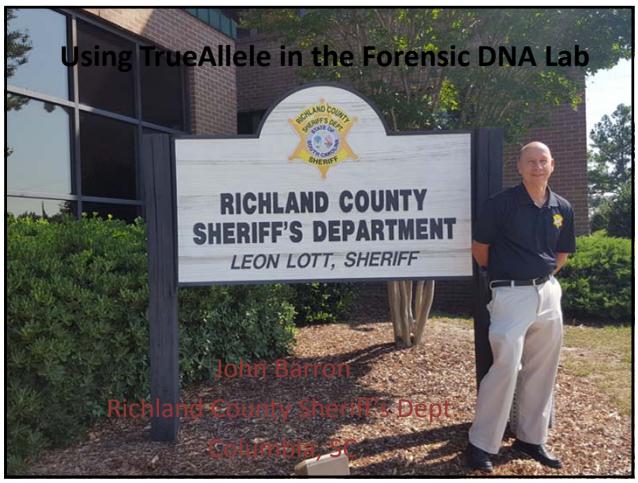
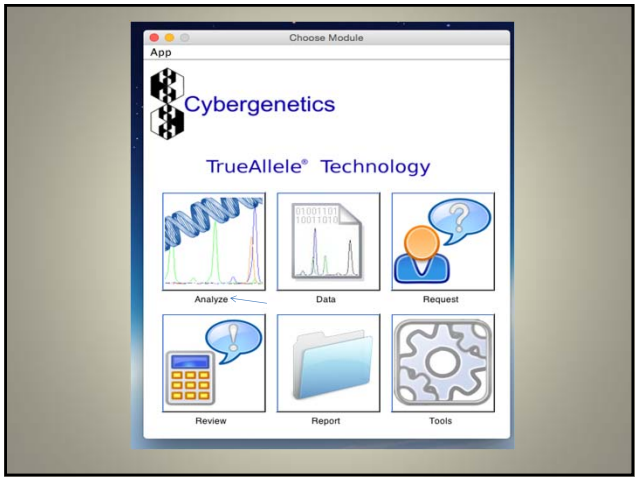
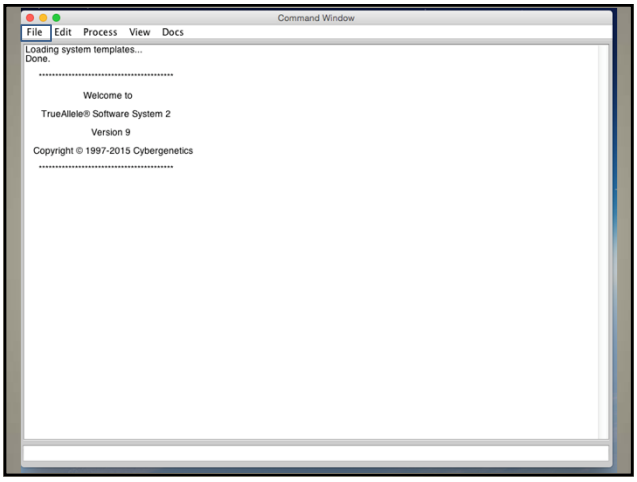
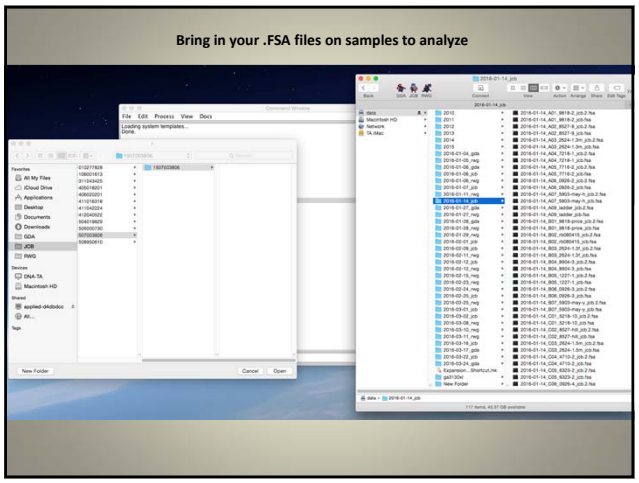


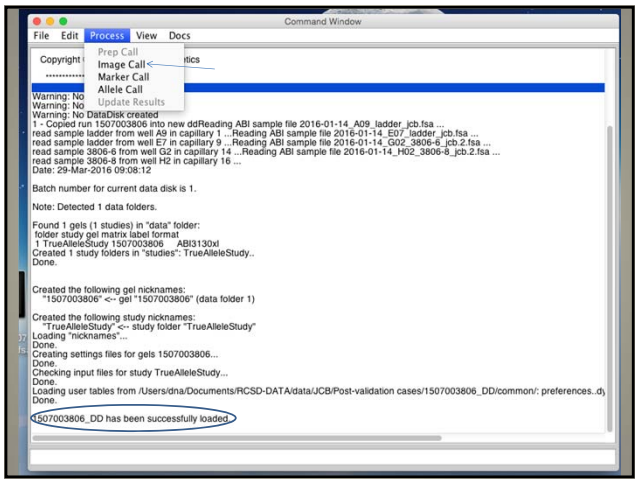
RCSD Validation Studies Highlights

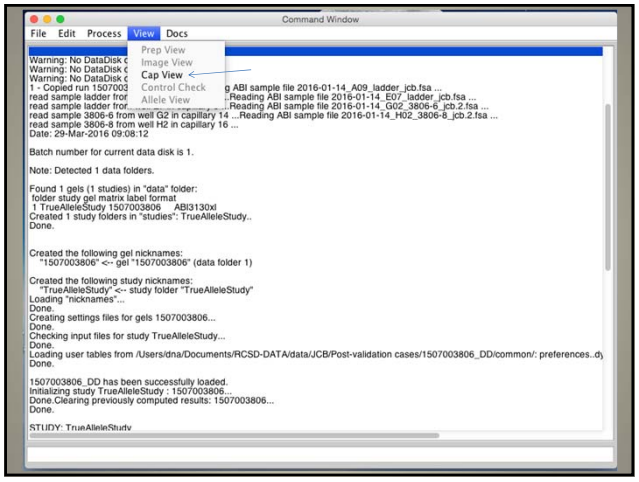
- Reproducible
- Correctly measured known mixture weights
- Correctly identified individuals present in various proportions up to 5 persons in mixture
- More sensitive than any threshold based mixture interpretation method
- Takes out biases by including all genotype probabilities all loci tested (if a locus or allele drops out it gets a low LR (<1) not just ignored)

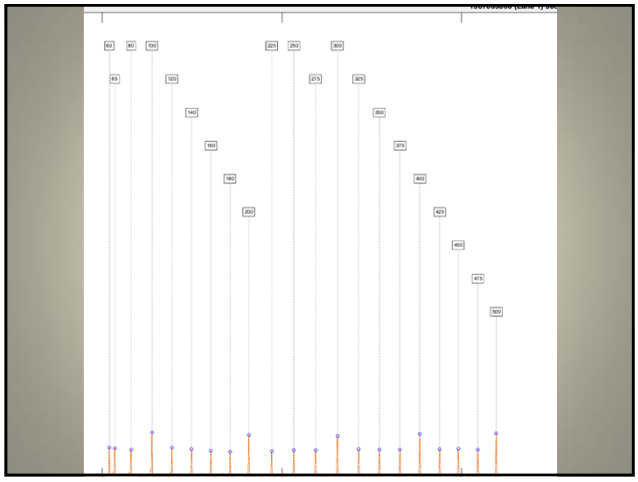


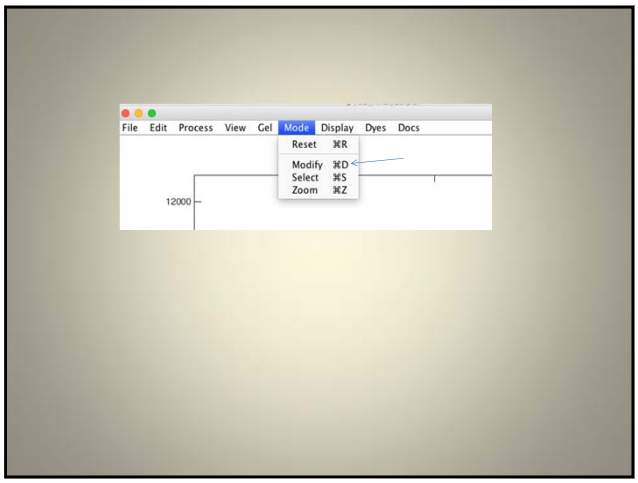


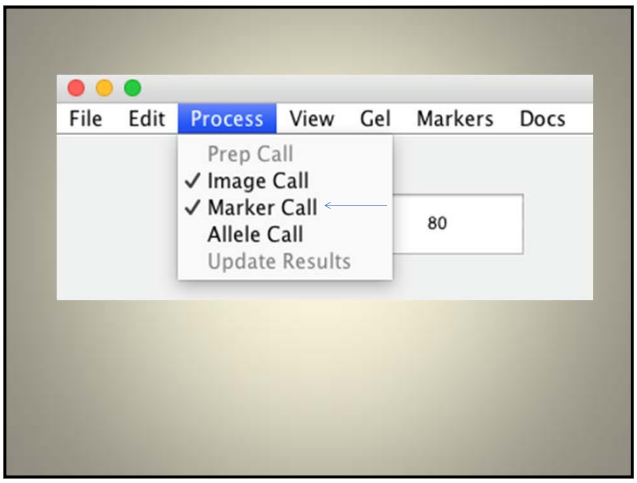

Bring in your .FSA files on samples to analyze

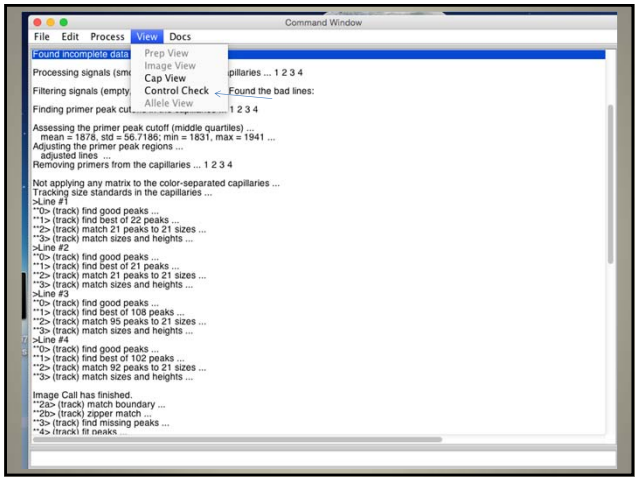

Command Window
File Edit Process View Docs
Prep Call
Image Call
Marker Call
Allele Call
Update Results
Warning: No DataDisk created
1 - Copied run 1507003806 into new ddReading ABI sample file 2016-01-14_A09_ladder_jcb.fsa ...
read sample ladder from well A9 in capillary 1 ...Reading ABI sample file 2016-01-14_E07_ladder_jcb.fsa ...
read sample ladder from well E7 in capillary 9 ...Reading ABI sample file 2016-01-14_G02_3806-8_jcb.2.fsa ...
read sample 3806-6 from well G2 in capillary 14 ...Reading ABI sample file 2016-01-14_H02_3806-8_jcb.2.fsa ...
read sample 3806-8 from well H2 in capillary 16 ...
Date: 29-Mar-2016 09:08:12
Batch number for current data disk is 1.
Note: Detected 1 data folders.
Found 1 gels (1 studies) in "data" folder:
folder study gel matrix label format
1 TrueAlleleStudy 1507003806 ABI3130xl
Created 1 study folders in "studies": TrueAlleleStudy..
Done.
Created the following gel nicknames:
"1507003806" <-- gel "1507003806" (data folder 1)
Created the following study nicknames:
"TrueAlleleStudy" <-- study folder "TrueAlleleStudy"
Loading "nicknames"...
Done.
Creating settings files for gels 1507003806...
Done.
Checking input files for study TrueAlleleStudy...
Done.
Loading user tables from /Users/dna/Documents/RCSD-DATA/data/JCB/Post-validation cases/1507003806_DD/common/: preferences..d
Done.
1507003806_DD has been successfully loaded.


Command Window
File Edit Process View Docs
Prep View
Image View
Cap View
Control Check
Allele View
Warning: No DataDisk c
Warning: No DataDisk c
Warning: No DataDisk c
1 - Copied run 1507003
read sample ladder fror
read sample ladder fror
g ABI sample file 2016-01-14_A09_ladder_jcb.fsa ...
.Reading ABI sample file 2016-01-14_E07_ladder_jcb.fsa ...
.Reading ABI sample file 2016-01-14_G02_3806-6_jcb.2.fsa ...
read sample 3806-6 from well G2 in capillary 14 ...Reading ABI sample file 2016-01-14_H02_3806-8_jcb.2.fsa ...
read sample 3806-8 from well H2 in capillary 16 ...
Date: 29-Mar-2016 09:08:12
Batch number for current data disk is 1.
Note: Detected 1 data folders.
Found 1 gels (1 studies) in "data" folder:
folder study gel matrix label format
1 TrueAlleleStudy 1507003806 ABI3130xl
Created 1 study folders in "studies": TrueAlleleStudy..
Done.
Created the following gel nicknames:
"1507003806" <-- gel "1507003806" (data folder 1)
Created the following study nicknames:
"TrueAlleleStudy" <-- study folder "TrueAlleleStudy"
Loading "nicknames"...
Done.
Creating settings files for gels 1507003806...
Done.
Checking input files for study TrueAlleleStudy...
Done.
Loading user tables from /Users/dna/Documents/RCSD-DATA/data/JCB/Post-validation cases/1507003806_DD/common/: preferences..d
Done.
1507003806_DD has been successfully loaded.
Initializing study TrueAlleleStudy : 1507003806...
Done.Clearing previously computed results: 1507003806...
Done.
STUDY: TrueAlleleStudy

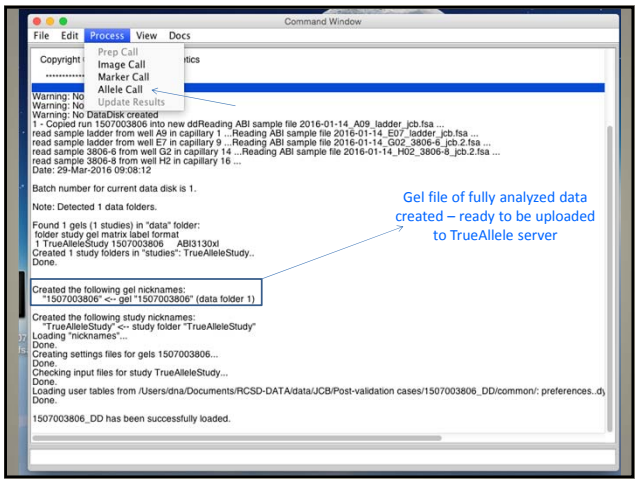


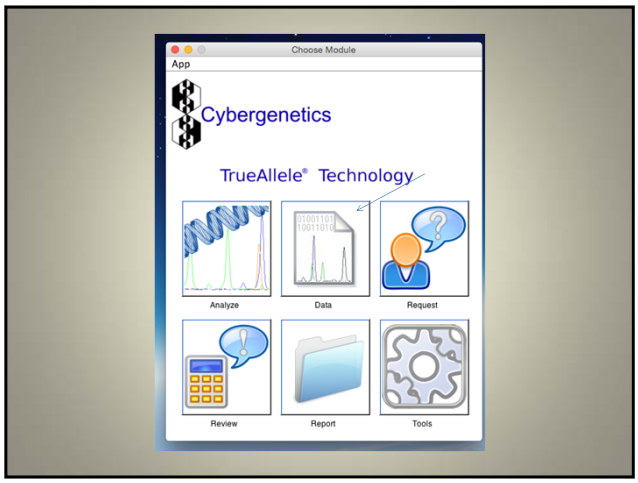

File Edit Process View Gel Mode Display Dyes Docs
Reset ⌘R
Modify ⌘D
Select ⌘S
Zoom ⌘Z
12000

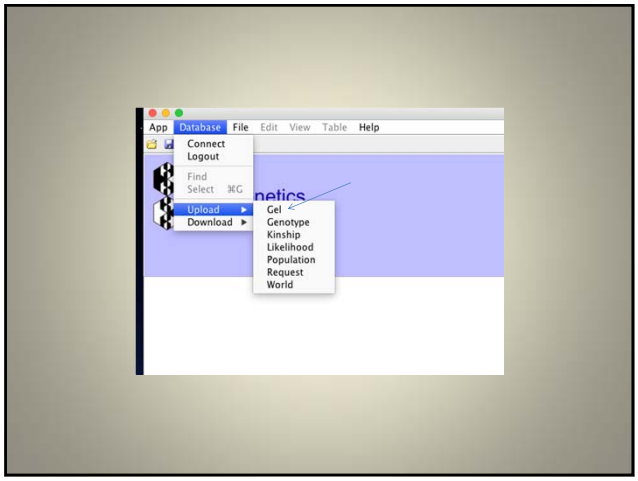

File Edit Process View Gel Markers Docs
Prep Call
✓ Image Call
✓ Marker Call
Allele Call
Update Results
80

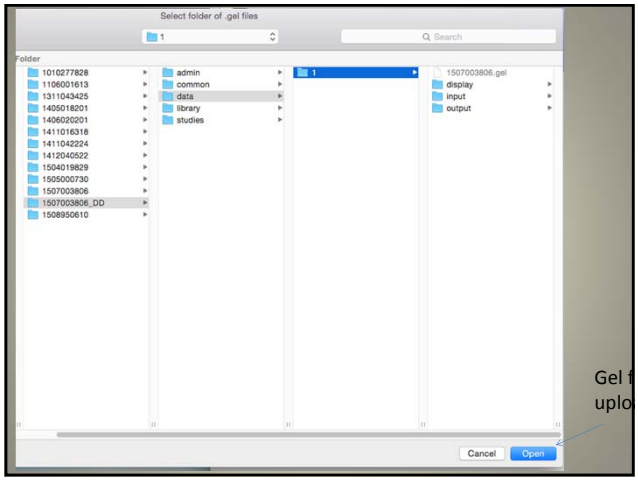

Command Window
File Edit Process View Docs
Prep View
Image View
Cap View
Control Check
Allele View
Found incomplete data
Found the bad lines:
Assessing the primer peak cutoff (middle quartiles) ...
mean = 1878, std = 56.7186; min = 1831, max = 1941 ...
Adjusting the primer peak regions ...
adjusted lines ...
Removing primers from the capillaries ... 1 2 3 4
Not applying any matrix to the color-separated capillaries ...
Tracking size standards in the capillaries ...
>Line #1
**0> (track) find good peaks ...
**1> (track) find best of 22 peaks ...
**2> (track) match 21 peaks to 21 sizes ...
**3> (track) match sizes and heights ...
>Line #2
**0> (track) find good peaks ...
**1> (track) find best of 21 peaks ...
**2> (track) match 21 peaks to 21 sizes ...
**3> (track) match sizes and heights ...
>Line #3
**0> (track) find good peaks ...
**1> (track) find best of 108 peaks ...
**2> (track) match 95 peaks to 21 sizes ...
**3> (track) match sizes and heights ...
>Line #4
**0> (track) find good peaks ...
**1> (track) find best of 102 peaks ...
**2> (track) match 92 peaks to 21 sizes ...
**3> (track) match sizes and heights ...
Image Call has finished.
**2a> (track) match boundary ...
**2b> (track) zipper match ...
**3> (track) find missing peaks ...

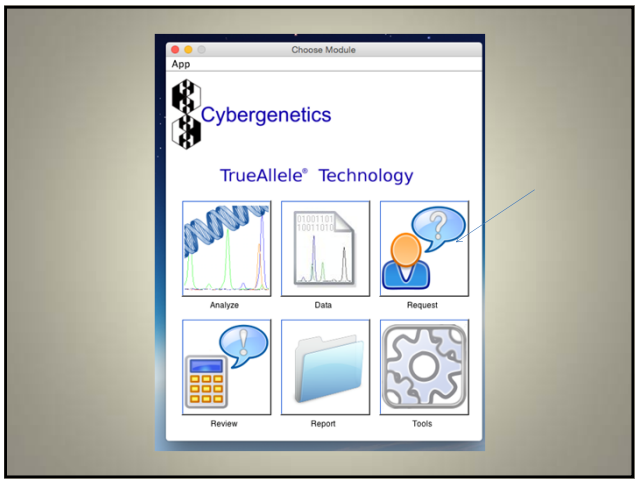
Control Check
File Edit Process View Gel Markers Docs
Start bp : 80
End bp : 89
1507003806: AMELO
bp
lane
Zoom
Darken
Brighten
Refresh
Size grid
Show dye
Rules
ladder interp
ladder missing
ladder overlay
negative peaks
outside marker window
positive missing
primer missing
Close
accept
Prev
Next

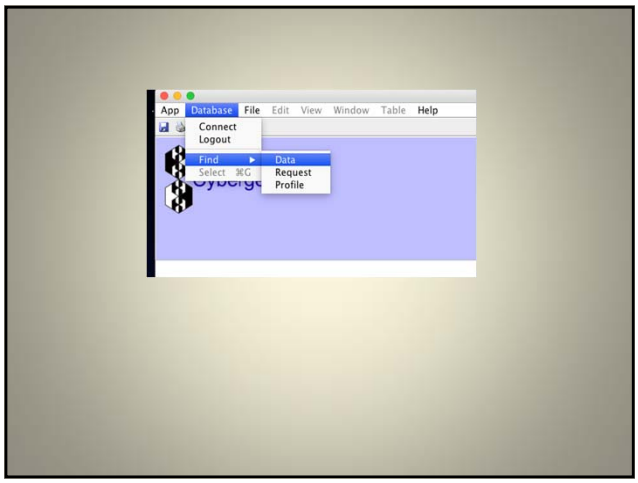

Command Window
File Edit Process View Docs
Prep Call
Image Call
Marker Call
Allele Call
Update Results
Gel file of fully analyzed data created – ready to be uploaded to TrueAllele server

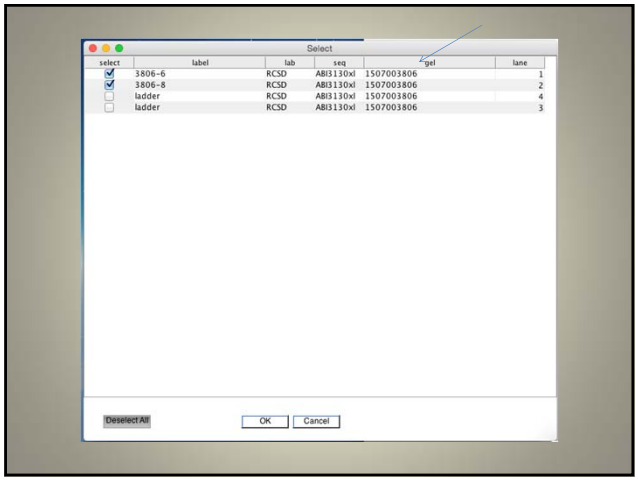

Choose Module
App
Cybergenetics
TrueAllele® Technology
Analyze
Data
Request
Review
Report
Tools

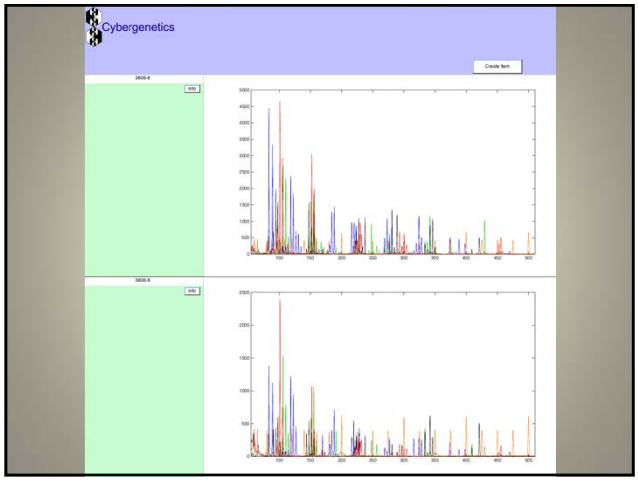

App Database File Edit View Table Help
Connect
Logout
Find
Select ⌘G
Upload
Download
Gel
Genotype
Kinship
Likelihood
Population
Request
World

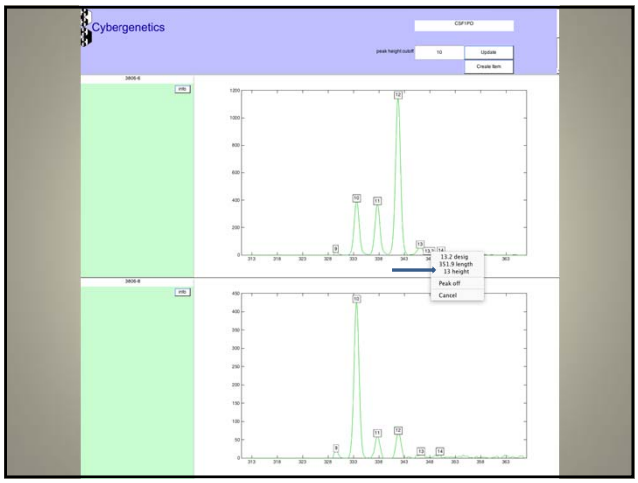

Select folder of .gel files
Folder
Cancel
Open

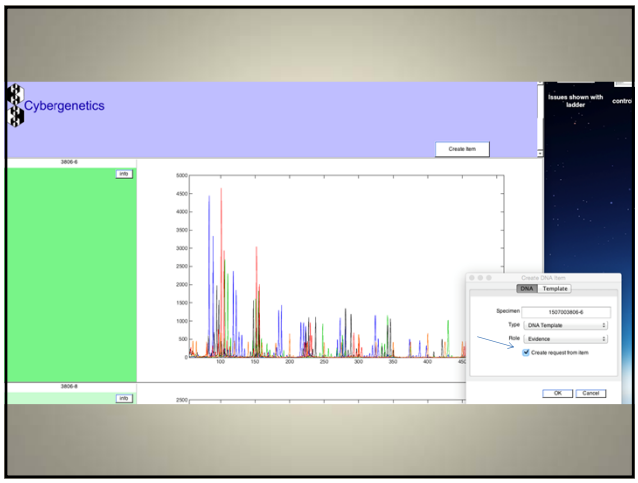

Choose Module
App
Cybergenetics
TrueAllele® Technology
Analyze
Data
Request
Review
Report
Tools

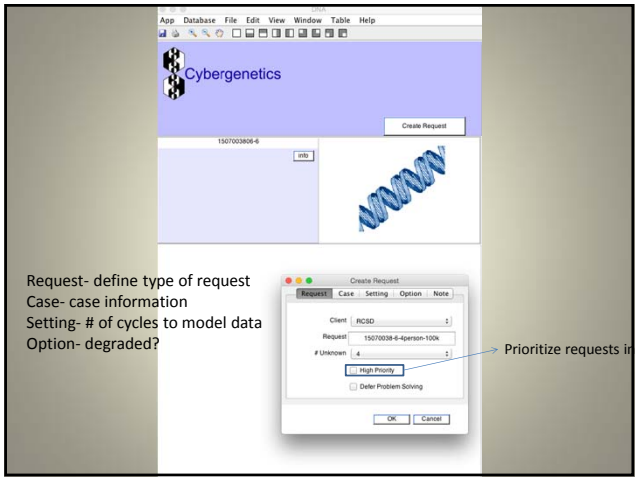

App
Database
File
Edit
View
Window
Table
Help
Connect
Logout
Find
Select ⌘G
Data
Request
Profile

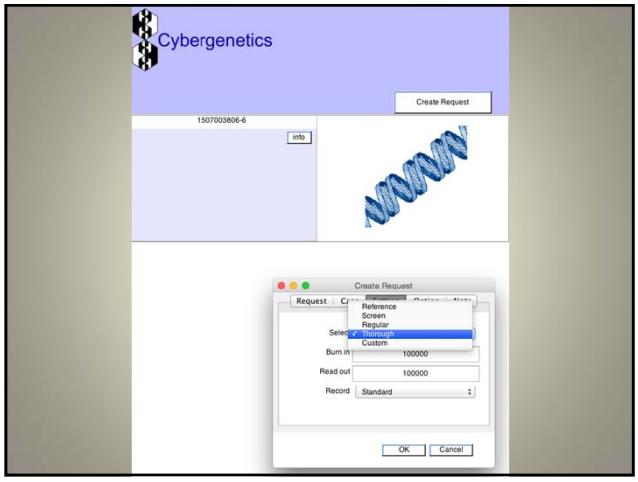

Select
select
label
lab
seq
gel
lane
3806-6 RCSD ABI3130xl 1507003806 1
3806-8 RCSD ABI3130xl 1507003806 2
ladder RCSD ABI3130xl 1507003806 4
ladder RCSD ABI3130xl 1507003806 3
Deselect All
OK
Cancel

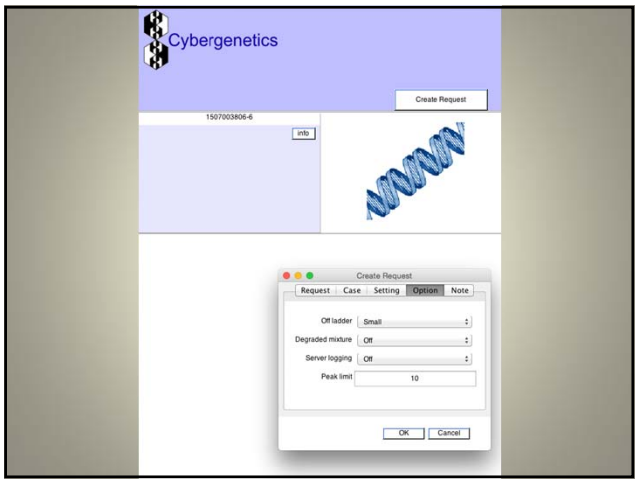

Cybergenetics
Create Item

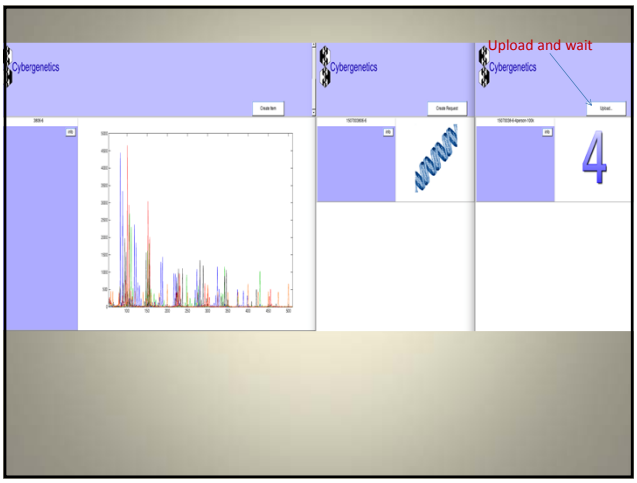

Cybergenetics
CSF1PO
peak height cutoff
10
Update
Create Item
3806-6
info
13.2 desig
351.9 length
13 height
Peak off
Cancel

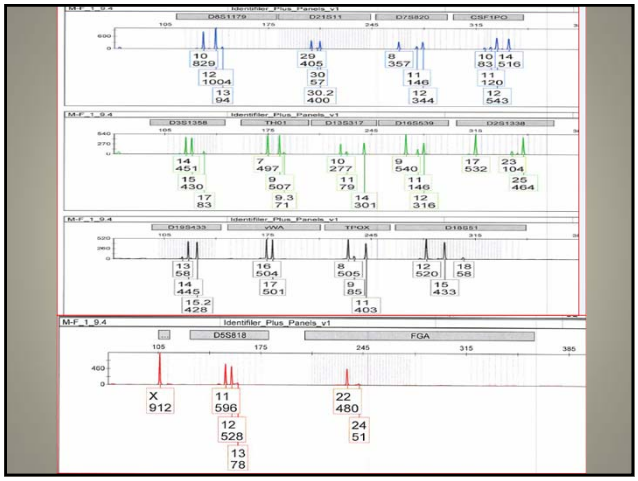

Cybergenetics
Create Item
3806-6
info
Issues shown with ladder
Create DNA Item
DNA
Template
Specimen
1507003806-6
Type
DNA Template
Role
Evidence
Create request from item
OK
Cancel

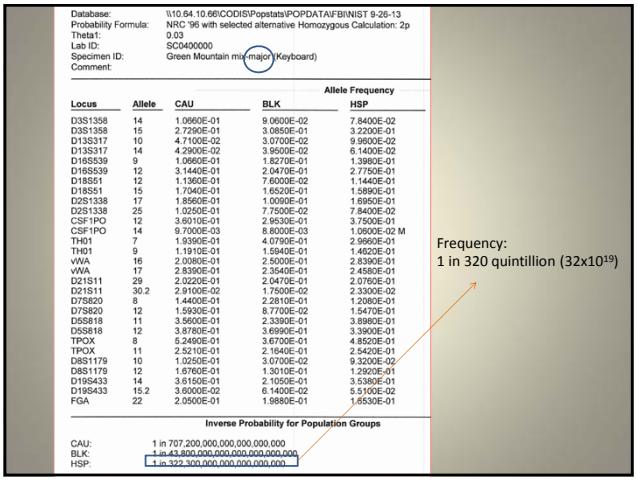

App Database File Edit View Window Table Help
Cybergenetics
Create Request
1507003806-6
info
Create Request
Request
Case
Setting
Option
Note
Client
RCSD
Request
15070038-6-4person-100k
Unknown
4
High Priority
Defer Problem Solving
OK
Cancel
Request- define type of request
Case- case information
Setting- # of cycles to model data
Option- degraded?
Prioritize requests in


Cybergenetics
Create Request
1507003806-6
info
Create Request
Request
Reference
Screen
Regular
Thorough
Custom
Burn in
100000
Read out
100000
Record
Standard
OK
Cancel









Database: \\10.64.10.66\CODIS\Popstats\POPDATA\FBI\NIST 9-26-13
Probability Formula: NRC '96 with selected alternative Homozygous Calculation: 2p
Theta1: 0.03
Lab ID: SC0400000
Specimen ID: Green Mountain mix (major) (Keyboard)
Comment:

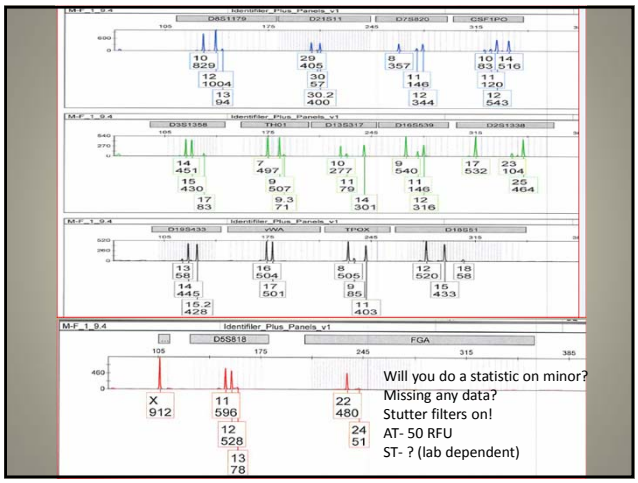
Allele Frequency

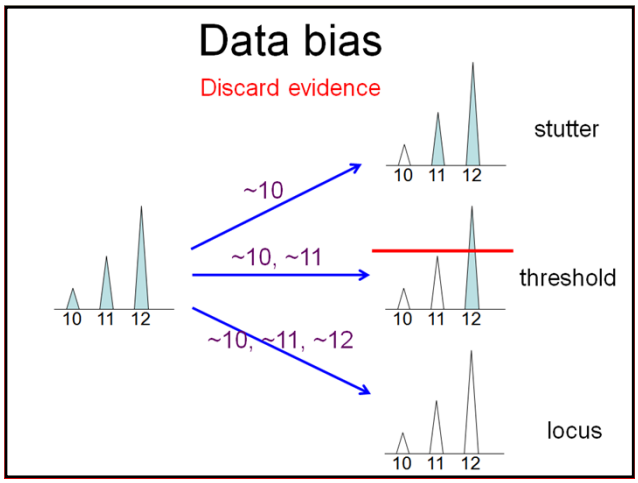
Locus	Allele	CAU	BLK	HSP
D3S1358	14	1.0660E-01	9.0600E-02	7.8400E-02
D3S1358	15	2.7290E-01	3.0850E-01	3.2200E-01
D13S317	10	4.7100E-02	3.0700E-02	9.9600E-02
D13S317	14	4.2900E-02	3.9500E-02	6.1400E-02
D16S539	9	1.0660E-01	1.8270E-01	1.3980E-01
D16S539	12	3.1440E-01	2.0470E-01	2.7750E-01
D18S51	12	1.1360E-01	7.6000E-02	1.1440E-01
D18S51	15	1.7040E-01	1.6520E-01	1.5890E-01
D2S1338	17	1.8560E-01	1.0090E-01	1.6950E-01
D2S1338	25	1.0250E-01	7.7500E-02	7.8400E-02
CSF1PO	12	3.6010E-01	2.9530E-01	3.7500E-01
CSF1PO	14	9.7000E-03	8.8000E-03	1.0600E-02 M
TH01	7	1.9390E-01	4.0790E-01	2.9660E-01
TH01	9	1.1910E-01	1.5940E-01	1.4620E-01
vWA	16	2.0080E-01	2.5000E-01	2.8390E-01
vWA	17	2.8390E-01	2.3540E-01	2.4580E-01
D21S11	29	2.0220E-01	2.0470E-01	2.0760E-01
D21S11	30.2	2.9100E-02	1.7500E-02	2.3300E-02
D7S820	8	1.4400E-01	2.2810E-01	1.2080E-01
D7S820	12	1.5930E-01	8.7700E-02	1.5470E-01
D5S818	11	3.5600E-01	2.3390E-01	3.8980E-01
D5S818	12	3.8780E-01	3.6990E-01	3.3900E-01
TPOX	8	5.2490E-01	3.6700E-01	4.8520E-01
TPOX	11	2.5210E-01	2.1640E-01	2.5420E-01
D8S1179	10	1.0250E-01	3.0700E-02	9.3200E-02
D8S1179	12	1.6760E-01	1.3010E-01	1.2920E-01
D19S433	14	3.6150E-01	2.1050E-01	3.5380E-01
D19S433	15.2	3.6000E-02	6.1400E-02	5.5100E-02
FGA	22	2.0500E-01	1.9880E-01	1.6530E-01

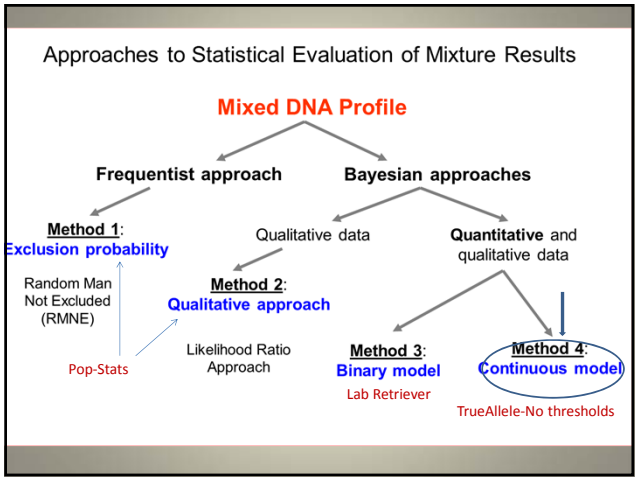
Inverse Probability for Population Groups

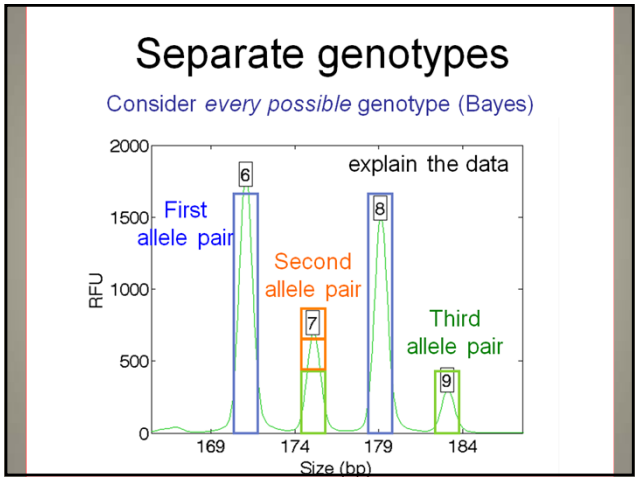
CAU: 1 in 707,200,000,000,000,000,000
BLK: 1 in 43,800,000,000,000,000,000,000
HSP: 1 in 322,300,000,000,000,000,000

Frequency:
1 in 320 quintillion (32×10^{19})


Will you do a statistic on minor?
Missing any data?
Stutter filters on!
AT- 50 RFU
ST- ? (lab dependent)


Data bias
Discard evidence
~10
~10, ~11
~10, ~11, ~12
stutter
threshold
locus
10 11 12

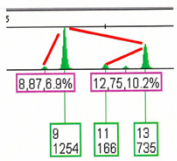

Approaches to Statistical Evaluation of Mixture Results
Mixed DNA Profile
Frequentist approach
Bayesian approaches
Method 1: Exclusion probability
Random Man Not Excluded (RMNE)
Pop-Stats
Qualitative data
Method 2: Qualitative approach
Likelihood Ratio Approach
Quantitative and qualitative data
Method 3: Binary model
Lab Retriever
Method 4: Continuous model
TrueAllele-No thresholds


Separate genotypes
Consider every possible genotype (Bayes)
explain the data
First allele pair
Second allele pair
Third allele pair
RFU
Size (bp)

Probabilistic Modeling of TA

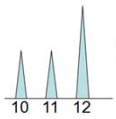
Mathematical Modeling of the Data → 50-100,000 Simulations (MCMC) → Probable ***Genotypes*** to explain the mixture

PHR, Mix Ratio, Stutter etc…



Genotypes	Probability
9,11	76%
11,11	15%
11,13	2%
8,11	2%
11,12	2%
9,9	1%
9,12	<1%
10,11	<1%
8,12	<1%
8,9	<1%

Data



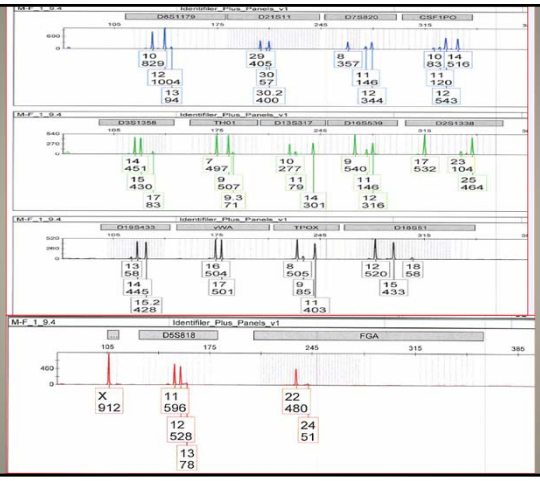
Now I look at my suspect!!

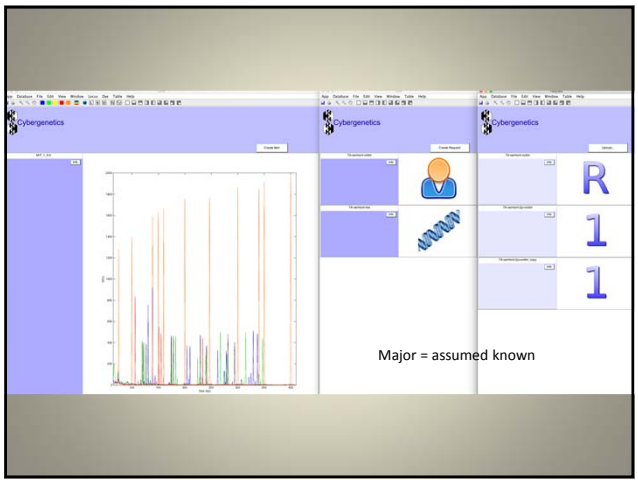
Genotype 2

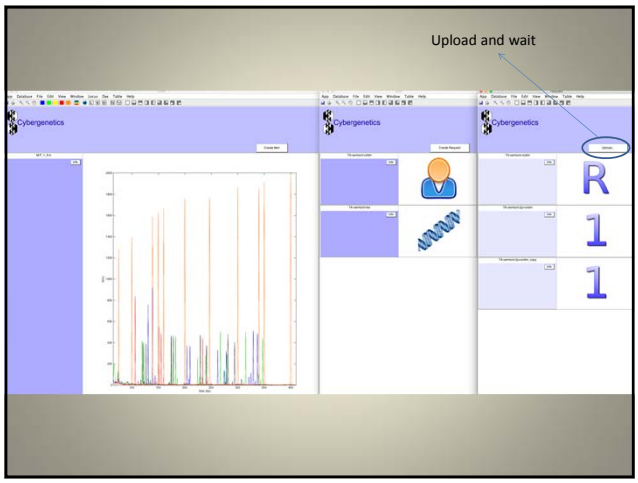
10, 10 @ 10%
10, 11 @ 10%
10, 12 @ 10%
11, 11 @ 10%
11, 12 @ 40%
12, 12 @ 20%

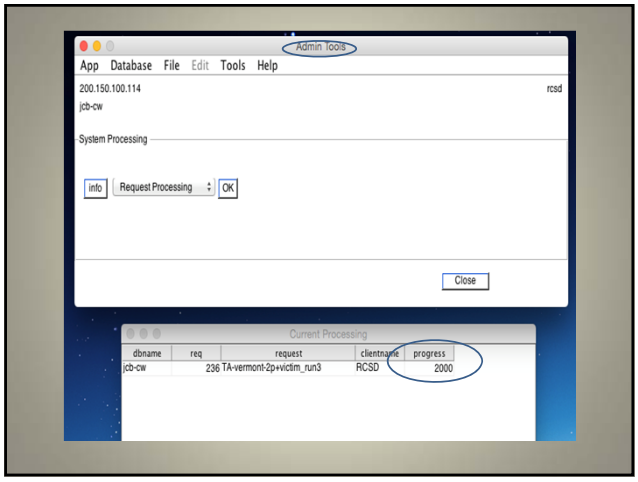
Compare with 11,12

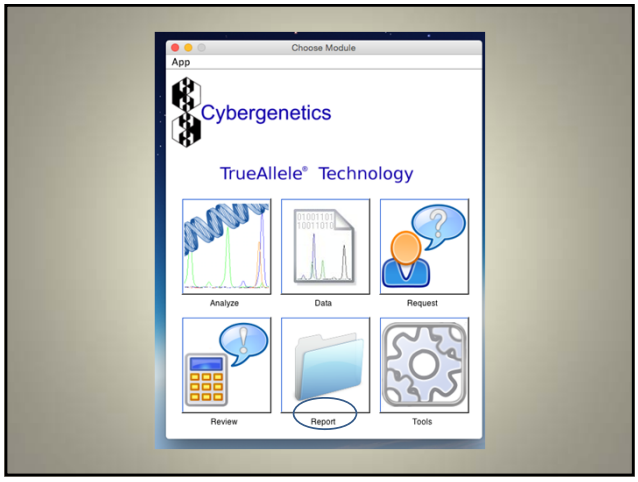
Match statistic = $\frac{\text{Prob\{match\}}}{\text{Prob\{coincidence\}}} = \frac{40\%}{4\%} = 10$

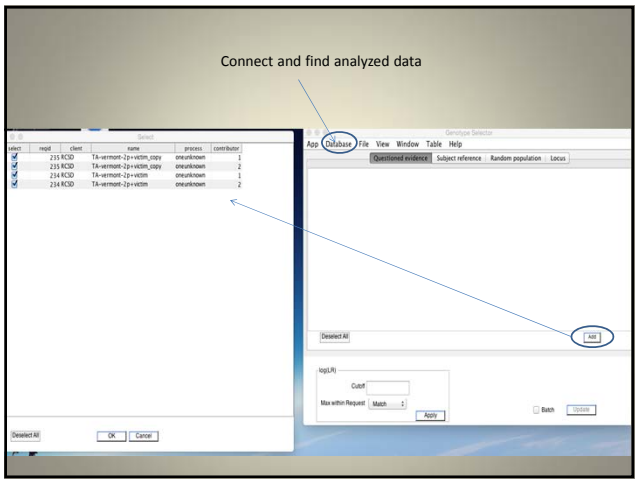


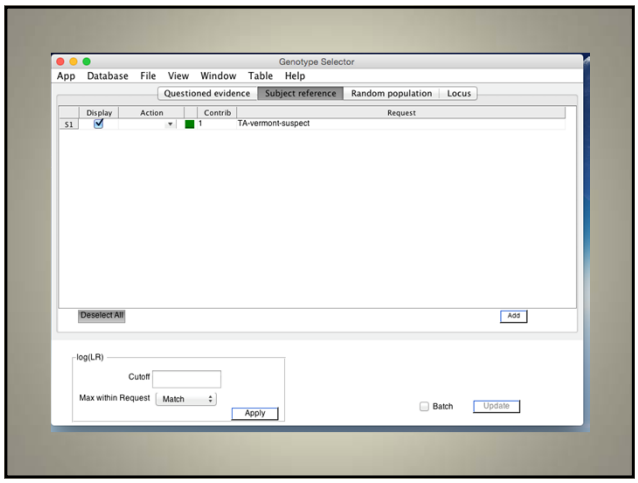

Cybergenetics
R
1
1
Major = assumed known

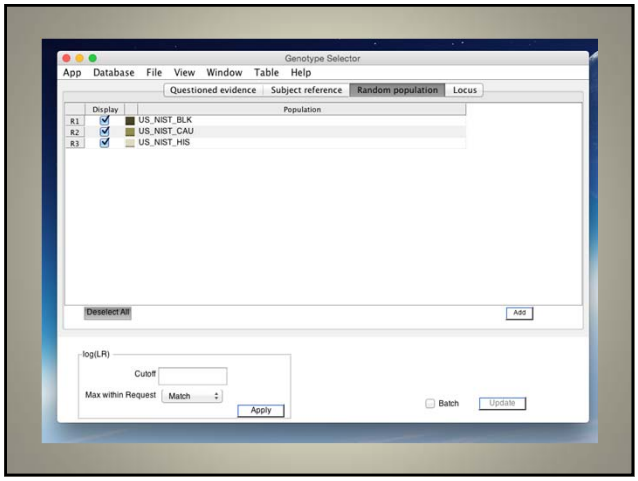

Upload and wait
Cybergenetics
R
1
1

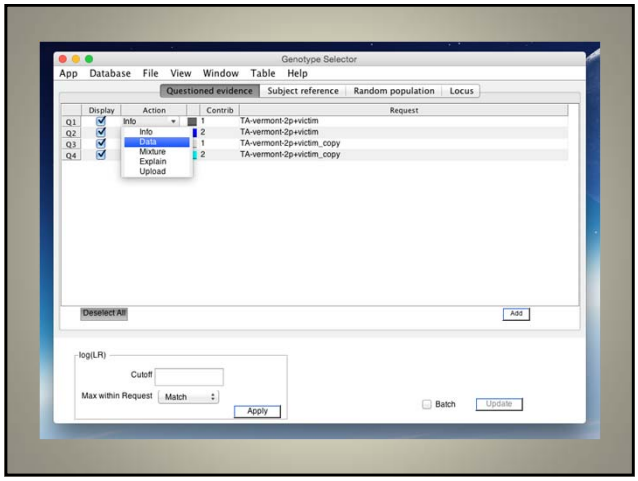

Admin Tools
App Database File Edit Tools Help
200.150.100.114
rcsd
jcb-cw
System Processing
Info
Request Processing
OK
Close
Current Processing
dbname
req
request
clientname
progress
jcb-cw
236
TA-vermont-2p+victim_run3
RCSD
2000

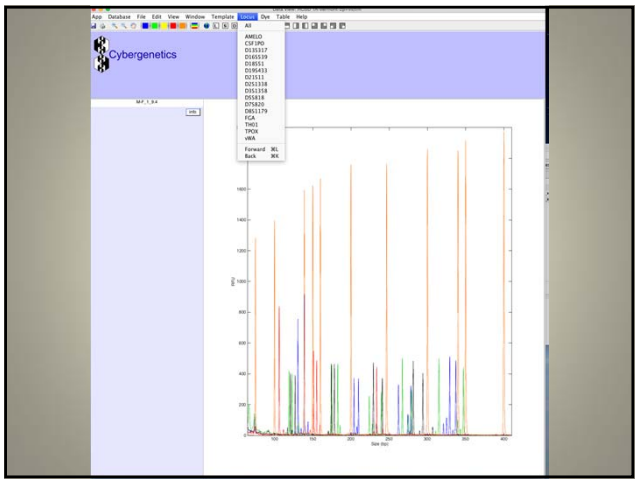

Choose Module
App
Cybergenetics
TrueAllele® Technology
Analyze
Data
Request
Review
Report
Tools

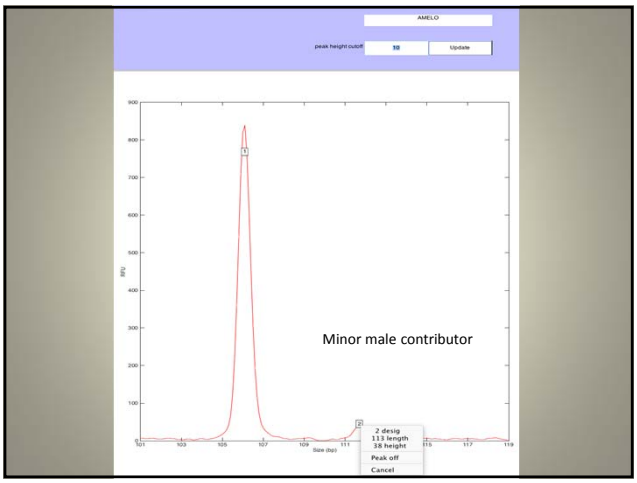

Connect and find analyzed data
Genotype Selector
App Database File View Window Table Help
Questioned evidence Subject reference Random population Locus
Deselect All
Add
log(LR)
Cutoff
Max within Request
Match
Apply
Batch
Update
OK
Cancel

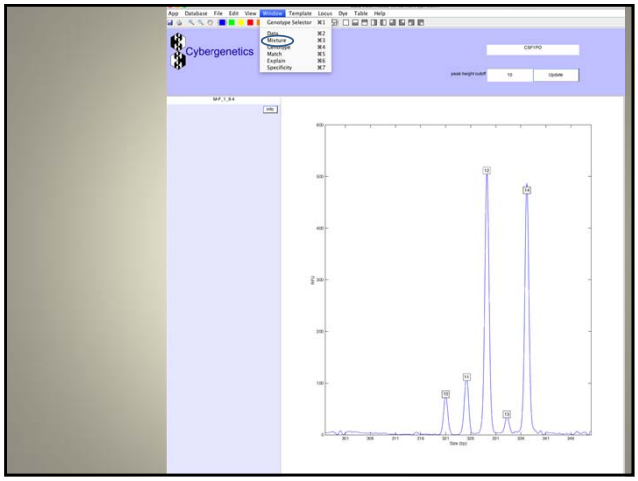

Genotype Selector
App Database File View Window Table Help
Questioned evidence Subject reference Random population Locus
Display Action Contrib Request
S1 1 TA-vermont-suspect
Deselect All
Add
log(LR)
Cutoff
Max within Request
Match
Apply
Batch
Update

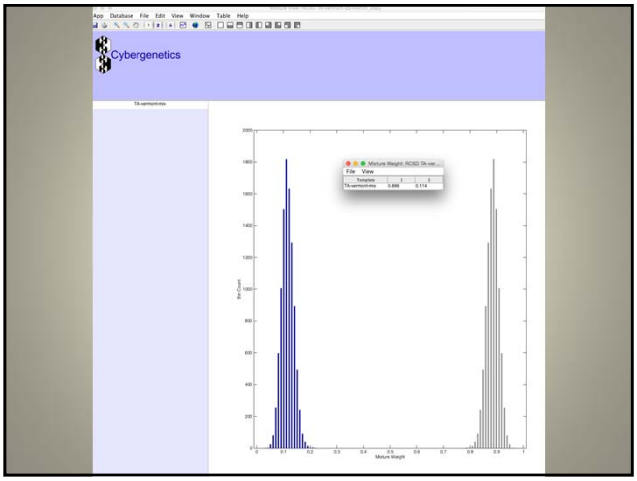

Genotype Selector
App Database File View Window Table Help
Questioned evidence Subject reference Random population Locus
Display Population
R1 US_NIST_BLK
R2 US_NIST_CAU
R3 US_NIST_HIS
Deselect All
Add
log(LR)
Cutoff
Max within Request
Match
Apply
Batch
Update

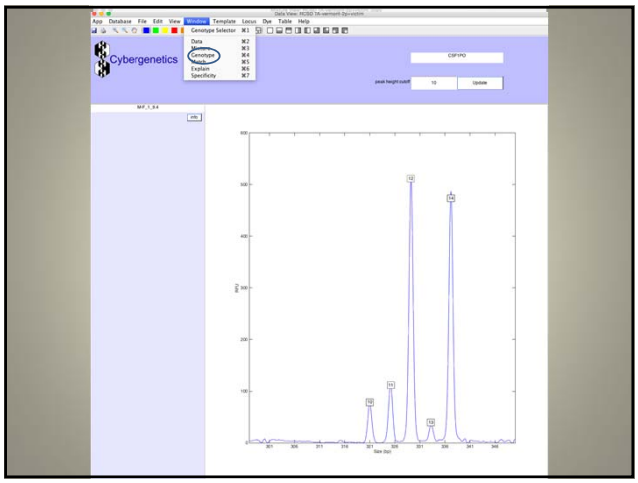

Genotype Selector
App Database File View Window Table Help
Questioned evidence Subject reference Random population Locus
Display Action Contrib Request
Q1 Info 1 TA-vermont-2p+victim
Q2 2 TA-vermont-2p+victim
Q3 1 TA-vermont-2p+victim_copy
Q4 2 TA-vermont-2p+victim_copy
Info
Data
Mixture
Explain
Upload
Deselect All
Add
log(LR)
Cutoff
Max within Request
Match
Apply
Batch
Update

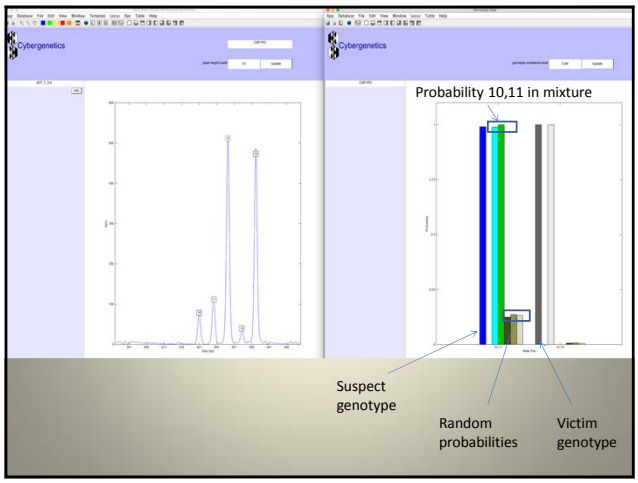

Cybergenetics

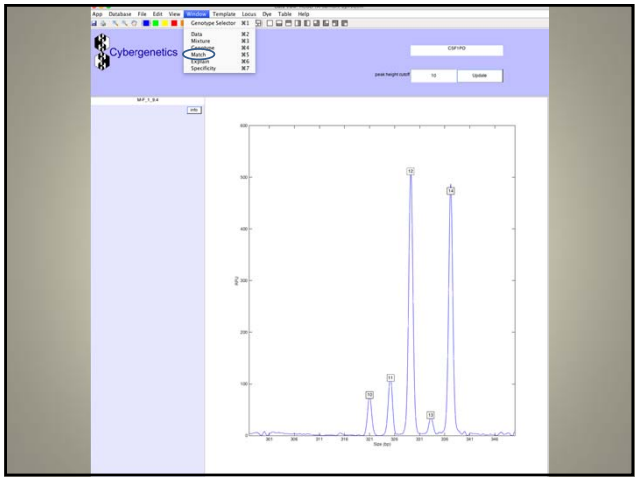

AMELO
peak height cutoff
10
Update
Minor male contributor
2 desig
113 length
38 height
Peak off
Cancel

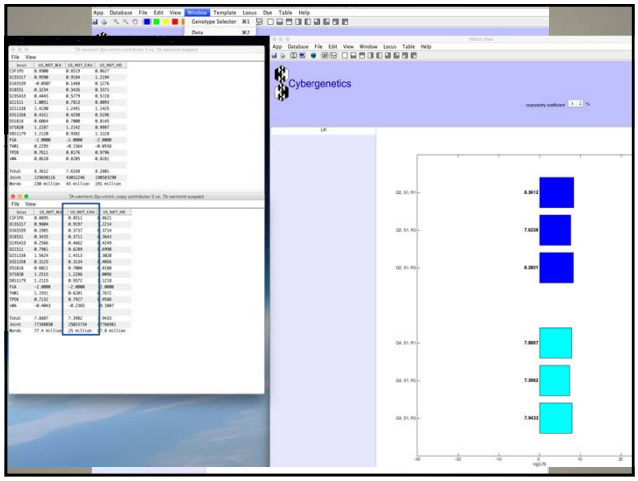

Cybergenetics
Genotype Selector
Data
Mixture
Genotype
Match
Explain
Specificity
CSF1PO
peak height cutoff
10
Update


Cybergenetics


Cybergenetics

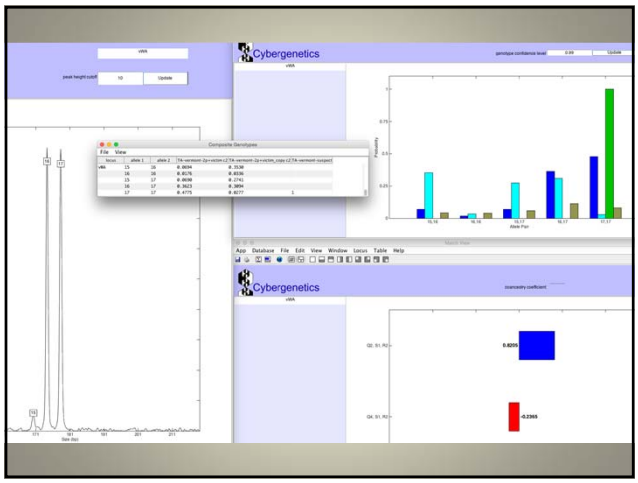

Cybergenetics
Probability 10,11 in mixture
Suspect genotype
Random probabilities
Victim genotype

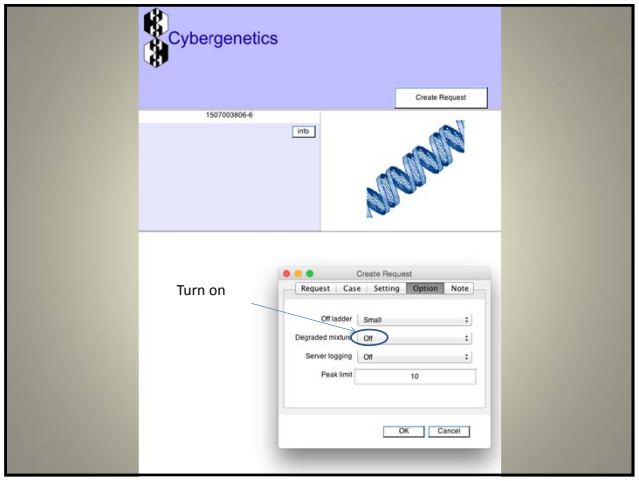

Cybergenetics

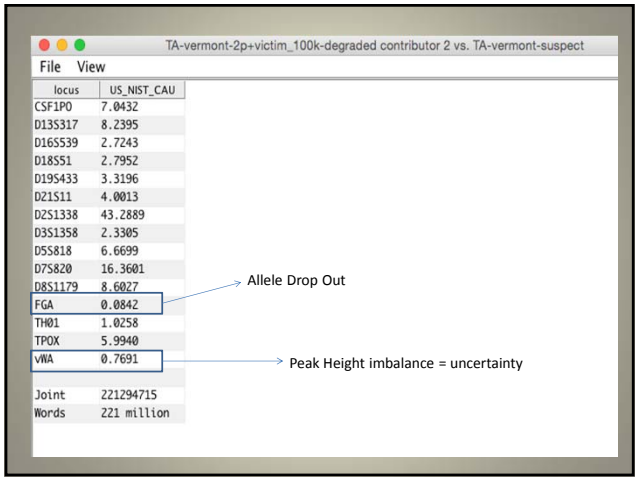

Cybergenetics

A match between a contributor to the mixture from the evidence (item) and the reference (name) is 25 million times more likely than a coincidental match to an unrelated person.

$$\frac{\text{Posterior probability (after analyzing data)}}{\text{Prior probability (before analyzing or random probability)}}$$







locus	US_NIST_CAU
CSF1PO	7.0432
D13S317	8.2395
D16S539	2.7243
D18S51	2.7952
D19S433	3.3196
D21S11	4.0013
D2S1338	43.2889
D3S1358	2.3305
D5S818	6.6699
D7S820	16.3601
D8S1179	8.6027
FGA	0.0842
TH01	1.0258
TPOX	5.9940
vWA	0.7691
Joint	221294715
Words	221 million

A match between a contributor to the evidence (Item) and the reference (suspect) is **221 million** times more probable than a coincidental match to an unrelated person.

Ask Further Questions or Get Copy of Slides:

John Barron
Phone: 803-576-3088
Email: jbarron@rcsd.net