Proc. Indian natn Sci Acad. **B69** No. 4 pp 525-538 (2003)

Genomics of the Human Y Chromosome and Molecular Diagnosis

ANU BASHAMBOO, HARLEEN MANGAT GIRAN, MD. ASIM AZFER and SHER ALI*
National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067

(Received on 14 January 2003; Accepted after revision on 6 May 2003)

The human Y chromosome carries a few functional genes as against a plethora of non-coding DNA sequences and shows a high degree of geographical and ethnic variations for a range of loci manifested as genetic polymorphisms. Y-chromosome linked sequence tagged sites (STS) and short tandem repeat (STR) marker systems offer infallible tool for gender identification, paternity testing, genome individualization and assessing male fertility status. Population-specific Y haplotypes and Single Nucleotide Polymorphisms (SNPs) are envisaged to be useful in establishing a correlation between diseased phenotypes with genetic polymorphisms. We discuss genomics of the human Y-chromosome and its possible applications in biology, medical and forensic sciences.

Key Words : Single nucleotide polymorphisms (SNPs), Y-linked marker, Diseased phenotype, Gender identification, DNA diagnosis

Introduction

The human Y chromosome enjoys genetic eminence by virtue of its haploid status and dominant role in sex-determination. There are two pseudoautosomal regions, PABY1 and PABY2 on the short (Yp) and long (Yq) arms of the Y chromosome respectively with their homologues on the X chromosome. Barring these regions, no other parts of the Y chromosome participate in the meiotic recombination thereby leaving about 95% of the human Y as non-recombining (NRY) region (Tilford et al. 2001). Consequently, the Y chromosome evolves along lineages, accumulating diversity through a repertoire of different mutational processes (Lahn et al. 2001). These lineages are distributed non-randomly among human populations as a spin-off of population dynamics, genetic drifts and socio-cultural factors (Parra et al. 1998, Ali & Hasnain 2002). Also, with the passage of time, genes specific for male differentiation seem to migrate to the Y chromosome (Delbridge et al. 1997). Normally, autosomal genes undergo DNA repair during meiotic recombination but for Y genes, in the absence of pairing (except for the PABY regions mentioned above), this repair mechanism is not operative. Multicopy and inert nature of some Y linked genes has been hampering their accurate mapping; nonetheless, this architectural strategy seems to protect their organizational and functional integrity (see also, Tilford et al. 2001). Y chromosome related genetic variations are seen not only amongst different populations but also between individuals. If these genetic variations were exploited to uncover the susceptibility of ethnic groups to a particular disease(s), it would help establishing molecular structure of a population. Prediction of high-risk groups will then compliment the demand of preventive medicine.

The human X and Y chromosomes originated from a pair of autosomes some 300 million years ago from the reptiles, long before mammals arose (Graves & Foster 1994). The genes on these chromosomes were subjected to environmental stimuli for sex determination, which is reflected even today in certain reptiles (Ganesh & Raman 1997). These genes acquired mutations over a period of time resulting in loss of responsiveness to

Correspondence Address: *E-mail:* sherali@nii.res.in; sheralib5@hotmail.com; *Tel:*(011) 2618 8315 *Ext.* 453; *Fax:*2616 2125; *Mobile:*9818004632

environmental cues. In mammals, sex chromosomes probably arose with the differentiation of *SRY* gene from *SOX3*, which is a structural homologue on the mammalian X chromosome (Stevanovic et al. 1993, Foster & Graves 1994). Expression studies have shown that *SRY* and SOX3 descended from a progenitor gene, with the more evolved *SRY* having gained and retained the male-determining function (Foster & Graves 1994). During the divergence of X and Y-chromosomes in mammals, remarkably minor changes took place in the X but rapid degeneration occurred in the Y chromosome (Graves 1995, Lahn & Page 1999). Further, blockwise mutations on the Y followed by largescale inversion of much of its portions gave rise to non-recombining regions (NRY) on the Y chromosome (Lahn & Page 1999). This gave rise to only one copy of many genes in males and two copies in females. This inequality of gene dosage was dealt with by X chromosome inactivation in the females. However, X-homologous genes on the NRY suggest their preservation on both NRY and X chromosome with male and female cells expressing two copies of such genes. For this, X-homologs of the NRY genes escaped X-inactivation (Lahn & Page 1997). It was also discovered that the genes common to the X and Y were functionally interchangeable. This finding has implications in Turner syndrome, a disorder in which females are born with only one X chromosome. It is speculated that Turner syndrome is caused by an inadequate expression of some X-Y common genes that escape X inactivation. Given that several X-NRY genes are involved in cellular housekeeping, it is likely that some Turner syndrome characteristics are due to inadequate expression of particular housekeeping genes. Thus, it would be of relevance if the X-homologous NRY genes were investigated in greater details as candidates for Turner syndrome. Owing to ethical and logistic constraints, however, it may not be possible to undertake expression studies on the up or down regulation of the Y-linked genes in human.

Recent work on the human Y chromosome has broadened our understanding on the complex organization of a number of genes and their possible biological roles (Tilford et al. 2001, Kuroda-Kawaguchi et al. 2001, Repping et al. 2002, Vogt & Fernandez 2003). The new information on the Y genes and specific insight into the molecular events that leads to male infertility are envisaged to be useful for genetic counseling to couples seeking *in vitro* fertilisation because of male-factor infertility.

Human Y Chromosome and Male Fertility

Approximately 5-10% of all men have severe defects in sperm production and in majority of the cases; the infertility cannot be attributed to known causes (idiopathic infertility). The vast majority of deletions (microdeletions) associated with complete absence of germ cells (azoospermia) or severe oligozoospermia ($<1x10^6$ sperm/ml) are cytogenetically undetectable. In rare cases, the latter group transmits the deletions to their male offspring. However, due to infertile phenotype, most deletions are *de novo* and lost subsequently from the population together with any associated NRY alleles. Studies indicate that between 10-15% of men with idiopathic azoospermia and 5-10% of men with oligozoospermia carry deletions of one or the other parts of the long arm of the Y chromosome (Krausz & McElreavey 1999). Approximately 60% of the azoospermic and severe oligozoospermic men are defined as idiopathic. With a conservative assumption that 5% of these cases are caused by Y chromosome deletions, the frequency of Y deletions in a general male population associated with infertility is estimated to be 2%x60%x5%=0.06%. Therefore, it may be assumed that 1 in 6,000 men carry a Y-deletion resulting in male infertility (Quintana-Murci et al. 2001). It has been known that fertility status of a person is independent of his overall health suggesting that genes involved in the regulation and control of fertility have well-defined roles in the genome. With the completion of the human genome project, in principle, it should be possible to uncover such Y-linked genes since their numbers are unlikely to be huge.

Thus far, a total of 107 genes (table 1) have been listed (http//gdbwww.gdb.org/gdbreports/GeneByChromosome.Y.alpha.html) and several of them have been characterized. Of the 46 characterized genes, 41 are found to transcribe exclusively in the testis whereas the other five express widely

Table 1. *Genes on Chromosome Y with their Accession and Map Positions*

Last Updated Sun Dec 29 230639 EST 2002
*after the Cytogenetic Location indicates HUGO approval

S.No	Symbol	Aliases	Accession ID	Cytogenetic Location
1	ALTE	Ac-like transposable element TRAMP, KIAA0785, H. s. 9933	GDB 9958980	Xp22.33-Xp22.33 Yp11-Yp11
2	AMELY	Hs.1238 AMGL (amelogenin on Y)	GDB119676	Yp11.2-Yp11.2*
3	SPRY3	Spry3 (Sprouty homolog 3, Drosophila)	GDB9835923	Xq-X-q Yq-Yq
4	ASMT	Acetylserotonin O-methyltransferase HIOMT	GDB136259	Xp22.3-Xp22.3* Yp11.3-Yp11.3*
5	ASMTL	Acetylserotonin O-methyltransferase-like (Hs.6315) Weekly similar to hydroxyindole-O-methyltransferase [H. sapiens]	GDB9956043	Xp22.3-Xp22.3 Yp11.3-Yp11.3
6	TGIF2LX	TGFB-induced factor 2-like, X-linked	GDB 11508016	Y-Y
7	AZF1	AZF (Azoospermia factor 1)	GDB119027	Yq11-Yq11*
8	AZF2	Azoospermia factor 2	GDB456131	Ycen-Yqter
9	CD24L4	CD24 antigen-like 4	GDB383841	Yq11-Yq11
10	CDY1	Chromodomain protein 1, Y chromosome	GDB9954846	Yq11.22-Yq11.22
11	CDY2	Chromodomain protein 2, Y chromosome	GDB9956669	Yq11.22-Yq11.22
12	CRLF2	Cytokine receptor-like factor 2, CRL2	GDB11499114	Xp22.3-Xp22.3 Yp11.3-Yp11.3
13	CSF2RA	CSF2R, Colony stimulating factor 2 receptor, alpha, low-affinity, (granulocyte-macrophage)	GDB118777	Xp22.32-Xp22.32 Yp11.3-Yp11.3
14	CSPG4LY	Chondroitin sulfate proteoglycan 4-like, Y-linked	GDB11505842	Y-Y
15	CYorf1	Chromosome Y open reading frame 1	GDB9865780	Yq11.22-Yq11.22
16	CYorf14	Chromosome Y open reading frame 14	GDB11510702	Y-Y
17	CYorf15A	Chromosome Y open reading frame 15A	GDB11510704	Y-Y
18	CYorf15B	Chromosome Y open reading frame 15B	GDB11510706	Yq11.22-Yq11.22
19	CYorf16	Chromosome Y open reading frame 16	GDB11510708	Y-Y
20	DAZ	Deleted in azoospermia, Hs.70936, SPGY	GDB635890	Yq12-Yq12* Yq11-Yq11
21	DAZ2	Deleted in azoospermia 2	GDB11501009	Yq11.21-Yq11.21
22	DAZ3	Deleted in azoospermia 3	GDB11501011	Yq11-Yq11
23	DAZ4	Deleted in azoospermia 4	GDB11501013	Yq11-Yq11
24	DBY	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide, Y chromosome, DDX3Y	GDB9956138	Yq-Yq
25	DXY-S155E	DNA segment on chromosome X and Y (Unique) 155 expressed sequence, XE7, XE7Y, Hs.21595	GDB9964054	Xp22.32-Xp22.32 Yp11.3-Y11.3
26	EIF1AY	Eukaryotic translation initiation factor 1A, Y chromosome	GDB9954847	Ypter-Yqter Yq11.22-Yq11.22
27	VCY	Variable charge, Y chromosome, BPY1	GDB9954845	Yq11-Yq11, Ypter-Yqter
28	GCY	Growth contrd, Y chromosome influenced, STA, TSY	GDB119267	Yq11-Yq11*
29	GOLGA-2LY	Golgi autoantigen, golgin subfamily a, 2-like, Y-linked	DGB11506083	Y-Y
30	HSFY	Heat shock transcription factor,Y-linked, HSF2L	GDB11506147	Y-Y
31	IL3RA	Interleukin 3 receptor, alpha (low affinity)	GDB128985	Xp22.3- Xp22.3* Yp11.3- Yp11.3*
32	IL9R	Interleukin 9 receptor	GDB134444	Xq28-Xq28* Yq12-Yq12
33	UTY	Ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	GDB9864459	Ypter-Yqter
34	MIC2	Hs.75467, Antigen identified by monoclonal antibodies 12E7, F21 and O13	GDB120184	Xp22.32-Xp22.32 Yp11.3-Yp11.3*
35	NLGN4Y	Neuroligin 4, Y linked	GDB11506338	Y-Y

36	USP9Y	Ubiquitin specific protease 9, Y Chromosome (fat facets-like Drosophila), (Drosophila fat facets related), DFFRY	GDB9839319	Yq11.2-Yq11.2
37	PABY	Pseudoautosomal boundary region, Y-link	GDB120259	Yp11.3-Yp11.3*
38	PCDH11Y	PCDH22, Protocadherin 11 Y-linked, Protocadherin 22, PCDHY	GDB11504539	Yp11.2-Yp11.2
39	TSPY	Hs.89644, Testis specific protein, Y-linked, Hs.2051	GDB120471	Yp-Yp* Ypter-Yp11.2
40	PPP2R3B	PPP2R3L, Protein phosphatase 2 (formerly 2A), regulatory subunit B'', beta, PPP2R3LY	GDB11507504	Xp22.3-Xp22.3 Yp11.3-Yp11.3
41	PRKY	Protein kinase, Y-linked, Hs.56336	GDB631715	Yp11.2- Yp11.2*
42	PRY	PTPN13-like, Y-linked, PTPN13LY	GDB9954842	Yq11.22-Yq11.22
43	RBMY1A1	RBM1, RBM2, RNA binding motif protein, Y chromosome family 1 menber A1, RNA binding motif protein 1, RNA binding motif protein 2, Hs. 2958, H. sapiens mRNA for YRRM1, YRRM2, YRRM1	GDB285393	Yq11.23-Yq11.23
44	RBMY1A2	RNA binding motif protein, Y chromosome, family 1, member A2	GDB9848784	Yq11.23-Yq11.23
45	TMSB4Y	TB4Y, Thymosin, beta 4, Y chromosome	GDB9954848	Y-Y
46	RBMY1C	RNA binding motif protein, Y chromosome, family 1, member C	GDB9848787	Yq11.23-Yq11.23
47	RBMY1G	RNA binding motif protein, Y chromosome, family 1, member G	GDB 9848789	Yq11.23-Yq11.23
48	RBMY1H	RNA binding motif protein, Y chromosome, family 1, member H	GDB9848791	Yq11.23-Yq11.23
49	RBMY2A	RNA binding motif protein, Y chromosome, family 2, member A	GDB9848793	Yq11.23-Yq11.23
50	RBMY2B	RNA binding motif protein, Y chromosome, family 2, member B	GDB9848795	Yq11.23-Yq11.23
51	TBL1Y	Transducin (beta)-like 1Y-linked, TBL1	GDB11508487	Yp11.2-Yp11.2
52	RPS4Y	Hs.90653, Ribosomal protein S4, Y-linked	GDB128052	Yp11.3-Yp11.3*
53	RPS4Y2	Ribosomal protein S4, Y-linked 2	GDB11507771	Y-Y
54	RVNP2	Retroviral sequences NP2	GDB119582	Ypter-Yqter*
55	ZFY	Zinc finger protein, Y-linked	GDB 120503	Yp11.3- Yp11.3*
56	TGIF2LY	TGFB-induced factor 2-like, Y-linked	GDB 11508018	Y-Y
57	XKRY	X Kell blood group precursor-related, Y-linked	GDB9954843	Yq11.22-Yq11.22
58	XGR	Expression of XG and MIC2 on erythrocytes, YG	GDB120533	Xp22.32-Xp22.32 Yp11.3-Yp11.3
59	VCY2	Variable charge, Y chromosome, 2BPY2	GDB 10797002	Yq11-Yq11
60	SLC25A6	ANT3, Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 6, Hs.74550, adenine nucleotide translocator 3 (liver), ANT3Y	GDB125184	Xp22.32-Xp22.32 Yp11.3-Yp11.3*
61	SMCY	HY, Smcy homolog, Y chromosome (mouse), SMC (mouse) homolog, Y chromosome, Histocompatibility Y antigen, Selected mouse cDNA on the Y, HYA, KIAA0234	GDB5875390	Ycen-Yq11.23*
62	SRY	Sex determining region Y, TDF, Hs.1992, Testis determining factor	GDB 125556	Yp11.3-Yp11.3*
63	TTTY1	Testis-specific transcript, Y-linked-1 TTY1	GDB 11508086	Y-Y
64	TTTY2	Testis-specific transcript, Y-linked-2, TTY2	GDB11508100	Y-Y
65	TTTY3	Testis-specific transcript, Y-linked 3,TTY3	GDB11508102	Y-Y
66	TTTY4	Testis-specific transcript, Y-linked 4,TTY4	GDB11508104	Y-Y
67	TTTY5	Testis-specific transcript, Y-linked 5,TTY5	GDB11508106	Y-Y
68	TTTY6	Testis-specific transcript, Y-linked 6,TTY6	GDB11508108	Y-Y
69	TTTY7	Testis-specific transcript, Y-linked 7,TTY7	GDB11508110	Y-Y
70	TTTY8	Testis-specific transcript, Y-linked 8,TTY8	GDB11508112	Y-Y
71	TTTY9	Testis-specific transcript, Y-linked 9,TTY9	GDB11508114	Y-Y
72	TTTY10	Testis-specific transcript, Y-linked 10	GDB11508088	Y-Y
73	TTTY11	Testis-specific transcript, Y-linked 11	GDB11508090	Y-Y
74	TTTY12	Testis-specific transcript, Y-linked 12	GDB11508092	Y-Y

75	TTTY13	Testis-specific transcript, Y-linked 13	GDB11508094	Y-Y
76	TTTY14	Testis-specific transcript, Y-linked 14	GDB11508096	Y-Y
77	TTTY15	Testis-specific transcript, Y-linked 15	GDB11508098	Y-Y
78	TTTY16	Testis-specific transcript, Y-linked 16	GDB11510269	Y-Y
79	TTTY17	Testis-specific transcript, Y-linked 17	GDB11510271	Y-Y
80	TTTY18	Testis-specific transcript, Y-linked 18	GDB11510273	Y-Y
81	TTTY19	Testis-specific transcript, Y-linked 19	GDB11510275	Y-Y
82	TTTY20	Testis-specific transcript, Y-linked 20	GDB11510277	Y-Y
83	TTTY21	Testis-specific transcript, Y-linked 21	GDB11510279	Y-Y
84	TTTY22	Testis-specific transcript, Y-linked 22	GDB11510281	Y-Y
85.	TTTY23	Testis-specific transcript, Y-linked 23	GDB11510283	Y-Y
86.	FAM8A4P	Family with sequence similarity 8, member A4 pseudogene	GDB11503326	Ypter-Yqter
87	RPS24P1	RPS24P, Ribosomal protein S24 pseudogene 1, Ribosomal protein S24 pseudogene	GDB10795337	Yp11-Yp11
88	XGPY	Xg pseudogene (Y-linked)	GDB636807	Yq11.21-Yq11.21
89	STSP	Steroid sulfatase (microsomal), pseudogene	GDB119605	Yq11-Yq11*
90	RBMY-1A3P	RNA binding motif protein, Y chromosome, family 1, member A3 pseudogene	GDB11510200	Y-Y
91	PPP1R-12BP	Protein phosphatase 1, regulatory (inhibitor) subunit 12B Y-linked pseudogene	GDB11507462	Y-Y
92	OA1P	Ocular albinism 1 (Nettleship-Falls) Y- linked pseudogene	GDB11506359	Y-Y
93	KALP	KAL-Y,ADMLY, Kallmann syndrome sequence pseudogene	GDB128310	Yq11-Yq11* Yq11.2-Yq11.2
94	CASKP	Calcium/calmodulin-dependent serine protein kinase (MAGUK family) pseudogene	GDB9954903	Yq11-Yq11
95	SLC25-A15P	Solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 15 Y-linked pseudogene	GDB11507833	Y-Y
96	ZNF381P	Zinc finger protein 381, Y-linked pseudogene	GDB11508227	Y-Y
97	SEDLP7	Spondyloepiphyseal dysplasia, late, pseudogene 7	GDB10013966	Yq11.23-Yq11.23
98	SEDLP6	Spondyloepiphyseal dysplasia, late, pseudogene 6	GDB10013965	Yq11.23-Yq11.23
99	SEDLP3	Spondyloepiphyseal dysplasia, late, pseudogene 3	GDB10013962	Yq11.23-Yq11.23
100	SEDLP4	Spondyloepiphyseal dysplasia, late, pseudogene 4	GDB10013963	Yq11.23-Yq11.23
101	SEDLP5	Spondyloepiphyseal dysplasia, late pseudogene 5	GDB10013964	Yq11.23-Yq11.23
102	ASSP6	Argininosuccinate synthetase pseudogene-6	GDB119020	Ypter-Yqter*
103	APXLP	Apical protein-like (Xenopus laevis) pseudogene	GDB9954905	Yq11-Yq11
104	ARSDP	Arylsulfatase D pseudogene	GDB9954909	Yq11.21-Yq11.21
105	ARSEP	Arylsulfatase E pseudogene	GDB9954907	Yq11.21-Yq11.21
106	ACTGP2	ACTL2, Actin, gamma pseudogene 2	GDB120537	Yq11-Yq11*
107	ADPRT-L1P	ADP-ribosyltransferase (NAD+; poly (ADP-ribose) polymerase)-like 1 Y-linked pseudogene	GDB11505603	Y-Y

including testis (Repping et al. 2002). Interestingly, many of the testis specific genes are present in multiple copies ranging from 1-6, in addition to a number of pseudogenes. Studies on a total of 23 testis specific transcripts (TTTY1-23) have shown that TTTY4 expresses in the prostate and testis and TTTY10 expresses in testis, prostate and fetal brain whereas TTTY14 expresses in testis, kidney and fetal brain (Repping et al. 2002). Thus, some of the transcripts originally thought to be testis specific are transcribed in somatic tissues as well suggesting their broader biological functions, in addition to a possible regulatory role in spermatogenesis.

Identification of *SRY* (sex-determining-region Y) gene (Sinclair et al. 1990) rekindled the interest in other genes involved in regulation of spermatogenesis and this led to the identification of the RBMY (RNA-binding motif) gene (Ma et al. 1993). The next human spermatogenesis gene to be identified was DAZ (deleted in azoospermia) (Sargent 1999). First it was thought to be a single-copy gene but subsequent analysis showed that like RBMY, it belongs to a member of a gene family. DAZ may be a misnomer since there is no correlation between testis histopathology and DAZ deletion. This view is supported by the fact that

some men with DAZ deletion have oligozoospermia. This may be due to the presence of its autosomal homologue DAZL on chromosome 3 (mouse chromosome 17). The *DAZ* cluster has probably only recently made the evolutionary jump to the Y chromosome (Thomson et al. 2000) since the same is found only in humans and old-world monkeys but not new-world primates (Shan et al. 1996, Seboun et al. 1997). Both RBMY and DAZ encode proteins that are gamete specific and related to HnRNPG protein. The *HnRNPG* is a well-conserved ubiquitous gene that regulates features of RNA metabolism such as packaging, transport to the cytoplasm, and splicing. DAZ protein is found in the cytoplasm of male and female gametes and may regulate RNA translational events. RBMY protein is found in the nucleus and may regulate RNA splicing events.

Microdeletions occur in 3-15% of azoospermic or oligozoospermic men and in 2% of the fertile men (Pryor et al. 1997). Azoospermia factor (AZF) has three forms (AZFa, AZFb and AZFc) occupying well-defined Yq11 region. In the proximal region Yq11.21, AZFa harbors two genes, *USP9Y* and *DBY* whereas in the distal region Yq11.23, AZFb and AZFc harbor 7 (*CDY2*,*EIF1AY, HRY, RBMY, SMCY, TTY5 and TTY6*) and 10 (*BPY2, CDY1, CSPG4LY, DAZ1, DAZ2, DAZ3, DAZ4, GOLGA2LY, TTY3 and TTY4*) genes, respectively (Vogt & Fernandez 2003). So far, a few Y genes such as RBMY, DAZ and DFFRY (Drosophila fat-facets-related Y) have been associated with spermatogenesis (Brown et al. 1998). However,recent studies have shown that more genes are involved in regulation of spermatogenesis. This is supported by the fact that of the 23 testis specific transcripts (TTTY1-23), TTTY3, 4, 5, 6, 9, 10, 13 and 14 of the proximal and distal regions of palindromic complexes encompassing AZF factors showed deletions in the patients with spermatogenic failure (Kuroda-Kawaguchi et al. 2001). Most common defect is the microdeletion encompassing the *AZFc* region carrying four active polymorphic forms of DAZ genes on the Y chromosome. Much less common are microdeletions of the *AZFa* carrying the *DFFRY* (also know as *USP9Y*) and *DBY* (dead box on the Y) genes and of the *AZFb* area carrying the *RBMY* gene (for gene's relative position, see Tilford et al. 2001, Vogt & Fernandez 2003). AZFc microdeletions have been transmitted to sons through natural conception (Kobayashi et al. 1994) and the use of assisted reproductive techniques (Silber et al. 1998). Interestingly, proximal and distal regions of the Y chromosome have been found to harbor 10 kb each of the proviral sequences (Sun et al. 2000) of the recently described HERV15 class of endogenous retroviruses (Löwer et al. 1996, Patience et al. 1997). Recombination between these proviruses has been implicated to account for most of the *AZFa* deletions (Sun et al. 2000). Irrespective of mechanisms involved, deletion of any one or all the three types of azoospermia factors, *AZFa, AZFb or AZFc* disrupts spermatogenesis (Sun et al. 1999, Sun et al. 2000). However, it is not clear if a particular population is more prone to a specific type of AZF deletion. Kuroki et al. (1999) detected DNA variation without Y chromosome microdeletions and classified the subjects into four groups according to their Y haplotypes (I, II, III, and IV). The haplotype II men have lower sperm count than men in the other categories, and men with azoospermia were more likely to have haplotype III. Variations in the sperm counts in different populations indicate genetic basis of fertility potential of a specific population. Thus, it may not be unreasonable to expect a correlation between a given population and particular type of AZF deletion. Such information would prove to be helpful for routine diagnosis of the cases with male infertility factor(s).

Y Chromosome Polymorphisms and DNA Diagnosis

The human Y chromosome carries a wide variety of polymorphic systems, which may be exploited for genetic mapping, linkage analysis, human identity testing and studying human evolution (See Ali & Gangadharan 2000, Ali & Hasnain 2002).

In forensic analysis, well-established autosomal hypervariable marker systems, in particular multi-locus (Gill et al. 1985a) and single-locus fingerprinting (Debenham 1992), and STR multiplexes (Urquhart et al. 1995) have been used. These highly informative marker systems have been adopted even for the development of DNA Database (Gill & Werrett 1990b). The Y-chromosomal hypervariable system in conjunction with autosomal DNA evidence is useful

in specialized cases, such as sample mixtures and in estimating the number of assailants in multiple rapes. This is possible because all the members of a patrilineage are expected to share a Y chromosome haplotype (Jobling et al. 1997).

In evolutionary studies, a set of highly variable Y-specific tri- and tetranucleotide STRs, known as the 'seven-locus haplotype' (*DYS19*, *DYS389I*, *DYS389II*, *DYS390*, *DYS391*, *DYS392*, *DYS393*) has been widely adopted (Hurles et al. 1999, Zerjal et al. 1997, Pérez-Lezaun et al. 1997, Hurles et al. 1998, Ruiz-Linares et al. 1999). This is further augmented by the trinucleotide *DYS388* and the pentanucleotide *DXYS156Y* marker systems (Foster et al. 1998). Recently, typing of the two-loci tetranucleotide marker *DYS389* useful for analyzing four individual blocks of repeats has been refined (Rolf et al. 1998). An average mutation rate has been calculated to be 0.21% (95% confidence interval limits 0.06 - 0.49%) per locus per generation from a study of nine Y-STRs in deep-rooting pedigrees (Heyer et al. 1997). For the detection of sequence diversity, the Y-STR marker system has been found to be comparable to that of the autosomal loci (Weber & Wong 1993). The detection of mutation by Y-STRs has implications in family studies and for discriminating relationships in forensic cases (Foster et al. 1998, Kayser et al. 1997, Heyer et al. 1997) because one father-son pair with confirmed paternity has been found to show mutations at two of the thirteen STRs (Kayser et al. 1997). Thus, for establishing accurate genetic identity, multiplex PCR based on sufficient number of Y-STR marker systems may be used. This is important because the haplotypes defined by Y STRs are highly variable. In European populations, a given seven-locus haplotype is shared by an average of only 1.5 times with unrelated males, giving rise to a discriminatory power of 66% (n=322) (de Knijff et al. 1997). In a German population (n=70) the system provided 90% discrimination, and this was increased to over 97% with the inclusion of the bi-allelic trinucleotide marker *DYS385* (Kayser et al. 1997). These marker systems clearly show Y chromosome sequence heterogeneity amongst different populations of Europe. On the basis of trinucleotide STRs frequency in the human genome as a whole and presence of about 30 MB size of the euchromatin on the Y chromosome, about 170 Y-linked STR markers are envisaged to be present in the human genome (Jobling et al. 1997). How are these 170 STRs distributed on the Y chromosome amongst the global population is not clear at this stage. Nonetheless, a growing consensus is that many more STR based polymorphisms are likely to be uncovered in due course of time and some of them may be population specific. For example, hybridisation of $(GATA)_n$ repeat probes to an arrayed cosmid library indicated presence of about 210 of these tetranucleotides, and in the process, six male-specific GATA STRs were isolated from the individual cosmids (White et al. 1999, see also *"short tandem repeat DNA internet database"* (http//www.cstl.nist.gov/biotech/strbase/). Similarly, analysis of 3.4 kb DYZ1 fraction showed 229 units of pentanucleotide "TTCCA" motifs from which a synthetic oligonucleotide probe OAT20Y carrying four such consecutive units was developed that showed male specific band pattern (Ali et al. 1992). The probe OAT20Y was developed using sequence information available from the GenBank. Thus, exponentially growing resources data of DNA sequence on the Y chromosome emerging from the Human Genome Project (http//www.ncbi.nlm.nih.gov/genome/seq/) would prove to be a good source for isolation of new loci. The fact that more than 95% of the Y chromosome is haploid representing non-recombining regions (NRY), sequences from any of these regions are expected to be male specific. Further, if these NRYs harbor short tandem repeat sequences, these may instantly be used as powerful marker(s) not only for gender identification but also for assessing their fate (preferential loss or gain) in the aberrant Y chromosome. Using computer program "tandem repeats finder", 926 kb of Y-chromosomal sequences were screened *in silico*, and 18 tri- to hexanucleotide repeats with arrays of 8 or more units were identified (Benson 1999). Of these, six polymorphic loci (one tri-, four tetra-, and one pentanucleotide) were found to uncover male-specific band. In a multiplex typing system, these have been found to show 3-5 alleles per locus. A combination of the seven-loci haplotype with the new markers described above is likely to allow near individual identification, though will not exclude direct line

male relatives. As described above, biallelic markers of low mutation rate are useful for evolutionary studies because they allow delineation of the lineages (haplogroups). About 40 such PCR based markers encompassing SNPs are currently available. The ongoing search for SNPs is likely to exceed several folds in due course of time that may lead to narrowing the search of rare haplotypes. Based on extensive SNP studies conducted in the fertile and infertile men on AZFc, putative origin of molecular heterogeneity in the DAZ gene family has been reported (Vogt & Fernandez 2003). This substantiates growing importance of Y chromosome sequence polymorphism and its application. A few biallelic Y markers display transcontinental polymorphism in many populations (e.g. the YAP *Alu* element insertion) (Hammer 1994) and the M9 base substitution (Underhill et al. 1997). At the other extreme, a few markers show apparent individual specificity. The SNP SRY-3225 provides a good example of C to T transition. Initially the derived allele was found in a single individual of Melanesian origin (Whitfield et al. 1995a) and, after the screening of several hundred individuals (Hurles et al. 1998, Hammer et al. 1998, Thomas et al. 1998), no further examples was found suggesting this to be a case of unique SNP. Between these extremes, there are many markers, which show a high degree of population or continental specificity and are likely to be useful in determination of the origin of population. For example, the haplogroup defined by the G allele at the sY81 SNP (Seielstad 1994) is present in about 60% of the sub-Saharan African Y chromosomes (Hammer et al. 1998) but was detected only in two of the 3600 European Y chromosomes examined. The haplogroup defined by the YAP *Alu* element insertion, but lacking any of the derived alleles on the YAP branch of the Y haplotype tree represents about 20% of the East and Central Asian chromosomes (Hammer et al. 1998). However, this is completely absent from the European samples. Thus, population specificity for some markers is clearly seen. In yet another example, more than half of Finnish Y chromosomes have unique SNP (Zerjal et al. 1997), which is completely absent from a set of over 2000 chromosomes examined from the British Isles, France, Iberia and Italy. With the typing of more SNPs on large population samples, prediction on the origin of population will become possible. This will facilitate establishing a correlation between a specific population and disease phenotype.

DNA Based Sex Tests

DNA based sex test is used for gender identification in forensic and archaeological work (Santos et al. 1998). Identification of gender, based solely on chromosome anatomy, becomes ambiguous in the absence of a reliable genotype-phenotype correlation. An individual with XX chromosome constitution but male phenotype presents a socio-biological dilemma. These sex-reversed individuals show a sexual phenotype that is discordant with their karyotype - for e.g. 46, XX males and 46, XY females. Interestingly, the underlying molecular causes for the discordance are many (Ogata & Matsuo 1995). Furthermore, two individuals with an apparently identical genotype may have distinctly different phenotype. Such anomalies in the genotype-phenotype correlation put the very definition of the prevalent XY sex-determination in humans at stake (Wilson & Erlandsson 1998). This issue may become crucial also in the field of sport where sex of an individual is to be clearly defined for statutory purposes. Several methods and markers systems are available with their merits and demerits (table 2). Since different probes and marker systems correspond to different regions of the Y chromosome, they may judiciously be exploited to ascertain loss/gain or sequence modulation of given region(s) of the Y chromosome. Sexing of the human DNA using multiplex approach may also be used for ascertaining presence/absence or Y chromosome mosaicism in cases with or without cytogenetically detectable Y chromosome (Gauri et al. 1996).

This is important because individuals with partially deleted, aberrant, or even absent Y- chromosomes (45, X0, Turner's Syndrome) are often found to exist in the general population (Gicquel et al. 1992). In such cases, a specific sex test (e.g. amelogenin) produces misleading results. Commonly, the cause for the occurrence of XX male, with a frequency of 1 in 25,000 individuals, is the translocation of the Y- chromosome specific sex-determining gene *SRY*, to the X-chromosome

Table 2. *Sex-Typing Markers*

S.No	Locus (References)	Primers	Product size	Comments
1	Amelogenin (Sullivan et al. 1993)	5'-CCCTGGGCTCTGTAAAGAATAGTG-3' 5'-ATCAGAGCTTAAACTGGGAAGCTG-3'	X = 106 bp Y = 112 bp	May be multiplexed with STRs
2	Amelogenin (Sullivan et al. 1993)	5'-ACCTCATCCTGGGCACCCTGG-3' 5'-AGGCTTGAGGCCAACCATCAG-3'	X = 212 bp Y = 218 bp	May be multiplexed with STRs or D1S80
3	Amelogenin (Eng et al. 1994)	5'-CTGATGGTTGGCCTCAAGCCTGTG-3' 5'-TAAAGAGATTCATTAACTTGACTG-3'	X = 977 bp Y = 788 bp	May be analyzed with agarose gels
4	Centromeric alphoid repeat (Nasser & Gallati 1995)	5'-TATTTGGACTCTCTCTGAGGA-3' (X3) 5'-TTCTACTACAAGGGTGTTGCA-3' (X4) 5'-GTGTATTCACCTCCGGGAG-3' (Y3) 5'-ACAAAAGGTTCAATTCTGTGAG-3' (Y4)	X = 157 bp Y = 200 bp	Both X- and Y-sequences may be co- amplified
5	Centromeric alphoid repeat (Lin et al. 1995)	5'-AATCATCAAATGGAGATTTG-3' (X1) 5'-GTTCAGCTCTGTGAGTGAAA-3' (X2) 5'-ATGATAGAAACGGAAATATG-3' (Y11) 5'-AGTAGAATGCAAAGGGCTC-3' (Y22)	X = 170 bp Y = 130 bp	Separate PCR reactions are to be performed
6	ZFX/ZFY zinc finger gene (Reynolds & Varlaro 1996)	5'-CTGGAGAGCCACAAGCTGAC-3' 5'-TTGCTGTGGACTGCCAAGAG-3'	X/Y = 209bp after *Hae*III digestion X = 172 +37 Y = 88 +84 +37	Reverse dot blot typing assay; may be used with AmpliType PM and HLA-DQA1
7	SRY 93 (Santos et al. 1998)	5'-ATAAGTATCGACCTCGTCGGAAG-3' 5'-GCACTTCGCTGCAGAGTACCGAAG-3'	Y = 93 bp	Can be multiplexed with amelogenin
8	DYZ1 (Ali et al. 1992)	5'TTCCATTCCATTCCATTCCA3' Single oligo probe (OAT20Y)	3.4 kb band after digestion of male DNA with *Hae*III	Southern Hybridization with OAT20Y oligo probe
9	Y-Heterochromatic Sequence (Bashamboo & Ali 2001)	5'CACCTCTCCACCTGCC 3'	Uncovers 513 bp male specific isomorphic band	MASA reaction with 33.15 repeat locus

(Andersson et al. 1986) without an accompanying translocation of the *AMELY* gene (Lau et al. 1989). Therefore, agarose gel-phenotype of an amelogenin sex test, like the karyotype, will be of the "female" type, although the individual's phenotype is male. Similarly, XY females having lost the *SRY* gene to the X-chromosome through reciprocal translocation also exist, although at a lower frequency than their XX male counterparts. These XY females are usually rendered with complete gonadal dysgenesis, also known as 'streak' gonads (Swyer syndrome). Inevitably, these XY females show the "male" amelogenin gel-phenotype in concordance with their karyotype. Absolute gonadal dysgenesis in 46,XY individuals can also be caused by base substitution mutations in the *SRY* gene. To complicate matters, XY females with an intact Y- chromosome and the *SRY* gene also exist. Femaleness in these cases is assumed to have resulted from other loss-of-function mutations downstream in a pathway responsible for sexual differentiation. These individuals exhibit a wide range of phenotypes owing to varying underlying causes. For example, in the X-linked recessive disorder, Androgen Insensitivity Syndrome (formerly known as testicular feminization, with a frequency of 1 in 20,000 individuals), 46, XY individuals develop testes, but can also have normal female external genitalia (Wachtel 1998). Moreover, 46, XY individuals with small deletions of 9p region show normal female or even ambiguous genitalia; and 46, XY individuals with a duplication of a part of the short arm of the X-chromosome show sex reversal with partial gonadal dysgenesis (Huret et al. 1988, Hoo et al. 1989). Though rare, XX males lacking the *SRY* gene also exist, which is presumably due to a downstream gain-of-function mutation in other gene(s) responsible for constitutive testis differentiation (see also Ali & Hasnain 2002).

Unlike many sex-reversed cases, individuals carrying deletions of the *AMELY* gene are present as completely 'normal' males in the population. The frequencies of such deletions is however, altered by genetic drift. It is now evident that gender analysis, based solely on one type of sex test may be inaccurate. To circumvent this problem, a PCR based analysis of *SRY* gene or simultaneously other sex tests (see table 1) are recommended. It would certainly be useful for interpreting and delineating translocations in XX males and XY females. Nonetheless, individuals that manifest sex reversal without gain or loss of function-mutation in the *SRY* gene would continue to remain an intractable problem.

Concluding Remarks

The future research focused on the global expression of Y-linked genes having their homologs on the X chromosome in normal subjects and patients with sex chromosome related anomalies (e.g. turner syndrome) would prove to be useful for molecular diagnosis leading to genetic counseling. Functionally, the Y chromosome is a coherent entity and owing to its strikingly exceptional status, enjoys genetic supremacy and therefore will continue to guide the genetic destiny of males.

Acknowledgment

Part of the work reported here was supported by a DBT grant no. BT/PR2225/Med/13077/2000 to SA and an Institutional core grant from the Department of Biotechnology, Government of India. We thank Dr Dinkar Sahal for critically going through the manuscript, Shri Khem Singh Negi for technical assistance and four anonymous referees for their most valuable comments on this manuscript.

References

Ali S and Hasnain S E 2002 Molecular Dissection of the Human Y Chromosome; *Gene* **283** 1-10

_____ , Gauri Verma and Shashi Bala 1992 A Synthetic oligodeoxyribonucleotide probe $(5'TTCCA\ 3')_n$ uncovers male specific hybridization patterns in the human genome; *Mole. Cell. Probes* **6** 521-525

_____ and Gangadharan S 2000 Differential evolution of coding and non-coding sequences in related vertebrates, Implications in probe design; *Proc. Ind. Nat. Sci. Acad.* (PINSA) **B66** 49-67

Andersson M, Page D C and de la Chapelle A 1986 Chromosome Y-specific DNA is transferred to the short arm of X chromosome in human XX males; *Science* **233** 786-788

Bashamboo A and Ali S 2001 Minisatellite associated sequence amplification (MASA) of the hypervariable repeat marker 33.15 reveals a male specific band in humans; *Mol. Cell Probes* **15** 89-92

Benson G 1999 Tandem repeats finder a program to analyze DNA sequences; *Nucl. Acids Res.* **27** 573-580

Brown G M, Furlong R A, Sargent C A, Erickson R P, Longepied G, Mitchell M, Jones M H, Hargreave T B, Cooke H J and Affara N A 1998 Characterisation of the coding sequence and fine mapping of the human DFFRY gene and comparative expression analysis and mapping to the Sxrb interval of the mouse Y chromosome of the Dffry gene; *Hum. Mol. Genet.* **7** 97-107

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, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider P M, Szibor R, Teifel-Greding J, Weichhold G and Roewer L 1997 Chromosome Y microsatellites population genetic and evolutionary aspects; *Int. J. Legal Med.* **110** 134-140

Debenham P G 1992 Probing identity - the changing face of DNA fingerprinting; *Trends Biotech.* **10** 96-102

Delbridge M L, Harry J L, Toder R, O'Neill R J, Ma K, Chandley A C and Graves J A 1997 A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome; *Nat. Genet.* **15** 131-136

Eng B, Ainsworth P and Waye J S 1994 Anomalous migration of PCR products using nondenaturing polyacrylamide gel electrophoresis the amelogenin sex-typing system; *J. Forensic Sci.* **39** 1356-1359

Foster E A, Jobling M A, Taylor P G, Donnelly P, de Knijff P, Mieremet R, Zerjal T and Tyler-Smith C 1998 Jefferson fathered slave's last child; *Nature* **396** 27-28

Foster J W and Graves J A 1994 An SRY-related sequence on the marsupial X chromosome, implications for the evolution of the mammalian testis-determining gene; *Proc. Natl. Acad. Sci. U.S.A.* **91** 1927-1931

Ganesh S and Raman R 1997 CvSox-4, the lizard homologue of the human SOX4 gene, shows remarkable conservation among the amniotes; *Gene* **196** 287-290

Gauri, Shashi Bala, Jamal Ahmad and Sher Ali 1996 Genomic distribution of 5'TTCCA 3' repeat motif and its diagnostic potential in the human Y Chromosome related anomalies; *Clin. Genet.* **50** 358-365

Gicquel C, Cabrol S, Schneid H, Girard F and Le Bouc Y 1992 Molecular diagnosis of Turner's syndrome; *J. Med. Genet.* **29** 547-551

Gill P and Werrett D J 1990b Interpretation of DNA profiles using a computerised database; *Electrophoresis* **11** 444-448

_____, Jeffreys A J and Werrett D J 1985a Forensic application of DNA 'fingerprints'; *Nature* **318**

Graves J A and Foster J W 1994 Evolution of mammalian sex chromosomes and sex determining genes; *Int. Rev. Cytol.* **154** 191-259

_____ 1995 The origin and function of the mammalian Y chromosome and Y-borne genes- an evolving understanding; *Bioessays* **17** 311-320

Hammer M F 1994 A recent insertion of an Alu element on the Y chromosome is a useful marker for human population studies; *Mol. Biol. Evol.* **11** 749-761

_____ , Karafet T, Rasanayagam A, Wood E T, Altheide T K, Jenkins T, Griffiths R C, Templeton A R, Zegura S L 1998 Out of Africa and back again nested cladistic analysis of human Y chromosome variation; *Mol. Biol. Evol.* **15** 427-441

Heyer E, Pismire J, Deletes P, Baker E and de Knijff P 1997 Estimating Y chromosome specific microsatellite mutation frequencies using deep rooting pedigrees; *Hum. Mol. Genet*. **6** 799-803

Hoo J J, Salafsky I S, Lin C C and Pinsky L 1989 Possible location of a recessive testis forming gene on 9p24 (Abstract); *Am. J. Hum. Genet.* **45** (suppl.) A78

Huret J L, Leonard C, Forestier B, Rethore M O, Lejeune J 1988 Eleven new cases of del (9p) and features from 80 cases; *Am. J. Med. Genet.* **25** 741-749

Hurles M E, Irven C, Nicholson J, Taylor P G, Santos F R, Loughlin J, Jobling M A and Sykes B C 1998 European Y-chromosomal lineages in Polynesia a contrast to the population structure revealed by mitochondrial DNA; *Am. J. Hum. Genet.* **63** 1793-1806

_____, Veitia R, Arroyo E, Armenteros M, Bertranpetit J, Pérez-Lezaun A, Bosch E, Shlumukova M, Cambon Thomsen A, McElreavey K, López de Munain A, Röhl 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 a Y-chromosomal DNA polymorphism; *Am. J. Hum. Genet*. **65** 1437-1448

Jobling M A, Pandya A and Tyler-Smith C 1997 The Y chromosome in forensic analysis and paternity testing; *Int. J. Legal Med.* **110** 118-124

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, Perez-Lezaun A, Piccinini A, Prinz M, Schmitt C, Schneider P M, Szibor R, Teifel Greding J, Weichhold G, de Knijff P and Roewer L 1997 Evaluation of Y-chromosomal STRs a multicenter study; *Int. J. Legal Med.* **110** 125-133

Kobayashi K, Mizuno K, Hida A, Komaki R, Tomita K, Matsushita I, Namiki M, Iwamoto T, Tamura S and Minowada S 1994 PCR analysis of the Y chromosome long arm in azoospermic patients, evidence for a second locus required for spermatogenesis; *Hum. Mol. Genet.* **3** 1965-1967

Krausz C and McElreavey K 1999 Y chromosome and male infertility; *Front Biosci.* **4** E1- E8

Kuroda-Kawaguchi T, Skaletsky H, Brown L G, Minx P J, Cordum H S, Waterston R H, Wilson R K, Silber S, Oates R, Rozen S and Page D C 2001 *The AZFc* region on the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men; *Nature Genet.* **29** 279-286

Kurdki Y, Iwamoto T, Lee J, Yoshiike M, Nozawa S, Nishida T, Ewis A A, Nakamura H, Toda T, Tokunaga K, Kotliarova S E, Kondoh N, Koh E, Namiki M, Shinka T and Nakahori Y 1999 Spermatogenic ability is different among males in different Y chromosome lineage; *J. Hum. Genet.* **44** 289-292

Lahn B T and Page D C 1997 Functional coherence of the human Y chromosome; *Science* **278** 675-680

_____ and _____1999 Four evolutionary strata on the human X chromosome; *Science* **286** 964-967

_____, Pearson N M and Jegalian K 2001 The human Y chromosome in the light of Evolution; *Nat. Rev. Genet.* **2** 207-217

Lau E C, Mohandas T K, Shapiro L J, Slavkin H C and Snead M L 1989 Human and mouse amelogenin gene loci are on the sex chromosomes; *Genomics* **4** 162-168

Lin Z, Kondo, T, Minamino, T, Ohtsuji, M, Nishigami, J, Takayasu, T, Sun, R and Ohshima T 1995 Sex determination by polymerase chain reaction on mummies discovered at Taklamakan desert in 1912; *Forensic Sci. Int.* **75** 197-205

Löwer R Löwer J and Kurth R 1996 The viruses in all of us: characteristics and biological significance of human endogenous retrovirus sequences; *Proc. Natl. Acad. Sci. USA* **93** 5177-5184

Ma K, Inglis J D, Sharkey A, Bickmore W A, Hill R E, Prosser E J, Speed R M, Thomson E J, Jobling M, Taylor K, Wolfe J, Cooke H J, Hargreave T B and Chandley A C 1993 A Y chromosome gene family with RNA-binding protein homology candidates for the azoospermia factor AZF controlling human spermatogenesis; *Cell* **75** 1287-1295

Nasser D and Liechti-Gallati S 1995 Sex determination of forensic samples by simultaneous PCR amplification of microsatellite DNA from both the X and Y chromosomes; *J. Forensic Sci.* **40** 239-241

Ogata T and Matsuo N 1995 Turner syndrome and female sex chromosome aberrations deduction of the principal factors involved in the development of clinical feature; *Hum. Genet.* **95** 607-629

Parra E J, Marcini A, Akey L, Martinson J, Batzer M A, Cooper R, Forrester T, Allison DB Deka R, Ferrell R E and Shriver M D 1998 Estimating African American admixture proportions by use of population-specific alleles; *Am. J. Hum. Genet.* **63** 1839-1851

Patience C, Wilkinson D A, Weiss R A 1997 Our retroviral heritage; *Trends Genet.* **13** 116-120

Pérez Lezaun A, Calafell F, Seielstad M, Mateu E, Comas D, Bosch E and Bertranpetit J 1997 Population genetics of Y-chromosome short tandem repeats in humans ; *J. Mol. Evol.* **45** 265-270

Pryor J L, Kent-First M, Muallem A, Van Bergen A H, Nolten W E, Meisner L and Roberts K P 1997 Microdeletions in the Y chromosome of infertile men; *N. Engl. J. Med.* **336** 534-539

Quintana-Murci L, Krausz C, Heyer E, Gromoll J, Seifer I, Barton D E, Barrett T, Skakkebaek N E, Rajpert-De Meyts E, Mitchell M, Lee A C, Jobling M A and McElreavey K 2001 The relationship between Y chromosome DNA haplotypes and Y chromosome deletions leading to male infertility; *Hum. Genet.* **108** 55-58

Repping S, Skaletsky H, Lange J, Silber S, Van der Veen F, Oates R D, Page D C, Rozen S 2002 Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure; *Amer. J. Hum. Genet.* **71** 906-922

Reynolds R and Varlaro J 1996 Gender determination of forensic samples using PCR amplification of ZFX/ZFY gene sequences; *J. Forensic Sci.* **41** 279-286

Rolf B, Meyer E, Brinkmann B and deKnijff P 1998 Polymorphism at the tetranucleotide repeat locus DYS389 in 10 populations reveals strong geographic clustering; *Eur. J. Hum. Genet.* **6** 583-588

Ruiz Linares A, Ortíz Barrientos D, Figuerca M, Mesa N, Múnera J G, Bedoya G, Vélez I D, García L F, Pérez Lezaun A, Bertranpetit J, Feldman M W and Goldstein D B 1999 Microsatellites provide evidence for Y chromosome diversity among the founders of the New World; *Proc. Natl. Acad. Sci. USA* **96** 6312-6317

Santos F R, Pandya A, and Tyler-Smith C 1998 Reliability of DNA-based sex tests; *Nature Genet.* **18** 103

Sargent C A, Boucher C A, Kirsch S, Brown G, Weiss B, Trundley A, Burgoyne P, Saut N, Durand C, Levy N, Terriou P, Hargreave T, Cooke H, Mitchell M, Rappold G A and Affara N A 1999 The critical region of overlap defining the AZFa male infertility interval of proximal Yq contains three transcribed sequences; *J. Med. Genet.* **36** 670-677

Seboun E, Barbaux S, Bourgeron T, Nishi S, Agulnik A, Egashira M, Nikkawa N, Bishop C, Fellous M, McElreavey K, Kasahara M and Algonik A 1997 Gene Sequence, localization, and evolutionary conservation of DAZLA, a candidate male sterility gene; *Genomics* **45** 227-235

Seielstad M T, Hebert J M, Lin A A, Underhill P A, Ibrahim M, Vollrath D and Cavalli-Sforza L L 1994 Construction of human Y-chromosomal haplotypes using a new polymorphic A to G transition; *Hum. Mol. Genet.* **3** 2159-2161

Shan Z, Hirschmann P, Seebacher T, Edelmann A, Jauch A, Morell J, Urbitch P and Vogt H 1996 A SPGY copy homologous to the mouse gene *Dazla* and the *Drosophila* gene *boule* is autosomal and expressed only in the human male gonad; *Hum. Mol. Genet.* **5** 2005-2011

Silber S J, Alagappan R, Brown L G and Page D C 1998 Y chromosome deletions in azoospermic and severely oligozoospermic men undergoing intracytoplasmic sperm injection after testicular sperm extraction; *Hum. Reprod.* **13** 3332-3337

Sinclair A H, Berta P, Palmer M S, Hawkins J R, Griffiths B L, Smith M J, Foster J W, Frischauf A M, Lovell-Badge R and Goodfellow P N 1990 A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif; *Nature* **346** 240-244

Stevanovic M, Lovell-Badge R, Collignon J and Goodfellow P N 1993 SOX3 is an X-linked gene related to SRY; *Hum. Mol. Genet.* **2** 2013-2018

Sullivan K M, Mannucci A, Kimpton C P and Gill P 1993 A rapid and quantitative DNA sex test fluorescence-based PCR analysis of X-Y homologous gene amelogenin. *BioTechniques* **15** 637-641

Sun C, Skaletsky H, Birren B, Devon K, Tang Z, Silber S, Oates R and Page D C 1999 An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y; *Nature Genet.* **23** 429-432

Sun C, Skaletsky H, Rozen S, Gromoll J, Nieschlag E, Oates R and Page D C 2000 Deletion of Azoospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses; *Hum. Mol. Genet.* **9** 2291-2296

Thomas M G, Skorecki K, Ben-Ami H, Parfitt T, Bradman N and Goldstein D B 1998 Origins of Old Testament priests; *Nature* **394** 138-140

Thomson R, Pritchard J K, Shen P, Oefner P J, Feldman, M W 2000 Recent common ancestry of human Y chromosomes evidence from DNA sequence data; *Proc. Nat. Acad. Sci.* **97** 7360-7365

Tilford C A, Kuroda-Kawaguchi T, Skaletsky H, Rozen S, Brown L G, Rosenberg M, McPherson J D, Wylie K, Sekhon M, Kukaba T A, Waterston R H and Page D C 2001 A physical map of the human Y chromosome; *Nature* **409** 943-945

Underhill P A, Jin L, Lin A A, Mehdi S Q, Jenkins T, Vollrath D, Davis R W, Cavalli Sforza L L and Oefner P J 1997 Detection of numerous Y chromosome biallelic polymorphisms by denaturing high-performance liquid chromatography; *Genome Res.* **7** 996-1005

Urquhart A, Oldroyd N J, Kimpton C P and Gill P 1995 Highly discriminating heptaplex short tandem repeat PCR system for forensic identification; *Biotechniques* **18** 116

Vogt P H and Fernandes S 2003 Polymorphic DAZ gene family in polymorphic structure of AZFc locus: Artwork or functional for human spermatogenesis? *APMIS* **111** 1-13

Wachtel S S 1998 X-linked sex-reversing genes; *Cytogenet. Cell Genet.* **80** 222-225

Weber J L and Wong C 1993 Mutation of human short tandem repeats; *Hum. Mol. Genet.* **3** 1123-1128

White P S, Tatum O L, Deaven L L and Longmire J L 1999 New male-specific microsatellite markers from the human Y chromosome; *Genomics* **57** 433-437

Whitfield L S, Sulston J E and Goodfellow P N 1995a Sequence variation of the human Y Chromosome; *Nature* **378** 379-380

Wilson J F and Erlandsson R 1998 Sexing of human and other primate DNA; *Biol. Chem.* **379** 1287-1288

Zerjal T, Dashnyam B, Pandya A, Kayser M, Roewer L, Santos F R, Schiefenhovel 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; *Am. J. Hum. Genet.* **60** 1174-1183