**Supplemental Data**

Supplemental Data 1.

Specimen preparation, measurement, and individual measurement error

All measured specimens came from museum collections, as the legally protected Na Černidlech site no longer yields *in situ* specimens. The prone specimens all lay parallel to bedding. Most were already fully exposed but some were delicately prepared with a pin to expose covered areas. After coating with ammonium chloride sublimate specimens were photographed directly with a Nikon D100 digital camera and macrolens, through a Nikon SMZ-U stereomicroscope with a Nikon CoolPix995 digital camera or with a Leica MZ16 stereomicroscope with a Leica DFC420 digital camera. The resulting images where digitized using the NIH ImageJ software package (Abràmoff et al., 2004) with the x and y coordinates recorded for each of a series of landmark points on each specimen (Fig. 1). Two-dimensional analysis was employed due to the differential extent of z dimension compaction seen among individuals (see main text).

A scale with half-millimeter divisions was included in each image. Shale preservation inevitably introduces compression-related shape variation into the analysis (see below and Hughes et al., 2014) and, given this, we used a simple estimate of measurement error aimed at capturing the relative magnitude of measurement error compared to within sample variance in size and shape. The raw landmark data was transformed into shape coordinates suitable for assessing differences between specimens using standard techniques (e.g. Hunda and Hughes, 2007). Five specimens of different size and number of thoracic segments were chosen for evaluating the degree of measurement error. These included MCZ114934, MCZ114935, MCZ114944, MCZ115987, and BMNH475176. All were taphonomic grade 2. Each of these five specimens was coated, mounted, photographed, and digitized, and this process was repeated ten times per specimen over the course of several days.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Specimen ID | morph | mean cranidial  length [mm] | standard deviation | coefficient of variation |
| BMNH475176 | 11 | 1.606419 | 0.013396 | 0.833878 % |
| MCZ115987 | 14 | 2.108004 | 0.022652 | 1.074578 % |
| MCZ114934 | 19 | 5.773783 | 0.017109 | 0.296314 % |
| MCZ114935 | 21 | 4.138166 | 0.015362 | 0.371233 % |
| MCZ114944 | 22 | 5.114652 | 0.025710 | 0.502675 % |

Size measurement error estimation from mounting, photographing, and digitizing ten times each of five specimens selected.

Shape variance was calculated using the IMP platform of H.D.S. See Supplementary Data 2 for details.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Specimen ID | morph | distance-based (Foote) disparity | 95% confidence (lower) | 95% confidence (upper) |
| BMNH475176 | 11 | 0.000672 | 0.000399 | 0.000799 |
| MCZ115987 | 14 | 0.000471 | 0.000287 | 0.000556 |
| MCZ114934 | 19 | 0.001471 | 0.000763 | 0.001708 |
| MCZ114935 | 21 | 0.000270 | 0.000125 | 0.000359 |
| MCZ114944 | 22 | 0.000344 | 0.000156 | 0.000423 |

Shape variance as represented by partial Procrustes distances from the mean form of five specimens selected for analysis of measurement error. Each of the five specimens was digitized ten times. Non-reflected 15 landmarks on the cephalon were used in the analysis. Confidence values calculated based on 1600 bootstrap resamples using DisparityBox7.

Mirroring and then averaging paired landmarks across the plane of bilateral symmetry has been a widely accepted method for landmark-based analysis (Fusco et al., 2004; Zelditch et al., 2004) partly because this approach can accommodate missing landmark data on one side only. As all specimens included in this analysis preserved all landmarks the use of reflection to increase sample size was not required for this dataset. Both reflected and non-reflected landmark data were used in the analyses, and the results were compared.

Abràmoff, M. D., P. J. Magalhães, and S. J. Ram. 2004. Image Processing with ImageJ. Biophotonics International, 11:36–42.

Hunda, B. R. and N. C. Hughes. 2007. Evaluating paedomorphic heterochrony in trilobites: the case of the diminutive trilobite *Flexicalymene* *retrorsa* *minuens* from the Cincinnatian Series (Upper Ordovician), Cincinnati region. Evolution and Development, 9:483–498.

Supplemental Data 2.

ANALYTICAL TOOLS

The Integrated Morphometrics package is a set of compiled software tools for displaying and analyzing 2-D landmark-based geometric morphometric data available at: (<http://www.canisius.edu/~sheets/morphsoft.html>)

IMP series programs used in the study:

1. Program CoordGen7 changes raw mm-scale x and y coordinates of specimens into coordinates of generalized least square Procrustes superimposition. Each specimen's coordinates are first translated to have its centroid at origin, then coordinates are scaled to unit centroid size, and specimens are rotated to minimize the summed squared distances between homologous landmarks (partial Procrustes distances).

2. Program DisparityBox7 measures morphological diversity of a group based on landmark data. In this study, each specimen's partial Procrustes distance from the mean shape is used as a unit of shape variance within an instar. The "bootstrap within-group disparity" option is used to calculate the variance and the 95% confidence interval.

3. Program PCAGen7 first translates the shape differences between individual specimen and the mean shape in a Procrustes superimposition into mathematically independent style of deformation (called the partial warp scores) that also sets the correct number of degree of freedom for statistical tests, and then runs principal component analysis on the partial warp scores. Each principal component is called a relative warp, and PCAGen7 measures which relative warp explains what percentage of the total shape variance.

4. In program Regress7, partial warp scores are regressed in a multivariate regression against ln centroid size to produce a vector of regression coefficients. Proportion of total shape variance explained by this allometry is then calculated, and significance of the multivariate regression is determined by bootstrapping method. Partial Procrustes distance can also be calculated through the program.

5. Program TwoGroup7 measures partial Procrustes distances between group mean shapes, and compares significance of within-group and between-group variances through bootstrapping method.

6. Program VecCompare7 tests significance of the difference between two vectors of regression coefficients by comparing the angle between the two groups and within-group angles through bootstrapping method. It could also calculate whether the growth vector is significantly different from isometry.

7. Program BigFix7 reflects paired homologous landmarks across an axial midline and calculate the average position for each paired landmarks.

Supplemental Data 3.

Specimen REPOSITORY INFORMATION

The principal holdings of *A. kokinckii* are as follows: National Museum, Prague (total holdings ca. 10,000 specimens), the Museum of Comparative Zoology at Harvard University (783 specimens), the Czech Geological Survey (ca. 200 specimens), the US National Museum of Natural History (55 specimens), the British Natural of Natural History (ca. 50 specimens), and the American Museum of Natural History (ca. 10 specimens). Many other museums throughout the world also hold representative collections.

Details on the specimen number, collector, and locality information (N=352).

1. National Museum of Prague (N=166)

NMP Lxxxxx - Lodenice-Cernidla (N=166)

2. Museum of Comparative Zoology at Harvard University (N=142)

MCZ 1xxxxx - Lodenice, Schary collection (N=142)

3. Czech Geological Survey (N=21)

CGS OTxxxx - Lodenice-Cernidla, Vanek collection (N=4)

CGS Pxxxx - Lodenice-Cernidla, Horný collection, 1976 (N=17)

4. US National Museum of Natural History (N=12)

USNM xxxxxx - Lodenitz (N=12)

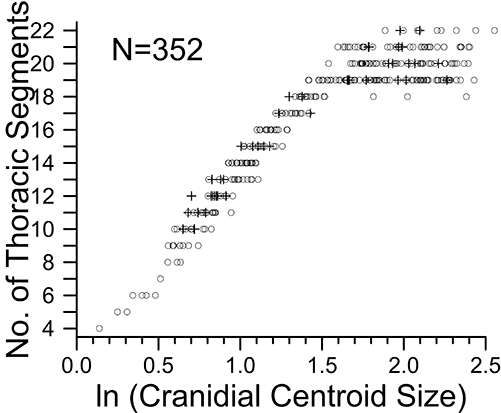
5. British Natural of Natural History (“Natural History Museum”) (N=11)

BMNH xxxxx.x - Lodenitz (N=11)

Supplemental Data 4.

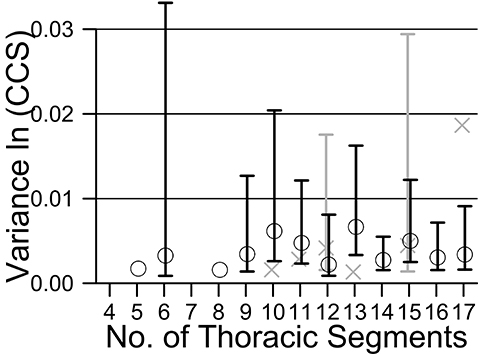
TAPHONOMIC GRADE and its effect on morphological variation

Taphonomic grade 1 included only near “perfect” specimens such as those with minor flaws unrelated to the landmarks positions, such as slight damage to the edge of the fossil, and a damaged eye surface etc, which apparently had no effect on landmark position. These specimens also appear to have suffered no compressional flattening. Grade 2 material are specimens that do not have apparent shape distortion of the x,y coordinates of landmark points, have the axes of the cranidia, pygidia, and thoracic segments in precise, parallel alignment, have all pairs of landmarks present, and possess only minor cracks, but are qualitatively slightly less well preserved than those within Grade 1, and in which we think minor flattening may have occurred.



Logarithm of cranidial centroid size for the two preservational grades of the new dataset showing that the two grades do not differ markedly in variance. Among the 15 landmarks of the cranidium, 6 paired homologous landmarks were reflected across an axial midline for the analysis and averaged. Crosses are Grade 1 material, and open circles indicate Grade 2 specimens.

This result suggests that Grade 1 and 2 materials are broadly comparable.

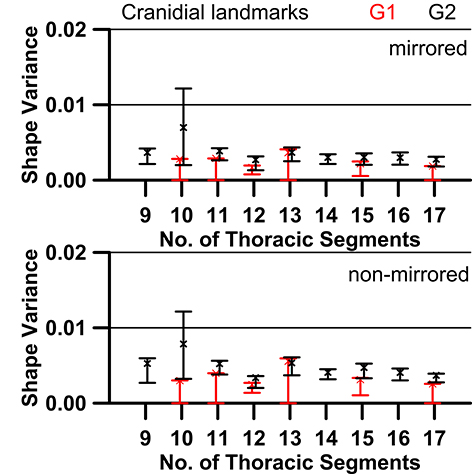


Variance of logarithm of cranidial centroid size for the two preservational grades of the new dataset suggesting that both grades show similar variance. Among the 15 landmarks of the cranidium, 6 paired homologous landmarks were reflected across an axial midline for the analysis and averaged, using the program BigFix7 in the morphometrics series IMP. Crosses are Grade 1 material, and open circles indicate Grade 2 specimens. Bars are 95% confidence intervals by the method of shortest unbiased confidence intervals (Sokal and Rohlf, 1981).

Accodingly, Grade 1 and 2 size variance cannot be distinguished at these sample sizes.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| morph | Grade 1 Variance | Grade 1 n | Grade 2 Variance | Grade 2 n | F | P |
| 10 | 0.001594 | 3 | 0.006161 | 9 | 3.864070 | 0.443592 |
| 11 | 0.002837 | 3 | 0.004755 | 13 | 1.676275 | 0.867515 |
| 12 | 0.004167 | 7 | 0.002203 | 8 | 1.891765 | 0.423577 |
| 13 | 0.001338 | 3 | 0.006648 | 14 | 4.966794 | **0.049956** |
| 15 | 0.004463 | 5 | 0.004993 | 14 | 1.118769 | 1.000000 |
| 17 | 0.018650 | 2 | 0.003393 | 12 | 5.496959 | 0.077717 |

Two-tailed F-test at 95% confidence level of differences in size variance between grade 1 and 2 for those morphs with sufficient sample size in both grades from the new dataset. Among the 15 landmarks in the cephalon, 6 paired homologous landmarks were reflected across an axial midline for the analysis. “Morph” refers to the number of thoracic segments, n to the number of individuals. No significant difference in variance is detected between the grade 1 and 2 materials except for the 13 thoracic segment number morph.



For the mirrored cranidial landmarks, shape variances of G1 and G2 overlap at 95% confidence level at meraspid degree 10, 11, 12, 13,15, and 17. For the non-mirrored data, they are slightly separated at meraspid degree 10 and 17, but amount of separation is minor at about one fifth of the standard error.

When the size variance of ln CCS (mirrored) value for each meraspid stage (10, 11, 12, 15, 17) is compared for G1 and G2, there was no significant difference. Comparison of the instar variance in the cranidial shape of specimens belonging to grades 1 and 2, computed via the partial Procrustes distance of individual specimen to the consensus of all specimens in Procrustes superimposition for each instar (Zelditch et al., 2004, p. 297–302), revealed that there was no significant difference.

Supplemental Data 5.

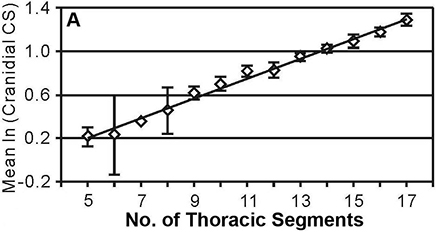
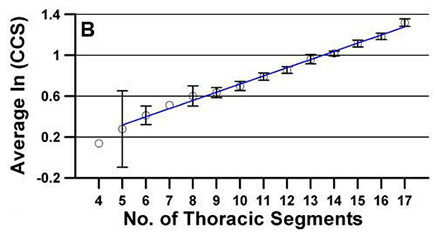
SPECIMENS IN THE SAMPLE AND ARTICULATION-BASED ONTOGENETIC STAGES

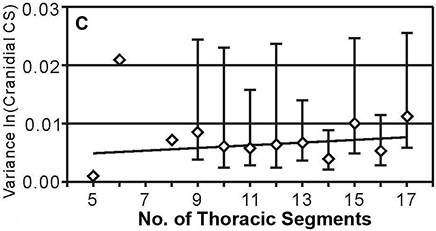
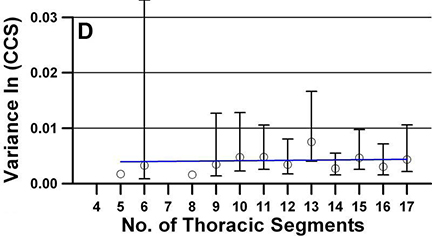
Since onset of the holaspid phase is defined by the cessation of thoracic segment addition but there are variations in number of thoracic segments in the holaspid phase for *A. koninckii* (Fusco et al., 2004), some of the smaller remaining 204 specimens (58.0%) with 18 or more thoracic segments will be late-stage meraspids. Hence below a certain size threshold (natural logarithm of cranidial centroid size [Bookstein, 1991] of 2.2 [see Fusco et al., 2004]) we cannot be confident that all specimens with more than 18 segments are holaspides. If the observed meraspid growth rate were to characterized the entire meraspid and hlaspid phases, the largest holaspid would represent an individual that had gone through about 32 molts since meraspid degree 0.

Bookstein, F. L. 1991. Morphometric Tools for Landmark Data. Cambridge University Press, New York, 435 p.

Supplemental Data 6.

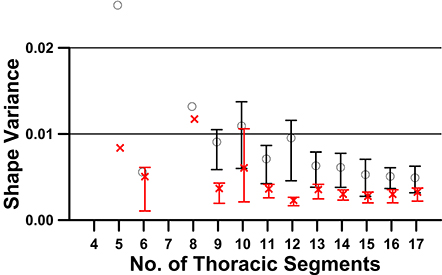
SIZE AND SHAPE VARIATION IN THE FUSCO ET AL. (2004) DATASET AND THE NEW DATASET COMPARED

Fusco et al., 2004 New Dataset

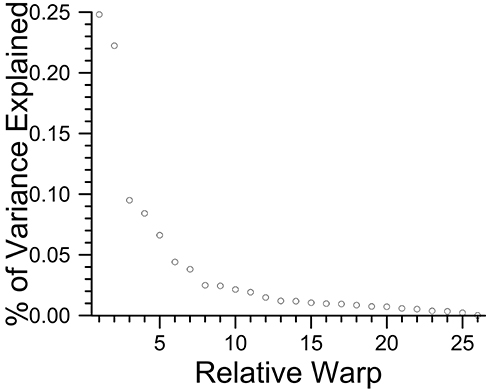
Growth (A,B) and variance (C,D) values for cranidial centroid size (CCS) for the previous (A,C) and new (B,D) datasets. The BigFix6 program from the IMP series was used to mirror fifteen landmarks into nine landmarks. Diamonds and circles are the mean logarithm of the CCS for each of the different degrees. Regression lines are shown. Bars in the variance plots are 95% confidence intervals (not calculated for stages with three or less specimens available), method of shortest unbiased confidence intervals (Sokal and Rohlf, 1981).



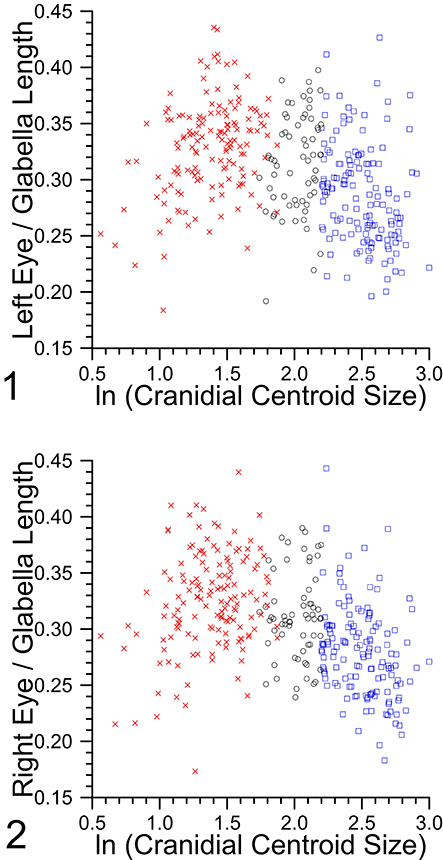
Shape variance values for the cranidium for the previous dataset and the new dataset. Among the 15 landmarks in the cranidium, 6 paired homologous landmarks were reflected across an axial midline for the analysis and averaged, using the program BigFix7 in the morphometrics series IMP. Red x’s represent the new dataset, and black circles indicate specimens from the previous dataset. Bars are 95% confidence intervals calculated based on 1600 bootstrap resamples using DisparityBox7.

Supplemental Data 7.

GROWTH AND VARIATION IN THE MERASPID AND HOLASPID CRANIDIUM



Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for all meraspid and holaspid specimens (N=352). Relative warp (RW) 1 and RW2 each explains 24.80% and 22.23% of total variance.



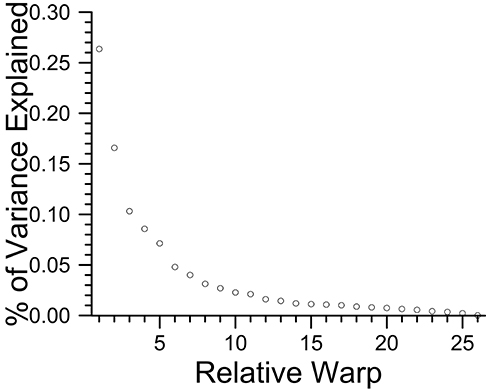
Bivariate plot of eye length-glabellar length ratio against logarithm of cranidial centroid size, red x's are merapid specimens, blue squares are holaspid specimens with ln CCS larger than 2.2, and black circles are specimens with 18 or more thoracic segments and with ln CCS smaller than 2.2: *1*, Left eye; *2*, Right eye.

Growth rate per molt of glabellar length is estimated by using the antilogarithm of the regression coefficient between the mean logarithm of the glabellar length at a certain meraspid stage and the stage number from meraspid degrees 9 through 17. The growth rate is 1.076 and confidence interval at the 95% level is from 1.070 to 1.082. Assuming similar exponential growth for the eye length, the growth rates for left and right eyes were 1.082 (1.072-1.092) and 1.0752 (1.065-1.086) with their confidence interval in parentheses. Therefore, at the 95% confidence level, growth rate for the glabellar length is not significantly different from that of eye length.

When the RMA regression method is used on the bivariate plots of logarithm of eye length against logarithm of glabellar length (Hammer and Harper, 2006, p. 151) for the same ontogenetic period as above, the slope values for the left eye (1.103; 1.000-1.203) and for the right eye (1.089; 0.937-1.210) were not significantly different from 1. Holaspid specimens also show the same results for the morph 19 (left 1.0833, 0.887-1.288: right 0.927, 0.765-1.075), the morph 20 (left 0.884, 0.670-1.081; right 0.892, 0.675-1.076), and the morph 21 (left 1.011, 0.753-1.225; right 1.018, 0.763-1.120). Specimens with logarithm of cranidial centroid size being above 2.2 were chosen for the analysis, and morphs 18 and 22 were excluded due to small sample size.

Supplemental Data 8.

GROWTH AND VARIATION IN THE MERASPID CRANIDIUM



Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks of meraspides from meraspid degree 4 through 17 (N=148). Relative warp (RW) 1, RW2, and RW3 each explains 26.35%, 16.58%, and 10.31% of total variance.

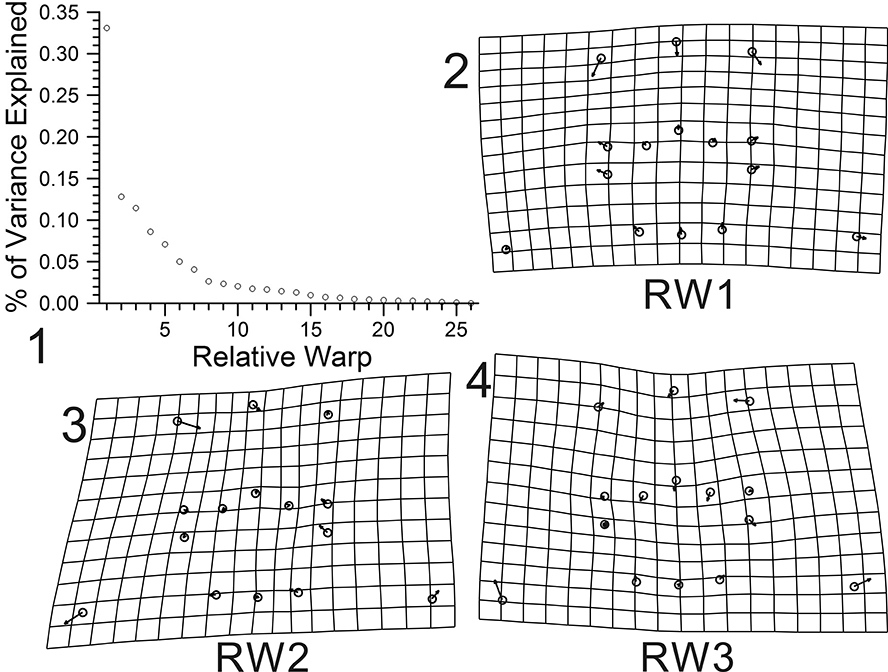
Comparsions of the partial Procrustes distance of cranidial landmarks from the reference specimen (the consensus of the three smallest specimens in the entire dataset) can only be detected when four consecutive meraspid stages are compared – see table below. Results also suggest there is no apparent size-related trend in the degree of morphological difference, as might be expected if the degree of instar-related shape change altered during meraspid ontogeny.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Group1 | Group2 | n | slope | min | max | p-value |
| 12 | 15 | 72 | 0.0479 | 0.0096 | 0.0861 | 0.0074 |
| 13 | 16 | 72 | 0.0498 | 0.0089 | 0.0907 | 0.0088 |
| 14 | 17 | 69 | 0.0573 | 0.0181 | 0.0965 | 0.0024 |
| 11 | 15 | 88 | 0.0381 | 0.0098 | 0.0665 | 0.0045 |
| 12 | 16 | 87 | 0.0521 | 0.0224 | 0.0817 | 0.0004 |
| 13 | 17 | 86 | 0.0396 | 0.0093 | 0.0817 | 0.0056 |
| 10 | 15 | 100 | 0.0242 | 0.0001 | 0.0484 | 0.0245 |
| 11 | 16 | 103 | 0.0432 | 0.0203 | 0.0660 | 0.0001 |
| 12 | 17 | 101 | 0.0431 | 0.0195 | 0.0667 | 0.0002 |
| 9 | 15 | 108 | 0.0276 | 0.0066 | 0.0486 | 0.0053 |
| 10 | 16 | 115 | 0.0310 | 0.0111 | 0.0510 | 0.0013 |
| 11 | 17 | 117 | 0.0383 | 0.0193 | 0.0572 | 0.0001 |
| 9 | 16 | 123 | 0.0330 | 0.0154 | 0.0506 | 0.0002 |
| 10 | 17 | 129 | 0.0300 | 0.0131 | 0.0468 | 0.0003 |
| 9 | 17 | 137 | 0.0316 | 0.0165 | 0.0467 | 0.0000 |

Significant relationships between Partial Procrustes distance with ln CCS among meraspid stages. Only the comparisions that show significant increase in partial Procrustes distance with ln CCS are included in the table. The "Group1" and "Group2" labels are the range of continuous meraspid stages in meraspid degree numbers. "n" is the sample size, and the "slope" is the regression coefficient of partial Procrustes distance (consensus of smallest three specimens as reference) against ln CCS. The "min" and "max" labels are the lower and upper 95% confidence limits for the slope, and "p-value" is from the significance test of the slope value.

Supplemental Data 9.

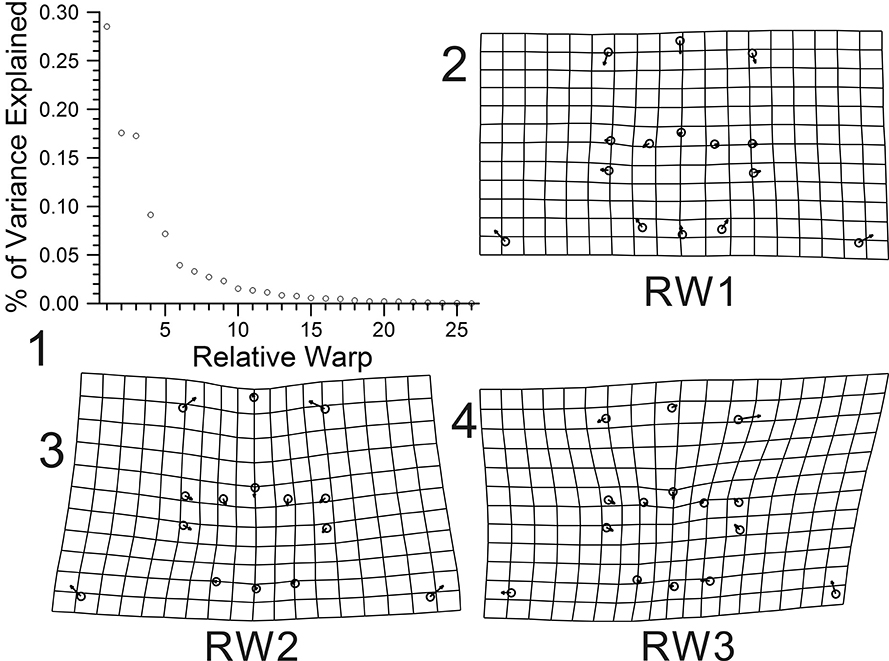
GROWTH AND VARIATION IN THE HOLASPID CRANIDIUM



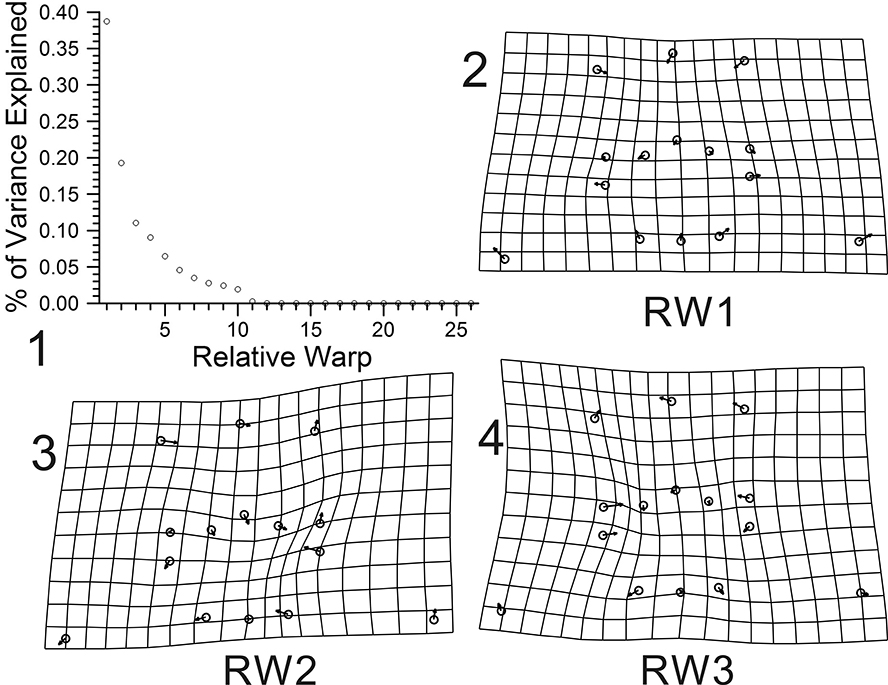
Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for holaspid specimens of morph 19 with its ln CCS value more than 2.2 (N=54): *1*, Relative warp (RW) 1 explains 33.09% of total variance. Likewise, RW2 12.80%, and RW3 11.45%; *2*, Shape variation related to relative warp 1 (depicting very subtle elongation of the cranidium); *3*, Shape variation related to relative warp 2 (depicting effects of shearing); *4*, Shape variation related to relative warp 3 (depicting size of the pleural region relative to the glabella and the palpebral lobes).



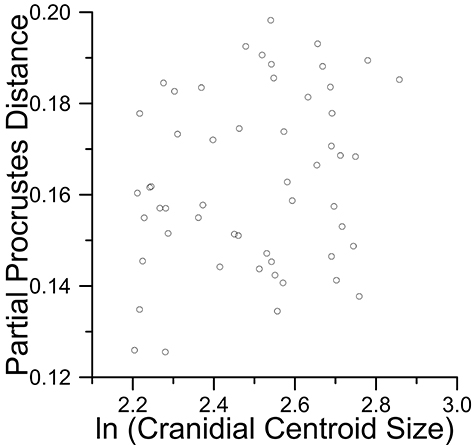
Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for holaspid specimens of morph 20 with its ln CCS value more than 2.2 (N=42): *1*, Relative warp (RW) 1 explains 27.06% of total variance. Likewise, RW2 22.54%, and RW3 11.97%; *2*, Shape variation related to relative warp 1 (depicting size of the pleural region relative to the glabella and the palpebral lobes); *3*, Shape variation related to relative warp 2 (depicting subtle elongation of the cranidium); *4*, Shape variation related to relative warp 3 (depicting effects of shearing, direction inversed).



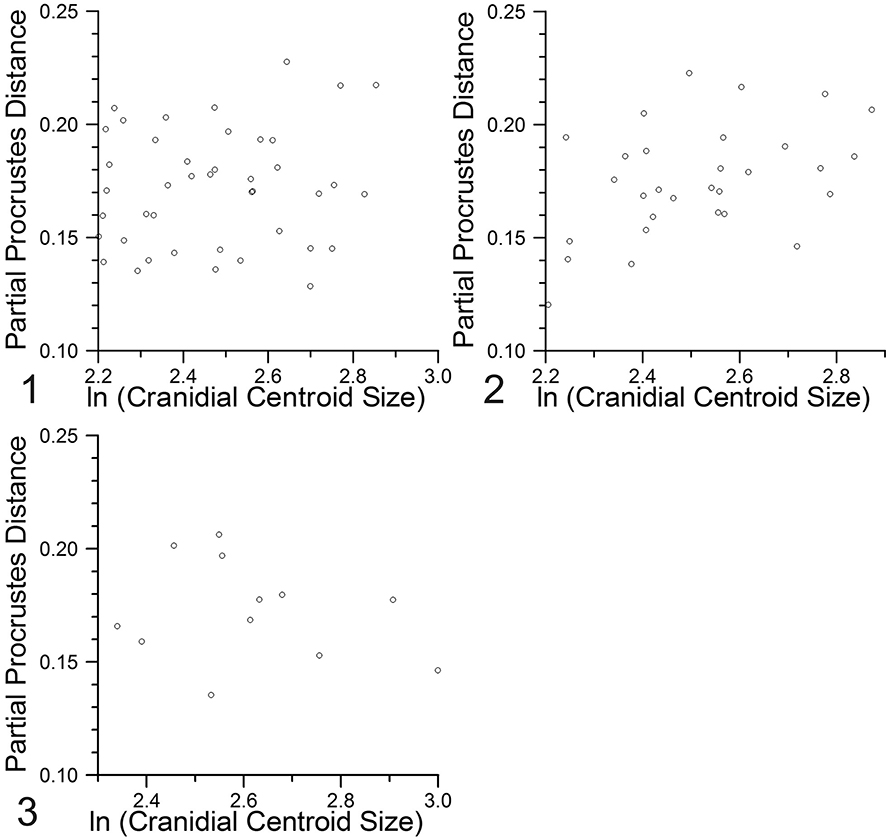
Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for holaspid specimens of morph 21 with its ln CCS value more than 2.2 (N=30): *1*, Relative warp (RW) 1 explains 28.50% of total variance. Likewise, RW2 17.58%, and RW3 17.27%; *2*, Shape variation related to relative warp 1 (depicting subtle elongation of the cranidium); *3*, Shape variation related to relative warp 2 (depicting size of the pleural region relative to the glabella and the palpebral lobes, direction inversed); *4*, Shape variation related to relative warp 3 (depicting effects of shearing).



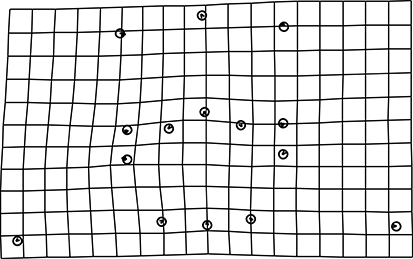
Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for holaspid specimens of morph 22 with its ln CCS value more than 2.2 (N=12): *1*, Relative warp (RW) 1 explains 38.72% of total variance. Likewise, RW2 19.29%, and RW3 11.04%; *2*, Shape variation related to relative warp 1 (depicting subtle elongation of the cranidium); *3*, Shape variation related to relative warp 2 (depicting effects of shearing); *4*, Shape variation related to relative warp 3 (depicting size of the pleural region relative to the glabella and the palpebral lobes, direction inversed).



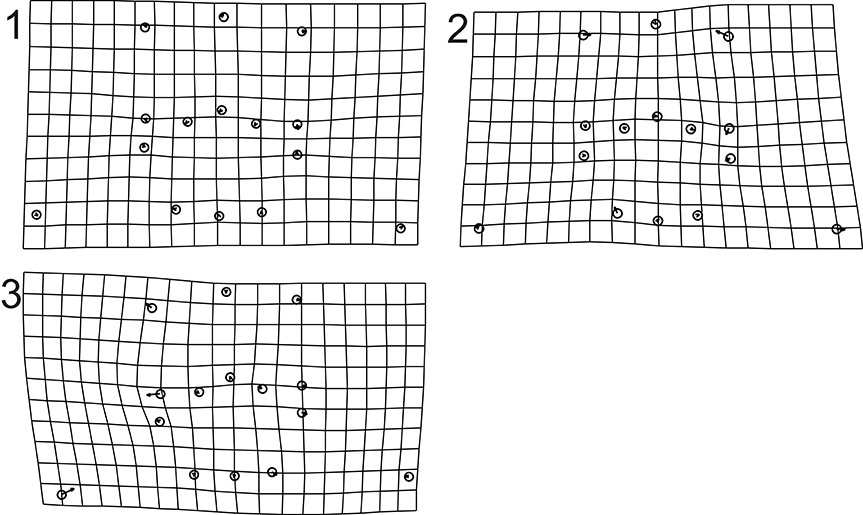
Partial Procrustes distance from the reference (mean shape of the smallest three specimens from the meraspid phase) of 15 cranidial landmarks for holaspid specimens of morph 19 with its ln CCS value more than 2.2 (N=54). Regression of partial Procrustes distance against logarithm of cranidial centroid size is not significant at the 95% confidence level (slope=0.0224, p=0.0543, r=0.2210)



Partial Procrustes distance from the reference (mean shape of the smallest three meraspid specimens) of 15 cranidial landmarks for holaspid specimens of morphs 20 to 22 with its ln CCS value more than 2.2: *1*, Morph 20 (N=42), regression of partial Procrustes distance against logarithm of cranidial centroid size is not significant at the 95% confidence level (slope=0.0209, p=0.2026, r=0.1325). *2*, Morph 21 (N=30), regression of partial Procrustes distance against logarithm of cranidial centroid size is significant at the 95% confidence level (slope=0.0566, p=0.008550, r=0.4292); *3*, Morph 22 (N=12), regression of partial Procrustes distance against logarithm of cranidial centroid size is not significant at the 95% confidence level (slope=-0.0263, p=0.7629, r=-0.2341).



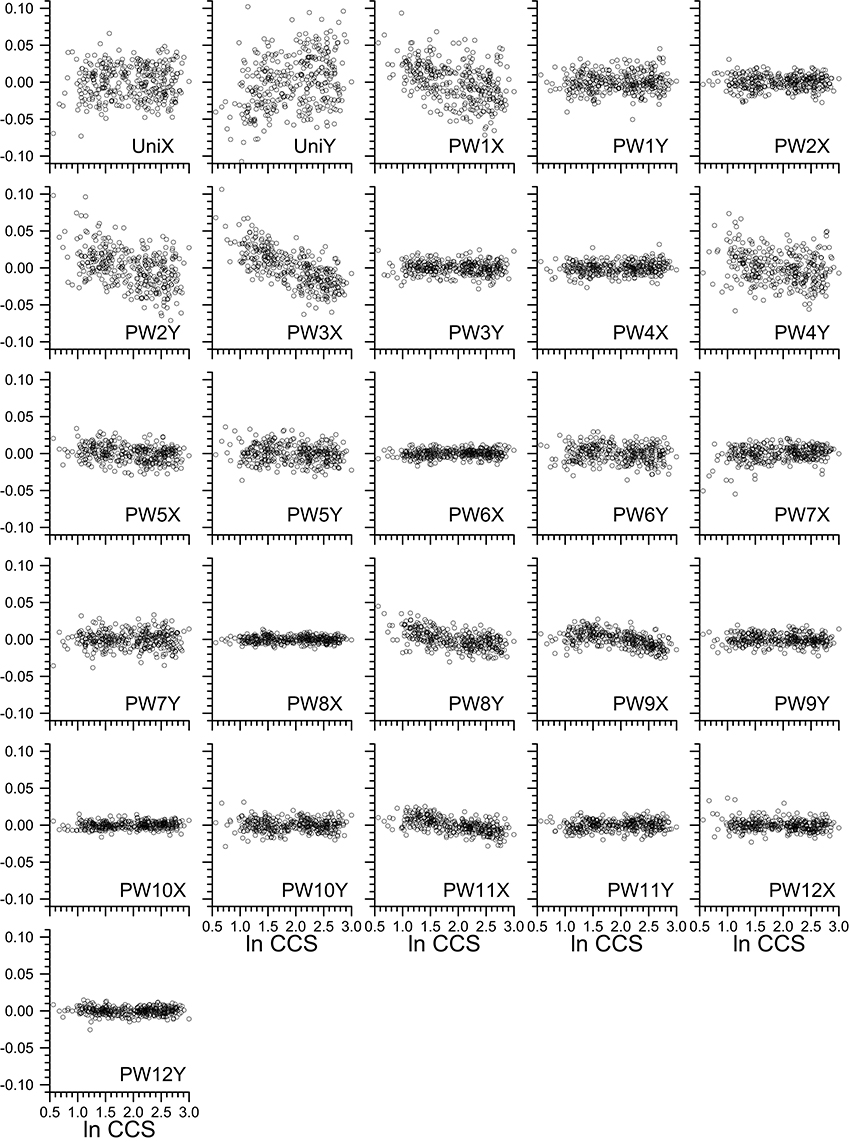
Thin-plate spline deformation grid of shape changes with growth of 15 cranidial landmarks for holaspid specimens of morph 19 with its ln CCS value more than 2.2 (N=54). Partial warp scores are regressed in a multivariate regression against ln centroid size, and total shape variance (based on summed squared residuals expressed in Procrustes units) explained by the allometry is not significant at the 95% confidence level (1.5045% of total variance explained, p=0. 578125 from 1600 bootstraps).



Thin-plate spline deformation grid of shape changes with growth of 15 cranidial landmarks for holaspid specimens of morphs 20 to 22 with its ln CCS value more than 2.2. Partial warp scores are regressed in a multivariate regression against ln centroid size, and percentages are total shape variance (based on summed squared residuals expressed in Procrustes units) explained by the allometry: *1*, Morph 20 (N=42), allometry not significant at the 95% confidence level (1.82% of total variance explained, p=0. 6525 from 1600 bootstraps); *2*, Morph 21 (N=30), allometry not significant at the 95% confidence level (4.49% of total variance explained, p=0. 2175 from 1600 bootstraps); *3*, Morph 22 (N=12), allometry not significant at the 95% confidence level (7.50% of total variance explained, p=0.56625 from 1600 bootstraps).

Supplemental Data 10.

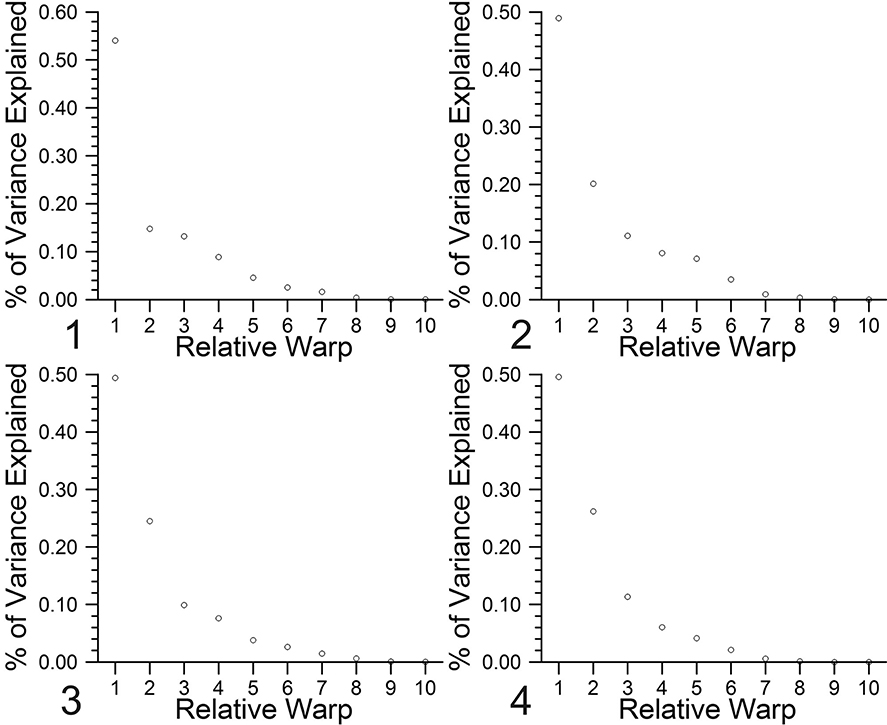
PARTIAL WARP SCORES FOR THE MERASPID AND HOLASPID CRANIDIUM



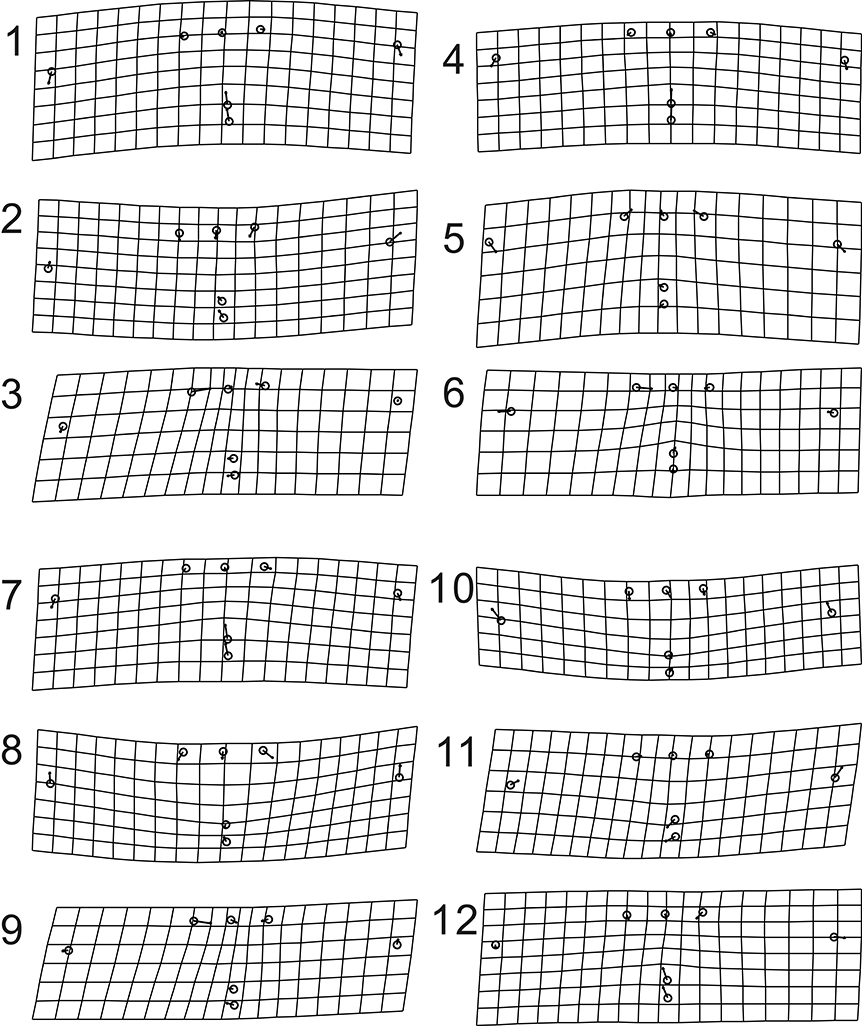
Bivariate plot of partial warp scores against ln CCS for the 15 cranidial landmarks of meraspides and holaspides. UniX and UniY are uniform components of X and Y, and PW#X and PW#Y are non-uniform components of X and Y.

Supplemental Data 11.

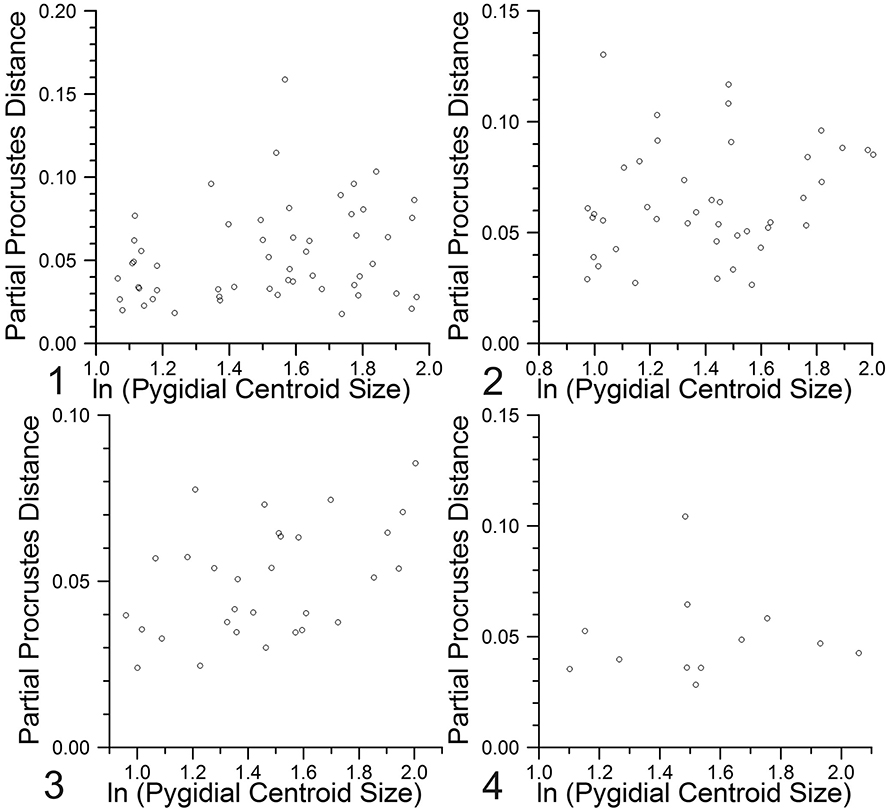
GROWTH AND VARIATION IN THE HOLASPID PYGIDIUM



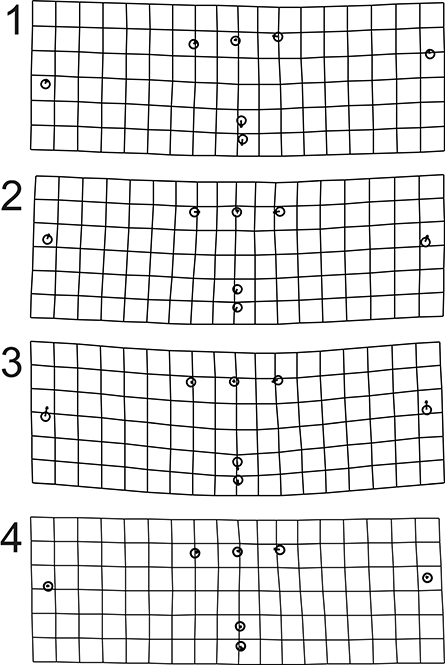
Bivariate plot of percentage of total variance explained by each relative warp of 7 pygidial landmarks for specimens of morphs 19 to 22 with its ln cranidial centroid size value being over 2.2. Percentage of total shape variance (based on summed squared residuals expressed in Procrustes units) explained by the allometry in parentheses: *1*, Morph 19 (N=54), Relative Warp (RW) 1 (54.03%), RW2 (14.78%), RW3 (13.19%); *2*, Morph 20 (N=42), RW1 (48.92%), RW2 (20.12%), RW3 (11.08%); *3*, Morph 21 (N=30), RW1 (49.38%), RW2 (24.51%), RW3 (9.90%); *4*, Morph 22 (N=12), RW1 (49.58%), RW2 (26.19%), RW3 (11.32%).



Bivariate plot of percentage of total variance explained by each relative warp of 15 cranidial landmarks for holaspid specimens of morphs 19 to 22 with its ln CCS value more than 2.2. Partial warp scores are regressed in a multivariate regression against ln pygidial centroid size, and percentage of total shape variance (based on summed squared residuals expressed in Procrustes units) explained by the allometry in parentheses: *1*–*3*, Morph 19 (N=54); *1*, Relative Warp (RW) 1 (54.03%); *2*, RW2 (14.78%); *3*, RW3 (13.19%); *4*–*6*, Morph 20 (N=42); *4*, RW1 (48.92%); *5*, RW2 (20.12%); *6*, RW3 (11.08%); *7*–*9*, Morph 21 (N=30); *7*, RW1 (49.38%); *8*, RW2 (24.51%); *9*, RW3 (9.90%); *10*–*12*, Morph 22 (N=12); *10*, RW1 (49.58%); *11*, RW2 (26.19%); *12*, RW3 (11.32%); *1*, *4*, *7*, *10*, Shape changes related to RW1 is arching of the whole pygidium and variations in the axis-direction distance between mid-posterior end of the pygidium and the anterolateral tips of the pygidium; *2*, *5*, *8*, *11*, RW2 represents variations in the relative length of the pygidium; *3*, *6*, *9*, *12*, RW3 represents variations in the width of the pygidial axis.



Partial Procrustes distance from the reference (mean shape of the smallest three specimens in each holaspid sample) of 7 pygidial landmarks for holaspid specimens of morphs 19 to 22 with its ln CCS value more than 2.2. Regression of partial Procrustes distance against logarithm of pygidial centroid size is significant for morphs 19 and 21. *1*, Morph 19 (slope=0.0241, p=0.0395, r=0.2412, N=54); *2*, Morph 20 (slope=0.0176, p=0.09248, r=0.2091, N=42); *3*, Morph 21 (slope=0.0262, p=0.0044, r=0.4649, N=30); *4*, Morph 22 (slope=0.0025, p=0.4577, r=0.0354, N=12).



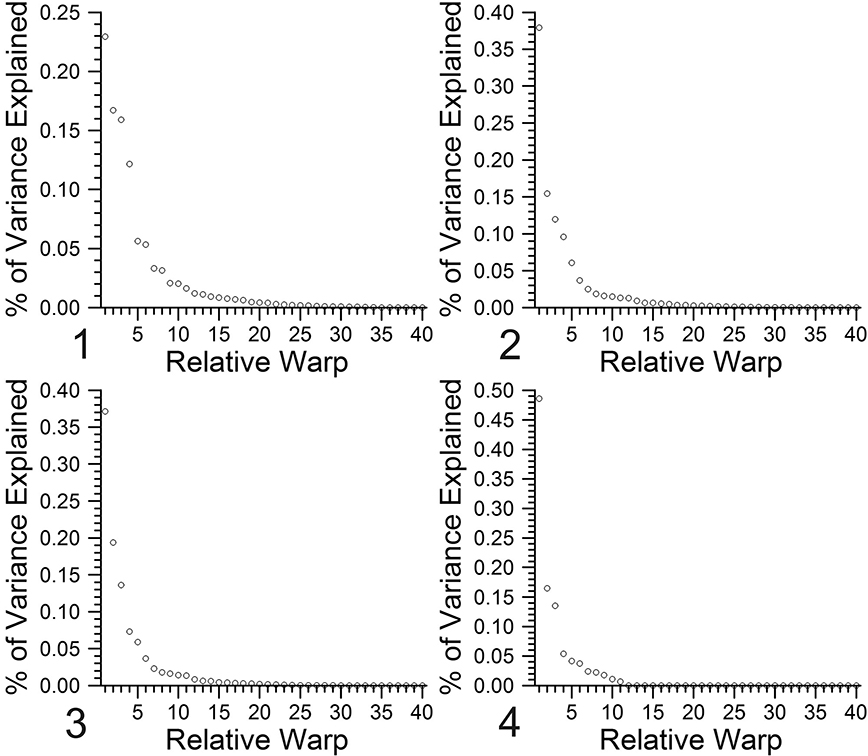
Thin-plate spline deformation grid of shape changes with growth of 7 pygidial landmarks for holaspid specimens of morphs 19 to 22 with its ln CCS value more than 2.2. Partial warp scores are regressed in a multivariate regression against ln pygidial centroid size. *1*, Morph 19, 9.40% of total shape variance (based on summed squared residuals expressed in Procrustes units) is explained by the allometry (p=0.001875 from 1600 bootstraps); *2*, Morph 20, 8.86% of total shape variance (based on summed squared residuals expressed in Procrustes units) is explained by the allometry (p=0.0075 from 1600 bootstraps); *3*, Morph 21, 15.32% of total shape variance is explained by the allometry (p=0.003750 from 1600 bootstraps); *4*, Morph 22, 3.47% of total shape variance is explained by the allometry (p=0.08675 from 1600 bootstraps).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Holaspides  (ln CCS > 2.2) | N | angle to  isometry | within  sample | Significant  difference |
| Morph 19 | 54 | 99.1° | 46.7° | yes |
| Morph 20 | 42 | 103.5° | 56.7° | yes |
| Morph 21 | 30 | 81.9° | 49.6° | yes |
| Morph 22 | 12 | 97.9° | 111.3° | no |

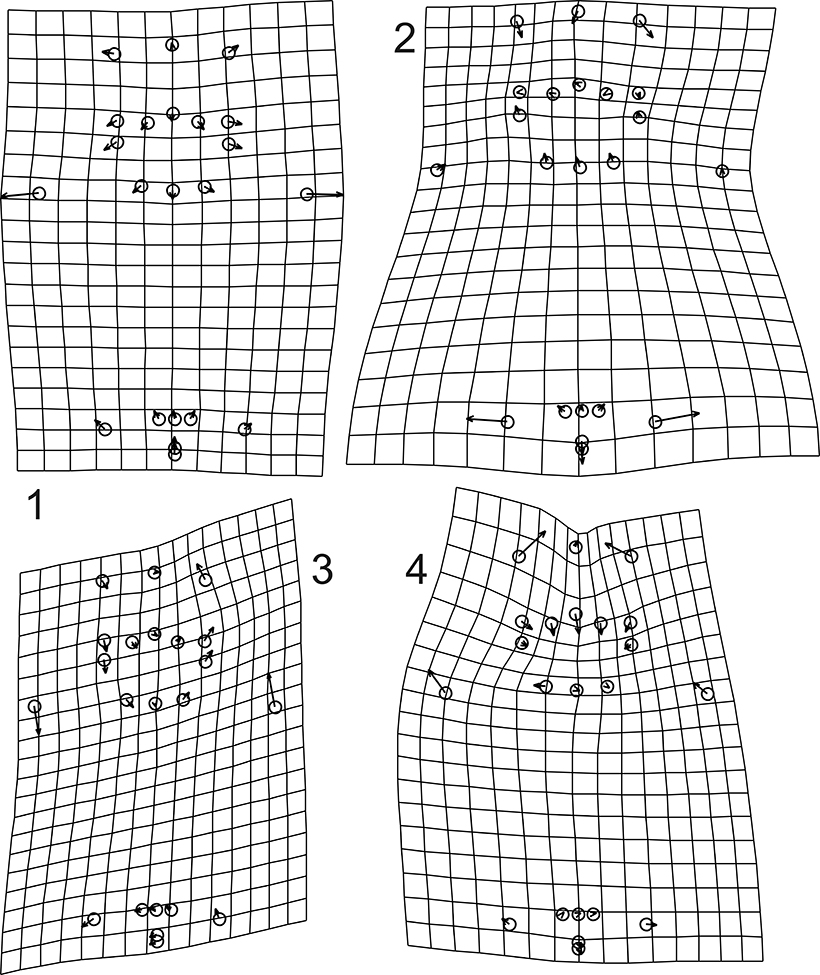
The vectors of regression coefficients (calculated from regressing partial warp scores in a multivariate regression against ln CCS; mean shape is used as a reference for the partial warp scores) for the 7 pygidial landmarks of holaspid specimens of morphs 19 to 22 are compared with isometry. Angle of the normalized vector within the sample and angle to isometry is compared at the 95% confidence level by 1600 bootstraps.

Supplemental Data 12.

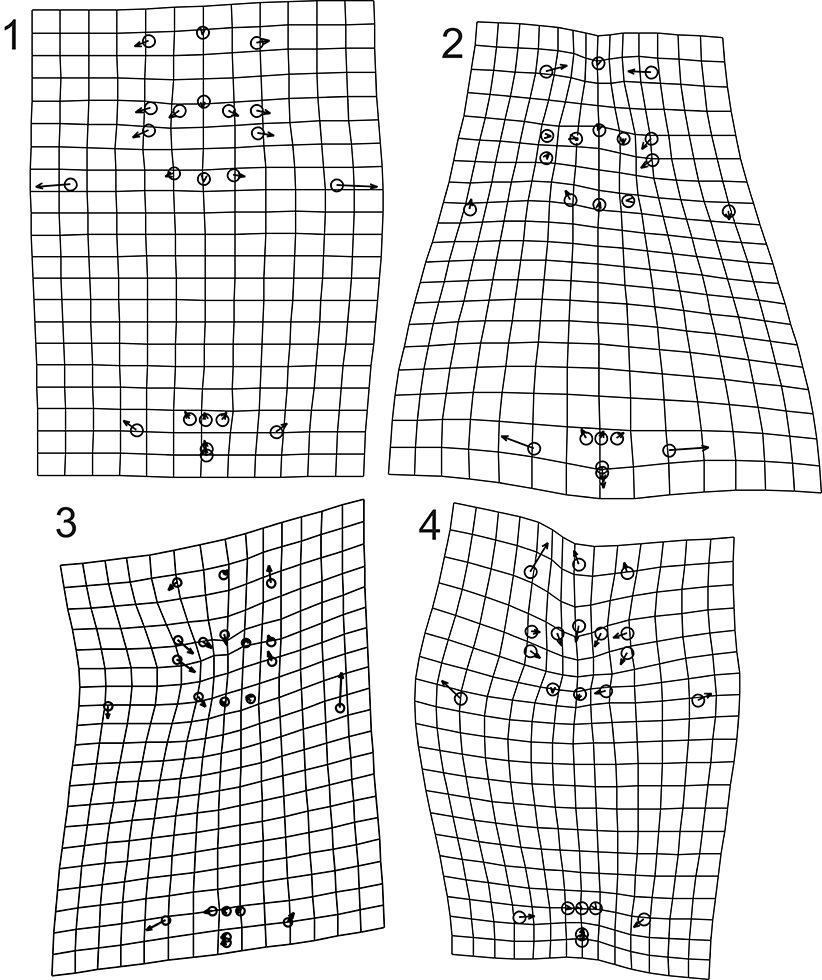
VARIATION AND GROWTH IN THE HOLASPID EXOSKELETON



Bivariate plot of percentage of total variance explained by each relative warp of 22 exoskeletal landmarks for specimens of morphs 19 to 22 with its ln cranidial centroid size value being over 2.2. Percentage of total shape variance (based on summed squared residuals expressed in Procrustes units) explained by the allometry in parentheses: *1*, Morph 19 (N=54), Relative Warp (RW) 1 (22.94%), RW2 (16.71%), RW3 (15.90%), RW4 (12.16%); *2*, Morph 20 (N=42), RW1 (37.93%), RW2 (15.45%), RW3 (11.96%), RW4 (9.60%); *3*, Morph 21 (N=30), RW1 (37.14%), RW2 (19.38%), RW3 (13.61%), RW4 (7.31%); *4*, Morph 22 (N=12), RW1 (48.57%), RW2 (16.47%), RW3 (13.51%), RW4 (5.37%).



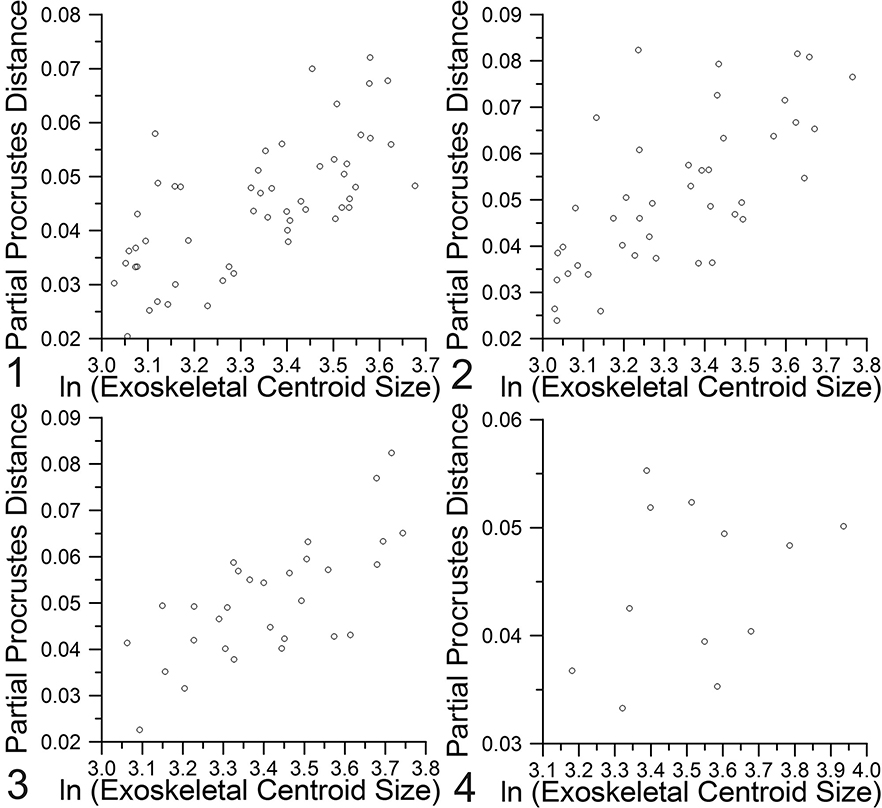
Thin-plate spline deformation grid of relative warps of 22 exoskeletal landmarks for holaspid specimens of morph 20 with its ln CCS value more than 2.2 (N=42): *1*, Shape variation related to relative warp 1 (37.93% of total variance explained); *2*, Shape variation related to relative warp 2 (15.45% of total variance explained, direction inversed); *3*, Shape variation related to relative warp 3 (11.96% of total variance explained); *4*, Shape variation related to relative warp 4 (9.60% of total variance explained).



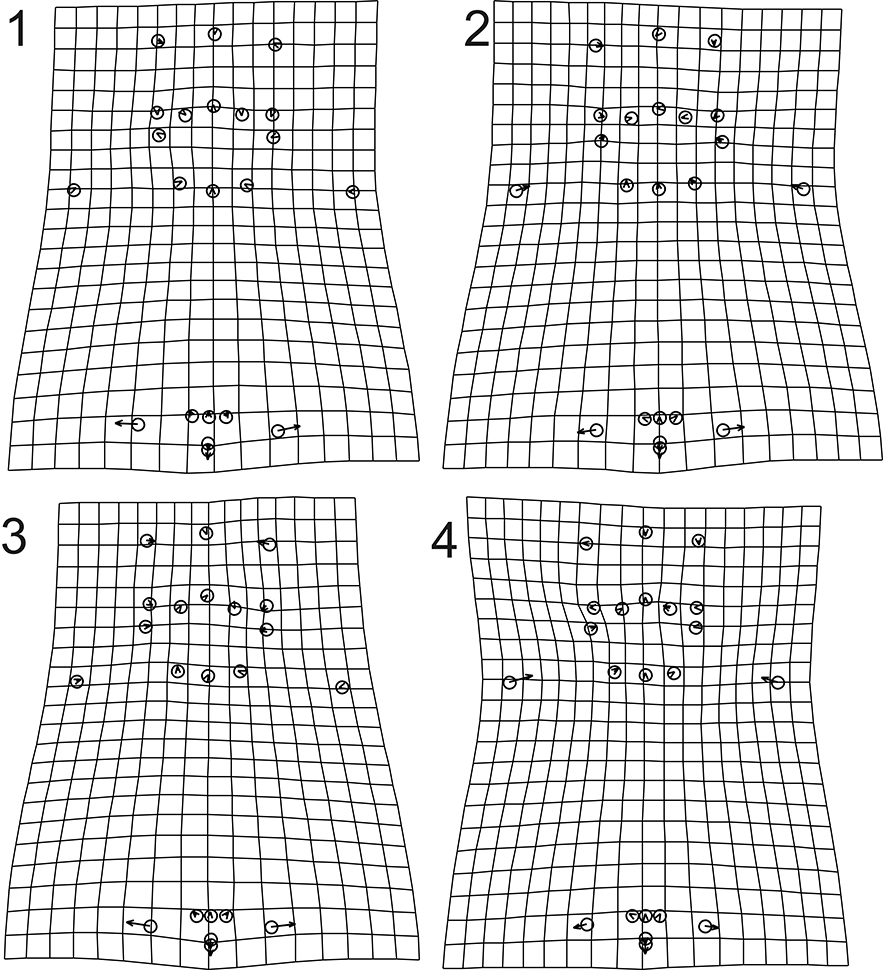
Thin-plate spline deformation grid of relative warps of 22 exoskeletal landmarks for holaspid specimens of morph 21 with its ln CCS value more than 2.2 (N=30): *1*, Shape variation related to relative warp 1 (37.14% of total variance explained); *2*, Shape variation related to relative warp 2 (19.38% of total variance explained); *3*, Shape variation related to relative warp 3 (13.61% of total variance explained, direction inversed); *4*, Shape variation related to relative warp 4 (7.31% of total variance explained).



Thin-plate spline deformation grid of relative warps of 22 exoskeletal landmarks for holaspid specimens of morph 22 with its ln CCS value more than 2.2 (N=12): *1*, Shape variation related to relative warp 1 (48.57% of total variance explained); *2*, Shape variation related to relative warp 2 (16.47% of total variance explained); *3*, Shape variation related to relative warp 3 (13.51% of total variance explained); *4*, Shape variation related to relative warp 4 (5.37% of total variance explained).



Partial Procrustes distance from the reference (mean shape of the smallest three specimens in each holaspid sample) of 22 exoskeletal landmarks for holaspid specimens of morphs 19 to 22 with its ln CCS value more than 2.2. Regression of partial Procrustes distance against logarithm of exoskeletal centroid size is significant for morphs 19 to 21 and it is not significant for the morph 22. *1*, Morph 19 (slope=0.0425, p<0.0001, r=0.6792, N=54); *2*, Morph 20 (slope=0.0533, p<0.0001, r=0.6861, N=42); *3*, Morph 21 (slope=0.0473, p<0.0001, r=0.7085, N=30); *4*, Morph 22 (slope=0.0105, p=0.1786, r=0.2977, N=12).



Thin-plate spline deformation grid of shape changes with growth of 22 exoskeletal landmarks for holaspid specimens of morphs 19 to 22 with its ln CCS value more than 2.2. Partial warp scores are regressed in a multivariate regression against ln exoskeletal centroid size. *1*, Morph 19, 14.79% of total shape variance (based on summed squared residuals expressed in Procrustes units) is explained by the allometry (p<0.000625 from 1600 bootstraps); *2*, Morph 20, 13.50% of total shape variance (based on summed squared residuals expressed in Procrustes units) is explained by the allometry (p<0.000625 from 1600 bootstraps); *3*, Morph 21, 15.46% of total shape variance is explained by the allometry (p<0.000625 from 1600 bootstraps); *4*, Morph 22, 13.50% of total shape variance is explained by the allometry (p=0.17 from 1600 bootstraps).

The vectors of regression coefficients (calculated from regressing partial warp scores in a multivariate regression against ln CCS; mean shape is used as a reference for the partial warp scores) for the 22 exoskeletal landmarks of holaspid specimens of morphs 19 to 22 are compared with isometry. Angle of the normalized vector within the sample and angle to isometry is compared at the 95% confidence level by 1600 bootstraps.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Holaspides  (ln CCS > 2.2) | N | angle to  isometry | within  sample | Significant  difference |
| Morph 19 | 54 | 74.5° | 35.2° | yes |
| Morph 20 | 42 | 83.8° | 45.0° | yes |
| Morph 21 | 30 | 85.9° | 49.4° | yes |
| Morph 22 | 12 | 91.1° | 78.8° | yes |