Supplement for

Sperm swimming speed and morphology differ slightly among the three genetic morphs in the ruff sandpiper (*Calidris pugnax*)

by Martin Bulla, Clemens Küpper, David B Lank, Jana Albrechtová, Jasmine L Loveland, Katrin Martin, Kim Teltscher, Margherita Cragnolini, Michael Lierz, Tomáš Albrecht, Wolfgang Forstmeier & Bart Kempenaers

Methods

S0 – Deviations from the *a priori* protocols

Sperm competition

We initially based our predictions about between-morph differences in sperm traits on the idea that the intensity of sperm competition might be stronger in Faeders and might thus have selected for more competitive sperm in this morph. However, we realized, that in ruffs the likelihood that beneficial mutations occur is several orders of magnitude lower in Faeders than in Independents. We now discuss this in the manuscript and also refrain from predictions about better sperm as what characterizes high quality sperm in birds is debated (see references in the main text).

Vas deference samples

We have not included the morphology data based on the vas deference samples, because (a) we did not have corresponding sperm velocity measurements, (b) we were unsure whether those sperm were fully developed and (c) we used a different method (ImageJ) to measure sperm morphology than the semi-automated Sperm sizer method that was developed later and that we used for the sperm collected by abdominal massage, i.e. the measures are not directly comparable.

Velocity measurements

We aimed at using curvilinear velocity as a measure of sperm speed, but report all three velocities provided by the sperm tracking software, for reasons described in section S2 below. We suggest that straight-line velocity is a better velocity measure in ruffs.

Coefficient of variation

Our main goal was to investigate how the within-male coefficient of variation in sperm length differs between the morphs. We did not use the between-male coefficient of variation per morph because of the unbalanced number of males per morph within our sample.

S1 - Housing

Individuals were housed in groups of varying size and different morph and sex composition (Table S0), in two outdoor aviaries (123 m^2 and 119 m^2) and in a complex of semi-outdoor aviaries divided into a large central space (117 m^2) and 24 smaller, adjacent aviaries (8.8 m^2 each). One aviary contained only males and in three aviaries male access to females was limited to 3.5h per day. Note that males in the male-only aviaries also produced sperm, and we observed homosexual matings in this setting. The semi-outdoor aviaries have a solid but transparent roof, and wire-mesh on at least one side, so that the birds experienced natural light and temperature cycles. All aviaries had natural grass, a few wooden logs, small elevated areas suitable for lekking, and at least one heatable water body (~1 m^2). Aviaries were cleaned daily and the birds were provided with a mixture of food pellets for waders, dried shrimp and live mealworms (Meghlys; www.meghlys.com) in dishes or spread over the grass.

Table S0 | Distribution of male morphs and their access to females across aviaries

Location	Access to ♀	Aviary	N ♀	N ♂	Independents	Satellites	Faeders
A	always	1	15	3	3		
	always	2	8	3	3		
	always	3	8	3	1	2	
	always	4	15	3	3		
B	3.5h/day	1	10	7	5	2	
	3.5h/day	2	10	10	8	2	
	none	3		8	2	5	1
	always	4	19	5	5		
	always	5	19	5			5
	always	6	19	5		5	
C	always	1-3	21	40	29	9	2

S2 - Velocities

During recording sessions, we observed that ruff sperm swim slower, i.e. cover shorter distances per unit of time, than passerine sperm (video examples). However, curvilinear velocity, our intended measure of velocity and a measure of sperm swimming speed typically used in studies on passerines (Laskemoen *et al.* 2010, Cramer *et al.* 2016, Opatová *et al.* 2016, Tomasek *et al.* 2017, Støstad *et al.* 2018, Schmoll *et al.* 2020), tracks the sperm's sideways "vibration" movements, and suggests that ruff sperm moves faster than passerine sperm (Fig. S1). Thus, in ruffs, the curvilinear velocity method likely does not reflect the actual velocity in terms of "distance covered per unit of time" and hence might not be the most appropriate velocity measure. Straight-line velocity seems more appropriate and reflects our observation that ruff sperm moves forward more slowly than zebra finch sperm. Straight-line velocity has also been linked to fertilization success in avian and non-avian taxa, and has been used in previous studies instead of or along with curvilinear velocity (reviewed in Pizzari *et al.* 2004, Denk *et al.* 2005, Helfenstein *et al.* 2009). Consequently, we report all three velocity measures provided by the sperm tracking software (curvilinear, straight-line, and average-path velocity). The three velocity measures are correlated (r = 0.39 - 0.91, Fig. S2). The within-male seasonal repeatability of sperm velocity, i.e. the percentage of the variation attributed to variation among males, was 24% for straight-line velocity, 35% for average-path velocity and 47% for curvilinear velocity (based on May and June velocity estimates of the same individuals; Fig. S3 and S4, Table S1).

S3 - Relationships between sperm components

In ruffs, sperm head length constitutes on average 23% of the total sperm length (range: 18 – 34%, N = 920 sperm from 92 males), the midpiece 17% (15 – 22%) and the tail 60% (45 – 65%). The head length reflects the length of the nucleus (r = 0.97), which makes up 82 – 96% of the head, while flagellum and total sperm length mainly reflect tail length (r = 0.95 and 0.89 respectively; Fig. S5). Head length did not correlate with midpiece length (r = 0.00) or tail length (r = -0.02), and midpiece length and tail length correlated weakly (r = 0.27; Fig. S5). The within-male within sperm-sample repeatability of sperm morphology measures varied from 25% to 60%, with acrosome length, the smallest part of the sperm, being the least repeatable (Fig. S3, Table S1). We did not measure between-sample repeatability in sperm components as avian sperm is highly heritable (Birkhead *et al.* 2005).

S4 – Inbreeding and Relatedness

Because of our captive population, we expect higher levels of inbreeding compared to males from the wild. Previous studies on birds and mammals have shown that inbred males have lower sperm velocity and a higher proportion of abnormal sperm than outbred males (Gomendio *et al.* 2000, Ala-Honkola *et al.* 2013, Heber *et al.* 2013, Opatová *et al.* 2016). However, there is no evidence that the morphology of normal-looking sperm (e.g., length, coefficient of variation) differs between inbred and outbred males (Mehlis *et al.* 2012, Ala-Honkola *et al.* 2013, Opatová *et al.* 2016). Based on these studies, we assumed that our sperm velocity measurements may be on average somewhat lower than those of free-living ruffs, whereas the morphological measurements likely reflect the variation in sperm morphology observed in the wild. Because the morphs interbreed in our aviary, inbreeding levels between morphs should be similar. To assess whether inbreeding and relatedness between individuals influenced our results, we estimated these traits based on genotypes of all males at 21 polymorphic microsatellite markers (Giraldo-Deck *et al.* 2022).

We quantified inbreeding as homozygosity-by-locus (Aparicio *et al.* 2006) using the 'GENHET' R-function version 3.1 (Coulon 2010) and calculated Pearson's correlation coefficients between sperm traits and homozygosity. Correlations were weak (mean r = -0.09, range: -0.26 to 0.07; Fig. S7), and particularly so for measures of velocity ($r_{curvilinear}$ = -0.04, $r_{straight\ line}$ = 0.07, $r_{average\ path}$ = 0.05), so we did not control the subsequent models for inbreeding.

To investigate whether the main results were confounded by relatedness between some of the individuals, we tested for a relatedness signal in the model residuals. We specified the model residuals as a new response variable in an intercept-only Bayesian linear regression fitted in STAN (Stan-Development-Team 2022) using the 'brm' function from the 'brms' R-package (Bürkner 2017, Bürkner 2018, Bürkner 2021) with the male relatedness matrix as a random effect. We constructed the relatedness matrix from the genotypes at 21 microsatellite loci using the 'coancestry' function from the 'related' R-package (Pew *et al.* 2015) and its "lynchli estimator", which correlated best with the expected values (although by a small margin). As the software sometimes estimates negative values for zero relatedness, negative relatedness values were assigned as zero. To make the relatedness matrix positive definite we added 0.1 to its diagonal. We used the default 'brm' priors, i.e. a flat prior for the intercept and a Student's t distribution for the standard deviation (Bürkner 2017, Bürkner 2018, Bürkner 2021). To decrease the possibility of divergent transitions threatening the validity of posterior samples, the target average proposal acceptance probability was increased to 0.99 or to 0.999 for the coefficients of variation (Bürkner 2017, Bürkner 2018, Bürkner 2021). Four Markov chains ran for 50,000 iterations each. For each chain, we discarded the first 25,000 iterations and sampled every 20th iteration, which resulted in a total of 5,000 samples (4 x 1,250) of model parameters.

We assessed the independence of samples in the Markov chain using graphic diagnostics and the convergence using the Gelman-Rubin diagnostics, which was 1 for all parameters, indicating model convergence (Brooks & Gelman 1998).

The relatedness matrix explained little variation in the residuals of models on swimming speed (10% for curvilinear, 5% for straight-line and 8% for average-path velocity) and close to zero variation in the residuals of all models on morphology (0.1%; Table S3). The alternative models using all velocity recordings or using individual sperm measurements gave similar results (Table S3). Importantly, the models without the relatedness matrix (i.e. intercept only model) fitted the residuals better than the models with the relatedness matrix; the estimated Bayes factor in favour of the model without the relatedness matrix ranged from seven to infinity (Lee & Wagenmakers 2014), and the posterior probability from 0.88 to 1 (mean = 0.97; Table S3). Thus, in the main text we report results from models without control for relatedness. In contrast, the relatedness matrix explained variable amounts of variation in the residuals from models on coefficients of variation in sperm morphology (7-77%, Table S3). As including the relatedness matrix did not change the conclusions (Fig. S8) and for consistency, we did not control for relatedness in the models described in the main text.

Figures

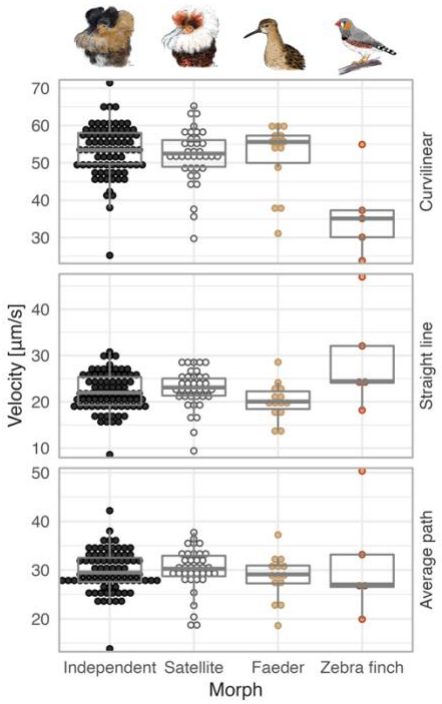

Figure S1 | Comparison of sperm swimming speed between ruff morphs and zebra finch. Dots represent velocity values for sperm of 4 ruff and 5 zebra finch males recorded in May, 46 ruff males recorded in June, and two values for 42 ruff males recorded in May and June. Boxplots depict median (horizontal line inside the box), the 25th and 75th percentiles (box) and the 25th and 75th percentiles ±1.5 times the interquartile range or the minimum/maximum value, whichever is smaller (bars). Five zebra finch males from a population at the Max Planck Institute for Biological Intelligence in Seewiesen were sampled in June along with the ruffs to ensure that the ruff motilities and velocities are not an artefact of our sampling method. The zebra finch sperm moved normally (see example), with velocity values well within the range of measurements from our laboratory (Opatová *et al.* 2016, Knief *et al.* 2017). Note that the zebra finches were sampled outside a breeding phase (no nest boxes available). Created with 'ggplot' function and dots stacked using 'geom_dotplot' function, both from the 'ggplo2' R-package (Wickham 2016). Illustrations by Yifan Pei under Creative Commons Attribution (CC BY 4.0).

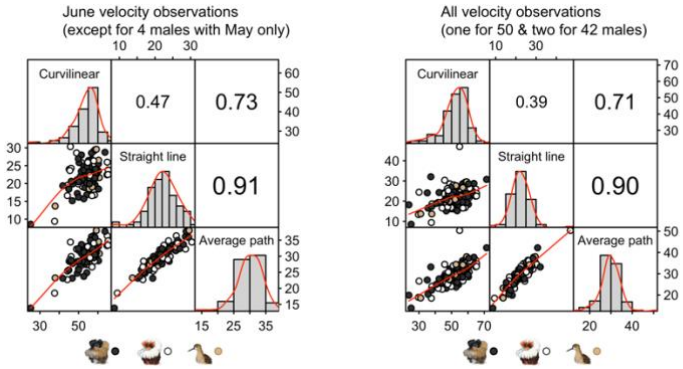

Figure S2 | Pairwise correlations among sperm velocity measures of ruffs. On the diagonal: histograms and density lines (red) for each variable. Above diagonal: Pearson's correlation coefficients with size highlighting the strength of the correlation. Below diagonal: bivariate scatterplots, with each dot representing June value per male (left), except for four males with May only values (N = 92), or all velocity observations (right; N = 134), dot color highlighting morph (black: Independents, white: Satellites, beige: Faeders) and red line representing loess-smoothed fit. Adapted from 'pairs.panels' function from 'psych' R-package (Revelle 2022). Ruff morph illustrations by Yifan Pei under Creative Commons Attribution (CC BY 4.0).

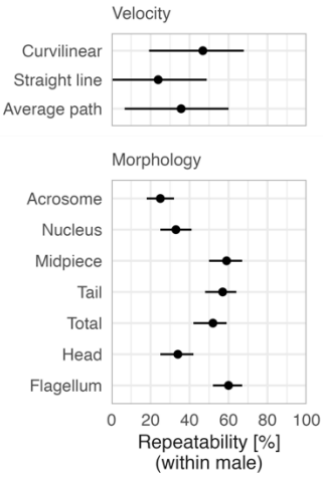

Figure S3 | Within-male repeatability of sperm traits in ruffs. Dots with bars represent repeatability estimates with 95%CI generated by the 'rpt' function from the 'rptR' R-package, without specifying fixed effects (Stoffel *et al.* 2017). For sperm velocity we used one measurement from May and one from June (N = 42 males). For sperm morphology, we used measurements of 10 sperm per male (N = 92 males). The last three morphological traits are composite traits (Total = Acrosome + Nucleus + Midpiece + Tail, Head = Acrosome + Nucleus, Flagellum = Midpiece + Tail). For precise estimates see Table S1.

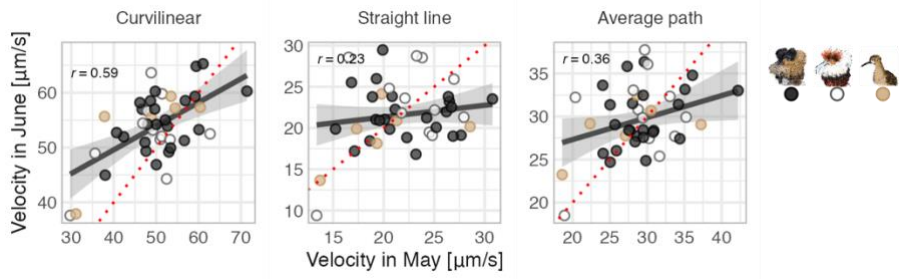

Figure S4 | Correlation between May and June sperm velocity of ruffs. Dots represent single males, dot color the morph (black: Independents, white: Satellites, beige: Faeders), lines with shaded areas linear model fits with 95%CIs generated by 'stat_smooth' function in 'ggplot2' R-package (Wickham 2016) using robust regression specified by 'rlm' function from 'MASS' R-package (Venables & Ripley 2002). '*r*' represents Pearson's correlation coefficient and dotted lines indicate equality, i.e. points above the line represent faster sperm in June, points below the line faster sperm in May. N = 42 males with velocity measured both in May and June (24 Independents, 12 Satellites, 6 Faeders).

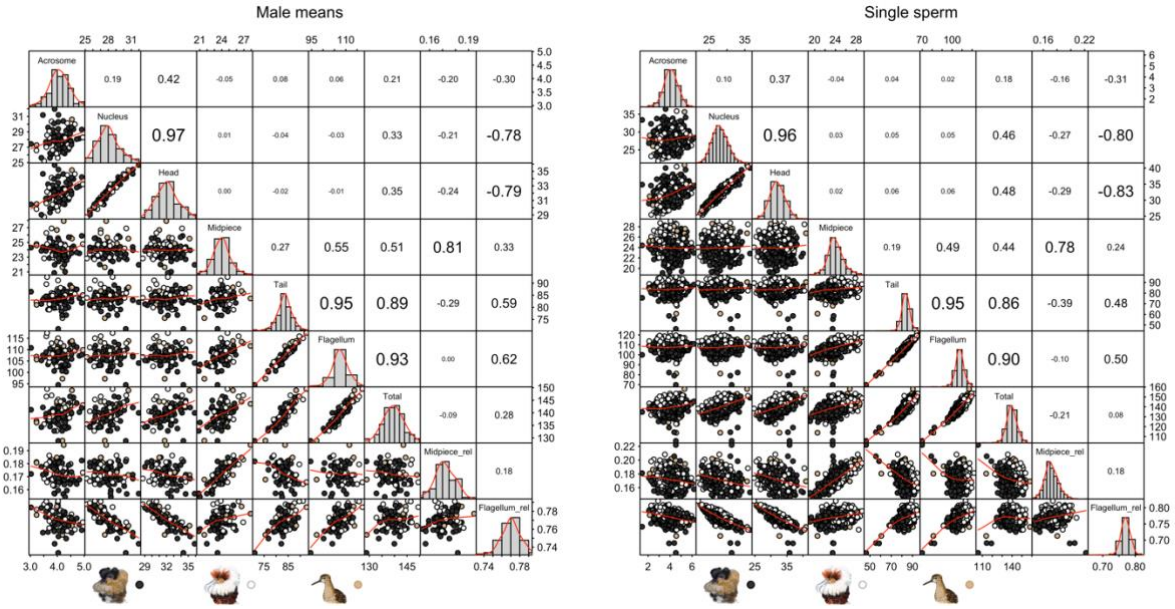

Figure S5 | Pairwise correlations among sperm morphological traits of ruffs. On the diagonal: histograms and density lines (red) for each variable. Above diagonal: Pearson's correlation coefficients with size highlighting the strength of the correlation. Below diagonal: bivariate scatterplots, with each dot representing average value per male (left; N = 92) or a single sperm value (right; N = 920), dot color highlighting morph (black: Independents, white: Satellites, beige: Faeders) and red line representing loess-smoothed fit. Adapted from 'pairs.panels' function from 'psych' R-package (Revelle 2022). Ruff morph illustrations by Yifan Pei under Creative Commons Attribution (CC BY 4.0).

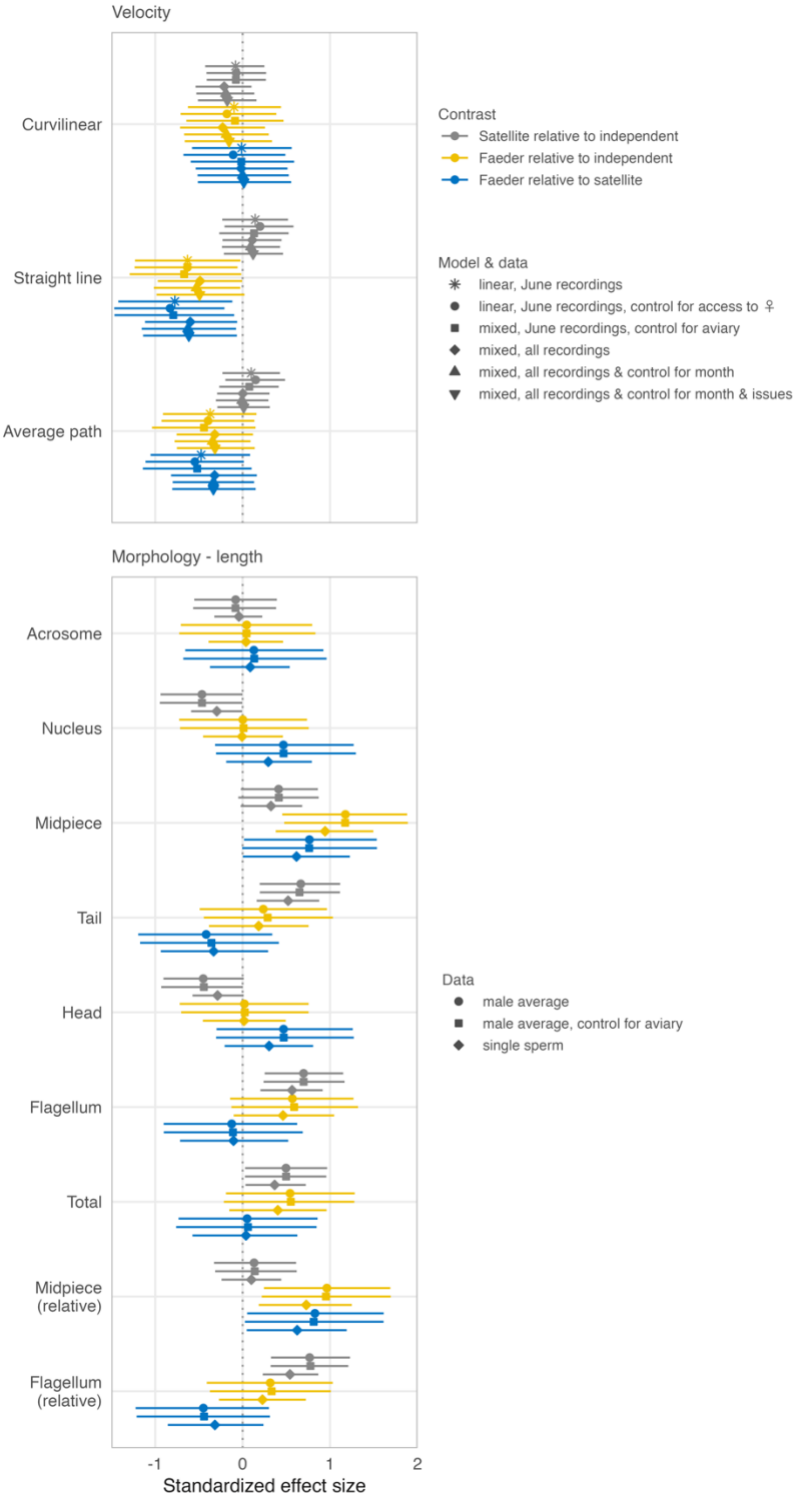

Figure S6 | Differences in sperm traits of ruff morphs according to model and data type. Shapes with bars represent estimated standardized effect sizes (medians) with their 95%CIs based on the joint posterior distribution of 5,000 simulated values generated from models by the 'sim' function from the 'arm' R-package (Gelman & Su 2021). Color highlights the between-morph differences – Satellite relative to Independent (grey), Faeder relative to Independent (yellow), Faeder relative to Satellite (blue) – and shape indicates type of model and data. For **velocity**, '*linear, June recordings*' indicates results based on a linear model fitted to the June velocity values of all males, except for four males for which only recordings from May were available; '*mixed, June recordings, control for aviary*' indicates results based on the same data but from a mixed-effect model that includes 'aviary' as a random intercept; '*mixed, all recordings*' indicates results based on a mixed-effect model fitted to all velocity values (including 42 males with a recording for both May and June) with male identity included as a random intercept. To control for '*access to ♀*' (none, temporal, continuous), '*month*' (May or June) and/or '*issues*' ((yes or no; indicating e.g. presence of faeces) in some models we included these terms as fixed effects. For **morphology**, '*male average*' indicates results based on linear models fitted to the average morphological value per male based on 10 sperm cells; '*male average, control for aviary*' indicates results based on the same data but from mixed-effect models containing aviary as a random intercept; '*single sperm*' indicates results based on mixed-effect models fitted to single sperm-cell measurements (10/male) and controlled for multiple sampling per male by including male identity as a random intercept. For both, velocity and morphology, the models from the main text Fig. 2 are listed first and in the above legends indicted as 'linear, June recordings' and 'male average'. Note that the main text estimates are similar to those from the alternative models.

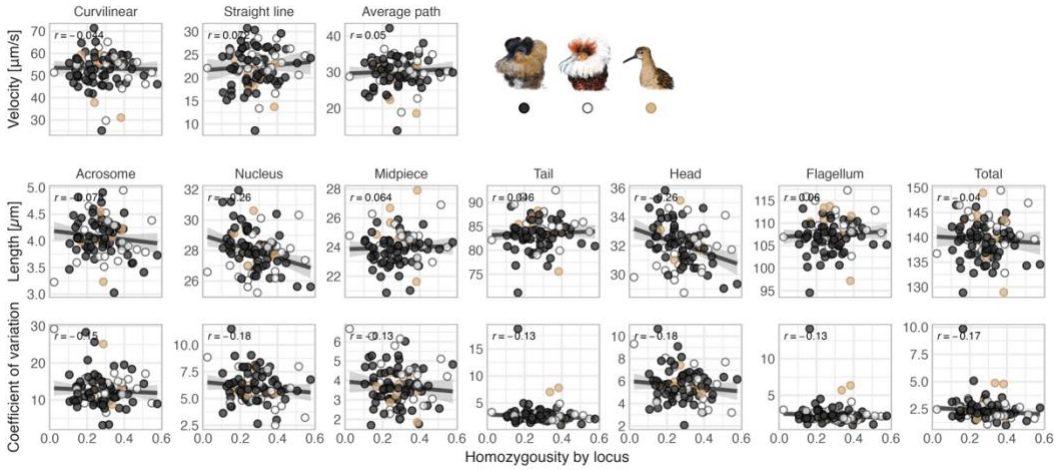

Figure S7 | Correlation between sperm traits and homozygosity by locus. Dots represent single male velocity values from June (except for four males with May only values) or male average lengths and coefficients of variation from 10 sperm cells (N = 92 males). Dot color indicates morph: black – Independent, white – Satellite, beige – Faeder. Lines with shaded area represent linear model fit with 95%CI generated by 'stat_smooth' function in 'ggplot2' R-package (Wickham 2016) using robust regression specified by 'rlm' function from 'MASS' R-package (Venables & Ripley 2002). '*r*' represents Pearson's correlation coefficient. Ruff morph illustrations by Yifan Pei under Creative Commons Attribution (CC BY 4.0).

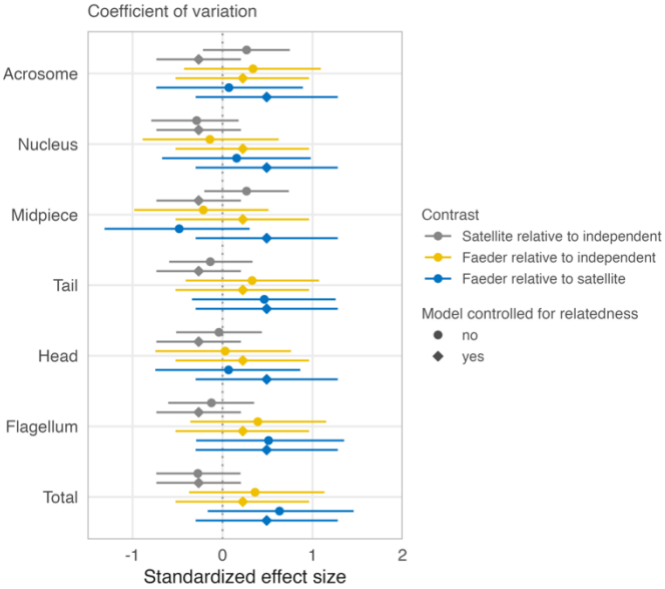

Figure S8 | Ruff morph differences in coefficients of variation of sperm traits with and without control for relatedness. Shapes with bars represent estimated standardized effect sizes with their 95%CIs, color the between-morph differences – Satellite relative to Independent (grey), Faeder relative to Independent (yellow), Faeder relative to Satellite (blue) – and shape indicates whether model was controlled for relatedness (diamond) or not (dot). Note that the estimates reported in the main text (Fig. 2, here indicated by dots) are similar to those from models controlled for relatedness (diamonds). Estimates (medians) and 95%CIs are based on the joint posterior distribution of 5,000 simulated values generated by the 'sim' function from the 'arm' R-package (Gelman & Su 2021) using the model outputs from Table S1 (dots) or generated by 'brm' function in 'brms' R-package (Bürkner 2017, Bürkner 2018, Bürkner 2021) with a vague (weakly informative) Gaussian priors centered on zero for the intercept and factor levels and half Cauchy priors for the standard deviations, including the error term - standard deviation of the residuals (diamonds). Using default 'brms' priors generated same results. For further details see Methods S1, for comparison of models with and without control for relatedness Table S3.

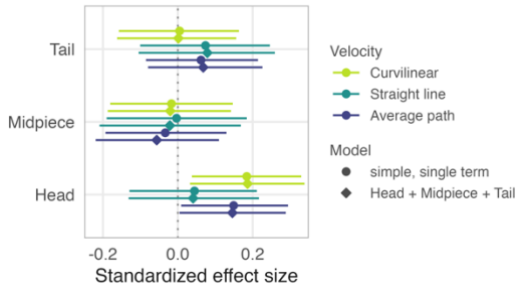

Figure S9 | Estimates from simple univariate models are similar to those from a model containing multiple traits. Shapes with bars represent estimated standardized effect sizes with their 95%CIs based on the joint posterior distribution of 5,000 simulated values generated from models by the 'sim' function from the 'arm' R-package (Gelman & Su 2021). Color indicates dependent variable, i.e type of velocity (green: curvilinear, turquois: straight line, purple: average path). Shape indicates source of estimate: from 'univariate' models with single morphological terms (head, midpiece or tail; dot) or a multivariate model containing all three terms (diamond). The models were controlled for number of tracked sperm (ln-transformed) and morph. The response (velocity) as well as the linear term/s (number of tracked sperm, head, midpiece and tail) were scaled (mean-centered and divided by standard deviation).

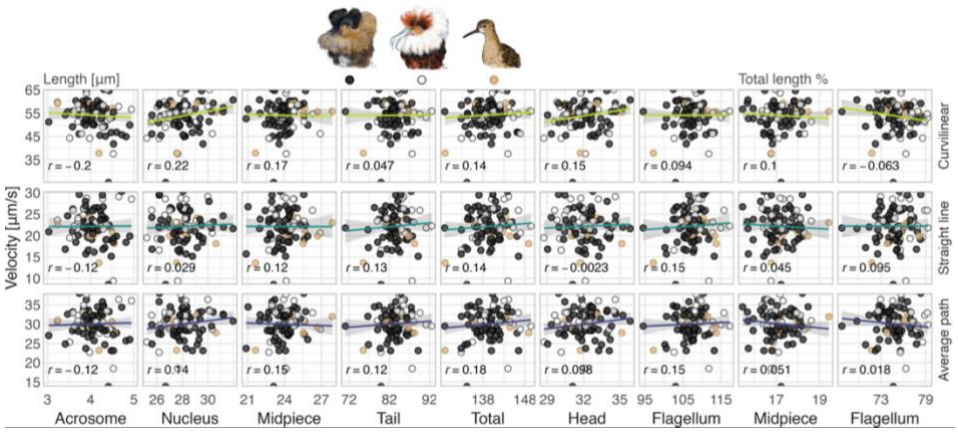

Figure S10 | Sperm swimming speed in relation to sperm morphology. Lines with shaded areas represent model predictions with their 95%CIs based on the joint posterior distribution of 5,000 predicted values generated from linear models, controlled for number of tracked sperm (ln-transformed) and morph (Table S6). '*r*' represents Pearson's correlation coefficient (ignoring morph, hence *r* may differ from model estimates). Dots represent data points based on single June-values for velocity (with exception of four males with May-values only) and average trait lengths of 10 sperm cells per male, dot color highlights morph. Ruff-morph illustrations by Yifan Pei under Creative Commons Attribution (CC BY 4.0).

Tables

Table S1 | Within male repeatability of ruff sperm traits

Trait	Specification	Estim	95%CI
Velocity	Curvilinear	4	19-6
	Straight line	2	0-4
	Average path	3	7-6
Length	Acrosome	2	18-3
	Nucleus	3	25-4
	Midpiece	5	50-6
	Tail	5	48-6
	Head	3	25-4
	Flagellum	6	53-6
	Total	5	42-5

Repeatability estimates with 95%CI generated by 'rpt' function from 'rptR' R-package (Stoffel *et al.* 2017) from two - May and June – velocity estimates per male (N = 42 males) and 10 sperm morphology measurements per male (N = 92 males).

Table S2 | Within and between observer repeatability of ruff sperm measurements

	Repeatability estimate (95%CI)	
Trait	**Within observer**	**Between observer**
Acrosome	92% (85.2 - 95.6)	90% (80.8 - 94.5)
Nucleus	98% (97.1 - 99.1)	98% (95.6 - 98.7)
Midpiece	99% (98.5 - 99.6)	99% (98.7 - 99.6)
Tail	97% (94.3 - 98.5)	99% (97.3 - 99.2)
Total	100% (99.4 - 99.8	99% (98.1 - 99.4)
Head	98% (95.6 - 98.8)	97% (94.3 - 98.4)
Flagellum	99% (97.5 - 99.3)	98% (96.4 - 98.9)

Repeatability estimates with 95%CI generated by the 'rpt' function from 'rptR' R-package (Stoffel *et al.* 2017) from 40 sperm measured twice by the same observer (within observer) or by two different observers (between observer). Note that the smallest part – acrosome – has the lowest repeatability, but the repeatability of head, the composite measure of acrosome and nucleus, is high.

Table S3 | Percentage of variation in 'residuals from original models' explained by relatedness and comparison of models without and with control for relatedness.

Data	Trait	Specification	Variance explained by relatedness	95% CI		Bays factor	Probability of model without relatedness matrix
June	Velocity	Curvilinear	10.1%	0.0%	44.7%	7	0.875
		Straight line	5.1%	0.0%	34.7%	10	0.913
		Average path	8.0%	0.0%	38.8%	8	0.886
Averages	Length	Acrosome	0.1%	0.0%	0.7%	Inf	1
		Nucleus	0.1%	0.0%	0.7%	Inf	1
		Midpiece	0.1%	0.0%	0.6%	Inf	1
		Tail	0.1%	0.0%	0.6%	Inf	1
		Total	0.1%	0.0%	0.6%	Inf	1
		Head	0.1%	0.0%	0.7%	Inf	1
		Flagellum	0.1%	0.0%	0.6%	Inf	1
Male value	Coefficient of variation	Acrosome	33.9%	0.3%	77.1%	1	0.583
		Nucleus	16.8%	0.1%	58.1%	3	0.767
		Midpiece	6.6%	0.0%	43.7%	7	0.874
		Tail	76.8%	24.0%	94.3%	0	0.1
		Total	47.8%	0.5%	85.6%	1	0.528
		Head	20.3%	0.1%	60.0%	2	0.724
		Flagellum	69.9%	3.4%	94.9%	0	0.278
All	Velocity	Curvilinear	1.8%	0.0%	15.6%	22	0.951
		Straight line	1.9%	0.0%	16.9%	15	0.937
		Average path	2.4%	0.0%	17.4%	16	0.948
All	Length	Acrosome	0.1%	0.0%	0.7%	84	0.989
		Nucleus	0.1%	0.0%	0.7%	94	0.99
		Midpiece	0.1%	0.0%	0.6%	132	0.992
		Tail	0.1%	0.0%	0.6%	118	0.991
		Total	0.1%	0.0%	0.6%	113	0.992
		Head	0.1%	0.0%	0.7%	90	0.989
		Flagellum	0.1%	0.0%	0.6%	143	0.993

Variance explained by relatedness with **95%CI** represents percentage of variance explained by relatedness matrix in an intercept only model fitted to residuals of the original models (Fig. 2, Table S5) in STAN (Stan-Development-Team 2022) using 'brm' function from 'brms' R-package (Bürkner 2017, Bürkner 2018, Bürkner 2021) with male relatedness matrix and male identification (in case of residuals from models on single values) as random effect. **Bayes factor** in favor of model without relatedness matrix and **probability of model without relatedness matrix** in comparison to a model with relatedness matrix. The Gelman-Rubin diagnostics was 1 for all models, indicating model convergence (Brooks & Gelman 1998). Note that for all cases, but coefficient of variation in Tail and Flagellum, the model without relatedness matrix fits residuals better than a model with relatedness matrix, which justifies our use of simple original models, not controlled for relatedness (Fig. 2). Importantly, controlling the original models on coefficient of variation for relatedness generated similar results (Fig S8).

Table S4 | AICc comparison of simple and quadratic effects of morphological traits on sperm velocity

Velocity	Trait	AICc of the model		ΔAICc	Akaike weight	Evidence ration
		Simple	Quadratic			
Curvilinear	Acrosome	208.44	210.59	2.15	0.03	32.33
	Nucleus	200.41	202.71	2.30	0.03	32.33
	Midpiece	209.02	209.89	0.87	0.06	15.67
	Tail	209.07	211.28	2.21	0.03	32.33
	Total	208.23	209.84	1.61	0.04	24
	Head	202.81	205.07	2.26	0.03	32.33
	Flagellum	209.07	211.33	2.26	0.03	32.33
	Midpiece relative	208.41	210.50	2.09	0.03	32.33
	Flagellum relative	204.84	206.93	2.09	0.03	32.33
Straight line	Acrosome	229.05	230.93	1.88	0.04	24
	Nucleus	228.75	231.05	2.30	0.03	32.33
	Midpiece	229.06	231.05	1.99	0.04	24
	Tail	228.30	230.55	2.25	0.03	32.33
	Total	228.25	228.88	0.63	0.07	13.29
	Head	228.77	231.05	2.28	0.03	32.33
	Flagellum	228.47	230.33	1.86	0.04	24
	Midpiece relative	228.67	230.60	1.93	0.04	24
	Flagellum relative	229.05	231.29	2.24	0.03	32.33
Average path	Acrosome	202.34	204.64	2.30	0.03	32.33
	Nucleus	198.15	200.49	2.34	0.03	32.33
	Midpiece	202.38	204.40	2.02	0.04	24
	Tail	201.80	204.14	2.34	0.03	32.33
	Total	200.74	201.93	1.19	0.05	19
	Head	198.34	200.68	2.34	0.03	32.33
	Flagellum	202.14	204.24	2.10	0.03	32.33
	Midpiece relative	200.68	201.97	1.29	0.05	19
	Flagellum relative	200.90	203.24	2.34	0.03	32.33

Each model was controlled for number of tracked sperm (ln-transformed) and morph. The response (velocity), number of tracked sperm and linear morphological term (in case of simple models) were scaled (mean-centered and divided by standard deviation). **Simple** represents AICc value for a model with a linear morphological term (scaled) and **Quadratic** AICc value for a model with a linear and quadratic morphological term (second polynomial). **ΔAICc** - the difference in AICc between the quadratic and simple model (i.e. positive values indicate a poorer fit of the quadratic model). **Akaike weight** - the weight of evidence that a quadratic model is the best approximating model, i.e. probability of the quadratic model. Note that the probability of the Simple model (with linear term only) is 1 – *wi* and hence 0.93-97. **Evidence ration** - the model weight of the simple model relative to the quadratic model, i.e. how many times is the simple model more likely than the quadratic model. All parameters confirm superiority of the simple models.

Table S5 | **Differences in sperm traits of ruff morphs**

		Estimate (95%CI)		
Response	**Predictor | Response:**	**Curvilinear**	**Straight line**	**Average path**
Velocity	Intercept (Independent)	0.03 (-0.16 - 0.22	0.01 (-0.19 - 0.22	0.00 (-0.17 - 0.18
	Count of tracked sperm	0.71 (0.56 - 0.86	0.60 (0.43 - 0.77	0.73 (0.59 - 0.87
	Morph (Satellite)	-0.09 (-0.43 - 0.26	0.15 (-0.24 - 0.51	0.09 (-0.23 - 0.41
	Morph (Faeder)	-0.10 (-0.63 - 0.47	-0.62 (-1.24 --0.04	-0.37 (-0.87 - 0.15
		Length	**Coefficient of variat**	
Acrosome	Intercept (Independent)	0.01 (-0.24 -0.27)	-0.10 (-0.36 -0.16)	
	Morph (Satellite)	-0.07 (-0.55 -0.40)	0.26 (-0.22 -0.74)	
	Morph (Faeder)	0.06 (-0.70 -0.80)	0.33 (-0.38 -1.08)	
Nucleus	Intercept (Independent)	0.12 (-0.13 -0.38)	0.09 (-0.16 -0.35)	
	Morph (Satellite)	-0.46 (-0.92 -0.01)	-0.29 (-0.76 -0.18)	
	Morph (Faeder)	-0.01 (-0.74 -0.75)	-0.13 (-0.89 -0.61)	
Midpiece	Intercept (Independent)	-0.21 (-0.46 -0.03)	-0.06 (-0.31 -0.21)	
	Morph (Satellite)	0.41 (-0.03 -0.86)	0.26 (-0.21 -0.73)	
	Morph (Faeder)	1.19 (0.46 -1.88)	-0.22 (-0.97 -0.55)	
Tail	Intercept (Independent)	-0.20 (-0.45 -0.04)	0.01 (-0.25 -0.27)	
	Morph (Satellite)	0.66 (0.20 -1.13)	-0.14 (-0.61 -0.33)	
	Morph (Faeder)	0.22 (-0.50 -0.96)	0.32 (-0.43 -1.06)	
Total	Intercept (Independent)	-0.18 (-0.43 -0.07)	0.04 (-0.22 -0.29)	
	Morph (Satellite)	0.49 (0.03 -0.96)	-0.26 (-0.74 -0.21)	
	Morph (Faeder)	0.54 (-0.22 -1.27)	0.35 (-0.40 -1.11)	
Head	Intercept (Independent)	0.12 (-0.13 -0.36)	0.01 (-0.26 -0.27)	
	Morph (Satellite)	-0.44 (-0.90 -0.02)	-0.04 (-0.52 -0.44)	
	Morph (Faeder)	0.02 (-0.72 -0.78)	0.04 (-0.73 -0.80)	
Flagellum	Intercept (Independent)	-0.24 (-0.49 -0.01)	0.00 (-0.27 -0.26)	
	Morph (Satellite)	0.70 (0.27 -1.15)	-0.12 (-0.60 -0.35)	
	Morph (Faeder)	0.57 (-0.15 -1.29)	0.40 (-0.35 -1.14)	
Midpiece	Intercept (Independent)	-0.12 (-0.38 -0.14)		
(relative)	Morph (Satellite)	0.14 (-0.32 -0.59)		
	Morph (Faeder)	0.97 (0.25 -1.71)		
Flagellum	Intercept (Independent)	-0.24 (-0.48 -0.01)		
(relative)	Morph (Satellite	0.77 (0.32 -1.22)		
	Morph (Faeder)	0.31 (-0.43 -1.02)		

The posterior estimates (medians) of the effect sizes with the 95% credible intervals (CI) from a posterior distribution of 5,000 simulated values generated from linear models by the 'sim' function from the 'arm' R-package (Gelman & Su 2021). Separate models were fitted for each of the three velocity-measures and for each morphological trait (mean length or coefficient of variation based on 10 sperm cells per male). Velocities, count of tracked sperm (ln-transformed) and morphological traits were scaled (mean centered and divided by standard deviation). Velocity represents June recording (with exception of four males with May recording only). N = 92 males.

Table S6 | **Ruff sperm velocity in relation to sperm morphology.**

		Estimate (95%CI)		
Model	**Predictor | Response:**	**Curvilinear**	**Straight line**	**Average path**
Acrosome	Intercept (Independent)	0.03 (-0.15 - 0.22	0.01 (-0.20 - 0.22	0.01 (-0.18 - 0.19
	Count of tracked sperm	0.70 (0.54 - 0.85	0.60 (0.43 - 0.77	0.74 (0.59 - 0.89
	Morph (Satellite)	-0.09 (-0.43 - 0.25	0.15 (-0.23 - 0.52	0.10 (-0.24 - 0.41
	Morph (Faeder)	-0.10 (-0.66 - 0.45	-0.62 (-1.24 --0.04	-0.38 (-0.92 - 0.18
	Acrosome	**-0.06 (-0.22 - 0.09**	**0.01 (-0.16 - 0.18**	**0.04 (-0.12 - 0.18**
Nucleus	Intercept (Independent)	0.01 (-0.18 - 0.18	0.01 (-0.20 - 0.22	-0.01 (-0.19 - 0.17
	Count of tracked sperm	0.71 (0.56 - 0.85	0.60 (0.42 - 0.76	0.73 (0.59 - 0.87
	Morph (Satellite)	0.01 (-0.31 - 0.35	0.17 (-0.21 - 0.57	0.16 (-0.16 - 0.48
	Morph (Faeder)	-0.10 (-0.61 - 0.42	-0.63 (-1.23 --0.03	-0.37 (-0.87 - 0.12
	Nucleus	**0.21 (0.07 - 0.36**	**0.05 (-0.13 - 0.21**	**0.15 (0.01 - 0.29**
Midpiece	Intercept (Independent)	0.03 (-0.16 - 0.21	0.02 (-0.20 - 0.23	0.00 (-0.19 - 0.18
	Count of tracked sperm	0.71 (0.56 - 0.87	0.60 (0.42 - 0.77	0.74 (0.58 - 0.89
	Morph (Satellite)	-0.07 (-0.43 - 0.29	0.15 (-0.23 - 0.54	0.11 (-0.24 - 0.44
	Morph (Faeder)	-0.08 (-0.64 - 0.47	-0.62 (-1.25 - 0.03	-0.33 (-0.89 - 0.23
	Midpiece	**-0.02 (-0.18 - 0.15**	**0.00 (-0.18 - 0.18**	**-0.03 (-0.20 - 0.13**
Tail	Intercept (Independent)	0.04 (-0.15 - 0.22	0.03 (-0.17 - 0.24	0.02 (-0.16 - 0.20
	Count of tracked sperm	0.71 (0.56 - 0.86	0.60 (0.43 - 0.76	0.73 (0.58 - 0.87
	Morph (Satellite)	-0.10 (-0.44 - 0.25	0.11 (-0.29 - 0.50	0.05 (-0.29 - 0.41
	Morph (Faeder)	-0.10 (-0.65 - 0.45	-0.64 (-1.25 --0.04	-0.38 (-0.91 - 0.12
	Tail	**0.01 (-0.15 - 0.16**	**0.07 (-0.10 - 0.25**	**0.07 (-0.08 - 0.22**
Total	Intercept (Independent)	0.04 (-0.14 - 0.23	0.03 (-0.18 - 0.24	0.02 (-0.16 - 0.21
	Count of tracked sperm	0.70 (0.55 - 0.85	0.59 (0.42 - 0.76	0.72 (0.57 - 0.86
	Morph (Satellite)	-0.12 (-0.46 - 0.23	0.11 (-0.27 - 0.51	0.05 (-0.29 - 0.39
	Morph (Faeder)	-0.13 (-0.67 - 0.42	-0.67 (-1.28 --0.07	-0.42 (-0.93 - 0.11
	Total	**0.07 (-0.08 - 0.22**	**0.08 (-0.10 - 0.25**	**0.10 (-0.05 - 0.25**
Head	Intercept (Independent)	0.01 (-0.17 - 0.19	0.01 (-0.20 - 0.23	-0.01 (-0.19 - 0.16
	Count of tracked sperm	0.72 (0.57 - 0.87	0.60 (0.43 - 0.76	0.74 (0.59 - 0.88
	Morph (Satellite)	0.00 (-0.34 - 0.35	0.17 (-0.24 - 0.56	0.16 (-0.16 - 0.50
	Morph (Faeder)	-0.11 (-0.64 - 0.41	-0.63 (-1.23 --0.01	-0.38 (-0.90 - 0.14
	Head	**0.18 (0.04 - 0.33**	**0.05 (-0.13 - 0.22**	**0.15 (0.01 - 0.30**
Flagellum	Intercept (Independent)	0.03 (-0.17 - 0.23	0.03 (-0.18 - 0.24	0.02 (-0.16 - 0.20
	Count of tracked sperm	0.71 (0.56 - 0.86	0.59 (0.42 - 0.76	0.72 (0.58 - 0.87
	Morph (Satellite)	-0.08 (-0.44 - 0.27	0.10 (-0.31 - 0.50	0.06 (-0.28 - 0.41
	Morph (Faeder)	-0.10 (-0.68 - 0.44	-0.66 (-1.27 --0.05	-0.39 (-0.91 - 0.13
	Flagellum	**0.00 (-0.16 - 0.16**	**0.07 (-0.12 - 0.24**	**0.05 (-0.11 - 0.20**
Midpiece	Intercept (Independent)	0.02 (-0.16 - 0.21	0.01 (-0.20 - 0.21	0.00 (-0.19 - 0.18
(relative)	Count of tracked sperm	0.72 (0.57 - 0.87	0.61 (0.44 - 0.78	0.75 (0.61 - 0.90
	Morph (Satellite)	-0.07 (-0.41 - 0.27	0.16 (-0.22 - 0.54	0.11 (-0.20 - 0.43
	Morph (Faeder)	-0.04 (-0.60 - 0.51	-0.58 (-1.20 - 0.05	-0.27 (-0.81 - 0.25
	Midpiece (relative)	**-0.06 (-0.22 - 0.10**	**-0.05 (-0.23 - 0.12**	**-0.10 (-0.25 - 0.05**
Flagellum	Intercept (Independent)	-0.01 (-0.20 - 0.18	0.01 (-0.20 - 0.22	-0.02 (-0.20 - 0.17
(relative)	Count of tracked sperm	0.73 (0.58 - 0.87	0.59 (0.43 - 0.76	0.74 (0.60 - 0.89
	Morph (Satellite)	0.04 (-0.31 - 0.39	0.15 (-0.26 - 0.56	0.16 (-0.19 - 0.51
	Morph (Faeder)	-0.05 (-0.59 - 0.49	-0.63 (-1.25 --0.01	-0.34 (-0.85 - 0.17
	Flagellum (relative)	**-0.16 (-0.32 - 0.00**	**0.00 (-0.18 - 0.18**	**-0.10 (-0.25 - 0.06**

The posterior estimates (medians) of the effect sizes with the 95% credible intervals (CI) from a posterior distribution of 5,000 simulated values generated from linear models by the 'sim' function from the 'arm' R-package (Gelman & Su 2021) as presented in main text Fig. 3. Separate models were fitted for each of the three velocity-measures and for each morphological trait (in bold) while controlling for number of tracked sperm (ln-transformed) and morph. Velocity, count of tracked sperm and morphological traits were scaled (mean centered and divided by standard deviation). N = 92 males.

References

Ala-Honkola O., D. J. Hosken, M. K. Manier, S. Lupold, E. M. Droge-Young, K. S. Berben, W. F. Collins, J. M. Belote, S. Pitnick 2013. Inbreeding reveals mode of past selection on male reproductive characters in *Drosophila melanogaster*. Ecol Evol **3**:2089-2102. https://doi.org/10.1002/ece3.625.

Aparicio J. M., J. Ortego, P. J. Cordero 2006. What should we weigh to estimate heterozygosity, alleles or loci? Mol Ecol **15**:4659-4665. https://doi.org/10.1111/j.1365-294X.2006.03111.x.

Birkhead T. R., E. J. Pellatt, P. Brekke, R. Yeates, H. Castillo-Juarez 2005. Genetic effects on sperm design in the zebra finch. Nature **434**:383-387. 10.1038/nature03420.

Brooks S., A. Gelman 1998. General methods for monitoring convergence of iterative simulations. J Comput Graph Stat **7**:434–455. https://doi.org/10.1080/10618600.1998.10474787.

Bürkner P.-C. 2017. Brms: an R package for Bayesian multilevel models using Stan. J Stat Softw **80**:1-28. https://doi.org/10.18637/jss.v080.i01.

Bürkner P.-C. 2018. Advanced Bayesian Multilevel Modeling with the R Package brms. The R **10**:395-411. https://doi.org/10.32614/RJ-2018-017.

Bürkner P.-C. 2021. Bayesian Item Response Modeling in R with brms and Stan. J Stat Softw **100**:1-54. https://doi.org/10.18637/jss.v100.i05.

Coulon A. 2010. genhet: an easy-to-use R function to estimate individual heterozygosity. Mol Ecol Resour **10**:167-169. https://doi.org/10.1111/j.1755-0998.2009.02731.x.

Cramer E. R., M. Alund, S. E. McFarlane, A. Johnsen, A. Qvarnstrom 2016. Females discriminate against heterospecific sperm in a natural hybrid zone. Evolution. https://doi.org/10.1111/evo.12986.

Denk A. G., A. Holzmann, A. Peters, E. L. M. Vermeirssen, B. Kempenaers 2005. Paternity in mallards: effects of sperm quality and female sperm selection for inbreeding avoidance. Behav Ecol **16**:825-833. https://doi.org/10.1093/beheco/ari065.

Gelman A., Y.-S. Su 2021. arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package version 1.12-2. http://CRAN.R-project.org/package=arm.

Giraldo-Deck L. M., J. L. Loveland, W. Goymann, B. Tschirren, T. Burke, B. Kempenaers, D. B. Lank, C. Küpper 2022. Intralocus conflicts associated with a supergene. Nat Commun **13**:1384. https://doi.org/10.1038/s41467-022-29033-w.

Gomendio M., J. Cassinello, E. R. Roldan 2000. A comparative study of ejaculate traits in three endangered ungulates with different levels of inbreeding: fluctuating asymmetry as an indicator of reproductive and genetic stress. Proc Biol Sci **267**:875-882. https://doi.org/10.1098/rspb.2000.1084.

Heber S., A. Varsani, S. Kuhn, A. Girg, B. Kempenaers, J. Briskie 2013. The genetic rescue of two bottlenecked South Island robin populations using translocations of inbred donors. Proc Biol Sci **280**:20122228. https://doi.org/10.1098/rspb.2012.2228.

Helfenstein F., M. Podevin, H. Richner 2009. Sperm morphology, swimming velocity, and longevity in the house sparrow *Passer domesticus*. Behav Ecol Sociobiol **64**:557-565. 10.1007/s00265-009-0871-x.

Knief U., W. Forstmeier, Y. Pei, M. Ihle, D. Wang, K. Martin, P. Opatová, J. Albrechtová, M. Wittig, A. Franke *et al.* 2017. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. Nature Ecology & Evolution **1**:1177-1184. https://doi.org/10.1038/s41559-017-0236-1.

Laskemoen T., O. Kleven, F. Fossøy, R. J. Robertson, G. Rudolfsen, J. T. Lifjeld 2010. Sperm quantity and quality effects on fertilization success in a highly promiscuous passerine, the tree swallow *Tachycineta bicolor*. Behav Ecol Sociobiol:1-11. https://doi.org/10.1007/s00265-010-0962-8.

Lee M. D., E.-J. Wagenmakers 2014. Bayesian cognitive modeling: a practical course. Cambridge: Cambridge University Press.

Mehlis M., J. G. Frommen, A. K. Rahn, T. C. M. Bakker 2012. Inbreeding in three-spined sticklebacks (*Gasterosteus aculeatus* L.): effects on testis and sperm traits. Biol J Linn Soc **107**:510–520. https://doi.org/10.1111/j.1095-8312.2012.01950.x.

Opatová P., M. Ihle, J. Albrechtová, O. Tomášek, B. Kempenaers, W. Forstmeier, T. Albrecht 2016. Inbreeding depression of sperm traits in the zebra finch *Taeniopygia guttata*. Ecol Evol **6**:295-304. https://doi.org/10.1002/ece3.1868.

Pew J., J. Wang, P. Muir, T. Frasier 2015. related: an R package for analyzing pairwise relatedness data based on codominant molecular markers. https://doi.org/10.1111/1755-0998.12323.

Pizzari T., P. Jensen, C. K. Cornwallis 2004. A novel test of the phenotype-linked fertility hypothesis reveals independent components of fertility. Proc Biol Sci **271**:51-58. 10.1098/rspb.2003.2577.

Revelle W. 2022. psych: Procedures for Personality and Psychological Research, version 2.2.5.

Schmoll T., G. Rudolfsen, H. Schielzeth, O. Kleven 2020. Sperm velocity in a promiscuous bird across experimental media of different viscosities. Proc Biol Sci **287**:20201031. https://doi.org/10.1098/rspb.2020.1031.

Stan-Development-Team 2022. RStan: the R interface to Stan. R package version 2.21.5. https://mc-stan.org/.

Stoffel M. A., S. Nakagawa, H. Schielzeth, S. Goslee 2017. rptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. Methods Ecol Evol **8**:1639-1644. https://doi.org/10.1111/2041-210x.12797.

Støstad H. N., A. Johnsen, J. T. Lifjeld, M. Rowe 2018. Sperm head morphology is associated with sperm swimming speed: A comparative study of songbirds using electron microscopy. Evolution **72**:1918-1932. https://doi.org/10.1111/evo.13555.

Tomasek O., J. Albrechtova, M. Nemcova, P. Opatova, T. Albrecht 2017. Trade-off between carotenoid-based sexual ornamentation and sperm resistance to oxidative challenge. Proc Biol Sci **284**. https://doi.org/10.1098/rspb.2016.2444.

Venables W. N., B. D. Ripley 2002. Modern Applied Statistics with S., Fourth ed. New York: Springer.

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York.