Comparative sperm proteomics in selected passerine birds reflects sperm morphology and mitochondrial metabolism

Authors: Otčenášková, Tereza, Stopková, Romana, Zemanová, Aneta, Míčková, Kristýna, Harant, Karel, et al.

Source: Journal of Vertebrate Biology, 72(23045)

Published By: Institute of Vertebrate Biology, Czech Academy of Sciences

URL: https://doi.org/10.25225/jvb.23045

Journal of Vertebrate Biology

Open Access

J. Vertebr. Biol. 2023, 72: 23045 DOI: 10.25225/jvb.23045

RESEARCH PAPER

Comparative sperm proteomics in selected passerine birds reflects sperm morphology and mitochondrial metabolism

Tereza OTČENÁŠKOVÁ[1], Romana STOPKOVÁ[1], Aneta ZEMANOVÁ[1], Kristýna MÍČKOVÁ[1], Karel HARANT[1], Oldřich TOMÁŠEK[1,2], Tomáš ALBRECHT[1,2,†,*] and Pavel STOPKA[1,†,*]

[1] *Department of Zoology, Faculty of Science, Charles University, BIOCEV, Praha and Vestec, Czech Republic; e-mail: pstopka@natur.cuni.cz, otcenast@natur.cuni.cz, romana.stopkova@natur.cuni.cz, anetazemanova@icloud.com, mickovkr@natur.cuni.cz*

[2] *Institute of Vertebrate Biology of the Czech Academy of Sciences, Brno, Czech Republic; e-mail: albrecht@ivb.cz, oltmsk@gmail.com*

Received 11 May 2023; Accepted 20 July 2023; Published online 11 September 2023

Abstract. Spermatozoa are the most variable cells across animal taxa. Phylogeny, speciation and postcopulatory sexual selection are typical factors that explain the sperm morphology variation in animals, and now these differences can also be explored on the level of genomic and proteomic differentiation. However, in non-model organisms, it is often difficult to employ these techniques because genomes are not yet available for most animal species, particularly for free-living songbirds (Passeriformes). Here, we employed label-free proteomics to generate proteomes in the zebra finch, a songbird species with an annotated genome and five wild-living songbirds representing five families within the Passerida clade, all with poorly known genomes. The results show that protein mapping of the new passerine proteomes to the zebra finch genome was successful, thus yielding highly similar protein identifications and a sufficient number of unique peptides in all the studied proteomes. Interestingly, while passerine sperm proteomes only partially reflect phylogenetic relationships between passerine families, midpiece length correlates with at least 59 proteins enriched in mitochondrial metabolism. Similar sperm proteomes seem to have evolved convergently across passerine lineages, potentially due to varying levels of sperm competition and marked variation in sperm sizes.

Key words: proteome, sperm competition, *Taeniopygia guttata*, *Acrocephalus palustris*, *Hirundo rustica*, *Cinclus cinclus*, *Chloris chloris*, *Phylloscopus collybita*

Introduction

Sperm is the male reproductive cell that exhibits exceptional morphology variation across animal taxa yet with a certain degree of phylogenetic conservatism (Pitnick et al. 2009, Kahrl et al. 2021). Apart from phylogenetic and speciation events, postcopulatory sexual selection imposed via sperm competition and cryptic female choice is considered the main evolutionary force responsible for such phenotypic diversity (Jamieson 1987, Pizzari & Parker 2009, Simmons & Fitzpatrick 2012, Vicens et al. 2017, Lifjeld et al. 2019). In mammals as well as birds, intraspecific variation in both sperm size and design are especially apparent in species with greater levels of promiscuity (Immler et al. 2008, Kleven et

* *Corresponding Author*

† *These authors jointly supervised this work*

al. 2008, Albrecht et al. 2013, Sandera et al. 2013). Due to the competition between sperm from different mating partners within the female reproductive tract, sperm traits conferring an advantage over rival sperm are under strong positive selection (Simmons & Fitzpatrick 2012).

In most passerine birds (Passeriformes), representatives of the largest avian clade, spermatozoa have a typical corkscrew shape with a helical membrane twisted around the acrosome and fused elongated mitochondria wound around the flagellum (Humphreys 1972, Jamieson 2007). Interestingly, morphological traits such as sperm head length and midpiece with flagellum length are likely under different selection forces. Whereas the length of the sperm head is rather evolutionarily conserved due to its involvement in the fine-tuned process of fertilisation and the negative impact of its increase on sperm swimming speed (Jamieson 2007, Rowe et al. 2015), flagellum length diverged rapidly in response to varying levels of sperm competition even across phylogenetically related taxa (Rowe et al. 2015). In general, a positive correlation exists between the size of the midpiece and flagellum (Lüpold et al. 2009) and the intracellular concentration of adenosine triphosphate (ATP) in passerines (Rowe et al. 2013). Nevertheless, in contrast to Galliformes (Froman & Feltmann 1998), in passerines, higher ATP levels were not found to translate into higher swimming velocity of spermatozoa (Rowe et al. 2013). This situation could be seen as contradictory evidence to the hypothesis that longer sperm are preferred during the postcopulatory phase of sexual selection in birds (Lüpold et al. 2009, Pizzari & Parker 2009) and there is a positive association between the intensity of sperm competition and total sperm length across passerines (Kleven et al. 2009). However, a possible explanation is that the total availability of ATP is utilised also in other functions different from motility (Rowe et al. 2013).

Unlike most songbirds (Lifjeld et al. 2019), the zebra finch *Taeniopygia guttata* is characteristic of infrequent sperm competition and remarkable inter-male variability in sperm design (Birkhead et al. 1989, 2005). Intraspecific comparison within the zebra finches of sperm with midpieces varying in length revealed higher ATP concentrations in shorter midpieces (Mendonca et al. 2018). Although there was a positive association between the midpiece size and total sperm length or sperm velocity in zebra finch males (Knief et al. 2017), this pattern was not detectable across 42 free-living passerine species (Kleven et al. 2009). In passerine sperm, the main metabolic pathway of ATP production is likely the oxidative phosphorylation (OXPHOS), considering that a fibrous sheath, a structure with bound glycolytic enzymes (Krisfalusi et al. 2006), appears to be absent (Sexton 1974, Jamieson 2007). Denser packing of OXPHOS complexes within a shorter and thicker midpiece results in higher ATP synthesis efficiency (Mendonca et al. 2018). For a broader understanding of selection on favoured sperm phenotype, morphological traits other than simple length measurements should be considered. Although rarely applied due to technological difficulties, high-resolution imaging methods such as electron microscopy represent an excellent tool for obtaining various multi-dimensional sperm measures. For example, using scanning electron microscopy, Støstad et al. (2018) showed that the more pronounced helical shape of the passerine sperm head generates greater forward propulsion, leading to a greater swimming speed. Thus, the sperm head morphology also contributes to sperm performance.

Given the complexity of postcopulatory sexual selection regulating fertilisation success, high-throughput proteomic approaches are increasingly applied in the reproductive biology of birds. Several studies show that the seminal fluid proteins modulate sperm performance, a process driven by sexual selection, especially in polyandrous avian species (Perry et al. 2013, McGraw et al. 2015). Notably, proteins of immune-related pathways (i.e. innate immunity response and antimicrobial defence) are likely under positive selection in both passerine (Rowe et al. 2020) and non-passerine species (Borziak et al. 2016, Słowińska et al. 2017). To date, only two studies providing protein profiles of sperm isolated from chicken (Soler et al. 2016) and zebra finch (Rowe et al. 2019) are available.

In this study, using label-free proteomics (nLC-MS/MS) we investigated the proteomic background of sperm morphology variation across six passerines representing six families of the Passerida clade, namely the marsh warbler (*Acrocephalus palustris*) – AP, European greenfinch (*Chloris chloris*) – CH, common chiffchaff (*Phylloscopus collybita*) – PC, white-throated dipper (*Cinclus cinclus*) – CC, and the barn swallow (*Hirundo rustica*) – HR, while the proteome of the zebra finch (*T. guttata*) – TG was treated as standard (Fig. 1A). This initial study aimed to: i) show the efficiency of the mapping of new proteins against the reference dataset in the identification of both similar and unique protein entries; and ii) reveal the

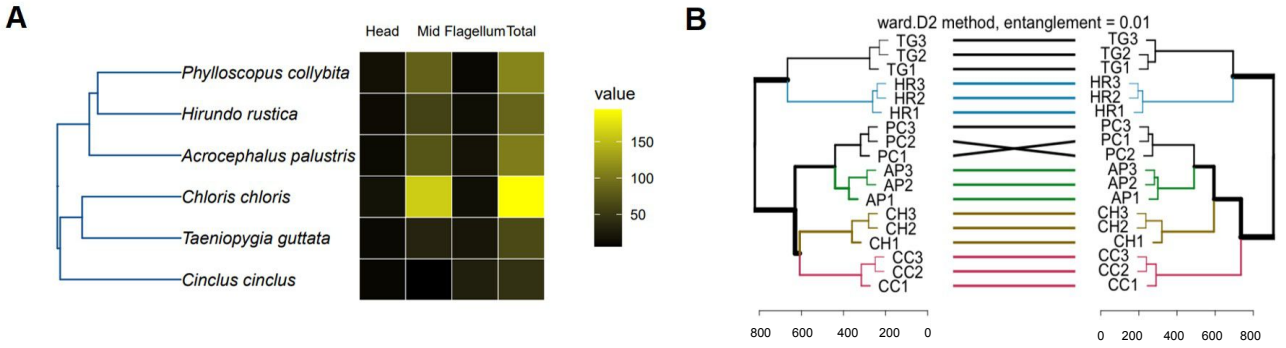

Fig. 1. Evolutionary history of the studied species and comparisons of sperm proteomes. The evolutionary history (A) was adopted from Jetz et al. (Jetz et al. 2012). Clustered image map (next to the branches) to visualise selected sperm dimensions (head, midpiece, length of flagellum, and the total sperm length in µm). Tanglegram between sperm (left) and seminal fluid (right) proteomes is in B. AP – *Acrocephalus palustris*; CC – *Cinclus cinclus*; CH – *Chloris chloris*; HR – *Hirundo rustica*; TG – *Taeniopygia guttata*.

potential of proteomic analysis of passerine sperm to explain the evolution of interspecific differences in sperm design and evaluate whether interspecific variation in sperm proteomes reflects the phylogeny (e.g. Jetz et al. 2012) or sperm morphology, i.e. mainly the midpiece size, because this structure containing mitochondria is the most variable character between our studied species and thus is our proxy for sperm competition which is driven by promiscuous mating.

Material and Methods

Subjects and sample collection

Our analyses used three males of six passerine species representing several families within the Passerida clade. Five species were free-living, while males from the captive population of zebra finches housed at the Institute of Vertebrate Biology bird breeding facility were used as the reference samples. Males of free-living species were released, immediately after which a sperm sample was obtained. Sampling was performed in May-June 2014-2015 in southern Bohemia, Czech Republic. Sperm samples were obtained by a standard massage of the cloacal protuberance (Wolfson 1952). Ejaculates (approx. 1 µl) were immediately dissolved in 20 µl of sterile phosphate-buffered saline (PBS). An aliquot (3 µl) was stored in 10% formalin for further inspection of sperm morphology and ejaculate quality (presence of red blood cells or debris), and the rest was deep frozen in liquid nitrogen. Only clean samples containing only sperm cells (based on the inspection of aliquotes) were used in further proteomic analysis and sperm measurements.

Sperm morphology

Formalin-fixed sperm samples were transferred with an automatic pipette onto a slide, smeared, air-dried and rinsed in dH_2O to wash off impurities. The images were captured on a digital camera (DP71, Olympus, Japan) mounted on a light microscope (BX51, Olympus, Japan). The measurements of three sperm parameters (head length, midpiece length and tail length) were made using the imaging software QuickPHOTO Industrial (Olympus, Japan). From these measurements, we calculated the total sperm length and flagellum length (the sum of midpiece length and tail length). For each individual, ten spermatozoa with normal sperm morphology were analysed, and the measured values were averaged for each species.

Proteomic analysis

Sperm samples (n = 18) were defrosted at room temperature, vortexed and a total of 10 µl transferred to new 0.5 ml tubes to which we added 20 µl of PBS. The samples were vortexed (1 min) and centrifuged (14,000 RCF) to obtain sperm cells (pellets: n = 18) and seminal fluids (supernatant: n = 18). Next, we added 100 µl of isolation buffer containing 8 M urea, thiourea, 1% Chaps, 20 mM DTT and dH_2O to each tube and left it at room temperature for one hour. Protein samples were precipitated with the ice-cold acetone (1:4) and centrifuged (14,000 RCF). This procedure was followed by a re-suspension of dried pellets in the digestion buffer (1% SDC, 100 mM TEAB – pH = 8.5). Protein concentration of each lysate was determined using the BCA assay kit (Fisher Scientific, Waltham, MA, USA). Cysteines in 20µg of proteins were reduced with a final concentration of 5 mM TCEP (60 °C for 60 min) and blocked with 10 mM MMTS (i.e. S-methyl methanethiosulfonate, 10 min at room temperature). Samples (n = 36) were cleaved with trypsin (i.e. 1/50, trypsin/protein) at 37 °C overnight. Peptides were desalted on a Michrom C18 column. Eluting peptide cations were converted to gas-phase ions by electrospray ionisation and analysed on a Thermo Orbitrap Fusion (Q-OT-qIT; Thermo Fisher, Waltham, MA, USA) using the same conditions and setup as

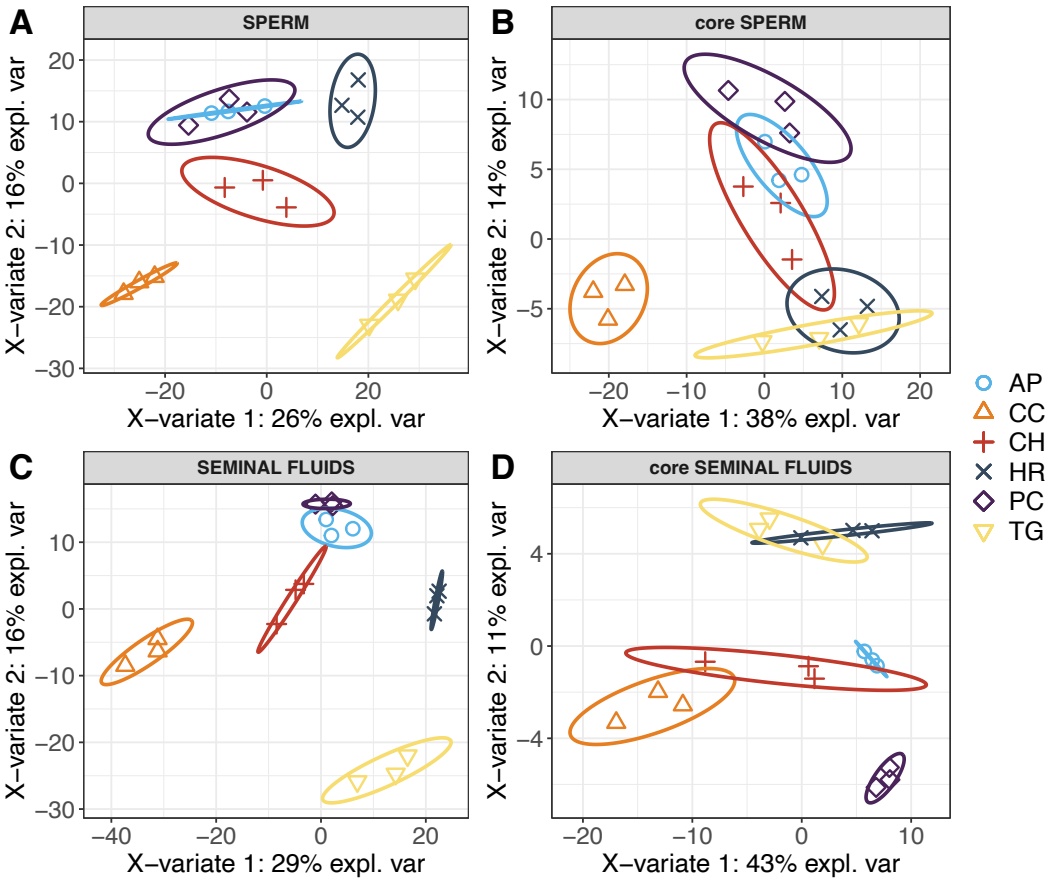

Fig. 2. Diverging patterns of passerine proteomes. Sparse Partial Least Squares Discriminant Analysis (sPLS-DA) revealed that all the studied passerines significantly separated from TG. Sperm proteomes are in A, B, seminal fluid proteomes are in C, D. Note that the phylogeny of the two major clades is well demonstrated by the second component (Comp2 – y-axis) in all discrimination. AP – *Acrocephalus palustris*; CC – *Cinclus cinclus*; CH – *Chloris chloris*; HR – *Hirundo rustica*; TG – *Taeniopygia guttata*.

previously described (Stopka et al. 2016, Kuntova et al. 2018). For example, survey scans of peptide precursors from 400 to 1,600 m/z were performed at 120 K resolution (at 200 m/z) with a 5 × 105 ion count target. Tandem MS was performed by isolation at 1.5 Th with the quadrupole, HCD fragmentation with normalised collision energy of 30, and rapid scan MS analysis in the ion trap. The MS2 ion count target was set to 104 and the max injection time was 35 ms. Only those precursors with charge states 2-6 were sampled for MS2. The dynamic exclusion duration was set to 45 s with a 10 ppm tolerance around the selected precursor and its isotopes. Monoisotopic precursor selection was turned on. The instrument was run in top speed mode with 2 s cycles.

Bioinformatics

All data were quantified using MaxQuant software version 1.6.34. The false discovery rate (FDR) was set to 1% for identifying all peptides and proteins. We set a minimum peptide length of seven amino acids. Enzyme specificity was set as C-terminal to Arg and Lys, allowing cleavage at proline bonds 98 and a maximum of two missed cleavages. Dithiomethylation of cysteine was selected as fixed modification, and N-terminal protein acetylation and methionine oxidation as variable modifications. The 'match between runs' feature of MaxQuant was used to transfer identifications to other LC-MS/MS runs based on their masses and retention time (maximum deviation 0.7 min). Quantifications were performed using the label-free algorithms 65 with a combination of unique and razor peptides. The Andromeda search engine was used for the MS/MS spectra search against the Uniprot database (UP000007754, 2021), with all duplicates being removed. Our raw dataset was LFQ normalised (Cox et al. 2014) with MaxQuant and is available as Table S1. From this point, all data manipulations, filtering, statistical analyses and plots were generated in R software (Crawley 2007). First, we removed all protein identifications that were detected but not quantified (zero abundances), all proteins with less than two unique peptides, and all contaminants from MaxQuant output files (porcine trypsin, bovine albumins, human keratins, etc.). Next, we removed singletons (rare cases) from

our data such that only those proteins produced by all three individuals in at least one species (e.g. TG) were passed to further analyses; otherwise, the whole row was deleted. This data format served for data exploration with the Sparse Partial Least Squares Discriminant Analysis (sPLS-DA) within the '*mixOmics*' package (Rohart et al. 2017). To make sure that interspecific variation is higher than the variation between replicates, we performed Spearman's rank correlations between individual replicates and between averaged values per species. Spearman's rank correlation was also used to find relationships between midpiece length and the abundance of individual proteins. Gene ontology (GO) searches were performed using Gene Set Enrichment Analysis (GSEA) within the '*clusterProfiler*' package (Yu et al. 2012) from Bioconductor, while significant GO terms were searched by using p-value cut-off $P = 0.1$ (P-value < 0.05 revealed only a few terms) within the latest chicken database (August 2022) – org.Gg.eg.db. We recognise that there are limitations to using the chicken database for passerine species. If we used a passerine database, we would struggle with another limitation: the presence of zebra finches in our samples and pure annotations. So, we opted for chicken, which is similarly phylogenetically distant to all our studied species. Also, this approach does not look at species-specific proteins due to data reduction. To add, we identified 97 significantly correlated proteins in sperm samples, which successfully mapped to 72 GO terms (i.e. 74%). Our final dataset has an approximately similar number of proteins in all the species: 1,198 and a reasonable number of unique peptides (> 2). With this approach, we probably omitted many reproductive proteins that evolve faster, but the data set thus enabled comparison between species, at least in the proteins they share. All plots and figures were generated in R using '*ggplot2*' and related packages (Wickham 2016).

Results and Discussion

Characterisation of sperm and seminal fluid proteomes

The proteomic dataset was generated with nLC-MS/MS from 36 samples containing 18 sperm samples and 18 fluid samples, obtained from six species with three replicates per species. The whole matrix extracted from label-free quantification (Cox et al. 2014) was LFQ normalised. The resulting dataset is based on 144,449 high-quality spectral matches to 2,384 protein identifications (Table S1). Of these, 88.2% were identified by multiple unique peptides shared between species (on average 7.2). These searches were conducted against *T. guttata* proteome (UP000007754, 2021) in the absence of the annotated genomes in the other species. We removed all rows with fewer unique peptides than two, all potential contaminants and all proteins identified but not quantified. Next, the matrix was split into the sperm matrix (n = 1,203) and seminal fluid matrix (n = 1,130), while both matrices contained a similar number of identifications. After the data reduction, we used 'normalizerDE' in R to perform all eight most common normalisations (e.g. log2, VSN, Quantile, GI, etc.) and used the Quantile normalisation simply because the within triplicate variation was the lowest. For the subsequent analyses, we used two datasets, the complete proteome and the so-called core proteome, which used only the proteins detected in all the studied samples/individuals.

The hierarchical clustering of sperm and seminal fluid proteomes revealed a surprisingly similar branching structure (entanglement = 0.01). This structure is, however, different from the phylogenetic branching in Fig. 1A. For example, genetically different TG and HR cluster together in Fig. 1B but not in Fig. 1A. Thus, the phylogeny was to some extent overridden by as yet unspecified ecological factors that have driven proteome evolution. Correlations between individual replicates were high (r > 0.8). In contrast, correlations between averaged TG replicates with other species were much lower (approx. 0.5). Particularly, Spearman's rank correlation between TG and other passerines was relatively low (SPERM: $r_{CH} = 0.51$, $r_{CC} = 0.29$, $r_{AP} = 0.46$, $r_{PC} = 0.48$, $r_{HR} = 0.52$; FLUIDS: $r_{CH} = 0.51$, $r_{CC} = 0.41$, $r_{AP} = 0.46$, $r_{PC} = 0.44$, $r_{HR} = 0.46$; $P < 2.2e\text{-}16$), while the correlation between biological replicates was generally high (SPERM: TG – 0.77-0.85, CH – 0.77-0.86, CC – 0.75-0.85, AP – 0.79-0.86, PC – 0.79-0.83, HR – 0.86-0.89; FLUIDS: TG – 0.78 for all comparisons, CH – 0.73-0.88, CC – 0.56-0.71, AP – 0.80-0.85, PC – 0.84-0.88, HR – 0.85-0.89; $P < 2.2e\text{-}16$). High correlations between biological replicates are also demonstrated in a tanglegram (Fig. 1B), where individuals cluster together within a particular species.

Sources of proteomic variation

We used the Sparse Partial Least Squares Discriminant Analysis (sPLS-DA) to assess the potential sources of variation in proteomes. In Fig. 2, we demonstrate using the Area Under Curve analysis (AUC) that all the five passerine species in this study have patterns significantly diverging from TG (e.g. for Fig. 2A – Comp1: AUC = 1, $P = 0.008$; Comp2: AUC = 1.0, $P = 0.008$). A similarly diverging pattern observed in sperm samples (Fig. 2A, B) was obtained with

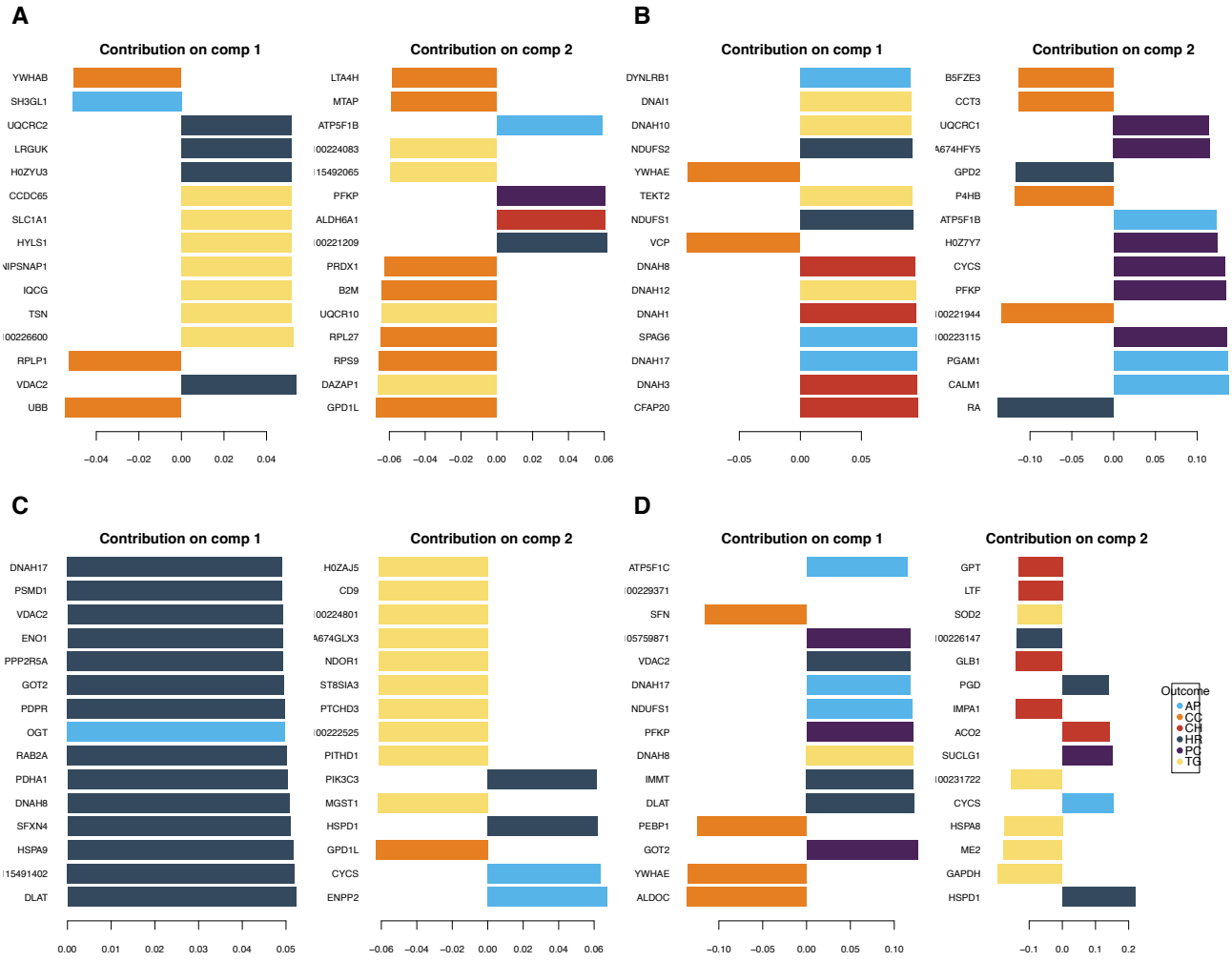

Fig. 3. The contribution of particular proteins to species differentiation is shown with loadings from sPLS-DA analysis. The colour coding is similar to in Fig. 2 and indicates the species in which the protein has the maximum mean value. Sperm proteomes are in A and B, while seminal fluid proteomes are in C and D (complete proteomes left, core proteomes right). The importance of proteins for separation between species is ranked from the bottom (the greatest) to the top. AP – *Acrocephalus palustris*; CC – *Cinclus cinclus*; CH – *Chloris chloris*; HR – *Hirundo rustica*; TG – *Taeniopygia guttata*.

sPLS-DA on the data from seminal fluids (Fig. 2C, D) (*P*-values for all comparisons are in Table S1). Note that sPLS-DA is separated by the second component (Comp2), the two major clades from Fig. 1A. TG, CH, and CC are well separated from AP, HR, and PC, reflecting species phylogenetic relatedness. Next, we investigated the key proteins responsible for separating given species and whether these proteins represent biologically relevant features such as particular sperm metabolic pathways or molecular

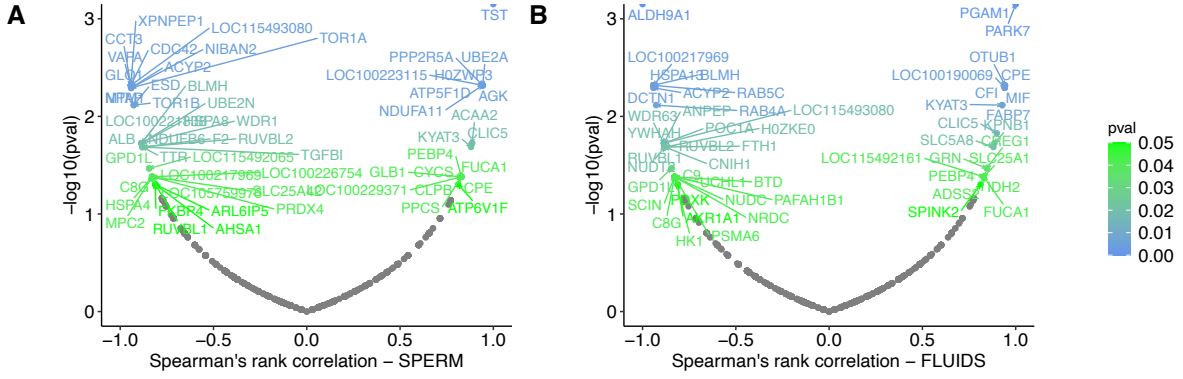

Fig. 4. Correlation of complete proteomes with sperm midpiece length. We performed Spearman's rank correlation between midpiece length and sperm proteomes (A) and fluids proteomes (B) averaged over species. Significantly correlated proteins are scaled from green ($P < 0.05$) to blue ($P < 0.01$).

functions. For this purpose, we visualised the top 15 loading values and thus the importance of proteins contributing to each principal component (Comp1 and Comp2) of the sPLS-DA model (Fig. 3). Proteins are identified as important based on the mean abundancy value in each species. Positivity and negativity of importance values (i.e. positive and negative contribution on each component) indicate species in which the protein is highly abundant; however, proteins with positive and negative importance values contribute equally to species separation. The most important result from this analysis revealed that sperm and seminal fluids have different proteins responsible for the differentiation, which thus provides evidence that the two fractions from the ejaculate samples were not biased by extensive cross-contamination. There is yet another message in the visualisations of loadings values; some species differ from others by lower protein expressions while others have higher expressions when compared to other species (do not mistake with negative and positive expression indicated by negative and positive values of importance, respectively). The same colour code representing species is used as in Fig. 2.

Highly important sperm proteins play important roles in sperm metabolism, e.g. ATP-dependent 6-phosphofructokinase (PFKP), phosphoglycerate mutase 1 (PGAM1), aldehyde dehydrogenase 6 family member A1 (ALDH6A1), glycerol-3-phosphate dehydrogenase NAD(+) (GPD1L), glycerol-3-phosphate dehydrogenase (GPD2), and in energetics, e.g. cytochrome c, somatic (CYCS), NADH-ubiquinone oxidoreductase subunit, mitochondrial (NDUFS1), NADH dehydrogenase iron-sulfur protein 2, mitochondrial (NDUFS2), cytochrome b-c1 complex subunit Rieske, mitochondrial (UQCRC1), cytochrome b-c1 complex subunit 2, mitochondrial (UQCRC2) and complex III subunit 9 (UQCR10) (Fig. 3A, B). Like spermatozoa, seminal fluids contain proteins with a high importance involved in energetics, e.g. CYCS, NDUFS1, metabolism, e.g. aconitate hydratase, mitochondrial (ACO2), acetyltransferase component of pyruvate dehydrogenase complex

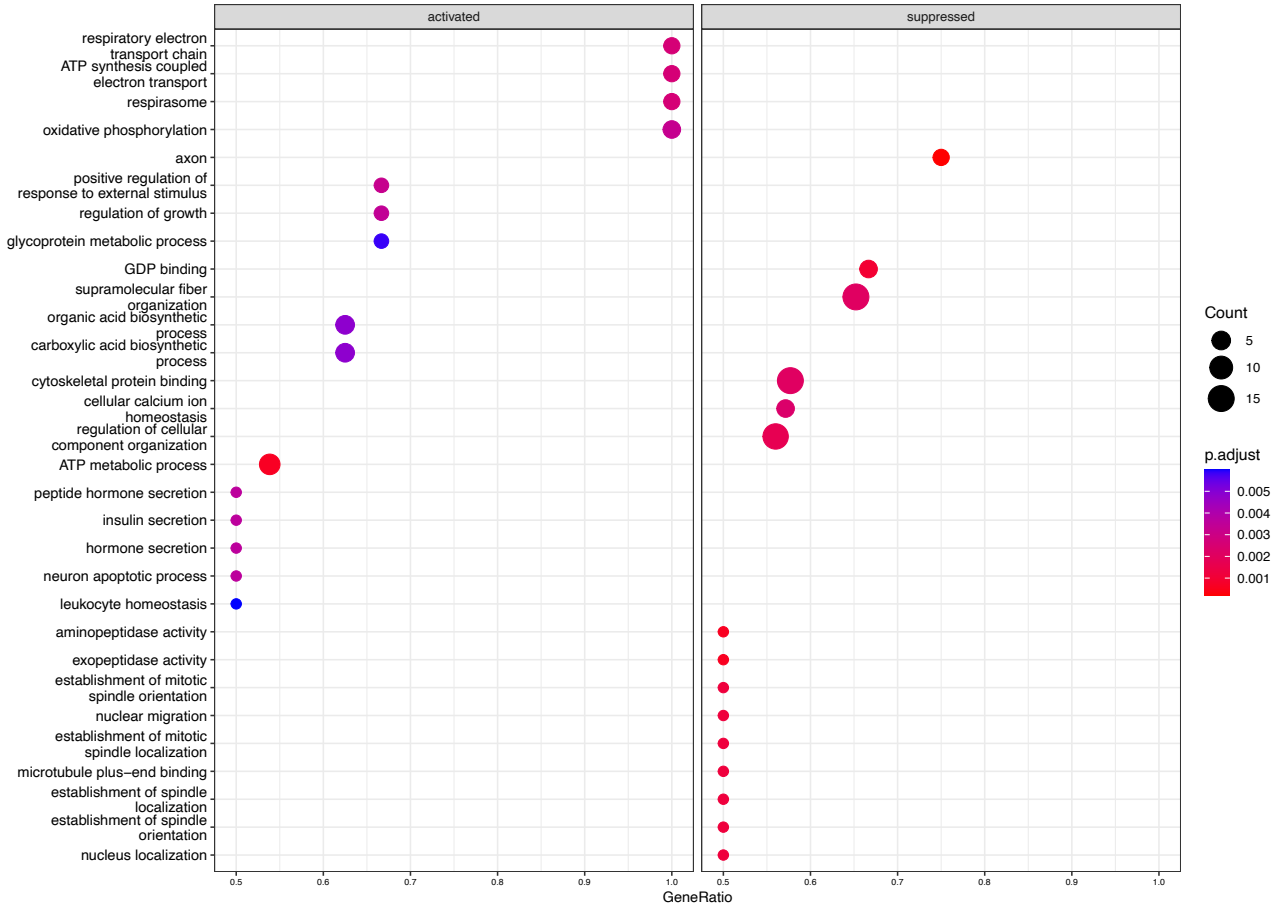

Fig. 5. Enriched gene ontology (GO) terms revealed with Gene Set Enrichment Analysis (GSEA). The colour of the dots represents the level of significance adjusted with Benjamini-Hochberg correction for multiple testing. The size of the dots indicates the number of proteins enriched in that term. Activated proteins (left) are those positively correlated with the midpiece length. Note that many of them are related to mitochondrial metabolism. Suppressed proteins (negatively correlated) are on the right.

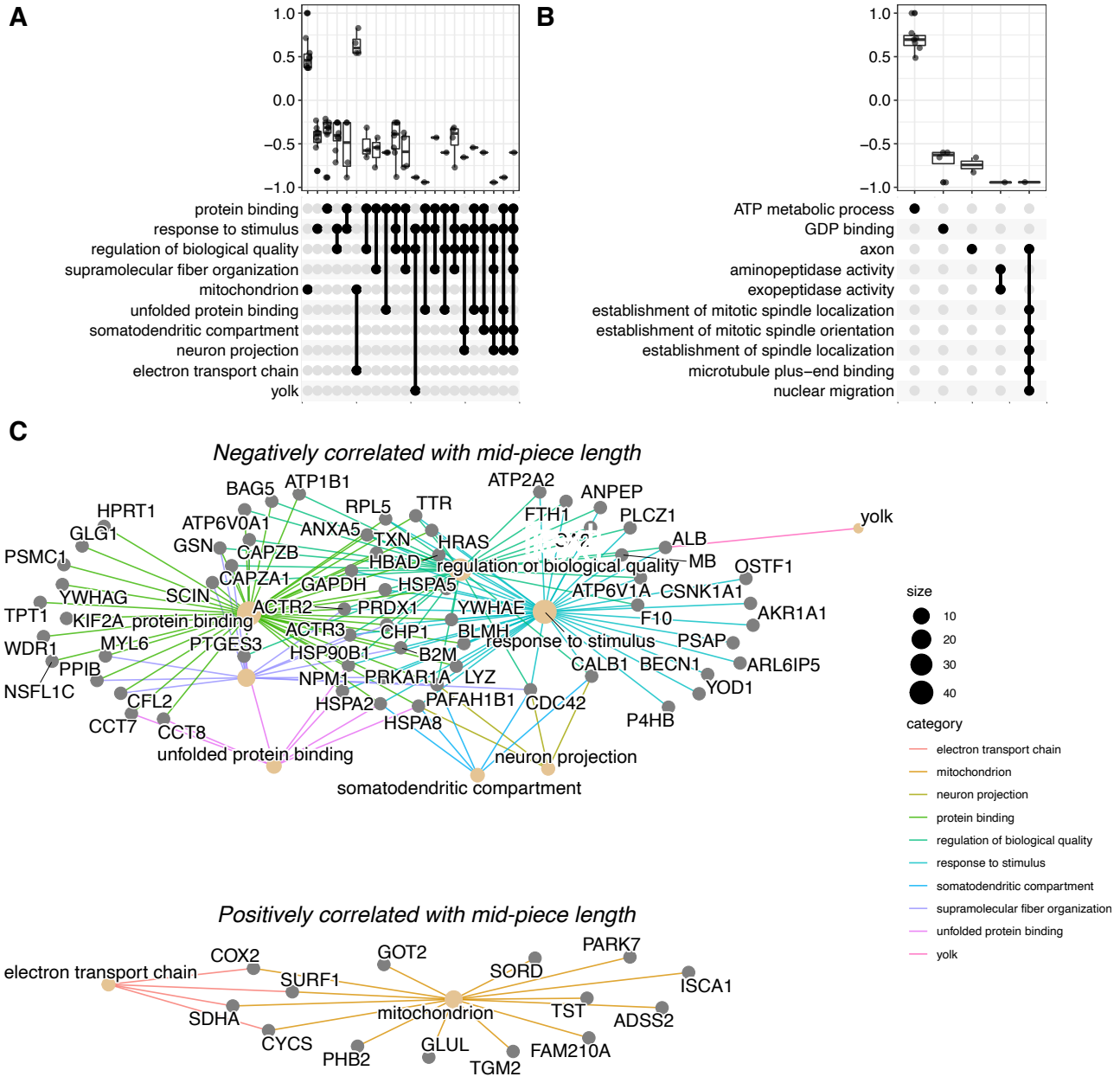

Fig. 6. GSEA analysis showing the prevalence of GO terms associated with mitochondria. The distribution of activated and suppressed proteins of the complete sperm (A) and seminal fluid (B) proteomes and their association with one or more GO pathways (see the intersection of black dots) is captured with UpSet plots. Network analysis in C shows the complex association between proteins and multiple annotation terms.

(DLAT), PFKP, fructose-bisphosphate aldolase (ALDOC), GPD1L, fertilisation, e.g. tetraspanin CD9, and in protein degradation, e.g. 26S proteasome non-ATPase regulatory subunit 1 (PSMD1) (Fig. 3C, D).

In detail, PFKP has positive importance values for both sperm datasets and the core proteome of seminal fluids in PC, a highly promiscuous species (EPP = 48% in (Lifjeld et al. 2019)). In the same species, CYCS is important for separating the core sperm proteome. On the other hand, proteins DLAT and heat shock protein (HSPD1), which are both mitochondrial, are positively separating seminal fluids of HR from other species regardless of the dataset used. Further, a similar contribution is observed for protein CYCS in differentiation of AP species. In CC, monogamous species with the shortest sperm tails examined is cytosolic GPD1L responsible for separation in complete proteomes, although with negative values of importance. Taken together, although the sperm design looks pretty much the same in all the six studied species, there has been a selective pressure that has driven biochemical sperm components to species-specific states.

Interestingly, mammalian acrosomes also contain mitochondrial and proteasomal proteins (Guyonnet et al. 2012, Otčenášková et al. 2023), presumably because redundant mitochondria are removed from the cytosol after midpiece formation, and the resulting proteins are ubiquitinated and deposited in acrosomes. Here, we provide new evidence that this could be also true for avian species. Thus, the presence of mitochondrial proteins in seminal fluids practically means that the similarity between sperm and fluid dendrograms in Fig. 1B is highly influenced by proteins occurring in both fractions. On the other side, the presence of acrosomal proteins, e.g. tetraspanin CD9, acrosin-binding protein (ACRBP) and zona pellucida-binding protein 2 (ZPBP2), in seminal fluids might provide important information on the acrosomal content. Noteworthy, these proteins, together with acrosin (ACR), tetraspanin CD81, hyaluronidase PH-20 (SPAM1) involved in sperm-oocyte interactions, have been recently detected in seminal fluid proteomes in chicken and two closely-related sparrow species (Borziak et al. 2016, Li et al. 2020, Rowe et al. 2020). Although the abundance of acrosomal proteins in seminal fluid is often explained by acrosome damage either during sample preparation (which might also be our case) or as an indicator of poor semen quality (Li et al. 2020), they still might be of biological significance.

Functional divergence of proteomes

Our above candidate-proteins approach revealed that many proteins detected in sperm proteomes are involved in metabolic pathways and mitochondria-related energy production. Moreover, our morphology analysis revealed that midpiece length is the most variable character between species. Thus, we used the mean sperm midpiece length for Spearman's rank correlation with individual proteins, also averaged over species. Figure 4 provides the graphical visualisation of the correlation structure. In general, sperm proteomes revealed more proteins correlating with midpiece length (59 out of 1,198) than fluids (12 out of 1,125). As expected, sperm proteins that significantly correlated with midpiece length are primarily involved in mitochondrial energetics (Fig. 4A). These included several subunits and components of electron-transport chain and ATP synthase complexes driving OXPHOS namely CYCS, NADH dehydrogenase 1 alpha subcomplex subunit 11 (NDUFA11), ATP synthase F1 subunit delta (ATP5F1D) and V-ATPase subunit (ATP6V1F). The strongest correlation, however, was assigned to thiosulfate sulfurtransferase (TST), which serves as a mitochondrial importer of rRNA ($r = 1$, $P < 0.0001$). Other correlated proteins function in diverse metabolic processes like lipid phosphorylation, fatty acid beta-oxidation, amino acid catabolism or acetyl-CoA biosynthesis. Interestingly, these categories were also enriched in other studies providing interspecific comparisons of sperm proteomes from mouse species exhibiting different levels of sperm competition (Vicens et al. 2017). It is also interesting that seminal fluids contain proteins that significantly correlate with midpiece length (Fig. 4B). The most positively correlated proteins include PGAM1, catalysing glycolysis and gluconeogenesis and Parkinsonism associated deglycase (PARK7) that mediates stress response (for both $r = 1$, $P < 0.0001$). Further, other proteins are involved in innate immune responses and inflammation, such as macrophage migration inhibitory factor (MIF), complement factor I (CFI), and progranulin (GRN), all being secreted into the extracellular space. Some proteins were similar to sperm and seminal fluid samples, which may reflect the known seminal fluid modulation of sperm performance (Perry et al. 2013, McGraw et al. 2015).

Important evidence that the variation between proteomes is driven by sperm competition – demanding high-energy input – comes from Gene Set Enrichment Analysis (GSEA) applied for gene ontology (GO) annotation. Rowe et al. (2019) provided the first proteome of the zebra finch spermatozoa and revealed that the most enriched sperm proteins function in biological processes linked to mitochondria. There is an overlap between GO terms detected in (Rowe et al. 2019) and our data. When looking at significantly enriched GO terms in this study, most of them are mitochondria-related (Fig. 5, 6). We show that for positively correlated proteins (termed as activated) from the complete sperm proteome, the most enriched terms include oxidative phosphorylation (GO:0006119), respiratory electron transport chain (GO:0022904) and ATP synthesis coupled electron transport (GO:0042773) from the Biological Process category and mitochondrion (GO:0005739) with mitochondrial intermembrane space (GO:0005758) representing the Cellular Component category (Fig. 5, 6A). A similar result is obtained with network analysis (Fig. 6C), showing a group of proteins that are strictly related to GO terms mitochondrion (GO:0005739) and electron transport chain (GO:0022900) and are not involved in another pathway. Note that some of them were found to positively correlate with midpiece length (Fig. 4). Notably, our data further support descriptive and comparative cross-species proteomics studies that

attempt to establish a core sperm proteome. Functional categories of proteins that are associated with sexual reproduction, flagellar movement and metabolic processes, e.g. mitochondria-related ATP production, are enriched in spermatozoa across a wide range of vertebrate and invertebrate taxa (Bayram et al. 2016, Whittington et al. 2017, 2019, Rowe et al. 2019, Garlovsky et al. 2022, Poignet et al. 2022). Moreover, the analysis of the whole seminal fluid proteome revealed that the fluids dataset is dominated by the ATP metabolic process (GO:0046034), which is linked to activated proteins (Fig. 6B). This further supports the previous argument that fluids might be a mixture of seminal fluids and acrosomal content. Negatively correlated (termed as suppressed) sperm (Fig. 5, 6A) and seminal fluids proteins (Fig. 6B) function, on the other side, in cellular processes like regulation of biological quality (GO:0065008), response to stimulus (GO:0050896), or have distinct molecular functions such as protein binding (GO:0005515), GDP binding (GO:0019003), aminopeptidase activity (GO:0004177) and exopeptidase activity (GO:0008238).

Broader context

Taking a broader view, we touched upon a fundamental question in reproductive biology and evolution: the relationships between the phylogeny, varying levels of sperm competition, sperm morphology (size), and proteomic variation. We found that each species differs from the other by species-biased groups of proteins and that these groups represent biologically relevant features such as mitochondrial energy production, metabolism, protein degradation and sperm-egg fusion. These features represent important building blocks of sperm cells in the view of sperm competition because mitochondria are the drivers of proper sperm movement while the acrosome-mediated gamete interaction and fusion are the key processes during fertilisation. Particular roles of these proteins were corroborated by gene ontology analysis, where most proteins that correlate with sperm midpiece length are involved in mitochondrial energetic metabolism. Although this result appears expected, it is not trivial. This is because some mammalian species (e.g. muroid rodents) have variable apical hooks containing proteins involved in fertility and acrosome reaction (Breed 2004, Immler et al. 2007, Johnson et al. 2007, Clift et al. 2009, Sandera et al. 2013, Hook et al. 2021) and, for example, resulting proteomes of mouse spermatozoa and acrosomes contain proteins enriched in energetic metabolism (Vicens et al. 2017, Otčenášková et al. 2023). In this study, the fertilisation process was not detected by GO analysis. However, it was detected with sPLS-DA techniques, which revealed that combining several techniques to extract relevant biological features is helpful. Although this study brings just facets of sperm divergence on the proteomic level, the results of fifty years of solving the enigma of sperm competition in animals (Parker 1970, 1992, Dixson 1993, Dixson & Anderson 2004, Gomendio et al. 2006, Immler 2008, Pizzari & Parker 2009, Vicens et al. 2017, Birkhead & Montgomerie 2020) are now explored with new OMICs techniques, which helps to detect important candidate proteins and their involvement in particular biological processes. These could be further explored on larger datasets to reconstruct the selective forces that shaped the evolution of sperm competition in animals.

Conclusions

We found that proteomic variation in sperm and seminal fluids from six passerine species partially follows their phylogeny. However, a group of important proteins is responsible for given species-specific patterns. These proteins are often linked to mitochondrial metabolism and oxygen transport, and thus, it is likely that this study, for the first time in passerines, linked genotype with sperm phenotype because, among the four measured sperm dimensions, the mitochondria-containing midpiece is also the most variable character across the six studied species. Despite the limitations of this study, such as the low number of species or the lack of appropriate passerine genomes for protein mapping, we were still able to demonstrate that selective forces that shaped the evolution of passerine sperm most likely acted upon mitochondrial metabolism and midpiece size and not on general sperm metabolism and morphology.

Acknowledgements

We are grateful to Pavel Talacko for running the MS equipment within the Proteomic Core Facility, BIOCEV, Faculty of Science, Charles University, Praha and Vestec. Funding: A grant from the Czech Science Foundation (GAČR) project No: 19-22538S funded the whole research. K. Míčková was funded by GAUK 1308120. The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation. Institutional Review Board Statement: the study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of the Faculty of Science, Charles University in Praha. All animal procedures were carried out in strict accordance with the law of the Czech Republic paragraph 17 no. 246/1992.

Author Contributions

T. Albrecht and P. Stopka conceived and designed the study. P. Stopka, T. Otčenášková and T. Albrecht wrote the first draft of the manuscript. T. Albrecht, O. Tomášek and K. Míčková collected the samples. K. Míčková measured the sperm samples. R. Stopková processed the samples for nLC-MS/MS and, with T. Otčenášková and A. Zemanová, helped to prepare the final datasets and GO analyses. K. Harant operated the MS instruments and helped to prepare datasets. All authors participated in writing and reviewed the final manuscript. The authors declare that they have no competing interests.

Literature

Albrecht T., Kleven O., Kreisinger J. et al. 2013: Sperm competition in tropical versus temperate zone birds. *Proc. R. Soc. B Biol. Sci. 280: 20122434.*

Bayram H.L., Claydon A.J., Brownridge P.J. et al. 2016: Cross-species proteomics in analysis of mammalian sperm proteins. *J. Proteomics 135: 38–50.*

Birkhead T.R., Hunter F.M. & Pellatt E.J. 1989: Sperm competition in the zebra finch, *Taeniopygia guttata*. *Anim. Behav. 38: 935–950.*

Birkhead T.R. & Montgomerie R. 2020: Three decades of sperm competition in birds. *Philos. Trans. R. Soc. B: Biol. Sci. 375: 20200208.*

Birkhead T.R., Pellatt E.J., Brekke P. et al. 2005: Genetic effects on sperm design in the zebra finch. *Nature 434: 383–387.*

Borziak K., Álvarez-Fernández A., Karr T.L. et al. 2016: The seminal fluid proteome of the polyandrous red junglefowl offers insights into the molecular basis of fertility, reproductive ageing and domestication. *Sci. Rep. 6: 35864.*

Breed W.G. 2004: The spermatozoon of Eurasian murine rodents: its morphological diversity and evolution. *J. Morphol. 261: 52–69.*

Clift L.E., Andrlikova P., Frolíková M. et al. 2009: Absence of spermatozoal CD46 protein expression and associated rapid acrosome reaction rate in striped field mice (*Apodemus agrarius*). *Reprod. Biol. Endocrinol. 7: 29.*

Cox J., Hein M.Y., Luber C.A. et al. 2014: Accurate proteome-wide label-free quantification by delayed normalisation and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics 13: 2513–2526.*

Crawley M.J. 2007: The R book. *Wiley Publishing, Chechester, UK.*

Dixson A.F. 1993: Sexual selection, sperm competition and the evolution of sperm length. *Folia Primatol. 61: 221–227.*

Dixson A.F. & Anderson M.J. 2004: Sperm behavior, reproductive physiology and sperm competition an male mammals. *Physiol. Behav. 83: 361–371.*

Froman D.P. & Feltmann A.J. 1998: Sperm mobility: a quantitative trait of the domestic fowl (*Gallus domesticus*). *Biol. Reprod. 58: 379–384.*

Garlovsky M.D., Sandler J.A. & Karr T.L. 2022: Functional diversity and evolution of the *Drosophila* sperm proteome. *Mol. Cell. Proteomics 21: 100281.*

Gomendio M., Martin-Coello J., Crespo C. et al. 2006: Sperm competition enhances functional capacity of mammalian spermatozoa. *Proc. Natl. Acad. Sci. U. S. A. 103: 15113–15117.*

Guyonnet B., Zabet-Moghaddam M., SanFrancisco S. & Cornwall G.A. 2012: Isolation and proteomic characterisation of the mouse sperm acrosomal matrix. *Mol. Cell. Proteomics 11: 758–774.*

Hook K.A., Wilke L.M. & Fisher H.S. 2021: Apical sperm hook morphology is linked to sperm swimming performance and sperm aggregation in *Peromyscus* mice. *Cells 10: 2279.*

Humphreys P.N. 1972: Brief observations on the semen and spermatozoa of certain passerine and non-passerine birds. *J. Reprod. Fertil. 29: 327–336.*

Immler S. 2008: Sperm competition and sperm cooperation: the potential role of diploid and haploid expression. *Reproduction 135: 275–283.*

Immler S., Calhim S. & Birkhead T.R. 2008: Increased postcopulatory sexual selection reduces the intramale variation in sperm design. *Evolution 62: 1538–1543.*

Immler S., Moore H.D., Breed W.D. & Birkhead T.R. 2007: By hook or by crook? Morphology, competition and cooperation in rodent sperm. *PLOS ONE 2: e170.*

Jamieson B.G.M. 1987: The ultrastructure and phylogeny of insect spermatozoa. *Cambridge University Press, Cambridge, UK.*

Jamieson B.G.M. 2007: Avian spermatozoa: structure and phylogeny. In: Jamieson B.G.M. (ed.), Reproductive biology and phylogeny of birds. *Science Publisher, Enfield, UK: 348–511.*

Jetz W., Thomas G.H., Joy J.B. et al. 2012: The global diversity of birds in space and time. *Nature 491: 444–448.*

Johnson P.M., Clift L.E., Andrlíková P. et al. 2007: Rapid sperm acrosome reaction in the absence of acrosomal CD46 expression in promiscuous field mice (*Apodemus*). *Reproduction 134: 739–747.*

Kahrl A.F., Snook R.R. & Fitzpatrick J.L. 2021: Fertilisation mode drives sperm length evolution across the animal tree of life. *Nat. Ecol. Evol. 5: 1153–1164.*

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: 2466–2673.*

Kleven O., Laskemoen T., Fossøy F. et al. 2008: Intraspecific variation in sperm length is negatively related to sperm competition in passerine birds. *Evolution 62: 494–499.*

Knief U., Forstmeier W., Pei Y. et al. 2017: A sex-chromosome inversion causes strong

overdominance for sperm traits that affect siring success. *Nat. Ecol. Evol. 1: 1177–1184.*

Krisfalusi M., Miki K., Magyar P.L. & O´Brien D.A. 2006: Multiple glycolytic enzymes are tightly bound to the fibrous sheath of mouse spermatozoa. *Biol. Reprod. 75: 270–278.*

Kuntova B., Stopková R. & Stopka P. 2018: Transcriptomic and proteomic profiling revealed high proportions of odorant binding and antimicrobial defense proteins in olfactory tissues of the house mouse. *Front. Genet. 9: 26.*

Li Y., Sun Y., Ni A. et al. 2020: Seminal plasma proteome as an indicator of sperm dysfunction and low sperm motility in Chickens. *Mol. Cell. Proteomics 19: 1035–1046.*

Lifjeld J.T., Gohli J., Albrecht T. et al. 2019: Evolution of female promiscuity in Passerides songbirds. *BMC Evol. Biol. 19: 169.*

Lüpold S., Calhim S., Immler S. & Birkhead T.R. 2009: Sperm morphology and sperm velocity in passerine birds. *Proc. R. Soc. B Biol. Sci. 276: 1175–1181.*

McGraw L.A., Suarez S.S. & Wolfner M.F. 2015: On a matter of seminal importance. *BioEssays 37: 142–147.*

Mendonca T., Birkhead T.R., Cadby A.J. et al. 2018: A trade-off between thickness and length in the zebra finch sperm mid-piece. *Proc. R. Soc. B Biol. Sci. 285: 20180865.*

Otčenášková T., Macíčková E., Vondráková J. et al. 2023: Proteomic analysis of the mouse sperm acrosome – towards an understanding of an organelle with diverse functionality. *Eur. J. Cell Biol. 102: 151296.*

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

Parker G.A. 1992: Snakes and female sexuality. *Evol. Biol. 355: 395–396.*

Perry J., Sirot L. & Wigby S. 2013: The seminal symphony: how to compose an ejaculate. *Trends Ecol. Evol. 28: 414–422.*

Pitnick S., Hosken D.J. & Birkhead T.R. 2009: Sperm morphological diversity. In: Pitnick S., Hosken D.J. & Birkhead T.R. (eds.), Sperm biology: an evolutionary perspective. *Academic Press, London, UK: 247–394.*

Pizzari T. & Parker G.A. 2009: Sperm biology: an evolutionary perspective. In: Pizzari T. & Parker G.A. (eds.), Sperm competition and sperm phenotype. *Academic Press, Lodnon, UK: 207–245.*

Poignet M., Baránková L., Reif J. et al. 2022: Sperm morphology and performance in relation to postmating prezygotic isolation in two recently diverged passerine species. *Sci. Rep. 12: 22275.*

Rohart F., Gautier B., Singh A. & Lê Cao K.-A. 2017: mixOmics: an R package for 'omics feature selection and multiple data integration. *PLOS Comput. Biol. 13: e1005752.*

Rowe M., Albrecht T., Cramer E.R. et al. 2015: Postcopulatory sexual selection is associated with accelerated evolution of sperm morphology. *Evolution 69: 1044–1052.*

Rowe M., Laskemoen T., Johnsen A. & Lifjeld J.T. 2013: Evolution of sperm structure and energetics in passerine birds. *Proc. R. Soc. B Biol. Sci. 280: 20122616.*

Rowe M., Skerget S., Rosenow M.A. & Karr T.L. 2019: Identification and characterisation of the zebra finch (*Taeniopygia guttata*) sperm proteome. *J. Proteomics 193: 192–204.*

Rowe M., Whittington E., Borziak K. et al. 2020: Molecular diversification of the seminal fluid proteome in a recently diverged passerine species pair. *Mol. Biol. Evol. 37: 488–506.*

Sandera M., Albrecht T. & Stopka P. 2013: Variation in apical hook length reflects the intensity of sperm competition in Murine rodents. *PLOS ONE 8: e68427.*

Sexton T.J. 1974: Oxidative and glycolytic activity of chicken and turkey spermatozoa. *Comp. Biochem. Physiol. 48: 59–65.*

Simmons L.W. & Fitzpatrick J.L. 2012: Sperm wars and the evolution of male fertility. *Reproduction 144: 519–534.*

Słowińska M., Nynca J., Arnold G.J. et al. 2017: Proteomic identification of turkey (*Meleagris gallopavo*) seminal plasma proteins. *Poult. Sci. 96: 3422–3435.*

Soler L., Labas V., Thélie A. et al. 2016: Intact cell MALDI-TOF MS on sperm: a molecular test for male fertility diagnosis. *Mol. Cell. Proteomics 15: 1998–2010.*

Stopka P., Kuntova B., Klempt P. et al. 2016: On the saliva proteome of the Eastern European house mouse (*Mus musculus musculus*) focusing on sexual signalling and immunity. *Sci. Rep. 6: 32481.*

Støstad H.N., Johnsen A., Lifjeld J.T. & Rowe M. 2018: Sperm head morphology is associated with sperm swimming speed: a comparative study of songbirds using electron microscopy. *Evolution 72: 1918–1932.*

Vicens A., Borziak K., Karr T.L. et al. 2017: Comparative sperm proteomics in mouse species with divergent mating systems. *Mol. Biol. Evol. 34: 1403–1416.*

Whittington E., Forsythe D., Borziak K. et al. 2017: Contrasting patterns of evolutionary constraint and novelty revealed by comparative sperm proteomic analysis in Lepidoptera. *BMC Genomics 18: 931.*

Whittington E., Karr T.L., Mongue A.J. et al. 2019: Evolutionary proteomics reveals distinct patterns of complexity and divergence between Lepidopteran sperm morphs. *Genome Biol. Evol. 11: 1838–1846.*

Wickham H. 2016: ggplot2: elegant graphics for data analysis. *Springer-Verlag, New York, USA.*

Wolfson A. 1952: The cloacal protuberance: a means for determining breeding condition in live male Passerines. *Bird-Banding 23: 159–165.*

Yu G., Wang L.G., Han Y. & He Q.Y. 2012: clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics 16: 284–287*

Supplementary online material

Table S1. Proteomes of sperm and seminal fluids. The table contains gene names, signal intensities, correlation coefficients and protein annotations (https://www.ivb.cz/wp-content/uploads/JVB-vol.-72-2023-Otcenaskova-T.-et-al.-Table-S1.xlsx). Complete dataset is available via ProteomeXchange with identifier PXD042116.