A general criterion for translating phylogenetic trees into linear sequences

Proposal (474) to South American Classification Committee

In most of the books, papers and check-lists of birds (e.g. Meyer de Schauensee 1970, Stotz et al. 1996; and other taxa e.g. Lewis et al. 2005, Haston et al. 2009, for plants), at least the higher taxa are arranged phylogenetically, with the "oldest" groups (Rheiformes/Tinamiformes) placed first, and the "modern" birds (Passeriformes) at the end. The molecular phylogenetic analysis of Hackett et al. (2008) supports this criterion. For this reason, it would be desirable that in the SACC list not only orders and families, but also genera within families and species within genera, are phylogenetically ordered.

In spite of the SACC's efforts in producing an updated phylogeny-based list, it is evident that there are differences in the way the phylogenetic information has been translated into linear sequences, mainly for node rotation and polytomies.

This is probably one of the reasons why Douglas Stotz (Proposal #423) has criticized using sequence to show relationships. He considers that we are creating unstable sequences with little value in terms of understanding of relationships. He added, "We would be much better served by doing what most taxonomic groups do and placing taxa within the hierarchy in alphabetical order, making no pretense that sequence can provide useful information on the branching patterns of trees. This would greatly stabilize sequences and not cost much information about relationships".
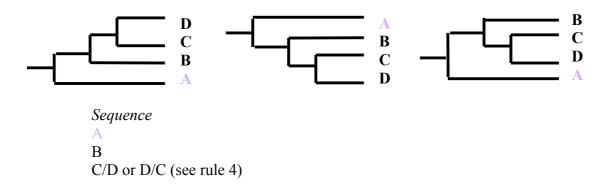
However, we encourage creating sequences that reflect phylogeny as much as possible at all taxonomic levels (although we agree with Douglas Stotz that the translation of a phylogenetic tree into a simple linear sequence inevitably involves a loss of information about relationships that is present in the trees on which the sequence is based). Furthermore, the idea that sequences cannot provide phylogenetic information is not restricted to species but is also pertinent to other taxonomic levels, such as genera, families, etc.

In this context, it appears necessary to unify the criteria used by all members of the Committee for translating phylogenetic trees into linear sequences. In an attempt to find a general criterion, we propose the following approach, to be applied over all taxonomic levels, based on five rules. This approach is valid only for those trees (or clades, e.g. *Campylorhamphus* in the Dendrocolaptinae tree -see Examples-) including all or nearly all the known species:
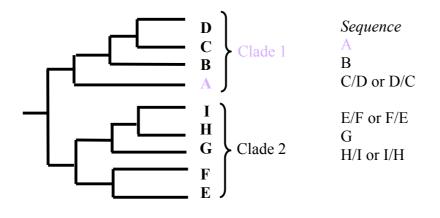
1) First-splitting taxon:
The taxon that splits first (presenting the lesser number of ancestors, that is, internal nodes*) is placed at the top of the sequence (taxon A). The same rule is applied to the next taxa, following the order of the branching pattern.

*Each node with descendants represents, in a rooted tree, the inferred most recent ancestor of those descendants. Nodes can be rotated without altering the relationships between taxa.
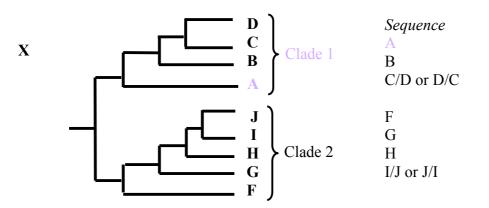
Sequence
A
B
C/D or D/C (see rule 4)

2) *Sister clades:*

For sister clades, the clade containing the taxon that splits first (taxon A, clade 1) is arranged first. After it, the next clade is listed (clade 2).

Sequence
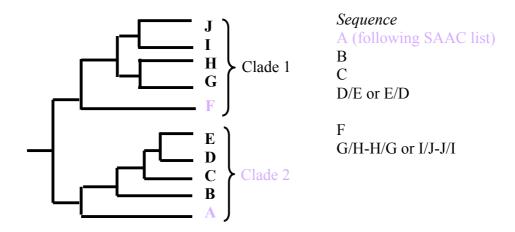A
B
C/D or D/C

E/F or F/E
G
H/I or I/H

When the number of nodes is the same between the first-splitting taxon in both sister clades (taxon A and F in X; A/B and F/G in Y), the less diverse clade is arranged first (clade 1).

Sequence
A
B
C/D or D/C

F
G
H
I/J or J/I

Y

Tree (Clade 1, Clade 2)	Sequence
E, D, C — Clade 1; B, A — Clade 1; K, J, I, H — Clade 2; G, F — Clade 2	A/B or B/A C D/E or E/D F/G or G/F H I J/K or K/J

Sequence
A/B or B/A
C
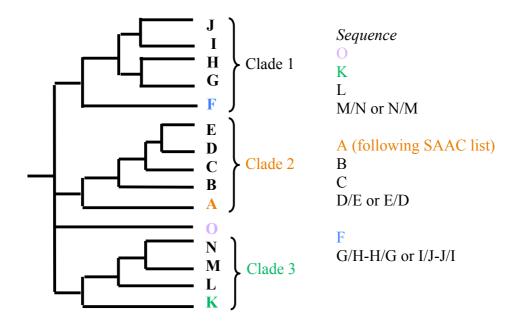D/E or E/D

F/G or G/F
H
I
J/K or K/J

When the number of nodes is the same between the first-splitting taxon (taxon F and A) in equally diverse sister clades, the order of the SACC list is followed for selecting the first taxon to be placed (e.g. taxon A, clade 2).

Sequence
A (following SAAC list)
B
C
D/E or E/D

F
G/H-H/G or I/J-J/I

3) Polytomies:

In the case of polytomies, the taxon that splits first (taxon O), or the clade containing the taxon that splits first, is placed at the top of the sequence. Then, if the number of nodes between the first-splitting in sister clades is equal (taxon F, A and K), the less diverse clade is arranged first (clade 3); if these clades are equally diverse (clade 1 and 2), the SACC list is followed to select the first taxon (e. g. taxon A, clade 2).

```
            J                           Sequence
            I                           O
            H  } Clade 1                K
            G                           L
            F                           M/N or N/M

            E
            D                           A (following SAAC list)
            C  } Clade 2                B
            B                           C
            A                           D/E or E/D

            O                           F
            N                           G/H-H/G or I/J-J/I
            M  } Clade 3
            L
            K
```
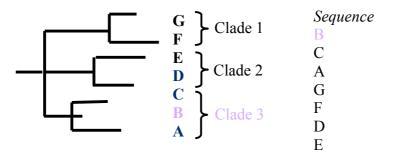
4) Sister taxa:

At the tips of the tree, when branch lengths are scaled (in phylograms), the taxon that has the shortest branch (that is, having the lowest amount of genetic change) is placed first (taxon B).

```
                                        Sequence
            B                           B
            A                           A


                                        Sequence
            C                           B
            B                           C
            A                           A
```
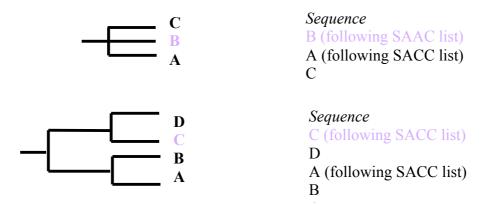
In polytomies, the clade with the taxon that has the shortest branch is placed first (taxon B, clade 3). Then, the clade with the shortest branched taxon follows (taxon G, clade 1) and so on (clade 2).

```
            G  } Clade 1                Sequence
            F                           B
            E                           C
            D  } Clade 2                A
            C                           G
            B  } Clade 3                F
            A                           D
                                        E
```

In equal terminal branches (in phylograms) or unscaled branch lengths (in cladograms), the SACC list is followed.

```
            B                           Sequence
                                        A (following SAAC list)
```

4

Sequence
B (following SAAC list)
A (following SACC list)
C

Sequence
C (following SACC list)
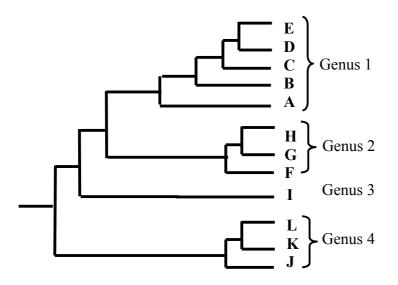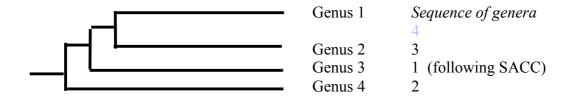D
A (following SACC list)
B

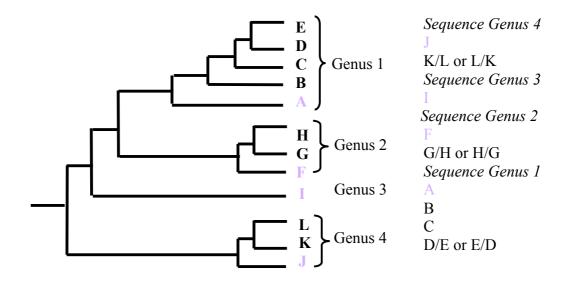## 5) *Trees comprising different levels of taxa*

In trees that analyze different levels of taxa, the rules explained above are applied within each taxonomic level. The figure shows a tree of genera and species.

For ordering genera, the species should be removed from the tree and the tree analyzed under the rules previously suggested.

	Genus 1	*Sequence of genera*
		4
	Genus 2	3
	Genus 3	1 (following SACC)
	Genus 4	2

Then, the species must be ordered within genera under the rules previously suggested.

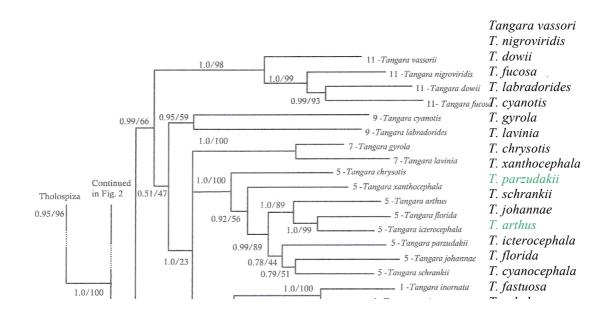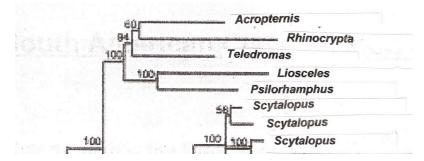|		Sequence Genus 4
E		J
D		
C	Genus 1	K/L or L/K
B		Sequence Genus 3
A		I
		Sequence Genus 2
H		F
G	Genus 2	G/H or H/G
F		Sequence Genus 1
I	Genus 3	A
		B
L		C
K	Genus 4	D/E or E/D
J		

EXAMPLES

In green, taxa whose position is defined following the order in the SACC list.

Tangara **(Sedano and Burns 2010)**

Sequence

Tangara vassori
T. nigroviridis
T. dowii
T. fucosa
T. labradorides
T. cyanotis
T. gyrola
T. lavinia
T. chrysotis
T. xanthocephala
T. parzudakii
T. schrankii
T. johannae
T. arthus
T. icterocephala
T. florida
T. cyanocephala
T. fastuosa

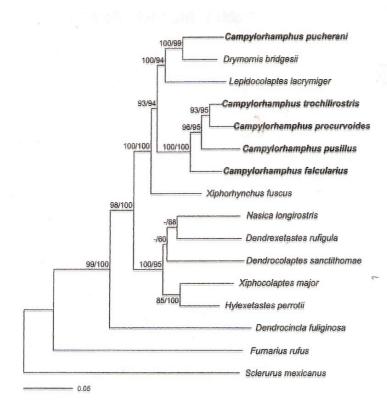**Rhinocryptidae (Ericson et al. 2010)**

Psilorhamphus
Liosceles
Teledromas
Acropternis
Rhinocrypta   7

Dendrocolaptinae (Claramunt et al. 2009)

　　In this example, it is possible to make the sequence of genera (except for *Sittasomus, Glyphorynchus* and *Deconichura* that were not included) and the sequence of species of *Campylorhamphus.*

Sequence of genera

Dendrocincla
Xiphorhynchus
Campylorhamphus
Lepidocolaptes
Drymornis
Drymothoxeres (pucherani)
Hylexetastes
Xiphocolaptes
Dendrocolaptes
Nasica
Dendrexetastes

Sequence of species of Campylorhamphus

C falcularius
C. pusillus
C. trochilirostris
C. procurvoides

Literature Cited

8

Claramunt, S., E. P. Derryberry, R. T. Chesser, A. Aleixo, and R. T. Brumfield. 2010. Polyphyly of Campylorhamphus with the description of a new genus for C. pucherani. Auk 127: 430-439.

Ericson, G.P., Olson S.L., Irestedt, M., Alvarenga H. and Fjeldså, J. 2010. Circumscription of a monophyletic family for the tapaculos (Aves: Rhinocryptidae): Psilorhamphus Kin and *Melanopareia* out. Journal of Ornithology DOI. 10.1007/s10336-009-0460-9.

Hackett, S.J, Kimball, R.T., Reddy, S., Bowie, R.C.K., Braun, E.L., Braun, M.J., Chojnowski, J.L., Cox, W.A., Han, K., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., and Yuri, T. 2008. A phylogenomic study of birds reveals their evolutionary history. Science, 320: 1763-1768.

Haston, E., Richardson, J.E., Stevens, P.F., Chase, M.W. and Harris, D.J. 2009. The Linear Angiosperm Phylogeny Group (LAPG) III: a linear sequence of the families in APG III. Botanical Journal of the Linnean Society, 161, 128–131.

Lewis, G.P., Schrire, B., Mackinder B. and Lock M. (eds). 2005. Legumes of the World. Royal Botanic Gardens, Kew.

Meyer de Schauensee, R. 1970. A guide of the birds of South America. Academy of Natural Science of Philadelphia. Pennsylvania.

Sedano, R.E., and Burns, K.J. 2010. Are the Northern Andes a species pump for Neotropical birds? Phylogenetics and biogeography of a clade of Neotropical tanagers (Aves: Thraupini). Journal of Biogeography 37: 325–343.

Stotz, D.F., Fitzpatrick, J.W., Parker, T.A.III and Moskovits, D.K. 1996. Neotropical birds: ecology and conservation. University of Chicago Press, Chicago.

General Literature

Avise, J.C. 2004. Molecular markers, Natural History and Evolution. Chapman and Hall, New York.

Hillis, D. M., Moritz, C. and Mable, B.K. 1996. Molecular Systematics. Sinauer Assoc., Inc. Sunderland, Massachusetts.

Manuel Nores (Centro de Zoología Aplicada); Noemí Cristina Gardenal, Paula Rivera, Raúl González Ittig, (Cátedra de Genética de Poblaciones y Evolución); María Jimena Nores (Instituto Multidisciplinario de Biología Vegetal). CONICET-Universidad Nacional de Córdoba. Argentina.

December 2010