

Cluster Analysis in Lipid Biomarker Studies: A Case of the Clyde Sea (Analisis Berkelompok dalam Kajian Penanda Bio Lipid: Kes di Laut Clyde)

MASNI MOHD ALI & STEPHEN M. MUDGE

ABSTRAK

Tiga puluh tiga sampel sedimen permukaan telah diambil di persekitaran Laut Clyde dan telah dianalisis bagi menentukan kandungan asid lemak, alkohol lemak dan sterol dengan menggunakan Kromatografi Gas-Spektrometri Jisim (KGSJ). Analisis berkelompok (kaedah Ward dengan pemalar korelasi) yang dilakukan secara berasingan ke atas sebatian asid lemak, alkohol lemak dan sterol ini menunjukkan sebatian-sebatian ini dikelompok mengikut sumber geokimia masing-masing. Walaubagaimanapun, dalam setiap analisis, sebatian-sebatian terbitan dari terestrial dapat dilihat berkelompok dengan sebatian-sebatian terbitan marin dan bakteria. Sebatian-sebatian terbitan dari bakteria boleh dijumpai dalam semua kelompok bagi analisis berkelompok yang dijalankan ke atas campuran ketiga-tiga kumpulan sebatian. Ini menunjukkan bahawa input daripada bakteria hadir bersama dengan input daripada marin dan terestrial.

Kata kunci: analisis berkelompok; penanda bio; asid lemak; alkohol lemak; sterol

ABSTRACT

Thirty-three surface sediment samples were collected around the Clyde Sea and were analyzed for fatty acids, fatty alcohols and sterols by the GC-MS detection. Cluster analysis (Ward's method with correlation coefficient) performed separately on fatty acids, fatty alcohols and sterols showed that they were grouped according to their geochemical sources. However, terrestrial derived compounds can be seen grouping with marine and bacterial derived compounds in every analysis. Bacterial derived compounds can be found in all clusters in the mixed compound cluster analysis showing that bacterial inputs occurred together with marine and terrestrial inputs.

Keywords: cluster analysis; biomarker; fatty acids; fatty alcohols; sterols

INTRODUCTION

Cluster analysis is a numerical technique for defining groups of related samples based on high similarity coefficients, computed between each pair of samples, which are then clustered. The main purpose of cluster analysis is to reduce the mass of data, without significant loss of information, to a tractable size, but in a way that emission patterns can be easily discerned. Thus, in most clustering procedures the nucleus of clusters (centroid) is formed by joining the samples with highest similarity and gradually admitting more samples as the similarity coefficient is lowered (Kaiser & Esterby 1991).

The Clyde Sea is a deep, partially enclosed basin, on the west coast of Scotland. The Clyde Sea communicates with the adjacent shelf sea only via exchange flow over a relatively shallow sill (depth of 45 m). The basin receives large inputs (60-70 m³s⁻¹) of freshwater from the River Clyde and other river sources, which enters mainly through the sea loch system to the north and from rivers along the Ayrshire coast (Poodle 1986).

Lipids diversity and specificity makes them useful compounds to study organic matter sources in marine environments. Therefore their characterisation, both at individual and group level, has been one of the main objectives in marine organic geochemistry (Wakeham and Lee, 1989). The analysis of fatty acids, fatty alcohols and sterols from the surface sediment samples of the Clyde Sea and cluster analysis will be discussed in this paper. Cluster analysis was employed to classify the study area into specific region, each having definite characteristics.

MATERIALS AND METHODS

Thirty-three samples were collected from the Firth of Clyde, Kilbrannan Sound, Gareloch, Loch Long, Loch Goil, Loch Striven, Loch Riddon and Loch Fyne in Scotland (Figure 1). A grab sampler from *RV Prince Madog* was used to collect these samples and scraped into glass jars.

Subsamples of the sediments were dried in a drying cabinet (60 °C). They were reweighed daily, until a constant

by electrons and they are separated according to their mass:charge (m/z) ratio. Masslab, which is a GC-MS program, was used to record and display the mass spectra.

The capillary column used for fatty alcohols and sterols was a non-polar, bonded phase BPX-5 (SGE) with 60 m length, 0.22 internal diameter and 0.25 μm film thickness. The temperature program used started at 60 °C, increasing at 15 °C min^{-1} to 300 °C, then at 5 °C min^{-1} to a maximum of 350 °C for 10 minutes. The MS was configured for electron impact ionisation of 70eV and a mass scan range of 45-585 m/z per second.

A BPX-70 (SGE) column was used for fatty acid methyl esters. It is a very polar, bonded phase column with a 50 m length, 0.33 mm internal diameter and 0.25 μm film thickness. The temperature program started at 80 °C, increasing at 40 °C min^{-1} to 160 °C, then at 0.5 °C min^{-1} to 170 °C and finally at 10 °C min^{-1} to a maximum of 250 °C for 8 minutes.

Fatty alcohols were quantified using a C18-TMS fatty alcohol standard and cholesterol-TMS was used to calibrate the sterols. Both of them were used as external standards and the calibration curves were obtained. Tricosanoic acid methyl ester (23:0) was used as an internal standard for the fatty acid methyl ester samples.

Standard method and techniques were adopted throughout. The efficiency of the whole extraction process was confirmed by repeat reflux of some sediment samples; no further fatty alcohols, sterols and fatty acids could be detected in these later extractions. Blanks and calibration standards were used throughout the GC injections. A blank was injected first and followed by the calibration standard. Five samples were injected afterwards, and followed by the blank and calibration standard again.

Random samples were extracted three times to test the reproducibility of the extraction. The mean and standard deviation of fatty acid 18:1 ω 7, saturated alcohol 18:0 and cholesterol from sample 1 in the Clyde Sea are $15236 \pm 416 \text{ ng g}^{-1}$, $582 \pm 46 \text{ ng g}^{-1}$ and $1768 \pm 100 \text{ ng g}^{-1}$ respectively. Therefore, the reproducibility of the extraction was found to be better than 90% for the three classes of compound. Procedural blanks were also analysed and no compounds of interest were measured in any sample. All glassware and Teflon-lined caps used in these analyses were rinsed with organic solvents prior to work.

CLUSTER ANALYSIS

Cluster analysis was carried out separately on fatty acids, fatty alcohols, sterols and mixed compounds (combination of fatty acids, fatty alcohols and sterols) using a Minitab statistical program (Release 12) (1996) suggests Ward's clustering Ward's clustering procedure with correlation coefficient distance was used throughout the analysis as it yields a larger proportion of correct classified observations. All the data was transformed into proportion with 0.001 added and then log transformed.

RESULTS AND DISCUSSION

GC-MS analysis of the samples led to the identification of 35 fatty acids, 25 fatty alcohols and 20 sterols in the sediments. The results are presented by compound class initially and then as combined markers for the clustering purposes.

FATTY ACIDS

Fatty acids are often the most abundant lipid type in sediments because they are abundant in most organisms. Therefore they can be used to differentiate sources that contribute to the sedimentary lipid. A total of 35 fatty acids were identified with concentrations ranged from 4 to 47710 ng g^{-1} dry weight. 16:0 and 18:0 fatty acids were the most two abundant compounds found from the 33 sampling sites.

Algae are a major source of fatty acids in most marine sedimentary environments. Homologous series of short chain saturated fatty acids ($<C_{20}$) are considered to originate from algae (Carrie et al. 1998, Rohjans et al. 1998). 16:0 and 18:0 fatty acids are commonly assigned to planktonic sources (Volkman et al. 1998). Polyunsaturated fatty acids are normally associated with phytoplankton. Chlorophyta (green algae) contain abundant C_{16} and C_{18} polyunsaturated fatty acids especially with the positional isomers ω 3 and ω 6 (Volkman et al. 1989, Carrie et al. 1998). Branched fatty acids (principally *iso* and *anteiso* odd chain length compounds) are commonly encountered in bacterial lipids (Parkes, 1987; Thoumelin et al. 1997). Therefore these acids are useful indicators of bacterial lipid contribution (Thoumelin et al. 1997, Carrie et al. 1998). Long chain fatty acids are generally considered as terrestrial plant biomarkers in sediments as they occur in surface waxes of higher plants (Thoumelin et al. 1997, Mudge et al. 1998, Volkman et al. 1998). Fatty acids with chain lengths up to 28:0 were quantified in the sediment samples. The most abundant fatty acids with more than twenty carbon atoms were the saturated 24:0 and 26:0.

The output of the cluster analysis is given as a dendrogram (Figure 2). There are two major cluster groups divided into 2 smaller subgroups. The first group in cluster group I (CGI) corresponds to branched fatty acids indicating bacterial markers origin. The other subgroup consists of marine and algal derived compounds. Meanwhile long chain fatty acids are grouped together within cluster group II (CGII) associating their terrestrial origin. Saturated 16:0 and 18:0 acids were also clustered together with this group. These two components are the most abundant fatty acids in this data set and are associated to aquatic source organisms (Carrie et al. 1998, Mudge et al. 1998). These results, therefore, suggest that this cluster group was also dominated by marine derived compounds.

FATTY ALCOHOLS

Fatty alcohols in marine sediments are primarily derived from wax esters, which are supplied by a variety of organisms. The primary sources of wax esters in marine

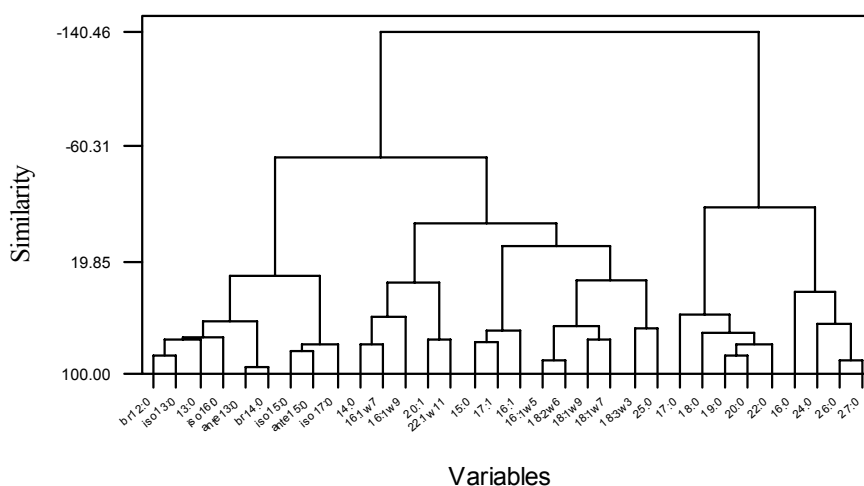


FIGURE 2. Cluster analysis of fatty acids showing the Ward's method for the Clyde Sea samples

environments are thought to be from marine zooplankton and terrestrial plants (Fukushima and Ishiwatari 1984, Mudge and Norris, 1997). Twenty-nine saturated and monounsaturated alcohols were found with concentrations ranged from 2 to 2216 ng g⁻¹ dry weight. In these samples, the C₂₂ had the highest mean concentration (864 ng g⁻¹ dry weight).

Fatty alcohols of all marine wax esters are invariably either saturated short chain alcohols especially C₁₄ and C₁₆, or monounsaturated alcohols especially 16:1, 18:1, 20:1 and 22:1 (Sargent et al. 1981, Rajendran et al. 1991). Therefore short chain alcohols are often used as marine indicators (Mudge and Norris, 1997, Mudge and Lintern, 1999). Another major source of wax esters is terrestrial higher plants. Fatty alcohols derived from terrestrial plants consist of long saturated carbon chain (>C₂₀) and they can be used to indicate terrestrial inputs (Fukushima and Ishiwatari, 1984, Mudge and Norris, 1997). Branched fatty acids and fatty alcohols usually result from bacterial metabolism (Parkes,

1987).

Figure 3 shows the dendrogram from fatty alcohols cluster analysis. The results are more complicated to interpret with marine, terrestrial and bacterial derived compounds clustered together with each other. There are 3 major cluster groups as shown in Figure 3. Cluster group I (CGI) corresponds to bacterial derived fatty alcohols with branched fatty alcohols clustered together. Greatest similarity can be seen at *iso*-C₁₇ and *anteiso*-C₁₇, suggesting that these compounds were the strongest indicators of bacterial inputs. Meanwhile cluster group II (CGII) consists of short chain fatty alcohols, therefore this cluster is associated with marine markers. Cluster group III (CGIII) is more complicated as marine, bacterial and terrestrial derived compounds are clustered in one group. Terrestrial derived compounds such as long chain fatty alcohols are also found to be clustered within cluster group I and cluster group II. On the other hand, monounsaturated fatty alcohols are grouped with cluster group I and cluster group III.

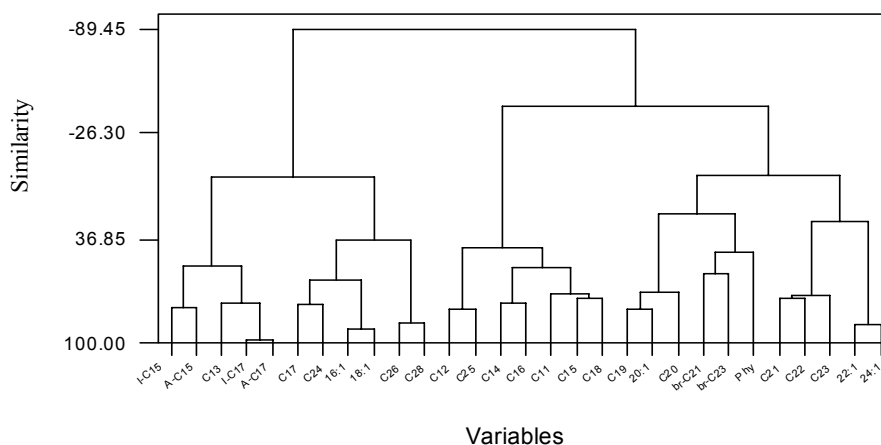


FIGURE 3. Cluster analysis of fatty alcohols showing the Ward's method for the Clyde Sea samples

STEROLS

Sterols have been used as tracers of inputs from marine and terrestrial plants and animals (Saliot et al. 1991, Mudge and Norris, 1997). Of the 20 sterols identified, cholesterol was the most abundant sterol and was present in all samples with mean concentration of 17102 ng g⁻¹ dry weight. However, there is no clear trend of cholesterol distribution throughout the sampling sites. Cholesterol originates from a variety of planktonic organisms of all trophic levels (Volkman, 1986, Volkman et al. 1987, Hinrichs et al. 1999, Mudge et al. 1999) with zooplankton being the major source.

β -sitosterol, campesterol and stigmasterol are the most common sterols in vascular plants (Volkman 1986) although they also have been found in phytoplankton. The relative distribution of these sterols has been proposed as a source marker for higher plants (Volkman 1986, Laureillard and Saliot 1993). While algae contain larger amounts of campesterol and stigmasterol, β -sitosterol dominates in vascular plants. Coprostanol (5 β -cholestan-3 β -ol) has been used as a sewage tracer in a variety of environments (Venkatesan and Kaplan 1990, Jeng and Han 1994, Leeming and Nichols 1998). Coprostanol is produced in the intestine of mammals by the bacterial transformation of cholesterol, therefore, it present in sewage contaminated waters and sediments.

The cluster analysis results of the sterols are shown in Figure 4. Altogether 4 clusters could be explained divided into two bigger cluster groups. Group I in the cluster group I (CGI) contains marine derived sterols such as st1, brassicasterol, cholesterol and dinosterol. Brassicasterol and cholesta-5,22(E)-dien-3 β -ol are the strongest indicator of marine sterols. They were clustered together with the greatest similarity. Meanwhile the second group include the terrestrial markers (stigmasterol and campesterol). The first group in cluster group II (CGII) indicates sewage derived markers such as coprostanol, epicoprostanol, st9

and cholesterol are clustered together. In the second group of cluster group II, again, the terrestrial derived compounds are clustered together within this group (β -sitosterol, st10 and C30 trienol).

MIXED COMPOUNDS

Two main cluster groups were identified from the dendrogram (Figure 5a). Within the first cluster group (CGI), there are two separate clusters. First cluster includes short chain fatty acids and fatty alcohols as well as branched fatty acids. 24 norcholesta-5,22(E)-dien-3 β -ol, cholesta-5,22(E)-dien-3 β -ol, brassicasterol and cholesterol clustered together. These compounds are abundant in marine organisms and in bacterial biomass. Meanwhile the second cluster (CGII) consists of marine derived fatty alcohols, fatty acids and sterols. It can be concluded that cluster group I contain a mixture of marine and bacterial derived compounds. The second cluster group includes the terrestrial derived compounds as well as sewage biomarkers such as coprostanol. Bacterial derived compounds can be seen in both clusters, showing that bacterial inputs occurred together with marine and terrestrial input especially for diagenesis processes.

Two main groups were distinguished in the Ward's method dendrograms for the sampling sites (Figure 5b). The first cluster group (CGI) includes samples mainly located in the open seas including samples collected from Loch Long, while the second cluster group consists of samples collected within the sea lochs especially from Loch Fyne. Generally, the cluster group I corresponds to samples characterised by marine and bacterial derived compounds such as short chain fatty acids and fatty alcohols, saturated 18:2 ω 6, cholesterol, st1, st3 and brassicasterol. The cluster group II (CGII) consists of samples collected within the lochs. These samples characterised by terrestrial and sewage derived compounds such as long chain fatty acids

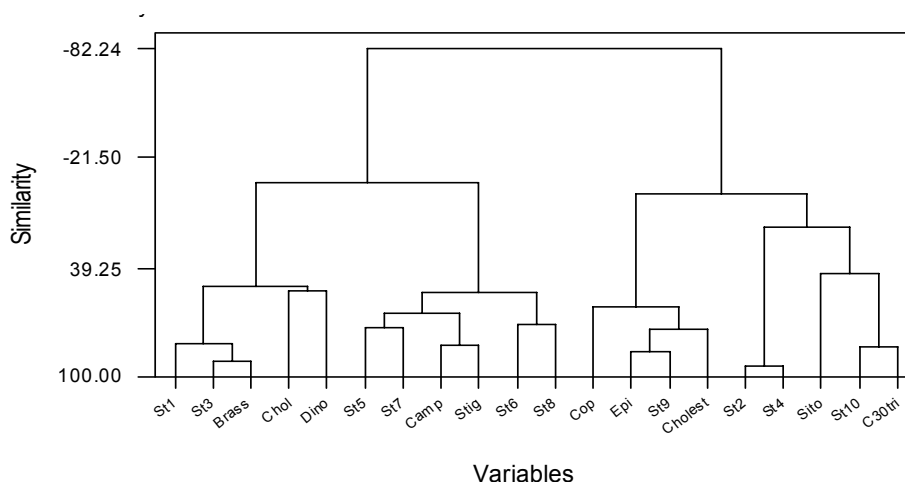


FIGURE 4. Cluster analysis of sterols showing the Ward's method for the Clyde Sea samples

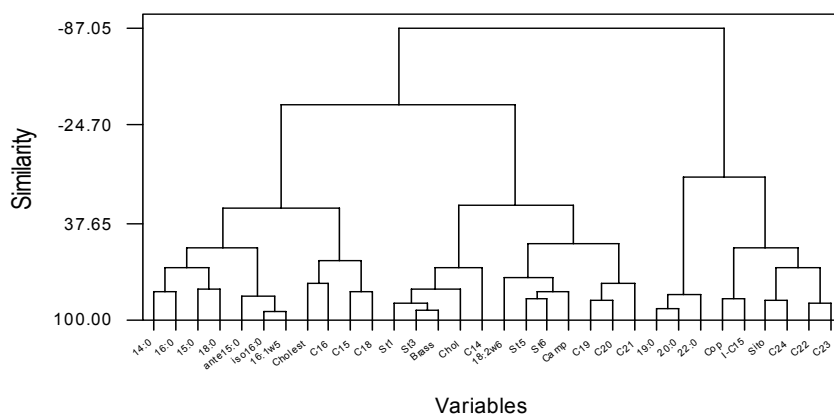


FIGURE 5A. Cluster analysis of mixed compounds showing the Ward's method for the Clyde Sea samples

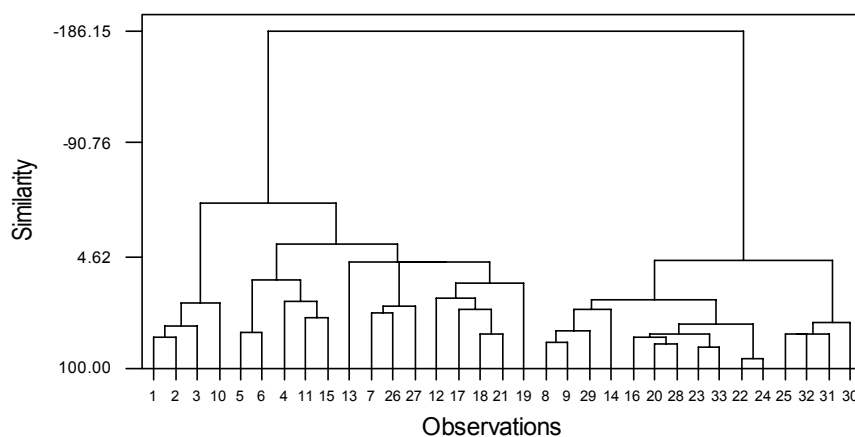


FIGURE 5B. Cluster analysis of sampling sites showing the Ward's method for the Clyde Sea samples

and fatty alcohols, sitosterol, coprostanol and cholestanol. However, two open sea samples (S8 and S9) were also clustered together with this group not with cluster group I.

CONCLUSION

Cluster analysis differentiated the study area into two main regions. The first cluster group defined mainly the open sea samples, which are characterised by marine derived compounds, while the second cluster group represented the loch samples with terrigenous and sewage inputs.

REFERENCES

- Carrie, R. H., Mitchell, L. & Black, K. D. 1998. Fatty acids in surface sediments at the Hebridean shelf edge, west of Scotland. *Organic Geochemistry*. 29: 1583-1593.
- Fukushima, K. & Ishiwatari, R. 1984. Acid and alcohol compositions of wax esters in sediments from different environments. *Chemical Geology*. 47: 41-56.
- Hinrichs, K. -U., Schneider, R. R., Müller, P. J. & Rullkötter, J. 1999. A biomarker perspective on paleoproductivity variations in two late Quaternary sediment sections from the Southeast Atlantic Ocean. *Organic Geochemistry*. 30: 341-366.
- Jeng, W. L. & Han, B. C. 1994. Sedimentary coprostanol in Kaohsiung Harbour and the Tan Shui Estuary, Taiwan. *Marine Pollution Bulletin*. 28: 494-499.
- Kaiser, K. L. E. & Esterby, S. R. 1991. Regression and cluster analysis of the acute toxicity of 207 chemicals to six species of biota and octanol/water partition coefficient. *Science of the Total Environment*. 109-110: 499-514.
- Laureillard, J. & Saliot, A. 1993. Biomarkers in organic matter produced in estuaries: a case study of the Krka estuary (Adriatic Sea) using the sterol marker series. *Marine Chemistry*. 43: 247-261.
- Leeming, R. & Nichols, P. D. 1998. Determination of the sources and distribution of sewage and pulp-fibre-derived pollution in the Derwent Estuary, Tasmania, using sterol biomarkers. *Marine and Freshwater Research*. 49: 7-17.
- Mudge, S. M., Bebianno, M. J., East, J. A. & Barreira, L. A. 1999. Sterols in the Ria Formosa Lagoon, Portugal. *Water Research*. 33: 1038-1048.
- Mudge, S. M., East, J. A., Bebianno, M. J. & Barreira, L. A. 1998. Fatty acids in the Ria Formosa Lagoon, Portugal. *Organic Geochemistry*. 29: 963-977.
- Mudge, S. M. & Lintern, D. G. 1999. Comparison of sterol

- biomarkers for sewage with other measures in Victoria Harbour, B.C., Canada. *Estuarine, Coastal and Shelf Science*. 48: 27-38
- Mudge, S. M. & Norris, C. E. 1997. Lipid biomarkers in the Conwy Estuary (North Wales, U.K.): a comparison between fatty alcohols and sterols. *Marine Chemistry*. 57: 61-84.
- Parkes, R. J. 1987. Analysis of microbial communities within sediments using biomarkers. In: Fletcher, M., Gray, T. R. G. and Jones, J. G. (Eds.). *Ecology of Microbial Communities*. Cambridge University Press, London, 440 pp.
- Pereira, M. G. S. 1999. *Bacterial degradation of linseed and sunflower oils in salt marsh sediments*. PhD Thesis. University of Wales, Bangor. 199p.
- Poodle, T. 1986. Freshwater inflows to the Firth of Clyde. *Proceedings of the Royal Society of Edinburgh, Section B* 90, 55-66.
- Rajendran, N., Fujiyoshi, E. & O. Matsuda. 1991. Fatty acids and fatty alcohols of zooplankton from coastal upwelling area of Hyuga Nada, Japan. *Bulletin of the Japanese Society of Fisheries*. 57: 2277-2284.
- Rohjans, D., Brocks, P., Scholz-Böttcher, B. M. & Rullkötter, J. 1998. Lipid biogeochemistry of surface sediments in the Lower Saxonian Wadden Sea, Northwest Germany, and the effect of the strong winter 1995-1996. *Organic Geochemistry*. 29: 1507-1516.
- Saliot, A., Laureillard, J., Scribe, P. & Sicre, M. A. 1991. Evolutionary trends in the lipid biomarker approach for investigating the biogeochemistry of organic matter in the marine environment. *Marine Chemistry*. 36: 233-248.
- Sargent, J.R., Gatten, R. R. & Henderson, R. J. 1981. Marine wax esters. *Pure and Applied Chemistry*. 53: 867-871.
- Thoumelin, G., Bodineau, L. & Wartel, M. 1997. Origin and transport of organic matter across the Seine estuary: Fatty acid and sterol variation. *Marine Chemistry*. 58: 59-71.
- Venkatesan, M.I. & Kaplan, I. R. 1990. Sedimentary coprostanol as an index of sewage addition in Santa Monica Basin, Southern California. *Environmental Science and Technology*. 24: 208-214.
- Volkman, J. K. 1986. A review of sterol markers for marine and terrigenous organic matter. *Organic Geochemistry*. 9: 83-99.
- Volkman, J. K., Barrett, S. M., Blackburn, S. I., Mansour, M. P., Sikes, E. L. & Gelin, F. 1998. Microalgal biomarkers: A review of recent research developments. *Organic Geochemistry*. 29: 1163-1179.
- Volkman, J. K., Farrington, J. W. & Gagosian, R. B. 1987. Marine and terrigenous lipids in coastal sediments from the Peru upwelling region at 15°S: Sterols and triterpene alcohols. *Organic Geochemistry*. 11: 463-477.
- Wakeham, S. G. & Lee, C. 1989. Organic geochemistry of particulate matter in the ocean: the role of particles in oceanic sedimentary cycles. *Organic Geochemistry*. 14: 83-96.

Stephen M. Mudge
 School of Ocean Sciences
 University of Wales, Bangor
 Menai Bridge
 Anglesey LL59 5AB
 United Kingdom

Masni Mohd Ali
 School of Environment and Natural Resource Sciences
 Faculty of Science and Technology
 Universiti Kebangsaan Malaysia
 43600 UKM Bangi, Selangor
 MALAYSIA