ISFG Prize Presentation

The use of the Y chromosome in forensic genetics – current practices and future perspectives

Lutz Roewer

Institut für Rechtsmedizin, Humboldt-Universität Berlin, Germany

There are a number of merits that qualify the Y chromosome as a special forensic genetic tool: the male specificity for most of its length, the absence of recombination which provides unambiguous male lineage's and the small effective population size that tends to create population specific allele distributions on the Y chromosomes.

Since the main goal of forensic genetics is individualization of persons or lineages of descent an analytical strategy for the male chromosome must enable the expert to differentiate between the majority of unrelated haplotypes. For this to achieve the choice of the sequence type and its variability (i.e. its mutation rate) as well as the number of individual sequences to be used for profiling is crucial. We have introduced a STR profile for the Y-chromosome consisting of 11 microsatellite sequences which is both informative for individualization purposes as well as for a genetic distance analysis of populations. The technical simplicity of the approach led to a rapid introduction of the technique in many of the forensic labs world-wide. Intense international collaboration facilitates the generation of large haplotype reference databases, most of them are online available and searchable (Europe: http://ystr.charite.de and USA: http://www.ystr.org/usa/). By use of haplotype specific parameters such as the molecular distance (which equals the minimum number of mutational steps separating two haplotypes) and the largest available haplotype databases a Bayesian approach to evaluate Y-STR haplotype matches has been proposed (Roewer et al. 2000, Krawczak 2001).

The recommendations of the ISFG (Gill et al., in press) state some basic principles on forensic analysis using Y-STR polymorphisms: the use of sequenced allelic ladders (now commercially available), the application of a repeat-based nomenclature and the use of suitable haplotype reference databases for statistical evaluation of matches.

Still a matter of research , but of the utmost interest is the potential of the Y-chromosome analysis to unravel the ethnological background of a given male profile. A dual approach – that using Y-STRs as well as Y-SNPs – probably renders the maximum amount of information about the descent of a male lineage typed in a forensic specimen (Kayser et al. 2001).

Address of correspondence:
Dr. Lutz Roewer, Institut für Rechtsmedizin,
Humboldt-Universität Berlin, Hannoversche Straße 6,
D-10115 Berlin, Germany,
Tel. 0049 (0)30 2093 7531; Fax: 0049 (0)30 2093 7240,
e-mail: lutz.roewer@charite.de

IL1

DNA Microarrays

R.G. Sosnowski, Ph.D.

Nanogen, 10398 Pacific Center Court
San Diego, CA 92121, USA

IL2

Continental and subcontinental distributions of mtDNA control region types

P. Forster[1,2]*, F. Calí[2], A. Röhl[3], E. Metspalu[4], O. Rickards[5], R. D'Anna[6], M. Mirisola[6], G. De Leo[6], G.F. De Stefano[5], A. Flugy[6], A. Salerno[6], G. Ayala[2], A. Kouvatsi[7], R. Villems[4], V. Romano[2,6]

[1]The McDonald Institute for Archaeological Research, University of Cambridge, England
[2]Laboratorio di Genetica Molecolare, Istituto OASI (I.R.C.C.S.), Troina, Italy
[3]Institute of Legal Medicine, University of Münster, Germany
[4]Estonian Biocentre and Dept. of Evolutionary Biology, Tartu University, Tartu, Estonia
[5]Dept. of Biology, University of Rome "Tor Vergata", Roma, Italy
[6]Dip. Biopatologia e Metodologie Biomediche, University of Palermo, Palermo, Italy
[7]Dept. of Genetics, Aristotle University, Thessaloniki, Greece

**These authors contributed equally to the present work.*

When the mtDNA profile of a crime scene matches that of a suspect, it is necessary to determine the probability of a chance match by consulting the frequencies of the identified allele in a "reference population". The ceiling principle suggests that that population should be chosen in which the allele of the suspect is found at highest frequency, in order to give the suspect the maximum benefit of doubt. Recently, we advocated the use of a worldwide mitochondrial database combined with a geographical information system to identify the regions of the world with the highest frequencies of matching mtDNA types. Here, we demonstrate that the alternative approach of defining a ceiling reference population on the basis of continent or phenotype (race) is too coarse for a non-negligible percentage of mtDNA control region types.

Corresponding author:
P. Forster
The McDonald Institute for Archaeological Research
University of Cambridge
Downing Street
Cambridge CB2 3ER
UK
Fax +44-1223-339285, e-mail pf223@cam.ac.uk

IL3

Y-chromosomal DNA variation and human population history

C. Tyler-Smith [1], T. Zerjal [1], Y. Xue [1,2], R. S. Wells [3], W. Bao [1,4], S. Zhu [1,4], R. Qamar [1,5], Q. Ayub [1,5], A. Mohyuddin [1,5] and collaborators

[1]Department of Biochemistry, University of Oxford, UK
[2]Department of Medical Biology, Harbin Medical University, China
[3]Wellcome Trust Centre for Human Genetics, University of Oxford, UK
[4]Institute of Genetics, Chinese Academy of Sciences, China
[5]Biomedical and Genetic Engineering Labs, Islamabad, Pakistan

The Y chromosome has unique genetic properties that make it particularly useful for some forensic and evolutionary purposes. It is haploid, male-specific and escapes recombination over most of its length, so it changes only by the gradual accumulation of mutations in distinct lineages. Its effective population size is one quarter of that of autosomes and the genetic mobility of males may be lower than that of females in many societies. The Y is therefore greatly affected by genetic drift and shows very high levels of geographical differentiation. More than two hundred slowly-evolving binary markers (mainly SNPs) and over 30 rapidly-evolving STRs are now available, and combinations of these can be highly informative.

The common patterns of Y-chromosomal variation are illustrated by studies of haplogroup distribution in Europe and STR variation in Pakistan. A collaborative study of 3,616 men from 47 populations in Europe and neighbouring countries was carried out using 11 binary markers known to be informative in this area (Rosser *et al.* (2000) *Am. J. Hum. Genet.* **67,** 1526–1543). Six common haplogroups were found and five of these showed clinal patterns of variation which were probably established by the entry of modern humans in the Palaeolithic and multiple but minor subsequent migrations. Large-scale STR analyses of these European samples are not yet available, but 711 men from 12 populations in Pakistan have been analysed with 16 STRs (Mohyuddin *et al.* (2001) *Forensic Sci. Int.* **118**, 141–146). Most haplotypes were found in single individuals and only two were present in more than 10 men, in each case largely or entirely from the same population.

We have now analysed 1,274 men from Central Asia, China and nearby countries with 16 binary markers and 16 STRs, and find significantly different patterns of Y variation. Large differences in haplogroup frequency are often seen between neighbouring populations, suggesting a high degree of genetic isolation and significant bottlenecks or founder effects. Single lineages with low levels of STR variation can be present at high frequency in individual populations. For example, different lineages in the Kazaks and Kyrgyz had coalescence times of ~ 700 years (with large uncertainties). In addition, one lineage represented almost 10% of the chromosomes found in a wide area stretching from the Pacific to the Caspian Sea. This lineage had a coalescence time of ~ 900 years. In contrast to the Kazak and Kyrgyz lineages, it was found in 16 different populations and has thus spread rapidly over a large region. Genealogical evidence from the Pakistani Hazara suggests that it may be that of Genghis Khan.

Address for correspondence:
Chris Tyler-Smith, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK. Tel. 00 44 (0)1865 275222; Fax. 00 44 (0)1865 275259; e-mail: chris@bioch.ox.ac.uk

IL4

PHYSICAL TRAITS

E. Healy

Dermatopharmacology, University of Southampton, United Kingdom

There is widespread variation in human hair, eye and skin colour, and this results from alterations in the amount and ratio of the two types of melanin pigment, brown black eumelanin and red yellow phaeomelanin, in these tissues. Furthermore, in addition to differences in constitutive (untanned) skin pigmentation, there is also variation between individuals in facultative pigmentation (the degree of cutaneous tanning following exposure to single or multiple doses of ultraviolet radiation). In the mouse, over 50 genetic loci are known to be involved in the control of pigmentation, and many patterns of altered coat colour have been demonstrated to result from monogenic abnormalities. In man, however, most inter-individual differences in pigmentation are estimated to result from polymorphisms at approximately 4 or 5 genetic loci. Pigmentary disorders resulting from pathological alterations in single genes, including oculocutaneous albinism type I *(tyrosinase)*, oculocutaneous albinism type II *(p-gene)*, piebaldism *(c-kit)*, etc. account for only a minority of the overall variation in human pigmentation.

Research during the last decade on the melanocortin 1 receptor *(MC1R)* gene has provided evidence for variants in this gene being an important determinant in a substantial amount of the phenotypic variation in normal human pigmentation. *MC1R* variants are associated with red hair and fair skin, and investigations employing cell transfections and transgenic mice have confirmed that certain variants are compromised in their ability to signal intracellularly via cAMP, and to promote the synthesis of eumelanin. The complete absence of non-synomous variants in African negroes, and their prevalence in white Caucasians suggests that most non-synonymous *MC1R* variants may result in lighter skin colour. Case control studies support a role for *MC1R* variants, when present in the heterozygous state, in causing fair skin type, and kindred studies suggest that the majority of red haired individuals have two variant *MC1R* alleles. However, over 30 *MC1R* variants have been identified to date, and further research is necessary to characterise the effects of the majority of the less frequently encountered variants, as well as to identify the other genes which determine variation in normal human pigmentation.

Address for correspondence:
Dr. Eugene Healy
Dermatopharmacology
Southampton General Hospital
Southampton, UK
e-mail: E.Healy@soton.ac.uk

IL5

Degraded DNA

H. Poinar

Max Planck Institute for Evolutionary Anthropology
Inselstrasse 22
D-04103 Leipzig
Germany
e-mail: poinar@eva.mpg.de

IL6

Statistical issues in Y-chromosomal microsatellite haplotyping

M. Krawczak

Institute of Medical Genetics, University of Wales College of Medicine, Cardiff, UK

A Bayesian approach is proposed to the estimation from existing databases of Y-chromosomal microsatellite (STR) haplotype frequencies. Starting from the assumption that the frequency of a given Y-STR haplotype follows a beta distribution, the method takes the pecularities of microsatellites into account by modelling the distribution parameters of a haplotype frequency as functions of the weighted inverse molecular distance between the haplotype of interest and all other haplotypes in the database. Here, molecular distance denotes the minimum number of mutational steps separating any two haplotypes. Estimates of the distribution parameters are obtainable from databases by exponential regression. The approach, termed "frequency surveying", draws upon standard population genetics theory and can be applied to any combination of markers located on the Y chromosome or in the mitochondrial genome. The method has been implemented as part of the Y-STR database website [www.ystr.org], maintained by the Institut für Gerichtliche Medizin, Humboldt-Universität, Berlin. Before the associated database comprising a variety of populations of European descend was made available for public forensic use, the extent of population stratification pertaining to the database was assessed through Analysis of Molecular Variance (AMOVA). A high level of genetic homogeneity was revealed, and any significant haplotype frequency differences observed were found to correlate with known demographic and historic features. Application of the "frequency surveying" method to a quality-controlled reference Y-STR haplotype database will prove useful for the evaluation of positive trace-donor matches in forensic casework. Problems arising in terms of a logically consistent interpretation of matching results are discussed.

Prof. Dr. rer. nat. Dr. rer. biol. hum. habil. Michael Krawczak,
Institute of Medical Genetics, University of Wales College of Medicine,
Heath Park, Cardiff CF14 4XN, UK,
Tel. 00 44 (0)29 2074 4957; Fax: 00 44 (0)29 2074 6551,
e-mail: krawczak@cardiff.ac.uk