

SOCIETAT CATALANA DE BIOLOGIA

FILIAL DE L'INSTITUT D'ESTUDIS CATALANS

COL·LOQUIS - XII i XIII

XII

**BIOLOGIA MOLECULAR  
I ECOLOGIA**



XIII

**MATEMÀTICA I BIOLOGIA**

BARCELONA

1979



**COL·LOQUIS  
DE LA  
SOCIETAT CATALANA DE BIOLOGIA**

**XII  
BIOLOGIA MOLECULAR  
I ECOLOGIA**

•

**XIII  
MATEMÀTICA I BIOLOGIA**



SOCIETAT CATALANA DE BIOLOGIA

FILIAL DE L'INSTITUT D'ESTUDIS CATALANS

COL·LOQUIS - XII i XIII

XII

**BIOLOGIA MOLECULAR  
I ECOLOGIA**



XIII

**MATEMÀTICA I BIOLOGIA**

BARCELONA

1979

**ISBN 84-7283-027-6**

**Dipòsit Legal: B. 6.409-1980**

***Tallers Gràfics A. Nàñez - Paris, 208 - Barcelona***

*El dia 20 de maig de 1976 tingué lloc el dotzè dels Colloquis de la SOCIETAT CATALANA DE BIOLOGIA, el qual fou dedicat a la Biologia Molecular i a l'Ecologia. La sessió fou celebrada al Palau Dalmases, Barcelona, sota la presidència del doctor CARLES BAS i PEIRED. En foren coordinadors els doctors JOAN-ANTON SUBIRANA i RAMON MARGALEF.*

This One



LTRA-CDR-492B





**XII**

**BIOLOGIA MOLECULAR  
I ECOLOGIA**



## DIFRACCIÓ DE RAIGS X DE COMPLEXOS CROMATÍNICS

per FERRAN AZORÍN i JOAN SUBIRANA

Departament de Química Macromolecular del C.S.I.C., Escola Tècnica Superior  
d'Enginyers Industrials de Barcelona. Universitat Politècnica de Barcelona

Han estat analitzades per difracció de raigs X mostres de complexos cromatínics de diverses procedències. Els diagrames de difracció obtinguts revelen:

1) Les fibres sotmeses a estudi estan compostes per dos tipus de components que coexisteixen en una mateixa mostra: per una banda el component globular responsable de la sèrie de màxims de difracció 11 nm, 5,5 nm, 3,7 nm i 2,7 nm; i d'altra banda el component fibrós responsable del màxim de difracció equatorial, la posició del qual varia amb la concentració.

2) El component globular és degut a la distribució repetitiva de sub-unitats espaiades 11 nm al llarg de les molècules de DNA. Es demostra la presència *in vivo* d'aquest tipus d'estructura.

3) El component fibrós està constituït per molècules de DNA paral·leles entre elles. La importància relativa d'aquest component en relació amb el globular varia segons la mostra, en tot cas, la presència *in vivo* d'aquesta estructura no és conclouent.

4) El màxim de difracció de 2,1 nm, no pertany al mateix sistema de difracció dels altres màxims, i està relacionat amb el primer *layer-line* de la forma B del DNA.

5) A concentracions en pes del 60 %, els diagrames canvien de sobte, es perd l'ordre i apareixen dos màxims a 8 nm i 3,7 nm, així com un de central.

De la presència dels màxims de 8 nm i 3,7 nm, en mostres submergides en glicerol es dedueix l'existència d'una diferenciació espacial entre regions riques en DNA i regions riques en proteïna.

De l'estudi del màxim central es dedueix que el component fibrós està organitzat en cilindres formats per diverses molècules de DNA, d'una longitud superior a 15 nm. A concentracions superiors al 60 % en pes aquests cilindres es trenquen i es redueix la seva longitud.

# ESTUDI SOBRE L'ORGANITZACIÓ DE LA CROMATINA I ANÀLISI DEL COMPLEMENT PROTEIC NUCLEAR DURANT LA DIFERENCIACIÓ CEL·LULAR

per E. ROCHA, L. MENGOD i LL. CORNUDELLA

Departament de Química Macromolecular del C.S.I.C.  
Universitat Politècnica de Barcelona

## INTRODUCCIÓ

Les cèl·lules eucariotes contenen ADN, una gran part del qual es troba localitzat en el nucli cel·lular protegit per membranes que el diferencien de la resta de fraccions i cossos cel·lulars. El nucli conté, a més, una gran quantitat de proteïnes que es troben involucrades amb grau distint de participació en els mecanismes de síntesi, procés i elaboració de molècules d'ARN. La molècula dúplex d'ADN resulta directament associada amb un sol grup específic de proteïnes nuclears —les histones— constituint complexos de naturalesa nucleoproteica. El conjunt d'aquests complexos configura el material genètic pròpiament dit, el qual esdevé empaquetat en l'interior del nucli i rep el nom de cromatina tot definint els caràcters de la cèl·lula i constituint el genoma distintiu cel·lular. Una característica específica dels organismes superiors és que solament una petita part del genoma expressa el seu potencial genètic, mentre que la resta es manté inactiu al llarg del cicle vital de la cèl·lula. Aquest darrer fet és clarament indicatiu de l'elevat nivell de complexitat en l'organització estructural de la cromatina, els components de la qual es disposen de manera que donen lloc a una estructura repetitiva, regularment espaiada que s'anomena nucleosoma (LEWIN, 1975) l'existència del qual ha estat evidenciada al microscopi electrònic (OLINS i OLINS, 1974) i reconeguda bioquímicament per mitjà de l'acció de diversos enzims nucleolítics que el fragmenten d'una forma regular (HEWISH i BURGOYNE, 1974; CLARK i FELSENFELD, 1974).

Donada l'estreta vinculació entre l'organització estructural i l'activitat genètica i el tenir en compte que aquesta darrera adquireix singular

intensitat durant el fenomen de la diferenciació cel·lular, l'estudi bioquímic d'aquesta interdependència al llarg del desenvolupament cel·lular esdevé de cabdal importància. L'objecte, doncs, del nostre treball ha consistit en l'estudi de l'organització de la cromatina de cèl·lules eucariotes en procés de diferenciació i ensem llur íntima relació amb el control de l'activitat i l'expressió genètiques. Com a sistema biològic s'han emprat cèl·lules germinals, principalment d'equinoderms, durant el procés diferenciatiu (espermatogènesi).

#### CONSIDERACIONS METODOLÒGIQUES I RESULTATS

L'objectiu que acabem d'esmentar ha predeterminat el següent plantejament experimental: (a) examen de l'acció de nucleases i llur cinètica en cromatina nadiua de cèl·lules espermatogèniques en diferents fases del cicle reproductiu; (b) aïllament i caracterització de les sub-unitats nucleoproteiques resultants de la digestió enzimàtica; (c) aïllament i anàlisi dels components de les partícules unitàries (ADN i complement proteic).

Per assolir aquestes fites, espècies mascles de *Holothuria tubulosa* van ésser recollides en etapes diverses al llarg del seu desenvolupament (CORNUDELLA, 1977). Un cop extirpat el teixit espermatogènic es procedí al fraccionament cel·lular i a la preparació de cromatina que es digeriria amb nucleasa micrococcal, separant-ne els productes de la digestió per centrifugació en gradients de sacarosa (ROCHA i CORNUDELLA, 1976). Paral·lelament, una alíquota de la cromatina nadiua era tractada en medi àcid per a extreure-li el complement proteic d'histones, que posteriorment eren caracteritzades electroforèticament en gels de poliacrilamida.

D'antuvi cal assenyalar la peculiar resistència mecànica del nucli cel·lular a la penetració enzimàtica, i és indispensable la preparació de suspensions de cromatina nuclear per tal que l'enzim pogués iniciar la solubilització de l'ADN per bé que amb una lentitud molt marcada.

D'altra banda, els diagrames de sedimentació dels digerits (fig. 1) obtinguts de gònades immadures palesen la presència d'unitats monomèriques o nucleosomes juntament amb agrupacions d'embalum superior. El més notable, però, és l'aparició d'un material que sedimenta lentament, una mena de subnucleosomes, la proporció del qual depèn de la concentració d'enzim. Aquest pic de material lleuger decreix de manera progressiva a mesura que el cicle espermatogènic avança fins a desaparèixer totalment de la cromatina pertanyent a esperma madur (fig. 2). És poc plausible que aquest material procedeixi de la degradació del nucleosoma donat que en condicions més enèrgiques de digestió les espècies multimèriques

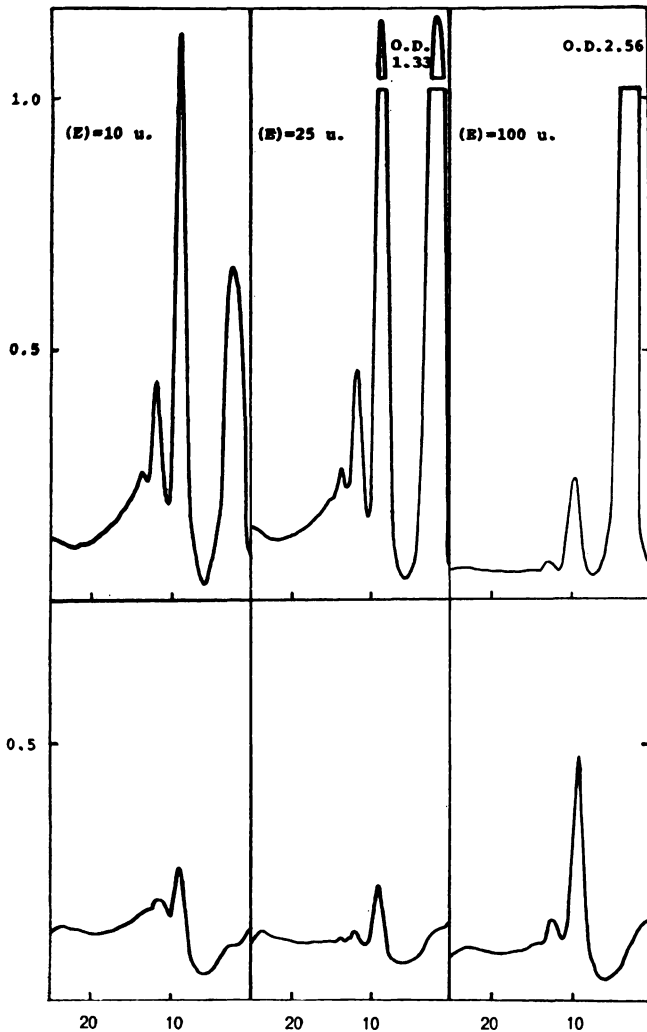


FIG. 1. — Anàlisi en gradients de sacarosa de cromatina nadiua de gònada immadura digerida amb 30 min. concentracions creixents de nucleasa micrococcal. *Diagrames superiors*: 10, 25 i 100 unitats òptiques d'enzim. *Diagrames inferiors*: Material soluble recuperat al rentar els sediments de cromatina no-digerida després del tractament amb les concentracions enzimàtiques esmentades

Absorbància a 254 nm  
Fracció n.º

donen lloc a monòmers sense que s'apreciï cap variació quantitativa del pic de sedimentació lenta en els gradients de sacarosa.

La conversió de formes multimèriques en nucleosomes és coherent amb la relació precursor-producte postulada. Ara bé, l'accessibilitat variable detectada de la nucleasa a la cromatina concomitant amb el procés de maduració germinal, així com la presència inicialment de subnucleosomes i llur posterior absència en les darreres fases de la espermatogènesi són demostratives de la generació de canvis importants en l'organització estruc-

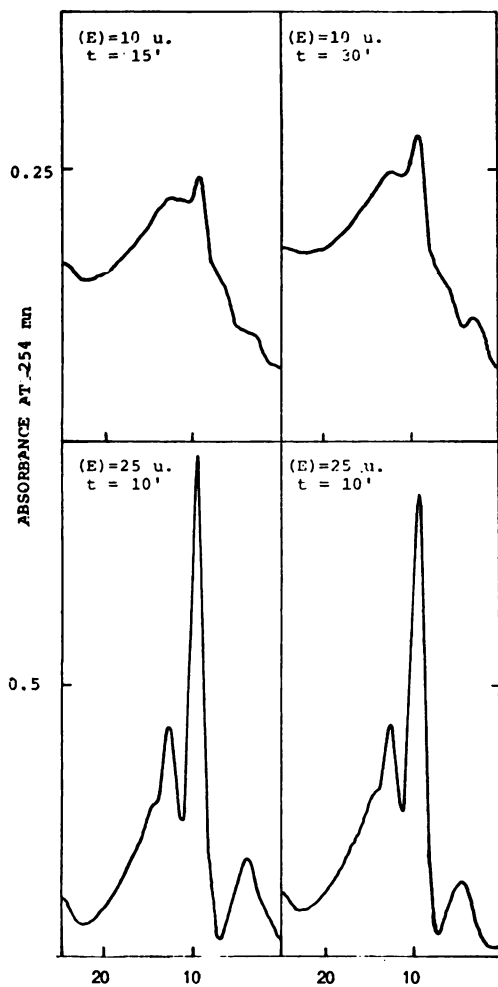


FIG. 2. — Anàlisi de sedimentació de cromatina digerida en etapes finals de la maduració. Després del tractament de preparacions de cromatina nadiua (diagrames superiors), el material insoluble es re-digerí amb el mateix enzim (diagrames inferiors). Les condicions d'hidròlisi enzimàtica es corresponen amb les indicades en la figura 1  
Absorbància a 254 mμ  
Fracció n.º

tural de la cromatina durant el desenvolupament a ben segur reflectint estats funcionals desiguals. En addició, l'anàlisi electroforètica de la dotació proteica nuclear (fig. 3) mostra com a tret que més ressalta l'aparició sobtada d'una petita proteïna, anomenada  $\phi_0$  (ROCHA i CORNUDELLA, 1976), en les etapes mitjanes del cicle, la proporció de la qual augmenta molt sensiblement en l'espermatozoide.

És indubtable que existeix una estreta correlació entre la generació d'aquesta proteïna i els canvis descrits del grau d'accessibilitat de l'enzim, en particular amb la desaparició de material subnucleosòmic a les darreries del procés germinatiu.



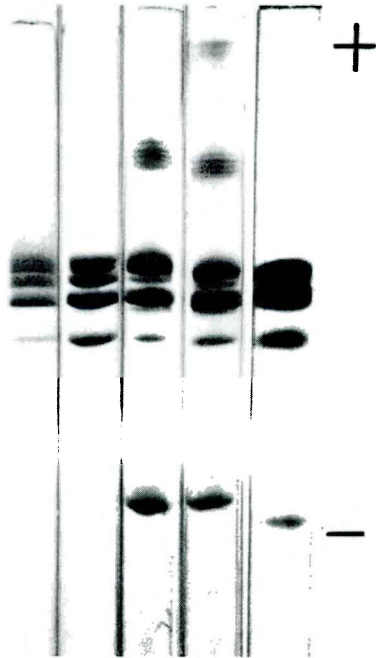


FIG. 3. — Anàlisi electroforètica en gels d'acrilamida d'histones totals extretes de gònades masculines en diversos períodes de desenvolupament germinal. Els electroferogrames corresponen d'esquerra a dreta als estats de: Activació de la gònada; creixement, maduració. Cal notar l'absència del component de ràpida mobilitat (proteïna  $\Phi_0$ ) en els estadis inicials de desenvolupament i la seva presència al llarg de l'espermiogènesi



## DISCUSSIÓ

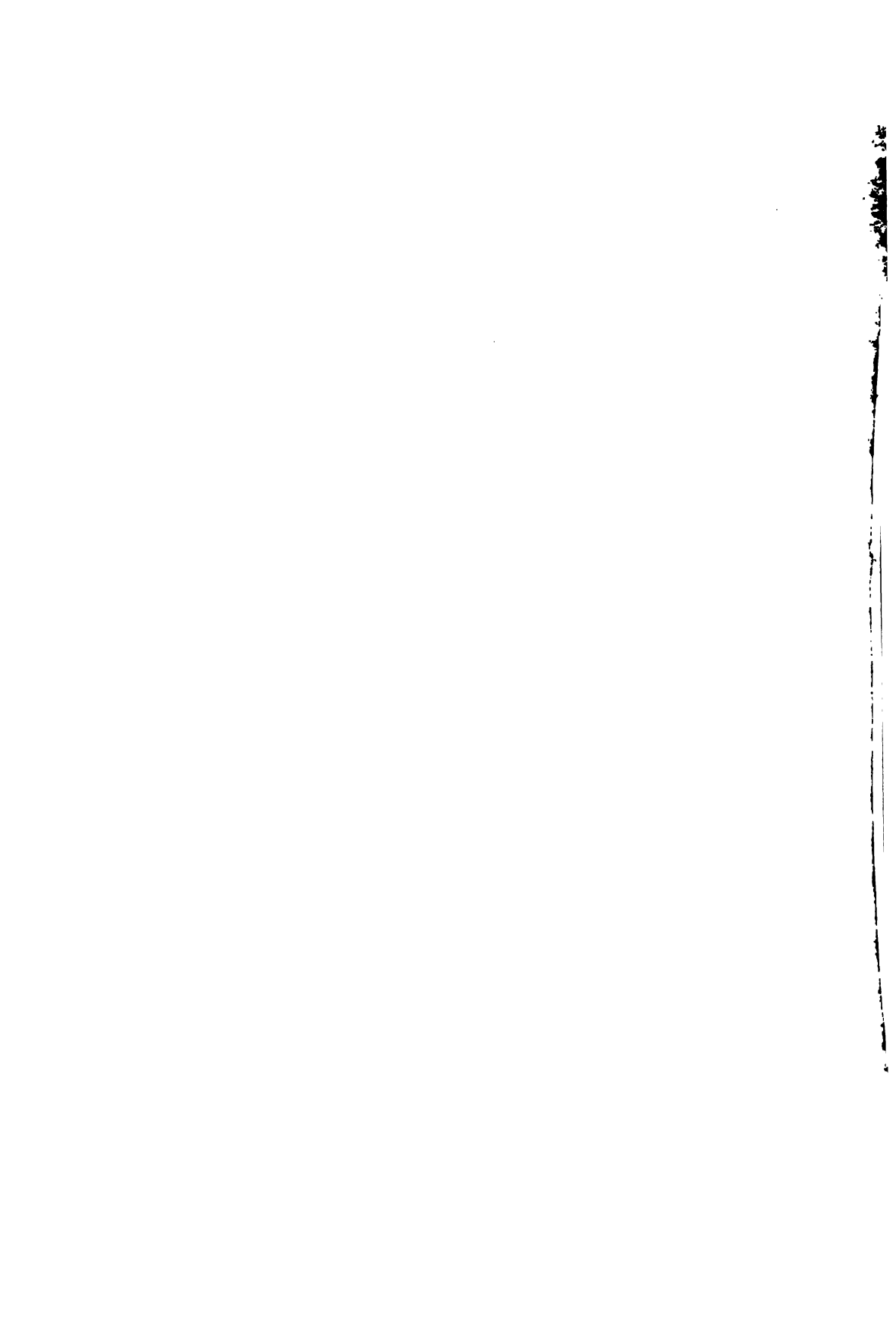
El nucli celular que deuria ésser normalment molt permeable a l'enzim, en l'organisme estudiat ofereix una resistència inusitada a la penetració enzimàtica. Aquest comportament decididament anòmal del nucli enfront de la nucleasa micrococcal implica una singular compacitat en l'organització de la membrana nuclear que redueixi la seva permeabilitat i entorpeixi l'accés de l'enzim. Aquesta conducta suggereix, a més, que el complex de nuclihistona, en les cèl·lules germinals posseeix un grau de cohesió que es tradueix en una accentuada ocultació de l'ADN.

La variació de l'accés enzimàtic a la cromatina evidenciada per l'anàlisi de sedimentació al llarg del desenvolupament, com també la desaparició gradual dels components de sedimentació lenta, particularment pesada en aquelles etapes de diferenciació caracteritzades per un grau uniforme en la condensació de la cromatina, són consistents amb unes modificacions o reordenaments estructurals que impliquin l'alteració de les regions de cromatina d'on inicialment s'originaven aquells darrers, de manera que esdevinguin en un moment donat aquestes regions inaccessibles per l'enzim.

Amb tota probabilitat, la presència de la proteïna  $\phi_0$ , detectada en les últimes fases espermatogèniques i especialment en l'esperma madura, deu ésser essencial en la inducció del procés de bloqueig esmentat. El petit volum molecular d'aquesta proteïna esdevé molt adient per a introduir-se en els llocs d'ancoratge de la nucleasa i dificultar així l'acció de l'enzim. L'observació, doncs, del fet que els nucleosomes i llurs agrupacions multimèriques siguin els productes exclusius de la digestió per nucleasa micrococcal al termini de l'espermatogènesi en la *Holothuria*, afavoreix la idea que el model proposat d'organització de l'ADN cromosòmic, basat en la distribució periòdica d'una sub-estructura repetitiva fonamental: el nucleosoma, es trobi restringit a formes cromatíniques genèticament inactives.

## BIBLIOGRAFIA

1. CLARK, R. J. i FELSENFELD, G.: *Biochemistry*, 13, 3622-2638 (1974).
2. CORNUDELLA, L.: *Avances de la Bioquímica*, Editat per L. Cornudella, C. F. de Heredia, J. Oro i A. Sols, Salvat Eds. Barcelona, pp. 39-50 (1977).
3. HEWISH, D. R. i BURGOYNE, L. A.: *Biochem. Biophys. Res. Commun.* 52, 504-510 (1973).
4. LEWIN, B.: *Nature* 254, 651-653 (1975).
5. OLINS, A. L. i OLINS, D. E.: *Science* 183, 330 (1974).
6. ROCHA, E. i CORNUDELLA, L.: *Biochem. Biophys. Res. Commun.* 68, 1073-1081 (1976).



# CARÀCTER DE LES DIFERÈNCIES ENTRE EL PLÀNCTON VEGETAL MARÍ I EL D'AIGUA DOLÇA

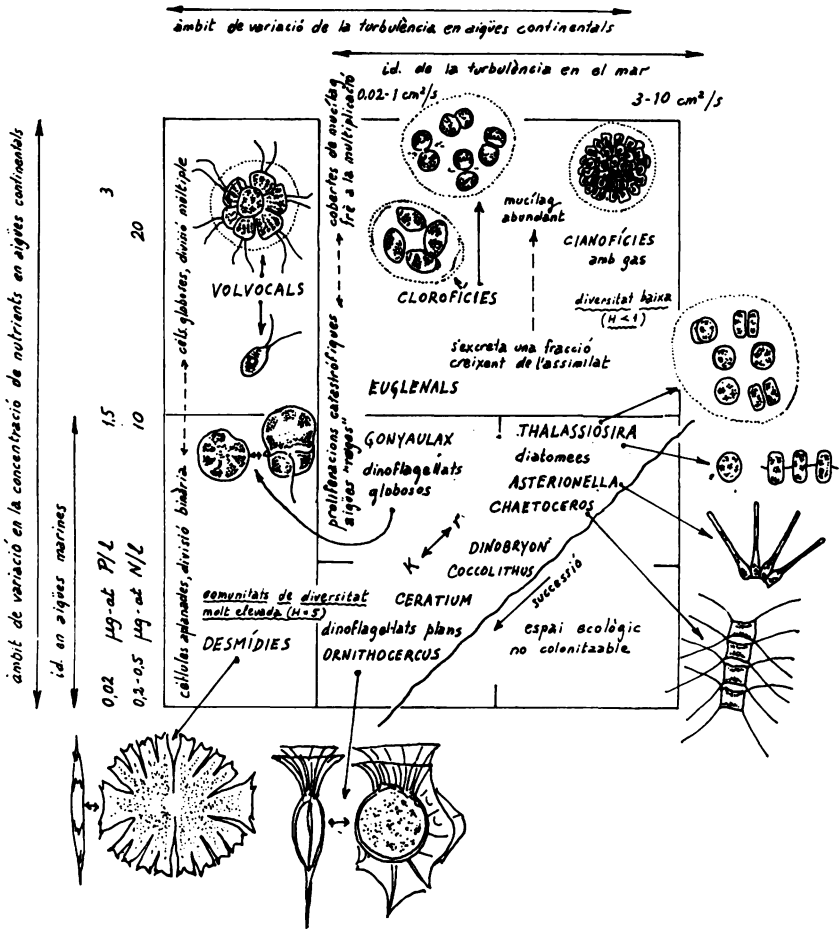
per RAMON MARGALEF

Cap del Departament d'Ecologia. Facultat de Biologia.  
Universitat de Barcelona

El fitoplàncton marí i el d'aigua dolça tenen un aspecte diferent. No es tracta solament de diferències en la importància relativa dels diversos grups taxonòmics, sinó també de diferències en la representació dels diversos grups biològics (*Lebensformen, life forms*); aquesta darrera diferenciació podria semblar més important a l'ecòleg que es fixa sempre més en la fisonomia de la vegetació que en la composició d'una flora. El fitoplàncton marí és dominat per diatomees, dinoflagel·lats, crisofícies i en menor quantitat, criptomonadals; en el plàncton de les aigües continentals, aquests grups hi són ben representats, i hi ha, a més, força proporció de clorofícies, de cianofícies i d'euglenals. En termes de grups biològics d'adaptació, l'aigua dolça es caracteritza per la presència d'una forta proporció d'organismes de divisió cel·lular múltiple i abundant producció de mucil·lages, la qual cosa porta sovint a la formació de colònies o cenobis flotants.

Un examen crític de les diferències basat en l'experiència acumulada en els dos àmbits, el marí i el continental, fa pensar que la major riquesa de grups o tipus biològics d'adaptació en el domini de l'aigua dolça o continental és una conseqüència senzilla de l'àmbit més gran de variació dels paràmetres ecològics significants dintre les aigües no marines que en les aigües oceàniques, les quals, tot i sent incomparablement més extenses, ofereixen unes condicions de vida notablement menys variades.

Si considerem els tipus biològics d'organismes planctònics, i llur encarnació en els diversos grups taxonòmics, damunt d'un espai ecològic ideal, definit —per simplificar la representació— solament per dos paràmetres que ens semblen d'importància indiscutible, la concentració de nutrients essencials i limitants, i la turbulència de l'aigua, veurem que l'àmbit marí és més petit que l'àmbit representat per les aigües continentals. En



Esquema dels principals tipus biològics del fitoplancton, i la seva expressió taxonòmica concreta, portat sobre un espai de variació dels dos factors més importants per la biologia del plancton (a més de la llum): nutrients (de baix a dalt) i turbulència (d'esquerra a dreta). El text es pot seguir sobre aquest esquema, on s'assenyalen o suggereixen moltes de les relacions o aspectes comentats. A més, la sageta que porta «successió» assenjala la direcció preferent de canvi en les curses successions planctòniques que acompanyen l'ús dels nutrients i l'estratificació de l'aigua. Damunt d'ella, la indicació K—r, posa de manifest la relació entre els canvis successional i les estratègies bàsiques reconegudes en ecologia

l'espai on els dos es superposen, poden haver-hi diferències de composició taxonòmica, però no hi ha diferències essencials en les síndromes d'adaptació dels organismes a la vida planctònica, sigui al mar, sigui a l'aigua dolça. En aquest espai de turbulència moderada a forta, i de capacitat nutritiva de baixa a moderada, les comunitats s'estenen entre les dues possibilitats freqüentment realitzades: 1) aigües fèrtils i turbulentes, en les quals són seleccionats organismes de forta taxa de multiplicació, i passius, com les diatomees; 2) aigües pobres i estratificades, en les quals acaben dominant poblacions molt disperses d'organismes nedadors, de baixa taxa de multiplicació, sovint aplanats i amb adaptacions que faciliten la difusió a l'entorn de la cèl·lula, no rarament ben defensats contra potencials depredadors, els representants més típics dels quals són els elegants dinoflagellats de les àrees oceàniques centrals, de morfologia qualificable de barroca. La combinació d'una forta turbulència amb una quantitat molt limitada de nutrients no és colonitzable, puix les poblacions es dispersen sense temps de multiplicar-se. L'altra combinació de turbulència feble amb gran fertilitat potencial, condueix al desenvolupament massiu de dinoflagellats de les «aigües roges»; tenen una forma més globosa que els dinoflagellats ja esmentats i són més rics en clorofil·la, com correspon a llur taxa de multiplicació més elevada. De vegades tanmateix massa elevada, puix porta a situacions catastròfiques per a la mateixa població de dinoflagellats i per a les poblacions acompanyants d'altres organismes. Existeix una selecció natural en aquest nivell, si en situacions d'alta fertilitat de les aigües, tenen valor algunes característiques que frenen la multiplicació? Hom es sentiria temptat de creure-ho, si es reflexiona sobre la base dels fets següents. Les algues excreten en forma soluble una fracció de la matèria orgànica formada com a resultat final de la fotosíntesi. És una fracció considerable que sovint es mou entre el 5 i el 30 % de l'assimilat. És variable fins i tot dintre una espècie, i depèn de factors diversos. La concentració de ions amoni en l'aigua n'és un; com més amoni, major fracció excretada. Els animals, que contribueixen a la concentració d'amoni, i per tant poden regular la producció de les algues amb les que sovint estan associats, de vegades molt íntimament. Com és obvi, l'abundant producció de matèria orgànica extracel·lular requereix una elevada concentració de nutrients. Les cobertes visibles de mucíl·lag no són més que una forma particular, feta de mol·lècules filiformes particularment llargues, d'aquella matèria orgànica excretada. S'ha debatut el sentit biològic de la presència de cobertes de mucíl·lag en les algues del plàncton, en relació amb un millorament de la flotabilitat o com a defensa enfront d'animals. La cosa no és clara. Solament val a dir que mentre les cobertes de mucíl·lag són molt freqüents en el plàncton d'aigua dolça, són raríssimes en el plancton marí i en el medi oceànic solament

se'n troben en algunes aigües d'una fertilitat excepcionalment alta, on viuen, per exemple, *Thalassiosira partheneia* i *Phaeocystis poucheti*. El que és cert, és que una coberta de mucíl·lag redueix la turbulència entorn la cèl·lula i fa nosa per una difusió ràpida dels nutrients. Podríem dir que, amb la coberta de mucíl·lag, una cèl·lula suspesa en una solució molt nutritiva, frena d'alguna manera la pròpia multiplicació. És en aquest sentit que he començat aquest paràgraf com ho he fet, pensant que la producció de mucíl·lag porta d'alguna manera a una major estabilització de les poblacions planctòniques, i aquesta estabilització s'ha d'entendre, obviament, no solament en relació amb la pròpia espècie, sinó també en relació amb els animals. Existeix informació convincent que porta a acceptar que les algues embolcallades per una coberta de mucíl·lag tenen una probabilitat major de passar vives a través del tub digestiu dels crustacis filtradors del plàncton.

Com a resum es pot dir que el plàncton vegetal marí representa la realització d'un àmbit menor de possibilitats que les que es troben realitzades en el plàncton de les aigües continentals. L'ampliació del que podem anomenar «espai ecològic» en aquestes depèn del fet que les aigües dolces són sovint molt més fèrtils del que poden ser-ho les marines. Una conseqüència de la fertilitat és la secreció de part de l'assimilat. És natural que les secrecions visibles, en forma de mucíl·lag, caracteritzin un tipus biològic que enriqueix el fitoplàncton d'aigua dolça i és gairebé absent en el medi marí.



# PRESENCIA DE COMPSOPOGON MONTAGNE (RHODOPHYCOPHYTA, BANGIOPHYCEAE) AL DELTA DE L'EBRE

per XAVIER FERRER \* i FRANCISCO A. COMÍN

Departament d'Ecologia. Facultat de Biologia. Universitat de Barcelona

## ABSTRACT

*Compsopogon MONTAGNE 1846 is cited for the first time in the Iberian Peninsula. It was identified as Compsopogon aeruginosus (J. Ag.) KÜTZING with the key of Krishnamurthy (1962). But some differences in the measurements (specimens from the Ebro delta are two or three times wider than the references), spine-like branches, etc., cast doubt on this determination.*

*It was found on the muddy bottom of a coastal lagoon, a channel and a little pond in the Ebro delta. The characteristics of the medium were: stagnant or very slow running waters; shallowness (0.2-1 m); oligohaline (0.1-5 gr Cl. l<sup>-1</sup>); pH: 7.5-8.7; alkalinity: 1.5-4.9 meq. l<sup>-1</sup>; temperature: 8.5-26 °C. Salinity appears to exert a great influence on the presence of Compsopogon. This species was abundant the months of October, November and December, while the cover of macrophytes was down.*

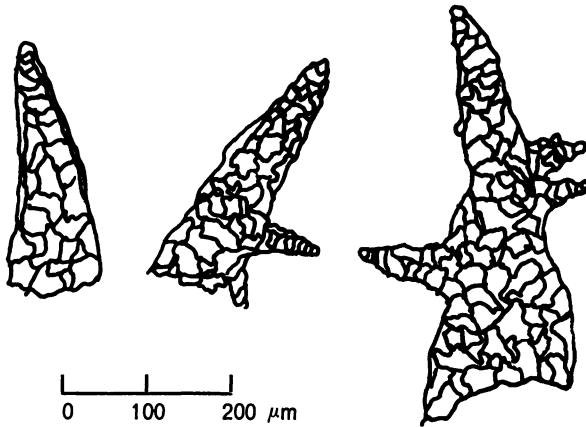
## INTRODUCCIÓ

El desembre de 1975 es trobà a l'estany de l'Encanyissada (delta de l'Ebre), la rodofícia *Compsopogon* Montagne. Aquest gènere té una distribució mundial comuna a les regions tropicals i subtropicals i ocasional a les àrees temperades, i ocupa aigües dolces o salabroses (KRISHNAMURTHY, 1962). A Europa es troben dues espècies: *C. coeruleus* i *C. corinaldii* restringides a unes poques localitats d'Anglaterra, de França, d'Itàlia i d'Alemanya (LUSINA, G., 1943; KRISHNAMURTHY, 1962; BOURRELLY, P., 1970). Referent a la Península Ibèrica, devers els anys cinquanta, MARGA-

\* Nova adreça: Departament de Zoologia. Facultat de Biologia. Universitat de Barcelona

LEF determinà una mostra d'aquest gènere procedent d'un estany dels voltants de Barcelona, recollida per un alemany (MARGALEF, com. pers.).

Els exemplars col·lectats al delta de l'Ebre han estat determinats emprant la clau dicotòmica de KRISHNAMURTHY (1962), que ens conduí a l'espècie *C. aeruginosus* Kützing 1849. La citació és força interessant puix segons la magra bibliografia relativa a *Compsopogon* de què disposem, seria la primera observació d'aquesta espècie a Europa. La seva distribu-



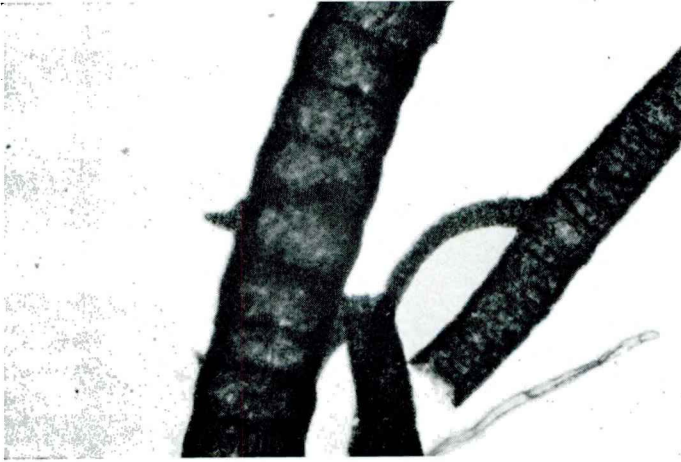
Espines de l'axis principal de *Compsopogon* del punt 1— Gola Nord del riet Zaida

Spine-like branches of the main axis in *Compsopogon* from 1— Gola Nord del riet Zaida

ció mundial s'extén sobre les regions tropicals i subtropicals del Nou Món, Filipines, Java i Israel. Tanmateix cal remarcar la diferència considerable d'alguns trets entre els nostres exemplars i els de KRISHNAMURTHY (1962). Els més palesos són referits a les dimensions (molt més grans els del Delta de l'Ebre) i a la presència d'espines compostes a l'axis principal. Tot això fa que malgrat considerar els exemplars com pertanyents a l'espècie *C. aeruginosus* posem en qüestió aquesta determinació fins que no quedi clar si els individus del Delta de l'Ebre són una altra espècie nova o una forma local més gran.

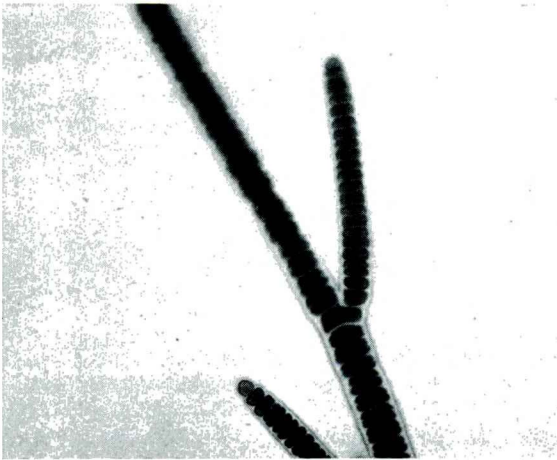
#### DESCRIPCIÓ DELS EXEMPLARS

A *Compsopogon* existeix una palesa diferenciació entre els filaments joves i els vells. Els joves són uniseriats i sense còrtex. Els vells, multiseriats, desenvolupen un còrtex d'1-2 capes de cèl·lules poligonals (KRISHNAMURTHY, 1962). El tret específic de *C. aeruginosus* és la presència d'espines a les porcions velles (KRISHNAMURTHY, 1962).



Filament jove sense còrtex i filament vell amb còrtex de cèl·lules poligonals i marcat anellament. Estació de mostreig: 3— l'Encanyissada

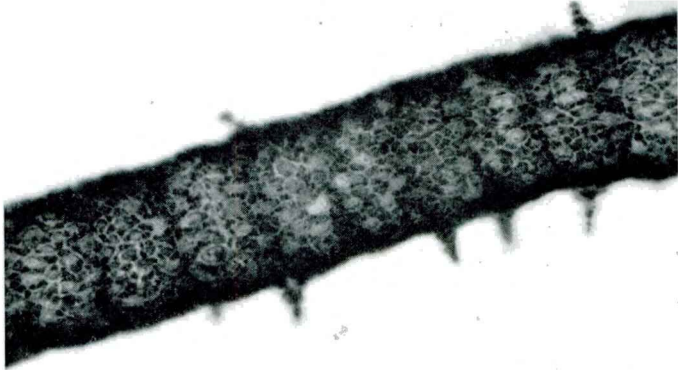
Young filament without cortex and old filament with a cortical investment of polygonal cells and marked septation. Sampling station: 3— l'Encanyissada



Filaments joves uniseriats. 3— l'Encanyissada

Uniseriate young filaments. 3— l'Encanyissada





Filament vell amb còrtex i espines. 3—l'Encanyissada  
Old filament with cortex and spine-like branches. 3— l'Encanyissada



En el cas del Delta de l'Ebre, els exemplars s'ajusten a la descripció d'aquest autor (llevat que no vam mirar el nombre de capes del còrtex) però les dimensions són entre 2 i 3 cops més grans. Compararem ara algunes mesures extretes d'exemplars de l'Encanyissada i de la Gola Nord del riet Zaida. El nombre de mesures per paràmetre ha estat de 10.

	KRISHNAMURTHY 1962	DELTA EBRE
Diàmetre de l'axis pral.	0,5 mm	1-1,5 mm
Diàmetre filam. uniser.	fins a 28 $\mu\text{m}$	fins a 25 $\mu\text{m}$
Cèl. uniseriades: h	6-16 $\mu\text{m}$	30-40 $\mu\text{m}$
Cèl. uniseriades: a	19-28 $\mu\text{m}$	110-140 $\mu\text{m}$
Cèl. uniseriades: forma	Discoïdals	Discoïdals-Rectang.

Es pot apreciar que les dimensions dels exemplars del delta de l'Ebre són considerablement més grans. Altres característiques són un color blau molt intens i un anellat molt conspicu.

Referent a la longitud de l'axis principal, del qual Krishnamurthy dóna la xifra de 20 cm, nosaltres no n'hem mesurat cap, però tenim mesures d'altres filaments que oscil·laven entre 20-45 cm amb una mitjana de 28 cm. Cal indicar també que les espines eren molt més grans i a l'axis principal moltes estaven ramificades tal com mostra el dibuix. Tal possibilitat d'espines compostes no és indicada per Krishnamurthy (1962).

#### ESTACIONS DE PRESA DE MOSTRES

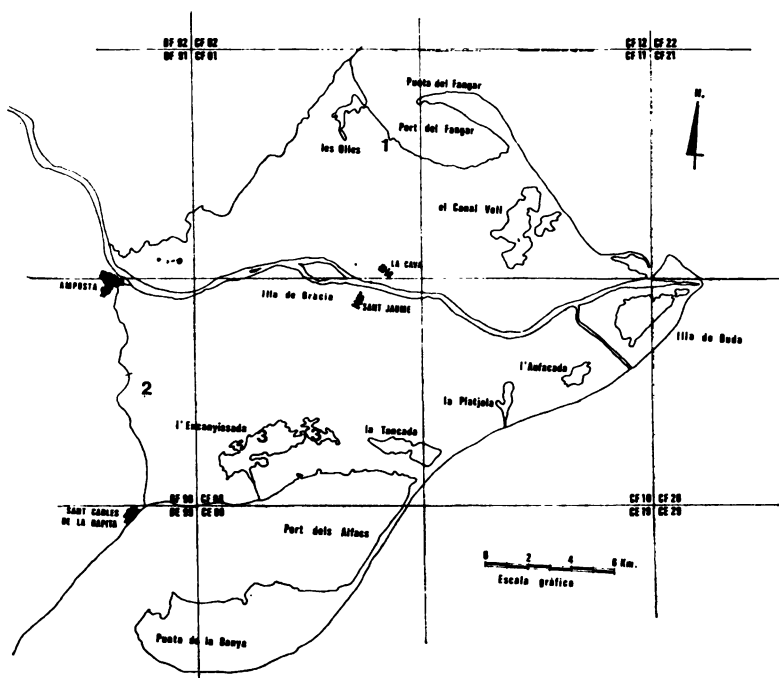
Tal com mostra el mapa, *Compsopogon* es trobà en tres estacions, les característiques de les quals pasem a descriure ara:

1. Gola Nord del riet Zaida — Antic canal, actualment és una bassa d'aigua salabroso-salada. Mostra agafada del fons, a uns 50 m del mar, el 8.10.76. (pH = 7,3. — Prof = 40 cm.)
2. Ullals del Prat del Notari — Basses d'aigua dolça que arriba per filtració subterrània. Corrent quasi nul. Mostra agafada del fons, tapissat totalment per aquesta rodofícia el 8.11.76. (pH = 7,2 — Prof = 60 cm.)
3. L'Encanyissada — Bassa de aigua salabrosa en contacte amb la mar. Mostres fixades sobre el fons fangós-sorrenc. Present de desembre de 1975 a febrer de 1976. Reapareixent de setembre de 1976 fins febrer de 1977. Durant 1976, el pH oscil·là entre 7,5 i 8,7 amb la mitjana durant els mesos de presència de *Compsopogon* de 8,1.

La temperatura varià entre 8,5 i 26 °C. Entre octubre i gener la mitjana es situà a 11,1 °C. (Prof: 0,20 — 1 m. Aigües molt tampoades amb una alcalinitat de 2,5 — 4,7 meq/l.)

### ECOLOGIA

La referirem a la tercera estació de presa de mostres: l'Encanyissada. De desembre de 1975 a maig de 1977, els autors han estudiat la limnologia d'aquesta bassa amb una periodicitat d'un mostreig mensual i de dos durant maig i juny.



*Mapa del Delta.* — Estacions de mostreig: 1— Gola Nord del riet Zaida; 2— Ullals del Prat del Notari; 3— l'Encanyissada. Quadrícula UTM

Sampling stations: 1— Gola N del riet Zaida. 2— Ullals del Prat del Notari. 3— l'Encanyissada. UTM system

L'estany té dos moments molt diferents. Entre maig i desembre, els canals de regadiu resten oberts i els conreus de l'arròs queden enaiguats. L'aigua és abocada després a l'estany, el qual esdevé quasi dolç; tenia 163-2146 mg Cl<sup>-</sup>/l. Dos mesos després de la collita de l'arròs (pels volts



de desembre) tanquen els canals i els conreus queden eixuts. El nivell de l'aigua a l'estany baixa entre 20-50 cm i la bassa esdevé progressivament salada, entre 1849-16 914 mg Cl/l. Referent als nutrients, els fosfats es troben en força quantitat; de setembre a febrer, una mitjana de 0,62 µgr at/l. i no arriba a ser factor limitant per a les poblacions. Altrament passa amb els nitrats que, durant els mesos d'agost, setembre i octubre, s'esgoten i actuen de limitants per certes espècies. A partir de novembre començaren a pujar i atenyeren el màxim durant desembre-gener-febrer.

La presència de *Compsopogon* a l'Estany de l'Encanyissada no és uniforme, ja que presenta preferència per una àrea (els abismes) especialment soma (40 cm de fondària mitjana) i també per les sortides dels canals que aboquen l'aigua a l'Encanyissada. Els Abismes presenten una vegetació de Caròfits i de *Najas marina* estacional de finals de juny a l'octubre, i desapareixen després totalment.

Els mesos de màxima cobertura de *Compsopogon* són octubre, novembre i desembre. Es troben també en petita quantitat, relegat a les boques dels canals, durant setembre, gener i febrer. Vistes les característiques limnològiques i la fenologia de l'espècie, veiem en primer lloc que la salinitat és un factor limitant per l'espècie, que ateny la màxima extensió quan les aigües són molt poc salabroses, la temperatura encara és càlida i no falten nutrients. La possibilitat de la competència és l'últim aspecte, ja que les susdites condicions es donen també durant maig i juny i no apareix *Compsopogon* (si més no, no l'hem trobat, encara que pot restar en el fons d'alguns canals). És simptomàtic la seva colonització del Abismes en un moment (octubre) on s'havien eliminat altres poblacions estivals d'algunes filamentoses (*Cladophora* sp.) i els altres macròfits dominants (*Najas marina*, *Nitellopsis obtusa* i *Chara* sp.) es descomponien ràpidament fins a desaparèixer a finals de novembre.

#### BIBLIOGRAFIA

1. BOILLOT, A.: 1958. «Sur la présence en France d'un *Compsopogon* (Rhodophyceae, Bangioideae)». *Bull. Soc. Phyc. France*, 4: 13-16.
2. BOURRELLY, P.: 1970. «Les Algues d'eau douce. Algues bleues et rouges». N. Boubee & Cie. Paris.
3. KRISHNAMURTHY, V.: 1962. «The morphology and taxonomy of the genus *Compsopogon* Montagne». *J. Linn. Soc. (Bot.)*, 58 (372): 207-222.
4. LUSINA, G.: 1943. «Il *Compsopogon corinaldii* nuovamente scoperto in Italia». *Ann. Bot.*, 22 (3): 85-89.
5. PUJALS, C.: 1967. «Presencia en la Argentina del género *Compsopogon* (Rhodophycophyta, Bangiophyceae)». *Com. Mus. Arg. Cienc. Nat. Bs. As. (Hydrobiol.)*, I (6): 47-57.



# COMPOSICIÓ I VARIACIÓ EN EL TERME DE DOS ANYS D'UNA COMUNITAT BENTÒNICA DEL RIU TER

per NARCÍS PRAT

Departament d'Ecologia. Facultat de Biologia.  
Universitat de Barcelona

## INTRODUCCIÓ

Hem estudiat una comunitat que es desenvolupa en les pedres del Ter en una zona de fort corrent. Durant dos anys hem observat l'evolució dels seus components, principalment la variació dels percentatges. Està situada aquesta zona del riu a uns 200 m. d'alçària sobre el nivell del mar prop de La Celler de Ter (La Selva). Hem estudiat principalment la fauna macrobentònica, la que restava filtrant la mostra per una malla de 250  $\mu$  i dintre d'ella especialment els insectes aquàtics. Degut a les diferències que presentaven vàrem separar la fauna superior de la inferior a cada pedra.

## RÉGIM HÍDRIC DE LA ZONA ESTUDIADA

La quantitat d'aigua que baixa pel riu ve regulada per l'embassament de Susqueda, de manera que segons sigui l'època o la pluviositat de l'any el cabdal és variable. Els dos anys de mostreig varen ésser molt variables, mentre 1974 va ésser un any sec i pràcticament només va baixar el cabdal mínim, l'any 1975 va ésser molt plujós i en certes èpoques baixava molta aigua; principalment per juny-juliol i a la tardor. Si tenim en compte que l'aigua venia del hipolimnion de l'embassament i es va produir una baixa de 6 °C en la temperatura de l'aigua podrem comprendre fàcilment els efectes sobre la fauna que veurem a continuació.

## COMPOSICIÓ DE LA COMUNITAT

Tal com hem dit vàrem diferenciar la part superior i la inferior de les pedres. La part superior és més homogènia que la inferior. Tot i això no vàrem poder evitar en moltes ocasions mostrejar les parts laterals per la qual cosa algunes espècies es presenten a ambdós cantons.

La part superior presenta un aspecte de color verd degut a les algues que creixen sobre les cases que fabrica el tricòpter *Psychomyia pusilla* aquestes cases recobreixen tota la pedra donant-se densitats larvàries de fins 300 larves per 100 cm<sup>2</sup>. Aquesta gran densitat es dona degut a la gran riquesa de nutrients que porta l'aigua i a la gran quantitat de materials en suspensió, materials, dels quals s'alimenta la larva tot filtrant-los. També hi viuen una gran quantitat de Quironòmids, Dípters, que es posen entre les cases a l'estiu, quan les pupes del tricòpter les abandonen. També un altre Tricòpter, *Hydyoptila*, es freqüent en aquesta part sobretot quan hi ha *Cladophora* desenvolupant-se, la qual cosa ens fa creure en una certa relació de l'alga i de l'insecte. Altres components són sempre força infreqüents (Taula I).

La part inferior és més variada quant a nombre d'espècies i la seva localització. Referint-nos a aquesta part hem de dir que es pot distingir de manera general (segons la inclinació i la localització de la pedra això varia) una zona lateral o marginal a on es produeix un cert barreig d'espècies. En aquesta zona (i a la part més superior) predominen tricòpters els quals fan la casa amb petits grans de sorra o de pedres com *Hydropsychae* i *Cheumatopsychae*, així com les larves errants de *Rhyacophila* que en el moment de la nimfosi construeixen una bossa posada dintre d'una d'aquestes cases fetes amb pedres, sempre molt més regulars que les de les dues espècies abans esmentades. En la composició de les cases poden formar-hi part closques de cargols, trossos de vidre, fils, trossos de plumes i materials més diversificats. En la part més central de la pedra es situen tricòpters que realitzen galeries sedoses per capturar el seu aliment, com *Polycentropus*, també hi ha uns efemeròpters amb les brànquies doblades cap a la part dorsal (*Ephemerella*). Entre els quironòmids cal destacar la presència dels tubs de *Tanytarsus*, no presents a la part superior.

## VARIACIÓ DE LA COMUNITAT EN EL TEMPS

Normalment la vida de les espècies observa d'aprop la variació de la temperatura, quan aquesta puja és quan els insectes augmenten el seu

ritme i les larves passen a pupes i després a adult. Segons l'aliment i la densitat de la població es poden donar una o diverses generacions anuals. En el nostre cas, amb una bona alimentació i una gran densitat de població es donen fins a tres generacions en alguns quironòmids i dues (encara que parcial la segona) al tricòpter *Psychomyia*, com veurem més tard.

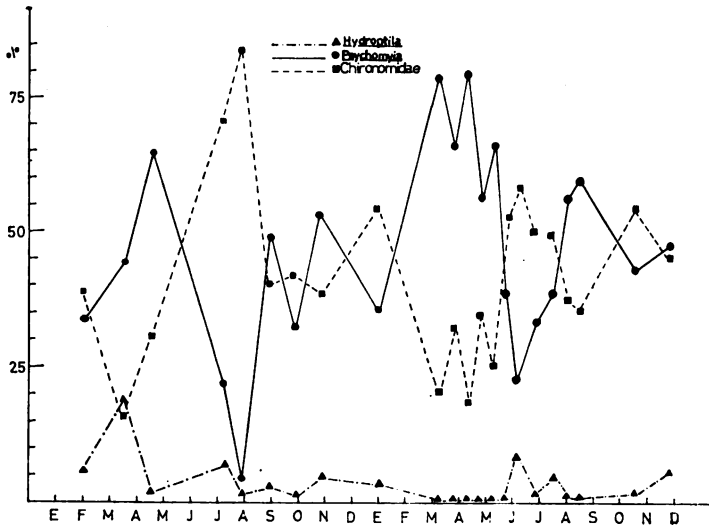


FIG. 1. — Variació dels percentatges de les espècies principals de la fauna superior de les pedres

A les Taules I i II es presenta la variació dels tants per cent de les espècies al llarg dels dos anys de mostreig segons sigui la part superior o inferior de les pedres. Es poden veure quines són les espècies o grups dominants i com a la part inferior la fauna és més variada.

A la fig. 1 hem representat la variació dels percentatges de la fauna de la part superior de les pedres, es veu una certa regularitat emmascarada el segon any per refredament de l'aigua que retorna la comunitat a una situació primitiva. En els primers moments els percentatges estan igualats, l'augment inicial del tant per cent de *Psychomyia* coincideix amb l'emergència (sortida d'adults) dels Quironòmids. Després en començar a l'abril les emergències del tricòpter, els tants per cents es van invertint i a l'agost són totalment oposats. En aquest moment s'ha acabat l'emergència de la primera generació de *Psychomyia* i les larves de Quironòmid de la segona generació són ja adultes emergint, i pel contrari es desenvolupa

TAULA I. — Variació del percentatge de les espècies de les pedres del riu Ter

MOSTRE DATA	1	2	3	4	5	6	7	8	9	10
<b>EFEMEROPTERS</b>										
<i>Baetis</i> sp.	19	3	3	—	—	0'17	0'64	1'29	0'16	1'1
<i>Ephemerella</i> sp.	—	4	—	1'5	+	1'03	—	—	—	—
<i>Caenis</i> sp.	—	2	1	—	—	—	—	0'11	0'16	—
<b>TRICOPTERS</b>										
<i>Rhyacophila</i> (larva)	1	0'5	—	+	—	0'17	—	—	0'8	0'1
<i>Hydroptila</i> (L+N)	8	6	19	—	7	2'97	3'33	2	5'16	4'1
<i>Hydropsychae</i> L	8	8	12	+	—	0'86	0'41	0'47	—	—
<i>Tinodes waeneri</i> L	—	—	+	+	—	2'75	—	—	—	—
<i>Psychomyia pusilla</i> L	59	34	44	49	11	4'70	47'27	32'5	53'06	35'1
N	—	—	—	16	11	0'34	5'18	7'68	—	—
<b>DIPTERS</b>										
Chironomidae L	2	39	17	31	70	83'4	40'28	41'13	37'25	54'1
N	—	—	—	—	—	2'24	0'51	2'83	1'45	1'1
Simuliidae	2	—	+	2	—	—	0'1	—	—	2'1
Tipulidae	—	1	—	—	—	0'17	0'2	0'11	1'45	0'1
Empididae	—	—	—	—	—	0'1	0'1	—	0'48	—
<b>MOLUSCS</b>										
<i>Ancylastrum fluviatile</i>	1	2	2	+	+	1'20	0'64	0'11	—	0'1

 TAULA II. — Variació del percentatge de les espècies de les pedres del riu Ter (pa  
veure)

MOSTRE	1	2	3	4	5	6	7	8	9	10
<b>EFEMEROPTERS</b>										
<i>Ephemerella</i> sp.	—	—	—	8	38	15'3	1'91	1'02	0'55	—
<i>Caenis</i> sp.	—	4	4	+	+	1'53	—	0'34	0'55	0'1
<i>Baetis</i> sp.	1'7	2	2	—	—	—	0'76	1'02	1'11	3'5
<b>TRICOPTERS</b>										
<i>Rhyacophila</i> L	—	0'5	2	+	+	—	0'38	—	0'27	2'1
N	—	—	—	—	—	—	0'76	1'02	—	1'1
<i>Hydroptila</i> (L+N)	—	—	2	+	+	1'53	0'38	—	0'27	1'1
<i>Hydropsychae</i> L	67	51	24	36	19	30'6	20'3	17'6	31'84	27'1
N	—	—	11	1	3	—	1'5	1'36	—	—
<i>Cheumatopsychae</i> L	12	6	15	22	4	—	24'8	35'83	30'72	16'1
N	—	—	—	—	5	—	—	0'34	—	—
<i>Tinodes waeneri</i> L	—	—	—	—	—	4'61	0'38	—	—	—
<i>Psychomyia pusilla</i>	—	—	+	8	+	—	1'14	5'44	7'26	1'1
<i>Polycentropus</i> L	—	9	4	—	1	—	11'49	1'02	4'18	0'1
N	—	—	—	—	—	—	3	—	—	—
<b>DIPTERS</b>										
Chironomidae L	12'3	28	13	14	25	30'6	20'6	26'27	18'43	33'1
N	—	—	—	—	—	—	1'14	1'70	0'55	5'1
Simuliidae	—	—	—	—	—	—	—	—	—	—
Tipulidae	—	—	—	—	—	—	—	—	—	—
Empididae	—	—	—	—	—	—	—	—	—	—
<b>COLEOPTERS (Ollimnius)</b>	—	—	—	—	—	—	—	0'68	—	—
<b>MOLUSCS</b>										
<i>Ancylastrum fluviatile</i>	3	2	6	2	2	15'3	3'8	—	1'11	3'1
<i>Pisidium</i>	4	0'5	17	1	2	1'53	4'98	4'43	0'55	0'1
<b>OSTRACODA</b>	—	—	—	—	—	—	—	—	—	—
<b>HIDRACARINA</b>	—	—	—	—	—	1	3'04	2'38	2'23	—
<b>HIRUDINEA</b>	—	—	—	—	—	3'07	—	—	—	—

t superior). + indica percentatge menor al 0'1 %. L vol dir larva i N pupa

	12	13	14	15	16	17	18	19	20	21	22	23
	25.3	5.4	20.4	3.5	18.5	1.6	28.6	11.7	28.7	9.8	12.10	30.11
5	—	—	—	—	—	0'59	0'33	—	1'08	2'04	0'56	1'44
	—	0'4	0'66	0'64	—	—	—	—	—	0'69	—	—
	—	—	—	—	—	—	—	—	—	—	—	—
	0'35	0'23	0'33	0'32	0'44	8'78	1'67	5'47	1'08	—	1'42	5'79
	—	—	—	—	—	—	—	—	—	—	—	—
0	65'49	77'98	56'81	66'5	38'65	23'75	34'11	39'72	56'14	52'71	42'73	47'82
	1'06	3'98	6'64	6'8	4'04	4'19	8'02	—	4'08	9'86	—	—
5	31'31	17'79	34'53	25'4	53'25	57'88	45'15	50'68	36'50	34'35	54'98	42'02
	1'06	—	0'99	—	0'22	0'79	6'68	1'36	0'54	—	—	2'17
	0'17	—	—	—	—	—	—	—	—	—	0'2	—
6	0'17	—	—	—	—	—	—	—	—	—	—	—
	—	—	—	0'32	—	—	0'33	—	—	0'34	—	—
	—	—	—	0'32	—	—	0'33	—	—	—	—	—

rior). + indica percentatge menor a 0'1 %. L vol dir larva i N pupa per les dates a I.

	12	13	14	15	16	17	18	19	20	21	22	23
5	3	6'41	11'3	4	—	1'31	1	1'69	4'95	6'14	5'10	2'66
	0'6	1'01	3'18	—	—	0'87	—	0'56	—	—	0'9	0'88
	0'9	0'33	1'81	1'6	0'56	—	—	—	3'49	1'11	2'70	0'59
	—	—	—	0'8	2'8	2'63	10	2'83	1'42	—	0'9	0'29
8	—	0'66	0'9	1'6	3'4	10'52	—	2'25	2'04	0'55	—	0'29
	—	—	—	—	0'56	1'31	—	0'58	0'87	—	0'3	—
3	12'76	15'87	5	10'4	8'46	8'77	2	17'46	11'14	21'78	24'92	34'90
3	0'30	2'36	2'72	—	3'38	1'31	—	3'94	—	1'11	0'9	—
1	12'61	15'53	7'72	16'8	10'15	9'21	—	2'25	7'57	6'14	3'03	3'84
	—	—	—	0'8	0'56	3'94	—	—	—	—	—	—
	19'66	12'1	9'09	16'8	5'64	10'08	6	3'38	4'66	2'79	4'20	1'77
3	0'75	1'35	0'45	—	—	—	27	12'99	—	—	2'70	4'43
3	—	1'35	0'45	—	—	—	6	—	—	—	1'20	0'29
7	36'63	30'06	36'31	36'8	46'22	27'31	9	22'57	46'35	55'86	37'50	38'14
5	4'20	5'06	6'15	5'06	5'07	2'5	3	8'47	2'33	1'67	3'63	3'73
	—	0'33	0'9	—	0'56	—	—	—	1'74	—	—	—
5	0'3	—	0'45	—	—	—	—	0'56	1'45	0'56	—	—
5	—	0'63	—	—	—	—	—	—	—	—	—	0'29
3	0'15	—	2'27	0'8	1'69	7'01	—	1'12	—	0'56	0'3	—
5	7'35	2'02	2'27	3'2	—	—	3'	0'56	2'04	—	3'93	4'73
3	0'15	2'02	1'81	—	0'56	5'26	3	3'38	0'58	—	—	—
	0'3	—	2'27	—	5'07	5'7	21	8'47	1'16	—	1'80	1'77
2	0'3	3'04	4'09	0'8	4'51	1'75	7	6'37	5'87	—	2'10	0'29
	—	—	—	—	—	—	—	—	—	—	1'5	0'29

Iupen les larves de la segona generació de tricòpters per la qual cosa els percentatges es tornen a invertir. A partir d'aquí hi ha un cert equilibri amb unes oscil·lacions que podrien ser causa de l'emergència d'alguns quironòmids de la tercera generació de larves (octubre-novembre) i d'algunes *Psychomyia* de la segona generació. Al gener del 1975 els percentatges tornen a estar equilibrats, i comença de nou el cicle el mes de març amb el mateix signe; aquest mes també comencen les emergències de Quironòmids i a l'abril-maig les de *Psychomyia*, però a mitjans de juny es produeix l'alliberament d'aigua freda per l'embassament, i això fa que la situació no arribi a l'extrem de l'any anterior i al juliol i a l'agost trobem uns percentatges molt diferents amb una situació de tipus més primaveral. Els augments són deguts a que des de finals de juliol a mitjans d'agost l'aigua que baixava ho feia amb temperatura més elevada per la qual cosa es va poder desenvolupar una quantitat de larves de la segona generació que potser abans de les altres crescudes de setembre varen donar lloc a una segona generació d'adults més minsa que l'any abans. Noves inundacions al setembre no van permetre que es desenvolupés massa la comunitat i al mostrejar de nou a l'octubre ja no hi ha havia pupes tornant els tants per cents a una situació equilibrada.

Veiem doncs com la regulació per l'embassament del cabdal del riu produeix uns efectes variables segons els anys. El conèixer la variació dels estats larvaris d'una de les espècies més comuns varem creure que ens aclariria els dubtes sobre el seu cicle encara que ens faltessin algunes mostres intermitges.

#### VARIACIÓ DELS ESTATS LARVARIS DE LA POBLACIÓ DE *Psychomyia pusilla*

A la fig. 2 s'expressa la variació en % de classes d'edat de les larves de *Psychomyia pusilla* i les pupes (N), determinats per la medició de l'amplada del cap. Es separen clarament 5 estadis larvaris a més de les pupes. En alguns casos es pot endevinar al 5<sup>e</sup> estadi dos grups de larves que podrien correspondre a mascles i a femelles.

Amb correlació amb els tants per cents observem que el primer any per juliol hi ha poques larves, el 50 % de la població són pupes, la mostra següent és encara la finiquitació de la primera generació. A les mostres de setembre, d'octubre i de novembre dominen en canvi els estadis joves. Potser algunes de les primeres larves de la segona generació nascudes a finals d'agost podrien per setembre o per octubre passar ja a ser adultes.

Al segon any observem com la població varia molt lentament en tenir mostres cada quinze dies. Poc a poc es va desplaçant l'equilibri cap a



les classes d'edat més grans fins que per juny ja gairebé totes estan a l'últim estadi, és aquí quan l'aigua baixa de temperatura, per això aquesta acumulació fa que el canvi de pupa a adult sigui més lent. Al mig de les mostres de juliol i d'agost es va comprovar que no hi havia pupes,

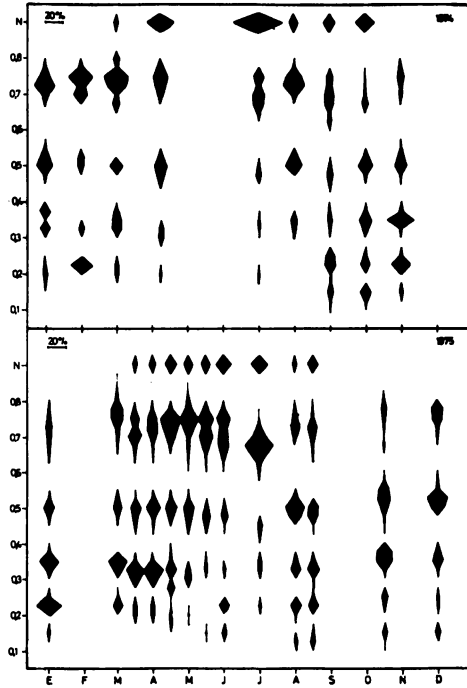


FIG. 2 — Variació en els dos anys de mostreig de les diferents classes d'estadis larvaris de *Psychomyia pusilla* en la part superior de les pedres del riu Ter

i va sortir després una segona generació de pupes potser provinents encara de la primera generació larvària. Cal observar també com els tants percents més grans corresponen a formes joves de la segona generació. Potser algunes d'aquestes arribin a adultes, encara que ens sembla que no és probable, a l'octubre i a finals de novembre la situació és ja com l'any abans, sense pupes.

## CONCLUSIONS

—La fauna aquàtica d'aquesta zona del Ter assoleix una gran densitat (fins a 1000 ind, per pedra: per 200/cm<sup>2</sup> aprox.) com a resultat de la gran quantitat d'elements nutritius que porta el riu, la qual cosa genera una gran producció primària per una part i sobretot un aport de matèries en suspensió.

—El poblament petrícola es diferencia clarament en dos comunitats, la superior i la inferior amb diferents espècies. Tanmateix hi ha un cert barreig a la zona lateral. La comunitat inferior està més estructurada i diversificada per la major protecció al corrent que té.

—El cicle anual de les espècies d'insectes ve, donat que no manca aliment, regulat per la temperatura i aquesta està directament influenciada pel cabdal regulat per l'embassament de Susqueda, d'aquí la importància de la regulació d'aquest sobre el riu. El refredament de l'aigua pot provocar que una part de la biomassa dels insectes no emergeixi.

—L'estudi dels estats larvaris de *Psychomyia* ens ha permès d'esbrinar-hi el nombre i també d'observar la influència de la regulació del riu per l'embassament.

—No cal oblidar que la sortida dels insectes del riu en suposa una pèrdua de la biomassa, la qual cosa és un sistema d'autodepuració del riu del qual se n'aprofiten altres depredadors com ocells i rat-penats que viuen prop del riu.

AGRAÏMENTS. — Voldria agrair al Dr. Margalef, cap del Departament d'Ecologia de la Universitat Central de Barcelona, les facilitats i l'ajuda per la realització d'aquest treball i al meu germà Josep l'ajuda per a l'obtenció d'algunes mostres.

# LA DESNATURALITZACIÓ PARCIAL DEL DNA EN N,N'DIMETILFORMAMIDA

per JORDI CORTADAS i JOAN A. SUBIRANA

Departament de Química Macromolecular del C.S.I.C.  
Universitat Politècnica de Barcelona

La desnaturalització del DNA per dissolvents orgànics (alcohols alifàtics, urea, formamida, N,N'dimetilformamida...) ha estat amplament estudiada per diferents autors. Se sap que aquest procés es influenciat per una sèrie de factors: naturalesa del dissolvent, força iònica, temperatura i composició mitjana en bases de la mostra de DNA. Hi ha, però, una sèrie de fets que no han estat aclarits. En el cas de la formamida s'ha vist que, en presència d'una determinada quantitat de sal, el procés de desnaturalització no és complert. Els mecanismes implicats en aquest fenomen no han estat estudiats.

El present treball ha consistit en l'estudi de l'estabilitat de DNAs de diferents fonts, enfront de la N,N'dimetilformamida, en funció de la força iònica. Hem tractat un DNA víric (bacteriòfag  $\phi$  29), un bacterià (*Micrococcus lisodeikticus*) i un d'un organisme superior (lletó de vedella). Els resultats obtinguts demostren que no tots posseeixen la mateixa estabilitat. El control dels diferents factors que intervenen en el procés ens ha permès poder donar, en una primera aproximació, una interpretació d'aquests resultats. Hem conclòs que el comportament dels DNAs tractats pot ser explicat a partir de la diferent solubilitat de cada un, i que depèn de la composició en bases i del grau d'heterogeneïtat del DNA.



# TRENCAMENT DE PROTAMINES AMB PEPSINA. SEPARACIÓ I CARACTERITZACIÓ DE LLURS PÈPTIDS

per AGNÈS JORDAN i JOAN A. SUBIRANA

Departament de Química Macromolecular del C.S.I.C. Escola Tècnica Superior  
d'Enginyers Industrials de Barcelona. Universitat Politècnica de Barcelona

Atesa la importància que té el coneixement de l'estructura primària de les proteïnes nuclears, s'ha projectat establir la seqüència de la protamina del caragol i del component  $\varphi_0$  de l'esperma d'holotúria.

En aquesta comunicació es descriurà el treball preliminar realitzat, en el qual s'han estudiat els següents aspectes: *a)* aplicació del mètode de G. Bretzel per digerir enzimàticament protamines amb pepsina, a la protamina del caragol i al component  $\varphi_0$ , *b)* separació per cromatografia de tamís molecular dels pèptids obtinguts en la digestió del component  $\varphi_0$ , *c)* caracterització d'aquests darrers pèptids per mitjà de llurs anàlisis d'aminoàcids ensems amb la determinació de llurs C i N terminals.

El treball realitzat ens ha dut a l'obtenció de 5 pèptids en la digestió de la protamina del caragol i de 4 en la del component  $\varphi_0$ . Posteriorment, s'ha aconseguit separar dos dels quatre pèptids obtinguts en la digestió del component  $\varphi_0$ , per mitjà de Biogel P-10. S'ha determinat la seva composició d'aminoàcids així com les seves N terminals pel mètode del DNS-Cl i C terminals per mitjà de digestions amb carboxipeptidases i finalment s'ha proposat un esquema de la col·locació d'aquests pèptids en el component  $\varphi_0$ .



**XIII**

**MATEMÀTICA I BIOLOGIA**





*Els dies 21 i 22 d'abril de 1977 tingué lloc la Reunió Científica Anual Plenària de la SOCIETAT CATALANA DE BIOLOGIA, que fou dedicada a "Matemàtica i Biologia". Les sessions foren celebrades al Departament d'Ecologia de la Facultat de Biologia de la Universitat de Barcelona.*

*La pàgina següent conté un index dels temes presentats. La trobada tingué un caire de reunió de treball on va jugar un paper important la discussió informal dels problemes plantejats. Aquest volum recull els originals que foren lliurats per a ésser publicats. Les discussions que els seguiren no pretenen reproduir paraula per paraula tot el que es va dir; intenten ser només una exposició concisa de les idees comentades.*



# RELACIONS ENTRE MITJANES I VARIÀNCIES DE MESURES DE BIOMASSA FITOPLANCTÒNICA

per MARTA ESTRADA i MIYARES

Institut d'Investigacions Pesqueres, Barcelona

## ABSTRACT

Mean-variance relationships of planktonic biomass measurements

*Mean values ( $m$ ) and variances ( $s^2$ ) of series of phytoplankton and particle counts, taken in the NW Africa upwelling region, followed the classical relationship:*

$$\log(m) = B \log(s^2) + A$$

*The same occurred when biomass measurements in terms of chlorophyll concentration or fluorescence were used. A simplified model shows that biomass distributions due to turbulent diffusion from a high density patch may originate such relationships between means and variances.*

## INTRODUCCIÓ

Les distribucions de freqüències i les relacions entre mitjana (castellà «media») i variàncies de sèries de comptatges d'organismes plantegen molts problemes interessants. L'estudi empíric de la correlació entre mitjana i variància per a dades d'aquest tipus indica que, la major part de les vegades, el núvol de punts s'ajusta a la relació:

$$\log(s^2) = A + B \log(m), \quad (1)$$

on  $s^2$  és la variància,  $m$  la mitjana i  $A$  i  $B$  són constants (TAYLOR, 1961).

Es demostra (FRONTIER, 1973) que la transformació que estabilitza la variància és de la forma:

$$g(x) = C f^{x-t-B/2} dt, \quad (2) \text{ on } C \text{ és una constant.}$$

D'ací es deriva, prescindint de constants aditives i factors multiplicatius, quer per a  $B \neq 2$ ,  $g(x) = x^{-B/2}$ . Per tant, si  $B = 1$  es té  $g(x) = \sqrt{x}$  (distribució de Poisson). Si  $B = 2$ ,  $g(x) = \int^x dt/t = \ln(x)$ , (distribució log-normal). A partir d'aquests resultats s'han proposat diverses explicacions teòriques de la relació (1).

En aquest treball es comenten les relacions entre mitjana i variàncies de seqüències de comptatges i altres mesures de biomassa del fitoplàncton emprades en oceanografia. La generalitat amb la qual s'obtenen relacions empíriques del tipus (1) fa pensar que es pot recórrer a mecanismes senzills per a interpretar, almenys parcialment, aquests resultats. Una possible hipòtesi es relaciona amb les distribucions a què dona lloc la difusió turbulenta d'una substància (o població d'organismes planctònics) continguda en un volum d'aigua. Per a simular aquest fenomen s'introdueix un simple model matemàtic i es comparen les relacions mitjana-variància obtingudes amb les trobades en condicions naturals.

## MATERIALS I MÈTODES

### *Dades de camp*

Les dades considerades en aquest treball foren obtingudes a la campanya ATLOR II del vaixell *Cornide de Saavedra*, portada a terme el març del 1973 a la zona d'afiorament del NO d'Àfrica.

Durant els transectes entre les estacions hidrogràfiques, aigua de 3 m de fondària es bombejava al laboratori de forma contínua. A cada minut s'enregistrava automàticament la temperatura de l'aigua, les concentracions de nitrit i nitrat i la fluorescència. Cada mitja hora es prenen mostres per a l'anàlisi del fitoplàncton, determinació de clorofilla i recompte de partícules de sis grups de mida amb un Coulter Counter. Si es considera una velocitat de creuer del vaixell de 8 nusos, l'interval de mostratge resulta aproximadament de 7,2 km.

Per a les mostres de fitoplàncton es fixaven, amb solució de Lugol, 120 ml de l'aigua bombejada. Un cop al laboratori, es deixaven sedimentar 100 ml de mostra en una cubeta de fons mòbil. Per comptar les espècies més nombroses s'observava, a gran augment (400x), un transecte del fons de la cubeta equivalent a uns 3 ml de mostra. Després s'examinava a petit augment (100x) tot el fons de la cubeta per a anotar els individus d'espècies menys abundants i de mida suficientment gran per a ser vistos. Per a fer els càlculs, les dades es varen expressar en cèl·lules per 100 ml. Aquesta estandarització fa que les variàncies de les espècies comptades en 3 ml quedin multiplicades pel factor  $(100/3)^2$ , mentre que les mitjanes

ho són només per  $(100/3)$ . Teòricament, això no afecta la pendent de la recta de regressió entre els logaritmes de les mitjanes i els de les variàncies, però sí la posició d'aquesta recta, o sigui, el terme independent. Es veu fàcilment que sí, en general:

$\log(s^2) = p \log(m) + q$ , on  $m$  és la mitjana,  $s^2$  la variància, i  $p$  i  $q$  són constants, i tenim

$$\begin{aligned} m_{st} &= (100/3)m_3; \quad s_{st}^2 = (100/3)^2 s_3^2 \quad \text{i} \\ \log(s_{st}^2) &= p \log(m_3) + q, \quad \text{es compleix} \\ \log(s_{st}^2) &= p \log(m_{st}) + (2-p) \log(100/3) + q. \quad (3) \quad \text{on} \end{aligned}$$

$m_3$ , i  $s_3^2$  són la mitjana i la variància dels comptatges fets a partir de 3 ml de mostra i  $m_{st}$  i  $s_{st}^2$  la mitjana i la variància després d'estandaritzar les dades a 100 ml. Resulta, però, que  $p$  és en general molt pròxima a 2 i el terme de (3) multiplicat per  $(2-p)$  esdevé pròxim a zero. Per això, en aquest treball preliminar s'han considerat espècies i sumes d'espècies estimades a partir de 3 o de 100 ml de volum sedimentat. Tampoc no s'han estudiat a part les espècies colonials. Més detalls sobre els mètodes de mostrejat i d'anàlisi poden trobar-se a ESTRADA (1974, 1976) i MARGALEF (1974).

### *El model*

Per a les simulacions amb ordinador es va emprar un model que, essencialment, reproduïda la sembra a l'atzar, en una àrea determinada, d'una sèrie de pertorbacions consistents en centres a partir dels quals difonia una variable (que podia representar, per exemple, una substància contaminant o una població de plàncton). Es tracta d'un model intencionadament simple, que intenta només il·lustrar els efectes de processos que poden actuar a la realitat, com és, en aquest cas, la difusió turbulenta. Cal però aclarir la interpretació d'algunes assumpcions fetes.

D'una banda, les simulacions suposen l'existència d'unes zones favorables on es desenvolupen de forma relativament ràpida altes densitats de població o taques de plàncton que actuen com a centres de difusió. En condicions naturals, hi ha molts mecanismes que poden donar lloc a l'aparició d'aquestes taques; per exemple, la ruptura d'ones internes o els fenòmens d'afforament, és a dir, la pujada a la superfície de volums d'aigua fonda rica en nutrients. Com a primera aproximació, en el model s'ha considerat que els centres eren puntuals en relació amb l'interval de mostrejat i que es distribuïen a l'atzar en un espai de dues dimensions.

D'altra banda, a les simulacions s'ha suposat que un cop formades les taques, els canvis de la població per reproducció o mortalitat eren

despreiables en front dels deguts a difusió. Una situació com aquesta és plausible si considerem que, inicialment, el creixement del fitoplàncton en un volum d'aigua fertilitzat és de caràcter exponencial, però ben aviat es veu frenat per l'esgotament dels nutrients o la manca de llum. L'estudi de les dades utilitzades en aquest treball (ESTRADA, 1976) suggeria aquestes assumpcions, però cal remarcar que la seva validesa depèn de les escales espacials i temporals que es considerin.

En un primer grup de simulacions s'utilitzà una equació de difusió de forma exponencial negativa; la intensitat de la pertorbació  $S$  en el punt  $j$  s'expressava per:

$$S_j = \sum_{i=1}^n C \exp(-r_i/W'), \quad (4)$$

$r_i$  és la distància des del centre  $i$  al punt de mostratge  $j$ ,  $n$  és el nombre de centres i  $C$  i  $W'$  són constants.

En un segon grup de simulacions es va emprar l'equació:

$$S_j(t) = \sum_{i=1}^n \frac{C}{W^2 t^2} \exp(-r_i^2/W^2 t^2), \quad (5)$$

Aquesta expressió va ésser deduïda (OKUBO, 1962) per OKUBO i PRITCHARD, i és una de moltes equacions semblants proposades per a representar la concentració teòrica ( $S_j$ ) en el temps  $t$  i el punt  $j$ , d'una substància que difon a partir d'un punt d'aportament instantani d'una quantitat  $C$  d'aquesta substància, per unitat de profunditat,  $r_i$  és la distància del punt d'aportament  $i$  al punt  $j$ ,  $t$  és el temps i  $W$  representa una velocitat de difusió.

En varen calcular les distribucions resultants de l'establiment de 1, 3 i 9 centres de difusió; per a preparar les figures s'utilitzaren les mitjanes i variàncies corresponents a tres ordenades fixes de l'àrea considerada en la simulació.

## RESULTATS

A la figura 1 es representa el núvol de punts corresponent a les mitjanes i variàncies de sèries de 15, 30 i 60 mostres consecutives de transectes fets els dies 17, 18 i 21 de març. Resultats semblants s'obtingueren amb registres d'altres dies, que no s'inclouen per no complicar les gràfiques. Com

es pot veure, existeix una correlació positiva entre el logaritme de la mitjana ( $\log m$ ) i el de la variància ( $\log s^2$ ); és interessant també assenyalar que la pendent d'una recta de regressió ajustada visualment al núvol de punts és més gran que 2, valor que correspondria a una distribució log-normal de freqüències de les dades originals.

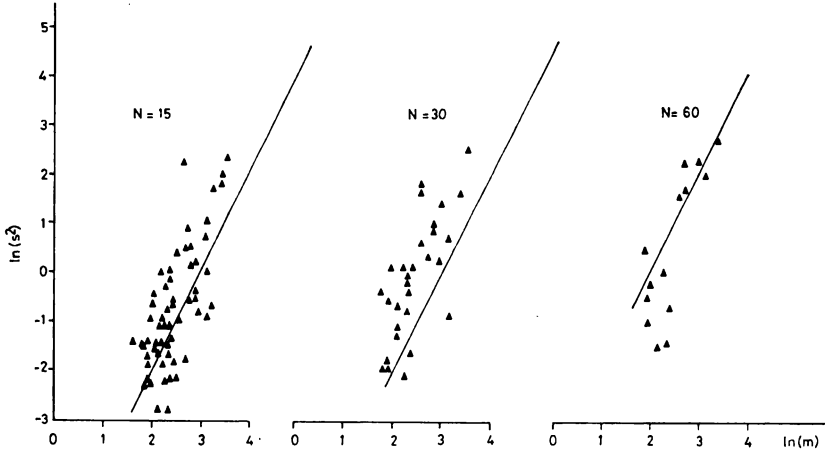


FIG. 1. — Mitjana ( $m$ ) i variàncies ( $s^2$ ) de grups de 15, 30 i 60 observacions ( $N$ ) successives de fluorescència preses a intervals d'un minut en transectes fets els dies 17, 18 i 21 de març. Les línies inclinades, de pendent, s'han posat com a referència

La figura 2 mostra la posició dels punts corresponents a les abundàncies individuals d'una sèrie d'espècies i als totals de diversos grups, per a les 213 mostres disponibles. S'han assenyalat també els valors obtinguts per als sis canals de partícules. Com es pot veure, a partir de  $m = 10$ , el núvol de punts s'ajusta bastant bé a una relació de la forma:

$$\log(s^2) = 5/3 \log(m) + K, \text{ amb } K = \text{constant.}$$

Per a valors de  $m$  inferiors a 10, la pendent de la línia de regressió s'aproxima a 1. Aquest efecte de pas d'una pendent pròxima a 2 (distribució log-normal) a una altra pròxima a 1 (POISSON) per a valors baixos d'abundància ha estat notat per FRONTIER (1973).

Amb la finalitat d'estudiar la influència de l'extensió de l'espai mostral en les relacions mitjana-variància, es varen repetir els càlculs després de dividir la seqüència de 213 mostres en grups de diverses mides. Les figu-

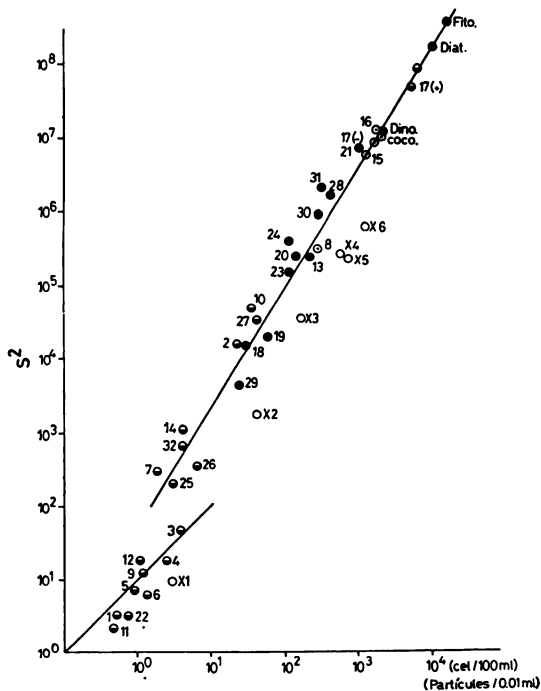


Fig. 2. — Punts (mitjana, variància) per a les abundàncies de diverses espècies i grups de fitoplàncton i per als comptatges de partícules corresponents a les 213 mostres preses durant la campanya. Les línies inclinades, de pendent 5/3 la superior i 1 la inferior, s'han posat com a referència. Espècies: 1. — *Blepharocysta splendormaris*; 2. — *Brachydidinium capitatum*; 3. — *Ceratium furca*; 4. — *C. fusus*; 5. — *C. kofoidii*; 6. — *C. tripos*; 7. — *Diplopsalis cf. asymmetrica*; 8. — *Exuviaella sp.*; 9. — *Gonyaulax fragilis*; 10. — *Gymnodinium sp.*; 11. — *Peridinium brochii*; 12. — *P. diabolus*; 13. — *Peridinium pl. sp.*; 14. — *Prorocentrum rostratum*; 15. — Dinoflagelats petits no identificats; 16. — Flagelats; 17. — *Coccolithus huxleyi*, (+) afegint les cèhules que correspondrien als coccolits, (—) sense afegir-les; 18. — *Helicosphaera pl. sp.*; 19. — *Syracosphaera sp.*; 20. — *Syracosphaera cf. pulchra*; 21. — *Amphora hyalina*; 22. — *Asteromphalus heptactis*; 23. — *Chaetoceros affinis*; 24. — *Ch. didymus*; 25. — *Ch. peruviana*; 26. — *Planktoniella sol*; 27. — *Rhizosolenia alata*; 28. — *R. fragilissima*; 29. — *R. imbricata*; 30. — *R. stolterfothii*; 31. — *Thalassiosira cf. parthenica*; 32. — *Coscinodiscus alborani*.

*Diat.*: nombre total de diatomes; *Dino.*: id. de dinoflagelats; *Coco.*: id. de coccolitoforals; *Fito.*: fitoplàncton total.  $X_1$ : nombre de partícules corresponent al canal 1 de treball del Coulter Counter (més grans de 34,6  $\mu\text{m}$ );  $X_2$ : id. canal 2 (21,8  $\mu\text{m}$ ) menys canal 1;  $X_3$ : id. canal 3 (13,7  $\mu\text{m}$ ) menys canal 2;  $X_4$ : id. canal 4 (8,6  $\mu\text{m}$ ) menys canal 3;  $X_5$ : id. canal 5 (6,8  $\mu\text{m}$ ) menys canal 4;  $X_6$ : id. canal 6 (5,4  $\mu\text{m}$ ) menys canal 5.

Espècies l'abundància de les quals s'ha estimat en examinar tota la mostra sedimentada (100 ml)).

⊙ Id. examinant una part equivalent a 3 ml del total sedimentat.

Id. examinant en general 3 ml i alguna vegada 100 ml de mostra, o bé sumes d'abundàncies de diferents espècies.

res 3 i 4 mostren la distribució d'una sèrie de punts corresponents a una part de les variables, preses com a exemple: La concentració de partícules al canal 5 ( $X_5$ ), clorofil·la, nombre de diatomes per mil·lilitre, nombre total de cèhules per mil·lilitre i nombre de *Ceratium furca* per 100 mil·lilitres. Per a les altres variables analitzades, els resultats foren semblants. S'obtenen núvols de punts bastant lineals, de pendent pràcticament igual a 2 en el cas de les cèhules i més gran en el de clorofil·la i partícules.



FIG. 3. — Punts (mitjana, variància) per a seqüències de diversa longitud (n = nombre d'observacions incloses en cada cas) de valors de concentració de clorofilla (clor. a, mg m<sup>-3</sup>) i de concentració de partícules corresponents al canal 5 (X<sub>5</sub>, partícules/o,01 ml). La línia inclinada té pendent 2 i s'ha posat per a referència

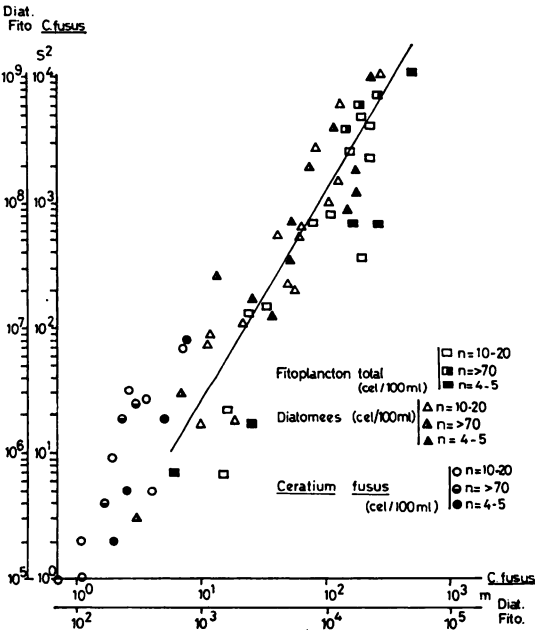
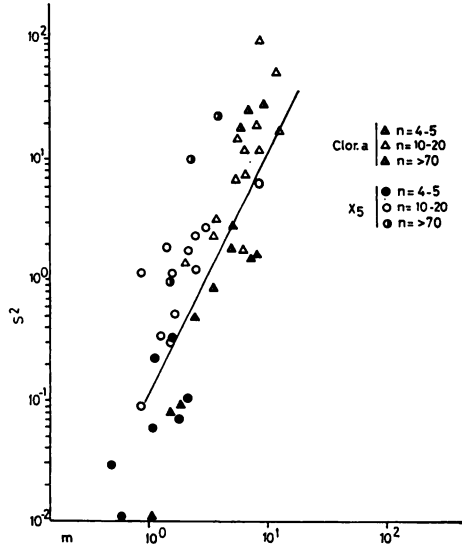


FIG. 4. — Punts (mitjana, variància) per a seqüències de diversa longitud (n = nombre d'observacions incloses en cada cas) de valors de concentració del fitoplankton total cèHules/ml, diatomes (cèHules/ml) i *Ceratium fusus* (cèHules/100 ml). La línia inclinada, de pendent 2, s'ha posat per a referència. A ordenades, l'escala de la dreta es refereix a *Ceratium fusus*, la de l'esquerra a la concentració total de fitoplankton i a la de diatomees

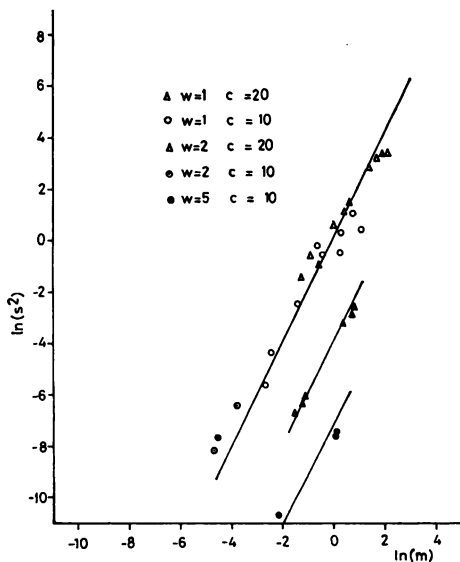
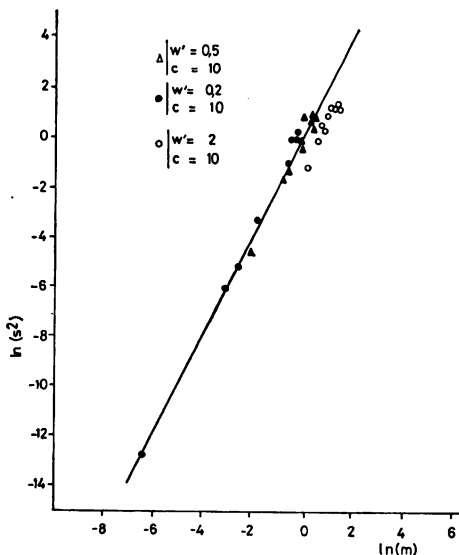


FIG. 5. — Mitjanes i variàncies de transectes presos sobre simulacions d'un procés de sembra a l'atzar de perturbacions amb distribució exponencial negativa a partir del seu punt d'origen (equació 3).  $W$ , indica la intensitat de difusió;  $C$ , l'amplitud de la perturbació;  $m$ , la mitja i  $s^2$ , la variància. La línia inclinada té pendent 2

FIG. 6. — Mitjanes i variàncies de transectes presos sobre simulacions fetes a l'emprar l'equació (4). S'ha suposat  $t$  (temps)=1.  $W$ , indica la intensitat del procés de difusió,  $C$ , l'amplitud de la perturbació;  $m$ , la mitja i  $s^2$ , la variància. La línia inclinada té pendent 2



Com es pot veure a les figures 5 i 6, el model de simulació proporciona també relacions lineals entre el logaritme de la mitjana i de la variància. La pendent dels núvols de punts és pròxima a 2 en tots els casos, però

la posició de les rectes de regressió varia per a diferents valors de  $W$  o  $W'$ , paràmetres que indiquen la intensitat del procés de difusió; en augmentar aquesta intensitat, disminueix la variància per a un mateix valor de la mitjana.

### DISCUSSIÓ

Una de les teories proposades per a explicar relacions del tipus (1) es basa en considerar la taxa de variació de biomassa en resposta a una pertorbació ambiental com a proporcional, en primera aproximació a la biomassa existent. Una distribució normal de la intensitat d'aquests estímuls donaria lloc a una distribució log-normal de les biomasses (CASSIE, 1962, MARGALEF, 1963). CASSIE exposà la hipòtesi d'una llei de probabilitat POISSON + log-normal; el nombre d'organismes per mostra seguiria una llei de POISSON, el paràmetre de la qual es distribuïria log-normalment en un domini d'espai o temps més ampli. FRONTIER (1973) aplicà aquesta hipòtesi a dades de concentració de zooplàncton però no va obtenir una coincidència acceptable amb les prediccions teòriques derivades de la idea de CASSIE i va arribar a la conclusió que la justificació de les lleis estadístiques observades havia de ser molt més complexa.

El pas progressiu de la llei de POISSON a una log-normal o a altres lleis com ho demostra el cas de valors de la pendent més grans que 2) tradueix únicament, segons FRONTIER, l'efecte del mostreig; en augmentar l'heterogeneïtat de l'espai mostral, la pendent ( $B$ ) tendiria a 2.

El fet que mitjanes i variàncies d'espècies molt diferents i àdhuc variables amb un elevat component no-vivent (com la concentració de partícules mesurada amb el *Coulter Counter*) s'ajustin a distribucions similars indica que, en les condicions de presa de mostres d'aquest estudi, les explicacions s'han de trobar en fenòmens molt generals. D'altra banda, en el cas de les mostres recollides en seqüència, el mateix tipus de relació empírica s'obté amb nombres d'observacions molt diferents. En el cas dels models, l'equació de difusió emprada no afecta les relacions trobades entre mitjana i variància, com ho mostra la comparació de les figures 5 i 6. El que és important, és que hi figurin termes de tipus exponencial respecte a la distància al centre de difusió. Aquestes consideracions suggereixen que en el cas de dades seqüencials de biomassa fitoplanctònica, l'actuació de processos de difusió turbulenta sobre nuclis d'alta densitat de població pot jugar un paper important en la generació de distribucions amb relacions empíriques senzilles entre mitjanes i variàncies.

## BIBLIOGRAFIA

1. CASSIE, R. M.: 1962. Frequency distribution models in the ecology of plankton and other organisms. *J. Anim. Ecol.*, 31: 62-92.
2. ESTRADA, M.: 1974. Photosynthetic pigments and productivity in the upwelling region of NW Africa. *Téthys*, 6(1-2): 247-260.
3. ESTRADA, M.: 1976. Estudios sobre poblaciones de organismos acuáticos en medio no uniforme. Tesis doctoral. Universitat de Barcelona.
4. FRONTIER, S.: 1973. Étude statistique de la dispersion du zooplancton. *J. exp. mar. Biol. Ecol.*, 12: 229-162.
5. MARGALEF, R.: 1963. Algunas regularidades en la distribución a escala pequeña y media de las poblaciones marinas de fitoplancton y en sus características funcionales. *Inv. Pesq.*, 23: 169-230.
6. MARGALEF, R.: 1974. Distribution du seston dans la région d'affleurement du nord-ouest de l'Afrique en mars 1973. *Téthys*, 6(1-2): 77-88.
7. OKUBO, A.: 1962. A Review of Theoretical Models for Turbulent Diffusion in the Sea. *Journal of the Oceanographical Society of Japan*. Vol. del 20.è Aniversari, pp. 286-320.
8. TAYLOR, L. R.: 1961. Aggregation, variance and the mean. *Nature*, 198: 732-735.

# EXPRESSIÓ SENZILLA DE LA REGULACIÓ EN SISTEMES DINÀMICS

per RAMON MARGALEF

Departament d'Ecologia. Facultat de Ciències.  
Universitat de Barcelona

En tractar d'explicar els processos de regulació en sistemes ecològics he trobat sempre molt efectiva una representació gràfica, amb les existències, material emmagatzemat o biomassa en abscesses, i les entrades i sortides en ordenades. Entrades i sortides estan representades per ratlles qual trajectòria mostra si són o no funció del material existent o de la biomassa. Les entrades i sortides es creuen en el punt estable d'acumulació de biomassa o de material en general.

El model s'introdueix amb diverses possibilitats senzilles, no restringides a sistemes vivents. Regulació de la temperatura de la terra (model B de la figura), regulació de la quantitat de fullaraca en un bosc (model A), o bé, de consuetud, a poblacions, que és l'aplicació que m'interessa desenvolupar.

És fàcil veure que una situació estacionària es dona quan les sortides són funcions de la biomassa o població existent d'un grau més elevat que no pas les entrades. Altrament la regulació no es possible i no ho és si natalitat i mortalitat es fan funció linear simple de la densitat de població. El mateix passa amb la migració; si immigració i emigració són funcions semblants de la població no hi ha regulació. Però és fàcil que l'emigració és més funció de la població que la immigració, i que les morts són més funció de la població que els naixements. A la llarga, aquest és el principi bàsic de regulació en sistemes naturals.

Aquesta discussió facilita el comprendre que les potències en què entra la població o la biomassa en els models matemàtics estil Volterra i Lotka, no tenen per què ser 0, 1 o 2, i és molt més versemblant que siguin fraccionaris. De fet, amb dades experimentals, la regressió de log naixements, o log morts, contra log població, mostra pendents diverses. Els valors diversos de la pendent de la recta de regressió són, en part, «vers», i, en

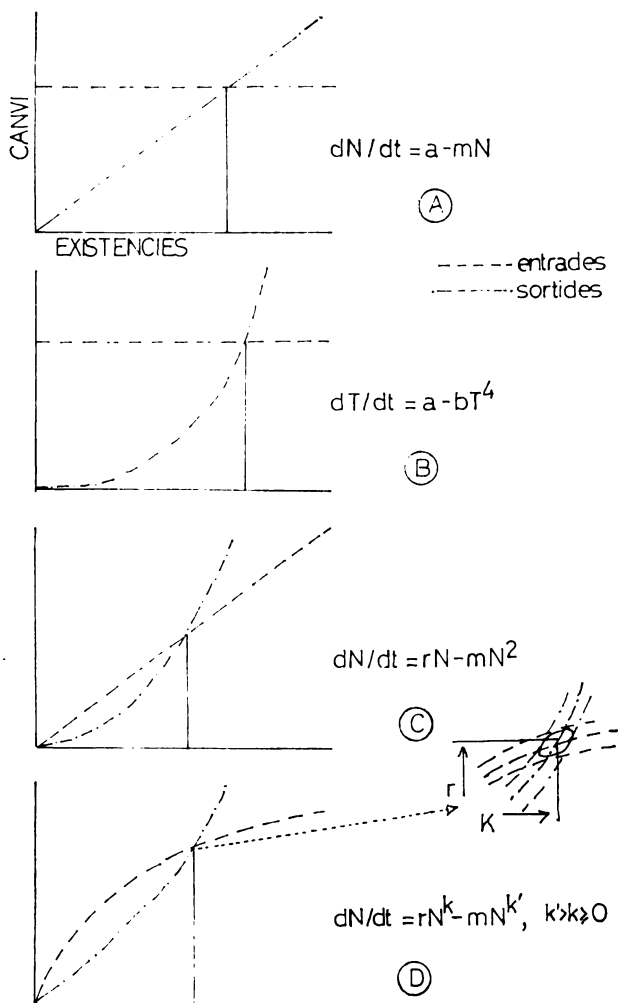


FIG. 1. — Quatre exemples de regulació d'existències, biomassa o població (en abscisses). La quantitat estacionària és donada en el punt on les entrades i les sortides s'equilibren (en ordenades, entrades a ratlletes; sortides, ratlla i punt). Representació gràfica i expressions matemàtiques d'ús corrent en ecologia. *A* correspondria a l'acumulació de fulleraça en un bosc, amb entrada constant i descomposició proporcional a la fulleraça existent. *B* representa la regulació de la temperatura a la Terra, amb entrada constant de radiació i emissió proporcional a la quarta potència de la temperatura. És un model aproximat que ignora la heterogeneïtat de la superfície de la Terra. *C* i *D* representen poblacions biològiques, i en *D* es suggereix l'interès d'una visió més exacta del punt d'encreuament i de la seva variabilitat, en relació amb les estratègies evolutives

part, una conseqüència estadística d'un menor grau de correlació en les entrades que en les sortides, el que és cert i té el mateix efecte. És a dir, l'essència de la regulació és una major predictabilitat en les sortides que en les entrades. Aquestes nocions no solament s'imposen a l'estudiant, sinó també al professor, que sovint ha llegit nocions molt diverses en llibres d'ús corrent i interminables polèmiques que no tenen massa sentit.

És possible donar un pas més vers l'abstracció, si es té present que, en la població d'un animal presa per exemple, les entrades (naixements) són

funció de la població de la mateixa presa, però les sortides són funció de la població de la presa ensem que de la població del depredador. Però la població del depredador és, al seu torn, funció de la població de la presa, de manera que podem escriure

$$dN/dt = \text{funció}(N) - \text{funció}(\text{funció}(N)), \text{ o } dN/dt = f(N) - f(f(N))$$

Expressions que potser són massa abstractes per ésser d'ús immediat, però introdueixen la noció de recurrència i el fet que les sortides tendeixen a dependre de l'estructura de la funció, més que dels valors de les variables implicades.

S'obre una altra perspectiva molt ampla en l'ús possible de funcions periòdiques, o de sumes de funcions periòdiques tant en les entrades com en les sortides, visualitzant possibles fluctuacions de la població com resultat de l'interferència d'aquells.

És possible portar a terme una anàlisi més detallada de la regió d'intersecció de les dues línies, sigui considerant les diverses derivades en aquesta regió, sigui la incertesa o la variabilitat de cada una de les corbes. D'una o altra manera es delimiten regions on «poden passar coses». Vull dir en termes d'evolució i d'estratègia de la evolució. En aquest sentit el gràfic que comento pot servir també per discutir un tema d'actualitat, el de les diverses estratègies evolutives al llarg de l'eix  $r-K$ . El quocient entre les ordenades i les abscisses, en cada punt, representa la relativa importància de les dues estratègies, de la  $r$  (entrades i sortides) i de la  $K$  (màxima població mantinguda).

## DISCUSSIÓ

### FLÓS

El problema és fer les equacions i seguir-les numèricament pas a pas. També caldria introduir el temps; l'estat del sistema pot canviar i donar lloc a respostes diferents.

### MARGALEF

Quan varia l'activitat metabòlica, per exemple, les línies que representen les entrades i les sortides es mouen d'alguna forma. Es poden estudiar amb detall les lleis de canvi. Els càlculs numèrics poden complicar-se molt, matemàticament.





# ASPECTES DE LA PROBLEMÀTICA DE LA DISTÀNCIA TAXONÒMICA

per JORDI OCAÑA i REBULL

Unitat Docent de Bioestadística. Facultat de Biologia.  
Universitat de Barcelona

## INTRODUCCIÓ

A la figura 2 hi ha un esquema, necessàriament simplificat que tracta de reflectir les relacions existents entre els mètodes matemàtics utilitzables per un investigador que vulgui estudiar les semblances (o diferències) entre unitats taxonòmiques orgàniques (OTUs) i establir classificacions. La problemàtica de la distància taxonòmica es pot situar en els esglaons (a) i (b) + (b').

Generalment es parteix d'una taula de dades (figura 1) on cada unitat taxonòmica queda caracteritzada com un element d'un espai de  $n$  dimen-

		$x_1$	$x_2$	...	$x_n$
O T U S	$o_1$	$x_{11}$	$x_{12}$	...	$x_{1n}$
	$o_2$	$x_{21}$	$x_{22}$	...	$x_{2n}$
	$o_m$	$x_{m1}$	$x_{m2}$	...	$x_{mn}$

FIG. 1. — Exemple de taula de dades

sions, de coordenades ( $x_{11}, \dots, x_{1n}$ ), i cada caràcter per un element d'un espai  $m$ -dimensional, amb coordenades ( $x_{11}, \dots, x_{m1}$ ). Però amb la matriu de dades no n'hi ha prou, cal considerar també quina és l'estructura d'aquests espais. Això depèn del tipus d'estudi realitzat, podríem dir de

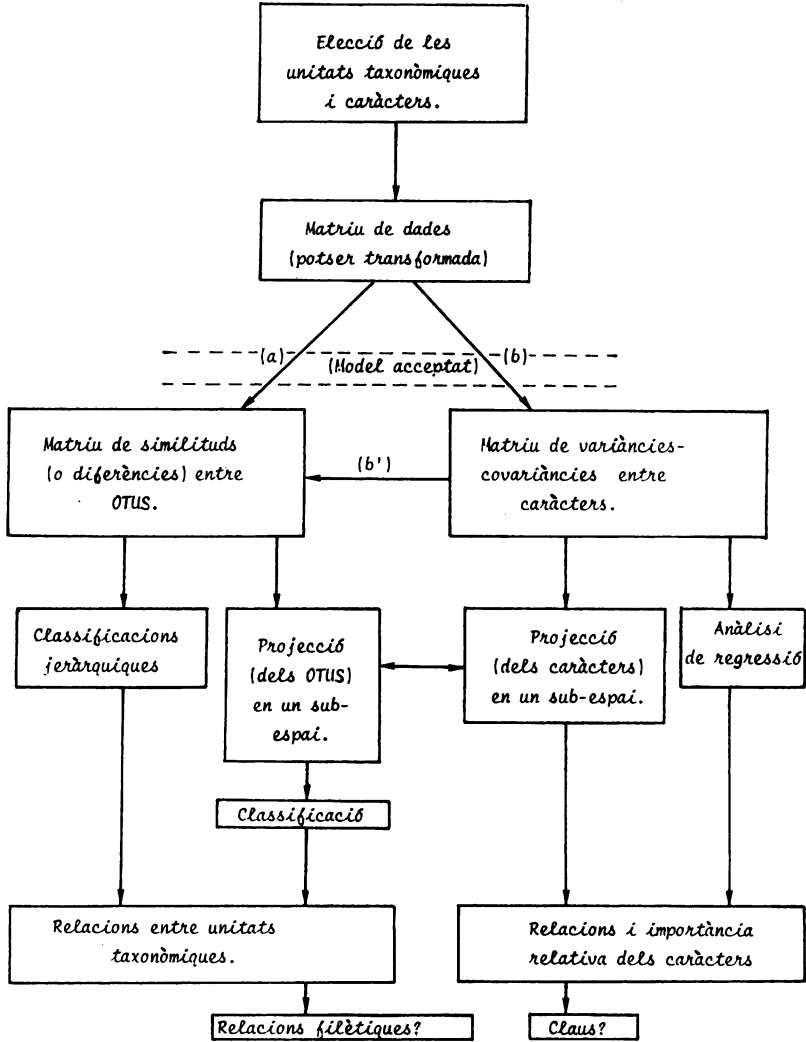


FIG. 2.

Esquema simplificat (basat en BOYCE, 1969) de les relacions existents entre els mètodes matemàtics utilitzables per a establir classificacions d'unitats taxonòmiques orgàniques (OTUs)

quin és el tipus de model inicial considerat més adient (si es poden considerar els caràcters independents, si es tracta de variables aleatòries amb distribució normal, etc.) i de la mateixa natura dels caràcters emprats (normalment es poden considerar variables aleatòries, però el seu «nivell de mesura» pot ésser ben diferent: mesures contínues, comptatges, indicadors de presència-absència, etc...).

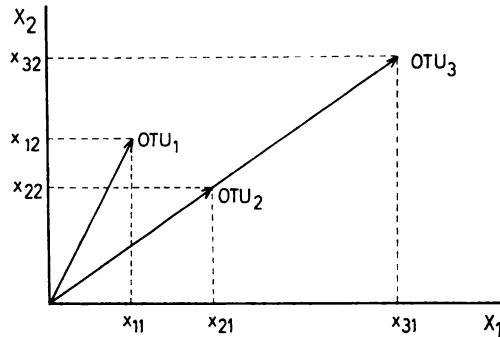


FIG. 3. — Esquema de la posició de tres OTUs en un espai de dues dimensions

Quan s'ha fixat la natura de l'espai de representació es pot intentar definir o escollir amb una certa precisió la distància taxonòmica més adequada.

La següent classificació dels índexs de semblança pot servir com un punt de referència en la discussió que a continuació detallem:

- Coeficients angulars
- Coeficients de distància
  - \*Al marge de l'estructura de l'espai
    - °No mètriques
    - °Mètriques
  - \*Índexs que tenen en compte (o ho proven) l'estructura de l'espai (normalment tenen propietats de mètrica).

#### COEFICIENTS ANGULARS

Un exemple característic n'és la distància de CAVALLI-SFORZA i EDWARDS (1967) i EDWARDS (1971). També ho són les expressions semblants als coeficients de correlació, quan s'utilitzen per relacionar OTUs. Els

darrers corresponen al cosinus de l'angle format pels dos OTUs, un cop representats en un cert espai. El primer correspon a la corda de l'arc associat a les dues unitats taxonòmiques representades en un espai transformat.

Aquests índexs reflexen les diferències degudes a l'orientació dels OTUs, no les degudes a la seva dimensió. Com es pot apreciar a la figura 3 la distància angular (angle, cosinus) entre OTU<sub>1</sub> i OTU<sub>2</sub> és la mateixa que entre OTU<sub>1</sub> i OTU<sub>3</sub>. Hom podria considerar que la «punta» d'OTU<sub>1</sub> està més allunyada d'OTU<sub>3</sub> que d'OTU<sub>2</sub>, ja que OTU<sub>3</sub> és més «llarg» i per tant esperar que la distància d'OTU<sub>1</sub> a OTU<sub>3</sub> fos més gran.

#### COEFICIENTS DE DISTÀNCIA

Es podria considerar com a tals aquells que pretenen expressar les dues components (angular, dimensional), junt amb una llarga llista d'índexs d'interpretació geomètrica difícil o impossible, ja que normalment no es tracta de mètriques, malgrat que puguin ésser útils en algun cas determinat. Els darrers inclourien índexs basats en la probabilitat de trobar valors iguals d'un caràcter, en observacions independents, a cada OTU: SNEATH (SOKAL i SNEATH, 1963), HEDRICK (1971); índexs basats en la Teoria de la Informació: KULLBACK (1968), ORLOCI (1969); índexs d'associació: JACCARD (SNEATH i SOKAL, 1973), GOWER (1971); distància de NEI (1970-71).

Dins de la classe dels índexs amb propietats de mètrica cal distingir clarament entre aquells que no tenen en compte l'estructura de l'espai (malgrat que la seva definició ja pressuposa una interpretació geomètrica dels OTUs, es suposa que aquests són punts d'un espai euclidi, expressats segons una base de vectors ortonormals, com si consideréssim els caràcters incorrelacionats i cadascú amb el mateix pes) i aquells que tracten de posar-la de manifest.

Les mètriques de Minkovski, de fórmula general

$$D_r = \left( \sum_{j=1}^n |x_{1j} - x_{2j}|^r \right)^{1/r}$$

donen lloc, quan  $r = 1$ , a la «distància segons les illes de cases a una ciutat» (per anar de l'OTU<sub>1</sub> a l'OTU<sub>2</sub>, es va «trencant» de dimensió en dimensió, figures 4a i 4b)

$$D_1 = \sum_{j=1}^n \left| x_{1j} - x_{2j} \right|$$

i quan  $r = 2$ , a la distància euclídia usual (figura 4a)

$$D_2 = \left( \sum_{j=1}^n (x_{1j} - x_{2j})^2 \right)^{1/2}$$

La distància  $D_1$ , dividida per el nombre de caràcters, correspon a la distància M.C.D. (*mean character difference*, CAIN i HARRISON, 1957). Una expressió semblant va ésser proposada per PREVOSTI (1974), com a distància genètica entre poblacions, caracteritzades per les freqüències dels al·lels de  $n$  caràcters ( $s_j$  al·lels pel caràcter  $j$ ,  $j = 1, \dots, n$ )

$$A = 1/2n \left( \sum_{j=1}^n \sum_{K=1}^{s_j} \left| p_{1jk} - p_{2jk} \right| \right)$$

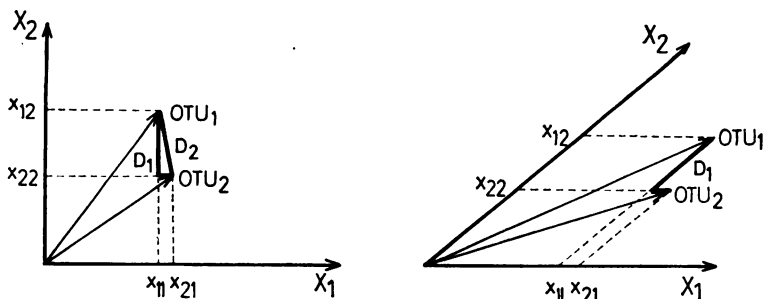
Com mostren les figures 4a i 4b, aquestes distàncies són independents de l'angle (correlacions) entre els vectors (variables aleatòries) que formen el sistema de referència, la qual cosa no passa amb  $D_2$  o amb  $1/nD_2$  (M.S.D. *mean square difference*) o amb la distància genètica de ROGERS (1972), que només tindria sentit en un sistema de referència com el de la figura 4a.

Certs índexs, com ara el coeficient de Pearson C.R.L. (*coefficient of racial likeness*) i la distància ji-quadrat de BENZECRI (1973), malgrat que no tinguin en compte les relacions entre caràcters, sí que ponderen d'alguna manera el «pes» de cada un d'ells. La distància ji-quadrat, utilitzada per comparar distribucions finites, s'expressa com

$$X^2 = \sum_{j=1}^n 1/p_{.j} (p_{1j} - p_{2j})^2$$

essent  $p_{.j}$  la probabilitat de que es presenti l'estat  $j$  (per exemple, l'al·lel  $j$ ) a la reunió de totes les poblacions estudiades. Aquesta distància té la propietat, realment interessant (BENZECRI, 1973), d'ésser pràcticament equivalent a la informació mútua (informació processada per un canal) entre les distribucions associades a les unitats taxonòmiques 1 i 2.

És interessant que les distàncies tinguin en compte les relacions entre caràcters, ja que si aquests són dependents, la informació continguda en un d'ells estarà en part ja expressada en els altres. Quan les unitats taxonòmiques han estat caracteritzades mitjançant variables aleatòries amb



FIGS. 4a-4b. — Esquema de la posició de dos OTUs en un espai de dos dimensions. a: eixos perpendiculars. b: eixos correlacionats

distribució conjunta normal multivariant, i quan la matriu de variàncies covariàncies és la mateixa a tots els OTUs, es pot utilitzar la distància de Mahalanobis.

$$\Delta^2 = (m_1 - m_2)' \Sigma^{-1} (m_1 - m_2)$$

essent

$$m_1 = (m_{11}, \dots, m_{1n})$$

$$m_2 = (m_{21}, \dots, m_{2n})$$

els vectors de mitjanes dels  $n$  caràcters, a les poblacions 1 i 2, i la matriu de variàncies covariàncies comunes.

Aquesta distància té certes propietats que la fan òptima. Això es pot deduir de raonaments basats en la raó de versemblança (MAHALANOBIS, 1936) o en raonaments basats en conceptes d'àlgebra lineal (per exemple, DEMPSTER, 1969). Es pot demostrar que té cura de les correlacions entre caràcters. Com en el cas de distribució multinormal, incorrelació implica independència estocàstica, el problema queda ben resolt.

Però quan les variables estan lligades per relacions no lineals (sovint les dades es poden considerar situades en hipersuperfícies, esfèriques per exemple) ja no té sentit considerar una distribució multinormal. Evidentment, tampoc en té quan es tracta de variables no contínues. En aquests casos s'ha fet servir la distància de Mahalanobis o distàncies formalment

anàlogues, adaptades a d'altres tipus de variables aleatòries (BALAKRISHNAN i SANGHVI, 1968; KURKZÒNSKI, 1970; OCAÑA, ALONSO i PREVOSTI, 1976). La validesa d'aquestes distàncies és dubtosa, car només tenen en compte les relacions lineals entre caràcters, això vol dir que la seva eficiència serà més gran quan el coeficient de correlació sigui més proper a la raó de correlació. En realitat també és dubtós que reflecteixin les relacions lineals ja que com la mitjana i la variància no es poden considerar independents, l'estimació de la matriu de variàncies-covariàncies comuns, és força problemàtica.

Probablement caldria definir una distància basada en una teoria matemàtica més general que l'àlgebra lineal, potser basada en la raó de correlació. Es tractaria de trobar una transformació que permetés passar d'un conjunt de variables dependents a un conjunt de variables independents. Potser la Teoria de la Informació tingui quelcom a dir al respecte.

## BIBLIOGRAFIA

1. BALAKRISHNAN, V.; SANGHVI, L. D.: (1968). Distances between populations on the basis of attribute data. *Biometrics*, 24: 859-866.
2. BENZECRI, J. P.: (1973). L'analyse des données. I. La Taxinomie. Ed. Dunod.
3. BOYCE, A. J.: (1969). Mapping Diversity: a comparative study of some numerical methods. *Proc. Coll. Num. Tax. Univ. St. Andrews*, September 1968), (1-31). A. J. Cole Ed. Academic Press.
4. CAIN, A. J.; HARRISON, G. A.: (1958). An analysis of the taxonomist's judgement of affinity. *Proc. Zool. Soc. Lond.*, 131: 85-98.
5. CAVALLI-SFORZA, L. L.; EDWARDS, A. W. F.: (1967). Phylogenetic analysis: models and estimation procedures. *Evolution*, 21: 550-570.
6. DEMPSTER, A. P.: (1969). Elements of continuous Multivariate Analysis. Addison-Wesley.
7. EDWARDS, A. W. R.: (1971). Distances between populations on the basis of gene frequencies. *Biometrics*, 27: 873-881.
8. GOWER, J. C.: (1971). A general coefficient of similarity and some of its properties. *Biometrics*, 27: 857-871.
9. HEDRICK, P. W.: (1971). A new approach to measuring genetic similarity. *Evolution*, 25: 276-280.
10. JACCARD, A.: (1973). Distances généalogiques et distances génétiques. *Cahiers d'Antropologie et d'Ecologie Humaine*, 1: 11-124.
11. KULLBACK, S.: (1968). Information Theory and Statistics. J. Wiley.
12. KURKZYNSKY, T. W.: (1970). Generalized distance and discrete variables *Biometrics*, 26: 525-534.
13. MAHALANOBIS, P. C.: (1936). On the generalized distance in Statistics. *Proc. Nat. Inst. Sci. India*, 2: 49-55.
14. NEI, M.: (1971). Interspecific gene differences and evolutionary time estimated from electrophoretic data on protein identity. *Amer. Nat.*, 105: 385-398.
15. NEI, M.: (1972). Genetic distance between populations, *Amer. Nat.*, 106: 283-292.
16. OCAÑA, J.; ALONSO, G. i PREVOSTI, A.: (1977). La estructura matemàtica de las poblaciones y la definición de distancias entre ellas. 2.º S.I.N.A.P.E. Universidade Estadual de Campinas. Brasil.

17. ORLOCI, L.: (1969). Information Theory models for hierarchic and non hierarchic classifications. *Proc. Coll. Num. Tax. Univ. St. Andrews*. September 1968 (148-164). Academic press.
18. PREVOSTI, A.: (1974). La distancia genética entre poblaciones. *Miscelanea Alcobé*: 109-118.
19. SNEATH, P. H. A.; SOKAL, R. R.: (1973). *Numerical Taxonomy*. W. H. Freeman.
20. SOKAL, R. R.; SNEATH, P. H. A.: (1963). *Principles of Numerical Taxonomy*. W. H. Freeman.

## DISCUSSIÓ

### CUADRAS

M. G. KENDALL introduí una geometria estocàstica que considerava distàncies geodèsiques. Utilitza mètodes molt complicats de geometria diferencial i exigeix recursos molt difícils d'anàlisi matemàtica. El problema de certes distàncies no euclídiess és que es relacionen amb matrius de «covariàncies» que tenen valors propis negatius. Les coordenades euclídiess per a representar poblacions, tenen, aleshores, unes coordenades reals i unes altres imaginàries, és a dir, la «distància» és negativa per alguns determinats eixos. Existeix, altrament un mètode de formació molt diferent, però amb més aplicació que la geometria estocàstica, que permet trobar una representació euclídia raonable de distàncies no euclídiess. És l'anàlisi de proximitats de Torgerson, Shepard, Kruskal, Caroll i altres.

### ALONSO

En aquest camp, caldria més comunicació entre biòlegs i matemàtics.



## ANÀLISI DE PERFILS

per M. DE RENZI <sup>1</sup>, J. SERRA KIEL <sup>2</sup> i C. CUADRAS <sup>3</sup>

1) Departament de Paleontologia de la Universitat de Barcelona

2) Departament d'Estratigrafia de la Universitat de Barcelona

3) Seminari Matemàtic de la Universitat de Barcelona

### INTRODUCCIÓ

L'anàlisi de perfils és un procediment molt interessant quan es tracta de comparar els tipus de gràfics dels quals parlarem a continuació. Moltes vegades s'ha de decidir si dos o més gràfics se'ls pot considerar coincidents o no i sovint aquesta decisió es pren arbitràriament; aleshores hom diu que hi ha equivalència o coincidència si ambdós gràfics presenten lleus diferències i, en canvi, no hi haurà coincidència si les diferències són fortes. Ara bé, el límit entre fort i lleu és totalment arbitrari i la qüestió de si una diferència forta o una diferència lleu es poden prendre com a significatives, es deixa totalment de banda.

Els naturalistes empren freqüentment gràfics i els comparen, i de la comparació surten conseqüències —d'ordre biològic o geològic— que poden ser incorrectes si les decisions es prenen arbitràriament. Només els procediments estadístics ens poden guiar a una elecció correcta quan se'ns presenten problemes de la naturalesa que acabem d'esmentar.

Els tipus de gràfics a comparar, mitjançant l'anàlisi de perfils, és el següent: tenim  $h$  poblacions, i a cada una definim una variable aleatòria  $X_i$  ( $p \times 1$ ) multinormal  $N(\mu_i, \Sigma)$  amb  $i = 1, 2, \dots, h$ . De cadascuna obtenim una mostra d'extensió  $N_i$  i estimem la  $\mu_i$  corresponent mitjançant el vector de mitjanes aritmètiques  $X_i$ . Si aleshores posem sobre uns eixos de coordenades les components de  $X_i$  com a abscisses igualment espaiades, i les mitges aritmètiques com a ordenades, a cada component li correspondrà una mitjana i obtindrem un punt; si unim tots aquests punts amb una línia trencada tindrem el perfil de les mitjanes de la població  $i$ -ésima. Això ens ho mostra la figura 1 i és ara quan estem en condicions de comparar els perfils corresponents de les  $h$  poblacions.

D'acord amb MORRISON (1967), el primer problema serà el de saber si els perfils són paral·lels o no. Si no ho són, podem concloure que

$$\vec{\mu}_1 \neq \vec{\mu}_2 \neq \dots \neq \vec{\mu}_h$$

Si hi ha paral·lisme, 2) pot ser que els perfils coincideixin o no. Donat el paral·lisme, provar que la suma d'esperances desconegudes de les components del perfil de cada població és igual per a totes les poblacions. Si coincideixen, la conseqüència és

$$\vec{\mu}_1 = \vec{\mu}_2 = \dots = \vec{\mu}_h$$

i la no-coincidència s'interpretaria com una transformació proporcional d'un vector en un altre

$$\vec{\mu}_i = k \vec{\mu}_j$$

Per últim, 3) si hi ha paral·lisme podem plantejar encara la hipòtesi d'igualtat de components dins de cada perfil.

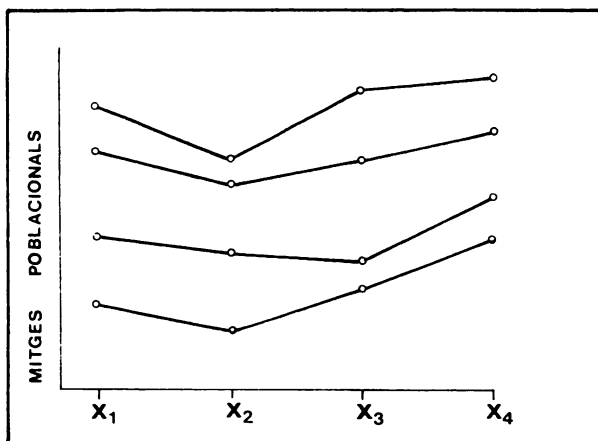


FIG. 1. — Exemple de perfils corresponents a quatre poblacions

Com es pot veure, es requereix provar primer la igualtat de matrius de covariància a partir de les estimacions mostrals. Les tres proves d'hipòtesis referides a l'anàlisi de perfils pròpiament dita, són casos particulars d'anàlisi multivariable de la variança i es troben descrits a MORRISON (1967) o a RAO (1973). Per a la prova d'aquestes hipòtesis, CUADRAS ha

fet ús dels programes CANG i CANP (CANON modificat) per ell desenvolupats (CUADRAS, CAMPA i MONTORIOL, 1972).

Comparar dos perfils, en definitiva, és visualitzar dos o més punts (cada punt un perfil) d'un espai p-dimensional. Aleshores, l'anàlisi de perfils pot tractar-se pels procediments de l'anàlisi canònica; en el nostre cas

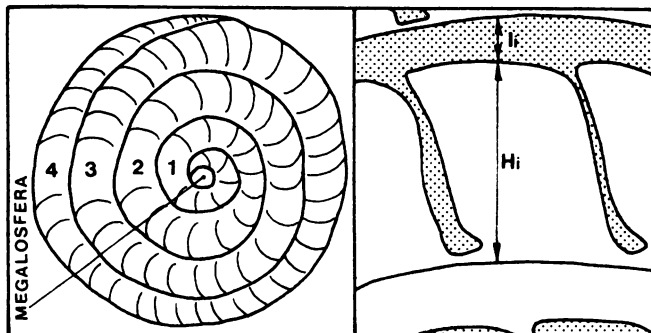


FIG. 2. — Esquema de les mesures preses en els *Nummulites*.  
FIG. 3. — Perfils de quatre poblacions de *Nummulites* de la Plana de Vic

concret, les tres hipòtesis a provar poden tenir una representació canònica adequada. En particular, la prova 2) de *coincidència o superposició* equivaldrà a una anàlisi canònica de poblacions amb la representació corresponent. Les altres proves 1) i 3) requereixen la formació de noves variables aleatòries, combinacions lineals de les primitives, les quals seran també multinormals i susceptibles de noves representacions canòniques (CUADRAS, 1973).

#### APLICACIÓ DE L'ANÀLISI DE PERFILS A L'ESTUDI SISTEMÀTIC DELS *Nummulites* DE LA PLANA DE VIC

Un mètode clàssic introduït per SCHAUB per a la sistemàtica i per a la filogènia dels *Nummulites* és l'estudi quantitatiu de l'ontogènia dels caràcters interns. Els *Nummulites* són animals que posseeixen una conquilla que creix sota un patró espiral i hom pren convencionalment cada volta com un estat de creixement. SCHAUB fa ús dels *Spirendiagram*, però REGUANT i CLAVELL (*in litt.*) fan una modificació i utilitzen una mesura d'alçada de cambra ( $H_i$ ) a cada volta d'espira, junt amb una mesura del corresponent gruix de la corda marginal ( $e_i$ ), tal com representem a la figura 2.

Aquests autors fan correspondre a cada volta  $H_1$  i  $H_1 + e_1$ , i cada població queda representada per un doble perfil de les mitjanes d'ambdues magnituds per volta d'espira. Dos de nosaltres (DE RENZI i SERRA) hem introduït una petita modificació en aquest sistema: En abscisses posem, consecutivament,  $H_1$ ,  $e_1$ ,  $H_2$ ,  $e_2$ , etc. Així obtenim gràfiques de perfils com la representada a la figura 3 i objecte de la nostra aplicació. A la figura 3 comparem, per anàlisi de perfils, quatre poblacions que es superposen en el temps amb evidència estratigràfica; les quatre procedeixen de la Plana de Vic i han estat atribuïdes a quatre unitats taxonòmiques diferents (REGUANT i CLAVELL, 1967) i són

(YSA)	<i>Nummulites</i>	<i>crusafonti</i>	REGUANT et CLAVELL
(YNP)	—	<i>puigsecensis</i>	REGUANT et CLAVELL
(Y 64-5)	—	<i>biedai</i>	SCHAUB
(Y 64-18)	—	<i>biedai complanatus</i>	REGUANT et CLAVELL

Es tractava de comparar els perfils corresponents a les cinc primeres voltes d'espira, el que comporta l'ús de 10 variables, com hom pot veure a la figura 3. Les mitjanes vénen representades pels vèrtex de les trenques dels perfils. La matriu de co-variances comunes estimada —prèvia prova d'homogeneïtat— és

	$H_1$	$e_1$	$H_2$	$e_2$	$H_3$	$e_3$	$H_4$	$e_4$	$H_5$	$e_5$
10 <sup>-4</sup> x	41.231	0.34191	8.0414	-0.50389	-0.63302	-1.5269	0.40190	2.6257	8.2602	-0.26300
		4.5629	-2.0423	1.6095	-0.67240	-0.096762	0.70494	0.054557	0.13839	0.87900
			31.948	0.57943	4.2030	2.4014	1.0239	2.1470	0.54872	1.6820
				6.0012	0.0058927	1.3870	-0.45853	0.70070	-1.0765	0.96383
					27.345	2.0465	3.5788	0.30740	2.1386	-1.4159
						8.8896	0.46174	1.3789	-0.24141	1.4823
							29.886	-1.3773	4.3075	-3.8836
								6.2935	0.68978	2.0128
									41.276	-4.6820
										12.720

La hipòtesi de paral·lisme s'acompleix per a les dues primeres poblacions (YSA i YNP), per separat de les altres dues (Y 64-5 i Y 64-18), però ambdós grups no són paral·lels entre ells. D'aquesta manera separem, per caràcters interns, dos grups de poblacions.

La prova de super-posició dels grups parcials no té tanta potència com quan tots els grups són paral·lels; aleshores, un de nosaltres (CUADRAS) ha fet una prova simultània i s'ha pogut veure que a les dues primeres poblacions (YSA i YNP) era més segur diferenciar llurs perfils paral·lels que a les dues darreres (Y 64-5 i Y 64-18). Per tant, provisionalment creiem que les dues primeres són diferenciables per l'ontogènia dels caràcters interns (mancaria analitzar mostres més extenses) mentre que

les altres dues no ho són. No obstant això, les quatre poblacions serien diferenciables estadísticament per altres caràcters, la qual cosa no exposem aquí perquè no ve al cas.

De tot això en podem treure algunes conclusions provisionals: la diferenciació estadística de tres de les quatre entitats taxonòmiques (YSA, YNP i Y 64-5) pels seus perfils, però no de les dues darreres (Y 64-5 i Y 64-18).

També hom podria pensar que mentre les poblacions no estan gaire separades en el temps, les magnituds internes es modifiquen totes proporcionalment, de manera que els perfils resulten separats i també paral·lels (YSA junt amb YNP), però que si estan massa properes en el temps (el cas de Y 64-5 i Y 64-18) primer es produeix un augment total de la grandària externa i de la megalosfera, però el perfil roman immòdicat.

Si les poblacions estan molt separades en el temps, el cas concret estudiat sembla mostrar una modificació dels perfils, que perden paral·lisme, així com també sembla haver-hi una tendència evolutiva a l'augment de les magnituds utilitzades per a la construcció del perfil.

#### BIBLIOGRAFIA

1. CUADRAS, C.: (1973). Análisis estadístico multivariante y representación canónica de funciones estimables. *Tesi doctoral*, 130 pp. Barcelona.
2. CUADRAS, C.; CAMPA, J. A. i MONTORIOL, J.: (1972). El programa CANON para IBM-360. *Acta Geol. Hispánica*, VII(1): 26-28.
3. MORRISON, D. F.: (1967). Multivariate statistical methods. *McGraw Hill Book Co. (Series in Probability and Statistics)*, XIII + 388 pp.
4. RAO, C. R.: (1973). Linear statistical inference and its applications. *John Wiley and Sons*, 2nd edition. 522 pp. New York.
5. REGUANT, S. i CLAVELL, E.: (1967). Descripción de algunos *Nummulites* afines al *N. perforatus* del Eoceno de Vic (Barcelona, *Notas y Comens. Inst. Geol. y Min. de España*, n.º 101-102, pp. 41-56, 4 lám. i 6 fig.
6. REGUANT, S. i CLAVELL, E. (*in litt.*): Contribución al conocimiento del valor estratigráfico de las *Assilina* (aplicación al Eoceno de San Vicente de la Barquera, Santander, España).

#### DISCUSSIÓ

J. WAGENSBERG

Es podria comparar la importància relativa del creixement i de la diferenciació.

RENZI

Això es podria aplicar a la ontogènia.

MARGALEF

L'estudi biomètric dels copèpodes s'inicià de manera semblant, per Kozminski, en la dècada dels anys 20. Feia moltes mesures a cada població i comparava els conjunts de mesures, o perfils. Com que les mesures estan correlacionades, en l'augmentar el nombre de variables es reforça la diferència. Aquest mètode en certs casos resultava en atribuir diferències significants a generacions successives. És millor acceptar d'entrada la correlació entre diferents mesures dintre el marc de l'alometria i acceptar provisionalment com significatius solament els canvis en coeficients d'alometria o en conjunts de coeficients d'alometria. Aquest és el problema amb què s'ha encarat Tecla Riera. Penso que seria interessant aplicar aquesta experiència al cas presentat pels foraminífers. En particular, el paral·lisme dels perfils podria ser millorat usant una transformació adient que tingués compte de l'allometria (per exemple, la transformació logarítmica de les dades inicials). Pregunto si es té idea de la velocitat de creixement, per si els canvis estacionals de temperatura o d'altres factors es poguessin reflectir en variacions, aparentment irregulars, de l'amplada de les espines. També si aquests foraminífers són perforats i si s'ha pensat d'utilitzar la densitat de perforacions com estima independent de factors ambientals, tal com s'ha fet en espècies actuals.

RENZI

En el cas que hem estudiat, les  $H$  i les  $e$  no estan correlacionades i les  $H$  de les diferents voltes tenen correlacions molt baixes.

MARGALEF

¿Quant temps triguen a fer una volta els equivalents actuals dels *Nummulites*?

RENZI

No se sap gaire.

FLÓS

Es podrien aplicar a cada paràmetre els mètodes d'anàlisi de sèries temporals i comparar els espectres.

**CUADRAS**

En aquest cas, la variació depèn del temps, però a l'anàlisi considerem els perfils de cada població que es comparen sense tenir en compte el temps.

**MARGALEF**

Poden comptar-se els foradets de la conquilla per unitat de superfície?

**RENZI**

És difícil; moltes vegades estan obliterats.

**MARGALEF**

Valdria la pena d'estudiar-ho.





# DETERMINISTIC MODEL FOR THE EVOLUTION OF A POPULATION IN A LIMITED ENVIRONMENT

per J. WAGENSBERG

Departament de Termologia. Facultat de Física.  
Universitat de Barcelona

## ABSTRACT

*Systems which are closed to matter transfer are specially interesting in population studies. A deterministic model for the time evolution of growth and extinction is outlined in this note and is tested with bacterial experimental data.*

## RESUM

*Els sistemes que són pròxims a la transferència de matèria són especialment interessants en els estudis de poblacions. Es formula un model determinístic per a l'evolució en el temps de creixement i d'extinció, que es demostra amb dades experimentals de bacteris.*

### *Deterministic model*

Let us represent the population function by  $n(t)$ ; it is the number of viable individuals of the specie under consideration. We represent its surroundings —the nutrients— by the magnitude  $M(t)$ . The working hypothesis will be that the evolution of  $n(t)$  results only from the interaction between each individual and its environment. Interactions between individuals will not be taken into account; we start in a certain sense from the idea of a «ideal gas population».

The first law for population growth was given by T. R. MALTHUS (1798) under conditions of unlimited sources:

$$\frac{\ddot{n}}{n} = k \tag{1}$$

where  $k$  is the growth rate «constant». According to (1) the population function will grow exponentially,

$$n(t) = n(0) \exp(kt) \quad (2)$$

In a closed system solution (2) is in accord with the observation only after a short period of adaptation and until a certain critical point  $M(t_0) = M_0$  of the evolution has been reached. After this point the ambient environment intervenes to inhibit growth. This situation can be expressed by means of a system of two differential equations:

$$\frac{\dot{n}}{n} = k \quad (3a)$$

$$\dot{k} = f(n) \quad \text{with } f(n) = 0 \quad \text{for } t > t_0 \quad (3b)$$

where (3b) should express the environment-population interaction through the function  $f(n) = g(M(n))$  for  $t > t_0$ . To establish a model means to suggest a concrete form for  $f(n)$ . If we consider the qualitative evolution of

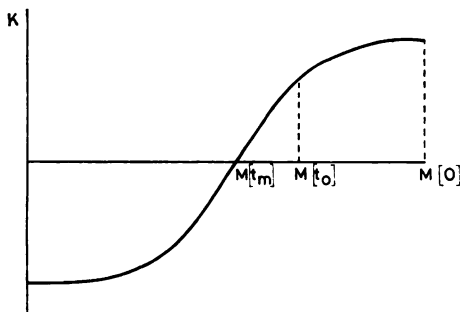


FIG. 1. — Qualitative evolution of the growth rate with the nutrient concentration

Evolució qualitativa de la taxa de creixement amb la concentració de nutrients

the growth rate  $k(t)$  as a function of the nutrients  $M(t)$  as shown in fig. 1, we can develop  $k(t)$  around the inflection point  $M(t_m) = M_m$ :

$$k(t) \sim -a + b M(t) \quad \text{with } a, b > 0 \quad (4)$$

where the point  $t_m$  is the maximum population state ( $n(t_m) = n_m$ ,  $k(t_m) = 0$ ). This approximation which becomes exact in the limit  $t \rightarrow t_m$ , is sufficient for our purposes. On the other hand, the instantaneous value  $M(t)$  of the nutrients (in which toxic effects are conceptually included)

is a direct function of the history of the population until time  $t$ . We will assume that this dependence has the linear form:

$$M(t) = M(t_0) - c \int_{t_0}^t n(t) dt \quad \text{with } c > 0 \quad (5)$$

Equations (4) and (5) allow us to rewrite system (3) in the form:

$$\frac{\dot{n}}{n} = k \quad (6a)$$

$$\dot{k} = -p n = > \quad \text{for } t > t_0 \quad (6b)$$

where the interaction constant  $p = bc$  is a positive number. This is a set of two differential equations of first order; it is equivalent to a second order differential equation for the population function:

$$\frac{\ddot{n}}{n} - \left(\frac{\dot{n}}{n}\right)^2 + p n = 0 \quad (7)$$

from which we derive the population function by direct integration:

$$n(t) = \frac{\mu^2}{2p} \left[ 1 - \left( \frac{1 - \alpha \exp \mu (t - t_0)}{1 + \alpha \exp \mu (t - t_0)} \right)^2 \right] \quad (8)$$

using the following initial conditions at  $t = t_0$ : a)  $n(t_0) = n_0$   
 b)  $k(t_0) = k_0$ , the parametres  $\mu$  and  $\alpha$  are given by:

$$\mu^2 = k_0^2 + 2 p n_0 \quad (9)$$

$$\alpha = \frac{1 - \left[ 1 - \frac{2pn_0}{\mu^2} \right]^{1/2}}{1 + \left[ 1 - \frac{2pn_0}{\mu^2} \right]^{1/2}} \quad (10)$$

Equation (6a) then yields the function  $k(t)$ :

$$k(t) = \mu \left[ \frac{1 - \alpha \exp \mu (t - t_0)}{1 + \alpha \exp \mu (t - t_0)} \right] \quad (11)$$

We can now obtain the dependence of the interaction with the initial conditions of the surroundings and the population. From (5) we have:

$$M(\infty) = 0 \quad ; \quad M(t_0) = c \int_{t_0}^{\infty} n(t) dt \quad (12)$$

where the constant  $c$  is the nutrient that one individual needs in order to live one unit of time. If we write the nutrients in «c» units  $m(t) = \frac{1}{c} M(t)$ , we derive from equations (8), (9), (10) and (12):

$$m_0 = \left[ \left( \frac{k_0}{p} \right)^2 + \frac{k_0}{p} \right]^{1/2} + \frac{k_0}{p} \quad (13)$$

and therefore, for a given environment ( $m_0$ ), the interaction depends on the quantity ( $n_0$ ) and the quality ( $k_0$ ) of the initial biomass in the linear form:

$$p = \frac{2}{m_0} \left[ \frac{n_0}{m_0} + k_0 \right] \quad (14)$$

#### *Experimental test and concluding remarks*

To make an experimental test of the model we need a more accessible form of eq. (8). We use the maximum population state ( $k(t_m) = 0$ ,  $n(t_m) = n_m$ ) in order to suppress  $k_0$  and to adjust the constant  $p$ . We rewrite then:

$$n(t) = n_m \left[ 1 - \left( \frac{1 - \alpha \exp \mu (t - t_0)}{1 + \alpha \exp \mu (t - t_0)} \right)^2 \right] \quad (15)$$

with

$$\alpha = \frac{1 - \left[ 1 - \frac{n_0}{n_m} \right]^{1/2}}{1 + \left[ 1 - \frac{n_0}{n_m} \right]^{1/2}} \quad (16)$$

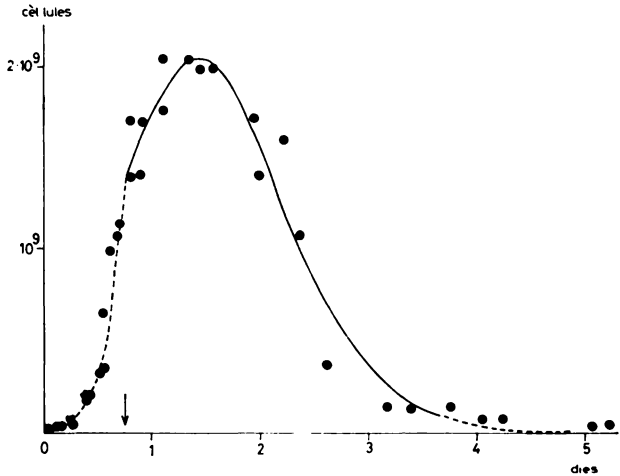
$$\mu = - \frac{\ln \alpha}{t_m - t_0} \quad (17)$$

$$p = \frac{\mu^2}{2 n_m} \quad (17)$$

Figure 2 shows a test of eq. (15) applied to a bacterial culture of *Citrobacter intermedium* C<sub>3</sub> (Clotet *et al.* 1968), inoculated in a hermetically closed vessel of 10 ml with 6.7 ml of liquid medium (Torra *et al.* 1976). The experimental points (dark circles) represent the number of

FIG. 2. — Equation (2) (...) and (15) (—) applied to a bacterial culture of *Citrobacter intermedium*. The arrow indicates the critical point  $t_0$

Equacions (2) (...) i (15) (—) aplicades a un cultiu bacterià de *Citrobacter intermedium*. La sageta indica el punt crític  $t_0$



viable bacteria in the vessel, determined by plate count from a serie of evolutions prepared identically at constant temperature (20 °C) (Torra *et al.* 1976). The evolution follows the Malthus' law after a latency phase of two hours until the critical instant  $t_0 = 18$  h:

$$n(t) = 0.15 \times 10^9 \exp (0.26 t) \text{ cells for } 2 < t < 18 \quad (18)$$

For  $t > 18$  h equation (15) is applied with the following values of the initial and maximum population states:

$$\begin{array}{lll} n_0 = 1.4 \times 10^9 \text{ cells} & \text{at} & t_0 = 18 \text{ h} \\ n_m = 2.1 \times 10^9 \text{ cells} & \text{at} & t_m = 34 \text{ h} \end{array}$$

We conclude that population function in a closed system has a different behavior before and after the critical point  $t_0$ . Before the critical point, the behavior is given by the Malthus' law (2). After the critical point the behavior follows solution (15). There is a discontinuity at  $t_0$  and this is in contrast with the behavior predicted from the so-called S-shaped curve for open systems (models of EIGEN, VERHULST'S, GOMPERTZ) (EIGEN, 1971) (GOEL, 1974, p. 74). This discontinuity has an important biological meaning. A minimum value of the nutrient concentration is needed in order that the population may grow freely following Malthus' law; on the other hand below this minimum nutrient concentration the interaction becomes important and it drives the population to total extinction, after reaching a single characteristic maximum.

#### BIBLIOGRAFIA

1. CLOTET, R.; GUINEA, J. i PARÉS, R.: (1968). «Segregación de aminoácidos por una cepa C. Intermedium». *Microbiología Española* 21, 155.
2. EIGEN, M.: (1971). «Selforganization of Matter and the evolution of Biological Macromolecules». *Die Naturwissenschaften* 58, 10, 465.
3. GOEL, N. S. i RICHTER-DYN, N.: (1974). «Stochastic Models in Biology» Academic Press, N. Y.
4. TORRA, V. i WAGENSBERG, J.: (1976). «Microcalorimetric study of C. Intermedium». *Conference de Thermocinetique A.F.C.A.T.* 2, 71.

#### DISCUSSIÓ

##### MARGALEF

Aquest procés, que depèn d'un impuls inicial i unes constants, podria aplicar-se al consum de petroli en estudis com els del Club de Roma.

##### J. WAGENSBERG

Potser sí, però aquest model està pensat per a experiments de laboratori. El model no inclou el paràmetre, decisió intel·ligent, malgrat que,

també és veritat, en el cas del petroli potser sigui irrelevant considerar-lo.

D. LURIÉ

Fins a quin punt és lineal el model?

J. WAGENSBERG

L'expressió:  $k(t) = -a + b M(t)$ , és certa en el punt d'inflexió; si s'en allunya va deixant de ser vàlida. Si es vol augmentar la validesa del sistema cal prendre més termes en el desenvolupament de  $k(t)$ . L'ideal seria cobrir el sistema del començament al final.

VALLESPINÓS

¿I el cas dels bacteris marins que has estudiat amb el microcalorímetre, on després d'una aparent extinció reapareixen processos metabòlics?

J. WAGENSBERG

Per al model, aixó fóra un miracle. Hauria de canviar qualitativament el sistema. En el cas dels bacteris marins l'he ajustat fins a la primera extinció.

MARGALEF

Després d'un impuls inicial, arriba un moment en que es redueix el canvi d'energia i poden introduir-se noves possibilitats.





# UNA MANERA DISCRETA D'OBSERVAR UN SISTEMA DEPREDADOR-PRESA

per BERNARD P. ZEIGLER i RAMON GUARDANS

Dep. Applied Mathematics. Weizmann Institute of Science Rehovot. Israel  
Departament Ecologia. Universitat de Barcelona.

## INTRODUCCIÓ

I. Considerant la família de models que descriuen la cinètica d'una població derivats de les equacions diferencials de Volterra-Lotka, és evident, i s'ha dit moltes vegades, que un dels problemes que plantegen és el fet de no considerar l'espai.

Quan hom intenta incloure l'espai en aquesta manera de descriure les poblacions, en principi es poden seguir dos camins, que si bé en últim terme poden resultar bastant similars, parteixen de punts diferents.

Podem considerar la descripció de la morfologia espacial d'un sistema (comunitat, població) com una funció contínua («superposada» a un espectre continu d'organització). Si describim apropiadament les característiques de les relacions entre les diferents unitats considerades a *un* determinat «nivell», podrem derivar les descripcions dels nivells «superiors» o «inferiors» (en alguna jerarquia d'escales) expandint, amb mètodes estrictament matemàtics, aquella descripció.

D'altra banda podem considerar l'esforç per imaginar una descripció de manera discontinua, és a dir, fixant uns nivells d'organització, unes escales sobre les quals elaborarem un determinat model.

Aquesta estratègia és la que hem seguit. És difícil tractar de comparar-la respecte a la primera en termes de eficiència o precisió. Més que per referència als resultats, té interès perquè ens permet de formalitzar el procés d'aproximació al problema.

Es tracta, doncs, de considerar el sistema d'una manera global. D'una banda tenim les dades experimentals i una sèrie de formalismes matemàtics, els models. D'altra banda uns recursos tècnics per instrumentar les possibles relacions (en aquest cas l'ordinador). Finalment, organitzant tot

això, hi ha la pregunta, l'interès per descriure o modelar d'alguna manera la morfologia espacial del sistema que considerem.

De cara a exposar les coses de la manera més clara possible, mirarem de resseguir la «història» del model amb el qual estem treballant ara.

II. El problema que considerem en aquest cas és la clàssica observació que un sistema de dos espècies, predador i presa, és en unes determinades condicions inestable, és a dir d'una durada molt restringida en el temps, i en altres condicions, en canvi, por arribar a ser estable, és a dir a tenir una vida virtualment infinita. En el primer cas, el sistema no és capaç d'internalitzar les oscil·lacions, en el segon, sí.

Evidentment, es pot formular la pregunta. ¿Quines són les diferències entre el dos sistemes?

Moltes dades fan pensar que un element important és la morfologia espacial del sistema, en una imatge lineal, podríem dir el seu grau d'organització de l'espai.

Restringint encara molt més el context experimental, per fixar-nos més aviat en l'elaboració del model, considerarem els experiments fets per HUFFAKER (1958-63) amb taronges i dues espècies d'àcars (apèndix).

Com a recursos teòrics de descripció utilitzarem la formalització del procés de modelació proposada per ZEIGLER (1976) i les idees de MAYNARD SMITH (1974) sobre la importància dels processos discrets.

Considerarem, tal com hem dit, diversos nivells d'aproximació amb una certa autonomia formal i teòrica.

Un punt de vista, consisteix a considerar tot l'univers experimental com un conjunt de taques («patches»), i que poden trobar-se en diferents estadis.

Es descriu també, una forma de migració entre les taques.

Finalment, es considera tot l'univers, tot el conjunt de taques entrelligades per la migració, com un sistema tancat.

En el nostre cas, les taques responen a les taronges dels experiments de HUFFAKER.

En considerar el conjunt de taques hem d'imaginar una manera de descriure les relacions en l'espai i el temps. Les relacions en el temps pendran la forma d'un cicle que inclou els diferents estadis possibles per a la taca. Les relacions en l'espai vindran donades per la migració, com veurem més endavant.

Els estadis que considerem per a les taques són els següents:

- I Inocupable, quan la cèl·lula és buida, sense aliment.
- O Ocupable, quan la cèl·lula no és colonitzada, però té aliment.
- P Ocupada per preses però lluny encara de la capacitat de suport de la taca, és a dir, amb aliment.

- P' Ocupada per preses, però després de l'exhauriment de l'aliment.
- P Predadors i preses presents, les preses estan en l'estadi P.
- P' Predadors i preses presents, les preses estan en l'estadi P'.

Aquests estadis es poden relacionar gràficament tal com veiem a la figura 1. De les transicions entre uns i altres en parlarem amb més detall en tractar de les migracions.

Fins aquí tenim el que podríem anomenar model qualitatiu del comportament de les cèl·lules o taques.

Ara ens interessa de poder trobar una manera de descriure el comportament quantitatiu d'aquestes taques.

Per a fer això ens hem valgut de les equacions diferencials de *VOLTERRA* i *LOTKA*.

Si anomenem  $x$  la població de presa sobre una taca,  $y$  la població de predador i  $r$  a la quantitat d'aliment present (en el nostre cas mesurada en fraccions de superfície de taronja), tindrem:

$$\frac{dx}{dt} = (b \text{ pos}(r) - d)x - cxy \quad \dots (1a)$$

$$\frac{dy}{dt} = - (d'y + c'xy) \quad \dots (1b)$$

$$\frac{dr}{dt} = - ux \text{ pos}(r) \quad \dots (1c)$$

on

- $b$  natalitat de la presa
- $\text{pos}(r)$  és una funció que pren el valor 1 s'hi ha aliment, i el valor 0 si no n'hi ha
- $d$  mortalitat de la presa
- $c$  coeficient d'interacció predador-presa
- $d'$  mortalitat del predador
- $c'$  coeficient d'interacció presa-predador.
- $u$  utilització de l'aliment

Utilitzant el llenguatge *CSMP* vàrem simular aquestes equacions a l'ordinador per trobar els paràmetres més ajustats als valors obtinguts per *HUFFAKER* sobre una sola taronja. Vàrem obtenir els següents valors:

- b 0,550 o 0,420 (segons els experiments, ja que va utilitzar dos tipus de taronja)  
 d 0,300  
 c 0,050  
 d' 0,300  
 c' 0,005  
 u  $1,3 \times 10^{-5}$  o  $1,06 \times 10^{-5}$  (corresponents a les dues b)

Pel que fa a les migracions entre les taques o cèl·lules considerem que, les preses *només* migren a partir del moment que s'acaba l'aliment (i que simultàniament assolixen la màxima densitat) i que els predadors comen-

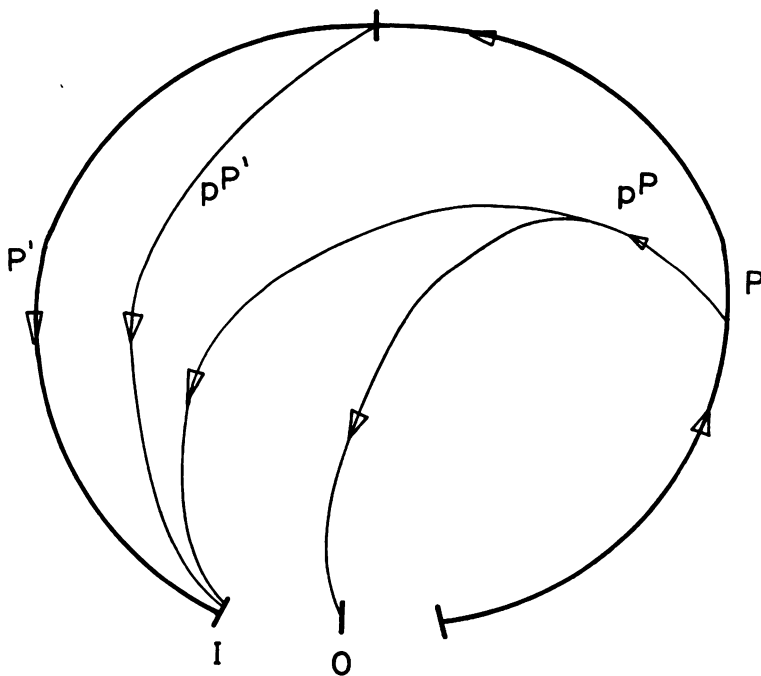


FIG. 1. — Esquema dels estats possibles per a cada taca o cèl·lula

çaran a emigrar en arribar a la seva màxima densitat, és a dir, quan les preses comencin a minvar. La migració és considerada a l'atzar («random walk»), controlada per uns paràmetres que veurem més endavant.

Ara bé, és evident que aquesta descripció de les cèl·lules que resulta quantitativament bastant satisfactòria, si pretenem estendre-la directa-

ment a tot l'univers experimental resulta d'una complexitat i d'un volum insostenible. I aquí és on fem una filigrana que crec que pot resultar interessant teòricament i pràctica.

Es tracta de codificar la informació que tenim, teòricament a les equacions i el model de la figura 1, però «traduint-la» de forma que pugui ser utilitzada de manera senzilla i clara a l'hora de considerar el sistema globalment. Això es pot fer de dues maneres, analíticament i numèrica; les dues es complementen per obtenir aquesta versió «condensada» de la informació que tenim.

Per exemple, considerem la cèl·lula o taca colonitzada només per preses, sense depredadors (és a dir 1a i 1c amb  $y = 0$ ). Per un valor inicial de  $r$  que anomenarem  $r(0)$  ( $r$  al temps 0) positiu podem resoldre analíticament les equacions i obtenir:

$$r(t) = r(0) - \frac{u}{a} (x(t) - x(0)) \quad \dots(2a)$$

$$x(t) = x(0) e^{(b-d)t} \quad \dots(2b)$$

on

$r(t)$  i  $x(t)$  són la quantitat d'aliment i de presa en un moment donat  $t$ .

El temps que triga l'aliment en exhaurir-se és a dir, en prendre el valor 0 ve donat per l'eq. 2.

$$\tau = \frac{1}{2} \ln \left( \frac{(b-d)r}{ux(0)} + 1 \right) \quad \dots(3a)$$

i el valor de la població de presa en aquell moment serà

$$x(\tau) = \frac{(b-d)r(0)}{u} + x(0)$$

El model només ha de «recordar» els valors de  $r$ ,  $x$ ,  $y$  per a cada cèl·lula. Si en algun moment  $t$ , una presa migra a una cèl·lula no ocupada i amb una quantitat d'aliment  $r$ , l'eq. 3a) amb  $x(0) = 1$ ,  $r(0) = r$ , ens serveix per determinar quan ocorrerà la migració de la població d'aquella cèl·lula, que serà el moment  $t + \tau$ , (recordem que les preses migren només en exhaurir-se l'aliment). Quan arribi el moment de la migració, l'eq. 3b) s'utilitzarà per «posar al dia» la població de presa i fer o el valor de  $r$ . La desaparició d'aquella colònia de preses quedarà fixada

pel moment  $t = \frac{1}{2} \ln x(\tau)$ .

A partir d'aquí, a cada migració subsegüent es posa al dia l'estat de la cèhula.

Es pot fer veure que aquest algorisme *reprodueix* exactament el comportament del model anterior. Per afegir-hi el predador, en principi es fa el mateix; ara bé, la «posta al dia» de les cèhules i la determinació

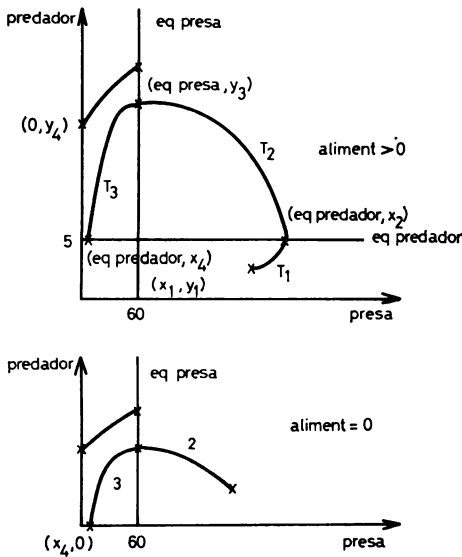


FIG. 2. — Variació de les poblacions de depredador i presa segons les equacions (1)

en el temps dels pròxims esdeveniments no es poden calcular analíticament.

Es poden, però, obtenir acurades aproximacions a partir de les trajectòries generades per les eq. 1.

Per exemple, considerem el cas que les preses han colonitzat una cèhula però no s'han acabat l'aliment. Aleshores amb les eq. 1a i 1b podem fàcilment obtenir les isoclines pels valors d'equilibri de presa i predador, o més exactament, de les equacions:

$$x^* = \text{presa en equil.} = \frac{d'}{c'}$$

$$y^* = \text{pred. en equil.} = \frac{c}{a}$$

Aquestes isoclines divideixen el pla en regions com podem veure a la figura 2.

Una trajectòria típica, iniciada per una migració d'un predador es segmenta en una fase inicial de creixement de les dues poblacions ( $\tau_1$ ), una fase de creixement de predador i decreixement de la presa ( $\tau_2$ ) i una fase de decreixement de les dues poblacions ( $\tau_3$ ). Seguint les observacions de HUFFAKER suposem que els predadors migren en assolir la seva població màxima, és a dir, al final de la fase ( $\tau_2$ ). Al final de la fase ( $\tau_3$ ) quan s'arriba al mínim de presa, es pot donar una desaparició de la colònia si els predadors són prou nombrosos o s'assoleix la isoclina amb el valor d'equilibri per a la presa. Les preses que queden a la cèhula prenen part subsegüentment a la interacció normal d'aliment i presa.

La determinació dels temps («scheduling») i les postes al dia dels censos per cadascuna de les possibles trajectòries (segons l'arribada de 1, 2, 3, 4 o 5 predadors) es fan amb el programa CSMP a partir de les taules generades de la simulació d'aquestes trajectòries amb les eq. 1 i els paràmetres obtinguts (fig. 3).

Aquestes taules són força interessants; que sapiguem és el primer estudi tan global de la dinàmica del model VOLTERRA LOTKA.

La base teòrica per a la representació en esdeveniments discrets es pot trobar a ZEIGLER (1977).

Per fer els càlculs s'ha utilitzat el llenguatge SIMSCRIPT 115 i un ordinador IBM 370.

El procediment de treball és el següent: Les cèhules vénen definides per tres coordenades.

Hi ha tres matrius que «recorden» a cada moment els valors de  $x$ ,  $y$  i  $r$  per a cadascuna de les cèhules. A més es disposa d'una llista d'esdeveniments ordenada cronològicament.

Queden encara, en la descripció del model dos punts importants per considerar. Un és el renovament de l'aliment i l'altre el control de les migracions.

La renovació de l'aliment es fa seguint les instruccions del que va fer HUFFAKER en el seu experiment, que consisteix a crear una sèrie de grups amb igual nombre de cèhules (taronges) distribuïdes a l'atzar en l'espai, i cada grup es renovava al cap de 11 dies. Si una taronja al moment que havia de ser renovada tenia una població d'àcars considerable ( $< 100$ ) no es renovava i s'afegia al grup següent. Si n'hi havia menys, els individus es retornaven a l'univers però es canviava la taronja.

Finalment descriurem breument com hem tractat el problema de la migració.

Considerem, basant-nos en les observacions de HUFFAKER, que les preses migren només quan s'acaba l'aliment ( $r = 0$ ). Aleshores es procedeix de la manera següent:

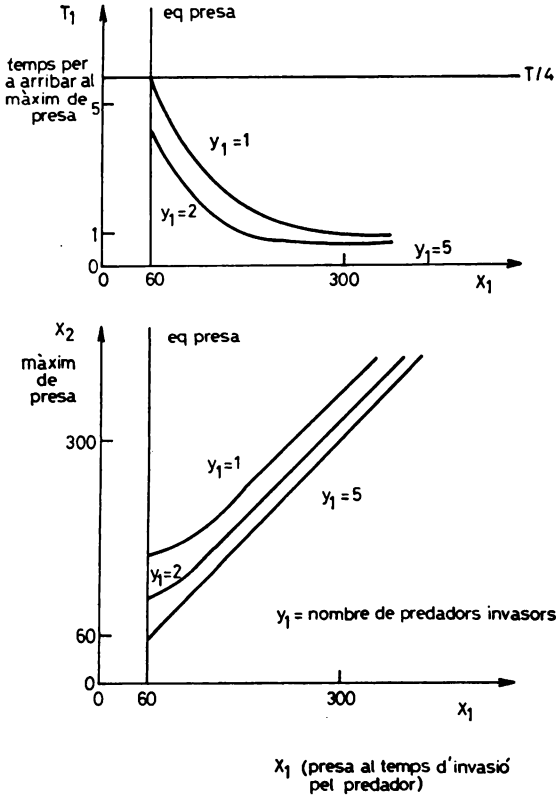


FIG. 3. — Trajectòries obtingudes amb el model de les equacions (1), utilitzant paràmetres ajustats a les dades de Huffaker

1. Quan s'exhaureix l'aliment de la població present la taronja es deixa allí una fracció d'aquesta població (*pyremain*) que queda subjecta a la cinètica descrita per eq. 1.
2. De les preses que emigren ( $(1-pyremain)$  vegades la població) es considera que una fracció (*pysurvive*) aconsegueix d'arribar a una altra cèlula. La resta es perd a efectes del sistema.
3. A cada individu emigrat li correspon un temps de recerca ( $T_s$ ) pres a l'atzar d'una distribució exponencial que té per valor mitjà un tercer paràmetre.
4. La cèlula d'«arribada» designada per aquell individu es calcula a partir d'una distribució normal (independentment per a cada dimensió), que té per mitjana una coordenada de la cèlula d'origen i per variància  $T_s \cdot pydif_1$  i  $T_s \cdot pydif_2$  per l'horitzontal i la vertical respectivament;  $dif_1$  i  $dif_2$  són els valors de difusió corresponents.



5. Un cop transcorregut el temps  $T_s$ , s'afegeix l'individu a la població de la cèl·lula designada, si a aquella cèl·lula hi ha aliment, es queda allí, si no, amb probabilitat  $p_{ysurvive}$  se'l torna a enviar al pas 2 per a una segona migració, i amb probabilitat  $(1-p_{ysurvive})$  mor.

La migració dels predadors es porta a terme de la mateixa manera amb diferents paràmetres i les úniques diferències són:

- 1' La migració de predadors comença en assolir-se un màxim en la seva densitat.
- 5' Un cop ha transcorregut el temps  $T_s + retard$  s'afegeix l'individu a la població assignada a no ser que la població de presa estigui per sota del valor d'equilibri de la presa, moment en què es considera que a efectes del predador la cèl·lula és buida.

X	1	2	3	4	5	6
Y						
1	13	22	21	23	10	3
2	10	36	73	35	14	6
3	21	73	242	86	15	4
4	22	48	62	39	17	7
5	8	14	18	13	8	4
6	4	11	7	7	2	0

FIG. 4. — Exemple de distribució espacial de 1000 individus emigrats d'una de les cèl·lules

El retard es va introduir com a hipòtesi a partir de les dades experimentals. És lògic pensar que hi hagi un retard entre la arribada del predador i el moment que comença a produir efectes sobre la població de presa i a reproduir-se.

A la figura 4 podem veure un exemple del tipus de distribució que aquest procediment genera en l'espai.

Tot aquest model es tradueix a un programa d'ordenador amb el qual hem simulat de manera francament satisfactòria els resultats obtinguts per HUFFAKER, tant en sentit qualitatiu com quantitatiu.

En aquest moment, estem buscant els paràmetres de migració més escaients per als universos més complexos i hem obtingut, pel que fa a per-

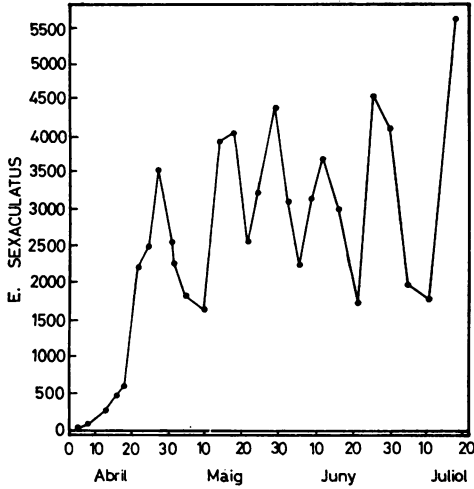


FIG. 5. — Mitja taronja a  $\frac{3}{4}$  parts dels llocs. Els altres s'ocupen amb boles de goma. Es comença amb 20 femelles en una sola taronja. En aquest cas no hi ha predador. Total 40 cèl·lules (taronges)

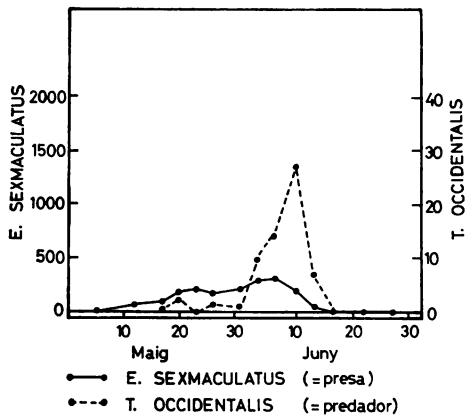
sistència i oscil·lacions de les poblacions, valors bastant pròxims als donats per l'autor.

Més endavant tenim intenció de publicar els resultats de manera ordenada. Compararem també els «patterns» espacials obtinguts amb els trobats experimentalment.

### III. Per acabar voldríem insistir en alguns punts d'ordre general.

Tal com dèiem al començament, creiem que el major interès de tota aquesta feina consisteix no tan en els «valors» o resultats obtinguts a par-

FIG. 6. —  $\frac{1}{20}$  de superfície de taronja a cada lloc. Hi ha barreres de vaselina. Es comença amb 10 femelles en 2 taronges, i 11 dies més tard 2 femelles de predador a una d'aquestes taronges. Total 120 cèl·lules



tir del model, sinó en el que hi ha darrera de tot això que hem presentat com un exemple concret, és a dir, el procediment de dividir l'aproximació a la descripció d'un sistema biològic. Aquesta divisió i la relativa autonomia formal i teòrica de les diferents parts o punts de vista poden resultar molt eficients.

En el nostre cas, les «cèl·lules», individualment, es tracten de manera determinista, però en canvi la migració és probabilística. Les restriccions a la cinètica de les cèl·lules i de les migracions són en aquest cas eminentment matemàtiques, per la senzilla raó que la formació i informació dels que hem construït aquest model era matemàtica i biològica o ecològica, però se'n poden introduir també d'ordre o forma física termodinàmica o més estrictament biològica.

Es pot pensar que tot això és «massa matemàtic» i poc biològic o ecològic. Hi estem bastant d'acord per a aquest exemple concret, però hom pensa que una equació diferencial tampoc no és gaire «matemàtica», i que si ha arribat a tenir un valor ecològic com a eina no és pas perdent la seva «matemàticitat» sinó posant-hi més biologia.

Hi ha diverses perspectives des de les quals aquesta eina pot resultar útil, d'una banda, tal com dèiem, la possibilitat d'introduir a diferents «nivells» diferents tipus de restriccions o descripcions segons la informació i el bagatge teòric de què disposem. D'altra banda, aquesta mateixa «agilitat» i independència de les diferents «parts» pot suposar un avantatge a l'hora d'aprofitar de manera eficient les dades.

Per exemple un cas en el qual es podria utilitzar aquesta eina amb algun profit seria el de la idea de successió. Sembla clar que diferents sistemes entrelligats per la successió es podrien descriure en un tipus de imatge com la que hem vist a la figura 1. Les restriccions i relacions

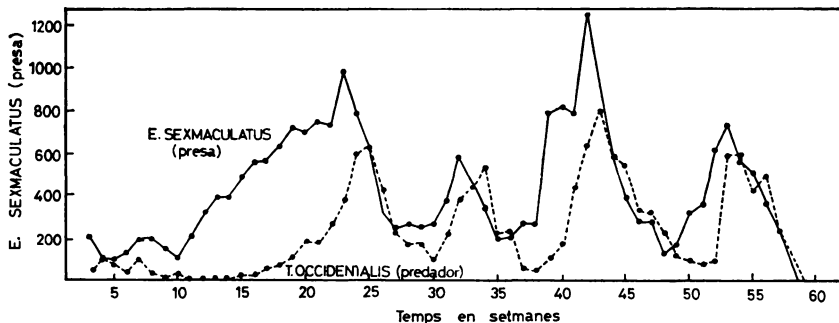


FIG. 7. — 1/20 de superfície de taronja a cada lloc. Es comença amb una femella a cadascuna de 63 taronges triades a l'atxar; 28 dies després es posa una femella de predador sobre 21 de les 63 taronges. Total 252 cèl·lules

entre els diversos estats, en l'espai i el temps, es podrien elaborar de manera semblant. Evidentment, això queda lluny, però no creiem que sigui impensable.

Finalment, creiem que tornant al problema original de la descripció de la morfologia espacial dels sistemes, es podria pensar a utilitzar aquesta manera de fer les coses per cercar una possible formalització que consideraria la morfologia en l'espai i el temps, no referida a *una* organització única, sinó a les organitzacions relatives a cada unitat. Caldria considerar no la morfologia (vista des de «fora»), sinó el que podríem anomenar transparència del sistema, l'organització condicionant per a cada element, unitat o nivell.

## APÈNDIX

A fi que es pugui veure més clarament la relació entre les dades experimentals i el model construït, afegim un breu resum dels dos articles de HUFFAKER sobre els quals ens hem basat.

Això pot servir per veure més clarament el tipus d'aproximació que hem seguit i el context en el d'aquestes dades es poden considerar.

Tal com hem dit, l'estructura bàsica dels experiments consisteix en una sèrie de distribucions espacials de taronges que es colonitzen amb àcars.

Les taronges poden tenir diferents valors nutritius, més exactament la superfície de les taronges que es deixa a l'abast de les preses és controlable. Aquest control s'exerceix senzillament tapant amb paper o parafina tota una part de la taronja i deixant descoberta només la que interessa.

Les preses i els predadors es col·loquen inicialment sobre les taronges en quantitats i distribucions que varien segons els experiments.

El nombre i distribució en l'espai de les taronges també varia amb els experiments; en alguns casos es col·loquen barreres entre les taronges que dificulten selectivament el pas de predadors o preses. No sempre totes les posicions són ocupades per taronges, de vegades s'inclouen boles de goma.

Les figures 5, 6, 7 donen una idea d'alguns dels experiments més característics.

## BIBLIOGRAFIA

1. HUFFAKER, C. B.: (1958). Experimental studies in predation: Dispersion factors and predator-prey oscillations. *Hilgardia* 27(14): 143-383.
2. HUFFAKER, C. B.; SHEA, E. P. & HERMAN, S. G.: (1973). Experimental studies on predation: Complex dispersion and level of food in an acarine predator-prey interaction. *Hilgardia* 54(9): 305-329.
3. LEVIN, S. A.: (1976). Spatial patterning and the structure of ecological communities, some mathematical questions in Biology VII. *A.M.S.*
4. MAYNARD-SMITH, J.: (1974). Models in Ecology. *Cambridge Univ. Press.*
5. ZEIGLER, B. P.: (1976). Theory of modelling and simulation. *Wiley, New York.*
6. ZEIGLER, B. P. (1977a). Structuring the organization of partial system. *Int. J. Gen. Sys.*
7. ZEIGLER, B. P. (1977b). Persistence and patchiness of predator-prey systems induced by discrete event population exchange mechanisms. *J. Theor. Biol.*
8. ZEIGLER, B. P. & BARTO, A. G.: (1977). Alternative formalisms for bio and eco-system modelling. *Simulation.*



# MÈTODES DE REPRESENTACIÓ DE DADES I LA SEVA APLICACIÓ A LA BIOLOGIA

per CARLES M. CUADRAS

Unitat Docent de Bioestadística. Facultat de Biologia.  
Universitat de Barcelona

## ABSTRACT

**Methods of data representation and their application in Biology**

*Several multivariate methods of data representation widely used in Biology are presented and compared. Their utilization in the analysis of biological data is discussed and several examples of application are given.*

## INTRODUCCIÓ

En aquest treball es comenten les tècniques per a representar dades, d'aplicació més corrent a Biologia. Aquestes tècniques multivariades formen un important capítol de l'anomenada «Geometria de l'Estadística», el qual té per finalitat analitzar i interpretar el material estadístic mitjançant representacions gràfiques.

La informació d'entrada en una anàlisi de dades és una matriu de dades  $X$

		Variables			
		$X_1$	$X_2 \dots$	$X_n$	
Individus	1	$X_{11}$	$X_{12} \dots$	$X_{1n}$	$x = (x_{ij})$
	2	$X_{21}$	$X_{22} \dots$	$X_{2n}$	
	$\vdots$		....		
	R	$X_{k1}$	$X_{k2} \dots$	$X_{kn}$	

on  $x_{ij}$  és el valor observat de la variable  $X_j$  sobre l'individu  $i$ .

Normalment,  $(X_i)$  són mesures biomètriques assimilades a variables aleatòries i els individus són  $k$  representants d'una espècie,  $k$  races geogràfiques o  $k$  espècies (etc.). Però de vegades la distinció entre espècies i variables no és prou clara. Els individus poden ésser espècies de zooplàncton i les «variables» els llocs geogràfics on han estat trobades, essent aleshores  $x_{ij}$  l'abundància (freqüència) de l'espècie  $i$  en la localitat  $j$ .

La informació de sortida en una anàlisi de dades és una representació euclídia dels individus en  $d = 2$  dimensions que descriu les analogies i diferències entre ells. Les representacions en  $d=3$  dimensions són menys freqüents.

En taxonomia numèrica, la informació de sortida és una representació gràfica anomenada dendrograma que permet establir una classificació jeràrquica entre els individus.

#### DISTÀNCIES ESTADÍSTIQUES

Donada la matriu de dades  $X$ , cada individu  $i$  el podem representar com un punt  $P_i$  de l'espai euclidià  $R^n$ , de coordenades

$$P_i : (x_{i1}, x_{i2}, \dots, x_{in})$$

Les distàncies entre aquests punts  $P_1, \dots, P_k$  informen sobre les esmentades analogies i diferències entre els individus. Cal doncs introduir una distància entre els individus. Hi ha moltes maneres de fer-ho. La més corrent és definir la distància<sup>2</sup> euclídia.

$$d^2(i, j) = d^2(P_i, P_j) = \sum_{h=1}^n (x_{ih} - x_{jh})^2$$

Però aquesta distància té alguns inconvenients: pressuposa que les variables són independents i queda alterada per canvis d'escala en les variables. Això no s'esdevé, en canvi, per a la distància de Mahalanobis,

$$D^2(i, j) = (P_i - P_j)' \cdot C^{-1} \cdot (P_i - P_j)$$

on  $C^{-1}$  és la inversa de la matriu  $C$  de covariàncies entre les variables.  $D^2$  té en compte les correlacions entre les variables i és independent de l'escala de mesura (metres, centímetres, etc.) de cada variable. Quan existeixen relacions lineals entre les variables, es pot prendre la distància:

$$D^2(i, j) = (P_i - P_j)' \cdot C \cdot (P_i - P_j)$$



on  $C^*$  és una  $g$ -inversa de  $C$  (la inversa de  $C$  no existeix), que té les mateixes propietats i no depèn de la  $g$ -inversa calculada.

Tant  $d$  com  $D$  són distàncies que provenen d'un producte intern entre les variables.

En efecte,

$$d^2(i, j) = s_{ii} + s_{jj} - 2 s_{ij} \text{ essent } s_{ij} = \sum_{h=1}^n x_{ih} x_{jh}$$

$$D^2(i, j) = P_i' \cdot C \cdot \bar{P}_i + P_j' \cdot C \cdot \bar{P}_j - 2 P_i' \cdot C \cdot \bar{P}_j$$

és a dir, són distàncies associades a productes interns de matrius  $I$  (identitat) i  $C^*$  respectivament, entre les variables  $X_1, \dots, X_n$ .

Aquestes dues distàncies són les que posen menys problemes de representació en dimensió reduïda. La distància  $D$  formulada de diferents maneres, ha estat emprada en genètica. Veure PREVOSTI (1974), OCAÑA (1975).

De vegades té interès utilitzar distàncies del tipus

$$\bar{d}(i, j) = \sum_{h=1}^n |x_{ih} - x_{jh}|$$

que no provenen d'un producte intern. Aleshores, la representació euclídia de les dades es fa seguint un camí diferent.

#### ANÀLISI DE COMPONENTS PRINCIPALS

Adoptem ara la distància euclídia  $d$ . Si  $T$  és una matriu ortogonal d'ordre  $n \times n$ , la transformació

$$Y = X T'$$

dóna una nova matriu de dades  $Y$ , que defineix una configuració de punts

$$Q_i : (y_{i1}, \dots, y_{in})$$

tal que  $d^2(P_i, P_j) = d^2(Q_i, Q_j)$ .

Els  $k$  individus estan aleshores també representats pels punts  $Q_i$ . La distància entre dos individus en relació a les  $d$  primeres dimensions és

$$d^2 (Q_i, Q_j)_d = \sum_{h=1}^n (y_{ih} - y_{jh})^2$$

Doncs bé, l'anàlisi de components principals obté la matriu  $T$  que fa que aquestes  $d$  dimensions siguin les més rellevants possibles. Es demostra que si  $T$  conté els  $n$  vectors propis de la matriu de covariàncies  $C$

$$C = T D T' \quad D = \text{diag} (\lambda_1, \dots, \lambda_n)$$

ordenats de forma creixent pels valors propis de  $C$ , aleshores

$$\sum_{i,j=1}^n d^2 (Q_i, Q_j)_d = \text{màxima}$$

és a dir, la dispersió, mesurada per la suma de distàncies entre els punts, és màxima en dimensió  $d$ .

La dispersió global és proporcional a  $(\lambda_1 + \dots + \lambda_n) = \text{Traça } (C)$ . Si  $(\lambda_1 + \dots + \lambda_d)$  absorbeix una part important (el 75 % per exemple) la representació en dimensió  $d$  és adequada.

La matriu de dades  $Y = X T$

		Components		
		Y <sub>1</sub>	Y <sub>2</sub> ...	Y <sub>n</sub>
Individus	1	y <sub>11</sub>	y <sub>12</sub> ...	y <sub>1n</sub>
	2	y <sub>21</sub>	y <sub>22</sub> ...	y <sub>2n</sub>
	⋮	.....		
	k	y <sub>k1</sub>	y <sub>k2</sub> ...	y <sub>kn</sub>

és tal que la variabilitat de les columnes decreix d'esquerra a dreta. Si prenem  $d=2$  agafarem les 2 primeres columnes, o sigui, les coordenades

$$(y_{11}, y_{12}), (y_{21}, y_{22}), \dots, (y_{k1}, y_{k2})$$

i farem la representació respecte a dos eixos ortogonals (Fig. 1).

A més a més, les noves variables

$$Y_j = t_{j1} X_1 + \dots + t_{jn} X_n \quad j=1, \dots, n$$

anomenades components principals, verifiquen: 1) estan incorrelacionades dues a dues, 2) tenen variàncies respectivament màximes

$$\text{Var} (Y_1) = \lambda_1 \cong \text{Var} (Y_2) = \lambda_2 \cong \dots \cong \text{Var} (Y_n) = \lambda_n$$

De vegades les components tenen interessants interpretacions biològiques.

Aquest tipus d'anàlisi es pot fer també diagonalitzant la matriu de correlacions  $R$ . Els resultats són diferents. Cal utilitzar  $C$  quan les variables són raonablement comparables (exemple: talles biomètriques mesurades en centímetres). En cas contrari (variables de naturalesa diferent) cal utilitzar  $R$ , que significa treballar amb variables sense dimensió física.

L'anàlisi de components principals va ésser introduït per HOTELLING (1933). RAO (1964) va escriure un esclaridor treball sobre la seva utilització i aplicació. S'han fet moltes aplicacions a la Biologia. Vegeu, per exemple, el treball de CHARDY, GLEMAREC i LAUREC (1976) aplicat a l'ordenació de comunitats bentòniques. Els autors comparen els resultats que dona l'anàlisi de components amb les anàlisis de coordenades principals i de correspondències.

*Exemple:* En un treball realitzat per ROMERO (1978) es pretèn tipificar diverses comunitats vegetals del massís de Collserola (Barcelona). Es tenen  $n=14$  formes biològiques, que fan el paper de variables (1=macrofaneròfits laurifolis, 2=macrofaneròfits escleròfils, ..., 9=camèfits suculents, ..., 13=geòfits, 14=hidròfits), i  $k=33$  inventaris d'espècies, amb les característiques generals següents: comunitats ruderals (24,25), prats secs (3,8,9,21,32,33), alzinars (1,2,6,12,17,19,29), matolls (15,30,31), diverses comunitats de caducifolis, relacionats amb la presència de microclimas humits (14,18,5,10,26), i carritxars (a la vora d'un estany, 16,17).

El valor  $x_{ij}$  que cada variable (forma biològica) fa correspondre a cada individu (inventari) és el nombre d'espècies de la forma biològica trobades a l'inventari que es feia sobre superfícies de  $10 \times 10$  m<sup>2</sup>. Per exemple, a l'inventari 6 (efectuat en un alzar) vàrem trobar 10 espècies de la forma 3, és a dir,  $X_{63} = 10$ .

La fig. 1 és la representació per anàlisi de components principals amb dimensió  $d=2$ . Els 2 primers eixos expliquen el 66,8 % de la variància total. Les conclusions que es treuen són:

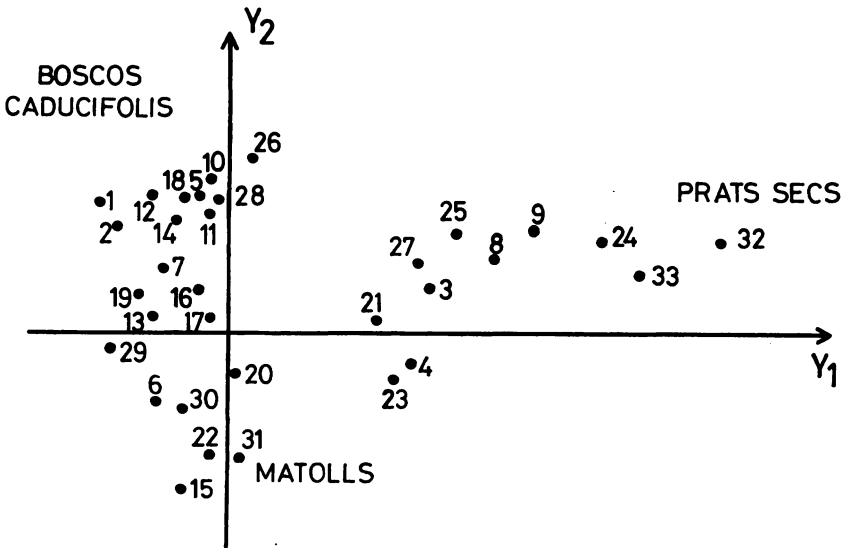


FIG. 1. — Representació per anàlisi de components principals de 33 inventaris d'espècies vegetals del massís de Collcerola (Barcelona)

1) La fisonomia dels inventaris ben definits queda ben reflectida; 2) Les comunitats intermèdies queden ben identificades. Aquests híbrids apareixen també en lloc intermedi de la gràfica; 3) Observant els alzinars, es veu que 1 i 2 tenen forta influència de caducifolis, mentre que 29 i 6 són alzinars degradats, amb un estrat arbustiu cada cop més important, en detriment de l'estrat arbori.

La interpretació dels eixos s'obté buscant les correlacions entre les variables inicials i les components principals. La correlació entre la variable  $X_i$  i la component  $Y_j$  és

$$\rho(X_i, Y_j) = \frac{t_{ij}}{\sigma_i} \sqrt{\lambda_j}$$

essent  $\sigma_i$  la desviació típica de  $X_i$ . Així, destaca la correlació 0,841 entre la primera component i la forma 12 (teròfits: plantes de cycle vital no superior a un any i que travessen l'estació desfavorable en forma de llavor). El primer eix s'interpreta com una dimensió que representa la riquesa de la comunitat en plantes anuals. La segona component correlaciona amb les formes 5 i 6 (nanofaneròfits planifolis i aciculifolis) i el segon eix sembla expressar la importància de l'estrat arbustiu.

## ANÀLISI DE COORDENADES PRINCIPALS

La tècnica per trobar les components principals és particularment adequada quan les variables són contínues.

Suposem ara que les variables  $X_1$  són dicotòmiques, basades en absència (—) o en presència (+) de caràcters qualitius. Un individu queda aleshores caracteritzat per les presències o les absències del  $n$  caràcters.

$$\begin{array}{cccccc} X_1 & X_2 & X_3 & \dots & X_n \\ i & + & - & + & \dots & + \end{array}$$

La informació útil és aleshores el nombre de caràcters presents sobre els  $n$  caràcters estudiats. Per representar els  $k$  individus, amb aquests tipus de dades, és molt útil l'anàlisi de coordenades principals introduïda per GOWER (1966).

L'associació entre els individus  $i, j$ , s'obté de la taula de freqüències

$$\begin{array}{c} \begin{array}{c} i \\ + \quad - \\ \hline \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array} \\ \hline \begin{array}{c} j \\ + \\ - \end{array} \end{array} \quad n = a + b + c + d$$

que conté el nombre de caràcters comuns  $a$ , el nombre de caràcters no comuns,  $d$ , etc. L'associació es mesura mitjançant un coeficient de similitat  $s_{ij}$  que verifica, en general, les propietats

$$0 \leq s_{ij} \leq 1$$

$$s_{ij} = 0 \quad \text{si} \quad c + b = n \quad (\text{discrepància total})$$

$$s_{ij} = 1 \quad \text{si} \quad a + b = n \quad (\text{concordància total})$$

La similaritat  $s_{ij}$  dona el grau de semblança entre  $i, j$ , en relació als  $n$  caràcters. Exemples de coeficients de similaritat són

$$\text{SOKAL i MICHENER: } \frac{a+d}{n}$$

$$\text{SOKAL i SNEATH: } \frac{a}{a+2(b+c)}$$

$$\text{JACCARD: } \frac{a}{a+b+c}$$

$$\text{RUSSELL i RAO: } \frac{a}{n}$$

Un cop el biòleg ha escollit el coeficient que millor expressa la similaritat entre els  $k$  individus, es forma la matriu d'associacions.

$$S = \begin{pmatrix} s_{11} & \dots & s_{1k} \\ s_{k1} & \dots & s_{kk} \end{pmatrix}$$

Definim ara la distància<sup>2</sup> entre  $i, j$

$$d_{ij}^2 = d^2(i, j) = s_{ii} + s_{jj} - 2s_{ij}$$

Quan  $s_{ij}$  és una similaritat del tipus descrit, és

$$d_{ij}^2 = 2(1 - s_{ij})$$

Aquesta distància és adequada ja que

$$\begin{aligned} d_{ii} &= 0 & i &= i, \dots, k \\ d_{ij} &= 0 & \text{si } s_{ij} &= 1 \text{ (similaritat total entre } i, j) \\ d_{ij} &= 2 & \text{(màxima distància) si } s_{ij} &= 0 \text{ (similaritat nulla).} \end{aligned}$$

El mètode de GOWER consisteix a trobar una matriu de dades

$$\begin{array}{l} \text{Individu} \\ 1 \\ 2 \\ \vdots \\ k \end{array} \begin{array}{|c|} \hline \begin{array}{cccc} y_{11} & y_{12} & \dots & y_{1k} \\ y_{21} & y_{22} & \dots & y_{2k} \\ \dots & \dots & \dots & \dots \\ y_{k1} & y_{k2} & \dots & y_{kk} \end{array} \\ \hline \end{array} \quad Y = (y_{ij})$$

que defineixi una configuració de punts en  $R^k$ , de manera que la seva distància euclídia coincideixi amb  $d_{ij}$

$$d_{ij}^2 = \sum_{h=1}^k (y_{ih} - y_{jh})^2$$

Agafant aleshores les dues primeres coordenades, tindrem una representació dels individus en dimensió  $d=2$ . Per obtenir aquestes coordenades es forma la matriu  $T=(t_{ij})$  essent

$$t_{ij} = s_{ij} - \bar{s}_i - \bar{s}_j + \bar{s}$$

on  $s_i$  és la mitjana de la  $i$ -èsima fila de  $S$ ,  $\bar{s}_j$  és la mitjana de la  $j$ -èsima columna,  $\bar{s}$  és la mitjana de tots els elements de  $S$ .  $T$  verifica

$$\text{rang}(T) = \text{rang}(S) - 1$$

i per tant té almenys, un valor propi nul. Les columnes de la matriu  $Y$  són els vectors propis de  $T$ , calculats de manera que

$$\sum_{h=1}^k y_{hj}^2 = \lambda_j \quad \lambda_j \text{ valor propi de } T$$

és a dir, que la seva norma sigui el valor propi corresponent. Si els vectors propis s'ordenen segons l'ordre decreixent dels valors propis, la dispersió de les  $d$  primeres columnes és màxima, en el sentit de la secció anterior.

*Exemple:* CANTON i SANCHO (1976) apliquen aquesta tècnica per a classificar 57 soques del gènere *Pseudomonas* en relació a 39 proves (glucosa, manitol..., producció de fluoresceïna), utilitzant el coeficient de similaritat de Jaccard.

En aquest cas,  $k=57$ ,  $n=39$ . La representació en dimensió  $d=3$  es troba a la Fig. 2. Els autors formen 7 grups de soques, que identifiquen totalment o parcialment, amb diverses espècies (*aeruginosa*, *mendocina*, *putrida*, etc.).

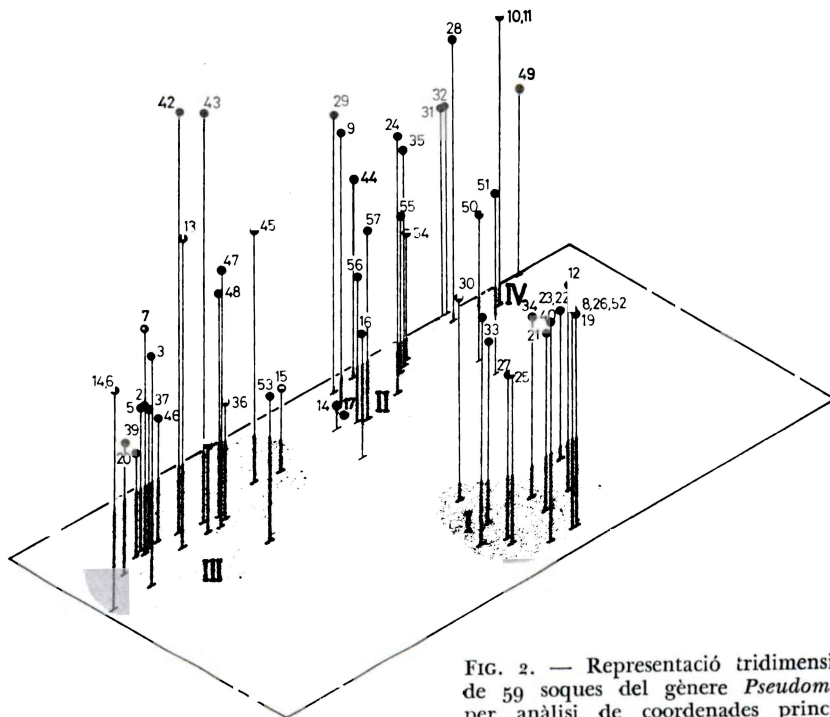


FIG. 2. — Representació tridimensional de 59 soques del gènere *Pseudomonas* per anàlisi de coordenades principals

ANÀLISI DE CORRESPONDÈNCIES

Suposem, tot seguit, que les nostres dades responen a dos criteris de classificació, els quals convencionalment podem anomenar de caràcters i poblacions, i que tenim una taula de contingència de freqüències

		Caràcters						
		A <sub>1</sub>	A <sub>2</sub>	...	A <sub>n</sub>			
Poblacions	1	f <sub>11</sub>	f <sub>12</sub>	...	f <sub>1n</sub>	f <sub>1.</sub>		
	2	f <sub>21</sub>	f <sub>22</sub>	...	f <sub>2n</sub>	f <sub>2.</sub>	f <sub>i.</sub> = $\sum_h f_{ih}$	
	⋮	.....						f <sub>j.</sub> = $\sum_h f_{hj}$
	k	f <sub>k1</sub>	f <sub>k2</sub>	...	f <sub>kn</sub>	f <sub>k.</sub>		
		f <sub>.1</sub> f <sub>.2</sub> ... f <sub>.n</sub>						



on  $f_{ih}$  és la freqüència de la població  $i$ , caràcter  $A_h$ . Les diferències entre dues poblacions queden reflectades per les diferències entre les distribucions de freqüència dels seus caràcters. Assignem, aleshores, a les poblacions les coordenades

$$\text{Població } i: \left( \frac{f_{i1}}{f_{i.}}, \frac{f_{i2}}{f_{i.}}, \dots, \frac{f_{in}}{f_{i.}} \right) \quad i = 1, \dots, k$$

que donen la distribució de probabilitats dels caràcters dins cada població  $i$ .

Per comparar dues poblacions, es defineix la distància xi-quadrat

$$d^2(i, j) = \sum_{h=1}^n \frac{1}{f_{.h}} \left( \frac{f_{ih}}{f_{i.}} - \frac{f_{jh}}{f_{j.}} \right)^2$$

Aquesta distància és també igual a

$$d^2(i, j) = \sum_{h=1}^n \left( \frac{f_{ih}}{f_{i.} \sqrt{f_{.h}}} - \frac{f_{jh}}{f_{j.} \sqrt{f_{.h}}} \right)^2$$

i per tant, si formem la nova taula de dades X

	$A_1$	$A_2$	$\dots$	$A_n$
1	$\frac{f_{11}}{f_{1.} \sqrt{f_{.1}}}$	$\frac{f_{12}}{f_{1.} \sqrt{f_{.2}}}$	$\dots$	$\frac{f_{1n}}{f_{1.} \sqrt{f_{.n}}}$
⋮	$\dots \dots \dots$			
k	$\frac{f_{k1}}{f_{k.} \sqrt{f_{.1}}}$	$\frac{f_{k2}}{f_{k.} \sqrt{f_{.2}}}$	$\dots$	$\frac{f_{kn}}{f_{k.} \sqrt{f_{.n}}}$

veiem que la distància Xi-quadrat es una distància euclídia ordinària entre  $k$  punts de  $\mathbb{R}^n$ , representats per la taula X.

La distància xi-quadrat reflecteix la diferència de les proporcions relatives dels caràcters a través de cada població. El divisor  $f_{.h}$  pondera la presència de petites desviacions dels caràcters de petita freqüència, i permet augmentar la seva importància.

L'anàlisi factorial de correspondències consisteix a aplicar una anàlisi de components principals a la matriu de dades  $X$ , tal com expliquem a la secció 3. La nova matriu de dades  $Y = XT$  defineix les coordenades de les  $k$  poblacions. Agafant les  $d$  primeres, tindrem una representació que tindrà màxima dispersió en dimensió  $d$ .

Fins aquí, aquest tipus d'anàlisi no és més que aplicar una anàlisi de components principals per a representar les poblacions, però agafant la distància xi-quadrat en lloc de la distància euclídia. L'interès d'aquest mètode resideix en el fet que també podem representar els caràcters amb referència a les seves diferències de distribució respecte a les poblacions. En efecte, es defineix la distància xi-quadrat entre els caràcters  $A_i, A_j$

$$d^2(A_i, A_j) = \sum_{h=1}^k \frac{1}{f_{h.}} \left( \frac{f_{hi}}{f_{.i}} - \frac{f_{hj}}{f_{.j}} \right)^2$$

que reflecteix la diferència de les proporcions relatives de les poblacions a través de cada caràcter. De manera anàloga, doncs, per anàlisi de components principals trobarem una matriu de dades  $Z$  que ens donarà les coordenades dels  $n$  caràcters. Agafant les  $d$  primeres, tindrem una representació que tindrà màxima dispersió en dimensió  $d$ .

Un altre avantatge és la possibilitat de representar simultàniament caràcters i poblacions, amb referència a uns mateixos eixos de coordenades. En efecte, si les coordenades en dimensió  $d$  són

$$\text{Població } i : (y_{i1}, y_{i2}, \dots, y_{id})$$

$$\text{Caràcter } A_j : (z_{j1}, z_{j2}, \dots, z_{jd})$$

es demostra la següent relació entre la  $h$ -èsima coordenada de la població  $i$ , i les  $h$ -èsimes coordenades dels  $n$  caràcters

$$y_{ih} = \frac{1}{\sqrt{\lambda_h}} \left( \frac{f_{i1}}{f_{.1}} z_{1h} + \dots + \frac{f_{in}}{f_{.n}} z_{nh} \right) \quad i = 1, \dots, k$$

on  $\lambda_h$  és el valor propi número  $h$ . Amb altres paraules, el punt que representa la població  $i$  és un terme mitjà dels punts que representen els caràcters, ponderat per les probabilitats de presència en la població  $i$ . La proximitat, en la representació simultània, d'una població a un determinat grup de caràcters, indica que predominen en aquesta població.

La representació simultània de caràcters i poblacions havia estat esmentada per KENDALL i estudiada per COOMBS, BENNETT i HAYS, entre els

anys 1950 i 1960. Però no arribaria a ésser popular fins que BENZECRÍ va introduir l'anàlisi de correspondències l'any 1963.

*Exemple:* ALONSO (1975) fa un ampli estudi de la distribució geogràfica del polimorfisme cromosòmic de *Drosophila subobscura* utilitzant anàlisi factorial de correspondències. Sobre la taula de freqüències de 66 poblacions i  $8+3+7+23+11$  ordenacions dels 5 cromosomes A, J, E, U i O respectivament, fa una anàlisi de correspondències global, i diverses anàlisis parcials. Utilitzarem com il·lustració una de les anàlisis parcials, concretament la que agafa 13 poblacions i 3 inversions del cromosoma A. Les dades es donen en forma de percentatge a la taula 1.

TAULA 1. — Percentatges de freqüència de tres inversions del cromosoma A per a 13 poblacions de *Drosophila subobscura*

Poblacions	Inversions		
	A-ST	A-1	A-1
HELSINKI . . . . .	96.0	4.0	0.0
DROBACK . . . . .	78.4	16,2	5.4
HERIOT . . . . .	100.0	0.0	0.0
DALKEITH . . . . .	100.0	0.0	0.0
GRONINGEN . . . . .	80.0	16.0	4.0
FONTAINEBLEAU . . . . .	88.5	7.7	3.8
VIENA . . . . .	56.9	35.8	7.4
ZURICH . . . . .	67.8	24.4	7.8
FRUSKA-GORA (YUG.) . . . . .	36.1	55.6	8.3
LAGRASSE . . . . .	72.5	17.5	10.0
MONTPELLIER . . . . .	60.2	24.3	15.5
CARASCO . . . . .	50.0	31.8	18.2

L'anàlisi s'ha de fer, naturalment, sobre les freqüències originals. El resultat és la fig. 3.

Observem que Heriot i Dalkeith queden representades en un mateix punt, donada la seva distribució idèntica. També queda reflectida la similitat entre Drobak i Fontainebleau. La població Frusca-Gora és la que queda més al marge, degut a la influència de l'ordenació A-1, menys freqüent a les altres poblacions. Les proporcions de les 3 ordenacions queden molt ben reflectides a la gràfica; es veu que les poblacions queden més pròximes de les ordenacions que s'hi presenten més.

*Comparacions entre les anàlisis de components principals, coordenades principals i correspondències.*

En aquesta secció es comenta, prenent com exemple situacions concretes, quin és el mètode més adequat per a representar unes dades determinades.

El punt fonamental que cal tenir ben present és que la representació gràfica dels objectes (individus, espècies, etc.) no és més que una imatge

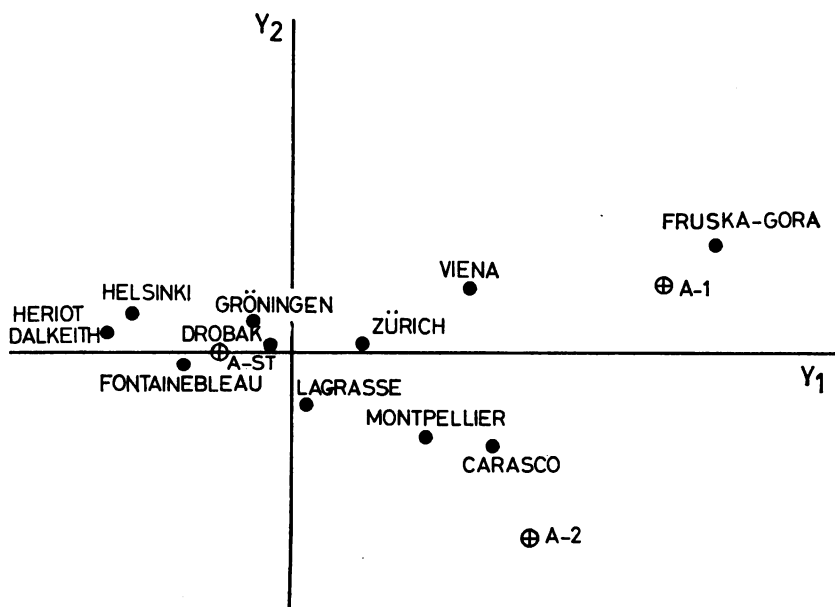


FIG. 3. — Anàlisi de correspondències sobre 13 poblacions europees de *Drosophila subobscura*, en relació a tres inversions del cromosoma A. Les inversions hi són també representades; reflecteixen la incidència que tenen a cada població

de la distància que definim entre objectes. Distàncies diferents donaran lloc a imatges diferents.

Suposem que volem comparar diverses races humanes amb referència a la distribució dels grups sanguinis. Si les freqüències per dues races són

	O	A	B	AB
Raça i	58	56	11	5
Raça j	68	64	12	6

la distància euclídica és 15,3, mentre que la distància xi-quadrat val 0. Aleshores, si apliquem l'anàlisi de components principals, les dues races sortiran més o menys diferenciades. Però, en canvi, obtindrem un mateix punt utilitzant l'anàlisi de correspondències. La diferència consisteix que en el primer cas es té en compte el nombre d'individus (130 i 150), és a dir, la grandària de la mostra influencia la discriminació. En el segon cas es compara la distribució de percentatges dels grups sanguinis, que és idèntica per a cada raça. En general, quan les dades s'assimilen a una taula de contingència o classificació doble de dos classes de característiques, és més apropiat l'anàlisi de correspondències. L'exemple donat a la secció anterior (poblacions X ordenacions cromosòmiques) és una correcta il·lustració d'això. L'anàlisi de components seria en canvi menys adequada perquè reflectiria el nombre de mosques de les poblacions, que és arbitrari ja que depèn en general del material experimental que aporten els diferents investigadors que han estudiat cada població.

Cal utilitzar l'anàlisi de components quan es volen comparar individus o espècies amb referència a caràcters biomètrics quantitius (llargada de cos, extremitats, pes, etc.). També és aconsellable aquesta anàlisi quan es volen comparar espècies respecte al nombre d'individus trobats en quadrats o zones de la mateixa extensió, obtinguts, de les mostres preses d'una gran extensió. Per exemple: nombre d'individus de cada una de  $k$  espècies de fitoplàncton trobats en  $n$  quadrats en el transcurs d'una campanya oceanogràfica. Si l'abundància de cada espècie és una dada important, l'anàlisi de components serà més adequat que el de correspondències. La primera component reflectirà, en general, la grandària o abundància de les espècies. Per poder donar aquesta interpretació cal que la primera component correlacioni positivament amb tots els quadrats (que fan el paper de variables). La forma de les poblacions es representa prenent la segona i tercera components.

L'exemple de representació d'inventaris de la secció 3 és una aplicació correcta d'anàlisi de components. Però en aquest cas la primera component no es podrà interpretar com un factor de grandària, ja que correlaciona negativament amb algunes formes. (Tinguem en compte que en l'exemple es representen els inventaris i no les formes.)

És important observar, també, que l'anàlisi de correspondències, donat el pes que es dóna a cada caràcter (o a cada quantitat), tendeix a destacar les poblacions que tenen una presència més singular.

Hi ha una altra manera de representar poblacions, fent una interpretació factorial de l'anàlisi de components principals. Suposem, per exemple, que tenim  $n$  espècies  $E_1, \dots, E_n$  i prenem com a dades el nom-

bre d'individus per espècie trobats en  $N$  mostres. Sigui  $R$  la matriu de correlacions entre les espècies i sigui

$$R = T D T' = A \cdot A' \quad A = D^{1/2}$$

( $T$ =matriu ortogonal amb els vectors propis,  $D$ =matriu diagonal amb els valors propis) una descomposició factorial de  $R$ . La matriu factorial  $A = (a_{ij})$

$$\begin{array}{c} E_1 \\ E_2 \\ \vdots \\ E_n \end{array} \left| \begin{array}{cccc} C_1 & C_2 & \dots & C_n \\ a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{array} \right.$$

dóna les saturacions entre les espècies ( $E_i$ ) i les components principals ( $C_i$ ). La saturació  $a_{ij}$  coincideix amb la correlació entre  $E_i$  i  $C_j$ . Aleshores es representen les poblacions agafant, com a coordenades de cada espècie, les fileres de  $A$

$$E_i: (a_{i1}, a_{i2}, \dots, a_{in}) \quad i = 1, \dots, n$$

Estudiem ara la distància<sup>2</sup> entre espècies:

$$d^2(E_i, E_j) = \sum_{h=1}^n (a_{ih} - a_{jh})^2 = \sum_{h=1}^n a_{ih}^2 + \sum_{j=1}^n a_{jh}^2 - 2 \sum_{h=1}^n a_{ih}a_{jh}$$

segons el teorema de THURSTONE, la correlació entre  $E_i, E_j$  és

$$\gamma_{ij} = \sum_{h=1}^n a_{ih} a_{jh}$$

i per tant,

$$d^2(E_i, E_j) = 2(1 - \gamma_{ij})$$

que coincideix amb la distància de GOWER construïda a partir d'un coeficient d'associació. Podem afirmar: l'anàlisi de components principals

que representen les espècies que utilitzen la matriu factorial, dona les mateixes distàncies que l'anàlisi de coordenades principals.

Però, si bé les representacions amb totes les dimensions coincideixen en els dos mètodes, les representacions amb dimensió reduïda no donen el mateix resultat. Per exemple, sigui la matriu de correlacions

$$R = \begin{pmatrix} 1 & 0.9 & 0 \\ 0.9 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Els resultats, aplicant components principals i coordenades principals són, respectivament,

$$A = \begin{pmatrix} 0,975 & 0 & 0,22 \\ 0,975 & 0 & -0,22 \\ 0 & 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 1,18 & 0,31 \\ 1,18 & -0,31 \\ -2,36 & 0 \end{pmatrix}$$

FIG. 4. — Representació de tres espècies per anàlisi de components sobre la matriu de correlacions

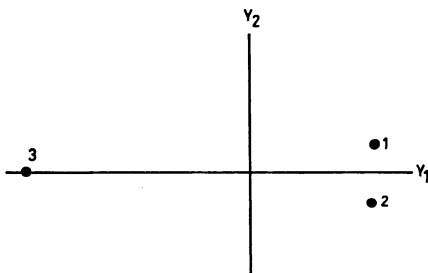
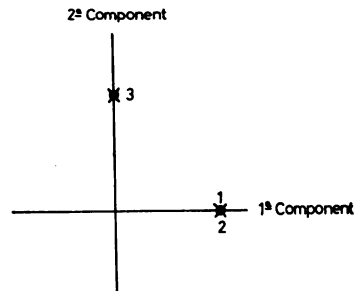


FIG. 5. — Representació de les mateixes espècies per anàlisi de coordenades principals

Les figures 4 i 5 són les corresponents representacions en dimensió 2.

De la Fig. 4 podem concloure que les espècies 1 i 2 estan molt lligades a la primera component i, per tant, molt relacionades entre sí;

l'espècie 3 defineix íntegrament la segona component. De la Fig. 5 podem concloure que les espècies 1 i 2 són molt pròximes entre sí, però allunyades de l'espècie 3. La Fig. 6 relaciona aquestes dues representacions.

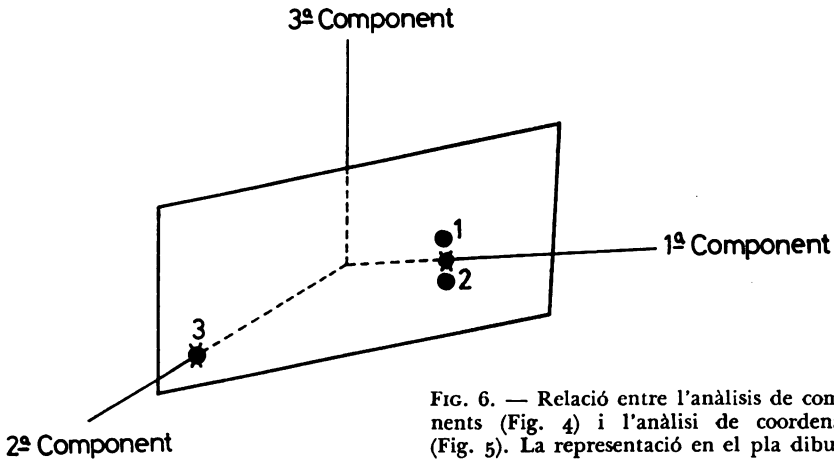


FIG. 6. — Relació entre l'anàlisi de components (Fig. 4) i l'anàlisi de coordenades (Fig. 5). La representació en el pla dibuixat es correspon amb la fig. 5

L'anàlisi de components tendirà a agrupar en direcció a un eix les variables fortament correlacionades amb una component, i en direcció oposada les variables amb correlació negativa. Unes variables poc correlacionades amb unes altres (3 respecte a 1 i 2, per exemple) quedaran representades ortogonalment. Com que la distància està basada en la correlació, no influeix la grandària (nombre d'individus) de l'espècie. Aquesta representació s'ha de interpretar segons els criteris de l'anàlisi factorial, que agrupen les variables (espècies) segons la seva relació amb uns determinats factors (per exemple: conjunts d'aigua definits per llur salinitat, temperatura, etc., en el cas de fitoplàncton), de la mateixa manera que les manifestacions de la personalitat es relacionen amb factors (neuroticisme, estabilitat, introversió, extroversió), fent un símil amb la Psicologia.

Encara que la distància global és la mateixa, l'anàlisi de coordenades representa distàncies amb màxima resolució en dimensió reduïda, en comptes de representar direccions. Les distàncies oscil·laran entre 0 (correlació 1) a 4 (correlació -1). Si el criteri d'aquesta anàlisi és maximitzar la suma de quadrats de les distàncies, el criteri de l'anàlisi de components és maximitzar la variància explicada per cada component. Les representacions en dimensió reduïda seran, en general, sem-



blants. De fet, l'anàlisi de coordenades principals ens aporta novetat metodològica quan la comparació entre espècies la fem sobre una matriu de similaritats calculada sobre variables dicotòmiques, ja que una anàlisi de components sobre aquesta matriu (que no és de correlacions) seria poc adequada.

Com a aplicació d'aquesta versió de l'anàlisi de components, vegeu ESTRADA (1975).

Estudiem ara la representació de dades de presència i d'absència de caràcters. Com hem vist a la secció 4, l'anàlisi de coordenades principals ens permet representar objectes en referència a una distància deduïda d'un coeficient de similaritat. Si codifiquem les dades en 0 (absència), en 1 (presència), la distància<sup>2</sup> euclídia és

$$d_{ij}^2 = b + c$$

Si escollim el coeficient de similaritat de SOKAL i MICHENER, la distància<sup>2</sup> deduïda d'aquest coeficient és

$$d_{ij}^2 = 2 \left( 1 - \frac{a+d}{n} \right) = \frac{2}{n} (b + c)$$

Tenim, en conseqüència, que l'anàlisi de components principals amb variables dicotòmiques és equivalent a una anàlisi de coordenades principals utilitzant el coeficient de SOKAL i MICHENER, llevat de la constant  $2/n$ , que no l'afecta, perquè tota representació de distàncies és relativa.

L'anàlisi de correspondències dóna també el mateix resultat si cada caràcter es presenta el mateix nombre  $t$  de vegades, perquè és fàcil veure que la distància xi-quadrat és

$$d^2(i, j) = \frac{1}{tn^2} (b + c)$$

Quan el nombre de presències del caràcter és variable, hi ha un factor de ponderació que, com ja hem esmentat, tendeix a destacar els objectes que tenen unes característiques que la majoria no tenen. Aleshores surten en llocs extrems de la gràfica.

Per concloure aquesta secció, tornem a repetir que el resultat de la representació gràfica depèn de la distància que definim, si bé és possible, en alguns casos, obtenir resultats semblants o àdhuc idèntics. Vegeu amb profit l'esmentat treball metodològic de CHARDY, GLEMAREC i LAUREC

(1976) per ampliar aquesta comparació dels tres mètodes de representació de dades en Biologia.

### TAXONOMIA NUMÈRICA

Donar una classificació ordenada dels fenòmens naturals és una de les tasques fonamentals de la ciència. Només fent una bona classificació és possible establir relacions entre la gran varietat de resultats de l'observació científica.

Tots els biòlegs coneixen molt bé el sistema de classificació sistemàtica de C. LINNEUS. En aquest sistema cada ser vivent té assignat un nom llatí amb el gènere i l'espècie.

El sistema taxonòmic de Linneus es pot descriure com una jerarquia organitzada a nivells, on les classes disjunctes a cada nivell constitueixen les anomenades *taxes*. A un nivell donat, les taxes constitueixen les categories. Així es parla de les categories «espècies», «gèneres», «famílies», «ordres», «classes», etc. La categoria gènere, per exemple, té diverses taxes: els gèneres corresponents a una família donada.

La taxonomia numèrica és un intent de construir classificacions naturals sobre la base de la semblança fenotípica entre els individus. Aquesta semblança fenotípica es mesura agafant un conjunt de caràcters significatius i calculant una matriu de similituds  $S$ , per algun dels procediments explicats a la secció anterior. Però en taxonomia numèrica és preferible treballar amb dissimilaritats. Una dissimilaritat és una distància que en general no verifica la propietat triangular. S'obté una dissimilaritat només posant

$$d_{ij} = 1 - s_{ij}$$

Però hi ha també altres maneres de definir dissimilaritats.

La dissimilaritat mesura les diferències fenotípiques entre dos individus o dues classes. Doncs bé, conegudes les dissimilaritats entre  $k$  individus, el resultat final d'una taxonomia numèrica és una jerarquia indexada o dendrograma, que és una representació gràfica de la classificació.

Els termes família, gènere, espècie, etc., tenen aleshores un significat més precís, ja que es parla de  $d$ -taxes o classes amb distància fenètica. Per exemple, la similitud entre les espècies 2 i 3 ( $d=0,3$ ) és més gran que entre les espècies 4 i 5 ( $d=0,5$ ). A partir de  $d=0,6$  es parla de gèneres i a partir de  $d=0,8$  es parla de famílies. Hi ha dues famílies (1,2,3) + (4,5)

i tres gèneres (1)+(2,3)+(4,5). La distància fenètica entre el gènere (1) i el gènere (2,3) és 0,8, etc. La distància fenètica  $d$  és l'índex de la jerarquia, com ara explicarem.

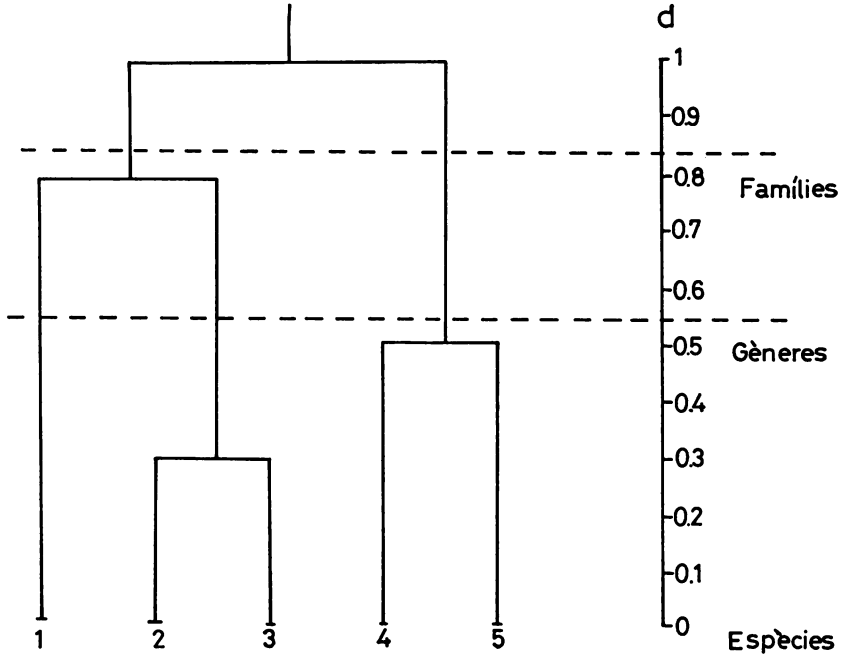


FIG. 7. — Exemple hipotètic de dendrograma;  $d$ =distància fenètica

Passem ara a donar els conceptes generals. Sigui  $E = (1, 2, \dots, k)$  un conjunt d'objectes (individus, espècies, etc.). Es diu que  $H$ , conjunt de parts de  $E$ , és una jerarquia, si:

- 1) Per tot  $h, h' \in H$  o bé  $h \subset h'$ , o  $h' \subset h$  o  $h \cap h' = \emptyset$
- 2) Tot  $h$  és reunió dels  $h'$  inclosos en  $h$
- 3)  $\{i\} \in H$  per tot  $i \in E$ ;  $E \in H$

Els elements de  $H$  s'anomenen classes.

Per exemple:

$$H = \left\{ \{1\}, \{2\}, \{3\}, \{4\}, \{5\}, \{2,3\}, \{4,5\}, \{1,2,3\}, \{1,2,3,4,5\} \right\}$$

és una jerarquia sobre  $E = \{1,2,3,4,5\}$ , representada en la fig. 7.

Una jerarquia  $H$  es diu que és indexada si existeix una aplicació  $d: H \rightarrow \mathbb{R}^+$  verificant

$$\begin{aligned} d(\{i\}) &= 0 && \text{per tot } i \in E \\ h \subset h' &\text{ implica } d(h) < d(h') \end{aligned}$$

essent la inclusió estricta.

Per exemple  $d(\{2,3\}) = 0,3$ ,  $d(\{4,5\}) = 0,5$ ,  $d(\{1,2,3\}) = 0,8$ , etc., és un índex sobre  $H$  que coincideix amb l'esmentada distància fenètica.

La representació d'una jerarquia indexada és el dendrograma, que permet establir una classificació jeràrquica entre els objectes de  $E$ .

Per trobar una tal jerarquia cal construir sobre  $E$  un tipus especial de dissimilaritat anomenada ultramètrica. Una ultramètrica verifica el següent axioma

$$u(i,j) \leq \sup \{u(i,t), u(j,t)\} \quad \text{per tot } i,j,t \in E$$

Es demostra fàcilment que aquest axioma és equivalent al fet que tot triangle (conjunt format per tres objectes) és isòsceles, i la base és el costat més petit.

Doncs, bé, coneguda una ultramètrica  $u$ , es pot construir una jerarquia indexada i recíprocament. La jerarquia  $H$  donat  $u$ , s'obté per un algorisme que va ajuntant els objectes més pròxims per la dissimilaritat  $u$ . Recíprocament, es construeix  $u$  a partir d'una jerarquia indexada posant

$$u(i, j) = d_h$$

si  $h$  és la classe més petita que conté  $i, j$ . Per exemple, de la fig. 3 tindríem:  $u(2,3) = 0,3$ ,  $u(1,3) = 0,8$ ,  $u(1,2) = 0,8$ . Fixem-nos que  $\{1,2,3\}$  és un triangle isòsceles amb base  $\{2,3\}$ .

Així doncs, podem construir una classificació jeràrquica si tenim una dissimilaritat ultramètrica definida sobre  $E$ . Aquest seria l'ideal del taxonomista. Però, per desgràcia, només trindrem, en general, una matriu de dissimilaritats  $\Delta = (\delta_{ij})$  on  $\delta_{ij}$  (que moltes vegades es calcula a partir d'una similaritat  $s_{ij}$ ) que no és pas una ultramètrica. Un algorisme de classificació consisteix en obtenir una distància ultramètrica partint d'una dissimilaritat  $\delta_{ij}$ . Aleshores la distància ultramètrica es representa mitjançant un dendrograma.

$$\Delta = \begin{pmatrix} \delta_{11} & \dots & \delta_{1k} \\ \dots & \dots & \dots \\ \delta_{k1} & \dots & \delta_{kk} \end{pmatrix} \xrightarrow{\text{Algorisme U}} \begin{pmatrix} u_{11} & \dots & u_{1k} \\ \dots & \dots & \dots \\ u_{k1} & \dots & u_{kk} \end{pmatrix} \longleftrightarrow \text{Dendrograma}$$

Matriu dissimilaritats                      Matriu ultramètrica

Cal que la ultramètrica no deformi gaire la dissimilaritat inicial. Vegem alguns d'aquests algorismes.

JOHNSON (1967) proposa obtenir la més gran de totes les ultramètriques inferiors a

$$u_{iM}(i,j) = \sup \{ u(i,j) \mid u(i,j) \leq \delta_{ij} \}$$

$u_{iM}$  és la ultramètrica inferior màxima. Equival a deformar tot triangle  $(i,j,t)$  prenent com a base el costat més petit i com a longitud dels costats iguals (tot triangle és isòsceles) la del següent costat més petit. La ultramètrica  $u_{iM}$  és única i és sistemàticament inferior a  $\delta$ . Com algorisme de classificació, havia estat introduït per Sneath l'any 1957.

JOHNSON (1967) proposa també la més petita de totes les ultramètriques superiors a  $\delta$

$$u_{sM}(i,j) = \inf \{ u(i,j) \mid u(i,j) \geq \delta_{ij} \}$$

$u_{sM}$  és una ultramètrica superior mínima i n'hi poden haver diverses. Aquesta ultramètrica és sistemàticament superior a  $\delta$ . El corresponent algorisme de classificació havia estat introduït per SÖRENSEN l'any 1948.

Les dos ultramètriques de JOHNSON són inferiors i superiors respectivament a la dissimilaritat inicial. Altres autors han proposat ultramètriques que siguin un terme mitjà entre aquestes dues. Vegeu ROHLF (1970).

El UPGMA (*Unweighted pair group method using method averages*) de SOKAL i SNEATH (1963) és un dels més utilitzats.

La taxonomia numèrica es va desenvolupar a partir del 1957 amb la publicació de diversos articles, en defensa d'aquest mètode, escrits pel microbiòleg britànic P. H. SNEATH i els entomòlegs americans C. D. MICHENER i R. K. SOKAL. Però, sobretot, es va desenvolupar després de l'obra *Principles of Numerical Taxonomy*, publicada l'any 1963, escrita per

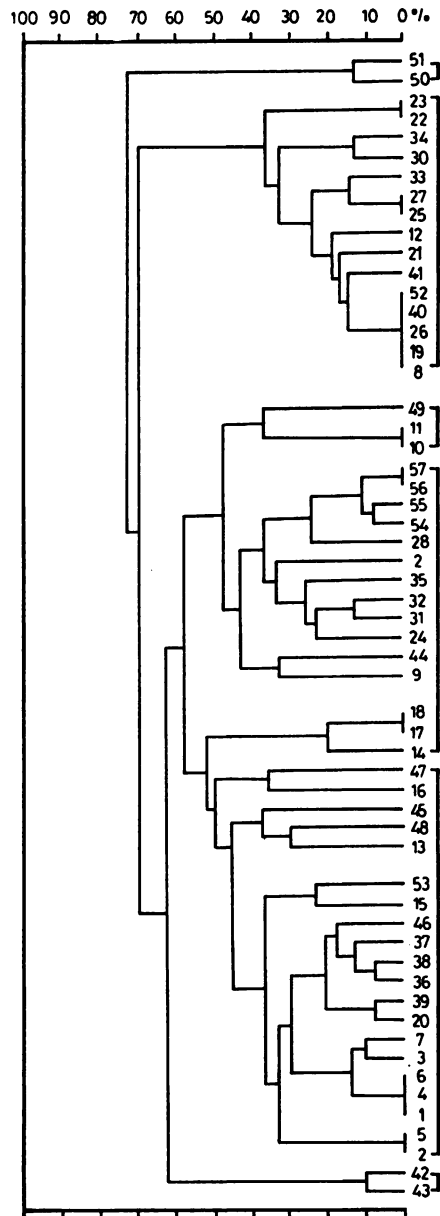


FIG. 8. — Dendrograma de 59 soques del gènere *Pseudomonas* obtingut pel algorisme UPGMA. L'index és la distància fenètica, que té les propietats de distància ultramètrica

SOKAL i SNEATH, on s'exposa l'estudi teòric de la classificació, incloent les seves bases, principis, procediments, i regles.

*Exemple:* Continuant l'exemple de la secció 4, E. CANTON i J. SANCHO (1976) realitzen també una taxonomia numèrica de les 59 soques del gènere *Pseudomonas*. El coeficient de similaritat de Jaccard el transformen en dissimilaritat posant

$$\delta_{ij} = 100 (1 - s_{ij})$$

Aquesta dissimilaritat oscil·la entre 0 (soques fenotípicament idèntiques) i 100 (soques completament oposades). El resultat d'aplicar el algorisme UPGMA a la matriu de dissimilaritats és el dendrograma de la fig. 8. Els autors comenten aleshores, els diferents grups trobats. (En realitat els autors treballen amb l'índex  $100 \cdot s_{ij}$ , que és complementari de  $\delta_{ij}$ ; el resultat és idèntic en interpretació.)

Per estudiar el grau de distorsió de la distància ultramèrica respecte la dissimilaritat inicial, es correlacionen els  $k(k-1)/2$  parells de distàncies ( $\delta_{ij}$ ,  $u_{ij}$ ). El coeficient de correlació obtingut s'anomena correlació cofenètica. Els autors troben una correlació cofenètica de 0.96. El dendrograma interpreta molt bé les dissimilaritats inicials.

#### ANÀLISI DE PROXIMITATS

Donats  $k$  objectes (individus, espècies, estímuls) i coneguda una determinada informació  $\Delta$  sobre les diferències entre els objectes, l'anàlisi de proximitats («multidimensional scaling» en anglès) és una tècnica que obté una configuració euclídia, formada per  $k$  punts  $P_1, \dots, P_k$ , de manera que les seves distàncies euclídies concordin amb la informació  $\Delta$ .

Aquesta informació pot venir donada de diferents maneres: ordenació entre les distàncies dels objectes, matriu de dissimilaritats o distàncies no euclídies, distàncies afectades d'errors aleatoris o d'aproximació numèrica, etc.

SHEPARD (1962 a,b) va formular la següent conjectura: Coneguda només l'ordenació entre les distàncies dels objectes

$$d(i_1, i_2) \leq d(i_3, i_4) \leq \dots \leq d(i_{m-1}, i_m) \quad (m = k(k-1)/2)$$

existeix una configuració euclídia tal que les seves distàncies verifiquen aproximadament aquesta ordenació.

SHEPARD no va donar una prova rigorosa de la seva conjectura, però la va il·lustrar amb nombrosos exemples i va formular un algorisme nu-

mèric que trobava aquesta configuració. Altres autors varen afirmar i desenvolupar la idea de SHEPARD, proposant mètodes de construcció (KRUSKAL, GUTTMAN, LINGOES, YOUNG, BENECRI, CARROLL, etc.).

Vegem-ne tres exemples. Siguin quatre objectes A,B,C,D, tals que les ordenacions entre les seves distàncies desconegudes siguin (poden haver-hi rangs iguals):

	A	B	C	D		A	B	C	D		A	B	C	D
A	0	1	2	3	A	0	1	1	1	A	0	1	1	1
B		0	1	2	B		0	1	1	B		0	1	1
C			0	1	C			0	0	C			0	1
D				0	D				0	D				0
	(1)					(2)					(3)			

Aleshores estarien representades per configuracions euclídiades en 1, 2 i 3 dimensions respectivament.

L'interès del mètode de SHEPARD és trobar una configuració en un espai de dimensió reduïda que reproduïxi aproximadament l'ordenació inicial. Aleshores podrem representar els *k* objectes gràficament, conservant les relacions de proximitat.

El sol coneixement de l'ordenació és el cas d'informació més pobre. Normalment es coneix una matriu de distàncies no euclídiades o dissimilaritats  $\Delta = (\delta_{ij})$ . El resultat final d'una anàlisi de proximitats és una matriu de dades *Y*,

$$\begin{array}{c}
 \begin{array}{c} 1 \\ 2 \\ \vdots \\ k \end{array} \left| \begin{array}{cccc}
 1 & 2 & \dots & d \\
 y^{11} & y^{12} & \dots & y^{1d} \\
 y^{21} & y^{22} & \dots & y^{2d} \\
 \dots & \dots & \dots & \dots \\
 y^{k1} & y^{k2} & \dots & y^{kd}
 \end{array} \right. P_i: (y_{i1}, \dots, y_{id})
 \end{array}$$

que defineix *k* punts de  $R^d$  tals que les seves distàncies<sup>2</sup>

$$d^2 (P_i, P_j) = d^2_{ij} = \sum_{h=1}^d (y_{ih} - y_{jh})^2$$



siguin semblants a les dissimilaritats  $S_{ij}$ . La matriu  $Y$  es pot calcular trobant primer la matriu  $S = (s_{ij})$  essent

$$s_{ij} = - \frac{1}{2} (\delta_{ij}^2)$$

i diagonalitzant a continuació la matriu  $T$ , pel procediment explicat en l'anàlisi de coordenades principals (vegeu GOWER, 1966). Però la matriu

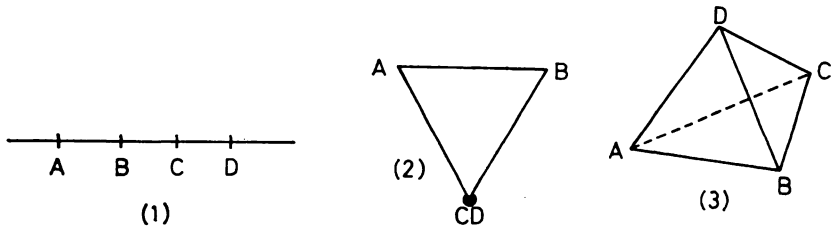


FIG. 9. — Representació mètrica en 1, 2 i 3 dimensions de 4 objectes, coneguts només els rangs (1), (2) i (3) entre les seves proximitats

$T$  pot tenir valors propis negatius i obtenir, doncs, una configuració poc concordant amb  $\Delta = (\delta_{ij})$ .

El mètode de KRUSKAL (1964) amplia el mètode de SHEPARD i dóna solucions a aquest problema. Consisteix a construir una nova dissimilaritat, anomenada disparitat, que sigui una funció monòtona creixent de  $\delta_{ij}$ , és a dir,

$$d_{ij} = f(\delta_{ij}) \quad \delta_{ij} \leq \delta_{i,j}, \rightarrow \hat{d}_{ij} \leq \hat{d}_{i'j'}$$

Aquesta transformació de  $\delta_{ij}$  conserva l'ordenació inicial. Utilitzant  $d_{ij}$  es troba una matriu de dades  $Y$  (pel procediment indicat) que dóna unes distàncies euclídiades  $d_{ij}$ . Com a mesura d'ajust entre  $d_i$  i  $d_{ij}$ , Kruskal defineix la quantitat

$$S = \left( \frac{\sum_{i < j} (d_i - d_{ij})^2}{\sum_{i < j} d_{ij}^2} \right)^{1/2}$$

anomenada *stress*. Verifica  $0 \leq S \leq 1$ , però es dóna en forma de percentatge. KRUSKAL desenvolupa un mètode adient per trobar  $d_{ij}$ ,  $Y$  i  $d_i$  de manera que quedi minimitzat el *stress*. Considera que la representació és bona si no supera el 5%.

Agafant les coordenades que conté Y, tindrem una representació dels objectes que conserva aproximadament les relacions de proximitat inicials, malgrat que la distància euclídia sigui una deformació de la dissimilaritat inicial.

El *stress* disminueix a mida que augmenta la dimensió *d*. Cal agafar *d* que doni un *stress* acceptable. Cal també calcular el coeficient de correlació ordinal de KENDALL entre les distàncies euclídies i les dissimilaritats originals, per mesurar el grau de concordància entre l'ordenació inicial i l'ordenació entre les distàncies dels punts representats.

*Exemple:* PREVOSTI (1974) proposa una distància genètica basada en les diferències d'ordenacions cromosòmiques. La distància entre dues poblacions és

$$D(1,2) = \frac{1}{2r} \sum_{j=1}^r \sum_{h=1}^{s_j} |p_{1jh} - p_{2jh}|$$

on *r* és el nombre de diferents cromosomes, *s<sub>j</sub>* és el nombre de diferents ordenacions en el cromosoma *j* (ordenacions dels seus gens), *p<sub>1jh</sub>* i *p<sub>2jh</sub>* són les freqüències de l'ordenació *h* del cromosoma *j* en les poblacions 1 i 2 respectivament. Aquesta distància verifica  $0 \leq D(1,2) \leq 1$ . Els seus avantatges com a distància genètica han estat discutits per l'autor.

Però aquesta distància no prové d'un producte escalar. Una configuració euclídia entre poblacions, que reproduïx les distàncies originals, es pot obtenir per anàlisi de proximitats.

Per a il·lustrar-ho, farem servir part de les dades del treball de PREVOSTI, OCAÑA i ALONSO (1975), que aplica aquesta distància per estudiar el polimorfisme cromosòmic de *Drosophila subobscura*. Les distàncies genètiques entre *k* = 6 poblacions del nord i centre d'Europa són:

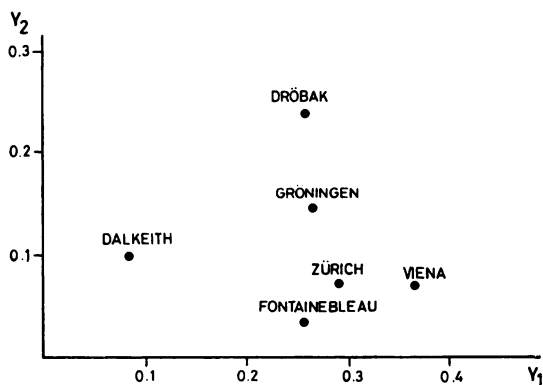
	1 Dröbak	2 Dalkeith	3 Gröningen	4 Fontainebleau	5 Viena	5 Zürich
1	0	0,307	0,152	0,271	0,260	0,235
2		0	0,276	0,225	0,370	0,300
3			0	0,150	0,187	0,112
4				0	0,195	0,120
5					0	0,128
6						0

Realitzant una anàlisi de proximitats amb dimensió  $d=2$ , per un algorisme iteratiu desenvolupat per CUADRAS (1979), s'obté que la funció monòtona creixent

$$f(\delta_{ij}) = d_{ij} = 1.062 \delta_{ij} - 0.04 \quad (\delta_{ij} = D(i,j))$$

dóna un *stress* del 4,4%. La fig. 10 és una representació de la configuració euclídia obtinguda.

FIG. 10. — Representació en dues dimensions de 6 poblacions de *Drosophila subobscura* per anàlisi de proximitats, amb un «stress» del 4,4%. Les distàncies representades es diferencien molt poc de la distància genètica inicial. Compareu amb la fig. 3



L'evolució dels *stress* en diferents dimensions va ésser:

Dimensió:	1	2	3	4	5
<i>Stress</i> :	40,0	4,4	1,27	0,24	0

El coeficient de correlació ordinal de KENDALL entre les distàncies dels punts de la fig. 2 i les distàncies originals va donar 0,9929. Les ordenacions entre ambdós conjunts de distàncies pràcticament coincideixen.

Cal observar que la matriu T calculada pels 6 punts no tenia cap valor propi negatiu. Això vol dir que existeix una configuració euclídia que reproduïx exactament les distàncies originals. Aquesta configuració que té dimensió 5, es pot obtenir per anàlisi de coordenades principals, amb resultats similars als obtinguts. La utilitat de l'anàlisi de proximitats es fa patent quan agafem moltes poblacions. Aleshores es presenta una influència notable dels valors propis negatius, i no existeix cap configuració euclídia, a menys que deformem monotònicament les distàncies inicials.

ANÀLISI CANÒNICA DE POBLACIONS

Malgrat que la distància  $D$  de Mahalanobis és la més perfecta per a representar dades, encara no l'hem utilitzada. El problema que significa utilitzar  $D$  resideix en el fet que, generalment, no es coneix la matriu de covariàncies  $C$  entre les variables. Si poguéssim conèixer  $C$  amb independència de la matriu de covariàncies de la taula de dades  $X$ , podríem representar els  $k$  individus aplicant la distància  $D$ , amb resultats més escaients que l'anàlisi de components principals, la qual utilitza la distància euclídia ordinària.

En l'anàlisi canònica de poblacions és possible fer-ho. Suposem que cada espècie (o raça, o grup, etc.), està representat per més d'un individu, i que la taula de dades és ( $k$  poblacions i  $p$  variables)

		Variables				
		$X_1$	$X_2$	...	$X_p$	
Població 1	$X_{111}$	$X_{121}$	...	$X_{1p1}$	}	$n_1$
	$X_{11n_1}$	$X_{12n_1}$	...	$X_{1pn_1}$		
⋮						
Població k	$X_{k11}$	$X_{k21}$	...	$X_{kp1}$	}	$n_k$
	$X_{k1n_k}$	$X_{k2n_k}$	...	$X_{kpn_k}$		

on  $n_t$  és el nombre d'individus de la població  $t$ . Prenem com a representant de la població  $t$  l'individu mitjà de coordenades

$$M_t: (\bar{x}_{t1}, \bar{x}_{t2}, \dots, \bar{x}_{tp}) \quad t = 1, 2, \dots, k$$

on  $x_{it}$  és la mitjana de la variable  $x_i$  dins de la població  $t$ . Aleshores podem estimar la matriu de covariàncies calculant

$$C = \frac{1}{N - k} \sum_{t=1}^k n_t S_t \quad (N = \sum_{t=1}^k n_t)$$

on  $S_t$  és la matriu de covariàncies dins la població  $t$ ,

$$S_t = (s_{ij}) \text{ essent } s_{ij} = \frac{1}{n_t} \sum (x_{tjh} - \bar{x}_{tj}) (x_{tjh} - \bar{x}_{tj})$$

La distància de Mahalanobis entre les poblacions  $i, j$  és

$$D^2(i, j) = (M_i - M_j)' C^{-1} (M_i - M_j)$$

El problema a resoldre és trobar una configuració euclídia de  $k$  punts, deduïda d'una taula de dades  $Y$ ,

Eixos canònics

	1	...	m
Població 1	y <sub>11</sub>	...	y <sub>1m</sub>
...			
Població k	y <sub>k1</sub>	...	y <sub>km</sub>

de manera que la distància euclídia coincideixi amb  $D$

$$\sum_{h=1}^n (y_{ih} - y_{jh})^2 = D^2(i, j)$$

S'anomenen eixos canònics els utilitzats per a representar les  $k$  poblacions. La dimensió màxima és  $m = \min(k-1, p)$ .

Per trobar la taula de coordenades canòniques  $Y$ , cal obtenir la matriu  $B = (b_{ij})$  essent

$$b_{ij} = \frac{1}{k} \sum_{h=1}^k (\bar{x}_{ih} - \bar{x}_i) \cdot (\bar{x}_{jh} - \bar{x}_j)$$

la «covariància» entre les variables  $X_i, X_j$ , agafant només les  $k$  mitjanes. Aleshores es busquen els  $m$  vectors propis de  $B$  respecte de  $C$  resolent l'equació matricial

$$B V_i = \lambda_i C V_i \quad (\lambda_i = \text{valor propi de } B \text{ respecte de } C)$$

Si  $V$  és la matriu dels vectors propis, aleshores

$$Y = MV$$

essent  $M$  la matriu que conté les  $k$  mitjanes per les  $p$  variables.

Finalment, prenent les  $d$  primeres coordenades, tindrem una representació en màxima dispersió en dimensió  $d$ . Es verifica

$$\sum_{i,j=1}^k \sum_{h=1}^d (y_{ih} - y_{jh})^2 = 2k (\lambda_1 + \dots + \lambda_d)$$

L'anàlisi canònica ha d'anar acompanyada d'un test multivariant de comparació de mitjanes i d'un test d'homogeneïtat entre les  $k$  matrius de covariàncies de les poblacions. Perquè l'anàlisi tingui sentit, cal que el primer test sigui significatiu i el segon, en canvi, no ho sigui. Vegeu CUADRAS (1979).

Com que les mitjanes de les poblacions tenen fluctuacions degudes al mostratge, cal trobar una regió confidencial per la representació canònica de cada individu mitjà  $M_i$ . La regió confidencial és un cercle ( $d=2$ ) o una esfera ( $d=3$ ), de radi

$$\frac{R_\alpha}{\sqrt{n_i}} \text{ essent } R^2_\alpha = F_\alpha \frac{(N-k)}{(N-k-p+1)}$$

$F_\alpha$  és el valor de la  $F$  de SNEDECOR tal que  $P(F > F_\alpha) = \alpha$  amb  $p$  i  $(N-k-p+1)$  graus de llibertat. Aleshores el coeficient de confiança és  $1 - \alpha$ . Generalment es pren  $1 - \alpha = 0,9$ , o sigui,  $\alpha = 0,1$ .

*Exemple:* PETITPIERRE i CUADRAS (1977) fan una classificació sistemàtica de 32 poblacions (trobades en localitats geogràfiques diferents) de coleòpters del gènere *Timarcha*. Per diferenciar els individus mascles, es prenen 5 mesures biomètriques del pretòrax i dels èlitres, i 3 mesures del diàmetre dels tarses de I, II i III parells de potes. Per les femelles es prenen només les primeres mesures.

La fig. 11 és una representació canònica de 32 poblacions de mascles. la fig. 12 és una representació de 29 poblacions de femelles (no es disposava de femelles per a tres de les poblacions). Hi havia uns quaranta exemplars per població.

Les distàncies representades en dimensió 2 són projecció de les distàncies de Mahalanobis amb dimensió  $m=8$ . Observant les dues gràfi-

FIG. 11. — Representació canònica de 32 poblacions de coleòpters mascles del gènere *Timarcha*. Les regions confidencials dels individus mitjans estan construïdes al 90 %

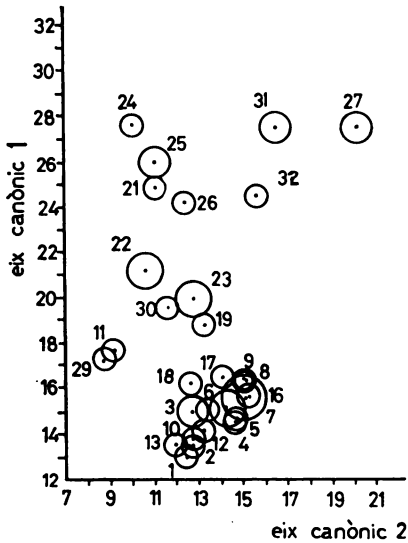
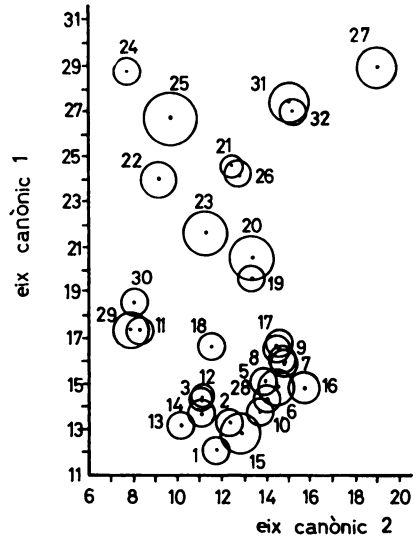


FIG. 12. — Representació canònica de 29 poblacions de coleòpters femelles del gènere *Timarcha*. Les regions confidencials pels individus mitjans estan construïdes al 90 %

ques, es pot veure que les poblacions 11, 20...32 es distingeixen com espècies diferenciades. En canvi les poblacions 1, 2, 10, 15,... formen un conglomerat que les fa interpretar com a races geogràfiques d'una mateixa espècie.

L'anàlisi canònica, en representar les espècies (o races, inventaris, etc.), amb referència a la distància de Mahalanobis, invariable per canvis d'escala i que té en compte les correlacions entre les variables, és la forma més objectiva i correcta de representar dades. Però, perquè es pugui aplicar, cal que tinguem una mostra d'individus representants de cada espècie.

#### ANÀLISI CANÒNICA GENERALITZADA

Podem mirar l'anàlisi canònica de poblacions com una representació gràfica dels nivells d'un disseny d'un sol factor. Quan les nostres dades depenguin de dos o més factors, haurem de realitzar tantes anàlisis canòniques dels nivells d'un factor determinat com nivells tenen els altres factors. En l'exemple de la secció anterior, hi ha dos factors: sexe i població, amb 2 i 29 nivells (ignorem ara les tres poblacions de mascles on no hi ha les corresponents femelles). Aleshores hem fet dues representacions canòniques: una per els mascles i l'altra per a les femelles. Però això ens condueix a algunes contradiccions. Per exemple, les poblacions 31 i 32 coincideixen en els mascles, però queden diferenciades en les femelles. Seria absurd distingir com a espècies diferents les femelles i com espècies iguals els mascles. Com es pot resoldre això?

Interpretem-ho com un problema d'anàlisi de la variància. La variabilitat total d'un disseny es descomposa en la variabilitat deguda als nivells de cada factor, la deguda a les interaccions i la variabilitat residual. Quan intervenen  $p > 1$  variables observades, aleshores hi ha una matriu de covariàncies total que es descomposa anàlogament amb suma de diverses matrius de covariàncies. Aleshores, si busquem els vectors propis de la matriu de covariàncies que dona la covariabilitat dels nivells d'un factor, respecte a la matriu de covariàncies residual (que és

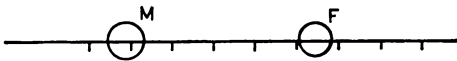


FIG. 13. — Representació canònica de mascles i femelles. Es pot interpretar com una estimació del grau de dimorfisme sexual

la que dona una estimació de la matriu de covariàncies  $C$  entre les variables), tindrem, per analogia amb l'anàlisi canònica de poblacions, una forma de representar els nivells d'un factor, eliminant la variabilitat deguda als altres factors i les interaccions. Àdhuc podrem representar també els altres factors.



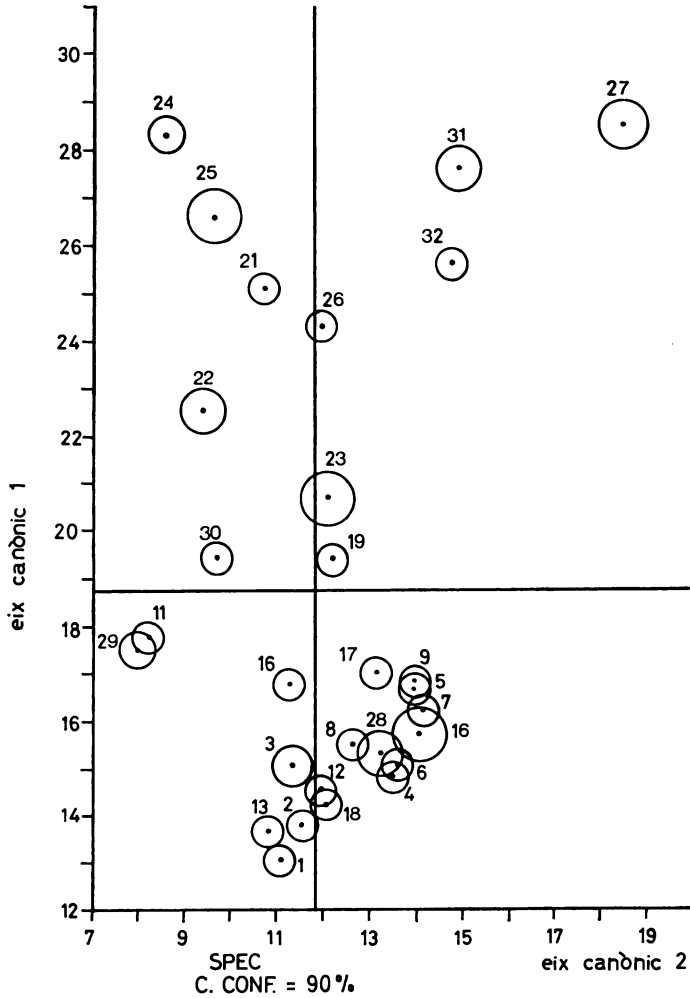


FIG. 14. — Representació canònica de 29 poblacions de coleòpters del gènere *Timarcha*. És una representació conjunta de mascles i femelles que elimina la covariabilitat deguda al dimorfisme sexual i a la interacció i resumeix en una sola les representacions fetes independentment per a mascles (Fig. 11) i femelles (Fig. 12)

El desenvolupament matemàtic d'aquesta generalització forma part de la teoria multivariant de funcions paramètriques estimables, i és massa complicat per poder-ho resumir aquí. Vegeu CUADRAS (1974).

Com una aplicació, vegem el resultat de representar les  $2 \times 29$  poblacions de mascles i femelles del gènere *Timarcha*. El primer factor,

o factor sexe, té 2 nivells. La fig. 13 és una representació canònica (unidimensional naturalment) dels mascles i de les femelles.

Aquesta figura es pot interpretar com una representació gràfica del grau de dimorfisme sexual. Les femelles serien, en general, més grans que els mascles, mentre que l'eix seria una dimensió de grandària.

La fig. 14 és una representació canònica dels 29 nivells dels segons factors, o sigui, de les 29 poblacions, havent eliminat la variabilitat deguda al sexe i a la interacció. Tenim així una representació única i objectiva de les 29 races geogràfiques o espècies. Així, les poblacions 31 i 32 deuen ésser considerades definitivament diferents.

Cal dir, també, que les distàncies representades són projecció d'una distància que és generalització de la distància de Mahalanobis. Té, doncs, les propietats esmentades d'invariabilitat per canvis d'escala en la mesura de les variables biomètriques, etc.

Altres aplicacions de l'anàlisi canònica generalitzada a la Biologia, i també a la Psicologia i a la Medicina, es poden veure a CUADRAS, PETITPIERRE i COLL (1977).

#### BIBLIOGRAFIA

1. ALONSO, G.: *Estudio de la distribución geográfica del polimorfismo cromosómico en Drosophila subobscura*. Tesina Fac. Biología. Univ. Barcelona. 275 pp. (1975).
2. CANTON, E. i SANCHO, J.: *Análisis Numérico de un grupo de Pseudomonas aeróbicos*. Microbiol. Española. 29, 59-73 (1976).
3. CHARDY, P.; GLEMAREC, M. i LAUREC, A.: *Application of inertia methods to benthic marine ecology. Practical Implications of the basic options*. Estuarine coastal marine science. 4: 179-205 (1976).
4. CUADRAS, C. M.: *Análisis discriminante de funciones paramétricas estimables*. Trab. Estad. Inv. Oper. 25: 3-31 (1974).
5. CUADRAS, C. M.: *Métodos de análisis multivariante*. Ed. Eunibar. Barcelona (en premsa) (1979).
- 5 bis. CUADRAS, C. M.; PETITPIERRE, E. i COLL, M. D.: *Generalized discriminant canonical analysis: Applications to Biology, Psychology and Medicine*. First World Conf. on Mat. at the Serv. of Man, Barcelona, Preprints (1977).
6. ESTRADA, M.: *Statistical consideration of some limnological parameters in Spanish reservoirs*. Ver. Internat. Verein. Limnol. 19: 1849-1859 (1975).
7. GOWER, J. C.: *Some distance properties of latent root and vector methods used in multivariate analysis*. Biometrika. 53: 325-338 (1966).
8. HOTTELLING, H.: *Analysis of a complex of statistical variables into principal components*. J. educ. Psychol. 24: 417-41 (1933).
9. JOHNSON, S. C.: *Hierarchical clustering schemes*. Psychometrika. 32: 241-254 (1967).
10. KRUSKAL, J. B.: *Non metric multidimensional scaling: a numerical method*. Psychometrika. 29: 115-129 (1964).
11. OCAÑA, J.: *Sobre la distancia genética*. Tesina Fac. Biología. Univ. Barcelona. 130 pp. (1975).

12. PETITPIERRE, E. i CUADRAS, C. M.: *The canonical analysis applied to the taxonomy and evolution of the genus Timarcha Latr. (Coleoptera, Chrysomelidae). Mediterranea.* 2: 13-28 (1977).
13. PREVOSTI, A.: *La distancia genética entre poblaciones.* Miscelánea Alcobé. 109-118 (1974).
14. PREVOSTI, A.; OCAÑA, J. i ALONSO, G.: *Distance between populations of Drosophila subobscura based on chromosome arrangement frequencies.* Theor. Appl. Genetics. 45: 231-241 (1975).
15. RAO, C. R.: *The use and interpretation of principal components analysis in applied research.* Sankhya A, 26: 328-358 (1964).
16. ROMERO, J.: *Análisis de componentes principales y una aplicación a la fitosociología.* Trabajo práctico de Bioestadística. Univ. Barcelona (inèdit) (1978).
17. ROHLF, F. J.: *Adaptative hierarchical clustering scheme.* Syst. Zool. 19 (1): 58-82 (1970).
18. SHEPARD, R. N.: *The analysis of proximities: multidimensional scaling with an unknown distance function. I.* Psychometrika. 27: 125-139 (1962 a).
19. SHEPARD, R. N.: *The analysis of proximities: multidimensional scaling with an unknown distance function. II.* Psychometrika. 27: 219-246 (1962 b).
20. SNEATH, P. H. A. i SOKAL, R. R.: *Numerical taxonomy.* W. H. Freeman Co. San Francisco. 573 pp. (1973).
21. SOKAL, R. R. i SNEATH, P. H. A.: *Principles of numerical taxonomy.* W. H. Freeman Co. San Francisco. 359 pp. (1963).

## DISCUSSIÓ

### FLÓS

Moltes vegades, quan es mesuren unes variables, no es té encara cap model fet. *A posteriori* se n'adopta un, però existeix una interacció contínua que porta a replantejar-lo sobre la marxa.

### J. WAGENSBERG

El problema és abans de l'estadística i després de la biologia. Consisteix en filtrar les dades en funció del model que es busca. S'han d'eliminar interaccions a fi que apareixin factors simples.

### FLÓS

Això pot ser molt difícil a la realitat. Pensem, per exemple, en el fitoplàncton. Tenim el moviment de l'aigua, l'absorció de nutrients, etc. Un model es complica de seguida. Molts paràmetres fisiològics no es coneixen i cal fer suposicions.

J. WAGENSBERG

Qualsevol fenomen té un grau de determinisme. Depèn del nivell que es consideri.

VALLESPINÓS

No cal ser nihilista. Sovint, el que es busca es només un model que doni resultats orientatius.

ALONSO

De fet, el model es tria quan es fa l'experiència. Per què no plantejar d'entrada el problema a un matemàtic?

MARGALEF

El nombre de dades no és mai prou important per a estimar les variables que entren en el model. El mostratge es pot simplificar, però el sentit biològic dels factors és sempre molt complicat. La tria dels paràmetres que es volen mesurar es fa en general per intuïció.

ESTRADA

D'altra banda, moltes vegades es mesuren paràmetres sense una adequada dimensió temporal. Per exemple, una mesura de la concentració de nutrients associada a una densitat de població de plàncton en un moment determinat, té poc valor per sí mateixa. És com un fòssil al qual cal buscar un context històric.

ESCARRÉ

És bo que les dades siguin cada vegada de tipus diferent. Així s'estableix un diàleg que obliga el matemàtic a espabilar-se.

CUADRAS

El matemàtic utilitza models que pugui elaborar. En FISHER, per exemple, havia d'imaginar quadrats llatins.

J. WAGENSBERG

No es pot admetre un model diferent cada vegada.

MARGALEF

El biòleg tendeix a uniformitzar la variabilitat. On és més alta pren més mostres. En oceanografia, per exemple, els punts de mostratge són més pròxims en sentit perpendicular a la costa. Falta una discussió crítica de com una teoria basada en la uniformitat pot servir de guia quan es tendeix a uniformitzar la variació.



## PARTICIPANTS

Són indicades les pàgines on comencen les aportacions originals (xifres en rodó) o les intervencions a les discussions (xifres en cursiva)

ALONSO, G. . . . .	64, 132
AZORÍN, F. . . . .	9
COMÍN, F. A. . . . .	21
CORNUDELLA, Ll. . . . .	11
CORTADAS, J. . . . .	35
CUADRAS, C. M. . . . .	64, 65, 71, 95, 133
ESCARRÉ, A. . . . .	132
ESTRADA, M. . . . .	43, 132
FERRER, X. . . . .	21
FLÓS, J. . . . .	70, 131
GUARDANS, R. . . . .	81
JORDAN, A. . . . .	37
LURIE, D. . . . .	79
MARGALEF, R. . . . .	17, 53, 70, 71, 78, 79, 132, 133
MENGOD, L. . . . .	11
OCAÑA, J. . . . .	57
PRAT, N. . . . .	27
RENZI, M. DE . . . . .	65, 69, 70, 71
ROCHA, E. . . . .	11
SERRA, J. . . . .	65
SUBIRANA, J. A. . . . .	9, 35, 37
VALLESPINÓS F. . . . .	79, 132
WAGENSBERG, J. . . . .	69, 73, 79, 131, 132, 133
ZEIGLER, B. . . . .	81





## T A U L A

<i>Difracció de raigs X de complexos cromàtics</i> , per F. AZORÍN i J. A. SUBIRANA . . . . .	9
<i>Estudi sobre l'organització de la cromatina i anàlisi del complement proteic nuclear durant la diferenciació cel·lular</i> , per E. ROCHA, L. MENGOD i LL. CORNUDELLA . . . . .	11
<i>Caràcter de les diferències entre el plàncton vegetal marí i el d'aigua dolça</i> , per R. MARGALEF . . . . .	17
<i>Presència de Compsopogon Montagne (Rhodophycophyta, Bangiophyceae) al delta de l'Ebre</i> , per X. FERRER i F. A. COMÍN . . . . .	21
<i>Composició i variació en el terme de dos anys d'una comunitat bentònica del riu Ter</i> , per N. PRAT . . . . .	27
<i>La desnaturalització parcial del DNA en N,N'dimetilformamida</i> , per J. CORTADAS i J. A. SUBIRANA . . . . .	35
<i>Trencament de protamines amb pepsina. Separació i caracterització de llurs pèptids</i> , per A. JORDAN i J. A. SUBIRANA . . . . .	37
<i>Relacions entre mitjanes i variàncies de mesures de biomassa fitoplanctònica</i> , per M. ESTRADA . . . . .	43
<i>Expressió senzilla de la regulació en sistemes dinàmics</i> , per R. MARGALEF . . . . .	53
<i>Aspectes de la problemàtica de la distància taxonòmica</i> , per J. OCAÑA . . . . .	57
<i>Anàlisi de perfils</i> , per M. DE RENZI, J. SERRA KIEL i C. M. CUADRAS . . . . .	65
<i>Deterministic model for the evolution of a population in a limited environment</i> , per J. WAGENSBERG . . . . .	73
<i>Una manera discreta d'observar un sistema depredador-presa</i> , per B. ZEIGLER i R. GUARDANS . . . . .	81
<i>Mètodes de representació de dades i la seva aplicació a la Biologia</i> , per C. M. CUADRAS . . . . .	95



ACABAT D'IMPRIMIR ALS  
TALLERS GRÀFICS A. NÚÑEZ  
DE BARCELONA  
EL DIA 28 DE MARÇ DEL 1980  
GRAVATS DE J. M. LLOVET





