Abstract
In the euchromatic part of the long arm of the human Y chromosome (Yq11) at least 13 Y genes encoding proteins and expressed in male germ cells were found in 3 distinct genomic Y regions frequently deleted in infertile men with idiopathic azoospermia, i.e., for unknown reasons no mature sperm were found in their semen fluid. Accordingly, they were designated as azoospermia factor (AZF) regions: AZFa, AZFb, and AZFc. Additionally, 10 Y genes called “testis-specific transcript Y” (TTTY) genes were mapped in the same AZF intervals. They belong to the long non-coding RNA gene pool in human germ cells because they seem to lack any protein-coding potential. Distinct chromatin regions in Yq11 overlapping with AZFb and AZFc are supposed to be involved in the premeiotic X and Y chromosome pairing and inactivation process controlling male germ cell meiosis. It can thus be assumed that the germ line function of the AZF loci in Yq11 may be not only based on the expression of some germ cell proteins, but also on the expression of some germ cell-specific TTTY transcripts and a locally dynamic and specific chromatin folding structure probably controlled by some germ cell-specific nuclear proteins.

Human Y Chromosome and Male Infertility: Forward and Back from Azoospermia Factor Chromatin Structure to Azoospermia Factor Gene Function

Peter H. Vogt · U. Bender · J. Zimmer · T. Strowitzki
Reproduction Genetics Unit, Department of Gynecological Endocrinology and Reproductive Medicine, University of Heidelberg, Heidelberg, Germany

The azoospermia factor (AZF) locus in Yq11 was first recognised by observation of microscopically visible terminal deletions of the long arm of the Y chromosome in 6 patients with azoospermia, i.e., with no mature sperm in their ejaculate [1]. They included the complete polymorphic heterochromatin block in distal Yq (Yq12) and were assumed to also affect the neighbouring distal euchromatin Y long arm region (Yq11.23). It has been suggested that some function of the distal Yq heterochromatin blocks human male fertility despite the commonly observed polymorphism length in Yq12 [2]. Careful molecular analyses then revealed more individual breakpoints extending along the complete euchromatic long Y arm region (Yq11) in infertile men with AZF-Yq11 aberrations, and histological analyses of their individual testicular pathology displayed a severe disruption of spermatogenesis at different development phases [3]. This suggested an essential function of AZF during premeiotic and postmeiotic human male germ cell development, and consequently that this genetic locus should encode more than 1 Y gene functionally involved in
the complex human male germ cell differentiation process.

Their location in Yq11 could first be distinguished after having established detailed molecular deletion maps of genomic Y DNA in a number of infertile patients with distinct aberrant Yq11 regions and collected in different research groups. These maps were based on the distinct location of Y-specific DNA probes selected from different Y chromosome clone libraries and the analyses of their presence or absence on DNA blots of a series of genomic DNA samples extracted from infertile men with microscopically visible Yq11 deletions and rearrangements [4–6]. On this molecular level, for the first time, we observed some interstitial microdeletions in Yq11 when screening azoospermic men with a normal-looking Y chromosome and a normal 46,XY karyotype [7, 8]. At that time, interstitial microdeletions were already found in comparable dynamic chromosome regions of different autosomes [9]. This therefore raised the question of whether the genomic region(s) of this AZF locus could be better focussed when mapping the location of more interstitial microdeletions in Yq11 on the Y chromosome of azoospermic men with a normal karyotype (46,XY).

This became possible after more detailed molecular interval maps of the human Y chromosome could be generated using the novel PCR technology with specific primer pairs which generate numerous sequence tagged sites (STSs) along the complete human Y chromosome [10–12].

In the first large STS deletion screening study using the genomic DNA samples of 370 infertile men with idiopathic azoospermia – all with a normal 46,XY karyotype – numerous interstitial Yq11 microdeletions were revealed, especially in distal Yq11. Of greatest interest was the identification of 3 microdeletions which cluster with similar breakpoints in 3 distinct molecular intervals in Yq11 [13]. They were “de novo” mutation events, i.e., occurring only in the patients’ Y chromosomes. Inspection of the testicular histology of these patients revealed disruption of their spermatogenesis at 3 distinct developmental phases. The 3 Yq11 microdeletions overlapped with the location of AZF in Yq11; they were tentatively coined “AZFa,” “AZFb,” and “AZFc,” respectively. A complete Sertoli cell-only (SCO) syndrome was observed in all patients with deletion of the complete AZFa interval mapped in proximal Yq11 (Yq11.21). Only Sertoli cells, but no germ cells, were visible in the tubules of their testis tissue sections. Arrest at meiosis was observed in the testis tissue of all patients with deletion of the complete AZFb interval mapped distal to AZFa in Yq11.22. The populations of spermatogonia and primary spermatocytes in all testis tubules analysed were in the normal range; however, no postmeiotic germ cells were identified. A gradual inhibition of the postmeiotic spermatid maturation process with individually variable testicular pathologies was found in patients with the most frequently occurring AZFc deletions mapped in distal Yq11 (Yq11.23). In most tubules, only Sertoli cells were identified, but in some tubules germ cells of different developmental stages were clearly visible and the occurrence of mature sperm cells, albeit in a low number, was observed repeatedly. This testicular pathology has also been described as SCO type II, or mixed atrophy [14].

Some patients with a “de novo” AZFc deletion were still able to produce small numbers of mature sperm in their ejaculate (i.e., “severe oligozoospermia”), while some patients were found to have inherited their AZFc deletions from their fathers [15–18]. It could therefore be concluded that the primary pathology of an AZFc deletion is most likely not an SCO syndrome but hypospermatogenesis [13]. Some more interstitial Y deletions were found in and outside of AZFa, b, and c in Yq11, but most of them were also present on the father’s Y chromosome [19]. This confirmed the presence of some polymorphic sequence regions in Yq11 as expected [2]. Polymorphic genomic mutations are, however, most likely not as-
associated with the occurrence of azoospermia in these infertile patients.

The molecular complexity of the genomic structure in Yq11 and the functional gene content of each AZF microdeletion became established after knowledge of the complete genomic Y sequence in Yq11 [20]. Twenty-three Y genes expressed in human testis and located in 1 of the AZF deletion intervals were registered. These include 13 protein-coding genes specifically expressed in testicular tissue (5 single copy, 8 multicopy) and 10 so-called testis-specific transcript Y (TTTY) genes (5 single copy, 5 multicopy; Fig. 1). At that time (2003), the TTTY genes were considered to be non-functional, although these RNAs contained distinct exon structures and significant lengths of “open reading frames” for potential protein coding were reported [20].

The human Y chromosome that was first sequenced is now called the “reference sequence” (http://www.ncbi.nlm.nih.gov/nuccore/NC_000024.10). Today we know that it is only 1 of more than 1,200 distinct Y sequences now analysed. They are spread with different rates worldwide in all human populations [21]. The Y reference sequence originates from the donation of a genomic DNA sample by an anonymous man in California with an unknown fertility status and belongs to Y lineage R1* [22]. All Y lineages could be schematically ordered in 1 large pedigree due to distinct patterns of single nucleotide variations combined with distinct lengths of STR (small tandem repeat) blocks found along the genomic sequence of the short and long Y arm. This indicates that they have evolved from an ancient common ancestor and form stable “haplogroups” due to the lack of meiotic crossing over events between the X and Y chromosome during their continuous evolution in the different human population pedigrees [23–25].

Here we aim to update the reader on the functional complexity of the AZF locus in Yq11. We begin with a comprehensive description of its putatively germ line-specific dynamic chromosome structure probably supported by expression of at least some TTTY genes, followed by a critical description of putative function(s) of the single AZFa, AZFb, and AZFc genes proposed to be required for the human spermatogenesis process. For the latter purpose, some PCR multiplex assays were developed which allow the detection of any single AZF gene deletion and to distinguish distinct AZF breakpoint border regions found on the Y chromosome of infertile men with respectively complete and partial AZFa, AZFb, and AZFc deletions [26]. They have been composed according to the guidelines provided by the European Academy of Andrology (EAA) and the European Molecular Quality Network (EMQN; www.emqn.org), respectively.

**Impairment of Dynamic Azoospermia Factor Chromosome Structure in Men with Infertility**

It has been shown in mice that pairing of the sex chromosomes in premeiotic male germ cells is required for their further differentiation during meiosis and postmeiotic spermatid development [27]. It starts during the DNA replication phase with B spermatogonia increasing their nuclei and becomes visible by long chromosome threads in the leptotene phase of spermatocytes. During this process, the Y chromosome decondenses and starts to pair with the X chromosome at their DNA homologous sequence domains called pseudoautosomal regions (PARs). They are located at the tip of the short and long chromosome arms (XpYp: PAR1; XqYq: PAR2) and display a high recombination frequency [28]. In humans, pairing of both sex chromosomes proceeds in a gradient beyond the PARs, including eventually the complete short Y arm and at least the proximal part of Yq11 (Fig. 2a) [29, 30]. These extensive premeiotic X-Y paring structures, which were confirmed by different research groups, can explain why induction of a high frequency of double strand breaks required for subsequent XY chromosome pairing and recombination in PAR1 and PAR2 also extends into the neighbouring
proximal non-homologous short Y arm [31]. It also explains the putative molecular origin of a number of distinct Xp-Yp translocations in mice and men, all causing male infertility.

In the late pachytene stage, the pairing structure of X and Y chromosomes then forms a compact chromatin structure also called the X-Y body [30], which is visible with a light microscope (Fig. 2b). The X-Y body displays no transcriptional activity and this meiotic sex chromosome inactivation (MSCI) is required for the subsequent meiotic division events [32]. Consequently, any disruption of this X-Y pairing gradient in male germ cells results in male infertility. Causative disrupting agents are mainly XY aneuploidies and structural rearrangements [33].

Recently, it has been proposed that distinct chromatin regions in Yq11 overlapping with the genomic AZFb and AZFc intervals might also be functional for this premeiotic X and Y chromosome pairing process. In premeiotic germ cells of azoospermic men with distinct AZF microdeletions, distortion of the X-Y pairing structure was observed with Y chromosomes containing AZFb [34] and AZFc microdeletions [35], and earlier with overlapping AZFb-c deletions [36]. AZFb deletions seem to interfere with the initiation of meiotic recombination [34, 36], whereas Y chromosomes with AZFc deletions displayed an extended zygotene stage and reduced condensation at pachytene [35]. These data suggest that meiotic arrest, the major testicular pathology of infertile
men with complete AZFb deletion, might not be mainly caused by any AZFb gene deletion, but rather by deletion of a locus-specific and dynamic chromatin structure located in the genomic AZFb and AZFc region.

This would revitalise an old hypothesis on the germ line function of AZF proposed to include not only functional AZF proteins, but also a specific AZF chromatin folding code controlling the condensation cycle of the Y chromosome in the male germ line [37, 38]. Assemblies of tandem repetitive sequence blocks organised in large locus-specific repetitive sequence structures like those found in AZFb and AZFc (Fig. 1) should indeed be able to fold the underlying chromatin in locus-specific domains [39, 40].

Three large heterochromatic sequence blocks, DYZ17, DYZ18, and DYZ19, were mapped in Yq11 by sequence analyses [20]. Two border the euchromatic Yq11 long arm region distal to the centromere (DYZ17) and proximal to the large and polymorphic Yq12 heterochromatin block (DYZ18), respectively. DYZ19 is located in AZFb (Fig. 1) and occasionally visible under the microscope as the small G(iemsa) dense band – Yq11.22 – of human metaphase spreads. Whereas DYZ17 and DYZ18 are composed by multiple tandem repetitive GGAAT sequence units, DYZ19 contains 3,200 tandem repetitive copies of a basic 125-nt long sequence unit in a 400-kb long heterochromatic sequence block [20]. These heterochromatic sequence blocks can be considered as putative candidate regions for a specific chromatin folding code involved in the general decondensation process of the Y long arm during premeiotic male germ cell development [41].

However, nothing is yet known about a putative dynamic decondensation-condensation cycle of the DYZ17/DYZ18 chromatin structures, nor that of DYZ19 in AZFb during the X-Y pairing process. Neither has it been determined whether the compact DYZ19 chromatin structure in AZFb has some control function for the expression of the most neighboured AZFb protein-coding...
genes, SMCY and EIF1AY, respectively (Fig. 1). This has recently been shown for some polymorphic heterochromatic sequence blocks located on the Y chromosome of Drosophila controlling expression of the neighbouring Y-encoded protein genes expressed in male germ cells [42].

In proximal AZFb, we probably found another putative locus-specific chromatin structure in the b(lue)5 and b(lue)6 amplicons forming the P4 palindrome. Both amplicons contain significant homology regions to a germ cell-specific transcript (dhMiF1) expressed on the lampbrush loops of the Y chromosome of D. hydei and designated earlier as the pYH6 sequence family [43]. Their major sequence cluster (pYH614, pYH34, pYH54, pYH64, pYH65) was found on both arms of the P4 palindrome [Keil et al., unpubl. data], which also include the protein-coding HSPY genes overlapping in their 3’UTR in antisense with the TTTY9 gene copies (Fig. 1) also expressed only in male germ cells [44]. Most interestingly, the multicopy pY6H65 sequence in P4 was also found in the other b amplicons (b1-b4) in AZFb and AZFc, respectively, suggesting a common evolutionary origin [Emig and Vogt, in preparation]. Future experiments will now have to analyse the chromatin structure dynamic in these b amplicons especially in the P4 palindrome. Lange et al. [45] reported that the P4 palindrome might be a hotspot for breakage and structural rearrangements to form dicentric chromosome structures.

**AZFa Locus and Gene Deletions and Male Infertility**

The genomic extension of the AZFa locus in proximal Yq11 can be estimated by its conserved chromatin domain with the same gene order in mouse, cat, porcine, and non-human primates [46]. It extends about 1,150 kb and is bordered by STS sY85 and sY89 (12,526,311 nts: start of sY85 and 13,675,643 nts: end of sY89), respectively, according to the recent human Y chromosome sequence edition: NC000024.10. It includes 3 protein encoding genes, DDX3Y, USP9Y, UTY, and 1 of the testis-specific transcription units, TTTY15 (Fig. 1).

An intrachromosomal recombination event between 2 homologous retroviral sequence blocks (HERV15yq1 and HERV15yq2) deletes 792 kb of this AZFa region [47–49]. No somatic pathology was observed in men with this large Y deletion; however, they had a severe testicular pathology. Men with AZFa deletion suffer from a complete germ cell aplasia, also called SCO syndrome. Consequently, it was proposed that at least 1 of the Y genes deleted in this genomic Y interval, USP9Y, DDX3Y, TTTY15, might already be involved in male germ cell development during early human embryogenesis. AZFa deletions are frequently observed in men with complete SCO syndrome (9–55%) [50–52].

To analyse which of the AZFa Y gene deletions is indeed associated with the observed severe testicular pathology, SCO patients with single gene “de novo” deletion events in AZFa would be most informative. The first “de novo” AZFa gene deletion was found with USP9Y. Its deletion was found to be associated with a postmeiotic disruption phase of spermatogenesis [53]. This was confirmed later by expression analysis of the USP9Y protein being predominantly in spermatids [46]. Natural transmission of some USP9Y mutations, including complete deletion in an Italian pedigree [54, 55], then revealed that disruption of the putative function of the USP9Y gene during the postmeiotic male germ cell development might not be essential for male fertility. Consequently, USP9Y could be excluded as a major AZFa gene, which is required to be functional during foetal germ cell development.

Single DDX3Y gene deletions were reported in some infertile men from northern Italy, and their testicular pathologies included the SCO syndrome [56]. Later observation of DDX3Y protein expression in foetal male germ cells revealed its initiation during the 17th gestation week [57]. Af-
ter birth its expression was concentrated in the so-called Ap(ale) spermatogonia, and after puberty in the cytoplasm of Ap and Ad(ark) spermatogonia located at the basal membrane [57, 58].

In xenograft experiments within the mouse germ line transfected with human transgenes, including an active DDX3Y gene, it was recently shown that the DDX3Y protein in this germ line supports the maintenance and differentiation of the germ line spermatogonia stem cell fraction after puberty [59]. These experiments significantly support a function of DDX3Y in human male germ cells also after puberty. The molecular targets of DDX3Y in human spermatogonia are not yet known. It might be involved as translational control factor for other still unknown germ cell transcripts with extended and structural complex 5′UTRs. This conclusion could be drawn because it was found for all other RNA helicases of the DDX3 protein subfamily conserved from yeast to men [60].

One might thus predict that any AZFa deletion including the DDX3Y gene on the human Y chromosome will indeed cause an early disruption of human spermatogenesis, which finally results in the complete absence of germ cells in the testis tubules after puberty.

It remains an open question whether expression of the third AZFa gene, UTY, can interfere with the testicular phenotype of an AZFa deletion including only USP9Y and DDX3Y. UTY belongs to the large family of H3K27me2/3 histone demethylases with a JmjC domain, which are known to be involved in male germ cell development of the mouse (JMJD1A, also known as JHDM2A) [61, 62] and JMJD1C [63]. JMJD3, the third member of the UTX/UTY subfamily of JmjC domain-containing proteins (also called KDM6B), was found to be expressed in mouse spermatogonia stem cells [64]. We recently found a pattern of UTY expression in foetal male germ cells and after puberty in spermatogonia at the basal membrane, comparable with that of the DDX3Y protein [Vogt, in preparation]. It would therefore be interesting to analyse a putative functional interaction of both these AZFa proteins in the premeiotic human male germ cells.

### AZFb Gene Deletions and Male Infertility

The molecular extension of complete AZFb deletions causing meiotic arrest were originally defined by absence of a set of genomic STS markers mapped schematically in linear order of a molecular deletion interval map [13]. Later sequence analysis revealed that the proximal border of this microdeletion was mainly located in a palindrome of 2 large repetitive sequence blocks with nearly identical sequence structures, coined P5 palindrome [65]. The distal border line was also found in a palindrome with 2 highly homologous 92-kb-long sequence blocks. This was embedded within a large repetitive sequence block called “yellow amplicon” located in the AZFc deletion interval (see the section relating to AZFc, below). It was designated P1.2 because the yellow amplicon was part of the large P1 palindrome in AZFc (Fig. 1); AZFb deletions thus overlap with AZFc deletions. In proximal AZFb, the break fusion events were proximally mapped 30 kb proximal to the centre of the P5 palindrome and distally 25 kb proximal to the centre of the P1.2 palindrome [65]. Small homologous sequence domains within these sequence regions suggest that AZFb deletions are caused by homologous recombination events like AZFa deletions. Classical AZFb deletions are therefore sometimes also designated as “P5/proximal-P1” or P5/P1.2 deletions [65]. However, non-homologous recombinations that also cause complete AZFb deletions and even larger microdeletions, additionally encompassing nearly the complete AZFc interval, have been identified [65]. The molecular extensions of these so-called AZFb+c deletions are not the sum of an AZFb and AZFc deletion, but are shorter because
the distal breakpoints in AZFc are located in the short P1.1 palindrome embedded in the distal yellow amplicon in AZFc. Classical AZFc deletions end in the more distally located repetitive sequence block called “b(1ue) 4” amplicon (Fig. 1).

Classical AZFb deletions caused by P5/P1.2 recombination encompass a 6.2-Mb-long genomic Y sequence including 6 protein-encoding Y genes, EIF1AY, HSFY (2 functional copies), PRY (2 functional copies), RBMY (6 functional copies), RP-S4Y, and SMCY (also known as JARID1D or KDM5D), which are deleted completely in patients with this Y chromosome aberration. Additionally, 1 copy of CDY2 and XKRY is located in the proximal AZFb P5 palindrome, and copies of the AZFc gene families, BPY1 (1 copy), DAZ (2 copies), and CDY1 (1 copy), are in the distal overlapping AZFb/c interval. Complete AZFb deletions also include 9 Y genes of the TTTY class (Fig. 1).

It has not yet been clarified whether any of these AZFb genes are causative agents for the testicular pathology because disruption of the chromatin folding structure in AZFb might be already sufficient to cause meiotic arrest of the patients’ germ cells (see the section on the AZF chromosome structure, above). To obtain a final answer to this basic question, PCR multiplex assays for diagnosing deletions of each single AZFb gene separately are obviously required [26, 66, 67].

This focus is feasible, especially in AZFb, because multiple genomic deletions were found in this genomic Y interval on the Y chromosome of infertile men which were significantly smaller than those of the “classical” AZFb and AZFc microdeletions that were described first [13]. Some of them were not associated with the infertility phenotype because they were also found as polymorphic length variants on the Y chromosome of fertile men from different populations [68]. Therefore, AZFb genes in these polymorphic AZFb subdomains could be excluded as causative agents for the meiotic arrest pathology. A short summary of single AZFb gene deletion results analysed so far is given here in alphabetical order.

EIF1AY protein expression was mainly found in the testis and ovary (http://www.genecards.org/). Like DDX3Y it is functional in the efficiency control of the ribosomal translation initiation process. However, single EIF1AY gene deletions associated with male infertility have not yet been found.

HSFY protein is expressed in many germ cells (spermatogonia, zygotic spermatocytes, elongated spermatids) and also in Sertoli cells [69]. Complete HSFY deletion (2 gene copies) were found to be inherited over many generations in some families with hypospermatogenesis in southern France [66]. They are therefore most likely not candidates for causing the major AZFb pathology (meiotic arrest).

PRY protein expression was found in spermatids and spermatozoa, and the PRY fraction positively staining spermatozoa increased from 1.5 to 51.2% in sperm samples with increased apoptotic DNA degradation [70]. This protein phosphatase is therefore assumed to be involved in the apoptotic degradation of non-functional spermatozoa. Since PRY protein was not detected in premeiotic germ cells, its contribution to the testicular pathology of men with AZFb deletion (meiotic arrest) is questionable. Single PRY gene deletions are difficult to identify [71]. The 2 functional gene copies have a high sequence homology because they are located in the P3 palindrome of AZFb (Fig. 1).

RBMY protein is expressed only in the male germ line and predominantly in premeiotic germ cells [72]. It probably provides spermatogonia and early spermatocytes with a germ cell-specific splicing cofactor interacting in a multiprotein complex with other splicing factors, such as 9G8, SRp20, SRp30c, and TRA-2β [72, 73]. It is interesting that its X homolog, RBMX, cannot functionally replace RBMY because only the RBMY RNA binding domain recognises stem-loop structures in RNAs with a CA/AUCA consensus sequence [74]. It can therefore be assumed that the RBMY protein structure evolved from that of...
RBMX by evolutionarily selection for a male germ line-specific splicing factor functional in human germ cells before meiosis [72]. Single RBMY gene deletions are difficult to assess because at least 6 functional gene copies are located in the P3 palindrome of AZFb near the PRY genes (Fig. 1).

RPS4Y2 is mainly expressed in testis tissue containing germ cells and in the prostate [75]. The gene evolved by duplication from RPS4Y1 located on the short Y arm during primate evolution [76]. Single RPS4Y2 gene deletions have not been found in infertile men. Whether and where the protein is required for human male germ cell development is therefore not yet known.

SMCY protein can form a complex with MSH5, a critical meiosis-regulatory protein, and immunohistochemical analyses with murine testicular tissue sections revealed co-localisation of both proteins at a specific meiotic stage of spermatogenesis [77]. Like RBMY, it can therefore be considered as a major candidate for causing the AZFb testicular pathology. However, SMCY (now designated KDM5Y because it is recognised as a functional histone demethylase) gene deletions have not yet been found in infertile men.

More informative in revealing the AZFb genes that cause meiotic arrest if disrupted has been the description of a “de novo” partial AZFb deletion on the Y chromosome of an infertile man in China who suffers from a severe oligoasthenoterato (OAT) syndrome [78]. It encompassed the complete deletion of HSFY, SMCY, EIF1AY, and RPS4Y2, but did not remove the PRY and RBMY gene copies, suggesting that only these 2 AZFb genes, but not HSFY, SMCY, EIF1AY, and RPS4Y2 are required for meiotic human germ cell development.

The description of another partial AZFb deletion also found in an infertile patient with OAT syndrome excluded the PRY genes as major AZFb genes because this deletion was found to be inherited from father to son and included the complete P2 and P3 amplicon in the distal AZFb region with 2 DAZ gene copies, both functional PRY copies, and 4 RBMY gene copies [79]. This leaves the possibility of a major AZFb gene now only open for RBMY, since on the Y chromosome of this OAT patient 2 RBMY gene copies were still present [79].

The exclusion of EIF1AY, HSFY, RPS4Y, and SMCY as major AZFb genes was confirmed by our own diagnostic experience in applying PCR multiplex AZF gene deletion assays. We found a partial AZFb deletion on the Y chromosome of a man with severe OAT syndrome, again encompassing EIF1AY, RPS4Y, and SMCY, but no RBMY genes and only 1 copy of PRY [Vogt et al., in preparation]. Nevertheless, we cannot yet exclude that the Y genes deleted may contribute to the observed aberrant OAT sperm phenotype of these patients.

More partial AZFb deletions associated with variable testicular histologies but not meiotic arrest and with distal breakpoints concentrated in the overlapping genomic AZFb-AZFc interval have been described in the literature [80–82]. Unfortunately, most of them did not address the basic question of whether the observed partial AZFb microdeletion was “de novo,” and thus only found on the patients’ Y chromosomes, or polymorphic because it was also found on the Y chromosomes of their fertile fathers. Nevertheless, no partial AZFb deletion was identified with complete deletion of all RBMY gene copies and a testicular pathology less severe than meiotic arrest.

We can thus surmise that only the RBMY gene family might have an essential function during meiosis for the human male germ cell. It probably controls some germ cell-specific splicing processes of transcripts occurring in germ cells predominantly before meiosis and required for the meiotic cell divisions. Moreover, partial RBMY gene deletion might be associated with the OAT syndrome (see above) because expression of the RBMY protein was also found in postmeiotic round and elongating spermatids and in the midpiece region of sperm [83].

In this context, it is worthwhile reporting the exceptional case of an infertile man with complete
AZFb+c deletion but with a testicular histology usually only found in patients with partial AZFb or AZFc deletions, since the postmeiotic marker genes, PRMI and TNP1, were found to be expressed and the patient produced a low number of mature sperm [84]. The haplogroup of this Y chromosome derived from the Y-L* lineage, which is rare in Western Europe and North American male populations, but predominant in Asia, Pakistan, and India. The authors therefore assume that some genetic modifier(s) must be present on the Y chromosome supporting the meiotic pairing process in the patient’s spermatocytes, and thereby probably replacing any functional requirement for an intact genomic AZFb region with some RBMY genes.

Based on this overview, we can assume that the genomic Y region encompassed by the AZFb microdeletion in the Y reference sequence with haplogroup R1* might be highly variable on the Y chromosomes derived from other Y lineages located predominantly only in Africa, Asia, or Australia, as already suggested earlier by Jobling et al. [68]. Indeed, some hot-spot regions for polymorphic Y chromosomal breaks and rearrangements were found especially in distal AZFb near the PRY and RBMY gene copies in or near the P3 palindrome. They result in a number of complete PRY deletions and partial RBMY gene copy deletions, which looks to be compatible with normal male fertility [85].

In summary, it might therefore be possible that the major testicular pathology, complete meiotic arrest, found in infertile patients with complete AZFb microdeletion is based primarily on the disruption of some specific chromatin structure in AZFb extending at this phase of male germ cell development to arrange a proper start of the pre-meiotic X-Y pairing structure. Expression of the germ cell-specific splicing factors encoded by the RBMY gene family during the same germ line phase might not be required for the subsequent meiotic cell divisions, but probably are functional by supporting the efficiency of this process.

**AZFc Gene Deletions and Male Infertility**

Similar considerations might help to reveal the functional contribution of the polymorphic gene content of the genomic AZFc region, sometimes inherited from fathers to sons [15, 16]. Complete AZFc microdeletions were found with high frequency (5–15%), however, only in genomic DNA samples of infertile men with a severe reduction in sperm number (<10 million per mL ejaculate) and including the OAT syndrome [22, 86–88]. These studies clearly show that complete AZFc deletions interfere with the spermatogenic maturation process, reducing its efficiency and causing hypospermatogenesis as the primary mutation effect. Age-dependent secondary mutation effects might – but must not – eventually cause a complete absence of mature spermatozoa (i.e., azoospermia) in the man’s semen fluid [13, 88].

This raised similar questions to those concerning the AZFb genes – which AZFc gene, if any, is functionally required for human male germ cell development, and do AZFc genes only improve the efficiency of this complex cellular differentiation process?

All AZFc gene families are composed by more than 1 gene copy, and partial AZFc deletions reducing the copy number were also found in fertile men. It has been shown that the impact of any partial AZFc deletion on the efficiency of the spermatogenic process mainly depends on the patient’s Y chromosomal haplogroup [86, 89–91]. The frequency of the different Y haplogroups found in a given human population might therefore be the result of a balancing selection aimed to optimise the male fertility function of the Y chromosomal AZF locus in each population, thus driving male reproductive fitness [92, 93]. The observed genomic deletion and amplification mechanisms in the locus-specific repetitive sequence blocks on the short and long arm of the Y chromosome in men, which create large palindromes, might be considered as efficient driving tools [94, 95].
The most prominent partial AZFc deletions became the so-called gr-gr deletions, meaning any partial AZFc deletion caused by recombination within the 3 green, g1, g2, g3, and 4 red, r1, r2, r3, r4, amplicons in AZFc [22, 98]. At least 14 independent recombination events in these amplicons have already been found in distinct Y lineages. Some seem to cause male infertility, while some do not [22, 96, 97]. It is therefore surprising that clinical protocols aimed to reveal a gr-gr recombination events interfering with patient fertility only use 1 STS marker (sY1291). This sY marker is located at the red2-grey1 border site [22, 98] and always deleted in men with a gr-gr partial AZFc deletion independent from his fertility status.

Another popular approach to distinguishing the heterogeneity of gr-gr recombinations in AZFc is the numbering of FISH signals specifically marking the number of green and red amplicons on the Y chromosome in the lymphocyte nuclei with green and red fluorescent spots visible under the microscope [98]. However, the large heterogeneity of recombination sites in the red and green amplicons now found on the Y chromosome of fertile men and infertile men in different human populations cannot be resolved with this diagnostic tool, but only by sequence analyses [99–101].

It has been shown that complete AZFc microdeletions do not interfere with the premeiotic male germ cell development; numbers of spermatogonia and spermatocytes in the testis tubules are comparable to those of fertile men [102]. It can therefore be concluded that expression of any AZFc gene is only required during the postmeiotic spermatid maturation process. In a population-based survey of 20,000 Y chromosomes derived from fertile and infertile men with different Y haplogroups it has been recently shown that the risk for male infertility increased 145-fold in men with complete AZFc microdeletion, whereas partial AZFc deletions designated gr-gr, b1-b3, and b2-b3, respectively, would only double the infertility risk depending on the Y haplogroup [101]. Similar results were previously found in a study of partial AZFc deletions in fertile and infertile men in Portugal [103], China [104], and France [105]. However, there might be some exceptions since another study in China claimed that b2-b3 partial AZFc deletions show a higher risk for spermatogenic failure than complete AZFc microdeletions [106], and a study in Italy claimed that gr-gr deletions in AZFc are significantly increased in men with a lower sperm count [107].

Most AZF deletion studies do not answer the question of whether the partial AZFc microdeletion identified has occurred “de novo,” i.e., is present only on the patient’s Y chromosome or is also present on the Y chromosome of the father or other fertile male family members. Therefore, it is difficult to assess which of these partial AZFc microdeletions diagnosed by the presence or absence of a couple of STS and single nucleotide variation markers has an impact on the clinical observed phenotype of the patient’s infertility. Since no study has described the exact break-fusion sequence of the proposed ampliconic recombination site, the question remains open regarding which of the AZFc genes, when expressed in the testis tubules, is really functionally required for the development of the human postmeiotic male germ cells. It would be most interesting in this context to reveal some spermatogenic impairment associated with failure of expression of the BPY2, CDY1, and DAZ gene copies in the postmeiotic germ cells of these patients, since these genes are known to be expressed during spermatid maturation and in mature sperm.

The BPY2 gene family contains 3 gene copies (BPY2.1–3) with an identical exon structure and were mapped to the green amplicons (g1–g3) in distal Yq11 (Fig. 1). They have evolved only recently during primate evolution and encode a basic protein of 13.9 kDa, found after meiosis in the nuclei of round spermatids and sperm [108]. It was therefore suggested that these basic proteins are functionally important for maintaining the sperm fertilisation capacity. This view gains support through a specific interaction of BPY2 protein with UBE3A, a ubiquitin protein ligase [109], and with VCY2IP-1 associated with the sperm tail microtubules [110].
The CDY genes (2 copies in AZFc in the yellow amplicons) were also shown to be expressed only in testis tissue and expression of the protein was mainly observed in spermatids [111]. CDY proteins are functionally histone acetyltransferases with a strong preference for histone H4 localised in the nuclei of maturing spermatids. This suggests that they functionally contribute to the specific histone hyperacetylation process in late spermatids, resulting in a more open chromatin structure, which is required to facilitate the spermatogenic histone replacement by transition proteins and protamines.

The number of DAZ gene copies in the red amplicons in AZFc can change between 2 and 8 copies on the Y chromosome of fertile men depending on the Y lineage [98, 112, 113]. Their exon structures are variable, especially in the number of their exon 7 variants, which are also called “DAZ repeats” [99, 114, 115]. DAZ repeats are known to be important RNA binding partners [116]. All DAZ gene copies present in the AZFc region were shown to become translated, forming a pattern of polymorphic protein variants expressed in the postmeiotic human male germ cells [117]. Alternative splicing of distinct DAZ transcripts can additionally increase the number of polymorphic DAZ proteins.

The most prominent DAZ protein interacting partner with a putative function in the postmeiotic male germ cells might be the DAZAP1 protein first expressed in midpachytene spermatocytes then in round and elongated spermatids shuttling between the nucleus and cytoplasm during postmeiotic germ cell maturation [118]. Since this protein specifically recognises the DAZ repeat (i.e., variable number of exon 7 sequence variants) encoded peptide region [116], its binding potential can be efficiently modulated by the number of DAZ repeat copies in each DAZ protein variant expressed after meiosis. In mice, DAZAP1 controls the postmeiotic splicing efficiency of CREM, CRISP2, and POT1a transcripts in germ cells [119]. However, DAZ proteins are absent in mice because the AZFc locus evolved on the Y chromosome first during primate evolution.

It is therefore most likely that the genomic Y region containing AZFc and its DAZ gene family in humans evolved and became established on the Y chromosome during a consecutive duplication-diversification process from their autosomal ancestors, DAZL on chromosome 3 and BOLL (also known as BOULE) on chromosome 2 [120], thereby extending and/or improving their functional profile in the male germ cells. Expression of BOLL and DAZL proteins was mainly found in germ cells before meiosis [121], whereas DAZ proteins were mainly expressed during the postmeiotic germ cell phase in late spermatids and spermatzoa [122].

Considering the great heterogeneity of DAZ gene deletions in fertile and infertile men (see above), we assume that DAZ proteins are not essential for the terminal differentiation process of male germ cells, but may be required for their optimal function. In immunostaining experiments using antibodies probably specific for DAZ2, we found DAZ immunoreactivity in the innermost layer of the mature sperm cell epithelium, preferably in their so-called residual bodies immediately prior to spermiation and in the tails of ejaculated spermatzoa [122]. From these findings, we can deduce a function for DAZ proteins in the RNA metabolism of late spermatids, presumably in the storage or transport of testis-specific mRNAs, the translation of which is repressed until the formation of mature spermatzoa.

**Functional Aspects of the Testis-Specific LNC Transcripts (TTTY) in AZFa, AZFb, and AZFc**

Sequence analysis of the human Y reference sequence also revealed a number of Y genes in the 3 AZF regions which are expressed only in testis tissue but supposed to have no protein-encoding potential, despite containing longer open reading frames [20]. They were designated as testis transcript unit Y (TTTY) genes and can be classified as belonging to the long non-coding (LNC) transcript genes. Thirteen TTTY genes were single copy, 12
TTTY genes were probably duplicated due to their location in palindromes; 2 TTTY genes, TTTY4 and TTTY17, had 3 copies. The map of those located in AZFa, AZFb, and AZFc is shown in Figure 1. In AZFc, we discovered by sequence comparison that TTTY4 and TTTY17 were both derived from ancient tandem repetitive sequence units related to the β-satellite sequence family. Interestingly, their exon structures could be compared to that of the embedded BPY2 genes, suggesting an evolutionary pathway for development of the protein-encoding BPY2 genes from an ancient β-satellite sequence unit on the Y chromosome of primates with development of the flanking TTTY4 and TTTY17 structures as stable intermediates.

One way to reveal a potential protein-coding potential in LNC genes would be inspection for the presence of some known functional protein domains in the open reading frames identified using the Prosite database (http://prosite.expasy.org/). Until now, we were not able to mark any functional protein domain in the exon structures of TTTY transcripts located in AZFb and/or AZFc.

The function of TTTY gene expression during human spermatogenesis might therefore be on the transcript level, like many other LNC genes with tissue-specific transcripts now mapped to the X chromosome and the autosomes [123]. Accordingly, TTTY transcription might be involved in opening and extending the AZFb chromatin structure during the premeiotic X-Y pairing process (see the above section relating to AZF chromosomes). An enhancer function of a testis-specific long non-coding RNA expressed only in mouse spermatocytes was recently found 3.4 kb upstream of the mouse testicular cell adhesion molecule1 (Tcam1) gene [124]. Major candidates in AZFb would be the TTTY9 copies in the P4 palindrome overlapping the distal part of the HSFY genes in opposite polarity, TTTY5, TTTY6, and TTTY13, in the P3 palindrome embedded in the RBMY and PRY gene copies, and TTTY4 and TTTY17 flanking the BPY2 gene copies (Fig. 1). TTTY13 transcription might also contribute to the observed instability of the human Y chromosome only in the P3 palindrome [45] because its transcription seems to be controlled by HERV (human endogenous retroviral)-related sequence domains belonging to the HERV-K14C sequence family [125].

References


31 Decapretrice F, Ojjarinde OA, Mitchell MJ, Burgoyne PS: Recombination between the mouse Y chromosome short arm and an additional Y short arm derived chromosomal segment attached distal to the X chromosome PAR. Chromosoma 2016;125:177–188.


58 Ditton HJ, Zimmer J, Kamp C, Rajpert-De Meyts E, Vogt PH: The AZFa gene DBY (DDX3Y) is widely transcribed but the protein is limited to the male germ cells by translation control. Hum Mol Genet 2004;13:2333–2341.


Human Y Chromosome and Male Infertility