DNA as a Biometric

Fundamentals of Identity Science

NSF Workshop on Fundamental Challenges for Trustworthy Biometrics

November 8, 2010
Dr. Peter M. Vallone
Biochemical Science Division
National Institute of Standards and Technology
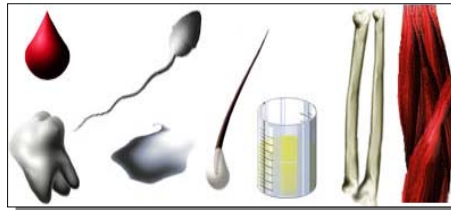
Outline

- Basics of DNA Typing

- DNA as a Biometric

General Characteristics of Genomic DNA

- Each individual has a unique DNA profile
  - with exception of monozygotic siblings

- Each person's DNA is the same in every cell
  - DNA from skin cells will match DNA from blood cells

- An individual's DNA profile remains the same throughout life

- Half of your DNA comes from your mother and half from your father
  - implications for determining kinship
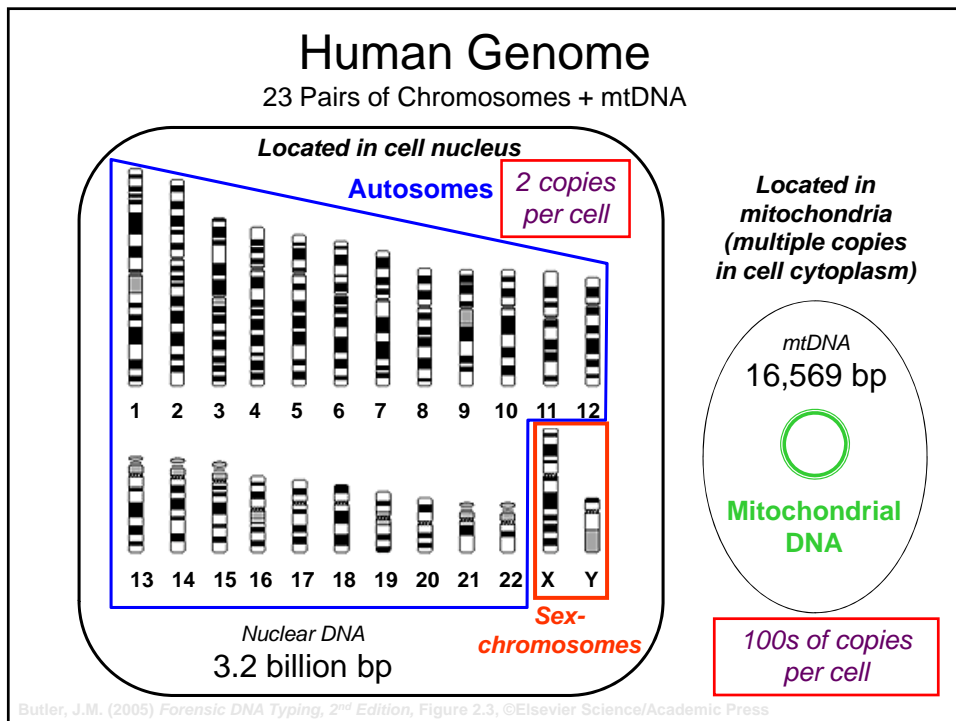

Sources of Biological Evidence

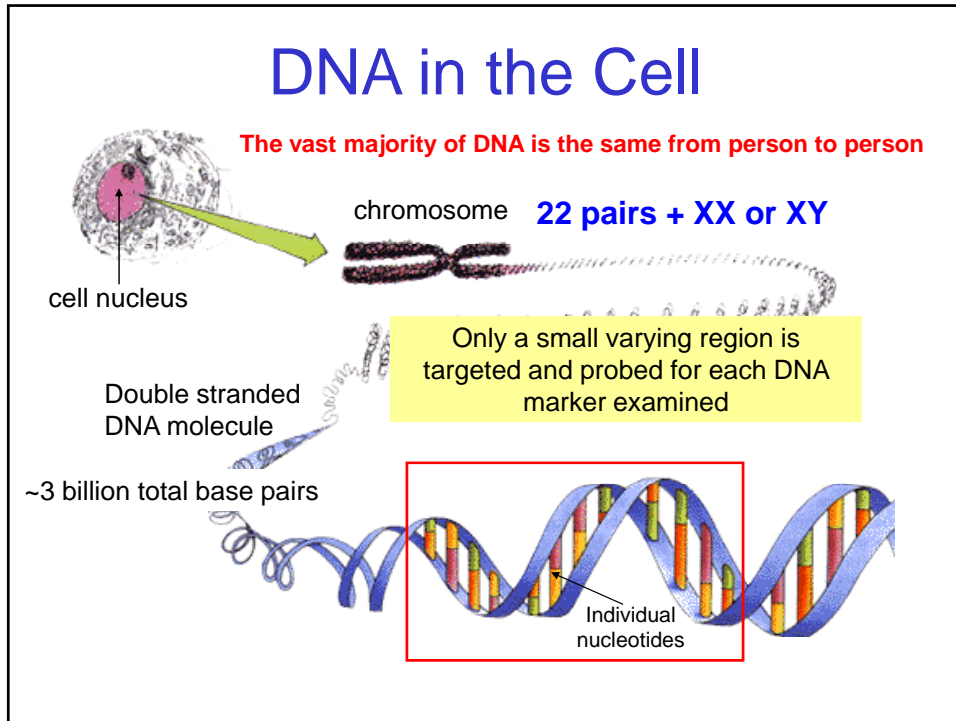- **Blood**
- **Semen**
- **Saliva**
- **Urine**
- **Hair**
- **Teeth**
- **Bone**
- **Tissue**

Blood Sample

Only a very small amount of blood is needed to DNA

best results with >100 cells, but DNA profiles can be recovered from fewer cells

DNA in the Cell

The vast majority of DNA is the same from person to person

chromosome   **22 pairs + XX or XY**

cell nucleus

Only a small varying region is targeted and probed for each DNA marker examined

Double stranded
DNA molecule

~3 billion total base pairs

Individual nucleotides

Human Genome

23 Pairs of Chromosomes + mtDNA

Located in cell nucleus

Autosomes | *2 copies per cell*

1 2 3 4 5 6 7 8 9 10 11 12

13 14 15 16 17 18 19 20 21 22 X Y

Nuclear DNA
3.2 billion bp

Sex-chromosomes

Located in mitochondria (multiple copies in cell cytoplasm)

mtDNA
16,569 bp

Mitochondrial DNA

100s of copies per cell

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition,* Figure 2.3, ©Elsevier Science/Academic Press

What Type of Genetic Variation?

- Sequence Variation

single nucleotide polymorphisms (SNPs)

insertions/deletions

GCTAGTCGATGCTC**(G/A)**GCGTATGCTGTAGC
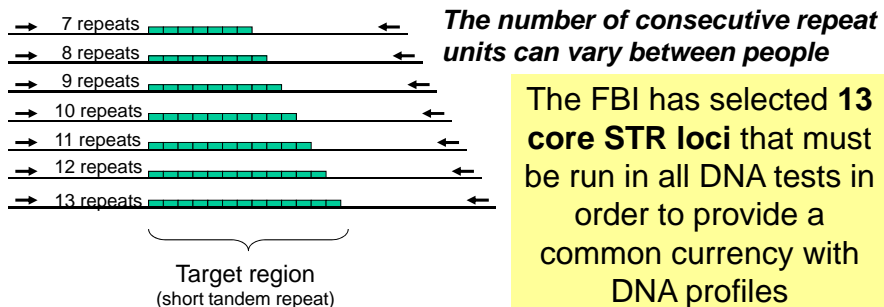
- Length Variation

short tandem repeats (STRs)
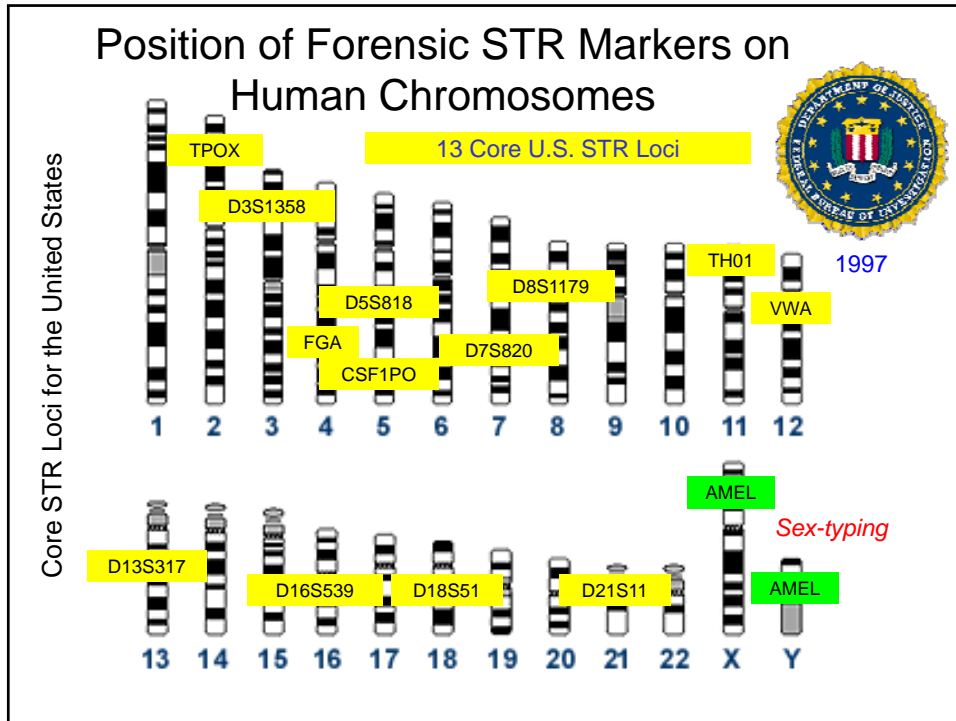
CTAGTCGT**(GATA)(GATA)(GATA)**GCGATCGT

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACA
GGTG**GATAGATAGATAGATAGATAGATAGATAGATAGATAGATA
GATAGATA**TCATTGAAAGACAAAACAGAGATGGATGATAGATACA
TGCTTACAGATGCACAC

= 12 GATA repeats ("12" is all that is reported)

→ 7 repeats
→ 8 repeats
→ 9 repeats
→ 10 repeats
→ 11 repeats
→ 12 repeats
→ 13 repeats

The number of consecutive repeat units can vary between people

The FBI has selected **13 core STR loci** that must be run in all DNA tests in order to provide a common currency with DNA profiles

Target region
(short tandem repeat)

Position of Forensic STR Markers on Human Chromosomes

Core STR Loci for the United States

13 Core U.S. STR Loci

TPOX

D3S1358

TH01 1997

D8S1179

VWA

D5S818

FGA

D7S820

CSF1PO

1 2 3 4 5 6 7 8 9 10 11 12

AMEL

Sex-typing

D13S317

D16S539 D18S51 D21S11 AMEL

13 14 15 16 17 18 19 20 21 22 X Y

Forensic DNA Testing

Probe subsets of genetic variation in order to differentiate between individuals (14 to 16 regions)

DNA typing must be done efficiently and reproducibly (information must hold up in court)
    Over 8 million profiles in the national FBI database

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

Applications

- Forensic cases: matching suspect with evidence
- Paternity testing: identifying father
- Missing persons investigations
- Military DNA "dog tag"
- Convicted offender DNA databases
- Mass fatalities
- Historical investigations
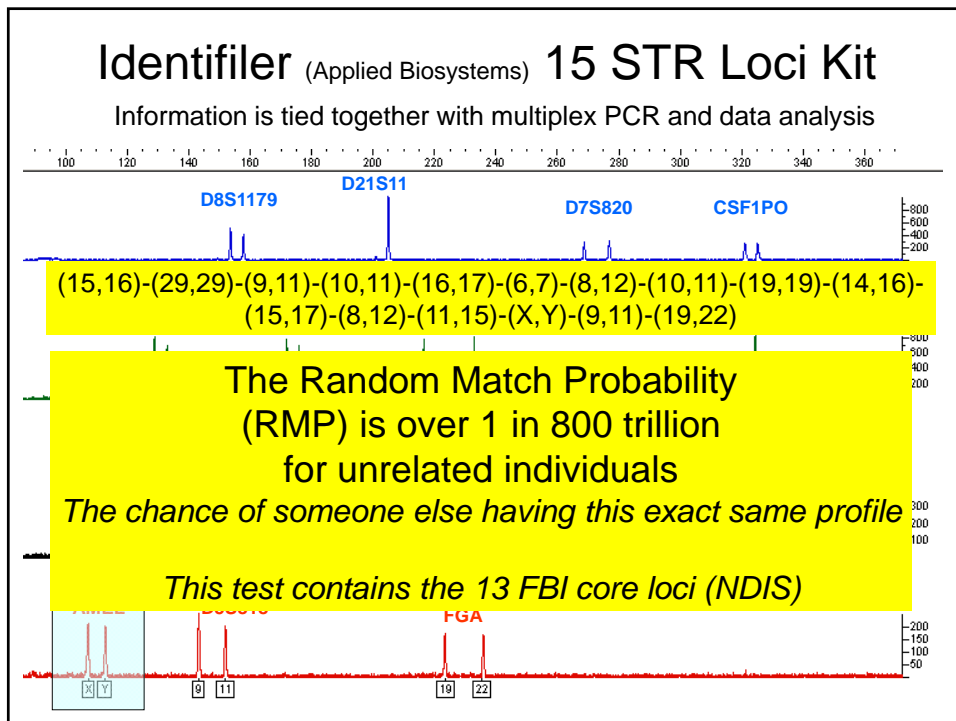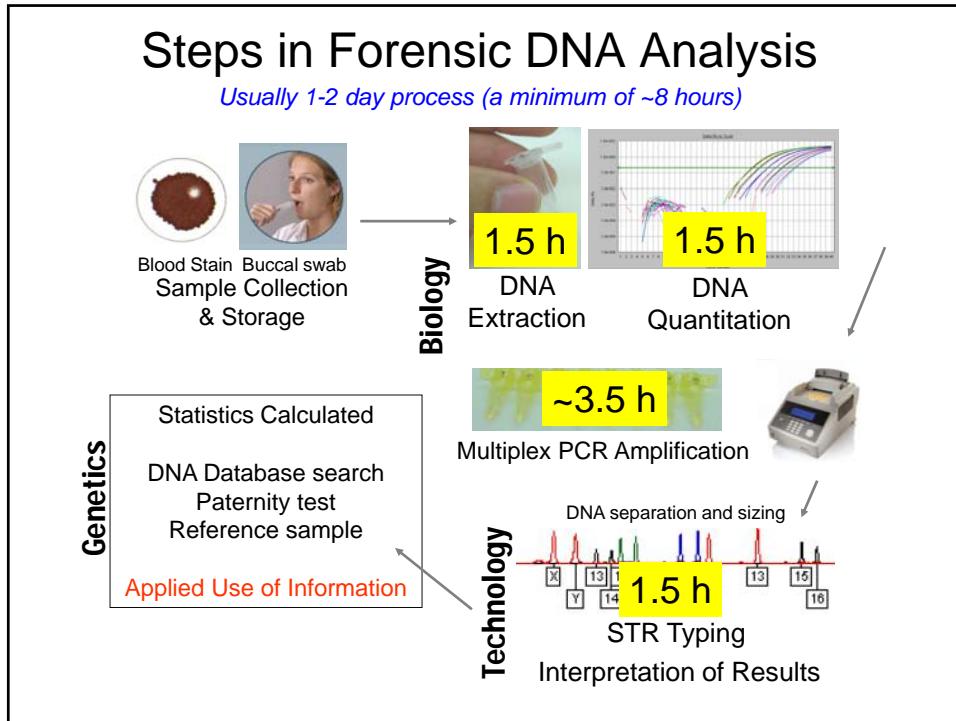- Genetic genealogy
- DNA as a biometric tool

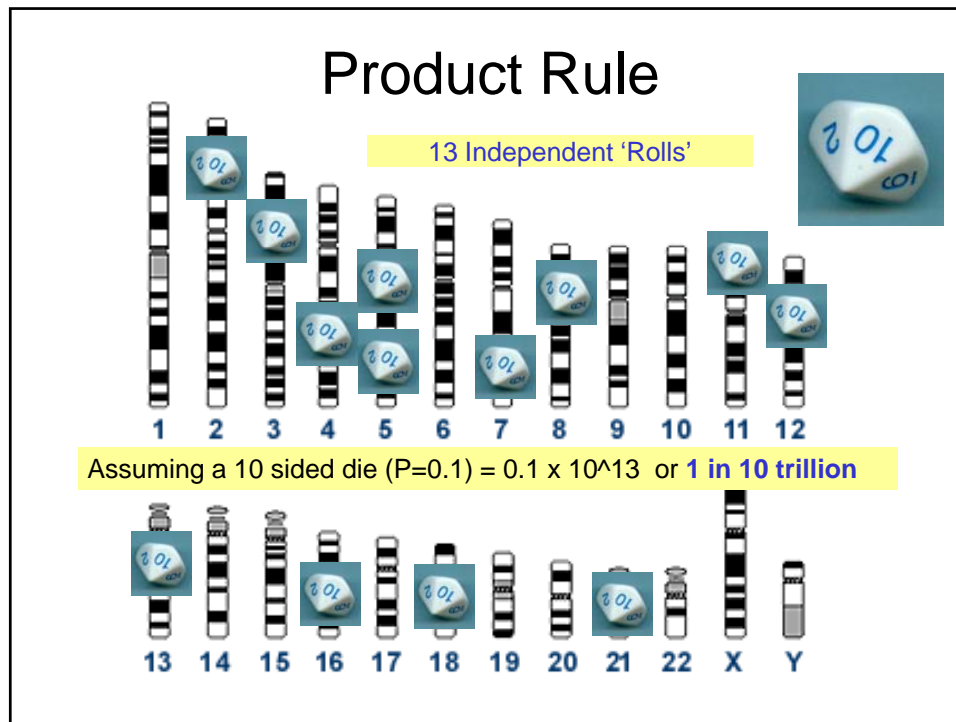# DNA Testing Requires a Reference Sample

A DNA profile by itself is fairly useless because it has no context…

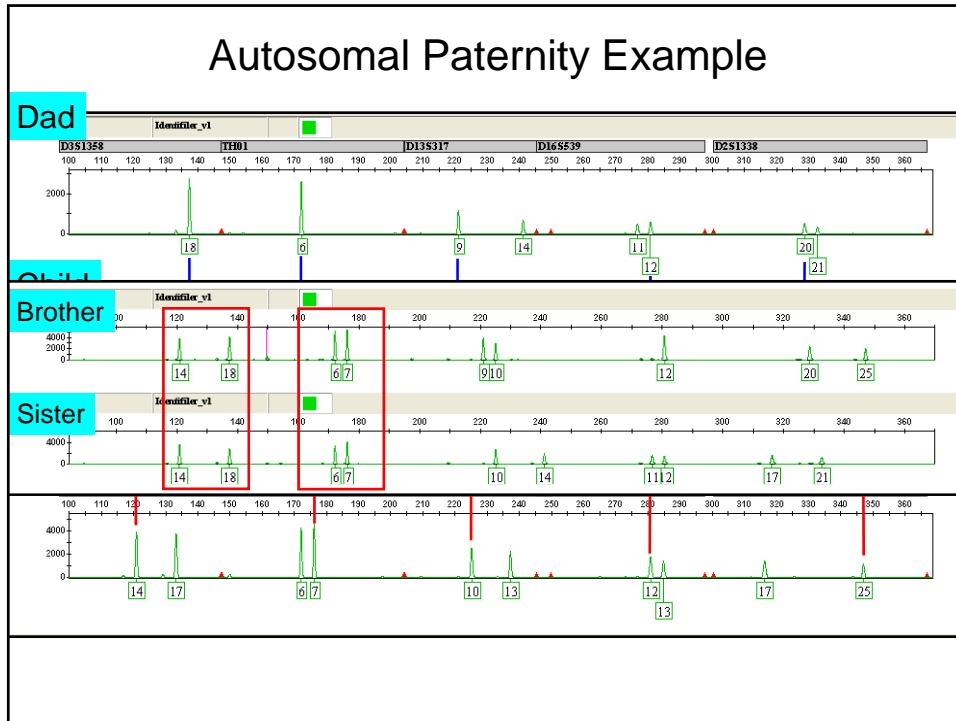DNA analysis for identity only works by comparison – **you need a reference sample**

Crime Scene Evidence compared to **Suspect(s)** (Forensic Case)
Child compared to **Alleged Father** (Paternity Case)
Victim's Remains compared to **Biological Relative** (Mass Disaster ID)
Soldier's Remains compared to **Direct Reference Sample** (Armed Forces ID)

Steps in Forensic DNA Analysis

Usually 1-2 day process (a minimum of ~8 hours)

Blood Stain Buccal swab
Sample Collection & Storage

Biology

1.5 h
DNA Extraction

1.5 h
DNA Quantitation

~3.5 h
Multiplex PCR Amplification

Statistics Calculated

DNA Database search
Paternity test
Reference sample

Applied Use of Information

Genetics

Technology

DNA separation and sizing

1.5 h
STR Typing
Interpretation of Results

Identifiler (Applied Biosystems) 15 STR Loci Kit

Information is tied together with multiplex PCR and data analysis

D8S1179   **D21S11**   **D7S820**   **CSF1PO**

(15,16)-(29,29)-(9,11)-(10,11)-(16,17)-(6,7)-(8,12)-(10,11)-(19,19)-(14,16)-(15,17)-(8,12)-(11,15)-(X,Y)-(9,11)-(19,22)

**The Random Match Probability
(RMP) is over 1 in 800 trillion
for unrelated individuals**
The chance of someone else having this exact same profile

This test contains the 13 FBI core loci (NDIS)

FGA

Product Rule

13 Independent 'Rolls'

Assuming a 10 sided die (P=0.1) = 0.1 x 10^13 or **1 in 10 trillion**

Kinship Testing

- DNA profiles can also be used to evaluate the probability of a specific familial relationship

- As a familial relationship becomes more distant, the ability of DNA to confirm the likelihood of that relationship decreases
  1. Parent-offspring
  2. Siblings
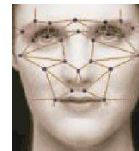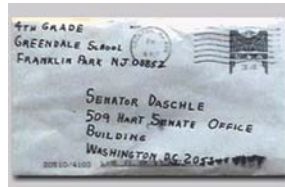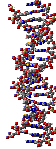 3. Half siblings = uncle/nephew = grandchild
  4. Cousins

Autosomal Paternity Example

DNA as a Biometric

Current Biometrics

Some commonly measured features

- Physical
  - Fingerprints (Palm/hand geometry)
  - Iris, retinal
  - Face
  - Odor/scent
  - DNA?

- Behavioral
  - Gait
  - Voice
  - Vein (IR thermogram)
  - Hand geometry
  - Handwriting

Characteristics of a Biometric

- Universality
  - each person should have the characteristic

- Uniqueness
  - is how well the biometric separates individuals from another

- Permanence
  - measures how well a biometric resists aging and variance over time

- Collectability
  - ease of acquisition for measurement

Jain, A. K.; Ross, Arun; Prabhakar, Salil (January 2004), "An introduction to biometric recognition", *IEEE Transactions on Circuits and Systems for Video Technology* **14th** (1): 4–20

Characteristics of a Biometric
(practical considerations)

- Performance
  - accuracy, **speed**, and robustness of technology used

- Acceptability
  - degree of approval of a technology

- Circumvention
  - ease of use of a substitute

Jain, A. K.; Ross, Arun; Prabhakar, Salil (January 2004), "An introduction to biometric recognition", *IEEE Transactions on Circuits and Systems for Video Technology* **14th** (1): 4–20

DNA Typing as a Biometric

Advantages	*Challenges*
• High level of accuracy (Gold Standard)	• Expensive
• Solid foundation of Forensic DNA Testing (pop stats, molecular biology, court acceptance, protocols, training, education)	• **Time consuming**
	• Sample collection (invasive, stability issues)
• Kinship determination (unique to DNA)	• Technical expertise required for analysis
• Potential use for: – Phenotype (traits; eye/hair color) – Biogeographical Ancestry	• Low level template, mixtures, PCR inhibition
	• Policy/Privacy/Ethical issues

Interest in Rapid DNA Typing

- DoD (field testing, rapid intelligence, mass fatalities)
- DHS (kinship determination, border security, immigration)
- DoJ (law enforcement, initial information)
- Industry (security, authentication)

- Each customer will have specific requirements
  - sample input
  - information output
  - degrees of 'accuracy'

  The time required for generating a STR profile will have to be significantly reduced

Goals for Rapid DNA Typing Systems

- Develop an integrated system capable of performing DNA testing in less than 1 hour

- Little user interaction (or experience)
- Rugged
- Robust

  Swab in…answer out

- Simple data interpretation
- 4-16 samples per run
- Disposable chips (with reagents on board)

Rapid DNA Typing Systems
Under Development

- Systems are currently under development and are not yet commercially available

- Network Biosystems (Woburn, MA)
  http://www.netbio.com
- ZyGEM and Lockheed Martin (Charlottesville,VA)
  http://www.zygem.com
- IntegenX (Pleasanton, CA)
  http://www.integenx.com
- Forensic Science Service (UK)
  http://www.forensic.gov.uk/

Use of DNA as a Biometric Tool American Academy of Forensic Science, Feb 22, 2010, Seattle, WA
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm
Biometrics Consortium Conference September, 2010 Tampa, FL
http://www.biometrics.org/bc2010/program.pdf

NIST Efforts with DNA Biometrics

- Developing rapid PCR protocols

- Evaluating kinship analysis software

- Support for external rapid DNA efforts

- Designing standards materials for device testing

- Testing prototype rapid DNA devices

Thank you for your attention!

Questions?

Peter.Vallone@nist.gov

301-975-4872

Acknowledgements

Erica Butts

Kristen O'Connor