

Minimal hierarchical analysis of global human Y-chromosome SNP diversity by PCR-RFLP

MURRAY P. COX^{1*}

¹Arizona Research Laboratories-Biotechnology, University of Arizona, 1041 East Lowell Street, Biosciences West (Room 246B), Tucson, AZ 85721, USA

Received 12 July 2005; accepted 15 September 2005

Abstract Phylogenetically informative binary polymorphisms on the Y-chromosome have become a standard source of markers for molecular anthropology studies. Although several analytical methods are currently available, most methodologies rely on specialized equipment or expensive reagents that research groups with a primarily anthropological focus often cannot afford. Here I present 15 Y-chromosome lineage assays based on a more widely accessible technique—Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP). In conjunction with eight pre-existing tests, these validated protocols allow the assignment of unknown men to every major branch of the global Y-chromosome tree, including the 18 haplogroups/paragroups defined by the Y-Chromosome Consortium. While not intended to supplant the high-throughput technologies adopted by advanced research units, these minimal protocols will allow smaller molecular anthropology groups to survey male variation in a broader range of human populations, thereby furthering our understanding of global human Y-chromosome diversity.

Key words: Y-SNP, binary polymorphism, haplogroup, screening

Introduction

Binary polymorphisms on the Y-chromosome are highly informative anthropological markers for reconstructing the prehistory of men (Underhill et al., 2000). The Y-Chromosome Consortium (YCC) has inferred a detailed tree of global Y-chromosome diversity from over 250 polymorphisms (YCC, 2002; Jobling and Tyler-Smith, 2003), and currently recognizes 18 primary branches called haplogroups or paragroups (labelled A to R). One or more unique polymorphisms characterize each major monophyletic clade, whose distributions are often geographically restricted. Most Y-chromosome studies first classify male samples to a major branch of the Y-chromosome tree, followed by more detailed classification via screening of additional single nucleotide polymorphisms (SNPs) and short tandem repeat (STR) sequences. This paper presents a set of minimal protocols that allow unknown men to be assigned to every major lineage defined by the YCC.

Various methodologies are currently available for screening binary polymorphisms on the Y-chromosome (Budowle, 2004), including DHPLC (Underhill et al., 1997), MALDI-TOF mass spectrometry (Paracchini et al., 2002), flow cytometry (Vallone and Butler, 2004), and SNaPshot minisequencing (Brion et al., 2004). However, such procedures typically utilize proprietary reagents or specialized equipment that are inaccessible to many small genetics laborato-

ries. For some time, a need has been apparent for lineage assays that utilize a less sophisticated methodology. Here I present 15 tests that employ a more widely accessible technique—Polymerase Chain Reaction-Restriction Fragment Length Polymorphism (PCR-RFLP). In conjunction with eight similar protocols published previously, these tests characterize the global array of Y-chromosome diversity, including the 15 haplogroups and three paragroups defined by the Y-Chromosome Consortium (Figure 1).

While not intended to supplant high-throughput methodologies, PCR-RFLP protocols that require only a minimal molecular biology setup fulfil two niche roles. Firstly, many existing multiplexes target Y-chromosome diversity from specific geographical regions (e.g. Africa and Europe: Brion et al., 2004). Singleplex PCR-RFLP tests will allow the screening of individuals derived from geographical regions outside the scope of an adopted multiplex. Secondly, PCR-

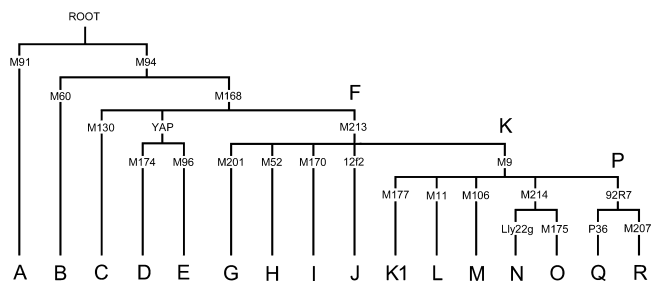


Figure 1. Parsimony tree of global Y-chromosome diversity (modified from Jobling and Tyler-Smith, 2003) showing the phylogenetic position of binary polymorphisms mentioned in this paper.

* Corresponding author. e-mail: mpcoc@email.arizona.edu
phone: +1-520-621 9791; fax: +1-520-626 8050

Published online 28 January 2006
in J-STAGE (www.jstage.jst.go.jp) DOI: 10.1537/ase.050712

RFLP assays will allow less well-resourced research groups to type Y-chromosome haplogroup diversity in many human populations that are currently only poorly studied.

The Y-SNP tests presented here exploit two related methods: conventional and forced PCR-RFLP. Conventional PCR-RFLP utilizes polymorphic SNPs that occur naturally within a restriction endonuclease recognition site. Simple presence or absence of the restriction site discriminates between Y-chromosomes that lack ('ancestral') or carry ('derived') the nucleotide state characteristic of that Y-chromosome lineage. Unfortunately, phylogenetically informative SNPs seldom occur within pre-existing restriction sites, thus necessitating genetic engineering of a novel restriction pattern. This forced PCR-RFLP approach utilizes one PCR primer with a single, predetermined nucleotide mismatch relative to the genuine DNA sequence. PCR amplification incorporates this mismatch primer into the final DNA product, in which the mismatch nucleotide lies adjacent to the haplogroup-defining SNP. This SNP forms an artificial restriction site together with the engineered mismatch nucleotide, and its character state (ancestral or derived) is distinguished by presence/absence of the novel restriction site.

Both PCR-RFLP variants require only basic experimental apparatus (e.g. a PCR thermal cycler, agarose gel electrophoresis equipment), and Y-chromosome lineages are easily assigned by visualizing discrete differences in DNA fragment size on an agarose gel. With the inclusion of ancestral and derived control DNAs, PCR-RFLP methodologies are a cheap and robust alternative to more costly methods of Y-SNP analysis.

Protocols for Typing Y-Chromosome Lineages

Test protocols for lineage-specific Y-chromosome SNPs were validated on control DNAs from the YCC cell line repository (YCC, 2002), a collection of 74 male samples representing a global range of Y-chromosome diversity (kindly provided by Dr Michael Hammer, University of Arizona, USA). All major clades of the Y-chromosome tree are represented by these control DNAs except haplogroup L. Genomic DNA (kindly provided by Dr Chris Tyler-Smith, Wellcome Trust Sanger Institute, UK) classified by DHPLC to haplogroup L-M20 was used to validate the haplogroup L test developed here. To generate sufficient DNA for extensive protocol testing and validation, whole genome amplification was performed on 5 ng aliquots of genomic DNA with Φ 29 DNA polymerase (GenomiPhi DNA Amplification Kit, Amersham Biosciences) yielding a final product of 3–4 μ g of whole genome amplified DNA (wgaDNA). SNP profiles generated from wgaDNA have greater than 99.8% fidelity relative to SNP profiles generated from unamplified genomic DNA (Hosono et al., 2003; Barker et al., 2004). No discrepancies have been detected by this research group.

PCR amplifications were carried out on 10–20 ng of DNA template with primers that flank short stretches of DNA containing a haplogroup-defining polymorphism. PCR reactions (25 μ l) incorporated 1 U of *Taq* DNA polymerase (Roche Molecular Biochemicals). *Taq* DNA polymerase lacks proofreading 3'-to-5' exonuclease activity, thereby preventing undesired 'correction' of engineered primer mis-

match nucleotides during the PCR reaction. The reaction buffer consisted of 2 mM Tris-HCl pH 7.5, 10 mM KCl, 100 μ M dithiothreitol, 10 μ M EDTA, 5% v/v glycerol, 2.5 mM MgCl₂, 200 μ M each dNTP, and 300 nM of the appropriate forward and reverse primers (Prologo Oligonucleotides). Increased efficacy of PCR amplification was facilitated by adding 160 ng/ μ l bovine serum albumin and 0–1.5% formaldehyde (Sarkar et al., 1990). Table 1 lists the sequences of primers that flank each Y-chromosome binary polymorphism together with the size of the amplified PCR products.

PCR amplifications were performed in 96-well thermal cyclers produced by Hybaid (OmniGene) or Techne (TC-512 Gradient PCR) using either standard or touchdown PCR protocols (Table 1). All PCR amplifications had an initial denaturation step of 94°C for 3 min, a final elongation step of 72°C for 5 min, and a final indefinite hold step of 4°C. Standard PCR conditions were 35–40 cycles of 30 s at 94°C, 30 s at the primer-specific annealing temperature (indicated in Table 1), and 60 s at 72°C. Touchdown PCR protocols were designed to avoid the need to optimize primer-annealing temperatures, and more importantly, to circumvent non-specific PCR products by favoring primer annealing at the intended binding site (Don et al., 1991). Three versions of the touchdown PCR protocols were adopted: TD-A, TD-B, and TD-C. TD-A conditions involved 18 cycles starting at 94°C for 20 s, 63°C for 30 s, and 72°C for 60 s with the annealing temperature decreasing by 0.5°C per cycle, followed by 22 cycles with an annealing temperature of 54°C. TD-B conditions involved 20 cycles starting at 94°C for 20 s, 58°C for 30 s, and 72°C for 60 s with the annealing temperature decreasing by 0.5°C per cycle, followed by 20 cycles with an annealing temperature of 48°C. TD-C conditions involved 18 cycles starting at 94°C for 20 s, 55°C for 30 s, and 72°C for 60 s with the annealing temperature decreasing by 0.5°C per cycle, followed by 22 cycles with an annealing temperature of 46°C.

Amplified products ranged from 88 to 709 bp in length (Table 1), and amplification of the correct fragment was confirmed by electrophoresis on 1–3% w/v SeaKem LE agarose gels (Cambrex). Aliquots of the unpurified PCR product (10 μ l) were incubated with an excess (2 U) of the appropriate restriction enzyme (Table 1) in total volumes of 20 μ l, under reaction conditions recommended by the manufacturer. Digest mixtures were electrophoresed on 1–3% w/v SeaKem LE agarose gels, and the observed fragments were sized by comparison with a 100 bp DNA ladder (DNA Molecular Weight Marker XIV, Roche Molecular Biochemicals). Ancestral and derived polymorphic states were determined by comparing observed fragment sizes with expected banding patterns (Table 1).

Evaluation of Y-Chromosome Lineage PCR-RFLP Protocols

Here I describe novel test procedures for 15 Y-chromosome binary polymorphisms that utilize a PCR-RFLP methodology. Together with pre-existing tests for eight additional binary polymorphisms, these protocols allow the assignment of unknown male samples to every major branch of the Y-

Table 1. Primers and polymorphic states for phylogenetically informative binary polymorphisms demarking major branches of the global Y-chromosome tree

Clade	Marker	Mutation (A→D)	Primers (forward/reverse)	PCR protocol [‡]	PCR size (bp)	Restriction endonuclease [†]	RFLP fragments (bp)		References [‡]	
							Ancestral	Derived by sequencing	Marker	Test
A	M91	A ₉ →A ₈ or A ₁₀	F: GAGCTTGGACTTTAGGACGG R: AAACCTTTAAAGGCACCTCTGGC	<i>o</i>	495 or 496	N/A	determined by sequencing	<i>l</i>	<i>o</i>	
B	M60	T insertion	F: GCACCTGGCTCATCATCT R: ATGTTCAATATGTTTCCAGGAG	56°C	389	<i>Mbo</i> I ²	241, 147	389	<i>l</i>	this paper
BCDEF	M94	C→A	F: CACATGGAGAACAGAGAAATGC R: CTTGTGAAATGTTGAAAAGTGG	56°C	405	<i>Mse</i> I ²	405	225, 180	<i>l</i>	this paper
CDEF	M168	C→T	F: AGTTTGAGGTAGAATACTGTTTGCT R: AATCTCATAGTCTCTGACTGFTC	55°C	473	<i>Hinf</i> I ⁴	234, 106, 81, 52	234, 187, 52	<i>l</i>	this paper
C	M130/RPS4Y ₇₁₁	C→T	F: GAGTGGAGGACTGTGAGA R: CCACAGAGATGGTGGGGTA	TD-B	528	<i>Bsr</i> YI ⁵	232, 155, 141	387, 141	<i>i</i>	<i>p</i>
DE	YAP	Alu ⁻ →Alu ⁺	F: CAGGGGAAGATAAGAAATA R: ACTGCTAAAAGGGATGGAT	<i>c, d</i>	455 or 150		150	455	<i>b</i>	<i>c, d</i>
D	M174	T→C	F: TCTTCTCCGTCACAGCAAAAATG R: GCAAAATGCACCCCTCACCTCT	<i>r, s</i>	123	<i>Fsp</i> BI ¹	96, 27	123	<i>l</i>	<i>r, s</i>
E	M96	G→C	F: GTGTAACTTGGAAAACAGGTTCTc R: TCCCAAGGGTACTGTTTTCCT	55°C	180	<i>Bsr</i> YI ⁵	151, 29	180	<i>m</i>	this paper
F	M213	T→C	F: TATAATCAAGTTACCAATTAATGGC R: TTTTGTAAACATTAATGGCAAA	TD-C	409	<i>Hsp</i> 92II ³	291, 118	409	<i>l</i>	this paper
G	M201	G→T	F: GTTCTCAGATCAATAATCCAGTAcC R: GTTCTGAATGAAAAGTTCAAAAGTCAAA	TD-A	226	<i>Bsr</i> YI ⁵	194, 32	226	<i>l</i>	this paper
H	M52	G→T	F: TAATACCTATAAGAATATGCTGTA R: GCTCAAGATTCATCAGTCAAGTC	TD-A	203	<i>Rsa</i> I ⁵	203	178, 25	<i>l</i>	this paper
I	M170	A→C	F: ATGTTTGTTCATAAATGTCAGC R: CACAACCCACACTGAAAAC	TD-B	118	<i>Hsp</i> 92II ³	99, 19	118	<i>l</i>	this paper
J	I2f2	Deletion	F: CTGACTGATCAAATGTTACAGATC R: GGATCCCTTCTTACACCTTATAC	<i>g, k</i>	88	N/A	88	absent	<i>a</i>	<i>g, k</i>
	Control amplicon		F: GACTGTGAGTGTGACTTGTGA R: GAAAGTGCCGTAATAAGTGTGA		112	N/A	present	present		
K	M9	C→G	F: GCAGATATAAAACCTTCAGG R: AAAACCTAACTTTGCTCAAGC	TD-C	340	<i>Hinf</i> I ⁴	181, 95, 64	245, 95	<i>f</i>	<i>h</i>
K1	M177/SRY ₀₃₈	C→T	F: TTAACATTCACAGGACCCAG R: GTTGTGGTCTCCTGTAAAG	56°C	421	<i>Hae</i> III ²	191, 119, 59, 52	191, 178, 52	<i>e</i>	this paper
L	M11	A→G	F: CCTCCCTCTCTGTTGATTTCTAc R: ATACAGCTCTCTCTGGGG	TD-B	148	<i>Msp</i> I ⁵	148	128, 20	<i>f</i>	this paper
M	M106	A→G	F: GGAGGCAACCTAAGAAAG R: AGGTGAACGCAATCTGTGCAI	58°C	572	<i>Taq</i> I ⁵	348, 224	224, 185, 163	<i>l</i>	this paper
NO	M214	A→G	F: ATGCCATGGTCCAAITGTACAGCA R: ACAGTGTGAGACACTGTCTGAAAACAc	TD-B	166	<i>Msp</i> I ⁵	166	137, 29	<i>l</i>	this paper
N	Liy22g	G→A	F: ATAGATGGGCTTCAITGAGI R: GATGTTGGCTTTACAGCTC	<i>n</i>	202	<i>Hind</i> III ⁵	202, 115, 87	202	<i>k</i>	<i>n</i>
O	M175	5 bp deletion	F: CTCTCTGAATCAGGCACATG R: TTAGCCCTTGATTTGACTGTAATCC	TD-B	172 or 167	<i>Eae</i> I ²	134, 38	167	<i>l</i>	this paper
P	92R7	C→T	F: GCCTATCTACTCAGTGAATTTCT R: GACCCGCTGTAGACTGACT	57°C	709	<i>Hind</i> III ⁵	709, 512, 197	709	<i>j</i>	<i>j</i>
Q	P36	G→T	F: GAGAGGATAATGGATAIATAc R: ATCCCATCTGTAACCTTATC	TD-C	123	<i>Taq</i> I ⁵	102, 21	123	<i>q</i>	this paper
R	M207	A→G	F: AGAAAAATCAGAAATATCCCTG R: CTGTTCCGCTGTACGAAATCT	TD-B	204	<i>Dra</i> I ⁴	127, 77	204	<i>l</i>	this paper

Nucleotide mismatches indicated by lower case letters in forward (F) and reverse (R) PCR primers. A primer annealing temperature indicates use of a standard PCR protocol; three touchdown PCR protocols (TD-A, TD-B, TD-C) are described in the text. Restriction digest fragment sizes in bold differentiate ancestral (A) and derived (D) states. References indicate a description of each marker, and if applicable, a previously published PCR-based test protocol.

[†] Restriction endonuclease suppliers: 1, Fermentas; 2, New England Biolabs; 3, Promega Corporation; 4, Q-BIOgene; 5, Roche Molecular Biochemicals.

[‡] References: *a*, Casanova et al. (1985); *b*, Hammer and Horai (1995a); *c*, Hammer and Horai (1995b); *d*, Hammer and Horai (1995b); *e*, Whitfield et al. (1997); *f*, Underhill et al. (1997); *g*, Zerjal et al. (1997); *h*, Hurles et al. (1998); *i*, Bergen et al. (1999); *j*, Hurles et al. (1999); *k*, Rosser et al. (2000); *l*, Shen et al. (2000); *m*, Underhill et al. (2000); *n*, Underhill et al. (2001); *o*, Underhill et al. (2001); *p*, Qamar et al. (2002); *q*, Y Chromosome Consortium (2002); *r*, Cordaux et al. (2004a); and *s*, Cordaux et al. (2004b).

chromosome tree, including the 15 haplogroups and three paragroups defined by the Y-Chromosome Consortium (YCC, 2002). These assays have been validated against fully-profiled YCC control wgaDNAs, including direct DNA sequencing of ancestral and derived controls, and are now used routinely to screen non-whole genome amplified genomic DNAs from Africa, Southeast Asia, and Oceania (Cox, 2003). For a practical application of these tests on real-world samples, see Cox and Lahr's (2006) analysis of Y-chromosome diversity in the Solomon Islands.

Each PCR primer set amplifies a single 88–709 bp DNA fragment. Mismatch primers undergo PCR amplification as efficiently as conventional primers. Two binary polymorphisms (N-Lly22g and P-92R7) occur within segmental duplications (Skaletsky et al., 2003), and these PCR products therefore represent a composite sequence from two Y-chromosome regions. Only one repeat carries the Lly22g and 92R7 polymorphisms, and thus only a portion of the amplified PCR product represents the phylogenetically informative marker. Consequently, these PCR products are not completely digested even when the restriction endonuclease site is present; uncleaved PCR product is always observed following restriction endonuclease digestion. However, ancestral and derived character states can always be assigned correctly for N-Lly22g and P-92R7 (Table 1). Restriction endonuclease cleavage is complete for the remaining assays.

Digested DNA fragments vary from 19 to 709 bp, but no single assay requires this range of DNA fragment sizes to be discriminated. All 23 Y-chromosome lineages can be assigned correctly by examining restriction fragments between 88 and 197 bp, and these fragment sizes are easily distinguished following standard agarose gel electrophoresis (Figure 2). Conventional PCR-RFLP typically generates larger DNA fragments, whereas forced PCR-RFLP tests—in which the mismatch primer forms part of the engineered restriction site—cleave only the primer sequence (19–32 bp) from the PCR product. Electrophoresis times should be increased with forced PCR-RFLP tests to discriminate between cleaved and uncleaved DNA fragments.

All binary polymorphisms described here, except P36, demark unique, monophyletic clades on the global Y-chromosome tree. P36 (a G→T polymorphism described originally as G→A; Karafet et al., 2005: p. 98) defines haplogroup Q, a lineage with North American and Central Asian affinities (Zegura et al., 2004). However, P36 recurs on a limited subset of African Y-chromosome backgrounds, including four San bushmen carrying lineage A2 Y-chromosomes (YCC DNAs 5, 22, 34 and 35). Nevertheless, haplogroups A and Q are readily distinguished with the protocols presented here. Lineage A Y-chromosomes can be identified by their A-M91 background; lineage Q Y-chromosomes can be identified by their P-92R7 background. A PCR-RFLP test is presented here for P-92R7, and although the A-M91 polymorphism cannot be characterized by PCR-RFLP, direct DNA sequencing distinguishes the ancestral state (a 9-residue polyadenine tract) from the derived state (an 8- or 10-residue polyadenine tract).

These 15 novel Y-chromosome binary polymorphism assays (in association with eight previously published PCR-

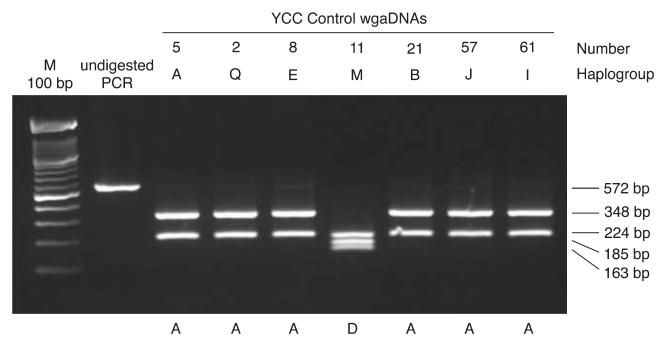


Figure 2. A PCR-RFLP test for Y-chromosome lineage M-M106. Lane 1 (left) is a DNA sizing ladder (marker) with 100 bp increments. Lane 2 shows an undigested 572 bp PCR product. Lanes 3–9 show *TaqI* digested PCR products from seven unrelated Y Chromosome Consortium DNAs (sample ID numbers and haplogroup assignments indicated above the gel; YCC, 2002). Six samples show two DNA fragments of 348 and 224 bp, and are therefore ancestral (A) for Y-chromosome haplogroup M. Only one sample shows the three DNA fragments of 224, 185, and 163 bp created by the second lineage M-specific *TaqI* restriction site, and therefore carries the derived (D) haplogroup M polymorphism.

based tests) allow full resolution of the major branches of the global Y-chromosome tree. Because many PCR products are smaller than 500 bp and can be amplified from fragmented and degraded DNA samples (unpublished data), some of these tests may find utility in forensic and ancient DNA applications. However, these protocols are intended primarily for modern anthropological surveys of Y-chromosome diversity. PCR-RFLP protocols require little experimental equipment and limited technical expertise. Hopefully these tests will encourage new research groups, perhaps previously dissuaded by the high costs traditionally associated with Y-chromosome screening, to implement a broader range of anthropological surveys, thereby contributing to our growing understanding of global Y-chromosome diversity.

Acknowledgments

I extend my appreciation to Dr Michael Hammer (University of Arizona) for providing control DNAs from the Y Chromosome Consortium's cell line repository; Dr Chris Tyler-Smith (Wellcome Trust Sanger Institute) for supplying haplogroup L DNAs and for guidance with Y-SNP typing protocols; and Dr. Peter Underhill (Stanford University) and Dr Matthew Hurles (Wellcome Trust Sanger Institute) for helpful discussion regarding Y-Chromosome typing methods. I also thank the Department of Zoology (University of Oslo), the Department of Biochemistry (University of Otago), and the Leverhulme Centre for Human Evolutionary Studies (University of Cambridge) for their patronage of this research. Funding from the Foundation for Research, Science and Technology (New Zealand), the University of Otago Research Committee (New Zealand), the Kon-Tiki Museum (Norway), and the Isaac Newton Trust (United Kingdom) supported this research.

References

- Barker D.L., Hansen M.S., Faruqi A.F., Giannola D., Irsula O.R., Lasken R.S., Latterich M., Makarov V., Oliphant A., Pinter J.H., Shen R., Slepstova I., Ziehler W., and Lai E. (2004) Two methods of whole-genome amplification enable accurate genotyping across a 2320-SNP linkage panel. *Genome Research*, 14: 901–907.
- Bergen A.W., Wang C.Y., Tsai J., Jefferson K., Dey C., Smith K.D., Park S.C., Tsai S.J., and Goldman D. (1999) An Asian-Native American paternal lineage identified by RPS4Y resequencing and by microsatellite haplotyping. *Annals of Human Genetics*, 63: 63–80.
- Bermisheva M.A., Viktorova T.V., and Khusnutdinova E.K. (2001) Polymorphism of Y-chromosomal diallelic loci in populations of the Volga–Ural region. *Russian Journal of Genetics*, 37: 833–837.
- Brion M., Sobrino B., Blanco-Verea A., Lareu M.V., and Carracedo A. (2004) Hierarchical analysis of 30 Y-chromosome SNPs in European populations. *International Journal of Legal Medicine*, 119: 10–15.
- Budowle B. (2004) SNP typing strategies. *Forensic Science International*, 146 Supplement 1: 139–142.
- Casanova M., Leroy P., Boucekkine C., Weissenbach J., Bishop C., Fellous M., Purrelo M., Fiori G., and Siniscalco M. (1985) A human Y-linked DNA polymorphism and its potential for estimating genetic and evolutionary distance. *Science*, 230: 1403–1406.
- Cordaux R., Auger R., Bentley G., Nasidze I., Sirajuddin S.M., and Stoneking M. (2004a) Independent origins of Indian caste and tribal paternal lineages. *Current Biology*, 14: 231–235.
- Cordaux R., Weiss G., Saha N., and Stoneking M. (2004b) The northeast Indian passageway: a barrier or corridor for human migrations? *Molecular Biology and Evolution*, 21: 1525–1533.
- Cox M.P. (2003) Genetic Patterning at Austronesian Contact Zones. Ph.D. dissertation, University of Otago, Dunedin, New Zealand.
- Cox M.P. and Lahr M.M. (2006) Y-chromosome diversity is inversely associated with language affiliation in paired Austronesian- and Papuan-speaking communities from Solomon Islands. *American Journal of Human Biology* (in press).
- Don R., Cox P., Wainwright B., Baker K., and Mattick J. (1991) ‘Touchdown’ PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, 19: 4008.
- Hammer M.F. (1994) A recent insertion of an Alu element on the Y Chromosome is a useful marker for human population studies. *Molecular Biology and Evolution*, 11: 749–761.
- Hammer M.F. and Horai S. (1995a) Y chromosomal DNA variation and the peopling of Japan. *American Journal of Human Genetics*, 56: 951–962.
- Hammer M.F. and Horai S. (1995b) Erratum: Y chromosomal DNA variation and the peopling of Japan. *American Journal of Human Genetics*, 56: 1512.
- Hosono S., Faruqi A.F., Dean F.B., Du Y., Sun Z., Wu X., Du J., Kingsmore S.F., Egholm M., and Lasken R.S. (2003) Unbiased whole-genome amplification directly from clinical samples. *Genome Research*, 13: 954–964.
- Hurles M.E., Irvén C., Nicholson J., Taylor P.G., Santos F.R., Loughlin J., Jobling M.A., and Sykes B.C. (1998) European Y-chromosomal lineages in Polynesians: a contrast to the population structure revealed by mtDNA. *American Journal of Human Genetics*, 63: 1793–1806.
- Hurles M.E., Veitia R., Arroyo E., Armenteros M., Bertranpetit J., Perez-Lezaun A., Bosch E., Shlumukova M., Cambon-Thomsen A., McElreavey K., Lopez de Munain A., Rohl A., Wilson I.J., Singh L., Pandya A., Santos F.R., Tyler-Smith C., and Jobling M.A. (1999) Recent male-mediated gene flow over a linguistic barrier in Iberia, suggested by analysis of Y-chromosomal DNA polymorphism. *American Journal of Human Genetics*, 65: 1437–1448.
- Jobling M.A. and Tyler-Smith C. (2003) The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics*, 4: 598–612.
- Karafet T.M., Lansing J.S., Redd A.J., Watkins J.C., Surata S.P.K., Arthawiguna W.A., Mayer L., Bamshad M., Jorde L.B., and Hammer M.F. (2005) Balinese Y-chromosome perspective on the peopling of Indonesia: genetic contributions from pre-Neolithic hunter-gatherers, Austronesian farmers, and Indian traders. *Human Biology*, 77: 93–113.
- Paracchini S., Arredi B., Chalk R., and Tyler-Smith C. (2002) Hierarchical high-throughput SNP genotyping of the human Y chromosome using MALDI-TOF mass spectrometry. *Nucleic Acids Research*, 30: e27.
- Qamar R., Ayub Q., Mohyuddin A., Helgason A., Mazhar K., Mansoor A., Zerjal T., Tyler-Smith C., and Qasim Mehdi S. (2002) Y-chromosomal DNA variation in Pakistan. *American Journal of Human Genetics*, 70: 1107–1124.
- Rosser Z.H., Zerjal T., Hurles M.E., Adojaan M., Alavantic D., Amorim A., Amos W., Armenteros M., Arroyo E., Barbujani G., Beckman G., Beckman L., Bertranpetit J., Bosch E., Bradley D.G., Brede G., Cooper G., Côte-Real H.B., de Knijff P., Decorte R., Dubrova Y.E., Evgrafov O., Gilissen A., Glisic S., Gölge M., Hill E.W., Jexiorowska A., Kalaydjieva L., Kayser M., Kivisild T., Kravchenko S.A., Krumina A., Kučinskas V., Lavinha J., Livshits L.A., Malaspina P., Maria S., McElreavey K., Meitinger T.A., Mikelsaar A.V., Mitchell R.J., Nafa K., Nicholson J., Nørby S., Pandya A., Parik J., Patsalis P.C., Pereira L., Peterlin B., Pielberg G., Prata M.J., Previderé C., Roewer L., Roots S., Rubinsztein D.C., Saillard J., Santos F.R., Stefanescu G., Sykes B.C., Tolun A., Vilems R., Tyler-Smith C., and Jobling M.A. (2000) Y-chromosomal diversity in Europe is clinal and influenced primarily by geography, rather than by language. *American Journal of Human Genetics*, 67: 1526–1543.
- Sarkar G., Kapelner S., and Sommer S. (1990) Formamide can dramatically improve the specificity of PCR. *Nucleic Acids Research*, 18: 7465.
- Shen P., Wang F., Underhill P.A., Franco C., Yang W.-H., Roxas A., Sun R., Lin A.A., Hyman R.W., Vollrath D., Davis R.W., Cavalli-Sforza L.L., and Oefner P.J. (2000) Population genetic implications from sequence variation in four Y chromosome genes. *Proceedings of the National Academy of Sciences of the United States of America*, 97: 7354–7359.
- Skaletsky H., Kuroda-Kawaguchi T., Minx P.J., Cordum H.S., Hillier L., Brown L.G., Repping S., Pyntikova T., Ali J., Bieri T., Chinwalla A., Delehaunty A., Delehaunty K., Du H., Fewell G., Fulton L., Fulton R., Graves T., Hou S.-F., Latrielle P., Leonard S., Mardis E., Maupin R., McPherson J., Miner T., Nash W., Nguyen C., Ozersky P., Pepin K., Rock S., Rohlfling T., Scott K., Schultz B., Strong C., Tin-Wollam A., Yang S.-P., Waterston R.H., Wilson R.K., Rozen S., and Page D.C. (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature*, 423: 825–837.
- Underhill P., Jin L., Lin A., Mehdi S., Jenkins T., Vollrath D., Davis R., Cavalli-Sforza L.L., and Oefner P. (1997) Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography. *Genome Research*, 7: 996–1005.
- Underhill P., Shen P., Lin A., Jin L., Passarino G., Yang W., Kaufman E., Bonne-Tamir B., Bertranpetit J., Francalacci P., Ibrahim M., Jenkins T., Kidd J., Mehdi S., Seielstad M., Wells S.R., Piazza A., Davis R., Feldman M., Cavalli-Sforza L.L., and Oefner P. (2000) Y chromosome sequence variation and the history of human populations. *Nature Genetics*, 26: 358–361.
- Underhill P.A., Passarino G., Lin A.A., Shen P., Mirazon Lahr M.,

- Foley R.A., Oefner P.J., and Cavalli-Sforza L.L. (2001) The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Annals of Human Genetics*, 65: 43–62.
- Vallone P.M. and Butler J.M. (2004) Y-SNP typing of U.S. African American and Caucasian samples using allele-specific hybridization and primer extension. *Journal of Forensic Science*, 49: 723–732.
- Whitfield L., Sulston J., and Goodfellow P. (1995) Sequence variation of the human Y chromosome. *Nature*, 378: 379–380.
- Y Chromosome Consortium (2002) A nomenclature system for the tree of human Y-chromosomal binary haplogroups. *Genome Research*, 12: 339–348.
- Zegura S.L., Karafet T.M., Zhivotovsky L.A., and Hammer M.F. (2004) High-resolution SNPs and microsatellite haplotypes point to a single, recent entry of Native American Y chromosomes into the Americas. *Molecular Biology and Evolution*, 21: 164–175.
- Zerjal T., Dashnyam B., Pandya A., Kayser M., Roewer L., Santos F.R., Schiefenhövel W., Fretwell N., Jobling M.A., Harihara S., Shimizu K., Semjidmaa D., Sajantila A., Salo P., Crawford M.H., Ginter E.K., Evgrafov O.V., and Tyler-Smith C. (1997) Genetic relationships of Asians and Northern Europeans, revealed by Y-chromosomal DNA analysis. *American Journal of Human Genetics*, 60: 1174–1183.