

Extracting evidence from forensic DNA analyses: future molecular biology directions

Bruce Budowle^{1,2} and Angela van Daal³

¹Department of Forensic and Investigative Genetics, University of North Texas Health Science Center, Ft. Worth, TX, USA, ²Institute of Investigative Genetics, University of North Texas Health Science Center, Ft. Worth, TX, USA, and ³Faculty of Health Science and Medicine, Bond University, Gold Coast, Queensland, Australia

BioTechniques 46:339-350 (April 2009 Special Issue) doi 10.2144/000113136

Keywords: forensic science; molecular biology; low copy number; SNPs; mini-STRs; automation; microfluidics; microbial forensics; databases

Molecular biology tools have enhanced the capability of the forensic scientist to characterize biological evidence to the point where it is feasible to analyze minute samples and achieve high levels of individualization. Even with the forensic DNA field's maturity, there still are a number of areas where improvements can be made. These include: enabling the typing of samples of limited quantity and quality; using genetic information and novel markers to provide investigative leads; enhancing automation with robotics, different chemistries, and better software tools; employing alternate platforms for typing DNA samples; developing integrated microfluidic/microfabrication devices to process DNA samples with higher throughput, faster turnaround times, lower risk of contamination, reduced labor, and less consumption of evidentiary samples; and exploiting high-throughput sequencing, particularly for attribution in microbial forensics cases. Knowledge gaps and new directions have been identified where molecular biology will likely guide the field of forensics. This review aims to provide a roadmap to guide those interested in contributing to the further development of forensic genetics.

Introduction

Forensic science has embraced the use of DNA molecular biology tools for diagnostic purposes more than any other scientific field. The discipline has been driven by the need for high-resolution human identity testing techniques. Over the past 20–25 years, forensic science has developed and implemented various robust and reliable DNA typing technologies (1–3). Successes have enabled the reliable typing of extremely minute quantities of DNA, with a resolving power such that, in many cases, the number of evidence-sample contributors can be reduced to a few individuals, if not just one source. In addition, forensic molecular biology tools are very reliable because of well-defined validation requirements (4,5).

Given the forensic field's maturity, it could be assumed that dramatic changes in technology will not be sought and only refinements will be embraced. There are fewer demands to meet technologically; in fact, the capability to routinely type samples such as a cigarette butt or a single strand of hair has exceeded the expectations of most scientists who first began using molecular biology tools to characterize forensic biological evidence. Rather than using restriction length polymorphism analysis by Southern blot–based hybridization methods (6–8), scientists in the field are now routinely using PCR-based methods coupled with automated fluorescent detection technologies (9–17). The use of offender and

forensic sample DNA databases contributes to reticence for change. These databases were developed to help investigate future crime and have been standardized on a core set of short tandem repeat (STR) or microsatellite loci (18,19). Because of the size of these databases [for example, there are >6 million reference profiles in the United States Combined DNA Index System (CODIS) database (20)], there is a substantial movement to maintain just the current core genetic marker repertoire. Additionally, because of the substantial resource outlay to validate molecular biology analytical systems, to equip a laboratory, and to educate and make proficient practitioners; as well as the efforts undertaken to gain admissibility in the courtroom (21), forensic scientists tend not to change sound methodologies quickly. One might predict, therefore, that there are not likely to be any dramatic changes in the molecular biology tools used in forensic science. Such a view, however, would be myopic because there are several areas where molecular biology could offer improvements to the capabilities of the forensic scientist. Indeed, it is incumbent on the forensic scientist to be vigilant and embrace new technologies that will benefit society by their ability to analyze more challenging samples in an effort to continue to exonerate the innocent, to enhance abilities to solve crime, and to identify missing persons.

With analysis success, there is motivation to attempt to analyze more difficult samples,

such as trace samples termed touch DNA or low copy number (LCN) (22,23). DNA databases may not have been exploited fully and could provide leads to new investigative questions. In addition, the recently developed field of microbial forensics will exploit high-resolution, high-throughput technologies beyond those needed for human identification. Therefore, the future of molecular biology in forensic science still promises to be dynamic.

Predicting the future is never exact, and fundamental leap technologies are not obvious. Thirty years ago, few if any would have predicted the PCR method and the impact it has had on molecular biology. In this review, we describe the primary gaps in the handling and analysis of forensic biological evidence that are being or are likely to be filled by molecular biology tools. The gaps are not unreasonable predictions; many are obvious needs that will drive development in the forensic science field for the next 5–10 years. The areas that will be addressed are: (i) improvements to the current limits of typing samples of limited quantity and quality; (ii) investigative information including phenotypic inference from a DNA sample and pharmacogenetic information for molecular autopsy, tissue type determination by expression analysis, and microbial forensics; (iii) microbial forensics; and (iv) automation with a focus on in-field testing.

Because of space limitations, the topics herein are only discussed briefly; readers

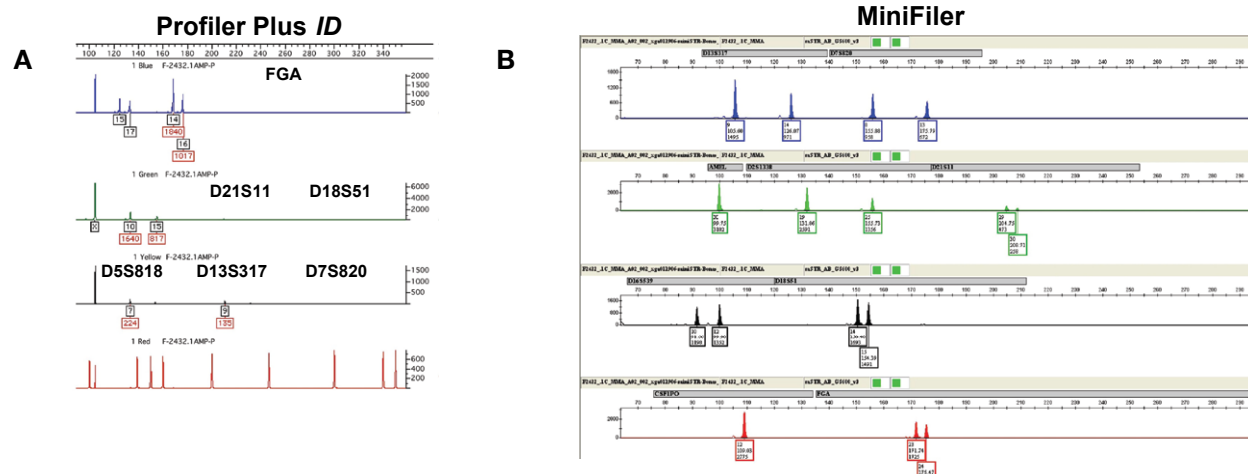


Figure 1. Comparison of STR results using standard kit format and a mini-STR format. A bone sample was analyzed with (A) the AmpFISTR Profiler Plus ID PCR Amplification Kit and (B) AmpFISTR MiniFiler PCR Amplification Kit. The loci FGA, D21S11, D18S51, D5S818, D13S317, and D7S820 failed to amplify sufficiently with the AmpFISTR Profiler Plus. All loci amplified with the AmpFISTR MiniFiler PCR Amplification Kit because of the reduced amplicon size.

should refer to the references (and their citations) for more details and other examples beyond those provided here. Hopefully, these gap assessments will help guide those who invest resources in forensic molecular biology diagnostics.

Improvements to typing-challenged samples

The success and widespread acceptance of DNA typing in forensic science is due partly to its sensitivity of detection (by the use of PCR) and an ability to analyze minute samples. However, when only a few template molecules are available for PCR, stochastic amplification will occur, and the degree of the effect is indirectly related to the number of template molecules (24,25). For short tandem repeat (STR) loci, whose alleles are based on a varied number of tandem repeats, the stochastic effects manifest as a substantial imbalance of two alleles at a given heterozygous locus, allelic dropout, and increased stutter.

Even with these stochastic vagaries making quality control more problematic, scientists are working towards enhancing the sensitivity of detection for LCN DNA typing (25–28), and it is predicted that this endeavor will continue. Low-quantity samples for human-remains identification and high-volume crimes (such as burglaries) abound and could provide leads in solving these crimes. Additionally, typing trace levels of DNA for tracking individuals or determining security breaches would further enhance counter-terrorism efforts.

Many strategies to enable the generation of LCN DNA profiles exist. These include simply increasing the number of PCR cycles (from 28 to 34); reducing the PCR volume; implementing a post-PCR cleanup step to concentrate the sample for analysis and to remove competitive ions; increasing the electrophoretic injection time; using nested

PCR; or using better signal-to-noise ratio fluorescent tags (24). With these methods for enhanced sensitivity, however, the same stochastic effects and concerns of contamination (termed allele drop-in) persist. Scientists have used multiple aliquots from one sample to introduce redundancy and assist in interpretation of any stochastic effects and potential contamination (25). The issue of contamination has been further addressed by building dedicated specialized laboratories and implementing protocols to reduce the risk of intralaboratory contamination (29).

Most of the effort directed to the application of LCN typing has involved these two approaches. There will be a need for contamination-free reagents and efficient decontamination procedures to support LCN analyses (30). Moreover, the limitations of LCN would be better addressed technically. Common sense dictates that splitting a sample into multiple aliquots exacerbates the stochastic problems of LCN typing (24): there are even fewer template molecules subjected to PCR amplification. Methodologies that reduce stochastic effects during PCR are, in the long run, more logical than sample dilution. Indeed, improvements in PCR should be pursued so that LCN samples will amplify and approximate the performance of samples containing 200 pg–1 ng of DNA (the robust template range for conventional STR typing). A potential approach to improve PCR robustness might be to test the use of PCRboost (Biomatrix, San Diego, CA, USA) to enhance PCR or to use additives in PCR that effectively concentrate the target and enzyme, such as volume excluders like polyethylene glycol and dextran sulfate.

Increasing the available template molecules obtained from LCN samples is another approach that should be considered. In fact, it is our own experience that DNA (e.g., from whole blood or buccal cells) collected on swabs or cotton swatches is not efficiently

removed from the collection device during extraction. Often there is more DNA still entrapped in the sample collection medium than was extracted (data not shown). Efforts should focus on more efficiently extracting DNA from current sample collection devices. More efficient sample recovery and extraction strategies (such as voltage-induced release and novel ion-exchange columns) may yield more template molecules. Alternatively, better collection devices should be developed that are more efficient at recovering samples at crime scenes, or that better aid in extraction by being inert to DNA or dissolving during extraction to completely free the DNA housed within.

Forensic samples can contain contaminants from the environment that inhibit PCR amplification. Even when there are sufficient template molecules for a conventional analysis, stochastic affects may occur because of the presence of PCR inhibitors. In essence, the effective number of template molecules for the PCR is diminished. DNA extracts that are purified of PCR inhibitors, or additives that can be used to counteract the effects of inhibitors, are approaches that will enhance typing reliability for compromised samples. For example, bovine serum albumin (BSA) is routinely added to the PCR with forensic samples. BSA has been shown to overcome the effects of some PCR inhibitors and has no deleterious affects during PCR with pristine samples (31). BSA's mechanism works by either binding the inhibitors or stabilizing the polymerase. However, BSA does not overcome all inhibitors; an additive that would overcome a wider range of PCR inhibitors would have obvious benefits for analyzing diverse unknown forensic samples. Alternatively, removal of the inhibitor could be considered.

Another approach that may show promise for limited-quantity DNA samples is the use of whole-genome amplification (WGA) (32,33). Ideally, the WGA method amplifies all the

Table 1. United States National DNA Index System Statistics (December 2008) (20)

Category	Number of Profiles
Convicted offender	6,398,874
Forensic	248,943
Arrestees	140,719
Missing person	519
Relatives of missing person	5032
Unidentified human remains	2283

DNA in a sample in an unbiased fashion, yielding substantially greater quantities of DNA that can be subsequently analyzed using standard forensic assays. WGA methods, however, are subject to some of the same stochastic effects that LCN typing encounters. One WGA technique known as rolling circle amplification (RCA) (34), which uses a circular DNA template, could possibly obviate some of the stochastic limitations. With a highly processive polymerase, RCA can yield microgram quantities of DNA from circular templates and—because of the phenomenon of strand displacement—produce many copies of the same target molecule. However, with the exception of intact mitochondrial DNA (mtDNA), circular DNA fragments do not occur naturally in humans, and DNA is often highly degraded in human remains. Therefore, to better exploit RCA, the fragmented DNA in an evidence sample could be circularized (35). One example of a potentially useful ligase is CircLigase (Epicentre Biotechnologies, Madison, WI, USA), which possesses a catalytic activity that circularizes single-stranded DNA. It offers the potential for RCA to be used on compromised DNA materials. In effect, highly fragmented DNA that was a poor substrate for PCR can be converted into circular DNA and hence become a good template for RCA.

Repairing the DNA of lesions that have occurred after exposure to environmental insults is potentially a viable approach for increasing the number of template molecules obtained from degraded DNA samples. Some DNA repair kits have begun to reach the scientific community, such as PreCR Repair Mix (New England Biolabs, Ipswich, MA, USA). WGA might then be possible by employing linear multiple strand displacement.

Improvements to typing low-quality samples

Mini-STRs

Forensic samples are often compromised to the point where obtaining DNA typing results is challenging or not possible. Highly degraded samples that contain DNA fragments that are too short in length cannot be used to generate amplicons that are longer in length. Whether for LCN (to increase the number of available

template molecules) or for conventional DNA typing (where the sample is degraded), reducing the size of the amplicons generated during the PCR will increase typing success by providing more efficient amplification.

The STR loci are currently the most informative genetic markers for identity testing. To improve success in STR typing with degraded DNA, increase sensitivity of detection, and enhance robustness or the quality of the assay, the PCR primers for the STR loci can be repositioned so they reside closer to the repeat (i.e., polymorphic) region (36). Thus, the amplified PCR products will be reduced in length, and if smaller than some of the fragmented DNA template molecules, genetic characterization of the sample may then be possible (Figure 1). Mulero et al. (37) described the conversion of eight STRs (D7S820, D13S317, D16S539, D21S11, D2S1338, D18S51, CSF1PO, and FGA) into mini-STRs and the development of the commercial AmpFISTR MiniFiler PCR Amplification Kit (Applied Biosystems, Foster City, CA, USA). The amplicon range for the mini-STRs spans 71–250 bp in length. Since the amplicon products of mini-STRs will overlap in size more so than those in conventional STR kits, the four fluorescent dye tag system was increased to five dyes. Thus, the mini-STRs could be accommodated into one multiplex analysis. It would make sense that all STR kits be reconfigured into mini-STR kits for routine analysis of forensic evidence. Only one attempt would be needed to obtain a DNA profile, if possible, for a degraded, limited-quantity sample. More efforts will continue to convert all current forensically relevant STRs into mini-STRs. To do so and to incorporate at least the core forensic STRs in one multiplex amplification kit, it is likely that fluorescent capillary electrophoresis-based systems will require additional tags or dyes to resolve more size-overlapping mini-STRs. Likely additional novel mini-STRs will be sought; some already have been recommended. For example, Gill et al. (38) described the acceptance of three new mini-STRs (D10S1248, D14S1434, D22S1045) into the European standard Interpol loci which now comprise 10 STR loci.

Single-nucleotide polymorphisms (SNPs)

A different class of genetic markers, known as single-nucleotide polymorphisms (SNPs),

can be typed in much smaller amplicons than those of mini-STRs, in theory as short as 50 bp (3). Because SNPs are bi-allelic, on a per-locus basis they are less informative for identity testing than a STR locus. Thus, a large battery of SNPs is required for typing to attain the same level of discrimination as that of the core forensic STR loci. An advantage for SNPs, however, is that substantial research and development currently is underway to improve analytical capabilities, possibly making large multiplex assays and complete automation feasible. Multiplexing—the ability to simultaneously type several genetic markers in one analysis—is key to forensic testing. More simultaneously typed markers will reduce consumption of often-limited precious forensic evidence so that in more cases some sample can remain if re-testing is requested. Multiplexing also yields more information regarding the source of the sample. Furthermore, because of fewer manipulations when only one analysis is carried out, the chance of intralaboratory contamination is reduced. The difficulty with multiplexing, however, is that as more markers are added to a multiplex, detection sensitivity begins to be compromised. More efforts will be focused on improving multiplexing capabilities through molecular biology approaches and by instrumentation so the large battery of required SNPs can be exploited and provide the power of discrimination currently afforded with STR kits.

Initially, the analytical platforms for SNP detection will likely be the same as those used for STR typing (i.e., fluorescent detection capillary electrophoresis instruments, which are currently used in forensic laboratories). However, novel platforms will be sought, such as the mass spectrometer (39–43). Mass spectrometry offers greater resolution, mass accuracy, and automation without the need for fluorescent labels. Mitochondrial DNA SNP analysis by mass spectroscopy already has been developed (40) and offers a number of benefits that include obtaining more variation than other SNP-based technologies, analysis of heteroplasmic profiles, quantitation, and automation. New instrumentation is constantly being developed and the choice for forensic analyses will ultimately depend on reliability, sensitivity of detection, throughput, and cost.

Since ~85% of human genomic variation is based on SNPs, there is an abundance of SNPs for human forensic identity testing. Research will continue to discover SNPs that are useful for the characterization of biological evidence. But not all SNPs are the same. Budowle and van Daal (3) categorized SNPs into (i) identity-testing SNPs for individualization, requiring high heterozygosity and low population heterogeneity; (ii) lineage-infor-

mative SNPs, sets of tightly-linked SNPs that function as haplotype markers to identify missing persons through kinship analyses; (iii) ancestry-informative SNPs for establishing high probability of an individual's biogeographical ancestry to indirectly infer some phenotypic characteristics for investigative lead value, requiring low heterozygosity and high population heterogeneity; and (iv) phenotype-informative SNPs for establishing high probability that an individual has a particular phenotypic characteristic such as skin color, hair color, or eye color, for investigative lead value. A fifth class of SNPs is those for pharmacogenetic investigations for determining the cause of death (44,45).

Predictably, efforts first will be in selecting identity-testing SNPs. These SNPs are likely to be incorporated into forensic molecular biology more readily because they have the broadest application and will enable the analysis of highly degraded human-remains samples, such as bones, teeth, and hair. Research is ongoing in selecting the core identity-testing SNPs (46–49). These identity-testing SNPs eventually will make their way into more routine casework because of desires to analyze LCN samples and for database searching for developing investigative leads (see "Database searches" section). The effort

will be to select a universal consensus set that applies to the majority of populations worldwide. To date, a few attempts to develop identity SNP panels have been reported, such as a 21-SNP panel (46), 52-SNP panel (49) by European Forensic Laboratories, and >40 SNPs described by Kidd et al. (47,48). These SNPs will be evaluated for population criteria of high heterozygosity and low population heterogeneity in a large number of relevant populations, linkage disequilibrium, and PCR design compatibility. If these criteria are met, then the selected panel will be universally useful for identity testing.

Lineage-based genetic markers are quite useful since the analysis of missing persons and unidentified human remains sometimes involve complex kinship scenarios. Typically, such analyses utilize genetic markers residing on the maternally-inherited mitochondrial genome and the paternally-inherited Y chromosome. Unfortunately, these lineage-based systems lack a high power of discrimination, so identity cannot be assigned with a high degree of confidence. Existence of haploblock structures in the human genome, revealed by using data from the International HapMap project (50–54), offers an opportunity to explore novel panels of SNP markers that would perform a similar function to

non-autosomal lineage-based genetic markers. Several tightly linked autosomal SNPs that are inherited together form a haplotype block. As a unit, the haploblock has higher discrimination power for kinship analysis than the individual SNPs within the block. Ge et al. (55) described selection criteria for candidate haploblocks to include linkage disequilibrium of SNPs comprising the block, low and high levels of population heterogeneity, and haploblock conformance to Hardy-Weinberg equilibrium expectations. Several haploblocks have been identified, and it is likely more will be identified through additional research efforts. Haploblock panels will enable highly discriminating assays best suited for relationship testing, familial assessment, and admixture analysis.

Investigative information

Phenotypic information from a DNA sample

When there is no suspect, SNPs that describe phenotypic traits would enable a genetic prediction of appearance for investigative leads to identify the perpetrator of a crime (3). If an individual's pigmentation, facial features and height can be predicted, investigators may be able to eliminate potential

NEW MICROCALORIMETRY PRODUCTS FROM TA INSTRUMENTS

...The World Leader in
Thermal Analysis, Rheology, and Microcalorimetry

NANO DSC

- High Sensitivity Scanning Calorimeter
- Protein Denaturation, Stability

NANO ITC

- High Sensitivity Titration Calorimeter
- Protein Binding, Interactions, Kinetics

TAM III

- High Sensitivity Isothermal Calorimeter
- Stability, Compatibility, Morphology



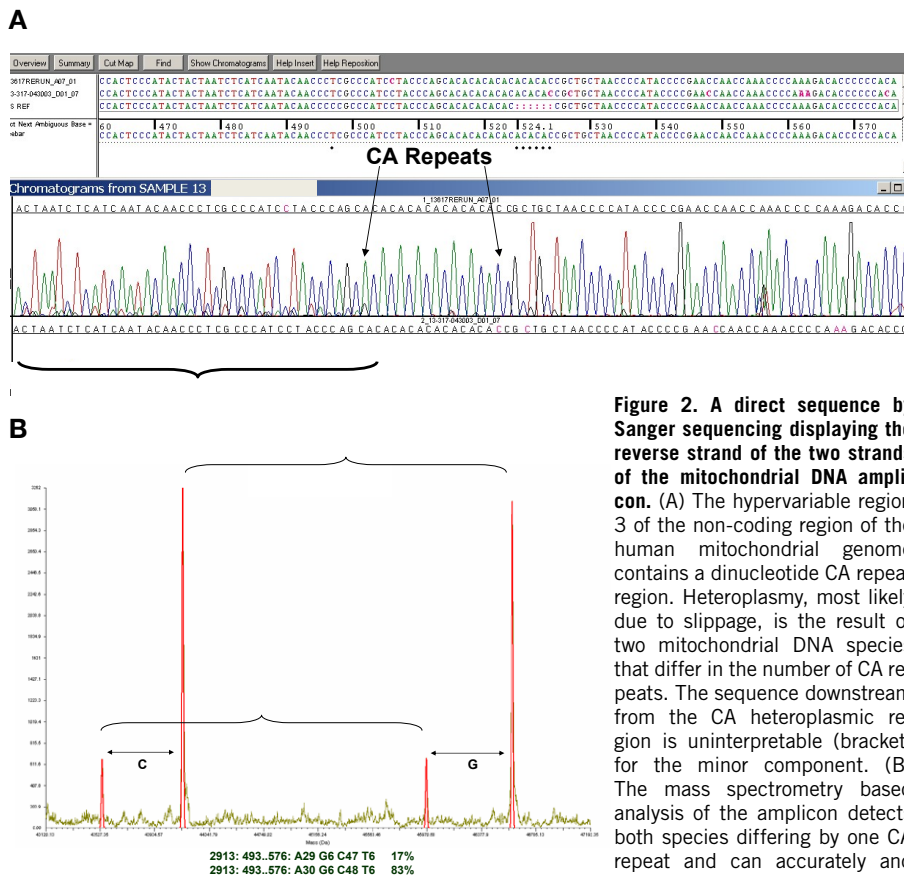


Figure 2. A direct sequence by Sanger sequencing displaying the reverse strand of the two strands of the mitochondrial DNA amplification. (A) The hypervariable region 3 of the non-coding region of the human mitochondrial genome contains a dinucleotide CA repeat region. Heteroplasmy, most likely due to slippage, is the result of two mitochondrial DNA species that differ in the number of CA repeats. The sequence downstream from the CA heteroplasmic region is uninterpretable (bracket) for the minor component. (B) The mass spectrometry based analysis of the amplicon detects both species differing by one CA repeat and can accurately and simultaneously measure the relative abundance of the components. The base composition and relative amounts for both species are displayed. Figure kindly provided by L. McCurdy, FBI Laboratory, Quantico, VA, USA.

Figure kindly provided by L. McCurdy, FBI Laboratory, Quantico, VA, USA.

suspects, focus their search, and at a minimum help confirm or refute the more refractory eyewitness description. Once a suspect is identified, his or her reference DNA profile can be compared with the evidence profile using standard DNA markers for inculpation or exculpation. The same phenotypic SNPs could be used to facilitate facial reconstructions for identifying missing persons. The research to discover phenotypic SNPs has identified a few good candidates, but more genome-wide scans will be needed to develop a battery of informative SNPs. Because these SNPs may reside anywhere within or around a gene, alternative analytical technologies will be needed than what will be used for identity testing and lineage SNPs. The position of the informative SNP that may infer a particular phenotype may not always be known a priori, so the assay(s) will need to be able to determine the full sequence of the gene(s) of interest; the same demand will likely hold for pharmacogenetic SNPs (see next section, below). It is likely that sequencing-based technologies will be better suited for scanning genes for informative phenotypic SNPs in forensic evidence. This sequencing approach, combined with whole genome screens and association studies will be used to identify the causal SNPs or SNPs in high linkage disequilibrium (LD) that can be used to predict phenotype.

Pharmacogenetic information from a DNA sample (molecular autopsy)

Genetic variation and its effects on metabolism can be applied to postmortem analysis to help resolve some cases initially believed to be suicides or classified as sudden, unexplained deaths, especially in cases where poisoning, incapacitation, inebriation, or certain diseases where pharmacotherapy is an essential treatment (such as epilepsy, depression, cardiac disease, or diabetes) are factors in the cause of death. Individuals vary in their response to drugs or physical exertion (44,45,56–58). Some people, for example, can metabolize a drug better than others due to pharmacogenetic SNPs in or around specific encoded enzymes (such as those in the human cytochrome P450 monooxygenase superfamily) (44,45,56,58). Those who have a genetic makeup that enables very rapid metabolism of a drug may receive no benefit from a certain administered dose. In contrast, those individuals who cannot metabolize the drug may be poisoned by accumulation or overdose. Rodriguez-Calvo et al. (59) recently reviewed the potential role for genetic analysis into the cause of sudden cardiac death (SCD). SCD is one of the most common causes of death in developed countries and even though it is a highly heterogeneous and variable penetrance group of diseases, some elucidation into

genetic associations and cardiac disease are emerging. For example, there may be explanations for sudden death in some cases of apparently healthy young people (i.e., <35 years of age). Scientists are exploring these metabolic differences among individuals and how they impact the cause of drug-related or unexplained deaths. Pharmacogenetic SNPs will eventually make their way into molecular autopsy protocols in pathology laboratories. An additional benefit is that pharmacogenetic analysis can help determine the cause and manner of death and may provide health information (certainly only via proper ethical disclosure practices) to at-risk relatives.

Expression analysis to determine tissue type

A matching DNA profile comprised of the core set of STR loci is very strong evidence regarding the source of a sample. Additional information regarding the tissue source of that sample can be useful: for example, determining whether the source of the DNA was from semen instead of saliva can help reconstruct how a sexual assault transpired. Crime scenes are rarely pristine and stains that are apparent may be human in origin or could be from other organic or inorganic sources. Being able to screen these samples for human origin and tissue specificity can reduce unnecessary DNA typing. Most presumptive and confirmatory serological tests for species specificity and tissue origin (limited to blood, semen, and saliva) are based on immunological or catalytic assays. Conventional serological methods of tissue identification are laborious, use diverse techniques, consume significant amounts of sample, and are costly. While the DNA in each tissue is essentially the same, the mRNA and protein profiles are substantially different. The differences in the proteins, which are the target of serological assays, account for the distinctive properties of the tissues. There are no confirmatory tests for some of the typically encountered tissues, such as saliva and vaginal secretions, making a serological approach to tissue identification problematic.

An alternative approach would be the use of low- to medium-density expression profiling for typing of the presence of mRNA species that are tissue-specific (60,61). Multiplex reverse transcription PCR (RT-PCR) methods for tissue identification for blood, saliva, semen, and vaginal secretions have some appeal because they can be assayed using the same platforms used for current DNA typing assays, can provide specificity for tissues of interest, and RNA can be recovered during DNA extraction thus reducing sample consumption. Work will continue on selecting genes that are expressed in only one tissue, developing assays that parallel DNA diagnostic methods, and determining the degree of stability of mRNA in aged and environmentally exposed samples.

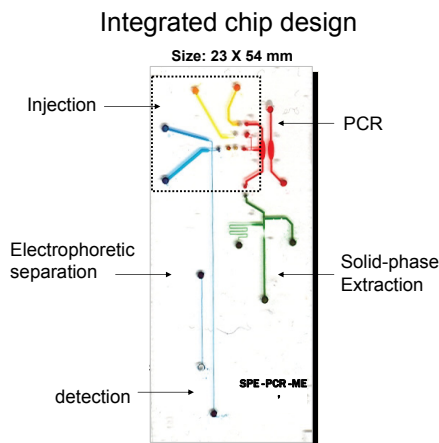


Figure 3. Design of an integrated chip for the entire DNA analysis process. Figure kindly provided by J.P. Landers, University of Virginia, Charlottesville, VA, USA.

Alternatively, a tissue's proteomic profiles could be determined by mass spectrometry. The method of choice will likely depend on which species (RNA or protein) is more stable and more abundant in forensic samples.

Database searches

A number of countries have established DNA databanks that contain DNA profiles from, at a minimum, convicted offenders and forensic samples from unsolved cases (18,19). These databases are designed to help solve future crimes or identify missing persons by providing genetic investigative leads. The United States' CODIS databank houses the largest number of DNA profiles compared with any other offender/forensic DNA database (Table 1). There are indices for crime-scene evidence, individuals convicted of felonies, arrestees (in some states), missing persons, human remains, and family members. Because of their success in providing investigative leads, these databases continue to increase in size and may provide additional information other than solely direct matching of DNA profiles for investigative leads.

In order for DNA profile databanks to be useful at a national (or international) level, standardization of the genetic markers used among laboratories was essential. In order to ensure comparability of DNA profiles across the United States, for example, the STR loci for characterizing DNA reference samples and forensic samples were standardized (18). Thirteen autosomal STR loci were selected as core markers for CODIS. (They are CSF1PO, FGA, THO1, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, and D21S11.) The number and specific loci vary between some countries but a core set is common to all forensic DNA databases. A DNA profile comprised of these thirteen STR loci often yields a strong investigative lead when a direct-match hit is obtained

through a database search. However, the full capability of developing investigative leads is not exploited by searching only for direct matches.

Although the database-searching algorithms are designed to facilitate obtaining direct matches (e.g., matching a forensic sample profile with a convicted-felon reference profile) some have sought to use the large profile archive to develop potential investigative leads by identifying possible relatives of the source of an evidentiary sample through kinship or familial inferences (62). On average, close relatives (i.e., parents, offspring, and siblings) share more alleles than do unrelated individuals. Therefore, despite a lack of a direct match in a database search, a partially matching profile still may be informative. While the use of this form of database searching—known as familial searching—is being debated on legislative and civil-rights grounds, some states and countries have proceeded and identifications of relatives have been obtained. It is likely that if more kinship associations ultimately result in solving crime, there will be more motivation to further exploit familial searching.

In addition to the need for better searching algorithms, certain limitations exist regarding the use of familial searching. When searching large databases with the thirteen core STR loci there will be a large number of fortuitous hits (possibly hundreds) that cannot be excluded as potential relatives. Moreover, it is likely that the top hits (i.e., the strongest associations) often will be with unrelated individuals. This phenomenon is due to the fact that the thirteen STR loci are not sufficiently resolving to be an efficient screen for one-to-one kinship analysis. Another exacerbating factor is that a mutation in one of the true relatives could appear as an exclusion. Searches have to tolerate a degree of "mismatch" and thus allow for more fortuitous candidates. To overcome this limitation, additional genetic markers are needed for more efficient searching of candidates. Since most individuals represented in these DNA databanks are males and close male relatives are being sought with familial searching, the use of genetic markers on the paternally inherited Y chromosome will substantially reduce the number of candidate hits. If familial searching becomes routine, then reference samples should be typed for Y STRs in addition to the core autosomal STRs (63). To further reduce the list of candidates, identity- and kinship-testing SNPs would be good markers for additional genetic characterization. SNPs also have very low mutation rates and because of their smaller amplicon size they may provide data on substantially degraded samples. A battery of SNPs with a high power of discrimination is desired. Commercial kits are needed that provide the reagents necessary to multiplex Y chromosome STRs with the core CODIS STRs or possibly combine a

large suite of SNPs (as would be developed for missing persons identifications) with the core CODIS STR loci. Being able to multiplex these markers would be an economic boon, enabling one analysis for both sets of markers and thus reducing labor and cost for typing reference database samples.

Microbial forensics and high-throughput sequencing

The threat of terrorist or criminal use of microorganisms and their toxins is a great concern for biodefense and biosecurity worldwide (64,65). The anthrax-letters attack of 2001 demonstrated the public's vulnerability to such attacks and the U.S. government's inability to forensically investigate the evidence for attribution purposes (66). This resulted in the birth of the field of microbial forensics. Microbial forensics is an evolving subdiscipline of forensic science for analyzing evidence from a bioterrorism act, biocrime, hoax, or an inadvertent release for attribution purposes (65). In many ways, microbial forensics is not a novel field; its bases and practices are derived from similar approaches established in public health and epidemiology (67). The difference between microbial forensics and epidemiology is that the former desires to further individualize a sample. Nonetheless, microbial forensic analyses must encompass sample handling, collection, preservation, method selection, casework analysis, interpretation of results, validation, and quality assurance.

Molecular genetics, genomics and informatics will be central to species/strain identification, virulence determination, pathogenicity characterization, and source attribution. The ultimate in source attribution is to be able to individualize a sample such that it can be traced to a unique source. That is unlikely with current capabilities and may not be possible in many cases because of the nature of microbiological samples. Epidemiologic investigations tend to focus on species and strain level resolution, which are helpful information for a microbial forensics investigation (67). However, forensic science endeavors to individualize samples: for the anthrax-letters attack, a multi-locus variable-number tandem repeat (VNTR) analysis technique was used to identify the *Bacillus anthracis* bacteria as that of the Ames strain. While the strain data appropriately focused the investigation toward laboratory sources, differentiating closely-related laboratory samples of the same strain was far more challenging. For future cases, technology is needed to facilitate identification of those unique SNPs, duplications, deletions, insertions, or rearrangements—if they exist—that will better individualize samples and help focus an investigation (68). Unlike human identification, where

a standardized core set of loci can be used to differentiate individuals, the microbial forensic marker(s) for individualization will be unknown and case-specific. Whole-genome sequencing is the preferred method for discovering genetic variation of forensic value (68,69). The most effective approach for comprehensive genetic variation discovery, which was used in the anthrax-letter investigations (69), has been by high-throughput shotgun sequencing exploiting Sanger sequencing (70). Though considered to be the gold standard of sequencing technology, this method is laborious, costly, has relatively low coverage, and exhibits sample bias problems. If whole-genome resequencing were desired for a repository of samples (of a few to thousands) the cost would be prohibitive.

Therefore, advances in sequencing technology are needed that increase accuracy and speed, reduce cost, and maximize efficiency for forensic analysis. Hybridization resequencing [such as the chip technology developed by Affymetrix (Santa Clara, CA, USA) (71)] enables an extremely large number of probings to be carried out simultaneously and would provide fast turnaround for typing results (71–73). But hybridization chip technology may not have the sensitivity of detection required for forensic applications. Matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF), which exploits the absolute mass of a nucleic acid molecule as an intrinsic property, offers advantages over hybridization and electrophoresis approaches (41,42): MALDI-TOF is not subject to the vagaries of electrophoretic anomalies and DNA secondary structure, and does not require labeling molecules for detection. However, it cannot be used to sequence whole genomes. Second-generation massively parallel sequencing technology, such as SOLiD (Applied Biosystems, Foster City, CA, USA), Genome Analyzer (Illumina, San Diego, CA, USA), and 454 GS-FLX (454 Life Sciences, Roche, Branford, CT, USA), and other related technologies offer rapid resequencing of whole bacterial genomes with high coverage (74–83). Third-generation single-molecule sequencing technologies, such as that in development by Pacific Biosciences (Menlo Park, CA, USA) and Helicos BioSciences Corporation (Cambridge, MA, USA), may supplant the current novel sequencing technologies (84). However, single-molecule approaches can have sampling issues that will need to be addressed and the current technologies, being more in the proof-of-concept phase, are far from robust. It is difficult to predict what technologies will be selected for microbial forensics, but low-cost, high-coverage, low-error, high-throughput sequencing of whole microorganism genomes will be a necessity for supporting development

of the most effective microbial forensics attribution assays.

Automation

The demands of generating, entering, and maintaining DNA profiles in a national DNA database have driven developments in automation. The number of reference samples from convicted felons, arrestees, detainees, and missing persons continues to increase, and the burden is such that these samples cannot continue to be typed and reviewed manually. Robotics and modified chemistries more amenable to automated processes have been developed to increase throughput and efforts will continue to improve automation efficiency (85–91).

Automation offers quality control, consistent results, and data management with lower operational costs. By removing the human component from the process, results tend to be more consistent and high-quality. Error is reduced primarily by minimizing the chance of sample switching and carryover contamination. Software developments enable tracking of sample handling throughout the process. Lower reagent volumes translate into fewer consumables and less waste.

Most automation has focused on the extraction of DNA from standard reference samples and some have extended the application to casework samples such as bone, hair, teeth, cigarette butts, and sperm. The robotic platforms vary and include the Tecan Genesis RSP 150/8 robotic workstation and the Tecan Freedom EVO liquid handling stations (Tecan, Mannedorf, Switzerland), the Biomek 2000 automation workstation (Beckman Coulter, Fullerton, CA, USA), the Plato 3000 robotic system (Rosys/Anthos AG, Hombrechtikon, Switzerland), and the BioRobot EZ1 System and BioRobot 8000 workstation (Qiagen, Dusseldorf, Germany), to name a few (85–90). The development and implementation of robotic workstations require alternative chemistries for extraction. Some parts of a manual extraction are not accommodated readily by a robotic system, such as organic solvent extraction, centrifugation and boiling. Solid-phase extraction chemistries, such as the DNA IQ (Promega, Madison, WI, USA) (92) and the Qiagen EZ1, QIA Symphony Investigator Kit, and QIAamp Investigator BioRobot Kit (Qiagen) have been adopted to facilitate automation of extraction (85, 87–89).

It is important to quantify the DNA and normalize the amount that is used in PCR to obtain more consistent typing results. Greenspoon et al. (87) used the same robotic platform (the Biomek 2000 Automation Workstation) for extraction, DNA quantitation, and PCR setup, thus automating three parts of the process prior to the PCR step. In addition, the remaining DNA

extracts are transferred directly to storage tubes for long-term archiving. Automation has been and will continue to be developed for the protocols encountered in the forensic laboratory. However, these robotic systems are macroscale approaches that have yet to capitalize on the potential benefits of microscale technologies (see “In-field testing” section).

Forensic analysis of mitochondrial DNA (mtDNA) often provides results for samples where nuclear autosomal marker analyses are difficult or impossible (such as old bones, teeth, and hair shafts) (21). Typing generally involves the PCR amplification of two short regions of mtDNA called hypervariable regions 1 and 2 (HV1 and HV2), followed by direct sequencing of the PCR products by Sanger sequencing. This process is laborious, time-consuming, and costly. Additionally, data analysis can be confounded by sequence artifacts, electrophoretic anomalies, the presence of heteroplasmy (i.e., the presence of more than one mitochondrial genome variant within an individual) and limited ability to quantify the components of a mixed sample. Recently, multiplex PCR followed by electrospray ionization time-of-flight mass spectrometry (ESI-TOF-MS) was demonstrated to be applicable for typing the hypervariable regions of human mtDNA, expanding the discriminating potential of an assay beyond that of specific SNP targeting (40). Additionally, heteroplasmic samples can be analyzed and the relative quantity of the components of mixed samples can be determined (Figure 2). The T5000 Biosensor (Ibis Biosciences Inc., Carlsbad, CA, USA) (39) combines robotic workstations for PCR and sample cleanup (i.e., de-salting) with mass spectrometry, so mtDNA typing can be performed with at least a 10-fold increase in throughput and a 5-fold decrease in reagent cost, with no loss in sensitivity and little or no loss in information compared with traditional sequencing. This platform holds promise for readily accommodating other genetic marker assays.

The areas where automation has yet to improve throughput sufficiently are at the front and back ends of the analysis. For reference samples, some success has been achieved because the sample format can be standardized. However, for casework, the sample types and the substrates on which they reside vary substantially, making it difficult to standardize the initial sample preparation. A blood stain may reside on a non-porous car bumper or on porous wood, a bone sample requires pulverization, a semen sample may reside on a vaginal swab, a clothing item, and so on. Automating the sample preparation of casework materials is likely to be the most challenging endeavor for forensic scientists.

The back end of the process is the interpretation of results. Algorithms will be needed

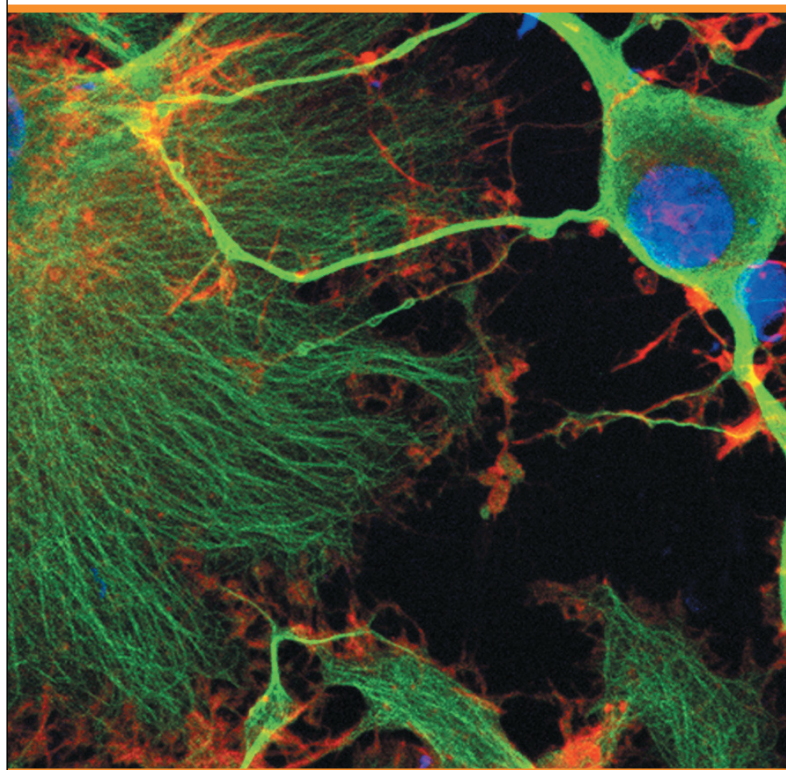
to facilitate this very labor-intensive step. Currently, two qualified scientists are required to manually read a DNA profile (whether STRs or sequences). Expert systems are being developed to replace one if not both of the scientists for typing STR reference samples entered into DNA databanks (93,94). Such efforts will continue and the design of expert systems will be attempted for interpreting the more challenging casework samples (93–95).

The platforms currently used in forensic laboratories serve their purposes but are still microfluidics-based systems requiring relatively large volumes of reagents and generating relatively large volumes of waste. Additionally, they tend to be modular. Most of these robotic systems automate parts of the analytical process. An individual is needed to move the microplate from one robotic system carrying out one function (e.g., extraction) to another (e.g., PCR), and eventually to the capillary electrophoresis instrument. Integration of all facets from extraction to detection has yet to be realized. However, micro-fabricated devices offer the possibility of automating the entire analytical process and freeing the analyst to carry out other tasks.

In-field testing

There is some interest in the ability to perform DNA diagnostics at the crime scene. For microbial forensics and public health, the need is paramount to be able to determine the presence of microorganisms that are harmful to humans. A biocrime scene requires investigators to wear protective equipment, making it difficult to work for prolonged periods of time. If the scene could first be determined safe (as in the case of a hoax, for example), this onerous requirement for sample collection could be omitted. The instrumentation for pathogen detection should be portable, not just transportable. The diagnostic capability should have a high degree of sensitivity of detection and be able to detect a wide range of known harmful pathogens, as well as the genes that confer pathogenicity, in case genetic engineering was used to modify an otherwise harmless microorganism. Microfluidics has appeal because it enables molecular biology analyses to be carried out on miniaturized platforms that integrate all aspects of the analysis from sample preparation to nucleic acid typing [e.g., the lab-on-a-chip concept (96, 97)]. Additional potential benefits of microfluidics include reduced sample consumption and reagents (lowering cost), less waste, better thermodynamics during the PCR (that possibly could reduce stochastic effects with limited template), and less contamination (being an integrated closed system) (96–98). It is conceivable that throughput would increase by decreasing analysis time and exploiting parallel processing. Analysts would also be freed from some manual processes that are still encumbered with macrofluidic manipulations. Development in this area is exemplified by the research efforts at the Landers laboratory (University of Virginia, Charlottesville, VA, USA), which has demonstrated that a wide range of samples and even differential extraction (i.e., isolating sperm DNA from DNA from other cell types) can be accommodated in a microfluidic format. Thus, the macroscale of samples and the microscale of extraction analysis requirements can be bridged. In addition, they have developed an integrated system that enables the entire process from sample extraction through electrophoresis to be carried out therein (96, 97, 99–101) (Figure 3).

There are advocates for a portable field testing microfluidic device for performing human identification DNA typing at the crime scene to rapidly identify suspects. Presumably, this would be by generating a profile and immediately searching a DNA database for developing an investigative lead. It certainly would not be used for eliminating lingering suspects: even if the perpetrator remained at the crime scene, obtaining a reference sample would require probable cause and thus is not amenable to rapid response. A significant concern would be the possible contamination of evidence by reference samples in a suboptimally controlled environment. The crime scene is a chaotic environment and it is important to control the scene, and efforts should be focused on proper collection of



BioTechniques Weekly

Keep up-to-date with this free weekly electronic newsletter delivered every Thursday to *BioTechniques* subscribers. Coverage includes: **Science news, Recent top grants, New products, Employment opportunities, Event calendar**, and more!

Not a subscriber? Sign up free at
www.BioTechniques.com/newsletters



evidence and to minimize its contamination. If DNA typing were performed at the crime scene, then qualified practitioners would have to be deployed, since expertise is required for DNA typing and interpretation of the generated DNA profiles. This deployment would reduce the throughput of an already backlogged laboratory: scientists would be occupied going to and from crime scenes and could only work one case at a time, and there would not be enough qualified personnel to analyze DNA at multiple, simultaneous crime scenes. One solution is for scientists to remain remote and the profiles transmitted electronically. This approach still does not address the need for properly trained practitioners to process the samples and carry out the analytical portion of the assay. Again, the front-end sample preparation is perhaps the biggest hurdle to overcome, since crime-scene samples present themselves in myriad manners and these macro-samples may not be readily amenable to microfluidic processing. However, a microdevice may be useful at the crime scene for collection and sample storage.

Rarely is the time to move a sample from the crime scene to the laboratory an impediment. However, casework continues to increase while manpower does not, concomitantly. Fully integrated automated systems hold promise for increasing throughput by freeing an analyst from several manual aspects of the process so he or she can focus on other, more demanding processes. Automation is essential to address the increasing demand for casework analysis and for generating and entering samples into national DNA databases.

Conclusion

There still are a number of gaps that need to be addressed in forensic biology. We have identified some of these areas where further development is needed: improving the current limits on typing samples of low quantity and quality; improving the efficiency of sample recovery and extraction; converting current STRs to mini-STRs; selecting and validating new mini-STRs; selecting and validating a variety of SNPs for different applications; enhancing multiplexing; developing automation for high throughput; developing expert systems for data interpretation; developing sequencing capabilities for screening microorganism genomes; and field testing. There are likely other gaps as well. We did not address identification beyond that of humans and microorganisms. Plant and animal forensic genetics may have additional requirements that molecular biology may resolve. The future of molecular biology for forensic science will be exciting and dynamic. There is still much to achieve and molecular biology developments will be essential for

assisting in solving crimes and identifying missing persons.

Acknowledgments

The authors declare no competing interests.

References

- Budowle, B. and A.J. Eisenberg. 2007. Forensic Genetics, p.501-517. *In* D.L. Rimooin, J.M. Connor, R.E. Pyeritz, and B.R. Korf (Eds.), Emery and Rimooin's Principles and Practice of Medical Genetics, 5th ed. Vol. 1. Elsevier, Philadelphia.
- Budowle, B., J.V. Planz, R. Campbell, and A.J. Eisenberg. 2005. Molecular diagnostic applications in forensic science, p.267-280. *In* G. Patrinos and W. Ansorge (Eds.), Molecular Diagnostics. Elsevier, Amsterdam.
- Budowle, B. and A. van Daal. 2008. Forensically relevant SNP classes. *BioTechniques* 44:603-610.
- Budowle, B., H.A. Deadman, R.S. Murch, and F.S. Baechtel. 1988. An introduction to the methods of DNA analysis under investigation in the FBI Laboratory. *Crime Lab. Dig.* 15:8-21.
- Scientific Working Group on DNA Analysis Methods (SWGDM). 2004. Revised Validation Guidelines. *Forensic Sci. Commun.* 6 (http://www.fbi.gov/hq/lab/fsc/backissu/july2004/standards/2004_03_standards02.htm).
- Budowle, B. and F.S. Baechtel. 1990. Modifications to improve the effectiveness of restriction fragment length polymorphism typing. *Appl. Theor. Electrophor.* 1:181-187.
- Gill, P., A.J. Jeffreys, and D.J. Werrett. 1985. Forensic application of DNA fingerprints. *Nature* 318:577-579.
- Jeffreys, A.J., V. Wilson, and S.L. Thein. 1985. Hypervariable minisatellite regions in human DNA. *Nature* 314:67-73.
- Budowle, B., R. Chakraborty, A.M. Giusti, A.J. Eisenberg, and R.C. Allen. 1991. Analysis of the VNTR locus D1S80 by the PCR followed by high-resolution PAGE. *Am. J. Hum. Genet.* 48:137-144.
- Budowle, B., J.A. Smith, T. Moretti, and J. DiZinno. 2000. DNA Typing Protocols: Molecular Biology and Forensic Analysis. Eaton Publishing Company/BioTechniques Books, Natick, MA.
- Collins, P.J., L.K. Hennessy, C.S. Leibelt, R.K. Roby, D.J. Reeder, and P.A. Foxall. 2004. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433, and amelogenin: the AmpFISTR Identifiler PCR Amplification Kit. *J. Forensic Sci.* 49:1265-1277.
- Green, R.L., I.C. Roinestad, C. Boland, and L.K. Hennessy. 2005. Developmental validation of the Quantifiler Real-Time PCR Kits for the quantification of human nuclear DNA samples. *J. Forensic Sci.* 50:809-825.
- Kayser, M., P. de Knijff, P. Deltjes, M. Krawczak, M. Nagy, T. Zerjal, A. Pandya, C. Tyler-Smith, and L. Roewer. 1997. Applications of microsatellite-based Y chromosome haplotyping. *Electrophoresis* 18:1602-1607.
- Kimpton, C.P., P. Gill, A. Walton, A. Urquhart, E.S. Millican, and M. Adams. 1993. Automated DNA profiling employing multiplex amplification of short tandem repeat loci. *PCR Methods Appl.* 3:13-22.
- Krenke, B.E., A. Tereba, S.J. Anderson, E. Buel, S. Culhane, C.J. Finis, C.S. Tomsey, J.M. Zachetti, et al. 2002. Validation of a 16-locus fluorescent multiplex system. *J. Forensic Sci.* 47:773-785.
- Moretti, T.R., A.L. Baumstark, D.A. Defenbaugh, K.M. Keys, and B. Budowle. 2001. Validation of short tandem repeats (STRs) for forensic usage: performance testing of fluorescent multiplex STR systems and analysis of authentic and simulated forensic samples. *J. Forensic Sci.* 46:647-660.
- Wilson, M.R., J.A. DiZinno, D. Polansky, J. Replogle, and B. Budowle. 1995. Validation of mitochondrial DNA sequencing for forensic casework analysis. *Int. J. Legal Med.* 108:68-74.
- Budowle, B., T.R. Moretti, S.J. Niezgoda, and B.L. Brown. 1998. CODIS and PCR-based short tandem repeat loci: law enforcement tools, pp. 73-88. *Proceedings from the Second European Symposium on Human Identification 1998*. Promega Corporation, Madison.
- Martin, P.D., H. Schmitter, and P.M. Schneider. 2001. A brief history of the formation of DNA databases in forensic science within Europe. *Forensic Sci. Int.* 119:225-231.
- CODIS—NDIS Statistics. Federal Bureau of Investigation. (<http://www.fbi.gov/hq/lab/codis/clickmap.htm>).
- Budowle, B., M.W. Allard, M.R. Wilson, and R. Chakraborty. 2003. Forensics and mitochondrial DNA: applications, debates, and foundations. *Annu. Rev. Genomics Hum. Genet.* 4:119-141.
- Findlay, I., R. Frazier, A. Taylor, and A. Urquhart. 1997. Single cell DNA fingerprinting for forensic applications. *Nature* 389:555-556.
- van Oorschot, R.A. and M.K. Jones. 1997. DNA fingerprints from fingerprints. *Nature* 387:767.
- Budowle, B., D.L. Hobson, J.B. Smerick, and J.A. Smith. L. 2001. Low copy number: consideration and caution, p. 1-9. *Proceedings from the Twelfth International Symposium on Human Identification 2001*, Promega Corporation, Madison. (<http://www.promega.com/geneticidproc/ussymp12proc/contents/budowle.pdf>).
- Gill, P., J. Whitaker, C. Flaxman, N. Brown, and J. Buckleton. 2000. An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA. *Forensic Sci. Int.* 112:17-40.
- Barbaro, A., G. Falcone, and A. Barbaro. 2000. DNA typing from hair shaft. *Prog. Forensic Genet.* 8:523-525.
- Kloosterman, A.D. and P. Kersbergen. 2003. Efficacy and limits of genotyping low copy number (LCN) DNA samples by multiplex PCR of STR loci. *J. Soc. Biol.* 197:351-359.
- Wiegand, P. and M. Kleiber. 1997. DNA typing of epithelial cells after strangulation. *Int. J. Legal Med.* 110:181-183.
- Shaw, K., I. Sesardic, N. Bristol, C. Ames, K. Dagnall, C. Ellis, F. Whittaker, and B. Daniel. 2008. Comparison of the effects of sterilization techniques on subsequent DNA profiling. *Int. J. Legal Med.* 122:29-33.
- Caddy, B., D.R. Taylor, and A.M.T. Lincare. 2008. A review of the science of low template DNA analysis, p. 1-35. (http://police.homeoffice.gov.uk/publications/operational-policing/Review_of_Low_Template_DNA_1.pdf?view=Binary).
- Hochmeister, M.N., B. Budowle, U.V. Borer, U.T. Eggmann, C.T. Comey, and R. Dirnhofer. 1991. Typing of DNA extracted from compact bone tissue from human remains. *J. Forensic Sci.* 36:1649-1661.
- Dean, F.B., S. Hosono, L. Fang, X. Wu, A.F. Faruqi, P. Bray-Ward, Z. Sun, Q. Zong, et al. 2002. Comprehensive human genome amplification using multiple displacement amplification. *Proc. Natl. Acad. Sci. USA* 99:5261-5266.

33. Dettler, J.C., J.M. Jett, S.M. Lucas, E. Dalin, A.R. Arellano, M. Wang, J.R. Nelson, J. Chapman, et al. 2002. Isothermal strand-displacement amplification applications for high-throughput genomics. *Genomics* 80:691-698.
34. Polidoros, A.N., K. Pasentsis, and A.S. Tsiftaris. 2006. Rolling circle amplification-RACE: a method for simultaneous isolation of 5' and 3' cDNA ends from amplified cDNA templates. *BioTechniques* 41:35-42.
35. Nunez, A.N., M.F. Kavlick, J.M. Robertson, and B. Budowle. 2008. Application of circular ligase to provide template for rolling circle amplification of low amounts of fragmented DNA, p. 1-7. Nineteenth International Symposium on Human Identification 2008, Promega Corporation, Madison. (<http://www.promega.com/geneticidproc/ussymp19proc/oralpresentations/Nunez.pdf>).
36. Hellmann, A., U. Rohleder, H. Schmitter, and M. Wittig. 2001. STR typing of human telogen hairs: a new approach. *Int. J. Legal Med.* 114:269-273.
37. Mulero, J.J., C.W. Chang, R.E. Lagacé, D.Y. Wang, J.L. Bas, T.P. McMahon, and L.K. Hennessy. 2008. Development and validation of the AmpFISTR MiniFiler PCR Amplification Kit: a MiniSTR multiplex for the analysis of degraded and/or PCR inhibited DNA. *J. Forensic Sci.* 53:838-852.
38. Gill, P., L. Fereday, N. Morling, and P.M. Schneider. 2006. The evolution of DNA databases—recommendations for new European STR loci. *Forensic Sci. Int.* 156:242-244.
39. Ecker, D.J., R. Sampath, C. Massire, L.B. Blyn, T.A. Hall, M.W. Eshoo, and S.A. Hofstadler. 2008. Ibis T5000: a universal biosensor approach for microbiology. *Nat. Rev. Microbiol.* 6:553-558.
40. Hall, T.A., B. Budowle, Y. Jiang, L. Blyn, M. Eshoo, K.A. Sannes-Lowery, R. Sampath, J.J. Drader, et al. 2005. Base composition analysis of human mitochondrial DNA using electrospray ionization mass spectrometry: a novel tool for the identification and differentiation of humans. *Anal. Biochem.* 344:53-69.
41. Honisch, C., Y. Chen, C. Mortimer, C. Arnold, O. Schmidt, D. van den Boom, C.R. Cantor, H.N. Shah, and S.E. Gharbia. 2007. Automated comparative sequence analysis by base-specific cleavage and mass spectrometry for nucleic acid-based microbial typing. *Proc. Natl. Acad. Sci. USA* 104:10649-10654.
42. Köster, H., K. Tang, D.J. Fu, A. Braun, D. van den Boom, C.L. Smith, R.J. Cotter, and C.R. Cantor. 1996. A strategy for rapid and efficient DNA sequencing by mass spectrometry. *Nat. Biotechnol.* 14:1123-1128.
43. Sampath, R., T.A. Hall, C. Massire, F. Li, L.B. Blyn, M.W. Eshoo, S.A. Hofstadler, and D.J. Ecker. 2007. Rapid identification of emerging infectious agents using PCR and electrospray ionization mass spectrometry. *Ann. N. Y. Acad. Sci.* 1102:109-120.
44. Lunetta, P., A. Levo, A. Mannikko, A. Pentilla, and A. Sajantila. 2002. Death in a bathtub revisited with molecular genetics: a victim with suicidal traits and a LQTS gene mutation. *Forensic Sci. Int.* 130:122-124.
45. Sistonen, J., S. Fuselli, J.U. Palo, N. Chauhan, H. Padh, and A. Sajantila. 2009. Pharmacogenetic variation at CYP2C9, CYP2C19, and CYP2D6 at global and microgeographic scales. *Pharmacogenet. Genomics* 19:170-179.
46. Dixon, L.A., C.M. Murray, E.J. Archer, A.E. Dobbins, P. Koumi, and P. Gill. 2005. Validation of a 21-locus autosomal SNP multiplex for forensic identification purposes. *Forensic Sci. Int.* 154:62-77.
47. Kidd, K.K., A.J. Pakstis, W.C. Speed, E.L. Grigorenko, S.L. Kajuna, N.J. Karoma, S. Kungulilo, J.J. Kim, et al. 2006. Developing a SNP panel for forensic identification of individuals. *Forensic Sci. Int.* 164:20-32.
48. Pakstis, A.J., W.C. Speed, J.R. Kidd, and K.K. Kidd. 2007. Candidate SNPs for a universal individual identification panel. *Hum. Genet.* 121:305-317.
49. Sanchez, J.J., C. Phillips, C. Børsting, K. Balogh, M. Bogus, M. Fondevila, C.D. Harrison, E. Musgrave-Brown, et al. 2006. A multiplex assay with 52 single nucleotide polymorphisms for human identification. *Electrophoresis* 27:1713-1724.
50. Daly, M.J., J.D. Rioux, S.F. Schaffner, T.J. Hudson, and E.S. Lander. 2001. High-resolution haplotype structure in the human genome. *Nat. Genet.* 29:229-232.
51. Gabriel, S.B., S.F. Schaffner, H. Nguyen, J.M. Moore, J. Roy, B. Blumenstiel, J. Higgins, M. DeFelice, et al. 2002. The structure of haplotype blocks in the human genome. *Science* 296:2225-2229.
52. International HapMap Consortium. 2003. The International HapMap Project. *Nature* 426:789-796.
53. McVean, G.A., S.R. Myers, S. Hunt, P. Deloukas, D.R. Bentley, and P. Donnelly. 2004. The fine-scale structure of recombination rate variation in the human genome. *Science* 304:581-584.
54. Patil, N., A.J. Berno, D.A. Hinds, W.A. Barrett, J.M. Doshi, C.R. Hacker, C.R. Kautzer, D.H. Lee, et al. 2001. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science* 294:1719-1723.
55. Ge, J., H. Xi, B. Budowle, and R. Chakraborty. 2008. Haplotype block: a new type of forensic DNA markers. Poster presented at the Eighteenth Meeting of the International Association of Forensic Sciences, New Orleans, LA.
56. Baudhuin, L.M., L.J. Langman, and D.J. O'Kane. 2007. Translation of pharmacogenetics into clinically relevant testing modalities. *Clin. Pharmacol. Ther.* 82:373-376.
57. Smith, G., M.J. Stubbins, L.W. Harries, and C.R. Wolf. 1998. Molecular genetics of the human cytochrome P450 monooxygenase superfamily. *Xenobiotica* 28:1129-1165.

GO FISHING!



SEQUENCE CAPTURE

HybSelect™

CATCH YOUR REGION OF INTEREST FOR YOUR NEXT GENERATION SEQUENCING!

- Excellent SNP detection
- Deep sequence coverage
- Automation minimizes hands-on time
- Simple and streamlined workflow

HYBSELECT NOW AVAILABLE

- as a full-service from febit Analytical Services
- as the newest application for the Geniom RT Analyzer



READ, WRITE, UNDERSTAND THE CODE OF LIFE

Europe: phone +49 6221 6510-300 • info@febit.eu
Americas: phone +1 781 391 4360 • info@febit.com

www.febit.com

58. Zhou, S.F., Y.M. Di, E. Chan, Y.M. Du, V.D. Chow, C.C. Xue, X. Lai, J.C. Wang, et al. 2008. Clinical pharmacogenetics and potential application in personalized medicine. *Curr. Drug Metab.* 9:738-784.
59. Rodríguez-Calvo, M.S., M. Brion, C. Allegue, L. Concheiro, and A. Carracedo. 2008. Molecular genetics of sudden cardiac death. *Forensic Sci. Int.* 182:1-12.
60. Juusola, J. and J. Ballantyne. 2005. Multiplex mRNA profiling for the identification of body fluids. *Forensic Sci. Int.* 152:1-12.
61. Zubakov, D., M. Kokshoorn, A. Kloosterman, and M. Kayser. 2009. New markers for old stains: stable mRNA markers for blood and saliva identification from up to 16-year-old stains. *Int. J. Legal Med.* 123:71-74.
62. Bieber, F.R., C.H. Brenner, and D. Lazer. 2006. Human genetics. Finding criminals through DNA of their relatives. *Science* 312:1315-1316.
63. Budowle, B., J.V. Planz, R. Chakraborty, T.F. Callaghan, and A.J. Eisenberg. 2006. Clarification of statistical issues related to the operation of CODIS, p. 1-20. Seventeenth International Symposium on Human Identification 2006, Promega Corporation, Madison. (<http://www.promega.com/geneticidproc/ussymp17proc/oralpresentations/budowle.pdf>).
64. Breeze, R.G., B. Budowle, and S. Schutzer. 2005. *Microbial Forensics*, p. 1-425. Academic Press, Amsterdam.
65. Budowle, B., S.E. Schutzer, A. Einseln, L.C. Kelley, A.C. Walsh, J.A.L. Smith, B.L. Marrone, J. Robertson, and J. Campos. 2003. Building microbial forensics as a response to bioterrorism. *Science* 301:1852-1853.
66. Morse, S.A. and B. Budowle. 2006. Microbial forensics-application to bioterrorism preparedness and response. *Infect. Dis. Clin. North Am.* 20:455-473.
67. Morse, S.A. and A.S. Khan. 2006. Epidemiologic investigation for public health, biodefense, and forensic microbiology, p. 157-171. *In* R.G. Breeze, B. Budowle, and S. Schutzer (Eds.) *Microbial Forensics*. Academic Press, Amsterdam.
68. Budowle, B., M.D. Johnson, C.M. Fraser, T.J. Leighton, R.S. Murch, and R. Chakraborty. 2005. Genetic analysis and attribution of microbial forensics evidence. *Crit. Rev. Microbiol.* 31:233-254.
69. Read, T.D., S.N. Peterson, N. Tourasse, L.W. Baillie, I.T. Paulsen, K.E. Nelson, H. Tettelin, D.E. Fouts, et al. 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. *Nature* 423:81-86.
70. Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463-5468.
71. Lipshutz, R.J., D. Morris, M. Chee, E. Hubbell, M.J. Kozal, N. Shah, N. Shen, R. Yang, and S.P. Fodor. 1995. Using oligonucleotide probe arrays to access genetic diversity. *BioTechniques* 19:442-447.
72. Cutler, D.J., M.E. Zwick, M.M. Carrasquillo, C.T. Yohn, K.P. Tobin, C. Kashuk, D.J. Mathews, N.A. Shah, et al. 2001. High-throughput variation detection and genotyping using microarrays. *Genome Res.* 11:1913-1925.
73. Drmanac, S., D. Kita, I. Labat, B. Hauser, C. Schmidt, J.D. Burczak, and R. Drmanac. 1998. Accurate sequencing by hybridization for DNA diagnostics and individual genomics. *Nat. Biotechnol.* 16:54-58.
74. Adessi, C., G. Matton, G. Ayala, G. Turcatti, J.J. Mermoud, P. Mayer, and E. Kawashima. 2000. Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res.* 28:E87.
75. Brenner, S., M. Johnson, J. Bridgham, G. Golda, D.H. Lloyd, D. Johnson, S. Luo, S. McCurdy, et al. 2000. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat. Biotechnol.* 18:630-634.
76. Brenner, S., S.R. Williams, E.H. Vermaas, T. Storck, K. Moon, C. McCollum, J.I. Mao, S. Luo, et al. 2000. In vitro cloning of complex mixtures of DNA on microbeads: physical separation of differentially expressed cDNAs. *Proc. Natl. Acad. Sci. USA* 97:1665-1670.
77. Holt, K.E., J. Parkhill, C.J. Mazzoni, P. Roumagnac, F.X. Weill, I. Goodhead, R. Rance, S. Baker, et al. 2008. High-throughput sequencing provides insights into genome variation and evolution in *Salmonella Typhi*. *Nat. Genet.* 40:987-993.
78. Margulies, M., M. Egholm, W.E. Altman, S. Attiya, J.S. Bader, L.A. Bemben, J. Berka, M.S. Braverman, et al. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376-380.
79. Oh, J.D., H. Kling-Bäckhed, M. Giannakis, J. Xu, R.S. Fulton, L.A. Fulton, H.S. Cordum, C. Wang, et al. 2006. The complete genome sequence of a chronic atrophic gastritis *Helicobacter pylori* strain: evolution during disease progression. *Proc. Natl. Acad. Sci. USA* 103:9999-10004.
80. Quail, M.A., I. Kozarewa, F. Smith, A. Scally, P.J. Stephens, R. Durbin, H. Swerdlow, and D.J. Turner. 2008. A large genome center's improvements to the Illumina sequencing system. *Nat. Methods* 5:1005-1010.
81. Van Tassel, C.P., T.P. Smith, L.K. Matukumalli, J.F. Taylor, R.D. Schnabel, C.T. Lawley, C.D. Haudenschild, S.S. Moore, et al. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat. Methods* 5:247-252.
82. Wheeler, D.A., M. Srinivasan, M. Egholm, Y. Shen, L. Chen, A. McGuire, W. He, Y.J. Chen, et al. 2008. The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452:872-876.
83. Wicker, T., E. Schlagenhauf, A. Graner, T.J. Close, B. Keller, and N. Stein. 2006. 454 sequencing put to the test using the complex genome of barley. *BMC Genomics* 7:275.
84. Braslavsky, I., B. Hebert, E. Kartalov, and S.R. Quake. 2003. Sequence information can be obtained from single DNA molecules. *Proc. Natl. Acad. Sci. USA* 100:3960-3964.
85. Anslinger, K., B. Bayer, B. Rolf, W. Keil, and W. Eisenmenger. 2005. Application of the BioRobot EZ1 in a forensic laboratory. *Leg. Med.* 7:164-168.
86. Fréreau, C.J., C.M. Lett, J. Elliott, C. Yensen, and R.M. Fournery. 2008. Automated processing of forensic casework samples using robotic workstations equipped with nondisposable tips: contamination prevention. *J. Forensic Sci.* 53:632-651.
87. Greenspoon, S.A., K.L. Sykes, J.D. Ban, A. Pollard, M. Baisden, M. Farr, N. Graham, B.L. Collins, et al. 2006. Automated PCR setup for forensic casework samples using the Normalization Wizard and PCR Setup robotic methods. *Forensic Sci. Int.* 164:240-248.
88. Greenspoon, S.A., J.D. Ban, K. Sykes, E.J. Ballard, S.S. Edler, M. Baisden, and B.L. Covington. 2004. Application of the BioMek 2000 Laboratory Automation Workstation and the DNA IQ System to the extraction of forensic casework samples. *J. Forensic Sci.* 49:29-39.
89. Hanselle, T., M. Otte, T. Schnibbe, E. Smythe, and F. Krieg-Schneider. 2003. Isolation of genomic DNA from buccal swabs for forensic analysis, using fully automated silica-membrane purification technology. *Leg. Med.* 5:S145-S149.
90. Parson, W. and M. Steinlechner. 2001. Efficient DNA database laboratory strategy for high throughput STR typing of reference samples. *Forensic Sci. Int.* 122:1-6.
91. Tereba, A., L. Flanagan, P. Mandrekar, and R. Olson. 2004. A new, rapid method to separate sperm and epithelial cells. *Profiles in DNA* 7:8-10.
92. Mandrekar, P., B.E. Krenke, and A. Tereba. 2001. DNA IQ: the intelligent way to purify DNA. *Profiles in DNA* 4:16.
93. Kadosh, K., B.E. Kozłowski, L.A. Biega, and B.W. Ducean. 2004. Validation study of the TrueAllele automated data review system. *J. Forensic Sci.* 49:1-8.
94. Roby, R.K. and A.D. Cristen. 2007. Validating expert systems: Examples with the FSS-i3 expert systems software. *Profiles in DNA* 10:13-15.
95. Pendleton, J. and T.W. Wang. Gilbert, k., and Lucas, C. 2004. Probability distribution of allele bands for multi-person STR mixtures. *Proceedings from the Fifteenth International Symposium on Human Identification 2004*, Promega Corporation, Madison. (<http://www.promega.com/geneticidproc/ussymp15proc/posterpresentations/12Wang.pdf>).
96. Horsman, K.M., J.M. Bienvenue, K.R. Blasier, and J.P. Landers. 2007. Forensic DNA analysis on microfluidic devices: a review. *J. Forensic Sci.* 52:784-799.
97. Easley, C.J., J.M. Karlinsey, J.M. Bienvenue, L.A. Legendre, M.G. Roper, S.H. Feldman, M.A. Hughes, E.L. Hewlett, et al. 2006. A fully integrated microfluidic genetic analysis system with sample-in-answer-out capability. *Proc. Natl. Acad. Sci. USA* 103:19272-19277.
98. Yeung, S.H., P. Liu, N. Del Bueno, S.A. Greenspoon, and R.A. Mathies. 2009. Integrated sample cleanup-capillary electrophoresis microchip for high-performance short tandem repeat genetic analysis. *Anal. Chem.* 81:210-217.
99. Norris, J.V., K. Manning, S.J. Linke, J.P. Ferrance, and J.P. Landers. 2007. Expedited, chemically enhanced sperm cell recovery from cotton swabs for rape kit analysis. *J. Forensic Sci.* 52:800-805.
100. Wen, J., C. Guillo, J.P. Ferrance, and J.P. Landers. 2007. Microfluidic-based DNA purification in a two-stage, dual-phase microchip containing a reversed-phase and a photopolymerized monolith. *Anal. Chem.* 79:6135-6142.
101. Wen, J., L.A. Legendre, J.M. Bienvenue, and J.P. Landers. 2008. Purification of nucleic acids in microfluidic devices. *Anal. Chem.* 80:6472-6479.

Received 4 February 2009; accepted 3 March 2009.

Address correspondence to Bruce Budowle, Ph.D., Director of the Institute of Investigative Genetics Department of Forensic and Investigative Genetics, 3500 Camp Bowie Boulevard, Ft. Worth, TX, 76107. USA. email: bbudowle@hsc.unt.edu