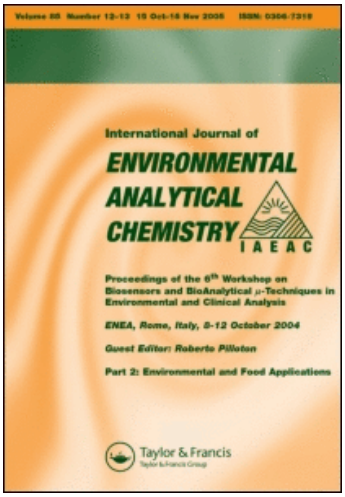

International Journal of Environmental Analytical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713640455

Assessment of Organic Source Contributions in Coastal Waters by Principal Component and Factor Analysis of the Dissolved and Particulate Hydrocarbon and Fatty Acid Contents

Joan O. Grimalt[a]; Joaquim Olivé[a]; Josep I. Gómez-Belinchon[a]
[a] Department of Environmental Chemistry, (CID-CSIC), Barcelona, Spain

To cite this Article Grimalt, Joan O. , Olivé, Joaquim and Gómez-Belinchon, Josep I.(1990) 'Assessment of Organic Source Contributions in Coastal Waters by Principal Component and Factor Analysis of the Dissolved and Particulate Hydrocarbon and Fatty Acid Contents', International Journal of Environmental Analytical Chemistry, 38: 2, 305 — 320
To link to this Article: DOI: 10.1080/03067319008026934
URL: http://dx.doi.org/10.1080/03067319008026934

Intern. J. Environ. Anal. Chem., Vol. 38, pp. 305–320
Reprints available from the publisher
Photocopying permitted by license only

ASSESSMENT OF ORGANIC SOURCE CONTRIBUTIONS IN COASTAL WATERS BY PRINCIPAL COMPONENT AND FACTOR ANALYSIS OF THE DISSOLVED AND PARTICULATE HYDROCARBON AND FATTY ACID CONTENTS

JOAN O. GRIMALT, JOAQUIM OLIVÉ and JOSEP I. GÓMEZ-BELINCHON

Department of Environmental Chemistry (CID-CSIC), Jordi Girona, 18–26, 08034-Barcelona, Spain

Principal component (PCA) and factor analysis (FA) are evaluated for the interpretation of the information contained in large datasets resulting from the study of environmental samples by gas chromatography (GC) and GC coupled to mass spectrometry (GC-MS). A case involving the identification and quantitation of 64 variables (hydrocarbons and fatty acids) in 87 water samples (dissolved and particulate fractions) of a coastal system (Ebre Delta) has been selected for examination.

PCA has evidenced important differences between the dissolved and particulate materials, as well as between the particulates collected in the bays and those obtained in the river and channels. PCA has also allowed the identification of outlier samples in the dissolved fraction. Independent application of FA to each of these groups has provided a useful method for the characterization of diverse algal, terrestrial, microbial and anthropogenic inputs. Direct correspondences between these source inputs and factor loadings have provided a selection of representative components of each contribution in the coastal system.

KEY WORDS: Hydrocarbons, fatty acids, water particulates, water dissolved fraction, principal component analysis, factor analysis, coastal waters.

INTRODUCTION

The use of chromatographic methods for the analysis of lipid components in aquatic environments currently provides extended multivariate datasets containing a wealth of information on the system under study. From these data, handling, unravelling and interpretation are generally simplified by selection of limited groups of components for examination. Experience from previous studies based on the "molecular marker" concept[1–6] facilitate useful criteria for the choice of components. However, in this type of approach, most of the results rely on the portion of data which "a priori" is expected to be interpretable, being other possibly relevant remaining information easily underestimated. Thus, although interesting conclusions from the geochemical and environmental standpoints can

be obtained, it may always be questioned whether or not bias is introduced at the step of compound selection.

In this situation, the application of multivariate data handling methods seems to be adequate for the assessment of the representativity of the compounds selected. However, these techniques have been rarely considered in studies of lipid components in recent environments. Furthermore, in some of the few applications,[7–9] the limited number of samples analyzed precludes any generalization of the results obtained.

In the context of our current investigations on the inputs of organic matter in the Ebre River Delta,[10–14] we have undertaken a study of the aquatic lipid material (dissolved and particulate) and multivariate analysis methods have been systematically used for data interpretation. The results corresponding to the composition of hydrocarbons and fatty acids are presented here. They encompass the quantitation of 64 compounds in 87 water samples, including the dissolved and particulate fractions and representing a dataset of 5,568 values. A preliminary description of the composition of these samples have been reported elsewhere.[10] This paper is devoted to the application of principal component analysis (PCA) and factor analysis (FA) to the interpretation of the results.

MATERIAL AND METHODS

A detailed description of the water sampling system is given elsewhere.[15] In summary, it encompasses a priming circuit, a non-contaminating Teflon impeller pump, a 142 mm d. stainless steel filter holder and a sorption column. The water particulate fraction is collected on a microfibre filter (rated pore size 0.5 um) and then the organic dissolved phase is adsorbed on Amberlite XAD-2 resin contained in a 20 cm × 2.7 cm i.d. glass cylinder. Collection of particles larger than 1 mm is prevented by placing a stainless steel grid at the inlet of the pumping tube.

Volumes of about 100 liters were collected for each sample at a flow rate of 500 ml/min. According to previous investigations[15] these sampling conditions, and particularly the total sampled volume, provide consistent results with those obtained by liquid-liquid extraction for hydrocarbons and fatty acids in seawater. Filters were changed whenever rates dropped below 50% of normal.

The procedures for lipid isolation and fractionation are described in detail elsewhere.[10] In short, filters were Soxhlet-extracted with (1:2) methanol-methylene chloride and the Amberlite columns were introduced into a continuous percolation refluxing apparatus[16] and extracted with (9:1) acetone-water for four hours. All filter and column extracts were hydrolyzed with 6% KOH/MeOH and the neutral and acidic fractions were successively recovered by *n*-hexane extraction, the latter after acidification (pH 2) with aq. HCl 6N. The acidic fractions were esterified with a solution of BF_3 in methanol. The hydrocarbons were separated from the neutrals by column chromatography according to a previous established method.[17]

Gas chromatographic analysis (GC) was performed with a Carlo Erba FTV 4160 GC instrument, equipped with a flame ionization detector and a splitless injector. A column of 25 m × 0.25 mm i.d. coated with SE-54 was used and

hydrogen was the carrier gas. Gas chromatography-mass spectrometry (GC-MS) was carried out with a Hewlett-Packard 5995 instrument coupled to an HP 300 data system. The chromatographic conditions were the same as described above except that helium was used as carrier gas. Further details on the instrumental operating conditions are given elsewhere.[10,15]

The resolved components of the hydrocarbon fraction were quantitated by comparison with an external standard mixture of *n*-C_{14}, *n*-C_{22}, *n*-C_{32} and *n*-C_{36}. The unresolved GC envelopes were measured by planimetry and quantitated by reference to *n*-C_{22}. All these operations were carried out semi-automatically using a Hewlett-Packard 86 microprocessor equipped with a digital planimeter. Fatty acids were quantitatively determined with reference to a mixture of *n*-heptadecanoate and *n*-heneicosanoate methyl esters.

RESULTS AND DISCUSSION

Sampling Sites and Lipid Composition

According to our previous experience in the area[10-14] seven sampling sites have been selected for the study of the dissolved and particulate hydrocarbons and fatty acids. These include (see Figure 1) the main river flow, one adjacent channel (irrigation), one channel draining into Alfacs Bay (the largest) and the two bays, where surface and bottom waters ($\sim$3 m deep) have been collected for the study of their stratified water columns. Sampling has been carried out every two months covering a period of one and a half years. The samples obtained are listed in Table 1.

GC and GC-MS of the hydrocarbon and fatty acid fractions have allowed the identification and quantitation of 64 components (33 and 31, respectively). These are listed in Table 2 where their concentration ranges, means and standard deviations for the dissolved and particulate phases are also indicated. The mean values for *n*-alkanes and UCM are of the same order of magnitude as preliminary data previously reported by us in the same area.[13] These concentrations of hydrocarbons and fatty acids are also of the same order of magnitude as those found in other coastal zones.[18-23]

In general terms the concentration values are higher in the particulate phase. This trend contrasts with the regularly larger amount of organic carbon in the dissolved phase[24,25] and suggests that the particulate organic carbon (POC) is enriched in labile components with respect to the dissolved organic carbon (DOC). In this sense, the comparison of the mean concentration values of Table 2 shows that this partitioning is especially important for the planktonic and terrestrial components, pointing to a primary association of most biogenic hydrocarbons and fatty acids with the particulate matter. Conversely, the anthropogenic hydrocarbons are either more uniformly distributed.

On the other hand, the diverse concentration ranges observed for each of these components compel data standardization. Thus, the mean values of Table 2 have been subtracted from their corresponding component concentration set and the

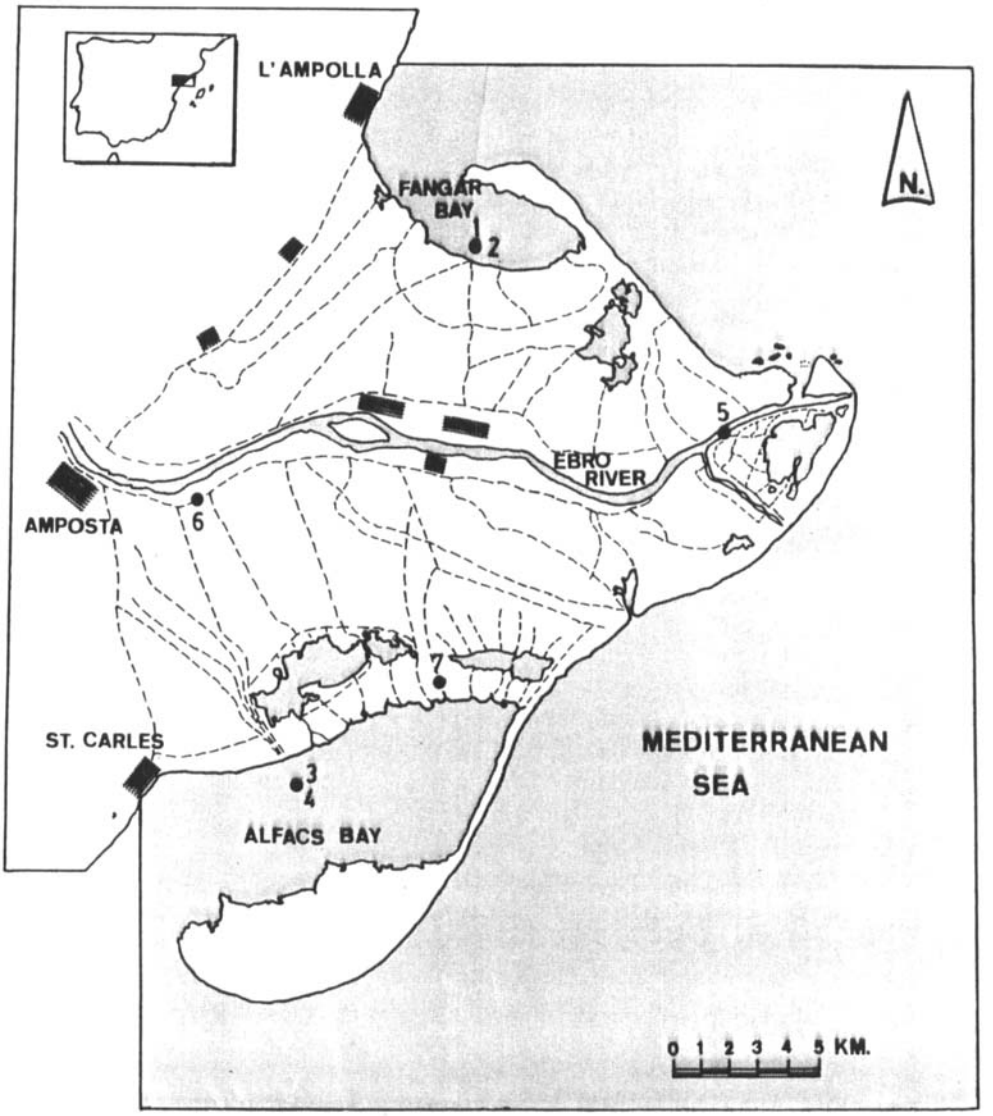

Figure 1 Map describing the location of the sampling sites selected in the Ebre Delta. 1) Fangar Bay. Surface water (brackish). 2) Fangar Bay. 3 meter deep water (marine). 3) Alfacs Bay. Surface water (brackish). 4) Alfacs Bay. 3 meter deep water (marine). 5) River bed. Surface water (fresh). 6) Irrigation channel (freshwater). 7) Drainage channel (freshwater).

results have been divided by the standard deviation. Therefore, the variance-covariance matrix to be calculated from these observations is equivalent to the correlation matrix and this is what will be considered for multivariate analysis in the present study.

Sample Grouping and Outliers

One of the first aspects to be considered in most multivariate analysis methods

Table 1 List of samples collected in the Ebre Delta (D, dissolved material; P, particulate fraction). Sampling sites refer to Figure 1.

Sampling site	*Mar 86*	*May 86*	*Jul 86*	*Sep 86*	*Nov 86*	*Jan 87*	*Mar 87*	*May 87*
1	D+P	D+P	D+P	D+P	D+P	D+P	D+P	P
2	D+P	D+P	D+P	D+P	D+P	D+P	D+P	P
3		D+P	D+P	D+P	D+P	D+P	D+P	
4		D+P	D+P	D+P	D+P	D+P	D+P	
5	D+P	D+P	D+P	D+P	D+P	D+P	D+P	P
6		D+P	D+P	D+P	D+P			P
7		D+P	D+P	D+P	D+P			P

refers to the uniformity of the population of samples for examination. At this respect, straightforward application of PCA to the samples listed in Table 1 shows important differences between the group of dissolved and particulate samples. This is illustrated in Figure 2 by representation of the first three component scores corresponding to this calculation. In both diagrams (PC2 vs PC1 and PC3 vs PC2) the values of the samples corresponding to the dissolved fraction cluster in a narrow area, whereas those from the particulates are widely spread. These results reflect the higher temporal and spatial variability of the particulates and suggest that, in principle, both groups of samples should be analyzed separately.

The application of PCA only to the particulate fraction evidence additional major unhomogeneities. This is again illustrated in Figure 2 from the diagrams of the first three component scores. The samples corresponding to particulate material collected in the bays cluster in a limited area showing no significant differences between surface and bottom waters. In contrast, the samples from the river (5) and channels (6, 7) extend over the whole graph. The differences are, however, less defined than in the previous case so that the contrast between dissolved and particulate samples is in fact higher than the variability of particulate material from different locations, although a higher range of variation in the molecular lipid contents of the particulates from the river and channels with respect to that of the bays is observed. Independent multivariate analysis of the two groups of samples is again recommended.

PCA of the dissolved fraction shows a different type of problem in relation with homogeneity. This is illustrated in Figure 3 where sample projections of the corresponding three principal components are plotted. One sample (irrigation channel, September 86) outlies with respect to PC2. When this is excluded from the calculations a rather uniform distribution can be observed in relation with this second component.

The presence of this outlier sample is also reflected in variations of the composition of the principal components (see Figure 4). In terms of environmental significance, the main change concerns the different contribution behaviour of *n*-heptadecane (variable No. 5), *n*-heneicosane (No. 11) and *n*-tricosane (No. 13). When all dissolved samples are included these three *n*-alkanes show a distinct trend with respect to the other *n*-alkane homologs. They contribute to PC2 and are depleted in PC1. A close examination of the hydrocarbon composition of the

Table 2 List of lipid components determined in the waters of the Ebre Delta. (concentration in ng/l; 0 = below 0.01 ng/l for hydrocarbons and 0.1 ng/l for fatty acids).

	Dissolved				*Particulated*			
	Min	*Max*	*Mean*	*S.D.*	*Min*	*Max*	*Mean*	*S.D.*
1. *n*-tetradecane	0	0.63	0.024	0.098	0	0.30	0.015	0.056
2. *n*-pentadecane	0	2.6	0.082	0.41	0	3.1	0.13	0.53
3. *n*-hexadecane	0	2.2	0.094	0.35	0	1.5	0.10	0.29
4. norpristane	0	0.097	0.016	0.025	0	0.5	0.026	0.090
5. *n*-heptadecane	0	10	0.44	1.7	0	93	6.0	18
6. pristane	0	1.7	0.093	0.26	0	3.0	0.50	0.82
7. *n*-octadecane	0	1.6	0.13	0.24	0	12	1.2	2.1
8. phytane	0	0.32	0.039	0.052	0	4.7	0.60	0.97
9. *n*-nonadecane	0.017	0.70	0.12	0.14	0	18	1.8	3.4
10. *n*-eicosane	0.012	0.98	0.14	0.20	0	23	2.3	4.1
11. *n*-heneicosane	0.028	13	0.60	2.0	0	92	5.7	14
12. *n*-docosane	0.066	1.7	0.31	0.32	0	22	3.1	4.1
13. *n*-tricosane	0.047	7.6	0.50	1.2	0	72	6.8	12
14. *n*-tetracosane	0.045	1.9	0.34	0.40	0.2	58	5.8	9.3
15. *n*-pentacosane	0.033	2.3	0.41	0.48	0.3	54	7.7	9.6
16. *n*-hexacosane	0.054	2.3	0.39	0.44	0.4	44	5.7	7.6
17. *n*-heptacosane	0.070	3.7	0.74	0.82	0.4	64	12	16
18. *n*-octacosane	0.073	3.8	0.66	0.79	0.2	42	4.9	7.0
19. *n*-nonacosane	0.11	8.3	1.5	1.7	0.5	170	25	41
20. *n*-triacontane	0.067	7.4	1.0	1.5	0.1	31	4.1	5.7
21. *n*-hentriacontane	0.10	10	1.4	1.9	0.4	100	13	19
22. *n*-dotriacontane	0.083	8.0	0.96	1.5	0.1	15.5	2.4	3.0
23. *n*-tritriacontane	0.072	8.7	1.1	1.7	0.2	19	4.1	4.4
24. *n*-tetratriacontane	0.046	6.2	0.71	1.1	0	31	2.2	4.8
25. *n*-pentatriacontane	0.013	3.5	0.41	0.56	0	6.5	1.3	1.6
26. *n*-hexatriacontane	0	2.3	0.35	0.54	0	5.7	0.74	1.2
27. *n*-heptatriacontane	0	1.2	0.19	0.26	0	1.2	0.18	0.30
28. *n*-octatriacontane	0	0.17	0.013	0.035	–	–	–	–

29. *n*-nonatriacontane	0	0.11	0.008	0.020	–	–	–	–
30. UCM*	5.2	205	35	39	0	1200	190	270
31. Ftalate	0	8.1	0.70	1.5	0.2	32	5.6	6.6
32. Squalene	0.20	31	3.2	6.1	0	18	1.9	3.4
33. long-chain alkylbenzenes	0	66	1.9	10	0	25	0.72	3.8
34. methyl *iso*-tetradecanoate	0	4.0	0.32	0.95	0	260	57	70
35. methyl *n*-tetradecenoate	0	18	3.2	3.6	0	300	26	60
36. methyl *n*-tetradecanoate	1.0	170	62	48	88	16000	2400	3800
37. methyl *iso*-pentadecanoate	0	7.6	3.1	1.9	0	830	270	230
38. methyl *anteiso*-pentadecanoate	0	11	1.9	2.1	0	650	140	130
39. methyl *n*-pentadecanoate	0.1	33	13	7.8	1.0	5000	300	760
40. methyl *iso*-hexadecanoate	0	99	9.1	18	0	4100	440	920
41. total methyl *n*-hexadecenoates	0.9	230	71	62	170	60000	5500	11000
42. methyl *n*-hexadecanoate	1.5	940	180	170	250	50000	6500	10000
43. methyl *iso*-heptadecanoate	–	–	–	–	0	180	12	34
44. methyl *anteiso*-heptadecanoate	–	–	–	–	0	250	17	50
45. methyl *n*-heptadecanoate	0	33	4.3	6.1	0	1700	95	260
46. total methyl *n*-octadecatrienoates	0.2	410	98	110	15	6600	950	1400
47. methyl *n*-octadecadienoate	0.4	210	34	35	15	2700	550	650
48. methyl *n*-octadec-9-enoate	0.7	4000	230	610	64	17000	1500	3300
49. methyl *n*-octadec-11-enoate	0	64	14	14	61	2800	1000	600
50. methyl *n*-octadecanoate	0	80	27	17	22	5100	700	860
51. methyl *n*-nonadecanoate	0	5.9	0.90	1.6	0	645	39	110
52. methyl *n*-cosapentaenoate	1.3	760	130	150	40	16000	1500	3300
53. methyl *n*-eicosanoate	0	5.4	1.2	1.7	0	1100	40	160
54. methyl *n*-heneicosanoate	0	3.5	0.29	0.68	0	445	31	94
55. methyl *n*-docosahexaenoate	0.3	260	65	74	0	4200	600	1000
56. methyl *n*-docosanoate	0	180	7.1	28	0	1000	83	190
57. methyl *n*-tricosanoate	0	13	1.3	2.9	0	610	35	98
58. methyl *n*-tetracosanoate	0	115	6.3	19	0	1200	250	280
59. methyl *n*-pentacosanoate	0	36	4.4	7.8	0	340	17	54
60. methyl *n*-hexacosanoate	0	82	5.6	15	0	500	29	82
61. methyl *n*-heptacosanoate	0	11	0.86	2.5	0	62	2.3	10
62. methyl *n*-octacosanoate	0	100	7.0	22	0	435	20	72
63. methyl *n*-nonacosanoate	0	8.9	0.74	1.9	0	34	1.0	5.3
64. methyl *n*-triacontanoate	0	130	6.2	22	0	79	4.7	17

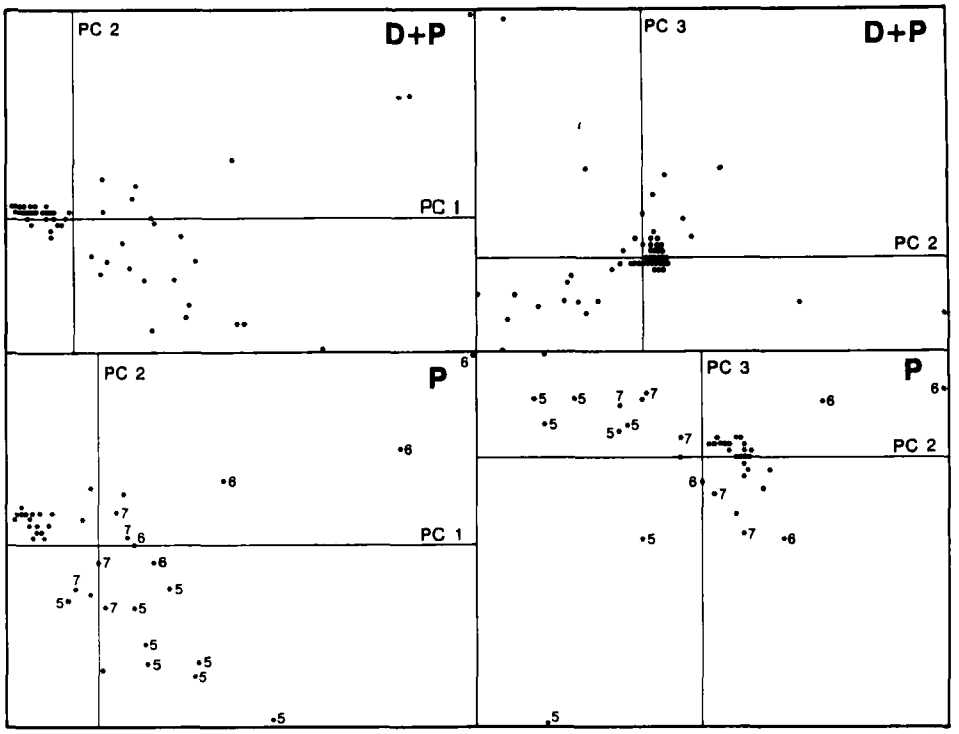

Figure 2 Principal component scores of the PCA study corresponding to the dissolved (·) and particulate (x) phases (D + P), and to the particulate (x) phase only (P). Numbers refer to the stations in Figure 1.

sample collected at the irrigation channel in September 86 showed that this was constituted by high amounts of an unusual distribution of *n*-alkanes and *n*-alkenes reflecting a specific algal input. The high concentration of these hydrocarbons determine the influence of this sample over PC2 which, as it will be discussed below, essentially reflects algal contributions.

The exclusion of this outlier sample from the calculations results in a contribution trend of these three hydrocarbons (Nos. 5, 11 and 13) similar to that observed for the rest of *n*-alkanes (see Figure 4). This corresponds to a more representative description of the pricipal components of the dissolved fraction of the delta. However, in this latter case, two samples outlie with respect to PC1 (River Bed, March 86 and Fangar Bay, bottom, March 86) and it can also be questionned whether they should be excluded. We have carried out PCA calculations without inclusion of these two samples and no major differences were observed between the resulting principal components and those represented in Figure 4 (D-1 case). Accordingly, in the multivariate analyses for the characterization of the source inputs in the dissolved fraction of the Delta waters only the sample collected at the irrigation channel in September 86 has been excluded (D-1 case).

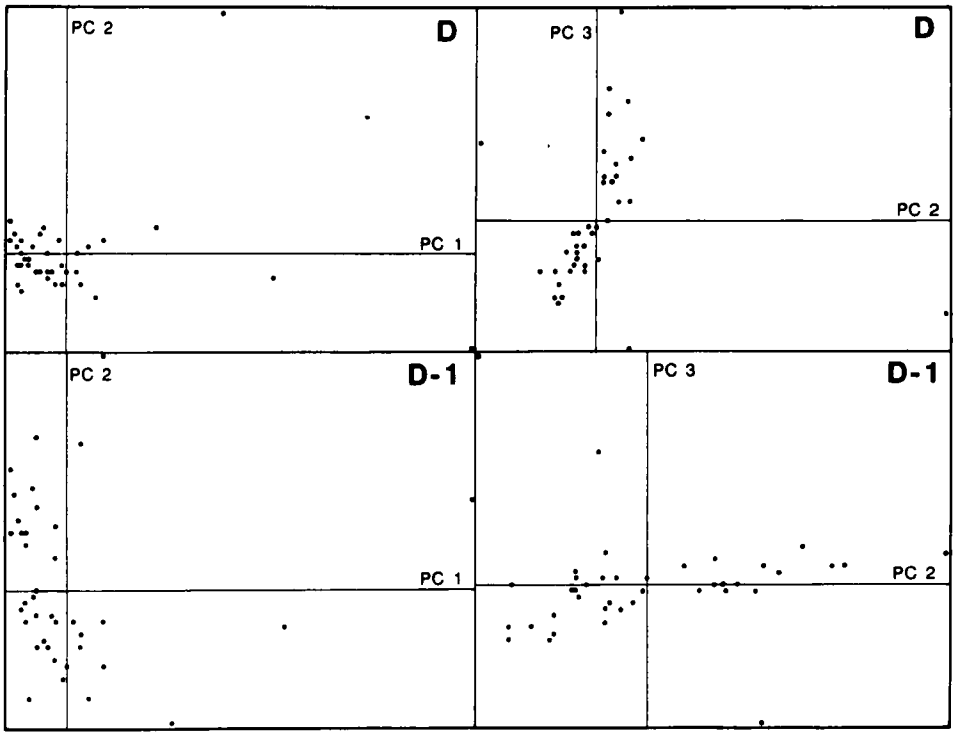

Figure 3 Principal component scores of the PCA study corresponding to the dissolved fraction of the waters from the Ebre Delta (D). D-1 correspond to the scores resulting from the exclusion of an outlier sample (irrigation channel, September 86).

Source Input Characterization

The eigenvalues of the three first principal components corresponding to the dissolved fraction (D-1 case; Figure 4) represent up to 63% of the total variance (PC1 37%, PC2 16% and PC3 10%). For the description of about 90%, 95% and 99% of the total variance of the dataset eleven, fourteen and twenty-two components are respectively required. Similar results are obtained when performing PCA on the particulate fraction. The dispersion of the total variance between so many principal components is a consequence of the high variability of hydrocarbons and fatty acids in environmental samples, especially in water samples. This problem is currently observed provided that a reasonable number of samples is included in the calculations. Studies involving small sample numbers have resulted in descriptions of high percentages of the total variance with a limited number of principal components. Obviously, in these cases the results are not likely to be representative of the system under study.

This important dispersion effect suggests that source input assessment may be easier using multivariate methods devoted to the description of the portion of the total variance that is shared by the variables. Accordingly, FA has been applied in this study. For the particulates, two separate calculations have been performed,

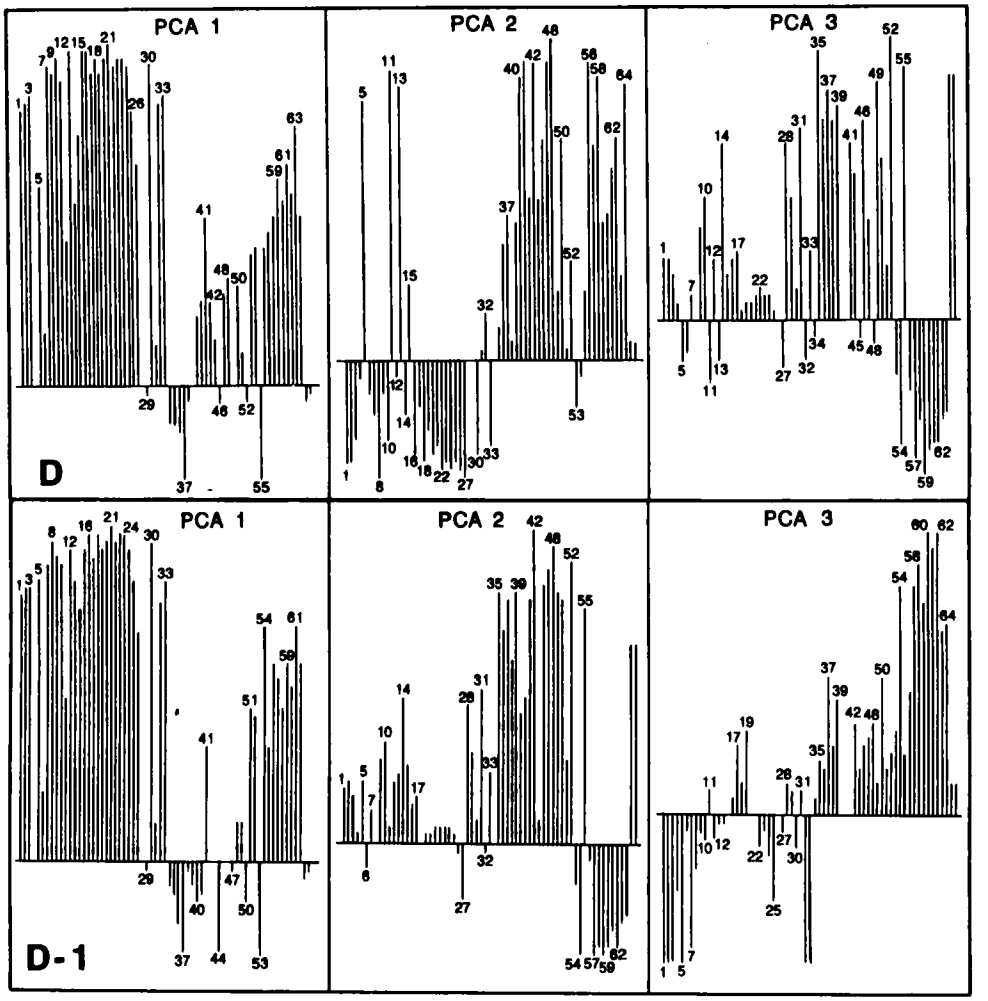

Figure 4 Representation of the principal components resulting from the PCA study of the dissolved fraction (D). D-1 correspond to the principal components obtained after exclusion of an outlier sample (irrigation channel, September 86). Numbers refer to Table 2.

one corresponding to the samples collected at the bays and the other to the samples obtained in the river and the channels.

Dissolved Phase

The three main varimax rotated factor loadings obtained in the FA study of the dissolved fraction (D-1 case) are represented in Figure 5. The former factor is loaded by most of the *n*-alkanes (variables Nos. 1–3, 5, 7, 9–27 in Table 2), the UCM (No. 30), the LAB (No. 33) and phytane (No. 8), representing a set of components essentially related to anthropogenic contributions. The concurrence of these *n*-alkanes with the UCM reflects their preferential association with petro-

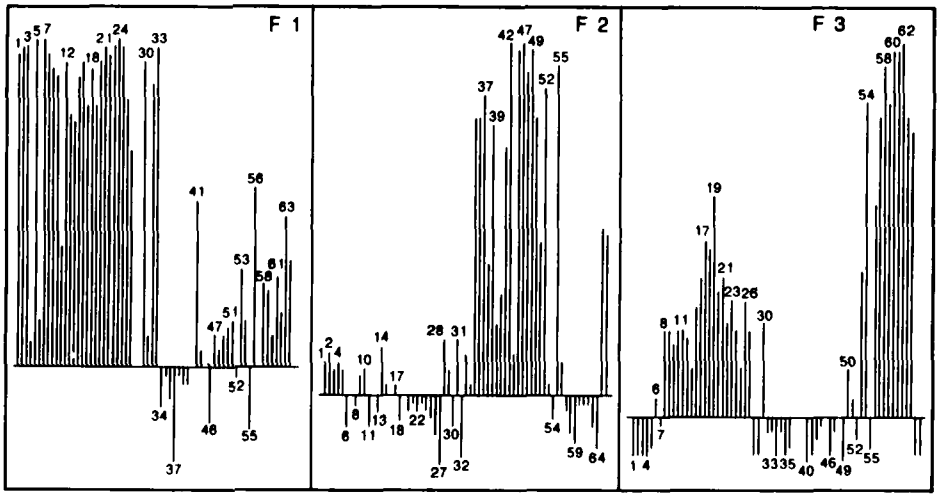

Figure 5 Representation of the three main factor loadings resulting from the FA study of the samples from the dissolved fraction. One outlier (irrigation channel, September 86) is excluded. Numbers refer to Table 2.

genic inputs. The lower loading of *n*-heneicosane (No. 11) may correspond to a problem of quantitation because this hydrocarbon coelutes with 2,6,10,14-tetramethyl-7-(3-methylpentyl)-pentadecatriene,[13] a compound of biogenic origin that occurs in some of the analyzed samples. The high loading of phytane (No. 8) is in agreement with its known petrogenic origin in recent environments. On the other hand, the absence of pristane (No. 6) corresponds to an occurrence related to biogenic (zooplankton) sources. Finally, the high loading of the LAB (No. 33), an indicator of urban effluents,[26] points to a geographical coincidence of the urban and petrogenic sources in the delta. In this respect, it is interesting to note that squalene (No. 32), an indicator of anoxic bacterial activity, is also heavily loading on this factor which suggests a possible production by bacterial decomposition of domestic sewage.

The second factor is loaded by a series of C_{14}–C_{22} fatty acids (Nos. 35–37, 39, 41, 42, 46–52 and 55) representing autochthonous biogenic inputs. The high load of *n*-eicosapentaenoic (No. 52) and *n*-docosahexaenoic (No. 55) acids, two polyunsaturated fatty acids which may originate from phyto-[27] and zooplankton,[28] represents an interesting feature. The absence of other zooplankton markers in this group, such as pristane[29] suggests their association with algal production. Bacterial inputs are not represented significantly in this factor. There is only an important load of *iso*-pentanoic acid (No. 37), but the contribution of other *iso*- and *anteiso*-acids is very small.

The major loadings of the third factor correspond to the C_{21}–C_{30} *n*-alkanoic acids (Nos. 54, 56–64). These are common constituents of the cuticular waxes of higher plants[30] and their occurrence in coastal environments is associated with terrestrial inputs.[31] In addition to these components, there is a significant load of a group of C_{25}–C_{33} *n*-alkanes (Nos. 15–23) which is characterized by the higher proportion of the odd carbon numbered homologs. C_{25}–C_{33} odd carbon

numbered *n*-alkane distributions are indicative of contributions from higher plants.[32] The occurrence of this group of *n*-alkanes in factor 3 indicates that a portion of the hydrocarbons in the dissolved fraction correspond to terrestrial vegetation and, consistently, is found loading in this third factor. In short, the three main varimax rotated factor loading of the dissolved fraction (D-1 case) indicate that the variance of hydrocarbons and fatty acids is affected by anthropogenic (petroleum + urban) (F1), algal (F2) and higher plant (F3) inputs.

Particulate Phase

The application of FA to the particulate samples collected in the bays have resulted in the identification of five distinct source inputs, the corresponding factor loads are displayed in Figure 6. The first factor is essentially loaded by a series of C_{14}–C_{22} saturated and unsaturated fatty acids (Nos. 34–37, 39, 41, 42, 46–50, 52, 55 and 57), representing a distribution of algal origin.[22] The second factor is mainly loaded by a group of C_{27}–C_{37} *n*-alkanes (Nos. 17–27). As described for the dissolved fraction, odd-to-even carbon number predominated *n*-alkanes in the C_{25}–C_{37} range are indicative of contributions from higher plants. The high load of both even and odd carbon numbered hydrocarbons involves that, independently of their chain length, the *n*-alkanes of this range are related to terrestrial sources. In this situation, due to the effect of variable standardization, all these *n*-alkanes must load equally independently of their concentration in the source material.

The third factor is loaded by a modal mixture of C_{21}–C_{26} *n*-alkanes (Nos. 11–16). *n*-Alkane distributions is this range lacking odd-to-even carbon number may be produced by bacterial activity,[33–35] although in some cases they correspond to specific petroleum spillages.[36] The low load of the UCM (No. 30) supports a bacterial origin for these *n*-alkanes. The UCM is, in contrast, loading heavily in factor four, along with pristane (No. 6), phytane (No. 8) and the C_{15}–C_{20} *n*-alkanes (Nos. 2, 3, 5, 7, 9 and 10). These features, especially the high load of the UCM and phytane indicate that this factor describes the petrogenic contributions to the bays. Finally, factor five is loaded by diverse *iso* and *anteiso* fatty acids (Nos. 40, 43 and 44), *n*-eicosanoic acid (No. 53) and the polyunsaturated octadecanoic acids (Nos. 46 and 47). The high load of branched fatty acids suggest that factor five also represents microbial contributions.[37,38] However, the absence of C_{21}–C_{26} *n*-alkanes indicates that the bacterial inputs described by both factors are related to distinct origins.

The four main factor loadings obtained in the FA study of the particulate samples collected at the main river flow and in the channels are displayed in Figure 7. Factor one is loaded by a series of C_{14}–C_{22} fatty acids (Nos. 36, 39–42, 45–48, 52, 54 and 55) that again correspond to a mixture of algal origin. The high load of *n*-pentadecane (No. 2), *n*-hexadecane (No. 3) and *n*-heptadecane (No. 5) represents a distinct feature in comparison with the factor loadings related with algal contributions described above. Factor two is mainly loaded by diverse C_{17}–C_{29} *n*-alkanes (Nos. 7–10, 12, 14–19) and the UCM (No. 30). In principle, the high load of the UCM relates F2 with petrogenic contributions. However, the presence of *n*-nonacosane (No. 19) and *n*-hentriacontane (No. 21) in higher

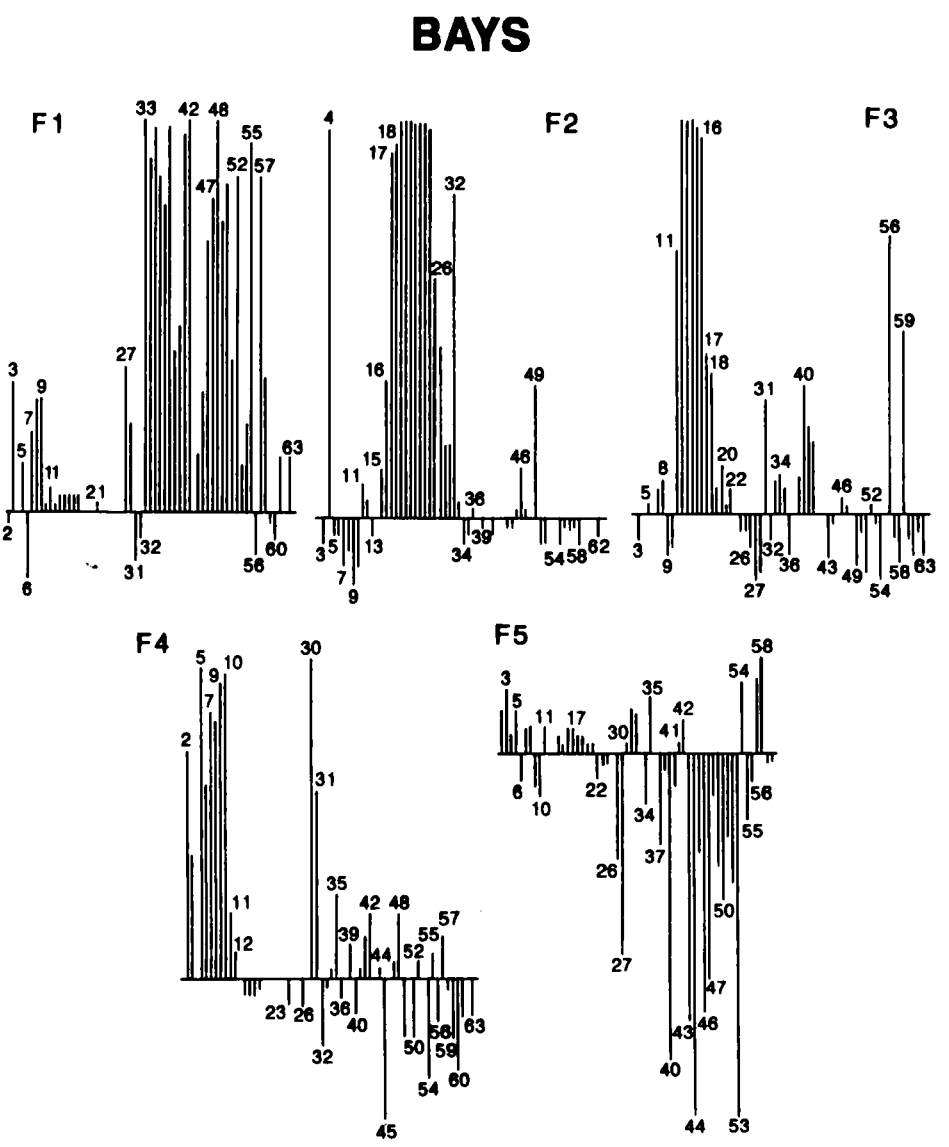

Figure 6 Representation of the five main factor loadings resulting from the FA study of the particulate samples collected at the bays. Numbers refer to Table 2.

proportion than *n*-octacosane (No. 18) or *n*-triacontane (No. 20) indicates that terrestrial contributions are also reflected in this factor. Higher plant waxes are however better represented in factor three, which is highly loaded by C_{22}–C_{28} even carbon number fatty acids (Nos. 56, 60 and 62). Nevertheless, bacterial components are also important since *iso-* and *anteiso-* heptadecanoic acids are present in high proportion (Nos. 43 and 44). Finally, factor four is loaded by a series of hydrocarbons (Nos. 5, 11, 13, 16, 24 and 32) and fatty acids (Nos. 34, 45, 48, 51

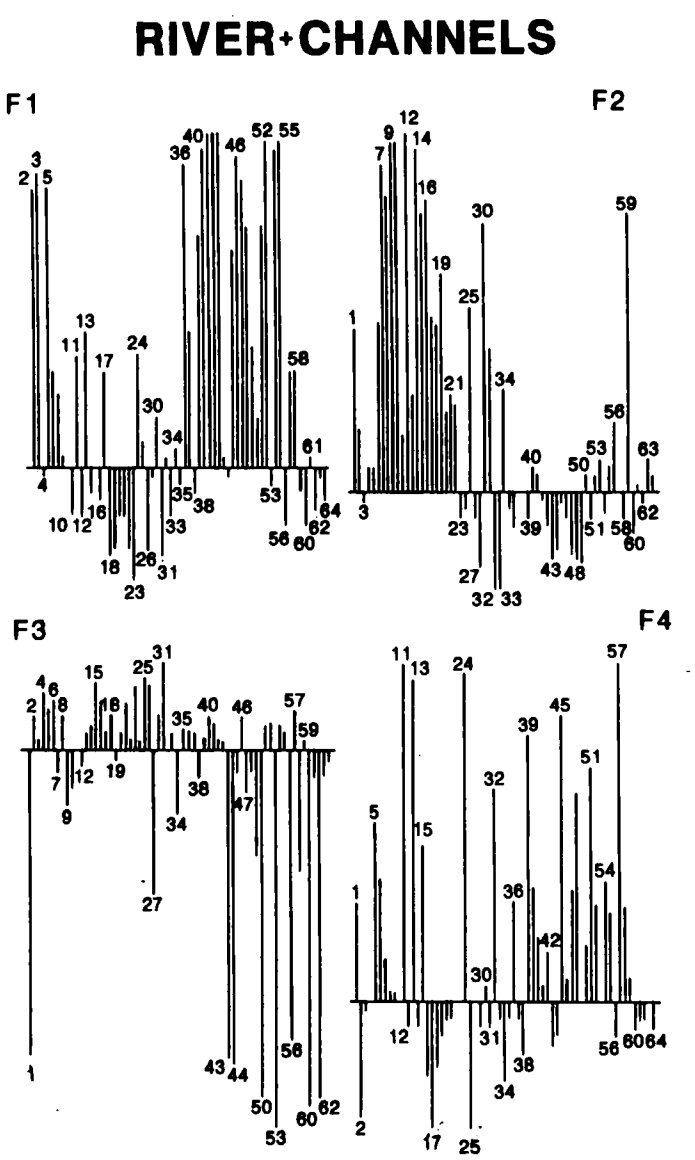

Figure 7 Representation of the four main factor loadings resulting from the FA study of the particulate samples collected at the river and the channels. Numbers refer to Table 2.

and 57) that, as described in the preceeding section, essentially correspond to a specific input from determined algal species.

An important feature of the application of FA to the study of the particulates from the river and channels is that the correspondence between source inputs and main factor loadings is not so closely established as in the previous case. Repeated calculations, each using a different number of factors, have allowed to exclude factor size as responsible. Conversely, the problem may be related with the lack of

uniformity of the population of samples. Thus, PCA of the particulate fraction already revealed that the samples from the irrigation channel and the main river flow show different distribution trends when representing PC1, PC2 and PC3 scores (see Figure 2). The degree of definition of the factor loadings is likely limited by this lack of uniformity, the global analysis is compelled to generalization when the composition of diverse samples is only affected by specific sources. This type of problem is currently observed in situations where the variables largely outnumber the samples, as it is the common case in organic geochemistry.

CONCLUSIONS

FA is a useful technique for the assessment of organic source inputs in aquatic environments. Direct correspondences between factor loadings and source inputs can be obtained providing a selection of components that, in the system under study, are representative of each contribution. These correspondences represent an easy way of simplification of the information contained in the large datasets usually obtained by chromatographic analysis of environmental samples. However, FA may also give rise to biased results. In all cases, a preliminary evaluation of the population of samples for unhomogeneous grouping and outliers is needed. This can be performed be means of PCA, provided that the dataset selected for study is composed by a sufficient number of samples.

Accordingly, in the present study of hydrocarbons and fatty acids in coastal waters, PCA has evidenced that important differences may be observed between the dissolved and the particulate fraction, and between the particulates collected in the bays and those obtained from the river and channels. PCA has also allowed the identification of outlier samples in the dissolved fraction.

Independent application of FA to these three groups of samples has allowed the characterization of the following organic source contributions:

a) *Dissolved phase*: F1) Anthropogenic (petroleum + urban) inputs: C_{14}–C_{34} *n*-alkanes, UCM, LAB and phytane. F2) Algal contributions: Diverse C_{14}–C_{22} saturated and unsaturated fatty acids. F3) Higher plant wax materials: C_{22}–C_{30} *n*-alkanoic acids.

b) *Particulates from the bays*: F1) Algal inputs: C_{14}–C_{23} *n*-alkanoic and *n*-alkenoic acids. F2) Terrestrial contributions: C_{25}–C_{37} *n*-alkanes. F3) Microbial products: C_{21}–C_{26} *n*-alkanes. F4) Petrogenic contamination: C_{15}–C_{20} *n*-alkanes, UCM, pristane and phytane. F5) Other bacterial components not correlated with F3: Some linear (*n*-$C_{21:0}$), branched (C_{16} and C_{17}) and unsaturated ($C_{18:3}$ and $C_{18:2}$) fatty acids.

c) *Particulates from the river and channels*: F1) Algal contributions: C_{14}–C_{22} saturated and unsaturated fatty acids along with C_{15}–C_{17} *n*-alkanes. F2) Mixed anthropogenic and terrestrial inputs: C_{17}–C_{29} *n*-alkanes and the UCM. F3) Mixed terrestrial and microbial contributions: C_{22}–C_{28} even carbon number fatty acids and *iso* and *anteiso*-heptadecanoic acids. F4) Specific input from a determined group of algal species.

Acknowledgements

We thank Prof. Joan Albaigés (Department of Environmental Chemistry, CID-CSIC) for his helpful comments. One of us (JIGB) thanks CSIC for a post-doctoral fellowship. Financial support from EROS-2000 is also acknowledged.

References

1. R. B. Johns, *Biological Markers in the Sedimentary Record, Methods in Geochemistry and Geophysics* (Elsevier, Amsterdam, 1986).
2. A. Farran, J. Grimalt, J. Albaigés, A. V. Botello and S. A. Macko, *Mar. Pollut. Bull.* **18,** 284 (1987).
3. J. Grimalt, J. M. Bayona and J. Albaigés. In: *VIèmes Journées d'études sur les pollutions marines en Mediterranée* CIESM, Luzerne 1984, 533 (1986).
4. J. J. Boon, J. W. de Leeuw, J. W. Hoek and V. d. G. Vosjan, *J. Bacteriol.* **129,** 1183 (1977).
5. S. C. Brassell and G. Eglinton. In: *Advances in Organic Geochemistry 1981* (M. Bjoroy *et al.*, ed.) (John Wiley, Chichester, 1983), pp. 684 (1983).
6. L. Chuecas and J. P. Riley, *J. Mar. Biol. Assoc. UK.* **49,** 97 (1969).
7. P. M. Shaw and R. B. Johns, *Org. Geochem.* **10,** 951 (1986).
8. P. M. Shaw and R. B. Johns, *Org. Geochem.* **9,** 237 (1986).
9. M. A. Sicre, J. L. Paillasseur, J. C. Marty and A. Saliot, *Org. Geochem.* **12,** 281 (1988).
10. J. I. Gomez-Belinchon, R. Llop, J. O. Grimalt and J. Albaigés, *Mar. Chem.* **25,** 325 (1988).
11. J. M. González, J. Grimalt and J. Albaigés, *Mar. Chem.* **14,** 61 (1983).
12. R. W. Risebrough, B. W. de Lappe, W. Walker II, B. R. T. Simoneit, J. Grimalt and J. Albaigés, *Mar. Pollut. Bull.* **14,** 181 (1983).
13. J. Albaigés, J. Grimalt, J. M. Bayona, R. Risebrough, B. de Lappe and W. Walker II, *Org. Geochem.* **6,** 237 (1984).
14. J. O. Grimalt, J. I. Gomez-Belinchon, R. Llop and J. Albaigés, *Chemosphere* **17,** 1893 (1988).
15. J. I. Gómez-Belinchon, J. O. Grimalt and J. Albaigés, *Environ. Sci. Technol.* **22,** 677 (1988).
16. M. Ehrhardt, *ICES Techniques in Marine Environmental Sciences* No. 14 (1987).
17. M. Aceves, J. O. Grimalt, J. Albaigés, F. Broto, L. Comellas and M. Gassiot *J. Chromatogr.* **436,** 503 (1988).
18. Chr. Osterroht and G. Petrick, *Mar. Chem.* **11,** 55 (1982).
19. K. A. Burns and J. P. Villeneuve, *Geochim. Cosmochim. Acta* **47,** 995 (1983).
20. S. A. Macko, J. K. Winters and P. L. Parker, *Mar. Pollut. Bull.* **13,** 174 (1982).
21. B. W. de Lappe, R. W. Risebrough and W. Walker II, *Can. J. Fish. Aquat. Sci.* **40,** 322 (1983).
22. G. Kattner, G. Gercken and K. Eberlein, *Mar. Chem.* **14,** 149 (1983).
23. G. Kattner, G. Gercken and K. O. Hammer, *Mar. Chem.* **14,** 163 (1983).
24. R. A. Kerr and J. G. Quinn, *Mar. Chem.* **8,** 217 (1980).
25. R. Pocklington and F. C. Tan, *Geochim. Cosmochim. Acta* **51,** 2579 (1987).
26. R. P. Eganhouse, D. L. Blumfield and I. R. Kaplan, *Environ. Sci. Technol.* **17,** 523 (1983).
27. R. G. Ackman, C. S. Tocher and J. McLachlan, *J. Fish Res. Board Can.* **25,** 1603 (1968).
28. J. R. Sargent and R. F. Lee, *Mar. Biol.* **31,** 15 (1975).
29. M. Blumer, M. M. Mullin and D. W. Thomas, *Science* **140,** 974 (1963).
30. P. E. Kolattukudy, *Lipids* **5,** 259 (1970).
31. J. W. Farrington and J. W. Quinn, *Geochim. Cosmochim. Acta* **37,** 259 (1973).
32. G. Eglinton and R. J. Hamilton, *Science* **156,** 1322 (1967).
33. R. W. Johnson and J. A. Calder, *Geochim. Cosmochim. Acta* **37,** 1943 (1973).
34. J. Grimalt, J. Albaigés, A. A. Z. Douabul and H. T. Al-Saad, *Naturwissenschaften* **72,** 35 (1985).
35. J. Grimalt, E. Torras and J. Albaigés, *Org. Geochem.* **13,** 741 (1988).
36. C. Lee, S. G. Wakeham and J. W. Farrington, *Mar. Chem.* **13,** 181 (1983).
37. G. J. Perry, J. K. Volkman, R. B. Johns and H. J. Bavor Jr., *Geochim. Cosmochim. Acta* **43,** 1715 (1979).
38. J. K. Volkman, R. B. Johns, F. T. Gillan, G. J. Perry and H. J. Bavor, *Geochim. Cosmochim. Acta* **44,** 1133 (1980).