Evolution of sperm structure and energetics in passerine birds

Melissah Rowe, Terje Laskemoen, Arild Johnsen and Jan T. Lifjeld

Natural History Museum, University of Oslo, 0318 Oslo, Norway

Spermatozoa exhibit considerable interspecific variability in size and shape. Our understanding of the adaptive significance of this diversity, however, remains limited. Determining how variation in sperm structure translates into variation in sperm performance will contribute to our understanding of the evolutionary diversification of sperm form. Here, using data from passerine birds, we test the hypothesis that longer sperm swim faster because they have more available energy. We found that sperm with longer midpieces have higher levels of intracellular adenosine triphosphate (ATP), but that greater energy reserves do not translate into faster-swimming sperm. Additionally, we found that interspecific variation in sperm ATP concentration is not associated with the level of sperm competition faced by males. Finally, using Bayesian methods, we compared the evolutionary trajectories of sperm morphology and ATP content, and show that both traits have undergone directional evolutionary change. However, in contrast to recent suggestions in other taxa, we show that changes in ATP are unlikely to have preceded changes in morphology in passerine sperm. These results suggest that variable selective pressures are likely to have driven the evolution of sperm traits in different taxa, and highlight fundamental biological differences between taxa with internal and external fertilization, as well as those with and without sperm storage.

1. Introduction

Post-copulatory sexual selection is thought to be an important driver of evolutionary change in sperm traits. In particular, sperm competition is thought to favour the evolution of longer sperm because longer sperm are presumed to be able to swim faster, and thus be capable of gaining access to sperm-storage organs or reaching and fertilizing ova before shorter sperm from rival males [1,2]. This idea is generally well supported by empirical studies showing that both sperm size and velocity are typically greater in species experiencing higher levels of sperm competition [3–7], and that sperm swimming speed is a major determinant of fertilization success in a range of taxa [2]. Moreover, comparative studies show that sperm velocity is influenced by sperm size: longer sperm swim at higher speeds [5,8,9]. Though seemingly more controversial at the intraspecific level [10], recent studies also reveal a correlation between sperm length and velocity within species [11–13]. Thus, there is mounting evidence for a functional link between sperm structure and performance. Our understanding of the mechanisms underlying this link, however, remains limited.

Theoretical models predict three ways in which sperm size might translate into faster-swimming sperm. First, larger midpieces are suggested to have a higher mitochondrial loading (i.e. greater number/volume of mitochondria), and thus have an increased capacity for producing energy to power the flagellum [14]. Next, increased flagellum length is suggested to increase sperm propulsive forces and generate more thrust [15]. Finally, a longer flagellum relative to head length is suggested to translate into faster swimming speeds because a relatively longer flagellum is capable of overcoming drag generated by the head [10]. Importantly, because a longer flagellum requires more energy to generate power, increases in flagellum length should be accompanied by increases in midpiece length [14]. Such positive phenotypic correlations among the size of the midpiece and flagellum have been reported in both
mammals [16] and birds [9,17]. Thus, sperm structure and swimming speed appear to be linked via energy availability; indeed, energy availability has been proposed to explain the observed association between sperm size and speed in a variety of taxa, including birds [9].

To date, support for the idea that longer sperm swim faster because they have more energy is limited. However, intraspecific studies in fish show that sperm swimming speeds are correlated with initial sperm adenosine triphosphate (ATP) concentrations [18,19]. Furthermore, Vladić et al. [20] demonstrated a positive correlation between sperm midpiece size and ATP concentration. Taken together, these results suggest that a bigger midpiece affords a higher ATP content, which translates into faster-swimming sperm. Similar relationships between sperm ATP content and sperm motion have also been observed in rats (Rattus norvegicus) [21] and chickens (Gallus gallus) [22], suggesting that energy reserves may provide the link between sperm structure and performance in a range of taxa. In contrast, evidence from comparative studies remains indirect, and studies examining the relationships between sperm size, speed and energetics are lacking.

Under the assumption that sperm ATP content does influence sperm swimming speed, and given the fertilization advantage of fast sperm under competitive conditions [2], we might also expect a relationship between sperm competition and sperm ATP levels. Indeed, intraspecific comparisons in bluegill (Lepomis macrochirus) and Atlantic salmon (Salmo salar), two species that display alternative male reproductive tactics, revealed that sperm ATP concentration was related to sperm competition risk: males experiencing a greater risk of sperm competition had higher levels of intracellular ATP [18,23,24]. Additionally, in a comparison between humans (Homo sapiens) and chimpanzees (Pan troglodytes), Anderson et al. [25] showed that sperm competition was associated with sperm bioenergetics: chimpanzees showed a significantly higher mitochondrial membrane potential both before and after capacitation. However, robust comparative studies are necessary to test whether selection imposed through sperm competition shapes variation in the intracellular ATP concentration of sperm.

Finally, in a study of African cichlids, Fitzpatrick et al. [5] showed that evolutionary changes in sperm size were likely to be contingent on earlier increases in motility. Moreover, the authors suggested that, because evolutionary increases in ATP (e.g. increases in mitochondrial efficiency or mitochondrial loading) could increase sperm velocity, sperm energetics were probably the initial target of selection owing to sperm competition [5]. Importantly, this implies that evolutionary changes in the intracellular ATP content of sperm occurred earlier in evolutionary time relative to changes in total sperm length, an idea that is currently untested.

Sperm ATP is generated via two metabolic pathways: oxidative phosphorylation (OXPHOS) and glycolysis. Across taxa, sperm can show considerable metabolic flexibility [26]. For example, in eutherian mammals, sperm can be glycolysis dominant, OXPHOS dominant, intermediate between the two pathways or flexible [27]. OXPHOS appears to be the predominant pathway for energy metabolism in birds, whereas glycolysis appears limited [22,28], although a degree of interspecific variation in glycolic activity has been observed [29,30]. Moreover, sperm ATP content is correlated with both sperm mobility (i.e. the net movement of sperm against resistance at body temperature) and the fertilizing ability of sperm in galliformes (i.e. chickens; turkeys, Meleagris gallopavo [22,30]). Consequently, mitochondrial respiration appears crucial for avian sperm motility.

In passerines, the midpiece houses a single fused mitochondrion wound helically around the sperm flagellum [31], (see [32] for an exception), suggesting that increases in OXPHOS could be achieved by increases in midpiece size. Importantly, sperm appear to be immotile prior to ejaculation in birds [33], while, after ejaculation, sperm ascension of the vagina and entrance to sperm-storage tubules (SSTs) is powered by oxidation of endogenous fatty acids [34]. Under the assumption that intracellular ATP provides the immediate source of energy for sperm motility [35], high intracellular ATP concentrations are predicted to translate into faster swimming speeds.

Here, we examine the relationships between sperm size, speed and intracellular ATP concentration using data from 23 passerine species. This approach allowed us to test the hypothesis that longer sperm swim faster because they have more available energy by asking the following questions. Do sperm with longer midpieces have more intracellular ATP? Do sperm with higher concentrations of intracellular ATP swim faster? Furthermore, we tested for an association between sperm competition and sperm ATP levels. In addition to variation in size, variation in mitochondrial structure may influence a cell’s capacity for ATP production [36]. Consequently, evolutionary changes in ATP concentration or mitochondrial efficiency (i.e. the amount of stored ATP per unit of mitochondrion) might be expected to lead to faster-swimming sperm. Thus, we asked whether sperm competition was related to either sperm ATP concentration (i.e. nmol ATP/10^6 sperm) or mitochondrial efficiency of sperm (i.e. ATP concentration/midpiece length). Finally, we modelled evolutionary change in sperm energy stores and morphometry, in order to examine whether changes in sperm ATP preceded changes in overall sperm size, and to understand how selection may have acted over evolutionary time to shape contemporary variation in sperm structure and performance.

2. Material and methods

(a) General methods and sperm collection

We collected samples from wild populations of 23 passerine species during the breeding season (April–June) 2011 (see electronic supplementary material, table S1 for species and sample sizes). Birds were trapped using mist nets and song playback at sites in southern Norway, and sperm samples were collected via cloacal massage [37]. Semen was collected in a 10 µl capillary tube and immediately used for analysis of intracellular ATP concentration. Any samples contaminated with faeces (determined by visual inspection) were discarded.

(b) Determination of ATP concentration

ATP concentration was determined using a luciferase-based ATP bioluminescent assay kit (ATP Bioluminescence Assay Kit HS II, Roche). Immediately after collection, fresh semen was diluted in 112 µl of dilution buffer. Next, a 100 µl aliquot of diluted semen was transferred to a new microcentrifuge tube and mixed with 100 µl of cell lysis reagent. The mixture was then vortexed and incubated at room temperature for 5 min. The resulting cell lysate was centrifuged at 12,000g for 2 min, and the supernatant recovered and immediately snap-frozen in liquid N_2.
The frozen samples were stored at $-80^\circ$C until analysis. Finally, a separate $50 \mu$l aliquot of the original sperm dilution was fixed in 20 $\mu$l of 5 per cent formalin solution and the density of sperm in each sample determined using a Makler counting chamber.

To measure ATP concentration, we first standardized the density of sperm contributing to ATP content of each sample by adding a variable amount of dilution buffer so that the final sperm density was $1 \times 10^{8}$ per ml. Next, 50 $\mu$l of luciferase reagent was added to 50 $\mu$l of sample (via auto-injection), and bioluminescence measured over 10 s following a 1 s delay. Bioluminescence was measured using a Varioskan Flash (Thermo Fisher Scientific). All samples were run in triplicate and the ATP content of each sample was calculated by comparison with a standard curve. Bioluminescence values across triplicate samples were significantly repeatable within individuals ($R = 0.998$, $p < 0.0001$), so we calculated an average value for each sample. Results were then converted to nmol ATP per 108 sperm. For a subset of males ($n = 18$), we performed repeat assays and found that ATP measurements between different microplates were also highly repeatable ($R = 0.997$, $p < 0.0001$). Mean ATP concentrations were ln-transformed prior to analysis to normalize the distribution.

(c) Sperm morphology and velocity analysis

We obtained measurements of sperm morphology and velocity obtained from fresh sperm samples collected from wild populations in southern Norway during 2006–2010; these individuals were sampled from the same location, or in the same region as, those sampled in 2011. As before, birds were trapped using mist nets and song playback, and sperm was collected by cloacal massage.

For morphology, sperm samples were fixed in 5 per cent formalin solution. Sperm morphometric data were obtained from digital images captured at magnifications of 160 $\times$ or 320 $\times$ (DM6000 B Leica digital microscope) using digital image analysis (Leica APPLICATION SUITE v. 2.6.0 R1). For each individual, 10 morphologically normal sperm were analysed to obtain measurements ($\pm 1.0 \mu$m) of the following sperm traits: (i) head length, (ii) midpiece length, (iii) flagellum length and (iv) total length. For each sperm trait, we used the means within individuals ($n = 5–90$; electronic supplementary material, table S1) to calculate the mean for each species. Mean sperm lengths were ln-transformed prior to analysis to normalize the distribution.

We measured sperm performance immediately after ejaculate collection. Fresh sperm samples were diluted in pre-heated (40°C) Dulbecco’s Eagle Medium (DMEM, Invitrogen). Next, 3–5 $\mu$l of the diluted semen was placed in a pre-heated microscopy counting chamber (depth 20 $\mu$m; Leja, Nieuw-Vennep, the Netherlands) mounted on a MiniTherm slide warmer (Hamilton Thorne Inc) maintained at a constant temperature of 40°C. Sperm movement was then recorded using a phase contrast microscope (CX41, Olympus, Japan) connected to a digital video camera (HDR-HC1C, PAL, Sony, Japan).

The resulting videos were analysed using computer-assisted sperm analysis (HTM-CEROS sperm tracker, CEROS v. 12, Hamilton Thorne Research) following Kleven et al. [4]. Briefly, we recorded straight-line velocity (VSL; i.e. average velocity on a straight line between the start and end points of the sperm track), curvilinear velocity (VCL; i.e. velocity over the actual sperm track) and average path velocity (VAP; i.e. average velocity over a smoothed sperm track) for each male. Tracks were filtered to exclude the potential effects of drift (VAP $< 10 \mu$m s$^{-1}$, VSL $< 5 \mu$m s$^{-1}$) and data obtained from non-sperm particles. Additionally, only individuals with tracks from 20 or more motile sperm cells were included in the analysis. To quantify sperm velocity, we calculated the mean values of VCL, VAP and VSL for each male ($n = 5–84$; electronic supplementary material, table S1) to calculate mean values for each species.

We chose to use VCL (hereafter referred to as sperm velocity) instead of VAP or VSL as a measure of sperm swimming speed because under in vitro conditions (i.e. in the absence of an attractant or cell guidance mechanism) sperm movement may not be linear [38]. Moreover, nonlinear sperm motion is commonly observed in in vitro assays (M.R., personal observation). Thus, the actual point-to-point track (i.e. VCL) provides a more realistic measure of sperm velocity. Nonetheless, these parameters are strongly intercorrelated (all $R > 0.84$, $p < 0.0001$), and analyses using VAP and VSL returned qualitatively similar results (data not shown). Previous studies show that sperm velocity is highly repeatable between males within species, and that mean values for each species are representative [4].

(d) Index of sperm competition

We used two estimates to infer the level of sperm competition risk: relative testis size and sperm length variation. Relative testis size, a commonly used proxy for sperm competition in avian studies [6,17], was estimated by including both (ln-transformed) testes mass and body mass as independent variables in statistical models. Data on testes mass and body mass were obtained from museum sources, from published literature [39] or from males collected (under licence) during the breeding season. Data on sperm length variation were calculated from sperm as the between-male coefficient of variation ($CV_{b,m} = \text{s.d./mean} \times 100$), adjusted for small sample size according to the formula adjusted $CV_{b,m} = (1 + 1/4n) \times CV_{b,m}$ [40]. Adjusted $CV_{b,m}$ (hereafter referred to as simply spermCV) provides a proxy measure of extra-pair paternity, and thus levels of sperm competition, in passerine birds [41].

(e) Phylogeny

We constructed a phylogenetic topology of the 23 species based on recent literature. Because branch length information cannot be inferred from published trees utilizing varying methods, we estimated branch length for this topology using PAUP v. 4.0b10 [42] based on sequence data from three mitochondrial loci: cytochrome c oxidase subunit 1 (COI), cytochrome b (cytb) and NADH dehydrogenase II (ND2). State frequencies were estimated via maximum-likelihood with a general time-reversible (GTR) nucleotide substitution model, and the heterogeneity of nucleotide substitution rates among sites was approximated by a discrete gamma distribution ($\Gamma_\alpha$) and an assumption of invariant sites (I; see the electronic supplementary material for further details).

(f) Statistical analysis

We performed analyses on all 23 species, with the exception of analysis including sperm velocity, where data for the European greenfinch (Carduelis chloris) were lacking and analysis was performed on 22 species. Non-normal data distributions were ln-transformed to meet the parametric requirements of the statistical models. All repeatabilities (ANOVA-based) were calculated using the R code written by Nakagawa & Schielzeth [43].

To determine whether there were associations between sperm size, velocity and intracellular ATP concentration, and to test whether sperm ATP levels (ATP concentration and mitochondrial efficiency) were associated with sperm competition, we performed generalized least-squares (GLS) regressions within a phylogenetic framework. Importantly, this approach accounts for the non-independence of data owing to shared ancestry of species [44,45]. In addition, the GLS method allows the estimation of the scaling parameter $\lambda$, which assesses the degree of phylogenetic association in traits: $\lambda = 0$ indicates no phylogenetic association, $\lambda = 1$ indicates complete phylogenetic association. We used likelihood-ratio tests to compare the model where $\lambda$ assumes its maximum-likelihood
of the likelihood-ratio tests (first position: against energetics preceded changes in morphometry using BAYESCONTINUOUS in BAYESTRAITS v. 2.0 [44,48,49]. This software implements a GLS approach and allows for the analysis of continuously varying traits. We used this package to investigate both the tempo of trait evolution and to assess whether traits show any dominant direction of evolutionary change.

We first evaluated directionality in trait evolution by determining whether a random walk model (model A) described the data significantly better than a directional random walk model (model B). Model A is a standard constant-variance random walk model (i.e. Brownian motion) incorporating a single parameter: the instantaneous variance of evolution. Model B accounts for the same variance of evolution as model A, but incorporates an additional parameter (δ) that measures the regression of the trait of interest against total path length from the root to the tips of the tree. This parameter describes the tendency for directional evolutionary change. These models incorporate ancestral state reconstruction using Markov chain Monte Carlo methods; thus, we report both the estimated ancestral state (α) and the slope (β) of the regression of path length on each trait value for both sperm size (i.e. midpiece length, total sperm length) and sperm ATP concentration.

After evaluating directional trends, we used the best-fitting model to assess the tempo (δ) of evolutionary change in sperm traits. δ can detect whether the rate of trait evolution has accelerated or slowed over time, and determines whether character change is concentrated at the root or towards the tips of a phylogeny. Values of δ < 1 can indicate temporally early evolution of a trait (i.e. change concentrated at the root of the phylogeny), whereas values of δ > 1 indicate temporally late evolution of a trait (i.e. change concentrated towards the tips of the phylogeny). Finally, values of δ = 1 are interpreted as gradual change over the phylogeny [48]. For each sperm trait, δ was estimated (maximum-likelihood value) using the best-fitting model (i.e. model A versus model B above), and then this model was compared against a null model where δ = 1.

All models were run using the tree detailed above as input, and for each run rate parameter values were explored to find acceptance rates when running Markov chains of between 20 and 40 per cent [49,50]. The Markov chain was run for 1 050 000 generations, sampled every 1000th generation after a burn-in of 50 000 generations. All priors were set as hyperpriors with a uniform distribution. Each run was repeated three times in order to ensure the stability of the harmonic mean estimator. Finally, we compared and selected models in BAYESTRAITS using Bayes factors (BF [51]) based on the harmonic mean estimator. Values of BF > 10 indicate strong support for the model with higher BF value against the models with values of BF < 1. Analyses were performed using R v. 2.14.1 [46], the R package ‘ape’ [47] and code written by R. Freckleton (University of Sheffield).

Next, we examined the evolution of sperm size and intracellular ATP concentrations to determine whether changes in energetics preceded changes in morphometry. For each sperm trait, ATP concentration was also positively related to flagellum length, but in this instance the relationship was not statistically significant (table 1). In passerines, the sperm midpiece and
flagellum are morphologically associated (i.e. the midpiece is wound helically along the flagellum for much of its length), and our finding that flagellum length and ATP concentration are positively associated is likely to reflect this underlying association. We therefore examined the relationship between ATP concentration and the length of the section of the flagellum that is not associated with the midpiece. In this instance, we found that ATP concentration was negatively associated with the length of the midpiece-free flagellum (table 1).

We found no relationship between sperm velocity and the ATP concentration of sperm (table 1, figure 1b), nor was sperm velocity related to mitochondrial efficiency of sperm (table 1). However, in our dataset, sperm velocity was not related to sperm total length \( t = 0.25, p = 0.81 \), midpiece length \( t = -0.07, p = 0.95 \) or the ratio of head to flagellum length \( t = 0.20, p = 0.85 \). Given that a relationship between sperm size and speed is a key assumption of the hypothesis that longer sperm swim faster because they have more available energy, we generated a subset of species \( n = 15 \), electronic supplementary material, table S3), we found that ATP concentration was not associated with relative testis size (table 2). Moreover, neither ATP concentration nor mitochondrial efficiency were related to sperm competition when estimated as spermCV (table 2). Finally, using a subset of species for which extra-pair paternity data were available \( n = 15 \), electronic supplementary material, table S3), we found that ATP concentration was not related to the rates of extra-pair paternity in a population \( t = 0.85, p = 0.41, \lambda = 0.37 \).

In the analysis of trait evolution comparing model A (Brownian random walk) and model B (directional random walk), we found support for directional change in all traits examined: sperm ATP concentration, midpiece length and total sperm length (table 3). Therefore, the model assuming a directional random walk was used for all subsequent analyses. For analysis of evolutionary tempo \( \delta \), the model in which \( \delta \) was allowed to take its maximum-likelihood value \( \delta = 2.55 \) was accepted, suggesting that sperm ATP concentration has undergone increased trait evolution towards the tips of the phylogeny (i.e. temporally late evolution). In contrast, for both midpiece length and total sperm length the model in which \( \delta \) was allowed to take its maximum-likelihood value did not differ from the null model (table 3), suggesting that evolutionary change in both sperm size traits has occurred gradually over time.

### Table 2. Regression analysis controlling for phylogeny of sperm ATP concentration [ATP] and mitochondrial efficiency in relation to sperm competition estimated by (i) relative testis mass and (ii) spermCV. Superscripts following the \( \lambda \) estimates indicate significance levels of the likelihood-ratio tests (first position: against \( \lambda = 1 \); second position: against \( \lambda = 0 \); n.s.: not significant, \( * p < 0.05 \)).

<table>
<thead>
<tr>
<th>sperm trait</th>
<th>predictor</th>
<th>slope</th>
<th>t</th>
<th>p-value</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ATP]</td>
<td>testis mass</td>
<td>0.34</td>
<td>0.38</td>
<td>0.71</td>
<td>0.99(^{a,s}), n.s.</td>
</tr>
<tr>
<td></td>
<td>body mass</td>
<td>-1.56</td>
<td>-1.59</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>[ATP]</td>
<td>spermCV</td>
<td>0.54</td>
<td>0.91</td>
<td>0.38</td>
<td>0.99(^{a,s}), n.s.</td>
</tr>
<tr>
<td>mitochondrial efficiency</td>
<td>testis mass</td>
<td>0.16</td>
<td>0.21</td>
<td>0.84</td>
<td>0.99(^{a,s}), n.s.</td>
</tr>
<tr>
<td></td>
<td>body mass</td>
<td>-1.28</td>
<td>-1.54</td>
<td>0.15</td>
<td></td>
</tr>
</tbody>
</table>

### Table 3. Evolutionary analysis of sperm intracellular ATP concentration and sperm size using BAYESCONTINUOUS (BAYESTRAITS v. 2.0). Maximum likelihood estimate (MLE) of evolutionary tempo \( \delta \), ancestral state \( \alpha \) and directional trend \( \beta \).

<table>
<thead>
<tr>
<th>characters</th>
<th>( \delta )</th>
<th>BF(^a) against ( \delta = 1 )</th>
<th>interpretation</th>
<th>( \alpha )</th>
<th>( \beta )</th>
<th>BF(^b)</th>
<th>[R] versus [D]</th>
<th>interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sperm [ATP]</td>
<td>2.55</td>
<td>8.02**</td>
<td>temporally late change</td>
<td>0.026</td>
<td>1.05</td>
<td>2.82**</td>
<td>[D] increasing</td>
<td></td>
</tr>
<tr>
<td>midpiece length</td>
<td>1.54</td>
<td>1.22(^{a,s})</td>
<td>gradual change</td>
<td>10.64</td>
<td>87.09</td>
<td>7.21**</td>
<td>[D] increasing</td>
<td></td>
</tr>
<tr>
<td>total sperm length</td>
<td>1.51</td>
<td>0.75(^{a,s})</td>
<td>gradual change</td>
<td>11.57</td>
<td>92.76</td>
<td>3.30*</td>
<td>[D] increasing</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Interpretation of Bayes factors [51]: BF < 2: no evidence against the null hypothesis (i.e. n.s.); BF = 2–6: positive evidence against the null hypothesis (*); BF 6–10: strong evidence against the null hypothesis (**).
4. Discussion

(a) Relationships between sperm size, speed and energetics

Comparative studies suggest that longer sperm swim faster in a range of taxa [5,8], including passerine birds [9] (but see also [4]). Moreover, Lüpold et al. [9] showed that, relative to other cell components, midpiece length is especially variable and is the strongest predictor of sperm velocity. Since the mitochondrion of the midpiece supplies the sperm with energy [27], it is often assumed that longer sperm swim faster because they have more energy. Although results from in vitro studies must be interpreted with caution, our results suggest that while sperm with longer midpieces have higher levels of intracellular ATP, greater energy reserves do not translate into faster-swimming sperm. Although somewhat counter to predictions from theory [14], these results may be understood in light of the reproductive biology of birds (i.e. internal fertilization, sperm storage), and may be explained by biological processes, such as differential energy utilization and synthesis, or perhaps by issues of taxonomic sampling.

Post-copulatory sexual selection promotes the evolution of sperm traits that enhance fertilization success. However, selection is unlikely to act on any one sperm trait in isolation. Instead, several aspects of sperm performance (e.g. velocity, longevity, viability) are likely to be important predictors of sperm competitive ability, and thus be under selection, especially in taxa with internal fertilization and sperm storage. In birds, sperm enter SSTs, where they may remain for extended periods [53]. The amount of time necessary for sperm to reach SSTs is unknown, but this process is unlikely to be immediate. Furthermore, models of avian sperm storage suggest that sperm remain motile while in storage in order to maintain position in the tubules against a fluid current generated by SST epithelial cells [54]. Under such conditions, the ability of sperm to remain motile for extended periods may be important for male fertilization success, as increases in sperm lifespan would provide a male with greater representation in the pool of fertilizing sperm for a longer period of time [2,55].

Pizzari & Parker [2] suggested that one way in which sperm longevity may be determined is the rate at which metabolically active sperm use their energy reserves. Importantly, the strength of selection for longevity may vary across species. For example, under conditions of temporal variation in paternity success linked to variation in sperm motile performance [53], interspecific variation in sperm storage duration (which may in turn be linked to variation in clutch size or female mating rate) is likely to shape the evolution of sperm longevity and energy use. Thus, sperm from different species may differ in the rate at which they use available ATP. The notion that slower-moving sperm use their energy reserves at a slower rate is supported by studies showing a negative association between sperm velocity and longevity [13,56]. Thus, the results of the current study may be explained by differential energy utilization, as it would imply that initial swimming velocity is not determined by the total pool of available intracellular ATP but by the rate at which sperm use such energy resources. Such a pattern has been shown for human spermatozoa, where ATP content, while not associated with initial swimming speeds, is associated with a decline in sperm performance over time [57]. In chickens, sperm motility is dependent upon calcium (Ca$^{2+}$) cycling through the mitochondria, which is in turn dependent upon extracellular sodium (Na$^{+}$) reserves, and fowl sperm conserve ATP by preventing active transport of Ca$^{2+}$ and Na$^{+}$ in sperm cells [58]. A similar mechanism may underlie the results of this study; interspecific variation in Ca$^{2+}$ cycling may determine the rate at which passerine sperm cells use energy, and thus determine variation in sperm longevity across species. Finally, in addition to its role as an energy source, ATP is essential for several cellular and biochemical processes (e.g. protein phosphorylation in cell signalling, cofactor regulation of protein function [59]). Consequently, the total pool of ATP resources may reflect energy available for a range of functions, and not simply energy availability for sperm motion.

An alternative, but not necessarily mutually exclusive explanation of our results is that it arises as an artefact of taxonomic sampling. Although support for the notion that longer sperm swim faster has grown in recent years, the idea remains somewhat contentious [10], and empirical studies have yielded conflicting results. In birds, for example, Lüpold et al. [9] showed that sperm velocity was positively associated with sperm length and the ratios between sperm components, while Kleven et al. [4] found no relationship between sperm velocity and any metric of sperm size. Such inconsistencies may be explained by differences in taxonomic coverage and variation in selection pressures faced by different taxonomic groups (cf. [17]), or differences in experimental methodology (e.g. temperature of medium). However, we found that ATP concentration was not associated with sperm swimming speed across several data subsets (see the electronic supplementary material), suggesting that the lack of association between ATP concentrations and sperm velocity may be somewhat robust to issues of taxonomic sampling, and that higher concentrations of intracellular ATP do not necessarily translate into higher swimming speeds in passerine sperm.

Whether differences in energy utilization or issues of taxonomic sampling explain the results of the current study, our findings highlight the complexity of the relationship between sperm form and function. More specifically, while a link between sperm structure and performance is both logical and highly probable, the combined results of this and other studies suggest that such a link is unlikely to be as simple as ‘long sperm swim fast’. Instead, numerous factors are likely to influence the relationship between sperm structure and swimming speed. For example, selection for other aspects of performance (e.g. longevity, endurance [5]) and for structural properties other than length (e.g. shape [60]) are likely to influence sperm velocity. Moreover, external factors (e.g. female reproductive tract environment [61]) are likely to further influence actual sperm swimming velocity under in vivo conditions. Thus, selection is likely to select for an optimal morphometry (size, shape) that maximizes multiple aspects of sperm performance that influence fertilization success of males.

The results of our study also offer some insight into the relative importance of the metabolic pathways used by passerine sperm to generate ATP. In sperm cells generally, the question of whether the main source of energy is OXPHOS or glycolysis remains a topic of much discussion [62]. OXPHOS occurs in the mitochondria, whereas glycolysis occurs primarily in enzymes bound to the sperm
flagellum [62]. The debate over which metabolic pathway is the primary source of energy has been most extensively studied in mammals, for which it is argued that, despite the greater efficiency of ATP generation via OXPHOS, glycolysis is necessary to support sperm motility because ATP may not be able to sufficiently diffuse from the midpiece to the distal segments of the flagellum [62,63]. Given that, in passerines, the midpiece is intricately associated with the flagellum for much of its length, the problem of ATP diffusion along the flagellum may be less relevant to avian taxa. To date, studies of the biochemical pathways of energy synthesis in birds are limited to galliform species, for which it is suggested that OXPHOS is the primary pathway of ATP synthesis [22,28]. In the current study, we found that intracellular ATP concentration was positively associated with midpiece size, but negatively associated with the length of the portion of the flagellum not associated with the midpiece (i.e. midpiece-free flagellum length). Taken together, these results indicate that OXPHOS may also be the predominant pathway of energy metabolism in passerine birds.

(b) Sperm competition and sperm ATP concentration
Sperm motile performance is a major determinant of fertilization success under conditions of sperm competition [64,65]. Under the assumption that sperm energetics influence sperm velocity, we might expect a relationship between sperm ATP levels and sperm competition. In our dataset, however, we found no association between sperm competition and either ATP concentration or the mitochondrial efficiency of sperm. Low interspecific variation in the risk of sperm competition faced by males is one potential explanation for these results. However, species included in this study represent a broad range of sperm competition levels as inferred from both extra-pair (EP) paternity levels (EP young, 9–37%; EP broods, 23–72%; electronic supplementary material, table S3) and relative testis size (percentage body mass: range 0.7–4.7%; electronic supplementary material, table S3). Alternatively, sperm traits may evolve as an adaptation to the female reproductive tract environment [66,67], especially in species with sperm storage. In pheasants, Immler et al. [68] showed that sperm velocity was negatively related to the duration of sperm storage (see also [4]), but independent of the risk of sperm competition. Given that sperm energetics probably influence sperm longevity, the effects of selection imposed by the female reproductive environment may confound a direct link between sperm competition and sperm ATP concentrations.

(c) Evolutionary trajectories of sperm size and ATP concentration
Finally, we examined the evolutionary trajectories of sperm size and intracellular ATP concentrations, and modelled the tempo of evolutionary change in these traits in order to explore the notion that changes in sperm energetics preceded changes in sperm morphology. Bayesian modelling revealed that both sperm size and intracellular ATP concentrations show evolutionary trajectories that differ significantly from Brownian motion. These findings suggest that sperm characters have been selected for increasing trait values: sperm have become longer and ATP concentrations have increased over evolutionary time. In addition, our analyses revealed that sperm size and ATP content show distinct patterns of trait evolution. Specifically, evolutionary change in sperm size appears to have occurred gradually over time, whereas evolutionary change in ATP content appears to have accelerated as time has progressed.

The signature of accelerating evolution indicates temporally late changes in sperm ATP levels. Thus, while changes in sperm size have occurred gradually over time, changes in sperm energetics, at least in terms of total ATP concentrations, appear to be concentrated towards the tips of the phylogeny, and represent species-specific responses. Importantly, these contrasting patterns of evolutionary change in sperm size and energetics fail to support the idea that changes in sperm energetics preceded changes in total sperm length. Moreover, these findings suggest that changes in morphology may actually have occurred prior to changes in intracellular ATP concentrations. These patterns of evolutionary change suggest that early increases in midpiece length were not associated with the majority of change in ATP content, and that subsequent increases in ATP were relatively independent of changes in midpiece length. Such changes in the energetic state of a sperm cell independent of overall mitochondrion size might be achieved via changes to mitochondrial structure. For example, both an increase in the surface area of the inner membrane (achieved via an increase in cristae density) and an increase in the density of OXPHOS complex molecules on the inner membrane would increase the capacity for ATP synthesis [36]. Additionally, increases in mitochondrial membrane potential or the volume of the inner matrix, perhaps achieved by increases in mitochondrial volume resulting from changes in mitochondrion diameter, could lead to increases in ATP production [25,36].

The evolutionary patterns inferred from our analyses contrast with recent suggestions that changes in sperm energetics may represent an early evolutionary transition in the evolution of sperm morphology in fishes [5]. The reason for this discrepancy is unclear, but there are tremendous biological differences between the taxa studied (i.e. cichlid fish versus passerine birds). Most notably, cichlids are externally fertilizing organisms, while passerine birds are internal fertilizers with the capacity for sperm storage. Consequently, the fertilizing environment experienced by sperm from these two groups is widely divergent (i.e. freshwater versus female reproductive tract). This suggests that selection pressures are likely to vary enormously between these two groups, and that selection is likely to lead to differences in sperm phenotype. For example, the duration of sperm motility tends to be very short in externally fertilizing fishes (e.g. less than 1 min [69]), whereas sperm can remain motile for extended periods in passerine birds (e.g. more than 24 h; M.R., personal observation; see also [70]). Thus, it is not surprising that the evolutionary trajectories of sperm phenotype differ dramatically between these taxa.

5. Conclusions
Our results suggest that sperm midpiece size reflects the energy status of a sperm cell in passerine birds. However, our results also highlight the complexity of the relationship between sperm structure and performance. Under conditions of sperm competition, selection is likely to select for an optimal sperm morphometry that maximizes multiple aspects of
sperm performance; thus, while sperm form and function are linked, the relationship is not as simple as ‘longer sperm swim faster’ because they have more energy. Our results do not suggest that ATP is an unimportant energy source; rather they suggest that interspecific variation in ATP levels may not be the primary determinant of variation in sperm velocity among passerine birds, and that factors such as selection for multiple aspects of sperm performance and the female environment need to be considered when attempting to understand the relationship between sperm form and function. Finally, differences between our study and previous studies examining the evolutionary trajectories of sperm phenotype suggest that variable selective pressures are likely to have driven the evolution of sperm traits in different taxa, and highlight fundamental biological differences between taxa.

We are grateful to Johan Nylander and Andrew Meade for phylogenetic and statistical advice. We also thank Gary Burness, Steve Dorus, Jostein Gohl, Montse Gomendio, Mar González-Barroso, Lars Erik Johannessen, Oddmund Kleven, Eduardo Rial, Eduardo Roldan, Maxi Tournmente and Elaine Wilkins for useful discussions, and Tommaso Pizzari and two anonymous reviewers for valuable comments on the manuscript. This project was funded by the Research Council of Norway. Data are deposited in Dryad. All procedures performed in this study conform to the legal requirement for animal research in Norway, and males were caught under licences issued by the Norwegian Directorate for Nature Management.

References
