Increased Efficiency of Forensic Y-STR Analysis with the New PowerPlex® Y23 System

Lutz Roewer
Dept. Forensic Genetics
Institute of Legal Medicine and Forensic Sciences
Charité – Universitätsmedizin Berlin
Applications of Y chromosome analysis

- Crime cases (mixed stain analysis)
- "Familial search" (USA) (Verification of near matches)
- Disaster victim identification (patrilinear relation)
- Prediction of origin (geographical, ethnic)
- Kinship testing (patrilinear relationship)
- Genealogical research
- Reconstruction of human history
- Archeogenetics
Why is a high resolution of Y-STR analysis so important?

- homicide case in Berlin in 2003
- spurious male DNA from the fingernails of the female victim (mixed stain)
- mass screening (558 males compared)
- one man, person A, matched (12 loci, Powerplex Y), $f = 5.7 \times 10^{-5}$, in Africa: $1.7 \times 10^{-3}$ (YHRD)
- this man was excluded with 1 of 9 additional Y-STRs
- one year later the true perpetrator, person B, was arrested and he confessed
- complete 21/21 match of the Y-STR profiles
- the man is a Caribbean of African origin

---

<table>
<thead>
<tr>
<th>Loci</th>
<th>DYS19</th>
<th>389I</th>
<th>389II</th>
<th>390</th>
<th>391</th>
<th>392</th>
<th>393</th>
<th>385ab</th>
<th>438</th>
<th>439</th>
<th>437</th>
<th>446</th>
<th>447</th>
<th>448</th>
<th>449</th>
<th>463</th>
<th>464</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fingernails</td>
<td>14</td>
<td>12</td>
<td>28</td>
<td>25</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>14,14</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>19</td>
<td>23</td>
<td>19</td>
<td>32</td>
<td>21</td>
<td>15,16</td>
</tr>
<tr>
<td>Person A</td>
<td>14</td>
<td>12</td>
<td>28</td>
<td>25</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>14,14</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>20</td>
<td>23</td>
<td>19</td>
<td>32</td>
<td>21</td>
<td>15,16</td>
</tr>
<tr>
<td>Person B</td>
<td>14</td>
<td>12</td>
<td>28</td>
<td>25</td>
<td>11</td>
<td>11</td>
<td>13</td>
<td>14,14</td>
<td>11</td>
<td>11</td>
<td>15</td>
<td>19</td>
<td>23</td>
<td>19</td>
<td>32</td>
<td>21</td>
<td>15,16</td>
</tr>
</tbody>
</table>

Exclusion in only 1 of 21 Y-STR loci!
Discrimination capacity (DC) in different YHRD population sets (release 37), 7-17 loci
Search for Y-STRs with extremely high mutation rates (Hypermutable Y-STRs or rapidly mutating Y-STRs)

Complete individualisation with RM Y-STRs possible?

(Ballantyne et al. 2010)
Improving global and regional resolution of male lineage differentiation by simple single-copy Y-chromosomal short tandem repeat polymorphisms

Mark Vermeulen a,1, Andreas Wollstein a,b,1, Kristiaan van der Gaag c, Oscar Lao a, Yali Xue d, Qiuju Wang d,e, Lutz Roewer f, Hans Knoblauch g, Chris Tyler-Smith d, Peter de Knijff c, Manfred Kayser a,*

a Department of Forensic Molecular Biology, Erasmus University Medical Center, Rotterdam, The Netherlands
b Cologne Center for Genomics and Institute of Genetics, University of Cologne, Germany
c Forensic Laboratory for DNA Research, Department of Human and Clinical Genetics, Leiden University Medical Center, The Netherlands
d The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK
e Department of Otolaryngology, Head & Neck Surgery, Chinese PLA General Hospital, Beijing, China
f Abteilung für Forensische Genetik, Institut für Rechtsmedizin und Forensische Wissenschaften, Charité - Universitätsmedizin Berlin, Germany

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
<table>
<thead>
<tr>
<th>Locus</th>
<th>Minimal Haplotype (9)</th>
<th>Powerplex Y (12)</th>
<th>Yfiler (17)</th>
<th>Powerplex Y23</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS19</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS389I</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS389II</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS390</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS391</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS392</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS393</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS385ab</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS437</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS438</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS439</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>DYS448</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS456</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS458</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS635</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>YGATAH4</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS481</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS533</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS549</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS570</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS576</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DYS643</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

New markers to close the gap between autosomal and Y chromosomal STR analysis?
Different patterns of inheritance

**Autosomal Markers**

13 CODIS STR Loci

22 pairs of autosomes (passed on in part, from all ancestors)

**Lineage Markers**

12 or 17 Y-STRs

Y-Chromosome (passed on complete, but only by sons)

mtDNA control region

Mitochondrial (passed on complete, but only by daughters)
The new loci: DYS 570

- TTTC
- Average repeat number: 17.6
- Number of alleles: 11
- Diversity: 0.86
- Mutation rate $1.2 \times 10^{-2}$


©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
DYS 576

- AAAG
- Average repeat number: 17.3
- Number of alleles: 7
- Diversity: 0.82
- Mutation rate $1.4 \times 10^{-2}$

DYS 481

- CTT
- Average repeat number: 23.3
- Number of alleles: 11
- Diversity: 0.9
- Mutation rate: $4.4 \times 10^{-3}$

DYS 643

- CTTTT
- Average repeat number: 11.1
- Number of alleles: 9
- Diversity: 0.82
- Mutation rate: $4.7 \times 10^{-4}$

DYS 533

- TATC
- Average repeat number: 11.2
- Allele number: 6
- Diversity: 0.72
- Mutation rate: $2.7 \times 10^{-3}$

DYS 549

- GATA
- Average repeat number: 12.1
- Number of alleles: 5
- Diversity: 0.72
- Mutation rate: $1.4 \times 10^{-3}$

Allelic Ladder Y23

Overlapping size ranges possible

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
<table>
<thead>
<tr>
<th>Locus</th>
<th>Length alleles (YHRD r39)</th>
<th>Allelic ladder Y23</th>
</tr>
</thead>
<tbody>
<tr>
<td>DYS19</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>DYS389I</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>DYS389II</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>DYS390</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>DYS391</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>DYS392</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>DYS393</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>DYS385ab</td>
<td>45</td>
<td>50</td>
</tr>
<tr>
<td>DYS437</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>DYS438</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>DYS439</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>DYS448</td>
<td>26</td>
<td>11</td>
</tr>
<tr>
<td>DYS456</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>DYS458</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td>DYS635</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>YGATAH4</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>DYS481</td>
<td>-</td>
<td>18</td>
</tr>
<tr>
<td>DYS533</td>
<td>-</td>
<td>12</td>
</tr>
<tr>
<td>DYS549</td>
<td>-</td>
<td>13</td>
</tr>
<tr>
<td>DYS570</td>
<td>-</td>
<td>21</td>
</tr>
<tr>
<td>DYS576</td>
<td>-</td>
<td>15</td>
</tr>
<tr>
<td>DYS643</td>
<td>-</td>
<td>30</td>
</tr>
</tbody>
</table>

Missing intermediate alleles
The hexameric locus DYS448 has frequent intermediate length alleles

**Table Of Non-Uniform Alleles**

<table>
<thead>
<tr>
<th>N</th>
<th>Allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>10.2</td>
</tr>
<tr>
<td>1</td>
<td>10.2-10</td>
</tr>
<tr>
<td>24</td>
<td>16.2</td>
</tr>
<tr>
<td>1</td>
<td>17.1</td>
</tr>
<tr>
<td>2</td>
<td>17.2</td>
</tr>
<tr>
<td>1</td>
<td>17.2-19.20</td>
</tr>
<tr>
<td>2</td>
<td>19.20</td>
</tr>
<tr>
<td>1</td>
<td>18.1</td>
</tr>
<tr>
<td>3</td>
<td>19.0</td>
</tr>
<tr>
<td>2</td>
<td>18.8</td>
</tr>
<tr>
<td>21</td>
<td>19.20</td>
</tr>
<tr>
<td>6</td>
<td>19.21</td>
</tr>
<tr>
<td>14</td>
<td>19.2</td>
</tr>
<tr>
<td>1</td>
<td>19.3</td>
</tr>
<tr>
<td>1</td>
<td>19.4</td>
</tr>
<tr>
<td>5</td>
<td>20.21</td>
</tr>
<tr>
<td>1</td>
<td>20.22</td>
</tr>
<tr>
<td>1</td>
<td>20.1</td>
</tr>
<tr>
<td>2</td>
<td>20.2</td>
</tr>
<tr>
<td>2</td>
<td>20.3</td>
</tr>
<tr>
<td>2</td>
<td>20.4</td>
</tr>
<tr>
<td>2</td>
<td>21.2</td>
</tr>
</tbody>
</table>

Frequency in 41,000 alleles (r.39 der YHRD from 10.2.12)
Problematic systems, e.g. DYS19 improved

Powerplex Y23
307-348 bp

YFiler
173-209 bp
Crime scene evidence (mixed, LT DNA)

Autosomal analysis

Mixed female/male profile

Autosomal male profile clearly identifiable

Report

National Police Database

Y chromosomal analysis

Male profile

Database (YHRD) frequency

Direct matching possible?

Report

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
Charité Pilot study to evaluate the workflow: 40 sexual crimes with only „touch“ DNA (2011)

40 cases

- 8 cases
  - Individual male autosomal profile

- 10 cases
  - Informative YSTR Profile

- 22 cases
  - Inconclusive

Report with YHRD match statistics

Number of informative profiles doubled

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
Gain of information

NGMSElect – 3 : 1

YFiler – 2 alleles called

Powerplex Y23 – 7 alleles
Haplotype with 7 alleles: -,12,30,22,11,-,-,-,16,-,-,16,20,-
f = 7.3 \times 10^{-5} \ (1.5 \times 10^{-5} - 2.1 \times 10^{-4})$, direct matching possible
Now 100,000 Haplotypes in the YHRD (r.39)
The adapted YHRD search mask for the Y23 panel

Please note: The database size will vary based on the loci you have entered.

- 7 loci haplotype (DYS19, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393): 9881 haplotypes
- 9 loci haplotype (+ DYS385a/b): 36084 haplotypes
- 11 loci haplotype (+ DYS438, DYS439): 70997 haplotypes
- 12 loci haplotype (+ DYS437): 51454 haplotypes
- 17 loci haplotype (+ DYS448, DYS456, DYS458, DYS635, YGATH4): 39339 haplotypes

Y-SNPs:
- 121 Y-SNP branches (defined by 131 Y-SNP markers)
- 8506 haplotypes with Y-SNP information
Evaluation of the „PowerPlex® Y 23 Systeme“ prototype (Promega Corp.)

- Alpha test series -
  (October/November 2011)
The 23 Y-STR loci of the PowerPlex® Y 23 system

Composition:
→ 12 loci of PowerPlex® Y
→ *5 loci of AmpFEL®STR® YFiler™
→ *6 new loci (Vermeulen et al. 2009)
  → with 2 RM-YSTRs:
    • DYS570 (1,24x10^{-2})
    • DYS576 (1,43x10^{-2})

Characteristics
• High diversity of single markers (0,75-0,92)
• Increased discrimination capacity of the haplotype

Alpha test

I. Sensitivity study
II. Mixture study
III. Casework study

- Reference: AmpF™STR® YFiler™ Kit (Applied Biosystems)
- 23 Y-STRs (17 known from the AmpF™STR® YFiler™ + 6 simple single-copy Y-STR marker (ssYSTR))
I. Sensitivity study

• 5 dilutions of male genomic DNA (31.2 pg - 500 pg)
• Repeat analysis

• Analysis criteria:
  • concordance
  • sensitivity
  • Peak balance within the multiplex-PCR

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
I. Sensitivity study

<table>
<thead>
<tr>
<th>DNA amount</th>
<th>Allele calling PowerPlex® Y23</th>
<th>Allele calling AmpFISTR® YFiler™</th>
</tr>
</thead>
<tbody>
<tr>
<td>500pg</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>250pg</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>125pg</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>62.5pg</td>
<td>On average 1 allelic drop-out</td>
<td>Partial profile + high background</td>
</tr>
<tr>
<td>31.2pg</td>
<td>on average up to 4 allelic drop-outs with increased background</td>
<td>No profile</td>
</tr>
</tbody>
</table>

Observations:

✓ concordance between the kits
✓ PowerPlex® Y 23 System is significantly more sensitive
I. Sensitivity study

Observations:

- more sensitive (on average doubled peak-sizes)
- improved balance (especially for the new 6 systems)
II. Mixture study

i. variable male DNA in 5 dilutions (c♂ = 31,2pg – 500 pg) and constant female DNA (c♀=400ng)

ii. constant male DNA (c♂=500pg) and variable female DNA (c♀ = 400 – 10ng)

• Analyses in replicate

• Analysis criteria:
  • Conkordance
  • Sensitivity
  • Influence of excessive amounts of female DNA
  • Peak balance within the multiplex-PCR
Excess of ♀ DNA has no influence on the ♂ profile. High sensitivity & specificity retained.

**PowerPlex®Y 23**

- Clean male profile with 6400 fold excess of ♀ DNA

**AmpFLESTR® YFiler™**

- 62.5 pg ♂ DNA + 400 ng ♀ DNA
- Clean male profile with 3200 fold excess of ♀ DNA

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
III. Casework study

131 difficult „touch DNA“ stains of sexually motivated crime cases (mostly unbalanced female/male mixtures)

Analysis criteria:
- sensitivity
- peak balance
- assessment of all peaks with an area >500 rfu (cut-off)
- Reference AmpF/STR YFiler™ kit (Applied Biosystems)
Case example 1: Contact stain (T-Shirt, chest area, swab)

**PowerPlex® Y 23**

23 out of 23 systems successfully amplified

**AmpFLESTR® YFiler™**

12 / 17 systems successfully amplified
Casework example 2: contact stain (skirt, swab)

**PowerPlex® Y 23**

Observations:
- additional peaks → sensitivity increased, additional donor or „drop in“?
- drop-outs with Powerplex® Y compared to AmpFELSTR® YFiler™

21 out of 23 systems called

**AmpFELSTR® YFiler™**

17 / 17 called
Casework example 3: contact stain (skin, swab)
Evaluation of casework stains

Observation:

✓ On average less drop-outs with Powerplex Y23
  - Median Y23 = 35% (~8 of 23 alleles)
  - Median YFiler = 12% (~2 of 17 alleles)
✓ sensitive → cut-off value increased
Evaluation of single Y-STR markers

• “Difficult“ systems of the AmpF®STR® YFiler™ are DYS19, DYS456, DYS439 and Y-GATA-H4

Observations:
✓ DYS19 & DYS439 in PowerPlex® Y improved
• Still problematic: DYS456 & Y-GATA-H4 in Powerplex® Y and AmpF®STR® YFiler™
• Large fragments drop out first
Mutation study (preliminary)

- 27 males in 18 genealogical trees, 46 meioses

Observation:
- 2 mutations in 46 meioses
DYS576 → 1 mutation in 46 meiotic transfers

- DYS576 mutation rate: $1.43 \times 10^{-2}$ (Ballantyne et al. 2010)
DYS549 $\rightarrow$ 1 mutation in 46 meiotic transfers

DYS549 mutation rate: $1.38 \times 10^{-3}$
(Vermeulen et al. 2009)
Discussion of the Y23 alpha test

- Expansion by 6 highly variable loci
  - Increased discrimination → less coincidental matches
- Improvement of primer design
  - Increased sensitivity
  - Much improved peak balance among markers in the multiplex
  - Higher cut-off values possible
  - Optimization of difficult systems (e.g. DYS19) of the AmpF\$STR® YFiler™ kit
- Rapidly mutating systems (RM-YSTRs)
  - increased differentiation of patrilineages
- shorter PCR time

✓ concordance between AmpF\$STR® YFiler™ and PowerPlex®Y 23 allele calls
✓ Excess of female DNA has no influence on sensitivity and specificity of the male profiling

Problems of PowerPlex®Y 23:
- Still low-performing systems: Y-GATA-H4 und DYS456
- Drop-in observed in LT DNA analysis
Outlook

- Generation of population databases for the 23-loci haplotype format
- Supply of frequency data
- Adaptation of existing Y-STR Haplotype databases (YHRD, U.S. Consolidated Y-STR Database), inclusion of new loci in the search mask (socket 3,000 haplotypes)

www.yhrd.org

©Charité – Universitätsmedizin Berlin, Dept. Forensic Genetics 2012
Further Reading

Manual YHRD, download at www.yhrd.org

Technical and interpretation guidelines for Y-STR analysis