A two-step protocol for the detection of rearrangements at the AZFc region on the human Y chromosome

Y.-W. Lin1,2*, C.-L. Hsu2* and Pauline H. Yen1,2,3

1Graduate Institute of Life Sciences, National Defense Medical Center and 2Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan
3To whom correspondence should be addressed at: Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan.
E-mail: pyen@ibms.sinica.edu.tw

The AZFc region on the human Y chromosome consists mainly of very long direct and inverted repeats and is prone to rearrangement. Although deletion of the entire AZFc is found only in subfertile men, duplications and deletions of portions of AZFc as well as inversions are quite common and represent major polymorphisms of the Y chromosome. Several methods are available to detect these rearrangements, and each has its own advantages and limits. We designed a two-step PCR protocol to study the polymorphic structure of AZFc. The first PCR determines the copy number of the Deleted in Azoospermia (DAZ) genes within AZFc using the autosomal DAZ-Like gene as a dosage control, and the results could be verified by dosage Southern blot analyses. The second PCR simultaneously detects five sequence tagged sites (STSs) that are either present or absent in the various AZFc partial deletions. One of the STSs, sY1291, was found to be polymorphic in size due to varying lengths of a poly-T stretch. A combination of the DAZ dosage PCR and the 5-STS multiplex PCR reaction detects most, if not all, deletions and duplications at AZFc. It offers a simple and reliable way to screen large populations for AZFc rearrangements and study their effects on male fertility.

Key words: AZFc/DAZ/male infertility/Y chromosome

Introduction

The human Y chromosome plays a major role in spermatogenesis. It contains numerous gene families that are expressed specifically in the testis (Lahn and Page, 1997; Skaletsky et al., 2003). Deletion of these genes often causes spermatogenic defects and is a major cause of male infertility (Krausz et al., 2003). A vast majority of Y-chromosome microdeletions fall within three regions on the long arm, designated AZFa, AZFb and AZFc, with AZFc being the most frequently deleted (Vogt et al., 1996). The AZFc region spans 3.5 Mb and contains 21 genes, including four Deleted in Azoospermia (DAZ) genes that are implicated in the regulation of mRNA translation (Reijo et al., 1995; Saxena et al., 2000; Kuroda-Kawaguchi et al., 2001; Collier et al., 2005). AZFc consists almost entirely of very long direct and inverted repeats and is prone to rearrangement (Figure 1, ‘A’). Indeed, numerous structures at AZFc have been identified. The first AZFc structure, the prototype structure, was derived from a Y chromosome belonging to the haplogroup R1* (hgrR1*) and appears to be present in many other Y haplogroups (Figure 1, ‘A’: Repping et al., 2003a; Vogt, 2005). AZFc with inversions caused by recombination between the b2 and b3 or the g1-r1-r2 and g3-r4-r3 inverted repeats were subsequently identified in several Y haplogroups (Figure 1, ‘I’ and ‘L’: Repping et al., 2004a). The AZFc deletion, resulting from non-allelic homologous recombination between the b2 and b4 repeats, is present in about 1 in 4000 men (Figure 1, ‘B’: Kuroda-Kawaguchi et al., 2001; Page, 2004). Partial deletions of the AZFc region, with the retention of two of the DAZ genes, are much more common (Fernandes et al., 2002, 2004; de Vries et al., 2002; Repping et al., 2003a, 2004; Machev et al., 2004; Ferlin et al., 2005; Giachini et al., 2005; Hucklebroich et al., 2005; Lynch et al., 2005). These deletions can result from recombination between the g1-r1-r2 and g2-r3-r4, or the b1 and b3 direct repeats within the prototype AZFc structure (Figure 1, ‘C’ and ‘E’; Fernandes et al., 2002; de Vries et al., 2002; Repping et al., 2003a). They can also occur on Y-chromosome variants carrying inversions (Figure 1, ‘J’ and ‘M’: Fernandes et al., 2004; Machev et al., 2004; Repping et al., 2004a). In addition to deletions, different duplications at AZFc have also been reported. Duplication can occur on a chromosome with partial AZFc deletion and generate a Y-chromosome variant with four DAZ genes but missing some markers (Figure 1, ‘D’ and ‘N’: Repping et al., 2003a, 2004a). It can also occur on the original Y chromosome with or without inversion and generate Y-chromosome variants with six or eight DAZ genes (Figure 1, ‘F’, ‘G’, ‘H’, ‘K’ and ‘O’: Lin et al., 2005; Writzl et al., 2005), though only the gr/gr duplication (Figure 1, ‘F’) has been characterized (Lin et al., 2005). The structure around AZFc is therefore highly polymorphic. Complete deletions of AZFc can be readily detected by PCR or Southern blotting, but partial deletions or duplications of AZFc are more difficult to identify. Several methods have been developed and used either alone or in combination to detect these partial deletions or duplications. They include Southern hybridization (Fernandes et al., 2002, 2004; Lin et al., 2005), fluorescence in situ hybridization (FISH) (Saxena et al., 2000; Repping et al., 2003a,b, 2004a) and PCR assays based on sequence family variants (SFVs) (Saxena et al., 2000; Fernandes et al., 2002; de Vries et al., 2002; Machev et al., 2004), plus-minus STSs (Repping et al., 2003a, 2004a) and DAZ copy number (Machev et al., 2004; Writzl et al., 2005). Each of them has its own advantages and limits.
We now report a simple two-step PCR protocol to investigate the structures at AZFc. The first PCR determines the copy number of the DAZ gene, and the second PCR reaction detects the presence of five STS markers around AZFc. Although the protocol cannot distinguish all identified structures at AZFc, it offers a simple and effective way for large population screening to identify individuals with AZFc dele-
tions and duplications and to study the effects of the rearrangements on male fertility.

**Materials and methods**

**DNA samples**

Individuals and cell lines with zero, two and six DAZ genes were described previously (Najmabadi et al., 1996; Lin et al., 2005). DNA samples with unknown DAZ copy number were obtained from the Taiwan Supercontrol Biobank that maintains a collection of DNA samples and cell lines of Han Chinese in Taiwan (Hung et al., 2005).

**DAZ dosage PCR assay**

The PCR primers were designed from the 3' UTR of DAZ-Like (DAZL) (accession number U66078). PrDAZ109 (5'–gtaaagaaaatatgcagagag) contained a mismatch with the DAZ sequence, whereas PrDAZ110 (5'–ggtattatccctaggtacc) was a perfect match. The reaction contained 20 mM Tris–HCl (pH 8.8 at 25°C), 10 mM KCl, 10 mM (NH4)2SO4, 0.1% Triton X-100, 2 mM MgSO4, 0.2 μM each of the primers, 0.1 mM each of the dNTPs, 50 ng of genomic DNA and 0.5 units of Taq DNA polymerase in a final volume of 10 μl. PCR amplification was carried out in a Biometa TGradient thermocycler (Biometrax, Goettingen, Germany) with an initial heating at 95°C for 5 min, followed by 25 cycles of 95°C for 20 s, 49°C for 20 s and 72°C for 30 s and a final incubation at 72°C for 5 min. The products were separated by electrophoresis on 2% agarose gels (a 1 : 1 mixture of Seakem LE agarose and NuSieve 3 : 1 agarose). The gels were stained with EtBr (1 μg/ml) and mixed in a 1 : 2 ratio before being 32P-labelled by random priming using Chinese in Taiwan (Hung et al., 2005). The PCR primers PrDAZ109 and PrDAZ110 were designed from the 3' UTR regions. They amplified fragments of 331 and 245 bp from DAZL and DAZ, respectively, which could be readily separated on 2% agarose gels (Figure 2). We first optimized the PCR conditions using DNA samples with known DAZ copy numbers (Lin et al., 2005). PCR amplification for 25 cycles gave the most satisfactory results. Samples with two, four, and six DAZ genes were readily distinguished by comparing the relative intensities of the DAZ and DAZL fragments either visually or by densitometry (Figure 2A). Increasing the PCR cycles to 30 resulted in enhanced amplification of the smaller DAZ fragment and diminished distinction between samples with different DAZ genes (Figure 2B). Owing to small variation in PCR conditions, the DAZ/DAZL signal ratio for the same DNA sample varied between different sets of reactions. It was therefore important to include controls with known DAZ copy number in each set of PCR reactions for comparison. We used the optimized condition to test over 100 DNA samples with unknown DAZ copy number and identified several cases with two and six DAZ genes (Figure 2C).

**Multiplex PCR assay**

The reaction amplified five AZFc markers, sY1161, sY1191, sY1201, sY1206 and sY1291, and a control gene pair ZFX/ZFY. We used published primers for sY1161, sY1201 and ZFX/ZFY (Repping et al., 2003a; Simoni et al., 2004) and changed one or both primers for sY1191 (reverse only), sY1206 (both) and sY1291 (both) to amplify sY1191a, sY1206a and sY1291a (‘a’ denotes altered). The new primers were sY1191R2, 5'-ccacacatgccgtagag; sY1206R2, gattacggtcctagttag; sY1291F2, gattcacggtgactaggctgag (from AF334537); sY1291R2, aatgggagataaagag) contained a mismatch with the sequence, whereas PrDAZ110 (5'–gtaaagttctgcaacg; sY1206F2, aggaggca gagattgatctc and sY1206R2, tagaaga-gaca tgcgtggcc. The concentrations of the primers were adjusted to give PCR products of comparable intensities. They were 1 μM each for ZFX/ZFY, sY1161, sY1191a and sY1206a, 2 μM for sY1291a and 10 μM for sY1201. PCR reactions were carried out using Taq DNA polymerase master mix (AMP140303; Amplipon, Copenhagen, Denmark) under the following conditions: preheated at 95°C for 15 min, followed by 34 cycles of 95°C for 30 s; 63°C for 90 s; 72°C for 60 s and a final extension at 72°C for 10 min. The products were analysed on 2% agarose gels as described above.

**Dosage Southern blot**

Isolation of genomic DNA and Southern blotting were carried out as previously described (Lin et al., 2005). Two 1.4 kb fragments containing the 3' UTRs of DAZ and DAZL were PCR amplified from the respective cDNA clones using primers PrDAZ110 (5'–ggtattatccctaggtacc) and PrDAZ115 (5'–agttgcatcagttagcatcagag) that amplified DAZ and DAZL fragments were gel purified and mixed in a 1 : 2 ratio before being 32P-labelled by random priming using the Prime-a-Gene labelling system (Promega, Madison, WI, USA). Hybridization signals were detected on a Typhoon 9410 Variable Mode Imager (Amerham, Piscataway, NJ, USA) and quantified using the line analysis and graphic display programs.

**Results and discussions**

Most of the structures identified at AZFc can be differentiated by the copy number of the DAZ genes, which are embedded in the red repeats, and the presence or absence of five STS markers sY1161, sY1191, sY1291, sY1206 and sY1201 (Figure 1; Repping et al., 2003a). We developed a dosage PCR assay to determine the DAZ copy number using the autosomal DAZL gene as an internal dosage control. DAZ and DAZL share 89% sequence similarity (Saxena et al., 1996; Yen et al., 1996). The PCR primers PrDAZ109 and PrDAZ110 were designed from the 3' UTR regions. They amplified fragments of 331 and 245 bp from DAZL and DAZ, respectively, which could be readily separated on 2% agarose gels (Figure 2). We first optimized the PCR conditions using DNA samples with known DAZ copy numbers (Lin et al., 2005). PCR amplification for 25 cycles gave the most satisfactory results. Samples with two, four, and six DAZ genes were readily distinguished by comparing the relative intensities of the DAZ and DAZL fragments either visually or by densitometry (Figure 2A). Increasing the PCR cycles to 30 resulted in enhanced amplification of the smaller DAZ fragment and diminished distinction between samples with different DAZ genes (Figure 2B). Owing to small variation in PCR conditions, the DAZ/DAZL signal ratio for the same DNA sample varied between different sets of reactions. It was therefore important to include controls with known DAZ copy number in each set of PCR reactions for comparison. We used the optimized condition to test over 100 DNA samples with unknown DAZ copy number and identified several cases with two and six DAZ genes (Figure 2C).
We further verified the results of the DAZ dosage PCR assay by dosage Southern blotting using again the autosomal DAZL as an internal dosage control. The probe contained a mixture of the 3' UTRs of DAZ and DAZL. On Southern blots containing DNA samples digested with NsiI, the probe detected a 4.7 kb DAZ fragment and a 3.0 kb DAZL fragment (Figure 3). The relative hybridization signals of the two fragments showed good correlation with the copy number of DAZ genes in the DNA samples with known DAZ copy number. The ratios were used to determine the copy numbers of the unknown samples, and the results were in complete agreement with those determined by the dosage PCR assay.

So far, the most commonly used method to detect AZFc partial deletions is the plus/minus STSs-based PCR assay that detects the presence of five markers around AZFc (Figure 1; Repping et al., 2003a). When each marker is assayed individually, it is sometimes difficult to tell whether a finding of no PCR product is due to deletion or PCR failure. We thus developed a multiplex PCR, similar to the one used to detect the various AZF deletions (Simoni et al., 2004). We included the X/Y homologous gene pair ZFX and ZFY as a quality control for PCR amplification and changed the primers for sY1191, sY1206 and sY1291 so that the six PCR products in a reaction were well separated on 2% gels (Figure 4). The new STSs were designated sY1191a, sY1206a and sY1291a ('a' for altered) to distinguish from the original STSs. Multiplex PCR reactions on samples with various deletions gave the same results as those determined by the single PCR reactions using the original primers. They also showed that the sY1291a fragments (as well as sY1291 fragments, data not shown) were of various sizes. Marker sY1291 spans the junction of the red and the grey repeat (Figure 1), and its GenBank sequence (accession number G72340) contains a stretch of 39T (nt 46 to nt 94) in the red repeat portion. We determined the DNA sequences of several sY1291a fragments and found that the observed size polymorphism stemmed from differences in the number of T's in the stretch that ranged from 20 to 37.

Here we report a two-step protocol consisting of a DAZ dosage PCR assay and a 5-STS multiplex PCR assay. The dosage PCR assay detects both deletions and duplications, regardless of the underlying mechanisms. It is simple and does not require additional restriction digestions such as in the SFV-based assays. It is also much easier to perform than a similar dosage PCR assay recently reported by Machev et al. (2004). In that assay, the PCR products of DAZ (184 bp) and DAZL (187 bp), amplified from intron 10, differ only by three base pairs and require long sequencing gels to separate. In addition, the DAZL fragment in that assay is polymorphic in size due to a 40 bp insertion in some individuals, making the dosage comparison less straightforward. Our dosage PCR assay may not be sensitive enough to distinguish odd number (such as three) DAZ genes from even number genes (such as two and four). On the basis of the current models of the AZFc structures and the molecular mechanisms underlying the various deletions and duplications, it would be rare for a Y chromosome to carry odd number DAZ genes. Although there were earlier reports of Y chromosomes with odd number DAZ genes, they need to be re-evaluated in the light of our new knowledge on the AZFc structure (Glaser et al., 1998; Moro et al., 2000). The 5-STS multiplex PCR assay we developed detects five AZFc markers in a single reaction and significantly reduces the workload. In addition, it incorporates ZFX/ZFY as a PCR quality control and can readily distinguish failure of PCR amplification from genuine deletions. Our assay is an improvement over a similar multiplex PCR assay that detects only the AZFc markers (Lynch et al., 2005). No PCR products in that assay could mean PCR failure, or that the individual lacks the markers because he is an XX male or has a Yq terminal deletion. Our protocol nonetheless has some limitations. It cannot distinguish Y chromosome with the prototype AZFc structure (Figure 1, 'A(1)') and the ones with inversions (Figure 1, 'I' and 'L'), nor can it distinguish the various duplications with six or eight AZFc genes (Figure 1, 'F', 'G', 'H', 'K' and 'O'). Delineation of these structures would require multicolour FISH analyses. So far, the frequencies of the complete AZFc deletion (Figure 1, 'B'), the gr/gr deletion (Figure 1, 'C') and the g1/g3 (b2/b3) deletion (Figure 1, 'J' and 'M') and their effects on spermatogenesis have been widely studied (Reijo et al., 1995; Fernandes et al., 2002, 2004; de Vries et al., 2002; Repping et al., 2003a, 2004a; Machev et al., 2004; Page, 2004; Ferlin et al., 2005; Hucklebroich et al., 2005; Lynch et al., 2005). Much less is known about those of the remaining structures, partly due to the difficulty of identifying them (Repping et al., 2003a, 2004a; Lin et al., 2005; Wirtzl et al., 2005).

Our two-step protocol offers a simple and reliable way to screen large populations for these structures and has the potential to identify new Y-chromosome variants carrying structures that do not fit the current models. Once found, these variants can be subjected to further structural analysis using other methods to investigate the nature of the rearrangements and the underlying mechanisms.
Acknowledgements
We thank the Taiwan Supercontrol Biobank for providing DNA samples for our study. The work was supported by the Institute of Biomedical Sciences in Academia Sinica.

References

Submitted on February 8, 2006; accepted on March 15, 2006