Human Identity Testing Using DNA

DOC Headquarters - Commerce Research Library
12 March 2015

Peter M. Vallone, Ph.D.
Leader, Applied Genetics Group
Biomolecular Measurement Division

NIST
National Institute of Standards and Technology
U.S. Department of Commerce

Topics

- What is human identity testing?
  - Aka 'Forensic DNA Testing'
- Common questions
- NIST's role
- The future of human identity testing

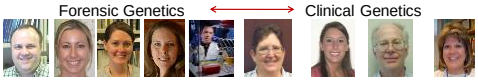
Applied Genetics Group

Group Leader
Peter Vallone

Advancing technology and traceability through quality genetic measurements to aid work in Forensic and Clinical Genetics

A core competency of our group is the application of *nucleic acid-based methods*
PCR – Genotyping – Sequencing – Real-time PCR – Digital PCR - DNA based SRMs

Forensic Genetics ⟷ Clinical Genetics



Mike Coble | Becky Steffen | Erica Romsos | Katherine Gettings (Postdoc) | Kevin Kiesler | Margaret Kline | Jo Lynne Harenza (Postdoc) | Dave Duewer (Data analysis support) | Patti Rohmiller (Office Manager)

General Characteristics of Genomic DNA

Each individual has a unique DNA profile
- with exception of monozygotic siblings[1]

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

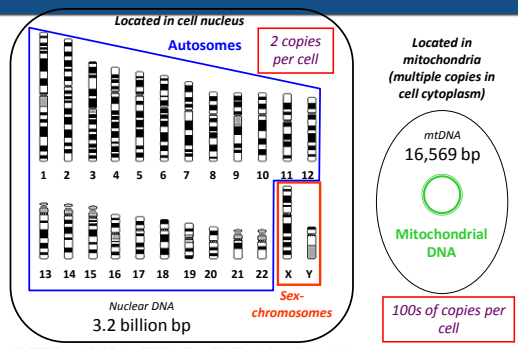
An individual's DNA profile remains the same throughout life

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

[1]Weber-Lehmann et al., Finding the needle in the haystack: Differentiating "identical" twins in paternity testing and forensics by ultra-deep next generation sequencing Forensic Science International: Genetics 9 (2014) 42–46

Human Genome

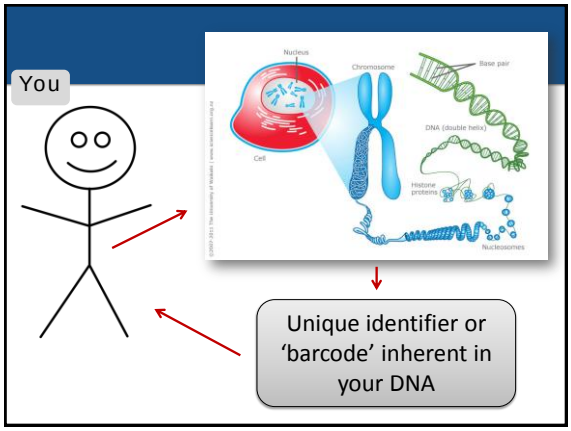
23 Pairs of Chromosomes + mtDNA



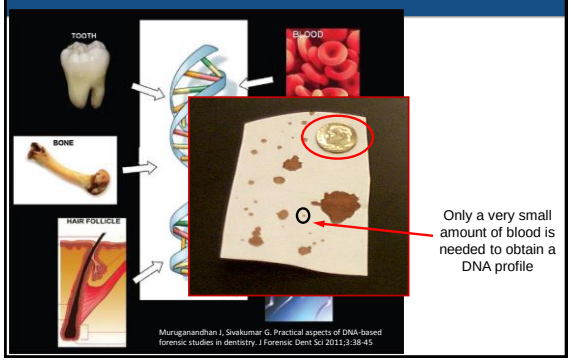
Human Identity Testing

- Probe subsets of genetic variation in order to differentiate between individuals
  - ≈13 to 24 regions in the human genome
- DNA typing must be done efficiently and reproducibly (information must hold up in court)
  - Over 13 million profiles in the national FBI database[1]
- Typically, we are *not* looking at genes – little/no information about ancestry, predisposition to disease, or phenotypic information (facial features, eye color, height, hair color) is obtained → **evolving**

[1]http://www.fbi.gov/about-us/lab/biometric-analysis/codis/codis-and-ndis-fact-sheet



Sources of Biological Evidence



Applications of Human Identity Testing

- Forensic cases: matching suspect with evidence
- Paternity testing: identifying father
- Missing persons investigations
- Military DNA "dog tag"
- National DNA database (FBI)
- Mass disasters: accounting for remains
- Historical investigations
- Genetic genealogy

What Type of Genetic Variation?

- Sequence Variation
  - single nucleotide polymorphisms (SNPs)
  - insertions/deletions

GCTAGTCGATGCTC**[G/A]**GCGTATGCTGTAGC

- Length Variation
  - **short tandem repeats** (STRs)

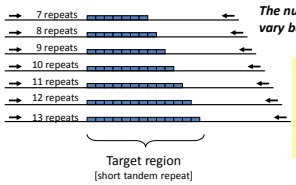
CTAGTCGT**[GATA][GATA][GATA]**GCGATCGT

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACAGGTG**GATAGATAGATAGA**
TAGATAGATAGATAGATAGATAGATAGATAGATATCATTGAAAGACAAAACAGAGATGGAT
GATAGATACATGCTTACAGATGCACAC

= 12 GATA repeats ("12" allele is reported)



The number of consecutive repeat units can vary between people

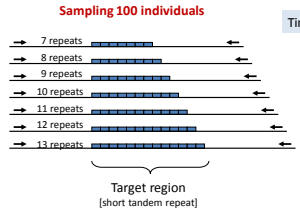
The frequency of these repeats observed in the general population have been sampled and are used for the statistical representation of a DNA profile

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACAGGTG**GATAGATAGATAGA**
TAGATAGATAGATAGATAGATAGATAGATAGATATCATTGAAAGACAAAACAGAGATGGAT
GATAGATACATGCTTACAGATGCACAC

Sampling 100 individuals



Times observed	Frequency	1 in …
5	5%	20
12	12%	8.3
22	22%	4.5
32	32%	3.1
20	20%	5
7	7%	14.3
2	2%	50
= 100	= 100%	

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

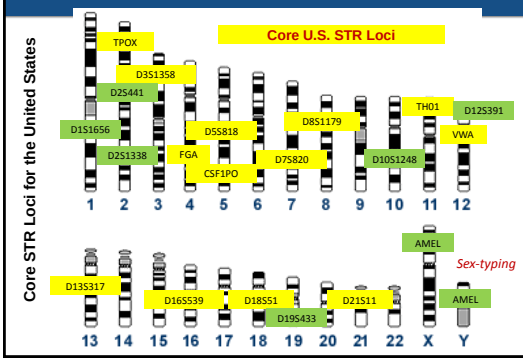
TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACAGGTGGATAGATAGATAGA
TAGATAGATAGATAGATAGATAGATAGATAGATATCATTGAAAGACAAAACAGAGATGGAT
GATAGATACATGCTTACAGATGCACAC

Marker	1	2	3	4	5	6	7	8	9	10	11	12	13	Full Profile
Frequency	0.2	0.04	0.12	0.33	0.1	0.57	0.31	0.22	0.15	0.27	0.3	0.11	0.09	1.48134E-10
1 in ...	5.00	25.00	8.33	3.03	10.00	1.75	3.23	4.55	6.67	3.70	3.33	9.09	11.11	6,750,627,188

Mock example to illustrate the product rule
Assuming independence, the frequency of each marker can be multiplied

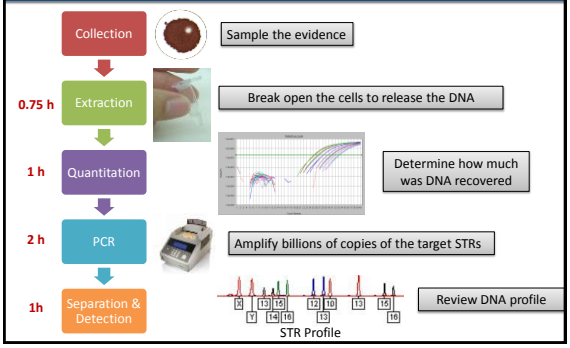
The chance of an **unrelated individual** having this exact same profile is 1 in 6,750,627,188

Position of Forensic STR Markers on Human Chromosomes



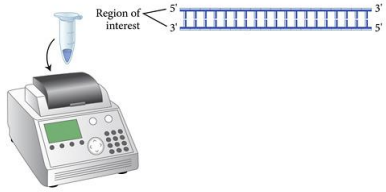
Steps in Forensic DNA Analysis

Usually 1-2 day process (a minimum of ~5 hours)

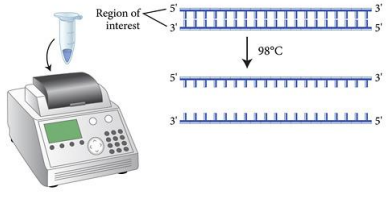


PCR

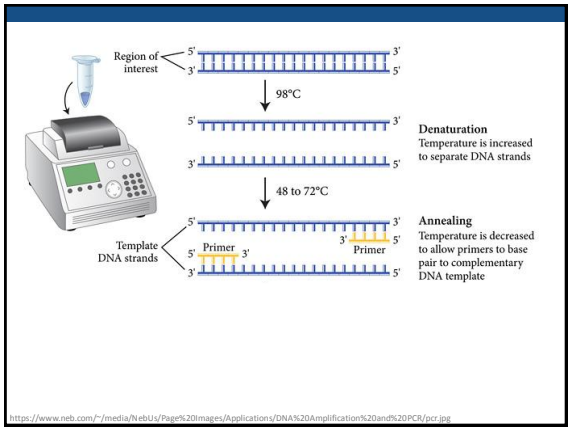
- Polymerase Chain Reaction
- A means to create billions of exact copies of a specific region of the genome

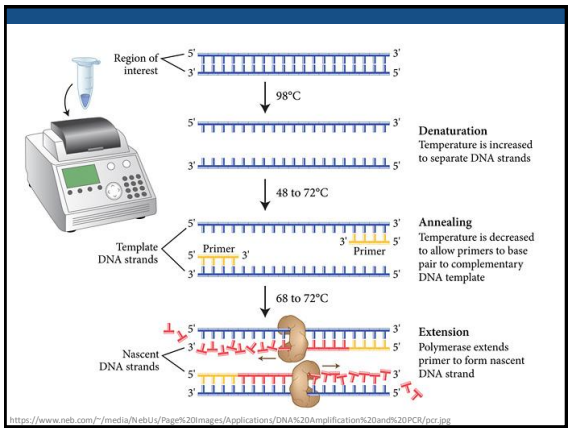


https://www.neb.com/~/media/NebUs/Page%20Images/Applications/DNA%20Amplification%20and%20PCR/pcr.jpg

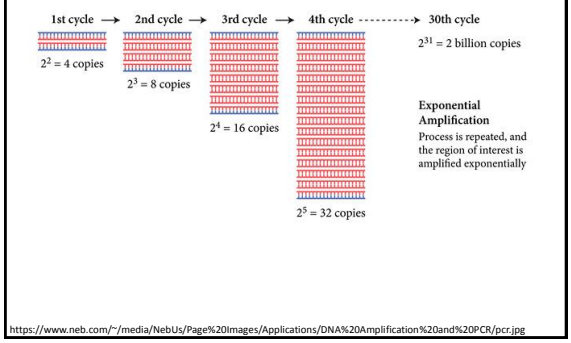


https://www.neb.com/~/media/NebUs/Page%20Images/Applications/DNA%20Amplification%20and%20PCR/pcr.jpg

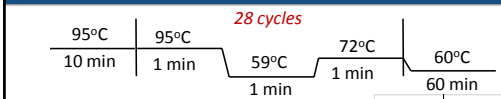




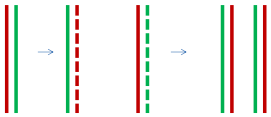
Exponential Amplification



PCR Thermal Cycling Profile



Duplex DNA denatures into 2 single strands
Polymerase and NTPs create new compliments
Repeat!

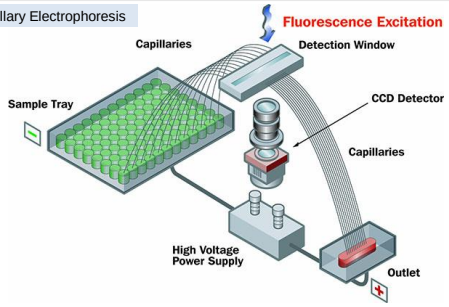


Polymerase and NTPs

Cycles	Copies
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
28	268,435,456.00

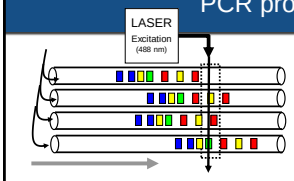
Separation and Detection of PCR products

Capillary Electrophoresis



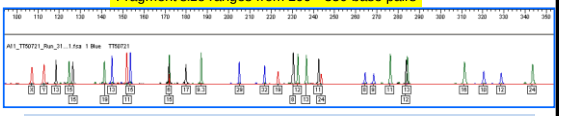
http://www.philekorea.co.kr/kor/data/cheditor4/1305/a49f0e4c8103cb5277053a548368c0d8_nC5e7v8vWJZS.jpg

Separation and Detection of PCR products



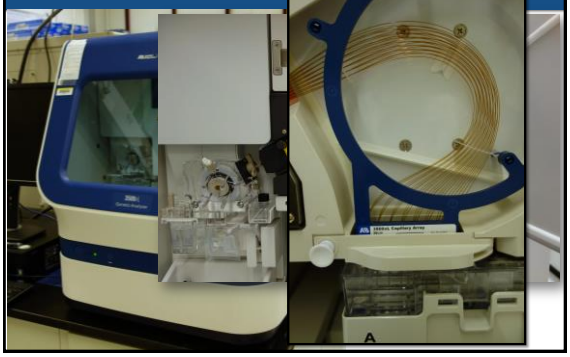
The labeled fragments are separated (based on size) and detected on a gel or capillary electrophoresis instrument
~1 hour or less

Fragment size ranges from 100 - 350 base pairs

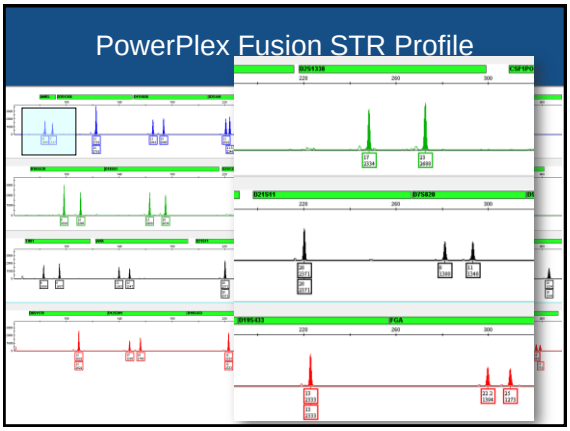


Peaks represent labeled DNA fragments separated by electrophoresis
This 'profile of peaks' is unique for an individual – a DNA type

Capillary Electrophoresis



PowerPlex Fusion STR Profile



PowerPlex Fusion STR Kit

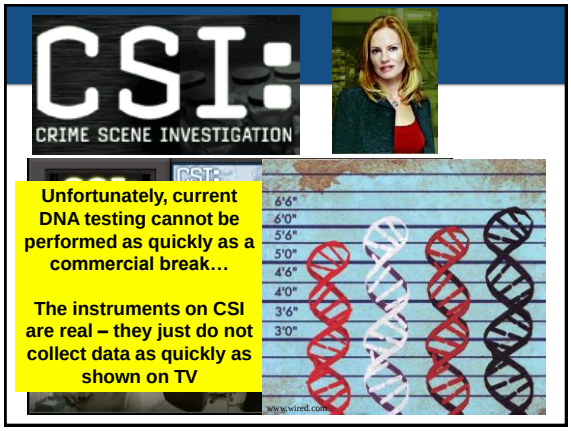
Amel	{X,Y}
D3S1358	{15,15}
D1S1656	{11,13}
D2S441	{11,11.3}
D10S1248	{8,12}
D13S317	{11,12}
Penta E	{8,14}
D16S539	{9,12}
D18S51	{15,18}
D2S1338	{17,23}
CSF1PO	{8,11}
Penta D	{9,11}
TH01	{6,9}
VWA	{13,15}
D21S11	{28,28}
D7S820	{8,11}
D5S818	{9,10}
TPOX	{9,12}
DYS391	{10}
D8S1179	{15,15}
D12S391	{17,19}
D19S433	{13,13}
FGA	{22.2,25}
D22S1045	{16,17}

Multiplying the frequency of each genotype at each locus gives us the Random Match Probability (RMP) of 7.81×10^{-39} for **unrelated individuals**

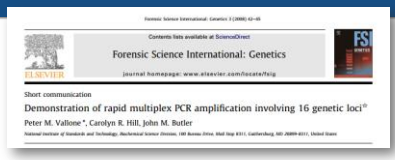
*The chance of an **unrelated individual** having this exact same profile is 1 in…*

128,040,973,111,396,000,000,000,000,000 ,000,000,000

This test contains the FBI core STR markers

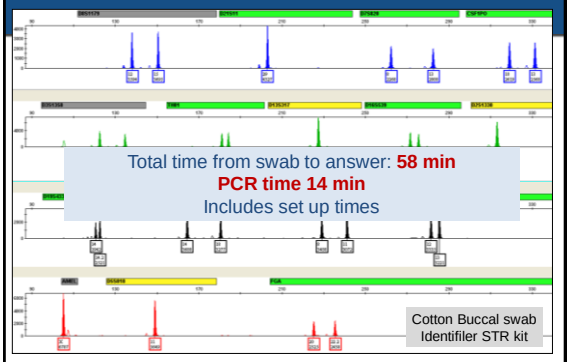


NIST and Rapid PCR


Forensic Science International: Genetics
Short communication
Demonstration of rapid multiplex PCR amplification involving 16 genetic loci
Peter M. Vallone*, Carolyn R. Hill, John M. Butler

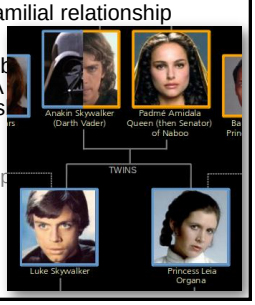
- Up until 2008 PCR amplification times required approximately 3 hours
- Utilizing new (faster) DNA polymerases and rapid PCR thermal cyclers we demonstrated results in 36 minutes
- Enabling faster commercial STR typing kits (40 min) and fully integrated 'Rapid DNA' typing instruments (swab to profile in < 2 hours)

Example of a Rapid STR Profile



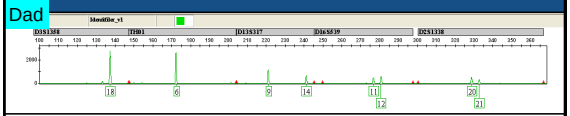
Kinship Testing

- DNA profiles can also be used to evaluate the probability of a specific familial relationship
- As a familial relationship b
  distant, the ability of DNA
  likelihood of that relations
  1. Parent-offspring
  2. Siblings
  3. Half siblings = uncle/nep
  4. Cousins

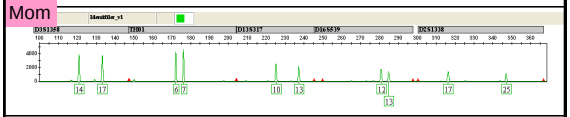


Paternity Example

Dad

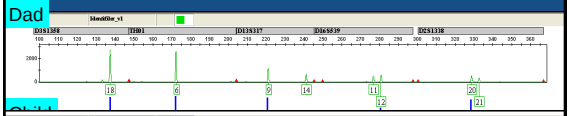


Mom

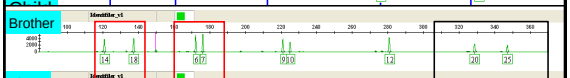


Paternity Example

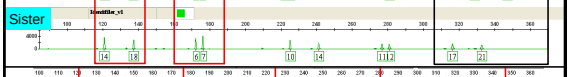
Dad

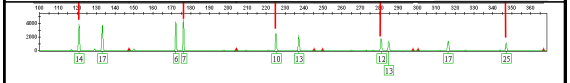


Brother

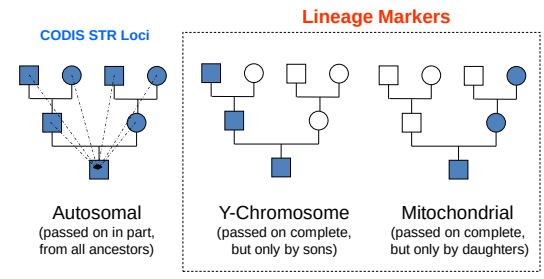


Sister





Different Inheritance Patterns



Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.1, ©Elsevier Science/Academic Press

NIST's Role in Human Identity Testing

- Reference materials
  - To ensure confidence in results
- Research
  - New Technologies
- Training and Education
  - Workshops, papers, talks

SRM 2391c: PCR-Based DNA Profiling Standard

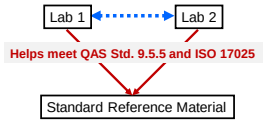
9.5.5 The laboratory shall check its DNA procedures annually or whenever substantial changes are made to a procedure against an appropriate and available ***NIST standard reference material or standard traceable to a NIST standard****.*

http://www.fbi.gov/about-us/lab/biometric-analysis/codis/qas_testlabs

SRM 2391c: PCR-Based DNA Profiling Standard

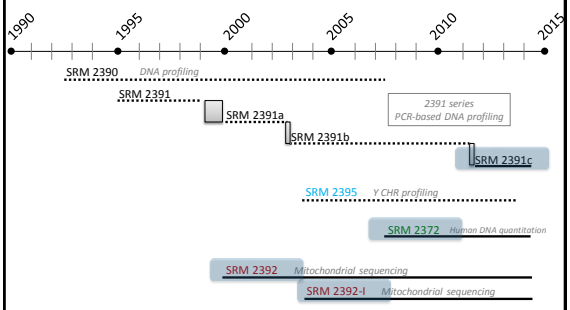
- Components A through D: DNA extracts in liquid form
- Components E and F: DNA spotted on paper
- Certified values for STR alleles based on DNA sequencing

Genomic DNAs characterized for the CODIS core loci and Y-STRs

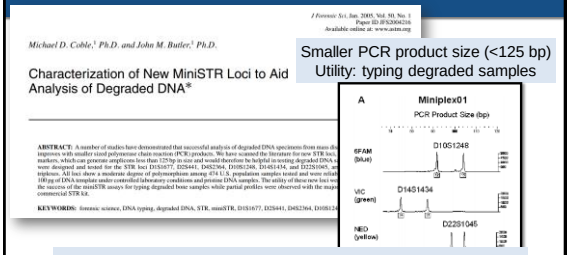


Calibration with SRMs enables confidence in comparisons of results between laboratories

NIST Forensic SRM Timeline



miniSTRs


Michael D. Coble,[1] Ph.D. and John M. Butler,[1] Ph.D.

Characterization of New MiniSTR Loci to Aid Analysis of Degraded DNA*

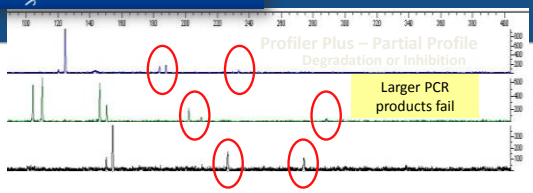
Smaller PCR product size (<125 bp)
Utility: typing degraded samples

Initial work developing miniSTRs was started for WTC identifications
John Butler (NIST), Bruce McCord (FIU), Bode Technology Group (Lorton, VA)

Technology adopted by U.S. commercial STR kit vendors
(Life Tech: MiniFiler, Promega: S5)

Bode Technology. Advancing Identity Technologies

World Trade Center – Phase I Summary



12,392	Bone samples processed
3,405	Full profiles (13 STR loci)
2,143	High partial profiles (≥7 STR loci)
2,670	Low partial profiles (<7 STR loci)
4,174	No loci

Over 6800 profiles

miniSTRs are helping here

Final 20% of WTC victims identified were based on a miniSTR technique pioneered at NIST

Additional STR Loci

- The U.S. is increasing the number of core STR markers required in the national DNA database from 13 to 20

Core 13:
1. TPOX
2. D3S1358
3. D5S818
4. FGA
5. CSF1PO
6. D7S820
7. D8S1179
8. TH01
9. vWA
10. D13S317
11. D16S539
12. D18S51
13. D21S11

Additional 7:
14. D1S1656
15. **D2S441**
16. D2S1338
17. **D10S1248**
18. D12S391
19. D19S433
20. **D22S1045**

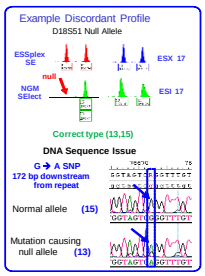
+ Amelogenin

Three of the new STR markers characterized by NIST

Coble, M.D. and Butler, J.M. (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA. J. Forensic Sci. 50: 43-53.

STR Kit Testing at NIST

- When new commercial STR multiplex kits are developed, prototypes are often sent to NIST for testing
- These kits are tested on 665 highly characterized population samples to determine if there are any incorrect genotypes
- The cause of any discordances are confirmed by DNA sequencing and reported to the commercial companies prior to the kits being released to the forensic community



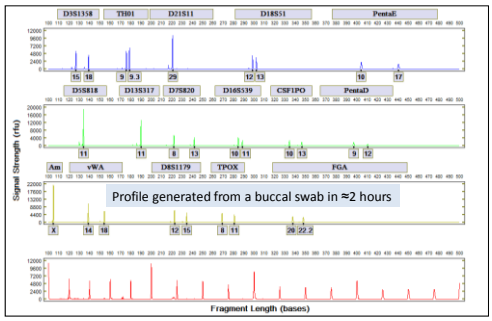
The Future of human identity testing

Rapid DNA

- Fully integrated instruments capable of generating a DNA profile
- 'Swab in profile out' in less than 2 hours
- No special training required
- Potential applications
  - Booking station
  - In the field
  - Border crossing

http://integenx.com/ (RapidHit 200)
http://netbio.com/ (DNA Scan)

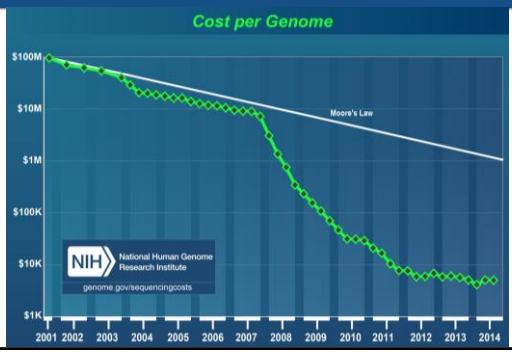
Rapid DNA Profile
Fully integrated process



Next Generation Sequencing

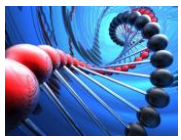
High throughput DNA sequencers capable of sequencing million to billions of bases of DNA per day are now commonplace

Next Generation Sequencing



Non-forensic applications

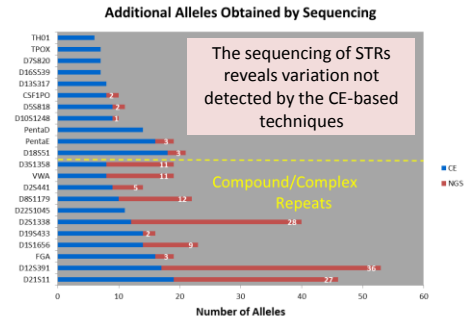
- Clinical research
- Inherited disease
- Reproductive health
- Cancer – gene fusion
- Rare variants
- Pre-implantation (genetic screening)
- Transplant medicine (HLA)
- Microbiomics/Metagenomics
- Gene expression | RNA seq
- Public health
- Ancient DNA
- NIPT (non-invasive prenatal testing)



Use of NGS for forensic applications

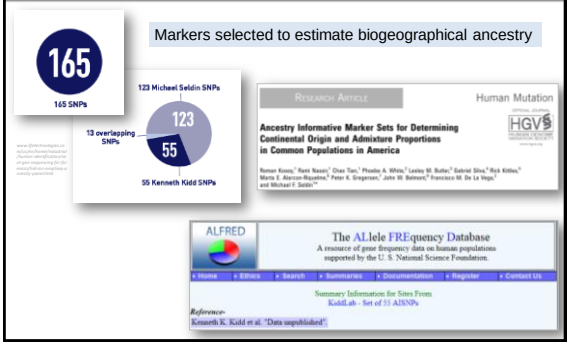
- Forensically relevant markers (SNPs)
  - newer human identity applications
  - biogeographical ancestry, externally visible traits, complex kinship
- Going in depth into STR loci and beyond
  - STRs are useful for legacy (databases)
  - SNPs within STRs identify 'sub-alleles'

Sequencing STRs

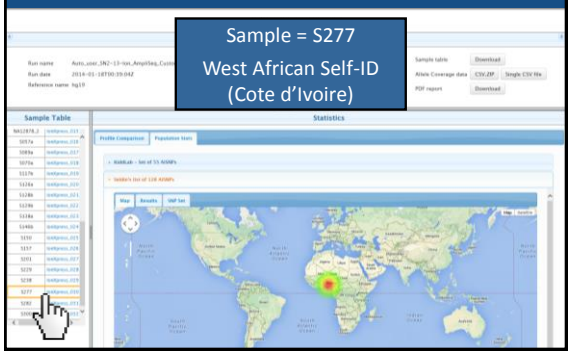


HID-Ion Ampliseq Ancestry Panel

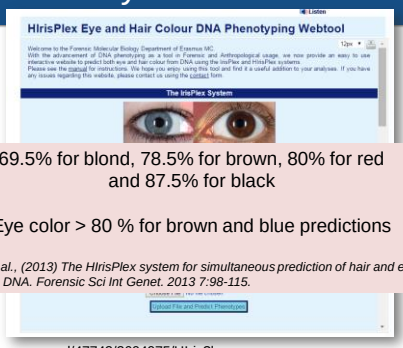
Ion Torrent PGM



Ancestry SNP NGS Assay



Hair and Eye Color Prediction



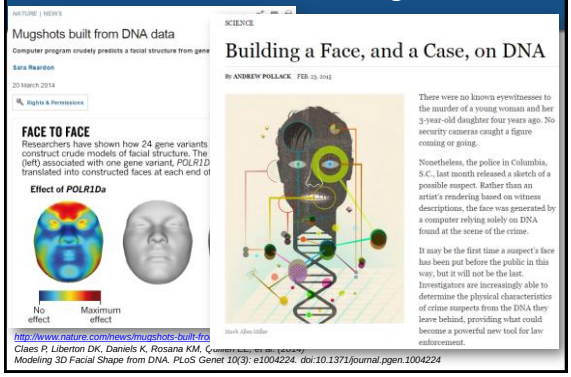
69.5% for blond, 78.5% for brown, 80% for red and 87.5% for black

Eye color > 80 % for brown and blue predictions

Walsh S et al., (2013) The HIrisPlex system for simultaneous prediction of hair and eye colour from DNA. Forensic Sci Int Genet. 2013 7:98-115.

http://www.erasmusmc.nl/47743/3604975/HIris?lang=en

Facial Modeling



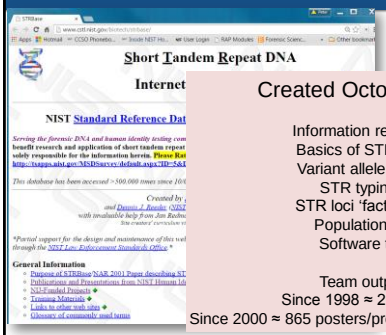
http://www.nature.com/news/mugshots-built-fro

Claes P, Liberton DK, Daniels K, Rosana KM, Quillen EE, et al. (2014) Modeling 3D Facial Shape from DNA. PLoS Genet 10(3): e1004224. doi:10.1371/journal.pgen.1004224

NIST Resources

STRBase ("Star – base")

http://www.cstl.nist.gov/strbase/ Or Google STRBase



Created October, 1997

Information related to:
Basics of STR typing
Variant allele reports
STR typing kit
STR loci 'fact sheets'
Population data
Software tools

Team outputs:
Since 1998 ≈ 200 papers
Since 2000 ≈ 865 posters/presentations/workshops

Textbooks Written by Dr. John Butler

Applied Genetics Group

Group Leader
Peter Vallone

Advancing technology and traceability through quality genetic measurements to aid work in Forensic and Clinical Genetics

A core competency of our group is the application of *nucleic acid-based methods*
PCR – Genotyping – Sequencing – Real-time PCR – Digital PCR - DNA based SRMs

Forensic Genetics ⟷ Clinical Genetics

Mike Coble | Becky Steffen | Erica Romsos | Katherine Gettings (Postdoc) | Kevin Kiesler | Margaret Kline | Jo Lynne Harenza (Postdoc) | Dave Duewer (Data analysis support) | Patti Rohmiller (Office Manager)

Acknowledgements

The Applied Genetics Group

Funding

NIST: "Forensic DNA"

FBI: "DNA as a Biometric"

DHS: "Rapid DNA and Kinship support"

NIJ: Interagency agreements

peter.vallone@nist.gov

301-975-4872