

**ALTERED GENE EXPRESSION SIGNATURE OF EARLY STAGES
OF THE GERM LINE SUPPORTS THE PREMEIOTIC ORIGIN OF
HUMAN SPERMATOGENIC FAILURE**

Journal:	<i>Andrology</i>
Manuscript ID:	ANDR-2014-0045.R1
Manuscript Type:	Original Article
Date Submitted by the Author:	18-Mar-2014
Complete List of Authors:	Bonache, Sandra; IDIBELL, Human Molecular Genetics Group Algaba, Ferrán; Fundació Puigvert, Pathology Department Franco Miranda, Eladio; Hospital de Bellvitge, Urology Department Bassas, Lluís; Fundació Puigvert, Andrology Service LARRIBA, SARA; Fundacio IDIBELL, Human Molecular Genetics Group
Key Words:	gene expression, spermatogenesis, infertility, arrays, azoospermia

1
2
3 1
4
5 2 **ALTERED GENE EXPRESSION SIGNATURE OF EARLY STAGES OF THE**
6
7 3 **GERM LINE SUPPORTS THE PREMEIOTIC ORIGIN OF HUMAN**
8
9 4 **SPERMATOGENIC FAILURE**
10
11
12 5

13
14 6 **Running title:** Early germ line gene expression in impaired sperm production
15
16 7

17
18 8 **Authors:** Sandra Bonache^{1#}, Ferran Algaba², Eladio Franco³, Lluís Bassas⁴, Sara
19
20 9 Larriba^{1*}
21
22

23 10
24
25 11 ¹Human Molecular Genetics Group-IDIBELL, 08908 L'Hospitalet de Llobregat,
26
27 12 Barcelona, Spain

28
29 13 ²Pathology Department-Fundació Puigvert, Universitat Autònoma de Barcelona 08025
30
31 14 Barcelona, Spain

32
33 15 ³Urology Service-Hospital Universitari de Bellvitge, 08907 L'Hospitalet de Llobregat,
34
35 16 Barcelona, Spain

36
37 17 ⁴Laboratory of Seminology and Embryology, Andrology Service-Fundació Puigvert,
38
39 18 08025 Barcelona, Spain

40
41 19 # Current address: Oncogenetics Laboratory, University Hospital Vall d'Hebron, 08035
42
43 20 Barcelona, Spain
44
45

46
47 21
48
49 22
50
51 23 ***Correspondence address**

52 24 Tel. +34 93 260 74 25 ext. 73 38; Fax.+34 93 260 74 14; e-mail: slarriba@idibell.cat
53
54
55
56 25
57
58
59
60

1
2
3 26 **ABSTRACT (300 words)**
4

5 27 The molecular basis of spermatogenic failure (SpF) is still largely unknown.
6
7 28 Accumulating evidence suggests that a series of specific events such as meiosis, are
8
9 29 determined at the early stage of spermatogenesis. The present study aims to assess the
10
11 30 expression profile of pre-meiotic genes of infertile testicular biopsies that might help to
12
13 31 define the molecular phenotype associated with human deficiency of sperm production.
14
15 32 An accurate quantification of testicular mRNA levels of genes expressed in
16
17 33 spermatogonia was carried out by RT-qPCR in individuals showing SpF due to germ
18
19 34 cell maturation defects, Sertoli cell-only syndrome (SCO) or conserved spermatogenesis
20
21 35 (CS). Additionally, the gene expression profile of SpF was compared with that of
22
23 36 testicular tumour (GCT), which is considered to be a severe developmental disease of
24
25 37 germ cell differentiation. Protein expression from selected genes was evaluated by
26
27 38 immunohistochemistry.
28
29
30

31
32 39 Our results indicate that SpF is accompanied by differences in expression of
33
34 40 certain genes associated with spermatogonia in the absence of any apparent
35
36 41 morphological and/or numerical change of this specific cell type. In SpF testicular
37
38 42 samples, we observed down-regulation of genes involved in cell cycle (*CCNE1* and
39
40 43 *POLD1*), transcription and post-transcription regulation (*DAZL*, *RBM15* and *DICER1*),
41
42 44 protein degradation (*FBXO32* and *TM9SF2*) and homologous recombination in meiosis
43
44 45 (*MRE11A* and *RAD50*) which suggests that the expression of these genes is critical for a
45
46 46 proper germ cell development. Interestingly, a decrease in the *CCNE1*, *DAZL*, *RBM15*
47
48 47 and *STRA8* cellular transcript levels was also observed, suggesting that the gene
49
50 48 expression capacity of spermatogonia is altered in SpF contributing to an unsuccessful
51
52 49 sperm production. Altogether, these data point to the spermatogenic derangement being
53
54 50 already determined at, or arising in, the initial stages of the germ line.
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

51
52
53
54
55

Key words: gene expression, testis, early stages of germ line, spermatogenic failure, male infertility

For Peer Review

56 INTRODUCTION

57 Spermatogenesis is a highly orchestrated developmental process by which
58 spermatogonia develop into mature spermatozoa. During the course of spermatogenesis
59 the three major forms of cell cycle are represented: mitosis of primitive spermatogonia;
60 two rounds of meiosis, from primary spermatocytes to haploid round spermatids; and
61 differentiation including structural and nuclear changes to generate mature spermatids
62 and spermatozoa. These processes are unique in male germ cell differentiation and
63 depend on precise, developmental stage- and germ cell type- specific gene expression.
64 However, the regulatory network that confers specific germ line gene expression in
65 mammals is not properly understood, especially at the mitotic and meiotic stages.
66 Understanding the regulatory step is essential for determining the molecular
67 requirements for the progression of spermatogenesis, and thus for understanding male
68 infertility which is often based on lack of replication of spermatogonia or meiotic
69 blockade.

70 Accumulating evidence suggests that a series of specific events during
71 spermatogenesis, such as meiosis and morphological changes, are determined at the
72 early stage of spermatogenesis. Spermatogonia, and specifically the type B
73 spermatogonia, should be an important preparation stage for meiosis. These data are
74 supported by the description of activation or up-regulation of many genes during this
75 specific germ cell stage (Guo *et al.*, 2004) and the generation of recombinant mouse
76 models of spermatogonia-expressed genes exhibiting severe defects in meiosis (Wang *et al.*
77 *al.*, 2001). Many of these genes codify germ cell specific proteins involved in
78 transcriptional or post-transcriptional regulation of gene expression (Wang *et al.*, 2001).

79 Furthermore, cellular interactions between germ line and somatic components of
80 the testicular seminiferous tubule, where spermatogenesis takes place, are essential to

1
2
3 81 achieve germ cell development, and thus for maintaining male fertility. The relevance of
4
5 82 these cellular interactions is supported by physiological events (for review, see
6
7 83 (Mclachlan *et al.*, 2007).
8

9
10 84 Some studies have used microarray technology to characterize the transcriptional
11
12 85 profile in germ and somatic cells at different steps of testicular development (Chalmel *et*
13
14 86 *al.*, 2007; Diederichs *et al.*, 2005; Namekawa *et al.*, 2006; Pang *et al.*, 2003; Schlecht *et*
15
16 87 *al.*, 2004; Schultz *et al.*, 2003; Sha *et al.*, 2002; Shima *et al.*, 2004). We have used
17
18 88 information from cDNA microarrays and mouse models to focus mainly on genes
19
20 89 specifically or preferentially expressed in immature germ cells in mammals, to
21
22 90 determine whether early gene expression changes are associated with subsequent
23
24 91 spermatogenic disorders and thus, male infertility. Additionally, the gene expression
25
26 92 profile of germ cell maturation failure was compared with that seen in testicular tumour,
27
28 93 which is considered to be a developmental disease of germ cell differentiation, in order
29
30 94 to give additional clues about the functional pathways involved in spermatogenic
31
32 95 derangement.
33
34
35

36 96
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 97 **MATERIALS AND METHODS**

4
5 98 **Subjects of study**

6
7 99 Our study recruited nineteen patients (range 30-49 yr) infertile due to spermatogenic
8
9 100 failure (SpF) at different germ cell stages, with a phenotype consistent with non-
10
11 101 obstructive azoospermia or severe oligozoospermia (<5 million sperm per mL). In
12
13 102 addition, fourteen patients with Sertoli cell-only syndrome (SCO) phenotype or germ
14
15 103 cell aplasia, were studied as controls of somatic cells (negative controls), and seventeen
16
17 104 infertile patients (range 32-50 yr), who were diagnosed with obstructive azoospermia
18
19 105 (as a consequence of congenital absence of vas deferens or previous vasectomy) and
20
21 106 showed conserved spermatogenesis (CS) were studied as gene expression controls of a
22
23 107 complete spermatogenic process (positive controls). Patients were selected on the basis
24
25 108 of the testicular histological pattern of >20 tubules from the same testicular section;
26
27 109 only testicular samples with a homogeneous histological pattern were included in the
28
29 110 study.

30
31
32
33
34 111 An extra group of individuals (n=5) diagnosed with germ cell tumour (GCT)
35
36 112 (range 29-46 yr) were analyzed. Two GCT samples were histologically classified as
37
38 113 carcinoma *in situ* (CIS) whereas the remaining three samples were classified as non-CIS
39
40 114 or GCT of advanced stages: one as classic seminoma, one as embryonal carcinoma and
41
42 115 one as mixed germ cell tumour (80% embryonal carcinoma; 20% classic seminoma).

43
44
45 116 Infertile individuals were selected from men referred for couple infertility to the
46
47 117 Andrology Service of the Fundació Puigvert, whereas GCT samples were recruited
48
49 118 from the Andrology Service of the Fundació Puigvert and the Urology Service of the
50
51 119 Hospital Universitari de Bellvitge. The study was approved by the Institutional Review
52
53 120 Board of both Centres, and all the participants signed a written informed consent.
54
55
56
57
58
59
60

1
2
3 121 The clinical procedures for infertile patients included medical history, physical
4
5 122 examination, semen analyses [performed in accordance with World Health Organization
6
7 123 guidelines (World Health Organization, 1999)] and hormonal study. The routine genetic
8
9 124 study for all non-obstructive samples included karyotype and analysis of chromosome Y
10
11 125 microdeletions, the latter performed according to the European guidelines (Simoni *et*
12
13 126 *al.*, 1999; Simoni *et al.*, 2004). Men with a chromosomal aberration or a Y-chromosome
14
15 127 microdeletion were not included in the study.
16
17
18
19

20 21 129 **Testicular samples**

22
23 130 Testicular biopsies from infertile men were obtained when necessary to confirm the
24
25 131 clinical diagnosis and for sperm retrieval (TESE) and cryopreservation purposes. Each
26
27 132 specimen was divided into three aliquots, one piece (≈ 10 -20 mg) was fixed in Bouin's
28
29 133 solution and reserved for histological analysis, a second aliquot (≈ 100 -200 mg)
30
31 134 processed for sperm extraction and the third (≈ 10 mg) was immediately transferred to
32
33 135 liquid nitrogen and stored at -80°C until analysis for gene expression experiments.
34
35

36
37 136 Referring to GCT, testicular samples were obtained directly after orchidectomy
38
39 137 and macroscopic pathological evaluation. For gene expression studies, one tissue
40
41 138 fragment was taken from the tumour portion of the testis and was immediately frozen at
42
43 139 -80°C .
44

45
46
47 140

48 141 **Histological analysis**

49
50 142 An assessment of spermatogenic status and the severity of the alteration were performed
51
52 143 after hematoxylin-eosin staining of paraffin samples from infertile patients (5- μm
53
54 144 sections) by quantification of specific germ cells (spermatogonia, spermatocytes I, round
55
56 145 spermatids and elongated spermatids) and Sertoli cells. The average number per tubule
57
58
59
60

1
2
3 146 was calculated after analysis of at least 15-20 cross-sectioned tubules per testis. A
4
5 147 modified Johnsen score (JS) count (Schulze *et al.*, 1999) was calculated on the basis of
6
7 148 the number of different cell types per tubule and infertile samples were classified as CS,
8
9 149 SpF-HS (hypospermatogenesis), SpF-MA (meiotic arrest) and SCO (Table 1).
10

11 150

14 151 **RNA extraction and cDNA synthesis**

16 152 Total RNA was obtained from the testicular biopsy using Absolutely RNA Miniprep Kit
17
18 153 (Stratagene, La Jolla, CA), following the manufacturer's instructions. The quality of
19
20 154 RNA [28S/18S ratio and RNA Integrity Number (RIN)] was assessed using the Agilent
21
22 155 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). Testicular RNA from
23
24 156 the five groups of study (SCO, SpF-MA, SpF-HS, GCT and CS) showed similar quality
25
26 157 values: both 28S/18S ratio and RIN presented no significant differences among them (p
27
28 158 = 0.056 and $p = 0.072$, respectively). Mean 28S/18S ratio value between all samples
29
30 159 (mean±SD) was 1.17±0.18 and mean RIN value was 7.77±0.62.
31
32

33
34 160 Single-stranded cDNA was obtained by reverse transcription (RT) of 500 ng of
35
36 161 RNA, using random primers and the High Capacity cDNA Reverse Transcription Kit
37
38 162 (AB, Foster City, CA). The resulting cDNA solution was aliquoted and stored at -20°C
39
40 163 until use.
41
42

43 164

45 165 **Gene expression quantification and statistical evaluation**

46
47 166 Quantitative real-time PCR (qPCR) assays were performed by means of the application
48
49 167 of the PCR arrays on micro fluidic cards (MFC), using 384-well TaqMan® Low
50
51 168 Density Arrays (TLDA) on an Applied Biosystems 7900HT Fast Real-Time PCR
52
53 169 System (AB, Foster City, CA). The 48-gene format MFC (47 experimental assays and 1
54
55 170 TLDA amplification control, *18S*) allowed simultaneous measurement of 34 target
56
57
58
59
60

1
2
3 171 genes that were selected based on a preferential expression in spermatogonia among
4
5 172 germ cells (n=26) and/or in Sertoli cells (n=8) -information obtained from cDNA
6
7 173 microarrays and mouse models bibliography data- (Supplemental Table 1), three marker
8
9 174 genes of the presence and/or function of spermatid (*PRMI*), Leydig (*INSL3*) and myoid
10
11 175 cells (*SI00A6*) and 10 potential reference genes. Genes and the corresponding assays on
12
13 176 demand used for the setup of the TLDA are listed in Supplemental Table 2. Selected
14
15 177 target genes are involved in different functional pathways (Supplemental Table 1 and
16
17 178 Supplemental Table 2).

18
19
20
21 179 Patient and control samples were always analysed as paired samples in the same
22
23 180 analytical run in order to exclude between-run variations. Additionally, a calibrator
24
25 181 sample was included in all the plates to compare the change in expression of a nucleic
26
27 182 acid sequence against the expression in all samples in the same study. Real-time PCR
28
29 183 data (Ct values) were pre-processed and stored in SDS 2.2 software (AB, Foster City,
30
31 184 CA).

32
33
34 185 Expression stability of the candidate reference gene/s was calculated with the
35
36 186 GeNorm software (Vandesompele *et al.*, 2002), in order to select the most stable
37
38 187 reference genes and improve normalization of target genes. GeNorm software calculates
39
40 188 the gene expression stability value M of multiple candidate genes as the average pair-
41
42 189 wise variation of a particular gene compared with all other candidate reference genes.
43
44 190 Lower M values indicate genes with less expression variation among samples.
45
46 191 Therefore, target gene expression was calculated relative to the expression of *PGKI* and
47
48 192 *PGMI* reference genes for SpF and control samples, whereas *PGKI* and *PPIA*
49
50 193 combination was selected as the most appropriate for GCT and controls. They showed
51
52 194 no statistical differences in absolute expression levels between groups (Kruskal-Wallis
53
54
55
56
57
58
59
60

1
2
3 195 test) (Supplemental Figure 1) and low M-value (GeNorm software indicating stable
4
5 196 expression among samples.
6

7 197 Thus, raw data (Ct values) were normalised to the two reference genes and
8
9 198 relative quantification (RQ) values were calculated using the qBase program
10
11 199 (Hellemans *et al.*, 2007). and the $2^{-\Delta\Delta Ct}$ strategy. The Mann-Whitney U test was used to
12
13 200 evaluate differences in relative expression of target genes in each patient group or
14
15 201 subgroup compared with controls. Multiple test adjustment was applied by using
16
17 202 Bonferroni correction.
18
19

20
21 203 Pearson product-moment correlation coefficients were calculated to determine
22
23 204 the correlation between the expression ratios of the target genes and the various
24
25 205 histological parameters in patient groups and controls.
26

27 206 All statistical analyses were performed using the SPSS software version 12
28
29 207 (Lead Technologies, Chicago, USA)
30
31

32 208

33 34 209 **Immunohistochemistry**

35
36 210 Tissue sections were prepared from Bouin-fixed, paraffin-embedded fragments of
37
38 211 testicular biopsies. For this study, the following commercially available polyclonal
39
40 212 rabbit antibodies were used: DAZL, HPA019777, Sigma-Aldrich, Inc; CCNE1 (C-19):
41
42 213 sc-198, Santa Cruz Biotechnology, Inc; CDKN1C (C-20):sc-1040, Santa Cruz
43
44 214 Biotechnology, Inc; DLK1 (H-118): sc-25437, Santa Cruz Biotechnology, Inc.
45
46

47 215 Immunohistochemistry was performed using the Dako EnVision+ kit (DAKO,
48
49 216 Hamburg, Germany) in conjunction with the Dako autostainer, according to the
50
51 217 instructions provided by the manufacturer.
52

53
54 218 Endogenous peroxidase was quenched by incubation in 0.5% hydrogen
55
56 219 peroxide. Dilutions of primary antibodies were adjusted at 1:300 to 1:100 to optimize
57
58
59
60

1
2
3 220 the results. Secondary goat anti-rabbit antiserum was coupled to a labelled polymer-
4
5 221 HRP, and staining was done with DAB and haematoxylin-eosin. Incubation with non
6
7 222 immune serum was used as a negative control (data not shown). Stained sections were
8
9 223 evaluated in bright-field microscopy (Axioskop 40, Zeiss, Göttingen, Germany) and
10
11 224 images captured with a Nikon Coolpix 5400 digital camera. The immunoexpression of
12
13
14 225 proteins was determined on tissue samples from at least six different individuals
15
16 226 showing each of the spermatogenic phenotypes.

17
18 227

19
20
21 228
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1
2
3 229 **RESULTS**

4
5 230 **Quantitative determination of spermatogenic status in defective spermatogenesis**

6
7 231 Paraffin-embedded testicular specimens from infertile patients and controls were
8
9 232 available for histological quantification. Considering the heterogeneity of human
10
11 233 testicular pathologies, we needed a detailed definition of the spermatogenic status of
12
13 234 each sample in order to acquire high quality data related to germ cell specific
14
15 235 transcriptional changes. To this end, the average number of Sertoli cells and specific
16
17 236 germ cell types per tubule were determined in haematoxylin and eosin stained testicular
18
19 237 sections and additionally the JS value was calculated (Table 1).

20
21
22
23 238 Using this strategy we confirmed the diagnosis of SCO (JS score 1-2) and CS (JS
24
25 239 score 9-10) phenotypes. With respect to SpF patients, seven of them presented
26
27 240 maturation arrest at primary spermatocyte level (SpF-MA) (JS score 4-6) and twelve
28
29 241 presented hypospermatogenesis (SpF-HS) (JS score 7-8) (Table 1).

30
31
32 242 Round and elongated spermatids were absent in 4 out of 7 SpF-MA samples
33
34 243 whereas the other three had very low values: round spermatids ranged between 2-4 per
35
36 244 tubule and elongated spermatids ranged between 1-2 per tubule suggesting a 80-90% of
37
38 245 meiotic arrest. This result confirms that only histological phenotypes with a defined and
39
40 246 highly homogeneous pattern of individual tubules were included in the study.

41
42
43 247 The number of Sertoli cells showed a near two-fold increase in the group SCO
44
45 248 (21.03 ± 7.00) compared with the CS samples (13.69 ± 3.12). Interestingly, the number
46
47 249 of spermatogonia and Sertoli cells showed no significant differences ($p = 0.318$ and p
48
49 250 $=0.447$ respectively) among SpF-MA, SpF-HS and CS groups (Table 1).

50
51
52 251

53
54 252 **Gene expression related to impaired spermatogenesis**

55
56 253 *PRMI expression confirms the histological phenotype*
57
58
59
60

1
2
3 254 In order to confirm the histological quantification of samples, we first analyzed the
4
5 255 expression of *PRMI*, the marker gene for the presence of spermatids. As expected,
6
7 256 negligible transcript level values were found in SCO and 100% MA samples; very low
8
9 257 values were determined in incomplete MA samples and decreased levels in HS samples
10
11 258 when compared with CS controls, showing a fold decrease in expression of 2.24×10^4
12
13 259 (SCO), 42.59 (SpF-MA) and 4.45 (SpF-HS) ($p < 0.001$). The absence of quantifiable
14
15 260 *PRMI* expression values in GCT samples confirmed the absence of mature germ cells.
16
17

18 261 Our *PRMI* gene expression results consistently agreed with the histological
19
20 262 diagnosis of all samples included in the study.
21
22

23 263

24 264 *Expression behaviour of Leydig and myoid cells in spermatogenic disorders*

25 265 Referring to the expression of the marker genes for the presence and/or function of
26
27 266 Leydig (*INSL3*) and myoid cells (*S100A6*), no statistically significant differences were
28
29 267 found between SpF-HS and CS groups ($p=0.245$ and $p=0.059$ respectively) and between
30
31 268 SpF-MA and CS samples ($p=0.065$ and $p=0.075$ respectively), however these genes
32
33 269 were significantly over-expressed in SCO group when compared with CS samples
34
35 270 ($p=0.001$ and $p=0.000$ respectively), probably attributable to the absence of germ cells;
36
37 271 since total testis samples are analysed, the loss of germ cells enrich the relative
38
39 272 contribution of the remaining somatic cells in SCO phenotype.
40
41

42
43 273 GCT samples also showed no statistically significant differences in *INSL3* and
44
45 274 *S100A6* expression when compared with the CS control group ($p=0.401$ and $p=0.542$
46
47 275 respectively)
48
49

50 276

51
52 277 *Most of the spermatogonia-transcriptionally associated genes analyzed show a*
53
54 278 *preferential germ line expression*
55
56
57
58
59
60

1
2
3 279 Spermatogonia-associated genes included in the study have been previously described
4
5 280 as having a preferential gene expression in the early germ line stages among germ cells
6
7 281 (Supplemental Table 1). The assessment of expression of these genes in the SCO
8
9 282 phenotype would provide important new information about whether these genes can be
10
11 283 expressed also in somatic cells of the testis.

14 284 Selective or preferential germ cell expression was confirmed for 21 out of the 26
15
16 285 spermatogonia-related genes of our study. First, the negligible transcript level values
17
18 286 found for *CCNE1*, *DAZL*, *RBM15* and *STRA8* in complete SCO samples supported their
19
20 287 selective germ cell lineage gene expression. Furthermore, statistically significant
21
22 288 reduced transcript values in SCO compared with CS suggested a preferential germ cell
23
24 289 expression: *ATM*, *BARD1*, *CCND1*, *CCNF*, *DICER1*, *E2F3*, *FBXO32*, *c-KIT*, *MRE11A*,
25
26 290 *POLA1*, *POLD1*, *RAD50*, *SIRT6* and *TM9SF2* presented a very significant fold-change
27
28 291 decrease in SCO ranging from 1.46 to 6.10 ($p < 10^{-4} - 10^{-6}$) whereas *BAX*, *CDKNIC* and
29
30 292 *XPA* showed less marked decrease, with a fold-change ranging from 1.22 to 1.55
31
32 293 ($p < 0.05$) (Supplemental Table 3).

36 294 However, for the five remaining spermatogonia-associated genes studied, a
37
38 295 preferential expression in the germ cell lineage could not be demonstrated. *CDKN1A*
39
40 296 and *DLK1* were found significantly over-expressed in SCO samples when compared
41
42 297 with CS controls ($p = 0.003$ and $p < 0.002$ respectively), showing a fold-change increase
43
44 298 of 1.77 and 3.99 respectively. Also, *BMPRIA*, *SMAD3* and *VEGFA* were not
45
46 299 differentially expressed between SCO and CS samples ($p > 0.05$) (Supplemental Table
47
48 300 3). Taken together, these results indicate a substantial expression of these five genes in
49
50 301 testicular somatic cells.
51

52
53
54 302

55
56 303 *Differential gene expression in defective and conserved spermatogenesis*
57
58
59
60

1
2
3 304 We then looked for differences in target gene expression between SpF samples and CS
4
5 305 controls.

6
7 306 For the 21 spermatogonia-associated genes preferentially expressed in the germ
8
9 307 line, we identified 13 differentially expressed genes and grouped them into three gene
10
11 308 clusters by their expression behaviour (Figure 1). In cluster I, both SpF-MA and SpF-
12
13 309 HS phenotypes were associated with notable significant decreases in transcript levels of
14
15 310 *BARD1*, *CCNE1*, *DAZL*, *FBXO32*, *RBM15* and *TM9SF2* genes ($p<0.002$), the reduction
16
17 311 of expression being more pronounced in the SpF-MA phenotype. Cluster II contained
18
19 312 genes significantly decreased in the SpF-HS ($p<0.05$) but not in the SpF-MA phenotype
20
21 313 (*CCNF* and *E2F3*). Cluster III included all those genes differentially expressed in SpF-
22
23 314 MA but not in SpF-HS phenotype: *DICER1*, *MRE11A*, *POLD1* and *RAD50* expression
24
25 315 levels were decreased whereas *CCND1* was increased in the SpF-MA ($p<0.05$) when
26
27 316 compared with controls. *CDKN1C* presented a similar pattern of gene expression
28
29 317 deregulation to *CCND1*, although the difference in expression was not statistically
30
31 318 significant (CS: 0.51 ± 0.15 ; HS: 0.61 ± 0.54 $p=0.394$; MA: 1.04 ± 0.73 $p=0.172$)
32
33 319 (Supplemental Table 3).

34
35
36
37
38 320 Interestingly, significant positive correlation coefficients (Pearson's correlation
39
40 321 $r\geq 0.6$; $p<0.0001$) were found between the number of elongated spermatids and the
41
42 322 transcription levels of genes from cluster I (Supplemental Table 4.A), suggesting that
43
44 323 these changes in gene expression could be of physiological relevance.

45
46
47 324 In regard to the spermatogonia and somatic cell-expressed genes, *CDKN1A* was
48
49 325 significantly increased in the SpF-MA and SCO samples ($p<0.05$). *DLK1* was found
50
51 326 highly over-expressed in SpF samples, although the differences were only statistically
52
53 327 significant for the SpF-HS and SCO ($p=0.001$) and not the SpF-MA phenotype
54
55 328 (probably due to a high standard deviation value) when compared with controls (Figure
56
57
58
59
60

1
2
3 329 2; Supplemental Table 3).When considering the Sertoli cell-preferentially expressed
4
5 330 genes, *SCIN* and *SLC4A11* were found very significantly decreased in MA and HS
6
7 331 samples when compared with CS. SCO samples, although presenting an increased
8
9 332 number of Sertoli cells, showed reduced levels of *SPAG7*, *SCIN* and *SLC4A11*
10
11 333 compared with CS controls ($p<0.006$) (Figure 2; Supplemental Table 3). The expression
12
13 334 of *FASLG* mRNA in SpF samples was increased compared with CS, although no
14
15 335 statistically significant changes in the average of *FASLG* expression were observed
16
17 336 among groups due to high standard deviation values in the infertile groups (CS control:
18
19 337 1.10 ± 0.71 ; HS: 1.79 ± 1.19 ; MA: 3.44 ± 2.87).
20
21
22
23

338

339 *Reduced cellular expression levels of germ cell-specific genes in SpF*

24
25 339
26
27 340 We additionally analyzed the transcript levels per cell of spermatogonia-associated
28
29 341 genes with a selective germ cell expression, in SpF subgroups compared to CS controls
30
31 342 in order to exclude the differences in gene expression due to changes in testicular
32
33 343 cellularity and to determine whether transcript level per cell is also altered in SpF.
34
35 344 Selective germ cell expression of *CCNE1*, *DAZL*, *RBM15* and *STRA8* was previously
36
37 345 confirmed as negligible transcript level values were found in SCO samples as
38
39 346 previously described, furthermore, *DAZL* and *STRA8* were previously described to be
40
41 347 expressed in spermatogonia but not in somatic tissues (Wang *et al.*, 2001). Values of
42
43 348 transcript amount per cell, in arbitrary units, were obtained for each testicular sample by
44
45 349 dividing the *CCNE1*, *DAZL*, *RBM15* and *STRA8* expression values by either the
46
47 350 proportion of the spermatogonia (Figure 3e, f, g and h), as it is the germ cell stage that
48
49 351 predominantly expresses *CCNE1*, *DAZL*, *RBM15* and *STRA8* or by the proportion of the
50
51 352 spermatogonia plus spermatocytes (Figure 3i, j, k and l), as the meiotic germ cells
52
53 353 potentially express these genes although at much lower levels (GermSAGE,
54
55
56
57
58
59
60

1
2
3 354 <http://germsage.nichd.nih.gov>; GermOnline <http://www.germonline.org>), present in a
4
5 355 seminiferous tubule of the sample. Significant differences in cellular transcript levels
6
7 356 were additionally found for *CCNE1*, *DAZL*, *RBM15* genes between SpF-HS patients
8
9 357 and controls and between SpF-MA patients and controls, when considering either the
10
11 358 proportion of spermatogonia or the proportion of spermatogonia and spermatocytes in
12
13 359 the tubule ($p \leq 0.002$). Interestingly, cellular transcript levels for *STRA8* were found
14
15 360 statistically decreased in SpF-MA when compared with CS when considering either the
16
17 361 proportion of spermatogonia or the proportion of spermatogonia and spermatocytes in
18
19 362 the tubule ($p \leq 0.002$). These results suggest that the decreased tissular expression levels
20
21 363 in SpF are not attributable to a decreased number of spermatogonia or spermatocytes in
22
23 364 the tubule but to a reduced number of transcripts in immature germ cells. Furthermore,
24
25 365 the decreased cellular *STRA8* expression levels observed in SpF-MA suggest that the
26
27 366 number of genes whose expression is altered in immature germ cells might be higher
28
29 367 than that initially observed in the whole tissue with meiotic arrest.

30
31
32
33
34 368 Strikingly, the decrease cellular expression observed in patients was
35
36 369 accompanied by a higher severity in spermatogenic impairment, and cellular expression
37
38 370 levels of *CCNE1*, *DAZL*, *RBM15* and *STRA8* genes were highly significantly and
39
40 371 positively correlated with the number of elongated spermatids in the tubule (Pearson's
41
42 372 correlation range; $r = 0.68-0.87$; $p < 0.0001$) (Supplemental Table 4.B.)
43
44

45 373

46 47 374 *Gene expression pattern in Germ cell tumours*

48
49 375 Relative expression values of fourteen spermatogonia preferentially- expressed genes
50
51 376 were found to have extremely significant differences in expression between GCT and
52
53 377 CS samples: *ATM*, *BARD1*, *CCNE1*, *CDKN1C*, *DAZL*, *DICER1*, *E2F3*, *FBXO32*,
54
55 378 *MRE11A*, *RAD50*, *RBM15*, *SIRT6*, *STRA8* and *TM9SF2* ($p < 0.002$). Less consistent
56
57
58
59
60

1
2
3 379 statistical differences were found for four additional genes *BAX*, *CCND1*, *POLD1* and
4
5 380 *XPA* ($p<0.05$). All these differentially expressed genes, with the exception of *BAX*, were
6
7 381 found to be under-represented in GCT samples compared with controls (Table 2;
8
9 382 Supplemental Table 3).

11 383 Regarding the somatically expressed genes; ten genes were found to have
12
13 384 statistical differences in expression: *AMHR2*, *BCL3*, *SCIN*, *SMARCA1*, *SOX9*, *SPAG7*,
14
15 385 *VEGFA* ($p<0.002$) *BMPRIA*, *CDKN1A* and *FASLG* ($p<0.05$). *FASLG* was found to be
16
17 386 over-expressed, whereas the other nine genes were under-expressed, in GCT samples
18
19 387 compared with controls (Table 2; Supplemental Table 3).

22 388 Some differences in expression behaviour were found when GCT samples were
23
24 389 divided into CIS and non-CIS samples (Supplemental Table 3): over-expression of the
25
26 390 *KIT* gene (an established marker for early-stage GCT) in CIS samples consistently
27
28 391 agreed with the histological diagnosis of samples. No difference in expression was
29
30 392 found for *BARD1*, *BMPRIA*, *CCND1*, *CDKN1A*, *CDKN1C*, *SMAD3* and *XPA* in CIS
31
32 393 samples compared with CS controls, whereas their transcript values were found highly
33
34 394 decreased in non-CIS samples ($p<0.002$). *DLKI* was significantly over-expressed in
35
36 395 CIS samples ($p=0.012$), contrary to the marked under-expression found in non-CIS
37
38 396 samples ($p<0.002$). Thus, this set of genes is, somehow, associated with different stages
39
40 397 in tumour progression.
41
42
43
44

45 398

47 399 **Expression signature by functional categories**

49 400 In order to obtain some clues about the functional pathways that are affected in testis
50
51 401 with SpF, genes were grouped into functional clusters according to the process they are
52
53 402 involved in: cell proliferation, apoptosis/cell cycle, meiosis, DNA repair, transcription
54
55 403 regulation, post-transcriptional regulation and degradation. The functional expression
56
57
58
59
60

1
2
3 404 signature was compared with that obtained from the GCT samples (Supplemental Table
4
5 405 3).

6
7 406 Some of the spermatogonia-preferential transcripts differentially represented
8
9 407 between the SpF subgroups and controls encode proteins involved in the regulation of
10
11 408 the mitotic and meiotic cell cycle (cyclins, cyclin –dependent-kinase-inhibitors and
12
13 409 DNA polymerases). *CCNE1* and *POLD1* were significantly under-expressed, this being
14
15 410 particularly noticeable for *CCNE1*, whereas *CCND1* was over-expressed in the MA
16
17 411 samples. In GCT only the *CCNE1* transcript value of this group of genes was
18
19 412 significantly decreased compared with controls. Furthermore, we observed decreased
20
21 413 expression levels of several genes that encode for putative RNA-binding proteins (such
22
23 414 as the germ cell specific *DAZL* and *RBM15*), other proteins essential for production of
24
25 415 miRNAs (such as *DICER1*), proteins involved in the protein degradation pathway
26
27 416 (*FBXO32* and *TM9SF2*) and also proteins implicated in the homologous recombination
28
29 417 in meiosis (*MRE11A* and *RAD50*) in both SpF and GCT groups.

30
31
32
33
34 418 There were no differences in expression of genes such as *BMPRIA*, *c-KIT* and
35
36 419 *VEGFA* (cell population proliferation), *BAX* gene (apoptosis), *ATM*, *SIRT6* and *XPA*
37
38 420 (DNA repair genes) observed between SpF and CS groups, in contrast to the observed
39
40 421 expression alteration of these genes in GCT.

41
42 422 The expression of all the Sertoli-specific expressed genes, with the exception of
43
44 423 *SLC4A11* was significantly affected in the GCT samples, whereas only those genes
45
46 424 involved in cell cycle/apoptosis, *SLC4A11* and *SCIN*, were significantly decreased in
47
48 425 the SpF meiotic altered samples.

49
50
51 426

52
53
54 427 **Protein expression**
55
56
57
58
59
60

1
2
3 428 We sought to determine whether the changes in transcript levels would correlate with
4
5 429 modifications at the protein level. At the same time we aimed to determine whether the
6
7 430 suggested alterations in gene expression affected expression levels of encoded proteins
8
9 431 in the germ line. We focused first on robust differences in expression levels which
10
11 432 might be discerned by immunohistochemistry and chose *CCNE1* and *DAZL* as examples
12
13 433 of spermatogonia-associated genes specifically expressed in the germ line. Secondly, we
14
15 434 selected proteins whose coding genes presented an upregulated expression in SpF
16
17 435 samples such as *DLK1* and *CDKN1C* (Figure 4).
18
19

20
21 436 *CCNE1* immunostaining (A-D; Figure 4) in CS samples was observed mainly in
22
23 437 both the nucleus and the cytoplasm of postmeiotic secondary spermatocytes/round
24
25 438 spermatids, and in the in nucleus of spermatogonia. Sertoli cells showed less intense
26
27 439 *CCNE1* expression in the nucleus; some staining of the cytoplasm of interstitial Leydig
28
29 440 cells and Sertoli cells was also seen, which tended to increase in samples with MA and
30
31 441 SCO; in the germ line of SpF samples *CCNE1* expression was exclusively detected in
32
33 442 the spermatogonia and not detected in primary spermatocytes and elongated spermatids,
34
35 443 when present. The *DAZL* protein (E-H, Figure 4) was almost exclusively present in the
36
37 444 cytoplasm of primary pachytene spermatocytes in CS samples and therefore expression
38
39 445 was negative in testis with SCO. Overall, immunoexpression in testis sections of both
40
41 446 *CCNE1* and *DAZL* protein decreased within seminiferous tubules, in the germ line, as
42
43 447 spermatogenic damage progressed, showing good correlation with RNA expression.
44
45

46
47 448 *CDKN1C* (I-L; Figure 4) in CS samples was seen in the cytoplasm of Sertoli
48
49 449 cells and Leydig cells; the immunostaining for Leydig cells was more intense in MA
50
51 450 and SCO samples; peritubular cells were positive in SCO; however in SpF samples
52
53 451 some spermatogonia and primary spermatocytes also exhibited moderate expression of
54
55 452 *CDKN1C*, leading to a global increase in *CDKN1C* expression in biopsies with SpF.
56
57
58
59
60

1
2
3 453 The immunoreactivity of DLK1 (M-P; Figure 4) in CS samples seemed to be restricted
4
5 454 to the cytoplasm of a few Leydig cells, which were more frequently stained in MA, and
6
7 455 more so in SCO suggesting a relevant contribution of Leydig cells to DLK1 expression.
8
9 456 Overall, immunoexpression of both CDKN1C and DLK1 protein in testis sections,
10
11 457 preferentially in somatic cells for DLK1, increased as spermatogenic damage
12
13 458 progressed, showing good correlation with RNA expression.
14
15

16 459
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1
2
3 460 **DISCUSSION**

4
5 461 The aim of this study was to assess early testicular transcriptional changes that could be
6
7 462 involved in human severe deficiency of sperm production. We focused our attention
8
9 463 mainly on the expression profile of premeiotic germ cells as this is a key step in male
10
11 464 germ cell maturation. The accurate quantification of testicular mRNA levels in SpF by
12
13 465 RT-qPCR experiments led to the identification of differences in expression of certain
14
15 466 genes associated with spermatogonia in the absence of any apparent morphological
16
17 467 and/or numerical change of this specific cell type.

18
19
20 468 The gene expression profile in SpF can be used as a basis for identification of
21
22 469 candidate genes that contribute to spermatogenic impairment. SpF and GCT expression
23
24 470 signature comparison could additionally give some clues about the molecular
25
26 471 mechanisms underlying the origin of these alterations. Our data indicate that, in the SpF
27
28 472 patients, a large proportion of spermatogonia-preferentially expressed genes exhibited
29
30 473 reduced testicular expression levels when compared with CS individuals. As expected,
31
32 474 the number of genes whose expression was altered as well as the magnitude of increase
33
34 475 or decrease of gene expression in GCT was even higher, possibly related to the fact that
35
36 476 the germ line in testicular tumour has undergone a dedifferentiation process
37
38 477 representing an extreme situation of gene expression deregulation of spermatogenic
39
40 478 impairment. Interestingly, gene expression signatures of both phenotypes, SpF and
41
42 479 GCT, share some aberrant patterns of gene expression supporting the idea that the
43
44 480 participation of these genes is essential for physiological germ cell development. In
45
46 481 contrast, other genes are differentially affected in both pathological groups suggesting
47
48 482 that they contribute to the phenotype and could be used as potential molecular markers.

49
50 483 The cellular complexity of the testis is an inherent problem which should be
51
52 484 taken into account when studying gene expression profiles in this organ. As the
53
54
55
56
57
58
59
60

1
2
3 485 pathological seminiferous tubules lack germ cells to varying degrees, changes in gene
4
5 486 expression at the tissue level can reflect changes in the capability for transcribing the
6
7 487 mRNA in a specific cell type as well as changes in the cell type composition in
8
9 488 pathological testis. The absence of significant differences in the spermatogonia and
10
11 489 Sertoli cell number among the SpF-MA, SpF-HS and CS groups in our study indicates
12
13 490 that transcript levels cannot be attributed to the presence or absence of these specific
14
15 491 cell types. Additionally, we are aware of the fact that the levels of different mRNAs
16
17 492 could change as the proportion of immature germ cells is different in tubules in
18
19 493 conserved spermatogenesis, in which all stages of germ cell are present, and in those in
20
21 494 maturation arrest, where only some stages are present. In this condition, we assume that
22
23 495 there would be less mRNA from most spermatogonia-expressed genes in CS testis than
24
25 496 in SpF testis. Our study shows that most of the differentially expressed genes showed
26
27 497 higher expression in testis with conserved spermatogenesis suggesting that the reduction
28
29 498 of target genes could not be attributable to either the spermatogonia cell number or to
30
31 499 different proportion of this germ cell stage in the tubule, but to real differences in the
32
33 500 expression capability of the cell. Furthermore, we describe reduced cellular expression
34
35 501 levels of four germ line specific genes in SpF samples supporting this premise.
36
37
38
39

40 502 The spermatogonia-related genes whose transcripts were differentially
41
42 503 represented between SpF subgroups and controls included genes involved in specific
43
44 504 functional pathways. A first group of genes encodes proteins involved in the regulation
45
46 505 of the mitotic and meiotic cell cycle such as cyclins, cyclin-dependent-kinase-inhibitors
47
48 506 and DNA polymerases indicating that the regulation of this specific functional gene
49
50 507 cluster in the initial stages of spermatogenesis is critical for further differentiation and
51
52 508 meiosis of germ cells. Additionally, decreased expression levels of several genes
53
54 509 encoding putative RNA-binding proteins (such as the germ cell specific *DAZL* and
55
56
57
58
59
60

1
2
3 510 *RBM15*), other proteins essential for production of miRNAs such as *DICER1*, and also
4
5 511 proteins involved in the protein degradation pathway (*FBXO32* and *TM9SF2*) in both
6
7 512 SpF and GCT groups, underlie the complexity of post-transcriptional control in
8
9 513 proliferation and differentiation of germ cells. *MRE11A* and *RAD50*, involved in
10
11 514 homologous recombination in meiosis, also showed altered expression in meiotic
12
13 515 blockade. Unexpectedly, we observed no difference in testicular expression for *STRA8*,
14
15 516 a well known gene involved in meiotic cell cycle, participating in chromosome pairing
16
17 517 and in the process that leads to stable commitment to the meiotic cycle (Mark *et al.*,
18
19 518 2008), in SpF group nor in SpF-MA subtype samples when compared with controls.
20
21 519 Interestingly, an additional statistically significant reduction in the expression levels of
22
23 520 germ cell-specific genes per spermatogonia was observed in MA and HS when
24
25 521 compared to CS samples, demonstrated for *CCNE1*, *DAZL* and *RBM15* genes, being
26
27 522 more pronounced in the MA phenotype suggesting that the expression capacity in
28
29 523 immature germ cells correlates with the severity of testicular damage. More
30
31 524 intriguingly, a decreased cellular expression level was even observed for *STRA8* in MA
32
33 525 pattern, suggesting that in MA phenotype, premeiotic cellular expression could be
34
35 526 affected for a larger number of genes. The remarkable correlation coefficient between
36
37 527 the *CCNE1*, *DAZL*, *RBM15* and *STRA8* transcript levels per cell and the number of
38
39 528 elongated spermatids in the testicular tubule additionally underlines the determinant role
40
41 529 of premeiotic *CCNE1*, *DAZL*, *RBM15* and *STRA8* expression in the progression of the
42
43 530 spermatogenic process.
44
45
46
47
48

49
50 531 In previous studies, the reduction of gene expression in SpF patients has been
51
52 532 mainly attributed to the decreased number of germ cells that specifically express the
53
54 533 gene of interest (i.e. significantly lower *DAZL* mRNA concentrations were previously
55
56 534 found in testes of non-obstructive azoospermic men (Lin *et al.*, 2001)), although germ
57
58
59
60

1
2
3 535 cell quantification was not properly performed. Here, we demonstrate that the changes
4
5 536 in expression observed among groups could be not exclusively explained by the
6
7 537 immature germ cell number but the contribution of the reduced cellular expression of
8
9 538 i.e. *DAZL* mRNA in spermatogenic impairment should be also taken into account.

11 539 Protein data on nonobstructive testicular tissue corroborate our mRNA
12
13 540 expression results: CCNE1 and *DAZL* protein decreased within seminiferous tubules, in
14
15 541 the germ line, as spermatogenic damage progressed. The reduced transcript levels of
16
17 542 other genes expressed in spermatogonia as the ones involved in piRNA processing
18
19 543 machinery such as *PIWIL2* and *TDRD1* in SpF, as we recently described (Heyn *et al.*,
20
21 544 2012) further supports the role of a proper gene expression in early germ line stages for
22
23 545 a successful sperm production.

24
25
26
27 546 The expression levels of genes participating in cell population proliferation,
28
29 547 mitochondria-mediated apoptosis and DNA repair (assuring the maintenance of genome
30
31 548 integrity) are in general maintained in SpF, contrary than in GCT, supporting the idea
32
33 549 that they are similarly processed in meiotic derangement and in conserved
34
35 550 spermatogenesis. However, there is one exception possibly due to other regulatory
36
37 551 pathways: the levels of spermatogonia-specific full-length *BARD1* transcript (Feki *et*
38
39 552 *al.*, 2004; Irminger-Finger *et al.*, 2001), involved in germ cell apoptotic events, could be
40
41 553 repressed in SpF by the high levels of FSH in spermatogenic failure samples (Feki *et*
42
43 554 *al.*, 2004).

44
45
46
47 555 Although the transcription profile of spermatogonia-associated genes in SpF is
48
49 556 different from that in GCT, some functional clusters are affected in both phenotypes:
50
51 557 genes with functions in cell cycle, transcription and post-transcriptional regulation and
52
53 558 protein degradation. Meanwhile, other spermatogonia-expressed genes encoding
54
55 559 proteins involved in cell proliferation, apoptosis and DNA repair pathways are not
56
57
58
59
60

1
2
3 560 affected in SpF, suggesting that in spermatogenic failure, although the abnormal
4
5 561 initiation of the meiotic process is already determined in these immature germ cells,
6
7 562 these cells maintain their activity related to mitosis and cell proliferation.
8

9
10 563 Regarding the Sertoli-specific expressed genes, the absence of differences in
11
12 564 gene expression of six out of eight genes studied in the SpF-MA and SpF-HS
13
14 565 phenotypes suggest that, in spermatogenic derangement, the functions of Sertoli cells
15
16 566 are in general maintained, but not those involved in cell cycle/apoptosis (*SLC4A11* and
17
18 567 *SCIN*) related to germ cell support. Previous studies have shown that, in mouse,
19
20 568 chemically-induced germ cell depletion can alter expression of several Sertoli cell genes
21
22 569 (Jonsson *et al.*, 1999; Maguire *et al.*, 1993; O'shaughnessy *et al.*, 2008) demonstrating
23
24 570 that germ cells regulate Sertoli cell activity by means of the regulation of Sertoli cell
25
26 571 gene expression. Here we observed a similar pattern in a pathological naturally-
27
28 572 occurring phenotype. We observed a gradual decrease of *SLC4A11* and *SCIN* transcript
29
30 573 levels attributable to the progressive depletion of germ cell stages (SCO>MA>HS>CS).
31
32 574 It is noteworthy that even in the presence of a two-fold increase of the number of Sertoli
33
34 575 cells as a compensatory phenomenon in SCO, the decrease in expression of certain
35
36 576 Sertoli cell genes becomes much more noticeable in the total absence of germ cells.
37
38
39

40 577 Some recent microarray studies have assessed global gene expression analysis in
41
42 578 testicular biopsies from infertile men in order to identify the genes critical for
43
44 579 spermatogenesis (Chalmel *et al.*, 2012; Von Kopylow *et al.*, 2010; Ellis *et al.*, 2007;
45
46 580 Feig *et al.*, 2007; Rockett *et al.*, 2004; Fox *et al.*, 2003). In these studies specific germ
47
48 581 cell transcription patterns are inferred from infertile testicular phenotypes in men and a
49
50 582 pattern of significantly decreased regulated genes has been attributed to the degree of
51
52 583 spermatogenic failure and the loss of specific stages of germ cells. We provide data
53
54 584 suggesting that the molecular basis for severe spermatogenic impairment is more
55
56
57
58
59
60

1
2
3 585 complex than initially proposed. In SpF the immature germ cells present an altered and
4
5 586 decreased transcriptional pattern of certain genes, and thus the number of genes
6
7 587 associated with these cells could be underestimated from microarray studies of infertile
8
9 588 samples. Furthermore, our results should be helpful to better interpret microarray or
10
11 589 future NGS transcriptome studies. Elucidation of a more extensive transcriptional
12
13 590 profile with the detailed analysis of testicular cellular composition could be important in
14
15 591 understanding the molecular mechanisms that underlie male infertility.

16
17
18 592 It is conspicuous that most of the altered spermatogonia-related genes are
19
20 593 involved in essential processes during spermatogenesis and aberrant expression is often
21
22 594 associated with spermatogenic defects. Whether the observed differential expression
23
24 595 profiles represent the cause or consequence of maturation arrest remains to be
25
26 596 elucidated. Considering the heterogeneous aetiologies and highly individual molecular
27
28 597 causes which may underlie spermatogenic failure in humans, the molecular changes
29
30 598 described here may represent common symptoms, but may also reflect early
31
32 599 dysfunction events affecting germ cells which may causally contribute to the pathology.
33
34 600 This data should be useful in delineating the patterns of gene expression involved in
35
36 601 male germ cell maturation deficiency, which may contribute to understanding male
37
38 602 infertility.

39
40
41 603 In summary, our study provides evidence that the premeiotic stage of germ cell
42
43 604 differentiation, exhibits associated patterns of gene expression deregulation in
44
45 605 spermatogenic impairment, which is more severe in meiotic arrest. This altered gene
46
47 606 expression pattern is observed despite there being no apparent morphological and/or
48
49 607 numerical change observed in this early stage of the germ cell population. In our
50
51 608 previous study, the spermatocyte capacity to express meiosis-related genes was
52
53 609 observed to be markedly reduced in spermatogenic failure, contributing to meiosis
54
55
56
57
58
59
60

1
2
3 610 impairment (Terribas *et al.*, 2010). Our present data demonstrate that the low
4
5 611 spermatogenic efficiency in infertile men is accompanied not only by meiotic but also
6
7 612 by premeiotic events in spermatogenesis, which contribute to spermatogenic blockade.
8
9 613 Furthermore, the differences in expression during the initial stages of spermatogenesis
10
11 614 in SpF-MA individuals suggest that this phenotype is already determined or arises in the
12
13 615 premeiotic stages of the germ line.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1
2
3 617 **ACKNOWLEDGEMENTS**
4

5 618 We are indebted to the patients who participated in this study. We thank Begoña
6
7 619 Hurtado for qBase analysis support, Xavier Solé for suggestions about statistical
8
9 620 analysis and Harvey Evans for the revision of the English text.
10
11

12
13 621

14
15 622 **FUNDING**
16

17
18 623 This work was supported by grants from the *Fondo de Investigaciones*
19
20 624 *Sanitarias/Fondo Europeo de Desarrollo Regional* (FIS/FEDER) [grant number
21
22 625 PI05/0759, PI09/1727, PI12/00361] and the *Generalitat de Catalunya* [grant number
23
24 626 2009SGR01490]. S.L. is sponsored by the Researchers Stabilization Program from the
25
26 627 Spanish National Health System [CES09/020]. S.B. was supported by the *Fondo de*
27
28 628 *Investigaciones Sanitarias-Instituto de Salud Carlos III* (FIS-ISCiii) [CA06/0055].
29

30
31 629

32
33
34 630 **AUTHOR'S CONTRIBUTION**
35

36
37 631 S.B. performed the RNA experiments and analysed the data. F.A. performed
38
39 632 immunohistochemistry and histological interpretation of data. E.F. provided samples
40
41 633 and clinical data. L.B. performed clinical assessment, provided samples, and critically
42
43 634 reviewed the manuscript. S.L. conceived and designed the experiments, supervised the
44
45 635 analysis of data and wrote the manuscript.
46
47

48
49 636

50
51 637
52
53
54
55
56
57
58
59
60

1
2
3 638 **FIGURE LEGENDS**

4
5 639 **Supplemental Figure 1.** Absolute expression levels of candidate reference genes (a.) in
6
7 640 SCO, MA, HS and CS groups and (b.) in GCT and CS groups.

8
9 641 ○, □, outlying values. * $p < 0.05$, (a.) Kruskal-Wallis test and (b.) Mann-Whitney test.

10
11 642 SCO: Sertoli cell only syndrome; MA: maturation arrest at spermatocyte stage; HS:
12
13 643 hypospermatogenesis; CS: conserved spermatogenesis

14
15 644

16
17
18 645 **Figure 1.** Spermatogonia-preferentially expressed genes whose relative expression
19
20 646 values were statistically altered in patient SpF group (MA and HS sub-phenotypes)
21
22 647 compared with CS controls. Expression levels relative to *PGK1/PGM1* are shown.

23
24 648 * $p < 0.05$, Mann-Whitney test; ** $p < 0.002$, Mann-Whitney test and Bonferroni
25
26 649 correction.

27
28
29 650 MA: maturation arrest at spermatocyte stage; HS: hypospermatogenesis; CS: conserved
30
31 651 spermatogenesis.

32
33 652

34
35
36 653 **Figure 2.** Testicular somatic cell-preferentially expressed genes that showed differences
37
38 654 in gene expression in infertile SCO, MA and HS patients relative to CS controls.
39
40 655 Expression levels relative to *PGK1/PPIA* are shown..

41
42 656 * $p < 0.0083$, Mann-Whitney test and Bonferroni correction.

43
44
45 657 SCO: Sertoli cell only syndrome; MA: maturation arrest at spermatocyte stage; HS:
46
47 658 hypospermatogenesis; CS: conserved spermatogenesis.

48
49 659

50
51 660 **Figure 3.**

52
53 661 Tissular expression profiling of *CCNE1* (a), *DAZL* (b), *RBM15* (c) and *STRA8* (d) by
54
55 662 quantitative real-time qPCR in testis with conserved spermatogenesis (CS),

56
57
58
59
60

1
2
3 663 hypospermatogenesis (HS) and maturation arrest at the spermatocyte (MA) Expression
4
5 664 levels relative to *PGKI* and *PGMI* are shown. Expression per cell profiling of *CCNE1*,
6
7 665 *DAZL*, *RBM15* and *STRA8* displayed as expression ratio per spermatogonium (x100) (e,
8
9 666 f, g, h) and expression per cell profiling of *CCNE1*, *DAZL*, *RBM15* and *STRA8*
10
11 667 displayed as expression ratio per spermatogonium and spermatocyte (x100) (i, j, k, l).
12
13 668 The horizontal bar indicates median value. Significant differences from the control are
14
15 669 indicated: * $p < 0.05$; ** $p < 0.005$.
16
17
18
19

20
21 **Figure 4.**

22
23 672 Immunohistochemical localization of selected proteins in sections of human testes with
24
25 673 different phenotypes. From left to right, first column (A,E,I,M) shows sections of CS,
26
27 674 second column (B,F,J,N) corresponds to SpF-HS, third column (C,G,K,O) represents
28
29 675 SpF-MA, and the fourth column (D,H,L,P) displays SCO pattern. *CCNE1* protein
30
31 676 staining is shown in panels A-D, *DAZL* in E-H, *CDKN1C* in I-L, and *DLK1* expression
32
33 677 in M-P, were stained Leydig cells are indicated by arrows. See explanation of the
34
35 678 cellular localization of different proteins in the text. Original magnification was X400
36
37 679 for panels A-L, and X200 for M-P. Scale bar in A and M = 100 μm .
38
39
40

41 680
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

681 **REFERENCES**

682

683 Chalmel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, Gattiker A,
684 Schulze W, Jegou B, Kirchhoff C & Primig M (2012) Global human tissue
685 profiling and protein network analysis reveals distinct levels of transcriptional
686 germline-specificity and identifies target genes for male infertility. *Hum Reprod*,
687 27, 3233-3248.

688 Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker
689 A, Moore J, Patard JJ, Wolgemuth DJ, Jegou B & Primig M (2007) The
690 conserved transcriptome in human and rodent male gametogenesis. *Proc Natl
691 Acad Sci U S A*, 104, 8346-8351.

692 Diederichs S, Baumer N, Schultz N, Hamra FK, Schrader MG, Sandstede ML, Berdel
693 WE, Serve H & Muller-Tidow C (2005) Expression patterns of mitotic and
694 meiotic cell cycle regulators in testicular cancer and development. *Int J Cancer*,
695 116, 207-217.

696 Ellis PJ, Furlong RA, Conner SJ, Kirkman-Brown J, Afnan M, Barratt C, Griffin DK &
697 Affara NA (2007) Coordinated transcriptional regulation patterns associated
698 with infertility phenotypes in men. *J Med Genet*, 44, 498-508.

699 Feig C, Kirchhoff C, Ivell R, Naether O, Schulze W & Spiess AN (2007) A new
700 paradigm for profiling testicular gene expression during normal and disturbed
701 human spermatogenesis. *Mol Hum Reprod*, 13, 33-43.

702 Feki A, Jefford CE, Durand P, Harb J, Lucas H, Krause KH & Irminger-Finger I (2004)
703 BARD1 expression during spermatogenesis is associated with apoptosis and
704 hormonally regulated. *Biol Reprod*, 71, 1614-1624.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 705 Fox MS, Ares VX, Turek PJ, Haqq C & Reijo Pera RA (2003) Feasibility of global
4
5 706 gene expression analysis in testicular biopsies from infertile men. *Mol Reprod*
6
7 707 *Dev*, 66, 403-421.
8
9
10 708 Guo R, Yu Z, Guan J, Ge Y, Ma J, Li S, Wang S, Xue S & Han D (2004) Stage-
11
12 709 specific and tissue-specific expression characteristics of differentially expressed
13
14 710 genes during mouse spermatogenesis. *Mol Reprod Dev*, 67, 264-272.
15
16 711 Hellemans J, Mortier G, De Paepe A, Speleman F & Vandesompele J (2007) qBase
17
18 712 relative quantification framework and software for management and automated
19
20 713 analysis of real-time quantitative PCR data. *Genome Biol*, 8, R19.
21
22
23 714 Heyn H, Ferreira HJ, Bassas L, Bonache S, Sayols S, Sandoval J, Esteller M & Larriba
24
25 715 S (2012) Epigenetic disruption of the PIWI pathway in human spermatogenic
26
27 716 disorders. *PLoS One*, 7, e47892.
28
29
30 717 Irminger-Finger I, Leung WC, Li J, Dubois-Dauphin M, Harb J, Feki A, Jefford CE,
31
32 718 Soriano JV, Jaconi M, Montesano R & Krause KH (2001) Identification of
33
34 719 BARD1 as mediator between proapoptotic stress and p53-dependent apoptosis.
35
36 720 *Mol Cell*, 8, 1255-1266.
37
38
39 721 Jonsson CK, Zetterstrom RH, Holst M, Parvinen M & Soder O (1999) Constitutive
40
41 722 expression of interleukin-1alpha messenger ribonucleic acid in rat Sertoli cells is
42
43 723 dependent upon interaction with germ cells. *Endocrinology*, 140, 3755-3761.
44
45 724 Lin YM, Chen CW, Sun HS, Tsai SJ, Hsu CC, Teng YN, Lin JS & Kuo PL. (2001)
46
47 725 Expression patterns and transcript concentrations of the autosomal DAZL gene
48
49 726 in testes of azoospermic men. *Mol Hum Reprod*. 7, 1015-1022.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 728 Maguire SM, Millar MR, Sharpe RM & Saunders PT (1993) Stage-dependent
4
5 729 expression of mRNA for cyclic protein 2 during spermatogenesis is modulated
6
7 730 by elongate spermatids. *Mol Cell Endocrinol*, 94, 79-88.
8
9
10 731 Mark M, Jacobs H, Oulad-Abdelghani M, Dennefeld C, Feret B, Vernet N, Codreanu
11
12 732 CA, Chambon P & Ghyselinck NB (2008) STRA8-deficient spermatocytes
13
14 733 initiate, but fail to complete, meiosis and undergo premature chromosome
15
16 734 condensation. *J Cell Sci*, 121, 3233-3242.
17
18 735 McLachlan RI, Rajpert-De Meyts E, Høei-Hansen CE, de Kretser DM & Skakkebaek
19
20 736 NE (2007) Histological evaluation of the human testis--approaches to optimizing
21
22 737 the clinical value of the assessment: mini review. *Hum Reprod*, 22, 2-16.
23
24 738 Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD & Lee JT
25
26 739 (2006) Postmeiotic sex chromatin in the male germline of mice. *Curr Biol*, 16,
27
28 740 660-667.
29
30 741 O'Shaughnessy PJ, Hu L & Baker PJ (2008) Effect of germ cell depletion on levels of
31
32 742 specific mRNA transcripts in mouse Sertoli cells and Leydig cells.
33
34 743 *Reproduction*, 135, 839-850.
35
36
37 744 Pang AL, Taylor HC, Johnson W, Alexander S, Chen Y, Su YA, Li X, Ravindranath N,
38
39 745 Dym M, Rennert OM & Chan WY (2003) Identification of differentially
40
41 746 expressed genes in mouse spermatogenesis. *J Androl*, 24, 899-911.
42
43
44 747 Rockett JC, Patrizio P, Schmid JE, Hecht NB & Dix DJ (2004) Gene expression
45
46 748 patterns associated with infertility in humans and rodent models. *Mutat Res*,
47
48 749 549, 225-240.
49
50
51 750 Schlecht U, Demougin P, Koch R, Hermida L, Wiederkehr C, Descombes P, Pineau C,
52
53 751 Jegou B & Primig M (2004) Expression profiling of mammalian male meiosis
54
55
56
57
58
59
60

- 1
2
3 752 and gametogenesis identifies novel candidate genes for roles in the regulation of
4
5 753 fertility. *Mol Biol Cell*, 15, 1031-1043.
6
7 754 Schultz N, Hamra FK & Garbers DL (2003) A multitude of genes expressed solely in
8
9 755 meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive
10
11 756 targets. *Proc Natl Acad Sci U S A*, 100, 12201-12206.
12
13
14 757 Schulze W, Thoms F & Knuth UA (1999) Testicular sperm extraction: comprehensive
15
16 758 analysis with simultaneously performed histology in 1418 biopsies from 766
17
18 759 subfertile men. *Hum Reprod*, 14 Suppl 1, 82-96.
19
20
21 760 Sha J, Zhou Z, Li J, Yin L, Yang H, Hu G, Luo M, Chan HC & Zhou K (2002)
22
23 761 Identification of testis development and spermatogenesis-related genes in human
24
25 762 and mouse testes using cDNA arrays. *Mol Hum Reprod*, 8, 511-517.
26
27 763 Shima JE, McLean DJ, McCarrey JR & Griswold MD (2004) The murine testicular
28
29 764 transcriptome: characterizing gene expression in the testis during the progression
30
31 765 of spermatogenesis. *Biol Reprod*, 71, 319-330.
32
33
34 766 Simoni M, Bakker E, Eurlings MC, Matthijs G, Moro E, Muller CR & Vogt PH (1999)
35
36 767 Laboratory guidelines for molecular diagnosis of Y-chromosomal
37
38 768 microdeletions. *Int J Androl*, 22, 292-299.
39
40
41 769 Simoni M, Bakker E & Krausz C (2004) EAA/EMQN best practice guidelines for
42
43 770 molecular diagnosis of y-chromosomal microdeletions. State of the art 2004. *Int*
44
45 771 *J Androl*, 27, 240-249.
46
47 772 Terribas E, Bonache S, Garcia-Arevalo M, Sanchez J, Franco E, Bassas L & Larriba S
48
49 773 (2010) Changes in the expression profile of the meiosis-involved mismatch
50
51 774 repair genes in impaired human spermatogenesis. *J Androl*, 31, 346-357.
52
53
54 775 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A &
55
56 776 Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR
57
58
59
60

1
2
3 777 data by geometric averaging of multiple internal control genes. *Genome Biol*, 3,
4
5 778 RESEARCH0034.
6
7 779 von Kopylow K, Kirchhoff C, Jezek D, Schulze W, Feig C, Primig M, Steinkraus V &
8
9 780 Spiess AN (2010) Screening for biomarkers of spermatogonia within the human
10
11 781 testis: a whole genome approach. *Hum Reprod*, 25, 1104-1112.
12
13 782 Wang PJ, McCarrey JR, Yang F & Page DC (2001) An abundance of X-linked genes
14
15 783 expressed in spermatogonia. *Nat Genet*, 27, 422-426.
16
17 784 World Health Organization (1999) Laboratory Manual for the Examination of Human
18
19 785 Semen and Sperm-Cervical Mucus Interaction, p. Cambridge University Press,
20
21 786 New York.
22
23
24
25 787
26 788
27
28 789
29
30 790
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1 Quantitative histological evaluation of the testicular samples included in the study

Histological pattern	Number of samples	Spermatogoniae	Spermatocytes I	Round spermatids	Elongated spermatids	Sertoli cells	JS
SCO	14	0	0	0	0	21.03 ± 7.00	1-2
SpF-MA	7	20.28 ± 4.38	23.30 ± 11.73	2.27 ± 3.10	0.52 ± 0.81	13.87 ± 3.01	4-6
SpF-HS	12	20.53 ± 6.68	29.77 ± 10.44	16.48 ± 8.07	5.70 ± 4.81	15.74 ± 5.61	7-8
CS	17	21.96 ± 4.27	32.30 ± 6.64	25.38 ± 12.04	21.14 ± 7.07	13.62 ± 3.11	9-10

The mean number ± SD of different type of cells/tubule are given for each Johnsen score classified group

Abbreviations: SCO: Sertoli cell only syndrome, SpF-MA: spermatogenic failure diagnosed with meiotic arrest, SpF-HS: spermatogenic failure diagnosed with hypospermatogenesis, CS: conserved spermatogenesis, JS: Johnsen score

Pre Peer Review

Table 2. List of genes differentially expressed in GCT

Gene name	Relative expression values (PGK1/PPIA)		fold-change	<i>p</i> -value
	CS group	GCT group		
Spermatogonia target genes that are under-expressed				
ATM	1.05 ± 0.25	0.32 ± 0.18	-3.28	0.000076**
BARD1	1.02 ± 0.15	0.55 ± 0.25	-1.85	0.002202**
CCND1	0.73 ± 0.17	0.35 ± 0.36	-2.08	0.030987*
CCNE1	1.04 ± 0.25	0.16 ± 0.12	-6.50	0.000076**
CDKN1C	0.49 ± 0.13	0.19 ± 0.13	-2.58	0.000911**
DAZL	1.04 ± 0.20	0.06 ± 0.08	-17.33	0.000076**
DICER1	1.13 ± 0.22	0.21 ± 0.12	-4.90	0.000076**
E2F3	1.00 ± 0.14	0.58 ± 0.18	-1.72	0.000304**
FBXO32	0.82 ± 0.18	0.17 ± 0.18	-4.82	0.000076**
MRE11A	1.09 ± 0.25	0.29 ± 0.19	-3.75	0.000076**
POLD1	1.14 ± 0.36	0.59 ± 0.38	-1.93	0.030987*
RAD50	1.09 ± 0.26	0.16 ± 0.12	- 6.81	0.000076**
RBM15	0.99 ± 0.28	0.12 ± 0.06	-8.25	0.000076**
SIRT6	1.00 ± 0.21	0.36 ± 0.13	-2.77	0.000076**
STRA8	0.78 ± 0.31	0.13 ± 0.16	-6.00	0.000304**
TM9SF2	0.93 ± 0.14	0.24 ± 0.18	-3.87	0.000076**
XPA	1.06 ± 0.21	0.43 ± 0.41	-2.46	0.006304*
Spermatogonia target genes that are over-expressed				
BAX	0.97 ± 0.14	1.19 ± 0.18	+1.22	0.019291*
Somatic cell target genes that are under-expressed				
AMHR2	1.01 ± 0.25	0.20 ± 0.26	-5.05	0.000076**
BCL3	0.93 ± 0.29	0.45 ± 0.15	-2.06	0.000076**
BMPRI1A	0.97 ± 0.10	0.65 ± 0.28	-1.49	0.014962*
CDKN1A	0.97 ± 0.35	0.40 ± 0.38	-2.42	0.030987*
SCIN	1.15 ± 0.45	0.06 ± 0.05	-19.16	0.000076**
SMARCA1	1.02 ± 0.14	0.23 ± 0.29	-4.43	0.000076**
SOX9	0.95 ± 0.21	0.18 ± 0.23	-5.27	0.000076**
SPAG7	0.98 ± 0.12	0.38 ± 0.24	-2.58	0.000076**
VEGFA	0.85 ± 0.23	0.40 ± 0.19	-2.12	0.001443**
Somatic-cell target genes that are over-expressed				
FASLG	0.71 ± 0.84	3.89 ± 2.95	+5.48	0.002850*

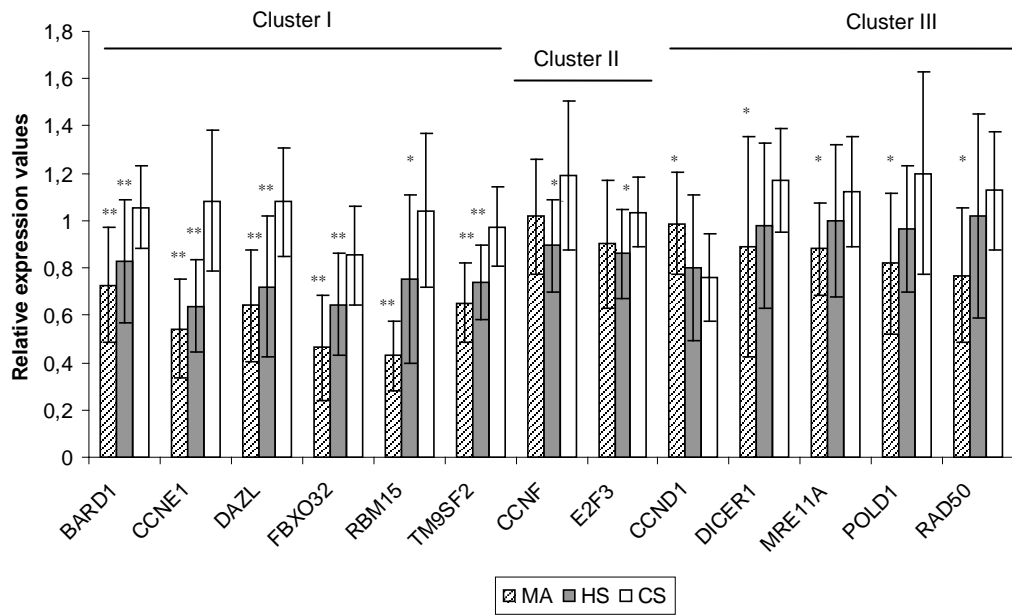
Normalisers are described in brackets.

* $p < 0.05$, ** $p < 0.002$, Mann-Whitney test.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

Figure 1

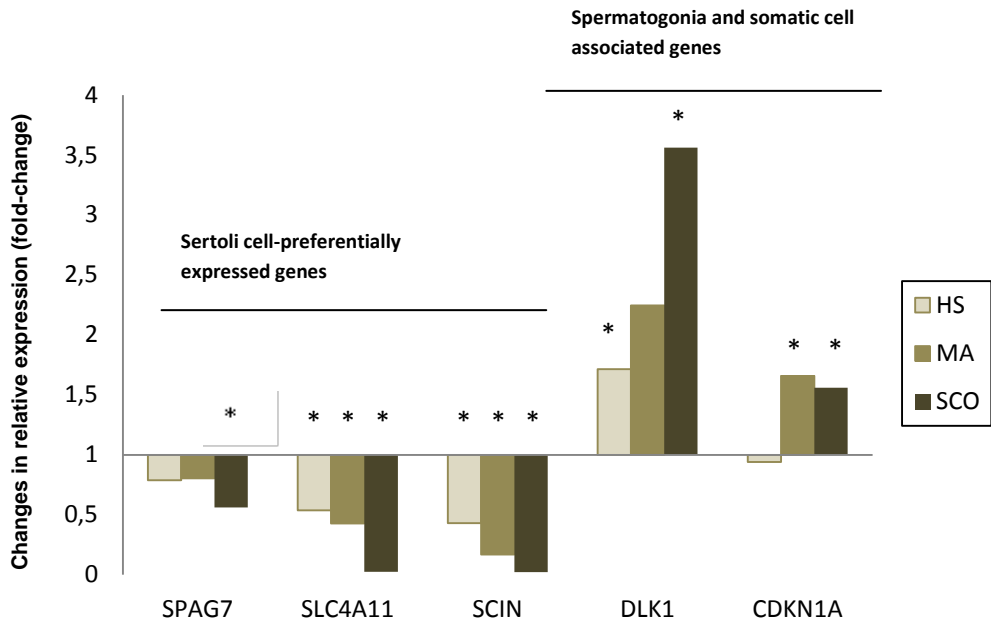


Peer Review

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

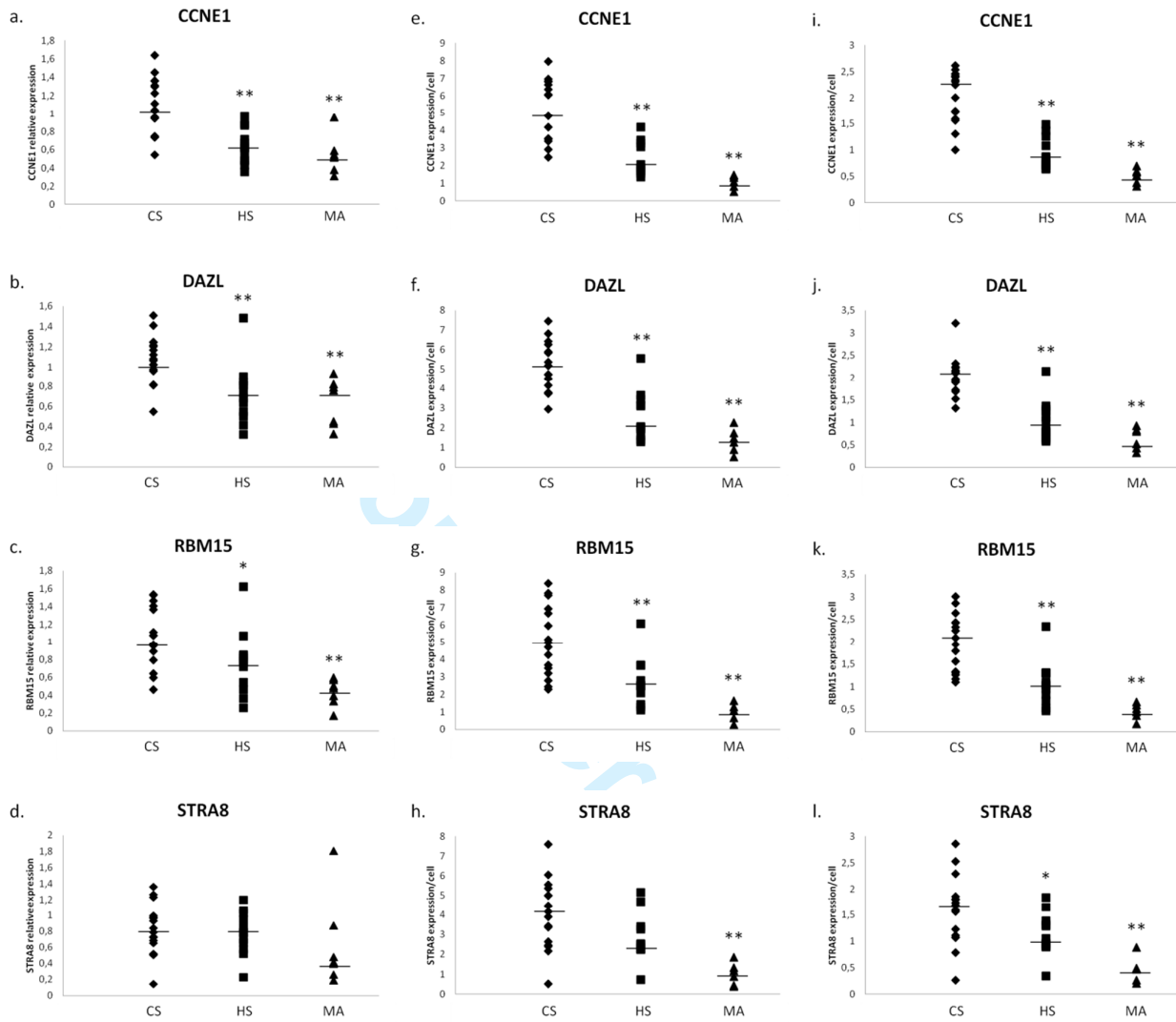
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2



Peer Review

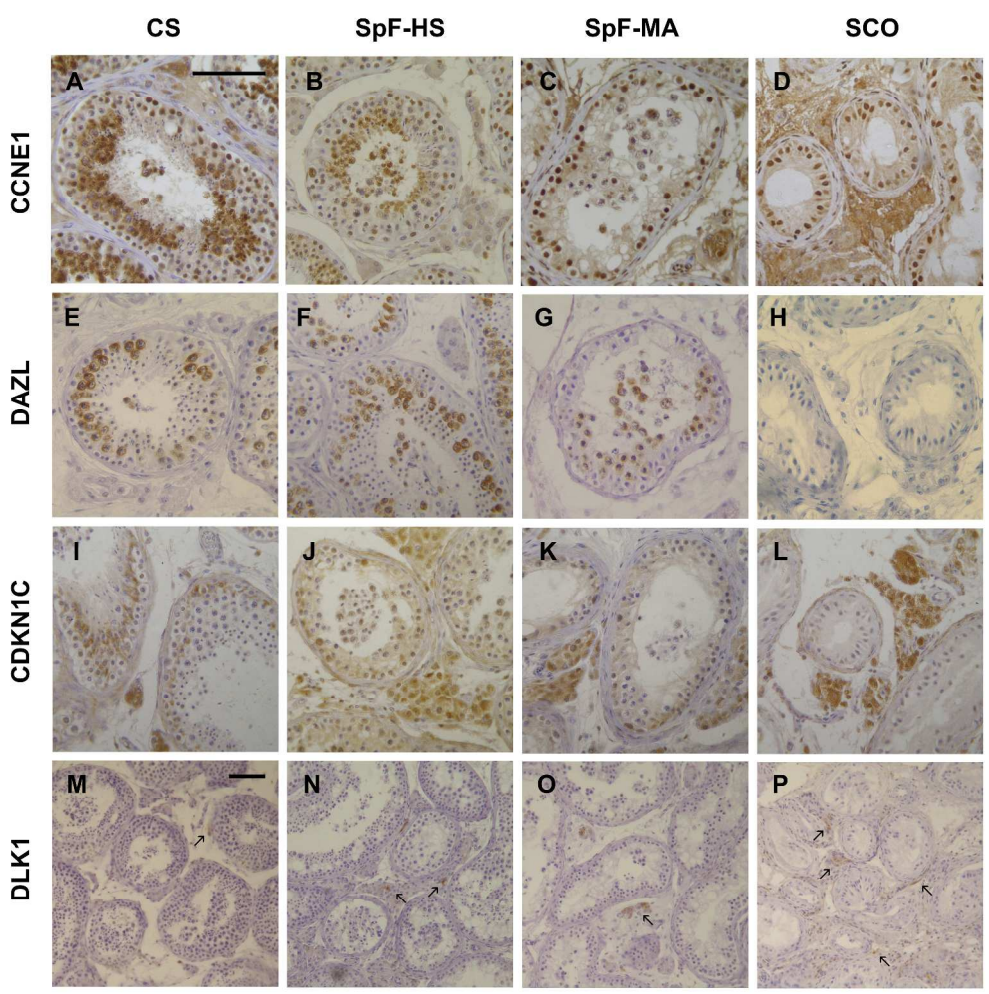
Figure 3



ew

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



ew

1
2
3 1
4
5 2 **ALTERED GENE EXPRESSION SIGNATURE OF EARLY STAGES OF THE**
6
7 3 **GERM LINE SUPPORTS THE PREMEIOTIC ORIGIN OF HUMAN**
8
9 4 **SPERMATOGENIC FAILURE**
10
11
12 5

13
14 6 **Running title:** Early germ line gene expression in impaired sperm production
15
16 7

17
18 8 **Authors:** Sandra Bonache^{1#}, Ferran Algaba², Eladio Franco³, Lluís Bassas⁴, Sara
19
20 9 Larriba^{1*}
21
22

23 10
24
25 11 ¹Human Molecular Genetics Group-IDIBELL, 08908 L'Hospitalet de Llobregat,
26
27 12 Barcelona, Spain

28
29 13 ²Pathology Department-Fundació Puigvert, Universitat Autònoma de Barcelona 08025
30
31 14 Barcelona, Spain

32
33 15 ³Urology Service-Hospital Universitari de Bellvitge, 08907 L'Hospitalet de Llobregat,
34
35 16 Barcelona, Spain

36
37 17 ⁴Laboratory of Seminology and Embryology, Andrology Service-Fundació Puigvert,
38
39 18 08025 Barcelona, Spain

40
41 19 # Current address: Oncogenetics Laboratory, University Hospital Vall d'Hebron, 08035
42
43 20 Barcelona, Spain
44
45

46
47 21
48
49 22
50
51 23 ***Correspondence address**

52 24 Tel. +34 93 260 74 25 ext. 73 38; Fax.+34 93 260 74 14; e-mail: slarriba@idibell.cat
53
54
55
56 25
57
58
59
60

26 **ABSTRACT (300 words)**

27 The molecular basis of spermatogenic failure (SpF) is still largely unknown.
28 Accumulating evidence suggests that a series of specific events such as meiosis, are
29 determined at the early stage of spermatogenesis. The present study aims to **reveal**
30 **assess the expression profile of** pre-meiotic genes **expression-changes** of infertile
31 testicular biopsies that might help to define the molecular phenotype associated with
32 human deficiency of sperm production. An accurate quantification of testicular mRNA
33 levels of genes expressed in spermatogonia was carried out by RT-qPCR in individuals
34 showing SpF due to germ cell maturation defects, Sertoli cell-only syndrome (SCO) or
35 conserved spermatogenesis (CS). Additionally, the gene expression profile of SpF was
36 compared with that of testicular tumour (GCT), which is considered to be a severe
37 developmental disease of germ cell differentiation. Protein expression from selected
38 genes was evaluated by immunohistochemistry.

39 Our results indicate **that** SpF is accompanied by differences in expression of
40 certain genes associated with spermatogonia in the absence of any apparent
41 morphological and/or numerical change of this specific cell type. In SpF **testicular**
42 samples, we observed down-regulation of genes involved in cell cycle (*CCNE1* and
43 *POLD1*), transcription and post-transcription regulation (*DAZL*, *RBM15* and *DICER1*),
44 protein degradation (*FBXO32* and *TM9SF2*) and homologous recombination in meiosis
45 (*MRE11A* and *RAD50*) which suggests that the expression of these genes is critical for a
46 **successful-proper** germ cell development. **Interestingly, a decrease in the *CCNE1*,**
47 ***DAZL*, *RBM15* and *STRA8* cellular transcript levels was also observed, suggesting that**
48 **the gene expression capacity of spermatogonia is altered in SpF contributing to an**
49 **unsuccessful sperm production.** Altogether, these data point to the spermatogenic
50 derangement being already determined at, or arising in, the initial stages of the germ

1
2
3 51 line, ~~although these cells maintain their transcriptional activity related to mitosis and~~
4
5 52 cell proliferation.

6
7 53

8
9 54 **Key words:** gene expression, testis, early stages of germ line, spermatogenic failure,
10
11 55 male infertility

12
13 56

14
15 57
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

58 INTRODUCTION

59 Spermatogenesis is a highly orchestrated developmental process by which
60 spermatogonia develop into mature spermatozoa. During the course of spermatogenesis
61 the three major forms of cell cycle are represented: mitosis of primitive spermatogonia;
62 two rounds of meiosis, from primary spermatocytes to haploid round spermatids; and
63 differentiation including structural and nuclear changes to generate mature spermatids
64 and spermatozoa. These processes are unique in male germ cell differentiation and
65 depend on precise, developmental stage- and germ cell type- specific gene expression.
66 However, the regulatory network that confers specific germ line gene expression in
67 mammals is not properly understood, especially at the mitotic and meiotic stages.
68 Understanding the regulatory step is essential for determining the molecular
69 requirements for the progression of spermatogenesis, and thus for understanding male
70 infertility which is often based on lack of replication of spermatogonia or meiotic
71 blockade.

72 Accumulating evidence suggests that a series of specific events during
73 spermatogenesis, such as meiosis and morphological changes, are determined at the
74 early stage of spermatogenesis. Spermatogonia, and specifically the type B
75 spermatogonia, should be an important preparation stage for meiosis. These data are
76 supported by the description of activation or up-regulation of many genes during this
77 specific germ cell stage (Guo *et al.*, 2004) and the generation of recombinant mouse
78 models of spermatogonia-expressed genes exhibiting severe defects in meiosis (Wang *et al.*
79 *al.*, 2001). Many of these genes codify germ cell specific proteins involved in
80 transcriptional or post-transcriptional regulation of gene expression (Wang *et al.*, 2001).

81 Furthermore, cellular interactions between germ line and somatic components of
82 the testicular seminiferous tubule, where spermatogenesis takes place, are essential to

1
2
3 83 achieve germ cell development, and thus for maintaining male fertility. The relevance of
4
5 84 these cellular interactions is supported by physiological events (for review, see
6
7 85 (Mclachlan *et al.*, 2007).
8

9
10 86 Some studies have used microarray technology to characterize the transcriptional
11
12 87 profile in germ and somatic cells at different steps of testicular development (Chalmel *et*
13
14 88 *al.*, 2007; Diederichs *et al.*, 2005; Namekawa *et al.*, 2006; Pang *et al.*, 2003; Schlecht *et*
15
16 89 *al.*, 2004; Schultz *et al.*, 2003; Sha *et al.*, 2002; Shima *et al.*, 2004). We have used
17
18 90 information from cDNA microarrays and mouse models to focus mainly on genes
19
20 91 specifically or preferentially expressed in immature germ cells in mammals, to
21
22 92 determine whether early gene expression changes are associated with subsequent
23
24 93 spermatogenic disorders ~~that could potentially lead and/or contribute to defective or~~
25
26 94 ~~absent spermatogenesis~~—and thus, male infertility. Additionally, the gene expression
27
28 95 profile of germ cell maturation failure was compared with that seen in testicular tumour,
29
30 96 which is considered to be a developmental disease of germ cell differentiation, in order
31
32 97 to give additional clues about the functional pathways involved in spermatogenic
33
34 98 derangement.
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 100 **MATERIALS AND METHODS**

4
5 101 **Subjects of study**

6
7 102 Our study recruited nineteen patients (range 30-49 yr) infertile due to spermatogenic
8
9 103 failure (SpF) at different germ cell stages, with a phenotype consistent with non-
10
11 104 obstructive azoospermia or severe oligozoospermia (<5 million sperm per mL). In
12
13 105 addition, fourteen patients with Sertoli cell-only syndrome (SCO) phenotype or germ
14
15 106 cell aplasia, were studied as controls of somatic cells (negative controls), and seventeen
16
17 107 infertile patients (range 32-50 yr), who were diagnosed with obstructive azoospermia
18
19 108 (as a consequence of congenital absence of vas deferens or previous vasectomy) and
20
21 109 showed conserved spermatogenesis (CS) were studied as gene expression controls of a
22
23 110 complete spermatogenic process (positive controls). Patients were selected on the basis
24
25 111 of the testicular histological pattern of >20 tubules from the same testicular section;
26
27 112 only testicular samples with a homogeneous histological pattern were included in the
28
29 113 study.

30
31
32
33
34 114 An extra group of individuals (n=5) diagnosed with germ cell tumour (GCT)
35
36 115 (range 29-46 yr) were analyzed. Two GCT samples were histologically classified as
37
38 116 carcinoma *in situ* (CIS) whereas the remaining three samples were classified as non-CIS
39
40 117 or GCT of advanced stages: one as classic seminoma, one as embryonal carcinoma and
41
42 118 one as mixed germ cell tumour (80% embryonal carcinoma; 20% classic seminoma).

43
44
45 119 Infertile individuals were selected from men referred for couple infertility to the
46
47 120 Andrology Service of the Fundació Puigvert, whereas GCT samples were recruited
48
49 121 from the Andrology Service of the Fundació Puigvert and the Urology Service of the
50
51 122 Hospital Universitari de Bellvitge. The study was approved by the Institutional Review
52
53 123 Board of both Centres, and all the participants signed a written informed consent.
54
55
56
57
58
59
60

1
2
3 124 The clinical procedures for infertile patients included medical history, physical
4
5 125 examination, semen analyses [performed in accordance with World Health Organization
6
7 126 guidelines (World Health Organization, 1999)] and hormonal study. The routine genetic
8
9 127 study for all non-obstructive samples included karyotype and analysis of chromosome Y
10
11 128 microdeletions, the latter performed according to the European guidelines (Simoni *et*
12
13 129 *al.*, 1999; Simoni *et al.*, 2004). Men with a chromosomal aberration or a Y-chromosome
14
15 130 microdeletion were not included in the study.
16
17
18
19

20 132 **Testicular samples**

21
22
23 133 Testicular biopsies from infertile men were obtained when necessary to confirm the
24
25 134 clinical diagnosis and for sperm retrieval (TESE) and cryopreservation purposes. Each
26
27 135 specimen was divided into three aliquots, one piece (≈ 10 -20 mg) was fixed in Bouin's
28
29 136 solution and reserved for histological analysis, a second aliquot (≈ 100 -200 mg)
30
31 137 processed for sperm extraction and the third (≈ 10 mg) was immediately transferred to
32
33 138 liquid nitrogen and stored at -80°C until analysis for gene expression experiments.
34
35

36
37 139 Referring to GCT, testicular samples were obtained directly after orchidectomy
38
39 140 and macroscopic pathological evaluation. For gene expression studies, one tissue
40
41 141 fragment was taken from the tumour portion of the testis and was immediately frozen at
42
43 142 -80°C .
44

45
46 143

47 144 **Histological analysis**

48
49
50 145 An assessment of spermatogenic status and the severity of the alteration were performed
51
52 146 after hematoxylin-eosin staining of paraffin samples from infertile patients (5- μm
53
54 147 sections) by quantification of specific germ cells (spermatogonia, spermatocytes I, round
55
56 148 spermatids and elongated spermatids) and Sertoli cells. The average number per tubule
57
58
59
60

1
2
3 149 was calculated after analysis of at least 15-20 cross-sectioned tubules per testis. A
4
5 150 modified Johnsen score (JS) count (Schulze *et al.*, 1999) was calculated on the basis of
6
7 151 the number of different cell types per tubule and infertile samples were classified as CS,
8
9 152 SpF-HS (hypospermatogenesis), SpF-MA (meiotic arrest) and SCO (Table 1).
10

11 153

14 154 **RNA extraction and cDNA synthesis**

15 155 Total RNA was obtained from the testicular biopsy using Absolutely RNA Miniprep Kit
16
17 156 (Stratagene, La Jolla, CA), following the manufacturer's instructions. The quality of
18
19 157 RNA [28S/18S ratio and RNA Integrity Number (RIN)] was assessed using the Agilent
20
21 158 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). Testicular RNA from
22
23 159 the five groups of study (SCO, SpF-MA, SpF-HS, GCT and CS) showed similar quality
24
25 160 values: both 28S/18S ratio and RIN presented no significant differences among them (p
26
27 161 = 0.056 and $p = 0.072$, respectively). Mean 28S/18S ratio value between all samples
28
29 162 (mean±SD) was 1.17±0.18 and mean RIN value was 7.77±0.62.
30
31

32
33
34 163 Single-stranded cDNA was obtained by reverse transcription (RT) of 500 ng of
35
36 164 RNA, using random primers and the High Capacity cDNA Reverse Transcription Kit
37
38 165 (AB, Foster City, CA). The resulting cDNA solution was aliquoted and stored at -20°C
39
40 166 until use.
41
42

43 167

45 168 **Gene expression quantification and statistical evaluation**

46
47 169 Quantitative real-time PCR (qPCR) assays were performed by means of the application
48
49 170 of the PCR arrays on micro fluidic cards (MFC), using 384-well TaqMan® Low
50
51 171 Density Arrays (TLDA) on an Applied Biosystems 7900HT Fast Real-Time PCR
52
53 172 System (AB, Foster City, CA). The 48-gene format MFC (47 experimental assays and 1
54
55 173 TLDA amplification control, *18S*) allowed simultaneous measurement of 34 target
56
57
58
59
60

1
2
3 174 genes that were selected based on a preferential expression in spermatogonia among
4
5 175 germ cells (n=26) and/or in Sertoli cells (n=8) -information obtained from cDNA
6
7 176 microarrays and mouse models bibliography data- (Supplemental Table 1), three marker
8
9 177 genes of the presence and/or function of spermatid (*PRMI*), Leydig (*INSL3*) and myoid
10
11 178 cells (*SI00A6*) and 10 potential reference genes. Genes and the corresponding assays on
12
13 179 demand used for the setup of the TLDA are listed in Supplemental Table 2. Selected
14
15 180 target genes are involved in different functional pathways (Supplemental Table 1 and
16
17 181 Supplemental Table 2).

18
19
20
21 182 Patient and control samples were always analysed as paired samples in the same
22
23 183 analytical run in order to exclude between-run variations. Additionally, a calibrator
24
25 184 sample was included in all the plates to compare the change in expression of a nucleic
26
27 185 acid sequence against the expression in all samples in the same study. Real-time PCR
28
29 186 data (Ct values) were pre-processed and stored in SDS 2.2 software (AB, Foster City,
30
31 187 CA).

32
33
34 188 Expression stability of the candidate reference gene/s was calculated with the
35
36 189 GeNorm software (Vandesompele *et al.*, 2002), in order to select the most stable
37
38 190 reference genes and improve normalization of target genes. GeNorm software calculates
39
40 191 the gene expression stability value M of multiple candidate genes as the average pair-
41
42 192 wise variation of a particular gene compared with all other candidate reference genes.
43
44 193 Lower M values indicate genes with less expression variation among samples.
45
46 194 Therefore, target gene expression was calculated relative to the expression of *PGKI* and
47
48 195 *PGMI* reference genes for SpF and control samples, whereas *PGKI* and *PPIA*
49
50 196 combination was selected as the most appropriate for GCT and controls. They showed
51
52 197 no statistical differences in absolute expression levels between groups (Kruskal-Wallis
53
54
55
56
57
58
59
60

1
2
3 198 test) (Supplemental Figure 1) and low M-value (GeNorm software (~~Vandesompele et~~
4
5 199 ~~al., 2002~~) indicating stable expression among samples.
6

7 200 Thus, raw data (Ct values) were normalised to the two reference genes and
8
9 201 relative quantification (RQ) values were calculated using the qBase program
10
11 202 (Hellemans *et al.*, 2007). and the $2^{-\Delta\Delta Ct}$ strategy. The Mann-Whitney U test was used to
12
13 203 evaluate differences in relative expression of target genes in each patient group or
14
15 204 subgroup compared with controls. Multiple test adjustment was applied by using
16
17 205 Bonferroni correction.
18
19

20 206 Pearson product-moment correlation coefficients were calculated to determine
21
22 207 the correlation between the expression ratios of the target genes and the various
23
24 208 histological parameters in patient groups and controls.
25
26

27 209 All statistical analyses were performed using the SPSS software version 12
28
29 210 (Lead Technologies, Chicago, USA)
30
31

32 211

33 212 **Immunohistochemistry**

34
35 213 Tissue sections were prepared from Bouin-fixed, paraffin-embedded fragments of
36
37 214 testicular biopsies. For this study, the following commercially available polyclonal
38
39 215 rabbit antibodies were used: DAZL, HPA019777, Sigma-Aldrich, Inc; CCNE1 (C-19):
40
41 216 sc-198, Santa Cruz Biotechnology, Inc; CDKN1C (C-20):sc-1040, Santa Cruz
42
43 217 Biotechnology, Inc; DLK1 (H-118): sc-25437, Santa Cruz Biotechnology, Inc.
44
45

46
47 218 Immunohistochemistry was performed using the Dako EnVision+ kit (DAKO,
48
49 219 Hamburg, Germany) in conjunction with the Dako autostainer, according to the
50
51 220 instructions provided by the manufacturer.
52

53
54 221 Endogenous peroxidase was quenched by incubation in 0.5% hydrogen
55
56 222 peroxide. Dilutions of primary antibodies were adjusted at 1:300 to 1:100 to optimize
57
58
59
60

1
2
3 223 the results. Secondary goat anti-rabbit antiserum was coupled to a labelled polymer-
4
5 224 HRP, and staining was done with DAB and haematoxylin-eosin. Incubation with non
6
7 225 immune serum was used as a negative control (data not shown). Stained sections were
8
9 226 evaluated in bright-field microscopy (Axioskop 40, Zeiss, Göttingen, Germany) and
10
11 227 images captured with a Nikon Coolpix 5400 digital camera. The immunoexpression of
12
13 228 proteins was determined on tissue samples from at least six different individuals
14
15 229 showing each of the spermatogenic phenotypes.
16
17
18
19
20
21 230
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60 231

For Peer Review

1
2
3 232 **RESULTS**

4
5 233 **Quantitative determination of spermatogenic status in defective spermatogenesis**

6
7 234 Paraffin-embedded testicular specimens from infertile patients and controls were
8
9 235 available for histological quantification. Considering the heterogeneity of human
10
11 236 testicular pathologies, we needed a detailed definition of the spermatogenic status of
12
13 237 each sample in order to acquire high quality data related to germ cell specific
14
15 238 transcriptional changes. To this end, the average number of Sertoli cells and specific
16
17 239 germ cell types per tubule were determined in haematoxylin and eosin stained testicular
18
19 240 sections and additionally the JS value was calculated (Table 1).

20
21
22
23 241 Using this strategy we confirmed the diagnosis of SCO (JS score 1-2) and CS (JS
24
25 242 score 9-10) phenotypes. With respect to SpF patients, seven of them presented
26
27 243 maturation arrest at primary spermatocyte level (SpF-MA) (JS score 4-6) and twelve
28
29 244 presented hypospermatogenesis (SpF-HS) (JS score 7-8) (Table 1).

30
31
32 245 Round and elongated spermatids were absent in 4 out of 7 SpF-MA samples
33
34 246 whereas the other three had very low values: round spermatids ranged between 2-4 per
35
36 247 tubule and elongated spermatids ranged between 1-2 per tubule suggesting a 80-90% of
37
38 248 meiotic arrest. This result confirms that only histological phenotypes with a defined and
39
40 249 highly homogeneous pattern of individual tubules were included in the study.

41
42
43 250 The number of Sertoli cells showed a **near** two-fold increase in the group SCO
44
45 251 (21.03 ± 7.00) compared with the CS samples (13.69 ± 3.12). Interestingly, the number
46
47 252 of spermatogonia and Sertoli cells showed no significant differences ($p = 0.318$ and p
48
49 253 $=0.447$ respectively) among SpF-MA, SpF-HS and CS groups (Table 1).

50
51
52 254

53
54 255 **Gene expression related to impaired spermatogenesis**

55
56 256 *PRMI expression confirms the histological phenotype*
57
58
59
60

1
2
3 257 In order to confirm the histological quantification of samples, we first analyzed the
4
5 258 expression of *PRMI*, the marker gene for the presence of spermatids. As expected,
6
7 259 negligible transcript level values were found in SCO and 100% MA samples; very low
8
9 260 values were determined in incomplete MA samples and decreased levels in HS samples
10
11 261 when compared with CS controls, showing a fold decrease in expression of 2.24×10^4
12
13 262 (SCO), 42.59 (SpF-MA) and 4.45 (SpF-HS) ~~fold decrease in expression compared with~~
14
15
16 263 ~~controls~~ ($p < 0.001$). The absence of quantifiable *PRMI* expression values in GCT
17
18 264 samples confirmed the absence of mature germ cells.

19
20
21 265 Our *PRMI* gene expression results consistently agreed with the histological
22
23 266 diagnosis of all samples included in the study.

24
25 267

26
27 268 *Expression behaviour of Leydig and myoid cells in spermatogenic disorders*

28
29 269 Referring to the expression of the marker genes for the presence and/or function of
30
31 270 Leydig (*INSL3*) and myoid cells (*SI00A6*), no statistically significant differences were
32
33 271 found between SpF-HS and CS groups ($p = 0.245$ and $p = 0.059$ respectively) and between
34
35 272 SpF-MA and CS samples ($p = 0.065$ and $p = 0.075$ respectively), however these genes
36
37 273 were significantly over-expressed in SCO group when compared with CS samples
38
39 274 ($p = 0.001$ and $p = 0.000$ respectively), probably attributable to the absence of germ cells;
40
41 275 since total testis samples are analysed, the loss of germ cells ~~changes the~~ ~~enrich the~~
42
43 276 relative contribution of the remaining somatic cells in SCO phenotype.

44
45
46 277 GCT samples also showed no statistically significant differences in *INSL3* and
47
48 278 *SI00A6* expression when compared with the CS control group ($p = 0.401$ and $p = 0.542$
49
50 279 respectively)

51
52
53 280
54
55
56
57
58
59
60

1
2
3 281 *Most of the spermatogonia-transcriptionally associated genes analyzed show a*
4
5 282 *preferential germ line expression*

6
7 283 Spermatogonia-associated genes included in the study have been previously described
8
9 284 as having a preferential gene expression in the early germ line stages among germ cells
10
11 285 (Supplemental Table 1). The assessment of expression of these genes in the SCO
12
13 286 phenotype would provide important new information about whether these genes ~~are-can~~
14
15 287 ~~be~~ expressed also in somatic cells of the testis.

16
17
18 288 Selective or preferential germ cell expression was confirmed for 21 out of the 26
19
20 289 spermatogonia-related genes of our study. First, the negligible transcript level values
21
22 290 found for *CCNE1*, *DAZL*, *RBM15* and *STRA8* in complete SCO samples supported their
23
24 291 selective germ cell lineage gene expression. Furthermore, statistically significant
25
26 292 reduced transcript values in SCO compared with CS suggested a preferential germ cell
27
28 293 expression: *ATM*, *BARD1*, *CCND1*, *CCNF*, *DICER1*, *E2F3*, *FBXO32*, *c-KIT*, *MRE11A*,
29
30 294 *POLA1*, *POLD1*, *RAD50*, *SIRT6* and *TM9SF2* presented a very significant fold-change
31
32 295 decrease in SCO ranging from 1.46 to 6.10 ($p < 10^{-4} - 10^{-6}$) whereas *BAX*, *CDKN1C* and
33
34 296 *XPA* showed less marked decrease, with a fold-change ranging from 1.22 to 1.55
35
36 297 ($p < 0.05$) (Supplemental Table 3).

37
38
39 298 However, for the five remaining spermatogonia-associated genes studied, a
40
41 299 preferential expression in the germ cell lineage could not be demonstrated. *CDKN1A*
42
43 300 and *DLKI* were found significantly over-expressed in SCO samples when compared
44
45 301 with CS controls ($p = 0.003$ and $p < 0.002$ respectively), showing a fold-change increase
46
47 302 of 1.77 and 3.99 respectively. Also, *BMPRIA*, *SMAD3* and *VEGFA* were not
48
49 303 differentially expressed between SCO and CS samples ($p > 0.05$) (Supplemental Table
50
51 304 3). Taken together, these results indicate a substantial expression of these five genes in
52
53 305 testicular somatic cells.
54
55
56
57
58
59
60

306

307 *Differential gene expression in defective and conserved spermatogenesis*

308 We then looked for differences in target gene expression between SpF samples and CS
309 controls.

310 For the 21 spermatogonia-associated genes preferentially expressed in the germ
311 line, we identified 13 differentially expressed genes and grouped them into three gene
312 clusters by their expression behaviour (Figure 1). In cluster I, both SpF-MA and SpF-
313 HS phenotypes were associated with notable significant decreases in transcript levels of
314 *BARD1*, *CCNE1*, *DAZL*, *FBXO32*, *RBM15* and *TM9SF2* genes ($p < 0.002$), the reduction
315 of expression being more pronounced in the SpF-MA phenotype. Cluster II contained
316 genes significantly decreased in the SpF-HS ($p < 0.05$) but not in the SpF-MA phenotype
317 (*CCNF* and *E2F3*). Cluster III included all those genes differentially expressed in SpF-
318 MA but not in SpF-HS phenotype: *DICER1*, *MRE11A*, *POLD1* and *RAD50* expression
319 levels were decreased whereas *CCND1* was increased in the SpF-MA ($p < 0.05$) when
320 compared with controls. *CDKN1C* presented a similar pattern of gene expression
321 deregulation to *CCND1*, although the difference in expression was not statistically
322 significant (CS: 0.51 ± 0.15 — 0.75 ± 0.18 ; HS: 0.61 ± 0.54 0.80 ± 0.31 $p = 0.394$; MA:
323 1.04 ± 0.73 0.98 ± 0.21 $p = 0.172$) (Supplemental Table 3).

324 Interestingly, significant positive correlation coefficients (Pearson's correlation
325 $r \geq 0.6$; $p < 0.0001$) were found between the number of elongated spermatids and the
326 transcription levels of genes from cluster I (Supplemental Table 4.A), suggesting that
327 these changes in gene expression could be of physiological relevance.

328 In regard to the spermatogonia and somatic cell-expressed genes, *CDKN1A* was
329 significantly increased in the SpF-MA and SCO samples ($p < 0.05$). *DLK1* was found
330 highly over-expressed in SpF samples, although the differences were only statistically

1
2
3 331 significant for the SpF-HS and SCO ($p=0.001$) and not the SpF-MA phenotype
4
5 332 (probably due to a high standard deviation value) when compared with controls (Figure
6
7 333 2; Supplemental Table 3).

8
9 334 When considering the Sertoli cell-preferentially expressed genes, *SCIN* and
10
11 335 *SLC4A11* were found very significantly decreased in MA and HS samples when
12
13 336 compared with CS. SCO samples, although presenting an increased number of Sertoli
14
15 337 cells, showed reduced levels of *SPAG7*, *SCIN* and *SLC4A11* compared with CS controls
16
17 338 ($p<0.006$) (Figure 2; Supplemental Table 3). The expression of *FASLG* mRNA in SpF
18
19 339 samples was increased compared with CS, although no statistically significant changes
20
21 340 in the average of *FASLG* expression were observed among groups due to high standard
22
23 341 deviation values in the infertile groups (CS control: 1.10 ± 0.71 ; HS: 1.79 ± 1.19 ; MA:
24
25 342 3.44 ± 2.87).

26 27 343 28 29 344 *Reduced cellular expression levels of germ cell-specific genes in SpF*

30
31 345 We additionally analyzed the transcript levels per cell of spermatogonia-associated
32
33 346 genes with a selective germ cell expression, in SpF subgroups compared to CS controls
34
35 347 in order to exclude the differences in gene expression due to changes in testicular
36
37 348 cellularity and to determine whether transcript level per cell is also altered in SpF.
38
39 349 Selective germ cell expression of *CCNE1*, *DAZL*, *RBM15* and *STRA8* was previously
40
41 350 confirmed as negligible transcript level values were found in SCO samples as
42
43 351 previously described, furthermore, *DAZL* and *STRA8* were previously described to be
44
45 352 expressed in spermatogonia but not in somatic tissues (Wang *et al.*, 2001). Values of
46
47 353 transcript amount per cell, in arbitrary units, were obtained for each testicular sample by
48
49 354 dividing the *CCNE1*, *DAZL*, *RBM15* and *STRA8* expression values by either the
50
51 355 proportion of the spermatogonia (Figure 3e, f, g and h), beingas it is the germ cell stage
52
53
54
55
56
57
58
59
60

1
2
3 356 that predominantly expresses *CCNE1*, *DAZL*, *RBM15* and *STRA8* or by the proportion
4
5 357 of the spermatogonia plus spermatocytes (Figure 3i, j, k and l), ~~being the latter as~~ the
6
7 358 meiotic germ cells ~~stage that~~ potentially expresses these genes although at much lower
8
9 359 levels (GermSAGE, <http://germsage.nichd.nih.gov>; GermOnline
10
11 360 <http://www.germonline.org>), present in a seminiferous tubule of the sample. Significant
12
13 361 differences in cellular transcript levels were additionally found for *CCNE1*, *DAZL*,
14
15 362 *RBM15* genes between SpF-HS patients and controls and between SpF-MA patients and
16
17 363 controls, when considering either the proportion of spermatogonia or the proportion of
18
19 364 spermatogonia and spermatocytes in the tubule ($p \leq 0.002$). Interestingly, cellular
20
21 365 transcript levels for *STRA8* were found statistically decreased in SpF-MA when
22
23 366 compared with CS when considering either the proportion of spermatogonia or the
24
25 367 proportion of spermatogonia and spermatocytes in the tubule ($p \leq 0.002$). These results
26
27 368 suggest that the decreased tissular expression levels in SpF are not attributable to a
28
29 369 decreased number of spermatogonia or spermatocytes in the tubule but to a reduced
30
31 370 ~~expression capability of the number of transcripts in~~ immature germ cells. Furthermore,
32
33 371 ~~the decreased cellular *STRA8* expression levels observed in SpF-MA suggest that the~~
34
35 372 ~~number of genes whose expression is altered in immature germ cells might be higher~~
36
37 373 ~~than that initially observed in the whole tissue with meiotic arrest.~~

374 Strikingly, the decrease cellular expression observed in patients was
375 accompanied by a higher severity in spermatogenic impairment, and cellular expression
376 levels of *CCNE1*, *DAZL*, *RBM15* and *STRA8* genes were highly significantly and
377 positively correlated with the number of elongated spermatids in the tubule (Pearson's
378 correlation range; $r=0.68-0.87$; $p < 0.0001$) (Supplemental Table 4.B.)

379

380 *Gene expression pattern in Germ cell tumours*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 381 Relative expression values of fourteen spermatogonia preferentially- expressed genes
4
5 382 were found to have extremely significant differences in expression between GCT and
6
7 383 CS samples: *ATM*, *BARD1*, *CCNE1*, *CDKN1C*, *DAZL*, *DICER1*, *E2F3*, *FBXO32*,
8
9 384 *MRE11A*, *RAD50*, *RBM15*, *SIRT6*, *STRA8* and *TM9SF2* ($p<0.002$). Less consistent
10
11 385 statistical differences were found for four additional genes *BAX*, *CCND1*, *POLD1* and
12
13 386 *XPA* ($p<0.05$). All these differentially expressed genes, with the exception of *BAX*, were
14
15
16 387 found to be under-represented in GCT samples compared with controls (Table 2;
17
18 388 Supplemental Table 3).

19
20
21 389 Regarding the somatically expressed genes; ten genes were found to have
22
23 390 statistical differences in expression: *AMHR2*, *BCL3*, *SCIN*, *SMARCA1*, *SOX9*, *SPAG7*,
24
25 391 *VEGFA* ($p<0.002$) *BMPRIA*, *CDKN1A* and *FASLG* ($p<0.05$). *FASLG* was found to be
26
27 392 over-expressed, whereas the other nine genes were under-expressed, in GCT samples
28
29 393 compared with controls (Table 2; Supplemental Table 3).

30
31
32 394 Some differences in expression behaviour were found when GCT samples were
33
34 395 divided into CIS and non-CIS samples (Supplemental Table 3): over-expression of the
35
36 396 *KIT* gene (an established marker for early-stage GCT) in CIS samples consistently
37
38 397 agreed with the histological diagnosis of samples. No difference in expression was
39
40 398 found for *BARD1*, *BMPRIA*, *CCND1*, *CDKN1A*, *CDKN1C*, *SMAD3* and *XPA* in CIS
41
42 399 samples compared with CS controls, whereas their transcript values were found highly
43
44 400 decreased in non-CIS samples ($p<0.002$). *DLK1* was significantly over-expressed in
45
46 401 CIS samples ($p=0.012$), contrary to the marked under-expression found in non-CIS
47
48 402 samples ($p<0.002$). Thus, this set of genes is, somehow, associated with different stages
49
50 403 in tumour progression.

51
52
53
54 404

55
56 405 **Expression signature by functional categories**
57
58
59
60

1
2
3 406 In order to obtain some clues about the functional pathways that are affected in **testis**
4
5 407 **with** SpF, genes were grouped into functional clusters according to the process they are
6
7 408 involved in: cell proliferation, apoptosis/cell cycle, meiosis, DNA repair, transcription
8
9 409 regulation, post-transcriptional regulation and degradation. The functional expression
10
11 410 signature was compared with that obtained from the GCT samples (Supplemental Table
12
13 411 3).

14
15
16 412 Some of the spermatogonia-preferential transcripts differentially represented
17
18 413 between the SpF subgroups and controls encode proteins involved in the regulation of
19
20 414 the mitotic and meiotic cell cycle (cyclins, cyclin –dependent-kinase-inhibitors and
21
22 415 DNA polymerases). *CCNE1* and *POLD1* were significantly under-expressed, this being
23
24 416 particularly noticeable for *CCNE1*, whereas *CCND1* was over-expressed in the MA
25
26 417 samples. In GCT only the *CCNE1* transcript value of this group of genes was
27
28 418 significantly decreased compared with controls. Furthermore, we observed decreased
29
30 419 expression levels of several genes that encode for putative RNA-binding proteins (such
31
32 420 as the germ cell specific *DAZL* and *RBM15*), other proteins essential for production of
33
34 421 miRNAs (such as *DICER1*), proteins involved in the protein degradation pathway
35
36 422 (*FBXO32* and *TM9SF2*) and also proteins implicated in the homologous recombination
37
38 423 in meiosis (*MRE11A* and *RAD50*) in both SpF and GCT groups.

39
40
41 424 There were no differences in expression of genes such as *BMPRIA*, *c-KIT* and
42
43 425 *VEGFA* (cell population proliferation), *BAX* gene (apoptosis), *ATM*, *SIRT6* and *XPA*
44
45 426 (DNA repair genes) observed between SpF and CS groups, in contrast to the observed
46
47 427 expression alteration of these genes in GCT.

48
49
50 428 The expression of all the Sertoli-specific expressed genes, with the exception of
51
52 429 *SLC4A11* was significantly affected in the GCT samples, whereas only those genes
53
54
55
56
57
58
59
60

1
2
3 430 involved in cell cycle/apoptosis, *SLC4A11* and *SCIN*, were significantly decreased in
4
5 431 the SpF meiotic altered samples.
6

7 432

8
9 433 **Protein expression**

10 434 We sought to determine whether the changes in transcript levels would correlate with
11
12 435 modifications at the protein level. At the same time we aimed to determine whether the
13
14 436 suggested alterations in gene expression affected expression levels of encoded proteins
15
16 437 in the germ line. We focused first on robust differences in expression levels which
17
18 438 might be discerned by immunohistochemistry and chose *CCNE1* and *DAZL* as examples
19
20 439 of spermatogonia-associated genes specifically expressed in the germ line. Secondly, we
21
22 440 selected proteins whose coding genes presented an upregulated expression in SpF
23
24 441 samples such as *DLK1* and *CDKN1C* (Figure 4).
25
26
27
28

29 442 *CCNE1* immunostaining (A-D; Figure 4) in CS samples was observed mainly in
30
31 443 both the nucleus and the cytoplasm of postmeiotic secondary spermatocytes/round
32
33 444 spermatids, and in the in nucleus of spermatogonia. Sertoli cells showed less intense
34
35 445 *CCNE1* expression in the nucleus; some staining of the cytoplasm of interstitial Leydig
36
37 446 cells and Sertoli cells was also seen, which tended to increase in samples with MA and
38
39 447 SCO; in the germ line of SpF samples *CCNE1* expression was exclusively detected in
40
41 448 the spermatogonia and not detected in primary spermatocytes and elongated spermatids,
42
43 449 when present. The *DAZL* protein (E-H, Figure 4) was almost exclusively present in the
44
45 450 cytoplasm of primary pachytene spermatocytes in CS samples and therefore expression
46
47 451 was negative in testis with SCO. Overall, immunoexpression in testis sections of both
48
49 452 *CCNE1* and *DAZL* protein decreased within seminiferous tubules, in the germ line, as
50
51 453 spermatogenic damage progressed, showing good correlation with RNA expression.
52
53
54
55
56
57
58
59
60

1
2
3 454 CDKN1C (I-L; Figure 4) in CS samples was seen in the cytoplasm of Sertoli
4
5 455 cells and Leydig cells; the immunostaining for Leydig cells was more intense in MA
6
7 456 and SCO samples; peritubular cells were positive in SCO; however in SpF samples
8
9 457 some spermatogonia and primary spermatocytes also exhibited moderate expression of
10
11 458 CDKN1C, leading to a global increase in CDKN1C expression in biopsies with SpF.
12
13 459 The immunoreactivity of DLK1 (M-P; Figure 4) in CS samples seemed to be restricted
14
15 460 to the cytoplasm of a few Leydig cells, which were more frequently stained in MA, and
16
17 461 more so in SCO suggesting a relevant contribution of Leydig cells to DLK1 expression.
18
19 462 Overall, immunoexpression of both CDKN1C and DLK1 protein in testis sections,
20
21 463 preferentially in somatic cells for DLK1, increased as spermatogenic damage
22
23 464 progressed, showing good correlation with RNA expression.
24
25
26
27 465
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 466 **DISCUSSION**

4
5 467 The aim of this study was to ~~reveal~~-assess early testicular transcriptional changes that
6
7 468 ~~are~~-could be involved in human severe deficiency of sperm production. We focused our
8
9 469 attention mainly on the expression profile of premeiotic germ cells as this is a key step
10
11 470 in male germ cell maturation. The accurate quantification of testicular mRNA levels in
12
13 471 SpF by RT-qPCR experiments led to the identification of differences in expression of
14
15 472 certain genes associated with spermatogonia in the absence of any apparent
16
17 473 morphological and/or numerical change of this specific cell type.

18
19
20 474 The gene expression profile in SpF~~shown by the altered genes~~ can be used as a
21
22 475 basis for identification of candidate genes ~~involved in the aetiology of~~that contribute to
23
24 476 spermatogenic impairment. SpF and GCT expression signature comparison could
25
26 477 additionally give some clues about the molecular mechanisms underlying the origin of
27
28 478 these alterations. Our data indicate that, in the SpF patients, a large proportion of
29
30 479 spermatogonia-preferentially expressed genes exhibited reduced testicular expression
31
32 480 levels when compared with CS individuals. As expected, the number of genes whose
33
34 481 expression was altered ~~as well as the magnitude of increase or decrease of gene~~
35
36 482 ~~expression~~ in GCT was even higher, possibly related to the fact that the germ line in
37
38 483 testicular tumour has undergone a dedifferentiation process representing an extreme
39
40 484 situation of gene expression deregulation of spermatogenic impairment. Interestingly,
41
42 485 gene expression signatures of both phenotypes, SpF and GCT, share some aberrant
43
44 486 patterns of gene expression supporting the idea that the participation of these genes is
45
46 487 essential for physiological germ cell development. In contrast, other genes are
47
48 488 differentially affected in both pathological groups suggesting that they contribute to the
49
50 489 phenotype and could be used as potential molecular markers.
51
52
53
54
55
56
57
58
59
60

1
2
3 490 The cellular complexity of the testis is an inherent problem which should be
4
5 491 taken into account when studying gene expression profiles in this organ. As the
6
7 492 pathological seminiferous tubules lack germ cells to varying degrees, changes in gene
8
9 493 expression at the tissue level can reflect changes in the capability for transcribing the
10
11 494 mRNA in a specific cell type as well as changes in the cell type composition in
12
13 495 pathological testis. The absence of significant differences in the spermatogonia and
14
15 496 Sertoli cell number among the SpF-MA, SpF-HS and CS groups in our study indicates
16
17 497 that transcript levels cannot be attributed to the presence or absence of these specific
18
19 498 cell types. Additionally, we are aware of the fact that the levels of different mRNAs
20
21 499 could change as the proportion of immature germ cells is different in tubules in
22
23 500 conserved spermatogenesis, in which all stages of germ cell are present, and in those in
24
25 501 maturation arrest, where only some stages are present. In this condition, we assume that
26
27 502 there would be less mRNA from most spermatogonia-expressed genes in CS testis than
28
29 503 in SpF testis. Our study shows that most of the differentially expressed genes showed
30
31 504 higher expression in testis with conserved spermatogenesis suggesting that the reduction
32
33 505 of target genes could not be attributable to either the spermatogonia cell number or to
34
35 506 different proportion of this germ cell stage in the tubule, but to real differences in the
36
37 507 expression capability of the cell. Furthermore, we describe reduced cellular expression
38
39 508 levels of four germ line specific genes in SpF samples supporting this premise.
40
41
42
43
44

45 509 The spermatogonia-related genes whose transcripts were differentially
46
47 510 represented between SpF subgroups and controls included genes involved in specific
48
49 511 functional pathways. A first group of genes encodes proteins involved in the regulation
50
51 512 of the mitotic and meiotic cell cycle such as cyclins, cyclin-dependent-kinase-inhibitors
52
53 513 and DNA polymerases indicating that the regulation of this specific functional gene
54
55 514 cluster in the initial stages of spermatogenesis is critical for further differentiation and
56
57
58
59
60

1
2
3 515 meiosis of germ cells. Additionally, decreased expression levels of several genes
4
5 516 encoding putative RNA-binding proteins (such as the germ cell specific *DAZL* and
6
7 517 *RBM15*), other proteins essential for production of miRNAs such as *DICER1*, and also
8
9 518 proteins involved in the protein degradation pathway (*FBXO32* and *TM9SF2*) in both
10
11 519 SpF and GCT groups, underlie the complexity of post-transcriptional control in
12
13 520 proliferation and differentiation of germ cells. *MRE11A* and *RAD50*, involved in
14
15 521 homologous recombination in meiosis, also showed altered expression in meiotic
16
17 522 blockade. Unexpectedly, we observed no difference in testicular expression for *STRA8*,
18
19 523 a well known gene involved in meiotic cell cycle, participating in chromosome pairing
20
21 524 and in the process that leads to stable commitment to the meiotic cycle (Mark *et al.*,
22
23 525 2008), in SpF group nor in SpF-MA subtype samples when compared with controls.
24
25 526 ~~The reduced transcript levels of other genes expressed in spermatogonia as the ones~~
26
27 527 ~~involved in piRNA processing machinery such as *PIWIL2* and *TDRD1* in SpF, as we~~
28
29 528 ~~recently described (Heyn *et al.*, 2012) further supports the role of a proper gene~~
30
31 529 ~~expression in early germ line stages for a successful sperm production.~~ Interestingly, an
32
33 530 additional statistically significant reduction in the expression levels of germ cell-
34
35 531 specific genes per spermatogonia was observed in MA and HS when compared to CS
36
37 532 samples, demonstrated for *CCNE1*, *DAZL* and *RBM15* genes, being more pronounced in
38
39 533 the MA phenotype suggesting that the expression capacity in immature germ cells
40
41 534 correlates with the severity of testicular damage. More intriguingly, a decreased cellular
42
43 535 expression level was even observed for *STRA8* in MA pattern, suggesting that in MA
44
45 536 phenotype, premeiotic cellular expression could be affected ~~in a lesser extend~~ for a
46
47 537 larger number of genes. The remarkable correlation coefficient between the *CCNE1*,
48
49 538 *DAZL*, *RBM15* and *STRA8* transcript levels per cell and the number of elongated
50
51 539 spermatids in the testicular tubule additionally underlines the determinant role of
52
53
54
55
56
57
58
59
60

1
2
3 540 premeiotic *CCNE1*, *DAZL*, *RBM15* and *STR48* expression in the progression of the
4
5 541 spermatogenic process.
6

7
8 542 In previous studies, the reduction of gene expression in SpF patients has been
9
10 543 mainly attributed to the decreased number of germ cells that specifically express the
11
12 544 gene of interest (i.e. significantly lower *DAZL* mRNA concentrations were previously
13
14 545 found in testes of non-obstructive azoospermic men (Lin *et al.*, 2001)), although germ
15
16 546 cell quantification was not properly performed. Here, we demonstrate that the changes
17
18 547 in expression observed among groups could be not exclusively explained by the
19
20 548 immature germ cell number but the contribution of the reduced cellular expression of
21
22 549 i.e. *DAZL* mRNA in spermatogenic impairment should be also taken into account.
23

24
25 550 Protein data on nonobstructive testicular tissue corroborate our mRNA
26
27 551 expression results: *CCNE1* and *DAZL* protein decreased within seminiferous tubules, in
28
29 552 the germ line, as spermatogenic damage progressed. The reduced transcript levels of
30
31 553 other genes expressed in spermatogonia as the ones involved in piRNA processing
32
33 554 machinery such as *PIWIL2* and *TDRD1* in SpF, as we recently described (Heyn *et al.*,
34
35 555 2012) further supports the role of a proper gene expression in early germ line stages for
36
37 556 a successful sperm production.
38

39
40 557 The expression levels of genes participating in cell population proliferation,
41
42 558 mitochondria-mediated apoptosis and DNA repair (assuring the maintenance of genome
43
44 559 integrity) are in general maintained in SpF, **contrary than in GCT**, supporting the idea
45
46 560 that they are **similarly normally**-processed in meiotic derangement **and in conserved**
47
48 561 **spermatogenesis**. However, there is one exception possibly due to other regulatory
49
50 562 pathways: the levels of spermatogonia-specific full-length *BARD1* transcript (Feki *et*
51
52 563 *al.*, 2004; Irminger-Finger *et al.*, 2001), involved in germ cell apoptotic events, could be
53
54
55
56
57
58
59
60

1
2
3 564 repressed in SpF by the high levels of FSH in spermatogenic failure samples (Feki *et*
4
5 565 *al.*, 2004).

6
7 566 Although the transcription profile of spermatogonia-associated genes in SpF is
8
9
10 567 different from that in GCT, some functional clusters are affected in both phenotypes:
11
12 568 genes with functions in cell cycle, transcription and post-transcriptional regulation and
13
14 569 protein degradation. Meanwhile, other spermatogonia-expressed genes encoding
15
16 570 proteins involved in cell proliferation, apoptosis and DNA repair pathways are not
17
18 571 affected in SpF, suggesting that in spermatogenic failure, although the abnormal
19
20 572 initiation of the meiotic process is already determined in these immature germ cells,
21
22 573 these cells maintain their activity related to mitosis and cell proliferation.

23
24
25 574 Regarding the Sertoli-specific expressed genes, the absence of differences in
26
27 575 gene expression of six out of eight genes studied in the SpF-MA and SpF-HS
28
29 576 phenotypes suggest that, in spermatogenic derangement, the functions of Sertoli cells
30
31 577 are in general maintained, but not those involved in cell cycle/apoptosis (*SLC4A11* and
32
33 578 *SCIN*) related to germ cell support. Previous studies have shown that, in mouse,
34
35 579 chemically-induced germ cell depletion can alter expression of several Sertoli cell genes
36
37 580 (Jonsson *et al.*, 1999; Maguire *et al.*, 1993; O'shaughnessy *et al.*, 2008) demonstrating
38
39 581 that germ cells regulate Sertoli cell activity by means of the regulation of Sertoli cell
40
41 582 gene expression. Here we observed a similar pattern in a pathological naturally-
42
43 583 occurring phenotype. We observed a gradual decrease of *SLC4A11* and *SCIN* transcript
44
45 584 levels attributable to the progressive depletion of germ cell stages (SCO>MA>HS>CS).
46
47 585 It is noteworthy that even in the presence of a two-fold increase of the number of Sertoli
48
49 586 cells as a compensatory phenomenon in SCO, the decrease in expression of certain
50
51 587 Sertoli cell genes becomes much more noticeable in the total absence of germ cells.
52
53
54
55
56
57
58
59
60

1
2
3 588 Some recent microarray studies have assessed global gene expression analysis in
4
5 589 testicular biopsies from infertile men in order to identify the genes critical for
6
7 590 spermatogenesis (Chalmel *et al.*, 2012; Von Kopylow *et al.*, 2010; Ellis *et al.*, 2007;
8
9 591 Feig *et al.*, 2007; Rockett *et al.*, 2004; Fox *et al.*, 2003). In these studies specific germ
10
11 592 cell transcription patterns are inferred from infertile testicular phenotypes in men and a
12
13 593 pattern of significantly decreased regulated genes has been attributed to the degree of
14
15 594 spermatogenic failure and the loss of specific stages of germ cells. We provide data
16
17 595 suggesting that **the molecular basis for severe spermatogenic impairment is more**
18
19 596 **complex than initially proposed.** In SpF the immature germ cells present an altered and
20
21 597 decreased transcriptional pattern of certain genes, and thus the number of genes
22
23 598 associated with these cells could be underestimated from microarray studies of infertile
24
25 599 samples. **Furthermore, our results should be helpful to better interpret microarray or**
26
27 600 **future NGS transcriptome studies. Elucidation of a more extensive transcriptional**
28
29 601 **profile with the detailed analysis of testicular cellular composition could be important in**
30
31 602 **understanding the molecular mechanisms that underlie male infertility.**
32
33
34
35

36 It is conspicuous that most of the altered spermatogonia-related genes are
37
38 604 involved in essential processes during spermatogenesis and aberrant expression is often
39
40 605 associated with spermatogenic defects. Whether the observed differential expression
41
42 606 profiles represent the cause or consequence of maturation arrest remains to be
43
44 607 elucidated. Considering the heterogeneous aetiologies and highly individual molecular
45
46 608 causes which may underlie spermatogenic failure in humans, the molecular changes
47
48 609 described here may represent common symptoms, but may also reflect early
49
50 610 dysfunction events affecting germ cells which may causally contribute to the pathology.
51
52 611 This data should be useful in delineating the patterns of gene expression involved in
53
54
55
56
57
58
59
60

1
2
3 612 male germ cell maturation deficiency, which may contribute to understanding male
4
5 613 infertility.
6

7 614 In summary, our study provides evidence that the premeiotic stage of germ cell
8
9 615 differentiation, exhibits associated patterns of gene expression deregulation in
10
11 616 spermatogenic impairment, which is more severe in meiotic arrest. This altered gene
12
13 617 expression pattern is observed despite there being no apparent morphological and/or
14
15 618 numerical change observed in this early stage of the germ cell population. In our
16
17 619 previous study, the spermatocyte capacity to express meiosis-related genes was
18
19 620 observed to be markedly reduced in spermatogenic failure, contributing to meiosis
20
21 621 impairment (Terribas *et al.*, 2010). Our present data demonstrate that the low
22
23 622 spermatogenic efficiency in infertile men is accompanied not only by meiotic but also
24
25 623 by premeiotic events in spermatogenesis, which contribute to spermatogenic blockade.
26
27 624 Furthermore, the differences in expression during the initial stages of spermatogenesis
28
29 625 in SpF-MA individuals suggest that this phenotype is already determined or arises in the
30
31 626 premeiotic stages of the germ line.
32
33
34
35

36 627
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 628 **ACKNOWLEDGEMENTS**
4

5 629 We are indebted to the patients who participated in this study. We thank Begoña
6
7 630 Hurtado for qBase analysis support, Xavier Solé for suggestions about statistical
8
9 631 analysis and Harvey Evans for the revision of the English text.
10
11

12 632

13
14
15 633 **FUNDING**
16

17
18 634 This work was supported by grants from the *Fondo de Investigaciones*
19
20 635 *Sanitarias/Fondo Europeo de Desarrollo Regional* (FIS/FEDER) [grant number
21
22 636 PI05/0759, PI09/1727, PI12/00361] and the *Generalitat de Catalunya* [grant number
23
24 637 2009SGR01490]. S.L. is sponsored by the Researchers Stabilization Program from the
25
26 638 Spanish National Health System [CES09/020]. S.B. was supported by the *Fondo de*
27
28 639 *Investigaciones Sanitarias-Instituto de Salud Carlos III* (FIS-ISCiii) [CA06/0055].
29
30

31 640
32
33

34 641 **AUTHOR'S CONTRIBUTION**
35

36
37 642 S.B. performed the RNA experiments and analysed the data. F.A. performed
38
39 643 immunohistochemistry and histological interpretation of data. E.F. provided samples
40
41 644 and clinical data. L.B. performed clinical assessment, provided samples, and critically
42
43 645 reviewed the manuscript. S.L. conceived and designed the experiments, supervised the
44
45 646 analysis of data and wrote the manuscript.
46
47
48

49 647
50

51 648
52
53
54
55
56
57
58
59
60

1
2
3 649 **FIGURE LEGENDS**

4
5 650 **Supplemental Figure 1.** Absolute expression levels of candidate reference genes (a.) in
6
7 651 SCO, MA, HS and CS groups and (b.) in GCT and CS groups.

8
9 652 ○, □, outlying values. * $p < 0.05$, (a.) Kruskal-Wallis test and (b.) Mann-Whitney test.

10
11 653 SCO: Sertoli cell only syndrome; MA: maturation arrest at spermatocyte stage; HS:
12
13 654 hypospermatogenesis; CS: conserved spermatogenesis

14
15 655

16
17
18 656 **Figure 1.** Spermatogonia-preferentially expressed genes whose relative expression
19
20 657 values were statistically altered in patient SpF group (MA and HS sub-phenotypes)
21
22 658 compared with CS controls. Expression levels relative to *PGK1/PGM1* are shown.

23
24
25 659 * $p < 0.05$, Mann-Whitney test; ** $p < 0.002$, Mann-Whitney test and Bonferroni
26
27 660 correction.

28
29 661 MA: maturation arrest at spermatocyte stage; HS: hypospermatogenesis; CS: conserved
30
31 662 spermatogenesis.

32
33 663

34
35
36 664 **Figure 2.** Testicular somatic cell-preferentially expressed genes that showed differences
37
38 665 in gene expression in infertile SCO, MA and HS patients relative to CS controls.
39
40 666 Expression levels relative to *PGK1/PPIA* are shown..

41
42
43 667 * $p < 0.0083$, Mann-Whitney test and Bonferroni correction.

44
45 668 SCO: Sertoli cell only syndrome; MA: maturation arrest at spermatocyte stage; HS:
46
47 669 hypospermatogenesis; CS: conserved spermatogenesis.

48
49 670

50
51 671 **Figure 3.**

52
53 672 Tissular expression profiling of *CCNE1* (a), *DAZL* (b), *RBM15* (c) and *STRA8* (d) by
54
55 673 quantitative real-time qPCR in testis with conserved spermatogenesis (CS),

1
2
3 674 hypospermatogenesis (HS) and maturation arrest at the spermatocyte (MA) Expression
4
5 675 levels relative to *PGKI* and *PGMI* are shown. Expression per cell profiling of *CCNE1*,
6
7 676 *DAZL*, *RBM15* and *STRA8* displayed as expression ratio per spermatogonium (x100) (e,
8
9 677 f, g, h) and expression per cell profiling of *CCNE1*, *DAZL*, *RBM15* and *STRA8*
10
11 678 displayed as expression ratio per spermatogonium and spermatocyte (x100) (i, j, k, l).
12
13 679 The horizontal bar indicates median value. Significant differences from the control are
14
15 680 indicated: * $p < 0.05$; ** $p < 0.005$.
16
17
18
19

20
21 682 **Figure 4.**

22
23 683 Immunohistochemical localization of selected proteins in sections of human testes with
24
25 684 different phenotypes. From left to right, first column (A,E,I,M) shows sections of CS,
26
27 685 second column (B,F,J,N) corresponds to SpF-HS, third column (C,G,K,O) represents
28
29 686 SpF-MA, and the fourth column (D,H,L,P) displays SCO pattern. *CCNE1* protein
30
31 687 staining is shown in panels A-D, *DAZL* in E-H, *CDKN1C* in I-L, and *DLK1* expression
32
33 688 in M-P, were stained Leydig cells are indicated by arrows. See explanation of the
34
35 689 cellular localization of different proteins in the text. Original magnification was X400
36
37 690 for panels A-L, and X200 for M-P. Scale bar in A and M = 100 μm .
38
39
40

41 691
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

692 **REFERENCES**

693

694 Chalmel F, Lardenois A, Evrard B, Mathieu R, Feig C, Demougin P, Gattiker A,

695 Schulze W, Jegou B, Kirchhoff C & Primig M (2012) Global human tissue

696 profiling and protein network analysis reveals distinct levels of transcriptional

697 germline-specificity and identifies target genes for male infertility. *Hum Reprod*,

698 27, 3233-3248.

699 Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker

700 A, Moore J, Patard JJ, Wolgemuth DJ, Jegou B & Primig M (2007) The

701 conserved transcriptome in human and rodent male gametogenesis. *Proc Natl*702 *Acad Sci U S A*, 104, 8346-8351.

703 Diederichs S, Baumer N, Schultz N, Hamra FK, Schrader MG, Sandstede ML, Berdel

704 WE, Serve H & Muller-Tidow C (2005) Expression patterns of mitotic and

705 meiotic cell cycle regulators in testicular cancer and development. *Int J Cancer*,

706 116, 207-217.

707 Ellis PJ, Furlong RA, Conner SJ, Kirkman-Brown J, Afnan M, Barratt C, Griffin DK &

708 Affara NA (2007) Coordinated transcriptional regulation patterns associated

709 with infertility phenotypes in men. *J Med Genet*, 44, 498-508.

710 Feig C, Kirchhoff C, Ivell R, Naether O, Schulze W & Spiess AN (2007) A new

711 paradigm for profiling testicular gene expression during normal and disturbed

712 human spermatogenesis. *Mol Hum Reprod*, 13, 33-43.

713 Feki A, Jefford CE, Durand P, Harb J, Lucas H, Krause KH & Irminger-Finger I (2004)

714 BARD1 expression during spermatogenesis is associated with apoptosis and

715 hormonally regulated. *Biol Reprod*, 71, 1614-1624.

- 1
2
3 716 Fox MS, Ares VX, Turek PJ, Haqq C & Reijo Pera RA (2003) Feasibility of global
4
5 717 gene expression analysis in testicular biopsies from infertile men. *Mol Reprod*
6
7 718 *Dev*, 66, 403-421.
- 9 719 Guo R, Yu Z, Guan J, Ge Y, Ma J, Li S, Wang S, Xue S & Han D (2004) Stage-
10
11 720 specific and tissue-specific expression characteristics of differentially expressed
12
13 721 genes during mouse spermatogenesis. *Mol Reprod Dev*, 67, 264-272.
- 16 722 Hellemans J, Mortier G, De Paepe A, Speleman F & Vandesompele J (2007) qBase
17
18 723 relative quantification framework and software for management and automated
19
20 724 analysis of real-time quantitative PCR data. *Genome Biol*, 8, R19.
- 23 725 Heyn H, Ferreira HJ, Bassas L, Bonache S, Sayols S, Sandoval J, Esteller M & Larriba
24
25 726 S (2012) Epigenetic disruption of the PIWI pathway in human spermatogenic
26
27 727 disorders. *PLoS One*, 7, e47892.
- 29 728 Irminger-Finger I, Leung WC, Li J, Dubois-Dauphin M, Harb J, Feki A, Jefford CE,
30
31 729 Soriano JV, Jaconi M, Montesano R & Krause KH (2001) Identification of
32
33 730 BARD1 as mediator between proapoptotic stress and p53-dependent apoptosis.
34
35 731 *Mol Cell*, 8, 1255-1266.
- 38 732 Jonsson CK, Zetterstrom RH, Holst M, Parvinen M & Soder O (1999) Constitutive
39
40 733 expression of interleukin-1alpha messenger ribonucleic acid in rat Sertoli cells is
41
42 734 dependent upon interaction with germ cells. *Endocrinology*, 140, 3755-3761.
- 45 735 Lin YM, Chen CW, Sun HS, Tsai SJ, Hsu CC, Teng YN, Lin JS & Kuo PL. (2001)
46
47 736 Expression patterns and transcript concentrations of the autosomal DAZL gene
48
49 737 in testes of azoospermic men. *Mol Hum Reprod*. 7, 1015-1022.
- 51
52 738
- 53
54
55
56
57
58
59
60

- 1
2
3 739 Maguire SM, Millar MR, Sharpe RM & Saunders PT (1993) Stage-dependent
4
5 740 expression of mRNA for cyclic protein 2 during spermatogenesis is modulated
6
7 741 by elongate spermatids. *Mol Cell Endocrinol*, 94, 79-88.
8
9
10 742 Mark M, Jacobs H, Oulad-Abdelghani M, Dennefeld C, Feret B, Vernet N, Codreanu
11
12 743 CA, Chambon P & Ghyselinck NB (2008) STRA8-deficient spermatocytes
13
14 744 initiate, but fail to complete, meiosis and undergo premature chromosome
15
16 745 condensation. *J Cell Sci*, 121, 3233-3242.
17
18
19 746 McLachlan RI, Rajpert-De Meyts E, Hoei-Hansen CE, de Kretser DM & Skakkebaek
20
21 747 NE (2007) Histological evaluation of the human testis--approaches to optimizing
22
23 748 the clinical value of the assessment: mini review. *Hum Reprod*, 22, 2-16.
24
25 749 Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD & Lee JT
26
27 750 (2006) Postmeiotic sex chromatin in the male germline of mice. *Curr Biol*, 16,
28
29 751 660-667.
30
31
32 752 O'Shaughnessy PJ, Hu L & Baker PJ (2008) Effect of germ cell depletion on levels of
33
34 753 specific mRNA transcripts in mouse Sertoli cells and Leydig cells.
35
36 754 *Reproduction*, 135, 839-850.
37
38
39 755 Pang AL, Taylor HC, Johnson W, Alexander S, Chen Y, Su YA, Li X, Ravindranath N,
40
41 756 Dym M, Rennert OM & Chan WY (2003) Identification of differentially
42
43 757 expressed genes in mouse spermatogenesis. *J Androl*, 24, 899-911.
44
45 758 Rockett JC, Patrizio P, Schmid JE, Hecht NB & Dix DJ (2004) Gene expression
46
47 759 patterns associated with infertility in humans and rodent models. *Mutat Res*,
48
49 760 549, 225-240.
50
51
52 761 Schlecht U, Demougin P, Koch R, Hermida L, Wiederkehr C, Descombes P, Pineau C,
53
54 762 Jegou B & Primig M (2004) Expression profiling of mammalian male meiosis
55
56
57
58
59
60

- 1
2
3 763 and gametogenesis identifies novel candidate genes for roles in the regulation of
4
5 764 fertility. *Mol Biol Cell*, 15, 1031-1043.
6
7 765 Schultz N, Hamra FK & Garbers DL (2003) A multitude of genes expressed solely in
8
9 766 meiotic or postmeiotic spermatogenic cells offers a myriad of contraceptive
10
11 767 targets. *Proc Natl Acad Sci U S A*, 100, 12201-12206.
12
13
14 768 Schulze W, Thoms F & Knuth UA (1999) Testicular sperm extraction: comprehensive
15
16 769 analysis with simultaneously performed histology in 1418 biopsies from 766
17
18 770 subfertile men. *Hum Reprod*, 14 Suppl 1, 82-96.
19
20
21 771 Sha J, Zhou Z, Li J, Yin L, Yang H, Hu G, Luo M, Chan HC & Zhou K (2002)
22
23 772 Identification of testis development and spermatogenesis-related genes in human
24
25 773 and mouse testes using cDNA arrays. *Mol Hum Reprod*, 8, 511-517.
26
27 774 Shima JE, McLean DJ, McCarrey JR & Griswold MD (2004) The murine testicular
28
29 775 transcriptome: characterizing gene expression in the testis during the progression
30
31 776 of spermatogenesis. *Biol Reprod*, 71, 319-330.
32
33
34 777 Simoni M, Bakker E, Eurlings MC, Matthijs G, Moro E, Muller CR & Vogt PH (1999)
35
36 778 Laboratory guidelines for molecular diagnosis of Y-chromosomal
37
38 779 microdeletions. *Int J Androl*, 22, 292-299.
39
40
41 780 Simoni M, Bakker E & Krausz C (2004) EAA/EMQN best practice guidelines for
42
43 781 molecular diagnosis of y-chromosomal microdeletions. State of the art 2004. *Int*
44
45 782 *J Androl*, 27, 240-249.
46
47 783 Terribas E, Bonache S, Garcia-Arevalo M, Sanchez J, Franco E, Bassas L & Larriba S
48
49 784 (2010) Changes in the expression profile of the meiosis-involved mismatch
50
51 785 repair genes in impaired human spermatogenesis. *J Androl*, 31, 346-357.
52
53
54 786 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A &
55
56 787 Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR
57
58
59
60

1
2
3 788 data by geometric averaging of multiple internal control genes. *Genome Biol*, 3,
4
5 789 RESEARCH0034.
6
7 790 von Kopylow K, Kirchhoff C, Jezek D, Schulze W, Feig C, Primig M, Steinkraus V &
8
9 791 Spiess AN (2010) Screening for biomarkers of spermatogonia within the human
10
11 792 testis: a whole genome approach. *Hum Reprod*, 25, 1104-1112.
12
13 793 Wang PJ, McCarrey JR, Yang F & Page DC (2001) An abundance of X-linked genes
14
15 794 expressed in spermatogonia. *Nat Genet*, 27, 422-426.
16
17 795 World Health Organization (1999) Laboratory Manual for the Examination of Human
18
19 796 Semen and Sperm-Cervical Mucus Interaction, p. Cambridge University Press,
20
21 797 New York.
22
23
24
25 798
26 799
27
28 800
29
30 801
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60