

— Cladogram and Information —

比較生物学に於ける情報の「要約」

How is information integrated into a coherent summary? (p. 183)

(1) 2-taxon Cladogram (p. 183 ff.)

Cladistic Analysis に於ける「情報」とは何か? 前節のはじめで、「Component」の定義を

Component: Unit of information corresponding to a statement of general synapomorphy

と与えた。だから、ここでいう「情報」とは少なくとも [General] Synapomorphy と関わりがあることは明らかである。では具体的に何なのか? Synapomorphy とは、ある派生形質 [状態] の共有を意味しているから、Synapomorphy に関する情報とは、ある派生形質がどの程度共有されているか、又はどの程度出現 (occurrence) するかということである。ある形質が派生的であるか否かは独立にテストできる仮説である。従って「情報」は次のように定義される。

Information: Pattern of Character Occurrence

つまり、形質が何であるか (gill, lung, eye, ...) とは全く関係なく、その形質の出現パターンだけが Cladistic Analysis に於ける情報となり得る。だから、たとえば

2種 A, B について、(+) が形質の存在 (+) と欠除 (-) species A B

が右図のようになっているならば、形質 1 について character 1 + -

1) A には存在し、 2 - +

2) B には存在しない、 3 + +

というのが、その形質のもつ情報である。 ~~ここでいう~~ 4 - -

~~形質の存在 (+) と欠除 (-) とは、Apomorphy 或は Plesiomorphy とは関係なく考える必要はない。只、その出現パターンだけが問題なのである。~~

2-taxon の場合、形質の出現パターンは前表の 4通りしかない。したがって、その 4形質は個々の形質というよりむしろその出現パターンを示す形質の集合と考えられる。そこでこの形質の集合を "Character-

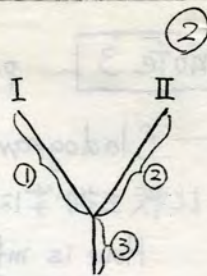
Type" と呼ぶ。 2-taxon の character-type species A B C D

は右図の 4通り (A-D) である。 character I + - + -

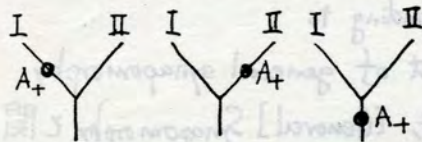
type とは information-type に他ならない。 II - + + -

次に、上で述べた Character-type (Information) は、Cladogram に於てどのように表現されているか？ Cladogram に於ける 形質の出現 或は 欠除が、internode (line) 上で  $A^+$  或は  $A^-$  の

2-taxon Cladogram



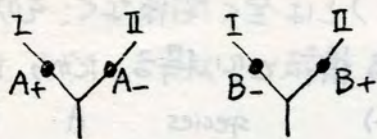
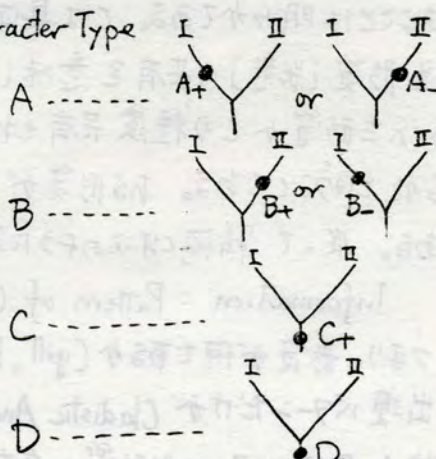
方に表わされるとして、2-taxon cladogram の場合、右図に示される 3本の line があることになる。これら が意味するものは次の通りである。(cf. Fig. 3.13)



左の2つは、形質  $A^+$  が I または II に出現したことを示す。最も右は形質  $A^+$  が I と II の両方に出現したことを示す。

従って、前頁の 4 Character Types は、右の如く示される。特に、A、B タイプは 2つの方法で示されるが、どちらかが秀れているという理由はない。そこで、これら 2つを併記して

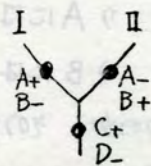
Character Type



と書けば、極めて正確な表現になるのである。

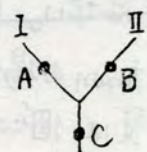
以上 4つの Character-Types をまとめて書けば、右図になる。そこには、形質の出現に7-ンつまり

- + : 形質の存在 (Positive Occurrence)
- : 形質の欠除 (Negative Occurrence)



(cf. Fig. 3.14.2)

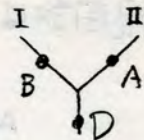
が正確に示されている。この図に於ては Positive 及び Negative Occurrences は同等の地位をもつ。これでは存在する (+) の形質に着目して、Information を表現するようになるか？ これは右上の図から、 $A^-$ 、 $B^-$ 、 $D^-$  を省いたもので、右下の如



(cf. Fig. 3.14.1)

になる (+ は明らかであるから省略した)。この図は A、B、C の形質が、どの taxon に存在するかを示しているのである。(p. 84 の議論も参照) 同様に

欠除している(-)の形質に着目して、左図の表現とする(-は省略)こともできる。+に着目するか-に着目するかは全く主観的な選択であるが、少なくとも直感的に、+に着目した方が解釈が容易であることは確かである。とにかく、これは主観(美観?)の問題である。



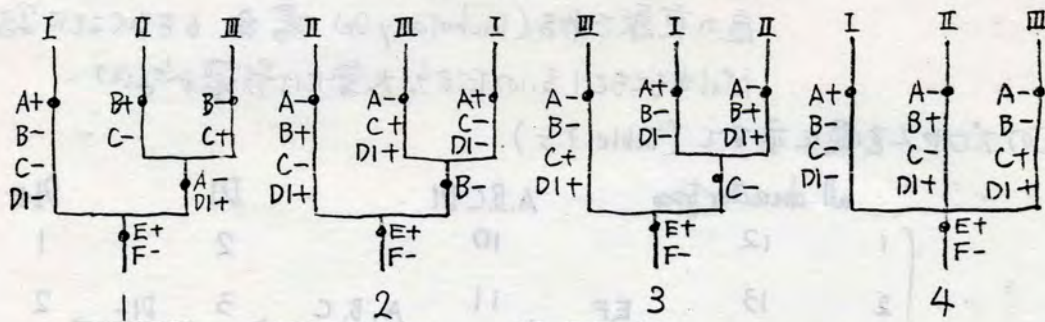
(3) 3-taxon Cladogram (p. 184 ff.)

2-taxonの場合、Cladogramは1つしかないが、3-taxonの場合4つのcladogramがあるから、一組の形質があるとき、このcladogramを選ぶか、またその選ぶ基準をどうするかという問題が生じてくる。

3-taxonの場合、可能なCharacter-Typeは次の8つである。

いま、A, B, C, D1, E, F (Table 3.3) の6形質を考えると、下の4 cladograms が可能である。(Fig. 3.16)

	A	B	C	D1	D2	D3	E	F
I	+	-	-	-	+	+	+	-
II	-	+	-	+	+	-	+	-
III	-	-	+	+	-	+	+	-



問題は、上の4つのcladogramのどれを選ぶか? またその選択基準をどうするか? である。直感的に見て1のcladogramがよさそう(p. 185~186の議論)であるが、はっきりとした基準を与えることができるか?

One way to judge each possibility is simply to count the number of character occurrences. (p. 186)

つまり、それぞれのcladogramについて positive (+) 及び negative (-) occurrences がいくつあるかを数えあげるのである。そして、

The difference in occurrences indicates the degree to which the various cladograms integrate, or generalize, the original information. (p. 188)

であるから、最も出現回数の少ないcladogramが最良(即ち情報を最もコンパクトにまとめている)と判断されるのである。上の1~4のそれぞれについて、そ

の出現回数 (Table. 3.4) 次のようになる。

Cladogram	positive						negative						total
	A+	B+	C+	D+	E+	F+	A-	B-	C-	D-	E-	F-	
1	1	1	1	1	1	0	1	2	2	1	0	1	12
2	1	1	1	2	1	0	2	1	2	1	0	1	13
3	1	1	1	2	1	0	2	2	1	1	0	1	13
4	1	1	1	2	1	0	2	2	2	1	0	1	14

従って、Cladogram 1 はその total occurrences が最も小さいから、最良と判断される。これは直感的に選んだものと一致する。ではこの出現回数の差は、何に由来するのか？ これを調べるため、いくつかの Character Types の削除を行う。

- 1) E, F の除去: E, F はすべての Cladogram に同等の貢献をしていない。
- 2) A, B, C の除去: positive occurrences については、すべての Cladogram においてスコアが同じである。negative occurrences についても、同程度の貢献である (Trichotomy (4) の場合、6 を 2 に減らすことになるが total がもっとも 1 多いのだから、大勢には影響がない)

このプロセスを順に示す (Table. 3.5)

Cladogram	all character types	A, B, C, D, E, F	D, E, F	D	D+
1	12	10	2	1	1
2	13	11	3	2	2
3	13	11	3	2	2
4	14	12	3	2	2

Cladogram 2, 3 への変換: E, F eliminated (10 → 11), A, B, C eliminated (11 → 12)  
 Cladogram 3 への変換: D eliminated (11 → 10)

つまり、cladogram のスコアの差は D タイプの [positive] occurrences の差を反映している。

It is apparent that the cladograms differ with respect only to character-type D, which appears fewer times in cladogram (3.16) 1 than in cladograms (3.16) 2-4. (p. 188)

上から、次のことがわかる。

- 1) Information を最も効率的に要約する最良の cladogram を得るには、Character Type D の [positive] occurrences のみを調べ、そのスコアが最小のものを選ぶべき。
- 2) Negative occurrences は cladogram の選択に於て、何の役にも立たない。

上で見たとおり、negative occurrences は全スコアに同じ貢獻しかしていないから、全て削除される。つまりこれらは "uninformative" (p.188) である。

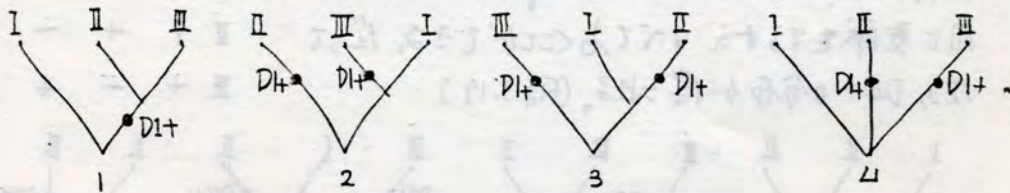
Synapomorphy の観点から更に negative occurrences について調べる (p.188~189) と上の Character-set について、Synapomorphy となり得る共有形質に次の 4つである。

Cladogram	1	D1+, A- - - - -	(II III)	} Synapomorphy による grouping
	2	B- - - - -	(I III)	
	3	C- - - - -	(I II)	

ところが、上で調べたように最適 cladogram は 1 であるから (II, III) の group だけが真である。ということは B-, C- による grouping は "false" ということになる。つまり、ある Character-set について negative occurrences に基づく grouping は任意の cladogram を容認する (A- → 1; B- → 2; C- → 3)。従って negative occurrences は "uninformative" であるばかりでなく "falsely informative" でもある (p.189)。これは形質の存在 (D1+) による "A-group" は natural であり、形質の欠乏 (B-, C-) による "not A group" は unnatural であるという Eldredge and Cracraft [1980] の議論と一致しているのである。

従って、以上をまとめると、total occurrences の判断基準を用いる限り、上の Character-set に対しては、D1+ を 2-taxon synapomorphy とする Cladogram 1 が真かれてくることになる。そこで、以下では Character Type D のみに着目して議論を定めることにする。 1982.10.19.

上の character-set について D1+ の出現のみに着目すると次のようになる。

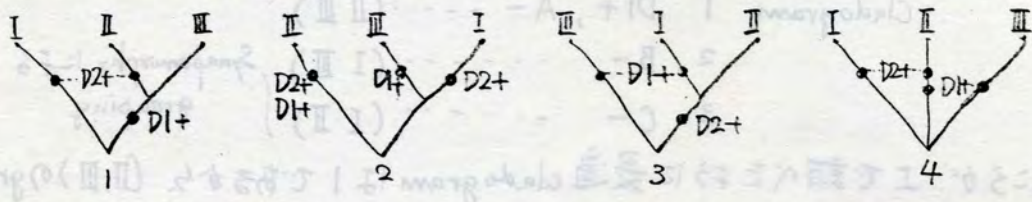


確かに Cladogram 1 のとき D1 の positive occurrences は 1 で最小だから、それは最適な cladogram といえる。

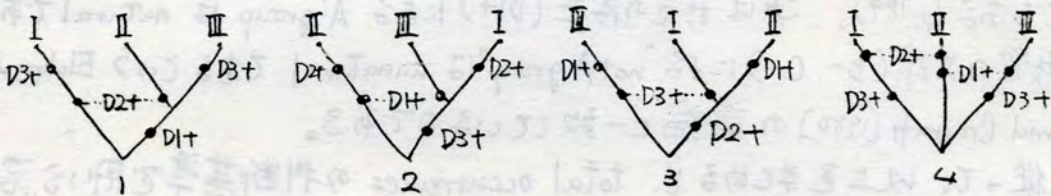
Conflicting Characters (p.189ff.): D1に加えて D2 を character set に入れると、D1 と D2 の positive occurrence は不一致 (conflict) を起こす。(Table 3.6)

Character-Types A, B, C 及び E, F は上述の議論により全て削除できる。その結果、下に示す clado-grams 及び character-type D1, D2 の occurrences が得られる (Fig. 3.17)

	A	B	C	D1	D2	E	F
I	+	-	-	-	+	+	-
II	-	+	-	+	+	+	-
III	-	-	+	+	-	+	-



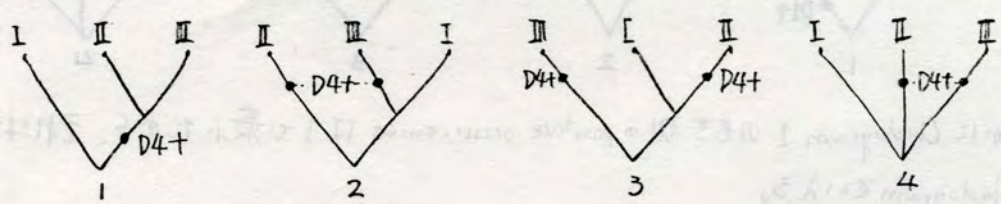
Cladogram 1と3は occurrences の数が3で等しく、character-set の情報を同程度に要約している。2は (I III) という "False" Group を示すから拒否される。結果として、Conflicting character がある場合 trichotomy 4 を採用するしかない。更に D3 を加えると (Table 3.7) 次のようになる。(Fig. 3.18)



Cladogram 1, 2, 3 は情報を同程度にまとめているから、この場合も4の trichotomy をえらぶしかない。

D1 と同じ917の D4 を考える (Table 3.8); このとき D1, D2, D3 は合わせると、各 cladogram に 2+, 1- の同じ貢献をするからすべて省くことができる。従って、次の D4+ の分布が得られる。(Fig. 3.19)

	D1	D2	D3	D4
I	-	+	+	-
II	+	+	-	+
III	+	-	+	+



この場合、1が最大の cladogram である。

This example illustrates that, for a given sample of information containing conflict in positive occurrences, various cladograms are possible, which differ more or less among themselves in their efficiency in integrating the information. (p.191)

Dタイプの特徴は 2-taxon level での "General Character" と考えられる。そして、それら  
が持つ意味は、それらを共有する 2-taxa が "Real Group" を作るという Natural  
Order に関する Prediction 或は Expectation of Generality (p.16) である。つまり  
Type-D character は Real Group を定義 (Predict) しているのである。

"Introduction" に於て、Incongruence の生ずる理由として次の3つが挙げられた (p.30)

We can summarize all these arguments by saying simply that if the  
order we seek in nature exists, incongruent characters indicate only that  
we have erred by considering as a general character features that are

- 1) two or more general characters,
- 2) single general characters used at the wrong level, or
- 3) not general characters at all.

そして、上の3番目の理由は "Absence of Character" として説明されている。

There is an infinite number of attributes that are lacking in any organism;  
general characters must refer to the presence of attributes, not to their absence  
(p.29)

ここからも、negative occurrences による grouping が無力であることが示されるので  
ある。Real Group を定義する General Character は即ち Synapomorphy である。

ここにきて、はじめて、Dタイプの Positive Occurrences が Synapomorphy 仮説であ  
ることがわかるのである。そして、上で引用した General Character の Incongruence に  
関する説明はそのままたの議論に適用される。(p.195 ff. Fig.3.21)

D-type (positive occurrences) ≡ General Character ≡ Synapomorphy

D-type の conflicting <sup>positive</sup> occurrences は複数個の cladogram を許すが、そのうち1つだけ  
が true であるのだから、残りは false つまり (real group を定義する) positive occur-  
rences ではないということになる (p.195) このことは D-type の特徴を再点検して、generality  
に関する判断がまちがいのならば、その D-type の特徴は false であるということになる (p.196,  
Fig.3.21)。このプロセスの結果

To the extent that conflicting positive occurrences can be studied and re-  
interpreted, conflicting occurrences disappear — if not in fact, at least  
in one's best judgement. (p.199)

Natural Order に対する確信がこれらの議論の根拠にある。

1982. 10. 20.