

The role of the Y chromosome in male infertility

Nabeel A. Affara

It was suggested by Ronald Fisher in 1931 that genes that benefit the male (including those required for spermatogenesis) would accumulate on the Y chromosome. Following the discovery that microdeletions of the Y chromosome were associated with diverse spermatogenic phenotypes, at least three intervals that contain one or more genes controlling male germ-cell differentiation have been identified in humans. These intervals, named *AZF_a*, *AZF_b* and *AZF_c*, have been mapped, cloned and examined in detail for the presence of functional genes. In this review, I have discussed the genes that map to the *AZF* intervals and the evidence indicating which ones are the most likely candidates underlying Y-linked male infertility. In addition, I have considered the analysis of key intervals on the mouse Y chromosome, where it provides comparative data supporting the role of a candidate gene in an infertility phenotype.

It is estimated that as many as 10% of couples worldwide suffer from infertility or reduced fertility, and that in approximately half of these cases this results from defective spermatogenesis (Ref. 1). In 60% of these infertile men, the failure to produce mature germ cells (i.e. azoospermia) or the formation of low numbers of sperm (i.e. oligozoospermia) can be ascribed a genetic aetiology (Ref. 2). Few of the loci associated with male infertility have been mapped in humans; however, several genetic models of defective germ-cell development and differentiation have been described in mice (Refs 3, 4, 5). The schematic in Figure 1 summarises the main features of spermatogenesis in humans.

Deletions of the Y chromosome define the location of key genes in spermatogenesis

The involvement of the human Y chromosome in male infertility was originally suggested by cytogenetically detectable terminal deletions of the long arm of the chromosome, in which the entire Yq heterochromatin was lost (Ref. 6). This type of deletion does not appear to impair spermatogenesis in some men, but leads to infertility in others. The simple explanation for these observations is that there is a key locus (or loci) close to the boundary between genetically inert heterochromatin and Yq euchromatin. In some men, the removal of this locus by

Nabeel A. Affara

Reader in Molecular Genetics and Genomics, University of Cambridge, Department of Pathology, Tennis Court Road, Cambridge, CB2 1QP, UK. Tel: +44 (0)1223 333 777; Fax: +44 (0)1223 333 777; E-mail: na@mole.bio.cam.ac.uk

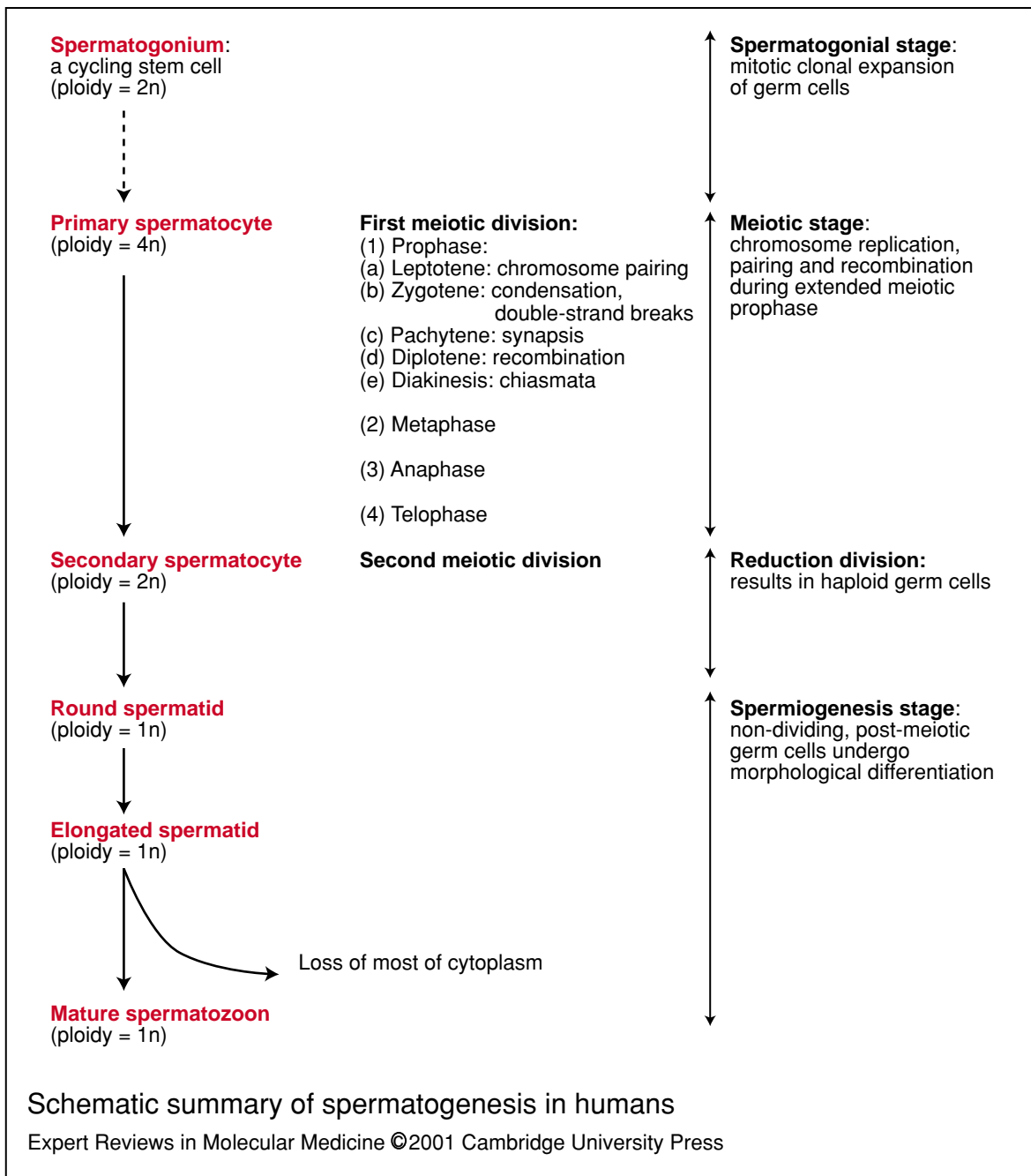


Figure 1. Schematic summary of spermatogenesis in humans. Spermatogenesis is the formation and development of spermatozoa, and includes spermatocytogenesis and spermiogenesis. Spermatogonia are cycling stem cells, which become type A and type B committed spermatogonia before giving rise to spermatocytes. After two meiotic divisions, spermatocytes develop into spermatids, which in turn develop into spermatozoa. These mature male germ cells develop in the seminiferous tubules of the testes, and consist of a head, a body, and a tail that provides propulsion. The head consists mainly of chromatin (i.e. nucleic acid and associated histone protein packed into nucleosomes) (fig001nac).

more-extensive deletions (which cannot be resolved by cytogenetic analysis) causes infertility. However, it could be argued that such large

structural changes to the Y chromosome (the Yq heterochromatin accounts for at least half of the long arm) might perturb normal pairing and

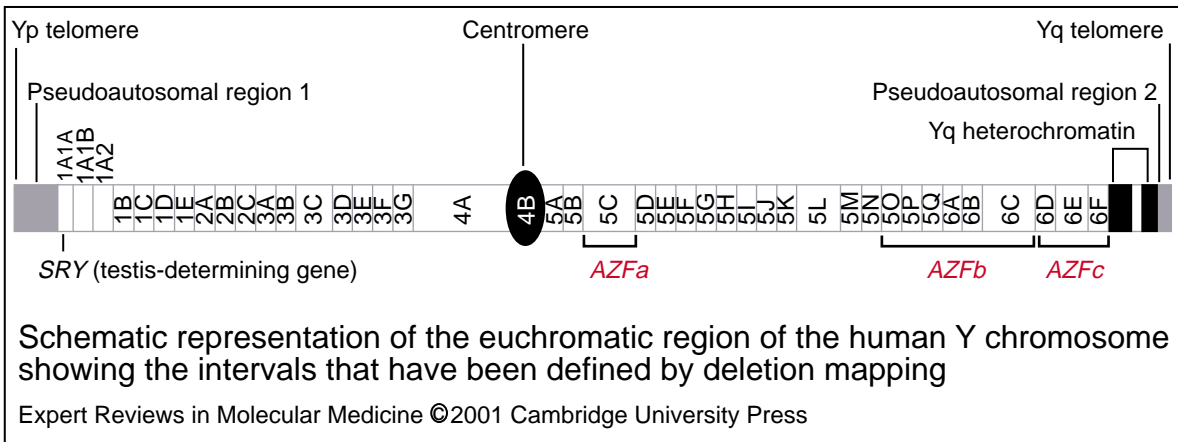


Figure 2. Schematic representation of the euchromatic region of the human Y chromosome showing the intervals that have been defined by deletion mapping. The chromosome is divided into deletion intervals, beginning with 1A1A, which contains the testis-determining gene *SRY*. Interval 4B is the centromere. The *AZFa*, *AZFb* and *AZFc* intervals are marked by brackets. Several spermatogenic phenotypes are associated with deletions in the *AZF* chromosomal segments. Azoospermia (the failure to produce mature spermatozoa), oligozoospermia (the production of reduced numbers of mature spermatozoa) and Sertoli cell only syndrome (the lack of detectable germ cells within the seminiferous tubules) are all associated with all three intervals, whereas meiotic and maturation arrest is associated with the *AZFb* and *AZFc* intervals. The pseudoautosomal regions (shaded gray) are those segments of the Y chromosome that are homologous to the X chromosome and pair with the X chromosome during meiosis in males. The solid black sections at the distal end of the Yq arm of the chromosome represent the Y heterochromatin, which is composed of two tandem repeats (**fig002nac**).

segregation with the X chromosome during meiosis, and that this, rather than the loss of a key locus, is the cause of spermatogenic failure in these men.

Definitive genetic evidence that loci on the long arm of the human Y chromosome are required for the differentiation of germ cells has come from studies using molecular markers (such as sequence-tagged sites) that have been mapped at high density and ordered along the Y chromosome. At least three distinct, non-overlapping intervals, each associated with variable degrees of spermatogenic impairment, have been defined by numerous deletion-mapping studies (Refs 7, 8, 9, 10, 11, and references cited therein). These intervals, named *AZFa*, *AZFb* and *AZFc* for azoospermia factors a, b and c (see Fig. 2), indicate that at least three different loci on the long arm of the Y chromosome are critical for germ-cell differentiation. Similar deletion-analysis studies have defined critical intervals on the mouse Y chromosome. These studies have been reviewed elsewhere (Ref. 12), and have been considered here only if they are relevant to the analysis of the human Y chromosome.

Genotype–phenotype correlations

It is possible to make genotype–phenotype correlations in those cases where a patient's testicular biopsy material is available, thus allowing the spermatogenic phenotype to be directly examined. Such studies of patients who have deletions of *AZFa*, *AZFb* or *AZFc* have revealed considerable variations in the nature and degree of spermatogenic failure.

Deletions of the *AZFa* region (the least common class of deletion) are most frequently associated with azoospermia manifested by Sertoli cell only syndrome (SCOS) and less often with oligozoospermia (Refs 8, 13, 14). SCOS is characterised by the complete absence of germ cells within seminiferous tubules. Two SCOS phenotypes have been described: SCOS I, where none of the tubules in a histological section contains germ cells; and SCOS II, where some of the tubules are devoid of germ cells and others contain reduced numbers (Ref. 8). The SCOS I phenotype is likely to arise from a primary defect in early spermatogenesis, whereas the SCOS II phenotype is believed to develop at a later stage of differentiation as a secondary effect of a primary defect that affects spermatogenesis.

Deletions of the *AZFb* region are associated with azoospermia, oligozoospermia and normozoospermia (normal spermatogenesis; Refs 8, 9, 15). Azoospermia is accompanied not only by SCOS, but also by the arrest of meiosis and maturation at the spermatocyte and spermatid stages of male germ-cell differentiation.

Deletions of the *AZFc* interval are associated with azoospermia, severe-to-mild oligozoospermia and the production of insufficient mature sperm to enable reproduction (Refs 8, 16, 17). As for deletions of *AZFb*, the azoospermia reflects a spectrum of spermatogenic phenotypes from SCOS to meiotic and maturation arrest.

From these genotype–phenotype correlations, it is evident that similar or identical deletions in patients can cause different impairments in spermatogenesis. There could be several explanations for this phenotypic variation. First, it is notoriously difficult to map the Y chromosome because of the accumulation on this chromosome of amplified families of DNA sequences. It is, therefore, difficult to be sure that a marker has only one location on the chromosome. Clearly, this could lead to errors when deletions are scored: what might appear to be similar deletions could, in fact, be different. Second, variation between individuals might reflect the modification of phenotype as a consequence of X–Y mosaicism in the germline, resulting in some germ cells possessing an intact Y chromosome. Such mosaicism might not be present in the lymphocytes that are used to prepare DNA from these patients. Third, variation in phenotype could also reflect the presence, within the genetic background of the individual, of alleles of modifying genes that are able to compensate for the loss of function of Y-chromosome-linked genes.

The Y chromosome has been mapped in considerable detail, using a combination of deletion-mapping studies that define specific intervals, and a series of overlapping yeast artificial chromosome (YAC), P1-derived artificial chromosome (PAC) and bacterial artificial chromosome (BAC) clones that cover the entire active, euchromatic DNA (Refs 18, 19, 20). This has enabled a detailed search for candidate genes within the clones covering the key intervals that are associated with the disease phenotypes. A great deal of effort has been placed on the identification of candidate genes that map to the *AZFa*, *AZFb* and *AZFc* intervals; these

mapped genes are considered below in historical order.

Candidate genes that map to *AZF* intervals

The *AZFb* interval

The first Y-chromosome-linked gene to be isolated as a strong candidate for a role in spermatogenesis is from the *AZFb* interval. Ma and colleagues (Ref. 21) described two transcripts mapping to this region that are part of a larger gene family located on the Y chromosome. The transcripts encoded potential RNA-binding proteins related to the autosomal gene *HnRNP G*, and contained a single N-terminal RNA-binding motif (known as the RNP motif) and four tandem repeats of 37 amino acid residues (known as the SRGY box, which is rich in serine, arginine, glycine and tryptophan residues) in the central portion of the proteins. This family of genes, initially known as *YRRM* (for Y RNA recognition motif) but now named *RBM* (for RNA-binding motif), contains at least 30 genes and pseudogenes, which fall into different subclasses that map to several locations on the short and long arms of the human Y chromosome, including the *AZFb* interval (Ref. 22). Similar studies in mice have shown that there are multiple copies of the *Rbm* gene family on the short arm of the Y chromosome; however, unlike the human *RBM* gene family, they are clustered at a single location just proximal to the *Sry* sex-determining gene (see later, and Refs 23, 24). In both species, it is not clear how many of the genes within the family are functional. Analysis of genomic and complementary DNA (cDNA) clones and reverse transcriptase-polymerase chain reaction (RT-PCR) products indicates that there could be six subclasses of *RBM* genes in humans (*RBMI–RBMVI*, renamed *RBMY1–RBMY6*); only members of the *RBMY1* subclass are transcribed (Ref. 25). The pattern of expression of the *RBMY1* genes is complex. Nucleotide substitutions and alternative splicing result in several different transcripts that might lead to the synthesis of protein isoforms that have three or four SRGY boxes (Ref. 25). This complexity makes functional analysis difficult.

Antibodies to *RBM* show that, in human testes, the protein is localised to the nucleus, confined to germ cells and is present at all stages of spermatogenesis except in elongating spermatids (Ref. 26). The analysis of *RBM* protein in testicular sections from infertile men indicated

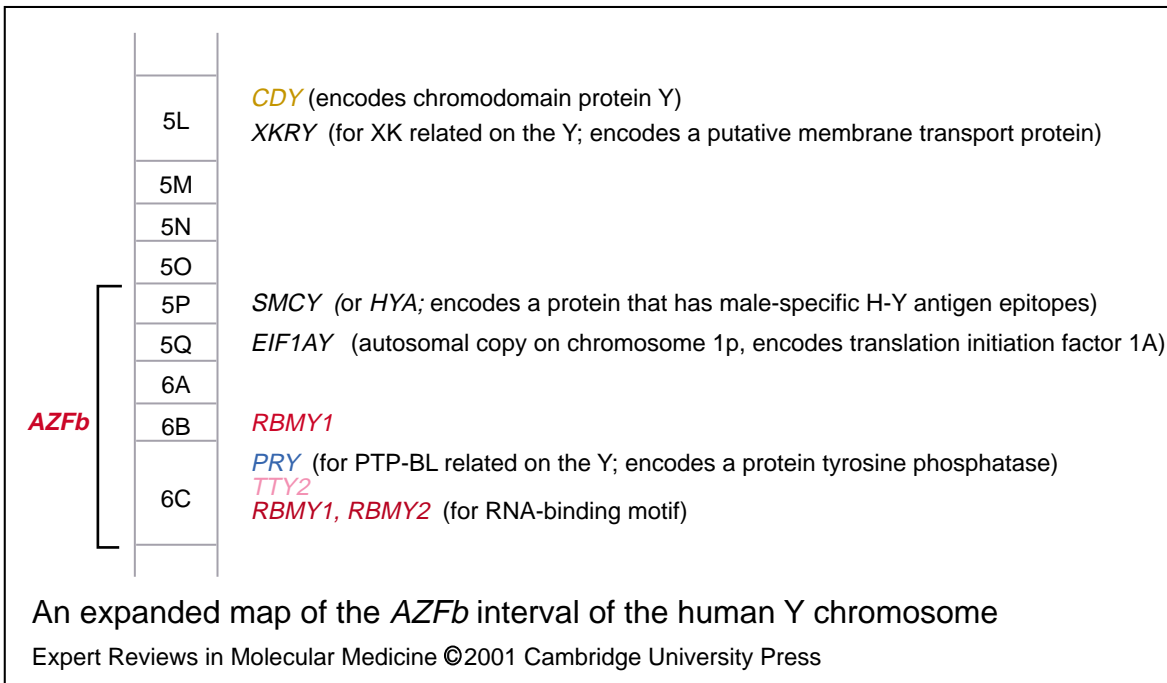


Figure 3. An expanded map of the *AZFb* interval of the human Y chromosome. Genes, in addition to *RBM1* sequences, that map to this region are shown. The different colours represent deletion intervals (fig003nac).

a correlation between active protein expression and the presence of the *AZFb* region, particularly interval 6B on the Y chromosome deletion map (Ref. 26; see Fig. 3). This is an important result because it suggests that the key *RBM* genes lie in the *AZFb* interval and that other *RBM* loci on the human Y chromosome do not encode proteins that are essential for spermatogenesis. In the mouse, *Rbmy* transcripts were also shown to have a pattern of expression that is specific to germ cells. However, unlike the human gene(s), expression is switched off in pachytene spermatocytes and can be detected in elongating spermatids (Ref. 27).

It is likely that *RBMY1* functions as an RNA-binding protein and its location in the nucleus in association with splicing factors (Ref. 28) implies that it might have a role in RNA metabolism, perhaps in RNA processing. Post-transcriptional and post-translational control is important in spermatogenesis (for review, see Ref. 29), especially during post-meiotic stages, where transcriptional activity has almost ceased. *RBMY1* is present throughout most of spermatogenesis, indicating that the protein has a role at many stages. However, because *AZFb* deletions abolish the expression of *RBMY1* and cause meiotic arrest, this protein does not appear

to be essential for germ cells to reach meiosis but does seem to be necessary for them to pass successfully to the haploid state. In mice, the deletion of *Rbmy* is associated with maturation defects, indicating that the protein might be essential for the post-meiotic maturation of germ cells and that it could be required for post-transcriptional control (Ref. 27).

Comparative analysis of *RBM* and related genes in different eutherian mammals and marsupials has revealed details of the evolution of this family of genes (Ref. 30). The original data indicated that the Y-borne genes arose by the translocation of an ancestral, autosomal *HnRNP G* gene onto the Y chromosome; this was followed by independent amplification events in different species. Two amplification events can be distinguished: (1) an internal amplification of the exon that encodes the SRGY box in humans, giving rise to four tandem repeats; and (2) amplification of the whole gene, leading to the multicopy family. In mice, the *Rbm* gene on the Y chromosome has been amplified, but not the SRGY box. In some marsupials, it appears that the gene has been amplified on the Y chromosome, but that there is a single copy of the SRGY box. Because different members of the human gene family

contain the four, tandem SRGY-box repeats, these findings suggest that the amplification of the SRGY box occurred after the divergence of the human and mouse lineages and before amplification of the gene on the Y chromosome. That *RBMY1* genes have an important function is suggested by the fact that they are conserved on the Y chromosome of both eutherian mammals and marsupials. This finding is consistent with these genes having a role in an essential function such as spermatogenesis.

The recent discovery of a widely expressed *RBMY1* homologue that contains introns on the X chromosome both of mice (Ref. 31) and of humans (at Xq26, and named *RBMX*; Ref. 32), and that the autosomal loci (on chromosome 6 in humans and 14 in mice) represent X-derived pseudogenes (Refs 31, 32) has altered our thinking about the evolution of this gene family. It is clear that these genes date from the ancestral homologues that gave rise to the sex chromosomes and that the autosomal loci have been generated by retrotransposition. The locus on the Y chromosome has been amplified and selected for a role in spermatogenesis during evolution.

The fact that the *RBMY1* genes are a multicopy family makes it difficult to provide direct genetic and functional evidence that they are critical to spermatogenesis. There is strong, indirect evidence to support their involvement in germ-cell differentiation, based primarily on the correlation between the *AZFc* interval and expression of the gene encoding the *RBMY1* protein in germ cells. Figure 3 shows the region of the Y chromosome in which *AZFc* deletions have been mapped and the genes that are mapped in this interval. Several questions arise when attempting to correlate genotype to phenotype. How many *RBM* genes are involved in the critical deletions that abolish *RBM* protein production? What contribution (if any) do other genes mapped to the *AZFc* interval make to spermatogenic phenotypes that are associated with this region? What other genes are encompassed by the *AZFc* interval? Two X–Y homologous genes that are expressed ubiquitously have been mapped to the *AZFc* region: *SMCY*, which encodes a protein that has male-specific H-Y antigen epitopes (Ref. 33); and *EIF1AY*, which encodes a translation initiation factor (Ref. 34). Two other transcripts, *TTY2* and *PRY*, both members of gene families that are amplified on the Y chromosome, have also been

assigned to the *AZFc* interval (Ref. 34). No functional open reading frames are present in *TTY2* transcripts; *PRY* is related to *PTP-BL*, which encodes a protein tyrosine phosphatase.

How can these questions be resolved? Refined deletion mapping using a larger group of patients is one approach to isolate single genes within critical intervals. Mutation analysis in non-deletion cases of male infertility would provide direct genetic evidence of the role of a gene in an associated phenotype. This is more complex in multicopy gene families because it requires the analysis of all family members. In the case of *RBM* genes, analysis could be restricted to genes mapping to the critical region that is required for protein expression. Functional analysis using mouse models is another approach, whereby different strategies can be used to inactivate the relevant genes and to observe the phenotypic consequences, or to attempt phenotypic rescue by introducing functional genes into a mouse in which the *Rbmy* genes have been deleted.

The *AZFc* interval

The first suggestion of a candidate gene in the *AZFc* interval came from a systematic exon-trapping analysis (involving the use of a special vector system that can identify and isolate exonic sequences from genomic DNA) of DNA clones covering about 500 kilobases (kb) of the region. Reijo and colleagues (Ref. 17) isolated expressed exons that formed part of a gene, named *DAZ* (for deleted in azoospermia), that contains an N-terminal RNA-recognition motif and a repetitive internal structure composed of a series of seven tandem repeats of 72 base pairs. These features indicate that the gene encodes an RNA-binding protein that might have a regulatory role in RNA metabolism. Initially, *DAZ* was thought to be a single-copy gene, and the only functional gene in the *AZFc* interval (Ref. 17). However, subsequent analysis of the *DAZ* locus has demonstrated that it represents a clustered, amplified gene family and that the number of tandem repeats varies considerably between family members; there is also polymorphism within the population in the number and sequence of *DAZ* repeats (Refs 35, 36). The fact that a diverse spectrum of *DAZ* transcripts is produced suggests that all members of the gene family might be transcribed (Ref. 36).

DAZ is expressed specifically in testes (Refs 17, 37) and is closely related to a fertility gene, named *boule*, in *Drosophila* (Ref. 38). Mutations

that disrupt the function of the *boule* gene cause azoospermia. Thus, it seems reasonable to expect that *DAZ* could be involved in the *AZFc* phenotype(s), although it is not necessarily the only gene. The protein encoded by *DAZ* is found mainly in post-meiotic germ cells (late spermatids and sperm; Ref. 39) and this might indicate that it has a role in post-transcriptional control during the transcriptionally inactive stages of germ-cell differentiation.

An autosomal homologue of *DAZ* that contains a single *DAZ* repeat has been described in both humans and mice, but no *DAZ*-related sequences have been found on the mouse Y chromosome. The human gene (named *DAZLA*), which maps to chromosome 3p24, is expressed specifically in the testes and, at lower levels (detectable by RT-PCR), in the ovaries (Refs 35, 40, 41, 42). This observation has prompted the suggestion that the *DAZLA* gene might be involved in autosomal recessive forms of male infertility. The mouse gene (named *Dazla* for *DAZ* like, autosomal) maps to chromosome 17 and is also expressed in male germ cells and the female gonads (Refs 43, 44) but, unlike *Rbm*, is located in the cytoplasm rather than in the nucleus (Ref. 45). The temporal pattern of *Dazla* expression indicates that it is present throughout spermatogenesis and, thus, might be critical for all stages of germ-cell differentiation. Reijo and colleagues (Ref. 44) suggested that this might explain the range of phenotypes that is associated with *AZFc* deletions; however, it is not yet known if the transcript from the human Y chromosome has the same pattern of expression and, thus, it is premature to extrapolate from observations made for the mouse autosomal gene. Disruption of the mouse *Dazla* gene has provided functional evidence that the gene plays a critical role in the production of both male and female germ cells (Ref. 45). The loss of *Dazla* function completely abolishes gamete production, demonstrating that it is essential for successful gametogenesis.

As mentioned previously, the *DAZ* gene family differs from the *RBM* gene family, in that no *DAZ* genes have been found on the Y chromosome of mice or any non-primate species studied (Refs 35, 42, 43). Furthermore, there is no evidence of the existence of homologous sequences on the X chromosome in humans and mice. *DAZ* sequences are relatively recent arrivals to the Y chromosome, having appeared on the non-recombining region of the primate

Y chromosome after the divergence of the New World monkeys but before the splitting of the lineage that gave rise to Old World monkeys. It has been suggested that this occurred by direct transposition from the autosomal locus to the non-recombining part of the Y chromosome, and subsequent amplification (Ref. 35). Thus, this amplified gene family has a very different evolutionary history to that of the *RBM* gene family. Two points are worth noting in relation to this finding. First, Lahn and Page (Ref. 34) have suggested that amplified genes on the Y chromosome were recruited from an autosomal homologue. However, the *DAZ* and *RBM* gene families illustrate that this is not the case, and that a second class of Y-chromosome-linked, amplified gene family can arise from an ancestral homologue of the X chromosome. This might also be the case for other Y-linked, amplified gene families. Second, genes that are homologous between the X and Y chromosomes, not just Y-specific genes, must be considered as candidates that might be involved in the regulation of spermatogenesis.

Although the *DAZ* gene would appear to be involved in the *AZFc* phenotype, efforts to confirm this are dogged by the same problems that confront a functional analysis of the *RBM* genes. No discrete mutations have been found in *DAZ* genes that are associated with infertility, and the existence of a gene family hinders efforts to establish such genetic evidence. The picture becomes even more complex as a consequence of new data showing that the *DAZ* gene family is not the only one to map to the *AZFc* region (Ref. 34). A recent map of this region is summarised in Figure 4, which shows that *AZFc* deletions occur in at least three other gene families: *BPY2* transcripts encode a basic protein of unknown function, *PRY* is mentioned above, and *CDY* contains a chromodomain (chromatin-binding domain) and a potential CoA-substrate catalytic domain. Members of the *CDY1* gene family might modify chromatin during spermatogenesis. Stuppia and colleagues (Ref. 46) have reported deletions of the *AZFc* region that cause infertility. These do not involve the deletion of *DAZ* genes but might remove other genes, such as *CDY1* and *BPY1*. This finding suggests that, in addition to *DAZ*, other loci in the *AZFc* interval contribute to fertility, and this might partially explain the diverse phenotypes that are associated with different deletions.

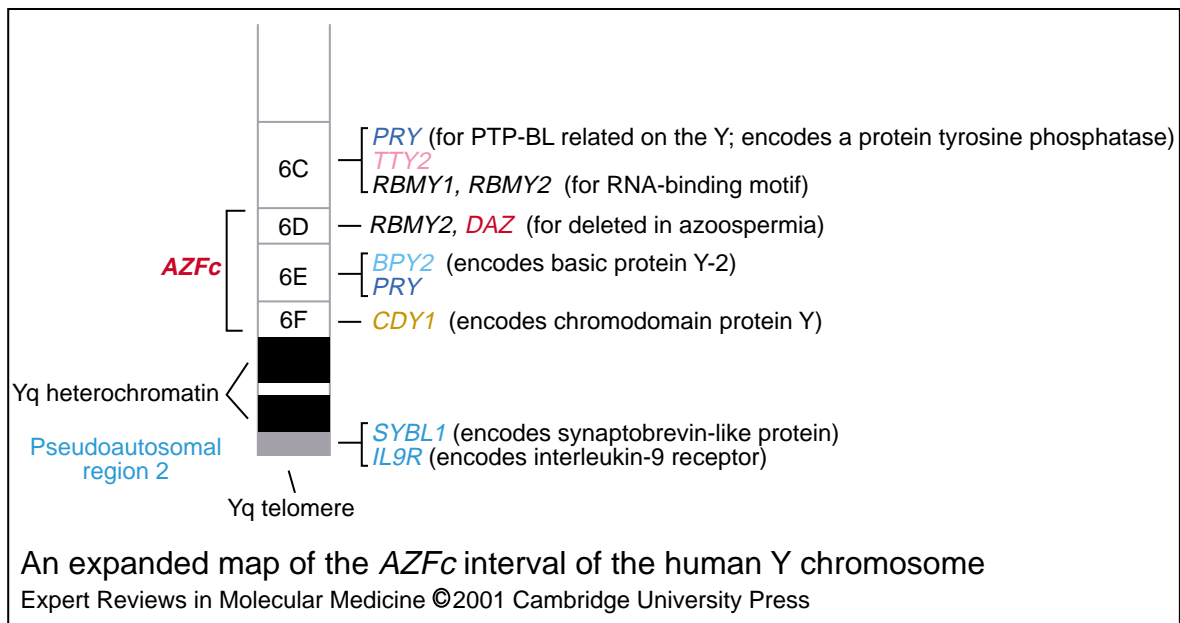


Figure 4. An expanded map of the AZFc interval of the human Y chromosome. Genes, in addition to those of the *DAZ* family, that map to this region are shown. The different colours represent different gene families (fig004nac).

Despite their distinct evolutionary histories, the parallels between *RBM* and *DAZ* genes are striking: both genes encode RNA-binding proteins and are likely to be involved in RNA metabolism; both genes have undergone internal amplification of a single repeat motif in primates; both genes have undergone gene amplification to give rise to a Y-linked multicopy family; and, finally, both genes are expressed in a testis-specific manner.

The AZFa interval

The *AZFa* interval is the most proximal region of the long arm of the human Y chromosome to be associated with male infertility. Several deletion studies have defined a segment of approximately 400 kb that correlated with the *AZFa* phenotype in Yq11.2 (Refs 8, 47, 48). The frequency of deletions defining *AZFa* is lower than that of both *AZFb* and *AZFc*. With the exception of one study (Ref. 49), all the laboratories that have reported *AZFa* deletions agree that 1–2% of Y-chromosome deletions involve the *AZFa* region. Recently, just over 1 megabase (Mb) of DNA spanning the *AZFa* interval has been sequenced by the Whitehead Institute for Medical Research/MIT Center for Genome Research, which has allowed the accurate positioning of breakpoints in relation to functional genes that have been

mapped to the region. These genes include the *DFFRY* gene (Refs 34, 47, 48, 50), the *DBY* gene (Ref. 34) and the *UTY* gene (Refs 34, 51). All three genes have ubiquitously expressed X homologues (*DFFRX*, *DBX* and *UTX*) that map in the Xp11.3–p11.4 interval and escape X-inactivation (i.e. both X-linked copies are transcriptionally active in females). The *DFFRY* (the Y-linked homologue of *DFFRX*, the *Drosophila* fat facets related X gene; it is also known as *USP9Y*) gene is also expressed ubiquitously and encodes a ubiquitin-specific hydrolase that is closely related to the *Drosophila* gene *faf* (for fat facets; Ref. 52). The *faf* gene is involved in photoreceptor and oocyte development, and the ubiquitin-dependent degradation of protein, where it targets specifically the D-Jun transcription factor and other members of the signal transduction cascade involved in regulating D-Jun activity (Ref. 53). The *DBY* (for DEAD box on the Y) gene encodes a protein containing the DEAD box motif that is characteristic of RNA helicases and, therefore, might have a role in RNA metabolism. Two *DBY* transcripts are evident, one of which is found specifically in testes (Ref. 34). The function of the product of the *UTY* (for ubiquitous transcribed tetratricopeptide repeat gene on the Y) gene is unknown, but it cannot be involved in the *AZFa* phenotype because it is expressed in all patients

studied to date. One further transcribed sequence, designated AZFaT1, has been assigned to the deletion interval that is just proximal to the *DFFRY* gene (Ref. 48). The complete sequence of this transcript remains to be established and, thus, it is not known whether it represents a functional gene. Detailed analysis of the sequence covering the *AZFa* interval by Sargent and colleagues (Ref. 48) and Sun and colleagues (Ref. 54) has also identified several pseudogenes that are related to gene sequences mapping to Xp22.3.

Our present knowledge of the map of the *AZFa* region is summarised in Figure 5, which shows the location of key deletion breakpoints in four infertile patients (Ref. 48) and the relationship of this region to homologous regions of the X chromosome. The position of the breakpoints indicates the involvement of the *DFFRY* and *DBY* genes in the *AZFa* phenotype. Of the four patients, one (SAYER) has a milder oligozoospermic phenotype. This patient has a deleted *DFFRY* gene and a transcriptionally active *DBY* gene (Ref. 48), whereas both these genes are deleted in the other three, suggesting that the more-severe SCOS phenotype is produced only if both of these genes are deleted. More patients must be studied to investigate this possibility, particularly cases where there is genetic evidence of point mutations affecting both genes. A recent study of 500 infertile men described one case of a deletion of 4 base pairs in the *DFFRY* gene that affected a splice donor site in intron 7 and led to premature termination of the polypeptide chain (Ref. 54). This patient's phenotype was similar to that of SAYER, supporting the theory that the *DFFRY* and *DBY* genes might act synergistically to enable full fertility. None of the 500 patients studied by Sun and colleagues (Ref. 54) showed evidence of mutations in the *DBY* gene.

The mouse *Sxr^b* region shares homology with the human *AZFa* interval

A spermatogenic phenotype in mice (known as *Sp^y*) that resembles the *AZFa* phenotype in humans is associated with the deletion of the *Sxr^b* region of the short arm of the mouse Y chromosome (Ref. 55). In this phenotype, spermatogonial proliferation is blocked at an early stage, and few germ cells progress beyond this stage of differentiation. The mouse orthologue of the *DFFRY* gene (i.e. *Dffry*) is located in the *Sxr^b* region (Ref. 47). Recently, comparative mapping studies have revealed that the mouse orthologues

of *DBY* (i.e. *Dby*) and *UTY* (i.e. *Uty*) are also located in the *Sxr^b* region, and that the centromere-to-telomere order (*DFFRY-DBY-UTY*) is the same on both the mouse and human chromosomes. This finding indicates that the *Sxr^b* interval in mice and the proximal Yq11.2 region in humans represent a conserved, syntenic segment (Ref. 56, and see Fig. 6). Furthermore, this suggests that these genes represent an ancient organisation on the Y chromosome that predates the divergence of the human and mouse lineages, and further supports the idea that shared, conserved genes may underlie the spermatogenic phenotypes observed in both species. Southern blot analysis of mouse genomic DNA using a human AZFaT1 probe has revealed no evidence for the conservation of the AZFaT1 sequence on the mouse Y chromosome (M. Mitchell, INSERM, Marseille, France, pers. commun.).

Studies of the expression of the *Dffry* (Ref. 47) and *Dby* (Ref. 56) genes in the mouse have shown that, in contrast to *DFFRY* and *Dby*, *Dffry* is transcribed specifically in testis messenger RNA (mRNA), thus making it a good candidate for involvement in the *Sp^y* and *AZFa* phenotypes. Temporal expression studies of testis mRNA post-partum have shown that *Dffry*, but not *Dby*, transcripts correlated with the appearance of germ cells, and are first evident 7.5–10.5 days post-partum, when type A and B spermatogonia and preleptotene and leptotene spermatocytes are present (Refs 47, 48). Furthermore, the analysis of testis mRNA from XX*Sxr^a* mice, which are devoid of germ cells, showed that the expression of *Dffry* was not detectable in the absence of germ cells, whereas the expression of *Dby* was unaffected (Ref. 48). Thus, the expression of *Dffry* is either confined to the germ cells or occurs in supporting cells in a germ-cell-dependent manner. The fact that the expression of *Dby* is independent of the presence of germ cells within the testis does not rule out the expression of *Dby* at low levels either in these cells during the early stages of differentiation or in supporting cell lineages that are important for correct germ-cell differentiation. It is possible that the combined action of both these genes also underlies the *Sp^y* phenotype, in a similar manner to that suggested for the human genes and the full *AZFa* phenotype.

Intriguingly, the *Ube1y* gene maps within the *Sxr^b* deletion interval. It, too, is an X–Y homologous gene and the Y chromosome copy is

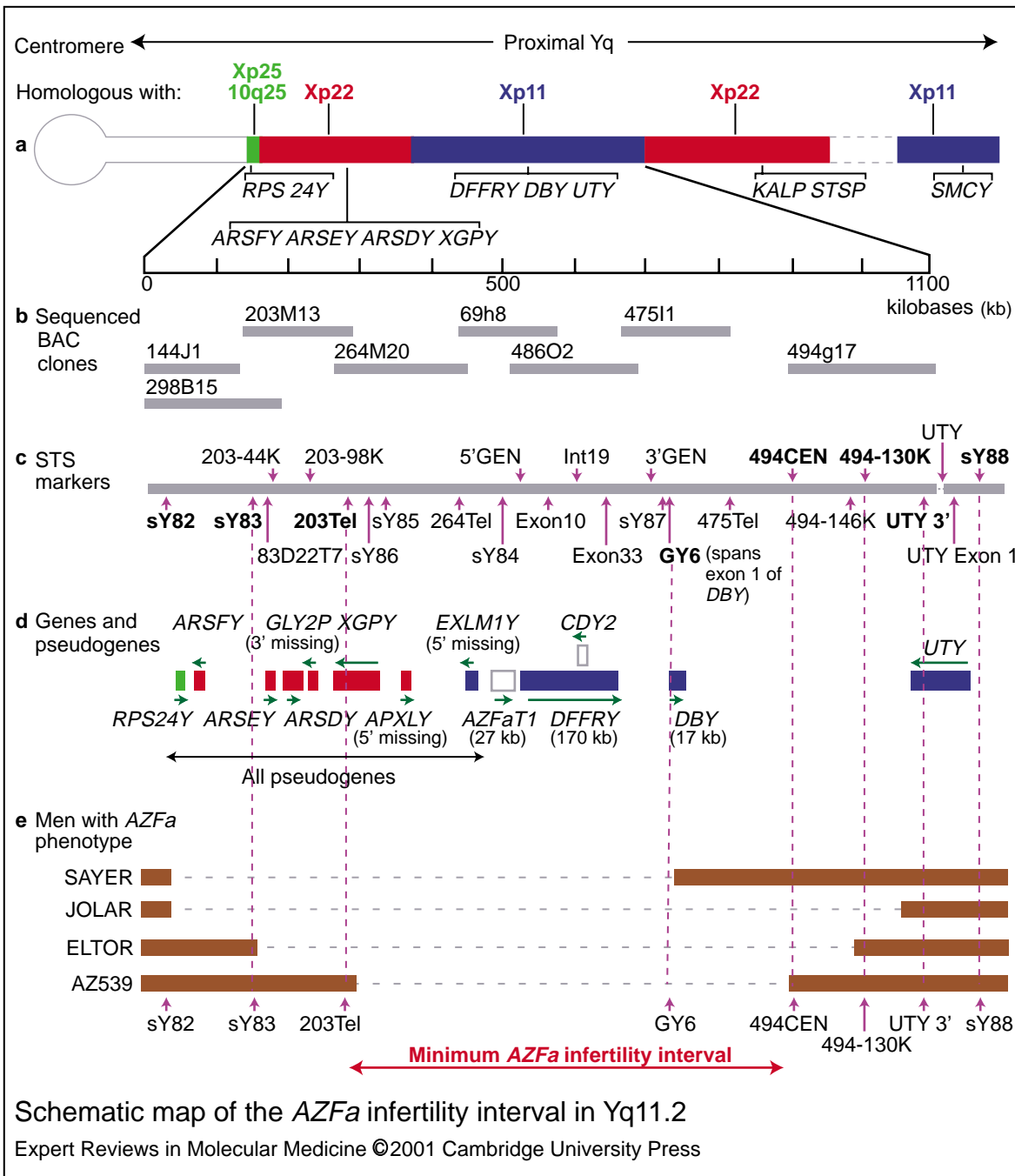


Figure 5. Schematic map of the AZFa infertility interval in Yq11.2. The map shows, from top to bottom: (a) the pattern of homology of this part of the Y chromosome to regions of the X chromosome; (b) the tiling path of the bacterial artificial chromosome (BAC) clones used to provide about 1 Mb (megabase) of DNA sequence; (c) a selection of STS markers (sequence-tagged sites) spanning the interval; (d) genes and pseudogenes mapped in relation to STS markers and BAC clones; and (e) the position of deletion breakpoints found in infertile patients who have a deletion in the AZFa region, defined using STS markers mapping across the interval. The STS markers are stretches of unique genomic sequence for which a polymerase chain reaction assay has been developed. They, therefore, act as markers for a unique site within the genome (fig005nac).

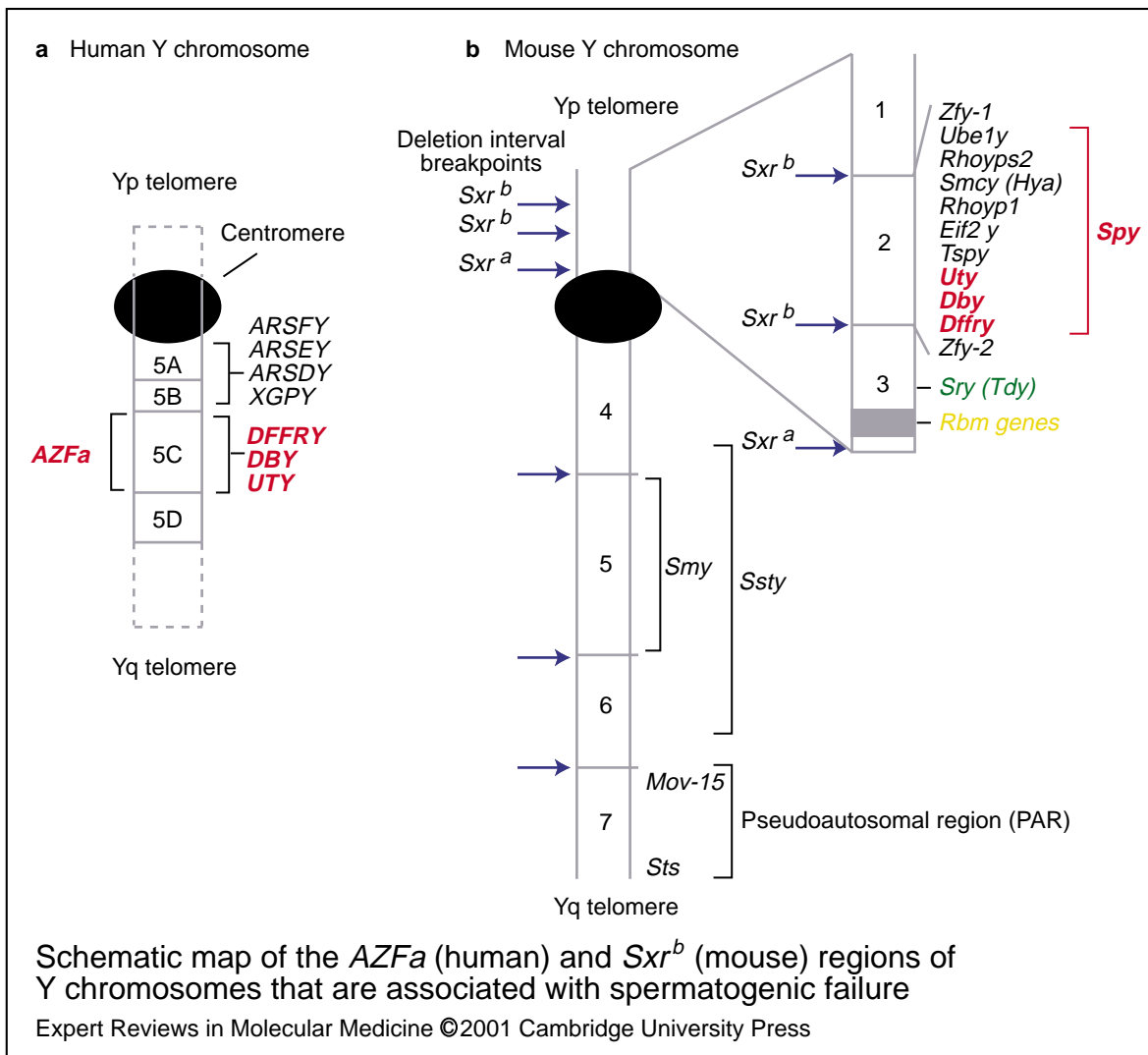


Figure 6. Schematic map of the AZFa (human) and Sxr^b (mouse) regions of Y chromosomes that are associated with spermatogenic failure. The comparison illustrates the conserved order of the human *DFFRY/DBY/UTY* and mouse *Dffry/Dby/Uty* genes, shown in red and bold. *DFFRY*, for *Drosophila fat facets* related gene; *DBY*, for Dead box on the Y; *UTY*, for ubiquitous transcribed tetratricopeptide repeat gene on the Y; *Smy*, for sperm head morphology; *Spy*, blocks spermatogonial proliferation (**fig006nac**).

expressed in a testis-specific manner (Refs 57, 58). The proteins Ube1y and Dffry are required for the addition and removal of ubiquitin from proteins, respectively (Ube1y is a ubiquitin-activating enzyme; Dffry is a ubiquitin hydrolase). Ubiquitination marks a protein for degradation, which occurs after ubiquitin has been removed, just before entry into the 26S proteasome. The protein encoded by *DFFRY* or *Dffry* belongs to the class of ubiquitin hydrolases that are associated with the 26S proteasome and remove ubiquitin as proteins enter the complex for degradation. This promotes efficient ubiquitin-dependent

degradation. It is possible that Ube1y and Dffry act together in germ cells to permit fine control over the turnover of one or more proteins that are necessary for successful spermatogenesis.

Why have spermatogenesis genes accumulated and become amplified on the Y chromosome?

The molecular genetic analysis of the mammalian Y chromosome has revealed several distinct themes: (1) genes involved in reproductive performance appear to have accumulated on this chromosome; (2) some of these genes have

been recruited from autosomal and X-linked homologues during evolution, whereas others are derived from ancestral progenitors dating from the proto-X and proto-Y chromosomes; and (3) there is a propensity for genes to become amplified on the Y chromosome.

In 1931, Fisher argued that, what he termed, male-benefit genes would accumulate on the Y chromosome (Ref. 59). Where this provided a reproductive advantage to the male, selection would have acted to maintain these genes in a genetically isolated region of the genome, thus favouring the suppression of recombination for most of the length of the Y chromosome and the transmission of an unchanged chromosome in the male line. Such genes are not necessarily concerned exclusively, or directly, with the production of germ cells; they could also encode genetic functions that enhance the reproductive performance of the male, such as size and attractiveness to a mate. Muller and Charlesworth have both argued that the rapid degeneration of genes in the non-recombining region of the Y chromosome follows from the failure to reconstruct defective genes through recombination (Refs 60, 61). Thus, once established, deleterious mutations in genes on the Y chromosome become fixed irreversibly in the male population (known as Muller's Ratchet) and will accumulate, leading to the degeneration of most of the genes on the Y chromosome. Rice demonstrated this process elegantly in a *Drosophila* model system in which the removal of a chromosomal segment from the recombination system caused the rapid degeneration of genes contained within it (Ref. 62).

These unique features perhaps partially explain why the genes on the Y chromosome have become amplified. Thus, the selection of multiple copies of a Y-linked gene that is critical for spermatogenesis might be favoured because it increases the amount of the protein product which, owing to the accumulation of deleterious mutations, is produced in insufficient amounts. In contrast, the absence of pairing and recombination for members of an amplified gene family producing an altered gene product might reduce selection against the gene and lead to its retention. Other selective pressures favouring the amplification of genes that are involved in spermatogenesis might also arise from the need to produce large amounts of

protein or a combination of variant proteins that act synergistically to ensure efficient gametogenesis.

Future directions

Several challenges now face those researching spermatogenesis genes on the Y chromosome. Biochemical and functional studies of those genes that have been identified in the *AZFa*, *AZFb* and *AZFc* regions are under way to gain insight into their roles in regulating germ-cell differentiation. What are the nucleic acid targets of the RBM and DAZ proteins? What are their precise biochemical functions? Do they interact with other proteins that are involved in RNA metabolism? What protein(s) does DFFRY regulate? For those genes that are X–Y homologous, is there a difference in function between the X-linked genes and the Y-linked genes? Several other multicopy gene families on the Y chromosome have not yet been studied in depth. How many of these are involved in spermatogenesis? How have they arrived on the Y chromosome? Do they follow the *DAZ* model or the *RBM* model? What is their function? The field is moving fast and will expand rapidly in the next few years to provide the beginnings of a molecular understanding of the genetic pathways controlling male germ-cell differentiation.

There are two main clinical implications of this work. First, a description of the genes that cause some cases of infertility provides the basis for diagnosis at the molecular level. Thus, patients can be screened for deletions of the Y chromosome to establish the causative lesion. Second, an understanding of the structure and function of the genes involved in male infertility and with which genes they interact will facilitate the development of potential therapies. Based on an understanding of the biochemical function(s) and pathways with which these genes are involved, it may be possible to develop drugs that permit spermatogenesis to proceed normally.

Acknowledgements and funding

Nabeel Affara is supported by grants from both the Biotechnology and Biological Sciences Research Council and the Wellcome Trust, and would like to thank Michael Mitchell (INSERM, Marseille, France) and Carole Sargent (University of Cambridge, UK) for peer review of the manuscript and their helpful suggestions.

References

- 1 Vaessen, M. (1984) Childlessness and Infecundity. WFS Comparative Studies, Note 2, Series 31, Cross National Summaries, International Statistic Institute, Voorburg, The Netherlands
- 2 Lilford, R. et al. (1994) Case-control study of whether subfertility in men is familial. *Bmj* 309, 570-573, PubMed ID: 94372746
- 3 Chubb, C. (1989) Genetically defined mouse models of male infertility. *J Androl* 10, 77-88, PubMed ID: 89233971
- 4 Lamb, D.J. and Niederberger, C.S. (1994) Animal models that mimic human male reproductive defects. *Urol Clin North Am* 21, 377-387, PubMed ID: 94337480
- 5 Okabe, M., Ikawa, M. and Ashkenas, J. (1998) Male infertility and the genetics of spermatogenesis. *Am J Hum Genet* 62, 1274-1281, PubMed ID: 98302852
- 6 Tiepolo, L. and Zuffardi, O. (1976) Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Hum Genet* 34, 119-124, PubMed ID: 77070291
- 7 Vogt, P. et al. (1992) Microdeletions in interval 6 of the Y chromosome of males with idiopathic sterility point to disruption of AZF, a human spermatogenesis gene. *Hum Genet* 89, 491-496, PubMed ID: 92340073
- 8 Vogt, P.H. et al. (1996) Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* 5, 933-943, PubMed ID: 96414299
- 9 Pryor, J.L. et al. (1997) Microdeletions in the Y chromosome of infertile men. *N Engl J Med* 336, 534-539, PubMed ID: 97164732
- 10 Kent-First, M. et al. (1999) Defining regions of the Y-chromosome responsible for male infertility and identification of a fourth AZF region (AZFd) by Y-chromosome microdeletion detection. *Mol Reprod Dev* 53, 27-41, PubMed ID: 99245845
- 11 Elliott, D.J. and Cooke, H.J. (1997) The molecular genetics of male infertility. *Bioessays* 19, 801-809, PubMed ID: 97443093
- 12 Burgoyne, P.S. (1998) The role of Y-encoded genes in mammalian spermatogenesis. *Semin Cell Dev Biol* 9, 423-432, PubMed ID: 99033111
- 13 Qureshi, S.J. et al. (1996) Polymerase chain reaction screening for Y chromosome microdeletions: a first step towards the diagnosis of genetically-determined spermatogenic failure in men. *Mol Hum Reprod* 2, 775-779, PubMed ID: 97383715
- 14 Foresta, C. et al. (1997) Y-chromosome deletions in idiopathic severe testiculopathies. *J Clin Endocrinol Metab* 82, 1075-1080, PubMed ID: 97255229
- 15 Girardi, S.K., Mielnik, A. and Schlegel, P.N. (1997) Submicroscopic deletions in the Y chromosome of infertile men. *Hum Reprod* 12, 1635-1641, PubMed ID: 97452370
- 16 Simoni, M. et al. (1997) Screening for deletions of the Y chromosome involving the DAZ (Deleted in AZoospermia) gene in azoospermia and severe oligozoospermia. *Fertil Steril* 67, 542-547, PubMed ID: 97205571
- 17 Reijo, R. et al. (1995) Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* 10, 383-393, PubMed ID: 95400318
- 18 Foote, S. et al. (1992) The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* 258, 60-66, PubMed ID: 93068245
- 19 Jones, M.H. et al. (1994) A set of ninety-seven overlapping yeast artificial chromosome clones spanning the human Y chromosome euchromatin. *Genomics* 24, 266-275, PubMed ID: 95213016
- 20 Affara, N. et al. (1996) Report of the Second International Workshop on Y Chromosome Mapping 1995. *Cytogenet Cell Genet* 73, 33-76, PubMed ID: 96244593
- 21 Ma, K. et al. (1993) A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor AZF controlling human spermatogenesis. *Cell* 75, 1287-1295, PubMed ID: 94094323
- 22 Prosser, J. et al. (1996) Degeneracy in human multicopy RBM (YRRM), a candidate spermatogenesis gene. *Mamm Genome* 7, 835-842, PubMed ID: 97032533
- 23 Laval, S.H. et al. (1995) Y chromosome short arm-Sxr recombination in XSxr/Y males causes deletion of Rbm and XY female sex reversal. *Proc Natl Acad Sci U S A* 92, 10403-10407, PubMed ID: 96036093
- 24 Elliott, D.J. et al. (1996) An RBM homologue maps to the mouse Y chromosome and is expressed in germ cells. *Hum Mol Genet* 5, 869-874, PubMed ID: 96414293
- 25 Chai, N.N., Salido, E.C. and Yen, P.H. (1997) Multiple functional copies of the RBM gene

- family, a spermatogenesis candidate on the human Y chromosome. *Genomics* 45, 355-361, PubMed ID: 98008925
- 26 Elliott, D.J. et al. (1997) Expression of RBM in the nuclei of human germ cells is dependent on a critical region of the Y chromosome long arm. *Proc Natl Acad Sci U S A* 94, 3848-3853, PubMed ID: 97268659
- 27 Mahadevaiah, S.K. et al. (1998) Mouse homologues of the human AZF candidate gene RBM are expressed in spermatogonia and spermatids, and map to a Y chromosome deletion interval associated with a high incidence of sperm abnormalities. *Hum Mol Genet* 7, 715-727, PubMed ID: 98167859
- 28 Elliott, D.J. et al. (1998) Dynamic changes in the subnuclear organisation of pre-mRNA splicing proteins and RBM during human germ cell development. *J Cell Sci* 111, 1255-1265, PubMed ID: 98215736
- 29 Hecht, N.B. (1998) Molecular mechanisms of male germ cell differentiation. *Bioessays* 20, 555-561, PubMed ID: 98390244
- 30 Delbridge, M.L. et al. (1997) A human candidate spermatogenesis gene, RBM1, is conserved and amplified on the marsupial Y chromosome [published erratum appears in *Nat Genet* 1997 Apr; 15(4): 411]. *Nat Genet* 15, 131-136, PubMed ID: 97172969
- 31 Mazeyrat, S. et al. (1999) RBMY evolved on the Y chromosome from a ubiquitously transcribed X-Y identical gene. *Nat Genet* 22, 224-226, PubMed ID: 99318091
- 32 Delbridge, M.L. et al. (1999) The candidate spermatogenesis gene RBMY has a homologue on the human X chromosome. *Nat Genet* 22, 223-224, PubMed ID: 99318090
- 33 Agulnik, A.I. et al. (1994) A novel X gene with a widely transcribed Y-linked homologue escapes X-inactivation in mouse and human. *Hum Mol Genet* 3, 879-884, PubMed ID: 95038739
- 34 Lahn, B.T. and Page, D.C. (1997) Functional coherence of the human Y chromosome. *Science* 278, 675-680, PubMed ID: 98022381
- 35 Saxena, R. et al. (1996) The DAZ gene cluster on the human Y chromosome arose from an autosomal gene that was transposed, repeatedly amplified and pruned. *Nat Genet* 14, 292-299, PubMed ID: 97051923
- 36 Yen, P.H., Chai, N.N. and Salido, E.C. (1997) The human DAZ genes, a putative male infertility factor on the Y chromosome, are highly polymorphic in the DAZ repeat regions. *Mamm Genome* 8, 756-759, PubMed ID: 97468274
- 37 Menke, D.B., Mutter, G.L. and Page, D.C. (1997) Expression of DAZ, an azoospermia factor candidate, in human spermatogonia. *Am J Hum Genet* 60, 237-241, PubMed ID: 97136593
- 38 Eberhart, C.G., Maines, J.Z. and Wasserman, S.A. (1996) Meiotic cell cycle requirement for a fly homologue of human Deleted in Azoospermia. *Nature* 381, 783-785, PubMed ID: 96267004
- 39 Habermann, B. et al. (1998) DAZ (Deleted in AZoospermia) genes encode proteins located in human late spermatids and in sperm tails. *Hum Reprod* 13, 363-369, PubMed ID: 98216918
- 40 Yen, P.H., Chai, N.N. and Salido, E.C. (1996) The human autosomal gene DAZLA: testis specificity and a candidate for male infertility. *Hum Mol Genet* 5, 2013-2017, PubMed ID: 97123509
- 41 Shan, Z. et al. (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, PubMed ID: 97123508
- 42 Seboun, E. et al. (1997) Gene sequence, localization, and evolutionary conservation of DAZLA, a candidate male sterility gene [published erratum appears in *Genomics* 1997 Oct 15; 45(2): 477]. *Genomics* 41, 227-235, PubMed ID: 97288520
- 43 Cooke, H.J. et al. (1996) A murine homologue of the human DAZ gene is autosomal and expressed only in male and female gonads. *Hum Mol Genet* 5, 513-516, PubMed ID: 96254982
- 44 Reijo, R. et al. (1996) Mouse autosomal homologue of DAZ, a candidate male sterility gene in humans, is expressed in male germ cells before and after puberty. *Genomics* 35, 346-352, PubMed ID: 96299784
- 45 Ruggiu, M. et al. (1997) The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. *Nature* 389, 73-77, PubMed ID: 97433329
- 46 Stuppia, L. et al. (1998) A quarter of men with idiopathic oligo-azoospermia display chromosomal abnormalities and microdeletions of different types in interval 6 of Yq11. *Hum Genet* 102, 566-570, PubMed ID: 98316778
- 47 Brown, G.M. et al. (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, PubMed ID: 98046012

- 48 Sargent, C.A. et al. (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, PubMed ID: 99435579
- 49 Foresta, C., Ferlin, A. and Moro, E. (2000) Deletion and expression analysis of AZFa genes on the human Y chromosome revealed a major role for DBY in male infertility. *Hum Mol Genet* 9, 1161-1169, PubMed ID: 20231543
- 50 Jones, M.H. et al. (1996) The *Drosophila* developmental gene fat facets has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2 [published erratum appears in *Hum Mol Genet* 1997 Feb; 6(2): 334-5]. *Hum Mol Genet* 5, 1695-1701, PubMed ID: 97081753
- 51 Greenfield, A. et al. (1996) An H-YDb epitope is encoded by a novel mouse Y chromosome gene. *Nat Genet* 14, 474-478, PubMed ID: 97099462
- 52 Fischer-Vize, J.A., Rubin, G.M. and Lehmann, R. (1992) The fat facets gene is required for *Drosophila* eye and embryo development. *Development* 116, 985-1000, PubMed ID: 93202020
- 53 Isaksson, A. et al. (1997) The deubiquitination enzyme fat facets negatively regulates RTK/Ras/MAPK signalling during *Drosophila* eye development. *Mech Dev* 68, 59-67, PubMed ID: 98092101
- 54 Sun, C. et al. (1999) An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat Genet* 23, 429-432, PubMed ID: 20047771
- 55 Sutcliffe, M.J. and Burgoyne, P.S. (1989) Analysis of the testes of H-Y negative XOS^{xrb} mice suggests that the spermatogenesis gene (*Spy*) acts during the differentiation of the A spermatogonia. *Development* 107, 373-380, PubMed ID: 90214522
- 56 Mazeyrat, S. et al. (1998) The mouse Y chromosome interval necessary for spermatogonial proliferation is gene dense with syntenic homology to the human AZFa region. *Hum Mol Genet* 7, 1713-1724, PubMed ID: 98409500
- 57 Mitchell, M.J. et al. (1991) Homology of a candidate spermatogenic gene from the mouse Y chromosome to the ubiquitin-activating enzyme E1. *Nature* 354, 483-486, PubMed ID: 92086048
- 58 Odoriso, T. et al. (1996) Transcriptional analysis of the candidate spermatogenesis gene *Ube1y* and of the closely related *Ube1x* shows that they are coexpressed in spermatogonia and spermatids but are repressed in pachytene spermatocytes. *Dev Biol* 180, 336-343, PubMed ID: 97105847
- 59 Fisher, R.A. (1931) The evolution of dominance. *Biol Rev* 6, 345-368
- 60 Muller, H.J. (1964) The relation of recombination to mutational advance. *Mut Res* 1, 2-9
- 61 Charlesworth, B. (1978) Model for evolution of Y chromosomes and dosage compensation. *Proc Natl Acad Sci U S A* 75, 5618-5622, PubMed ID: 79074872
- 62 Rice, W.R. (1994) Degeneration of a nonrecombining chromosome. *Science* 263, 230-232, PubMed ID: 94112549

Further reading, resources and contacts

Mitchell, M.J. (2000) Spermatogenesis and the mouse Y chromosome: specialisation out of decay. *Results Probl Cell Differ* 28, 233-270, PubMed ID: 20091746

Bishop, C.E. and Mitchell, M.J. (1999) Mouse Y chromosome. *Mamm Genome* 10, 962, PubMed ID: 99431661

Krausz, C. and McElreavey, K. (1999) Y chromosome and male infertility. *Frontiers Biosci* 4, 1-8

The British Andrology Society (BAS) is a multidisciplinary group of scientists and clinicians working in the fields of human and mammalian reproduction who have an interest in the male. Research plays a prominent role within the society; areas of interest include: spermatogenesis, semen analysis, sperm function, fertilisation, contraception and cryopreservation.
<http://www.repromed.org.uk/bas>

The Society for the Study of Fertility promotes the study of the biological, medical and agricultural aspects of fertility, particularly by holding conferences at which scientists, medical and veterinary clinicians and other research workers can meet, present information and exchange and discuss ideas germane to the understanding and regulation of fertility. The proceedings are published in the *Journal of Reproduction and Fertility*.
<http://www.ssf.org.uk>

The British Fertility Society (BFS) comprises clinicians, scientists, nurses, counsellors and other interested parties working in the fields of human fertility and subfertility, reproductive medicine and reproductive biology throughout the UK. The BFS holds two scientific meetings each year and publishes its own journal, *Human Fertility*, which is distributed to members.
<http://www.britishfertilitysociety.org.uk>

Features associated with this article

Figures

Figure 1. Schematic summary of spermatogenesis in humans (fig001nac).

Figure 2. Schematic representation of the euchromatic region of the human Y chromosome showing the intervals that have been defined by deletion mapping (fig002nac).

Figure 3. An expanded map of the *AZFb* interval of the human Y chromosome (fig003nac).

Figure 4. An expanded map of the *AZFc* interval of the human Y chromosome (fig004nac).

Figure 5. Schematic map of the *AZFa* infertility interval in Yq11.2 (fig005nac).

Figure 6. Schematic map of the *AZFa* (human) and *Sxr^b* (mouse) regions of Y chromosomes that are associated with spermatogenic failure (fig006nac).

Citation details for this article

Nabeel A. Affara (2001) The role of the Y chromosome in male infertility. *Exp. Rev. Mol. Med.* 3 January, <http://www-ermm.cbccu.cam.ac.uk/01002319h.htm>