Three Types of DNA Forensic Issues

- **Single Source:** DNA profile of the evidence sample providing indications of it being of a single source origin
- **Mixture of DNA:** Evidence sample DNA profile suggests it being a mixture of DNA from multiple (more than one) individuals
- **Kinship Determination:** Evidence sample DNA profile compared with that of one or more reference profiles is to be used to determine the validity of stated biological relatedness among individuals

- Interpretation of a result:
- 1. Non-match - exclusion
- 2. Inconclusive - no decision
- 3. Match - estimate frequency

What is an Exclusion?

Single Source: DNA profiles of the evidence and reference samples differ from each other at one or more loci; i.e., barring sample mix-up and/or false identity of samples, reference individual is not the source of DNA found in the evidence sample

DNA Mixture: Reference DNA profile contains alleles (definitely) not observed in the evidence sample for one or more loci; i.e., reference individual is excluded as a part contributor of the mixture DNA of the evidence sample

Kinship: Allele sharing among evidence and reference samples disagrees with the Mendelian rules of transmission of alleles with the stated relationship being tested

What is an Inclusion?

Single Source: DNA profiles of the evidence and reference samples are identical at each interpretable locus (also called DNA Match); i.e., reference individual may be the source of DNA in the evidence sample

DNA Mixture: Alleles found in the reference sample are all present in the mixture; i.e., reference individual can not be excluded as a part contributor of DNA in the evidence sample

Kinship: Allele sharing among evidence and reference samples is consistent with Mendelian rules of transmission of alleles with the stated relationship being tested; i.e., the stated biological relationship cannot be rejected

When is the Observation at a Locus Inconclusive?

- Compromised nature of samples tested failed to definitively exclude or include reference individuals
- May occur for one or more loci, while other loci typed may lead to unequivocal definite inclusion/ exclusion conclusions
- Caused often by DNA degradation (resulting in allele drop out), and/or low concentration of DNA (resulting in alleles with low peak height and/or area) for the evidence sample

Statistical Assessment of DNA Evidence

- Needed most frequently in the inclusionary events
- (Apparent) exclusionary cases may also be sometimes subjected to statistical assessment, particularly for kinship determination because of genetic events such as mutation, recombination, etc.
- Loci providing inconclusive results are often excluded from statistical considerations
- Even if one or more loci show inconclusive results, inclusionary observations of the other typed loci can be subjected to statistical assessment

Exclusion vs Match

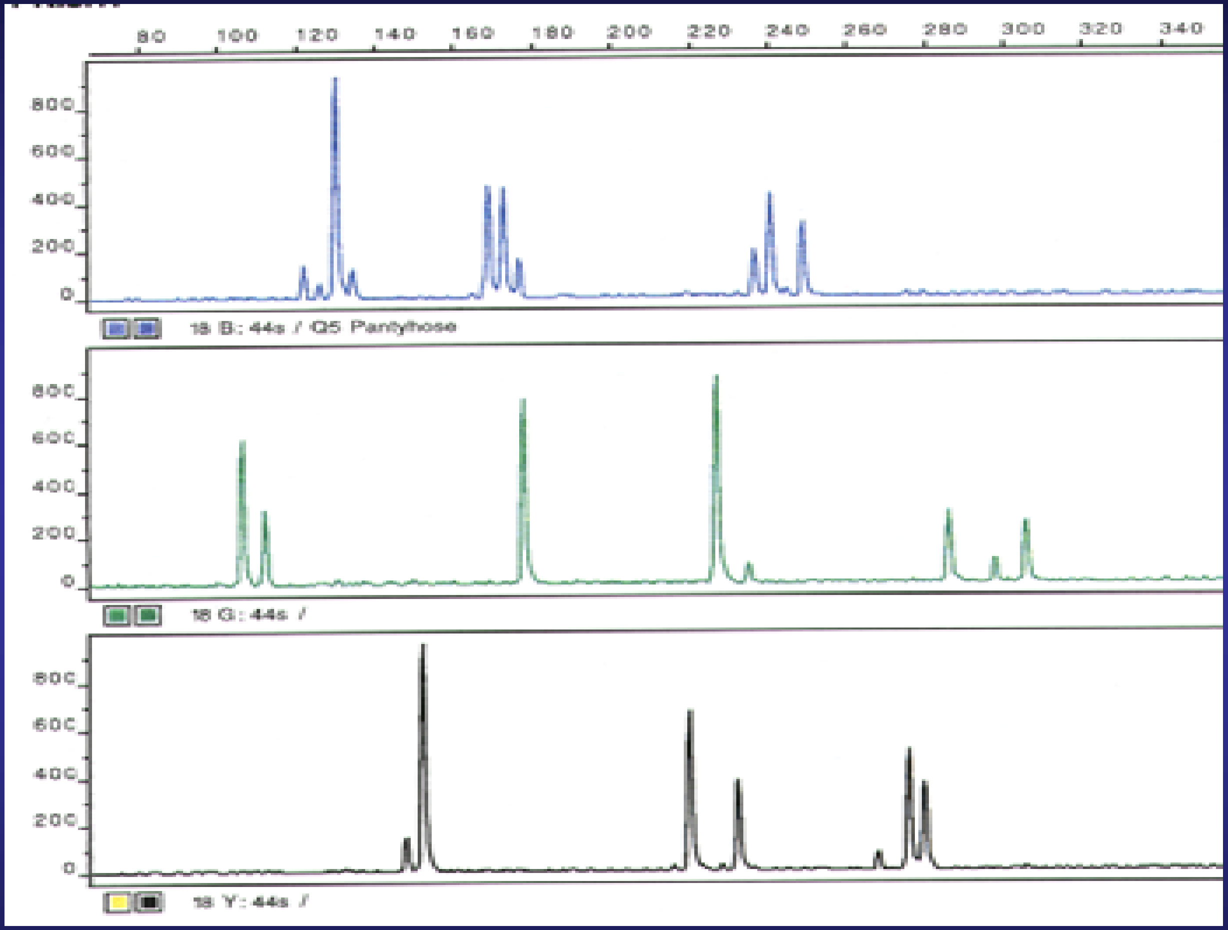
- Exclusion – numbers are not needed
- Match - requires a numerical estimate (weight of evidence)

Statistical Analysis

About “Q” sample

- “K” matches “Q”
- Who else could match “Q"
- partial profile, mixtures

Mixtures



80
100
120
140
160
180
200
220
240
260
280
300
320
340
800
600
400
200
0
18 B: 44s / Q5 Pantyhose
18 G: 44s /
18 Y: 44s /

What Constitutes a DNA Mixture?

- Presence of three or more alleles at several loci typed

Rationale: Typically a single individual displays one or two alleles

However, exceptions do occur – e.g., three allele profiles at the TPOX locus

What Constitutes a DNA Mixture?

- Notable imbalance in alleles at a locus

Rationale: Peak height/area difference, intensity difference, etc. caused by different contributions of amount of DNA from contributors of the sample.

Common Mixture Interpretations

- Profiles of the tested persons can explain all alleles detected in the mixture sample
- Profiles of the tested persons do not explain the alleles detected in the mixture sample
- Mixture contains unexplained alleles beyond the ones present in the profiles of the known persons profiled
- Major and minor contributors
- Interpretable and uninterpretable

Mixture Interpretation

Alleles considered first must meet interpretation guidelines for data analysis

- RFU Threshold
- Peak ratios/percent stutter
- “Match” criteria

Calculations for mixed profiles are based on same statistical assumptions used for calculating single contributor profiles

DAB
Statistics

- Alternate methods for assessing weight
- Rarely is there only one approach
- Philosophy, experience, legal system
- Practicality, available data, assumptions
- Simplistic approaches are acceptable
- Convey to fact finder
- Probability of Exclusion or LR
- Bayesian inferences
- NRC II Report

Two common questions can be asked:

- How often would a random person be excluded as a contributor of the observed mixture? (Exclusion Probability)

- What statistical support is there for postulated hypotheses on the origin of the mixture? (Likelihood Ratio)

Frequentist Approach of Statistical Assessment for DNA Mixture

- When the evidence mixture DNA profile fails to exclude a reference sample as a part contributor, and more commonly a set of reference samples together explains all alleles seen in the mixture, one or more of the following questions are answered:
- How often a random person would be excluded as a part contributor of the mixture sample? – also called Exclusion Probability, the complement of which is the inclusion probability, giving the expected chance of Coincidental Inclusion
- (Note: This answer is based on the data on the evidence sample alone, without any consideration of the profiles of the reference samples)
- With a stipulation on the number of contributors, how often a random person's DNA, mixed with that of one or more of the reference persons, would provide a mixture profile as seen in the evidence sample, given that the reference persons are also part contributors of the DNA mixture
- (Note: This answer considers data on the profiles of evidence sample as well as those of the reference samples stipulated to be part contributors)

Hardy - Weinberg Equilibrium

A_1A_1	A_1A_2	A_2A_2
$p_1{}^2$	$2p_1p_2$	$p_2{}^2$

freq(A_1) = p_1

freq(A_2) = p_2

	A_1	A_2
A_1	$p_1{}^2$ A_1A_1	p_1p_2 A_1A_2
A_2	p_1p_2 A_1A_2	$p_2{}^2$ A_2A_2

$$(p_1 + p_2)^2 = p_1{}^2 + 2p_1p_2 + p_2{}^2$$

D3S1358 = 16, 16 (homozygote)

Frequency of 16 allele = 0.3071

When same allele:

Genotype Frequency $= p^2$
(for now!)

Genotype freq = 0.3071 x 0.3071 = 0.0943

VWA = 15, 17 (heterozygote)

Frequency of 15 allele = 0.2361

Frequency of 17 allele = 0.1833

When heterozygous:

Frequency = 2 X allele 1 freq X allele 2 freq

($2pq$)

Genotype freq = 2 x 0.2361 x 0.18331 = 0.0866

If you do not understand it,
then do not present it

Statistics

Mixtures

Likelihood Ratios

P_E

Probability of Exclusion

Not as powerful as LR

Exclusion Probability

How often a random person would be excluded as a contributor of the observed DNA mixture?

Exclusion Probability

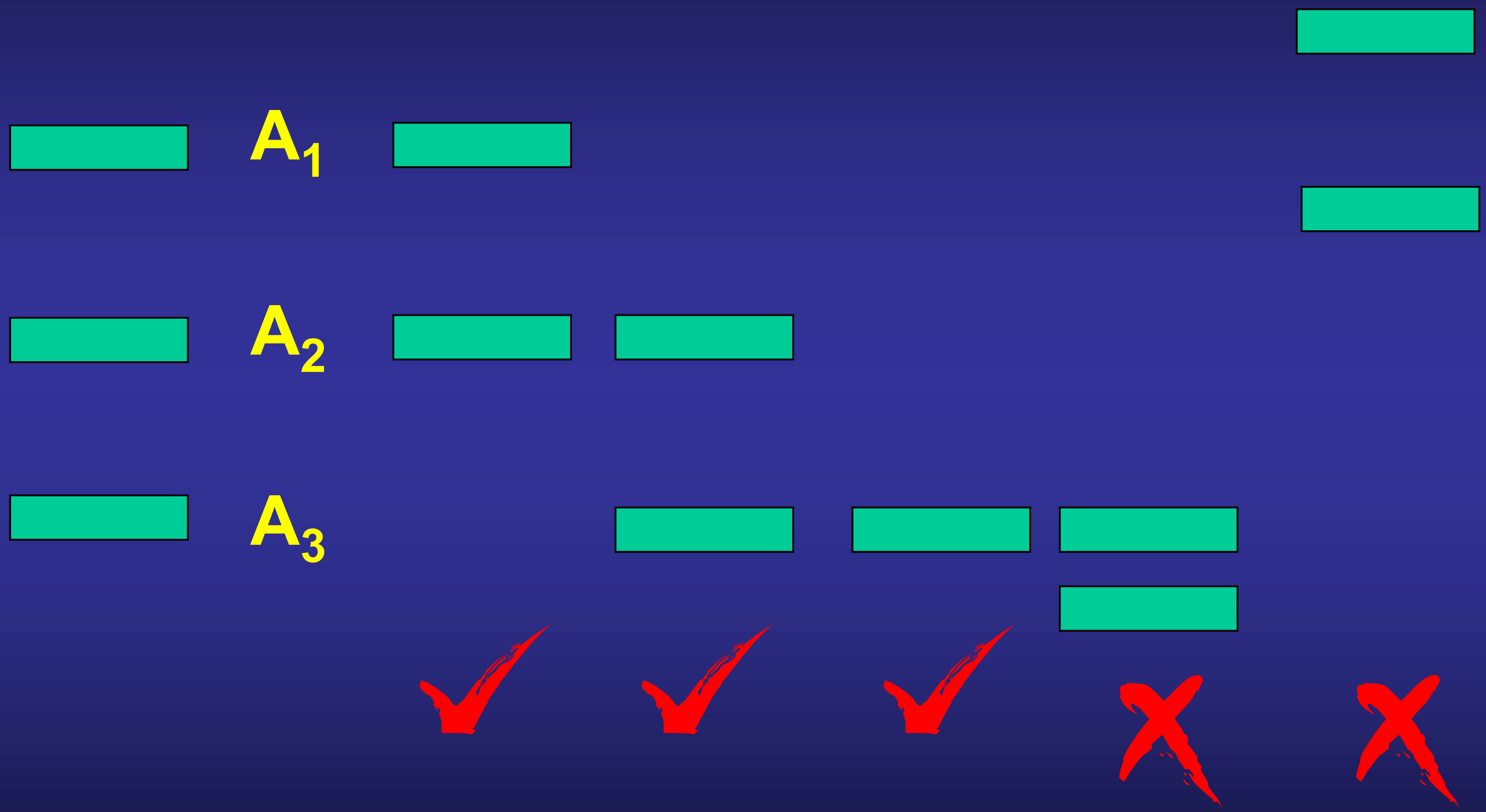
- Power of the DNA testing panel for excluding potential noncontributing individuals from the profile
- Based on the alleles in the mixture…not the profiles of potential contributors

Suspect Evidence

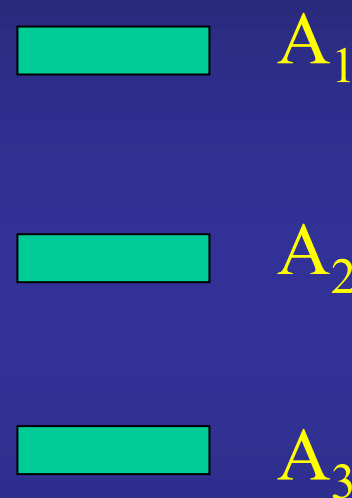
No assumption - # of contributors

Three Allele Scenario

Evidence



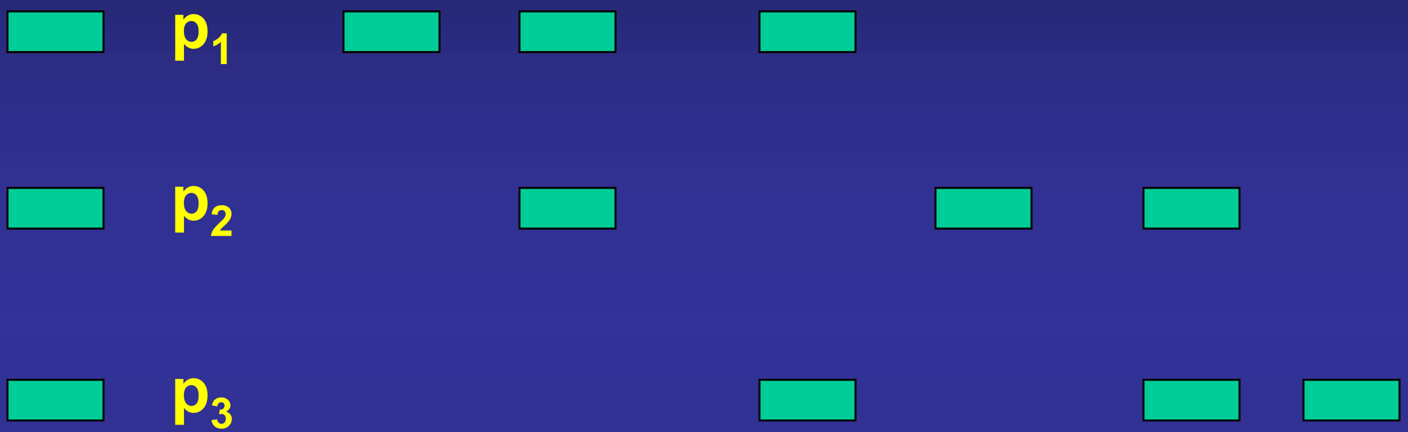
Evidence



$A_1 + A_2 + A_3 = p$ $\qquad$ $q = 1 - p$

$$P_E = 2pq + q^2$$

$$p_1 + p_2 + p_3 = p_c$$

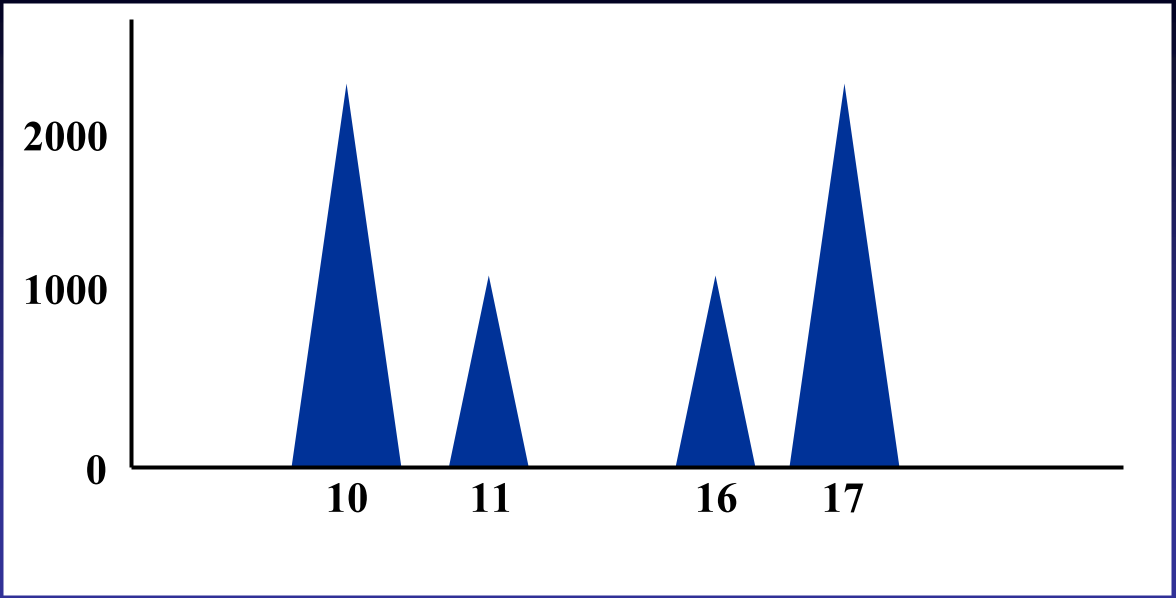


$$(p_1 + p_2 + p_3)^2 = (p_c)^2$$

$$p_1^2 + 2p_1p_2 + 2p_1p_3 + p_2^2 + 2p_2p_3 + p_3^2 = p_c^2$$

Combined Probability of Exclusion

$$P_{Ec} = 1 - (1-P_{E1})(1-P_{E2})(1-P_{En})$$



What is approach for P_E?

Even though we can interpret profiles as a 10,17 and 11,16

Likelihood Ratio

- compares the probabilities of a given observation under two different hypotheses

- a relatively subtle concept and LR computations should be explained with care

- a useful concept and used widely in statistics

Likelihood Ratio

With two (mutually exclusive) hypotheses, say H_1 and H_2, the likelihood ratio (LR) is the ratio of probabilities of observing the same data under H_1 and H_2 , giving

$$LR = Prob.\ (Data \mid H_1) / Prob.\ (Data \mid H_2).$$

Meaning of LR:

LR < 1: Data less well supported by H_1, compared with H_2

LR = 1: Data equally well supported by H_1 and H_2

LR > 1: Data better supported by H_1, compared with H_2

Assumptions for LR

- Independence
- No appreciable substructure
- All contributors of the same race
- All unrelated
- No allele dropout
- No intensity differences
- Defined hypotheses

Mixtures
Likelihood Ratio

1. H_p - prosecution hypothesis
2. H_d - defense hypothesis
3. Mutually exclusive
4. Probabilities ---- $LR = H_p/H_d$

Interpretation of mixture depends on circumstances of the case

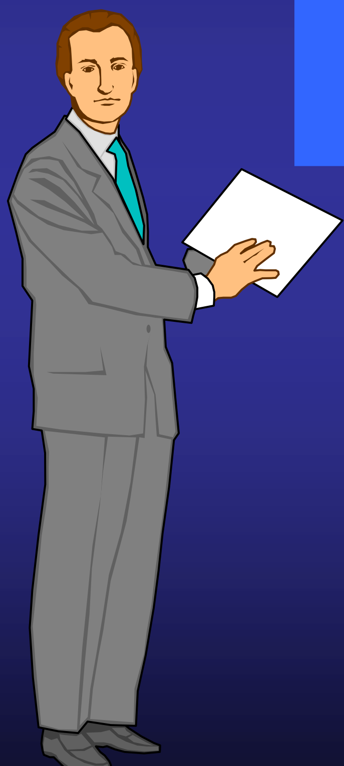
Four Allele Scenario

Four Alleles

Two Match the victim - A_3 A_4

Two Match the suspect - A_1 A_2

Mutually Exclusive Hypotheses

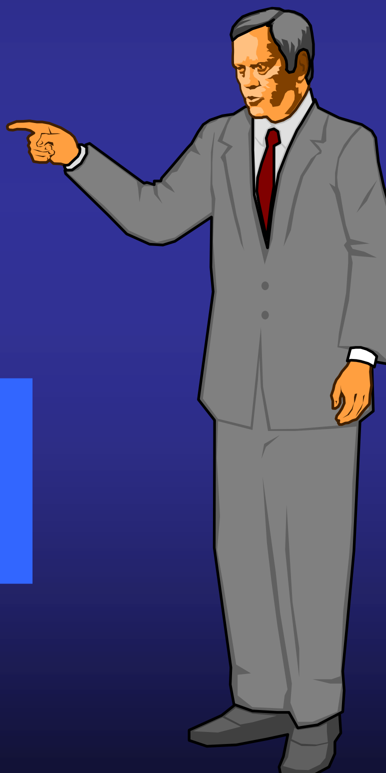


Prosecution

DNA is from victim and suspect

Defense

DNA is from victim and unknown person



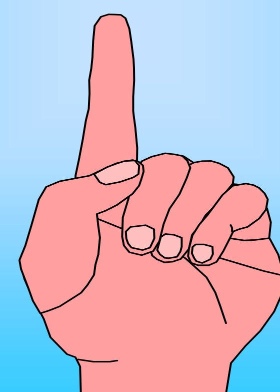
Explain the evidence by the hypotheses

H_p

1

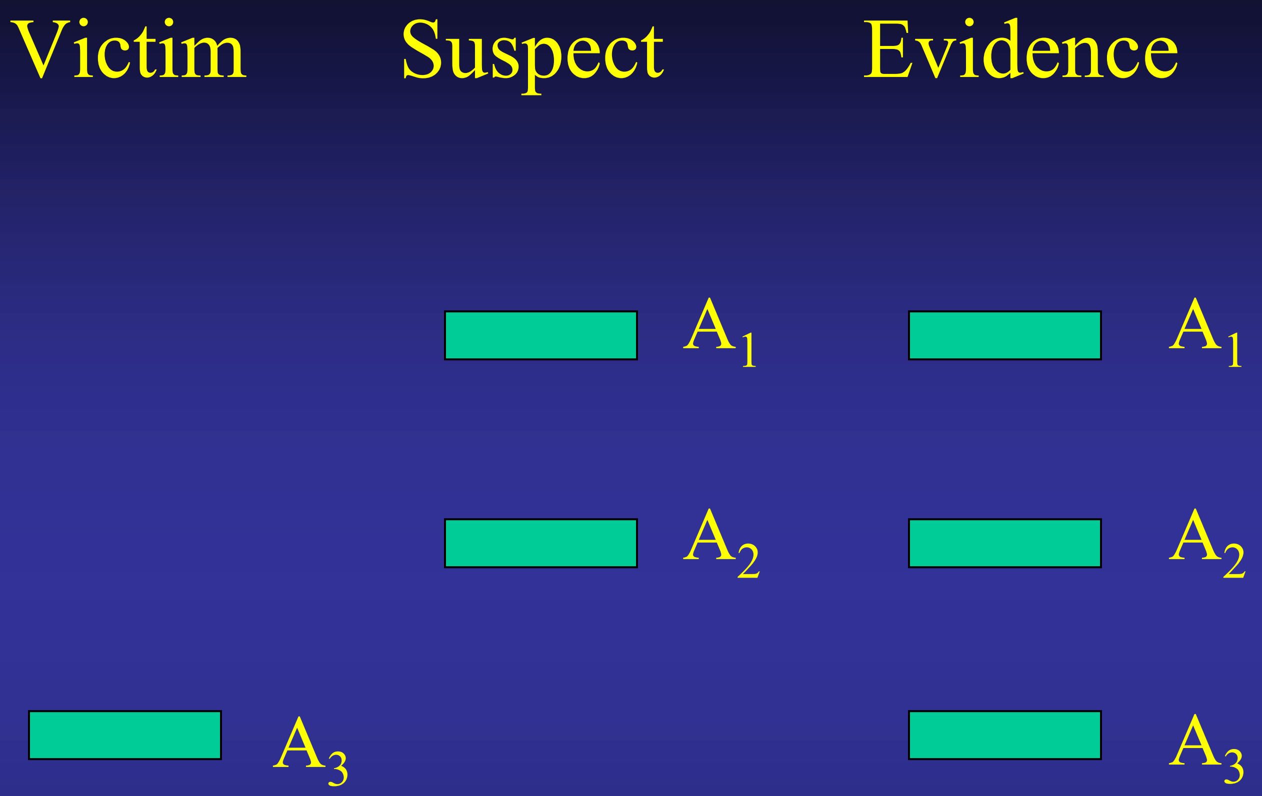
H_d

$2p_1p_2$



$$LR = H_p/H_d$$

$$1/\ 2p_1p_2$$



Three Allele Scenario

Three Alleles

Victim is homozygote - A_3A_3

Two Match the suspect - A_1A_2

Mutually Exclusive Hypotheses

Prosecution

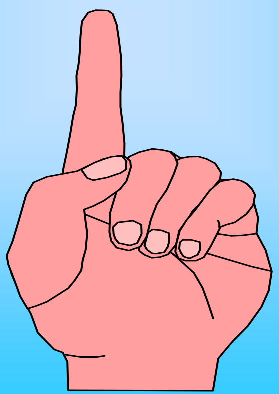
DNA is from victim
and suspect

Defense

DNA is from victim
and unknown person

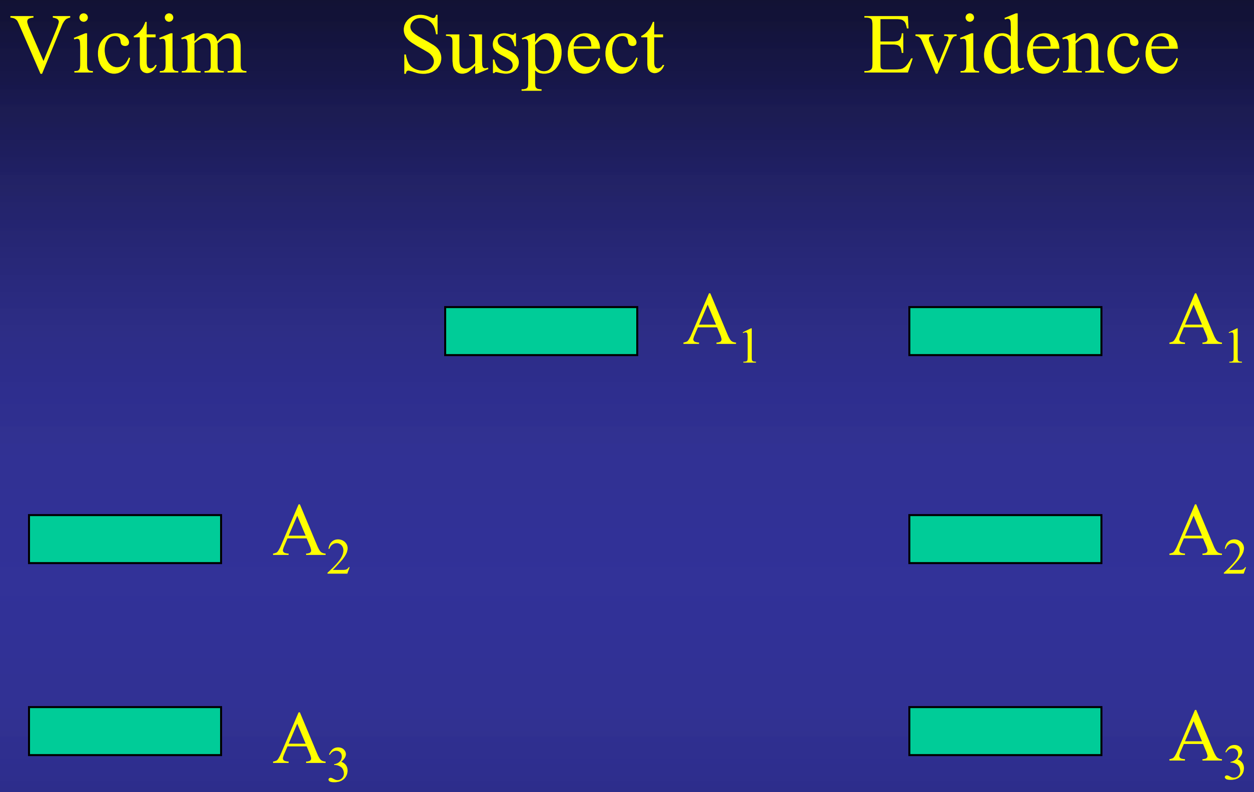
Explain the evidence by the hypotheses

H_p	H_d
1	$2p_1p_2$



$$LR = H_p/H_d$$

$$1/\ 2p_1p_2$$



Three Allele Scenario

Three Alleles

Victim is heterozygote - A_2A_3

Suspect is homozygote- A_1A_1

Mutually Exclusive Hypotheses

Prosecution

DNA is from victim and suspect

Defense

DNA is from victim and unknown person

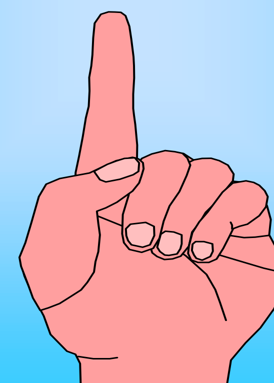
Explain the evidence by the hypotheses

H_p

1

H_d

?



Three possible genotypes
can explain the evidence

Given that the
Victim is heterozygote - A_2A_3

The possible genotypes to
explain the evidence:

A_1A_1, A_1A_2, A_1A_3

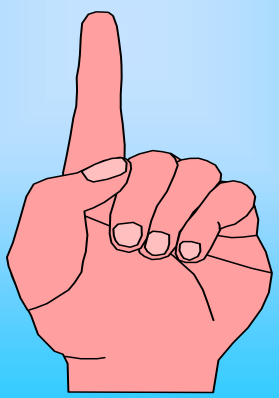
A_1A_1 A_1A_2 A_1A_3

p_1^2 $2p_1 p_2$ $2p_1 p_3$

$$P_1^2 + 2p_1 p_2 + 2p_1 p_3$$

Explain the evidence by the hypotheses

H_p	H_d
1	$p_1^2 + 2p_1 p_2 + 2p_1 p_3$



$$LR = H_p / H_d$$

$$1 / p_1^2 + 2p_1 p_2 + 2p_1 p_3$$

Victim	Suspect	Evidence	H_p	H_d
A_1A_1	A_1A_1	A_1A_1	1	p_1^2
A_1A_2	A_1A_2	A_1A_2	1	$p_1^2 + 2p_1 p_2 + p_2^2$
A_1A_1	A_2A_2	A_1A_2	1	$2p_1 p_2 + p_2^2$

HOW TO EXPRESS THE LIKELIHOOD RATIO COMPUTATIONS

- When making statements on statistical inferences remember express only the genetic data
- Not dealing with issues of "chance"
- To do so would involve Bayesian inferences which include Prior Probabilities…for which genetic data offer little assistance

Likelihood Ratio is how many times more likely it is to see this evidence under hypothesis #1 compared to hypothesis #2

$$LR = H_p/H_d$$

$$1.0/0.11 = 9.09$$

Compared with the prosecution's hypothesis (H_p), the defense scenario (H_{d1}) is

9-times less well-supported!

Alternatively

The observed mixture profile is 9-times more likely to occur under the scenario that it is a mixture of DNA from the victim and suspect, as opposed to the scenario that it originated from a mixture of DNA from the victim and an unrelated unknown person.

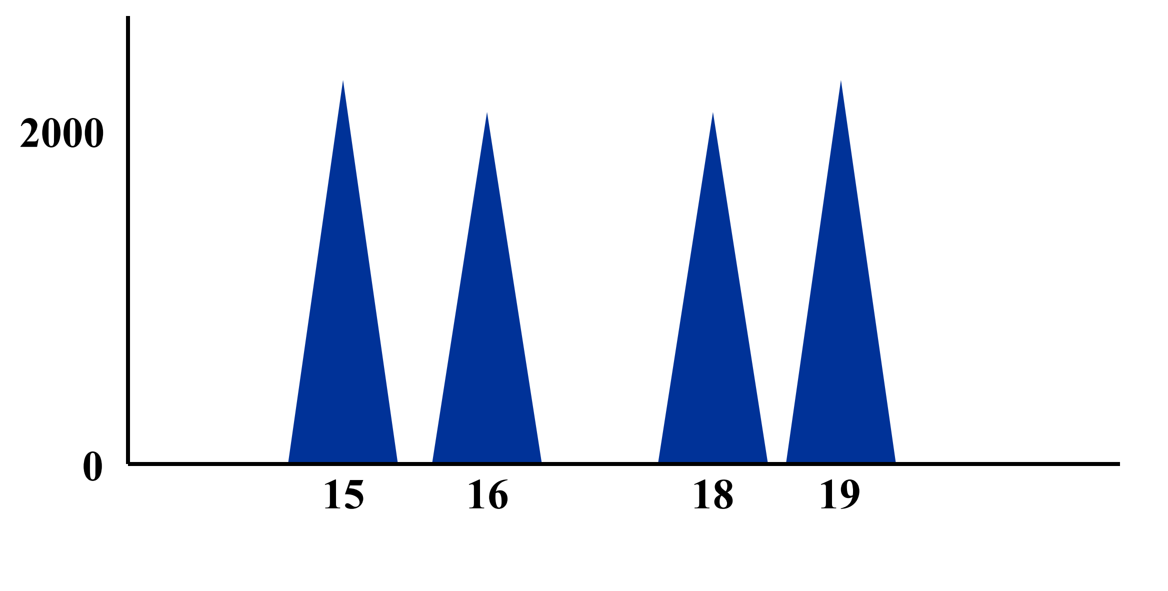
Synthesis of Three Approaches of Statistical Assessment

- Frequency approach provides the probability of the observed DNA evidence (unconditional as well as conditional) under a given stipulated hypothesis
- Likelihood Ratio (LR) contrasts such probabilities for two mutually exclusive hypotheses
- In Bayesian approach, with the use of prior probability, LR is transformed to obtain the relative odds of one hypothesis against another given the DNA data of the evidence (and that from known persons tested)

Synthesis of Three Approaches (Contd.)

- The three approaches are built on one another, and hence, it is inaccurate to say one is wrong and the others are correct
- LR, without the transformation with the use of the prior probability, may be incorrectly interpreted as the answer of the Bayesian computation, but the numerator and denominator of LR can be stated with frequentist's interpretation to avoid the error of reverse conditioning
- The prior probability of the Bayesian approach generally comes from non-DNA evidence, and hence, their assumptions are untestable from DNA data

Problems



12 …	0.010
13 …	0.025
14 …	0.070
15 …	0.215
16 …	0.230
17 …	0.190
18 …	0.150
19 …	0.070
20 …	0.030
21 …	0.010

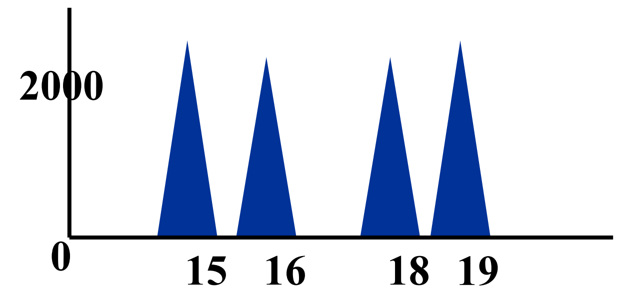
Calculate probability of exclusion for this mixture

Calculate LR for this mixture

a. Frame hypotheses

b. Two contributors – males

c. Only one suspect – type 16,19

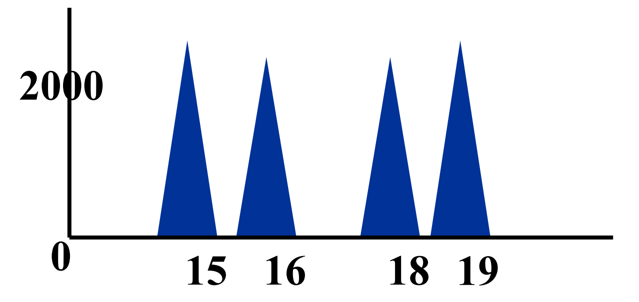


15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate probability of exclusion for this mixture

$$PE \quad = 1 - (0.215 + 0.230 + 0.150 + 0.070)^2$$

$$PE \quad = 1 - (.665)^2 = 0.558 \text{ or } 55.8\%$$



15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate LR

What are possible genotype combinations?

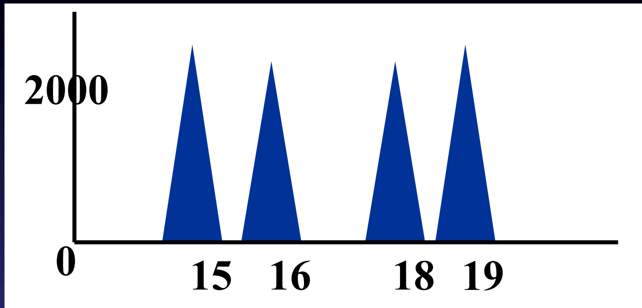
Explain the evidence
Prosecution

Suspect – 16,19
Unknown – 15,18

$$\frac{1}{2p_{15}p_{18}}$$

Defense

15,16 & 18,19
15,18 & 16,19
15,19 & 16,18
16,18 & 15,19
16,19 & 15,18
18,19 & 15,16



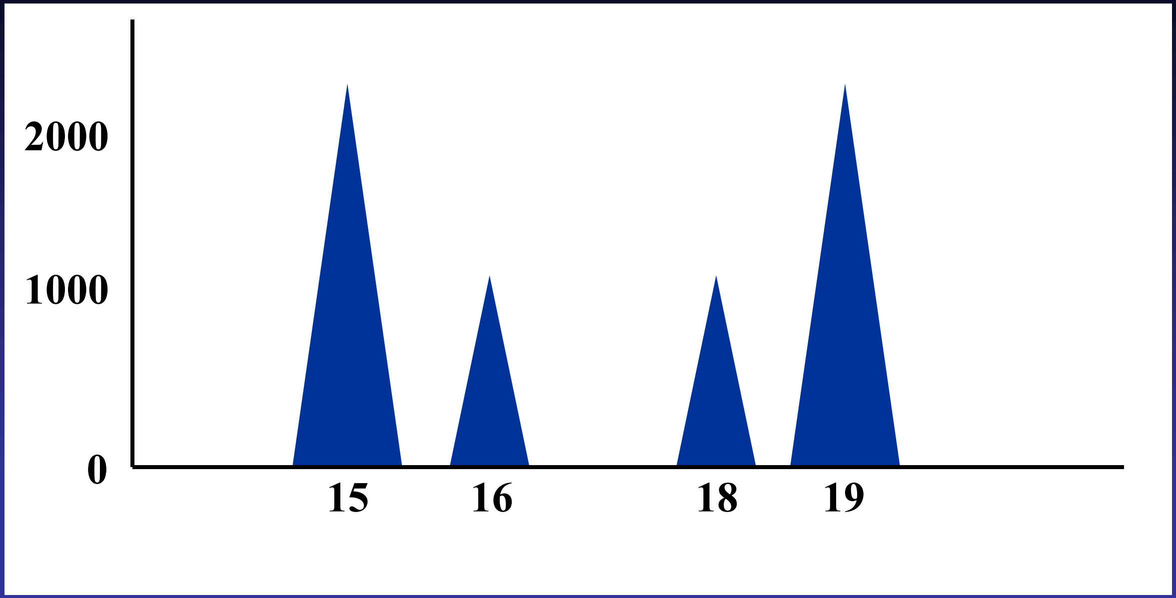
15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate LR

15,16 & 18,19	$2p_{15}\,p_{16} \times 2p_{18}\,p_{19}$
15,18 & 16,19	$2p_{15}\,p_{16} \times 2p_{18}\,p_{19}$
15,19 & 16,18	$2p_{15}\,p_{19} \times 2p_{16}\,p_{18}$
16,18 & 15,19	$2p_{16}\,p_{18} \times 2p_{15}\,p_{19}$
16,19 & 15,18	$2p_{16}\,p_{19} \times 2p_{15}\,p_{18}$
18,19 & 15,16	$2p_{18}\,p_{19} \times 2p_{15}\,p_{16}$
	$24p_{15}\,p_{16}p_{18}\,p_{19}$

$$LR = \frac{2p_{15}p_{18}}{24p_{15}\,p_{16}p_{18}\,p_{19}} = \frac{1}{12\,p_{16}\,p_{19}}$$

LR = 1/ 12 x 0.230 x 0.070 = 1/.1932 = 5.2



12 …	0.010
13 …	0.025
14 …	0.070
15 …	0.215
16 …	0.230
17 …	0.190
18 …	0.150
19 …	0.070
20 …	0.030
21 …	0.010

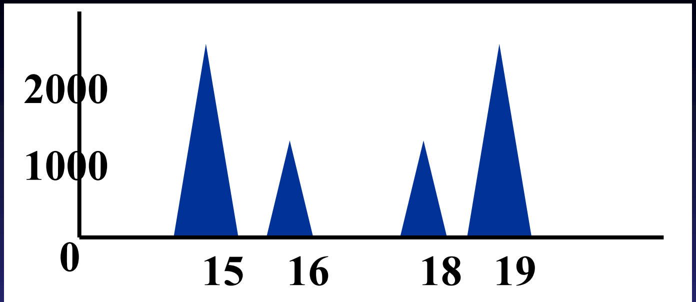
Calculate probability of exclusion for this mixture

Calculate LR for this mixture

a. Frame hypotheses

b. Two contributors

c. Victim – type 16,18; suspect – type 15,19



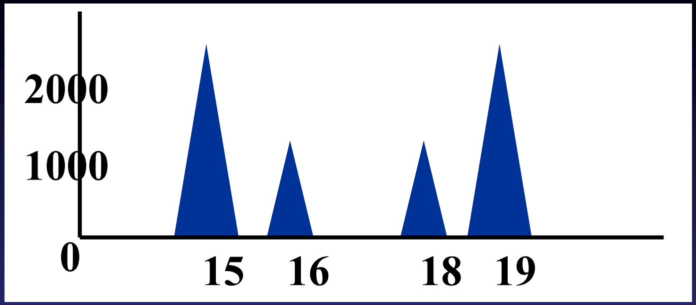
15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate probability of exclusion for this mixture

$$PE = 1 - (0.215 + 0.230 + 0.150 + 0.070)^2$$

$$PE = 1 - (.665)^2 = 0.558 \text{ or } 55.8\%$$

However, these may be treated each as single source



15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate LR
Victim – type 16,18; suspect – type 15,19

Explain the evidence

Suspect – 15,19
Victim – 16,18

1

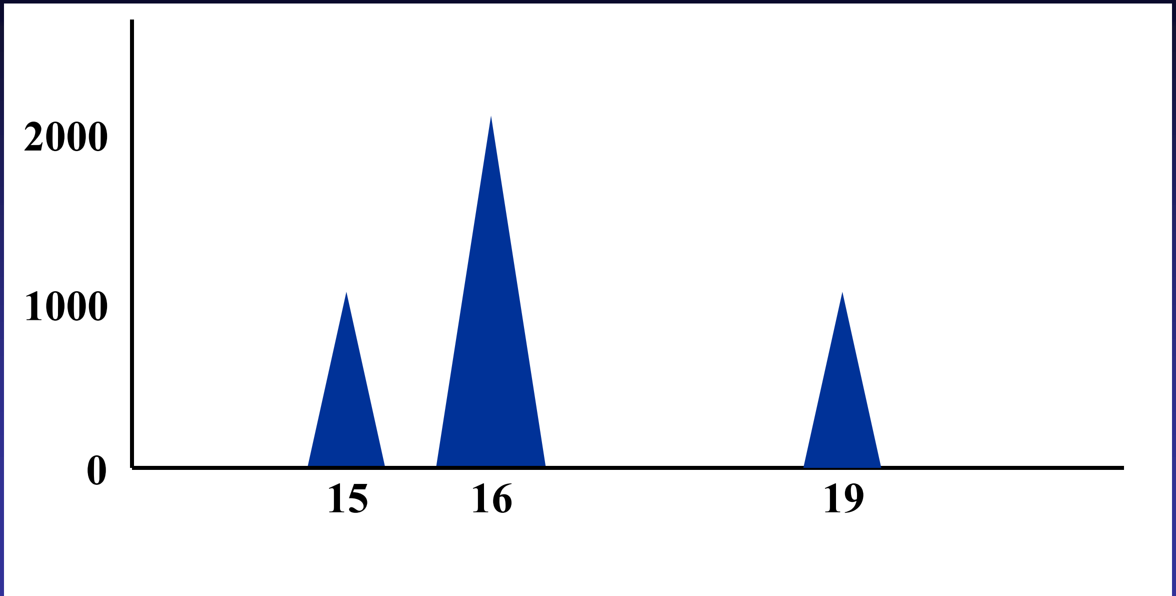
What are possible genotype combinations?

$16,18_v$ & $15,19_u$

$1 \times 2p_{15}\, p_{19}$

$$LR = \frac{1}{2\, p_{15}\, p_{19}}$$

LR = 1/ 2 x 0.215 x 0.070 = 1/.0301 = 33.2



12 … 0.010
13 … 0.025
14 … 0.070
15 … 0.215
16 … 0.230
17 … 0.190
18 … 0.150
19 … 0.070
20 … 0.030
21 … 0.010

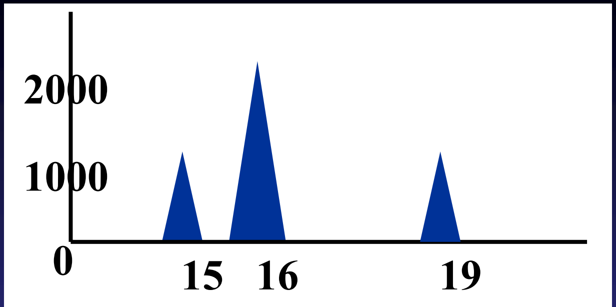
Calculate probability of exclusion for this mixture

Calculate LR for this mixture

a. Frame hypotheses

b. Two contributors

c. Victim – type 15,16; suspect – type 16,19

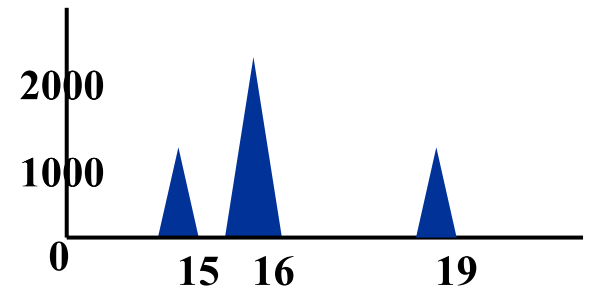


15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate probability of exclusion for this mixture

$$PE = 1 - (0.215 + 0.230 + 0.070)^2$$

$$PE = 1 - (.515)^2 = 0.735$$ or 73.5%



15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate LR
Victim – type 15,16; suspect – type 16,19

Explain the evidence

Suspect – 16,19
Victim – 15,16

1

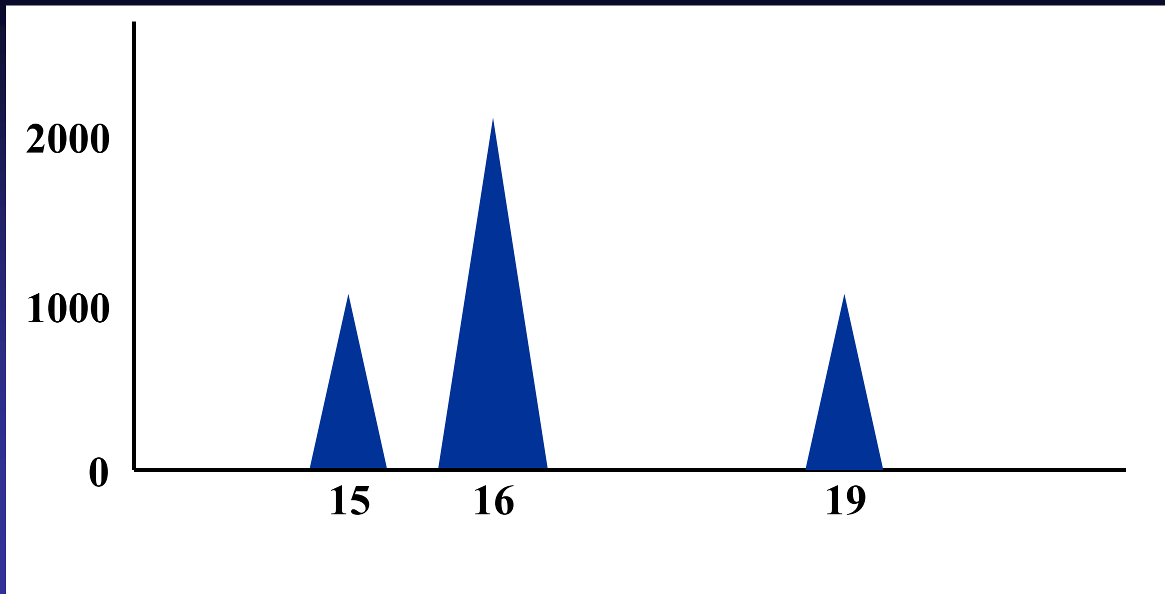
What are possible genotype combinations?

$15,16_v$ & $16,19_u$

$1 \times 2p_{16}\,p_{19}$

$$LR = \frac{1}{2\,p_{16}\,p_{19}}$$

LR = 1/ 2 x 0.230 x 0.070 = 1/.0322 = 31.1



12 … 0.010
13 … 0.025
14 … 0.070
15 … 0.215
16 … 0.230
17 … 0.190
18 … 0.150
19 … 0.070
20 … 0.030
21 … 0.010

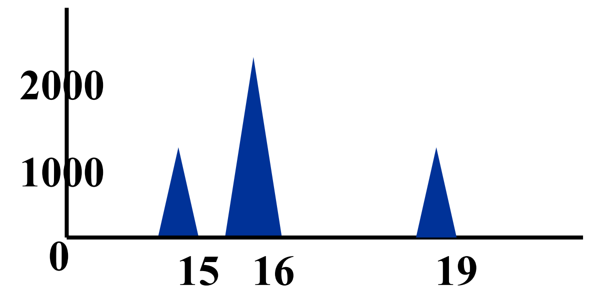
Calculate probability of exclusion for this mixture

Calculate LR for this mixture

a. Frame hypotheses

b. Two contributors

c. Victim – type 15,16; suspect – type 15,19



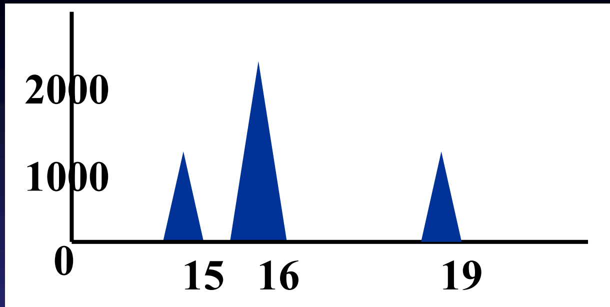
15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Calculate probability of exclusion for this mixture

Victim – type 15,16; suspect – type 15,19

$$PE = 1 - (0.215 + 0.230 + 0.070)^2$$

$$PE = 1 - (.515)^2 = 0.735 \text{ or } 73.5\%$$



Calculate LR

15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Suppose defense says DNA is from two unknown individuals

Victim – type 15,16, suspect – type 16,19

Explain the evidence

Suspect – 16,19
Victim – 15,16

1

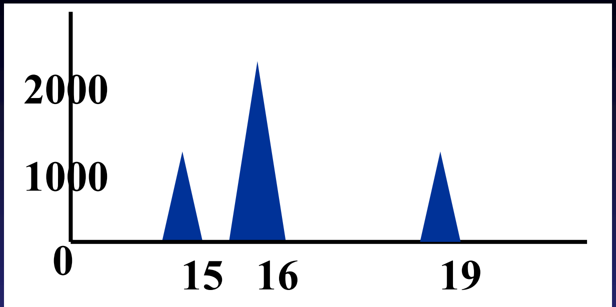
What are possible genotype combinations?

$15,16_u$ & $16,19_u$
$15,19_u$ & $16,16_u$
$16,19_u$ & $15,16_u$
$16,16_u$ & $15,19_u$

$$LR = \frac{1}{2(2p_{15}\,p_{16} \text{ x } 2p_{16}\,p_{19} + 2p_{15}\,p_{19} \text{ x } p_{16}^{2})}$$

$\mathbf{2 \text{ x } 2p_{15}\,p_{16} \text{ x } 2p_{16}\,p_{19}}$
$\mathbf{2 \text{ x } 2p_{15}\,p_{19} \text{ x } p_{16}^{2}}$

$$LR = 1/0.00954 = 104.8$$



Calculate LR

15 … 0.215
16 … 0.230
18 … 0.150
19 … 0.070

Suppose defense says the DNA is from two unknown individuals and one does not use quantitative data from electropherogram

Victim – type 15,16; suspect – type 16,19

Explain the evidence

Suspect – 16,19
Victim – 15,16

1

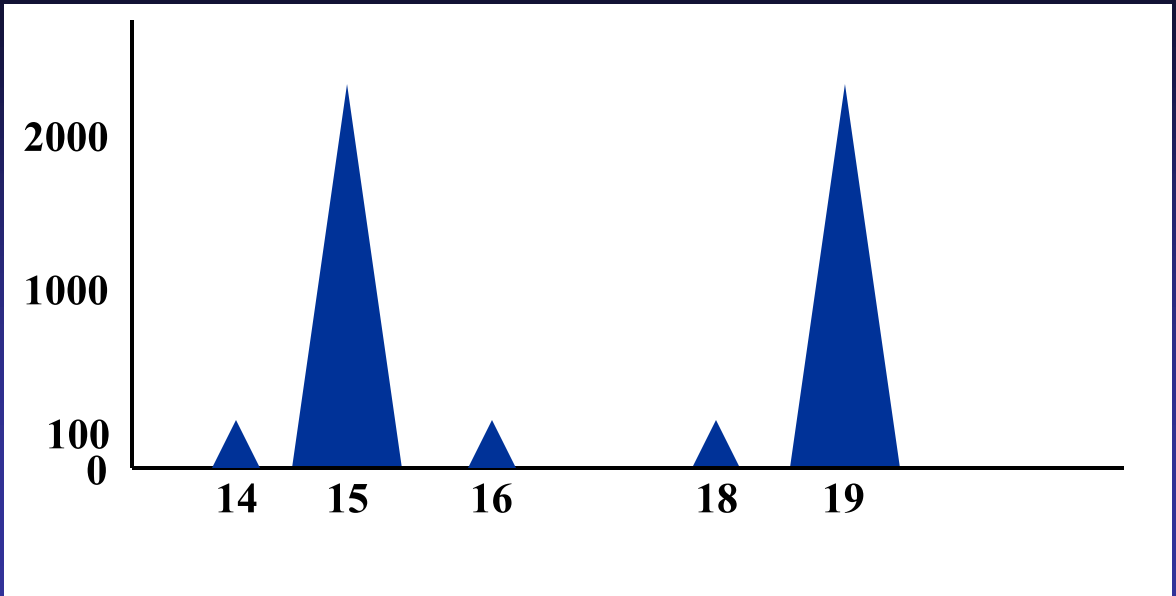
What are possible genotype combinations?

H_d

Unk 1	Unk 2	
$A_{15}A_{16}$	$A_{15}A_{19}$	$2p_{15}p_{16} \times 2p_{15}p_{19}$
$A_{15}A_{16}$	$A_{16}A_{19}$	$2p_{15}p_{16} \times 2p_{16}p_{19}$
$A_{15}A_{16}$	$A_{19}A_{19}$	$2p_{15}p_{16} \times p_{19}^2$
$A_{15}A_{19}$	$A_{15}A_{16}$	$2p_{15}p_{19} \times 2p_{15}p_{16}$
$A_{15}A_{19}$	$A_{16}A_{19}$	$2p_{15}p_{19} \times 2p_{16}p_{19}$
$A_{15}A_{19}$	$A_{16}A_{16}$	$2p_{15}p_{19} \times p_{16}^2$
$A_{16}A_{19}$	$A_{15}A_{16}$	$2p_{16}p_{19} \times 2p_{15}p_{16}$
$A_{16}A_{19}$	$A_{15}A_{19}$	$2p_{16}p_{19} \times 2p_{15}p_{19}$
$A_{16}A_{19}$	$A_{15}A_{15}$	$2p_{16}p_{19} \times p_{15}^2$

$A_{15}A_{15}$	$A_{16}A_{19}$	$p_{15}^2 \times 2p_{16}p_{19}$
$A_{16}A_{16}$	$A_{15}A_{19}$	$p_{16}^2 \times 2p_{15}p_{19}$
$A_{19}A_{19}$	$A_{15}A_{16}$	$p_{19}^2 \times 2p_{15}p_{16}$

$$12p_{15}p_{16}p_{19}(p_{15} + p_{16} + p_{19})$$



12 …	0.010
13 …	0.025
14 …	0.070
15 …	0.215
16 …	0.230
17 …	0.190
18 …	0.150
19 …	0.070
20 …	0.030
21 …	0.010

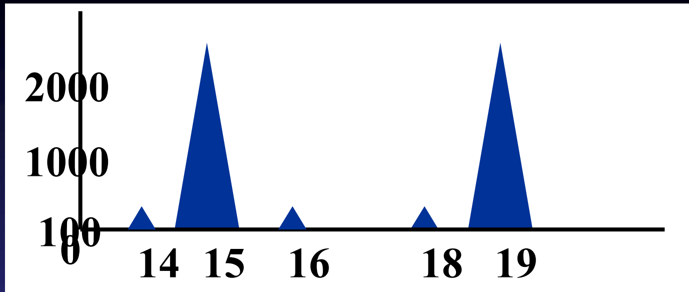
Calculate probability of exclusion for this mixture

Calculate LR for this mixture

a. Frame hypotheses

b. Two contributors

c. Victim – type 15,19; suspect – type 16,19



Calculate PE

12 … 0.010
13 … 0.025
14 … 0.070
15 … 0.215
16 … 0.230
17 … 0.190
18 … 0.150
19 … 0.070
20 … 0.030
21 … 0.010

Victim – type 15,19; suspect – type 16,19

How to interpret?

$$PE = 1 - (p_{15} + p_{16} + p_{19})^2$$

or

$$PE = 1 - (p_{14} + p_{15} + p_{16} + p_{18} + p_{19})^2$$

or

$$PE = 1 - 2p_{16}$$

or

Treat as single source - $p_{15}\, p_{19}$