

# Hierarchical clustering in dinosaur baraminology studies

Matthew Cserhati, Brian Thomas, and Joel Tay

A recent baraminology study places the bird species *Archaeopteryx lithographica* together with the terrestrial dromaeosaurid theropod *Velociraptor mongoliensis*. This despite the two species being created on Day 5 and Day 6 of Creation Week, respectively. This data set was entropy-filtered and re-analyzed, yet the same error persisted. To correct this, the distance matrix from the BDIST results of the filtered data sets were transformed into a proximity matrix and *hierarchical* clustering was applied to it. This way *A. lithographica* clustered away from theropod dinosaurs such as *V. mongoliensis*, together with other bird species such as *Anchiornis huxleyi* to form a bird cluster. Using two separate data sets, one or two tyrannosaurid clusters were predicted. One more theropod cluster, as well as a dromaeosaurid, troodontid, and oviraptor cluster, was predicted as well as the family Ornithomimidae. These clusters can be viewed as putative baramins.

Genesis describes flying creatures (עוף, or “oph” in Hebrew), such as bats and birds, separately from terrestrial animals, such as most reptiles created on Day 5 and Day 6 of Creation Week, respectively. Reptiles are defined as mainly terrestrial animals that lack feathers and hair, and also have scaly skin and lay soft-shelled eggs. Since God created these two distinctly demarcated animal groups on separate days, they necessarily must belong to separate baramins, because they are the results of separate creative acts of God. Pterosaurs are flying reptiles, so they would have been created on Day 5 of Creation Week, together with birds and bats. Conversely, several bird species do not fly, such as rails and ostriches. These birds may either have been created to be incapable of flight, or, more likely, lost flight capability altogether.

A relatively recent baraminology paper by McLain *et al.* using the BDIST algorithm incorrectly lumped the bird species *Archaeopteryx lithographica* into the same cluster with terrestrial animals, such as dinosaurs, for example *Velociraptor mongoliensis*.<sup>1</sup> *Archaeopteryx* has the anatomy of a flying bird, thus it was created on Day 5. Therefore, it cannot be classified as a terrestrial dinosaur.

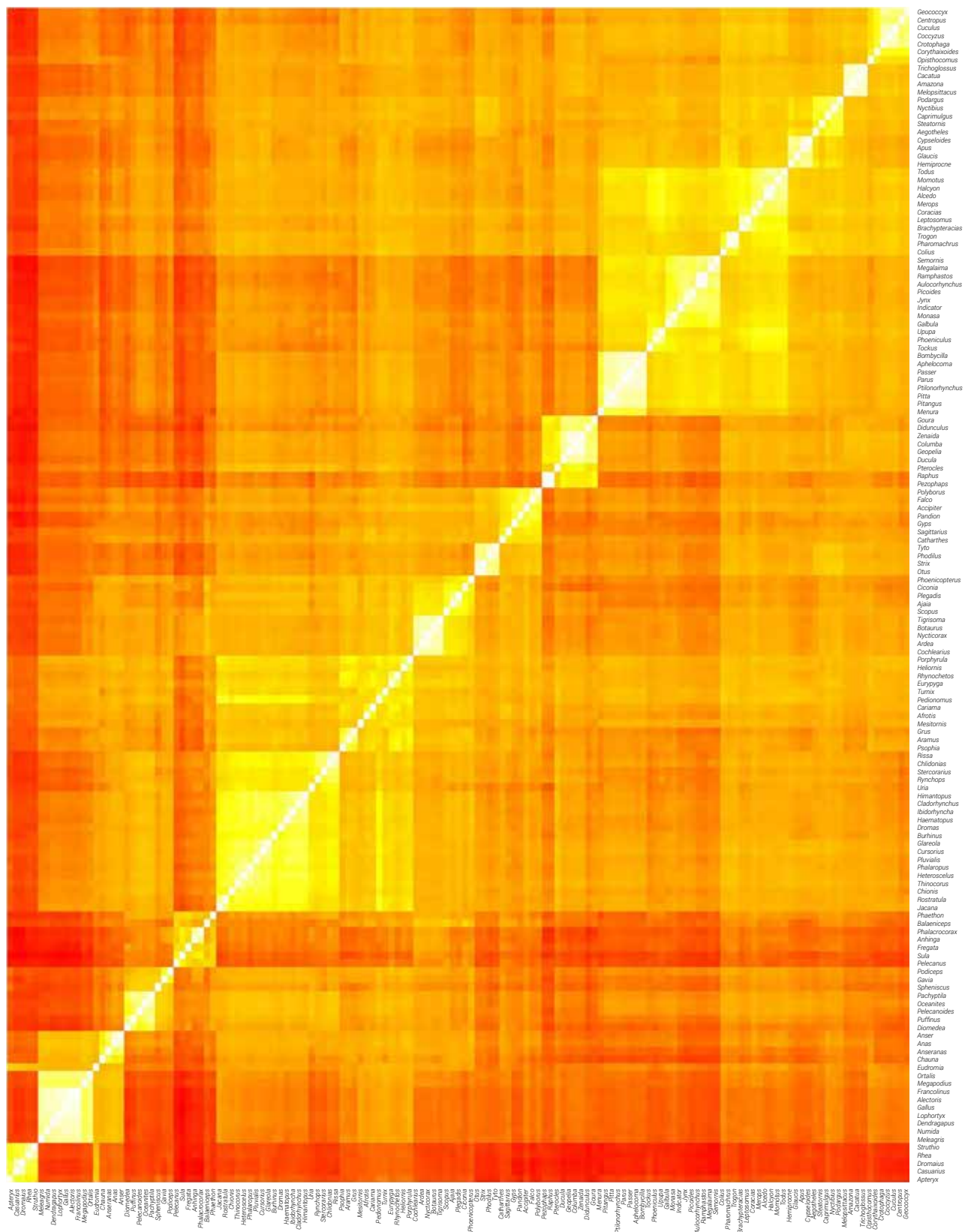
It is highly unlikely that *V. mongoliensis* also would have had feathers, even though some scientists point to bony knobs on its ulna as quill knobs.<sup>2</sup> An alternative and defensible position suggests that these tiny structures on *Velociraptor* are unlikely to be quill knobs, but rather are attachment sites for anchoring tendons or connective tissue. They are more shallow, more irregularly spaced, smaller in proportion to body size, and positioned elsewhere on the ulna than living bird quill knobs. Figure 1A, in Turner reference 2, compares the ulnae of *Velociraptor* and turkey vulture (part C of the figure). Quill knobs are not necessarily indicative of feathers—they

are possible indicators of a *particular type of flight*. Only certain flying birds that experience high stress while flying require quill knobs. For example, the albatross lacks them due to its gliding style of flight.

Pennaceous feathers have never been found buried together with *V. mongoliensis* or *Dakotaraptor*, meaning that it is unlikely that they would have had quill knobs on the bones of their forearms.<sup>3</sup> Some non-avian vertebrates also have knobs which serve as attachment points for structural elements such as fascia. Quill knobs appear in certain birds which undergo high-stress flying. The quill knobs of *Dakotaraptor* are ovoid and irregularly shaped, and are also unevenly spaced.<sup>4</sup> The quill knobs of *Dineobellator* are also broad, rounded, and irregularly spaced, as opposed to the quill knobs of *Cathartes* (turkey vulture), seen in part C of figure 1 of Jasinski, 2020.<sup>5</sup>

What many scientists believe is evidence for protofeathers in dinosaurs falls short under scrutiny. After more careful examination, filamentous structures in the skin of *Sinosauropteryx* proved to be parts of a unified structural element—possibly a crest which runs from the neck down to the tail.<sup>3</sup> Even a ceratopsian dinosaur, *Psittacosaurus*, and the heterodontosaurid *Tianyulong* had such filamentous elements, even though these species are ornithischians, which are not in the alleged lineage of dinosaurs evolving into birds.<sup>6</sup> Such collagen fibres, called pycnofibres, are also present on the membrane-like wing structures of several pterosaurs, such as *Jeholopterus* and *Nemicolopterus*. However, such structures would cause drag during flight for these creatures, and accelerate heat dissipation when wet due to rain. In comparison, bats also have membranous, hairless wings and need to expend more energy heating themselves in the rain.<sup>6</sup>

The discovery that seemingly identical structures have also been found in pterosaurs has caused some to assume



**Figure 1.** Heat map of proximity matrix derived from the BDIST analysis of the filtered Livezey and Zusi data set. Lighter colours represent proximity values close to 1, denoting species within the same baramin. Darker colours represent proximity values close to 0, denoting species from different baramins.

that pterosaurs were also covered in a dense feathery covering, but this does not explain why they are also found in marine creatures. These pterosaur ‘pynofibres’ are very likely just collagen and keratin structures in their partially degraded state. Pynofibres are indistinguishable from what some identified as protofeathers in dinosaurs. Conversely, if these pterosaurs were not feathered, neither were dinosaurs.<sup>7</sup>

One of the problems with feathered dinosaur claims is that partially decayed skin collagen is often mistaken for downy feathers. Lingham-Soliar’s two-volume work, *The Vertebrate Integument*, is widely regarded as the most detailed work on the taphonomy of integumentary structures. He demonstrates that similar fuzzy structures have been found on decaying sharks and dolphins and even fossilized marine reptiles and flying reptiles. “All reported cases of protofeathers in dinosaurs, marine reptiles, and flying reptiles alike, show a striking similarity to the structure and levels of organization of dermal collagen. The proposal that these fibres are protofeathers is dismissed.” Thus, while creationists may hypothesize that *Velociraptor* had volant ancestors, and thus created on Day 5, this has to be much better established.

#### Overview of the BDIST method and hierarchical clustering

The BDIST method is a morphology-based baraminology algorithm, which quantitatively compares the similarities and differences between the traits of a set of input species.<sup>8</sup> The algorithm calculates the baraminic distance between any two species in the following way:  $d_{ij} = m_{ij}/n_{ij}$ , where  $n_{ij}$  is the number of comparable characteristics between species  $i$  and  $j$ ,  $m_{ij}$  is the number of mismatching characteristics, and  $d_{ij}$  is the distance between the two species.

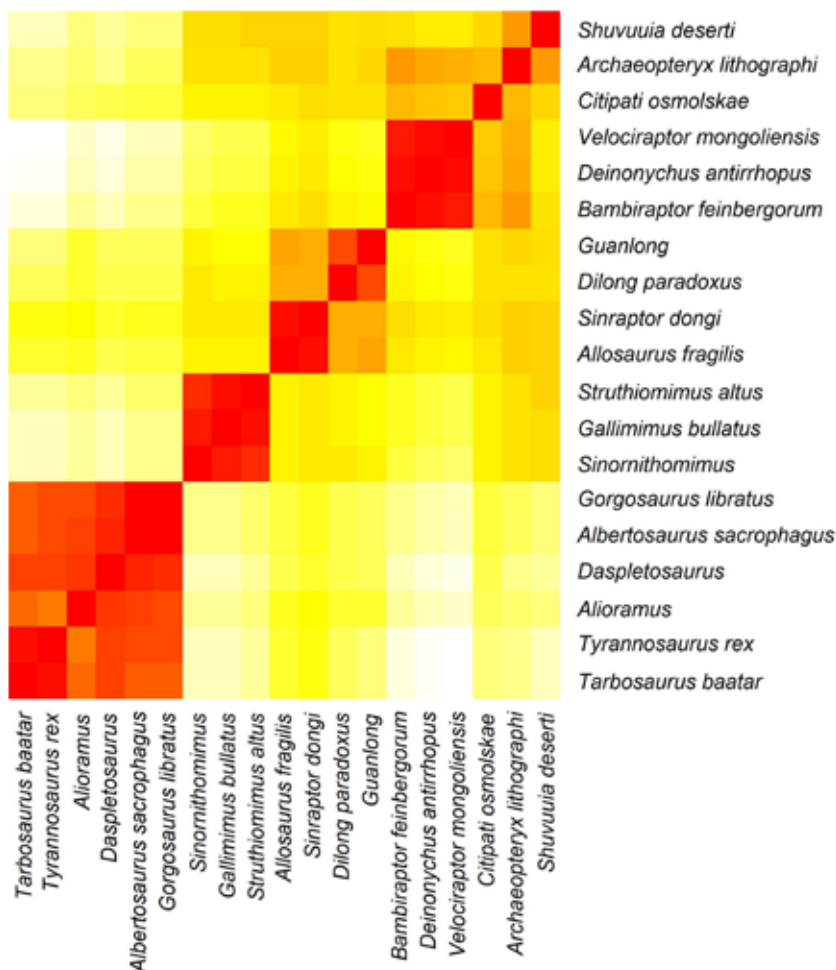
As a hypothetical example, if we compare 100 characteristics for a horse and a zebra, two animals from the same baramin, we’ll find that they differ in only five characteristics, thus  $d_{\text{horse,zebra}} = 5/100 = 0.05$ . If we compare a horse with a lion, we’d find bigger differences, let’s say 40. Thus the distance  $d_{\text{horse,lion}} = 40/100 = 0.4$ . These distance values are then put into a matrix, which can then be analyzed by clustering algorithms, such as the k-means algorithm, which belongs to the family of hierarchical clustering algorithms.<sup>9</sup>

**Table 1.** List of putative clustering, which is the result of hierarchical clustering of the proximity matrix, derived from the entropy-filtered data set of Livezey and Zusi

cluster	species	min	mean	max	stdev	p-value
5 (Coraciiformes)	5	0.897	0.921	0.967	0.023	2.09E-11
6 (Galliformes)	7	0.922	0.955	0.981	0.017	2.11E-33
7 (Psittaculidae)	4	0.957	0.962	0.967	0.004	1.60E-28
11 (Passeriformes)	7	0.936	0.955	0.98	0.012	1.17E-48
13 (Apodiformes)	4	0.904	0.929	0.967	0.022	7.23E-07
15 (Ardeidae)	5	0.933	0.947	0.967	0.011	1.12E-16
16 (Piciformes)	7	0.878	0.91	0.984	0.024	5.54E-26
20 (Caprimulgiformes)	3	0.884	0.891	0.904	0.011	3.33E-04
22 (Palaeognathae cluster)	4	0.869	0.898	0.953	0.035	3.02E-06
24 (Cuculiformes)	5	0.9	0.927	0.959	0.022	7.42E-12
26 (Ciconiiformes)	7	0.863	0.896	0.95	0.022	6.20E-24
27 (Charadriiformes I)	4	0.861	0.9	0.948	0.028	6.11E-06
29 (Charadriiformes II)	5	0.867	0.902	0.942	0.025	1.51E-10
31 (Columbidae)	4	0.938	0.958	0.981	0.021	1.29E-07
48 (Procellariiformes)	3	0.918	0.929	0.946	0.015	6.42E-04
50 (Strigiformes)	4	0.917	0.936	0.976	0.023	7.12E-07

Hierarchical clustering involves partitioning a set of  $N$  entities, such as species, into  $K$  partitions (clusters), where  $1 \leq K \leq N$ . Hierarchical clustering methods fall into one of two main categories: bottom-up (or agglomerative) or top-down methods. The result of bottom-up algorithms is a nested hierarchy of clustered entities. Top-down approaches involve optimally splitting all the entities into  $K$  clusters.

Each hierarchical clustering method uses a dissimilarity and a linkage function to cluster the data. Dissimilarity functions are used to calculate the distance between the entities. These functions include squared Euclidean distance, Manhattan distance, or  $1 - |PCC|$ , where PCC stands for the Pearson Correlation Coefficient. In this way, the distance measurement (or a derivation thereof) used by the BDIST algorithm can be used in further analyses using a hierarchical clustering algorithm. The linkage function dictates how pairwise entities (species, clusters or both) are connected to one another to form clusters.<sup>10</sup> Linkage functions include Ward’s algorithm, single, complete, and average linkage.



**Figure 2.** Heat map of the proximity matrix from the BDIST results of the filtered Brusatte data set. Lighter colours represent proximity values close to 1, denoting species within the same baramin. Darker colours represent proximity values close to 0, denoting species from different baramins.

The relationships and distances between entities can be depicted on a dendrogram. The number of clusters can be determined by cutting branches at a certain height in the dendrogram. In baraminology, these clusters ideally correspond to the created kinds. Other algorithms, such as *k*-Means or PAM (Partition Around Medoids) clustering, can be used to cluster species hierarchically. The *k*-Means algorithm is a vector quantization method that is frequently used for cluster analysis by grouping objects since it quickly finds the centres of the clusters. The algorithm prefers groups with low variance and similar size. The PAM algorithm discovers objects called medoids, that are centrally located in clusters. The algorithm tries to minimize the distance between the objects and their closest medoid.

#### Previous studies and goal

Another baraminology methodology paper, which applied data filtering based on entropy measurement, attempted to

correct this over-clustering tendency in some of the data used by McLain *et al.*<sup>11</sup> It helped resolve baramins into more anatomically sensible (less inclusive) groupings, but it did not separate *A. lithographica* from terrestrial dinosaurs in two data sets from Lee *et al.*<sup>29</sup>, Brusatte *et al.*<sup>12</sup>, van der Reest *et al.*<sup>13</sup>, and Lamanna *et al.*<sup>14</sup>

Over-clustering and misclustering may result from two different causes. The first cause may be due to the use of pre-filtered data sets obtained from evolutionary studies. When creating data sets, individual characters may be chosen, and values may be assigned to them based on one's presuppositions. Evolutionary data sets might emphasize certain traits which support one or another evolutionary trajectory. For example, bird feathers may be conflated with dinosaur dermal collagen.<sup>15</sup> Therefore, if biased data goes in, we get biased results from the algorithm, regardless how well it was programmed.

It is extremely difficult—even impossible—to solve this problem, because it would involve manually evaluating and filtering possibly several hundred characters over dozens of species. However, the data set may be improved using an entropy-filtering algorithm. In addition, we can try to

alleviate this problem by using a different clustering algorithm than the one used by BDIST.

The goal of this study is to apply another clustering algorithm in an attempt to get more accurate results from analyzing the filtered data sets of the previous baraminology study. The algorithm is outlined in the Materials and Methods section, followed by the results.

## Materials and Methods

### Data

Distance matrixes were built following the methods of Brusatte, van der Reest and Lamanna studies<sup>12–14</sup> and the results of this study are available on github ([github.com/csmaty/EntropyFilter2](https://github.com/csmaty/EntropyFilter2)). The Livezey and Zusi data set is available at [morphobank.org/index.php/Projects/DownloadProjectPage/project\\_id/510](http://morphobank.org/index.php/Projects/DownloadProjectPage/project_id/510). It was downloaded



and recoded from letter characters to numerals (i.e. a=1, b=2, etc.) for use in the BDIST algorithm.

All supplementary data files are available online at [github.com/csmaty/dino\\_reanalysis](https://github.com/csmaty/dino_reanalysis). The Cytoscape file for creating figure 5 (dino\_baramins.cys) is also available at this web address.

## Methods

The proximity matrix was derived using the following simple equation:

$$p_{i,j} = 1 - d_{i,j}$$

where  $d_{i,j}$  stands for the baraminic distance between two species  $i$  and  $j$  as calculated by the BDIST algorithm, and where  $p_{i,j}$  is the corresponding proximity value for the same two species. This transformed matrix was visualized as a heatmap with the ‘heatmap’ command in R, using the ‘average’ clustering method. The ‘cutree’ command was used to determine clusters for the data sets based on the estimated number of clusters.

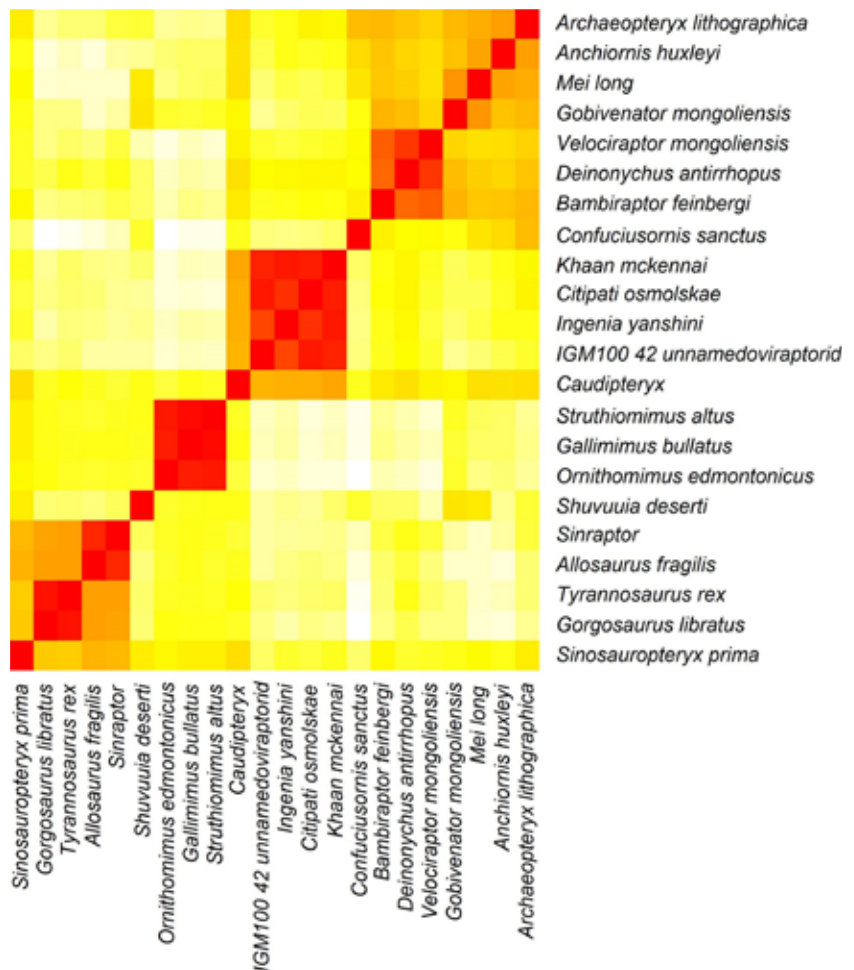
A Hopkins clustering measurement value above 0.5 corresponds to a data set which is acceptable for cluster analysis. A value above 0.75 denotes a good quality data set for such analysis.<sup>16</sup> Version 3.6.0. of R was used.<sup>17</sup> Cytoscape 3.7.2. was used to generate figure 6.

The BDIST algorithm was used for analysis of the Livezey and Zusi data set, available at [coresci.org/bdist.html](http://coresci.org/bdist.html).<sup>18,19</sup> A relevance cut-off of 0.95 was used in the analysis. The EntropyFilter.R R script was used, and is available at [github.com/csmaty/EntropyFilter2](https://github.com/csmaty/EntropyFilter2).

## Results

### Analysis of the Livezey and Zusi bird data set

To ensure accurate results with the data we are analyzing, we decided to run the algorithm on a data set, which can serve as a sort of gold standard. This data set contains morphological measurements of birds whose baraminic status is well known. Birds have been studied extensively in baraminology studies, and baraminologists can say with a high level of confidence which baramin(s) these birds belong to, for example that of Galliformes, or the ‘landfowl’ kind.<sup>20</sup>



**Figure 3.** Heat map of the proximity matrix from the BDIST results of the filtered van der Reest data set. Lighter colours represent proximity values close to 1, denoting species within the same baramin. Darker colours represent proximity values close to 0, denoting species from different baramins.

The validity of this kind has been supported by biblical evidence, hybridization data, and ortholog protein content.<sup>21</sup>

A morphological data set containing data for 2,954 characters for a wide variety of 188 bird species assembled by Livezey and Zusi<sup>22</sup> was analyzed. The BDIST software was used to analyze this data set using a relevance cut-off value of 0.95. The result can be seen in supplementary figure 1. The stress graph (supplementary figure 2) also shows minimal stress at only two dimensions. This result does an excellent job at illustrating the need to improve the BDIST algorithm with a better clustering method.

Thus, the data set was entropy filtered, using the Entropy Filter R script, described in the Materials and Methods section. A maximum ratio of undetermined values was set to 25% for both rows and columns. The minimum entropy value was set to 0.25. During this process, the percent of undetermined values dropped from 22.1% to 1.6%. The mean character entropy rose from 0.32 to 0.57. Of the 188 initial

**Table 2.** Difference in clustering using the BDIST method and hierarchical clustering for the entropy-filtered Brusatte data set

Species	BDIST cluster no.	Hierarchical cluster no.
<i>Albertosaurus sarcophagus</i>	1	1
<i>Alioramus</i>	1	1
<i>Daspletosaurus</i>	1	1
<i>Gorgosaurus libratus</i>	1	1
<i>Tarbosaurus baatar</i>	1	1
<i>Tyrannosaurus rex</i>	1	1
<i>Allosaurus fragilis</i>	2	2
<i>Dilong paradoxus</i>	2	2
<i>Guanlong</i>	2	2
<i>Sinraptor dongi</i>	2	2
<i>Archaeopteryx lithographi</i>	3	3
<i>Bambiraptor feinbergorum</i>	3	4
<i>Deinonychus antirrhopus</i>	3	4
<i>Velociraptor mongoliensis</i>	3	4
<i>Citipati osmolskae</i>	3	5
<i>Gallimimus bullatus</i>	4	6
<i>Sinornithomimus</i>	4	6
<i>Struthiomimus altus</i>	4	6
<i>Shuvuuia deserti</i>	3	7

**Table 3.** Statistics for the hierarchical clustering for the entropy-filtered Brusatte data set

baramin	species	min	mean	max	st. dev.	p-value
1 (tyrannosaurid1)	6	0.824	0.904	1	0.046	3.33E-21
2 (tyrannosaurid2)	4	0.743	0.811	0.973	0.098	1.00E-03
4 (dromaeosaurids)	3	0.959	0.973	0.986	0.014	1.17E-23
6 (Ornithomimidae)	3	0.932	0.955	0.973	0.021	1.98E-09

species, 147 (78%) remained, with 1,136 of the 2,954 characters (38.4%) also remaining, corresponding to a 3.3-fold data reduction. The filtered data set was then inputted into the BDIST algorithm. The results can be seen in supplementary figure 3: entropy filtering resulted in two larger and three smaller clusters. However, as can be seen in the stress graph in supplementary figure 4, there is minimal stress at 49 dimensions. Clearly, the clustering method must be improved.

Therefore, the distance matrix resulting from the BDIST analysis on the filtered data set was transformed into a proximity matrix according to the procedure outlined in Materials and Methods. Hierarchical clustering was applied to this matrix, which can be seen in the baraminic heatmap in figure 1.

In total, 70 putative clusters were discovered, among them many small groups, as can be seen in the baraminic heatmap. The species and cluster number for the 147 species can be found in supplementary file 1 online, seen in table 1. Overall, 16 statistically significant putative clusters were discovered using hierarchical clustering. These clusters correspond to several orders, superfamilies or families of birds. For example, the superfamily Galliformes was found in the analysis of the proximity data set.<sup>21</sup> Several groups within the Caprimulgiformes apobaramin also emerged. These include the families *Caprimulgus*+*Nyctibius*+*Podargus*. The genus *Steatornis* and *Aegotheles* formed their own cluster. In contrast with previous baraminology studies, *Podargus* didn't belong to the same cluster as *Steatornis*.<sup>23</sup> This may be due to the fact that a large number of species were taken from a large number of groups in this study.

The Brusatte data set

To assess the cluster tendency of our data sets, the Hopkins clustering statistic was used. For the proximity data set of the Brusatte results we calculated a

value of 0.871. This means that this data set can very well be used for cluster analysis. In the following, we will refer to groups of species found by our algorithm as clusters. Hypothetically, these clusters should correspond to baramins, but can only be labelled as such tentatively, until supported by further evidence.

Following this method, seven clusters were found. The proximity matrix is displayed in figure 2. Here we can see four larger clusters, and three singleton species, *Archaeopteryx lithographica*, *Citipati osmolskae*, and *Shuvuuia deserti*. Whereas in the previous analysis these three species were lumped together with the dromaeosaurids, they now stand separately.

Next, we compared how this clustering relates to the old analyses (table 2). As we can see, 14 of the 19 species (73.7%) were assigned to the same group. In table 3 the four largest clusters are listed with at least three members. These are tyrannosaurids 1 and 2, dromaeosaurids, and the family Ornithomimidae. The four largest clusters (ones with more than two members) have statistically significant p-values. The proximity matrix and the clusters and stats are available online in supplementary file 2.

The van der Reest data set

Next, we wanted to see how well we could reclassify species from the van der Reest data set. The Hopkins clustering statistic for the proximity matrix used for this data set was 0.746, which is fair quality for clustering, but somewhat worse than the Brusatte data set. Using hierarchical clustering revealed 10 clusters, which are presented as a proximity matrix (heatmap in figure 3). In the heatmap we can see four larger groups—species belonging to clusters 1, 3, 5, and 7 listed in supplementary data file 2. These four larger clusters also have statistically significant p-values.

**Table 4.** Difference in clustering using the BDIST method and hierarchical clustering for the entropy-filtered van der Reest data set

Species	BDIST cluster no.	Hierarchical cluster no.
<i>Allosaurus fragilis</i>	1	1
<i>Gorgosaurus libratus</i>	1	1
<i>Sinraptor</i>	1	1
<i>Tyrannosaurus rex</i>	1	1
<i>Anchiornis huxleyi</i>	2	2
<i>Archaeopteryx lithographica</i>	2	2
<i>Bambiraptor feinbergi</i>	2	3
<i>Deinonychus antirrhopus</i>	2	3
<i>Velociraptor mongoliensis</i>	2	3
<i>Caudipteryx</i>	3	4
<i>Citipati osmolskae</i>	3	5
<i>IGM100 42 unnamed dromaeosaurid</i>	3	5
<i>Ingenia yanshini</i>	3	5
<i>Khaan mckennai</i>	3	5
<i>Confuciusornis sanctus</i>	2	6
<i>Gallimimus bullatus</i>	4	7
<i>Ornithomimus edmontonicus</i>	4	7
<i>Struthiomimus altus</i>	4	7
<i>Gobivenator mongoliensis</i>	2	8
<i>Mei long</i>	2	8
<i>Shuvuuia deserti</i>	5	9
<i>Sinosauropteryx prima</i>	1	10

**Table 5.** Statistics for the hierarchical clustering for the entropy-filtered van der Reest data set

baramin	species	min	mean	max	st. dev.	p-value
1 (tyrannosaurid)	4	0.816	0.977	0.867	0.076	1.10E-4
3 (dromaeosaurid)	3	0.883	0.937	0.904	0.029	1.34E-4
5 (oviraptor)	4	0.922	0.968	0.954	0.019	8.23E-21
7 (ornithomimid)	3	0.964	0.987	0.973	0.012	1.12E-11

**Table 6.** Difference in clustering using the BDIST method and hierarchical clustering for the entropy-filtered Lamanna data set

Species	BDIST cluster no.	Hierarchical cluster no.
<i>Velociraptor mongoliensis</i>	1	1
<i>Herrerasaurus ischigualastensis</i>	1	2
<i>Citipati osmolskai</i>	2	3
<i>Conchoraptor gracilis</i>	2	3
<i>Ingenia yanshini</i>	2	3
<i>Khaan mckennai</i>	2	3
<i>Nemegtomaia barsboldi</i>	2	3
<i>Rinchenia mongoliensis</i>	2	3
<i>Yulong mini</i>	2	3
<i>Zamyn Khondt oviraptorid</i>	2	3
<i>Anzu wyliei</i>	3	4
<i>Archaeopteryx lithographica</i>	1	5
<i>Avimimus portentosus</i>	4	6
<i>Caudipteryx zoui</i>	5	6
<i>Incisivosaurus gauthieri</i>	5	6

**Table 7.** Statistics for the hierarchical clustering for the entropy-filtered Lamanna data set

baramin	species	min	mean	max	st. dev.	p-value
5 (oviraptor)	8	0.768	0.863	0.977	0.054	1.21E-29

Here also, we compared our new clustering patterns with those previously observed (table 4). Cluster statistics are presented in table 5. Ten out of 22 species (45.5%) classifications were the same in both the old and new analysis. The four largest clusters have statistically significant p-values at the 5% level. We found one of the same putative clusters as when we re-analyzed the Brusatte data set, a dromaeosaurid cluster and a new cluster of oviraptorids. However, in the re-analysis of this data set, we found that tyrannosaurids

form one large cluster instead of two clusters in the re-analysis of the Brusatte data set. This might be due to the selection of characteristics in the individual data sets.

The proximity matrix and the clusters and stats are available online in supplementary file 3 for this analysis.

The Lamanna data set

Lastly, we also re-analyzed the Lamanna data set. For this data set, the Hopkins clustering statistic was 0.739, which is fair clustering quality. Six clusters were determined. The proximity matrix is depicted in figure 4, showing one large cluster with eight species, together with either pairs of species or single species. The large cluster corresponds to an oviraptorid cluster, which was also found in the analysis of the van der Reest data set.

The old and new clustering can be seen in table 6. Twelve out of 15 classifications (80%) were shown to be the same in both the old and new analyses. Cluster statistics can be seen in table 7, showing that only one cluster with more than two members was determined, and has a statistically significant p-value. The proximity matrix and the clusters and stats are available online in supplementary file 4.

## Discussion

In our current study, we re-analyzed three dinosaur data sets from a previous analysis which employed entropy filtering, but which still gave obviously incorrect results (i.e. lumping *A. lithographica* together with dromaeosaurs). In order to correct these results, the BDIST pipeline was modified. Thus, a large-scale morphology data set for 188 bird species was entropy-filtered, used as input for the BDIST algorithm, and the resulting distance matrix was transformed into a proximity matrix, which was then subjected to hierarchical clustering. This method corrected the over-clustering, which was present in the BDIST results of both the initial, raw results as well as even the entropy-filtered BDIST results.

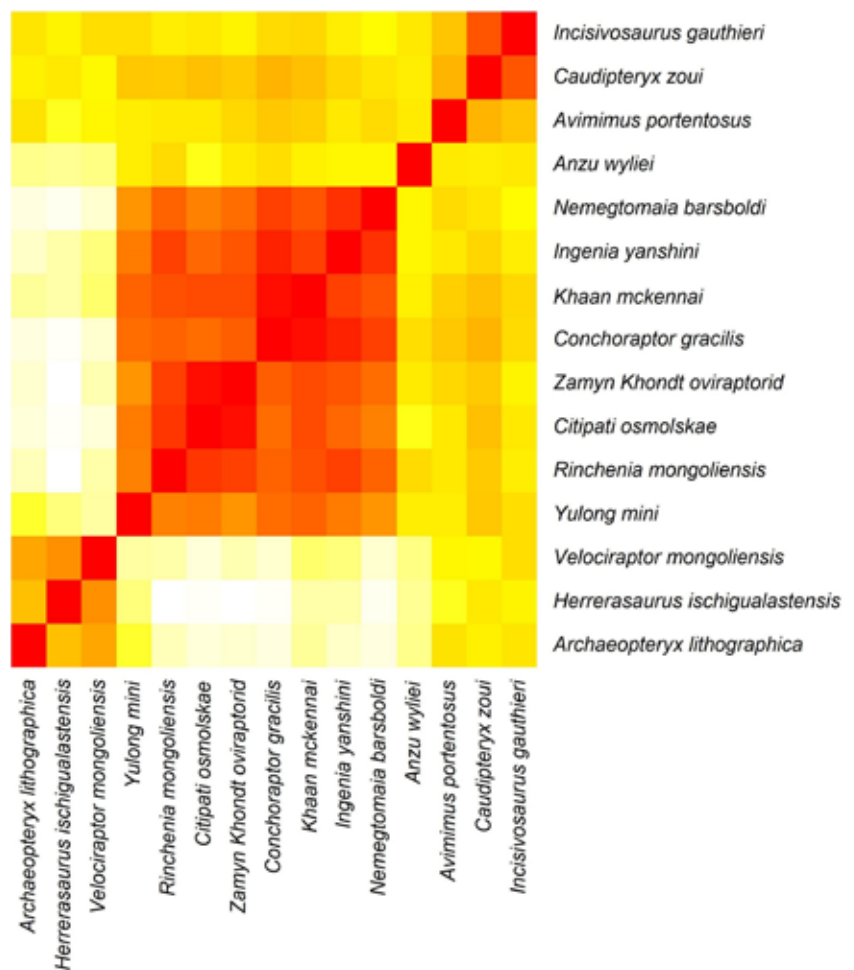


The present analysis succeeded in properly classifying the incorrect classifications coming from our previous analysis of the three dinosaur data sets. We observed a total of seven novel putative baraminic relationships based on the clustering we found in our results (depicted in figure 5). These are the tyrannosaurids (all species put together), dromaeosaurids, troodontid 1, oviraptor 1, theropod 2, ornithomimid, and *Archaeopteryx*, which includes *A. lithographica* and *A. huxleyi*.

In all three studies *A. lithographica* does not cluster with *Bambiraptor feinbergi*, *Deinonychus antirrhopus*, *Herrerasaurus ischigualastensis*, and *Velociraptor mongoliensis*. These dromaeosaurids consistently form a cluster in all three modified data sets. None of the latter four fossils have feather impressions or keratin signatures associated with them. Inferences based on *A. lithographica* anatomy strongly suggest that at the very least, it is a gliding bird, if not fully capable of flight. In particular, the classic defining bird trait is feathers, and *Archaeopteryx*'s asymmetrical feather impressions<sup>24</sup> and feather protein remnant signatures<sup>25</sup> are consistent with flight-capable remiges for the extinct creature.

However, the lack of a keel in *Archaeopteryx* implies a much smaller pectoralis major muscle than modern powered-flight birds such as the passerines. Why would a bird have feathers to fly but no muscles to power flight? From a creationist perspective it shows design features for gliding, not for flying. The modern roadrunner (*Geococcyx*) may approximate the form and function of *Archaeopteryx*. However, *Geococcyx* has no wing claws. Whereas *Geococcyx* uses terrain to elevate itself, perhaps *Archaeopteryx* used its wing claws and other features in more arboreal settings. Like roadrunners, once elevated, the bird could glide downward. This again supports the idea that certain characters should be given more weight than others. For example, with birds, the presence of feathers should have much more weight than the size of bones, which can be up- or down-scaled genetically.

In the van der Reest study, *A. lithographica* clustered together with *Anchiornis huxleyi*. Both of these species have a similar wing structure, with slender feather shafts, meaning they

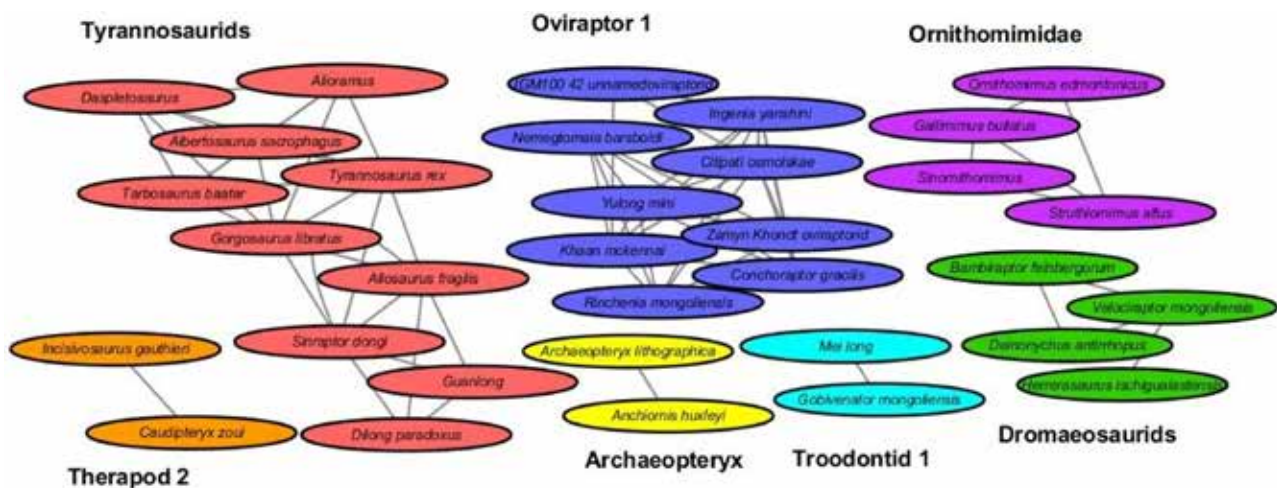


**Figure 4.** Heat map of the proximity matrix from the BDIST results of the filtered Lamanna data set. Lighter colours represent proximity values close to 1, denoting species within the same baramin. Darker colours represent proximity values close to 0, denoting species from different baramins.

had similar flight mechanisms.<sup>26</sup> In contrast, *Confuciusornis sanctus* has elongated primary feathers, with short covert feathers, resembling modern birds.<sup>27</sup>

Nine species all make up a larger cluster, throughout the three studies: *Citipati osmolskae*, *Conchoraptor gracilis*, IGM100-42 unnamed oviraptorid, *Ingenia yanshini*, *Khaan mckennai*, *Nemegtomaia barsboldi*, *Rinchenia mongoliensis*, *Yulong mini*, and *Zamyn Khondt*. All of these are oviraptors, which are theropod dinosaurs with shortened, toothless skulls.<sup>28–30</sup>

*Gallimimus bullatus*, *Ornithomimus edmontonicus*, *Sinornithomimus*, and *Struthiomimus altus* all cluster together. They are all members of the theropod family of dinosaurs called Ornithomimidae. *Mei long* and *Gobivenator mongoliensis* are both troodontids. *Anzu wyliei* and *Avimimus portentosus* belong to separate oviraptor baramins. *Caudipteryx zoui* and *Incisivosaurus gauthieri* are members of the same theropod cluster. However, feathered dinosaur



**Figure 5.** Baraminic relationships between seven baramins representing 34 species from the BDIST results of the combined Brusatte, van der Reest, and Lamanna data sets.

doubt Feduccia and colleagues treated *Caudipteryx* as a flightless bird, perhaps analogous to flightless cormorants of the Galápagos islands, on the basis of precise arrangement of flight feathers on its hand.<sup>31</sup> However feathered it was, its wings were too small for flight, but its robust legs showed possible adaptation to terrestrial habitats.<sup>32</sup> Jones *et al.* noted its anterior centre of mass resembled those of cursorial birds.<sup>33</sup> By grouping *Caudipteryx* among theropods, our modified BDIST method has not captured the separation of *Caudipteryx* as a bird and not a reptile. This is likely due to the similarities in gross morphology, if not in important details like feathers versus not feathers, between theropod and *Caudipteryx* anatomy.

*Shuvuuia deserti* and *Sinosauropteryx prima* are also theropods but belong to separate clusters. This result compares well with the assignment of *Sinosauropteryx* to the compsognathid dinosaurs particularly on the basis of papulose skin scale—not feather—preservation on its tail.<sup>34</sup> Chiappe also grouped *Compsognathus* with *Velociraptor* among theropods based on its cursorial anatomy and curved, serrated teeth that birds do not have.<sup>35</sup>

Lastly, there is a discrepancy between the way tyrannosaurids are classified according to the Brusatte and the van der Reest data sets. Whereas the re-analysis of the van der Reest data puts *Allosaurus fragilis* and *Sinraptor* into the same baramin as *Tyrannosaurus rex*, the Brusatte re-analysis puts them into a separate cluster, together with *Guanlong* and *Dilong paradoxus*. It could be that the two data sets were put together using different characters, or there may be more than one tyrannosaurid baramin. This is not uncommon, since previous baraminology studies have predicted multiple turtle and squid baramins.<sup>36,37</sup>

In contrast with the tyrannosaurids, this study successfully predicted the dromaeosaurid, oviraptor, and Ornithomimidae clusters in the re-analysis of the three data sets.

This means that the modified algorithm finds more consistent results than the BDIST method alone.

## Conclusion

Baraminology studies may sometimes give results that might intuitively seem incorrect. For example, in the previous analysis of three dinosaur data sets, *Archaeopteryx lithographica* was classified as a member of a terrestrial dinosaur baramin, when it has anatomy for flight. From the Bible, we know that birds were created on separate days than terrestrial dinosaurs, therefore they must belong to separate baramins. Feathers also provide a clear anatomical separation of Class Aves from Class Reptilia. Applying the extended entropy filter + the hierarchical clustering algorithm to the BDIST method corrected this. Furthermore, this method may be used as a possible tool in order to predict putative baramins more accurately.

## References

- McLain, M.A., Petrone, M., and Speights, M., Feathered dinosaurs reconsidered: new insights from baraminology and ethnotaxonomy; in: Whitmore, J.H. (Ed.), *Proceedings of the Eighth International Conference on Creationism*, Creation Science Fellowship, Pittsburgh, PA, p. 508, 2018.
- Turner, A.H., Makovicky, P.J., and Norell, M.A., Feather quill knobs in the dinosaur *Velociraptor*, *Science* **317**(5845):1721, 2007.
- Sarfati, J., Feathered dinos: no feathers after all! *J. Creation* **26**(3):8–10, 2012.
- DePalma, R.A., Burnham, D.A., Martin, L.D., Larson, P.L., and Bakker, R.T., The first giant raptor (Theropoda: Dromaeosauridae) from the Hell Creek Formation, *Paleontological Contributions* **2015**(14):1–16, 2015.
- Jasinski, S.E., Sullivan, R.M., and Dodson, P., New Dromaeosaurid dinosaur (Theropoda, Dromaeosauridae) from New Mexico and biodiversity of Dromaeosaurids at the end of the Cretaceous, *Scientific reports* **10**(1): 2020.
- Feduccia, A., *Riddle of the Feathered Dragons*, Yale University Press, New Haven, CT, p. 258–59, 2012.
- Tay, J., Feathered pterosaurs: ruffling the feathers of dinosaur evolution, *J. Creation* **33**(2):93–98, 2019.

8. Robinson, D.A. and Cavanaugh, D.P., A Quantitative approach to baraminology with examples from the catarrhine primates, *CRSQ* 34(4):196–208, 1998.
9. Wood, T.C. and Murray, M.J., *Understanding the Pattern of Life: Origins and organization of the species*, Broadman & Holman, Nashville, TN, 2003.
10. Kimes, P.K., Liu, Y., Hayes, D.N., and Marron, J.S., Statistical significance for hierarchical clustering, *Biometrics* 73(3):811–821, 2017.
11. Cserhati, M., Baraminology data filtering method based on entropy measurement and its application in dinosaur and cephalopod data sets, *J. Creation* 33(3):55–65, 2019.
12. Brusatte, S.L., Lloyd, G.T., Wang, S.C., and Norell, M.A., Gradual assembly of avian body plan culminated in rapid rates of evolution across the dinosaur-bird transition, *Current Biology* 24:2386–2392, 2014.
13. van der Reest, A.J. and Currie, P.J., Troodontids (Theropoda) from the Dinosaur Park Formation, Alberta, with a description of a unique new taxon: implications for deinonychosaur diversity in North America, *Canadian J. Earth Sciences* 54(9):919–935, 2017.
14. Lamanna, M.C., Sues, H.D., Schachner, E.R., and Lyson, T.R., A new large-bodied oviraptorosaurian theropod dinosaur from the latest Cretaceous of western North America, *PLoS ONE* 9(3):e92022, 2014.
15. Cserhati, M., and Tay, J., Comparison of morphology-based and genomics-based baraminology methods, *J. Creation* 33(3):10–15, 2019.
16. Lawson, R.G. and Jurs, P.C., New index for clustering tendency and its application to chemical problems, *J. Chemical Information and Computer Sciences* 30(1):36–41, 1990.
17. [cran.r-project.org/bin/windows/base/old/3.6.0](https://cran.r-project.org/bin/windows/base/old/3.6.0)
18. Wood, T.C., BDIST software, v. 1.0. Center for Origins Research and Education, Bryan College, Distributed by the author, 2001.
19. Wood, T.C., Visualizing baraminic distances using classical multidimensional scaling, *Origins (GRI)* 57:9–29, 2005.
20. Lightner, J.K., An initial estimate of avian Ark kinds, *ARJ* 6 (2013):409–466, 2013.
21. Ahlquist, J. and Lightner, J.K., Strategies for more clearly delineating, characterizing, and inferring the natural history of baramins I: establishing baraminic status, with application to the order galliformes (class: Aves), *CRSQ* 56:97–104, 2019.
22. Livezey, B.C. and Zusi, R.L., Higher-order phylogeny of modern birds (Theropoda, Aves: Neornithes) based on comparative anatomy: I—Methods and characters, *Bulletin of the Carnegie Museum of Natural History* 37:1–544, 2006.
23. Cserhati, M. and Alquist, J., Baraminology suggests cryptic relationships among Caprimulgiformes, *J. Creation* 33(3):58–68, 2019.
24. Wellnhofer, P., *Archaeopteryx: The Icon of Evolution*, Verlag Dr. Friedrich Pfeil, München, 2009.
25. Bergmann, U., Morton, R.W., Manning, P.L., Sellers, W.I., Farrar, S., and Huntley, K.G., Archaeopteryx feathers and bone chemistry fully revealed via synchrotron imaging, *Proc. Natl Acad. Sci. USA* 107(20):9060–9065, 2010.
26. Longrich, N.R., Vinther, J., Meng, Q., Li, Q., and Russell, A.P., Primitive wing feather arrangement in *Archaeopteryx lithographica* and *Anchiornis huxleyi*, *Curr Biol.* 22(23):2262–2267, 2012.
27. Nudds, R.L. and Dyke, G.J., Narrow primary feather rachises in *Confuciusornis* and *Archaeopteryx* suggest poor flight ability, *Science* 328(5980):887–889, 2010.
28. Yun, C.G., Heyuanninae clade nov., a replacement name for the oviraptorid subfamily “Ingeniinae” Barsbold, 1981, *Zootaxa* 4671(2):zootaxa.4671.2.11, 2019.
29. Lee, S., Lee, Y.N., Chinsamy, A., Lü, J., Barsbold, R., and Tsogtbaatar, K., A new baby oviraptorid dinosaur (Dinosauria: Theropoda) from the Upper Cretaceous Nemegt Formation of Mongolia, *PLoS One* 14(2):e0210867, 2019.
30. Lü, J., Chen, R., Brusatte, S.L., Zhu, Y., and Shen, C., A Late Cretaceous diversification of Asian oviraptorid dinosaurs: evidence from a new species preserved in an unusual posture, *Sci Rep.* 6:35780, 2016.
31. Feduccia, A., Theagarten, L.S., and Hinchliffe, J.R., Do feathered dinosaurs exist? Testing the hypothesis on neontological and paleontological evidence, *J. Morphology* 266:125–166, 2005.
32. Talori, Y.S. and Zhao, J.-S., Aerodynamics of soft flapping wings of *Caudipteryx*; in: Yu, H., Liu, J., Liu, L., Ju, Z., Liu, Y., and Zhou, D. (Eds.), *Intelligent Robotics and Applications, ICIRA 2019, Lecture Notes in Computer Science*, vol. 11742, Springer, Cham, Switzerland, pp. 155–170, 2019.
33. Jones, T.D., Farlow, J.O., Ruben, J.A., Henderson, D.M., and Hillenius, W.J., Cursoriality in bipedal archosaurs, *Nature* 406:716–718, 2000.
34. Lingham-Soliar, T., The evolution of the feather: scales on the tail of *Sinosauropteryx* and an interpretation of the dinosaur’s opisthotonic posture, *J. Ornithology* 154:455–463, 2013.
35. Chiappe, L.M., Climbing *Archaeopteryx*? A response to Yalden, *Archaeopteryx* 15:109–112, 1997.
36. Hennigan, T., An initial estimate toward identifying and numbering the Ark turtle and crocodile kinds, *ARJ* 7(2014):1–10, 2014.
37. O’Micks, J., A preliminary cephalopod baraminology study based on the analysis of mitochondrial genomes and morphological characteristics, *ARJ* 11:193–204, 2018.

**Matthew Cserhati** has a Ph.D. in biology and a B.Sc. in software development.. He has been an active creationist for 19 years and takes a great interest in molecular biology. He has published a number of articles in Journal of Creation. Matthew is currently studying at Greenville Presbyterian Theological Seminary.

**Brian Thomas** earned a Masters degree in biotechnology in 1999 from Stephen F. Austin State University in Texas, and a Ph.D. in paleobiochemistry from the University of Liverpool in 2019. He taught biology and chemistry as an adjunct professor, then later as an assistant professor at Dallas area universities. He is the author of *Dinosaurs and the Bible* and co-author of *Creation Basics and Beyond* and *Guide to Dinosaurs*. Since 2008 he has contributed hundreds of online science news and magazine articles as the Science Writer at the Institute for Creation Research in Dallas, USA.

**Joel Tay** has a diploma in biotechnology from Ngee Ann Polytechnic, a B.Sc. in both evolutionary biology and genetics from the University of Queensland, a Master of Divinity (Theology) from Trinity Theological College, and a Master of Theology (Systematic Theology) from Southwestern Baptist Theological Seminary. He is currently a speaker for Creation Ministries International, US.