Towards a Reliable Forensic and Population Genetic Use of Chromosome Y Microsatellites

Peter de Knijff[1], Manfred Kayser [2], Lutz Roewer [2]
[1] MGC-Dept. of Human Genetics, Leiden University Medical Center, PO Box 9503, Leiden, The Netherlands.
[2] Institut für Rechtsmedizin, Humboldt-Universität Berlin, Hannoversche Strasse 6, D-10115 Berlin, Germany.

SUMMARY

Chromosome Y microsatellites are extremely useful in delineating closely related human populations. Also, more distantly related populations were found to share Y-microsatellite haplotypes, probably identical by state (thus as a consequence of mutational processes), not identical by descent. Especially for forensic applications of chromosome Y microsatellites this latter observation could have major implications.

A more detailed analysis showed that there were differences in the discriminative power of individual Y-microsatellites. Therefore, we identified the sequence structure of all commonly used Y microsatellites, determined their position on the Y chromosome, and analysed their mutation model and mutation frequency. Several population genetic studies and actual forensic case-work revealed the power of these markers, when used with common sense.

Taken together, our analyses illustrate the resolving power of the various Y-microsatellites, in addition to autosomal and mtDNA studies. Still, for a proper understanding, more fundamental research is necessary to fully understand these loci.

INTRODUCTION

Chromosome Y microsatellites or short-tandem-repeats (STR's) seem to be ideal markers to delineate differences between human populations for several reasons: (i) They are transmitted in uniparental (paternal) fashion without recombination, (ii) They are very sensitive for genetic drift, and (iii) They allow a simple highly informative haplotype construction. Also for forensic applications this ability to differentiate distinct Y chromosomes makes Y-STR's an advantageous addition to the well characterized autosomal STR's. For a number of forensic applications Y-STR's could be superior to autosomal STR's. Especially in rape cases where (i) the differential lysis was unsuccessful, (ii) the number of sperm cells is very low, (iii) due to vasectomy epithelial cells instead of sperm cells from the ejaculate of the perpetrator have to be analysed, and (iv) the perpetrator, due to a familial relationship shares many autosomal bands with the victim, Y-STR's could provide crucial evidence. Also, in the case of male-male rape or rape cases with multiple perpetrators Y-STR's could lead to essential qualitative evidence. In all such cases Y-STR's facilitates a simple and reliable exclusion of suspects.

Here we will focus on the population aspects of Y-STR's only. Their forensic use will be discussed by Roewer *et al.*, following this paper. To date, the number of population genetic studies using Y-STR's is still rather limited. Therefore we explored the use of the Y-STR loci DYS19, DYS389, DYS390, DYS391, DYS392, DYS393, and DXYS156 by means of a number of distinct studies, the outcome of which will also serve their forensic application.

RESULT

First, 200 males from four distinct Dutch geographical regions were genotyped. For some, but not all, Y-STR's we observed marked differences in the genotype frequency distribution between the four regions. As illustrated in Figure 1, 20 from 42 (47.6%) pairwise allele frequency comparisons between regions were significantly different ($p<0.01$). The majority of these differences were found between the northern and southern Dutch regions. Only 4 out of 150 different Y-STR haplotypes were shared between all four regions (Figure 2). For three of these haplotypes the distribution was significantly different. These results confirm our previous conclusion that Y-STR loci are ideal markers for comparing populations at a micro-geographic scale (1).

We speculated that this could be due to recurrent mutation rather then allele sharing by descent. To provide formal proof for this we identified Y-STR haplotypes of a number of Canadian males all descending from French immigrants several centuries ago (2). This study revealed an average Y-STR mutation frequency of 0.21 % (or 1 in 500 meiosis). This seems sufficiently high to explain the above described population affinity observations. Still, this is only a first mutation frequency estimate. Evidently, additional research has to be performed in order to confirm our estimate.

From the population studies we noted that some Y-STR's were very powerful in separating human populations. Especially DYS389 revealed marked differences between some populations. After redesigning the PCR strategy for this locus, we were able to identify four distinct variable repeat stretches (A-D) in this single locus (Figure 3). Based on the variability in three of these four variable stretches we were able to demonstrate marked differences between human populations with this single locus only (3, see Figure 4). Similar differences between regions were also observed for DXYS156 and DYS390 (4,5) but not for the other Y-STR's.

We subsequently explored the use of Y microsatellites in actual case work. Two rape cases were studied. In the first (Figure 5) extremely low amounts of male specific DNA could be detected by means of amelogenin typing, In the second case (Figure 6) no male specific amelogenin peak could be detected at all, probably due to preferential amplification. In both cases, prominent peaks could be detected for chromosome Y microsatellites. In both cases, matching suspects could be found, enabling the prosecution to continue these cases.

DISCUSSION

For a proper understanding and reliable use of Y-STR's we initiated a series of studies. These studies provided valuable information and clearly illustrate their usefulness for the analysis of genetic affinities between human populations. However, for a wider forensic application of Y-STR's, there is still a long way to go before they are as robust in their routine use as most commonly used autosomal STR's. First of all, the PCR protocols currently in use have to be improved. This is especially important when we want to avoid the erroneous detection of a- specific non-Y chromosome bands. These are still regularly encountered by us. In addition, although several multiplex PCR protocols are now described, there is still sufficient room for improvement.

A second important aspect relates to the statistical interpretation of forensic Y-based evidence. As with mtDNA, Y-loci are transmitted uniparentally without recombination (7,8). This has the disadvantage that only distinct Y chromosomes, not individual males, can be identified when sufficient markers are identified, since all paternally related males will have identical Y-STR haplotypes. For this same reason, it will be very difficult to correlate a certain Y-haplotype with the ethnic affiliation of a single individual. In a multi-cultural society with an colonial history (such as the Dutch one), it is e.g. not uncommon to observe a Y-haplotype of African origin combined with an Asian mtDNA-lineage in the same phenotypical white Caucasian male. The regularly outspoken wish to use Y-loci to identify ethnic affiliation of a single individual therefore remains wishful thinking. At most one could think of the identification of certain ethnicity-related Y-lineages.

Not less important is the availability of a chromosome Y-haplotype database. As with mtDNA sequence databases, this will be essential for a correct forensic use of Y-loci. Currently, we have provided over 50 laboratories worldwide with a set of allelic ladders and PCR protocols for most of the above mentioned Y-STR's. This ensures a uniform allelic designation, at least at these laboratories. It is our aim to create an international database of individual Y-STR haplotypes, on the basis of data provided to us by these laboratories. In the near future such data will be available at our URL (see methods). In the mean time we would appreciate receiving any kind of constructive remarks on this topic.

METHODS

All essential information on the use of the Y-STR's have been published by us and others (9-12) and can be found at our Internet site:

http://ruly70.medfac.leidenuniv.nl/~fldo/

This site will, in the future, also contain individual Y-STR haplotype data. Please note that sequence information at this site is still not submitted. All sequences differ from those present in Genbank, mainly due to sequence errors in the deposited sequences. All sequences at our site are based on at least 5 males from each of three populations for almost all observed alleles.

ACKNOWLEDGEMENTS

We want to express our gratitude to B. Rolf, P. Forster, B. Brinkmann, T. Karafet and M. Hammer for sharing (yet) unpublished information. In addition we want to thank Patrick Dieltjes, Rene Mieremet and Thirsa Kraayenbrink for excellent technical support and Chris Tyler-Smith, Mark Jobling and Evelyne Heyer for the fruitful cooperation and the many stimulating discussions on the use of chromosome Y STR's.

REFERENCES

1. Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P. Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related populations. Hum Mol Genet 1996; 5:1029-1033. (with corrigendum in Hum Mol Genet 1997; 6:828).
2. Heyer E, Puymirat J, Dieltjes P, Bakker E, de Knijff P. Estimating Y-chromosome-specific microsatellite mutation frequencies using deep rooting pedigrees. Hum Mol Genet 1997; 6: 799-803.

3. Rolf B, Meyer E, Brinkmann B, de Knijff P. Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering. Eur J Hum Genet 1998; In press.
4. Karafet T, de Knijff P, Wood E, Ragland J, Clark A, Hammer MF. Different patterns of variation at the X- and Y-linked microsatellite loci DXYS156X and DXYS156Y in human populations. Human Biology 1998; In press.
5. Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B. Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. Mol Biol Evol 1998; In press.
6. Nielsen R. A likelihood approach to populations samples of microsatellite alleles. Genetics 1997; 146:711-716.
7. Jobling MA, Pandya A, Tyler-Smith C. The Y chromosome in forensic analysis and paternity testing. Int J Leg Med 1997; 110:118-124.
8. Jobling MA, Tyler-Smith C. Father and sons - the Y chromosome and human evolution. Trends Genet 1995; 11:449-456.
9. de Knijff P, Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Piccinini A, Perez-Lezaun A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, Roewer L. Chromosome Y microsatellites: population genetic and evolutionary aspects. Int J Legal Med 1997; 110: 134-140.
10. Kayser M, Caglià A, Corach D, Fretwell N, Gehrig C, Graziosi G, Heidorn F, Herrmann S, Herzog B, Hidding M, Honda K, Jobling M, Krawczak M, Leim K, Meuser S, Meyer E, Oesterreich W, Pandya A, Parson W, Penacino G, Piccinini A, Perez-Lezaun A, Prinz M, Schmitt C, Schneider PM, Szibor R, Teifel-Greding J, Weichhold G, de Knijff P, Roewer L. Evaluation of Y chromosomal STRs: a multicenter Study. Int J Legal Med 1997; 110:125-133.
11. Redd AJ, Clifford SL, Stoneking M. Multiplex DNA typing of short-tandem-repeat loci on the Y chromosome. Biol Chem 1997; 378:923-927.
12. Prinz M, Boll K, Baum H, Shaler B. Multiplex of Y chromosome specific STR's and performance for mixed samples. Forensic Sci Int 1997; 85:209-218.

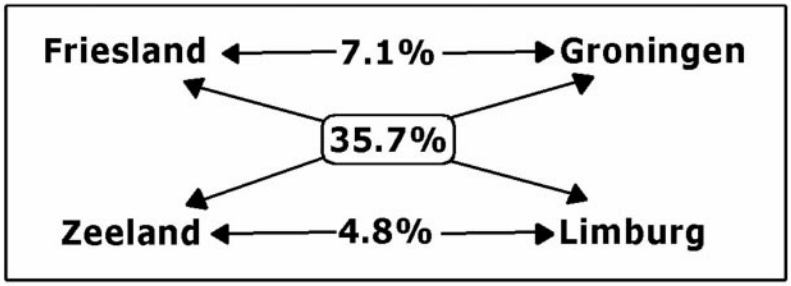

Figure 1. Summary of pairwise Y-STR allele frequency distributions between four distinct Dutch regions. From the 42 comparisons, 20 (47.6 %) were significantly different. Of these, 15 (35.7) were found between northern and southern Dutch regions.

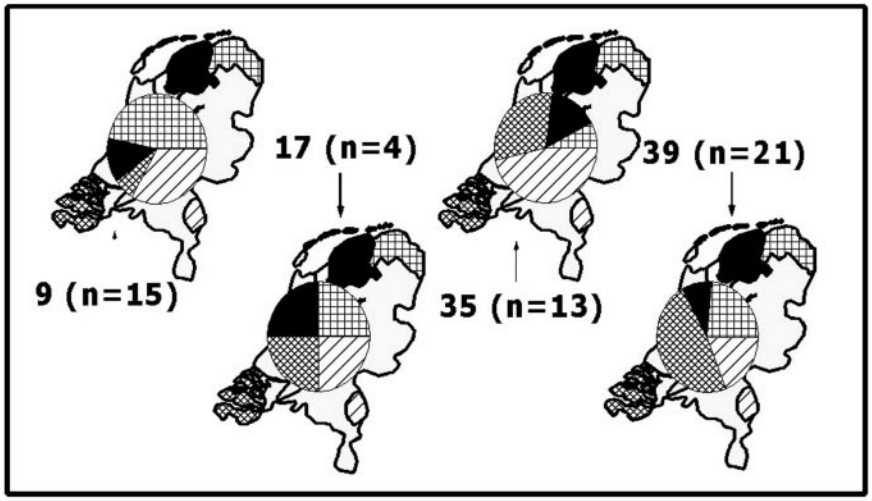

Figure 2. Regional haplotype frequency distribution for the 4 Y-STR haplotypes shared between all four Dutch regions. Haplotypes 9, 35 and 39 revealed markedly different frequencies whereas haplotype 17 was found in single males in all four area's.

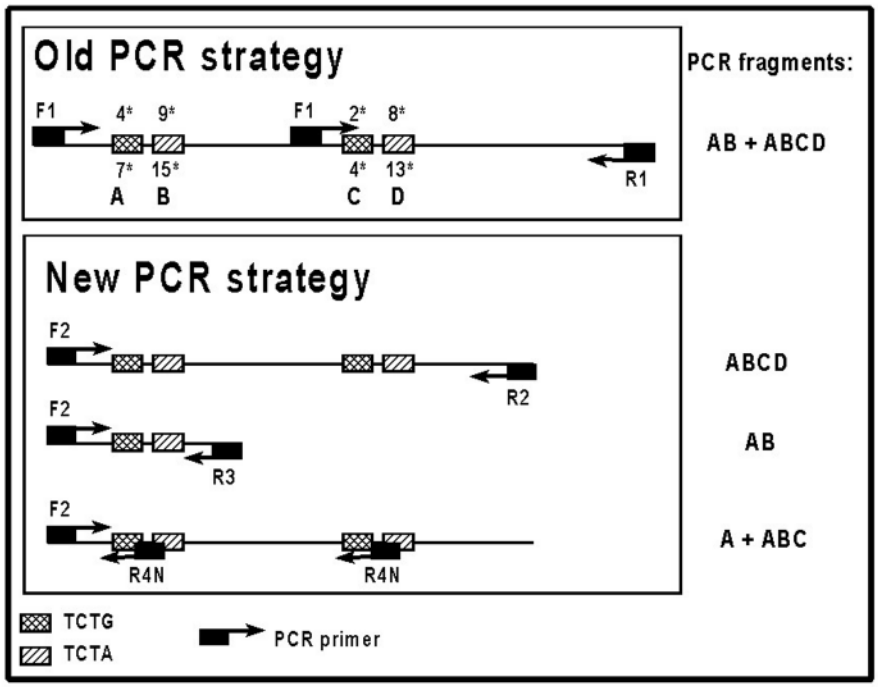

Figure 3. Schematic outline of PCR strategy's for DYS389. The new approach allows the identification of variability of all four repetitive stretches (A-D) of this locus. Detailed protocols can be found at our URL (see methods).

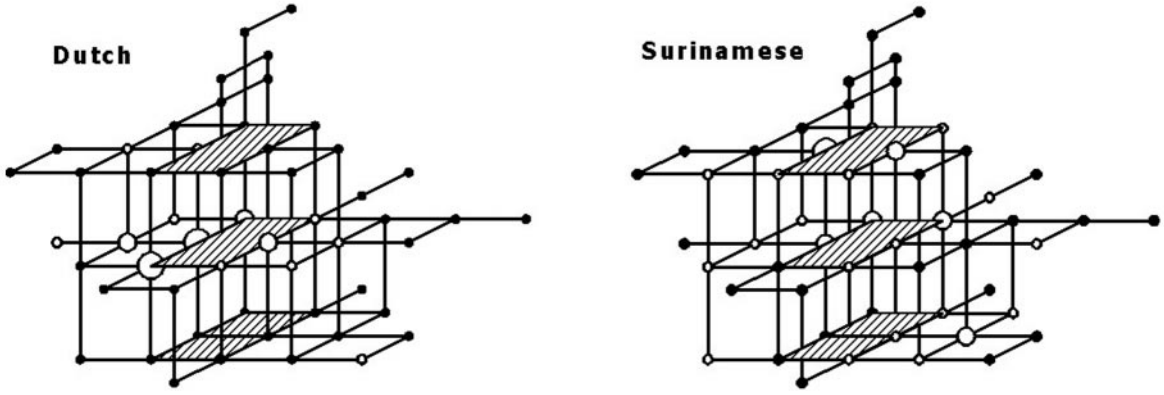

Figure 4. Minimum spanning networks showing the results of DYS389 finetyping. The three different planes (bottom-top, left-right, front-back) represent variability in the first (A), second (B) and fourth (D) variable repeat unit in DYS389. All males in this study had a third (C) repeat length of 3 units. The backbone of this network is based on fine-typing males from 10 distinct male population samples. Each circle represents a single DYS389 haplotype observed among these (>500) males. All distinct haplotypes could be connected with each other assuming single mutation steps in one of the three variable repeats. Results from two samples (Dutch and Surinamese males) are shown. Open circles indicate DYS389 haplotypes observed in the specific population, with size proportional to their frequency. Closed circles represent haplotypes observed at least once in the total study but not in the specific population. Note that virtually all Dutch DYS389 haplotypes can be found in the second top-bottom plane (indicating no variability in the A-stretch) whereas the genetically very diverse Surinamese population showed DYS389 haplotypes in three top-bottom planes (variability in the A-stretch), illustrating the power of this single locus haplotyping in differentiating human populations.

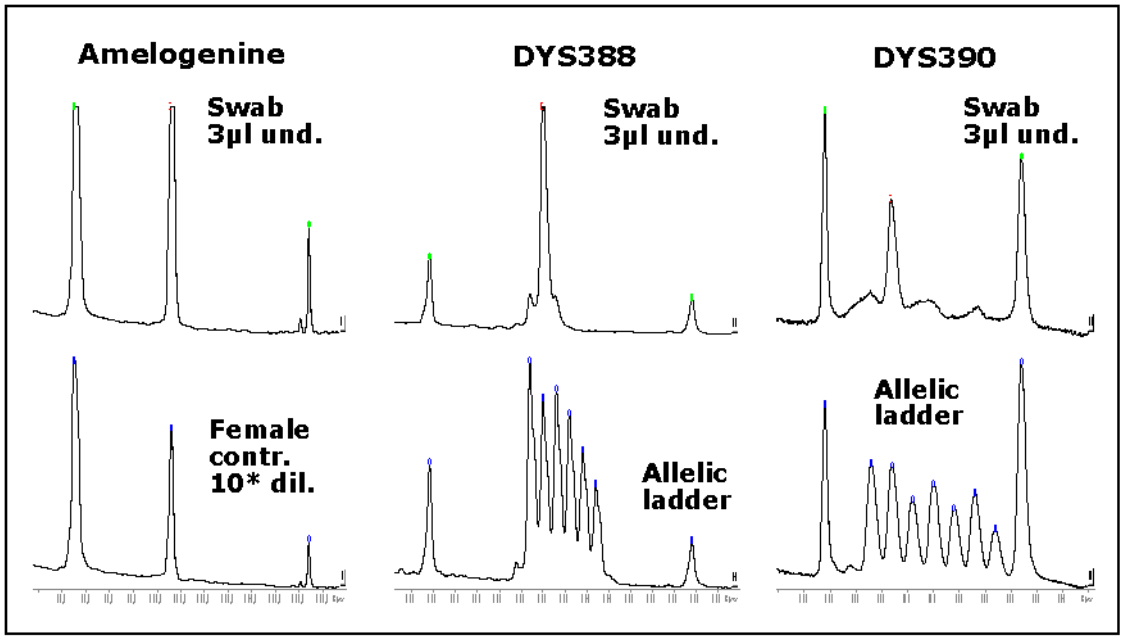

Figure 5. Typing of amelogenin, HUMTH01 and DYS19 in a rape case with low amount of sperm-derived DNA. Note the small 112 bp amelogenin peak in the overloaded vaginal swab-sample.

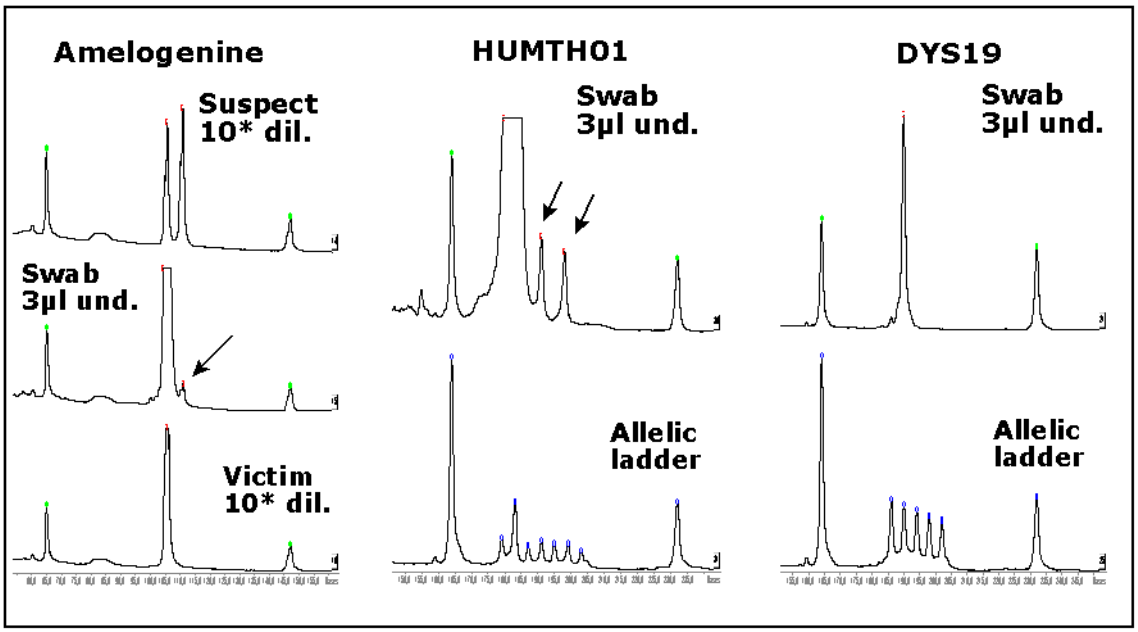

Figure 6. Another example of typing amelogenin and Y-microsatellites in a rape case. No 112 bp amelogenin peak could be detected. Still, prominent Y-peaks were obtained for DYS388 and DYS390.