

Screening for biomarkers of spermatogonia within the human testis: a whole genome approach

Kathrein von Kopylow¹, Christiane Kirchhoff¹, Davor Jezek², Wolfgang Schulze¹, Caroline Feig¹, Michael Primig³, Volker Steinkraus⁴, and Andrej-Nikolai Spiess^{1,*}

¹Department of Andrology, University Hospital Hamburg-Eppendorf, Hamburg, Germany ²University of Zagreb, School of Medicine, Zagreb, Croatia ³Inserm, U625, GERHM; IFR-140; Université de Rennes 1, Rennes, France ⁴Dermatologikum Hamburg, Hamburg, Germany

*Correspondence address. E-mail: a.spiess@uke.de

Submitted on November 26, 2009; resubmitted on January 27, 2010; accepted on January 29, 2010

BACKGROUND: A key step in studying the biology of spermatogonia is to determine their global gene expression profile. However, disassociation of these cells from the testis may alter their profile to a considerable degree. To characterize the molecular phenotype of human spermatogonia, including spermatogonial stem cells (SSCs), within their cognate microenvironment, a rare subtype of human defective spermatogenesis was exploited in which spermatogonia were the only germ cell type.

METHODS: The global expression profile of these samples was assessed on the Affymetrix microarray platform and compared with tissues showing homogeneous Sertoli-cell-only appearance; selected genes were validated by quantitative real-time PCR and immunohistochemistry on disparate sample sets.

RESULTS: Highly significant differences in gene expression levels correlated with the appearance of spermatogonia, including 239 best candidates of human spermatogonally expressed genes. Specifically, fibroblast growth factor receptor 3 (FGFR3), desmoglein 2 (DSG2), E3 ubiquitin ligase c-CBL (casitas B-cell lymphoma), cancer/testis antigen NY-ESO-1 (CTAG1A/B), undifferentiated embryonic cell transcription factor 1 (UTF1) and synaptosomal-associated protein, 91 kDa homolog (SNAP91) were shown to represent specific biomarkers of human spermatogonia.

CONCLUSIONS: These biomarkers, specifically the surface markers FGFR3 and DSG2, may facilitate the isolation and enrichment of human stem and/or progenitor spermatogonia and thus lay a foundation for studies of long-term maintenance of human SSCs/progenitor cells, spermatogonial self-renewal, clonal expansion and differentiation.

Key words: spermatogonia / gene expression / FGFR3 / desmoglein 2 / UCHL1

Introduction

In the testis, germ line stem cells can be identified as a subpopulation of 'undifferentiated spermatogonia' that reside in the basal layer of the seminiferous epithelium. Recent investigations described the isolation of a subpopulation of human testicular cells with stem cell-like features (Conrad *et al.*, 2008; Kossack *et al.*, 2009; Golestaneh *et al.*, 2009), although definite evidence in respect to the underlying cell type remains to be defined. If sufficient amounts of such cells could be collected and/or amplified from a small amount of human testis tissue, infertility and fatal diseases may be cured in male patients with a biopsy of their own testis in the future. However, the relatively low number of stem cells estimated to be present in the adult testis

brings into question their potential for clinical application. Moreover, culture conditions and growth factors essential for self-renewal, proliferation and differentiation of spermatogonial stem cells (SSCs) have largely been studied in animal model systems (Guan *et al.*, 2006; Seandel *et al.*, 2007), although little information is as yet available concerning factors that may affect fate determination of human spermatogonia.

An additional concern is the ability of human germline-derived stem cells to generate teratomas after transplantation into immunodeficient mice (Conrad *et al.*, 2008; Golestaneh *et al.*, 2009). Thus, if cell-based autologous therapies are to be developed, the extent to which the adult testis-derived stem cells *in vitro* resemble normal human SSCs at the molecular level is a critical issue. As a foundation,

comprehensive analysis of the molecular phenotype of human spermatogonia is required, including factors necessary for the isolation and long-term maintenance of human SSCs/progenitor cells. An analysis of the global gene expression pattern of human spermatogonia-derived cell suspensions in comparison with human embryonic stem cells and human adult germ-line stem cells was recently published (Conrad *et al.*, 2008), reporting differences in biomarker expression depending on the source of the cells and on their cellular microenvironment. Thus, disassociation of SSCs/progenitor cells from their niche and/or cell culture for days prior to analysis may have altered their gene expression pattern compared with spermatogonia in the *in vivo* situation.

To complement earlier results obtained from isolated SSCs/progenitor cells, we used an alternative strategy by studying the expression profile of these cells within their cognate microenvironment. We recently reported stage-specific global gene expression profiles of human male germ cells within their cognate microenvironments (Feig *et al.*, 2007; Spiess *et al.*, 2007). The selection of defined subtypes of defective spermatogenesis, combined with the analysis of individual biopsy samples, was a crucial step in this approach. In the present study, we employed microarray analysis to reveal the global expression patterns of adult human testes in which spermatogonia (SPG) were the only germ cell type. The global expression pattern in this rare subtype of defective spermatogenesis was compared with the more frequently observed condition of male infertility termed 'Sertoli-cell-only appearance' (SCO). Significant differences in gene expression levels between these two states correlated with the appearance of spermatogonia. This approach allowed us to dissect the gene expression pattern of human spermatogonia within their 'niche', at the same time avoiding potential influences on the gene expression signature that might arise from removing spermatogonia from their biological context.

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 October 2005. Informed consent and Ethic Committee Approval was obtained (OB/X/2000), and the study conducted in accordance with the guide lines of the 'Helsinki Declaration'. Tissues used for the microarray analysis were taken simultaneously for therapeutic testicular sperm extraction (TESE) and diagnostic purposes as described previously (Jezek *et al.*, 1998; Schulze *et al.*, 1999; Feig *et al.*, 2007). Five samples with highly homogeneous SCO characteristic and three samples with spermatogonia as the single germ cell type (SPG) were selected from a cohort of 700 biopsies. Their 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 test-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). Fetal tissues were obtained from abortion specimens collected for diagnostic purposes and deposited in the tissue bank of the University of Zagreb Medical School. The tissues included were fetal testis from Week 19 ($n = 2$), Week 22 ($n = 1$), Week 24 ($n = 1$) and Week 28 ($n = 1$). The study was approved by the ethical committee of the University of Zagreb Medical School (No.

04-1130-2006). The true fetal age was deduced from the gestational age (e.g. the time after the first day of the last menstrual period), supplemented with physiological data (e.g. fetal weight, fetal length, organ weight, etc.).

Microarray analysis of pathological subtypes

Tissue samples of rice grain size were removed and immediately submerged in RNAlater (Ambion, Austin, TX, USA). Total RNA was extracted in RNeasy columns (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 (Model 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 biotin labeling of cRNA was performed following the Eukaryotic Sample and Array Processing Manual (Affymetrix) essentially as described in Feig *et al.* (2007). The cRNA was transferred to Human Genome U133 Plus2.0 Array GeneChips (Affymetrix) and incubated for 16 h at 60 U/min at 45°C on a rotator (Affymetrix Hybridization Oven 640). Arrays were washed and stained by using a streptavidin-phycoerythrin conjugate. To increase signal strength, the antibody amplification protocol was used (EukGE-WS2v4; Eukaryotic Sample and Array Processing Manual). GeneChips were processed with a HP GeneArray Scanner (Affymetrix) by using default settings. DAT image files of the microarrays were generated using Microarray Analysis Suite 5.0 (MAS; Affymetrix).

Data analysis and statistics

Genes were filtered that had significant expression values across all samples and average expression higher than the fifth percentile of all experiments. The experiments were normalized by quantile normalization (*R* package *affy*, www.bioconductor.org). In compliance with the MIAME guide lines (Brazma *et al.*, 2001), raw and processed data files of all specimens were deposited in the Gene Expression Omnibus (GSE18997).

The unfiltered normalized data were subjected to bootstrap hierarchical clustering analysis (Manhattan distance, average metric) in order to reveal grouping structures and their stability based on the complete gene expression profile. Additionally, the same data were interrogated by principle component analysis and 3-dimensional display of the first three components to describe the main variances in the data. All clustering was conducted with the freely available TIGR MeV v4.4 software (<http://www.tm4.org/mev.html>).

To filter genes that pertain to differential expression between the SCO and SPG subtypes, we employed a Bayesian *t*-test that accounts for increasing variance in the low intensity region of the data (Smyth, 2004) in conjunction with a Benjamini-Hochberg multiple testing correction (Benjamini and Hochberg, 1995). The complete analysis above was done within the freely available *R* statistical environment (www.r-project.org) using the *limma* extension package. 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/>). Protein interaction networks of the upregulated differential genes were built using the Cytoscape software and the MiMI plugin (Gao *et al.*, 2009), which interrogates several different databases in respect to protein-protein interactions.

Quantitative real-time PCR

Reverse transcription from total RNA was carried out either with SuperscriptTMII RNaseH—Reverse Transcriptase or with 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 0.5 μM in a total volume of 20 μl. Primer sequences and amplicons sizes are available in Supplementary data, 1A. Cycling conditions were: 40 to 60 cycles of 10 s at 95°C, 20 s at 60°C and 30 s at 72°C. Special cycling conditions were used for the undifferentiated embryonic cell transcription factor 1 (UTF1) transcript with denaturation for 10 s at 95°C, annealing for 10 s at 64°C and elongation for 30 s at 80°C. Due to the GC-rich template, this reaction was furthermore resolved with the addition of 1.7 M Betaine. Transcript levels were normalized to ribosomal protein S27, whose expression was found to be unaltered over the different samples. Fold differences were calculated by use of the five-parameter sigmoidal model (*qpcR* package, Spiess et al., 2008). PCR efficiency was calculated by an R script employing the window-of-linearity method (Ramakers et al., 2003). Products were analyzed on an Agilent Bioanalyzer (Model 2100; Agilent Technologies, Palo Alto, CA, USA) and their identity was verified by sequence analysis. We validated the most interesting candidates by quantitative real-time PCR on a completely independent test set of 12 samples (9 SCO, 3 SPG) taken from the same cohort of 700 samples to avoid the problems of biological redundancy and sample dependency. Methodological agreement between the results obtained from microarray analysis and qPCR was calculated by a Bland-Altman plot using the log-transformed ratios obtained from averaging all replicates (Bland and Altman, 1999).

Immunohistochemistry

Validation via immunohistochemistry (IHC) was carried out with tissue samples from normogonadotrophic patients with obstructive azoospermia. For each antibody, tissue sections of two different patients were tested. Sections were prepared from Modified Davidson's fluid-fixed, paraffin-embedded testicular biopsies (Latendresse et al., 2002). Antibodies (mouse monoclonal and rabbit polyclonal) and dilutions are listed in Supplementary data, 1C. Rabbit polyclonal PGP9.5 (Z5116, Dako, Hamburg, Germany) served as positive control and rabbit immunoglobulin fraction (X0936 Dako, Hamburg, Germany) as negative control. Antigen localization was achieved using a two-step immunoperoxidase staining method (Envision plus polymer System, Dako, Hamburg, Germany) and/or a three-step immunoperoxidase staining method (ABC Staining System sc-2017, Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA). Specificity of immunostaining was confirmed by both omission of primary antibody and staining of parallel sections with antibodies directed against an irrelevant antigen. Stained sections were evaluated by bright-field microscopy (Axiovert 100, Zeiss, Germany) and images captured with an Axiocam ICc3 digital camera (Zeiss, Germany).

Results

The histology of the testis samples included in the microarray analysis was extrapolated from examinations of parallel biopsies from the same testis as part of the routine diagnostic work-up. Samples included belonged to two distinct testicular histologies: the first subtype showed sporadic tubules containing spermatogonia on a background of tubuli with predominantly SCO appearance; the second subtype showed a homogeneous SCO appearance in which no germ cell-containing tubuli were observed (examples are shown in Supplementary data, 3).

We aimed to quantify the impact of gene expression changes in testicular tissue when spermatogonia were present within the seminiferous

epithelium. Gene expression data were interrogated without prefiltering, based on the premise that if there were substantial changes in the testicular transcriptome these would then be evident using the complete gene expression profiles. Eliminating genes near background and positive/negative controls resulted in expression data of 54 613 probe sets reflecting 20 515 genes. We clustered the data by bootstrap hierarchical clustering which not only explores the grouping structure but also the internal stability of all clusters. All samples clustered completely in their own groups (SCO and SPG) with a cluster stability of 100% for each sample (Fig. 1A). This was a first indication of major changes in testicular gene expression when spermatogonia were present. As another approach, the dataset was subjected to principle component analysis and 3-dimensional display of the first three principle components (Fig. 1B). Corroborating the findings above, the samples belonging to SCO and SPG subtypes could be well separated into distinct clusters along the first principle component, demonstrating that the major expression differences between the two groups were related to the two different pathological subtypes.

We then filtered the differential genes by using a Bayesian t-test. About 339 genes with a *P*-value <0.05 were selected after a Benjamini-Hochberg multiple testing correction (Supplementary data, 2). From these 339 candidate genes, 70% of genes were significantly upregulated more than 2-fold when spermatogonia were present, whereas only 6 genes (1.7%) were downregulated more than 2-fold (Fig. 2). Maximum fold-induction was observed for testis-specific protein Y-linked 1 (TSPY1, 82-fold), transketolase 1 (TKTL1, 63-fold) and clusters of several MAGE (variants B2, A3, A6, A12, C1, A4) and GAGE (variants 1, 7, 4, 2, 3) cancer-testis antigens (10–30-fold). On the basis of a recent compilation of 153 cancer/testis (CT) antigens (Caballero and Chen, 2009), our list of spermatogonia-expressed genes contained at least 66 of CT antigens which are normally expressed in germ line cells but also activated in a wide range of cancer types.

For a further functional over-representation analysis of the differential genes, we focused on the 239 genes that were upregulated more than 2-fold between SCO and SPG under the assumption that this group is dominated by genes that are specific to the appearance of spermatogonia and not by genes that alter their expression in other cell types such as Sertoli cells or those of the interstitial compartment. Functional over-representation analysis of the 239 genes revealed a high significance ($P < 1E-7$) of genes involved in cell cycle, mitosis and sexual reproduction (Table 1).

Finally, our approach was to uncover regulatory networks that coincide with the appearance of spermatogonia. Using the same set of genes as above, a network was built from the protein interactions of the differential genes and known interaction partners that were not included in our list (Supplementary data, 4). To condense this complex network to a more comprehensible subset, we additionally built a network omitting all other known interactions outside of our list (Fig. 3). Four regulatory subnetworks were identified that dominate within the complete network, namely cell cycle control, FGFR3 receptor signaling, RNA processing and BCL6 signaling.

For the validation of our microarray data by real-time PCR we selected 11 spermatogonia-associated genes using an independent cohort of 12 different testis samples (3 SPG, 9 SCO). Ratios of gene expression between SCO and SPG obtained from microarray data and real-time PCR data are given in Supplementary data, 1B.

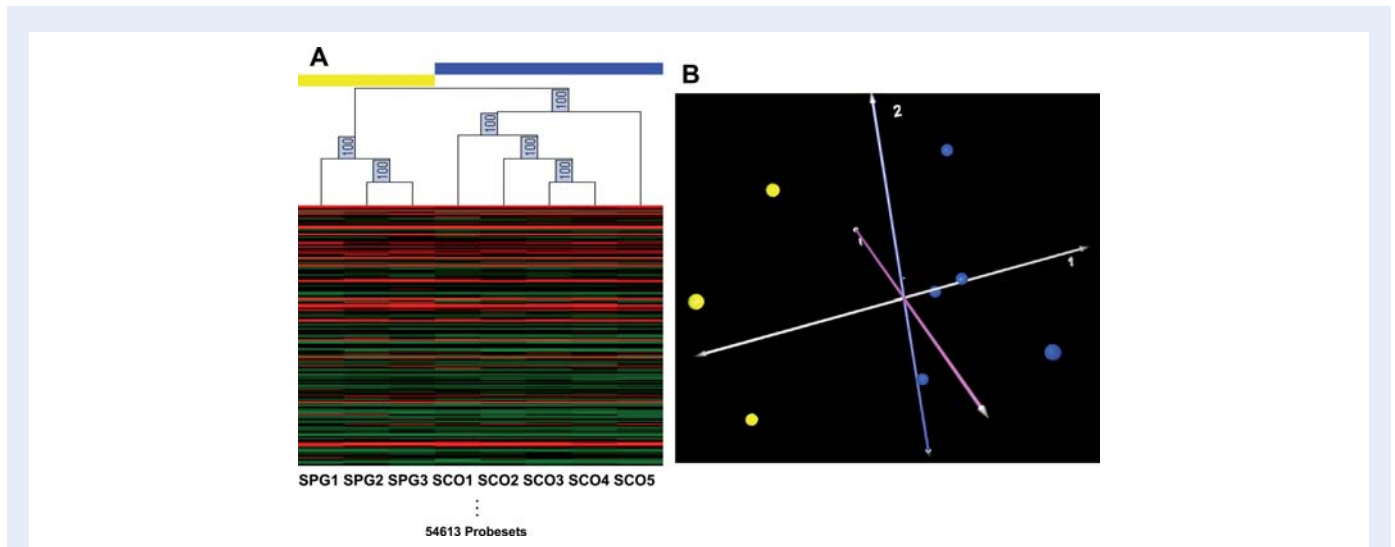


Figure 1 Classification of testicular biopsies with SCO syndrome (SCO1-5) and presence of spermatogonia (SPG1-3) by gene expression profiling. Five samples with SCO syndrome and three samples with spermatogonial presence within the tubular compartment were subjected to gene expression profiling with the Affymetrix Hu133Plus2 microarray system. Samples were clustered based on the expression values of all 54 613 probesets from the microarray without pre-selection of differential genes by (A) bootstrap hierarchical clustering (Manhattan distance, average metric) which calculates the cluster stability of the samples (given as percent values on the leaf nodes of the dendrogram) and (B) principle component analysis and 3-dimensional display of the first three components. Color coding of the samples is blue for SCO and yellow for SPG.

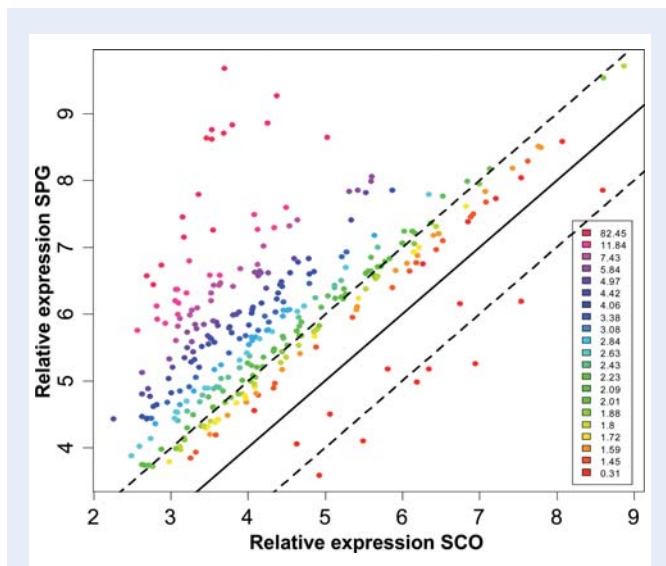


Figure 2 Ratio plot displaying dominant up-regulation of genes in the presence of spermatogonia (SPG) compared with SCO syndrome.

The plot displays the 339 most differential genes, with the majority of genes being upregulated (97.5%). The straight black line pertains to 'no differential expression' although the dashed lines reflect the 2-fold (0.5-fold) ratio borders. Color coding of the spots correlates with the magnitude of differential expression, as detailed in the color legend on the right. Ratios were calculated using the mean of all replicates.

We employed a Bland–Altman plot (Bland and Altman, 1999) to describe the concordance of the ratios obtained from microarrays to those obtained with quantitative real-time PCR. The overall agreement was very high, resulting in all validated genes to be within a narrow 95% limits of agreement (Fig. 4). It was not possible to

Table 1 Functional over-representation analysis using the upregulated genes between SCO and SPG

Function	Number of genes	P-value
M-Phase	20	6.2E−9
Spermatogenesis	18	4.2E−9
Mitosis	17	5.1E−8
Cell cycle checkpoint	6	6.9E−2
Dorso-ventral axis formation	4	2.6E−1
p53-signaling	4	7.1E−1
Ubiquitin-mediated hydrolysis	4	9.8E−1

SCO, Sertoli-cell only; SPG, spermatogonia.

calculate ratios from the real-time PCR outcome for two other validated genes (UTF1 and ESX1) because their expression in patients with SCO was below the detection level.

We wanted to know whether the increased transcript levels were reflected on the protein level. In most somatic tissues transcription and translation are closely linked, making detection of an mRNA a strong indicator of its pattern of protein synthesis. However, male germ cells rely on posttranscriptional regulatory mechanisms on various stages of development. We first searched the Human Protein Atlas portal (Ponten *et al.*, 2008; <http://www.proteinatlas.org>) to obtain protein expression data on the 239 best candidate genes with fold inductions >2 (Supplementary data, 2). Immunohistochemical data were available for 122 genes from this gene list (Supplementary data, 2), representing a coverage of nearly 30% of genes. On the basis of the testis tissues shown in the Atlas, the corresponding proteins seemed to be germ cell-specific in their vast majority; protein expression in spermatogonia was supported in more than 75% of cases. In

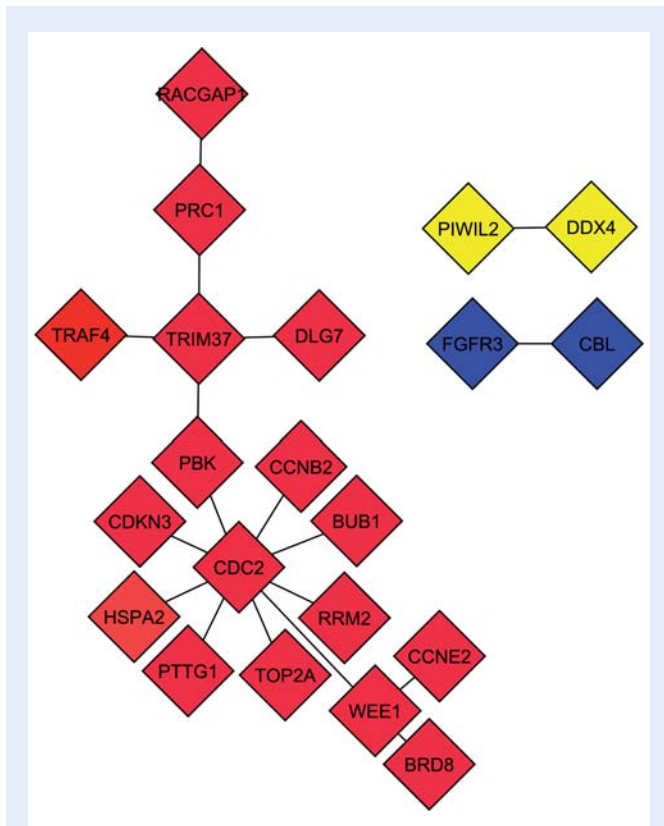


Figure 3 A spermatogonia-specific protein interaction network based on 239 upregulated genes condensing only the subset of genes that have an interaction within this group. Starting from the genes upregulated between SCO syndrome and spermatogonial presence, a network was built merging protein interaction data of several databases. Color coding reflects cell cycle (red), PIWIL2/DDX4 interaction (yellow) and FGFR3 signaling (blue). This is essentially a reduced version of the complete spermatogonial protein network found in Supplementary data, 4.

particular, FGFR3, DSG2, CBL, CTAG1A/B (NY-ESO-1) and SNAP91 proteins seemed to be specifically expressed in spermatogonia. In addition, we searched the Human Protein Atlas as a source for *in silico* data mining of additional proteins which might be mainly expressed or even restricted to spermatogonia. A keyword search revealed spermatogonia-selective protein expression also for DNAJB6, FMRI, MAP4K4, NAP1L4, PHF13 and UCHL1/PGP9.5. PGP 9.5, an ubiquitin C-terminal hydrolase (UCHL1), was already previously employed in various animal models as a marker to identify and isolate spermatogonia (Luo *et al.*, 2006; Rodriguez-Sosa *et al.*, 2006). UCHL1 was also found in the cytoplasm of human spermatogonia (He *et al.*, 2009) and was therefore selected here as an independent positive control (see below).

IHC was performed on tissue sections of normal adult human testis to further validate our microarray data and to confirm protein expression in normal spermatogonia. Although not a quantitative method, IHC is currently the only practical way of detecting proteins in small fixed tissue samples. We focused on genes exhibiting robust differences in expression levels and chose fibroblast growth factor receptor 3 (FGFR3), desmoglein 2 (DSG2), E3 ubiquitin ligase c-CBL (casitas B-cell lymphoma), cancer/testis antigen NY-ESO-1 (CTAG1A/B), undifferentiated embryonic cell transcription factor 1 (UTF1) and synaptosomal-associated protein, 91 kDa homolog (SNAP91) from our list. Employing the PGP9.5 antiserum as a positive control, a particularly intense immunostaining was observed in germ cells attached to the basement membrane of the seminiferous tubules, taken to represent A_{dark} and A_{pale} spermatogonia (Fig. 5), although some staining of Leydig cells and nerve fibers was also seen. Employing antibodies against FGFR3, DSG2, UTF1 and CBL seemed to be restricted to spermatogonia, although CTAG1A/B (NY-ESO-1) and SNAP91 were occasionally also seen in primary spermatocytes, but not in post-meiotic cells or in testicular somatic cells (Fig. 5). Regarding the subcellular distribution,

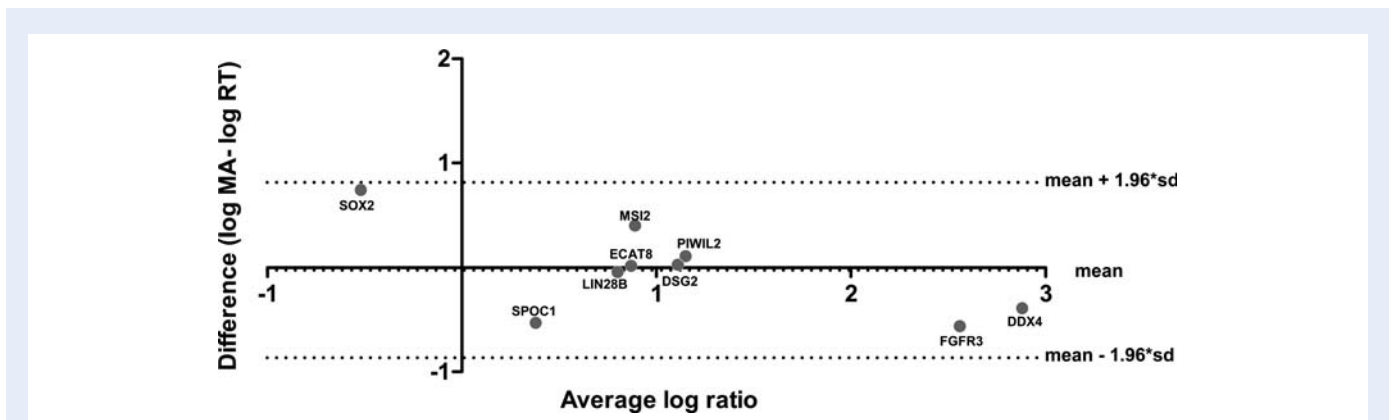


Figure 4 Bland–Altman plot for comparison of gene expression ratios obtained from microarray analysis (MA) and real-time PCR (RT). The plot displays the log ratios (abscissa) in relation to the differences in the logs between the two methods (ordinate). Grey dots represent the log transformed data for each validated gene. All ratios are found within the narrow 95% limits of agreement (dashed lines), estimated by the mean difference ± 1.96 standard deviation of the differences. All ratios were calculated from the mean of the replicates.

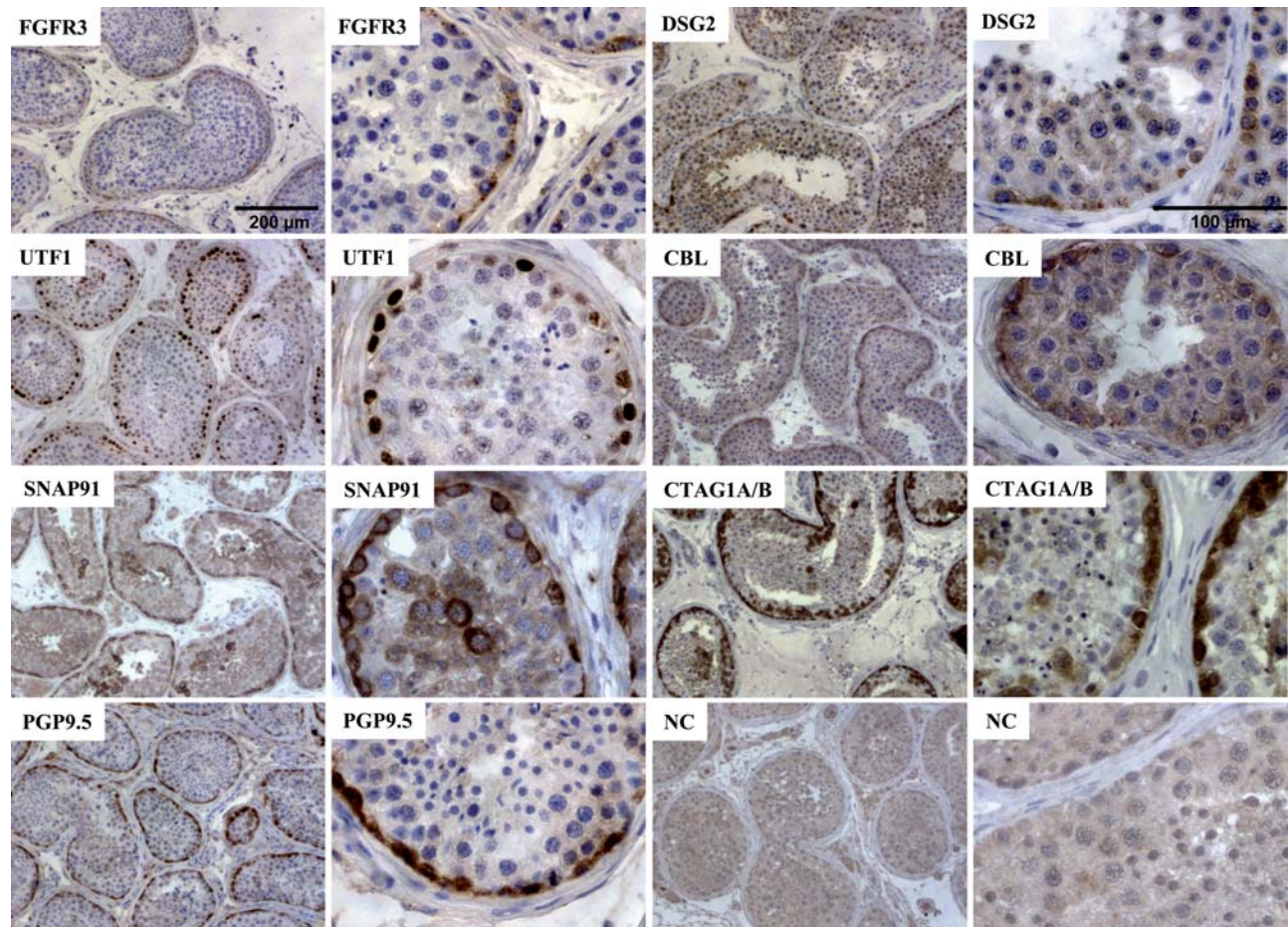


Figure 5 Immunohistochemical localization of proteins in normal human testis shows specific or enriched expression in spermatogonia for FGFR3, DSG2, UTF1, CBL, SNAP91, CTAG1A/B (NY-ESO-1) and UCHL1/PGP9.5 (positive control) as compared with normal serum (NC). Sections were counterstained with Mayers hemalaun solution. Magnification of 100-fold (left) and 400-fold (right) is shown. Scale bars represent 100 and 200 μm , respectively.

immunoreactivity of FGFR3, DSG2, CBL, CTAG1A/B (NY-ESO-1) and SNAP91 was seen in the cytoplasm and/or outer membrane of spermatogonia, although the staining pattern for UTF1 was nuclear. No specific staining of spermatogonia was observed when the primary antibodies were replaced by normal rabbit serum (Fig. 5).

To establish the potential expression of FGFR3, DSG2 and UTF1 in spermatogonial precursors we checked whether some of these markers were present in germ cells during human fetal testis development. Testes were taken from five male fetuses, in which the gonadal appearance was unremarkable, and endocrine organs involved in reproductive development (e.g. the pituitary, thyroid and adrenal glands) appeared healthy at the histological level. The gestational ages of these fetuses had been estimated to be 19 through 28 weeks, i.e. a time period after the germ cells have entered the mitotic arrest stage. In the examples shown in Fig. 6, testicular cords were lined with germ cells and Sertoli cells, and were surrounded by several layers of peritubular cells. Germ cells (pro-spermatogonia) were clearly distinct by their large and round nuclei and pale-stained cytoplasm. Frequently, one or more prominent nucleoli could be noted. The immunopositive reaction of FGFR3 and

DSG2 was associated with cells at the periphery in a majority of cords. Both, FGFR3 and DSG2 were immunolocalized to the cytoplasm and/or outer membrane of a subset of cells which, based on their morphology and association, were regarded as fetal germ cells. Thus, DSG2 protein expression was associated with the presence of germ cells and was observed long before the formation of the blood-testis-barrier. Neither the somatic cells in the seminiferous tubules (Sertoli and myoid cells) nor the interstitial Leydig cells reacted with the antibodies. Different from the continued expression of FGFR3 and DSG2 proteins, the same specimens were negative for UTF1 (data not shown; compare *Kristensen et al., 2008*). Also, no staining was observed when the primary antibodies were replaced by normal rabbit serum (data not shown).

Discussion

The strength of the current study is that spermatogonia, including stem cells, were analyzed directly in their cognate microenvironment in the human testis. The stem cell 'niche', or specialized microenvironment, is now recognized as one of the major contributors to control the

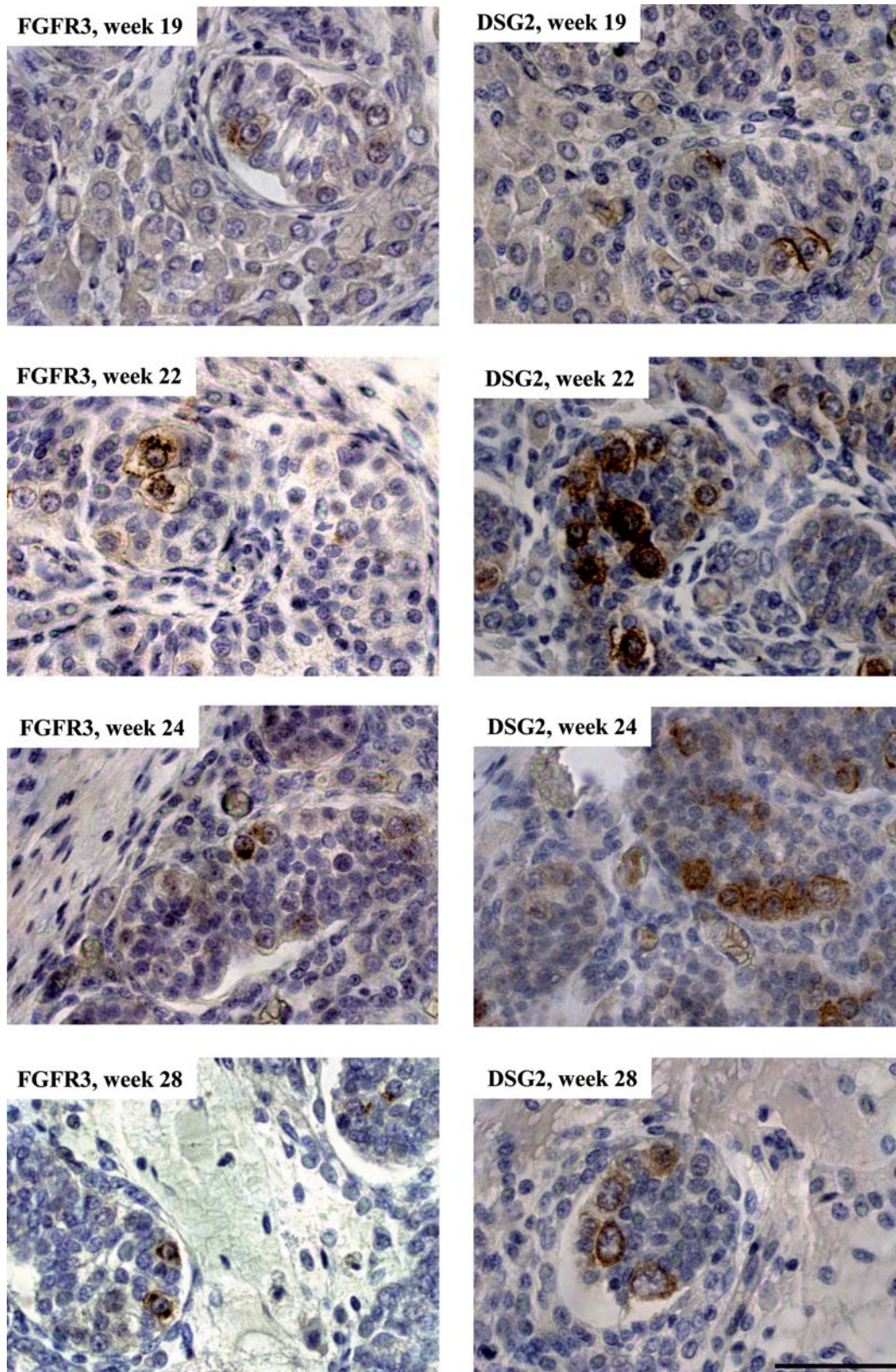


Figure 6 IHC of FGFR3 (left) and DSG2 (right) in fetal testis tissue at 19, 22, 24 and 28 weeks of gestation shows specific expression in spermatogonial precursors. Sections were counterstained with Mayer's hemalaun solution. Magnification of 400-fold is shown. Scale bar represents 50 μm .

balance of spermatogonial differentiation and self-renewal *in vitro* and *in vivo* (for review, see Hess *et al.*, 2006). It is important to consider that the SSC niche is composed of Sertoli cells and basement extracellular matrix, and possibly peritubular myoid and interstitial cellular components around the seminiferous tubules. Despite spermatogonia comprising only a minor proportion of the total testicular cell composition, microarray analysis and clustering revealed a significant change of the testicular transcriptome in the presence of spermatogonia, indicating that this germ cell types' specific function has a complex and specific transcriptional correlate. Prominent functions that pertain to spermatogonia are cell cycle, FGFR3 signaling, RNA processing and BCL6 signaling. Indeed, activating mutations in FGFR3 have been suggested to promote clonal expansion of spermatogonia in congenital disorders and testicular tumors (Juul *et al.*, 2007; Goriely *et al.*, 2009). Identification of a regulatory network for FGFR3 within spermatogonia and its specific protein expression tallies with the observations of bFGF being an important factor for successful spermatogonial culture (Conrad *et al.*, 2008; Golestaneh *et al.*, 2009) and of a spermatogenesis-preserving gain-of-function mutation in FGFR3 in a non-mosaic Klinefelter patient (Juul *et al.*, 2007). Two genes that play a prominent role in the regulation of FGF receptor signaling are upregulated in biopsies containing spermatogonia, therefore suggesting a vital role for a FGFR3 regulatory network in these cells. This applies to the Cas-Br-M (murine) ecotropic retroviral transforming sequence (CBL), whose mouse homologue was shown to be a negative regulator of FGFR signaling by forming a ternary complex with FRS2 and GRB2 (Wong *et al.*, 2002). Specific staining of spermatogonia with the CBL antibody as part of the immunohistochemical validation indicates that the human CBL protein could have a similar function in spermatogonia. A further putative member of the FGFR3 regulatory network is the synaptosomal-associated protein SNAP91, which was likewise expressed in spermatogonia and is known to be part of a heterotetrameric protein complex responsible for protein sorting (Baust *et al.*, 2008).

The importance of ubiquitin-mediated protein degradation in human spermatogonia is corroborated by our results of spermatogonia-specific expression of UCHL1/PGP9.5. In adult testis, UCHL1 cytoplasmic expression was predominantly found by IHC in Type A spermatogonia. The same localization was observed in boar, and its asymmetric distribution implicated with maintenance versus differentiation of SSCs (Luo *et al.*, 2009). Hence, this pattern suggests a role of PGP9.5 in human spermatogonia as a prelude for differentiation.

Of our differential list, 47 genes are known to interact with BCL6, which functions as a stabilizer for spermatocytes by protecting these against stressors (Kojima *et al.*, 2001). Transcript levels of BCL6 itself do not seem to be significantly elevated, as it is not included in our list of differentially expressed genes. However, BCL6B, another member of the same gene family, holds a pivotal role in the maintenance of SSCs *in vitro* and support of active spermatogenesis (Oatley *et al.*, 2006). The large number of differentially expressed and potentially spermatogonia-specific genes within our list that have known interactions with BCL6 supports a substantial weight within the network of spermatogonial regulation.

Although adult germ-line stem cells obtained *in vitro* are most likely derived from spermatogonia (Conrad *et al.*, 2008; Golestaneh *et al.*, 2009; Kossack *et al.*, 2009), their precise origin remains to be established. Culture conditions that are known to support the long-term

proliferation of SSCs can be used to generate ES-like colonies from adult human testis (Mizrak *et al.*, 2009). Moreover, it is also possible to obtain pluripotent cells from the human testis that are presumably not of the germ cell lineage (Gonzalez *et al.*, 2009). To ensure that no other germ cell types are the source of ES-like colonies, precise protocols for the isolation of SSCs are a major prerequisite. Unfortunately, no unique human or mouse surface markers are currently available to address this question. A selection of suitable surface markers assisting in the enrichment/purification of a spermatogonial fraction would therefore be an advance in the characterization of spermatogonial subpopulations. In addition to FGFR3, we identified DSG2, a calcium-binding transmembrane glycoprotein, as being localized to the spermatogonial membrane, although with some additional cytoplasmic staining. In the adult testis, it is thought to be part of the blood-testis-barrier (Cheng and Mruk, 2009). During tumorigenesis, loss of DSG2 protein expression contributes to an aggressive phenotype (Brennan *et al.*, 2007). Previous observations revealed that important DSG2 functions are desmosome-independent and are needed for embryonal stem cell and early embryo survival (Eshkind *et al.*, 2002), suggesting roles of DSG2 that transcend desmosomal adhesive function.

One of the limitations of our experimental design is that it might also identify genes that are activated in the pathological testis and that are not necessarily expressed by normal stem/progenitor spermatogonia. Although the various pathologies modify the germ cell and somatic environment of the testis, such data can still be used to investigate the expression and function of candidate genes in normal testes (Orwig *et al.*, 2008). Interestingly, and supporting our approach, we found an overlap of 46% between our list of differential genes and gene expression data obtained from comparing prepubertal spermatogonia and testicular somatic cells isolated by micromanipulation (Wu *et al.*, 2009). By using highly homogeneous pathological subtypes we have established a list of over 300 differential genes whose expression correlates significantly with the presence of spermatogonia in the human testis. These data should be useful in defining spermatogonial function within the biological niche as well as enabling a comparison of *in vivo* and *in vitro* gene expression changes of spermatogonia.

Supplementary data

Supplementary data are available at <http://humrep.oxfordjournals.org/>.

Acknowledgements

The antibody against NY-ESO-1 was kindly provided by Dr G. Spagnoli. We would like to thank Viviana Kozina for fruitful discussions.

Funding

This work was supported by a grant of the German Research Association DFG to K.v.K. (contract number FOR 1041) and by Serono Deutschland, on behalf of Dr Wilma Bilger. FC was supported by the Biozentrum Basel.

References

- Baust T, Anitei M, Czupalla C, Parshyna I, Bourel L, Thiele C, Krause E, Hoflack B. Protein networks supporting AP-3 function in targeting lysosomal membrane proteins. *Mol Biol Cell* 2008;**19**:1942–1951.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J Roy Stat Soc B* 1995;**57**:289–300.
- Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;**8**:135–160.
- Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC et al. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat Genet* 2001;**29**:365–371.
- Brennan D, Hu Y, Joubes S, Choi YW, Whitaker-Menezes D, O'Brien T, Uitto J, Rodeck U, Mahoney MG. Suprabasal Dsg2 expression in transgenic mouse skin confers a hyperproliferative and apoptosis-resistant phenotype to keratinocytes. *J Cell Sci* 2007;**120**:758–771.
- Caballero OL, Chen YT. Cancer/testis (CT) antigens: Potential targets for immunotherapy. *Cancer Sci* 2009;**100**:2014–2021.
- Cheng CY, Mruk DD. An intracellular trafficking pathway in the seminiferous epithelium regulating spermatogenesis: a biochemical and molecular perspective. *Crit Rev Biochem Mol Biol* 2009;**44**:245–263.
- Conrad S, Renninger M, Hennenlotter J, Wiesner T, Just L, Bonin M, Aicher W, Buhning HJ, Mattheus U, Mack A et al. Generation of pluripotent stem cells from adult human testis. *Nature* 2008;**456**:344–349.
- Eshkind L, Tian Q, Schmidt A, Franke WW, Windoffer R, Leube RE. Loss of desmoglein 2 suggests essential functions for early embryonic development and proliferation of embryonic stem cells. *Eur J Cell Biol* 2002;**81**:592–598.
- Feig C, Kirchhoff C, Ivell R, Naether O, Schulze W, Spiess AN. A new paradigm for profiling testicular gene expression during normal and disturbed human spermatogenesis. *Mol Hum Reprod* 2007;**13**:33–43.
- Gao J, Ade AS, Tarcea VG, Weymouth TE, Mirel BR, Jagadish HV, States DJ. Integrating and annotating the interactome using the MiMI plugin for cytoscape. *Bioinformatics* 2009;**25**:137–138.
- Golestaneh N, Kokkinaki M, Pant D, Jiang J, Destefano D, Fernandez-Bueno C, Rone JD, Haddad BR, Gallicano GI, Dym M. Pluripotent Stem Cells Derived from Adult Human Testes. *Stem Cells Dev* 2009;**18**:1115–1126.
- Gonzalez R, Griparic L, Vargas V, Burgee K, Santacruz P, Anderson R, Schiewe M, Silva F, Patel A. A putative mesenchymal stem cells population isolated from adult human testes. *Biochem Biophys Res Commun* 2009;**385**:570–575.
- Goriely A, Hansen RM, Taylor IB, Olesen IA, Jacobsen GK, McGowan SJ, Pfeifer SP, McVean GA, Meyts ER, Wilkie AO. Activating mutations in FGFR3 and HRAS reveal a shared genetic origin for congenital disorders and testicular tumors. *Nat Genet* 2009;**41**:1247–1252.
- Guan K, Nayernia K, Maier LS, Wagner S, Dressel R, Lee JH, Nolte J, Wolf F, Li M, Engel W et al. Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* 2006;**440**:1199–1203.
- He Z, Kokkinaki M, Jiang J, Dobrinski I, Dym M. Isolation, characterization, and culture of human spermatogonia. *Biol Reprod* 2010;**82**:363–372.
- Hess RA, Cooke PS, Hofmann MC, Murphy KM. Mechanistic insights into the regulation of the spermatogonial stem cell niche. *Cell Cycle* 2006;**5**:1164–1170.
- Jezek D, Knuth UA, Schulze W. 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* 1998;**13**:1230–1234.
- Juul A, Aksglaede L, Lund AM, Duno M, Skakkebaek NE, Rajpert-De Meyts E. Preserved fertility in a non-mosaic Klinefelter patient with a mutation in the fibroblast growth factor receptor 3 gene: case report. *Hum Reprod* 2007;**22**:1907–1911.
- Kojima S, Hatano M, Okada S, Fukuda T, Toyama Y, Yuasa S, Ito H, Tokuhisa T. Testicular germ cell apoptosis in Bcl6-deficient mice. *Development* 2001;**128**:57–65.
- Kossack N, Meneses J, Shefi S, Nguyen HN, Chavez S, Nicholas C, Gromoll J, Turek PJ, Reijo-Pera RA. Isolation and characterization of pluripotent human spermatogonial stem cell-derived cells. *Stem Cells* 2009;**27**:138–149.
- Kristensen DM, Nielsen JE, Skakkebaek NE, Graem N, Jacobsen GK, Rajpert-De Meyts E, Leffers H. Presumed pluripotency markers UTF-1 and REX-1 are expressed in human adult testes and germ cell neoplasms. *Hum Reprod* 2008;**23**:775–782.
- Latendresse JR, Warbritton AR, Jonassen H, Creasy DM. Fixation of testes and eyes using a modified Davidson's fluid: comparison with Bouin's fluid and conventional Davidson's fluid. *Toxicol Pathol* 2002;**30**:524–533.
- Luo J, Megee S, Rathi R, Dobrinski I. Protein gene product 9.5 is a spermatogonia-specific marker in the pig testis: application to enrichment and culture of porcine spermatogonia. *Mol Reprod Dev* 2006;**73**:1531–1540.
- Luo J, Megee S, Dobrinski I. Asymmetric distribution of UCH-L1 in spermatogonia is associated with maintenance and differentiation of spermatogonial stem cells. *J Cell Physiol* 2009;**220**:460–468.
- Mizrak SC, Chikhovskaya JV, Sadri-Ardekani H, van Daalen S, Korver CM, Hovingh SE, Roepers-Gajadien HL, Raya A, Fluiter K, de Reijke TM et al. Embryonic stem cell-like cells derived from adult human testis. *Hum Reprod* 2010;**25**:158–167.
- Oatley JM, Avarbock MR, Telaranta AI, Fearon DT, Brinster RL. Identifying genes important for spermatogonial stem cell self-renewal and survival. *Proc Natl Acad Sci USA* 2006;**103**:9524–9529.
- Orwig KE, Ryu BY, Master SR, Phillips BT, Mack M, Avarbock MR, Chodosh L, Brinster RL. Genes involved in post-transcriptional regulation are overrepresented in stem/progenitor spermatogonia of cryptorchid mouse testes. *Stem Cells* 2008;**26**:927–938.
- Ponten F, Jirstrom K, Uhlen M. The Human Protein Atlas—a tool for pathology. *J Pathol* 2008;**216**:387–393.
- Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* 2003;**339**:62–66.
- Rodriguez-Sosa JR, Dobson H, Hahnel A. Isolation and transplantation of spermatogonia in sheep. *Theriogenology* 2006;**66**:2091–2103.
- Schulze W, Thoms F, Knuth UA. Testicular sperm extraction: comprehensive analysis with simultaneously performed histology in 1418 biopsies from 766 subfertile men. *Hum Reprod* 1999;**14**:82–96.
- Seandel M, James D, Shmelkov SV, Falcatori I, Kim J, Chavala S, Scherr DS, Zhang F, Torres R, Gale NW et al. Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* 2007;**449**:346–350.
- Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 2004;**3**: Article3.
- Spiess AN, Feig C, Schulze W, Chalmel F, Cappallo-Obermann H, Primig M, Kirchhoff C. Cross-platform gene expression signature of human spermatogenic failure reveals inflammatory-like response. *Hum Reprod* 2007;**22**:2936–2946.
- Spiess AN, Feig C, Ritz C. Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. *BMC Bioinformatics* 2008;**9**:221.
- Wong A, Lamothe B, Lee A, Schlessinger J, Lax I. FRS2 alpha attenuates FGF receptor signaling by Grb2-mediated recruitment of the ubiquitin ligase Cbl. *Proc Natl Acad Sci USA* 2002;**99**:6684–6689.
- Wu X, Schmidt JA, Avarbock MR, Tobias JW, Carlson CA, Kolon TF, Ginsberg JP, Brinster RL. Prepubertal human spermatogonia and mouse gonocytes share conserved gene expression of germline stem cell regulatory molecules. *Proc Natl Acad Sci USA* 2009;**106**:21672–21677.