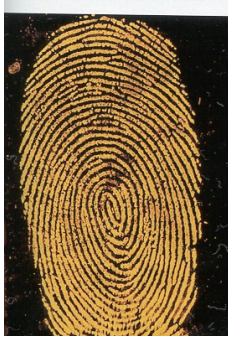
JS 111 Advanced Fluorescence applications- Mitochondrial DNA and Y chromosome markers – Single Nucleotide Polymorphisms



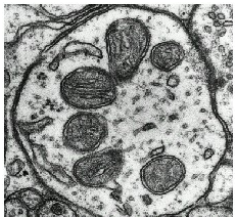
I. Announcements and Assignments
   a. Dr. John DeHaan- Fire Debris arson expert 11/29
   b. Assignments- Reading and Articles

II. Mitochondrial DNA
   a. Biology of mitochondria
   b. DNA sequencing

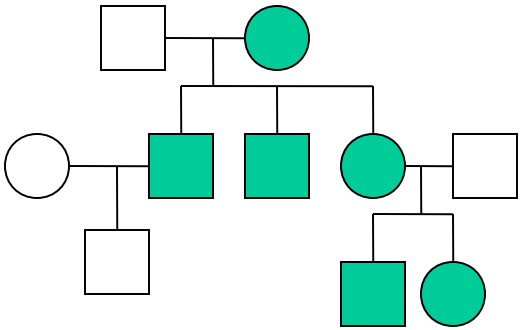
I. Y Chromosome markers : Intro to Y chromosomes- Types of Y polymorphism

II. Single nucleotide polymorphisms (SNPs)
   a. Why SNPs? Intro to Single Nucleotide Polymorphisms (SNPS)
   b. Applications of SNPs
   c. Detection Technologies for Y SNPs in Forensics: Primer Extension, Pyrosequencing, Light Cycling, Mass Spec
   d. Bead based assays-Luminex
   e. Universal Arrays and Bacterial Identification
   f. SNPs vs STRs or SNPs and STRs

Why mtDNA SNPs?



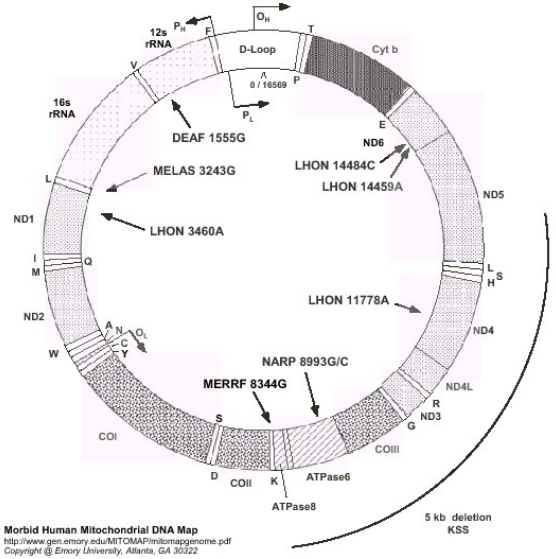
- **Well characterized and studied (population, evolutionary, medical and forensic studies)**



- **Uniparental maternal inheritance**
  missing persons

- **Relatively small size (16kb) and high copy number - low quantity/quality samples (hair, bone, teeth- ancient/degraded)**

- **Implicated in maternally inherited diseases : diabetes, deafness, hypertrophic cardiomyopathy and myopathy**



Assignments and Announcements

- Announcements-
  - Criminalist Isha Brown Weds 9th May- CA DOJ DNA Databank
  - Assignments-
  - Butler Chapters 8-11 Inman, 16, Appendices IV&V, Inman10-11
  - ***Read Article Butler Y chromosome review article posted to the web- Write a 500 word summary with 3Q and 3A – FOR 5 points extra credit***
  - ***Hand in assignment by weds 9th May***

Mitochondrial DNA regions used in forensics

- Hypervariable regions- also known as D-loop or control regions involved in the replication of mtDNA
- MtDNA is in very high copy number in every cell. There are many cells per sample and therefore many more copies than nuclear DNA that has only 1 per cell
- Most forensic laboratories utilize DNA sequencing to analyze mitochondrial DNA polymorphisms

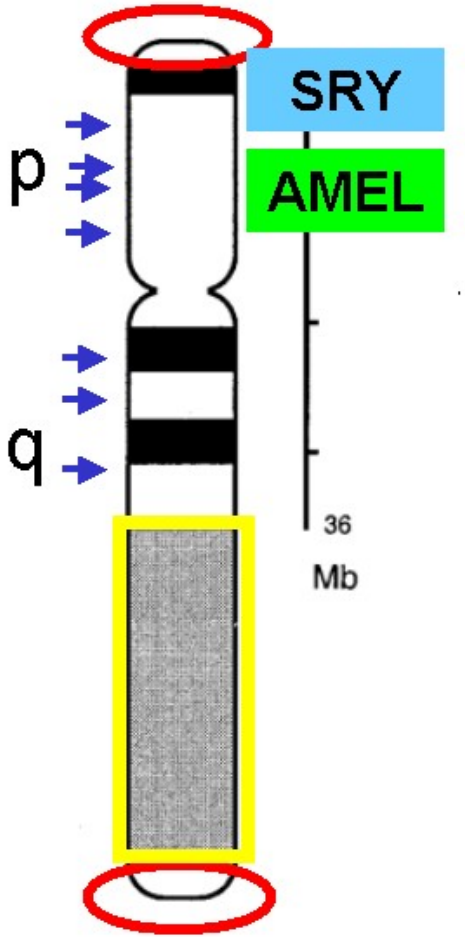
- Intro to Y chromosomes- Types of Y polymorphism
- Intro to Single Nucleotide Polymorphisms (SNPS)
  - Definitions
  - Why SNPs?
  - Applications of SNPs
- Detection Technologies for Y SNPs in Forensics
  - Primer Extension, Pyrosequencing, Light Cycling, Mass Spec
  - Bead based assays-Luminex
    - Universal Arrays and Bacterial Identification
- SNPs vs STRs or SNPs and STRs
  - Either/Or
  - Why SNPs?

Cycle sequencing : PCR in the presence of "bad" dNTPs – dideoxynucleoside triphosphates

- Synthesize DNA in the presence of some Dideoxy nucleotides without a 3-OH
- Building a railroad with some tracks that do not have connectors
- End result is a complete set of fragments that represent every base in the DNA strand
- See animation

Overview of the Y Chromosome

(add picture of Y Chrom from Chris Tyler –Smith's)



Nucleic Acids Res. 28(2), e8 (2000)

- Paternally inherited
- Represents 2% of the human genome
- ~60 Mb in length, 2.5Mb on tips recombine with the X
- 95% of the Y is non-recombining
- Y SNP Consortium - Over 4193 SNPs on the Y chromosome http://ycc.biosci.arizona.edu/

Why study the Y chromosome?

- **Population Genetics[1]**
- **Evolutionary and Genealogical studies[2]**
- **Molecular Ecology[3]**
- **Infertility studies[4]**
- **Forensics[5]**

1 Kivisild et al. 2003. Am J Hum Genet Feb;72(2):313-32 **Mountain JL 2002 Genome Res Nov;12(11):1766-72 SNPSTRs**: empirically derived, rapidly typed, autosomal haplotypes for inference of population history and mutational processes.

2 http://www.oxfordancestors.com/

3 Hellborg L et al. 2003. Mol Ecol Jan;12(1):283-91 Y chromosome conserved anchored tagged sequences (YCATS) for the analysis of mammalian male-specific DNA

4 Kostiner, D.R. et al (1998) Male infertility: analysis of the markers and genes on the human Y chromosome. Hum. Reprod. 13, 3032-3038.

5 Lareu M, Puente J, Sobrino B, Quintans B, Brion M, Carracedo A. 2001 The use of the LightCycler for the detection of Y chromosome SNPs.Forensic Sci Int. 2001 May 15;118(2-3):163-8..Ewis AA, Lee JW, Kuroki Y, Shinka T, Nakahori Y. 2002. Yfm1, a multicopy marker specific for the Y chromosome and beneficial for forensic, population

YY in forensics?

- Bad Boys: 98% of violent crime is committed by men
- Sexual Assault Evidence Screening: Rapid screening of sexual assault evidence : “male specific”- so no differential
- Mixtures: Especially with very low copy male DNA in mixtures. May assist in determining single or multiple donors in difficult mixtures
- No spermatozoa: Aspermic samples: Sibille I, et al. 2002 Forensic Sci Int. 2002 Feb 18;125(2-3):212-6.
- Missing persons/Paternity: paternal lineage reference samples

Polymorphisms on the Y

- Binary (biallelic) Markers
  - SNPs (single nucleotide polymorphisms)
  - YAP (Y *Alu* polymorphism)
- Microsatellites – STR's
  - Tetranucleotide repeats such as DYS19, DYS385, DYS388, DYS390, DYS391, etc.
- Minisatellites - MSY1

Definitions
What is a SNP?

Single Nucleotide Polymorphisms

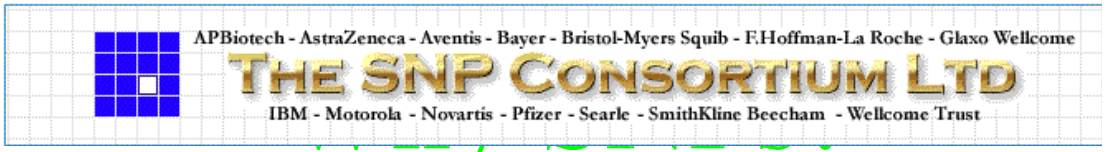
Point mutation GAATCCTCCATCT
GAATCCACCATCT

Deletion GAATCCTCCATCT
GAATCC-CCATCT

Insertion GAATCCT-CCATCT
GAATCCTCCCATCT

Most study Bi-allelic SNPs

GAATCCTCCATCT
GAATCCACCATCT



- Extremely Well Studied- Used in virtually every molecular field
- Huge menu: The SNP Consortium (http://snp.cshl.org/)
- The menu of Y SNPs includes over 4193 available Y SNPs (Nature 2001. 409:928)
- Contrast to under 100 available Y STRs (http://www.cstl.nist.gov/biotech/strbbase)

- Multiplexing capability
- “Easy” to score- on/off and Automate

Le SNP Menu

A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms

Nature 2001. 409:928

The International SNP Map Working Group*

* *A full list of authors appears at the end of this paper.*

We describe a map of 1.42 million single nucleotide polymorphisms (SNPs) distributed throughout the hu an average density on available sequence of one SNP every 1.9 kilobases. These SNPs were primarily disc The SNP Consortium and the analysis of clone overlaps by the International Human Genome Sequencing integrates all publicly available SNPs with described genes and other genomic features. We estimate that exon (coding and untranslated regions), and 85% of exons are within 5 kb of the nearest SNP. Nucleotide across the genome, in a manner broadly consistent with a standard population genetic model of human hi SNP map provides a public resource for defining haplotype variation across the genome, and should help important genes for diagnosis and therapy.

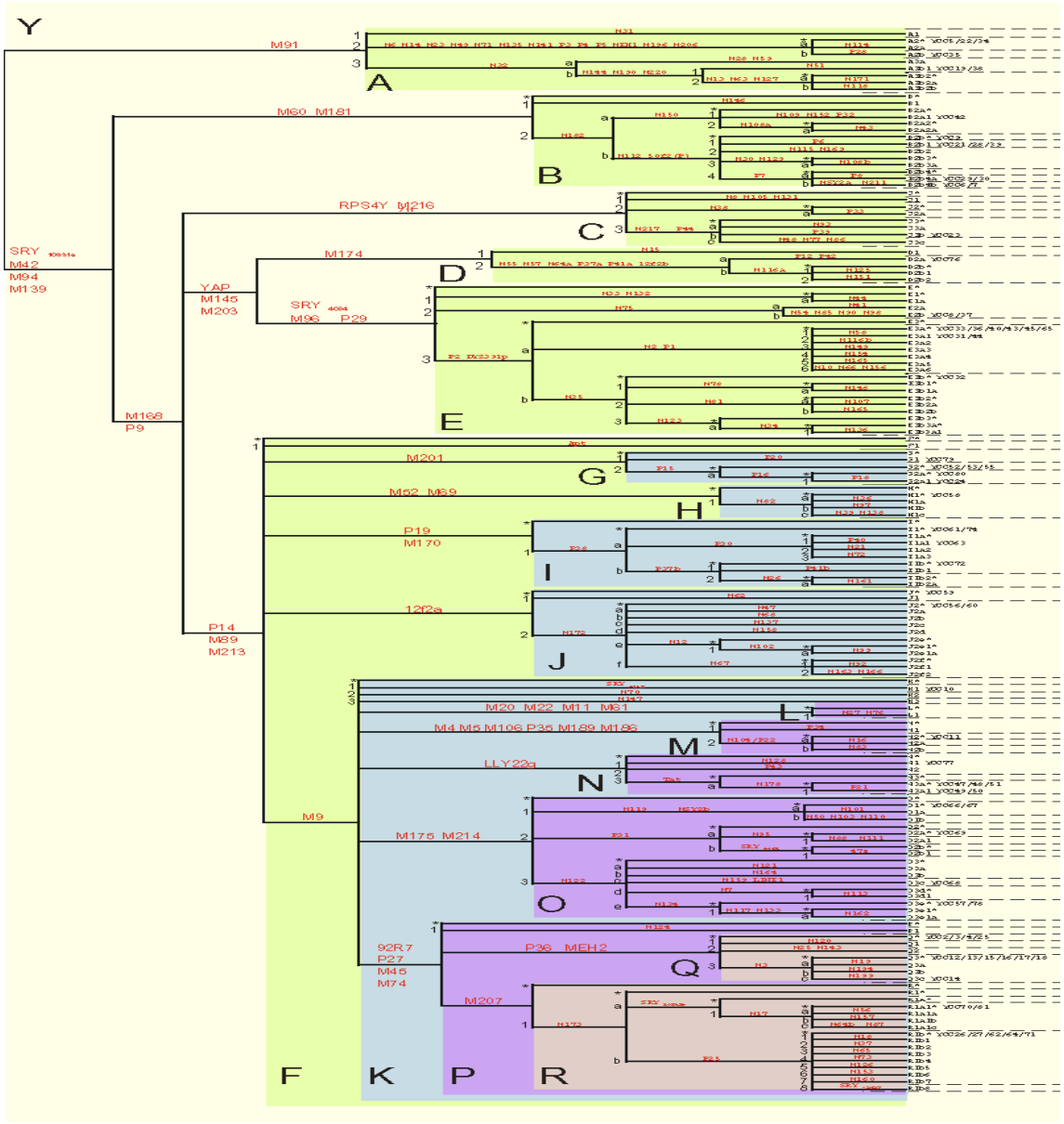
Table 1 SNP distribution by chromosome

Chromosome	Length (bp)	All SNPs		TSC SNPs	
		SNPs	kb per SNP	SNPs	kb per SNP
1	214,066,000	129,931	1.65	75,166	2.85
2	222,889,000	103,664	2.15	76,985	2.90
3	186,938,000	93,140	2.01	63,669	2.94
4	169,035,000	84,426	2.00	65,719	2.57
5	170,954,000	117,882	1.45	63,545	2.69
6	165,022,000	96,317	1.71	53,797	3.07
7	149,414,000	71,752	2.08	42,327	3.53
8	125,148,000	57,834	2.16	42,653	2.93
9	107,440,000	62,013	1.73	43,020	2.50
10	127,894,000	61,298	2.09	42,466	3.01
11	129,193,000	84,663	1.53	47,621	2.71
12	125,198,000	59,245	2.11	38,136	3.28
13	93,711,000	53,093	1.77	35,745	2.62
14	89,344,000	44,112	2.03	29,746	3.00
15	73,467,000	37,814	1.94	26,524	2.77
16	74,037,000	38,735	1.91	23,328	3.17
17	73,367,000	34,621	2.12	19,396	3.78
18	73,078,000	45,135	1.62	27,028	2.70
19	56,044,000	25,676	2.18	11,185	5.01
20	63,317,000	29,478	2.15	17,051	3.71
21	33,824,000	20,916	1.62	9,103	3.72
22	33,786,000	28,410	1.19	11,056	3.06
X	131,245,000	34,842	3.77	20,400	6.43
Y	21,753,000	4,193	5.19	1,784	12.19
RefSeq	15,696,674	14,534	1.08		
Totals	2,710,164,000	1,419,190	1.91	887,450	3.05

Length (bp) is from the public Genome Assembly of 5 September 2000. Density of SNPs on each chromosome is influenced by the amount of available genome sequence included in the Genome Assembly, depth of overlap coverage from TSC reads and clone overlaps, and the underlying heterozygosity (Table 2). Data are presented for the entire dataset (All SNPs) and for those from the SNP consortium (TSC SNPs), as the latter are more evenly spaced than those from clone overlaps.

The Y Chromosome Consortium Map

Genome Research (2002) 12: 339-348



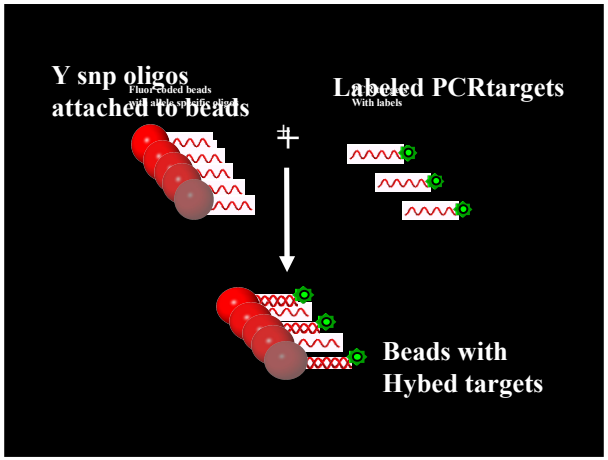
Other Applications of SNPs (aside forensics)

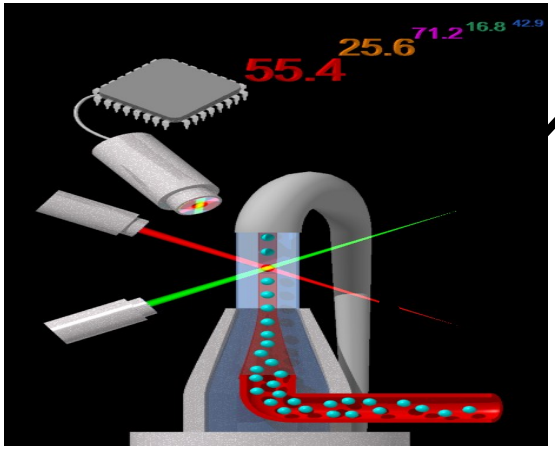
- **Medical Diagnostics**
  - **Tissue typing- HLA DQ alpha typing**
  - **Cystic Fibrosis**
  - **Inflammatory panels**
  - **Neuro-psychiatric illnesses**
  - **Cancers**
  - **Chronic degenerative diseases**
- **Pharmacogenomics - Predictive Pharmacology**
  - **Association of genotype to drug response**
  - **Genetic population studies of patients and their responses to treatment**
  - **Personalized Medicine**
- **Genetic Linkage studies- SNP Haplotyping**

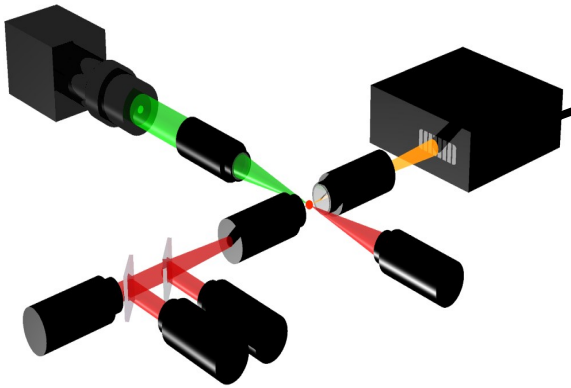
- **Detection Technologies for Y SNPs in Forensics**

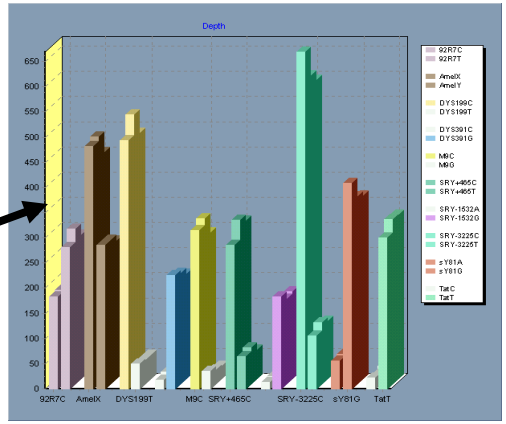
  - **Primer Extension- SNapShot- aka minisequencing. Dugan et al. 2003**
  - **Pyrosequencing- Ballantyne, J. 2003 AAFS**
  - **Light Cycling- Roche - Lareu M, et al. 2001 The use of the LightCycler for the detection of Y chromosome SNPs.Forensic Sci Int. 2001 May 15;118(2-3):163-8.**
  - **Quadruopole MS- Eckenrode et al. 2003 AAFS**
  - **Bead based assays-Luminex, Marligen Biosciences. Carlson et al 2002**

SNPs on the Luminex

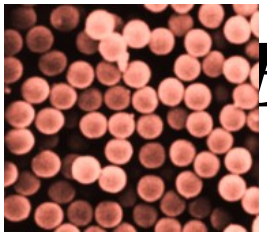




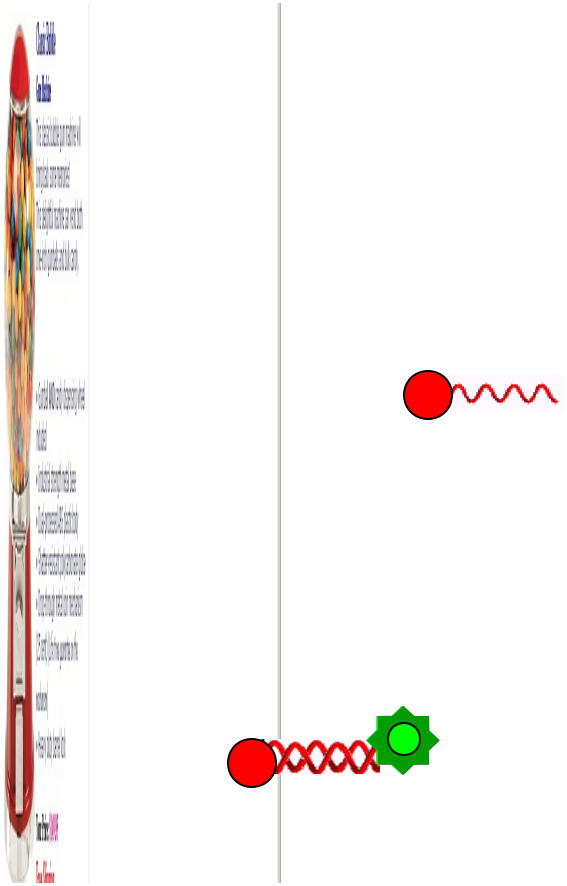




Butler, J. et al. 2003



Bead Based Assay- Luminex 100
DNA Gumballs



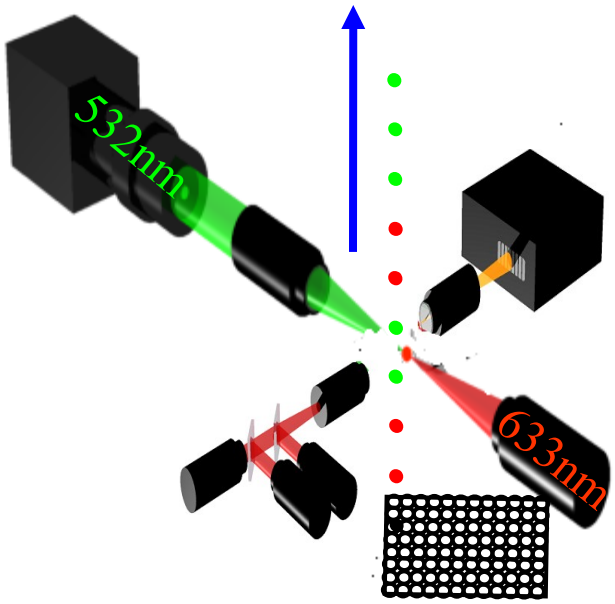
- **Internal Spectral Address™ R:IR ratio (gumball color) identifies each of the assays (probe?)**

- **Reporter fluorescence on the surface (target?) is quantified**

Recap of technology

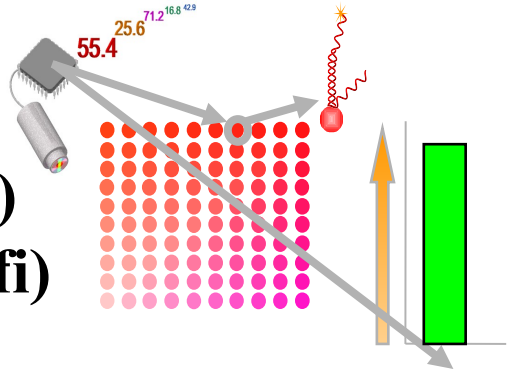
1- Beads flow single file past two lasers

- **633nm excites 2 dyes in beads**
- **532nm excites dye on target if there**



2- Detectors capture:

- **R:IR ratio →** **SNP probe ?**
- **Scatter →** **Single bead?**
- **Reporter fluor →** **SNP target ?**

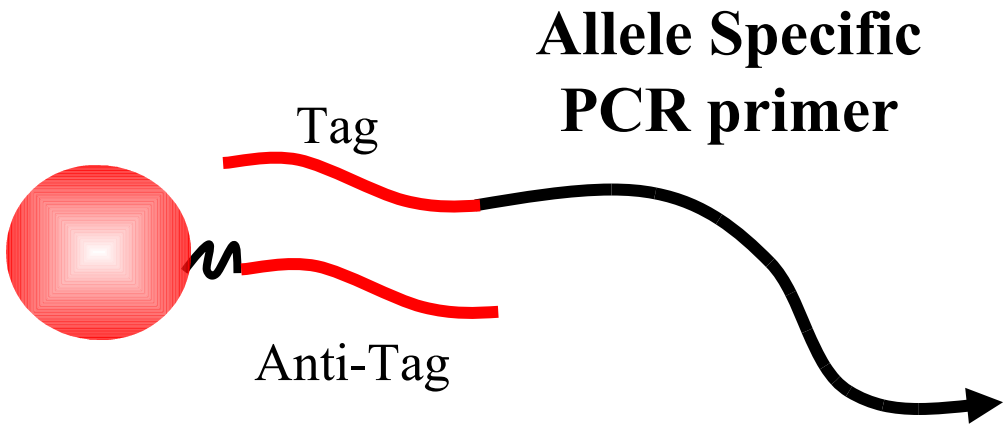


3- Digital signal processor:

- **Collects, processes and saves the data (csv)**
- **Records median fluorescence intensity (mfi)**

Universal Arrays: Tm Bioscience

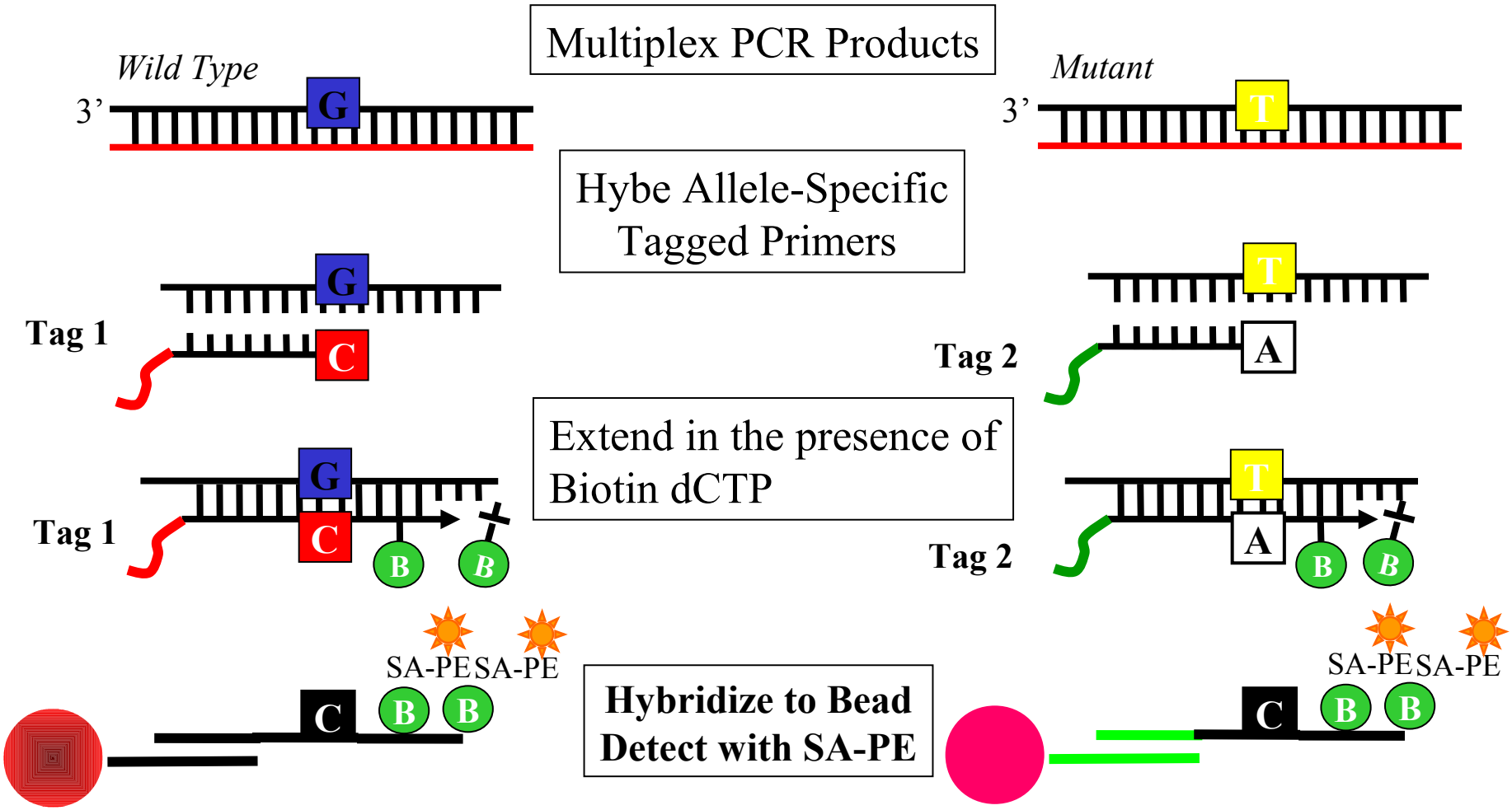
- **100 unique tags**
- **No C: : 75% A/T + 25% G**
- **24-mers (six 4-bp motifs)**
- **Isothermal (± 2˚C)**
- **Minimal cross-talk**



Universal arrays and Primer Extension – SNP Detection Format Alternatives

- Reproducibility: Pre-coupled capture probes eliminates any conjugation variability
- Flexibility: Bead-capture/probe sets for any loci
- Specificity: Primer-extension enhanced
- Multiplexing: 100 capture probes are isothermal (Tm $\pm$ 2^{o}C) for Tm Bioscience beads

Primer Extension (allele specific), Tm Universal arrays and Luminex 100



Luminex SNP Applications, Kits or Publications

• Bacterial ID	Ye et al. 2001 Hum Mut. 17:305
• Biodefense	Los Alamos National Laboratory
• Conservation Genetics	UC Davis- BML
• Cystic Fibrosis Testing	Dunbar et al. 2000. Clin Chem 46: 1498
• Forensics	Carlson et al. 2002 ISHI
• Environmental Microb.	Spiro et. Al. 2000. AEMicrobiol. 66:4258
• Plant Gene Expression	Yang et al. 2001 Genome Res. 11: 1888
• Haplotyping	www.polygenyx.com
• HLA Testing	www.onelambda.com
• Human Identity Testing	www.marligen.com
• Oncology	www.mutlimetrix.com
• Paternity testing	www.luminexcorp.com
• Trichosporon spp	University of Miami/www.miraibio.com
• Thrombophilia	www.luminexcorp.com
• Universal Arrays	Tm Bioscience
• Virology Assays	Smith et al. 1998. Clin Chem 44:2054

Applications: 1- Bacterial ID

HUMAN MUTATION 17:305–316 (2001)

METHODS

Fluorescent Microsphere-Based Readout Technology for Multiplexed Human Single Nucleotide Polymorphism Analysis and Bacterial Identification

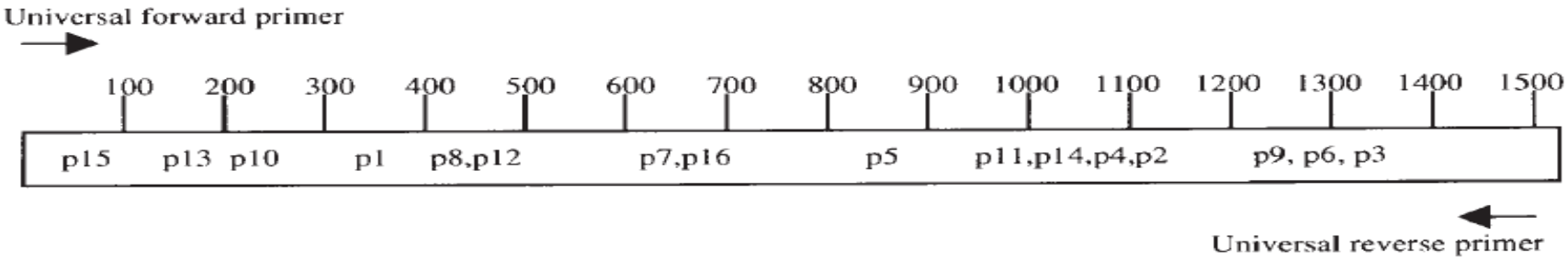
Fei Ye,[1*] May-Sung Li,[1] J. David Taylor,[1] Quan Nguyen,[2] Heidi M. Colton,[3] Warren M. Casey,[3] Michael Wagner,[2] Michael P. Weiner,[1] and Jingwen Chen[1]

[1]*Department of Genomic Sciences, Glaxo Wellcome Research and Development, Research Triangle Park, North Carolina*
[2]*Department of Human Genetics, Glaxo Wellcome Research and Development, Research Triangle Park, North Carolina*
[3]*Department of Analytical Sciences, Glaxo Wellcome Research and Development, Research Triangle Park, North Carolina*

Ye et al. 2001. Human Mutation 17:305
Ahmadian A and J Lundeberg. 2002. A brief History of Genetic Variation Analysis. Biotechniques. 32:1122-1137.
Entire Issues dedicated to SNP technology and Applications

Bacterial Identification using 16S rDNA SNPs

Ye et al. 2001 Human Mutation.17:305-316



B

Bacteria strains — **Identification patterns generated by primer extension assays**

	Probe 1	Probe 2	Probe 3	Probe 4	Probe 5	Probe 6	Probe 7	Probe 8	Probe 9	Probe 10	Probe 11	Probe 12	Probe 13	Probe 14	Probe 15	Probe 16
Bacillus circulans	A	C	C	C	A	G	G	#	#	#	#	#	#	#	#	#
Bacillus cereus	A	C	C	C	A	G	G	#	#	#	#	#	C	#	#	#
Bacillus stearothermophilus	A	C	C	C	A	#	G	#	#	#	#	#	#	#	#	#
Bacillus subtilis	A	C	C	C	G	#	G	#	#	#	#	#	#	#	#	#
Listeria monocytogenes	A	C	T	C	G	G	A	#	T	#	#	#	#	#	#	#
Staphylococcus xylosus	A	C	T	C	G	T	G	#	#	#	#	G	#	#	#	#
Staphylococcus epidermidis	A	C	T	C	G	T	A	#	#	C	#	A	#	#	#	#
Staphylococcus aureus	A	C	T	C	G	T	A	#	#	T	#	#	#	#	#	#
Escherichia coli	G	C	T	T	#	#	#	G	#	#	#	#	#	#	G	C
Shigella dysenteriae	G	C	T	T	#	#	#	G	#	#	#	#	#	#	#	C
Klebsiella pneumoniae	G	C	T	T	#	#	#	C	#	#	#	#	#	#	#	#
Salmonella enteritidis	G	C	T	T	#	#	#	T	#	#	#	#	#	#	#	T
Enterobacter cloacae	G	C	T	T	#	#	#	T	#	#	#	#	#	#	#	#
Ralstonia picketti	G	C	C	C	#	#	#	#	T	#	#	#	#	#	#	#
Burkholderiones cepacia	G	C	C	C	#	#	#	#	G	#	#	#	#	G	#	#
Pseudomonas aeruginosa	G	T	T	C	#	#	#	#	T	#	A	#	#	#	#	#
Pseudomonas fluorescens	G	T	C	C	#	#	#	#	T	#	#	#	#	#	#	#

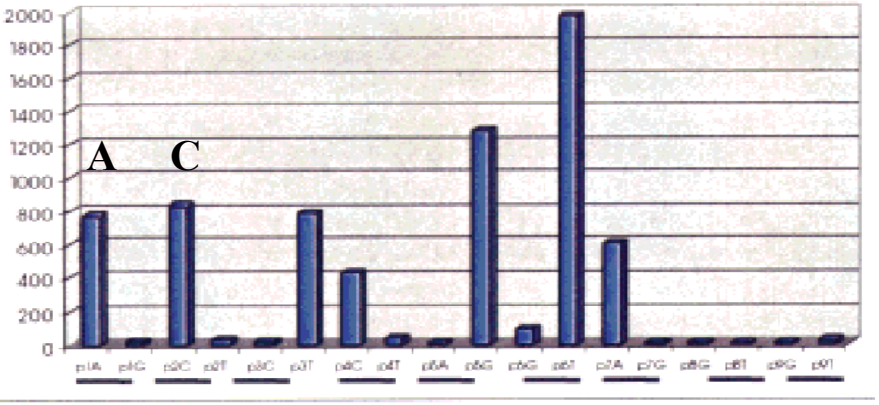
FIGURE 4. Design of the capture probes for multiplex ASPE or SBCE assays using 16S rDNA. **A:** Physical locations of the 16S rDNA probes. Based on the multi-alignment of different bacterial 16S rDNA sequences, 16 conserved regions were chosen for SBCE and ASPE assays. For SBCE reactions, probes are designed such that the 3′ end of the primers terminates one base 5′ to the variable site. For the ASPE assay, a pair of probes was designed such that the 3′ end differs from each other at the variable site. The locations of the probes are listed in Table 1. **B:** Polymorphic patterns in multiplexed SBCE and ASPE assays. Bacterial species can be divided into 17 groups based on their unique readout patterns.

ASPE vs SBCE on 16S rDNA SNPs

ASPE

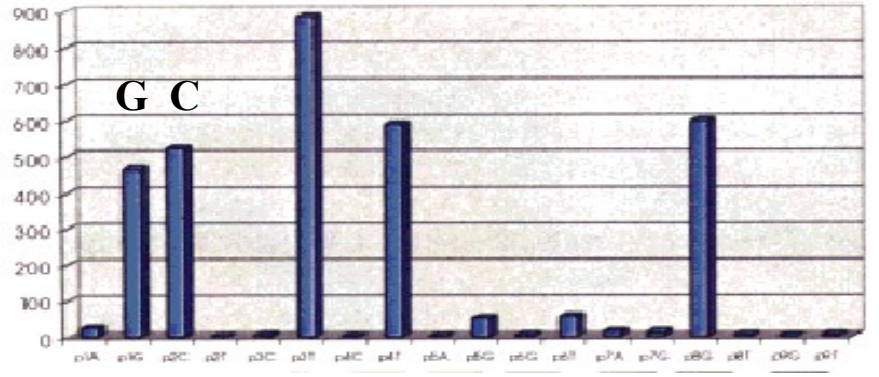
Staphylococcus aureus (ATCC 25923)

Pattern: p1-A/p2-C/p3-T/p4-C/p5-G/p6-T/p7-A/p8-/p9-

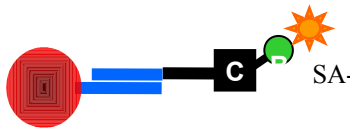


Escherichia coli (ATCC 26)

Pattern: p1-G/p2-C/p3-T/p4-T/p5-/p6-/p7-/p8-G/p9-

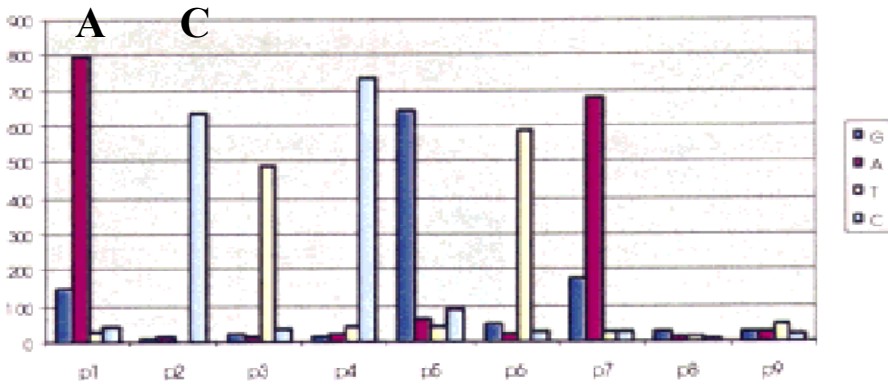


SBCE



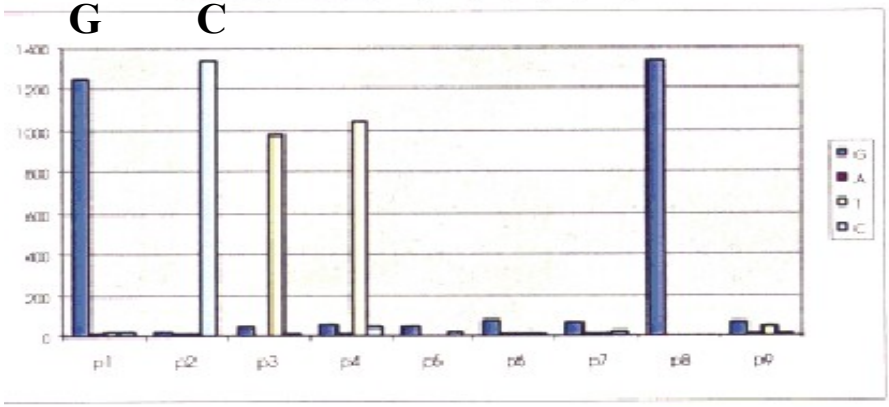
Staphylococcus aureus (ATCC 25923)

Pattern: p1-A/p2-C/p3-T/p4-C/p5-G/p6-T/p7-A/p8-/p9-



Escherichia coli (ATCC 26)

Pattern: p1-G/p2-C/p3-T/p4-T/p5-/p6-/p7-/p8-G/p9-



Ye et al. 2001 Hum. Mut.17:305-316

SNPs vs STRs

	Advantages	Disadvantages
STRs	Higher PD Multiplexes available Databases established Familiar instrumentation	Limited abundance Stutter Extremely degraded samples
SNPs	Abundance SNP Consortium High-throughput automation Highly degraded samples	Lower PD-need 50-100 Mixture limited New instrumentation Databases not established for forensics Validated kits missing

SNPs and STRs

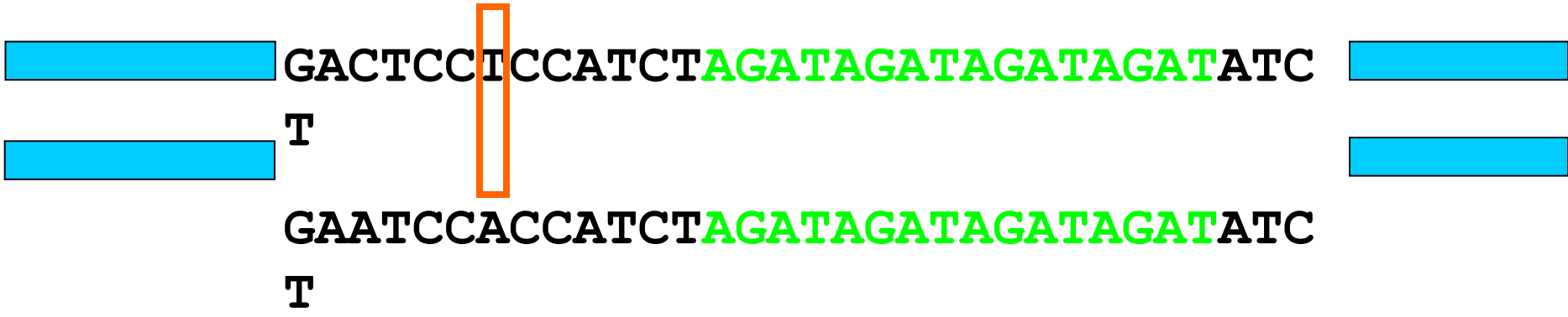
- No STR results- eg 911 samples benefit from SNP typing
- SNPs utilized as a rapid screen – Used to exclude.

- SNPs as additional markers when STRs don't provide sufficient discrimination (mass disasters where total families are lost- relatives with high numbers of shared alleles)
- SNPSTRs?

SNPs and STRs

SNPSTRs: "Each such segment includes one or more single nucleotide polymorphisms (SNPs) and exactly one short tandem repeat (STR) locus"

Genome Res. 2002 Nov;12(11):1766-72. SNPSTRs: empirically derived, rapidly typed, autosomal haplotypes for inference of population history and mutational processes. Mountain JL, Knight A, Jobin M, Gignoux C, Miller A, Lin AA, Underhill PA. Department of Anthropological Sciences, Stanford, California 94305, USA.mountain@stanford.edu



Summary

- **MtDNA** –Well studied, HV regions - degraded DNA Maternal lineage reference samples missing persons databases- Dideoxysequencing detection
- **Y chromosome markers -** 98% of violent crime by males, useful on mixtures and sexual assault evidence, aspermic individuals and missing persons
- **Why Single Nucleotide Polymorphisms (SNPS)**
  - Well studied, Huge selection, multiplexed and automated
  - Primer Extension, Pyrosequencing, Light Cycling, Mass Spec, Bead based assays-Luminex
- **SNPs vs STRs or SNPs and STRs**
  - Either/Or Why SNPs?