

A biopsy sample reduction approach to identify significant alterations of the testicular transcriptome in the presence of Y-chromosomal microdeletions that are independent of germ cell composition

Heike Cappallo-Obermann · Kathrein von Kopylow · Wolfgang Schulze · Andrej-Nikolai Spiess

Received: 7 May 2010 / Accepted: 15 July 2010 / Published online: 29 July 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract Y-chromosomal microdeletions (YCMD) are the major genetic cause of male infertility. To date, it is not known which global changes are induced by the presence of AZFc or AZFb + c deletions in the human testicular transcriptome. We investigated this question by microarray analysis in which we had to eliminate the ‘germ cell effect’, i.e., the dominating effect of germ cell transcripts due to the quantitative difference in germ cell composition in samples with/without YCMD. This problem was tackled by selecting 26 samples from an initial cohort of 34 samples by their homogeneity in respect to cellular composition as obtained from gene expression clustering. This way, the ‘germ cell effect’ was minimized, and a distinct ‘deletion effect’ became more apparent. Several hundred genes are influenced by YCMD as shown on the three different phenotypes hypospermatogenesis, meiotic arrest, and Sertoli-cell only syndrome. We validated on an independent cohort of samples five genes by quantitative real-time PCR that are expressed in germ cells or the somatic compartment and which are exclusively altered by the presence of YCMD. We conclude that the deletion of Y-chromosomal genes has a significant effect on spermatogenesis by modulating the transcriptional network of the germ cell and somatic compartment.

Introduction

Y-chromosomal microdeletions (YCMD) are the major genetic cause for primary spermatogenic failure. 5–10% of all infertile men can be attributed to a YCMD genotype (Ferlin et al. 2006) although there is also a frequency of 1% in men without fertility problems (Kent-First et al. 1999). Three recurrently deleted regions are implicated in YCMD, namely the *azoospermia factor* regions AZFa, AZFb, and AZFc located in Yq11. Recent sequencing and mapping of the Y-chromosome has led to the understanding that most of the YCMD are the consequence of de novo homologous rearrangements of repeated sequence blocks at the time of recombination (Skaletsky et al. 2003). No single genes within these regions can be linked to spermatogenic failure as each of the deletions is accompanied by the removal of several genes (Krausz et al. 2006), and deletion patterns are so variable in detail that they can hardly be related to a specific phenotype. Although reports exist of an azoospermic phenotype correlating to mutations in the USP9Y gene of the AZFa locus (Sun et al. 1999), this is still in question (Luddi et al. 2009). Linking YCMD to a specific spermatogenic phenotype is further complicated by the observation of partial YCMD (Hucklenbroich et al. 2005) and the development of at least 153 haplogroups during human evolution, so that the functional contribution of the AZF region to spermatogenesis may vary significantly in men with different haplotypes (Vogt 2005).

A minimum of 14 protein-encoding Y gene families localize to AZF regions (Vogt et al. 2008), namely two for AZFa, seven for AZFb, four for AZFc, and one for AZFb + c. The most common deletions pertain to the AZFc (60%), AZFb (16%), and AZFb + c (14%) regions (Cram et al. 2006), but there is evidence for a potential bias between different populations (Simoni et al. 2008).

Electronic supplementary material The online version of this article (doi:10.1007/s00439-010-0865-9) contains supplementary material, which is available to authorized users.

H. Cappallo-Obermann · K. von Kopylow · W. Schulze · A.-N. Spiess (✉)
Department of Andrology, University Hospital
Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany
e-mail: a.spiess@uke.de

Several studies indicate that the expression profile of genes residing within the AZFc/AZFb regions is specific to and essential for individual developmental stages of human spermatogenesis (Kleiman et al. 2007; Lardone et al. 2007; Shinka et al. 2004; Sato et al. 2006; Stouffs et al. 2004). In this report, we interrogated human testicular biopsies from men with hypospermatogenesis (HYS), meiotic arrest (MA), and Sertoli-cell-only syndrome (SCO) by global gene expression profiling, comparing phenotypes presenting no or AZFc/AZFb + c deletions. This approach has the advantage that not only transcripts of the Y-linked genes are under investigation, but one may discover alterations in gene expression that are up-/downstream of the AZF genes and therefore regulatory networks in which AZF genes play a putative role. Unveiling possible effects of AZF deletions on the testicular transcriptome is feasible since only expression data in respect to single AZF genes exist to date, and gene expression profiling of human spermatogenesis has already given vital insight to stage-specific differences in cases of spermatogenic impairment (Lian et al. 2009; Okada et al. 2008; Ellis et al. 2007; Spiess et al. 2007). A major obstacle that had to be overcome was to identify transcriptional regulation resulting from the ‘deletion effect’ and not from the dominating effect of altered germ cell composition as a consequence of YCMD (Ferras et al. 2004), herein termed as ‘germ cell effect’. This problem was minimized by (1) reducing an initial cohort of 34 samples with and without YCMD to a subset of 26 samples that featured similar germ cell composition based on their selection by gene expression profiling, (2) enrichment of differentially expressed genes for the somatic compartment by an intersection approach, and finally, (3) validation of the absence of any residual ‘germ cell effect’ in these samples by querying independent datasets of gene expression in human spermatogenesis.

Materials and methods

Patients and testicular biopsies

Testicular biopsies were obtained from patients presenting at the Department of Andrology, University Hospital Hamburg-Eppendorf, Germany, between August 2004 and August 2006. Informed consent and Ethic Committee Approval was obtained (OB/X/2000), and the study was conducted in accordance with the guidelines of the ‘Helsinki Declaration’. Tissues used for the microarray analysis were taken independently for therapeutic testicular sperm extraction (TESE), RNA extraction, and diagnostic purposes as described previously (Jezek et al. 1998). In total, we selected seven samples with hypospermatogenesis

(4 controls and 3 AZFc), 14 samples with meiotic arrest (10 controls, 3 AZFc, and 1 AZFbc) and 13 samples with Sertoli-cell-only syndrome (7 controls, 5 AZFc, and 1 AZFbc) from a cohort of 700 biopsies. Selection criteria were based on homogeneous histology within the same group, in which the histology was extrapolated from examinations of parallel biopsies from the same testis as part of the routine diagnostic work-up. As a final selection step, results from a trial-TESE were taken into account to exclude any discrepancies between the morphological classification of the first biopsy and the spermatogenic activity in a second biopsy from another area of the same testis (compare Feig et al. 2007).

Screening of YCMD deletions by polymerase chain reaction (PCR)

All samples used in this work were screened for AZFa, AZFb, or AZFc deletions according to the ‘Best practice guidelines for molecular diagnosis of y-chromosomal microdeletions’ (Simoni et al. 2004). Amplicons of the multiplex PCR were analyzed by electrophoresis on MetaPhorTM agarose gels (3%, Cambrex Corporation) in TAE buffer by running for 16 h at 25 V and subsequent ethidium bromide staining.

Microarray analysis of testicular biopsies

Briefly, at least four fragments per testis were processed for histological analysis, post-surgical trial-TESE, and cryopreservation. An additional sample of rice grain size (approximately 30 mg) was submerged in 2 ml RNeasy[®] (Ambion, Austin, TX, USA) for RNA extraction and subsequent gene expression profiling. Total RNA was extracted in RNeasy (Qiagen, Hilden, Germany) and re-purified on RNeasy columns (Qiagen, Hilden, Germany) according to the manufacturers’ protocols. Purity and integrity (28S/18S ratio) were assessed by loading 200 ng aliquots onto RNA 6000 nano assay chips using an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). Only samples with an RNA integrity number higher than 7.5 (RIN, Agilent software) were included in the analyses. Amplification and labeling of cRNA was done according to the manufacturer’s protocol (GE Healthcare, Piscataway, NJ, USA). 10 µg of cRNA were hybridized to a CodeLinkTM Human 20 K Bioarray (GE Healthcare) containing 19,881 gene-specific probes. Arrays were stringency washed, stained with Cy5TM-Streptavidin (GE Healthcare) and washed again according to the manufacturer’s instructions. Slides were dried by centrifugation and immediately scanned on a 428TM Array Scanner (Affymetrix) using Jaguar 2.0 Software. Images were analyzed and quantified using the CodeLinkTM Expression

Analysis Software v4.1 (GE Healthcare) and rigorously quality controlled for hybridization artefacts.

Data analysis and statistics

Genes were filtered that had significant expression values across all samples and average expression higher than the 5th percentile of all experiments. The experiments were normalized by quantile normalization (*R* package *affy*, <http://www.bioconductor.org>). In compliance with the MIAME guidelines (Brazma et al. 2001), raw and processed data files of all samples were deposited in the Gene Expression Omnibus (GSE21613).

The top 200 genes based on their variance across samples or genes that were obtained by statistical filtering were subjected to hierarchical clustering analysis (Manhattan distance, average metric). Additionally, the complete dataset without prior filtering was interrogated by principle component analysis and three-dimensional display of the first three components to visualize the main structure in the data. All clustering was conducted with the freely available TIGR MeV v4.5 software (<http://www.tm4.org/mev.html>). After biopsy sample selection, the top 200 genes based on variance were analyzed in respect to their cluster stability by bootstrap hierarchical clustering (Manhattan distance, average metric).

To select genes whose expressions are altered by AZFb/AZFb + c deletions, genes were filtered by *t*-tests between HYS/HYS-AZFc, MA/MA-AZFc, and MA/MA-AZFb + c. We used an equal-variance *t*-test as homogeneity of variance for the different groups, which was found for 91, 90, and 93% of all 19,227 genes from the complete data set, respectively. The obtained *p* values were subjected to a Benjamini–Hochberg multiple testing correction (Benjamini and Hochberg 1995). An additional *t*-test was conducted between HYS and SCO subtypes (without deletions) to filter a set of genes that define transcriptional changes throughout the complete human spermatogenesis ('germ cell effect'), and an intersection between this list and those obtained from the AZFc/AZFb + c deletion analysis was made. Furthermore, to enrich for genes exhibiting a 'deletion effect', this overlap was removed from the initial AZFc/AZFb + c lists. Finally, to validate the enrichment of AZFc/AZFb + c affected genes with reduced 'germ cell effect', these genes were matched (either by identical probe ID's or HUGO gene abbreviations) with expression data from two different independent microarray analyses of spermatogenic stages using a completely different cohort of samples (GSE4797, NCBI Gene Expression Omnibus; Feig et al. 2007; E-TABM-234, EMBL ArrayExpress; Spiess et al. 2007).

The complete analysis above was done within the freely available *R* statistical environment ([\[r-project.org\]\(http://www.r-project.org\)\). Over-representation analysis of functional groups using gene ontology \(GO\) terms was done using the HUGO gene abbreviations and the DAVID 2008 server \(<http://david.abcc.ncifcrf.gov/>\).](http://www.</p></div><div data-bbox=)

Quantitative real-time PCR (qPCR)

Reverse transcription from total RNA was carried out with the SuperScript[®] III First-Strand Synthesis SuperMix for qRT-PCR as described by the manufacturer (Invitrogen).

Quantitative RT-PCR (qRT-PCR) was performed on a LightCycler[™] (Roche, Basel, Switzerland) using the SYBR[®]Premix ExTaq[™] (Takara Bio Inc.) with working procedures according to the manufacturer's instructions, but with primer concentrations of 0.4 μ M in a total volume of 20 μ l. Primer sequences and amplicon sizes are available in Supplemental Data 4. Cycling conditions were: 30 s 95°C initial denaturation, 30–40 cycles of 10 s at 95°C, 20 s at 60–64°C annealing temperature, and 30 s at 72°C. Transcript levels were normalized to ribosomal protein S27 transcript, whose expression was found to be highly stable over the different samples by microarray expression values. Fold differences were calculated by use of the five-parameter sigmoidal model (Ritz and Spiess 2008). Products were analyzed in respect to correct amplicon size on an Agilent Bioanalyzer (Model 2100; Agilent Technologies, Palo Alto, CA, USA). We validated the most interesting candidates by quantitative real-time PCR on a completely independent test set of 13 samples.

Results

A major issue to resolve in this work was the discovery and selection of genes that are affected by AZFc/AZFb + c deletions in a background of transcriptional changes that are dominated by the depletion of specific germ cell types in the different spermatogenic stages when the deletions are present. To eliminate the dominating effect of germ cell-specific transcript levels within the comparison of two (or more) biopsies, it is an indispensable prerequisite to start with samples that present a homogeneous germ cell phenotype. Figure 1 exemplifies this problem using samples from our initial set. When comparing biopsies presenting continuously hypospermatogenesis with those that have an AZFc deletion and a 'mixed atrophy' phenotype with predominantly Sertoli-cell-only tubules (upper panel), any statistical filtering method will select genes specific to all germ cells that differ between the two states, i.e., spermatogonia, primary/secondary spermatocytes, and round/elongated spermatids. Furthermore, any clustering method will result in groups defining the subtypes of spermatogenic failure. In contrast, if one employs samples

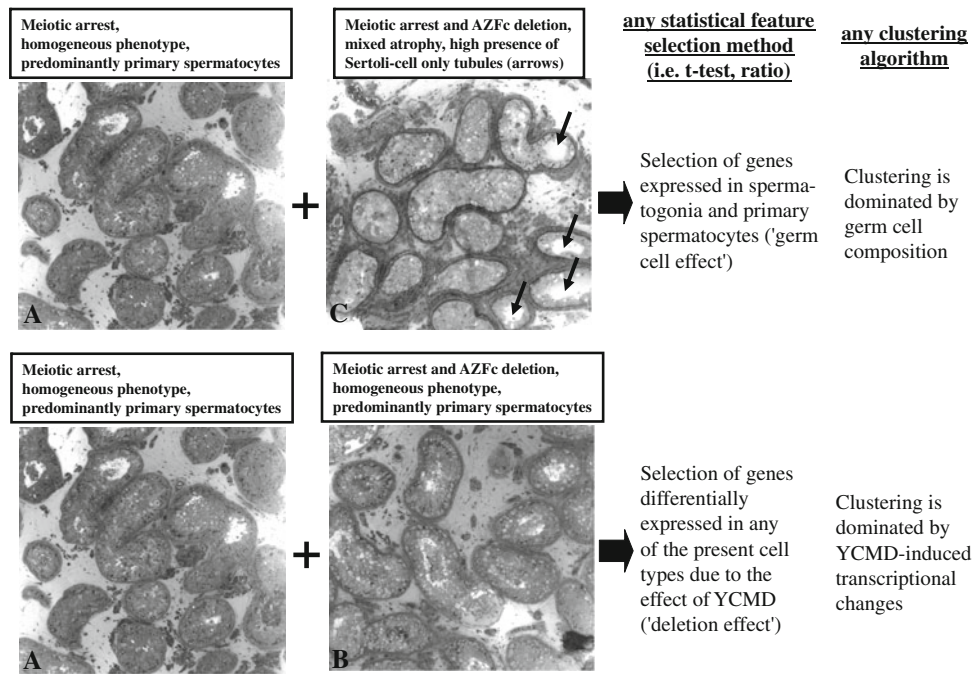


Fig. 1 Definition of the ‘germ-cell effect’ problem involved in filtering differential genes influenced by Y-chromosomal microdeletions. Based on the semi-thin section micrographs of three samples from our collection (A: MA_8; B: MA_11_c; C: MA_13_c; see Figs. 2, 3) which illustrate the difference in relative germ cell composition, the main problem can be stated as follows: if we approach to identify differential transcription as a consequence of the

YCMD, then using testicular biopsies that have a significant difference in relative germ cell composition despite having the same cell types will result in germ cell-specific transcripts as a consequence of the relative difference in transcript levels (*upper panel*). A very high homogeneity in respect to relative germ cell quantity is the prerequisite to identify transcripts that are independent of the ‘germ cell effect’ (*lower panel*)

with similar germ cell composition (*lower panel*), the dominating ‘germ cell effect’ is minimized, and a ‘deletion effect’ might be more apparent.

Our initial set of 34 biopsies (Supplemental Data 1, ‘initial set’) had been classified on the basis of germ cell types that are present in the seminiferous tubules. We grouped the top 200 variant genes by hierarchical clustering (Fig. 2, upper panel). Additional information was acquired about the relative quantity of germ cell composition found in the histological examination of the biopsies (*lower panel*), illustrating the paradigm shift from the most differentiated germ cell type to that which is most frequent: the separation of groups in the dendrogram correlates not so much with initial histological classification but significantly more with the dominant germ cell type present. For instance, sample SCO_10_c clusters within the MA samples, which tallies with the observation of tubules containing spermatocytes (MA) being the most frequent observation in this sample. A comparison of the dendrogram with the histology for all samples clearly shows (and without exception) that all samples cluster according to their quantitatively prevailing germ cell type. This is an important observation as many of the initial samples with YCMD cluster in groups distinct from their most differentiated germ cell type. To complement these results, we

employed principle component analysis (PCA) on the 34 samples using all genes. As PCA is a data reduction technique, one can visualize the first three main variances (principle components) within the gene expression data by three-dimensional display (Supplemental Data 2). Corroborating the dominant effect of relative germ cell composition on the grouping outcome, the main variance of gene expression within the 34 samples is again defined by the most frequent germ cell type (spread along the first principle component PCA1). A similar observation had been made by Ellis et al. (2007) using a mixture of arrest types with variable etiology. The second and third principle components (PCA2 and PCA3) separate a mixture of germ cell composition and YCMD status. It must be noted, however, that PCA1-3 only explained 65% of the total variance found in the data, and therefore other, yet undescribed effects seem likely. We subjected the top 200 variant genes to GO (gene ontology) term analysis and found a highly significant over-representation of terms such as spermatogenesis ($p = 2E-14$), mitosis ($p = 5E-6$) and meiosis ($p = 1E-3$). These results strengthen once again the fact that germ cell transcription is the dominating effect found in our initial sample set. Furthermore, the findings above indicate that the clustering of gene expression profiles from testicular biopsies is extremely sensitive to the

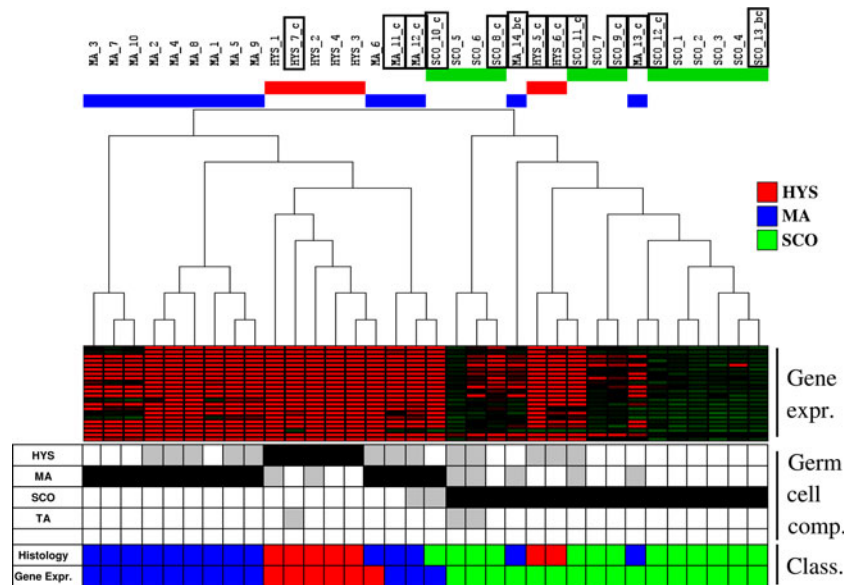


Fig. 2 Visualizing the appropriate pre-selection of testicular biopsies based on gene expression similarity. Testicular biopsies were classified based on their germ cell composition. The gene expression data of all 34 samples were grouped by hierarchical clustering (Pearson correlation, average distance; *red* high expression; *green* low expression), and the membership of each of the samples to the resulting clusters were analyzed. Additionally, the relative germ cell composition of each sample obtained from semi-thin micrographs was noted (*black square* dominating germ cell type within the different

seminiferous tubules; *gray square* less frequent of germ cell type). Finally, a classification of the samples is given, based on either their germ cell classification or on their clustering. Important is here that the samples cluster according to their dominant germ cell type, especially in cases of mixed atrophy in the presence of YCMD deletions. Samples containing AZF deletions are highlighted by *black rectangles*. *HYS* hypospermatogenesis, *MA* meiotic arrest, *SCO* Sertoli-cell only, *TA* tubule atrophy

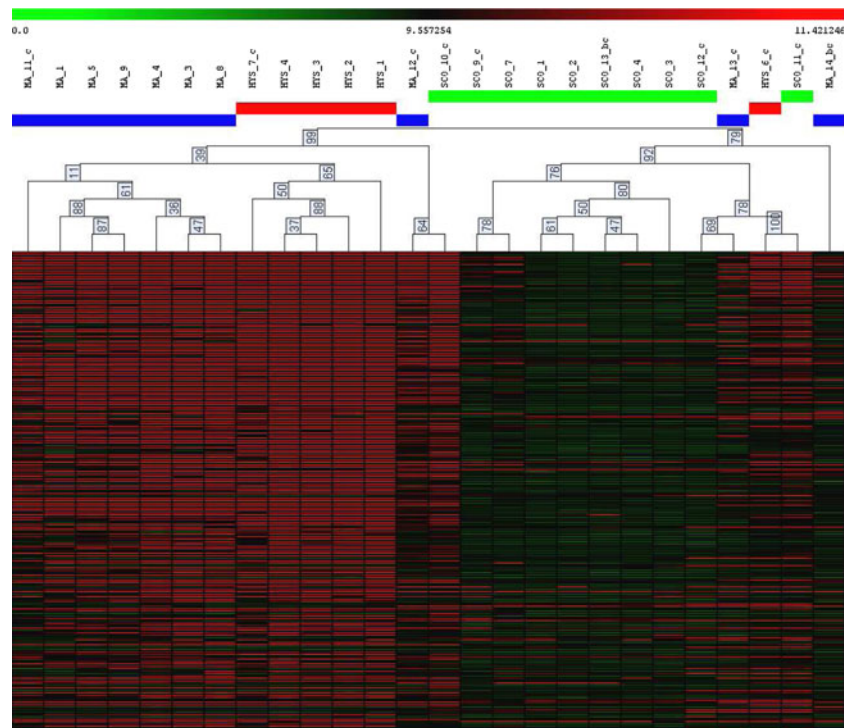
relative germ cell composition, a property that enables us to select homogeneous samples from the clustering outcome.

Based on the grouping of samples in Figs. 2 and 3 we selected 26 samples for further analysis (Supplemental Data 1, ‘selected set’). We clustered the top 200 variant genes from the selected subset of 26 samples by bootstrap hierarchical clustering to acquire some additional information about cluster stabilities (Fig. 3). Similar to the previous grouping structure of the initial set of 34 samples, we obtained three main clusters defining the different pathological conditions HYS, MA, and SCO as the most variant genes correlate also here with the germ cell state. In contrast to the initial clustering, samples with YCMD now grouped with the corresponding germ cell phenotypes without deletion and this with relatively high cluster stability. The procedure therefore resulted in a subset in which the difference in germ cell composition between samples of the same arrest type but with absence/presence of YCMD was minimized.

We conducted *t*-tests with subsequent Benjamini–Hochberg multiple testing correction on the MA and SCO phenotypes of this subset between normal/AZFc samples. For the HYS group, we were only able to calculate a ratio between the averaged normal samples and one sample with

AZFc due to the fact that possible replicates had been sorted out by the sample selection procedure. We filtered 214 genes with fourfold up/down-regulation for the HYS/HYS-AZFc group, 495 genes with $p_{\text{corr}} < 0.05$ for the MA/MA-AZFc group, 22 genes with $p_{\text{corr}} < 0.05$ for the SCO/SCO-AZFc group, and 18 genes with $p_{\text{raw}} < 0.05$ for the SCO/SCO-AZFb + c group (Supplemental Data 1, ‘HYS AZFc’, ‘MA AZFc’, ‘SCO AZFc’, and ‘SCO AZFbc’). This way, we obtained lists of differential gene expression between normal samples and those with either AZFc or AZFbc deletions specific to each of the three different spermatogenic failure types and putatively enriched for the ‘deletion effect’. In terms of gene numbers, the largest difference could be attributed to the MA phenotype with 495 genes that account for roughly 3% of genes that were under investigation. A further interest was to obtain a set of genes that show differential expression irrespective of the spermatogenic state. Interesting to note here is that for both HYS and MA, *DAZ2* was found to be the transcript with the highest ratio of downregulation (~ 20 -fold), which demonstrates that our gene expression data mirrors the depletion of *DAZ* transcripts as a consequence of the AZFc deletion. However, the probes used on our array platform were not able to discriminate between the four isoforms of *DAZ* (*DAZ1-4*). Furthermore, no overlap was observed

Fig. 3 Hierarchical clustering of the selected sample subset. Biopsy samples with names as indicated above were selected from Fig. 2 based on their homogeneous germ cell composition. The top 200 variant genes were clustered by bootstrap hierarchical clustering (Pearson correlation, average distance), yielding the percent cluster stabilities as noted in the boxes at the leaf nodes. The color coding is the same as in Fig. 2. Note the separation of clusters according to the presence of YCMD for each germ cell phenotype



between all three failure types, suggesting that the majority of affected genes correlate with specific germ cell expression.

Gene ontology analysis of differential genes within HYS and MA types in the presence of an AZFc deletion (SCO failed to be analyzed due to the small gene number) indicates functional differences in dependence of the phenotype (Table 1). Nevertheless, terms such as ‘spermatogenesis’, ‘sexual reproduction’, and ‘meiosis’, although having been decreased by the selection procedure by 2–10 orders of magnitude in p values, could still be attributed to the fact that our selection procedure was not optimal. We added another step of minimizing the ‘germ cell effect’ by eliminating all genes by an intersection approach from the obtained lists that are differentially expressed throughout human spermatogenesis, based on 2,568 genes obtained from t -tests between HYS and SCO normal samples, in analogy to Feig et al. (2007) (Supplemental Data 1, ‘Germ cell genes’). Intersections are given in Fig. 4. We removed these intersections from the initial HYS and MA gene lists, and reanalyzed in respect to GO term over-representation (Table 1, terms in bold). This resulted in the elimination of the term ‘sexual reproduction’ from the HYS group and a significant reduction of genes that correlate with the terms ‘spermatogenesis’ and ‘meiosis’ as shown on the increase of p values at 4–8 orders of magnitude. Several other GO terms such as ‘negative regulation of glucocorticoid synthesis’ and ‘killer cell-mediated cytotoxicity’ appeared after elimination of intersecting genes. An overlap analysis

of the reduced sets resulted in seven genes that are up/downregulated in the presence of AZFc in HYS as well as MA phenotype (Supplemental Data 1, ‘HYS AZFc vs. MA AZFc’).

In a last step, we queried the ‘germ cell-depleted’ gene lists on two gene expression datasets of human spermatogenesis that were obtained from a completely different cohort of samples as the analysis of microarray data has the inherent problem of ‘overtraining’ and thus can lack generality (Zervakis et al. 2009). The remaining 178 genes of the reduced HYS dataset (Supplemental Data 1, ‘HYS wo GSG’) showed no significant trend of expression changes throughout spermatogenesis (Supplemental Data 3, upper panel). The analysis of the reduced MA dataset of 232 genes (Supplemental Data 1, ‘MA wo GSG’) revealed a small increase from SCO to MA that was not statistically significant on the averaged values (Supplemental Data 3, lower panel; Codewalk platform: 3.8-fold, $p = 0.27$; Affymetrix platform: 2.23-fold; $p = 0.08$), which may be interpreted as a remnant and non-removable ‘germ cell effect’ although other interpretations are possible (see “Discussion”).

A transcriptional deletion map for Y-chromosomal genes residing within the deleted loci was built, demonstrating that the deletions were mirrored almost completely on the transcriptional level, and that our microarray analysis was sensitive enough to reveal the absence of transcripts as a consequence of the deleted genes (Fig. 5a, b). However, the general impression of Y-chromosomal

Table 1 Over-representation analysis of genes affected by the presence of AZFc or AZFb + c microdeletions

| | Down | Up |
|---------------|--|---|
| HYS | | |
| Gene ontology | Sexual reproduction (2E-3)/ absent Positive regulation of muscle contr. (3E-3)/(6.3E-2) Neg. regulation of glucocorticoid synthesis (5.9E-3) | Neg. regulation of cell cycle (3E-4)/(3E-4) Nucleosome assembly (4E-3)/(4E-3) Heart development (5E-3)/(5E-3) |
| KEGG | Nitrogen metabolism (7E-2)/(7E-2) | Calcium signaling (2E-3)/(2E-3) Systemic lupus erythematosus (7E-3)/(7E-3) GnRH signaling (4E-2)/(4E-2) |
| Biocarta | | Phospholipase C-epsilon pathway (7E-2)/(7E-2) CFTR and B2AR pathway (7E-2) |
| TFBS | IRF1 (1E-4)/(2E-5) | CMYB (3E-4)/(3E-4) |
| MA | | |
| Gene ontology | M-Phase (7E-11)/(4E-3) Meiosis (7E-7)/(6E-2) Spermatogenesis (1E-6)/(7E-2) | Regulation of immune response (1E-5)/(4E-6) Fatty acid metabolism (2E-3)/(2E-2) |
| KEGG | Folate biosynthesis (1E-2)/ absent DNA replication (9E-2)/ absent Killer cell-mediated cytotoxicity (6E-2) | Focal adhesion (1E-2)/(4E-2) Asthma (2E-2)/(2E-2) |
| Biocarta | Role of BRCA1/2/3 in cancer (6E-3)/(2E-2) Activation of PKA (9.2E-1)/ absent | |
| TFBS | CDPCR3HD (6E-5)/(3E-1) | HNF4 (1.5E-4)/(3E-4) |

Genes that are differentially regulated in the presence of AZFc deletions in samples presenting hypospermatogenesis (HYS) or meiotic arrest (MA) were analyzed in respect to functional over-representation. Analysis was done for gene ontology, KEGG/Biocarta pathways, and transcription factor binding sites (TFBS). Numbers in parenthesis reflect the corresponding *p* values. Functional terms in bold letters represent the *p* values after removing the overlap of germ cell-enriched expression (Fig. 5) from the respective groups. If terms were removed by this procedure, they are denoted as 'absent', and new terms are also given in bold letters

transcript levels was relatively heterogeneous in the different samples. We also noticed a complete absence of any signal for NLGN4Y transcripts in sample SCO_13_bc in contrast to the second sample with AZFb + c deletion, MA_14_bc (Supplemental Data 1, 'SCO AZFbc'). To confirm absence of this gene, which is located proximal to the classic AZFb boundary defined by Vogt et al. (1996), we mapped a possible deletion in this region by using gene-specific primers (Supplemental Data 4) for NLGN4Y, its proximal genes VCY and TMSB4Y, and two STS markers proximal and distal of the NLGN4Y gene (sY95 and sY98). For this sample, the complete chromosomal region for AZFb + c reaching proximally to sY95 was deleted (Fig. 5c–e), and therefore, NLGN4Y transcripts are undetectable due to the presence of a relatively rare massive AZFb + c deletion (Silber et al. 1998), also referred to as 'cytogenetic abnormality' (Krausz et al. 1999).

To validate our microarray data, we conducted quantitative real-time PCR on selected genes using 13 independent samples with/without deletions. All transcripts were normalized to the S27a transcript, which was shown in recent microarray investigations (compare GEO Database Accessions GSE4797, GSE6023, GSE18997) and also this study (GSE21613) to be highly stable expressed in

testicular tissue with different states of spermatogenic impairment. It was therefore selected as a good normalization gene to prevent wrong interpretations of expression changes due to differences in the amount of input cDNA in the real-time PCR reactions (Supplemental Data 5A + B). We confirmed a strong downregulation for HIV-1 Rev-binding protein (HRB; 7.5-fold array, 2.8-fold qPCR) in samples with hypospermatogenesis (HYS_7_c). For this sample, there were several other transcripts with more than tenfold upregulation in the presence of an AZFc deletion (marked in red in Supplemental Data 1), but we failed to verify several of these by qPCR (data not shown). Here, the results for genes with a very high upregulation must be taken critically as we cannot identify hybridization artefacts when replicates are lacking. In contrast to this, other genes with interesting functions in respect to AZF deletions that had prominent down-/upregulation in samples with meiotic arrest or Sertoli-cell-only syndrome (meiotic arrest: breast cancer antigen NY-BR-1, apolipoprotein E; Sertoli-cell only: ring finger protein 126; bradykinin receptor B1) could be validated in respect to the direction of expression changes by qPCR, most probably because replicates were available. The ratios obtained from microarray data and qPCR analysis differed slightly, a

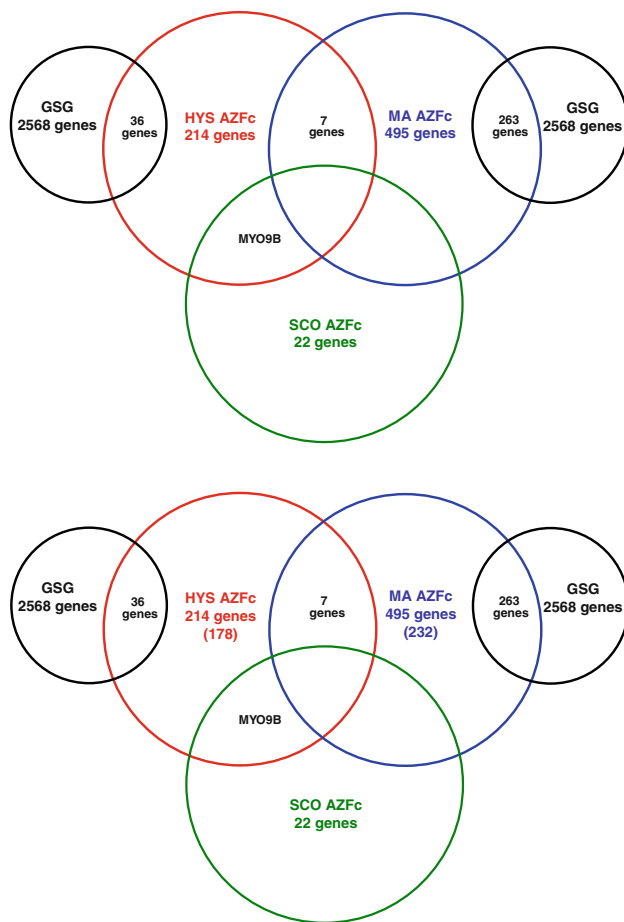


Fig. 4 Venn diagram of gene expression overlap between the filtered groups. Genes affected by the AZFc deletion between the different spermatogenic states (*HYS* hypospermatogenesis, *MA* meiotic arrest, *SCO* Sertoli-cell only) were analyzed in respect to overlap between the groups. Only seven genes were concordant between *HYS* and *MA* subtypes, and no overlap between all three groups was found. Furthermore, the overlap between *HYS* and our obtained list of highly statistically significant genes involved in general germ cell-specific gene expression ('GSG', see "Materials and methods" and Supplemental Data 1) was only 16.8%, but markedly higher in the case of *MA* (53.1%). Color coding same as in Fig. 2

common finding that may be attributed to the effect of 'ratio compression' for microarray data (Wang et al. 2006).

Discussion

It is necessary to dissect deletion-induced disturbances of transcriptional networks from those changes that are a result of relative transcript levels when the germ cell composition is different between non-deleted and deleted samples. We and others have found that the phenotype of 'mixed atrophy', i.e., a high heterogeneity in tubular germ cell composition within the same biopsy, is a common finding in patients with YCMD (Sigg and Hedinger 1981;

Luetjens et al. 2002; Ferras et al. 2004). This quantitative difference in cell type-specific transcripts is significantly shielding and obscuring any deletion-induced effects for downstream analysis with sensitive methods such as microarray analysis, leading to gene expression patterns that reflect only cellular composition of the samples (Okada et al. 2008; Ellis et al. 2007; Feig et al. 2007). This classification, which is based exclusively on the dominating cell type, diverges from any classification using conventional histological descriptions. After enrichment of the 'deletion effect', we obtained a relatively high number of differential genes when YCMD are present. If using these as an indicator of "deletion effect strength", we obtained for the different testicular phenotypes the order $SCO < HYS < MA$. This could be the consequence of decreased *SCO* differential expression pertaining mostly to the somatic compartment of the testis as germ cells are lacking, and meiosis being most strongly perturbed as haploid gamete formation is severely modulated by *DAZ* (Kee et al. 2009; Geoffroy-Siraudin et al. 2007).

The magnitude of transcriptional changes was the highest for those genes that reside within the respective deleted region, i.e., *DAZ* for AZFc deletions and some of the AZFb region located genes (*SMCY*, *EIF1AY*, *CYorf14*, *NLGN4Y*). This anticipated result tallies with the transcriptional changes observed in YCMD samples based on PCR methodology (Kleiman et al. 2007; Lardone et al. 2007). Nevertheless, several hundred other genes were down-/upregulated in the presence of YCMD, although the magnitude of transcriptional changes was mainly in a region of 2–5-fold, indicating a transcriptional system with a possible high redundancy that would result in the high heterogeneity of spermatogenic arrests in the presence of YCMD as frequently observed (Foresta et al. 2001). The precise mechanisms for the relative extensive amount of genes with altered transcription in the presence of YCMD remain to be resolved. However, the disturbance of the testicular transcriptional network by YCMD is most likely a consequence of the lacking RNA-/DNA-binding functionality of the deleted genes which is important for any transcriptional network. For instance, *DAZ*-family members have been shown to bind to specific transcripts in vivo and regulate pre-meiotic translation of important factors such as the mouse *VASA*-homolog *MVH* (Reynolds et al. 2005), which in turn plays a major role in the processing of *PIWI*-interacting RNAs (Kuramochi-Miyagawa et al. 2010). RNA binding has also been observed for the splicing-associated *RBM1Y* (Elliott et al. 2000), while other genes residing within the AZFb locus exhibit transcription factor activity (*HSFY*; Tessari et al. 2004) or are involved in ribosome assembly (*VCY*; Zou et al. 2003). Hence, impairment of the testicular transcriptional network by altering the function of transcription factors through

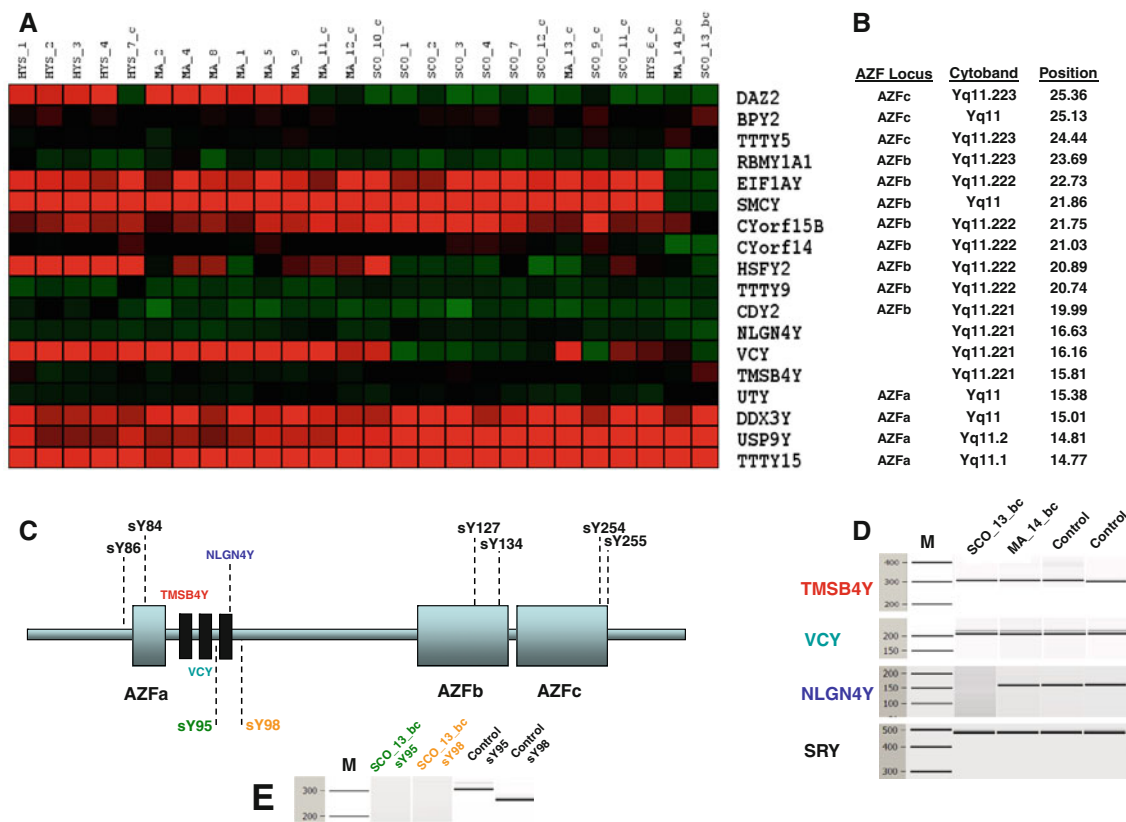


Fig. 5 Compilation of AZFc/AZFb + c deletion-induced transcriptional changes of Y-chromosomal genes. **a** Heatmap representation of Y-chromosomal gene expression for the complete subset of 26 samples. **b** Localization ('Position' given in Megabases) of Y-chromosomal genes that were present on the arrays. **c** Primer localizations used to map the extent of AZFb + c deletion in sample

SCO_13_bc. **d** Electropherogram (Bioanalyzer) after PCR using gene-specific primers as indicated. **e** Mapping the extent of the AZFb + c deletion with primers for sY95 (proximal to NLGN4Y) and sY98 (distal to NLGN4Y). SCO* and SCO_bc* are independent controls. *M* size marker

modulation of their mRNAs or their translation could play a putative role.

From our set of real-time PCR-validated genes, HRB (HIV-1 Rev-binding protein) could be confirmed to be downregulated approximately threefold in the presence of an AZFc deletion. This finding is interesting in respect to HRB's vital function for acrosome formation (Kang-Decker et al. 2001). Due to the lack of replicates in this group, these results should be viewed critically. For the phenotype of meiotic arrest in the presence of AZFc deletions, we confirmed ANKRD30A (breast cancer antigen NY-BR-1) to be downregulated approximately threefold, while APOE (apolipoprotein E) was upregulated approximately twofold. While the former is expressed almost exclusively in primary spermatocytes (Theurillat et al. 2008), the latter is largely expressed in the Leydig cells (Law et al. 1997), suggesting modulations of the interstitial compartment in the presence of AZFc deletions. For ring finger protein 126 (RNF126), which was validated to be approximately twofold upregulated in SCO samples with AZFc, no function in testis is

known to date. In case of AZFb + c deletions with SCO phenotype, we validated BDKRB1 (bradykinin receptor B1) as one of the few transcripts with increased transcript levels (approximately fourfold). As BDKRB1, at least in the rat model, elevates intracellular calcium levels in peritubular cells after bradykinin stimulation (Wennemuth et al. 2003), involvement of the lamina propria is also likely.

In this study, we did not include other factors such as the patients' haplotypes, first, because we wanted to identify transcriptional changes independent of haplotype background, and second, because there are indications that YCMD formation is more or less a stochastic event (Quintana-Murci et al. 2001). Nevertheless, future work should investigate the effect of haplotype background on global testicular expression profiles. Another drawback of our procedure is that it might mask interesting germ cell genes that are influenced by YCMD, and that a significantly higher number of microarray experiments might be necessary to reveal further or more subtle effects of YCMD-induced expression changes.

The obtained list of several hundred differential genes may serve as a starting point for a more detailed analysis of YCMD-induced testicular changes, either in germ cells or the somatic compartment. Moreover, the identified transcriptional networks that are disturbed might be pharmaceutically targeted to reconstitute impaired spermatogenesis in men with YCMD offering potential for male infertility treatment.

Note added During the review process of this paper, we became aware of an investigation with a similar topic that was published by Gatta et al. (2010). In this study, transcriptional changes due to Y-chromosomal microdeletions were investigated by microarray analysis of testicular biopsies without pre-selection based on clustering to minimize the quantitative ‘germ cell effect’. The results in respect to differential genes from this and our study differ substantially.

Acknowledgments Financial support for this work was in part from grant Sp721/1-3 of the German Research Foundation to ANS and HCO. We would like to thank Prof. Christiane Kirchoff for fruitful discussions and Dr. Caroline Feig for technical assistance.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 57:289–300
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 29:365–371
- Cram DS, Osborne E, McLachlan RI (2006) Y chromosome microdeletions: implications for assisted conception. *Med J Aust* 185:433–434
- Elliott DJ, Bourgeois CF, Klink A, Stevenin J, Cooke HJ (2000) A mammalian germ cell-specific RNA-binding protein interacts with ubiquitously expressed proteins involved in splice site selection. *Proc Natl Acad Sci USA* 97:5717–5722
- Ellis PJ, Furlong RA, Conner SJ, Kirkman-Brown J, Afnan M, Barratt C, Griffin DK, Affara NA (2007) Coordinated transcriptional regulation patterns associated with infertility phenotypes in men. *J Med Genet* 44:498–508
- Feig C, Kirchoff C, Ivell R, Naether O, Schulze W, Spiess AN (2007) A new paradigm for profiling testicular gene expression during normal and disturbed human spermatogenesis. *Mol Hum Reprod* 13:33–43
- Ferlin A, Arredi B, Foresta C (2006) Genetic causes of male infertility. *Reprod Toxicol* 22:133–141
- Ferras C, Fernandes S, Marques CJ, Carvalho F, Alves C, Silva J, Sousa M, Barros A (2004) AZF and DAZ gene copy-specific deletion analysis in maturation arrest and Sertoli cell-only syndrome. *Mol Hum Reprod* 10:755–761
- Foresta C, Moro E, Ferlin A (2001) Y chromosome microdeletions and alterations of spermatogenesis. *Endocr Rev* 22:226–239
- Gatta V, Raicu F, Ferlin A, Antonucci I, Scioletti AP, Garolla A, Palka G, Foresta C, Stuppia L (2010) Testis transcriptome analysis in male infertility: new insight on the pathogenesis of oligo-azoospermia in cases with and without AZFc microdeletion. *BMC Genomics* 11:401
- Geoffroy-Siraudin C, Akin-Seiffer I, Metzler-Guillemain C, Ghalmoun-Slaïmi R, Bonzi MF, Levy R, Guichaoua MR (2007) Meiotic abnormalities in patients bearing complete AZFc deletion of Y chromosome. *Hum Reprod* 22:1567–1572
- Hucklenbroich K, Gromoll J, Heinrich M, Hohoff C, Nieschlag E, Simoni M (2005) Partial deletions in the AZFc region of the Y chromosome occur in men with impaired as well as normal spermatogenesis. *Hum Reprod* 20:191–197
- Jezek D, Knuth UA, Schulze W (1998) Successful testicular sperm extraction (TESE) in spite of high serum follicle stimulating hormone and azoospermia: correlation between testicular morphology, TESE results, semen analysis and serum hormone values in 103 infertile men. *Hum Reprod* 13:1230–1234
- Kang-Decker N, Mantchev GT, Juneja SC, McNiven MA, van Deursen JM (2001) Lack of acrosome formation in Hrb-deficient mice. *Science* 294:1531–1533
- Kee K, Angeles VT, Flores M, Nguyen HN, Reijo Pera RA (2009) Human DAZL, DAZ and BOULE genes modulate primordial germ-cell and haploid gamete formation. *Nature* 462:222–225
- Kent-First M, Muallem A, Shultz J, Pryor J, Roberts K, Nolten W, Meisner L, Chandley A, Gouchy G, Jorgensen L, Havighurst T, Grosch J (1999) Defining regions of the Y-chromosome responsible for male infertility and identification of a fourth AZF region (AZFd) by Y-chromosome microdeletion detection. *Mol Reprod Dev* 53:27–41
- Kleiman SE, Yogev L, Hauser R, Botchan A, Maymon BB, Paz G, Yavetz H (2007) Expression profile of AZF genes in testicular biopsies of azoospermic men. *Hum Reprod* 22:151–158
- Krausz C, Quintana-Murci L, Barboux S, Siffroi JP, Rouba H, Delafontaine D, Souleyreau-Therville N, Arvis G, Antoine JM, Erdei E, Taar JP, Tar A, Jeandidier E, Plessis G, Bourgeron T, Dadoune JP, Fellous M, McElreavey K (1999) A high frequency of Y chromosome deletions in males with nonidiopathic infertility. *J Clin Endocrinol Metab* 84:3606–3612
- Krausz C, Degl’Innocenti S, Nuti F, Morelli A, Felici F, Sansone M, Varriale G, Forti G (2006) Natural transmission of USP9Y gene mutations: a new perspective on the role of AZFa genes in male fertility. *Hum Mol Genet* 15:2673–2681
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Takamatsu K, Chuma S, Kojima-Kita K, Shiromoto Y, Asada N, Toyoda A, Fujiyama A, Totoki Y, Shibata T, Kimura T, Nakatsuji N, Noce T, Sasaki H, Nakano T (2010) MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev* 24:887–892
- Lardone MC, Parodi DA, Valdevenito R, Ebensperger M, Piottante A, Madariaga M, Smith R, Pommer R, Zambrano N, Castro A (2007) Quantification of DDX3Y, RBMY1, DAZ and TSPY mRNAs in testes of patients with severe impairment of spermatogenesis. *Mol Hum Reprod* 13:705–712
- Law GL, McGuinness MP, Linder CC, Griswold MD (1997) Expression of apolipoprotein E mRNA in the epithelium and interstitium of the testis and the epididymis. *J Androl* 18:32–42
- Lian J, Zhang X, Tian H, Liang N, Wang Y, Liang C, Li X, Sun F (2009) Altered microRNA expression in patients with non-obstructive azoospermia. *Reprod Biol Endocrinol* 7:13

- Luddi A, Margollicci M, Gambera L, Serafini F, Cioni M, De Leo V, Balestri P, Piomboni P (2009) Spermatogenesis in a man with complete deletion of USP9Y. *N Engl J Med* 360:881–885
- Luetjens CM, Gromoll J, Engelhardt M, Von Eckardstein S, Bergmann M, Nieschlag E, Simoni M (2002) Manifestation of Y-chromosomal deletions in the human testis: a morphometrical and immunohistochemical evaluation. *Hum Reprod* 17:2258–2266
- Okada H, Tajima A, Shichiri K, Tanaka A, Tanaka K, Inoue I (2008) Genome-wide expression of azoospermia testes demonstrates a specific profile and implicates ART3 in genetic susceptibility. *PLoS Genet* 4:e26
- Quintana-Murci L, Krausz C, Heyer E, Gromoll J, Seifer I, Barton DE, Barrett T, Skakkebaek NE, Rajpert-De Meyts E, Mitchell M, Lee AC, Jobling MA, McElreavey K (2001) The relationship between Y chromosome DNA haplotypes and Y chromosome deletions leading to male infertility. *Hum Genet* 108:55–58
- Reynolds N, Collier B, Maratou K, Bingham V, Speed RM, Taggart M, Semple CA, Gray NK, Cooke HJ (2005) Dazl binds in vivo to specific transcripts and can regulate the pre-meiotic translation of Mvh in germ cells. *Hum Mol Genet* 14:3899–3909
- Ritz C, Spiess AN (2008) qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 24:1549–1551
- Sato Y, Yoshida K, Shinka T, Nozawa S, Nakahori Y, Iwamoto T (2006) Altered expression pattern of heat shock transcription factor, Y chromosome (HSFY) may be related to altered differentiation of spermatogenic cells in testes with deteriorated spermatogenesis. *Fertil Steril* 86:612–618
- Shinka T, Sato Y, Chen G, Naroda T, Kinoshita K, Unemi Y, Tsuji K, Toida K, Iwamoto T, Nakahori Y (2004) Molecular characterization of heat shock-like factor encoded on the human Y chromosome, and implications for male infertility. *Biol Reprod* 71:297–306
- Sigg C, Hedinger C (1981) Quantitative and ultrastructural study of germinal epithelium in testicular biopsies with “mixed atrophy”. *Andrologia* 13:412–424
- Silber SJ, Alagappan R, Brown LG, Page DC (1998) Y chromosome deletions in azoospermic and severely oligozoospermic men undergoing intracytoplasmic sperm injection after testicular sperm extraction. *Hum Reprod* 13:3332–3337
- Simoni M, Bakker E, Krausz C (2004) EAA/EMQN best practice guidelines for molecular diagnosis of y-chromosomal microdeletions. State of the art 2004. *Int J Androl* 27:240–249
- Simoni M, Tuttmann F, Gromoll J, Nieschlag E (2008) Clinical consequences of microdeletions of the Y chromosome: the extended Munster experience. *Reprod Biomed Online* 16:289–303
- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, Rohlfing T, Scott K, Schultz B, Strong C, Tin-Wollam A, Yang SP, Waterston RH, Wilson RK, Rozen S, Page DC (2003) The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423:825–837
- Spiess AN, Feig C, Schulze W, Chalmel F, Cappallo-Obermann H, Primig M, Kirchhoff C (2007) Cross-platform gene expression signature of human spermatogenic failure reveals inflammatory-like response. *Hum Reprod* 22:2936–2946
- Stouffs K, Lissens W, Verheyen G, Van Landuyt L, Goossens A, Tournaye H, Van Steirteghem A, Liebaers I (2004) Expression pattern of the Y-linked PRY gene suggests a function in apoptosis but not in spermatogenesis. *Mol Hum Reprod* 10:15–21
- Sun C, Skaletsky H, Birren B, Devon K, Tang Z, Silber S, Oates R, Page DC (1999) An azoospermic man with a de novo point mutation in the Y-chromosomal gene USP9Y. *Nat Genet* 23:429–432
- Tessari A, Salata E, Ferlin A, Bartoloni L, Slongo ML, Foresta C (2004) Characterization of HSFY, a novel AZFb gene on the Y chromosome with a possible role in human spermatogenesis. *Mol Hum Reprod* 10:253–258
- Theurillat JP, Zurrer-Hardi U, Varga Z, Barghorn A, Saller E, Frei C, Storz M, Behnke S, Seifert B, Fehr M, Fink D, Rageth C, Linsenmeier C, Pestalozzi B, Chen YT, Knuth A, Jager D, Moch H (2008) Distinct expression patterns of the immunogenic differentiation antigen NY-BR-1 in normal breast, testis and their malignant counterparts. *Int J Cancer* 122:1585–1591
- Vogt PH (2005) AZF deletions and Y chromosomal haplogroups: history and update based on sequence. *Hum Reprod Update* 11:319–336
- Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, Kiesewetter F, Kohn FM, Schill WB, Farah S, Ramos C, Hartmann M, Hartschuh W, Meschede D, Behre HM, Castel A, Nieschlag E, Weidner W, Grone HJ, Jung A, Engel W, Haidl G (1996) Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* 5:933–943
- Vogt PH, Falcao CL, Hanstein R, Zimmer J (2008) The AZF proteins. *Int J Androl* 31:383–394
- Wang Y, Barbacioru C, Hyland F, Xiao W, Hunkapiller KL, Blake J, Chan F, Gonzalez C, Zhang L, Samaha RR (2006) Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. *BMC Genomics* 7:59
- Wennemuth G, Blocher S, Schill WB, Aumuller G, Monsees TK (2003) Bradykinin increases intracellular calcium levels in rat testis peritubular cells via the B2 receptor subtype. *Br J Pharmacol* 138:351–358
- Zervakis M, Blazadonakis ME, Tsiliki G, Danilatu V, Tsiknakis M, Kafetzopoulos D (2009) Outcome prediction based on microarray analysis: a critical perspective on methods. *BMC Bioinformatics* 10:53
- Zou SW, Zhang JC, Zhang XD, Miao SY, Zong SD, Sheng Q, Wang LF (2003) Expression and localization of VCX/Y proteins and their possible involvement in regulation of ribosome assembly during spermatogenesis. *Cell Res* 13:171–177