Human Reproduction, Vol.32, No.11 pp. 2170–2177, 2017
Advanced Access publication on September 29, 2017 doi:10.1093/humrep/dex288

human reproduction

ORIGINAL ARTICLE *Andrology*

Genomic fragmentation and extrachromosomal telomeric repeats impact assessment of telomere length in human spermatozoa: quantitative experiments and systematic review

P. Kurjanowicz[1,2,*], S. Moskovtsev[2,3], and C. Librach[1,2,3]

[1]Department of Physiology, University of Toronto, 1 King's College Circle, Toronto, Ontario, Canada M5S 1A8 [2]CReATe Fertility Centre, 790 Bay St., Suite 1100, Toronto, Ontario, Canada M5G 1N8 [3]Department of Obstetrics and Gynaecology, University of Toronto, 123 Edward St. Suite 1200, Toronto, Ontario, Canada M5G 1E2

*Correspondence address. CReATe Fertility Centre, 790 Bay St., Suite 1100, Toronto, Ontario, Canada M5G 1N8.
E-mail: pamm.chan@mail.utoronto.ca

Submitted on August 23, 2016; resubmitted on July 19, 2017

STUDY QUESTION: Can differences in DNA isolation alter assessment of sperm telomere length (spTL) and do they account for conflicting results in the literature on spTL and male fertility?

SUMMARY ANSWER: DNA isolation methods preferentially include or exclude short, extrachromosomal (EC) telomere-specific sequences that alter spTL measurements, and are responsible for a proportion of the disparity observed between investigations.

WHAT IS KNOWN ALREADY: The relationship between spTL and male fertility has become an active area of research. The results across investigations, however, have been discordant, generating a need to critically evaluate the existing body of knowledge to guide future investigations.

STUDY DESIGN, SIZE, DURATION: Quantitative experiments determined the effect of DNA isolation on the integrity of sperm DNA and measures of spTL, while a systematic analysis of the current literature evaluated the effect of DNA isolation and study design on experimental outcomes.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Two DNA isolation methods were compared: Genomic Tips which isolate 'High Molecular Weight' (HMW) DNA exclusively, and QIAamp® DNA Mini which isolates 'Total' genomic DNA irrespective of size. DNA quality was assessed via field inversion gel electrophoresis (FIGE) and spTL was measured via terminal restriction fragment analysis. In addition, major databases in medicine, health and the life sciences were subject to a targeted search, and results were independently screened according to defined exclusion/inclusion criterion. Findings from primary articles were analyzed for concordance and study designs were compared across six moderator variables (sample size, participant age, fertility status, semen fraction, telomere population and type of analysis).

MAIN RESULTS AND THE ROLE OF CHANCE: HMW DNA spTL was significantly longer than spTL measured from total DNA ($P < 0.01$), indicating that Total DNA contained short, EC telomeric repeats that shifted downstream assessment towards shorter spTL. HMW DNA spTL reflected the length of intact, chromosomal telomeres. Major findings on spTL showed the greatest concordance amongst studies that implemented HMW DNA isolation prior to spTL assessment. Studies that utilized Total DNA varied in concordance, but outcomes were similar if (i) a comparative analysis was applied or (ii) a sample size threshold of 81 was achieved for correlative analysis.

LIMITATIONS, REASONS FOR CAUTION: Chromosomal and EC telomeric DNA were distinguished based on outcomes of HMW DNA isolation and size. Further experiments are required to determine the nature and function of these two types of telomeric sequences.

WIDER IMPLICATIONS OF THE FINDINGS: This study reveals a dramatic impact of upstream DNA processing and study design on measurements of spTL, which accounts for conflicting results in the literature. Future assessments of spTL should incorporate independent detection of chromosomal and EC telomeric DNA and specific experimental planning.

STUDY FUNDING/COMPETING INTERESTS: This study was funded by CReATe Fertility Centre, Toronto, Ontario, Canada. The authors have declared no conflict of interest.

REGISTRATION NUMBER: N/A.

Key words: Telomere / Telomere Shortening / Telomere Homeostasis / Spermatozoa / Sperm Head / DNA Fragmentation / DNA Damage

Introduction

Infertility affects ~1 in 8 couples, with male factors accounting for up to half of cases (Rowe and World Health Organization., 2000; Centers for Disease Control and Prevention [CDC], 2013). Diagnostic tests of male fertility in current practice (i.e. semen volume, sperm count, motility, morphology and in some centers, DNA fragmentation) bear limited predictive value for ART outcomes, as the etiology behind abnormal sperm parameters remains unknown in up to 45% of cases (Jungwirth *et al.*, 2012). Despite significant improvements in IVF to bypass semen disorders, the live birth rate associated with identifiable male disorders remains <39% (Rhouma *et al.*, 2003). Further developments in our understanding of sperm physiology and clinical testing of the male are crucial to improvement of ART outcomes. Interest in sperm telomere length (spTL), its potential relationship with semen disorders, and its ability to predict reproductive outcomes has grown rapidly in recent years.

Telomeres are repetitive, non-coding regions of DNA ($TTAGGG_n$) located at the ends of linear chromosomes that set a replicative limit on cell division, and play important roles in chromosome protection, stabilization and tethering (Hayflick *et al.*, 1979; Levy *et al.*, 1992; Griffith *et al.*, 1999; Morimoto *et al.*, 2012). Telomeres in human sperm exhibit unique properties compared to somatic cells: some studies have suggested that telomeres in spermatozoa are significantly longer (Kozik *et al.*, 1998; Kimura *et al.*, 2008; Ferlin *et al.*, 2013), while cross-sectional studies have suggested that spTL increases with age (Allsopp *et al.*, 1992; Aston *et al.*, 2012; Baird *et al.*, 2006). This is in stark contrast to the age-associated telomere erosion observed in somatic cells, which has been linked to decreased longevity and increased disease-susceptibility (Blackburn, 1991; Greider and Blackburn, 1996). SpTL may have an intergenerational effect, since older fathers, presumably having longer spTLs at the time of fertilization, sire offspring with longer leukocyte TLs (Aviv, 2012). Paternal inheritance of long telomeres is a proposed mechanism of positive selection driven by resilient males (Eisenberg, 2011; Aviv, 2012; Eisenberg *et al.*, 2012). Conversely, severely truncated spTL may account for some cases of idiopathic infertility (Baird *et al.*, 2006).

To reconcile discordant results on the relationship between spTL and male fertility, we compared all studies and replicated common experimental methods, and found marked differences in study design and upstream DNA processing prior to spTL assessment, which contributes to differences in outcomes.

Materials and Methods

Participant recruitment and sample processing

This study received institutional research ethics board approval from the University of Toronto, Canada (#29211). Men undergoing semen analysis or IVF treatment at CReATe Fertility Centre (Toronto, Ontario, Canada) provided informed consent to donate excess blood and semen to research following clinical procedures (n = 161). Semen samples were analyzed according to World Health Organization (WHO) guidelines (Cooper *et al.*, 2010). DNA fragmentation was measured via sperm chromatin structure assay (SCSA) (Evenson *et al.*, 1978). Semen samples acquired post diagnostic testing (PDT) were subject to one-layer density gradient centrifugation using PureSperm40™ to isolate all sperm. Samples acquired post IVF (PIVF) underwent two-layer density gradient centrifugation using PureSperm40™/PureSperm80™ to isolate only motile sperm. Whole blood was centrifuged to isolate buffy coat leukocytes. Clinical parameters were collected from patient records: age, days of sexual abstinence, seminal pH and volume, and sperm concentration, motility, DNA fragmentation and morphology.

DNA isolation

Two methods of DNA isolation were compared: Genomic Tips® (G-Tips) (QIAGEN, Cat. #10223) and QIAamp DNA Mini Kit® (QIAamp) (QIAGEN, Cat. #51104) which respectively isolate 'high molecular weight' (HMW) and 'total' genomic DNA. Each method was used according to the manufacturer's recommendations, with incorporation of dithiothreitol (DTT, 80 mM) and proteinase K (250ug/ml) treatments during cell lysis (QIAGEN – User Developed Protocol 2 for Sperm). DNA was eluted in TE buffer (10 mM Tris-Cl, pH 7.5; 1 mM EDTA) and stored at −20°C. QIAamp or G-Tip DNA isolation was applied to PDT samples. PIVF samples were divided in half and subjected to each of the two methods. Leukocyte DNA was isolated using the QIAamp DNA mini-kit according to manufacturer's recommendations, followed by elution and storage conditions similar to that for sperm.

Assessment of genomic DNA integrity

Integrity of undigested genomic DNA was assessed via field inversion gel electrophoresis (FIGE). A 0.4% agarose gel (15 cm × 10 cm) was run at 1.5 V/cm for 6 h to separate large DNA fragments, and one of two DNA ladders: GeneRuler 1 kb DNA Ladder (range 0.25–10 kb, ThermoFisher Scientific #SM0312) or GeneRuler High Range DNA Ladder (range 10–48 kb, ThermoFisher Scientific #SM1351). The High Range DNA Ladder confirmed that HMW undigested sperm DNA was >20 kb in size, the longest spTL expected. DNA from the TeloTAGGG Telomere Length Assay (Roche, Cat. #12209136001) was included as a somatic cell control in addition to leukocyte DNA. All DNA was stained with SYBR Safe™ DNA Gel Stain (ThermoFisher Scientific #S33102) and gels were imaged via UV excitation on a MiniBIS Pro™ (DNR Bio-Imaging Systems).

Telomere length assessment

To assess mean spTL, sperm DNA isolated with each method (total, n = 8; HMW, n = 9) was subjected to terminal restriction fragment (TRF) analysis, the gold standard for absolute TL determination. DNA was digested using MnlI/RsaI restriction enzymes according to Baird *et al.* (2006) to exclude sub-telomeric regions that can vary in length, followed by Southern blotting with reagents from the TeloTAGGG Telomere Length Assay (Roche, Cat. 12209136001) and gel modifications specific for sperm

(Kimura et al., 2010). Densitometric quantification was performed using AlphaEaseFC software to determine mean spTL ($spTL_n$).

Statistical analyses

Clinical patient parameters were compared using the Student's *T*-test (mean ± SEM, two-tailed, 95% confidence interval). Median spTL and interquartile ranges (IQR) for Total and HMW DNA samples ($spTL_{QIAamp}$ vs. $spTL_{G\text{-}Tip}$) were subjected to Whitney-Mann U tests with p-value of <0.01 for significance.

Systematic review of studies on spTL and male fertility

To identify studies on spTL and male fertility, a comprehensive systematic review was conducted. Pre-selected articles were used to identify keywords and Medical Subject Headings (MeSH) to search the US National Library of Medicine MeSH database. Keywords ('Sperm*' and 'Telomere or Telomeres') and MeSH headings ('Spermatozoa/ or Sperm Head/' and 'Telomere/ or Telomere Shortening/ or Telomere Homeostasis') were used to search four major online databases in medicine, health and the life sciences: MEDLINE, EMBASE, Scopus and Web of Science, to identify all relevant articles published up to 7 July 2017. Search results were refined to include studies with both animals and humans, and exclude studies limited to animals only. The search and selection strategy is shown in Fig. 1. Records generated by the search were screened by two independent reviewers according to the defined inclusion/exclusion criterion outlined in Table I.

Categorical analysis of study design and experimental methods: a narrative synthesis

A narrative analysis was conducted according to the 'Cochrane Consumers and Communication Review Group: Data synthesis and analysis' guidelines (Ryan, 2013). Narrative analysis was deemed appropriate as the data met three criteria: (i) inclusion of different study designs that were not well-suited to aggregation, (ii) capture of a wide variety of interventions (in this case, experimental methodologies and patient populations) and (iii) no quantitative data that could be directly pooled across studies. Data were extracted from each study, with particular focus on key 'moderator variables' that differed across studies. Moderator variables are variables that can be expected to modulate the main effects between groups and explain differences in the direction and size of an effect across studies (Popay et al., 2006). In the preliminary synthesis, findings across studies were compiled and studies were grouped according to whether the results were 'Supportive' or 'Discordant'. Concordance was calculated as:

$$\text{Concordance} = n_{\text{studiesinagreement}} / n_{\text{totalstudies}}.$$

Relationships in the data were then explored in the form of a 'Categorical analysis' that involved methodological triangulation of (i) each major finding, (ii) moderator variables and (iii) subcategories within each moderator variable (Supplementary Fig. S1). The maximum age gap (MAG) between men in each study was calculated as the difference in range (Maximum Age – Minimum Age) or approximated from the standard deviation using the 'Range Rule of Thumb' (Range ≈ Standard Deviation × 4), and grouped along 10-year intervals.

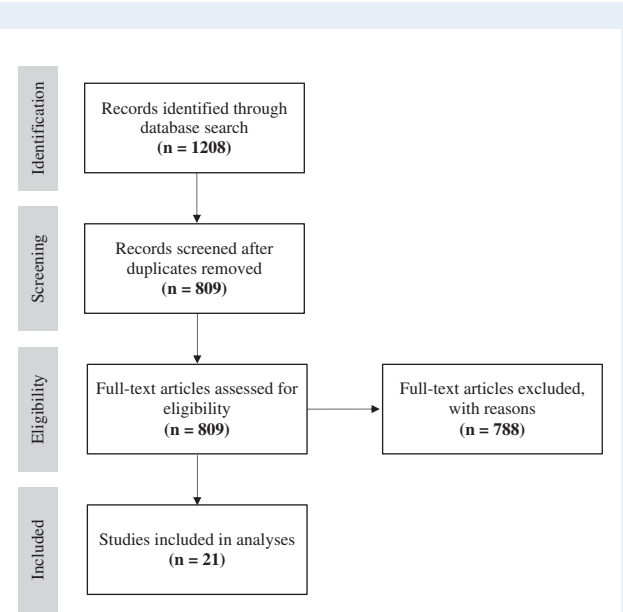

Figure 1 Flow diagram of search strategy used in systematic review of the literature.

Results

Inclusion of fragmented sperm DNA is method-specific and contributes short telomeric sequences to downstream assessment of spTL

PDT samples subjected to Total DNA isolation ($n = 57$) had semen parameters within the normal range (Table II), and genomic DNA was successfully isolated from all samples (57/57; 100% efficiency). FIGE revealed that undigested DNA ($n = 8$) was heterogeneous in fragment size (<1.5 – >10 kb) (Fig. 2, A; 1–8). SpTL ($n = 8$) ranged from 2.38 to 6.71 kb (median 3.20 kb, IQR: 2.50, 3.95). HMW DNA isolation was performed on another set of PDT samples ($n = 104$), and the robustness of the method for future application on both normal and abnormal semen samples was tested. These samples had a significantly lower mean sperm concentration (38.8 ± 41.15 vs 66.6 ± 35.45, $P < 0.001$) and motility (34.6 ± 20.76 vs 48.5 ± 13.09, $P < 0.001$) compared to Total DNA samples, but were still within the normal range overall (Table II). HMW DNA isolation did not successfully yield DNA for all samples (55/104; 52.9% efficiency). The decline in efficiency was not attributed to the increased number of abnormal samples because no significant difference in any clinical parameters were

Table I Inclusion and exclusion criterion applied in systematic literature review to identify relevant articles on sperm telomere length (spTL), tests of male fertility and reproductive outcomes.

Inclusion	Exclusion
• Primary article • Measures of mature spTL • Direct comparison between mature spTL and male fertility potential (i.e. clinical tests of male health, semen parameters, fertility outcomes and/or offspring health)	• Review article • Measures of telomerase activity or expression • Studies of germline precursor cells • Studies in only non-human species

Table II Clinical parameters of sperm samples obtained Post Diagnostic Testing (PDT) and Post IVF (PIVF) subjected to Total DNA and HMW DNA isolation.

	Normal reference ranges[1,2]	PDT						PIVF
		Total DNA	HMW DNA	*P*	Yes (+)	No (−)	*P*	
n		57	104		55	49		8
Age		36.9 ± 5.73	38 ± 7.26	0.457	38.0 ± 0.8	38.0 ± 1.17	0.992	34.3 ± 4.41
Abstinence (days)		3.6 ± 2.27	3.1 ± 1.26	0.089	3.1 ± 0.2	3.1 ± 0.16	0.798	3.0 ± 1.22
pH	7.2–7.8	8.2 ± 0.21	8.1 ± 0.24	0.496	9.9 ± 0.08	8.1 ± 0.04	0.072	3.5 ± 2.22
Seminal volume (ml)	1.5	3.6 ± 1.09	3.6 ± 1.61	0.94	3.7 ± 0.2	3.5 ± 0.24	0.371	59.1 ± 22.5
Sperm concentration (M/Ml)	≥15	66.6 ± 35.45	38.8 ± 41.15	*0.0008	37.0 ± 5.37	40.5 ± 5.9	0.660	48.2 ± 14.19
Sperm motility (%)	≥32	48.5 ± 13.09	34.6 ± 20.76	*0.0005	31.6 ± 2.53	37.3 ± 3.06	0.160	12.2 ± 4.8
DNA fragmentation (DFI%)	≤30	20.9 ± 13.9	20.3 ± 12.7	0.830	21.5 ± 1.61	19.2 ± 1.89	0.345	43.3 ± 6.83
DNA isolation efficiency		100%	52.9%					

The mean ± SD are shown. Samples that successfully yielded HMW DNA are represented as 'Yes (+)' or 'No (−)'. Overall, all experimental groups had semen parameters within the normal reference range.
[1]Rowe and World Health Organization., 2000.
[2]Zini and Libman, 2006—'fair to good fertility potential'.
*Significance <0.01.

observed between samples that successfully or unsuccessfully yielded HMW DNA (Table II). Undigested HMW genomic DNA (n = 9) was greater than 48 kb and homogenous in size (Fig. 2, C; 1–9). SpTL (n = 9) ranged from 9.27–14.95 kb (median 12.89 kb, IQR: 12.18, 14.14). Median spTL of HMW sperm DNA was significantly greater than that of Total DNA (P < 0.01), confirming that HMW DNA isolation excluded short fragments of telomeric DNA that were retained with Total DNA isolation. This effect was unique to sperm, as Total leukocyte DNA was HMW (>10 kb) and homogeneous size (Fig. 2, A; L). To further confirm that differences in DNA integrity and spTL were method-specific, not patient- or fraction-specific, Total and HMW DNA isolation was repeated on motile sperm obtained post IVF treatment (PIVF). PIVF samples had semen parameters within the normal range (Table II). Again, the size of Total DNA (n = 8) was heterogeneous, while HMW DNA was homogeneous (≥48 kb) (Fig. 3).

Differences between studies on spTL and male fertility are a direct result of DNA isolation method-selection and moderator variable disparity

A comprehensive review of the literature identified 21 relevant full-text articles. The data extraction process revealed six key moderator variables: sample size, fertility status, age gap between study participants, semen fraction analyzed, telomere population assessed (i.e. combination of telomere assay and DNA isolation method), and analysis type (i.e. correlative versus comparative) (Supplementary Table SI).

Preliminary synthesis identified a total of 13 major findings across studies (Supplementary Table SII). Of the major findings, four showed 100% concordance (Finding #1–4), two were mostly concordant but had studies with both positive and negative data (Finding #5–6), two were supported by only one study (Finding #12–13), and five ranged in concordance from 58–90% (Finding #7–11). Generally, concordance amongst the disputed findings (#7–11) was high (71–90%). The most disputed finding was Finding #11 ('SpTL increases with male age', 58%).

The first phase of the categorical analysis is shown in Supplementary Table SIII. This analysis revealed five studies that assessed spTL using HMW sperm DNA: Allsopp *et al.*, 1992; Kozik *et al.*, 1998; Baird *et al.*, 2006; Aston *et al.*, 2012; and Shuyuan *et al.*, 2015. No other moderator variables were shared (Supplementary Table SIV). These studies overlapped for Findings 1, 5 and 11, and were all 'Supportive' and in agreement with one another (Supplementary Table SIII, indicated by asterisks '*').

Discordant findings were contributed by six studies: Thilagavathi *et al.*, 2013 (Frequency = 4), Turner and Hartshorne, 2013 (Frequency = 5), Rocca *et al.*, 2016 (Frequency = 1), Yang *et al.* (2015b) (Frequency = 1), Mishra *et al.*, 2016 (Frequency = 1), and Zhao *et al.*, 2016 (Frequency = 1). In these studies, spTL was assessed on Total DNA, or the study did not provide sufficient information for determination (Supplementary Table SV). No other moderator variables were shared. However, since some 'Supportive' studies also assessed spTL in Total DNA, it was not the sole contributor of discordant results, and further analysis was conducted.

The second phase of the categorical analysis narratively described the comparisons within and between groups (Supplementary Table SVI). Of the findings that were not disputed, Finding #1 was the most thoroughly investigated (n = 6 studies) and robust, in that each study was markedly different in experimental design but had the same conclusion. Findings #2–3 were supported by only a few medium- to large-scale studies; finding #4 was supported by three small studies; and findings #5–6 were only disputed by studies that had both positive and negative data.

For findings that were more controversial (Findings #7–11), supportive studies implemented stricter approaches to data analysis via application of comparative over correlative analyses, or had a sample size ≥81 when conducting correlative analyses (Findings #7–10). This effect is clearly illustrated by Santiso *et al.*, 2010; and Thilagavathi *et al.*, 2013: with sample sizes <81, indirect comparative analysis in the data showed that spTL was significantly different in sperm with high versus low DNA fragmentation, and between fertile and infertile patients who had significantly different sperm count and motility. When correlative analysis was applied, no relationship was observed. The

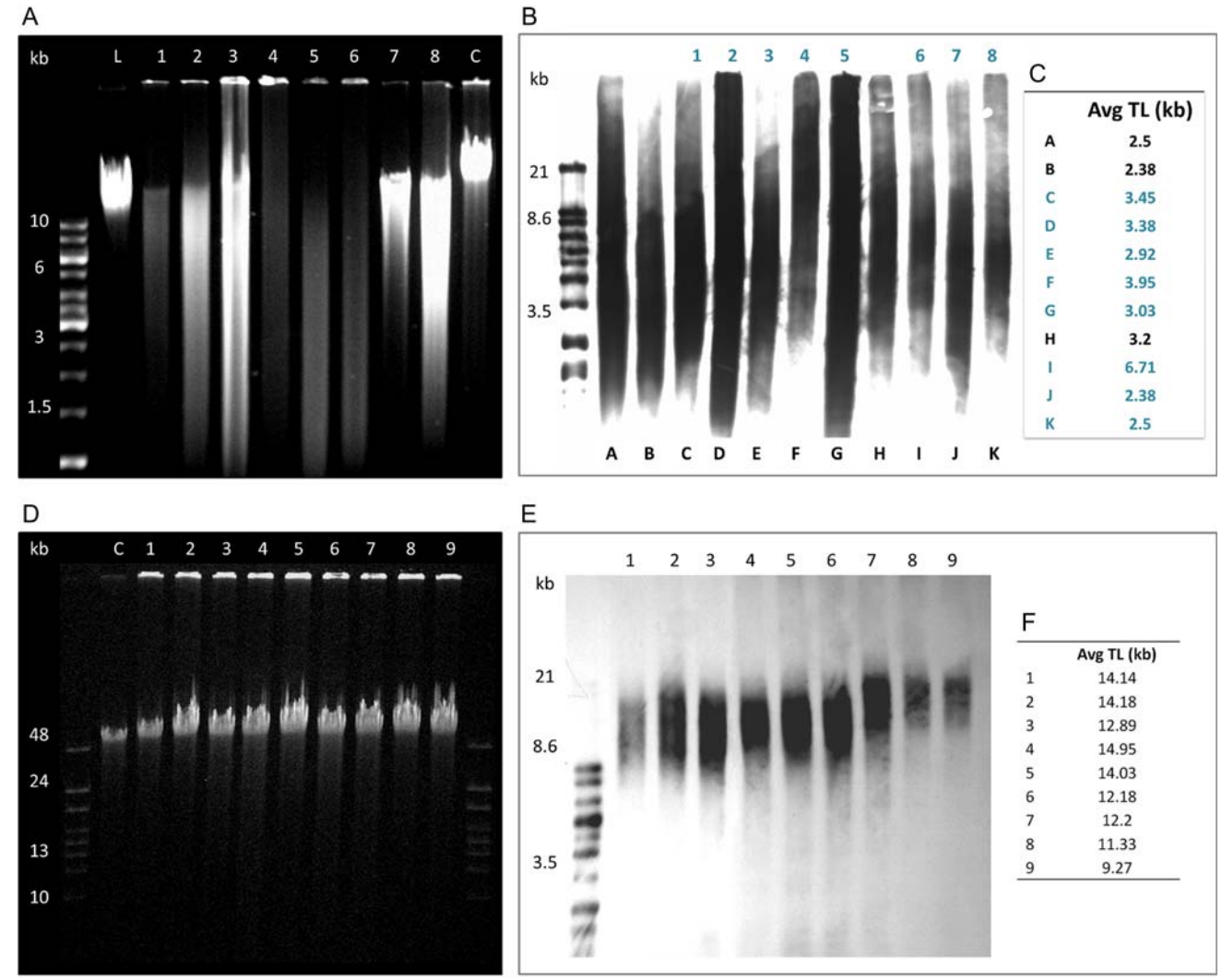

Figure 2 Field inversion gel electrophoresis (FIGE) to assess DNA integrity and terminal restriction fragment (TRF) analysis to measure average telomere length of Total DNA and High Molecular Weight (HMW) DNA. (**A**) FIGE of Total DNA ($n = 8$); L: leukocyte DNA, 1–8: sperm DNA, C: control DNA from TeloTAGGG telomere length assay (Roche). Total sperm DNA was heterogeneous in size. Leukocyte and control DNA was homogeneous and of HMW (>10 kb). (**B**) TRF analysis on Total DNA ($n = 11$). Lanes 1–8 correspond to samples shown in A. (**C**) Average telomere length (TL) of samples from B. Average TL ranged from 2.38–6.71 kb (median 3.20 kb, IQR: 2.50, 3.95). (**D**) FIGE of HMW DNA ($n = 9$); C: control DNA from TeloTAGGG telomere length assay, 1–9: sperm DNA. Sperm DNA was homogeneous in size and of HMW (≥ 48 kb). (**E**) TRF analysis of HMW DNA ($n = 9$). Lanes 1–9 correspond to samples shown in D. (**F**) Average TL of samples from D. Average TL ranged from 9.27 to 14.95 kb (median 12.89 kb, IQR: 12.18, 14.14). Average TL was significantly greater in HMW sperm DNA versus Total DNA ($P < 0.01$).

influence of moderator variables on outcomes of Finding #11 were more complex. Again, all studies that conducted comparative analyses were supportive. For correlative analyses, when pooled populations of healthy volunteers and/or infertility patients with normal semen analyses across a 20–50-year age span were tested, spTL was positively associated with age. When infertility patients (normal and abnormal semen analyses) were pooled, the relationship was only observed in men with a MAG of 15–18 years, and not 24–34 years, demonstrating that the relationship between spTL and age was confounded by inclusion of infertile patients. *P*-values of studies that contained more healthy men showed a stronger significance ($P \leq 0.02$) compared to those from studies with more infertile men ($P = 0.036$–0.05).

Discussion

We were interested in whether there was a difference in assessment of spTL that resulted from upstream DNA isolation, due to the intrinsic presence of DNA fragmentation in the sperm genome. Specifically, we were interested in whole genomic (i.e. 'Total') and HMW DNA isolation, which were both applied by previous investigations. We tested two commercial DNA isolation techniques: one that isolated Total DNA, regardless of size (QIAamp®), and one that isolated only HMW DNA (G-Tip). We then measured the outcomes of spTL assessment. QIAamp was directly implemented in seven previous studies (San Gabriel *et al.*, 2006; Ferlin *et al.*, 2013; Yang *et al.* (2015a);

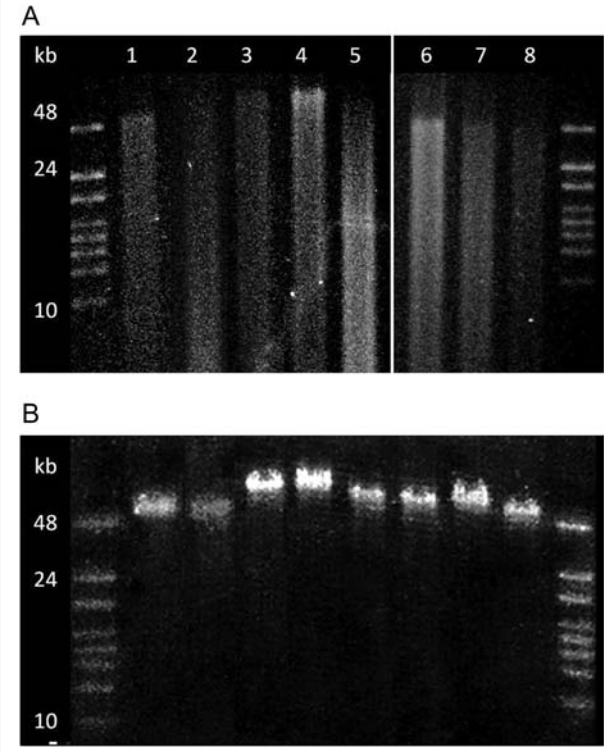

Figure 3 FIGE to assess integrity of eight paired sperm samples (obtained post IVF treatment) that underwent total and HMW DNA isolation. (**A**) Total DNA is heterogeneous in size. (**B**) HMW DNA is homogeneous in size ≥48 kb.

Yang *et al.* (2015b); Cariati *et al.*, 2016; Rocca *et al.*, 2016; Zhao *et al.*, 2016), while four studies assessed spTL in Total DNA using other methods (Moskovtsev *et al.*, 2010; Turner and Hartshorne, 2013; Antunes *et al.*, 2015; Ling *et al.*, 2016). The G-Tip method was directly applied by one study (Aston *et al.*, 2012), but its principle is similar to traditional phenol-chloroform-based techniques that target large DNA molecules, as employed by four studies (Allsopp *et al.*, 1992; Kozik *et al.*, 1998; Baird *et al.*, 2006; Shuyuan *et al.*, 2015).

We found that Total sperm DNA was heterogeneous in fragment size and that downstream assessment of mean spTL was below the expected size range published by other groups using TRF (8–20 kb) (Allsopp *et al.*, 1992; Kozik *et al.*, 1998; Baird *et al.*, 2006; Kimura *et al.*, 2008; Aston *et al.*, 2012) and significantly shorter than spTL assessed on HMW DNA. The differences in DNA integrity and spTL occurred as a direct result of the upstream DNA isolation method, and was not attributed to differences between patients or the sperm fraction assessed. The effect was unique to sperm. From this, we concluded that spTL obtained from Total sperm DNA includes short telomeric repeats that shift mean spTL towards shorter telomeres, and are not associated with HMW DNA. In other words, two distinct DNA and telomere populations are present in the sperm genome: HMW DNA that contains only intact chromosomal (Chr) telomeres, and pooled DNA that contains both Chr telomeres and extrachromosomal (EC) telomeric-repeats. Short, EC telomeric-repeats may occur as a by-product of global DNA fragmentation that includes breakdown of telomeres, or may be the products of telomere trimming (t-circles), which has previously been reported in mature sperm (Pickett *et al.*, 2011). Further research is required to determine the nature and function of these two types of telomeric sequences.

A comprehensive review of the literature returned 21 scientific journal articles that measured spTL and its relation to current clinical tests of male fertility and reproductive outcomes. To compare and contrast studies, a categorical analysis was performed in the form of a 'Narrative Synthesis': an emerging form of research—traditionally applied in Economic and Social research, and more recently adopted by major health research organizations such as Cochrane, to inform international health policy and practice. Narrative syntheses aim to qualitatively assess evidence from multiple studies to identify patterns in cases where quantitative meta-analysis is less suitable (e.g. differences in study design or experimental methodology that generate quantitative data that should not be aggregated, as observed here) (Ryan, 2013). In the future, when more investigations with similarity in moderator variables are available, experimental data can be subject to meta-analysis.

Our analyses revealed a number of major findings that did not encounter opposition in the literature: spTL distributions demonstrate inter-individual variation, spTL is not related to morphology, spTL is greater than somatic cell TL, and longer spTL is associated with positive indicators of male health, lower levels of reactive oxygen species (ROS) production in semen, and positive indicators of offspring health. Although these findings were not disputed, it was important to evaluate the quality of the current evidence. The most robust, undisputed finding was that spTL demonstrates detectable inter-individual variability; this is an important consideration, as such variability would contribute to a difference in experimental outcomes, and may or may not have been accounted for in all experimental designs. For other undisputed findings, the current evidence is based on less than four studies, and outcome measures often included very broad indicators of offspring and male health that did not directly overlap between studies (i.e. embryo quality, ongoing pregnancy, offspring leukocyte telomere length, BMI and environmental toxin metabolites). We contend that although these findings are not currently disputed, more large-scale studies with similar experimental designs and outcome measures would solidify the evidence behind these claims.

Findings that were more controversial included: a positive association between spTL and age, fertility status, and normal results on clinical tests of sperm quality (i.e. DNA fragmentation, motility and concentration). Discordant findings were mainly contributed by six studies (Thilagavathi *et al.*, 2013; Turner and Hartshorne, 2013; Rocca *et al.*, 2016; Yang *et al.* (2015b); Mishra *et al.*, 2016; Zhao *et al.*, 2016). Discordant studies had no moderator variables in common apart from spTL assessment using Total DNA. We did not, however, conclude that this variable alone was responsible for discordant results, as a number of studies that used Total DNA also yielded supportive results.

Further analysis revealed that supportive studies implemented stricter approaches to data analysis by applying comparative analyses instead of correlative analyses. All studies that conducted comparative analyses showed a positive, significant difference, regardless of sample size. Correlative analysis, on the other hand, had a lower detection rate, and required a larger sample size. We found that a sample size of ≥81 was required to detect a statistically significant positive

relationship using correlative analyses, and studies that showed no relationship had sample sizes below this threshold.

The relationship between spTL and male age proved to be the most disputed and complex. Again, all studies that conducted comparative analyses were supportive, but an important interaction between male age and fertility status was also observed: when men within the 'normal' fertility spectrum (pooled populations of healthy volunteers and/or infertility patients with normal semen analyses) across a 20–50-year age span were tested, spTL was positively associated with age and easily detected. When the infertility patient population was introduced (pooled normal and abnormal semen analyses), the relationship was only observed across a narrow age span of 15–18 years, and not at 24–34 years. In other words, the influence of the relationship between infertility and spTL confounded the relationship between spTL and age, rendering it detectable only within age-matched individuals. This effect was captured by the increasing significance of p-values in studies that contained more healthy men compared to those with more infertile men.

Lastly, to reflect our experimental findings on the effect of DNA isolation on spTL assessment, we identified five studies that assessed spTL using HMW DNA (Allsopp *et al.*, 1992; Kozik *et al.*, 1998; Baird *et al.*, 2006; Aston *et al.*, 2012; Shuyuan *et al.*, 2015). When the findings of the studies overlapped, the results were supportive of a relationship and in agreement with one another: spTL distributions demonstrate inter-individual variation, was greater than somatic cell TL, and increased with age. Agreement between these studies demonstrates that spTL assessment on HMW DNA has a very strong influence on experimental outcomes, and generates consistent results. The potential effect of other moderator variables such as sample size, age, fertility status, and semen fraction utilized for experimentation was dramatically reduced.

Overall, our findings highlight that spTL assessment is altered by upstream isolation of HMW or Total DNA, which yield the spTL of Chr only or pooled EC and Chr telomeres, respectively. Future investigations will be targeted towards determining whether EC telomere fragments can be separated from HMW sperm DNA to assess Chr spTL independently from EC telomeric DNA, and to determine whether EC telomeric repeat fragments are t-circles. Both investigations are crucial to understanding telomeres in human spermatozoa and their potential role in male fertility. Our results also demonstrate that the outcomes of spTL investigations are heavily dependent on specific moderator variables, particularly the telomere population assessed (Chr vs. Chr/EC), and type of analysis applied (comparative vs. correlative). These findings are important to consider in future investigations of spTL.

Supplementary data

Supplementary data are available at *Human Reproduction* online.

Acknowledgements

The authors would like to thank: Dr L. Shlush and Dr M. Kimura for their technical advice in the experimental component; Dr B. Cox and E. Lenton for guidance on categorical analyses and systematic review; and P. Gelis for abstract screening. We thank Dr T. Brown, Dr M. Allcott, Dr A. Gauthier-Fisher, T. Barretto, K. Quach, P. Vyas and our reviewers for critical comments on the manuscript.

Authors' roles

P.K. participated in the design, execution, and analyses of the study, and drafted the manuscript. S.M. and C.L. supervised the project, participated in critical discussions and revised the manuscript.

Funding

This study was funded by the CReATe Fertility Centre (Toronto, Canada).

Conflict of interest

None declared.

References

Allsopp RC, Vaziri H, Patterson C, Goldstein S, Younglai EV, Futcher AB, Greider CW, Harley CB. Telomere length predicts replicative capacity of human fibroblasts. *Proc Natl Acad Sci USA* 1992;**89**:10114–10118.

Antunes DMF, Kalmbach KH, Wang F, Dracxler RC, Seth-Smith ML, Kramer Y, Buldo-Licciardi J, Kohlrausch FB, Keefe DL. A single-cell assay for telomere DNA content shows increasing telomere length heterogeneity, as well as increasing mean telomere length in human spermatozoa with advancing age. *J Assist Reprod Genet* 2015;**32**:1685–1690.

Aston KI, Hunt SC, Susser E, Kimura M, Factor-Litvak P, Carrell D, Aviv A. Divergence of sperm and leukocyte age-dependent telomere dynamics: implications for male-driven evolution of telomere length in humans. *Mol Hum Reprod* 2012;**18**:517–522.

Aviv A. Genetics of leukocyte telomere length and its role in atherosclerosis. *Mutat Res* 2012;**730**:68–74.

Baird DM, Britt-Compton B, Rowson J, Amso NN, Gregory L, Kipling D. Telomere instability in the male germline. *Hum Mol Genet* 2006;**15**:45–51.

Blackburn EH. Structure and function of telomeres. *Nature* 1991;**350**:569–573.

Cariati F, Jaroudi S, Alfarawati S, Raberi A, Alviggi C, Pivonello R, Wells D. Investigation of sperm telomere length as a potential marker of paternal genome integrity and semen quality. *Reprod Biomed Online* 2016;**33**:404–411.

Centers for Disease Control and Prevention. (2013). Assisted reproductive technology: fertility clinici success rates report. Retrieved from http://www.cdc.gov/art/pdf/2013-report/art-2013-fertility-clinic-report.pdf

Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT *et al*. World Health Organization reference values for human semen characteristics. *Hum Reprod Update* 2010;**16**:231–245. doi:210.1093/humupd/dmp1048. Epub 2009 Nov 1024.

Eisenberg DTA. An evolutionary review of human telomere biology: the thrifty telomere hypothesis and notes on potential adaptive paternal effects. *Am J Human Biol* 2011;**23**:149–167.

Eisenberg DTA, Hayes MG, Kuzawa CW. Delayed paternal age of reproduction in humans is associated with longer telomeres across two generations of descendants. *Proc Natl Acad Sci U S A* 2012;**109**:10251–10256.

Evenson DP, Witkin S, deHarven E, Bendich A. Ultrastructure of partially decondensed human spermatozoal chromatin. *J Ultrastruc Res* 1978;**63**:178–187.

Ferlin A, Rampazzo E, Rocca MS, Keppel S, Frigo AC, De Rossi A, Foresta C. In young men sperm telomere length is related to sperm number and parental age. *Human Reproduction* 2013;**28**:3370–3376.

Greider CW, Blackburn EH. Telomeres, telomerase and cancer. *Sci Am* 1996;**274**:92–97.

Griffith JD, Comeau L, Rosenfield S, Stansel RM, Bianchi A, Moss H, de Lange T. Mammalian telomeres end in a large duplex loop. *Cell* 1999;**97**:503–514.

Hayflick L. The cell biology of aging. *J Invest Dermatol* 1979;**73**:8–14.

Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, Krausz C. European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol* 2012;**62**:324–332. doi:310.1016/j.eururo.2012.1004.1048. Epub 2012 May 1013.

Kimura M, Cherkas LF, Kato BS, Demissie S, Hjelmborg JB, Brimacombe M, Cupples A, Hunkin JL, Gardner JP, Lu X *et al*. Offspring's leukocyte telomere length, paternal age, and telomere elongation in sperm. *PLoS Genet* 2008;**4**:e37.

Kimura M, Stone RC, Hunt SC, Skurnick J, Lu X, Cao X, Harley CB, Aviv A. Measurement of telomere length by the Southern blot analysis of terminal restriction fragment lengths. *Nat Protoc* 2010;**5**:1596–1607. doi:1510.1038/nprot.2010.1124. Epub 2010 Sep 1592.

Kozik A, Bradbury EM, Zalensky A. Increased telomere size in sperm cells of mammals with long terminal (TTAGGG)(n) arrays. *Mol Reprod Dev* 1998;**51**:98–104.

Levy MZ, Allsopp RC, Futcher AB, Greider CW, Harley CB. Telomere end-replication problem and cell aging. *J Mol Biol* 1992;**225**:951–960.

Ling X, Zhang G, Chen Q, Yang H, Sun L, Zhou N, Wang Z, Zou P, Wang X, Cui Z *et al*. Shorter sperm telomere length in association with exposure to polycystic aromatic hydrocarbons: Results from the MARHCS cohort study in Chongqing, China and *in vivo* animal experiments. *Environ Intl* 2016;**95**:79–85.

Mishra S, Kumar R, Malhotra N, Singh N, Dada R. Mild oxidative stress is beneficial for sperm telomere length maintenance. *World J Methodol* 2016;**6**:163–170.

Morimoto A, Shibuya H, Zhu X, Kim J, Ishiguro KI, Han M, Watanabe Y. A conserved KASH domain protein associates with telomeres, SUN1, and dynactin during mammalian meiosis. *J Cell Biol* 2012;**198**:165–172.

Moskovtsev SI, Willis J, White J, Mullen JB. Disruption of telomere-telomere interactions associated with DNA damage in human spermatozoa. *Syst Biol Reprod Med* 2010;**56**:407–412.

Pickett HA, Henson JD, Au AY, Neumann AA, Reddel RR. Normal mammalian cells negatively regulate telomere length by telomere trimming. *Hum Mol Genet* 2011;**20**:4684–4692.

Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, Britten N, Roen K, Duffy S. Guidance on the conduct of narrative synthesis in systematic reviews. *ESRC Method Program* January 2006. DOI:10.13140/2.1.1018.4643.

Rhouma KB, Miled EB, Attallah K, Marrakchi H, Khouja H, Sakly M. Successful pregnancies after using immotile spermatozoa from ejaculate, epididymis and testis. *Eur J Obstet Gynecol Reprod Biol* 2003;**108**:182–185.

Rocca MS, Speltra E, Menegazzo M, Garolla A, Foresta C, Ferlin A. Sperm telomere length as a parameter of sperm quality in normozoospermic men. *Hum Reprod* 2016;**3.1**:1158–1163.

Rowe PJ, World Health Organization. *WHO Manual for the Standardized Investigation, Diagnosis and Management of the Infertile Male*. Cambridge: Cambridge University Press [for] World Health Organization, 2000.

Ryan R, Cochrane Consumers and Communication Review Group: Data Synthesis and Analysis. http://cccrg.cochrane.org, June 2013 (21 December 2016, date last accessed).

San Gabriel M, Zhang X, Zini A. Estimation of human sperm gene-specific deoxyribonucleic acid damage by real-time polymerase chain reaction analysis. *Fertil Steril* 2006;**85**:797–799.

Santiso R, Tamayo M, Gosalvez J, Meseguer M, Garrido N, Luis Fernandez J. Swim-up procedure selects spermatozoa with longer telomere length. *Mutat Res* 2010;**688**:88–90.

Shuyuan L, Changjun Z, Haiying P, Xiaoqin H, Hao S, Keqin L, Kai H, Jiayou C, Zhaoqing Y. Association study of telomere length with idiopathic male infertility. *Yi Chuan* 2015;**37**:1137–1142.

Thilagavathi J, Kumar M, Misha SS, Venkatesh S, Kumar R, Dada R. Analysis of sperm telomere length in men with idiopathic infertility. *Arch Gynecol Obstet* 2013;**287**:803–807.

Turner S, Hartshorne GM. Telomere lengths in human pronuclei, oocytes and spermatozoa. *Mol Hum Reprod* 2013;**19**:510–518.

Yang Q, Zhao F, Dai S, Zhang N, Zhao W, Bai R, Sun Y. Sperm telomere length is positively associated with the quality of early embryonic development. *Human Reprod* 2015a;**30**:1876–1881.

Yang Q, Zhang N, Zhao F, Zhao W, Dai S, Liu J, Bukhari I, Xin H, Niu W, Sun Y. Processing of semen by density gradient centrifugation selects spermatozoa with longer telomeres for assisted reproduction techniques. *Reprod Biomed Online* 2015b;**31**:44–50.

Yang Q, Zhao F, Hu L, Bai R, Zhang N, Yao G, Sun Y. Effect of paternal overweight or obesity on IVF treatment outcomes and the possible mechanisms involved. *Sci Rep* 2016;**6**:29787.

Zhao F, Yang Q, Shi S, Luo X, Sun Y. Semen preparation methods and sperm telomere length: density gradient centrifugation versus the swim up procedure. *Sci Rep* 2016;**6**:39051.

Zini A, Libman J. Sperm DNA damage: importance in the era of assisted reproduction. *Curr Opin Urol* 2006;**16**:428–434.