Forensic DNA Typing Workshop

John M. Butler, PhD
U.S. National Institute of Standards and Technology

http://www.fge.chiapas.gob.mx/congresoforense/default.asp

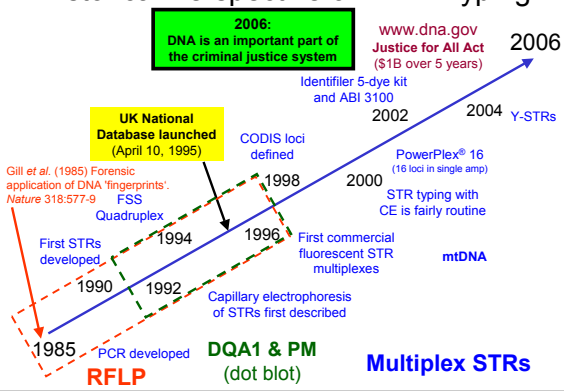
Presentation Outline

- DNA Analysis and STR Typing
  - Butler, J.M. (2006) Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci.* 51(2): 253-265.

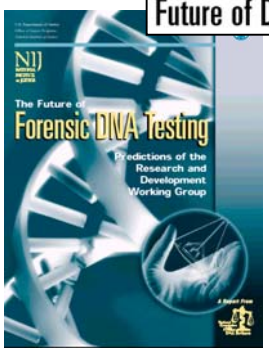
BREAK

- Capillary Electrophoresis
  - Butler, J.M., Buel, E., Crivellente, F., McCord, B.R. (2004) Forensic DNA typing by capillary electrophoresis: using the ABI Prism 310 and 3100 Genetic Analyzers for STR analysis. *Electrophoresis*, 25: 1397-1412.

Historical Perspective on DNA Typing



National Commission on the Future of DNA Evidence



http://www.ojp.usdoj.gov/nij/pubs-sum/183697.htm

- Report published in Nov 2000
- Asked to estimate where DNA testing would be 2, 5, and 10 years into the future

Conclusions

STR typing is here to stay for a few years because of DNA databases that have grown to contain millions of profiles

National DNA Index System (NDIS)

FBI LABORATORY

CODIS
COMBINED DNA INDEX SYSTEM

http://www.fbi.gov/hq/lab/codis/index1.htm

Combined DNA Index System (CODIS)

Launched in October 1998 and now links all 50 states
Used for linking serial crimes and unsolved cases with repeat offenders
Convicted offender and forensic case samples along with a missing persons index
Requires 13 core STR markers
>36,000 investigations aided nationwide as of June 2006
Contains more than 3.4 million DNA profiles

Applications for Human Identity Testing

- Crime solving – **matching suspect with evidence…**
- Accident victims – **after airplane crashes…**
- Soldiers in war – **who is the "unknown" soldier…**
- Paternity testing – **who is the father…**
- Inheritance claims – **who gets the money…**
- Missing persons investigations – **who's body…**
- Convicted felons databases – **cold cases solved…**

Involves generation of DNA profiles usually with the same core STR (short tandem repeat) markers and then MATCHING TO REFERENCE SAMPLE

DNA Testing Requires a Reference Sample

A DNA profile by itself is fairly useless because it has no context…

DNA analysis for identity only works by comparison – **you need a reference sample**

Crime Scene Evidence compared to **Suspect(s)** (Forensic Case)
Child compared to **Alleged Father** (Paternity Case)
Victim's Remains compared to **Biological Relative** (Mass Disaster ID)
Soldier's Remains compared to **Direct Reference Sample** (Armed Forces ID)

Basis of DNA Profiling

The genome of **each individual is unique** (with the exception of identical twins) and **is inherited from parents**

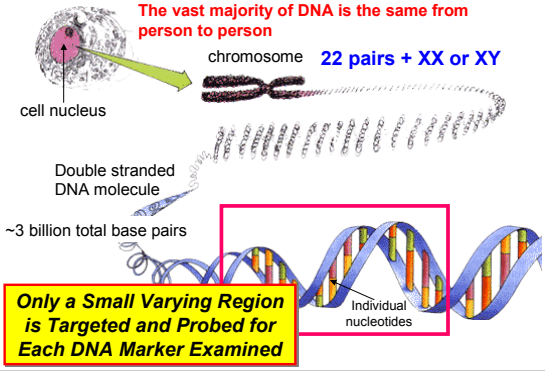
Probe subsets of genetic variation in order to differentiate between individuals (statistical probabilities of a random match are used)

DNA typing must be **performed efficiently and reproducibly** (information must hold up in court)

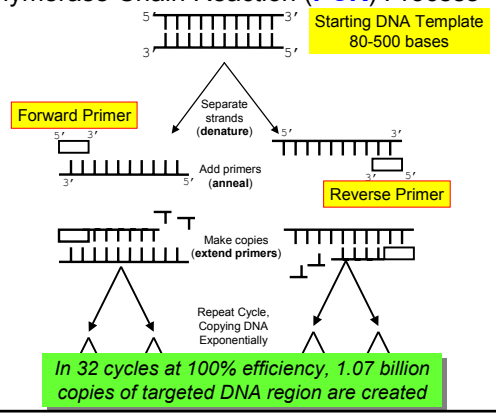
Current standard DNA tests **DO NOT look at genes** – little/no information about race, predisposal to disease, or phenotypical information (eye color, height, hair color) is obtained

DNA in the Cell

The vast majority of DNA is the same from person to person



Polymerase Chain Reaction (PCR) Process



Advantages for STR Markers

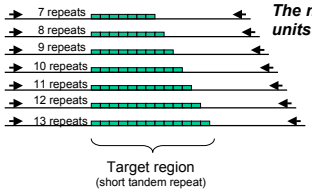
- Small product sizes are generally compatible with degraded DNA and PCR enables recovery of information from small amounts of material
- Multiplex amplification with fluorescence detection enables high power of discrimination in a single test
- Commercially available in an easy to use kit format
- Uniform set of core STR loci provide capability for national and international sharing of criminal DNA profiles

Short Tandem Repeat (STR) Markers

An accordion-like DNA sequence that occurs between genes

TCCCAAGCTCTTCCTCTTCCCTAGATCAATACAGACAGAAGACA
GGTG**GATAGATAGATAGATAGATAGATAGATAGATAGATAGA**
TAGATAGATATCATTGAAAGACAAAACAGAGATGGATGATAGAT
ACATGCTTACAGATGCACAC

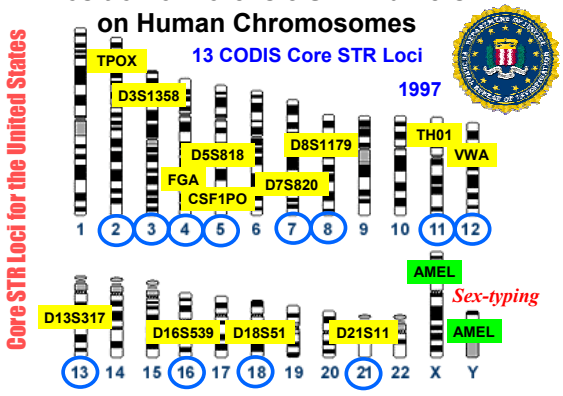
= 12 GATA repeats ("12" is all that is reported)



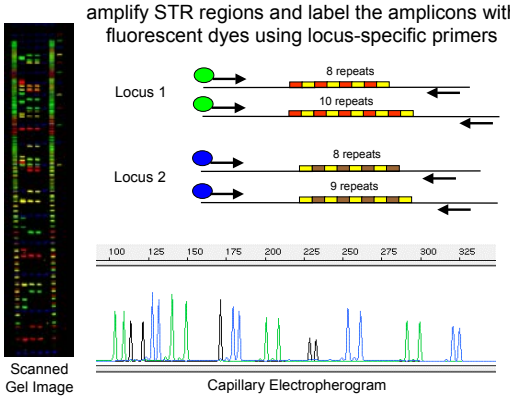
The number of consecutive repeat units can vary between people

The FBI has selected **13 core STR loci** that must be run in all DNA tests in order to provide a common currency with DNA profiles

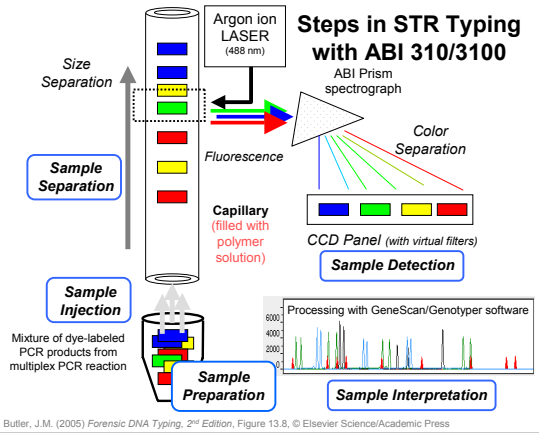
Position of Forensic STR Markers on Human Chromosomes



The polymerase chain reaction (PCR) is used to amplify STR regions and label the amplicons with fluorescent dyes using locus-specific primers



Steps in STR Typing with ABI 310/3100

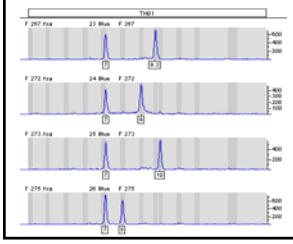


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 13.8, © Elsevier Science/Academic Press

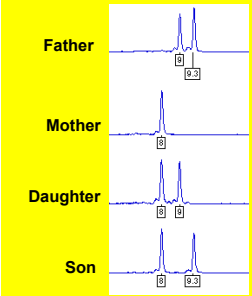
Variation of STRs Among Individuals

An allelic ladder, which is a mixture of common alleles, is used to convert DNA size to STR repeat number

Unrelated Individuals

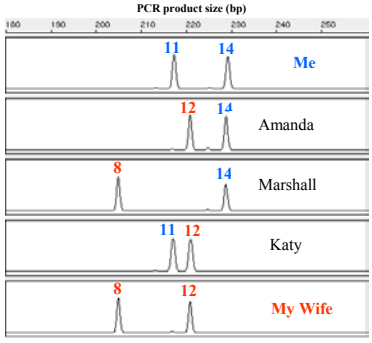


Familial Relationships Can Be Tracked

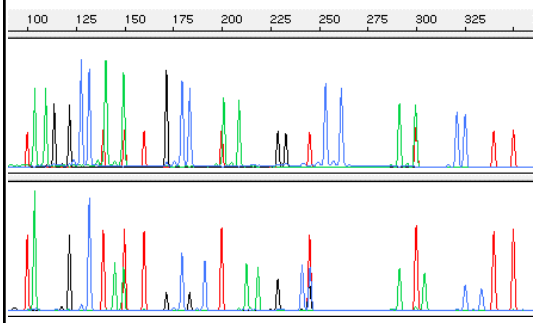


PATERNITY TESTING

Family Inheritance of STR Alleles (D13S317)



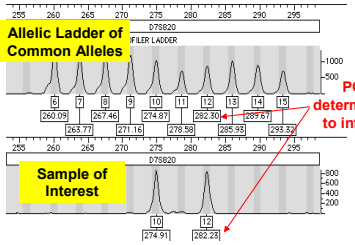
Internal Size Standards (Labeled in a Different Dye Color) Are Used to Calibrate Each Analysis



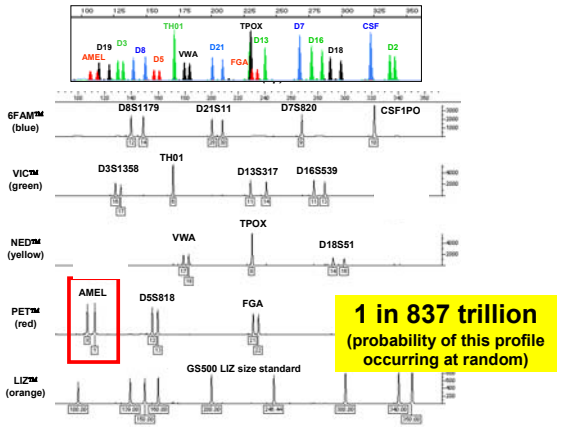
Red peaks are from the internal size standard GS500 ROX

STR Genotyping by Comparison to an Allelic Ladder

- Genotypes are generated by comparison of PCR product sizes of the STR alleles with allelic ladders composed of common variants observed in the population



Shaded bins are +/- 0.5 bp so for a D7S820 amplicon to be designated an allele "12" it must be in the range of 281.80 to 282.80 bp (since the allele 12 in the allelic ladder is 282.30 bp)



Types of STR Repeat Units

Requires size based DNA separation to resolve different alleles from one another

- **Di**nucleotide (CA)(CA)(CA)(CA)
- **Tri**nucleotide (GCC)(GCC)(GCC)
- **Tetra**nucleotide (AATG)(AATG)(AATG)
- **Penta**nucleotide (AGAAA)(AGAAA)
- **Hexa**nucleotide (AGTACA)(AGTACA)

Short tandem repeat (STR) = microsatellite = simple sequence repeat (SSR)

Categories for STR Markers

Category	Example Repeat Structure	13 CODIS Loci
Simple repeats – contain units of identical length and sequence	(GATA)(GATA)(GATA)	TPOX, CSF1PO, D5S818, D13S317, D16S539
Simple repeats with non-consensus alleles (e.g., TH01 9.3)	(GATA)(GAT-)(GATA)	TH01, D18S51, D7S820
Compound repeats – comprise two or more adjacent simple repeats	(GATA)(GATA)(GACA)	VWA, FGA, D3S1358, D8S1179
Complex repeats – contain several repeat blocks of variable unit length	(GATA)(GACA)(CA)(CATA)	D21S11

These categories were first described by **Urquhart *et al.* (1994) *Int. J. Legal Med.* 107:13-20**

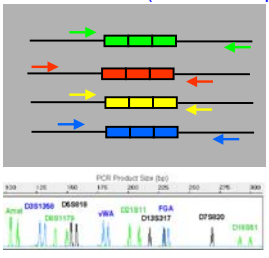
How many STRs in the human genome?

- The efforts of the Human Genome Project have increased knowledge regarding the human genome, and hence there are many more STR loci available now than there were 10 years ago when the 13 CODIS core loci were selected.
- **More than 20,000 tetranucleotide STR loci have been characterized in the human genome** (Collins et al. An exhaustive DNA micro-satellite map of the human genome using high performance computing. Genomics 2003;82:10-19)
- There may be more than a million STR loci present depending on how they are counted (Ellegren H. Microsatellites: simple sequences with complex evolution. Nature Rev Genet 2004;5:435-445).
- STR sequences account for approximately 3% of the total human genome (Lander et al. Initial sequencing and analysis of the human genome. Nature 2001;409:860-921).

Butler, J.M. (2006) Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci.* 51(2):253-265.

Multiplex PCR

(Parallel Sample Processing)

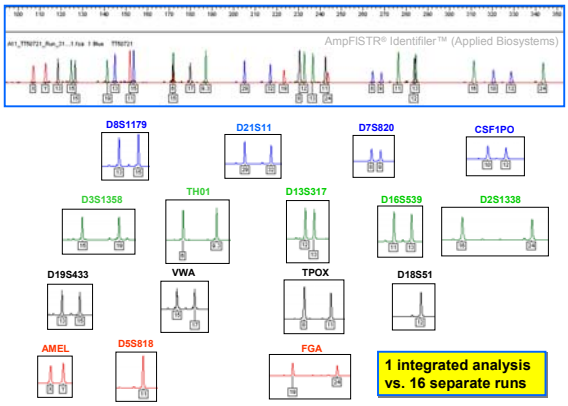


- **Compatible primers are the key to successful multiplex PCR**
- **STR kits are commercially available**
- **15 or more STR loci can be simultaneously amplified**

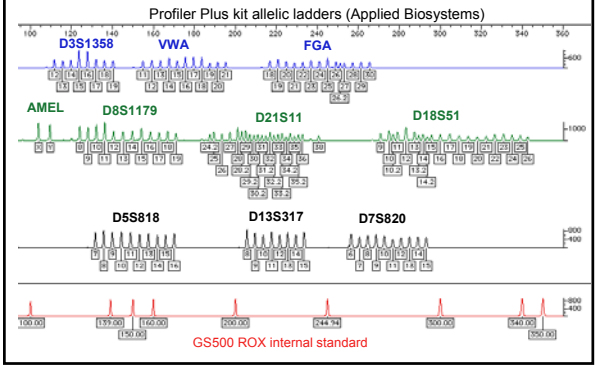
Challenges to Multiplexing
- primer design to find compatible primers (no program exists)
- reaction optimization is highly empirical often taking months

Advantages of Multiplex PCR
- –Increases information obtained per unit time (increases power of discrimination)
- –Reduces labor to obtain results
- –Reduces template required (smaller sample consumed)

Information is tied together with multiplex PCR and data analysis



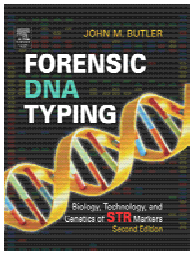
Companies Supply **Allelic Ladders** in STR Kits to **Aid Interlaboratory Consistency**



Biological "Artifacts" of STR Markers

- Stutter Products
- Non-template nucleotide addition
- Microvariants
- Tri-allelic patterns
- Null alleles
- Mutations

Chapter 6 covers these topics in detail



Stutter Products

- Peaks that show up primarily one repeat less than the true allele as a result of strand slippage during DNA synthesis
- Stutter is less pronounced with larger repeat unit sizes (dinucleotides > tri- > tetra- > penta-)
- Longer repeat regions generate more stutter
- Each successive stutter product is less intense (allele > repeat-1 > repeat-2)
- Stutter peaks make mixture analysis more difficult

STR Alleles with Stutter Products

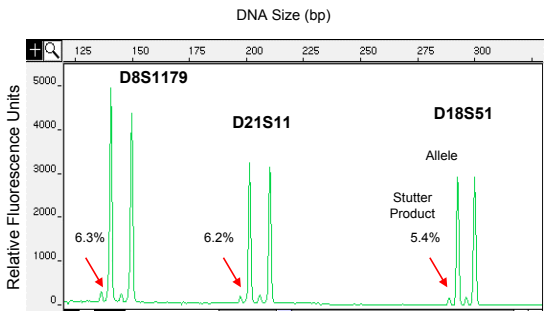
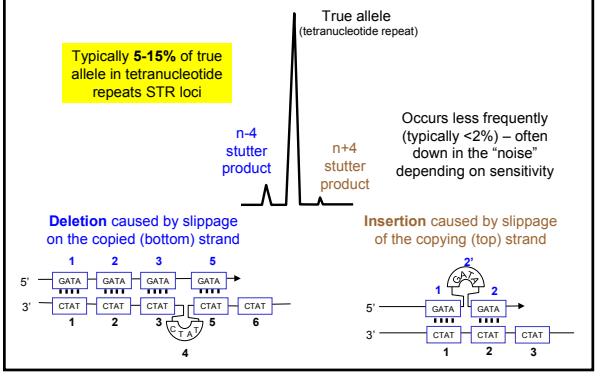


Figure 6.1, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

Stutter Product Formation

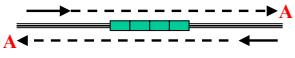
Repeat unit bulges out when strand breathing occurs during replication

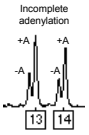


Non-Template Addition

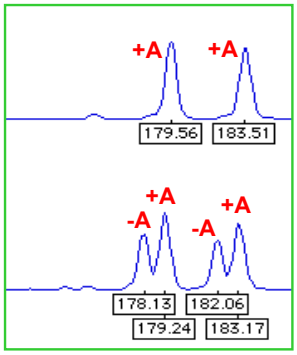
- Taq polymerase will often add an extra nucleotide to the end of a PCR product; most often an "A" (termed "adenylation")
- **Dependent on 5'-end of the reverse primer; a "G" can be put at the end of a primer to promote non-template addition**
- Can be enhanced with extension soak at the end of the PCR cycle (e.g., 15-45 min @ 60 or 72 ºC) – to give polymerase more time
- Excess amounts of DNA template in the PCR reaction can result in incomplete adenylation (not enough polymerase to go around)

Best if there is NOT a mixture of "+/- A" peaks
(desirable to have full adenylation to avoid split peaks)





Impact of the 5' Nucleotide on Non-Template Addition



5'-ACAAG…

Last Base for Primer Opposite Dye Label

(PCR conditions are the same for these two samples)

5'-CCAAG…

Promega includes an ATT sequence on the 5'-end of many of their unlabeled PP16 primers to promote adenylation
see Krenke *et al.* (2002) *J. Forensic Sci.* 47(4): 773-785
http://www.cstl.nist.gov/biotech/strbase/PP16primers.htm

Higher Levels of DNA Lead to Incomplete Adenylation

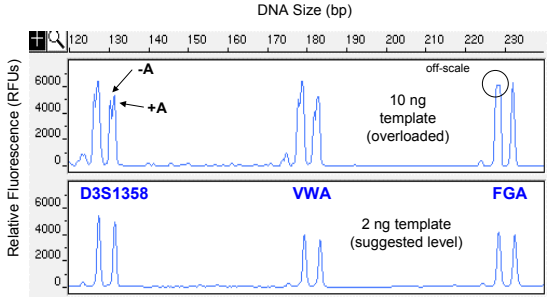


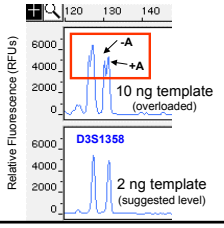
Figure 6.5, J.M. Butler (2005) *Forensic DNA Typing*, 2nd Edition © 2005 Elsevier Academic Press

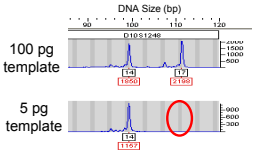
Impact of DNA Amount into PCR

Reason that DNA Quantitation is Important Prior to Multiplex Amplification

Generally 0.5 – 2.0 ng DNA template is best for STR kits

- Too much DNA
  - Off-scale peaks
  - Split peaks (+/-A)
  - Locus-to-locus imbalance
- Too little DNA
  - Heterozygote peak imbalance
  - Allele drop-out
  - Locus-to-locus imbalance

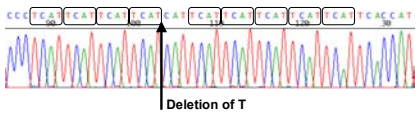




Stochastic effect when amplifying low levels of DNA produces allele dropout

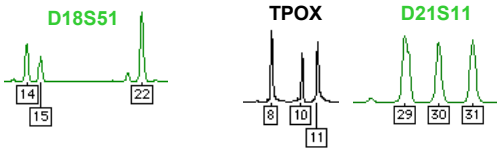
Microvariant "Off-Ladder" Alleles

- Defined as alleles that are not exact multiples of the basic repeat motif or sequence variants of the repeat motif or both
- Alleles with partial repeat units are designated by the number of full repeats and then a decimal point followed by the number of bases in the partial repeat (Bar *et al. Int. J. Legal Med.* 1994, 107:159-160)
- Example: **TH01 9.3 allele**: $[TCAT]_4$ -CAT $[TCAT]_5$



Three-Peak Patterns

Clayton *et al.* (2004) A genetic basis for anomalous band patterns encountered during DNA STR profiling. *J Forensic Sci.* 49(6):1207-1214



"Type 1"

Sum of heights of two of the peaks is equal to the third

Most common in D18S51 and …..

"Type 2"

Balanced peak heights

Most common in TPOX and D21S11

Variant Alleles Cataloged in STRBase

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

Off-Ladder Alleles

328 total variants reported as of 10/04/05

- CSF1PO (15)
- D2S1338 (5)
- D3S1358 (23)
- D5S818 (7)
- D7S820 (23)
- D8S1179 (6)
- D13S317 (13)
- D16S539 (12)
- D18S51 (32)
- D19S433 (12)
- D21S11 (24)
- FESFPS (1)
- FGA (83)
- F13A01 (1)
- HUMTH01 (11)
- Penta D (23)
- Penta E (16)
- TPOX (14)
- VWA (7)

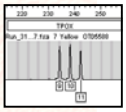
Currently **328** at 13/13 CODIS loci + F13A01, FES/FPS, Penta D, Penta E, D2S1338, D19S433

Tri-Allelic Patterns

80 total patterns reported as of 11/03/05

- CSF1PO (3)
- D3S1358 (4)
- D5S818 (2)
- D7S820 (6)
- D8S1179 (5)
- D13S317 (4)
- D16S539 (4)
- D18S51 (9)
- D21S11 (7)
- FGA (12)
- FES/FEP (1)
- HUMTH01 (1)
- TPOX (13)
- VWA (9)

Currently **80** at 13/13 CODIS loci + FES/FPS



Null Alleles

- Allele is present in the DNA sample but fails to be amplified due to a **nucleotide change in a primer binding site**
- Allele dropout is a problem because a heterozygous sample appears falsely as a homozygote
- Two PCR primer sets can yield different results on samples originating from the same source
- This phenomenon impacts DNA databases
- Large concordance studies are typically performed prior to use of new STR kits

For more information, see J.M. Butler (2005) *Forensic DNA Typing, 2nd Edition*, pp. 133-138

Concordance between STR primer sets is important for DNA databases

PowerPlex 16

17,18

DNA Database

17,17

Allele Dropout

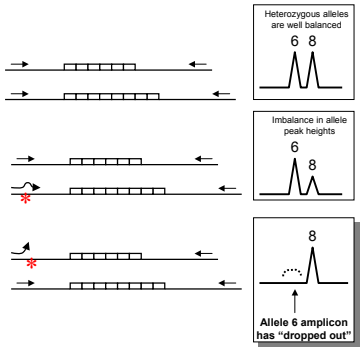
Identifiler

Allele 18 drops out

Search results in a false negative (miss samples that should match)

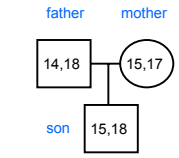
Reduced match stringency is a common solution

Impact of DNA Sequence Variation in the PCR Primer Binding Site

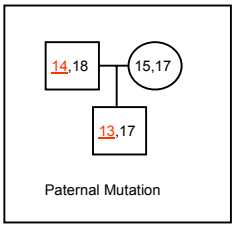


Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 6.9, ©Elsevier Academic Press

Mutation Observed in Family Trio



Normal Transmission of Alleles (No Mutation)



Paternal Mutation

Butler, J.M. (2001) *Forensic DNA Typing*, Figure 6.9, ©Academic Press

STR Measured Mutation Rates

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

	STR Locus	Maternal Meioses (%)	Paternal Meioses (%)	Either Parent	Total Mutations	Rate
13 CODIS core loci	CSF1PO	70/179,353 (0.04)	727/504,342 (0.14)	303	1,100/683,695	0.16%
	FGA	134/238,378 (0.06)	1,481/473,924 (0.31)	495	2,110/712,302	0.30%
	TH01	23/189,478 (0.01)	29/346,518 (0.008)	23	75/535,996	0.01%
	TPOX	16/299,186 (0.005)	43/328,067 (0.01)	24	83/627,253	0.01%
	VWA	133/400,560 (0.03)	907/646,851 (0.14)	628	1,668/1,047,411	0.16%
	D3S1358	37/244,484 (0.02)	429/336,208 (0.13)	266	732/580,692	0.13%
	D5S818	84/316,102 (0.03)	537/468,366 (0.11)	303	924/784,468	0.12%
	D7S820	43/334,886 (0.01)	550/461,457 (0.12)	218	811/796,343	0.10%
	D8S1179	54/237,235 (0.02)	396/264,350 (0.15)	225	675/501,585	0.13%
	D13S317	142/348,395 (0.04)	608/435,530 (0.14)	402	1,152/783,925	0.15%
	D16S539	77/300,742 (0.03)	350/317,146 (0.11)	256	683/617,888	0.11%
	D18S51	83/130,206 (0.06)	623/278,098 (0.22)	330	1,036/408,304	0.25%
	D21S11	284/258,795 (0.11)	454/306,198 (0.15)	423	1,161/564,993	0.21%
	Penta D	12/18,701 (0.06)	10/15,088 (0.07)	21	43/33,789	0.13%
	Penta E	22/39,121 (0.06)	58/44,152 (0.13)	55	135/83,273	0.16%
	D2S1338	2/25,271 (0.008)	61/81,960 (0.07)	31	94/107,231	0.09%
	D19S433	22/28,027 (0.08)	16/38,983 (0.04)	37	75/67,010	0.11%
	F13A01	1/10,474 (0.01)	37/65,347 (0.06)	3	41/75,821	0.05%
	FES/FPS	3/18,918 (0.02)	79/149,028 (0.05)	None reported	82/167,946	0.05%
	F13B	2/13,157 (0.02)	8/27,183 (0.03)	1	11/40,340	0.03%
	LPL	0/8,821 (<0.01)	9/16,943 (0.05)	4	13/25,764	0.05%
	SE33 (ACTBP2)	0/330 (<0.30)	330/51,610 (0.64)	None reported	330/51,940	0.64%

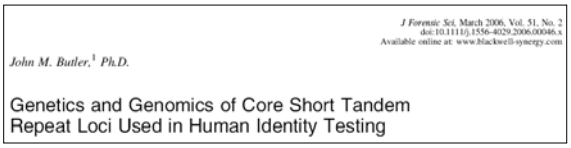
*Data used with permission from American Association of Blood Banks (AABB) 2002 Annual Report.

Summary of STR Mutations

Mutations impact paternity testing and missing persons investigations but not forensic direct evidence-suspect matches…

- Mutations happen and need to be considered
- Usually 1 in ~1000 meioses
- Paternal normally higher than maternal
- VWA, FGA, and D18S51 have highest levels
- TH01, TPOX, and D16S539 have lowest levels

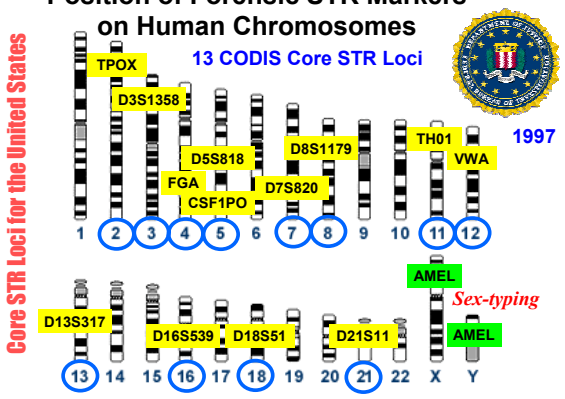
Review Article on Core STR Loci



Journal of Forensic Sciences 2006, 51(2): 253-265

- Reviews STR kits, genomic locations, mutation rates, potential genetic linkage, and known variant alleles for autosomal STR and Y-STR loci
- **Covers characteristics of 18 autosomal loci (13 core CODIS loci, D2, D19, Penta D, Penta E, SE33) and 11 SWGDAM-recommended Y-STR loci**

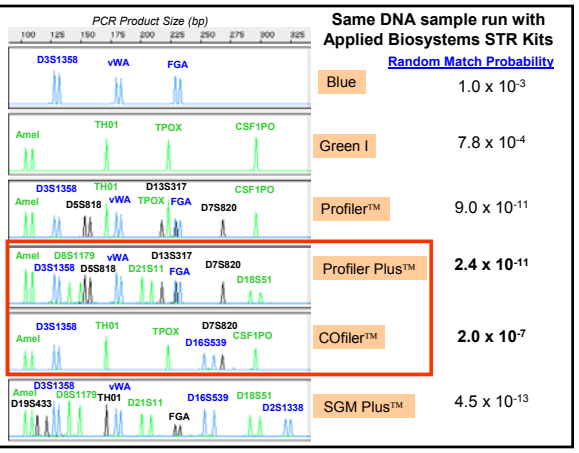
Position of Forensic STR Markers on Human Chromosomes



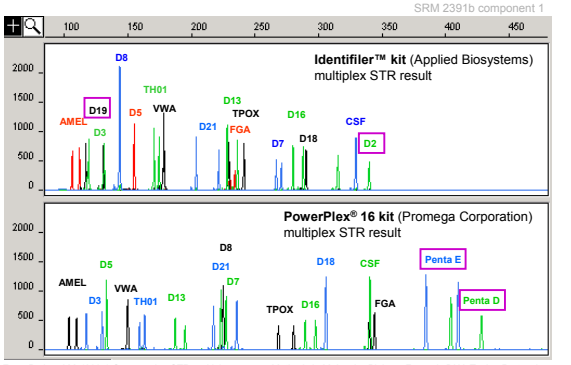
Characteristics of Core STR Loci

Locus	Chromosomal Location	Physical Position (May 2004; NCBI build 35)	Repeat Motif	Observed Alleles
TPOX	2p25.3 thyroid peroxidase, 10th intron	Chr 2; 1.472 Mb	GAAT	4-16
D3S1358	3p21.31	Chr 3; 45.557 Mb	[TCTG][TCTA]	8-21
FGA	4q31.3 alpha fibrinogen, 3rd intron	Chr 4; 155.866 Mb	CTTT	12.2-51.2
D5S818	5q23.2	Chr 5; 123.139 Mb	AGAT	7-18
CSF1PO	5q33.1 c-fms proto-oncogene, 6th intron	Chr 5; 149.436 Mb	TAGA	5-16
D7S820	7q21.11	Chr 7; 83.433 Mb	GATA	5-16
D8S1179	8q24.13	Chr 8; 125.976 Mb	[TCTA][TCTG]	7-20
TH01	11p15.5 tyrosine hydroxylase, 1st intron	Chr 11; 2.149 Mb	TCAT	3-14
VWA	12p13.31 von Willebrand Factor, 40th intron	Chr 12; 5.963 Mb	[TCTG][TCTA]	10-25
D13S317	13q31.1	Chr 13; 81.620 Mb	TATC	5-16
D16S539	16q24.1	Chr. 16; 84.944 Mb	GATA	5-16
D18S51	18q21.33	Chr 18; 59.100 Mb	AGAA	7-40
D21S11	21q21.1	Chr 21; 19.476 Mb	Complex [TCTA][TCTG]	12-41.2

Butler, J.M. (2006) Genetics and genomics of core STR loci used in human identity testing. *J. Forensic Sci.* 51(2): 253-265



Commercial STR 16plex Kits



From Butler, J.M. (2005) Constructing STR multiplex assays. *Methods in Molecular Biology: Forensic DNA Typing Protocols* (Carracedo, A., ed.), Humana Press: Totowa, New Jersey, 297: 53-66.

Value of STR Kits

Advantages

- Quality control of materials is in the hands of the manufacturer (saves time for the end-user)
- Improves consistency in results across laboratories – same allelic ladders used
- Common loci and PCR conditions used – aids DNA databasing efforts
- Simpler for the user to obtain results

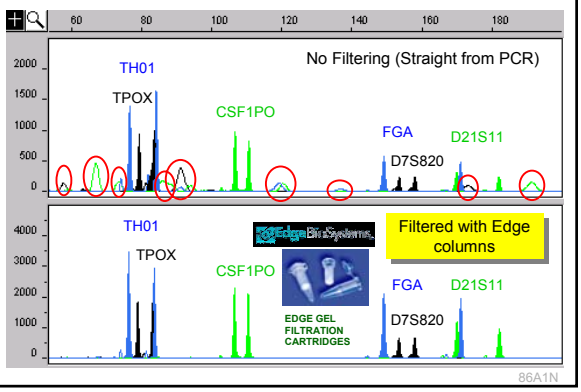
Disadvantages

- Contents may not be completely known to the user (e.g., primer sequences)
- Higher cost to obtain results

Primer Synthesis and Dye Blobs

- Oligonucleotide primers are synthesized from a 3'-to-5' direction on solid-phase supports using phosphoramidite chemistry
- The fluorescent dye is attached at 5'end of the primer (it is the last component added)
- The coupling reaction at each step of primer synthesis is not 100%, which can lead to some minor level impurities
- Left-over dye molecules that are not removed by post-synthesis purification can be carried through the PCR amplification step and injected onto the capillary to produce "dye blobs" or "dye artifacts" in CE electropherograms (wider than true allele peaks)

Problems with Dye Artifacts from Fluorescent Primers



Impact of Degraded DNA Samples

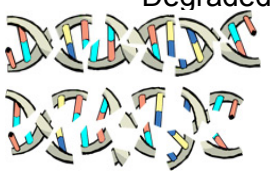
- Comparison to a phone number (string of 13 numbers)

001-301-975-4049

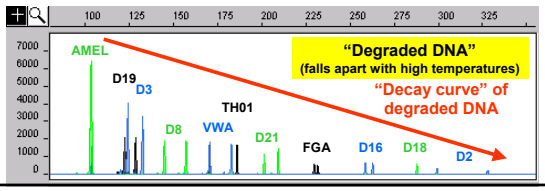
- If you only had "4049"…this information would be of limited value since it is not as specific (and could match other phone numbers from different area codes)
- DNA profiles are essentially a string of numbers – **if the DNA is damaged, then the string of numbers is shorter and less informative…**

----------4049 or ----301-9-------

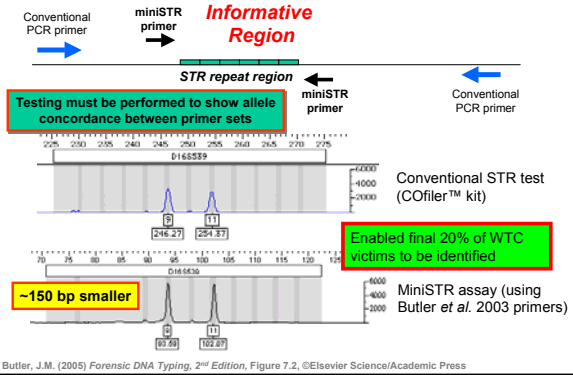
Degraded DNA



Larger segments of DNA cannot be recovered when DNA molecules have fragmented into small pieces (caused by heat, water, or bacteria)



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

Timeline for miniSTRs

and Demonstrating the Value of Using Reduced Size Amplicons for Degraded DNA

- 1994 – FSS finds that smaller STR loci work best with burned bone and tissue from Branch Davidian fire
- 1997 – New primers developed for time-of-flight mass spectrometry to make small STR amplicons
- 2001 – Work at NIST and OhioU with CODIS STRs; BodePlexes used in WTC investigation starting 2002
- 2004 – Work at NIST with **non-CODIS (NC) miniSTRs**
- 2006 – Applied Biosystems to release a 9plex miniSTR kit

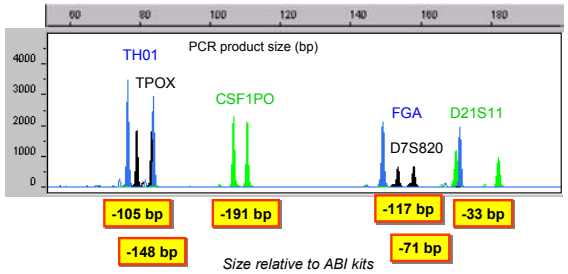
http://www.cstl.nist.gov/biotech/strbase/miniSTR/timeline.htm

J. Forensic Sci. Sept 2003 issue

J Forensic Sci, September 2003, Vol. 48, No. 5
Paper ID JFS2003043_485
Available online at: www.astm.org

John M. Butler,[1] Ph.D.; Yin Shen,[2,3] Ph.D.; and Bruce R. McCord Ph.D.[2]

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



Size relative to ABI kits

The International Commission on Missing Persons (ICMP) is Now Using miniSTRs

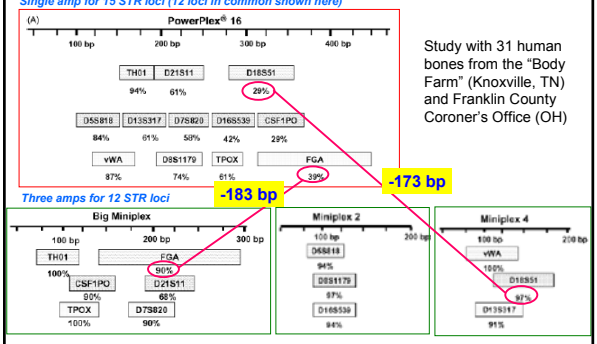
100s of bones are tested each week with miniSTRs to help in the re-association of remains

Miniplex 02
D21S11, D13S317, D7S820, CSF1PO, vWA and D8S1179

(Tom Parsons, personal communication)

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

Single amp for 15 STR loci (12 loci in common shown here)



Study with 31 human bones from the "Body Farm" (Knoxville, TN) and Franklin County Coroner's Office (OH)

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

EDNAP Exercise on Degraded DNA

ARTICLE IN PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

ELSEVIER

Forensic Science International

Forensic Science International xxx (2005) xxx–xxx

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. Briõn [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

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

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)

Identifying Victims of Mass Disasters

Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Chapter 24

POLICY FORUM
EPIDEMIOLOGY

DNA Identifications After the 9/11 World Trade Center Attack

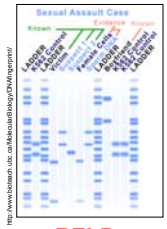
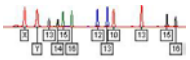
Leslie G. Biesecker,* Joan E. Bailey-Wilson, Jack Ballantyne, Howard Baum, Frederick R. Bieber, Charles Brenner, Bruce Budowle, John M. Butler, George Carmody, P. Michael Conneally, Barry Duceman, Arthur Eisenberg, Lisa Forman, Kenneth K. Kidd, Benoit Leclair, Steven Niezgoda, Thomas J. Parsons, Elizabeth Pugh, Robert Shaler, Stephen T. Sherry, Amanda Sozer, Anne Walsh

***Science* (2005) 310: 1122-1123**

Highly Degraded DNA Was Obtained from the Human Remains Recovered

Largest Forensic Case in History
~20,000 bone fragments were processed
>6,000 family reference samples and personal effects samples were analyzed

The DNA Field Moves Forward…

The Past	The Present	The Future
		
RFLP	STRs	miniSTRs
500 – 25,000 bp	100 - 500 bp	50 - 150 bp

STRBase

Short Tandem Repeat DNA Internet Database

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

General Information	Forensic Interest Data	Supplemental Info
•Intro to STRs (downloadable PowerPoint)	•FBI CODIS Core Loci	•Reference List >2500
•STR Fact Sheets	•DAB Standards	•Technology Review
•Sequence Information	•NIST SRMs 2391	•Addresses for Scientists
•Multiplex STR Kits	•Published PCR Primers	•Links to Other Web Sites
•Variant Allele Reports	•Y-Chromosome STRs	•DNA Quantitation
•Training Slides	•Population Data	•mtDNA
	•Validation Studies	•New STRs
	•miniSTRs	

New information is added regularly…

BREAK

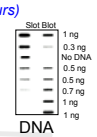
Steps in DNA Analysis

Steps Involved
- Collection
- Specimen Storage
- Extraction
- Quantitation
- Multiplex PCR
- STR Typing
- Interpretation of Results
- Database Storage & Searching
- Calculation of Match Probability

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

Blood Stain Buccal swab
Sample Collection & Storage

DNA Extraction



DNA Quantitation

Biology

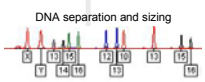
Multiplex PCR Amplification

Genetics

If a match occurs, comparison of DNA profile to population allele frequencies to generate a case report with probability of a random match to an unrelated individual

DNA Database Search

Technology



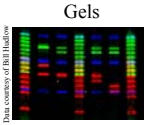
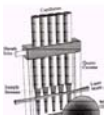
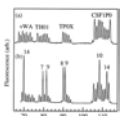
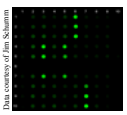
STR Typing

Male: 13,14-15,16-12,13-10,13-15,16

Interpretation of Results

STR Typing Technologies

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

Gels	Capillary Electrophoresis	Capillary Arrays
		
Data courtesy of Bill Hudlow		
J. Forensic Sci. (1998) 43: 1168-1180	*Electrophoresis.* (1998) 19: 86-93	*Nucleic Acids Res.* (1999) 27: e36
Microchip CE	**Mass Spectrometry**	**Hybridization Arrays**
		
	www.sequenom.com	Data courtesy of Jim Schumm
PNAS (1997) 94: 10273-10278	*Int. J. Legal Med.* (1998) 112: 45-49	*Nucleic Acids Res.* (2000) 28: e17

Typical Instruments Used for STR Typing

GeneAmp 9700

Thermal Cycler for PCR Amplification

Capillary electrophoresis instruments for separating and sizing PCR products

single capillary

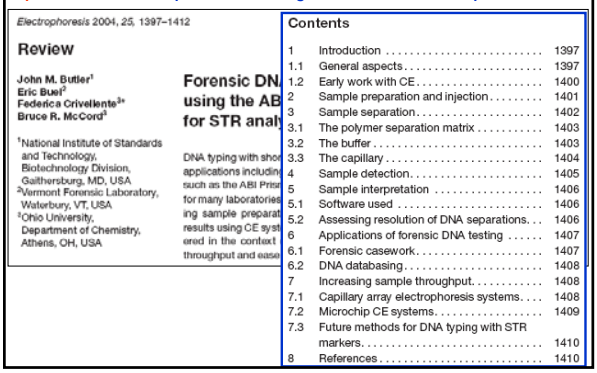
ABI 310

16-capillary array

ABI 3100

Review Article on STRs and CE

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



Electrophoresis 2004, *25*, 1397–1412

Review

John M. Butler[1]
Eric Buel[2]
Federica Crivellente[3]*
Bruce R. McCord[3]

[1]National Institute of Standards and Technology, Biotechnology Division, Gaithersburg, MD, USA
[2]Vermont Forensic Laboratory, Waterbury, VT, USA
[3]Ohio University, Department of Chemistry, Athens, OH, USA

Forensic DN
using the AB
for STR anal

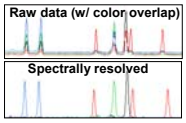
DNA typing with sho
applications includin
such as the ABI Prisr
for many laboratories
ing sample preparat
results using CE syst
ered in the context
throughput and ease

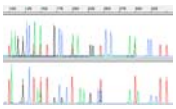
Contents

Analytical Requirements for STR Typing

Butler *et al.* (2004) Electrophoresis 25: 1397-1412

- Fluorescent dyes must be **spectrally resolved** in order to distinguish different dye labels on PCR products
- PCR products must be **spatially resolved** – desirable to have single base resolution out to >350 bp in order to distinguish variant alleles
- High **run-to-run precision** – an internal sizing standard is used to calibrate each run in order to compare data over time

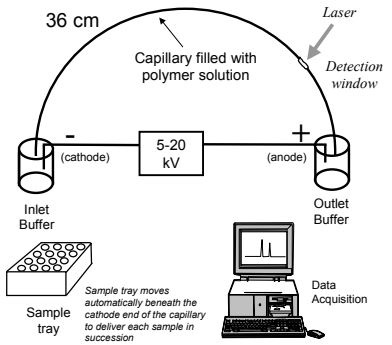




Why Use CE for DNA Analysis?

1. **Injection, separation, and detection are automated.**
2. **Rapid separations are possible**
3. **Excellent sensitivity and resolution**
4. **The time at which any band elutes is precisely determined**
5. **Peak information is automatically stored for easy retrieval**

Components of a Capillary Electrophoresis System



Butler, J.M. (2001) *Forensic DNA Typing*, Figure 9.3, ©Academic Press

In the early 1990s the real question was how to transition from a gel to a capillary

- Cross-linked acrylamide gel filled capillaries were tried first
  - Reusable?
  - Bubble formation
  - Themal degradation
- Alternative was to not use a gel at all
  - Refillable sieving polymers
  - However, resolution was poor early on

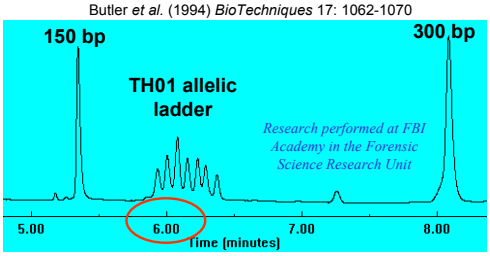
Early Work with CE and STRs

- Barry Karger's group (1988-1990)
  - Utilized gel-filled capillaries to separate ssDNA
  - Introduced sieving polymers in the form of linear polyacrylamide to separate restriction digests
- Beckman P/ACE 2050 is introduced in 1992 as the first commercially available CE coupled to a laser to enable fluorescence detection
- John Butler and Bruce McCord (1993-1995)
  - **First STR typing with single color CE using intercalating dyes and dual bracketing internal size standards**
- Rich Mathies' group (1995)
  - First STR typing with multi-color CE (and multi-capillary) using dye-labeled primers
- **ABI 310 is introduced in July 1995 as the first commercially available multi-color CE**

First Rapid STR Typing with Capillary Electrophoresis

Single color detection with dual internal size standards

Butler *et al.* (1994) *BioTechniques* 17: 1062-1070



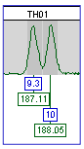
Performed in December 1993

Technology Implementation Takes Time – the FBI did not start running casework samples using STRs and CE until January 1999

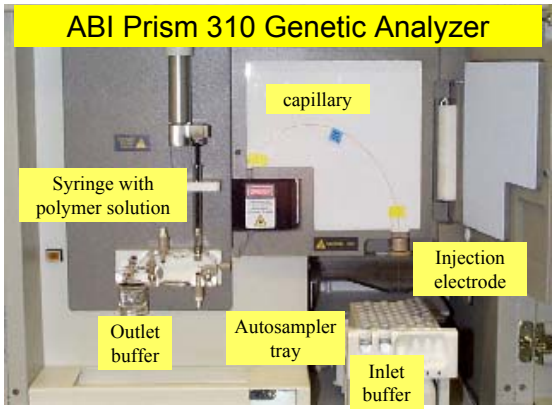
Requirements for Reliable STR Typing

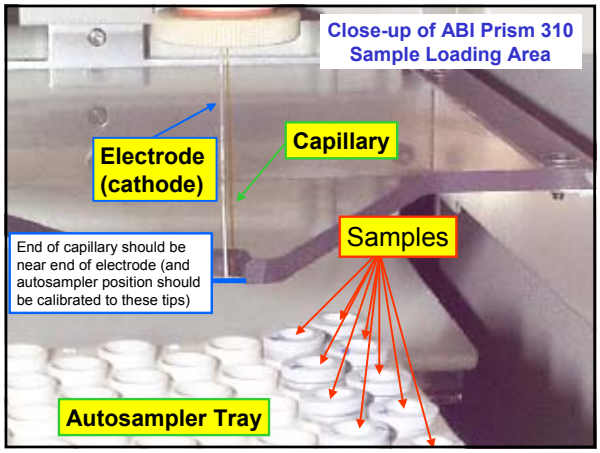
Butler *et al.* (2004) Electrophoresis 25: 1397-1412

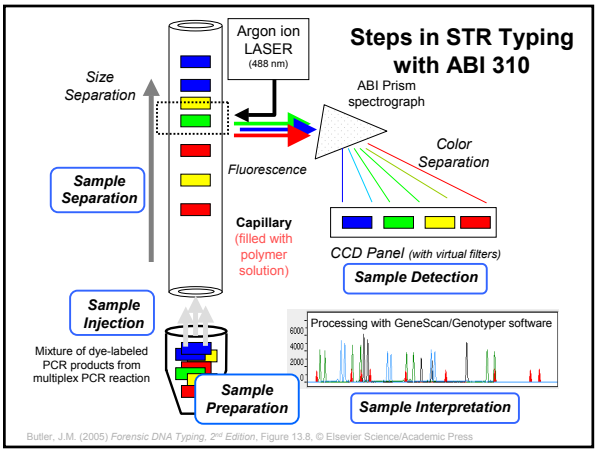
- Reliable sizing over a 75-500 bp size region
- High run-to-run precision between processed samples to permit comparison of allelic ladders to sequentially processed STR samples
- Effective color separations of different dye sets used to avoid bleed through between 4 or 5 different colors
- Resolution of at least 1 bp to >350 bp to permit reliable detection of microvariant alleles



ABI Prism 310 Genetic Analyzer







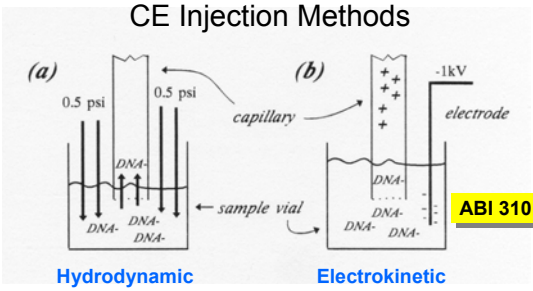
Butler, J.M. (2005) *Forensic DNA Typing, 2nd Edition*, Figure 13.8, © Elsevier Science/Academic Press

Process Involved in 310/3100 Analysis

- Injection
  - electrokinetic injection process (formamide, water)
  - importance of sample stacking
- Separation
  - Capillary – 50um fused silica, 43 cm
  - POP-4 polymer – Polydimethyl acrylamide
  - Buffer - TAPS pH 8.0
  - Denaturants – urea, pyrolidinone
- Detection
  - fluorescent dyes with excitation and emission traits
  - CCD with defined virtual filters produced by assigning certain pixels

Injection

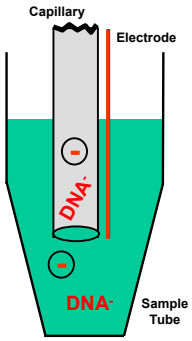
CE Injection Methods



Ulfelder K. J.; McCord, B. R. (1996) Capillary Electrophoresis of DNA, In *Handbook of Capillary Electrophoresis* (Landers, J., ed.), CRC Press: NY, pp. 347-378.

Butler, J.M. (1997) Effects of sample matrix and injection on DNA separations. *Analysis of Nucleic Acids by Capillary Electrophoresis* (Heller, C., ed.), Vieweg: Germany, Chapter 5, pp. 125-134

Electrokinetic Injection Process



Amount of DNA injected is inversely proportional to the ionic strength of the solution

Salty samples result in poor injections

Sample Conductivity Impacts Amount Injected

$$[DNA_{inj}] = \frac{Et(\pi r^2)\,(\mu_{ep} + \mu_{eof})[DNA_{sample}]\,(\lambda_{buffer})}{\lambda_{sample}}$$

$[DNA_{inj}]$ is the amount of sample injected

E is the electric field applied

t is the injection time

r is the radius of the capillary

μ_{ep} is the mobility of the sample molecules

μ_{eof} is the electroosmotic mobility

$[DNA_{sample}]$ is the concentration of DNA in the sample

λ_{buffer} is the buffer conductivity

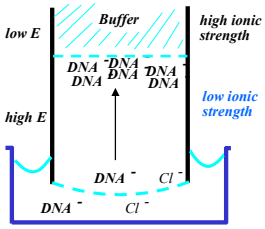
λ_{sample} is the sample conductivity

Cl- ions and other buffer ions present in PCR reaction contribute to the sample conductivity and thus will compete with DNA for injection onto the capillary

Butler *et al.* (2004) Electrophoresis 25: 1397-1412

Two Major Effects of Sample Stacking

1. Sample is preconcentrated. Effect is inversely proportional to ionic strength
2. Sample is focused. Ions stop moving in low electric field
3. Mobility of sample = μ_{ep} = velocity/ electric field



Typical Sample Preparation for ssDNA

1. Perform PCR with dye-labeled primers
2. Dilute 1 µL PCR product with 24 µL deionized formamide; add 1 µL ROX-labeled internal sizing standard
3. Denature 2 minutes at 95 °C with thermocycler
4. Cool to 4 °C in thermocycler or ice bath
5. Sample will remain denatured for at least 3 days

Comments on Sample Preparation

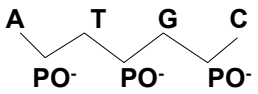
- Use high quality formamide (<100 µS/cm)!
  - ABI sells Hi-Di formamide
  - regular formamide can be made more pure with ion exchange resin
- Deionized water vs. formamide
  - Biega and Duceman (1999) *J. Forensic Sci.* 44: 1029-1031
  - *Crivellente, Journal of Capillary Electrophoresis* **2002,** *7 (3-4),* 73-80.
  - water works fine but samples are not stable as long as with formamide; water also evaporates over time…
- Denaturation with heating and snap cooling
  - use a thermal cycler for heating and cold aluminum block for snap cooling
  - heat/cool denaturation step is necessary only if water is substituted for formamide...

Separation

DNA and Electrophoresis

"From a practical point of view it is disappointing that electrophoresis cannot be used to fractionate or analyze DNA's on the basis of size" Olivera, Biopolymers 1964, 2, 245

$\mu_{ep} = q/6\pi\eta r$ **small ions with high charge move fastest**

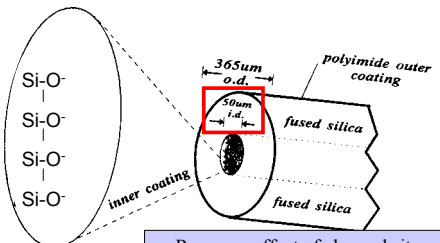


As size increases so does charge!

Separation Issues

- Capillary wall coating -- dynamic coating with polymer
  - Wall charges are masked by methyl acrylamide
- Electrophoresis buffer –
  - Urea for denaturing and viscosity
  - Buffer for consistent pH
  - Pyrolidinone for denaturing DNA
  - EDTA for stability and chelating metals
- Polymer solution -- POP-4 (but others work also)
- Run temperature -- 60 °C helps reduce secondary structure on DNA and improves precision. (Temperature control affects DNA sizing)

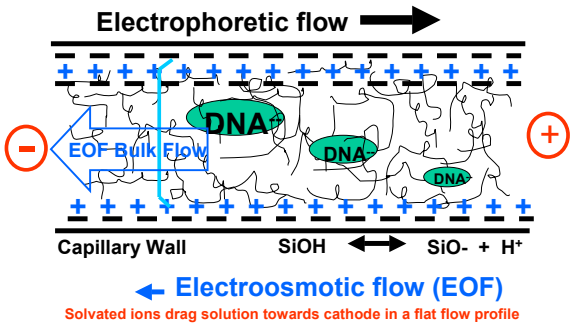
Capillary Coating



Removes effect of charged sites-eliminates EOF, sample adsorption

Dynamic coating of charged sites on fused silica capillary is accomplished with POP-4 polymer

Capillary Wall Coatings Impact DNA Separations



Solvated ions drag solution towards cathode in a flat flow profile

What is in POP-4 and Genetic Analyzer Buffer?

© 1997 Oxford University Press — Nucleic Acids Research, 1997, Vol. 25, No. 19 3925–3929

Improved single-strand DNA sizing accuracy in capillary electrophoresis

Barnett B. Rosenblum*, Frank Oaks, Steve Menchen and Ben Johnson

PE Applied Biosystems, 850 Lincoln Centre Drive, Foster City, CA 94404, USA

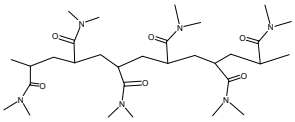
Received May 29, 1997; Revised and Accepted August 6, 1997

See also Wenz *et al.* (1998) *Genome Research* 8: 69-80

POP-4 (4% poly-dimethylacrylamide, 8 M urea, 5% 2-pyrrolidinone)

US Patent 5,552,028 covers POP-4 synthesis

Running buffer contains 100 mM TAPS and 1 mM EDTA (adjusted to pH 8.0 with NaOH) *TAPS = N-Tris-(hydroxymethyl)methyl-3-aminopropane-sulfonic acid*



How to Improve Resolution?

1. **Lower Field Strength**
2. **Increase Capillary Length**
3. **Increase Polymer Concentration**
4. **Increase Polymer Length**

All of these come at a cost of longer separation run times

Detection

Detection Issues

- Fluorescent dyes
  - spectral emission overlap
  - relative levels on primers used to label PCR products
  - dye “blobs” (free dye)
- Virtual filters
  - hardware (CCD camera)
  - software (color matrix)

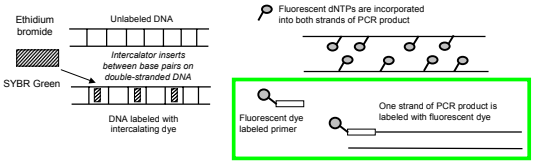
Filters determine which wavelengths of light are recorded from the CCD camera

Laser Used in ABI 310

- Argon Ion Laser
- 488 nm and 514.5 nm for excitation of dyes
- 10 mW power
- Lifetime ~5,000 hours (1 year of full-time use)
- Cost to replace ~$5,500
- Leads to highest degree of variability between instruments and is most replaced part
- Color separation matrix is specific to laser used on the instrument

Methods for Fluorescently Labeling DNA

- Intercalating Dyes (post-PCR)
- Dye-labeled nucleotide insertion during PCR
- Dye-labeled primer insertion during PCR

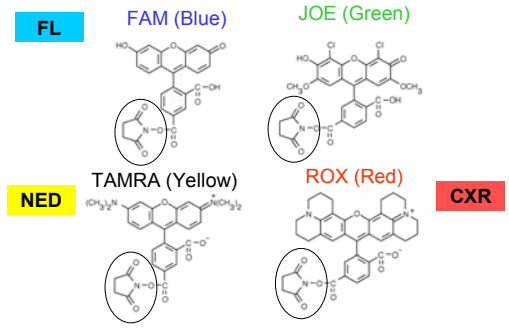


Butler, J.M. (2001) *Forensic DNA Typing*, Figure 10.2, ©Academic Press

Fluorescent Labeling of PCR Products

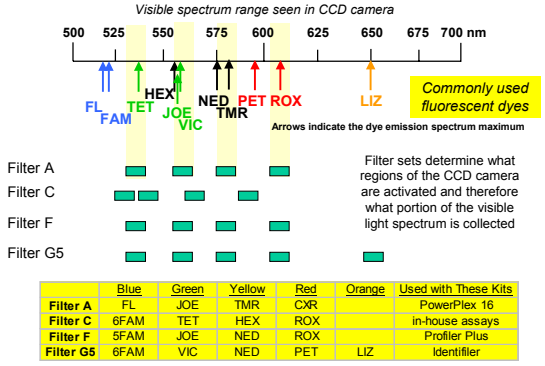
- Dyes are attached to one primer in a pair used to amplify a STR marker
- Dyes are coupled to oligonucleotides (primers) through NHS-esters and amine linkages on the 5'end of the primer: Dye-$(CH_2)_6$-primer
- Dye-labeled oligonucleotides are incorporated during multiplex PCR amplification giving a specific color "tag" to each PCR product
- PCR products are distinguished using CCD imaging on the 310

Fluorescent Dyes Used in 4-Color Detection



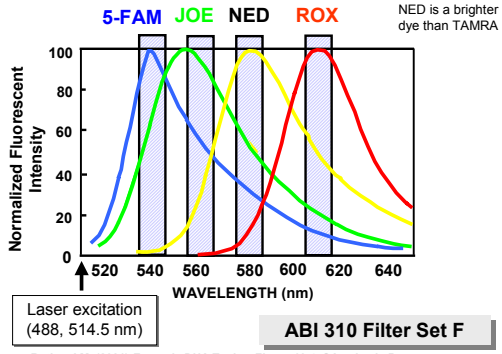
Butler, J.M. (2001) *Forensic DNA Typing*, Figure 10.3, ©Academic Press

Virtual Filters Used in ABI 310



	Blue	Green	Yellow	Red	Orange	Used with These Kits
Filter A	FL	JOE	TMR	CXR		PowerPlex 16
Filter C	6FAM	TET	HEX	ROX		in-house assays
Filter F	5FAM	JOE	NED	ROX		Profiler Plus
Filter G5	6FAM	VIC	NED	PET	LIZ	Identifiler

Fluorescent Emission Spectra for ABI Dyes

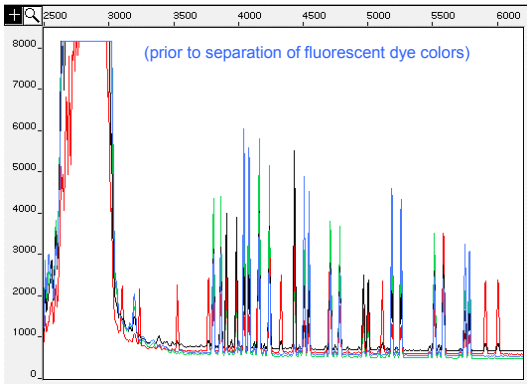


Butler, J.M. (2001) *Forensic DNA Typing*, Figure 10.4, ©Academic Press

Please Note!

- There are no filters in a 310
- Its just the choice of pixels in the CCD detector
- All the light from the grating is collected
- You just turn some pixels on and some off

Raw Data from the ABI Prism 310



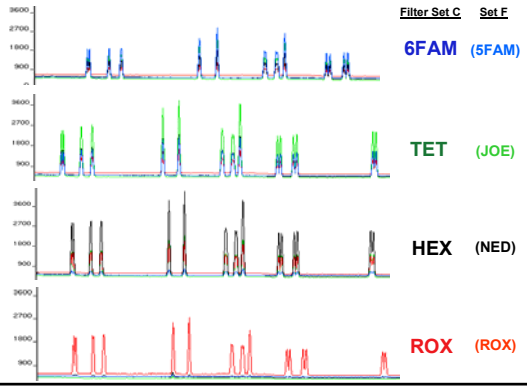
Why Make a Matrix?

The matrix is the solution to a problem:
What's the contribution at any given wavelength (filter set) from each dye ?

There are 4 dyes

- **Remember algebra from high school?**
- **To solve a problem with 4 unknowns, you need 4 equations**

Matrix Standards (Raw Data)



For Example

$I_{540} = bx_b + gy_b + yz_b + rw_b$ intensity of blue
$I_{560} = bx_g + gy_g + yz_g + rw_g$ intensity of green
$I_{580} = bx_y + gy_y + yz_y + rw_y$ intensity of yellow
$I_{610} = bx_r + gy_r + yz_r + yw_r$ intensity of red

Where
b is the %blue labeled DNA
g is the %green labeled DNA, etc.

x,y,z,w are the numbers in the matrix (sensitivity to each color)

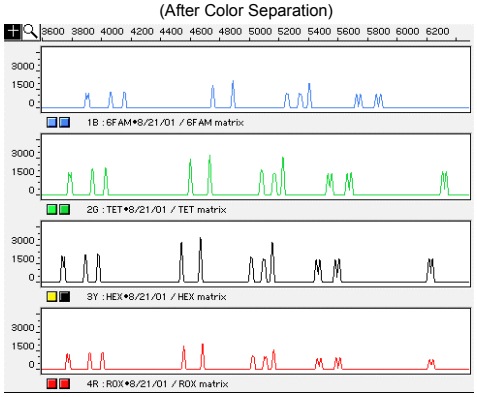
P0P4STRM0DF

Reactions

	B	G	Y	R
B	1.0000	0.8502	0.1380	0.0009
G	0.8300	1.0000	0.7622	0.0051
Y	0.6416	0.8324	1.0000	0.1102
R	0.4493	0.6484	0.7851	1.0000

If you solve xyzw for each dye individually
Then you can determine dye contribution for any mixture

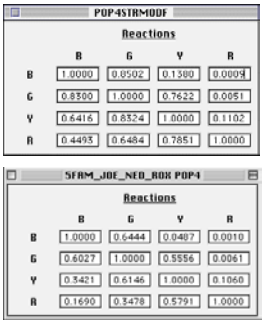
Matrix Standards
(After Color Separation)



Comments on Matrices
(Multi-Component Analysis)

- Make sure that the right filter set and matrix are applied when collecting data
- You can always apply another matrix to a sample collected on the ABI 310 but it must be run with the right filter set (matrix must be run first with ABI 3100)
- It is important to update matrices on a regular basis (depending on use) due to differences in laser power over time
- A good indication of when to run a new matrix is the observation of pull-up between dye colors when peaks are smaller than ~4,000 RFUs

Same Dye Set and Filter F with Different ABI 310s



Instrument lasers make a big difference

Injection List in Data Collection Software

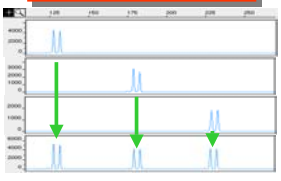
- Lists samples to be analyzed (repeats can be easily performed)
- Sets virtual filter on CCD camera
- Sets electrophoresis time and voltage
- Sets injection time and voltage
- Sets run temperature
- If desired, sample analysis can be set up for automatic matrix color separation and sizing with internal standards using defined analysis parameters

Ways to Increase Sample Throughput

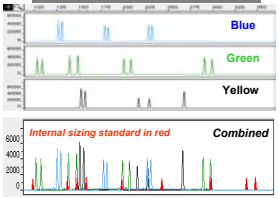
- Run more gels (FMBIO approach)
- Increase speed of single sample analysis (microchip CE systems)
- Multiplex fluorescent dyes of different colors (higher level PCR multiplexes)
- **Parallel separations using capillary arrays** (e.g., ABI 3100 or 3130)
- New detection technologies (MALDI-TOF mass spectrometry)

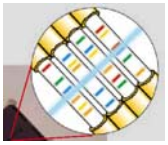
Methods for Parallel Sample Processing

Multiplex by Size



Multiplex by Dye Color

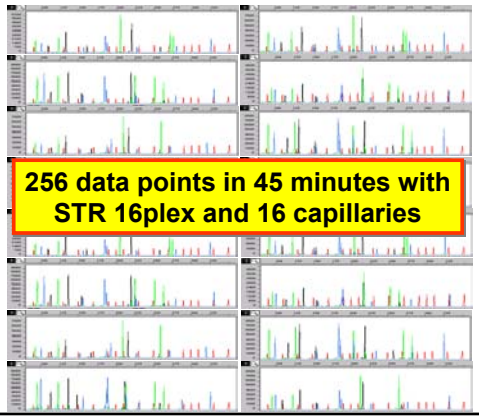




Multiplex by Number of Capillaries

ABI 3100: 16 capillaries
ABI 3730: 96 capillaries
ABI 3100 *Avant*: 4 capillaries

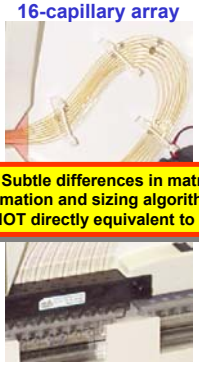
High-Throughput STR Typing on the ABI 3100 (16-capillary array)



256 data points in 45 minutes with STR 16plex and 16 capillaries

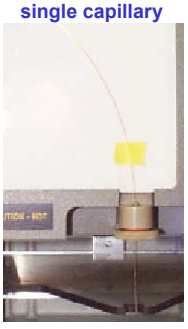
Increasing Sample Throughput with Parallel Processing

ABI 3100
16-capillary array

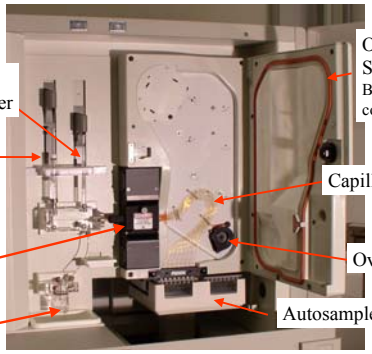


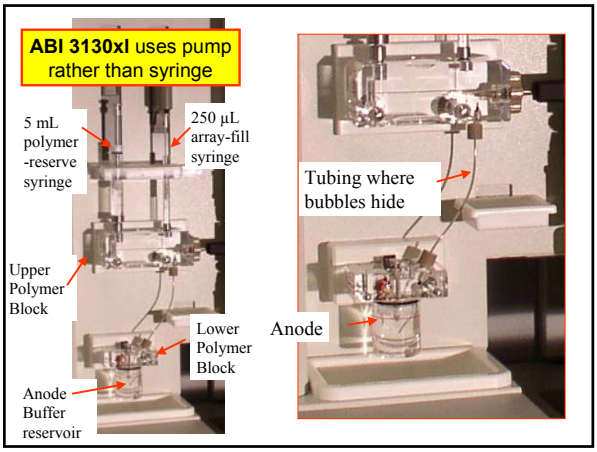
Subtle differences in matrix formation and sizing algorithms – NOT directly equivalent to 310

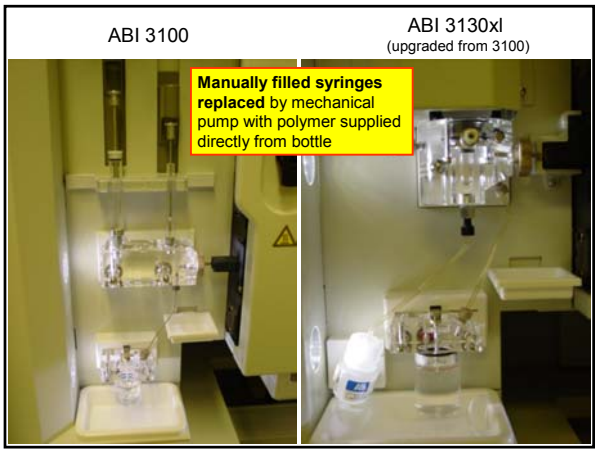
ABI 310
single capillary

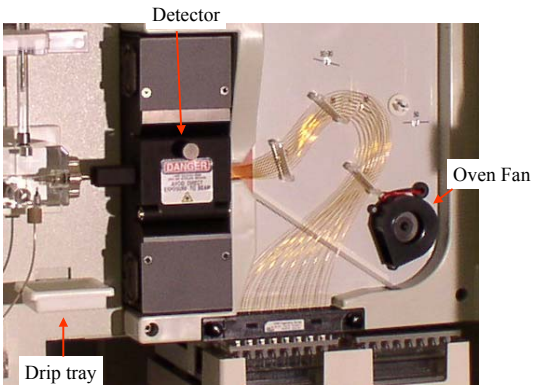


Inside the 3100

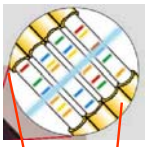




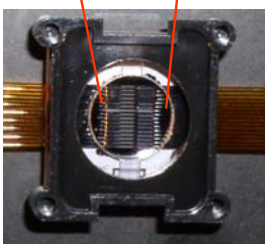


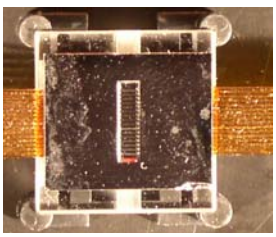


ABI 3100 Array Detection



16 Capillary Array detection cell





Spatial Calibration

Performed after:

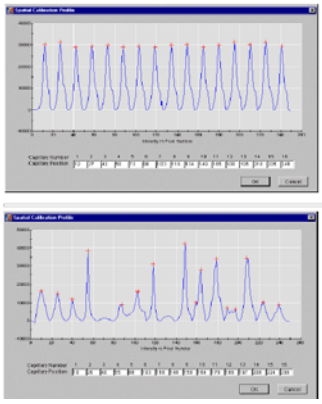
Installing or replacing a capillary array

Removal of the array from the detection block,

(Due to the design, to remove the upper polymer block for cleaning you must remove the Array from the detection window)

Information Provided:

Position of the fluorescence from each capillary on the CCD



Spatial Results

Good Results

Bad results
Try again

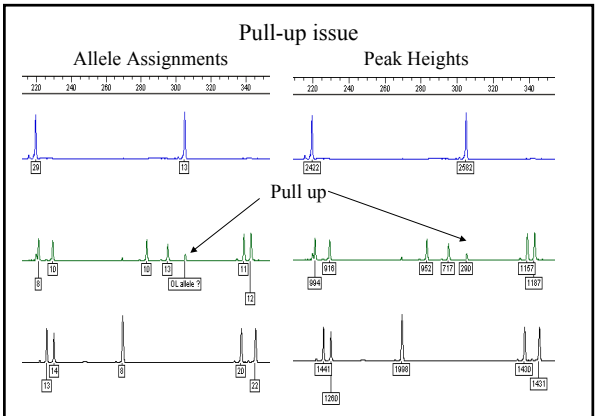
Maintenance of ABI 3100

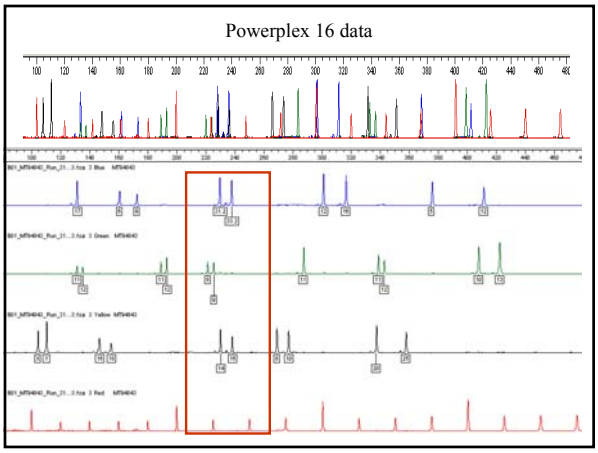
- Syringe – leaks cause capillary to not fill properly
- Capillary storage & wash – **it dries, it dies!**
- Pump block – cleaning helps insure good fill
- Change the running buffer regularly

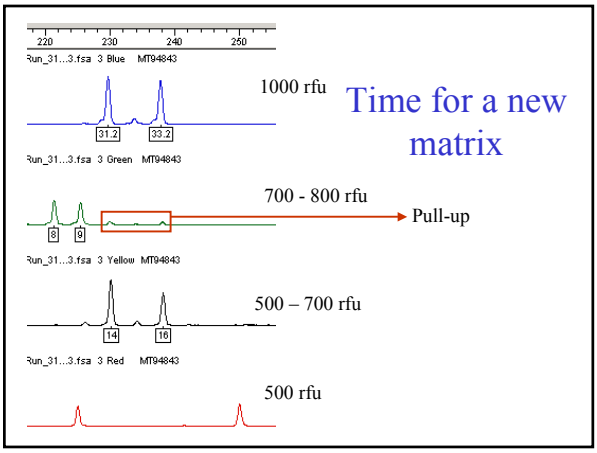
YOU MUST BE CLEAN AROUND A CE!

Spectral Calibration

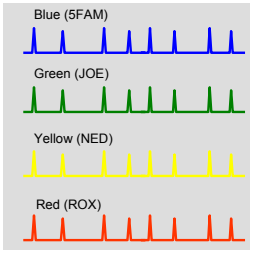
- Performed:
  - New dye set on the instrument
  - After Laser or CCD camera has been realigned
  - You begin to see a decrease in the spectral separation (pull-up, pull-down).
- You must have a valid separation matrix on the instrument prior to running samples.





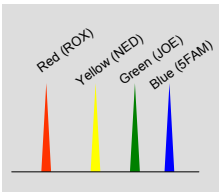


(A) **ABI 310 Matrix Samples**



Separate samples run for each dye color
Each sample contains multiple peaks
All peaks labeled with the same dye color

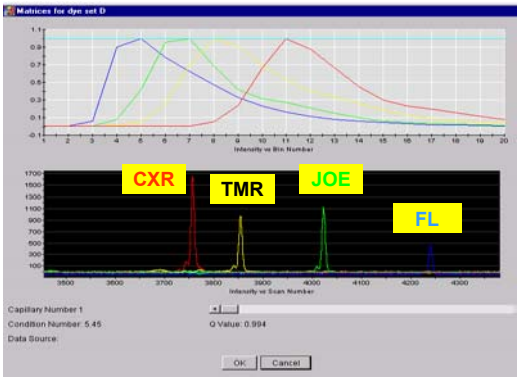
(B) **ABI 3100 Matrix (Spectral Calibration) Sample**



Single sample run containing all dye colors
Only one peak per dye color
Injected into each capillary of the array

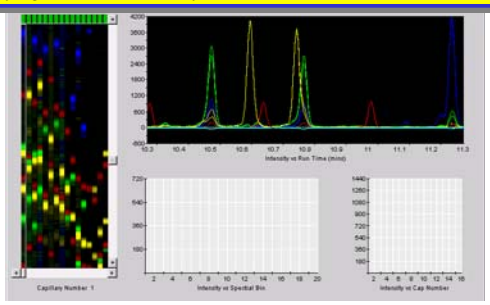
A separate spectral calibration file is created for each capillary

Defining the Matrix on the ABI 3100

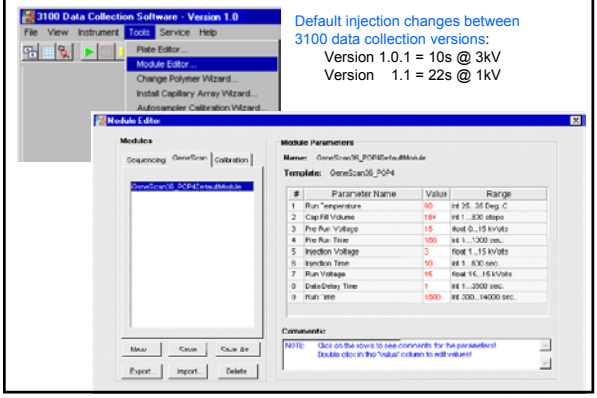


Data from ABI 3100 During the Run

Matrix is applied during the data collection so if there is a problem, the sample must be REINJECTED **after** a new matrix is applied rather than applying a new matrix to any raw data as can be done on the ABI 310…



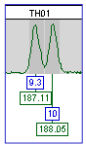
Parameters in Run Modules



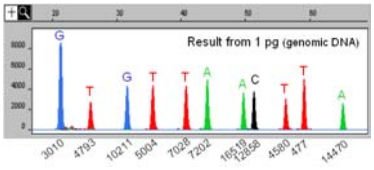
Default injection changes between 3100 data collection versions:
Version 1.0.1 = 10s @ 3kV
Version 1.1 = 22s @ 1kV

NIST ABI 3100 Analysis Using POP6 Polymer

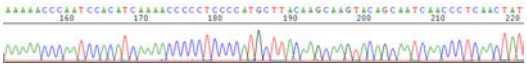
High Resolution STR Typing



SNaPshot SNP Typing
(Coding Region mtSNP 11plex minisequencing assay)



mtDNA Sequencing (HV1)



Conclusions

DNA typing by capillary electrophoresis involves:

1) The use of entangled polymer buffers

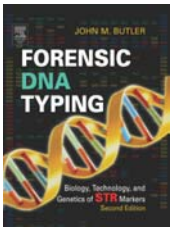
2) Injection by sample stacking

3) Multichannel laser induced fluorescence

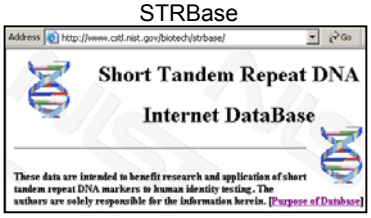
4) Internal and external calibration

If you want to know more…

- *Forensic DNA Typing: Biology and Technology behind STR Markers*
- NIST website: http://www.cstl.nist.gov/biotech/strbase
- John Butler email: john.butler@nist.gov



STRBase



Acknowledgments

NIST Human Identity Project Team

John Butler (Leader) | Margaret Kline | Pete Vallone | Jan Redman | Amy Decker | Becky Hill | Dave Duewer

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

Many wonderful collaborators from industry, university, and government laboratories.

Bruce McCord (Florida International University) for many of the slides

Spanish Translation: Lilliana Moreno (Florida International University)

Thank you for your attention…

Our team publications and presentations are available at:
http://www.cstl.nist.gov/biotech/strbase/NISTpub.htm

Questions?

http://www.cstl.nist.gov/biotech/strbase
john.butler@nist.gov