Thresholds and Modeling Strategies

Thresholds

The most straightforward way to interpret an STR profile is with the use of thresholds.

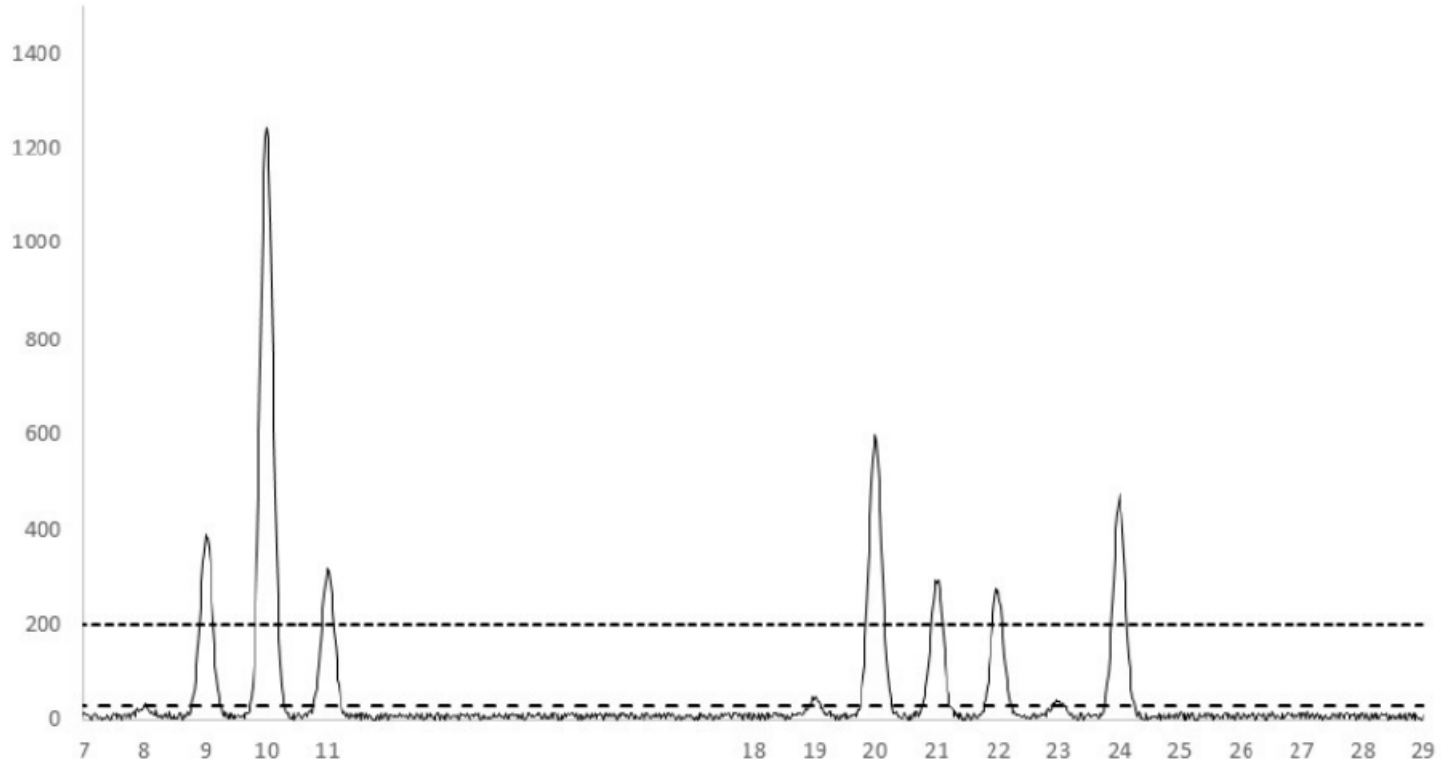
- **High thresholds**: will reduce the number of artifacts and remove a lot of background noise. However, it may potentially lead to a number of drop-outs.

- **Low thresholds**: will detect more authentic alleles, but have a higher probability of showing drop-ins.

Thresholds

An *analytical threshold* (AT) is usually set as a limit above which method response is interpreted as an authentic allele.

Additional stutter thresholds can help improve mixture profile interpretation (e.g. $5 - 15\%$ of the main allele).



Weight of Evidence

An STR profile obtained from a crime scene sample can be compared to a person of interest, and it may be found that this person cannot be excluded. An 'inclusion' may be reported, but is practically worthless without some expression on the strength of this evidence.

Likelihood Ratio

The forensic scientist is concerned with assigning a value to the likelihood ratio

$$\mathsf{LR} = \frac{\Pr(G_C|G_S, H_p, I)}{\Pr(G_C|G_S, H_d, I)},$$

which is equivalent to the reciprocal of the *profile probability* under simplified settings:

$$\mathsf{LR} = \frac{1}{\Pr(G_C|H_d, I)} = \frac{1}{p},$$

although the *match probability* is a more relevant quantity:

$$\mathsf{LR} = \frac{1}{\Pr(G_C|G_S, H_d, I)}.$$

Match Probabilities

Recall the match probabilities for homozygotes:

$$\Pr(AA|AA) = \frac{[3\theta + (1-\theta)p_A][2\theta + (1-\theta)p_A]}{(1+\theta)(1+2\theta)}$$

$$= p_A^2 \qquad (\text{if } \theta = 0),$$

and for heterozygotes:

$$\Pr(AB|AB) = \frac{2[\theta + (1-\theta)p_A][\theta + (1-\theta)p_B]}{(1+\theta)(1+2\theta)}$$

$$= 2p_A p_B \qquad (\text{if } \theta = 0).$$

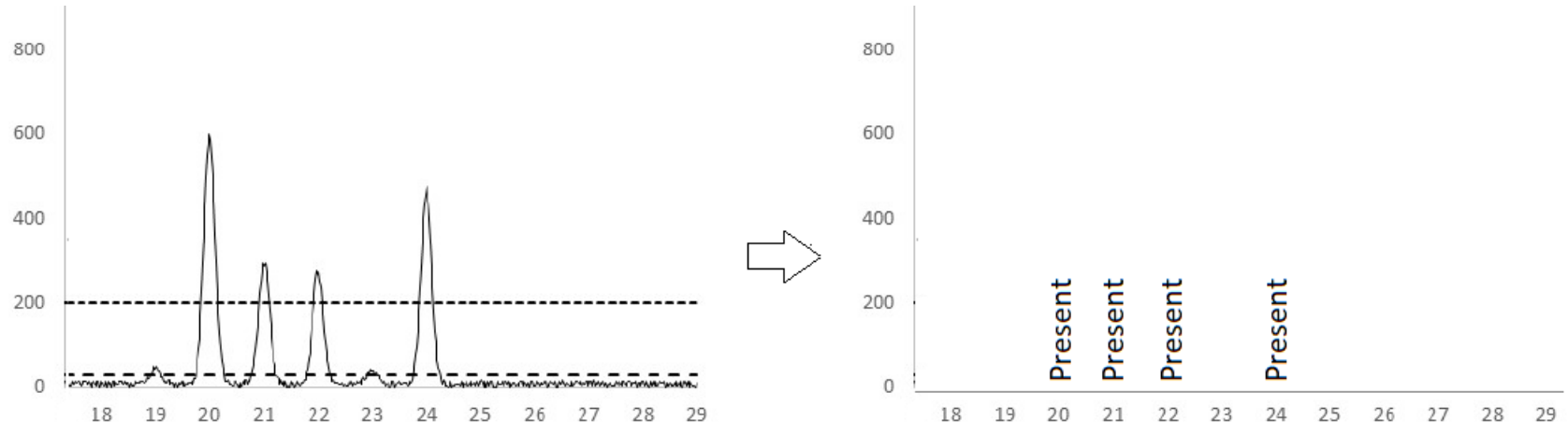
LR Modeling

Different approaches can be used to assess the likelihood ratio:

- Binary model
- Semi-continuous model
- Continuous model

Binary Model

A binary model limits interpretation of DNA profiles to qualitative allele callings only, without any attempt to infer the underlying genotypes (i.e. each are regarded as equally likely).



Single-locus LRs can be calculated and combined across loci via multiplication.

Semi-continuous Model

A semi-continuous model retains the simplicity of binary methods, but combines this with probabilistic modeling of known phenomena such as drop-ins and drop-outs.

These models will be of value when quantitative data is not available (e.g. old cases may only consist of allelic profiles).

Alleles carried by (hypothesized) contributors may not be detected in the evidence or vice versa. Drop-out and drop-in probabilities allow us to consider such situations.

Semi-continuous Model - Drop-out

For simplicity, consider a single-source profile evaluated while allowing for drop-out only in the crime scene profile G_C, as it will commonly be the stain that is of limited quantity or quality.

Two drop-out probabilities are usually considered: the probability D that an allele of a heterozygote drops out and the probability D_2 that both alleles of a homozygote drop out, with $D_2 < D^2$.

Assuming that drop-out is independent over alleles and markers, for $G_C = A$ and $G_S = AB$ the LR becomes:

$$\mathsf{LR} = \frac{\Pr(G_C|G_S, H_p)}{\Pr(G_C|G_S, H_d)} = \frac{D(1-D)}{(1-D_2)P_{AA} + D(1-D)\sum_{Q\neq A} P_{AQ}}$$

Semi-continuous Model - Drop-in

Let C denote the probability that a single allele has dropped in at a particular locus. If drop-ins at different loci are mutually independent and furthermore also independent of any drop-outs, for $G_C = A$ and $G_S = AB$:

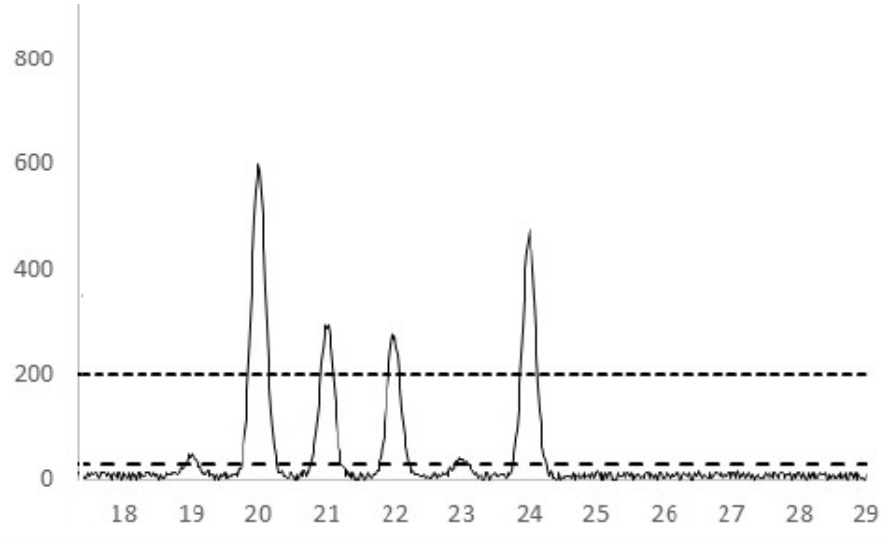
- Under H_p, $\Pr(G_C|G_S, H_p) = D(1-D)(1-C)$

- Under H_d, G_S can be AA, AQ, QQ or QQ' with $Q, Q' \neq A$ such that $\Pr(G_C|G_S, H_d) = (1-D_2)(1-C)P_{AA} + \sum_{Q \neq A}[D(1-D)(1-C)P_{AQ} + D_2 C p_A^* P_{QQ}] + \sum_{Q,Q' \neq A} D^2 C p_A^* P_{QQ'}$

Multiple drop-ins in a profile may be better interpreted as an additional (unknown) contributor.

Continuous Model

The key point of a fully continuous model is that it considers peak heights as a continuous variable.



Donor 1	Donor 2	Weights (Qualitative)	Weights (Quantitative)
20, 21	22, 24	1	0.05
20, 22	21, 24	1	0.05
20, 24	21, 22	1	0.75
21, 22	20, 24	1	0.05
21, 24	20, 22	1	0.05
22, 24	20, 21	1	0.05