Mitogenomic Baraminology Analysis and the Search for the Baraminic Demarcation Values Among Class Mammalia

Matthew Cserhati

Key Words: boundary of kind mitochondrial genome, molecular baraminology, sequence similarity, taxonomy

Accepted for publication February 25, 2026

Abstract

Molecular baraminology is a subdiscipline within the science of baraminology. Baraminology in general deals with how to classify organisms into different kinds, or baramins. Molecular baraminology uses a sequence similarity cutoff value to either join two species into the same baramin or separate them into two different baramins. Many mitochondrial and chloroplast genome baraminology studies have been performed, but a specific cutoff value has not been determined. To define this cutoff value, one can measure the range of sequence similarity values among species within the same kind across a large dataset.

In this study, the mitochondrial genomes of 1,005 mammalian species were compared to one another and stored in an online database. For any kind, the lowest sequence similarity value was taken as a possible cutoff value. The distribution of all such minimum sequence similarity values was analyzed over various kinds at the taxonomic level of the genus, family, or order.

For all three taxonomic levels the distribution of minimum sequence similarity values was normal, without showing multimodality. The correlation between the minimum sequence similarity and the number of genera/species in the family or the genus was weakly negative. The correlation between the minimum sequence similarity and the number of species in the order was strong and negative. Since the level of the kind rarely reaches that of the order, only genera and families were examined in more detail. The lowest minimum sequence similarity values for genera and families were 83.2% and 75%, respectively. Since the level of the kind for the mammal species used in the database was between the genus and the family, the sequence similarity cutoff value whereby two species can be classified into either the same or different baramins is within the range of 75–83.2%.

Also, the distribution of mtDNA sequence similarity values for several hybrid mammals was analyzed to help determine the sequence similarity cutoff value. Hybridization is a strong indication that two species belong to the same kind. The minimum sequence similarity found in this analysis was 86.6%.

Furthermore, a cumulative proportion function curve was plotted to depict the proportion of mtDNA sequence similarity values over series of sequence similarity values. Two inflection points between the values of 75–85% indicate that this might be the cutoff range that determines whether two species belong to the same or different kinds.

Lastly, an online app called the Mitogenome Database was developed that allows users to perform their own mitochondrial DNA-based baraminology studies.

Simple Summary

This research explores molecular baraminology, a field that uses genetic data to classify organisms into their original "created kinds" based on Biblical interpretation. By analyzing the mitochondrial genomes of over 1,000 mammalian species, the study seeks a specific sequence similarity cutoff to distinguish between kinds. Data from various taxonomic levels and hybridization cases suggest that species within the same kind typically share a genetic similarity between 75% and 83.2%. The author concludes that the "kind" usually aligns with the biological rank of genus or family rather than the broader category of order. Additionally, the study introduces an online database and analysis tool designed to help other researchers perform similar mitochondrial DNA comparisons. These findings aim to provide a standardized, empirical framework for identifying the boundaries of biological diversity.

Introduction

Baraminology is the study of taxonomy based on a Biblical, literal interpretation of the book of Genesis. In Genesis 1:21, we read: "So God created great sea creatures and every living thing that moves, with which the waters abounded, *according to their kind*, and every winged bird according to its kind. And God saw that it was good." A kind is a group of species that are continuous with one another and discontinuous with all other species. To use a technical term, a kind is also called a 'baramin.' Baraminology is the study of these created kinds, and the goal of baraminology is to determine the membership of these kinds.

Baraminology has come quite a long way in developing algorithms, methods, and software to determine the species membership of the kind. Such studies include the Gene Content (GC) method (O'Micks, 2017), the Whole Genome K-mer Signature (WGKS) method (Cserhati, 2020), and mitochondrial and chloroplast genome-based studies (Robinson, 1997; O'Micks, 2018; Cserhati, 2023a, 2023b, 2023c, 2023d, 2024a, 2024b; Di Martino et al., 2024). Morphology-based baraminology tools include the online BARaminology and Cluster AnalYsis (BARCLAY) software (Wood, 2021). However, a specific critical cutoff value remains ever elusive for mitogenome and chloroplast genome studies. This minimum sequence similarity cutoff value predicts that two species belong to the same holobaramin (a group that includes all species in a kind) if their sequence similarity (SS) value is above this limit. Species with an SS value below this cutoff value belong to separate holobaramins.

The reason this cutoff value has not been discovered is likely due to several reasons. In prokaryotes, for example, the gene structure and content lower the shared gene content of two prokaryotic organisms, despite the possibility that they may belong to the same kind. In addition, gene pleiotropy (a single gene influencing multiple distinct traits) may make two organisms look very dissimilar, despite their high common genetic content. God could also have created several externally similar-looking kinds, but which have different genetics (Cserhati and Carter, 2020).

The purpose of the present analysis is to calculate the minimum sequence similarity (MSS) value for mammalian groups having mitochondrial DNA (mtDNA) in the National Center for Biotechnology Database (NCBI). The analysis was performed over the ranks of genus, family, and order since the level of the kind is somewhere within these limits. The MSS value over all mammalian genera, families, and orders should be a good empirical calculation of the baraminic cutoff value for the kind because it is a highly diverse class of animals.

Hybridization is the strongest evidence of two animals belonging to the same kind. In light of this, several mammalian hybrids were analyzed to see how similar they are based on their mitogenome sequences.

Materials and Methods

The mtDNA sequences for a total of 1,005 mammalian species were downloaded from the NCBI database. Using a shell script, each of the complete 1,005 mammalian mtDNA sequences was compared to one another using the bl2seq program. These values were uploaded into a SQLite3 table. The NCBI taxonomy data (including NCBI taxonomy ID, species, genus, family, order, class, phylum, kingdom, and domain information) for all 1,005 mammals were downloaded from https://ftp.ncbi.nih.gov/pub/taxonomy/new_taxdump/new_taxdump.tar.gz. Taxonomic data was extracted from the file rankedlineage.dmp. This data was then uploaded into a SQLite table called taxa. Various SQLite3 queries were run on these tables to calculate results.

Table I. General statistics describing the distribution of minimum sequence similarity values over 57 mammalian families.

Taxonomic Rank	Min SS	Mean SS	Median SS	Max SS
Genus	83.2%	91.2%	91.7%	98.6%
Family	75%	84.4%	84.7%	95.5%
Order	72.5%	78.7%	78.4%	88.3%

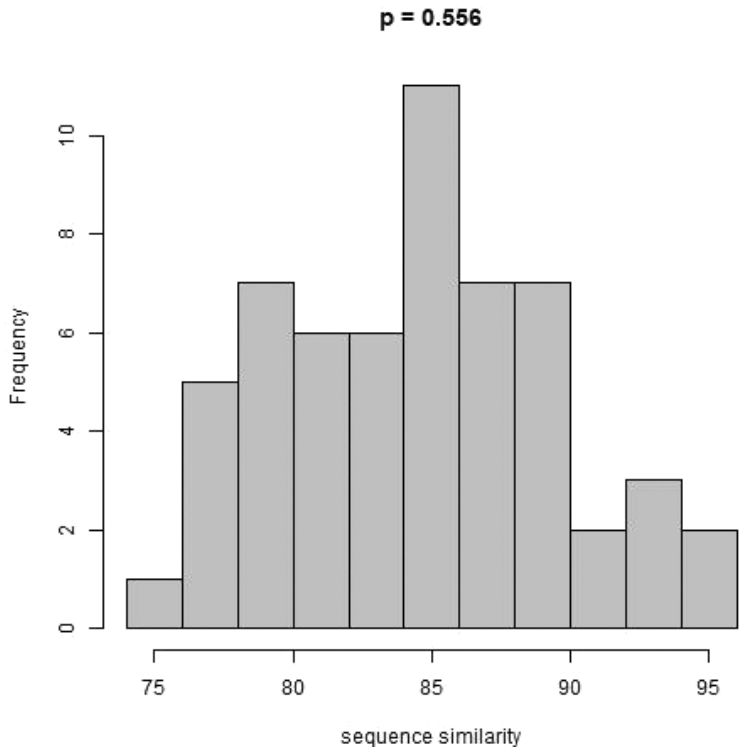

Figure 1. Histogram of all MSS values calculated at the family level.

Only genera, families, and orders with at least three species members were included in the analysis. The Shapiro test was run on the set of minimum sequence similarity values and displayed in a histogram and a linear regression plot. The hist and the plot commands were used in R (version 4.3.1) to depict the histogram and the linear regression plot. The diptest method from the diptest package was used to test sequence similarity values for multimodality.

Supplementary files are available on Zenodo at https://zenodo.org/records/13431371. The Mitogenome Database was programmed in R as part of a shiny app for the online Molecular Baraminology Analysis Tool (MBAT) app at https://molbar.shinyapps.io/molbar.

Results and Discussion

Minimum sequence similarity values across families

In Table I, the MSS values broken down across 57 mammalian families range from 75% to 95.5%. Therefore, the minimum MSS value is 75% for families, which, according to this study, would be the empiric cutoff value that decides whether two species belong to the same or different holobaramins (if the kind is equivalent to the family). Figure 1 depicts a histogram of all MSS values at the family level. The p-value of the Shapiro test is 0.556, which indicates that the distribution is close to normal. However, the histogram does appear to be slightly trimodal, with peaks at MSS values of 79%, 85%, and 93%. However, the Hartigan dip test had a D-value of 0.035, and a p-value of 0.902, which indicates multimodality is likely an artifact.

A question of interest was are the number of species or genera in a family and the MSS values correlated? The correlation value for number of species in a family is -0.343, which indicates a moderate negative effect (see Figure 2A). In general, a larger number of species per family corresponds to lower MSS values and increases the range of MSS values. The correlation value for the number of genera in a family is –0.36, (Figure 2B) where more genera per family also produces lower MSS values and increases the range of MSS values.

In cats (Felidae), the MSS value is 87.1% which includes 15 species in 8 genera (*Acinonyx*, *Catopuma*, *Felis*, *Lynx*, *Otocolobus*, *Prionailurus*, *Neofelis*, *Panthera*), all belonging to the cat kind. In dogs (Canidae), the MSS value is 84.9%, covering 9 species from 5 genera (*Canis*, *Nyctereutes*, *Otocyon*, *Urocyon*, *Vulpes*). In horses (Equidae), the MSS value is 92.9%, encompassing 9 species from a single genus (*Equus*).

Cserhati (2021) studied the molecular and morphological baraminology

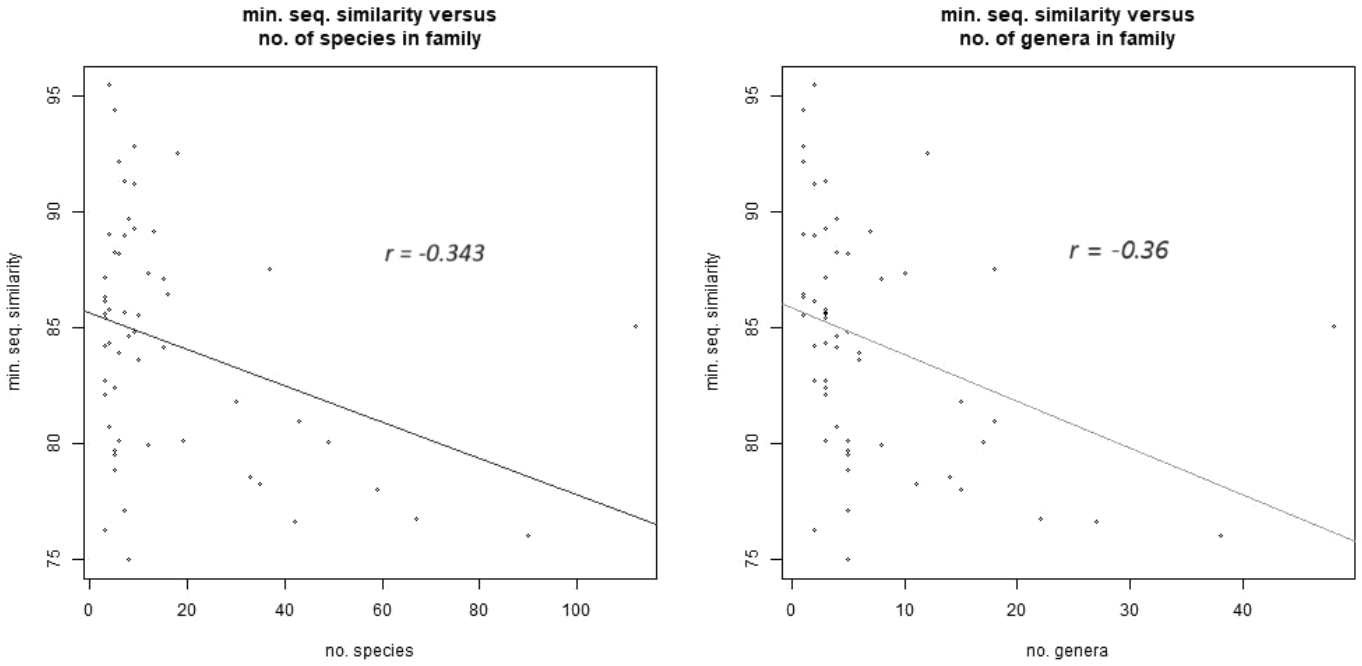

Figure 2. A. MSS according to the number of species in a family. B. MSS according to the number of genera in a family.

Table II. MSS values from several well-known and studied mammal families.

Family	No. species	No. genera	MSS value
Canidae	9	5	84.9%
Equidae	9	1	92.9%
Felidae	15	8	87.1%
Hominidae	8	4	84.7%
Pteropodidae	43	18	81.0%
Rhinolophidae	16	1	86.5%
Vespertilionidae	59	15	78.1%

of bats and found four different kinds. Of these, Pteropodidae (megabats), Rhinolophidae (horseshoe bats), and Vespertilionidae (vesper, or nosed bats) (Table II) had data in the Mitogenome Database. These families had a MSS value of 81%, 86.5%, and 78.1%, respectively. From Table II the lowest MSS value is 78.1%, which is greater than the 75% calculated for all mammalian families in the database. Table II lists the number of species and genera for seven well-known mammalian holobaramins. For statistics for all 117 families in the Mitogenome Database, see Supplementary File 4.

Genus level

At the genus level, the range of MSS values was somewhat higher, between 83.2% and 98.6%. The spread of MSS values is bimodal, with a Shapiro's p-value of 0.066, which is barely above the 5% cutoff limit (Figure 3). According to the dip test, the D-value is 0.035, with a p-value of 0.573, which is not bimodal. This suggests the genus is not equivalent to the kind. However, some genera do coincide with the kind, such as the genus *Homo*, including modern and archaic humans, such as Neanderthals, Denisovans, and *Homo heidelbergensis* (Lightner and Cserhati, 2019; Cserhati, 2023b). Based on this and previous baraminological analyses (Lightner and Cserhati, 2019), the great ape kind should be designated as the family Pongidae, and the human kind should be designated as Hominidae. The correlation between the number of species within a genus and the MSS value is -0.374 (Figure 4), where lower MSS values coincide with greater genus diversity.

Order level

Things are slightly different at the level of the order. Out of 25 orders, only 15 had more than two species so the

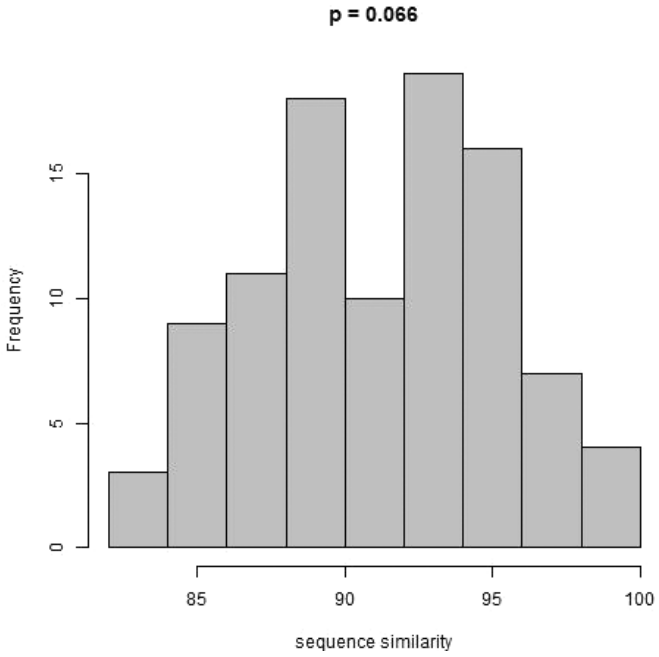

Figure 3. Histogram of all MSS values calculated at the genus level.

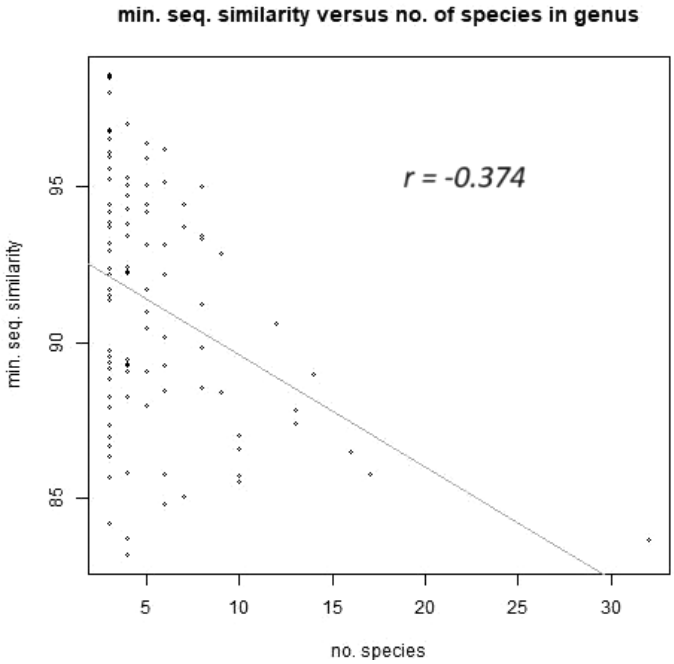

Figure 4. MSS according to the number of species in a genus.

Table III. General statistics for SS values according to genus, family, and order.

Taxonomic rank	Number of taxa	Number of taxa with > 2 species	Shapiro p-value	Correlation
Genus	491	97	0.066	-0.374
Family	117	57	0.556	-0.343
Order	25	15	0.358	-0.634

results from the analysis of the order do not have enough statistical power. The MSS value range is 72.5% to 88.3%. The distribution of SS values is normal, with a Shapiro's p-value of 0.358 (Figure 5). The D-value of the dip test is 0.107, with a p-value of 0.13 (indicating that the data is not multimodal). However, the correlation between the number of species in an order and the MSS value is -0.634, which is a strong correlation (Figure 6) where an order with a large species diversity will possess a lower MSS value.

Mitogenome Analysis Tool app

The mtDNA sequence similarity values have been made available in an online R Shiny app for other researchers to analyze. The Molecular Baraminology Analysis Tools (MBAT) app has a new section on Mitogenome Analysis. This section has two tools: a Taxonomy Database and a Heatmap Generation tool. The app is available at https://molbar.shinyapps.io/molbar.

The user can browse the Taxonomy Database to see what kinds of organisms are available in the app. A total of 6,336 vertebrates have their taxonomic information listed in the database. Table IV lays out the number of species from various tetrapod groups with sequence similarity data in the database. Taxonomic information includes the organism's full taxonomic classification, including species, genus, family, order, class, phylum, kingdom, and domain, as well as its NCBI taxonomy ID. The user can organize the data based on any of the eight taxonomic levels or locate specific data utilizing a search tool (Supplementary Figure 1). The heatmap tool allows the user to select a set of species to generate a heatmap, a sequence similarity matrix, and an SS value histogram (Supplementary Figures 2A–C).

The user has two options to select a species. The first option is to select a species from the drop-down list and then select an MSS value cutoff between 0 and 1. The app will then search for all other species that have an MSS value with the selected species that are greater than or equal to the

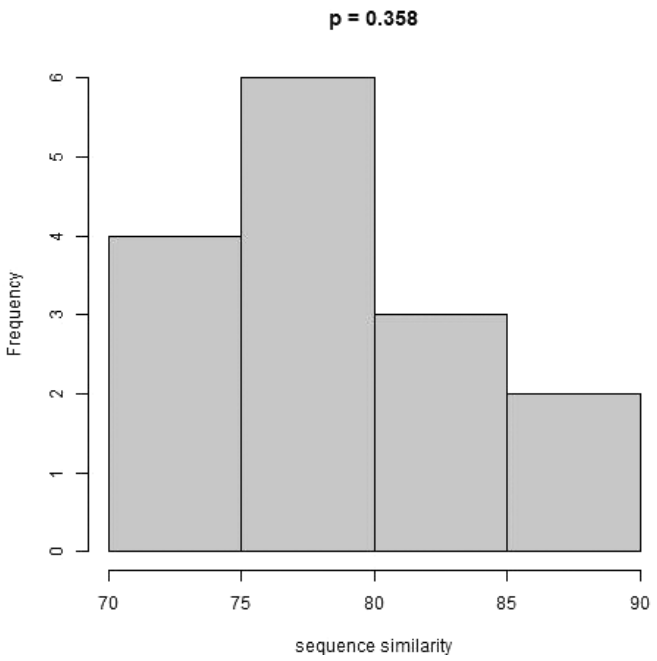

Figure 5. Histogram of all MSS values calculated at the order level.

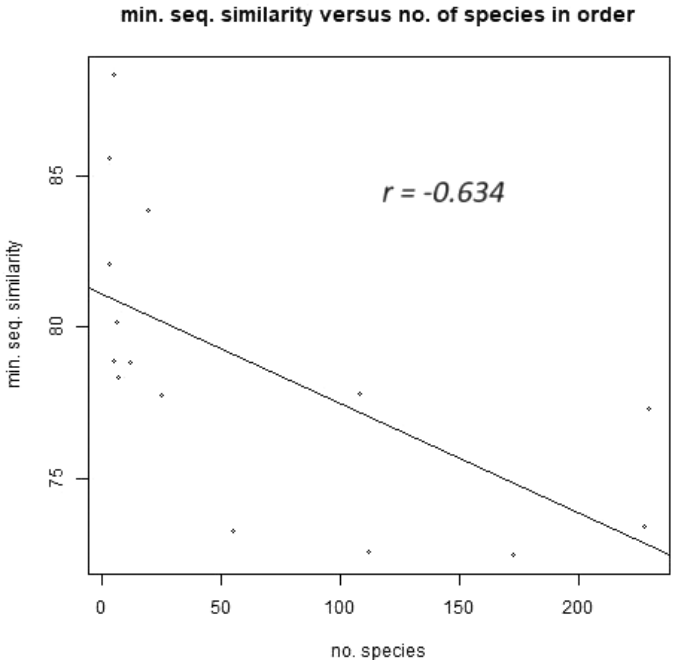

Figure 6. MSS according to the number of species in an order.

Table IV. The number of species with data in the Mitogenome Database from various tetrapod groups.

Group	Number of species	Number of genera
Mammalia	1005	491
Aves	663	371
Lepidosauria	311	166
Amphibia	216	96

cutoff. The second option is to list all species the user wants to examine, and the SS values will be downloaded from the database. The app then creates an SS value matrix for all selected species pairs (Supplementary Figure 2B). The matrix is represented as a heatmap (Supplementary Figure 2A).

A histogram will also show the SS values for all selected species (Supplementary Figure 2C). The Shapiro t-test is performed on the SS matrix, and the p-value will be displayed. P-values above 5% indicate that the SS value distribution is normal. Otherwise, the SS values depart from the normal distribution. This is because if the SS values are from two different kinds, there will be two types of SS values. First, SS values will describe sequence similarity between species of the same kind. These SS values will be high. In contrast, SS values between the two kinds will be low because of the dissimilarity between the two kinds. A bimodal histogram may depict such a dual distribution of SS values, indicating the presence of more than one kind in the data. The SS values depicted in the heat map are also shown at the bottom of the result screen (Supplementary Figure 2C).

MSS of various hybrid animals

Hybridization provides the strongest evidence that two animals belong to the same kind. Therefore, the parental species of several well-known hybrid species were analyzed to see how similar their mtDNA sequences are. This can also be used to measure the MSS value that separates from non-kind. However, this cutoff value is not absolute, since certain species that may be able to hybridize are unable to do so because of various biological factors. These include chromosome mismatching, or incompatibility of reproductive cells due to mutations in their surface receptors, making it impossible for them to fuse. Examples include the mule, which is a hybrid of *E. caballus* and *E. asinus*. An example of gamete incompatibility is when receptor proteins on the surface of egg cells in the

Table V. Mitochondrial SS between 18 mammalian hybrids.

Hybrid	Parental species 1	Parental species 2	Sequence similarity
cama	*Camelus dromedarius*	*Lama glama*	86.55
coywolf	*Canis latrans*	*Canis lupus*	95.61
geep	*Capra hircus*	*Ovis aries*	91.71
huarizo	*Vicugna pacos*	*Lama glama*	98.29
jagger	*Panthera onca*	*Panthera tigris*	91.42
jugapard	*Panthera onca*	*Panthera pardus*	92.52
kunga	*Equus hemionus*	*Equus asinus*	94.98
liger	*Panthera leo*	*Panthera tigris*	91.15
liguar	*Panthera leo*	*Panthera onca*	92.75
lipard	*Panthera leo*	*Panthera pardus*	94
mule	*Equus asinus*	*Equus caballus*	93.56
pizzly bear	*Ursus maritimus*	*Ursus arctos*	94.08
seal hybrid	*Phoca vitulina*	*Halichoerus grypus*	96.18
snow lion	*Panthera uncia*	*Panthera leo*	93.05
snow tiger	*Panthera uncia*	*Panthera tigris*	91
tigard	*Panthera tigris*	*Panthera pardus*	91.01
yakalo	*Bos grunniens*	*Bison bonasus*	94.01
yakalo	*Bos grunniens*	*Bison bison*	97.41

sea urchin genus *Echinometra* allow only conspecific sperm cells to bind via their binding protein (Palumbi, 2009). This means that the real SS cutoff may be lower due to these false negative cases.

These hybrids, their parental species, and their mtDNA SS are listed in Table V. The lowest SS is 86.6%, which corresponds to the hybrid 'cama,' a hybrid between the dromedary camel (*Camelus dromedarius*) and the llama (*Lama glama*).

Cumulative Distribution Function Plot

As a third approach, all mtDNA SS values calculated between all possible mammalian species pairs was visualized in a cumulative distribution function (CDF) plot (see Figure 7). To create the CDF plot, all SS values are ranked from lowest to highest. The x-axis shows the SS value, from 70% to 100%. The y-axis shows the proportion of all SS values that are less than or equal to the corresponding x-axis value (an SS value). No SS values are less than or equal to 70%, whereas all SS values are less than or equal to 100%. Since the SS value (the x-axis) gradually increases, so does the CDF value (the y-axis). The CDF curve strictly monotonically increases. It is interesting to note in Figure 7 that the CDF curve has two inflection points. There is a larger one between 75 and 80%, corresponding to a cumulative proportion of around 50%. Another, smaller inflection point is close to 85%, with a cumulative proportion of around 97.5%. This could indicate that the SS cutoff lies between this SS range (75–85%).

Conclusion

After all these considerations, can anything definite be said about an empirical cutoff value to differentiate species into baramins? As we saw from the example of the family Hominidae, families may contain more than one baramin. This would suggest that we look at the genus, which is one taxonomic level lower than the family. This way, the cutoff SS value would be 83.2%, the lowest MSS value among all genera in this study.

On the other hand, an individual kind may comprise more than one genus. This would mean we should take the lower bound of the MSS value range for families. This value is 75%. While some baramins correspond to the level of the order, such as Crocodylia (Cserhati, 2023c), this is generally not the case.

As a reference, the list of baraminology studies from Wood (2016) was analyzed, and the number of holobaramins corresponding to various taxonomical levels predicted by these studies, ranging from the genus to the superorder, was tallied. The result can be seen in Table VI. In this list, if a holobaramin was made up of more than one family, it was designated as a superfamily. If it contained more groups than a single order, it was considered a superorder. The list of holobaramins from the Wood study can be seen in Supplementary File 4. Overall, the low statistical support for the MSS value of 72.5% for order implies that for mammals, the kind is either at the family or genus level.

When actual animal hybrids are examined (which we can assume to be from the same kind), the MSS is 86.6%, which is above the cutoff limit range that was calculated by looking at groups of genera and families. Only 18

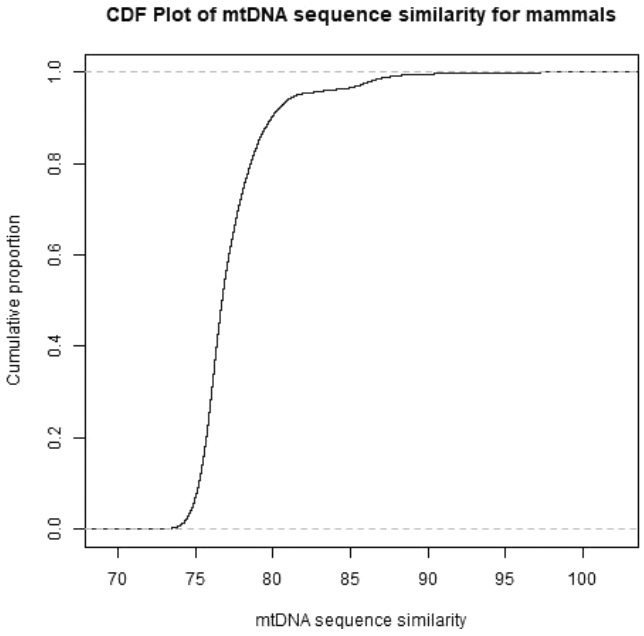

Figure 7. Cumulative Distribution Function Plot of all calculated sequence similarity values in the study.

Table VI. The number of holobaramins found in various baraminology studies listed in Wood (2016).

Taxonomic level	Number of holobaramins
Tribe	1
Genus	1
Subfamily	8
Family	46
Superfamily	9
Suborder	3
Order	1
Superorder	1

hybrids were examined, so with more species, this number might change. Based on the cumulative sequence similarity proportion plot, this range lies between 75–85%.

The result of this study indicates that the SS cutoff value for mtDNA studies is between 75–83.2% for mammals. While the MSS of various genera or families might be greater than these limits, these values were determined by taking the minimum MSS (the MMSS) value over all groups at a specific taxonomic level (genus and order). Any pair of species that are at least this similar based on their mitogenome sequence similarity may be considered to belong to the same baramin. Conversely, two species that are less similar than this may be considered to belong to separate baramins. This value could change with the inclusion of more data into the Mitogenome Database.

While this data is for mammals only, it would be interesting to study SS values for more animal groups, chloroplast SS from plants, and genome similarity for prokaryotes. Prokaryotic reproduction is very different from that of eukaryotes, so the results from this study might be valid only for eukaryotes, or even only for animals. Separate estimates might have to be made for non-animal groups.

The Mitogenome Database will serve as a useful tool for future molecular baraminology studies. The user can select from 6,336 tetrapod species for mitogenome studies from the app's Taxonomy Database. The user can also use the Heatmap Generation tool to run analysis on a selection of species. This includes all species from a selected genus or family. The user can also select a single species and retrieve all species pairs with an SS value above a selected limit. The user can then generate the SS-based heatmap, depict the resulting SS distribution on a heatmap, and depict the SS matrix with the selected species. The heatmap, the histogram, and the SS matrix can all be downloaded. Other baraminologists are encouraged to use this resource.

References

Cserhati, M. 2020. 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. 2021. Molecular and morphological analysis predict four bat baramins. *Journal of Creation* 35(1): 117–127.

Cserhati, M. 2023a. Chloroplast genome-based baraminology study of Liliales. *Creation Research Society Quarterly* 60(2): 84–96.

Cserhati, M. 2023b. Baraminic placement of *Homo heidelbergensis* based on molecular data. *Creation Research Society Quarterly* 59(3): 151–159.

Cserhati, M. 2023c. Baraminic analysis of Crocodylia based on mitochondrial DNA similarity. *Answers Research Journal* 16: 259–263.

Cserhati, M. 2023d. Molecular baraminology of marine and freshwater fish. In J.H. Whitmore (editor), *Proceedings of the Ninth International Conference on Creationism*, pp. 181–205. Cedarville University International Conference on Creationism, Cedarville, OH.

Cserhati, M. 2024a. Baraminology of Cucurbitaceae based on chloroplast genome analysis. *Answers Research Journal* 17: 613–619.

Cserhati, M. 2024b. Chloroplast genome-based molecular baraminology analysis of Proteales. *Creation Research Society Quarterly* 60(2): 84–96.

Cserhati, M., and Carter, R. 2020. Hierarchical clustering complicates baraminological analysis. *Journal of Creation* 34(3): 64–73.

Di Martino, O., McGehee, G., and Cserhati, M. 2024. Chloroplast Genome-Based Baraminology of the Order Myrtales. *Answers Research Journal* 17: 637–645.

Lightner, J.K., and Cserhati, M. 2019. The uniqueness of humans is clearly demonstrated by the Gene-Content Statistical Baraminology Method. *Creation Research Society Quarterly* 55(3): 132–141.

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

O'Micks, J. 2018. A preliminary cephalopod baraminology study based on the analysis of mitochondrial genomes and morphological characteristics. *Answers Research Journal* 11: 193–204.

Palumbi, S.R. 2009. Speciation and the evolution of gamete recognition genes: Pattern and process. *Heredity* 102(1): 66–76.

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

Wood, T. 2016. A list and bibliography of baraminology studies. *Journal of Creation Theology and Science Series B: Life Sciences* 6: 91–101.

Wood, T. 2021. Baraminology by cluster analysis: A response to Reeves. *Answers Research Journal* 14: 283–302.