_v2



Analyzing the Project

- Select All (Ctrl A) to analyze the project
  - -This will prompt you to save the project. Name the project (e.g., SampleSet# / SimSet #_Date_ initials)
  - -Once the project is analyzed, the green arrows will no longer be visible
- Select All files and go to the Size Match Editor
  - -From the Size Match Editor, you will be checking all of your Internal Lane Standard (ILS) peaks
  - -It is helpful to drag the Size Match Editor Box to the top left corner of your screen, and in the bottom right corner, pull the box to the size of the screen
  - -A completely blue box around a sample will allow you to use the arrow keys to go up and down
  - -A highlighted blue box will not allow you to view your samples with the arrow key

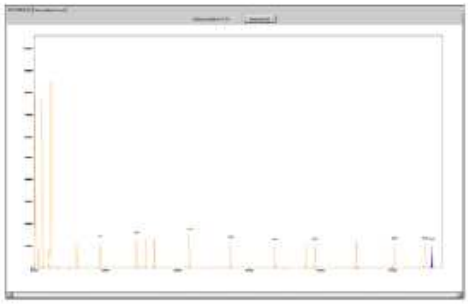




Problems with Size Standards

- Make sure you have the correct size standard selected
- If you are missing the 75 or the 450 peak, you will need to reanalyze your ***entire*** project with expanded parameters

What is wrong with this standard?



Problems with Size Standards

IF…	THEN…
Basepair sizes for peaks 50*, 75, 100, 139, 150, 160, 200, 300, 340, 350, 400, 450	
Peak heights are above defined threshold	Size standard (ILS) signal is acceptable. Continue
Peak shape is broadened and/or many intermediate sizes called	ILS may be unacceptable. Re-inject
Incorrect assignments exist for any peak(s)	If the incorrectly assigned peak(s) is needed to incorporate the local southern sizing method, re-inject sample
Standard peak(s) is unable to be redefined	Unknown fragments cannot be analyzed for that region. Re-inject or re-run

Raw Data Review

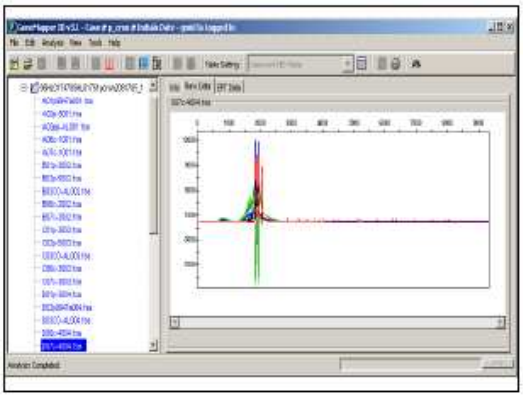
- In order to look at raw data, select all sample files to View>Raw Data (Ctrl F2)

-The window functions in a similar fashion to the Size Match Editor

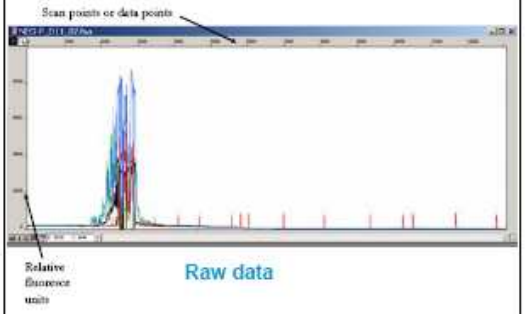
- The off-scale value for multi-capillary instruments is not a standard set value, like the 8191 on the 310. It will vary depending on the sample, run, capillary etc. In addition, if a spike, pull-up, or overlapping peaks are present, it may trigger the off-scale flag for samples less than 8000.

-Remember: OS in samples tab means there is an off-scale peak w/in the size standard region (75-500)

-OS in the genotypes tab means there is an off-scale region w/in the marker range



Data & size calling range



Off-scale Samples

- Peak height must be within the acceptable range for the instrument being used. **For samples with raw data ~7400RFUs on the 3130 data is interpreted with caution.**
- Multicapillary instruments may produce off-scale data less than 7400, be sure to examine the raw data. The use of the Off-scale (OS) PQV is a helpful indicator of off-scale data within a sample and locus.
- The sample may be interpreted with caution, but it can be re-injected at a decreased injection time.

Process Quality Values (PVQ) Samples page

PQV Column	Description
SQO	Size Quality Over-ride
SFNF	Sample File not Found
OS	Off Scale
SQ	Size Quality

Process Quality Values (PVQ)
Genotype page

PQV Column	Description
AE	Allele Edit
ADO	Allele Display Overload
AN	Allele Number
Bin	Out of bin allele

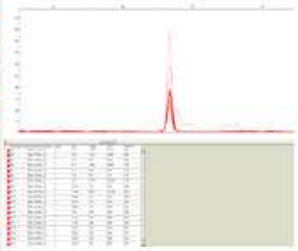
Displaying Plots

- Highlight the desire sample(s)
- Analysis>Display Plots
  - Samples with low sizing quality value can not be displayed
- To zoom in/out: Right click on X or Y axis>Highlight Zoom to…> enter desired value
  - Apply to all electropherograms
- Hold pointer over X or Y axis, what till magnifying glass appears, right click and drag

250 Basepair Peak

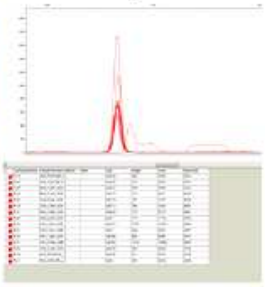
- The 250 basepair peak for each sample should be check to confirm the migration is not >±1bp.
- You can look at this data by going to Display plots and reviewing Orange dye only
- If the data is outside range – consult Supervisor/Instructor.
  - Data may need to be re-run or broken into smaller projects

250 Basepair Peak Evaluation



In display plots you can look zero in on only the 250 peak of all samples over-laid on top of each other

250 Basepair Peak Evaluation



When you filter on "show select rows" the table will show the base pair size of just the 250 peak

Deciphering Artifacts from the True Alleles

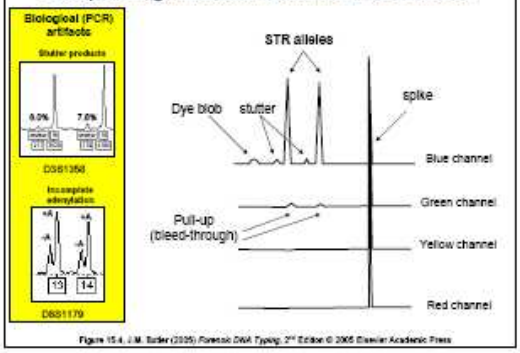


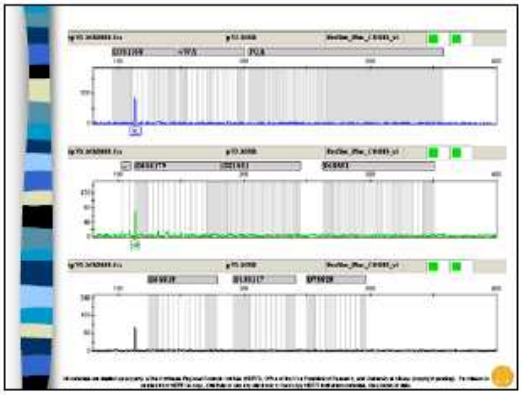
Figure 15.4, J.M. Butler (2005) Forensic DNA Typing, 2nd Edition © 2005 Elsevier Academic Press

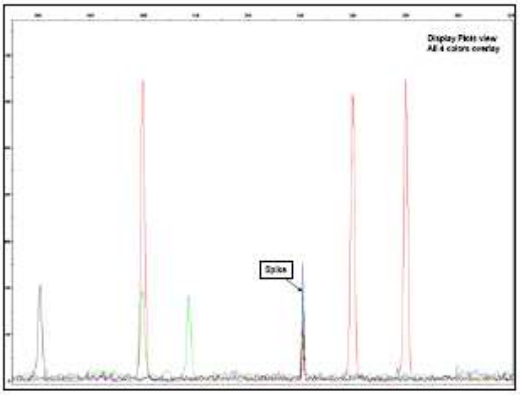
Artifacts

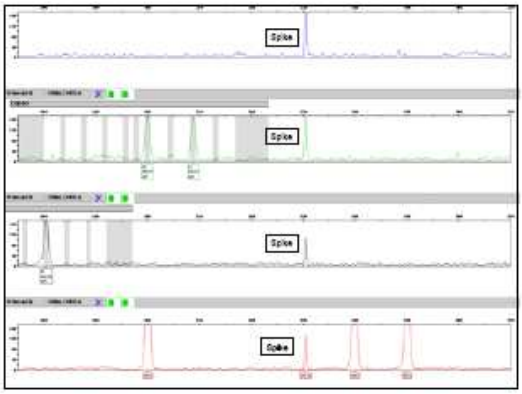
- Peaks may be labeled when they are not actual alleles – Artifacts
- Artifacts may be edited
- Edited peaks meeting specific criteria must be documented

Spikes

- Peaks are approximately the same bp size and may be present in all 5 colors
- A spike in <5 colors can be viewed by overlaying bins
  - If necessary re-run

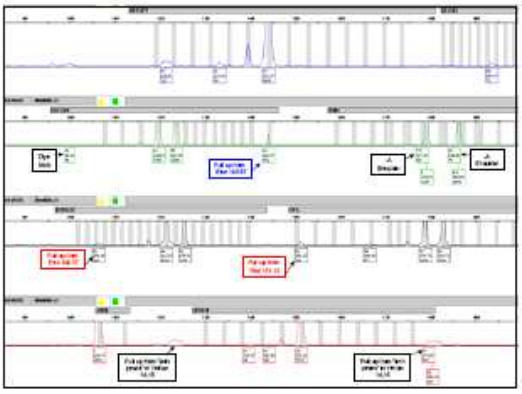


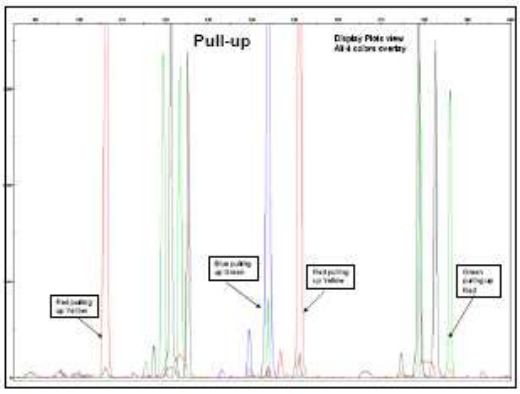


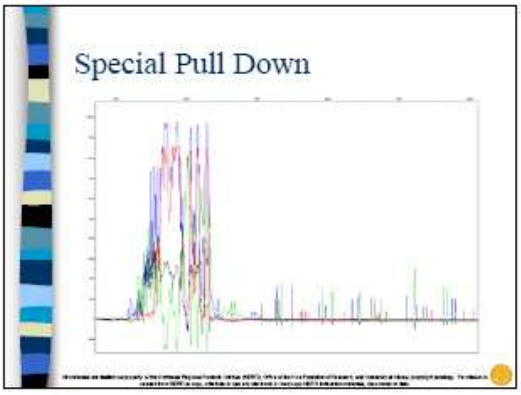


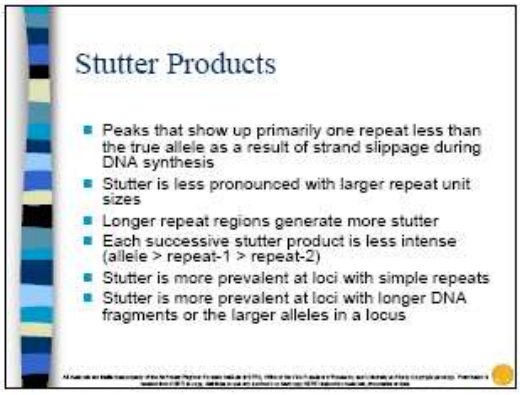
Pull up

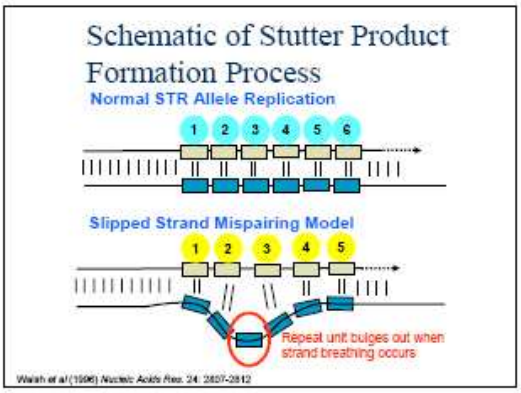
- Spectral overlap
- Pull up peaks can be approximately the same bp size as true alleles in another color

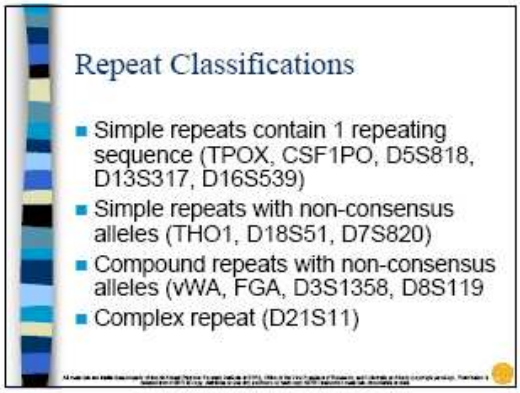




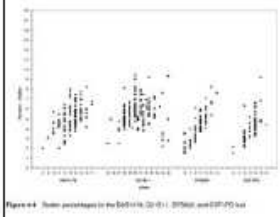








AB Validated stutter per allele

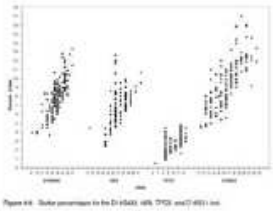


- CSF1PO – simple repeat – $[AGAT]_N$
- D7S820 – simple repeat – $[GATA]_N$
- D8S1179 – compound repeat – $[TCTR]_N$

D21S11 – complex repeat – $[TCTA]_N[TCTG]_N\{[TCTA]_3TA[TCTA]_3TCA[TCTA]_2TCCATA\}[TCTA]_NTA\ TCTA$

Applied Biosystems AmpFISTR® Identifiler™ Users Manual Part Number 4323291 Rev. C

AB Validated stutter per allele



- D18S51 – simple non-consensus repeat $[AGAA]_N$
- TPOX – simple repeat $[AATG]_N$
- vWA - Compound non-consensus repeat TCTA[TCTG]3-4[TCTA] $_N$

Applied Biosystems AmpFISTR® Identifiler™ Users Manual Part Number 4323291 Rev. C

Different types/degrees of Stutter

- If the template strand bulges out during the replication you will get stutter in the -4 position
- If the complimentary strand bulges out, you will get stutter in the +4 position
- Excessive DNA template can also lead to stutter in the -8 position

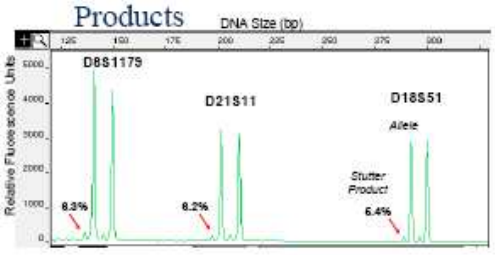
Stutter

- Majority of stutter is n-4 from true allele
- Interpret with caution if there is an indication of mixture
- We will use validated stutter values
- Forward stutter n+4 is usually seen when you look at data 50 rfu's and lower
- Extremely overloaded samples may have n-8 stutter

Stutter (continued)

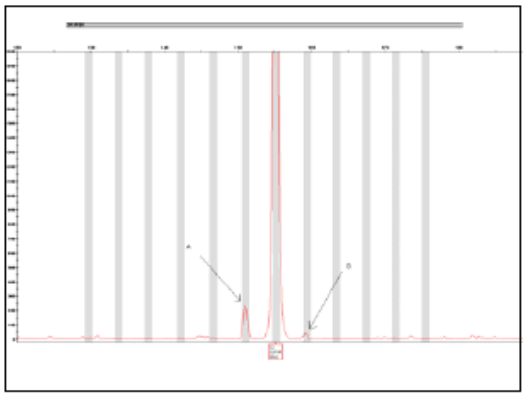
- The amount of the shorter (stutter) product is measured as a percentage of the height of the main peak
- Dividing the RFU value of the shorter signal by the RFU value of the main peak produces the stutter percentage
- The maximum percentage of stutter product has been identified for each STR locus
  - GMID will remove the allele label for signals within this maximum percentage
  - GMID computes the stutter percentage by dividing the RFU value of the stutter signal by the RFU difference in the two peaks

STR Alleles with Stutter Products



Expected Stutter Values

- Through validation studies, it is possible to determine maximum stutter values expected at a loci
- GenemapperID has adjustable stutter filter settings - peaks under max stutter value are not labeled

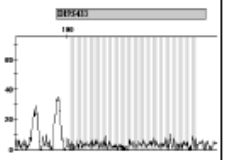


Stutter Cut-off Values

LOCUS	9947A RESULTS	% CUT OFF
D8S1179	13	11
D21S11	30	12
D7S820	10,11	8
CSF1PO	10,12	8
D3S1358	14,15	10
THO1	8,9.3	4
D13S317	11	9
D16S539	11,12	12
D2S1338	19,23	11
D19S433	14,15	13
vWA	17,18	11
TPOX	8	5
D18S51	15,19	15
Amelogenin	X	NA
D5S818	11	9
FGA	23,24	11

Dye Artifact peak

- Caused by unincorporated dye from primers, that can migrate into the range of interest.
- Morphology typically does not resemble a peak (broader and poorly shaped)



Shoulder peak

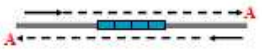
- Caused by broadening of a peak and appears approximately 1-3 bp smaller or bigger than the main allele

Raised baseline

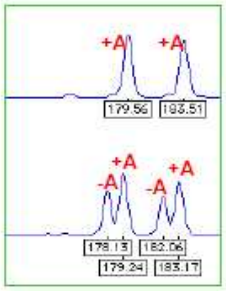
- Caused by fluctuations of the electrical current
- May occur in one or all five colors
- Edit noting base pair range and color

Non-template Addition

- Taq polymerase will often add an extra nucleotide to the end of a PCR product; most often an "A"
- Can be enhanced with extension soak at the end of the PCR cycle (e.g., 15-45 min @ 60 or 72 ºC)
- Can be reduced with new polymerase
- Best if there is NOT a mixture of "+/- A" peaks



-A product



- Caused by incomplete adenylation
- Appears approximately 1 base pair less than true allele

Artifacts List

TYPE of ARTIFACT	GMID EDITING ABBREVIATION
Stutter above cut-off value/elevate	ESTR
N-1/Minus A/Split peaks	N-1
Pull-up	PUB, PUG, PUY, PUR or PUO
Raised baseline	RBL
Spike	SPKE
Dye Blob	DYEB
Miscellaneous artifact	MISC
Inconclusive typing result	INC (must include a handwritten note on each page of the EPG regarding why the result was deemed inconclusive)

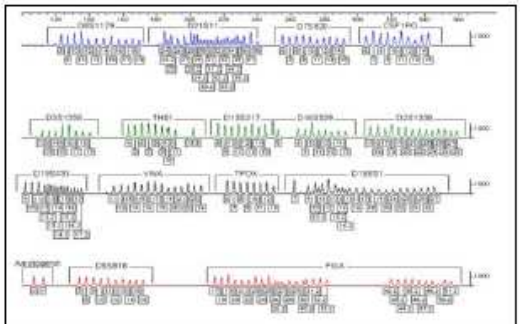


Figure 5-1 Genotyper® software plot of the AmpFℓSTR Identifiler Allelic Ladder, indicating the designation for each allele. These results were obtained on an ABI Prism 310 Genetic Analyzer

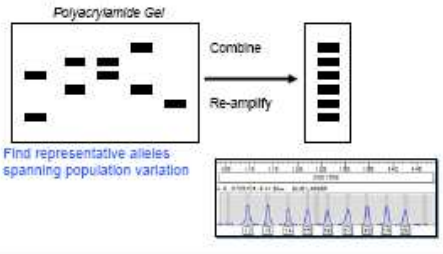
Applied Biosystems Identifiler Users Manual Part Number 4323291 Rev. C, page 144

Allelic Ladder

- Contains common alleles and some variants
- The common alleles are taken from a pool of individuals which will result in alleles in the allelic ladder having different peak heights (the more common the allele, the higher the peak in the ladder)
- Combines all loci and dyes in one amplified sample for each system (PP/CO)

Allelic Ladder Formation

Separate PCR products from various samples amplified with primers targeted to a particular STR locus



Allelic Ladder (continued)

- Base pair sizes of the alleles from samples are determined using the Internal Lane Size Standard (ILS) and then compared to the bp sizes from alleles in the ladder (also sized with ILS) to obtain the allele call
- Each electrophoresis run must contain a working allelic ladder for each system being used
- Vendors provide amplified DNA of the ladder
- Allelic ladders are included in PCR Kits; will vary vendor to vendor

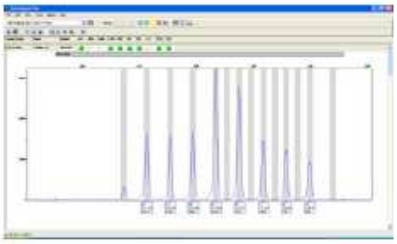
Discrete Alleles

- Amplification of STRs produce alleles of discrete sizes
- Discrete sized alleles allow for the construction of allelic ladders containing fragments of known (common) alleles at a given locus
- Allelic ladders permit simplified interpretation of results
- Multiplexing STR loci offers unlimited potential for a highly discriminatory test for human identification

Allelic Ladders
Check allelic ladders to assure:

IF...	THEN...
All peak heights are above threshold, and all alleles are present and well-resolved	Allelic ladder is acceptable
Peaks below threshold, or alleles missing or not well-resolved	Remove ladder – reanalyze If all are bad, repeat CE run (discuss with instructor)
A spike is present, in close proximity to an allele but has not affected allele assignment	Ladder may be used
A spike is present, in close proximity to an allele and does affected allele assignment	Remove ladder – reanalyze If all are bad, consult supervisor, you may be able to use another ladder

Virtual Alleles



Virtual Alleles what are they?

- Not all known alleles are present in ladder
- The software has sizes programmed in for known alleles that are not in the ladder
- You will get allele calls for these "virtual alleles"
- If your peak does not correspond to virtual alleles or represented ladder alleles
  - You get an OL allele call
    - May be a variant or a sizing error you have to determine which it is

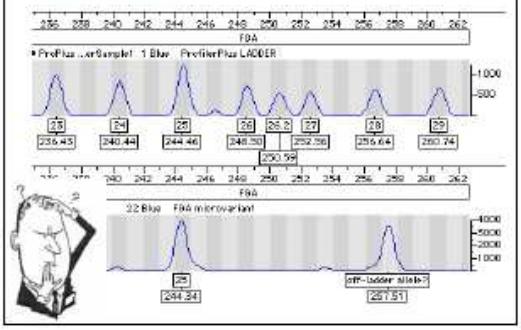
Off-Ladder Alleles

- Alleles detected that are not present in the ladder
- Sometimes called variant alleles
- Off-ladder alleles sometimes result when there are migration differences between alleles injection to injection; this can be corrected by sizing with the allelic ladder closest to the sample in the run
- If the off-ladder is reanalyzed but still not labeled, the sample must be repeated on another CE run or second amplification to show reproducibility

Off-Ladder Alleles (continued)

- If off-ladder allele is reproducible, it will be used for interpretation and designated as follows:
- Off-ladder alleles within the ladder
  - Designated with #repeats. bases of partial repeat
    - Example: at D21; a 29.2 has 29 repeats + 2 additional bases
- Off-ladder alleles outside the ladder
  - Alleles are shorter than the shortest allele in the ladder
  - Alleles are longer than the longest allele in the ladder
  - Designated as <shortest allele OR >largest allele
    - Example: at FGA; either a <18 or a >30 would be possible
- Allelic ladders for real off-ladder alleles will be required in the case file and must be from the same run as the sample

Microvariant



Negative Controls

Kit negatives, reagent blanks and LIZ standard

- Select negative controls in display plot
- Zoom the Y axis to approximately 100 rfus
- If the presence of contamination is determined in negative control samples, evaluate all samples. A contamination investigation must be initiated

Negative Controls

Kit negatives, reagent blanks and LIZ standard

IF...	THEN...
No alleles are present	Controls are acceptable, proceed
There are potential alleles above threshold	If peaks fall with a bin, and are not artifacts, the control is unacceptable. Assess if re-setup, re-injection is required
After re-injection potential alleles remain	Control is unacceptable. Notify supervisor and Re-setup all samples associated with that blank
After re-setup and re-injection potential alleles no longer remain	Controls are acceptable, proceed

Positive Controls

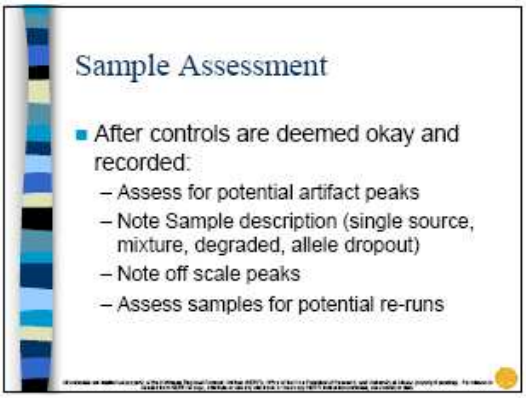
Amplification positives

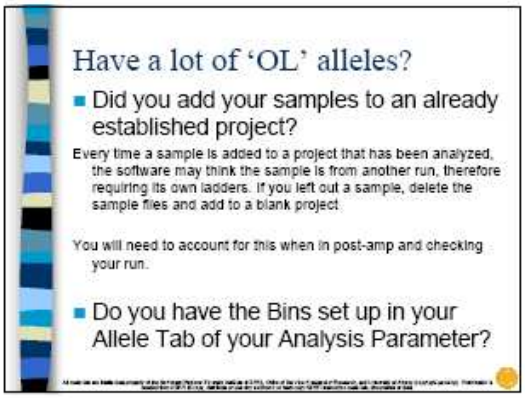
- The alleles of the Positive amplification control alleles should be checked against established values

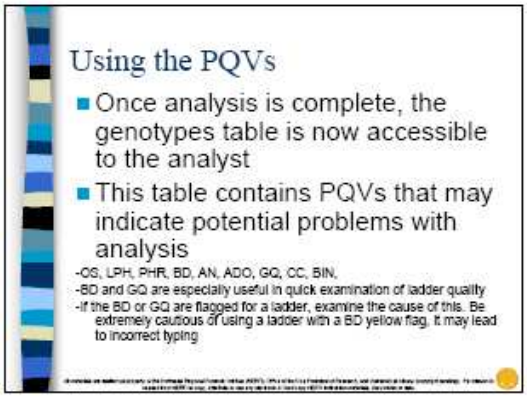
Positive Controls

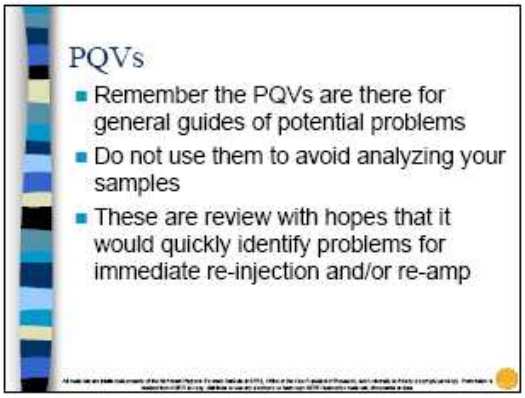
Extraction and Amplification positives

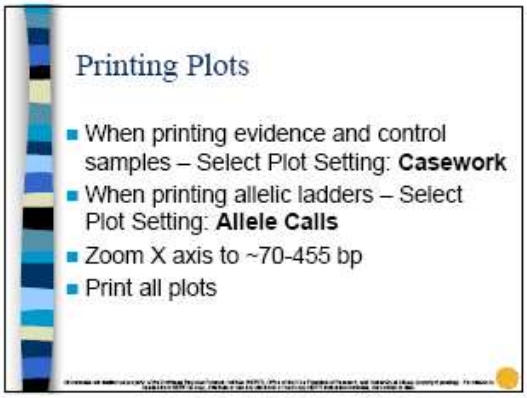
IF...	THEN...
Peak heights are above threshold rfus Fragment sizes are established values	Controls are acceptable, proceed
Peaks are below threshold rfus or improper allele calls are make	•Positive control peak signals are unacceptable –Repeat using longer injection time –The entire amplification will need to be repeated











Appendix 5: Capillary Electrophoresis Examination:

UNIVERSITY AT ALBANY
State University of New York

Northeast Regional Forensic Institute

Capillary Electrophoresis Workshop
Final Exam

Name _____ Date: _____

Answer the following short answer questions. (80 points total – 2 points each unless otherwise noted)

1. What is the name of the process that describes how the DNA enters the capillary?
    a. Sample Stacking
    b. Electrokinetic Injection
    c. Electro-osmotic Flow (EOF)
    d. None of the above

2. Which of the following helps keep DNA denatured:
    a. Urea
    b. 2-pyrrolinone
    c. Formamide
    d. All of the above

3. The migration of the DNA through the polymer is primarily based on fragment size since it has:
    a. Many A-T base pair sequences
    b. A constant charge-to-mass ratio
    c. A variable charge-to-mass ratio
    d. A high salt concentration in the sample solution

4. Which of the following should be performed at least on a weekly basis?
    a. Change the buffer
    b. Change the capillary
    c. Re-align the laser
    d. Change the polymer

5. Which of the following is performed on an "as needed" basis on the 3130xl:
    a. Spectral calibration
    b. Spatial calibration
    c. POP-4 change
    d. Buffer change

(10)

Page 1 of 9

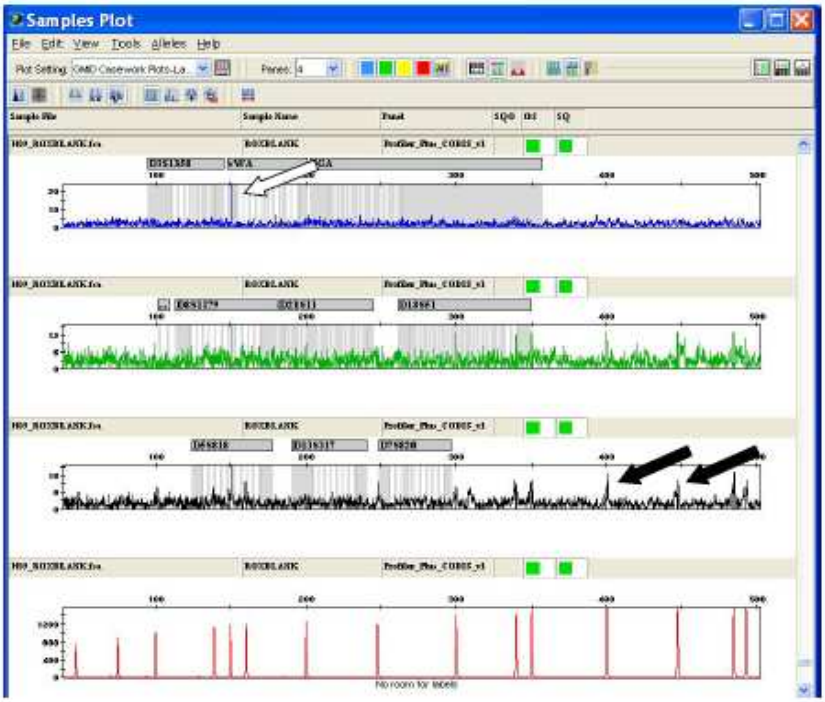
6. What are the components that are added to the master mix when loading Identifiler samples onto the CE?
    a. ROX and LIZ
    b. Formamide and LIZ
    c. Formamide and ROX
    d. Identifiler Ladder and Formamide

7. What is a function of the formamide during capillary electrophoresis?
    a. Keep the DNA denatured
    b. Provide an internal size standard for each of the samples
    c. Dilute the salts in the sample solution
    d. Both a and c

8. What are the DNA fragments within a sample compared to when determining their base pair sizes?
    a. ROX or LIZ
    b. Allelic ladder
    c. Formamide
    d. Buffer

9. What are the DNA fragments within a sample compared to when determining their allele designations?
    a. ROX or LIZ
    b. Allelic ladder
    c. Formamide
    d. Buffer

Questions 10 and 11 refer to the following electropherograms



10. What is the white arrow pointing to?
    a. Spike
    b. Pull-up
    c. PCR artifact
    d. –A

11. What are the black arrows pointing to?
    a. Spike
    b. Pull-up
    c. PCR artifact
    d. –A

(4)

12. The following can be used to help monitor the precision of a run:
    a. 150 bp ROX
    b. 200 bp ROX
    c. 300 bp ROX
    d. 250 bp ROX

13. Why is it important that the DNA remain in single strand conformation during electrophoresis?
    a. Better resolution
    b. Better interaction with the POP-4
    c. Less –A will occur
    d. Both a and b
    e. Both a and c

14. What term best describes the process where DNA fragments from the sample become collected in the beginning of the capillary in a sharp band, before being separated, due to the increased ionic strength of the buffer inside the capillary.
    a. Electrokinetic injection
    b. Sample Stacking
    c. Constant charge-to-mass ratio
    d. None of the above

15. What calibration on the 3130xl allows the instrument to map the area of each capillary onto the camera (in other words, it tells the camera from which capillary the fluorescent signal is coming from)?
    a. Matrix
    b. Spectral
    c. Autosampler calibration
    d. Spatial

16. What allows the CE instrument to correct for the overlap in the fluorescence emission spectra of the dyes?
    a. Matrix
    b. Spectral
    c. Spatial
    d. Both a and b
    e. Both a and c

17. What phenomenon could you see if the matrix or spectral calibration is not working properly?
    a. –A
    b. Stutter
    c. Peak shoulder
    d. Pull-up

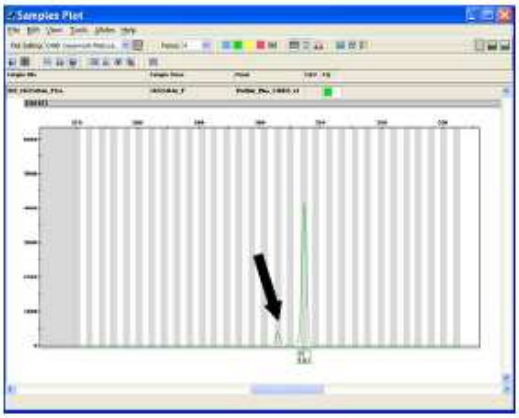
(12)

18. What sizing method does Genemapper® ID use to size peaks?
    a. Local Southern
    b. Global Southern
    c. Cubic Spline
    d. 2nd Order least squares

19. Using the following numbered steps, what is the correct order of events that occur on the capillary electrophoresis instrument?
    1 – Sample stacking into a precise, focused band
    2 - DNA fragments separated by size using the polymer in the capillary
    3 - Emissions of the fluorescent tags are captured by the CCD camera
    4 - Laser excitation at the capillary window
    5 - Sample injected into capillary

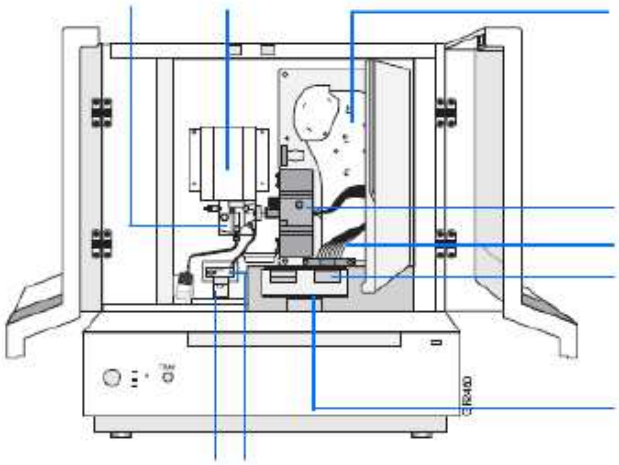
    a. 2-4-3-1-5
    b. 5-1-2-4-3
    c. 1-5-3-4-2
    d. 3-4-5-2-1

20. Review the following GeneMapper® ID image. What common PCR occurrence is the arrow pointing to?
    a. Stutter
    b. –A
    c. Dye artifact
    d. Spike



(6)

21. (18 points) On the diagram below of the 3130, identify the following parts:
    a. Detection cell block
    b. Buffer and water reservoirs
    c. Anode buffer reservoir
    d. Pump block
    e. Lower pump block
    f. Oven
    g. Autosampler
    h. Polymer delivery pump (PDP)
    i. Capillary array

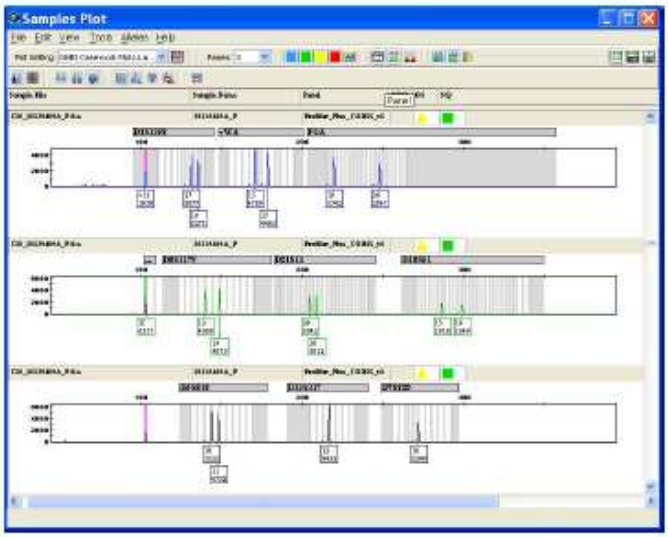


(18)

22. What factor is the electrokinetic injection dependent on?
    a. Ionic strength of the sample
    b. Ionic strength of the buffer within the capillary
    c. Both a and b
    d. Neither a nor b

23. What does the Local Southern Method use to determine the peak sizes in a DNA sample?
    a. 2 peaks above and 2 peaks below the fragment of interest
    b. 2 peaks above and 1 peak below the fragment of interest
    c. 1 peak above and 2 peaks below the fragment of interest
    d. 1 peak on either side of the fragment of interest

24. What is the most likely reason for the <12 allele present in the first blue locus on the electropherogram below?
    a. A true off-ladder allele
    b. –A
    c. Stutter
    d. Pull-up



(6)

25. Why is it important to have a low salt concentration in your sample solution?
    a. Less competition for the DNA molecules upon injection into the capillary
    b. To lessen the possibility of hypertension in the instrument
    c. It is important for sample stacking
    d. Both a and b
    e. Both a and c

26. Which of the following describe the capillaries used on the 3130xl for forensic purposes, as they are received from the manufacturer? Circle all that apply (4 pts).
    a. 47 cm in total length
    b. 36 cm from the inlet to the detection window
    c. Made of fused silica
    d. Coated on the inside
    e. Polyimide on the outside for added stability

27. What does the "4" stand for in POP-4?
    a. 4% concentration of polyacrylimide
    b. 4th modification of the formula
    c. Concentration of urea
    d. None of the above

28. DNA migrates toward the:
    a. Cathode
    b. Anode
    c. Negative electrode
    d. Both a and c

29. Which of the following describes the method of measuring the difference in the excitation and emission spectra of the fluorescent tags on the DNA fragments?
    a. Electrokinetic injection
    b. Sample Stacking
    c. Stokes Shift
    d. Multi-component analysis

30. What temperature is the electrophoresis run at?
    a. 50°C
    b. 55 °C
    c. 60 °C
    d. 65 °C

31. What data is along the left side of the electropherogram
    a. Relative fluorescent units
    b. Data points
    c. Scan points
    d. Base pair size (16)

BONUS QUESTION:

1. What does the acronym LASER stand for? (3)

Appendix 6: Capillary Electrophoresis Course Evaluations with Feedback

UNIVERSITY AT ALBANY
State University of New York

CAPILLARY ELECTROPHORESIS & DATA ANALYSIS
Jan. 11-15 & 8, 2010

Attendee Course Evaluation

Please fill out this evaluation form. Rate each of the aspects of this course by circling the appropriate number. Comment as appropriate.

Course Overall

	Poor				Excellent	Comments/Suggestions
A	1	2	3	4 (67%)	5 (33%)	Geared towards a less experienced crowd – could be translated for the audience. Examples and software time were great, we could have moved faster.

Speaker: Lucy Davis

	Poor				Excellent	Comments/Suggestions
B	1	2	3 (33%)	4 (33%)	5 (33%)	Lectures could be more focused.

Speaker: Jamie Belrose

	Poor				Excellent	Comments/Suggestions
D	1	2	3	4 (33%)	5 (66%)	Great pace and comfortable with material.

Quality of Audio-Visual

	Poor				Excellent	Comments/Suggestions
E	1	2	3	4 (66%)	5 (33%)	

Quality of Handouts

	Poor				Excellent	Comments/Suggestions
F	1	2	3	4 (50%)	5 (50%)	

Facilities

	Poor				Excellent	Comments/Suggestions
G	1	2	3	4 60%	5 40%	Was nice having it at our own lab. Did not take away from work for a week.

Please Comment On:

Did the material presented meet your professional expectations?
Yes - 6

What topic or topics did you find most useful to you?
Panels and settings, how they were set for lab already.
Practice with determining various artifacts
Theory / background lectures were great.
Trouble-shooting, intro to GMID-X
All of it
Theory and interpretation

Would you recommend this course to other scientists?
Absolutely
Yes - 5

General comments:
Thank you
Thanks for your time and expertise – the lectures were great!
I feel I really learned a lot. The class helped clarify a lot of things I "thought" I knew. I am no longer afraid of the software. Thank you!
The simulated data sets were very good for practice.

Suggestions for future classes:
PCR
Consider shortening to 3 or 3.5 days.

Thank you for your input!

UNIVERSITY AT ALBANY
State University of New York

CAPILLARY ELECTROPHORESIS & DATA ANALYSIS

Jan. 25-29, 2010

Attendee Course Evaluation

Please fill out this evaluation form. Rate each of the aspects of this course by circling the appropriate number. Comment as appropriate.

Course Overall

	Poor				Excellent	Comments/Suggestions
A	1	2	3	4 (22%)	5 (78%)	More hands-on with 3130xl; less time devoted to GMID A lot of very useful information, very well executed. Very informative and enjoyable. Flexible enough to accommodate different backgrounds. I learned more in this class than I expected to.

Speaker: Lucy Davis

	Poor				Excellent	Comments/Suggestions
B	1	2	3	4 (11%)	5 (89%)	Very good lecturer, made it very interesting. Builds good rapport with students, flexible and interactive. Very knowledgeable as well as friendly. Made course adjustments based on needs of the students. Presented the material well. Very knowledgeable and patient. Resourceful and very well spoken.

Speaker: Jamie Belrose

	Poor				Excellent	Comments/Suggestions
D	1	2	3 (11%)	4 (11%)	5 (78%)	Very well versed and knew and understood the chemistry. Really helped me understand the how's and why's of kits and instrument. Very knowledgeable of the course materials. Presented the material very well. Too fast. Top notch instructor and very helpful. Extremely knowledgeable and focused. Answered questions clearly and thoroughly. Very good instructor, able to explain complicated concepts very well.

Quality of Audio-Visual

	Poor				Excellent	Comments/Suggestions
E	1	2	3	4 (13%)	5 (88%)	The presentations were very well organized. I really enjoyed the notes, very thorough. Overheads were very helpful and aided in the presentation of material. Made it really easy to follow along.

Quality of Handouts

	Poor				Excellent	Comments/Suggestions
F	1	2	3	4 (13%)	5 (88%)	Made it really easy to follow along and to serve as a reference source.

Continued on back

Facilities

	Poor				Excellent	Comments/Suggestions
G	1	2	3	4 38%	5 63%	Easy access to all amenities. Campus was beautiful; I had enough space to work.

Please Comment On:

Did the material presented meet your professional expectations?

Yes: 3
Yes, it was a very good review of my training as well as more detailed in the areas I had issues with.
It exceeded my expectations. Unlike other workshops, there was a lot of explanation here and I'm walking out feeling like I know a lot more.
Personally, I expected more hands-on with the 3130xl. However, I did benefit greatly from the topics discussed.
Yes, very helpful: 2
Yes. It explained the overall process well and the instructors explained things completely and fully.

What topic or topics did you find most useful to you?

RT-PCR: 3
CE / 3130 / Electrophoresis theory: 5
GMID: 1
Explanation of local southern method and how software allows us to interpret data.
As much as I dislike theory, it was educational and interesting.
Quantifiler and qPCR lecture helped me a lot too.
Hands-on in lab and software.
Sample stacking and spikes.

Would you recommend this course to other scientists?

Yes: 5
Absolutely. If I was not employed I'd take the 4-month course.
Absolutely – even the most experienced analysts would learn so much valuable information.
Yes! Especially newbies that have limited experience. It fills in all the blank spots left over from site specific training.
Yes indeed.

General comments:

Very satisfied.
I appreciate the education opportunity that you provided for me. Thank you. Hope you find more grant money. I think this program provides a beneficial niche in the field of DNA.
Jamie and Lucy are both so knowledgeable and did a great job of explaining everything.
Thank you for sharing your wealth of knowledge.
Lectures on the CE and 3130 very helpful. qPCR lecture was a bonus.
I appreciate the opportunity to take part in the training – it was great.
I found the class to be very helpful. I feel more confident on the subject of CE. Thank you.

Suggestions for future classes:

Have it in the spring!
More hands-on time with 3130xl
Instruction and practice with GMID-IDX
Maybe a class on interpretation of complex mixtures.
I hope this course will be able to continue – very valuable to new analysts like myself + to experienced analsyst. The experienced analysts stated that they learned a lot of valuable information.
Maybe first day of computer work, Lucy or Jamie can have us following along with our laptops.
Do a sim set together and then separate.

Thank you for your input!

UNIVERSITY AT ALBANY
State University of New York

CAPILLARY ELECTROPHORESIS & DATA ANALYSIS

Feb. 8-12, 2010

Attendee Course Evaluation

Please fill out this evaluation form. Rate each of the aspects of this course by circling the appropriate number. Comment as appropriate.

Course Overall

	Poor				Excellent	Comments/Suggestions
A	1	2	3	4 (50%)	5 (50%)	Very specific and would highly recommend this course to forensic biologist. Very thorough and informative Very informative about theory/practices. In depth coverage and detail.

Speaker: Lucy Davis

	Poor				Excellent	Comments/Suggestions
B	1	2	3	4	5	NA – death in family

Speaker: Jamie Belrose

	Poor				Excellent	Comments/Suggestions
D	1	2	3	4 (50%)	5 (50%)	Clear and well spoken. Sufficiently answered all questions. Superb.

Quality of Audio-Visual

	Poor				Excellent	Comments/Suggestions
E	1	2	3 (13%)	4 (50%)	5 (38%)	Rich in information, very complete with detail specific to principles and applications; really enjoyed Jeopardy! Clear and new views of old material provided a nice change of pace.

Quality of Handouts

	Poor				Excellent	Comments/Suggestions
F	1	2	3 (25%)	4 (25%)	5 (50%)	Followed audio/visuals well. Small font on slides makes it hard to read on handout.

Facilities

	Poor				Excellent	Comments/Suggestions
G	1	2	3	4	5	

Please Comment On:

Did the material presented meet your professional expectations?
Exceeded expectations.
Yes, it touched up on many of our practices used in the lab.
Yes – 4
Yes, the information presented was very detailed and informative.

What topic or topics did you find most useful to you?
Electrophoresis and CE theory
The theory behind 3130 (capillary electrophoresis)
The detailed information that is more specific than Butler's book & more forensic application focused than ABI Quantifiler
Theory and concepts of instrumentation and procedures
All topics were helpful.
Theory, applications, and Jeopardy!

Would you recommend this course to other scientists?
Absolutely.
Yes - 6

General comments:
Good content. The Jeopardy game was fun and helped me to learn.
Thank you for the on-site visit.
Very clear. Informational and concise.
I would like to have this course again, at least for review. But I enjoyed this very much.

Suggestions for future classes:
None. This was very fun, interactive, and informative.

Thank you for your input!

Appendix 7: RT-PCR Course Information:

REAL-TIME PCR THEORY & CHEMISTRIES

GIVEN BY: NORTHEAST REGIONAL FORENSIC INSTITUTE
HOSTED BY: WESTCHESTER CO. CRIME LABORATORY
FUNDED BY: THE NATIONAL INSTITUTE OF JUSTICE
~ funding covers course instruction only

JANUARY 6 – 8, 2010; SPACES ARE LIMITED!

Instructors: *Bruce McCord, Ph.D., Lucy A. Davis, & Jamie L. Belrose*

COURSE DESCRIPTION

This 2.5 day course is designed to provide the theoretical and practical background necessary to perform quantitative PCR and data analysis. Today's advanced technologies have led to an exponential number of cases being submitted to crime labs for DNA testing. For these techniques to be successful, it is imperative that biological evidence be analyzed and interpreted effectively. *Those already using RT-PCR are welcome to attend the one-day advanced lecture by Dr. Bruce McCord.*

INTRODUCTORY LECTURE

WEDNESDAY, JANUARY 6, 12:30 – 4:00PM: NERFI

- Theory of Polymerase Chain Reaction (PCR)
- Why Quantitate DNA??? FBI QAS Standard 9.4 (effective: 7.1.2009)
- Several Methods of Quantitation (Yield Gels, Spectrophotometry, UV-Vis, & Sybr Green)
- How the amount of input DNA can/will impact PCR - adding too much or too little…
- Theory of Real-Time PCR
- Advantages of Quantitative PCR
- Fluorescent Detection (Stokes Shift)

THURSDAY, JANUARY 7, 8:00AM – 4:00PM: NERFI

- Kits from Applied Biosystems: Quantifiler™ Human, Quantifiler™ Y Human Male & Quantifiler® Duo
- Taqman® Probes & 5' Nuclease Activity
- Fluorescence Resonant Energy Transfer (FRET)
- Instrument Optics (Tungsten Lamp, Virtual Filters & Charged-Coupled Device Camera)
- Definitions: Threshold, Cycle Threshold, Internal PCR Control, Slope, Y-intercept, & Correlation Coefficient
- AB 7500 System Sequence Detection Software (SDS) v1.2.3 and HID Real-time PCR Analysis Software v1.0
- Absolute Quantification (Standard Curve)
- Troubleshooting outlier standards & changes in slope and Y-intercept
- Review of major concepts
- Multiple choice exam and issuance of completion certificates.

ADVANCED LECTURE

FRIDAY, JANUARY 8, 8:00AM - 4:00PM: MCCORD

- Multicopy methods for real time detection – Sybr Green, Plexor
- The über -male and other interesting issues with Y STRs
- How RT-PCR could be used to evaluate and decide whether to conduct autosomal vs. Y STRs.
- Cautionary notes
- Low copy number and real time PCR.
- Application of realtime PCR for setting laboratory thresholds
- Precision vs. Accuracy, variability with standards, and how labs might incorporate the use of SRMs.
- Validation issues
- Inhibition and mechanisms of how inhibition occurs (causative factors) and how to possibly overcome this dilemma.
- IPC Ct for detecting inhibition
- Efficiency/shape of the amplification plot for detecting inhibition
- Application of DNA melt curves for inhibition detection
- How RT-PCR could be used for screening samples for probative value.
- Methods to detect DNA degradation by real time PCR
- Short answer exam and issuance of completion certificates.

Appendix 8: RT-PCR Introductory Presentation

Real-Time PCR

NERFI / NIJ
Presented by: Jamie L. Belrose, Lucy A. Davis
Created by: NERFI Staff

1

Purpose of Quantitation

- Wide range of amounts of DNA at crime scene
- Inhibition of samples by either substrate or sample itself
- Narrow tolerance of PCR systems input DNA
  - 1 – 2 ng DNA typical Identifiler manual specifies
    - 0.5 – 1.25 ng
  - Too little → dropout of information, unequal amplification of target alleles
  - Too much → generation of artifacts, inhibition of alleles
- Is the DNA human?

2

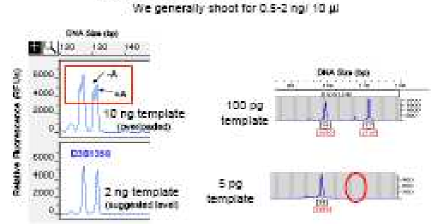
Purpose of Human-Specific DNA Quantitation

- All sources of DNA are extracted when biological evidence from a crime scene is processed to isolate the DNA present.
- Thus, non-human DNA such as bacterial, fungal, plant, or animal material may also be present in the total DNA recovered from the sample along with the relevant human DNA of interest.
- For this reason, the FBI QAS **Standard 9.3 (9.4 in 7/09 revised standards) requires human-specific DNA quantitation** so that appropriate levels of human DNA can be included in the subsequent PCR amplification.
- **Multiplex STR typing works best with a fairly narrow range of human DNA** – typically 0.5 to 1.25 ng of input DNA works best with commercial STR kits.

3

Impact of DNA Amount into PCR

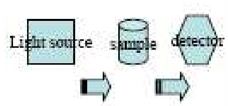
We generally shoot for 0.5-2 ng/ 10 µl



4

UV/Spectrophotometer

- Variable wavelength light source
- 260nm DNA
- 270nm phenol
- 280nm protein
- Absorbance detection



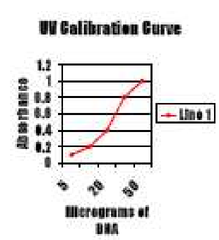
Can determine quality and quantity of DNA

Can't determine if human DNA

5

UV Methodology

- Run calibration standards
- Prepare calibration curve
- Measure unknown 260nm
- Compare to curve



6

UV Spectroscopy Ratios

- Abs 280/260 = $\geq$ 1.8 protein contamination
- Abs 270/260 = $\geq$ 1.2 phenol contamination

7

Advantages of UV Quantitation

- Quick
- Accurate if have clean solutions.
- Assessment of phenol or protein contamination

8

Disadvantages of UV Quantitation

- Needs lots of DNA at least 250ng.
- Not human specific
- No information on quality of DNA

9

Yield Gel Quantitation

- Agarose gel electrophoresis
- 1% agarose with ethidium bromide.
- 1x TAE buffer
- DNA phosphate backbone negatively charged

10

Yield Gel Apparatus

11

Yield Gel continued

- ~200V for 8 min applied across gel
- DNA travels toward positive anode
- Large fragments move slowest – sieving of agarose gel.
- Ethidium bromide intercallates with double stranded DNA.

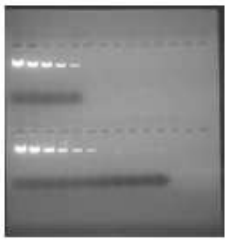
12

Yield Gel Detection

- The more molecules of DNA – the more Ethidium bromide you have associated with the DNA band.
- Compare intensity of band with standard bands
- High Molecular weight DNA is quantitated. Degraded DNA shows up as streak.
- Not human specific – viral and bacterial DNA looks the same as human
- Standards are Lambda DNA, not human DNA

13

Yield Gel Results



14

Yield Gel Advantages

- Quick and Easy!
- Low cost
- Consumes little sample
- Sensitive to approx. 1 ng of DNA
- Measures quality of DNA

15

Slot Blot Procedure

- Denature DNA
- Transfer DNA to charged nylon membrane
- Hybridize primate specific biotinylated probe
- Add HRP-SA conjugate
- Detection

16

Slot Blot Quantitation Hybridization Based Assay using Human Specific DNA probes

- QuantiBlot® Human DNA Quantitation System
  Detection methods:
  Colorimetric with biotinylated oligonucleotide labeled probes with streptavidin-horse radish peroxidase conjugate, color substrates
  Chemiluminescently with biotinylated oligonucleotide labeled probes with streptavidin-horse radish peroxidase conjugate, light-emitting substrate

17

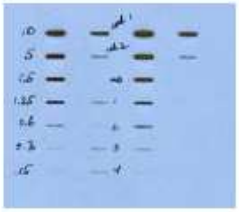
Slot Blot Apparatus

18

Membrane Development

19

Slot Blot Results



20

Slot Blot Advantages

- Consumes little sample
- Works on single stranded and double stranded DNA
- Human specific
- Accurate with degraded DNA to a point

21

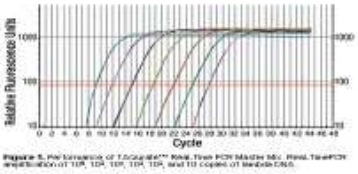
Disadvantages of Slot Blot Detection

- Time consuming
- No indication of DNA quality
- Expensive (cheaper than real time PCR)
- Technique dependent

22

Real Time PCR Quantitation

qPCR



23

Real Time PCR Advantages

- Automated
- Very sensitive and wide dynamic range
  0.023 ng to 50 ng/µl
- Evaluates inhibition
- Very reliable and robust
- Minimal sample handling and setup

24

History

- RT-PCR (qPCR)
  - Developed by Higuchi in 1993
  - Used a modified thermal cycler with a UV detector and a CCD camera
  - Ethidium bromide was used as intercalating reporter → as [dsDNA] increased fluorescence increased
- First paper on qPCR:
  - Higuchi, R.; Fockler, C.; Dollinger, G.; Watson, R. "Kinetic PCR analysis: real-time monitoring of DNA amplification reactions" Biotechnology (NY). 1993 Sep;11(9):1026-30
- Warning: RT-PCR also means reverse transcriptase PCR which is used when working with RNA. So real-time PCR is often abbreviated qPCR (quantitative)

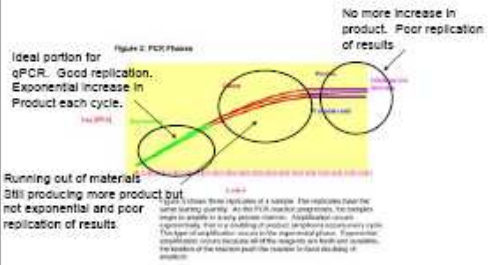
25

PCR Amplification

- What is PCR?
  - Polymerase Chain Reaction
    - Uses template DNA, oligonucleotide primers, deoxynucleotide triphosphates (dNTP's), & thermostable DNA polymerase to replicate selected regions of DNA
    - Exponential amplification (doubling with each cycle)

26

PCR Phases



27

Real-Time Quantitative (qPCR)

- What is real-time PCR?
  - qPCR works the same way as conventional PCR with the addition of fluorescent probes to quantitate DNA concentration.
  - Accumulation of product can be monitored with every cycle = "real-time".
- How can we measure the concentration of DNA in a sample with fluorescence?
  - To answer this question, we must know some of the fundamental concepts involved with fluorescence and its detection.

28

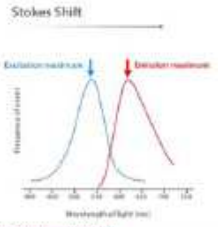
Detecting Fluorescence

- Four required components:
  - 1) Excitation Source –
    - Tungsten-halogen lamp in Applied Biosystems 7000/7500 real-time instruments
    - Argon laser in Applied Biosystems 7900 real time instrument
  - 2) Fluorophore (fluorescent molecule)
    - Fluorescent dyes from Applied Biosystems kits
  - 3) Wavelength filters
    - Separate emission from excitation photons
    - Extremely important for correct detection levels of the fluorophore of interest
  - 4) Detector
    - CCD (charged coupled device) camera detects signals proportional to intensity. Extremely sensitive, capture 70% of incident light as opposed to 2% with photographic camera
    - Detects emission photons
    - Records output as a signal value or photograph

29

Fluorescence

- Light of a certain wavelength excites fluorophores
- Emission of fluorescent light occurs at another, higher (longer) wavelength
- If different wavelengths of light are used to excite, intensity of light emitted will change



http://

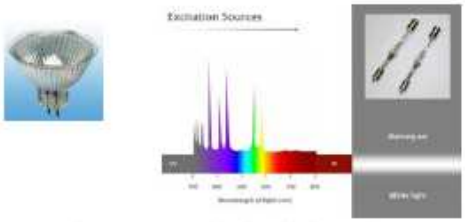
30

Stokes Shift

Different dyes excite at same wavelength but emit at different wavelengths allowing us to distinguish them

31

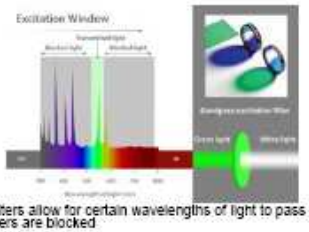
Excitation Source



- Tungsten-halogen lamp emits white light with peaks at varying intensity across the electromagnetic spectrum
- Filters needed to isolate the wavelengths of interest

32

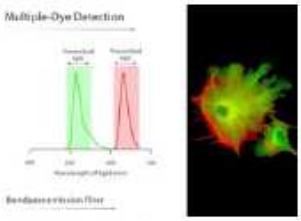
Filters



- Optical filters allow for certain wavelengths of light to pass through while others are blocked
- Several bandpass emission filters are used in real-time quantitation because of the multiple dye system used

33

Multiple-Dye Detection



- Although the graphic is of cell staining, the same principle is applied to fluorescently labeling the DNA products from qPCR

34

Forensic Kits Available

- Applied Biosystems
  - Quantifiler™
  - Quantifiler™ Y
  - Quantifiler Duo ™
- Promega
  - Plexor HY ™

Quantifiler™ Human DNA Quantitation Kit

- Amplicon is 62 bp
- Probe is to Human telomerase reverse transcriptase gene (hTERT) chromosome 5
- Nontranslated region (intron)
- Diploid

36

Applied Biosystems Quantifiler™ Kits Assay: Two Major Components

- Human Target Specific Assay
  - FAM labeled probe (autosomal or Y-chromosome)
- Internal PCR Control (IPC)
  - VIC labeled probe
  - 10,000 copies/rxn of synthetic non-human template
  - Inhibitor Detection

37

Dyes used in Quantifiler™

- FAM – blue human DNA reporter dye
- VIC – green IPC reporter dye
- ROX – red passive reference dye
  - Used in all samples – minimizes pipetting variation and well to well optical differences.
  - Color represented on screen is not necessarily same as dye - can set plots to any color.

38

Passive Reference Dye

- The Quantifiler kit also includes a passive reference dye – ROX.
- The information gained from including this dye is used to account for the instrument's shortcomings.
- The laser excitation across the 96-well plate is not equal.
- This is due to the plate's square shape and bulb's round shape. The wells in the middle of the plate receive more excitation energy than those around the outside.
- The ROX information is used to normalize the data for all wells of the plate.

39

Quantifiler™ Kits Use TaqMan® Probe Technology

- Primers used to amplify a particular region
- Fluorescently labeled probe binds to target
- When target is amplified, probe is hydrolyzed, releasing the fluorescent tag
- Instrument measures the cycle to cycle changes in fluorescent signal
- Quantity of fluorescence measured is directly related to target DNA concentration

40

TaqMan® Probe

- The Quantifiler™ probe is autosomal specific and is complimentary for a single copy target within the genome
- The target region is an intron within the human telomerase reverse transcriptase gene (hTERT) on chromosome 5 between the primers
- 5' end of the probe contains a fluorescent dye – FAM, known as the reporter dye
- 3' end of the probe contains a non-fluorescent dye, known as the quencher dye
- 3'end also contains a minor groove binder

41

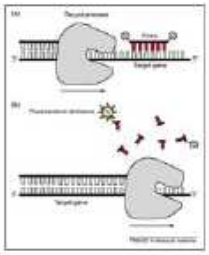
How TaqMan® Probe Works

- The probe has a fluorescent dye attached to the 5' end (reporter) and a non-fluorescent dye attached to the 3' end (quencher)
- While the reporter and quencher dyes are in close proximity to one another there is an energy transfer from the reporter (high energy) to the quencher (low energy), resulting in no fluorescence
- This phenomenon is called Fluorescent Resonance Energy Transfer or FRET

42

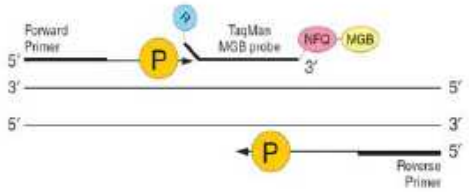
How TaqMan® Probe Works

- AmpliTaq Gold DNA Polymerase has 5' nuclease activity. The 5' nuclease activity of the enzyme acts upon the surface of the template to remove obstacles downstream of the growing amplicon that may interfere with its generation.
- This coupled with the FRET (Fluorescent Resonance Energy Transfer) makes it possible to detect PCR amplification in real-time.



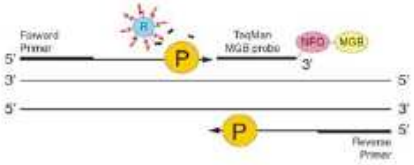
43

How TaqMan® Probe Works



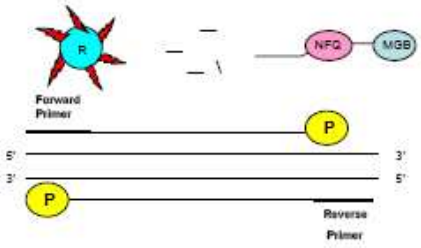
44

How TaqMan® Probe Works



45

How TaqMan® Probe Works



46

TaqMan Probe® Minor Groove Binder

- The Quantifiler™ probe is very short (12-18bp). This allows the PCR amplicon to be very small (62bp). Small amplicons mean a more efficient PCR and a greater chance of quant'ing degraded DNA.
- The Minor Groove Binder has two additional roles:
  - First, it anchors the small probe to the DNA; it does this by binding in the "minor" groove.
  - Second, it increases the annealing temperature of the probe. The probe needs to have a higher annealing temperature than the primers, otherwise there would be annealing and extension before the probe got a chance to sit down.

47

Real-Time PCR Instrument Fluorescence Detection

- The instrument directs light into each well using:
  - Applied Biosystems Prism 7000/7500 – tungsten-halogen lamp
  - Applied Biosystems Prism 7900 – argon laser
- The light passes through optical adhesive cover (or strip) on plate and excites fluorescent dyes in each well.
- A system of lenses, filters, and a dichroic mirror focuses the fluorescence into a charge-coupled device (CCD) camera
  - Fluorescence emission between 500nm and 660nm is separated based on wavelength into a predictably spaced pattern across the CCD camera.
- The instrument contains SDS software that applies algorithms to the fluorescence detected from the CCD camera to determine relative fluorescence units (RFU)

48

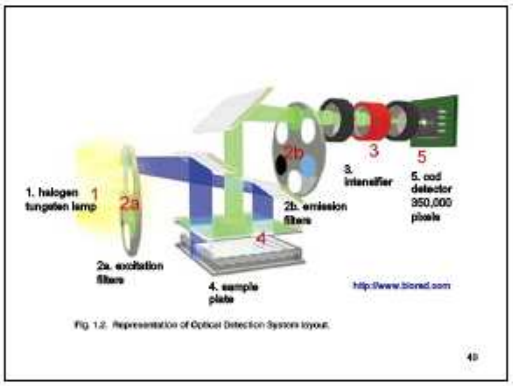


Fig. 1.2. Representation of Optical Detection System layout.

49

Summary of TaqMan® Probe and Quantitation

- Start with a sample containing an unknown amount of DNA
- Quantifiler™ kit contains primers and probe that will target and attach to the hTERT gene
- Through the 5' nuclease activity of the polymerase and FRET, the amount of DNA template will be detected by the instrument based on the amount of fluorescent emission after each PCR doubling of the DNA
- Using software and the standard curve, the fluorescent emissions are converted to give a quantity of DNA

50

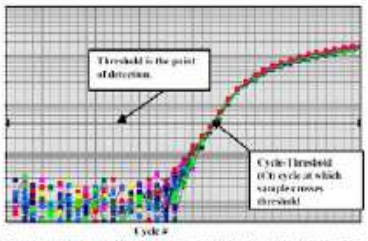
Cycle Threshold

Cycle Threshold (Ct)

- The **Threshold** is the level of detection or the point at which a reaction reaches a fluorescent intensity above what is considered to be background.
  - The Threshold is set at 0.2 as per Applied Biosystems
  - The Threshold line is set in the exponential phase of the amplification for the most accurate reading
- The cycle number at which the samples reaches this threshold level is called the **Cycle Threshold (Ct)**.

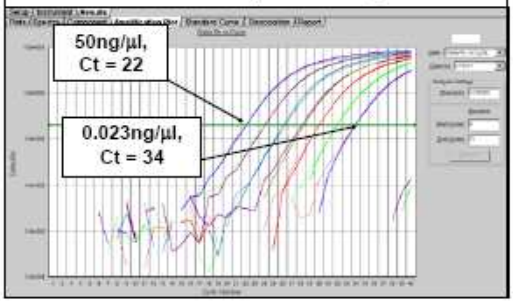
51

What is C_T?



52

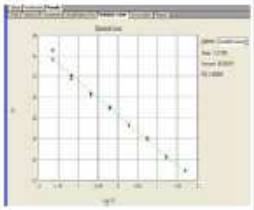
The smaller the Ct… the more initial template DNA you have



How do we get an actual DNA concentration?

Standard Curve:

- Several known concentrations of DNA (standards) are run on the 96-well plate along with your samples and the data is used to generate a standard curve (Ct vs. log concentration).
- The fluorescence measured from the unknown samples is then compared to this standard curve and the initial concentration is interpolated.



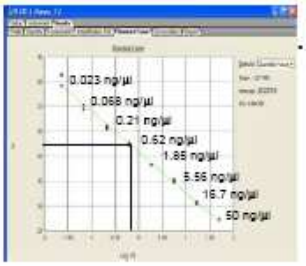
54

Standard Curve

- Made using purified human, male DNA
  - Known concentration of 200 ng/µl
- Standard is serially diluted (3-fold difference between each concentration) to give 8 concentrations ranging from 50 ng to 23 pg per µl
- Duplicates of each dilution of standard are added to analysis plate
- The Ct value for the different concentrations of standard is determined based on the number of cycles it takes for the fluorescence to reach the threshold

55

Standard Curve



- A standard curve is formed by plotting the log of the known concentration of standards against the Ct value for that concentration

  Example:

  0.62 ng of DNA would have a Ct value of 29

56

Using standard curve to determine unknown concentrations

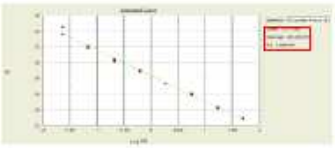
- Standards:
  - We input concentration of each dilution into the Sequence Detection Software (SDS)
  - Ct value for each dilution is measured by the instrument when cycle is reached where fluorescence crosses the threshold
- Unknowns:
  - Ct value for each unknown is measured by the instrument
  - Concentration is determined based on data inputted for standards:
    - Software compares Ct values of samples to Ct values of standards
    - Software takes input data of standard concentrations to determine concentrations of samples → quantities of DNA

57

Examining the Standard Curve

$C_T = m[\log(Qty)] + b$

- Why use log? Answer: to obtain a straight line
- **m** is the slope
- **b** is the y intercept
- **Qty** is the quantity of starting DNA



58

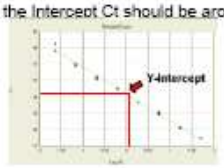
Examining the Curve

- Slope = steepness of the line
- Slope value (m) indicates amplification efficiency for the assay
  - -3.32 indicates 100% efficiency
  - Range of -2.9 – -3.3 considered acceptable for Quantifiler kit

59

Examining the Curve

- **Y-Intercept** indicates the expected Ct value for a sample with a quantity of 1ng/µl
  - because logConc = log (1) = 0
- For Quantifiler the Intercept Ct should be around 29.



60

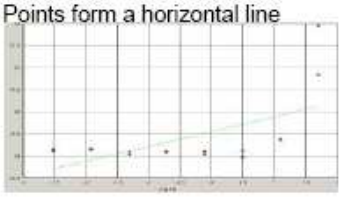
Examining the Curve

- R^2 should be ≥ 0.99 (correlation coefficient)
  - Value of 1 means curve perfectly fits data points
  - < 0.98 indicates a problem:
    - Task and quantity entered for wrong detector
    - Incorrect quantity entered for standards
    - Pipetting trouble
    - Other issues?
    - If you see this, don't worry we will figure it out!

62

Resulting Curve 1

- Each point is outside the expected region of the standard curve
- Points form a horizontal line



63

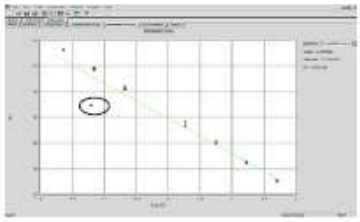
Possible Problem: Wrong Detector

- Task and quantity entered for IPC detector rather than Quantifiler™ detector
  - The Quantifiler™ Human detector's Task should be "Standard" NOT the IPC's Task

64

Resulting Curve 2

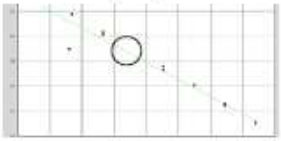
- One or more points lie(s) outside the standard curve



65

Possible Problem: Incorrect Quantity Entered

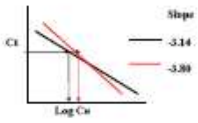
- In the diagram on the previous page, 0.062 was entered as the value for the outlying standard
- The correct value is 0.62 (circled area would be expected area for point)



66

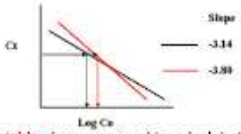
Change in Slope

- An acceptable slope of -3.14
- An unacceptable slope of -3.80
- Looking at the same Ct value for a theoretical sample – let's say 29 – a different quantity of DNA is calculated based on the different slopes in the standard curves
- What does this mean in terms of our data?



67

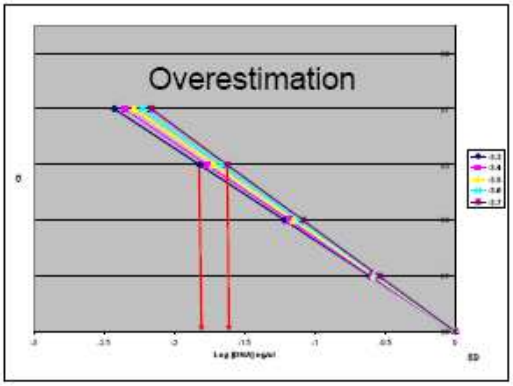
Downstream Issues



- If the unacceptable slope was used to calculate the amount of sample needed for amplification, the quantity would be overestimated to the left of the y-intercept (underestimated to the right of the y-intercept)
- Therefore, less sample and thus less DNA would be added for PCR amplification and allelic dropout is a possibility (sample overload if point to the right of the y-intercept → pull-up in profile)

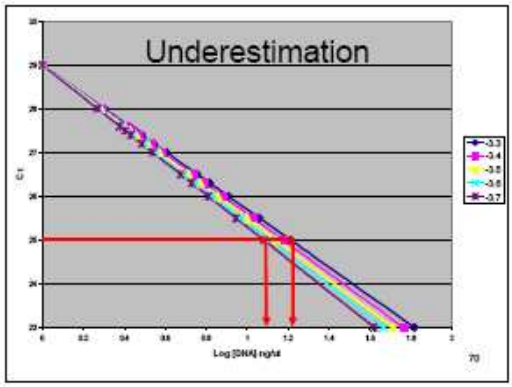
68

Overestimation



69

Underestimation



70

Resulting Quant values

Calculated DNA concentrations for various Ct and slope values holding Y=29
Ct= M*log[DNA] + Y Ct=cycle threshold, Y is intercept

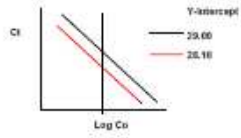
Y=29	Slope Values				
	-3.30	-3.40	-3.50	-3.60	-3.70
Ct values					
23	65.79	55.17	51.79	46.42	41.84
25	16.30	15.01	13.88	12.88	12.00
25.5	11.50	10.70	10.00	9.38	8.82
26	8.11	7.62	7.20	6.81	6.47
26.3	6.66	6.22	5.81	5.62	5.37
26.5	5.72	5.44	5.16	4.95	4.74
27	4.04	3.87	3.73	3.59	3.47
27.3	3.51	3.38	3.27	3.16	3.07
27.4	3.35	2.99	2.87	2.78	2.71
27.5	2.85	2.75	2.66	2.61	2.54
27.6	2.66	2.56	2.51	2.45	2.39
28	2.01	1.97	1.93	1.90	1.86
29	1.00	1.00	1.00	1.00	1.00
31	0.25	0.26	0.27	0.28	0.29
33	0.06	0.07	0.07	0.08	0.08
35	0.02	0.02	0.02	0.02	0.02
37	0.00	0.00	0.01	0.01	0.01

Calculated DNA values in ng/µl for each slope and Ct value

71

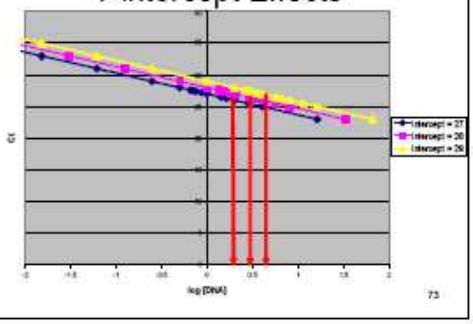
Change in Y-Intercept

- If the slope of the standard curve remains the same, but the y-intercept changes… (both values below are acceptable)
- The whole curve shifts
- The same Ct value leads to a different concentration calculated
- Can lead to the same downstream errors as poor slope



72

Y intercept Effects



73

Out of Range Slope Causes

- Generally slope issues are a result of inaccurate pipetting
  - Could be inaccurate pipetting
  - Could be inaccurate pipettor
  - Could be both!
  - Generally a problem with preparation of calibration standards
- Less likely – something wrong with kit

74

Y intercept out of range causes

- Could be DNA standard is not at stated concentration
- Could be inaccurate pipetting
- Could be inaccurate pipettor
- Generally a problem with calibration standards preparation
- It is a good practice to establish an acceptable range for Y intercept

75

Putting It All Together

- Let's go back and talk about the kit

76

Quantifiler™ Kit Components

- Primer/Probe Mix with Internal PCR Control (IPC)
  - PCR primers – one set for hTERT, one set for IPC sequence
  - FAM (Quantifiler™) and VIC (IPC) labeled TaqMan® probes
- TaqMan® 2X Universal PCR Master Mix
  - PCR reagents
  - ROX-passive reference dye
- DNA standard
  - Human male DNA, 200ng/µl

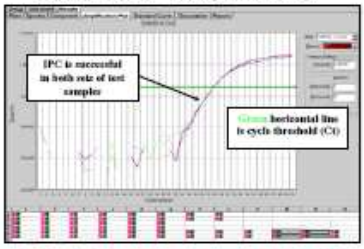
77

Internal PCR Control (IPC)

- In the primer mix
- Synthesized DNA strand – sequence not found in nature
- Detects inhibition, instrument malfunction, or problems with assay setup
- Warns of impending PCR and STR analysis failure

78

Internal PCR Control



Both IPC assays were successful

79

Internal PCR Control

Amplification failure

80

Summary of IPC Results

- The failure of both the samples and IPC to amplify most likely means an inhibitor was present in either the sample itself or in the material from which the sample was taken.
- Failure to load the necessary reagents (i.e. "Master Mix) would also lead to these irregular lines which do not cross the Ct.
- Additionally, an IPC failure could also indicate an instrumentation or assay setup error or the presence of no DNA combined with an amplification error in the IPC.
- As you are quickly learning, a failed IPC is not necessarily indicative of just one issue and a sample with a failed IPC should be diluted and re-quantitated.

81

Quantifiler™ Kits Procedure

- 12.5µl TaqMan® 2X Universal PCR Reaction Mix
- 10.5µl Primer/Probe Mix
- 2µl template

82

Quantifiler™ Kits Procedure

- Universal Cycling Parameters
  - 95°C 10 minutes
  - 40 cycles of
    - 95°C 15 seconds
    - 60°C 60 seconds
- ABI PRISM™ 7500 SDS

83

Quantifiler™ Y Human Male Quantification Kit

- Amplicon 64 bp
- Sex-determining region Y gene (SRY)
- Nontranslated region
- Haploid

84

Quantifiler™ Y Human Male DNA Quantification Kit

Y-chromosome kit useful tool for samples suspected of containing male DNA

- Sexual assault samples
- Crime scene mixtures
- Provides additional information for mixture interpretation
- Provides quant info for Y STR typing

http://

85

Quantifiler™ Y Human Male DNA Quantification Kit Conclusions

- Sensitive
  - Proof of male DNA present on intimate sample
  - Offers novel method of mixture detection

86

Calibrations

- Why are calibrations of the instrument necessary?
  - Because we are processing multicomponent (many dyes) data.
- Background calibration
  - Determines the amount of background fluorescence by measuring the raw spectra generated from a plate with PCR buffer.
    - Background electronic signal
    - Contaminants in the sample block
    - Plastic consumables…e.g. plates and caps
- Pure dye calibration
  - Collects and stores spectral data from pure dye standards
- Regions of Interest (ROI)
  - Maps the position of wells on the sample block so the software can trace increases in fluorescence to specific individual wells.

87

Calibrations

- Raw data is collected between 500 and 660nm.
- During data collection, the SDS software uses algorithms to determine the contribution of each dye.
  - The background component (stored in the BG calibration file) is subtracted from the raw data.
  - The information in the pure dye file is used to determine the contribution of each dye to the measured spectrum.
    - Measured spectrum = a(FAM)+b(VIC)+c(ROX)+d(BG)+MSE
      - *MSE = Mean Standard Error*

88

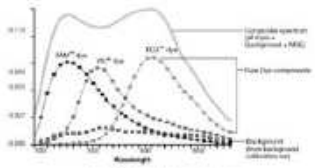
Calibrations

- The MSE indicates how closely the collective multicomponent spectrum conforms to the raw data.
  - Similar to R^2 on the standard curve.
  - Are the data be collecting along the lines of what one would expect?

89

Calibrations

Composite Spectrum



This is a composite spectrum from a single well.

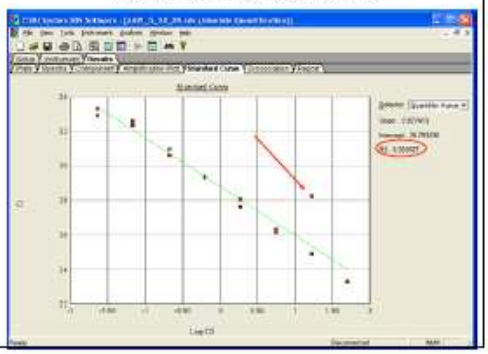
90

Calibrations

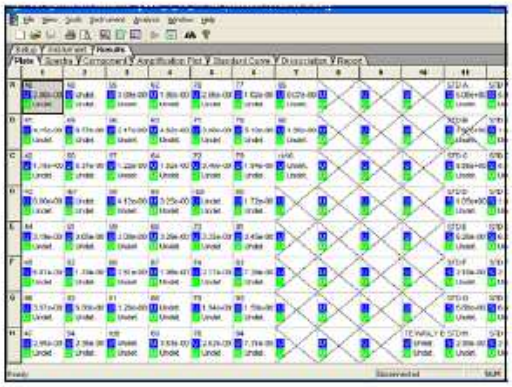
Normalization of Reporter Signals

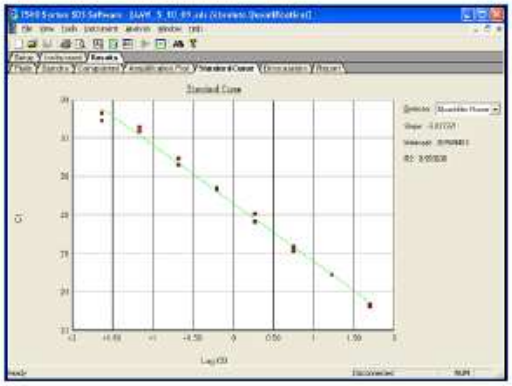
- The software normalizes each reporter signal by dividing it by the fluorescent signal of the passive reference dye, ROX.
- ROX is present at the same concentration in all wells as part of the master mix.
- By normalizing, the software can account for minor variations in signal caused by pipetting inaccuracies – thus allowing for more accurate well-to-well comparisons.

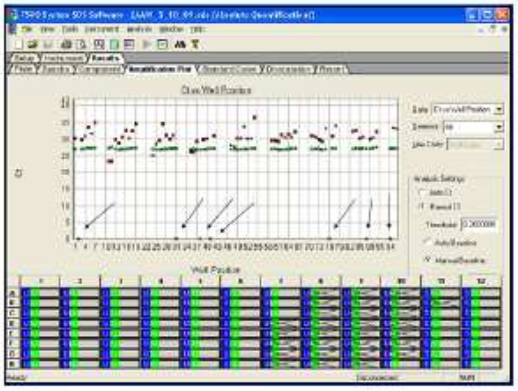
Curve Associated with Bad standard in Well B11

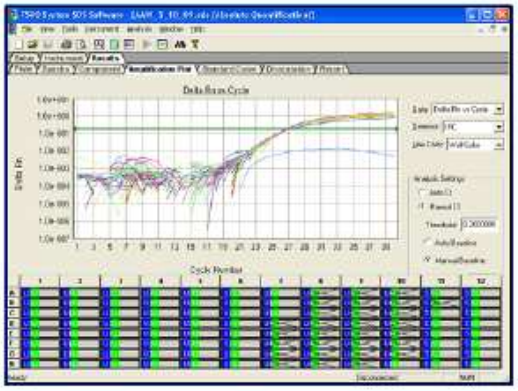


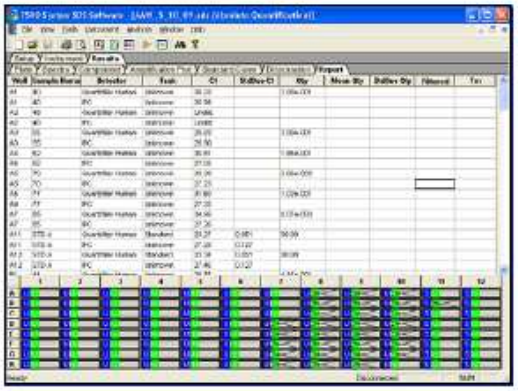
Bad Standard in well B11 Example

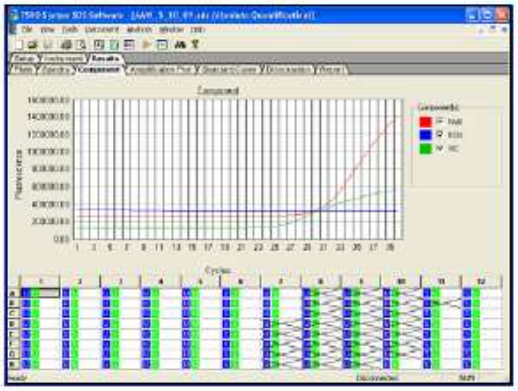




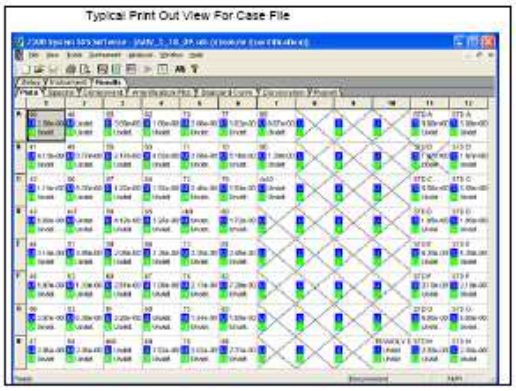




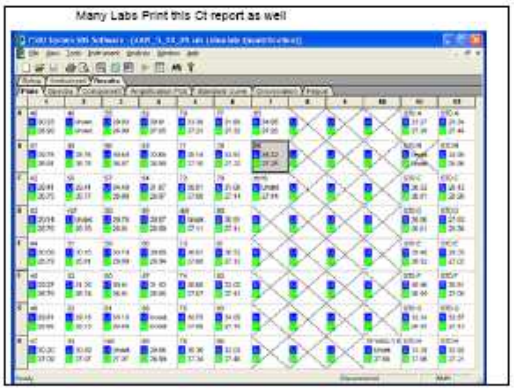


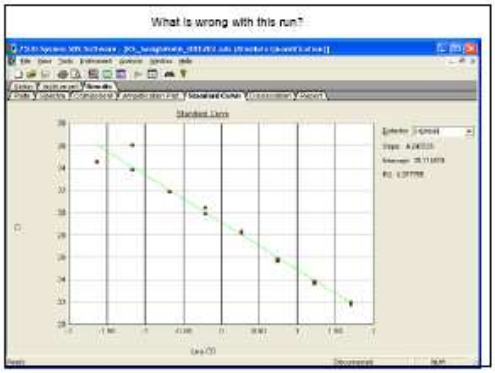


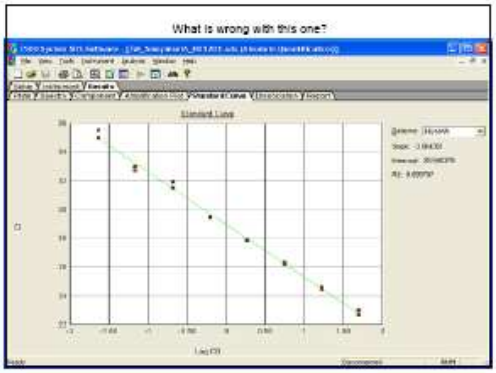
Typical Print Out View For Case File



Many Labs Print this Ct report as well







References

- Butler, John M. "Forensic DNA Typing" 2nd Edition. Elsevier Academic Press, 2005.
- Applied Biosystems training June 6-8, 2006.
- AB Quantifiler Kits User's Manual
- www.appliedbiosystems.com
- Information also gathered from Power Point slides created by Jamie Belrose, Lucy Davis, Andy Wist, Jacki Higgins, Kristin Shelling, and Margaret Terrill.
- Photos by Andy ☺

Appendix 9: Advanced RT-PCR Presentations

A comparison between Plexor and Quantifiler Duo

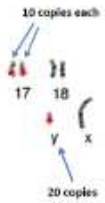
Heather LaSalle, George Duncan, Bruce McCord,
FIU Department of Chemistry
International Forensic Research Institute
Miami, FL 33199
mccordb@fiu.edu 305-348-7543

Introduction

- Quantifiler Duo (ABI) and Plexor HY(Promega) are two approaches to the analysis of DNA mixtures
- Both are qPCR kits that determine the relative difference between male and female DNA
- Quantifiler duo is a Taqman based assay that contains a Y probe, an autosomal probe and a control
- Plexor is a multilocus primer quenching based assay that also contains a Y probe, an autosomal probe and a control.
- The main differences between the two result from copy number, allele size and the issue of probes vs primers for defining specificity.

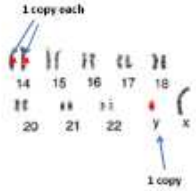
Plexor HY Setup

- Plexor probes multi copy variant CNV (n=10) at locus RNU2 on chromosome 17 for autosomal DNA. Amplicon size = 99bp fluorescein. Detect at 520 nm)
- Because there are 2 copies of the autosomal locus, there is effective equivalence between the number of copies of the autosomal and the Y locus in Plexor./
- Plexor also probes a multicopy variant CNV (n=20) at locus TSPY Amplicon size = 133bp . Detect at 560 nm
- An internal positive control exists at 610 bp. Amplicon size is 150 bp. Detect at 610 nm
- There is a internal reference IC5



Duo Setup

- Duo probes a single copy gene . The target is located in the RPPH1 gene (Ribonuclease P RNA component H1) on chromosome 14 (location: 14q11.2)
- Duo also probes a the Y chromosome at SRY locus with a target length of 130 bp. Detect at 520nm location: Yp11.).
- The internal positive control also has a length of 130 bp. Detect at 580 nm
- Rox is added as a passive reference 610 nm
- The common lengths permits a consistant response to inhibition and a reasonable correlation to a typical minimum sized STR.



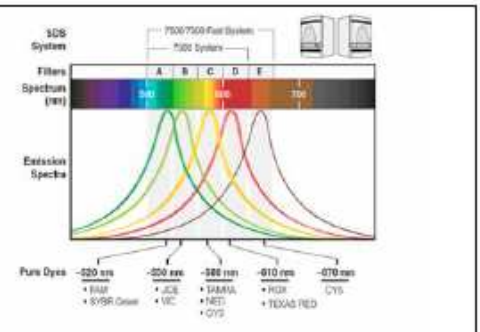
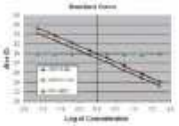


Figure 1-7 Example of an emission spectrum

Plexor users manual

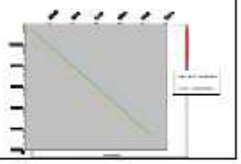
Duo Calibration Curve

- 50 ng to 23pg single copy



Plexor HY Calibration Curve

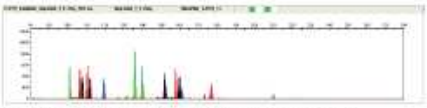
- 50 ng to 3pg multi copy



Approximate Sample Ratio (male:female)	Expected Total Human (ng/µL)	Quantifiler® Duo Total Human (ng/µL)	Expected Male (ng/µL)	Quantifiler® Duo Male (ng/µL)
1:100	12.8	9.60	0.128	0.145
1:50	11.3	9.05	0.226	0.225
1:25	8.5	6.84	0.340	0.321
1:10	7.0	5.61	0.699	0.625
1:10	1.0	0.850	0.091	0.085
1:5	1.0	0.760	0.167	0.152
1:1	1.0	0.790	0.395	0.401

Table 1. Mixture study results showing strong correlation between expected and obtained quantification values for male and total human DNA using the Quantifiler® Duo Kit.

Shelly Steadman and Sarah Geering
Sedgewick County Regional Forensic Science Center



Amplification results for a sample degraded with Dnase 1 that was quantified using Duo. Both Identifiler and miniflier were used. The data show that in spite of the degradation of larger sized alleles, lower alleles amplified properly with Identifiler (1ng template) Larger alleles were recovered with Miniflier (0.25ng template

An Analysis of Single and Multicopy Methods for DNA Analysis by Real Time PCR

Heather LaSalle
Broward Sheriff's Office

Introduction

- The goal of this study was to compare results between Quantifiler duo and Plexor HY.
- The hypothesis is that Plexor HY, a multicopy system should produce improved precision at lower levels of template
- However, it was also expected that the multicopy Y estimate for Plexor should be less precise than the single copy Quantifiler duo due to potential mutations affecting copy number in the genome.

Experimental design

- Initial experiments involved a cross comparison of external standards for each kit.
- Each kit was calibrated with its specific standards and then these samples were run again as unknowns to test reproducibility of the measurement.
- In addition, these standards were run with both chemistries.

Preparation of Standards

- A set of standards is made by serial dilution and run in duplicate
- The Plexor standards are made from a 50ng/ul pooled standard. The range is 50ng/ul, 10ng/ul, 2ng/ul, 0.4ng/ul, 0.08ng/ul, 0.016ng/ul and 0.0032ng/ul.
- The Duo standards are made from a 200ng/ul pooled standard. The range is 50ng/ul, 16.7ng/ul, 5.56ng/ul, 1.85ng/ul, 0.62ng/yl, 0.21ng/ul, 0.068ng/ul and 0.023ng/ul.
- Both the Plexor standards and Duo standards were run as unknowns and standards on multiple runs.

Precision of Plexor Standards

Plexor Standards as Unknowns of Plexor	Plate 1 Avg	Plate 2 Avg	Plate 3 Avg	Average of all 3	Standard Deviation
50 Human Standard	47	48	48	49.8	0.7
50 Human Standard Y	50.87	51	50	45.8	9.8
10 Human Standard	8.7	8.63	10.23	9.28	0.92
10 Human Standard Y	8.4	10.77	11.33	10.6	2.2
2 Human Standard	2.13	2.63	2.1	2.08	0.05
2 Human Standard Y	2.1	2.45	2.6	2.58	0.86
0.4 Human Standard	0.46	0.38	0.51	0.45	0.11
0.4 Human Standard Y	0.40	0.29	0.53	0.41	0.15
0.08 Human Standard	0.084	0.15	0.03	0.108	0.065
0.08 Human Standard Y	0.076	0.031	0.03	0.085	0.020
0.016 Human Standard	0.029	0.027	0.075	0.023	0.011
0.016 Human Standard Y	0.010	0.0107	0.018	0.01580	0.00564
0.0032 Human Standard	0.0036	0.0065	0.0026	0.0027	0.0026
0.0032 Human Standard Y	0.0048	0.0037	0.0022	0.0034	0.0017

Plexor Standards were run in triplicate on each plate

Precision of Plexor Standards

Plexor Standards as Unknowns on Duo	Plate 1	Plate 2	Average	Standard Deviation
50 Human Standard	49.5	51.5	50.5	1.4
50 Human Standard Y	51.9	54.6	53.3	1.9
10 Human Standard	8.02	8.14	8.08	0.085
10 Human Standard Y	8.20	8.82	8.51	0.42
2 Human Standard	1.5	3.1	2.3	1.1
2 Human Standard Y	1.5	3.1	2.3	1.1
0.4 Human Standard	0.29	0.60	0.46	0.26
0.4 Human Standard Y	0.35	0.59	0.47	0.16
0.08 Human Standard	0.083	0.101	0.110	0.012
0.08 Human Standard Y	0.083	0.220	0.160	0.012
0.016 Human Standard	0.0220	0.016	0.018	0.0034
0.016 Human Standard Y	0.015	0.030	0.029	0.011
0.0032 Human Standard	UND	0.004	0.002	N/A
0.0032 Human Standard Y	UND	UND	UND	N/A

Precision of Duo Standards

Duo Standards as Unknowns of Duo	Plate 1 Avg		Plate 3 Avg	Average of all 3	Standard Deviation
50 Human Standard	87		55	56	20
50 Human Standard Y	64		55	58	18
16.7 Human Standard	12.6		16.6	16.1	3.6
16.7 Human Standard Y	12.9		16.2	15.6	3.4
5.56 Human Standard	4.6		5.4	5.0	1.3
5.56 Human Standard	4.5		5.4	5.2	1.2
1.85 Human Standard	1.07		1.60	1.60	0.26
1.85 Human Standard Y	1.02		1.62	1.65	0.24
0.62 Human Standard	0.63		0.55	0.58	0.13
0.62 Human Standard Y	0.66		0.53	0.59	0.15
0.21 Human Standard	0.212		0.217	0.212	0.007
0.21 Human Standard Y	0.225		0.219	0.218	0.008
0.068 Human Standard	0.079		0.069	0.072	0.020
0.068 Human Standard Y	0.065		0.063	0.060	0.020
0.023 Human Standard	0.025		0.026	0.025	0.0046
0.023 Human Standard Y	0.024		0.023	0.020	0.0023

Precision of Duo Standards

Duo Standards as Unknowns on Plexor	Plate 1	Plate 2	Average	Standard Deviation
50 Human Standard	61	106	81	28
50 Human Standard Y	38	53	45	13
16.7 Human Standard	17	23	20.0	4.2
16.7 Human Standard Y	17	18	17.5	0.71
5.56 Human Standard	5.5	6.7	6.10	0.85
5.56 Human Standard	5.4	5.6	5.50	0.14
1.85 Human Standard	2.3	1.8	2.05	0.35
1.85 Human Standard Y	1.9	1.6	1.80	0.42
0.62 Human Standard	0.64	0.70	0.60	0.050
0.62 Human Standard Y	0.62	0.63	0.59	0.005
0.21 Human Standard	0.23	0.23	0.23	0
0.21 Human Standard Y	0.16	0.13	0.16	0.043
0.068 Human Standard	0.071	0.073	0.072	0.0014
0.068 Human Standard Y	0.070	0.076	0.070	0.0021
0.023 Human Standard	0.020	0.027	0.032	0.0085
0.023 Human Standard Y	0.017	0.017	0.017	0

Finish with sig figs

Results

- Results show good concordance between the two kits for both the Plexor and Quantifiler Duo standards at all levels
- Precision between the two kits was comparable with Duo being a little better at higher concentrations and Plexor better at low concentrations
- Plexor appeared to have a lower LOD as expected

Part two:

Validation tests with NIST standards

- Three NIST standards (#2372) were examined to check accuracy and sensitivity
  - Component A is a single source male that should quantitate as 52.4ng/ul according to NIST.
  - Component B is a pooled female sample that should quantitate as 53.6ng/ul according to NIST
  - Component C is a pooled mixture of males and females that should quantitate as 54.3ng/ul according to NIST
- Standards were examined directly and also using a serial dilution.

Data from NIST standards

	Expected from NIST	Quantifiler Human Avg. Result	Average Plexor Result	Auto/ Y ratio	Average Duo Result	Auto/ Y ratio
QSRM A Auto	52.4	51.62	32		59.53	
QSRM A Y			63.67	0.50	56.82	1.05
QSRM B Auto	53.6	63.92	60		57.56	
QSRM B Y			N/A	N/A	N/A	N/A
QSRM C Auto	54.3	70.52	41.375		33.44	
QSRM C Y			62.25	0.66	23.53	1.42

Results

- The Plexor HY results appear to show an anomalously low human quantification relative to the male quant, particularly for the single source male
- The Duo results are more consistent for the single source male sample but are just as poor as Plexor for the mixed sample
- Results for these high level samples however may suffer from inaccuracy – further tests are needed.

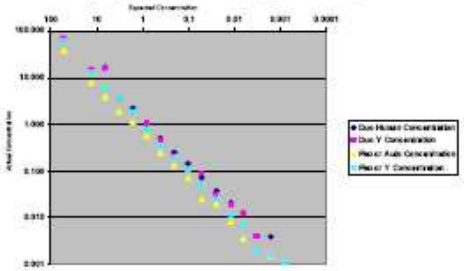
To examine these potential differences with more precision a series of dilutions were tested

- NIST Std A (single male) using both Plexor HY and Duo
- Standard A was diluted by halves starting with 13.1 ng/uL
- QSRM A = 52.4
  - Dilution values in ng/ul 13.1, 6.55, 3.275, 1.6375, 0.818, 0.409, 0.204, 0.102, 0.051, 0.025, 0.012, 0.0064, 0.0032, 0.0016, 0.0008, 0.0004, and 0.0002
- Standard A was run in triplicate with each chemistry
- Each diluted sample was amplified with real time PCR once with each chemistry

Plot of Experimental vs Actual Values

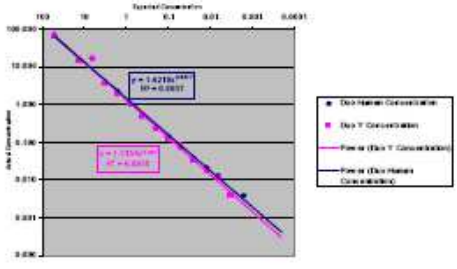
NIST QSRM A Duo and Plexor HY

(Duo produces a relatively higher estimate)



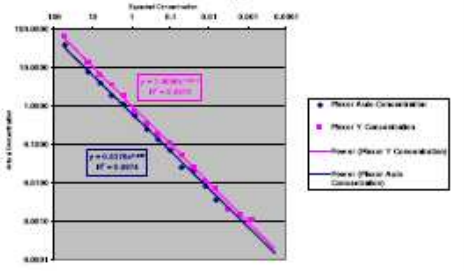
Plot of Experimental vs Actual Values

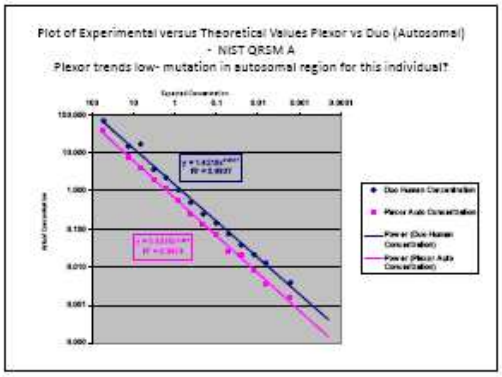
QSRM A (Duo Autosomal vs Y) Results show similar slope

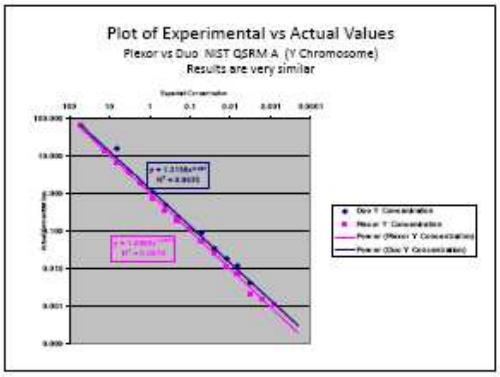


Plot of Experimental Quantities versus Actual values for QSRM A (Plexor HY -Auto and Y)

Y trends higher than autosomal- indicates extra male copy or loss of autosomal DNA copy

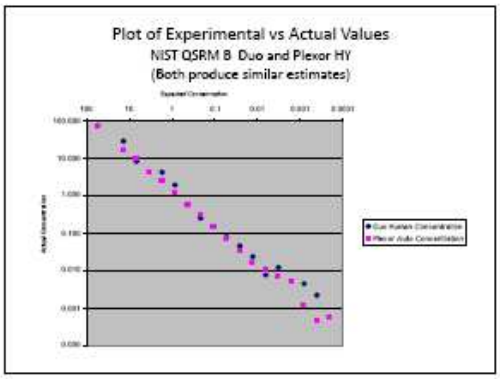


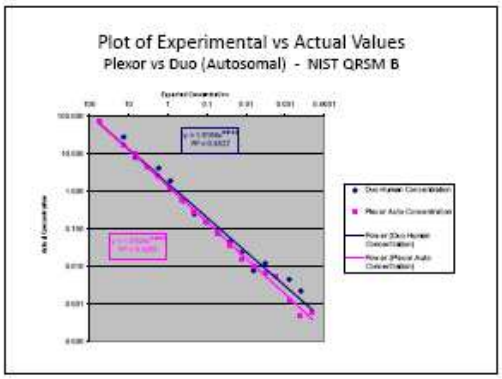




To examine these potential differences with more precision a series of dilutions were tested

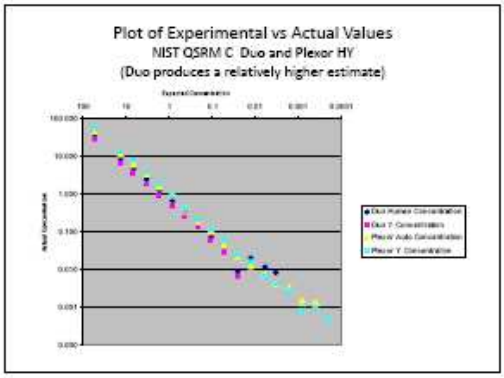
- NIST Std B (Pooled female) using both Plexor HY and Duo
- Standard B was diluted by halves starting with 13.4 ng/uL
- QSRM B = 53.6
  - Dilution values in ng/ul 13.4, 6.7, 3.35, 1.675, 0.8375, 0.418, 0.209, 0.104, 0.052, 0.026, 0.013, 0.0065, 0.0032, 0.0016, 0.0008, 0.0004, and 0.0002
- Standard B was run in triplicate with each chemistry
- Each diluted sample was amplified with real time PCR once with each chemistry

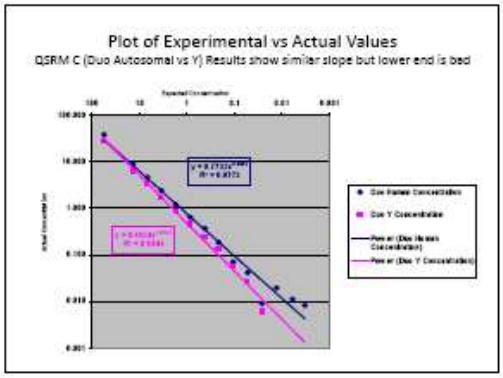


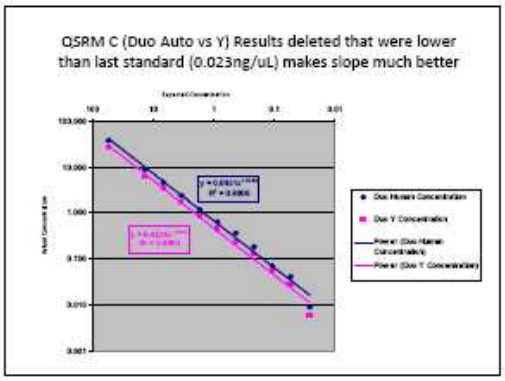


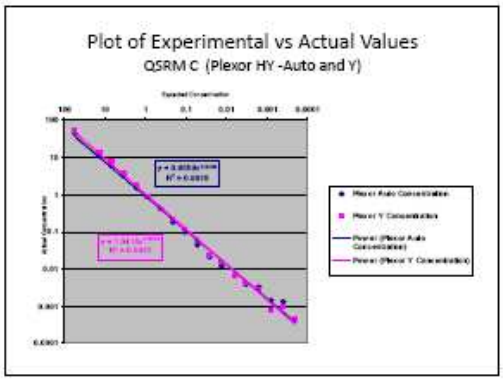
To examine these potential differences with more precision a series of dilutions were tested

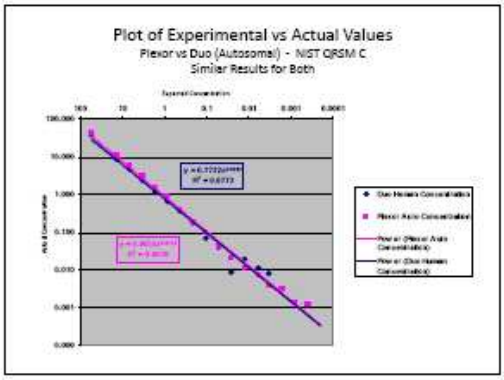
- NIST Std C (Pooled mixture of males and females) using both Plexor HY and Duo
- Standard was diluted by halves
- QSRM C = 54.3
  - Dilution values in ng/ul 13.575, 6.78, 3.39, 1.69, 0.848, 0.424, 0.212, 0.106, 0.053, 0.026, 0.013, 0.006, 0.0033, 0.0016, 0.0008, 0.0004, and 0.0002
- Standard C was run in triplicate with each chemistry
- Each diluted sample was amplified with real time PCR once with each chemistry

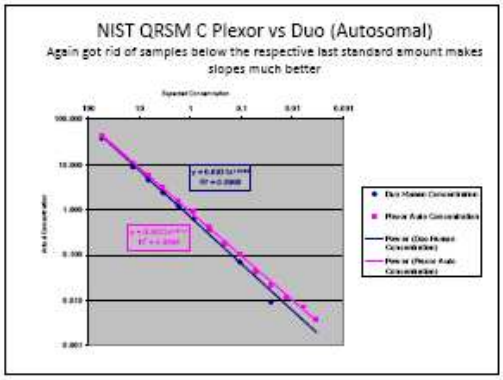


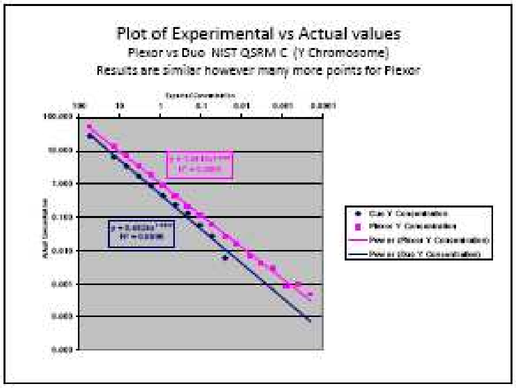


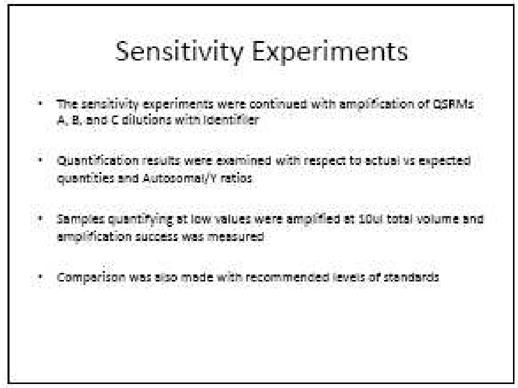


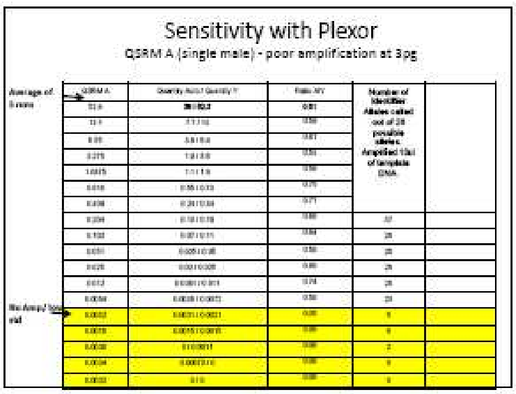


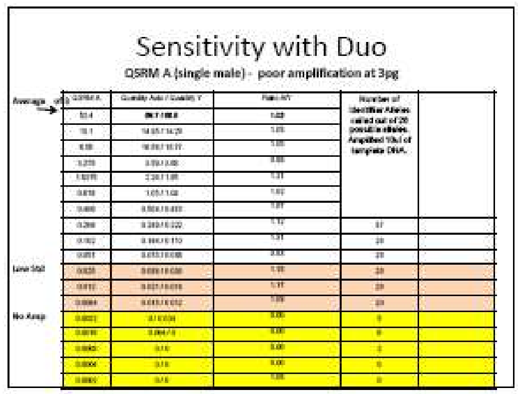


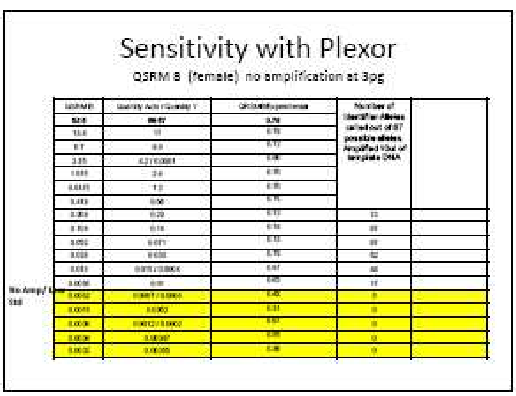


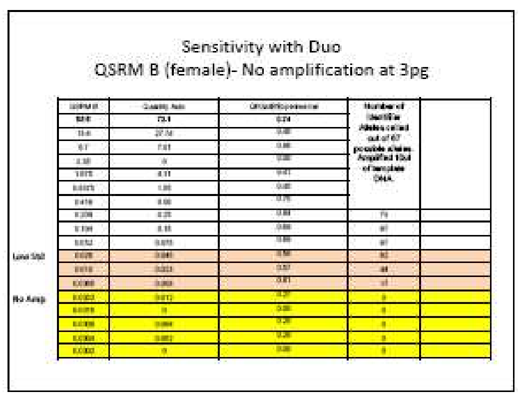






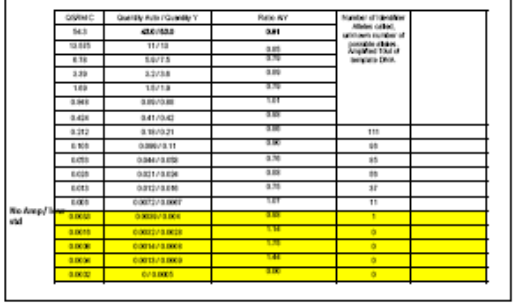






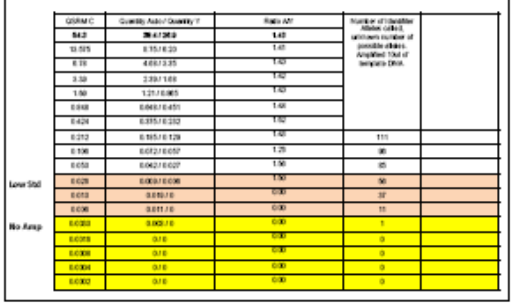
Sensitivity with Plexor

QSRM C (mixed males and females) - No amplification at 3pg



Sensitivity with Duo

QSRM C (mixed males and females) no Amp at 3 pg



Ratio Results (NIST QSRMs A-C)

Quantities from 13ng – 6.5 pg/uL*

Note: In all cases, expected ratio is 1.00

	Plexor a/y SRM A	Duo a/Y SRM A	SRMB/ Plexor Female	SRMB/ Duo Female	Plexor a/y SRM C	Duo a/y SRM C
Average	0.64	1.09	0.75	0.65	0.88	1.44*
Std dev.	0.10	0.12	0.06	0.16	0.10	0.10*

* Duo could not detect Y ratios below 26pg for SRM C

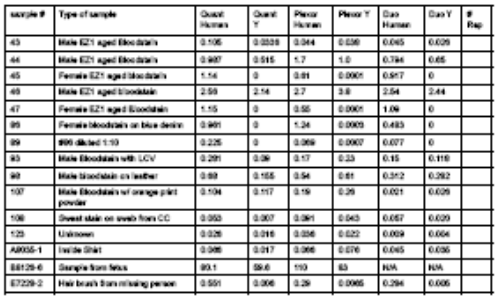
Conclusions on sensitivity

- Both kits showed a high degree of linearity over the tested range
- Precision for both kits showed RSDs ranging from 7-16% for Y ratios and for QSRM B/experimental of 8% (plexor) and 24% (duo)
- Accuracy was relatively poor and sample dependant, indicating potential sequence effects, Neither kit did well in this respect. In particular there appeared to be problems with Plexor (autosomal) for QSRMA and with Duo for QSRMC (Y)
- Sensitivity was better for Plexor as expected, although Duo performed better than expected below its suggested range.

Casework Samples

- To further test the two systems a set of casework samples were investigated
- Quantifiler Human, Y Duo, and Plexor HY were compared
- Data is presented below

Casework Samples



Casework Samples

sample #	Type of sample	Quant Human	Quant Y	Plexor Human	Plexor Y	Duo Human	Duo Y
E7229-7	Cigarette filter contact	0.0065	0	0.0043	0.0042	0.003	0.006
E7229-9	Left glove- inside	0.0165	0	0.023	0.01	0.007	0.003
E7229-12	Swabbing of mask	0.0079	0.0039	0.01	0.012	0.005	0
E7229-13	Cutting of mask	0.122	0.058	0.11	0.13	0.067	0.052
W6207-22	Passenger seat back blood	0.0569	0.0435	0.054	0.06	0.033	0.034
W6011-2	Knife handle	0.415	0.759	0.52	0.11	0.266	0.095
W7192-25	NW bedroom wall swab blood	0.454	0.506	0.23	0.56	0.302	0.366
W7192-26	Metal shelf blood	0.413	0.39	0.29	0.54	0.302	0.329
W7192-50	NW bedroom wall swab blood	0.195	0.105	0.24	0.16	0.107	0.057
W7192-54	Roof storage of pickup blood	0.576	0	0.69	0.0001	0.706	0
W7192-57	Masking tape blood	0.0245	0	0.048	0.0002	0.706	0
W6070-44	Trash room floor blood	0.243	0	0.22	0.0002	0.305	0
W6070-45	Exterior trash chute door blood	0.22	0	0.16	0.0001	0.193	0
W6070-46	Interior trash chute door blood	0.19	0	0.16	0.0003	0.147	0

Casework Samples

sample #	Type of sample	Quant Human	Quant Y	Plexor Human	Plexor Y	Duo Human	Duo Y
1325	Vag swab –non sperm fraction	1.3	0.031	2.2	0.033	0.989	0.012
1325	Vag swab –sperm fraction	0.274	0.022	0.265	0.038	0.09	0.055
1356	Vag swab –non sperm fraction	1.65	0.0193	1.85	0.034	1.07	0.019
1356	Vag swab –sperm fraction	0.25	0.093	0.485	0.235	0.125	0.107
1356	Vag swab –non sperm fraction	0.974	0.0359	1.35	0.0357	0.783	0.039
1356	Vag swab –sperm fraction	0.299	0.052	0.26	0.14	0.145	0.094
1366	Vag swab –non sperm fraction	1.76	0	2.15	0.0304	1.97	0
1366	Vag swab –sperm fraction	1.01	0.533	1.1	0.69	1.09	0.57
87238-1SF	Vag swab –sperm fraction	6.15	1.32	8.4	2	4.34	1.41
87238-2SF	Vag swab –sperm fraction	3.94	1.41	17	3.6	2.98	1.92
05342HE9NS	Vag swab –non sperm fraction	23.5	0	11	0.37	10.86	0.422
05342HE8NS	Cervical swab non sperm fraction	5.01	0.01	3.5	0.011	4.42	0.011
05342HE5SF	Cervical swab - sperm fraction	0.534	0.287	0.46	0.62	0.539	0.459

Conclusions casework

- Results need to be statistically analyzed but generally track each other
- A two treatment statistical test can be used.
- Further testing is needed to determine the utility of the A/Y ratio in predicting quality of results.

Other Issues to be resolved

1. The issue of Amplicon Size and recovery of degraded DNA - DUO amplicon sizes are 140 A and 130 Y Plexor sizes differ.

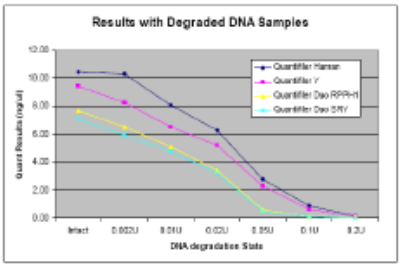
Kit	Gene	Target Location	Amplicon Length
Quantiflex Human	Human telomerase reverse transcriptase gene	5p15.33	62 base
Quantiflex Y	Sex-Determining region Y gene	Yp11.3	64 base
Quantiflex IPC	synthetic sequence		79 base[1]
Quant Duo Human Target	Ribonuclease P RNA Component H1 (RPPH1)	14q11.2	140 base
Quant Duo Human Male Target	Sex-Determining region Y (SRY)	Yp11.3	130 base
Quant Duo IPC	synthetic sequence		130 base
Plexor HY Human Autosomal	Human SNP2 locus (Multicopy)	17 Chromosome	90 base
Plexor HY Y	Testis-specific protein, Y-encoded (TSPY) locus (Multicopy)	Y Chromosome	133 base
Plexor HY IPC	Novel sequence		150 base

Amplicon size for different kits

Kit	Gene	Target Location	Amplicon Length
Quantiflex Human	Human telomerase reverse transcriptase gene	5p15.33	62 base
Quantiflex Y	Sex-Determining region Y gene	Yp11.3	64 base
Quantiflex IPC	synthetic sequence		79 base[1]
Quant Duo Human Target	Ribonuclease P RNA Component H1 (RPPH1)	14q11.2	140 base
Quant Duo Human Male Target	Sex-Determining region Y (SRY)	Yp11.3	130 base
Quant Duo IPC	synthetic sequence		130 base
Plexor HY Human Autosomal	Human SNP2 locus (Multicopy)	17 Chromosome	90 base
Plexor HY Y	Testis-specific protein, Y-encoded (TSPY) locus (Multicopy)	Y Chromosome	133 base
Plexor HY IPC	Novel sequence		150 base

Effect of Amplicon size- ABI kits

Target size QHuman 62 bp – 2copies, QY 64-1 copy Dhuman 140, DY 130

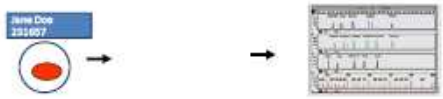


AFDAA winter meeting 2008 - ABI

Investigation of the effects of sample degradation and inhibition in forensic DNA typing with reference to QPCR

Bruce McCord
Professor Analytical/Forensic Chemistry
FIU Department of Chemistry
International Forensic Research Institute
Miami, FL mccordb@fiu.edu 305-348-7543

Large Multiplex Kits provide Efficient and Rapid Analysis of Convicted Offender Samples



But what about degraded DNA ?

Skeletal material being prepped for extraction

Such samples present a special challenge

DNA Degradation

1. polymer hydrolyzes (nucleic acids break apart

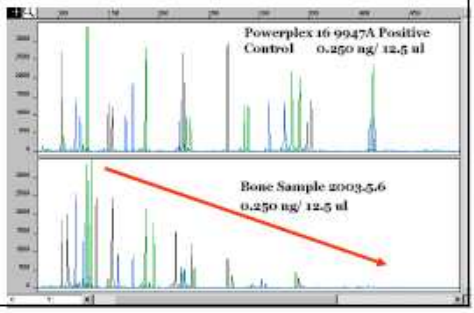
2. Pyrimidine dimers (bases X-link)

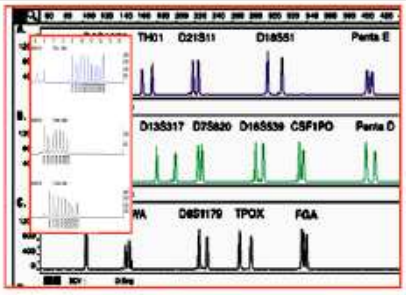
3. Chemical oxidation (bases become unreadable)

Thymine glycol

DNA Degradation

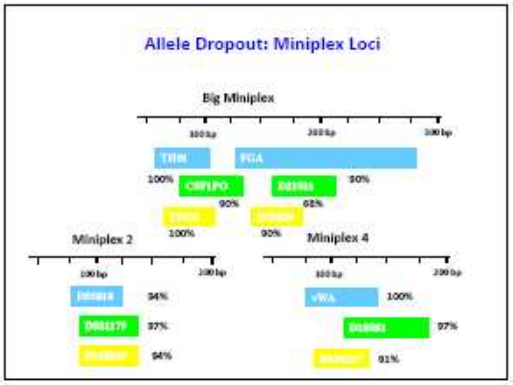
Note loss of intensity of larger alleles

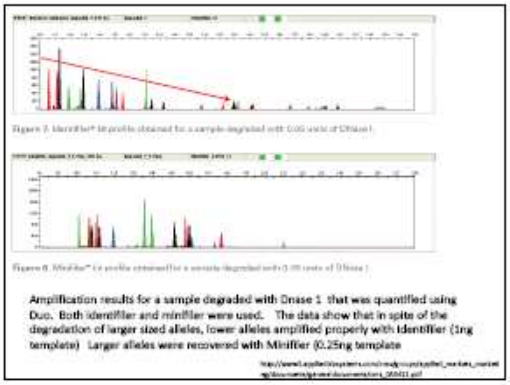


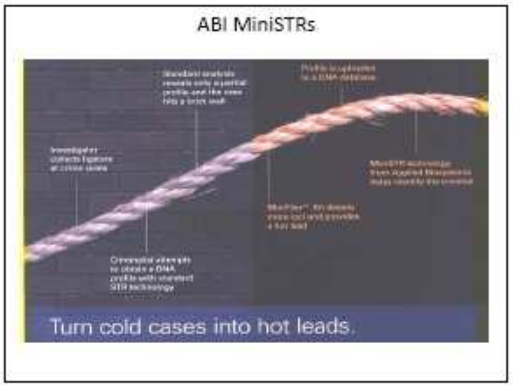


Miniplex 1 vs Powerplex 16

Allele Dropout: Standard DNA kit
Promega Powerplex 16





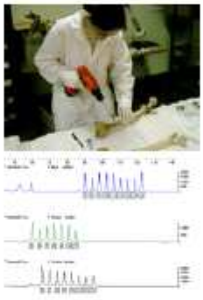


Non-Traditional Samples

- The obvious application for Mini-STRs are samples which are degraded, difficult or of low copy
- We have performed experiments to examine their sensitivity and
  - Capability to amplify DNA from bone and hair
  - Capability to amplify DNA from highly degraded samples
  - Big question: How to determine quality and quantity of low level DNA?
  - Answer QPCR?

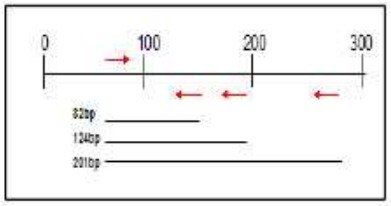
Work in FIU Laboratory- with assistance of Vermont Crime Lab

- Development of miniplex STRs for degraded DNA typical sizes 60-120 bp.
- Slot blot works poorly on these samples
- So develop a series of different primers to selectively amplify degraded dna

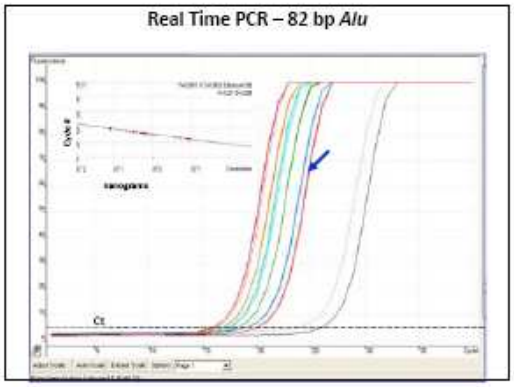


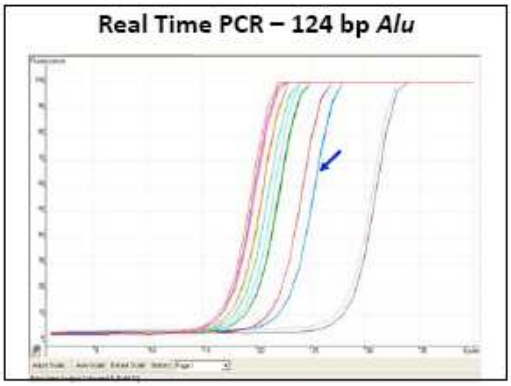
Determination of DNA Quality by qPCR

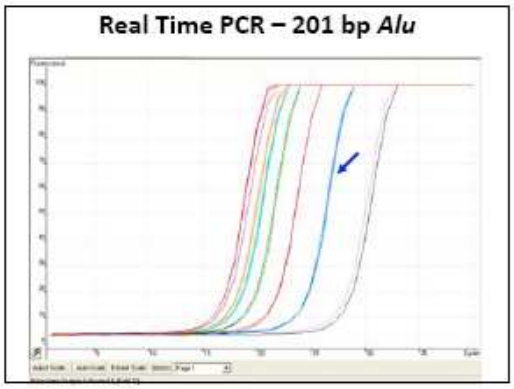
A/uYa5 Primers – Nicklas and Buel

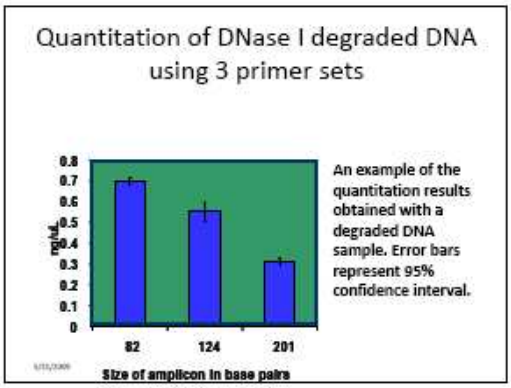


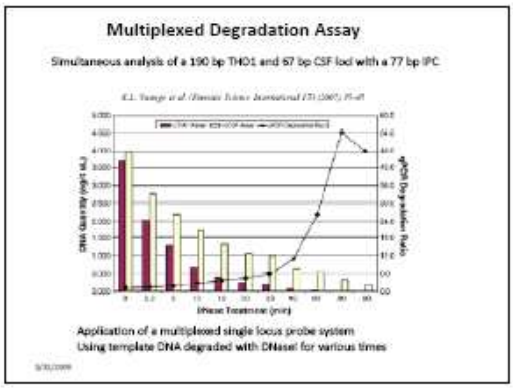
Primer design

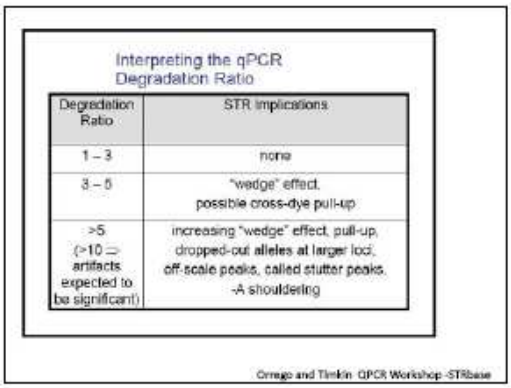


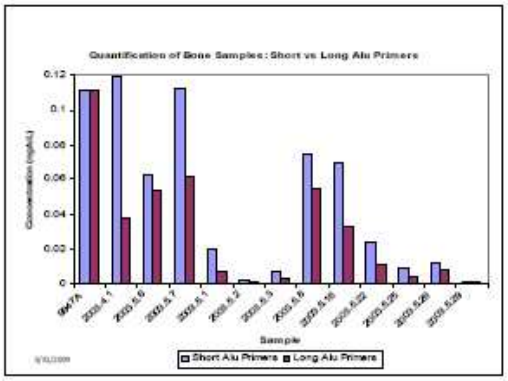


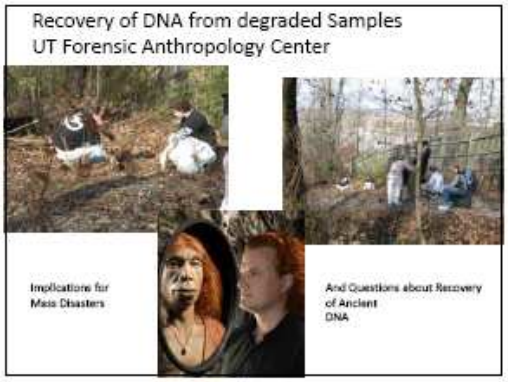


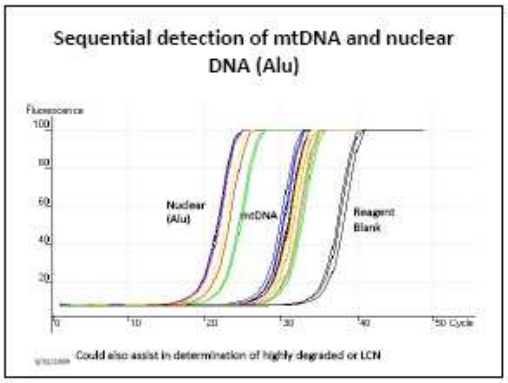


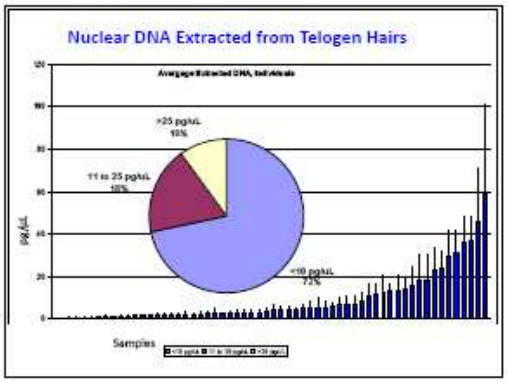


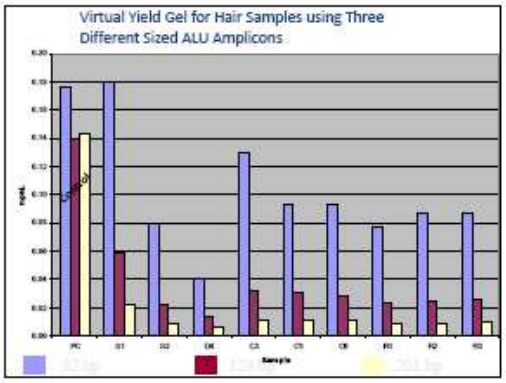




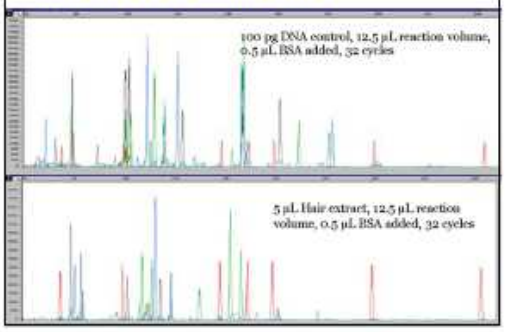




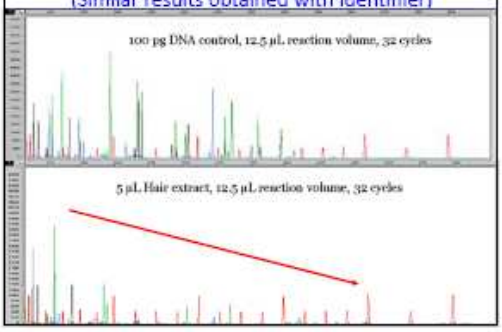




Amplified Sample - Miniplex



Amplified Sample – Powerplex (Similar results obtained with Identifiler)



Results for telogen hairs

- <100 pg – Mini 2 (D5, D8, D16)
  - 60% 2 loci (D5 and D16)
  - 80 % at least 1 locus (D16)
- 100-500 pg Mini 2 +Mini 4 (vWA, D18, D13)
  - 70% 3 or more loci
  - 30% 1 loci or less
- >500 pg all 3 sets
  - 70% 4 loci or more
  - 40% 6 loci or more

Sometimes you can run telogen hairs

The Down Side of MiniSTRs

MiniSTRs were developed to access degraded DNA.

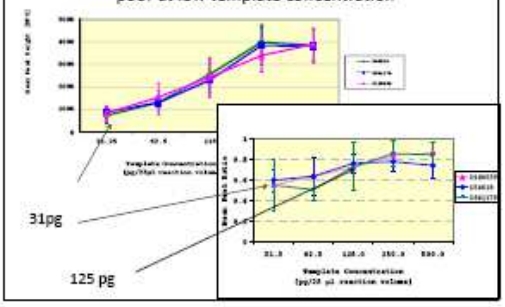
They do not solve the inherent low copy limitation of the PCR.

Instead because of their sensitivity, they complicate it.

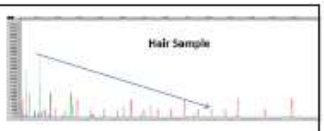
1. Short PCR primers amplify better.
2. Better amplification means laboratories can access LCN without pushing cycles
3. At low copy a scientist cannot express a strong opinion about how DNA arrived at the site where it was recovered. This DNA could just as easily come from pre or post deposition as it could come from the suspect.

MiniStrs show the same effect

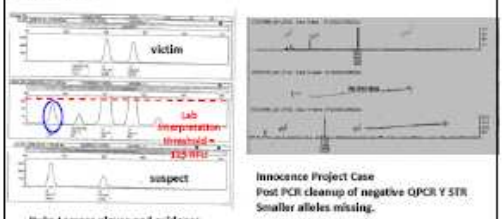
In spite of the improved sensitivity, peak balance is poor at low template concentration



This amplification needs MiniSTRs



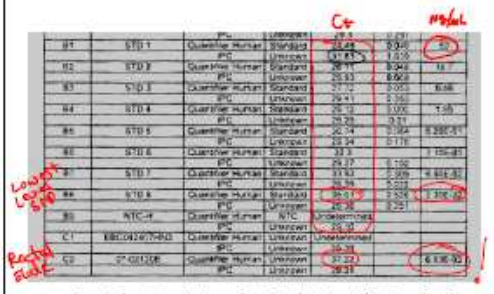
These do not:



Innocence Project Case
Post PCR cleanup of negative QPCR Y STR
Smaller alleles missing.

Duke Lacross player and evidence

What to look for in the QPCR results.



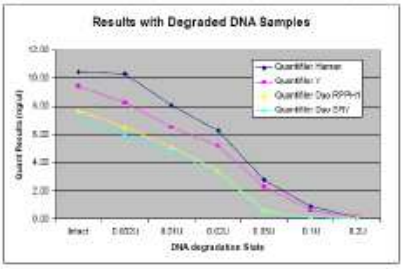
Sample is at a low level, nominally 6pg/uL but outside the calibration

Amplicon size for different kits

Kit	Gene	Target Location	Amplicon Length
Quantifiler Human	Human telomerase reverse transcriptase gene	5p15.33	62 base
Quantifiler Y	Sex-determining region Y gene	Yp11.3	64 base
Quantifiler IPC	synthetic sequence		79 base[1]
Quantifiler Duo Human Target	Ribonuclease P RNA Component H1 (RPPH1)	14q11.2	143 base
Quantifiler Duo Human Male Target	Sex-determining region Y (SRY)	Yp11.3	139 base
Quantifiler Duo IPC	synthetic sequence		130 base
Plexor HY Human Autosomal	Human RNPC1 locus (Multicopy)	17 Chromosome	70 base
Plexor HY Y	Selenoprotein protein, T-encoded (TSPY) locus (Multicopy)	Y Chromosome	133 base
Plexor HY IPC	Novel sequence		179 base

Effect of Amplicon size- ABI kits

Target size QHuman 62 bp – 2copies, QY 64 bp-1 copy QHuman 140bp, QY 130 bp



AFDAA winter meeting 2008 - ABI

Conclusions

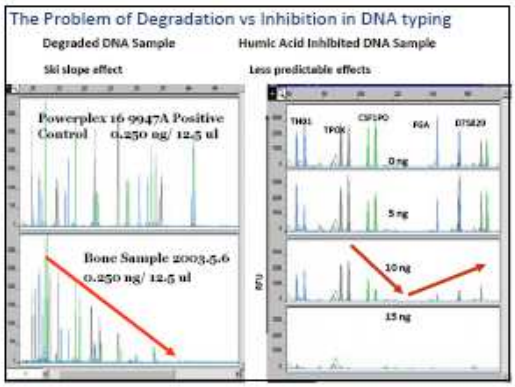
- MiniSTRs are for degraded DNA
- Validation data reveals a robust and sensitive multiplex amplification
- Virtual yield gel using qPCR helpful for proper analytical results
- Stochasitic effects still occur for samples under 125pg
- Improved results are possible for bone and telogen hair
- Degradation is still a problem

Inhibition and degradation

Which way to turn?

QPCR and Inhibition

- Inhibition Mechanisms
- Effects of inhibition on STR profiles
- A real time assay for inhibitors
- The effect of inhibition on Quantifiler and duo
- Plexor and inhibition

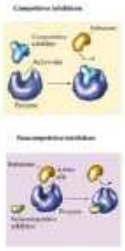


The Issue:

- With increasing interest in the forensic community in the interpretation of compromised samples and mixtures, we need to be able to better interpret electropherograms in court
  - We need to determine the relative effects of DNA degradation and inhibition on peak height ratios.
  - We need to understand the combinatorial effects of different inhibitors
  - We need to understand the environmental aspects of degradation and soil inhibition
  - We need to explore the interpretation of low level mixtures in the presence of a major contributors

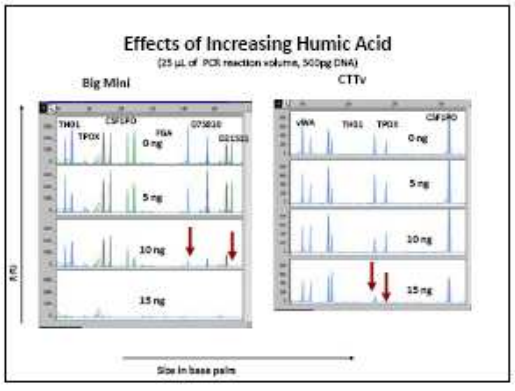
PCR Inhibitors

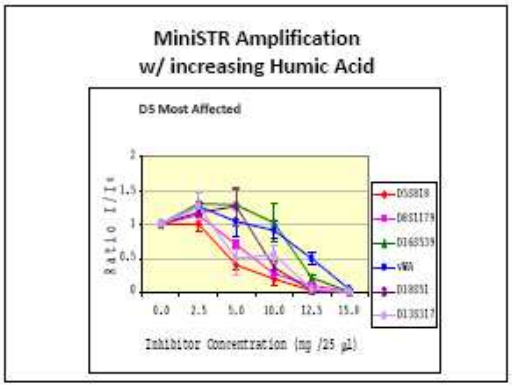
- Substances co-extracted in forensic samples that affect amplification of template DNA
- Theories
  - Inhibitors bind with the polymerase
  - Inhibitors interfere with polymerase by binding to DNA
  - Polymerase affected during primer extension



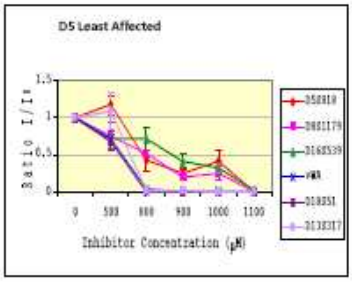
PCR Inhibitors

- In the PCR process, the enzyme moves along the DNA strand, adding complementary bases
- If inhibitors are present the PCR process fails- why?
- In our studies, the failure seems to be more a function of sequence than of amplicon size.
- Mechanisms for inhibition may very however, and in many cases- soil and bone- for example,
  both degradation and inhibition may be present



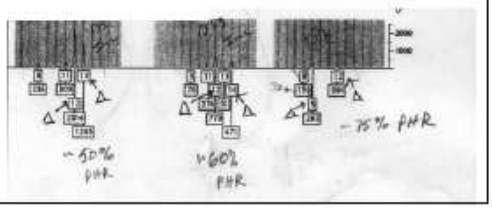


MiniSTR Amplification w/ increasing calcium



The Problem with Mixtures interpretation in the face of set PHRs, inhibition and degraded DNA

This case resulted in a hung jury. Defense expert claimed low PHR indicated exclusion. Prosecution claimed inconclusive result
Lab Protocol: loci must be above threshold and within 66% of each other to be included as major contributor. Defense confused this. Inconclusive ≠ Exclusion!
Defendants profile is D5(12,15) D13(12,14) D7(9,12)



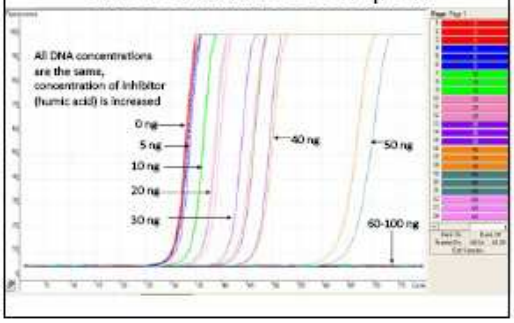
Results – Inhibition Thresholds MiniSTRs

Inhibitor	Big Miniplex	Miniplex 2	Miniplex 4
Hematin	1 μM	0.8 μM	0.8 μM
Indigo	280 μM	320 μM	300 μM
Melanin	0.16 ng/μL	0.2 ng/μL	0.2 ng/μL
Humic Acid	0.5 ng/μL	0.6 ng/μL	0.6 ng/μL
Collagen	24 ng/μL	32 ng/μL	24 ng/μL
Calcium	1100 μM	1100 μM	800 μM

PCR Inhibition: Observations

- Inhibitors act in many ways – The most worrisome are those which co-extract with DNA
- These inhibitors produce various effects on data including- peak balance problems, locus specific dropout, enhanced stutter, and poor sensitivity
- Mechanisms appear to vary with type of inhibitor and sequence of amplicon
- It is important to understand concentration effects and mechanisms so that inhibition cannot be confused with degradation, dropout and mixture effects

Effect of Inhibitors on qPCR



Tests for PCR inhibition using Realtime PCR with high resolution melt

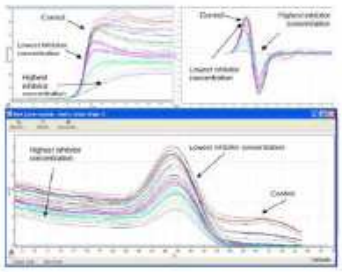
- Compare inhibition for a single locus (TH01) with primers of various lengths and melting temperatures
- Determine the effect of length and sequence on PCR inhibition
- Examine melt temperatures of amplicons in inhibited samples
- Classify inhibitors by effect on PCR
- Determine mechanism of different PCR inhibitors

Experimental Design

- Locus – HUMTH01 STR
- DNA – Homozygous 9.3 allele
- Primers
  - 3 lengths (~100, 200, 300 bp)
  - 3 Tm (58, 60, 62° C)
  - Amplified product (one product)
- qPCR conditions
  - Standard conditions for *Alu* quantification (Nicklas et al. 2003)
    - No BSA
    - Lower Primer Concentration
    - Reduced Taq
    - Annealing temperature appropriate for Tm (5 ° C < Tm)
- Inhibitors
  - Calcium, humic acid, hematin, collagen, melanin, tannic acid

Results - Calcium

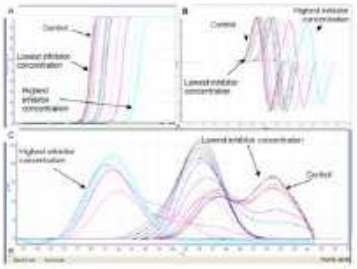
- No shift in take off cycle
- No change in melting curve
- Efficiency of amplification affected
- No difference for size or Tm



Conclusion: Taq Inhibitor

Results-Humic Acid

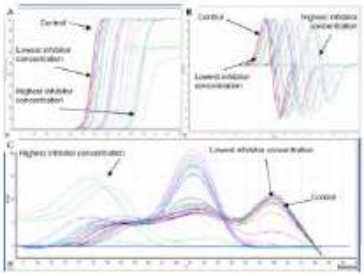
- Effects Ct
- No efficiency change
- Melt curve effects



Conclusion: Inhibits by binding to DNA template

Results - Melanin

- Effects Ct
- weak effect on efficiency
- melt curve effects



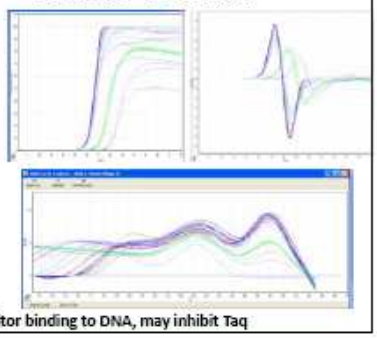
Conclusion: Inhibits by binding to DNA template, may affect Taq

Results - Hematin

Strong Ct Shift some effects on efficiency

Some loss of product

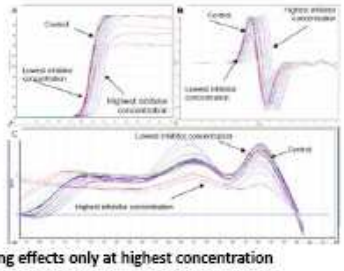
Evidence of binding DNA esp. at high conc.



Conclusion: Inhibitor binding to DNA, may inhibit Taq

Results – Tannic Acid

- Increase in take off cycle
- No change in efficiency
- Minimal effect on melt curve



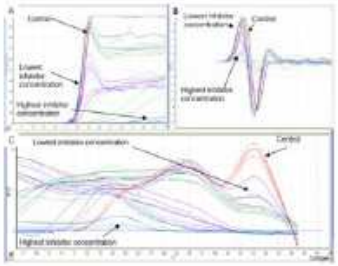
Conclusion: Binding effects only at highest concentration
Weaker effect than Humic acid or Melanin

Results - Collagen

Strange loss in signal with increased [inhibitor]

Minimal effect on takeoff cycle

Evidence of binding DNA And efficiency changes

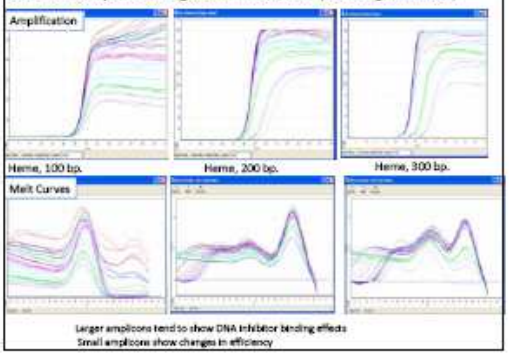


Conclusion: Inhibitor binding to DNA, inhibits Taq

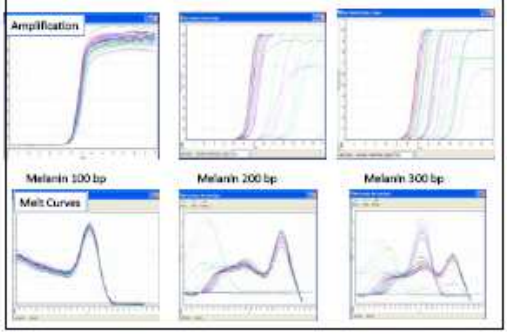
Size Effects

- Inhibition tests examined both concentration and amplicon size of THO1 target. For size:
- Hematin showed effeciency changes for short amplicons and Ct effects for larger amplicons.
- Melanin shows Ct effects and melt curve effects for larger amplicons.
- Results indicate Ct effects are reduced for smaller amplicons for these inhibitors. This indicates a potential advantage for these types of inhibitors with MiniSTRs.

Effect of amplicon length on inhibition (incressing conc. Heme)



Melanin results at various amplicon sizes (THO1)
Inhibitor binding occurs with larger amplicon sizes



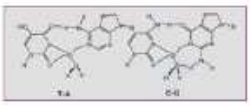
Important point for these results

- Just as MiniSTRs may be useful in amplifying DNA that is affected by inhibitors, short amplicons may be unable to pick up inhibitors that only affect large STRs
- Inhibitors that bind Taq will always be detected regardless of amplicon size.

Transition metal ions

Metal cations present in degraded samples

represent a different type of inhibition

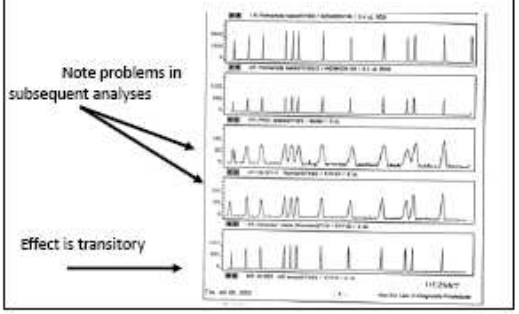


Zn^{2+}, Co^{2+}, and Ni^{2+} form DNA-metal ion complexes, termed M-DNA.at pH conditions above 8,

These cations produce severe effects in CE injection and analysis

Hartzell and McCord, *Electrophoresis*, 2006

Effect of contaminant in reference sample



Types of Inhibitors

- Length dependent
  - Largest amplicon least affected by inhibition
    - Collagen
    - Affects signal – possible dye quencher
    - Possible binding to DNA
- Tm/Sequence dependent kinetics
  - Humic Acid - 60 Tm least affected
  - Tannic Acid* – 58 Tm least affected
- Change in melt curve/DNA binding
  - Humic acid, collagen
  - Larger amplicons with melanin and hematin

Types of inhibitors 2

- Taq inhibition
  - Calcium
- Length dependent/primer extension
  - Smallest amplicon least affected by inhibition
    - Hematin
    - Melanin
    - Tannic Acid
- DNA intercalation
  - Metalized DNA (transition metals Zn, Co)
  - Certain dyes

Conclusions on Inhibitors

- Preliminary studies demonstrate that inhibition and degradation may produce different patterns of allele dropout
- Inhibition can
  - Reduce Taq availability (intensity decrease)
  - Reduce apparent concentration (takeoff cycle rise)
  - Affect specific loci/bind template (tm effects)
  - Affect Taq processivity (amplicon size effect)
  - Intercalate DNA (affect CE injection)

Clean-Up of PCR Inhibitors

- Bovine serum albumin (BSA)
  - Relieves PCR inhibition by making enzyme more efficient and binding certain inhibitory compounds
- Low-melting temperature agarose/sephadex/filtration
  - Relieves PCR inhibition by capturing large polymers like DNA, releasing smaller inhibitory compounds
- Electrophoretic Purification
  - Inhibitors move at different rates under applied fields
- Addition of higher concentrations of Taq polymerase
  - Overwhelm inhibitors that bind to taq
- Dilution of Sample
  - DNA still amplifies, inhibitors are less concentrated and bind to taq and/or other reaction components
- Destruction of inhibitors w/ NaOH

Improved Workflow

ABI Product Literature

Utilizing Quantifiler duo as an aid to evidence processing

- Duo gives the quantity of human autosomal DNA permitting proper dilution of sample prior to amplification
- Duo permits the quantification of Y DNA, and reveals its presence.
- The internal control sequence permits the determination of the presence of certain inhibitors
- The ratio of male to autosomal DNA permits the determination of the presence of a M/F mixture and the likelyhood of success of mixture interpretation. It also indicates if a sample should be processed for Y STRs.

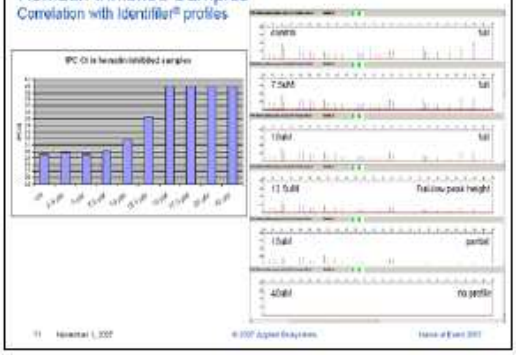
Table 5-3 Interpreting IPC amplification results

Duo Human (VIC® dye) and/or Duo Male (FAM™ Dye)	Duo IPC (NED™ Dye)	Interpretation
No amplification	Amplification	Negative result - no human DNA detected
No amplification	No amplification	Invalid result
Amplification (low C_T and high ΔR_n)	No amplification or C_T higher than 31	IPC result inconclusive
Amplification (High C_T and low ΔR_n)	No amplification or C_T higher than 31	PCR Inhibition

Note: Positive amplification occurs when the C_T value for the detector is <40. Because samples contain unknown amounts of DNA, a large range of C_T values is possible. The IPC system template DNA is added to the reaction at a fixed concentration, therefore, the NED C_T should range between 28 and 31, with a variation of 1 C_T across the standard curve samples.

Hematin Inhibited Samples

Correlation with Identifiler® profiles



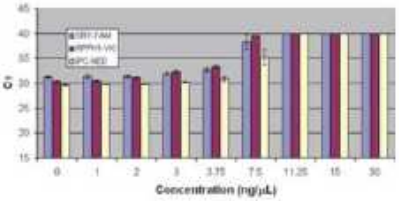
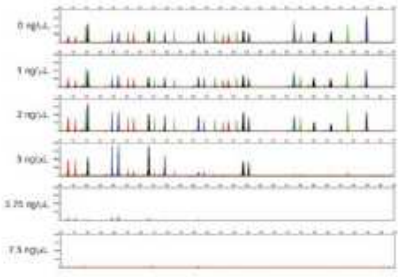


FIG. 4—*Inhibitor study: C_T values for RPPH1, SRY, and IPC targets for inhibited samples containing 0.5 ng/µL DNA and humic acid at final concentrations of 0, 1.0, 2.0, 3.0, 3.75, 7.5, 11.25, 15, and 30 ng/µL in the qPCR.*

Effect of increasing concentrations of humic acid on the amplification of Identifiler (1ng template DNA)



Conclusions

- MiniSTRs are for degraded DNA. They work especially well for hair and bone.
- Similarly QPCR amplicons are sensitive to degradation
- Realtime PCR combined with DNA melt curve analysis can be used to detect PCR inhibition
- Some inhibitors are amplicon size sensitive. This may cause trouble in detecting inhibition using QPCR internal positive controls.

Thank you

Acknowledgements

- Jan Nicklas and Eric Buel
- Jiri Drabek
- Denise Chung, Kerry Opel
- Nancy Tatarek, Lee Meadows Jantz
- John Butler, Yin Shen
- George Duncan

- Major support provided by
- The National Institute of Justice

QPCR and Low copy template

Bruce McCord
Professor Analytical/Forensic Chemistry
FIU Department of Chemistry
International Forensic Research Institute
Miami, FL mccordb@fiu.edu 305-348-7543

QPCR and LCN DNA

- Introduction to LCN
- Drop out and drop in
- qPCR assays for LCN detection
- Sensitivity issues

Spinal Tap Video

- The problem of instrument sensitivity
- Exists everywhere and is fundamental to the concept of signal to noise

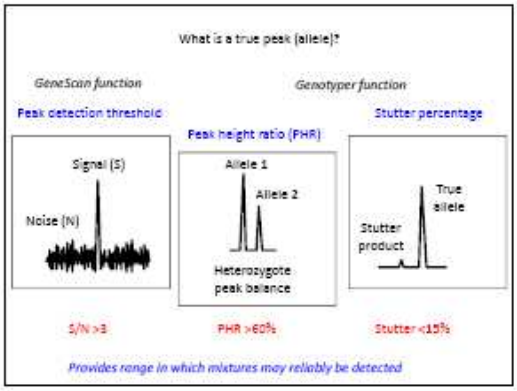
Amounts of DNA Required

1985-1995	RFLP/VNTRs	50 ng – 1000 ng
1991-present (fits since 1998)	PCR/STRs	0.5 – 2 ng
1999-present	LCN/STRs	<0.1 ng

LCN extends the range of samples that may be attempted with DNA testing

Setting thresholds for the ABI 310/3100

- Where do current ideas on instrument thresholds for the ABI 310/3100 come from?
- How do I set these values in my laboratory?
- Why might they vary from one instrument to the next?
- How do these thresholds affect data interpretation?
- How to deal with issue in a way to ensure a conservative result?



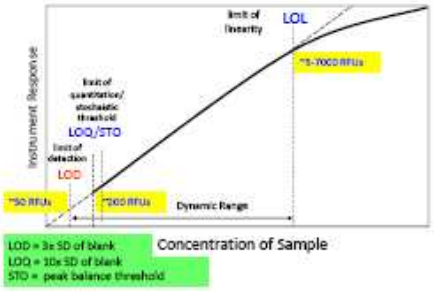
Fundamental ideas behind threshold Settings for the ABI310/3100

Detection Limit: 3x the standard deviation of the noise.
Estimated using 2x peak to peak noise. (approximately 35 - 50 RFUs)
Peaks below this level may be random noise

Limit of Quantitation: 10x the standard deviation of the noise
Estimated using 7x peak to peak noise (150-200 RFUs)
Below this point estimates of peak area or height are unreliable and may not be reliable indicators of mixture ratios

Stochastic Threshold: Level of quantifiable DNA below which peaks can show severe imbalance (peak height ratios below 60%) Approximately 150 -200 RFUs. (always greater than the LOQ) Variance in peak height ratio is the sum of variance due to the stochastic amplification and instrumental noise.

Useful Range of an Analytical Method



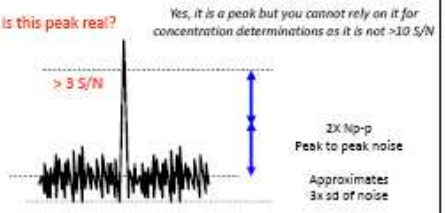
The Scientific Reasoning behind the Concept of an Analytical Threshold
(limit of detection)

- This is fundamentally an issue of reliability
- For a peak intensity below the LOD there is a very real chance that such a signal is the result of a random fluctuation
- This is known as pulling data out of thin air.

Abracadabra! It's an allele

Limit of Detection

- Typically 3 times the signal-to-noise (based on standard deviation of the noise)



The Scientific Reasoning behind the LOQ/Stochastic threshold

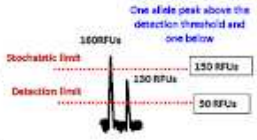
- With peak intensity below the LOQ, you have significant variation in height from one sample to the next.
- Similarly due to stochastic fluctuation in peak height ratios, interpreting data below the stochastic threshold presents the real problem of allele dropout.
- Bottom line: . Input levels of DNA should be sufficiently high to avoid straddle data. Mixture interpretation must be done cautiously on low level data as peak intensities are highly variable.

How low can you go?

Types of Results at Low Signal Intensity
(Stochastic amplification potential)

Straddle Data

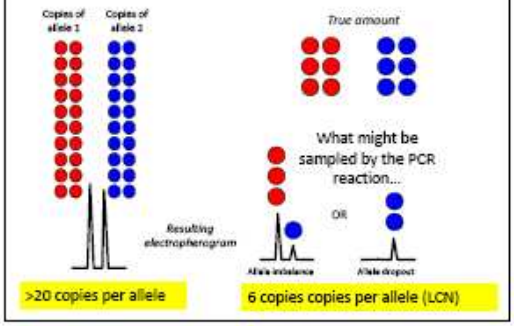
- Only one allele in a pair is above the laboratory stochastic threshold



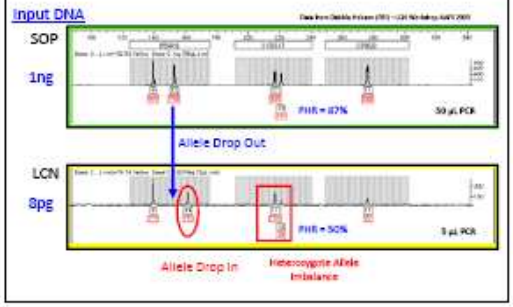
At low levels of input DNA, the potential for straddle data is high.
The issue is best avoided by reamping the sample at higher input DNA
Otherwise straddle data makes locus inconclusive

Straddle data may be caused by degredation, inhibition and low copy issues.

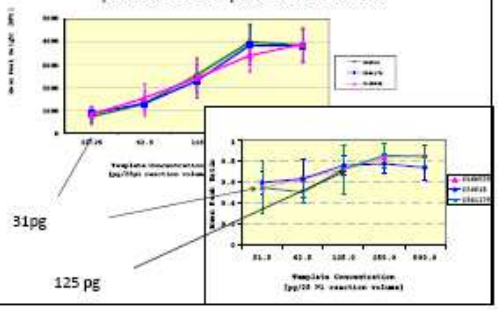
Stochastic Statistical Sampling



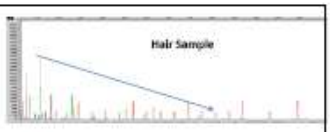
Comparison of STR Kit Amplification SOP with LCN Using the Same DNA Donor



MiniStrs show the same effect
In spite of the improved sensitivity, peak balance is poor at low template concentration



This sample needs MiniSTRs



These do not:

Innocence Project Case
Numerous problems with allele dropout

Duke Lacrosse player and evidence

Interesting effects with low copy data

Table 1 Details of analysis

Number of single cells analysed		226
Results obtained	206 (91%)	
Amplification failure		20 (9%)
Full STR profile		114 (50%)
Acceptable profile (amelogenin,>4 STRs)		144 (64%)
Partial profile (1-4 STRs)		62 (27%)
Surplus alleles*		28 (12%)
False alleles**		11 (5%)
Allele dropout		88 (39%)

*Additional allele present in conjunction with true alleles.

**Additional allele in place of true allele. Extra allelic peaks could be caused by contamination, somatic mutation or PCR-generated non-allelic peaks. We never see more than two additional peaks in a profile or in 10 negatives, minimising the possibility of cellular contamination

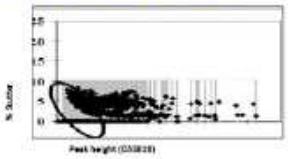
Findley *et al.* (1997) *Nature* article

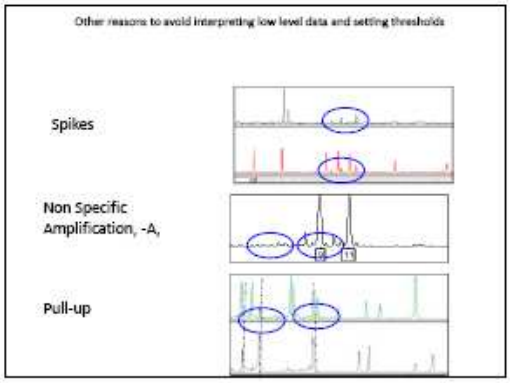
Straddle data may be low copy!

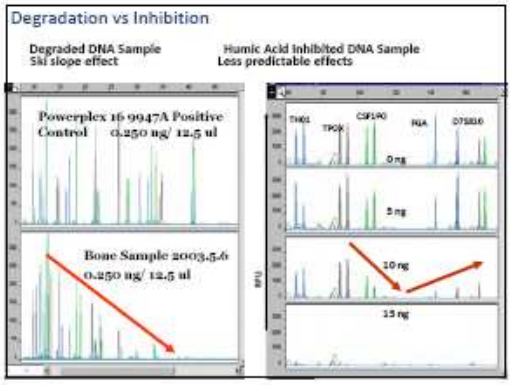
Another problem is stutter

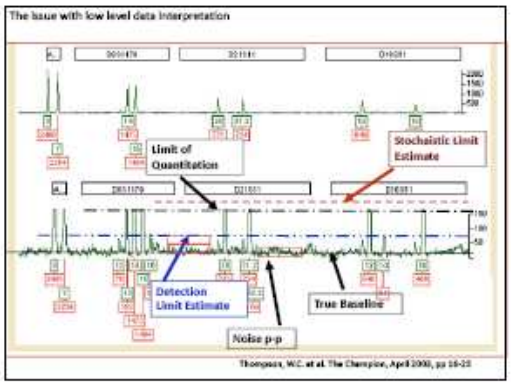
- Stutter increases at low copy #

Leclair et. al (2004) Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. J. Forensic Sci. 49(5): 968-980





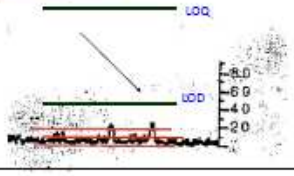




So how to set thresholds?

- First determine the analytical threshold for your particular laboratory using the signal intensity from one or several CE systems

Analytical threshold for this instrument is approx. 50 RFUs, LOQ is approximately180RFUs.

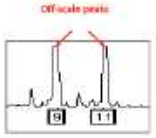


Next determine the dynamic range

1. Perform a series of amplifications of 5 different samples with 5.00, 2.00, 1.00, 0.50, 0.25, 0.13, 0.06, 0.03 ng DNA
2. Use your laboratory quantification system, your thermal cycler, and your 310.
3. Determine the average and standard deviation of each set of samples
4. Your dynamic range is the range of concentrations that are not overloaded. Overload point is where peaks flat top. (this can be checked by ex

Limit of Linearity (LOL)

- Point of saturation for an instrument detector so that higher amounts of analyte do not produce a linear response in signal
- In ABI 310 or ABI 3100 detectors, the CCD camera saturates leading to flat-topped peaks.



Stutter and noise may be artificially enhanced

TWGDAM validation of AmpFlSTR Blue

Wallin *et al.* (1998) *J. Forensic Sci.* 43(4): 854-870

Determination of Minimum Sample

- Goal: avoid situations where peak imbalance results in only one detectable allele from a heterozygous pair.
- Perform serial dilution (1ng- 8pg) of 2 control samples which were heterozygous at all 3 loci
  - Samples above 125pg had peak height RFUs above 150
  - Below 125pg peak heights were not significantly above background
  - At 31 pg peaks were very low or undetectable
- "Peaks below 150 RFU should be interpreted with caution"

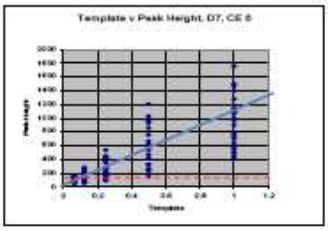
Sensitivity of Detection

Moretti et al, JFS, 2001, 46(3), 661-676

- Different 310 instruments have different sensitivities; determination of analytical thresholds should be performed following in-house studies
  - Variations in quantitation systems
  - Variations in amplification systems
  - Variations in instrument sensitivity
- Peaks with heights below the threshold should be interpreted with caution
  - Caution should be used before modification of
    - Amplification cycles
    - Electrophoretic conditions
- Peak heights are also a function of sample condition/ PCR, extraction, inhibition, degradation

Sensitivity Study:

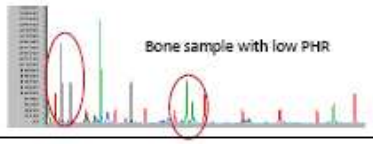
Profiler Plus (Debbie Hobson-FBI)



Observation: Peak height variation increases with concentration
Therefore: its difficult to assess the quantity of DNA solely by peak height

Next set the stochaistic threshold

- Stochastic Threshold – the signal intensity at which a particular quantity of DNA can no longer reliably be detected
  - Reliability can be defined by an increase in the standard deviation of peak height intensity or an increase in the standard deviation of signal intensity or both.
  - The stochastic threshold should be greater than the limit of quantitation

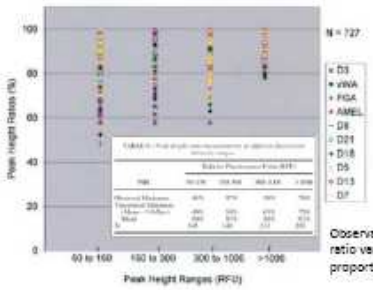


Sensitivity Study

(Debbie Hobson-FBI)

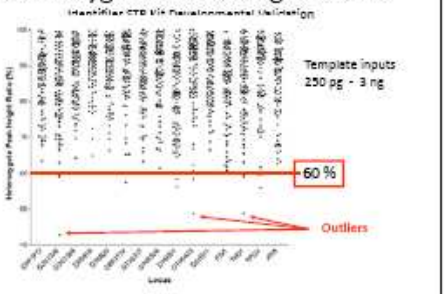
- 25 Individuals
  - 63 pg to 1 ng amplifications with Profiler Plus and Cofiler
  - amplicon run on five 310s
  - GeneScan Analysis threshold sufficient to capture all data
  - GenoTyper: category and peak height
- Import data into Excel
  - peak height ratios determined for heterozygous data at each locus

Heterozygote Peak Height Ratios



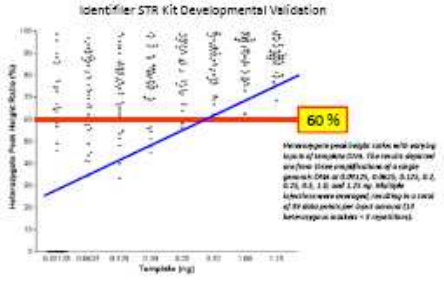
Observation: Peak height ratio variation inversely proportional to input DNA

Heterozygote Peak Height Ratios



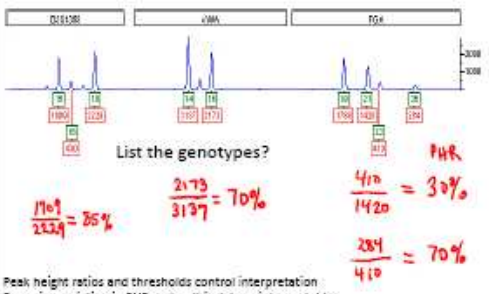
Heterozygote Peak Height Ratios

Identifiler STR Kit Developmental Validation



Mixture Example

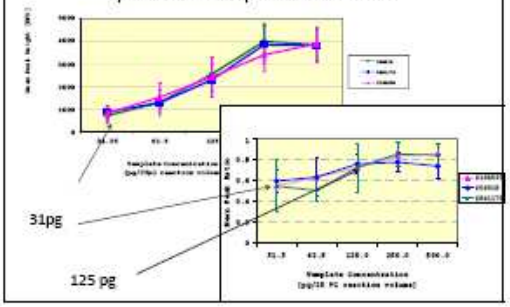
The proper setting of thresholds underlies rules for mixture interpretation.



Peak height ratios and thresholds control interpretation
Excessive variation in PHR makes this data uninterpretable

MiniStrs show the same effect

In spite of the improved sensitivity, peak balance is poor at low template concentration



MiniSTRs and LCN.

MiniSTRs were developed to access degraded DNA.

They do not solve the inherent low copy limitation of the PCR.

Instead because of their sensitivity, they complicate it.

1. Short PCR primers amplify better.
2. Better amplification means laboratories can access extremely low levels of DNA
3. At such levels (1-20 cells) a scientist cannot express a strong opinion about how DNA arrived at the site where it was recovered.
4. This DNA could just as easily come from pre or post deposition as it could come from the suspect.

The bottom line:

1. Low signal levels are bad because:
   a. They may indicate low copy # DNA = inconsistent or misleading results
   b. They often coincide with peak imbalance
   c. PCR and instrumental artifacts appear at these levels
2. Relying on signal level to determine DNA quantity can be misleading
   a. There is wide variation in signal strength of amplified DNA
   b. Inhibitors and mixtures complicate interpretation
      1. peak imbalance can occur even in single source samples due to inhibition and degradation
      2. instruments can vary in sensitivity

Fuzzy Logic in Data Interpretation

- The ABI 310 is a dynamic system
- Sensitivity varies with
  - Allele size
  - Injection solvent
  - Input DNA
  - Instrument factors
  - Presence of PCR inhibitors
  - Gel matrix
- Thus interpretation must be conservative and data from these studies yields guidelines, not rules. These guidelines must be based on in-house validation. In addition the interpretation and its significance cannot be dissociated from the overall facts of the case.

Proceeding with Testing when "No DNA" Detected

If the qPCR results indicate that there is no detectable DNA, will you stop testing or will you proceed with attempting STR typing?

- The practice of proceeding even with a "no result" Quantiblot was because the STR typing assay was more sensitive than the quantification method.
- Real time PCR should have solved this problem but it hasn't.

qPCR Assays Are Also Impacted by Stochastic Sampling in the LCN Region

Note the larger spread in these dilution series points for the LCN samples (16 pg) because of stochastic sampling

Quantifiler Kits User's Manual, p. 6-17

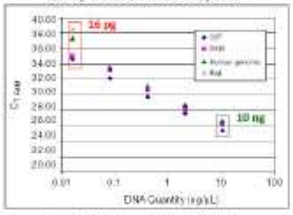


Figure 6-4 Sensitivity using the Quantifiler® Kit

Remember that DNA quantitation assays are also impacted by stochastic problems and may not be extremely reproducible on the low end, i.e., <100 pg

TECHNICAL NOTE

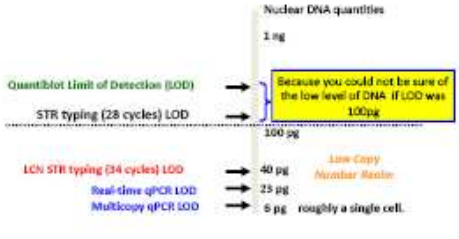
Catherine M. Cupples,[1] M.S.; Jarrod B. Champagne,[1] M.S.; Kristen E. Lewis,[1] M.S.; and Tracey Dawson Cruz,[2] Ph.D.

STR Profiles from DNA Samples with "Undetected" or Low Quantifiler™ Results

ABSTRACT: [illegible]

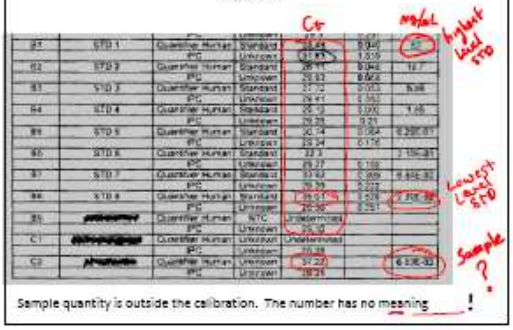
KEYWORDS: [illegible]

Difference in DNA Quantitation Capability vs. STR Typing Sensitivity



A typical amplification uses 10µL thus lowest quantity detected by qPCR that can be amplified with certainty is 250 pg single copy or 80pg multicopy

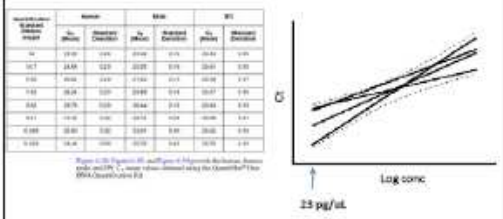
QPCR is a better way to assess quantity but be careful!.



Sample quantity is outside the calibration. The number has no meaning

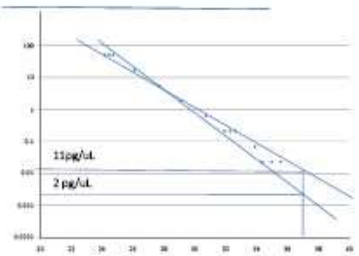
The Issue of calibration

ie: where exactly is 7 pg?



At levels below 23pg/uL you are below the calibration line and making it impossible to determine a true value. What is more, your lowest datapoint varies 2X as much.

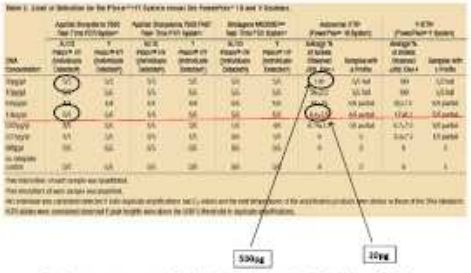
Effect of uncentainty in estimate



Outside the calibration curve, A 5 fold difference in the estimate is a reasonable possibility simply due to uncertainty in the slope of the line. The software reports the number as 6.83 pg/uL. How can it be accurate to three places? 6.83pg/uL?????? It could be 11 just as easily as it could be 2.

The difference between detection of amplification and reliability of result

Multilocus probes.



The Instrument reports 1.9pg/uL the analyst puts in 10 uL the result? An amplification of 20pg – Surely LCN. And the results show a partial profile.

RESEARCH HIGHLIGHTS

Ethics watch

LCN DNA: PROOF BEYOND REASONABLE DOUBT?

Low-copy-number (LCN) DNA forensic profiling has led to successful criminal prosecutions, including in the Peter Falconio case in Australia and the murder of the Swedish foreign minister, Anna Lindh. However, the technique has serious limitations, and few jurisdictions have followed the United Kingdom in accepting it as evidence in court. The discrediting of the LCN DNA evidence in the Omagh trial, which led the UK police to temporarily suspend their use of the method, has prompted further questioning of this technique and some scientists are claiming that criminal convictions based upon LCN DNA will soon start troubling the appeal courts.

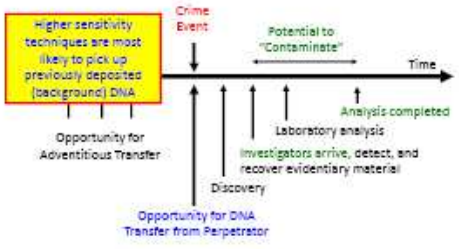
Issues:

Was evidence collected with LCN in mind?

Is LCN evidence reliable?

Does the obtained profile result from the evidence?

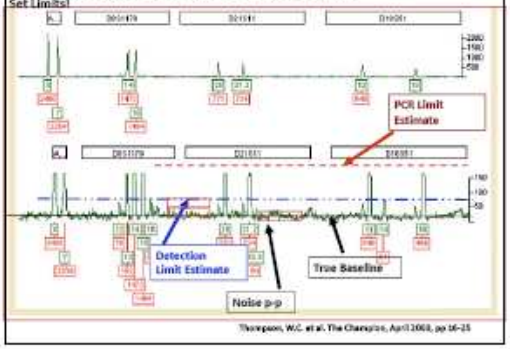
The problem with LCN DNA is that you cant be sure if it is from the case or if it is from some other event.



Adapted from Gill, P. (2002) BioTechniques 32(2): 366-385, Figure 5

How to interpret low level STR data safely?

Set Limits!



Thompson, W.C. et al. The Champion, April 2003, pp 16-25

Conclusions

- Be conservative in interpretation
  - Set thresholds based on signal to noise and stochaistic amplification (2 thresholds). Base these numbers on controlled in-house experiments
  - Understand that different instruments may vary in sensitivity – set thresholds high enough to encompass this variation
  - Understand that even with such guidelines issues such as degradation and inhibition can skew results.
- Leave room for the facts of the sample in your interpretation

Challenge with Being Able to Go Lower In DNA Quantitation Measurements

- Multi-copy marker (e.g., Alu assay) will be better than a single copy (e.g., Quantifiler) with qPCR of low quantity DNA samples
- qPCR enables measurement of lower amounts of DNA but…
- Going into the low copy number realm introduces new challenges
  - Interpretation of mixtures
  - Defining thresholds for different dyes and amplification systems
  - Defining the difference between investigative data and reliable "court-worthy" data

qPCR for DNA Quantitation

– will it lead more labs into LCN?
or
Are we already there and about to find out?

When properly used real time PCR can provide clearer information on absolute quantities of DNA

Why?

- Most laboratories will use amplification results to provide information on DNA thresholds
- But: The reason qPCR was developed is that using endpoint detection to determine quantity is notoriously imprecise
- Depending on amplification conditions, it is possible to produce full profiles from a single cell- 8pg.
- So wouldn't it be better to use qPCR to do so?

What else can go wrong?

Yarr,Take care mates!

- Most validation studies are performed on pristine samples derived from clean sources.
- DNA degradation will result in dropped alleles from larger sized amplicons
- DNA inhibition will result in dropped alleles from any location and the effects are difficult to predict
- Inhibition and degradation can produce stochastic effects – peak balance issues and allele dropout.

Fuzzy Logic in Data Interpretation

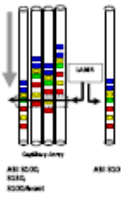
- The ABI 310/3100 is a dynamic system
- Sensitivity varies with
  - Allele size
  - Injection solvent
  - Input DNA
  - Instrument factors
  - Presence of PCR inhibitors
  - Gel matrix

- Thus interpretation must be conservative and data from these studies yields guidelines, not rules. These guidelines must be based on in-house validation. In addition the interpretation and its significance cannot be dissociated from the overall facts of the case.

Instrument factors

1. Because only signal is measured (RFUs) in forensic DNA analysis, many labs find that one instrument or another is more sensitive
2. There are also differences in sensitivity based on injection parameters, capillary illumination (single vs multiple) and laser intensity
3. Lastly the variation in qPCR sensitivity affects the output of any system
4. These differences should be corrected by proper setting of threshold parameters and/or adjustment of qPCR results.

310 vs 3100



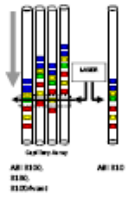
Sample
310 1.5uL in 24 uL formamide
31xx 1uL in 10uL formamide

Injection
310 5s@15kV = 75kVs
3130 (4 cap) 5s@3kV = 15kVs
3100(16cap) 10s@ 3kV = 30kVs

Irradiation
310 direct
3130 (4 cap) side
3100 (16 cap) both sides

Bottom line: you would expect to see
1. an approximate 3 fold difference in rfus between a 310 and a 3130 (4 cap)
2. an approximate 2 fold difference between a 310 and a 3130xl (16 cap)

Additional Issues



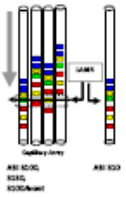
Threshold (ABI)
310 50 RFUs
31xx 30 RFUs

Stochastic
310 150 RFUs
31xx 90 RFU

Dynamic Range
310 4500
311 31xx 3500

Bottom line: 310 will appear more sensitive with a wider dynamic range unless proper validations are performed.

Bottom Line



Validate each class of instrument and expect differences in sensitivity/ signal to noise

Compensate for differences by choosing appropriate thresholds

Validate at 2 or more injection levels so that injection time can be increased- remembering that longer injections risk drifting into LCN regime

Calling thresholds involve sensitivity, dynamic range and the necessity to avoid LCN data

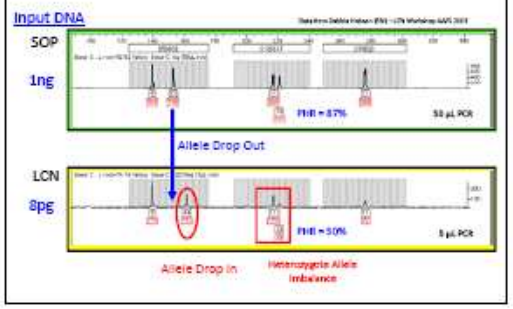
Issues with Data below the Stochastic Threshold

- PCR artifacts and stutter become prevalent
- Low levels of bleed through are possible
- Instrument spikes are more numerous
- -A peaks may appear
- Dye blobs become more significant in overall e-gram
- Low level 2nd contributors may show peaks

So why examine low level data at all?

- Detection of straddle data in which one allele is above threshold and the other is below
- Detection of the presence of low level mixtures
- Clues to the presence of inhibited samples or poor injections
- Aids in determination if a suspect is excluded as a contributor

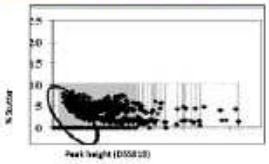
Comparison of STR Kit Amplification SOP with LCN Using the Same DNA Donor



Stutter and Peak Height Ratios change at LCN

- Stutter increases with allele size:
- Stutter increases at low copy #

Leclair et. al (2004) Systematic analysis of stutter percentages and allele peak height and peak area ratios at heterozygous STR loci for forensic casework and database samples. J. Forensic Sci. 49(5): 968-980



RESEARCH HIGHLIGHTS

Nature Reviews Genetics | AOP, published online 19 March 2008; doi:10.1038/nrg2362

Ethics watch

LCN DNA: PROOF BEYOND REASONABLE DOUBT?

Low copy number (LCN) DNA forensic profiling has led to successful criminal prosecutions, including in the Peter Falconio case in Australia and the murder of the Swedish foreign minister, Anna Lindh. However, the technique has serious limitations, and few jurisdictions have followed the United Kingdom in accepting it as evidence in court. The discrediting of the LCN DNA evidence in the Omagh trial, which led the UK police to temporarily suspend their use of the method, has prompted further questioning of this technique and some scientists are claiming that criminal convictions based upon LCN DNA will soon start troubling the appeal courts.

Issues:

Was evidence collected with LCN in mind?

Is LCN evidence reliable?

Does the obtained profile result from the evidence?

Low copy number situations exist in many samples

- In a 1:1 mixture, each DNA source is at LCN when the total amount of DNA in the amplification reaction is ~ 0.125 ng.
- In a 1:9 mixture, the minor component could be at LCN **even when the total amount of DNA in the amplification is 1 ng.**

Robin Cotton, AAFS 2003 LCN Workshop
"Are we already doing low copy number (LCN) DNA analysis?"

Findley *et al.* (1997) *Nature* article

Table 1 Details of analysis

Number of single cells analysed		226
Results obtained	206 (91%)	
Amplification failure		20 (9%)
Full STR profile		114 (50%)
Acceptable profile (amelogenin,>4 STRs)		144 (64%)
Partial profile (1-4 STRs)		62 (27%)
Surplus alleles*		26 (12%)
False alleles**		11 (5%)
Allele dropout		88 (39%)

*Additional allele present in conjunction with true alleles.

**Additional allele in place of true allele. Extra-allelic peaks could be caused by contamination, somatic mutation or PCR-generated non-allelic peaks. We never saw more than two additional peaks in a profile or in 10 together, minimizing the possibility of cellular contamination. When surplus alleles were observed we considered the locus, but not the profile, uninformative. We observed allele dropout in 39% of cells at a rate of ~10% in each allele. If two cells are analysed then the risk of allelic dropout and misinterpretation in cells is reduced to 1%. If three cells 0.1%, and so on. Wild-card designations and conservative statistical criteria are needed to ensure that evidential value can be properly assessed.

Some interpretational guidelines with LCN

- At least two* PCR amplifications from the same DNA extract

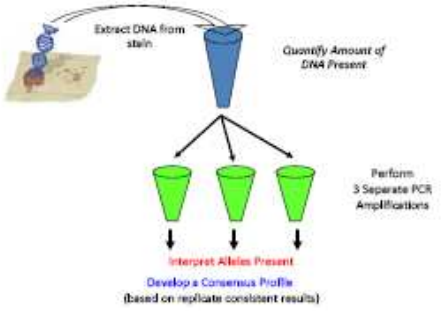
  *five is better; results are investigative
- An allele cannot be scored (considered real) unless it is present at least twice in replicate samples
- Extremely sterile environment is required for PCR setup to avoid contamination from laboratory personnel or other sources

Meatloaf Principle

- I want you
  I need you
  But -- there aint no way Im ever gonna love you
  Now dont be sad
  cause two out of three aint bad
  – Meatloaf

- You see an allele twice in 3 runs
- What if the the 4th measurement shows no allele?
- Is seeing an allele 50% of the time a measure of reliability. Is 66% ok?

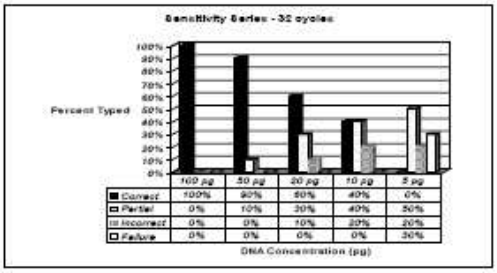
Typical LCN Procedure



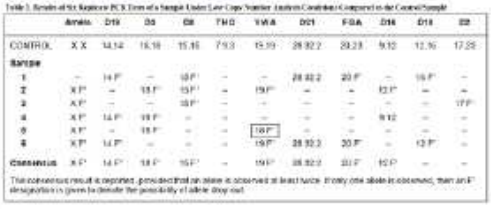
Catch 22

- Note the Catch 22. Are two amplifications of 50pg better than 1 of 100pg?
- Are 3 amplifications of 17pg better than one of 50?
- Data shows that the lower the amount of the DNA amplified the more likely allele dropout and false alleles occur
- This somewhat calls in to question the idea that a sample should be split and run multiple times

Problems with Obtaining Correct Allele Calls at Low DNA Levels



Replicate LCN Test Results from FSS



F used to designate that allele drop-out of a second allele cannot be discounted when only a single allele is observed (OCME uses "2")

Low Copy Number Limitations (cont):

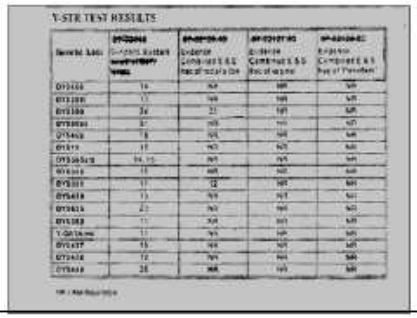
From Bruce Budowle (2005) 1st International Human Identification E-Symposium

- Tissue source cannot be determined
- DNA may not be relevant – casual contact/transfer
- If victim and suspect have any common access...
- Old cases may not be viable – handling
- Not for post conviction analysis
- Rarely useful for database searching
- An intellegence tool

http://www.e-symposium.com/humid/archive/drbrucebudowle.php

A Precautionary Tale

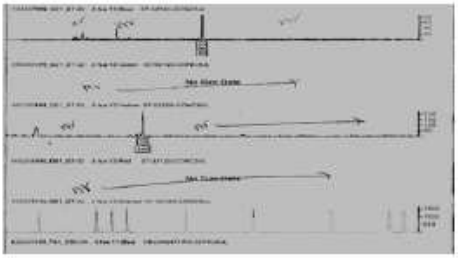
What does this mean?



What to look for in the QPCR results.

Sample is at a low level, nominally 6pg/uL but outside the calibration

Examine these results YFiler. Are these true alleles?



Look at height, relative position and ROX ladder

The Data

- Nuclear Results
  - Quantifiler
    - DNA detected below calibration
  - Identifiler
    - No results
  - Y Quantifiler
    - No result
  - Y STRs
    - One peak
    - Post PCR concentrate – 2 alleles

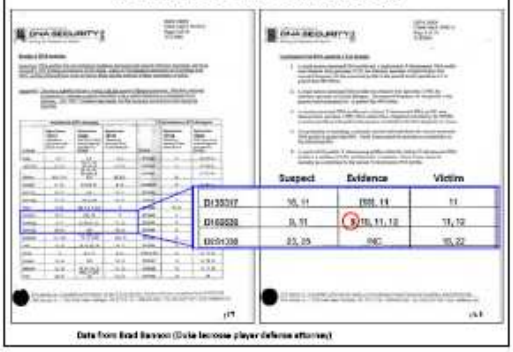
The report

- No nuclear profile due to insufficient or excessively degraded DNA
- Suspect is excluded based on results for 2 of 17 Y STR markers.
- Huh !?:
- My comments
  - 1. The result is clearly at low copy
  - 2. The pattern of alleles is not consistent with degradation as the cause of dropout.
  - 3. At low copy a scientist cannot express a strong opinion about how DNA arrived at the site where it was recovered. This DNA could just as easily come from thin air as it could come from the suspect.
- Bottom line: Why was this sample even run?

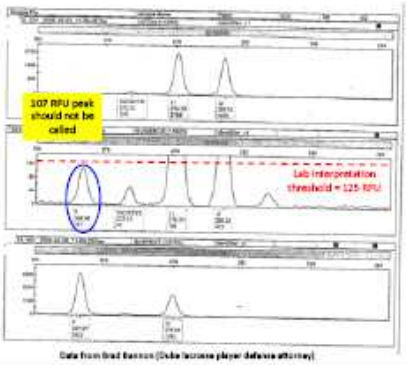
The bottom line:

1. Low signal levels are bad because:
   a. They may indicate low copy # DNA = inconsistent or misleading results – ie you cant determine when the DNA was placed
   b. They often coincide with peak imbalance
   c. PCR and instrumental artifacts appear at these levels
2. Relying on signal level to determine DNA quantity can be misleading
   a. There is wide variation in signal strength of amplified DNA, luckily the worst occurs at higher levels
   b. Inhibitors and mixtures complicate interpretation
      1. peak imbalance can occur even in single source samples due to inhibition and degradation
      2. instruments can vary in sensitivity

May 12, 2006: DNA Security Report



Data from Brad Benson (Duke lacrosse player defense attorney)



Data from Brad Benson (Duke lacrosse player defense attorney)

Other Issues – Real Time PCR

3. Real Time PCR follows the same rules as LCN DNA. Low levels of DNA will suffer from stochastic fluctuation, reducing precision, both for the sample and the calibration curve
4. Y chromosomal DNA suffers the most from LCN as there is exactly half as much DNA present in each cell
5. Multicopy real time PCR will provide lower detection thresholds

Conclusions

- Be conservative in interpretation
  - Set thresholds based on signal to noise and stochaistic amplification (2 thresholds). Base these numbers on controlled in-house experiments
  - Understand that different instruments may vary in sensitivity – set thresholds high enough to encompass this variation
  - Understand that even with such guidelines issues such as degradation and inhibition can skew results.
- Leave room for the facts of the sample in your interpretation

Y STRs and QPCR

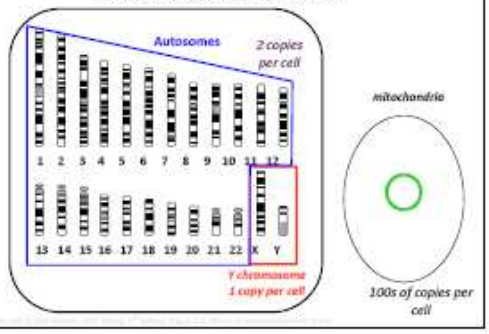
Bruce McCord
Professor Analytical/Forensic Chemistry
FIU Department of Chemistry
International Forensic Research Institute
Miami, FL 33199
mccordb@fiu.edu 305-348-7543

What if there is a mixture with a large amount of female DNA and a small amount of male DNA?

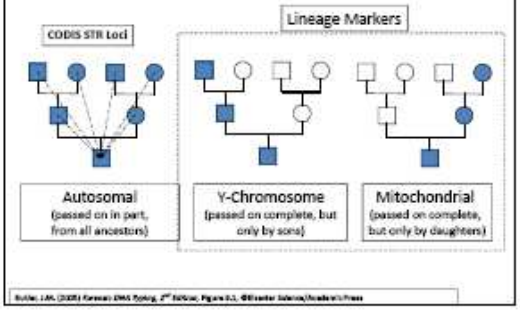
- Differential extractions may not work if
  - Large differences in DNA quantity between male and female exist in rape cases
  - Vasectimized males
  - Saliva samples
  - Such samples cry out for a male specific marker

Human Genome

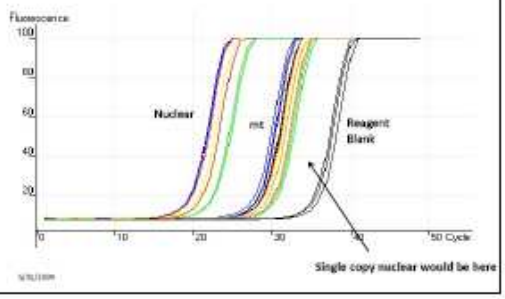
23 Pairs of Chromosomes + mtDNA



Different Inheritance Patterns



Sequential detection of mtDNA and multicopy nuclear DNA (Alu)



Role of Y-STRs (and mtDNA)

Compared to Autosomal STRs

- Autosomal STRs provide a higher power of discrimination and are the preferred method whenever possible
- Due to capabilities for male-specific amplification, Y-chromosome STRs (Y-STRs) can be useful in extreme female-male mixtures (bite marks, touch samples, etc.)
- Y STR typing can be very sensitive as the kit is designed to detect mixtures of DNA

Important Questions Can Be Answered with Y Chromosome Tests...

Apologies to George Lukas, John Butler and other B movie heroes

Y Chromosome STRs

Advantages:

1. Specific to males (inherited solely through the male-line descendants. Non recombining.
2. Differential extraction in rape kits is unnecessary.
3. Men commit the vast majority of violent crimes.
4. Because they don't recombine, Y chromosomes retain a unique genological record of mutations in their junk DNA

1. Like mtDNA, statistics are based on database size.
2. Mixed profiles can bs difficult to interpret statistically
3. There are odd issues with the Y chromosome: duplications, delections are possible – le
4. When sample limited choices must be made- autosomal, Y, mtDNA?

When will Y testing be used

1. When regular STR testing fails indicates little male DNA is present
2. When male saliva is collected from a female body
3. In a mixed blood stain – male cuts himself in a stabbing of a female
4. Fingernail clippings, touch samples
5. Ligatures around strangulation victim.

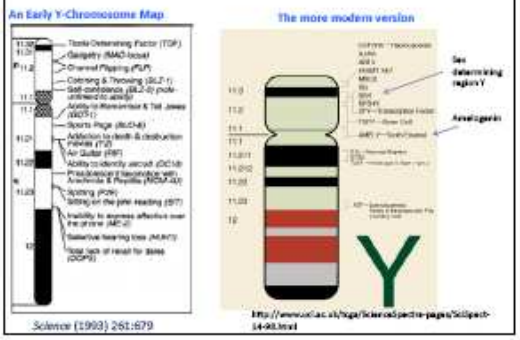
Sample collection will be important in such cases and QPCR will assist greatly in determining value

THE HUMAN Y CHROMOSOME:
AN EVOLUTIONARY MARKER COMES OF AGE
Mark A. Jobling & Chris Tyler-Smith
Nature Reviews Genetics (2003) 4, 598-612

Abstract

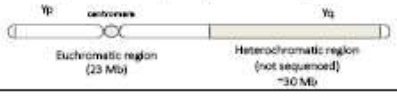
- Until recently, the Y chromosome seemed to fulfill the role of juvenile delinquent among human chromosomes — rich in junk, poor in useful attributes, reluctant to socialize with its neighbors and with an inescapable tendency to degenerate.
- The availability of the near-complete chromosome sequence, plus many new polymorphisms... now provide new avenues for investigating human evolution. Y-chromosome research is growing up.

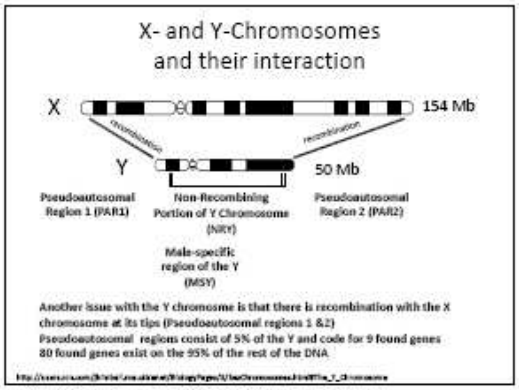
Traits found on the Y - Chromosome

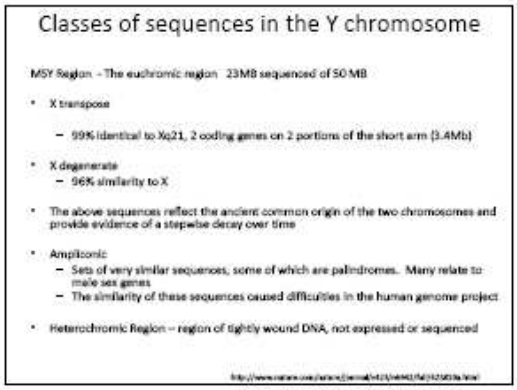


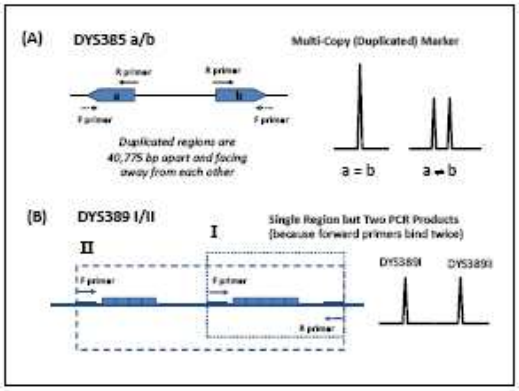
Issues surrounding the Y chromosome

- ca 50 Mbases long (third smallest after 21 and 22
- Contains 90 -300 genes
- 23 Mbases have been sequenced, the rest consists of highly repetitive DNA, difficult to sequence (Heterochromic region)
- 95% does not combine with the X chromosome, 2 regions at the tips however do, PAR1 and PAR2 (Pseudo autosomal region)
- The middle portion is known as the male specific region
  - It used to be called the non recombining region, but in fact it recombines with itself, and much sequence homology exists via back and forth gene conversion
- A variety of genetic diseases exist in which an X chromosome is damaged. This creates few problems to the mother but can be deadly to a male son. (hemophilia is an example)





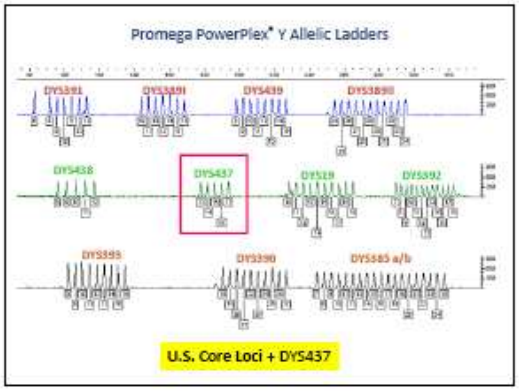


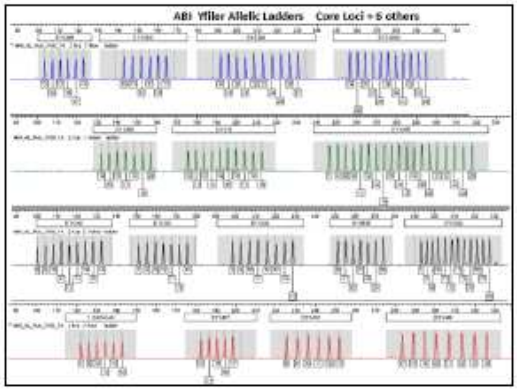


11 PCR products
9 primer sets

Core Y-STR Characteristics

STR Marker	Position (Mb)	Repeat Motif	Allele Range	Mutation Rate
DYS393	3.17	AGAT	8-17	0.05%
DYS19	10.12	TAGA	10-19	0.20%
DYS391	12.54	TCTA	6-14	0.40%
DYS439	12.95	AGAT	8-15	0.38%
DYS389 I/II	13.05	[TCTG] [TCTA]	9-17 / 24-34	0.20%, 0.31%
DYS438	13.38	TTTTC	6-14	0.09%
DYS390	15.71	[TCTA] [TCTG]	17-28	0.32%
DYS385 a/b	19.19, 19.23	GAAA	7-28	0.23%
DYS392	20.97	TAT	6-20	0.05%





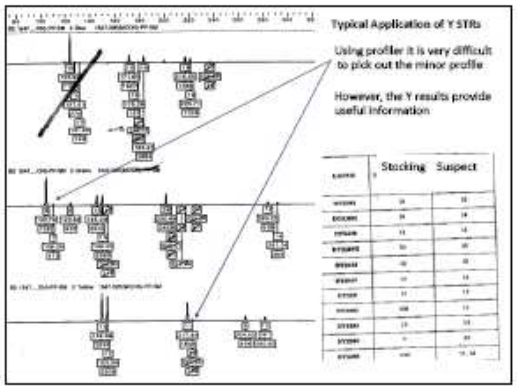
Difficult Questions...

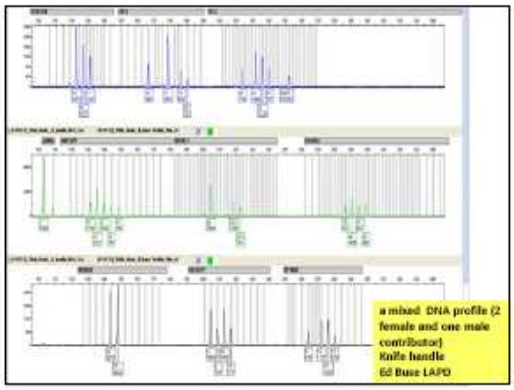
- Which database(s) should be used for Y-STR profile frequency estimate determination?
- Are any of the current forensic Y-STR databases truly adequate for reliable estimations of Y-STR haplotype frequencies?
  - Some individuals share identical Y-STR haplotypes due to recurrent mutations, not relatedness...
  - Is the database a random collection reflecting Y-STR haplotype frequencies of the population?
  - Is the Y-STR haplotype frequency relevant for the population of the suspect?

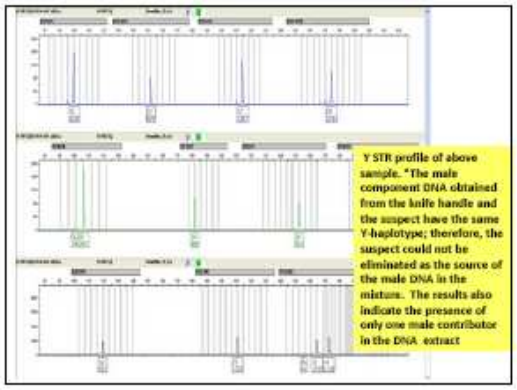
The Meaning of a Y-Chromosome Match

Conservative statement for a match report:

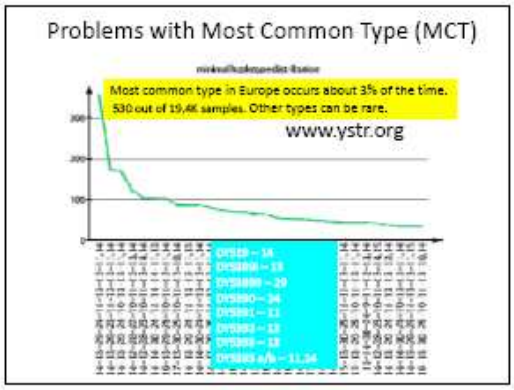
The Y-STR profile of the crime sample matches the Y-STR profile of the suspect (at *xxx* number of loci examined). Therefore, ***we cannot exclude the suspect*** as being the donor of the crime sample. In addition, we cannot exclude all patrilineal related male relatives and an unknown number of unrelated males as being the donor of the crime sample.

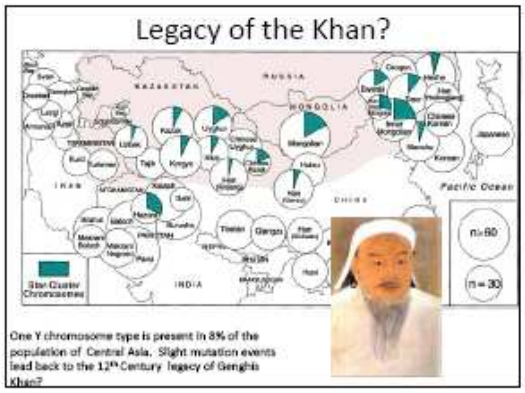


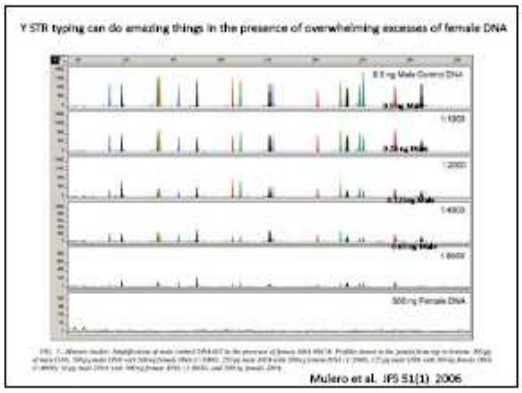


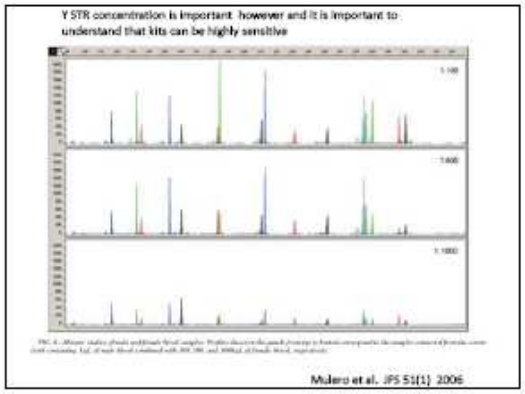


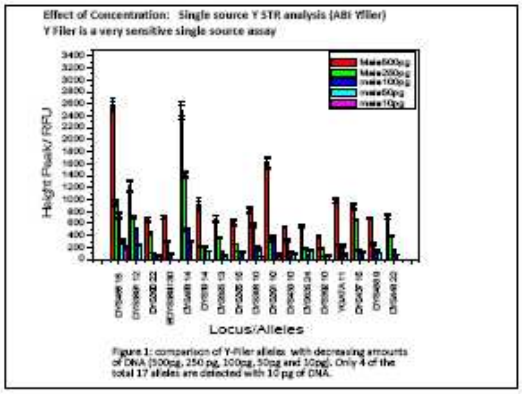
Problems with Most Common Type (MCT)

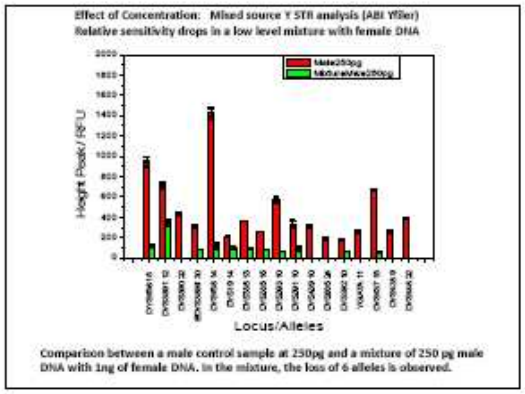












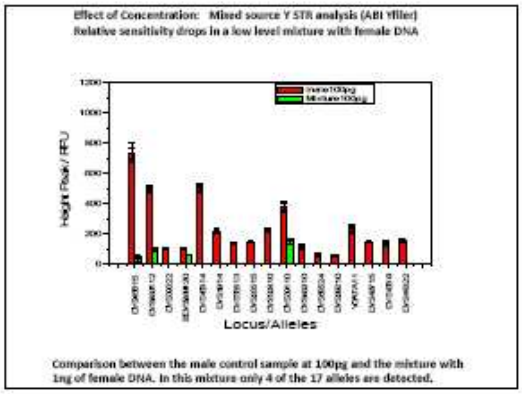


TABLE 4—*Amplification of male and female mixtures by using the Y-PLEX™ 12 kit.*

		Peak Height ratio of the Alleles for Male Sample in a Mixture Sample Containing DNA from Male and Female*					
Locus		1:0	1:100	1:200	1:400	1:600	1:800
DYS392		5337	1406	642	493	467	347
DYS390		2457	1692	1589	1257	1467	1963
DYS385 a/b		2839	1547	1282	843	1105	838
		3212	1705	1136	885	841	655
DYS393		3512	3081	3201	4277	3472	4012
DYS3891		1852	944	959	786	911	1003
DYS391		1733	1615	1167	805	1578	1206
DYS3901		1418	625	422	314	213	229
Amelogenin	X	1427	>5000	>5000	>5000	>5000	>5000
	Y	1466	176	93	75	<75	<75
DYS19		2221	1757	631	351	505	389
DYS439		1419	1630	1222	691	900	1075
DYS438		1936	548	386	195	251	184

* The mixture samples were prepared by taking 0.5 ng of male DNA and increasing the quantity of female DNA to obtain the indicated ratio.

Thewalt et al. Y Chromosome STR system, Y-Plex 12 kit for forensic casework: Development and validation. (FS 2005, 48)

So the Big Question:
When to Use Y? And how will QPCR help?

Modern QPCR kits – Quantifiler Duo and Plexor HY will provide a ratio of Autosomal to Male DNA

Laboratories will need to evaluate these kits in combination with their sensitivity thresholds

Issues will be:

1. What"s the Question? Is male DNA present ? Or Is a mixture present at some ratio?
2. Type of sample: differential extraction, bitemark, fingernail, touch /digital penetration sample
3. System sensitivity for autosomal mixtures ie: at what M/F ratio does it become impossible to recover/interpret a mixed profile? (note this ratio will change with total input DNA and A/Y ratio.
4. Availability of validated Y STR typing: in - house or outsourced? Would you ever not perform an autosomal STR profile?
5. Precision and sensitivity of estimate. Single copy assay may produce more precise ratios while multicopy assays will be more sensitive

Using Y quantification in casework

- Use it as a presumptive test for the presence of male DNA. - find and amplify the most probabtive samples
- Use it for an estimate of amplification success. If Y ratio is above 10% then Autosomal STR analysis may provide a useful result.
- Use it to estimate the amount of DNA template to get a useful Y STR result.

Improved Workflow

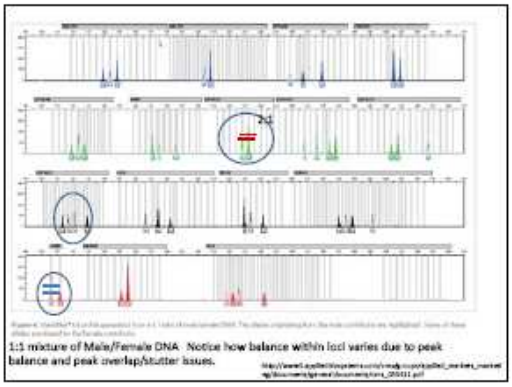
ABI Product Literature

TABLE 2. QUANTIFILER® DUO KIT RESULTS OBTAINED FROM AN APPROXIMATELY 1:1 MIXTURE OF MALE AND FEMALE GENOMIC DNA AND THE CORRESPONDING SINGLE SOURCE REFERENCE SAMPLES

Sample	Male DNA (SRY) Quantity (ng/µl)	Human DNA (RPPH1) Quantity (ng/µl)	Male:Female DNA Ratio
Male Reference	0.228	0.238	1:0.04
Evidence Sample	0.229	0.507	1:1.21
Female Reference	ND	0.217	0:1

ND = Not detected

Examining this sample, One can conclude that an approximate 1: 1 mixture exists.



1:1 mixture of Male/Female DNA. Notice how balance within loci varies due to peak balance and peak overlap/stutter issues.

From the Duo validation paper and Manual

Calculation of Male to Female DNA Ratio

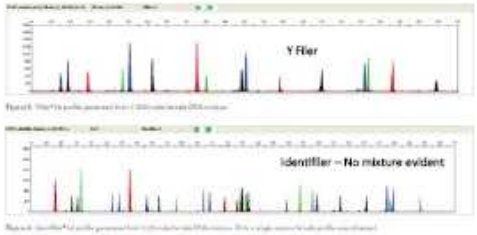
The Quantiflier® Duo kit provides the quantity of human and human male DNA in biological samples. From these values, one can calculate the ratio of male and female DNA using the following equation.

$$\text{Male DNA:Female DNA Ratio} = \frac{\text{Male DNA}}{\text{Male DNA}} \cdot \frac{(\text{Human DNA} - \text{Male DNA})}{\text{Male DNA}}$$

or

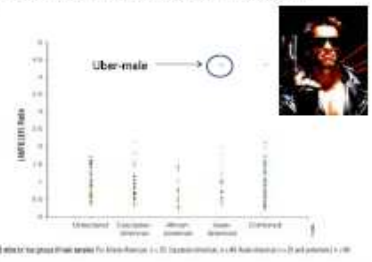
$$\text{Male DNA:Female DNA Ratio} = 1 : (\text{Human DNA} - \text{Male DNA})/\text{Male DNA}$$

All quantities in the above equations are ng/µL. This ratio determines the extent of the mixture, which is useful for making the choice of STR analysis method: autosomal STRs or Y-STRs.



The recovery of a male profile (upper egram) when no mixture is evident with Identifiler. (lower egram) Duo showed a 1/1000 male/female mixture.

Plexor also permits the relative detection of autosomal to Y DNA. Because the test uses multilocus probes, care must be taken to evaluate the possibility of Y chromosomal effects. Similar concern may exist with Duo, however such mutational events may be difficult to track.



Conclusions

- The Y Chromosome is a complex and interesting piece of DNA
- Y STR typing while not as valuable as autosomal can provide results and statistics based on the counting method
- Y DNA quantification can be used as a presumptive test for evidence screening,
- to determine mixtures and if autosomal DNA will work
- to determine the quantity of male DNA for Y STR analysis

Acknowledgements

- Heather LaSalle
- Silvia Zoppis
- George Duncan
- Eric Buel

Appendix 10: RT-PCR Introductory Examination

UNIVERSITY AT ALBANY
State University of New York

Northeast Regional Forensic Institute

"Providing Forensic Science Education for the Criminal Justice Community"
http://www.albany.edu/nerfi

NERFI: Introductory Real-Time Workshop

Name: _____
Date: _____

1. What is the purpose of the quantitation step?
5 pts

2. What are the components of the Quantifiler Kit?
5 pts

Reaction Mix –
Human DNA Standard –
Human Primer Mix –

3. Define C_T?
5 pts

4. What is the relationship between C_T and the concentration of input DNA?
5 pts

5. What can the IPC tell you?
5 pts

6. In the real-time quantification, a sample shows no amplification curve for either the sample or IPC. What could this indicate? How would you proceed?
5 pts

A. No DNA in sample - either re-extract more sample or call the analysis done.
B. Forgot to add IPC - re quant with IPC added
C. Inhibited sample - dilute out inhibitors with water or TE and re-quant
D. DNA in sample is degraded - end analysis on sample.

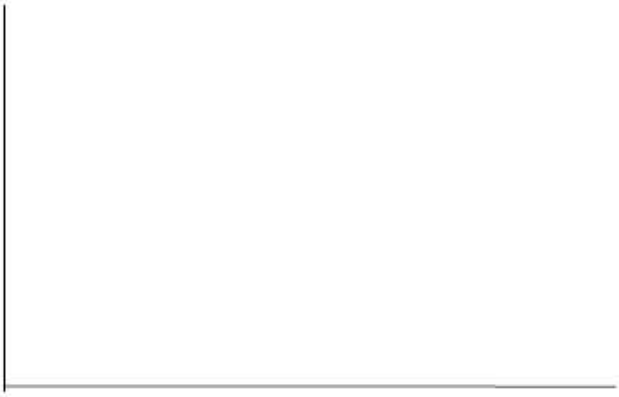
7. **In the real-time quantification, a sample shows no amplification curve for sample, but the IPC amplification curve is fine. What could this indicate? How would you proceed?**
5 pts

   A. Amplification failure – check 7500 and reagents, rerun
   B. Sample is inhibited - dilute out inhibitors with water or TE and rerun
   C. There is no human DNA in your sample - rerun quant or re-extract different part of stain.
   D. Standards made incorrectly - remake standards and rerun.

8. **The minor groove binder associated with the Quantifiler probe plays three roles, what are they?**
5 pts

1. **9.) What region of genomic DNA is amplified using the AB Quantifiler Human Real-time PCR Kit? Where is this region located? What size are the resulting amplicons?**
5 pts

10. Draw a standard curve and answer the following questions.
10pts



a. Label where the 50ng standard is located?
b. And the 23pg standard?
c. Describe the three criteria used to evaluate the standard curve:
   Slope:

   Y-intercept:

   Correlation Coefficient:

d. What values are ***actually*** measured by the real-time instrument?

e. How is the concentration determined for an unknown sample?

Appendix 11: RT-PCR Advanced Examination

Assessment Advanced issues in Real Time PCR Name_____

1. Explain the issue of PCR efficiency and its effects on real time measurement.

2. Why are calibrations based on the log of the concentration of template?

3. Explain the difference between Plexor and Quantifiler Duo

4. What is the effect of amplicon size on the ability to perform real time PCR

5. Explain two different applications for real time PCR melt curves

6. For what particular casework are Autosomal/Y ratios particularly useful

7. Explain the issues of Y STR sensitivity. Why are Y STRs assumed to be more sensitive if Y based real time results are poorer?

8. Explain this statement: Inhibitors affect PCR melt curves, efficiency and Ct and final product concentration

9. Why cant you quantitate DNA below its calibration curve?

10. Under what circumstances could a real time PCR result produce misleading results?

Appendix 12: RT-PCR Course Evaluations with Feedback

UNIVERSITY AT ALBANY
State University of New York

REAL-TIME PCR THEORY & CHEMISTRIES

Jan. 6, 7 & 8, 2010

Attendee Course Evaluation

Please fill out this evaluation form. Rate each of the aspects of this course by circling the appropriate number. Comment as appropriate.

Course Overall

	Poor				Excellent	Comments/Suggestions
A	1	2	3 19%	4 50%	5 31%	A lot of material in a short amount of time. Thorough look at RT-PCR & comparison of systems. Skip some of the elementary information. Familiar with introductory material, a few topics were informative. I wish Dr. McCord had more time!

Speaker: Lucy Davis

	Poor				Excellent	Comments/Suggestions
B	1	2	3 40%	4 45%	5 15%	Good concise overview of history of DNA quantitation. Spent a lot of time correctly what she previously said. Also spoke too slowly. Excellent at describing complex concepts and making them understandable.

Speaker: Bruce McCord

	Poor				Excellent	Comments/Suggestions
C	1	2	3 8%	4 42%	5 50%	Very interesting. Make sure slides match handouts, very distracting. 3 Informative slides provided multiple uses of Real-Time PCR.

Speaker: Jamie Belrose

	Poor				Excellent	Comments/Suggestions
D	1	2	3 10%	4 70%	5 20%	Seemed comfortable and very knowledgeable with subject. Speak slower. 3 Good explanation of calibration of real-time system. Very knowledgeable with subject matter.

Quality of Audio-Visual

	Poor				Excellent	Comments/Suggestions
E	1	2	3 19%	4 69%	5 12%	Some slides hard to read. Some slides had small font making it difficult to see. Some slides had too much information.

Quality of Handouts

	Poor				Excellent	Comments/Suggestions
F	1	2 4%	3 35%	4 46%	5 15%	Information out of sequence with handouts. 7 A little hard to read. 3 Small print.

Continued on back

Facilities

	Poor				Excellent	Comments/Suggestions
G	1	2	3 19%	4 62%	5 19%	Thank you for coming to see us.

Please Comment On:

Did the material presented meet your professional expectations?

8th only – Not particularly – I thought the material was covered in a far too hurried of a fashion. Too much information covered at too rapid of a pace to learn anything, followed by a far too difficult exam based on how the material was presented.
Yes, a lot information was provided.
Yes – exceeded. Great information.
8th only - Yes, I enjoyed McCord's lecture.
Yes. 13
Sure
Yes, I needed a refresher on the fundamentals of qPCR and this worked quite nicely.
Yes, very helpful information.
Yes, at time a bit over my head – but informative.

What topic or topics did you find most useful to you?

The information on qPCR inhibitors.
Melt curves, Plexor vs. Quant Duo, inhibition – everything really. 2
Inhibitors, DNA degradation – mini STRs, stochastic effects & thresholds.
The inhibition (Ct, Tm, etc.) 2
LCN 3
Background and theory of real-time and other possible uses of the system.
Problems with low level RT-PCR results & downstream applications
Understanding why it is necessary / the lab can benefit from other technologies of RT-PCR in analyzing low level samples.
Understanding more about inhibitors and how to read the different graphs on the RT-PCR instrument.
Troubleshooting, advances for future, current hot topics like low-copy.
LCN topics as it relates to interpretation of DNA profiles, mixtures and how to approach samples that fall outside the dynamic range of detection.
Taq-man explanation.
The theory of RT-PCR
There were interesting and applicable points in each of the talks. 2
Real-Time software
Although there weren't many – I enjoyed actual case examples to relate to the data and findings better.

Would you recommend this course to other scientists?

Yes. Absolutely. I would attend again too.
Yes. 17
Depends on their level of education – first 2 days only good for those just entering the field. Day 3 good for everyone.
Yes, good mix of overview and new topics.
Yes, it is very up to date.

General comments:

Lots of information in a small amount of time.
Worth the trip from Albany!
Thank you!
I don't think the tests need to be short answer. They should be brief and multiple choice.
Perhaps a little more time to finish the material. One of the talks jumped around a lot making it difficult to follow.
McCord's lecture could have been longer and the introductory material shorter.
Very informative.

Suggestions for future classes:

Mixture interpretation
Know your audience.
Maybe some time for troubleshooting and lab issues / casework.
CE Theory and Chemistry.

Appendix 13: Leadership Assessment Announcement

NERFI
The Northeast Regional Forensic Institute

Northeast Regional Forensic Institute
University at Albany
1400 Washington Avenue
Albany, New York 12222
Phone 518-437-3791
www.albany.edu/NERFI

John W. Hicks, Interim Director

Leadership Assessment: Developing the Next Generation of Leaders

Two-day Forensic Science Manager Workshops

January 14-15, 2010 **February 18-19, 2010**

March 18-19, 2010 **April 1-2, 2010**

Sponsors: Northeast Regional Forensic Institute (NERFI)
National Institute of Justice (NIJ), Grant Numbers 2008-DD-BX-K301, 2008-DNBXK-173

Instructors: Dr. Wendy S. Becker
Dr. Edward J. Pavur, Jr.

Location: Northeast Regional Forensic Institute, Albany, NY

Methods: Group discussion, case study, brainstorming, role play, readings, review of literature, surveys, integration of ideas, next steps

The ASCLD-LAB accreditation requires leaders to respond to a multitude of management issues that arise in labs. But many labs have limited opportunities and resources to provide leader development and training.

The purpose of this session is to provide leaders with an overview of tools and guidance for assessing and developing employees for supervisory and management positions. The interactive session is targeted to lab directors, supervisors, technical leaders and QA/QC managers.

Session facilitators will draw upon "best practices" research in the assessment, development and retention of forensic lab personnel. Topics include talent acquisition and retention, and staff development. Specific examples will be drawn from the facilitators' research including a national survey of public sector labs, a case study of laboratory recruitment and retention, use of a forensic advisory panel for development of forensic intellectual capital, and examples developing human resource metrics in labs. The session will be highly interactive and focused on topics of interest to session participants.

For reservations, contact John Hicks: johnhicks08@aol.com
Space limited to 10 for each workshop.

Appendix 14: Leadership Assessment Curriculum

Leadership Assessment: Developing the Next Generation of Leaders
February 18-19, 2010

Sponsored by the NorthEast Regional Forensic Institute and NIJ

Dr. Wendy Becker and Dr. Ed Pavur, Instructors

Agenda

Day 1 9:00 am – 5:00 pm

- Welcome – John Hicks, NERFI Director
- Introductions
- *A Case Study of Forensic Scientist Turnover* – Breakout and discussion
- Break
- Job Analysis: From Scientist to Supervisor – PP Presentation & discussion
- Lunch
- Challenges and Opportunities -- Discussion
- *Managing Intellectual Capital* – Breakout and discussion

Day 2 8:30 am – 3:30 pm

- NYSP FIC tour
- Leadership Skill Development – PP Presentations
- Lunch
- Leader Influencing Behaviors

Cited Works

Becker, WS, WM Dale. 2003. "Strategic Human Resource Management in the Forensic Science Laboratory." *Forensic Science Communications* 5(4).
http://www.fbi.gov/hq/lab/fsc/backissu/oct2003/2003_10_research01.htm

Becker, WS, WM Dale. 2007. "Critical Human Resource Issues: Scientists Under Pressure." *Forensic Science Communications* 9(2).
http://www.fbi.gov/hq/lab/fsc/backissu/april2007/research/2007_04_research02.htm

Dale, WM, WS Becker. 2004. "A Case Study of Forensic Scientist Turnover." *Forensic Science Communications* 6(3).
http://www.fbi.gov/hq/lab/fsc/backissu/july2004/research/2004_03_research04.htm

Dale, WM, WS Becker. 2005. "Managing Intellectual Capital." *Forensic Science Communications* 7(4).
http://www.fbi.gov/hq/lab/fsc/backissu/oct2005/research/2005_10_research02.htm

Yukl, GA, WS Becker. 2006. "Effective Empowerment in Organizations." *Organization Management Journal. Linking Theory & Practice: EAM White Papers*. 3(3), 208-228.

Yukl, G. 2009."Leading organizational learning: Reflections on theory and research." *The Leadership Quarterly*. 20, 49-53.

Department of Labor. http://online.onetcenter.org/

Appendix 15: Leadership Assessment Presentations

How Does the Supervisor or Manager of the Lab Differ from the Scientist?

Discussion Draft

Tables Derived in Part from the O*NET Database

Table 1. Forensic Analyst Activities

1. Documenting/Recording Information
2. Identifying Objects, Actions, and Events
3. Updating and Using Relevant Knowledge
4. Getting Information
5. Interacting With Computers
6. Making Decisions and Solving Problems: Analyzing Scenes, Samples and Situations
7. Communicating with Persons Outside the Organization
8. Interpreting the Meaning of Information for Others
9. Evaluating Information to Determine Compliance with Standards
10. Scheduling Work and Activities

Table 2. Forensic Lab Supervisor Activities

All of the activities for the analyst, plus:

11. Analyzing Data or Information: Principles, Patterns, and Facts
12. Communicating with Supervisors, Peers, or Subordinates: Receiving and Providing Information
13. Processing Information: Compiling, auditing, verifying

Leadership Assessment:
Developing the Next Generation of Lab Managers

Dr. Wendy S. Becker
Dr. Edward Pavur
Northeast Regional Forensic Institute (NERFI)
February 18-19, 2010

Leadership

Is leadership an attribution or impression; or is it related to effectiveness, emotion, or morality?

Leadership means

helping your group be more successful.

3

Leadership in the broad view has two aspects:

- initiation of structure, or *concern for production*;

and

- consideration, or *concern for people*.

4

Skill Practice

We will use
Leadership Challenges we have faced
to identify approaches to leadership problems
that will be productive.

Skill Practice Methods

- Case studies
- Analysis
- Practice (role-plays)
- Integration of Principles

What common challenges do leaders face in the crime lab?

- Selecting new group members.
- Evaluating performance; certifying competence.
- Contributing to planning and review for activities, objectives, and budgets.
- Implementing unpopular decisions.
- Making decisions.

What common challenges do leaders face in the crime lab?

- Listening to concerns; reacting to requests.
- Resolving conflicts.
- Helping the group meet objectives.
- Developing talent.

What common challenges do leaders face in the crime lab?

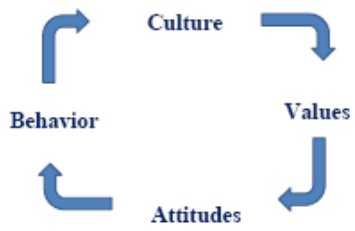
- Balancing mentoring with lab caseload, backlogs, and efficiency.
- Coordinating lab work with the objectives of outside agencies.

What other common challenges for leaders have we missed?

Attitudes and Performance

Are they related?

Influences on Work Behavior



Attitudes and Performance

- Work attitudes have a reliable relation to job performance.
- Judge 2001, *Psychological Bulletin* analyzed 25 years of research, including 11,155 articles.
- Results showed that **job satisfaction** predicts **individual performance.**
- *r* = .18 raw, .30 corrected, *n* = 54,417.

Individual Employees

General Work Attitudes	Job Performance
satisfaction involvement commitment engagement	efficiency quality productivity loyalty

Engagement and Business Unit Performance.

- Harter 2002, *Journal of Applied Psychology*
- Reviewed results from 7,939 business units and found **engagement** related to **composite business unit performance.**
- *r* = .22 raw, .54 to .63 corrected.

Engagement and Business Unit Performance

- Harter, Wagner, & Hunter (2007). *12: The Elements of Great Managing.*
- Used the Gallup archives, 1930s to 2006
- 10 million interviews from 124 countries, in 45 languages.
- Identified 23,910 business units with measures of engagement and business unit performance.

First, they found
key differences in engagement.

- Engaged: These people look out for you.
- Not Engaged: These people coast.
- Actively Disengaged: These people get even.
- Over time, people become less engaged.

Engagement and Tenure

	Tenure: 0 years	Tenure: 10 years
Engaged	39%	29%
Not engaged	49%	51%
Actively disengaged	12%	20%

Engagement is related to:

- **High customer ratings**
- **High productivity**
- **High profitability**
- Better safety
- Slightly greater creativity
- Low cheating and theft
- Low turnover
- Low absenteeism

Second, they found
key differences in managers

- Effective managers get the work done with the people they have,
- they don't try to change their personalities,
- but they capitalize on the skills that exist, not on what the managers wished their employees could do.

Effective Managers

- Effective managers focus on work elements that the manager can influence.
- For example, rather than trying to make Mary a more conscientious person, the manager recognizes Mary's best work.

Third, they found
some work conditions were related to employee engagement.

How does this apply to the forensic lab?

- Forensic Lab Managers might be able to influence the attitudes held by some of the employees in their labs.
- As a Manager, you can focus on a few high-impact areas for the Forensic Scientists who work for you.

For Example?

- Initiate structure
  - **Quality**
  - **Progress**
  - **Self development**
- Consideration
  - **Recognition**
  - **Valuing opinions**
  - **Caring about employees as people**

"My Coworkers are Committed to Doing Quality Work."

Encourage Coworkers to be Committed to Doing Quality Work.

"We value the same things, and those things lead to quality results."

What examples do you have?

What opportunities exist in your lab for improving attitudes?

Leadership Perspectives

Leadership means

helping your group
be more successful.

Leadership in the broad view has two aspects:

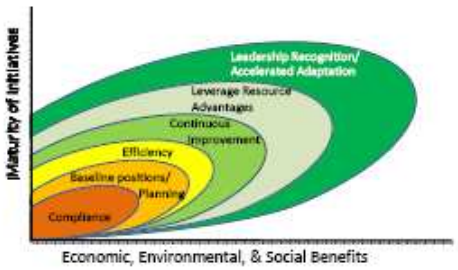
- initiation of structure,
  or *concern for production*;

and

- consideration,
  or *concern for people*.

Leadership and Results

Grow from Compliance to Quality to Innovation



Classic Leadership Viewpoints

- Henri Fayol
  - Administration performs five Functions.
- Mary Parker Follett
  - "Responsibility resides in the situation, not in people."
- Frederick Taylor
  - Management is responsible for systematic operations.

Most Influential

Three specialist groups ranked 71 key contributors to US Management. Specialist affiliations:

1) Business History and Economics,
2) Academy of Management, and
3) Management History.

Wren, D., & Hay, R. (1977). Management historians and business historians: Differing perceptions of pioneer contributors. *Academy of Management Journal, 20*, 470-476.

Who contributed most to American business and management thought and practice in the past 200 years?

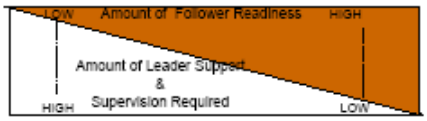
Number of first place votes:

51 Frederick Taylor
9 Chester Bernard
6 Thomas Edison
4 Henry Ford

Situational Leadership Theory

Tannenbaum and Schmidt (1973)
Hersey and Blanchard (1988)

A contingency theory that focuses on followers' readiness; the more "capable" the followers (the more willing and able) the less the need for leader support and supervision.



General Leader Functions

Gary A. Yukl (1987)

- Covariance among managers' behaviors (factor analysis)
- Similar content/purpose of managers' duties (judgment classification)
- Theory of managers' functions (deductive analysis)

Yukl (medium view)	General Description (broad view)	Stogdill, Ohio State (macro view)
Supporting	Participation, Motivation, Support	Consideration
Consulting		
Delegating		
Recognizing		
Rewarding		
Motivating		
Managing Conflict, Building Teams	Achievement, Constructive conflict [Follett]	
Developing		
Clarifying	Management Functions: Plan, Organize, Command, Coordinate, Control [Fayol] Organize and Monitor [Taylor]	Initiation of Structure
Planning & Organizing		
Problem Solving		
Informing		
Monitoring		
Representing	External liaison	
Networking, Interfacing		

Political Behavior

Influencing or attempting to influence the allocation of rewards in an organization.

- Legitimate
- Illegitimate

Illegitimate Political Behavior

- Deliberate, ruthless manipulation
- Extreme pressure
- Games
- Debasement

Games are a repeated series of exchanges used by an actor on a target, in which the surface intentions differ from the underlying intentions. The exchanges seem plausible, but contain a hidden agenda.

The actor may exert influence in order to manipulate, control, weaken, or retaliate against the target.

What Makes Some Kinds of Influence "Illegitimate?"

- Malevolence
- Ambition outweighs judgment
- "My own interests are more important than those of the organization."

Power, Influence

- The ability to get someone to do something you want.
- Bases of Power:
  - Formal, Legitimate
  - Informal
  - Coercive, Reward
  - Expert
  - Referent (charisma, status)

Influence Tactics

1. Rational Persuasion
2. Pressure/Assertiveness
3. Upward Appeal
4. Exchange
5. Ingratiation
6. Coalition Building
7. Inspirational appeal
8. Consultation

Influence Tactics
Long-term Effectiveness in the US

Higher
- Legitimacy
- Rational persuasion
- Inspirational appeals
- Consultation

Mid to Lower
- Exchange
- Personal appeals
- Ingratiation
- Pressure
- Coalitions

Loyalty and The Management Role

Managers owe a high degree of loyalty to the organization

Faithless Servant Doctrine

"[a]n agent is held to uberrima fides in his dealings with his principal, and if he acts adversely to his employer in any part of the transaction, or omits to disclose any interest which would naturally influence his conduct in dealing with the subject of the employment, it amounts to such a fraud upon the principal, as to forfeit any right to compensation for services." Murray v. Beard, 102 N.Y. 505, 508 (1886).

Astra USA, Inc. v. Bildman, Case No.SJC-10361 (Mass. Oct. 5, 2009).

The chief executive committed misconduct that included tax evasion, widespread sexual harassment of employees, and attempts to cover up the harassment.

Initial Case

A MA Superior Court jury on January 30, 2006 found Lars P.E.Bildman liable to his former employer, Astra USA, Inc. (Astra), for fraud, conversion, waste of corporate assets, breach of fiduciary duty, and sexual harassment of Astra employees, and awarded Astra damages in the aggregate amount of $1,040,812.

Result of Appeal

Supreme Judicial Court Boston, MA (SJC-10361) on October 5, 2009, granted Astra recovery of compensation it paid to Bildman during the period of his disloyalty -- $5,599,097 in salary and $1,180,000 in bonuses.

Leadership

Decision Making

Vroom-Yetton-Jago Normative Decision Model

Should my employees participate in this decision?

Participation

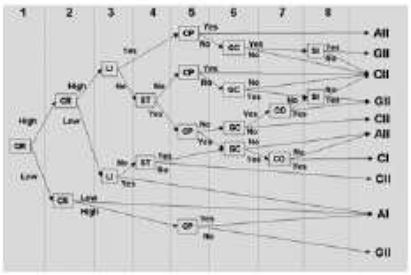
- Can increase acceptance of decisions
- Can uncover resistance
- Can improve quality of decisions
- Requires time and effort
- Can create conflict
- Can result in frustration
- Can be stressful

Decision Features 1-5

- Is technical quality important (QR)?
- Is employee commitment important (CR)?
- Do you have the information to make a high quality decision on your own (LI)?
- Is the problem well structured (ST)?
- If you made the decision alone, would employees accept it (CP)?

Decision Features 6-8

- Do employees share the organizational goals in solving the problem (GC)?
- Are employees likely to be in conflict about solutions (CO)?
- Do employees have enough information to make a high quality decision (SI)?



Decision Making Style	Description
Autocratic I (AI)	Leader solves the problem alone using information that is readily available to him/her
Autocratic II (AII)	Leader obtains additional information from employees, then makes decision alone. Employees don't supply or evaluate solutions, and may or may not be informed of the decision.
Consultative I (CI)	Leader shares problem with employees individually, and asks for information and evaluation. Employees do not meet as a group, leader makes decision alone. Decision may or may not reflect group influence.
Consultative II (CII)	Leader shares problem with employees as a group, but makes decision alone. Decision may or may not reflect group influence.
Group II (GII)	Leader meets with group to discuss situation. Leader focuses and directs discussion, but does not decide. Group makes final decision. Leader accepts group decision.

Leadership Assessment: Developing the Next Generation of Lab Managers

Dr. Wendy S. Becker
Dr. Edward Pavur
Northeast Regional Forensic Institute (NERFI)
February 18-19, 2010

Strengthening Forensic Science in the United States: A Path Forward

"Scores of talented and dedicated people serve the forensic science community, performing vitally important work. However, they are often constrained by lack of adequate resources, sound policies, and national support.

It is clear that change and advancements, both systematic and scientific, are needed in a number of forensic science disciplines to ensure the reliability of work, establish enforceable standards, and promote best practices with consistent application."

Committee on Identifying the Needs of the Forensic Sciences Community; Committee on Applied and Theoretical Statistics; National Research Council

---

Purpose and Importance

- ASCLD-LAB accreditation requires leaders to respond to a multitude of management issues that arise in labs.
- Labs have limited opportunities and resources to provide leader development and training.

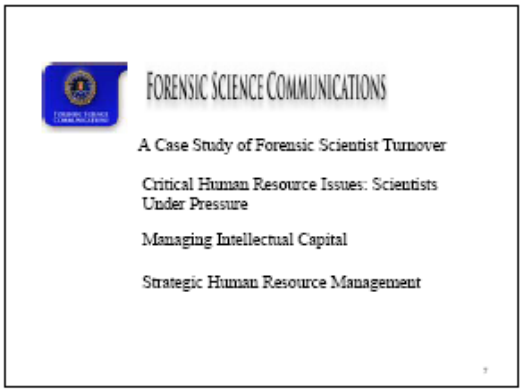
Purpose and Importance

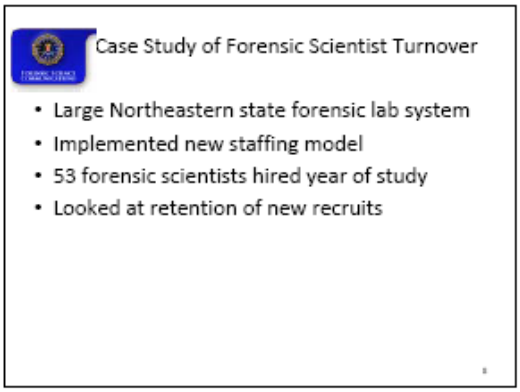
- To identify management tools to help develop managers of technology-intensive jobs
- Participants: lab directors, supervisors, technical leaders, QA/QC managers
- "Best practices" research in the assessment, development and retention of forensic lab personnel. Topics include talent acquisition, retention and development

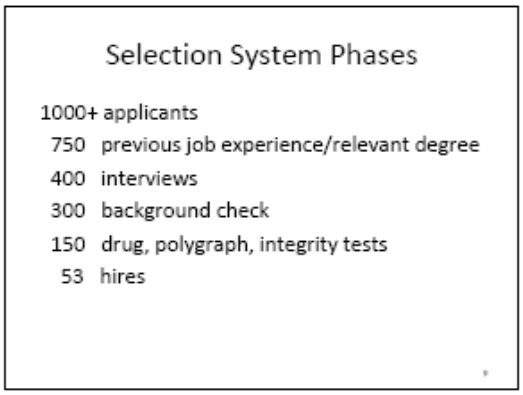
---

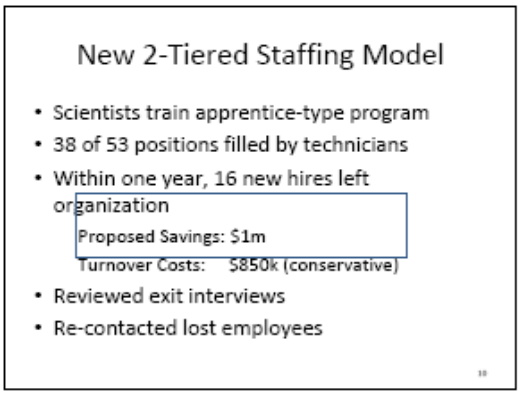
Agenda

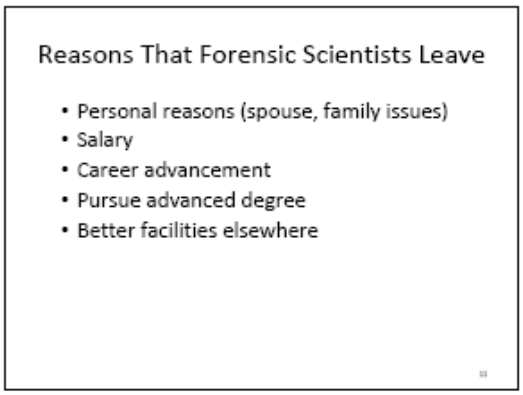
- Thursday, Feb 18
- Friday, Feb 19

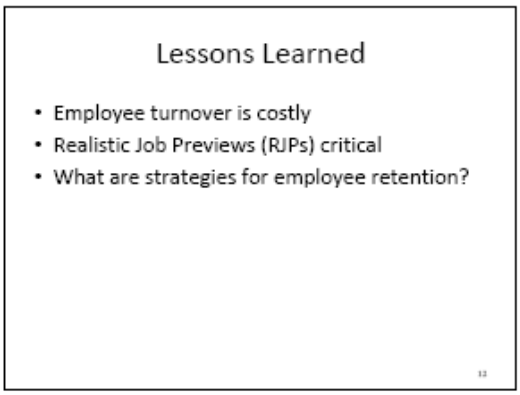












Skill Practice

Critical Human Resource Issues: Scientists Under Pressure

National Survey

Document basic staffing issues in public crime labs

Retention strategies

Lab capacity / level of outsourcing

Performance pressure on scientists / capacity

Method

- Web-based survey
- 250 American Society of Crime Laboratory Directors (ASCLD)
- 46 items grouped into six sections: demographics, caseload, recruitment, turnover, retention and performance issues
- Reminder email two weeks after the initial request
- Follow-up phone interviews

Survey Results

- 55 responses (22%)
- Populations served range from 59,000 to 22 million
- Average system 5.4 labs
- Pay (non-supervisor scientists) $28,800-116,000; average $59,087.
- Age of oldest case 0 to 480 months; average 28 months.

Survey Results …	Percentage	
	Yes	No
Do you … measure productivity?	89.8	10.4
Provide career development opportunities?	89.4	10.6
Need more tools to increase productivity and quality?	90.0	10.0
Have interns?	89.4	10.6
Are retention methods successful?	77.6	22.4
Lose scientists to public sector?	90.0	10.0
Lose scientists to private sector?	89.8	10.0
Have multiple labs?	89.8	10.2
Maintain a DNA convicted offender database?	89.8	10.1
Have shadow programs for students?	89.7	10.3
Is turnover a problem?	90.0	10.0
Would send more cases to private labs if had the funding?	28.8	71.1
Have a sufficient number of scientists needed?	21.2	78.8
Have shortage of applicants?	9.8	90.0

Pressure to Perform Increases With # Cases

Forensic scientists are pressured to complete cases in a timely manner	.282*
Forensic scientists are pressured to complete cases too quickly	.391**
Forensic scientists are pressured to get a particular result	.355**

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Lessons Learned

- First national survey to attempt to identify critical staffing issues faced in public forensic science labs
- Serious staff shortages
- Labs need additional scientists to meet our standard of one scientist per 30,000 population.
- Relationship between current staff capacity and amount of outsourcing cases to private labs.

Lessons Learned

- Interestingly, in many cases one would predict that as productivity increases, the pressure to complete a case would decrease.
- However, we found relationship between capacity and pressure to perform
- Vicious cycle; as capacity increases … having time/resources decrease:
  - having the proper equipment to do the job
  - having enough time to perform the job
  - having adequate resources to do the job
  - having enough time to prepare for courtroom testimony

Discussion

- Two major reasons for scientist turnover
  - personal reasons and salary
- This is cause for concern, as public labs typically have extensive recruitment phases; in larger organizations, recruitment and selection take as long as 12 months to cycle through.
- Identified retention strategies
  - hire people with a link to the local area
  - provide flexible work hours
  - further education
  - train scientists in multiple disciplines
  - opportunity to transfer to other labs

List of possible lab measures

- Cases/items analyzed per laboratory
- Cases/items analyzed per scientist
- Rate of local, state, and national ballistics hits in the National Integrated Ballistics Imaging Network (NIBIN) per firearms examiner and per capita of service region
- Rate of local, state, and national latent fingerprint hits in the Automated Fingerprint Identification System (AFIS) per fingerprint examiner and per capita of service region
- Rate of local, state, and national DNA hits in the Combined DNA Index System (CODIS) per DNA scientist and per capita of service region
- Number of technical support personnel
- Total cost of analyses per case and per item
- Total cost of casework
- Turnover of line staff and supervisors
- Quality system inventory, including:
  - Number of corrective actions
  - Number of types and frequency of corrective actions per discipline over time
  - Number of errors per case, per item
  - Timeliness of analyses
  - Total backlog
  - Complaints, both internal and external

Source: Durk & Peerbo, 2005

Leadership Research

Most Common Leadership Problems Public Sector

Cost cutting/staff reduction
Workload increase/workforce freeze and reduction
Excessive health care costs and sick leave
Training program evaluation
EEO compliance
Affirmative action/diversity programs -- line management resistance
Computer reporting programs
Competitive salaries in technical areas
Benefit administration
Computer training
Safety programs
Insufficient funding
Childcare
Terminating ineffective employees
Injury compensation costs
Union demands
Invalid performance appraisals
Ineffective pay for performance

(Source: Bernardin, 2010)

Skill Practice

As Leaders We Are Not Always Successful

- ***Attempted*** leadership – Person A attempts to change person B's behavior
- ***Successful*** leadership – Person B changes his/ her behavior as a function of person A's efforts
- ***Effective*** leadership – Person B changes behavior as a result of person A's efforts; person B more satisfied, better rewarded and attains a goal of mutual importance to person A and B

(Source: Bass, 1960)

Think of the differences between a time …

- … when you ***attempted to lead*** a person or group
- …when you ***successfully changed the behavior of*** a person or group
- … when you ***effectively changed the behavior of*** a person or group

Leadership – *More Complex* than Managing or Supervising

- "Manager" and "Supervisor" are job titles
  - Imply tasks or duties
  - Describe "what" needs to be done
- Leadership refers to social-psychological aspect of the role of supervisor/managers
  - Describes "how" tasks/duties carried out
  - Implies "change in behavior"
  - Depends on "followers" executing plans

Most people can become effective leaders in the right circumstances

- Leadership is "more than" possession of specific traits
- Modern theories of leadership *blend* duties of managers with effective leadership

18 Broad Dimensions of Managerial Responsibility

Guiding, directing, motivating subordinates	Communicating effectively and keeping others informed
Training, coaching, developing subordinates	Representing the organization to the public
Delegating	Persisting to reach goals
Influencing/selling	Handling crises
Maintaining good working relationships	Organizational commitment
Coordinating subordinates and other resources to get the job done	Monitoring and controlling resources
Planning & organizing	Technical proficiency
Decision making and problem solving	Administration and paperwork
Staffing	Collecting and interpreting data

(Source: Borman & Brush, 1993)

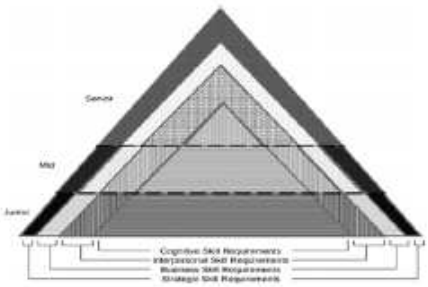
Leadership/Management Connection: Broad Dimensions

- Note that 6/18 dimensions related to leadership
  - *Effectively influencing* others
- Borman & Brush (1993) research based on wide range occupational settings: police departments, manufacturing, hospitals, universities, armed services, high-tech firms

Leadership/Management Connection: Skill Requirements

- *Strategic*
  - Highly conceptual skills needed to take a systems perspective to understand complexity, deal with ambiguity and influence others
  - Skills: visioning, systems perception, identification of downstream consequences, identification of key causes, problem identification, solution appraisal, objective evaluation
- *Business*
  - Specific skills related to specific functional context
  - Skills: management of material, personnel, financial resources, operations analysis
- *Interpersonal*
  - Interacting with and influencing others
  - Skills: social perceptiveness, coordination, negotiation, persuasion
- *Cognitive*
  - Foundation of leadership
  - Skills: collecting, processing, disseminating info; learning (speaking, active listening), writing, reading comprehension; adapting; critical thinking

(Source: Mumford et al., 2007)



Skill Requirements at Managerial Levels (Mumford et al. 2007)

Leadership/Management Connection: Skill Requirements

- As one moves up the managerial ladder:
  - *cognitive and business skills become less important*
  - *interpersonal and strategic skills become more important*
  - *both business and strategic skills increase with experience*

Managing Intellectual Capital

- If you can Measure it, you can Manage it
- Forensic Advisory Board

Recommendations

Employee Performance Measures

1. Cases/items analyzed per scientist/per project team/per laboratory
2. Ratio of local, state, and national ballistic hits in the National Integrated Ballistics Imaging Network per firearms examiner and per capita of service region
3. Ratio of local, state, and national latent fingerprint hits in the Automated Fingerprint Identification System per fingerprint examiner and per capita of service region
4. Ratio of local, state, and national DNA hits in the Combined DNA Index System per DNA scientist and per capita of service region
5. Ratio of technical support personnel per capita of service region (1 FS/50,000 pop)
6. Total cost of analyses per case and per item ($750/standard DNA profile)
7. Total cost of errors, for example, rework
8. Employee turnover
9. Quality system measures, including:
   a. Number of corrective actions
   b. Number of types and frequency of corrective actions per discipline over time
   c. Number of errors per case, per item
   d. Timeliness of analyses
   e. Total backlog

Combined DNA Index System (CODIS) Statistics

Source: fbi.gov/hq/lab/codis/stats

Forensic Advisory Board

- Simulates private sector Board; creates surrogate for natural movement
- Steve Kerr (Goldman Sachs) moved managing directors onto boards, created joint venture with Harvard to give people board training, set up action experiments, job shadowing (ex: Hewlett-Packard)
- Intellectual capital from experienced active and retired professionals various disciplines & academia; regular audits using ASCLD/LAB criteria;
- Review meetings as follow-up for corrective action; regular phone calls, emails, creates new organizational culture of collaboration between staff and advisory group

Implications/Future Research

- Increased reliance on private labs raises several critical questions.
  - Trends toward outsourcing and privatization impact public labs?
  - Will public labs outsource only the routine, redundant cases? Or will continued budget shortfalls ultimately bring about a downsizing or even the demise of public labs?
  - Should a different pay structure be implemented for DNA analyses that take these issues into consideration?
  - Should a price be put on DNA analyses that can include or exclude a defendant or free a convicted offender?
- We hope to continue dialogue in the forensic science community for these important issues.

Organization Management Journal — Effective Empowerment in Organizations

Condition	Hallmarks	Examples
Organizational Structure	Highly centralized formal structure; low cost, standard product or service	Decentralized and low formalization; customized, highly differentiated product or service
Organizational culture	Reliable, efficient operations do not allow mistakes; internal politics, criticism of new ideas; destructive internal competition; avoidance of risk; overemphasis on status quo	Flexibility, learning, participation; fair, constructive judgment of ideas; reward and recognition; mechanisms for developing new ideas; an active flow of ideas; and shared vision
Job design	Simple, repetitive tasks with technology dictating workflow; brief customer transactions that take place in a short time interval	Complex, non-routine and challenging tasks; flexible technology; repeated customer interaction in a continuing relationship
Access to resources	Resources are scarce or non-existent	Access to appropriate resources, funds, materials, facilities, and information

Employee rewards and ownership	None or very little	Employees shareholders or co-owners & invested in organization's success
Employee traits and skills	Low achievement motivation; low self confidence; and an external locus of control orientation	Low skill employees benefit more from empowerment efforts; employees with high need for achievement; high self confidence and self efficacy; and an internal locus of control orientation
Autonomy	Employees lack freedom in deciding how work is done and lack control over work	Employees have freedom deciding what work to do & how to do it; employees have a sense of control over work
Mutual trust	Low	High
Leaders as role models	Leaders do not model empowering behaviors	Leaders serve as role models, set appropriate goals, support work group, value individual contributions, show confidence

Guidelines for Participative Leadership

How to diagnose decision situations

- Evaluate how important the decision is
- Identify people with relevant knowledge or expertise
- Evaluate likely cooperation by participants
- Evaluate likely acceptance without participation
- Evaluate whether it is feasible to hold a meeting

How to Encourage Participation

- Encourage people to express their concerns
- Describe a proposal as tentative
- Record ideas and suggestions
- Look for ways to build on ideas and suggestions
- Be tactful in expressing concerns about a suggestion
- Listen to dissenting views without getting defensive
- Try to utilize suggestions and deal with concerns
- Show appreciation for suggestions

What to Delegate

- Tasks that can be done better by a subordinate
- Tasks that are urgent but not high priority
- Tasks that are relevant to a subordinate's career
- Tasks of appropriate difficulty
- Both pleasant and unpleasant tasks
- Tasks that are not central to the manager's role

How to Delegate

- Specify responsibilities clearly
- Provide adequate authority and specify limits of discretion
- Specify reporting requirements
- Ensure subordinate acceptance of responsibility
- Inform others who need to know
- Monitor progress in appropriate ways
- Arrange for the subordinate to receive necessary information
- Provide support and assistance, but avoid reverse delegation
- Make mistakes a learning experience

General Guidelines for Empowering Managers

- Involve people in decisions that affect them
- Clarify goals and objectives and explain how the work is related
- Delegate responsibility and authority for important work activities
- Take into consideration individual differences in ability and motivation
- Provide access to relevant information
- Provide the resources needed for new work responsibilities
- Realign management systems consistent with empowerment principles
- Remove bureaucratic constraints and unnecessary controls
- Express confidence and trust in people
- Provide coaching and advice on a timely basis
- Encourage and support initiative and problem solving
- Recognize important contributions and achievements
- Ensure that rewards are commensurate with new responsibilities
- Ensure accountability for the ethical use of power

How Leaders Enhance Organizational Learning

- Question traditional methods; look for more effective methods
- Articulate inspiring vision to gain support for innovations
- Encourage & facilitate acquisition new skills
- Help develop shared mental models about cause-effect relationships
- Encourage social networks facilitate shared knowledge
- Help people recognize when important learning has occurred
- Gain external support funding major initiatives
- Encourage experiments
- Encourage teams' after activity reviews
- Encourage recognition when initiative is failing & should be aborted
- Create decentralized subunits with authority

Source: Yukl, G. (2009) Leading organizational learning. *Leadership Quarterly*.

Influence Tactics

- Rational persuasion
- Exchange
- Inspirational appeal
- Legitimating
- Apprising
- Pressure
- Collaboration
- Ingratiation
- Consultation
- Personal appeals
- Coalition

(Source: Yukl, 2009)

References

Bass, B. M. (1990). *Leadership, psychology, and organizational behavior*. New York: Harper & Row.

Becker, W. S. (2006). In the crime lab. *The Industrial-Organizational Psychologist, 43, 4*, 21-27.

Becker, W. S., & Dale, W. M. (2007). Critical human resource issues: Scientists under pressure. *Forensic Science Communications, 9*, 2.

Becker, W. S., Dale, W. M., Lambert, A., & Magnus, D. (2005). Forensic lab directors' perceptions of staffing issues. *Journal of Forensic Sciences, 50, 5*, 1255-1257.

Becker, W. S., & Dale, W. M. (2003). Strategic human resource management in the forensic science laboratory. *Forensic Science Communications, 5*, 4.

Bernardin, H. J. (2010). *Human Resource Management: An Experiential Approach* (5th edition). Boston: McGraw-Hill Irwin.

Borman , W.C. & Brush, D.H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance, 6*, 1-21.

Dale, W. M. & Becker, W. S. (2007). *The Crime Scene: How Forensic Science Works*. NYC: Kaplan.

Dale, W.M.& Becker, W.S.(2005). Managing intellectual capital. Forensic Science Communications, *7*, 4.

References (cont'd)

Dale, W. M., & Becker, W. S. (2004). A case study of forensic scientist turnover. *Forensic Science Communications, 6*, 4.

Dale, W. M., & Becker, W. S. (2003). Strategy for staffing forensic scientists. *Journal of Forensic Sciences, 48*, 2, 465-466.

fbi.gov/hq/lab/codis/map

Mumford T.V., Campion M.A., & Morgeson F.P. (2007). The leadership skills strataplex: Leadership skill requirements across organizational levels *Leadership Quarterly*, 18 (2), 154-166.

National Academy of Sciences (2009). *Strengthening Forensic Science in the United States: A Path Forward*. Committee on Identifying the Needs of the Forensic Sciences Community; Committee on Applied and Theoretical Statistics, National Research Council.

Yukl, G. (2009). *Leadership in organizations*. Upper Saddle River, N.J.: Pearson Prentice Hall.

Yukl, G. (2009) Leading organizational learning. *Leadership Quarterly*.

Yukl, G., & Becker, W. S. (2006). Effective empowerment in organizations. *Organization Management Journal, 3*, 3, 210-231

Appendix 16: Leadership Assessment Course Evaluations with Feedback

Northeast Regional Forensic Institute (NERFI)
Funded by NIJ

Leadership Assessment: Developing the Next Generation of Leaders
A Two-Day Workshop for Forensic Managers
January 14-15, 2010

Participant Event Evaluation

Please fill out this program evaluation. Rate each of the aspects of this workshop by circling the appropriate number. Comment as appropriate. **9 Responses**

Program Overall

	Poor				Excellent	Comments/Suggestions
A	1	2	3	4	5	We had a great group – the chemistry was great Excellent program Great ideas Liked ….format in group setting Well-thought out. Wished we had more time, maybe an additional day for touring NY Crime Lab More time should have been available for theory of leadership. Maybe theory covered first
	# Responses			2	7	
	%			22%	78%	

Speakers

	Poor				Excellent	Comments/Suggestions
B	1	2	3	4	5	Both professional & knowledgeable. Very good use of time. Superb speakers. Each of the speakers were very knowledgeable and effective presenters. Highly specialized in leadership development Very professional and great knowledge in management. Will go back to my office to try the new ideas gleaned from this workshop. Will keep in touch.
	# Responses			1	8	
	%			11%	89%	

Quality of Audio-Visual

	Poor				Excellent	Comments/Suggestions
C	1	2	3	4	5	Appropriate More videos to show good/bad leadership followed by discussion
	# Responses		1	1	7	
	%		11%	11%	78%	

Quality of handouts

	Poor				Excellent	Comments/Suggestions
D	1	2	3	4	5	Papers were very helpful. I had a bit of trouble figuring out what slide we were on in the handout. So many slides missing from handouts. The presentations jumped from slide to slide in totally different handouts. Great resources of published articles I can use. Very informative and helpful.
	# Responses		1	2	6	
	%		11%	22%	67%	

Facilities

	Poor				Excellent	Comments/Suggestions
A	1	2	3	4	5	COLD! Very good Except for cold temps! Comfortable room Proximity to the airport for the hotel and university is superb!
	# Responses		1	1	7	
	%		11%	11%	78%	

Please comment on:

Did the material presented meet your professional expectations?

Yes – Very helpful
Yes – great. I would highly recommend it to the other labs.
The material did, for the most part, meet my expectations.
Yes
Yes. Will give me more options, to implement some of the ideas in my work place.
Yes. But I am sure the material will improve with further courses.
Yes. I'm walking away armed with some valuable management tools to use. I have optimism they will work.
Yes. The speakers, Dr. Wendy Becker and Dr. Ed Pavur were most effective and professional speakers. Enjoyed the interaction with the participants. Great experience.
Yes. Wish there was a little more time on 2nd day.

What topic or topics did you find most useful/interesting/novel?

All of it.
Leadership topics and solutions to the management issues/challenges brought forth by the participants.
Gary Yukl's theory
I can use these in…my work.
The last role plays.
Role play
Visit to Albany lab was excellent! Thanks for re-arranging the workshop schedule to accommodate the tour.
Material on influencing others

General comments:

I truly enjoyed the training.
Friendly staff, great workshop
This course could be very helpful to people aspiring to be a leader/manager instead of more experienced managers
The tactics list should be provided & discussed after the role plays. Such discussion would be more useful.
Group discussion was very helpful.
I'm going to recommend this to my colleagues.
Thank you to Dr. Becker and Dr. Pavur. Thank you, John & Katherine. Thank you, NIJ for the opportunity given us. Is there a Part II. We'll come back.
Keep tour of crime lab included! Like group projects/interaction & to hear from each other, like smaller group size. Could use a management class each year as a refresher & for general discussion.

Suggestions for improvement:

Thanks
Turn up the heat!
Another day/24 hours addition to this workshop will work better.
Add 1 day. More on Evaluating Employee Performance
Having more practice & leadership examples/case scenarios with suggestions on the best way to handle certain situations.
Temperature of room! Print agenda. X-tra paper for notes or tell people to bring paper/pens

Thanks for your input!

UNIVERSITY AT ALBANY
State University of New York

Northeast Regional Forensic Institute (NERFI)
Funded by NIJ

Leadership Assessment: Developing the Next Generation of Leaders
A Two-Day Workshop for Forensic Managers
February 18-19, 2010

Participant Event Evaluation

Please fill out this program evaluation. Rate each of the aspects of this workshop by circling the appropriate number. Comment as appropriate. **10 Responses**

Program Overall

	Poor			Excellent		Comments/Suggestions
A	1	2	3	4	5	Very useful topics covered
	# Responses			1	9	
	%			10%	90%	

Speakers

	Poor			Excellent		Comments/Suggestions
B	1	2	3	4	5	Very knowledgeable! Very well-presented. They know their stuff The instructors prompted very thought-provoking discussions in a very safe environment.
	# Responses			2	8	
	%			20%	80%	

Quality of Audio-Visual

	Poor			Excellent		Comments/Suggestions
C	1	2	3	4	5	Short, often didn't match hand-outs Some PP do not match handouts
	# Responses		1	2	7	
	%		10%	20%	70%	

Quality of handouts

	Poor			Excellent		Comments/Suggestions
D	1	2	3	4	5	Some PP do not match handouts
	# Responses			2	8	
	%			20%	80%	

Facilities

	Poor			Excellent		Comments/Suggestions
E	1	2	3	4	5	Room temperature fluctuates from cold to hot – not consistent Would have liked to have food at conference site
	# Responses		1	2	7	
	%		11%	20%	70%	

Please comment on:

Did the material presented meet your professional expectations?
Yes x 7
Absolutely. The workshop was way above my expectations

What topic or topics did you find most useful/interesting/novel?
Being able to network and discuss issues that other state/county/private labs face
The tools for discussing situations on how to get what you need (rational persuasion, etc.) Influence tactics
Having the chance to hear firsthand, the issues of other managers and how they lead in their department was very educational/informative
Types of ways to influence people
Psychology of being a leader, role playing
Interactions with other participants was very useful when discussing work issues
Information on influence tactics
Retention of employees/morale
Influence tactics & role playing with the group

General comments:
It was very helpful that so many in the class are in a similar situation to me. Thank you!
Excellent course, would like to see similar offerings
Great two day workshop
Wendy [Becker] is an excellent active listener. The focus on the issues the attendees are personally facing makes the class more pragmatic but less instructional/theory/content based.
Loved the interaction with other people & learn how they deal w/ situations, as well as techniques used un their labs
Breakout sessions – great time to have one on one's & share experiences

Suggestions for improvement:
More time. Maybe add a day – many issues will be generated by the participants & more time given to problem solve these issues
Less break-out sessions – more presentations
Longer

Thanks for your input!

UNIVERSITY AT ALBANY
State University of New York

Appendix 17: First Responder Training for Schenectady Police Department

Northeast Regional Forensic Institute
1400 Washington Avenue, Biology 225
Albany, New York 12222
http://www.albany.edu/nerfi

Schenectady Police Department
Crime Scene Analysis Evidence Collection & Packaging

Instructors: John Hicks, Lucy A. Davis & Jamie L. Belrose

Course Description
This course is designed to give police officers and other law enforcement professionals the basic knowledge of crime scene analysis and review of forensic evidence including biological fluids they need to keep up with today's technological advances. Today's advanced technology has led to an exponential number of cases being submitted to the crime lab. For this technique to be successful, it is imperative that the evidence be collected efficiently. Missteps at this first stage can, and most likely will, have deleterious effects on the laboratory analyses. By properly collecting and packaging evidence you can increase the probabilities of obtaining a successful forensic result.

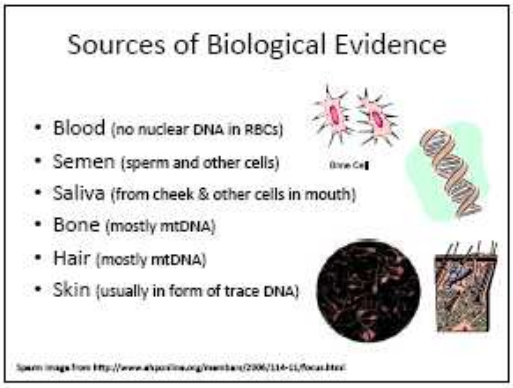
Course Outline

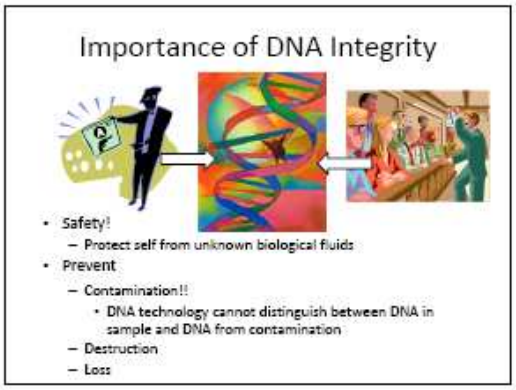
8:00 AM	Welcoming and Introductions
8:15 AM	**Crime Scenes & Crime Scene Evidence**: Things to look for and what is most helpful to the crime lab.
9:15 AM	*Break*
9:30 AM	**Crime Scenes & Crime Scene Evidence - continued**
11:15 AM	**Crime Scenes & Crime Scene Evidence - continued**
12:00 AM	*Lunch*

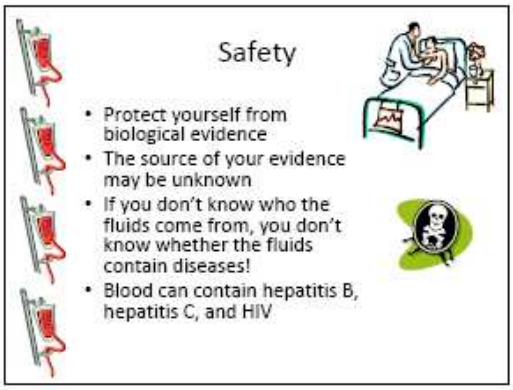
References:
Crime Scene Investigation: A Guide for Law Enforcement, January 2000, Technical Working Group on Crime Scene Investigation: National Institute of Justice; NCJ 178280

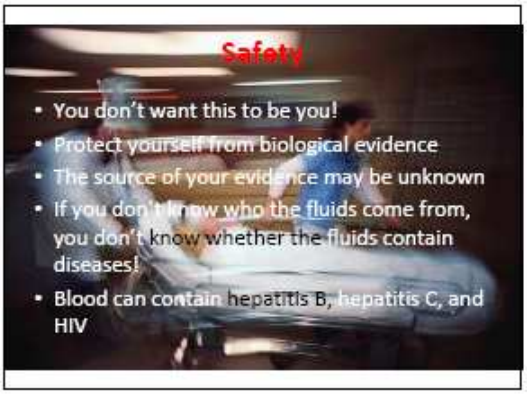
Crime Scene Investigation: A Reference for Law Enforcement, June 2004, Technical Working Group on Crime Scene Investigation: National Institute of Justice; NCJ 200160

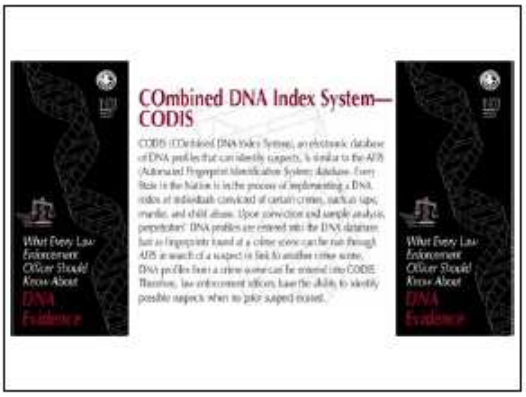
Appendix 18: First Responder Presentations











What we're talking about today

- Evidence
  - Collection
  - Crime scene
- Types of evidence
  - Biological
    - DNA
  - Trace
    - Hairs and Fibers

Forensic Science

- Public interest
  - Edgar Allen Poe, Sherlock Holmes
  - Quincy
  - And now CSI, etc.
- Intrigued by darker side of humanity
- Puzzles, clues, solutions

Role of the Forensic Laboratory

- To use science to support or deny the victim or suspect's story
  - The laboratory needs to know the total history of what happened:

    Number of suspects, previous sexual intercourse, exact orifices involved, time between assault and examination

We're all 'Pig Pen'

- Like 'Pig Pen', we carry our micro-environments with us wherever we go. When our environment comes into contact with another, transfers occur.
- These transfers are indications of association.

Example of transfer

You own a dog

You pet the dog

You go to work

You sit in your chair

Transfer Example

Hairs from the dog transfer to you

They transfer from you to your work area...

Like your chair...

Direct Source
(A to B)

Indirect Source
(A to B to C)

Crime Scenes

- Confirm what is involved
  - Always try to take a minimum of 2 people
- Inventory consumables in crime scene kit
  - Evidence bags, batteries, solutions, film
- Try to write the final report within 24 hours of completion of the scene

Planning

- Stop
- Look
- Listen

Do not go in without a plan.
Take a methodical approach

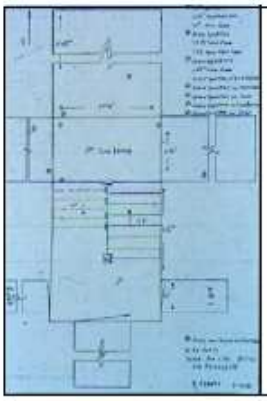
Collection of evidence

- From body outward
  - Good for outdoor scenes
  - Take the search in grids or circle patterns
  - Let the evidence found led you further

- Toward the body
  - Better for enclosed areas
  - Preferred if trace evidence is critical
  - Better if scene has been secured prior to you arrival

Photographs

- General to specific
- 90 degree angle
- Without scale
- Then with scale
- Remove lens cap and check roll advance
- Document what is on each roll



Drawings

- Best way to document scene
- Be specific with measurements
- Note directions
- Can be used with legends of collected exhibits
- Does not have to be high tech

TWGSCI Guidelines

- Report can be downloaded from the web: www.ncjrs.org
- Call 800-851-3420
- Request – Research report "Crime Scene Investigation – A Guide to Law Enforcement

Footwear

- Everyone wears them
- If they don't, they're leaving *foot*prints
- ***HUGE*** variation in styles, designs, and types
- Quality often dependent on substrate
  - Material in which impression is made
- We all walk differently—creates wear
- Outsole materials wear differently—creates patterning
- What we walk on varies and damages sole uniquely

Soles vary radically

Finding footwear impressions at the crime scene

- Be aggressive in searching for them
  - Often overlooked or damaged
- Failure relates to
  - Thinking none will be found
  - Incomplete searches of scene
  - Searches after trampling
  - Shoe and surface combination not conducive to leaving impression
  - Intentionally destroyed by culprits
  - Environmental damage (rain, snow, etc)

Factors of wear on the sole

- Person's foot type
- Occupation
- Habits
- Weight and body type
- Shoe construction
- Style
- Materials
- Surfaces walked upon

Left and right shoes don't wear the same

Types of impressions

- Dry
  - When shoes track across a dirty surface, the bottom of the soles will accumulate a coating of residue. If they then track on to a relatively clean surface, that residue will be deposited in the form of footwear impressions
- Wet
  - made when a shoe sole is wet or damp constitute another category of impression.
- Depression
  - The term 'depressed mark' is used to describe the evidence an item of footwear leaves when making contact with a deformable surface, such as soil, sand and snow. It is synonymous with the term 'three-dimensional impression'.

Collection—taking casts

Oblique lighting

Electrostatic lifting

- Applies electric charge to plate
- Lifts dust/dirt/debris
- Retained through surface charge until treated
- Can be powered through stun gun

Electrostatic lifting

Print lifted

Document before

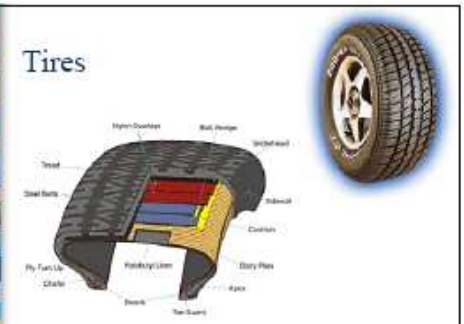
Results of examination

- Individualization
- Elimination
- Linking to crime, other scenes
- Location of impression
- Corroboration or rebuttal of alibis
- Classification, shoe size
- Number of perpetrators

Class characteristics

- Nike Air Dynamic Flight shoes
  - USA total all sizes 227,544
  - Size 13, Black 8,136
  - White 7,032
  - Total 15,168
  - Distribution to Georgia
    - All sizes 6,570
    - Size 13 433
- Can't use as statistics but good example

Tires



Tire print examinations

- Tread design
  - Thousands of designs (~11,000)
    - Passenger vehicles, trucks, off-road, motorcycle, etc.
- Tire size
- Track width
- Wheelbase
- Turning diameter
- Re-treaded/replacements

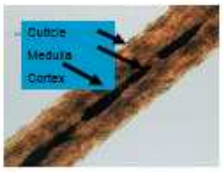
Not all tire tracks are left on pavement

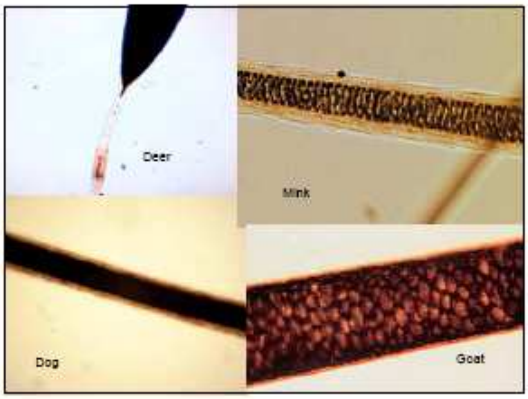
Trace evidence - Hair and Fibers

- Can be easily lost or missed
- Get on your hands and knees and look
  - Document location each are collected
- After search is completed you may go back and do vacuum sweeping or tape lifts

What can be determined from a hair?

- Human vs Animal
- Race
- Body Area
- Damage/Disease
- Comparison to a known sample





What area of the body is it from?

- **Head**
- **Pubic**
- Facial
- Limb
- Chest
- Axial (armpit)

These carry the most information for microscopic comparisons

What is the person's ancestry?

- African ancestry
- Asian ancestry
- European ancestry

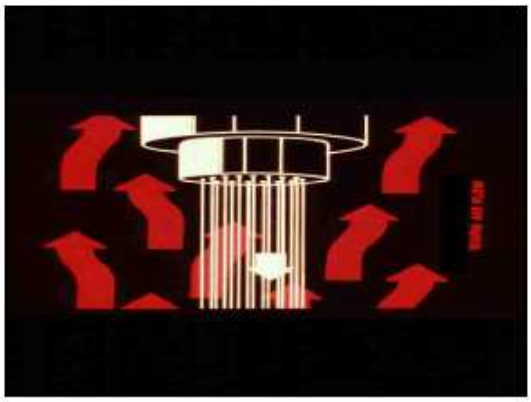
- Estimation of ancestry based on hair characteristics may not correlate with genetic or self-identity

Why Fibers Make Good Evidence

- everywhere (readily available)
- we're in constant contact with them
- designed for a purpose (variation)
- fungible (life-span)
- fashionable (variation)
- easily tested (limited characteristics)
- nearly infinite combinations of traits
  - Especially color

Examples of fibers

- Natural
  - Wool, cotton, silk, linen
- Manufactured
  - Rayon, acetate, lyocell
- Synthetic
  - Nylon, polyester, kevlar



Fiber characteristics

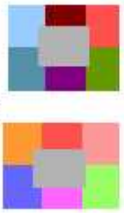
- Cross-section
- Color
- Polymer type
- Optical properties
- Voids/inclusions
- Diameter
- Other properties (metamerism)

Perception/detection of color requires:

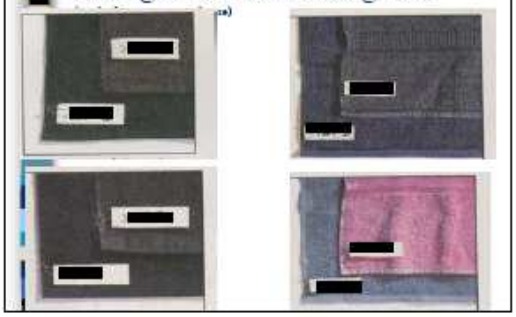
- light source
- object
- observer (human or instrumental)

Simultaneous Contrast

- The eye generates contrast to heighten perceptual borders
  - Important evolutionarily
- Simultaneous contrast changes apparent lightness and chromatic color in a direction opposite to those of the inducing background
  - **Complementary colors**



Changes due to washing 50x



Significance

- Examples to think of...
  - Department store
  - Public transit
- What are the chances of two non-associated textiles matching?

Textile Labeling

Manufacturer Identification

Fiber name and % composition

Care instructions

Country of origin

RN 9265

50% Merino wool
30% viscose rayon
16% microdenier polyester
4% Spandex

Dryclean (A)

Made in USA of imported fabric from Australia

Type of wool fiber
Type of rayon fiber
Type of polyester

Care symbol

Name of the country in which the fabric was produced

What Is Soil?

- Mixture of organic and inorganic material
- May range from 100% inorganic (sand) to nearly 100% organic (peat)
- Inorganic part is minerals
- Organic part is decayed plant and animal material and is sometimes called ***humus***
- Many labs put soil in a big generic class and say, "It's all the same."

Glass

- Tempered
  - Auto windows
  - Shower doors
- Non-tempered
  - Laminated glass
    - windshield
  - Architectural
    - windows
- Headlights
- Containers

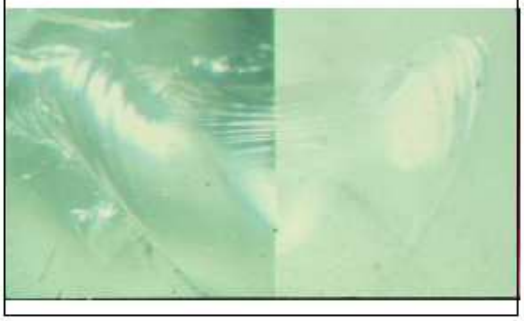
Forensic Significance

- 700 types of glass in common use today
- Can be individualized if fracture match can be made
- Tiny particles are class evidence
  - Analytically classified
    - Optical properties
    - Elemental content

Types of Analysis of Glass

- Physical Match
  - non-tempered
- Comparison
  - Physical properties
  - Chemical properties
  - Optical properties
- Impacts
  - Direction of blow
  - What caused the damage

Broken Windshield Fracture Match



Glass breakage

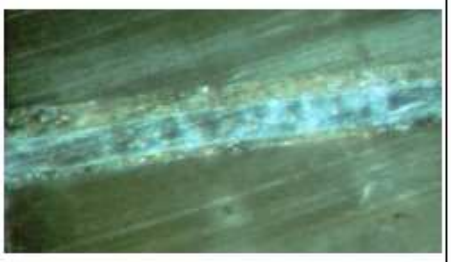
Radial fractures form initially. These run out from the point of impact on the opposite surface to the applied force. Then concentric fractures form. These run between the radial fractures but on the same side of the glass as the impact.

Most of the glass is projected forward, in the direction of the blow. But some of the fragments will be projected backwards, towards the person breaking the window, up to a distance of 3 meters – known as 'blowback'. Anyone standing close enough when the window breaks will be covered in tiny fragments of glass which will stick to their clothing and their hair.

Introduction - Paint

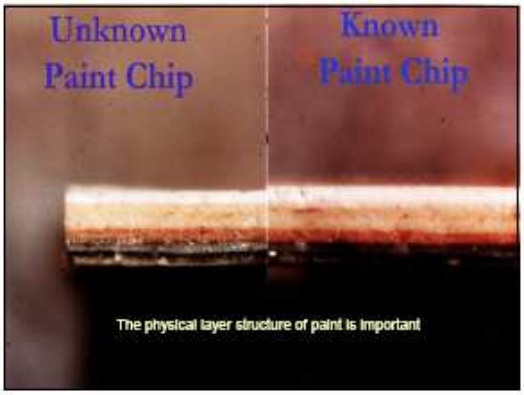
- Chemistry of paint very complex
- Forensic role is to compare paint evidence from scene to known source
- Individual evidence only in cases where large enough paint samples exist to physically match

Paint Cross Section



Unknown Paint Chip

Known Paint Chip



Tape

- May be source of fingerprints
- Hairs and fibers may adhere to the adhesive portion
- Chemical composition of tape adhesive and backing

What happens at a crime scene?

- Identification of blood
  - Documentation of the scene and examination performed
    - Photograph everything
    - Use drawings, charts, everything
  - Evaluation of stains location
    - Use of alternate light sources or chemical enhancers
    - Very important in blood spatter interpretation
  - Field identification of blood
    - Presumptive tests

Collection of blood samples

- Collect it all
  - Call the laboratory if there is a question
  - Evaluate the location of the stain and probative value
  - If in doubt of method of collection, cut it up and bring in the entire object
  - Always use clean utensils when collecting anything

Collection of blood samples

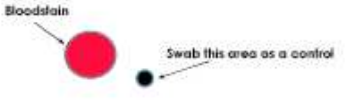
- Scrapings
  - Use a piece of paper, then fold the paper
  - Do not scrape directly into an envelope
  - Be cautious of this type of collect due to aerosols, and bionazardous materials
- Swabbing
  - Keep the stain concentrated
  - Cotton applicators (Q-tips)
  - Better to collect it and not use then lose the sample

Collection of blood samples - Controls

- Reagent controls
  - Use distilled or sterile water
  - Bottle water will do - *only if there is nothing else available*
  - Do collect the sample of whatever you use

Collection of blood samples - Controls

- Substrate controls
  - Collected from area adjacent to stain



Packaging, Transportation, Shipping

- Air dry all samples
- Liquid blood standards
  - EDTA – purple top tube
- Store in paper bags
- Keep it cool
  - Frozen, refrigerator, air condition
- Transport to laboratory as soon as possible
- Documentation

Chemical Tests

- Testing of visible stains
- Evaluation of results
  - These tests say that the stain is blood only - no species identification
  - A non-reaction does not mean that blood was not there

Chemical Tests

- Testing of Invisible stains
  - Luminol
    - Reacts better the older the stains
    - Be cautious of false positives
  - Amino black
    - Permanently stains the object
  - Flora-scene
- Evaluation of results
  - Presumptive tests - not definitive identification
  - Sensitive for small amounts of blood
  - Very good to lead you to larger stains

Luminol

Pre-luminol Post-luminol

Can you test blood after it has been fingerprinted?

- Chemical/substrate combinations
  - Super glue
    - used on glass, latex, metal and plastic
  - Ninhydrin
    - used only with paper
  - Super glue + Rhodamine
    - used on glass, plastic and latex
  - Super glue + Rhodamine + black powder
    - used on glass, plastic and latex

What is Blood

- Red Blood Cells - RBC
  - Approximately 4 million in a single drop of blood
  - Basic cell of analysis in conventional serological testing
  - Does not contain DNA

What is Blood

- White Blood Cells - WBC
  - Approximately 7000 in a single drop of blood
  - Does contain DNA - is the cell of that analysis

Species Origin

- Is it animal or human?
  - Interesting point in special cases
- How much do I need
  - Conventional serology
  - DNA analysis

Semen Identification

- Seminal Fluid
  - The liquid portion of semen
    - Chemical/enzymatic test can be performed to identify this
    - This can be used to screen for the presence of semen
    - Can be visualized by and alternate light source
    - UV light source, Woods light

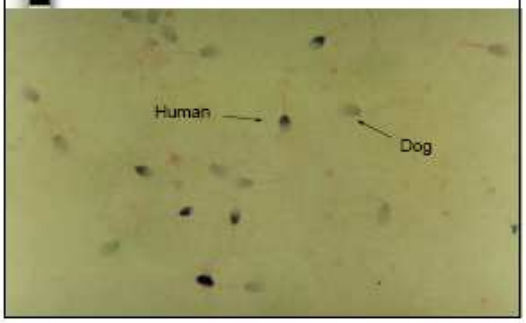
Semen Identification

- Spermatozoa
  - The solid part of semen
    - Can be used to differentiate between human and different animals
    - The portion of semen that contains DNA

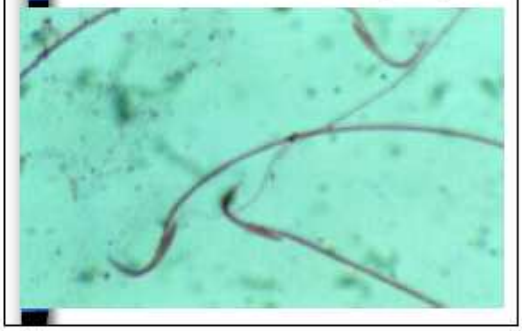
Identification of Semen

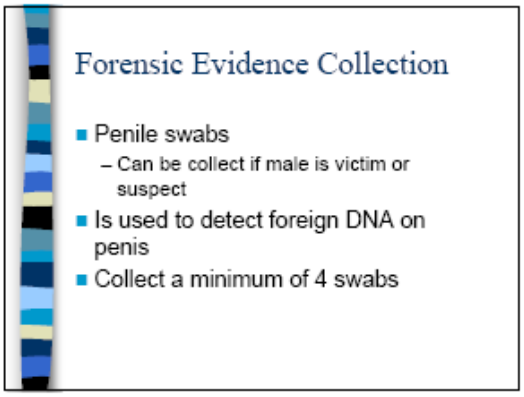
- Primary indicator of sexual activity
- Begin with the vaginal swabs and work our way out
  - Vaginal swabs, external swabs, underpants, sheets, bedding, other
- With successful analysis of semen, limited further analysis will be performed

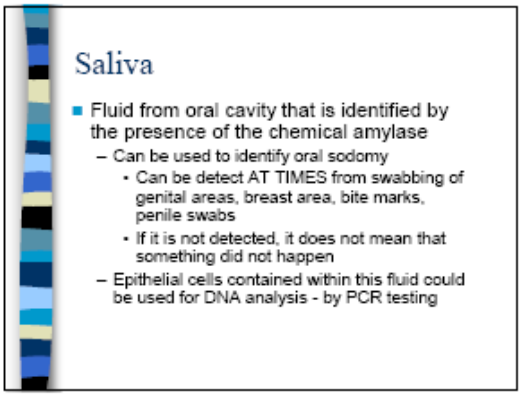
Different issues in cases

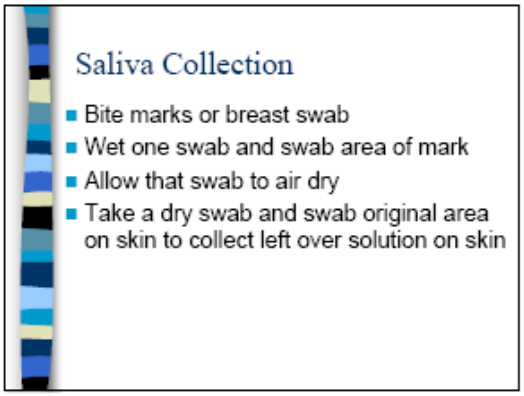


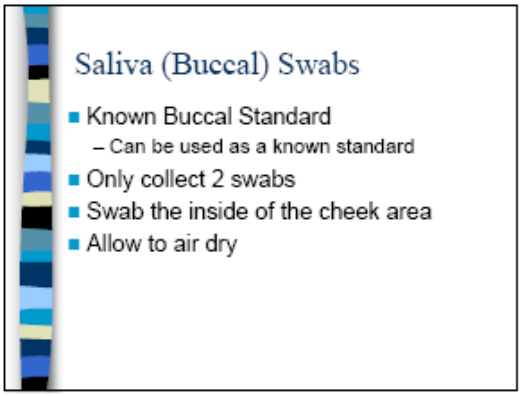
You never know what the story may be!

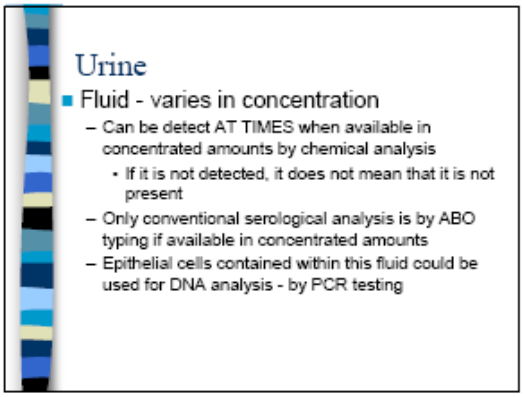


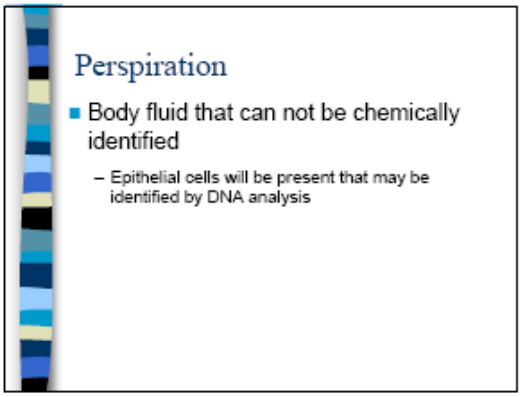










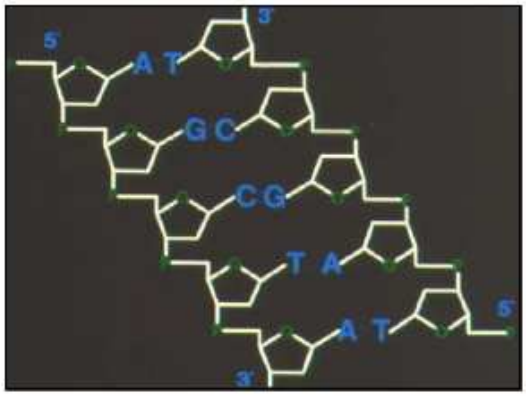


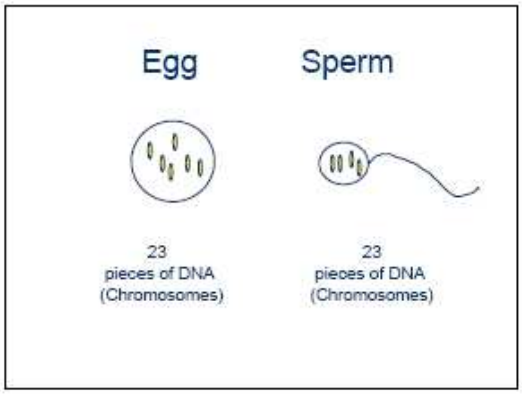
Other Physiological Fluids

- Feces
  - If it looks like it, it smells like it, it possibly is
  - Very little you can do to identify or type
- Tissue
  - Can be identified microscopically
  - Can be typed by conventional enzymatic testing
  - Can be typed by DNA analysis

DNA
Deoxyribonucleic Acid

- Fundamental building block of all living organisms
- Very stable chemical molecule
- No two people except for identical twins have the same type

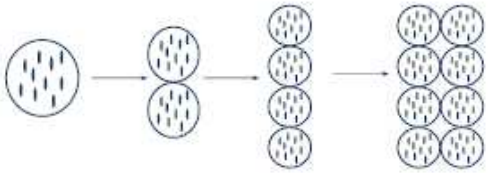




Single Cell

Contains 46 Chromosomes

The cell copies its DNA and divides into 2 identical cells



Then repeats the process over and over

Baby Rapist is born

The DNA in every cell of his body is identical to the DNA that started in that first sperm and egg combined.

Definitions

- DNA – Deoxyribonucleic Acid
- Locus – A single location (site) on the DNA chain
- Loci – Plural of locus
- D7S820 – A location on the DNA chain of Chromosome 7. It was the 820 site identified sequentially
- vWA – Another location on the DNA chain that is referred to by it's common name, and not designation

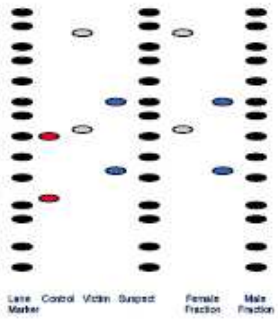
Differential Extraction

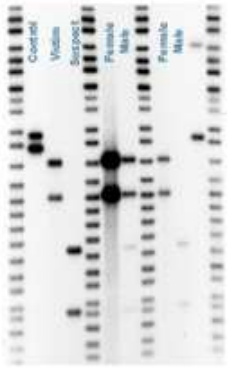
- Vaginal swabs will contain both the female's epithelial cells and male's sperm cells
- Separates sperm DNA from the female epithelial DNA
- From one piece of evidence we now have two DNA samples

Different types of Evidence

- Condoms thrown in trash, on floor, in toilets, in pockets!

Match at Locus D5S818





Autorad of the of DNA profile at D5S818.

The male fraction DNA profile matches the suspects DNA profile at locus D5S818

The frequency of occurrence of that match is

1 person out of 25

What do those numbers mean?

10 x 30 x 19 x 15 x 18 x 25 x 12 x 8 x 10 x 15 x 20 x 13 x 27

equals

3,889,360,800,000,000

1 person in 3.8 Quadrillion

What do those numbers mean?

10 x 30 x 19 x 15 x 18 x 25 x 12 x 8 x 10 x 15 x 20 x 13 x 27

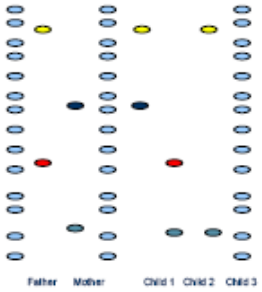
equals

3,889,360,800,000,000

6,000,000,000

(world population)

Paternity typing



Best Samples to use for identifying unknown bodies or bloodstains

- Father of the missing person
- Mother of the missing person
- Brothers and sisters of the missing person
- Children of the person accompanied by a sample from the other parent of the children

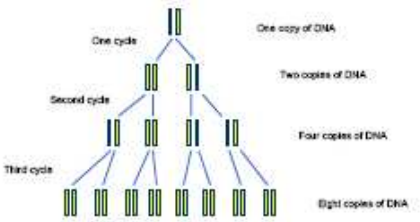
RFLP - Restriction Fragment Length Polymorphism

- Used in the United States since 1987
- Accepted in every states legal system
- The technique has been found to be very reliable and stable
- Only a couple of private labs can do

PCR
Polymerase Chain Reaction

- Make copies of the DNA until you have enough to analyze
- Based on the same procedures that your body uses when it copies the DNA so the cell can divide
- Applied to forensic cases in 1991

Polymerase Chain Reaction



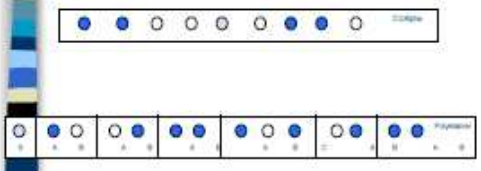
Biological specimens you can analyze with PCR

- Very small blood stains
- Semen from vasectomized males
- Stamps and envelopes
- Cigarette butts
- Hairs with a root end
- You name it, its worth a try

Types of analysis by PCR

- DQAlpha - PolyMarker DQA1-PM
- Short Tandem Repeats STR's

DQAlpha - PolyMarker



DQAlpha - PolyMarker

Frequencies of occurrences will be between

1 person out of 1000

to

1 person out of 100,000

DQAlpha - PolyMarker

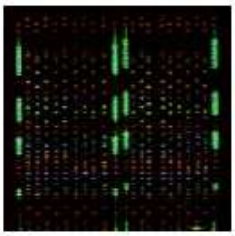
- Has been fully validated
- Used on forensic cases since 1991
- Interpretation of mixed samples
- No laboratory can to this type of analysis any longer

DQAlpha - PolyMarker

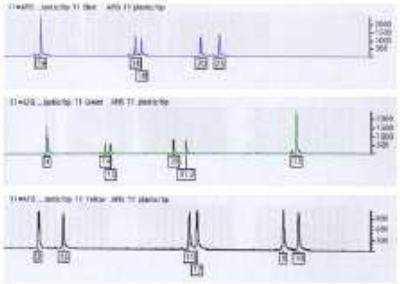
- Interpretation of mixed samples
- This is where a lot of the overturned DNA cases are coming from

STR's

- Easier to visualize mixed samples
- The standard DNA analysis used by all forensic laboratories
- Fully accepted in the courts
- Allows you to analyze up to 16 different genetic loci simultaneously



Techniques are rapidly advancing.



Report Formats - Match

Sample	Victim	Suspect	Vaginal swab - non-sperm	Vaginal swab - sperm
D3S1358	14, 15	13, 18	14, 15	13, 18
vWA	17, 18	11, 14	17, 18	11, 14
FGA	20, 23	24, 26.2	20, 23	24, 26.2
D8S1179	9, 12	13, 13	9, 12	13, 13
D21S11	26, 30	31.2, 33	26, 30	31.2, 33
D16S51	11, 18	15, 24	11, 18	15, 24
D5S818	8, 12	10, 15	8, 12	10, 15
D13S317	10, 10	12, 12	10, 10	12, 12
D7S820	8, 11	13, 14	8, 11	13, 14
Amelogenin	X, X	X,Y	X, X	X,Y

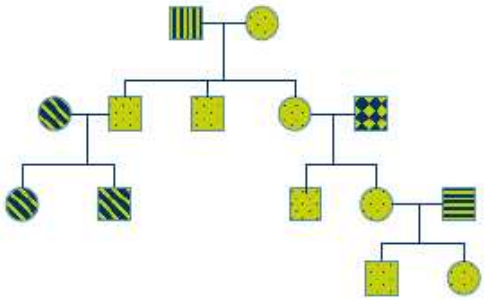
Y – STRs

- DNA testing for samples that contain a mixture of male and female DNA
- Only the male DNA type is identified, therefore mixtures can be resolved
- Statistics are lower than standard STRs, like mitochondrial statistics

Mitochondrial DNA Analysis

- Highly specialized analysis
- Good for bones and hair shafts
- Only a few laboratories perform the analysis - FBI, Armed Forces laboratory, and a few private labs
- Regional Public Labs – Arizona, Minnesota, New Jersey, Connecticut
- Cost - $2000
- Maternally inherited

Maternal Inheritance



Contamination

The curse of the Forensic Laboratory DNA Section

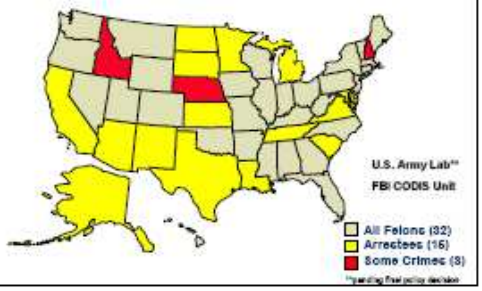
What can you do to help prevent Contamination

What can you do to help prevent contamination

- Be aware
- Gloves and protective clothing
- Clean everything
- Collect controls
- Documentation

All States Have Offender Legislation

or Federal Legislation
(as of June 2009)



CODIS Architecture

Database Statistics

As of June 2009

- Total number of profiles: > 7,409,920
  - Total Forensic profiles: > 272,452
  - Total Convicted Offender Profiles: > 7,137,468
- http://www.fbi.gov/hq/lab/html/codis1.htm

How does it work?

DNA profile is developed from the semen sample

Profile is placed in the CASEWORK database

Search is performed to see if any samples with in the database match

Case work database

Contains DNA profiles generated from case analysis. These samples are from "unknown" samples

Match between 2 casework samples

Unsolved case to unsolved case

Unsolved case to solved case

Date rape cases

How does it work?

DNA profile is developed from sample collected from a convicted offender

Profile is placed in the OFFENDER database

Search is performed to see if any unknown samples with in the database match

Offender database

Contains DNA profiles taken from persons convicted of specific criminal offenses