Hannah Jones
Evolution
December 5, 2008

Abstract:
Cetacea is an order that contains many questions regarding
the evolutionary relationships between taxa. An anatomical
analysis and analysis of alpha lactalbumin, and 12s DNA
data is conducted to determine the evolutionary
relationships among the cetacean taxa. Morphological data
and molecular data are used to determine evolutionary
relationships among 13 different taxa both extinct and
extant. Morphological characters are derived and character
states are described to demonstrate homologies within
anatomical features. Molecular data is used to further the
complexity of the findings, and give more complete data
sets to the extant taxa. Using both of these data types,
cladograms are created demonstrating the data in a visual
manner. Conclusions supporting morphological and molecular
data are drawn in the form of two concluding trees.

Introduction:
        The morphological research on Cetacea, conducted by
students of Georgia Southern University, was supported with
funds provided by the National Science Foundation (DEB-
0640361 to J. Geisler). Using skulls, skull casts, and DNA
sequences of extinct and extant taxa, morphological data
indicating homologous features, is analyzed to show
evolutionary relationships linking the species within
Cetacea. Morphological and molecular data was collectively
analyzed, a method that had not been attempted previously.
In the past either one or the other were used. These
characters will be used to determine homologies using
cladistic analysis. The 50 morphological characters as
well as the 850 molecular characters will be used to map
and better understand homologous features between the
cetaceans.

Methods:
        There are 13 taxa represented in this report. In the
process of determining the evolutionary relationships
between the taxa, anatomical differences were described as
characters and varying states were defined. Then
observations of the various taxa were conducted and the
data was transferred to a matrix. The characters and there
states were then documented electronically and pictures

were taken to digitally enhance our information and insure repeatability and accuracy.

Analyzing the data, using cladistics programs, is the next step in determining the morphological history of the cetaceans I have been studying. The programs I have used are Winclada and TNT. Winclada is a useful program in creating cladograms; I used it initially when analyzing my qualitative morphological data (Nixon, K. 1999). You can not put actual numbers or ratios in the program. The initial tree produced is from the qualitative characters. This gives us a basis of comparison, and initially gives us feed back on the taxa in which their positions vary. TNT(Goloboff, Farris and Nixon,2008) is the more suitable program to use. This program allows us to manipulate the data easier and it accepts both quantitative and qualitative data allowing a more complete analysis. Within this program weighting, scaling, and branch support can be accomplished, providing more derived accurate trees.

Discussion:

Within winclada first put in the original matrix derived from the original characters. This cladogram is the discrete, unscaled, unweighted tree. Place *Pakicetus* first designating it as the outgroup, meaning it is the most primitive. This provides us with a network of trees that will be created based on the outgroup taxa. I do not agree with this cladogram because a few of the taxa such as *Xenorophoid* and *A. patrius,* we know to be extinct should be closer to the outgroup, and others such as *D.leucas, G. macrorhynchus, P. blainvillei, and P. gangetica* should be more distant from the outgroup. The position of the blow hole is one distinguishing characteristic that can be used, in confidence, to determine the relative evolutionary position of the taxa. This indicates the cladogram is incorrect in comparison with the data we know to be correct. Since the tree did not place *Xenorophoid* and *A. patrius* close to the outgroup we know it is inaccurate and we have conflicting data, or not enough data. This is the only tree found, and because it is not an acceptable tree we cannot draw any conclusions using only this data.

The cladogram shows the following taxa branching in order: Pakicetus, G. vogtlensis, D. leucas, G. macrorhynchus, P. blainvillei, P. gangetica, Xenorophoid, A. patrius, S. fluviatilis, T. truncatus, L. obliquidens, D. delphis, L. borealis.
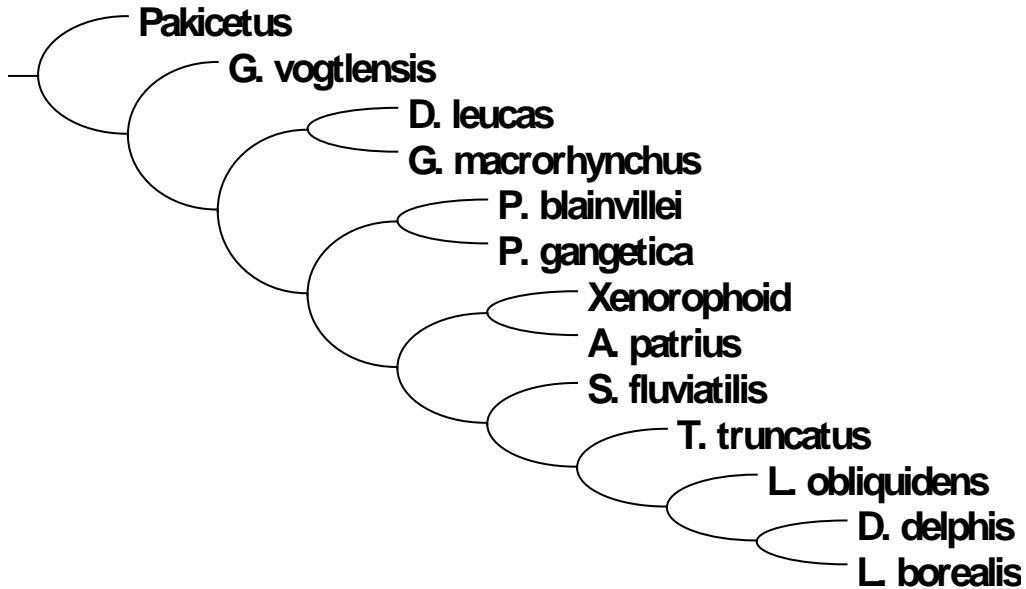
     The TNT program proved to be much more helpful. I still chose the most primitive taxa to be the outgroup. In my data set I was now able to use all of the data I gathered including the quantitative data. This analysis also includes the character scaling. Character scaling is a method of analysis that allows us to use quantitative and qualitative data in the same analysis. The formula allows the characters to be put into numbers that have the same maximum union cost for each character. The character states then have the same union cost between each character state. The formula to determine scaling is (x-min)/max-min. The x in this formula is the value of the character state for the taxa. The minimum value is the lowest value among the character state, meaning the maximum value is the highest value. This is the only tree found, but the cladogram is much more acceptable. *Pakicetus, G. vogtlensis, Xenorophoid, and A. patrius* are all evolutionally in the correct positions respectively. Other taxa such as *P. gangetica and P. blainvillei* resemble each other in that they have small skulls, the blowhole is in the same relative position on the skull, this allows me to conclude they are in good relative standings with the rest of the tree. *D. leucas and G. macrorhynchus* branch from the same node leading us to believe they are evolutionarily close. The shear size of these two taxa is a major character similarity and probably had a lot to do with the relative positions of the specimens. Reguardless I accept this tree as my best tree.

```
                    Pakicetus
                        G. vogtlensis
                            Xenorophoid
                            A. patrius
                            P. gangetica
                                P. blainvellei
                                    D. leucas
                                    G. macrorhynchus
                                    S. fluviatilis
                                        T. truncatus
                                            L. obliquidens
                                                D. delphis
                                                L. borealis
```
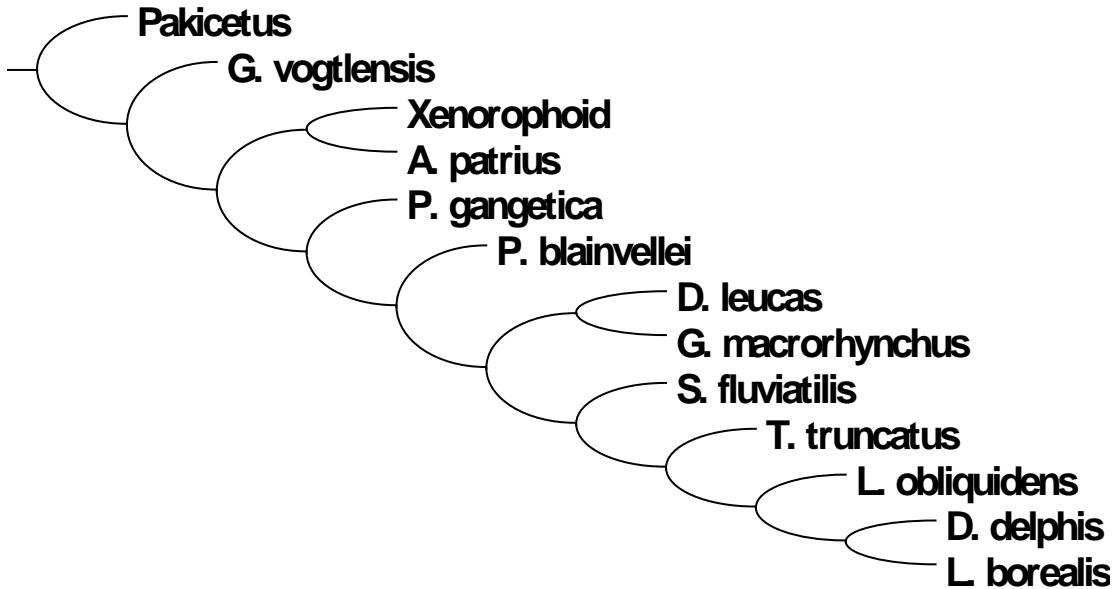
When I used implied weighting I tested K values 1-10. Between those K values my tree remained consistent, so I decided to keep the tree. Implied weighting tells us, with respect to our data sets, which characters are better fit than others. Only one tree was found, and it reflected the previous tree leading me to believe it was a reliable tree.

```
                    Pakicetus
                        G. vogtlensis
                            Xenorophoid
                            A. patrius
                            P. gangetica
                                P. blainvellei
                                    D. leucas
                                    G. macrorhynchus
                                    S. fluviatilis
                                        T. truncatus
                                            L. obliquidens
                                                D. delphis
                                                L. borealis
```
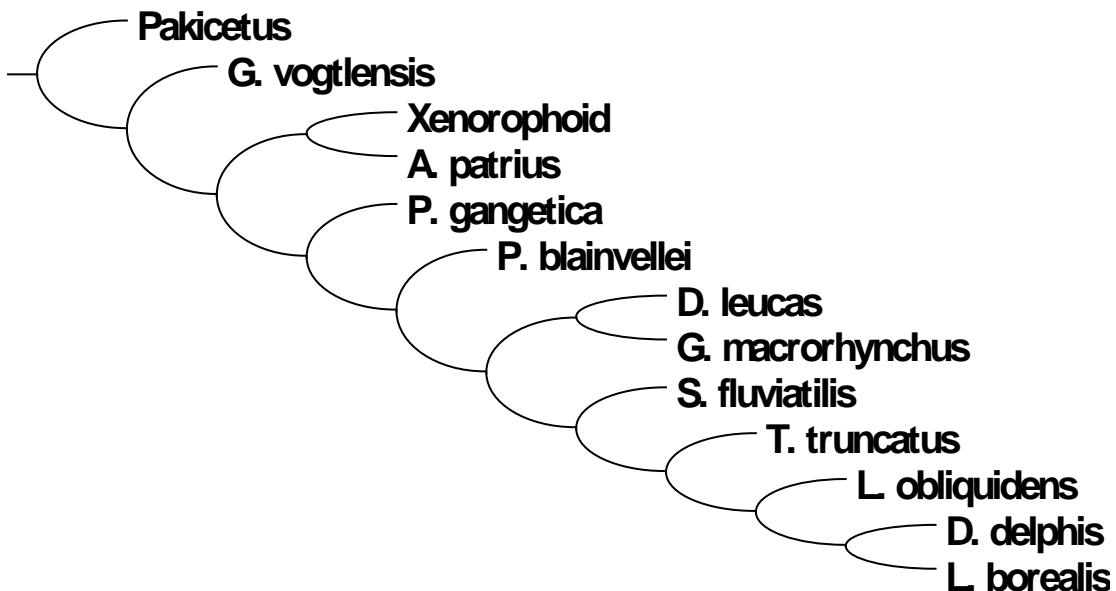
After analyzing the morphological data we received molecular data. The decision was made to apply both the morphological and the molecular data through TNT. The molecular data is placed at the end of the data set and run in TNT. Because so many more characters were added, the

cladogram became skewed. When the molecular data was run,
only two trees were received. These trees were far removed
form the trees I previously gained, and the first four
extinct taxa wee not located close to the out group. In the
first node however the maximum length was 0 and the data
became skewed as a result. I then changed the settings and
collapsed the node 0= X max. After collapsing the node
other trees were produced. From these cladograms I can
manipulate the trees and analyze the data in different ways
depending on our tree.

    This is the strict consensus as well as the majority
rule tree. This tree has become inaccurate when molecular
data is added. *L. borealis* is displaced the greatest. It
should be removed from the base of the cladogram, and
replaced in the more extant. *D. leucas, G. macrorhynchus,
and S. fluviatillis* are in the same relative position to
one another. This leads me believe that they are closely
related.

```
         ┌─ Pakicetus
       ┌─┤
       │ └─ L. borealis
     ┌─┤
     │ └─ Xenorophoid
     │ └─ A. patrius
     │   └─ G. vogtlensis
     │     └─ D. delphis
     │       └─ D. leucas
     │         └─ G. macrorhynchus
     │           └─ S. fluviatillis
     │             └─ P. blainvillei
     │               └─ L. obliquidens
     │                 └─ G. macrorhynchus
     │                 └─ T. truncatus
```
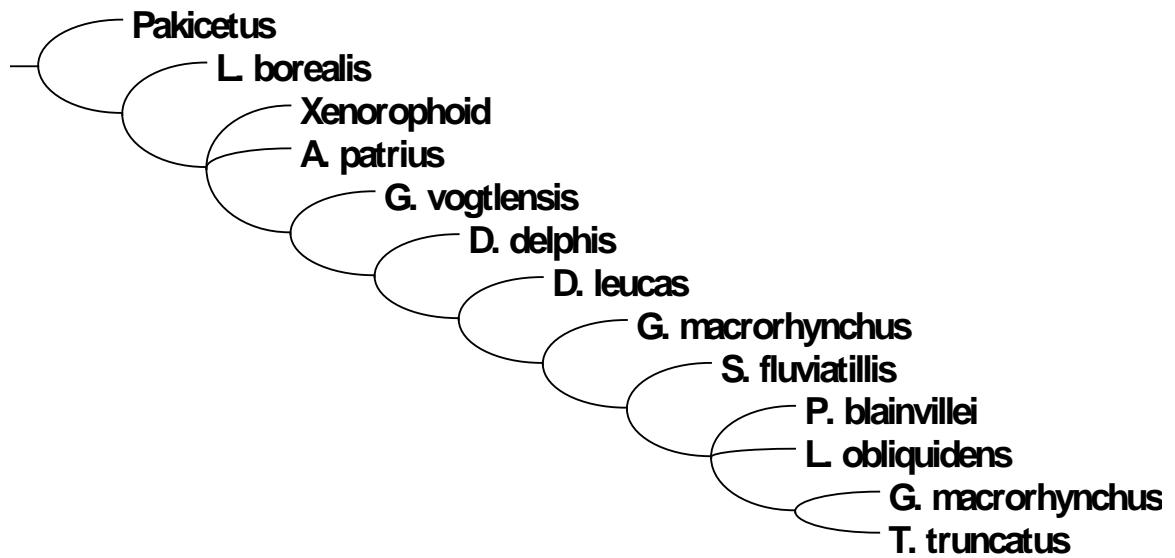
    The agreement subtree for this data set is my favorite
tree. The extinct taxa are in the right order in respect
to the outgroup, and the others follow a logical ordering
with more detailed relationships shown at the most derived
portion of the cladogram. It did not eliminate any taxa
which is a good sign of accuracy.

Pakicetus
G. vogtlensis
Xenorophoid
A. patrius
P. gangetica
P. blainvillei
D. leucas
G. macrorhynchus
S. fluviatillis
L. obliquidens
L. borealis
D. delphis
T. truncatus
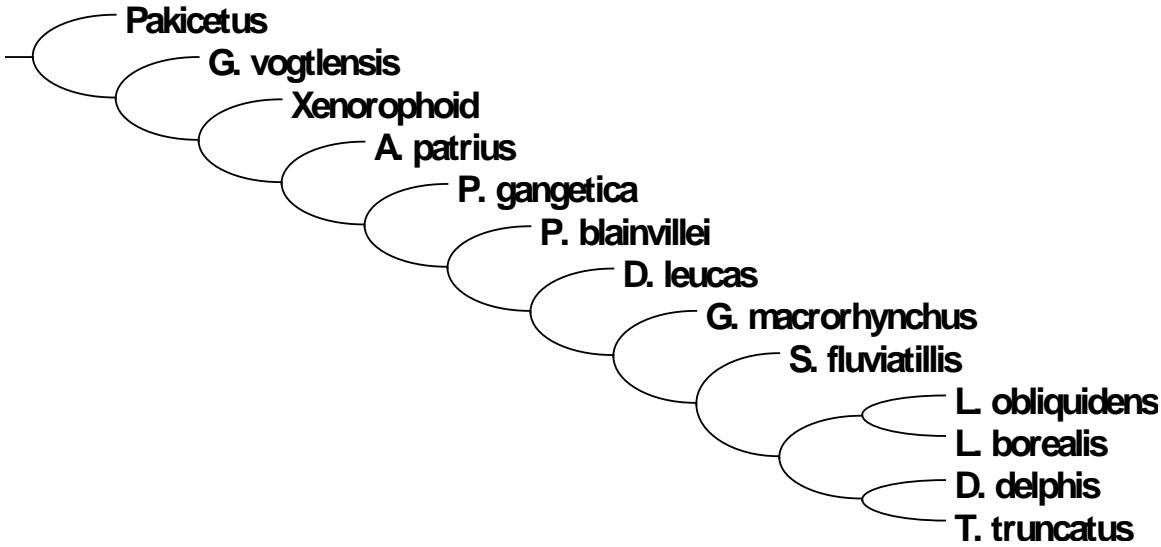
Branch support is the final method used to direct the assembly of the cladogram and describe which taxa are more similar and which are more different. When conducting branch support you take the length of the node, minus the length of the shortest most parsimonious tree.

Pakicetus
G. vogtlensis
Xenorophoid
A. patrius
P. gangetica
P. blainvillei
D. leucas
G. macrorhynchus
S. fluviatillis
L. obliquidens
L. borealis
D. delphis
T. truncatus
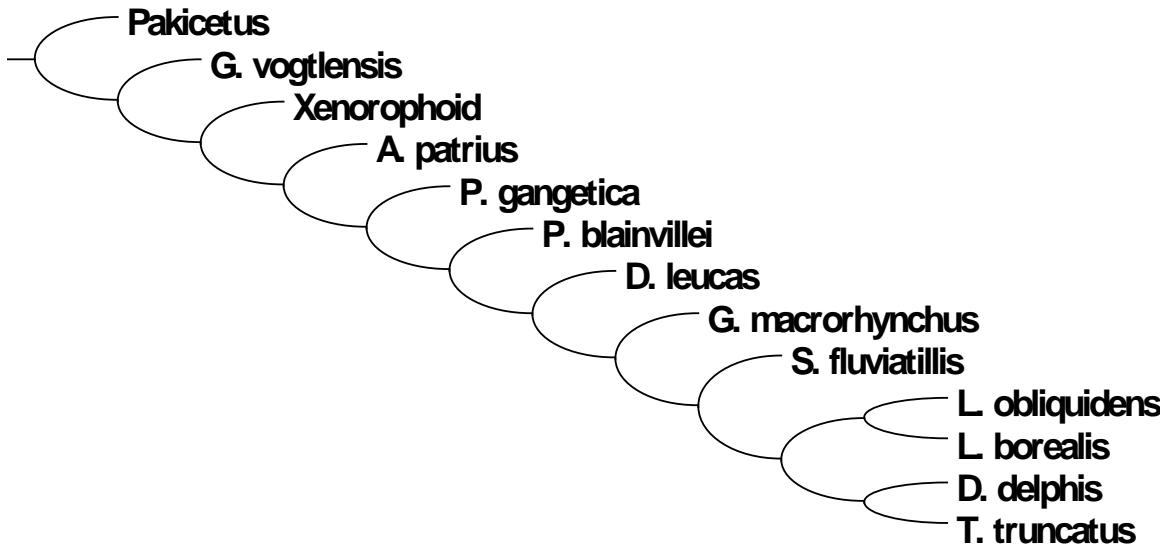
The bigger the number, when doing branch support is the most supported length in the tree. This means the relationship between *G. macrorhynchus* and all the other taxa is the most supported by the characters. *A. patrius*, *Xenorophoid* and all other taxa are also supported greatly. From the node including *P. gangetica* and all the other taxa to the node including *A. patrius, Xenorophoid* and all other taxa there is a significant jump in the branch support indicating a better supported node when the characters supported by *A. patrius, Xenorophoid* and all other taxa are included. The nodes displaying the most complexity are

T.truncatus and D. delphis with a branch support of 1.725, and *L. borealis and L. obliquidens* with a branch support of 1.632. This indicates a close relationship between these taxa, which is supported by high branch support numbers. Adding *S. fluviatillis* increases the branch support to 2.157 indicating when you add the data from *S. fluviatillis* the portion of the clade is more supported than what it originally was. The drop in the branch support number from 3 *(G. macrorhynchus + all taxa above) to (D. leucas + all taxa above) indicates* a drop in data support for the taxa relationships displayed within the cladogram.

Branch Support for the strict consensus tree analyzing both morphological and molecular data.

Taxa included	Node value
T.truncatus + D. delphis	1.725
L. borealis + L. obliquidens	1.632
S. fluviatillis + all taxa above	2.157
G. macrorhynchus + all taxa above	3
D. leucas + all taxa above	.6
P. blainvillei + all taxa above	.039
P. gangetica + all taxa above	.039
A. patrius + Xenorophoid + all taxa above	2
G. vogtlensis + all other taxa above	0

Pakicetus
G. vogtlensis
Xenorophoid
A. patrius
P. gangetica
P. blainvellei
D. leucas
G. macrorhynchus
S. fluviatilis
T. truncatus
L. obliquidens
D. delphis
L. borealis

Pakicetus
G. vogtlensis
Xenorophoid
A. patrius
P. gangetica
P. blainvillei
D. leucas
G. macrorhynchus
S. fluviatillis
L. obliquidens
L. borealis
D. delphis
T. truncatus
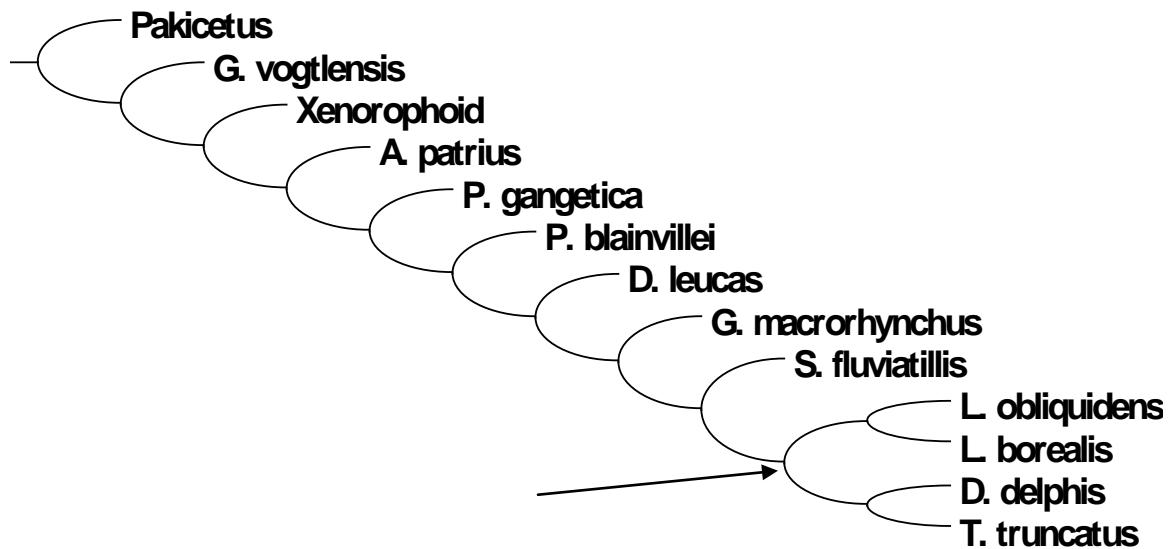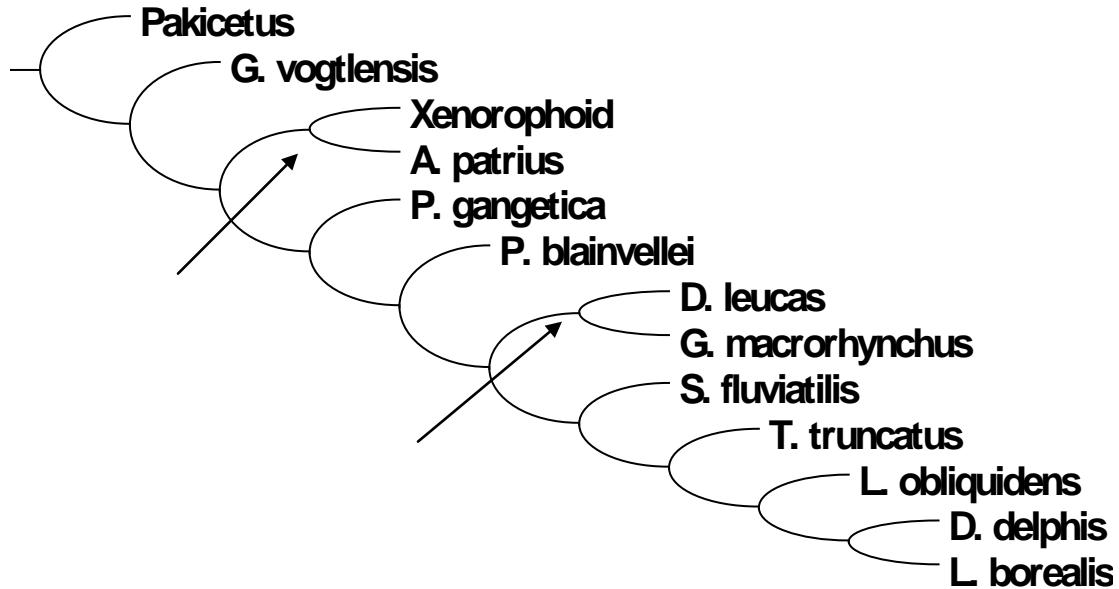
Conclusions:

  In conclusion, portions of each of the best cladograms depict the tree that is best supported. In the first cladogram, analyzing the morphological data, the taxa Pakicetus, G. vogtlensis, P.gangetica, P. blainvillei, and S. fluviatllis do not change positions relative to the other taxa. This leads me to believe they are in the right place with respect to the matrix. Other taxas however group together in one or the other cladograms. *L. borealis and L. obliquidens; and T.truncatus and D. delphis* are in nodes together off a more central node in the cladogram displaying both the morphological and the molecular data.

I like this position of these taxa because it shows more
complexity within the taxa and variance in the characters.
This variance is most likely due to the molecular data
bias. The extinct taxa do not have molecular data sets and
therefore are not as complex as the extant ones. With that
being said the opposite is true for the extinct taxa in the
first cladogram due to the more obvious differences in the
characters. Therefore the anatomical analysis is more
conclusive in determining morphology of extinct taxa, and
molecular data is more relevant to the extant taxa.

Materials:
The materials used for acquiring morphological data:
-Caliper (used for measuring quantitative characters)
-Camera (documentation of characters and states)
-Excel (organization of data)
-Ruler and string (used to determine accurate measurements)
-Calculator (to calculate ratios within continuous
characters)
-Casts of skulls

The materials used to acquire DNA data:
-Michael "Rocky" McGowen, from UC Riverside (incite on DNA
data, and supplier of 12s (ribosomal gene) DNA data)
-ClustalW website: http://align.genome.jp/ (used to acquire
alph lactalbumin DNA data, and align DNA data sequences)

The materials used to analyze morphological and DNA data:
-Winclada (used in the initial analysis of qualitative
morphological characters)
-TNT (used for the analysis of all of the morphological
data discrete and continuous; and DNA data alpha
lactalbumin and 12s)

Character list:
1. Tooth count (Total). Figure 1. a) 0-20 (0); b) 21-50 (1);
   c) 51-90 (2); d) 91-120 (3).
2. Extension of teeth vs. length of rostrum. Measuring from
   antorbital notch to the most anterior tip of the rostrum.
   .100-.659 (0); .660-.909 (1); .910-1.20 (2).
3. Shape of mandible. Figure 2. a) Y-shaped (0); b) V-shaped
   (1).
4. Rostral width vs. length.0-.20 (0); .21-.40 (1); .41-.50
   (2); .51-.70 (3); .71-.90 (4)
5. Length of rostrum vs. length of skull. 0-.50 (0); .51-.60
   (1); .61-.70 (2).

6. Length of mandible [mm]. Figure 3. a) 0.00-350.99 (0); b) 351.00-450.99 (1); c) 451.00-700.00 (2).
7. Length of mandible suture [mm]. Figure 4. a) 0.00-60.99 (0); b) 61.00-300.00 (1).
8. Shape of nasals. Figure 5. a) Square/rectangular (0); b) tear drop (1); c) triangular (2); d) rounded (3); e) other (4).
9. Dorsal profile of rostrum. Figure 6. a) Convex (0); b) concave (1).
10. Habitat. Fresh (0); saline (1); both (2).
11. Posterior, dorsal extension of maxilla into dorsal lateral portion of external bony nares. Figure 7. a) None (0); b) intermediate (1); c) dramatic (2).
12. Size of zygomatic process. Dorsoventrally and transversely. Figure 8. a) Thin (0); b) thick (1).
13. Geologic age. Early Eocene (0); middle Eocene (1); Oligocene (2); recent (3).
14. Dorsal position of external bony nares on the rostrum. Figure 9. a) Front (0); b) middle (1); c) top (2).
15. Teeth structure. Figure 10 a) multicusped (0); b) uniform (1).
16. Tooth shape. Figure 11 a) Bulky (0); b) peg-like (1).
17. Mesorostral canal located on dorsal side of rostrum. Figure 12 a) Open (0); b) slightly fused (1); c) mostly fused (2).
18. Nuchal crest. Figure 13 a) Absent (0); b) Present throughout (1).
19. Sagittal crest on occiput. Figure 14 a) Lack (0); b) present (1).
20. Lambdoidal crest. Figure 15 a) Slightly elevated (0); b) sharp (1).
21. Extension of nasals above braincase. Figure 16 a) No (0); b) yes (1).
22. Shape of temporal fossa. Figure 17 a) Rounded (0); b) elongate (1).
23. Condyle texture. Figure 18 a) Smooth (0); b) rough (1).
24. Mandibular notch. Figure 19 a) Lack (0); b) presence (1).
25. Medial contact of pterygoids along sagittal plane. Figure 20 a) No (0); b) yes (1).
26. Palate. Figure 21 a) Flat (0); b) vaulted (1).
27. Maxillary crest. Figure 22 a) Lack (0); b) presence (1).
28. Condyle shape, measured transversly. Figure 23 Longest/biggest axes. a) Y (0); b) X (1).

29. Occipital overlap of frontal. Figure 24 a) None (0); b) slight (1); c) dramatic (2).
30. Antorbital notch. Figure 25 a) Lack (0); b) slight (1); c) deep (2).
31. Squamosal length from anterior to posterior vs.length of skull. .00-.25 (0); .26-.39 (1); .40-.50 (2); .51-.60 (3).
32. Lateral wall groove on mandible. Figure 26 a) Lack (0); b) presence (1).
33. Mandibular foramen. Figure 27 Medial wall. a) Straight (0); b) rounded (1); b) Pointed anteriorly (2).
34. Mandibular fossa. Figure 28 a) Uniform (0); b) expands posteriorly (1).
35. Shape of posterior ventral edge of vomer, sutured to basisphenoid. Figure 29 a) Square (0); b) Deep V (1); c) Shallow V (2).
36. Ventral posterior tip of pterygoid. Figure 30 a) Points medially (0); b) laterally (1).
37. Premaxilla widens as approaches nasals. No(0), does not widen; Yes(1), widens.
38. Presence of premaxillary foramen. No(0); yes(1)
39. Paroccipital process is the most posterior part; extends posterior to exoccipital condyles. Strongly extends(0); slightly extends(1); no extension (2).
40. Intertemporal constriction. Yes(0), constricted; no, not constricted(1).
41. Dorsal condoyloid fossa. Deep(0); shallow(1).
42. Nasals touch. yes(0), no(1).
43. Occiput shape. concave(0); convex(1).
44. Shape of jugular notch. wide(0); narrow(1).
45. Exoccipital condyle flush with occiput. no(0); yes(1).
46. Exoccipital condyle in line with foramen magnum. yes(0); no(1).
47. Antero nasal bulge. absent(0), present(1).
48. Presence of beak formed from premaxilla. Lack(0); present(1).
49. Coronoid ramus. Lack(0); present(1).
50. Exoccipital condyle flush with occiput. Yes(0); no(1).

Data Matrix including continuous and discrete characters and the continuous characters are scaled.

Continuous:

Pakicetus	?	?	?	?	?	?	?
S._fluviatilis	0.611	0.566	0.395	0.447	0	0.103	0.103
D._leucas	0	0.508	1	0.053	0.272	0.107	0.333

```
G._macrorhynchus 0      0      1.405 0      0.423 0.054 0.256
D._delphis       0.278 0.449 0.479 0.377 0.117 0.047 0
P._blainvillei   1      0.588 0      0.807 0.114 0.605 0.590
L._obliquidens   0.522 0.719 0.709 0.160 0.073 0      0.205
L._borealis      0.311 0.447 0.566 0.324 0.117 0.052 0
G._vogtlensis    0.044 1      0.525 0.254 1      0.999 ?
P._gangetica     0.444 0.154 0.128 1      0.316 0.891 1
Xenorophoid      ?      0.699 0.226 0.484 ?      ?      0.538
A._patrius       ?      ?      ?      ?      ?      ?      ?
T._truncatus     0.333 0.639 0.572 0.258 ?      ?      0.103
```

Discrete:
```
Pakicetus
?4?0210000?11??1010?01?0???????????20?0?00?
S._fluviatilis
10021132112011000101002102121111111121110110
D._leucas
11012132101001011100000212120011011210110100
G._macrorhynchus
11110132100001111100002212121111011210101000
D._delphis
14011032112000000111000102010??010112111100
P._blainvillei
020111321120010101001012221211110102001?011
L._obliquidens
10110032111010000101000112101101011210111100
L._borealis
14011032111010001111001102100001011210111100
G._vogtlensis
04012111000001010100100?10?001001020001000
P._gangetica
04000132012000010001100220011??11010101000000
Xenorophoid
?4010122???0010101?000012??????0002000?000
A._patrius
?4??0022?????10?110?00?1????11??0??00000000
T._truncatus
14011032110011010101001100??011?101121011110
```

References:

"Cladistics, The International Journal of the Willi Hennig Society": Goloboff, P., Farris, J., & Nixon, K. 2008. TNT, a free program for phylogenetic analysis. Cladistics 24:774-786.

Geisler, J. and McKenna, M. 2007. New Mesonychian Mammal from Mongolia. *Acta   Palaentologica*. 52 (1), p.210-212.

Geisler, J. and Sanders, A. 2003. Morphological Evidence for the Phylogeny of Cetacea.   *Journal of Mammalian Evolution*. 10(1/2), p.23-129.

May-Collado, L. and Agnarsson, I. 2006. Cytochrome *b* and Bayesian inference of whale   phylogeny. *Molecular Phylogenetics and Evolution*. 38, p. 344-354.

"Multiple Sequence Alignment by ClustalW". GenomeNet. Kyoto University   Biomechanics Center. 1991. *http://algin.genome.jp*.

Nixon, K. 1999. WinClada ver. 0.9.99m8.1. *cladistics.com*.

O'Leary, A. and Uhen, M.D. 1999. The time origin of whales and the role of behavorial   changes in terrestrial-aquatic transition. *Paleobiology*. 25(4), p. 534-556.

Figure 1: Tooth count
Fig 1 a (D. leucas) [0-20]

Fig 1 b (D. delphis) [21-50]

Fig 1 c (S. fluviatillis) [51-90]

Fig 1 d (P. blainvellei) [91-120]
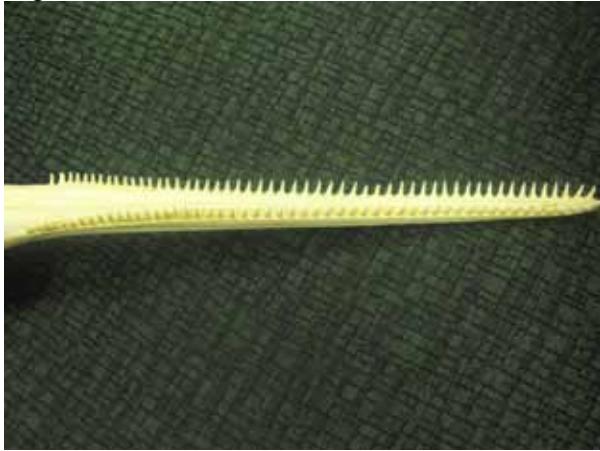
Figure 2: Shape of mandible
Fig 2 a (P. blainvellei) [Y-shaped]

Fig 2 b (G. macrorhynchus) [V-shaped]

Figure 3: Length of mandible
Fig 3 a (L. borealis) [0.00-350.99]

Fig 3 b (P.gangetica) [351.00-450.99]
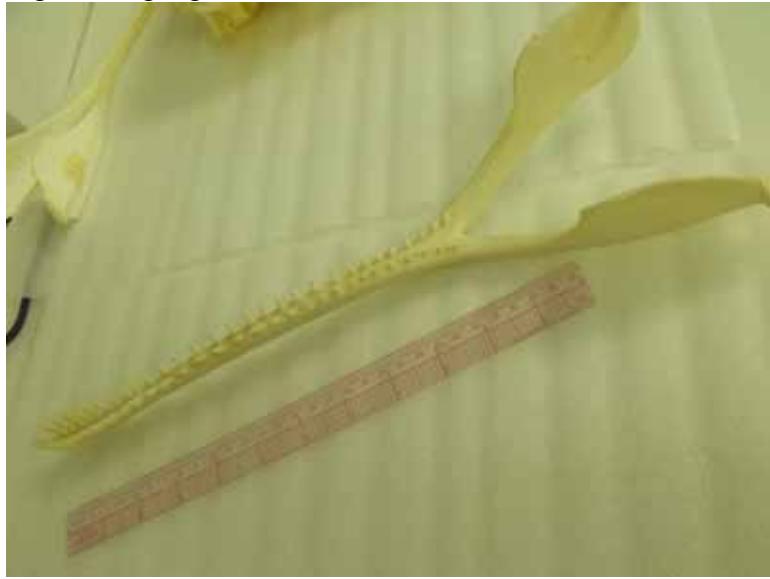
Fig 3 c (G. vogtlensis) [451.00-700.00]

Figure 4: Length of mandible suture
Fig 4 a (P. blainvellei) [0.00-60.99]

Fig 4 b (L. obliquidens) [61.00-300.00]
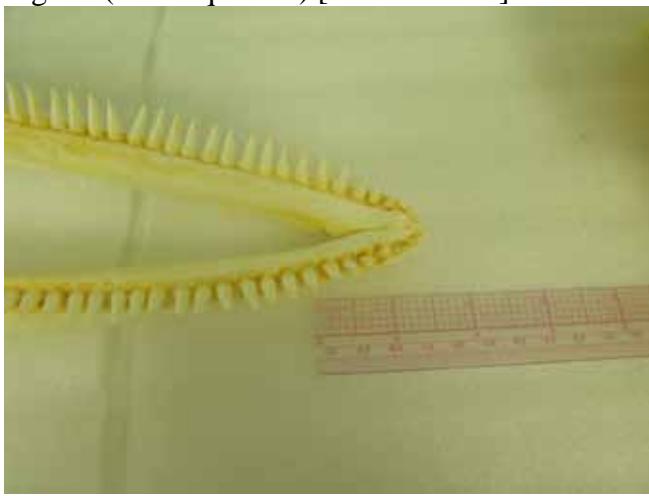
Figure 5: Shape of nasals
Fig 5 a (L. obliquidens)[Square/rectangular]
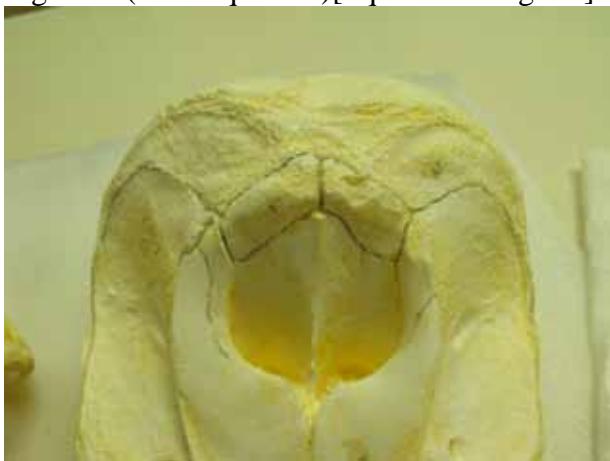
Fig 5 b (G. macrorhynchus)[teardrop]
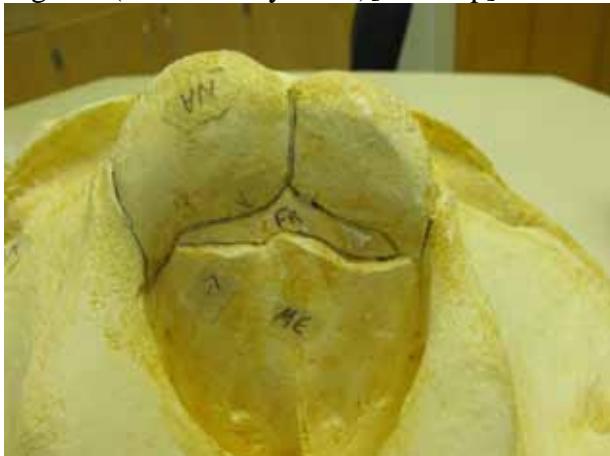
Fig 5 c (P. blainvellei) [triangular]

Fig 5 d (S. fluviatillis) [rounded]

Fig 5 e (D. delphis) [other]

Figure 6: Dorsal profile of rostrum
Fig 6 a (P. blainvellei) [convex]

Fig 6 b (G. macrorhynchus) [concave]

Figure 7: Posterior, dorsal extension of maxilla into dorsal lateral portion of external bony nares.

Fig. 7 a (G. macrorhynchus) [none]

Fig 7 b (D. leucas) [intermediate]

Fig 7 c (S. fluviatillis) [dramatic]

Figure 8: Size of zygomatic process
Fig 8 a (L. obliquidens) [thin]

Fig 8 b (P. gangetica) [thick]

Figure 9: Dorsal position of external bony nares
Fig 9 a (Pakicetus) [front]

Fig 9 b (G. vogtlensis) [middle]

Fig 9 c (L. borealis) [top]

Figure 10: Teeth structure
Fig 10 a (G. vogtlensis) [multicusped]

Fig 10 b (L. obliquidens) [uniform]

Figure 11: Tooth shape
Fig 11 a (G. macrorhynchus) [bulky]

Fig 11 b (L. borealis) [peg-like]
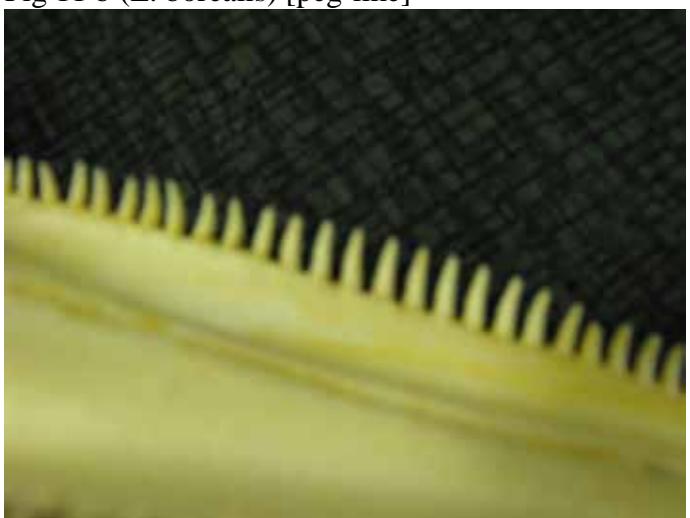
Figure 12: Mesorostral canal
Fig 12 a (T. truncatus) [open]

Fig 12 b (D. leucas) [slightly fused]

Fig 12 c (S.fluviatillis) [fused]

Figure 13: Nuchal crest
Fig 13 a (G.vogtlensis) [absent]

Fig 13 b (T. truncates) [present throughout]

Figure 14: Sagittal crest on occiput
Fig 14 a (D. leucas) [lack]
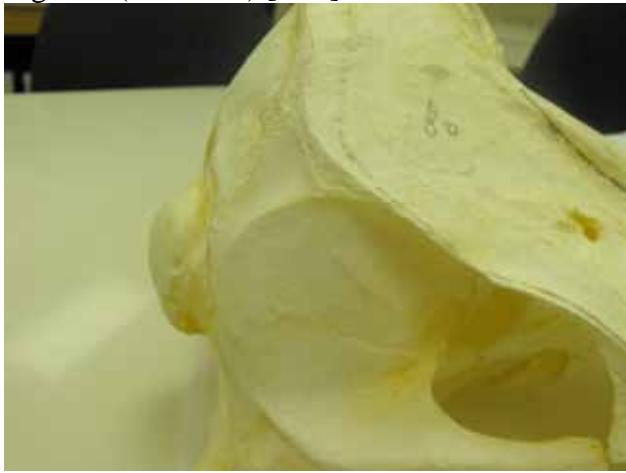
Fig 14 b (S. fluviatillis) [present]
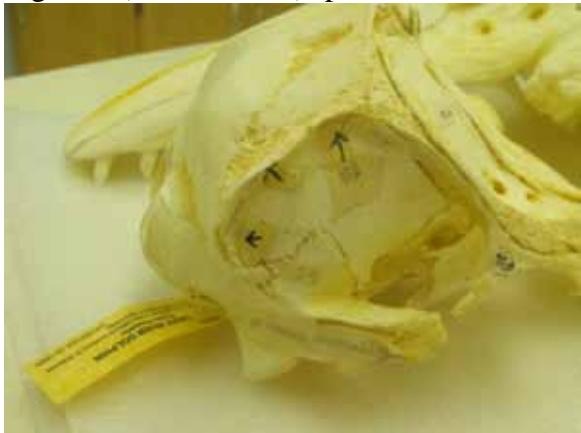
Figure 15: Lambdoidal crest
Fig 15 a (L. Obliquidens) [slightly elevated]

Fig 15 b (S. fluviatillis) [sharp]

Figure 16: Extension of nasals above braincase
Fig 16 a (L.obliquidens) [no]

Fig 16 b (G. macrorhynchus) [yes]

Figure 17: Shape of temporal fossa
Fig 17 a (A. patrius) [rounded]

Fig 17 b (L. borealis) [elongate]

Figure 18: Condyle texture
Fig 18 a (L. obliquidens) [smooth]
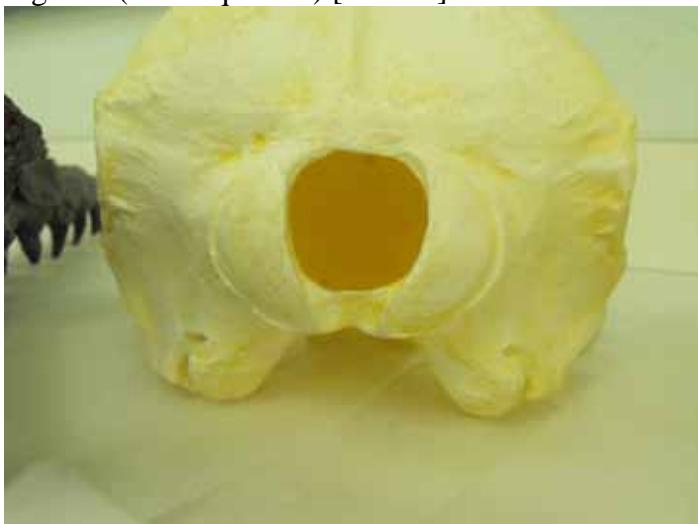
Fig 18 b (G. macrorhynchus) [rough]
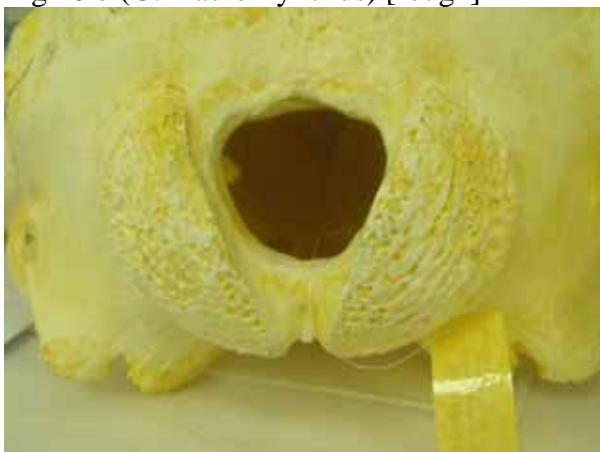
Figure 19: Mandibular notch
Fig 19 a (P. gangetica) [lack]

Fig 19 b (S. fluviatillis) [present]
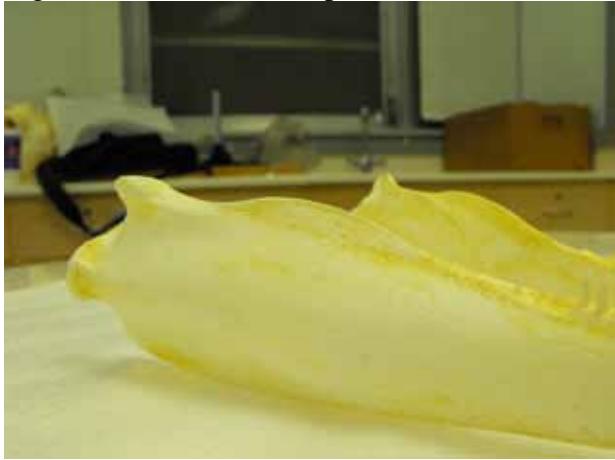
Figure 20: Medial contact of pterygoids along sagittal plane
Fig 20 a (G. macrorhynchus) [no]

Fig 20 b (L. borealis) [yes]

Figure 21: Palate
Fig 21 a (S. fluviatilis) [flat]

Fig 21 b (D. leucas) [vaulted]

Figure 22: Maxillary crest
Fig 22 a (L. borealis) [lack]

Fig 22 b (P. gangetica)[present]

Figure 23: Condyle shape
Fig 23 a (L. borealis) [Y]

Fig 23 b (G. vogtlensis) [X]

Figure 24: Occipital overlap of frontal
Fig 24 a (D. leucas) [none]

Fig 24 b (P. blainvellei) [slight]

Fig 24 c (G. macrorhynchus)[dramatic]

Figure 25: Antorbital notch
Fig 25 a (G. vogtlensis) [lack]

Fig 25 b (D. delphis) [slight]

Fig 25 c (P. gangetica) [deep]

Figure 26: Lateral wall groove on mandible
Fig 26 a (S. fluviatillis) [lack]

Fig 26 b (P. gangetica) [present]

Figure 27: Mandibular foramen
Fig 27 a (P.gangetica) [straight]

Fig 27 b (G. vogtlensis) [rounded]

Fig 27 c (L. borealis) [pointed anteriorly]

Figure 28: Mandibular fossa
Fig 28 a (P. gangetica) [uniform]
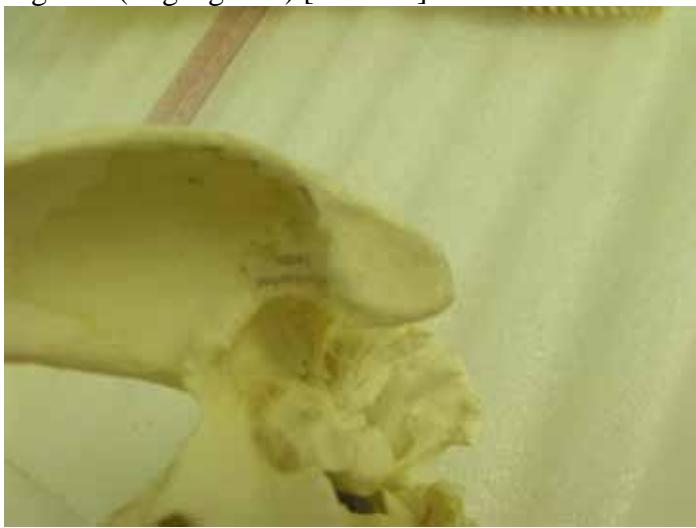
Fig 28 b (D. leucas) [expands posteriorly]

Figure 29: Shape of posterior ventral edge of vomer, sutured to basisphenoid
Fig 29 b (P. gangetica) [square]

Fig 29 c (P. blainvillei) [deep V]
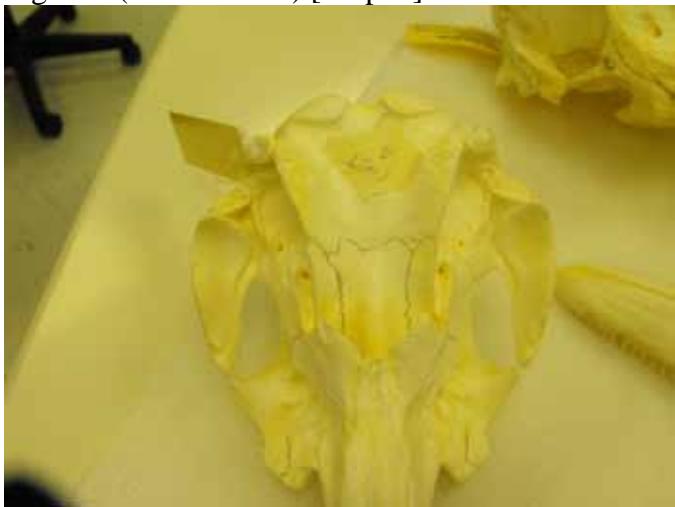
Fig 29 a (D. delphis) [shallow V]

Figure 30: Ventral posterior tip of pterygoid
Fig 30 a (L. borealis) [points medially]

Fig 30 b (S. fluviatillis) [points laterally]