National Forensic Science Technology Center

Mitochondrial DNA Workshop

Michael D. Coble, PhD

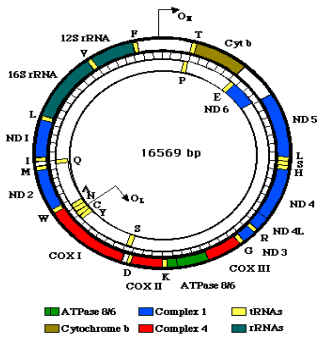
March 13-15, 2006

Goals and Objectives

- Overview and theory behind mtDNA analysis
- The science behind mtDNA sequencing
- Forensic casework applications of mtDNA (validation and examples)
- Tools for mtDNA screening – Linear Arrays
- Emerging mtDNA technologies – mtDNA genome sequencing, species identification, dHPLC for resolving mixtures.
- Summary and **Questions**

Terry Melton – 2005 Forensic e-symposium

Review



Coding Region
13 polypeptides
2 rRNAs
13 tRNAs

All necessary for OXPHOS

Highly compact (few intergenic spaces)

http://www.mitomap.org/

mtDNA as a Forensic Tool

Advantages of Using mtDNA

- Maternal Inheritance
- Lack of Recombination
- High Copy Number
- Cases where:
  - DNA is degraded
  - Only maternal references are available
  - Samples with little or no Nuclear DNA
    - » Shed hairs
    - » Fingernails
    - » Old bones

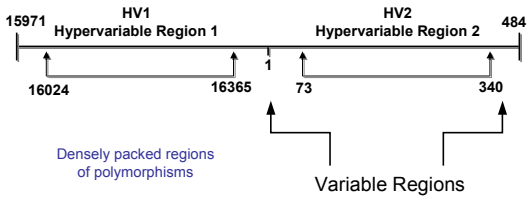
mtDNA as a Forensic Tool

Disadvantages of Using mtDNA

- Maternal Inheritance – You have many!
- Not a unique identifier – cannot multiply frequencies of polymorphisms – mtDNA is one linked marker.
- Some mtDNA types are common in the population.

Control Region (16024-576)

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



Emerging mtDNA Technologies

- Screening methods with SNPs.
- Increased discrimination using coding region sequence information.
- Non-Sequencing Strategy (Affymetrix mtDNA Chip).
- Species Identification.
- Low Copy nucDNA for high throughput screening.

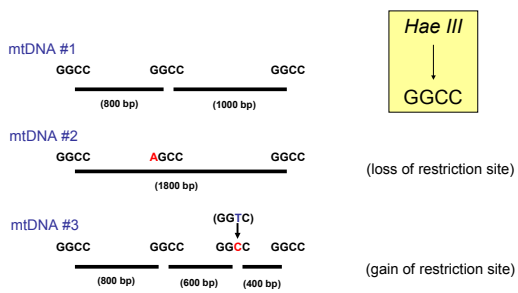
A Brief Sidestep…

- mtDNA as a genetic tool… "mitogenomics"
- The lack of apparent recombination, and high mutation rate make mtDNA an excellent tool for studying human evolution.
- Some of these insights have been useful for the mtDNA forensic scientist.

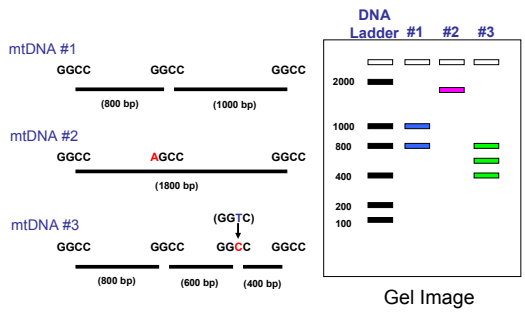
Methods for Measuring mtDNA Variation

- 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)

RFLP Analysis

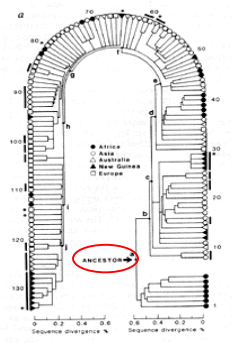


RFLP Analysis



Gel Image

mtDNA as a Genetic Marker

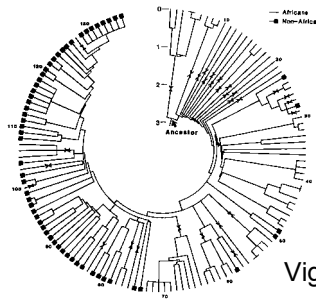


Cann *et al.* 1987

High res. RFLP analysis of 134 mtDNA types from 148 individuals ~ 370 different restriction sites per individual.

"Out of Africa"

mtDNA as a Genetic Marker



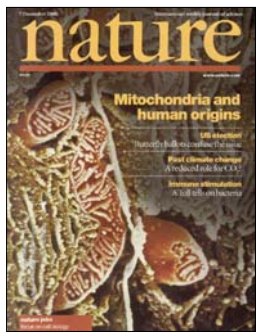
Control Region Sequence Analysis of 189 individuals

Vigilant *et al.* 1991

mtDNA as a Genetic Marker

- Templeton (1992) *Science* – Found phylogenetic trees that were more parsimonious than Vigilant *et al.* **AND** these trees did not suggest an "Out of African" origin.
- More sequence data and better tree-building methods confirmed the OOA hypothesis (Penny *et al.* 1995; Watson *et al.* 1997)

mtDNA as a Genetic Marker



Ingman *et al.* (2000)

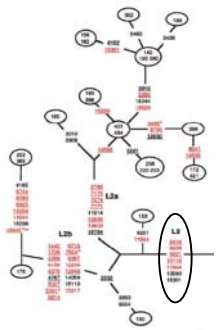
53 entire genome sequences from diverse global populations.

Confirmation for OOA.

mtDNA as a Genetic Marker

- RFLP variation has revealed continent-specific polymorphisms for classifying mtDNAs.
- Haplotype – the mtDNA sequence variations within an individual (e.g. your HV1/HV2 type).
- Haplogroup – a group of related haplotypes. These form monophyletic clades on a phylogenetic tree.

mtDNA Haplogroups

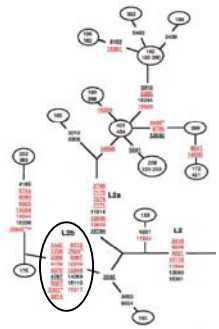


Each haplogroup cluster is defined by a set of ***specific, shared*** polymorphisms.

In this cluster, all individuals belonging to the African L2 haplogroup share a set of 7 SNPs in the coding region.

Herrnstadt et al. AJHG (2002)

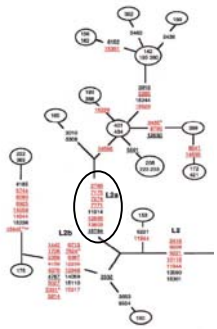
mtDNA Haplogroups



In this cluster, all individuals belonging to the African L2b sub-haplogroup share a set of 17 SNPs in the coding region.

Herrnstadt et al. AJHG (2002)

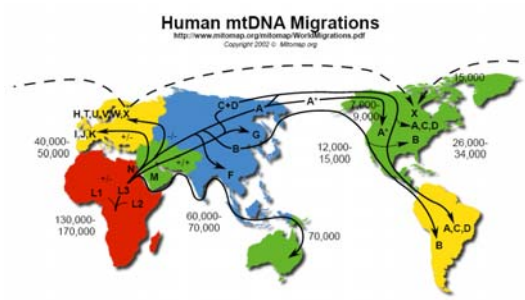
mtDNA Haplogroups



In this cluster, all individuals belonging to the African L2a sub-haplogroup share a set of 8 SNPs in the coding region.

Herrnstadt et al. AJHG (2002)

Human Migration Model



mtDNA Haplogroups (HV1/HV2)

- J - 16069 C-T 16126 T-C 73 A-G 295 C-T
- T - 16126 T-C 16294 C-T 73 A-G
- V - 16298 T-C 72 T-C
- L3e3 - 16223 C-T 16265 A-C 73 A-G 150 C-T 195 T-C

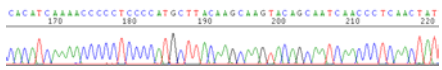
Generally, very good concordance between CR and coding haplogroups

Macaulay *et al.* (1999) *AJHG* **64:** 232-249
Allard *et al.* (2002) *JFS* **47:** 1215-1223
Brandstatter *et al.* (2004) *IJLM* **118:** 294-306

Screening Assays for mtDNA Typing

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.

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

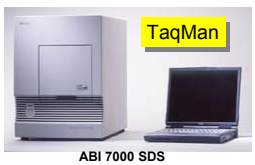
See Budowle *et al.* (2004) *Forensic Sci. Rev.* 16:21-36 for a review of some SNP typing technologies

SNP Typing Instrumentation

Multi-Color Capillary Electrophoresis (ABI 310 or 3100)

PCR & primer extension

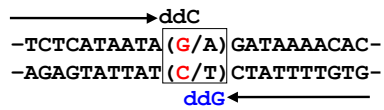
TaqMan



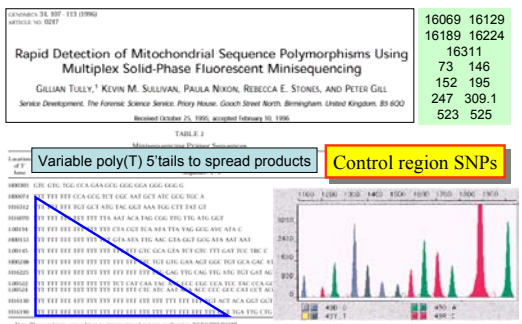
ABI 7000 SDS

SNP Extension Primer Design

- Must anneal to DNA template with 3'end of primer next to SNP site
- Can anneal to either top strand or bottom strand
- Should have uniform annealing temperature (by lengthening 5'end of SNP primer)
- Should not form significant hairpins or dimers with other SNP or PCR primers



Early Multiplex SNP Detection Work

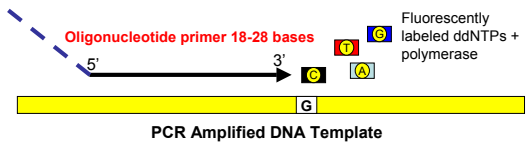


Allele-Specific Primer Extension

SNP Primer is extended by one base unit

ABI PRISM® SNaPshot™ Multiplex System

"tail" used to vary electrophoretic mobility



Products can be electrophoretically separated on an ABI 310, 3100

Allele-Specific Primer Extension

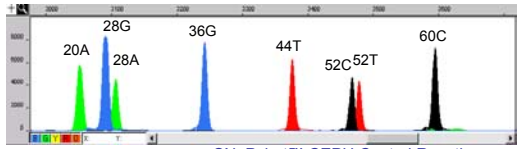
The use of "tailed" SNP primers allows for multiplexing in the SNaPshot assay

Sequences for 11 SNP primers

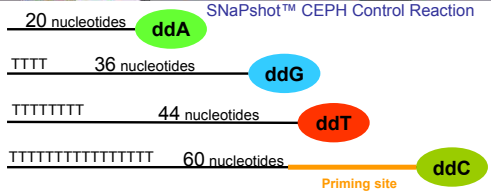
TGTTGGATCAGGACATCCC	19/19
TTTTTCAGAAGTGAAAGGGGGC	18/22
TTTTTTTTTTACTAAGAAGAATTTTATGGA	20/30
TTTTTTTTTTTTTAGACCCAGCTACGCAAAATC	20/34
TTTTTTTTTTTTTTTTTTTTGACACGTACTACGTTGTAGC	20/38
TTTTTTTTTTTTTTTTTTTTTTTCCACAACACTTTCTCGGCCT	20/42
TTTTTTTTTTTTTTTTTTTTTTTTTTGTGGGCTATTTAGGCTTTATG	22/46
TTTTTTTTTTTTTTTTTTTTTTTTTTTTGCAGCCATTCAAGCAATCCTATA	23/50
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTAGAACTGGAATAAAAGCTAG	25/54
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCCTCCCACTCCCATACTAC	20/58
TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAATGATGGTTGTCTTTGG	21/62

Template binding sequence – black
Tailed sequence for fragment separation - blue

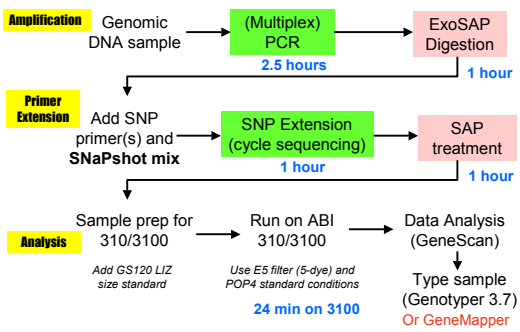
Allele-Specific Primer Extension



SNaPshot™ CEPH Control Reaction



Protocol with SNaPshot™ "Kit"



Use of Haplogroup Defining mtSNPs

Int J Legal Med (2003) 117: 291–298
DOI 10.1007/s00414-003-0395-2

ORIGINAL ARTICLE

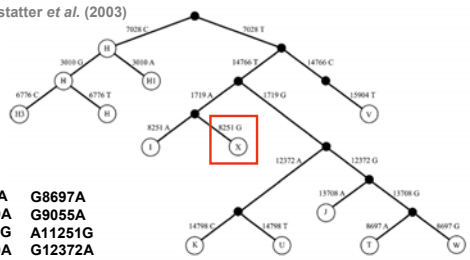
Anita Brandstätter · Thomas J. Parsons · Walther Parson

Rapid screening of mtDNA coding region SNPs for the identification of west European Caucasian haplogroups

Identifies coding region SNP to classify the 9 major Western European haplogroups (plus 2 sub-haplogroups).

Use of Haplogroup Defining mtSNPs

Brandstatter *et al.* (2003)



G709A	G8697A
G1719A	G9055A
A1811G	A11251G
G3010A	G12372A
T6365C	G13708A
T6776C	C14766T
C7028T	T14798C
G8251A	C15904T

16 mtSNPs run in two SNaPshot 8plex reactions

Use of Haplogroup Defining mtSNPs

Haplogroups Defined by Control Region mtSNPs in SWGDAM Caucasian Samples (n = 1771)

Allard *et al.* (2002) *J. Forensic Sci.* 47(6):1215-1223

- **H** (46%): 73A
- **U** (15.6%): 16270T
- **T** (10.5%): 16126C, 16294T
- **J** (10%): 16069T, 16126C, 295T
- **K** (8.9%): 16224C, 16311C
- **I** (2%): 16223T, 199C, 204C, 250C
- **V** (1.9%): 16298C, 72C
- **W** (1.9%): 16223T, 189G, 195C, 204C, 207A
- **X** (1.6%): 16189C, 16223T, 16278T, 195C
- **M** (1.9%): 16223T, 16298C

If a **G is observed at 8251**, then the sample can be classified as a member of **haplogroup X** so the following control region SNPs should be expected: 16189C, 16223T, 16278T, 195C

Use of Haplogroup Defining mtSNPs

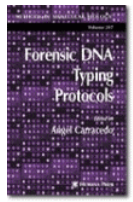
Advantages

Sensitive – 1 pg genomic DNA
Short amplicons – degraded DNA
Multiplexed PCR – conserves template
POD – 88.6% among 277 unrelated Austrian Caucasians

Disadvantages

No "kit" – need to order, validate primers.
Probability of random match among Caucasians is ~11.4%

Other SNP Assays



Methods in Molecular Biology
Volume: 297
(2005)

SNaPshot Typing of Mitochondrial DNA Coding Region Variants

Antonio Salas
Beatriz Quintáns
Vanesa Álvarez-Iglesias

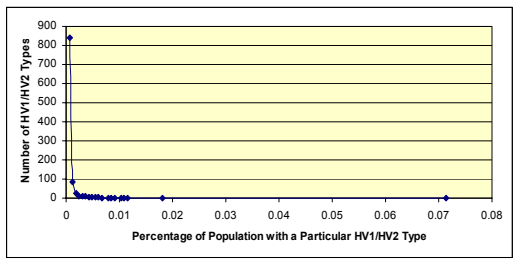
Typing of mitochondrial DNA coding region SNPs of forensic and anthropological interest using SNaPshot minisequencing.

Quintáns et al. (2004) *Forensic Sci Int.*

Increased discrimination using coding region sequence information

mtDNA Population Distribution

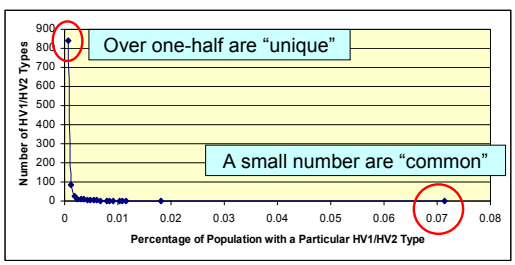
Caucasians (n=1665)



Coble *et al.* (2004)

mtDNA Population Distribution

Caucasians (n=1665)



Coble *et al.* (2004)

Framing the Problem

- 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.

A Case Example

- September 15, 1943 - B17F Bomber returning from a mission to Port Moresby, New Guinea

A Case Example

- The plane crashes in the Owen Stanley Mountain range due to “adverse weather.”
- Subsequent searches proved negative.
- 11 crewmen declared non-recoverable on July 22, 1949.

A Case Example

- October 9, 1992 - A private company helicopter discovers crash site.
- mtDNA testing reveals that 3/11 crewmen share the same HV type (263 A-G, 315.1 C).
- Further VR testing could distinguish 1 of the 3 crewmen (16519 T-C). However, 2 crewmen still matched.

A Case Example

- Partial dental records were used to associate 3 teeth among the 2 crewmen matching in the CR.
- One L femur could not be associated with either crewmen, and was buried in a grave containing group remains

Strategy for SNP Identification

- Sequence the entire genome of unrelated individuals sharing common HV1/HV2 types in the Caucasian population (focus on 18 of 22 common types that occur at a frequency of 0.5% or greater).

Ethical Considerations

- ~265 characterized diseases associated with mtDNA mutations in the coding region (Mitomap – www.mitomap.org)
- To avoid having forensic testing from evolving into genetic counseling, we decided to focus on neutral SNPs in the mtGenome.

SNPs for Discrimination

- Non-coding sites in the control region (outside of HV1/HV2).
- Non-coding "spacer" regions throughout the mtGenome.
- Silent mutations in protein coding genes.

SNPs for Discrimination

- Practical application – A set of SNP sites that can be rapidly assayed to provide maximal discrimination.
- Avoids further sequencing.
- Allele Specific Primer Extension – small amplicons, multiplexed - can conserve template, run on standard instrumentation.

Common mtDNA Haplogroups

Com	Haplo	Seq (+ CRS)
31	H1	CRS
25	H2	152 C
11	H3	16129 A
8	H4	16263 C
12	H5	16304 C
11	H6	73 G
7	H7	16162 G 16209 C 73 G

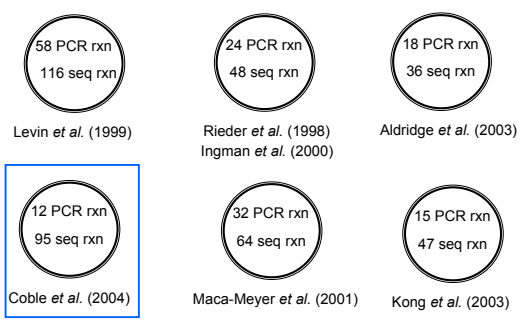
Length Variation in HV2 C-stretch – ignored (Stewart *et al.* (2001)

Common mtDNA Haplogroups

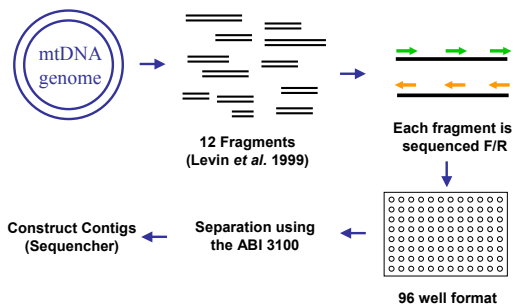
Com	Haplo	Seq (+ CRS)
15	J1	16069 T 16126 C 73 G 185 A 228 A 295 T
6	J2	16069 T 16126 C 73 G 228 A 295 T
12	J3	16069 T 16126 C 73 G 185 A 188G 228 A 295 T
3	J4	16069 T 16126 C 16145 A 16172 C 16222 T 16261 T 73 G 242 T 295 T
20	T1	16126 C 16294 T 16296 T 16304 C 73 G
10	T2	16126 C 16163 G 16186 T 16189 C 16294 T 73 G 152 C 195 C
8	T3	16126 C 16294 T 16296 T 73 G
25	V1	16298 C
14	K1	16224 C 16311 C 73 G 146 C 152 C
8	K2	16093 C 16224 C 16311 C 73 G
7	K3	16224 C 16311 C 73 G

241 total genomes from 18 common HV1/HV2 types (~14% of the total database)

Strategies for Whole mtGenome Analysis



Sequencing Strategy



The Nature of the SNPs

- Would the SNPs that resolve one group be useful for resolving other closely related groups?

Com	Haplo	Seq (+ CRS)
31	H1	CRS
25	H2	152 C
11	H3	16129 A
8	H4	16263 C
12	H5	16304 C
11	H6	73 G
7	H7	16162 G 16209 C 73 G

"Hot Spots"

The Nature of the SNPs

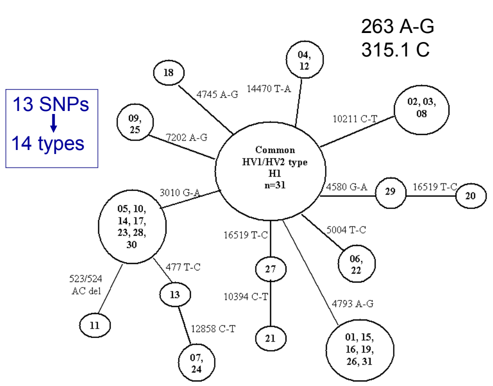
- Are resolving SNPs **slow and rare?** Did these SNPs arise once during the evolution of a haplogroup?

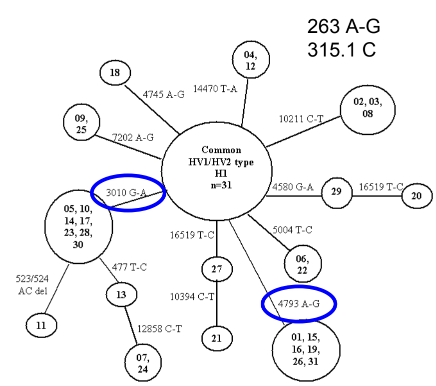
OR…

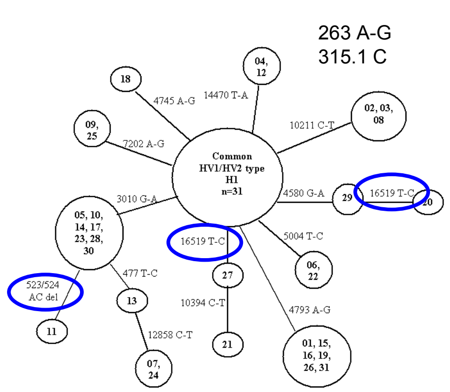
- Are resolving SNPs "universally" **fast hot spots**, useful for all haplogroups (L, M, N)?

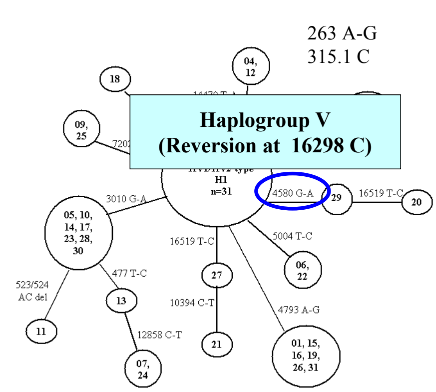
OR….

- Are resolving SNPs a combination of the two?

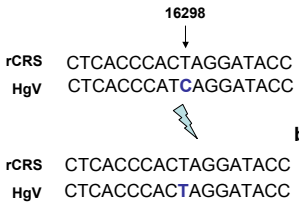






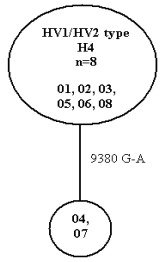


Reversion

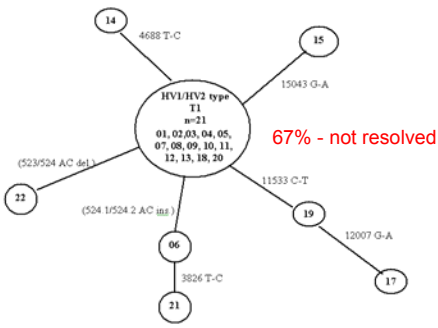


Sequence with HgV background…appears as a HgH (H1)

H4 - rCRS + 16263 T-C



T1 – Low Resolution

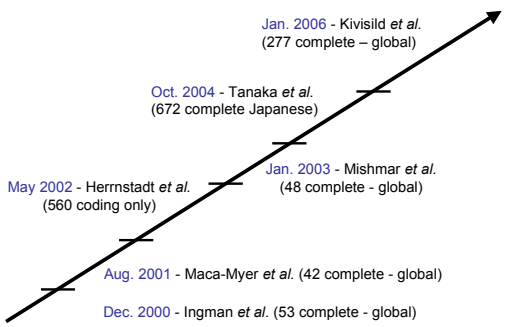


67% - not resolved

Brute-Force Sequencing

- Why not used information from the literature??
- Prior to 1999, only a handful of whole genome sequences in GenBank. Most of the mtDNA coding region data was from RFLP studies (assays ~ 20% of the genome)

mtGenomics

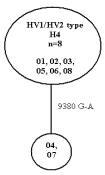


mtDB - Human Mitochondrial Genome Database

- Max Ingman (Uppsala University, Sweden)
- 1622 complete sequences and 839 coding region sequences.
- 2461 coding region sequences.

mtDB - Human Mitochondrial Genome Database

Anderson		2461 Sequences							
Posn.	Base	A	G	C	T	Gap	Location	Codon	Position
9380	G	13	2448				coIII	58	3



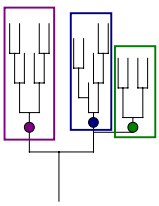
9380 G-A has only been observed in 10/2215 (0.45%) coding regions… would not be a good candidate if one was "trolling" for discriminating SNPs

Problem – very few common types in global DB

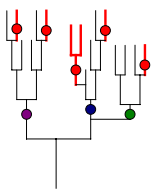
Summary

- 241 mtGenomes – 420 polymorphic sites in the coding region.
- 32/241 (13%) – matched one or more individuals over the entire mtGenome (0/12 H5 individuals matched; 4/8 H7 individuals matched).
- Homoplasies – common in HV1/HV2.

Homoplasy – Parallel Substitutions



3 basal polymorphisms that define3 clusters



"Red" mutation has occurred multiple times on the tree

Summary

- Percentage of sites that varied ranged from 1.0% (16S rRNA) to 6.6% (non-coding regions outside of the control region).
- ATP Synthase 8 (4.8%) and ATP Synthase 6 (3.7%) showed the greatest non-synonymous variation in the protein coding genes.

Gene	Length	Synonymous	Nonsynonymous	Total	% NonSyn.
ND1	956	14	8	22	36.4%
ND2	1,042	25	11	36	30.6%
CO1	1,542	29	9	38	23.7%
CO2	684	14	4	18	22.2%
ATP8	207	3	5	8	62.5%
ATP6	681	7	20	27	74.1%
CO3	784	14	4	18	22.2%
ND3	346	5	2	7	28.6%
ND4L	297	5	1	6	16.7%
ND4	1,378	30	7	37	18.9%
ND5	1,812	39	15	54	27.8%
ND6	525	8	7	15	46.7%
CYB	1,141	23	15	38	39.5%
Total	11,341	216	108	324	33.1%

SNPs for Forensic Discrimination

- 59 SNPs – that met our criteria (neutral, shared, non-redundant).
  - 49 – Protein coding (silent)
  - 8 – Control Region (outside HV1/2)
  - 1 – Non-coding spacer region
  - 1 – 16S rRNA*

* 3010 G-A

SNPs for Forensic Discrimination

A	B	C	D	E	F	G	H
477	477	72	482	4808	64	3826	64
3010	3010	513	5198	5147	4745	3834	4688
4580	3915	4580	6260	9380	10211	4688	11377
4793	5004	5250	9548	9899	10394	6293	12795
5004	6776	11719	9635	11914	10685	7891	13293
7028	8592	12438	11485	15067	11377	11533	14305
7202	10394	12810	11914	16519	14470	12007	16519
10211	10754	14770	15355		14560	12795	
12858	11864	15833	15884		16390	15043	
14470	15340	15884	16368		14869	16390	
16519	16519	16519				16519	
H1	H2 H3 H6	V1 H5	J1 J2 K2 K3	J4 T2 T3 H4	V1 H1 H2 H3	J1 J3 T1	K1

SNPs for Forensic Discrimination

A	B	C	D	E	F	G	H
477	477	72	482	4808	64	3826	64
3010	3010	513	5198	5147	4745	3834	4688
4580	3915	4580	6260	9380	10211	4688	11377
4793	5004	5250	9548	9899	10394	6293	12795
5004	6776	11719	9635	11914	10685	7891	13293
7028	8592	12438	11485	15067	11377	11533	14305
7202	10394	12810	11914	16519	14470	12007	16519
10211	10754	14770	15355		14560	12795	
12858	11864	15833	15884		16390	15043	
14470	15340	15884	16368		14869	16390	
16519	16519	16519				16519	
H1	H2 H3 ~~H6~~	V1 H5	J1 J2 K2 ~~K3~~	J4 T2 T3 ~~H4~~	V1 H1 ~~H2 H3~~	J1 J3 T1	K1

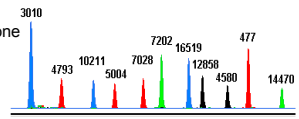
SNPs for Forensic Discrimination

A	B	C	D	E	F	G	H
477	477	72	482	4808	64	3826	64
3010	3010	513	5198	5147	4745	3834	4688
4580	3915	4580	6260	9380	10211	4688	11377
4793	5004	5250	9548	9899	10394	6293	12795
5004	6776	11719	9635	11914	10685	7891	13293
7028	8592	12438	11485	15067	11377	11533	14305
7202	10394	12810	11914	16519	14470	12007	16519
10211	10754	14770	15355		14560	12795	
12858	11864	15833	15884		16390	15043	
14470	15340	15884	16368		14869	16390	
16519	16519	16519				16519	
H1	H2 H3 ~~H6~~	V1 H5	J1 J2 K2 ~~K3~~	J4 T2 T3 ~~H4~~	V1 H1 ~~H2 H3~~	J1 J3 T1	K1

The SNaPShot™ Platform

Locus	SNP Primer Sequence	Length
3010-F	TCAGAAGTGAAAGGGGGC	18/n.s
4793-R	TTTTTTTTGTTGGATCAGGACATCCC	19/26
10211-R	TTTTTTTTTTACTAAGAAGAATTTTATGGA	20/30
5004-F	TTTTTTTTTTTTTTAGACCCAGCTACGCAAAATC	20/34
7028-F	TTTTTTTTTTTTTTTTTTGACACGTACTACGTTGTAGC	20/38
7202-F	TTTTTTTTTTTTTTTTTTTTTTCCACAACACTTTCTCGCCT	20/42
16519-R	TTTTTTTTTTTTTTTTTTTTTTTTGTGGGCTATTTAGGCTTTATG	22/46
12858-F	TTTTTTTTTTTTTTTTTTTTTTTTTTTTGCAGCCATTCAAGCAATCCTATA	23/50
4580-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGTTAGAACTGGAATAAAAGCTAG	25/54
477-F	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCCCTCCCACTCCCATACTAC	20/58
14470-R	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGGAATGATGGTTGTCTTTGG	21/62

Rebecca Just
Dr. Peter Vallone



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

SNPs for Forensic Discrimination

18 common HV1/HV2 types, 241 individuals

+8 Multiplexes (59 SNPs)
105 types (55 "unique")

+8 Multiplexes (with AC indel)
112 types (64 "unique")

6-fold improvement!

The Nature of the SNPs

- Are resolving SNPs **slow and rare?** Did these SNPs arise once during the evolution of a haplogroup?

OR…

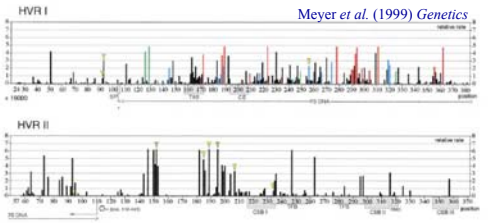
- Are resolving SNPs "universally" **fast hot spots**, useful for all haplogroups (L, M, N)?

SNPs for Forensic Discrimination

A	B	C	D	E	F	G	H
477	477	72	482	4808	64	3826	64
3010	3010	513	5198	5147	4745	3834	4688
4580	3915	4580	6260	9380	10211	4688	11377
4793	5004	5250	9548	9899	10394	6293	12795
5004	6776	11719	9635	11914	10685	7891	13293
7028	8592	12438	11485	15067	11377	11533	14305
7202	10394	12810	11914	16519	14470	12007	16519
10211	10754	14770	15355		14560	12795	
12858	11864	15833	15884		16390	15043	
14470	15340	15884	16368		14869	16390	
16519	16519	16519				16519	
H1	H2 H3 ~~H6~~	V1 H5	J1 J2 K2 ~~K3~~	J4 T2 T3 ~~H4~~	V1 H1 ~~H2 H3~~	J1 J3 T1	K1

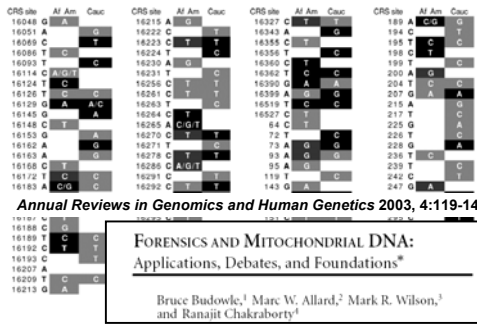
Mutation Rate Analysis in the mtDNA Control Region

Mutation rate heterogeneity – the variation of mutation rates among sites.



Meyer *et al.* (1999) *Genetics*

Mutation Rate Analysis in the mtDNA Control Region



***Annual Reviews in Genomics and Human Genetics* 2003, 4:119-141**

FORENSICS AND MITOCHONDRIAL DNA:
Applications, Debates, and Foundations*

Bruce Budowle,[1] Marc W. Allard,[2] Mark R. Wilson,[3] and Ranajit Chakraborty[4]

Mutation Rate Analysis in the mtDNA Coding Region

Previous Assumptions (I)

Adam Eyre-Walker *et al.* (1999) *Proc. R. Soc. Lond B*. Using partial DNA sequences of the human mtDNA genome (filled with errors), this group observed a significant amount of recurrent mutations (homoplasy) in their data.

Conclusion – **Recombination!** (between paternal and maternal mtDNA)

Mutation Rate Analysis in the mtDNA Coding Region

- Eyre-Walker *et al.* assume mutation rate ***Homo***geneity…
- "There is no evidence of variation in the mutation rate."
- (Mostly discredited for their poor data choice and method of calculating LD)

Mutation Rate Analysis in the mtDNA Coding Region

Previous Assumptions (II)

Herrnstadt *et al.* (2002) *AJHG* – 560 coding region sequences.

"One important result to emerge from these studies is the ***relatively large number of sites*** at which ***homoplasic events*** have occurred."

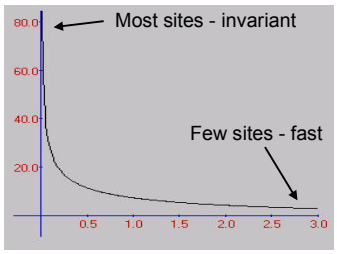
(Referring to their Table 2)

Mutation Rate Analysis in the mtDNA Coding Region

- Yao *et al.* (2003) *AJHG* – in response to an Amerindian paper filled with sequence errors.
- "***Homoplasy*** in the coding region is ***much less*** than in the control region and may have ***only a few*** hot spots (see, e.g., table 2 of Herrnstadt *et al.* [2002])"

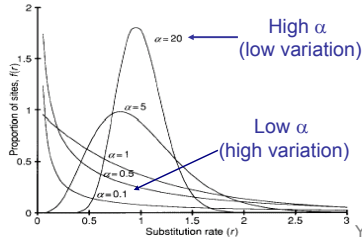
How is Rate Variation Measured?

- Control region rates follow a negative binomial distribution (gamma distribution).



How is Rate Variation Measured?

- The SHAPE of the curve (α) is inversely related to the amount of heterogeneity



Yang, 1996

Methods to Determine α

- Parsimony analysis of phylogenetic trees (646 coding region sequences).
- Count the number of character changes mapped upon the MPT to determine the relative mutation rate.
- Calculate the α parameter using the method of Yang and Kumar (1996).

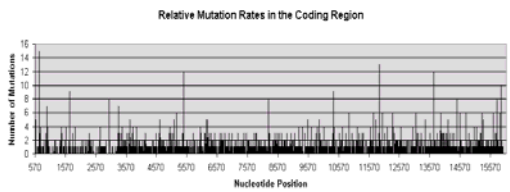
Results

- Analysis of 646 coding region genomes

	Parsimony		NJ	
Data Set (# genomes)	Tree Length	α estimation	Tree Length	α estimation
Ingman HV1 (53)	144	0.2091	144	0.2081
Ingman Control Region (53)	273	0.0038	281	0.0036
Ingman Coding Region (53)	588	0.0075	588	0.0074
Ingman Full Data (53)	873	0.0050	876	0.0067
Total Coding Data (646)	2352	0.0086	2353	0.0083

Extreme rate variation exists in the coding region

Relative Mutation Rates



The Mutation Rate Spectrum

- How does the estimated mutation rate spectrum compare to the forensically informative SNPs?
- Are all of the forensic SNPs mutational "hot-spots?"

Mutation Rates and the 8 Multiplex SNP Panels

Length	Character	Gene	codon	241 Caucasians
15	709	12S	*	Yes
13	11914	ND4	3	Yes-SNP
12	5460	ND2	1	Yes
12	13708	ND5	1	Yes
10	15924	tRNA(thr)	*	Yes
9	1719	16S	*	Yes
9	10398	ND3	1	Yes
8	3010	16S	*	Yes-SNP
8	8251	COII	3	
8	14470	ND6	3	Yes-SNP
8	15784	CYTB	3	
7	961	12S	*	
7	3316	ND1	1	
6	5237	ND2	3	Yes
6	10915	ND4	3	Yes
6	11719	ND4	3	Yes-SNP
6	12007	ND4	3	Yes-SNP
6	12346	ND5	1	
6	13105	ND5	1	Yes
6	13928	ND5	2	
6	14569	ND6	3	
6	14766	CYTB	2	
6	15301	CYTB	3	
6	15670	CYTB	3	
6	15884	CYTB	nc	Yes-SNP

Only 6 of the 59 SNPs are among the "fastest" sites

Mutation Rates and the 8 Multiplex SNP Panels

Length	Character	Gene	codon	241 Caucasians
15	709	12S	*	Yes
13	11914	ND4	3	Yes-SNP
12	5460	ND2	1	Yes
12	13708	ND5	1	Yes
10	15924	tRNA(thr)	*	Yes
9	1719	16S	*	Yes
9	10398	ND3	1	Yes
8	3010	16S	*	Yes-SNP
8	8251	COII	3	
8	14470	ND6	3	Yes-SNP
8	15784	CYTB	3	
7	961	12S	*	
7	3316	ND1	1	
6	5237	ND2	3	Yes
6	10915	ND4	3	Yes
6	11719	ND4	3	Yes-SNP
6	12007	ND4	3	Yes-SNP
6	12346	ND5	1	
6	13105	ND5	1	Yes
6	13928	ND5	2	
6	14569	ND6	3	
6	14766	CYTB	2	
6	15301	CYTB	3	
6	15670	CYTB	3	
6	15884	CYTB	nc	Yes-SNP

What about These highly polymorphic mutations?

Mutation Rates and the 8 Multiplex SNP Panels

Length	Character	Gene	codon	241 Caucasians
15	709	12S	*	Yes
13	11914	ND4	3	Yes-SNP
12	5460	ND2	1	Yes
12	13708	ND5	1	Yes
10	15924	tRNA(thr)	*	Yes
9	1719	16S	*	Yes
9	10398	ND3	1	Yes
8	3010	16S	*	Yes-SNP
8	8251	COII	3	
8	14470	ND6	3	Yes-SNP
8	15784	CYTB	3	
7	961	12S	*	
7	3316	ND1	1	
6	5237	ND2	3	Yes
6	10915	ND4	3	Yes
6	11719	ND4	3	Yes-SNP
6	12007	ND4	3	Yes-SNP
6	12346	ND5	1	
6	13105	ND5	1	Yes
6	13928	ND5	2	
6	14569	ND6	3	
6	14766	CYTB	2	
6	15301	CYTB	3	
6	15670	CYTB	3	
6	15884	CYTB	nc	Yes-SNP

Additional SNP panels with "hot spots" and non-synonymous SNPs have been developed (Rebecca Just, AFDIL)

A Case Example

Skeletal remains - "H1" in the HV1/HV2 region.

Thought to belong to one of two individuals…

(Smith or Jones)

Family references for Smith and Jones were obtained.

Smith Family	Jones Family
263 A-G	263 A-G
315.1 C	315.1 C

A Case Example

Skeletal remains - "H1" in the HV1/HV2 region.

Thought to belong to one of two individuals…

(Smith or Jones)

Remains tested for VR region: 477 T-C and 16519 T-C

Smith Family	Jones Family
263 A-G	263 A-G
315.1 C	315.1 C
477 T-C	
16519 T-C	16519 T-C

Question….

Can the Smith Family be excluded as a possible family reference for the skeletal remains?

Smith Family	Jones Family
263 A-G	263 A-G
315.1 C	315.1 C
477 T-C	
16519 T-C	16519 T-C

Question….

NO!

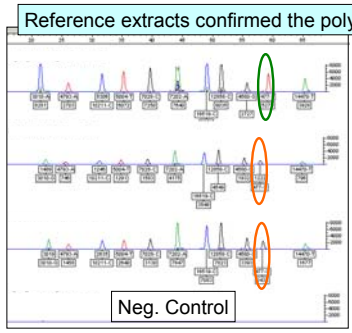
Only one mutation differs between the two families…

Smith Family	Jones Family
263 A-G	263 A-G
315.1 C	315.1 C
477 T-C	
16519 T-C	16519 T-C

INCONCLUSIVE

A Case Example

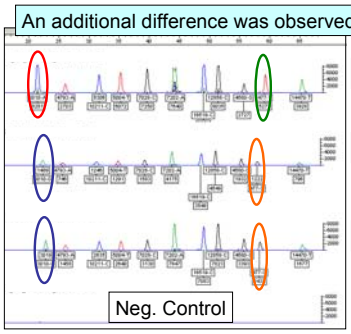
Reference extracts confirmed the polymorphism at 477



(R. Just, AFDIL)

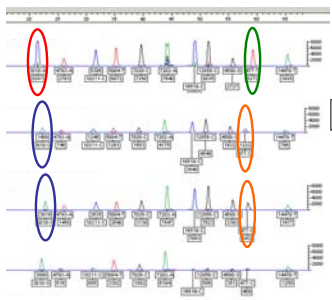
A Case Example

An additional difference was observed at position 3010



(R. Just, AFDIL)

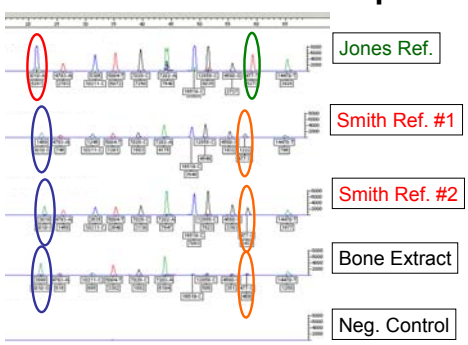
A Case Example



15uL Reaction; 0.07Units/uL Taq; 31 cycles --- 100 RFU cutoff

(R. Just, AFDIL)

A Case Example



(R. Just, AFDIL)

A Case Example

Smith Family	Skeletal Remains	Jones Family
263 A-G	263 A-G	263 A-G
315.1 C	315.1 C	315.1 C
477 T-C	477 T-C	16519 T-C
3010 A-G	3010 A-G	
16519 T-C	16519 T-C	

Remains – match exactly the Smith family, now **2 differences** from the Jones family – can be *excluded*.

Summary

- Purpose – Maximize Discrimination.
- A supplement to current HV1/HV2 testing.
- When the Forensic Scientist encounters a common type, select the most discriminating SNP panel.

Summary

- AFDIL – focused on sites that are not associated with the potential for phenotypic change.
- Most of the informative sites are rare, slow polymorphisms that are useful for discrimination in a particular common type.
- A few SNP sites may be useful for resolving common HV1/HV2 types from various backgrounds.
- Evaluation of non-synonymous sites that are not associated with diseases may also be useful for forensic discrimination… site-by-site evaluation (e.g. 3010 is very useful among HgH.

Publications

Michael D. Coble · Rebecca S. Just
Jennifer E. O'Callaghan · Ilona 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

Efforts with Coding Region Sequencing Applied to Human mtDNA Testing

- Tzen *et al.* (2001) *Forensic Sci. Int.* 120:204-209
  - Portions of mtATP6, mtATP8 among 119 Chinese individuals
- Andreasson *et al.* (2002) *Biotechniques* 32:124-133
  - Highly variable regions of mtDB among 190 Swedish individuals
- Lee *et al.* (2002) *Int. J. Legal Med.* 116:74-78
  - mtCyt B among 98 Korean individuals
- Lutz-Bonengel *et al.* (2003) *Int. J. Legal Med.* 117:133-142
  - mtATP6, mtATP8, mtND4 among 109 German individuals
- Poetsch *et al.* (2003) *Mitochondrion* 3:133-137
  - portions of tRNA K, ATP6, ATP8 among 180 German individuals
- Coble *et al.* (2004) *Int. J. Legal Med.*, 118:137-146
  - 241 complete mtGenomes among 18 common Cauc. HV1/HV2 types

Criticisms of Synonymous SNPs for Discrimination

Int J Legal Med (2005) 119: 314–315
DOI 10.1007/s00414-005-0543-y

LETTER TO THE EDITOR

B. Budowle · U. Gyllensten · R. Chakraborty · M. Allen

Forensic analysis of the mitochondrial coding region and association to disease

Budowle *et al.* (2005)

- [Coble and Vallone] have proposed that forensic analyses of the coding region [should] be restricted to synonymous substitutions [and] suggest that sequencing strategies for forensic analyses of the coding region of the mtDNA genome should be avoided [and] that only SNP-based systems should be employed.
- We disagree with this proposition [would] severely hamper the use of mtDNA in forensic testing.

Budowle *et al.* (2005)

- "by limiting the analysis only to synonymous polymorphisms that cannot have any phenotypic effect, ***a large part of the polymorphic positions (and thus forensically informative) would be excluded.***"

An Evaluation of Coding Region SNPs to Sequencing

- Data

How well will multiplex A perform among with random samples belonging to the most common haplotype?

What is the general utility of the SNPs for HgH proposed by Coble *et al.* (2004)?

24 Samples H:1 (263 A-G; 315.1C)

54 hgH Sequences (Achilli et al. 2004)

An Evaluation of Coding Region SNPs to Sequencing

- Methods

G3010A	A7202G
G4580A	C10211T
A4793G	C12858T
T5004C	T14470C
C7028T	

9 SNPs

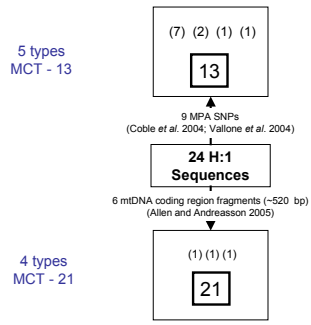
Coble *et al.* (2004) *Int J Legal Med* 118:137–146.
Vallone *et al.* (2004) *Int J Legal Med* 118:147–157.

Region (np)	Variable Position(s)
2782 - (2662-2762)	A2706G
4275 - (4303-4363)	T4336C
8665 - (8689-8780)	G8697A
	T8705C
10362 - (10385-10484)	A10398G
	T10463C
12673 - (12694-12784)	C12705T
15758 - (15777-15873)	C15833T

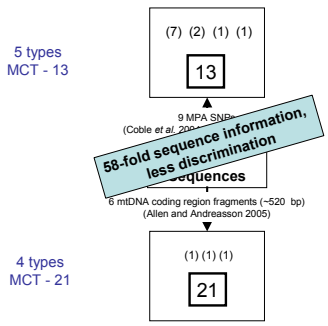
~520 bp

Allen and Andreasson (2005) *Methods Mol Biol* 297:179–196.

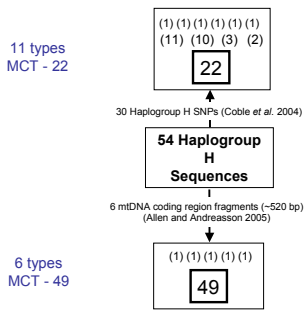
Results



Results



Results



Results

11 types
MCT - 22

(1) (1) (1) (1) (1) (1)
(11) (10) (3) (2)
22

Targeted SNPs performed even better than sequencing among non-common HgH types

6 mtDNA coding region fragments (~520 bp)
(Allen and Andreasson 2005)

6 types
MCT - 49

(1) (1) (1) (1) (1)
49

Why Did 9 SNPs Outperform Sequencing?

Region (np)	Variable Position(s)
2782 - (2662-2762)	A2706G
4275 - (4303-4363)	T4336C
8665 - (8689-8780)	G8697A
	T8705C
10362 - (10385-10484)	A10398G
	T10463C
12673 - (12694-12784)	C12705T
15758 - (15777-15873)	C15833T

Diagnostic Haplogroup T SNPs

Allen and Andreasson (2005) *Methods Mol Biol* 297:179–196.

Why Did 9 SNPs Outperform Sequencing?

Region (np)	Variable Position(s)
2782 - (2662-2762)	A2706G
4275 - (4303-4363)	T4336C
8665 - (8689-8780)	G8697A
	T8705C
10362 - (10385-10484)	A10398G
	T10463C
12673 - (12694-12784)	C12705T
15758 - (15777-15873)	C15833T

Diagnostic for SuperHaplogroup R

Allen and Andreasson (2005) *Methods Mol Biol* 297:179–196.

All Caucasians have the rCRS variant

What About Synonymous SNPs?

Are Non-Synonymous SNPs a rich source of variation in the mtDNA coding region that we are "throwing away?"

1. Global Variation - Location (mtDB – 2461 coding region sequences)

11,391 nucleotides in the coding region genes

1	2	3	
7,594	3,797		← potential variable sites
798	1,862		← observed variants
10.5%	49.0%		

2. Global Variation – "Polymorphism" (mtDB – 2461 coding region sequences)

114 variants at non-synonymous positions that occur at 1% or greater

111/114 (97%) belong can readily be described to mtDNA haplogroups (redundant and uninformative)

How Much Information is Lost?

African-derived Sequence

Haplogroup L0a1

"Hausa" (Ingman et al. 2000)

C	64	T	A	750	G	C	7028	T	G	11719	A
A	93	G	G	769	A	A	7146	G	G	11914	A
C	150	T	T	825	A	C	7256	T	G	12007	A
G	185	A	G	1018	A	G	7521	A	C	12705	T
A	189	G	C	1048	T	C	8428	T	A	12720	G
A	200	G	A	1438	G	C	8468	T	A	13105	G
T	236	C	A	2245	C	A	8566	G	A	13276	G
G	247	A	A	2706	G	C	8655	T	C	13506	T
A	263	G	G	2758	A	A	8701	G	C	13650	T
C	522	:	T	2885	C	A	8860	G	T	14308	C
A	523	:	C	3107	:	C	9042	T	C	14766	T
G	16129	A	C	3516	A	A	9347	G	C	15136	T
C	16148	T	C	3594	T	T	9540	C	A	15326	G
C	16168	T	T	3866	C	G	9755	A	G	15431	A
T	16172	C	A	4104	G	C	9818	T			
C	16187	T	C	4312	T	A	10398	G			
C	16188	G	T	4586	C	G	10589	A			
T	16189	C	A	4769	G	C	10664	T			
C	16223	T	T	5096	C	G	10688	A			
A	16230	G	G	5231	A	T	10810	C			
T	16311	C	T	5442	C	T	10873	C			
C	16320	T	G	5460	A	T	10915	C			
T	16362	C	C	5603	T	G	11176	A			
T	16519	C	T	6185	C	A	11641	G			

How Much Information is Lost?

27 non-synon/RNA mutations

35 synon mutations

"a large part of the polymorphic positions (and thus forensically informative) would be excluded."

A	750	G	C	7028	T	G	11719	A
G	769	A	A	7146	G	G	11914	A
T	825	A	C	7256	T	G	12007	A
G	1018	A	G	7521	A	C	12705	T
C	1048	T	C	8428	T	A	12720	G
A	1438	G	C	8468	T	A	13105	G
A	2245	C	A	8566	G	A	13276	G
A	2706	G	C	8655	T	C	13506	T
G	2758	A	A	8701	G	C	13650	T
T	2885	C	A	8860	G	T	14308	C
C	3107	:	C	9042	T	C	14766	T
C	3516	A	A	9347	G	C	15136	T
C	3594	T	T	9540	C	A	15326	G
T	3866	C	G	9755	A	G	15431	A
A	4104	G	C	9818	T			
C	4312	T	A	10398	G			
T	4586	C	G	10589	A			
A	4769	G	C	10664	T			
T	5096	C	G	10688	A			
G	5231	A	T	10810	C			
T	5442	C	T	10873	C			
G	5460	A	T	10915	C			
C	5603	T	G	11176	A			
T	6185	C	A	11641	G			

Coding Region

How Much Information is Lost?

All polymorphisms can be attributed to haplogroup L0a1 or as differences from the rCRS

High frequency is not necessarily a reliable indicator of "informativeness" in the coding region.

A	750 G	C	7028 T	G	11719 A
G	769 A	A	7146 G	G	11914 A
T	825 A	C	7256 T	G	12007 A
G	1018 A	G	7521 A	C	12705 T
C	1048 T	C	8428 T	A	12720 G
A	1438 G	C	8468 T	A	13105 G
A	2245 C	A	8566 G	A	13276 G
A	2706 G	C	8655 T	C	13506 T
G	2758 A	A	8701 G	C	13650 T
T	2885 C	A	8860 G	T	14308 C
C	3107 :	C	9042 T	C	14766 T
C	3516 A	A	9347 G	C	15136 T
C	3594 T	T	9540 C	A	15326 G
T	3866 C	G	9755 A	G	15431 A
A	4104 G	C	9818 T		
C	4312 T	A	10398 G		
T	4586 C	G	10589 A		
A	4769 G	C	10664 T		
T	5096 C	G	10688 A		
G	5231 A	T	10810 C		
T	5442 C	T	10873 C		
G	5460 A	T	10915 C		
C	5603 T	G	11176 A		
T	6185 C	A	11641 G		

Coding Region

Conclusions

- A selected SNP method out-performed a random sequencing protocol for increased discrimination.
- This method was developed to avoid additional sequencing, as often, the casework at AFDIL involves challenging cases where the quantity and quality of extract would prohibit an extensive post-HV1/HV2 sequencing strategy.

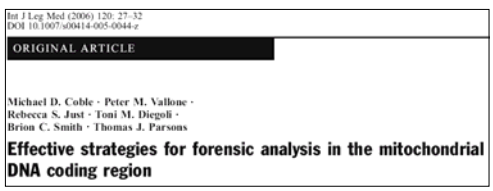
Conclusions

- Budowle *et al.* (2005) make several valid points about the usefulness of non-synonymous sites for discrimination, and we have made a careful evaluation about the potential use of these sites.
- However, many cases processed by AFDIL are publicly visible and involves large segments of the general population. The US military now has a policy of compulsory submission of a blood sample retained solely for the purposes of DNA identification, which is necessary in the face of military casualty.

Conclusions

- A conservative approach was developed, and this may or may not meet the needs of other forensic laboratories
- Some countries, such as Germany, have strict regulations the use of forensic testing that may reveal medical information… this has resulted in the call for disqualification of certain markers (e.g. X chromosome – see Szibor *et al.* 2005 *IJLM*).
- Need to weigh the costs and benefits for developing effective strategies to increase mtDNA discrimination.

More Information


Int J Leg Med (2006) 120: 27–32
DOI 10.1007/s00414-005-0044-z

ORIGINAL ARTICLE

Michael D. Coble · Peter M. Vallone · Rebecca S. Just · Toni M. Diegoli · Brion C. Smith · Thomas J. Parsons

Effective strategies for forensic analysis in the mitochondrial DNA coding region

http://www.cstl.nist.gov/biotech/strbase/pub_pres/Coble_IJLM_coding_mtSNPs.pdf

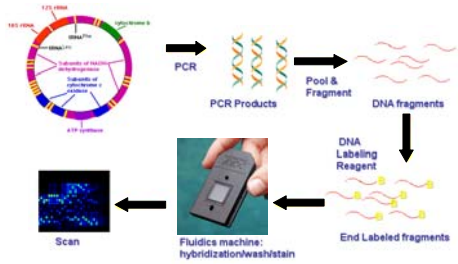
Non-Sequencing Strategy (Affymetrix mtDNA Chip)

Evaluation of New Technologies

Affymetrix GeneChip® Human Mitochondrial Resequencing Array version 2.0

Entire mtDNA genome plus additional redundant tiling for control region polymorphisms and mtDNA haplogroups

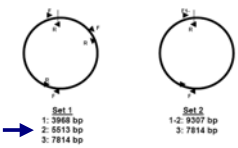
Overview of the Process



Linda Strausbaugh and John Jakupciak

Overview of the Process

Long PCR



	Forward Primer	Reverse Primer	Total BP
Mito1	ACATAGCACATTACAGTCAAATCCCTTCTCGTCCC	TGAGATTGTTTGGGCTACTGCTCGCAGTGC	3968
Mito2	TACTCAATCCTCTGATCAGGGTGAGCATCAAACTC	GCTTGGATTAAGGCGACAGCGATTTCTAGGATAGT	5513
Mito3	TCATTTTATTGCCACAACTAACCTCCTCGGACTC	CGTGATGTCTTATTTAAGGGGAACGTGTGGGCTAT	7814
Mito1-2	ACATAGCACATTACAGTCAAATCCCTTCTCGTCCC	ATTGCTAGGGTGGCGCTTCCAATTAGGTGC	9307

Tiling of the Array

Labeled PCR fragment

Forward CGTACTACACGACACGTACTACGTTGTACA

Reverse Tiling

TGATGTGCTGTGCATGATGCAACAT
TGATGTGCTGTGTATGATGCAACAT
TGATGTGCTGTGAATGATGCAACAT
TGATGTGCTGTGGATGATGCAACAT

25-mers with variable SNP in the 13th position

Tiling of the Array

Forward CGTACTACACGACACGTACTACGTTGTACA
X
TGATGTGCTGTGTATGATGCAACAT

Forward CGTACTACACGACACGTACTACGTTGTACA
X
TGATGTGCTGTGAATGATGCAACAT

Forward CGTACTACACGACACGTACTACGTTGTACA
X
TGATGTGCTGTGGATGATGCAACAT

Forward CGTACTACACGACACGTACTACGTTGTACA
|||
TGATGTGCTGTGCATGATGCAACAT

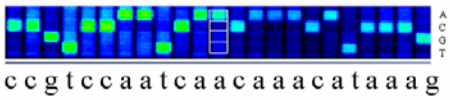
Tiling of the Array

Forward CGTACTACACGACACGTACTACGTTGTACA
|||
TGATGTGCTGTGCATGATGCAACAT

Sequence Tiled

Sequence Hybridized

Automated Analysis



The software can utilize 1-12 different thresholds (scores) for making calls

Threshold-1 is the most "liberal"
Threshold-12 is the most "conservative"

John Jakupciak

NIST Population Sample

46 differences between rCRS and PT83859.

Peter Vallone, John Jakupciak

Sequence Data Pos	Change	MitoChip Thresh - 1	MitoChip Thresh - 12
73	A - G		
249	A -del.	X	X
290	A -del.	X	X
291	A -del.	X	X
315.1	C - ins.	X	X
523	A -del.	X	X
524	C -del.	X	X
9,540	T - C	X	X
9,545	A - G		X
9,547	G - C	wrong	X
9,557	C - T		X
10,398	A - G		
10,399	C - G	wrong	X
10,400	C - T		
10,873	T - C	X	X
11,719	G - A	X	X
11,914	G - A	X	X
16,326	A - C	wrong	X
16,327	C - T		
16,519	T - C		
	Point µ	35/39 (90%)	33/39 (85%)
	Overall	35/46 (76%)	33/46 (72%)

(Only partial profile is shown)

Variants missed using either threshold

NIST Population Sample

46 differences between rCRS and PT83859.

Peter Vallone, John Jakupciak

Sequence Data Pos	Change	MitoChip Thresh - 1	MitoChip Thresh - 12
73	A - G		
249	A -del.	X	X
290	A -del.	X	X
291	A -del.	X	X
315.1	C - ins.	X	X
523	A -del.	X	X
524	C -del.	X	X
9,540	T - C	X	X
9,545	A - G		X
9,547	G - C	wrong	X
9,557	C - T		X
10,398	A - G		
10,399	C - G	wrong	X
10,400	C - T		
10,873	T - C	X	X
11,719	G - A	X	X
11,914	G - A	X	X
16,326	A - C	wrong	X
16,327	C - T		
16,519	T - C		
	Point µ	35/39 (90%)	33/39 (85%)
	Overall	35/46 (76%)	33/46 (72%)

Only partial profile is shown

Mis-hybridization at 10399 (between 10398 A-G and 10400 C-T) is called using threshold 1 for analysis.

This is not called when the threshold 12 is used.

NIST Population Sample

Peter Vallone, John Jakupciak

46 differences between rCRS and PT83859.

Sequence Data Pos	Change	MitoChip Thresh - 1	MitoChip Thresh - 12
73	A - G		
249	A -del.	X	X
290	A -del.	X	X
291	A -del.	X	X
315.1	C - ins.	X	X
523	A -del.	X	X
524	C -del.	X	X
9,540	T - C	X	X
9,545	A - G		X
9,547	G - C	wrong	X
9,557	C - T		X
10,398	A - G		
10,399	C - G	wrong	X
10,400	C - T		
10,873	T - C	X	X
11,719	G - A	X	X
11,914	G - A	X	X
16,326	A - C	wrong	X
16,327	C - T		
16,519	T - C		
	Point µ	35/39 (90%)	33/39 (85%)
	Overall	35/46 (76%)	33/46 (72%)

The 9545 A-G call is not made with threshold 12, but is called correctly with threshold 1.

Note close proximity of 9540 T-C, which was not called with either threshold.

NIST Population Sample

Peter Vallone, John Jakupciak

46 differences between rCRS and PT83859.

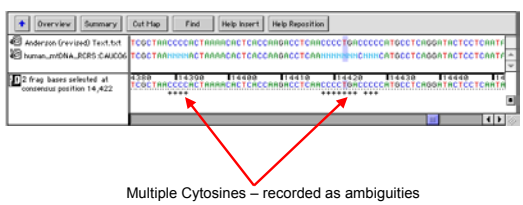
Sequence Data Pos	Change	MitoChip Thresh - 1	MitoChip Thresh - 12
73	A - G		
249	A -del.	X	X
290	A -del.	X	X
291	A -del.	X	X
315.1	C - ins.	X	X
523	A -del.	X	X
524	C -del.	X	X
9,540	T - C	X	X
9,545	A - G		X
9,547	G - C	wrong	X
9,557	C - T		X
10,398	A - G		
10,399	C - G	wrong	X
10,400	C - T		
10,873	T - C	X	X
11,719	G - A	X	X
11,914	G - A	X	X
16,326	A - C	wrong	X
16,327	C - T		
16,519	T - C		
	Point µ	35/39 (90%)	33/39 (85%)
	Overall	35/46 (76%)	33/46 (72%)

mtDNA chip has difficulties calling indels.

MitoChip Difficulties

Text file of sample compared to the rCRS

Score 12 – Caucasian sample 06



Multiple Cytosines – recorded as ambiguities

Affymetrix Mitochip

- Advantages
  - Cost evaluation compared to sequencing.
  - Higher throughput than manual sequencing.
  - Relative ease of use.
  - Potential for processing family references.
- Disadvantages
  - Variable thresholds can give differing results.
  - At present – requires ng quantity of template.
  - Long PCR – highly unlikely with degraded DNA.
  - Indels, heteroplasmy.

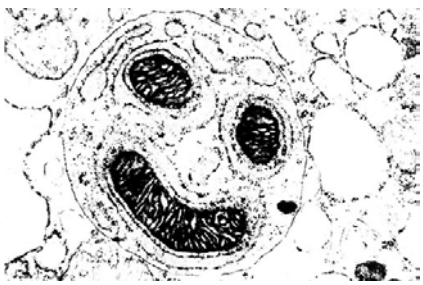
Acknowledgments

- Dr. John Butler (NIST)
- Dr. Tom Parsons (ICMP, formally AFDIL)
- Dr. Peter Vallone and Dr. John Jakupciak (NIST)
- Rebecca Just, Jodi Irwin, Jessica Saunier, Jennifer O'Callaghan, Ilona Letmanyi, and Christine Peterson (present and formerly of AFDIL)

Disclaimer

This project was funded by the National Institute of Justice through interagency agreement 2003-IJ-R-029 to the NIST Office of Law Enforcement Standards. Points of view are those of the authors and do not necessarily represent the official position or policies of the US Department of Justice. Certain commercial equipment, instruments and materials are identified in order to specify experimental procedures as completely as possible. In no case does such identification imply a recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that any of the materials, instruments, or equipment identified are necessarily the best available for the purpose.

QUESTIONS?



http://www-medlib.med.utah.edu/WebPath/jpeg2/EM003.jpg