

Zeitschrift: Archives des sciences [1948-1980]
Herausgeber: Société de Physique et d'Histoire Naturelle de Genève
Band: 26 (1973)
Heft: 1

Artikel: L'analyse de groupe en paléoécologie et en biostratigraphie : application aux ostracodes (Crustacea) miocènes
Autor: Carbonnel, G.
DOI: <https://doi.org/10.5169/seals-739916>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 16.02.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

L'ANALYSE DE GROUPE EN PALÉOÉCOLOGIE ET EN BIOSTRATIGRAPHIE

Application aux Ostracodes (*Crustacea*) miocènes

PAR

G. CARBONNEL ¹

TABLE DES MATIÈRES

<i>Résumé, abstract</i>	25
 <i>PREMIÈRE PARTIE</i>	
INTRODUCTION	27
1. DÉFINITION DE L'ANALYSE DE GROUPE	27
2. LE TAUX DE LIAISON ET SON ÉVALUATION QUANTITATIVE	28
2.1. Les coefficients d'association	28
2.1.1. Coefficient de Jaccard (J)	28
2.1.2. Autres coefficients	29
2.1.3. Coefficient de corrélation (r)	29
2.2. Avantages et inconvénients de l'emploi de ces coefficients	30
2.3. Résultats — Matrice de similarité	31
2.3.1. Matrice de similarité selon le mode R	34
2.3.2. Matrice de similarité selon le mode Q	34
3. L'ANALYSE DE GROUPE	35
3.1. Groupement par la moyenne des taux de liaison	36
Somme des variables de Spearman	36
3.2. Application pratique	36
3.2.1. 1 ^{er} cycle de calcul	36
3.2.2. 2 ^e cycle de calcul	37
3.3. Limite de la méthode	39
3.3.1. Dépassement du taux d'association d'un groupe sur l'autre (WPGM)	39
3.3.2. Méthode de groupement par paire pondérée (WPGM) et non pondérée (UPGM)	39

¹ Université LYON 1 - Claude Bernard — Département des Sciences de la Terre et Laboratoire associé au CNRS, 15 et 43 boulevard du 11-novembre-1918 — 69-VILLEURBANNE.

4. TRADUCTION GRAPHIQUE DES RÉSULTATS DE L'ANALYSE DE GROUPE: LE DENDOGRAMME	
4.1. Réalisation du dendogramme	40
4.2. Comparaison des dendogrammes entre eux — Méthode cophénétiq.	40
4.2.1. Phénons	41
4.2.2. Matrice cophénétiq.	41

DEUXIÈME PARTIE

1. LES CORRÉLATIONS PAR L'ANALYSE DE GROUPE SELON LE MODE Q — Application à l'étude des niveaux de prélèvements	43
1.1. Groupement des niveaux d'une coupe en formations	44
1.2. Groupement des niveaux et la notion d'instabilité paléoécologique (en lithofaciès constant)	45
1.3. Groupement des niveaux d'une coupe en lithofaciès variés — Influence du lithofaciès	47
1.4. Groupement des niveaux entre plusieurs coupes	49
1.4.1. Dans des bassins différents	49
1.4.2. Dans un même bassin	53
1.5. Groupement entre plusieurs coupes	55
1.5.1. Analyse écologique	55
1.5.2. Analyse biostratigraphique	56
2. LES CORRÉLATIONS PAR L'ANALYSE DE GROUPE SELON LE MODE R — Application à l'étude des associations d'espèces	56
2.1. Association(s) d'espèces dans une coupe	56
2.1.1. Relation entre les groupements de niveaux et les associations d'espèces	58
2.1.2. Stabilité des associations sur une coupe	58
2.1.3. Notion d'espèces dominantes et accessoires	60
2.1.4. Valeur du coefficient de corrélation cophénétiq. entre la matrice du coefficient de corrélation et celle de Jaccard	61
2.2. Association(s) d'espèces et les paramètres paléoécologiques	61
2.2.1. Liaison entre la salinité et les morphotypes de <i>Limnocythere pernesensis</i> CARBONNEL, 1972	61
3. CONCLUSIONS	64
3.1. Résultats généraux — Valeur de ces méthodes	64
3.1.1. Analyse de groupe selon le mode Q — Corrélation des niveaux	64
3.1.2. Analyse de groupe selon le mode R — Corrélation des espèces, biocénoses	65
3.2. Résultats locaux	65
3.2.1. Résultats stratigraphiques	65
3.2.2. Résultats paléogéographiques	66
3.3.3. Résultats paléoécologiques	66
BIBLIOGRAPHIE	66

RÉSUMÉ

Pour la première fois en langue française, les méthodes de calcul conduisant à la réalisation des dendogrammes sont longuement exposées dans la première partie.

Elles traitent :

- de l'obtention des matrices de similarité à partir de plusieurs coefficients (Jaccard, corrélation...)
- des méthodes de calcul propres à l'analyse de groupe (somme des variables de Spearman, variables pondérées ou WPGM et non pondérées ou UPGM)
- de la comparaison des dendogrammes entre eux par la méthode cophénétique.

Dans la seconde partie, on applique ces méthodes à partir de microfossiles miocènes (ostracodes) dans un but paléoécologique et biostratigraphique.

Elles débouchent sur les notions de stabilité et d'instabilité paléoécologiques et leur corollaire: les biocoénoses et thanathocoénoses.

Elles permettent également l'« établissement mathématique » des formations géologiques et des corrélations biostratigraphiques, à partir des dendogrammes des niveaux.

Enfin elles autorisent à tester l'influence des paramètres paléoécologiques, la salinité par exemple, sur les associations fauniques.

ABSTRACT

For the first time in French language, the methods of reckoning aiming at the realisation of dendograms are longly expounded in the first part.

They are dealing with :

- the obtention from several coefficients (Jaccard, correlation) of matrixes of similitude;
- the very reckoning methods for cluster analysis (sum of the variables of Spearman, weighted variables or WPGM and unweighted or UPGM)
- the comparison of the different dendograms thanks to the cophenetic method.

In the second part, these methods are applied to miocene microfossiles (ostracoda) with a paleoecological and biostratigraphical view.

They open on notions of equilibrium and unequilibrium in paleoecology and on their corolaries: biocenosis and thanathocenosis.

They also allow "the mathematical establishment" of geological formations and biostratigraphical correlations, from levels dendograms.

At last, they let test the influence of paleoecological parameters, salinity on faunic associations for instance.

PREMIÈRE PARTIE

INTRODUCTION

L'analyse de groupe (ou cluster analysis) constitue une méthode couramment utilisée dans diverses disciplines (en Microbiologie, en Botanique et en Entomologie) depuis les travaux fondamentaux de P. H. A. SNEATH et S. T. COWAN (1958), C. D. MICHENER et R. R. SOKAL (1957), etc... (voir la liste dans R. R. SOKAL et P. H. A. SNEATH 1963, pp. 259-260).

En Paléontologie, elle a été introduite plus récemment dans un but taxinomique (par exemple S. BRETSKY en 1970 et 1971). Son emploi s'est également développé en Micropaléontologie durant ces dix dernières années. R. L. KAESLER et R. F. MADDOCKS l'ont étendue aux ostracodes actuels en 1966. Le but de ce travail n'est pas de relater l'historique de ces méthodes, encore moins d'en envisager l'étude mathématique critique. Il est plutôt de tenter d'apporter un document en langue française, inexistant jusqu'alors, permettant d'appliquer l'analyse de groupe à la géologie stratigraphique.

Il s'agit seulement d'exposer ici les méthodes employées par l'auteur et leur application aux ostracodes miocènes dans un but stratigraphique et paléoécologique, à l'exclusion de toute utilisation taxinomique. On se reportera pour l'extension de l'analyse de groupe à la taxinomie numérique à R. R. SOKAL et P. H. A. SNEATH (1963) pour une étude générale et à R. L. KAESLER (1969) pour une application aux ostracodes.

1. DÉFINITION DE L'ANALYSE DE GROUPE

C'est une définition limitée au but poursuivi: stratigraphique et paléoécologique.

On soulignera d'abord qu'il s'agit d'une méthode exclusivement quantitative et de ce fait répliquable. C'est là qu'en réside l'intérêt. Dans l'orientation envisagée, elle quantifie, par des moyens appropriés, les liaisons existantes entre les espèces et (ou) les niveaux qui les contiennent. Ces moyens appropriés sont:

- le taux de liaison par emploi des coefficients de similarité;
- la moyenne des taux de liaison aboutissant au dendogramme;
- la comparaison des dendogrammes par la méthode cophénétique.

2. LE TAUX DE LIAISON ET SON ÉVALUATION QUANTITATIVE

Les différentes techniques pour évaluer la ressemblance sont représentées par trois types de coefficients:

- les coefficients d'association;
- le coefficient de corrélation;
- la distance (principalement réservée aux travaux de taxinomie numérique).

Seuls les deux premiers ont été utilisés ici. On consultera pour le troisième R. R. SOKAL et P. H. A. SNEATH (1963, pp. 143-153 et 349) et F. J. ROHLF ET et R. R. SOKAL (1965).

2.1. LES COEFFICIENTS D'ASSOCIATION

Ils sont multiples (*cf.* A. H. CHEETHAM et J. E. HAZEL, 1969). Un seul a été sélectionné: le coefficient de Jaccard. C'est le plus connu; il a été fréquemment utilisé dans les études antérieures.

2.1.1. Coefficient de Jaccard (*J*)

Sa formule s'écrit:
$$J = \frac{C}{N_1 + N_2 - C} \quad \text{où}$$

C = nombre d'espèces communes entre les deux échantillons (par exemple les niveaux de prélèvements);

N_1 = nombre d'espèces présentes dans l'échantillon n° 1;

N_2 = nombre d'espèces présentes dans l'échantillon n° 2.

Le nombre d'individus de chaque espèce n'est pas pris en considération. La valeur du coefficient varie de 1 à 0 suivant que toutes les espèces sont communes ou bien qu'aucune ne l'est. Elle est obtenue à partir des tableaux « classiques » de répartition verticale de la faune comme le tableau 1. On se reportera à G. CARBONNEL (1969, pp. 387 et suivantes) pour les détails pratiques de l'obtention du coefficient de Jaccard, tant manuel que sur ordinateur. Toutes les indications concernant la transcription des données brutes sur bordereau de transfert et la traduction sur carte perforée sont contenues dans cette étude. Le programme permettant son calcul a été publié pages 392-393 du même auteur et ne sera pas repris ici.

2.1.2. *Autres coefficients*

D'autres auteurs, en particulier, R. L. KAESLER (1966), A. H. CHEETHAM et J. E. HAZEL (1969) ont expérimenté quelques autres coefficients et étudié leurs limites mathématiques.

2.1.2.1. *Coefficient de Simpson et coefficients proches*

Jusqu'à présent, leur emploi est plus rare que celui du coefficient de Jaccard.

— coefficient de Simpson

sa formule s'écrit:
$$\frac{C}{N1}$$

(C représente le nombre d'événements communs entre les deux échantillons, $N1$ étant le plus petit des deux).

— coefficient de Dice

sa formule s'écrit:
$$\frac{2C}{N1 + N2}$$

(indices identiques à ceux de Jaccard, voir 2.1.1.)

— coefficient d'Otsuka

sa formule s'écrit:
$$\frac{C}{\sqrt{N1 N2}}$$

(indices identiques à ceux de Jaccard)

Ces deux coefficients sont à la fois très proches l'un de l'autre mais également du coefficient de Simpson. Ce dernier constitue en effet leur limite lorsque $N1 = N2$

2.1.2.2. *Simple Matching coefficient (SM)*

Sa formule s'écrit:
$$\frac{C + A}{(N1 + N2 - C) + A}$$

Les indices sont identiques à ceux de Jaccard à l'exclusion de A . Ce dernier représente le nombre de fois où l'évènement est absent dans les 2 échantillons comparés. Les inconvénients de son emploi sont analysés plus loin (voir 2.2.1.).

2.1.3. *Le coefficient de corrélation (r)*

sa formule s'écrit:

$$r = \frac{\sum_{i=1}^n (xi - \bar{x})(yi - \bar{y})}{\left[\sum_{i=1}^n (xi - \bar{x})^2 \sum_{i=1}^n (yi - \bar{y})^2 \right]^{1/2}}$$

A l'inverse du précédent il est basé sur le nombre réel d'individus de chaque espèce, dans chaque niveau ou prélèvement.

C'est une expression où :

x_i = le nombre d'individus de l'espèce 1 pour les niveaux de prélèvement étudié

y_i = le nombre d'individus de l'espèce 2 pour les niveaux de prélèvement étudié

\bar{x} = la moyenne du nombre d'individus de l'espèce 1 sur l'ensemble des niveaux de prélèvements

\bar{y} = la moyenne du nombre d'individus de l'espèce 2 sur l'ensemble des niveaux de prélèvements.

Sa valeur varie de -1 à $+1$. Lorsque $r = \pm 1$, les variables (donc les espèces) sont en dépendance fonctionnelle linéaire. Lorsque $r = 0$ les deux espèces sont indépendantes.

Pour connaître un moyen de calcul rapide et manuel de r on consultera, entre autres, R. R. SOKAL et *al.*, (1963, p. 297). Le programme permettant le calcul du coefficient de corrélation sur ordinateur n'a pas été publié. Il fera l'objet, ultérieurement, d'une publication, comme les autres programmes nécessaires à cette étude.

L'emploi du coefficient de corrélation a été prohibé par R. L. KAESLER (1966) dans le cadre de ses études écologiques. Un essai a néanmoins été tenté par M. JACOBZONE et G. CARBONNEL en 1971 dans un but paléoécologique. L'essai est renouvelé ici dans la même direction.

2.2. AVANTAGES ET INCONVÉNIENTS DE L'EMPLOI DE CES COEFFICIENTS

2.2.1. Coefficient de Jaccard

Son calcul envisage la présence ou l'absence des espèces, mais il ne retient que leur présence. Le nombre de fois où les deux espèces sont simultanément absentes n'est pas introduit dans sa formule comme c'est le cas pour d'autres coefficients de similitude (p. e. Simple Matching coefficient).

Comme l'a souligné en écologie R. L. KAESLER en 1966 et comme l'a repris en paléoécologie G. CARBONNEL en 1969, l'absence de deux espèces peut résulter de nombreux facteurs tels que la conservation des espèces, le niveau d'énergie des eaux, les conditions de fossilisation, l'altération secondaire. Ceux-ci n'ont rien de commun avec l'impossibilité biologique pour deux espèces de coexister en un lieu donné ce qui se traduirait alors par une absence significative.

2.2.2. Coefficient de corrélation

R. L. KAESLER (1966) estime que l'emploi du coefficient de corrélation est dépourvu de signification dans les études écologiques. Dans ce cas, on utilise selon

cet auteur le nombre total d'individus mais comme les prélèvements sont quantitativement variés on introduit ainsi des erreurs.

Cette position, d'un autre côté, tend à faire oublier que l'abondance d'une ou de plusieurs espèces au sein des niveaux de prélèvements est le reflet de la prédominance de ces espèces sur les autres dans leur biotope originel. Ces notions, essentielles en écologie comme en paléoécologie traduisent la réalité des espèces dominantes et des espèces accessoires.

Un essai, tenté plus loin (2^e partie 2.1.3.), montrera que l'on aboutit à des résultats contradictoires après emploi des coefficients de Jaccard et de corrélation sur un même exemple. Il faut se garder cependant de préjuger de la valeur relative de chacun.

2.3. RÉSULTATS — MATRICE DE SIMILARITÉ

Une fois les calculs effectués par utilisation du coefficient de Jaccard ou du coefficient de corrélation on obtient un ensemble de données « semi-élaborées » qui traduisent la valeur de la liaison existant entre les espèces des niveaux étudiés ou entre les niveaux renfermant ces mêmes espèces.

En effet, le tableau de répartition des données brutes (p.e. tableau de répartition verticale des espèces en fonction des niveaux, tabl. 1) peut être considéré comme une matrice de forme $n \times t$ où :

t = colonnes représentant les prélèvements

n = lignes représentant les espèces.

Dans un problème de ce genre on a l'habitude depuis R. R. SOKAL ET P. H. A. SNEATH (1963) de désigner les espèces sous le terme de: **Operational Taxonomic Units** ou **OTU**.

Une matrice de ce type peut être examinée selon les lignes, où les colonnes. On peut chercher les liaisons existantes entre les lignes c'est-à-dire entre les OTU contenues dans les différents prélèvements ou bien entre les colonnes, c'est-à-dire entre les prélèvements renfermant ces espèces (voir fig. 1).

	Prél.					Mode Q
	OTU	A	B	C N	
Mode R	1					
	2					
	3					
	•					
	•					
	n					

FIG. 1. — Matrice théorique des données brutes.

FORMATION PARARÉCIFALE DU CAP DES NAUTES (N)		FORMATION SAUMÂTRE DU ROUSSET (R)		FORMATION BIOCLASTIQUE DE CARRY (C)		FORMATION BIODÉTRITIQUE DE SAUSSET-LES-PINS (S)					FORMATIONS D'APRÈS P. ANDREIFF <i>et al.</i> 1971			
													N° des prélèvements	
18620	□													<i>Procythereis cf. deformis</i>
18581	•													<i>Loxoconcha hastata</i>
18585	•	•												<i>Schuleridea cf. oculata</i>
18588	•	•	•											<i>Hermanites gr. haidingeri</i>
18589	•	•	•											<i>Hemicyprideis helvetica nertheensis</i>
														<i>Cyamocytheridea cf. leptostigma</i>
														<i>Cytheridea acuminata caumontensis</i>
														<i>Costa tricosata subsp. I</i>
														<i>Hemicyprideis dacica grekoffi</i>
														<i>Loxoconcha sp.</i>
														<i>Callistocythere savoyonnei</i>
														<i>Miocyprideis aff. rara rara</i>
														<i>Krithe cf. "bartonensis"</i>
														<i>Hemicyprideis helvetica nertheensis</i> <i>tendance grekoffi</i>
														<i>Cytheretta sp. (ornée)</i>
														<i>Cytheretta cf. ovata</i>
														<i>Falunia sphaerulolineata</i>
														<i>Cytheretta cf. accedens</i>
														<i>Propontocypris ? solitaria</i>
														<i>Eocytheropteron bruggenense</i>
														<i>Hemicytherideis sp.</i>
														<i>Quadracythere cf. nodoreticulata</i>
1	2	1	2	3					FORMATIONS d'après l'analyse de groupe (voir fig. 9)					

• 1-2 valves ; □ 3-4 valves ; ▣ 5-10 valves ; ■ plus de 10 valves

TABLEAU I — Tableau de répartition verticale des ostracodes dans une coupe destiné à montrer les relations entre les formations obtenues avec ou sans analyse de groupe.

Une étude de liaison entre les lignes est appelée, de façon conventionnelle, analyse selon le mode **R**.

Une étude de liaison entre les colonnes est appelée, de façon conventionnelle, analyse selon le mode **Q**.

2.3.1. Matrice de similarité selon le mode *R* (fig. 2)

Elle est l'expression de la liaison 2 à 2 entre les espèces, par exemple entre l'espèce 1, 2 ou 1, 3 jusqu'à $n-1, n$. La traduction des données doit se faire sous la forme d'une matrice de similarité de dimensions $n \cdot n$. Dans le cas du coefficient de Jaccard la valeur de la liaison entre les espèces 1 et 2 est $J_{1,2}$. Sur la diagonale, la valeur de la liaison est la valeur de la liaison de l'espèce sur elle-même soit 1; de ce fait elle est négligée.

espèces	1	2	3	n
1	X	$J_{1,2}$	$J_{1,3}$		$J_{1,n}$
2	$J_{2,1}$	X			
3	$J_{3,1}$				
:					
:					
:					
n	$J_{n,1}$				

FIG. 2. — Matrice théorique de similarité pour les espèces (coefficient de Jaccard).

On remarque que la matrice de similarité est symétrique par rapport à la diagonale.

2.3.2. Matrice de similarité selon le mode *Q* (fig. 3)

Elle est l'expression de la liaison 2 à 2 entre les niveaux par exemple entre les niveaux A, B ou A, C jusqu'à $N-1, N$.

On obtient une matrice de similarité comparable à la précédente avec pour valeurs (cas du coefficient de Jaccard) les suivantes:

niveaux	A	B	C	N
A	X	$J_{A,B}$	$J_{A,C}$		$J_{A,N}$
B	$J_{B,A}$	X			
C	$J_{C,A}$				
.					
.					
.					
N	$J_{N,A}$				

FIG. 3. — Matrice théorique de similarité pour les niveaux (coefficient de Jaccard).

Rappelons que les valeurs réelles des éléments de ces matrices sont comprises entre 0 et 1 si l'on emploie le coefficient de Jaccard (J) et entre -1 et $+1$ dans le cas du coefficient de corrélation (r).

3. L'ANALYSE DE GROUPE

Après l'obtention de la matrice de similarité (fig. 2 ou 3), interviennent les méthodes de l'analyse de groupe proprement dite. Jusqu'à présent on envisageait seulement la liaison 2 à 2 entre les éléments de la matrice de similarité. Reprenons l'exemple de la figure 2; on connaît la valeur du taux de liaison entre les espèces 1 et 2 ($J_{1,2}$) mais on ignore celle entre le « groupe » [1, 2] et l'espèce 3. Le but de l'analyse de groupe est de combler cette lacune.

Différentes méthodes d'analyse de groupe ont été préconisées par les auteurs. On note parmi les plus employées :

— le groupement élémentaire

Chaque terme de la matrice de similarité (fig. 2 et 3) est placé suivant sa valeur dans une classe arbitrairement définie. Il est représenté par un figuré conventionnel (cf. CARBONNEL, 1969, p. 413 et 416). C'est le groupement le moins élaboré et le plus immédiat.

— le groupement par association simple

Les termes de la matrice de similarité sont associés à partir de la valeur la plus élevée que l'on observe. On l'abaisse successivement et régulièrement (cf. R. R. SOKAL et *al.*, 1963, p. 180) pour permettre à de nouveaux termes de se joindre au groupe déjà formé.

— le groupement par association complète — Méthode de Sørensen

Un terme de la matrice de similarité ne s'associe à 2 ou plusieurs autres termes de cette matrice que si la valeur de la liaison qui relie *chacun des termes* 2 à 2 dans le nouveau groupe est supérieur à la valeur retenue (cf. R. R. SOKAL et *al.*, 1963, p. 181).

— le groupement par la moyenne des taux de liaison (méthode de Sokal et Minchener *ibid.*, p. 182 et suiv.).

Ce dernier mode de groupement est le plus couramment utilisé. C'est également celui qui rend le mieux compte des phénomènes étudiés selon R. L. KAESLER (1967). Ces raisons nous conduisent à exposer les détails de la méthode d'autant qu'aucune étude en langue française ne l'a jamais réalisée jusqu'à présent.

3.1. GROUPEMENT PAR LA MOYENNE DES TAUX DE LIAISON. SOMME DES VARIABLES DE SPEARMAN

Nous emprunterons à R. R. SOKAL et P. H. A. SNEATH (1963, p. 305 et suiv.) la procédure de calcul.

Ce groupement base l'admission de chaque unité (OTU ou niveaux) dans un groupe sur la moyenne de similarité de cette unité avec le groupe. La formule générale employée est celle de Spearman. Elle a pour expression :

$$rqQ = \frac{\square qQ}{\sqrt{q + 2\Delta q} \sqrt{Q + 2\Delta Q}}$$

où $\square qQ$ = somme de toutes les corrélations entre les membres d'un groupe avec ceux de l'autre groupe.

Δq = somme de toutes les corrélations entre les membres du 1^{er} groupe

ΔQ = somme de toutes les corrélations entre les membres du 2^e groupe

q = nombre d'unités du 1^{er} groupe

Q = nombre d'unités du 2^e groupe

Si q ou $Q = 1$ (c'est-à-dire ne contient qu'une unité) la formule devient :

$$rqQ = \frac{\Sigma r x q'}{\sqrt{q' + 2\Delta q'}}$$

où $\Sigma r x q'$ = somme de toutes les corrélations entre les membres du groupe et l'unité

$\Delta q'$ = somme de toutes les corrélations entre les membres du groupe.

q' = nombre d'unités du groupe.

3.2. APPLICATION PRATIQUE

3.2.1. 1^{er} cycle de calcul

Soit une matrice¹ de similarité entre les espèces 1 à 6 résultant du coefficient de Jaccard.

¹ Valeurs réelles extraites de M. Jacobzone et G. Carbonnel, 1971.

espèces	1	2	3	4	5	6
1	<i>X</i>	0,41	<u>0,58</u>	<u>1</u>	<u>0,81</u>	<u>0,65</u>
2	0,41	<i>X</i>	<u>0,55</u>	0,41	<u>0,25</u>	<u>0,23</u>
3	0,58	<u>0,55</u>	<i>X</i>	0,58	0,54	0,50
4	<u>1</u>	0,41	<u>0,58</u>	<i>X</i>	<u>0,81</u>	0,61
5	<u>0,81</u>	0,25	<u>0,54</u>	0,81	<i>X</i>	0,58
6	0,65	0,23	0,50	0,61	0,58	<i>X</i>

FIG. 4. — 1^{re} matrice de similarité (coefficient de Jaccard) résultant du 1^{er} cycle de l'analyse de groupe.

Les différents termes de la procédure à employer sont :

- a) rechercher dans chaque colonne de la matrice la plus grande valeur de similarité (valeur soulignée dans la fig. 4)
- b) contrôler que la plus grande valeur observée dans une colonne se retrouve symétriquement par rapport à la diagonale dans la ligne correspondante; dans l'affirmative, on retient cette valeur; dans la négative, on la rejette.

1^{er} cas: colonne 1, ligne 4: valeur 1 } à retenir
 colonne 4, ligne 1: valeur 1 }

2^e cas: colonne 3, ligne 1: valeur 0,58 } à rejeter
 colonne 1, ligne 4: valeur 1 }

- c) on considère que les unités [1-4], en l'occurrence des espèces, forme un groupe associé suivant la valeur 1 et on rejette le « groupe » 1,3.
- d) on réalise les opérations *a*, *b*, *c*, pour la totalité de la matrice de similarité de la figure 4. On aboutit, dans ce cas, à l'obtention d'un seul groupe d'espèces associées [1-4], les autres espèces (2, 3, 5, 6) ne l'étant ni entre elles ni avec le groupe [1-4].
- e) c'est à ce moment qu'il convient de calculer qu'elle sera la valeur de la liaison existant entre le groupe [1-4] et les espèces 2, 3, 5, 6. On emploie le coefficient de Spearman défini plus haut.

3.2.2. 2^e cycle de calcul. Obtention de la seconde matrice de similarité par emploi du coefficient de Spearman

Dans l'exemple choisi, $q = 1$, la formule qui s'applique est donc :

$$rqQ = \frac{\Sigma r x q'}{\sqrt{q' + 2\Delta q'}}$$

On calculera successivement la liaison existante entre :

[1-4] et 2

[1-4] et 3

[1-4] et 5

[1-4] et 6

La nouvelle valeur du taux liant le groupe [1-4] à 2 est la somme des liaisons de chacun des termes soit en décomposant :

$$1-2 = 0,41 \text{ valeur lue dans la matrice précédente (fig. 4)}$$

$$4-2 = 0,41 \quad \text{id.}$$

$$\text{d'où } \Sigma rxq' = 0,41 + 0,41 = 0,82$$

$$q' = 2 \text{ (2 unités dans le groupe)}$$

$$\Delta q' = 1 \text{ valeur de la liaison du groupe [1-4] lue dans la matrice de la figure 4}$$

$$\sqrt{q' + 2\Delta q'} = \sqrt{2 + 2}$$

$$rqQ = \frac{0,82}{\sqrt{4}} = 0,41$$

On renouvelle ces opérations pour chacun des autres termes de la matrice: 3, 5 et 6. La nouvelle matrice obtenue a pour éléments ¹:

	<i>1'</i>	2	3	5	6
<i>1'</i>	X	0,41	0,58	0,81	0,61
2	0,41	X	0,55	0,25	0,23
3	0,58	0,55	X	0,54	0,54
5	0,81	0,25	0,54	X	0,58
6	0,61	0,23	0,54	0,58	X

FIG. 5. — 2^e matrice de similarité résultant du 2^e cycle de l'analyse de groupe.

On recommence alors « l'extraction » des groupes suivant la méthode indiquée au paragraphe 3.2.1. Elle conduit à l'obtention du groupe [1'-5]; les autres espèces 2, 3, 6, étant associées ni entre elles, ni avec le groupe [1'-5].

Rappelons la signification du groupe 1'-5. Le groupe [1'] résulte de l'association des espèces [1-4] (matrice de la fig. 4). Il est associé avec l'espèce 5 pour former le groupe [1'-5] de préférence à toute autre possibilité d'association parmi les espèces étudiées.

¹ Pour plus de commodité on appellera le groupe [1-4]: [1']; on réalisera la même simplification pour les autres groupes obtenus par la suite.

On continuera ainsi, jusqu'à épuisement des termes de la matrice de départ, à réaliser alternativement les opérations du 1^{er} cycle et du second cycle. Elles aboutiront à révéler l'existence des groupes $[1-4] = [1']$ avec une valeur de liaison égale à 1, $[1'-5] = [1'']$ avec une valeur de liaison égale à 0,81, $[1''-6] = [1''']$ avec une valeur de liaison égale à 0,63, etc.

On remarque que l'admission d'une nouvelle unité dans un groupe se produit, comme il est logique, par un abaissement successif du taux de liaison.

Ces calculs nécessitent, dès que les éléments de la première matrice de similarité sont supérieurs à 10, un ordinateur avec une mémoire de capacité élevée (CDC 3600, IRIS 80). Le programme de l'analyse de groupe par cette méthode sera publiée ultérieurement.

3.3. LIMITES DE LA MÉTHODE

Plusieurs imperfections limitent son emploi:

3.3.1. *Dépassement du taux d'association d'un groupe sur l'autre.*

Pour des raisons mal expliquées (l'étude mathématique complète de la fonction n'ayant pas encore été réalisée à ma connaissance), le taux de liaison d'un groupe peut prendre une valeur supérieure à celle qu'il devrait avoir: c'est l'« overlap » des auteurs anglo-saxons.

Ainsi, dans l'exemple du paragraphe 3.2. la liaison du groupe $[1''-6]$ pourrait s'effectuer avec une valeur légèrement supérieure à 0,81 alors quelle devrait être inférieure. Selon R. R. SOKAL et P. H. A. SNEATH (1963, p. 184), cela se réalise lorsque Δq et (ou) ΔQ sont très petits ou négatifs. Ce dépassement a été fréquemment observé dans l'application pratique (p.e. fig. 18).

3.3.2. *Méthodes de groupement par paire pondérée (WPGM) et non pondérée (UPGM)*

Elles sont plus connues sous le nom de WPGM (Weighted Pair Group Method) et UPGM (Unweighted Pair Group Method). La méthode WPGM est celle qui vient d'être exposée. La méthode UPGM utilise également le coefficient de Spearman. La différence entre les deux réside dans la matrice intermédiaire, choisie comme référence pour calculer les valeurs des taux d'association des nouveaux groupes lors du 2^e cycle.

Reprenons l'exemple précédent. Soit le groupe $[1''-6]$ résultant du groupement $[1'-5]$ avec 6. Dans la méthode WPGM, la valeur du taux de liaison $[1''-6]$ correspond à la valeur calculée $[1'-5]$ et 6 à partir de la matrice de la figure 5, (= 0,63).

Dans la méthode UPGM on retourne à la matrice initiale de la figure 4 pour calculer la valeur $1''-6$. On se rappelle que $[1'] = [1-4]$ et $[1''] = [1'-5]$.

Il faudra donc calculer l'expression suivante:

$$(1-4-5) 6 \text{ soit } 1-6, 4-6, 5-6.$$

Compte tenu des difficultés de programmation par rapport à l'autre méthode, cette variante n'a jamais été employée, bien que R. L. Kaesler la préconise dans ses divers travaux.

4. TRADUCTION GRAPHIQUE DES RÉSULTATS DE L'ANALYSE DE GROUPE: LE DENDOGRAMME.

4.1. RÉALISATION DU DENDOGRAMME

On porte en abscisse la valeur du taux d'association compris entre 0 et 1 ou -1 et $+1$ suivant le coefficient choisi. On inscrit en ordonné les groupes obtenus. L'écart entre les membres d'un groupe est arbitraire. On relie par un trait vertical, à l'aplomb de la valeur avec laquelle ils sont associés, les membres du groupe (fig. 6).

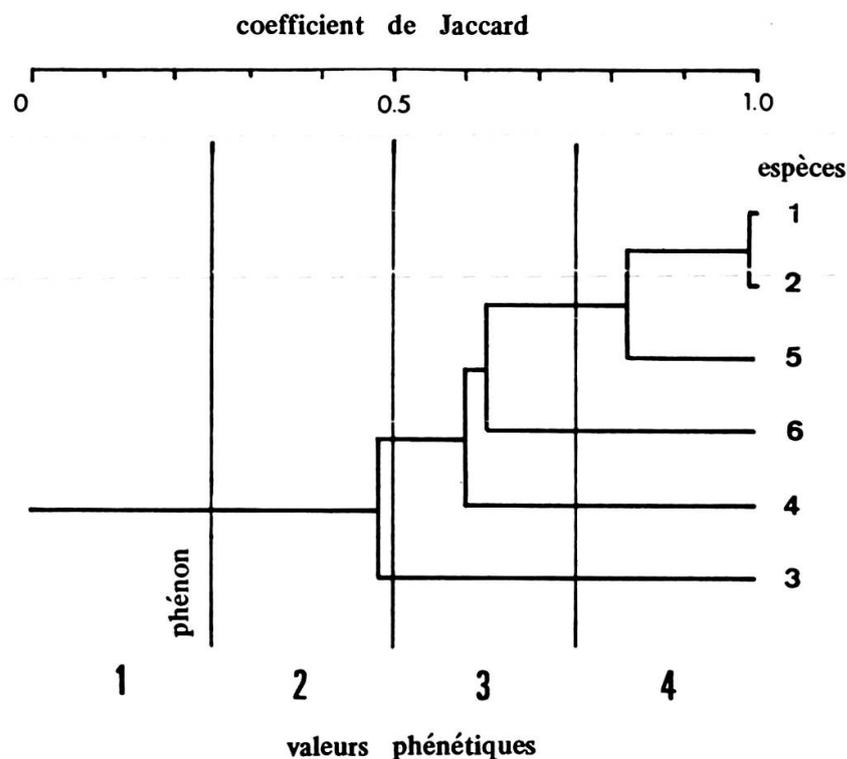


FIG. 6. — Dendrogramme obtenu à partir de la matrice du coefficient de Jaccard traité par la méthode WPGM et ses phénons.

La figure obtenue est appelée dendogramme. Il traduit en plan les liaisons multidimensionnelles existant entre les espèces ou les niveaux suivant le mode R ou Q. Il permet ainsi d'apprécier la valeur au-dessus de laquelle se groupent tels éléments entre eux. Lorsque le taux de liaison d'un membre a une valeur supérieure à celle du groupe auquel il doit se relier on le traduit conventionnellement par un trait interrompu (p.e. fig. 11).

4.2. COMPARAISON DES DENDOGRAMMES ENTRE EUX: MÉTHODE COPHÉNÉTIQUE.

4.2.1. Phénons.

La comparaison des dendogrammes entre eux est l'un des buts de l'analyse de groupe. L'emploi des différents coefficients ou des différentes méthodes (WPGM, UPGM) peut rendre cette comparaison indispensable. R. R. SOKAL et P. H. A. SNEATH (1963, p. 312) ont proposé une méthode de comparaison nommée méthode cophénétique. On trace à travers le dendogramme des lignes verticales ou phénons à intervalles réguliers et constants. Les classes cophénétiques ainsi délimitées sont égales et codées de 1 à n en partant de la valeur la plus faible du coefficient de similarité considéré. Le nombre de classes cophénétiques dépend du nombre d'éléments (OTU ou niveaux) du dendogramme. D'une façon générale, les dendogrammes comprenant moins de 10 éléments ne peuvent être subdivisées, toujours d'après les mêmes auteurs, en plus de 4 classes cophénétiques (fig. 6). Ceux qui en comportent plus de 100 peuvent être divisés au moins en 10 classes.

On caractérise alors un groupe par la valeur de son phénon. Le groupe [1-2-5] de la figure 6 a pour phénon 4 (cf. M. JACOBZONE et G. CARBONNEL, 1971, p. 170). On peut ainsi comparer la valeur qu'il aurait dans un autre dendogramme.

4.2.2. Matrice cophénétique — Coefficient de corrélation cophénétique

On établit pour chaque dendogramme une matrice cophénétique (fig. 7) en suivant la méthode explicitée dans le paragraphe précédent. L'exemple ci-dessous est réalisé à partir du dendogramme de la figure 6.

espèces →	1	2	3	4	5	6
↓						
1	X	4	2	3	4	3
2	4	X	2	3	4	3
3	2	2	X	2	2	2
4	3	3	2	X	3	3
5	4	4	2	3	X	3
6	3	3	2	3	3	X

FIG. 7. — Matrice cophénétique pour les espèces de la figure 6.

On établit une seconde matrice cophénétique concernant le second dendogramme dont on veut tester la ressemblance avec le premier. On calcule ensuite, pour la totalité des données, la valeur du coefficient de corrélation entre les deux dendogrammes. Il se nomme coefficient de corrélation cophénétique. Plus sa valeur est élevée, moins grande est la distorsion entre les deux dendogrammes et inversement.

DEUXIÈME PARTIE

1. LES CORRÉLATIONS PAR L'ANALYSE DE GROUPE SELON LE MODE Q

APPLICATION A L'ÉTUDE DES NIVEAUX DE PRÉLÈVEMENTS

On peut étudier par cette méthode la liaison existant entre les niveaux d'une coupe ou entre les niveaux de plusieurs coupes.

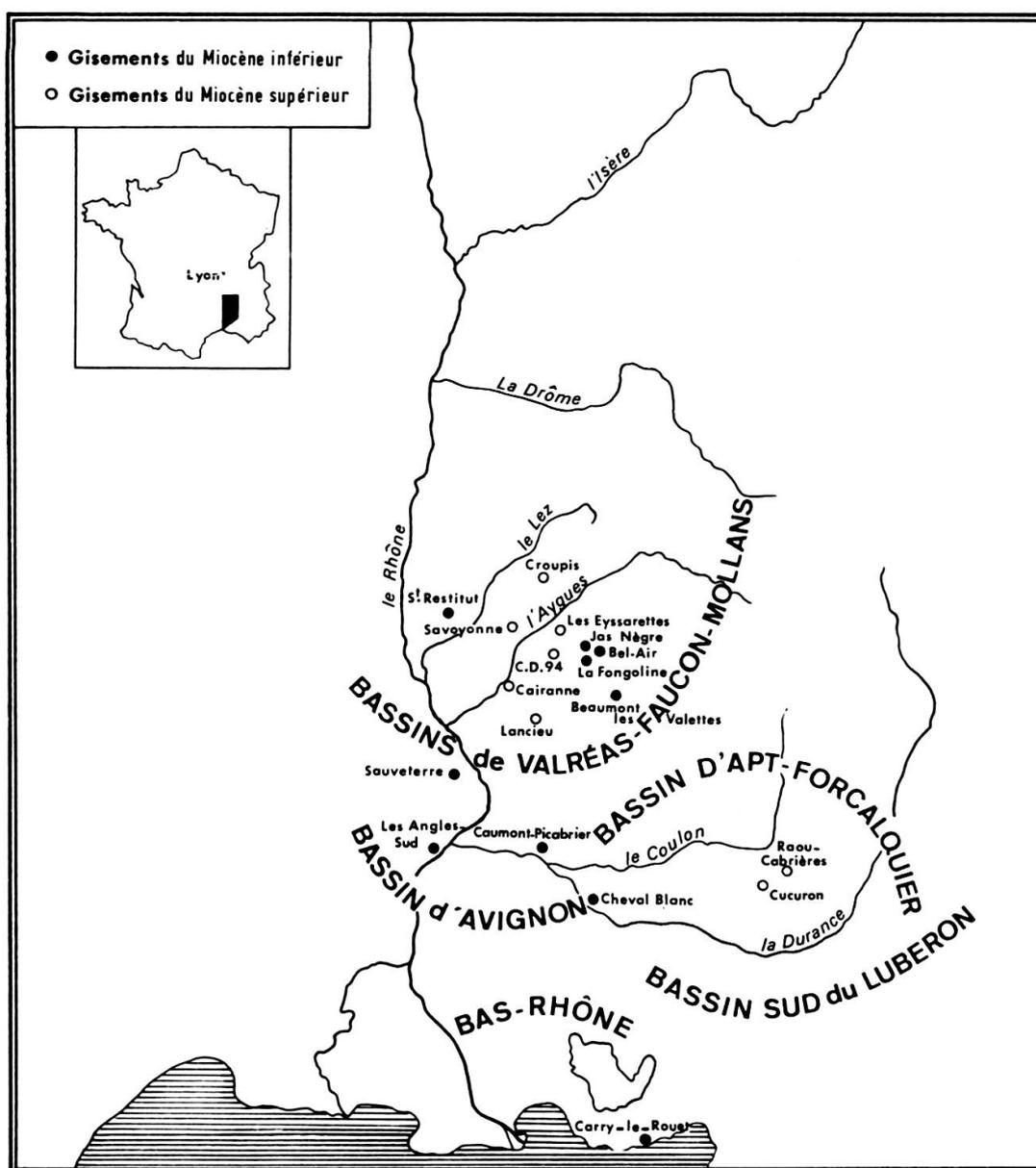


FIG. 8. — Localisation géographique des coupes et des gisements miocènes faisant l'objet de l'analyse de groupe.

1.1. GROUPEMENT DES NIVEAUX D'UNE COUPE EN FORMATION

L'exemple étudié concerne les niveaux de prélèvements de la coupe de Carry-le-Rouet, figure 8, (P. ANDREIFF, R. ANGLADA et *alii*, 1971, sous presse). Le dendogramme de la figure 9 a été obtenu à partir des données portées sur le tableau de

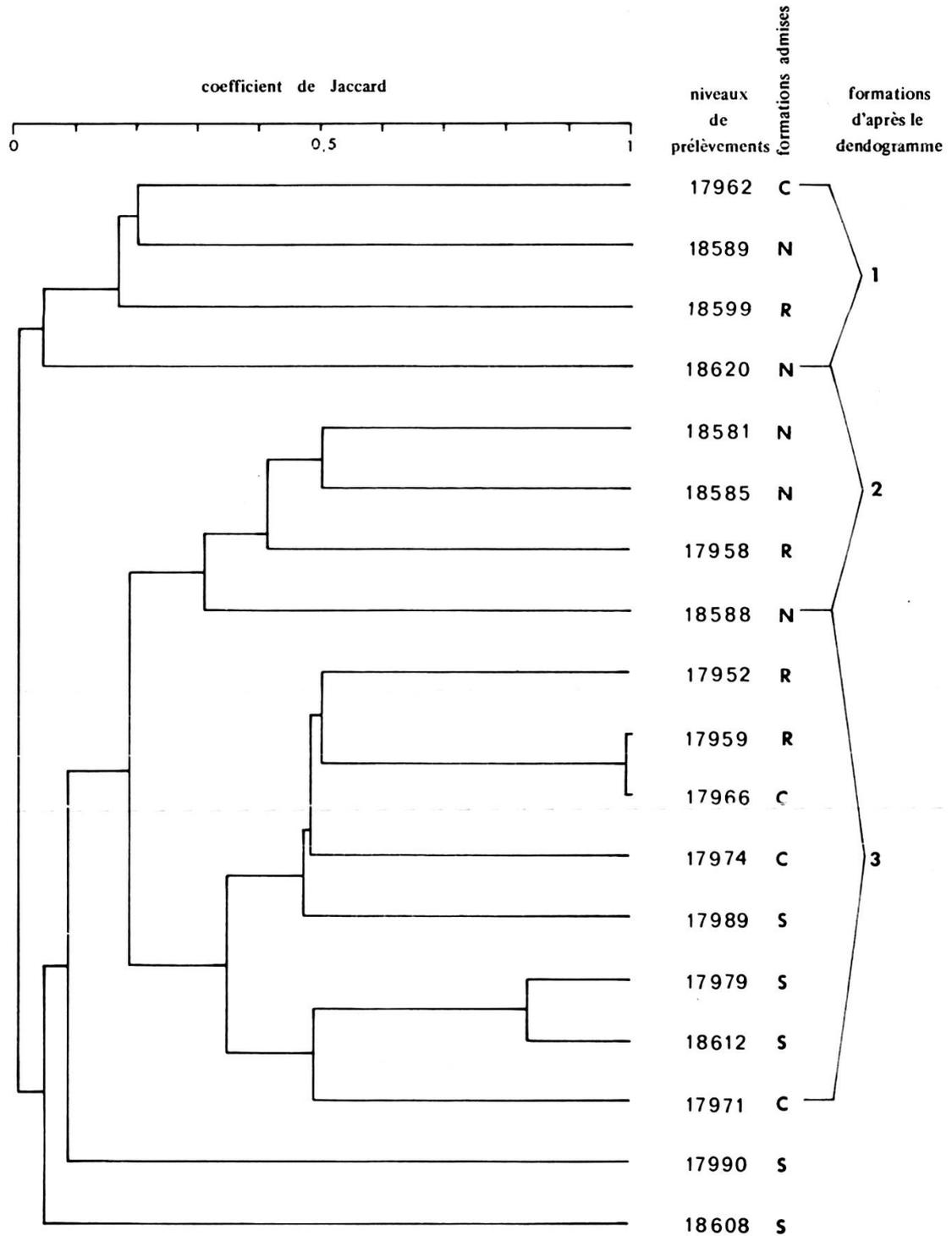


FIG. 9. — Dendogramme résultant du groupement des niveaux de la coupe de Carry-le-Rouet, d'après les ostracodes (Jaccard, WPGM).

répartition des ostracodes (tabl. 1). On a utilisé le coefficient de Jaccard et la méthode pondérée WPGM, pour traiter ce problème.

Pour une valeur du coefficient de Jaccard égale à 0,10, on peut séparer les niveaux de prélèvements en trois groupes, c'est-à-dire en trois formations. Le premier groupe comprend les niveaux 17962, 18589, 18599, 18620; le second comprend les niveaux 185881 à 18588 et le troisième les niveaux 17952 à 17971. Par comparaison, le découpage en formations basé sur les données lithologiques, sédimentologiques et fauniques est indiqué sur la figure 9. La concordance des résultats entre les deux méthodes est seulement approchée. Le premier et second groupe de niveaux correspondent, à l'exception du prélèvement 17962, à la formation pararéçifale du Cap des Nautes (N) et à la formation saumâtre du Rousset (R). Il semble impossible de les séparer par dendogramme sur la base de l'ostracofaune. L'autre groupe comprend, également mélangés, les niveaux rapportés à la formation bioclastique de Carry (C) et à la formation biodétritique de Sausset-les-Pins (S), sans qu'il soit possible, d'après l'analyse de groupe, de les séparer. Néanmoins, d'après cette méthode, les niveaux 17952 et 17959 au sommet de la formation saumâtre du Rousset, serait à rattacher à ce troisième groupe. C'est un ensemble de résultats que l'examen direct de la répartition des ostracodes ne laissait pas présumer. Dans cet exemple, l'analyse de groupe par la méthode Q, appliquée aux niveaux à ostracodes permet de découvrir, par le calcul seul, la réalité de deux grandes formations. On reviendra plus loin sur la signification paléoécologique de tels résultats.

1.2. GROUPEMENT DES NIVEAUX D'UNE COUPE ET LA NOTION D'INSTABILITÉ PALÉOÉCOLOGIQUE (EN LITHOFACIÈS CONSTANT)

Dans ce cas l'analyse de liaison entre les niveaux a porté sur les prélèvements de la coupe de Caumont-Picabrier. On se référera pour la répartition des ostracodes aux travaux de G. CARBONNEL (1969 et 1971). On se propose, comme précédemment, de chercher le degré de liaison entre les prélèvements numérotés de 1 à 13 du bas en haut de la coupe. Le dendogramme de la figure 10 obtenu par les mêmes méthodes que précédemment (coefficient de Jaccard et WPGM) permet de se rendre compte que la réunion des niveaux se réalise dans un certain désordre.

Le cas idéal (*cf.* J. E. HAZEL, 1970) consisterait à observer une liaison des niveaux dans l'ordre de leur succession verticale, ici de 1 à 13: cette succession dans l'ordre se produit partiellement pour les niveaux 4, 5, 6, 7 d'une part et 9, 10, 11 d'autre part. L'association de ces deux groupements par contre est quelconque puisque les niveaux 9, 10, 11 de l'un s'associent avec le niveau 4 de l'autre.

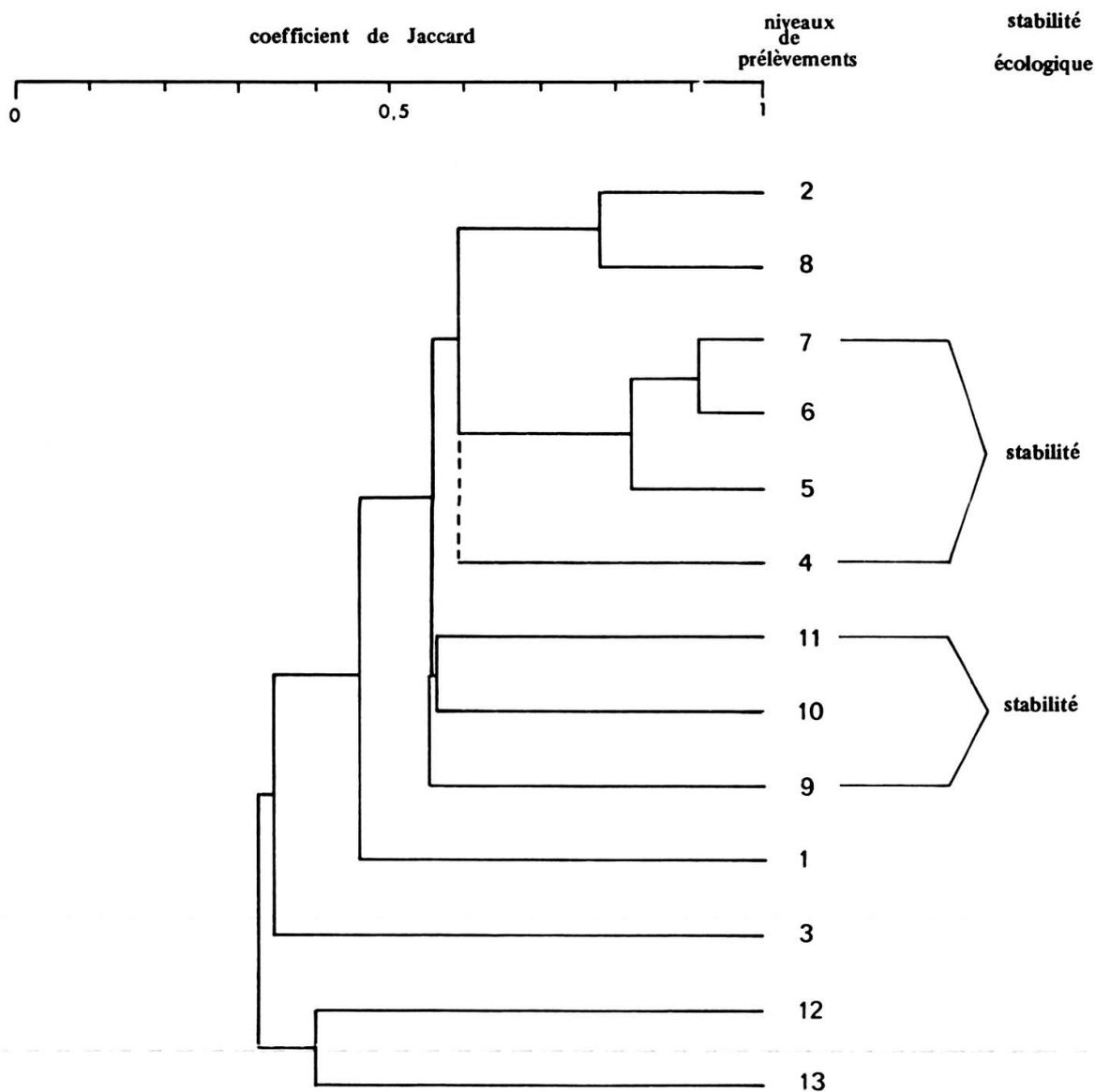


FIG. 10. — Dendrogramme résultant du groupement des niveaux de la coupe de Caumont-Picabrier d'après les ostracodes (Jaccard, WPGM).

L'anarchie du groupement des niveaux, par rapport à leur ordre de succession verticale, interdit d'envisager dans ce cas une utilisation stratigraphique des groupements obtenus. Elle impose au contraire de rechercher une explication paléécologique. L'exemple précédent nous laissait déjà pressentir la valeur paléécologique de l'analyse de groupe. Les niveaux s'associent en effet préférentiellement selon leur contenu faunique (ici, les ostracodes), reflets d'abord des conditions écologiques du biotope. A une échelle temporelle aussi faible (épaisseur de la coupe 7 m environ)

les affinités paléoécologiques prédominent. Elles peuvent expliquer les liaisons préférentielles observées entre les niveaux.

Ces résultats, décevants sur le plan stratigraphique, sont au contraire précieux sur le plan paléoécologique. Leur présence permet d'affirmer qu'ils existent des périodes de stabilité paléoécologiques 4 à 7, 9 à 11, entrecoupées de périodes d'instabilité.

D'un point de vue local, ils permettent en outre de rapprocher paléoécologiquement des niveaux aussi éloignés verticalement que 2 et 8.

L'ensemble de ces résultats ne pouvait se déduire de l'examen direct du tableau de répartition des espèces (tabl. 2). L'emploi de l'analyse de groupe se révèle ici supérieure aux méthodes classiques employées jusqu'à présent.

Un second exemple permettant d'apprécier les périodes de stabilité écologique au sein d'une coupe est fourni par le dendogramme de la figure 11 pour les niveaux de la coupe de St-Restitut (bassin de Valréas).

1.3. GROUPEMENT DES NIVEAUX D'UNE COUPE EN LITHOFACIÈS VARIÉS INFLUENCE DU LITHOFACIÈS

On s'intéresse dans ce cas à la liaison éventuelle entre l'ostracofaune et le lithofaciès. Dans l'exemple précédent, l'analyse a porté sur la liaison des niveaux au sein d'une formation considérée comme homogène d'après un examen sur le terrain.

Ici, les prélèvements ont été réalisés au sein de trois lithofaciès différents mais superposés de la coupe de St-Restitut (bassin de Valréas, fig. 8). On observe de bas en haut: un faciès sableux (S), un faciès calcaréo-argileux (MC) et un faciès calcaire (C). La répartition des ostracodes a été indiquée par G. CARBONNEL (1969, tabl. 9).

Dans l'hypothèse d'une réaction favorable entre le lithofaciès et l'ostracofaune les niveaux doivent se grouper en trois par affinités lithologiques. Dans le cas inverse on observera le désordre le plus complet.

Le dendogramme de la figure 11 (obtenu à partir du coefficient de Jaccard et de la méthode WPGM) permet de constater l'existence de trois groupements de niveaux, nommés respectivement 1, 2, 3.

Le groupement 1 comprend les niveaux 4 à 38
 » » 2 » » » 34 à 33
 » » 3 » » » 7 à 20

Dans le groupement 1 les niveaux appartiennent au lithofaciès sableux à l'exception des niveaux 38 et 40. Dans le groupement 2 les niveaux appartiennent au lithofaciès calcaréo-argileux ou calcaire. Dans le groupement 3 les niveaux appartiennent au lithofaciès sableux à l'exception des niveaux 27 et 28. Il ressort de cette analyse que l'ostracofaune réagit de façon différente suivant le lithofaciès. On ne peut

toutefois distinguer par cette méthode un groupement préférentiellement lié au lithofaciès calcaréo-argileux. Le pourcentage de CaCO_3 du milieu de sédimentation semble jouer un rôle moins prépondérant que sa granulométrie.

Cependant, l'ensemble des niveaux du lithofaciès sableux est scindé en deux groupes 1 et 2, d'après le dendrogramme de la figure 11. Aucune liaison de valeur

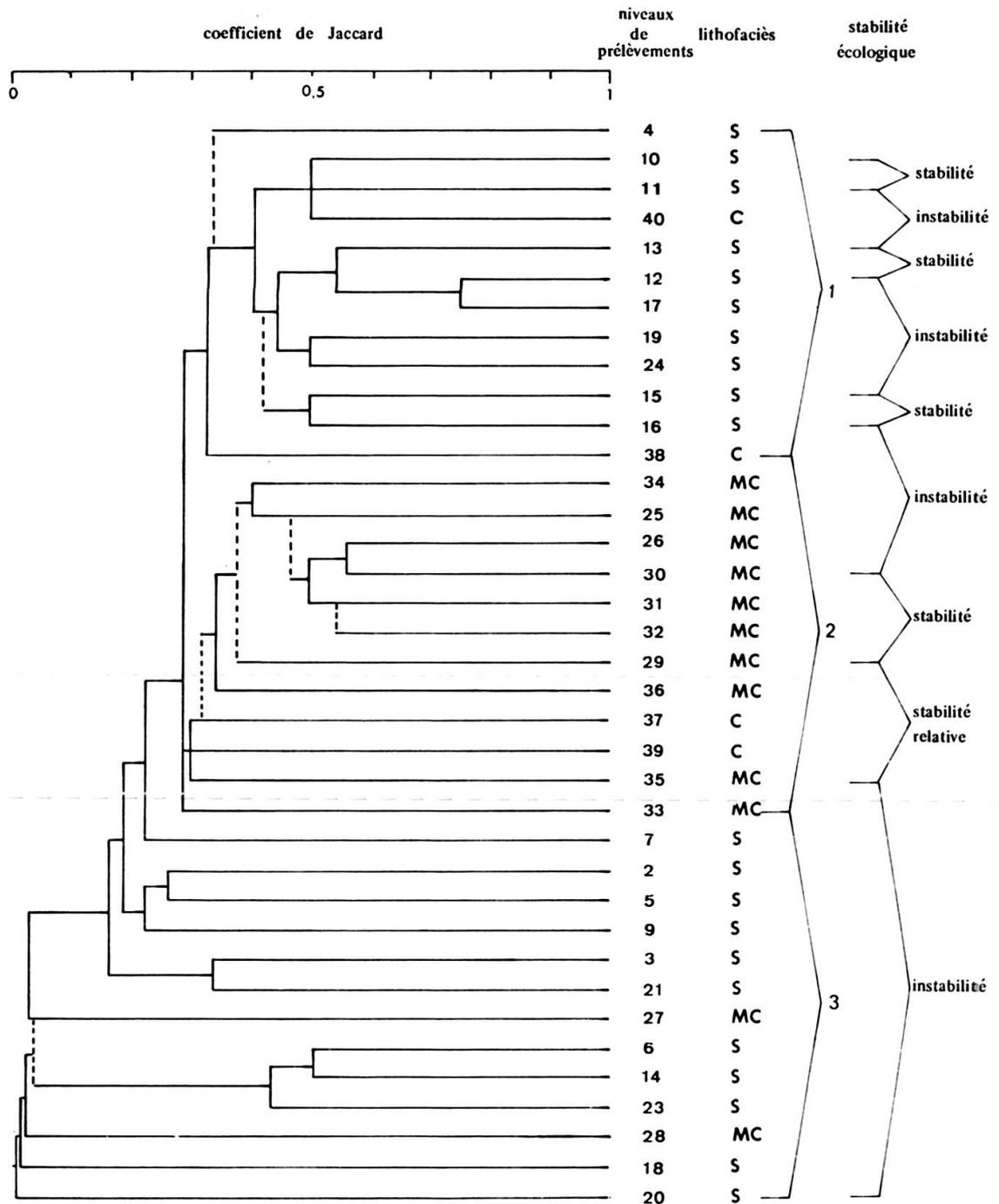


FIG. 11. — Dendrogramme résultant du groupement des niveaux de la coupe de St-Restitut d'après les ostracodes et montrant l'influence du lithofaciès (Jaccard, WPGM); S: faciès sableux « molassique »; MC: faciès marno-calcaire; C: faciès calcaire.

élevée n'existerait entre eux. Le facteur déterminant cette distribution est actuellement ignoré. Au sein de chacun des groupements la liaison entre les niveaux ne s'effectue pas dans l'ordre de leur succession verticale. On retrouve ici la notion définie dans l'exemple précédent. Cette observation tend à prouver que le lithofaciès n'est pas le seul facteur influant sur la répartition des ostracodes. Il n'est en effet qu'un des aspects résultant des conditions de milieu.

Remarque. — La liaison « anormale », d'après ce dendogramme (fig. 11), des niveaux 27 et 28, prélevés en milieu calcaréo-argileux, avec le lithofaciès sableux peut provenir du nombre réduit d'ostracodes récoltés (*cf.* CARBONNEL 1969, tabl. 9), le sédiment étant fortement induré.

1.4. GROUPEMENT DES NIVEAUX ENTRE PLUSIEURS COUPES

1.4.1. *Dans des bassins différents*

Il s'agit dans ce cas d'étudier la liaison pouvant exister entre les niveaux de plusieurs coupes relevées dans des bassins éloignés. Le bassin de Faucon-Mollans (fig. 8) comprend les coupes de la Fongoline (CARBONNEL, 1969, p. 271), de Jas-Nègre (*ibid.* p. 271) de Beaumont-les-Valettes (*ibid.* p. 275), et de Bel-Air (*ibid.* p. 268). Le bassin d'Avignon-Cavaillon comprend les coupes de Sauveterre (CARBONNEL 1969, p. 264), de Cheval-Blanc (*ibid.* p. 268), de Caumont-Picabrier (*ibid.* p. 262) et des Angles-Sud (*ibid.* p. 264).

Ces coupes, biostratigraphiquement synchrones d'après l'ostracofaune, les foraminifères et la macrofaune sont estimées appartenir, à tort ou à raison, à un lithofaciès similaire sinon proche. Certains géologues rhodaniens¹ récuse cette similitude de faciès basée selon eux sur une étude du terrain trop superficielle. En supposant que des analyses sédimentologiques confirment ainsi leur opinion, la similitude de faciès, alors apparente, ne peut en rien modifier les calculs conduisant à l'obtention du dendogramme (fig. 12). C'est un paramètre commun entre des coupes dont on cherche le degré de liaison. Sa présence ou son absence ne peut en aucune façon modifier le dendogramme. C'est seulement au niveau de l'explication, donc de la causalité que ce paramètre pourra ou non être invoqué. L'exemple du dendogramme de la figure 12 confirmera cette interprétation. Basés à l'origine sur l'identité du faciès, les résultats obtenus nécessitent de faire appel à d'autres facteurs pour les expliquer.

¹ L. David, Université LYON 1, communication orale.

Dans quelques cas, les niveaux de prélèvements de l'une des coupes s'associent de préférence entre eux plutôt qu'entre coupes différentes. On l'observe pour la majorité des niveaux des coupes de Caumont-Picabrier, de Cheval-Blanc et des Angles-Sud. Ceci implique l'existence de biotopes particuliers à chaque coupe.

On constate en outre (fig. 12) une certaine ségrégation des niveaux: d'un côté les niveaux correspondants aux coupes du bassin d'Avignon-Cavaillon s'associent entre eux; de l'autre la liaison s'effectue entre les niveaux des coupes du bassin de

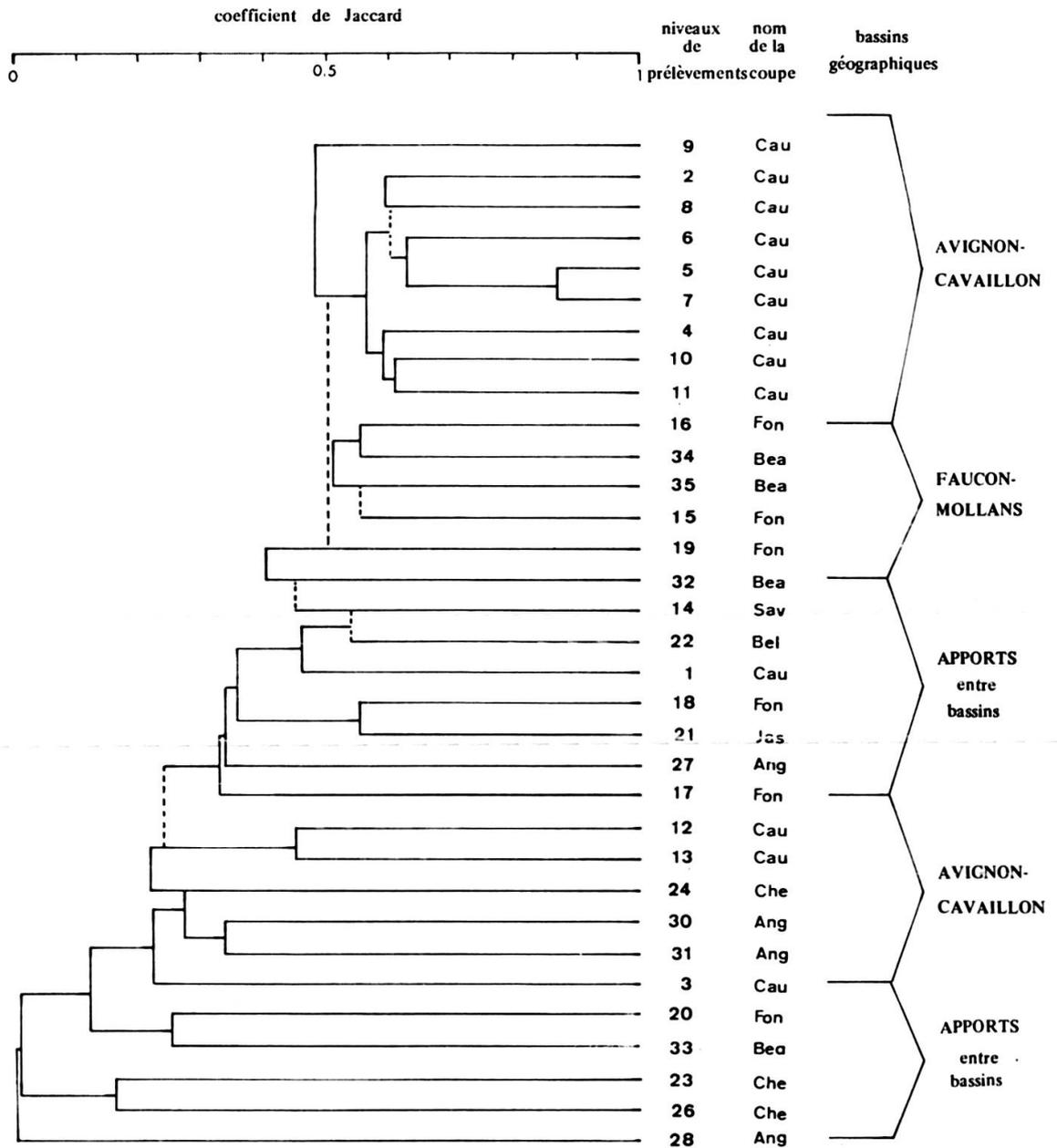


FIG. 12. — Dendrogramme résultant du groupement des niveaux des coupes de Caumont-Picabrier (Cau), de Sauveterre (Sau), des Angles-Sud (Ang), de la Fongoline (Fon), de Jas-Nègre (Jas), de Beaumont-les-Valettes (Bea) et de Cheval-Blanc (Che) d'après les ostracodes (Jaccard, WPGM).

Faucon-Mollans. C'est un résultat important et nouveau que seul peuvent révéler des méthodes de calcul analogues.

Il existe cependant des *liaisons désordonnées* entre les niveaux de plusieurs coupes traduisant des communications par apport faunique d'un bassin à l'autre, (par exemple les niveaux 1, 3, 12 et 13 de la coupe de Caumont-Picabrier). Ce « désordre » permet une tentative de reconstitution des liaisons existant à l'échelle verticale entre les niveaux des diverses coupes (cf. fig. 13 et 15).

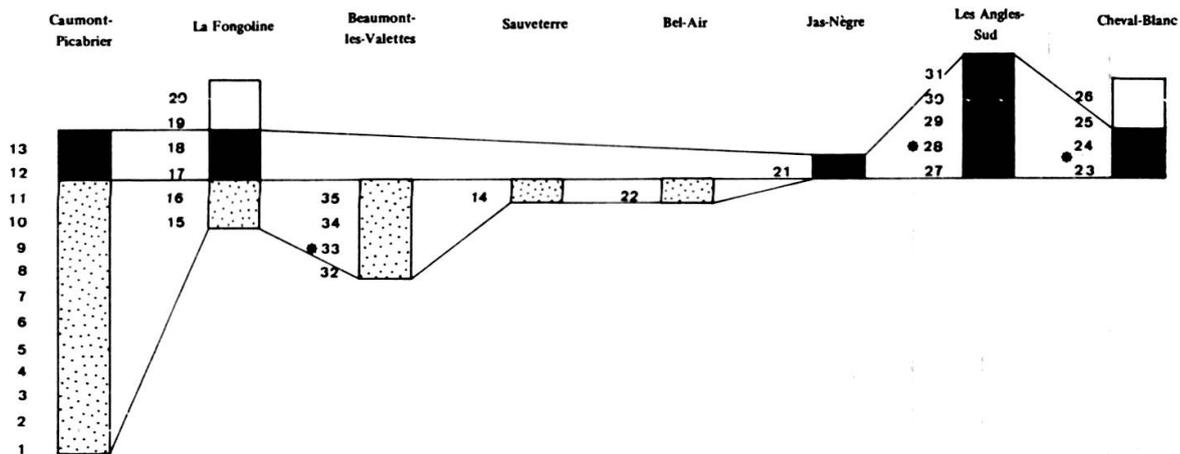


FIG. 13. — Essai de corrélation stratigraphique entre les différentes coupes d'après le dendrogramme de la figure 12 (épaisseur en mètres non respectée).

Méthode pratique

Il faut et il suffit de transcrire les corrélations obtenues par dendrogrammes sur les colonnes stratigraphiques de chaque coupe. On se souviendra que l'équivalence des niveaux sur un dendrogramme peut résulter d'une affinité écologique prépondérante et non pas seulement stratigraphique. Avant de poursuivre le raisonnement on se rappellera également que la corrélation biostratigraphique suppose une liaison exclusive des niveaux ainsi corrélés entre eux. A l'inverse dans l'analyse de groupe, chaque groupe élémentaire de niveaux (correspondant aux corrélations de valeurs les plus élevées) est forcément associé, ensuite, avec les autres à une valeur inférieure. Il sera donc nécessaire de se baser sur ces groupes élémentaires pour réaliser les corrélations stratigraphiques entre les coupes. Le dendrogramme de la figure 14 concernant les coupes de Cairanne, de la Savoyonne, de Croupis, de la R.D. 94, de Lancieu, de Roaix et des Eyssarettes (considérées comme synchrones du Miocène supérieur, CARBONNEL, 1969) nous servira d'exemple.

Les groupes élémentaires comprennent:

- les niveaux 1, 26, 30 biostratigraphiquement équivalents
- les niveaux 6, 14, 27, 11, 7, 17, 9, 15, 28, 29, 12, 13, 22 biostratigraphiquement équivalents entre eux mais non pas avec le groupe élémentaire précédent.

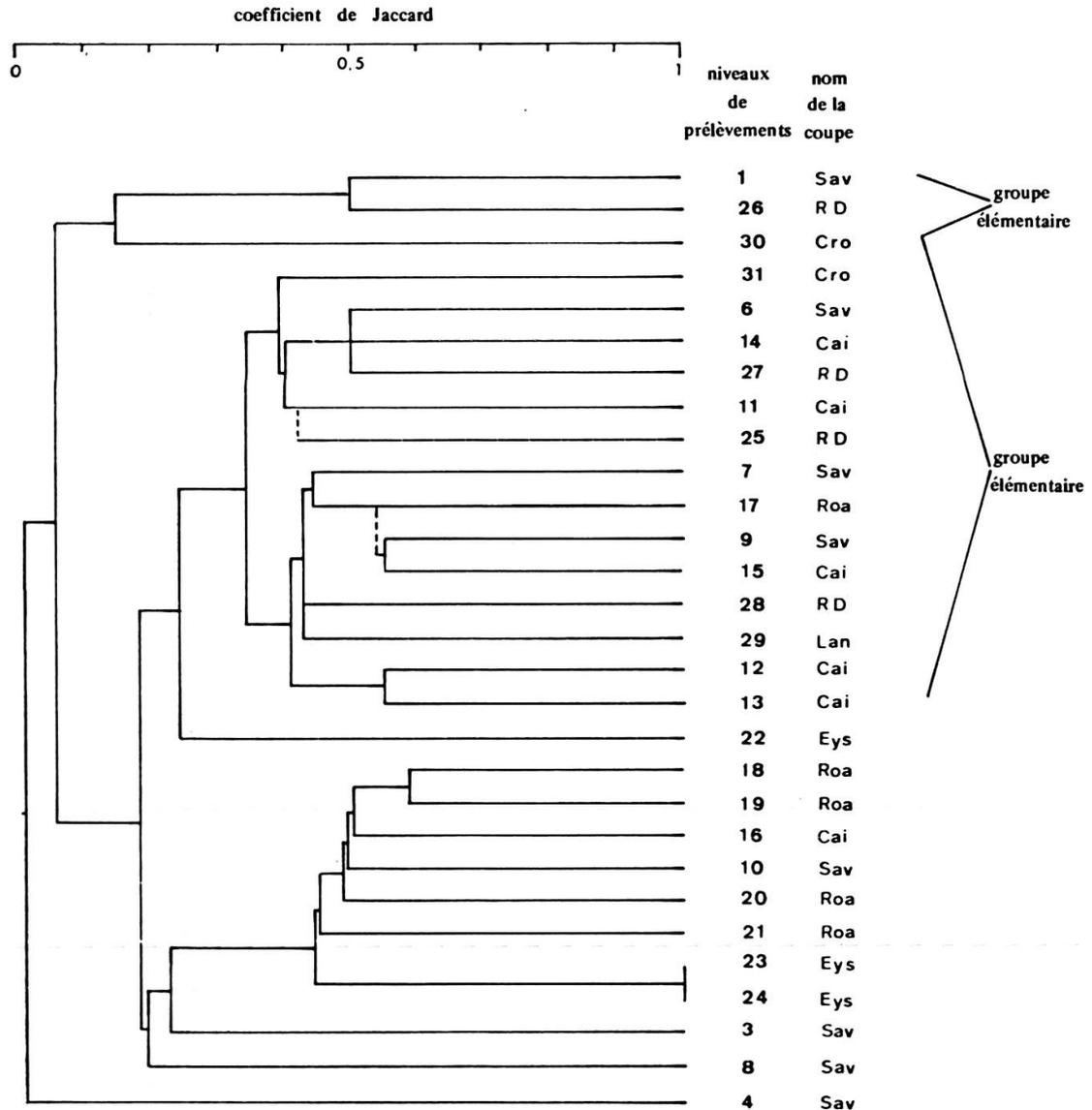


FIG. 14. — Dendrogramme résultant du groupement des niveaux des coupes de Roaix (Roa), des Eyssarettes (Eys), de Cairanne (Cai), de Lancieu (Lan), de la R.D. 94 (RD), de la Savoyonne (Sav) et de Croupis (Cro) d'après les ostracodes (Jaccard, WPGM).

Un exemple des difficultés ¹ rencontrées dans ce type d'interprétation est illustré par les niveaux 25 et 26 (de la coupe de la R.D. 94). Le niveau 25 est d'après le dendrogramme de la figure 14 inclus dans le second groupe élémentaire. Or le niveau 26 qui le surmonte appartient au premier groupe. Il est plus raisonnable de considérer

¹ Illustrées sur les figures 13 et 15 par un astérisque.

que les niveaux 25 et 26 sont réunis dans un seul groupe élémentaire: le premier. On peut ainsi situer biostratigraphiquement les coupes, les unes par rapport aux autres (*cf.* fig. 15).

Ainsi à l'échelle du bassin rhodanien l'un des résultats les plus importants concerne la position relativement basse de la coupe de Caumont-Picabrier par rapport aux autres coupes considérées.

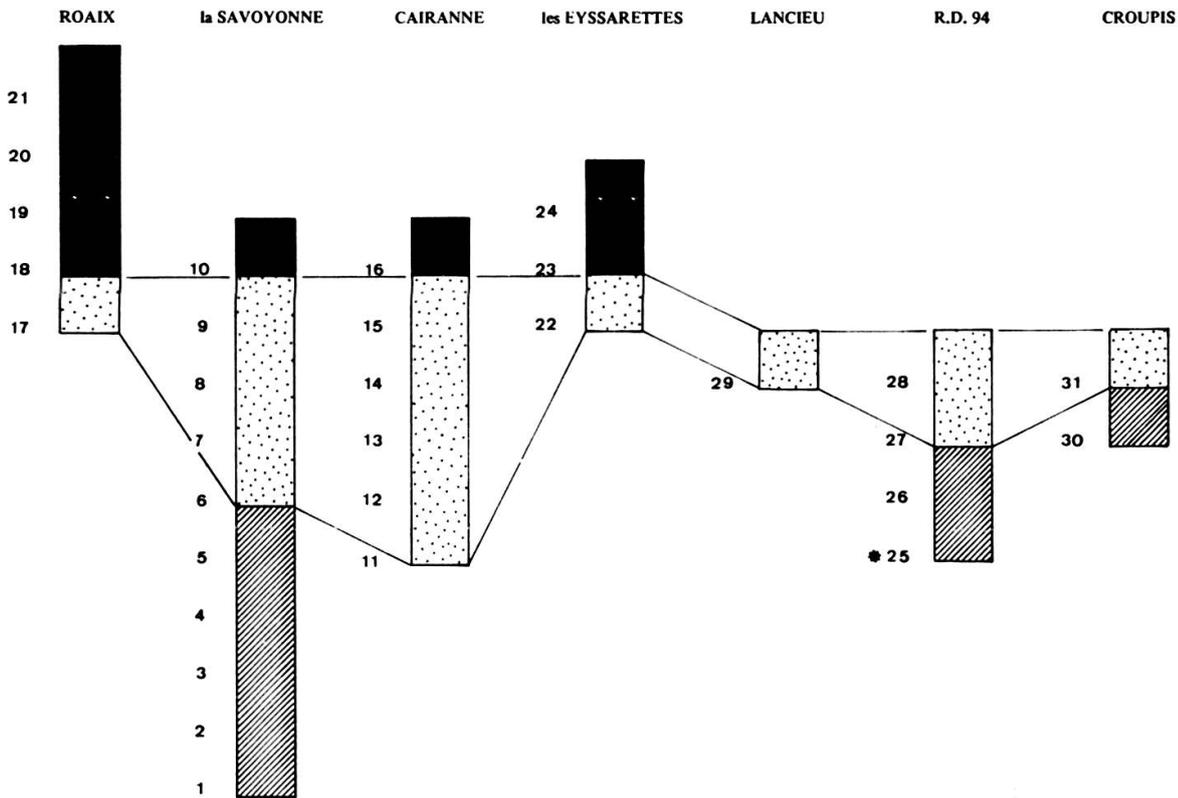


FIG. 15. — Essai de corrélation stratigraphique entre les différentes coupes d'après le dendrogramme de la figure 14 (épaisseur en mètres non respectée).

Cependant aucun moyen d'observation directe sur le terrain ne pourra confirmer ou infirmer la validité de cette méthode. C'est une lacune que l'exemple suivant s'efforcera de combler.

1.4.2. Dans un même bassin

Trois coupes ont été relevées sur le plateau des Courennes dans le bassin d'Apt et rapportées au Burdigalien. Elles sont numérotées 3, 10 et 18 (M. PHILIPPE, 1972). Cet auteur a suivi plusieurs niveaux repères permettant la corrélation des assises par observation directe.

L'analyse de groupe d'après l'ostracofaune et selon le mode Q, aboutit dans une seconde phase à proposer les corrélations de la figure 16.

On constate un décalage entre les corrélations calculées et celles observées. Ainsi l'intervalle compris entre les niveaux 7 à 9 de la coupe 18 est corrélé sur le terrain avec l'intervalle 4 à 5 de la coupe 10.

D'après l'analyse de groupe il serait corrélé avec l'intervalle 3-4 de la coupe 10. De même l'intervalle 3-4 de la coupe 10 est corrélé par observation directe avec l'intervalle 5-6 de la coupe 3. Au contraire, d'après les calculs, la corrélation interviendrait entre l'intervalle 3-4 de la coupe 10 et 4 à 6 de la coupe 3.

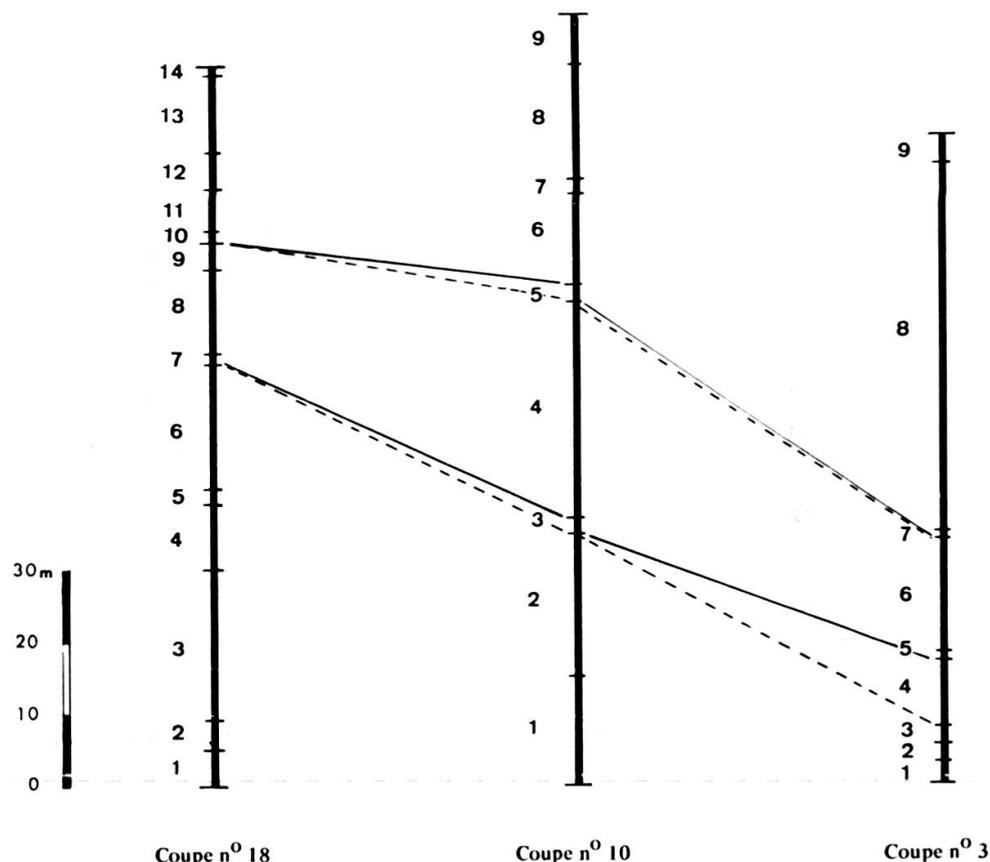


FIG. 16. — Essai de corrélation stratigraphique d'après l'analyse de groupe (Jaccard, W P G M) entre plusieurs coupes rapprochées, dans un même bassin, d'après les ostracodes (en trait interrompu). Comparaison avec les corrélations effectuées sur le terrain (en trait continu).

Un décalage constant en direction (ici vers le bas de la coupe) existe entre les corrélations obtenues par les deux méthodes. L'analyse de groupe, selon le mode Q traduit alors une plus grande identité écologique existant entre les intervalles ainsi corrélés, même chronologiquement légèrement décalés. On peut donc accepter l'ensemble de ces résultats sachant qu'une certaine imprécision persiste sur leur valeur chronostratigraphique.

1.5. GROUPEMENT ENTRE PLUSIEURS COUPES

On étudie le degré de liaison entre plusieurs coupes. L'exemple concerne les coupes de Cairanne, du C.D. 94, de la Savoyonne, de Raou-Cabrières, de Cucuron inférieur, de Cucuron supérieur et de Lancieu. Les trois premières sont situées (fig. 8) dans le bassin de Valréas-Faucon-Mollans, les trois suivantes dans le bassin Sud du Lubéron; la dernière est à la limite du bassin de Valréas et du bassin d'Apt (fig. 8).

Elles sont datées du Miocène supérieur (biozones C et D, CARBONNEL 1969, pp. 333-336) et considérées à grande échelle comme synchrones.

La plupart d'entre elles comprennent plusieurs niveaux de prélèvements. On a synthétisé la présence des ostracodes dans les prélèvements d'une coupe en un seul ensemble de données. La comparaison s'effectue alors entre sept ensembles de données correspondant aux sept coupes citées plus haut.

1.5.1. *Analyse écologique*

Le dendrogramme de la figure 17 permet de constater que toutes les coupes à l'exception de celle de Cucuron supérieur sont étroitement associées les unes aux autres et dans un certain ordre: d'abord celles du bassin de Valréas-Faucon-Mollans puis celles du bassin sud du Lubéron.

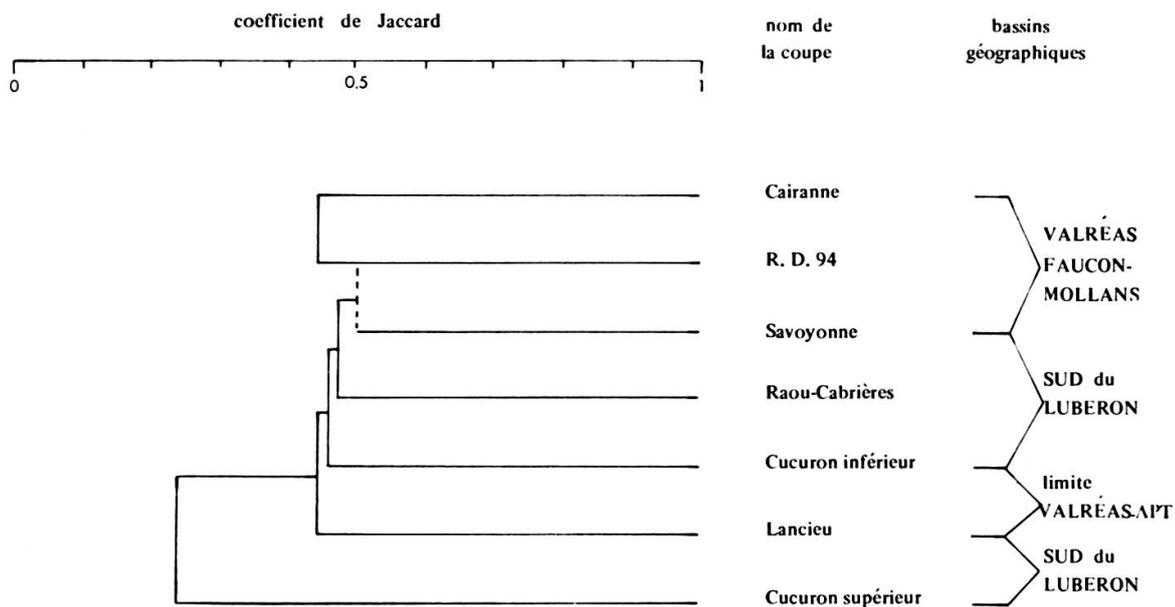


FIG. 17. — Dendrogramme résultant du groupement de quelques coupes du Miocène supérieur d'après les ostracodes (Jaccard, WPGM).

L'analyse de groupe détecte le déterminisme géographique qui semble prépondérant.

1.5.2. Analyse biostratigraphique

Pour une valeur du coefficient de Jaccard égale à 0,44 on peut distinguer deux groupes; le premier comprend toutes les coupes de Cairanne à Lancieu, le second correspond à la seule coupe de Cucuron supérieur. Les coupes du premier groupe sont équivalentes d'après le dendogramme. On peut considérer que cette équivalence s'applique également à leur position biostratigraphique, ce qui permet d'assurer aussi leur corrélation stratigraphique.

La coupe de Cucuron supérieur, associée au groupe précédent avec une valeur inférieure à 0,44 n'est donc pas équivalente biostratigraphiquement. Cette coupe située verticalement sur le terrain au-dessus de la coupe de Cucuron inférieur est donc biostratigraphiquement plus élevée que l'ensemble précédent. C'est là un résultat nouveau et important ignoré par les moyens d'investigation employés jusqu'à présent.

2. LES CORRÉLATIONS PAR L'ANALYSE DE GROUPE SELON LE MODE R

APPLICATION A L'ÉTUDE DES ASSOCIATIONS D'ESPÈCES

C'est l'étude de la liaison existant entre les différentes espèces (ou d'après la terminologie anglaise OTU) présentes dans une ou plusieurs coupes.

Les renseignements tirés de cette analyse concernent la composition des associations d'espèces ou ses modifications. Elle permet d'apprécier éventuellement l'action de certains paramètres écologiques.

2.1. ASSOCIATIONS D'ESPÈCES DANS UNE COUPE

La coupe de St-Restitut (CARBONNEL, 1969, p. 245) dont on étudie l'association des espèces d'ostracodes a fait l'objet de l'analyse des niveaux selon le mode Q (2^e partie, paragraphe 1.3.).

Le dendogramme de la figure 18 permet de séparer trois associations: A, B et C. L'association A renferme les espèces allant de *Cytheropteron sp. B* à *Pseudocytherura tricastensis* et l'association B les espèces allant de *Loxoconcha subovata* à *Costa edwardsii*. L'association C comprend celle allant de *Hemicythere triangularis triangularis* à *Eucytherura textilis tridentata*.



FIG. 18. — Dendrogramme résultant du groupement des ostracodes de la coupe de St-Restitut (Jaccard, WPGM).

On se reportera pour la répartition verticale des ostracodes au tableau 12 de G. CARBONNEL, 1969 et figure 1 du même auteur, 1971. Ces données sont évidemment analogues à celles utilisées plus haut pour l'étude du groupement des niveaux.

D'après l'analyse selon le mode Q, les périodes de stabilité écologique comprennent les niveaux 4 à 7 et 9 à 11 (paragraphe 1 et fig. 10). On peut alors sur la base des résultats ainsi obtenus calculer selon le mode R, les dendogrammes des associations d'espèces correspondant aux périodes de stabilité écologique ainsi définies. Les associations obtenues sont indiquées sur les dendogrammes des figures 19 et 20.

Période de stabilité paléoécologique 4 à 7.

Pour une valeur du coefficient de Jaccard de 0,50 le milieu comprend quatre associations :

1^{re} association à *A. haueri*, *N. helvetica*, *H. triangularis*

2^e association à *Cyth. sp. B*, *P. solida*

3^e association à *C. fournieri*, *Lox. punctatella*

4^e association à *C. boldi* seul

Période de stabilité paléoécologique 9 à 11.

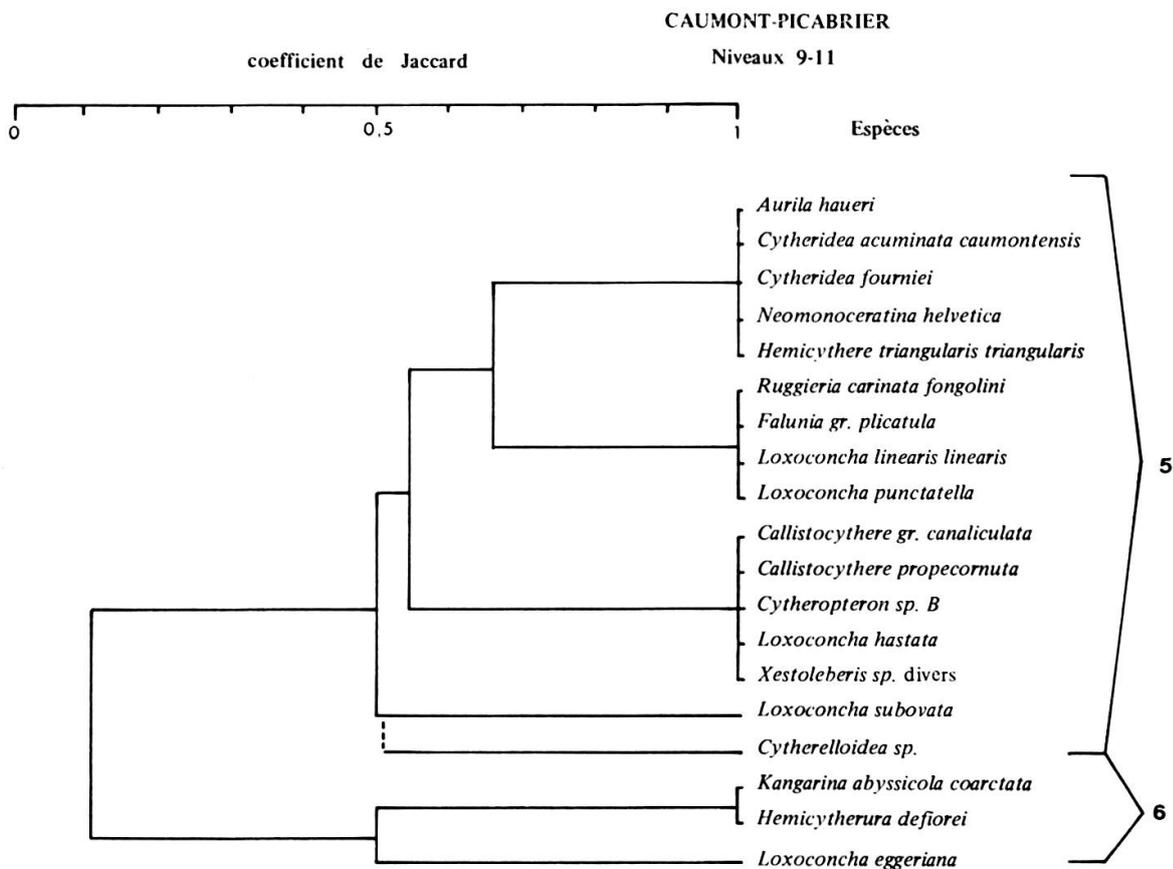


FIG. 20. — Dendrogramme résultant du groupement des ostracodes de la coupe de Caumont-Picabrier durant la période de stabilité paléoécologique des niveaux 9 à 11 (Jaccard, WPGM).

Pour une valeur du coefficient de Jaccard de 0,50 le milieu ne comprend plus que deux associations:

- 5^e association à *A. haueri*, *N. helvetica*, *Cyth. sp. B*, *C. fournieri*
 6^e association à *K. abyssicola*, *H. defiorei*, *L. eggeriana*.

On constate une condensation des associations de la période 4 à 7 durant cet intervalle. Elle aboutit à réunir en une seule association (assoc. n° 5) les associations n° 1, 2 et 3. Toute tentative d'explication est jusqu'à présent malaisée. Il est difficile d'envisager une modification du comportement écologique des espèces dans un intervalle de temps aussi réduit. On peut par contre proposer l'hypothèse suivante: les associations n° 1, 2 et 3 ont une réelle existence. Elles ont dû se succéder dans le temps de façon telle qu'aucun chevauchement de l'une sur l'autre n'ait pu se réaliser. En un lieu, celui des prélèvements 4 à 7, une seule association a existé à un moment donné. Durant l'intervalle 9 à 11, la majorité des espèces composant les trois associations précédentes a coexisté en un même lieu ce qui n'a plus permis de les séparer par la procédure utilisée.

2.1.3. Notion d'espèces dominantes et accessoires

Les dendogrammes reproduits jusqu'ici ont été obtenus à partir d'une matrice de base résultant de l'application du coefficient de Jaccard. On peut être tenté de la remplacer par le coefficient de corrélation (r). Cette substitution permettra de prendre en considération le nombre réel d'individus de chaque espèce. Le dendogramme de la figure 21 a été réalisé en suivant cette méthode. Il concerne les espèces de la coupe de Caumont-Picabrier considérée dans sa totalité (voir 2^e partie, paragraphe 1.2. fig. 8 et tabl. 2).

D'après le dendogramme, la première association comprend les espèces 1 à 18; la seconde les espèces 3 à 14.

D'après le tableau de répartition des espèces (CARBONNEL 1969, tabl. 12) les espèces dominantes sont:

<i>A. haueri</i>	<i>P. solida</i>
<i>C. fournieri</i>	<i>H. triangularis triangularis</i>
<i>F. groupe plicatula</i>	<i>C. acuminata caumontensis</i>
	<i>C. propecornuta</i>
	<i>Lox. eggeriana</i>
	<i>Lox. linearis linearis</i>

Les trois premières appartiennent à la première association, les suivantes à la seconde. Comme le coefficient de corrélation tient compte à la fois de la présence

et du nombre d'individus de chaque espèce on peut raisonnablement affirmer que les deux associations sont antagonistes, sur la totalité de la coupe; chacune contient son cortège d'espèces dominantes et d'espèces accessoires.

2.1.4. Valeur du coefficient de corrélation cophénétique entre la matrice du coefficient de corrélation et celle de Jaccard

L'analyse de groupe appliquée aux mêmes données mais à partir du coefficient de Jaccard, conduit à l'obtention d'un dendogramme différent de celui de la figure 21.

Il est particulièrement intéressant de suivre la position des espèces dominantes au sein des associations du nouveau dendogramme (fig. 22). Toutes (à l'exception de *P. solida*) se regroupent au sein d'une même association. C'est un résultat contraire à celui obtenu à partir du coefficient de corrélation.

Une comparaison effectuée entre les seules espèces dominantes par la méthode cophénétique (cf. 1^{re} partie, paragraphe 4.2.) permet de mieux apprécier les différences.

La matrice cophénétique conduit à une valeur de corrélation cophénétique très faible et négative: $-0,340$. Cette valeur traduit une distorsion très accentuée entre les deux groupes de résultats. Cependant, elle ne permet pas de choisir lequel des deux dendogrammes traduit le plus fidèlement les liaisons qui existaient au sein des biocoénoses.

L'emploi du coefficient de corrélation permet de conserver l'intégralité des données (présence de l'espèce et nombre d'individus). A l'inverse, le coefficient de Jaccard ampute les mêmes données de leur contenu quantitatif: à ce titre il assimile les espèces abondantes et rares à une même valeur, de ce fait il ne peut être d'un emploi universel.

2.2. ASSOCIATION(S) D'ESPÈCES ET PARAMÈTRE PALÉOÉCOLOGIQUES

Un exemple suffira à montrer l'intérêt que l'on peut tirer de la recherche du taux de liaison entre les associations d'espèces (ici les ostracodes) et des paramètres écologiques.

2.2.1. Liaison entre la salinité et les morphotypes de *Limnocythere pernesensis* CARBONNEL, 1972

G. CARBONNEL a attiré l'attention en 1972 sur l'existence de morphotypes lisses et ornés chez *Limnocythere pernesensis* dans les séries oligocènes du bassin d'Apt (fig. 8). Il a également montré la présence de tubercules pouvant apparaître sur ces morphotypes ce qui aboutit aux combinaisons suivantes: morphotype orné (o),

morphotype orné tuberculé (ot), morphotype lisse (l), morphotype lisse tuberculé (lt). Les sédiments ayant livré ces divers morphotypes fournissent également des représentants du genre *Neocyprideis*. L'analyse de groupe, par la méthode R, permet

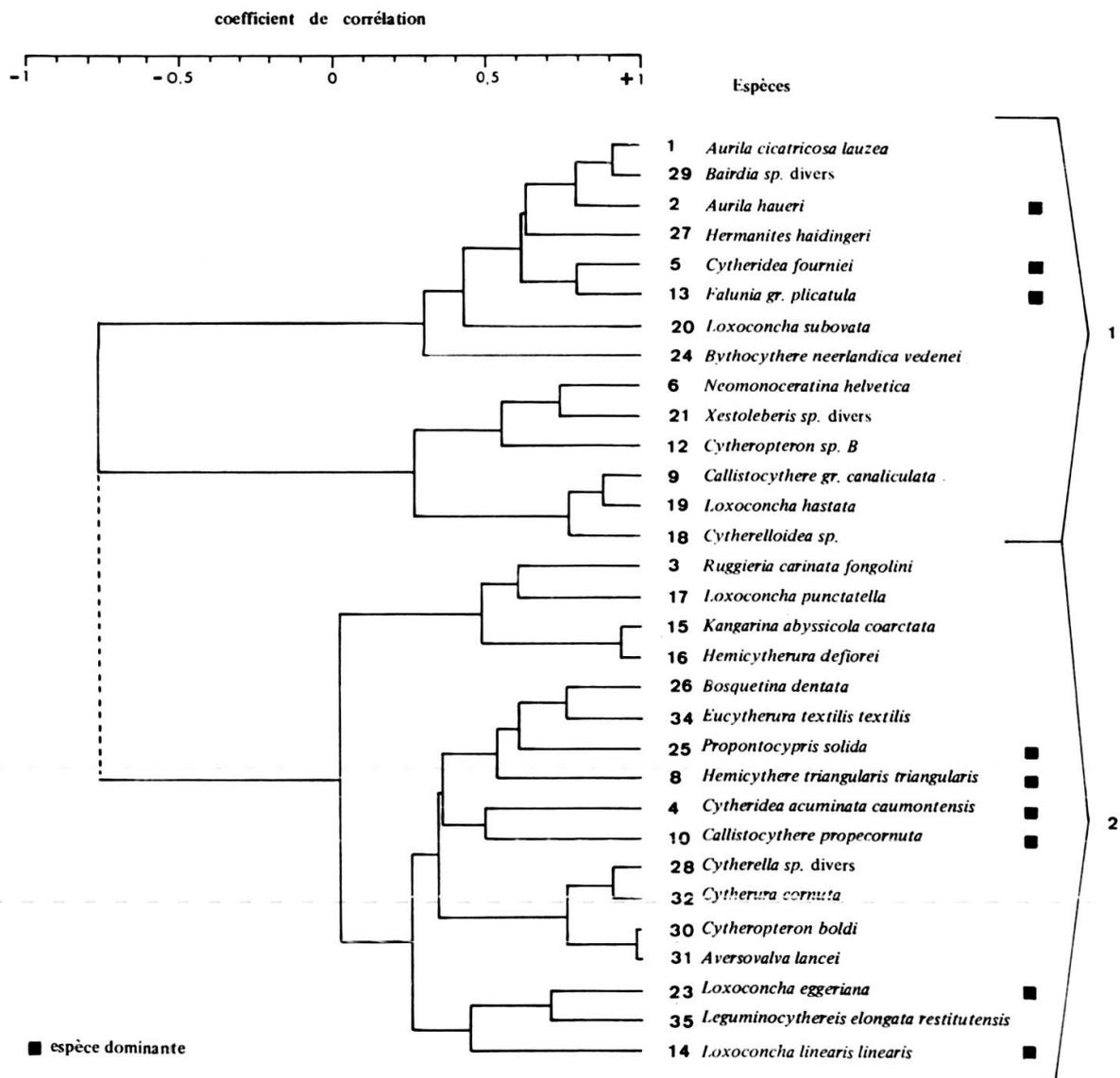


FIG. 21. — Dendrogramme résultant du groupement des ostracodes de la totalité de la coupe de Caumont-Picabrier (corrélacion r, WPGM).

alors l'étude du taux de liaison entre ces morphotypes et la présence, l'absence ou la dominance du genre *Neocyprideis*. Ce dernier, réputé saumâtre, constitue une référence de la chlorinité du milieu et joue ainsi le rôle d'un paramètre paléocéologique

Le dendrogramme de la figure 23 (extrait de G. CARBONNEL, 1972) obtenu à partir du coefficient de Jaccard traduit le degré de liaison entre les morphotypes et la chlorinité.

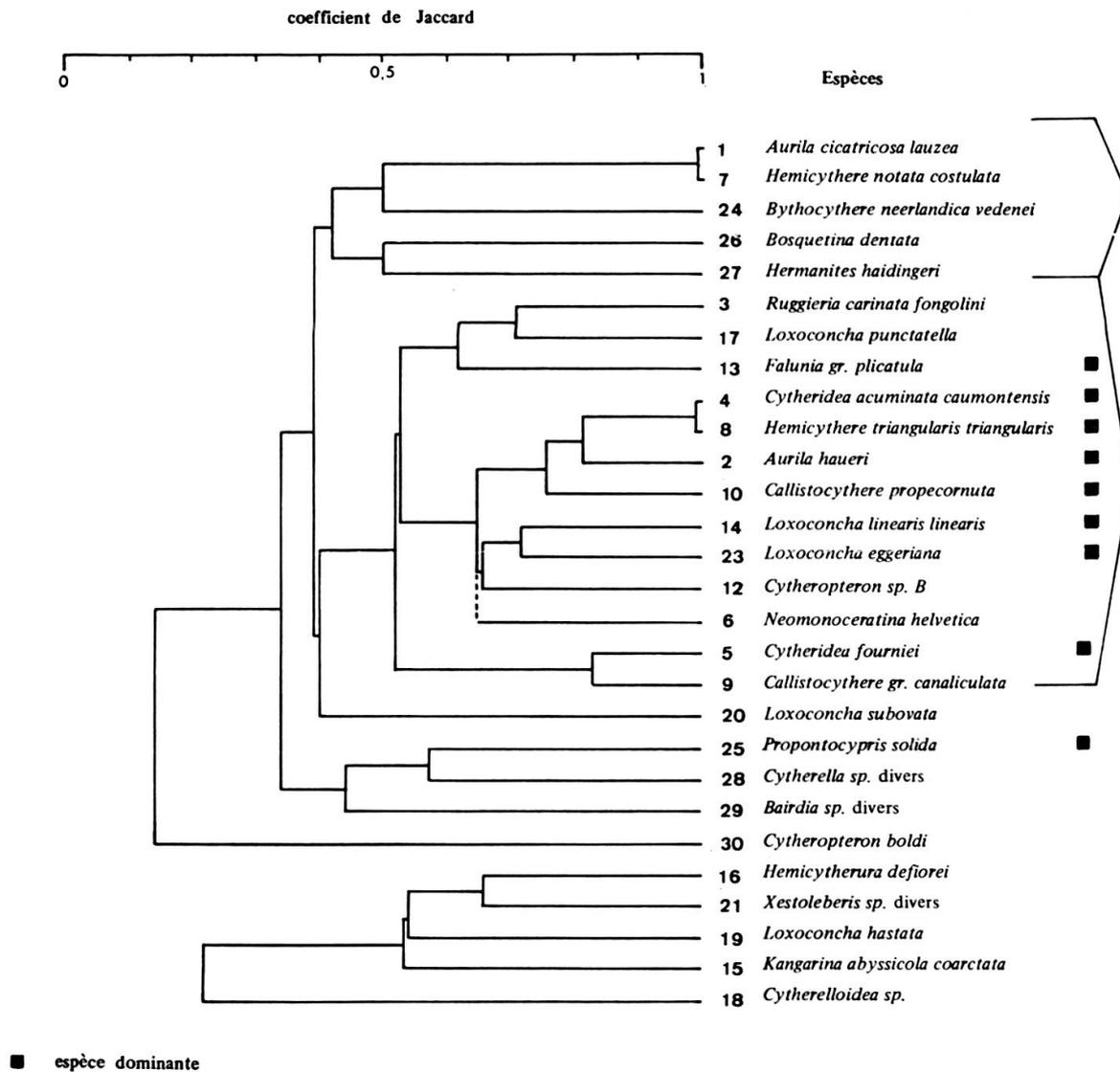


FIG. 22. — Dendrogramme résultant du groupement des ostracodes de la totalité de la coupe de Caumont-Picabrier (Jaccard, WPGM).

On constate que le morphotype lisse est associé à une population parmi laquelle le genre *Neocyprideis* n'est pas prédominant. Le morphotype lisse n'est donc pas inféodé à une chlorinité réduite.

l'établissement de corrélation biostratigraphique entre des coupes souvent très éloignées. L'interprétation des dendogrammes dans ce sens est très délicate. En effet si l'on travaille à une échelle très fine ceux-ci traduisent plus facilement les influences paléoécologiques que la succession stratigraphique des faunes. Ainsi malgré une marge d'erreur certaine, seule l'analyse de groupe permet d'assurer des corrélations à une aussi grande échelle. C'est là une des voies les plus spectaculaires et les plus enrichissantes que puisse procurer cette nouvelle méthode.

On peut également par l'analyse suivant le mode Q, apprécier l'influence du lithofaciès ou de la distribution géographique. Dans l'éventualité d'une liaison positive entre ces paramètres et les niveaux à ostracodes, ces derniers se groupent préférentiellement entre eux en respectant les affinités géographiques ou les lithofaciès. Contrairement à certains préjugés hâtifs, c'est seulement à ce stade que les paramètres invoqués pour rendre compte des dendogrammes sont susceptibles d'être pris en considération.

A l'avenir on aura la possibilité de ne pas se limiter à ces deux paramètres.

3.1.2. *Analyse de groupe selon le mode R. Corrélation des espèces, biocénoses*

Les associations obtenues sont souvent différentes suivant que l'on s'adresse à la totalité ou à une fraction de la coupe. Si l'on tient compte des renseignements fournis par l'analyse selon le mode Q, les ostracodes correspondant aux périodes de stabilité paléoécologique sont l'expression d'une ou de plusieurs biocénoses. A l'inverse, durant les périodes d'instabilité paléoécologique les associations traduisent à la fois les biocénoses et les thanatocénoses.

3.2. RÉSULTATS LOCAUX

3.2.1. *Résultats stratigraphiques*

— au Miocène inférieur (biozone A de CARBONNEL, 1969).

L'analyse suivant le mode Q a permis de préciser la position relativement basse de la coupe de Caumont-Picabrier et celle relativement élevée de la Fongoline et de Beaumont-les-Valettes.

— au Miocène supérieur (biozone C et D de CARBONNEL, 1969).

Mise en évidence des lacunes de sédimentation.

La même analyse souligne l'absence de sédimentation (voir fig. 17) correspondant à la coupe de Cucuron supérieur dans les bassins plus septentrionaux de Valréas, Faucon-Mollans. Cette constatation implique une durée plus longue de la biozone D dans le bassin sud du Lubéron que dans les autres bassins. Cette tendance est déjà esquissée dans la partie supérieure des coupes plus septentrionales (*cf.* fig. 15).

3.2.2. Résultats paléogéographiques

Les dendogrammes des niveaux (mode Q) montrent l'importance du facteur géographique. Cela est particulièrement sensible même pour des bassins aussi rapprochés que ceux d'Avignon-Cavaillon et de Faucon-Mollans.

3.3.3. Résultats paléoécologiques

D'une part il suffira de rappeler (*cf.* 2^e partie, 2.1.2.), durant les périodes de stabilité paléoécologique de la coupe de Caumont-Picabrier l'existence des biocénoses à:

1. *A. haueri*, *N. helvetica*, *H. haidingeri*
2. *C. fournieri*, *Lox. punctatella*
3. *Cytheropt. sp. B*, *Propont. solida*

D'autre part l'analyse de l'influence du lithofaciès sur les associations laisse penser que la granulométrie (ou l'énergie du milieu?) plus encore que le pourcentage de carbonate de calcium est le facteur déterminant du moins pour la répartition des ostracodes de la coupe de St-Restitut. Toutefois quelques-uns semblent liés, dans une certaine mesure, au faciès calcaréo-argileux. Il s'agit de:

- | | |
|---|---|
| 1. <i>Aurila cauditoides</i> | 9. <i>Flexus triebeli</i> |
| 2. <i>Costa edwardsii</i> | 10. <i>Hemicytherura defioerei</i> |
| 3. <i>Cuneocythere ariminensis</i> | 11. <i>Loxoconcha eggeriana</i> |
| 4. <i>Cytherella sp. divers</i> | 12. <i>Loxoconcha grignanensis</i> |
| 5. <i>Cytherelloidea sp. divers</i> | 13. <i>Loxoconcha linearis linearis</i> |
| 6. <i>Echinocythereis scabra</i> | 14. <i>Loxoconcha subvota</i> |
| 7. <i>Eucytherura textilis textilis</i> | 15. <i>Sclerochilus montbrisoni</i> |
| 8. <i>Falunia gr. plicatula</i> | |

BIBLIOGRAPHIE

On se référera à l'ouvrage de R.R. SOKAL et P. H. A. SNEATH pour une bibliographie complète traitant en particulier l'aspect mathématique de l'analyse de groupe.

- ANDREIEFF P., R. ANGLADA *et al.* (1971). Contribution à l'étude de l'Aquitanién. La coupe de Carry-le-Rouet (Bouches-du-Rhône, France). *Mém. Bur. Rech. géol. min.*, Paris, (sous presse).
- BRETSKY S. S. (1970). Phenetic and phylogenetic classifications of the *Lucinidae* (*Mollusca, Bivalvia*). *Bull. geol. Instn Univ. Upsala*, n.s., vol. 2, n° 2, pp.5-23, 2 pl., 4 fig., 5 tabl.
- BRETSKY S. S. (1971). Evaluation of the efficacy of numerical taxonomic methods: an example from the bivalve mollusks. *System. Zool.*, Washington, vol. 20, n° 2, p. 204-222, 2 fig., 5 tabl.
- CARBONNEL G. (1969). Les ostracodes du Miocène rhodanien. Systématique, biostratigraphie écologique, paléobiologie. Thèse. *Docum. Lab. Géol. Fac. Sci. Lyon*, n° 32, fasc. 1-2, 469 p., 48 fig., 57 tabl.

- CARBONNEL G. (1972). Le facteur lisse chez certains ostracodes tertiaires: un index de paléotempérature. Symposium Newark, (sous presse), 3 fig., 3 tabl., 2 pl.
- CHEETHAM A. H. et J. E. HAZEL (1969). Binary (presence-absence) similarity coefficients. *Journ. Paleontol.*, Tulsa, vol. 43, n° 5, pp. 1130-1136.
- HAZEL J. E. (1970). Binary coefficients and clustering in biostratigraphy. *Geol. Soc. Amer. Bull.*, New-York, v. 81, pp. 3237-3252, 7 fig., 1 tabl.
- JACOBZONE M. et G. CARBONNEL (1971). Coefficient de Jaccard et coefficient de corrélation. Application aux ostracodes miocènes. *Bull. Centre Rech. Pau-SNPA*, vol. 5 suppl., pp. 167-177, 4 fig., 1 tabl.
- KAESLER R. L. (1966). Quantitative re-evaluation of ecology and distribution of recent foraminifera and ostracoda of Todos Santos Bay, Baja California, Mexico. *Paleont. Contr. Kansas*, pap. 10, 50 p., 23 fig., 14 tabl.
- KAESLER R. L. (1967). Numerical taxonomy in invertebrate paleontology. Essays in Paleontology and stratigraphy, Raymond C. Moore Commemorative volume. *Dept. geol. spec. publ.*, Kansas, Lawrence, n° 2, pp. 63-81.
- KAESLER R. L. (1969). Numerical taxonomy of selected recent british Ostracoda. in The taxonomy, morphology and ecology of recent ostracoda, *Oliver and Boyd édit.*, Edinburgh, pp. 21-47, 14 fig., 3 tabl.
- KAESLER R. L. (1969). Ordination and character correlations of selected recent british Ostracoda. *Math. Geol. Univ. Kansas*, vol. 1, n° 1, pp. 97-111, 5 fig., 4 tabl.
- KAESLER R. L. (1969). Aspects of quantitative distributional paleoecology. *Comp. Applic. Earth Sc.*, New-York, pp. 99-120, 8 fig., 1 tabl.
- MADDOCKS R. F. (1966). Distribution patterns of living and subfossil podocopid ostracodes in the Nosy-Bé area, northern Madagascar. *Paleont. Contr.*, Topeka, Kansas, pap. 12, 72 p., 63 fig.
- MICHENER C. D. et R. R. SOKAL (1957). A quantitative approach to a problem in classification. *Evolution*, Lawrence, vol. II, pp. 130-162.
- PHILIPPE R. (1972). Le Burdigalien (Miocène) du plateau des Courennes (Vaucluse). Etude paléontologique et stratigraphique. *Docum. Lab. Géol. Fac. Sci. Lyon*, n° 51, 275 p., 50 fig., 12 tabl.
- ROHLF F. J. et R. R. SOKAL (1965). Coefficients of correlation and distance in numerical taxonomy. *Univ. Kansas Sc. Bull.*, Lawrence, vol. 45, n° 1, pp. 3-27, 12 fig., 2 tabl.
- SNEATH P. H. A. et S. T. COWAN (1958). An electro taxonomic survey of bacteria. *J. gen. microbiol.*, London, vol. 19, pp. 551-565.
- SOKAL R. R. et F. J. ROHLF (1962). The comparison of dendograms by objective methods. *Taxon*, Utrecht, vol. II, n° 2, pp. 33-39, 2 fig., 2 tabl.
- SOKAL R. R. et P. H. A. SNEATH (1963). Principles of numerical taxonomy. *Freeman and Co, édit.*, San Francisco, 359 p.
-

