

University of Groningen

## Increasing the accuracy and precision of relative telomere length estimates by RT qPCR

Eastwood, Justin R.; Mulder, Ellis; Verhulst, Simon; Peters, Anne

*Published in:*  
Molecular Ecology Resources

*DOI:*  
[10.1111/1755-0998.12711](https://doi.org/10.1111/1755-0998.12711)

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Eastwood, J. R., Mulder, E., Verhulst, S., & Peters, A. (2018). Increasing the accuracy and precision of relative telomere length estimates by RT qPCR. *Molecular Ecology Resources*, 18(1), 68-78.  
<https://doi.org/10.1111/1755-0998.12711>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).


The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Increasing the accuracy and precision of relative telomere length estimates by RT qPCR

Justin R. Eastwood<sup>1</sup>  | Ellis Mulder<sup>2</sup> | Simon Verhulst<sup>2</sup> | Anne Peters<sup>1</sup>

<sup>1</sup>School of Biological Sciences, Monash University, Clayton, Vic, Australia

<sup>2</sup>Groningen Institute for Evolutionary Life Sciences, University of Groningen, Groningen, The Netherlands

## Correspondence

Justin R. Eastwood, School of Biological Sciences, Monash University, Clayton, Vic, Australia.

Email: Justin.eastwood@monash.edu

## Funding information

Australian Research Council (ARC) Future Fellowship, Grant/Award Number: FT 110100505; Australian Research Council (ARC) Discovery Project, Grant/Award Number: DP150103595

## Abstract

As attrition of telomeres, DNA caps that protect chromosome integrity, is accelerated by various forms of stress, telomere length (TL) has been proposed as an indicator of lifetime accumulated stress. In ecological studies, it has been used to provide insights into ageing, life history trade-offs, the costs of reproduction and disease. qPCR is a high-throughput and cost-effective tool to measure relative TL (rTL) that can be applied to newly collected and archived ecological samples. However, qPCR is susceptible to error both from the method itself and pre-analytical steps. Here, repeatability was assessed overall and separately across multiple levels (intra-assay, inter-assay and inter-extraction) to elucidate the causes of measurement error, as a step towards improving precision. We also tested how accuracy, defined as the correlation between the “gold standard” for TL estimation (telomere restriction fragment length analysis with in-gel hybridization), could be improved. We find qPCR repeatability (intra- and inter-assay levels) to be at similar levels across three common storage media (ethanol, Longmire’s and Queen’s). However, inter-extraction repeatability was 50% lower for samples stored in Queen’s lysis buffer, indicating storage medium can influence precision. Precision as well as accuracy could be increased by estimating rTL from multiple qPCR reactions and from multiple extractions. Repetition increased statistical power equivalent to a 25% (single extraction analysed twice) and 17% (two extractions) increase in sample size. Overall, this study identifies novel sources of variability in high-throughput telomere quantification and provides guidance on sampling strategy design and how to increase rTL precision and accuracy.

## KEYWORDS

molecular aging, quantitative PCR, repeatability, statistical power, telomeres

## 1 | INTRODUCTION

Telomeres are of broad interest because they are associated with fundamental life history processes including ageing, longevity, reproduction and disease. Located at the end of eukaryotic chromosomes (Gomes et al., 2011), telomeres are composed of G-rich tandem DNA repeats (5'-TTAGGG-3')<sub>n</sub> and associated proteins which function to maintain chromosomal integrity and prevent DNA damage

(Blackburn, 1991; McClintock, 1941). Telomeres shorten with each cellular division due to the inability of DNA polymerases to completely replicate the terminal end of linear chromosomes (Levy, Allsopp, Futcher, Greider, & Harley, 1992), thus telomeres act as a buffer preventing the loss of coding DNA. In addition, telomeres are involved in mitosis and meiosis (alignment and segregation of chromosomes), DNA repair, gene expression and stress resistance (Aubert & Lansdorp, 2008; von Zglinicki, Bürkle, & Kirkwood, 2001).

Telomeres shorten until they reach a critical length, the outcome is either cell apoptosis or replicative senescence. Possibly because of this direct link with cellular senescence, average blood cell telomere length can act as indicator of remaining lifespan (e.g., Boonekamp, Simons, Hemerik, & Verhulst, 2013; Heidinger et al., 2012) and accumulated life stress (Monaghan & Haussmann, 2006). In ecological research, telomeres are used to shed light on a range of processes including chronological and biological ageing, life history trade-offs, reproductive success, environmental stress and disease (Asghar et al., 2015; Haussmann & Heidinger, 2015; Haussmann & Marchetto, 2010; Monaghan, 2010), and new applications undoubtedly will continue to arise. However, there are many concerns surrounding the reliability of telomere quantification (Aviv, 2008; Horn, Robertson, & Gemmill, 2010; Nakagawa, Gemmill, & Burke, 2004; Nussey et al., 2014; Verhulst et al., 2015).

The methodological techniques currently available to measure TL vary widely and have differing advantages and disadvantages (Nussey et al., 2014). The two primary methods that have been adopted in ecology are the Telomere Restriction Fragment length method (TRF) and a real-time kinetic quantitative PCR (qPCR) method (Cawthon, 2002; Harley, Futcher, & Greider, 1990). TRF measures the population of telomeres for each individual (genome wide telomere length) using restriction enzymes and labelling with radioactive oligonucleotide (Harley et al., 1990; Haussmann & Vleck, 2002). TRF is considered the "gold standard" for measuring telomere length (Nussey et al., 2014), and it can be applied across species. However, there are some aspects of the method that can be unappealing to ecologists. For example, TRF analysis requires more DNA, more specific collection and storage requirements, more specialized equipment and expertise, is lower throughput, and more time-consuming than alternative methods such as qPCR (Martin-Ruiz et al., 2015; Nussey et al., 2014).

In comparison with the other available methods, telomere quantification by qPCR is potentially high through-put and cost-effective, depending on its reliability in practice. Consequently, it has been widely adopted for telomere length estimation. The relative qPCR method uses the ratio between telomere repeat copy number versus a single-copy number control gene (Cawthon, 2002). An absolute qPCR method has also been developed and calculates telomere length using synthesized oligonucleotides as a standard (O'Callaghan, Dhillon, Thomas, & Fenech, 2008), but is not frequently applied. Telomere qPCR is useful in ecological research because the method requires less DNA than other methods and, most importantly, it can be utilized to take advantage of standard archived samples from longitudinal studies. However, telomere qPCR has several drawbacks. It requires identification of a control gene and separate optimization for each species. It only provides a single number (TRF yields a telomere length distribution for each sample) that cannot be compared between laboratories or within laboratories unless the same control samples are used. The qPCR method like most other methods (e.g., standard TRF, but unlike in-gel hybridization TRF) does not differentiate terminal from interstitial telomeric sequences (Foote, Vleck, & Vleck, 2013), and to what

extent this is a problem is species specific. Within-individuals though, this is not an issue if measuring relative change in TL because interstitial telomeres can be assumed to retain their length with age. More broadly qPCR can be prone to measurement error (Aviv et al., 2011), particularly if the method is not optimized to a high standard (Nakagawa et al., 2004; Nussey et al., 2014; Olsen, Bérubé, Robbins, & Palsbøll, 2012).

Random error can be a serious problem when using qPCR (Eisenberg, Kuzawa, & Hayes, 2015; Nussey et al., 2014; Olsen et al., 2012), particularly if the effects of interest are small relative to variation in telomere length, as is often the case because individual variation in telomere length is usually large at birth (Boonekamp, Mulder, Salomons, Dijkstra, & Verhulst, 2014; Factor-Litvak et al., 2016). Therefore, identifying factors that influence precision (repeatability) and accuracy (association with the true value) of the method are important. Telomere qPCR is typically validated by correlating the qPCR telomere estimates with those from a TRF method (Cawthon, 2002; Criscuolo et al., 2009), providing a proxy for the methods accuracy. Telomere qPCR precision is often reported as coefficient of variation or repeatability at the qPCR level, which includes intra-assay (between duplicate quantification cycle (Cq) values) and inter-assay (between qPCR runs) comparisons. Note, however, that it has recently become clear that the coefficient of variation is not an appropriate reliability measure for qPCR or TRF, and calculating repeatability has been suggested as an alternative (Eisenberg, 2016; Verhulst et al., 2015). TRF and qPCR reliability has mostly been estimated through repeated measurements of isolated DNA samples (Aviv et al., 2011; Martin-Ruiz et al., 2015), but reliability may also depend on preanalytical steps such as sample collection (timing and sample type) and storage (freezing, preservatives and duration; Koppelstaetter et al., 2005; Zanet et al., 2013; Dlouha, Maluskova, Kralova Lesna, Lanska, & Hubacek, 2014; Tolios, Teupser, & Holdt, 2015; Schmidt et al., 2016), and DNA extraction and storage (Boardman, Skinner, & Litzelman, 2014; Cunningham et al., 2013; Denham, Marques, & Charchar, 2014; Hofmann et al., 2014; Raschenberger et al., 2016; Seeker et al., 2016; Tolios et al., 2015). How blood sample storage media commonly used in ecological studies impact measures of telomere length is currently unknown but given that qPCR results are susceptible to DNA damage, differences are likely to occur between drastically different storage methods (e.g., between lysis buffers that dissolve the DNA, or those that leave the DNA contained within the cells such as ethanol or snap freezing; Nussey et al., 2014). How these preanalytical methods such as storage media or DNA extraction then influence telomere estimation precision is largely unknown and identifying these effects may therefore be a crucial step towards future gains in precision and accuracy of the qPCR method.

This study determines how preanalytical methods (storage medium and DNA extraction) affect accuracy and precision of telomere length measurements, and how these can be best optimized. To do so we used a qPCR method to measure relative telomere length (rTL) across sample storage media, DNA extractions, and repeated measures in a nested design, allowing the calculation of intra-assay,

inter-assay and inter-extraction repeatability across sample types (Figure 1). These repeatabilities were then used to assess how much precision is gained when using multiple measurements of rTL. The benefit of this approach is that it can identify methodological factors that reduce precision and determine where in the measurement process rTL estimation precision can be improved. In addition, we compared rTL to telomere length measured using a TRF method that is considered a “gold standard” and excludes interstitial telomere sequences from the estimates. Correlating rTL with TRF is useful because it measures how well qPCR matches the best method available and gives an estimate of accuracy, that is the extent random error and interstitial telomeric repeats together confound the association between rTL estimates and terminal telomere length. We then tested the extent to which averaging multiple measures from a single extraction or from two extractions increases precision and hence statistical power. These data allowed us to achieve our three main aims: (i) determine where random error in qPCR can be reduced and precision gained, (ii) determine how qPCR accuracy can be improved and by what degree, and (iii) determine how variation in common preanalytical steps influences rTL estimates.

## 2 | MATERIALS AND METHODS

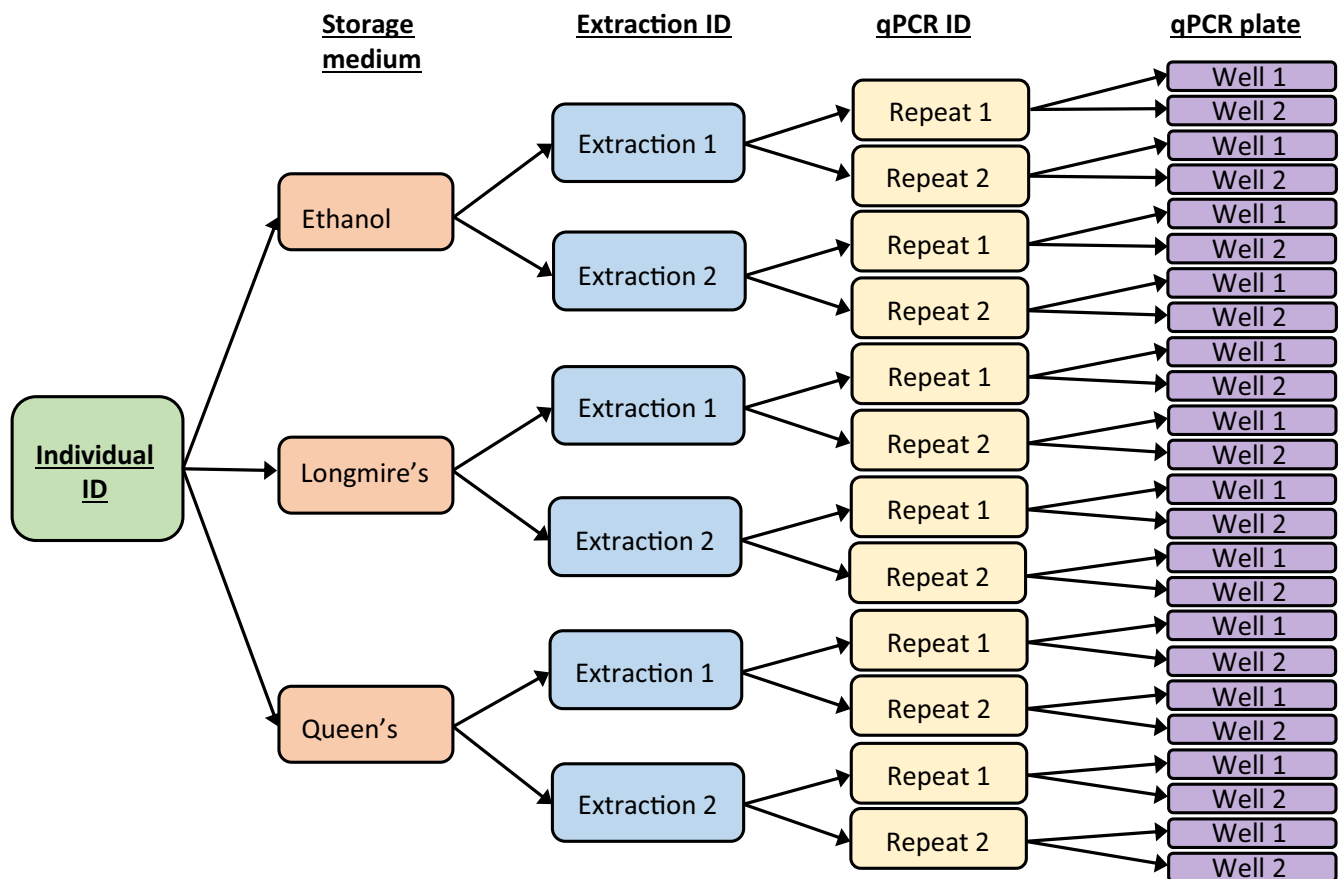
### 2.1 | Sample collection

Blood samples were collected from the purple-crowned fairy-wren (*Malurus coronatus*), at Mornington Wildlife Sanctuary—managed by Australian Wildlife Conservancy (Hidalgo Aranzamendi, Hall, Kingma, Sunnucks, & Peters, 2016; Kingma, Hall, Segelbacher, & Peters, 2009). After puncturing the brachial vein using a 25 gauge needle, approximately 70  $\mu$ l of blood was taken into heparinized capillary tubes and centrifuged at 13,793 g. Approximately 10  $\mu$ l of compacted red blood cells (excluding white blood cells) was then added to a storage medium, for different comparisons. Storage media and extraction methods vary and are presented below.

### 2.2 | Experiments

#### 2.2.1 | DNA extraction comparison

To determine the extent to which DNA extraction methods introduce variability in telomere qPCR measurements, we extracted DNA



**FIGURE 1** Design of the study. Each individual blood sample (Individual ID,  $n = 27$ ) was distributed across three storage (ethanol, Longmire's lysis buffer or Queen's lysis buffer). For each of the three media, DNA was extracted twice (Extraction ID). Each DNA extract was analysed twice according to our standardized qPCR protocol (qPCR ID), in duplicate per plate (qPCR plate). This design allowed repeatability to be calculated at each level: intra-assay, inter-assay and inter-extraction. It further allows an examination of the effect of repeated extractions, qPCR runs and number of wells on the precision of the rTL estimates [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

from  $n = 12$  individuals ( $n = 11$  individuals after quality control; stored in either Queen's lysis buffer, Longmire's lysis buffer or absolute ethanol) using three frequently used methods: QIAamp, Puregene and ammonium acetate precipitation (Eastwood et al., 2015; Longmire, Maltbie, & Baker, 1997; Seutin, White, & Boag, 1991). To optimize the extraction digestion temperature, we extracted blood from another  $n = 17$  samples ( $n = 15$  after quality control; stored as above) at two commonly used digestion temperatures (37°C and 56°C) using the QIAamp method.

The QIAamp DNA kit (Qiagen) was used in conjunction with the QIAcube HT instrument with the following modifications. From samples stored in Queen's or Longmire's lysis buffer we added 30  $\mu\text{l}$  to 220  $\mu\text{l}$  of ATL buffer and 30  $\mu\text{l}$  of Proteinase K; from samples stored in ethanol we added a speck of blood (air-dried to remove ethanol). Digestion was overnight at 37°C, based on the results from our comparison between incubation temperatures at both 37°C and 56°C (See Section 3, Fig. S1). The automated segment of this protocol included a second digestion step using a mixture containing AL buffer and absolute ethanol at a 1:1 ratio. The DNA was purified using AW1, AW2 and absolute ethanol, respectively, with a final AE buffer elution of 150  $\mu\text{l}$ .

The Puregene kit (Qiagen) protocol was used with the following modifications. The blood sample (as above) was added into 300  $\mu\text{l}$  of the Cell Lysis Solution (Qiagen) and 30  $\mu\text{l}$  of proteinase K for digestion at 37°C overnight. After cooling, 200  $\mu\text{l}$  of the Protein Precipitation Solution was mixed into the cell lysate, and the proteins were removed by centrifugation. 100% Isopropanol and 70% ethanol washes (gently mixed and centrifuged at 13,793 g) precipitated the DNA which was suspended in 100  $\mu\text{l}$  of the DNA hydration solution.

The third DNA extraction method used was an ammonium acetate precipitation protocol. Blood (as above) was added into 250  $\mu\text{l}$  of Cell lysis buffer (20 mM EDTA, 120 mM NaCl, 50 mM TRIS-HCL, 20% SDS) with 20  $\mu\text{l}$  of Proteinase K and incubated overnight at 37°C. Once the blood had been fully digested, 300  $\mu\text{l}$  of 4M ammonium acetate was added, mixed thoroughly and then centrifuged. Following ethanol precipitation (100% ethanol wash then a 70% ethanol wash), the DNA was stored in a Low Tris-EDTA buffer (10 mM Tris.HCL, 0.1 mM EDTA; pH 7.5–8.0).

Statistical analyses were completed using spss version 23 (IBM) unless stated otherwise. To test whether DNA extraction digestion temperature (37°C or 56°C) influenced rTL, we performed a linear mixed-effects model (LMM; restricted maximum likelihood) with individual identity as a random effect and extraction digestion temperature, blood storage media and an interaction between the two as fixed effects. We then repeated the analysis using average telomere Cq or average control Cq as dependent variables to test whether telomere length was affected by digestion temperature. To investigate whether DNA extraction influenced rTL estimates, we performed a LMM (restricted maximum likelihood) with individual identity as a random effect and DNA extraction method, blood storage media and an interaction between the two as fixed effects. In addition, we present rank correlations between our method of choice (QIAamp) and the other two DNA extraction methods (Fig. S3).

To choose the most appropriate DNA extraction method, we assumed that if error was introduced by pre-analytical methods it is likely to underestimate rTL, and hence, we consider a method with higher rTL estimates the better method. Of the three methods compared, the QIAamp method produced the highest rTL estimates, and therefore considered the best method (See Section 3; Fig. S2). This method was subsequently used to extract DNA from the samples included in our assessment of effects of storage media and repeatability.

## 2.2.2 | Effect of blood storage medium on rTL estimate and precision

We compared the effect of storage medium on rTL estimate and repeatability for three different storage media ( $n = 27$  individuals, totalling  $n = 81$  samples; see Figure 1 for experimental design): 100% absolute ethanol, Queen's lysis buffer and Longmire's lysis buffer. Individual blood samples (for collection method see above) were distributed across each of the three storage media at the time of collection and stored for 7 months at 4°C before analysis. DNA was extracted using the QIAamp method described above. Having a wide spectrum of aged individuals ensures that our estimates are not biased by age group. Here, individuals were aged between 0.3 and 9.5 years, which approximates the age spectrum of the population.

Repeatabilities and their confidence intervals were calculated using the reliability function in spss (Intraclass correlation coefficient; one-way random model). Each repeatability sample size was standardized at  $n = 48$ , based on complete selection of rTL estimates across extractions that passed quality control. All other levels (inter-assay and intra-assay) had  $n = 48$  selected randomly. To estimate the effect of averaging multiple rTL estimates (inter-assay and inter-extraction) on precision we calculated extrapolated repeatabilities using equation 37 in Nakagawa and Schielzeth (2010). To estimate combined repeatability across all levels, we performed a LMM with rTL as dependent variable and qPCR run, DNA extraction and individual as random intercept terms for each storage medium separately. This model yields information on the extent to which these factors explain variation in rTL estimates and allows the calculation of rTL repeatability whilst controlling for qPCR run and DNA extraction.

## 2.2.3 | Comparison of qPCR to pulse field gel electrophoresis TRF (PFGE-TRF)

For PFGE-TRF approximately, 35  $\mu\text{l}$  of whole blood from  $n = 34$  individuals ( $n = 32$  individuals after quality control; aged 0.1–7.9 years) was collected into chilled 2% EDTA buffer and kept at 4°C until snap frozen in Glycerol storage buffer (within 2 weeks) (Bauch, Becker, & Verhulst, 2013; Salomons et al., 2009). In addition, we also collected compacted red blood cells from these individuals in ethanol. To assess the accuracy of the qPCR method in *M. coronatus* telomere length estimates from PFGE-TRF and qPCR were

subsequently compared. We calculated a Pearson correlation coefficient between average PFGE-TRF telomere length estimates and rTL, (i) estimated once from a single extraction, (ii) two rTL estimates averaged from a single extraction and (iii) two rTL estimates averaged from two extractions. Using these correlation coefficients, we conducted power analyses to determine the benefit of measuring rTL multiple times in terms of required sample size to achieve a similar statistical power. The power analyses were conducted in R (R Core Team 2016) as implemented in the R package `pwr` (Champely, 2016) with a two-tailed test of significance level set to 0.05 and statistical power equal to 0.9.

## 2.3 | Telomere length estimation

### 2.3.1 | Telomere measurement by qPCR

The purity and concentration of DNA samples was assessed using a NanoDrop (ND-1000). DNA integrity was assessed by running all samples on a 1% Agarose gel (100V for 30 min). In this study, all DNA samples were of high integrity (i.e., each sample yielded a single band of high molecular weight). DNA was stored in its respective buffer at 4°C until the quality control checks and qPCR were completed for a period ranging between 1 to 4 weeks. Within each comparison, the DNA was treated identically.

rTL was measured using a qPCR protocol based on Criscuolo et al. (2009). Telomere primers were Tel1b 5'—CGG TTT GTT TGG GTT TGG GTT TGG GTT TGG GTT TGG GTT—3' and Tel2b 5'—GGC TTG CCT TAC CCT TAC CCT TAC CCT TAC CCT TAC CCT—3' (Criscuolo et al., 2009). The normalizing control gene used was glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and the primer sequences used were GT2-GAPDH-forward 5'—CCA TCA CAG CCA CAC AGA AG—3' and GT2-GAPDH-reverse 5'—TTT TCC CAC AGC CTT AGC AG—3' (Atema, van Oers, & Verhulst, 2013). Based on the Cq range and melt curve analysis, the GT2-GAPDH primer set was found to be more suitable for use in PCFW when compared to the GAPDH primers used in Criscuolo et al. (2009). GAPDH suitability as an appropriate single-copy control gene has been confirmed in multiple species (e.g., Atema et al., 2013; Barrett et al., 2012; Criscuolo et al., 2009; Sudyka et al., 2014) and was confirmed in *M. coronatus* by the presence of a single meltcurve with a tight peak in the qPCR and by demonstration of a single qPCR product after gel electrophoresis ( $n = 39$  individuals, 1.5% agarose gel, 100V for 30 min).

qPCR reactions were set up on 96-well plates using an EpMotion 5075 automated pipetting instrument (Eppendorf). The total volume for each reaction was 25  $\mu$ l and included: 12.5  $\mu$ l of SYBR Green I (Roche), 300 nM of both GAPDH primers or 400 nM of both telomere primers (Integrated DNA Technologies). For experimental samples, 10 ng of DNA was added into each reaction. qPCR plates were covered using LightCycler 480 sealing foils (Roche). We performed the qPCR reaction using a LightCycler 480 (Roche) with the telomere (95°C for 15 min, followed by 35 cycles of 15 s at 95°C, 30 s at 56°C, 30 s at 72°C) and GAPDH (95°C for 15 min, followed by 40 cycles of 15 s at 95°C, 30 s at 60°C, 30 s at 72°C) reactions run on

separate plates. The meltcurve analysis was performed at the end of each reaction to check for primer dimerization and the amplification of nontarget genes. Each plate contained a no template control (nuclease free water, Ambion) and a twofold serial dilution of the “golden sample” corresponding to 40, 20, 10, 5 and 2.5 ng of DNA. The “golden sample” was made up of DNA extracted using the QIAamp method and from multiple *M. coronatus* individuals. To normalize inter-assay variation, we used a value at 10 ng interpolated from the serial dilution linear regression line (this proved to yield slightly higher repeatabilities than the actual value at 10 ng, results not shown). All samples and controls were run in duplicate.

The standard curves included on each plate were also used for quality control purposes. All plates had standard curves within acceptable range. Coefficient of determination was  $> 0.99$  and efficiencies within of  $100 \pm 15\%$  (Telomere standard curve: mean = 1.90, standard deviation = 0.02; GAPDH standard curve: mean = 1.99, standard deviation = 0.04). The program `LINREGPCR` (version 2016) was used to calculate individual well qPCR efficiencies and Cq values (Ruijter et al., 2009). Samples were excluded if duplicate Cq values on the same plate were  $>0.5$  apart (3.8% were excluded explaining smaller sample sizes in results section; see Table S1). No exclusion criterion was applied to individual well efficiency because narrowing the efficiency range did not increase repeatability (see Table S2). For samples that met our quality criteria, rTL was calculated following equation 1 in Pfaffl (2001). In brief, individual well efficiency ( $E$ ) was raised to the power of the inter-plate control minus the well Cq value ( $\Delta Cq$ ). The resulting duplicate telomere, and separately, duplicate GAPDH values were averaged and divided by each other to create a single rTL value (Equation 1).

$$rTL = \frac{(E_{telomere})^{\Delta Cq_{telomere}}}{(E_{GAPDH})^{\Delta Cq_{GAPDH}}} \quad (1)$$

### 2.3.2 | Telomere measurement by PFGE-TRF

In order to estimate accuracy and validate the qPCR method for use in *M. coronatus*, we re-analysed  $n = 34$  samples using PFGE-TRF following published protocols that have been shown to work well for various avian species (Bauch et al., 2013; Salomons et al., 2009). In brief, to extract DNA from the snap frozen samples, the CHEF Genomic DNA Plug kit (Bio-Rad, CA, USA) was used with an overnight Proteinase K digestion at 50°C. A within plug restriction enzyme digestion followed using *Hind III* (60U), *Hinf I* (30U), and *Msp I* (60U) in  $10 \times$  NEB2 enzyme-buffer and incubated overnight at 37°C. The DNA was placed into a 0.8% pulse field certified agarose (Biorad) gel for pulse field gel electrophoresis (24 hr, 14°C at  $4.8 \text{ V cm}^{-1}$ , initial switch time 1 s, final switch time 25 s). Afterwards, the gel was dried and then hybridized overnight with the  $^{32}\text{P}$   $\gamma$ -ATP probe (5'—CCCTAA—3'). The labelled DNA and the  $^{32}\text{P}$  labelled size ladders (MidRange I PFG Marker (NEB) and DNA molecular weight marker XV (Roche)) were detected using a phosphor screen (PerkinElmer). Finally, telomere length size distributions were

quantified by densitometry in the program IMAGEJ (version 1.38 $\times$ ). Repeatability of PFGE-TRF method was assessed previously in Salomons et al. (2009), which reported a repeatability of 0.90 (Note: repeatability of PFGE-TRF was not assessed in *M. coronatus*).

### 3 | RESULTS

#### 3.1 | DNA extraction comparison

Digestion at 37°C typically resulted in a higher rTL than digestion at 56°C (Fig. S1a; 37°C: mean = 0.97, SD = 0.17; 56°C mean = 0.87, SD = 0.15; LMM:  $F = 8.84$ ,  $df = 14$ ,  $p = .01$ ,  $n = 30$  rTL estimates,  $n = 15$  individuals), whilst controlling for blood storage media (LMM:  $F = 0.14$ ,  $df = 12$ ,  $p = .87$ ). The interaction between digestion temperature and blood storage media was not included in the above LMM and was excluded because it was nonsignificant in the full model. GAPDH Cq values tended to be affected more by digestion temperature than telomere Cq values, suggesting the former to cause the overall effect of digestion temperature on rTL, although neither Cq effect reached statistical significance (see SI for details). DNA extraction method influenced rTL estimates (Fig. S2; LMM:  $F = 3.66$ ,  $df = 20$ ,  $p = .04$ ,  $n = 33$  rTL estimates,  $n = 11$  individuals) whilst controlling for blood storage media (LMM:  $F = 2.67$ ,  $df = 20$ ,  $p = .13$ ). The interaction between DNA extraction method and blood storage media was nonsignificant and was excluded from the above LMM. The QIAamp method produced 15% and 14% higher rTL estimates on average than both the ammonium acetate and Puregene methods, respectively (Fig. S2). Nonetheless, QIAamp rTL values were correlated with ammonium acetate ( $r = .95$ ) and Puregene methods ( $r = .72$ , for details see Fig. S3).

#### 3.2 | Effect of blood storage medium on rTL estimate and precision

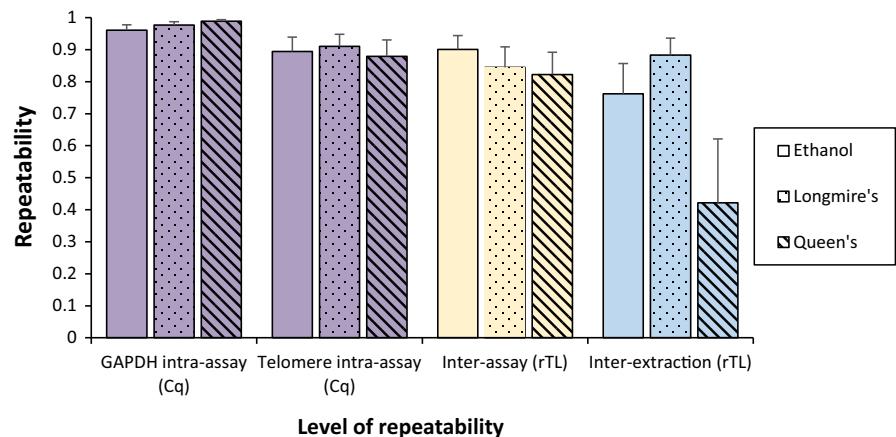
Overall rTL estimates were higher when the blood was stored in ethanol, whilst Longmire's and Queen's lysis produced telomere estimates in a similar range (Fig. S4; LMM:  $F = 52.86$ ,  $df = 130.03$ ,  $p < .001$ ). Inter-assay repeatability of rTL was similar between

ethanol, Longmire's and Queen's (Figure 2). However, there was large variation in inter-extraction rTL repeatability, with Longmire's and ethanol having a much higher ICC than Queen's (Figure 2). This was mirrored in overall repeatability of rTL calculated when controlling for qPCR run and DNA extraction in a LMM (Table S3). Intra-assay repeatability for GAPDH was similar across each of the blood storage media (see Figure 2 for repeatability estimates and confidence intervals). Intra-assay repeatability for telomere qPCR was lower than for GAPDH but was also similar across all storage media types.

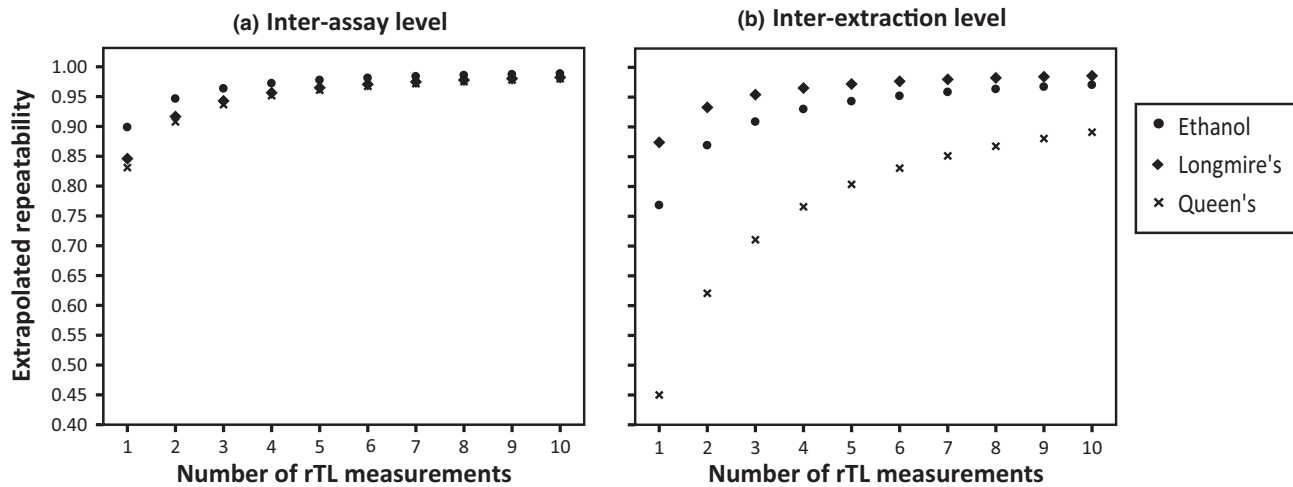
Extrapolated repeatability increases at a diminishing rate with each increase in the number of replicate measurements averaged (Figure 3). For example, Queen's lysis buffer had a repeatability that increased by 0.17 when averaging two extraction measurements and required approximately seven rTL measures from separate extractions to reach a repeatability  $>0.85$ . In contrast, at the intra-extraction level, extrapolated repeatability increases were much lower due to the already high repeatabilities. For example, ethanol, which was highly repeatable at the intra-extraction level, increased extrapolated repeatability by 0.05 when averaging two rTL measures.

#### 3.3 | Comparison of qPCR to PFGE-TRF

Correlation between TL calculated from PFGE-TRF and rTL-qPCR of the blood sample extracted from ethanol was high and significant (Figure 4a;  $r = .68$ ,  $n = 32$ ,  $p < .01$ ). Repeating the telomere qPCR and averaging the rTL estimates improved the correlation with TL from TRF (Figure 4b;  $r = .74$ ,  $n = 32$ ,  $p < .01$ ). A similar improvement in the correlation was achieved by re-extracting the same blood sample and averaging the two rTL estimate from the different extracts (Figure 4c;  $r = .73$ ,  $n = 32$ ,  $p < .01$ ). Using these correlation coefficients, we conducted power analyses to determine the benefit of measuring rTL multiple times in terms of required sample size to achieve a similar statistical power. The sample size required to find the correlation between rTL and TRF was 17.35 for a single rTL estimate, 13.9 for two rTL estimates and 14.8 for two rTL estimates from two separate extractions. This corresponds to an increase in statistical power equivalent to a 25% increase in sample size when



**FIGURE 2** Repeatability of duplicate rTL estimates varies with amplicon, storage medium and across levels of repeatability. The sample size for repeatability estimate was standardized to  $n = 48$ . Error bars represent 95% confidence intervals. Colours correspond to those in Figure 1 and represent different re-sampling levels. Patterns within each bar represent the particular storage medium used (see key) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]



**FIGURE 3** Extrapolated repeatability increases with the number of relative telomere length (rTL) measures taken, and the increase varies between storage media and type of replication (assay or extraction). The first value is the calculated repeatabilities at the (a) inter-assay, and (b) the inter-extraction levels, which were used to calculate extrapolated repeatability with increasing number of measurements. Combining such repeatability assessments and extrapolations is a simple method to help optimize sampling and analysis design according to costs and study constraints

using two measures from a single DNA extraction, and a 17% increase in sample size when using two measures from two DNA extractions.

## 4 | DISCUSSION

We studied the extent to which preanalytical methods (storage media and DNA extraction) affect accuracy and precision of rTL measurements obtained using qPCR, and the extent to which random measurement error can be reduced by replicating specific steps of the measurement protocol (Table 1). Our results showed that rTL estimates were influenced differentially by storage media, DNA extraction method and extraction digestion temperature, emphasizing the importance of methodological consistency. Averaging multiple rTL measures increases statistical power differently at different levels and for different storage media. Below we discuss these findings and make recommendations.

### 4.1 | rTL estimates

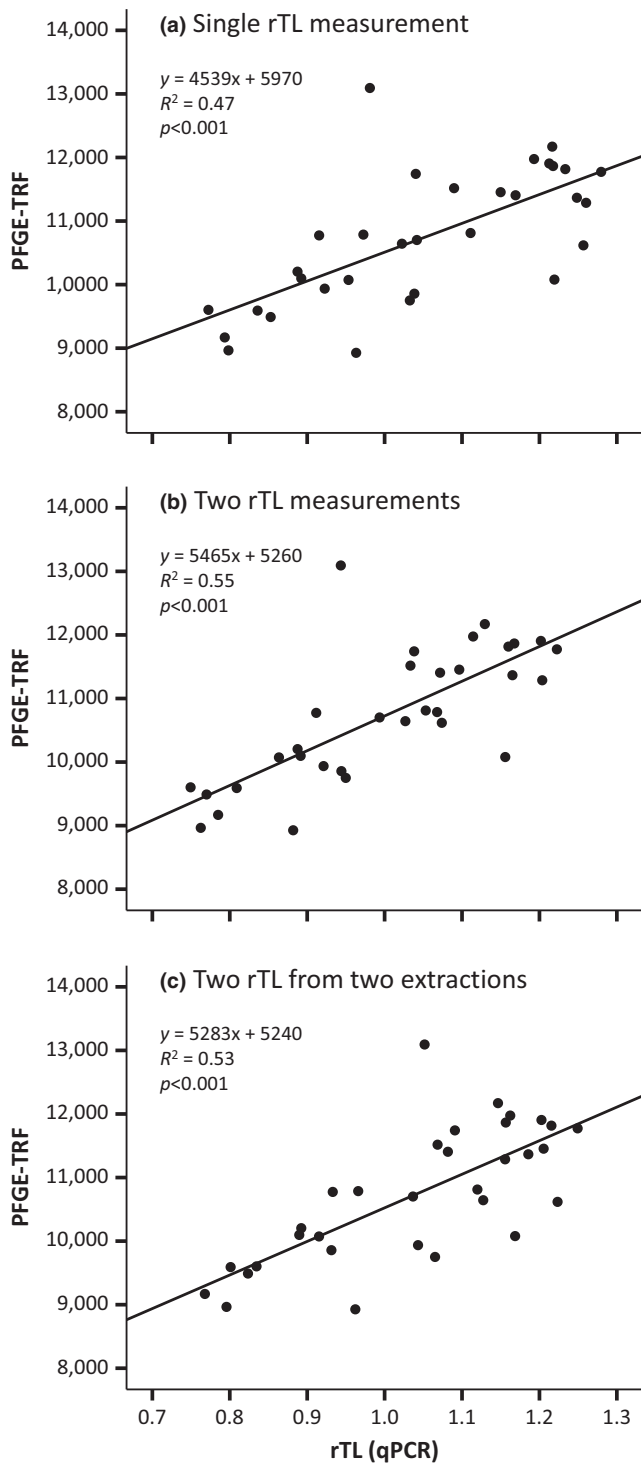
Estimates of rTL varied according to storage media and DNA extraction methods that are commonly used in ecology and evolution. rTL estimates are influenced by various preanalytical factors, such as source tissue (Reichert, Criscuolo, Verinaud, Zahn, & Massemin, 2013; Schmidt et al., 2016), the time period before analysis, DNA purity, integrity and storage (Denham et al., 2014; Dlouha et al., 2014; Zanet et al., 2013). Our study adds to a body of literature showing that DNA extraction also influences rTL estimates. However, our finding that QIAamp produced higher rTL estimates than both Puregene and ammonium acetate methods, contrasts with most previous evidence that column-based methods such as QIAamp tend

to produce lower rTL estimates (Cunningham et al., 2013; Hofmann et al., 2014; Raschenberger et al., 2016; Seeker et al., 2016; Tolios et al., 2015), although not always (Denham et al., 2014). Such differences could be due to variation in protocol, for example the use of a QIAcube HT instrument, or differences in degree of optimization or species differences. Thus, our data support calls for methodological consistency or suitable controls across studies (Horn et al., 2010; Nussey et al., 2014). As telomeric DNA is particularly sensitive to degradation which results in lower rTL estimates (Dlouha et al., 2014), it appears valid that methods yielding longer estimates are preferable (ethanol and QIAamp in this study).

### 4.2 | Assessing intra- and inter-assay precision

Precision (repeatability) was quantified at the intra-assay, inter-assay and inter-extraction levels. Telomere qPCR repeatability has thus far been primarily assessed at the level of the qPCR method itself, that is the intra-assay and inter-assay level. In our study, inter-assay repeatabilities for telomere estimation were within a similar range to what has previously been reported (0.77–0.93) (e.g., Asghar et al., 2015; Barrett et al., 2012; Parolini et al., 2015). At the intra- and inter-assay levels, there were no significant repeatability differences between storage media. As the largest effects at this level are presumably due to limited qPCR optimization and lack of DNA integrity, this lack of differences between storage media is in agreement with our quality control. However, to our knowledge, this is the first study to investigate inter-extraction repeatability. We found that repeatability was high, but significantly lower for Queen's lysis buffer at the inter-extraction level. Queen's lysis buffer was repeatable at the qPCR level and the lower precision only became evident at the inter-extraction level. As Longmire's lysis buffer had the highest precision, similar to that of ethanol storage, lysis itself is not likely to be the





**FIGURE 4** Congruence between qPCR rTL estimates and PFGE-TRF TL estimates is high and improves  $R^2$  by 17–13% when repeating the qPCR with the same or different DNA extract. Average TL estimated using TRF (PFGE-TRF) is plotted against (a) rTL estimated using a single measurement from one extraction. (b) rTL estimated using the average of two measurements from a single extraction. (c) rTL estimated using the average of two measurements from two extractions. Regression lines were calculated using ordinary least squares

primary cause. Queen's and Longmire's contain different lysing and protein denaturing reagents (n-Lauroylsarcosine and SDS, respectively Seutin et al., 1991; Longmire et al., 1997) and Longmire's has concentrations of both Tris and EDTA that are 10× greater than Queen's. Hence, one plausible but speculative explanation is that the DNA extraction method interacts differentially with storage media.

To determine to what extent averaging multiple estimates can increase precision, we calculated extrapolated repeatability at the inter-assay and inter-extraction levels. These calculations showed, not surprisingly, that averaging rTL estimates from multiple measures increases precision and these benefits were larger for the least repeatable (Queen's-stored) samples (inter-extraction level). A practical purpose for this approach is to determine how many extractions are needed to achieve high repeatability. For Queen's samples, two rTL measures from DNA extractions increased repeatability substantially, by 0.17, whilst seven extractions are needed to achieve a repeatability greater than 0.85. Because error reduces effect sizes by increasing the variance, and therefore statistical power, increasing precision increases statistical power (Kanyongo, Brook, Kyei-Blankson, & Gocmen, 2007). For example, the increase in repeatability from 0.76 to 0.86 when analysing a second extraction (ethanol at the inter-extraction level; Figure 3) would increase statistical power equivalent to a decrease in required sample size by 13% (Table 8 Kanyongo et al. (2007); one-way ANOVA, three groups, alpha = 0.05, power = 0.90). Hence, studies can increase statistical power by measuring rTL multiple times, particularly when sample size is constrained (increasing sample size usually has more effect on statistical power than increasing measurement precision).

#### 4.3 | Assessing accuracy: comparing qPCR to TRF

To assess accuracy of rTL estimation by qPCR, we compared rTL to average PFGE-TRF and found that the two were highly correlated, within a comparable range to other studies using birds (Crisuolo et al., 2009) or humans which have little interstitial telomeric repeats (Cawthon, 2002). The error (lack of correlation) in this instance may reflect measurement error in both techniques and fundamental difference in how TL is assessed, possibly related to the presence of interstitial telomeric repeats, which are included in qPCR rTL measurements, whereas PFGE-TRF as applied here measures exclusively terminal telomeres, that is at chromosome end (Foote et al., 2013).

Measuring telomere length multiple times and taking a mean value may increase accuracy. Indeed, the correlation between qPCR and PFGE-TRF improved by approximately 15% (Figure 2). The improvement in accuracy when averaging two rTL measures from a single extraction or averaging two rTL measures from two DNA extractions was similar in our sample. Hence, our data do not show any indication that increasing the number of extractions is any more beneficial than multiple rTL measures from a single extraction. To quantify the benefit of duplicating rTL estimates, we conducted a power analysis using the correlation coefficients between qPCR and

**TABLE 1** Summary of steps 1–3 used to identify the optimal procedure for measuring relative telomere length (rTL) using qPCR. Although the specific recommendations are likely to be relevant to the species used in this study (*Malurus coronatus*), this table represents a guide of decision rules for researchers wishing to optimize rTL-qPCR and highlights the importance of assessing repeatability at all levels

	Comparisons/approach	Decision rule	Recommendation for our system
Step 1	DNA extraction method (QIAamp, Puregene, ammonium acetate precipitation)	Longest rTL estimates	QIAamp
	DNA extraction digestion temperature (56°C or 37°C)	Longest rTL estimates	37°C
Step 2	Effect of blood storage medium on rTL estimate and repeatability (Ethanol, Longmire's lysis buffer and Queen's lysis buffer)	Longest rTL estimates High repeatability at the intra-assay, inter-assay and inter-extraction levels	Ethanol or Longmire's
	Improving repeatability using multiple measures	Depends on limiting factor: <ul style="list-style-type: none"> <li>Relative cost of extraction/qPCR vs. relative improvement with re-extracting/re-qPCR</li> <li>Sample number or re-sampling restrictions</li> <li>Low statistical power</li> </ul>	Multiple inter-extraction rTLs when using Queen's or low statistical power
Step 3	Comparison of rTL to PFGE-TRF	High correlation with "gold standard"	qPCR adequate
	Improving accuracy using multiple measures	Depends on limiting factor (see above)	Average multiple rTL measures when statistical power low

PFGE-TRF. Estimating rTL twice from a single extraction reduced the number of individuals required to be sampled by 25% to achieve the same statistical power as a single estimate. Studies that are limited by sample size or investigating small effects may thus quite simply be able to increase rTL accuracy by measuring rTL a second time.

#### 4.4 | Conclusions and recommendations

Our findings indicate that preanalytical steps can alter rTL estimates, accuracy and precision. Although the specific recommendations are only relevant to the species used in this study (*M. coronatus*), Table 1 lists the steps and decision rules that may impact results if not considered during optimization. Telomere qPCR was highly repeatable, with limited variation in intra- and inter-assay repeatability. However, assessing repeatability separately at the inter-extraction level revealed variation in repeatability for different storage media. Including inter-extraction comparisons is therefore critically important for the overall method precision and to estimate all potential sources of error. Our study showed that ethanol, a most commonly used storage medium in ecology, yields very good results when paired with an optimized assay and extraction protocol (Tables S1 and S2, Figs S1 and S2): ethanol-extracted DNA yielded the largest rTL estimates; ethanol has high precision that can be improved 10% by averaging a second extraction (Figure 3); ethanol-derived rTL estimates are accurate (high correlation with PFGE-TRF estimates) and accuracy can likewise be improved 7–9% by averaging two estimates, from one or two extractions (Figure 4).

Measuring rTL multiple times increases both the precision and accuracy of telomere qPCR. Sample sizes of free-living animals are not usually very large, suggesting substantial gains can be obtained from replicating measurements. Deciding whether to replicate measures depends on study constraints: although increasing sample size

is usually more effective in gaining statistical power, this might be costly, unethical or simply not possible. Given the significantly greater effect of repeating the extraction on precision gains, and the similar effect on gains in accuracy, we would recommend repeating extractions in such cases.

#### ACKNOWLEDGEMENTS

Research was approved by the Australian Bird and Bat Banding Scheme, the Western Australian Department of Environment and Conservation, the Australian Wildlife Conservancy, the Biological Sciences Animal Ethics Committee of Monash University (BSCI/2011/28) and the Department of Sustainability, Environment, Water, Population and Communities. Funding was provided by an Australian Research Council (ARC) Future Fellowship (FT 110100505) to AP; an ARC Discovery Project awarded to AP and SV (DP150103595) and Monash University. We would like to thank Nataly Hidalgo Aranzamendi, Niki Teunissen, Michael Roast and Marie Fan for collecting samples; Kaspar Delhey for statistical advice and two anonymous referees for insightful comments that improved the manuscript. Our special thanks to numerous volunteers during fieldwork and the support of staff at Australian Wildlife Conservancy's Mornington Wildlife Sanctuary.

#### DATA ACCESSIBILITY

All data presented in this manuscript are available from the DRYAD database (<https://doi.org/10.5061/dryad.7356h>).

#### AUTHOR CONTRIBUTIONS

J.E., S.V. and A.P. were involved in the study design. J.E. and E.M. optimized and performed the lab work. J.E. performed the data

analysis. J.E., S.V. and A.P. all contributed to interpreting the data and writing the manuscript.

## ORCID

Justin R. Eastwood  <http://orcid.org/0000-0002-5294-3321>

## REFERENCES

- Asgar, M., Hasselquist, D., Hansson, B., Zehndjiev, P., Westerdahl, H., & Bensch, S. (2015). Hidden costs of infection: Chronic malaria accelerates telomere degradation and senescence in wild birds. *Science*, *347*, 436–438.
- Atema, E., van Oers, K., & Verhulst, S. (2013). GAPDH as a control gene to estimate genome copy number in great tits, with cross-amplification in blue tits. *Ardea*, *101*, 49–54.
- Aubert, G., & Lansdorp, P. M. (2008). Telomeres and aging. *Physiological Reviews*, *88*, 557–579.
- Aviv, A. (2008). The epidemiology of human telomeres: Faults and promises. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *63*, 979–983.
- Aviv, A., Hunt, S. C., Lin, J., Cao, X., Kimura, M., & Blackburn, E. (2011). Impartial comparative analysis of measurement of leukocyte telomere length/DNA content by Southern blots and qPCR. *Nucleic Acids Research*, *39*, gkr634.
- Barrett, E. L. B., Boner, W., Mulder, E., Monaghan, P., Verhulst, S., & Richardson, D. S. (2012). Absolute standards as a useful addition to the avian quantitative PCR telomere assay. *Journal of Avian Biology*, *43*, 571–576.
- Bauch, C., Becker, P. H., & Verhulst, S. (2013). Telomere length reflects phenotypic quality and costs of reproduction in a long-lived seabird. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*, 20122540.
- Blackburn, E. H. (1991). Structure and function of telomeres. *Nature*, *350*, 569–573.
- Boardman, L. A., Skinner, H. G., & Litzelman, K. (2014). Telomere length varies by DNA extraction method: Implications for epidemiologic research—response. *Cancer Epidemiology and Prevention Biomarkers*, *23*, 1131.
- Boonekamp, J. J., Mulder, G. A., Salomons, H. M., Dijkstra, C., & Verhulst, S. (2014). Nestling telomere shortening, but not telomere length, reflects developmental stress and predicts survival in wild birds. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*, 20133287.
- Boonekamp, J. J., Simons, M. J. P., Hemerik, L., & Verhulst, S. (2013). Telomere length behaves as biomarker of somatic redundancy rather than biological age. *Aging Cell*, *12*, 330–332.
- Cawthon, R. M. (2002). Telomere measurement by quantitative PCR. *Nucleic Acids Research*, *30*, e47.
- Champely, S. (2016). *pwr: Basic Functions for Power Analysis*.
- Crisuolo, F., Bize, P., Nasir, L., Metcalfe, N. B., Foote, C. G., Griffiths, K., ... Monaghan, P. (2009). Real-time quantitative PCR assay for measurement of avian telomeres. *Journal of Avian Biology*, *40*, 342–347.
- Cunningham, J. M., Johnson, R. A., Litzelman, K., Skinner, H. G., Seo, S., Engelman, C. D., ... Boardman, L. A. (2013). Telomere length varies by DNA extraction method: Implications for epidemiologic research. *Cancer Epidemiology, Biomarkers & Prevention*, *22*, 2047–2054.
- Denham, J., Marques, F. Z., & Charchar, F. J. (2014). Leukocyte telomere length variation due to DNA extraction method. *BMC Research Notes*, *7*, 877.
- Dlouha, D., Maluskova, J., Kralova Lesna, I., Lanska, V., & Hubacek, J. A. (2014). Comparison of the relative telomere length measured in leukocytes and eleven different human tissues. *Physiological Research/Academia Scientiarum Bohemoslovaca*, *63*(Suppl 3), S343–S350.
- Eastwood, J. R., Berg, M. L., Spolding, B., Buchanan, K. L., Bennett, A. T. D., & Walder, K. (2015). Prevalence of beak and feather disease virus in wild *Platyercus elegans*: Comparison of three tissue types using a probe-based real-time qPCR test. *Australian Journal of Zoology*, *63*, 1–8.
- Eisenberg, D. T. A. (2016). Telomere length measurement validity: The coefficient of variation is invalid and cannot be used to compare quantitative polymerase chain reaction and Southern blot telomere length measurement techniques. *International Journal of Epidemiology*, *45*, dyw191.
- Eisenberg, D. T. A., Kuzawa, C. W., & Hayes, M. G. (2015). Improving qPCR telomere length assays: Controlling for well position effects increases statistical power. *American Journal of Human Biology*, *27*, 570–575.
- Factor-Litvak, P., Susser, E., Kezios, K., McKeague, I., Kark, J. D., Hoffman, M., ... Aviv, A. (2016). Leukocyte telomere length in newborns: Implications for the role of telomeres in human disease. *Pediatrics*, *137*, e2015392.
- Foote, C. G., Vleck, D., & Vleck, C. M. (2013). Extent and variability of interstitial telomeric sequences and their effects on estimates of telomere length. *Molecular Ecology Resources*, *13*, 417–428.
- Gomes, N. M. V., Ryder, O. A., Houck, M. L., Charter, S. J., Walker, W., Forsyth, N. R., ... Wright, W. E. (2011). Comparative biology of mammalian telomeres: Hypotheses on ancestral states and the roles of telomeres in longevity determination. *Aging Cell*, *10*, 761–768.
- Harley, C. B., Futcher, A. B., & Greider, C. W. (1990). Telomeres shorten during ageing of human fibroblasts. *Nature*, *345*, 458–460.
- Hausmann, M. F., & Heidinger, B. J. (2015). Telomere dynamics may link stress exposure and ageing across generations. *Biology Letters*, *11*, 20150396.
- Hausmann, M. F., & Marchetto, N. M. (2010). Telomeres: Linking stress and survival, ecology and evolution. *Current Zoology*, *56*, 714–727.
- Hausmann, M. F., & Vleck, C. M. (2002). Telomere length provides a new technique for aging animals. *Oecologia*, *130*, 325–328.
- Heidinger, B. J., Blount, J. D., Boner, W., Griffiths, K., Metcalfe, N. B., & Monaghan, P. (2012). Telomere length in early life predicts lifespan. *Proceedings of the National Academy of Sciences*, *109*, 1743–1748.
- Hidalgo Aranzamendi, N., Hall, M. L., Kingma, S. A., Sunnucks, P., & Peters, A. (2016). Incest avoidance, extrapair paternity, and territory quality drive divorce in a year-round territorial bird. *Behavioral Ecology*, *27*, 1808–1819.
- Hofmann, J. N., Hutchinson, A. A., Cawthon, R., Liu, C. S., Lynch, S. M., Lan, Q., ... Purdue, M. P. (2014). Telomere length varies by DNA extraction method: implications for epidemiologic research—letter. *Cancer Epidemiology and Prevention Biomarkers*, *23*, 1129–1130.
- Horn, T., Robertson, B. C., & Gemmill, N. J. (2010). The use of telomere length in ecology and evolutionary biology. *Heredity*, *105*, 497–506.
- Kanyongo, G., Brook, G., Kyei-Blankson, L., & Gocmen, G. (2007). Reliability and statistical power: How measurement fallibility affects power and required sample sizes for several parametric and nonparametric statistics. *Journal of Modern Applied Statistical Methods*, *6*, 81–90.
- Kingma, S. A., Hall, M. L., Segelbacher, G., & Peters, A. (2009). Radical loss of an extreme extra-pair mating system. *BMC Ecology*, *9*, 15.
- Koppelstaetter, C., Jennings, P., Hochegger, K., Perco, P., Ischia, R., Karoszkka, H., & Mayer, G. (2005). Effect of tissue fixatives on telomere length determination by quantitative PCR. *Mechanisms of Ageing and Development*, *126*, 1331–1333.
- Levy, M. Z., Allsopp, R. C., Futcher, A. B., Greider, C. W., & Harley, C. B. (1992). Telomere end-replication problem and cell aging. *Journal of Molecular Biology*, *225*, 951–960.

- Longmire, J. L., Maltbie, M., & Baker, R. J. (1997). *Use of "lysis buffer" in DNA isolation and its implication for museum collections*. Lubbock, TX: Museum of Texas Tech University.
- Martin-Ruiz, C. M., Baird, D., Roger, L., Boukamp, P., Kronic, D., Cawthon, R., ... von Zglinicki, T. (2015). Reproducibility of telomere length assessment: An international collaborative study. *International Journal of Epidemiology*, *44*, 1673–1683.
- McClintock, B. (1941). The stability of broken ends of chromosomes in *Zea mays*. *Genetics*, *26*, 234–282.
- Monaghan, P. (2010). Telomeres and life histories: The long and the short of it. *Annals of the New York Academy of Sciences*, *1206*, 130–142.
- Monaghan, P., & Haussmann, M. F. (2006). Do telomere dynamics link lifestyle and lifespan? *Trends in Ecology & Evolution*, *21*, 47–53.
- Nakagawa, S., Gemmell, N. J., & Burke, T. (2004). Measuring vertebrate telomeres: Applications and limitations. *Molecular Ecology*, *13*, 2523–2533.
- Nakagawa, S., & Schielzeth, H. (2010). Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. *Biological Reviews*, *85*, 935–956.
- Nussey, D. H., Baird, D., Barrett, E., Boner, W., Fairlie, J., Gemmell, N., ... Monaghan, P. (2014). Measuring telomere length and telomere dynamics in evolutionary biology and ecology. *Methods in Ecology and Evolution*, *5*, 299–310.
- O'Callaghan, N., Dhillon, V., Thomas, P., & Fenech, M. (2008). A quantitative real-time PCR method for absolute telomere length. *BioTechniques*, *44*, 807–809.
- Olsen, M. T., Bérubé, M., Robbins, J., & Palsbøll, P. J. (2012). Empirical evaluation of humpback whale telomere length estimates; quality control and factors causing variability in the singleplex and multiplex qPCR methods. *BMC Genetics*, *13*, 77.
- Parolini, M., Romano, A., Khoraiuli, L., Nergadze, S. G., Caprioli, M., Rubolini, D., ... Giulotto, E. (2015). Early-life telomere dynamics differ between the sexes and predict growth in the barn swallow (*Hirundo rustica*). *PLoS ONE*, *10*, e0142530.
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*, e45.
- R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Raschenberger, J., Lamina, C., Haun, M., Kollerits, B., Coassin, S., Boes, E., ... Kronenberg, F. (2016). Influence of DNA extraction methods on relative telomere length measurements and its impact on epidemiological studies. *Scientific Reports*, *6*, 25398.
- Reichert, S., Criscuolo, F., Verinaud, E., Zahn, S., & Massemin, S. (2013). Telomere length correlations among somatic tissues in adult zebra finches. *PLoS ONE*, *8*, e81496.
- Ruijter, J. M., Ramakers, C., Hoogaars, W. M., Karlen, Y., Bakker, O., van den Hoff, M. J., & Moorman, A. F. (2009). Amplification efficiency: Linking baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Research*, *37*, e45–e45.
- Salomons, H. M., Mulder, G. A., van de Zande, L., Haussmann, M. F., Linskens, M. H. K., & Verhulst, S. (2009). Telomere shortening and survival in free-living corvids. *Proceedings of the Royal Society of London B: Biological Sciences*, *276*, rspb20090517.
- Schmidt, J. E., Sirman, A. E., Kittilson, J. D., Clark, M. E., Reed, W. L., & Heidinger, B. J. (2016). Telomere correlations during early life in a long-lived seabird. *Experimental Gerontology*, *85*, 28–32.
- Seeker, L. A., Holland, R., Underwood, S., Fairlie, J., Psifidi, A., Ilska, J. J., ... Nussey, D. H. (2016). Method specific calibration corrects for DNA extraction method effects on relative telomere length measurements by quantitative PCR. *PLoS ONE*, *11*, e0164046.
- Seutin, G., White, B. N., & Boag, P. T. (1991). Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology*, *69*, 82–90.
- Sudyka, J., Arct, A., Drobniak, S., Dubiec, A., Gustafsson, L., & Cichoń, M. (2014). Experimentally increased reproductive effort alters telomere length in the blue tit (*Cyanistes caeruleus*). *Journal of Evolutionary Biology*, *27*, 2258–2264.
- Tolios, A., Teupser, D., & Holdt, L. M. (2015). Preanalytical conditions and DNA isolation methods affect telomere length quantification in whole blood. *PLoS ONE*, *10*, e0143889.
- Verhulst, S., Susser, E., Factor-Litvak, P. R., Simons, M. J., Benetos, A., Steenstrup, T., ... Aviv, A. (2015). Commentary: The reliability of telomere length measurements. *International Journal of Epidemiology*, *44*, dyv166.
- Zanet, D. L., Saberi, S., Oliveira, L., Sattha, B., Gadawski, I., & Côté, H. C. F. (2013). Blood and dried blood spot telomere length measurement by qPCR: Assay considerations. *PLoS ONE*, *8*, e57787.
- von Zglinicki, T., Bürkle, A., & Kirkwood, T. B. L. (2001). Stress, DNA damage and ageing—an integrative approach. *Experimental Gerontology*, *36*, 1049–1062.

## SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

**How to cite this article:** Eastwood JR, Mulder E, Verhulst S, Peters A. Increasing the accuracy and precision of relative telomere length estimates by RT qPCR. *Mol Ecol Resour*. 2018;18:68–78. <https://doi.org/10.1111/1755-0998.12711>