

동물학논단

DELTA: 반날개(Coleoptera: Staphylinidae)의 분류형질 데이터베이스(Taxonomic Character Database)



안 기 정

1985~1987 고려대학교 생물학과 (석사)
 1995 University of Kansas 곤충학과 (박사)
 1995~1997 University of Kansas 자연사박물관 곤충학부
 (박사후 연구원)
 1997~현재 충남대학교 생물학과 교수

조 영 복

1984 한남대학교 생물학과 (석사)
 1989 경북대학교 생물학과 (박사)
 현재 한남대학교 자연사박물관 강사

반날개 무리(딱정벌레목)들은 현재 약 40,000여 종이 넘게 기록되어 과(Family)수준에서 가장 종 다양성이 높은 분류군 중 하나로 알려져 있다. 이들은 많은 종수와 종들간의 외형적 유사성 (superficial similarity)으로 인하여 안정된 분류체계를 건설하는데 있어서 분류학자들에게 많은 어려움을 주고 있다. DELTA(Description Language for Taxonomy)는 이러한 분류학적 문제점을 크게 해소해줄 뿐만 아니라 데이터베이스를 쉽게 수정할 수 있고, 부가적인 데이터 역시 편리하게 추가될 수 있어 최적의 형질 데이터베이스 구축 시스템으로 사용되고 있다.

CSIRO(Commonwealth Scientific and Industrial Research Organization, Australia)의 Dallwitz 등에 의하여 지난 20여년간 연구 개발되고 있는 컴퓨터프로그램 시스템(computer program system)인

DELTA는 형질(character), 형질상태(character state), 그리고 images 데이터베이스를 작성하는데 매우 효과적인 프로그램 시스템으로 알려져 왔다. 또한 DELTA는 TDWG(International Taxonomic Databases Working Group)에 의해 데이터베이스 교환을 위한 표준으로 설정되어 있다. 이미 전세계에서 다양한 종류의 생물(viruses, 갑각류, 곤충류, 어류, 균류, 식물등)들의 분류에 이용되고 있다. 이 시스템은 여러 프로그램(CONFOR, DELFOR, KEY, INTKEY 등)들로 구성되어 있고, DELTA-format data는 기본적으로 분류형질 데이터베이스인 세 개의 파일 - CHARS, ITEMS, SPECS - 외에 기재, 검색표 등을 작성하는데 필요한 여러 파일(TONAT, TOKEY 등)들로 이루어져 있으며 이들은 프로그램 실행에 있어 서로간에 밀접한 연관성을 갖고 있다.

일반적으로 이들 기본적인 파일(CHARS, ITEMS, SPECS)들은 워드프로세서를 이용하여 아래의 예와 같은 형식으로 작성되어야 한다. 예를 들어 본 프로그램에서 지정된 기호(*; / 등)들은 다른 기호로 대치되어서는 안되고, 각 파일의 형식을 임의로 변경하면 문제점이 발생한다. 따라서 처음 사용하는 연구자는 기존의 완성된 데이터를 모방하며 익숙해지는 것이 가장 효율적인 DELTA 이용 방법이다. 또한 데이터 입력시 단순히 워드 프로세서를 이용할 경우의 불편함을 해결하기 위해 Taxasoft, Diana, DMS(Delta Menu System) 등의 데이터입력 프로그램이 개발되었다. 그리고 이들 파일의 문제점을 발견하고 지적해주는 기능을 하는 프로그램(Check)이 DELTA 시스템에 포함되어 있다.

현재까지 반날개무리 80분류군과 278형질로 구성된 DELTA-format 데이터가 완성되어 있다. 이를 예로 하여 DELTA-format에 필요한 기본적인

파일들을 설명하고자 한다. 지면관계상 몇몇 분류군과 형질들만을 예로 설명한다.

<CHARS 파일의 예>

* SHOW Rove Beetles (Coleoptera: Staphylinidae).
Revised Fri 18 Sept 18:04:58 1998 → 제목과 데이터가 완성된 시간

* CHARACTER LIST → 형질과 형질상태; 형질 1은 3개의 형질상태로 이루어짐을 알 수 있다.

- #1. <biomes>/
 - 1. intertidal/
 - 2. riparian/
 - 3. terrestrial <implicit>/
- #2. <intertidal habitat>/
 - 1. rocky intertidal zone/
 - 2. sandy intertidal zone/
- #3. <rocky intertidal zone>/
 - 1. under beach wrack/
 - 2. from debris along beach/
 - 3. inside empty barnacles/
 - 4. from rock crevice/
- #4. <sandy intertidal zone>/
 - 1. under beach wrack/
 - 2. from debris along beach/
 - 3. under boulder along beach/
- #5. <habitat type>/
 - 1. associated with other animals/
 - 2. associated with living plants/
 - 3. associated with fungi/
 - 4. from decaying plant materials/
 - 5. associated with carrion/
 - 6. associated with animal dung/
 - 7. from soil/

----- 중략 -----

#278. <type species>/

<ITEM 파일의 예>

* SHOW Rove Beetles (Coleoptera: Staphylinidae).
Revised Fri 18 Sept 18:04:58 1998 → 제목과 데이

터가 완성된 시간

* ITEM DESCRIPTIONS → 분류군과 형질; Adinopsis는 분류군명, Cameron은 명명자, 그리고 1,2는 형질 1의 형질상태 2를 의미한다. 즉 위의 CHARS 파일로부터 Adinopsis는 riparian이라는 것을 알 수 있다.

```
# Adinopsis <Cameron>/
1,2 17,6 19,2 20,1,6-3 21,3 22,3 24,1 25,1 27,3
34,3 35,1 40,1 45,2 49,1 53,3 54,1 55,1 56,1
57,1 61,1 62,2 63,1 64,1 69,3 75,2 76,2 77,3
78,2 80,2 81,1 83,5 84,2 88,2 91,4 92,2 94,2
95,4 97,3 101,3 103,1 104,2 105,1 106,1 108,4
109,2 112,1 113,3 114,1 115,2 116,1 120,1 121,3
128,2 132,2 138,1 139,3 141,1 142,1 147,1 148,1
149,3 150,3 151,3 154,1 155,1 164,1 166,6 167,1
169,1 178,1 179,1 181,1 185,2 186,1 191,3 198,2
203,1 204,1 206,3 209,1 213,2/3 214,1 215,1
277<[I]A. bicornis['I] Klm, [I]A. drakei['I]>
278<[I]A. rufobrunnea['I] Cameron>
```

<SPECS 파일의 예>

* SHOW Rove Beetles (Coleoptera: Staphylinidae).
Revised Fri 18 Sept 18:04:58 1998 → 제목과 데이터가 완성된 시간

* NUMBER OF CHARACTERS 278 → 형질수가 278

* MAXIMUM NUMBER OF STATES 10 → 형질의 최대형질상태가 10

* MAXIMUM NUMBER OF ITEMS 80 → 분류군수가 80

* CHARACTER TYPES 20,RN 23,RN 42,IN 44,IN 111,RN 172,IN 174,IN 176,IN 275-278,TE → 형질의 종류; RN(real number)은 몸의 길이 등과 같은 크기, IN(interger)은 털의 수와 같은 개수, TE(text)은 말로 설명되는 형질을 나타낸다.

* NUMBERS OF STATES 1,3 3,4 4,3 5,7 6,5 7,4 9,3 11,4 12,3 15-16,5 17,9 18,4 19,5 21-22,4 24,3 27-29,3 33,3 34,4 36,3 45,4 46,3 53-56,3 60-61,3 62,4 63,6 64-65,3 69-71,3 72,4 73-74,3 76,3 77,5 79,3 81-82,3 83,6 84,4 86-88,3 89,4 91,4

92-93,3 95,4 96-98,3 101-104,3 105,5 108,4 112,3
113-114,4 115-116,3 120-121,7 122,3 124-125,3
128,4 129,3 131,3 132,4 136,3 139,3 141,4 147,4
149,3 150,4 151,3 153,4 154,3 155-156,4 161,4
164-165,3 166,6 181,3 183-184,3 191,4 198,3 202,3
206-207,3 209,3 210,4 213,6 214,3 215,6 223-224,3
227,3 231,5 236,3 240,5 243,4 244,3 246,4 247,3
250,3 256,3 269,3 271,3 274,3 → 형질상태의 수;
1번 형질은 3개의 형질상태를 갖는다.

* **LINK CHARACTERS** 1-14 17-22 23-33 34-35
36-48 49-62 63-64 65-74 75-79 80-87 88-104
105-110 111-133 134-137 138-139 140-150 151-152
153-155 156-164 166-190 191-215 216-224 225-273
274-276 → 서로 연관성이 있는 형질들; 1-14 형질
은 모두 서식처에 관련된 형질.

* **NEW PARAGRAPHS AT CHARACTERS** 17
23 49 63 111 156 166 191 216 225 274 277 278
→ 새로운 문단이 시작되는 형질들; 기재시 17번
형질에서 새로운 문단 시작.

* **IMPLICIT VALUES** 1,3 14,2 16,5 18,2 24,2
26,2 28,1 29,3 30,2 31,2 32,1 36,1 37,2 38,1 39,2
40,2 41,1 43,1 50,1 51,1 52,2 57,2 58,2 59,2 60,3
65,3 66,1 67,1 69,1 72,1 73,1 74,1 76,3 78,1 79,2
87,1 90,2 96,1 106,2 107,2 108,1 110,1 117,1 118,1
119,2 122,1 123,1 129,2 130,1 132,1 133,1 135,1
138,1 140,2 141,4 142,2 143,2 144,1 146,1 152,2
153,1 156,1 157,2 158,1 159,2 160,1 161,2 162,2
163,2 165,1 167,2 169,2 171,1 173,1 175,1 177,1
180,2 182,2 183,3 185,1 187,1 188,1 189,2 190,2
192,2 196,2 197,1 199,1 200,2 201,2 202,3 203,2
205,2 206,1 207,3 208,2 210,1 211,2 212,2 216,1
218,1 219,1 220,1 221,1 223,3 224,2 225,1 226,1
228,1 230,1 234,1 239,1 241,1 242,1 245,1 249,1
257,1 260,1 263,1 266,1 268,1 270,1 273,1 → 분류
군 기재시 꼭 필요치는 않은 형질상태; 1,3은 육
상에 서식하는 종류를 의미하나 대부분의 반날개
무리들이 육상생활을 하므로 기재할 때 설명할
필요가 없다.

* **INAPPLICABLE CHARACTERS** 1,1:5-15
1,2:5-15 1,3:2-4 2,1:4 2,2:3 5,1:10-15 5,2:6-9:12-15
5,3:6-11:15 5,4:6-14 5,5/6/7:6-15 6,1:8-9 6,2:7:9
6,3:7-8 6,4:7-9 6,5:7-9 10,2:11 32,1:33 41,1:42

46,1:47-48 46,2:48 46,3:47 49,2:57-59 69,2/3:70-72
80,1:81 80,2:82 85,1:86 88,2/3:89 92,1:95 92,2/3:93
97,2/3:98 101,2:102 101,3:102 103,2:104 103,3:104
123,1:124-127 124,1:126-127 124,2:125:127 124,3:
125-126 130,1:131 135,1:136-137 136,2:137 136,3:
137 146,2:154-155 148,2:149-152:154-155 160,2:161
166,6:184 178,2:179-181 185,1:186 193,1:194-195
194,2:195 203,2:204 211,1:212 215,1:216 221,1:222
225,1:226-231:234-273 226,1:227 228,1:229 230,1:
231 231,1:233 231,2/4/5:232-233 231,3:232 234,1:
235-238 235,1:237-238 235,2:236:238 239,1:240
242,1:243-244 243,2/3/4:244 245,1:246-248 246,1:
248 246,2:247-248 246,3:247 246,4:247-248 249,1:
250-256 250,1:255-256 250,2:251-254:256 250,3:
251-255 251,2:252 252,2:253 252,1:254 257,1:
258-259 258,2:259 260,1:261-262 261,2:262 263,1:2
64-265 264,2:265 266,1:267 268,1:269-272 270,1:
271-272 271,1/2:272 → 적용될 수 없는 형질;
1,1:5-15는 1번 형질이 조건대에 서식하는 무리들
이므로 5-15형질은 적용될 수 없다.

위에서 설명한 반날개 무리들을 이용한 데이터
파일(chars, items, specs)들을 예로 하여 DELTA-
format 데이터가 어떻게 분류학적 연구에 이용되
는지를 살펴보자.

1. 분류군(신종, 미기록종등)의 기재 (Natural- Language Descriptions)

DELTA 시스템의 CONFOR 프로그램에 의하여
위의 DELTA-format 데이터가 각각의 분류군으로
기재된다. 기재를 위하여 이들 3 data files (chars,
items, specs)외에 필요한 tonat 파일은 다음과 같
이 구성된다.

- * SHOW Translate into natural language
- * HEADING Rove Beetles
(Coleoptera: Staphylinidae).
- * LISTING FILE tonat.lst *PRINT FILE tonat.prt
- * INPUT FILE specs
- * TRANSLATE INTO NATURAL LANGUAGE
- * REPLACE ANGLE BRACKETS

- * OMIT CHARACTER NUMBERS
- * OMIT INAPPLICABLES
- * OMIT PERIOD FOR CHARACTER 277-278
- * OMIT TYPESETTING MARKS
- * PRINT WIDTH 78
- * ITEM SUBHEADINGS
- #1. HABITAT:
- #17. Body:
- #23. HEAD:
- #49. ANTENNAE:
- #63. MOUTHPARTS:
- #111. THORAX:
- #156. ELYTRA:
- #166. LEGS:
- #191. ABDOMEN:
- #225. MALE SECONDARY SEXUAL CHARACTERISTICS:
- #274. DISTRIBUTION:
- #277. SPECIES EXAMINED:
- #278. TYPE SPECIES:
- * INPUT FILE chars
- * PRINT HEADING
- * INPUT FILE items

DELTA에 의해 작성된 여러 분류군 중 *Adinopsis* 를 예로 들면 기재는 다음과 같이 정리된다.

Adinopsis Cameron

HABITAT : Riparian. Body: Fusiform; uniformly reddish or reddish brown; length 1.6-3 mm; microsetae short, number reduced; microsculpture obsolete, subglabrous.

HEAD : Strongly deflexed into vertical plane; macrosetae small, difficult to distinguished from microsetae; temples broadly rounded from behind eyes to base. Compound eye moderate, 0.5-0.9 times as long as tempora; setae between facets present. Frons triangularly protruding and elevated; most to all head setae directed anteriorly.

ANTENNAE : Eleven articles; antennomere 3 distinctly shorter than antennomere 2; antennomere

4 elongate; antennomere 5 elongate; antennomere 10 elongate; antennomere 11 with papilla; antennomeres 5-10 distinctly decreasing in relative lengths apically; antennae extended to more or less 1/2 length of elytra.

MOUTHPARTS : Labrum transverse; more than 40 setae in addition to basic setal pattern. With two large subapical teeth. Maxilla with 4 articles; maxillary palpi elongate, 1.0-1.5 times of length of head; maxillary palpal segment 3 spindle-shaped; maxillary palpal segment 4 shorter than half of maximum width of 3. Lacinia elongate; myllaenine type; setae few only on mesal surface; galea with several setae only on mesal surface. Labial palpus with 3 articles; ligula very long, more than 1.5 times length of basal palpal segment; divided into 2 lobes; lobes parallel-sided; apical 0.9 or more split; medial setae absent; median pseudopore distinctly absent; lateral pseudopore distinctly present; about 6-12 present on each lateral field. Mentum almost trapezoidal, medium size, anterior margin emarginate; greatest width much greater than apex of gula; antero-lateral angles prolonged into short spinose processes; seta v absent.

THORAX : Strongly convex; antero-lateral margin moderately rounded; postero-lateral margin angulate, corners sharp, forming angle of 90 degree or greater; anterior margin more or less truncate; posterior margin more or less sinuate; pubescence at midline all directed posteriorly; lateral setae directed posteriorly and obliquely laterally; microsculpture obsoletely reticulate; pronotal length to elytral length ratio about 1.0. Hypomera delimited from disk by marginal ridge or carina; not visible in lateral view. Medial carina strong, complete to apex of mesosternal process; with numerous

microsetae present, especially in apical half; middle coxae narrowly separated (separation less than 0.15 total length meso- and metasternal processes); intercoxal processes separated by distinct suture; mesosternal process extended to more than 0.8 times to total length of mesocoxae; apex

pointed. Isthmus virtually to completely absent. Metasternal process extended to less than 0.3 times to total length of mesocoxae; apex rounded.

ELYTRA : Outer latero-apical angle markedly or deeply sinuate.

LEGS : Tarsal formula 2-2-2; numerous spines on front tibiae; numerous

spines on middle tibiae present; one empodial setae present; empodial setae distinctly longer than claws; ventral plate of hind coxa large.

ABDOMEN : Tapered from base to apex; tergites subglabrous; posterior margin of tergites III-VI pectinate; with two types of teeth; gland opening on tergite VII absent; apical margin of tergite VIII; tergite X finger-shaped, or rectangularly elongate apical part; tergite X with medial setae more or less long, setiform, unmodified; tergites not impressed.

SPECIES EXAMINED : *A. bicornis* Klm, *A. drakei*

TYPE SPECIES : *A. rufobrunnea* Cameron

2. 검색표 작성(Key Creation)

DELTA 시스템의 KEY 프로그램에 의하여 위의 DELTA-format 데이터가 검색표로 변환된다. 이를 위하여 필요한 tokey 파일은 다음과 같이 구성된다.

- * SHOW Translate into KEY format
- * HEADING Rove Beetles
(Coleoptera: Staphylinidae).
- * LISTING FILE tokey.lst
- * INPUT FILE specs
- * TRANSLATE INTO KEY FORMAT
- * EXCLUDE CHARACTERS
- * USE NORMAL VALUES
- * KEY STATES
- * KEY OUTPUT FILE kchars
- * INPUT FILE chars
- * KEY OUTPUT FILE kitems
- * INPUT FILE items

작성된 검색표의 일부분을 예로 들면 결과는 다음과 같다.

--- 중 략 ---

- 5(4). Setae very numerous, densely arranged; numerous spines on front tibiae present; frontal suture absent; maxilla with 4 and pseudo-palpal segment; labial palpus with 3 articles with pseudosegment*Aleochara*
Setae less numerous, few to many, more or less loosely and irregularly scattered; numerous spines on front tibiae absent; frontal suture present; maxilla with 4 articles; labial palpus with 3 articles*Oxypoda*
- 6(2). Temples moderately rounded; neck absent; two long medial setae absent 7
Temples broadly rounded from behind eyes to base; neck present, distinct to very distinct, 2/3 to 1/2 head width; two long medial setae present*Blepharhymenus*
- 7(6). Head almost as broad as pronotum; antennomere 5 elongate; antennomere 10 more or less quadrate; antennae extended to more or less 1/2 length of elytra*Ilyobates*
Head distinctly narrower than pronotum; antennomere 5 transverse; antennomere 10 transverse; antennae extended to more or less posterior margin of pronotum
.....*Aleochara*

--- 중 략 ---

3. 계통분석을 위한 자료행렬로의 변환 (Phylogenetic Analysis)

DELTA 시스템의 CONFOR 프로그램에 의하여 위의 DELTA-format 데이터가 계통분석을 위한 자료행렬로 변환된다. 이를 위하여 필요한 tohen 파일은 다음과 같이 구성된다. 이외에 nexus 파일 등으로의 전환(tonex)이 가능하다.

- * SHOW Translate into HENNIG86 format
- * HEADING Rove Beetles

(Coleoptera: Staphylinidae).
 * PRINT FILE tohen.prt
 * INPUT FILE specs
 * TRANSLATE INTO HENNIG FORMAT
 * PRINT CHARACTER LIST
 * OMIT TYPESETTING MARKS
 * COMMENT. EXCLUDE CHARACTERS
 * COMMENT. KEY STATES
 * INPUT FILE chars
 * OUTPUT FILE hendata
 * INPUT FILE items

몇몇 분류군을 예로 들면 다음과 같은 자료행렬을 얻을 수 있다.

```
xread
'Aleocharine genera (Coleoptera: Staphylinidae)'
266 79
Adinopsis
1????????????451122001202110?2001010001???0
001200001120100200?2???001121110?41
??01?131?1302???2?01001310020100010200????11
0?10?0??0210010?00022210000101011100
50?0?000001012?10001121???1010112001221001
1100??0000?210????????????????????
????????????????????
Agaricochara
2???2?????200?462?01101?02110?2?0101100????0
001?0011112?102200?0?011120010?020
??0021?1?1201???????0110?0?2210001?000???010
?00?1?0211110?02013110130101011110
21?1?0000????12?0?001101???10?011121?1021?011
?120?0000?210????????????????????
????????????????????
--- 중 략 ---
Xenodusa
2???00????????48212020100211122001011002???
0001000011121300200?022000121000?110
??20?101?1?02???01010110?0220200010201021?31
0?0010?0013110?02013110130101011110
11?1?000001121200?001100???10101121?1021?011
?130?00011200????????????????????
```

```
????????????????????????????
;
proc / ;
```

4. Interactive Identification Guide

DELTA 시스템의 INTKEY 프로그램에 의하여 위의 DELTA-format 데이터가 interactive key를 위한 데이터로 변환된다. 이를 위하여 필요한 toint 파일은 다음과 같이 구성된다. 또한 분류군과 형질의 images를 key에 사용하기 위해서는 프로그램 INTIMATE의 도움이 필요하다.

```
* SHOW Translate into INTKEY format
* HEADING Rove Beetles
(Coleoptera: Staphylinidae) #DATE
* LISTING FILE toint.lst
* INPUT FILE specs
* INPUT FILE cnotes
* INPUT FILE cimages
* INPUT FILE timages
* TRANSLATE INTO INTKEY FORMAT
* OMIT TYPESETTING MARKS
* INTKEY OUTPUT FILE ichars
* INPUT FILE chars
* INTKEY OUTPUT FILE iitems
* INPUT FILE items
```

interactive key를 이용한 Computerized Illustrated Identification Guide는 지면으로는 거의 설명이 불가능하므로 아쉽게도 생략해야 할 것 같다. 이들은 특정지역의 동식물상 조사, 그리고 revisional work, monographic work 등의 출판물을 완성하는데 최적의 자료로 이용된다. 최근에 Lucid, Linnaeus II 등 interactive identification guide를 완성하는 여러 프로그램이 개발되고 있으나 DELTA의 INTKEY와 INTIMATE를 이용한 출판물이 최고로 평가되고 있다.

보다 자세한 프로그램에 대한 내용은 동물분류학회의 분류학회보 제24호를 참고하시고 DELTA 시스템의 역사, registration, installation, 등의 행정적인 내용과 여러 분류군에 사용된 예, 참고문헌

등은 DELTA web 페이지(<http://biodiversity.bio.uno.edu/delta>)를 방문하시고 또한 전세계 DELTA 사용자들의 discussion list인 delta-L도 관심 있으신 분들께는 도움이 되리라 생각된다.

"I have evolved a rule that has proved useful for myself and might be for others

*not born with championship potential: -----.
Be a hunter and explorer, not a problem solver. Perhaps the strategy can never work for track, with one distance and one clock. But it serves wonderfully well at the shifting frontiers of science."*

– E. O. Wilson. 1995. Naturalist. p.123 –