High frequencies of Non Allelic Homologous Recombination (NAHR) events at the AZF loci and male infertility risk in Indian men

Deepa Selvi Rani1, Singh Rajender2, Kadupu Pavani2, Gyaneshwer Chaubey3, Avinash A. Rasalkar1, Nalini J. Gupta4, Mamta Deendayal5, Baidyanath Chakravarty4 & Kumarasamy Thangaraj1

Deletions in the AZoospermia Factor (AZF) regions (spermatogenesis loci) on the human Y chromosome are reported as one of the most common causes of severe testiculopathy and spermatogenic defects leading to male infertility, yet not much data is available for Indian infertile men. Therefore, we screened for AZF region deletions in 973 infertile men consisting of 771 azoospermia, 105 oligozoospermia and 97 oligoteratozoospermia cases, along with 587 fertile normozoospermic men. The deletion screening was carried out using AZF-specific markers: STSs (Sequence Tagged Sites), SNVs (Single Nucleotide Variations), PCR-RFLP (Polymerase Chain Reaction - Restriction Fragment Length Polymorphism) analysis of STS amplicons, DNA sequencing and Southern hybridization techniques. Our study revealed deletion events in a total of 29.4% of infertile Indian men. Of these, non-allelic homologous recombination (NAHR) events accounted for 25.8%, which included 3.5% AZFb deletions, 2.3% AZFbc deletions, 6.9% complete AZFc deletions, and 13.1% partial AZFc deletions. We observed 3.2% AZFa deletions and a rare long AZFabc region deletion in 0.5% azoospermic men. This study illustrates how the ethnicity, endogamy and long-time geographical isolation of Indian populations might have played a major role in the high frequencies of deletion events.

Microdeletions of the AZFa, AZFb and AZFc regions on the Y chromosome are frequent genetic causes of severe testiculopathy and spermatogenic defects leading to male infertility1–6. The deletion of AZFa removes a 0.792 Mb region, which includes two single-copy candidate genes USP9Y (DFFRY) and DDX3Y (DBY), is found to have phenotypical consequences in the male germline7–12. The AZFb region contains palindromes P2 to P5, as well as the proximal part of P1. The deletion of AZFb originates from homologous recombination between the palindromes P5/proximal P1, which removes a 6.2 Mb fragment13 along with multiple copies of the genes CDY2, EIF1AY, PRY, RBMY1, SMCY, TTY5, TTY6 and is reported to result in abnormal spermatogenesis14. The deletion of the combined AZFb regions (P4/distal P1 and P5/distal P1) removes 24 genes, most of which are present in multiple copies, and has been associated with male fertility12.

Homologous recombination between b2 and b4 deletes the complete AZFc region of 3.5 Mb, including the genes BPY2, CDY1, CSPG4LY, DAZ, GOLGA2LY, TTY3, TTY4, in variable numbers of copies, and results in spermatogenic failure and male infertility15,16. AZFc deletions (including partial AZFc) are more common, because of non-allelic homologous recombination (NAHR) events occurring between the highly homologous repeated sequences (same orientation) present in the AZFc region12–16. As a result, the deletion of AZFc may cause spermatogenic failure leading to azoospermia/severe oligozoospermia/oligoratozoospermia12,13,15–18. Two candidate genes, DAZ - deleted in azoospermia (DAZ; 400003), and CDY1 – chromodomain Y (CDY1; 400016) in the AZFc region are critical and required for spermatogenesis and their associations with male infertility are well-studied19–22. The DAZ gene family consists of 4 copies in two clusters of doublets; cluster I contains DAZ1/DAZ2 and the cluster II contains DAZ3/DAZ423, which all encode putative RNA-binding proteins. The CDY1

1CSIR-Centre for Cellular and Molecular Biology, Hyderabad, India. 2CSIR-Central Drug Research Institute, Lucknow, India. 3Department of Zoology, Banaras Hindu University, Varanasi, India. 4Institute of Reproductive Medicine, Salt Lake, Kolkata, India. 5Infertility Institute and Research Centre, Hyderabad, India. Correspondence and requests for materials should be addressed to K.T. (email: thangs@ccmb.res.in)
gene family consists of 2 functional copies, one within the DAZ cluster (CDY1a) and the other at the distal end of the DAZ cluster (CDY1b). Four major deletion combinations of these two genes (DAZ1/DAZ2 + CDY1a, DAZ1/ DAZ2 + CDY1b, DAZ3/DAZ4 + CDY1a and DAZ3/DAZ4 + CDY1b) were reported.

The two partial deletions of the AZFc region, namely gr/gr and b1/b3, both remove ~1.6 Mb of the AZFc region. The gr/gr deletion is identified by the deletion of the STS marker sY1291, and the b1/b3 deletion is identified by the absence of additional STS markers sY1191, sY1197, sY1161 and sY1291. Both the gr/gr and b1/b3 deletions retain two copies of the DAZ gene. The gr/gr deletion is the most common deletion type, and is caused by recombination events between the amplicons g1-r1-r2 and g2-r3-r4. The two other types of partial deletions, which result from inversions followed by gr/gr deletions or vice versa, are b2/b3 and b3/b4. The b3/b4 followed by gr/gr deletion also removes a 1.6 Mb segment of AZFc but differs in breakpoints. The b2/b3 deletion removes a 1.8 Mb segment of AZFc, which is identified by the deletion of the STS marker sY1191. Since the b2/b3 deletion is larger than the gr/gr deletion, it may increase the risk of complete AZFc deletion.

Some studies have also reported other rare deletion patterns, proving that the AZFc segment is highly polymorphic. In addition to deletion events, a few duplication events that generate Y chromosome variants with six or eight DAZ copies in the AZFc region have also been reported. Partial deletions of the AZFc region are common and have been extensively studied. However, the impact of all such partial AZFc deletions on male infertility is still a matter of debate. Some studies have shown that gr/gr deletions are fixed on specific haplogroup backgrounds using major bi-allelic markers. Some case/control studies have reported significant biases in distribution of haplogroups indicating that a particular haplogroup was at higher risk for infertility. However, these association studies lack homogeneity due to the geographical origin/environmental factors or of small sample size. Although several studies reported that the deletion of AZF regions on Y chromosome is associated with male infertility, there is no comprehensive study to correlate the deletion of AZF regions with infertility among Indian men. Genetic isolation and endogamy, which are widespread in Indian populations, can play major roles in introducing novel causal variations. Therefore, we undertook the present study to test the following hypotheses: a) whether deletion events of AZF regions on the Y chromosome in the diverse Indian population are associated with infertility; b) if partial deletions of the AZFc region are risk factors for spermatogenic failure among idiopathic infertile India men; c) whether the AZFc partial deletions associated with spermatogenic defects are due of lack of DAZ and CDY1 copies; and d) if any specific Y chromosome haplogroup is associated with any type of AZf deletion type and infertility.

Results

Microdeletions of AZFa, AZFb and AZFc regions on the Y chromosome. Our study revealed a total of 29.4% of AZF regions deletions (Table 1) on the Y chromosome (Fig. 1A), in infertile men. Of these, non-allelic homologous recombination (NAHR) events accounted for 25.8%, which include the AZFb deletions of 3.3% (P5/proximal P1) (Fig. 1B), both AZFbc deletions of 2.3% (P4/distal P1 and P5/distal P1) (Fig. 1B), the complete AZFc deletions of 6.9% (b2/b4) (Fig. 1C), and the partial AZFc deletions of 13.1% (b1/b3 (2.7%) (Fig. 1E); gr/gr (5.1%) (Fig. 1FII,FIII), b2/b3 (3.7%) (Fig. 1GII,GIII); b3/b4 (1.5%) (Fig. 1HII,HIII)]. The deletion of the AZFa region was observed in 31 infertile men (3.2%), consisting of 27 azoospermia (3.5%), 3 oligozoospermia (2.9%) and one oligotestaospermia man (1%) (Table 1). The deletion of AZFb region (P5/proximal P1) was detected in 34 infertile men (3.5%), of which 30 were azoospermic (3.9%) and 4 were oligozoospermic (3.8%) (Table 1). The deletions of both AZFbc regions (P4/distal P1 and P5/distal P1) with the absence of STS markers in both AZFb and AZFc regions were identified in 22 azoospermic men (2.3%). A total of 67 infertile men (6.9%) showed deletion of the complete AZFc region (b2/b4) (Fig. 1C), of which 59 were azoospermic (7.7%), 6 were oligozoospermic (5.7%) and 2 were oligotestaospermic (2.1%) (Table 1). A very rare long Yq deletion removing all of the three AZFabc regions (absence of STS markers in AZFa, AZFb and AZFc regions) was detected in 5 azoospermic men (0.5%). Importantly, none of these deletions AZFa, AZFb, AZFc, AZFbc or AZFabc was observed in the control men, signifying the importance of these AZF regions in spermatogenesis.

Partial AZF region deletions. We identified partial AZFc deletions in a total of 127 infertile (13.1%) and 12 fertile normospermic control men (2.0%), namely b1/b3 (Fig. 1E), gr/gr (Fig. 1FII,FIII), b2/b3 (Fig. 1GII,GIII), and b3/b4 (Fig. 1HII,HIII) (Table 1). The b1/b3 deletion (partial AZFc deletion removing ~1.6 Mb DNA fragment) was detected in 26 (2.7%) infertile cases and one normospermic fertile man (0.2%) (Table 1). Of the 26 infertile men with b1/b3 deletions, 23 were azoospermic (3.0%), 2 were oligozoospermic (1.9%), one oligotestaospermic (1%) (Table 1), suggesting its importance in spermatogenesis.

The gr/gr (partial AZFc) deletion was detected in 50 infertile (5.1%) and 9 fertile control men (1.5%) (Fig. 1F; Table 1). The gr/gr (partial AZFc) deletion arises from 3 patterns of non-allelic homologous recombination events (NAHR) that occur between the amplicons g1/g2, r1/r3, and r2/r4. We detected 38 out of 50 infertile men (3.9%) and 6 out of 9 fertile men (10.2%) (Table 1) with the absence of the sY1291, DAZ cluster I (DAZ1 + 2), and CDY1a, which show the gr g2 (Fig. 1FII) and r1/r3 (Fig. 1FIIID) NAHR patterns. Of the 38 infertile men, 25 were azoospermic (3.2%), 11 were oligozoospermic (10.5%) and 2 were oligotestaospermic (2%) (Table 1). The remaining 12 out of 50 infertile men (1.2%) and 3 out of 9 controls (0.5%) (Table 1) identified as gr/gr with the removal of sY1291, DAZ cluster II (DAZ3 + 4) and a copy of CDY1a show the r2/r4 NAHR pattern (Fig. 1FIII). Of the 12 infertile men, 11 were azoospermic (1.4%), and one oligozoospermic (0.9%) (Table 1).

Interestingly, we also observed two more inversions, b2/b3 (Fig. 1G) and b3/b4 (Fig. 1H) followed by gr/gr deletions or vice versa. The b2/b3 inversion (Fig. 1G) removes a 1.8 Mb segment of the AZFc region and was identified in 36 infertile cases (3.7%) and 2 controls (0.34%) (Table 1). The b2/b3 inversion was also found to follow 3 patterns of NAHR, which occur between the amplicons g1/g3, r2/r3, and r1/r4. The 11 (1.1%) out of 36 (3.7%) infertile men with b2/b3 deletions (consisting of 10 azoospermia 1.2% and one oligozoospermia 0.9%) (Table 1)
AZFa, AZFb, AZFc and AZFabc deletions on Y chromosome

8 gr/gr (r2/r4), deletions of
6 (b1/b3 deletions
5 AZFc (b2/b4) 59 7.7 46.96 0.083
4 AZFb,c 22 2.9 17.03 0.029
3 AZFa 27 3.5 20.7 0.036
2 AZFb 30 3.9 23.36 0.040
1 AZFa,c 5 0.7 <5 0.007
AZF deletions 143 18.7 157.6 0.228

The AZFa Partial deletions

(a) The b1/b3 deletions
6 b1/b3 deletions of Y1191, Y1197, Y1161, Y1291, DAZ1 + 2, CDY1a 23 3 15.19 0.031
(b) The gr/gr deletions
7 gr/gr (g1/g2, r1/r3), deletions of Y1291, DAZ1 + 2, CDY1a 25 3.2 7.37 0.034
8 gr/gr (r2/r4), deletions of Y1291, DAZ3 + 4, CDY1a 11 1.4 2.74 0.015
(c) The b2/b3 inversion followed by gr/gr deletion or vice versa
9 b2/b3 (g1/g3), Y1191, Y1206, DAZ3 + 4, CDY1a 10 1.3 <5 0.013
10 b2/b3 (r1/r4), Y1191, DAZ1 + 2, CDY1a 17 2.2 8.4 0.023
11 b2/b3 (r1/r4), Y1191, DAZ3 + 4, CDY1a 7 1 <5 0.009
12 b2/b3 (g1/g3) deletions 34 4.5 21.38 0.046
(d) The b3/b4 inversion followed by gr/gr deletion or vice versa
12 b3/b4, DAZ1 + 2, CDY1b 9 1.2 <5 0.012
13 b3/b4, DAZ3 + 4, CDY1b 8 0.8 <5 0.008
14 b3/b4, gr/gr deletions 15 2 11.55 0.0198 0.00688

The AZF partial deletions (or) gr/gr deletions (g1/g3 + b2/b3 and b3/b4 inversion followed by gr/gr deletions) (5.1 + 3.7 = 13.1 = 10.4) 101 10.4 39.75 0.116 0.0001 11 1.87
Total AZF partial deletions (b1/b3 + g1/g3 + b2/b3 and b3/b4 inversion followed by gr/gr deletions) (2.7 + 5.1 + 3.7 + 13.1 = 26.6) 127 13.1 54.66 0.150 0.0001 12 1.04
The total AZF deletions (both complete and partial deletions) (139 + 272 = 29.4%) 286 29.4 177.21 0.416 0.0001 12 1.04

Table 1. Different types of deletion events observed at AZF locus in infertile and control men.

show the g1/g3 NAHR pattern (Fig. 1G). 18 out of 36 infertile (17 azoospermic and 1 oligospermic man) and 2 fertile control men (0.34%) with the b2/b3 deletion were identified with the r2/r3 (Fig. 1GII) NAHR pattern of the AZFc region. The remaining 6 out of 15 azoospermic men (0.6%) with the deletions of the AZF regions in infertile men, whereas no such deletion was observed among controls. This further strengthens the idea that the classical Yq microdeletions are a cause of spermatogenic failure in the Indian idiopathic infertile men. Our study revealed a very high frequency of deletion events (a total of 29.4%) in Indian infertile men, compared to other populations. We observed 16.4% of classical Yq microdeletions, and these varied greatly in frequency among the populations, mainly due to the ethnic background, geographical

Discussion

Yq microdeletions are well-established causative factors for quantitative decline of spermatozoa and can lead to spermatogenic failure. In the present study, we identified very high frequencies of classical Yq microdeletions of the AZF regions in infertile men, whereas no such deletion was observed among controls. This further strengthens the idea that the classical Yq microdeletions are a cause of spermatogenic failure in the Indian idiopathic infertile men. Our study revealed a very high frequency of deletion events (a total of 29.4%) in Indian infertile men, compared to other populations. We observed 16.4% of classical Yq microdeletions, and these varied greatly in frequency among the populations, mainly due to the ethnic background, geographical
region or case-control selection criteria. The AZFc region is extremely rich in repetitive sequences and is organized as amplicons, and therefore a number of possible partial AZFc deletions (gr/gr, b1/b2, b2/b3, b3/b4) are proposed to be important risk factors for spermatogenic failure and male infertility. However, a few partial AZFc deletion studies have failed to show any such association with male infertility.

In two meta-analyses, one consisting of seven studies reported significant association of gr/gr deletions with less motile sperm with low sperm count, and another comprising of 18 case-control studies also established a strong relationship between gr/gr deletion and male infertility. A few independent studies have also reported that the gr/gr deletion was more common among infertile men with azoo/oligozoospermia than in men with normozoospermia, suggesting that the deletion might be a significant risk factor for spermatogenic failure. However, others failed to show any phenotypic impact of gr/gr deletions on spermatogenic failure. Therefore, it is extremely important to study the frequencies of AZFc partial deletions and their association with male infertility.

Figure 1. (A) Schematic representation of the human Y chromosome structure. (B) Schematic diagram showing NAHR events between the PS/proximal P1, P4/distal P1, P5/distal P1 and b2/b4 amplicons. (C) Schematic diagram showing the location of direct and inverted repeat sequences, transcription units and STSs markers of the AZFc region, containing the genes and transcription units.BPY2, CDY1, CSPG4LY, DAZ, GOLGA2LY, TTY3 and TTY4 in variable numbers of copies. (D) Schematic picture showing the complete deletion of AZFc results from recombination between the b2 and b4 amplicons, which removes 3.5 Mb of DNA. (E) Schematic diagram showing the b1/b3 deletion removes a ~1.6 Mb segment of the AZFc region. (F) Schematic diagram showing the gr/gr deletion, it is the most common deletion type caused by the recombination events between the amplicons g1-r1-r2 and g2-r3-r4. (G) Schematic diagram showing the b2/b3 inversion followed by gr/gr deletion or vice versa remove 1.8 Mb segment of AZFc region. (H) Schematic diagram showing the b3/b4 inversion followed by gr/gr deletion, or vice versa, removes a ~1.6 Mb segment of the AZFc region. Deleted regions are shown as faded, and the possible combinations of inversion and recombination events are given.
fertility in Indian idiopathic infertile men. Our study showed that gr/gr deletions are more frequent among men with oligozoospermia (11.4%) than azoospermia (4.6%) and than in oligoteratozoospermia (2.1%); as expected, the prevalence is very low in controls (1.53%) (Table 1), suggesting that these partial deletions might also be a significant risk factor for spermatogenic failure in Indian idiopathic infertile men. Therefore, ethnic-specific differences in gr/gr deletion frequencies and their association with infertility are evident. The previous studies have suggested that the gr/gr deletion frequency in the patient group was higher in the Asians (~10%) compared to the Europeans (~4.5%)71. Nevertheless, it is yet to be clarified whether the partial deletions (gr/gr, b1/b2, b2/b3, b3/b4) and their association is because of the lack of DAZ copies or due to other intervening genes that are also deleted. Some studies have suggested that the putative deletions of the BPY2 and CDY1 genes were associated differentially with the distinct DAZ gene copy deletions that affect sperm pathology leading to infertility25,72. A few studies have also reported that the gr/gr deletion was neutral because of unknown compensatory mechanisms that had rescued the deleterious gr/gr deletion effect25,42. Some individual reports have observed b1/b3 deletions in their population15,73. We in the present study detected 2.7% of b1/b3 deletions in Indian infertile men but 0.17% in controls, (P = 0.0002) (Table 1) suggesting a strong association of this deletion type with infertility in Indian idiopathic infertile men. Our study even detected b2/b3 (3.7%) and b3/b4 (1.5%) deletions including the deletion of two copies of DAZ and a copy of CDY1. The b2/b3 deletions removing 1.8 Mb DNA segment of AZFc region was reported in few studies16,58,69,74–76. The increased length of the b2/b3 deletion may raise the risk of complete AZFc deletion (b2/b4)16,42. We observed the b2/b3 deletion in 36 infertile men (3.7%) and 2 fertile men (0.34%), showing a (P < 0.0001) statistically significant difference between cases and controls (Table 1). The prevalence of the b2/b3 deletion in the present study was greater than reported previously in the Italian, Moroccan and North Indian populations57,58,61,74 but lower than in the Han-Chinese population (9.2%)42,57,58,75 or in Indian Dravidian men (7.21%)73.

Previous studies have reported that the gr/gr deletion was fixed in haplogroups D2b and Q1 in the Japanese and Chinese populations, respectively15,37,42. In the Northern Eurasian population, the b2/b3 partial deletion was fixed with haplogroup N16. However, some other studies have proposed that the b2/b3 deletion is different in different haplogroups58,69,74. Haplogrouping of 973 infertile and 587 normozoospermic fertile men with 24 Y chromosome binary markers40,41 in the present study revealed 8 major haplogroups, of which H1a-M82 and R1a-M17 were the two major haplogroups among both infertile and fertile men (Fig. 2). In India, the overall frequencies of haplogroups H1a-M82 and R1a-M17 were reported to be 40% and 17%, respectively77. Our results are also consistent with the general trend of Indian populations, where H1a-M82 is the most frequent haplogroup. These 8 haplogroups are common throughout India and are present among all the four major linguistic families8. Haplogroup H1a-M82 is an autochthonous haplogroup, whereas R1a-M17 is shared with the West Eurasian populations. Our previous studies have revealed that the people of Indian subcontinent are unique in their origin and differ significantly from the rest of the world in terms of their genetic affinities and disease

| Types of deletions | Category | C | D | F | H1 | J2a* | L1* | D2* | R1* | R1a* | R1b* | IG | Category Total | Total |
|-------------------|----------|---|---|---|----|------|-----|------|-----|------|-----|---|---------------|-------|
| AZF deletions     | Azoo     | 6 | 1 | - | 62 | 2 | 1 | 1 | 2 | 1 | 9 | 42 | 6 | 10 | 143 | 159 |
|                  | OligoOT  | - | - | - | 7 | - | 1 | - | 1 | - | 1 | 4 | 1 | 1 | 16 |
| b1/b2 deletions  | Azoo     | 1 | 1 | - | 11 | - | 1 | - | 1 | - | - | 6 | - | - | 23 | 26 |
|                  | OligoOT  | - | - | - | 2 | - | - | - | - | - | - | 1 | - | - | 3 |
| gr/gr deletions  | Azoo     | - | - | 1 | 23 | - | 1 | 1 | - | - | 8 | 1 | 1 | - | 36 | 50 |
|                  | OligoOT  | - | - | - | 6 | - | - | - | - | - | - | 5 | - | - | 14 | 14 |
| b2/b3 deletions  | Azoo     | 1 | 1 | - | 16 | - | 1 | - | 1 | - | 11 | 1 | 1 | - | 34 | 36 |
|                  | OligoOT  | - | - | - | 1 | - | - | - | - | - | - | 1 | - | - | 2 |
| b3/b4 deletions  | Azoo     | - | - | - | 5 | - | - | - | 1 | - | - | 2 | - | - | 9 | 12 |
|                  | OligoOT  | - | - | - | 2 | - | - | - | - | - | - | 1 | - | - | 3 |
| Total             | All      | 8 | 3 | 1 | 137 | 2 | 4 | 2 | 7 | 1 | 11 | 83 | 10 | 14 | 283 | 283 |
| Fertile samples with deletions | - | - | 0 | 7 | - | - | - | - | - | - | 5 | - | - | 12 | 12 |
| Infertile samples without deletions | 10 | 2 | 2 | 256 | 10 | 22 | 9 | 27 | 11 | 50 | 220 | 36 | 35 | 690 | 690 |
| Fertile samples without deletions | 13 | 3 | 3 | 263 | 11 | 14 | 1 | 11 | 8 | 30 | 146 | 20 | 52 | 575 | 575 |

Figure 2. Top: Y-chromosomal phylogenetic tree showing the markers tested and the haplogroups they define. Bottom: Distributions and frequencies of haplogroups in fertile and infertile men with deletion and without deletions.
susceptibility. Therefore, heterogeneity in terms of the haplogroups observed among Indians and the Chinese are not surprising.

Though our study included a substantial total sample size, subsamples such as oligozoospermic and oligoteratozoospermic patients remain small, which will have limited some statistical inferences. Even after the analysis of the Y chromosome partial deletions, the etiology remains unknown in a large proportion of the infertile men. Further analyses of genes, not only of the Y chromosome and the autosomes are required to understand the genetic causes of male infertility in a greater percentage of the idiopathic infertile cases.

Conclusions

Our study revealed a very high frequency of AZF deletion events in Indian infertile men (29.4%) compared to other populations. We observed 16.4% of AZF region deletions exclusively in infertile men, consisting of AZFa (3.2%), AZFb (3.5%), AZFc (6.9%), AZFbc (2.3%) and AZFab (0.5%). However, these deletion frequencies differ greatly in different populations, mainly due to the ethnic background/case-control selection criteria. We also identified partial AZFc deletions (gr/gr, b1/b3, b2/b3, b3/b4) in 127 infertile men (13.1%). Therefore, ethnic-specific differences in the AZFc partial deletion frequencies and their association with infertility are also evident. We found that gr/gr deletions are more frequent among oligozoospermic (11.4%), than azoospermic (4.6%) or oligoteratozoospermic (2.1%) patients, and that as expected the prevalence is very low in controls (1.53%), suggesting that these partial deletions might be a significant risk factor for spermatogenic failure (low sperm counts) in Indian idiopathic infertile men. Some studies have suggested that AZFc partial deletions are fixed in specific haplogroups. However, in the present study, we found that the distribution of haplogroups was not different between cases and controls with/without deletions, and that all deletions were rare in controls, suggesting that haplogroup has no role in determining risk associations in Indian infertile men. Thus, in our study we found some AZF deletion events that explained the infertility in these idiopathic infertile men. Indian populations are unique in their origin and have been practicing endogamy for the last two thousand years, and therefore it is important to add a study of the frequencies of AZF deletions on the Y chromosome and their association with fertility in Indian idiopathic infertile men to similar studies from other parts of the world.

Materials and Methods

Ethical statement and samples of infertile and fertile men. The Institutional Ethical Committees (IECs) of the participating institutes approved the study. The experiments were carried in accordance with the relevant guidelines and regulations approved for research on human samples. All the experimental protocols were approved by the IEC of the Centre for Cellular and Molecular Biology (CCMB). Before blood sample collection, the subjects underwent detailed medical and physical examinations. Informed written consents were obtained from all of 973 infertile and 587 fertile control men. The blood samples of 973 infertile men, consisting of 771 azoospermia (complete absence of sperm), 105 oligozoospermia (low sperm count) and 97 oligoteratozoospermia (low sperm count with abnormal shape and size) patients, were collected from the Genetic Clinic, Institute of Reproductive Medicine, Kolkata, India. The blood samples of the remaining 40 oligoteratozoospermic men were collected from the Infertility Institute and Research Centre, Hyderabad, India. In both the hospitals, a team of doctors (urologists and andrologists) performed detailed clinical investigations, which included semen analyses, and recorded complete case histories. In the hospitals, the blood samples were subjected to karyotyping and endocrinological assays, such as for follicle-stimulating hormone (FSH), luteinizing hormone (LH), testosterone (T), prolactin (PRL) and thyroid-stimulating hormone (TSH). Patients included in the study did not exhibit any obstructions, pelvic injury or major illness, karyotype abnormalities or endocrinological defects.

The 587 fertile normozoospermic control men with matched ethnic backgrounds had normal semen parameters (>20 × 10⁶ sperm/ml semen fluid with normal motility and morphology), according to the World Health Organization guidelines and normal levels of inhibin B, testosterone T, LH and FSH. They volunteered themselves to be included in this study as controls. About 5.0 ml of blood were collected from 537 and 50 infertile men from 2 hospitals: a) The Genetic Clinic, Institute of Reproductive Medicine, Kolkata, India and b) The Infertility Institute and Research Centre, Hyderabad, India, respectively. In addition, all the controls had fathered at least one child, each with proven paternity by STR-based DNA fingerprinting (Profiler Plus; Applied Biosystem, Foster City, USA), and were enrolled in the study after obtaining written consents. DNA was isolated from all the blood samples of infertile and fertile control men using the method published elsewhere.

Polymerase Chain Reaction (PCR). Primer sequences of the STS, and SNV markers were obtained from (www.ncbi.nlm.nih.gov/entrez/) and synthesized using an ABI394 oligo-synthesizer (Perkin Elmer, Foster City, California, USA). The Polymerase Chain Reaction (PCR) was performed in 0.2 ml thin-walled tubes using 50 ng of DNA, 10 μM of the STS primers mentioned above, 100 μM dNTPs, 10X PCR buffer containing 1.5 mM MgCl₂, and 2 units of AmpliTaq Gold (Perkin Elmer). Amplification was carried out in a MJ Research Thermal Cycler (Waltham, MA 02451, USA) using the amplification conditions: 94 °C for 5 minutes, 35 cycles at 94 °C for 45 seconds, 60 °C for 45 seconds and 72 °C for 1 minute, followed by the final extension at 72 °C for 5 minutes. The PCR products were size fractionated using 2% agarose gel electrophoresis and detected by staining with ethidium bromide.

Mapping of the AZFa, AZFb, AZFc complete (b2/b4), and AZFc partial deletions (gr/gr, b1/b2, b2/b3, b3/b4 including DAZ and CDY1 gene CNVs). The AZFa region deletion was detected using STSs markers: sY82, sY83, sY84, sY86, sY740, sY741, sY742, sY743, sY746, sY615, DBY and USP9Y. The AZFb deletion was identified by screening of STS markers: sY198, sY100, sY113, sY121, sY124, sY127, sY128, sY130, sY134, sY142, sY143, sY145 and sY146. To define the AZFc complete (b2/b4), and partial (gr/gr, b1/b2, b2/b3) deletions, we used STSs markers: sY153, sY158, sY242, sY254, sY255, sY1258, sY1161, sY1197, sY1191, sY1291, sY1206.
and sY1201 present within the amplicons. Further, to detect the presence/absence of particular DAZ gene copies in the AZFc region, we used multiple approaches. We first directly sequenced the PCR amplified products of following 3 additional STS markers sY587, sY581 and sY586 allowing the differentiation of DAZ gene copies. (B1) The sequence electropherogram shows allele ‘T’ in the left panel, allele ‘C’ in the right panel and both alleles in the middle panel. (B2) Schematic representation of the sY587 amplicon digested with restriction enzyme Dgrl. (B3) The first lane of both panels is the digested amplicon of fertile controls. The second lane of the first panel shows the absence of the 73 bp and 122 bp fragments, suggesting the deletion of DAZ1/DAZ2. The second lane of the second panel shows the absence of the 195 bp fragment, suggesting the deletion of DAZ3/DAZ4. (C1) The sequence electropherogram shows allele ‘C’ in the left panel, allele ‘T’ in the right panel and both alleles in the middle panel. (C2) Schematic representation of the sY581 amplicon digested with restriction enzyme Sau3A. (C3) The first lane of both panels is the digested amplicon of fertile men. The second lane of the first panel shows the absence of the 189 bp fragment, which suggests the deletion of DAZ1/DAZ4. The second lane of the second panel shows the absence of the 130 bp fragment, suggesting the deletion of DAZ2/DAZ3. (D1) The sequence electropherogram shows allele ‘C’ in the right panel and both ‘C’ and ‘T’ alleles in the left panel. (D2) Schematic representation of the amplicon of sY586 digested with TaqI. (D3) The first lane of the panel is the digested amplicon of fertile man. The second lane shows the absence of the 301 bp fragment, suggesting the deletion of DAZ2. (E) PCR-RFLP of CDY1-specific SNV; CDY1-7750 was amplified and digested with PvuII and size-fractionated in a 2.0% agarose gel. CDY1b has a PvuII restriction site, but CDY1a does not have a PvuII restriction site. Lanes 1, 2, 3 and 14 – showing the intact CDY1a and the digested CDY1b. Lanes 4, 7, 8 and 12 – showing the presence of only CDY1a. Lanes 5, 6, 9, 10, 11, 13 and 15 – showing the presence of only CDY1b. Lane 16 undigested DNA and lane M marker.
Southern hybridization. Southern hybridization was carried out to further confirm the DAZ gene copy deletions, using representative DNA samples (2–4) of different deletion combinations of DAZ gene copies (wherever sufficient DNA was available) (Fig. 4). About 5.0 ug of DNA from chosen infertile men was digested separately with EcoRI and TaqI restriction enzymes. After completion of digestions, the DNA samples were size fractionated using 1.0% agarose gel and then transferred onto a Hybond N+ nylon membrane (Amersham Pharmacia, Buckinghamshire, United Kingdom) by capillary transfer, using 0.4 N NaOH. The membrane was further hybridized at 65 °C with a DAZ-specific hybridization probe 49f, radiolabeled with 32P (BRIT, Jonaki, India) in 0.5 M phosphate buffer and 7% SDS. After hybridization, excess probe was washed from the membrane with three changes of solution containing 2X SSC and 1% SDS at 65 °C for 45 minutes. Washed blots were exposed to a phosphor imager screen (Fuji, Japan) and images were acquired after 2 hrs (Fig. 4).

Identification of Y-chromosomal haplogroups. All the infertile (973) and fertile control (587) men were haplogrouped using 24 Y chromosome binary markers40,41. PCR was carried out for all 24 binary markers, the amplified products were then directly sequenced using Sanger sequencing, and the haplogroups were assigned based on the sequence.

Statistical analysis. The types of deletion combinations observed in our study were tabulated (Table 1), and comparisons were made between each category versus control using biostatistical tools available online (http://faculty.vassar.edu/lowry/VassarStats.html). To confirm the results, statistical tests were repeated at least twice. P values less than 0.05 were considered as statistically significant changes.

References
1. Lahn, B. T. & Page, D. C. Functional coherence of the human Y chromosome. Science 278, 675–680 (1997).
2. Krausz, C. & Degl’Innocenti, S. Y chromosome and male infertility: update, 2006. Frontiers in bioscience: a journal and virtual library 11, 3049–3061 (2006).
3. Vogt, P. H. et al. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. Human molecular genetics 5, 933–943 (1996).
4. Simoni, M., Tuttelmann, F., Gromoll, J., & Nieschlag, E. Clinical consequences of microdeletions of the Y chromosome: the extended Munster experience. Reproductive biomedicine online 16, 289–303 (2008).
5. Krausz, C., Forti, G. & McElreavey, K. The Y chromosome and male fertility and infertility. International journal of andrology 26, 70–75 (2003).
6. Skakel, H. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423, 825–837, https://doi.org/10.1038/nature01722 (2003).
7. Kamp, C. et al. High deletion frequency of the complete AZFa sequence in men with Sertoli-cell-only syndrome. Molecular human reproduction 7, 987–994 (2001).
8. Sun, C. et al. Deletion of azoospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses. Human molecular genetics 9, 2291–2296 (2000).
9. Sun, C. et al. An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. Nature genetics 23, 429–432, https://doi.org/10.1038/ng.0889 (1999).
10. Ferlin, A., de Vries, J. W., Machev, N., Moro, E. & Rossi, A. CDY1 analysis in infertile patients with DAZ deletions. Human molecular genetics 9, 2291–2296 (2000).
11. Ferlin, A., Moro, E., Rossi, A. & Foresta, C. Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism. Journal of medical genetics 37, 752–758 (2000).
12. Repping, S. et al. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nature genetics 10, 383–393, https://doi.org/10.1038/ng.0883 (1995).
13. Habermann, B. et al. DAZ (Deleted in AZoospermia) genes encode proteins located in human late spermatids and in sperm tails. Human reproduction 13, 363–369 (1998).
14. Repping, S. et al. Y-chromosomal subdeletions in infertile men. Journal of medical genetics 37, 252–256, https://doi.org/10.1136/jmg.2004.022111 (2004).
15. Repping, S. et al. Phenotypic variation within European carriers of the Y-chromosomal gr/gr deletion is independent of Y-chromosomal background. Journal of medical genetics 46, 21–31, https://doi.org/10.1136/jmg.2008.059915 (2009).
16. Fernandez, S. et al. High frequency of DAZ1/DAZ2 gene deletions in patients with severe oligozoospermia. Molecular human reproduction 8, 286–298 (2002).
17. Fernandez, S. et al. A large AZFc deletion removes DAZ3/DAZ4 and nearby genes from men in Y haplogroup N. American journal of human genetics 74, 180–187, https://doi.org/10.1086/381132 (2004).
18. Machev, N. et al. Sequence family variant loss from the AZFc interval of the human Y chromosome, but not gene copy loss, is strongly associated with male infertility. Journal of medical genetics 41, 814–825, https://doi.org/10.1136/jmg.2004.022111 (2004).
19. Ferlin, A. et al. The human-specific region of the Y chromosome is a mosaic of discrete sequence classes. Nature genetics 42, 209–213, https://doi.org/10.1038/ng.02583 (2010).
20. Giachini, C. et al. The gr/gr deletion(s): a new genetic test in male infertility? Journal of medical genetics 42, 497–502, https://doi.org/10.1136/jmg.2004.028191 (2005).
21. Lahn, B. T. & Page, D. C. Retention of autosomal mRNA yielded testis-specific gene family on human Y chromosome. Nature genetics 21, 429–433, https://doi.org/10.1038/ng7771 (1999).
22. Lahn, B. T. et al. Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. Proceedings of the National Academy of Sciences of the United States of America 99, 8707–8712, https://doi.org/10.1073/pnas.082248899 (2002).
23. Saxena, R. et al. Four DAZ genes in two clusters found in the AZF region of the human Y chromosome. Genomics 83, 1046–1052, https://doi.org/10.1016/j.ygeno.2003.12.018 (2004).
24. Repping, S. et al. High mutation rates have driven extensive structural polymorphism among human Y chromosomes. Nature genetics 38, 463–467, https://doi.org/10.1038/ng1754 (2006).
25. Blanco, P. et al. Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism. Journal of medical genetics 37, 252–258 (2000).
26. Repping, S. et al. Divergent outcomes of intrachromosomal recombination on the human Y chromosome: male infertility and recurrent polymorphism. Journal of medical genetics 37, 752–758 (2000).
27. Repping, S. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature genetics 42, 209–213, https://doi.org/10.1038/ng.02583 (2010).
28. Giachini, C. et al. The gr/gr deletion(s): a new genetic test in male infertility? Journal of medical genetics 42, 497–502, https://doi.org/10.1136/jmg.2004.028191 (2005).
29. Hucklebroich, K. et al. Partial deletions in the AZF region of the Y chromosome occur in men with impaired as well as normal spermatogenesis. Human reproduction 20, 191–197, https://doi.org/10.1093/humrep/dch558 (2005).
30. Lynch, M. et al. The Y chromosome gr/gr subdeletion is associated with male infertility. Molecular human reproduction 11, 507–512, https://doi.org/10.1093/molehr/gah191 (2005).
31. de Vries, J. W. et al. Reduced copy number of DAZ genes in subfertile and infertile men. Fertility and sterility 77, 68–72 (2002).
32. de Vries, J. W. et al. Clinical relevance of partial AZFc deletions. Fertility and sterility 78, 1209–1214 (2002).
33. Vogt, P. H. AZF deletions and Y chromosome haplogroups: history and update based on sequence. Human reproduction update 11, 319–338, https://doi.org/10.1093/humupd/dm017 (2005).
34. Stouffs, K. et al. Do we need to search for gr/gr deletions in infertile men in a clinical setting? Human reproduction 23, 1193–1199, https://doi.org/10.1093/humrep/den069 (2008).
35. Yang, Y. et al. Differential effect of specific gr/gr deletion subtypes on spermatogenesis in the Chinese Han population. International journal of andrology 33, 745–754, https://doi.org/10.1111/1365-2065.2009.01015.x (2010).
36. Lu, Y. et al. Polymorphisms associated with the DAZ genes on the human Y chromosome. Genomics 86, 431–438, https://doi.org/10.1016/j.ygeno.2005.07.003 (2005).
37. Piertz, K., Zorn, B. & Peterlin, B. Copy number of DAZ genes in infertile men. Fertility and sterility 84, 1522–1525, https://doi.org/10.1016/j.fertnstert.2005.06.021 (2005).
38. Jobling, M. A. & Tyler-Smith, C. The human Y chromosome: an evolutionary marker comes of age. Nature reviews. Genetics 4, 998–612, https://doi.org/10.1038/nrg1124 (2003).
39. Consortium, Y. C. A nomenclature system for the tree of human Y-chromosomal binary haplogroups. Genome research 12, 339–348, https://doi.org/10.1101/gr.217602 (2002).
40. Lu, C. et al. The b2/b3 subdeletion shows higher risk of spermatogenic failure and higher frequency of complete AZF deletions than the gr/gr subdeletion in a Chinese population. Human molecular genetics 18, 1122–1130, https://doi.org/10.1093/hmg/ddn427 (2009).
41. Ferlin, A. et al. Y chromosome haplogroups and susceptibility to testicular cancer. Molecular human reproduction 13, 615–619, https://doi.org/10.1093/molehr/gam052 (2007).
42. Jobling, M. A. & Tyler-Smith, C. Human Y-chromosome variation in the genome-sequencing era. Nature reviews. Genetics 18, 485–497, https://doi.org/10.1038/nrg.2017.36 (2017).
Acknowledgements
We express our deep condolence on the passing away of our mentor, Dr. Lalji Singh. We thank Dr. Chris Tyler-Smith for providing probe – 49f, and editing the manuscript. KT was supported by FTT project (MLP0113) fund from the Council of Scientific and Industrial Research (CSIR), Government of India; and the Department of Biotechnology (GAP0514), Government of India. GC was supported by National Geographic explore grant HJ3-182R-18.

Author Contributions
D.S.R. and K.T. conceived and designed the experiments. D.S.R., S.R. and K.P. performed AZF-specific STS, SNV, PCR-RFLP, Southern analysis and DNA sequencing. D.S.R., G.C. and A.A.R. carried out Y chromosome haplogroup analyses. B.C., N.J.G. and M.D. carried out clinical evaluation, cytogenetic analysis, and contributed the cases and control samples. D.S.R. and K.T. have prepared Figure 1, 3 and 4. G.C. prepared the Figure 2. D.S.R. drafted the manuscript with input from K.T. S.R. and G.C. provided valuable suggestions to the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-42690-0.

Competing Interests: The authors declare no competing interests.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2019