

THE RED PANDA AND CSERHATI

This document contains the translation into English of the 18 blog entries about the red panda and Cserhati from my website
<https://creationismeweersproken.blogspot.com/> .

1	THE SCIENTIFIC SITUATION	2
2	LITERATURE SURVEY	10
3	A PHYLOGENY IS A MOBILE	14
4	TWO PAPERS	17
5	THE INTRODUCTION OF THE BMC GENOMICS ARTICLE	19
6	WHOLE GENOME K-MER SIGNATURE	30
7	WGKS IS NOT SUITABLE FOR PHYLOGENY	40
8	CLUSTERING	46
9	CLUSTERS ON WGKS DATA	49
10	PROCYONIDAE WGKS?	55
11	CLUSTERING MTDNA	58
12	PHYLOGENETIC TREES BASE ON mtDNA	64
13	THE BMC GENOMICS ARTICLE	73
14	THE INTRODUCTION OF THE CRSQ ARTICLE	76
15	DATA AND RESULTS IN THE CRSQ ARTICLE	80
16	CONCLUSIONS AND DISCUSSION IN THE CRSQ ARTICLE	97
17	BARAMIN	105
18	CSERHATI TRIES TO ANSWER	113

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-1-de.html>

THE RED PANDA AND CSERHATI (1): THE SCIENTIFIC SITUATION

On November 5, 2022, Jan van Meerten wrote a web post on his website 'Oorsprong' with the title: "**Scientist solves (creationist) biosystematic riddle of the red panda (*Ailurus fulgens*)**".

That web post made some strange claims:

"Incidentally, not only creationists have problems with the classification of the red panda. Naturalists are also not sure where to classify the beast. Does the beast belong to the bears (Ursidae), the raccoons (Procyonidae), or the skunks (Mephitidae)?"

"The most important conclusion we can draw from this study is that at the whole genome level, *A. fulgens* belongs to the clade of mustelids, not bears or skunks. (Cserhati)"

Jan van Meerten apparently did not take the trouble to look at wikipedia:

https://en.wikipedia.org/wiki/Red_panda . Wikipedia summarizes what science thinks about the red panda: the red panda does not belong to the bears, not to the raccoons, not to the skunks and not to the mustelids. The red panda belongs to itself: *Ailurus fulgens* is the only species of the family Ailuridae.

The Ailuridae family belongs to the superfamily Musteloidea, as do three other families: the skunks Mephitidae, the raccoon family Procyonidae, and the mustelid family Mustelidae. The only question is what the relations are within the superfamily Musteloidea: to what family or group or families is the family Ailuridae most closely related? Molecular evolutionary biology work over the last 20-25 years has identified three possibilities for the relationships of the families within the superfamily Musteloidea. The three possibilities are spelled out in wikipedia and shown in the diagram below.

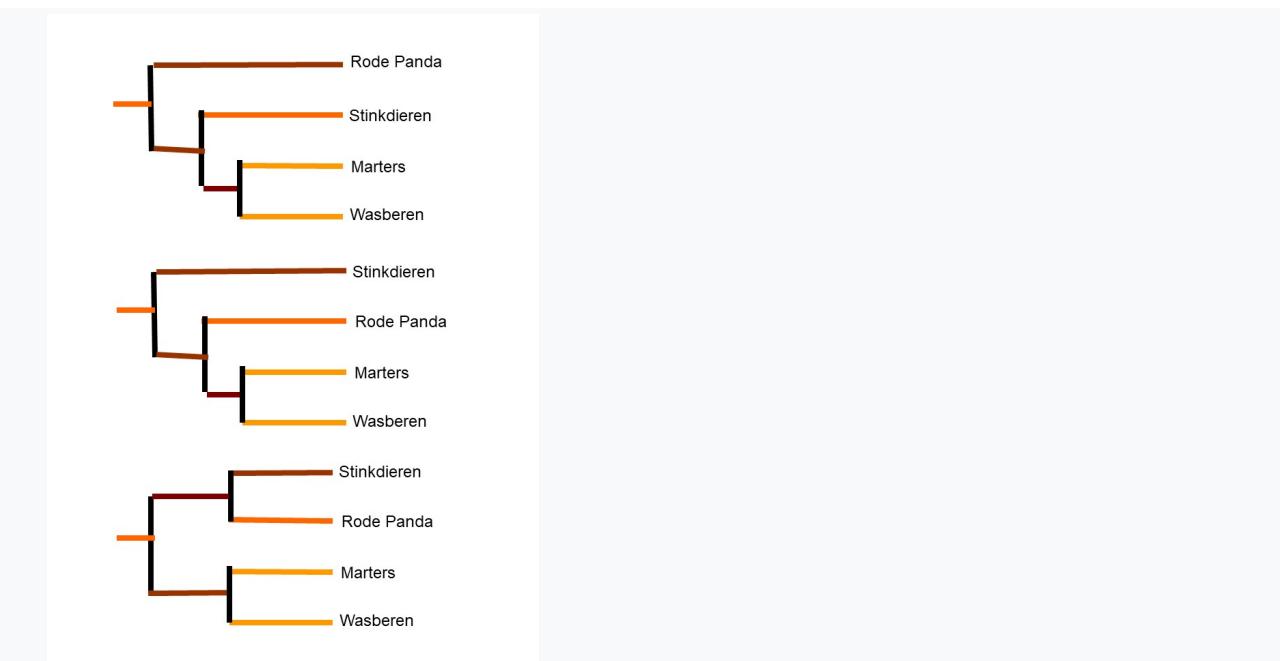


Figure 1 The three options in Wikipedia for the phylogeny of the Musteloidea. (stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

All scientific studies find that the mustelids and raccoons are more closely related to each other than either is to the red panda or the skunks - mustelids and raccoons are sister groups. But what is the relationship of mustelids + raccoons to the red panda and the skunks? Is the red panda the sister group of the skunks + mustelids+ raccoons (top), or of the mustelids + raccoons (middle) or of the skunks (bottom)?

The first possibility for the phylogeny of the Musteloidea can be found in Flynn *et al* (2005).

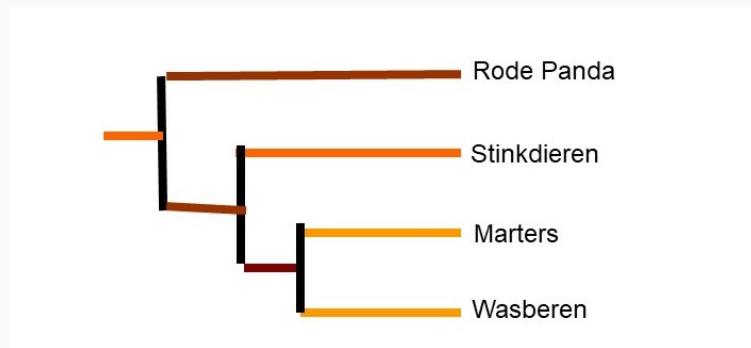


Figure 2 schematic representation of the classification of the Musteloidea as found in Flynn *et al* (2005). The marten families Mustelidae and the raccoon families Procyonidae are close relatives. The skunk family Mephitidae is as related to mustelids as it is to raccoons: the skunks family is the sister group to the Mustelidae + Procyonidae. The red panda, of the Ailuridae family, is as closely related to skunks as it is to martens and raccoons.(stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Flynn *et al* (2005) used a DNA sequence of 6243 bp from six genes to classify 76 species from the order Carnivora. The Carnivora consist of two large groups: the Feliformia and the Caniformia. The Caniformia fall into two groups: the canid family Canidae and a group called Arctoidea. The Arctoidea fall into two groups: the bear family Ursidae and a group with two more groups: the marine carnivorans Pinnipedia on the one hand and the superfamily Musteloidea on the other. The Musteloidea includes the red panda.

Flynn *et al* (2005) found most support for a first split within the Musteloidea between the red panda and the three remaining families (Figure 2, Figure 1 top), and a possibility for a first split between skunks + panda vs. mustelids + raccoons (Figure 1 bottom).

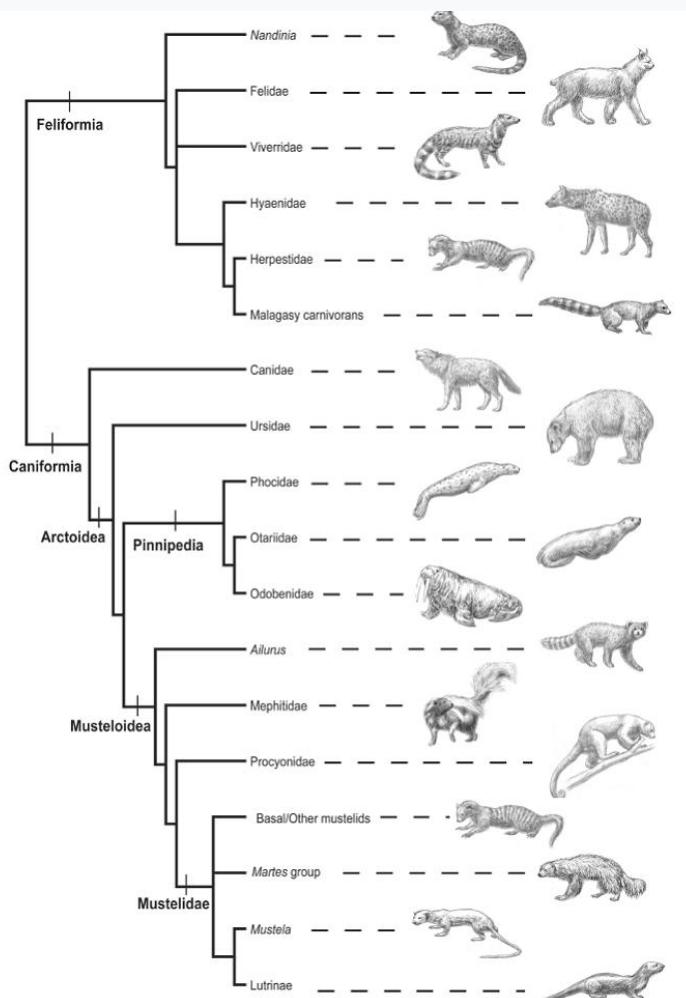


Figure 3 The phylogenetic tree of the Musteloidea based on molecular data of extant species. ; figure 5 from Flynn et al (2005)

The second possibility for the phylogeny of the Musteloidea can be found, among others, in Law *et al* (2018).

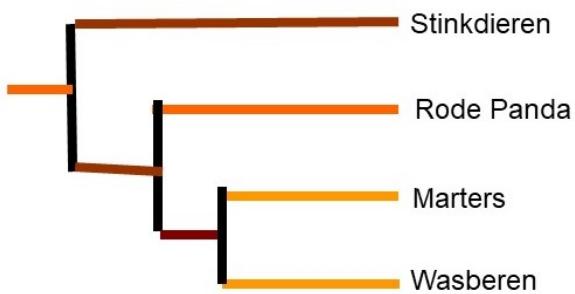


Figure 4 Schematic representation of the classification of the Musteloidea as found in Law *et al* (2018). The marten families Mustelidae and the raccoon families Procyonidae are close relatives. The red panda, of the Ailuridae family, is as closely related to martens as it is to raccoons. The skunk family Mephitidae is as closely related to the red panda as it is to the mustelids and raccoons.(stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Law *et al* (2018) were interested in the evolution of phenotype and life history in the Musteloidea. First Law *et al.* established a phylogenetic tree of the extant species of the superfamily Musteloidea based on all available molecular data and using Bayesian analysis. Then, Law *et al* combined this phylogenetic tree with available fossils. This yields a time-calibrated phylogenetic tree of Musteloidea, providing mean divergence times for the division of the families and the groups within the families.

The molecular phylogenetic tree of extant Musteloidea species gives as first split the divergence of skunks (Mephitidae) from the other Musteloidea. Figure 5 shows the phylogenetic tree of extant Musteloidea as found by Law *et al* (2018).

But when the fossils are included, the red panda and skunks emerges as sister group: as in figure 1 at the bottom. Figure 6 shows the phylogenetic tree of the living and fossil Musteloidea as found by Law *et al* (2018).

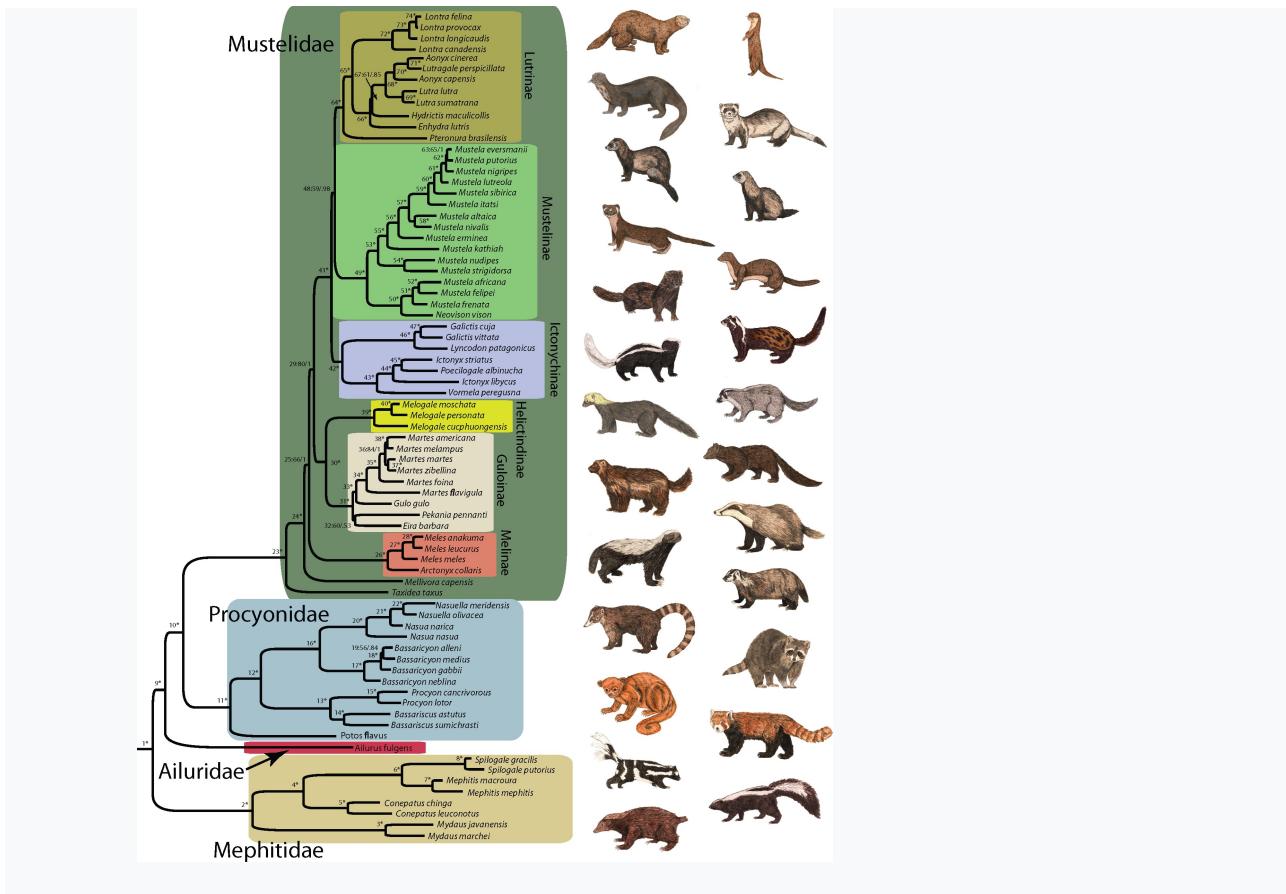


Figure 5 The phylogenetic tree of the Musteloidea based on molecular data from living species; figure 1 from Law et al (2018).

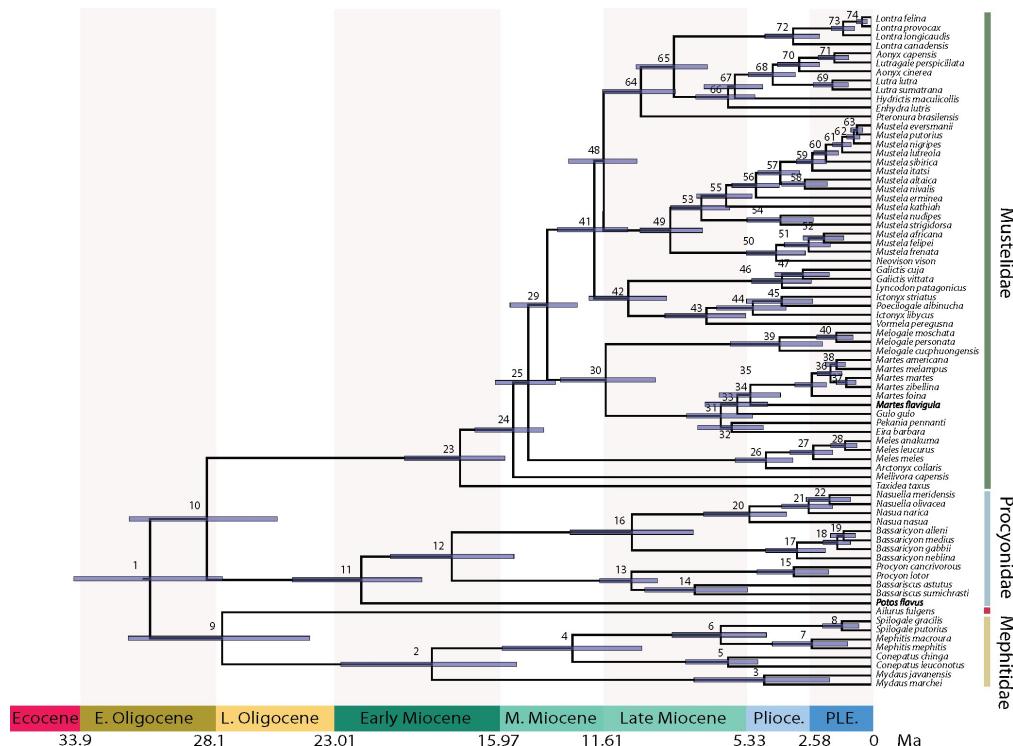


Figure 6 The phylogenetic tree of the Musteloidea based on molecular data from living species and known fossils.; figure 2 from Law et al (2018).

The third possibility for the phylogeny of the Musteloidea can be found in Hassanin *et al* (2021), and as a minor possibility in Flynn *et al* (2005) and Law *et al* (2018).

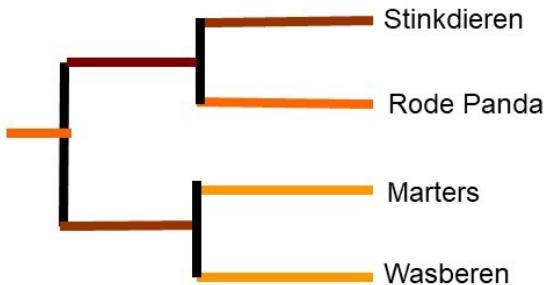


Figure 7 Scheme of the classification of the Musteloidea as found in Hassanin *et al* (2021). The marten families Mustelidae and the raccoon families Procyonidae are close relatives. The red panda, the family Ailuridae, and the skunk family Mephitidae are close relatives. The red panda and the skunks are as closely related to mustelids as they are to raccoons. (stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Hassanin *et al* (2021) performed the most comprehensive molecular study of the order Carnivora to date. Hassanin *et al* used the DNA sequence of the complete mtDNA of 220 species of Carnivora. They sequenced the mtDNA of some species themselves, while the mtDNA sequence of other species was already in the database. In total, they used DNA sequences from 2442 mitochondrial genomes, 220 species and 14892 base pairs per species.

Hassanin *et al* (2021) used Bayesian analysis to find the mtDNA phylogenetic tree of the Carnivora.

This largest molecular study of the phylogeny of the Musteloidea to date yields the red panda and the skunk family as sister groups. Previous studies had difficulty rejecting this possibility, but the large study by Hassanin *et al* (2021) is decisive: the possibility at the top of figure 1 has the best credentials.

All scientific studies indicate that the marten family Mustelidae and the raccoon family Procyonidae are sister groups. The red panda, family Ailuridae, and the skunk family Mephitidae as sister group have most support.

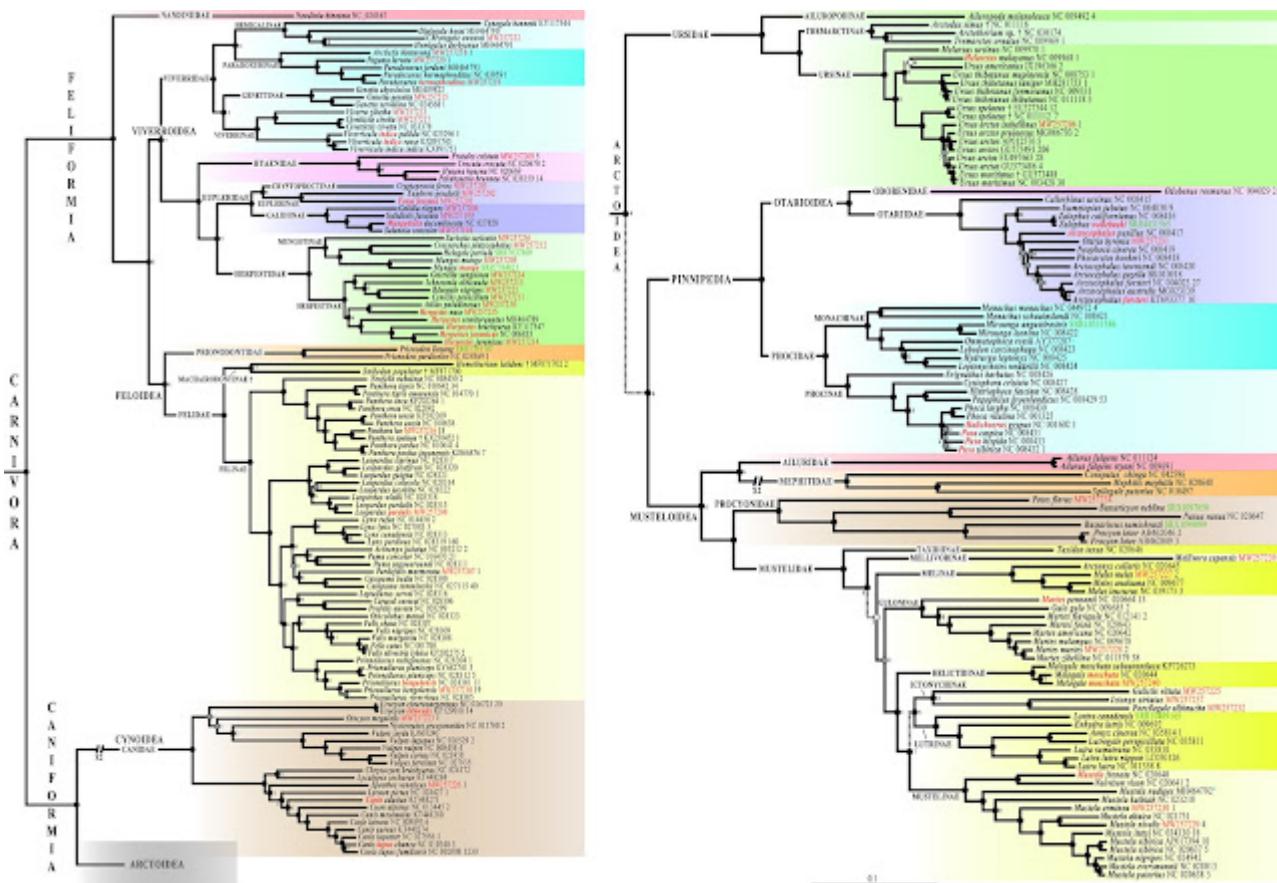


Figure 8 Phylogenetic tree of the Carnivora based on mtDNA. Figure 2 from Hassanin *et al* (2021).

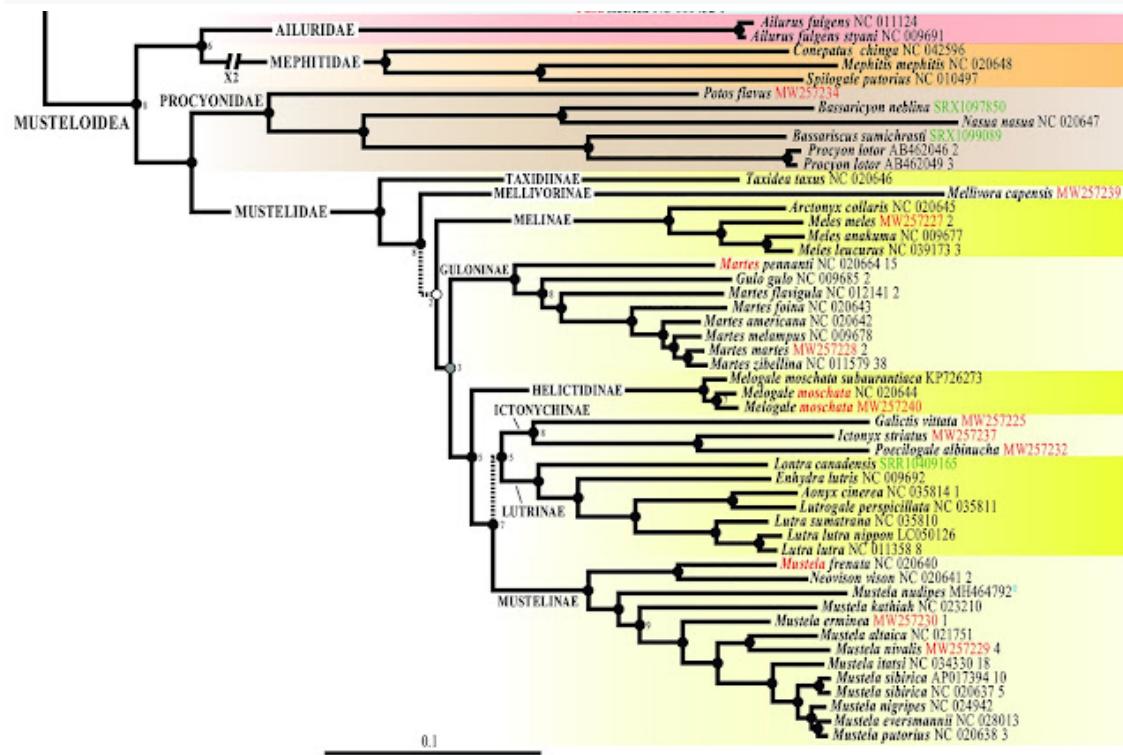


Figure 9 Phylogenetic tree of the Musteloidea according to their mtDNA. Detail of Figure 2 from Hassanin *et al* (2021).

Other studies do not give a different picture. This is the scientific state of affairs. The scientific state of affairs is in stark contrast to:

"Incidentally, not only creationists have problems with the classification of the red panda. Naturalists are also not sure where to classify the beast. Does the beast belong to the bears (Ursidae), the raccoons (Procyonidae), or the skunks (Mephitidae)?"

The red panda does not belong to the Ursidae bears, not to the Procyonidae raccoons, and not to the Mephitidae skunks. For evolutionary biologists, Jan van Meerten's question is 20 years out of date.

The scientific state of affairs is in complete contradiction with what Cserhati writes according to Jan van Meerten:

"The most important conclusion we can draw from this (Cserhati's) study is that at the whole genome level, A. fulgens belongs to the clade of mustelids, not bears or skunks."

The red panda is not a mustelid, it does not belong to the family Mustelidae, although it does belong to the superfamily Musteloidea.

<https://oorsprong.info/wetenschapper-lost-creationistisch-biosystematisch-raadsel-van-de-rode-panda-ailurus-fulgens-op/>

<https://evolutiebiologie.blogspot.com/2010/10/het-raadsel-van-de-rode-panda.html>

Flynn, J.J.; Finarelli, J. A.; Zehr, S.; Hsu, J. & Nedbal, M.A. (2005). Molecular phylogeny of the Carnivora (Mammalia): Assessing the impact of increased sampling on resolving enigmatic relationships. *Systematic Biology*. 54: 317–337.
doi:10.1080/10635150590923326.

Law, C.J.; Slater, G.J. & Mehta, R.S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. *Systematic Biology*. 67: 127–144. doi:10.1093/sysbio/syx047..

Hassanin, A.; Veron, G.; Ropiquet, A.; van Vuuren, B. J.; Lécu, A.; Goodman, S. M.; Haider, J.; Nguyen, T. T. (2021). Evolutionary history of Carnivora (Mammalia, Laurasiatheria) inferred from mitochondrial genomes. *PLOS ONE*. 16 (2): e0240770.

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-2-overzicht.html>

THE RED PANDA AND CSERHATI (2): LITERATURE SURVEY

In blog post 1 of this series it emerged that there are three scientific possibilities for the classification within the superfamily Musteloidea. The mustelids Mustelidae and the raccoon-like Procyonidae are always sister groups. The differences between the possibilities are in the placement of the red panda family Ailuridae and the skunk family Mephitidae

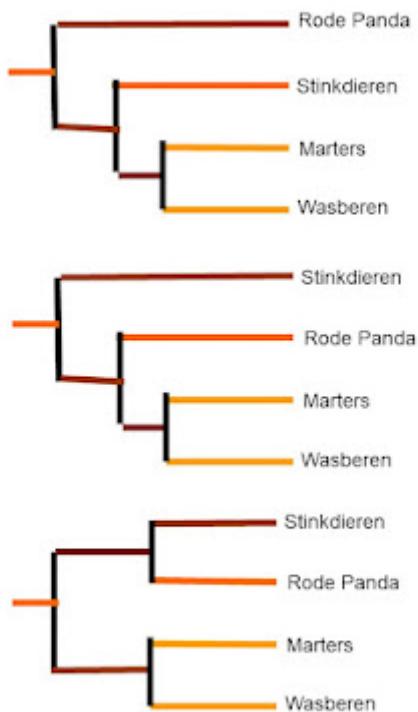


Figure 1 Three possibilities for the phylogenetic tree of the Musteloidea.(stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Here I present a survey of the literature, how often each of these possibilities has been found - without claiming to be complete.

The first possibility is that the first split in the Musteloidea is found between red panda and skunks + raccoons + mustelids.

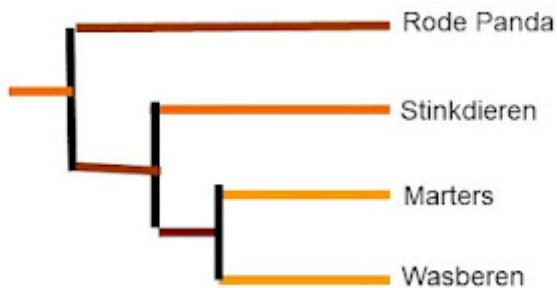


Figure 2 First split in the Musteloidea between red panda and skunks + raccoons + martens.
(stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Flynn, J.J., Nedbal, M.A., Dragoo, J.W., & Honeycutt, R.L. (2000) Whence the red panda? Mol Phylogenet Evol. 17:190–199.

Flynn, J.J.; Finarelli, J.A.; Zehr, S.; Hsu, J. & Nedbal, M.A. (2005). Molecular phylogeny of the Carnivora (Mammalia): Assessing the impact of increased sampling on resolving enigmatic relationships. Systematic Biology. 54: 317–337.

Fulton, T.L, Strobeck, C. (2007) Novel phylogeny of the raccoon family (Procyonidae: Carnivora) based on nuclear and mitochondrial DNA evidence. Molecular Phylogenetics and Evolution 43:1171–7.

The second possibility is that the first split in the Musteloidea is between skunks and red panda + raccoons + mustelids.

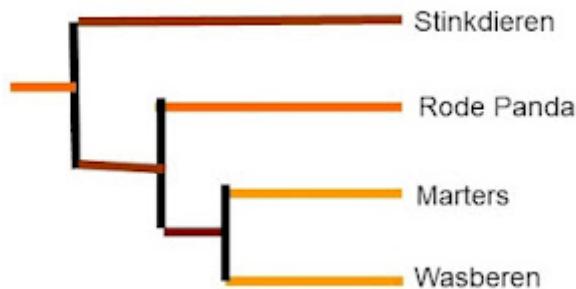


Figure 3 First split in the Musteloidea between skunks and red panda + raccoons + martens.
(stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Fulton, T.L, Strobeck, C. (2007) Novel phylogeny of the raccoon family (Procyonidae: Carnivora) based on nuclear and mitochondrial DNA evidence. *Molecular Phylogenetics and Evolution* 43:1171–7.

Sato, J.J., Wolsan, M., Minami, S., Hosoda, T., Sinaga, M.H., Hiyama, K., Yamaguchi, Y. & Suzuki, H. (2009) Deciphering and dating the red panda's ancestry and early adaptive radiation of Musteloidea. *Molecular Phylogenetics and Evolution* 53 (2009) 907–922

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

Yu, L., Luan, P.T., Jin, W, Ryder, O.A., Chemnick, L.G., Davis, H.A.& Zhang, Y.P. (2011) Phylogenetic Utility of Nuclear Introns in Interfamilial Relationships of Caniformia (Order Carnivora) *Systematic Biology* 60:175–187;

Law, C.J.; Slater, G.J. & Mehta, R.S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. *Systematic Biology*. 67: 127–144.

The third possibility is that the first split in the Musteloidea is between skunks + red panda and raccoons + mustelids. This possibility is often mentioned as a less substantiated possibility in articles that present evidence for one of the first two possibilities.

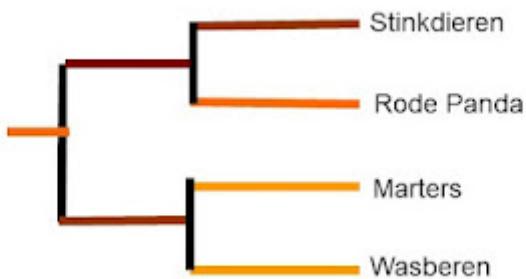


Figure 4 First split in the Musteloidea between skunks * red panda and raccoons + martens . (stinkdieren = skunks; rode panda = red panda; marters = marten family; wasberen = raccoon family)

Flynn, J.J., Nedbal, M.A., Dragoo, J.W, & Honeycutt, R.L. (2000) Whence the red panda? *Mol Phylogenet Evol.* 17:190–199.

Delisle, I. & Strobeck, C. (2005) A phylogeny of the Caniformia (order Carnivora) based on 12 complete protein-coding mitochondrial genes. *Molecular Phylogenetics and Evolution* 37: 192–201

Sato, J.J., Wolsan, M., Suzuki, H., Hosoda, T., Yamaguchi, Y., Hiyama, K., Kobayashi, M., Minami, S. (2006) Evidence from nuclear DNA sequences sheds light on the phylogenetic relationships of Pinnipedia: single origin with affinity to Musteloidea. *Zoological Science* 23:125–46.

Fulton, T.L, Strobeck, C. (2007) Novel phylogeny of the raccoon family (Procyonidae: Carnivora) based on nuclear and mitochondrial DNA evidence. *Molecular Phylogenetics and Evolution* 43:1171–1177.

Law, C.J.; Slater, G.J. & Mehta, R.S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. *Systematic Biology* 67: 127–144..

Jin, Z., Xu, H., Li, D. Xie, M., Zhang, M., Ni, Q., & Yao, Y. (2019) Complete mitochondrial genome of red panda (*Ailurus fulgens*) and its phylogenetic analysis. *Mitochondrial DNA part B* 4: 2339-2340

Hassanin, A.; Veron, G.; Ropiquet, A.; van Vuuren, B. J.; Lécu, A.; Goodman, S. M.; Haider, J.; Nguyen, T. T. (2021). Evolutionary history of Carnivora (Mammalia, Laurasiatheria) inferred from mitochondrial genomes. *PLOS ONE*. 16 (2): e0240770.

Hassanin A, Veron G, rA, Jansen van Vuuren B, Lécu A, Goodman SM, et al. (2021) Correction: Evolutionary history of Carnivora (Mammalia, Laurasiatheria) inferred from mitochondrial genomes. *PLoS ONE* 16(3): e0249387.

This third possibility for the classification of the superfamily Musteloidea, the red panda as family Ailuridae as a sister group of the family Mephitidae, now has most support. The study by Hassanin et al (2021) is the most comprehensive study of the phylogeny of the order Carnivora to date. Moreover, this possibility for classification was also found earlier, as a less substantiated possibility in studies that ended up with one of the other two possibilities.

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-3-een.html>

THE RED PANDA AND CSERHATI (3): A PHYLOGENY IS A MOBILE

This post may be superfluous, but anyway.

How to read a phylogeny? A phylogeny looks a bit like a number of bifurcating square rakes, with names on the right. Names on the right is the easiest layout. We could also put the names on the left, but that is a bit more difficult to read. Or we could put the names at the top, but then we might run into trouble with the page width.

Therefore, a set of left to right bifurcating branching rakes is the most convenient layout on paper.

Hang branching rakes from the ceiling and you have a mobile: a mobile can rotate round any vertical axis.



A phylogeny can rotate round any horizontal axis.

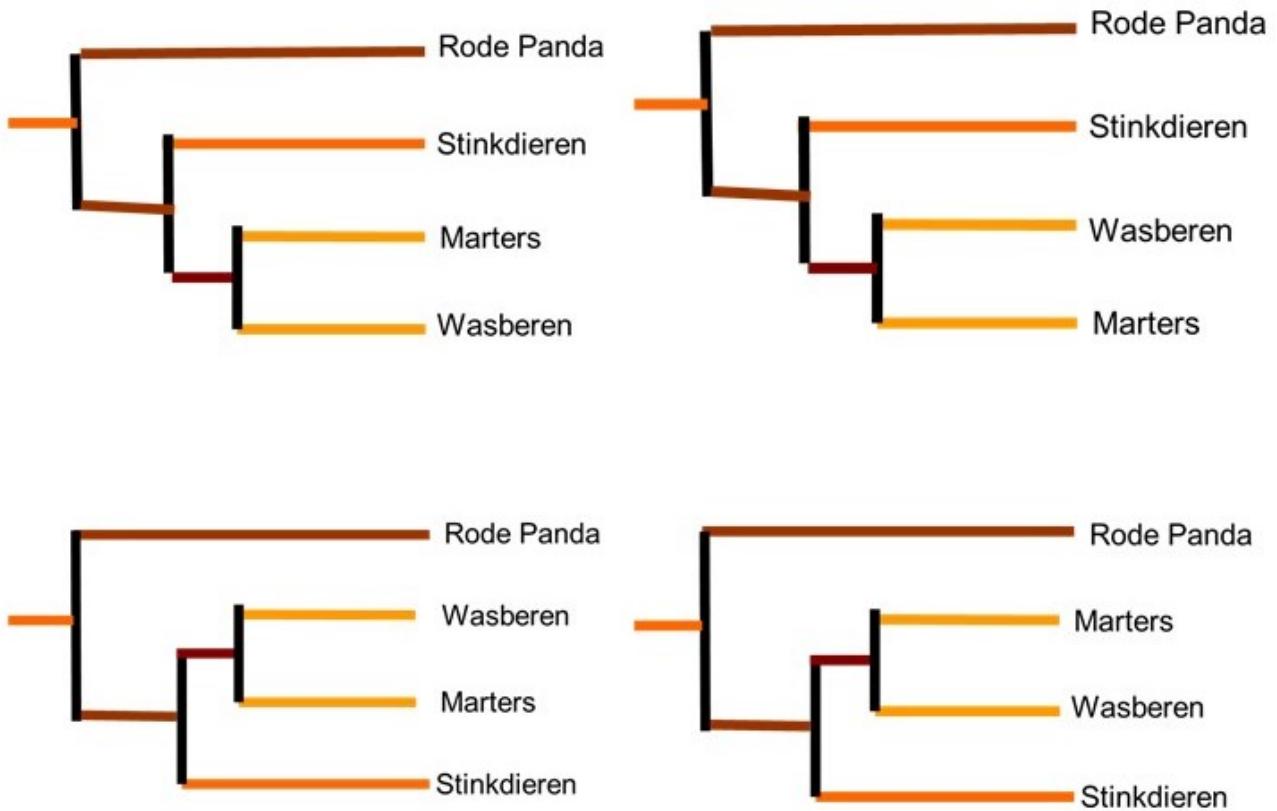


Figure 1. Identical phylogeny in four layouts. (Rode panda -= red panda; stinkdieren = skunks; marters = mustelids; wasberen = procyonids).

The figure gives the same phylogeny four times. The sequence of animal groups on the right is just the layout.

How to read such a phylogeny?

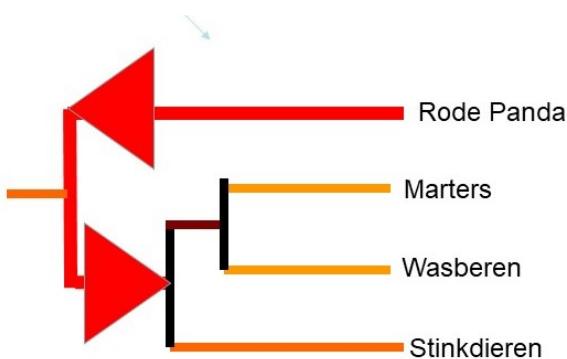


Figure 2 Road map

Pick any species.

Start going left along its line

Go to the left along that line until arriving at a vertical line.

Follow the vertical line until you find the first line branching on the right.

Go along that line to the right.

At the end of that line you'll find the sister group to your chosen species. The sister group might itself be a species, or it might be a group of species.

Warning: don't go with the lay-out! Here, declaring the marten family the sister group of the red panda would be a gross mistake!

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-4-twee.html>

THE RED PANDA AND CSERHATI (4): TWO PAPERS

On November 5, 2022, Jan van Meerten wrote a web post on his website Oorsprong with the title: "**Scientist solves (creationist) biosystematic riddle of the red panda (*Ailurus fulgens*)**". In his web post, Jan van Meerten referred to two articles, both by M. Cserhati.

One of the articles has been published in the scientific literature, in April 2021:

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228
<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-021-07531-3>

The other article is in a creationist journal, in Fall 2021:

Cserhati, M., 2021, Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis, Creation Research Society Quarterly 58 (2): 76-84
<https://www.creationresearch.org/classification-of-the-enigmatic-red-panda-ailurus-fulgens-based-on-molecular-baraminology-based-analysis>

These two articles both have a part that uses a WGKS method and a part that uses mtDNA for use in red panda classification. These parts are very similar between the articles.

The BMC genomics article is authored by Matyas Cserhati, the CSRQ article by Matthew Cserhati: a translation from Hungarian Mátyás to English. According to LinkedIn, Matthew/Matyas Cserhati holds a PhD in Bioinformatics from the University of Szeged in Hungary and was a full-time speaker for Creation Ministries International from February 2019 to May 2020.

Creation Ministries International details his involvement in creationism:
<https://creation.com/matthew-cserhati> and <https://creation.com/dr-matthew-cserhati-cv> citing quite a number of creationist articles by his hand.

The BMC Genomics article only lists a home address and email address, no work address. There is nothing in this article about funding: who paid for the work?

The CSRQ article gives no address at all, but acknowledges the funding:

"Acknowledgements

This paper is a part of Creation Research Society Grant #62 and part of the analysis was performed on the new CRS server based at Arizona Christian University"

It might be presumed this was also the funding for the BMC Genomics article.

Jan van Meerten (website 'Oorsprong') thinks so too:

*"As far as I'm concerned, this is **the** way to do creation research. First, obtain a fund to do a detailed study within the creation paradigm. Then publish the results in a standard naturalistic-scientific journal. Then further develop the results in a creation science journal."*

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-5-de.html>

THE RED PANDA AND CSERHATI (5): THE INTRODUCTION OF THE BMC GENOMICS ARTICLE

The article by Matyas Cserhati in BMC Genomics has the title: '**A tail of two pandas— whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*)'**

This suggests that the article will be about both the red panda and the giant panda. Cserhatis' abstract in BMC Genomics gives a different impression:

*Background: The red panda (*Ailurus fulgens*) is a riddle of morphology, making it hard to tell whether it is an ursid, a procyonid, a mustelid, or a member of its own family. Previous genetic studies have given quite contradictory results as to its phylogenetic placement.*

This tells us two things: the article is about the red panda *Ailurus fulgens*, rather than about the giant panda *Ailuropoda melanoleuca*, and presumably we are getting a genetic study to solve an outstanding morphological problem.

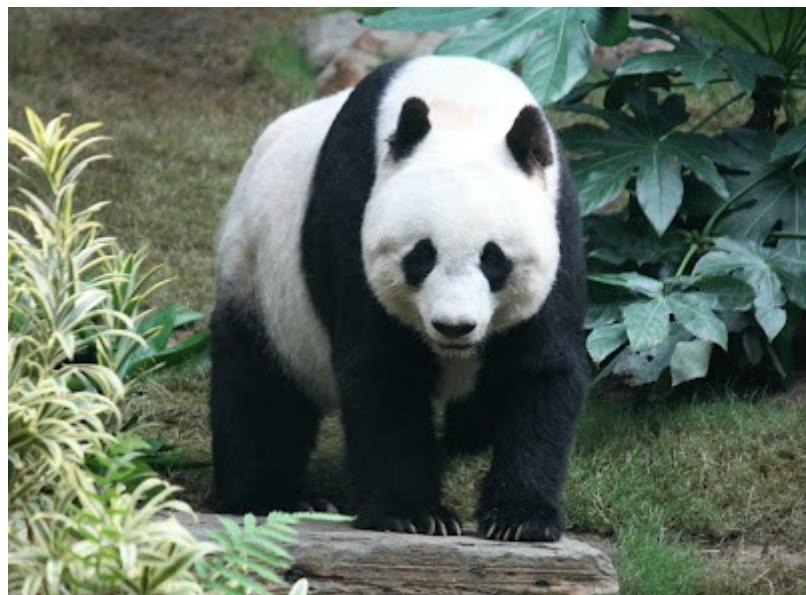


Figure 1 the giant panda *Ailuropoda melanoleuca*



Figure 2 the red panda *Ailurus fulgens*

The introduction of an article states more extensively than the abstract what the problem is and also what has been said about that problem before.

1

Cserhati begins by saying that the giant panda and the red panda have been considered related species. Cserhti writes: 'think' not 'thought'.

Some researchers think Ailurus fulgens is a relative of the giant panda (Ailuropoda melanoleuca) based on several physical characteristics. These include an almost exclusive diet of bamboo (both species eat meat on occasion), and have an enlarged radial sesamoid bone, which they use to process bamboo [1, 2].

What do these two articles [1] and [2] say about the relation between the two pandas?

Article [1], Flynn et al (2000), provides a historical overview of taxonomic work on both pandas. The giant panda belongs to the family of bears, Ursidae, as suggested on the basis of morphology by Davis (1964), on the basis of immunology by Sarich (1973), and on the basis of proteins, immunology, chromosomes, and a rough measure of DNA similarity by O'Brien et al (1985). In 1985, the placement of the giant panda with bears was considered a great success for molecular methods. After 1985, the placement of the giant panda is no longer a topic of discussion: the giant panda is a bear. .

What about the red panda?. Flynn et al (2000) mention a number of possibilities for the classification of the red panda. Their abstract says:

...whether it (the red panda) should be placed with the bears (ursids),,raccoons (procyonids), musteloids or as a monotypic lineage of uncertain phylogenetic affinities

Some of these proposals were already out of circulation in 2000: the red panda as a sister group to the giant panda (1943, 1982) or the red panda as a relative of the bears ((1973, 1989, 1993, 1994). Flynn et al (2000) do not elaborate on this: they only give a historical overview, and in 2000 the relationship of giant panda and red panda was from a molecular point of view not a possibility.

Can a relationship between red panda and giant panda be found in article [2], Hu et al (2017)? The article by Hu et al (2017) is about convergence between the red panda and the giant panda; in other words, about their similarity in their diet of bamboo and their similarity in their false thumb despite being unrelated. Convergence means similarity but never relatedness. Hu et al (2017) write:

The giant panda belongs to the family Ursidae, whereas the red panda belongs to the family Ailuridae within the superfamily Musteloidea.

Hu et al (2017) state in so many words that the red panda and the giant panda are not related; and also where the red panda has been placed since at least 2000, with the Musteloidea.

Cserhati is not very accurate in citing these articles [1] and [2]. Neither gives the red panda as a relative of the giant panda.

2

Cserhati proceeds:

According to other opinions, *A. fulgens* has been classified as a member of the family Procyonidae (raccoons).

Cserhati does not give a reference here, but in Flynn et al (2000) we come across a number of references, starting with Geoffroy-Saint-Hilaire and Cuvier (1825) with "closely resembling a raccoon (procyonid)". Those references in Flynn et al (2000) show that raccoon and red panda are more often referred to as sister groups than that the red panda would be within the raccoon family Procyonidae. Only Slattery et al (1995) arrive at the placement of the red panda within the raccoon family, in two of their three analyzes of the same material. In later studies, the red panda is no longer found to belong to the raccoon family. Flynn et al (2000) cite earlier work by Flynn & Nedball (1998): they placed the red panda in the superfamily Musteloidea, but not in the

raccoon family Procyonidae. Flynn et al (2000) also end up in the superfamily Musteloidea with placement of the red panda in the family Ailurida

3

It would have been logical if Cserhati after citing Flynn et al (2000) would have proceeded with the results of Flynn et al (2000). Cserhati however proceeds in a different direction:

*For example, Peng et al. classify *A. fulgens* either as a mustelid, placing them next to the American marten (*Martes americana*), or as a mephitid, next to the striped skunk (*Mephitis mephitis*). This was based on the analysis of 13 concatenated mitochondrial proteins, based on neighbor-joining (NJ) and maximum likelihood (ML) phylogenetic methods, respectively [7].*

Peng et al (2007) present a phylogenetic analyses of the giant panda, against the background of quite a number of more or less related species. The superfamily Musteloidea is represented by one species per family: the red panda for itself as family Ailuridae, the raccoon for the Procyonidae, the striped skunk for the Mephitidae and the American marten for the Mustelidae. Peng gives two analyses.

1 The NJ method:

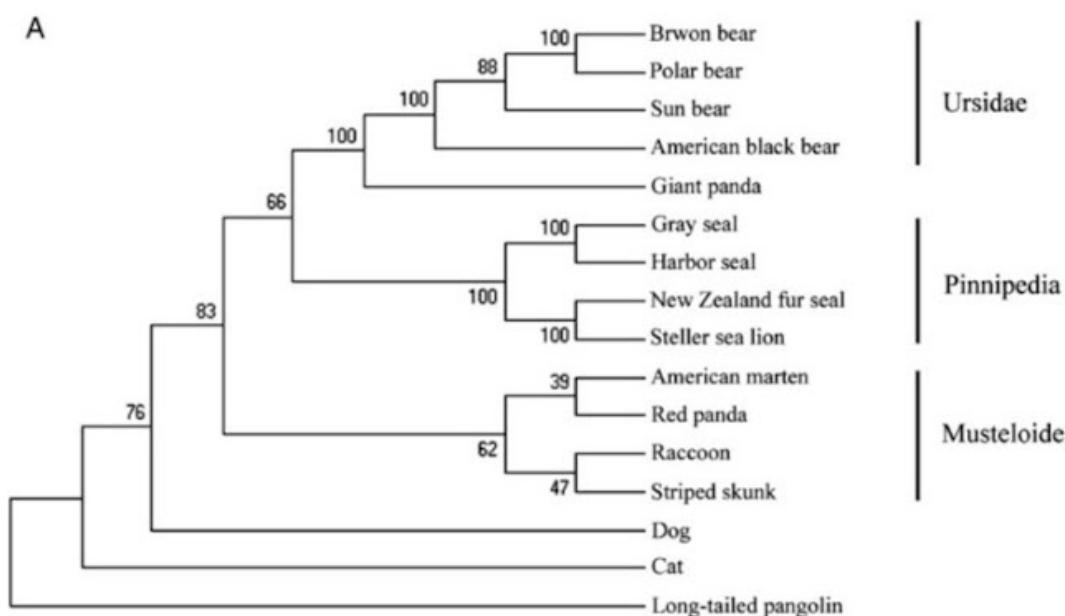


Figure 3. Figure 2A from Peng et al (2007) : The Ailuridae (represented by the red panda) and the Mustelidae (represented by the American marten) as sistergroups.

The NJ analysis places the red panda as the sistergroup of the American marten; that is, the Ailuridae as the sistergroup of the Mustelidae. The American marten is the only species used from the family Mustelidae; it is therefore impossible to conclude that the red panda belongs to the family Mustelidae: “*classify A. fulgens ... as a mustelid*” is an erroneous interpretation.

2 The ML method:

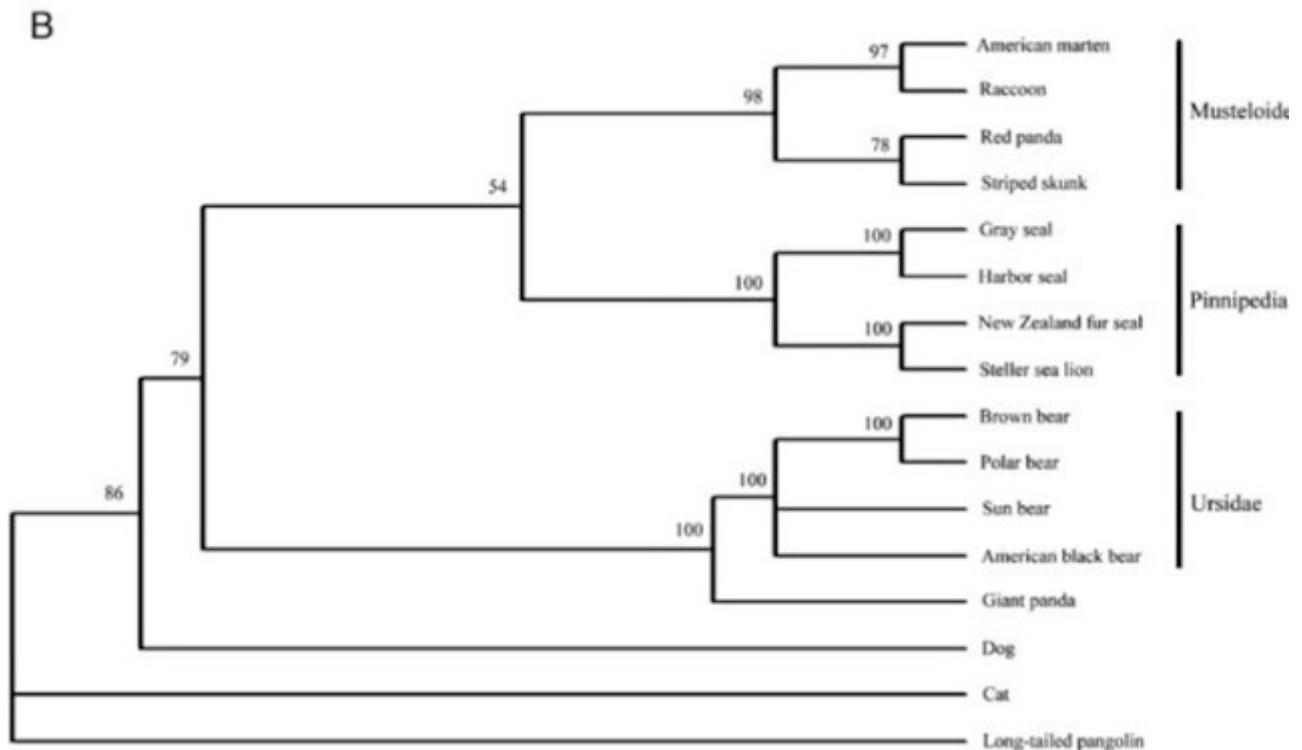


Figure 4. Figure 2B from Peng et al (2007) The Ailuridae (represented by the red panda) and the Mephitidae (represented by the striped skunk) as sistergroups.

The ML analysis places the red panda as the sistergroup of the striped skunk; that is, the Ailuridae as the sistergroup of the Mephitidae . The striped skunk is the only species used from the family Mephitidae; it is therefore impossible to conclude that the red panda belongs tot het family Mephitidae: “*classify A. fulgens ... as a mephitid*” is an erroneous interpretation. Both interpretations by Cserhati of the results of Peng et al (2007) show faulty reasoning: a basic lack of understanding of phylogenetic trees.

Cserhati proceeds with:

Flynn et al. also found that A. fulgens is neither an ursid, nor a procyonid, nor a mephitid, but a mustelid [1].

Cserhati now refers to the conclusions of Flynn et al (2000), the article Cserhati borrowed his historical description from. However, Flynn et al (2000) give their conclusion in their abstract as:

Combined phylogenetic analyses reject the hypotheses that the red panda is most closely related to the bears (ursids) or to the raccoons (procyonids). Rather, evidence from nucleotide sequences strongly support placement of the red panda within a broad Musteloidea (*sensu lato*) clade, including three major lineages (the red panda, the skunks [mephitids], and a clearly monophyletic clade of procyonids plus mustelids [Musteloidea *sensu stricto*, excluding skunks])

Flynn et al (2000) place the red panda as family Ailuridae in the superfamily Musteloidea; the red panda is a musteloid. Cserhati pretends that Flynn et al place the red panda in the family Mustelidae; if so, the red panda would be a mustelid.

Cserhati does not know or does not understand that musteloid and mustelid have two very different meanings. A musteloid belongs to the superfamily Musteloidea. A mustelid belongs to the family Mustelidae.

Pretending Flynn et al (2000) called the red panda a mustelid demonstrates great ignorance of taxonomy on the part of Cserhati.

Cserhati proceeds with:

Yu and Zhang studied introns 4 and 7 from the nuclear gene β -fibrinogen (FGB) as well as the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) in 17 species from the order Carnivora. In their results these researchers found that A. fulgens is most closely related to procyonids based on analysis of intron 4 of the FGB gene. But when intron 7 was analyzed, it clustered towards ursids. Classification based on the ND2 gene A. fulgens clustered with mustelids, but these results had poor bootstrapping support. When the two introns were combined with analysis of the genes IRBP and TTR, A. fulgens was closest to mustelids [9].

Cserhati presents the results of Yu & Zhang (2006) per intron.

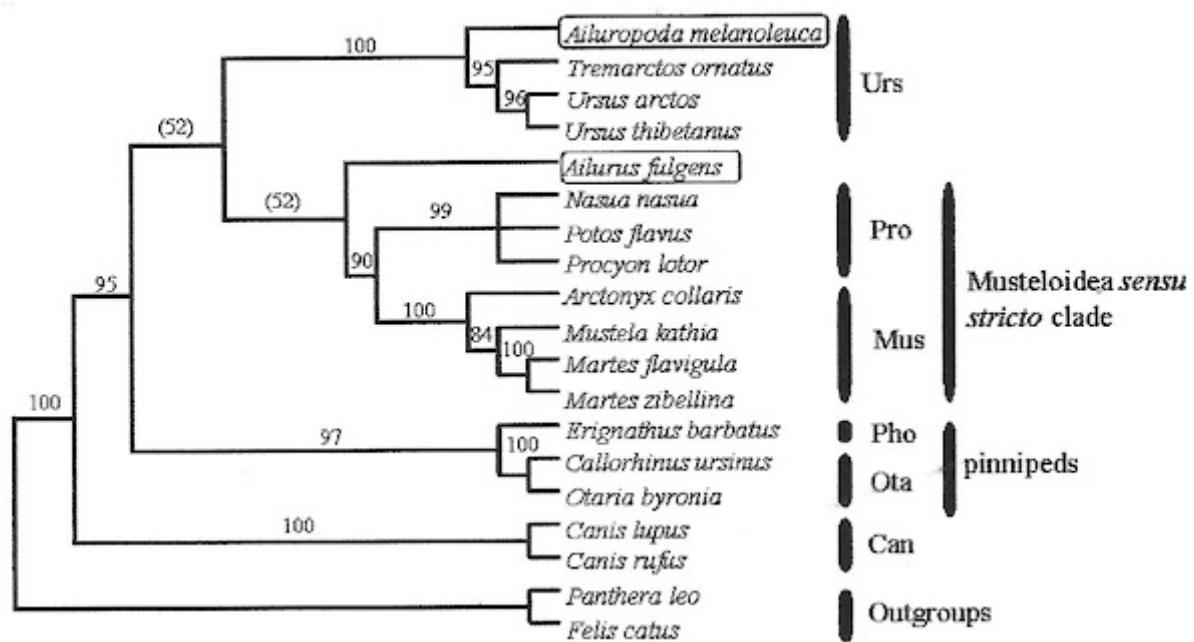


Figure 5 Figure 1 Yu & Zhang (2006): results intron 4

Yu & Zhang's figure 1 shows a phylogenetic tree based on nuclear b-fibrinogen intron 4; the red panda proves the sistergroup of the Procyonidae + Mustelidae (in the figure Musteloidea *sensu stricto*). The layout positions the red panda next to the Procyonidae, but that is only the layout of the tree.

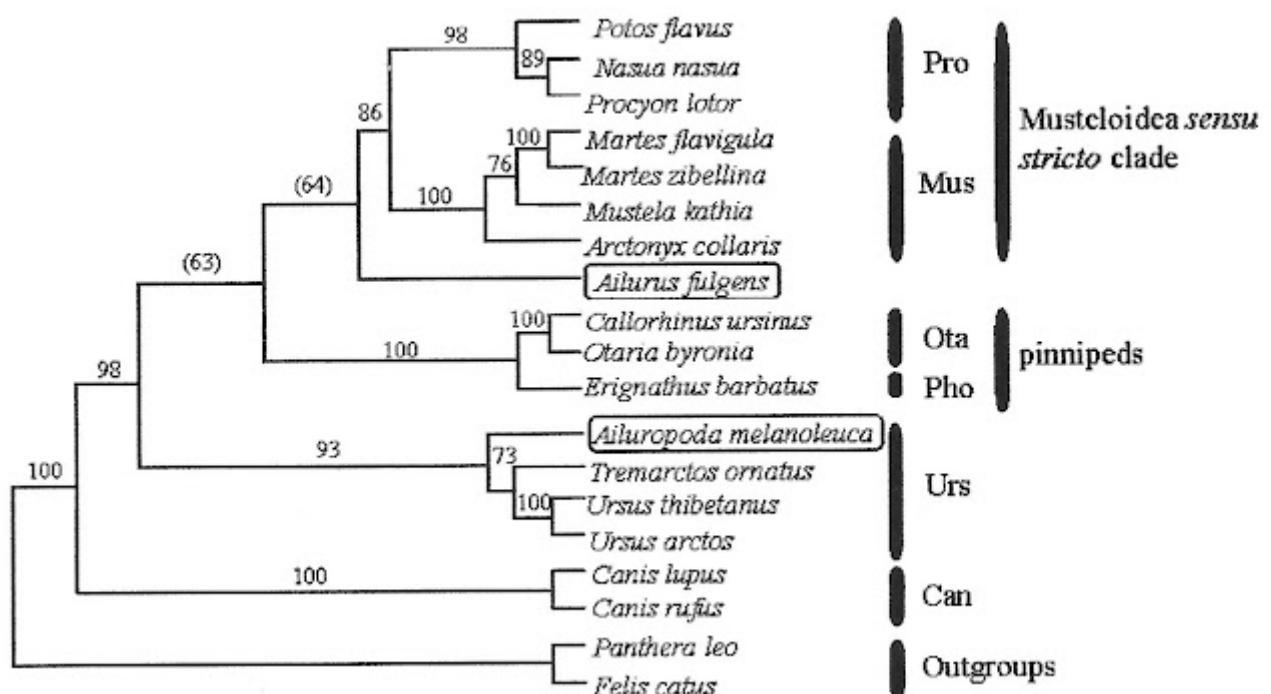


Figure 6. Figure 3 Yu & Zhang (2006): results ND2 gen

In Yu & Zhang's figure 3, het ND2 gen, the red panda is again the sistergroup of the Procyonidae + Mustelidae (Musteloidea *sensu stricto*) in the phylogenetic tree. The layout positions the red panda next to the Mustelidae, but that is only the layout of the tree.

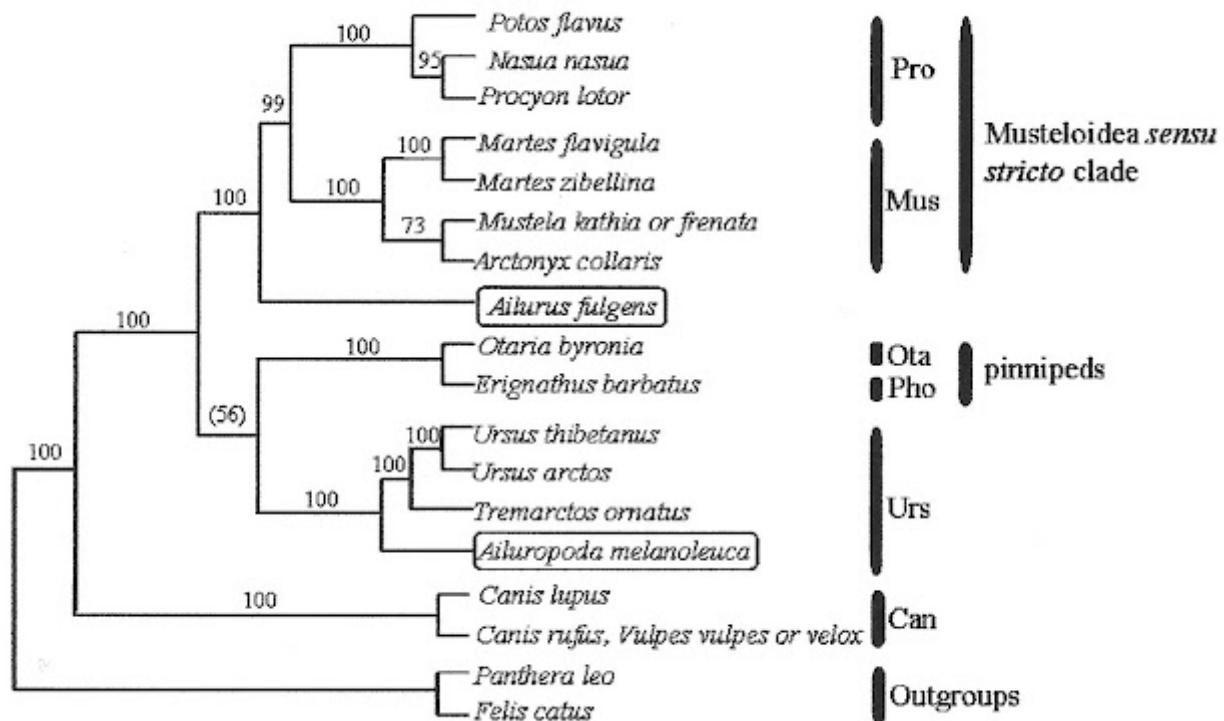


Figure 7. Figure 4 Yu & Zhang (2006): combined analysis

In Yu & Zhang's figure 4, the analysis of the combined IRBP, TTR, b-fibronogen intron 4 and 7 sequence data), the red panda is again the sistergroup of the Procyonidae + Mustelidae (Musteloidea *sensu stricto*) in the phylogenetic tree. The layout positions the red panda next to the Mustelidae, but that is only the layout of the tree.

Cserhati three times confuses the lay-out of the phylogenetic tree with the results of the phylogenetic tree. That is a pretty elementary error.

6

Cserhati proceeds with:

Sato et al. , and found that A. fulgens clusters together with procyonids and mustelids, and not with mephitids (skunks and stink badgers) [10].

Sato et al (2009) found that the red panda is the sister group of Musteloidea *sensu stricto*, i.e. of the raccoon family Procyonidae and the marten family Mustelidae together. Sato et al (2009) did not use clustering, but Bayesian analysis.

7

Cserhati proceeds with:

Intron analysis is useful, since these sequences are not under selection pressure. An analysis of 22 Kbp of nuclear intron sequences from 16 carnivore species groups A. fulgens with Musteloidea sensu stricto (Mustelidae+ Procyonidae) to the exclusion of mephitids [Yu et al 2011]. These results, however, contradict results coming from mtDNA analyses [DeLisle et al 2005].

Compare the text in Yu et al (2011):

Phylogenetic analyses of the more than 22 kb data set of noncoding intron DNA provided unambiguously strong support for the grouping of Musteloidea sensu stricto and Ailuridae to the exclusion of Mephitidae. ... This result is in contradiction to the mt studies (Ledje and Arnason 1996a, b; Delisle and Strobeck 2005; Arnason et al. 2007), but in agreement with the nuclear studies (Fulton and Strobeck 2006; Sato et al. 2009).

Cserhati almost literally copied Yu et al (2011), but ignores the message.

Summary:

This introduction was written by someone without any background in taxonomy and phylogeny. As a result, there are gross errors in taxonomy and phylogeny. Cserhati misrepresents the results of Flynn et al (2000), Peng et al (2007), Yu and Zhang (2006). This shows elementary unfamiliarity with taxonomy and with classifying species based on molecular data.

Adopting the abstract and historical part of the introduction by Flynn et al (2000) as a problem statement in 2021 shows no familiarity with the subject of 'red panda' or the subject of 'phylogeny'. Cserhati has a very limited grasp of the literature. The cited articles on phylogeny are at least 10 years old - newer ones are missing. Frequently cited articles such as Flynn et al (2005) with 253 citations, Eizirik et al (2010) with 142 citations or Law et al (2018) with 42 citations are missing from the bibliography.

Flynn et al (2000), Yu and Zhang (2006), Sato et al (2009) and Yu et al (2011) list the families Procyonidae and Mustelidae as sister groups of each other,

together called the Musteloidea *sensu stricto*. In addition, all the articles mentioned give the Musteloidea *sensu lato* as a superfamily with the four families Ailuridae, Mephitidae, Procyonidae and Mustelidae. For Cserhati, too, the question should therefore have been what the order of the Ailuridae, the Mephitidae and the Musteloidea *sensu stricto* is within the Musteloidea *sensu lato*.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

Davis, D. D. (1964). The giant panda: A morphological study of evolutionary mechanisms. Fieldiana Zool. Mem. 3: 1–339.

Sarich, V. (1973) The Giant Panda is a Bear. Nature 245: 218–220.

O'Brien, S., Nash, W., Wildt, D. et al. A molecular solution to the riddle of the giant panda's phylogeny. Nature 317, 140–144 (1985).

Peng R, Zeng B, Meng X, Yue B, Zhang Z, Zou F. The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*). Gene. 2017;397:76–83.

Yu, L., Zhang, YP. Phylogeny of the caniform carnivora: evidence from multiple genes. Genetica 127, 65–79 (2006).

Sato JJ, Wolsan M, Minami S, Hosoda T, Sinaga MH, Hiyama K. Deciphering and dating the red panda's ancestry and early adaptive radiation of Musteloidea. Mol Phylogen Evol. 2009;53(3):907–22.

Yu L, Luan PT, Jin W, Ryder OA, Chemnick LG, Davis HA, Zhang YP. Phylogenetic utility of nuclear introns in interfamilial relationships of Caniformia (order Carnivora). Syst Biol. 2011;60(2):175–87.

Delisle I, & Strobeck C. A phylogeny of the Caniformia (order Carnivora) based on 12 complete protein-coding mitochondrial genes. Mol Phylogen Evol. 2005;37(1):192–201

Flynn, J. J.; Finarelli, J. A.; Zehr, S.; Hsu, J. & Nedbal, M. A. (2005). Molecular phylogeny of the Carnivora (Mammalia): Assessing the impact of increased sampling on resolving enigmatic relationships. Systematic Biology. 54: 317–337.

Law, C. J.; Slater, G. J. & Mehta, R. S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using

Molecular and Fossil-Based Methods. Systematic Biology. 67: 127–144.

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

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-6-whole.html>

THE RED PANDA AND CSERHATI (6): WHOLE GENOME K-MER SIGNATURE

Cserhati indicates in the BMC Genomics article that he doubts the classification of the red panda on the basis of morphology. Since he has talked more about DNA sequences than about morphology, he will mean that he (also) finds the classification based on DNA subject to doubt. All the classifications based on DNA that he has mentioned are based on relatively little DNA (for 2021 standards). For example, Flynn et al (2000) used the DNA sequence of four genes. That was the year 2000, of course, and then no more DNA sequences were available

Cserhati prefers to use the whole genome for red panda classification: a whole genome study. That is increasingly what is happening. For example, De Ferran et (2022) did not search for orthologous genes in their eleven species of otters, but used genome fragments as found during genome sequencing for DNA comparison.

Cserhati's preference for using the entire genome is therefore perfectly understandable. The method used by Cserhati to characterize the whole genome is *Whole Genome K-mer Signature*, abbreviated as WGKS.

There are two questions: what is WGKS? And how useful is WGKS

First: what is WGKS? This is addressed in this post. How useful WGKS is for species classification will be discussed in the next installment.

In the Methods section of the BMC Genomics article, Cserhati writes:

The WGKS algorithm that was used in the analysis is an alignment-free k-mer sequence comparison method. These methods involve the statistical comparison of k-mers between species.

A k-mer is a segment of DNA k bp long,

The k-mer signature is simply a list of all k-mers ordered in lexicographical order from AA ... A to TT ... T, together with their score values. For a given value k, there are 4^k possible k-mers. Thus, the k-mer signature also corresponds to a vector of 4^k numbers. Since octamers were analyzed, this corresponds to 65,536 possible octamers.

Cserhati says: count all k-mers eight bases long - octamers - , and then find their scores. I'll give an example in two parts: counting octamers and finding scores.

1 Counting octamers and the correlation between octamer numbers

DNA has four bases: ACGT. An octamer, a DNA sequence 8 base pairs long, can show all possible sequences from AAAAAAAA to TTTTTTTT. Four possibilities for place 1, four possibilities for place 2, and so on. That means $4^8 = 65536$ possibilities. A computer walks along the genome, and reads sequentially which sequence of 8 basepairs is found.

In a DNA sequence;

gagtgggcagcactccaaataccgttaagctggagcctcggt

the consecutive octamers are:

from base 1: gagtgggc

from base 2: agtgggca

from base 3: gtgggcag

and so on. The computer counts the number of times an sequence of 8 bases occurs. In this example of a short DNA sequence, each sequence of 8 bases occurs once. The count defines the k-mer signature; with 8 bases it is called the octamer signature.

In DNA from two related species one would expect to find approximately the same distribution of 8-base sequences, octamers.

However, species differ not only in important DNA but also in unimportant DNA. For example, in the length of a repeat 'ac' - acacacacac or acacacacacacacacacaca, or a difference in the number of LINE1 elements. For example, a house mouse has hundreds of LINE1 elements (only a few of which are active as transposons (Jachowitz et al 2017)), and another species of mouse could have thousands of LINE1 elements. Such a repetition of the same

sequence as a plurality of LINE1 elements or a plurality of simple repeats will have a major influence on the distribution of octamers in a WGKS

Example of octamer counts:

A long example follows here, and I constructed it to show the influence of repetitive DNA.

Consider the following five sequences. In the first two sequences, a long or short 'ac' repeat is inserted relative to the third sequence. The fourth sequence differs from the first sequence in the part after the long 'ac' repeat. The fifth sequence has the shorter 'ac' repeat, and is otherwise the same as the fourth sequence.

>seq 1

>seq 2

gagtgggcagcaacacacacacacacactccaaataccgttaagctggaggcctcggt

>seq_3

gagtgggcagcactccaaataccgttaagctggagcctcggt

>seq_4

gagtgccggcaggcaacacacacacacacacacacacacacactttctggtccccacagact
cagagaga

>seq_5

gagtgggcagcaacacacacacacactttctggccccacagactcagagaga

Inserting points to bring out the 'ac' repeat:

>seq_1

>seq_2

gagtgggcagca...acacacacacacacac.....ctccaaataccgtta
agctggagcctcggt

>seq_3

gagtgggcagca.....ctccaaataccgtta
agctggagcctcggt

>seq_4

gagtgggcagca...acacacacacacacacacacacacacacacac...tcttctggtcccc
cagactcagagaga

>seq_5

gagtgggcagca...acacacacacacacacacacacacacacac...tcttctggtcccc
cagactcagagaga

A simple program, CLUSTAL, alignes the sequences, and provides two types of trees:

CLUSTAL O(1.2.4) multiple sequence alignment

seq_4	gagtgggcagcaacacacacacacacacacacacacactttctggtccc--	58
seq_5	Gagtgggcagcaacacacaca-----cacacacactttctggtccc--	42
seq_1	gagtgggcagcaacacacacacacacacacacacacacacacacaccc---aaataccgt	57
seq_2	gagtgggcagcaacacacac-----acacacacaccc---aaataccgt	41
seq_3	gagtgggcagca-----ctcc---aaataccgt	25

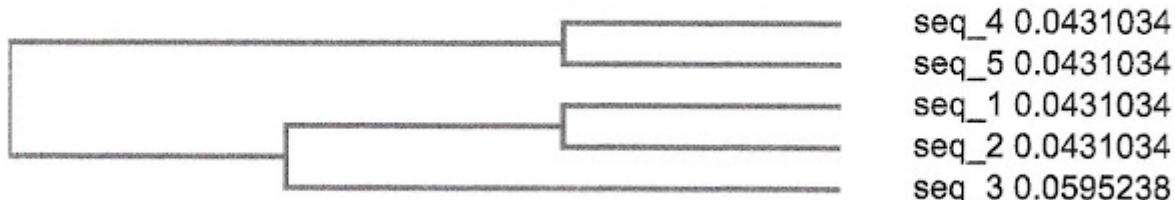
***** * * *

seq_4	-----cacagactcagagaga	74
seq_5	-----cacagactcagagaga	58
seq_1	taagctggagcctcggt----	74
seq_2	taagctggagcctcggt----	58
seq_3	taagctggagcctcggt----	42

** *** *

Phylogram

Branch length: Cladogram Real



Guide Tree

Figure 1. Graphic representation as Phylogram

Phylogenetic Tree

This is a Neighbour-joining tree without distance corrections.

Branch length: Cladogram Real

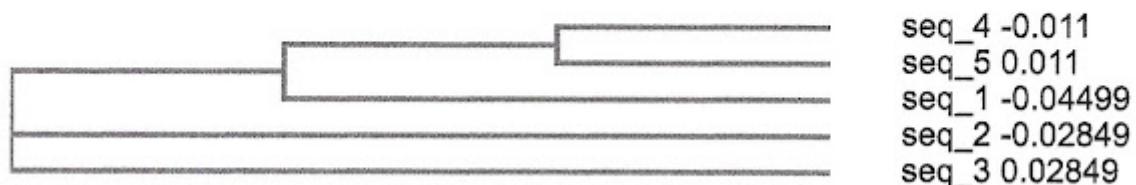


Figure 2. Graphic representation as Phylogenetic tree

In the phylogram the first three sequences are placed together, in the phylogenetic tree by Neighbor Joining sequences 4 and 5 are joined by the first sequence. In that case, the long repeat sends sequence 1 to sequences 4 and 5.

How does WGKS respond to repeats?

Octamer counts have been executed for these five sequences: 83 different octamers have been found. Of these, 81 occur zero or once in a sequence. The remaining two are *acacacac* and *cacacaca*, and occur multiple times. The octamer signatures begin:

string	seq 1	seq 2	seq 3	seq 4	seq 5
aaataccg	1	1	1	0	0
aacacaca	1	1	0	1	1
aagctgga	1	1	1	0	0
aataccgt	1	1	1	0	0
acacacac	13	5	0	14	6
acacacct	1	1	0	0	0
acacactc	0	0	0	1	1
acacctcc	1	1	0	0	0
acacttt	0	0	0	1	1
acagactc	0	0	0	1	1

The correlation matrix of the counts shows that sequence 3 deviates from the other four sequences:

	seq 1	seq 2	seq 3	seq 4	seq 5
seq 1	1	0.905	0.020338	0.915432	0.758678
seq 2	0.905	1	0.262668	0.693422	0.511118
seq 3	0.020338	0.262668	1	-0.27067	-0.43957
seq 4	0.915432	0.693422	-0.27067	1	0.935843
seq 5	0.758678	0.511118	-0.43957	0.935843	1

	seq 1	seq 2	seq 3	seq 4	seq 5
seq 1	1	0.992977	-0.07366	0.985514	0.964082
seq 2	0.992977	1	-0.00813	0.963535	0.936946
seq 3	-0.07366	-0.00813	1	-0.1939	-0.25839
seq 4	0.985514	0.963535	-0.1939	1	0.992437
seq 5	0.964082	0.936946	-0.25839	0.992437	1

$$Sc = (x-1)/(x+1)$$

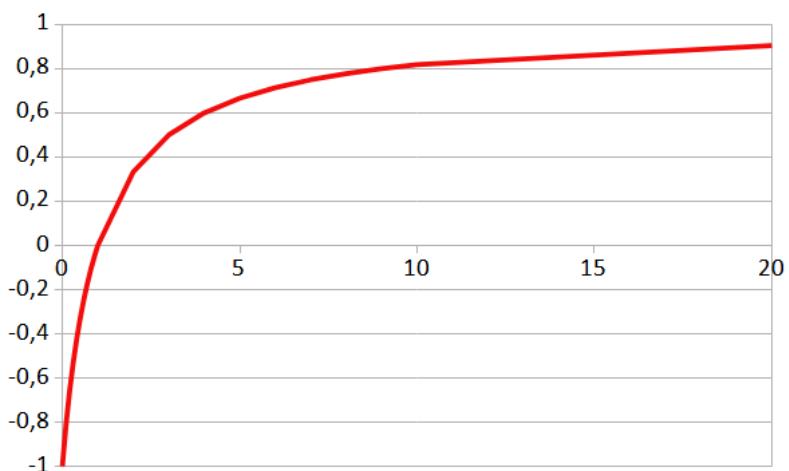


Figure 3 Score $Sc = (x-1)/(x+1)$ as function of x

Logarithmic x-axis

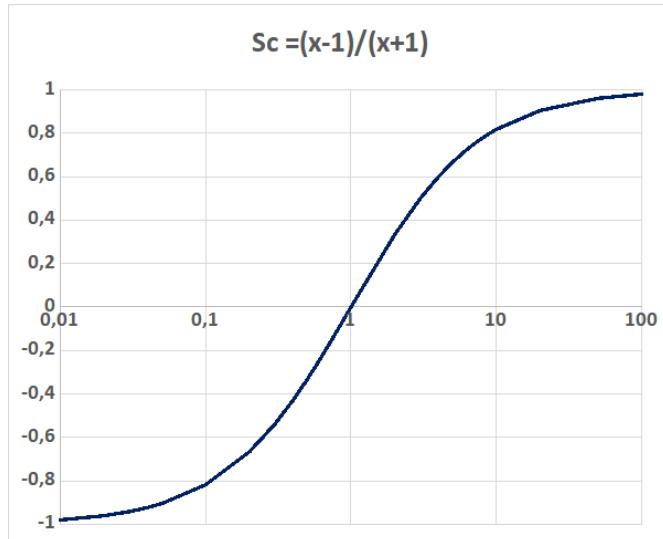


Figure 4 Score $Sc = (x-1)/(x+1)$ as function of x

The score seems rather sensitive to deviations between observed number O and expected number E for fairly small deviations from expected, and insensitive to large deviations from expected.

ii) The second approach looks at the absolute deviation from expected: now genome size becomes an important factor.

There are $4^8 = 65536$ different octamers. Genome size is N base pair. A rough first approximation has expected number of an octamer at $E=N \times 4^{-8}$. Observed number of that octamer is now $O=E+n$.

The score is now given by $Sc = (E+n-E)/(E+n+E) = n/(2E+n)$. Genome size and absolute deviation from expected appear in this score.

The next figure has scores for genome sizes $N=10^6$, $N=10^7$, $N=10^8$, $N=10^9$ (separate lines); the absolute difference between observed and expected ranges from 10^1 to 10^5 and is plotted at the x-axis.

With a large genome, in the order of 10^9 bp, the absolute deviation between observed and expected must be large for the score to change significantly. With a smaller genome, almost all changes lead to high scores.

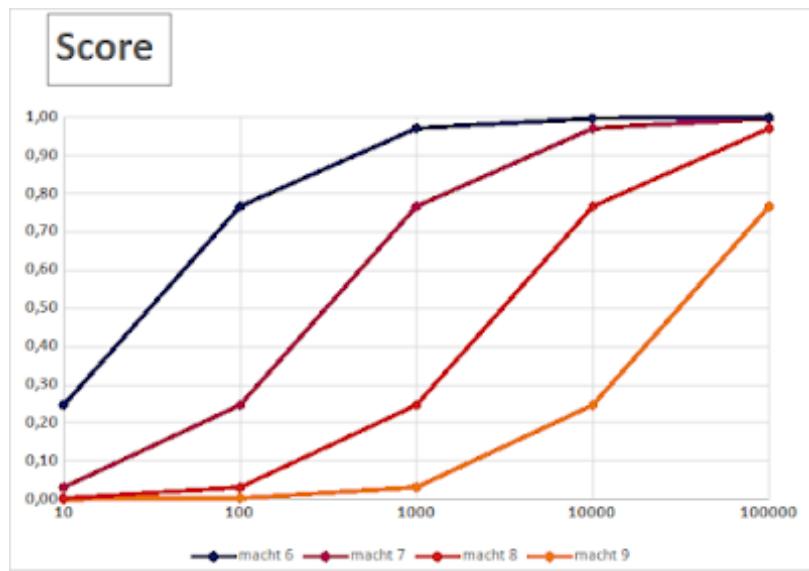


Figure 5. The score as a function of difference between observed and expected for different genome sizes, $N = 10^6$ (blue) to $N=10^9$ (orange)

Small duplications therefore have hardly any effect on the score, while transposon frequencies have major effects. Differences in frequency of transposons such as LINE1 with 15-20% of the genome will have a major influence on the genome, because such transposons occur in such large numbers.

Cserhati writes:

Even if the genome is partially or completely duplicated, then the score value will not change. This is because both the Observed and Expected values will increase by the proportion that the duplicated genome is compared to the pre-duplication genome

This is not correct. The expected number E of an octamer will depend less on the presence of duplications than the observed number O , if only part of the genome is duplicated. The influence of the non-duplicated part of the genome will predominate in E .

3 Take home message

The octamer score $Sc = (O-E)/(O+E)$ is nonlinear making it doubtful whether this is a workable measure of any genome trait.

In a correlation matrix based on WGKS, differences in large amounts of repetitive DNA will have a major impact.

When working with octamer signatures from related species, we have a fairly similar octamer pattern derived from informative DNA; differences in repetitive DNA have a major impact on the correlations between the signatures of the species against that similar background of informative DNA.

When working with octamer signatures of species that are far apart in their phylogeny, we expect a higher influence of the difference in octamer pattern from informative DNA. Against this background of more difference due to informative DNA, the influence of repetitive DNA can diminish, disappear or, when using a large number of species from a group, average out.

A phylogeny on WGKS octamer patterns might perhaps approximate the main lines of a phylogeny, but it cannot be expected that WGKS octamer patters produce an accurate phylogeny on smaller scale comparisons – as with species in a family or even families within a superfamily.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-021-07531-3>

Jachowicz, J.W., Bing, X., Pontabry, J., Bošković,A., Rando, O.J., & Torres-Padilla, M-J (2017) LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. Nature Genetics 49: 1502

de Ferran, V., Figueiró, H., de Jesus Trindade, F, en 17 anderen , & Eizirik, E. (2022) Phylogenomics of the world's otters. Current Biology 32; 3650–3658,

The Python script motif_analysis_k-1.py at github.com/csmatyi/motif_analysis was used to generate WGKS profiles

https://github.com/csmatyi/motif_analysis

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-7-wgks-is.html>

THE RED PANDA AND CSERHATI (7): WGKS IS NOT SUITABLE FOR PHYLOGENY

Cserhati prefers use of the entire genome for species classification, rather than a number of well-characterized genes. The method used by Cserhati to characterize the whole genome is Whole Genome K-mer Signature, abbreviated as WGKS. Cserhati clearly thinks that this WGKS method can be used to arrive at the proper classification of the red panda.

While the WGKS algorithm may not be a sensu stricto phylogenetic algorithm, it can still be used to classify species, based on their WGS into different groups. (BMC Genomics)

Cserhati uses two techniques for red panda and giant panda classification, a phylogenetic tree and clustering.

This post is about the phylogenetic tree; clustering will be discussed later.

Cserhati uses a WGKS data set of 28 species

*To this end, the Whole Genome Kmer Signature (WGKS) algorithm [15] is used to analyze the genomes of five bear species, eleven cat species and ten species from the family Mustelidae (weasels, otters, martens, and badgers), *Spilogala gracilis*, a mephitid species, as well as the red panda *Ailurus fulgens*, making 28 species in total.*

The five species of the bear family are the giant panda and four species of the genus *Ursus*. It is clear here that Cserhati takes the classification of the giant panda with the bears for granted.

The 28 species are listed with their scientific name and their English name in the following table:

familie	subfamilie	soort	nederlandse naam
Mustelidae		<i>Mustela erminea</i>	Ermine
Mustelidae		<i>Mellivora capensis</i>	Honey badger
Mustelidae	otter	<i>Pteronura brasiliensis</i>	Giant otter***
Mustelidae	otter	<i>Enhydra lutris</i>	Sea otter

Mustelidae		<i>Taxidea taxus</i>	American badger
Mustelidae		<i>Neovison vison</i>	American mink
Mustelidae	otter	<i>Lontra canadensis</i>	Northern river otter
Mustelidae		<i>Mustela putorius furo</i>	Ferret
Mustelidae		<i>Gulo gulo</i>	Wolverine
Mustelidae	otter	<i>Lutra lutra</i>	European otter
Ailuridae		<i>Ailurus fulgens</i>	Red panda
Mephitidae		<i>Spilogale gracilis</i>	Western spotted skunk
Ursidae		<i>Ursus thibetanus</i>	Asiatic black bear
Ursidae		<i>Ursus arctos</i>	Brown bear
Ursidae		<i>Ursus americanus</i>	Grizzly
Ursidae		<i>Ursus maritis</i>	Polar bear
Ursidae		<i>Ailuropoda melanoleuca</i>	Giant panda
Felidae		<i>Lynx canadensis</i>	Canadian lynx
Felidae		<i>Felis catus</i>	House cat
Felidae		<i>Puma concolor</i>	Puma
Felidae		<i>Lynx pardinus</i>	Iberian lynx
Felidae		<i>Prionailurus bengalensis</i>	Leopard cat
Felidae		<i>Panthera onca</i>	Jaguar
Felidae		<i>Panthera pardus</i>	Panther
Felidae		<i>Panthera leo</i>	Lion
Felidae		<i>Acinonyx jubatus</i>	Cheetah
Felidae		<i>Panthera tigris</i>	Tiger
Felidae		<i>Felis nigripes</i>	Black footed cat

Cserhati uses the WGKS data to create a phylogenetic tree with the UPGMA method. UPGMA is the simplest and oldest method to construct a phylogenetic tree from sequence data. The major drawback of UPGMA is that this method assumes that the rate of change in the sequence is the same over time and across all lines, ie over the entire phylogenetic tree. That is by no means always the case, and UPGMA is seldomly used anymore.

In his figure 2 (here figure 1) Cserhati gives the UPGMA phylogenetic tree for his 28 species. The layout in the figure has the same order of species from top to bottom as in the table above.

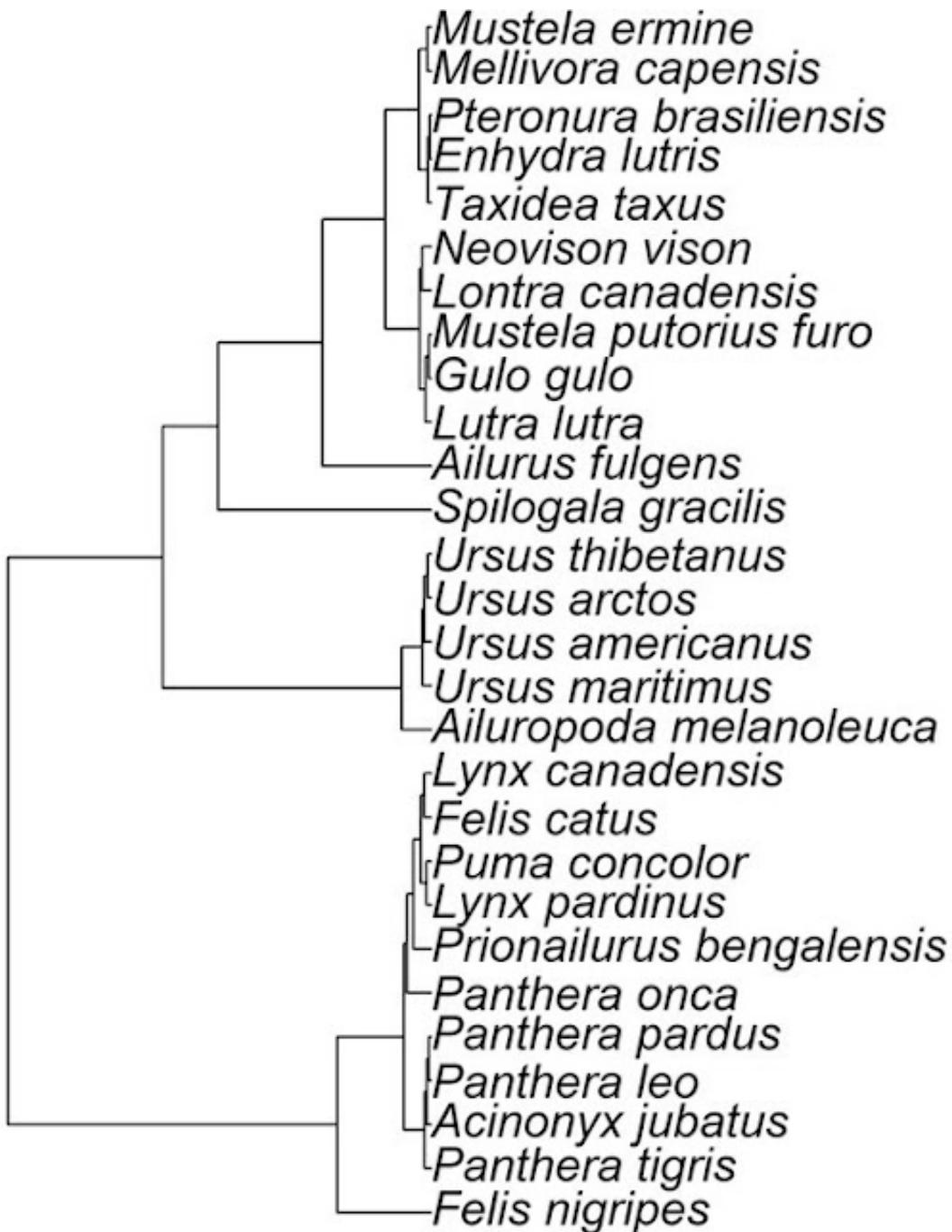


Figure 1. Phylogeny by UPGMA methode, WGKS data: Figure 2 Cserhati BMC Genomics. Horizontal line length corresponds to differnce found.

We clearly see three major groups in Figure 1. The first split is the cats Felidae against all other species, the second split is the bears Ursidae against the Musteloidea. Within the Musteloidea we have the Mephitidae splitting off first, and a sister group relationship of the red panda Ailuridae with the Mustelidae. This corresponds to the phylogenetic tree of Law et al (2018) (in blog post Cserhati 1), in the absence of the raccoon family Procyonidae.

No problems with the major groups, but problems surface in the phylogenetic trees for the families

Let's look at the cats first:

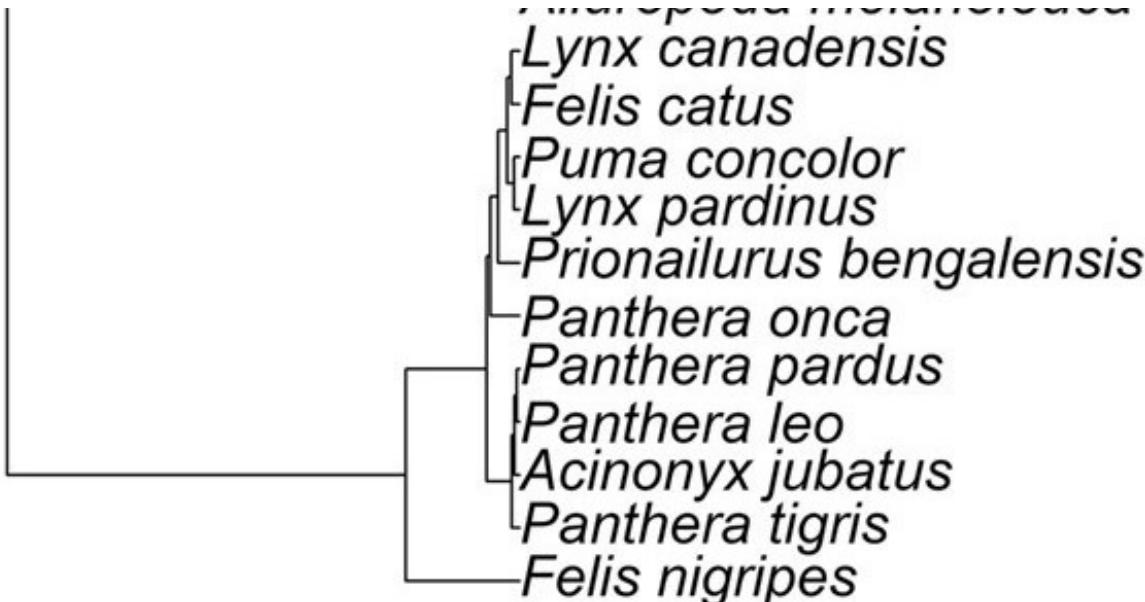


Figure 2 detail of Figure 1, the cats

The phylogeny in Figures 1 and 2 indicates that the black-footed cat *Felis nigripes* is equally related to all other cat species: equally related to the lion, the lynx as to the domestic cat *Felis catus*; equally related to the big cats as to a cat species from the same genus *Felis*. That can't be right.

Moreover tiger, lion and panther are grouped together with the cheetah. That can't be right: the cheetah *Acinonyx* has never been counted among the 'big cats'. The jaguar *Panthera onca*, on the other hand, is here separated from the other 'big cats' of the genus *Panthera*. The Canadian lynx gets the domestic cat as a closest relative, but the Iberian lynx gets the puma as closest relative. Something is going very wrong here.

The species of the mustelid family Mustelidae are messed up too, as much as the cat species:

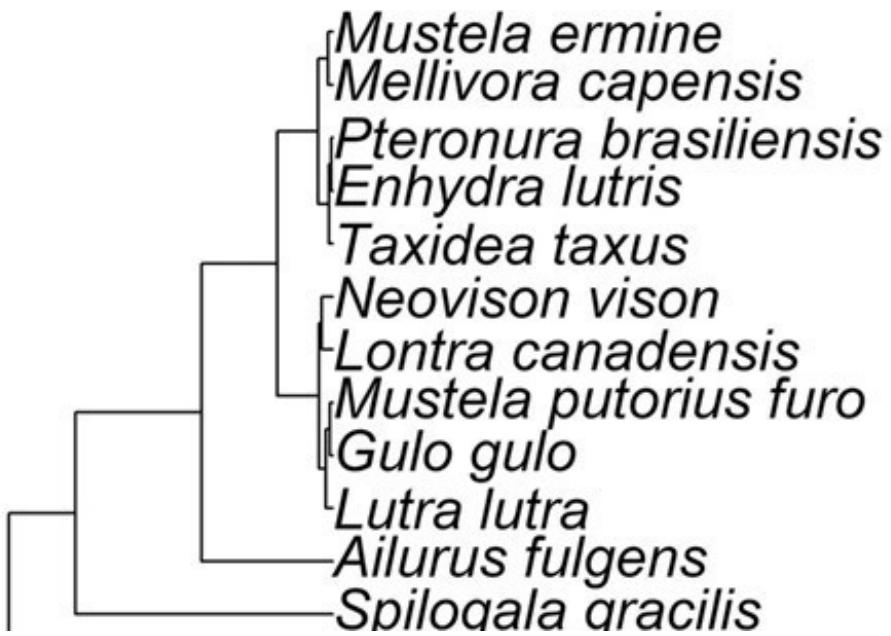


Figure 3. Detail Figure 1, superfamily Musteloidea

The phylogeny in Figures 1 and 3 indicates that the European otter *Lutra lutra* is more closely related to the wolverine and the ferret than to the other three otters. The ferret and ermine, both from the genus *Mustela*, are thrown apart in the first split within the family Mustelidae. One does not need to have a background in biology to see that the positions of the otters and weasels are messed up. The otters and *Mustela* clearly show that this phylogenetic tree cannot represent the relationships between the species in a correct way.

Altogether, the UPGMA phylogenetic tree on Whole-Genome K-mer Signatures shows well-known results among the major pattern of families and superfamilies. On a coarse level WGKS classifieds correctly. Within a family, WGKS cannot be used to assess relatedness. On a more detailed level, WGKS gives junk. It is impossible to say where the transition lies between 'coarse is correct' and 'fine is junk'.

What does Cserhati say about his Figure 2?

Based on this evidence, A. fulgens would belong to mustelids as a monophyletic group. This can also be seen well in Fig. 2, which shows the UPGMA-based phylogenetic tree for the 28 species in the whole genome analysis.

Not so. Note the long horizontal line length between *Ailurus* and the species of the Mustelidae family In fact, Cserhati's Figure 2 shows the red panda as the sister group of the family Mustelidae; it does not show the red panda to belong to that family.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), *BMC Genomics* 22: 228

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-021-07531-3>

<https://en.wikipedia.org/wiki/UPGMA>

Law, C. J.; Slater, G. J. & Mehta, R. S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. *Systematic Biology*. 67: 127–144.

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-8-clustering.html>

THE RED PANDA AND CSERHATI (8): CLUSTERING

Cserhati employs two techniques for red panda and giant panda placement based on the Whole Genome K-mer Signatures: a phylogenetic tree and clustering.

Cserhati does not put much emphasis on the phylogenetic tree based on his WGKS data of 28 species, but gives ample attention to clustering. In his analysis he creates a correlation matrix, the pairwise correlations of the octamer signatures of the species. He displays this matrix in a 'heat map' in which the size of the correlation is shown on a light-dark scale. Based on this matrix, Cserhati searches for clusters.

The last step (in the analysis) involves visualizing the PCC in a heatmap and using clustering algorithms to detect monophyletic groups.

"Using clustering algorithms to detect monophyletic groups".

Really? You're serious?

Is it possible to find or define groups of common descent, monophyletic groups, from clustering?

1 Clustering describes similarity, not phylogeny

In clustering, we have a large number of independent individuals (persons, schools, cars, pieces of music, countries) each with a number of characteristics. In clustering we look for 'who looks like whom', whether groups can be defined so that each individual within the group resembles every other individual in that group more than an individual in the other group. Such a group is called a cluster. A cluster indicates similarity. Lineage is not an issue in clustering. Nor is monophyly: that is common descent, not clustering.

2 A phylogeny yields clusters

If we have a phylogeny, and choose species from it: eg 5 monkeys, 5 rodents, 5 cattle, 5 bats, and run a clustering program, we are guaranteed (here) 4 clusters, each with 5 species. We introduce existing groups, and when existing

known groups are put into a clustering program, those known groups will be found.

In the phylogeny of the carnivore order Carnivora, we have the cat family Felidae, bear family Ursidae, and the superfamily Musteloidea. When Cserhati submits these 28 species of carnivores belonging to these monophyletic groups to clustering, he finds three clusters: cats, bears and Musteloidea. Cserhati put species from three monophyletic groups in a cluster analysis, and of course found those monophyletic groups.

3 A monophyletic group gives a monophyletic cluster

Not the other way round. A group of species is not monophyletic because they cluster together, they cluster because they are monophyletic.

4 Monophyly cannot be concluded based on clustering, .

'Cluster' and 'monophyletic group' will often coincide when biological species are being clustered, but when clustering some seal species with *Poiana leightoni*, *Poiana richardsonii* (Africa linsangs), *Prionodon linsang* and *Prionodon pardicolor* (Asian linsangs) we (presumably) get two clusters, one with the seals and one with the linsangs. This while the Asian linsangs and the African linsangs do not belong to the same family or superfamily. Such a linsang cluster is heterogeneous. Nothing in clustering per se tells us that a cluster would be monophyletic.

5 Clustering and Input

The number of clusters depends on the scope of the input. Among the 28 species with WGKS data from Cserhati, there is a cluster of cats, a cluster of bears and a cluster of Musteloidea. There are 12 species in the Musteloidea cluster and this cluster does not split any further. In Cserhati's analysis on mtDNA there are 37 species of Musteloidea, and the Musteloidea split into 4 clusters: viz the four monophyletic families Mephitidae, Ailuridae, Procyonidae and Mustelidae - giving clusters because they are a monophyletic family. The 10 species of Mustelidae do not split further into clusters, but just entering many Mustelidae species can lead to different clusters.

6 A phylogeny is hierarchical, clustering is not

Clustering cannot discern a hierarchical structure in the data. The hierarchy in the classification of the living creatures appears only on repeating the clustering input of different scopes (see example under point 5). This means that clustering of biological groups does not give a good representation of the hierarchical structure of the animal world.

7 Clusters have nothing to do with relatedness

Clusters only give the optimal split of the data, nothing about relatedness, neither within nor between clusters. Relatedness follows from the phylogeny.

For examples, see points 4 and 5. The family Mustelidae is monophyletic and the species are related, but if you put otters, weasels and martens in a clustering program you will find several clusters. The families of the monophyletic superfamily Musteloidea are related, but in a clustering program they emerge as a cluster given sufficient scope of input. If only all cats species are put in a cluster program, you get a cluster of 'big cats' and a cluster of 'little cats'. The 'big cats' and the 'little cats' are related, even though they end up in different clusters

8 Clustering is a statistical trick, not a biological classification

A phylogeny is biology, clustering statistics.

Summarizing, "*using clustering algorithms to detect monophyletic groups*" shows no insight into clustering or phylogeny or biology. Clustering in the same cluster cannot be used to conclude to monophyly and relatedness. Clustering in two clusters cannot be used to conclude the species in the separate clusters are not related.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), *BMC Genomics* 22: 228

<https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12864-021-07531-3>

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-9-clusters-op.html>

THE RED PANDA AND CSERHATI (9): CLUSTERS ON WGKS DATA

Cserhati uses two techniques for red panda and giant panda placement based on the Whole Genome K-mer Signatures: a phylogenetic tree and clustering.

Cserhati does not place much emphasis on the phylogenetic tree based on his WGKS data, but gives ample attention to clustering of the 28 species. Based on clustering, Cserhati says in 'Results and Discussion':

Based on this evidence, A. fulgens would belong to mustelids as a monophyletic group.

In 'Conclusion' Cserhati says:

In conclusion, A. fulgens possibly belongs to Mustelidae, based on the analysis of the WGKS.

The UPGMA phylogenetic tree on WGKS data indicates that the red panda is the sister group of the marten family Mustelidae, but does not belong to the marten family. These statements by Cserhati come from his clustering of the WGKS data.

What does Cserhati do with the WGKS data?

In his analysis he makes a correlation matrix of pairwise correlations of the WGKS data of the species. He displays this matrix in a 'heat map' in which the size of the correlation is shown on a light-dark scale.

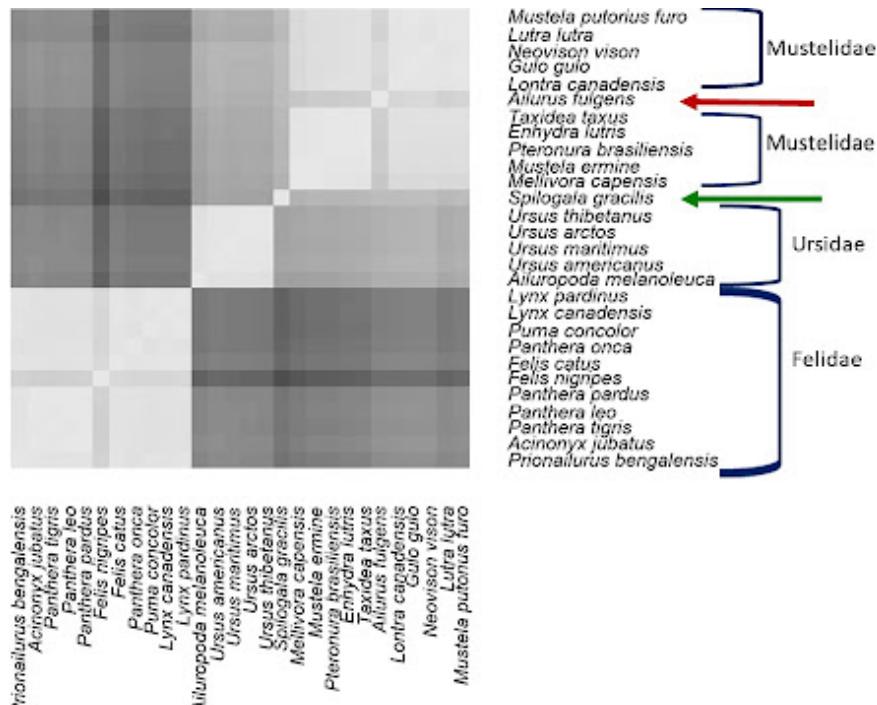


Figure 1 Heat map: pairwise correlations on WGKS data. The order on the x-axis from left to right is the same as the order on the y-axis from bottom to top. The south-west to north-east diagonal gives identity. Red arrow: red panda *Ailurus fulgens*; green arrow skunk *Spilogale gracilis*. Blue: family names. Lighter is higher correlation between species. This is Cserhati fig 1 BMC Genomics with group names added.

How the species order on both axes came about is not mentioned in Cserhati's BMC Genomics article. It is clear that the cats are grouped together, and also the bears and Musteloidea are grouped together. The heatmap thus provides a rough but clear first access to the hierarchical classification of the animals.

It is clear that the red panda differs from the species of the family Mustelidae: we see a 'Finnish flag' pattern: the correlations of the red panda with the species of the marten family are lower than the correlations between the species of the family Mustelidae themselves . That's why we see those dark stripes. But why is the red panda in the middle of the mustelids? No explanation from Cserhati.

Let's take another look at the heatmap. Fortunately, Cserhati has put the numbers of the correlation matrix in one of the supplementary files of the article, so that a colored version can be made in Excel:

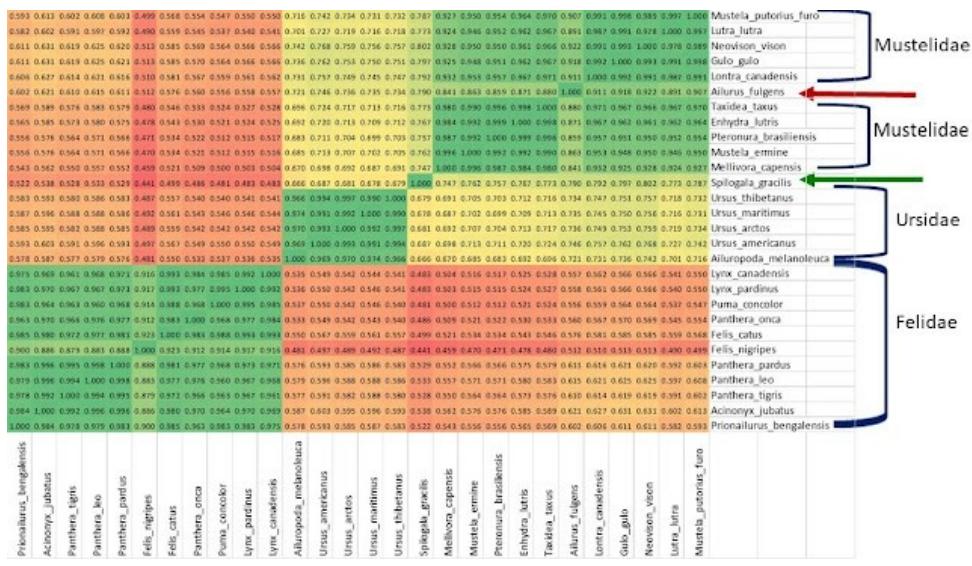


Figure 2 Heatmap according to Cserhati on pairwise correlations on WGKS data. The order on the x-axis from left to right is the same as the order on the y-axis from bottom to top. The south-west to north-east diagonal gives identity. Greener is higher correlation between species, redder is lower correlation between species Red arrow: red panda *Ailurus fulgens*; green arrow skunk *Spilogale gracilis*. Blue: family names.

This is Cserhati fig 1 BMC Genomics with group names added, and colored

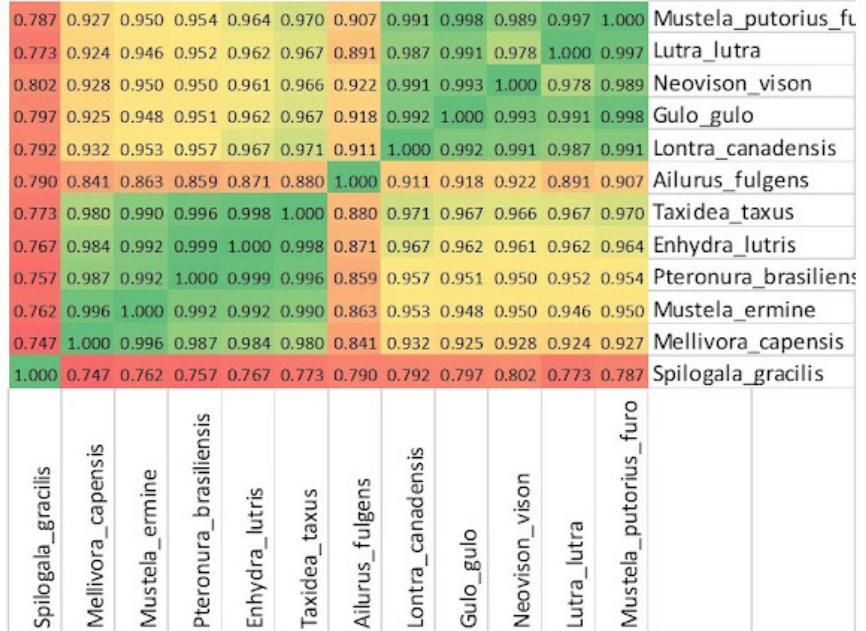


Figure 3 Heatmap according to Cserhati on pairwise correlations on WGKS data of the superfamily Musteloidea. The order on the x-axis from left to right is the same as the order on the y-axis from bottom to top. The south-west to north-east diagonal gives identity. Greener is higher correlation between species, redder is lower correlation between species. This is Cserhati fig 1 BMC Genomics with group names added, and colored

It is clearly visible that the red panda has lower (redder) correlations with the species of the marten family than the species of the marten family have among each other. It is not explained why the red panda *Ailurus fulgens* is placed amongst the species the marten family Mustelidae. Cserhati says:

A. fulgens clearly clusters together with the mustelids, although on average, it has a lower mean PCC value compared to all the other species, 0.89 ± 0.03 , whereas mustelids have a mean PCC value of 0.95 ± 0.04 .

This difference is not too significant.

How significant is "not too significant"? Considering the correlations as independent numbers, we get the mean of all correlations of the red panda with the species of the marten family as 0.89 ± 0.03 , and the mean of all correlations of the species of the marten family among themselves (without red panda) 0.97 ± 0.2 . This is a significant difference. In a two-tailed t-test with unequal variance, this difference is found with a probability of $P = 1.44 * 10^{-6}$. A one-tailed t-test with equal variance gives $P = 3.41 * 10^{-14}$. The WGKS correlations of the red panda with mustelid species differ significantly from the WGKS correlations between mustelids. It remains unclear why Cserhati finds a significance of $P = 1.44 * 10^{-6}$ or $P = 3.41 * 10^{-14}$ "not too significant".

Cserhati uses a clustering program on the data in the correlation matrix, and finds three or four clusters. The cats and bears give clear clusters. The Musteloidea with skunk, red panda and 10 species of mustelids can be considered as a cluster, but Cserhati prefers the skunk on its own and a cluster of 11 species, the 10 species of mustelids with the red panda.

Why is visible in the heatmaps for all species used, as soon as the correlations are sorted by their difference with the red panda:

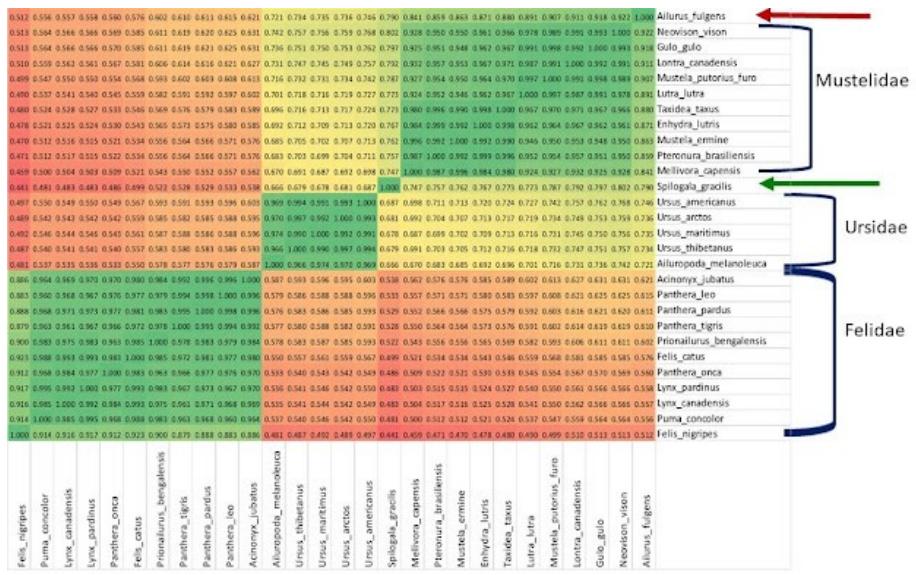
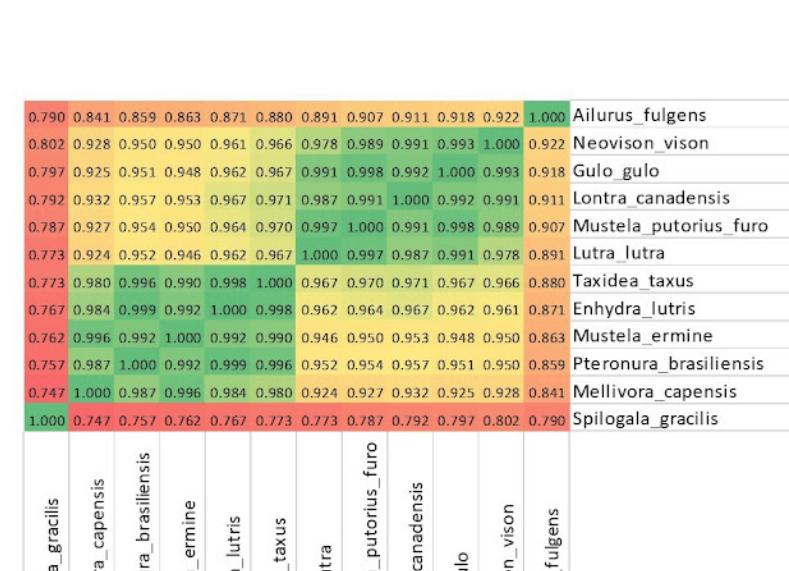


Figure 4 Heatmap on pairwise correlations on WGKS data from Cserhati. The correlations are sorted by similarity to the red panda, which is now far right on the x-axis and highest on the y-axis. The order on the x-axis from left to right is the same as the order on the y-axis from bottom to top. The south-west to north-east diagonal gives identity. Greener is higher correlation between species, redder is lower correlation between species
Red arrow: red panda *Ailurus fulgens*; green arrow skunk *Spilogale gracilis*.
Blue: family names



When we look at the superfamily Musteloidea in more detail in a heatmap, it is clear that setting the skunk apart is obvious. Not much remains of a cluster of 10 species of mustelids + the red panda: the red panda clearly differs from the mustelids (figure 5). The two groups within the mustelid family are the two groups found within the mustelids in the phylogenetic tree with UPGMA on the WGKS data. (That mess in which the otters and the weasels end up as strange bed fellows.)

All in all:

Although the red panda clusters with the species of the mustelid family when a remote group as the cats are present, the red panda differs from the mustelids; exactly as in the phylogeny on the same data. It is clear from the phylogeny that the red panda and the marten family are monophyletic, but as sister groups: not because the red panda belongs to the marten family Mustelidae. When Cserhati says:

Based on this evidence, A. fulgens would belong to mustelids as a monophyletic group.

the "*monophyletic*" is correct, but "*would belong to mustelids*" not. After all, the family Ailuridae and the family Mustelidae are sistergroups, and that means monophyletic.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-10-wasberen.html>

THE RED PANDA AND CSERHATI (10): PROCYONIDAE WGKS?

In the introduction to his BMC Genomics article, Cserhati cites Flynn et al (2000), Yu and Zhang (2006), Sato et al (2009) and Yu et al (2011). The four articles give the raccoon families Procyonidae and mustelid families Mustelidae as sister groups of each other, together called Musteloidea *sensu stricto*. In addition, all the articles mentioned list the Musteloidea *sensu lato* as a superfamily with the four families red panda Ailuridae, skunks Mephitidae, Procyonidae and Mustelidae.

For Cserhati, too, the question should therefore have been what the mutual order of the Ailuridae, the Mephitidae and the Musteloidea *sensu stricto* is within the Musteloidea *sensu lato*.

Cserhati emphasizes the importance of Whole Genome data.

The advantages of using a genomics-based algorithm to analyze the WGS of these organisms is that it takes all the information present in the WGS, as opposed to just a handful of genes, utilized in gene studies

A WGKS profile for all families within the Musteloidea, the four families red panda Ailuridae, skunks Mephitidae, Procyonidae and Mustelidae, was therefore indicated in order to find the place of the red panda. The raccoons are missing from Cserhati's WGKS analysis. Why?

One answer is that no whole-genome study of the raccoon family was available. At least, no whole genome study of the raccoon family was available when Cserhati wrote his BMC Genomics article. Publishing history for the BMC Genomics article is given as: Received: 18 May 2020 Accepted: 14 March 2021.

The only paper I can find on whole genome sequencing for the raccoon family appeared when Cserhati's BMC Genomics paper was under review, online December 17, 2020:

Tsuchija et al, 2021. Whole-Genome Sequencing of Procyonids Reveals Distinct Demographic Histories in Kinkajou (*Potos flavus*) and Northern Raccoon (*Procyon lotor*) Genome Biology and Evolution 13, January 2021, evaa255, Published online: 17 December 2020

Alertness would have been required to notice this in time to add the Procyonidae WGKS to the BMC Genomics paper. However, time was quite sufficient for Procyonidae WGKS to be added to the CRSQ paper.

However, does it matter that the raccoons are missing from the WGKS analysis?

The raccoons are always the sister family of the mustelid family. The possible divisions of the superfamily Musteloidea are:

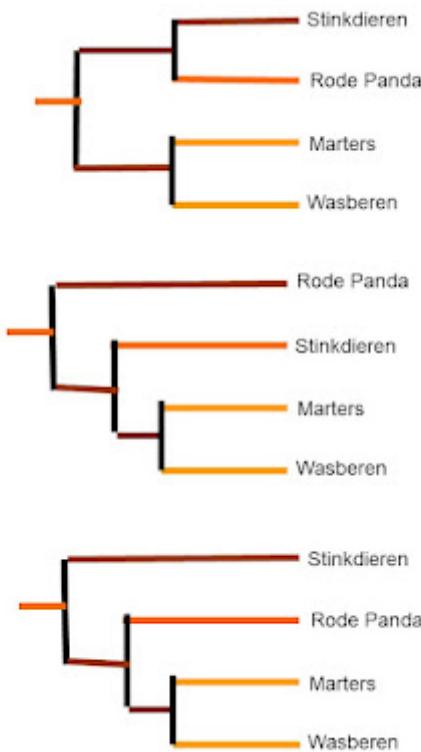


Figure 1 The three distinct configurations of the phylogenetic tree of the Musteloidea. If the Procyonidae are absent, the three configurations are still distinct. (Wasberen = Procyonids; Marters = Mustelids; Rode Panda = Red Panda; Stinkdieren = Mephitids)

In the absence of a species of the raccoon family, a sister group relationship of the red panda Ailuridae with the Mustelidae appears in the lowest diagram for the Musteloidea phylogenetic tree. This agrees with the phylogenetic tree of

Law et al (2018). This means that the raccoons are not strictly necessary to see what placement of the red panda yields the WGKS data.

However, the presence of one or more species of the raccoon family could have prevented a sister group relationship between the red panda and the mustelid family from being misinterpreted as if the red panda belonged to the mustelids.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

M.T.N. Tsuchiya, R.B. Dikowkow, K.P Koepfli, P.B Frandsen, L.L Rockwood, J.E Maldonado, 2021. Whole-Genome Sequencing of Procyonids Reveals Distinct Demographic Histories in Kinkajou (*Potos flavus*) and Northern Raccoon (*Procyon lotor*) Genome Biology and Evolution, Volume 13, Issue 1, January 2021, evaa255; Published online: 17 December 2020

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-11-clustering.html>

THE RED PANDA AND CSERHATI (11): CLUSTERING MTDNA

In addition to WGKS, Cserhati uses mitochondrial DNA to classify the red panda. He uses the full mtDNA as reported in GenBank for 52 species: 15 species and subspecies of the bear family, the two subspecies of the red panda, three species of skunks, 30 species of the mustelid family, and now also two species of the family of the raccoon family, the raccoon itself and the coati *Nasua nasua*.

Cserhati again starts by making a heat map (BMC Genomics Figure 3) to show the correlation matrix between the species. Cserhati writes:

Three larger clusters and two smaller clusters are visible in the heat map.

That can hardly be right.

It looks more like two really big clusters; or eight or even nine clusters.

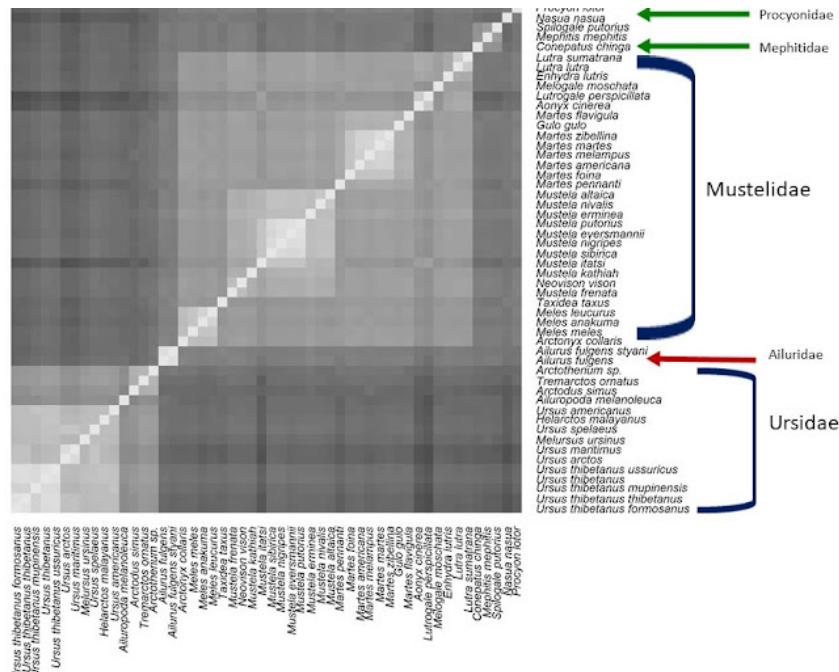


Figure 1. Figure 3 BMC Genomics. Heatmap showing the size of the correlations in mtDNA between the species. The red pandas, family Ailuridae, are plotted here far from the skunks Mephitidae and the raccoon and coatis Procyonidae. Lighter is higher correlation, darker is lower correlation.

Cserhati gives the matrix with correlations in mtDNA between the 52 species in Additional File 2. I copied that correlation matrix in Excel, and sorted it so that the red pandas, the raccoon, the coatis and the skunks are next to each other, and then used the high-low color option in Excel. The following version of Cserhati's matrix emerges:

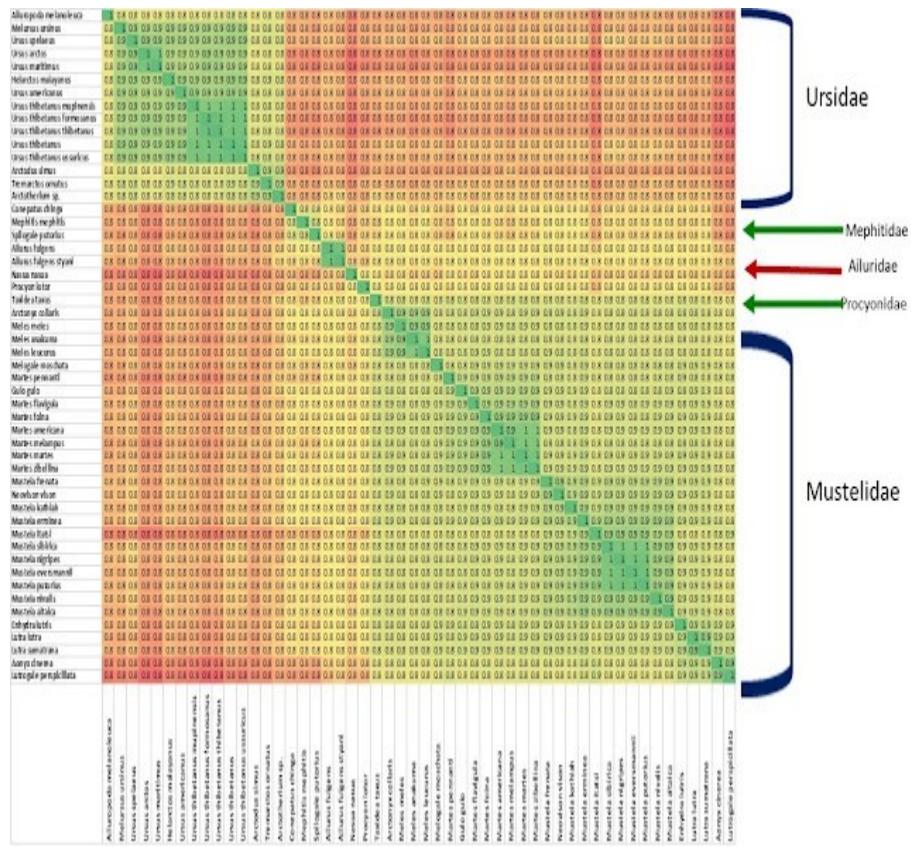


Figure 2. Heatmap showing the size of the correlations in mtDNA between the species. Greener is high correlation, redder is low correlation, yellow in between. Figure 3 of Cserhati BMC genomics with the species in different order, and different color.

The first thing to notice is the division into two large groups: the bears Ursidae and the superfamily Musteloidea with the four families Mephitidae, Ailuridae, Procyonidae and Mustelidae. The contrast yellow / red makes the Musteloidea visible in the heatmap.

The Musteloidea as an important group is reflected in the clustering data given by Cserhati in his Table 3. The first split into clusters is between the Ursidae and the Musteloidea. Two clusters is what the data say.

	cluster	species	min	mean	max	stdev	p-value	neglog
<i>Ursidae</i>								
<i>Ursidae</i>	1	15	0.811	0.88	0.989	0.048	5.03E-41	40.298
<i>Musteloidea</i>								
<i>Mustelodea</i>	2-5	37	0.769	0.837	0.981	0.037	3.30E-185	184.481
<i>Mephitidae</i>	2	3	0.83	0.838	0.849	0.01	0.0117376	1.93
<i>Procyonidae</i>	3	2	0.803	0.803	0.803	NA	1.90E-17	16.721
<i>Ailuridae</i>	4	2	0.98	0.98	0.98	NA	1.96E-122	121.708
<i>Mustelidae</i>	5	30	0.822	0.858	0.981	0.029	1.70E-201	200.769

Table 3 of Cserhati with minor modification from Additional file 2. Min, Max refer to lowest and highest correlation of a group. Mean and StDev to mean and standard deviation of a group's correlations. P-value must refer to the clusters.

Cserhati moves from two to five clusters, each representing a family. Why does he stop at five clusters? Why doesn't Cserhati come up with clusters within the family Mustelidae?

Look again at the Mustelidae: their correlations are colored just a little greener, not as yellowish as the correlations with the other species of the superfamily Musteloidea. Within the Mustelidae we see beautiful green areas. Perhaps clusters within the Mustelidae? The Mustelidae are a large family with subfamilies. The weasel subfamily gives a cluster as good as the skunk family, for example.



Figure 3. Heatmap showing the size of the correlations in mtDNA between the species. Greener is high correlation, redder is low correlation, yellow in between. Families and subfamilies have now been given a box. Figure 3 from Cserhati BMC Genomics with the species in different order. (beren = bears; stinkdieren = skunks; wasberen = raccoon family; dassen = badgers; martens = martens; wezels= weasels; otters = otters)

How does Cserhati arrive at five clusters?

Not from statistics. Look at Additional file 5: Figure S3 with legend by Cserhati:

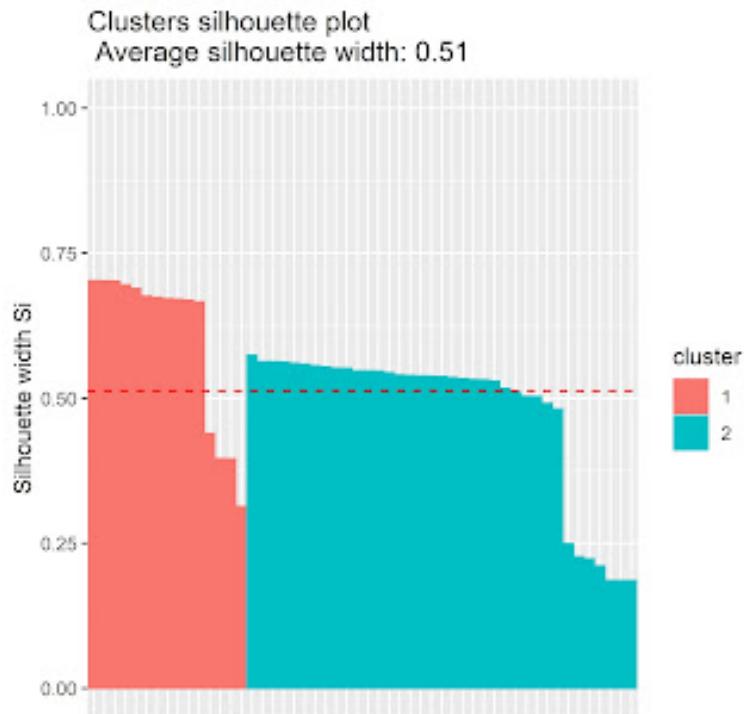


Figure 4. Additional figure 5 BMC Genomics. Plot showing the mean silhouette width according to the number of clusters for the mitochondrial data, based on the 'silhouette' method. The maximum average silhouette width is 0.51 for two clusters.

Additional file 5 Figure S3 in the BMC genomics article gives two clusters—the bear family and the superfamily Musteloidea.

Compare this figure with Additional Figure S3 from the CRSQ article, below. Cserhati's other article, but about the same data. The optimal number of clusters is two. Five clusters is not better than six, seven, eight or ten clusters: all are perfectly good possibilities. I've indicated eight clusters in Figure 3 here, emphasizing the subfamilies within the Mustelidae; five clusters is not a better solution than eight clusters.

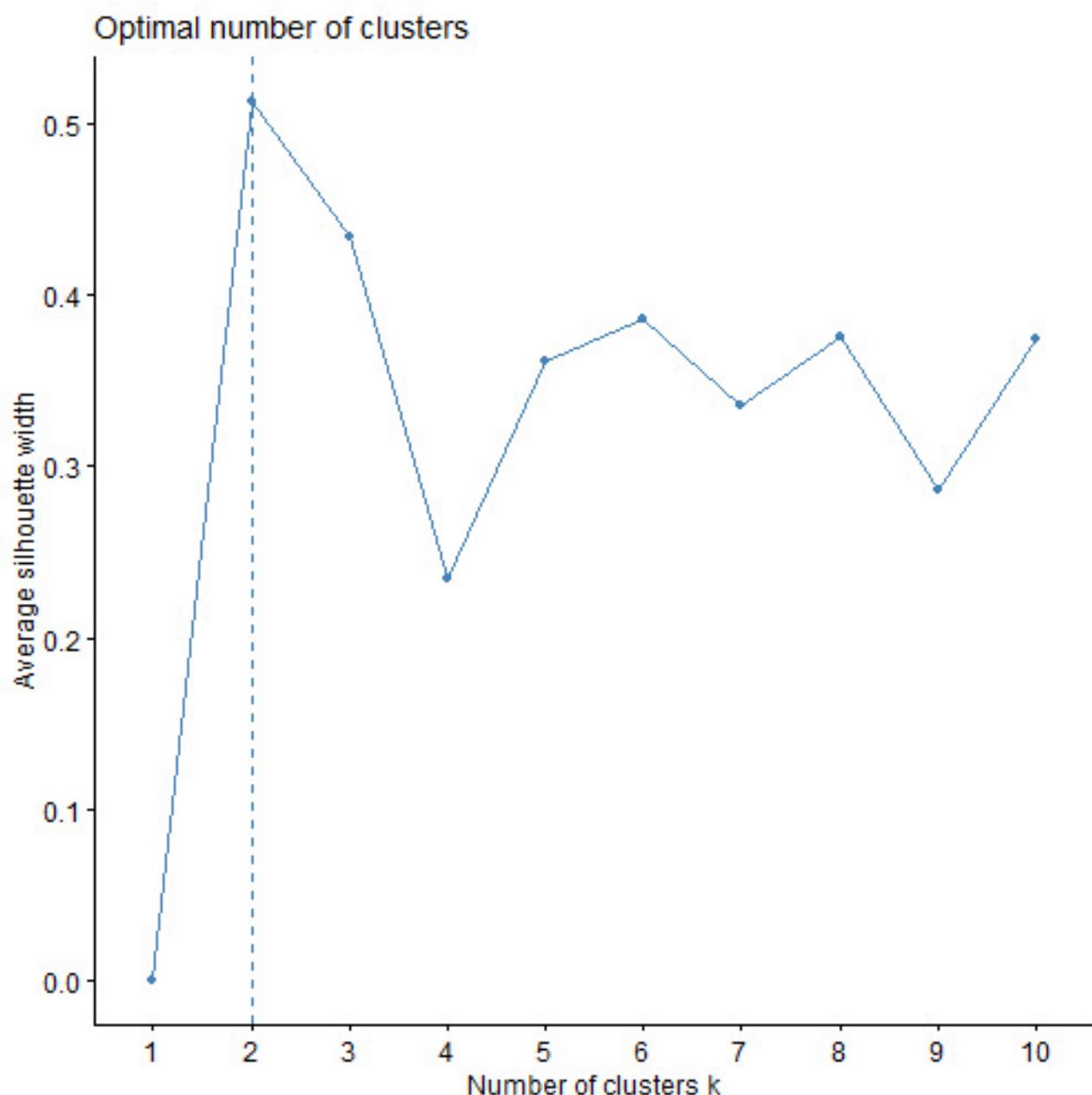


Figure 5 Additional Figure S3 from the CRSQ. Cluster width at different numbers of clusters, and optimal number of clusters, for the mitochondrial data.

Statistically, Cserhati should have stayed with two clusters, family Ursidae and superfamily Musteloidea. Based on clustering, eight clusters is just as good a solution as five clusters.

It might be that Cserhati prefers five clusters because there are species from five families in the data. If so, the clustering was superfluous - the result was known.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-12.html>

THE RED PANDA AND CSERHATI (12): PHYLOGENETIC TREES BASE ON mtDNA

Cserhati presents three 'hierarchical trees' based on mitochondrial DNA. He uses the complete mtDNA as reported in GenBank for 52 species: 15 species and subspecies of the bear family, the two subspecies of the red panda, three species of skunks, 30 species of the mustelid family, and now also two species of the raccoon family, the raccoon itself and the coati.

Cserhati uses three different methods to obtain his 'hierarchical trees'. Cserhati does not call his results a phylogenetic tree or phylogeny, see his captions to his figures.

So hierarchical trees, with three phylogenetic methods: UPGMA, NJ and ML. Phylogenetic trees, whatever Cserhati avoids to call them.

UPGMA and NJ use the differences = distances, between sequences.

UPGMA is the simplest and oldest method to construct a phylogenetic tree from sequence data. The major drawback of UPGMA is that this method assumes that the rate of change of the sequences is the same over time and across all lineages, ie over the entire phylogenetic tree. UPGMA produces an implicitly rooted tree.

In his Figure 4, Cserhati provides a UPGMA phylogenetic tree

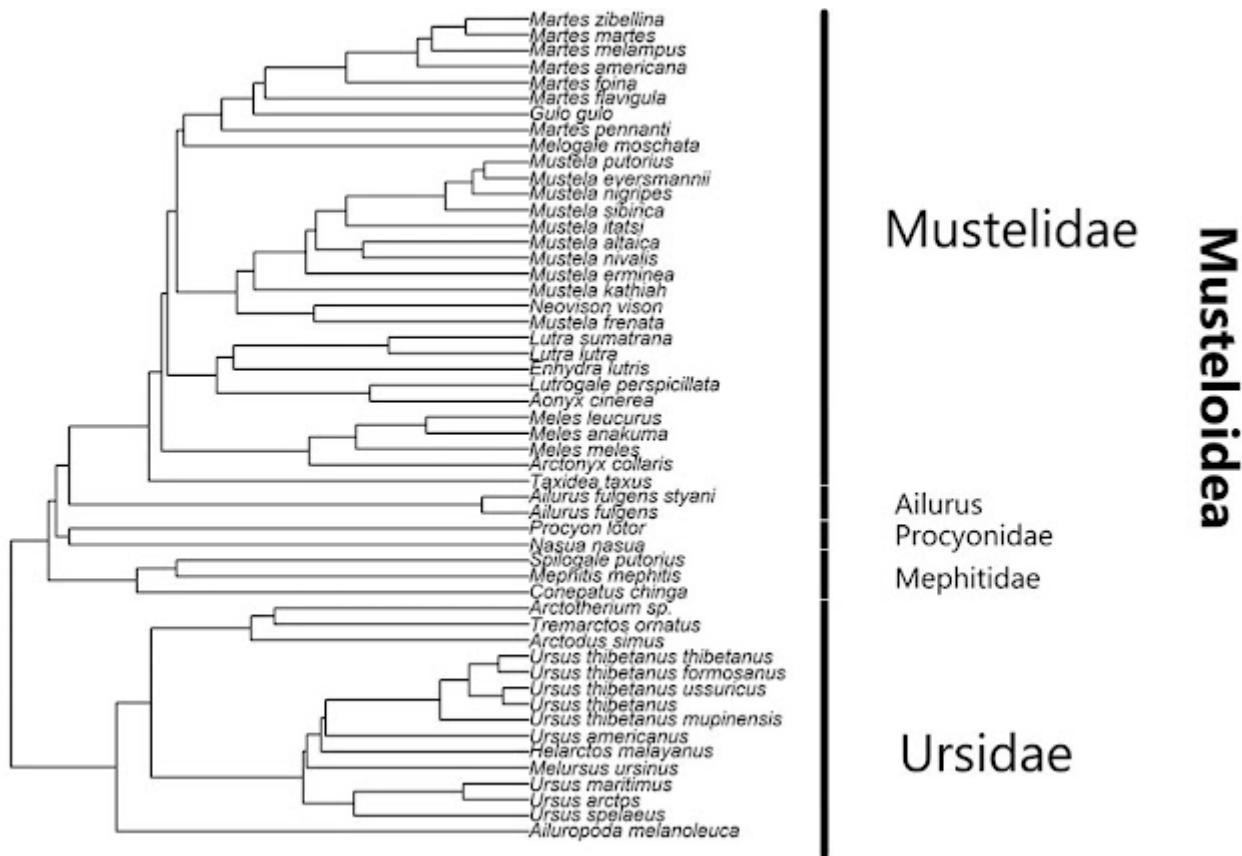


Figure 1: Figure 4 from Cserhati BMC Genomics. UPGMA phylogenetic tree.

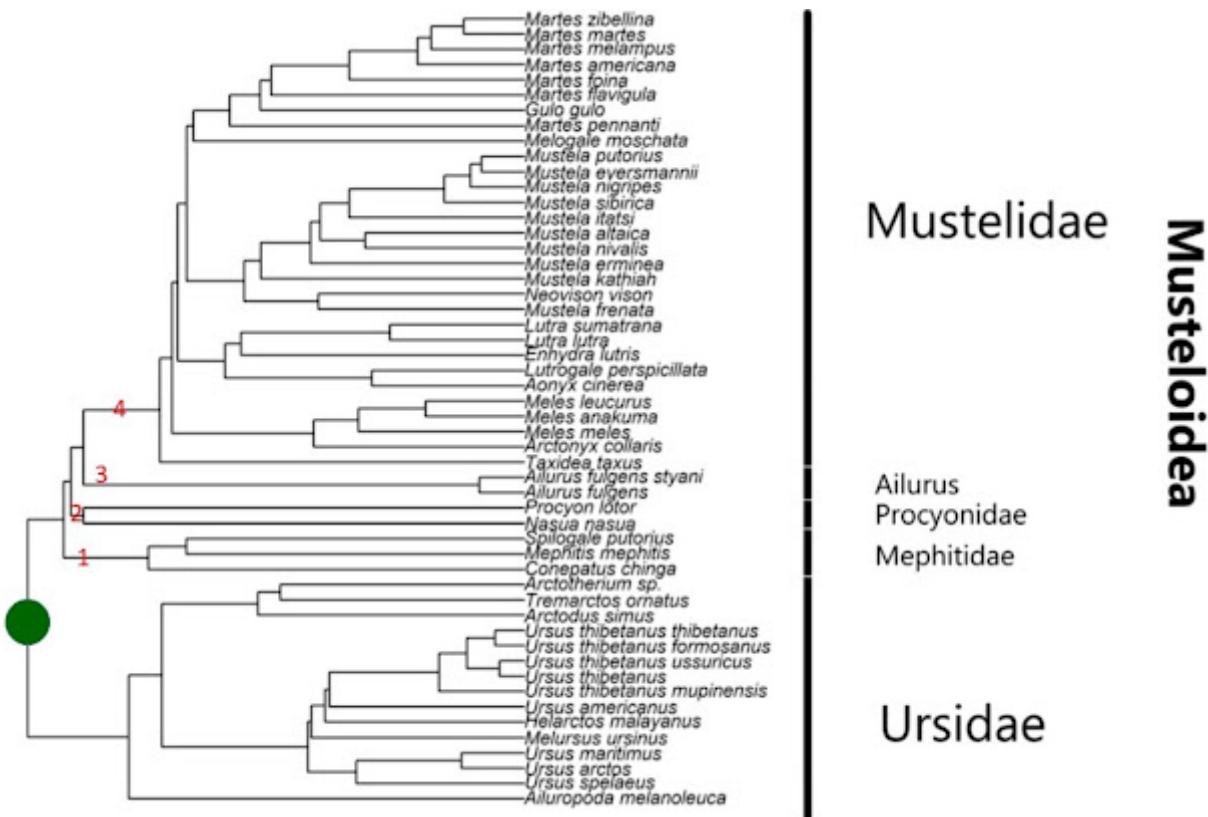


Figure 2: Figure 4 from Cserhati BMC Genomics sequence of splits within Musteloidea indicated.

The legend to his Figure 4 as given by Cserhati is:

UPGMA-based hierarchical tree for the 52 species analyzed in the mtDNA study, based on sequence identity metrics. Mustelids and ursids form two large clades, and mephitids, procyonids forming two small groups. Ailurus fulgens and Ailurus fulgens styani appear either to form their own clade, or loosely associate with mustelids.

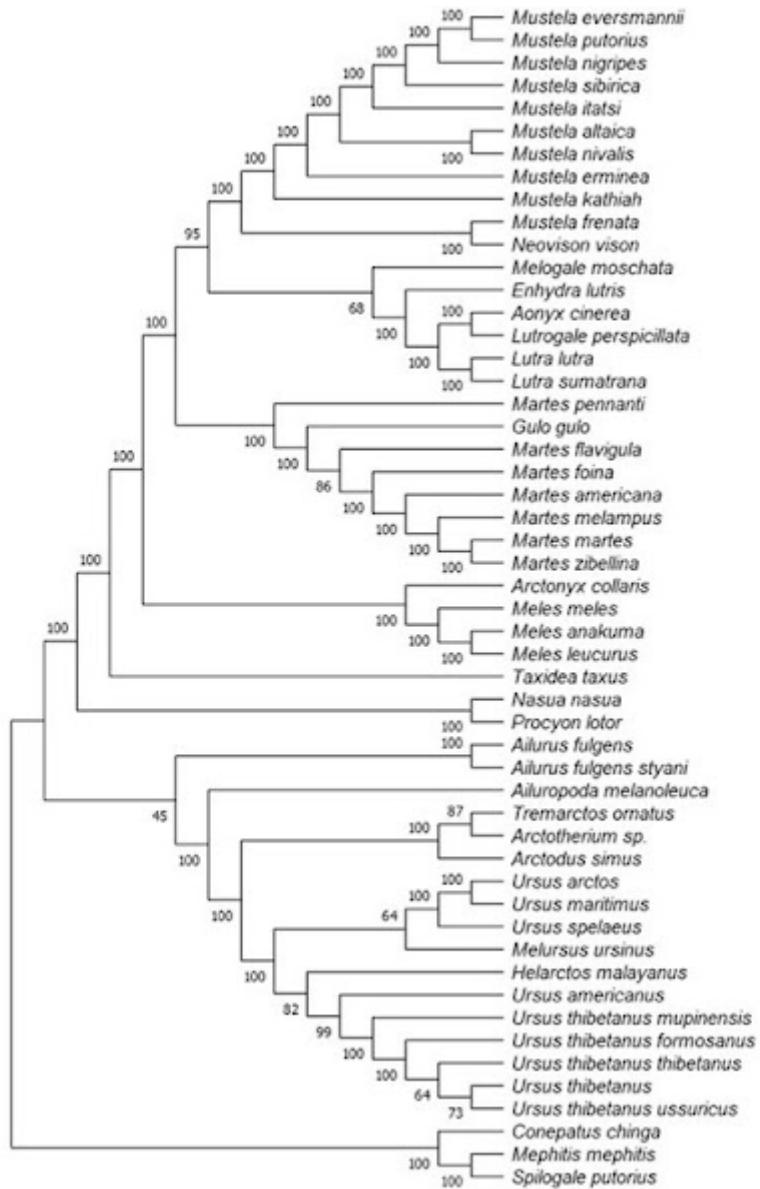
“ or loosely associate with mustelids.”

Very loosely associated.

This Figure 4 clearly shows that the family Ailuridae, with the two red panda subspecies *Ailurus fulgens* and *Ailurus fulgens styani*, appears as the sister group of the family Mustelidae. So their own clade, here as a sister group to the mustelids. Remarkably, the Ailuridae appear as a sister group to the mustelids in the presence of the Procyonidae, the raccoon family; this has no precedent in the literature.

Neighbor-Joining takes the distances between sequences, pairs the two sequences with the smallest distance, and continues with the pair as if it were a taxonomic unit. NJ always searches where the smallest distance can be found. The NJ method gives unrooted trees, without any assumption about rate of change over time.

An NJ program might plot an unrooted tree as a cladogram: the program then inserts a fictitious root to be able to establish a lay-out. It is up to the investigator using NJ to ensure that the tree is rooted, by providing an outgroup. Cserhati writes that he uses the default settings, and that means, no outgroup. An unrooted tree was plotted as a cladogram: the lengths of the horizontal lines are not a representation of the number of differences between the species. The layout gives on its own no conclusion about the phylogeny. Cserhati indicates how well substantiated the splits are, but that does not establish the phylogeny.



Mustelidae

Procyonidae Ailurus

Ursidae

Mephitidae

Figure 3: Figure 5 Cserhati BMC Genomics. Neighbour Joining phylogenetic tree plotted as cladogram. The horizontal distance between successive splits is the same; only the topology of the tree is of importance. .

Cserhati's description of his result shows that he does not grasp the difference between a layout and a phylogenetic tree:

Mustelidae forms a well-defined clade, with almost all branch points supported with a bootstrap value of 100. Nasua nasua and Procyon lotor form a smaller clade right next to Mustelidae. The three mephitids, Conepatus chinga, Mephitis mephitis and Spilogale putorius also form a small clade, well separated from the other clades. The NJ method places Ailurus next to Ursidae, suggesting that they possibly form a monophyletic group. However, the node connecting Ailurus with Ursidae only has a bootstrap value of 45.

A phylogenetic tree is a mobile, and we can start flipping the Mephitidae, the skunks, upwards. We also have an unrooted tree, so we can put a root

between the bears Ursidae as outgroup and all the other species, all of which belong to the Musteloidea. The low bootstrap value of the node between bears and *Ailurus* also argues in favor of placing the root between *Ailurus* and bears.

We take the bears as an outgroup to the Musteloidea, and walk along the lines to find the successive splits in the Musteloidea. Then one of the well-known phylogenetic trees of the Musteloidea emerges, namely the tree in Flynn et al (2000).

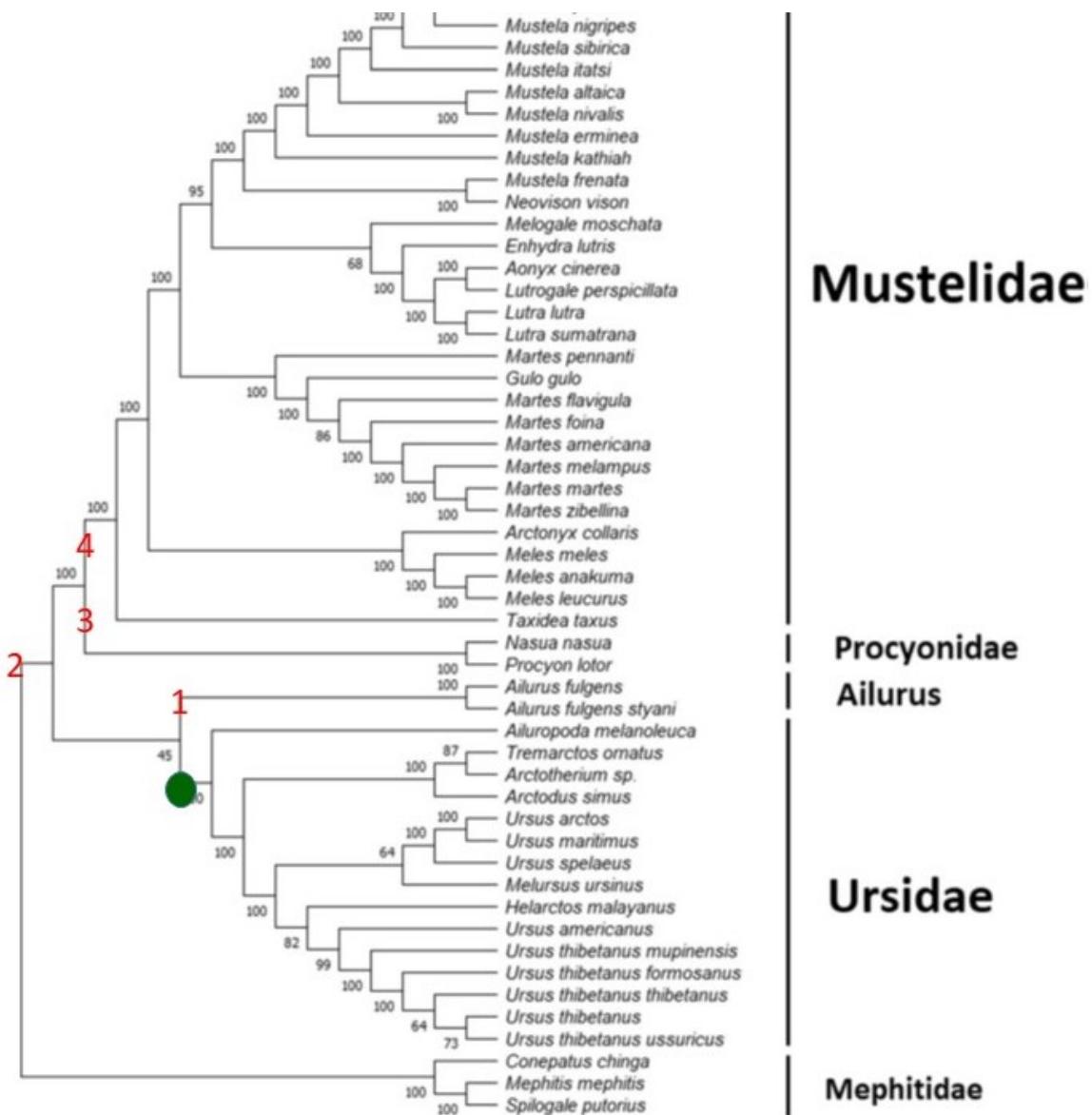
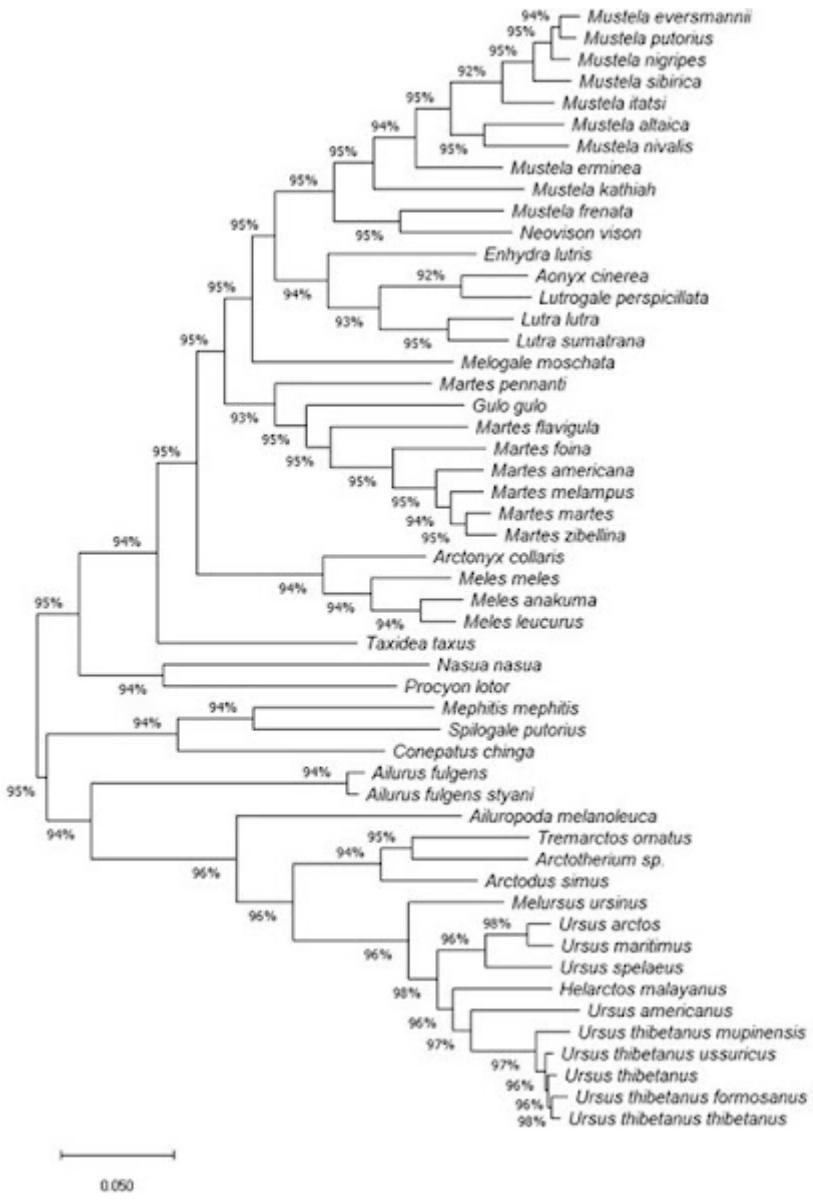


Figure 4: Figure 5 Cserhati BMC Genomics rooted between bear family and superfamily Musteloidea; successive divisions in the Musteloidea numbered.

In blog post 2, a list of articles is presented that found this classification of the Musteloidea. The outcome of the NJ phylogenetic tree of Cserhati is part of this. Cserhati is misled by the layout that comes out of his default program.



Figure 5. Classification of the Musteloidea in Figure 4 with bears as outgroup.



Mustelidae

Procyonidae

Mephitidae

Ailurus

Ursidae

Figure 6. Figure 6 from Cserháti BMC Genomics. Maximum Likelihood phylogenetic tree. The horizontal distance between successive splits now represents genetic distance

Cserháti's description of his result now indicates that the red panda may form a clade of its own:

Here Mustelidae, Procyonidae, and Mephitidae all form their own clades with a likelihood value of at least 94%. As opposed to the NJ tree, here Ailurus is separated from Ursidae suggesting that it might form its own clade as well.

The two red panda species form their own clade, also with a likelihood values of 94% (Figure 6).

The topology of the bears and the red panda is the same in Cserhati's NJ tree as in Cserhati's ML tree: the difference is that now the genetic distance is visible. Cserhati's "opposed" and "separated" is based on insufficient understanding of layout versus topology of the phylogenetic trees.

Of course we can consider the bears again as an outgroup to the Musteloidea, and place the root of the phylogenetic tree between bears and Musteloidea.

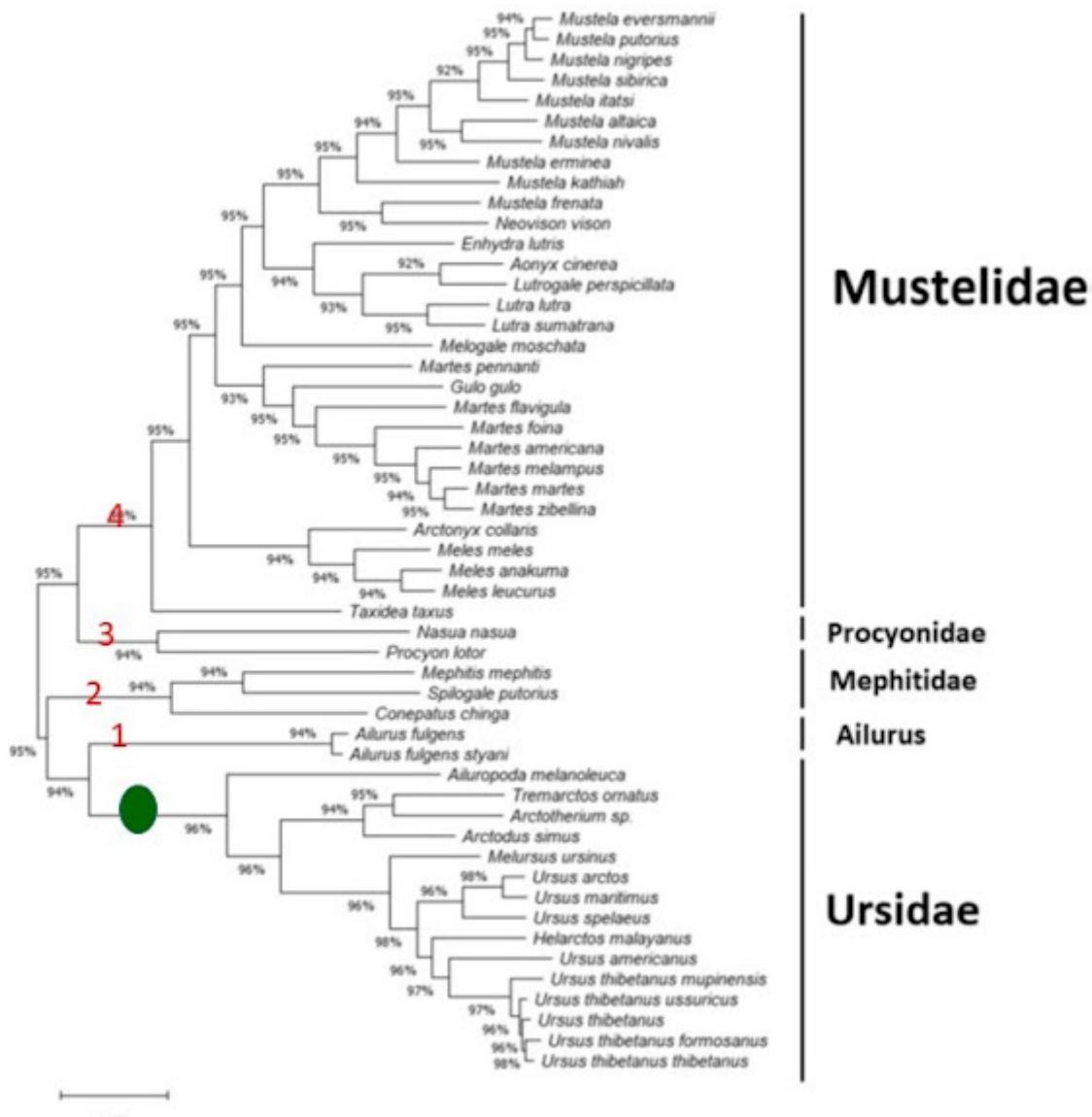


Figure 7. Figure 6 Cserháti BMC Genomics. Maximum Likelihood phylogenetic tree rooted between bear family and superfamily Musteloidea; successive divisions in the Musteloidea numbered.

Following the lines of the ML phylogenetic tree to find the successive splits in the Musteloidea yields the same classification as the NJ phylogenetic tree, namely the one in Flynn et al (2000).

Cserhati provides three phylogenetic trees on mtDNA of 15 bears and 37 musteloids. In all three trees the red panda forms its own group, a clade, the family Ailuridae. The red panda as a separate family was proposed in 1996 by Ledje and Arnason; since then, the red panda's status as a family Ailuridae has not been debated (except by Cserhati).

UPGMA gives a different phylogenetic tree, with the red panda Ailuridae and not the raccoon family Procyonidae as sister group of Mustelidae. The NJ phylogenetic tree and the ML phylogenetic tree give the Procyonidae as a sister group to the Mustelidae, as commonly found. The NJ and ML trees also provide a classification within the Mustelidae, in the placement of otters, weasels, badgers and martens, that is consistent with results elsewhere in the literature.

Cserhati concludes:

*The mtDNA results as well as the maximum likelihood tree appear to place *Ailurus fulgens* into a monophyletic group.*

???

Perhaps: The clustering of the 52 species on their mtDNA and all three phylogenetic analyses of those 52 species on their mtDNA give the family Ailuridae, the red panda, as its own monophyletic group.

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

Ledje C, & Arnason U. (1996) Phylogenetic relationships within caniform carnivores based on analyses of the mitochondrial 12S rRNA gene. Journal of Molecular Evolution 43:641–649.

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-13-het-bmc.html>

THE RED PANDA AND CSERHATI (13): THE BMC GENOMICS ARTICLE

In the BMC Genomics article, Cserhati uses two datasets and two types of methods to find the best classification, the phylogenetic placement, of the red panda. The natural classification of the species represents their relationship, and relationships between species and families are always interesting.

The two types of methods used by Cserhati are clustering and phylogenetic methods. The two data sets are Whole Genome K-mer Signatures of 28 species, and mtDNA sequences of 52 species.

1 Clustering

Cserhati is looking for the "phylogenetic placement" of the red panda (as he says in the second sentence of his abstract) and then clustering is not a useful method. Clustering is not a phylogenetic method and does not provide a phylogeny, and therefore no "phylogenetic placement". Clustering cannot be used to conclude monophyly of a group or groups. Cserhati says in Methods: "...*using clustering algorithms to detect monophyletic groups*", which is wrong, as clustering is not able to detect monophyly.

The discussion of the results of the WGKS data states:

Ailurus fulgens clearly clusters together with the mustelids,

Based on this evidence ["*Ailurus fulgens* clearly clusters together with the mustelid"], *Ailurus fulgens* would belong to mustelids as a monophyletic group.

The first sentence is not controversial but not interesting for the classification of the red panda *Ailurus fulgens*. The second sentence is wrong in two ways. First, monophyly cannot be derived from clustering. Second, "monophyly" and "belonging to the same family" are not the same thing.

2 Phylogenetic methods on the two datasets.

2.1 WGKS

One phylogenetic analysis method, UPGMA, was applied to the WGKS data set with 28 species.

The UPGMA phylogenetic tree mainly shows that WGKS is not very useful for phylogeny (BMC Genomics Figure 2). The very large differences, those between the cat family, bear family and the superfamily Musteloidea, are found correctly. The split found within the Musteloidea is among the possibilities, given all other studies on the red panda. Within the marten family, the weasels and the otters are intermingled. Within the cat family the 'big cats' and the 'little cats' are intermingled. For the more related species, the UPGMA analyses on WGKS gives garbage. In the Abstract, Cserhati writes:

(WGKS) Being a genomics-based algorithm, it also reduces stochastic error to a minimum.

Judging by his results in figure 2, this is not correct. Stochastic error is high.

2.1 mtDNA

On the 52-species mtDNA dataset, Cserhati applied three phylogenetic methods.

The UPGMA method gives a placement of marten family and red panda as sister group; marten family Mustelidae and Ailuridae as sistergroups are not found anywhere in the literature, in the presence of species of the raccoon family.

Cserhati interprets the phylogenetic trees from the NJ analysis and the ML analysis according to the layout, showing a total lack of understanding of phylogenetic trees. Surely it should have alarmed Cserhati that the superfamily Musteloidea did not appear in his interpretation.

With the bear family as outgroup both NJ and ML analyses yield the same pattern within the superfamily Musteloidea: the first split is that of red panda Ailuridae against skunks + raccoons + marten family, the second split is that of skunks against raccoons + marten family. The raccoons and the marten family appear as sister groups. This is a well-known relationship pattern within the superfamily Musteloidea, found in a number of scientific papers.

Cserhati appears sloppy in his presentation of the results for the mtDNA dataset. The Abstract states:

A Maximum Likelihood tree suggests that A. fulgens and Ursidae form a monophyletic group, although the bootstrap value is weak.

Ailurus fulgens and Ursidae a monophyletic group is Cserhati's interpretation of the NJ phylogenetic tree, not the ML tree. He mixes up his own results.

3 It is striking that in Cserhati's analysis of mtDNA the word 'outgroup' never appears.

Even more striking is that the idea 'sistergroup' does not occur at all. Two sistergroups are monophyletic. Where Cserhati wants to use 'monophyly' to argue that the red panda belongs to the marten family Mustelidae, that monophyly shows a sistergroup relationship between the families Ailuridae and Mustelidae. Monophyly does not mean belonging to the same family.

4 In the Abstract Cserhati says:

Conclusions: The main conclusion that we can draw from this study is that on a whole genome level Ailurus fulgens possibly belongs to the mustelid clade,

This has no basis. It derives from the WGKS clustering, but 'cluster' and 'clade' are confused.

5 The introduction is teeming with errors. Almost all articles on phylogeny are cited incorrectly.

6 As I wrote on Panda's Thumb :

"One problem: how did a paper as bad as this ever get through review and published?"

Cserhati, M., 2021, A tail of two pandas – whole genome k-mer signature analysis of the red panda (*Ailurus fulgens*) and the Giant panda (*Ailuropoda melanoleuca*), BMC Genomics 22: 228

<http://pandasthumb.org/archives/2022/12/a-tale-of-two-papers.html>

<https://creationismeweersproken.blogspot.com/2023/01/de-rode-panda-en-cserhati-14-de.html>

THE RED PANDA AND CSERHATI (14): THE INTRODUCTION OF THE CRSQ ARTICLE

The article by Matthew Cserhati in Creation Research Society Quarterly (2021) has the title: '*Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis*'.



Figuur 1. De rode panda *Ailurus fulgens*

Creation Research Society and Baraminology: this is a different world from BMC Genomics.

1 What says Genesis?

The introduction of the CRSQ article has as its first sentence:

Genesis 1:20–25 describes how God created fish, birds, and land-animals each according to their kinds.

Fish, the great whale, and birds are mentioned in Genesis 1: 20-23, the land animals in Genesis 1:24-25:

25 And God made the beast of the earth after his kind, and cattle after their kind, and every thing that creepeth upon the earth after his kind: and God saw that it was good. (King James Version)

Creationists have their own interpretation of the Hebrew word translated as 'kind', 'after his kind'. That Hebrew word is *mîn*, and it occurs almost exclusively in this kind of context. The word is a collective, and indicates diversity:

Thus the biblical text emphasizes the diversity of life – plants and animals – with which God filled the sky, the sea, and the dry land he had created.

Consistent with the basic message of Genesis 1, the emphasis rests upon God's creation of life in all its abundance and diversity. (*Biologos*)

The point is variety, diversity: 'as they are', 'as they live', as herbivores, predators, frugivores, climbers and diggers. According to their nature. 'Kind' is a word with a broad meaning in non-creationist English.

2 'Kind' and cluster

The first sentence of the CRSQ article paraphrases Genesis.

The next two sentences are:

Species within a kind were originally capable of breeding with one another but are incapable of breeding with species from another kind. Hence, there is continuity between species within a kind, and discontinuity between two separate kinds.

In the CRSQ article we are dealing with this creationist extension of Genesis. The indeterminate word 'kind' is developed into a central technical concept:

In technical terms, a kind is called a 'baramin, ... Thus, molecular baraminology is the study of the created kinds from a molecular biology perspective. A 'holobaramin' denotes all species which constitute a single baramin.

Followed by:

God probably created holobaramins with only one single species, but some kinds diversified more than others over time.

None of this is in Genesis, but it is standard creationism. It is so standard that it seems that creationists think Genesis says: God created 'kinds'.

The question for creationists is how to know which species belong in a 'kind', 'baramin' or 'holobaramin'. The idea is to look for a group that is continuous among members of the group, and discontinuous with other groups:

Hence, there is continuity between species within a kind, and discontinuity between two separate kinds.

Hennigan (2009) has a glossary of creationist terminology:

Continuity – Biologically meaningful similarity between organisms that embrace all types of biological characteristics.

Discontinuity – Biologically meaningful differences between organisms.

This leaves 'biologically meaningful' to be specified.

From the definition of 'kind', species must be related within a 'kind'. Continuity in traits must go back to common descent within a 'kind'. Discontinuity seems to imply unrelated.

Therefore:

- Baramin were created independently, and species in the same baramin are related
- Baramin were created independently, and species in different baramin are unrelated

The two questions for creationists are:

- How do you show that there is relationship within a baramin?
- How do you show that there is no relationship between baramin?

Proving no relationship exists is not so easy. Take sheep for example. The species *Ovis aries* is the domestic sheep, all breeds, and belongs to livestock, cattle. Cattle and wild animals were created separately (Genesis 1:25), what implies that domestic sheep must be placed in a different baramin than wild sheep. One species of wild sheep, the Asian mouflon, is regarded as the ancestor of the domesticated species. Apart from the domestic sheep species and Asian mouflon, the genus *Ovis* contains quite a few sheep species; and all kinds of species in the subfamily Caprinae are also called 'sheep'. Quite similar sheep species, but wild animals, not cattle. How do you show biologically that the domestic sheep belongs in another baramin, as it is cattle?

Creationists prefer to use clustering to look for biological groups: statistically, species in a cluster are more similar than species in different clusters.

Similarity within a cluster is a different idea of "continuous" than "biologically meaningful" or relatedness.

Clustering tells you nothing about relatedness and little about biology. Nevertheless, for creationists cluster and continuity and relatedness coincide without further discussion; and being in different clusters and discontinuity and unrelated coincide without further discussion. Here is a great leap across the gap between statistics and biology.

3 Classification red panda

Cserhati pretends ambiguity about the placement of the red panda in the (evolutionary biological) scientific literature to be a good reason to look at the classification of the red panda. Cserhati mentions classification of the red panda with the raccoons, with the giant panda, as one's own family or in the vicinity of the raccoons, skunks and martens as possibilities. Much of what Cserhati writes in the introduction to the CRSQ article is also included in the introduction to the BMC Genomics article. I discussed that introduction earlier (blogpost 5).

In the CRSQ article, Cserhati elaborates on the placement of the red panda among the raccoon family, a possibility mentioned in 1995 by Slattery & O'Brien and in 1997 by Dragoo & Honeycutt. The study by Slattery & O'Brien (1995) compared the giant panda and four other species of bears, and four species of the raccoon family, to the red panda using different outgroups. The red panda is placed within the raccoon family (2x) or as a sister group of the raccoon family. Dragoo & Honeycutt (1997) provide a phylogenetic tree of their own data, without the red panda, combined with literature data with the red panda, and find weak evidence for red panda placement in the raccoon family. Both studies are old, and not very convincing.

Recent molecular phylogenetic work on the classification of the Carnivora and the red panda is not mentioned in this CRSQ article, as it was not in the BMC Genomics article.

Cserhati indicates that no previous creationist work focused on the red panda. The red panda only appears in the general baraminology study on mammals by Thompson & Wood (2018). The data used by Thompson & Wood come from a study of the skull and teeth of the raccoon family and the red panda, with two skunk species and a marten as an outgroup (Ahrens 2012). Ahrens (2012) preferred the red panda within the raccoon family Procyonidae, on grounds of dentition. However, Thompson & Wood's analysis places the red panda outside the Procyonidae

Cserhati, M., 2021, Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis, Creation Research Society Quarterly 58 (2): 76-84

<https://www.creationresearch.org/classification-of-the-enigmatic-red-panda-ailurus-fulgens-based-on-molecular-baraminology-based-analysis>

<https://biologos.org/articles/the-meaning-of-min-in-the-hebrew-old-testament>

Hennigan, T. 2010. The case for holobaraminic status in bears (family Ursidae) and the implications within a creation model of ecology. CRSQ 46(4):271–283.

Chessa et al, 2009. Revealing the history of sheep domestication using retrovirus integrations. Science 324:532-536

Slattery, J.P., and S.J. O'Brien. 1995. Molecular phylogeny of the red panda (*Ailurus fulgens*). The Journal of Heredity 86(6):413–422.

Dragoo, J.W., and R.L. Honeycutt. 1997. Systematics of mustelid-like carnivores. Journal of Mammalogy 78(2):426–441.

Thompson, C., and T.C. Wood. 2018. A survey of Cenozoic mammal baramins. In Proceedings of the Eighth International Conference on Creationism, J.H. Whitmore (editor), pp. 217–221, A1-A83 (appendix). Creation Science Fellowship, Pittsburgh, PA.

Ahrens, H. 2012. Craniodontal characters and the relationships of Procyonidae (Mammalia: Carnivora). Zoological Journal of the Linnean Society 164: 669–713.

<https://creationismeweersproken.blogspot.com/2023/02/de-rode-panda-en-cserhati-15-de-data-en.html>

THE RED PANDA AND CSERHATI (15): DATA AND RESULTS IN THE CRSQ ARTICLE

In his Creation Research Society Quarterly article, Cserhati uses three types of data to find the red panda's place among other species: Whole Genome K-mer Signature, mitochondrial DNA sequence, and amino acid sequence in the protein cytochrome-b.

The first two sets of data, WGKS and mtDNA, are the same as in his BMC Genomics article. For the cytochrome-b amino acid sequence, Cserhati uses new data. Data on the following numbers of species/subspecies are used:

family	WGKS	mtDNA	aa cytochrome-b
cats Felidae	11	-	25
bears Ursidae	5	15	5
Martens Mustelidae	10	30	17
raccoons Procyonidae	0	2	1
skunks Mephitidae	1	3	1
red panda Ailuridae	1	2	2
	28	52	51

The raccoon family is absent from the WGKS data, although Cserhati could have collected data for two species. The paper by Tsuchiya et al in *Genome Biology and Evolution* detailing the kinkajou and raccoon genomes was published in January 2021. This would have been in time for use in the CRSQ article in Fall 2021.

WGKS

From the WGKS data, Cserhati calculates a correlation matrix between the species. This is the same correlation matrix as in the BMC Genomics article, as

evidenced by the two Additional Files. Cserhati provides heatmaps of the correlation matrix in both articles: the same figure, but colored differently.

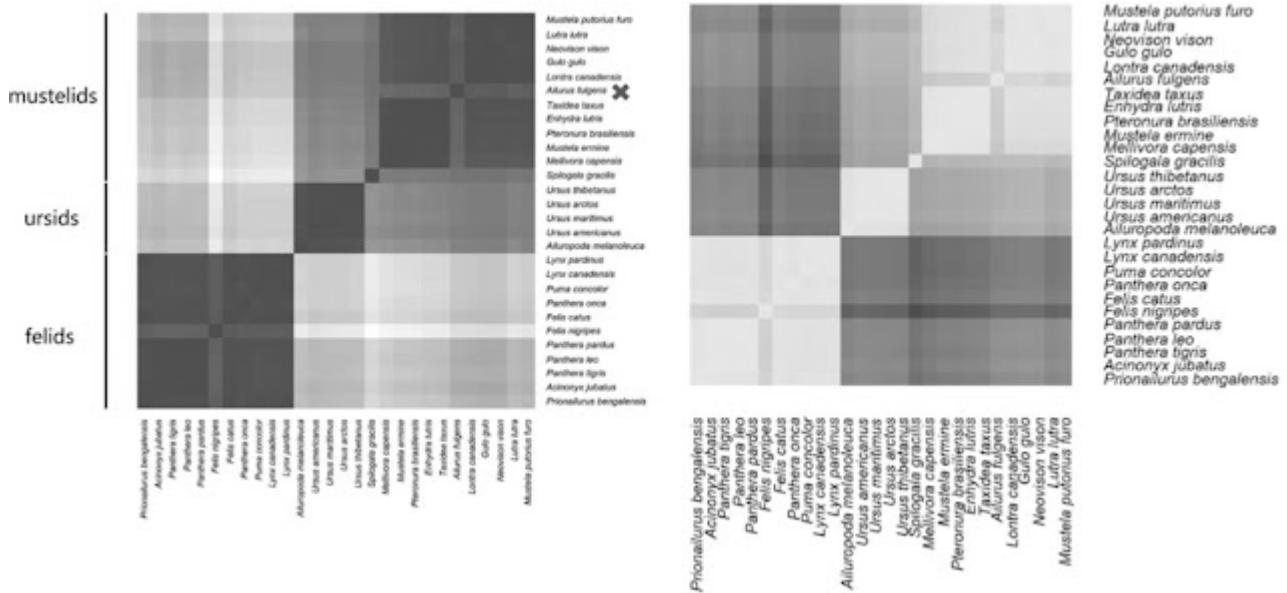


Figure 1. Heatmaps of the correlations on the WGKS data: left as in CRSQ article, right as in BMC Genomics article. The gray scale is reversed, otherwise it is the same figure.

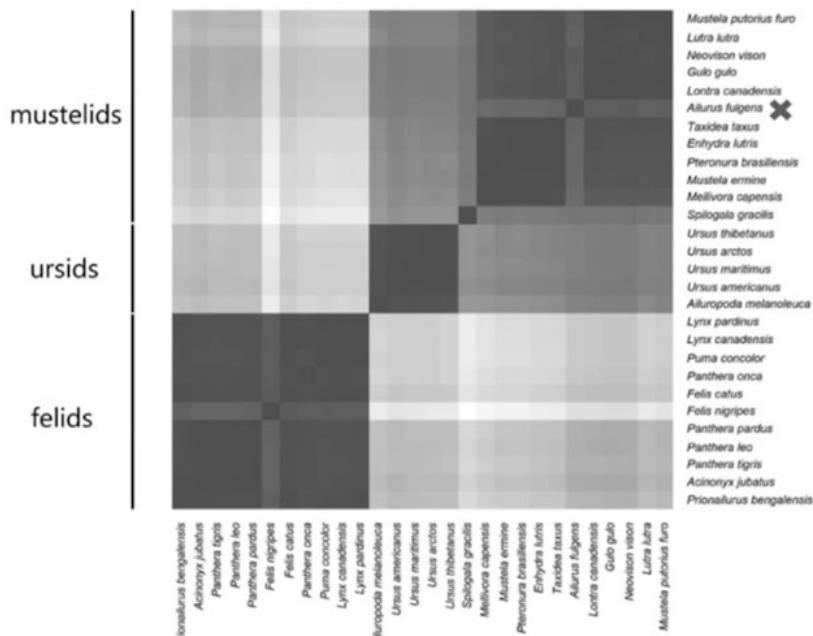


Figure 2: Heatmap indicating the size of the correlations in WGKS between the species. Darker is more similar. The order of species on the axes is the same from top to bottom as from right to left; the southwest-northeast diagonal gives correlation 1, the correlation of the species with itself. The families are indicated on the left: from top to bottom marten family, bear family, cat family. Note that here both the western spotted skunk *Spilogale gracilis* and the red panda *Ailurus fulgens* are reckoned to be mustelids, to belong to the marten family Mustelidae rather than to the superfamily Musteloidea. Figure 2 CRSQ 2021.

The location of the red panda *Ailurus fulgens* is indicated on the right, in the middle of the mustelids. This leads to a kind of 'Danish flag' pattern. The only species of the skunk family Mephitidae in these dates, *Spilogale gracilis*, can be found between the bears and the martens, the block next to the 'Danish flag'.

In this figure, Cserhati indicates families: felids, the cat family Felidae, ursids, the bear family Ursidae, and mustelids, the marten family Mustelidae.

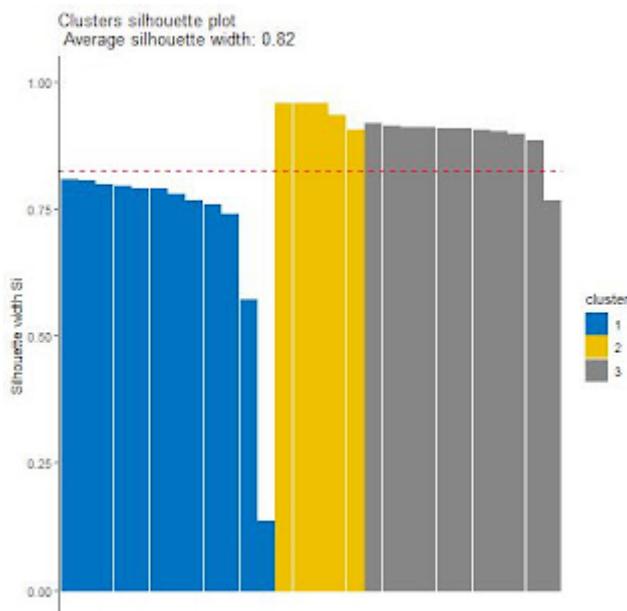
Apparently, the red panda and *Spilogale* are counted in the marten family in the figure.

In the text, Cserhati however says that the three clusters found in the WGKS data are the cats, the bears and the superfamily Musteloidea:

... felids covering the family Felidae (eleven species), ursids covering the family Ursidae (five species), and musteloids, a superfamily including Mephitidae and Mustelidae (twelve species).

Two points about Cserhati's 'three clusters' result deserve notice.

The first point of interest is that Cserhati arrives at three clusters in this CRSQ article, but arrives at four clusters with exactly the same results in the BMC Genomics article . In the BMC Genomics article, the skunk *Spilogale gracilis* is set apart as the fourth cluster. Here in the CRSQ article, *Spilogale gracilis* is included in the large cluster Musteloidea. Cserhati gives silhouette plots for 3 clusters and for 4 clusters: four clusters if *Spilogale* is held to be a separate cluster. The zero silhouette width for *Spilogale* indicates that clustering in the adjacent cluster is better. Therefore, the silhouette plots show better clustering with three clusters.



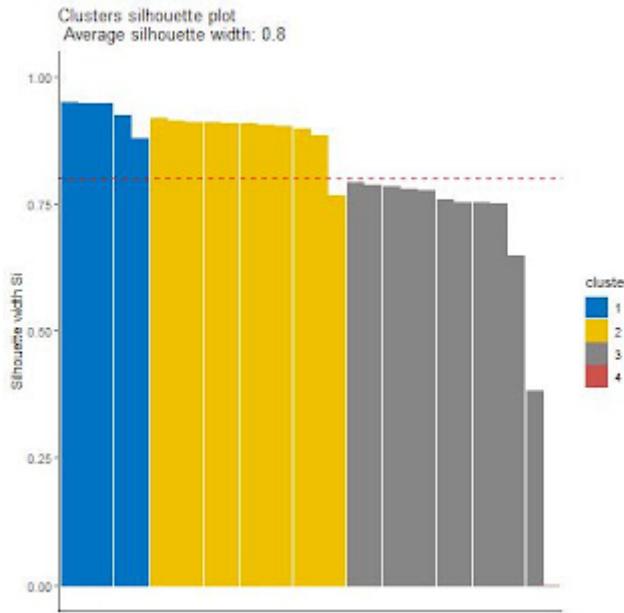


Figure 3. Silhouette plots for 3 clusters (above, CRSQ supplementary figure 1) and for 4 clusters (below CRSQ supplementary figure 2). Cserhati gives no description for the x-axis. The five species of the bear family are yellow above, blue below. The eleven species of the cat family above gray and below yellow. The twelve species of the superfamily Musteloidea are blue above and grey/red below.

The second point of interest is that Cserhati classifies the red panda in the marten family Mustelidae - something he should have shown, but here assumes without any justification. The red panda is different from all mustelid species: as can clearly be seen in the heatmap

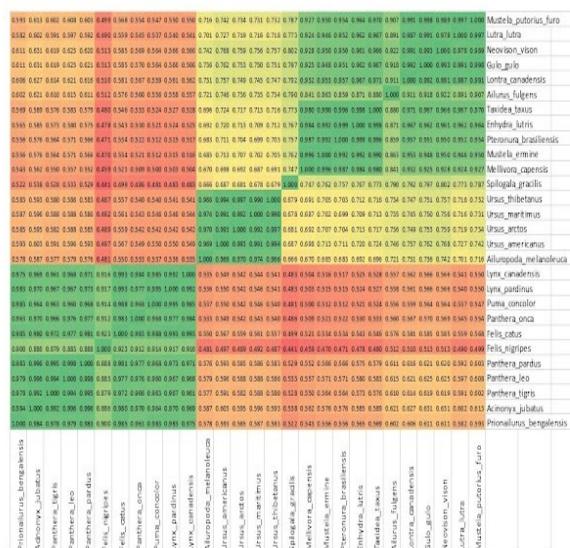


Figure 4. Colored WGKS heatmap from CRSQ paper: greener is higher correlation, redder is lower correlation. Minimum Correlation: .44. This heatmap has exactly the same correlations as black and white Figure 2 here.

Figures 2 and 4 clearly show the major difference in the correlations is between the cats on the bottom left and all other species. To clarify differences within the superfamily Musteloidea, a heatmap with just the correlations within the Musteloidea is called for.

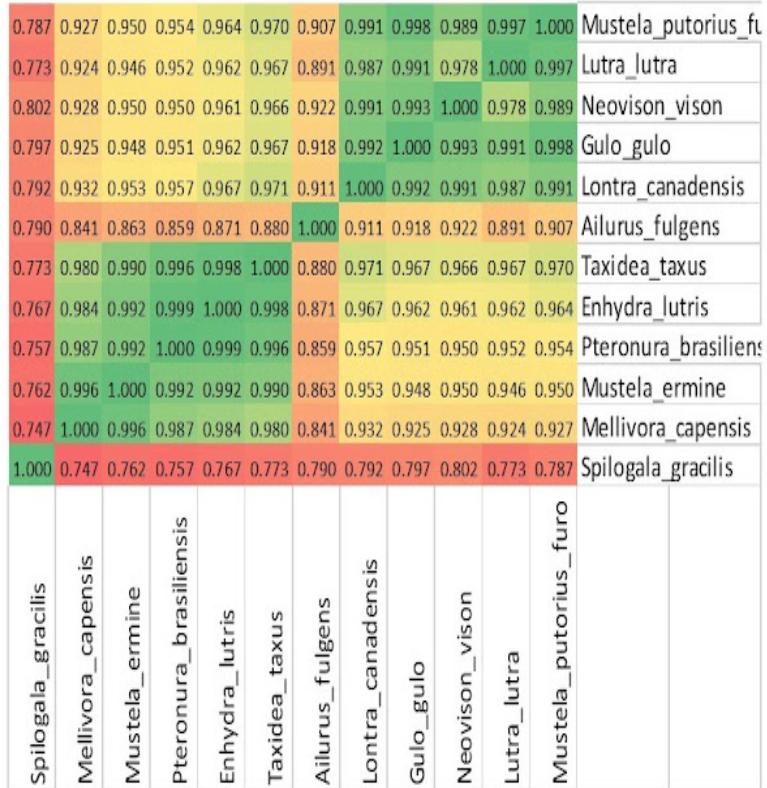


Figure 5. Heatmap of the Musteloidea, no cats and bears. The order of the species is the same as in Figure 2 and Figure 4. Now the lower correlation between the red panda *Ailurus fulgens* and the species of the marten family is clearly visible.

The difference is statistically significant (blog post 9). The red panda may cluster within the superfamily Musteloidea, but it is not belong to the family Mustelidae.

mtDNA sequences

Cserhati calculates a correlation matrix between the species from the mitochondrial DNA sequence data. This is the same correlation matrix as in the *BMC Genomics* article. Cserhati provides a heat map of the correlation matrix: the same heatmap as in *BMC Genomics*, but colored differently and with the species in different order.

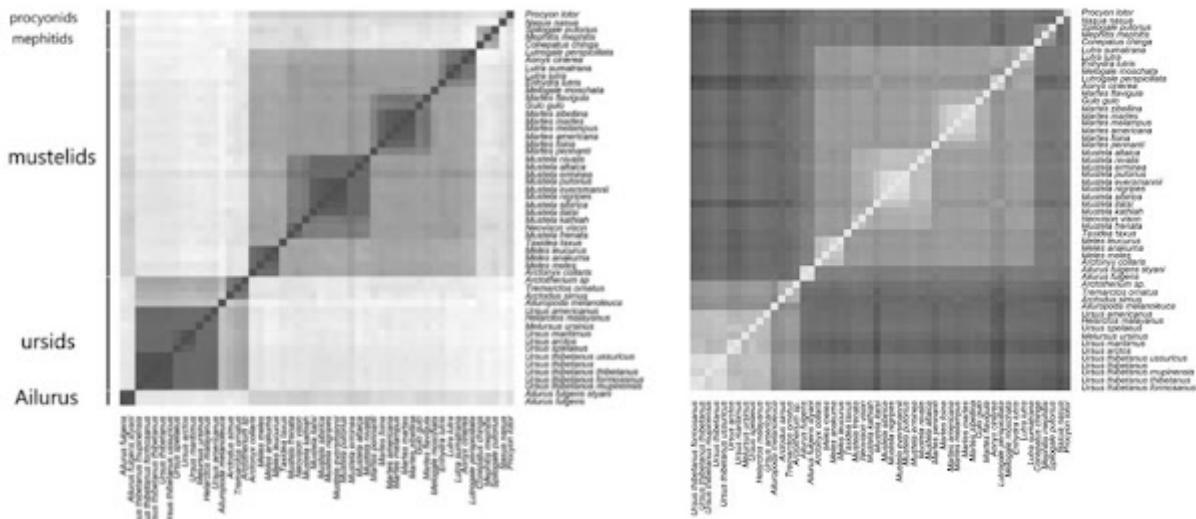


Figure 6. Heatmaps of the correlations on the mtDNA data: left as in CRSQ article, right as in BMC Genomics article. The gray scale is reversed. The two subspecies of the red panda are placed at the very bottom left in the CRSQ article, and can be found in the BMC Genomics article as the small light block between the two larger light blocks.

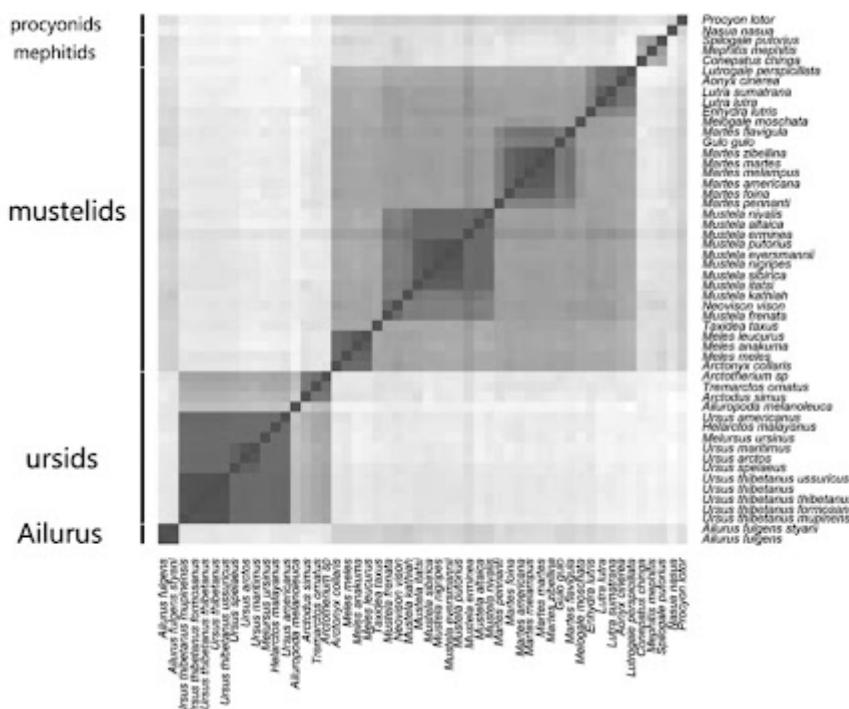


Figure 7:
Heatmap of the correlations in mtDNA sequence between species. Darker is more similar. The order of species on the axes is the same from top to bottom as from right to left; on the diagonal southwest -northeast is the correlation with the species itself. The two red panda subspecies are at the bottom left. Figure 3 CRSQ 2021

Cserhati comments:

In this heatmap we can see two main, large clusters as well as three smaller ones.

The bears, ursids, and the marten family, mustelids, will presumably be meant by 'the two large clusters'. The interpretation with two large clusters and three small clusters is made more intuitive by the placement of the two red panda subspecies in the lower left (*Ailurus*) and the placement of the two raccoons (procyonids) and the three skunks (mephitids) in the upper right. As a consequence, it is difficult to see the correlations of the red panda with the mustelids, mephitids and procyonids: light gray along the lower edge, indicating the two red panda subspecies are not well placed in this figure.

The correlations in the heatmap can be found in Additonal File 2 of the article. I've changed the order of the families along the axes, and brightened up the correlation matrix for a clearer red-green heatmap.

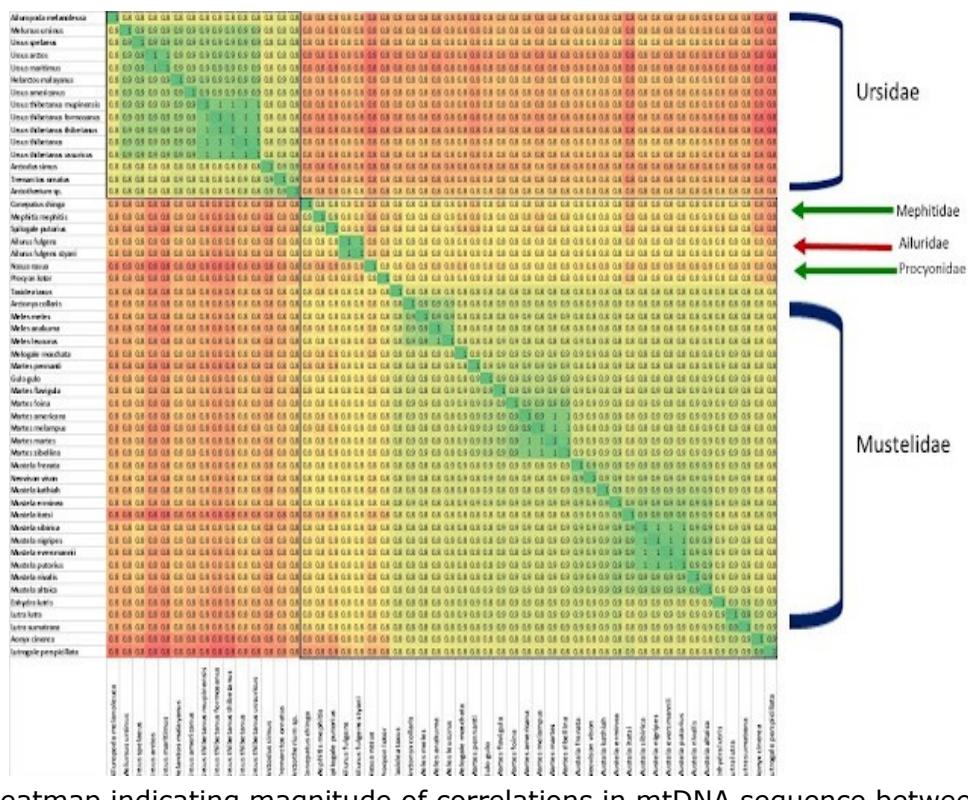


Figure 8. Heatmap indicating magnitude of correlations in mtDNA sequence between species. Greener is more similarity, redder is more difference. Minimum Correlation: .75. The order of species on the axes is the same from top to bottom as from right to left; on the diagonal northwest-southeast is the correlation with the species itself. Figure 3 CRSQ 2021 with reordering of the families and in different colors.

The resulting heatmap provides a clearer view. There are two main groups, one with the bear family Ursidae on the top left and one with the superfamily Musteloidea on the bottom right (including the families Mephitidae skunks,

Ailuridae red panda, and Procyonidae raccoons). Within the Musteloidea more variation is found than within the bears. The Musteloidea show families and subfamilies.

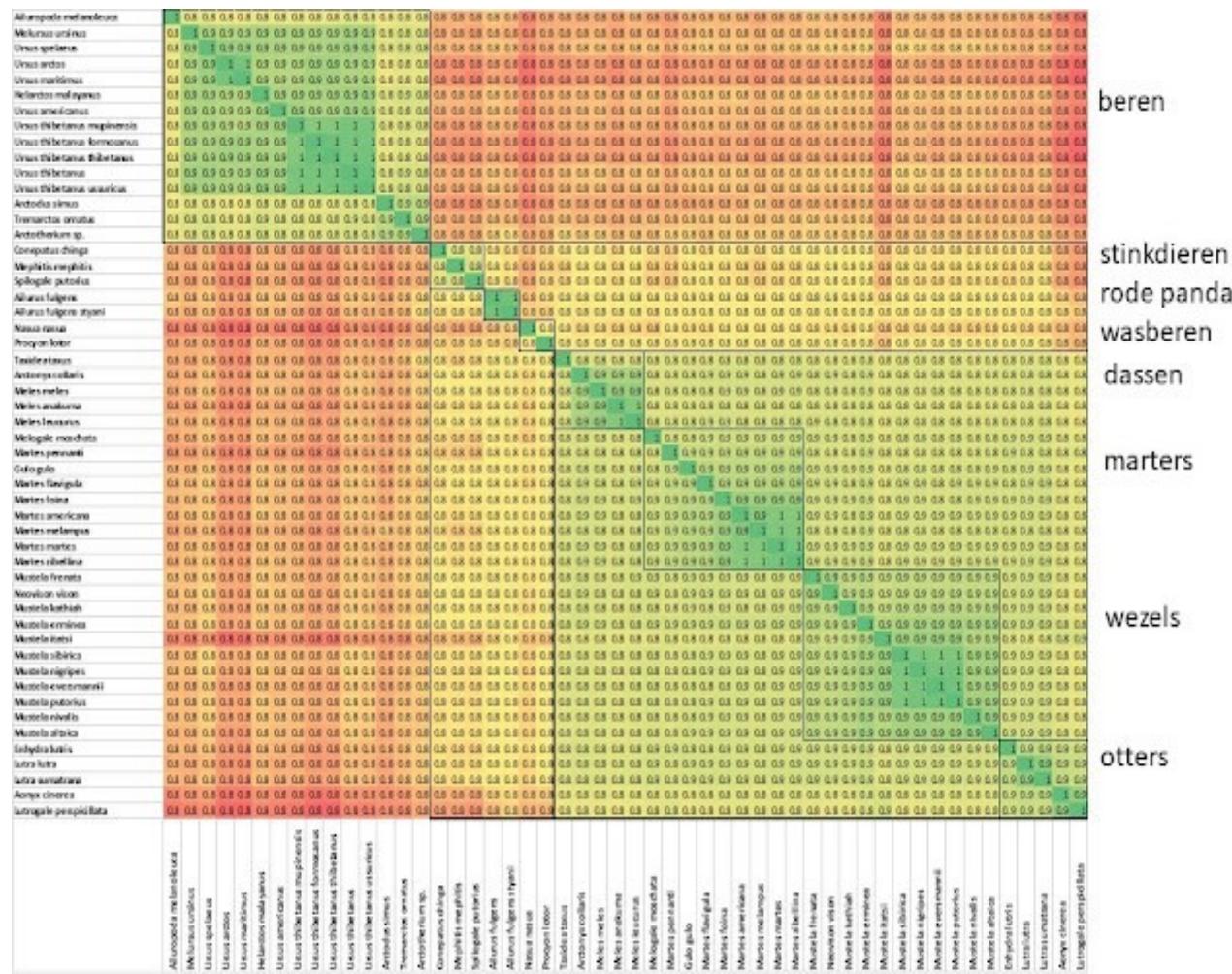


Figure 9. Heatmap from figure 8 with families and subfamilies within the Musteloidea indicated. Beren = bears; stinkdieren = skunks; rode panda = red panda; wasberen = raccoons; marters = martens; wezels = weasels; otters = otters.

Some differences show up between the mustelid family Mustelidae and the other three families in the superfamily Musteloidea. It's a little more orange in those lanes than in the square. The differences can be put into focus.

A heat map of the superfamily Musteloidea on its own emphasizes the differences. In such a heatmap, the major differences between bears and Musteloidea have disappeared. The differences within the Musteloidea become visible with continued heatmap magnification and focussing.

Figure 10 Identical correlation matrix as in the figures 7, 8 en 9 emphasizing the differences in the correlations within the Musteloidea. Beren = bears; stinkdieren = skunks; rode panda = red panda; wasberen = raccoons; martens = martens; wezels = weasels; otters = otters.

Figure 12 Same correlationmatrix as in figures 7, 8, 9 and 10, emphasizing differences within the family Mustelidae.

Would the otters stand a bit apart? After all, otters are biologically meaningfully different from the other species here - swimmers and piscivores.

Cserhati interprets figure 7 (his figure 3) as if the five known taxonomic families appear from the clustering.

In this heatmap we can see two main, large clusters as well as three smaller ones.

There is no statistical underpinning of five clusters. Cserhati says five clusters, but then a silhouette plot would be expected that makes that clear. The silhouette plots given by Cserhati show something else.

The mtDNA part of the BMC Genomics article contains a curious stand-alone sentence among its description of the phylogenetic trees:

Supplementary figure 3 shows the average silhouette width according to the number of clusters, with an average silhouette width of 0.51 for two clusters.

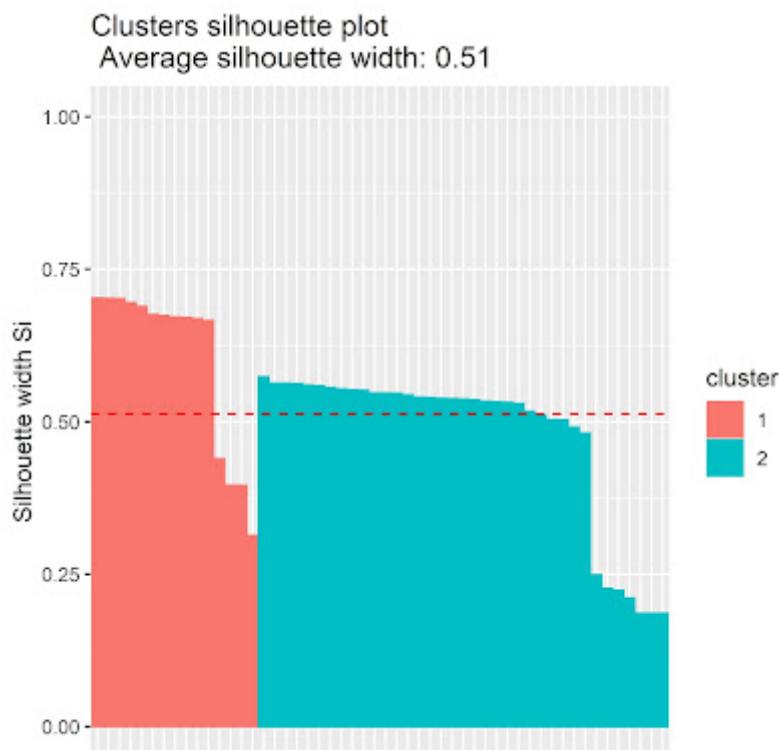


Figure 13. Supplementary figure 3 from the BMC Genomics article, silhouette plot with two clusters Left ursids, right Musteloidea. In this silhouette plot the 30 species of the marten family and the 7 species of Musteloidea that do not belong to the marten family can be recognized.

Cserhati has made silhouette plots for different numbers of clusters, as can be seen in Supplementary figure 3 of the CRSQ article.

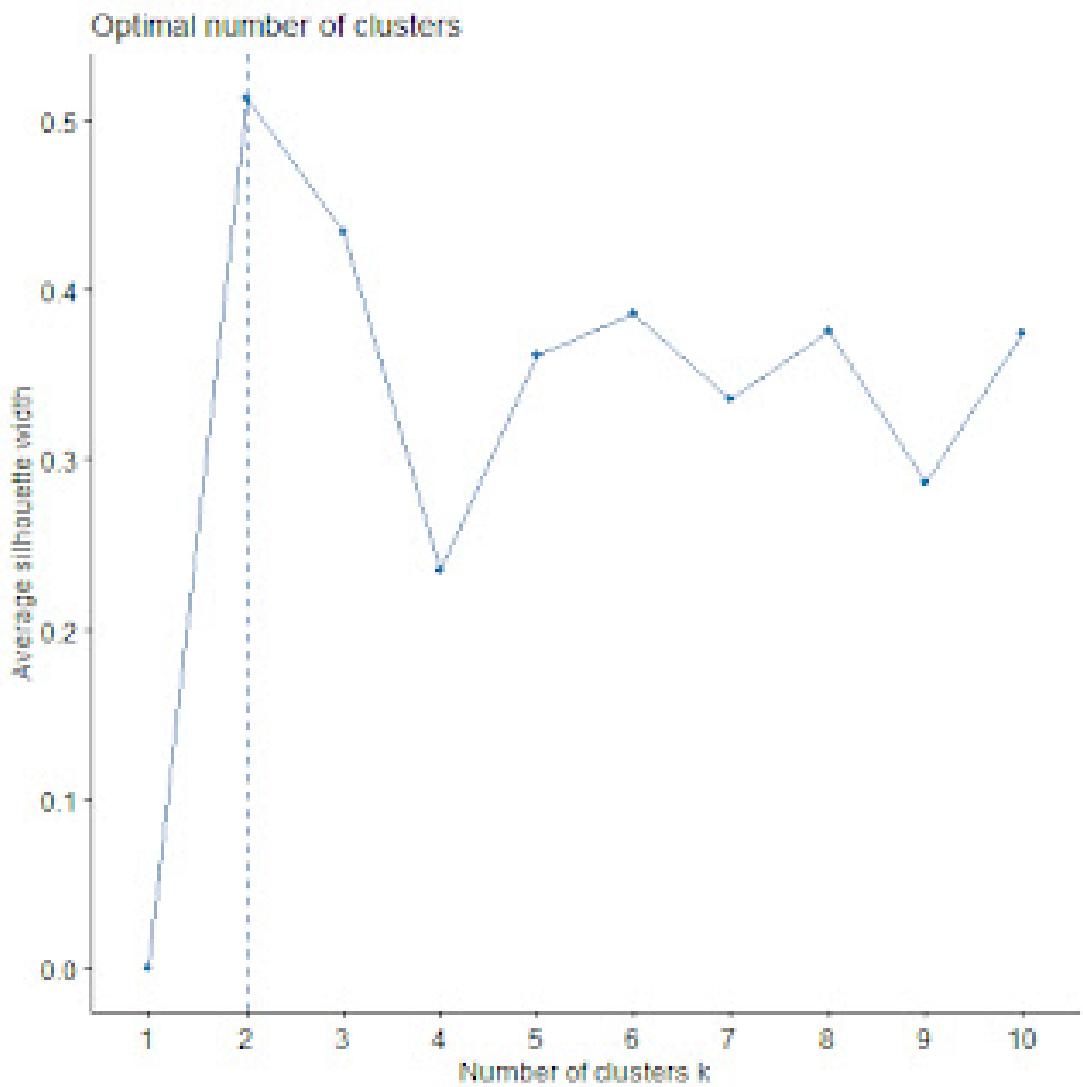


Figure 14. Supplementary figure 3 van het CRSQ artikel shows the average silhouette width according to the number of clusters, with an average silhouette width of 0.51 for two clusters as maximum. The optimal number of clusters is 2.

In fact, Cserhati finds two major clusters, namely the bear family Ursidae and the superfamily Musteloidea. He also indicates this in his table with the statistical data of the clusters, Table 3 of the BMC Genomics article.

Five clusters, the five classical families Ursidae bears, Mephitidae skunks, Ailuridae red panda, Procyonidae raccoons, Mustelidae mustelids is not better than eight clusters. Eight clusters with the four families Ursidae bears, Mephitidae skunks, Ailuridae red panda, Procyonidae raccoons and four subfamilies, Melinae badgers, Guloninae martens, Lutrinae otters and Mustelinae weasels within the marten family is also possible, as good as five clusters.

Amino acid sequence in the protein cytochrome b.

The gene for the protein cytochrome b is found on the mitochondrion. Cserhati decides to use the amino acid sequence of this protein for species clustering.

Cytochrome-b is a structurally conservative protein which does not mutate freely. It is often used to infer phylogenetic relationships between organisms (Meyer, 1994), and can so be used in baraminology studies as well.

Cytochrome b was one of the first genes for which good techniques and good data became available. We are talking about popularity 30 years ago, before DNA sequences of the entire mitochondrial DNA became available. Since the, cytochrome b has fallen out of favour (Meyer1994). Genbank and the NCBI database have thousands of entries; finding data is not difficult.

Cserhati compares the amino acid sequence in 25 cats, 5 bears (but not the giant panda), the two subspecies of the red panda, the striped skunk *Mephitis mephitis* alone in its family, the coati *Nasua nasua* representing the raccoon family, and 17 species of mustelids . The heatmap that Cserhati gives of the correlation matrix for these species is given in the following figure.

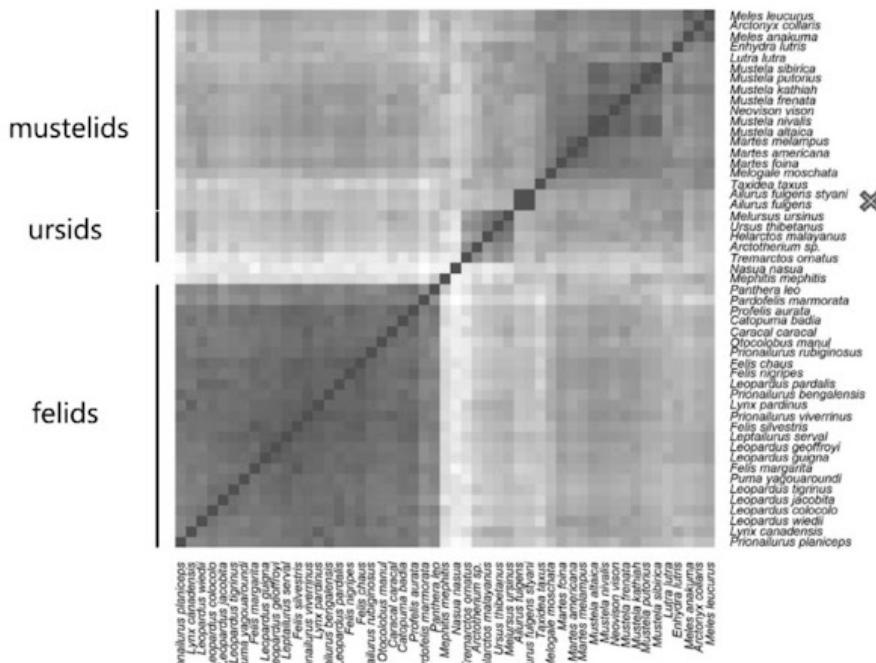


Figure 15: Heatmap indicating magnitude of correlations in cytochrome b amino acid sequence between species. Darker is more similar. The order of species on the axes is the same from top to bottom as from right to left; on the diagonal zw-no is the correlation with the species itself. The groups are indicated on the left. The two subspecies of the red panda are indicated on the right. The coati and the skunk are in this layout between the group of bears and the group of cats. Figure 4 CRSQ 2021.

Cserhati writes:

*In the heatmap (Figure 4) we can see three main clusters and two species by themselves, *N. nasua* and *M. mephitis*. In the lower left there is a tight cluster of 25 felids. In the upper right there is a cluster of five ursid species, and another with 19 mustelid species.*

No silhouette plots are provided to substantiate three clusters, of cats, bears and mustelids + pandas, and two non-clustered species. A silhouette plot should have made clear why no cluster Musteloidea is reported. Given the heatmap, a cluster of Musteloidea is a possibility.

The red panda is without a doubt counted among the mustelids without any arguing why. As can be seen from the heatmap (fig 15 here, fig 4 Cserhati) this is not obvious. The two subspecies of the red panda clearly form a block next to the species of the family Mustelidae; the correlations of the red panda with the mustelids are again lower than the correlations between the mustelids themselves, as can be seen from the gray values.

This becomes clearer in a red-green heatmap. The correlations can be found in Additional File 3 of the article. I've changed the order of the families along the axes, and brightened up the correlation matrix for a clearer heatmap.

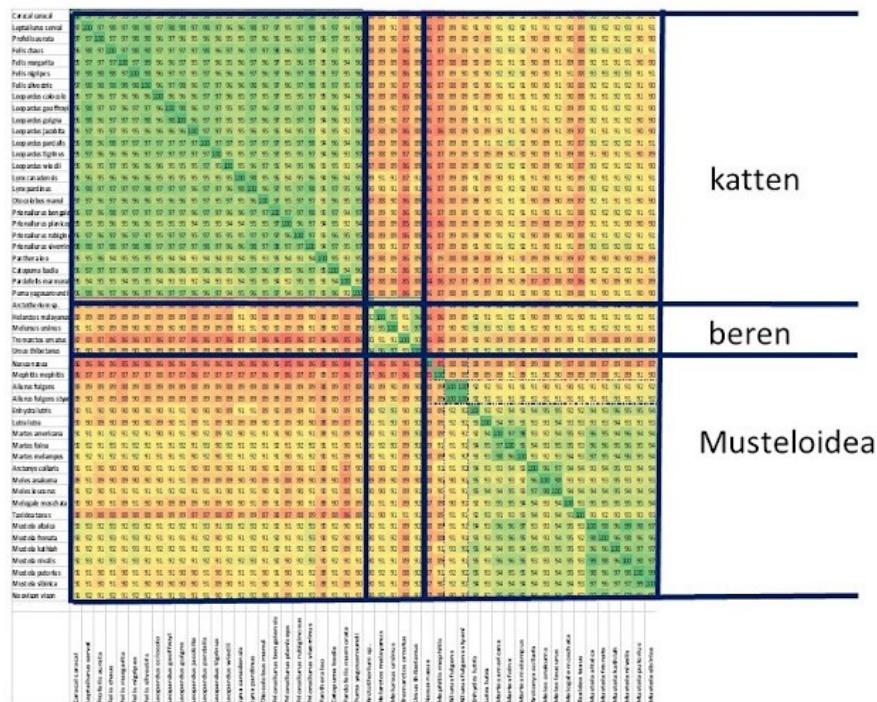


Figure 16. Heatmap indicating magnitude of correlations in cytochrome b amino acid sequence between species. Greener is more similar, redder is more different. Minimum Correlation: .84. The order of species on the axes is the same from top to bottom as from right to left; on the diagonal nw-zo is the correlation with the species itself. Figure 4 CRSQ 2021 with change of order. (katten = cats; beren = bears)

The cat group is homogeneous, the few bears are also homogeneous, and the superfamily Musteloidea mustelids, coati, skunks and red pandas appear to be quite heterogeneous. The coati is the topmost in the Musteloidea, followed by the skunk, and then the two subspecies of red pandas. Coati and skunks differ. Coloring only the superfamily Musteloidea.

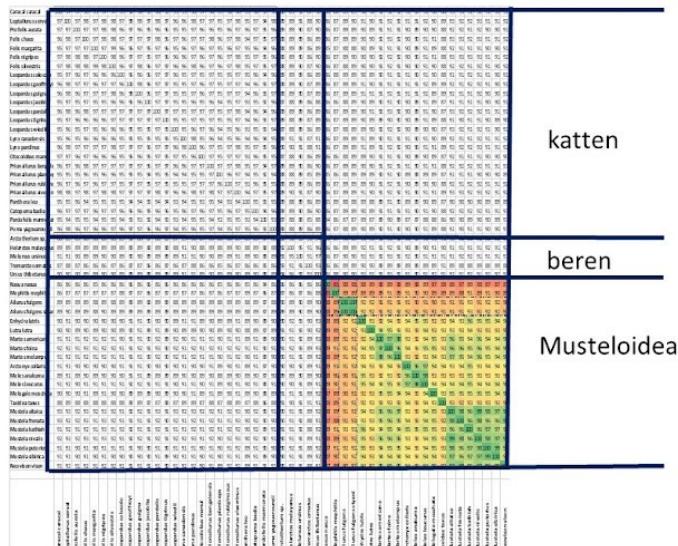


Figure 17. Figure 16, Musteloida colored green = similar red=different.(katten = cats; beren = bears)

Coloring only red panda and family Mustelidae:

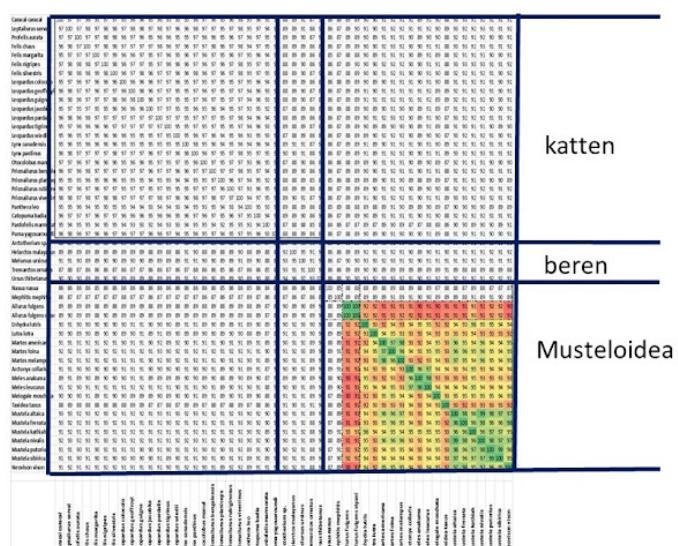


Figure 18. Figure 16 red panda and Mustelidae colored green = similar red = different. (katten = cats; beren = bears)

The large difference of the coati *Nasua nasua* from the other species of the superfamily Musteloidea is the striking feature of the heatmap..

Cserhati uses a strange selection of species to determine whether the red panda belongs to the mustelids based on cytochrome b. Why 25 species of cats? Those cats don't matter much in the analysis; at most they drive the non-cats in a heat map to colors that are closer together. If we want to know where to place the red panda within the superfamily Musteloidea, we have to use species from the Musteloidea.

Cserhati does not give his Fasta file with amino acid sequences for cytochrome b, but he does give the numbers of the accessions he used. I looked up the amino acid sequence of cytochrome b for mustelids, bears, red panda, skunk, coati, and seven species of cats in the NCBI database. I supplemented this with the amino acid sequence of another skunk, and ten species of the raccoon family Procyonidae. This gives the following heatmap:

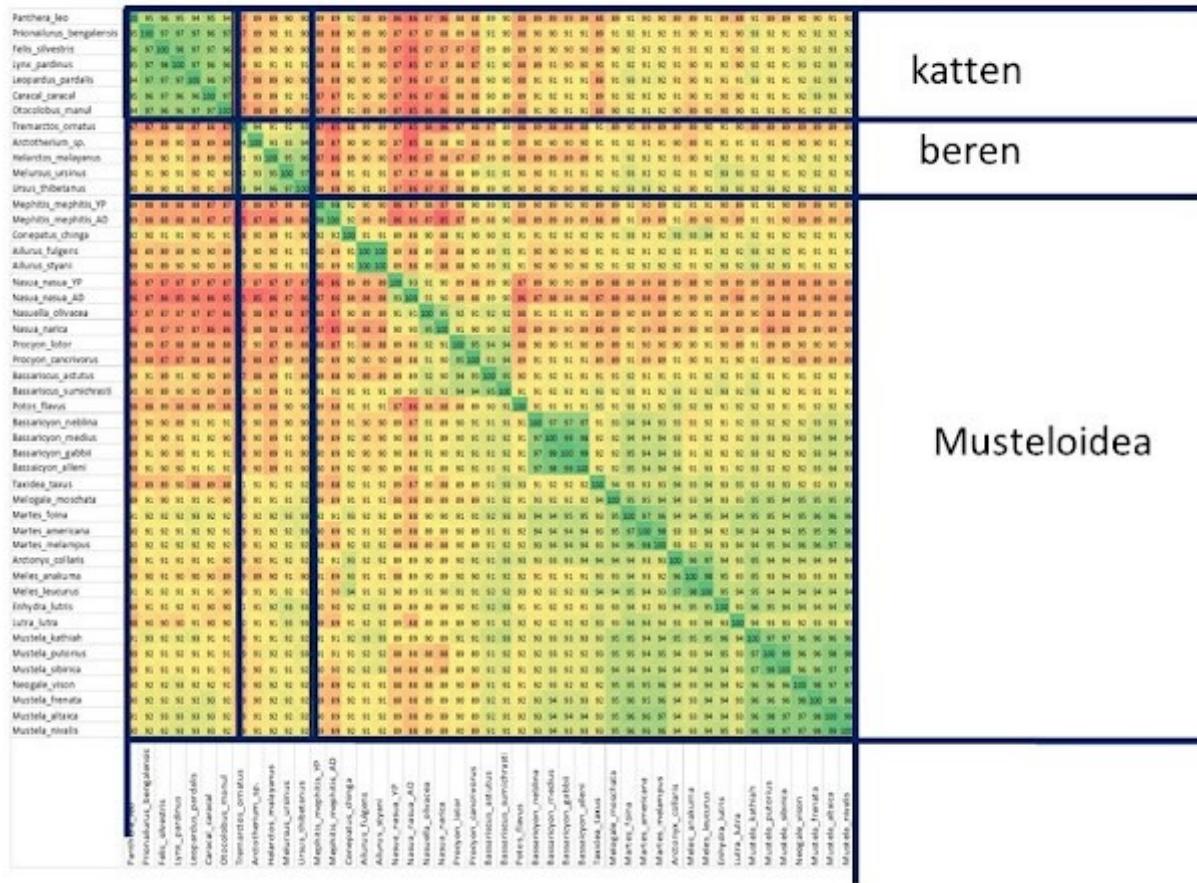


Figure 19. Correlationmatrix of cytochrome b amino acid sequences: fewer cats, more Musteloidea. Green = similar, red = different. Minimum correlation: .85. Katten = cats; beren = bears.

The superfamily Musteloidea proves quite heterogeneous. In more detail:

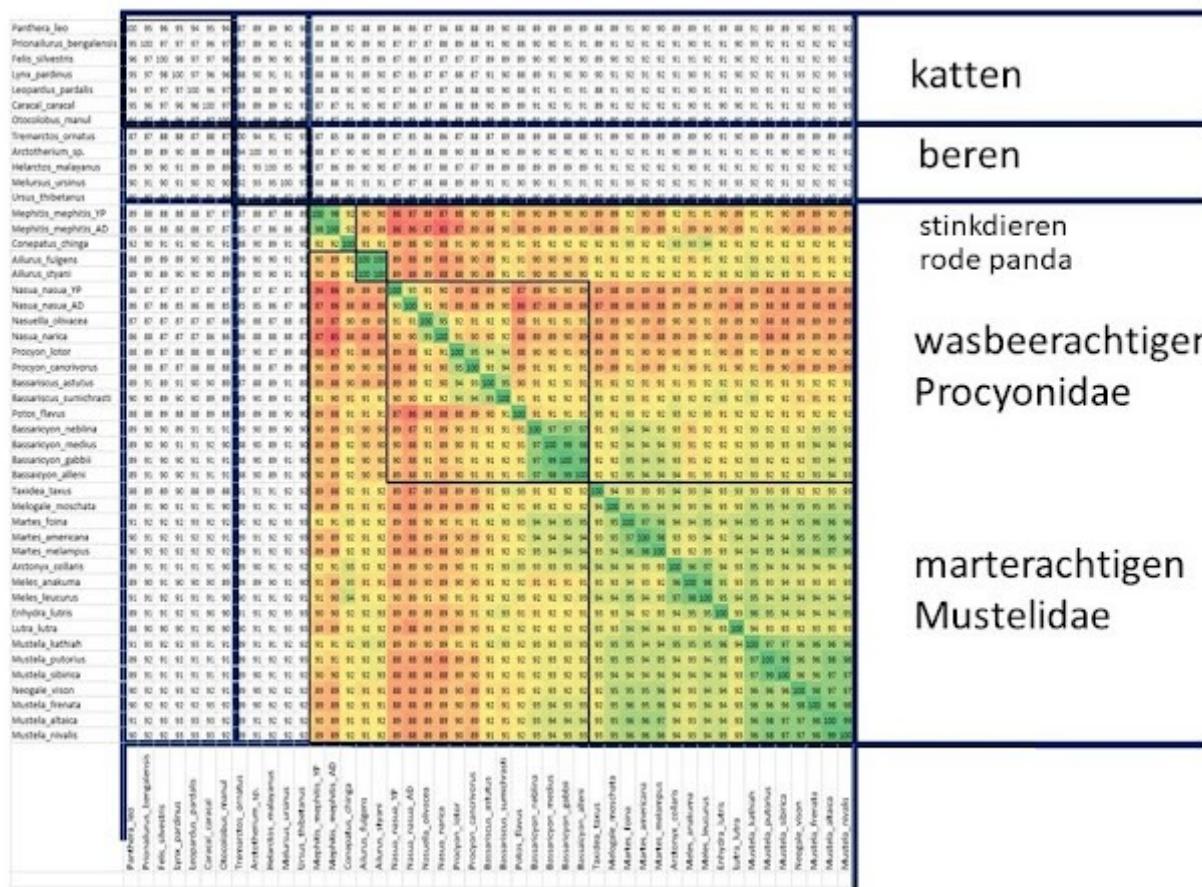


Figure 20. Correlationmatrix of cytochrome b amino acid sequences: fewer cats, more Musteloidea. Green = similar, red = different. Minimum correlation: .85. Katten = cats; beren = bears. stinkdieren = skunks;

Within the family Procyonidae, some species are quite similar to the family Mustelidae, but other species are very different from the Mustelidae. One of the species that differs greatly is the coati *Nasua nasua*. The choice of only the coati in the Cserhati heat map makes the red panda more closely resemble the mustelids than is warranted across the whole of the Procyonidae.

Conclusions

The WGKS clustering gives three groups: the cat family Felidae, the bear family Ursidae and the superfamily Musteloidea.

The mtDNA clustering gives two groups: the bear family and the superfamily Musteloidea.

The cytochrome b clustering gives a cluster cats, a cluster bears and is unclear on the Musteloidea.

The heatmaps indicate that the red panda does not fall within the family Mustelidae in any of these three analyses.

It can also be seen that clustering does not accurately represent biology. The mtDNA analysis gives two clusters, but the heatmap clearly shows how much detail is lost during clustering.

Sloppiness:

Reporting on the mtDNA heatmap Cserhati writes in BMC Genomics "*Three larger clusters and two smaller clusters are visible in the heat map.*" (pg 4-5, flanking figure 1; reporting on the same data in CRSQ, Cserhati writes "*In this heatmap we can see two main, large clusters as well as three smaller ones.*", pg 80

Page 78: whole genome sequence for Procyonidae is mentioned as downloaded

Page 79, Table 1: *Spilogale gracilis* in cluster 3, the cats

Page 81: "... the WGKS analysis, where the average silhouette width was 0.8 for classifying *M. mephitis* ..." It is *Spilogale gracilis* in the WGKS analysis, not *Mephitis mephitis*.

In Figure 2, the text is not in accordance with the figure. The figure should have given 'musteloids' on the left, not 'mustelids'.

Cserhati, M., 2021, Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis, Creation Research Society Quarterly 58 (2): 76-84

M.T.N Tsuchiya, R.B Dikow, K.P Koepfli, P.B Frandsen, L.L Rockwood, J.E Maldonado, 2021. Whole-Genome Sequencing of Procyonids Reveals Distinct Demographic Histories in Kinkajou (*Potos flavus*) and Northern Raccoon (*Procyon lotor*) Genome Biology and Evolution, Volume 13, Issue 1, January 2021, evaa255, Published online: 17 December 2020

Hennigan, T. 2010. The case for holobaraminic status in bears (family Ursidae) and the implications within a creation model of ecology. CRSQ 46(4):271–283.

Meyer, A. 1994. Shortcomings of the cytochrome b gene as a molecular marker. Trends in Ecology & Evolution 9(8):278–280

<https://creationismeweersproken.blogspot.com/2023/02/de-rode-panda-en-cserhati-16-conclusies.html>

THE RED PANDA AND CSERHATI (16): CONCLUSIONS AND DISCUSSION IN THE CRSQ ARTICLE

In his CRSQ article "*Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis*" Cserhati aims to place the red panda in a baramin on the basis of molecular methods:

Thus, molecular baraminology is the study of the created kinds from a molecular biology perspective.

An important point is that 'kinds', baramin, were created separately. For creationists it follows from 'created separately' that differences between the 'kinds' must still be visible today.

Hence, there is continuity between species within a kind, and discontinuity between two separate kinds

not *ex hypothesi*, but *per axioma*.

A 'kind' is identified by finding a 'holobaramin', a hypothesis for a baramin.

Cserhati translates the idea of 'continuous' into statistics, into groups formed by statistical clusters.

All (three) of these groups show statistically significant continuity within themselves and discontinuity with all other species in this study.

Cserhati's question is whether the red panda clusters consistently with other species. If so, the red panda can be assigned to a holobaramin with these species.

Conclusion WGKS analyse

The WGKS analysis produced three clusters, the cat family Felidae, the bear family Ursidae and the superfamily Musteloidea. Cserhati considers these clusters to be hypotheses for holobaramin:

Based on this result, three putative baramins can be defined: felids covering the family Felidae (eleven species), ursids covering the family Ursidae (five species),

and musteloids, a superfamily including Mephitidae and Mustelidae (twelve species).

The red panda is apparently moved into the family Mustelidae, without further explanation. Not the family Mustelidae, but the superfamily Musteloidea is here considered a holobaramin.

Conclusion mtDNA analysis

Cserhati arrives at five clusters based on mtDNA; the five clusters correspond to the five families of the standard taxonomy: bears, skunks, raccoons, red panda and mustelids. That should have led to five holobaramin, with the red panda as its own holobaramin, but Cserhati doesn't draw this conclusion. He searches for literature that seems to support the placement of the red panda among the mustelids, and gives two references.

- 1 *However, some authors have found that A. fulgens is related to mustelid species. For example, Peng et al. (2017) placed A. fulgens next to Martes americana, the American marten in an analysis of 13 concatenated mtDNA proteins.*

Peng et al (2017) use the superfamily Musteloidea as an outgroup for their study of bears, with one species from each of the four families of the Musteloidea: the American marten for the Mustelidae, the striped skunk for the Mephitidae, the raccoon for the Procyonidae and the red panda for the Ailuridae. In one of their two analyses, Peng et al (2017) list the red panda *Ailurus fulgens* as related to the American marten, but as a sister group of the American marten, not as a mustelid. In the other analysis, Peng et al (2017) find the red panda as a sister group of the striped skunk. Peng et al (2017) is an evolutionary study, and then relatedness is not limited to the family, but relatedness also exists between families.

- 2 *Based on a study of the 12S rRNA, the 16S rRNA, and cytochrome-b, Flynn et al. (2000) also classified A. fulgens as a mustelid.*

Flynn et al (2000) classified the red panda not as a mustelid, but as a musteloid. Flynn et al (2000) wrote:

Rather, evidence from nucleotide sequences strongly support placement of the red panda within a broad Musteloidea (*sensu lato*) clade, including three major lineages (the red panda, the skunks [mephitids], and a clearly monophyletic clade of procyonids plus mustelids [*sensu stricto*, excluding skunks])

Cserhati is inaccurate in reading and makes errors in citing.

Cserhati searches in vain for support for his idea that the red panda belongs to the mustelids; such support cannot be found in these two articles.

Conclusion amino acid sequence in the protein cytochrome b.

Cserhati arrives at three clusters: cats, bears and mustelids + red panda; and two separate species, the striped skunk and the coati.

All three larger clusters show statistically significant continuity among themselves and discontinuity with all other clusters.

No silhouette plots are provided to substantiate three clusters of cats, bears and mustelids + pandas and two non-clustered species. As a result, it is also unclear whether mustelids + red panda can be considered a holobaramin.

Three different conclusions from three analyses

Cserhati is faced with three different outcomes in his three analyses. He tries to reconcile them in line with his third finding, that of cytochrome b, a clustering of mustelids with the red panda.

1 comparing results for WGKS and cytochrome b

For the WGKS data, Cserhati now mentions the possibility that there are four clusters, the three clusters (bears, cats and mustelids + red panda) and the Western spotted skunk *Spilogale gracilis* from the family Mephitidae separately. This was the possibility he gave in the BMC Genomics article, but did not mention in this CRSQ article. For cytochrome b Cserhati gives the same three clusters, and now the coati *Nasua nasua* of the family Procyonidae and the striped skunk *Mephitis mephitis* of the family Mephitidae separately.

2 comparing results for mtDNA and cytochrome b

The gene for cytochrome b is located on the mitochondrion. Cserhati wants to compare changes in part of the mitochondrion with changes in the mitochondrion as a whole; and comparing differences in an amino acid sequence with differences in a DNA sequence. Of course, there is much less difference between species in the amino acid sequence than in the DNA of cytochrome b. In fact, there is so little difference between the species that the amino acid sequence of cytochrome b hardly provides enough differentiation between the species to work with. This can be clearly seen in the low resolution and fuzzy patterns in figure 4 of the CRSQ article (figures 15 and 16

blog post 15). A difference between the mtDNA and the cytochrome b analysis does not occur because mtDNA has too many mutations, but because the "structurally conservative" (page 80) cytochrome b protein gives too few changes for a clear analysis. Table 1 shows that there are relatively few differences between species in cytochrome b.

TABEL 1	minimum correlation	mean correlation
WGKS	0.666	0.838
mtDNA	0.751	0.816
cytochrome b	0.852	0.921

3 comparing results for mtDNA and WGKS

Cserhati mentions the possibility that mtDNA does not provide good material for an analysis because it mutates too quickly. To demonstrate rapid mtDNA mutation, he compares the mitochondrial DNA of cytochrome b with an exon of the same length of a gene RAG1 (Recombination Activating Gene 1) that is located in the nuclear genome. It is not mentioned why RAG1 was chosen. More mutations have occurred in the DNA for cytochrome b than in the DNA for RAG1.

This is another indication that mtDNA mutates faster than nuclear DNA.

This is not formulated correctly. Cserhati's comparison is of two protein-coding DNA sequences, one on the mitochondrion and one in the nuclear DNA. Thus, the comparison of mutation rate is not for all DNA in mitochondria and nucleus, but only for protein-coding DNA.

This difference in mutation rate is a well-known phenomenon. The textbook 'Fundamentals of Molecular Evolution' (1990) states on page 86 that the mutation rate in mammalian mtDNA genes is about 10x higher than in nuclear genes (exact quote at the end here). This has since been found dozens of times. A reference to a textbook would have sufficed.

However, then Cserhati says:

This can explain why the WGS, mtDNA, and cytochrome-b results are divergent.

If there were more changes in the mtDNA than in WGKS, the mean cross-species correlation and the minimum cross-species correlation should be greater for WGKS than for mtDNA. Table 1 shows the mean and minimum of the correlations between all species pairs of bears and Musteloidea in the three

analyses. WGKS and mtDNA have about the same range in interspecies correlations.

Although WGKS is a kind of measure over the entire genome, it only depends to a small extent on base pair mutations in protein coding genes. The mutation rate in the nuclear protein coding genome is not decisive for the interspecies differences reflected in the WGKS.

The difference between the detailed heatmap of the mtDNA analysis and the coarse scale heatmap of the WGKS analysis is not because there are more mtDNA mutations. The WGKS results have less resolution than the mtDNA analysis, because there are only few species in the WGKS analysis. Therefore, it is difficult to find clusters within the superfamily Musteloidea.

family	WGKS	mtDNA
Ursidae	5	15
Mustelidae	10	30
Procyonidae	0	2
Mephitidae	1	3
Ailuridae	1	2
	17	52

Alternative conclusion

Another possibility would have been to regard the superfamily Musteloidea as a holobaramin. Cserhati chooses not to do that.

The superfamily Musteloidea as holobaramin is the first outcome of the WGKS analysis. It is also the result of the mtDNA analysis, if we keep the two clusters bears and Musteloidea as indicated by the silhouette plots (see blog post 15). For the amino acid sequence of cytochrome b, it is impossible to say how good the possibility of a cluster of Musteloidea is, in the absence of silhouette plots. The heatmap leaves open the possibility of a Musteloidea cluster.

The discussion of the CRSQ article

Cserhati prefers his interpretation of the WGKS analysis.

The WGKS analysis seems to bring the strongest results, since it is a global analysis of the entire genome. According to this analysis, A. fulgens belongs to the

mustelid holobaramin. Also, *M. mephitis* could either belong to the mustelid holobaramin, or it could either belong to another holobaramin, due to the minimal differences in average silhouette width values.

As part of the results of the WGKS analysis, Cserhati gave a Musteloidea baramin (pages 78, 79) based on clustering of 10 species from the family Mustelidae, the red panda *Ailurus fulgens* and the western spotted skunk *Spilogale gracilis*. (*Mephitis mephitis* represents the family Mephitidae in the other two analyses; Cserhati is confused about his own data). If *Spilogale gracilis* is removed from that cluster, the red panda will continue to cluster with the 10 species of mustelids.

Cserhati is allowed to call any cluster a holobaramin – as holobaramin is a fantasy term -, but the problem is ‘mustelid’. Calling the cluster ‘mustelid’ suggests that the red panda is part of the Mustelidae family, and Cserhati fails to demonstrate that. A clustering does not give a classification.

Cserhati gives a reference to a study by Nie et al (2002); Cserhati presumably intended to indicate this citation to argue a close relationship between red panda and the family Mustelidae.

*Genomically, *A. fulgens* shares several apomorphic chromosome fusions with mustelids, namely F2+C1p and A1p+C1q (Nie, 2002). However, *A. fulgens* differs in several other chromosomal rearrangements indicating that it diverged early from the mustelids.*

Nie et al (2002) examined the chromosomes of the domestic cat, the red panda and five species of the marten family. Nie et al (2002) considered the domestic cat, the red panda and the five mustelids to belong to three different families. In a study based on these five species the mustelids and the red panda are more similar to each other than any of these are to the domestic cat; as is obvious given the phylogenetic tree of the order Carnivora. The domestic cat comes from the main division Feliformia of the order Carnivora, the mustelids and the red panda from the alternative main division, the Caniformia. Procyonidae and Mephitidae are not present in the study by Nie et al (2002) - and therefore nothing whatsoever can be said about how related the red panda and mustelids are. The study by Nie et al (2002) cannot be used to place the red panda closer to the mustelids than to the other families of the superfamily Musteloidea. Cserhati’s phrase “indicating that it (red panda) diverged early from the mustelids” is from Cserhati, not from Nie et al (2002).

There is another possible conclusion: that mtDNA gives the best results.

Alternatively, A. fulgens could be the only known member of its own holobaramin, as mentioned in the Introduction and supported by the mtDNA results.

Cserhati interpretation of five clusters in mtDNA correspond to the scientific classification: species from five families were present in the mtDNA analysis. In fact, the mtDNA heatmap shows the neatest results.

Summary and Conclusion

Based on all of these considerations, it is likely that A. fulgens belongs to the mustelid holobaramin, and not the ursid holobaramin.

The problem here is the use of scientific terms such as 'mustelid' and 'ursid'. A 'mustelid' holobaramin contains a cluster of the family Mustelidae and the family Ailuridae. Everyone is free to provide clusters with names of their own choice: everyone is free to call a cluster of choice a holobaramin. After all, clustering is not biology, but statistics, and clustering is different from classification. Only, the suggestion is here that the scientific family Mustelidae is defined differently, namely including the red panda *Ailurus fulgens*. That would pose a scientific problem.

Cserhati, M., 2021, Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis, Creation Research Society Quarterly 58 (2): 76-84

Hennigan, T. 2010. The case for holobaraminic status in bears (family Ursidae) and the implications within a creation model of ecology. CRSQ 46(4):271–283.

Li, W-H, & D.Graur (1990) Fundamentals of Molecular Evolution. Sinauer Ass, Inc. pg 86:"The synonymous rate of substitution in mammalian mitochondrial genes has been estimated to be 5.7×10^{-8} substitutions per synonymous site per year (Brown et al 1982). This is about 10 x the value for synonymous substitutions in nuclear protein-coding genes. "

Peng, R., B. Zeng, X. Meng, B. Yue, Z. Zhang, and F. Zou. 2017. The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*). Gene 397(1–2):76–83.

Flynn, J.J., M.A. Nedbal, J.W. Dragoo, and R.L. Honeycutt. 2000. Whence the red panda? Molecular Phylogenetics and Evolution 17(2):190–99.

Nie, W., J. Wang, P.C. O'Brien, B. Fu, T. Ying, M.A. Ferguson-Smith, and F. Yang. 2002. The genome phylogeny of domestic cat, red panda and five mustelid species revealed by comparative chromosome painting and G-banding. *Chromosome Research* 10(3):209–222.

<https://creationismeweersproken.blogspot.com/2023/03/de-rode-panda-en-cserhati-17-baramin.html>

THE RED PANDA AND CSERHATI (17): BARAMIN

Cserhati advances several proposals for 'baramin' in his CRSQ article, and provides an interesting comment about recognizability of baramin.

Proposal 1 Musteloidea als baramin (pp 78, 79), first interpretation of the WGKS data

Proposal 2 A mustelid cluster of the red panda with the Mustelidae species, an interpretation of the WGKS and cytochrome b analyses (pp 81, 82)

Proposal 3 The red panda as a separate baramin, as indicated by the mtDNA analysis (page 82)

Proposal 4 How do you recognize 'lineage' versus 'baramin'? (p.81)

Ad Proposal 1

The Musteloidea as baramin appear in the first interpretation of the WGKS results, but this possibility is not further discussed.

Ad Proposal 2

In the WGKS revised interpretation and the cytochrome b interpretation, the red panda and the mustelids cluster together. Cserhati considers them to belong to one baramin. So, he considers the red panda to be related to the mustelids (by definition of baramin), and not related to the raccoons and the skunks. Cserhati realizes that the red panda can be found not inside but next to the Mustelidae family in the heatmaps of these two analyses, somewhat on the outside of the group.

Cserhati proposes an explanation for the red panda being somewhat different from the mustelids, visible as a lower correlation.

*The reason for the low mean PCC (correlation) value between *A. fulgens* and all other mustelids might be its geographic isolation in the mountainous areas of Nepal, India, and China. This could have allowed for greater genetic change to take place.*

It remains to be seen whether the geography of the red panda is quite so isolated, in red panda terms, and whether geographic isolation would explain the difference of the red panda from the mustelids. The hog badger *Arctonyx*, from around the same area as the red panda, falls neatly within the mustelid family. The sea otter *Enhydra* can also be called geographically isolated, but also falls neatly within the mustelids. All in all, although geographic isolation may lead to a separate genetic lineage, it is not certain that this accounts for differences between mustelids and the red panda. In any case, no 'greater genetic change' – that is not documented –, only 'different genetic change'.

Ad Proposal 3

The red panda as a separate holobaramin is a possible conclusion according to the mtDNA heatmap.

Alternatively, A. fulgens could be the only known member of its own holobaramin

The red panda subspecies are the only living representatives of the Ailuridae family according to science. This is supported by the heatmap of mtDNA in the CRSQ article.

Cserhati considers a holobaramin with only one species to be possible.

When a taxon, as in this case, a holobaramin loses a large portion of its constituent species, during a mass extinction, such as the Genesis Flood, it loses its capability to re-diversify after the extinction.

This statement of Cserhati leaves loose ends.

Genesis 7:14 says that of the unclean animals according to their kind (KJV) one pair, a male and a female, go into the Ark. All predators are unclean, so the family Mustelidae with 65 or so now living species and the family Ailuridae with one species would have been both present on the Ark with one pair for their baramin. In the case of the Mustelidae, this would involve a re-diversification to the most species-rich and diverse family within the order Carnivora, and in the case of the Ailuridae, the family would be stuck at one species.

Cserhati thinks it's possible that with a large loss of species in a mass extinction, the remaining species are somehow unlucky and have little genetic variation left for diversification. The question then arises whether the Ailuridae or the Mustelidae lost "a large portion of its constituent species" during a mass extinction. It should be possible to verify this using fossils of the families

Ailuridae and Mustelidae. Fossils of both families are known. Nothing in the fossil record of these families indicates a mass extinction.

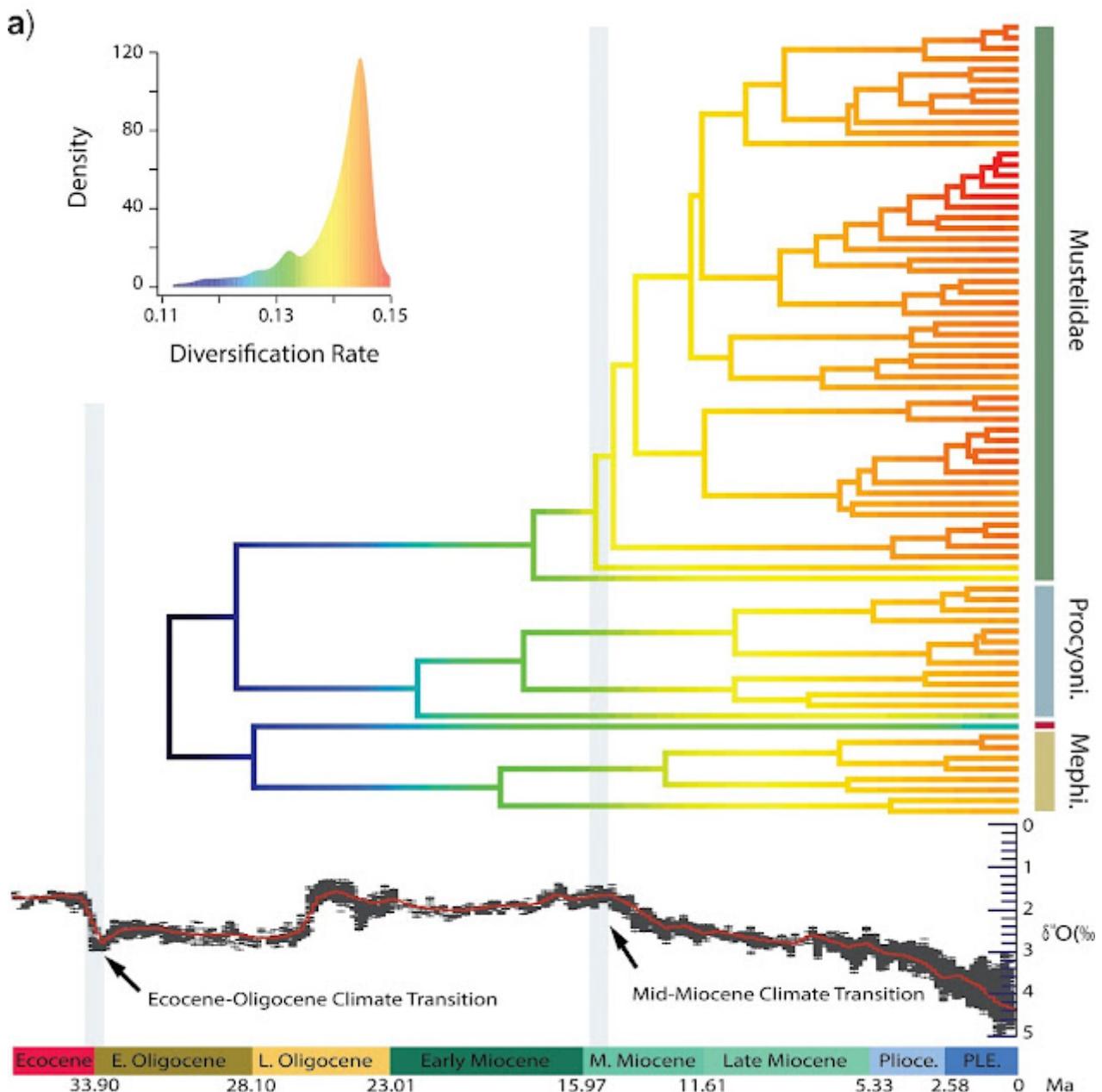


Figure 1. Diversification in fossil Musteloidea. Red, in the lay-out tussen Mephi. en Procyon., Ailuridae are indicated . Figuur 3a van Law et al (2017)

Proposed creationist End-Flood boundaries are the Pliocene-Pleistocene boundary and the Cretaceous-Paleogene boundary. At the Pliocene-Pleistocene boundary, species diversity in the Mustelidae and Ailuridae is close to current (see figure 1). At the Cretaceous-Paleogene boundary, the orders of mammals and therefore the families Ailuridae and Mustelidae did not exist.

Cserhati gives an example of inability for extensive speciation after a mass extinction:

When a taxon, as in this case, a holobaramin loses a large portion of its constituent species, during a mass extinction, such as the Genesis Flood, it loses its capability to re-diversify after the extinction. Such taxa include ..., and parareptiles (MacDougall, 2019).

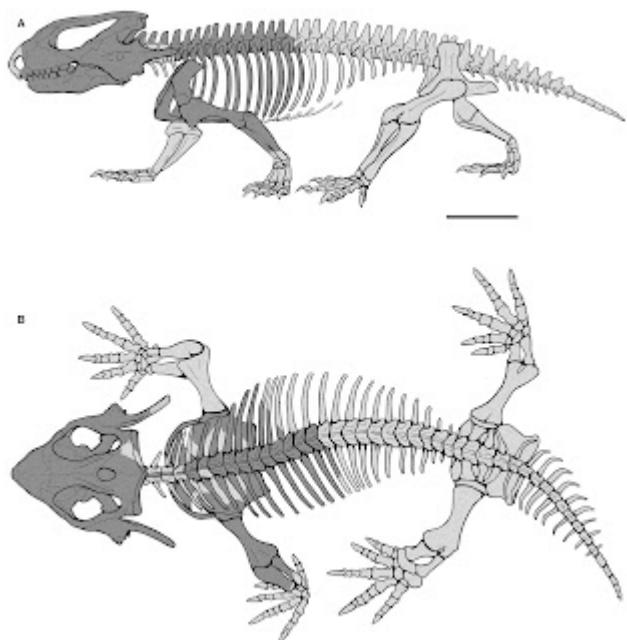


Figure 2 Fossil skeleton parareptile Kapes

The group Parareptiles differs from living reptiles and their associated fossils. The Parareptiles existed from the end of the Carboniferous 306 million years ago to the end of the Triassic period 201 million years ago. They survived the Permian-Triassic mass extinction, but with greatly reduced diversity. The phenomenon cited by Cserhati exists. But it's not such a good example for a creationist or for extinction by the Genesis Flood. The group Parareptiles originated during the Genesis Flood, and died out again during the Genesis Flood.

Ad Proposal 4

Cserhati gives some argument why not to stick with the mtDNA results.

A faster mutation rate means larger differences in the mtDNA sequence in contrast with the nuclear genome. This may reveal lineages within a kind, which might appear to be separate baramins.

Baramin or lineages within baramin surface using mtDNA. Now what are lineages, and what are baramin? The last sentence of Cserhati above is a very interesting comment (p. 81):

lineages within a kind, which might appear to be separate baramins.

This implies that baramin cannot be recognized. After all, it is always possible to find lineages. Hennigan (2009) lists creationist definitions to identify baramin:

Continuity – Biologically meaningful similarity between organisms that embrace all types of biological characteristics.

Discontinuity – Biologically meaningful differences between organisms.

Let us see how these ideas work out.

The red panda eats bamboo, is a vegetarian. The mustelids are carnivorous, with some degree of omnivory in some species of badgers. That is a 'biologically meaningful difference'. Then the red panda would represent a baramin whether or not the red panda clusters with the mustelids.

Is a cluster a baramin or a lineage? With clustering you can find 'lineages' and 'baramin'. Take only the family Mustelidae, and start clustering. A cluster of 'otters' appears, the subfamily Lutrinae. A baramin? Within that subfamily Lutrinae there are lineages, namely the genera. The sea otter stands alone as a species in its genus, with a different lifestyle and different teeth than all other otters. These are 'biologically meaningful differences' between the sea otter and other otters. Perhaps the genus *Enhydra* is a baramin with only one species?

In Figure 2 of the CRSQ article (Figure 3 here) we see subdivisions and subdivisions in the heat map of the WGKS correlations. The hierarchy in classification of the animal kingdom is clearly visible. The cats stand out as a dark group at the bottom left. The bears and the mustelids together color dark gray - aren't we dealing with two baramin, a cat baramin and a bear+superfamily Musteloidea baramin? Rather than with three baramin, cats, bears and Musteloidea? The bears and the mustelids both belong to the taxonomic group Arctoidea, while the cats do not. (There is no figure for the WGKS data with number of clusters on the x-axis and silhouette width on the y-axis to see what the optimal number of clusters is).

The group Parareptiles differs from living reptiles and their associated fossils. The Parareptiles existed from the end of the Carboniferous 306 million years ago to the end of the Triassic period 201 million years ago. They survived the

Permian-Triassic mass extinction, but with greatly reduced diversity. The phenomenon cited by Cserhati exists. But it's not such a good example for a creationist or for extinction by the Genesis Flood. The group Parareptiles originated during the Genesis Flood, and died out again during the Genesis Flood.

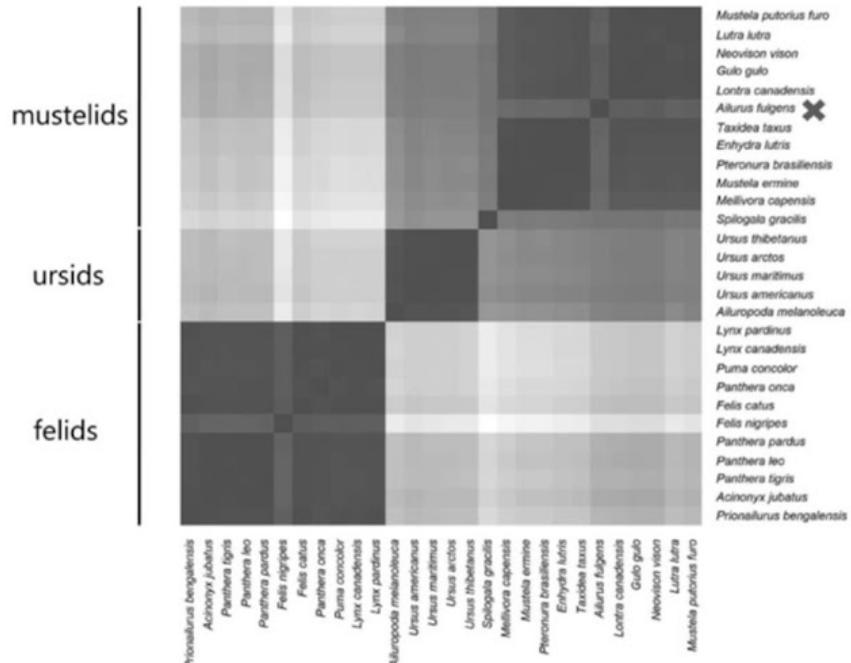


Figure 3 WGKS heatmap with hiérarchical pattern

The dog family Canidae and the Arctoidea together form the group Caniformia, one of the two main groups within the order Carnivora. The other main group is the Feliformia. There are many 'biologically meaningful differences' between the main groups Feliformia and Caniformia. Are the Feliformia and the Caniformia baramin? Or the order Carnivora? Orders differ in important 'biologically meaningful differences'. Perhaps an order is a baramin?

Orders can also be clustered. Take three orders, and three species from each order. The correlations*100 in mtDNA of the nine species are:

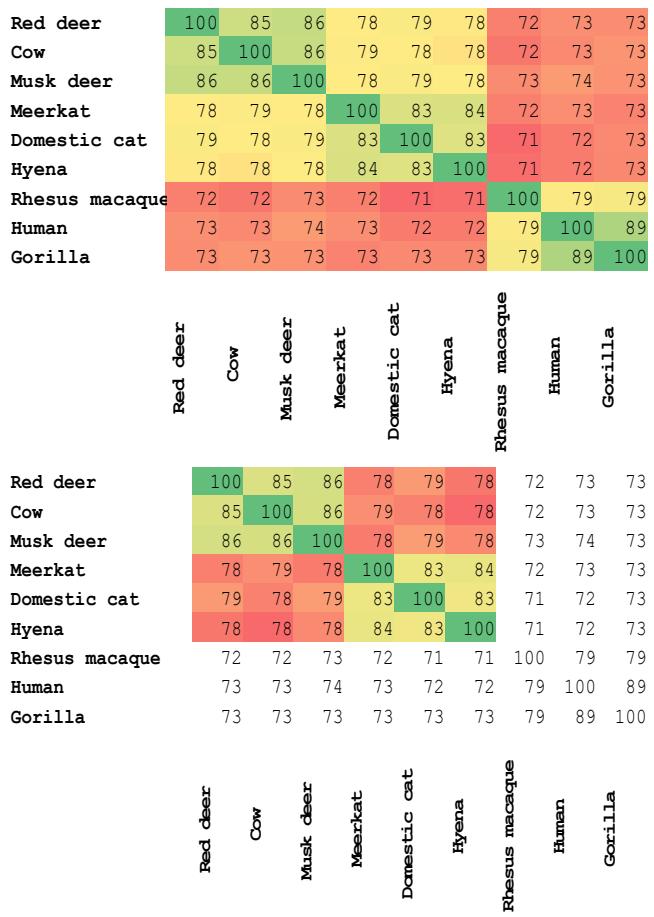


Figure 4.Nine species from three orders in three clusters; correlations * 100 mtDNA

Two groups are visible in the upper heatmap: in mammalian taxonomy, those groups are the Laurasiatheria, upper left, and the Euarchontoglires, lower right. Two baramin? Those Laurasiatheria show two groups. Or should we think of three groups, from top left to bottom right, the orders ungulates Artiodactyla, predators Carnivora and apes/monkeys Primates? Three baramin? Or eight families, with eight baramin? Or nine baramin, despite the clustering?

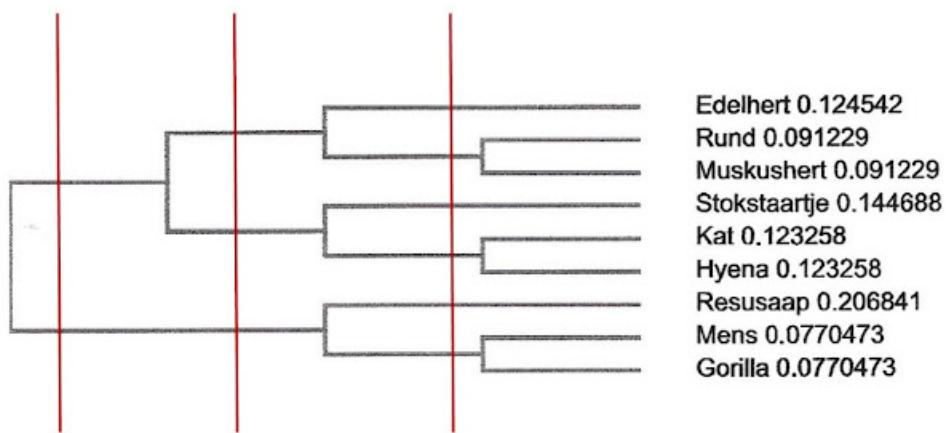


Figure 5 Hierarchical classification of the nine species. Top three species ungulates, middle

three species carnivores, lower three species primates. This is a diagram, the differences between the species are in the numbers on the right, not in the line length. The red lines indicate how many groups will be found on a taxonomic level: 2,3 or 6 - compare the heatmap.

The animal kingdom is arranged hierarchically: this is how the animal kingdom is structured. Hierarchical classification is according to 'biologically meaningful similarity'. Hierarchical division allows 'biologically meaningful differences' to emerge. Clustering produces random patterns if data are hierarchical in nature - a cluster can be chosen at any level in the hierarchy (see vertical lines). Hence, clustering cannot tell what is a baramin and what is a lineage.

In fact, the hierarchical structure of life makes a fiction of the whole idea that baramin can be found by clustering.

Cserhati, M., 2021, Classification of the Enigmatic Red Panda (*Ailurus fulgens*) Based on Molecular Baraminology-Based Analysis, Creation Research Society Quarterly 58 (2): 76-84

Hennigan, T. 2010. The case for holobaraminic status in bears (family Ursidae) and the implications within a creation model of ecology. CRSQ 46(4):271–283.

Law, C. J.; Slater, G. J. & Mehta, R. S. (2018). Lineage Diversity and Size Disparity in Musteloidea: Testing Patterns of Adaptive Radiation Using Molecular and Fossil-Based Methods. Systematic Biology. 67: 127–144.
doi:10.1093/sysbio/syx047.

<https://en.wikipedia.org/wiki/Parareptilia>

https://commons.wikimedia.org/wiki/File:Kapes_full_skeleton.jpg

<https://creation.com/disagreements-on-the-post-flood-boundary>

<https://creationismeweersproken.blogspot.com/2023/04/de-rode-panda-en-cserhati-18-cserhati.html>

THE RED PANDA AND CSERHATI (18): CSERHATI TRIES TO ANSWER

On November 5, 2022, Jan van Meerten wrote an article on his Dutch website 'Oorsprong' with the title: **Wetenschapper lost (creationistisch) biosystematisch raadsel van de rode panda (*Ailurus fulgens*) op.** This translates as: "**Scientist solves (creationist) biosystematic riddle of the red panda (*Ailurus fulgens*)**". A Dr. M. Cserhati had investigated the possible taxonomic position of the red panda, and published the results in the scientific journal BMC Genomics and in the creationist journal Creation Research Society Quarterly.



figure 1

Jan van Meerten described the outcome of the research as: "The results of this study show that the red panda belongs to the mustelid family (Mustelidae)." This was based on what Cserhati wrote in his BMC Genomics article: 'The main conclusion we can draw from this research is that at the whole genome level, *A. fulgens* belongs to the mustelid clade.' And on what Cserhati wrote in his CRSQ article: 'Probably the red panda belongs to the holobaramin of the mustelids.'

One doesn't need to know much about the red panda to know that "red panda a mustelid" can't be right. In 2010 I had looked at the molecular taxonomy of the red panda, and the classification of the red panda turned out to be not completely clear, but not too puzzling either. A club of four families form the superfamily Musteloidea: the martens, raccoons, skunks and the red panda (the red panda itself as a family). In 2020, all studies arrived at the combination (martens, raccoons), but then? Is it (red panda , (skunks , (martens , raccoons))), or (skunks , (red panda , (martens , raccoons))))? An amusing little problem: what exactly is the order of that quick split between skunks, red panda, and martens+raccoons? But the red panda a mustelid? Never in all of the scientific literature. There's something odd about what Cserhati is doing.

So I looked into Cserhati's writings a bit further, and wrote a 1000 word comment on the Panda's Thumb website (December 2, 2022). Cserhati joined the discussion on Panda's Thumb and wrote a 4000 word comment on Jan van Meerten's website 'Oorsprong' (December 13, 2022). Then in 2023 I wrote a long series on the scientific literature concerning the red panda and all the errors in Cserhati's two articles, on my website "creationisme weersproken blogspot" ('creationism contradicted blogspot').

Now Cserhati's piece on the website "Oorsprong" is due for comment.

Cserhati on 'Oorsprong' I: Until Figure 1

Cserhati's first 141 words, that is, his title and the sentences above figure 1, contain eight erroneous statements. Statements by Cserhati are in **bold**.

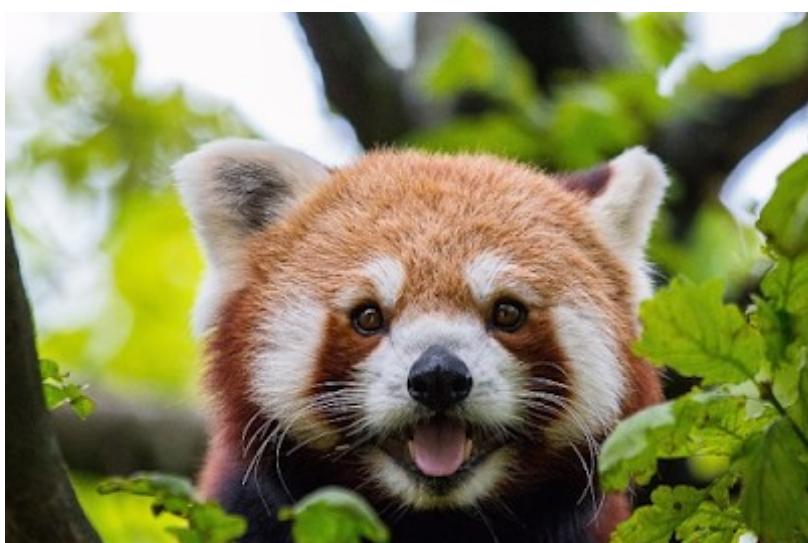


Figure 2: Figure 1 Cserhati on Oorsprong

1 The title "**Evolutionists on Panda's Thumb Attack Well-Respected Science Journal**" gives the impression that the article in Panda's Thumb is an attack on BMC Genomics. In contrast, it is critical of both of Cserhati's pieces, and that has nothing to do with the journal BMC Genomics. It has to do with the extremely poor work Cserhati has done. It is true that the question: "*One problem: how did a paper as bad as this ever get through review and published?*" appears on Panda's Thumb in the section on the BMC Genomics article, but it also applies to the CRSQ article.

2 '**Some have classified it** (the red panda) **as a cat...**' This is wrong. No one has ever classified the red panda as belonging to the cat family Felidae. It is not even claimed article by Cserhati in any of his two articles. The CRSQ article says: "*Geoffroy-Saint-Hilare and Cuvier originally classified the red panda as a member of the **raccoon** family (Procyonidae), although they ended up calling it Ailurus, because of its somewhat feline appearance*". (bold added gdj) The red panda was classified with the raccoons Procyonidae upon its first description. That was in 1825, and since then the poor red panda has been stuck with a name that means "wild cat" in ancient Greek.

3 "Some have classified it as a cat, **some as a relative of small carnivorous animals, such as minks or weasels, ...**" This is wrong. No one has ever classified the red panda with the minks and weasels or with the family that the minks and weasels belong to. Minks and weasels belong to the Mustelinae subfamily of the Mustelidae family. In the introduction to the BMC Genomics article, Cserhati gives three references for his idea that the red panda belongs to the mustelids: one misquoted - it says musteloid and Cserhati reads mustelid - , and two misinterpreted.

4 "Some have classified it as a cat, some as a relative of small carnivorous animals, such as minks or weasels, **yet others claim that it is a relative of the giant panda, *Ailuropoda melanoleuca*.** .." This is wrong. Past and present are mixed up. Cserhati's source, Flynn et al (2000), makes clear that this should be 'claimed' and not 'claim'. Flynn et al (2000) write about the older literature: "*(red panda) sister taxon to the giant panda either with uncertainty about their broader relationships (Segall, 1943) or with the panda clade as sister taxon to ursids (Ginsburg, 1982)*". References from 1943 and 1982 are a bit long ago to use the present tense in 2022.

5 " Similar to the duck-billed platypus, **biologists do not know where to put this animal,** ..." . This is wrong. The red panda has been placed in the superfamily Musteloidea for at least 25 years, in a quick split between skunks, red panda, and martens+raccoons. The question was how to resolve the small

distances between the families (see blog post 2). In fact, with the article by Hassanin et al (2021), the placement of the red panda seems decided. Cserhati could have known that at the end of 2022. In 2020 or thereabouts when Cserhati wrote his articles, it was already established that the red panda belonged to the Musteloidea *sensu lato* but not to the Musteloidea *sensu stricto*. So, there were only the three possibilities that I mentioned in blog 2.

6 "Similar to the duck-billed platypus, biologists do not know where to put this animal,". This is wrong. The platypus is staged for no valid reason. The platypus is the only living member of the platypus family (Ornithorhynchidae) and suborder platypus (Platypoda), within the order Monotremata of mammals. Neatly classified, with a known evolutionary history. For unbiased information see wikipedia; for a biased uninformative story see the creation.com entry Cserhati alludes to

7 "Molecular characterization of this species based on small sets of genes also did not fare any better, and evolutionists contradicted one another again in further attempts to classify this animal." This is wrong. Flynn et al (2000) had "a total of 3450 bp ... for 17 carnivoran species" - few by current standards, but they arrived at a classification consistent with the most recent classification, that of Hassanin et al (2021) based on much more material.

8 "Molecular characterization of this species based on small sets of genes also did not fare any better, and evolutionists contradicted one another again in further attempts to classify this animal." This is wrong. The red panda placement differences were never major - it's just about the red panda's place within the superfamily Musteloidea. No one makes the red panda a relative of, for example, the bear cat (*Arctictis binturong*), although the bear cat may resemble the red panda in terms of appearance, lifestyle and distribution area. The bear cat belongs to the civet family Viverridae. Cserhati tries to make a major disagreement out of a minor uncertainty.

Cserhati 'Oorsprong' II. Table 1

After the eight errors above the first figure, we get:

"To illustrate the fact that evolutionists themselves cannot make heads or furry red tails of where to place the red panda, see Table 1 below to see how various evolutionist researchers tried to classify this mysterious animal."



Figure 3 Figure 2 Cserhati on Oorsprong

Cserhati's Table 1 does not tabulate the classifications scientists arrived at, but Cserhati's erroneous interpretation of their scientific work.

Cserhati's Table 1.

Researcher(s)	Analysis	Conclusion
Peng et al	13 mitochondrial genes	Either mustelid or mephitid
Fulton and Strobeck	3 nuclear and 3 mitochondrial genes	Related to <i>Mephitis mephitis</i>
Yu and Zhang	Introns 4 and 7 of nuclear FGB gene, ND2 mitochondrial gene	Sister to procyonids
Sato et al	5.5 Kbp segment containing 5 nuclear genes	Closer to mustelids or procyonids

Cserhati cites four articles and succeeds in drawing four erroneous conclusions.

Table 1 line 1

'Either mustelid or mephitid'

The article by Peng et al (2007) deals with the placing of the giant panda in the bear family Ursidae, with a whole series of more or less related species as background, as outgroups. The Musteloidea are part of the outgroup. For the Musteloidea are represented by one species per family: the red panda for itself as family Ailuridae, the raccoon for the Procyonidae, the striped skunk for the Mephitids and the American marten for the Mustelidae. Peng provides two analyses.

This is the NJ method.

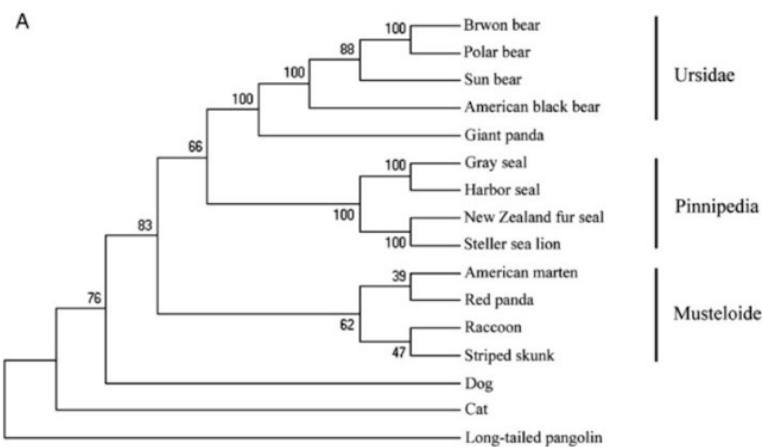


Figure 4 Figure 2A Peng et al (2007)

Here the red panda forms the sister group to the American marten, but since there is only one species of the family Mustelidae – the American marten-, it cannot be concluded that the red panda belongs to the Mustelidae: "*mustelid*" is incorrectly reasoned.

This is the ML method:

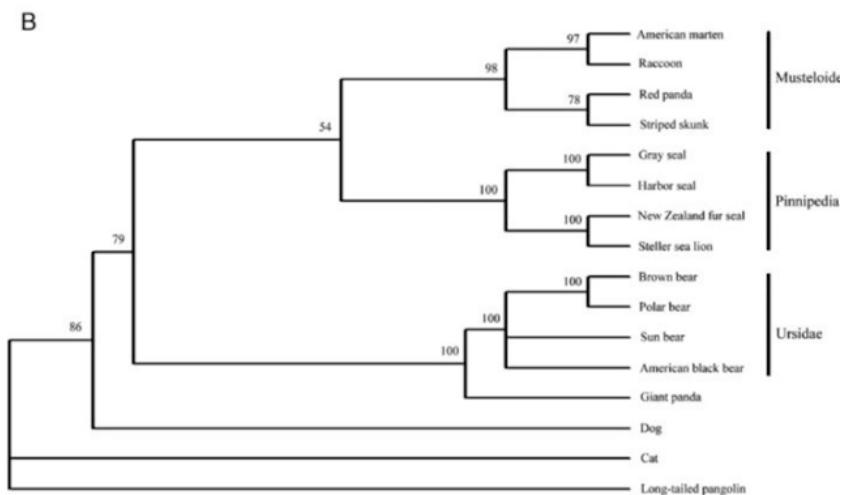


Figure 5 Figure 2B Peng et al (2007)

Here the red panda forms the sister group to the striped skunk, but since there is only one species of the family Mephitidae, it cannot be concluded that the red panda belongs to the Mephitidae: "mephitid" is incorrectly reasoned. It is a basal error in reasoning.

The red panda, as family Ailuridae, appears to be the sister family of the Mustelidae or the Mephitidae. No way the red panda would be part of the family Mustelidae or Mephitidae, as 'mustelid or mephitid' suggests. Cserhati doesn't understand the sistergroup idea. Cserhati misreads the article by Peng et al (2007). Cserhati's citation of the work of Peng et al is also wrong: in the BMC Genomics article as well as the CRSQ article and on Origin, Cserhati gives the date '2017' while the article dates from 2007.

Table 1 line 2 'Related to *Mephitis mephitis*'

Fulton & Strobeck (2007) is a study of the phylogeny within the raccoon family Procyonidae, with the red panda, two marine Carnivora (Pinnipedia) and the wolf as outgroup.

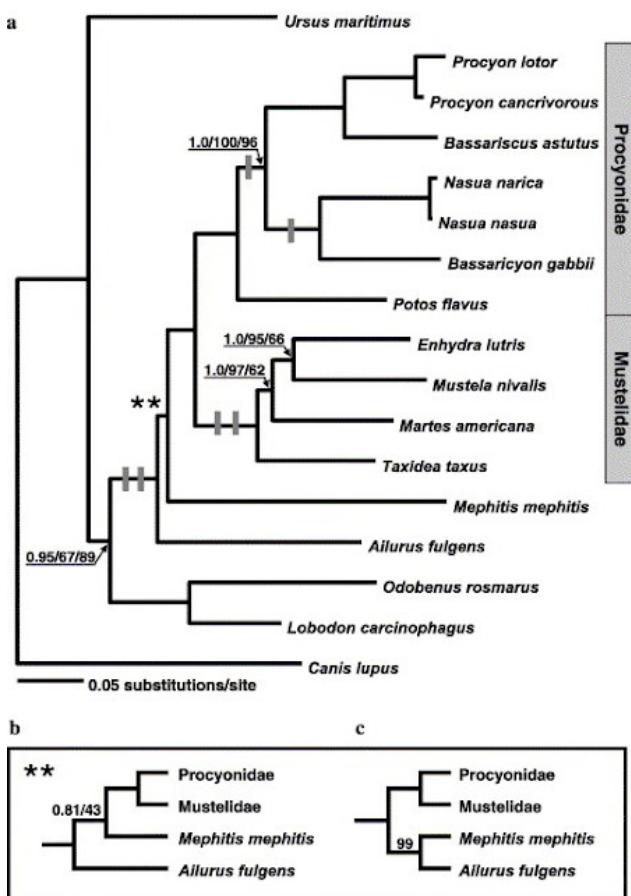


Figure 6 Figure 1 Fulton & Strobeck (2007)

Fulton & Strobeck (2007) provide several analyses. Maximum Likelihood and Bayesian give the red panda Ailuridae as a sister group of Mephitidae + Mustelidae + Procyonidae (their fig 1a,b); Maximum Parsimony lists the red panda as a sister group to the Mephitidae (their fig 1c).

"Related to *Mephitis mephitis*" - At most there is a sister-group relationship between the family Ailuridae and the family Mephitidae. The preference of Fulton & Strobeck (2007) goes to fig. 1ab, with the red panda family Ailuridae as a sister group of the three other families within the Musteloidea.

Note that in the layout of figures 1abc of Fulton & Strobeck (2007) *Ailurus fulgens* is placed next to *Mephitis mephitis*, even when the phylogenetic tree differs. Perhaps Cserhati just looked at the lay-out.

Table 1 line 3 'Sister to procyonids'

Yu & Zhang (2006) provide a phylogeny of the Caniformia, one of the two main groups of the order Carnivora.

In Table 1 on the website 'Oorsprong', Cserhati gives a different interpretation of the results of Yu & Zhang than in his introduction to the BMC Genomics article. Both interpretations are wrong.

In his introduction to the BMC Genomics article, Cserhati says:

"Yu and Zhang studied introns 4 and 7 from the nuclear gene β -fibrinogen (FGB) as well as the mitochondrial gene NADH dehydrogenase subunit 2 (ND2) in 17 species from the order Carnivora. In their results these researchers found that A. fulgens is most closely related to procyonids based on analysis of intron 4 of the FGB gene. When the two introns were combined with analysis of the genes IRBP and TTR, A. fulgens was closest to mustelids [9]."

In his introduction to the BMC Genomics article, Cserhati confuses three times the layout of the phylogenetic tree with the classification according to the phylogenetic tree. (See blog 5 of this series for the figures). That's quite a basic mistake.

In Table 1 on 'Oorsprong', Cserhati says: "*Introns 4 and 7 of nuclear FGB gene, ND2 mitochondrial gene - Sister to procyonids*". On 'Oorsprong', Cserhati seems to have forgotten what he said in the introduction to the BMC Genomics article, but again Cserhati confuses the layout of a phylogenetic tree with the classification according to a phylogenetic tree.

Yu & Zhang (2006) give for "*Introns 4 and 7 of nuclear FGB gene, ND2 mitochondrial gene*" the red panda as a sister group of the Musteloidea *sensu*

stricto: the mustelid family Mustelidae and the raccoon family Procyonidae together.

Table 1 line 4

'Closer to mustelids or procyonids'

Sato et al (2009) is the only one of the four papers listed that really deals with red panda placement. Sato et al (2009) give the following phylogenetic tree:

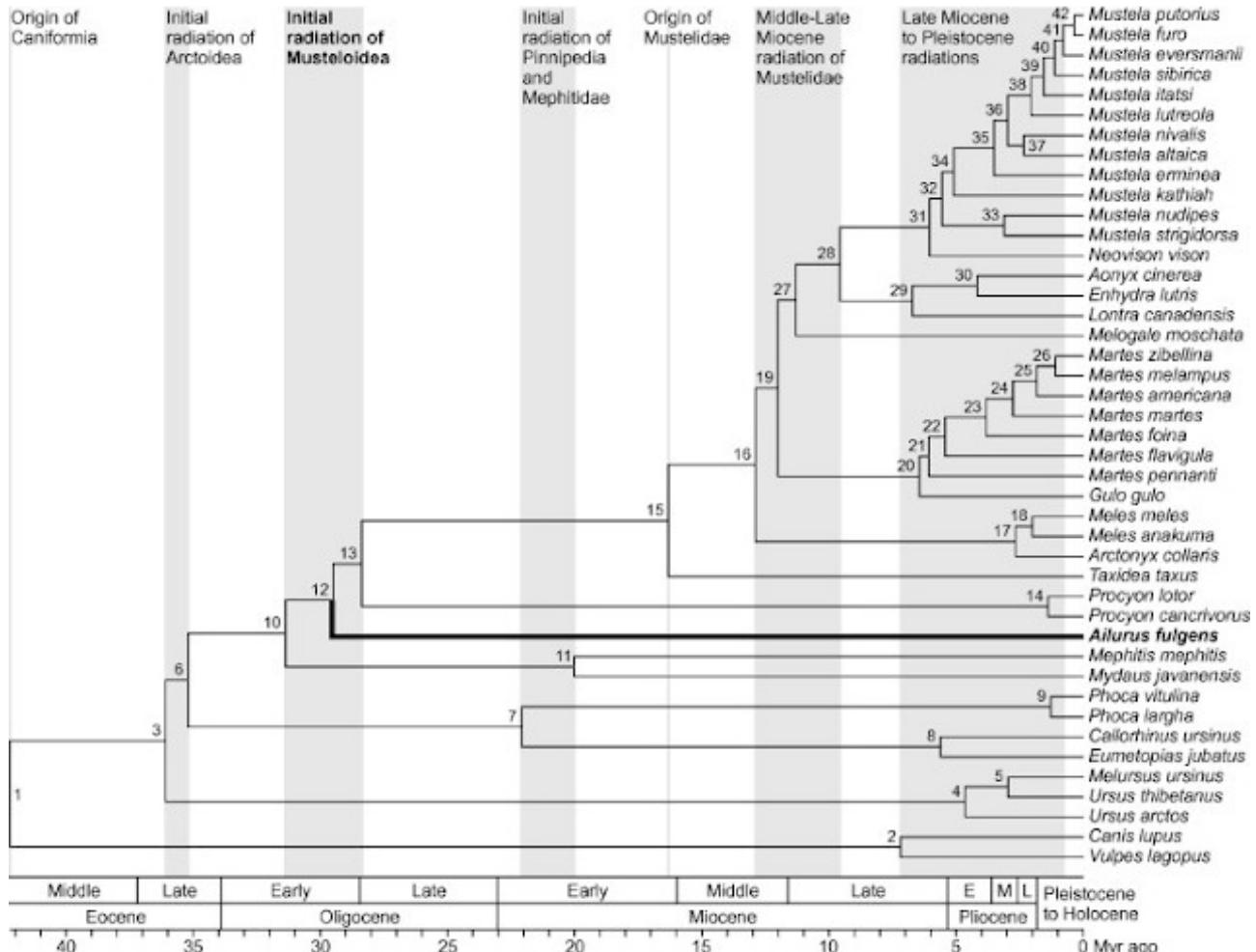


Figure 7 Figure 3 Sato et al (2009)

The conclusions by Sato et al (2009) are:

All our phylogenetic analyses, both the probabilistic and parsimony ones, strongly supported a close relationship between the red panda and a clade containing Mustelidae and Procyonidae to the exclusion of Mephitidae (Fig. 1).

The close affinity of the red panda with the mustelids and procyonids to the exclusion of mephitids had been first hypothesized in our previous study (Sato et al., 2006) and a few months later was independently confirmed by Fulton and Strobeck (2006).

Sato et al (2009) list the red panda as the sister species of Musteloidea *sensu stricto*, the families Mustelidae and Procyonidae together. To quote the conclusion of Sato et al (2009) as "Closer to mustelids or procyonids" is pure lack of understanding, not grasping sistergroups. "Closer to mustelids or procyonids" is quite wrong. What would Cserhati think Sato's "a clade containing Mustelidae and Procyonidae" means?

Yu & Zhang (2006) and Sato et al (2009) give the same phylogenetic tree, a phylogenetic tree that also occurs elsewhere in the literature; see blog 2 for that. Fulton & Strobeck (2007) give a slightly different one often found phylogeny (see again blog 2).

Table 1 was meant: "*To illustrate the fact that evolutionists themselves cannot make heads or fury red tails of where to place the red panda, see Table 1 below to see how various evolutionist researchers tried to classify this mysterious animal.*"

Not so. On the contrary Table 1 is extremely useful *to illustrate the fact* that Cserhati is not able to read the scientific literature.

Note these four article are published in 2007, 2006, 2007 and 2009. Cserhati is either ignorant of more recent scientific work on the classification of the red panda, or he chooses not to cite recent papers.

Cserhati on 'Oorsprong' III

Section III contains 1245 irrelevant words.



Figure 8: Figure 3 Cserhati on Oorsprong

Cserhati on 'Oorsprong 'IV: Cserhati's comment on post and discussion on the Panda's Thumb

After figure 3, in section IV, Cserhati gets to the point and starts discussing blog post and the first part of the discussion that appeared on the website Panda's Thumb. Cserhati gives four propositions which he says can be found on the Panda's Thumb. Cserhati seems to disagree with those four statements. The Panda's-Thumb-statements-according-to-Cserhati are printed here in italics / bold.



Figure 9

IV A

1. The red panda and the marten family (Mustelidae) are sister groups. Ailurus fulgens is not a member of Mustelidae.

I wrote on *Panda's Thumb*: "Nowhere in his (Cserhati's) analyses does the red panda appear within the marten group. At most we see a monophyletic group of the marten and panda families." Seemingly Cserhati disagrees, as he marks this statement for comment.

Cserhati says in his BMC Genomics article:

"The main conclusion that we can draw from this study is that on a whole genome level *A. fulgens* possibly belongs to the mustelid clade,"

"Based on this evidence, *A. fulgens* would belong to mustelids as a monophyletic group."

"In conclusion, *A. fulgens* possibly belongs to Mustelidae, based on the analysis of the WGKS."

Cserhati's conclusion is pertinently wrong. Not only that, but this conclusion is based on disregarding his own results. Cserhati's own results show that the red panda forms a group of its own.

The red panda as a sister group of the Mustelidae follows directly from Cserhati's BMC Genomics figure 2 (UPGMA phylogenetic tree on WGKS, blog post 7) and figure 4 (UPGMA on mtDNA, blog post 12). The NJ and ML analyses (BMC Genomics figure 5 and figure 6) give the red panda as an independent group even with Cserhati's strange interpretation; and if the bears are taken as an outgroup – as they should have been in a correct analysis – , the red panda emerges as an independent group and sister group to the remaining three families within the Musteloidea. (See blog post 12 for further explanation). These four phylogenetic trees, Figures 2, 4, 5, 6 from the BMC Genomics article, give the actual results of Cserhati.

However, Cserhati relies entirely on the heat map of the WGKS correlations. There is no valid justifiable reason for this. First, to be valid reasoning a heat map should be suitable for classification; on the contrary, a heatmap is not suitable for classification. Second, to be valid reasoning the red panda should fall within the family Mustelidae in the heatmap; however, the red panda is not included within the mustelids in the heat map. Cserhati gives the heatmap of the WGKS correlations as his figure 1 from the BMC Genomics article and as figure 4 on Origin. This figure clearly shows that the correlations of the red panda with the species of the Mustelidae are lower than the correlations between the species of the marten family amongst each other. Cserhati agrees: "*It is true that the red panda has a mean correlation value of 0.89 with the mustelids, whereas the mean correlation among mustelids is 0.95±0.04*". The difference is significant, as I said on Panda's Thumb and when discussing the BMC Genomics results. Cserhati does not deny that. Cserhati does not understand this precludes inclusion of the red panda in the marten family Mustelidae.

Cserhati takes a completely different turn, with contorted reasoning. He refers to the difference between the black-footed cat *Felis nigripes* and the other ten species of the cat family. In his heatmap of WGKS correlations, the black-footed cat appears to differ strongly from the other cats: even significantly different.

The black-footed cat is undoubtedly a species of the cat family Felidae.

Then Cserhati reasons:

- (i) apparently under WGKS a great difference between a species A and the other species of its family is possible;
- (ii) then a great difference between a species B and known species of a family means that species B will belong to the same family as those other species.

According to that sort of reasoning, an animal species B might belong to a family of plants.

Somewhere in the discussion on Panda's Thumb, Cserhati says: "*Since we know from biology that *F. nigripes* belongs to the cat baramin, then why can we then not infer that the red panda is a mustelid, with its smaller distance?*" The answer is of course: "Because, if A implies B, it is not necessary that B implies A."

In the BMC Genomics article Cserhati remarks a propos the WGKS difference of the black footed cat and the other cats: "*Based on this (black footed cat) evidence, *A. fulgens* would belong to mustelids as a monophyletic group.*" It is not 'evidence' at all, far from it. Nowhere in his two articles has Cserhati any evidence that the red panda might be a mustelid, could belong to the family Mustelidae. On the contrary, the phylogenetic trees in Figures 2, 4, 5, 6 from the BMC Genomics article, phylogenetic trees that give the actual results of Cserhati, clearly contradict any such idea..

Cserhati protests on the website 'Oorsprong' and in the discussion on Panda's Thumb that I did not go into his story about the black-footed cat . I had, of course, read it, but considered this tortuous and twisted reasoning irrelevant given the clear results of the rest of the analyses.

However, the black-footed cat WGKS results lead to quite a different result. The black-footed cat raises the question of what correlations in WGKS actually refer to and how useful they are for phylogeny. Apparently under WGKS a great difference between a species A and the other species of its family is possible; this does not spell well for WGKS as a method. The UPGMA phylogenetic tree on the WGKS clearly shows that WGKS leads to classification artifacts (BMC Genomics figure 2). Cserhati should have noted that. The artefacts indicate WGKS is not valid procedure for classification.

Cserhati does not provide a data file for WGKS (Whole Genome Octamer Scores) for his 28 species with the supplementary files posted on the internet as far as I can see. That prevents anyone from gaining insight into how those octamer scores behave. It is necessary to know how the octamer scores behave in order to make heads or tails of the strange position of the black-footed cat in the WGKS analysis. In the WGKS analysis, the black-footed cat

Felis nigripes differs strongly from the domestic cat *Felis catus*. *Felis nigripes* is basal to all other cats in this analysis (figure 2 BMC Genomics). That is something that needs explanation, minimally some attention. Especially since Cserhati always highly praises his 'whole genome' method, without any justification that is apparent or available.

The problem with Whole Genome Octamer Scores is that it is an unvalidated method with unknown sensitivities. A neat validation study, for example a Mitochondrial DNA Octamer score, or an Octamer Score profile of a circumscribed part of the genome, is needed to see how useful octamer scores are for phylogeny. Now we have to make do with Cserhati's claim that the octamer scores are a good measure of the entire genome. Why should we take that at Cserhati's word? At a minimum, the data, the octamer scores, must be available on the internet. The question is to what extent octamer scores not so much represent the important part of the genome, but above all microsatellites, repeats and defective transposons. (Blog post 6)

Cserhati should have done a study of the influence of repeats, microsatellites, transposon numbers, on his octamer scores. Without that, the question is whether a correlation between octamer scores indicates something other than similarity or difference in number of repeats. Classifying species by number of repeats is not an obvious or immediately reliable method for phylogeny.

IV B

2. In the mtDNA study there is no outgroup, why was one not included?

On Panda's Thumb I wrote the following about the mtDNA part of the BMC Genomics article.

"The mtDNA sequences are subjected to phylogenetic analysis by UPGMA, Maximum Likelihood and Neighbour Joining. No outgroup is mentioned, and no outgroup seems to have been specified.

The three phylogenetic trees on mtDNA differ in the placement of the families. The mess seems the result of the lack of outgroup. When plotting an unrooted tree in a square format, the program picks the plot. The pandas do not appear within the marten family in any of the three phylogenetic trees."

So I clearly objected to the way Cserhati uses evolutionary phylogenetic methods in the BMC Genomics article. In his answer to that obvious objection,

Cserhati only talks about baramin and baraminological methods: not a word about his phylogenetic studies in the BMC Genomics article, while the objection was undoubtedly about those.

It is unclear whether Cserhati understands why an outgroup is used. Cserhati begins with a correct description of outgroup: "*outgroup species or groups to compare the group under study to a species or group of species that are known to be unrelated to the group under study.*" Such a correct definition can be found in any textbook.

After that, however, confusion sets in: "*.. since the groups that I used (bears, cats, and martens) follow the "traditional classification" the use of an outgroup may be superfluous.*" That is total nonsense. It shows that Cserhati does not understand what an outgroup is for: to put the groups of the "traditional classification" in the correct order in a hierarchical classification.

Then something very strange happens. After giving a correct description of an outgroup Cserhati does not know how the concept is applied in any phylogenetic analysis:

"De Jong wishes to impose her evolutionary view of classification on the species examined in this study, looking for an outlier species more basal or primitive than the species in the study".

Three jinxes in this opinion by Cserhati. First, Cserhati changes the definition of outgroup. Second, Cserhati presents four phylogenetic trees in his BMC Genomics article, and in presenting a phylogenetic tree he is dealing with the "evolutionary view of classification" whether he recognizes it or not, and whether he wants to or not. Third, I wasn't asking for an "*outlier species more basal or primitive*" (an idiotic description for outgroup), I was asking Cserhati for a neat analysis of his own mtDNA material.

Cserhati uses Neighbor Joining and Maximum Likelihood as phylogenetic methods. "*Both trees were constructed using the MEGA-X software [29], with parameters set to default values.*" (page 9 of 12). If you run the data through a program on the default you'll get an unrooted phylogenetic tree for both NJ and ML.

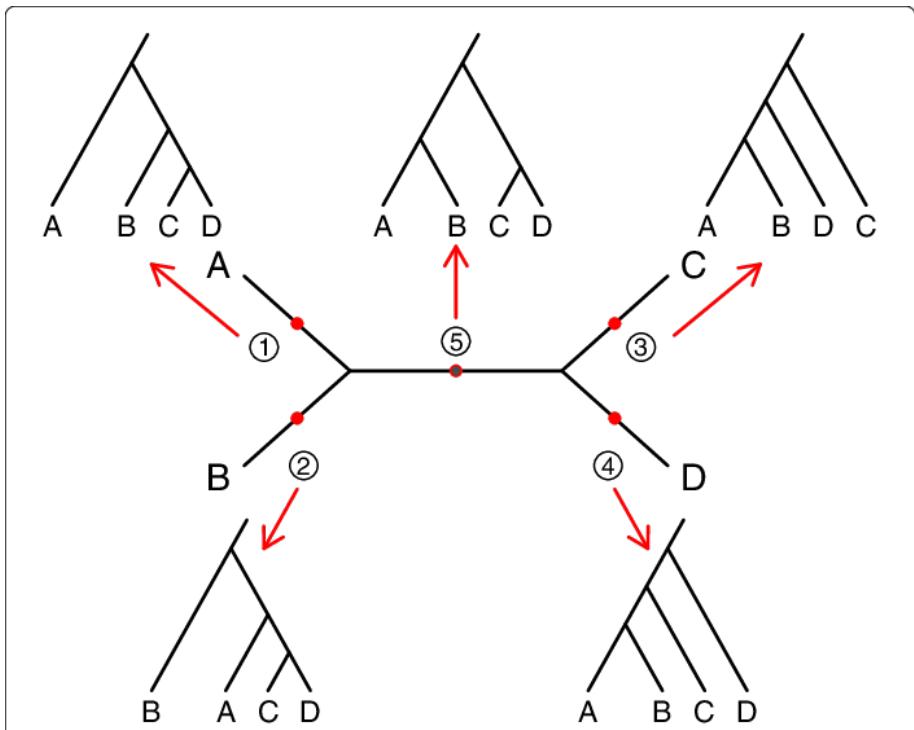


Figure 10. An unrooted tree of four Operational Taxonomic Units A B C D, and the five rooted trees associated with them. The place of the root determines the degree of relatedness.

An unrooted tree tells something but not much. With an unrooted tree mutual distances between the species are known, but not the way in which the species are related. And if you don't know how they are related, you can't classify properly.

Rooting is necessary for a good phylogenetic tree, and rooting is achieved by designating an outgroup. Here the bears are suitable as the outgroup to the superfamily Musteloidea, to find a proposal for a classification within the Musteloidea. Cserhati should have been able to tell the software to use an outgroup. Now Cserhati gets messy results, because he mistakes the efficient way the program plots an unrooted tree for the outcome of his tree. Cserhati fails to distinguish layout from result.

With the bears as an outgroup, both the NJ and ML analyzes show the red panda as a sister group to the remaining three families within the Musteloidea (See blog post 12 for further explanation).

There is a deeper problem here: a rooted phylogenetic tree provides a hierarchical classification. And no matter how you look at it, a hierarchical classification suggests evolution. That is why creationists prefer to use clustering, because clustering does not lead to a hierarchical classification.

IV C

3. The discordance between the mitochondrial DNA results and the genome shows that baraminology is useless.

Cserhati asserts this claim can be found in my Panda's Thumb post or in the discussion that follows. I can't find where that claim is made.

Searching for 'discordance' yields only one occurrence of 'discordance' a comment:

"The BMC Genomics paper admits that the results are tentative and that there may be differences between mtDNA and the whole genome due to nuclear-mitochondrial discordance."

That comment is written by Cserhati. No one else says anything about nuclear-mitochondrial discordance. (By the way, it's the CRSQ paper where Cserhati talks about nuclear-mitochondrial discordance, not the BMC Genomics paper; Cserhati should have known that.)

There is no evidence of nuclear-mitochondrial discordance in Cserhati's data. First, the UPGMA on WGKS and the UPGMA on mtDNA give the same higher level classification within the Musteloidea. Thus, the direct comparison of Cserhati's 'nuclear' and his 'mitochondrial' phylogenetic tree show no discordance in red panda classification. Second, Cserhati claims that the mtDNA sequences show more differences than the nuclear genome as represented by WGKS. That is not true. Table 2 here shows the mean and minimum of the correlations between all species pairs of bears and Musteloidea in the WGKS and mtDNA analyses. WGKS and mtDNA have about the same range in correlations between species (blog post 16). This means that there is no difference between genomic divergence between species as in WGKS and mtDNA divergence between species.

TABEL 2	minimum correlation	mean correlation
WGKS	0.666	0.838
mtDNA	0.751	0.816

IV D

4. Baraminology is a fictional construct, and there is nothing real behind it.

In the discussion that followed my item on The Panda's Thumb website, a certain Henry J commented: "Well, given that "baramin" is a fictional concept, why can't they (creationists) just make up something and leave it at that." This is the only time "fictional" appears in that post on The Panda's Thumb.

So, not baraminology, but the baramin are called 'fictional'.

Baramin as creationists use that concept has little to do with Genesis 1. "Bara" (ברא) is the verb 'to create'. Creationists have their own interpretation of the Hebrew word which is translated as 'kind', 'according to its nature' in the traditional Dutch translations. That word is *mîn*, and it occurs almost exclusively in this kind of context. The word is a collective, and indicates diversity:

Thus the biblical text emphasizes the diversity of life – plants and animals – with which God filled the sky, the sea, and the dry land he had created.

Consistent with the basic message of Genesis 1, the emphasis rests upon God's creation of life in all its abundance and diversity. (biologos)

The point is variety: 'as they are', 'as they live', namely herbivores, predators, frugivores, climbers and diggers. That is their 'nature'.

Baramin is a new word created by creationists to substitute a specific idea instead of the wide Biblical word 'kind'. Cserhati's interpretation of Genesis is creationist fiction: "*Of course, within an individual baramin species are related to one another That is because Genesis 1:11, 12, 21, 24, and 25 declare that kinds form reproductive communities.*" Do Genesis 1:11, 12, 21, 24, and 25 really say so? Is what Cserhati says found in Scripture?

The King James Version says:

11 And God said, Let the earth bring forth grass, the herb yielding seed, and the fruit tree yielding fruit after his kind, whose seed is in itself, upon the earth: and it was so.

12 And the earth brought forth grass, and herb yielding seed after his kind, and the tree yielding fruit, whose seed was in itself, after his kind: and God saw that it was good.

21 And God created great whales, and every living creature that moveth, which the waters brought forth abundantly, after their kind, and every winged fowl after his kind: and God saw that it was good.

24 And God said, Let the earth bring forth the living creature after his kind, cattle, and creeping thing, and beast of the earth after his kind: and it was so.

25 And God made the beast of the earth after his kind, and cattle after their kind, and every thing that creepeth upon the earth after his kind: and God saw that it was good.

Genesis says nothing whatsoever about species within a 'kind', or that 'kinds' form closed reproductive communities. That kind of detail is non-Biblical modern creationism. It takes a high degree of imagination to get the creationist baramin out of Genesis. It is misuse of the Bible.

All in all, baramin are extra-biblical fiction.

Baraminology has two components: a statistical component, clustering, and a fictional component, the baramin, holobaramin, monobaramin, apobaramin. There exists no way whatsoever to relate a statistical cluster to Genesis. That is the fiction part. Remains the question whether statistical clustering is in any way informative or useful in scientific classification of species. Genesis is not relevant there. Creationists spend much effort on clustering, seemingly in the confused preconception that clustering will find baramin.

Cserhati doesn't like my calling baraminology a waste of effort because it does nothing but repeat evolutionary taxonomy. Cserhati denies that baraminology repeats evolutionary taxonomy. Cserhati should pay more attention.

(i) In the article '*A survey of Cenozoic mammal baramins*' (Thompson & Wood 2018, cited by Cserhati), Thompson & Wood give the following result: "*Based on the successful analyses, we identified 59 putative holobaramins, 49 of which corresponded to families, seven to subfamilies (or portions of families), two to superfamilies (or multiple families), and one (Sirenia) to an infraorder (Table 3)*". Traditional evolutionary taxonomic groups are declared baramin.

- (ii) Lightner et al (2012) give the families of the mammals according to the modern molecular evolutionary classification as baramin.
- (iii) McLain et al (2018) arrive at the known dinosaur groups as baramin in feathered dinosaurs, after many unnecessary detours with repeated clustering.
- (iv) Cserhati & Carter (2020) give in their Table 1 what they call a known 'true cluster' in addition to their own clusters based on mtDNA and WGKS. That 'true cluster' is the family according to the classic, well-known, evolutionary, classification.

The animal kingdom is classified hierarchically, because the animal kingdom is structured hierarchically. The hierarchical structure of life makes a fiction of the whole idea that baramin could be found by clustering. Cserhati knows that: see the article by creationists Cserhati & Carter (2020), entitled '**Hierarchical clustering complicates baraminological analysis**'. Cserhati & Carter(2020) do not provide a solution to their problem. On the contrary. Eight significant quotes from Cserhati & Carter (2020):

- (i) *One major issue is the hierarchical structure of species relationships.*
(pg 64)
- (ii) *Yet the greatest problem facing baraminology might well be the hierarchical structure of life* (pg 66)
- (iii) *This simply means that cladistics can assemble any group of organisms into a hierarchy, regardless of whether they are truly related or not. And baraminology could be considered as just a mathematical expression of cladistics* (pg 66)
- (iv) *Baraminology was designed to elucidate the real patterns, but it had no way of making absolute judgments about the boundaries between kinds* (pg 66)
- (v) *Hierarchical clustering problems also present themselves when selecting species for baraminology studies. Species must be chosen at the right taxonomic level. If species are chosen too broadly, different clustering algorithms could put different species together which don't belong to the same baramin. Conversely, if species are chosen from one single baramin, the algorithms will still find clusters* (pg 67)
- (vi) *Thus, the promise of baraminology has not yet manifested itself. We do not yet have a way to make statistical determinations of group membership.* (pg 67)

(vii) *This study highlights one of the main problems of baraminology: the subjective classification of species into baramins* (pg 71)

(viii) *The fact that life follows a hierarchical pattern makes it hard to delimit the created kinds* (pg 71)

In those clear words, Cserhati & Carter (2020) conclude: **baraminology doesn't work**. There is no method to recognize baramin. Although Cserhati & Carter (2020) know that baraminology does not work, they stick to baramin. The conclusion that they are chasing a chimera turns out to be impossible for them. Baramin as fiction is Verboten.

In his 'Oorsprong' article, Cserhati claims that a baramin can be found by starting from two directions: start with a large group of beasts and divide them into clusters, or start with single species and make larger clusters. I have reproduced Cserhati's figure and his caption here.

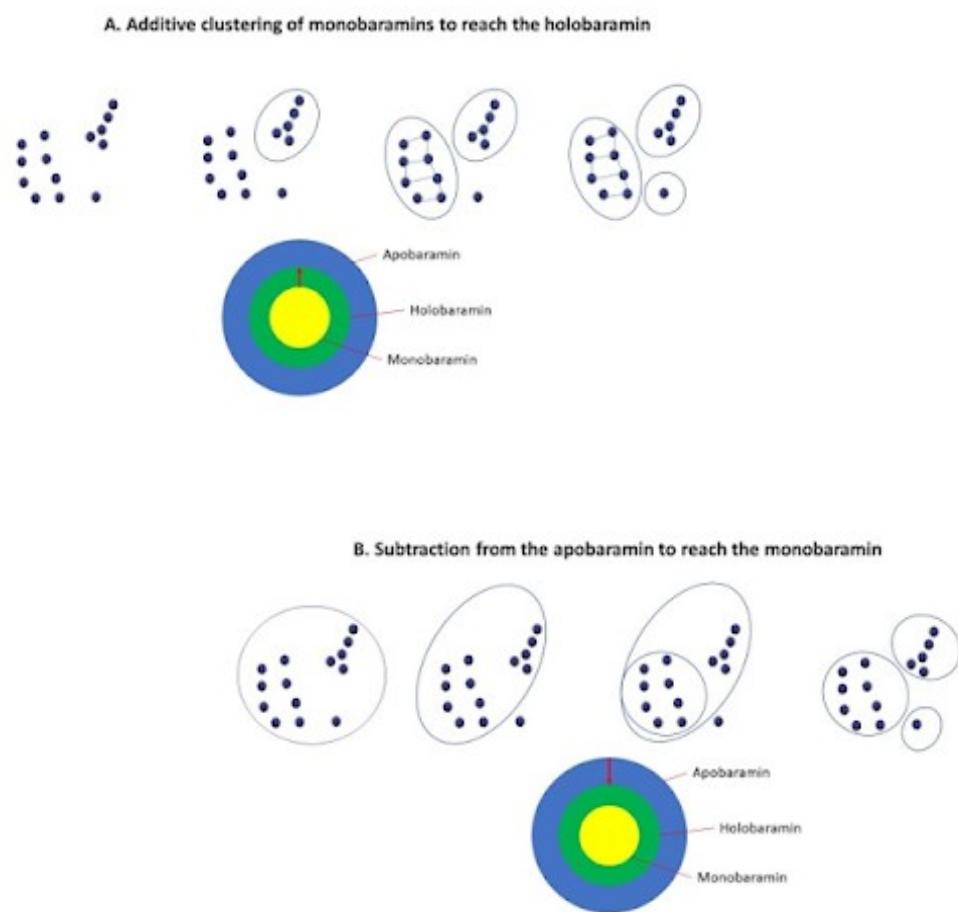


Figure 6. The process of successive approximation. A. Starting from a larger group of species (the apobaramin) we drill down to find the holobaramin. B. Starting from individual monobaramins, we use clustering to approximate the holobaramin. When both processes meet we have the holobaramin.

Species are not dots in a plane, and this recipe, first from bottom to top and then from top to bottom until the two directions meet, does not work – precisely because the animal kingdom is hierarchically structured.

As a guide I take how Cserhati & Carter (2020) indicate how baramin should behave: "*Species within the same kind should show continuity with one another on a morphological and a molecular level. They should also show discontinuity with all other species outside their kind.*" This is consistent with Hennigan (2010) and seems standard creationism to me.

I'll walk along the hierarchical structure of the animal kingdom, first from top to bottom, and then from bottom to top. I'll use the placental mammals as example. The hierarchical classification of the mammals is shown in Figure 7.

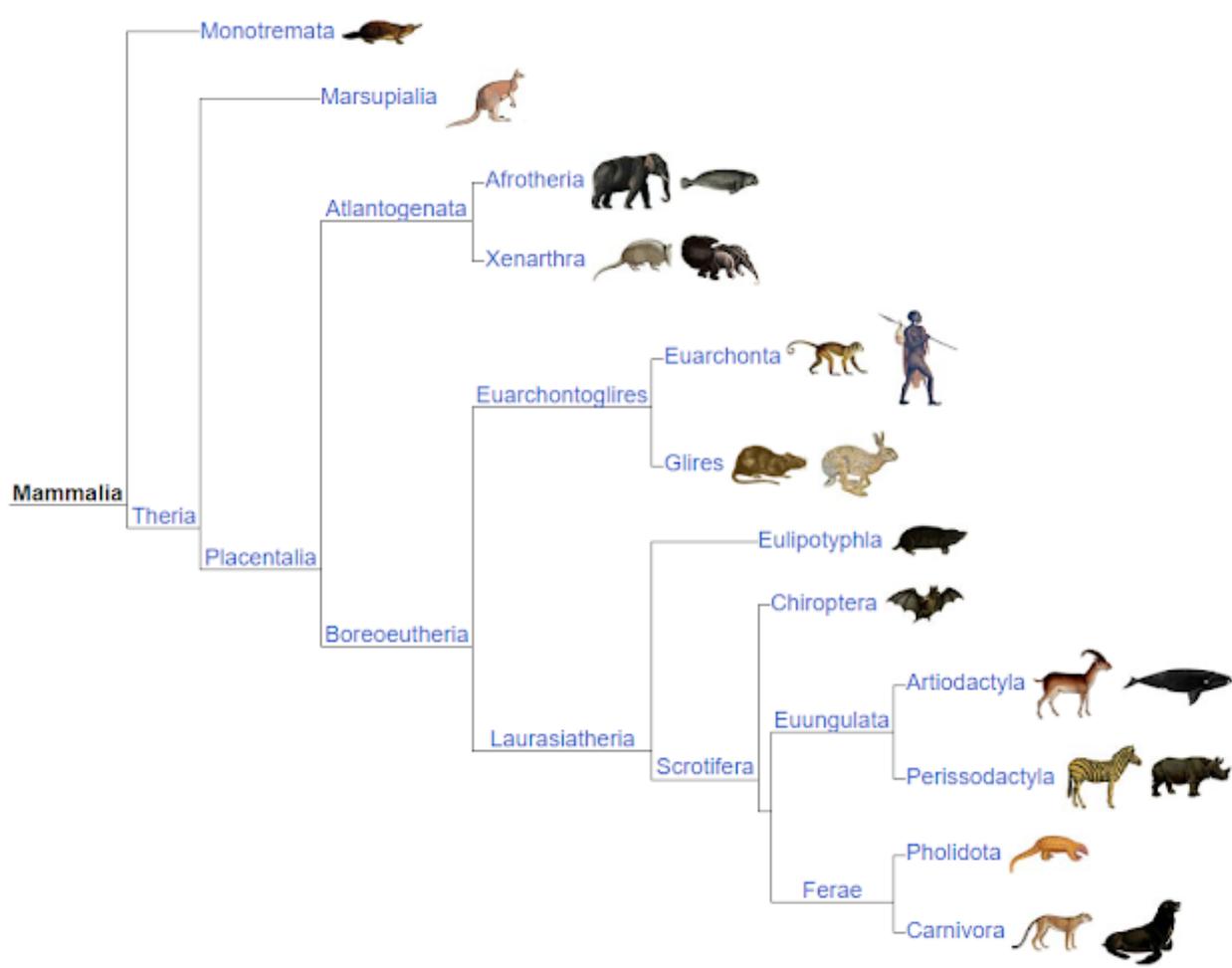


Figure 11. The hierarchical classification of the mammals. After Lv et al (2021), figure from en.wikipedia.org/mammal

The Theria are a group with internal continuity. Theria species have hair, differentiated dentition (incisors, canine, premolars, molars), have three ossicles in the middle ear, one bone in the lower jaw, and a relatively large brain. In addition, the Theria are molecularly genetically continuous relative to non-Theria.

In most characters the placental mammals are continuous with the marsupials and among themselves – hair, milk, etc, see under Theria. The placental mammals all have a placenta in reproduction, and are discontinuous in this characteristic with the marsupials. The placental mammals form a continuous group molecularly genetically to the marsupials, which form a different group molecularly genetically.

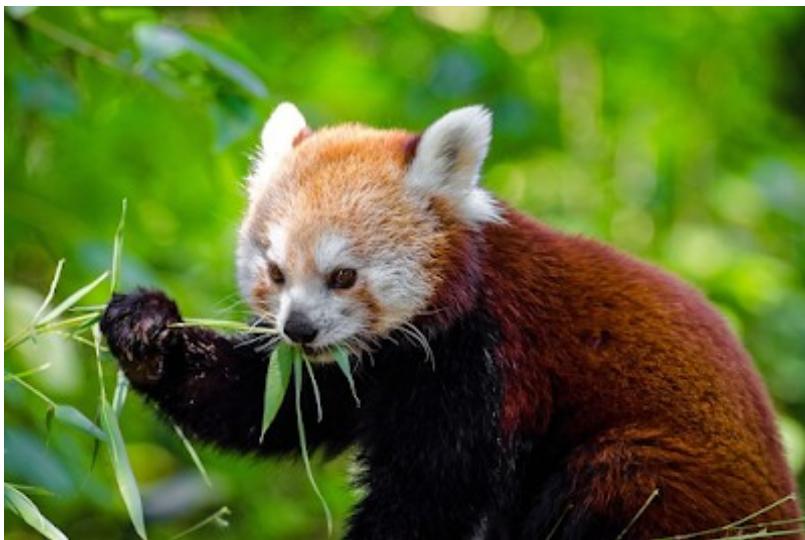
The placental mammals can be divided into two major groups, group Atlantogeneta and group Boreoeutheria, on molecular genetic grounds. They are continuous as placental mammal and Theria, and discontinuous on molecular grounds.

And so forth and so on. At least 15 steps down to the sea otter, in genus *Enhydra* of the subfamily Lutrinae, always with continuity at the level you look at and discontinuity below that level.

The other way, bottom to top, is also possible. Starting with the sea otter, merge the otters into the subfamily Lutrinae - a continuous group; merge the Lutrinae with the subfamily Mustelinae – together a continuous group; merge Lutrinae + Mustelinae with subfamily Gulolinae – together a continuous group; add the whole of these three subfamilies together, in three steps, with three more subfamilies, and there appears the family Mustelidae — a continuous group. Climb the stairs according to the layout, and at every level of the stairs and above you have continuity. At each level in a hierarchical format, there is a continuous group.

In a hierarchical format, it is possible to move from top to bottom and bottom to top without ever getting stuck halfway: there is no privileged level. Never a privileged level appears where top-to-bottom and bottom-to-top join together: one can always go up and down in a hierarchical classification, based on the existing hierarchical structure in the animal kingdom.

That is why the idea baramin is fiction.



V Assessment

The problem with Cserhati's two papers on the red panda is that it is such bad work. That is more important for the BMC Genomics article than for the CRSQ article: BMC Genomics is a scientific journal.

Cserhati neglects his own results: according to his own results the red panda forms an independent group of its own.

Cserhati prefers and adheres to his WGKS method without any justification: the UPGMA phylogenetic tree on the WGKS results clearly shows that WGKS is not a suitable method for classifying species – it gives artifacts – , but Cserhati does not even look at this UPGMA placement of the red panda.

Cserhati ignores his own statistical results: his silhouette plots indicate different numbers of clusters than he wishes to report.

Cserhati bases his main conclusion on WGKS clustering: but the silhouette plot indicates a Musteloidea cluster, not a Mustelidae+Ailuridae cluster.

Cserhati confuses the square layout of his unrooted phylogenetic NJ and ML trees with the phylogenetic trees itself.

Cserhati makes egregious errors in reading the literature; see his introduction to the BMC Genomics article and his Table 1 in the 'Oorsprong' article.

Cserhati does not cite recent scientific literature on the classification of the red panda.

How did papers as bad as these ever get through review and published?

<https://oorsprong.info/wetenschapper-lost-creationistisch-biosystematisch-raadsel-van-de-rode-panda-ailurus-fulgens-op/>

<https://evolutiebiologie.blogspot.com/2010/10/het-raadsel-van-de-rode-panda.html>

<http://pandasthumb.org/archives/2022/12/a-tale-of-two-papers.html>

<https://oorsprong.info/evolutionists-on-pandas-thumb-attack-well-respected-science-journal/>

Peng, R., Zeng, B., Meng, X., Yue, B., Zhang, Z., and Zou, F. (2017) The complete mitochondrial genome and phylogenetic analysis of the giant panda (*Ailuropoda melanoleuca*). *Gene* 397:76–83

Fulton, T.L., and Strobeck, C. (2007) Novel phylogeny of the raccoon family (Procyonidae: Carnivora) based on nuclear and mitochondrial DNA evidence. *Mol Phylogenet Evol.* 43:1171–1177 (2017 moet 2007 zijn)

Yu, L., and Zhang, Y.P. (2006) Phylogeny of the caniform carnivorae: evidence from multiple genes. *Genetica* 127(1–3):65–79

Sato, J.J., Wolsan, M., Minami, S., Hosoda, T., Sinaga, M.H., and Hiyama, K. (2009) Deciphering and dating the red panda's ancestry and early adaptive radiation of Musteloidea. *Mol Phylogenet Evol.* 53(3):907–22

Thompson, C., & T.C. Wood. 2018. A survey of Cenozoic mammal baramins. In Proceedings of the Eighth International Conference on Creationism, J.H. Whitmore (editor), pp. 217–221, A1-A83 (appendix). Creation Science Fellowship, Pittsburgh, PA.

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

Lightner, J.K. 2012 Mammalian Ark Kinds. *Answers Research Journal* 5:151–204

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

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

Lv, X., et al. 2021. Diverse phylogenomic datasets uncover a concordant scenario of laurasiatherian interordinal relationships. *Molecular Phylogenetics and Evolution* 157 (2021) 107065