ORIGINAL ARTICLE

Measuring sperm swimming performance in birds: effects of dilution, suspension medium, mechanical agitation, and sperm number

Emily R. A. Cramer[1,2] · Melissah Rowe[1,3] · Fabrice Eroukhmanoff[3] · Jan T. Lifjeld[1] · Glenn-Peter Sætre[3] · Arild Johnsen[1]

Abstract

Sperm swimming performance, including swimming speed and the proportion of motile cells, may strongly affect fertilization success. However, little is known about how methodological factors affect in vitro measurement of these parameters. We compare the swimming performance of sperm from House Sparrows (*Passer domesticus*) and Spanish Sparrows (*Passer hispaniolensis*) in two standard suspension media, at two different dilutions, and with different degrees of cell agitation. Further, we conduct a resampling analysis to investigate sample size effects. Sperm performance was generally reduced when sperm were diluted, or when suspended in phosphate-buffered saline (PBS) rather than a medium containing additional nutrients. Sperm performance was particularly low when they were diluted more and suspended in PBS, suggesting that seminal fluid may provide compounds that enhance performance but that these are less available following dilution. Mechanical agitation of the cells by vigorous pipetting increased the proportion of motile cells. Between-male repeatability, assessed on a single sample measured in multiple conditions, was moderate and significant, suggesting that similar results may be obtained regardless of the methodology used to assess sperm motion. We found no evidence of biased results when low numbers of cells per male were used in analysis, though precision increased substantially as sample size increased from five to 20 or more cells per male. We recommend using the same suspension media and similar sperm concentrations and levels of agitation, to the greatest degree possible, and including as many individuals as possible in analyses, even when some individuals are represented by few sperm cells.

Keywords Postcopulatory sexual selection · Sperm competition · Cryptic female choice · Experimental design · Sperm behavior

Zusammenfassung

Messung der Schwimmfähigkeit von Spermien bei Vögeln: Verdünnungseffekte, Suspensionsmedium, mechanische Durchmischung und Spermienzahl Die Schwimmfähigkeit der Spermien, einschließlich der Schwimmgeschwindigkeit

---

✉ Emily R. A. Cramer
  ERC25@cornell.edu

1   Natural History Museum, University of Oslo, P.O. Box 1172, Blindern, 0318 Oslo, Norway

2   Present Address: Premedical Education Department, Weill Cornell Medicine Qatar, P.O. Box 24144, Doha, Qatar

3   Department of Biosciences, Center for Ecological and Evolutionary Synthesis, University of Oslo, P.O. Box 1066, Blindern, 0316 Oslo, Norway

und dem Anteil an beweglichen Zellen, könnte den Befruchtungserfolg stark beeinflussen. Jedoch ist wenig darüber bekannt, wie die methodischen Faktoren die in vitro-Messungen dieser Parameter beeinflussen. Wir vergleichen die Schwimmfähigkeit der Spermien von Haus- und Weidensperlingen (*Passer domesticus* und *P. hispaniolensis*) in zwei verschiedenen Standardsuspensionen, in zwei verschiedenen Verdünnungen und unter unterschiedlich starker Durchmischung (Agitation) der Zellen. Weiterhin führten wir eine wiederholte Probenanalyse durch, um von der Stichprobengröße abhängige Effekte zu untersuchen. Die Spermienleistung reduzierte sich grundsätzlich bei der Verdünnung der Spermien oder bei Zugabe einer phosphatgepufferten Salzlösung (PBS; engl. phosphate-buffered saline) anstelle eines zusätzlich mit Nährstoffen angereicherten Mediums (DMEM). Die Spermienleistung war besonders gering bei einer stärkeren Verdünnung oder bei der Zugabe von PBS, was vermuten lässt, dass die Samenflüssigkeit bestimmte Komponenten zur Steigerung der Schwimmfähigkeit beinhaltet, welche jedoch nach einer Verdünnung nur noch im geringeren Maße zur Verfügung stehen. Eine mechanische Agitation der Zellen durch starkes Pipettieren erhöhte den Anteil an beweglichen Zellen. Die Wiederholbarkeit der Messungen zwischen den untersuchten Männchen, basierend auf einzelnen Proben, welche unter verschiedenen Bedingungen gemessen wurden, war moderat und signifikant. Dies lässt vermuten, dass gleichartige Ergebnisse unabhängig von der für die Messung der Spermienmobilität angewandten Methode erzielt werden können. Wir haben keine Anhaltspunkte dafür gefunden, dass eine geringe Spermienanzahl pro Männchen die Analyseergebnisse beeinflusst. Jedoch nahm bei der Erhöhung der Stichprobengröße von fünf auf 20 oder mehr Zellen pro Männchen auch die Präzession zu. Wir empfehlen - sofern möglich - die Anwendung des gleichen Suspensionsmediums, einer gleichartigen Spermienkonzentration und einen gleichbleibenden Durchmischungsgrad. Weiterhin sollten so viele Individuen wie möglich in einer Analyse berücksichtigt werden, auch wenn einige Individuen nur durch wenige Spermienzellen repräsentiert werden.

Introduction

In many bird species, females copulate with multiple males (Griffith et al. 2002), which implies that sperm and ejaculate characteristics can strongly affect male reproductive success, via competition between sperm from different males and/ or via cryptic female choice for certain sperm traits (Parker 1970; Eberhard 1996). Accordingly, there has been a recent surge of interest in examining the evolution of post-copulatory traits in birds, particularly wild passerines (e.g., Calhim et al. 2007; Lifjeld et al. 2010; Rowe et al. 2013). Of primary consideration has been the morphology of sperm cells, the speed at which sperm swim, and associations between morphology and swimming speed (Lüpold et al. 2009; Kleven et al. 2009; Cramer et al. 2015), with swimming speed often hypothesized to be a key functional aspect of sperm performance (Bennison et al. 2014). Because of the putative importance of sperm swimming speed to fertilization success, it is important to assess methodologies for measuring sperm swimming speed, as has been done with sperm morphology (Laskemoen et al. 2007; Schmoll et al. 2016).

Most studies on passerine sperm have recorded videos of swimming cells through a microscope, then tracked individual cells through multiple video frames using specialized commercial or open-access software (e.g., Wilson-Leedy and Ingermann 2007; Lüpold et al. 2009; Kleven et al. 2009). To conduct such studies, ejaculates must first be diluted sufficiently that individual cells can be visualized on the video recording, without diluting to a point where too few cells are visible. In this study, we focused on four factors that are likely to arise in any study performing video recording of sperm cells: the choice of the suspension medium,

the degree to which the ejaculate is diluted, the vigor with which sperm cells and the suspension medium are mixed before being deposited on the microscope slide, and the number of individual sperm cells tracked per male. Degree of dilution and type of suspension medium affect sperm swimming parameters in several mammal species (Farrell et al. 1996; Rijsselaere et al. 2003), and studies optimizing video recording have been conducted in some wild vertebrates (Fasel et al. 2015; Humann-Guilleminot et al. 2018). However, systematic data are currently lacking for passerine birds. Here, we compare swimming performance in two suspension media, two degrees of dilution, and two levels of agitation testing the sperm of two closely related species, the House Sparrow (*Passer domesticus*) and the Spanish Sparrow (*Passer hispaniolensis*). We then use a resampling analysis to investigate the impact of including different numbers of sperm cells per male.

Methods

Study subjects and general methods

We conducted experiments on 25 April 2014 using populations of House and Spanish Sparrows that had been in captivity in Oslo, Norway since 2010 (see full details in Cramer et al. 2014). Birds were confirmed to be in breeding condition, as 19 of the 23 males captured produced sperm samples, and courtship and copulation behaviors were regularly observed. For all experiments, we collected sperm samples into a capillary tube via cloacal massage (Wolfson 1952) and deposited the sample onto a sheet of Parafilm;

this procedure typically yielded at least 3 µL of sample per male. Sperm were then immediately pipetted into suspension media for experiment 1 (see details below), where we tested the effects of suspension medium and degree of dilution in a fully crossed design. The effects of mechanical agitation via pipetting were examined separately in experiment 2, which we conducted immediately after experiment 1, using excess from the experiment 1 treatment that we judged to have the best sperm cell density for video analysis. Therefore, we conducted experiments 1 and 2 sequentially on each sample, before progressing to the next male (Figs. S1, S2). Excess sperm samples were placed in 300 µL of 5% formaldehyde to allow for morphological analysis, and these samples and video recordings were accessioned to the University of Oslo Natural History Museum sperm collection (Table S1).

For both experiments, we filmed sperm swimming behavior using a HDR-HC1E Sony camera attached to an Olympus CX41 microscope with a heated Tokai Hit TP-S glass stage (TP-S, Gendoji-cho, Shizuoka, Japan). All solutions and glassware were pre-warmed to 40 °C before contact with the sperm, and were maintained at 40 °C throughout recording. We used Leja four-chambered microscope slides (20-µm depth; Nieuw-Vennep, Netherlands), which fill via capillary action.

Experiment 1: effects of dilution and suspension medium

For the different suspension media, we compared Dulbecco's modified Eagle medium (DMEM), a commonly used medium for sperm swimming analysis; e.g., Kleven et al. (2009) to phosphate-buffered saline (PBS), a medium that lacks the sugars and amino acids in DMEM, but that has been used as a neutral medium in several studies (e.g., Laskemoen et al. 2008; Cramer et al. 2014, 2016a, b). For the degrees of dilution, we created one treatment (termed "concentrated") where sperm cells were as dense as possible while still allowing analysis software to identify separate sperm cells. The other treatment (termed "diluted") was a 2:7 dilution of the concentrated treatment. We chose this dilution ratio to match methods used in previous studies (Cramer et al. 2014, 2016a, b) and to achieve a substantial increase in dilution while still obtaining movement data on a large number of sperm cells. The exact concentration of sperm cells from each sample, however, was not determined.

To conduct experiments, we pipetted 1.2 µL sperm from the Parafilm into 74 µL of suspension medium (DMEM or PBS). After mixing thoroughly by pipetting up and down four to six times, we transferred 17 µL of this concentrated sperm solution into 42 µL of the same medium and mixed this by pipetting up and down four to six times. To equalize pipetting between the concentrated and dilute treatments, we then re-pipetted the concentrated sperm four to six times

[note that our dilution scheme creates approximately equal end volumes in the concentrated (57 µL) and dilute (59 µL) treatments, so that pipetting an equal number of times should have similar effects in each treatment]. We then repeated this procedure with the other medium. Finally, we loaded 2.9 µL of each mixture into a chamber on a four-chambered Leja slide, and filmed sperm swimming in each chamber. We alternated which solution was mixed first, and we rotated which treatment was loaded into each slide chamber among males, to reduce potential biases. Mixing cells and loading the slide fully took approximately 1.5 min.

Each slide chamber was filmed in four different locations, to increase the number of sperm cells filmed. To allow us to investigate how sperm behavior changed over time, filming of the different slide chambers was interspersed (i.e., in the order a b c d d c b a a b c d d c b a, where each letter represents a slide chamber; Fig. S2). Approximately 5 s elapsed between each successive period of filming, as we switched chambers, found filming locations without air bubbles or other imperfections or contaminants, and paused to ensure a sufficiently long still period for analysis (approximately 1 s). We therefore recorded the single sample in all four treatments (DMEM concentrated, DMEM diluted, PBS concentrated, and PBS diluted) in rapid succession, with filming typically lasting approximately 1.5 min in total.

Experiment 2: effects of mechanical agitation

Immediately after filming a sample in experiment 1, we chose the treatment from experiment 1 that we judged to be best for video analysis, based on visual inspection of cell density on the video recording, and we used that treatment as the source of sperm for experiment 2. In all but two cases, the diluted treatment from experiment 1 was judged to have a better cell density for analysis and thus was the source of sample for experiment 2; to simplify analysis, we excluded those two experiments where concentrated sperm was used. By chance, we chose only DMEM-diluted sperm from Spanish Sparrow males, but chose both DMEM and PBS diluted sperm from House Sparrow males.

For the low agitation treatment, we transferred 2.9 µL from the original treatment tube to one chamber on a four-chambered Leja slide. These cells had been in suspension at 40 °C without agitation during filming for experiment 1, with a total duration of approximately 3.5 min between collecting the sample and beginning to video record for experiment 2. For the high-agitation treatment, we mixed the solution in the same tube by pipetting up and down five to ten more times, with the pipette volume set above 15 µL. We then loaded 2.9 µL of suspended sperm onto an adjacent chamber on the microscope slide and filmed each chamber in an interspersed fashion, as above (e.g., a b b a a b), in three to five locations (filming in more locations when we judged

that we had filmed fewer cells). While the low-agitation treatment was thus always loaded onto the microscope slide a few seconds before the high-agitation treatment, we rotated which treatment was filmed first by alternating the assignment of treatments to slide chambers.

Analysis of sperm swimming velocity and the proportion of motile sperm

All videos were analyzed with the software Hamilton Thorne CEROS II Sperm Analyzer (Hamilton Thorne Research, Beverly, MA). We analyzed 0.5 s of video, at a frame rate of 50 Hz, from each of the recording locations. To exclude air bubbles and contaminants, we excluded all detections with an elongation score > 50 from the data set. Detections with a straight-line velocity < 25 or average-path velocity < 30 were typically cells moving by drift, rather than swimming. These cells were considered immotile in calculating the proportion of motile cells and were excluded from analyses on sperm swimming speed (see also Cramer et al. 2014).

For our velocity measure, we used the curvilinear velocity (VCL), which follows the motions of the sperm cells most closely, following the logic of Laskemoen et al. (2010) that more-derived measurements may be less informative in in vitro calculations. To exclude inaccurate tracks (for example, tracks where the software switched sperm cells between successive detections), we excluded tracks that failed to detect a cell at each successive timepoint, where track straightness was less than 80 or track linearity was less than 35. Only tracks with at least ten detection points were included. Further, we excluded a track if any single movement between successive detections was greater than five interquartile ranges for the other movement distances in that track.

Statistical analysis

A summary of statistical analyses is given in Table S2. To assess treatment effects, we constructed linear mixed models. Models concerning sperm swimming speed used the measures from individual sperm cells as data points, while models on the proportion of motile cells used the proportion of motile cells in each recording location as data points. Velocity results were similar when averages for each recording location were used instead of individual cells (not shown). We included a random effect of slide chamber nested within male identity to account for having multiple data points per chamber per male. In velocity measures, we further included a random effect of filming location nested within chamber, to account for possible non-independence of cells filmed together; this was not possible for the proportion of motile cells as each filming location was used as a data point. The nested random effect structure significantly

improved model fit in all cases [assessed via a likelihood ratio test on models fit with restricted maximum likelihood (REML) estimation, and full parameterization of fixed effects (Zuur et al. 2009)]. The models initially included an interaction term between time since the beginning of recording and other variables of interest (below) because previous work (Cramer et al. 2016a, b) shows that sperm behavior changes over the time it takes to obtain video recordings for each sample. Specifically, for experiment 1, we began with a four-way interaction between degree of dilution, type of suspension medium, species and time since the beginning of filming, as well as all constituent lower order interactions (analysis 1, Table S2). For experiment 2, as indicated above, we began with an unbalanced subset of dilution and suspension medium treatments. Therefore, we began with a three-way interaction between agitation treatment, suspension medium, and time (as well as constituent pairwise interactions), and pairwise interactions of species with agitation treatment and with time. A four-way interaction was not possible because all Spanish Sparrow males were recorded in DMEM for experiment 2 (analysis 2, Table S2). To reduce the issue of "cryptic" multiple testing in model selection, we followed Forstmeier and Schielzeth (2011)'s recommendation to first compare the global model (with all fixed effects and interactions) to a null model. As the likelihood ratio tests were significant in all cases ($p < 0.0001$, not shown), we proceeded to simplify models by removing non-significant interactions ($p > 0.05$). We began by removing the highest order interactions, until only significant interactions, or lower order interactions supporting a higher order significant interaction, remained. Finally, we applied false discovery rate correction to F-test results for each test (Verhoeven et al. 2005; Forstmeier and Schielzeth 2011). We base interpretations on corrected values and report the raw values. Because model residuals approximated normality for both velocity and the proportion of motile cells, we modeled each response variable as normally distributed.

During the analysis, we noticed that the initial values of proportion motile for experiment 2 seemed higher than the same males' final values for experiment 1, which were measured approximately 30 s earlier. To test whether this effect was real, and to assess whether it also occurred in velocity, we calculated the average velocity of sperm cells and the overall proportion of motile cells, in the second half of experiment 1 and the first half of experiment 2 (analysis 3, Table S2). We compared these values using paired nonparametric Wilcoxon tests. Only data from the appropriate treatment in experiment 1 were included (i.e., the treatment that was used for experiment 2).

For analyses 1–3, we included data only from males with at least five well-tracked motile cells in the first and second half of the video recording for velocity models, and only data from males with at least 15 total cells detected in both

the first and second half of the video recordings for proportion of motile models. These cut-off values were chosen to include data from a large number of males, while also having data from a moderate to high number of cells for each male. To best allow within-male comparisons, only males that met these criteria for all experimental conditions were included.

To assess whether males' overall sperm performance rank relative to other males was robust to different measurement conditions, we analyzed repeatability, defined as the percent of variance that could be attributed to a random effect of male identity in a mixed-effects model (Nakagawa and Schielzeth 2010; analysis 4, Table S2). For this analysis, in order to simplify the dataset, we calculated the average swimming speed or total proportion of motile cells in each of the six experimental conditions (four combinations of dilution and suspension medium from experiment 1 and agitation combinations from experiment 2). We used only males for which there were at least 20 detected cells (proportion motile) or at least ten well-tracked motile cells (VCL), and fit a linear mixed-effects model with REML estimation in the package nlme (Pinheiro et al. 2013). Following the recommendation of Zuur et al. (2009) that all possible fixed effects of interest should be included when assessing significance of random effects, we included an interaction between species and the six-category variable describing the recording conditions. To determine the significance of repeatability, we compared the mixed-effect model to a model including only the fixed effects using a likelihood ratio test (Zuur et al. 2009; Pinheiro et al. 2013). Note that only a single sample was measured per male. Further, we calculated the coefficient of variation (SD/mean $\times$ 100) for the proportion of motile cells and for VCL among males. Bias-corrected and accelerated confidence intervals around the coefficient of variation were calculated using 1000 replicates using the package boot (Canty and Ripley 2017).

Finally, we conducted a resampling analysis to assess the impact of the number of sperm cells recorded for each male on sperm swimming speed analysis, by drawing cells from each male from both the concentrated DMEM and PBS treatments of experiment 1 (analysis 5, Table S2). For each resampled data set, we assessed repeatability as above, with a three-way fixed effect interaction between species, time, and suspension medium, and constituent pair-wise interactions. Additionally, for each resampled data set, we compared males' mean speed in DMEM and PBS using a paired t-test. Finally, we calculated the difference between the resampled mean and the male's grand mean (including all measured cells) in the DMEM treatment only (chosen arbitrarily). In order to include individuals with a large number of cells to resample from, and to minimize additional variation that we would need to account for, we included only males with at least 150 cells detected in both the concentrated PBS and concentrated DMEM treatments in experiment 1. We did not include recordings from diluted treatments nor recordings from experiment 2. We randomly sampled with replacement 5, 10, 20, 30, 40, 50, 100, or 125 cells from each treatment and each male, 100 times for each sample size.

All analyses were conducted in R (3.3.0, R Development Core Team). Unless otherwise noted, mixed-effects models were constructed using package lme4 (Bates et al. 2014) with statistical significance assessed via package lmerTest (Kuznetsova et al. 2014). Marginal (r_m^2) and conditional (r_c^2) r^2-values were calculated in package MuMIn (Barton 2016), which reflect variation explained by fixed effects only and by fixed and random effects together, respectively. Trend lines (with 95% confidence intervals for visualization purposes) and other graphs were constructed in ggplot2 (Wickham 2009). Normality of residuals was assessed visually, following Zuur et al. (2009).

Results

Experiment 1: effects of dilution and suspension medium (analysis 1)

Effects on sperm speed

Swimming speed depended on the interaction between suspension medium and concentration (Fig. 1; $F_{1,205.1} = 10.06$, $p = 0.002$; $n = 8$ House Sparrow and four Spanish Sparrow males, 9201 cells, $r_m^2 = 0.21$, $r_c^2 = 0.39$; Table S3). At the start of filming (i.e., estimated intercept from the statistical models), swimming speed was 6–9 μm/s slower in the diluted PBS treatment than in any of the other treatments ($|t| > 2.7$, $p < 0.01$), while differences among other treatments were not significant ($|t| < 0.81$, $p > 0.4$). Spanish Sparrow sperm swimming speed tended to be reduced in PBS compared to the reduction in PBS for House Sparrows, though this difference was not significant following correction for multiple testing (corrected $p = 0.054$).

Sperm swimming speed declined over time for all treatment combinations (all $|t| > 8.2$, $p < 0.001$), and the rate of decline over time depended on the suspension medium (interaction $F_{1,205.4} = 5.78$, $p = 0.02$). Swimming speed declined more quickly in PBS than in DMEM ($t_{205.4} = 2.4$, $p = 0.02$), and tended to decline faster when sperm were diluted than when concentrated, though this effect was not significant following correction for multiple testing (corrected $p = 0.054$).

Effects on the proportion of motile cells

For the proportion of motile cells, the simplified model included a three-way interaction between time, suspension medium, and concentration (Fig. 1, $F_{1277.58} = 5.52$,

Fig. 1 Effects over time of suspension medium and the degree of dilution on sperm swimming speed (**a**, **b**) and the proportion of motile cells (**c**, **d**) in House (**a**, **c**) and Spanish (**b**, **d**) Sparrows. Each individual was represented in all four treatments, allowing within-individual tests. *Dark grey lines* Concentrated treatment, *light grey lines* dilute treatment, *solid lines* cells suspended in Dulbecco's modified Eagle medium (DMEM), *dashed lines* cells suspended in phosphate-buffered saline (PBS). *Shaded areas* 95% Confidence intervals
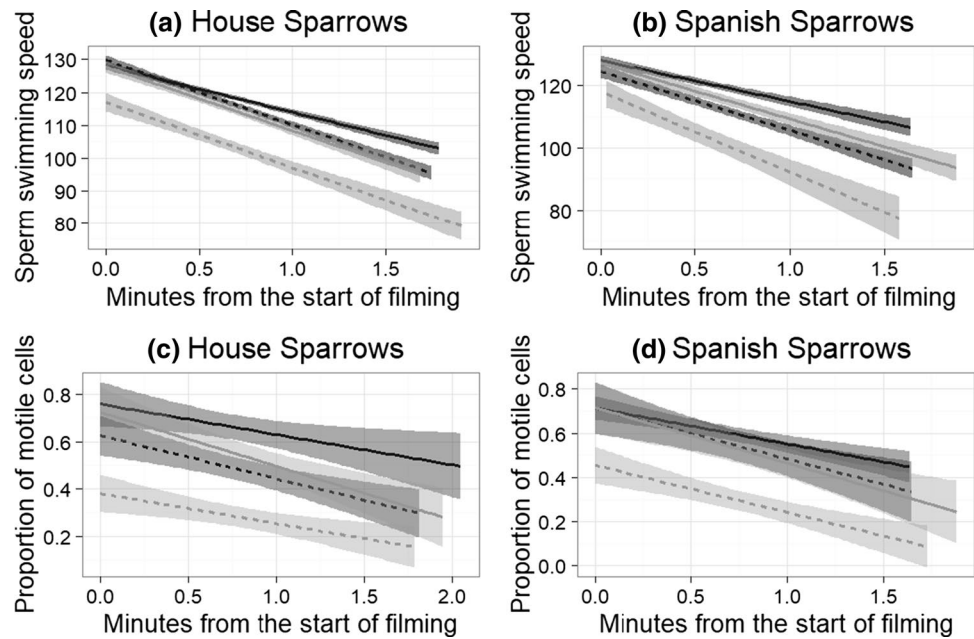
$p = 0.02$) as well as significant pairwise interactions of species with time ($F_{1,272.84} = 4.42$, $p = 0.01$; $n = 12$ House Sparrow and five Spanish Sparrow males, 343 filming locations, $r_{m}^2 = 0.44$, $r_{c}^2 = 0.79$; Table S4). The proportion of motile cells was lower in dilute than concentrated treatments when cells were suspended in PBS ($t_{113.8} = -5.97$, $p < 0.001$) but not in DMEM ($t_{122.3} = -1.02$, $p = 0.31$). Cells suspended in DMEM had a higher proportion motility than cells suspended in PBS for both degrees of dilution ($t > 2.3$, $p < 0.03$).

The proportion of motile cells decreased over time in all treatments ($|t| > 4.5$, $p < 0.001$). Decline in the proportion of motile cells over time was greater in Spanish Sparrows than House Sparrows ($t_{272.8} = 2.48$, $p = 0.01$). Decline in the proportion of motile cells was faster in dilute cells suspended in DMEM compared to diluted cells suspended in PBS ($t_{276.7} = 2.48$, $p = 0.01$). Other pairwise comparisons were not significant ($|t| < 1.4$, $p > 0.15$), though the decline tended to be faster in diluted than concentrated cells suspended in DMEM ($t_{276.9} = -1.93$, $p = 0.055$).

Experiment 2: effects of mechanical agitation (analysis 2)

Effects on sperm swimming speed

The simplified model included a significant pairwise interaction between time and suspension medium (Fig. 2; $F_{1,99.0} = 10.89$, $p = 0.001$) and between time and species ($F_{1,85.9} = 20.27$, $p < 0.001$, $n = 9$ House Sparrow and four Spanish Sparrow males, 2749 cells; $r_{m}^2 = 0.17$, $r_{c}^2 = 0.42$; Table S5). The effect of agitation level was not significant ($F_{1,87.1} = 0.2$, $p = 0.66$). Initial sperm swimming speed

was faster for Spanish Sparrow than for House Sparrow ($t_{33.58} = 5.41$, $p < 0.001$), and was faster for sperm suspended in PBS ($t_{40.0} = 2.37$, $p = 0.02$). Sperm swimming speed declined over time in all treatments ($|t| > 4.0$, $p < 0.001$) except for House Sparrow sperm in DMEM ($t_{86.0} = -0.51$, $p = 0.61$). Comparisons between suspension media in this experiment were between-male tests.

Effects on the proportion of motile cells

The final model included significant pairwise interactions between suspension medium and time ($F_{1,84.51} = 5.52$, $p = 0.02$); and species and time ($F_{1,84.49} = 8.34$, $p = 0.004$; $n = 9$ House Sparrow and five Spanish Sparrow males, 15 analysis frames; $r_{m}^2 = 0.23$, $r_{c}^2 = 0.73$; Table S6). The initial proportion of motile cells was higher in the more highly agitated treatments ($t_{12.55} = 3.78$, $p = 0.002$). The proportion of motile cells declined more quickly in Spanish Sparrows than in House Sparrows ($t_{84.5} = 2.89$, $p = 0.004$), and more quickly in PBS than in DMEM for both species ($t_{84.5} = 2.35$, $p = 0.02$). Specifically, the proportion of motile cells did not decline significantly over time for cells suspended in DMEM ($t_{85.2} = 0.01$, $p = 0.99$), but did decline significantly for cells suspended in PBS ($t_{84.2} = -2.86$, $p = 0.005$).

Comparison of experiments 1 and 2 (analysis 3)

In paired tests, the proportion of motile cells at the beginning of experiment 2 was higher than at the end of experiment 1 (Wilcoxon $V = 96$ and $V = 104$, $p < 0.005$, $n = 14$ males; Fig. 3), despite the apparent decline over time in cell motility during filming in experiment 1 and the fact that we

Fig. 2 Effects over time of mechanical agitation on sperm swimming speed (**a**, **b**) and the proportion of motile cells (**c**, **d**) in House (**a**, **c**) and Spanish (**b**, **d**) Sparrows. *Dark grey lines* High agitation level, *light grey lines* low agitation level; *solid lines* cells suspended in DMEM, *dashed lines* cells suspended in PBS. Each individual sample was measured for high and low levels of agitation, but only in a single suspension medium and degree of dilution for this experiment; Spanish Sparrows samples were tested only in DMEM. *Shaded areas* 95% confidence intervals. For abbreviations, see Fig. 1
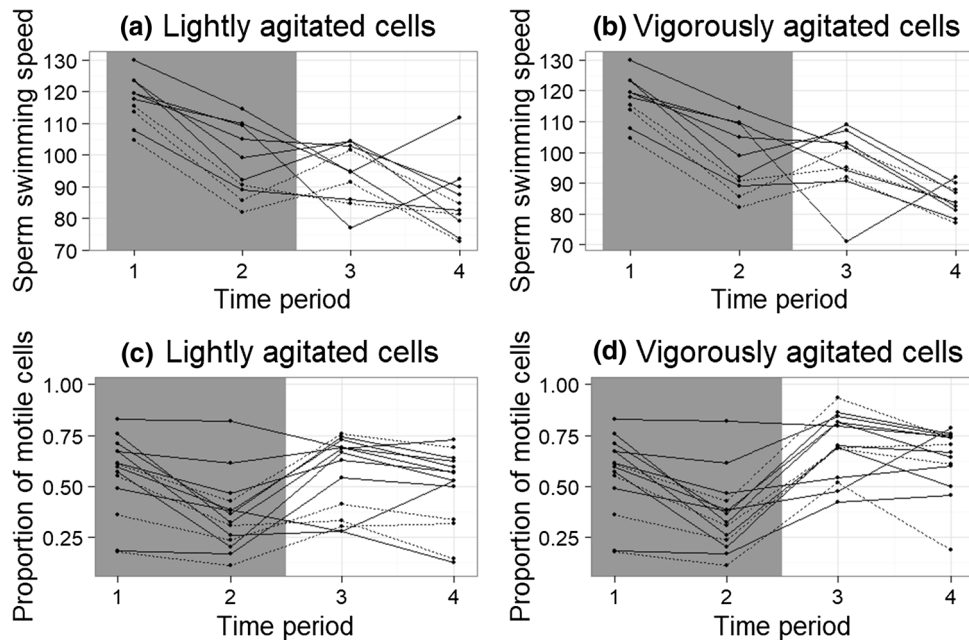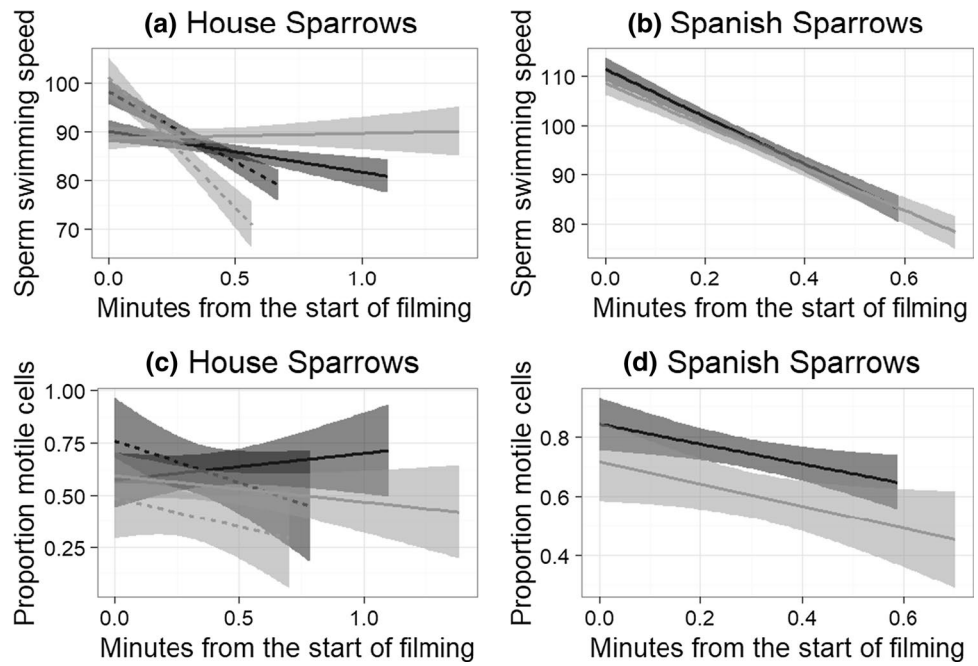

Fig. 3 Overall changes in sperm swimming speed (**a**, **b**) and proportion of motile cells (**c**, **d**) for lightly agitated (**a**, **c**) and highly agitated (**b**, **d**) treatments, combined across experiments 1 (*grey shading*) and 2 (*unshaded*). Data points for individual males are connected by *lines*. *Solid lines* Cells suspended in DMEM, *dotted lines* cells suspended in PBS. To simplify visualization, we averaged values within four time periods: the first and second half of experiment 1 (*grey shading*, periods 1 and 2, with each half approximately 1-min duration), and the first and second half of experiment 2 (*no shading*, periods 3 and 4, with each half approximately 30-s duration). Approximately 30 s elapsed between the end of data collection for experiment 1 and the beginning of experiment 2. For this figure and analysis, each sample is represented in both agitation treatments but only for a single concentration and suspension medium, such that time periods 1 and 2 are the same for lightly and vigorously agitated cells. For abbreviations, see Fig. 1

filmed experiment 2 later in time than experiment 1. We saw no similar "recovery" of cells in velocity: swimming speed in the second half of experiment 1 did not differ significantly from swimming speed in the first half of experiment 2 (Wilcoxon $V = 22$ and 30, $p > 0.6$; Fig. 3).

Consistency of male ranking across recording conditions (analysis 4)

Male identity (which is equivalent to sperm sample identity in our dataset) explained 52.5% of the variation in mean sperm swimming speed and 59.4% of the variation in the proportion of motile cells across treatments. These repeatability scores were significant (VCL: likelihood ratio 31.62, $p < 0.001$, $n = 18$ males, 102 recordings; proportion motile, likelihood ratio 36.96, $p < 0.001$; $n = 19$ males, 111 recordings). In this dataset, samples were measured in all six treatments for 18 (proportion motile) and 13 males (VCL); in five treatments for four males (VCL), in four treatments for one male (VCL) and in three treatments for one male (proportion motile).

The between-male CV for the proportion of motile cells in highly agitated samples was notably lower than that for the other treatments, though 95% confidence intervals overlapped with those for the other tested conditions (Table 1).

Resampling analysis: impact of number of cells analyzed (analysis 5)

Including a larger number of cells per male per treatment increased the precision of estimated values of between-male repeatability in sperm swimming speed, of the difference in swimming speed between PBS and DMEM, and of the mean speed for each male, but there was no apparent bias when low numbers of cells were included (Fig. 4). In the full data set from which resampling was conducted, between-male repeatability in sperm swimming speed was low but significant (with each male represented by a single sample measured in two conditions; likelihood ratio 757, $p < 0.0001$, 14.87% of variance attributable to male identity, $n = 12$ males and 7174 sperm cells). For each number of cells resampled, the mean repeatability across the 100 resamples closely approximated the value from the full data set (Fig. 4a). In resampled datasets, repeatability was significant ($p < 0.05$) for 56/100 tests using five cells per male per treatment, for 96/100 tests using ten cells per male per treatment, and for all tests using more than ten cells per male per treatment. In the full data set, swimming speed was higher in DMEM than in PBS ($t_{11} = 2.791$, $p = 0.02$, mean difference = 4.84 μm/s, based on the mean value for each male across all cells in concentrated treatments, experiment 1). Similarly, cells tended to be faster in DMEM than in PBS, for all resampled sets except for some of the smaller data subsets where cells tended to be faster in PBS (nine five-sperm sets, two ten-sperm sets, and one 20-sperm set, all with non-significant t-tests; Fig. 4b). t-test results were significant and in the expected direction for 17, 24, 38, 44, 52, 57, 73, and 79 of 100 tests (for five, ten, 20, 30, 40, 50, 100, and 125 cells per male per treatement, respectively). The precision of the estimated mean VCL for each male improved with increasing sample size of cells (Fig. 4c).

Discussion

Here, in two congeneric species of passerine birds, we show that sperm performance differs depending on how a sample is prepared for measurement. Specifically, sperm swimming velocity and the proportion of motile cells was typically higher when sperm were concentrated (having been diluted to a lesser degree) and when cells were suspended in a medium that contained nutrients (DMEM), compared to a medium without nutrients (PBS), a result that has also been seen in some mammal species (Farrell et al. 1996; Rijsselaere et al. 2003). This result might suggest an important role for other components of the ejaculate, such as seminal fluid proteins, sugars, or ions, in affecting sperm performance, since these other factors would also have been diluted by the suspension medium. Factors such as pH and the concentration of calcium ions are known to affect in vitro sperm performance in poultry (Holm and Wishart 1998; Wishart and Wilson 1999), highlighting the ability of sperm to respond to their chemical environment. A similar interactive effect was observed in a study on rooster (*Gallus gallus*

Table 1 Coefficient of variation (*CV*) and 95% confidence intervals (*CI*) for sperm performance among House and Spanish Sparrow males, for different measurement conditions

Treatment	CV for swimming speed (CI, n males)	CV for proportion motile (CI, n males)
DMEM, concentrated	5.46 (3.94–7.73, $n = 18$)	33.52 (20.09–73.74, $n = 19$)
DMEM, dilute	7.12 (5.54–9.31, $n = 18$)	39.57 (27.16–59.82, $n = 19$)
PBS, concentrated	7.08 (6.49–10.25, $n = 18$)	38.60 (22.83–71.07, $n = 19$)
PBS, dilute	10.03 (7.03–15.29, $n = 16$)	42.69 (31.73–64.08, $n = 18$)
Low agitation	7.27 (6.25–9.11, $n = 15$)	37.54 (27.26–50.98, $n = 18$)
High agitation	7.76 (6.58–9.42, $n = 17$)	21.28 (13.12–35.91, $n = 18$)

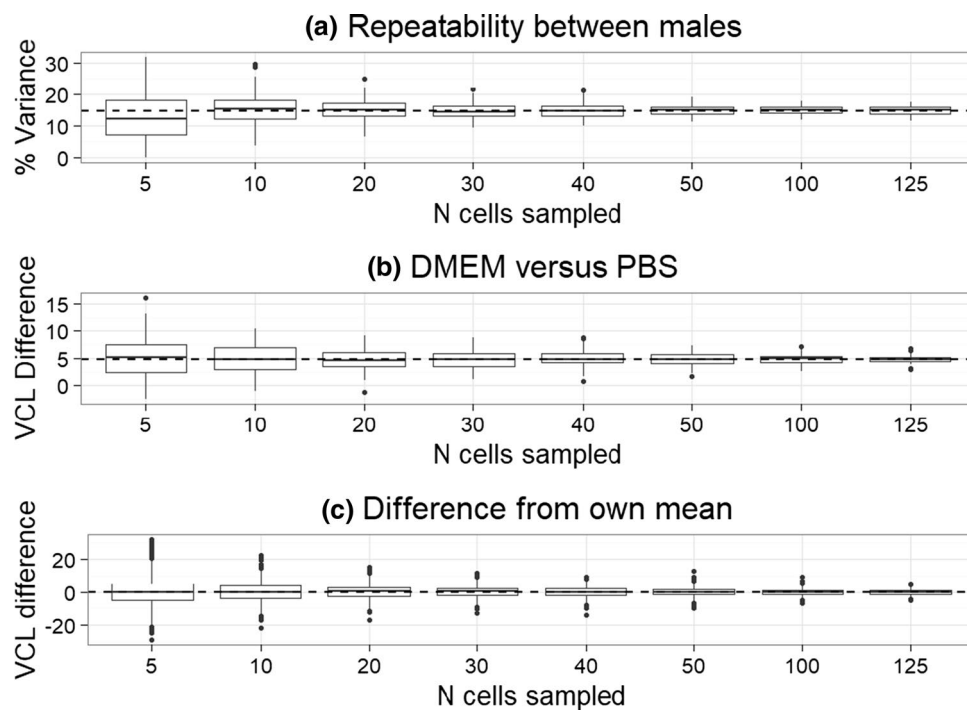*DMEM* Dulbecco's modified Eagle medium, *PBS* phosphate-buffered saline

Fig. 4 Results from resampling cells from 12 males, from concentrated treatments in experiment 1. *Box* and *midline* show the 25th, 50th, and 75th percentiles; *whiskers* extend to the most extreme value within 1.5 interquartile ranges beyond the 25th and 75th percentiles, and outliers beyond that value are *dots*. **a** Repeatability was assessed as the percent of variance attributable to male identity (equivalent to sample identity) in a mixed-effect model that also included a three-way interaction among fixed effects of species, time, and treatment. *Horizontal line* indicates repeatability in the full dataset including all measured cells for those males. **b** Estimated difference between mean swimming speed in PBS and in DMEM (concentrated cells only) in paired testing. *Horizontal line* indicates the paired difference in the full dataset. **c** Difference between the mean swimming speed of a male's sperm in DMEM for a randomly chosen subset of cells compared to its overall mean in DMEM. For abbreviations, see Fig. 1

domesticus) sperm using a different measure of sperm quality; more diluted sperm exhibited lower sperm performance, particularly when dilution was conducted in a saline solution similar to PBS (Parker and McDaniel 2006). While PBS appeared to be a harsh medium in our study, sperm from Bluethroats (*Luscinia svecica*) swim faster in PBS than in a medium derived from conspecific blood plasma (Laskemoen et al. 2008). Which suspension medium and degree of dilution is most representative of conditions sperm face within the female reproductive tract is unclear, making it difficult to determine which sperm measurements are the most biologically relevant.

Surprisingly, we found that mechanical agitation increased the proportion of motile sperm cells. Samples that were subjected to a high degree of agitation before being applied to the microscope slide showed a higher proportion of motile cells, compared to cells from the same sample that had been subjected to minimal agitation. Moreover, the proportion of motile cells was higher at the beginning of experiment 2, when cells were newly introduced onto a microscope slide (which involved agitation as the cells were pipetted onto the slide and filled the chamber via capillary action, as well as additional agitation for cells in the high-agitation

treatment), compared to the same sample at the end of experiment 1, by which time cells had been on the microscope slide for several minutes. This latter result could also have been partly due to the different conditions cells experienced during incubation in the microscope slide chamber, compared to in the microcentrifuge tube where excess sample was stored between experiments (e.g., differences in availability of oxygen). While the biological relevance, if any, of this responsiveness to mechanical agitation is unknown, we can speculate that mechanical agitation of the cells during ejaculation could increase sperm movement as they enter the female reproductive tract; or that mechanical agitation of sperm cells stored in the female's sperm storage organs, for example when a fully shelled egg is laid, facilitates the exit of sperm from storage [as hypothesized by Grigg (1957), though biochemical stimulation also plays a role in release (e.g., Ito et al. 2011; Hiyama et al. 2014)].

The effects of methodological factors on sperm swimming speed and the percentage of motile sperm were not always parallel, for example, with mechanical agitation affecting the proportion of motile cells but not the swimming speed of cells. Similarly, velocity and the proportion of motile cells showed different treatment effects in studies

on how sperm respond to conspecific versus heterospecific female fluids (Cramer et al. 2016b), and in methodological optimization in mammalian studies (Farrell et al. 1996). Together these results strongly suggest that conditions that affect whether an individual sperm cell is motile or not will not necessarily also affect the rapidity with which it moves.

In this study, we standardized the degree to which we diluted a sample, rather than standardizing the final concentration of sperm cells in the video recording. Which approach provides more reliable and/or biologically relevant results remains to be determined. Moreover, different experimental approaches tended to result in differing levels of between-male variation in sperm performance, which may affect statistical power to detect patterns. Nonetheless, sperm performance was moderately repeatable across experimental treatments among males. In this study we examined only a single sample per male, but previous work on the same captive populations showed low but significant between-male repeatability in sperm swimming speed in a dataset that included multiple samples per male, collected weeks to years apart and recorded following different protocols (Cramer et al. 2015). The same study found near-zero between-male repeatability for the proportion of motile cells. Lower repeatability in studies that examine multiple samples per male may be expected if sperm performance changes depending on, for example, social, environmental, or physiological factors that can differ among sampling events.

Resampling analysis showed little evidence for a bias when low cell numbers were used, and the precision of estimates did not improve dramatically when cell numbers increased beyond 20 cells per male. In contrast, recommendations from studies on livestock and humans suggest that at least 200 cells per male should be assessed (ESHRE Andrology Special Interest Group 1998). While we agree that including more cells per male is beneficial, we have found it exceptionally difficult to record large numbers of cells for each individual in some species, particularly under field conditions and when trying to assess speed in multiple experimental treatments (personal observation). For example, when designing our resampling analysis, we initially included only males with 200 cells per treatment (rather than 150, resulting in a data set of nine, rather than 12, males). However, in this reduced data set, the paired comparison between DMEM and PBS became marginally non-significant, presumably due to the lower number of males (data not shown). Given the lack of bias with low sperm numbers, we argue that it may be better for studies on evolutionary biology and behavioral ecology to include more males, despite having low cell numbers for some males, than to exclude males with few cells.

This study highlights the need to be consistent within a study in suspension medium, degree of dilution, and the vigor of pipetting applied to samples. Other studies performed in domestic mammals further suggest that factors such as the type of microscope slide is important (e.g., Hoogewijs et al. 2012; Gloria et al. 2014). However, given the significant and moderate repeatability between males across experimental treatments, results of between-male studies examining how sperm characteristics relate to factors such as timing in the breeding season (Cramer et al. 2013b), male age and ornamentation (Sætre et al. 2018), and paternity success (Cramer et al. 2013a; Edme et al. 2017) should be robust to methodological decisions, as long as researchers are consistent within a study. While in vitro conditions fail to capture many aspects of the complexity of biological reality (e.g., the architecture of the oviduct, the viscosity and biochemical milieu of the female reproductive tract) and thus may not provide an accurate picture of sperm behavior (Lüpold and Pitnick 2018), we currently lack the technology to readily conduct in vivo experiments in vertebrates in which fertilization is internal such as birds. Nonetheless, we suggest that, until such technology becomes available, we can still capture ecologically and evolutionarily relevant information from carefully conducted in vitro studies.

References

Barton K (2016) MuMIn: multi-model inference. R package version 1(15):6

Bates D, Maechler M, Bolker B, Walker S (2014) Lme4: linear mixed-effects models using Eigen and S4. R package version 1:1–5

Bennison C, Hemmings N, Slate J, Birkhead T (2014) Long sperm fertilize more eggs in a bird. Proc R Soc B 282:20141897

Calhim S, Immler S, Birkhead TR (2007) Postcopulatory sexual selection is associated with reduced variation in sperm morphology. PLOS ONE 2:e413

Canty A, Ripley B (2017) Boot: bootstrap R (S-Plus) functions. R package version 1:3–20

Cramer ERA, Laskemoen T, Kleven O et al (2013a) No evidence that sperm morphology predicts paternity success in wild House Wrens. Behav Ecol Sociobiol 67:1845–1853

Cramer ERA, Laskemoen T, Kleven O, Lifjeld JT (2013b) Sperm length variation in House Wrens *Troglodytes aedon*. J Ornithol 154:129–138

Cramer ERA, Laskemoen T, Eroukhmanoff F et al (2014) Testing a post-copulatory pre-zygotic reproductive barrier in a passerine species pair. Behav Ecol Sociobiol 68:1133–1144

Cramer ERA, Laskemoen T, Stensrud E et al (2015) Morphology-function relationships and repeatability in the sperm of *Passer* sparrows. J Morphol 276:370–377

Cramer ERA, Ålund M, McFarlane SE et al (2016a) Females discriminate against heterospecific sperm in a natural hybrid zone. Evolution 70:1844–1855

Cramer ERA, Stensrud E, Marthinsen G et al (2016b) Sperm performance in conspecific and heterospecific female fluid. Ecol Evol 6:1363–1377

Eberhard W (1996) Female control: sexual selection by cryptic female choice. Princeton University Press, Princeton

Edme A, Zoba P, Opatová P et al (2017) Do ornaments, arrival date, and sperm size influence mating and paternity success in the Collared Flycatcher? Behav Ecol Sociobiol 71:3

ESHRE Andrology Special Interest Group (1998) Guidelines on the application of CASA technology in the analysis of spermatozoa. Hum Reprod 13:142–145

Farrell PB, Foote RH, McArdle MM et al (1996) Media and dilution procedures tested to minimize handling effects on human, rabbit, and bull sperm for computer-assisted sperm analysis (CASA). J Androl 17:293–300

Fasel NJ, Helfenstein F, Buff S, Richner H (2015) Electroejaculation and semen buffer evaluation in the microbat *Carollia perspicillata*. Theriogenology 83:904–910

Forstmeier W, Schielzeth H (2011) Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. Behav Ecol Sociobiol 65:47–55

Gloria A, Contri A, Carluccio A et al (2014) The breeding management affects fresh and cryopreserved semen characteristics in *Melopsittacus undulatus*. Anim Reprod Sci 144:48–53

Griffith SC, Owens IPF, Thuman KA (2002) Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol Ecol 11:2195–2212

Grigg GW (1957) The structure of stored sperm in the hen and the nature of the release mechanism. Poult Sci 36:450–451

Hiyama G, Matsuzaki M, Mizushima S et al (2014) Sperm activation by heat shock protein 70 supports the migration of sperm released from sperm storage tubules in Japanese quail (*Coturnix japonica*). Reproduction 147:167–178

Holm L, Wishart GJ (1998) The effect of pH on the motility of spermatozoa from chicken, turkey and quail. Anim Reprod Sci 54:45–54

Hoogewijs MK, Vliegher SPDE, Govaere JL et al (2012) Influence of counting chamber type on CASA outcomes of equine semen analysis 44:542–549

Humann-Guilleminot S, Blévin P, Azou-Barré A et al (2018) Sperm collection in Black-legged Kittiwakes and characterization of sperm velocity and morphology. Avian Res 9:1–12

Ito T, Yoshizaki N, Tokumoto T et al (2011) Progesterone is a sperm-releasing factor from the sperm-storage tubules in birds. Endocrinology 152:3952–3962

Kleven O, Fossøy F, Laskemoen T et al (2009) Comparative evidence for the evolution of sperm swimming speed by sperm competition and female sperm storage duration in passerine birds. Evolution 63:2466–2473

Kuznetsova A, Brockhoff PB, Christensen RHB (2014) lmer test: tests for random and fixed effects for linear mixed effects models (lmer objects of lme4 package). R package version 2:0–11

Laskemoen T, Kleven O, Fossøy F, Lifjeld JT (2007) Intraspecific variation in sperm length in two passerine species, the Bluethroat *Luscinia svecica* and the Willow Warbler *Phylloscopus trochilus*. Ornis Fenn 84:131–139

Laskemoen T, Fossøy F, Rudolfsen G, Lifjeld JT (2008) Age-related variation in primary sexual characters in a passerine with male age-related fertilization success, the bluethroat *Luscinia svecica*. J Avian Biol 39:322–328

Laskemoen T, Kleven O, Fossøy F et al (2010) Sperm quantity and quality effects on fertilization success in a highly promiscuous passerine, the tree swallow *Tachycineta bicolor*. Behav Ecol Sociobiol 64:1473–1483

Lifjeld JT, Laskemoen T, Kleven O et al (2010) Sperm length variation as a predictor of extrapair paternity in passerine birds. PLOS ONE 5:e13456

Lüpold S, Pitnick S (2018) Sperm form and function: what do we know about the role of sexual selection? Reproduction 155:R229–R243

Lüpold S, Calhim S, Immler S, Birkhead TR (2009) Sperm morphology and sperm velocity in passerine birds. Proc R Soc B 276:1175–1181

Nakagawa S, Schielzeth H (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. Biol Rev Camb Philos Soc 85:935–956

Parker G (1970) Sperm competition and its evolutionary consequences in the insects. Biol Rev 45:525–567

Parker HM, McDaniel CD (2006) The immediate impact of semen diluent and rate of dilution on the sperm quality index, ATP utilization, gas exchange, and ionic balance of broiler breeder sperm. Poult Sci 85:106–116

Pinheiro J, Bates D, DebRoy S et al (2013) nlme: linear and nonlinear mixed-effects models. R package version 3:1–113

R Development Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rijsselaere T, Van Soom A, Maes D (2003) Effect of technical settings on canine semen motility parameters measured by the Hamilton-Thorne analyzer 60:1553–1568

Rowe M, Laskemoen T, Johnsen A, Lifjeld JT (2013) Evolution of sperm structure and energetics in passerine birds. Proc R Soc B 280:20122616

Sætre CLC, Johnsen A, Stensrud E, Cramer E (2018) Sperm morphology, sperm motility and paternity success in the bluethroat (*Luscinia svecica*). PLOS ONE 13:e0192644

Schmoll T, Sanciprian R, Kleven O (2016) No evidence for effects of formalin storage duration or solvent medium exposure on avian sperm morphology. J Ornithol 157:647–652

Verhoeven KJF, Simonsen KL, McIntyre LM (2005) Implementing false discovery rate control: increasing your power. Oikos 108:643–647

Wickham H (2009) ggplot2: elegant graphics for data analysis. Springer, New York

Wilson-Leedy JG, Ingermann RL (2007) Development of a novel CASA system based on open source software for characterization of zebrafish sperm motility parameters. Theriogenology 67:661–672

Wishart GJ, Wilson YI (1999) Temperature-dependent inhibition of motility in spermatozoa from different avian species. Anim Reprod Sci 57:229–235

Wolfson A (1952) The cloacal protuberance: a means for determining breeding condition in live male passerines. Bird 23:159–165

Zuur AF, Ieno EN, Walker NJ et al (2009) Mixed effects models and extensions in ecology with R. Springer, New York