Training on STR Typing Using Commercial Kits and ABI 310/3100

Margaret C. Kline, Janette W. Redman,
John M. Butler

National Institute of Standards and Technology

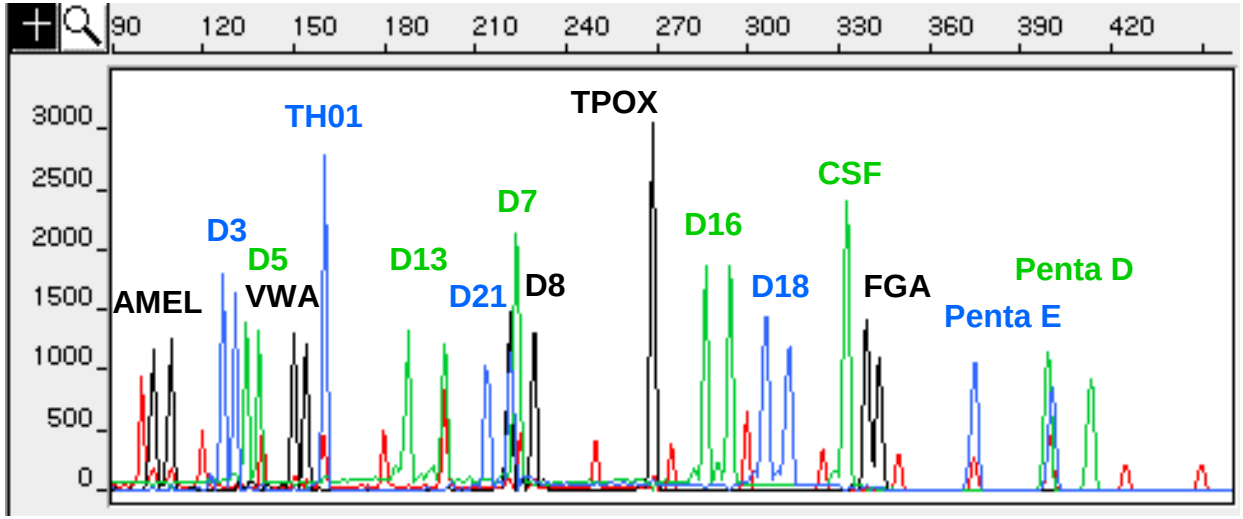
October 22-26, 2001

Human Identity Testing

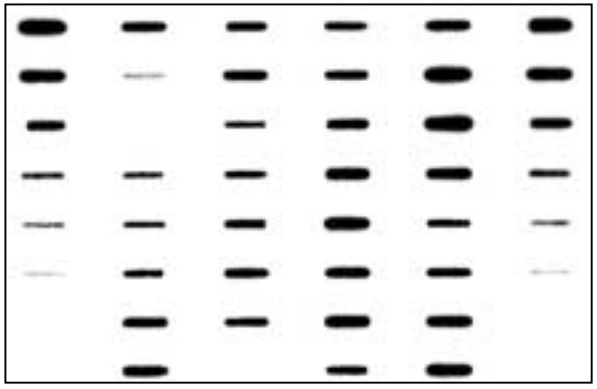
- Forensic cases -- **matching suspect with evidence**
- Paternity testing -- **identifying father**
- Historical investigations
- Missing persons investigations
- Mass disasters -- **putting pieces back together**
- Military DNA “dog tag”
- Convicted felon DNA databases

Overview of Steps Involved in DNA Typing

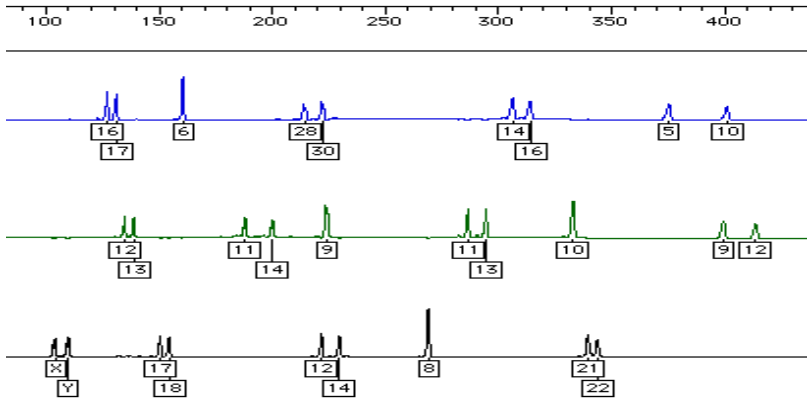
Blood Stain



PCR Amplification with Fluorescent STR Kits and Separation with Capillary Electrophoresis

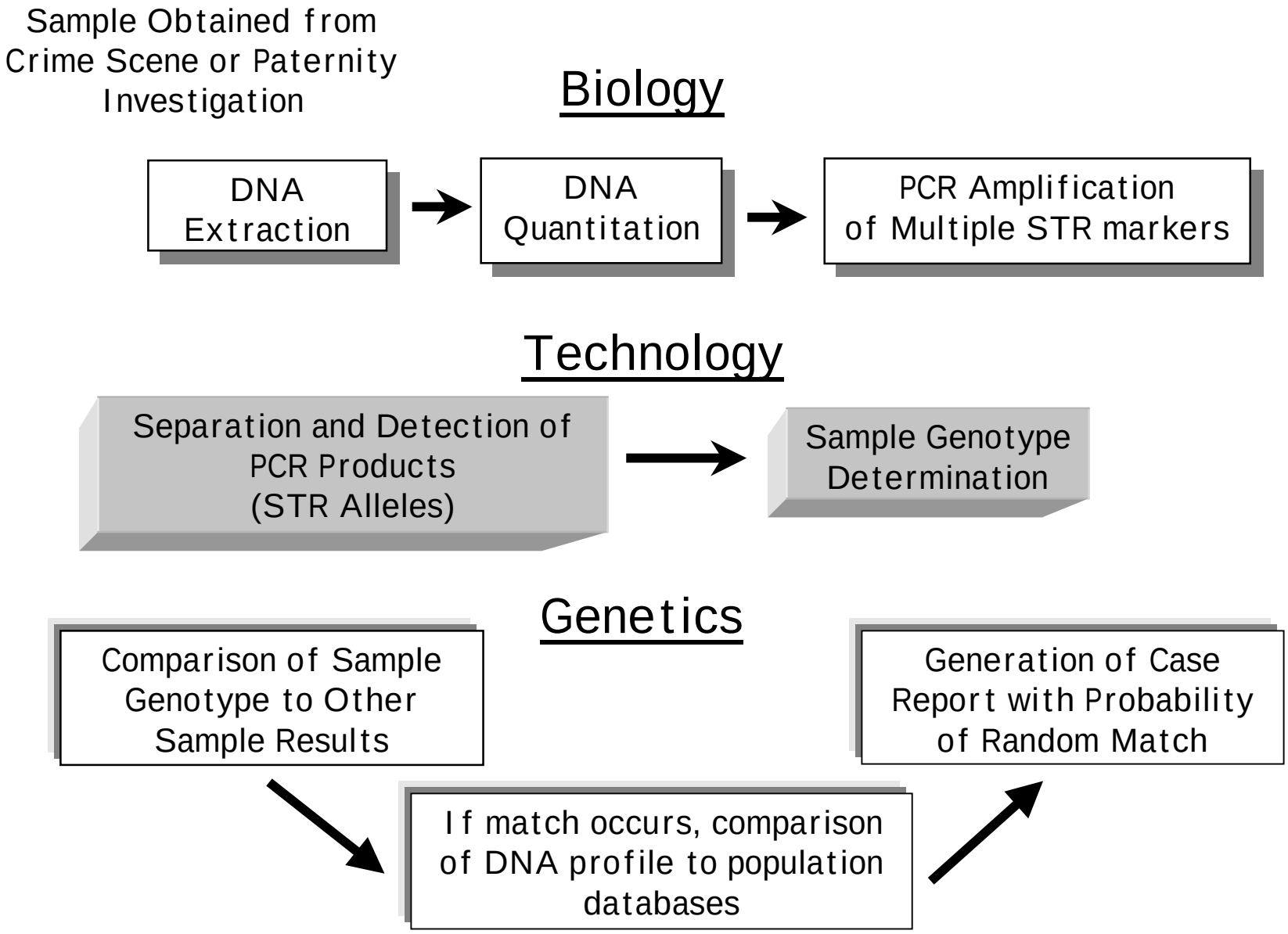


DNA Quantitation using Slot Blot



Genotyping by Comparison to Allelic Ladder

Steps in Sample Processing



Calculation of DNA Quantities in Genomic DNA

Important values for calculations:
1 bp = 618 g/mol A: 313 g/mol; T: 304 g/mol; A-T base pairs = 617 g/mol
G: 329 g/mol; C: 289 g/mol; G-C base pairs = 618 g/mol

1 genome copy = ~3 x 10^9 bp = 23 chromosomes (one member of each pair)

1 mole = 6.02 x 10^{23} molecules

Standard DNA typing protocols with PCR amplification of STR markers typically ask for 1 ng of DNA template. **How many actual copies of each STR locus exist in 1 ng?**

1 genome copy = (~3 x 10^9 bp) x (618 g/mol/bp) = 1.85 x 10^{12} g/mol

= (1.85 x 10^{12} g/mol) x (1 mole/6.02 x 10^{23} molecules)

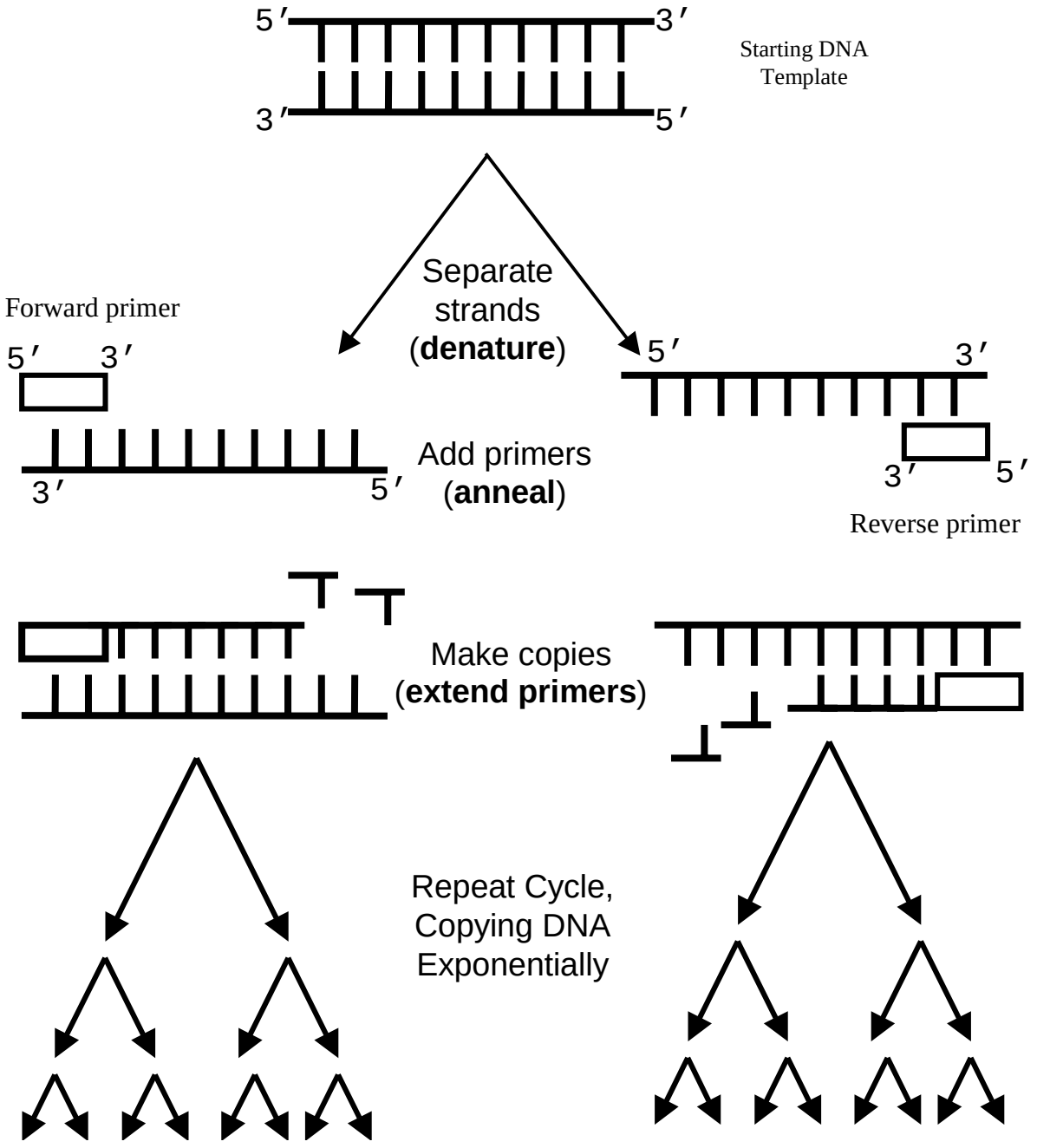
= 3.08 x 10^{-12} g = **3.08 picograms (pg)**

Since a diploid human cell contains two copies of each chromosome, then

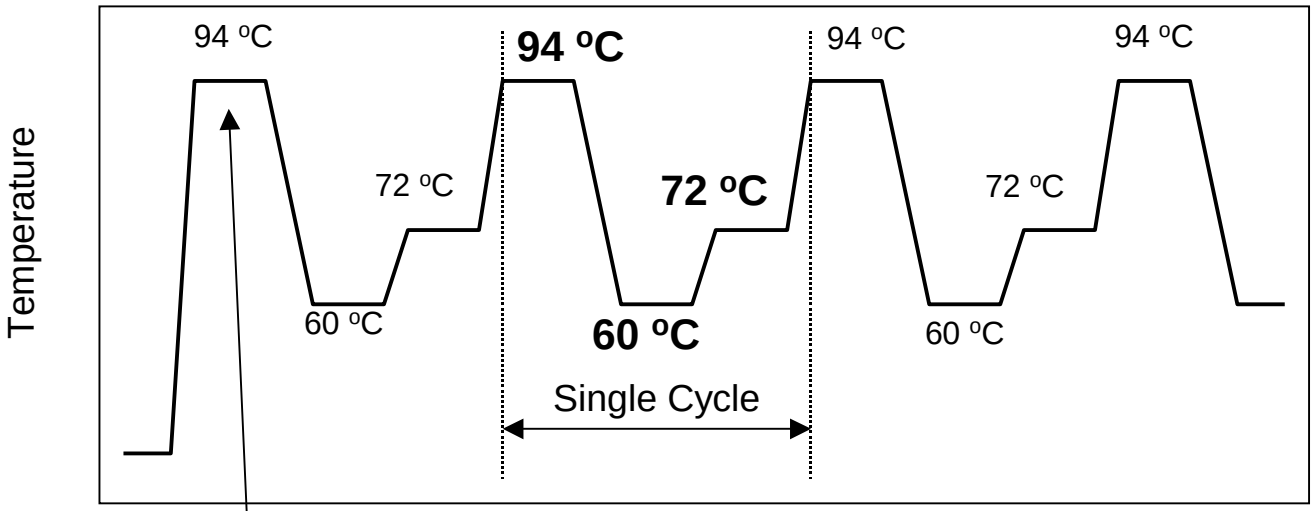
each diploid human cell contains ~6 pg genomic DNA

∴ 1 ng genomic DNA (1000 pg) = ~333 copies of each locus **(2 per 167 diploid genomes)**

PCR Process


5′
3′
3′
5′
Starting DNA Template
Separate strands (**denature**)
Forward primer
5′
3′
3′
5′
Add primers (**anneal**)
5′
3′
3′
5′
Reverse primer
Make copies (**extend primers**)
Repeat Cycle, Copying DNA Exponentially

Thermal Cycling Temperatures



Typically 25-35 cycles performed during PCR

The denaturation time in the first cycle is lengthened to ~10 minutes when using AmpliTaq Gold to perform a “hot-start” PCR

Thermal Cycling Parameters

Step in Protocol	AmpFlSTR® kits (Applied Biosystems)	GenePrint® STR kits (Promega Corporation)
Initial Incubation	95 °C for 11 minutes	95 °C for 11 minutes
Thermal Cycling	**28 cycles**	**30 cycles[a]**
Denature	94 °C for 1 minute	94 °C for 30 seconds (cycle 1-10) 90 °C for 30 seconds (cycle 11-30)
Anneal	59 °C for 1 minute	60 °C for 30 seconds
Extend	72 °C for 1 minute	70 °C for 45 seconds
Final Extension	60 °C for 45 minutes	60 °C for 30 minutes
Final Soak	25 °C (until samples removed)	4 °C (until samples removed)

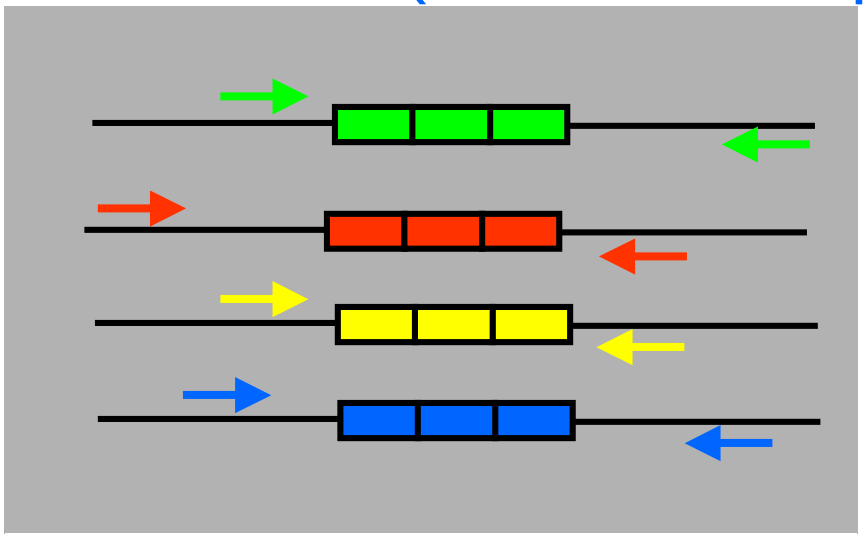
[a] The first 10 cycles are run with a denaturation temperature of 94 °C and the last 20 cycles are run at 90 °C instead. The Promega PowerPlex 1.1 and 2.1 kits also use specific ramp times between the different temperature steps that differ from the conventional 1 °C/second.

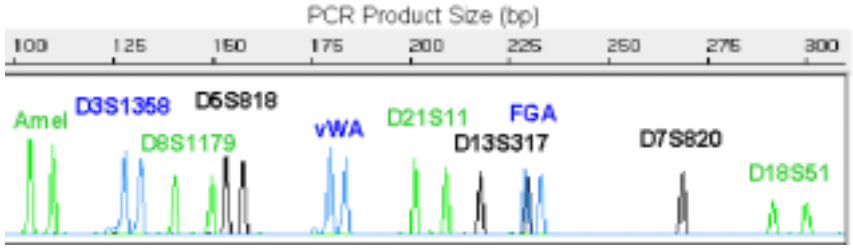
Advantages of PCR

- Minute amounts of DNA template may be used from as little as a single cell.
- DNA degraded to fragments only a few hundred base pairs in length can serve as effective templates for amplification.
- Large numbers of copies of specific DNA sequences can be amplified simultaneously with multiplex PCR reactions.
- Contaminant DNA, such as fungal and bacterial sources, will not amplify because human-specific primers are used.
- Commercial kits are now available for easy PCR reaction setup and amplification.

Multiplex PCR

(Parallel Sample Processing)





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

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)

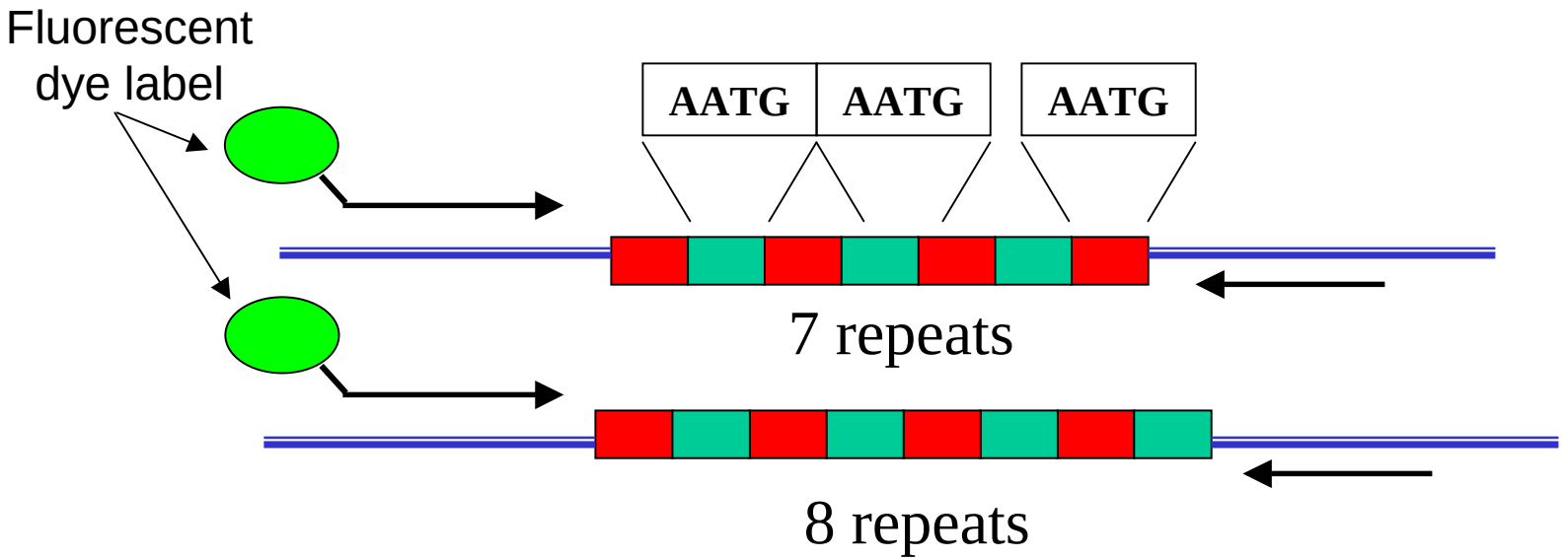
Potential Pitfalls of PCR

- The target DNA template may not amplify due to the presence of PCR inhibitors in the extracted DNA
- Amplification may fail due to sequence changes in the primer binding region of the genomic DNA template
- Contamination from other human DNA sources besides the forensic evidence at hand or previously amplified DNA samples is possible without careful laboratory technique and validated protocols

Tips for Avoiding Contamination

- Pre- and post-PCR sample processing areas should be physically separated.
- Equipment, such as pipettors, and reagents for setting up PCR should be kept separate from other lab supplies, especially those used for analysis of PCR products.
- Disposable gloves should be worn and changed frequently.
- Reactions may also be set up in a laminar flow hood, if available.
- Aerosol-resistant pipet tips should be used and changed on every new sample to prevent cross-contamination during liquid transfers.
- Reagents should be carefully prepared to avoid the presence of any contaminating DNA or nucleases.
- Ultraviolet irradiation of laboratory PCR set-up space when the area is not in use and cleaning workspaces and instruments with isopropanol and/or 10% bleach solutions help to insure that extraneous DNA molecules are destroyed prior to DNA extraction or PCR set-up

Short Tandem Repeats (STRs)



the repeat region is variable between samples while the flanking regions where PCR primers bind are constant

Homozygote = both alleles are the same length

Heterozygote = alleles differ and can be resolved from one another

Primer positions define PCR product size

STR Repeat Nomenclature

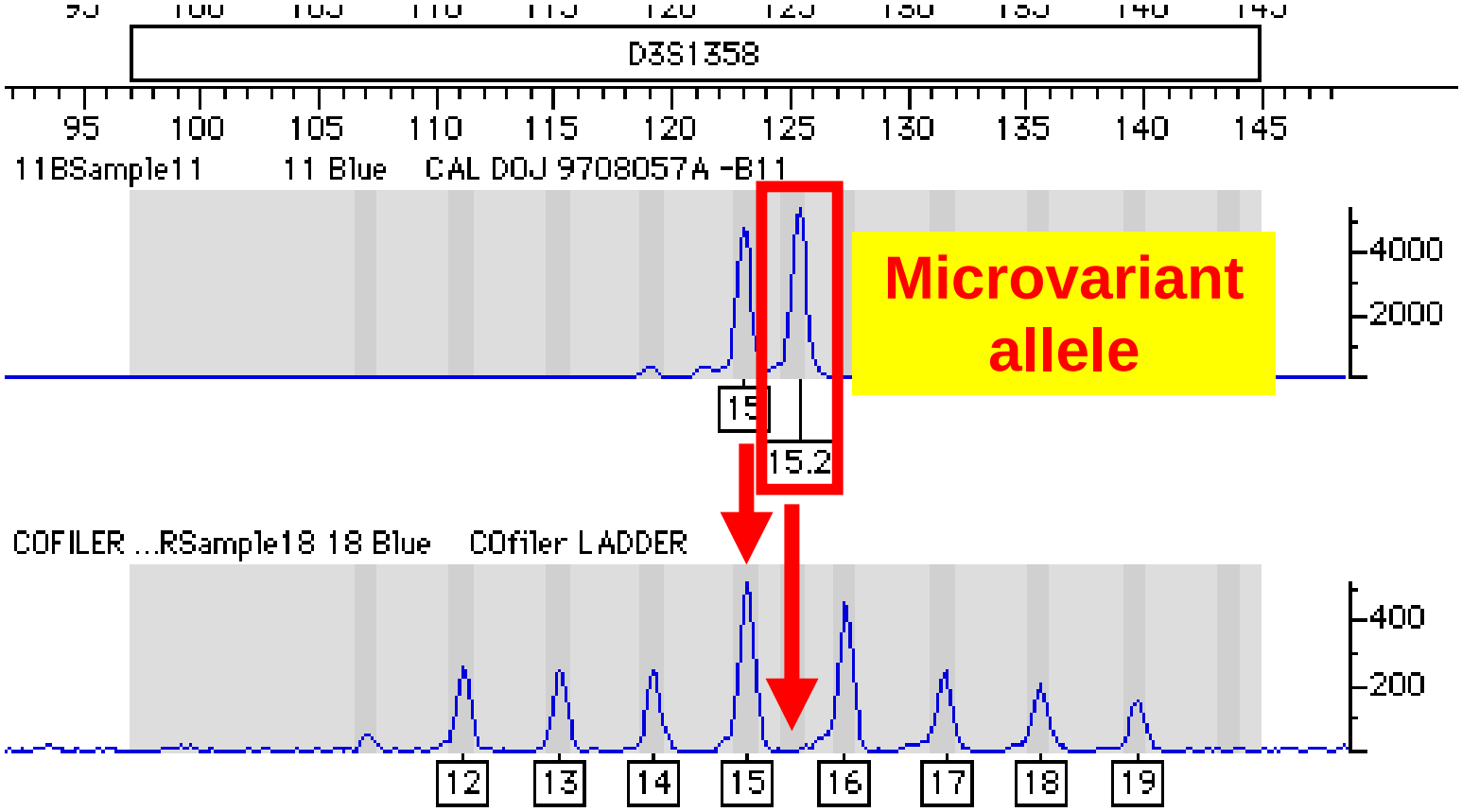
International Society of Forensic Haemogenetics (ISFH) -- *Int. J. Legal Med.* (1997) 110:175-176

- For sequences within genes, use the coding strand
- For other sequences, select the first GenBank database entry or original literature description
- Define the repeat sequence which will provide the largest number of consecutive repeats
- If two sequences are repeated, include both motifs in determining the repeat number
- Microvariants: should be designated by the number of complete repeats and the number of base pairs of the partial repeat separated by a decimal point (Int. J. Legal Med. 1994, 107:159-160) *e.g. TH01 allele 9.3*

Why STRs are Preferred Genetic Markers

- Rapid processing is attainable
- Abundant throughout the genome
- Highly variable within various populations
- Small size range allows multiplex development
- Discrete alleles allow digital record of data
- Allelic ladders simplify interpretation
- PCR allows use of small amounts of DNA material
- Small product size compatible with degraded DNA

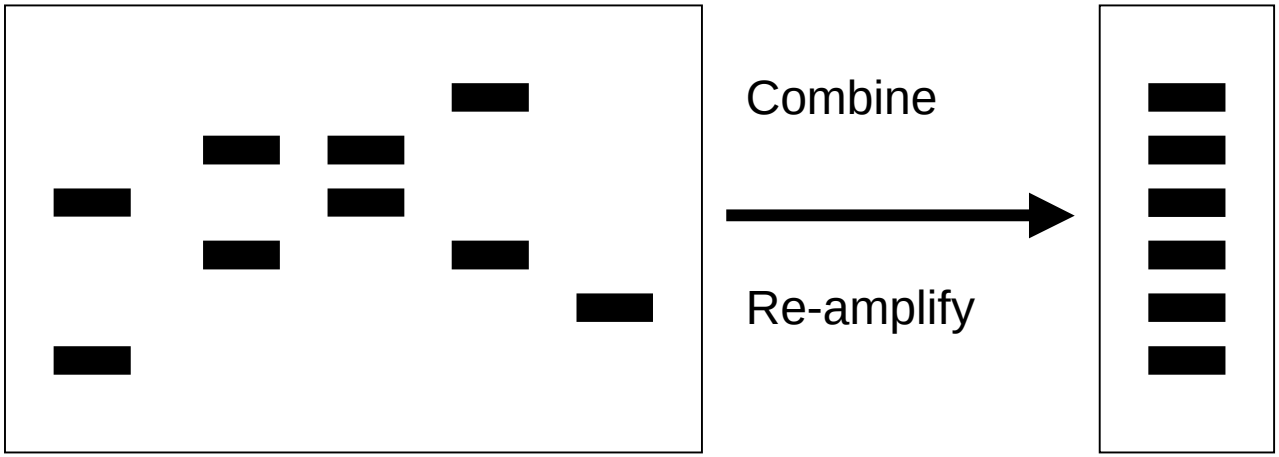
STR genotyping is performed by comparison of sample data to allelic ladders



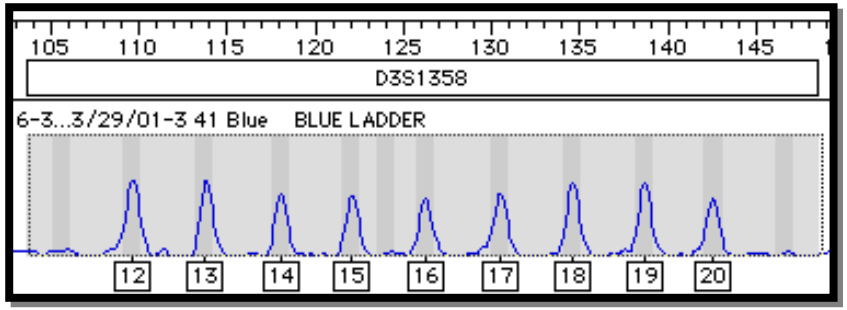
Allelic Ladder Formation

Separate PCR products from various samples amplified with primers targeted to a particular STR locus

Polyacrylamide Gel

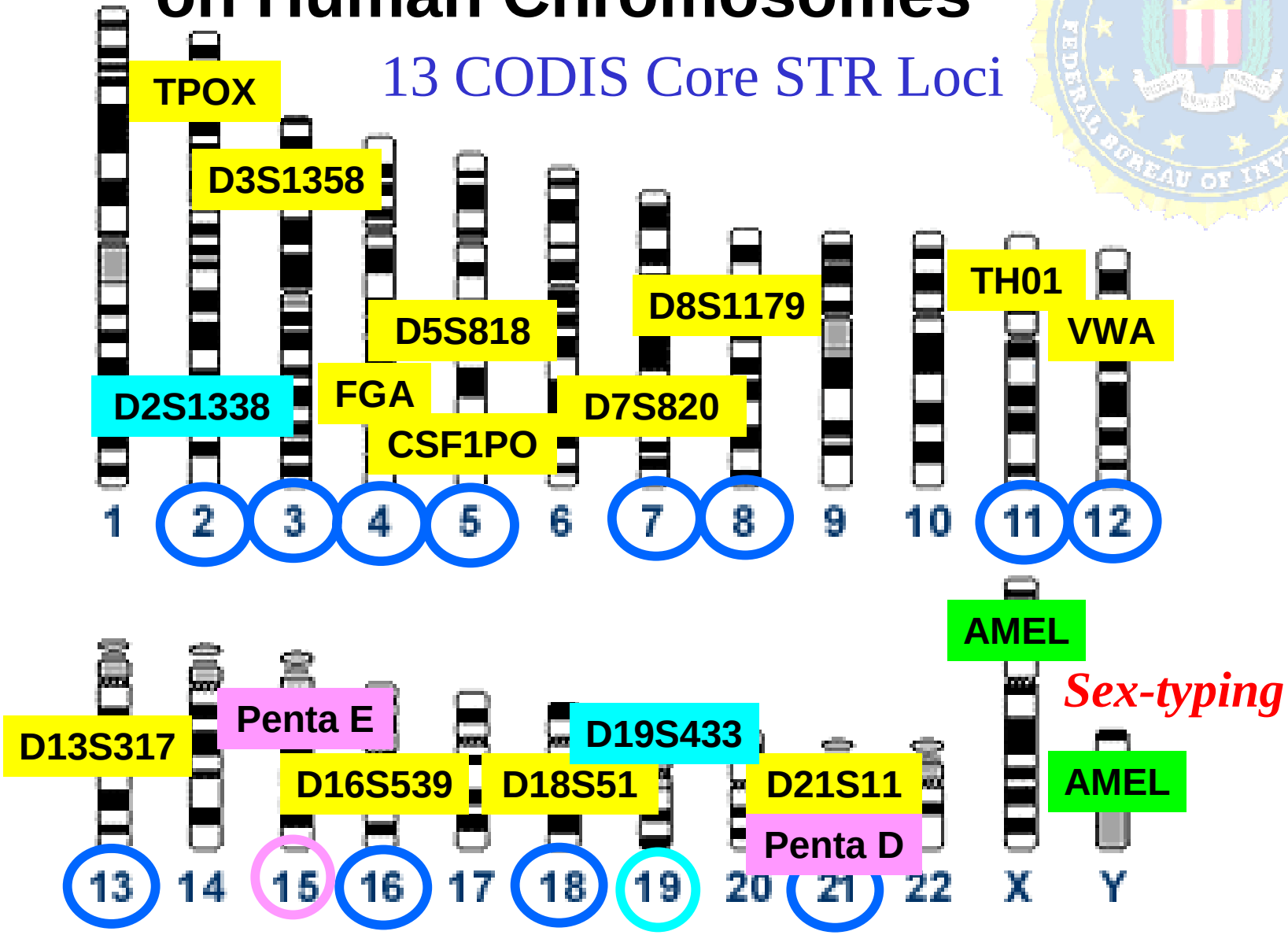


Find representative alleles spanning population variation



Position of Forensic STR Markers on Human Chromosomes

13 CODIS Core STR Loci



Information on 13 CODIS STRs

Locus Name	Chromosomal Location	Repeat Motif ISFH format	GenBank Accession	Allele in GenBank	Allele Range	Number of Alleles Seen*
CSF1PO	5q33.3-34	TAGA	X14720	12	6-16	15
FGA	4q28	CTTT	M64982	21	15-51.2	69
TH01	11p15.5	TCAT	D00269	9	3-14	20
TPOX	2p23-pter	GAAT	M68651	11	6-13	10
VWA	12p12-pter	[TCTG][TCTA]	M25858	18	10-24	28
D3S1358	3p	[TCTG][TCTA]	NT_005997	18	9-20	20
D5S818	5q21-31	AGAT	G08446	11	7-16	10
D7S820	7q11.21-22	GATA	G08616	12	6-15	22
D8S1179	8	[TCTA][TCTG]	G08710	12	8-19	13
D13S317	13q22-31	TATC	G09017	13	5-15	14
D16S539	16q24-qter	GATA	G07925	11	5-15	10
D18S51	18q21.3	AGAA	L18333	13	7-27	43
D21S11	21q21	Complex [TCTA][TCTG]	AP000433	29	24-38	70

* *Butler, Forensic DNA Typing, Appendix 1*

Probability of a Random Match Using 13 CODIS STR Markers

STR Marker	#Alleles	Random match probability (FBI Caucasian)
CSF1PO	11	0.112
FGA	19	0.036
TH01	7	0.081
TPOX	7	0.195
VWA	10	0.062
D3S1358	10	0.075
D5S818	10	0.158
D7S820	11	0.065
D8S1179	10	0.067
D13S317	8	0.085
D16S539	8	0.089
D18S51	15	0.028
D21S11	20	0.039
	Product	0.00000000000001683
	One in	594,059,679,247,540
		1 in 594 trillion

Commercial STR Kits

<u>Kit Contents</u>:

Allelic Ladders for Genotyping

PCR Component Mix

Primer Mix

Positive Control DNA Sample

Cost to User: $15-30 per DNA sample tested

Currently 2 Suppliers: Applied Biosystems and Promega Corporation

Value of STR Kits

Advantages

- Quality control of materials is in the hands of the manufacturer
- 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

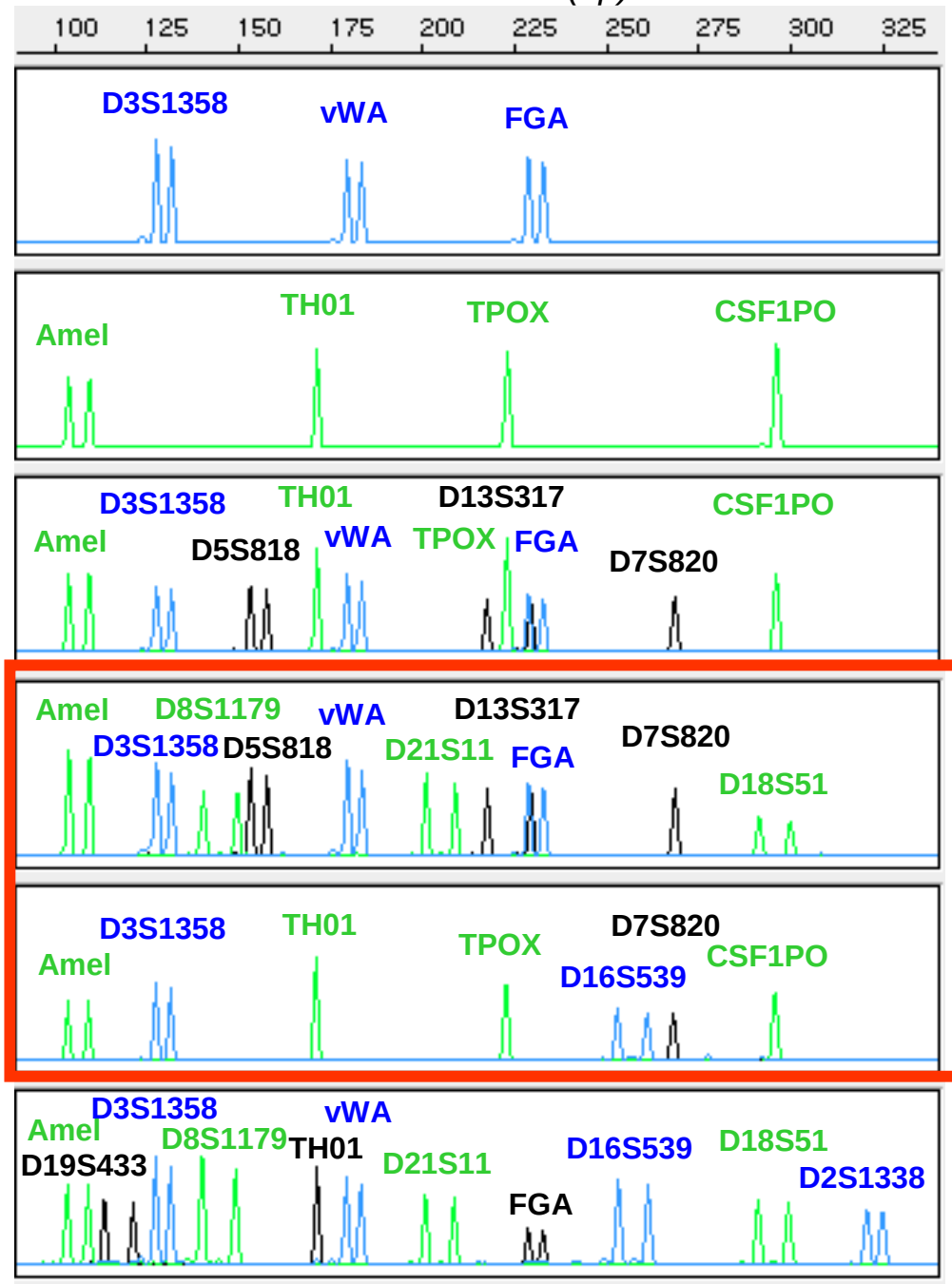
FSS: 5X higher cost with SGM Plus kit

Commercially Available STR Kits

Name	Source	Release Date	STR Loci Included	*Power of Discrimination
TH01, TPOX, CSF1PO monoplexes	Promega	Feb 1993	TH01, TPOX, CSF1PO	**1:410**
AmpFlSTR® Blue	Applied Biosystems	Oct 1996	D3S1358, VWA, FGA	**1:5000**
AmpFlSTR® Green I	Applied Biosystems	Jan 1997	Amelogenin, TH01, TPOX, CSF1PO	**1:410**
CTTv	Promega	Jan 1997	CSF1PO, TPOX, TH01, VWA (vWF)	**1:6600**
FFFL	Promega	Jan 1997	F13A1, FES/FPS, F13B, LPL	**1:1500**
GammaSTR	Promega	Jan 1997	D16S539, D13S317, D7S820, D5S818	**$1:1.8x10^4$**
PowerPlex™ (version 1.1 and 1.2 later)	Promega	Jan 1997	CSF1PO, TPOX, TH01, VWA, D16S539, D13S317, D7S820, D5S818	**$1:1.2x10^8$**
AmpFlSTR® Profiler™	Applied Biosystems	May 1997	D3S1358, VWA, FGA, Amelogenin, TH01, TPOX, CSF1PO, D5S818, D13S317, D7S820	**$1:3.6x10^9$**
AmpFlSTR® Profiler Plus™	Applied Biosystems	Dec 1997	D3S1358, VWA, FGA, Amelogenin, D8S1179, D21S11, D18S51, D5S818, D13S317, D7S820	**$1:9.6x10^{10}$**
AmpFlSTR® COfiler™	Applied Biosystems	May 1998	D3S1358, D16S539, Amelogenin, TH01, TPOX, CSF1PO, D7S820	**$1:8.4x10^5$**
AmpFlSTR® SGM Plus™	Applied Biosystems	Feb 1999	D3S1358, VWA, D16S539, D2S1338, Amelogenin, D8S1179, D21S11, D18S51, D19S433, TH01, FGA	**$1:3.3x10^{12}$**
PowerPlex™ 2.1	Promega	Jun 1999	D3S1358, TH01, D21S11, D18S51, VWA, D8S1179, TPOX, FGA, Penta E	**$1:8.5x10^{10}$**
PowerPlex™ 16	Promega	May 2000	CSF1PO, FGA, TPOX, TH01, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, Penta D, Penta E, amelogenin	**$1:1.8x10^{17}$**
AmpFlSTR® Identifiler™	Applied Biosystems	May 2001	CSF1PO, FGA, TPOX, TH01, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11, D19S433, D2S1338, amelogenin	**$1:2.1x10^{17}$**

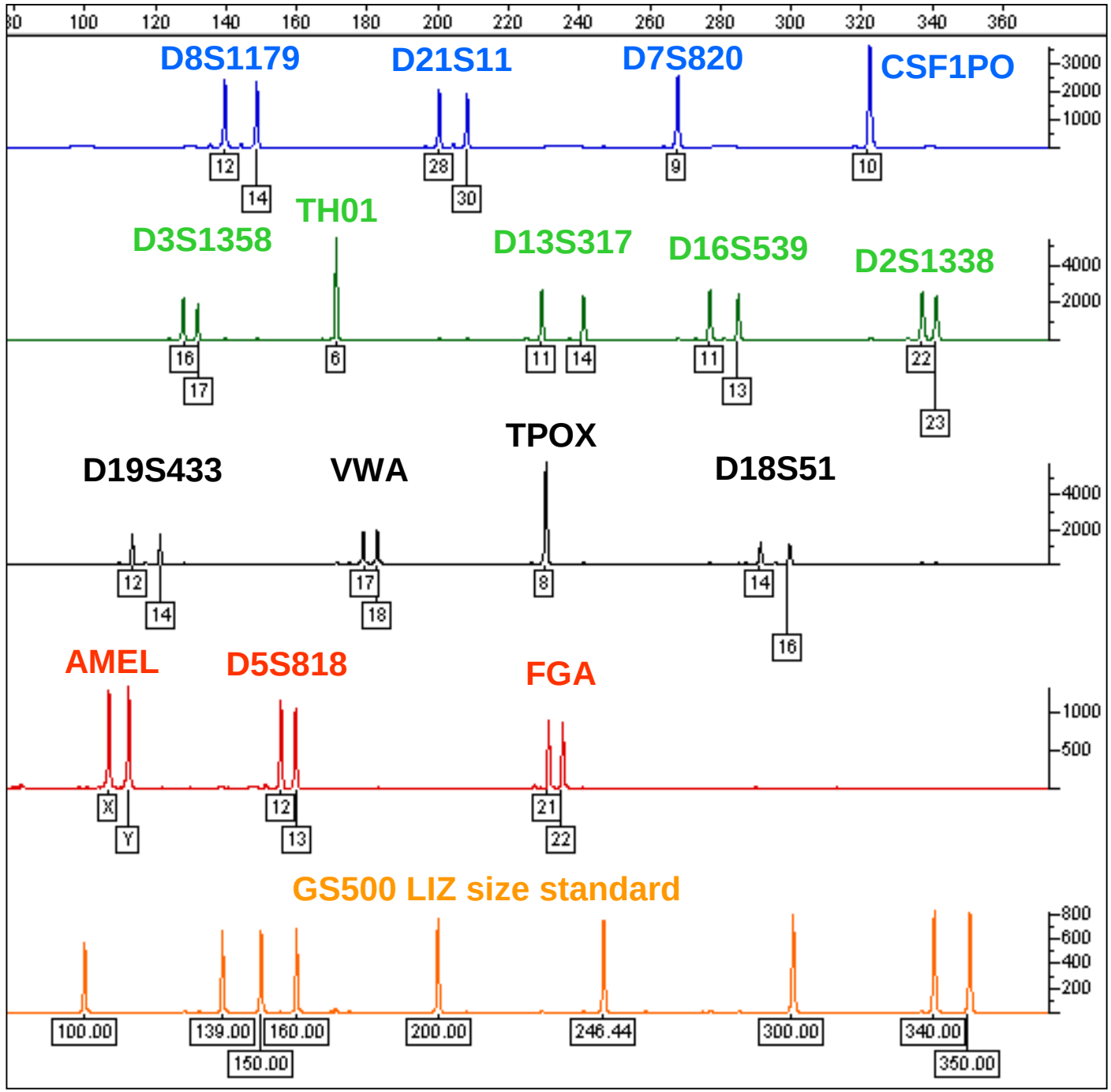
*Caucasian population (rounded to 2 significant figures)

Same DNA Sample Run with Each of the ABI STR Kits



	Power of Discrimination
Blue	1:5000
Green I	1:410
Profiler™	$1:3.6 \times 10^{9}$
Profiler Plus™	$1:9.6 \times 10^{10}$
COfiler™	$1:8.4 \times 10^{5}$
SGM Plus™	$1:3.3 \times 10^{12}$

AmpFlSTR®
Identifiler™



Requirements for Accurate STR Typing

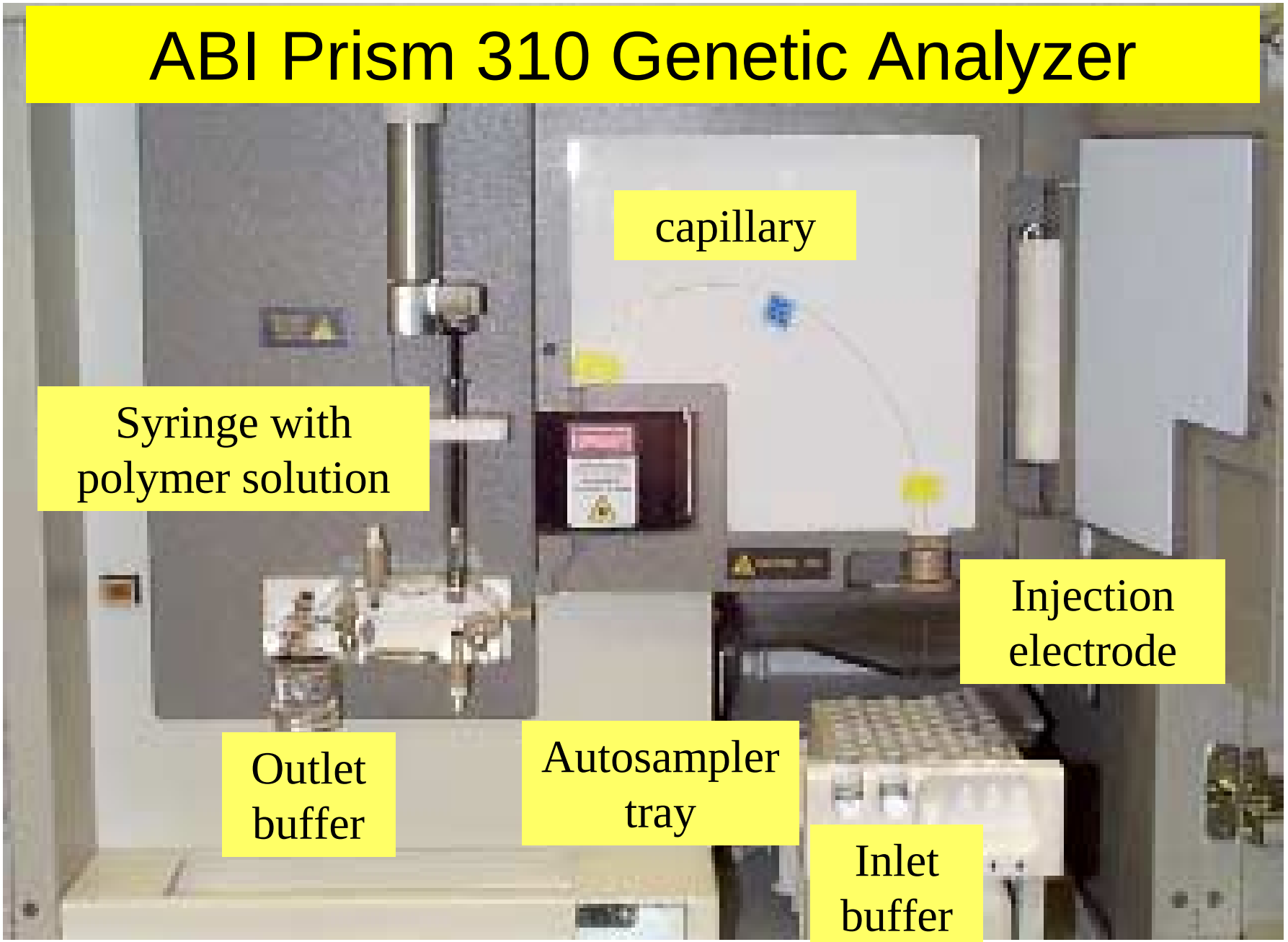
- High precision (to permit comparison of allelic ladders to sequentially processed STR samples)
- Color separation of different dye sets used (to avoid bleed through between different colors)
- Resolution of at least 1 bp to >300 bp (to detect microvariants)
- Reliable sizing over 75-450 bp region

Accurate typing can be achieved with ABI 310

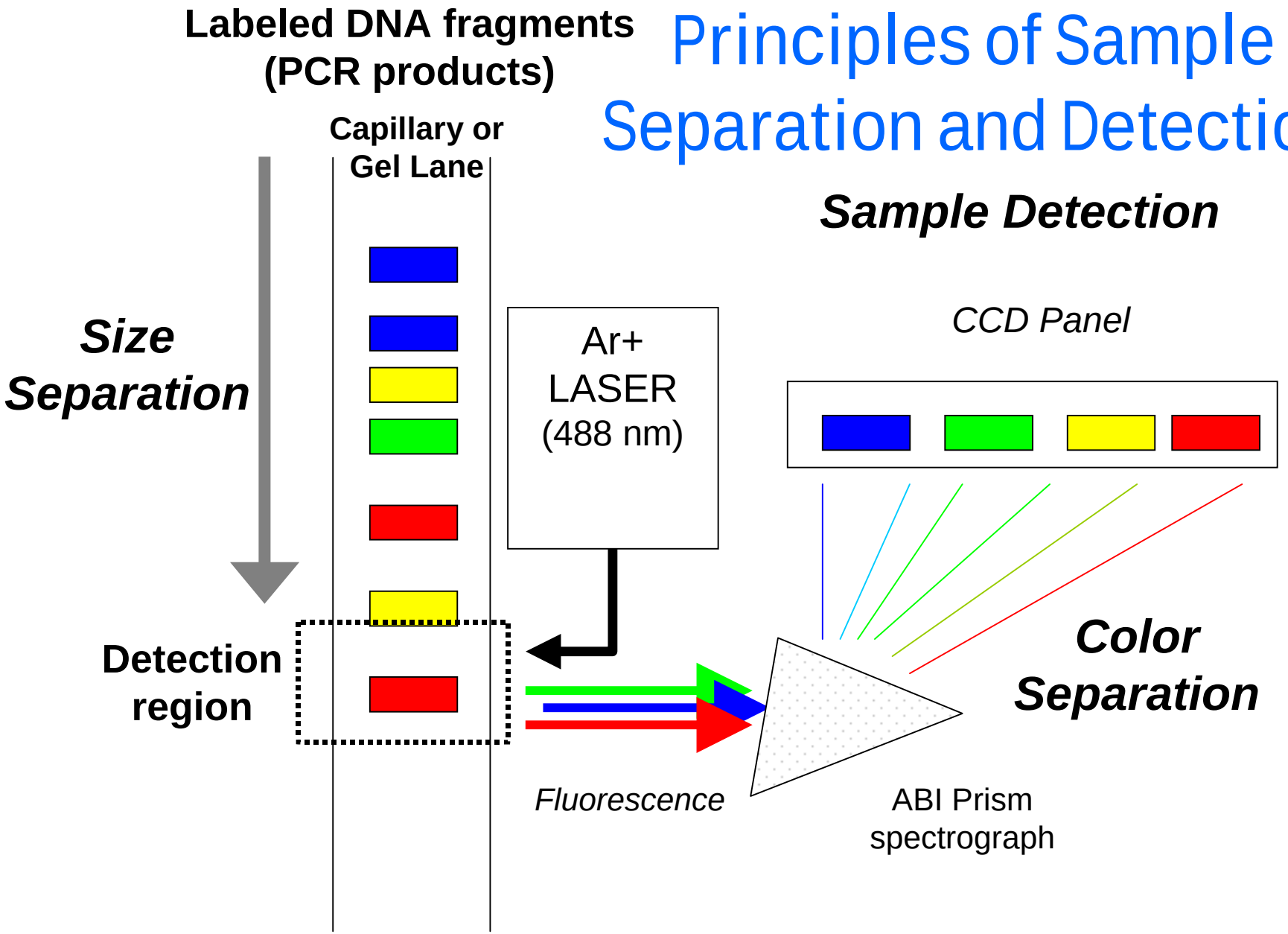
Components of ABI 310

- Chemistry
  - STR kits, fluorescent dyes, matrix samples, capillary, buffers, polymer, formamide
- Hardware
  - CCD camera, laser, electrodes, pump block, hot plate for temperature control, autosampler
- Software
  - Data collection, color separation, peak sizing & calling, genotyping, stutter removal

ABI Prism 310 Genetic Analyzer



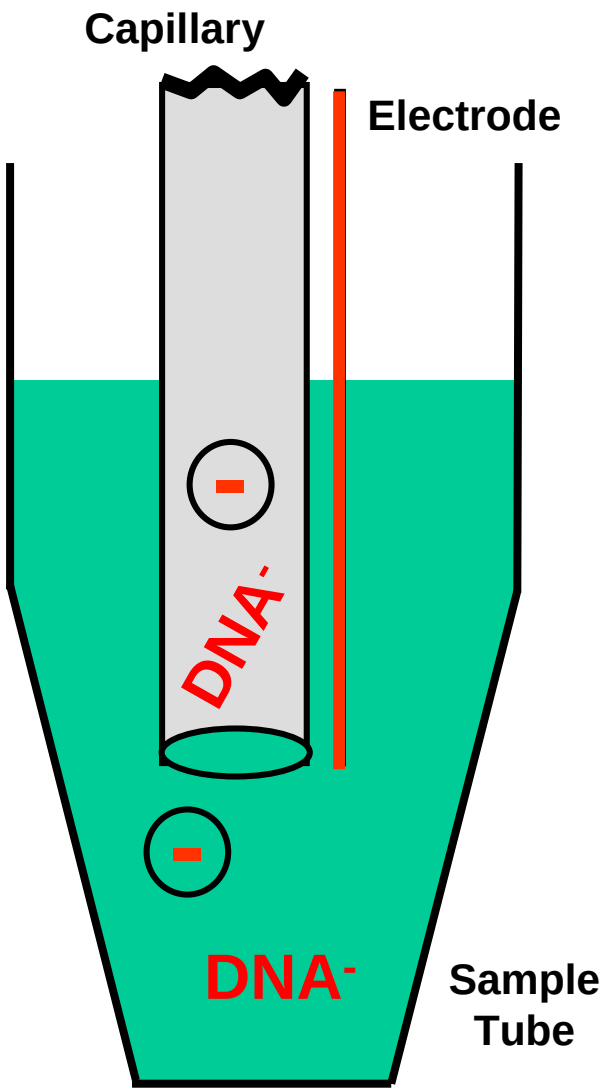
Principles of Sample Separation and Detection



Chemistry Involved

- Injection
  - electrokinetic injection process
  - importance of sample preparation (formamide)
- Separation
  - capillary
  - POP-4 polymer
  - buffer
- Detection
  - fluorescent dyes with excitation and emission traits
  - virtual filters (hardware/software issues)

Electrokinetic Injection Process



$$Q = \frac{\pi r^2 c_s(\mu_{ep} + \mu_{eo})Et\lambda_b}{\lambda_s}$$

Q is the amount of sample injected

r is the radius of the capillary

c_s is the sample concentration

E is the electric field applied

t is the injection time

λ_s is the sample conductivity

λ_b is the buffer conductivity

μ_{ep} is the mobility of the sample molecules

μ_{eo} is the electroosmotic mobility

Rose *et al* (1988) *Anal. Chem*. 60: 642-648

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
  - water works fine but samples are not stable as long as with formamide
- Denaturation with heating and snap cooling
  - use 480 cycler for heating (holds 0.5 mL tubes) and cold aluminum block for snap cooling (instead of ice)
  - heat/cool denaturation step is not always necessary...

Separation Issues

- Run temperature -- 60 ^{o}C helps reduce secondary structure on DNA and improves precision
- Electrophoresis buffer -- urea in running buffer helps keep DNA strands denatured
- Capillary wall coating -- dynamic coating with polymer
- Polymer solution -- POP-4

DNA Separation Mechanism

- Size based separation due to interaction of DNA molecules with entangled polymer strands
- Polymers are not cross-linked (as in slab gels)
- “Gel” is not attached to the capillary wall
- Pumpable -- can be replaced after each run
- Polymer length and concentration determine the separation characteristics

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 collected onto the CCD camera

Laser Used in ABI 310

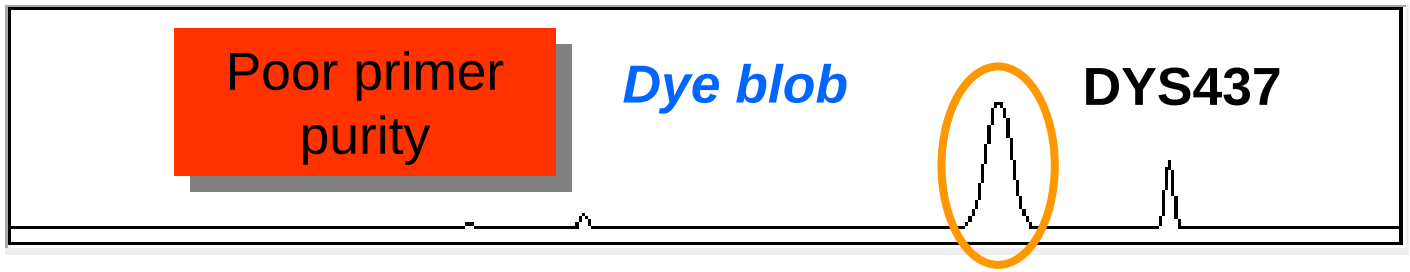
- Argon Ion Laser
- 488 nm and 514.5 nm for excitation of dyes
- 10 mW power
- Lifetime ~5,000 hours (1 year of full-time use)
- Cost to replace ~$5,500
- Leads to highest degree of variability between instruments and is most replaced part
- Color separation matrix is specific to laser used on the instrument

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 usually through a 6-carbon spacer --- Dye-$(CH_2)_6$-primer
- Dye-labeled oligonucleotide is incorporated into PCR product during multiplex PCR amplification giving a specific color "tag" to each PCR product
- Dyes can be spectrally distinguished using virtual filters and CCD imaging to yield different colored peaks in ABI 310 electropherogram

Dye Blobs

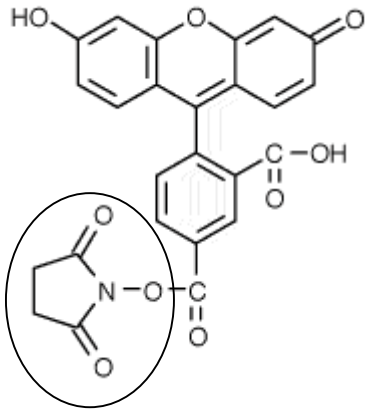
- Free dye (not coupled to primer) can be injected into the CE capillary and interfere with detection of true STR alleles
- Dye blobs are wider and usually of less intensity than true STR alleles (amount depends on the purity of the primers used)
- Dye blobs usually appear at an apparent size that is unique for each dye (e.g., HEX ~170 bp)



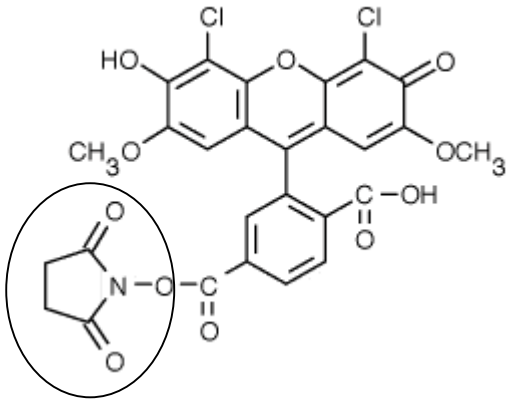
Fluorescent Dyes Used in 4-Color Detection

FL

FAM (Blue)

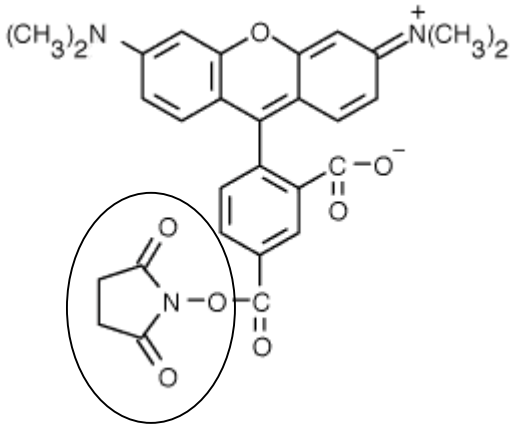


JOE (Green)

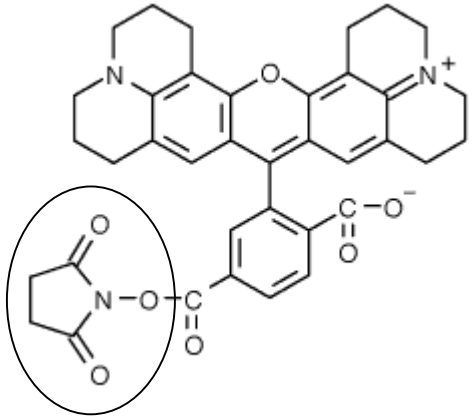


NED

TAMRA (Yellow)

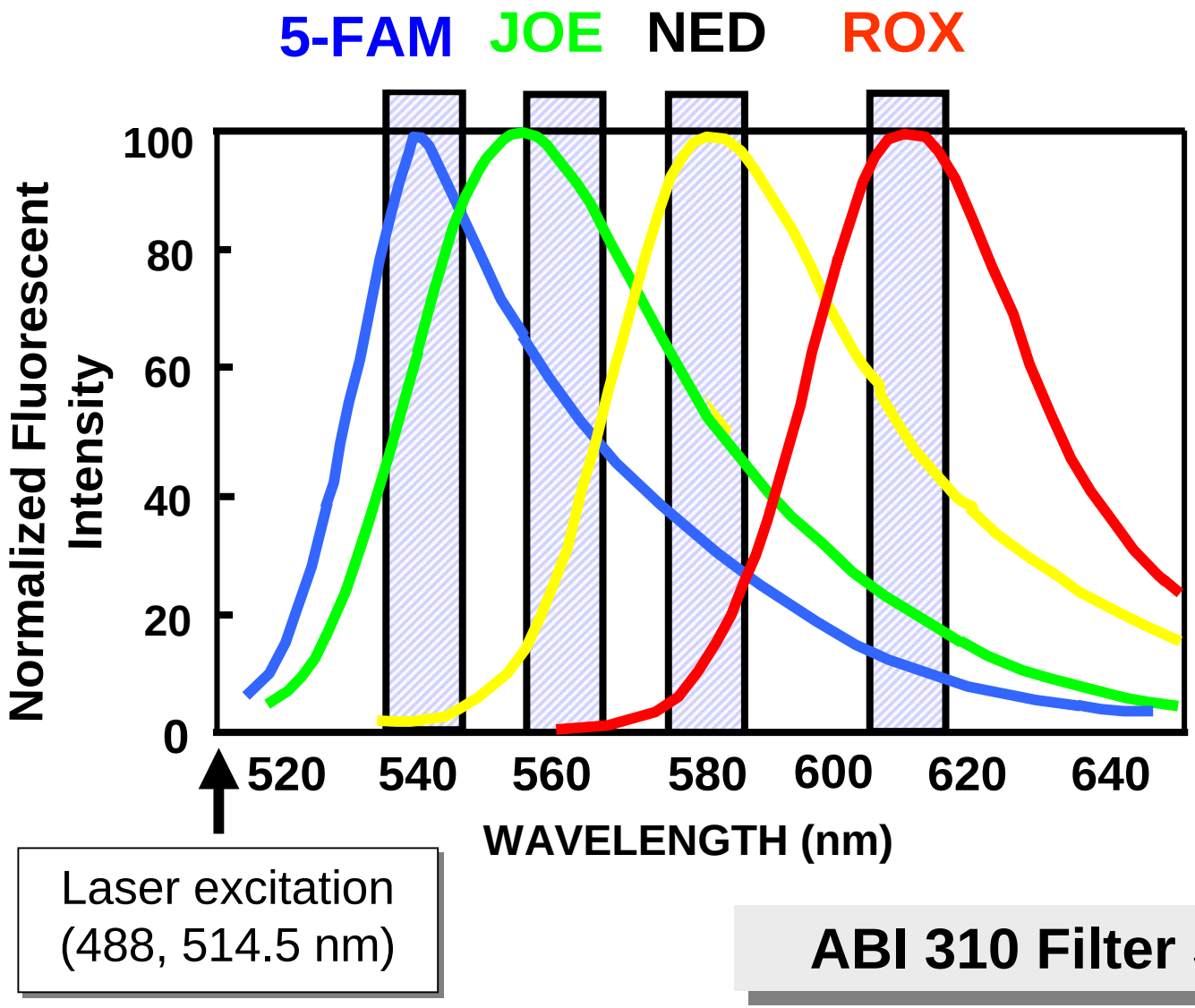


ROX (Red)



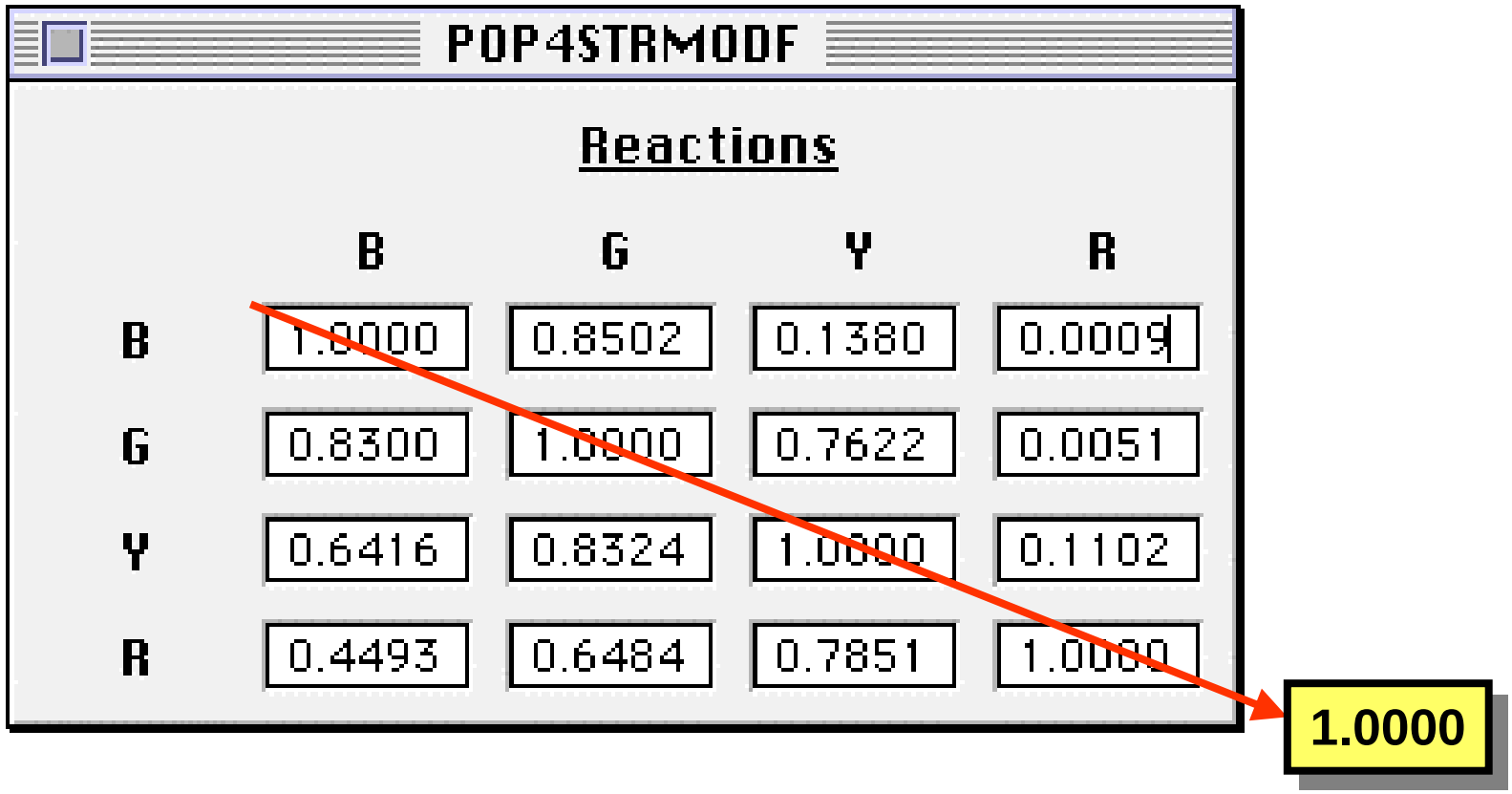
CXR

Fluorescent Emission Spectra for ABI Dyes



ABI 310 Filter Set F

Matrix File Table from an ABI 310



These values are used by the GeneScan® Analysis Software to separate the various dye colors from one another. The letters B, G, Y, and R represent the dye colors Blue, Green, Yellow, and Red, respectively.

Same Dye Set and Filter F with Different ABI 310s

POP4STRMODF

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

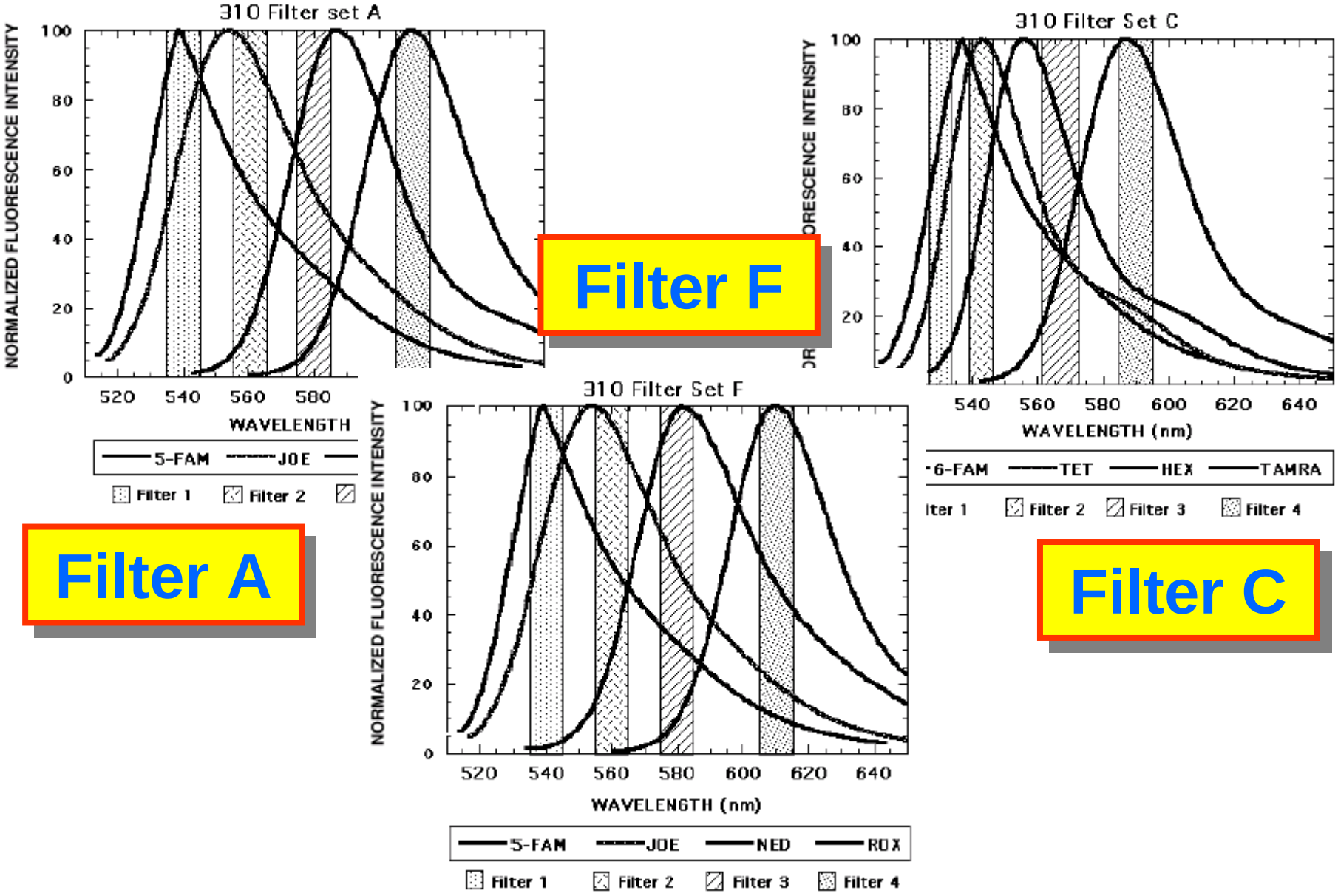
5FAM_JOE_NED_ROX POP4

Reactions

	B	G	Y	R
B	1.0000	0.6444	0.0487	0.0010
G	0.6027	1.0000	0.5556	0.0061
Y	0.3421	0.6146	1.0000	0.1060
R	0.1690	0.3478	0.5791	1.0000

Instrument lasers make a big difference

Primary Filter Sets on the ABI 310



If Wrong Filter Is Used…

Y STR 10plex labeled with FAM, TET, HEX dyes

Filter A

Filter F

Filter C