SYMPOSIUM

Divergent Genital Morphologies and Female–Male Covariation in Watersnakes

J. F. Greenwood[*], G. Lara Granados[*], S. M. Secor[†], B. D. Todd [iD][‡], I. Showalter[§], B. P. Hedrick [iD][¶] and P. L. R. Brennan [iD][*,1]

[*]Department of Biological Sciences, Mount Holyoke College, South Hadley, MA 01075, USA; [†]Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487, USA; [‡]Department of Wildlife, Fish, and Conservation Biology, University of California, Davis, CA 95616, USA; [§]Biology Department, University of Massachusetts–Amherst, Amherst, MA 01003, USA; [¶]Department of Cell Biology and Anatomy, Louisiana State University Health Sciences Center, New Orleans, LA 70012, USA

[1]E-mail: pbrennan@mtholyoke.edu

Synopsis   Genital evolution can be driven by diverse selective pressures. Across taxa we see evidence of covariation between males and females, as well as divergent genital morphologies between closely related species. Quantitative analyses of morphological changes in coevolving male and female genitalia have not yet been shown in vertebrates. This study uses 2D and 3D geometric morphometrics to quantitatively compare the complex shapes of vaginal pouches and hemipenes across three species of watersnakes (the sister taxa *Nerodia fasciata*, *N. sipedon*, and a close relative *N. rhombifer*) to address the relationship between genital morphology and divergence time in a system where sexual conflict may have driven sexually antagonistic coevolution of genital traits. Our pairwise comparisons of shape differences across species show that the sister species have male and female genitalia that are significantly different from each other, but more similar to each other than to *N. rhombifer*. We also determine that the main axes of shape variation are the same for males and females, with changes that relate to deeper bilobation of the vaginal pouch and hemipenes. In males, the protrusion of the region of spines at the base of the hemipene trades off with the degree of bilobation, suggesting amelioration of sexual conflict, perhaps driven by changes in the relative size of the entrance of the vaginal pouch that could have made spines less effective.

Introduction

Genitalia display diverse ornamentation and shape features across taxonomic groups, revealing a variety of evolutionary patterns (Eberhard 1985; House and Simmons 2005; Orbach et al. 2018). Sexual selection—including intrasexual competition, sexually antagonistic coevolution, and cryptic female choice—is thought to be the dominant driving force acting on genital diversification (House and Simmons 2005). Rapidly divergent male genital morphology across closely related species has been studied extensively, while the same divergence pattern in female genital structures has only recently been explored (Puniamoorthy, Kotrba, and Meier 2010; Showalter et al 2014; Orbach et al. 2018; Simmons and Fitzpatrick 2019). Sexually antagonistic coevolution (SAC) is an important driver behind genital diversity (Arnqvist and Rowe 2005; Brennan et al. 2007; Brennan and Prum 2015) and involves the rapid evolution of antagonistic traits in males and females resulting from their differing evolutionary interests (Arnqvist and Rowe 2005). However, despite selection for antagonistic traits derived from sexual conflict, genitalia must interact mechanically and fit together for copulation to be feasible. For this reason, we expect some degree of shape correspondence between the sexes despite the possibility that SAC may also be acting on genital traits (Brennan and Prum 2015).

Here, we investigate genital shape variation in three species of watersnakes within the Genus *Nerodia* (natricine snakes native to North America) to examine how divergence time may correlate with genital morphology in a group where sexual conflict may have influenced genital traits. *Nerodia fasciata* and *N. sipedon* (the southern watersnake and the northern watersnake), diverged 3.9 mya, while their more distant relative *N. rhombifer* (the Diamond-Backed Watersnake), diverged from the last common ancestor with these two about 6.5 mya (Brandley et al. 2010), and we predict that the two species that more recently shared a common ancestor would have genitalia that are more similar to each other than to the third species with which they diverged more distantly. We also predict that males and females covary in their genital shape, even if some genital traits have evolved as a result of sexual conflict, to make copulation feasible. This type of complex relationship between divergence time, and covariation under sexual conflict has been previously reported in beetles (e.g., *Ohomopterus* ground beetles, Sota and Nagata 2008).

The vagina in snakes, also known as a vaginal pouch, is the female copulatory organ that consists of a cloacal opening leading to a cloacal chamber that branches off into two oviducts. The vaginal pouch interacts with the male intromittent organ during copulation, plays a functional role in giving birth and ovipositing, and it is known to vary in morphology among snake species (Sanchez-Matinez et al. 2007; Aldridge et al. 2011; Siegel et al. 2012), and watersnakes specifically (Showalter et al. 2014). The male hemipenes are the intromittent organs that become everted and inflated during copulation and typically have a number of external elaborations such as spines, chalices, and folds (Dowling and Savage 1960). Snakes have two hemipenes kept inside out in the posterior end of the tail, and males use one at a time during copulation (Dowling and Savage 1960).

Hemipenes are considered to be species-specific amongst squamates (e.g., Jadin and Parkhill 2011; Andonov et al. 2017). In *Nerodia*, males have a bilobed hemipene tip and sharp calcified spines at the base of the hemipene shaft (Andonov et al. 2017). Sharp basal spines on the hemipenes of snakes function to perforate vaginal tissue and act as an anchoring mechanism (Pope 1941; Edgren 1953). Because spines are found at the base of the hemipenes of *Nerodia*, sexual conflict over copulation may be found where males increase the chances of sperm delivery by firmly attaching to the females, while females may be injured from spines penetrating their vaginal mucosa, possibly causing delays before additional mating events can occur, and/or being forced to receive more sperm than what they may otherwise choose. An example of such conflict over copulation duration has been described in the red-sided

garter snake (*Thamnophis sirtalis parietalis*), where experimental manipulation of genitalia has revealed that males with a clipped basal spine copulate for shorter periods of time and transfer a smaller copulatory plug and fewer sperm to females (Friesen et al. 2014; 2016).

Further evidence of the relationship between sexual conflict over copulation duration and hemipene morphology has been suggested for New World natricine snakes, where hemipenes were classified according to whether they were simple or bilobed (King et al. 2009). Multiple shape transitions are found in this group, but no clear association emerged between hemipene shape and copulation duration (King et al. 2009). However, King et al. (2009) did not investigate the potential role of basal spines in conflict over copulation duration (e.g., Friesen et al. 2014), and classified shape in only two distinct categories, while it is likely that different species have different degrees of bilobation.

In addition to these hemipene features that likely indicate an evolutionary history of sexual conflict, females mate with multiple males in both *N. rhombifer* and *N. sipedon* (Prosser 2002; King et al. 2010), which increases the potential conflict because it would be more advantageous to males if females were monogamous (Arnqvist and Rowe 2005). Aggressive behaviors of males towards females during copulation have been observed in field studies of *N. sipedon*, which include forceful jerking of the female and surrounding males (Weatherhead et al. 1995). However, forced copulation is not likely in watersnakes since females must give males access to their cloaca to begin copulation (Barry et al. 1992), there is a loose structure of mating aggregations (Weatherhead et al. 1995), and large female size would suggest females can control mating (Kissner et al. 2005). Therefore, the conflict is most likely about the duration of copulation, and potential harm incurred by females from the hemipene spines as is the case in red-sided garter snakes (Friesen et al 2014; 2016).

Sexual conflict over copulation may have acted as a driving force for divergent genitalia between species of watersnakes, resulting in a pattern of SAC which can be traced across evolutionary time by comparing *N. rhombifer* genital shape to that of more recently diverged species *N. fasciata* and *N. sipedon* (e.g., Lawson 1987; McVay, Flores-Villela, and Carstens 2015). Significant differences in vaginal pouch shape across reproductive *N. fasciata* and *N. sipedon* females have been reported (Showalter et al. 2014), but the differences in genital shape of *Nerodia* hemipenes have not been quantified. We use 2D and 3D geometric morphometrics (GM), which have been previously applied to genital morphological studies with success (Showalter et al. 2014; Orbach et al. 2018; Hedrick et al. 2019; Orbach et al. 2020; Lara-Granados et al. submitted), to quantify
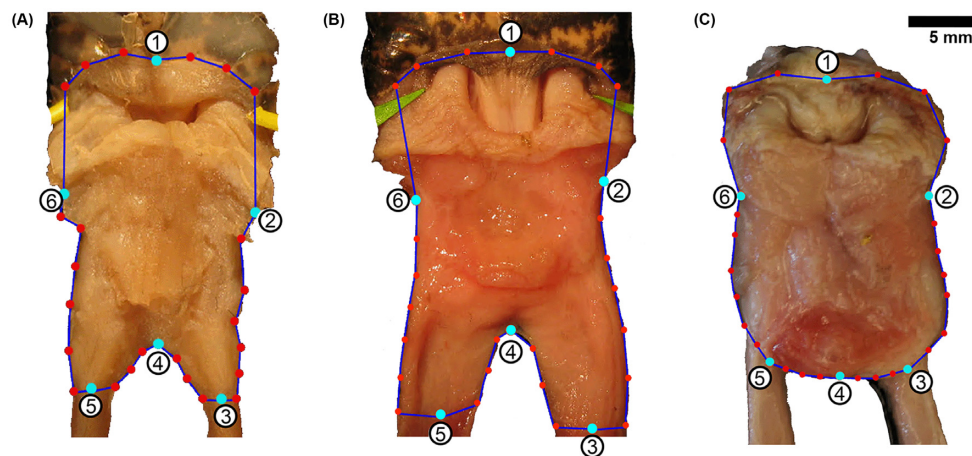
Fig. 1 *Nerodia* vaginal pouch landmark configuration. Numbered points represent homologous anchoring landmark placement and red points represent semi-landmark placement on *N. fasciata* **(A)**, *N. sipedon* **(B)**, and *N. rhombifer* **(C)**.

male and female watersnake genital shape both inter- and intraspecifically. Additionally, we compare our 2D results with the 3D results of *N. rhombifer* female vaginal pouches from a different study (Lara-Granados et al. submitted), to determine which aspects of shape variation we may be missing by using 2DGM.

Methods

Specimen collection

N. fasciata (females, n = 30, males, n = 20) and *N. sipedon* (females, n = 33, males, n = 20) genitalia were previously dissected and photographed with a scale by Showalter et al. (2014). In brief, *N. fasciata* and *N. sipedon* were taken from established non-native populations in California in 2011 where they were trapped under the California Department of Fish and Wildlife collecting permits #SC-11197 to BD Todd, and #802,046–02 to RN Reed of the US Geological Survey. Both trapping and euthanasia were performed according to UC Davis Institutional Animal Care and Use Committee protocol #2011–16553 (full details in Showalter et al. 2014). *Nerodia rhombifer* (females, n = 29, males, n = 22) specimens were acquired in collaboration with the University of Alabama Department of Biological Sciences and collected by S Secor under Alabama collecting permit number 2018101271668680.

Specimen dissection and preservation

Prior to dissection, we measured snout-to-vent length (SVL) to the nearest mm (Tables S1 and S2). We dissected specimens to examine and measure their gonads to the nearest mm using calipers. Males were determined to be reproductive if they had coiled, well developed *vas deferens*. Females were deemed non-reproductive if they had small eggs (<2 mm). Females

with yolk-filled or fertilized eggs in their oviducts were considered reproductive.

To generate 2D data for vaginal shape and size, we cleaned the connective tissue surrounding the vaginal pouch, removed any remnants of the intestines, and cut the oviducts close to their connecting point with the vaginal pouch. We laid the vaginal pouches on their dorsal side and photographed them next to a scale using Canon EOS Rebel T5i, following previous protocols used in Showalter et al. (2014).

We inflated the left hemipenes of male snakes with warmed petroleum jelly following standard herpetological techniques (e.g., Jadin and King, 2012). We preserved the removed hemipenes suspended in a solution of 10% formalin.

Vaginal pouch shape

We used two-dimensional geometric morphometrics (2DGM) to quantitatively determine the shape of 92 vaginal pouches (*N. fasciata*; n = 30, *N. rhombifer*; n = 29, *N. sipdeon*; n = 33) using a combination of anchoring landmarks (n = 6) and semi-landmarks (n = 6 curves) (Fig. 1). Photographs of *N. fasciata* and *N. sipedon* females were previously acquired by Showalter et al. (2014), but were re-landmarked by J Greenwood using a modified landmarking scheme developed after adding *N. rhombifer* to the sample, eliminating inter-observer landmarking error. Landmarks were placed based on the following criteria: (1) The midpoint of the cloacal opening where the cloacal scale transitions into vaginal tissue, based on change in texture and coloration (2) the intersection of vaginal pouch with cloacal opening on the right side, (3) the midpoint where thick vaginal tissue meets oviduct tissue on the right side based on texture and change in coloration (vaginal pouch is white and thick compared to pink and thin

oviduct tissue), (4) the midpoint in the portion of vaginal tissue at the point of furthest bifurcation, (5) the midpoint where thicker vaginal tissue meets oviduct tissue on the right side based on texture and change in coloration, (6) the intersection of the vaginal pouch with cloacal opening on the left side. Landmarking required zooming in the image to place each point. To further reduce bias during manual landmark placement, specimens were randomly ordered when landmarking rather than landmarking all individuals of a single species before moving on to the next species. Landmarks and semi-landmarks were then placed in tpsDIG2 (Rohlf, 2006).

The shape data were uploaded to R v. 4.04 (R Core Development Team, 2021) for generalized Procrustes analysis (GPA) using geomorph v. 3.3.2 (Adams et al. 2021) to align and scale all landmark configurations in relation to centroid size (Zelditch et al., 2012). Semi-landmarks were slid using the bending energy criterion (Perez et al. 2006). GPA data were submitted to PCA to examine overarching trends in morphospace of the vaginal pouch and assess potential grouping of species. We only assessed shape trends in principal components (PCs) that contained greater than 10% of total variance since additional PCs accounted primarily for outlier specimens. We ran Procrustes ANOVAs assessing the relationship between vaginal pouch shape, log10-transformed SVL, species, and reproductive state. Each Procrustes ANOVA used 1000 iterations. First, we assessed the relationship between shape and species only to determine if the species differed in shape. We used the function "pairwise" in RRPP v. 0.6.2 (Collyer and Adams, 2018; 2019) to determine which species were different from which other species. Prior to evaluating how our factors related to one another in more complex Procrustes ANOVAs, we tested whether our factors were correlated. We used a two-sample t-test of log10-transformed SVL and reproductive state to determine whether reproductive individuals were statistically larger than non-reproductive individuals. We then used an ANOVA with a Tukey post-hoc test to determine whether there was a relationship between species and log10-transformed SVL. Both of these preliminary analyses were plotted as violin plots using ggplot2 v. 3.3.5 to visualize trends. Following these analyses, we ran a Procrustes ANOVA assessing the relationship between shape and species after first factoring out SVL and reproductive state (shape ∼ SVL + reproductive state + species).

We assessed inter-specimen variation caused by intra-individual landmarking error following Singleton (2002) and Lockwood et al. (2002). We landmarked a single *N. sipedon* specimen (NS023) 11 times using tpsDIG2, which constituted an error subset. We ran a GPA which included the full dataset and the error subset. We then used a principal component analysis (PCA) to qualitatively assess whether the error subsample grouped closely in morphospace and thus that landmarking error did not strongly impact overarching shape trends in morphospace. To further address error, the average of the PC1, PC2, PC3, and PC4 (constituting 84.1% of total shape variance) scores were taken for the error sample. Then the square root of the sum of squared distances between the average of the error sample and PC1–4 for each individual specimen were taken to generate the Euclidean distances between the error sample and original specimens. We then plotted the distribution of these distances as a histogram to show that the error samples were more similar in shape to one another than to other individuals.

Comparison OF 2D and 3D Results for vaginal pouches

We additionally examined differences between our 2D vaginal pouch data and 3D data collected for *N. rhombifer* as part of a different study (Lara Granados et al. in press). In brief, models of the lumen of the female *N. rhombifer* were scanned and made into 3D models that were analyzed using auto3dgm (with 1000 landmarks) (see Lara Granados et al. in press for details). The same protocols were carried out for female data as those described from the 3D shape data of male hemipenes in this study. We then performed a GPA on the subset of *N. rhombifer* data and used PCA to examine differences in morphospace trends. Similar sets of Procrustes ANOVAs described for the vaginal pouches above were run for the *N. rhombifer* 2D data subset and in 3D (Lara Granados et al. submitted). Finally, we performed a two-block partial least squares (PLS) analysis with the 2D data in one block and the 3D data in the other block to examine whether the two sets of data were strongly associated (Rohlf and Corti, 2000). We could not assess vaginal pouch shape in *N. sipedon* or *N. fasciata* because the vaginal pouches were already preserved in 10% Formalin and we could not make molds of the lumen.

Hemipene shape

We used an EXScan Pro 2X Plus laser scanner and Shining 3D EXScan Pro software (v. 3.2) to create 3D object files for the left hemipene of each male specimen (Fig. 2). We fixed hemipenes at the base and placed them on the center of a rotating table and scanned each specimen at three angles to fully capture the complex three-dimensional shape. Data were then exported as obj files, which were imported into MeshMixer to create a flat base for every specimen using the "plane cut" feature.
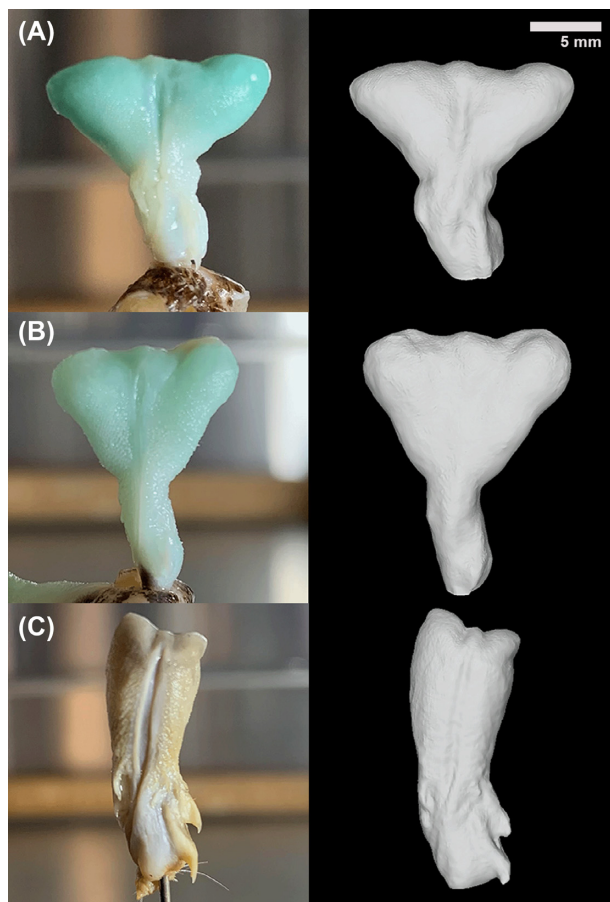
Fig. 2 Sulcate view of inflated left hemipenes of *N. fasciata* **(A)** *N. sipedon* **(B)** and *N. rhombifer* **(C)** next to their corresponding 3D models. 3D models are based on laser scans of tissue specimens with the removal of inconsistencies at the basal portion of the model.

We uploaded 62 male 3D models (*N. fasciata*; n = 20, *N. rhombifer*; n = 22, *N. sipdeon*; n = 20) into 3D Slicer as individual object files to be analyzed through Automated 3D Geometric Morphometrics (auto3Dgm) implemented in SlicerMorph software (Boyer et al. 2015; Rolfe et al. 2021). The auto3Dgm method involves the automatic placement of pseudolandmarks on the surface of 3D models followed by the computing of correspondences and differences between individuals within the sample, thereby eliminating landmarking error, and decreasing subjectivity in landmark placement. For this reason we did not carry out error estimates on landmarking hemipenes as we did with vaginal pouches. The analysis computed pairwise distances between each pseudolandmark on every 3D hemipene specimen (Gunnel et al. 2018). We assigned 100 landmarks for phase 1 and 800 landmarks for phase 2 and used the phase 2 data for our analyses.

Similar to the female vaginal pouch data, the male hemipene shape data were first evaluated in mor-

phospace for overarching shape trends using PCA. As with the female data, only PCs that accounted for more than 10% of total shape variation were considered. We first assessed the relationship between species and log10-transformed SVL using an ANOVA and Tukey post-hoc test. Results were visualized using a violin plot as with the female data. To assess whether there was a significant relationship between reproductive state and hemipene shape, we examined only *N. fasciata* because they were the only species for which we had a similar number of reproductive (n = 9) and non-reproductive (n = 11) males. We evaluated reproductive state in morphospace using PCA and then ran a Procrustes ANOVA of shape and reproductive state using 1000 iterations. Following this, we ran a Procrustes ANOVA for the entire male hemipene dataset evaluating differences in hemipene shape, log10-transformed SVL, and species (shape $\sim$ SVL + species) in *geomorph* and calculated pairwise comparisons in *RRPP*.

Results

Vaginal pouch shape

In the full female dataset, only the first three PCs accounted for more than 10% of total shape variation (39.2%, 24.2%, and 15.4%, respectively; Table S3). Along PC1, the entrance of the vagina is wider relative to the cranial end at the positive end of the axis with the opposite trend being true at the negative end of the axis (Fig. 3). The three species are not separated out clearly along PC1, although *N. rhombifer* does cluster on the negative end of PC1. Non-reproductive individuals tended towards the positive end of PC1 and reproductive individuals tended towards the negative end of PC1. Unlike PC1, PC2 clearly separates all three species in morphospace with minimal overlap between *N. rhombifer* and the other species and moderate overlap between *N. fasciata* and *N. sipedon*. PC2 is characterized by a shift in the degree of cranial bifurcation of the vaginal pouch. In the negative values of PC2, where both *N. fasciata* and *N. sipedon* are located (Fig. 3), the vaginal pouches are deeply bifurcated. Conversely, *N. rhombifer* is located on the positive end of PC2 and has minimal bifurcation. PC3 accounts for a much smaller degree of shape change than either PC1 or PC2 and is characterized by a slight broadening of the entrance of the vagina and center of the vaginal pouch with minimal change to the cranial end of the pouch. Species are not discriminated along PC3 (Fig. S1).

The three snake species differed in size (log10-transformed SVL, F = 18.86, df = 2, p < 0.001) (Fig. 4A), and larger snakes were more likely to be reproductive (log10-transformed SVL, t = -10.12, df = 70.91, p < 0.001) (Fig. 4B). Tukey post-hoc tests revealed *N.*
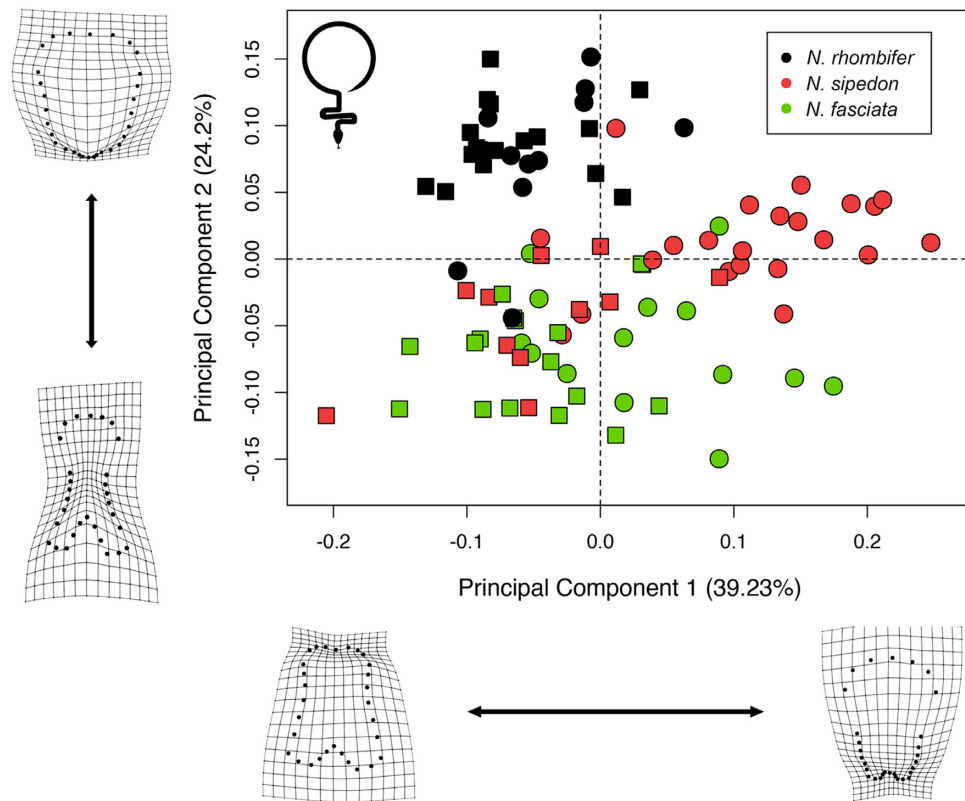
Fig. 3 Principal component analysis (PCA) of vaginal shape. PC1 shows the transition from a wider, square pouch (negative values) to a narrower, curved, bifurcated pouch (positive values). PC2 shows a transition from bifurcated, narrow pouch (bottom) to a wider overall shape particularly at the cloacal entrance, with very little bifurcation (top). Squares represent reproductive specimens while circles represent non-reproductive specimens. Deformation grids show specimen shapes located at the extreme ends of each PC.

rhombifer being larger than *N. fasciata* and *N. sipedon* (both p-values < 0.001). However, *N. fasciata* and *N. sipedon* did not differ significantly in size (p = 0.325). The model examining the relationship between vaginal pouch shape, log10-transformed SVL, reproductive state, and species (shape ∼ SVL + reproductive state + species) revealed that all factors were significant (Table 1; Table S4). SVL accounted for 23.5% of shape variation, reproductive state accounted for 4.9% of shape variation, and species accounted for 16.2% of shape variation. Pairwise tests revealed that all three species were significantly different from one another in vaginal pouch shape (p < 0.001). The pairwise distances between means were lowest for *N. fasciata* and *N. sipedon* (d = 0.09), were relatively higher for *N. rhombifer and N. sipedon* (d = 0.11), and were highest for *N. fasciata* and *N. rhombifer* (d = 0.15). These results mirror those found in the PCA.

The intra-observer error analysis for the vaginal pouches revealed a low degree of scatter for the error subset within the morphospace (Fig. S2A; Table S5). This suggests that intra-observer landmarking error is unlikely to have significantly impacted morphospace trends. To quantitatively measure the error, we plotted

a frequency distribution of the Euclidean distances between the mean of the error sample and each individual specimen, including all error replicates (Table S6). The error replicates had Euclidean distances that were closer to the mean of the error dataset than to any other individual specimen in the analysis (Fig. S2B).

Comparison OF 2D VS. 3D Results for vaginal pouches

The results of the Procrustes ANOVA analysis and the morphospace trends in PCA for female *N. rhombifer* were similar in 2D and 3D (data for 3D analyses in Lara Granados et al. in press). SVL was significantly associated with vaginal pouch shape in both methods (2D: F = 3.21, P = 0.024; 3D: F = 3.32, P < 0.001), whereas reproductive status was not (2D: F = 1.43, P = 0.223; 3D: F = 1.40, P = 0.122). Examination of PC1 and PC2 showed similar shape variation with both methods (Fig. S4). In both 2D and 3D data, PC1 describes variation in the aspect ratio of the vagina, with vaginal pouches on one end being narrower than at the other end. PC2 also similarly describes variation between a squarer looking vaginal pouch on one end, and a wider and more
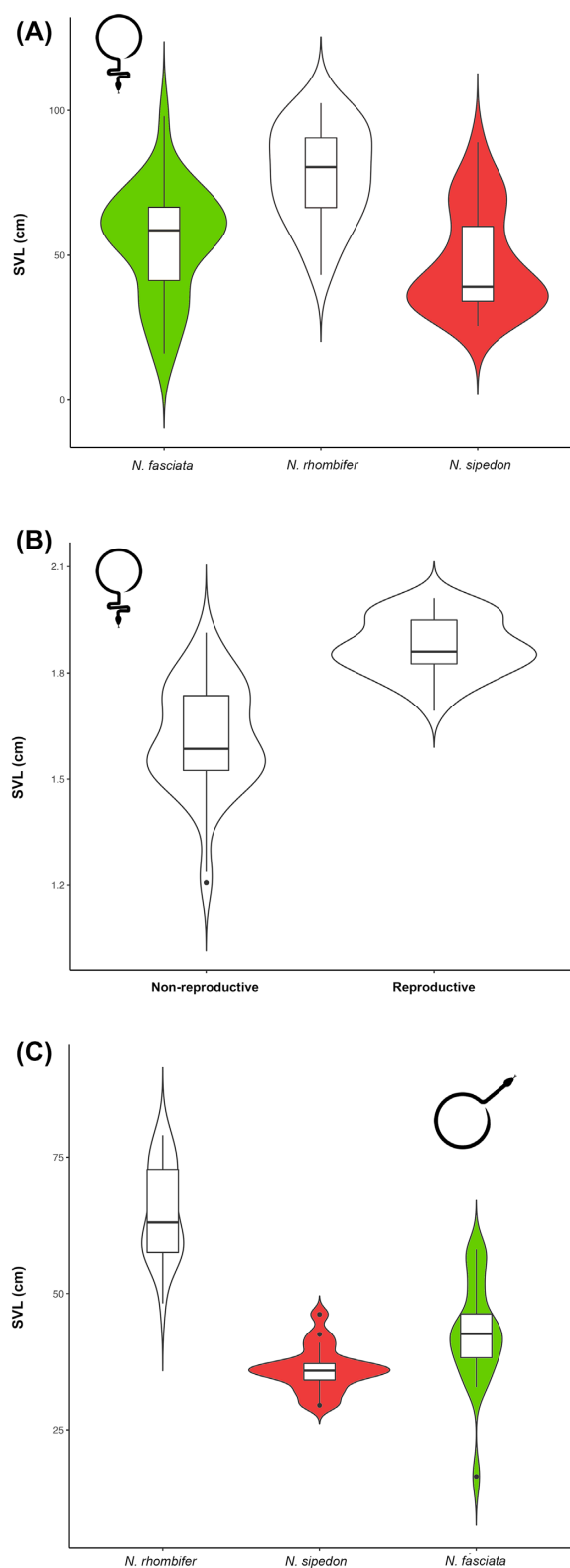
Fig. 4 Violin plots showing relationship between SVL and species for females **(A)** and reproductive status **(B)** across all female specimens (n = 92). Violin plots showing the relationship between SVL and species for males **(C)**

bifurcated pouch on the opposite end. However, the separation of reproductive and non-reproductive females was more obvious in PC2 than in PC1 in the 2D data, with most reproductive females falling along the positive end of PC2, whereas it was more obvious along the PC1 in 3D data, with reproductive females falling along negative values of PC1. The lateral aspect of the vaginal pouch was narrower for reproductive versus non-reproductive females in 3D, and this view was missing from our 2D study. Despite strong similarities in biological signal across the two datasets, the PLS analysis suggested limited association between the 2D and 3D data (rPLS = 0.651, P = 0.445).

Hemipene shape

Before examining trends in hemipene shape across all three species, we first evaluated the relationship between shape and reproductive state in *N. fasciata*. For *N. rhombifer*, we had mostly reproductive males, in *N. sipedon* we had mostly non-reproductive males, and only in *N. fasciata* we had fairly equal numbers of reproductive and non-reproductive males. Plotting the *N. fasciata* data as reproductive and non-reproductive in morphospace revealed no obvious trend, with strong overlap between the two groups (Fig. S3). A Procrustes ANOVA shows that hemipene shape and reproductive state were not significantly correlated with one another in *N. fasciata* (F = 0.515, P = 0.73). Therefore, we are unlikely to be missing important trends by not considering reproductive state in the multi-species hemipene dataset.

In the full male dataset, only PC1 and PC2 accounted for > 10% of total shape variation (60.3% and 12.76% respectively). The PCA revealed distinct clustering based on species (Fig. 5). Along PC1, specimens of *N. rhombifer* are highly concentrated in the negative region where the hemipene is more cylindrical, lacks marked bilobation, and has a wide shaft, widest at the base that marks the position of the large and curved basal spines observed in *N. rhombifer*. In contrast, *N. fasciata* and *N. sipedon* remain mostly on the positive end of this axis that describes deep bilobation and much narrower shafts with no widening of the base, reflecting relatively small and straight basal spines that lie against the shaft. There is limited separation of species along PC2, but *N. sipedon* remains mostly within the positive quadrant whereas *N. rhombifer* clumps toward the middle and *N. fasciata* toward the negative end. Negative values of PC2 are representative of a wider overall hemipene shape in both the basal and apical regions as well as from sulcal and lateral views. *Nerodia fasciata* occupies this negative quadrant of PC2 whereas

Table 1 Procrustes ANOVA and pairwise comparisons output for all females. SVL = Snout-vent length. Repro = reproductive state.

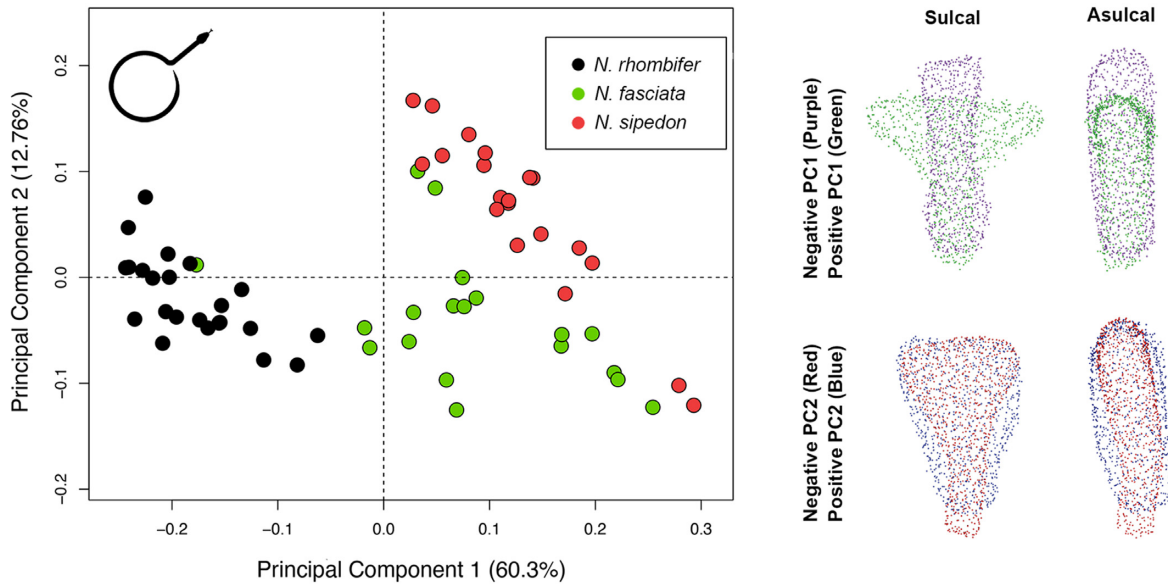| Model: Vaginal Pouch Shape $\sim$ SVL + Reproductive State + Species							
	Df	SS	MS	Rsq	F	Z	Pr(>F)
SVL	1	0.495	0.495	0.235	37.034	5.685	**0.001**
Repro	1	0.104	0.104	0.049	7.781	3.719	**0.001**
Species	2	0.341	0.17	0.162	12.744	6.07	**0.001**
Residuals	87	1.163	0.013	0.553			
Total	91	2.103					

Fig. 5 Principal component analysis of hemipene shape. 3D clusters show hemipenal shape variation in PC1 and PC2 from cranial and lateral perspective. The purple point cloud corresponds with the sub-cylindrical hemipene shape at the negative end of PC1. The green point cloud corresponds with bilobed hemipene shape at the positive end of PC1. The red point cloud corresponds with the elongated hemipene base shown at the negative end of PC2. The blue point cloud corresponds with the wide base and widened width at the positive end of PC2.

N. sipedon occupies the positive values where overall hemipene shape is narrower (Fig. 5).

All three species were significantly different in SVL from one another at a 95% confidence interval with *N. rhombifer* being the largest, *N. fasciata* being intermediate, and *N. sipedon* being the smallest (Fig. 4C, log10-transformed SVL, df = 2, F = 58.35, $P < 0.001$). Tukey post hoc tests revealed that *N. rhombifer* was significantly larger than the other two species ($P < 0.001$), but that *N. fasciata* and *N. sipedon* were not significantly different in size ($P = 0.057$). When examining the relationship between hemipene shape, SVL, and species (shape $\sim$ SVL + species), we found that both factors were significantly correlated with shape ($p < 0.001$) (Table 2; Table S8). SVL accounted for 22.1% of shape variation and species accounted for 32.1% of shape variation. All three species significantly differed in shape based on pairwise comparisons, with *N. fasciata* and *N. sipedon* being the most similar to one another (d = 0.139, $P = 0.003$),

and *N. rhombifer* being significantly different from both other species (P-values < 0.001). As with the vaginal pouch data, these results mirror those found in the PCA.

Discussion

We show quantitatively that the shape of the vaginal pouch and the hemipenes of three species of watersnakes differs according to their divergence time, with *N. fasciata* and *N. sipedon* being more similar to one another than to the more phylogenetically distant *N. rhombifer*. We show that female vaginal pouches in different species are significantly different from each other in the relative size of the vaginal pouch opening compared to the cranial end of the pouch, with most *N. rhombifer* having proportionally similar widths in the openings and cranial ends and *N. fasciata* and *N. sipedon* having larger openings and relatively narrower

Table 2 Procrustes ANOVA and pairwise comparisons output for all males.

Model: Hemipene Shape ~ SVL + Species							
	Df	SS	MS	Rsq	F	Z	Pr(>F)
SVL	1	0.572	0.572	0.221	28.037	3.329	**0.001**
Species	2	0.829	0.415	0.321	20.317	4.463	**0.001**
Residuals	58	1.184	0.02	0.458			
Total	61	2.585					

cranial ends. *Nerodia rhombifer* clearly separates from the other two species in that it has a minimally bifurcated vaginal pouch, whereas both *N. fasciata* and *N. sipedon* have deep bifurcation. In the hemipenes, the main axis of variation relates to the degree of bilobation, with *N. rhombifer* having no distinct lobes and a wider shaft specially at the base where the basal spines are located, and the other two species having large, bifurcated lobes and narrower shafts that reflect a reduction of the basal spines. These results contradict the hemipene classification of King et al. (2009), where *N. rhombifer* is classified as bilobed, and suggest that bilobation and basal spines are two independent mechanisms that allow males to anchor to the female during copulation. This work provides the basis for further comparative work on snakes that can help determine if the evolutionary rate of genitalia corresponds to evolutionary time in this group.

From these morphological observations and previously reported observations of mating conflict in *Nerodia*, we can speculate how changes in their genitalia may have evolved. Males with a simple, more cylindrical hemipene may require larger and/or more curved spines to more effectively anchor inside the female, especially if she tries to dislodge the male quickly during copulation (e.g., King et al. 2009; Friesen et al. 2014). If the spines cause harm to the females, selection on females may limit the damage of copulatory wounding by making the spines less effective. A widening of the opening of the pouch may make anchoring of the hemipene spines more difficult, so that males require an additional anchoring mechanism, in this case bilobation. Females then also become bifurcated because of likely reduced harm of the anchoring using lobes compared to the large basal spines. If a bilobed shape is effective at keeping the male anchored, the basal spines may become reduced in size, and become more straight. Such a change would benefit females if it reduces harm brought by basal spines anchoring in her vaginal mucosa. Finding evidence for this phenomenon would require closer examination of copulatory mechanics and basal spine morphology across a wider range of *Nerodia* species. Our current hypothesis is supported by the observation that *N. erythrogaster*, a species sister to *N.*

fasciata and *N. sipedon* that shared a last common ancestor more recently than with *N. rhombifer* (4.1 mya vs. 6.5 mya) has deeply bilobed hemipenes, but also retains the larger and curved basal spines seen in *N. rhombifer* (Cadle, 2011). This suggests that these two traits may evolve independently.

Further differences in the relative size of the vaginal pouch opening between *N. fasciata* (narrower) and *N. sipedon* (wider), could be a potential female-driven mechanical barrier to hybridization. Considering the cloaca is the point of entrance for the hemipene (not just the anchoring point to the spines), this divergence in vaginal pouch widening could indicate a female-evolved reproductive isolation mechanism to prevent interpopulation mating that is worth investigating in a future study. Current evidence suggests that hybridization of these two species in contact zones is common, but *N. fasciata* genes dominate in the hybrids (Mebert, 2008).

Our findings revealed distinct genital shape covariation in males and females across closely related species of *Nerodia*, and are consistent with the patterns of species-specific genital diversity often observed in squamates (Brennan and Prum 2015). The data showed significant shape differences across species pairs for both vaginal and hemipenal shape, where shape changes in one sex correspond with changes in the other. In this case, bilobation was found in both males and females of *N. fasciata* and *N. sipedon* and not in either sex of *N. rhombifer* (Figs. 3, 5). While we could not directly compare genital size between females and males of *N. sipedon* and *N. fasciata* because we used 2D vs 3D methods to quantify their shape, we found that the there are no differences in genital size in female and male *N. rhombifer* in 3D (Lara-Granados et al., in press), further suggesting coevolution to allow for mechanical fit during copulation.

In our comparison of female vaginal pouches in *N. rhombifer* for 2D data in this study and 3D data in Lara Granados et al. (in press), we found that the main aspects of shape variation were similar for both methods, revealing variation in the aspect ratio of the vagina, as well as the degree of bifurcation. Similarly, the Procrustes ANOVAs also showed significant differences in

vaginal pouch shape associated with SVL and not with reproductive state (although note that SVL and reproductive state were strongly correlated in both sets of analyses). This means that when computing resources are limited, 2DGM analysis may suffice for studies of the vaginal pouch and that biological signal is not lost, a result similarly found by Cardini and Chiapelli (2020) in horse crania. However, our low correlation between 2D and 3D data demonstrates that 2D and 3D data are not interchangeable and though the biological signals are the same across our 2D and 3D data, 2D is not fully capturing shape—a result also found by Cardini and Chiapelli (2020). The missing element from the 2D analysis was the observation that the cranial end of the vaginal pouch becomes narrower in its lateral aspect in more reproductive females, likely to facilitate sperm transfer (Lara Granados et al. submitted). This observation may or may not be consistent across species. Generally, the approximation of 3D structures into 2D is not always perfect and 2D and 3D analyses on the same data may generate differing trends (e.g., Hedrick et al. 2019). Given the large number of studies using each methodology, it is important that future studies establish when and how to employ 2D approximations of 3D structures.

We could not include reproductive status for all species in our analyses because we had differing sample compositions in each species. We had fairly equal numbers of reproductive and non-reproductive males only in *N. fasciata*, and the single-factor Procrustes ANOVA analysis of reproductive status and hemipene shape in *N. fasciata* did not show significant differences associated with reproductive status (Fig. S3). This allowed us to rule out reproductive status as a confounding variable influencing our main three species analyses. In the future, it would be beneficial to have larger sample sizes that include equal representation of reproductive and non-reproductive males of each species to more effectively analyze the relationship of reproductive status to male hemipene shape across *Nerodia*.

In addition to reinforcing previous findings in both squamate and general genital evolution studies, this research represents the first attempt to measure hemipene morphology across species using automated landmarking methods. The structural similarities in *N. fasciata* and *N. sipedon* hemipenes led us to believe they would not have significant shape differences, and although our findings contradicted this prediction, they confirm the ability of automated 3DGM to capture complex shape variation that is difficult to estimate with the naked eye. Another advantage of auto3dgm is that it eliminates the need to landmark by hand, thereby eliminating observer errors related to landmark placement (Rolfe et al. 2021). However, despite this shortcoming of 2D landmarking, our study also demonstrates the repeatability

of 2DGM since we confirmed previously found results for female genital variation in *N. fasciata* and *N. sipedon* (Showalter et al., 2014).

This study adds to the growing research on female genital variation and explores shape covariation in the genitalia of males and females, suggesting a potential scenario for a coevolutionary arms race between males and females. We also quantify genital shape differences amongst species of snakes who diverged at different times from their last common ancestor showing that divergence time affects genital divergence. Further work in a larger comparative context would help elucidate how widespread the patterns we observe may be.

Acknowledgments

Funding

Supplementary data

Supplementary data available at *ICB* online.

Data Availability Statement

The data and code underlying this article will be shared on reasonable request to the corresponding author.

References

Adams DC, Collyer ML, Kaliontzopoulou A, Balken E. 2021. Geomorph: Software for geometric morphometric analyses. R package version 3.3.2. https://cran.r-project.org/package = geomorph.

Aldridge RD, Siegel DS, Aurélien M, Chabarria RE. 2011. Female reproductive anatomy: cloaca, oviduct, and sperm storage. Chapter 9. In: Aldridge RD, Sever DM, editors. Reproductive Biology and Phylogeny of Snakes. 1st ed. Reproductive biology and phylogeny Boca Raton, FL: CRC Press. p. 347–409.

Andonov K, Natchev N, Kornilev YV, Tzankov N. 2017. Does Sexual Selection Influence Ornamentation of Hemipenes in Old World Snakes? Anat Rec 300:1680–94.

Arnqvist G, Rowe L. 2005. Sexual conflict, Monographs in behavior and ecology. Princeton (NJ): Princeton University Press.

Barry FE, Weatherhead PJ, Philipp DP. 1992. Multiple paternity in a wild population of northern water snakes, *Nerodia sipedon*. Behav Ecol Sociobiol 30:193–9.

Boyer DM, Winchester JM, Glynn C, Puente J. 2015. Detailed Anatomical Orientations for Certain Types of Morphometric Measurements Can Be Determined Automatically With Geometric Algorithms: Automatic Alignment for Measurement. Anat Rec 298:1816–23.

Brandley MC, Guiher TJ, Pyron RA, Winne CT, Burbrink FT. 2010. Does dispersal across an aquatic geographic barrier obscure phylogeographic structure in the diamond-backed watersnake (*Nerodia rhombifer*)? Mol Phylogenet Evol 57:552–60.

Brennan PLR, Prum RO, McCracken KG, Sorenson MD, Wilson RE, Birkhead TR. 2007. Coevolution of Male and Female Genital Morphology in Waterfowl. PLoS One 2:e418.

Brennan PLR, Prum RO. 2015. Mechanisms and Evidence of Genital Coevolution: The Roles of Natural Selection, Mate Choice, and Sexual Conflict. Cold Spring Harb Perspect Biol 7:a017749.

Cadle JE. 2011. Hemipenial morphology in the North American snake genus *Phyllorhynchus* (Serpentes: Colubridae), with a review of and comparisons with natricid hemipenes. Zootaxa 3092:1.

Cardini A, Chiapelli M. 2020. How flat can a horse be? Exploring 2D approximations of 3D crania in equids. Zoology 139:125746.

Collyer ML, Adams DC. 2018. RRPP: An r package for fitting linear models to high-dimensional data using residual randomization. Methods Ecol Evol 9:1772–9.

Collyer ML, Adams DC. 2019. RRPP: Linear Model Evaluation with Randomized Residuals in a Permutation Procedure. https://CRAN.R-project.org/package = RRPP

Dowling HG, Savage JM. 1960. A guide to the snake hemipenis: a survey of basic structure and systematic characteristics. Zoologica: scientific contributions of the New York Zoological Society 45:17–28.

Eberhard WG. 1985. Sexual Selection and Animal Genitalia. Cambridge (MA): Harvard University Press.

Edgren RA. 1953. Copulatory adjustment in snakes and its evolutionary implications. Copeia 1953: 162–4.

Friesen CR, Uhrig EJ, Mason RT, Brennan PLR. 2016. Female behaviour and the interaction of male and female genital traits mediate sperm transfer during mating. J Evol Biol 29:952–64.

Friesen CR, Uhrig EJ, Squire MK, Mason RT, Brennan PLR. 2014. Sexual conflict over mating in red-sided garter snakes (*Thamnophis sirtalis*) as indicated by experimental manipulation of genitalia. Proc R Soc B 281:20132694.

Gunnell GF, Boyer DM, Friscia AR, Heritage S, Manthi FK, Miller ER, Sallam HM, Simmons NB, Stevens NJ, Seiffert ER. 2018. Fossil lemurs from Egypt and Kenya suggest an African origin for Madagascar's aye-aye. Nat Commun 9:3193.

Hedrick BP, Antalek-Schrag P, Conith AJ, Natanson LJ, Brennan PLR. 2019. Variability and asymmetry in the shape of the spiny dogfish vagina revealed by 2D and 3D geometric morphometrics. J Zool 308:16–27.

House CM, Simmons LW. 2005. The evolution of male genitalia: patterns of genetic variation and covariation in the genital sclerites of the dung beetle *Onthophagus taurus*: The evolution of genitalia in a dung beetle. J Evol Biol 18:1281–92.

Jadin RC, King RB. 2012. Ontogenetic Effects on Snake Hemipenial Morphology. Journal of Herpetology 46:393–5.

Jadin RC, Parkhill RV. 2011. Hemipenis descriptions of *Mastigodryas* (Serpentes: Colubridae) from northern Middle America, with comments on the use of hemipenial data in phylogenetics. Herpetology Notes 4:207–10.

King RB, Grayburn WS, Duvall MR, Wusterbarth TL, Burghardt GM. 2010. Phylogenetically widespread multiple paternity in New World natricine snakes. Herpetological Conservation and Biology 5:86–93.

King RB, Jadin RC, Grue M, Walley HD. 2009. Behavioural correlates with hemipenis morphology in New World natricine snakes. Biol J Linn Soc 98:110–20.

Kissner KJ, Weatherhead PJ, Gibbs HL. 2005. Experimental assessment of ecological and phenotypic factors affecting male mating success and polyandry in northern watersnakes, *Nerodia sipedon*. Behav Ecol Sociobiol 59:207–14.

Lawson R. 1987. Molecular Studies of Thamnophiine Snakes: 1. The Phylogeny of the Genus *Nerodia*. Journal of Herpetology 21:140.

Lockwood CA, Lynch JM, Kimbel WH. 2002. Quantifying temporal bone morphology of great apes and humans: an approach using geometric morphometrics. J Anat 201:447–64.

McVay JD, Flores-Villela O, Carstens B. 2015. Diversification of North American natricine snakes. Biol J Linn Soc 116:1–12.

Mebert K. 2008. Good species despite massive hybridization: genetic research on the contact zone between the watersnakes *Nerodia sipedon* and *N. fasciata* in the Carolinas, Mol Ecol 17:1918–29.

Orbach DN, Brennan PLR, Hedrick BP, Keener W, Webber MA, Mesnick SL. 2020. Asymmetric and spiraled genitalia coevolve with unique lateralized mating behavior. Sci Rep 10(1): 1–8.

Orbach DN, Hedrick BP, Würsig B, Mesnick SL, Brennan PLR. 2018. The evolution of genital shape variation in female cetaceans. Evolution 72:261–73.

Perez SI, Bernal V, Gonzalez PN. 2006. Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. J Anat 208:769–84.

Pope CH. 1941. Copulatory adjustment in snakes. Field Museum of Natural History. Field Museum of Natural History, Zoology Series 24:249–52.

Prosser MR. 2002. Genetic analysis of the mating system and opportunity for sexual selection in northern water snakes (*Nerodia sipedon*). Behavioral Ecology 13:800–7.

Puniamoorthy N, Kotrba M, Meier R. 2010. Unlocking the "Black box": internal female genitalia in Sepsidae (Diptera) evolve fast and are species-specific. BMC Evol Biol 10: 275.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rohlf FJ, Corti M. 2000. Use of Two-Block Partial Least-Squares to Study Covariation in Shape. Syst Biol 49:740–53.

Rohlf FJ. 2006. TpsDig.

Rolfe S, Pieper S, Porto A, Diamond K, Winchester J, Shan S, Kirveslahti H, Boyer D, Summers A, Maga AM. 2021. SlicerMorph: An open and extensible platform to retrieve, visualize and analyse 3D morphology. Methods Ecol Evol 12: 1816–25.

Sánchez-Martínez PM, Ramírez-Pinilla MP, Miranda-Esquivel DR. 2007. Comparative histology of the vaginal-cloacal region

in Squamata and its phylogenetic implications: Phylogenetic implications of the cloacal histology in Squamata. Acta Zoologica 88:289–307.

Showalter I, Todd BD, Brennan PLR. 2014. Intraspecific and interspecific variation of female genitalia in two species of watersnake: Variation of Female Genitalia. Biol J Linn Soc 111:183–91.

Siegel DS, Miralles A, Trauth SE, Aldridge RD. 2012. The phylogenetic distribution and morphological variation of the 'pouch' in female snakes. Acta Zoologica 93:400–8.

Simmons LW, Fitzpatrick JL. 2019. Female genitalia can evolve more rapidly and divergently than male genitalia. Nat Commun 10:1312.

Singleton M. 2002. Patterns of cranial shape variation in the Papionini (Primates: Cercopithecinae). J Hum Evol 42:547–78.

Sota T, Nagata N. 2008. Diversification in a fluctuating island setting: rapid radiation of *Ohomopterus* ground beetles in the Japanese Islands. Philosophical Transactions of the Royal Society B: Biological Sciences 363:3377–90.

Weatherhead PJ, Barry FE, Brown GP, Forbes MRL. 1995. Sex ratios, mating behavior and sexual size dimorphism of the northern water snake, *Nerodia sipedon*. Behav Ecol Sociobiol 36:301–11.

Zelditch M, Swiderski D, Sheets D. 2012. Geometric morphometrics for biologists: a primer. 2nd ed London: Academic Press.