Creation Research Society Quarterly 2022 58:175–192

Statistics, Baraminology, and Interpretations: A Critical Evaluation of Current Morphology-Based Baraminology Methods

Harry Sanders and Matthew Cserhati*

Abstract

Baraminology has been a hotbed of creationist research for the last two decades. Dozens of studies have been published, most using a single methodology, called BDIST. However, there has yet to be a thorough evaluation of the methodology, though a few have raised concerns about its results. This paper reviews the body of work for the most popular statistical baraminology method and finds that it depends on numerous false assumptions and is prone to deliver false results. The method is characterized by evolutionary assumptions and a lack of critical evaluation of the secular literature. Creation science needs to pursue a more accurate method of determining baraminic relationship. Several ideas are suggested, as well as potential revisions to the dominant methodology that might help it return more accurate results.

Key Words: Baraminology, baramins, BDIST, characters, discontinuity

Introduction

Biological diversity has long been a subject of much speculation. It was this diversity that Darwin attempted to explain without a Creator in his *Origin of Species*, first published in 1859 (Darwin, 1872). However, Darwin wrote in absence of any understanding of heredity which Mendel had yet to publish (Mendel, 1866). It was easy for Darwin to conceive of one basic type of animal changing into another over long periods of time. Such is no longer the case with our modern knowledge of genetics (Watson and Crick, 1949), information theory (Gitt, 2005) and genetic entropy (Sanford, 2014). However, creation science still must undertake the task of explaining and understanding both the Biblical kinds, and the diversification that has happened within them since they were created. To do this, a field of creation science has been created called baraminology. The name for this field of study comes from two Hebrew words and one Greek word. The word בָּרָא (barah) means 'he created', and the word מִין (min) means 'kind', or 'type.' The Greek word λογος means 'science.' Therefore, 'baraminology' is the scientific study of the created kinds.

A Brief History of Baraminology

Baraminology has long been a focus of creation science efforts. One of the founders of the Creation Research Society was Dr. Frank Marsh who coined the term 'baramin' to describe created

* Harry F. Sanders, III, Moore's Hill, IN, hsanders@answersingenesis.org
Dr. Matthew F. Cserhati, Rosemead, CA, csmatyi@protonmail.com
Accepted for publication October 18, 2021

kinds in 1941 (Marsh, 1941). Marsh also published works numerous times on the topic of the baramin, from 1964 to 1992. Marsh's ideas served as the foundation for discontinuity systematics (ReMine, 1990) and modern baraminology (Wise, 1990).

Wise built his baraminology on the discontinuity systematics of ReMine, who coined the major terms used in baraminology today such as 'apobaramin,' 'holobaramin,' and 'monobaramin.' According to ReMine, the holobaramin was the created unit of reproduction, described in Genesis 7:2–3, 19–20, and Genesis 1:11, 21–22, and 24, though ReMine deliberately built his system to avoid any reference to the Bible (ReMine, 1990). The apobaramin is a group containing multiple holobaramins, and a monobaramin was a group containing taxa united by ancestry, but which also potentially shared ancestry with other taxa within a holobaramin.

Wise built on this and coined the term "neo-creationist orchard" to explain the discontinuity of baramins as well as premiering the term "baraminology" to name the study of created kinds (Wise, 1990). Wise greatly expanded ReMine's criteria for identifying baramins. For example, he incorporated cladistic methodology to determine baramins, and visualized them using "baraminograms." He further recommended the use of cladistics and the construction of the most parsimonious tree to identify homoplasies, traits which were shared across groups but are not ancestral. He argued that the greater the homoplasies between two groups, the more likely they belonged to different holobaramins (Wise, 1992).

Prior to Wise, statistics had never been formally proposed as a method of baraminology, though the cladistics methods Wise advocated were hardly new. Most baraminologists had been content with Marsh's hybridization criteria (Marsh, 1964; Siegler, 1974) though a few had proposed using DNA in the future when such data became more available (Frair, 1967). However, Wise's statements about using cladistics opened the door for more statistics-based baraminology methods which have come to dominate the field today. A glossary of terms used in this paper are available in Supplementary Glossary 1 at https://github.com/csmatyi/bdist_review.

Methodology

The BDC

The original baraminic distance correlation (BDC) of Robinson and Cavanaugh (1998a) has developed into the baraminic distance and multidimensional scaling (BDIST) of Wood (2008a). For sake of simplicity, BDIST will be used to refer to the suite of statistical baraminology methods that have been curated by Wood on the Core Academy of Science website.

While methods have changed over time, the core idea of measuring baraminic distance has remained central to the methodology. The statistical calculation that measures the number of mismatched characters between taxon pairs is $d_{ij}=m_{ij}/n_{ij}$ where m is the number of characters that do not match between taxa i and j, and n is the number of comparable characters between the taxa. This coefficient produces a fraction which is said to represent the baraminic distance d (Robinson and Cavanaugh, 1998a). The implied assumption is that the more characters in common a taxon pair share, the more closely related they are.

Assuming that all similarity reveals relationship is problematic. It is an explicitly evolutionary assumption, often phrased as 'common design is equivalent to common ancestry.' When evolutionists look at traits that are similar, they assume that the traits are descended from a common ancestor unless there is reason not to (Brooks and McLennan, 1991). This assumption is not always true. Homologous characters can appear in unrelated organisms, which the evolutionists recognize, and account for by appealing to convergent evolution. However, the first assumption when presented with a similar set of characters is that they are the result of ancestry, not convergence.

It is worth pointing out at this stage that statistical baraminologists assume that all similarities are created equal. As one reviewer helpfully pointed out, all creationist statistical baraminology models assume that most shared features will be a result of design in the original baramins, not common ancestry. While this is correct, the equations and algorithms they use do not know this. The equations and algorithms assume all similarity is the same. Therefore, the results of the equations will reflect this assumption.

Wise (1990) points out that phylogenetics, cladistics, and phenetics (aka taximetrics, i.e., observable traits) all cannot see discontinuity, a key element of modern baraminology, and lists several reasons why these methods are not helpful. Yet Robinson and Cavanaugh (1998a) freely admit to using phenetics derived from others as the basis of their classification. Since phenetics is classification based on similarity, by default the BDC, which Robinson and Cavanaugh introduced, assumes that all similarities are the same.

At its core, the BDC and phylogenetics make similar assumptions. While in some cases homologies are ancestral, in others they are artifacts of design. Fully retractable claws, which are among the diagnostic traits of felids (except cheetahs) are a result of common ancestry. Common forelimb structure such as found in humans, apes, whales, and bats is an artifact of design. Assuming that *all* homologous characters are a result of ancestry rather than design as the BDC model does, is predicated on the evolutionary story being true. It also assumes what it is trying to prove, a logi-

It would be valid to argue as follows:

If organisms are the same kind. then we would expect to see continuity.

We see discontinuity. therefore they are not the same kind.

It is however, not logically valid to argue as follows:

If organisms are not the same kind. then we expect to see discontinuity

We see discontuinity. therefore they are not the same kind.

This is affirming the consequent and can be demonstrated in simple terms this way:

If it is raining. the streets are wet.

The streets are wet. therefore it is raining.

Obviously this is not necessarily true. There are a myriad of reasons the streets could be wet. from a broken water main to a child with a garden hose. The same is true of discontinuity and continuity. There are multiple explanations for the existence of both from similar habitats to selective breeding. The discontinuity argument logically cannot demonstrate baraminic relationship because it is based on a fallacy.

Box 1

cal fallacy known as "circular reasoning" or "begging the question" (See Box 1). While Robinson and Cavanaugh (1998a) were explicit in pointing out that discontinuity needed to be looked for, the BDC equation itself assumes that all similarities are created equal. Discontinuity must be determined after the BDC equation is performed. The BDC does not take into account homoplasy.

Even leaving aside the evolutionary assumption of ancestry, the BDC unaided cannot determine discontinuity in 45% of all cases (Wood, 2012). Obviously, since character choice is a major factor, the BDC cannot be expected to determine discontinuity every time, but 55% success rate is hardly reliable. Other tools have been introduced into the BDC, beginning with ANOPA (Cavanaugh and Wood, 2002) and culminating with bootstrapping (Wood, 2008a). More recently, Wood (2020), has debuted an updated version of statistical baraminology, termed BARCLAY that still relies on the BDC model, but updated from one statistical coefficient to another. However, the fact that the basic method fails to determine discontinuity in nearly half of its applications does not inspire confidence in its ability to accurately define baramins.

Critical evaluation of datasets

The BDIST method is statistically driven using mainly morphological characters to determine baramins. Because statistical applications are only as good as the information they are provided, proper data selection is crucial to ensure that correct baramins are produced. It has been suggested that creationists collect their own data rather than simply using the method to reinterpret the data of others (Wood, 2002). While this has been done at least once (Sanders, 2016) for all practical purposes, every study published using the BDIST method has relied almost exclusively on data collected and interpreted by evolutionists. As noted below, these datasets are not unbiased. It would be very easy, for example for an evolutionist to assemble a dataset containing characters we share with chimps, (large brain size, body hair, opposable thumbs, etc.) to demonstrate we are related to chimps, while ignoring the differences (chromosome number, presence/absence of a tail, etc.).

While this reliance on evolutionists to obtain data is somewhat inevitable, given the budget and access constraints faced by many creation scientists, it is important to critically review the data for bias, and filter out improperly defined characters. Unfortunately this is not done and some authors even argue that a particular bias must be demonstrated before considering the possibility of a general bias (Wood, 2011a). Unfortunately, this is an inaccurate view of the evolutionary community, as even members of the community itself have admitted (Winsor, 1994; Todd, 1999). Because evolutionists are not unbiased, as they themselves admit and demon-

strate, it seems reasonable that their bias will affect their datasets. Wood (2011a) even provides an example of this bias from the dataset he used in his paper. Most of the time, however, there is no attempt to determine bias in the data set and it is simply treated as an unbiased set of facts. All this does not mean these data sets cannot be used, but it does mean critical examination is required. The BDIST method does filter out unknown characters if enough taxa in question are missing the character, but this is not critical evaluation of the dataset. It is an evaluation of the completeness, not the accuracy of the data.

Relevance of characters

Proponents of the method might argue that the relevance statistic determines which characters are used is a critical examination of the data. This fails for two reasons. It was not the original purpose of the relevance measurement. The original purpose of the relevance statistic was to measure the completeness of the dataset (Robinson and Cavanaugh, 1998a). The equation is (a_i=x/n) where a_i stands for the relevance of the character in question, *x* represents the number of organisms in the dataset where such a character exists, and *n* equals the total number of organisms in the dataset. The equation has since been used to exclude characters that did not hit a movable, arbitrary relevance cut-off (Wood, 2002). This movable arbitrariness is the second failure of the statistic. Because the relevance statistic does not examine the characters themselves, but merely determines how complete the data set is with respect to the members of the dataset, it cannot be used as a critical evaluation tool. In other words, the relevance statistic does not determine whether the character states in the data set were measured correctly, or in an unbiased fashion. It merely determines what percentage of the characters are present in all taxa contained in the dataset.

There appears to be some confusion on what the relevance statistic does. Because it is calculated using an algorithm, it is blind to the types of data it is presented. All it sees is a sequence of numbers and symbols. If enough taxa have a number instead of a symbol, that character is retained for analysis. Relevance is not impacted by the desire for holism. If it is given a dataset where certain dental characteristics are present for all or most taxa under analysis, but skeletal or morphological characteristics are only rarely found, it will retain the dental characters, but remove the skeletal or morphological ones. In effect, relevance can have the effect of *reducing* holism.

The relevance cut-off originally proposed was 95% (Robinson and Cavanaugh, 1998a), but Wood lowered this to 90% so that some datasets would give better results (Wood, 2008b) and further to 75% in his examination of human fossils (Wood, 2010). Further fossil studies have also used this 75% cutoff (see Aaron, 2014; Garner and Asher, 2018, for examples). A philosophical justification for this drop has not been presented. Fossils are much more difficult to obtain data from than live organisms as they are often disarticulated and lack soft morphological traits. While it is important to increase sample size, lowering the relevance cut-off does not just increase the number of useable characters. It also increases the potential for statistical noise and increases the potential for inaccuracy. By dropping the relevance cut-off twenty percent to increase the number of characters available for fossils, BIDST creates the possibility of an artificially enlarged baramin. The inverse is also true. If the relevance cut-off is too high, the result could be too many baramins. There needs to be a consistent relevance cut-off, preferably one determined through an appeal to an absolute standard, not an arbitrary one. If an arbitrary cut-off must be used, then let it be immovable for all types of data. As the cut-off is currently used, it gives the appearance that researchers are picking a relevance that suits their preferred outcome.

While rejecting characters which are not present in a high enough percentage of the taxa is useful, it does not constitute a critical examination of the dataset as a critic might object. This requires examining multiple datasets collected by different research teams to determine either a consensus or average value of the characters in question. In effect, the creation scientist is taking all the information available in the literature and synthesizing it into one dataset. Obviously no perfect dataset can be obtained in this way. The researcher would need to know all the characters of a given set of organisms to craft a perfect dataset. However, by synthesizing data, the likelihood of errors or biases changing the results is decreased.

Critical evaluation does not mean taking existing data sets and compiling them into a matrix. A matrix will still contain any errors or biases committed by the original authors. Nor does it mean taking multiple individual datasets and comparing their results. One of the reviewers of this paper has objected that Wood (2010) satisfied the criteria for critical evaluation of the data set because he used multiple data sets. In that instance, the results are being critically evaluated, but the data itself is not. In other words, the data was accepted as accurate, run through BDIST, then the results were evaluated. This adds a layer of error potential between the data and the evaluation. This is also true of Wood's work with turtles (2005) and with Felidae (2008b). In fact, the datasets on Felidae were nearly identical as the latter one was copied from the earlier one, with a few additional characters. Neither contained any attempt to average multiple authors' characters, nor synthesize a mean dataset. Therefore, they do not meet the requirements for

critical evaluation of the data. A mean value for common characters will reduce any potential bias, though it may not completely eliminate it.

It is important to realize that the datasets being used in baraminology studies are created by people with biases. As one reviewer was kind enough to point out, this has been demonstrated. McConnachie and Brophy (2008) examined a dataset of Galliformes characters with 102 characters and sixty taxa. Careful work with the BDC and MDS suggested four baramins. Hybridization data linked three of the four groups. This study is very illustrative of the power of bad or biased data. Had McConnachie and Brophy stopped with the BDIST, they would have published results that created Biblically incorrect baramins. Because they did not, their paper demonstrates that the data cannot be implicitly relied upon or considered unbiased.

Alternatively, the baraminologist could take measurements of the characters themselves. While performing the measurements may be impossible due to any number of circumstances, crosschecking is much easier for most organisms, with perhaps the exception of some more obscure groups and fossils. If it is impossible to critically examine the dataset, it is better to admit lack of knowledge than make incorrect pronouncements based on potentially flawed data.

Bootstrapping

Undoubtedly practitioners of the BDIST system will appeal to bootstrapping to justify the data selection. However, bootstrapping, by Wood's own admission merely points to how sensitive the dataset is to random changes in it (Wood, 2008a). By convention, 90% or higher is considered significant. While it is difficult to philosophically justify a particular number as a cutoff for data strength, leaving it open to interpretation does not solve the issue either, as two individuals could look at the same data and come to different conclusions depending on what number they chose to use to represent strong or weak data. As stated with the relevance cut-off, if an arbitrary cut-off must be used, let it be consistent. In this instance, it is consistent, but with relevance it is not and there is no attempt to justify why one is fixed and the other is moveable.

Bootstrapping is an artifact of cladistics that BDIST has borrowed. While not necessarily a bad thing, borrowing from the evolutionists must be done with careful critical evaluation, to ensure that the good is separated from the bad. While some of its practitioners have acknowledged criticisms of the method from the evolutionists themselves, they spent little time rebutting them or examining them, instead choosing to simply accept the consensus of bootstrapping's value uncritically, and apply it to baraminology (Wood, 2008a). They have not addressed the underlying assumptions of the method, nor more than cursorily rebutted the criticisms of the method.

There are, however, serious issues with bootstrapping. It has four underlying assumptions as listed by Felsenstein, who is perhaps the primary popularizer of the method (Felsenstein, 1985):

1. Characters have been selected in an unbiased fashion by the taxonomist;
2. Characters have evolved independently;
3. Character evolution is randomly determined;
4. The phylogenetic program used generates the correct phylogeny and the most parsimonious tree is correct.

All of these assumptions are problematic. Everyone carries a bias and applies that bias to character selection (Winsor, 1994). Often certain characters are deliberately excluded from the dataset for a variety of reasons, which has the potential to bias the data (Sanderson, 1995). Other authors view bootstrapping as implicitly biased and a bad estimate of repeatability (Hillis and Bull, 1993). While in limited situations bootstrapping could be considered a measure of accuracy, bootstrap values cannot be compared from study to study as, due to each study's implicit bias, they will vary unpredictably. Further, as later commentators on his work point out, Felsenstein himself admitted that not all characters arise independently, even assuming evolution to be true (Kluge and Wolf, 1993). Further, the assumption of maximum parsimony is not something evolution is compelled to produce (Ridley, 1986). Parsimony is not even the only method cladists use. Character evolution is also not always random. Diversification in bird body sizes, for example, sometimes has been shown to be non-random (Maurer, 1998). This may explain why phylogenetic trees fail to predict diversity in characters except among close relatives (Scotland et al., 2014), as phylogenetics assumes random character evolution. Because its underlying assumptions are inconsistent with baraminology, bootstrapping fails as a measure of result robustness.

One reviewer of this paper has objected that, because bootstrapping is not being used for evolutionary purposes, its underlying assumptions can be ignored. This is incorrect. If statistical baraminology is going to use bootstrapping, it needs to meet the underlying assumptions of the method, or the bootstrapping results cannot be trusted. Wood used just that logic in replacing the Pearson coefficient with the Spearman coefficient in the updated BARCLAY method (Wood, 2020). As the reviewer pointed out, the assumptions of the bootstrapping model are not included in statistical baraminology. Therefore relying on bootstrapping to justify character selection within statistical baraminology is unjustified.

Character selection

Characters are the underlying force that drives the statistics behind taxonomic methods. The characters are the data

being interpreted by statisticians using the statistics. Thus, selecting characters is critically important. However, with minimal exception, creation scientists never select the characters. The characters come preselected from the evolutionary community in their datasets, mostly accumulated for phylogenetic purposes. Therefore, it is crucial to know how they select their characters and what a character means to them.

To the evolutionist, a character can be any feature of an organism from morphological and anatomical, to ecological and behavioral (Gemeinholzer, 2008). Characters are defined as either evolutionarily stable or volatile, depending on whether they change with time or not and as either ancestral or derived depending on which is assumed to have appeared first. Characters are chosen to fit the purpose of the study at hand (Wiley, 1981) and different character types often produce discordant phylogenies (Ridley, 1986). This makes any result obtained by a statistical analysis of characters suspect as it is difficult to be certain that the correct characters were selected.

As an example of some of the problems with character sampling, an analysis of character selection across 512 phylogenetic studies discovered that systematists were usually very vague when it came to why they selected traits. Worse, they found that taxonomists were selecting characters differently and assuming that everyone else was making selections the same way they were (Poe and Wiens, 2000). Applied to baraminology, such a statement serves as a warning that characters from the same organisms can be vastly different depending on who is doing the measurements. The potential for bias is why critical examination, as detailed above, needs to be performed on data sets. Further, unless it is explicitly detailed in the paper, it should not be assumed that the secularist has used the same methods the reader would have used to obtain their data.

After characters have been selected, they are coded into a matrix where the state of each character is represented numerically. Coding can be done multiple ways, and each comes with its own problems (Strong and Lipscomb, 1999). The code can be represented multiple ways with ones for presence and zeroes for absence, or, more commonly, one number will represent one state of a character, (usually the presumed ancestral state), while a second number will represent a second state and so on (i.e., 1 = round, 2 = square, 3 = flat) (Gemeinholzer, 2008). These numbers are then converted into phylogenetic trees either by graphing by hand or, more commonly, using a computer program.

There are significant issues with a character-based system. One issue is how to assign discrete values to continuous character values, such as length. A larger issue is that numeric characters do not distinguish between traits which are diagnostic and those that are not. For example, the presence of mammary glands is diagnostic of a mammal, yet has the same weight as a non-diagnostic trait, such as teeth attached to the palate, which occurs in both reptiles and fish. This means traits which are truly unique to an organism or group of organisms, and thus could indicate discontinuity, can easily be lost in an unweighted system. They may simply be unable to overcome the background noise generated by the other traits which they either share or lack in common with another group in the analysis. This can lead to inaccurate clustering.

The best-case scenario would be to cover as many characters in as many species as possible. Adding more characters increases the resolution of a study and also decreases stochastic error. However, eliminating systematic error with more characters is not guaranteed. Adding more taxa to a study could make the reconstruction of species relationships more accurate. However, species selection must be even and not be skewed towards some groups over others. If one can choose between more characters or more taxa, choosing more taxa is a better option, at least in phylogenetic theory (Heath, 2008). Experiments need to be done to confirm this in baraminology.

The practitioners of the BDIST recognize that there is an issue with character selection. It was for this purpose bootstrapping was brought into the model, to see whether characters had been reliably selected (Wood, 2008a). However, as noted above, bootstrapping is rife with problems and cannot provide evidence of correct character selection.

Continuity and Discontinuity

The fundamental tenet of statistical baraminology states that species within a baramin are continuous with one another, and discontinuous with all other species. According to the BDIST, both continuity and discontinuity are necessary to define a baramin. Species must be shown to be continuous with one another, and discontinuous with all other species. Continuity is considered additive, not subtractive, evidence. Instead of demonstrating continuity, discontinuity between organisms must be demonstrated to show they are not in the same baramin. Just because continuity cannot be demonstrated does not mean discontinuity is present (Wood et al., 2003). Formally defined, discontinuity is "a significant, holistic difference between two organisms. Two organisms that are discontinuous with respect to each other are found in separate potentiality regions" (Wood et al., 2003).

Discontinuity is demonstrated in potentiality regions. These are particular regions of biological character space. Any possible character takes up a dimension of character space and has a unique position. Each unique design occupies a single point in this multi-dimensional space. Potentiality regions are areas within this space where various organismal designs can be found. Outside these

potentiality regions, no designs can exist. These empty spaces form the borders of discontinuity. Finding these borders in multidimensional space is the work of the BDIST.

There are numerous underlying assumptions here. The most obvious is that discontinuity is a primary tool for identifying the baramin. This raises a question: how is this assumption different than the homology argument evolutionists like to wield? By definition, *discontinuity is difference*. This difference must be significant and holistic (Wood and Murray, 2003). Therefore, similarity within groups discontinuous from other groups is accepted as evidence of relationship.

A reviewer has objected that it is possible that similarity can exist within a baramin without being a product of ancestry. In some baramins, this might be possible, depending on how many members of the baramin were originally created and how they bred post-Fall. In the time since the Fall, particularly with the harsh Flood bottleneck, it seems unlikely, but the premise can be granted as theoretically possible. However, for any baramins that were on the Ark, all members of the baramin must share ancestry by default. Therefore, any similarity in those baramins will be a result of ancestry. Further, Wood et al. (2003) admitted that baramins were to be inferred from similarity. Therefore, at minimum, BDIST assumes similarity is important in determining baraminic relationship.

This argument is a twist on the homology argument, except it also suffers from the problem of character selection. The discontinuity argument assumes that continuity or lack of discontinuity is evidence for ancestry right up to the point where it is not. It is very difficult to define exactly where discontinuity begins without making character selections have weight. This is because characters are not all created equal. Thus, the anal scent glands, which can be viewed as evidence for continuity between Mustelidae and Mephitidae are not weighted as such and are given the same weight as other data that makes them discontinuous, such as DNA which is what was used to separate them, as they once were the same family (Eizirik et al., 2010). However, weighting one character over another introduces the possibility of picking weighted characters in an incorrect or arbitrary fashion which opens the model to criticism. This paradox is unescapable. In fact, one would have to know everything about the species under study to devise the perfect weighting scheme for characters. In reality, this is impossible.

A reviewer has objected that BDIST can correctly determine discontinuity without weighing characters. This argument is flawed because he assumes that BDIST produces an accurate measure of discontinuity, then uses that to argue that the BDIST has produced accurate measures of discontinuity. This is circular logic. If BDIST's accuracy is tested without assuming that it produces accurate results, as will be shown below, it will be demonstrated that it does not always produce an accurate measurement of discontinuity.

Worse, this argument assumes that discontinuity is only a prediction of creation science. Evolution predicts discontinuity as well, though not to the extent that a creation model does. Evolutionists have appealed to discontinuity to explain everything from the difference between modern humans and Neanderthals (Bertorelle et al., 2003) to the difference between two sub-species of the Ural field mouse, *Sylvaemus uralensis* (Chelomina and Atopkin, 2010). Discontinuity can be accommodated within the evolutionary models, something baraminologists stated as far back as 2009 (Wood and Garner, 2009). The only major difference is that evolution expects organisms can cross between potentiality regions over long periods of time.

On a philosophical level, discontinuity fails as well. A reviewer has claimed that discontinuity is measured not at the character level, but at the organismal level. This claim is innaccurate for several reasons. First, the BDIST relies on the BDC which measures how many taxa have a given character, then maps the average as distances in three dimensional space. In essence what this does is get an average, not of the whole organism, but of the characters in the dataset. Given that many disparate organisms share similar anatomical and morphological parts, what is really being mapped is not organismal similarity, but character similarity. Organismal continuity and/or discontinuity is inferred from character similarity or difference. It is not a direct result of the BDIST.

Second, traits themselves do not necessarily predict what the organism is. Given the following traits: aerobic respiration, streamlined body, dorsal fin, caudal fin, pectoral fins, dorsal spine, scales, lack an operculum, spiracle. That list of traits could apply to any number of species of shark. There is no way to know simply from reading it that it was written specifically about *Squalus acanthias*, the spiny dogfish, though the mention of a dorsal spine might provide a clue. Organisms are more than just a list of their parts. The arrangement of their parts matters. Because BDIST separates the parts from their arrangement, it is not able to measure continuity and discontinuity at the organismal level. It can only measure continuity and discontinuity at the level of the trait, which may not hold for the whole organism. While the practitioners of the BDIST may believe that the method is working at the organismal level, that is not, in fact, what is going on. They are making inferences about the organismal level, from the character level.

Holism and Significance

While the desire for statistical significance, which is applied to this model, is certainly understandable, in practice,

it is meaningless. As Wood et al. (2003) point out, because significance is highly dependent on character selection, it was considered less valuable than the holism of the dataset. It is quite possible to have statistically significant results that are inaccurate if the data is limited or biased. Holism was meant to mitigate this issue. However, holism has not been used as intended.

Holism is the idea inherent within the BDIST methodology that datasets should include multiple types of characters (Wood et al., 2003). Ideally, this means that datasets containing different types of morphological data, molecular characters, ecological data, biochemical characters and so on are used. In practice however, this is not the case. Both Wood (Wood, 2002) and Robinson and Cavanaugh (Robinson and Cavanaugh, 1998a) strongly recommended not using molecular characters as they produce incorrect baraminic relationships. It is worth pointing out here that using multiple datasets does not constitute holism. Multiple data *types* are meant to represent holism.

Since most baraminologists are simply recycling secular datasets, it is often very difficult to use a holistic dataset. However, the further baraminology has progressed, the less holism has been applied. A reviewer has objected that holism is still alive and well because multiple datasets are being used. This is spurious for multiple reasons. First, as noted with Wood's (2008b) work on Felidae, many times the datasets are derived from one another and thus use the same characters with a few additions or subtractions. Such datasets add nothing to holism. Second, even if the datasets are not derived from one another, DNA data is rarely used if ever, and ecology data is uncommon. The most common types of data are morphometric and skeletal. By the BDIST's own definition of holism, this is not holistic.

The first use of BDC used four types of data: morphological, ecological, chromosomal, and molecular (Robinson and Cavanaugh, 1998a). Each of the character types had a range of different sub-types. For example, morphology included craniodental, and axial and appendicular skeleton character states. Thus, even when the dataset was revised to remove the molecular characters, the morphological characters were still holistic within the morphology character type. However, it did not take long before holism began to decline. Molecular data mostly went unused after 2002 due to Wood's strong recommendation not to use it (Wood, 2002).

By 2010, holism was all but dead. Wood (2010) published a BDIST analysis that clustered *Homo habilis*, *Homo rudolphensis*, and *Australopithecus sediba* into the human baramin based on craniodental and craniometric characters. This prompted a response by Menton et al. (2010) pointing out that the anatomy of *A. sebida* makes it an australopithecine rather than human.

Wood and O'Micks followed a similar pattern by tentatively assigning *Homo naledi* to the human baramin in 2016 based on craniodental characters, though, in fairness, this was all that was available at the time (Wood, 2016; O'Micks, 2016a). Thus, in spite of dental characters having been shown to be unreliable characters in hominids (Wood, 2013), O'Micks then followed up using a slightly modified BDIST by examining postcranial data and concluded that *H. naledi* was probably not human (O'Micks, 2017). In several follow-up papers, O'Micks demonstrated that, if BDIST is used, *H. naledi* is not in the human baramin, despite strong objections from Wood (O'Micks 2017a; 2017b; Wood, 2017).

Data Reduction Techniques Used in Baraminology

Baraminology studies ideally use data sets with large numbers of characters. Each character is regarded as a separate spatial dimension. As such, each species in a baraminology study is represented by a single point in n-dimensional hyperspace. Therefore, baraminologists can detect baramins in this n-dimensional space based on clustering patterns. Depicting species in one, two, or three dimensions is easy. But beyond three dimensions, this becomes complicated, because we are trying to depict n-dimensional clusters in only three or less dimensions.

The visualization of baramins have used several techniques, including Principle Component Analysis (PCA), Analysis of Patterns (ANOPA), and Multidimensional Scaling (MDS). The goal of these algorithms is to represent n-dimensional data (n being the number of characters in a baraminology data set) in only two or three dimensions. ANOPA and MDS approach these problems in different ways.

PCA

PCA is a widely used data reduction technique, which helps in understanding relationships between traits and discovering patterns within the data. It reduces a larger number of variables into a smaller set of variables with minimal information loss. In baraminology studies, variables represent traits in a baraminology data set. PCA selects the top few (usually 2–4) variables (called principal components), which explain the most variability in the data set, and plots species in 2D or 3D space along the axes of the principal components.

When using PCA there is a tradeoff concerning the number of traits used. If all traits are used in an analysis, the risk of overfitting and false positives increases. However, if only the top few traits are used, which explain the most variation, then information can be lost (Sainani, 2014). This is even more so if the data is highly variable, and the top few principal components can capture a lot less variability. That is why other techniques, such as ANOPA or MDS, are preferred (Wood and Murray, 2003).

ANOPA

For an n-dimensional data set, ANOPA defines a 'centroid' point which represents the average of all data points (species) in the data set. Then a vector t_0 is constructed between the centroid and the 'outlier point,' defined as the species farthest away from the centroid. The axis between the centroid and species i is the first dimension. Next, a vector d_2 is constructed perpendicularly from t_0 to point x_i. The angle α_r defines the angle of rotation between t_0 and d_2 for any given species. This way each species can be represented by three values: d_{t0}, d_2, and α_r. All species in the study can be depicted in a two- or three-dimensional plot using any two or three of these values. One dimensional ANOPA is also possible, if only the distance between the species and the centroids is analyzed. Clusters of species can then be determined based on the way they group in two- or three-dimensional space (Cavanaugh and Sternberg, 2004).

However, ANOPA has been criticized. It doesn't provide any means for determining discrete, discontinuous groups (Bolnick, 2006). For example, a paper using ANOPA found that around 20,000 species of the family Asteraceae were discontinuous from all other organisms (Cavanaugh and Wood, 2002). A solution would be to try out any kind of clustering algorithm on three-dimensional data to resolve this problem.

MDS

MDS is a general statistics technique which aims at determining and depicting similarity between objects in a quantitative manner. In baraminology, MDS starts out from a character matrix with n species and m characters. A dissimilarity (or proximity) value, $\delta_{i,j}$ is calculated for each pair of species. Thus, the original m x n character matrix of m rows and n columns (generally called the configuration matrix) is transformed into an n x n proximity matrix. If $n > 3$, then the n species cannot be depicted in only a two or three dimensional MDS plot, therefore the proximity matrix must undergo dimension reduction. In the final output, species are depicted as objects, for example in three-dimensional space, with $d_{i,j}$ representing the baraminic distance between species i and j (De Leeuw, 2000). The inherent problem with MDS and other dimension reduction (scaling) algorithms is that information is lost and distorted during this process. The distortion of information is called stress. The stress function captures the amount of stress as the data undergoes during reduction:

$$S = \sqrt{\frac{\sum_{i,j}(\delta_{i,j} - d_{i,j})^2}{\sum_{i,j} d_{i,j}^{\ 2}}}$$

It must be emphasized that scaled distances only approximate the true baraminic distances. The stress value describes the goodness of fit between these two sets of values (Wood, 2001). A stress value is calculated for each value of $k > 0$, where k represents the number of dimensions. Where the stress value is the lowest, that is the optimal number of dimensions. These stress values can be depicted on what is known as a 'scree' or 'elbow plot.' If the optimal value of k is much greater than three, there is a high chance that clusters will be distorted in the 3D MDS plot.

Dissimilarity values and distance values should show linear correlation; they would be equal in the case of a perfect fit. Smaller stress values correspond to a better fit. In general, stress values less than 0.1 correspond to a good fit. Where the elbow plot shows minimal stress is where the best fit exists between baraminic distances and scaled distances. Visually speaking, this is the 'elbow' in the elbow plot.

What must be said about MDS is that it is a descriptive technique, and statistical inference is almost completely absent from it. An MDS plot does not determine baraminic relationships. Sometimes a minimal stress value is not evident. Furthermore, objects will be placed on the MDS plot based on the primary dimension, which may or may not represent the clusters well, and like ANOPA, interpretation of clusters is subjective (Hout et al., 2013).

Testing BDIST

An internal critique

If the BDIST system is accurate, it should, given valid data, be able to detect both continuity and discontinuity within a dataset. Wilson (2020, personal communication) suggested it could be possible using the BDIST to detect discontinuity within breeds of domestic dogs. Thus, a test was devised, using data obtained from Jordana et al. (2006). Their dataset consisted of twenty-five dog taxa without an outgroup. The dataset consisted of 42 morphological and behavioral characters. These numbers are in line with other BDIST analysis such as that of Wood (2014) and Ingle and Aaron (2015).

Using a 0.95 relevance cut-off, all taxa and all characters were retained for analysis. Results were also obtained for cutoff values from 0.75 to 0.9, but they were very much like the results for a cutoff of 0.95. Based on this analysis, six separate groups are proposed (See Figure 1). The first consists largely of working dogs such as the Saint Bernard, Great Dane, the retrievers, and so on. There are nine breeds in this group and they divide fairly cleanly into two subgroups, with the Labrador Retriever and Pointer providing the overlap between the groups. Some of these breeds are also discontinuous with other groups.

The second group is very small, consisting of just the Beagle and Basset Hound. The third group consists of the herding dogs, including the German Shepherd and Shetland Sheepdog. Both of these groups are stand-alone groups. The German Shepherd does positively

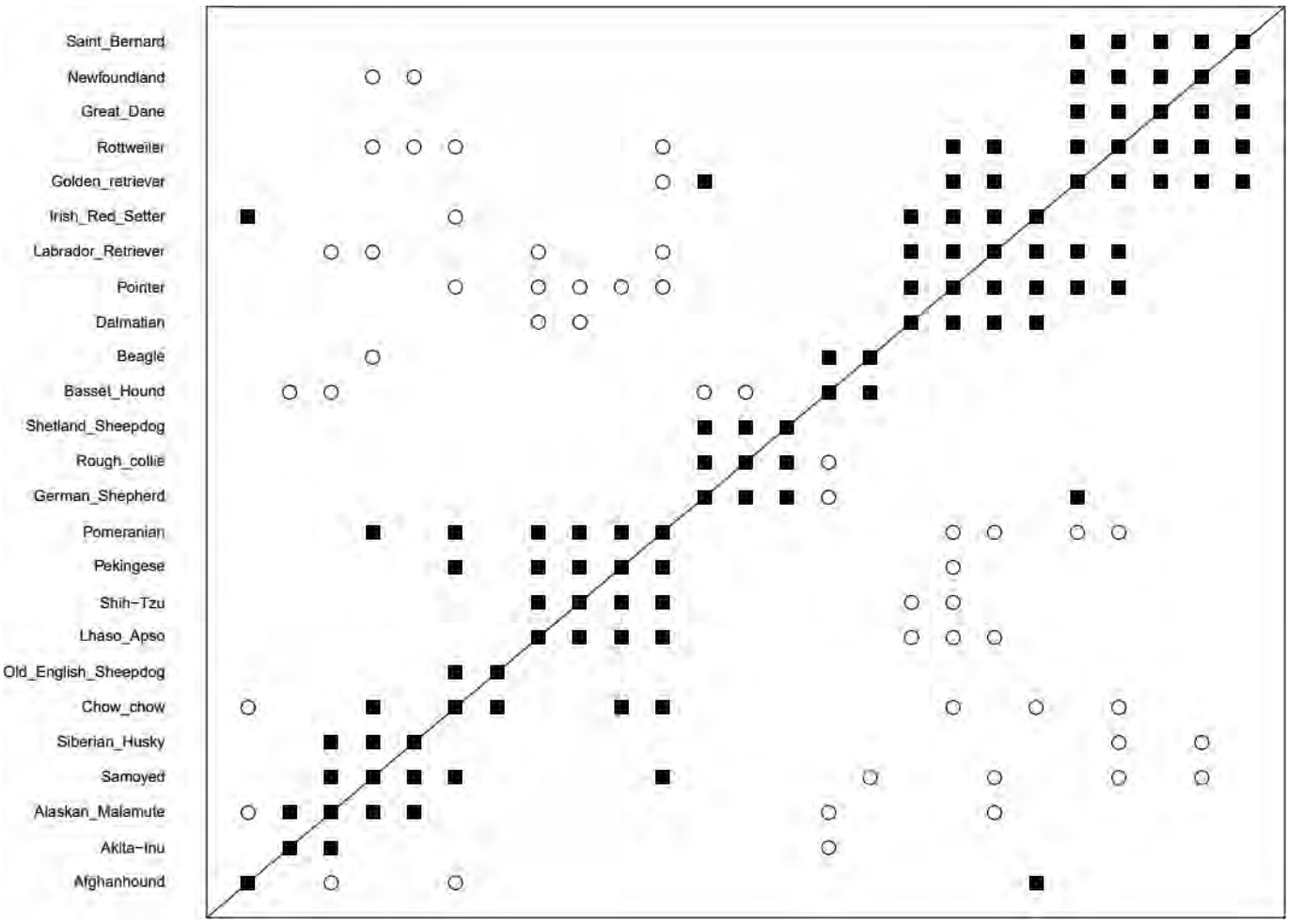

Figure 1. A baraminic distance correlation of the taxa from Jordana et al. and their clustering according to BDIST. Significant positive correlation indicated by filled squares. Significant negative correlation indicated by unfilled circles.

correlate with the Golden Retriever of the working dog group, but not with any other members of the working dog group.

The fourth group consists largely of the smaller toy dogs including the Shih-Tzu, Pekinese, and Pomeranian. The Pomeranian and Pekinese also correlate with the Chow Chow of group five so it is possible group five, which consists of just two breeds, could be merged into group four.

Group six consists of the cold-weather dogs such as the Siberian Husky, Alaskan Malalmute, and so on. Interestingly, the Chow-Chow also correlates with this group, leaving open the possibility that groups four through six could be merged. However, given MDS results (discussed below), this seems unlikely to be correct.

Perhaps more interesting than the potential groupings, given these are all breeds of domestic dogs, is what the BDIST views as significant negative correlation. Twenty-six taxon pairs are considered significantly negatively correlated. Bootstrap values are universally moderate to poor, with only a few in the 80% range and none above 90%. Keep in mind, these are domestic dog breeds. They are members of the same species, *Canis familiaris*, yet the BDIST is showing them as negatively correlated.

Multidimensional scaling is not definitive here, as the taxa show in a cloud of varying different distances, though there are some small apparent clusters (see Figure 2). Minimum stress occurs at nine dimensions. As Wood (2001) points out, this high-dimensional number indicates poor fit between the various data groups. Given how poorly most of the groups are defined, this is not surprising. However, with the strong correlation between the Beagle

and Basset Hound and relatively robust bootstrapping results within that group, they could be tentatively placed in their own baramin and leave the remainder of the groups unresolved.

A reviewer has argued that the dogs cluster based on the MDS and thus the BDIST is perfectly validated by these results. This argument is purely subjective. In Figure 2, there are several potentiality regions splitting small clusters. While it is theoretically possible to argue that what Figure 2 indicates is a single cluster, it is equally valid to argue that there are multiple clusters with small breaks between them. MDS is a subjective tool. The reviewer may choose to interpret the results as he pleases, but given how poorly the data fits and the strong negative correlations between taxa, his interpretation seems less likely than the one presented here.

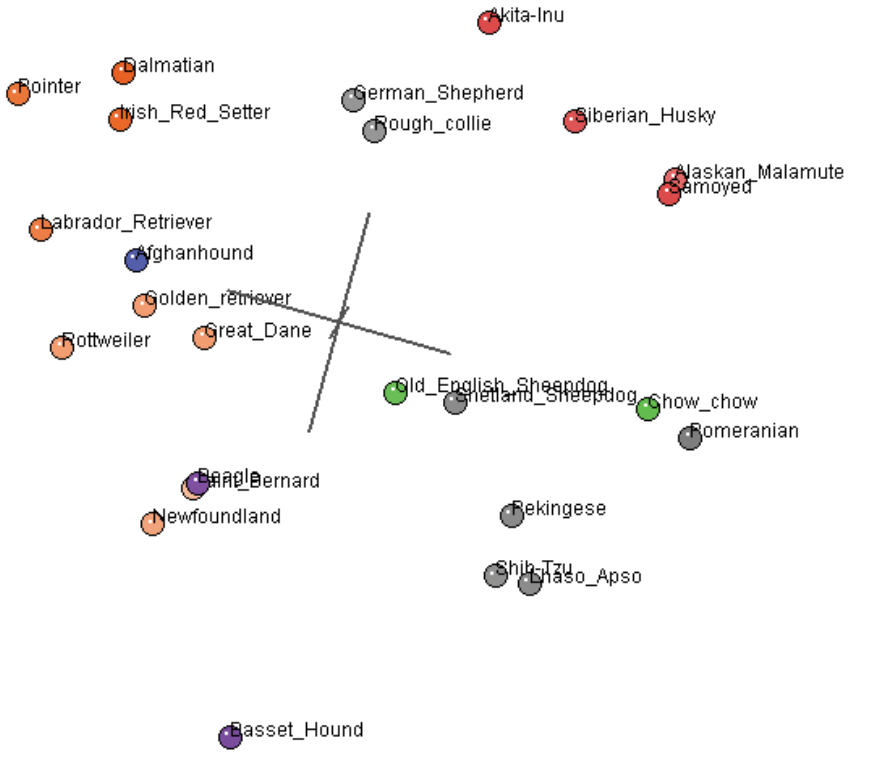

Figure 2. The MDS for the Jordana et al. dataset showing seven clusters. These clusters correspond to the clusters in the graphical output of Supplementary Data File X (Jordana et al., 1999).

Discussion

Of course, the inference drawn from these results, i.e., that different dog groups are unrelated, is incorrect. Empirical evidence from genetics, breeding records and morphology shows that these dogs are members of the same kind. In fact, they are the same species. There is no logical, empirical or Biblical reason to place these dogs in different baramin. Yet the BDIST results lend themselves to such an inference. That inference would be incorrect, but would be accepted if this was a group of organisms about which little was known.

The above analysis illustrates the problems with BDIST and really most purely statistical analysis. They are subjective on multiple levels. The first level of subjectivity is with the data itself, and the potential for bias in collection. The second is within the BDIST calculations, as the equations make assumptions that are not warranted. The third is within the MDS results as these are wholly subjective. The fourth is within the boostrapping convention, as the convention sets the cut-off and is thus subjective. The fifth is with the published baramins as they are interpretations of subjective results. None of these are empirical. While baraminology is a historical science, an empirical foundation would be helpful in ensuring robust results.

It could be argued that the BDIST results are not presenting the toy dogs as a holobaramin, but rather a monobaramin within a larger holobaramin. This is more logical than claiming holobaraminic status certainly. The issue is, many of the dog breeds in question show discontinuity between one another, not similarity. Since as a helpful reviewer pointed out, monobaramins are defined by similarity, not ancestry, it is impossible to argue for a single monobaraminic status here. Thus, explaining this as a monobaramin does not work.

Such aberrant results raise questions about the authenticity of the methodology. There are four options here, none of which are likely to salvage the current iterations of statistical baraminology as a workable methodology. The first option is there is some issue with the data. While this is a possibility, this opens the door to questioning *every* result of the BDIST that has borrowed data from the secular literature. BDIST must assume that the data collected from secular sources is always accurate, else its results are potentially flawed. This would mean that no baraminology study that has used secular data can be deemed empirical, even if it gives expected results. If only expected results are accepted from the method, then confirmation bias is at work and the results cannot be supported logically.

A second option requires rejecting the BDIST methodology. Given its prevalence, the general reluctance to change, and the reviews of this paper, this seems unlikely to occur, at least immediately. As will be shown below, BDC should have been reworked or reconsidered after its first use. Instead, it has become the foundation for the dominant model of baraminology.

The third logical option is accepting the above-presented results as valid. Obviously, this is incorrect. Multiple holobaramins within a single species is something that does not work either with empirical evidence or a Biblical understanding of a baramin.

A fourth potential option leaves some room to potentially salvage the methodology. It could be claimed that the above results are invalid because the taxa in question have undergone generations of artificial selection, making it difficult to determine true discontinuity. This is perhaps the best argument, and one worthy of examining.

Artificial selection, while it does produce a great deal of phenotypic variation as mutations and recessive traits are selected for, does not change a dog into something fundamentally different. For example, even though the Lhasa Apso and the Dalmatian share little in common, it is clear to any observer that both are dogs. The essential nature of the dog kind has not been changed (Joubert, 2011). Thus, discontinuity should not be observed within the canid kind, yet BDIST claims to detect it.

Other incorrect results

Were the above results the only instance of the BDIST producing demonstrably false results, it could be overlooked as an aberration. However, Wood (2012) admitted that the BDC can only detect discontinuity slightly more than half the time. The BDC and its later descendant program, the BDIST, have never been reliably able to determine created kinds.

The very first use of the BDC should have been a warning that all was not well with the methodology. The first use was an attempt to separate catarrhine primates from humans (Robinson and Cavanaugh, 1998a). This is a key test for the validity of the method because we know, Biblically, that apes and humans are not relatives (Genesis 1:26). The original dataset consisted of 204 characters, of which, over half were molecular.

The results were demonstrably false. Humans could not be reliably distinguished from most of the other hominoid primates. Only morphological and ecological data could make such a distinction, which required a post-hoc removal of over half the data from the dataset. Interestingly, no one has repeated this analysis with the newly revised BDIST, though ANOPA was applied to this group with some success (Cavanaugh, 2004). This early failure was an unheeded warning. Instead, it was interpreted as evidence that the genetic data did not need to be considered and the method remained unquestioned.

This first study was followed up the same year by a second study examining the felids. In this study, ecological data was unable to establish discontinuity between felids and the outgroups, while molecular data was (Robinson and Cavanaugh, 1998b). The authors readily accepted the molecular data here because it did not contradict the Scriptures and pointed them to the most logical conclusion, that felids are a holobaramin.

This inconsistency however, between studies, is alarming. It sets up the possibility of a bias accusation because the authors are selecting only data that provides the results they want. Also, illustrating how important character selection is, a later study indicated the felids were, in fact, not discontinuous from non-felids, though the author was quick to defer to the earlier study (Wood, 2008b). Because the results were contradictory, that one study determining felids were a distinct kind and the determined felids were not a distinct kind, it raises questions. Either there is an issue with one or both datasets, or there is an issue with the method.

A reviewer has raised the point that doing multiple studies with felids is an example of statistical baraminology being repeatable. This represents a fundamental misunderstanding of the word "repeatable" in science. When an experiment is deemed "repeatable" it means its results can be replicated by other researchers using similar methods. The felid results use the same method, yet gave disparate results. This is not replication in any scientific sense. It is a repeat of the study, but it does not replicate the results of the original, and, in fact, produces contradictory results. Either felids are a single created kind or they are part of a larger created kind. The two results are fundamentally incompatible.

Despite the change into the BDIST, statistical baraminology has continued to produce Biblically troubling results. A study of the Anserinae (geese and ducks) revealed discontinuity between Cygnini (swans) and the rest of the Anserini (Wood, 2008b). The study was quite robust, with 160 characters used, and low 3D stress for the MDS. Discontinuity surrounded both baramins in the BDIST display results, though MDS displayed a tetrahedral shape. The problem was, there are numerous known hybrids between Cygnini and Anserini, which seems to run afowl of the Biblical implications about bringing forth after their kind (Genesis 1:21). To Wood's credit, he acknowledged this issue and did not break them apart, but the method had failed another test.

A reviewer has objected to the above example, claiming that tetrahedral geometry of the MDS was the reason Wood did not split the Anserini and thus the BDIST worked as intended. This argument is inaccurate because MDS is a purely subjective tool that has no empirical power. Thus MDS cannot be empirically informative in any sense.

Further, the only previous example of tetrahedral shapes in an MDS plot came from Wood's study of Sulidae (gannets and boobies), where Wood (2005) post hoc rejected the dataset because it gave him unexpected results. While it is possible that the tetrahedral shape does indicate that there are issues with the data, this is far from confirmed, especially given the subjectivity of MDS. Rejecting the Anserini results based on tetrahedral geometry is also a post-hoc decision. If the method works, then these post-hoc rejections should not be required on a regular basis. This goes back to the lack of critical evaluation of the dataset. Because there is no critical evaluation, post hoc rejection of results is necessitated.

In 2010 an evolutionary scientist applied BDIST to a dataset of fossil theropods and birds (Senter, 2010). Using an aspect of the BDIST, multidimensional scaling, he was able to demonstrate common ancestry between birds and theropods. Predictably, this prompted a strong response from Wood, which the secular journal permitted him to publish. Wood (2011b) correctly pointed out that multidimensional scaling was not used to determine membership, merely to visualize relationship between taxa. While he conceded that Senter's analysis did lead to a continuum between Mesozoic birds and coelurosaurs, Wood presented an improved analysis of the dataset that pointed to discontinuity between birds and dinosaurs.

Senter (2011) then responded, admitting his original error and, using an improved version of the dataset, found eight very broad "kinds" of dinosaurs, all of which could be related using morphological intermediates. The analysis was robust, using over 100 taxa and nearly 400 characters. Senter actually apologized to Wood and Cavanaugh, who helped him with the study, for, in his mind, utterly destroying creation science. While Senter did not destroy creation science, his work was enough to undermine the BDIST. Senter demonstrated that *scientists can make the BDIST say anything they want it to say*. This subjectivity strips away any vestige of scientific respectability from BDIST. While the BDIST practitioners have published numerous dinosaur and fossil bird baraminologies since Senter's paper (see Cavanaugh, 2011; Wood et al., 2011; Garner et al., 2013), no one has attempted to directly rebut the Senter paper using an unedited BDIST and, indeed, some papers have used BDIST to reach similar conclusions (Doran et al., 2018; McLain et al., 2018).

Can the BDIST method be improved?

Diagnostic features

Another issue is the lack of diagnostic features, which can potentially cloud the analysis. This happens when, for many characters, there are only very few character states, and/or the great majority of species are in the same state for these characters. This kind of character uniformity makes it very likely that many species will match over many characters. This will decrease the baraminic distance between many pairs of species and will lump them together. If species are found in similar habitats and have similar diets, much of their anatomy and physiology may be the same even if they are not from the same baramin. It might be helpful to weigh features generally accepted by the taxonomic community as diagnostic of a group higher than non-diagnostic traits. An example would be increasing the weight of retractable claws in felids since they are much more valuable to identifying cats than habitat choice might be. Experiments would need to be done to determine exact weighing schemes of course.

The issue with weighing characters however, is the potential for bias. However, every attempt to analyze the data uses a weighing scheme, even when all characters are weighed the same (Wheeler, 1986). It is simply conventional to accept equal weighing. It would be useful to experiment with weighing diagnostic characters to see if the BDIST can be improved in that manner, though such would need to be done carefully.

If not BDIST, then what?

BDIST advocates might object that without it, there is no comprehensive baraminology method. This statement is incorrect. There are several other methods available. A system was developed to measure and determine kinds for Answers in Genesis' Ark Encounter and published by Lightner et al. (2011). This system was based mostly on morphology and the cognitum concept developed by Sanders (2010). However, it also incorporated the Marsh's hybridization ideas as the gold standard, based on numerous Biblical passages (incl. Genesis 1:11–12; 8:17–19), as well as using genetic data and some statistics. This is the most holistic method proposed to date, as well as one of the simplest to use, as it is far less labor intensive than some other systems.

There are, however, problems with this method. Organisms that appear similar to one person, may not to another. Further, Wood (2006) points out that hybridization works as an additive, not subtractive, criterion. Lack of hybridization data is not evidence of the organisms being different kinds. While this can potentially be overcome using embryological studies as suggested by Lightner (2007), it remains an issue at present due to lack of data. Perhaps embryological studies would be a good place to focus future baraminological research.

Molecular baraminology

Alternatively, molecular baraminology methods could be proposed. Online databases, such as the NCBI database/EBI/UniProt all contain millions, even billions of sequences. By 2014, the entire genomes of around 14,000 organisms

had been sequenced (Land et al., 2015). By 2015, morphological traits comprised only a mere 2% of all characters in secular phylogenetic analyses. The number of nucleotide positions in molecular data sets has also increased greatly (see Figure 3; Lee and Palci, 2015). So much molecular data has been produced that we do not yet have the resources to even process all of it. It would be helpful to begin incorporating this untouched body of data into baraminology.

Although morphology studies will never go away, the area of molecular baraminology stands wide open before us. Up until recently, some baraminology studies have only examined mitochondrial DNA sequences (Robinson, 1997) or small sequences of DNA (Robinson and Cavanaugh, 1998). The mtDNA is only 16–20 Kbp, only a small fragment compared to the Mbp/Gbp-sized genomes of more complex organisms. Much more information can be extracted from the genome in order to assist in the determination of baraminic membership. Besides these, a handful of studies have appeared, which have examined whole proteomes, and which have analyzed and compared the whole genome sequence of different species (Yaugh, 2017; Cserhati, 2019; Cserhati and Alquist, 2019; Cserhati, 2020a).

The advantages of molecular baraminology studies are strong, compared to morphology studies (see Figure 3). As discussed previously, morphology studies are confounded many times by convergence. It is a well-known fact that the genotype largely defines the phenotype. The genotype contains the core information of the makeup of an organism. Whatever change is made in the genome is followed by changes in the phenotype (Cserhati and Tay, 2019). Furthermore, as opposed to morphology studies, where individual characters are subjectively interpreted, genome sequences constitute the raw hereditary material of an organism. Besides being less subjective, genetic studies are also rich in character data (i.e., the number of genes or nucleotides), and thus greatly reduces stochastic error (Heath, 2008). This, however, assumes that the genetic sequences were sequenced objectively and are not contaminated in anyway, which is not always the case.

A relatively new baraminology tool, the Gene Content Method (GCM) compares the overlapping orthologous protein content of different species. The Jaccard Coefficient Value (JCV), defined as the proportion of common orthologous proteins to all proteins between two species, is a measure of how similar two species are genetically. Genes code for proteins, which themselves fulfill functions in the cell, and which express the phenotype of an organism. Therefore, if two species have more genes/proteins in common, they likely belong to the same baramin. A smaller common gene/protein content is an indication that they belong to different baramins (O'Micks, 2017).

However, the coding sequence makes up only 1–2% of the genome, and there is much more genetic information to make use of than this. The ENCODE project found that virtually 100% of the human genome is associated with some biological function (ENCODE, 2012). That is why the Whole Genome K-mer Signature method was developed, in order to do just this (Cserhati, 2020b). It has the added advantage over the GCM in that protein sequences must be experimentally verified (otherwise researchers must rely on lower quality hypothetical or predicted proteins), whereas the whole genome sequence is generally the first thing that researchers determine when studying an organism. Whole genome sequences are thus easier to come by and provide holistic information about an organism.

The weaknesses of molecular studies are: (1) if the proteome is incomplete, or (2) if the genome sequence is low coverage, or (3) has a large percentage of undefined bases. This is akin to the problem of morphology studies which involve fragmented fossil remains (see Figure 3). Thus, care must be taken when viewing the results of molecular studies to ensure that overreliance is not placed on results of insufficient data. Further, the GCM allows users to select their preferred number of clusters. This has huge potential to bias the results and

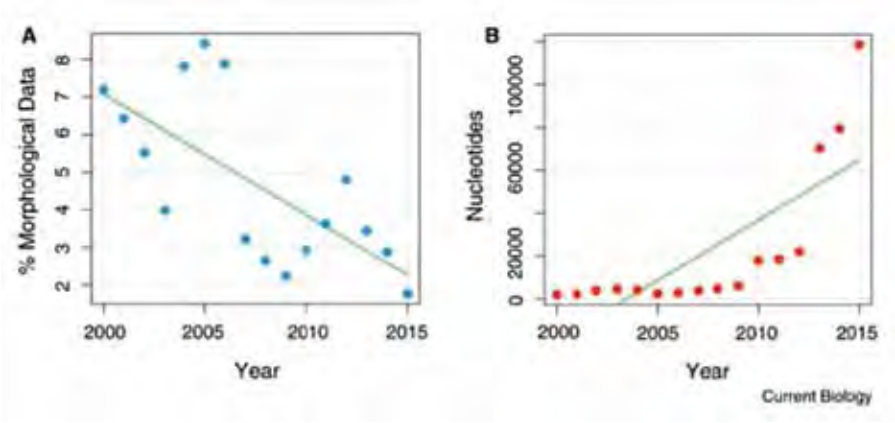

Figure 3. The decline of morphological data sets, and the rise of genetic information used in phylogenetic studies between 2000 and 2015. By 2015, only 2% of studies used morphological characters, whereas the number of nucleotides involved in these studies is increasing exponentially. Figure taken from Lee and Palci (2015).

result in the same selection of data to fit a desired result as the BDIST. Neither method is the answer as yet.

There is a place to combine molecular and morphological studies. Given that the essential nature of an organism has not changed from its creation, despite many variations it has undergone since then (Joubert, 2011), its morphology is indeed informative. Baraminology needs to adapt to the genetic revolution in science and begin incorporating DNA in its analysis, as well as developing new methods to replace the existing ones.

BARCLAY

Recently Wood (2020) has debuted a new baraminology algorithm termed BARCLAY. It has updated from the Pearson coefficient to the Spearman coefficient and incorporated Jaccard distance, along with two new output statistics, PAM and FANNY (Kaufman and Rousseeuw, 1990). The core of the model is still the BDC combined with bootstrapping and multidimensional scaling. Rerunning the domestic dog data presented earlier produced very similar results to those presented above for the traditional BDC. The criticisms presented of the BDIST remain valid for BARCLAY as do those presented against the GCM.

Summary

While the BDIST remains the most popular baraminology method, an honest assessment shows that it is rife with problems. The foundational calculation, the BDC, is flawed because all similarity is considered to be evidence of baraminic relationship. This rules out the possibility of homology being an artifact of design as it will be in many instances in a Biblical model.

Because statistical studies such as the BDIST are highly dependent on the data they are fed, character selection is key to ensuring accurate results. Yet baraminologists almost never select their own data, preferring to rely on the potentially biased character sets produced by evolutionists, which has led to discordant results in the past and undoubtedly will again. However, in molecular studies the DNA sequence is available and the same for all.

The discontinuity/continuity criterion is also limited. There is no reason to *a priori* assume that discontinuity/continuity is a primary criterion of the baramin, and such an assumption brings in a form of the homology problem. Discontinuity should exist between baramins, and continuity should exist within baramins, but it is not necessarily the *a priori* criterion which defines baramins. Attempting to define baramins holistically has its issues and, in practice, has been all but abandoned as a criterion.

The BDIST also fails an internal critique. It demonstrates significant negative correlation between domestic dog breeds as shown above. Further, its published results present contradictory and unbiblical results. It has been used to suggest evolution in dinosaurs and, undoubtedly, could be used to demonstrate evolution in other organisms as well if an evolutionist took the time to do so. As such, it is time for the BDIST to be critically re-evaluated. Changes need to be made to the model, be it diagnostic-character weighing or some other restructuring of the algorithm. BARCLAY is not the answer as it retains all the flaws inherent with BDIST. As it is currently constructed, the model simply does not work.

There is a wealth of potential in baraminology. New methods are rapidly developing. Ideas of Wise, ReMine, and Marsh could also be revisited. It is important to explore that potential and ensure baraminology does not become tied to a system which has a history of producing incorrect results.

References

Aaron, M. 2014. Discerning tyrants from usurpers: A statistical baraminological analysis of Tyrannosauroidea yielding the first dinosaur holobaramin. *Answers Research Journal* 7:463–481.

Bertorelle, G., D. Caramelli, C. Lalueza-Fox, C. Vernesi, M. Lari, A. Casoli, F. Mallegni, B. Chiarelli, I. Dupanloup, J. Bertranpetit, and G. Barbujani. 2003. Evidence for a genetic discontinuity between Neandertals and 24,000 year-old anatomically modern Europeans. *Proceedings of the National Academy of Sciences* 100:6593–6597.

Bolnick, D.I. 2006. ANOPA: 'Statistical' systematics for young-Earth creationists, *Reports of the National Center for Science Education* 26:22–31.

Brooks, D.R., and D.H. McLennan. 1991. *Phylogeny, Ecology and Behavior*. University of Chicago Press, Chicago, IL.

Cavanaugh, D.P. 2004. The cataharrine primates revisited: An ANOPA study. In Sanders, R.W. (editor). *'Discovering the Creator' Proceedings of the Third BSG (Baraminology Study Group) Conference*, R2:13.

Cavanaugh, D.P. 2011. An ANOPA study of coelurosaurian theropods. In Sanders, R.W. (editor). *Proceedings of CBS Annual Conference* 1:1.

Cavanaugh, D.P., and R.V. Sternberg. 2004. Analysis of morphological groupings using ANOPA, a pattern recognition and multivariate statistical method: A case study involving centrarchid fishes. *Journal of Biological Systems* 12(2):137–167.

Cavanaugh, D.P., and T.C. Wood. 2002. A baraminological analysis of the tribe Heliantheae *sensu lato* (Asteraceae) using analysis of pattern (ANOPA). *Occasional Papers of the Baraminological Study Group* 1:1–11.

Chelomina, G.N., and D.M. Atopkin. 2010. Molecular genetic evidence of a deep phylogenetic discontinuity between the Asian and European races of pygmy wood mouse based on the mitochondrial cytochrome *b* gene variation. *Molecular Biology* 44(5):699–708.

Cserhati, M. 2019. Baraminology data filtering method based on entropy measurement and its application in dinosaur and

cephalopod data sets. *Journal of Creation* 33(3):55–65.

Cserhati, M. 2020a. A new baraminology method based on Whole Genome K-mer Signature analysis and its application to insect classification. *Journal of Creation* 34(1):86–95.

Cserhati, M. 2020b. Do lizards and snakes form separate apobaramins? *Journal of Creation* 34(1):96–101.

Cserhati, M., and J. Alquist. 2019. Baraminology suggests cryptic relationships among Caprimulgiformes. *Journal of Creation* 33:58–68.

Cserhati, M., and J. Tay. 2019. Comparison of morphology-based and genomics-based baraminology methods. *Journal of Creation* 33(3):10–15.

Cserhati, M., P. Xiao, and C. Guda. 2019. K-mer based motif analysis in insect species across *Anopheles*, *Drosophila* and *Glossina* genera and its application to species classification. *Computational and Mathematical Methods in Medicine* 2019:425–479. doi: 10.1155/2019/4259479.

Darwin, C. 1872. *On the Origin of Species by Means of Natural Selection, or the Preservation of the Favored Races in the Struggle for Life*. 6th edition. John Murray, London, UK.

De Leeuw, J. 2000. Multidimensional Scaling, UCLA Dept. of Statistics Papers, 1–9.

Doran, N.A., M.A. McLain, N. Young, and A. Sanderson. 2018. The Dinosauria: Baraminological and multivariate patterns. In Whitmore, J.H. (editor). *Proceedings of the Eighth International Conference on Creationism*. 8:404–457. Creation Science Fellowship, Pittsburgh, PA.

Eizirik, E., W.J. Murphy, K.Koepfli, W.E. Johnson, J.W. Dragoo, R.K. Wayne, and S.J. O'Brien. 2010. Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Molecular Phylogenetics and Evolution* 56(1):49–63.

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome, *Nature* 489(7414):57–74.

Frair, W. 1967. Some molecular approaches to taxonomy. *Creation Research Society Quarterly* 4:18–22.

Garner, P.A., T.C. Wood, and M. Ross. 2013. Baraminological analysis of Jurassic and Cretaceous Avialae. In *Proceedings of the Seventh International Conference on Creationism*. Volume 7, Article 7. Creation Science Fellowship, Pittsburgh, PA.

Garner, P.A., and J. Asher. 2018. Baraminological analysis of Devonian and Carboniferous tetrapodomorphs. In Whitmore, J.H. (editor). *Proceedings of the Eighth International Conference on Creationism* 8:458–471. Creation Science Fellowship, Pittsburgh, PA.

Gemeinholzer, B. 2008. Phylogenetic networks. In Junker, B.H., and F. Schreiber (editors), *Analysis of Biological Networks*. Wiley-Interscience, Hoboken, NJ.

Gitt, W. 2005. *In the Beginning Was Information*. Master Books, Green Forest, AR.

Heath, T.A., D.J. Zwickl, J. Kim, and D.M. Hillis. 2008. Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. *Systematic Biology* 57(1):160–166.

Hillis, D.M., and J.J. Bull. 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. *Systematic Biology* 42(2):182–192.

Hout, M.C., M.H. Papesh, and S.D. Goldinger. 2013. Multidimensional scaling. *WIREs (Wiley Interdisciplinary Reviews) Cognitive Science* 4(1):93–103.

Ingle, M.E., and M. Aaron. 2015. A baraminic study of the blood flukes of family Schistosomatidae. *Answers Research Journal* 8:327–337.

Jordana, J., X. Manteca, and O. Ribo. 1999. Comparative analysis of morphological and behavioral characters in the domestic dog and their importance in the reconstruction of phylogenetic relationships in canids. *Genetics and Molecular Biology* 22(1):49–57.

Joubert, C. 2011. Created kinds and essential natures: A Biblical and philosophical response to evolutionists. *Answers Research Journal* 4:103–112.

Kaufman, L., and P.J. Rousseeuw. 1990. *Finding Groups in Data*. John Wiley & Sons, New York, NY.

Kluge, A.G., and A.J. Wolf. 1993. Cladistics: What's in a word? *Cladistics* 9(2):183–199.

Land, M., L. Hauser, S.R. Jun, I. Nookaew, M.R. Leuze, T.H. Ahn, T. Karpinets, O. Lund, G. Kora, T. Wassenaar, S. Poudel, and D.W. Ussery. 2015. Insights from 20 years of bacterial genome sequencing. *Functional & Integrative Genomics* 15(2):141–161.

Lee, M.S.Y., and A. Palci. 2015. Morphological phylogenetics in the genomic age. *Current Biology* 25(19):R922-R929.

Lightner, J.K. Identification of species within the cattle monobaramin (kind). *Journal of Creation* 21(1):119–122.

Lightner, J.K., T. Hennigan, G. Purdom, and B. Hodge. 2011. Determining the Ark kinds. *Answers Research Journal* 4:195–207.

McConnachie, M., and T.R. Brophy. 2008. A baraminological analysis of the landfowl (Aves: Galliformes). *Occassional Papers of the Baraminology Study Group* 11:9–10.

McLain, M.A., M. Petrone, and M. Speights. 2018. Feathered dinosaurs reconsidered: New insights from baraminology and ethnotaxonomy. In Whitmore, J.H. (editor). *Proceedings of the Eighth International Conference on Creationism* 8:472–515. Creation Science Fellowship, Pittsburgh, PA.

Marsh, F.L. 1941. *Fundamental Biology*. Self-published, Lincoln, NE.

Marsh, F.L. 1964. The Genesis kinds in the modern world. *Creation Research Society Quarterly* 1:30–38.

Maurer, B.A. 1998. The evolution of body size in birds. I. Evidence for non-random diversification. *Evolutionary Ecology* 12:925–934.

Mendel, G. 1866. Experiments in plant hybridization. *Proceedings of the Natural History Society of Brunn* 4:3–47.

Menton, D.N., A. Habermehl, and D.A. DeWitt. 2010. Baraminological analysis places *Homo habilis*, *Homo rudolpfensis*,

and *Australopithecus sebida* in the human holobaramin: Discussion (of Wood, T.C. 2010). *Answers Research Journal* 3:153–158.

O'Micks, J. 2016a. Preliminary baraminological analysis of *Homo naledi* and its place within the human baramin. *Journal of Creation Theology and Science Series B: Life Sciences* 6:31–39.

O'Micks, J. 2016b. *Homo naledi* probably not part of the human holobaramin based on baraminic re-analysis including postcranial evidence. *Answers Research Journal* 9:263–272.

O'Micks J. 2017a. Rebutal to "Reply to O'Micks concerning the geology and taphonomy of the *Homo naledi* site" and "Identifying humans in the fossil record: A further response to O'Micks." *Answers Research Journal* 10:63–70.

O'Micks, J. 2017b. Further evidence that *Homo naledi* is not a member of the human holobaramin based on measurements of vertebrae and ribs. *Answers Research Journal* 10:103–113.

O'Micks, J. 2017c. Baraminological classification based on gene content similarity measurement. *Creation Research Society Quarterly 54(1)*:27–37.

Poe, S., and J.J. Wiens. 2000. Character selection and the methodology of morphological phylogenetics. In Wiens, J.J. (editor). *Phylogenetic Analysis of Morphological Data*. Smithsonian Institution Press, Washington, D.C.

ReMine, W. 1990. Discontinuity systematics: A new methodology of biosystematics relevant to the creation model. In Walsh, R.E. (editor). *Proceedings of the Second International Conference on Creationism* 2:207–213. Creation Science Fellowship, Pittsburgh, PA.

Ridley, M. 1986. *Evolution and Classification*. Longman Group Limited, London, UK.

Robinson, D.A. 1997. A mitochondrial DNA analysis of the Testudine apobaramin. *Creation Research Society Quarterly* 33(4):262–272.

Robinson, D.A., and D.P. Cavanaugh. 1998a. A quantitative approach to baraminology with examples from the catarrhine primates. *Creation Research Society Quarterly* 34(4):196–208.

Robinson, D.A., and D.P. Cavanaugh. 1998b. Evidence for a holobaraminic origin of cats. *Creation Research Society Quarterly* 35(1):1–14.

Sainani, K.L. 2014. Introduction to principal components analysis. *American Academy of Physical Medicine & Rehabilitation* 6(3):275–278.

Sanders, R.W. 2010. A quick method for developing a cognitum system exemplified using flowering plants. *Occasional Papers of the BSG (Baraminology Study Group)* 16:1–63.

Sanders, R.W. 2016. Evidence for holobaraminic status of the Verbenaceae (Verbena family). Journal of Creation Theology and Science, Series B: Life Science 6:81–90.

Sanderson, M.J. 1995. Objections to bootstrapping phylogenies: A critique. Systematic Biology 44(3):299–320.

Sanford, J. 2014. *Genetic Entropy*. FMS Publications, Livonia, NY.

Scotland, R.W., S. Kelly, and R. Grenyer. 2014. Phylogenetic trees do not reliably predict feature diversity. *Diversity and Distributions* 20(5):600–612.

Senter, P. 2010. Using creation science to demonstrate evolution: Application of a creationist method for visualizing gaps in the fossil record to a phylogenetic study of coelurosaurian dinosaurs. *Journal of Evolutionary Biology* 23(8):1732–1743.

Senter, P. 2011. Using creation science to demonstrate evolution 2: Morphological continuity within Dinosauria. *Journal of Evolutionary Biology* 24(10):2197–2216.

Siegler, H.R. The magnificence of kinds as demonstrated by canids. *Creation Research Society Quarterly* 11:94–97.

Strong, E.E., and D. Lipscomb. 1999. Character coding and inapplicable data. *Cladistics* 15(4):363–371.

Todd, S.C. A View from Kansas on that Evolution Debate. *Nature* 401:423.

Watson, J.D., and F.H. Crick. 1953. A Structure for Deoxyribose Nucleic Acid. *Nature* 171(4356):737–738.

Wheeler, Q.D. 1986. Character weighting and cladistic analysis. *Systematic Zoology* 35(1):102–109.

Wiley, E.O. 1981. *Phylogenetics*. John Wiley & Sons, New York, NY.

Winsor, M.P. 1994. The Lessons of History. In Scotland, R.W., D.J. Siebert, and D.M. Williams. *Models in Phylogeny Reconstruction*. Clarendon Press, Oxford, UK.

Wise, K. 1990. Baraminology: A young Earth creation biosystematic method. In Walsh, R.E. (editor). *Proceedings of the Second International Conference on Creationism* 2:345–358. Creation Science Fellowship, Pittsburgh, PA.

Wise, K. 1992. Practical baraminology. *Creation Ex Nihilo Technical Journal* 6:122–137.

Wood, T.C. 2002. A baraminology tutorial with examples from the grasses (Poaceae). *Technical Journal* 16(1):15–25.

Wood, T.C. 2005. Visualizing baraminic distances using classical multidimensional scaling. *Origins* 57:9–29.

Wood, T.C. 2006. The current status of baraminology. *Creation Research Society Quarterly* 43(3):149–158.

Wood, T.C. 2008a. Baraminic distance, bootstraps, and BDISTMDS. *Occasional Papers of the Baraminology Study Group* 12:1–17.

Wood, T.C. 2008b. *Issues in creation: Animal and plant baramins*. Wipf & Stock, Eugene, OR.

Wood, T.C. 2009. Natura Facit Saltum: The Case for Discontinuity. In Wood, T.C., and P.A. Garner. *Genesis Kinds: Creationism and the Origin of Species*. Wipf & Stock, Eugene, OR.

Wood, T.C. 2010. Baraminological analysis places *Homo habilis*, *Homo rudolphensis*, and *Australopithecus sebida* in the human holobaramin. *Answers Research Journal* 3:71–90.

Wood, T.C. 2011a. Baraminology, the image of God, and *Australopithecus sebida*. *Journal of Creation Theology and Science, Series B: Life Science* 1:6–14.

Wood, T.C. 2011b. Using creation science to demonstrate evolution? Senter's strategy

revisited. *Journal of Evolutionary Biology* 24(4):914–918.

Wood, T.C. 2012. What can baraminic distance correlation really tell us? In Francis, J. (editor). *Proceedings of CBS Annual Conference* 2:5–6.

Wood, T.C. 2013. The value of dental characters for resolving the baraminic status of *Australopithecus sebida*. In Francis, J. (editor). *Proceedings of CBS Annual Conference* 3:4–5.

Wood, T.C. 2014. Two kinds of varmints: Preliminary baraminological analysis of racoons (*Procyon lotor*) and opossums (*Didelphis virginiana*). *Journal of Creation Theology and Science, Series B: Life Science* 4:12–18.

Wood, T.C. 2016. An evaluation of *Homo naledi* and "early" *Homo* from a young age creationist perspective. *Journal of Creation Theology and Science, Series B: Life Sciences* 6:14–30.

Wood, T.C. 2017. Identifying humans in the fossil record: A further response to O'Micks. *Answers Research Journal* 10:57–62.

Wood, T.C. 2020. Baraminology and cluster analysis. *Answers Research Journal*, in press.

Wood, T.C., and Murray, M.J. 2003. *Understanding the Pattern of Life: Origins and Organization of the Species*. Broadman & Holman, Nashville, TN.

Wood, T.C., M. Ross, and P.A. Garner. 2011. Preliminary baraminological analysis of Jurassic and Cretaceous Avialae. In Sanders, R.W. (editor). *Proceedings of CBS Annual Conference* 1:25–26.

Wood, T.C., K.P. Wise, R. Sanders, and N. Doran. 2003. A refined baramin concept. *Occasional Papers of the BSG (Baraminology Study Group)* 3:1–14.

Yaugh, A. 2017. Baraminological analysis of a set of Archaea species based on genomic data. *Creation Research Society Quarterly* 53:140–154.