STRs vs. SNPs:

Thoughts on the Future of Forensic DNA Testing

John M. Butler
Michael D. Coble and Peter M. Vallone

U.S. National Institute of Standards and Technology

International Symposium on the Forensic Sciences (ANZFSS 2006)
April 4, 2006 – Fremantle, Australia

Presentation Outline

- Why consider SNPs for human identity testing?
- Work with SNPs at NIST
- Recent work with SNPs by others
- Direct comparisons of SNPs and STRs
- miniSTR work at NIST
- Score card: SNPs vs. STRs/miniSTRs

Reasons Often Given for Considering SNPs in Human Identity Testing…

- Use on degraded samples (WTC), low copy number, or telogenic (shed) hairs
- **Lower mutation rate** (Paternity testing)
- Easier data interpretation (no microvariants or stutter products)
- Amenable to high throughput analysis

Issues to be Addressed with SNPs

- **Power of Discrimination**
  - How many SNPs = 1 STR ?
- Multiplex-ability (robust 50plex < 1ng DNA ?)
- Population databases
- Many different platforms for SNP typing
- Unique interpretation issues – mixtures
- Validation
- Sensitivity
- Assay cost

Possible Allele Combinations

STRs

ATGCTA**(GATA)$_n$**GACTAC

Alleles	Genotypes
7	7,7
8	7,8 8,8
9	7,9 8,9 9,9
10	7,10 8,10 9,10
11	7,11 8,11 9,11
12	7,12 8,12 9,12
13	7,13 8,13 9,13
14	7,14 8,14 9,14
15	7,15 8,15 9,15

. . .

45 possible genotypes

SNPs

ATGCTA**(C/T)**GACTAC

Alleles	Genotypes
C	CC
T	TT
	CT

3 possible genotypes

DQA1 + PolyMarker

SNP Typing Platforms

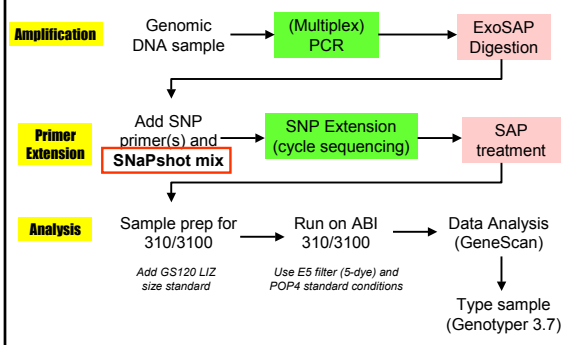
- RT-PCR (TaqMan, Light Cycler, Molecular Beacon)
- ASPE (SNaPshot, Orchid UHT, MALDI, FP)
- Mass Spectrometry (Electrospray)
- Sequencing
- Flow Cytometry (Luminex)
- Pyrosequencing
- Ligation (SNPplex, Illumina)
- Invader assay
- ARMS assay (FSS)
- RFLP

Sensitivity, multiplexing, accurate typing

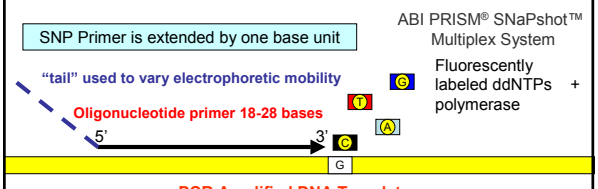
ASPE = allele-specific primer extension

Budowle 2004 FSI 139-142; Sobrino et al., 2005 FSI (epub); Dixon et al., 2005 FSI (epub)

Protocol with SNaPshot™ "Kit"



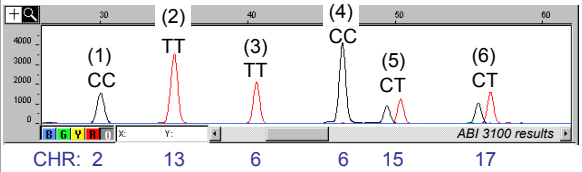
Allele-Specific Primer Extension



ddNTP	Dye label	Color
A	dR6G	Green
C	dTAMRA	Black
G	dR110	Blue
T	dROX	Red

25 Cycles
96°C 10s
50°C 5s
60°C 30s

6-plex SNP Assay Using SNaPshot



Extension primers for 6-plex

(1) TTTTTAGCTCCTAATTTCTTGATGGG
(2) TTTTTTTTTTCATCTGATGCCATGAGAAAGC
(3) TTTTTTTTTTTTTTGTTCTGCTTTAATACAAAACCAG
(4) TTTTTTTTTTTTTTTTTTTTTTTTATAAAGGGCAGAATGAGGATTA
(5) TTTTTTTTTTTTTTTTTTTTTTTTTAGAAAGTATCTTGCAAAAGGTCCA
(6) TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATAATCACAGCTTTTTTCTCCCAA

Utility of SNP Markers

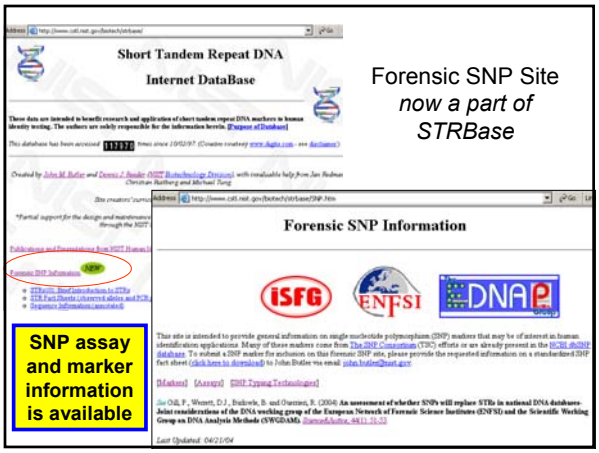
Replace Autosomal STRs?

"It is unlikely that SNPs will replace STRs as the preferred method of testing of forensic samples in the near to medium future."

Specialized applications

mtDNA – coding region and linear arrays
Y-SNPs – lineage, population study, sample discrimination
Autosomal SNPs – highly degraded samples, shed hairs, physical characteristics, ethnic/geographical determination

Gill, P., Werret, D.J., Budowle, B., and Guerreri, R. (2004) *Science & Justice* 44: 51-53

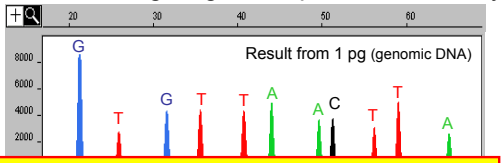


Forensic SNP Site
now a part of STRBase

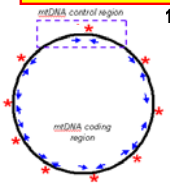
SNP Work at NIST

- mtDNA coding region SNPs
  - Will selected SNPs aid resolution of common HV1/HV2 types?
- Y-SNPs
  - Potential for ethnicity prediction?
- Autosomal SNPs
  - Multiplexing limitations?
  - Capability of mixture detection/interpretation?
  - Performance with degraded DNA or LCN templates?

mtDNA Coding Region 11-plex ASPE Assay



Now being used by Armed Forces DNA Identification Laboratory (AFDIL) for mtDNA casework



11-plex PCR and 11-plex SNP detection

Sites are polymorphic in Caucasians (H1) and useful in resolving most common HV1/HV2 types

Multiplex PCR used to co-amplify all regions of interest at once

PCR product sizes kept under 200 bp to enable success with degraded DNA samples

Vallone *et al.* (2004) *Int J Legal Med* 118: 147-157

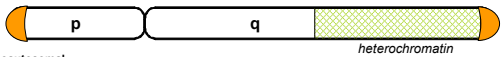
Forensic Utility of Y Chromosome SNPs

Y chromosome markers are useful in mixed male - female samples

Haplogroups are non-randomly distributed among populations therefore potential exists for predicting population of origin

Low mutation rate of SNPs 2 x 10^{-8} per base per generation

Typed 51 Y-SNPs using ASPE and Marligen (Luminex beads)



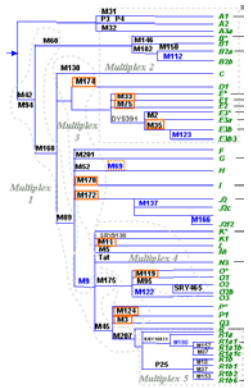
Pseudoautosomal region

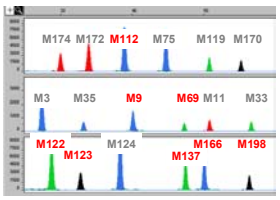
>250 Y-SNPs described

Pseudoautosomal region

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

Typing 51 Y-SNPs

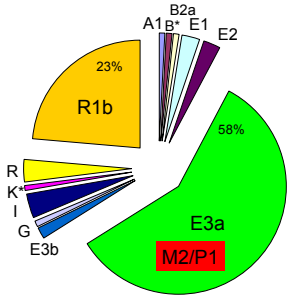




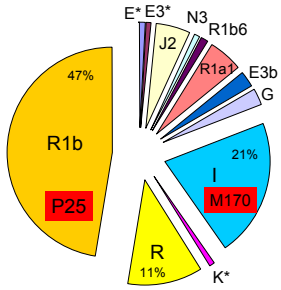
51 Y-SNPs
115 African Americans
114 Caucasians

Concordant typing results for over 19 loci (>3,800 allele calls)

Y-SNP haplogroups for 115 African Americans



Y-SNP haplogroups for 114 Caucasians



18 of 46 different haplogroups observed in 229 males

Vallone and Butler (2004) *J Forensic Sci.* 49: 723-32

Variation was observed in 24 of the 51 Y-SNPs

Publication on U.S. Groups with Y-SNPs

J. Forensic Sci. 2004; 49(4): 723-732

J Forensic Sci, July 2004, Vol. 49, No. 4
Paper ID JFS2003303
Available online at: www.astm.org

Peter M. Vallone,[1] Ph.D. and John M. Butler,[1] Ph.D.

Y-SNP Typing of U.S. African American and Caucasian Samples Using Allele-Specific Hybridization and Primer Extension*

Different technologies yield the same Y-SNP type
Full concordance was observed between hybridization and primer extension technologies on 18 different Y-SNPs (>3,800 allele calls)

Y-SNPs will have limited value for individualizing a sample
18 different types observed in 229 individuals

Current Y-SNPs appear to have limited value for ethnic differentiation in U.S. populations (with the exception of M2 that is only found in African Americans and not in Caucasians)

Forensic Utility
51 Y-SNPs versus 1 Y-STR

For N = 211 male samples

	51Y-SNPs	Y-STR DYS464
Amount of sample consumed	10ng	1ng
Number for types observed	**18**	**62**
Analysis	Multiple	1 reaction
Degraded samples	+	?

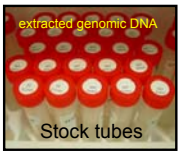
As a stand alone forensic assay
1 Y-STR is better than 51 Y-SNPs

Standard U.S. Population Dataset

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

260 Caucasians, 260 African Americans, 140 Hispanics, 3 Asians = **663 males**

DNA extracted from whole blood (anonymous; self-identified ethnicities) received from Interstate Blood Bank (Memphis, TN) and Millennium Biotech Inc. (Ft. Lauderdale, FL)



To date: (>100,000 allele calls)

Identifiler (15 autosomal markers + Amelogenin) (10,608)
Roche Linear Arrays (HV1/HV2 10 regions) (6,630)
Y STRs 22 loci—27 amplicons (17,388)
Y STRs 27 new loci (14,535)
Yfiler kit 17 loci (11,237)
Y SNPs 50 markers on sub-set of samples (11,498)
Orchid 70 autosomal SNPs on sub-set (13,230)
miniSTR testing-new loci and CODIS concordance (9,228)
New miniSTR loci – for 11 loci, 7,293 genotypes
mtDNA full control region sequences by AFDIL

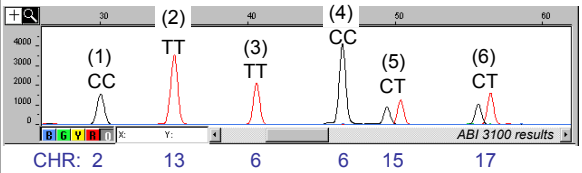
Genotypes with various human identity testing markers

Autosomal SNP characteristics

- 70 Loci – sites from Orchid – C/T bi-allelic
- Present on 20 of 22 autosomal CHR (3,16,X,Y)
- Amplicon size range 59 - 108 bp (average 69)
- Markers are typed by allele-specific primer extension assays (ABI SNaPshot)
- Level of multiplexing (6- 12-plexes)
- Web page for SNP site info

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

6-plex SNP Assay

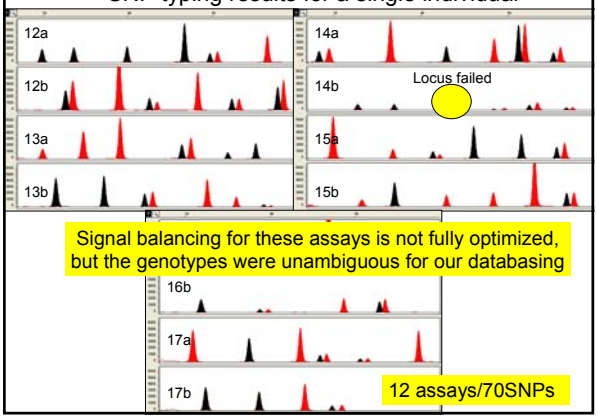


CHR: 2 13 6 6 15 17

Extension primers for 6-plex

(1) TTTTTAGCTCCTAATTTCTTGATGGG
(2) TTTTTTTTTTCATCTGATGCCATGAGAAAGC
(3) TTTTTTTTTTTTTTTGTTCTGCTTTAATACAAAACCAG
(4) TTTTTTTTTTTTTTTTTTTTTTTTATAAAGGGCAGAATGAGGATTA
(5) TTTTTTTTTTTTTTTTTTTTTTTTTAGAAAGTATCTTGCAAAAGGTCCA
(6) TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCATAATCACAGCTTTTTTCTCCCAA

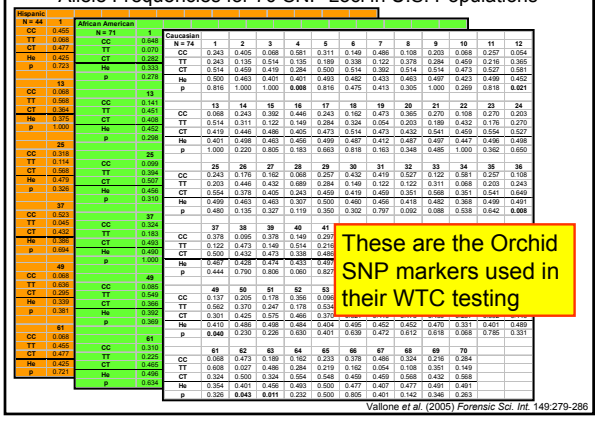
SNP typing results for a single individual



Signal balancing for these assays is not fully optimized, but the genotypes were unambiguous for our databasing

12 assays/70SNPs

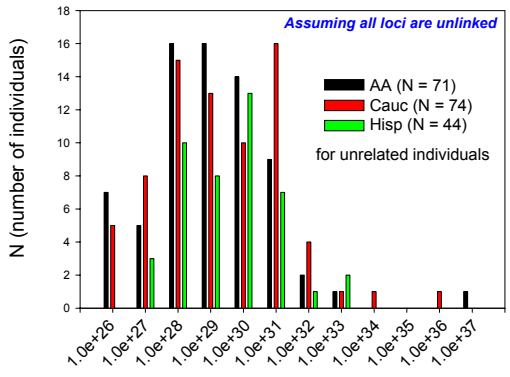
Allele Frequencies for 70 SNP Loci in U.S. Populations



These are the Orchid SNP markers used in their WTC testing

Vallone *et al.* (2005) *Forensic Sci. Int.* 149:279-286

Probability of a Random Match using 70 SNPs



SNP Assay Results

70 were typed for 189 U.S. samples (self identified ethnicities)
74 Caucasians + 71 African Americans AA + 44 Hispanics

Total of 13,230 possible genotypes

42 Samples were re-injected to confirm ambiguous results (99.7 %) success rate on first pass

Results described in manuscript *(Vallone, P.M., Decker, A.E., Butler, J.M. (2005) Forensic Sci. Int.,* 149:279-286*)*

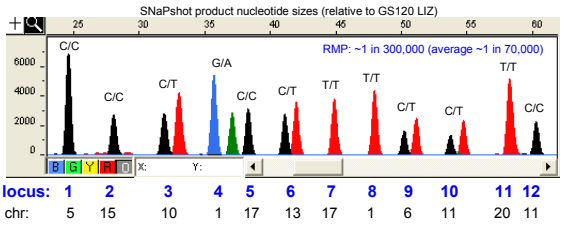
"Best" 12 loci combined into a 12plex SNP assay

Vallone et al. (2006) *Progress in Forensic Genetics, in press*

Autosomal 12plex SNP Assay

12plex PCR and 12plex ASPE from 0.5 ng DNA template

all PCR products <90 bp



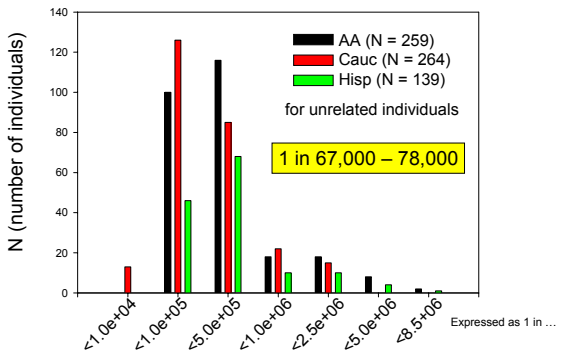
locus:	1	2	3	4	5	6	7	8	9	10	11	12
chr:	5	15	10	1	17	13	17	1	6	11	20	11

"Best" 12 SNPs selected from 70 originally tested;
>0.45 heterozygosity in U.S. Caucasians, Hispanics, and African Americans

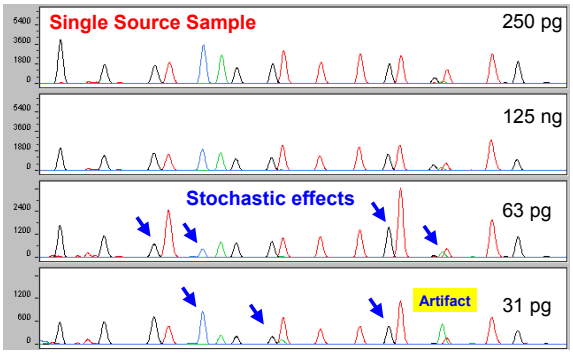
Vallone, P.M., Decker, A.E., Butler, J.M. (2005) Allele frequencies for 70 autosomal SNP loci with U.S. Caucasian, African American, and Hispanic Samples. *Forensic Sci. Int.* 149:279-286.

Population data has been collected on >1,000 samples (662 U.S. and 375 world samples) **from 10 different population groups**

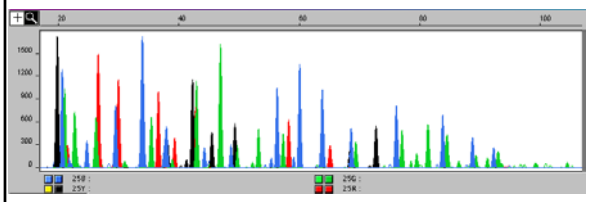
Probability of a Random Match using NIST 12plex



Sensitivity Study with NIST 12plex

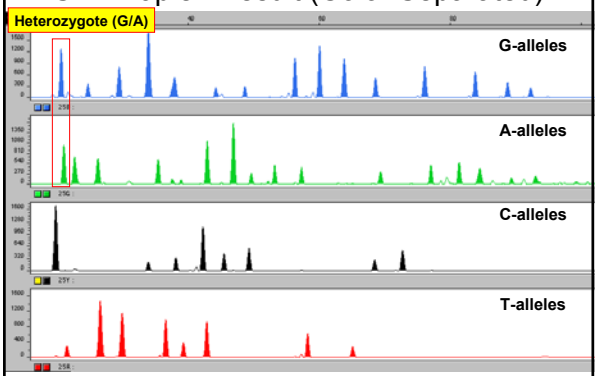


29plex SNP Assay Developed by SNP*for*ID
(EDNAP Study Results Generated at NIST)



Single source sample – yet challenging to interpret quickly…

SNP 29plex Result (Color Separated)

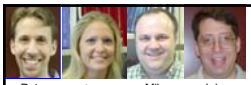


NIST AA03 result from EDNAP study (SNPforID 29plex assay)

NIST Work Conclusions

- mtSNPs: Coding region SNPs can fulfill a useful role for separating common HV1/HV2 mitotypes.
- Y-SNPs: Y-SNPs will have limited utility for individualizing a sample. Determination of ethnic origin may be challenging for U.S. samples.
- Autosomal SNPs: 12plex assay shows some promise for typing degraded samples and shed hairs. Would be used in conjunction with STR kits rather than in place of them.

NIST Work with SNP Loci



Pete Vallone | Amy Decker | Mike Coble | John Butler

- **mtDNA coding region SNP** 11plex assay
  - Vallone et al. (2004) *Int. J. Legal Med.* 118: 147-57
- U.S. population information with 50 **Y-SNPs**
  - Vallone *et al.* (2004) *J. Forensic Sci.* 49: 723-732
- U.S. population frequencies with 70 **autosomal SNPs**
  - Vallone *et al.* (2005) *Forensic Sci. Int.* 149: 279-286
- Construction of 12plex autosomal SNP assay
  - Vallone *et al.* (2006) *Progress in Forensic Genetics 11*

>1,000 samples examined from 10 populations

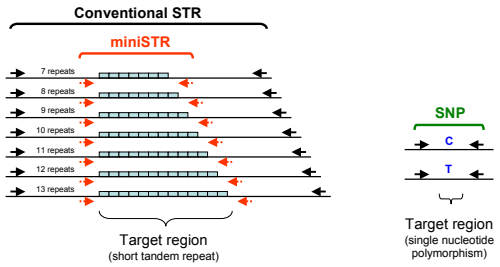
- Creation of Forensic SNP Information website on STRBase
  - see Gill et al. *Science&Justice* 44(1): 51-53

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

(Selected) SNP Work by Others

- **SNP*for*ID** (http://www.snpforid.org) – team of five European labs working on selection of useful autosomal SNP loci, developing multiplex assays, and collecting population data
- Manfred Kayser's group – seeking minimal set of SNP loci for distinguishing ethnicities
  AJHG 2006 78: 680-690
- Ken Kidd's group – seeking "best" forensic SNPs
  FSI article, in press

Comparison of STRs and SNPs



Larger target region (miniSTR targets same region)
More possible variants than SNPs
Only need a moderate number of STR markers
Range of sizes examined (e.g., 28 bp spread if 4 bp/repeat)
Mixtures easier to detect and decipher

Smaller target region
Fewer possible variants
Need more SNP markers
Constant size examined

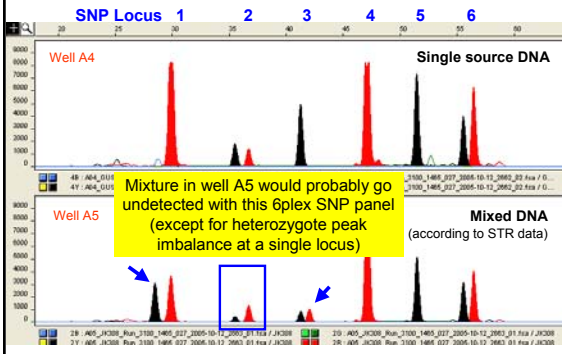
STR Markers Ability to Detect DNA Mixtures

Green channel from Identifiler STR tests

All 5 loci shown exhibit 3 alleles suggesting that a DNA mixture from at least two individuals is present

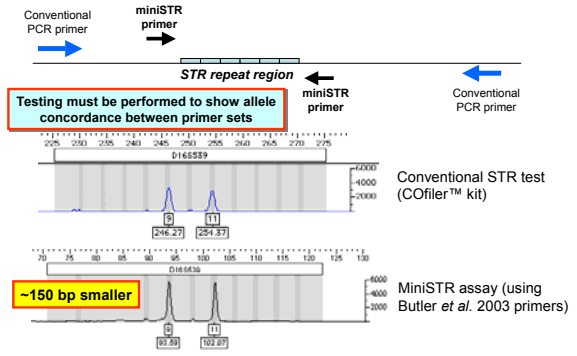
STRs permit fairly easy mixture identification due to the number of possible alleles and the high heterozygosities of loci used for human identity testing

SNP 6plex with the Same DNA Samples



SNP panel 15a

A miniSTR is a reduced size STR amplicon that enables higher recovery of information from degraded DNA samples

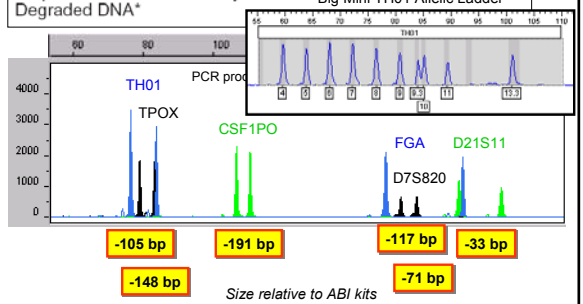


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 7.2, ©Elsevier Science/Academic Press

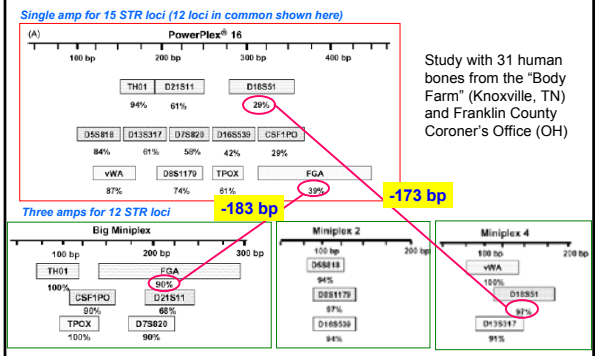
J. Forensic Sci. Sept 2003 issue

The Development of Reduced Size STR Amplicons as Tools for Analysis of Degraded DNA*

Describes new primer sequences for all CODIS loci and initial assays developed



Comparison of PCR Amplification Success Rates with Commercial Kit vs. miniSTR Assays



Opel K. L.; Chung, D. T.; Drábek, J.; Tatarek, N. E.; Jantz, L. M.;. McCord, B.R. (2006) The application of miniplex primer sets in the analysis of degraded DNA from human skeletal remains, *J. Forensic Sciences* 51(2): 351-356

miniSTRs

Advantages

- Better success with degraded DNA (compared to larger PCR products present in commercial STR kits)
- Better success with low amounts of DNA (due to more efficient PCR amplification compared to larger PCR products)
- Better capacity for handling mixed DNA samples than SNPs (due to more alleles being possible)
- Concordance to STR loci in commercial kits is possible

Disadvantages

- Not all commonly used STRs can be made significantly smaller—thus **new loci will be needed**
- Cannot multiplex as many loci due to size constraints
- **No commercial kit (yet)**
- STR flanking region mutations may make results disconcordant (e.g., D13 and VWA deletions)

Characterization of New miniSTR Loci

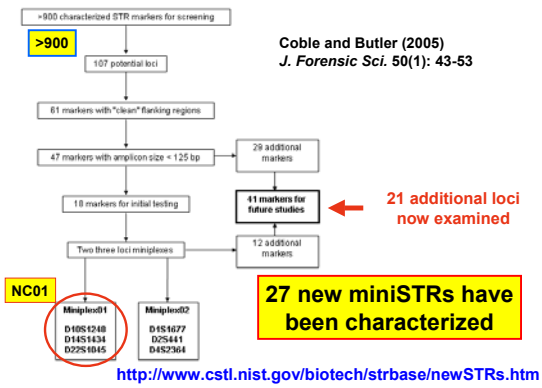
"Computer Work"

Candidate STR marker selection (e.g. Marshfield Clinic Center of Medical Genetics) → Pull down sequence data from the web (e.g. NCBI) → Identify Chromosome Location (e.g. Human BLAT Search) → Screen for PCR Primers (e.g. Primer3) → Test primers for Multiplex-ability (e.g. AutoDimer - NIST)

"Laboratory Work"

Test Markers on Population samples | Sequence homozygotes to determine allele sizes | Build Macros for Genotyping | Construct Allelic Ladders

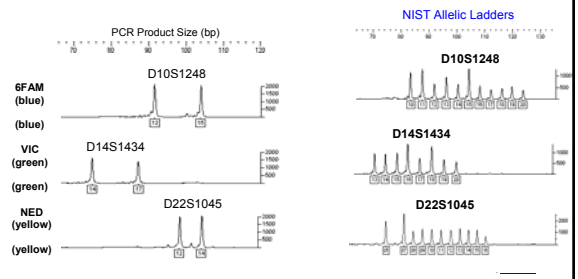
Initial Testing Results with Potential miniSTR Loci



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

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

Miniplex "NC01"



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

Coble and Butler (2005) Characterization of new miniSTR loci to aid analysis of degraded DNA *J. Forensic Sci.* 50(1): 43-53

Direct Comparisons of SNPs and STRs on Degraded DNA Templates

- World Trade Center DNA Investigation
  - A panel of 70 SNPs run on >15,000 bone extracts by Orchid Cellmark; no additional identifications made
  - **Reduced size STR markers (miniSTRs) aided last 20% of WTC DNA identifications**
- EDNAP Degraded DNA Study
  - Organized by Lindsey Dixon and Peter Gill (UK FSS); involved 9 labs testing artificially degraded blood and saliva stains
  - **miniSTRs outperformed SNPs**

EDNAP Exercise on Degraded DNA

ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Forensic Science International xxx (2005) xxx–xxx

Forensic Science International

www.elsevier.com/locate/forsciint

Analysis of artificially degraded DNA using STRs and SNPs—results of a collaborative European (EDNAP) exercise

L.A. Dixon [a,*], A.E. Dobbins [a], H.K. Pulker [a], J.M. Butler [b], P.M. Vallone [b], M.D. Coble [b], W. Parson [c], B. Berger [c], P. Grubwieser [c], H.S. Mogensen [d], N. Morling [d], K. Nielsen [d], J.J. Sanchez [d], E. Petkovski [e], A. Carracedo [f], P. Sanchez-Diz [f], E. Ramos-Luis [f], M. Brión [f], J.A. Irwin [g], R.S. Just [g], O. Loreille [g], T.J. Parsons [g], D. Syndercombe-Court [h], H. Schmitter [i], B. Stradmann-Bellinghausen [j], K. Bender [j], P. Gill [a]

MiniSTR primer mixes and allelic ladders were provided by NIST

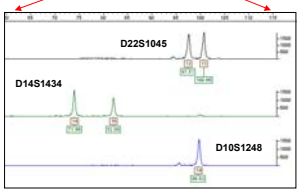
MiniSTR performance on degraded DNA samples

Blood Stain – 2 Weeks

Allelic drop out at D16 and FGA Failure at D18

SGM+ 32 cycles

NC01 32 cycles



Dixon *et al.*, *FSI, in press*

Recent Article Advocating miniSTRs

They recommend that miniSTRs "be adopted as the way forward to increase both the robustness and sensitivity of analysis."

ELSEVIER

Forensic Science International 156 (2006) 242–244

Forensic Science International

www.elsevier.com/locate/forsciint

Short communication

The evolution of DNA databases—Recommendations for new European STR loci

Peter Gill [a,*], Lyn Fereday [b], Niels Morling [c], Peter M. Schneider [d]

[a] Forensic Science Service, Birmingham, UK
[b] Forensic Science Service, London, UK
[c] Department of Forensic Genetics, Institute of Forensic Medicine, University of Copenhagen, Denmark
[d] Institute of Legal Medicine, University of Cologne, Germany

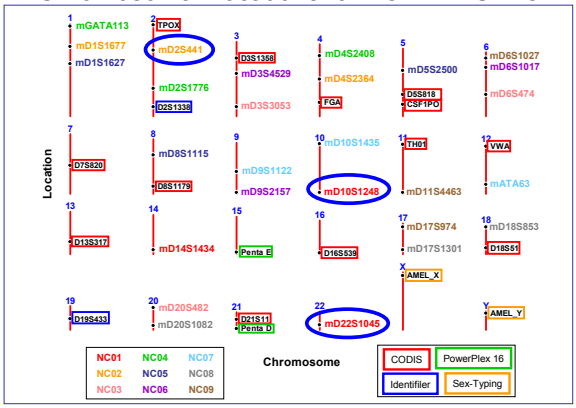
They recommend that European laboratories adopt three new mini-STR loci, namely: D10S1248, D14S1434 and D22S1045. (D14 now replaced by D2S441)

Comparison of STR Locus Variability

Locus	N	Heterozygosity (Overall)	Rank	Size Range (bp)
FGA	659	0.886	1	196 - 352 (ProPlus)
D2S1338	659	0.882	2	288 - 340 (SGM+)
D18S51	659	0.876	3	264 - 344 (ProPlus)
D9S2157	*661*	*0.844*	*4*	*71 - 101*
D21S11	659	0.844	5	186 - 244 (ProPlus)
ATA63 (D12)	*659*	*0.829*	*6*	*76 - 106*
vWA	659	0.826	7	152 - 212 (ProPlus)
D7S820	659	0.806	8	253 - 293 (ProPlus)
D19S433	659	0.803	9	106 - 140 (SGM+)
EDNAP/ ENFSI suggested markers D10S1248 (NC01)	663	0.792	10	83 - 123
D22S1045 (NC01)	*663*	*0.784*	*11*	*76 - 109*
D2S441 (NC02)	660	0.774	12	78 - 110
D8S1179	659	0.774	13	123 - 171 (ProPlus)
D16S539	659	0.766	14	233 - 273 (CoFiler)
D10S1435	663	0.766	15	82 - 139
D3S1358	659	0.763	16	97 - 145 (ProPlus)
D2S1776	654	0.763	17	127 - 161
D3S4529	660	0.761	18	111 - 139
D6S474	648	0.761	19	107 - 135
D5S2500	664	0.747	20	85 - 125
⋮	⋮	⋮	⋮	⋮
TPOX	659	0.707	34	213 - 249 (CoFiler)
D20S1082	*664*	*0.696*	*35*	*73 - 100*
D14S1434 (NC01)	663	0.696	36	70 - 98

<150 bp

Chromosomal Locations for New miniSTRs

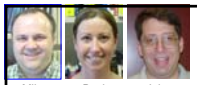


Characterization of New miniSTRs

Source	Number of Samples	Population
Coble and Butler (2005)	164	African Americans
Coble and Butler (2005)	170	U.S. Caucasians
Coble and Butler (2005)	140	U.S. Hispanics
Asamura et al. (2005)	142	Japanese
Yong et al. IJLM, (2006) in press	185	Chinese
Yong et al. IJLM, (2006) in press	182	Malaysian
Yong et al. IJLM, (2006) in press	178	East Indian
Prof. Ranieri Domenici, (2006) in press	100	Italian
Gergely Nagy, in preparation	200	Hungarian
Gergely Nagy, in preparation	200	Romanian
	226	Chinese
	238	Malaysian
	238	East Indian
	total = 2363	

2363 samples run with NC01 loci so far…

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



Mike Coble | Becky Hill | John Butler

miniSTRs for Degraded DNA

- Original miniSTR paper with CODIS loci, D2, D19, Penta D, Penta E
  - Butler *et al.* (2003) *J. Forensic Sci.* 48: 1054-1064
- Many CODIS loci are too big and make poor miniSTRs
- New miniSTRs and assays: **NC01**, NC02
  - Coble, M.D. and Butler, J.M. (2005) *J. Forensic Sci.* 50:43-53
- New **miniSGM** miniplex: AMEL, TH01, FGA, D18, D16, D2
- EDNAP/ENFSI degraded DNA study coordinated by Peter Gill
- Creation of miniSTR information on STRBase

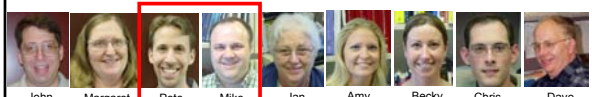
http://www.cstl.nist.gov/biotech/strbase/miniSTR.htm

Score Card

	STRs/miniSTRs	SNPs
Success with Degraded DNA	✓	✓
Power of Discrimination	✓	✗
Mixture Det./ Interpretation	✓	✗
Other Applications: Ethnicity Estimation, Physical Traits, etc.	✗	✓

Acknowledgments

Funding from interagency agreement 2003-IJ-R-029 between NIJ and the NIST Office of Law Enforcement Standards



John Butler | Margaret Kline | Pete Vallone | Mike Coble | Jan Redman | Amy Decker | Becky Hill | Chris DeAngelis | Dave Duewer

Past and Present Collaborators (also funded by NIJ):
Mike Hammer and Alan Redd (U. AZ) for Y-chromosome studies
Tom Parsons, Rebecca Just, Jodi Irwin (AFDIL) for mtDNA coding SNP work
Sandy Calloway (Roche) for mtDNA LINEAR ARRAYs
Bruce McCord and students (FL Int. U.) for miniSTR work
Marilyn Raymond and Victor David (NCI-Frederick) for cat STR work
Artie Eisenberg and John Planz (U. North Texas) for miniSTR testing on bones
Murray Brilliant (U. AZ) for phenotype markers
Ken Kidd (Yale U.) for SNP typing population samples