

1 **Changes in the Expression Profile of the Meiosis-Involved Mismatch Repair**
2 **(MMR) Genes in Impaired Human Spermatogenesis**

3

4 **Running title:** MMR mRNA in impaired sperm production

5

6 **ERNEST TERRIBAS,^{a, b} SANDRA BONACHE,^a MARTA GARCÍA-ARÉVALO,^a**
7 **JOSVANY SÁNCHEZ,^c ELADIO FRANCO,^d LLUÍS BASSAS,^c SARA LARRIBA^a**

8 *From: a. Medical and Molecular Genetics Center-Fundació IDIBELL, L'Hospitalet de*
9 *Llobregat, Barcelona, b. Present address: Institut de Medicina Predictiva i Personalitzada del*
10 *Càncer (IMPPC), Badalona, Barcelona, c. Andrology Service-Fundació Puigvert, Barcelona,*
11 *and d. Urology Service-Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona,*
12 *Spain*

13

14 **Correspondence and Reprint requests:** Sara Larriba, Medical and Molecular Genetics Center -
15 Fundació IDIBELL, Hospital Duran i Reynals, Gran Vía s/n km. 2.7, 08907 L'Hospitalet de
16 Llobregat, Barcelona, Spain. Tel. +34 932 60 74 25; Fax:+34 932 60 74 14; E-mail address:
17 slarriba@idibell.cat

18

19 **Financial support:** This work was supported by grants from FIS/FEDER (PI05/0759), Red de
20 Centros de Genética Clínica y Molecular ISCiii (C03/07) and the Generalitat de Catalunya
21 (2005SGR00018). S.L. and S.B. are supported by FIS-ISCiii (CP03/00088 and CA06/0055
22 respectively).

23

24 **ABSTRACT:** DNA Mismatch Repair (MMR) genes have been described to participate in
25 crossover events during meiotic recombination, which is, in turn, a key step of spermatogenesis.
26 This evidence suggests that MMR family gene expression may be altered in infertile men with
27 defective sperm production. In order to determine the expression profile of MMR genes in
28 impaired human spermatogenesis , we performed transcript levels analysis of MMR genes
29 (*MLH1*, *MLH3*, *PMS2*, *MSH4* and *MSH5*), and other meiosis-involved genes (*ATR*, *HSPA2*,
30 *SYCP3*) as controls, by real time reverse transcription-polymerase chain reaction (RT-PCR) in
31 testis from 13 patients with spermatogenic failure, 5 patients with primary germ-cell tumors and
32 10 controls with conserved spermatogenesis. Correlation of the expression values with the
33 histological findings was also performed.

34 The MMR gene expression values, with the exception of *PMS2*, are significantly
35 decreased in men with spermatogenic failure. The pattern of MMR reduction correlates with the
36 severity of damage, being maximum in maturation arrest. Specifically, expression of the testicular
37 *MSH4* gene could be useful as a surrogate marker for the presence of intratesticular elongated
38 spermatid in patients with non-obstructive azoospermia, contributing to predict the viability of
39 assisted reproduction. Interestingly, a reduction in the *MSH4* and *MSH5* transcript concentration
40 per spermatocyte was also observed. The decreased expression level of other meiosis-specific
41 genes, such as *HSPA2* and *SYCP3*, suggests that the spermatocyte capacity to express meiosis-
42 related genes is markedly reduced in spermatogenic failure, contributing to meiosis impairment
43 and spermatogenic blockade.

44

45 **Key Words:** *MLH1*, *MLH3*, *PMS2*, *MSH4*, *MSH5*, gene expression profiling, spermatocyte,
46 impaired sperm production

47 ***Introduction***

48 Mammalian spermatogenesis is a developmental process in which male germ cells undergo
49 mitotic proliferation, meiotic division and differentiation to produce a haploid gamete for sexual
50 reproduction. One major difference between mitosis and meiosis is the formation of a
51 proteinaceous structure called the synaptonemal complex (SC), which allows synapsis and
52 recombination between the two homologous chromosomes during meiotic prophase I. The axial
53 elements, the SC component also known as lateral elements after synapsis of the homologues,
54 begin to form between the two sister chromatids of each chromosome in leptotene cells. Synapsis
55 of homologous chromosomes usually begins even before axial element formation is complete and
56 involves alignment of homologues, connection of the two axial elements and formation of two
57 additional components of the SC, the central element and the transverse filaments. The process is
58 completed at the pachynema (Schmekel and Daneholt, 1995). The SC also contributes to the
59 crossing over at sites along the SC known as recombination nodules (Carpenter, 1987). Crossing
60 over, crucial for homologous recombination, occurs when two nonsister chromatids of the four
61 homologous chromatids cut and exchange equal segments, ensuring a correct segregation of
62 homologous chromosomes. Failure to segregate the appropriate haploid complement of
63 chromosome can have disastrous consequences by generating aneuploid gametes with the
64 potential to cause subsequent developmental anomalies or fetal loss (Koehler et al, 1996).
65 Alternatively, errors in recombination can activate checkpoint mechanisms resulting in meiotic
66 arrest and sterility (Gonsalves et al, 2004; Smith and Nicolas, 1998).

67 DNA Mismatch Repair (MMR) family proteins, consisting of the MutS and MutL
68 proteins in eukaryote organisms, have been evidenced to have a determinant role in DNA repair
69 after replication errors and their malfunction can lead to cancer in mammals. Studies in yeast and
70 mammals revealed that some members of this family participate in the meiotic recombination

71 process either correcting the potential mismatched bases of the heteroduplex DNA molecule after
72 recombination or promoting crossover events. Among these members there are three MutL
73 homologues (MLH1, MLH3 and PMS2) and two MutS homologues (MSH4 and MSH5) [see
74 (Kolas and Cohen, 2004; Surtees et al, 2004) for review]. Moreover, MSH4 and MSH5 are
75 meiosis-specific proteins crucial for reciprocal recombination but have no apparent mismatch
76 repair activity.

77 Several studies in yeast carrying disruptions in these five MMR genes showed reduced
78 meiotic crossing over and high frequency of postmeiotic segregation, demonstrating their role in
79 meiotic recombination [see (Kolas and Cohen, 2004; Surtees et al, 2004) for review].

80 To understand the role of MMR genes in DNA repair, cancer predisposition and meiosis,
81 several knockout mouse lines have been generated demonstrating a critical role of the above
82 mentioned MutS and MutL homologues in mammalian meiotic recombination [see (Kolas and
83 Cohen, 2004; Surtees et al, 2004) for review]. In *Mlh1*^{-/-} and in *Mlh3*^{-/-} male mice, germinal
84 differentiation is clearly arrested at pachytene and no mature sperm is produced. *Pms2*^{-/-} mice,
85 however, are able to produce spermatozoa, although they are aberrant due to an abnormal
86 chromosome synapsis in meiosis, causing sterility. Disruption of *Msh4*, as well as of *Msh5*, gene
87 also results in sterility due to an anomalous chromosome synapsis and meiotic failure.

88 These MMR family members have been reported to be highly expressed in mammalian
89 testicular tissue (Plevova et al, 2005; Santucci-Darmanin et al, 2002; Her et al, 2001; Kneitz et al,
90 2000; Paquis-Flucklinger et al, 1997; Bocker et al, 1999; Her and Doggett, 1998; Her et al, 1999),
91 preferentially in spermatocytes where meiotic recombination takes place. Human *MLH1* gene is
92 up-regulated relative to other stages in leptotene/zygotene cells, and decreases its expression in
93 pachytene nuclei (Marcon et al, 2008). MLH1 protein localizes on the SC (Anderson et al, 1999;
94 Oliver-Bonet et al, 2005), appearing by the early-mid pachytene transition and gradually

95 decreasing as spermatocyte progress through late pachynema (Oliver-Bonet et al, 2005; Santucci-
96 Darmanin et al, 2000). Mouse *Mlh3* transcripts and protein are found in zygotene and pachytene
97 spermatocytes (Santucci-Darmanin et al, 2002). *Pms2* transcripts and protein are found elevated
98 in mitotically proliferating spermatogonia and in leptotene and zygotene spermatocytes, declining
99 in early-mid pachytene (Richardson et al, 2000; Santucci-Darmanin et al, 2002). Mouse *Msh4*
100 gene is selectively expressed in spermatocytes from leptotema up to pachynema (Santucci-
101 Darmanin et al, 2002; Santucci-Darmanin et al, 2001), and the MSH4 protein presents a similar
102 distribution in mouse and human spermatocytes (Kneitz et al, 2000; Oliver-Bonet et al, 2005;
103 Santucci-Darmanin et al, 2000). The expression of MSH5 protein starts after early primary
104 spermatocytes and ends with elongated spermatids (Bocker et al, 1999).

105 This strong evidence of the requirement of *MLH1*, *MLH3*, *PMS2*, *MSH4* and *MSH5* genes
106 for a correct recombination process during yeast and mammalian meiosis suggests that some of
107 these genes may have a role in the regulation of spermatogenesis. Hence, we hypothesize their
108 expression may be affected in infertile men with spermatogenic failure and/or in men diagnosed
109 with germ-cell tumors. The aim of our study is to analyze the expression levels of *MLH1*, *MLH3*,
110 *PMS2*, *MSH4* and *MSH5* genes in testicular tissue from infertile patients and in primary germ-cell
111 tumors using quantitative real-time RT-PCR and evaluate the relationship between gene
112 expression levels and patients' testicular phenotypes. The results of this study will help us to
113 understand whether potential changes in MMR gene expression play a major role in the
114 impairment of sperm production.

115

116 ***Materials and Methods***117 *Patients and controls*

118 Our study recruited 13 infertile patients (mean age, 33 yr; range 27-40 yr) due to severe
119 spermatogenic failure (SpF), with a phenotype consistent with non-obstructive azoospermia or
120 severe oligozoospermia (<5 million sperm per mL), comprising patient group 1. Six of these
121 patients were diagnosed with maturation arrest at spermatocyte level (MA) and 7 with
122 hypospermatogenesis (HS) phenotype. Among MA samples submitted for histological analysis
123 (Table 1), 4 out of 5 showed arrest at the pachytene stage and 1 at the leptotene/pachytene stage.
124 Patient group 2 was formed by 5 men diagnosed with germ-cell tumor (GCT) (30 yr; 19-44 yr).
125 In addition, 10 infertile patients diagnosed with obstructive azoospermia (32 yr; 23-42 yr), who
126 showed conserved spermatogenesis (CS), were studied as controls (Table 1). Both patient group 1
127 and controls were selected from men referred for couple infertility to the Andrology Service of
128 the Fundació Puigvert and samples for patient group 2 were recruited from the Urology Service
129 of the Hospital Universitari de Bellvitge. The study was approved by the Institutional Review
130 Board of both Centers, and all the participants were informed and gave written consent to the
131 procedures of the study.

132 The clinical procedures for infertile patients included anamnesis, physical examination,
133 semen analyses [performed in accordance with World Health Organization guidelines (World
134 Health Organization, 1999)] and hormonal study. Concentrations of FSH reflected in general the
135 findings of testicular histology, although some patients showing blockade of primary
136 spermatocyte or hypospermatogenesis had normal FSH (Table 1). Spermograms included
137 volume, pH, sperm concentration, motility, vitality, morphology and fructose and citrate levels in
138 seminal plasma. The presence of normal vas deferens was assessed by scrotal palpation. The

139 testicular biopsy was obtained when necessary to confirm the clinical diagnosis and for sperm
140 retrieval (TESE) and cryopreservation purposes.

141 The routine genetic study for all samples included karyotype and analysis of chromosome
142 Y microdeletions, the latter performed according to the European guidelines (Simoni et al, 1999;
143 Simoni et al, 2004). Men with a chromosomal aberration or a Y-chromosome microdeletion were
144 not included in the study.

145

146 *Testicular samples*

147 Testicular biopsies from infertile men were obtained under local anesthesia through a small
148 incision. Each specimen was divided into three aliquots, one piece (≈ 10 -20 mg) was fixed in
149 Bouin's solution and reserved for histological analysis, a second aliquot (≈ 100 -200 mg) was
150 processed for sperm extraction and the third (≈ 10 mg) was immediately transferred to liquid
151 nitrogen and stored at -80°C until analysis for gene expression experiments.

152 Referring to GCT, testicular samples were obtained directly after orchiectomy and
153 macroscopic pathological evaluation. For gene expression studies, one tissue fragment was taken
154 from the tumoral portion of the testis and was immediately frozen at -80°C .

155

156 *Histological analysis*

157 Fixed testicular biopsies were cut in 5- μm sections and stained with haematoxylin-eosin.

158 Assessment of spermatogenic status was performed by quantification of specific germ cells, that
159 is spermatogoniae, primary spermatocytes, round spermatids and elongated spermatids, and of
160 Sertoli cells. The average number of cells per tubule was calculated after analysis of at least 15-20
161 cross-sectioned seminiferous tubules per testis. The number of elongated spermatids counted per

162 tubule and the yield of spermatozoa extracted per 100 mg of the matched samples during the same
163 procedure showed a Pearson's $R=0.775$ ($p<0.001$), confirming that the cell count in the tissue
164 sections was representative of the testicular histology found in the samples used for gene
165 expression (Silber et al, 1997). A modified Johnsen score (JS) count (Schulze et al, 1999) was
166 calculated on the basis of the number of different cell types per tubule. The mean diameter of the
167 seminiferous tubules was additionally evaluated (Table 1).

168

169 *RNA extraction and complementary DNA synthesis*

170 Total cellular RNA was extracted from the testicular biopsy using TriPure Isolation Reagent
171 (Roche Applied Science, Indianapolis, IN, USA), according to manufacturer's instructions. The
172 RNA samples were then quantified by measuring the absorbance at 260 nm using the Nanodrop
173 spectrophotometer. The quality of RNA [28S/18S ratio and RNA Integrity Number (RIN)] was
174 also assessed using the Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany).
175 Testicular RNA from the three groups of study showed similar quality values as both, 28S/18S
176 ratio and RIN, presented no significant differences among them ($p=0.450$, $p=0.190$, respectively).
177 Mean 28S/18S ratio value between all samples (mean \pm SD) was 1.2 ± 0.15 and mean RIN value
178 was 7.4 ± 0.49 .

179 Complementary DNA (cDNA) was synthesized with 1 μ g of RNA using random primers
180 and SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA) at 42°C for 50 min.

181 For quantitative real-time experiments, each RNA sample was submitted to two reverse
182 transcription (RT) reactions to minimize the variation of the experimental determination of
183 mRNA quantities due to RT efficiency. After RT reaction, 1:4 dilutions of cDNA samples were
184 prepared. The resulting cDNA aliquots were then stored at -20°C until use.

185
186 *Primers and quantitative PCR*
187 Hypoxanthine phosphoribosyl-transferase (*HPRT*), hydroxymethylbilane synthase (*HMBS*),
188 peptidylprolyl isomerase A-cyclophilin A (*PPIA*) and beta-2-microglobulin (*B2M*) were selected
189 as candidate reference genes for data normalization in quantitative PCR experiments. Ataxia
190 telangiectasia and Rad3 related (*ATR*), heat shock 70kDa protein 2 (*HSPA2*) and synaptonemal
191 complex protein 3 (*SYCP3*) were selected as control genes of the meiotic process. The sequences
192 of the forward and reverse primers used to amplify the human *MLH1*, *MLH3*, *PMS2*, *MSH4*,
193 *MSH5*, *ATR*, *HSPA2*, *SYCP3*, *HPRT* and *B2M* cDNA (Supplemental Table 1) were designed
194 using Oligo4.0 and Primer3 programs. Primers for MMR mRNA quantification were designed on
195 coding sequences that are present in multiple testicular transcript variants of the same target gene.
196 Primers for *HMBS* and *PPIA* cDNA amplification were previously described (Neuvians et al,
197 2005; Pluvinet et al, 2004). Primer specificity was first assessed prior to real-time PCR
198 experiments by agarose gel electrophoresis and sequencing the PCR product. It was later
199 confirmed in real-time PCR experiments by the analysis of the melting temperature, which is
200 product-specific.

201 Quantitative real-time PCR reactions were performed in a LightCycler 1.5 Instrument
202 (Roche Molecular Systems, Alameda, CA), using SYBR Green I dye and 1 μ l of the diluted
203 cDNA template in a total volume of 10 μ l. The products of the two previous RT reactions were
204 amplified twice each, to ensure best reproducibility. Negative controls without template were
205 included in each set of PCR assays as well as a sample of known gene expression copy number,
206 which was used as standard. In addition, patient group 1 and control group samples were always
207 analyzed as paired samples in the same analytical run in order to exclude between-run variations.

208 Details regarding temperatures and incubation times of amplification are shown in Supplemental
209 Table 1.

210 The Second Derivative Maximum Method for crossing point (Cp) determination from the
211 LightCycler Software 4.0 (Roche Diagnostics GmbH, Mannheim, Germany) and the standard
212 curve quantification method were used to calculate the raw expression values in all PCR samples.
213 Prior to data submission for gene expression normalization and for every gene, both target and
214 reference genes raw expression values were standardized, aiming for an accurate comparison of
215 the data and to avoid abundance differences between the genes. The gene expression ratio was
216 then calculated for each duplicated cDNA sample dividing the mean expression value of the
217 target gene by the mean expression value of the reference gene (both coming from the same RT
218 reaction) to normalize their expression for sample-to-sample differences in RNA input, RNA
219 quality and RT efficiency. Thus, two target/reference gene expression ratios were obtained for
220 each individual testicular sample, one of each coming from each RT reaction. The mean value
221 between these two ratios was submitted to statistical analysis.

222 Real-time PCR efficiencies for each gene of study were determined by measuring serial
223 1:2 dilutions of a cDNA sample in triplicate. Efficiencies were then calculated with LightCycler
224 Software 4.0 (Roche Diagnostics GmbH, Mannheim, Germany) according to the equation: $E =$
225 $10^{(-1/\text{slope})}$. Efficiency values ranged from 1.66 to 2.30 (Supplemental Table 1). To confirm
226 reproducibility and precision of real-time PCR experiments, intra-assay and inter-assay variation
227 were determined. Variation was measured as the coefficient of variation (CV) of Cp from the Cp
228 mean value. In the above mentioned RT-PCR runs, intra-assay variation ranged from 0.18% to
229 0.81%, confirming high reproducibility and precision. Inter-assay variation ranged from 0.65% to
230 3.08%.

231

232 *Data analysis*

233 Statistical analyses were performed using SPSS 12.0 software (LEAD Technologies, Inc, NJ,
234 USA). The nonparametric Mann-Whitney U test was used to analyze differences in absolute
235 expression levels of reference genes in both patient groups 1 and 2 compared to controls. The
236 Mann-Whitney U test was also used to evaluate differences in relative expression of target genes
237 in patient groups compared to controls. The nonparametric Kruskal-Wallis test was used to assess
238 differences in RNA quality between the three groups.

239 The gene-stability measure M was used to select the most stable reference genes and
240 improve the normalization of target genes. It is defined as the average pair wise variation of a
241 particular gene with all other genes, and for a given reference gene it is calculated by determining
242 the pair wise variation with all other reference genes as the standard deviation of the
243 logarithmically transformed expression ratios (Vandesompele et al, 2002).

244 Pearson product moment correlation coefficients were calculated to determine the
245 correlation between the expression ratios of the target genes and the different histological
246 parameters in patient group 1 and controls. Receiver operating characteristic (ROC) curve
247 analysis of the *MSH4* transcript was used for distinguishing between individuals with testicular
248 spermatids and those without. Accuracy was measured as the area under the ROC curve. The
249 threshold value was determined by Youden's index, calculated as sensitivity plus specificity – 1
250 (Skendzel and Youden, 1970). A *p*-value <0.05 was considered significant.

251

252 **Results**

253 *Assessment of alternative MMR transcripts in human testis*

254 Prior to primer design for gene expression quantification, we assessed the presence of previously
255 described alternatively spliced transcript variants in several testicular samples by conventional
256 RT-PCR and sequence analysis with the primers listed in Supplemental Table 1.

257 As regards the *MLH1* gene, three alternatively spliced transcript variants encoding
258 different isoforms have been previously described in several somatic tissues, although their full-
259 length nature have not been fully determined (Charbonnier et al, 1995). We found that a low
260 proportion of testicular transcripts lacked exons 10 and 11 (in CS as well as in MA and HS
261 tissues), while the other two *MLH1* isoforms were negatively expressed (Supplemental Figure 1).

262 Several full-length alternative MMR mRNAs have been described in somatic and
263 testicular tissues. We confirmed the testicular expression of the two *MLH3* mRNA isoforms,
264 *MLH3* variant 1 (NCBI RefSeq NM_001040108), and variant 2 (NM_014381), showing the latter
265 to contain an in-frame deletion of exon 7 that results in a shorter protein (Lipkin et al, 2000;
266 Santucci-Darmanin et al, 2002). Both the longest (variant 1, NM_000535) and shortest (variant 2,
267 NR_003085) *PMS2* transcripts were also found in testis. Expression of testicular *MSH4* isoform,
268 $\Delta hMSH4$, lacking exon 6 (Santucci-Darmanin et al, 1999), was also corroborated. Referring to
269 *MSH5*, several mRNA splice variants have been described, such as the one containing an in-
270 frame insertion of the last 51 bp of intron 6 resulting in the longest isoform (transcript variant 1,
271 NM_025259), a transcript variant containing three extra base pairs between exons 20 and 21
272 (variant 2, NM_172165) and two transcript variants resulting from the use of an alternative in-
273 frame splice donor site in exon 1, compared to variant 1 (variant 3, NM_002441 and variant 4,
274 NM_172166). We determined that *MSH5* variants 1, 2 and 4 were expressed, while variant 3 was

275 absent in testis. The presence of these alternative forms was additionally determined in the
276 infertile testicular tissues (Supplemental Figure 1).

277 Moreover, while assessing primer specificity, we detected six additional alternatively
278 spliced MMR mRNA forms expressed in testis although their full-length sequence was not fully
279 determined. These new spliced forms were a *MLH3* variant lacking exon 5, a *MLH3* variant
280 lacking both exons 5 and 7, an alternative *MSH4* mRNA containing an in-frame deletion (the first
281 97 pb of exon 6), an alternatively spliced *MSH4* mRNA lacking exon 19, a variant of *MSH5*
282 mRNA including intron 6 and a variant of *MSH5* mRNA lacking exon 9. Both *MLH3* isoforms
283 contained a deletion of the coding region that preserves the open reading frame and,
284 consequently, translation of these transcripts would result in shorter proteins. The deletion
285 observed in *MSH4* and *MSH5* isoforms, however, would result in a reading frameshift with a new
286 stop codon downstream, thus translation of these transcripts are predicted to result in truncated
287 proteins. All these new variants were confirmed to be present in lymphocytes [with the exception
288 of *MSH4* (-Exon 19)] as well as in both MA and HS testicular tissue (Supplemental Figure 1).

289 Based on these results, primers for MMR mRNA quantification were designed so as to be
290 able to amplify different testicular transcript variants of the same target gene, and thus global
291 differences in gene expression, affecting multiple splice variants, could be evaluated.

292

293 *Expression levels of candidate reference genes. Selection of a suitable normalizer*

294 Measurement of gene expression by real-time RT-PCR requires at least one proper internal
295 control gene for normalization purposes in order to achieve precise and reliable quantitative
296 expression results of the genes under study. We have assessed the suitability of four
297 housekeeping genes widely used as normalizers in the literature: *HPRT* (purine nucleotide

298 biosynthesis), *HMBS* (porphyrin metabolism), *PPIA* (protein folding) and *B2M* (immune
299 response) as candidate genes for normalization in both pathological and normal testicular tissue.

300

301 Spermatogenic failure

302 Figure 1a shows the mRNA levels of the four candidate reference genes in patient group 1 and
303 controls. When we first looked for differences in gene expression between both groups, non-
304 significant differences in absolute gene expression levels were observed for the four genes
305 analyzed (*HPRT*, $p=0.522$; *HMBS*, $p=0.077$; *PPIA*, $p=0.410$; *B2M*, $p=0.648$), suggesting that all
306 four genes were potentially suitable reference genes for expression normalization in testis tissue
307 from SpF and CS samples. However, in order to select the most stable control genes and improve
308 the normalization, the gene-stability measure M (Vandesompele et al, 2002), was calculated for
309 every reference gene. Lower M values indicate genes with less expression variation among
310 samples. M values for our reference genes resulted in 0.82 for *HPRT*, 0.72 for *HMBS*, 0.65 for
311 *PPIA* and 1.11 for *B2M*. Because *B2M* showed the highest M value and it was much more
312 elevated than the M value of the other three genes, it was excluded as a reference gene for further
313 analysis. Then M values were recalculated for the remaining genes, resulting in 0.80 for *HPRT*,
314 0.67 for *HMBS* and 0.69 for *PPIA*. Since all three reference genes displayed a similar M value, a
315 normalization factor (NF), based on the average of the expression levels of the most stable
316 reference genes – *HPRT*, *HMBS*, and *PPIA* –, was calculated for each sample as the arithmetic
317 mean of the selected reference gene expression values. This NF was selected as the suitable value
318 for normalization in order to calculate the relative expression of our target genes in SpF samples.

319

320 Germ-cell tumors

321 As regards the absolute expression levels of the reference genes in patient group 2 and control
322 samples (Figure 1b), significant differences in expression were found for *HPRT* ($p=0.001$), *PPIA*
323 ($p=0.019$) and *B2M* ($p=0.001$) when both groups were compared. *HMBS* expression showed non-
324 significant differences between tumoral and normal control testicular tissue ($p=0.953$), hence, it
325 was later used as the normalizer gene when studying GCT samples.

326

327 *Relative expression of MMR genes in patients and controls*

328 Spermatogenic failure

329 When studying patient group 1, normalized expression levels of MMR genes were calculated as
330 the ratio ‘target gene/NF’ for each sample (Figure 2a). Significant differences in gene expression
331 levels were observed between patients and controls for *MLH1* ($p=0.030$), *MLH3* ($p=0.015$),
332 *MSH4* ($p=0.000$) and *MSH5* ($p=0.004$). Non-significant differences were found for *PMS2*
333 ($p=0.088$) (Figure 2a, Supplemental Table 2).

334 In order to demonstrate the relevance of choosing an appropriate reference gene/s to
335 obtain reliable interpretation of target gene expression data, MMR gene expression ratios were
336 also calculated in patient group 1 and controls using each single reference gene as normalizer.
337 The difference of expression was statistically significant for *MSH4* and *MSH5* when data were
338 normalized with both single reference gene and NF. However, the reliability of the interpretation
339 of MutL homologues expression data was highly dependent on the applied normalizer. In this
340 case, similar results were obtained when using NF or *PPIA* as normalizers (Supplemental Table
341 2).

342 When the patient group 1 was divided into MA and HS subgroups, we observed that
343 reduction of expression was more pronounced in the maturation arrest phenotype. The percentage
344 of expression reduction of our target genes compared to controls ranged from 24% for *PMS2* to

345 60% for *MSH4* in MA patients and from 11% for *PMS2* to 34% for *MSH4* in HS patients (Figure
346 2b). When we compared the expression ratios between MA subgroup and CS controls, we
347 observed significant differences for *MLH1* ($p=0.016$), *MLH3* ($p=0.031$), *MSH4* ($p=0.002$) and
348 *MSH5* ($p=0.031$) and non-significant differences for *PMS2* ($p=0.118$). Interestingly, expression
349 ratios between HS phenotype and controls were found to be differentially expressed only for
350 *MSH4* ($p=0.005$) and *MSH5* ($p=0.014$), while MutL homologues, although their expression was
351 reduced in the HS subgroup, presented non-significant differences when compared to CS
352 controls: *MLH1* ($p=0.230$), *MLH3* ($p=0.070$) and *PMS2* ($p=0.230$).

353

354 Germ-cell tumors

355 Normalized expression levels of MMR genes in patient group 2 compared to control group were
356 calculated as the ‘target gene/*HMBS*’ expression ratio for each sample (Figure 3). Statistically
357 significant differences in relative gene expression levels were observed between tumoral and
358 normal testicular tissue in each gene (*MLH1*, $p=0.001$; *MLH3*, $p=0.008$; *PMS2*, $p=0.005$; *MSH4*,
359 $p=0.001$; *MSH5*, $p=0.001$).

360

361 *Correlation study between MMR gene expression profiles and histological parameters*

362 In order to assess whether there is an association between gene expression and tubular cell
363 number and confirm whether the results could be of physiological relevance, we performed the
364 correlation study between the five normalized gene expression ratios and several histological
365 parameters such as: seminiferous tubular diameter, number of each type of cell from the germ
366 line, Sertoli cell number and JS count (Figure 4a) in patient 1 and control groups. When referring
367 to the diameter of seminiferous tubules, it significantly correlated with *MLH3*, *MSH4* and *MSH5*.
368 When the total number of samples was considered, most of the histological parameters, with the

369 exception of Sertoli cell and spermatogonia number, positively correlate with *MLH1*, *MLH3*,
370 *MSH4* and *MSH5* expression levels. Interestingly, significant positive correlation coefficients
371 were found between the number of elongated spermatids and the transcription levels of the five
372 genes, being remarkable the correlation coefficient for *MSH4* ($r=0.815$) (Figure 4a). We
373 hypothesize that there may be a threshold level of *MSH4* transcripts related with the presence of
374 intratesticular elongated spermatid. The ROC curve analysis of *MSH4* expression levels indicates
375 that the threshold that gave the maximum true-positive fraction (sensitivity) and false-positive
376 fraction (1-specificity) was 0.917. At this threshold value, the sensitivity and the specificity for
377 predicting the presence of ≥ 1 elongated spermatid per tubule were 80 and 100%, respectively
378 (Figure 4b). The calculated area under the curve was 0.944, with a 95% confidence interval of
379 0.850 to 1. As comparison, the areas under the ROC curve of testicular volume and FSH
380 concentration were 0.726 (95% CI 0.477 – 0.976) and 0.327 (95% CI 0.028 – 0.627) respectively.

381

382 *Relative expression of other meiosis- involved genes in SpF patients and controls*

383 In order to determine whether testicular gene expression alteration in SpF patients exclusively
384 affected MMR genes or, on the contrary, it was a generalized phenomenon affecting other
385 meiosis-involved genes, we have performed an additional expression profile analysis of other
386 spermatocyte preferentially-expressed genes (Chalmel et al, 2007) such as *ATR* (cell cycle arrest
387 and DNA damage repair), *HSPA2* (male meiosis) and *SYCP3* (synaptonemal complex structure).

388 Normalized expression levels of meiosis-involved genes were calculated as the ratio
389 ‘target gene/NF’ for each sample (Figure 5). Significant differences in gene expression levels
390 were observed between patients and controls for *ATR* ($p=0.006$), *HSPA2* ($p=0.018$), and *SYCP3*
391 ($p=0.026$). When considering SpF subgroups, significant differences in gene expression levels
392 were observed for *ATR* in both MA and HS subgroups when compared to CS ($p=0.042$ and

393 $p=0.019$ respectively) and for *HSPA2* between MA and CS ($p=0.011$), whereas non-significant
394 differences were found for *HSPA2* between HS and CS ($p=0.161$) and for *SYCP3* in both MA
395 and HS patients comparing to CS ($p=0.073$ and $p=0.070$ respectively).

396

397 **Expression levels of germ cell-specific genes per cell**

398 We additionally analyzed the germ cell-specific transcript levels per spermatocyte cell in both
399 MA and HS subgroups compared to CS controls in order to obviate the differences in gene
400 expression due to changes in testicular cellularity and to determine whether spermatocyte gene
401 expression is altered in spermatogenic failure. Selective germ cell expression of *MSH4* and
402 *MSH5*, but not of *MLH1*, *MLH3* and *PMS2*, was previously confirmed as negligible transcript
403 level values were found in three complete Sertoli Cell-Only (SCO) samples (data not shown).
404 Values of transcript amount per cell, in arbitrary units, were obtained for each testicular sample
405 by dividing the expression ratio value by the proportion of primary spermatocytes [known to be
406 the germ cell stage that predominantly expresses *MSH4* and *MSH5* in the testis (Chalmel et al,
407 2007)] present in a seminiferous tubule of the sample (Figure 6a, 6b). Significant differences in
408 cellular transcript levels were found for *MSH4* between SpF patients and controls ($p=0.000$), MA
409 subgroup and controls ($p=0.001$) and HS subgroup and controls ($p=0.000$) and for *MSH5*
410 between SpF patients and controls ($p=0.000$), MA subgroup and controls ($p=0.002$) and HS
411 subgroup and controls ($p=0.001$).

412 In order to determine whether differences of expression per cell affected other germ cell-
413 specific genes involved in meiosis, *HSPA2* and *SYCP3* transcript levels per cell were also
414 determined in the groups of study, as these genes have been previously described to be
415 predominantly expressed in primary spermatocytes (Chalmel et al, 2007). Selective germ cell
416 expression of *HSPA2* and *SYCP3* was confirmed as negligible transcript level values were found

417 in three complete Sertoli Cell-Only (SCO) samples (data not shown). Values of transcript amount
418 per cell were determined as described for *MSH4* and *MSH5* (Figure 6c and 6d). Significant
419 differences were found for *HSPA2* and *SYCP3* between SpF patients and controls ($p=0.001$;
420 $p=0.000$ respectively), MA subgroup and controls ($p=0.001$; $p=0.002$ respectively) and HS
421 subgroup and controls ($p=0.023$; $p=0.005$ respectively).

422

423 ***Discussion***

424 The formation of a mature sperm is a very complex process involving the transcription of many
425 genes. Defects in essential genes can result in impaired sperm or no sperm at all. Defective
426 meiosis during spermatogenesis is one of the critical causes of severe sperm impairment,
427 although the details still remain unknown. In mammals, meiosis is a fundamental process that
428 allows a genetic exchange between maternal and paternal genomes (Nasmyth, 2002). To shed
429 some light on the expression behavior of the meiosis-involved MMR genes in the impairment of
430 sperm production, we evaluated testicular *MLH1*, *MLH3*, *PMS2*, *MSH4* and *MSH5* gene
431 expression in relation to meiosis alterations.

432 The present analytical strategy for gene expression quantification was very carefully
433 designed to minimize both external and internal influences on expression data and improve
434 experimental accuracy. Several factors were considered, including an acceptable RNA quality
435 without statistical differences between the groups under study, duplicates of RT and PCR
436 reactions, and low intra-assay and inter-assay variation values of PCR runs. As alternative
437 splicing, as well as, alternative transcriptional initiation and polyadenylation are the main
438 mechanisms for generating germ cell-specific and stage-dependent mRNAs (Eddy and O'Brien,
439 1998), multiple testicular transcript variants of the same target gene were considered for primer
440 design. Transcript amounts were measured by real-time RT-PCR analysis and data normalized to

441 suitable reference genes, which should show constitutive and stable expression levels in the
442 samples investigated. An appropriate reference gene/s was chosen for each experimental
443 condition affecting testicular tissue, being determinant especially when studying the biological
444 significance of small expression differences between groups (Supplemental Table 2).

445 Our data indicate that testicular expression levels of meiosis-involved MMR genes (with
446 the exception of *PMS2*) are significantly reduced in SpF patients compared to CS men. Moreover
447 the reduction is much more significant in the MA phenotype than in HS. The MMR gene
448 transcription efficiencies are even more reduced in GCT infertile individuals and specifically
449 MutS homologues mRNA expression levels were very low or almost negligible in these patients,
450 possibly related to the fact that the germ-line in testicular tumor has undergone a dedifferentiation
451 process.

452 Interestingly, a remarkable significant positive correlation coefficient was found between
453 the number of elongated spermatids and the transcription levels of *MSH4*. The testicular *MSH4*
454 expression ratio was able to accurately predict the presence of intratesticular elongated spermatid.
455 This could be potentially used as a surrogate marker for the presence of full spermatogenesis in
456 patients with non-obstructive azoospermia, especially those considering further attempts of
457 invasive testicular extraction after a first negative biopsy with fine needle sperm aspiration.

458 Previous data evidenced that these four MMR proteins (Mlh1, Mlh3, Msh4 and Msh5)
459 collaborate with each other as a complex in promoting meiotic recombination and crossing over,
460 initiated by the association of meiosis-specific MutS γ heterodimer (Msh4-Msh5) with the DNA at
461 zygonema and followed by the recruitment of the heterodimeric complex of MutL homologues
462 Mlh1-Mlh3 (MutL γ) at pachynema, stabilizing the interaction. Pms2, however, although capable
463 of heterodimerizing with Mlh1, does not have a direct function in crossing over (Kolas et al,

464 2005) but has a role in the regulation of the nuclear or cytoplasmic location of MLH3 in the cell
465 by competing with MLH3 for the interaction with MLH1 (Korhonen et al, 2007). This crossing
466 over-independent regulatory role of PMS2 supports the finding of the lack of expression
467 difference for *PMS2* but not for *MLH1*, *MLH3*, *MSH4* and *MSH5* in our testicular samples.

468 When studying gene expression profiles in the testis, an inherent problem to be taken into
469 account is the cellular complexity of this organ. Changes in gene expression at the tissue level
470 can reflect changes in the capability of transcribing the mRNA in a specific cell type as well as
471 changes in the cell-type composition or number. The reduction of MMR gene expression in SpF
472 patients could be partially explained by the decreased number of germ cells that specifically
473 express MMR genes in these individuals, in fact, a positive correlation between gene expression
474 and germ cells was determined (Figure 4a). Nevertheless, although the number of spermatocytes
475 per tubule was decreased in infertile samples when compared to controls, non-significant
476 differences among groups were found ($p=0.113$), thus changes in expression observed among
477 groups could be not exclusively explained by the spermatocyte cell number. Interestingly, an
478 additional statistically significant reduction in the expression levels of germ cell-specific genes
479 per spermatocyte was observed in MA and HS when compared to CS samples, demonstrated for
480 *MSH4* and *MSH5* genes (Figure 6), being more pronounced in the maturation arrest phenotype.
481 The histological pattern of testicular hypospermatogenesis may be related to some level of
482 maturation arrest in the tubules, which may explain the differences in MMR gene expression
483 profiles per cell of HS with that of the CS or MA group and should contribute to the
484 understanding of patterns of in vivo expression of MMR genes in male infertility of testicular
485 origin.

486 A decreased cellular expression level of other meiosis-involved genes, *HSPA2* and
487 *SYCP3*, was also detected and correlates with the severity of testicular damage, as occurred for

488 *MSH4* and *MSH5*. These data indicate that MMR gene expression alteration is the result of a
489 generalized phenomenon affecting spermatocyte gene expression capacity, and support the
490 hypothesis that the meiosis alteration may already be arising in early stages of spermatogenesis,
491 leading to a global reduction of the meiosis-involved gene expression contributing to
492 spermatogenic blockade. Protein data on non-obstructive testicular tissue corroborate our mRNA
493 expression results: maturational arrest tissue showed weak HSPA2 staining within spermatocytes
494 when compared to normal tissue by means immunofluorescence technique (Feng et al, 2001).

495 *MSH4/MSH5* heterodimer acts locally at sites of emerging recombination events.
496 Specifically, a role for *MSH4* in synapsis initiation and maintenance has been suggested as well
497 as in the determination of the recombination sites by attracting *MLH3/MLH1* (Oliver-Bonet et al,
498 2005). Antibodies against *MLH1* are used to identify the sites of meiotic recombination on
499 synaptonemal complex. Meiotic studies on the pachytene stage of spermatogenesis have
500 demonstrated that non-obstructive infertile men have impaired chromosome synapsis, a
501 significantly decreased frequency of recombination, and an increased frequency of chromosomes
502 completely lacking a recombination site (Sun et al, 2006). It is tempting to speculate that such
503 errors could be partially consequence of decreased expression levels of MMR genes in the
504 spermatocytes. Moreover, the defects of germ-cell MMR expression can increase the generation
505 of aneuploid gametes with potential consequences for fetal development, if the non-obstructive
506 individual is included in an assisted reproduction program.

507 In summary, we developed a reliable approach that allows the analysis of gene expression
508 in testicular biopsies taking into account the variability in testicular cellularity between control
509 and pathological infertile testis. By this method we describe a reduction of transcript
510 concentration of meiosis-involved MMR genes in patients with severe impaired sperm
511 production, especially in those with maturation arrest. The defects of transcript levels in SpF

512 seem to be a consequence of a global phenomenon, where the spermatocyte expression capability
513 is affected, contributing to spermatogenic blockade. Future studies of gene expression of early
514 cellular stages of spermatogenesis as well as the study of factors involved in regulating gene
515 expression in the spermatogenic process may help us to understand the molecular mechanisms
516 that regulate the correct initiation and progression of meiotic process. Moreover, these findings
517 contribute to the search and selection of the most valuable gene markers potentially useful as
518 additional tools for the detection of sperm production, *MSH4* as a marker of spermatogenesis, and
519 for predicting the viability of assisted reproduction.

520 ***Acknowledgements***

521 We are indebted to the patients who participated in this study. We thank JM Pujal for *B2M*
522 primers and LC analysis support, Laboratory of Pathology of the Fundació Puigvert for testicular
523 sections, E Guinó for her advice on statistical analysis and H Kruyer for the revision of the
524 English text.

525

526 **References**

- 527 Anderson LK, Reeves A, Webb LM, Ashley T. Distribution of crossing over on mouse
528 synaptonemal complexes using immunofluorescent localization of MLH1 protein.
529 Genetics. 1999; 151:1569-79.
- 530 Bocker T, Barusevicius A, Snowden T, Rasio D, Guerrette S, Robbins D, Schmidt C, Burczak J,
531 Croce CM, Copeland T and others. hMSH5: a human MutS homologue that forms a novel
532 heterodimer with hMSH4 and is expressed during spermatogenesis. Cancer Res. 1999;
533 59:816-22.
- 534 Carpenter AT. Gene conversion, recombination nodules, and the initiation of meiotic synapsis.
535 Bioessays. 1987; 6:232-6.
- 536 Chalmel F, Rolland AD, Niederhauser-Wiederkehr C, Chung SS, Demougin P, Gattiker A,
537 Moore J, Patard JJ, Wolgemuth DJ, Jegou B and others. The conserved transcriptome in
538 human and rodent male gametogenesis. Proc Natl Acad Sci U S A. 2007; 104:8346-51.
- 539 Charbonnier F, Martin C, Scotte M, Sibert L, Moreau V, Frebourg T. Alternative splicing of
540 MLH1 messenger RNA in human normal cells. Cancer Res. 1995; 55:1839-41.
- 541 Eddy EM, O'Brien DA. Gene expression during mammalian meiosis. Curr Top Dev Biol. 1998;
542 37:141-200.
- 543 Feng HL, Sandlow JI, Sparks AE. Decreased expression of the heat shock protein hsp70-2 is
544 associated with the pathogenesis of male infertility. Fertil Steril. 2001;76:1136-39.
- 545 Gonsalves J, Sun F, Schlegel PN, Turek PJ, Hopps CV, Greene C, Martin RH, Pera RA.
546 Defective recombination in infertile men. Hum Mol Genet. 2004; 13:2875-83.
- 547 Her C, Doggett NA. Cloning, structural characterization, and chromosomal localization of the
548 human orthologue of *Saccharomyces cerevisiae* MSH5 gene. Genomics. 1998; 52:50-61.
- 549 Her C, Wu X, Wan W, Doggett NA. Identification and characterization of the mouse MutS
550 homolog 5: Msh5. Mamm Genome. 1999; 10:1054-61.
- 551 Her C, Wu X, Bailey SM, Doggett NA. Mouse MutS homolog 4 is predominantly expressed in
552 testis and interacts with MutS homolog 5. Mamm Genome. 2001; 12:73-6.
- 553 Kneitz B, Cohen PE, Avdievich E, Zhu L, Kane MF, Hou H, Jr., Kolodner RD, Kucherlapati R,
554 Pollard JW, Edelman W. MutS homolog 4 localization to meiotic chromosomes is
555 required for chromosome pairing during meiosis in male and female mice. Genes Dev.
556 2000; 14:1085-97.

- 557 Koehler KE, Hawley RS, Sherman S, Hassold T. Recombination and nondisjunction in humans
558 and flies. *Hum Mol Genet.* 1996; 5 Spec No:1495-504.
- 559 Kolas NK, Cohen PE. Novel and diverse functions of the DNA mismatch repair family in
560 mammalian meiosis and recombination. *Cytogenet Genome Res.* 2004; 107:216-31.
- 561 Kolas NK, Svetlanov A, Lenzi ML, Macaluso FP, Lipkin SM, Liskay RM, Greally J, Edelmann
562 W, Cohen PE. Localization of MMR proteins on meiotic chromosomes in mice indicates
563 distinct functions during prophase I. *J Cell Biol.* 2005; 171:447-58.
- 564 Korhonen MK, Raevaara TE, Lohi H, Nystrom M. Conditional nuclear localization of hMLH3
565 suggests a minor activity in mismatch repair and supports its role as a low-risk gene in
566 HNPCC. *Oncol Rep.* 2007; 17:351-4.
- 567 Lipkin SM, Wang V, Jacoby R, Banerjee-Basu S, Baxevanis AD, Lynch HT, Elliott RM, Collins
568 FS. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite
569 instability. *Nat Genet.* 2000; 24:27-35.
- 570 Marcon E, Babak T, Chua G, Hughes T, Moens PB. miRNA and piRNA localization in the male
571 mammalian meiotic nucleus. *Chromosome Res.* 2008; 16:243-60.
- 572 Nasmyth K. Segregating sister genomes: the molecular biology of chromosome separation.
573 *Science.* 2002; 297:559-65.
- 574 Neuvians TP, Gashaw I, Sauer CG, von Ostau C, Kliesch S, Bergmann M, Hacker A, Grobholz
575 R. Standardization strategy for quantitative PCR in human seminoma and normal testis. *J*
576 *Biotechnol.* 2005; 117:163-71.
- 577 Oliver-Bonet M, Turek PJ, Sun F, Ko E, Martin RH. Temporal progression of recombination in
578 human males. *Mol Hum Reprod.* 2005; 11:517-22.
- 579 Paquis-Flucklinger V, Santucci-Darmanin S, Paul R, Saunieres A, Turc-Carel C, Desnuelle C.
580 Cloning and expression analysis of a meiosis-specific MutS homolog: the human MSH4
581 gene. *Genomics.* 1997; 44:188-94.
- 582 Plevova P, Sedlakova E, Zapletalova J, Krepelova A, Skypalova P, Kolar Z. Expression of the
583 hMLH1 and hMSH2 proteins in normal tissues: relationship to cancer predisposition in
584 hereditary non-polyposis colon cancer. *Virchows Arch.* 2005; 446:112-9.
- 585 Pluvinet R, Petriz J, Torras J, Herrero-Fresneda I, Cruzado JM, Grinyo JM, Aran JM. RNAi-
586 mediated silencing of CD40 prevents leukocyte adhesion on CD154-activated endothelial
587 cells. *Blood.* 2004; 104:3642-6.

- 588 Richardson LL, Pedigo C, Ann Handel M. Expression of deoxyribonucleic acid repair enzymes
589 during spermatogenesis in mice. *Biol Reprod.* 2000; 62:789-96.
- 590 Santucci-Darmanin S, Paul R, Michiels JF, Saunieres A, Desnuelle C, Paquis-Flucklinger V.
591 Alternative splicing of hMSH4: two isoforms in testis and abnormal transcripts in somatic
592 tissues. *Mamm Genome.* 1999; 10:423-7.
- 593 Santucci-Darmanin S, Walpita D, Lespinasse F, Desnuelle C, Ashley T, Paquis-Flucklinger V.
594 MSH4 acts in conjunction with MLH1 during mammalian meiosis. *Faseb J.* 2000;
595 14:1539-47.
- 596 Santucci-Darmanin S, Vidal F, Scimeca JC, Turc-Carel C, Paquis-Flucklinger V. Family of
597 SRY/Sox proteins is involved in the regulation of the mouse Msh4 (MutS Homolog 4)
598 gene expression. *Mol Reprod Dev.* 2001; 60:172-80.
- 599 Santucci-Darmanin S, Neyton S, Lespinasse F, Saunieres A, Gaudray P, Paquis-Flucklinger V.
600 The DNA mismatch-repair MLH3 protein interacts with MSH4 in meiotic cells,
601 supporting a role for this MutL homolog in mammalian meiotic recombination. *Hum Mol*
602 *Genet.* 2002; 11:1697-706.
- 603 Schmekel K, Daneholt B. The central region of the synaptonemal complex revealed in three
604 dimensions. *Trends Cell Biol.* 1995; 5:239-42.
- 605 Schulze W, Thoms F, Knuth UA. Testicular sperm extraction: comprehensive analysis with
606 simultaneously performed histology in 1418 biopsies from 766 subfertile men. *Hum*
607 *Reprod.* 1999; 14 Suppl 1:82-96.
- 608 Silber SJ, Nagy Z, Devroey P, Tournaye H, Van Steirteghem AC. Distribution of
609 spermatogenesis in the testicles of azoospermic men: the presence or absence of
610 spermatids in the testes of men with germinal failure. *Hum Reprod.* 1997; 12:2422-8.
- 611 Simoni M, Bakker E, Eurlings MC, Matthijs G, Moro E, Muller CR, Vogt PH. Laboratory
612 guidelines for molecular diagnosis of Y-chromosomal microdeletions. *Int J Androl.* 1999;
613 22:292-9.
- 614 Simoni M, Bakker E, Krausz C. EAA/EMQN best practice guidelines for molecular diagnosis of
615 y-chromosomal microdeletions. State of the art 2004. *Int J Androl.* 2004; 27:240-9.
- 616 Skendzel LP, Youden WJ. Systematic versus random error in laboratory surveys. *Am J Clin*
617 *Pathol.* 1970; 54:448-50.

- 618 Smith KN, Nicolas A. Recombination at work for meiosis. *Curr Opin Genet Dev.* 1998; 8:200-
619 11.
- 620 Sun F, Turek P, Greene C, Ko E, Rademaker A, Martin RH. Abnormal progression through
621 meiosis in men with nonobstructive azoospermia. *Fertil Steril.* 2006; 87: 565-71.
- 622 Surtees JA, Argueso JL, Alani E. Mismatch repair proteins: key regulators of genetic
623 recombination. *Cytogenet Genome Res.* 2004; 107:146-59.
- 624 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.
625 Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of
626 multiple internal control genes. *Genome Biol.* 2002; 3:RESEARCH0034.
- 627 World Health Organization. *Laboratory Manual for the Examination of Human Semen and*
628 *Sperm-Cervical Mucus Interaction.* 4th ed. Cambridge: Cambridge University Press, 1999.
629

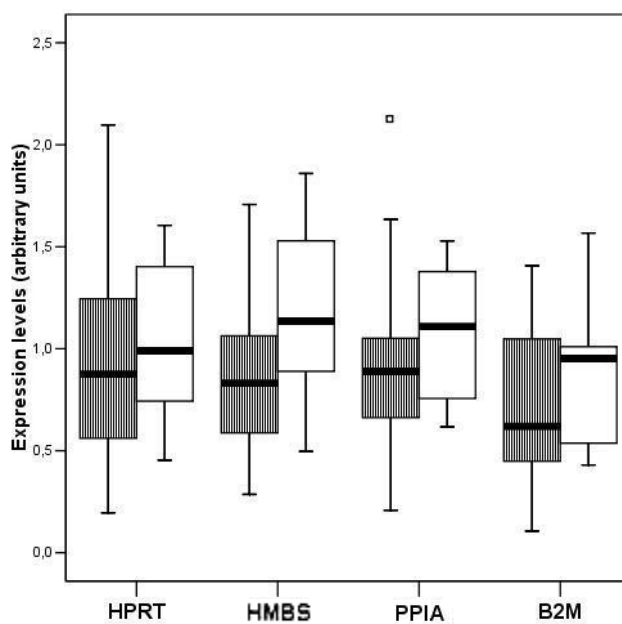
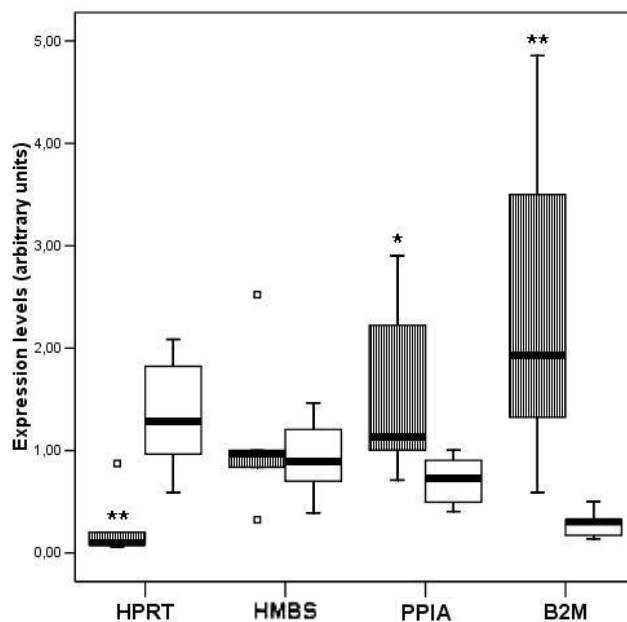
630 Table 1. *Phenotypical and histological description of the testicular samples of the study*

No. patient	Diagnosis	Histology	FSH (U/L)	Semen Sperm conc. (million/ml)	Tubular diameter	Spgonia	Spocyte I	Round Sptid	Elongated Sptid	Sertoli cells	Johnsen score
Patient group 1											
1	SpF (SA)	MA 95%	20.40	0.0	154.10	17.30	8.50	1.00	0.0	8.90	4.80
2	SpF (SA)	MA 100%	8.96	0.0	145.80	16.60	15.30	0.50	0.0	15.00	5.20
3	SpF (SSO)	MA 95%	15.30	0.005	179.50	27.85	35.10	6.65	0.40	13.75	6.20
4	SpF (SSO)	MA >80%	3.60	0.4	196.80	24.80	21.10	1.90	0.50	11.80	5.20
5	SpF (SSO)	MA >80%	n/a	0.004	184.25	21.20	26.15	19.10	0.35	17.05	6.85
6	SpF (SSO)	MA >80%	13.30	3.5	n/a	n/a	n/a	n/a	n/a	n/a	n/a
7	SpF (SSO)	HS hom	n/a	0.5	184.50	14.50	21.20	10.40	5.90	9.20	7.80
8	SpF (SSO)	HS hom	15.60	0.009	156.70	26.00	45.90	28.20	1.50	18.90	7.70
9	SpF (SSO)	HS hom	3.98	3.0	158.70	20.20	31.10	22.50	5.10	17.00	8.10
10	SpF (SSO)	HS hom	3.00	5.0	205.00	20.20	28.60	15.70	0.10	23.30	6.60
11	SpF (SSO)	HS hom	7.18	0.08	188.00	18.50	33.50	20.40	6.70	19.00	8.30
12	SpF (SSO)	HS mix	14.20	0.5	190.95	18.80	6.60	2.00	1.30	12.75	5.75
13	SpF (SSO)	HS hom	3.17	0.007	182.50	16.10	16.45	11.35	4.85	18.30	7.40
Patient group 2											
14	GCT	CSem	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
15	GCT	EC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
16	GCT	EC	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
17	GCT	MX	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
18	GCT	MX	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
Control group											
19	OA	CS	3.60	0.0	176.25	20.25	29.75	27.00	4.50	6.00	7.50
20	OA	CS	5.96	0.0	193.75	21.95	28.85	18.90	19.65	11.95	9.30
21	OA	CS	2.10	n/a	209.40	21.15	34.65	24.00	24.15	15.05	9.20
22	OA	CS	6.60	0.0	247.80	26.80	32.00	22.05	18.40	13.55	8.90
23	CBAVD	CS	1.89	0.0	192.50	25.55	42.25	30.75	30.30	15.75	9.85
24	CBAVD	CS	3.50	0.0	n/a	n/a	n/a	n/a	n/a	n/a	n/a
25	CBAVD	CS	5.90	0.0	197.75	20.10	27.35	29.80	19.10	11.85	9.00
26	CBAVD	CS	4.30	0.0	192.50	18.65	31.30	22.50	22.20	12.00	9.45
27	CUAVD	CS	6.40	0.0	220.15	25.70	38.15	23.30	23.95	15.60	9.30
28	CUAVD	CS	5.35	0.0	194.75	14.75	29.75	12.50	22.25	8.35	8.65

631
632
633 The mean number of the different type of cells per tubule is given in each group.
634 Abbreviations: Spgonia: spermatogonia; Spocyte: spermatocyte, Sptid: spermatid, SpF: spermatogenic
635 failure, SA: secretory azoospermia, SSO: severe secretory oligozoospermia, GCT, germ-cell tumor; OA,
636 obstructive azoospermia; CBAVD, congenital bilateral absence of the vas deferens; CUAVD, congenital
637 unilateral absence of the vas deferens. MA, maturation arrest; HS hom, homogeneous hypospermatogenesis;
638 HS mix: mixed hypospermatogenesis CSem, classic seminoma; EC, embryonic carcinoma; MX, mixed
639 germ-cell tumor (80% embryonic carcinoma; 20% classic seminoma); CS, conserved spermatogenesis.
640

641 **Figure Legends**

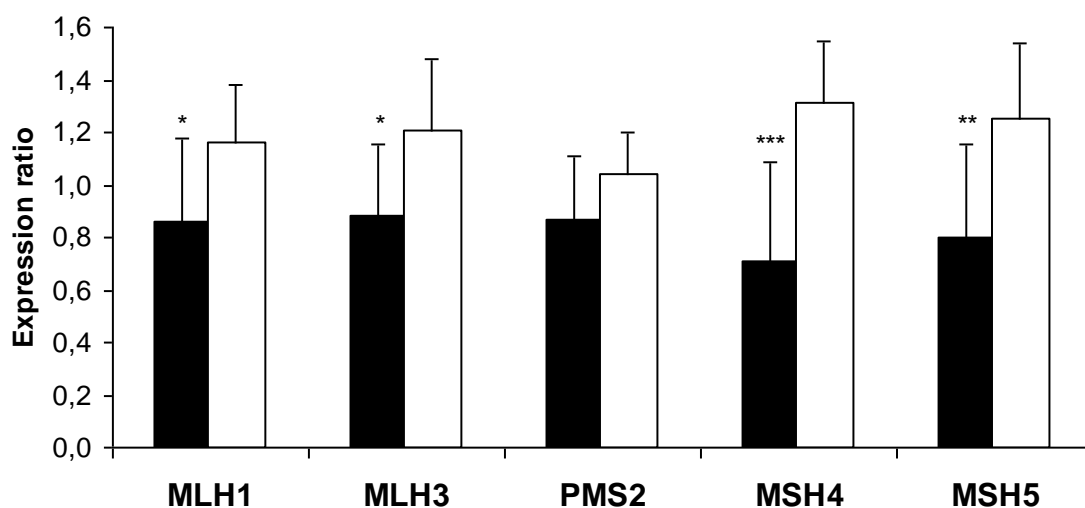
642 Figure 1. **a.)** Absolute expression levels of reference genes in patient group 1 (striated boxes) and
 643 control group (white boxes). □, outlying value (*B2M* outlying value in patient group 1, 5.87, is not
 644 represented). Non-significant differences were observed between the two groups for all genes **b.)**
 645 Absolute expression levels of reference genes in patient group 2 (striated boxes) and control
 646 group (white boxes). □, outlying values. Significant differences are indicated by *asterisks* (*,
 647 $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$)

648 **a.**666 **b.**

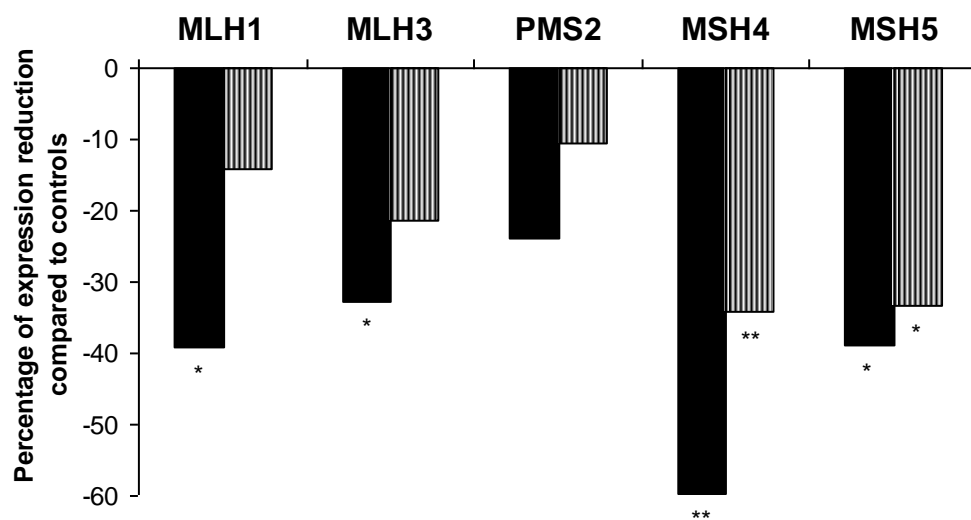
668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680

681 Figure 2. **a.)** Expression ratios of MMR genes in patients with spermatogenic failure (black bars)
 682 and controls (white bars) using NF as normalizer. Significant differences are indicated by
 683 *asterisks* (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$) **b.)** Percentage of expression reduction of MMR
 684 genes in maturation arrest (black bars) and hypospermatogenesis (striated bars) subgroups
 685 compared to controls. (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$)

686 **a.**

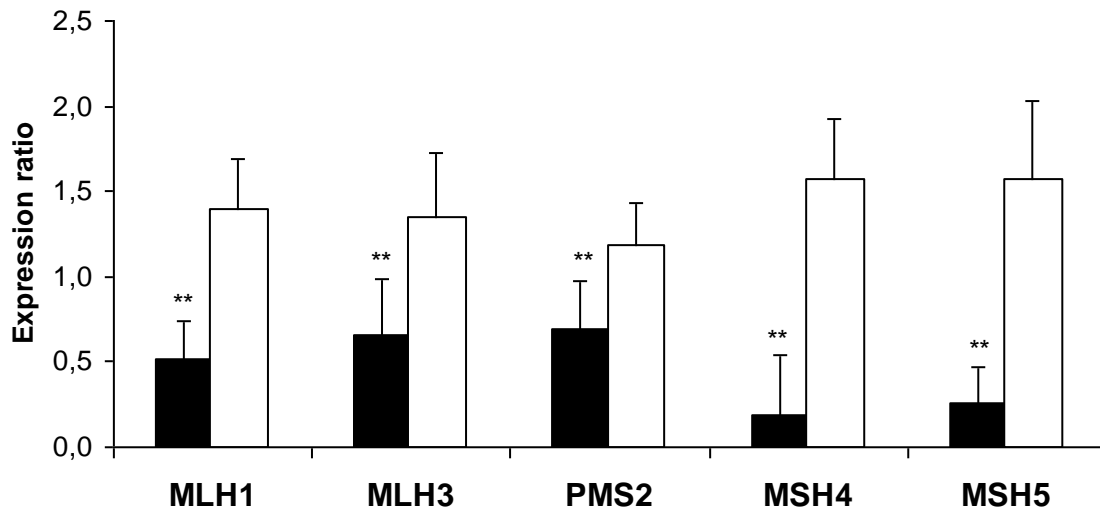


702 **b.**



719

720 Figure 3. Expression ratios of MMR genes in patients diagnosed with germ-cell tumor (black
721 bars) and controls (white bars) using *HMBS* as normalizer. Significant differences are indicated
722 by asterisks (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$)

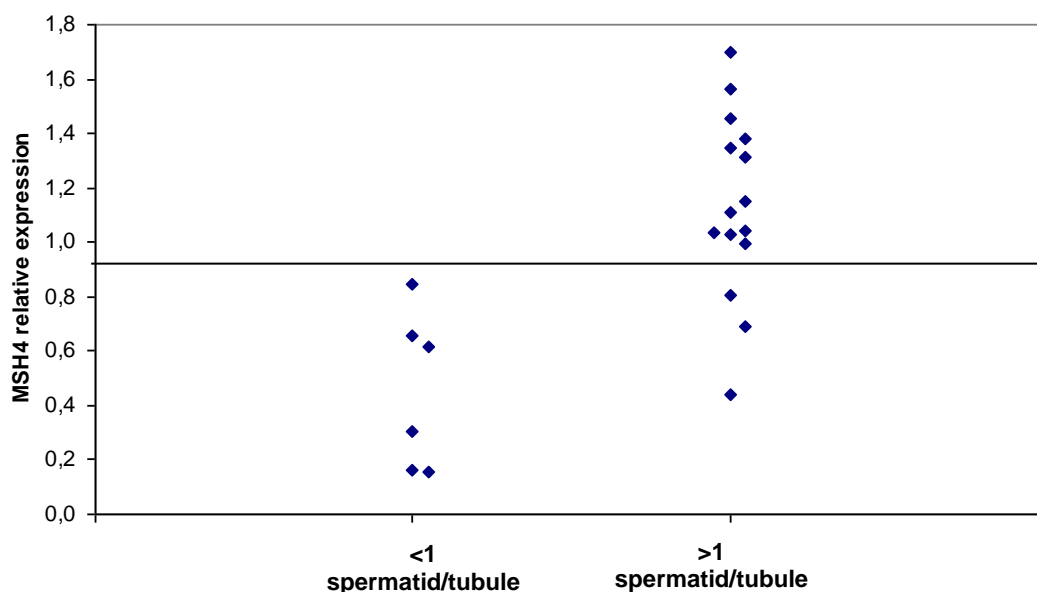


733 Figure 4. **a.)** Table: Pearson correlation coefficients and adjusted p -values (r ; p) between the
 734 expression ratios of the target genes and the different histological parameters for all the samples
 735 analyzed. Significant differences ($p < 0.05$) are indicated in bold. sp: spermatid, JS: Johnsen score.
 736 **b.)** Testicular *MSH4* expression ratio as a marker for spermatogenesis. *Horizontal line* indicates
 737 the *MSH4* transcript ratio threshold value -0.917- that predicts the presence of testicular
 738 elongated spermatid with a sensitivity of 80% and specificity of 100% (ROC curve analysis).

739 **A.**
 740

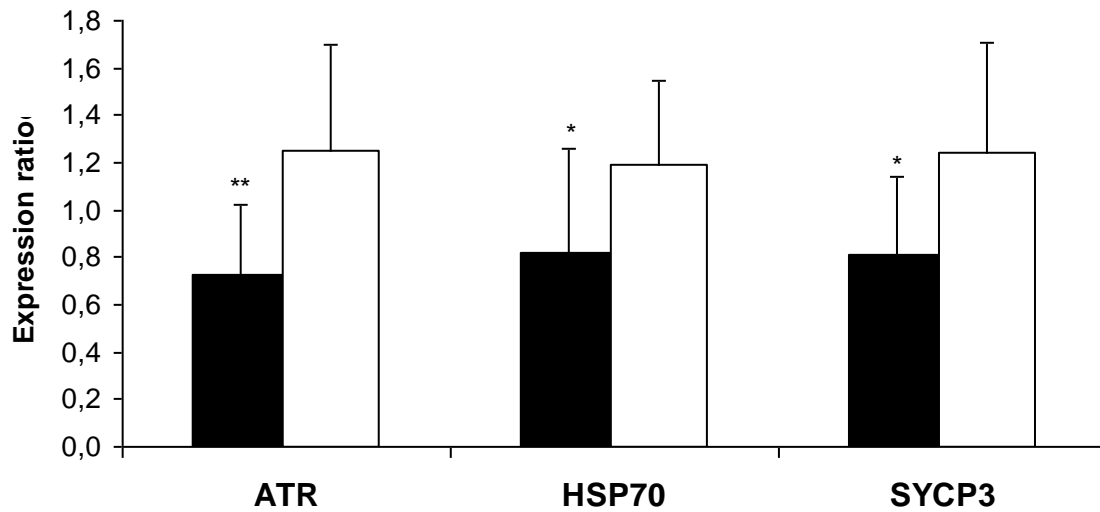
	Tubular Diameter	Spermatogoniae	Spermatocyte I	Round sp	Elongated sp	Sertoli cells	Total cell number	JS
MLH1	0.313; $p=0.166$	0.213; $p=0.354$	0.715; $p=0.000$	0.735; $p=0.000$	0.697; $p=0.000$	0.270; $p=0.236$	0.816; $p=0.000$	0.817; $p=0.000$
MLH3	0.467; $p=0.033$	0.010; $p=0.964$	0.498; $p=0.022$	0.411; $p=0.064$	0.744; $p=0.000$	0.014; $p=0.953$	0.583; $p=0.006$	0.746; $p=0.000$
PMS2	0.419; $p=0.059$	0.108; $p=0.643$	0.186; $p=0.420$	0.379; $p=0.091$	0.467; $p=0.033$	0.122; $p=0.598$	0.363; $p=0.078$	0.461; $p=0.035$
MSH4	0.448; $p=0.042$	0.184; $p=0.424$	0.731; $p=0.000$	0.785; $p=0.000$	0.815; $p=0.000$	0.090; $p=0.700$	0.851; $p=0.000$	0.923; $p=0.000$
MSH5	0.482; $p=0.027$	0.253; $p=0.268$	0.689; $p=0.001$	0.526; $p=0.014$	0.751; $p=0.000$	-0.035; $p=0.881$	0.717; $p=0.000$	0.729; $p=0.000$

741 **B.**
 742

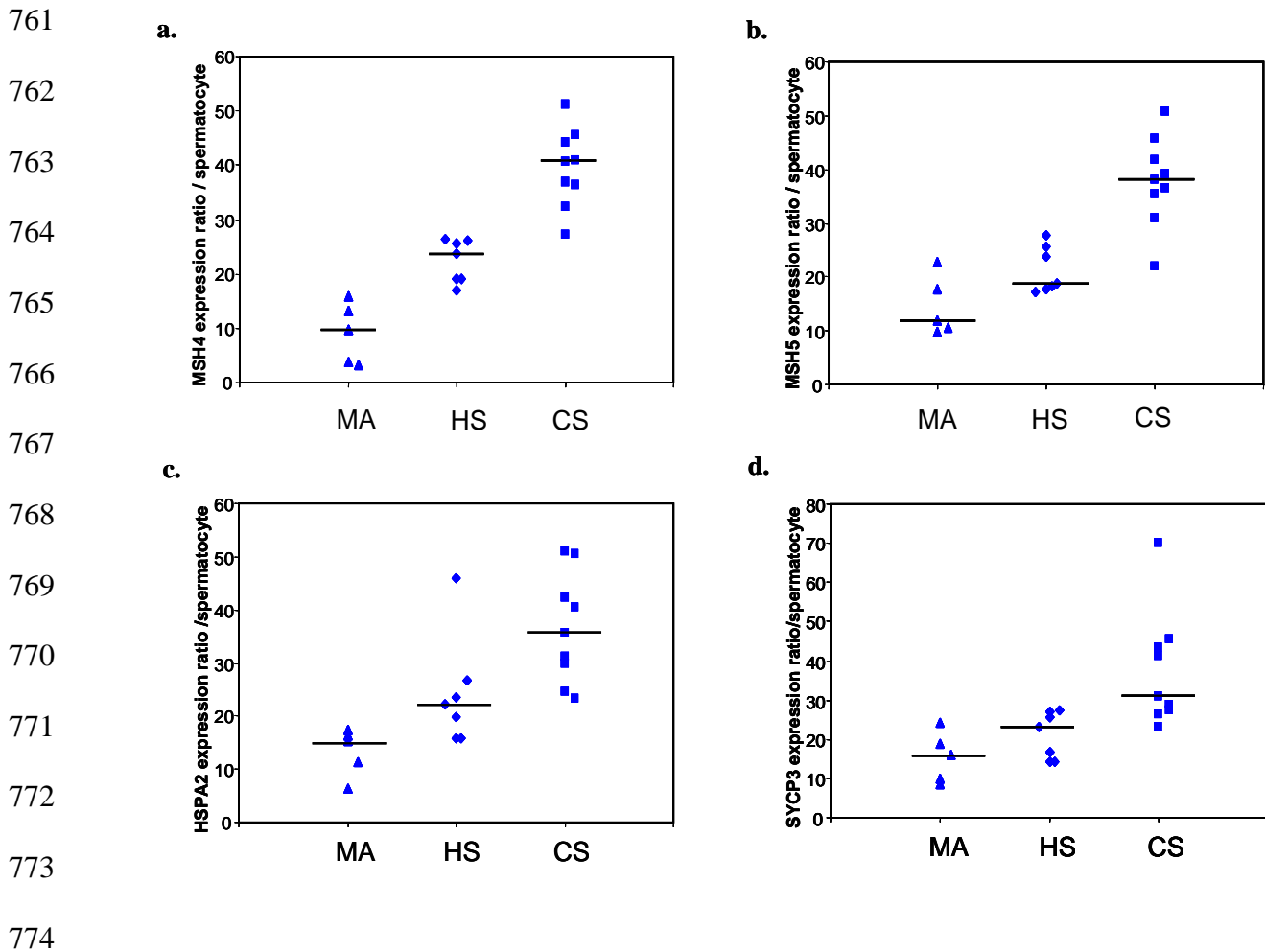


743
 744

745 Figure 5. Expression ratios of *ATR*, *HSPA2* and *SYCP3* control genes in patients with
746 spermatogenic failure (black bars) and controls (white bars) using NF as normalizer. Significant
747 differences are indicated by *asterisks* (*, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$)

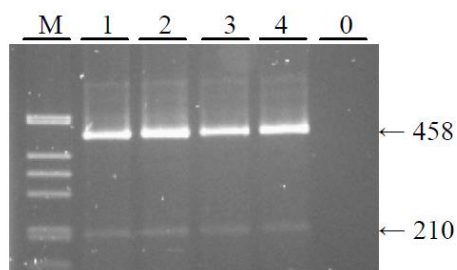


757 Figure 6. *MSH4* (a.), *MSH5* (b.), *HSPA2* (c.) and *SYCP3* (d.) expression ratio per spermatocyte
758 (x1000) of different testicular histological groups. Maturation arrest (filled triangle),
759 hypospermatogenesis (filled rhombus) and control (filled square) samples. *Horizontal lines*
760 indicate median values.

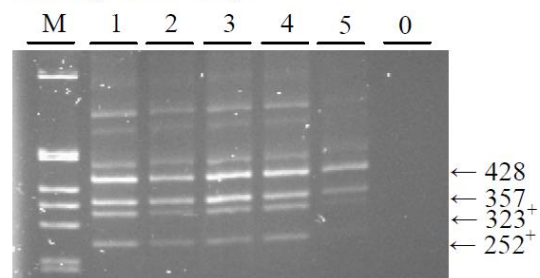


775 Supplemental Figure 1. RT-PCR analysis of alternative MMR transcripts in human testis on (A.)
776 a 2.5% agarose gel (nusieve/agarose 3:1) and on (B.) a 6% polyacrylamide gel. Primers for
777 cDNA amplification and description of variants corresponding to the different amplicon sizes are
778 detailed in Supplemental Table 1. *M*: molecular weight marker, *lanes 1* and *2*: CS testicular
779 samples, *lane 3*: HS testicular sample, *lane 4*: MA testicular sample, *lane 5*: lymphocytes sample
780 and *lane 0*: water (negative control). ⁺ Novel described transcript variant.

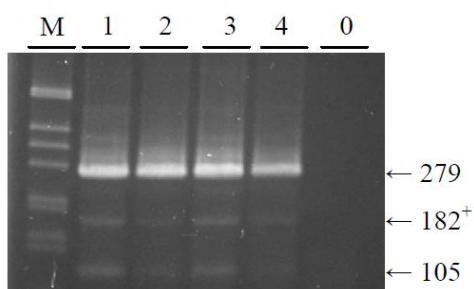
A. MLH1 (Exons 8 → 12)



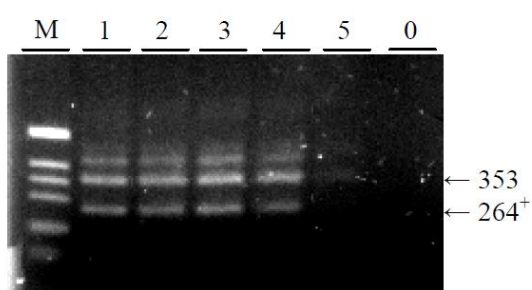
MLH3 (Exons 3 → 8)



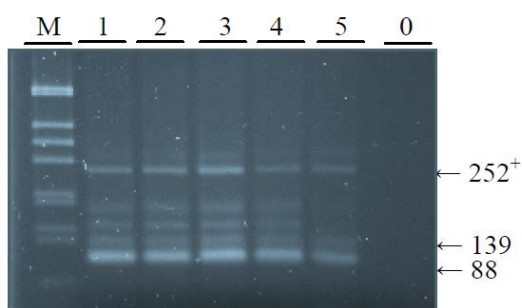
MSH4 (Exons 5 → 7)



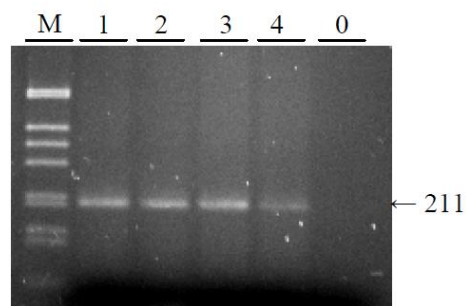
MSH4 (Exons 17 → 20)



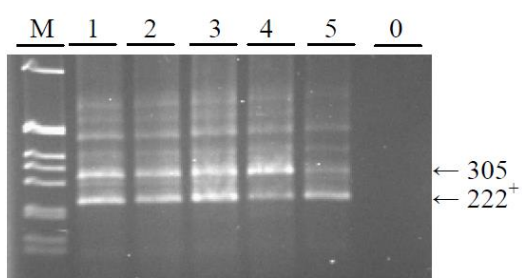
MSH5 (Exons 6 → 7)



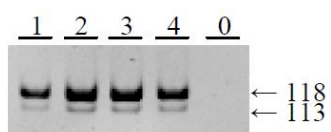
MSH5 (Exons 1a → 3)



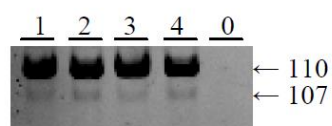
MSH5 (Exons 6 → 11)



B. PMS2 (5' UTR → exon 2)



MSH5 (Exons 20 → 21)



782 Supplemental Table 1. *Primers used for conventional and real-time PCR experiments*
783

Primers	Sequence	NCBI Gene ID	Exon spanning	Amplicon sizes (bp); transcript variants identified	PCR efficiency
<i>Conventional PCR amplification</i>					
MLH1					
Forward	5' TTC GCT CCA TCT TTG GAA ATG 3'	4292	8 → 12	458; MLH1	-
Reverse	5' CGA GGT CAG ACT TGT TGT GG 3'			210; MLH1 (-Exons10,11)	
MLH3					
Forward	5' ACC TTG TTC TTC CTT TCC TTC 3'	27030	3 → 8	428; MLH3 variant 1	-
Reverse	5' TTG TGC CTG TTG CTT CTC GT 3'			357; MLH3 variant 2 323; MLH3 (-Exon5) ⁺ 252; MLH3 (-Exons5,7) ⁺	
PMS2					
Forward	5' GGA GGG AAC TTT CCC AGT C 3'	5395	5' UTR → 2	118; PMS2 variant 1	-
Reverse	5' GAT GGA CTG ACT TCC GAT CAA 3'			113; PMS2 variant 2	
MSH4					
Forward	5' AGC AGA ATT CAG CAC TGT CCT 3'	4438	5 → 7	279; MSH4	-
Reverse	5' CGA AGT CGT CTA CTC CCT CCA 3'			182; MSH4 (-97pb Exon 6) ⁺ 105; ΔhMSH4	
MSH4					
Forward	5' GAA GAA GGT ATT GGC ATT TGT 3'	4438	17 → 20	353; MSH4	-
Reverse	5' TGT CTT TCC ATC TCA GGG GT 3'			264; MSH4 (-Exon19) ⁺	
MSH5					
Forward	5' ATT CCC TTT GAC TGC CTC CT 3'	4439	6 → 7	88; MSH5	-
Reverse	5' GGG GAC GCT GAC ATT ATA GT 3'			139; MSH5 variant 1 252; MSH5 (+intron6) ⁺	
MSH5					
Forward	5' TAG ACG CCA TCT TCA CAC GA 3'	4439	20 → 21	110; MSH5	-
Reverse	5' GTG GCA TTG TTC ACT GCT TTC 3'			107; MSH5 variant 2	
MSH5					
Forward	5' GGC GTT CTC CCA CCT GTA G 3'	4439	1a → 3	211; MSH5 variant 4	-
Reverse	5' ATT CCA CAG CAC ACA CAG ATG 3'				
MSH5					
Forward	5' ATT CCC TTT GAC TGC CTC CT 3'	4439	6 → 11	305; MSH5	-
Reverse	5' CAT GAG TCG GAC GTG TGA AC 3'			222; MSH5 (-Exon9) ⁺	
<i>Real-time PCR amplification</i>					
MLH1					
Forward	5' CTT CAC CCA GAC TTT GCT AC 3'	4292	11-12 → 13	421	1.72
Reverse	5' TTC CAC CAT TTC CAC ATC AGA 3'				
MLH3					
Forward	5' CGG TAG AAG ATG CCA CAG GT 3'	27030	2 → 3	312	1.68
Reverse	5' GAA GGA AAG GAA GAA CAA GGT 3'				
PMS2					
Forward	5' TCA GCA GGC ATC CGT GTA AG 3'	5395	6 → 8	283	2.07
Reverse	5' ACT GTC TGT CTG TTG AAC TCC 3'				
MSH4					
Forward	5' GCT TCA TCC TCA TCT GCG A 3'	4438	2-3 → 6	504	1.73
Reverse	5' GGC TGT CTG TTC ACT ACC C 3'				
MSH5					
Forward	5' GCG ACT GGC AGG TTC TCT AC 3'	4439	12 → 15	294	1.95
Reverse	5' CCA GAT TCT CCA GCT CCT TG 3'				
ATR					
Forward	5' ATG TTT GAA GAC GGT GTG CTC 3'	545	4 → 5	323	2.10
Reverse	5' TTA GAA GGG TTT AGA GAC GAG 3'				
HSPA2					
Forward	5' TGG TAG TGC CCG TGG TGC TT 3	3306	1 → 1	286	1.66
Reverse	5' GAT GGT GTT GGT GGG GTT CA 3'				
SYCP3					
Forward	5' GAT GTT ATT GAA GGG AAG ACT 3'	50511	2-3 → 6	271	2.30
Reverse	5' AAA TCC CAC TGC TGA AAC AAA G 3'				
HPRT					
Forward	5' ATT CTT TGC TGA CCT GCT G 3'	3251	3 → 6-7	268	1.76
Reverse	5' GCT TGC GAC CTT GAC CAT C 3'				
HMBS*					
Forward	5' AAC GGC GGA AGA AAA CAG 3'	3145	1-2 → 4-5	190	1.92
Reverse	5' TCC AAT CTT AGA GAG TGC A 3'				
PPIA*					
Forward	5' CTC CTT TGA GCT GTT TGC AG 3'	5478	1-2 → 5	325	1.81
Reverse	5' CAC CAC ATG CTT GCC ATC C 3'				
B2M					
Forward	5' CCA GCA GAG AAT GGA AAG TC 3'	567	2 → 3	269	2.02
Reverse	5' GAT GCT GCT TAC ATG TCT CG 3'				

784
785 * Primers for *HMBS* and *PPIA* cDNA amplification were previously described (Neuvians et al, 2005;
786 Pluvinet et al, 2004)
787 + Novel described transcript variant.
788

789 Supplemental Table 2. *p-values obtained while comparing*
 790 *MMR gene expression ratios between patient group 1 and*
 791 *controls using each single reference gene and NF as*
 792 *normalizers*
 793

	HPRT	HMBS	PPIA	B2M	NF
MLH1	0.003*	0.257	0.021*	0.208	0.030*
MLH3	0.026*	0.166	0.002*	0.067	0.015*
PMS2	0.010*	0.832	0.077	0.077	0.088
MSH4	0.000*	0.012*	0.000*	0.008*	0.000*
MSH5	0.006*	0.036*	0.002*	0.026*	0.004*

794

795 *significant ($p < 0.05$). Mann-Whitney U test.