Advanced Topics in Forensic DNA Analysis

mtDNA

New Jersey State Police
Training Workshop

Hamilton, NJ
December 5-6, 2006

Dr. John M. Butler
National Institute of Standards and Technology

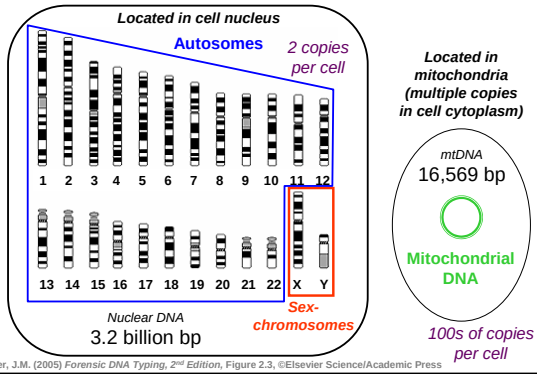
john.butler@nist.gov

Outline for This Section

- Lineage Markers
- mtDNA background and fundamentals
- HV1 & HV2 sequence and interpretation issues
- Tools for mtDNA screening – LINEAR ARRAYs
- Emerging mtDNA technologies – mtDNA genome sequencing for increased discrimination, mtDNA micro-chip technology

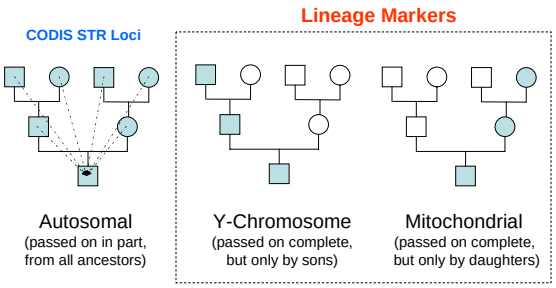
Human Genome

23 Pairs of Chromosomes + mtDNA



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

Different Inheritance Patterns



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

Mitochondrial DNA Inheritance

Sons receive mtDNA from their mothers but do not pass it on to their offspring

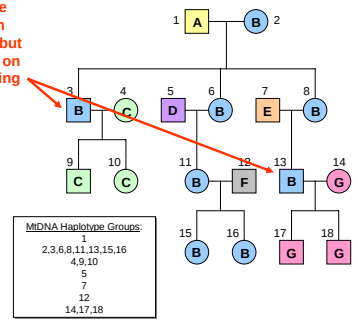


Figure 10.2, J.M. Butler (2005) *Forensic DNA Typing,* 2nd Edition © 2005 Elsevier Academic Press

Role of Y-STRs and mtDNA

Compared to Autosomal STRs

- **Autosomal STRs provide a higher power of discrimination and <u>are the preferred method</u> whenever possible**
- **Due to capabilities for male-specific amplification**, Y-chromosome STRs (**Y-STRs**) can be useful in extreme female-male mixtures (e.g., when differential extraction is not possible such as fingernail scrapings)
- **Due to high copy number**, mitochondrial DNA (**mtDNA**) may be the only source of surviving DNA in highly degraded specimens or low quantity samples such as hair shafts

A mtDNA result is better than no result at all…

Lineage Markers: Y-STRs and mtDNA

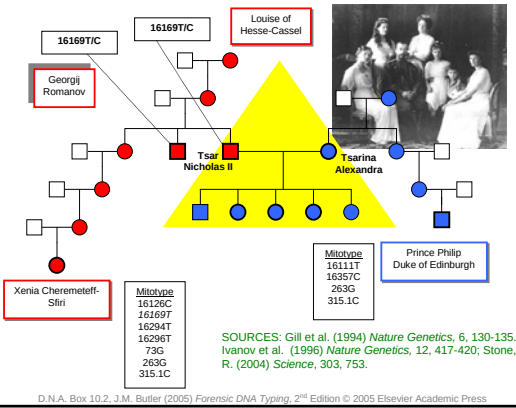
Advantages

- Extend possible reference samples beyond a single generation (benefits missing persons cases and genetic genealogy)
- Family members have indistinguishable haplotypes unless mutations have occurred

Disadvantages

- Lower power of discrimination due to no genetic shuffling with recombination
- Family members have indistinguishable haplotypes unless mutations have occurred

Identifying the Romanov Remains (the Last Russian Czar)



SOURCES: Gill et al. (1994) *Nature Genetics*, 6, 130-135.; Ivanov et al. (1996) *Nature Genetics*, 12, 417-420; Stone, R. (2004) *Science*, 303, 753.

D.N.A. Box 10.2, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Genetic Genealogy Companies

FamilyTreeDNA
http://www.familytreedna.com
http://www.dna-fingerprint.com

Sorenson Genomics
http://www.sorensongenomics.com

Relative Genetics
http://www.relativegenetics.com

oxford ancestors
http://www.oxfordancestors.com

DNA Heritage
http://www.dnaheritage.com

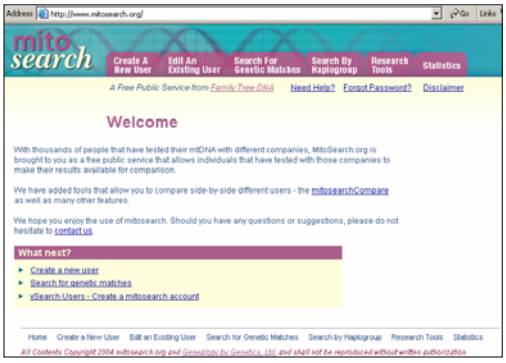
Ethno Ancestry
http://www.ethnoancestry.com

GeoGene
http://www.geogene.com

The rapidly growing field of genetic genealogy is expanding the use of mtDNA and Y-STRs.

FamilyTreeDNA

www.mitosearch.org



FamilyTreeDNA

www.Ysearch.org

Genetic Genealogy

http://www.isogg.org

"*The mission of the International Society of Genetic Genealogy is to advocate for and educate about the use of genetics as a tool for genealogical research, and promote a supportive network for genetic genealogists.*"

Famous DNA

http://www.isogg.org/famousdna.htm

Jesse James

In 1995, the infamous outlaw, Jesse James' body was exhumed for DNA testing. Samples of hair and other fragments from his first burial site at his home were also recovered and sampled, along with mtDNA samples extracted from two living James' relatives. All samples resulted in a perfect match, thus concluding that the body in Mt. Olivet Cemetery, is indeed that of Jesse James.

Name	mtDNA Haplogroup	mtDNA Sequence
Jesse James	T2	16126C, 16274A, 16294T, 16296T, 16304C

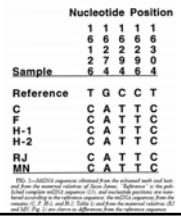
Jesse James mtDNA Results

Stone *et al.* (2001) *J. Forensic Sci.* 46(1):173-176

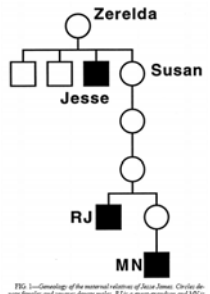
LAST WORD SOCIETY

Anne C. Stone,[1] Ph.D.; James E. Starrs,[2] LL.M.; and Mark Stoneking,[1] Ph.D.

Mitochondrial DNA Analysis of the Presumptive Remains of Jesse James*



Sample	Nucleotide Position 16126	16274	16294	16296	16304
Reference	T	G	C	C	T
C	C	A	T	T	C
F	C	A	T	T	C
H-1	C	A	T	T	C
H-2	C	A	T	T	C
RJ	C	A	T	T	C
MN	C	A	T	T	C



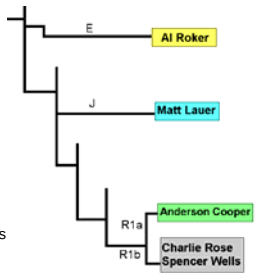
DNA Results from Some Famous People

Famous Haplogroups

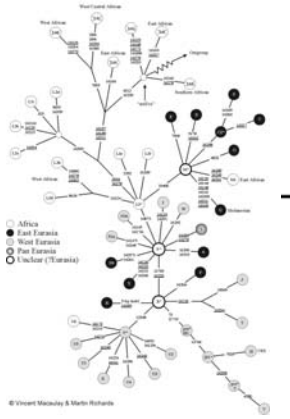
http://www.isogg.org/famoushg.htm

The following haplogroups for the hosts of the "Today Show" were aired 18 Nov 2005 during an interview with Dr. Spencer Wells. The Y-haplogroups for Dr. Wells and PBS host, Charlie Rose, were aired on the 23 Jan 2006 segment of the "Charlie Rose Show". Anderson Cooper's haplogroup was referenced on 21 Feb 2006 "Anderson Cooper 360" segment.

Y-SNP Haplogroups



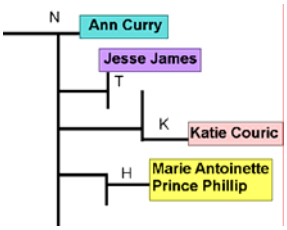
http://msnbc.msn.com/id/10095659/



http://www.stats.gla.ac.uk/~vincent/images/skeleton07-08-02.jpg

mtDNA Results

mtDNA Haplogroups



http://www.isogg.org/famoushg.htm

http://www.isogg.org/ancientdna.htm

Cheddar Man

In 1903, skeletal remains were found in a cave in Cheddar, England. The remains of a 23 year-old man, who was killed by a blow to the face, were discovered to be at least 9,000 years old. Ninety-four years after the discovery of "Cheddar Man", scientists were able to extract mitochondrial DNA from his tooth cavity.

Bryan Sykes, and his team at Oxford University distributed DNA test kits to local Cheddar schools, and a match was found to a local schoolteacher, Adrian Targett.

Name	mtDNA Haplogroup	mtDNA Sequence
Cheddar Man	U5a	16192T, 16270T

The New York Times nytimes.com

March 24, 1997

Tracing Your Family Tree to Cheddar Man's Mum

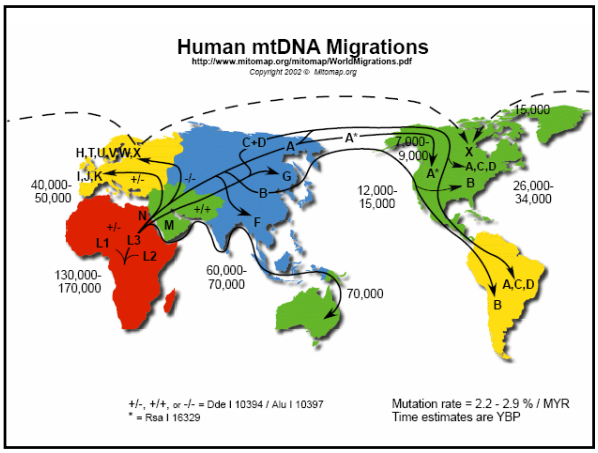
By SARAH LYALL

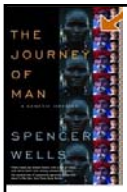
Until several weeks ago, Adrian Targett, a high school history teacher, didn't appear to have much in common with Cheddar Man, a 9,000-year-old pile of bones at the Natural History Museum in London.

Sure, Mr. Targett had heard of Cheddar Man, and had even visited the cave in this quaint Somerset village where his skeleton was found in 1903. But after a seemingly quixotic experiment in which scientists compared Cheddar Man's DNA to that of 20 local residents, Mr. Targett recently received a wholly unexpected piece of news: He is, it seems, related to Cheddar Man on his mother's side.

"I'm thinking of writing to the Marquess of Bath, who owns these caves, and saying, 'I'd like my cave back,' " Mr. Targett, 42, said over a meat pie and a pint in the local pub recently, considering the implications of having such a venerable relative. "All those times I'd visited this cave before, and I'd never realized I was going home."

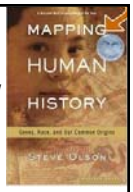
http://query.nytimes.com/gst/fullpage.html?res=9807EEDB133BF937A15750C0A961958260&sec=health&pagewanted=print





The Genographic Project

https://www3.nationalgeographic.com/genographic/



- Funded $50 million for 5 years by IBM and National Geographic
- Will gather and run DNA samples from ~100,000 people around the world with Y-SNPs and mtDNA
- For U.S. participants, Mike Hammer's lab is running 12 Y-STRs or sequencing mtDNA HV1

Perhaps the Real Reason Some Genetic Genealogy Is Performed…

"You don't look anything like the long haired, skinny kid I married 25 years ago. I need a DNA sample to make sure it's still you."

Y-Chromosome and Mitochondrial DNA Analysis

mitochondrial DNA

Over 350 mtDNA slides available on STRBase at
http://www.cstl.nist.gov/biotech/strbase/YmtDNAworkshop.htm

NEAFS 2006 Workshop
Rye Brook, NY
November 1, 2006

Northeastern Association of Forensic Scientists

Dr. John M. Butler
Dr. Michael D. Coble

john.butler@nist.gov
Michael.Coble@afip.osd.mil

Location and Copy Number of mtDNA

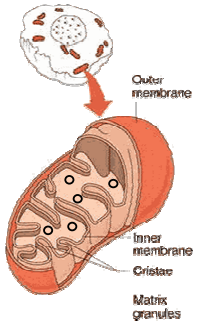
- Found within the mitochondria in the cellular cytoplasm.
- On average 4-5 copies of mtDNA molecules per mitochondria (range of 1-15 mtDNA copies).
- Number of mitochondria vary by cell type (e.g., muscles have more…).
- Generally, hundreds of mitochondria per cell.

Comparison of Human nucDNA and mtDNA

Characteristics	Nuclear DNA (nucDNA)	Mitochondrial DNA (mtDNA)
Size of genome	~3.2 billion bp	~16 569 bp
Copies per cell	2 (1 allele from each parent)	Can be > 1000
Percent of total DNA content per cell	99.75%	0.25%
Structure	Linear; packaged in chromosomes	Circular
Inherited from	Father and Mother	Mother
Chromosomal pairing	Diploid	Haploid
Generational recombination	Yes	No
Replication repair	Yes	No
Unique	Unique to individual (except identical twins)	Not unique to individual (same as maternal relatives)
Mutation rate	Low	At least 5–10 times nucDNA
Reference sequence	Described in 2001 by the Human Genome Project	Described in 1981 by Anderson and co-workers

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Table 10.1, ©Elsevier Science/Academic Press

Mitochondrial Morphology



Cytoplasmic organelle

Double membrane

Outer membrane – porin proteins for the transportation of materials.

Inner membrane – highly folded (increased surface area) and highly impermeable.

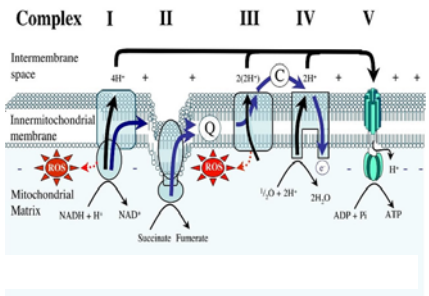
Inner Matrix – several copies of mtDNA

Mitochondrial Functions

Cellular Respiration – ATP production via oxidative-phosphorylation (OX-PHOS).

Apoptosis – programmed cell death
Steroid synthesis
Elongation of fatty acids
Oxidation of epinephrine (adrenaline)
Degradation of tryptophan
Heme synthesis
Heat production

Ox-Phos



http://myweb.uiowa.edu/bballard/Research%20Figures/OXPHOS.jpg

Mitochondrial Proteins Come from Nuclear Genes as Well as mtDNA Genes

Complex	I	II	III	IV	V
Enzyme	NADH-CoQ Reductase	Succinate-CoQ Reductase	CoQ-Cytochrome C Reductase	Cytochrome C Oxidase	ATP Synthase
Inhibitor	Rotenone Amytal	TTFA malonate	Antimycin A	Cyanide Carbon Monoxide Azide	Oligomycin
Nuclear DNA Subunits	~43	4	10	10	~14
mtDNA Subunits	7 ND1-6, ND4L	0	1 Cytochrome b	3 COX I, II, III	2 ATPase 6 ATPase 8

~81 subunits encoded by the nuclear genome

Forensic Focus

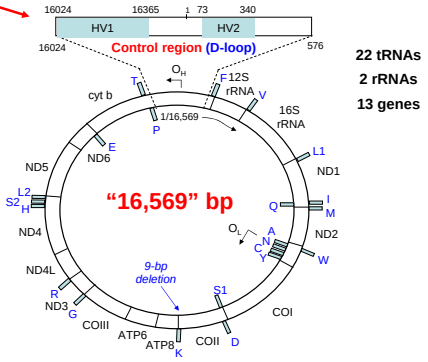


Figure 10.1, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

mtDNA Is Not Always 16,569 bp …

- Dinucleotide repeat at positions 514-524 (near end of control region)
  - Usually ACACACACAC or $(AC)_5$ in most individuals
  - Can vary from $(AC)_3$ to $(AC)_7$
- Other insertions and deletions may occur
  - 9 bp deletion (positions 8277 to 8285) in some individuals from Asia and Pacific Islands (haplogroup B) and Africans (haplogroup L)

Control Region (16024-576)

- 1,122 nucleotide positions
- Typically only **610 bases examined**
  - (HVI: 16024-16365; HVII: 73-340)

Coding Region (577-16023)

- 15,446 nucleotide positions
- Challenges with typing widely spaced SNPs
  - Multiplex PCR required
- Polymorphisms may have medical significance

"Heavy" vs. "Light" Strand

- The two strands ("inner" and "outer" loops) of mtDNA can be separated with an alkaline CsCl gradient.
- Heavy or H-strand contains a greater number of guanine nucleotides (largest molecular weight of the four nucleotides) – purine rich.
- Light or L-strand contains more C and T nucleotides and is thus physically lighter (pyrimidine rich).
- H-strand codes for 28 gene products while the L-strand is used to transcribe 8 tRNAs and the ND6 protein product.

Original Reference Sequence

- Human mtDNA was first sequenced in 1981 in Frederick Sanger's lab located in Cambridge, England.
- Authors for this paper (Nature 1981, 290:457-465) were listed in alphabetical order so Stan Anderson was the first author.
- This sequence has come to be referred to as the **"Anderson" sequence** (GenBank accession: M63933).
- This first sequence is sometimes called the **Cambridge Reference Sequence (CRS).**

Re-Sequencing of CRS

- The 1981 sequence was derived primarily from a placenta of an individual with European ancestry; however, some HeLa and bovine sequence was used to fill in gaps due to early sequencing procedures performed.
- Re-analysis of original placental material by Andrews et al. (1999) found 11 nucleotides that differed from Anderson et al. (1981) sequence.
- This **revised Cambridge Reference Sequence (rCRS)** is now the accepted standard for comparison.

Evaluation of Sequence Differences

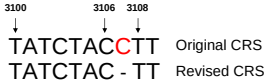
Between CRS (Anderson et al. 1981) and rCRS (Andrews et al. 1999)

Nucleotide Position	Region of mtGenome	Original CRS	Revised CRS	Remarks
3106–3107	16S rRNA	CC	C	Error
3423	ND1	G	T	Error
4985	ND2	G	A	Error
9559	COIII	G	C	Error
11335	ND4	T	C	Error
13702	ND5	G	C	Error
14199	ND6	G	T	Error
14272	ND6	G	C	Error (bovine sequence inserted)
14365	ND6	G	C	Error (bovine sequence inserted)
14368	ND6	G	C	Error
14766	cyt b	T	C	Error (HeLa sequence inserted)

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Table 10.3, ©Elsevier Science/Academic Press

Further Comparison of CRS and rCRS

- No differences seen between CRS and rCRS within the mtDNA control region.
- **The original CRS contained a "CC" at positions 3106-3107 but rCRS was found to possess only a single "C"**



- Thus, **rCRS is only 16,568 bp**!

Maternal Inheritance of mtDNA

- Fertilizing sperm contributes only nuclear DNA.
- Cellular components including the mitochondria in the cytoplasm come from the mother's ovum.
- Any sperm mitochondria that may enter a fertilized egg are selectively destroyed due to a ubiquitin tag added during spermatogenesis.
- Barring mutation, a mother passes her mtDNA type on to her children.

Maternal Inheritance of mtDNA

Note that mtDNA is not unique to an individual

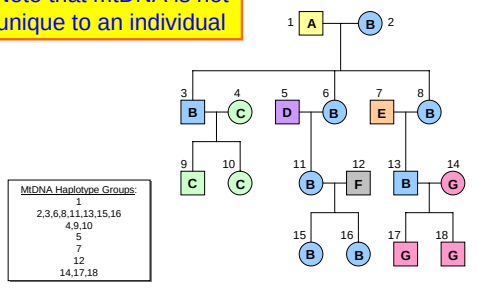


Figure 10.2, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

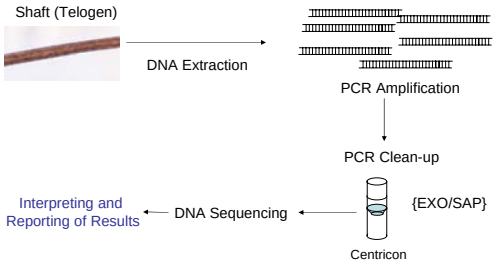
Summary – mtDNA Characteristics

- High copy number of mtDNA.
- Maternal inheritance of mtDNA.
- Lack of recombination.
- High mutation rate compared to single copy nucDNA.

Methods for Measuring mtDNA Variation

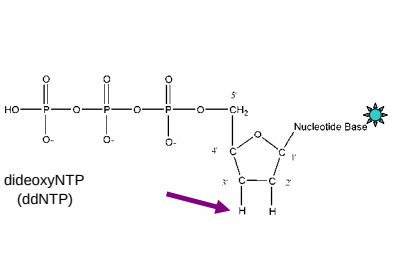
- Low-resolution RFLP (1980s)
- High-resolution RFLP (1990s)
- Sequence analysis of HV1 and HV2 within control region (1991-present)
- Sequence analysis of complete mtDNA genome (2000-present)

Process for Evaluation of mtDNA Samples



Science of DNA Sequencing

- Fred Sanger – developed the dideoxy method of sequencing in the 1970s… still used today.



Sanger Sequencing

DNA template 3′-TAAATGATTCC-5′
5′ → 3′
Primer anneals

Extension produces a series of ddNTP terminated products each one base different in length

Each ddNTP is labeled with a different color fluorescent dye

A
AT
ATT
ATTT
ATTTA
ATTTAC
ATTTACT
ATTTACTA
ATTTACTAA
ATTTACTAAG
ATTTACTAAGG

Sequence is read by noting peak color in electropherogram (possessing single base resolution)

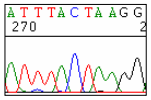


Figure 10.5, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Control Region (16024-576)

1,122 nucleotide positions

16024 16365 1 73 340

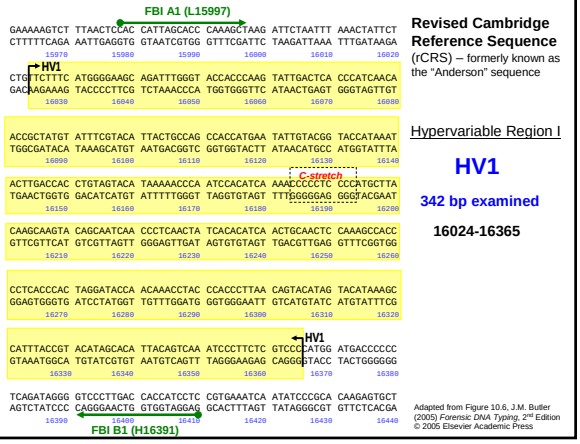
HV1 HV2

16024 **Control region (D-loop)** 576

$(AC)_3$
$(AC)_4$
$(AC)_5$
$(AC)_6$
$(AC)_7$

Forensic Focus

Typically only **610 bases examined**
- (HVI: 16024-16365; HVII: 73-340)



Revised Cambridge Reference Sequence (rCRS) – formerly known as the "Anderson" sequence

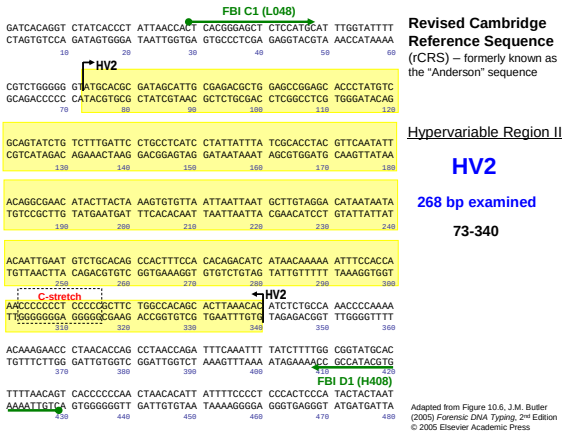
Hypervariable Region I

HV1

342 bp examined

16024-16365

Adapted from Figure 10.6, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press



Revised Cambridge Reference Sequence (rCRS) – formerly known as the "Anderson" sequence

Hypervariable Region II

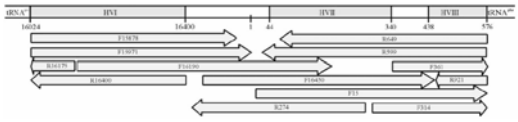
HV2

268 bp examined

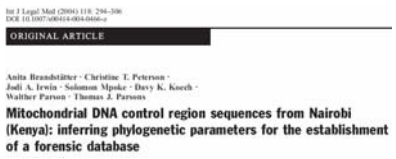
73-340

Adapted from Figure 10.6, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

AFDIL Primers Used for Control Region Amplification and Sequencing



13 primers – 11 for routine sequencing



ORIGINAL ARTICLE

Anita Brandstätter · Christine T. Peterson · Jodi A. Irwin · Solomon Mpoke · Davy K. Koech · Walther Parson · Thomas J. Parsons

Mitochondrial DNA control region sequences from Nairobi (Kenya): inferring phylogenetic parameters for the establishment of a forensic database

Why the Redundancy?

- Homopolymeric stretches of Cytosines (C-stretches).

16189T

HV1 **AAAACCCCCTCCCCATG**

5 Cs 4 Cs

16189C

AAAACCCCCCCCCCATG

10 Cs

Strand slippage can create 11+ tandem Cs

Challenges with Sequencing Beyond the Polymeric C-Stretches in HV1 and HV2

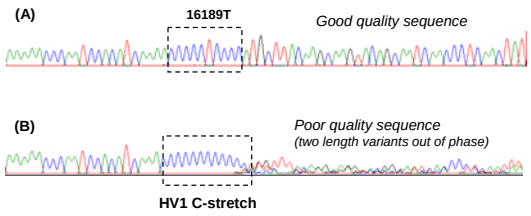
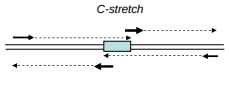


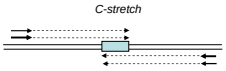
Figure 10.7, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Challenges with Sequencing Beyond the Polymeric C-Stretches in HV1 and HV2

Primer strategies typically used with C-stretch containing samples



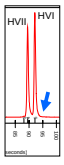
Use of internal primers

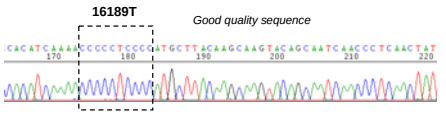


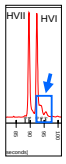
Double reactions from the same strand

Figure 10.7, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

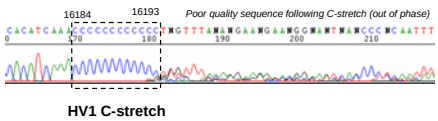
Analysis of HV1/HV2 PCR Products

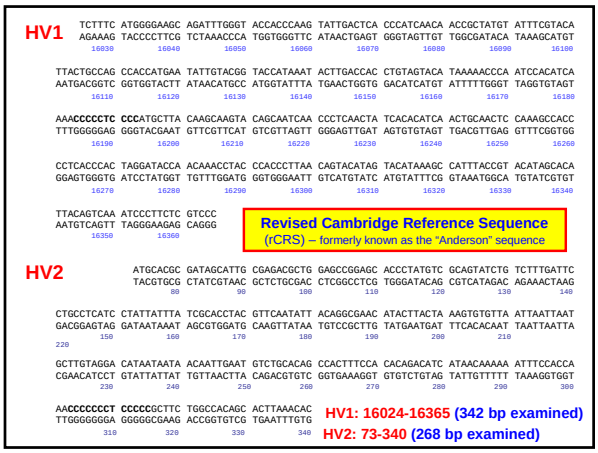






The Agilent electropherogram also gives an indication of the HVI "C-stretch" by the presence of extra peaks (see Butler *et. al.* (1998) *Electrophoresis* 19, 119-124).





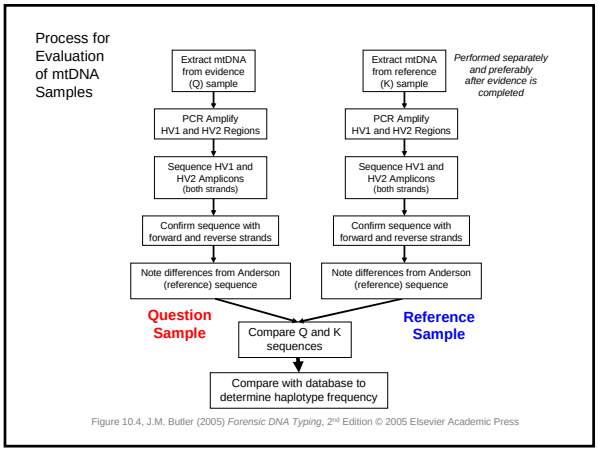
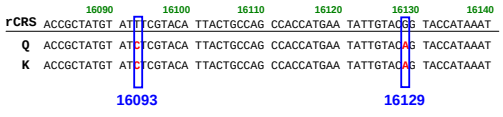


Figure 10.4, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Differences from Reference Sequence

mtDNA sequences from tested samples are aligned with the reference rCRS sequence (e.g., positions 16071-16140)



Differences are reported by the position and the nucleotide change (compared to the rCRS)

Sample Q	Sample K
16093C	16093C
16129A	16129A

Adapted from Figure 10.8, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Candidates for mtDNA Testing

- Shed hairs lacking root bulb or attached tissue
- Fragments of hair shafts.
- Aged bones or teeth that have been subjected to long periods of exposure.
- Crime scene stains or swabs that were unsuccessful for nuclear DNA testing.
- Tissues (muscle, organ, skin) that were unsuccessful for nuclear DNA testing.

Terry Melton – International Symposium on the Application of DNA Technologies in Analytical Sciences

mtDNA Testing on Hairs

- Human hair shafts contain very little DNA but because mtDNA is in higher copy number it can often be recovered and successfully analyzed
- **Melanin found in hair is a PCR inhibitor**

Important Publications:

- **Wilson, M.R., *et al.* (1995) Extraction, PCR amplification and sequencing of mitochondrial DNA from human hair shafts. Biotechniques 18(4): 662-669.**
  - **Tissue grinding method described by FBI Lab**
- **Melton *et al.* (2005) Forensic mitochondrial DNA analysis of 691 casework hairs. J. Forensic Sci. 50(1): 73-80.**
  - **Obtained a full or partial mtDNA profile for >92% of hairs tested**

The Mitotyping Experience

Terry Melton,[1] Ph.D.; Gloria Dimick,[1] M.S.; Bonnie Higgins,[1] M.S.; Lynn Lindstrom,[1,2] B.S.; and Kimberlyn Nelson,[1] Ph.D.

Forensic Mitochondrial DNA Analysis of 691 Casework Hairs*

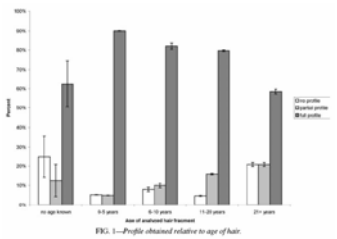


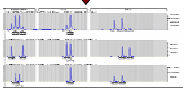
FIG. 1—*Profile obtained relative to age of hair.*

Journal of Forensic Science (2005) **50(1):** 73-80.

Nuclear DNA Analysis

Sample Collection

Laboratory



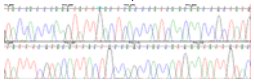
Profile generation

- 24-36 hours
- ~$100 per sample
- Use commercially available kits for processing

MtDNA Analysis

Sample Collection

Laboratory



Profile generation

- 1-6 weeks post-submission to the laboratory
- ~$1,000 per sample
- Custom designed primers

Challenges with mtDNA

- Data Interpretation
  - Heteroplasmy, Mixtures, Taq Error, and other Issues (Pseudogenes, etc…)
- DNA Database Sizes
  - Similar issues to Y-STRs but takes longer to generate mtDNA data than Y-STR haplotypes
- DNA Database Quality

Interpretational Issues - Heteroplasmy

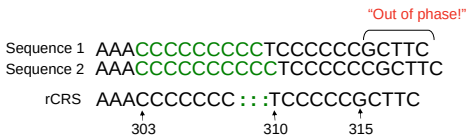
- Heteroplasmy – the presence of more than one mtDNA type in an individual (Melton 2004).
- Once thought to be rare, heteroplasmy exists (at some level) in all tissues (Melton 2004).
- Especially important in hair analysis (semi-clonal).

Some Interesting Papers on mtDNA Heteroplasmy

- Melton, T. (2004) Mitochondrial DNA heteroplasmy. *Forensic Science Reviews* **16:**1-20.
- Calloway *et al.* (2000) The frequency of heteroplasmy in the HVII region of mtDNA differs across tissue types and increases with age. *Am J Hum Genet.* **66(4):**1384-1397.
- Stewart *et al.* (2001) Length variation in HV2 of the human mitochondrial DNA control region. *Journal of Forensic Science* **46(4):**862-870.
- Sekiguchi *et al.* (2003) Inter- and intragenerational transmission of a human mitochondrial DNA heteroplasmy among 13 maternally-related individuals and differences between and within tissues in two family members. *Mitochondrion* **2(6):**401-414.
- Salas *et al.* (2001) Heteroplasmy in mtDNA and the weight of evidence in forensic mtDNA analysis: a case report. *Int J Legal Med.* **114(3):**186-190.
- Tully, L *et al.* (2000) A sensitive denaturing gradient-Gel electrophoresis assay reveals a high frequency of heteroplasmy in hypervariable region 1 of the human mtDNA control region. *Am J Hum Genet.* **67(2):**432-443.

Interpretational Issues - Heteroplasmy

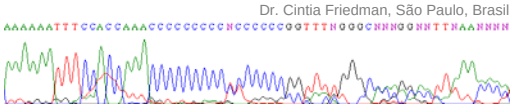
- Two types: Length (most common) and Point Heteroplasmy.



Sequence 1 has 9 Cs before 310T
Sequence 2 has 10 Cs before 310T

HV2 Length Heteroplasmy

Dr. Cintia Friedman, São Paulo, Brasil



"Out of phase!"

Sequence 1 AAACCCCCCCCCTCCCCCCGCTTC
Sequence 2 AAACCCCCCCCCCTCCCCCCGCTTC

Double coverage is important to determine sequences surrounding HV1, HV2, HV3 C-stretches.

Point Heteroplasmy

"Y" Pyrimidine

"Hotspot" for heteroplasmy

16093 (C/T)

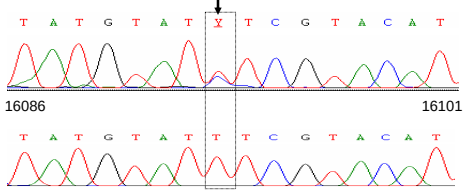
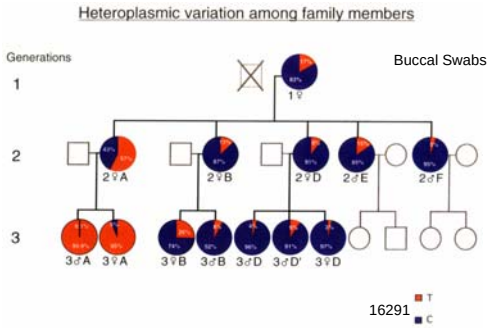


Figure 10.9, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Heteroplasmic Variation



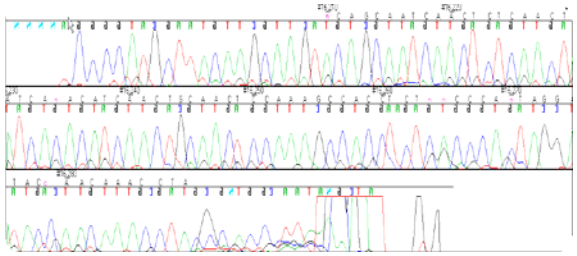
Sekiguchi *et al.* (2003)

Heteroplasmy

- Heteroplasmy can look a lot like a mixture, but is *typically* only present at one position in the CR.
- Verification and authenticity of heteroplasmy by a second extraction of the sample is required.

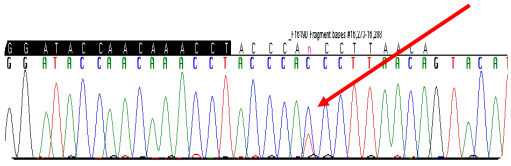
Poor Clean-Up

Poor clean-up of your amplification or sequencing product can cause pull-up, dye blobs, or other high background noise to appear in the EPGs.



Taq Error

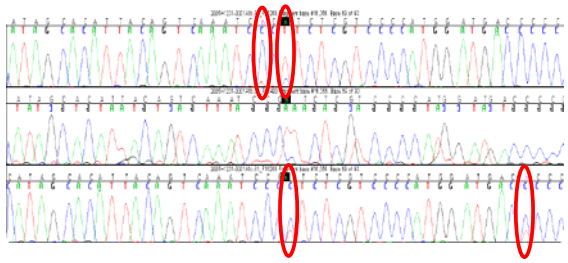
- Degraded template DNA and non-proofreading can cause mis-incorporation of bases at single positions.



- Re-amplification can correct the issue or show possible contamination.

Mixtures

- Samples can be truly mixed, either at collection, extraction or amplification.



Nuclear Pseudogenes

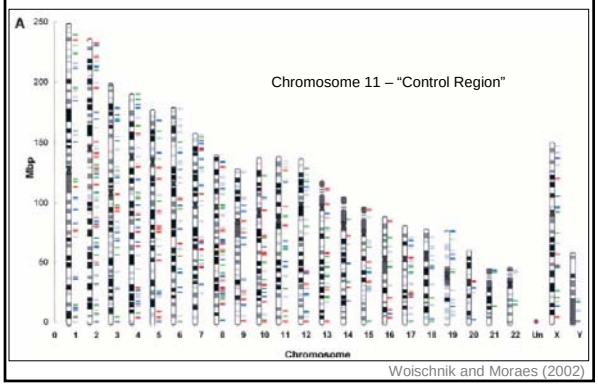
- Throughout history – movement of mtDNA genes into the nucleus.
- Nuclear Pseudogenes (nuclear-mitochondrial like sequence ***numts***) – could potentially be amplified, confounding interpretation. "Molecular Fossils"

Article — Genome Research (2002)

Pattern of Organization of Human Mitochondrial Pseudogenes in the Nuclear Genome

Markus Woischnik and Carlos T. Moraes[1]
Department of Neurology, University of Miami–School of Medicine, Miami, Florida 33136, USA

Nuclear Pseudogenes



Woischnik and Moraes (2002)

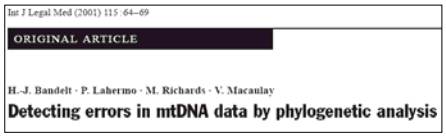
mtDNA Population Database: Size and Quality of Information

- Population databases are critical for estimating expected frequencies. The more, the better.

Database	# Profiles
African-American	1148
Afro-Caribbean	0
Sierra Leone	109
Caucasian	1655
Hispanic	686
Japan	163
Korea	182
Thailand	52
Navajo	146
Apache	180
Egypt	48
China/Taiwan	356
Guam	87
India	19 ←
Pakistan	8 ←
Total	4839

mtDNA Population Database: Size and Quality of Information

- Recently – mtDNA database quality has become an issue…


Int J Legal Med (2001) 115 :64–69

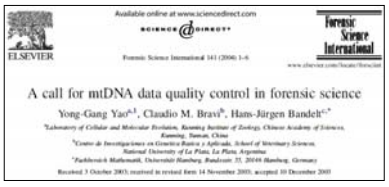
ORIGINAL ARTICLE

H.-J. Bandelt · P. Lahermo · M. Richards · V. Macaulay

Detecting errors in mtDNA data by phylogenetic analysis

Artificial recombinations and phantom mutations plague the quality of mtDNA data in population genetics, forensics, and clinical studies

mtDNA Population Database: Size and Quality of Information


Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Forensic Science International

A call for mtDNA data quality control in forensic science

Yong-Gang Yao[a,1], Claudio M. Bravi[b], Hans-Jürgen Bandelt[c,*]

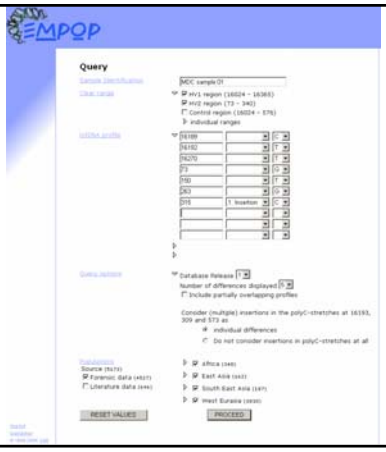
mtDNA database - USA.AFR.000942

HV1	HV2
16126-16187-16189-16223-16264 16270-16278-16293-16311-16519	73-249d-263-290d-291d 309.1C-315.1C-489
African haplogroup L1b	Asian haplogroup C1

Walther Parson
Institute of Legal Medicine, Innsbruck Medical University

Version 1.0 is online (16 October 2006)

http://www.empop.org



Reporting Statistics

- When “cannot exclude” is the interpretation, then a statistical estimate is needed in order to weigh the significance of the observed match
- Counting method is most common approach used and involves counting the number of times that a particular mtDNA haplotype (sequence) is seen in a database
- The larger the number of unrelated individuals in the database, the better the statistics will be for a random match frequency estimate.

Example Calculation of mtDNA Profile Frequency Estimate

The frequency (p) of observing a mtDNA profile (X) times in a database having a size of (N) is…

$p = X/N$

$p = 70/1665 = 0.042$

A 95% confidence interval can be determined by using a normal approximation of the binomial

$$p \pm 1.96 \sqrt{\frac{(p)\,(1-p)}{N}}$$

Upper Bound

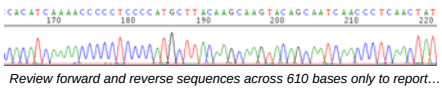
$$p \pm 1.96 \sqrt{\frac{(0.042)\,(0.958)}{1665}} = p \pm 1.96\,(0.0049) = 0.032 \text{ and } 0.052$$

Holland and Parsons (1999) *Forensic Sci. Rev.*

Tools for mtDNA Screening

Disadvantages to Sequencing

- Expensive
  - Primarily due to intensive labor in data analysis
- Error possibilities with more data to review
- Most information is not used



Review forward and reverse sequences across 610 bases only to report…

263G, 315.1C **Most common type: found in ~7% of Caucasians…**

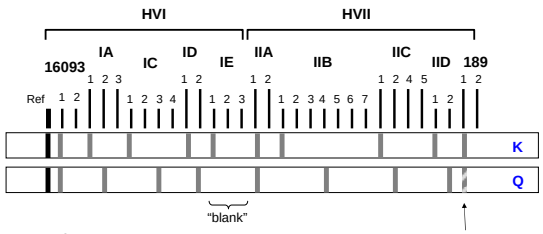
Advantages to Screening Methods

- Rapid results
- Aids in exclusion of non-matching samples
- Less labor intensive
- Usually less expensive
- Permits more labs to get involved in mtDNA

Screening assays are essentially a presumptive test prior to final confirmatory DNA sequencing.

Sequencing is necessary to certify that every position matches between a question and a known sample.

LINEAR ARRAY mtDNA Typing Strips: New Screening Method



Reported Types

K: 1-1-1-1-1-1-1-1-1

Q: 1-2-3-2-0-1-4-2-2-*w*1

If known (K) and question (Q) samples do not match, there is no need to involve the expense of mtDNA sequencing

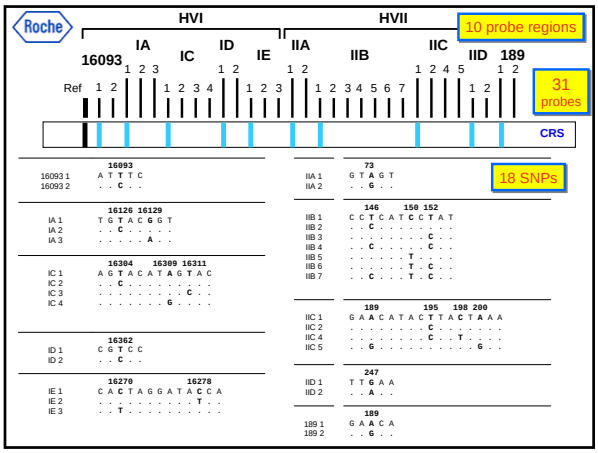
Figure 10.10, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

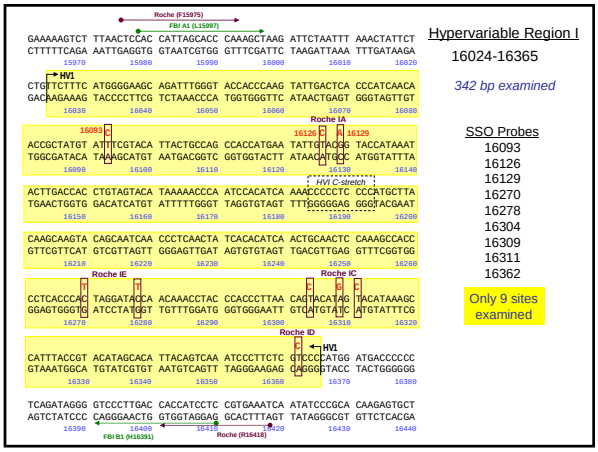
Methodologies for SNP Typing

High-tech
- **SNaPshot (minisequencing)**
- Luminex 100 allele-specific hybridization
- Pyrosequencing
- TaqMan
- Primer extension with time-of-flight mass spectrometry
- TagArray (SNPstream UHT)
- Affymetrix hybridization chip

Low tech
- **Reverse dot blot (LINEAR ARRAYs)**
- PCR-RFLP
- Allele-specific PCR

See Budowle *et al.* (2004) *Forensic Sci. Rev.* 16:21-36 for a review of some SNP typing technologies





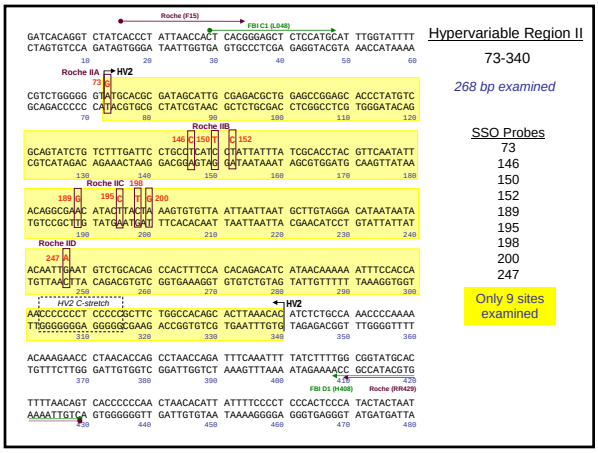
Hypervariable Region I

16024-16365

342 bp examined

SSO Probes
16093
16126
16129
16270
16278
16304
16309
16311
16362

Only 9 sites examined



Hypervariable Region II

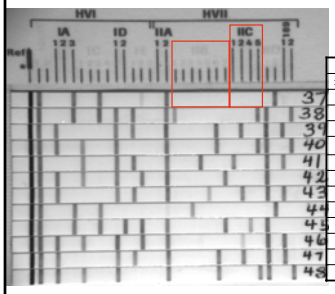
73-340

268 bp examined

SSO Probes
73
146
150
152
189
195
198
200
247

Only 9 sites examined

Data Interpretation for LINEAR ARRAYs



Analysis of probe results is still manual!

HVI						HVII			
16093	A	C	D	E	A	B	C	D	189
1	3	3	0	w2	2	0	0	2	0
1	3	3	1	1	2	1	5	1	2
1	1	0	1	1	2	6	2	1	1
1	1	1	1	0	2	6	5	1	2
1	0	1	1	1	2	4	1	1	1
1	1	1	1	2	2	7	0	1	0
1	1	0	1	1	2	5	2	1	1
1	3	3	0	0	w2	3	0	2	0
1	3	3	2	2	2	7	4	1	0
1	1	1	1	2	2	0	w2	1	1
1	1	0	1	1	2	6	2	1	1
1	1	1	1	0	2	6	5	1	2

Typing results from 50 ng of each PCR product

Margaret Kline, NIST

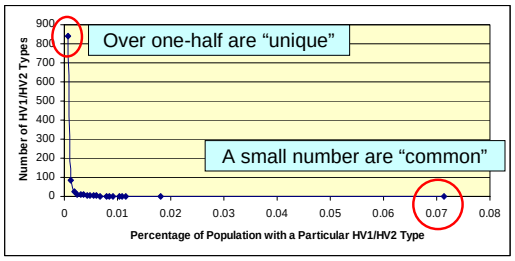
Sequencing – Increased Discrimination

# times haplotype observed	LINEAR ARRAY	HV1	HV1+HV2	control region	
1	**185**	**334**	**454**	**502**	
2	45	38	36	37	
3	18	11	11	10	
4	4	7	4	1	
5	4	4	6	5	
6	3	5	2	3	
7	1	7	3	·	
8	9	·	1	·	
9	2	1	·	·	
10	4	1	·	**1**	
11	1	2	·	·	
12	1	1	·	·	
17	·	·	**1**	·	
18	1	·	·	·	
23	1	·	·	·	
28	1	·	·	·	
40	·	**1**	·	·	
51	**1**	·	·	·	
HD	0.9869	0.9936	0.9982	0.9990	
% DC	42.19%	61.86%	77.78%	83.93%	total
# HT	**281**	**412**	**518**	**559**	**666**

Coble et al. (in review)

mtDNA Population Distribution

Caucasians (n=1665)

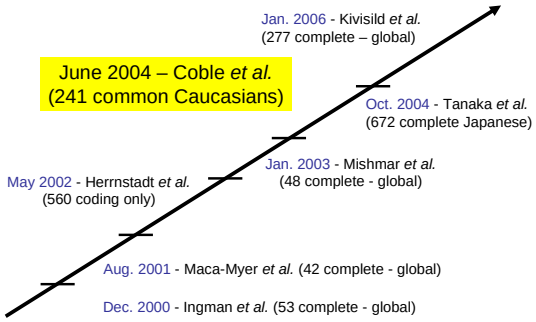


Coble *et al.* (2004)

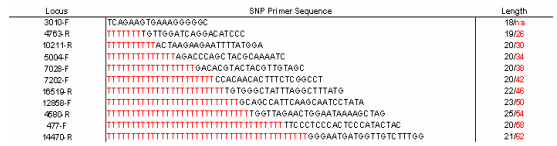
The Problem of Common mtDNA Types

- The greatest limitation for mtDNA testing lies with the small number of common types for which the power of discrimination is low.
- ~20% of the time, the Forensic Scientist encounters a HV1/HV2 type that occurs at greater than ~0.5% of the population.
- In database or mass fatality comparisons: multiple hits will occur for these common types.

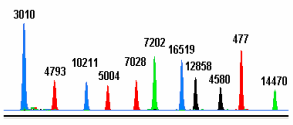
mtGenomics – Whole mtDNA Sequencing



The SNaPShot™ Platform



Rebecca Just
Dr. Peter Vallone



Vallone *et al.* (2004) *IJLM* **118:** 147- 157.

Publications

Michael D. Coble · Rebecca S. Just
Jennifer E. O'Callaghan · Bona H. Letmanyi
Christine T. Peterson · Jodi A. Irwin · Thomas J. Parsons

Single nucleotide polymorphisms over the entire mtDNA genome that increase the power of forensic testing in Caucasians

IJLM (2004) **118:** 137-146

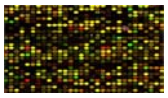
Peter M. Vallone · Rebecca S. Just · Michael D. Coble
John M. Butler · Thomas J. Parsons

A multiplex allele-specific primer extension assay for forensically informative SNPs distributed throughout the mitochondrial genome

IJLM (2004) **118:** 147- 157

http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

Steps in Running the Affymetrix Resequencing Array



Day 1

1. PCR
2. Desalting and Reaction Clean Up
3. Quantitation
4. Pooling
5. Fragment
6. Label
7. Hybridize

Day 2

8. Wash/Stain
9. Scan
10. Data Analysis

Dr. Peter Vallone, NIST – Presentation at DNA in Forensics, Innsbruck, Austria (Sept. 06)

Summary

- mtDNA is useful in forensic situations with limited or highly degraded DNA due to its high copy number
- Forensic applications typically examine 610 bp from the control region (HV1 & HV2)
- mtDNA sequencing is labor-intensive, but some screening methods are now available