Advanced Issues - Forensic DNA

Jack Ballantyne, PhD
National Center for Forensic Science
University of Central Florida

Forensic Science Training for Capital Defense Attorneys
A Wee Philosophical Preamble

- Nature of Science
- Forensic Science
- Theory of DNA Testing
- Bayes Theorem
- Bias
- CSI
- NRC Report 2009
Can science prove anything with certainty?

- Nothing is certain with certainty
- Impossible to prove cause and effect...only correlations (Hume)
  
  Feynman

- *The test of all knowledge is experiment.* Experiment is the *sole judge* of scientific “truth”
- It is *fundamentally impossible* to make a precise prediction of *exactly what will happen* in a given experiment......the same thing does not happen, that we can find only an average, statistically, as to what happens
Forensic science

• The solution of certain problems that arise in connection with the administration of justice. It is science acting in the service of the law.
  – Helps establish facts that may be at issue.
• Identification and comparison science
• Can it establish guilt or innocence?
Biological Evidence Transfer

scene

perpetrator

victim
DNA, for example, frankly doesn’t care if you did it. DNA doesn’t care if you didn’t do it. DNA knows exactly who you, your parents, and your children are, but has no opinion about it and no interest in being a friend or getting your votes. DNA knows it was you who left seminal fluid in someone, but is neither judgmental nor voyeuristic about how or why that deposit might have occurred. So I am far more inclined to trust DNA than the defendant on the witness stand, and it is a shame that DNA is too busy working crimes and pedigree disputes to reconstruct the history of the United States. If DNA had the time, I suspect we would find that most of what we presently believe about the past is tainted, perhaps shockingly so.

Patricia Cornwall, 1991 (Isle of Dogs)
99.9% of the DNA sequence is the same, and therefore 0.1% differs between individuals

#Variable sites = 3 million
#Genotypes = $2^{3\text{million}} = \text{infinity}$ (c.f. $2^{33} = 8$ billion)

Therefore, the DNA of no two individuals is the same (with the exception of identical twins?)

Nuclear DNA = $3 \times 10^9$ bp
3200 Mb
"Yes I can give you liposuction and a face-lift, Mrs Gribbs, but I must warn you that it won’t change the fact that basically way down deep and underneath it all you’re really a duck."

we’re stuck with our genes!
Bayes Theorem

Our belief in a hypothesis after seeing data is proportional to how well that hypothesis explains the data times our initial belief.

All hypotheses must be considered. Need computers to do this properly.

\[
\frac{\Pr(H_p | \text{Evidence})}{\Pr(H_d | \text{Evidence})} = \frac{\Pr(\text{Evidence} | H_p)}{\Pr(\text{Evidence} | H_d)} \times \frac{\Pr(H_p)}{\Pr(H_d)}
\]

Posterior Odds

Likelihood Ratio

Prior Odds

Find the probability of causes by examining effects.
Subjectivity and Contextual Bias and Error
Rapid DNA?
STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES

A PATH FORWARD

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

2009
Overview of the Rest of the Talk

- Y-STRs
  - Sexual assault investigations
- Body fluid ID
- Additional topics
  - Low template DNA analysis (LCN/LTDNA)
  - Bias
  - Mixtures
  - Time since deposition (TSD)
Y-Chromosome STRs
Classic View of Y-Chromosome

- TDF master gene
- Patrilineal inheritance
- No recombination in NRY
- Recombination in PAR
- Junk-rich, gene poor
Y Chromosome map

- Testis Determining Factor (TDF)
- Gadgetry (MAC-locus)
- Channel Flipping (FLP)

P

- Catching and Throwing (BLZ-1)
- Self Confidence (BLZ-2) (note unlinked with ability)

11.1

- Ability to remember and tell jokes (GOT-1)
- Sports Page (BUD-E)

11.21

- Addiction to death and destruction movies (T-2)

Q

- Ability to identify aircraft (DC10)
- Preadolescent fascination with Arachnida and Reptilia (MOM-4U)
- Spitting (P2E)
- Sitting on the John reading (SIT)

11.23

- Inability to express affection (ME-2)

12

- Selective Hearing Loss (HUH)
- Lack of Recall for dates (OOPS)

---

-Jane Gitschier, UCSF  Science, 261, 679 (Aug. 93)
Mutation rates:
Y Filer: $\sim 3 \times 10^{-3}$
RM: $\sim 2 \times 10^{-2}$
Reasons Y?

• Males
  – 80% of all violent crime
  – 95% of all sex offenses
  – (Criminal Victimization in United States, BJS 2001)

• When trying to determine the genetic profile of the male donor in a male/female DNA admixture (when F/M > 20, often >1000) and autosomal STR analysis fails (is not informative) or not possible

• Determination of number of semen donors

• Missing persons (MP)
  – haplotype of MP determined by typing male relative
    • son, brother, father, uncle, nephew
Reasons Y?

- Additional statistical discrimination
  - mixture/partial profiles
- Familial Searching
  - Reduce number of potential male relatives obtained by low stringency match of sample profile to offender database
- RM Y-STRs may differentiate father/son/grandfather etc
Kayser et al.: ‘minimal haplotype’

White et al.: novel YSTRs

UCF: begins Y-chromosome project

Ayub et al.: novel YSTRs

Prinz et al.: forensic validation (pentaplex)

Y-PLEX-5™ released (Reliagene)

Y-PLEX-12™ released (Reliagene)

YFiler™ released (Applied Biosystems)

SWGDAM core

1997

1998

1999

2000

2001

2002

2003

2004

PowerPlex Y® released (Promega)

Kayser: novel YSTRs

Y-STR Development Timeline
“Minimal Haplotype” Loci

(Kayser et. al. 1997):

- DYS 19    DYS 389I   DYS 389II   DYS 390,
  DYS 391    DYS 392    DYS 393    DYS 385 I/II

  - First used in Europe
  - Formed the basis for multiplex development in the U.S.
SWGDAM Core Loci

- Recommended in 2003

- DYS19  DYS389I  DYS389II  DYS390,  DYS391  DYS392  DYS393  DYS385 I/II

- Also: DYS438, DYS439
Commercially Available Y-STR Multiplex Kits

• **Reliagene**: Y-PLEX-5, -6, and -12
  – No longer available

• **ABI**: Yfiler™
  – 17 loci in a single reaction

• **Promega**: PowerPlex Y™
  – 12 loci in a single reaction
Novel Markers?
Novel Y-STR Markers

- **Promega: PowerPlex Y23™**
  - 23 loci in a single reaction
  - Not yet available
- RM loci
- Work done at NCFS
PowerPlex® Y23

- 17 Y-STR loci currently contained in commercially available Y-STR kit (Yfiler™)
- Six new highly discriminating Y-STR loci:
  - DYS481, DYS533, DYS549, DYS570, DYS576, DYS643
- Adds more evidential power to a forensic Y-STR analysis
RM loci

• With current Y-STRs, often a failure to distinguish between males of the same lineage
• Possible to utilize highly mutating Y-STR loci
  – Differences between males in same lineage
• 13 Y-STRs with the highest mutability recommended
  – “RM Y-STR” loci
• Likely to have significant impact on Y-STR analysis
  – Additional discriminatory power to currently used loci
  – In some cases differentiate male lineage relatives
Example (RM Loci)

Father

Son (Twin 1)

Son (Twin 2)
<table>
<thead>
<tr>
<th>Yfiler®</th>
<th>PowerPlex® Y</th>
<th>PowerPlex® Y</th>
<th>RM Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS19</td>
<td>DYS19</td>
<td>DYS19</td>
<td></td>
</tr>
<tr>
<td>DYS385a/b</td>
<td>DYS385a/b</td>
<td>DYS385a/b</td>
<td></td>
</tr>
<tr>
<td>DYS389I</td>
<td>DYS389I</td>
<td>DYS389I</td>
<td></td>
</tr>
<tr>
<td>DYS389II</td>
<td>DYS389II</td>
<td>DYS389II</td>
<td></td>
</tr>
<tr>
<td>DYS390</td>
<td>DYS390</td>
<td>DYS390</td>
<td></td>
</tr>
<tr>
<td>DYS391</td>
<td>DYS391</td>
<td>DYS391</td>
<td></td>
</tr>
<tr>
<td>DYS392</td>
<td>DYS392</td>
<td>DYS392</td>
<td></td>
</tr>
<tr>
<td>DYS393</td>
<td>DYS393</td>
<td>DYS393</td>
<td></td>
</tr>
<tr>
<td>DYS437</td>
<td>DYS437</td>
<td>DYS437</td>
<td></td>
</tr>
<tr>
<td>DYS438</td>
<td>DYS438</td>
<td>DYS438</td>
<td></td>
</tr>
<tr>
<td>DYS439</td>
<td>DYS439</td>
<td>DYS439</td>
<td></td>
</tr>
<tr>
<td>DYS448</td>
<td>DYS448</td>
<td>DYS448</td>
<td></td>
</tr>
<tr>
<td>DYS456</td>
<td>DYS456</td>
<td>DYS456</td>
<td></td>
</tr>
<tr>
<td>DYS458</td>
<td>DYS458</td>
<td>DYS458</td>
<td></td>
</tr>
<tr>
<td>DYS635</td>
<td>DYS635</td>
<td>DYS635</td>
<td></td>
</tr>
<tr>
<td>Y-GATA H4</td>
<td></td>
<td>Y-GATA H4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS481</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS533</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS549</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS570</td>
<td>DYS570</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS576</td>
<td>DYS576</td>
<td></td>
</tr>
<tr>
<td></td>
<td>DYS643</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF387S1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF399S1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS518</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS526a/b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS626</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS627</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF403S1a/b</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYF404S1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS449</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS547</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DYS612</td>
</tr>
</tbody>
</table>
## NCFS Developed Multiplexes

<table>
<thead>
<tr>
<th>Multiplex</th>
<th>Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>9</td>
</tr>
<tr>
<td>II</td>
<td>10</td>
</tr>
<tr>
<td>III</td>
<td>8, 1 InDel</td>
</tr>
<tr>
<td>IV</td>
<td>19</td>
</tr>
<tr>
<td>V</td>
<td>13</td>
</tr>
<tr>
<td>VI</td>
<td>14</td>
</tr>
<tr>
<td>VII</td>
<td>12</td>
</tr>
<tr>
<td>VIII</td>
<td>11</td>
</tr>
<tr>
<td>IX</td>
<td>7</td>
</tr>
<tr>
<td>X</td>
<td>6</td>
</tr>
</tbody>
</table>

= 109 Markers

*Additional 25 Loci screened and rejected → 34% of known loci evaluated
<table>
<thead>
<tr>
<th>Locus</th>
<th>Locus</th>
<th>Locus</th>
<th>Locus</th>
<th>Locus</th>
<th>Locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS19</td>
<td>DYS388</td>
<td>DYS463</td>
<td>DYS522</td>
<td>DYS588</td>
<td>DYS688</td>
</tr>
<tr>
<td>DYS385</td>
<td>DYS425</td>
<td>DYS464</td>
<td>DYS525</td>
<td>DYS590</td>
<td>DYS698</td>
</tr>
<tr>
<td>DYS389I/II</td>
<td>DYS426</td>
<td>DYS468</td>
<td>DYS527</td>
<td>DYS591</td>
<td>DYS703</td>
</tr>
<tr>
<td>DYS390</td>
<td>DYS434</td>
<td>DYS476</td>
<td>DYS531</td>
<td>DYS593</td>
<td>DYS707</td>
</tr>
<tr>
<td>DYS391</td>
<td>DYS435</td>
<td>DYS481</td>
<td>DYS533</td>
<td>DYS594</td>
<td>Y-GATA-A7.1</td>
</tr>
<tr>
<td>DYS392</td>
<td>DYS436</td>
<td>DYS484</td>
<td>DYS534</td>
<td>DYS596</td>
<td>Y-GATA-A7.2</td>
</tr>
<tr>
<td>DYS393</td>
<td>DYS441</td>
<td>DYS485</td>
<td>DYS535</td>
<td>DYS598</td>
<td>Y-GATA-A10</td>
</tr>
<tr>
<td>DYS394</td>
<td>DYS442</td>
<td>DYS488</td>
<td>DYS540</td>
<td>DYS607</td>
<td>YAP</td>
</tr>
<tr>
<td>DYS395</td>
<td>DYS443</td>
<td>DYS490</td>
<td>DYS543</td>
<td>DYS617</td>
<td></td>
</tr>
<tr>
<td>DYS396</td>
<td>DYS444</td>
<td>DYS494</td>
<td>DYS549</td>
<td>DYS622</td>
<td></td>
</tr>
<tr>
<td>DYS397</td>
<td>DYS445</td>
<td>DYS495</td>
<td>DYS552</td>
<td>DYS627</td>
<td></td>
</tr>
<tr>
<td>DYS398</td>
<td>DYS446</td>
<td>DYS497</td>
<td>DYS556</td>
<td>DYS630</td>
<td></td>
</tr>
<tr>
<td>DYS399</td>
<td>DYS447</td>
<td>DYS499</td>
<td>DYS557</td>
<td>DYS631</td>
<td></td>
</tr>
<tr>
<td>DYS400</td>
<td>DYS449</td>
<td>DYS503</td>
<td>DYS561</td>
<td>DYS634</td>
<td></td>
</tr>
<tr>
<td>DYS401</td>
<td>DYS450</td>
<td>DYS505</td>
<td>DYS562</td>
<td>DYS636</td>
<td></td>
</tr>
<tr>
<td>DYS402</td>
<td>DYS452</td>
<td>DYS508</td>
<td>DYS570</td>
<td>DYS638</td>
<td></td>
</tr>
<tr>
<td>DYS403</td>
<td>DYS453</td>
<td>DYS510</td>
<td>DYS572</td>
<td>DYS641</td>
<td></td>
</tr>
<tr>
<td>DYS404</td>
<td>DYS454</td>
<td>DYS511</td>
<td>DYS575</td>
<td>DYS643</td>
<td></td>
</tr>
<tr>
<td>DYS405</td>
<td>DYS455</td>
<td>DYS513</td>
<td>DYS576</td>
<td>DYS637</td>
<td></td>
</tr>
<tr>
<td>Y-GATA-C4</td>
<td>DYS459</td>
<td>DYS520</td>
<td>DYS578</td>
<td>DYS660</td>
<td></td>
</tr>
<tr>
<td>Y-GATA-H4</td>
<td>DYS462</td>
<td>DYS521</td>
<td>DYS587</td>
<td>DYS685</td>
<td></td>
</tr>
</tbody>
</table>
Sexual Assault Investigations
SEXUAL ASSAULT INVESTIGATION

• Typical case involves detection of semen stains on victim’s vaginal samples taken from the victim shortly after the incident
• As a result of post-coital drainage seminal stains are often found on victim’s underpants and such stains can be a good source of seminal material
SEXUAL ASSAULT INVESTIGATION

- Assume that a woman (victim, V) is raped by an individual (perpetrator, P) but shortly before this she has consensual intercourse with her husband/boyfriend (B)

- Before consideration of genetic marker results:
  - semen may only have come from P
  - semen may only have come from B
  - semen may comprise a mixture of semen from P and B
  - semen may not be present

- Absence of semen from P does not mean that P did not rape V
  - semen from P does not prove he raped V (could be consensual)
SEXUAL ASSAULT INVESTIGATION

• Semen stains on vaginal swab/underpants will be mixed with vaginal epithelia/secretions
  – testing will identify the genetic factors from V
• sperm lost from vaginal vault (normally not present after 72 hours), but seminal stains on underpants are stable (months or years)
  – vaginal swab best represents semen deposited at time of rape
• Genetic profiles of V, P and B may have alleles in common and masking of one component by another is possible
SEXUAL ASSAULT INVESTIGATION

• often no scientifically reliable method of determining no. of semen donors

• interpretation of genetic marker testing data in sexual assault evidence is multi-factorial and requires a case by case consideration of the meaning of the data generated
Problem

- some rape victims provide vaginal samples > 24-36 hours after the incident
- ability to obtain an autosomal STR profile of the semen donor from the living victim diminishes rapidly as the post-coital interval is extended (> 24-48h difficult, >72 h not normally possible)
- from classical forensic serology and reproductive biology studies
  - sperm persist in the post-coital vaginal canal up to 3 days after intercourse
  - from the medical literature sperm may be detectable up to 7 days after intercourse in the cervix
  - sperm are few in number after these extended intervals
Sperm Loss Over Time

1. vaginal lavage
2. vaginal drainage
3. normal intra-cervicovaginal sperm degradative changes
   – sperm become damaged and fragile
4. below analytical sensitivity of the test
5. during the differential extraction process within the laboratory
   – multiple manipulations of the sample
   – few remaining sperm lyse into female fraction
     » kinetics of the PCR process
     » majority component in an admixture will titrate out critical PCR reagents
     » female/male DNA ratio \( \geq 300/1 \)
     » failure to type the male component
Potential Solutions to Problem?

- use Y-STRs
  - theoretically can detect male in female/male admixtures
    - ‘ignores’ female component
  - no differential extraction
    - avoids unnecessary sample loss

- low copy number approach
  - add large quantities of DNA (300-450 ng)
    - to permit adequate sampling of small number of still persisting sperm
  - increased PCR cycle number (34-36 cycles)

- ensure cervicovaginal sampling
  - low to mid-vaginal sampling may not recover sperm after extended post coital interval
Not All Y-STRs Are Made Equal

- Y chromosome retains high level of sequence homology with the X chromosome
- Different Y-STR primers will possess different degrees of homology with the X chromosome
Early Work Conclusions

- it is possible to obtain the genetic profile of the semen donor in postcoital cervicovaginal samples recovered up to 4 d after intercourse
  - use of carefully selected Y-STR markers
    - chosen for their ability to detect the male donor in an overwhelming quantity of background female DNA
  - no differential extraction
  - 300-450 ng of input template DNA
  - increased cycle number (34-35 cycles)
  - cervicovaginal sampling
- such strategies may significantly impact the recovery and processing of rape evidence
Additional Work

- Number of DNA profile enhancement strategies employed
  - Cervical brushing
  - Differential versus non-differential DNA extraction
  - Post-PCR purification

- Standard manufacturers’ cycling conditions used

- Y-STR profiles
  - Full profiles routinely 3-5 days after intercourse
  - Profiles > 6 days after intercourse but mainly partial
  - Use of post PCR purification significantly improved ability to obtain profile, especially from the 5-6 day samples
  - Better profiles with differential lysis (sperm fraction)
  - 8 locus Y-STR profile from 7 day post-coital sample

- Approaching limit for sperm detection in cervix
Enhanced DNA Profiling for Detection of the Male Donor in Trace DNA Samples
Nested-PCR Strategy

Pre-amplification

MinElute purification to remove excess primers
### Nested PCR Pre-Amplification Multiplex

<table>
<thead>
<tr>
<th>Input</th>
<th>Pre-amp?</th>
<th>Avg RFU</th>
<th>Avg # Loci</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>&gt;150</td>
</tr>
<tr>
<td>15pg</td>
<td>with</td>
<td>4088</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>without</td>
<td>105</td>
<td>3</td>
</tr>
<tr>
<td>10pg</td>
<td>with</td>
<td>3727</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>without</td>
<td>91</td>
<td>1</td>
</tr>
<tr>
<td>5pg</td>
<td>with</td>
<td>4272</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>without</td>
<td>89</td>
<td>1</td>
</tr>
</tbody>
</table>

*Improvement in allele recovery and signal intensity with prior pre-amplification*

*Yfiler Amplification kit (ABI)*

*Averages based on data from 10 male individuals*
5pg Input Male DNA (Single Source)

Without Pre-Amp

With Pre-Amp
Without pre-amplification (1/17 alleles)

With pre-amplification (11/17 alleles)
Door Handle

Without pre-amplification (0/17 alleles)

With pre-amplification (17/17 alleles)
7 days after intercourse

Without pre-amplification
(5/17 alleles)

With pre-amplification
(16/17 alleles)
National U.S. Y-STR Database
Y-STR Database Goals

• To compile and consolidate Y-STR data from all available ‘legitimate’ sources
• To create a Y-STR Consortium comprised of stakeholders and data contributors from the forensic community
• Expand data
  – Type additional samples using core loci
• Provide custodial and managerial responsibility
• Develop quality indicators for data inclusion and submission
  – ‘Proficiency testing’ for labs who wish to contribute data
  – Screen data and remove duplicate & related samples
• Ensure allele-call consistency among different primer sets
• Provide accessibility and statistical data to the forensic community via the Internet
The Y-STR Consortium was formed at the 2006 AAFS Meeting in Seattle, WA to assist in sample consolidation and the design and development of the database.

Y-STR Consortium Members

- NCFS
  - Jack Ballantyne
  - Lyn Fatolitis

- Applied Biosystems
  - Lisa Calandro

- FBI
  - Bruce Budowle

- University of Arizona
  - Mike Hammer

- NIST
  - John Butler

- MN Dept of Public Health
  - Ann Marie Gross

- NYC OCME
  - Mecki Prinz

- University of North Texas
  - Arthur Eisenberg

- Promega
  - Curtis Knox

- ReliaGene
  - Sudhir Sinha

- Orchid Cellmark
  - Cassie Johnson

- NIJ
  - John Paul Jones
5. Statistical Interpretation

5.1. Y-STR loci are located on the nonrecombining part of the Y-chromosome and, therefore, should be considered linked as a single locus. A Y-STR database must consist of haplotype frequencies rather than allele frequencies. The source of the population database(s) used should be documented. Relevant population(s) for which the frequency will be estimated should be identified. A consolidated U.S. Y-STR database (http://usystrdatabase.org) has been established and should be used for population frequency estimation. A number of other Y-STR haplotype frequency databases exist online. (See available listing on the NIST [National Institute of Standards and Technology] STRBase Web site at http://www.cstl.nist.gov/biotech/strbase/y_strs.htm.)
Was made available on January 3, 2012
Comprised of 18,719 haplotypes
  – An additional 61 Yfiler haplotypes were uploaded
    • 35 Caucasian
    • 3 African American
    • 7 Asian
    • 16 Hispanic
  – San Diego Sheriff’s Crime Laboratory (39 samples)
  – Santa Clara County Crime Laboratory (22 samples)
Release 2.6 contains 15,395 complete PowerPlex Y (12-locus) haplotypes. Of these, 6194 haplotypes are unique (i.e., seen only once in the database) while 9201 haplotypes are seen more than once, giving a DP of 40.2%.

The Database contains 8548 complete Y-filer (17-locus) haplotypes. Of these, 6934 haplotypes are unique while 1614 haplotypes are seen more than once, giving a DP of 81.1%. 
Database Home Page: www.usystrdatabase.org

### US Y-STR Database

**Release:** 2.6 | **Last Updated:** 01/03/2012

#### Select Alleles

<table>
<thead>
<tr>
<th>Common Markers</th>
<th>DYS19</th>
<th>DYS385</th>
<th>DYS390</th>
<th>DYS391</th>
<th>DYS289I</th>
<th>DYS389II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>*</td>
<td>□</td>
<td>*</td>
<td>□</td>
<td>*</td>
<td>□</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>□</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>□</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>□</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>□</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>□</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>□</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Search By Ancestry

- All
- African American
- Asian
- Caucasian

**Queries Performed:** 86995
Since the release of the US Y-STR Database in January 2008, over 86,200 database search queries have been performed. Up until January of 2011, the database had an average of approximately 800 searches per month. Between January and August 2011 the average use increased to >8,000 searches per month. Since this time, the average is over 1200 searches per month.
Court Support

Frye Hearings

- State of California v Miszkewycz (county of Placer)
  - “..contends that the statistical analysis applied in this case is faulty because of an unreliable or unknown data base used to formulate the statistics”
  - “Court is satisfied that accepted scientific procedures and principals (sic) were properly used in this case” (After “Ms Caser’s” testimony)

- State of Kansas v Gonzalez
  - Judge Pokorny, Seventh Judicial District for the District of Kansas, Douglas County, KS
  - Challenge that over a period of six months the frequency of the evidence/suspect haplotype changed from 1/2717 to 1 in 1786
  - 4 October 2010: found that Y-STR database is fit for purpose and motion to deny/exclude Y-STR haplotype evidence denied

Peer Reviewed Journal Article

CA DOJ Y-STR Mixture Tool

Instructions:

1. Interpret your mixture to determine the alleles of possible contributors.
2. Using the drop-down menus, enter the interpreted alleles into the table on the "Profile" worksheet. 
   Note: If the list of alleles for a locus is interpreted as possibly incomplete (i.e., not representative of all possible contributors), the locus should be left blank.
2a. If an allele in the mixture is not present in the list below the table, enter it as a "New Variant" prior to entering it into the table. 
   A "wild card" entry (e.g., .999) could instead be used when there are multiple mixture alleles not present in a locus' list.
3. Click on the button-macro "Compare the mixture to the database.
   This will filter the database, leaving only those haplotypes that would be included as possible contributors to your mixture.
4. Counts of non-excluded haplotypes (k) and database sizes (N) are summarized in the table.
5. If you wish to view the filtered list of non-excluded haplotypes, click on the button-macro "View the filtered list."

Disclaimer: This is a BETA version of the Y-Mix Database Filter spreadsheet. Prior to its use in criminal and/or civil cases, users agree to either conduct their own validation of this spreadsheet or independently confirm the results on a case-by-case basis.

Steven.Myers@DOJ.ca.gov
Harris County Institute of Forensic Sciences Y-Mixture Tool

To use the template:

- Enable Macros when opening the file
- Enter in all alleles required for the profile to generate statistics
  - For locus DYS385, always enter in two alleles even if there is only one occurrence. For example, if the profile at locus DYS385 is 14, enter in 14 for allele 1 and 14 for allele 2.
  - Alleles must be entered from shortest (lowest) to largest (highest).
- If the analyst is generating statistics for a single source sample, use the single source search button, if the analyst is generating statistics for a mixture sample, use the Mixture Search button.
- A button will pop up to indicate that the search is complete.
- The statistics will be shown at the bottom of the tab/worksheet.

Disclaimer: The author of this software disclaims all warranties and conditions, whether expressed or implied, statutory or otherwise, in connection with use. User assumes full responsibility for quality control, testing, and determination of suitability for intended application or use. Author makes no warranty of fitness for a particular purpose and shall not be liable for any claim of damages resulting from loss of data, use of equipment, or for any special, incidental, indirect or consequential damages arising out of or in connection with the use or performance of this software.

<table>
<thead>
<tr>
<th>Locus</th>
<th>DYS456</th>
<th>DYS385</th>
<th>DYS390</th>
<th>DYS389I</th>
<th>DYS458</th>
<th>DYS19</th>
<th>DYS385</th>
<th>DYS393</th>
<th>DYS391</th>
<th>DYS439</th>
<th>YGATA_C4 or DYS635</th>
<th>DYS392</th>
<th>YGATA_H4</th>
<th>DYS437</th>
<th>DYS438</th>
<th>DYS448</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allele 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Statistics

<table>
<thead>
<tr>
<th></th>
<th>Frequency (95% upper CI)</th>
<th>95% Upper Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>African American</td>
<td>3128 in 6160 profiles</td>
<td>0.520277</td>
</tr>
<tr>
<td>Asian</td>
<td>442 in 960 profiles</td>
<td>0.496582</td>
</tr>
<tr>
<td>Caucasian</td>
<td>3306 in 6763 profiles</td>
<td>0.500750</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1465 in 3343 profiles</td>
<td>0.468049</td>
</tr>
<tr>
<td>Native American</td>
<td>585 in 983 profiles</td>
<td>0.628803</td>
</tr>
<tr>
<td>Total</td>
<td>8926 in 18199 profiles</td>
<td>0.497730</td>
</tr>
</tbody>
</table>

User Instructions: When typing in alleles, enter them in sequential order, entering the first allele in the Allele 1 row, the second allele in the Allele 2 row, etc.

If an allele falls outside of the allelic ladder range for a locus, enter in a < and the smallest recognized allele or a > and the largest recognized allele as appropriate.
Body Fluid Identification
The serology problem

Do we need serology?

YES we do!

But very few people are doing it!
Why the need to identify solid tissues?

• Crime scene
  – Tissue fragments
  – Trajectory of Bullet/Condition of Victim
    • skin, adipose vs. brain, heart
Conventional Methods of Body Fluid and Tissue Identification

- Protein-based
- Analysis performed in series
  - not multiplex

- Conventional histology
- Requires a highly trained histologist and/or pathologist

- Consumption of sample
- Labor-intensive
- Time-consuming
- Technologically diverse
  - Difficult to automate
Development of a novel multiplex analysis procedure for body fluid identification

- New parallel procedure should be compatible with current DNA analysis procedures
- Two possible routes:
  - Both expressed in tissue specific manner
    - **Proteins**
      - Proteomics: further developments necessary for multiplex analysis of complex, partially degraded protein mixtures
    - **RNA**
      - Molecular intermediate between DNA and protein
      - Better option: technologies for parallel analysis available now
The Multicellular Transcriptome - the collection of genes that are expressed within the constellation of differentiated cells that make up a body fluid.

<table>
<thead>
<tr>
<th>Tissue-Specific Genes</th>
<th>BLOOD STAIN</th>
<th>SEMEN STAIN</th>
<th>SALIVA STAIN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housekeeping genes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rare</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderately Abundant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abundant</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Determine type and abundance of expressed genes → identify tissue or body fluid
Advantages of mRNA Based Approach

- Greater specificity
- Simultaneous analysis of multiple markers and body fluids through common assay format
- Improved timeliness
- Decreased sample consumption
- Automation
- Ability to co-extract DNA and RNA
### Body Fluid Specific mRNA Markers

<table>
<thead>
<tr>
<th>Blood</th>
<th>Semen</th>
<th>Saliva</th>
<th>Vaginal</th>
<th>Menstrual</th>
<th>Skin</th>
<th>Housekeeping</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALAS2</td>
<td>PRM1</td>
<td>HTN3</td>
<td>MUC4</td>
<td>MMP7</td>
<td>LCE1C</td>
<td>B2M</td>
</tr>
<tr>
<td>SPTB</td>
<td>PRM2</td>
<td>STATH</td>
<td>HBD1</td>
<td>MMP10</td>
<td>LCE1D</td>
<td>UBC</td>
</tr>
<tr>
<td>PBGD</td>
<td>MSP</td>
<td>PRB4</td>
<td>ESR1</td>
<td>MMP11</td>
<td>LCE2D</td>
<td>UCE</td>
</tr>
<tr>
<td>CD3G</td>
<td>TGM4</td>
<td>SPRR3</td>
<td>CYP2B7P1</td>
<td>CK19</td>
<td>CCL27</td>
<td>GAPDH</td>
</tr>
<tr>
<td>HBB</td>
<td>PSA</td>
<td>SPRR1A</td>
<td>MYOZ1</td>
<td>PR</td>
<td>IL1F7</td>
<td>G6PDH</td>
</tr>
<tr>
<td>CASP2</td>
<td>SEMG1</td>
<td>KRT4</td>
<td>FUT6</td>
<td>LEFTY2</td>
<td>LOR</td>
<td>TEF</td>
</tr>
<tr>
<td>AM1CA1</td>
<td>SEMG2</td>
<td>KRT6A</td>
<td>DKK4</td>
<td>MSX1</td>
<td>CDSN</td>
<td></td>
</tr>
<tr>
<td>C1QR1</td>
<td></td>
<td>KRT13</td>
<td>SFTA2</td>
<td>SFRP4</td>
<td>KRT9</td>
<td></td>
</tr>
<tr>
<td>ALOX5AP</td>
<td></td>
<td></td>
<td>IL19</td>
<td></td>
<td>CST6</td>
<td></td>
</tr>
<tr>
<td>AQP9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>DSC1</td>
<td></td>
</tr>
<tr>
<td>C5R1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCFS2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MNDN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARHGAP26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GYP A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANK1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HBA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Various Multiplex Assays Developed
(no commercially available kits)

Example: NCFS 15-plex
**Vaginal Secretions**

- Lesser encountered body fluid?
  - Sexual assault vaginal swabs
    - Presence is obvious since swabs taken from vaginal canal
    - Identified from DNA profile of victim
- Identification may be crucial to some cases:
  - Assault with foreign object
    - Victim’s DNA on object – skin vs vaginal secretions?
  - Digital penetration
    - Victim’s DNA on suspect’s hands – skin vs vaginal secretions?
Mock Casework Samples

- Digital penetration ("finger" samples)
  - 2 collections
- Assault with object ("object" sample)
- Sexual assault
  - Samples collect from penis and underwear before and after intercourse
• Perceived to be the result of DNA obtained from shed skin cells
  – STR profiles can be obtained
  – Tissue source not typically confirmed
    • Lack of molecular based strategies for skin identification
    • Some claim that identification of tissue source of origin from trace biological material is not possible
    • Failure to identify tissue source could have undue influence on perception of the circumstances of the crime
Expression in Touch/Contact samples

- **Glove, female**
- **Fingerprint - glass**
- **Coffee Pot Handle**
- **Forehead**
- **Candy jar lid**
New LCE1C – touch/contact samples

- 100% detection in sample set (N=12) with the re-designed LCE1C primer set
- Small sample size → need to test additional samples

<table>
<thead>
<tr>
<th>Sample</th>
<th>Singleplex</th>
<th>HD Adv 4plex</th>
<th>LCE1C- re-designed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LCE1C</td>
<td>LCE1D</td>
<td>LCE2D</td>
</tr>
<tr>
<td>Cheek</td>
<td>4412</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arm</td>
<td>392</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leg</td>
<td>1775</td>
<td>61</td>
<td></td>
</tr>
<tr>
<td>Forehead</td>
<td>69</td>
<td>126</td>
<td></td>
</tr>
<tr>
<td>Candy jar lid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pencil</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mouse</td>
<td>2647</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phone - receiver (by ear)</td>
<td>1838</td>
<td>99</td>
<td></td>
</tr>
<tr>
<td>Phone - handle</td>
<td>1959</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Door handle - inside</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Door handle - out</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coffee pot handle</td>
<td>3171</td>
<td>377</td>
<td></td>
</tr>
</tbody>
</table>
Compatibility with DNA analysis pipeline

Co-Extraction

PREPARING FOR CASUAL FRIDAY AT THE GENOME LAB
Everything you look for and all that you perceive has a way of proving whatever you believe.
Subjectivity and bias in forensic DNA mixture interpretation

Itiel E. Dror a,b,⁎, Greg Hampikian c

a Institute of Cognitive Neuroscience, University College London (UCL), London, UK
b Cognitive Consultants International (CCI), London, UK
c Departments of Biology and Criminal Justice, Boise State University, USA

The design of “the first experimental study exploring DNA interpretation”

David H. Kaye a

Penn State University, Dickinson School of Law, Lewis Katz Building, University Park, PA 16802, United States

Expectations, contextual information, and other cognitive influences in forensic laboratories

Dear Editor:

The objections to cognitive bias in forensic science often center around:

1. Examiners are introduced to the facts of a case via the statistical analysis of evidence. This means that the examiner may assess the data in light of specific expectations.
2. Examiners often state that even if contextual influences and bias are indeed part of human cognition, they are nevertheless impartial because they conduct their work without pressure or expectation to decide specific outcomes.

Here Brauner's paper [4] provides much insight. As a forensic examiner he acknowledges that “information from investigators or media news reports may affect a forensic scientist's partiality in result interpretation.” Indeed, a recent survey of 35 forensic examiners [5] showed that 26% of respondents change their assessed probability of guilt after a change in interpretation of the evidence.
Bias and Subjectivity

- Confirmation or contextual bias: seeking and interpreting information in a way that it fits existing beliefs, expectation, hope or motivation
- Objectivity: “seeing” the universe exactly for what it is from a standpoint free from human perception and its influences, human cultural interventions, past experience and expectation of the result.
- Subjectivity: the specific discerning interpretations of any aspect of experiences. They are unique to the person experiencing them
  - subjectivity is the only way we have to experience the world, mathematically, scientifically or otherwise
Adjudicated case

- Gang rape in Georgia: one assailant testified against co-defendants in a plea bargain but other defendants denied any involvement
- State law requires corroboration: if suspects excluded or DNA inconclusive then testimony of admitted rapist would “most likely” not hold
- Took epgs/materials used by original DNA examiners (conclusion: suspect cannot be excluded)
- Presented to 17 qualified DNA examiners ‘context free’-working in an accredited governmental laboratory in North America
  - examined DNA profiles from mixture and from V and S (x3)
  - suspect 3 was point of interest (“cannot be excluded”)
  - had to give one conclusion (“cannot be excluded”, “excluded” or “inconclusive”)
- Results:
  - cannot be excluded = 1
  - excluded = 12
  - Inconclusive = 4
Conclusions

1. 17 examiners not consistent: subjectivity

2. Between the two labs (original “extraneous context” and the 17 member “context free”)
   - “It is possible that the domain irrelevant information may have biased their interpretation”
   - “Thus the extraneous context appears to have influenced the interpretation of the DNA mixture, however, it is always hard to draw scientific conclusions when dealing with methodologies involving real casework”
Acknowledgements

Graduate Students
- Jane Juusola (PhD)
- Ashley Hall (PhD)
- Michelle Alvarez (PhD)
- Erin Hanson (PhD)
- April Marrone (PhD)
- Debra Glidewell (MS)
- Jeffrey Ban (MS)
- Gigi Raker (MS)
- Stacey Smith (MS)
- Darlene Daniels (MS)
- Paulina Berdos (MS)
- Mindy Setzer (MS)
- Jeremy Fletcher (MS)

Graduate Students
- Susan Hastings (MS)
- Christine Sanders (MS)
- Katherine Press (MS)
- Pam Smith (MS)
- Micah Halpern (MS)
- Charly Parker (MS)

Undergraduate Research:
- Sami Nelson (PURE program)
- Ariana Albornoz (YES program)
- Michelle Josey (Honors in the major)

Collaborators
- Cordula Haas (Institute of Legal Medicine, Univ. of Zurich, Switzerland)
- Francesca DiPasquale, Holger Engel, Sascha Strauss, Helge Lubenow (QIAGEN)
- Richard Bisbing, Kirsten Primozic, Bianca Vigil (McCron)
- University of Tennessee (Pat Speck)
- Microcronics (John Gerdes)
- Cybergnetics (Mark Perlin)
We shall never cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.

T. S. Eliot

Thank you for your attention!