DNA Future Trends Technical Review/Workshop

Michael D. Coble

NIST Applied Genetics Group
National Institute of Standards and Technology
Gaithersburg, Maryland

Presentation Outline

- Marker Systems – STRs
- Marker Systems – SNPs
- Low Template (Copy Number) Testing
- Second (Next) Generation Sequencing
- Mixture Interpretation

Types of Genetic Variation

Length Variation

CODIS Loci are STRs

- **short tandem repeats (STRs)**

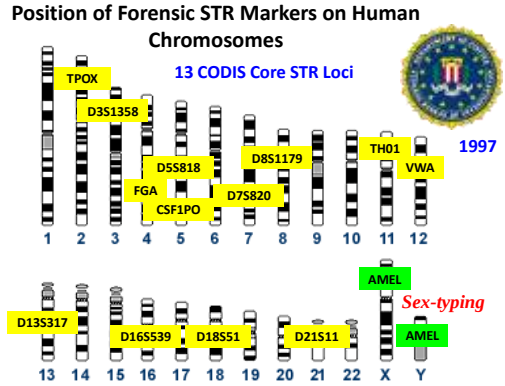
CTAGTCGT**(GATA)(GATA)(GATA)**GCGATCGT

Sequence Variation

- insertions/deletions
- **single nucleotide polymorphisms (SNPs)**

GCTAGTCGATGCTC**(G/A)**GCGTATGCTGTAGC

Position of Forensic STR Markers on Human Chromosomes



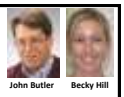
The 11 STR Loci Beyond the CODIS 13

STR Locus	Location	Repeat Motif	Allele Range*	# Alleles*
D2S1338	2q35	TGCC/TTCC	10 to 31	40
D19S433	19q12	AAGG/TAGG	5.2 to 20	36
Penta D	21q22.3	AAAGA	1.1 to 19	50
Penta E	15q26.2	AAAGA	5 to 32	53
D1S1656	1q42	TAGA	8 to 20.3	25
D12S391	12p13.2	AGAT/AGAC	13 to 27.2	52
D2S441	2p14	TCTA/TCAA	8 to 17	22
D10S1248	10q26.3	GGAA	7 to 19	13
D22S1045	22q12.3	ATT	7 to 20	14
SE33	6q14	AAAG‡	3 to 49	178
D6S1043	6q15	AGAT/AGAC	8 to 25	25

5 new European loci

**Allele range and number of observed alleles from Appendix 1, J.M. Butler (2011) Advanced Topics in Forensic DNA Typing: Methodology; ‡SE33 alleles have complex repeat structure*

New STR Loci Characterized



Coble and Butler (2005) *J. Forensic Sci.* 50(1):43-53

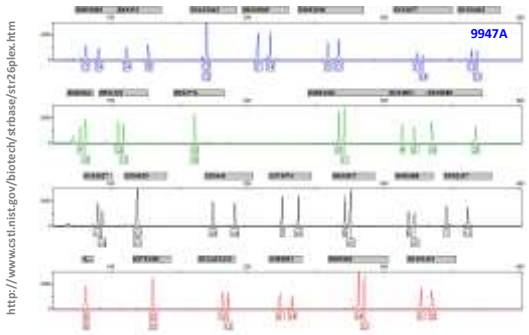
Characterization of New MiniSTR Loci to Aid Analysis of Degraded DNA*

Hill *et al.* (2008) *J. Forensic Sci.* 53(1):73-80

Characterization of 26 MiniSTR Loci for Improved Analysis of Degraded DNA Samples

NIST STR 26plex

Hill et al. (2009) *Journal of Forensic Sciences,* 54(5):1008-1015



Gender identification + 25 autosomal STR loci in a single amplification

SNP Typing


Pete Vallone

- Forensic scale SNP assays
  - ~10 - 50 SNP markers
  - Utility with low amounts of sample (< 1ng)
- High throughput sequencing and DNA/SNP microarrays
  - Thousands to millions of SNPs
  - Greater input amounts of DNA are required
  - Higher level computational methods are required (data storage and analysis)

Multiple approaches and technologies exist for SNP typing

SNP Classifications

- Individual Identification SNPs (IISNPs): SNPs that collectively give very low probabilities of two individuals having the same multi-locus genotype
- Ancestry Informative SNPs (AISNPs): SNPs that collectively give a high probability of an individual's ancestry being from one part of the world or being derived from two or more areas of the world
- Phenotype Informative SNPs (PISNPs): SNPs that provide a high probability that the individual has particular phenotypes, such as a particular skin color, hair color, eye color, etc.
- Lineage Informative SNPs (LISNPs): Sets of tightly linked SNPs that function as multi-allelic markers that can serve to identify relatives with higher probabilities than simple bi-allelic SNPs

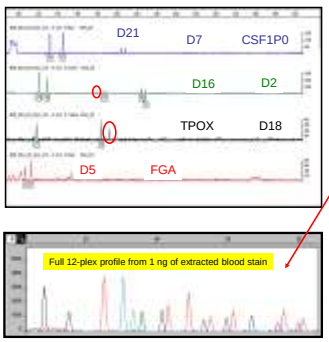
Budowle, B. and van, Daal. A. (2008) Forensically relevant SNP classes, *Biotechniques 44*, 603-8, 610.
Butler, J.M., Budowle, B., Gill, P., Kidd, K.K., Phillips, C., Schneider, P.M., Vallone, P.M., Morling, N. (2008) Report on ISFG SNP panel discussion. *Forensic Science International: Genetics Supplement Series (Progress in Forensic Genetics 12)* 1: 471-472.

Individual Identification SNPs

- Use for individual identification of a sample
- Power of Discrimination – how many SNPs are needed to match STRs?
- Can the assay amplify > 30 loci using a small amount of template DNA?
- Use on a degraded sample
- Issues with a mixture

General Criteria
- Low F_{ST} (not population specific)
- No linkage disequilibrium between SNPs or CODIS loci
- Amplicon size < 120 bp
- Minimum 30% heterozygosity
- Minimum distance of 100 kb between SNPs and neighboring genes

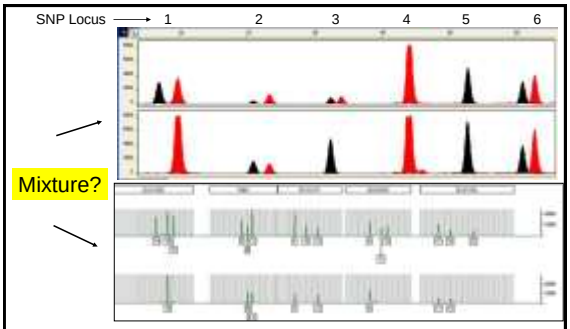
Typing an Aged Blood Stain



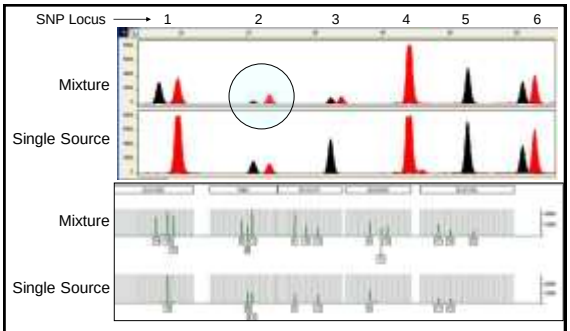
Identifiler genotyping result from a blood stain aged 15 years stored at room temperature. (stored on 903 paper, Chelex extracted)

The same sample extract as above typed by the 12-plex SNP assay.

11 different samples that gave partial profiles with Identifiler gave full profiles typed with the 12-plex assay.



Butler, J.M., Coble, M.D., Vallone, P.M. (2007) STRs vs SNPs: thoughts on the future of forensic DNA testing. *Forensic Sci. Med. Pathol.* 3: 200-205.



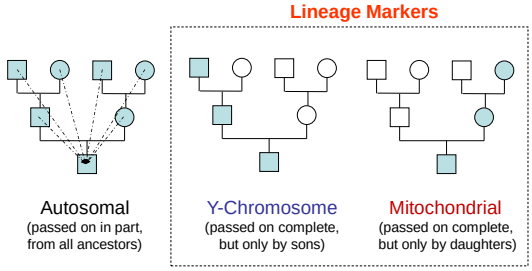
• Only SNP locus 2 (top panel) has an allele imbalance suggesting that a possible second contributor is present
• 5 STR loci shown from the green dye channel of the Identifiler kit contain three alleles making mixture detection straightforward

Butler, J.M., Coble, M.D., Vallone, P.M. (2007) STRs vs SNPs: thoughts on the future of forensic DNA testing. *Forensic Sci. Med. Pathol.* 3: 200-205.

Challenges of Using IISNPs for Forensic Testing

- U.S. and international databases consist of STR profiles
  - Is there a benefit to changing the DNA typing technology for databanking and routine casework?
- Mixture analysis using genome-wide arrays
  - Detection is possible but interpretation is still ongoing
- SNPs in linkage disequilibrium: match probability calculations?
- Sensitivity
  - Genome-wide arrays require >500 ng DNA
  - Non-biased whole genome amplification is necessary but not here yet
- Cost
  - Approximately $500 per sample for genome-wide arrays
  - Arrays cost the same amount for typing 1,000 SNPs or 1 million SNPs

Different Inheritance Patterns



Autosomal
(passed on in part, from all ancestors)

Y-Chromosome
(passed on complete, but only by sons)

Mitochondrial
(passed on complete, but only by daughters)

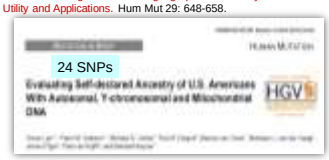
Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 9.1, ©Elsevier Science/Academic Press

Ancestry Informative SNPs

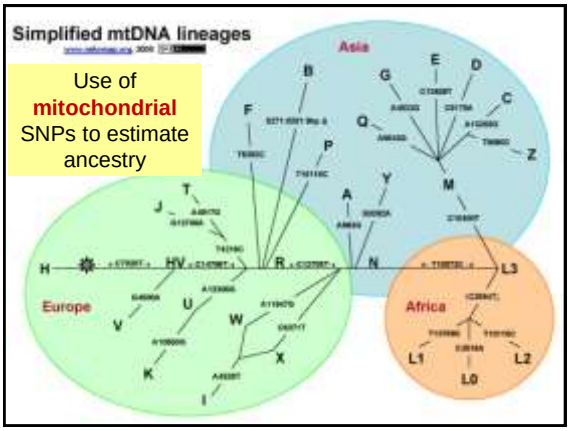
Lao O, van Duijn K, Kersbergen P, de Knijff P, Kayser M. (2006) Proportioning whole-genome single-nucleotide-polymorphism diversity for the identification of geographic population structure and genetic ancestry. Am J Hum Genet 78: 680-90. *10 SNPs*

Phillips, C., Salas, A., Sanchez, J.J., Fondevila , M., Gomez-Tato, A., Alvarez-Dios, J., Calaza, M., Casares de Cal, M., Ballard, M., Lareu, M.V., Carracedo, A. (2007) Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs. FSI: Genetics 1: 273-280. *34 SNPs*

Halder, I., Shriver, M., Thomas, M., Fernandez, J.R., Frudakis, T. (2008) A Panel of Ancestry Informative Markers for Estimating Individual Biogeographical Ancestry and Admixture From Four Continents: Utility and Applications. Hum Mut 29: 648-658.

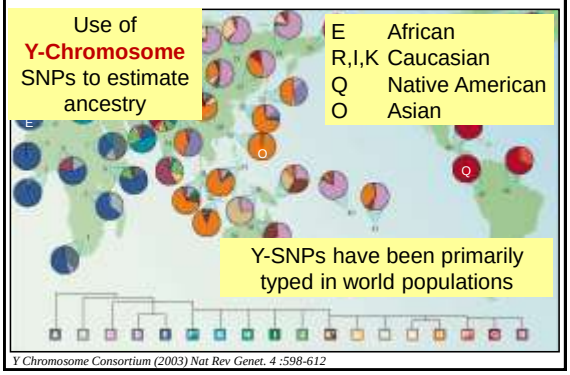


Ongoing and active area of research
Screening for sets of **autosomal** SNPs that will estimate ancestry
Global perspective



Global Distribution of Y Haplogroups

Use of
Y-Chromosome
SNPs to estimate ancestry



Y Chromosome Consortium (2003) Nat Rev Genet. 4 :598-612

Biogeographical Ancestry

Sample 1

Mitochondrial Hg
H
Europe

Y Chromosome Hg
R1b1b2
Western Europe

24 autosomal SNPs
94 % Caucasian
6 % (other)

Self Identified Ancestry
Caucasian

Biogeographical Ancestry

Sample 2

Mitochondrial Hg
L0a1a
Africa

Y Chromosome Hg
E
Africa/Southern Europe

24 autosomal SNPs
98 % Afr. Amer.
2 % (other)

Self Identified Ancestry
African American

Biogeographical Ancestry

Sample 3

Mitochondrial Hg
L3d1
Africa

Y Chromosome Hg
R1b1b2
Western Europe

24 autosomal SNPs
99 % Afr. Amer.
1 % (other)

Self Identified Ancestry
African American

Biogeographical Ancestry

Sample 4

Mitochondrial Hg
A2
Americas

Y Chromosome Hg
R1b1b2
Western Europe

24 autosomal SNPs
89 % Hispanic
7 % Asian
4 % Caucasian

Self Identified Ancestry
Hispanic

Biogeographical Ancestry

Sample 5

Mitochondrial Hg
H
Europe

Y Chromosome Hg
R1b1b2
Western Europe

24 autosomal SNPs
69 % Hispanic
28 % Caucasian
3 % Asian

Self Identified Ancestry
Caucasian

Summary of Ancestry SNPs

Success estimating Self Identified Ancestry

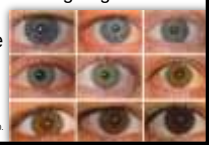
n	U.S. sampling	Mitochondrial	Y Chromosome	24 Autosomal
259	African American	93%	69%	98%
262	Caucasian	97%	84%	81%
49	Asians	99%	100%	100%
140	Hispanics	NA	NA	74%

The appropriate populations of interest need to be studied

Issues with 'recent' admixed populations (e.g. US Hispanics)

Phenotype Informative SNPs

- Predict an observable trait – definitively
  - Eye, hair, and skin color
  - Height, stature
- Some key pigmentation genes have been characterized
  - Wide range of pigmentation in humans
  - Multiple genes involved, complex phenotype
- Gene discovery and characterization is ongoing
- Should not be predictive of disease



Tully G. Genotype versus phenotype: human pigmentation. Forensic Sci Int Genet. 2007 1: 105-10.

Pigmentation Related Genes

Eye	Hair	Skin
ASIP	ASIP	ASIP
		DCT
		DRD2
		EGFR
HERC2	HERC2	HERC2
IRF4	IRF4	IRF4
	KITLG	KITLG
	MATP	MATP
	MC1R	MC1R
OCA2	OCA2	
		MYO5A
SLC24A4	SLC24A4	SLC24A4
SLC45A2		
	TPCN2	SLC24A5
TYR	TYR	TYR

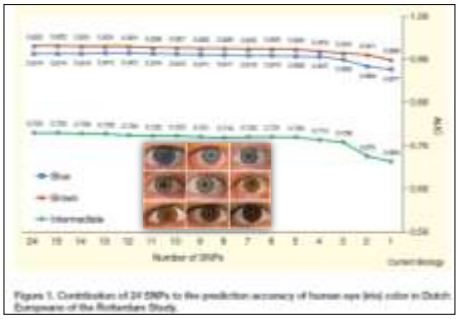
Overlap between eye, hair, and skin related genes

Potential for a panel of markers that could predict all three traits

Recent Work on PISNPs

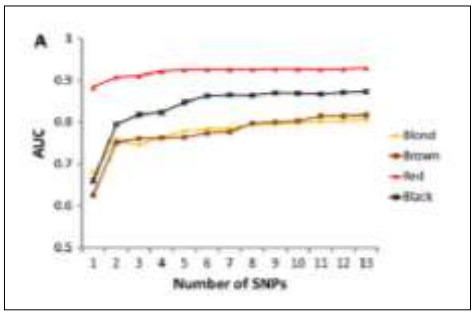
- Branicki W, Kayser M et al. (2011). Model-based prediction of human hair color using DNA variants. *Human Genetics;* DOI 10.1007/s00439-010-0939-8
- Walsh S., et al. (2010) IrisPlex: A sensitive DNA tool for accurate prediction of blue and brown eye colour in the absence of ancestry information. Forensic Sci. Int. Genet. (Epub)
- Mengel-From J., et al. (2010) Human eye colour and HERC2, OCA2 and MATP. Forensic Sci. Int. Genet. 5: 323-8
- Kayser M., Schneider P.M. (2009) DNA-based prediction of human externally visible characteristics in forensics: motivations, scientific challenges, and ethical considerations. Forensic Sci. Int. Genet. 3(3):154-61
- Liu F., et al. (2009). Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19:R192–R193

Prediction of Eye Color



Liu F., et al. (2009). Eye color and the prediction of complex phenotypes from genotypes, Curr. Biol. 19:R192–R193

Prediction of Hair Color

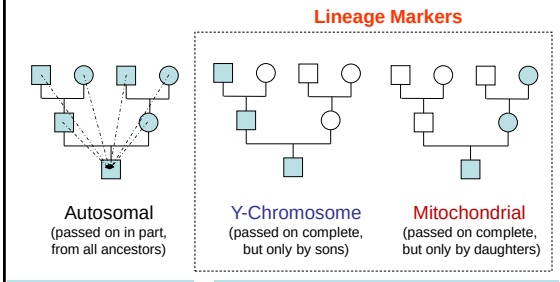


Branicki W, Kayser M et al. (2011). Model-based prediction of human hair color using DNA variants. *Human Genetics;* DOI 10.1007/s00439-010-0939-8

Lineage Informative SNPs

- Much of the LISNP literature focused on Y-chromosome and mitochondrial DNA SNPs
- With genome-wide arrays, autosomal SNP typing for lineage analysis is possible
- Looking for blocks of DNA that have been transmitted unchanged from one generation to the next
- Useful in evolutionary studies and kinship analysis

Different Inheritance Patterns



Can infer a specific relationship

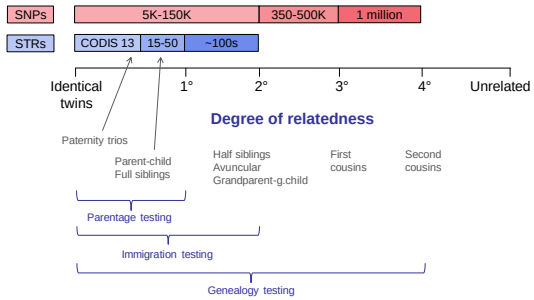
Cannot define a specific relationship, only paternal (Y) or maternal (mito) relatedness

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

Kinship Inference with Autosomal STRs or SNPs


Pete Vallone


Kristen O'Connor



Low Template DNA Testing


Becky Hill

Some Definitions of Low Template (LT) DNA

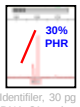
- Working with **<100-200 pg genomic DNA**
- Considered to be data below stochastic threshold level where PCR amplification is not as reliable (determined by each laboratory; typically 150-250 RFUs)
- Enhancing the sensitivity of detection (increasing PCR cycles, PCR product clean-up, increasing CE injection/voltage)
- Having too few copies of DNA template to ensure reliable PCR amplification (allelic or full locus drop-out)
- Can often be the minor component of mixture samples consisting of low level DNA template amounts

Stochastic (Random) Effects with LT-DNA

When Combined with Higher Sensitivity Techniques

Loss of True Signal (***False Negative***)

Heterozygote Peak Imbalance



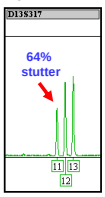
Identifiler, 30 pg DNA, 31 cycles

Allelic **Drop-out**

Identifiler, 30 pg DNA, 31 cycles

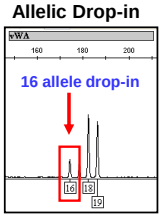
Gain of False Signal (***False Positive***)

Higher Stutter



Identifiler, 10 pg DNA, 31 cycles

Allelic Drop-in

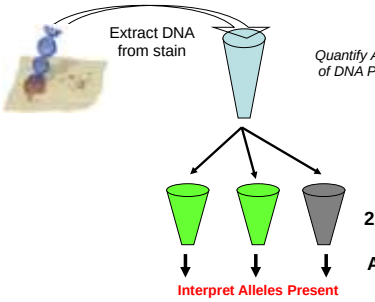


Identifiler, 10 pg DNA, 31 cycles

Suggestions for Optimal Results with LT-DNA

- Typically at least 2 – 3 PCR amplifications from the same DNA extract are performed to obtain consensus profiles
- An allele cannot be scored (considered real) unless it is present at least twice in replicate samples
- Extremely sterile environment is required for PCR setup to avoid contamination from laboratory personnel or other sources

Replicate Testing and Consensus Profiles



Develop a Consensus Profile
(based on replicate consistent results)

Replicate LT-DNA Test Results from FSS

Gill, P. (2002) Role of short tandem repeat DNA in forensic casework in the UK--past, present, and future perspectives. *BioTechniques* 32(2): 366-385.

Table 2. Results of Six Replicate PCR Tests of a Sample Under Low Copy Number Analysis Conditions Compared to the Control Sample

	Amelo	D18	D3	D8	THO	VWA	D21	FGA	D16	D18	D2
CONTROL	X X	14,14	16,16	10,10	7,9.3	16,19	28:32.2	20,23	9,12	12,16	17,23
Sample											
1	–	14 F	–	13 F	–	–	28:32.2	20 F	–	16 F	–
2	X F	–	16 F	13 F	–	19 F	–	–	12 F	–	–
3	X F	–	–	13 F	–	–	–	–	–	–	17 F
4	X F	14 F	16 F	–	–	–	–	–	9,12	–	–
5	X F	–	16 F	–	–	16 F	–	–	–	–	–
6	X F	14 F	–	–	–	19 F	28:32.2	20 F	–	12 F	–
Consensus	X F	14 F	16 F	13 F	–	19 F	28:32.2	20 F	12 F	–	–

The consensus result is reported, provided that an allele is observed at least twice. If only one allele is observed, then an F designation is given to denote the possibility of allele drop-out.

F' used to designate that allele drop-out of a second allele cannot be discounted when only a single allele is observed

Discussion

Response to Comment on "Low copy number typing has yet to achieve "general acceptance"" (Budowle et al., 2009, Forensic Sci. Int. Genetics: Supplement Series 2, 551–552) by Theresa Caragine, Mechthild Prinz

Bruce Budowle[a,b,*], Arthur J. Eisenberg[a,b], Angela van Daal[c]

Discussion

Further Comment on "Low copy number typing has yet to achieve "general acceptance" " by Budowle, B., et al, 2009. Forensic Sci. Int. Genetics: Supplement Series 2, 551–552

John Buckleton[a], Peter Gill[b,c]

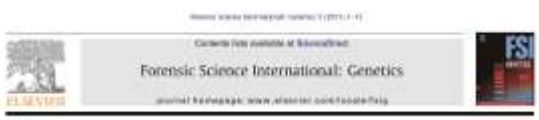
Letter to the Editor

Reply to Comments by Buckleton and Gill on "Low copy number typing has yet to achieve 'general acceptance'" by Budowle, B., et al., 2009. Forensic Sci. Int.: Genet. Suppl. Series 2, 551–552

Discussion

Comment on "A universal strategy to interpret DNA profiles that does not require a definition of low copy number" by Peter Gill and John Buckleton, 2010, Forensic Sci. Int. Genetics 4, 221–227

Bruce Budowle[a,*], Angela van Daal[b]



Editorial

Publications and letters related to the forensic genetic analysis of low amounts of DNA

"With the publication of this editorial, it is announced that the journal will not accept any further letters on this issue. However, publications are very welcome describing original scientific research contributing new data on the issues of validation and interpretation of results obtained from forensic genetic typing of DNA of low quantity and/or quality."

Next Generation Sequencing

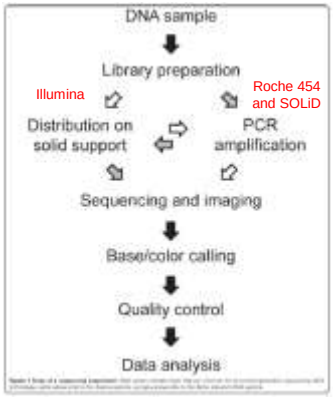
REVIEW Open Access

Next-generation sequencing technologies and applications for human genetic history and forensics

Ina C Borglund, Anna Wadenvik and Ann-Christine Syvänen*

Trends and upcoming technologies

- **Low-capacity sequencing systems**
- **Single molecule sequencing**
- **Bioinformatic analysis of sequence data**
- **Computer resources**



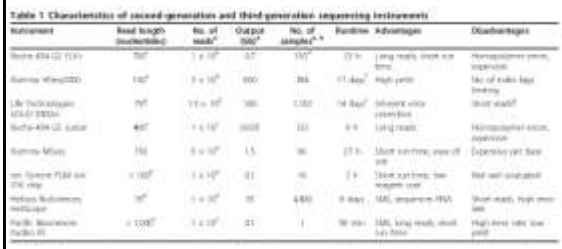


Table 1 Characteristics of second-generation and third-generation sequencing instruments

Instrument	Read length (nucleotides)	No. of reads[a]	Output (Gb)[a]	No. of samples[a,b]	Runtime	Advantages	Disadvantages
Roche 454 GS FLX+	700[a]	1×10^6	0.7	1[a]	23 h	Long reads, short run time	Homopolymer errors, expensive
Illumina HiSeq2000	150[a]	6×10^9	600	384	11 days[a]	High yield	No. of index tags limiting
Life Technologies SOLiD 5500xl	75[a]	3.5×10^9	300	1,152	14 days[a]	Inherent error correction	Short reads[a]
Roche 454 GS Junior	400[a]	1×10^6	0.035	1[c]	9 h	Long reads	Homopolymer errors, expensive
Illumina MiSeq	150	5×10^6	1.5	96	27 h	Short run time, ease of use	Expensive per base
Ion Torrent PGM 316 chip	< 100[a]	1×10^6	0.1	96	2 h	Short run time, low reagent cost	Not well evaluated
Helicos BioSciences HeliScope	30[a]	1×10^9	35	4,800	8 days	SMS, sequences RNA	Short reads, high error rate
Pacific Biosciences PacBio RS	> 1,000[a]	1×10^5	0.1	1	90 min	SMS, long reads, short run time	High error rate, low yield

NGS Studies with the Ychr

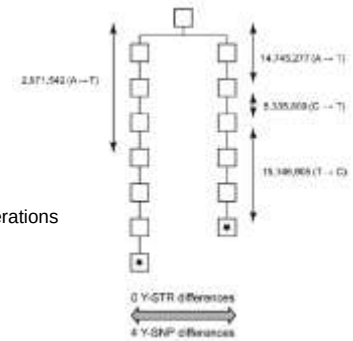
Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Review

The hare and the tortoise: One small step for four SNPs, one giant leap for SNP-kind

Yali Xue, Chris Tyler-Smith *



13 Generations

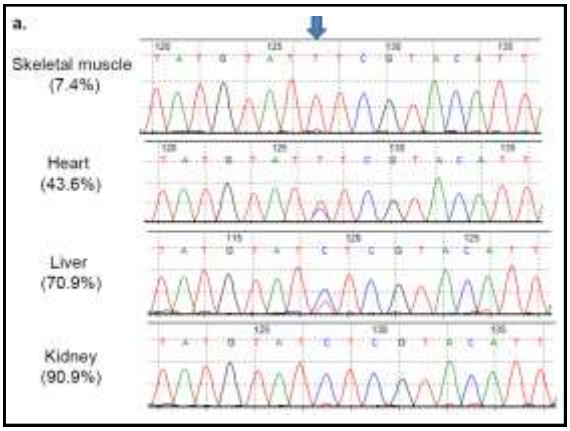
Xue and Tyler-Smith (2010)

NGS Studies with mtDNA

LETTERS

Heteroplasmic mitochondrial DNA mutations in normal and tumour cells

Yiping He[1], Jian Wu[1], Devin C. Dressman[1], Christine Iacobuzio-Donahue[2], Sanford D. Markowitz[3], Victor E. Velculescu[3], Luis A. Diaz Jr[3], Kenneth W. Kinzler[3], Bert Vogelstein[3] & Nicholas Papadopoulos[1]



Because mtDNA template molecules are so numerous in comparison with nuclear DNA template molecules, they are also useful for forensic applications. Previous studies have shown variations in the length of mononucleotide tracts in mtDNA from hair roots compared with blood[29,30]. Our new results clearly show that heteroplasmies affect the entire mitochondrial genome, are common in normal individuals and vary markedly from tissue to tissue. Thus an individual, and perhaps even a single cell, does not have a single mtDNA genotype. Instead, tissues have a mixture of genotypes, a few of which may be maternally inherited and the remaining ones the result of somatic mutations. This suggests caution in excluding identity on the basis of a single or small number of mismatched alleles when the tissue in evidence (such as sperm) is not the same as the reference tissue of the suspect (such as blood or hair).

Position	Allele 1	Allele 2	Skeletal muscle	Lung	
60	C	T*	<0.35	<0.35	
64	A	C*	**1.73**	<0.35	N.D. by Sanger Sequencing
72	C	T*	<0.35	<0.35	
73	G	A*	**2.27**	<0.35	N.D. by Sanger Sequencing
74	G	T*	<0.35	<0.35	
→ 189	G	A*	**9.77**	0.37	
408	A	T*	**3.64**	<0.35	N.D. by Sanger Sequencing
1983	C	T*	<0.35	<0.35	
6078	C	A*	0.82	1.23	
8021	G	A*	<0.35	1.42	
11090	C	A*	**1.63**	1.12	N.D. by Sanger Sequencing
14274	C	A*	0.41	<0.35	
16092	C	T*	<0.35	<0.35	
→ 16093	C	T*	**7.44**	**73.0**	
Total number of heteroplasmic variants			6	1	

Patient #	Age	Position	Allele 1	Allele 2
1	66	60	C	T*
1	66	72	C	T*
1	66	94	A	G*
2	77	60	C	T*
2	77	72	C	T*
2	77	94	A	G*
4	50	72	C	T*
5	35	72	C	T*
6	53	72	C	T*
6	53	94	A	G*
8	64	72	C	T*
9	42	60	C	T*
9	42	72	C	T*
9	42	94	A	G*
10	59	60	C	T*
10	59	72	C	T*

60, 72, 94
(Artifacts?)

FSI-Genetics, 6(1): 143-145

Short communication

Current Next Generation Sequencing technology may not meet forensic standards

Hans-Jürgen Bandelt[a,b,*], Antonio Salas[a,b]

[a] Department of Mathematics, University of Hamburg, 249 ath Hamburg, Germany
[b] Unidade de Xenética, Instituto de Medicina Legal, Facultade de Medicina, and Departamento de Anatomía Patolóxica e Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, 15782 Galicia, Spain

Patient 4

T 709 1888 4917 8697 10463 13368 14905 15607 15928 16294

4	72	T	C
4	73	A	G
4	200	A	G
4	263	A	G
4	709	G	A
4	750	A	G
4	1420	T	C
4	1438	A	G
4	1888	G	A
4	2141	T	C
4	4769	A	G
4	6249	G	A
4	6524	T	C
4	7028	C	T
4	8860	A	G
4	9137	T	C
4	11251	A	G
4	11719	G	A
4	11764	A	G
4	11812	A	G
4	12741	C	T
4	13368	G	A
4	14233	A	G
4	14687	A	G
4	14766	C	T
4	14905	G	A
4	15452	C	A
4	16126	T	C
4	16296	C	T
4	16324	T	C
4	16519	T	C
4	10726	G	A

Table 1
Deficiencies in the mtDNA sequences reported by He et al. [6].

Sample	HG[a]	Missed mutations (characterizing haplogroup)[b]
Patient 1	J1c3a1	15326 (H2a2), 2706 (H), 4216 (R2'T), 3010[d] (J1), 13934 (J1c3)
Patient 2	J1b1a	15326 (H2a2), 2706 (H), 4216 (R2'T), 3010[d] (J1), 16222 (J1b), 16261[c] (J1b), 5480[d] (J1b1)
Patient 3	J1c or J1c3a	15326 (H2a2), either 2706 (H) or 13934 (J1c3), 4216 (R2'T), 3010[d] (J1)
Patient 4	T2a1b1	15326 (H2a2), 2706 (H), 4216 (R2'T), 4917 (T), 8697 (T), 10463 (T), 15607 (T), 15928 (T), 16294[c] (T), 13965 (T2a), 13966 (T2a1b1)
Patient 5	N22	15326 (H2a2), 2706 (H), 16223[c] (R), 042 (N22), 16249[c] (N22)
Patient 6	U5a1	15326 (H2a2), 2706 (H), 13617 (U5), 16270 (U5), 16256[c] (U5a)
Patient 7	X2a2	15326 (H2a2), 2706 (H), 7028 (H), 16223[c] (R), 153 (X), 13966 (X), 16278[c] (X), 1719[d] (X2), 12397 (X2a'J), 8913 (X2a), 16213 (X2a)
Patient 8	J1c3a	15326 (H2a2), 2706 (H), 4216 (R2'T), 3010[d] (J1), 13934 (J1c3)
Patient 9	J1c	15326 (H2a2), 2706 (H), 4216 (R2'T), 3010[d] (J1), 185[c] (J1c)
Patient 10	H7c	15326 (H2a2), 4793 (H7)
CEPH 45₁	T2b3	15326 (H2a2), 1438 (H2), 14905 (T)
CEPH 45₂	H	15326 (H2a2), 1438 (H2)
CEPH1377₁	T1	15326 (H2a2), 1438 (H2), 14905 (T)
CEPH1377₂	K1b1a	15326 (H2a2), 1438 (H2), 152[c] (K1b1a)

FSI-Genetics, in press

Short communication

Current Next Generation Sequencing technology may not meet forensic standards

Hans-Jürgen Bandelt[a,b,*], Antonio Salas[a,b]

[a] Department of Mathematics, University of Hamburg, 249 ath Hamburg, Germany
[b] Unidade de Xenética, Instituto de Medicina Legal, Facultade de Medicina, and Departamento de Anatomía Patolóxica e Ciencias Forenses, Facultade de Medicina, Universidade de Santiago de Compostela, 15782 Galicia, Spain

"Before one can really set out to access to entire mtDNA genome data with relative ease for forensic purposes, one needs careful calibration studies under strict forensic conditions—or might have to wait for another generation."

Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy

Mitchell M. Holland[a],
Megan R. McQuillan[a],
Katherine A. O'Hanlon[a]

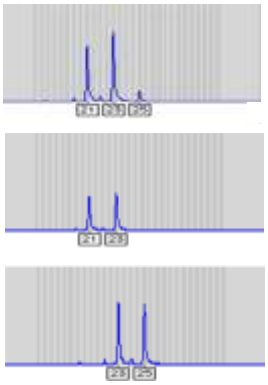
[a] Forensic Science Program, The Pennsylvania State University, University Park, Pa., USA

[b] National County Sheriffs Office, Forensic Laboratory, Normal, Illinois, USA

ONLINE
CMJ
CROATIAN MEDICAL JOURNAL

"The pyrosequencing approach results in poor resolution of homopolymeric sequences, and PCR/sequencing artifacts require a filtering mechanism similar to that for STR stutter and spectral bleed through. In addition, chimeric sequences from jumping PCR must be addressed to make the method operational."

DNA Mixture Interpretation

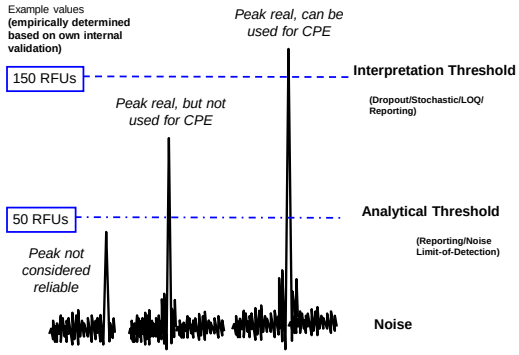


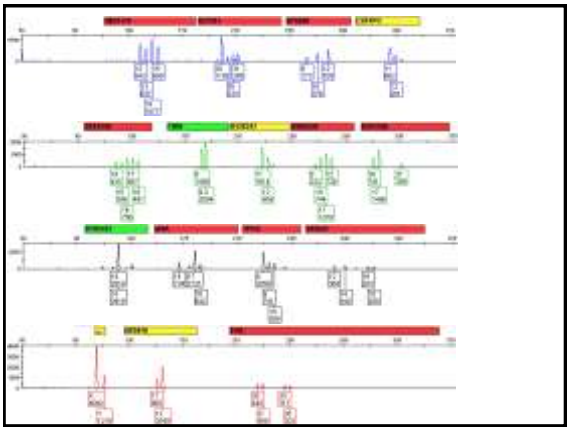
Evidence

Victim

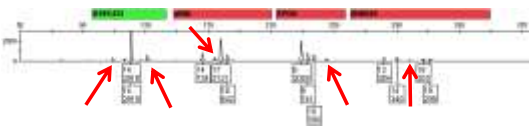
Suspect?

Different Thresholds

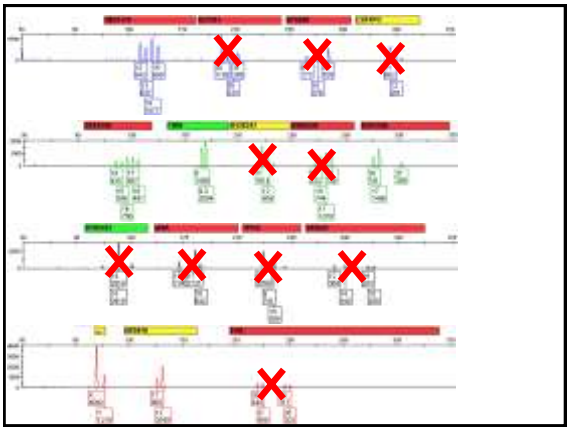




Alleles below the Stochastic Threshold



4.6.3. When using CPE/CPI (with no assumptions of number of contributors) to calculate the probability that a randomly selected person would be excluded/included as a contributor to the mixture, *loci with alleles below the stochastic threshold* ***may not be used*** for statistical purposes to support an inclusion.



Available Loci (CPI Stats)

- D8S1179 PI = 0.5927
- D3S1358 PI = 0.9704
- D2S1338 PI = 0.0658
- D5S818 PI = 0.5550

- CPI = 0.021

~2.1% of the Hispanic population can be included in this mixture

Expert Software for Mixture Analysis

ID-X Software

FSS-i3

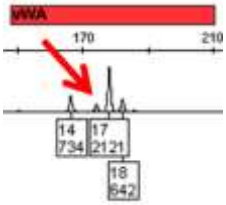
GeneMarker® HID

NicheVision

Software Limitations…

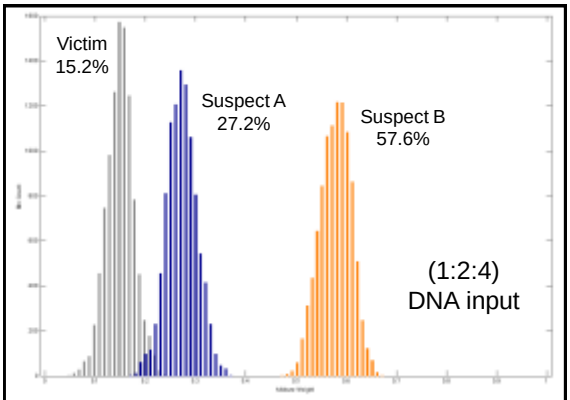
The "16" allele is considered an artifact (stutter) peak from allele 17 and is ignored.

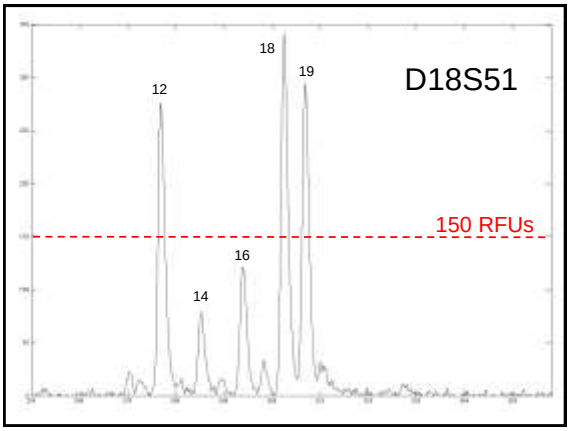
This marker is included in the CPI calculation.

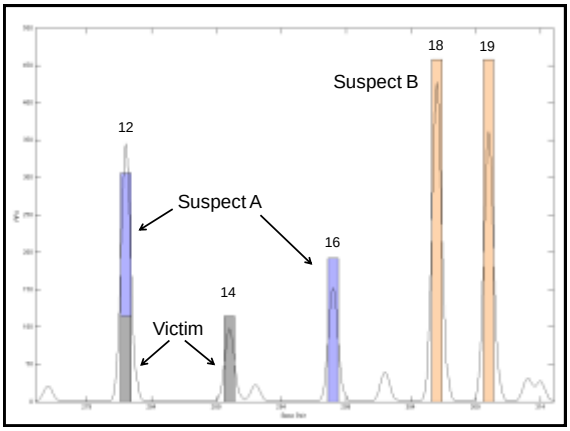


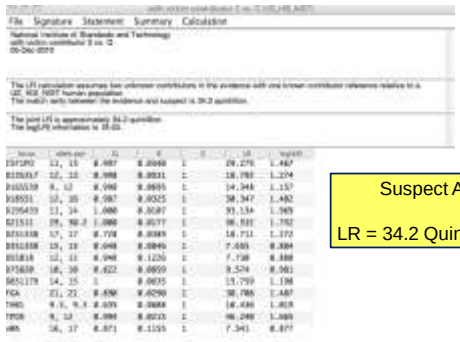
True Allele Software

"Markov Chain Monte Carlo Testing"





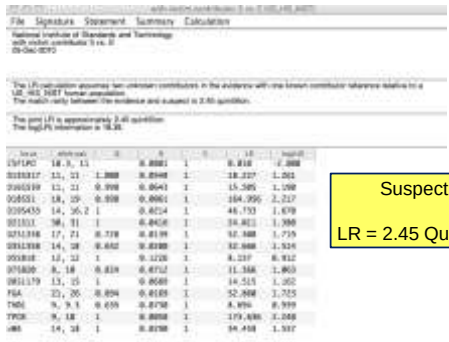




Suspect A

LR = 34.2 Quintillion

Million – Billion – Trillion – Quadrillion - Quintrillion



Suspect A

LR = 2.45 Quintillion

Million – Billion – Trillion – Quadrillion - Quintrillion

Benefits of Increased Information

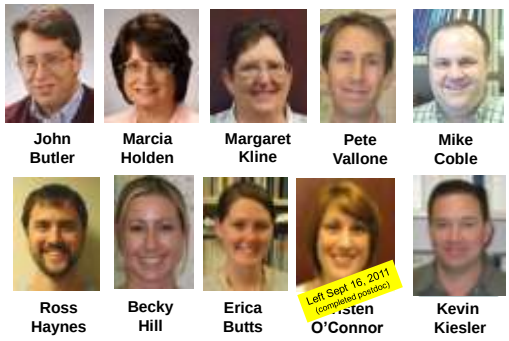
Manual Calculation (CPI)	Mixture Software (CPI)
1 in 2.1% included	1 in 2.6 million included

True Allele Software
(LR)

2.45 Quintillion and 34.2 Quintillion

NIST Applied Genetics Group

Group Leader



John Butler | **Marcia Holden** | **Margaret Kline** | **Pete Vallone** | **Mike Coble**

Ross Haynes | **Becky Hill** | **Erica Butts** | **Kristen O'Connor** | **Kevin Kiesler**

Questions?