



## Forensic DNA Typing and Prospects for Biometrics



Dr. John M. Butler  
NIST Biotechnology Division


May 12, 2004

### Examples of DNA in the News

- Saddam Hussein Identification
- Source of Cow with "Mad cow" Disease
- Scot Peterson Murder Trial
- Identification of WTC Victims
- "Thomas Jefferson fathered slave's children"




December 14, 2003

**"We got him!"**




DNA Profile



**Saddam Hussein's capture verified with DNA testing**

Source: www.cnn.com; The Scientist Dec 19, 2003

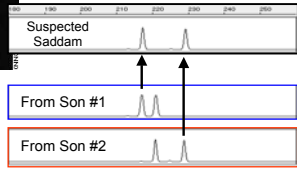
Captured December 13, 2003



**"A Forensic Paternity Test"**

*Y-chromosome tests were also used to verify male lineage...*

Portion of DNA Profile




Suspected Saddam

From Son #1

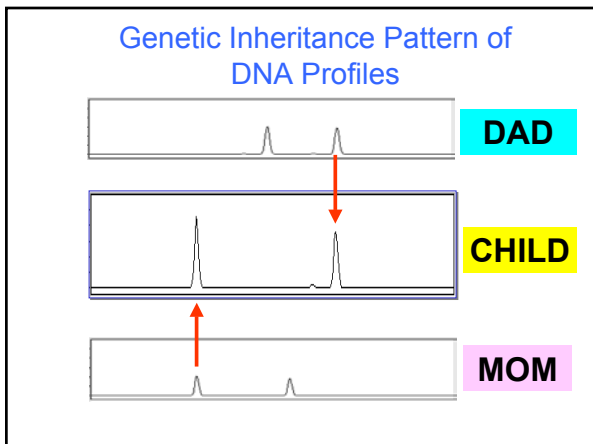
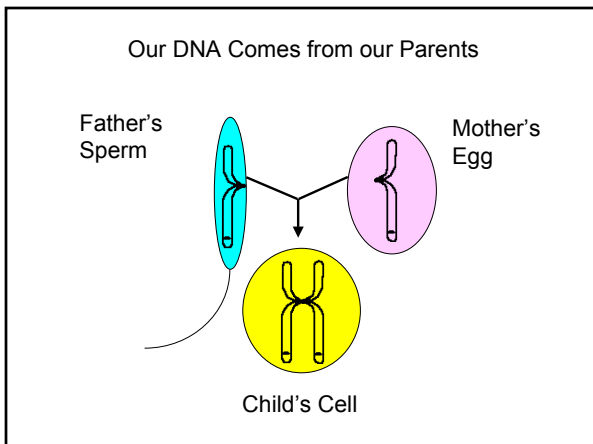
From Son #2

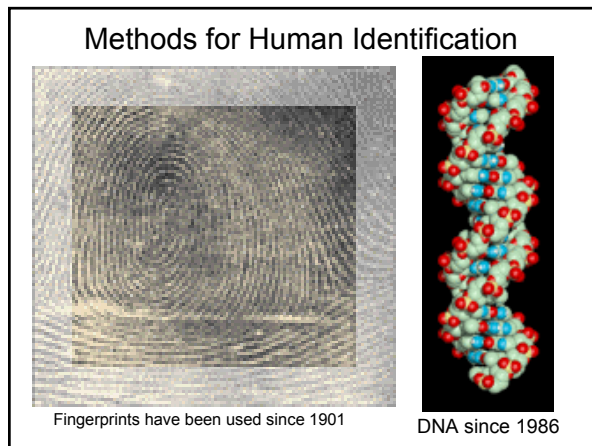
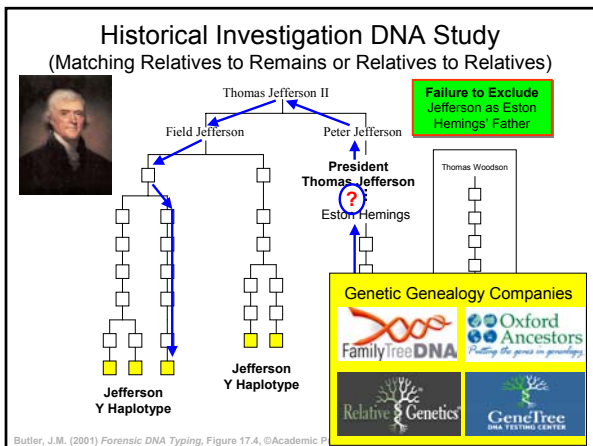
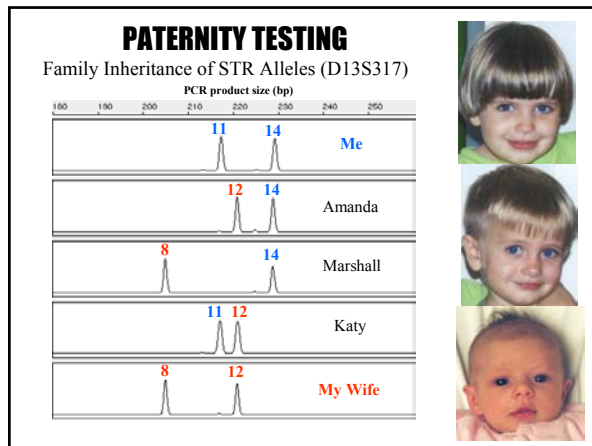
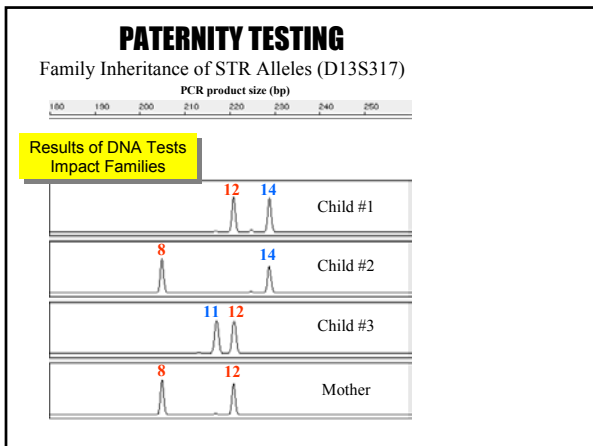
Killed July 22, 2003



**DNA Used as the Ultimate Biometric for Confirming Saddam's Identity**

Source: www.cnn.com Uday and Qusay Hussein





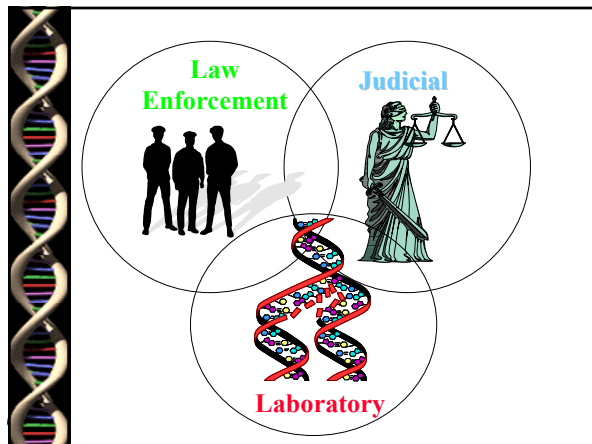
## Forensic DNA Testing

The genome of each individual is unique (with the exception of identical twins)

Probe subsets of genetic variation in order to differentiate between individuals

DNA typing must be done efficiently and reproducibly (information must hold up in court)

Typically, we are not looking at genes – little/no information about race, predisposal to disease, or phenotypical information (eye color, height, hair color) is obtained



### Applications for Human Identity Testing

- Forensic cases - **matching suspect with evidence**
- Paternity testing - **identifying father**
- Historical investigations
- Missing persons investigations
- Mass disasters - **putting pieces back together**
- Military DNA "dog tag"
- Convicted felon DNA databases

As DNA analysis has shown its usefulness, the number of samples gathered for testing purposes has gone up dramatically...

### Armed Forces DNA Repository



>4.5 million blood cards on file from members of U.S. military

Are being used to identify remains in case of combat casualties in Iraq



Located in Gaithersburg, Maryland



### Armed Forces DNA Identification Laboratory (Rockville, MD)

Identification of Skeletal Remains from Previous Military Conflicts

Identification of Michael J. Blassie - Vietnam Tomb of the Unknown

- >2100 Unaccounted for from Southeast Asia
- >8000 Unaccounted for from Korea
- >75,000 Unaccounted for from WWII



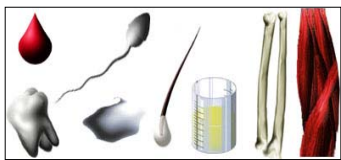
### Roles of Biological Evidence in Criminal Investigation

- Identify a person
- Exclude a suspect
- Link suspect, victim and crime scene
- Link weapon to victim
- Link witness to scene
- Prove or disprove an alibi
- Reconstruct the scene
- Provide investigative leads

**Innocence Project**  
Post-conviction DNA testing is used to exonerate falsely incarcerated individuals (143 so far)

### Sources of Biological Evidence

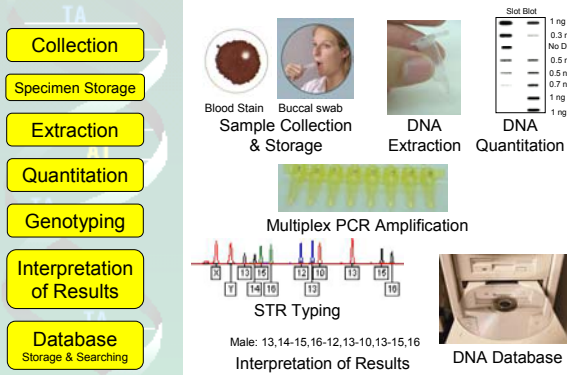
- Blood
- Semen
- Saliva
- Urine
- Hair
- Teeth
- Bone
- Tissue

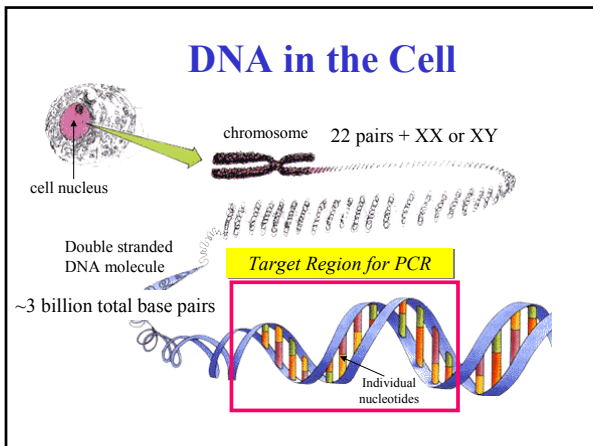


#### Blood stain

Only a very small amount of blood is needed to obtain a DNA profile

### Steps in DNA Analysis





### What Type of Genetic Variation?

- Length Variation  
short tandem repeats (**STRs**)  
CTAGTCGT(GATA)(GATA)(GATA)GCGATCGT

- Sequence Variation  
single nucleotide polymorphisms (**SNPs**)  
insertions/deletions  
GCTAGTCGATGCTC(G/A)GCGTATGCTGTAGC

### Basic Concepts

**PCR polymerase chain reaction** – method of amplifying a specific region of the genome – go from 1 to over a billion copies in about 2 hours

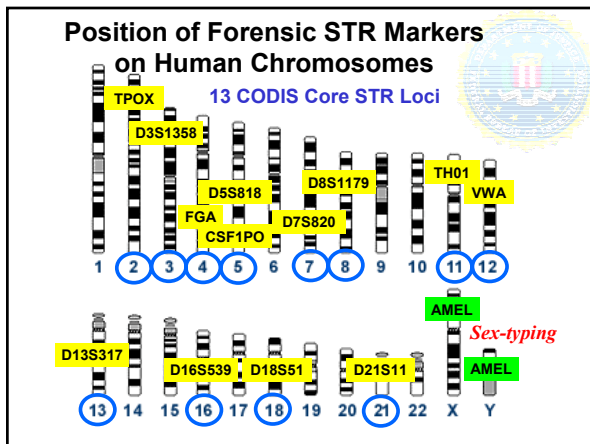
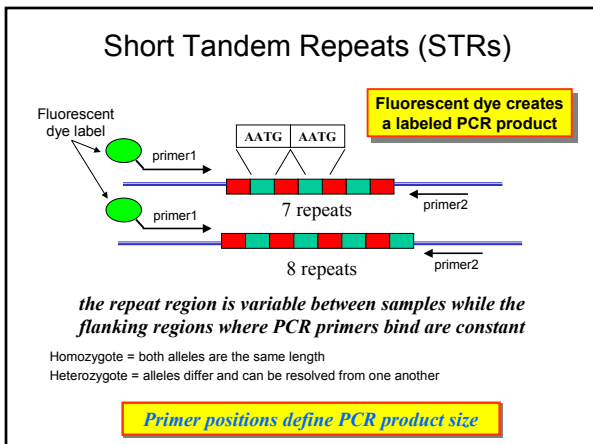
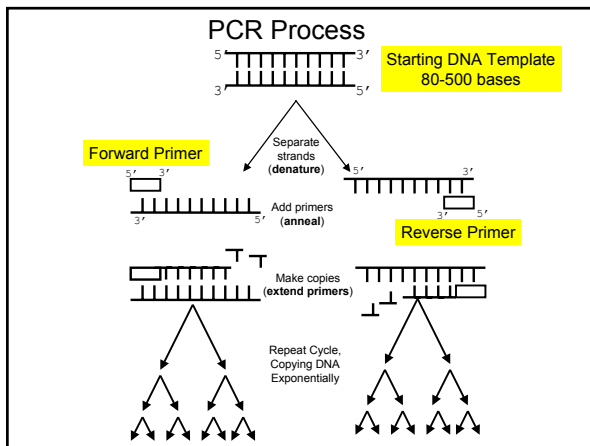
**Locus** region of the genome being examined

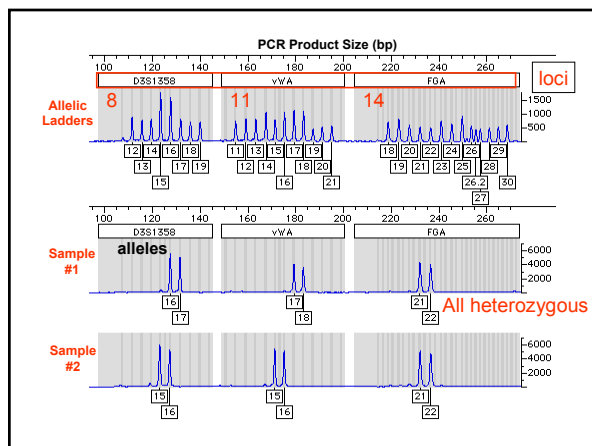
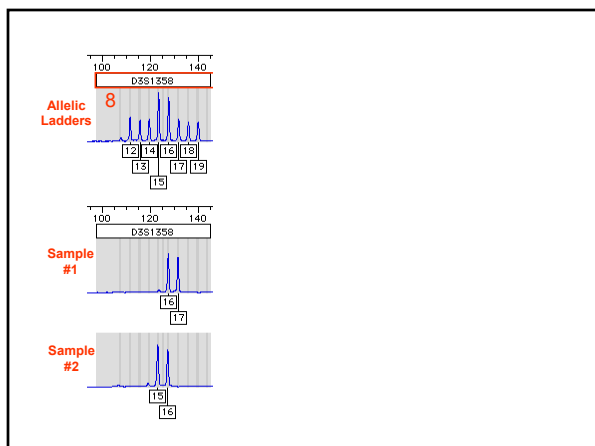
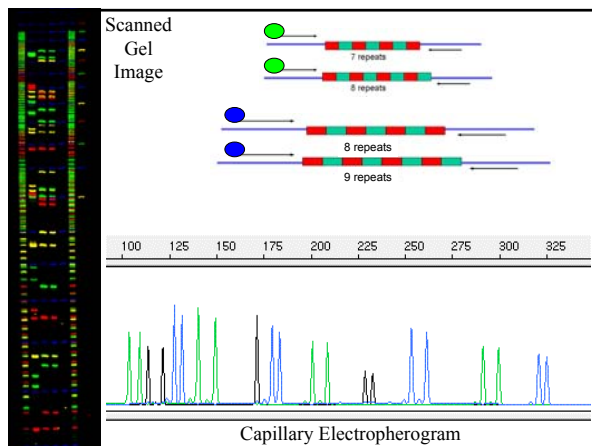
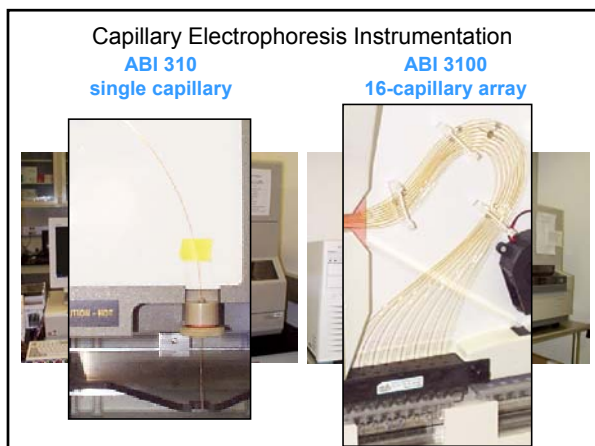
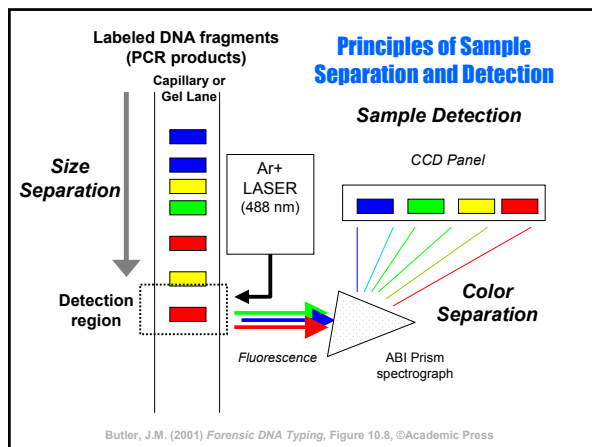
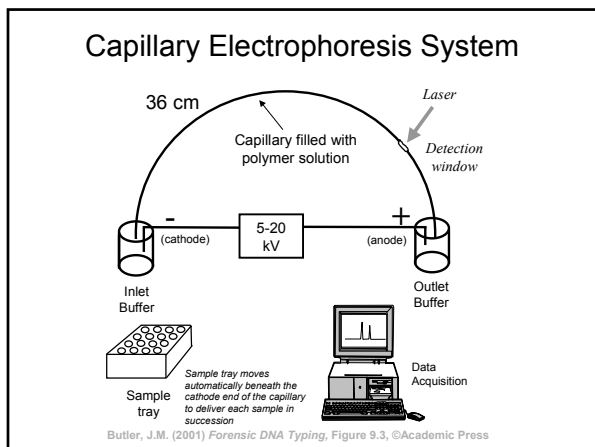
**Allele** the state of the genetic variation being examined  
(STRs = number of repeat units)  
(SNPs = base sequence at the site)

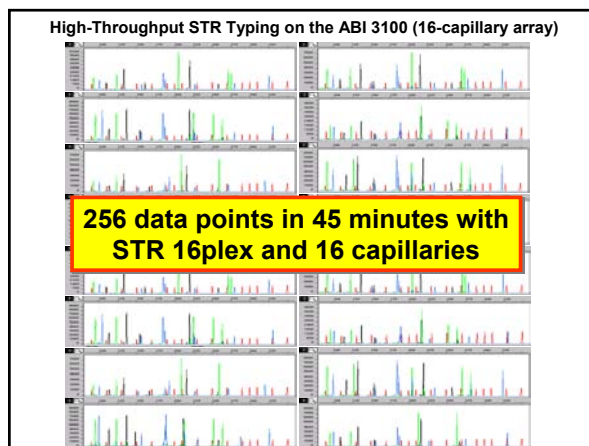
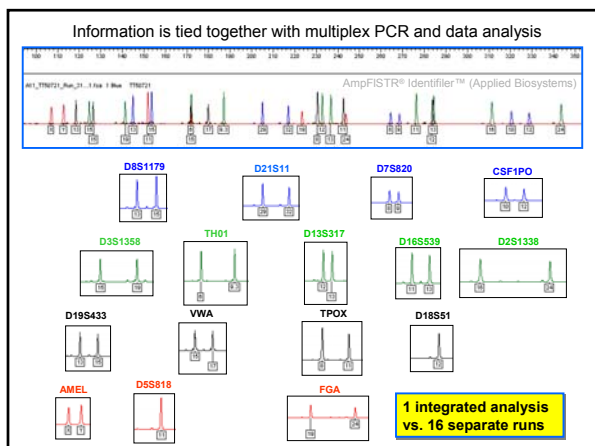
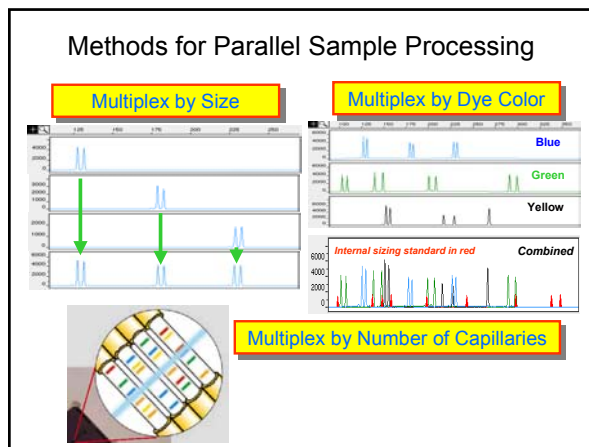
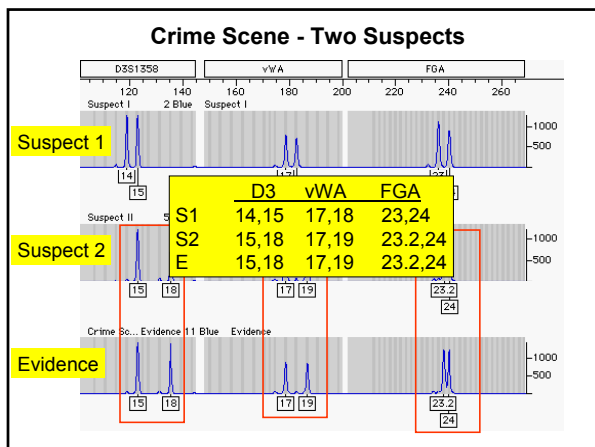
Chromosomes are paired so...

**Homozygous** – Alleles are identical on each chromosome

**Heterozygous** - Alleles differ on each on each chromosome







## DNA Statistics

For heterozygous loci

$$P = 2pq$$

P = probability; p and q are frequencies of allele in a given population

Example: For the locus D3S1358 and individual is 16,17 with frequencies of 0.2315 and 0.2118 respectively

$$P = 2(0.2315)(0.2118) = 0.0981 \text{ or } 1 \text{ in } 10.2$$

For independent loci, the genotype frequencies can be combined through multiplication...

Profile Probability =  $(P_1)(P_2)...(P_n)$   
 = 1 in a very large number...


### DNA Profile Frequency with all 13 CODIS STR loci

What would be entered into a DNA database for searching:  
 16,17-17,18-21,22-12,14-28,30-14,16-12,13-11,14-9,9-11,13-6,6-8,8-10,10

Locus	allele	value	allele	value	frequency, 1 in
D3S1358	16	0.2315	17	0.2118	10.20
VWA	17	0.2628	18	0.2219	8.57
FGA	21	0.1735	22	0.1888	15.26
D8S1179	12	0.1454	14	0.2015	17.07
D21S11	28	0.1658	30	0.2321	12.99
D18S51	14	0.1735	16	0.1071	26.91
D5S818	12	0.3539	13	0.1462	9.66
D13S317	11	0.3189	14	0.0357	43.92
D7S820	9	0.1478			43.28
D16S539	11	0.2723	13	0.1634	11.24
TH01	6	0.2266			18.83
TPOX	8	0.5443			3.35
CSF1PO	10	0.2537			15.09

The Random Match Probability for this profile in the FBI Caucasian population is 1 in 1.56 quadrillion ( $10^{15}$ )





### CODIS DNA Database

**Combined DNA Index System**

Used for linking serial crimes and unsolved cases with repeat offenders


Convicted offender and forensic case samples

Launched October 1998 and links all 50 states

Requires 13 core STR markers

Current backlog of >750,000 samples (millions of dollars are spent each year to reduce backlog)

All 50 states now require convicted offenders to submit a sample for DNA testing purposes



**16,160 Investigations Aided as of March 2004**

As of March 2004 the profile composition of the National DNA Index System (NDIS) is as follows:

Total number of profiles: 1,719,551  
 Total Forensic profiles: 78,475  
 Total Convicted Offender Profiles: 1,641,076


<http://www.fbi.gov/hq/lab/codis/clickmap.htm>

1,641,076 offenders  
 78,475 forensic  
 102 missing persons  
 175 CODIS labs in 50 states, FBI, US Army Crime Lab & Puerto Rico  
 16,160 investigations aided

**Results from Virginia**

Statistical Information	Total
Offender Profiles	185,387
Forensic Samples	3,068
Number of CODIS Labs	4
NDIS Participating Labs	4
Investigations Aided	1,773

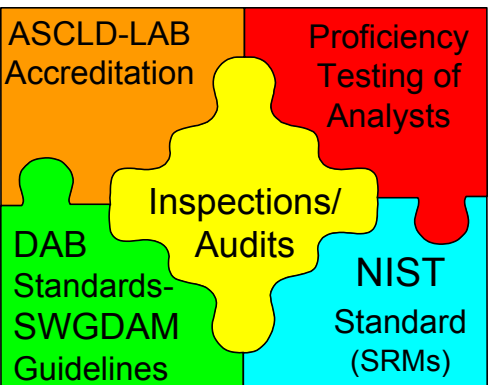
<http://www.fbi.gov/hq/lab/codis/va.htm>




**Our Human Identification Project Team**

- Prepares Standard Reference Materials (SRMs)
  - SRM 2391b PCR-based DNA Profiling Standard
  - SRM 2395 Human Y-Chromosome DNA Profiling Standard
- Creates databases with useful information
  - STRBase (<http://www.cstl.nist.gov/biotech/strbase>)
- Evaluates and develops new technologies
- Conducts interlaboratory testing
- Performs quality control testing for labs & companies

### Ensuring Accurate Forensic DNA Results




ASCLD-LAB Accreditation

Proficiency Testing of Analysts

Inspections/Audits

DAB Standards-SWGDAM Guidelines

NIST Standard (SRMs)



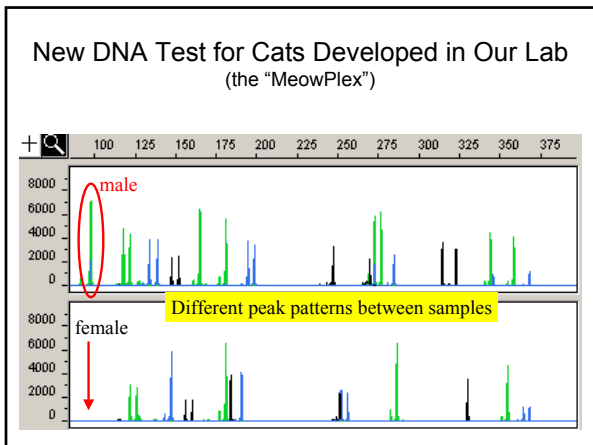
### STRBase Short Tandem Repeat DNA Internet Database

working with industry to develop and apply technology, measurements and standards

General Information	Forensic Interest Data	Supplemental Info
<ul style="list-style-type: none"> <li>Intro to STRs (downloadable PowerPoint)</li> <li>STR Fact Sheets</li> <li>Sequence Information</li> <li>Multiplex STR Kits</li> <li>Variant Allele Reports</li> </ul>	<ul style="list-style-type: none"> <li>FBI CODIS Core Loci</li> <li>DAB Standards</li> <li>NIST SRM 2391</li> <li>Published PCR Primers</li> <li>Y-Chromosome STRs</li> <li>Population Data</li> <li>Validation Studies</li> </ul>	<ul style="list-style-type: none"> <li>Reference List</li> <li>Technology Review</li> <li>Addresses for Scientists</li> <li>Links to Other Web Sites</li> </ul>

Standardized information formats

<http://www.cstl.nist.gov/biotech/strbase>

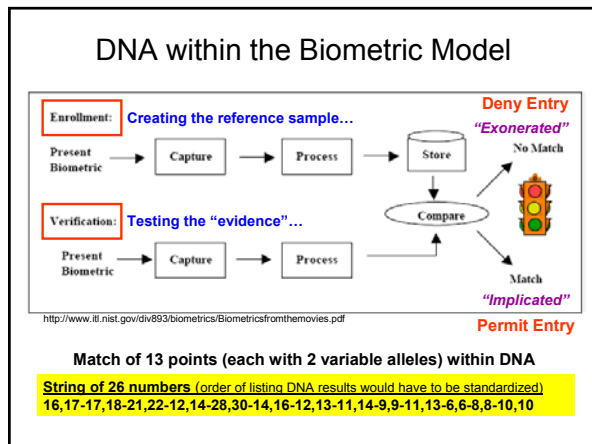


### Telling Cats Apart...

"Kitty Biometrics"

### Disclaimer

- Our laboratory performs research and development in human identification techniques involving DNA testing
- I am not an "expert" in biometrics and therefore do not understand all of the issues involved
- The following ideas are my own and in no way represent the official position of the National Institute of Standards and Technology (NIST)



### Issues with DNA as a Biometric Tool

**Individualization of everyone except identical twins**

- **Invasive collection of test sample** (swab of cheek cells or finger prick for blood)
- **Concern over genetic privacy** if sample is stored (would or could sample ever be used for other purposes?)
- **Expense for testing** (>\$20 per sample; forensic testing is >\$500)
- **Time required for testing** (minimum 4-5 hours with forensic markers)
- Database search/storage of DNA profile (who controls access?; would it ever be used for criminal investigations?)
- Development of reference databases for terrorists, etc. (how would a match be used?)
- Chance for falsifying your sample (e.g., movie GATTACA)

### Collection of Test Sample

The Buccal DNA Collector  
Collection of the DNA sample

The Bode Technology Group, Inc.  
A OnePoint Company

Used routinely in Virginia for felon arrestee DNA collection with 99.5% success (only 23 failures in first 6 months of use)

- Buccal (cheek cell) swab is not as invasive as a finger prick to draw blood
- DNA can be reliably extracted off of the paper or Q-tip swab



### Genetic Privacy Concerns

- The DNA profile itself is neutral (and uninformative)—just a string of numbers like a Social Security Number
- DNA markers used in forensics were selected to be neutral and are located away from or between genes rather than being part of gene products and therefore are not generally thought to be associated with any genetic disease
- Concern is really with the DNA sample collected—would it be retained and used for any other type of testing? (e.g., Armed Forces DNA Repository can only be used for identifying combat casualties)

### Time Required for Testing

*Now typically a minimum of 4-5 hours*

<b>Collection</b>	
<b>Extraction</b>	Could be <5 minutes
<b>Quantitation</b>	Not necessary if samples are uniform in amount
<b>Amplification</b>	Rapid thermal cycling to-date <u>done with singleplexes</u> ; typically 2-3 hours <b>Biggest problem is length of time for PCR (with multiplex amplification)</b>
<b>Genotyping</b>	DNA separations (STR analysis) of <5 minutes have been demonstrated; typically ~30 minutes
<b>Interpretation of Results</b>	Currently performed manually in most labs; expert systems are under development to enable rapid interpretation
<b>Database Storage &amp; Searching</b>	Search could be similar to fingerprint search in terms of speed

Comparison a DNA profile to a reference or database

Male: 13,14-15,16-12,13-10,13-15,16-.....

### Possible Ways DNA Could Be Used Now

As part of check before issuing visa for foreign visitors wanting to visit U.S.

- DNA samples could be shipped from each embassy to U.S.-based lab for testing
- Visa would not be issued until DNA profile was generated and searched against a national/international database
- DNA profile would be included in biometric passport for future confirmation of identity as needed

Confirming identity upon entry into the U.S. while the plane is in the air

- Would require DNA labs operating in foreign airports
- 200+ DNA profiles would be generated in ~8 hours (before a plane coming from Europe landed)



**In FY2003, 7,300,667 U.S. passports were issued**

[http://travel.state.gov/passport\\_statistics.html](http://travel.state.gov/passport_statistics.html)

### How close are we to GATTACA?



Gattaca Corp. is an aerospace firm in the future. During this time society analyzes your DNA and determines where you belong in life.

Ethan Hawke's character was born with a congenital heart condition which would cast him out of getting a chance to travel in space. So in turn he assumes the identity of an athlete named Jerome who has genes that would allow him to achieve his dream of space travel.

Hawke's character Vincent constantly passes genetic tests by diligently using samples of Jerome's hair, skin, blood and urine.

**Entry to secure locations controlled by rapid genetic profiling**

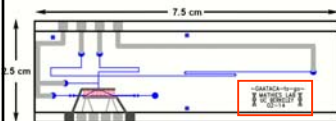
**Ability to predict future health risks based on genetic testing**

<http://www.imdb.com/title/tt0119177/>

1997 movie with a futuristic story of a genetically imperfect man and his seemingly unobtainable goal to travel in space.

**NOT THERE YET!**

### Portable Genetic Diagnostics Device

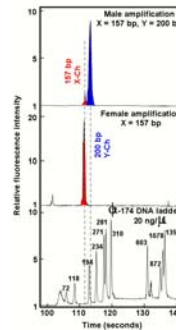


- Integrated system includes:**
- glass CE microchannels
  - PCR chamber
  - heater
  - temperature sensor
  - microfabricated electrodes
  - microfabricated valves



From Richard Mathies presentation at 14<sup>th</sup> International Symposium on Human Identification, Oct 2003

### Sex Determination PCR from human buccal cells



Multiplex PCR *directly* from human buccal cells without DNA extraction

Hotstart PCR protocol activates polymerase and lyses cells

20 cycles PCR after hotstart

Short "clean-up" injection removes debris, followed by diagnostic injection

Female genotype results in a single 157-bp product

Male genotype results in a 157-bp product (X) and a 200-bp product (Y)

Lagally et al., Lab-on-a-Chip, 1, 102 (2001)

**15 minutes for PCR amplification and detection**

From Richard Mathies presentation at 14<sup>th</sup> International Symposium on Human Identification, Oct 2003

### Virginia DNA Testing of Felon Arrestees

As of January 1, 2003, any individual arrested for a violent felony crime (Code of Virginia § 19.2-310.2:1) must provide a buccal sample for DNA analysis, with the resultant profile incorporated into the Virginia DNA Data Bank (Code of Virginia § 19.2-310.5).

#### Since January 2003

- Buccal swab collected upon arrest
- DNA sample processed within 72 hours
- DNA profile searched against state database (national database does not currently allow searches for individuals prior to conviction)
- If a match results, then arrestee is detained and later prosecuted
- From Jan 2003 – Dec 2003, VA processed 7,836 arrestee samples (not all analyzed) and scored 63 hits against their state database (*Profiles in DNA*, 2004, 7(1):3-5)



### If you want to know more...

- *Forensic DNA Typing: Biology and Technology behind STR Markers*
- NIST website: <http://www.cstl.nist.gov/biotech/strbase>
- John Butler email: [john.butler@nist.gov](mailto:john.butler@nist.gov)

