The **DIP-STR** marker system for the evaluation of extremely unbalanced mixture:

Interpretative framework and comparison with other marker systems

G. Cereda, A. Biedermann, D. Hall, F. Taroni
Problem of STR primers: the DNA that contributes less than 10% to a mixture cannot be detected, since the mayor contributor's genotype masks the minor contributor's one.

Solution

A new compound genetic marker formed by a Deletion/Insertion Polymorphism (DIP) linked to a STR polymorphism (D.Hall, V.Castella, University Center of Legal Medicine Lausanne and Géneve)

Introduction

Probative value of the DIP-STR results with respect to hypotheses of interest

Comparison with STR and Y-STR, under particular assumptions
**Why STR method is not working properly for unbalanced mixtures?**

STR primers are loci-specific and not allele-specific: these primers should anneal to both the markers of the major and to those of the minor contributor.

Below the threshold of **10%** the minor contributor's DNA is generally not detected, as it is masked by the major DNA.

**With the use of primers that are allele-specific,** each time the two contributors have different genotypes in some marker, the primers will anneal to different alleles.
Primers loci specific

Only the major contributor is detected!
Primers Allele specific

Allele B of the second contributor is detected!

Major contributor

Minor contributor

Primers specific of allele B

Gene Alleles

Gene Locus

Allele A

Allele B
The novelty of DIP-STR markers

Deletion/Insertion Polymorphisms (DIPs), are created by deletion or insertion of some nucleotides in the genome.

![Diagram showing S- and L-alleles with primers]

- **S-allele (Short):** 5'-ATGC\_TTAGGGCTGGATC-3'
- **L-allele (Long):** 5'-ATGC\_TAAT\_TTAGGGCTGGATC-3'

**Allele-specificity:** one can obtain the target amplification of one allele even in presence of a high amount of the other allele.

**DIPs are di-allelic:** low discriminating power.
The novelty of DIP-STR markers

Selection of DIP markers closer than 500bp to an STR marker.

The two markers are paired to form a superlocus:

**Advantages:**

- The DIP locus allows preferential amplification of all DIP-STR haplotypes that distinguishes the minor contributor's DNA from that of the major.
- The presence of an STR in the DIP haplotype provides high resolution for the DNA profiling.
- Same genotyping technique.
DIP-STR marker - How it works

Single Locus

Analysis on the mixture

Primers L-DIP + STR:

Major DNA profile

Primers S-DIP + STR:

Minor DNA profile
Informative cases

1) Both contributors Homozygous for opposite DIP alleles (HomL/HomS)

Result on the mixtures using S primers

S4-S5

S4

2) Only minor DNA heterozygous for DIP alleles (Hom/Het)

Result on the mixtures using S primers

S5

(L4 not detected)
Uninformative cases

1) Major DNA DIP heterozygous

2) Homozygous for same DIP alleles

Better to say: Less informative!
## In summary

<table>
<thead>
<tr>
<th>Major/Minor DIP genotype</th>
<th>N of Minor’s DNA haplotypes retrieved</th>
<th>Informative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hom/Hom (different DIP allele)</td>
<td>2 (if STR het)</td>
<td>Completely</td>
</tr>
<tr>
<td>Hom/Hom (different DIP allele)</td>
<td>1 (if STR hom)</td>
<td>Completely</td>
</tr>
<tr>
<td>Hom/Het</td>
<td>1 (regardless STR)</td>
<td>Partially</td>
</tr>
<tr>
<td>Hom/Hom (same DIP allele)</td>
<td>0 (regardless STR)</td>
<td>Partially</td>
</tr>
<tr>
<td>Het/Hom</td>
<td>0 (regardless STR)</td>
<td>No</td>
</tr>
<tr>
<td>Het/Het</td>
<td>0 (regardless STR)</td>
<td>No</td>
</tr>
</tbody>
</table>

- **Informative**
- **Partially informative**
- **Uninformative**
Evaluation of evidence: the LR approach

Hp: DNA of victim and **suspect** in the mixture
Hd: DNA of victim and **unknown person** in the mixture

Evidence:

Et: DNA profiling is performed on the recovered mixture.
Eg: Victim and Suspect DIP-STR genotypes are known.

$$LR = \frac{P(Eg, Et \mid Hp)}{P(Eg, Et \mid Hd)} = \frac{P(Et \mid Hp, Eg)}{P(Et \mid Hd, Eg)} \frac{P(Eg \mid Hp)}{P(Eg \mid Hd)}$$
Evaluation of evidence: the Bayesian approach

Hp: DNA of victim and **suspect** in the mixture

Hd: DNA of victim and **unknown person** in the mixture

\[ LR = \frac{P(E_t \mid H_p, E_g)}{P(E_t \mid H_d, E_g)} \]

Object Oriented Bayesian Network
SECOND PART: Comparison with STR and Y-STR

• Classical approach to extremely unbalanced mixtures:

Stain is recovered → STR: No evidence of a mixture

Otherwise → Y-STR:

Wrong evaluation → Impossible to know in advance

If one presumes mixture → Impossible to know in advance

Major female minor male

No results

Minor’s contributor’s Y-STR genotype

This can mean that there is no mixture or that the minor is not male!
SECOND PART: Comparison with STR and Y-STR

- **Novel** approach to **extremely unbalanced mixtures**:

Stain is recovered

The major is DIP-homozygous (S-S or L-L)

DIP-STR

Partial or complete minor genotype

No results at the specific marker considered

The major is DIP-heterozygous at

Possible to know in advance

P=6.12×10^{-5}

Using the allelic proportion of 9 DIP-STR marker in Suisse population

In order to have no results at all, major should be heterozygothe in all markers
Comparison of the LR distributions

100000 mixtures are simulated and the LR regarding their results are collected

Compare distribution of the LR for the obtained results

- Moderately unbalanced mixtures
  - DIP-STR vs STR
  - DIP-STR vs Y-STR
  - No comparison (only DIP-STR available)

- Female-male mixtures

- All other situations:

Point of view of the prosecution

Mixture=V+U

Point of view of the defence

Mixture=V+S
Comparison of the LR distributions

Under the prosecution’s point of view

$Hp: \text{Mixture} = V+S$

- $V \\ S$
- $V \rightarrow \text{Mixture}$
- $S \rightarrow \text{Mixture}$

$LR > 1$

LR in favour of $S$ being a contributor

The higher the LR the more interesting is the chosen method from the prosecution’s point of view.

Under the defence’s point of view

$Hd: \text{Mixture} = V+C2$

- $V \\ C2 \\ S$
- $V \rightarrow \text{Mixture}$
- $C2 \rightarrow \text{Mixture}$
- $S \rightarrow \text{Mixture}$

$LR > 1$

LR in favour of $S$ being a contributor

The higher the number of zero likelihood ratios which are obtained, the better is the attractive from the defence’s point of view.
**DIP-STR vs STR**

### Under the prosecution’s point of view

<table>
<thead>
<tr>
<th>Marker system</th>
<th>Min</th>
<th>1st Quantile</th>
<th>Median</th>
<th>Mean</th>
<th>3rd Quantile</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>DIP</td>
<td>0</td>
<td>2.228</td>
<td>3.201</td>
<td>3.367</td>
<td>4.324</td>
<td>13.706</td>
</tr>
</tbody>
</table>

Histograms of the distributions of log10(LR) for STR and DIP-STR

(moderately unbalanced mixtures)

**Better form the prosecution’s point of view**

### Under the defence’s point of view

<table>
<thead>
<tr>
<th>LR&lt;sub&gt;d&lt;/sub&gt;</th>
<th>Verbal equivalent in favor of Hd</th>
<th>DIP-STR markers</th>
<th>STR markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Exclusion</td>
<td>99.988%</td>
<td>100%</td>
</tr>
<tr>
<td>1-10</td>
<td>Limited support</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>10-100</td>
<td>Moderate support</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>100-1000</td>
<td>Moderately strong support</td>
<td>0.003%</td>
<td>0%</td>
</tr>
<tr>
<td>1000-10000</td>
<td>Strong support</td>
<td>0.007%</td>
<td>0%</td>
</tr>
<tr>
<td>&gt;10000</td>
<td>Very strong support</td>
<td>0.002%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Better form the defence’s point of view

**Tables 8 summarises the percentage of DIP-STR and STR LR values found for various intervals of probative strength.**

<table>
<thead>
<tr>
<th>LR&lt;sub&gt;d&lt;/sub&gt;</th>
<th>Verbal equivalent in favor of Hd</th>
<th>DIP-STR markers</th>
<th>STR markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Exclusion</td>
<td>99.988%</td>
<td>100%</td>
</tr>
<tr>
<td>1-10</td>
<td>Limited support</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>10-100</td>
<td>Moderate support</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>100-1000</td>
<td>Moderately strong support</td>
<td>0.003%</td>
<td>0%</td>
</tr>
<tr>
<td>1000-10000</td>
<td>Strong support</td>
<td>0.007%</td>
<td>0%</td>
</tr>
<tr>
<td>&gt;10000</td>
<td>Very strong support</td>
<td>0.002%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Under the prosecution’s point of view: better form for DIP-STR markers.

Under the defence’s point of view: better form for STR markers.

**Figure 32 shows the Tippett plots of the DIP-STR likelihood ratio distribution and the 3 different number of markers in the two kits, can be made in the case here. But even if one chooses the 9 STR markers which have more non-zero values (D8, D3, D1S, D12, VWA, D2S1, D18, FGA, D2S4) and we multiply them to the di markers (for balanced mixtures), because the proportion of likelihood ratio values with 0 is maximal. In principle, the ratio distributions with 9 STR markers are closer to the DIP-STR likelihood ratio distribution than the one with 16.**
DIP-STR vs Y-STR

Under the prosecution’s point of view

Conservative with respect to the strength of support obtained with values >0

Higher rate of LRs that would wrongly associate a suspect with a mixture. However, the LR for such cases would be more moderate than in case of the DIP-STR method.

Under the defence’s point of view

Conservative with respect to the number of times

Higher number LR=0, but with some possibility of a high LR against a suspect who has a genotype compatible with a mixture to which he is not a contributor.

Histograms of the distributions of log10(LR) for Y-STR and DIP-STR

(female-male mixtures)
Discussion and Conclusion

- DIP-STR is a novel technique with a high utility in the field of DNA unbalance mixture.
- DIP-STR results are of less impact if not accompanied with an interpretative model that permits to quantify the probative value of these results in the light of two competing hypotheses of interest (Hp and Hd).
- LR is a way to quantify the probative value.
- The OOBN built to interpret and evaluate the DIP-STR method is an intuitive and user-friendly means that permits to obtain the LR without any calculation effort.
COMPARISON

• Compare DIP-STR profiling technique with the classical STR profiling technique, used for “good” mixtures.
<table>
<thead>
<tr>
<th>Scenario</th>
<th>Prosecution’s point of view</th>
<th>Defence’s point of view</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moderately unbalanced mixtures:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STR vs DIP-STR</td>
<td>STR</td>
<td>STR</td>
</tr>
<tr>
<td>Extremely unbalanced mixtures +</td>
<td>DIP-STR</td>
<td>Y-STR or DIP-STR</td>
</tr>
<tr>
<td>Good gender combination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y-STR vs DIP-STR</td>
<td>DIP-STR</td>
<td></td>
</tr>
<tr>
<td>All other situations</td>
<td>DIP-STR</td>
<td></td>
</tr>
<tr>
<td>Only DIP-STR</td>
<td>DIP-STR</td>
<td></td>
</tr>
</tbody>
</table>
Impossible to know in advance in which situation we are

Moderately unbalanced mixtures:

Extremely unbalanced mixtures + Good gender combinations

All other situations
The use of DIP-STR markers could present an interest for all kind of DNA stains, independently of the use of STR markers.

DIP-STR markers can establish in advance if this method could be used.

LR distributions obtained under the defence's and the prosecution's point of view are not as marked as those of traditional STR markers.

New DIP-STR markers are currently researched, which may improve the likelihood ratio distributions under the various competing points of view in a near future.

Time and monetary constraint should be considered, too.