Validation of endogenous normalizing genes for expression analyses in adult human testis and germ cell neoplasms

T. Svingen*, A. Jørgensen, and E. Rajpert-De Meyts

Department of Growth and Reproduction, Copenhagen University Hospital (Rigshospitalet), Copenhagen DK-2100, Denmark

*Correspondence address. Tel: +45-3545-4330; Fax: +45-3545-6054; E-mail: terje.svingen@regionh.dk or erm@rh.dk

Submitted on March 3, 2014; resubmitted on April 1, 2014; accepted on April 14, 2014

ABSTRACT: The measurement of gene expression levels in cells and tissues typically depends on a suitable point of reference for inferring biological relevance. For quantitative (or real-time) RT–PCR assays, the method of choice is often to normalize gene expression data to an endogenous gene that is stably expressed across the samples analysed: a so-called normalizing or housekeeping gene. Although this is a valid strategy, the identification of stable normalizing genes has proved challenging and a gene showing stable expression across all cells or tissues is unlikely to exist. Therefore, it is necessary to define suitable normalizing genes for specific cells and tissues. Here, we report on the performance of a panel of nine commonly employed normalizing genes in adult human testis and testicular pathologies. Our analyses revealed significant variability in transcript abundance for commonly used normalizers, highlighting the importance of selecting appropriate normalizing genes as comparative measurements can yield variable results when different normalizing genes are employed. Based on our results, we recommend using RPS20, RPS29 or SRSF4 when analysing relative gene expression levels in human testis and associated testicular pathologies. OCT4 and SALL4 can be used with caution as second-tier normalizers when determining changes in gene expression in germ cells and germ cell tumour components, but the relative transcript abundance appears variable between different germ cell tumour types. We further recommend that such studies should be accompanied by additional assessment of histology and cellularity of each sample.

Key words: testis / germ cell cancer / qRT–PCR / reference genes / human

Introduction

Over the past two decades, quantitative RT–PCR (qRT–PCR) has become a popular method to determine the relative mRNA abundance of selected genes in tissues and cells. Constant technological improvements have made qRT–PCR applications more cost-efficient and reliable and hence more commonplace in research laboratories. Despite great improvements though, the most common qRT–PCR assays still rely heavily on upstream parameters, including input RNA handling, pre-PCR experimental design and a reliable way of normalizing input RNA levels across samples (Bustin and Nolan, 2004; Nolan et al., 2006).

The most frequent protocol used for determining RNA abundance is perhaps the semi-quantitative comparable cycle threshold (Ct), or \(2^{(\text{Ct})}\) method (Livak and Schmittgen, 2001; Schmittgen and Livak, 2008), which also adjusts for variability in input RNA levels between samples. This method allows for comparison of relative expression levels in different tissues, but does rely on certain, sometimes neglected, assumptions. In particular, (i) the PCR product is doubled for each cycle and (ii) the chosen normalizing gene(s) is stably expressed across all the biological replicates (Schmittgen and Livak, 2008).

Optimal PCR cycling efficiency is achieved by careful primer design, and some deviation from 100% amplification efficiency will not significantly affect the normalized expression data (Bookout and Mangelsdorf, 2003). Choosing a stably expressed reference gene is more difficult and using one that is not stable across samples can significantly affect the output (Tricarico et al., 2002; Bas et al., 2004; van den Bergen et al., 2009). This potential pitfall has steadily attracted more attention and numerous studies have identified suitable endogenous normalizing genes in specific cells and tissues. Unfortunately, many tissues have still not received this attention and the selection of normalizing genes is often based on anecdotal rather than empirical evidence.

For the human testis and associated testicular pathologies, endogenous reference genes have not been fully characterized. Since the testis comprises numerous different cell types including interstitial Leydig cells, seminiferous epithelium made up of peritubular and Sertoli cells, germ cells at both diploid and haploid stages and others (Svingen and Koopman, 2013), identifying a stable reference gene is a challenge. This is further complicated when comparing gene expression at various life stages or in disease, as the relative cellularity changes significantly during development and even more so under pathological...
conditions. Testicular germ cell tumours (TGCTs) are very heterogeneous and may largely comprise neoplastic germ cells (seminomas) but may also contain embryonic or somatic elements (embryonal carcinoma or teratoma).

An earlier study failed to identify stable reference genes in adult human testis and seminoma (SE) (Neuviens et al., 2005). For instance, many genes were up-regulated in SE relative to normal testis, including genes encoding several ribosomal proteins, GAPDH (glyceraldehyde-3-phosphate dehydrogenase) and ACTB (beta-actin), whereas UBC (ubiquitin C) was down-regulated. More recently, analyses of fetal human testis suggested that SDHA (succinate dehydrogenase complex subunit A, flavoprotein), TBP (TATA box-binding protein), PMM1 (phosphomannomutase 1) and B2M (β2 microglobulin) were expressed at relatively stable levels during fetal development (O’Shaughnessy et al., 2011). Interestingly, Sdha and Tbp have also been verified as good endogenous reference genes in fetal mouse gonads (Svingen et al., 2009; van den Bergen et al., 2009). However, it cannot simply be inferred from these observations that the genes are also stably expressed in adult testis or testicular pathologies. Finally, a third study on human testis, which compared the expression stability of reference genes in adult testis with that in ejaculate, showed that a few genes including ACTB and GAPDH varied between samples, but not to an extent that excluded them from being used as endogenous reference genes (Cavalcanti et al., 2011). However, interpreting these data is somewhat difficult as the cDNA samples were synthesized from a low and varying amount of input RNA.

Based on the obvious lack of empirically verified endogenous normalizing genes in adult human testis and testicular pathologies, we aimed to identify suitable genes for use in future studies. We selected nine putative normalizing genes in adult human testis and testicular pathologies, we aimed to input RNA. as the cDNA samples were synthesized from a low and varying amount of RNA.

**Materials and Methods**

**Human tissue samples**

The use of samples from the human tissue archive at the Copenhagen University Hospital (Permit number: H-I-2012-007) was approved by the regional Danish Committee for Medical Research Ethics. Tissue samples were obtained from orchidectomized testes with tumours, snap frozen in liquid nitrogen and stored at −80°C and included ‘normal’ testis tissues (normal morphology and complete spermatogenesis located in a testis with a tumour); three each of testis specimens containing (i) mixture of tubules with carcinoma in situ and normal tubules (NT/CIS), (ii) ~100% tubules with CIS (by histological assessment, all tubules found to contain CIS cells and no spermatogenesis or invasive tumour cells, although normal tubules outside the examined area cannot be excluded), (iii) embryonal carcinoma (EC), (iv) SEM, (v) Sertoli cell-only (notably from testis diagnosed with germ cell tumour) and (vi) teratoma. Of note, the NT/CIS category denotes testis tissue adjacent to TGCTs obtained from orchidectomy specimens and the relative percentage value of CIS to normal testis can vary considerably.

**RNA extraction and cDNA synthesis**

Total RNA was extracted from frozen tissues using the NucleoSpin RNA II purification kit (Macherey-Nagel, Düren, Germany) as per manufacturer’s recommendations. An additional two human adult testis RNA samples were purchased from Ambion (Life Technologies, Naerum, Denmark) and BioChain (Odense, Denmark). Before qRT–PCR synthesis, RNA quantity and quality were determined by Nano-drop spectrophotometry and Qubit RNA assays (Invitrogen, Naerum, Denmark). Only samples with corresponding RNA concentrations between the two assays were included in this study (eight samples were excluded).

A total of 1 μg RNA was used for the synthesis of cDNA as follows: an initial reaction volume of 10 μl containing 1 μg RNA in nuclease-free (DEPC-treated) H₂O and 0.5 μl primer mix (1 μg/μl dT20:1 μg/μl random hexamers at 5:1 ratio) was primed at 65°C for 1 min, then on ice for 1 min and 37°C for 1 min. Then 10 μl cDNA synthesis mix containing 4 μl 5X cDNA buffer (650 mM Tris—HCl, pH 8.3; 25 mM MgCl₂; 100 mM KCl), 0.5 μl 100 mM dNTP mix, 0.3 μl AMV reverse transcriptase (Affymetrix, USB product, Santa Clara, USA), and 5.2 μl DEPC-treated H₂O was added to the reaction and left at 42°C for 1 h. Enzyme reaction was inactivated by the addition of 80 μl 0.1% Triton-X before incubation at 95°C for 1 min. cDNA samples were stored at − 80°C.

**qRT–PCR and primer analyses**

qRT–PCR assays were performed with SYBR Green (Stratagene, Agilent Technologies, Waldbronn, Germany) on a Stratagene Mx3000P QPCR System in 20 μl reactions. Each reaction consisted of 10 μl 2x SYBR Green Master Mix, 1 μl cDNA, final concentration of 150 nM each forward and reverse primers and ddH₂O. Transcript levels were measured with two-step thermal cycling conditions: initiation at 95°C for 15 min, then 40 cycles of 95°C for 15 s and 60°C for 60 s, followed by dissociation curve measurements. Gene-specific primers were designed using the Universal Probe library tool (Roche Applied Sciences) and the NCBI Primer designing tool and are listed in Table I.

In silico analysis was performed to design primers with similar annealing temperatures (Tm ~60°C), without secondary structure formation, amplifying products of <200 nucleotides and spanning at least one intron—exon boundary. OCT4 primers do not span an intron as it was not possible to design OCT4 primers fulfilling all criteria. Therefore, further care should be made to ensure that cDNA is free of genomic DNA contamination, as amplification of this would not be avoided by rapid PCR cycling conditions nor be detected by melt curve analysis. Here, we performed routine checks of no-reverse transcriptase (−RT) cDNA reactions yielding no specific amplification. Following size verification by agarose gel electrophoresis and nucleotide sequencing of the amplicons (data not shown), standard curve analyses were performed to determine amplification efficiencies (E) of the primer pairs by the formula $E = 10^{(1/\text{slope})}$ converted to percentage-efficiency by $(E - 1) \times 100\%$ (Rasmussen, 2001). Standard curves were performed in triplicate reactions over five serial dilutions (1:10) of cDNA spiked with trace amounts of primer-specific amplicons. The slopes of the standard curve measurements were determined by the Stratagene software and are given in Table II together with primer pair efficiencies. All primer pairs had amplification efficiencies deemed acceptable for further analyses by the 2−ΔCt method (Livak and Schmittgen, 2001). Individual melt (or dissociation) curve analyses revealed single peaks > 75°C indicating specific amplification of a single product (Supplementary data, Fig. S1).

**Analytical methods**

The Stratagene Mx3000 software was used for the acquisition of qRT–PCR Ct values, and analyses of dissociation and standard curves. Relative gene expression was calculated using the 2−ΔΔCt method (Applied Biosystems...
Research Bulletin No. 2 P/N 4303859), and statistical significance determined by Student’s two-tailed (unpaired) t-test in Microsoft Excel. A $P$ value of $<0.05$ was considered statistically significant. Expression stability of normalizing genes was analysed using BestKeeper (Technical University of Munich, Germany) and NormFinder (Aarhus University Hospital Molecular Diagnostic Laboratory, Denmark) softwares.

BestKeeper make calculations at the Ct level to allow for parametric tests on normally distributed expression levels. It determines the coefficient of variance for each gene, gives the geometric mean of Ct values and calculates a Pearson coefficient of correlation ($r^2$), with corrections by PCR efficiency ($E$) values possible. Exclusions can be made by SD calculations and finally a weighted index of the most suitable normalizing genes across tested biological samples is produced (Pfaffl et al., 2004). Notably, BestKeeper assumes normal distribution of input data (Ct values) that are logarithmic in nature, hence using parametric methods of analysis.

In contrast to BestKeeper, NormFinder uses input data transformed into linear scale to calculate an expression stability value: a combined estimate of intra- and inter-group variations for each putative normalizing gene submitted for analysis. It takes into account likely heterogeneity between biological samples and circumvents bias caused by co-regulation of genes. It also allows for the immediate ranking of best normalizing genes based on expression variance both within and between biological groups (Andersen et al., 2004). However, the NormFinder algorithm does not display systematic inter-group variation for single reference genes and further, if several putative reference genes show similar systematic variation across biological samples, the mean expression will be biased towards the variant gene. This means that a less stable reference gene can be erroneously selected.

**Immunohistochemistry**

Immunohistochemistry (IHC) was carried out on 10 μm sections of tissue snap frozen in isopentane and mounted in Tissue-Tek OCT compound (Sakura, Alphen aan den Rijn, The Netherlands). After mounting on glass slides, sections were dried for 2 h at room temperature, treated with 0.5% $\text{H}_2\text{O}_2$ followed by blocking for 1 h in 2% non-immune goat serum (Zymed histostain kit; Life Technologies, Naerum, Denmark). Primary antibodies were diluted in blocking serum: mouse anti-OCT3/4 (1:100, Santa Cruz sc-5279), mouse anti-D2-40 (1:100, Dako M3619) and mouse anti-SALL4 (1:400, Santa Cruz sc-101147) then added to sections and left for 2 h at room temperature. Slides were incubated with biotinylated secondary goat

<table>
<thead>
<tr>
<th>Gene</th>
<th>Accession no.</th>
<th>Primer sequence (5′–3′)</th>
<th>Size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalizing genes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2M</td>
<td>NM_004048</td>
<td>F: AGTATGCGCTTGGGTTGGAAC</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: ATCCGAATCCTAAAGCCGATTC</td>
<td></td>
</tr>
<tr>
<td>GAPDH</td>
<td>NM_002046</td>
<td>F: TCAACGGACCTCTTATCAAGC</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: GGTGTTCAGAGGGTCTTACTC</td>
<td></td>
</tr>
<tr>
<td>PMMI1</td>
<td>NM_002676</td>
<td>F: CAGAAGCTACGAAATGAGTG</td>
<td>123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CCCTCTCGGGCACACACATA</td>
<td></td>
</tr>
<tr>
<td>RPLP0</td>
<td>NM_053275</td>
<td>F: TCTCAACCCCTGAGTTCTTGT</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CAAATCCGAGAGCCAGAATG</td>
<td></td>
</tr>
<tr>
<td>RPS20</td>
<td>NM_001023</td>
<td>F: CGCTCTGCTGTCGTTTCCA</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CCTTGGCTACTGACCGGA</td>
<td></td>
</tr>
<tr>
<td>RPS29</td>
<td>NM_001032</td>
<td>F: AACAAGCCCAAGTAAAAATC</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: ACAGATCCACAGGTTTAAATT</td>
<td></td>
</tr>
<tr>
<td>SDHA</td>
<td>NM_004168</td>
<td>F: GACCATGAAAGTACATGCAAG</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: GCTTCCATACGCAAACCTCAT</td>
<td></td>
</tr>
<tr>
<td>SRSF4</td>
<td>NM_005626</td>
<td>F: CTAAGAGTAGATCTCTGAGTT</td>
<td>104</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: GGAGCTCTTCTTCTTCT</td>
<td></td>
</tr>
<tr>
<td>TOP1</td>
<td>NM_003286</td>
<td>F: AAAGTTCAGATATTTGGCCCAAC</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: ATTCATGGTGCGACATTTTG</td>
<td></td>
</tr>
<tr>
<td>Marker genes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYP51A1</td>
<td>NM_000786</td>
<td>F: ATGCAGATTTGGATGGAGGTT</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CCGTATCTCCGGATGAGG</td>
<td></td>
</tr>
<tr>
<td>DDX4</td>
<td>NM_024415</td>
<td>F: AAGGCTTGGGAGAGCCAGAAG</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: AACATAGGAAGACATATGAGGGTTG</td>
<td></td>
</tr>
<tr>
<td>GATA4</td>
<td>NM_002052</td>
<td>F: GGAGCCCAAGAAGCTGAAATAA</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: GTTCTCTGAGTTTCTGGAAG</td>
<td></td>
</tr>
<tr>
<td>MMP2</td>
<td>NM_004530</td>
<td>F: CCCCCAAGGCCAACAGAGAG</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CTTCAACAGCAAAACAGGTGGT</td>
<td></td>
</tr>
<tr>
<td>OCT4</td>
<td>NM_002701</td>
<td>F: TACCTACTCCTGCTTTTCCC</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: CAAAAAACCGTGACCACAACT</td>
<td></td>
</tr>
<tr>
<td>SALL4</td>
<td>NM_020436</td>
<td>F: CCGCACCTGAATGGAAGGT</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td></td>
<td>R: AAGTTCAGAGGCTCCACAA</td>
<td></td>
</tr>
</tbody>
</table>

Table I Gene-specific primer sequences and corresponding amplicon sizes corresponding to human sequences obtained at NCBI.
anti-mouse immunoglobulin G (Invitrogen, Naerum, Denmark), followed by
colour development with aminoethyl-carbazole substrate from 4 to 20 min
depending on antibody and counterstained with Meyer’s haematoxylin.
Between incubation steps sections were washed in Tris-buffered saline.
Images were captured using a NanoZoomer 2.0 HT scanner (Hamamatsu
Photonics, Herrsching am Ammersee, Germany).

Estimation of cellular composition in frozen
tumour specimens

Frozen tissues were sectioned at 10 μm with four sections collected and
marked ‘front’. The tissue sample was released from the block, turned
180° and again mounted with OCT before an additional four sections at
10 μm were collected and marked ‘back’. The remainder of the tissue (4–
5 mm thickness) was cleared from OCT and RNA extracted as described
above. The ‘front’ and ‘back’ sections from each tissue were subjected to
IHC using antibodies against PDPN (podoplanin), OCT4 and SALL4, demar-
cating neoplastic cells within the tissue samples, allowing for percentage area
estimation and high correlation (r). The top three ranked genes were
RPS20, SALL4 and OCT4 as reference. A stability
value (S) was calculated for each gene within the 23 replicates and ranked
accordingly. The top three ranked genes were RPS29, SRSF4 and GAPDH,
RPLP0, RPS20 and RPS29) were between 20 and 23 cycles,
whereas the Ct values of genes expressed at low levels (PMM1, SDHA,
SRSF4 and TOP1) were between 26 and 30 cycles (Fig. 1A). When includ-
ing the average Ct values from testicular pathologies, we observed a
similar but slightly broader expression range (Fig. 1B).

To further assess the stability of gene expression across the biological
samples, the raw Ct values were analysed with the BestKeeper software.
Across all 23 biological replicates, the SD was < 1 for 5 of the genes and
> 1 for GAPDH, TOP1, B2M and PMM1, the latter 4 thus falling outside
the exclusion value (Table III). Using pair-wise correlation and regression
analysis to determine inter-gene relations, it was possible to rank the
genes relative to expression stability. Here, genes with the greatest vari-
ability across the biological replicates displayed the highest variation (CV)
and lowest correlation (r), whereas the most stable genes had low vari-
ation and high correlation (r approaching 1). Table III ranks the nine nor-
malizing genes according to these parameters, scoring RPS20, RPS29 and
RPLP0 as the most stably expressed and PMM1 as the least stable.

Next, we analysed the data set with NormFinder (Andersen et al.,
2004). Raw Ct values were converted to a linear scale by the 2^(-ΔCt)
method using the geometric mean of the five most abundantly expressed
genes (B2M, GAPDH, RPLP0, RPS20 and RPS29) as reference. A stability
value (S) was calculated for each gene within the 23 replicates and ranked
accordingly. The top three ranked genes were RPS29, SRSF4 and RPS20
and the lowest ranked genes were PMM1 and SDHA (Fig. 2). To be
deemed relatively stable, a gene should have a S value < 0.5, which
was the case for five of the nine tested genes. Since biological replicates
are heterogeneous, with both cellularity and expression profiles
expected to vary substantially, NormFinder was used to analyse expres-
sion stability between normal testis and specific pathologies (Table IV).
No single gene was very stable across all samples.

To analyse what effect the use of different reference genes can have on
final normalized results, we prepared cDNA from six new samples: three

### Table II Relative gene abundance and primer-specific amplification efficiencies for human testis cDNA.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Average Ct (mean ± s.d.)</th>
<th>Slope</th>
<th>R-value</th>
<th>E-value</th>
<th>Efficiency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalizing genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2M</td>
<td>21.45 ± 0.738</td>
<td>−3.301</td>
<td>0.996</td>
<td>2.01</td>
<td>101</td>
</tr>
<tr>
<td>GAPDH</td>
<td>22.54 ± 0.673</td>
<td>−3.451</td>
<td>0.900</td>
<td>1.95</td>
<td>95</td>
</tr>
<tr>
<td>PMM1</td>
<td>29.93 ± 1.092</td>
<td>−3.472</td>
<td>0.981</td>
<td>1.94</td>
<td>94</td>
</tr>
<tr>
<td>RPLP0</td>
<td>22.25 ± 0.685</td>
<td>−3.391</td>
<td>0.996</td>
<td>1.97</td>
<td>97</td>
</tr>
<tr>
<td>RPS20</td>
<td>20.59 ± 0.368</td>
<td>−3.454</td>
<td>0.997</td>
<td>1.95</td>
<td>95</td>
</tr>
<tr>
<td>RPS29</td>
<td>21.58 ± 0.611</td>
<td>−3.454</td>
<td>0.977</td>
<td>1.99</td>
<td>99</td>
</tr>
<tr>
<td>SDHA</td>
<td>26.45 ± 0.855</td>
<td>−3.346</td>
<td>0.977</td>
<td>1.99</td>
<td>99</td>
</tr>
<tr>
<td>SRSF4</td>
<td>28.48 ± 0.996</td>
<td>−3.386</td>
<td>0.992</td>
<td>2.02</td>
<td>102</td>
</tr>
<tr>
<td>TOP1</td>
<td>38.35 ± 0.843</td>
<td>−3.273</td>
<td>0.982</td>
<td>2.02</td>
<td>102</td>
</tr>
<tr>
<td>Marker genes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYPS1A1</td>
<td>—</td>
<td>−3.164</td>
<td>0.998</td>
<td>2.07</td>
<td>107</td>
</tr>
<tr>
<td>DDX4</td>
<td>—</td>
<td>−3.255</td>
<td>0.999</td>
<td>2.03</td>
<td>103</td>
</tr>
<tr>
<td>GATA4</td>
<td>—</td>
<td>−3.328</td>
<td>0.994</td>
<td>1.98</td>
<td>98</td>
</tr>
<tr>
<td>MMP2</td>
<td>—</td>
<td>−3.273</td>
<td>0.996</td>
<td>2.02</td>
<td>102</td>
</tr>
<tr>
<td>OCT4</td>
<td>—</td>
<td>−3.400</td>
<td>0.998</td>
<td>1.97</td>
<td>97</td>
</tr>
<tr>
<td>SALL4</td>
<td>—</td>
<td>−3.201</td>
<td>0.991</td>
<td>2.05</td>
<td>105</td>
</tr>
</tbody>
</table>
NTs and three ECs, again initially normalized at the input RNA level. We analysed the relative mRNA expression of three genes: SALL4, the cytochrome P450 family member CYP51A1 that is expressed strongly by germ cells (Rozman and Waterman, 1998) and the metalloproteinase member MMP2, also involved in spermatogenesis and reportedly up-regulated in TGCTs (Milia-Argeiti et al., 2012). The three genes were normalized against four different reference genes, SDHA, GAPDH, RPS20 and RPS29, representing low-, medium- and high-ranked normalizers (Fig. 3A). Compared with NT, SALL4 was up-regulated in EC regardless of normalizing gene, but only significantly relative to SDHA and RPS20. The increase in mRNA was nearly 3-fold greater when normalized to SDHA, a normalizer deemed the least stably expressed by our previous analyses.

CYP51A1 was significantly down-regulated in EC when measured by all four normalizing genes, but the fold change in expression varied between strategies. MMP2 was not statistically significantly altered, although a trend of down-regulation was evident when normalized against GAPDH, RPS20 and RPS29. The SDHA strategy suggested a reverse trend.

It has also been suggested that normalization by the comparative Ct method is most reliable when employing the geometric mean of several normalizing genes as opposed to a single reference gene. However, this requires selection of several stably expressed reference genes, as the inclusion of unstably expressed genes for normalization could make the output data less reliable. Nevertheless, to illustrate the point we also analysed the relative transcript abundance of SALL4, CYP5IA1 and MMP2 in EC versus NT as measured when normalized against one, two and three normalizing genes, respectively (Fig. 3B). As evidenced, SALL4 expression was statistically significantly different when normalized against RPS20, but not when normalized against the geometric mean of RPS20 and RPS29. When normalized against the geometric mean of three genes (RPS20, RPS29 and GAPDH), the statistical outcome was the same as for RPS20 only. Changes to statistical values were also observed for CYP5IA1 and MMP2.

A further complication when measuring gene expression levels in multi-cellular tissues, especially those where cellular composition changes over time, is to define whether changes in transcript levels represent real alterations in transcription within individual cells or are a result of relative changes in cell types/numbers within the tissue. One approach to address this is to use a stably expressed cell-specific gene as a second-tier normalizer. We calculated the relative percentage of
neoplastic cells to determine if OCT4, SALL4 or DDX4 could be used as germ cell-/neoplastic cell-specific normalizers in TGCTs. Tissue sections were stained using antibodies against PDPN, OCT4 and SALL4 and specimens were analysed by two antibodies each (Fig. 4A). NT samples were found to contain smaller elements of OCT4-positive neoplastic cells (Fig. 4, arrow), thus not representing pure normal testis, but rather experimentally relevant control tissues. The percentage areas occupied by neoplastic cells of SEM and EC samples were calculated by marking the PDPN-, OCT4- and/or SALL4-positive regions, and the total was added up and averaged between sections from the front

Table IV Expression stability within histology groups according to relative NormFinder stability values.

<table>
<thead>
<tr>
<th>Most stable</th>
<th>NT/CIS versus testis</th>
<th>CIS versus testis</th>
<th>EC versus testis</th>
<th>SEM versus testis</th>
<th>SCO versus testis</th>
<th>TER versus testis</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>GAPDH</td>
<td>SRSF4</td>
<td>TOP1</td>
<td>SDHA</td>
<td>B2M</td>
<td></td>
</tr>
<tr>
<td>RPLP0</td>
<td>RPS20</td>
<td>SRSF4</td>
<td>TOP1</td>
<td>RPLP0</td>
<td>GAPDH</td>
<td></td>
</tr>
<tr>
<td>RPS29</td>
<td>TOP1</td>
<td>RPLP0</td>
<td>RPS29</td>
<td>B2M</td>
<td>RPLP0</td>
<td></td>
</tr>
<tr>
<td>B2M</td>
<td>RPS20</td>
<td>SDHA</td>
<td>RPS29</td>
<td>B2M</td>
<td>TOP1</td>
<td></td>
</tr>
<tr>
<td>Least stable</td>
<td>PMMI</td>
<td>PMMI</td>
<td>PMMI</td>
<td>PMMI</td>
<td>TOP1</td>
<td></td>
</tr>
</tbody>
</table>

NT, ‘normal’ testis; CIS, carcinoma in situ; EC, embryonal carcinoma; SEM, seminoma; SCO, Sertoli cell-only; TER, teratoma.

Figure 3 Effects of different normalizing genes on measured relative mRNA abundance. Differences in relative fold expression of SALL4, CYP51A1 and MMP2 between adult testis with normal tubules (NT) and embryonal carcinoma (EC) when (A) normalized against four reference genes with different stability (SDHA, GAPDH, RPS20 or RPS29) and (B) normalized against a single reference gene (RPS20), against two reference genes (geometric mean of RPS20 and RPS29), and three reference genes (geometric mean of RPS20, RPS29 and GAPDH). Each bar represents mean ± s.e.m. of three biological replicates, with expression levels relative to NT mRNA set at 1. *P < 0.05; **P < 0.01; ***P < 0.001, all NT versus EC by Student’s two-tailed (unpaired) t-test.
and back of the tissues. Average neoplastic areas of CIS specimens were similarly calculated, but as shown by the representative sections in Fig. 4A, PDPN- and OCT4-positive CIS cells only make up a portion of the total cell count within these regions. Thus, according to area calculation, a 50% CIS will contain far fewer neoplastic cells than a 50% SEM specimen. Finally, although there was a general agreement between protein markers, one NT/CIS sample displayed discrepancy in neoplastic percentage with a much lower number obtained with OCT4 (4%) staining compared with PDPN (53%), as the latter also included larger neo-vascular structures. Likewise, one EC sample showed discrepancy in OCT4 (11%) and SALL4 (50%) positive areas, which could represent a tumour that contains more differentiated elements and has down-regulated OCT4 expression but maintained SALL4. As a final note, the front and back sections occasionally varied significantly, meaning that the midpiece used for RNA extraction could potentially vary significantly from our estimate. This highlights not only a weakness of this approach, but a general problem with assessing the cellularity of tumour samples and must be considered when interpreting data.

We performed additional qRT–PCR experiments measuring the relative expression levels of OCT4, SALL4, DDX4 and GATA4 relative to RPS20. X-axis denotes the relative neoplastic component (%) of the individual specimens, as calculated from PDPN, OCT4 and SALL4 IHC staining. Bars represent mean ± s.d. of triplicate reactions of single specimens.
area of neoplastic cells (Fig. 4B). As expected, OCT4 was slightly elevated in the one NT tissue where CIS elements had been observed. OCT4 mRNA remained very low in all CIS samples with an estimated proportion ≤ 4%, except for one sample in which it was significantly elevated (6-fold higher in the 57% CIS compared with the 4% CIS specimen). Surprisingly, the 100% SEM sample displayed a relative OCT4 expression similar to the 57% CIS specimen. Also, the EC specimens, which were estimated to contain around 14% OCT4-positive cells, showed variable mRNA levels within the EC group and also when compared with CIS and SEM percentages. At the mRNA level, the expression pattern of SALL4 was similar to OCT4, again surprisingly low in the 100% SEM sample compared with the other pathologies and NT. Therefore, the relative OCT4 or SALL4 transcript abundance was not directly comparable with relative OCT4 and SALL4 protein expression as estimated by number of positively stained cells across different TGCT pathologies. However, there was some agreement in OCT4 and SALL4 mRNA levels with total number of neoplastic cells within tumour subgroups.

DDX4, a germ cell marker, was significantly down-regulated in the tumour tissues. This renders DDX4 an unsuitable marker in TGCTs. Likewise, GATA4, a marker of Sertoli cells, was down-regulated in TGCT specimens. Interestingly, GATA4 appeared down-regulated in the NT with < 1% CIS compared with the purer NT specimen, potentially indicating a down-regulation in Sertoli cells adjacent to CIS cells. The absence of GATA in the SEM and EC specimens confirmed the loss of Sertoli cells in these pathologies.

**Discussion**

The determination of the absolute or relative expression level of a gene in any tissue typically requires a point of reference. In qRTPCR assays this point of reference is often an endogenously expressed gene that is unaltered in transcript abundance between biological replicates. Employing this strategy will not only correct for smaller variations in input RNA but also give a quantifiable number, at least relative to something known. However, if the reference gene is not stably expressed across samples, the data can be significantly skewed and, in worst-case scenarios, can give rise to flawed conclusions (Dheda et al., 2005).

Here, we have tested the stability of nine commonly employed normalizing genes in adult human testis and testicular pathologies, and ranked them by expression stability based on the statistical algorithms BestKeeper (Pfaffl et al., 2004) and NormFinder (Andersen et al., 2004). These algorithms were devised to overcome the conundrum of measuring the stability of a putative factor without having a stable reference to compare it against in the first place. Although both algorithms have inherent shortcomings because of assumptions such as there being no systematic variation in expression across samples, the combined use of both strengthen the basis for reference gene selection.

We were unable to identify a normalizing gene with only minor variation in expression across the complete panel of testicular pathologies. Given the heterogeneity between biological samples, this is not surprising and in line with previous studies (Neuvians et al., 2005; Cavalcanti et al., 2011). On the other hand, several of the normalizing genes proved to perform within acceptable level of variation, the best being the high-abundance ribosomal genes RPS20 and RPS29 and the low-abundance gene SRSF4.

In contrast to what has been shown for fetal human testis (O’Shaughnessy et al., 2011), our data suggest that B2M and SDHA are not very well suited as normalizing genes for adult testis tissue, particularly in testicular neoplasms. It has been suggested that another gene, TBP, is stably expressed in fetal testis of both mice and humans (Svingen et al., 2009; O’haughnessy et al., 2011); however, we failed to design optimal primers for our assays and chose not to include TBP in our study. Likewise, the X-linked gene HPR1 (hypoxanthine phosphoribosyltransferase 1), a popular normalizing gene, was excluded because a numerical increase in X chromosomes is frequently observed in TGCTs (Peltonäki et al., 1991; Rajpert-De Meyts et al., 2007) and because the gene has been shown to vary greatly between germ cell populations in mice (Svingen et al., 2009; van den Bergen et al., 2009). Finally, we would normally avoid reference genes from the short arm of chromosome 12 when normalizing RNA from TGCTs, as 12p polyplody or segmental amplification is frequently observed in these neoplasms (reviewed by Looijenga et al., 2003). Nevertheless, we included GAPDH as a control and, surprisingly, found it to be relatively stably expressed. We would still caution against the use of GAPDH though, as greater changes in transcript abundance could be encountered in individual TGCT specimens.

The challenge of identifying stable normalizing genes for multi-cellular tissues is even greater when the relative cellularity varies between samples. If a change in gene expression is observed, the question often becomes whether there is a real change within cells, or whether the observed change is simply a result of altered cell ratios within the tissue. To overcome this, researchers often employ a cell-specific gene as a second normalizer. This is a clever strategy, but it requires that this second pseudo-normalizing gene is also stably expressed between cells and tissue types. In an effort to establish whether this strategy is viable for human testis and testicular cancers, we characterised the relative expression of OCT4, SALL4 and DDX4 in specimens with a characterised neoplastic germ cell component. If any of these genes were stably expressed at the cellular level, we would expect for instance an EC specimen with 50% neoplastic cells to express twice as much gene-specific mRNA as a specimen with 25% neoplastic cells. Although we observed such a trend with OCT4 and SALL4 expression within specific neoplasms, for example CIS, SEM or EC, it was not the case when we compared different pathologies. Therefore, this approach should be used with great caution on these tissues and should always be accompanied by additional histological assessments. DDX4 was significantly down-regulated in neoplastic testis compared with the control and was nearly absent in SEM and EC, as also observed previously (Zeeman et al., 2002; Mitchell et al., 2014). Therefore, DDX4 is not a usable second-tier normalizing gene in TGCTs.

In view of our analysis of identical biological samples normalized with different reference genes, it is apparent that final normalized results and even conclusions can vary significantly and must always be considered both during study design and when comparing results between studies. Also, although a careful use of OCT4 or SALL4 with a delta-delta Ct measurement could be considered, our study indicates that there are significant changes to transcript abundance within single cells. Thus, a cell-specific normalizing gene for human germ cells or their neoplastic counterparts remains to be identified and would likely require the use of different strategies, such as RNA sequencing, on more carefully defined pathological tissue pieces.

In conclusion, our study has confirmed previous observations that the identification of stable normalizing genes in human testis and TGCTs is particularly challenging (Neuvians et al., 2005; Cavalcanti et al., 2011). We did however, identify RPS20, RPS29 and SRSF4 to be more stable
than other more popular normalizing genes, and would recommend their use in future studies. When practically feasible, we would also advocate the use of geometric mean of multiple (stable) reference genes when normalizing qRT–PCR data by the delta-Ct method. What is perhaps more important is the recognition that no single normalizing gene has to date been identified as particularly stable in testicular pathologies, which should be taken into consideration when interpreting data.

**Supplementary data**

Supplementary data are available at http://molehr.oxfordjournals.org/.

**Acknowledgements**

We thank Brian V. Hansen, Bonnie S. Häkansson, John E. Nielsen and Ana R. Nielsen for their valuable assistance with tissue handling, processing and histological assessments. We thank Lillian Sande for helpful comments on the manuscript.

**Authors’ roles**

T.S. conceived the study, performed experiments, analysed data and wrote the manuscript. A.J. participated in study design, data interpretation and in drafting the manuscript. E.R.-D.M. provided samples, participated in study design and in drafting of the manuscript. All authors approved the final version of the manuscript.

**Funding**

This work was supported by a Global Excellence Grant from the Capital Region of Denmark (to Department of Growth & Reproduction, Rigshospitalet), the Danish Cancer Society [grant nr. R40-A2127-11-S2 to ERDM] and the Childhood Cancer Foundation [grant nr. 2013-16 to ERDM].

**Conflict of interest**

The authors declare no conflicts of interest related to this work.

**References**


