UNITED STATES DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

BIOGENIC AND ANTHROPOGENIC ORGANIC MARKERS AS SOURCE DISCRIMINANTS AND SEDIMENT TRANSPORT INDICATORS IN SOUTH SAN FRANCISCO BAY, CALIFORNIA

Frances D. Hostettler
John B. Rapp
Keith A. Kvenvolden
Samuel N. Luoma

Open-File Report 88-408

U.S. Geological Survey
Menlo Park, California

ABSTRACT

Sediment samples from near-shore sites in south San Francisco Bay and from streams flowing into that section of the bay have been characterized in terms of their content of biogenic and anthropogenic molecular markers. The distributions, input sources, and applicability of these compounds in determining sediment movement are discussed. By means of inspection and multivariate analysis, the compounds were grouped according to probable input sources, and the sampling stations according to the relative importance of source contributions. A suite of polycyclic aromatic hydrocarbons (PAH's) dominated by pyrene, fluoranthene and phenanthrene, typical of estuarine environments worldwide, and suites of mature sterane and hopane biomarkers were found to be most suitable as background markers for the Bay. A homologous series of long-chain n-aldehydes (C_{12}-C_{32}) with a strong even-over-odd carbon number dominance in the upper range and the ubiquitous n-alkanes (n-C_{24}-C_{34}) with a strong odd-over-even carbon number dominance were utilized as terrigenous markers. Several ratios of these terrigenous and Bay markers were calculated for each station. These ratios and the statistical indicators from the multivariate analysis gave clear evidence of net southerly sediment transport in this area of the Bay where non-tidal currents have been considered to be minimal or non-existant.

INTRODUCTION

San Francisco Bay, which receives drainage from 40% of the land area of California, is subjected to vast inputs of chemicals in largely unknown amounts from land and river runoff, industrial, agricultural and sewage discharges, atmospheric fallout, harbor and shipping traffic, and other anthropogenic sources. The Bay has not been as thoroughly investigated in terms of organic contaminants as many other estuaries in the U. S., for example, its East Coast equivalent, Chesapeake Bay, the largest estuary on that coast (Soiffer, 1985). Studies that have been undertaken in San Francisco Bay include those by Nichols *et al.*, (1986) which chronicles the modification of the estuary by human activity, by Kvenvolden (1962) which discusses n-alkanes in bay sediments, and by Spies *et al.*, (1985 and 1987) and Thomson *et al.* (1984), which investigate contaminants in the bay and their effect on bay organisms or their use as source markers.

It has long been recognized that one way to investigate the geochemical setting and (or) history of an estuarine environment is through its sedimentary record (Wade and Quinn, 1979; Prahl, 1985; Readman *et al.*, 1986). This present study was undertaken as a preliminary investigation of selected categories of organic constituents of surficial sediments in an area of south San Francisco Bay (Fig. 1). We chose to focus on this limited area because of previous work done here by one of us (SNL) on trace metal concentrations in Bay sediment and organisms (Thomson *et al.*, 1984). The objectives of this study were: a) to survey sediment organic constituents, mainly hydrocarbons, including biomarkers and other biogenic lipids, in this localized part of south San Francisco Bay; b) to investigate any relationships of these hydrocarbons and specific trace metals; c) to differentiate hydrocarbon input from marine, terrigenous (plant), and anthropogenic sources in order to find markers for each source; and d) to investigate the suitability of these markers for determining sediment transport processes.

Figure 1 shows the primary study area. Seven sites (#1-7) were sampled, including a sample somewhat north of the other sites near the San Mateo Bridge (#1). Table 1 gives specific location information. We were especially interested in looking at trends in the distributions of hydrocarbons to the north and south of San Francisquito Creek where we expected to see substantial terrigenous (plant) input from the creek flow, and to the north and south of the sewage outfall where we were unsure what input could be identified but where Thomson *et al.*, (1984) clearly showed a high input of the trace metals Cu and Ag. In order to identify more clearly terrigenous plant input, we took a second set of samples (#8-14) from beds of streams flowing into this area of south San Francisco Bay.

EXPERIMENTAL

METHODS

Methods used in this study focused on the unbound organic constituents which are extractable with dichloromethane. The extracts of the bay sediments were separated into three fractions and the extracts of the upstream sediments into six fractions in order of increasing polarity using liquid chromatography (see below). The larger number of fractions in the upstream samples resulted from a subfractionation to obtain cleaner class separation of constituents. The hexane fraction contains aliphatic hydrocarbons, including several suites of biomarkers. Biomarkers are "molecular fossils" of natural products now in the sediments. These compounds retain their basic skeleton but are modified by time and temperature. The degree of their modification, their distributional patterns and structure give information on their maturity, source, depositional environment and other molecular parameters. The benzene fraction(s) contain aromatic and polycyclic aromatic hydrocarbons (PAH's), and possible S- or O-containing molecules, including molecules with a carbonyl functional group (aldehydes and possible ketones). The methanol fraction contains the most polar constituents, that is, the more polar N,S,O-containing compounds. The hexane and benzene eluates were analyzed by gas chromatography (GC) and gas chromatography/mass spectrometry (GC/MS) to identify constituents; the methanol eluates were checked for weight contribution only. Then the chromatograms of the fractions were analyzed by inspection and by multivariate analysis in order to determine how individual and groups of compounds varied with respect to each other. Wherever possible, the input source of the constituents was identified, and then ratios of representative compounds from terrigenous and bay input were used to indicate sediment transport.

Sample collection

Surficial sediment samples from 14 sites were collected by spooning the wet sediment into clean aluminum cans. Seven samples from along the South Bay perimeter were taken in April, 1986 just after the end of the rainy season, and seven samples were obtained from beds of streams flowing into this area of the bay (see Figure 1 and Table 1) during the dry summer season. Some of the upstream samples were only moist because of seasonal lack of water flow. The bay samples were mainly fine silt and sand, and the upstream samples fine sand and gravel. Five of the seven samples (#8,9,10,13 and 14) were collected in the hills above any urban areas; two (#11 and 12) were collected from near urban areas where local urban runoff would be expected. The samples were frozen as soon as possible after collection and then freeze-dried to remove the water.

Sample fractionation

One hundred grams (from bay sites) or 200-300 g (from upstream sites) was manually broken up as finely as possible with a spatula to approximately 30 mesh, then placed in a bottle and extracted serially with dichloromethane (DCM) on a wrist-action shaker: 200 ml for 2 hrs, 100 ml for 2 hrs, then washed with 100 ml for 15 min. Each time the bottles were centrifuged for 5 min at 1500 rpm and the supernatant decanted and filtered through glass wool. Each total extract was evaporated to 5-10 ml on a rotary evaporator. Elemental sulfur was removed by filtering the extract through HCl-activated copper powder. The solvent of the sulfur-free eluate (SFE) was exchanged for 5 ml hexane, and the SFE was then applied to a column layered with 5 g and 2.5 g activated silica gel (Davidson Nos. 923 and 62, respectively) and 2.5 g deactivated (5% water) alumina, all in hexane. The SFE flask was further rinsed by the sequential elution solvents. Two fractionation procedures were used: for the bay sediments (#1-7) three solvents, hexane, benzene and methanol, were used serially to elute the constituents, and three corresponding fractions (H, B and M) were collected; for the upstream samples (#8-14) the solvents used were hexane, 20, 40, and 60% benzene in hexane, benzene and methanol, and six corresponding fractions were collected. Aliquots (50 ul) of the SFE and the three or six fractions were removed for weight percentage calculations. Weights are given in Table 1 (weights for the upstream 20, 40, 60 and 100% benzene fractions are combined). A method blank was carried through each extraction procedure. The hexane and the benzene fractions were examined by GC and GC/MS.

Chromatography and Spectrometry

Gas chromatography of the fractions was done on a Varian Model 3700 Gas Chromatograph using a 30 m x 0.3 mm, DB-1 bonded phase, fused silica capillary column with flame ionization detection (FID). Conditions were: column temperatures, 90°C for 3 min followed by 4°C/min temperature increase to 310°C and held for 20 min; injection port and detector temperatures, 300°C; column pressure, 10 psi helium; splitless injection. Quantitation of n-alkanes (for calculation of CPI, the carbon preference index), isoprenoids and PAH's was carried out by

GC using external standards. Gas chromatography/mass spectrometry was done on a Hewlett-Packard Model 5995A GC/MS using a 30 m x 0.3 mm DB-5 bonded phase, fused silica capillary column. Identification of compounds was accomplished by comparison of obtained spectra with archived and literature spectra, and verified by comparison with known compounds in most cases. Only the lower molecular weight n-aldehydes were available as standards; therefore, for the higher homologs literature spectra were used, as well as retention time extrapolation and molecular ions. Biomarker identifications were made as in previous studies (Kvenvolden *et al.*, 1987). Selected-Ion-Monitoring (SIM) was done on the hexane fractions, monitoring m/z 191 for terpanes and triterpanes, 217 for steranes, and 57 for alkanes, using dwell times of 600, 600 and 10 msec respectively. Two temperature programs were used: 1) injection at 60°C, fast ramp to 90°C, then 6°/min to 300°C, holding for 10 min; and 2) injection at 150°C, fast ramp to 200°C, then 1°/min to 300°C.

Multivariate analysis

After chromatography the areas of the major peaks of each of the bay (#1-7) and the upstream (#8-14) hexane fractions and the bay benzene fractions were placed into three respective data matrices for statistical analysis (For relevant discussions on the use of multivariate analysis in geochemical studies, see Zumberge (1987), Shaw and Johns (1986), Øygard *et al.* (1984), Malinowski and Howery (1980), Kowalski (1977) and Davis (1986)). The data matrix used for the benzene fraction of the bay samples consisted of 32 compounds (known PAH's and unknowns which appear most frequently in the samples). Initially, Q mode factor analysis was used with cosine theta as the similarity index. Q mode analyzes the data with the objective of determining the relationships between sample sites. R mode factor analysis was then done with correlation coefficients used as the similarity index. R mode analyzes the data with the objective of determining the relationships between variables (compounds). In both types of analysis, Varimax Rotation was used to optimize the results. The "composition" of a factor is obtained from the factor loadings in R mode and from the factor scores in Q mode.

The data matrix used for the hexane fractions of the seven bay samples consisted of 42 compounds, and that of the seven upstream samples, 26 compounds (n-alkanes, isoprenoids and unknowns which appear most frequently in these fractions). Q mode analysis was used with cosine theta as the similarity index. Because >95% of the total information was explained by one factor in both cases, using unrotated principal components, Varimax Rotation was not employed.

Geochemical Ratios

The following calculations were made to show relationships among recent terrigenous (plant) input and background bay constituents of the sediments:

1. Carbon Preference Index, calculated according to the method of Cooper and

Bray (1963):

$$CPI = \frac{1}{2}\left(\frac{C_{25}+C_{27}+C_{29}+C_{31}+C_{33}}{C_{24}+C_{26}+C_{28}+C_{30}+C_{32}}+\frac{C_{25}+C_{27}+C_{29}+C_{31}+C_{33}}{C_{26}+C_{28}+C_{30}+C_{32}+C_{34}}\right)$$

(using absolute concentrations)

2. From Fraction B:

$$\frac{[Major\ \ terrestrial\ \ aldehyde\,]}{[Major\ \ background\ \ PAH\,]} = \frac{[n-C_{30}\ aldehyde\,]}{[Fluoranthene\,]}$$

(using GC total areas)

3. From Fraction H:

$$\frac{[Major\ \ terrestrial\ \ n-alkane\,,\ m\,/z\ \ 57]}{[Major\ \ background\ \ biomarker\,,\ m\,/z\ \ 191]} = \frac{[n-C_{29}]}{[\alpha,\beta\text{–}Hopane\,,C_{30}]}$$

(using GC/MS SIM ion areas)

RESULTS AND DISCUSSION

CHROMATOGRAPHY

The hexane and benzene fractions were examined in detail by GC and GC/MS. The following sections outline the types of compounds found in the two fractions and the information that can be obtained from them.

Hexane Fraction

Figure 2 shows the chromatography of a typical bay and a pristine upstream hexane extract, and Table 2 lists the identifications of constituents of this and later fractions. All of the hexane fractions are dominated by the C_{24}-C_{34} n-alkanes with a strong odd-over-even carbon number predominance, giving high CPI values (see Table 1). Values for CPI can theoretically range from about 20 down to 1, with the former representing distributions of original unaltered plant waxes, and 1 (no odd or even preference) representing a mixture that has been substantially modified, especially by time and temperature as, for example, in oils and other petroleum products, or by dilution with petroleum contaminants. The CPI can therefore be viewed as an approximate measure of the magnitude of the recent terrigenous contribution to the sediment.

The lower molecular weight n-alkanes and isoprenoids are present in bay fractions but at significantly lower concentrations than the higher molecular weight n-alkanes. Only a relatively minor amount of the isoprenoid pristane is present, suggesting little or no zooplankton contribution (Boehm, 1984). Low n-C_{15} and n-C_{17} abundances also indicate little marine phytoplankton or freshwater algal input (Wade and Quinn, 1979). There is, however, a possibility that some preferential degradation of the lower molecular weight n-alkanes is occurring and thus contributing to an apparent enhancement of the amount of the higher molecular weight terrigenous contribution (Brassell *et al.*, 1983).

A series of several sesquiterpenes (C_{15}) is present in the bay samples and, in lower concentrations, in the upstream samples. Because sesquiterpenes occur in the essential oils of higher plants and are therefore of terrigenous origin, it was of interest to further identify the compounds. The C_{15} compounds occur in both the hexane and benzene fractions according to their degree of aromaticity. Figure 3 shows the early eluting compounds on chromatograms from the GC/MS. We find three groups of sesquiterpenes, as indicated in Table 2. There are eight cycloalkene isomers, #1-8 in Fig. 3, molecular weight 204, four partially aromatized isomers, A-D, molecular weight 202, and one fully aromatic compound, cadalene, e (I in Appendix; the structures of all compounds with Roman numerals are shown in the appendix), molecular weight 198. The mass spectra of all the sesquiterpenes are shown in Figure 4. In the fractionation scheme used for the bay samples the cycloalkene isomers were split between the hexane and benzene fractions because of their mixed alkane-alkene character (see Figure 3, Station #4, hexane and benzene fractions). Fine-tuning the fractionation procedure for the upstream samples allowed all of the cycloalkene isomers to be eluted in the hexane fraction. Among all the samples there is one dominant cycloalkene (6), seven minor isomers (1-5,7,8) and traces of other isomers not shown. The benzene fractions of all samples contain the aromatic sesquiterpenes. The sesquiterpenes from both the hexane and benzene fractions are discussed in this section for continuity.

The mass spectra of the cycloalkene isomers all show molecular ions of m/z 204, several (1,3,5,6 and 7) with major ions at 161 and 189. The spectra of these are similar, but not clearly identical, to literature spectra of sesquiterpenoid cycloalkenes like longifolene (II)(Philp, 1985) or members of the cadinene family (III)(Albaiges *et al.*, 1984b). The cadinene family of compounds, with two double bonds endo- or exo- to the bicyclic system, has many naturally-occurring isomers. Without standards, positive identifications cannot be made on the compounds. However, the presence of cadinene-type sesquiterpenoids is further indicated by the presence of their partially and fully aromatized counterparts, for which they are the reported precursors (Simoneit and Mazurek, 1982). Cadinenes are easily degraded, and the degradation pathway leads first to the partially aromatized calamanene, B (IV in Appendix) and 5,6,7,8-tetrahydrocadalene, C (V in Appendix), and then to the fully aromatized cadalene, e (I in Appendix), all of which are seen in the benzene fraction of some of the upstream samples.

In the benzene fractions in addition to IV and V above, two other sesquiterpenoids at molecular weight 202 are seen in some samples. The first (A, Figure 3) shows a spectrum similar to that of cuparene (VI in Appendix) (Grantham and Douglas, 1980), also a constituent of a higher plant essential oil. The other member of this category (C, Figure 3) has not yet been identified.

Although all these compounds are clearly derived from resinous higher plants (Simoneit and Mazurek, 1982) and are therefore terrigenous markers, their relative lability makes interpretation of their distribution difficult. For example, the dominant cycloalkene member, 6, shows a maximum concentration in the bay at site #4, south of San Francisquito Creek (Fig. 3), but it is not highly concentrated upstream. The observation that the more reduced members of the cadinene family dominate in bay sediments, while the partially and fully aromatized members are seen upstream is contrary to the expected precursor-product relationship described above. Differing sedimentary environments (i.e., reducing or oxidizing) or the seasonality of the sediment collection could explain these results. Further work is necessary to determine the exact plant origin of these sesquiterpenes and, if standards are available, their structures and their interactions.

A significant Unresolved Complex Mixture (UCM) is also seen in the hexane fractions of the bay samples, but is absent in the upstream samples collected away from urban input. The UCM is a mix of chromatographically unresolvable components seen as a "hump" in the chromatogram (Fig. 2). It is usually considered indicative of petroleum contamination or biodegradation (Brassell and Eglinton, 1986).

The chromatograms of the most pristine upstream sediment sample extracts are considerably simpler than those of the bay samples, showing only the n-alkanes with a high odd-even predominance, no UCM, and only minimal sterane and hopane biomarkers (see below). This observation is not surprising, but because one of the purposes of this work is to differentiate different types of input to the bay sediments, it is important in distinguishing terrigenous input.

Biomarkers

The hexane extracts were also examined for sterane and hopane biomarkers. Figure 5 shows biomarker mass chromatography of a representative bay sample; all the bay samples were similar. This mass chromatography was done in the SIM mode, which focuses the mass spectrometer on only a small number of ion masses rather than an entire range, thus increasing the method sensitivity and allowing the monitoring of compounds which are present in low concentrations. We examined terpanes, especially the pentacyclic hopanes, at mass 191, steranes at mass 217, and for reference, n-alkanes at mass 57. In contrast to the n-alkanes which clearly show very recent terrigenous sources because of their strong odd-even predominance, the steranes and hopanes show the complex signature of mature or highly reworked constituents. The 17α(H),21β(H)-hopane (C_{30}), especially, is the mature end-member of the naturally-occurring 17β(H),21β(H)-hopane, as are the 17α(H),21β(H)-30-homohopanes with their C22-S/R-isomer ratio of 1.6 (Ensminger *et al.*, 1974). The steranes show high diasterane content and the complex equilibrium mixture of mature isomers. Because of the maturity of these two latter signatures and their consistency in the bay samples, and because the pristine upstream samples show minimal amounts of the mature isomers, we consider the terpanes and steranes to be background for the bay, most likely from petroleum contamination. A petroliferous origin is supported by work in other systems (Brassell and Eglinton, 1986; Farran *et al.*, 1987; Venkatesan *et al.*, 1980). Contrasting the SIM response of the dominant n-alkane, as a terrigenous marker, to that of the dominant hopane, as bay background, allows us to compare terrigenous input to background bay input (Table 1, Fig. 9).

Benzene Fraction

Chromatography of all of the bay benzene fractions are shown in Figure 6, with identifications in Table 2. The sesquiterpenes found here have been discussed with their counterparts in the section on the hexane fraction. In the bay samples two major suites of components are observed, the first a suite of PAH's indicated by letters, and the second a suite of straight-chain aldehydes, indicated by numbers which give their chain length (only the even-carbon-numbered constituents are labeled; the odd-carbon-numbered homologs are present but in much lower concentrations).

The PAH's throughout this area of the bay are dominated by pyrene, fluoranthene and phenanthrene, and as a group, their distribution is typical of that found all over the world in this type of an estuarine environment, both in highly populated areas like the Charles River in Boston and in remote areas (LaFlamme and Hites, 1978; Johnson *et al.*, 1985; Pruell and Quinn, 1985). These compounds are considered to be derived from pyrolytic sources like forest fires and anthropogenic combustion, and to be spread primarily by aeolian transport (LaFlamme and Hites, 1978). Because their distribution in estuarine environments is universal, for the purposes of this paper they are considered background Bay constituents. Although pyrene is the dominant PAH, fluoranthene is present in only slightly lower amounts, and is shown by multivariate analysis to be the most representative in the PAH assemblage. Thus fluroanthene is used in source ratio calculations. The PAH's from the pristine upstream sites have only minimal representation from the above group; rather, these sites are dominated by PAH's like cadalene (I) that, as discussed above, is considered to be a product of natural diagenesis of plant constituents (Prahl and Carpenter, 1983).

The aldehyde suite consists of normal-aldehydes from C_{12} to C_{32}. The entire range is not obvious in the bay samples, but when the upstream samples were subfractionated, the entire suite of aldehydes was separated in the 60% benzene fraction, and the lower homologs were identified (Figure 7). In the higher molecular weight range the aldehydes have a strong even-over-odd carbon number dominance. These compounds have been reported elsewhere (Prahl and Pinto, 1987) to be primary material from terrigenous plants, which our analyses of the upstream samples substantiate, but these compounds have not been previously reported in San Francisco Bay. Also, in Albaiges *et al.*, 1984a, these aldehydes were identified in lagoon sediments, but they were found in considerably lower concentrations than are present in our study, and as bound rather than unbound constituents. The aldehyde concentrations were low enough in the lagoon sediments that they were considered to be secondary diagenetic material, in contrast to the conclusions of Prahl and Pinto (1987).

Of major importance in this study is the respective variation of these two suites of compound at each of the different sampling sites. Two extremes exist in the bay sample set, sites 1 and 4. At #1, San Mateo Bridge, the background PAH's dominate, whereas at #4, south San Francisquito Creek, where we believe terrigenous input is highest in this sample set, the terrigenous aldehydes dominate. These two suites of compounds vary relative to one another throughout the South Bay sample set, and a ratio of the GC response of the most representative member of each suite, that is, the C_{30}-aldehyde to fluoranthene, gives us another comparison of terrigenous input to background bay input (see Table 1 and Fig. 8).

PARAMETER COMPARISONS

The three independent ratios of organic parameters discussed above were used to compare terrigenous input to background bay constituents in the bay sediments. Trends developed from absolute concentrations were not useful; absolute concentrations are known to be highly variable with fluctuations of two to three orders of magnitude (Farran *et al.*, 1987). Comparative data is considered to be more reliable (Brassell and Eglinton, 1986).

The first parameter is the Carbon Preference Index (CPI), the quantitation of the magnitude of the odd-over-even dominance of the C_{24} to C_{34} n-alkanes. The second parameter is a ratio from the benzene fractions of the prominant terrestrial aldehyde to fluoranthene, the most representative background PAH. The third parameter is a ratio from the SIM biomarker mass chromatography of the major terrestrial n-alkane n-C_{29} to the major background biomarker, $17\alpha(H),21\beta(H)$-hopane.

The highest values in this sample set for the three ratios come from the five most pristine upstream samples (Table 1). This result is expected, especially for CPI, because the n-alkanes in this range are known to originate from the leaf waxes of vascular land plants. Two upstream samples (#11 and #12) show lower values comparable to values observed in the bay; these samples were collected in an urban area where high anthropogenic pollution, i.e. petroleum product contamination, was expected. The lowest value occurs at station #1, the Bay site furthest from terrigenous input.

The results of the parameter comparisons for the South Bay sites alone are listed in Table 1 and graphed in Figure 8. For all three ratios there is a fairly broad but distinct range at most sites. But, at site #4, the site just south of San

Francisquito Creek, all three ratios show a clear maximum. San Francisquito Creek is the site where we expected to see the highest input of terrigenous components from the stream flow, but the fact that only the site south of the creek and not the north site showed the maxima is of interest. The hydrodynamics of south San Francisco Bay are not well established, but it is generally accepted that non-tidal, residual currents are not strong in the slowly circulating, semi-enclosed terminus of the south part of the Bay (Conomos, 1981). Nevertheless, these data and the multivariate analysis discussed below, all of which show a distinct bias in terrigenous input to the south of the major source of this input, namely San Francisquito Creek, give an indication of a consistent net southward transport. A second example of southerly bias is indicated below by the distribution of compounds of apparent sewage origin in the multivariate analysis. Both examples agree with the previous work (Thomson *et al.*, 1984) that showed the concentration of Cu and Ag trace metals maximizing south of their major input area, the sewage outfall.

MULTIVARIATE ANALYSIS

The above conclusions were verified and expanded by the statistical analysis of the chromatography of the hexane and benzene fractions. The data from the factor analysis was used in R mode to group the various constituents in the fractions according to source, and in Q mode to compare sample sites with respect to input sources. In the case of the hexane fractions, both the bay sample set and the upstream sample set needed only one factor (unrotated) to explain the great majority (>95%) of the total information. Within these factors, 88% (bay) and 94% (upstream) of the information was contained in only three compounds (the major plant wax alkanes, n-C_{27}, n-C_{29} and n-C_{31}), supporting the contention that the great majority of the material in the hexane fraction of the bay sediments has a terrigenous origin.

The analysis of the bay benzene fractions was more complex. Factor loadings for R mode, which demonstrate how the fraction constituents are distributed, are shown in Figures 9a and 9c; factor loadings for Q mode, which demonstrate how the seven stations are related with respect to these constituents, are shown in Figures 9b and 9d. It is fortuitous that the order of the factors is the same in both modes. Because of this, both Q mode and R mode will be discussed together for each factor. In the Q mode, four factors account for 99% of the total variance of the data:

FACTOR 1: This factor explains 34% of the total information in the Q mode. Figure 9a shows the R mode factor loadings of the constituents of Factors 1 and 2, and Table 3 gives identifications of the dominant compounds. Factor 1 is composed mainly of the low to medium molecular weight PAH's, phenanthrene, fluoranthene, pyrene, benzo(a)pyrene, and perylene. As discussed above, this is the suite of compounds found in many diverse parts of the world in estuarine environments, which we are considering to be background in the bay. Figure 9b shows that the sample with the highest loading in this factor (-0.83) is site #1, the northernmost site. The second highest loading (-0.67) is #7, the southernmost site. A high factor loading at a site means that a large amount of the total information for that site is explained by that factor. In quantitative terms, the square of the factor loading times 100 gives the percentage of the total information at that site explained by that factor (see factor 2 for an example).

FACTOR 2: This factor explains 35% of the total information in the Q mode. Figure 9a shows that the factor consists of a homologous series of n-aldehydes, with the C_{26}, C_{28}, C_{30} and C_{32} homologs being present in the highest concentrations. These aldehydes reflect terrigenous input. Figure 9b shows that the sample with the highest loading (-0.89) is site #4 (south of San Francisquito Creek) while #3 (north of San Francisquito Creek) has the second highest loading (-0.70). This result implies that: 1) the creek is the source of the aldehydes (which we confirmed with the upstream sediment samples), and 2) there is a net southward movement of material from the mouth of the creek (quantitatively, factor 2 explains $(-0.89)^2$ x 100 = 79% of the information in #4 and only $(-0.70)^2$ x 100 = 49% of the information in #3; if there were uniform mixing at the mouth, the values would be similar).

Figure 9b represents a large condensation of information. Sixty nine percent [34%(factor 1) plus 35%(factor 2)] of the total variation in the 224 data points (7 samples, 32 compounds) is represented in only two dimensions! Stations 1, 7, 3 and 4 all have >75% of their information explained by adding variable amounts of two end members (the estuarine background PAH's,factor 1, and the terrestrial aldehydes, factor 2). Stations 5, 6 and 2 have 59 to 44% of their information explained by factors 1 and 2 and thus need other factors to fully explain their total variance.

FACTOR 3: This factor explains 24% of the total information in the Q mode analysis. Figure 9c is a graph of the factor loadings of the compounds (R mode) for factor 3 vs. factor 4. Compounds with the highest loading in factor 3 are a highly correlated group of low to medium volatility unknown compounds. The unknown compounds elute from a DB-1 capillary column between 288° and 310° C, and their retention times indicate that they may be a homologous series. The compounds are most predominant at station #2 (loading = 0.74). The stations near the sewage outfall, # 5 and 6, are intermediate between #2 and a group with low loadings (#1, 3, 4 and 7). Presently we have no interpretation of this factor but are working on the identification of the unknowns.

FACTOR 4: This factor explains 7% of the information in the Q mode. This factor consists of a group of co-occurring high molecular weight PAH's (benzo(b)fluoranthene, benzo(e)pyrene, indeno(1,2,3-cd)pyrene, benzo(ghi)perylene and coronene). This group of compounds is most prominent (i.e., has the highest loading, 0.41) at station #6 (south of the sewage outfall) followed by #5 (0.33)(north of the sewage outfall) and #7 (0.32). This distribution (#6 > #5 = #7 > the rest of the samples) implies that: 1) either the sewage outfall is a source of these compounds, or components of the sewage discharge facilitate higher inclusions of these high molecular weight PAH's in the sediments; and 2) there is a net southward movement of material from the sewage outfall.

CONCLUSIONS

This preliminary survey of the hydrocarbon constituents of the sediments from an area of south San Francisco Bay has established the presence of several hydrocarbon markers or suites of markers that can be utilized as terrigenous and bay source indicators. One of these suites of markers is an homologous series of n-aldehydes with a strong even-over-odd carbon number dominance, reported here for the first time in San Francisco Bay. Calculated ratios of these markers and multivariate analysis are then used as indicators of input source bias and therefore sediment movement. In this area of the bay non-tidal currents have been considered to be weak or non-existant. Nevertheless, this study shows clear evidence for southward sediment transport from the major input source of its terrigenous sedimentary component, San Francisquito Creek, and implies southward sediment transport from the sewage outfall. These results are similar to the result seen in the aforementioned study with the trace metals, Cu and Ag, showing a southerly bias in sediment concentrations from their major input source, the sewage outfall.

REFERENCES

Albaiges, J., Algaba, J., and Grimalt, J., 1984a, Extractable and bound lipids in some lacustrine sediments: Organic Geochemistry, v. 6, p. 223-236.

Albaiges, J., Grimalt, J., Bayona, J.M., Risebrough, R., De Lappe, B., and Walker, W., II, 1984b, Dissolved, particulate and sedimentary hydrocarbons in a deltaic environment: Organic Geochemistry, v. 6, p. 237-248.

Boehm, P.D., 1984, Aspects of the saturated hydrocarbon geochemistry of recent sediments in the Georges Bank region: Organic Geochemistry, v. 7, p. 11-23.

Brassell, S.C., and Eglinton, G., 1986, Molecular geochemical indicators in sediments, *in* Sohn, M.L., ed., Organic Marine Geochemistry: ACS Symposium Series 305, p. 10-32.

Brassell, S.C., Howell, V.J., Gowar, A.P., and Eglinton, G., 1983, Lipid geochemistry of Cretaceous sediments recovered by the Deep Sea Drilling Project, *in* Bjorøy, M., and others, eds., Advances in Organic Geochemistry 1981: New York, John Wiley & Sons Limited, p. 477-484.

Cooper, J.E., and Bray, E.E., 1963, A postulated role of fatty acids in petroleum formation: Geochimica et Cosmochimica Acta, v. 27., p. 1113-1127.

Conomos, T.J., 1981, Properties and circulation of San Francisco Bay waters, *in* Conomos, T.J., ed., San Francisco Bay: The Urbanized Estuary: San Francisco, Pacific Division AAAS, p. 47-84.

Davis, J.C., 1986, Statistics and data analysis in geology (2d ed.): New York, John Wiley & Sons, 646 p.

Ensminger, A., van Dorsselaer, A., Spyckerelle, Ch., Albrecht, P., and Ourisson, G., 1974, Pentacyclic triterpenes of the hopane type as ubiquitous geochemical markers: origin and significance, *in* Tissot, B., and Bienner, F., eds., Advances in Organic Geochemistry 1973: Paris, Editions Technip, p. 245-260.

Farran, A., Grimalt, J., Albaiges, J., Botello, A.V., and Macko, S.A., 1987, Assessment of petroleum pollution in a Mexican river by molecular markers and carbon isotope ratios: Marine Pollution Bulletin, v. 18, p. 284-289.

Grantham, P.J., and Douglas, A.G., 1980, The nature and origin of sesquiterpenoids in some tertiary fossil resins: Geochimica et Cosmochimica Acta, v. 44, p. 1801-1810.

Johnson, A.C., Larsen, P.F., Gadbois, D.F., and Humason, A.W., 1985, The distribution of polycyclic aromatic hydrocarbons in the surficial sediments of Penobscot Bay (Maine, USA) in relation to possible sources and to other sites worldwide: Marine Environmental Research, v. 15, p. 1-16.

Kowalski, B.R., 1977, Chemometrics: theory and application: ACS Symposium Series 52, Washington, D.C., American Chemical Society, 288 p.

Kvenvolden, K.A., 1962, Normal paraffin hydrocarbons in sediments from San Francisco Bay, California: Bulletin of the American Association of Petroleum Geologists, v. 46, p. 1643-1652.

Kvenvolden, K.A., Rapp, J.B., Golan-Bac, M., and Hostettler, F.D., 1987, Multiple sources of alkanes in Quaternary oceanic sediment of Antarctica: Organic Geochemistry, v. 11, p. 291-302.

LaFlamme, R.E., and Hites, R.A., 1978, The global distribution of polycyclic hydrocarbons in recent sediments: Geochimica et Cosmochimica Acta, v. 42, p. 289-303.

Malinowski, E.R., and Howery, D.G., 1980, Factor analysis in chemistry: New York, John Wiley and Sons, 251 p.

Nichols, F.H., Cloern, J.E., Luoma, S.N., and Peterson, D.H., 1986, The modification of an estuary: Science, v. 231, p. 567-573.

Øygard, K., Grahl-Nielsen, O., and Ulvoen, S., 1984, Oil/oil correlation by aid of chemometrics: Organic Geochemistry, v. 6, p. 561-567.

Philp, R.P., 1985, Fossil fuel biomarkers, applications and spectra: Methods in Geochemistry and Geophysics: v. 23, Amsterdam, Elsevier, 294 p.

Prahl, F.G., and Carpenter, R., 1983, Polycyclic aromatic hydrocarbon (PAH)-phase associations in Washington coastal sediment: Geochimica et Cosmochimica Acta, v. 47, p. 1013-1023.

Prahl, F.G., 1985, Chemical evidence of differential particle dispersal in the southern Washington coastal environment: Geochimica et Cosmochimica Acta, v. 49, p. 2533-2539.

Prahl, F.G., and Pinto, L.A., 1987, A geochemical study of long-chain n-aldehydes in Washington coastal sediments: Geochimica et Cosmochimica Acta, v. 51, p. 1573-1582.

Pruell, R.J., and Quinn, J.G., 1985, Geochemistry of organic contaminants in Narragansett Bay sediments: Estuarine, Coastal and Shelf Science, v. 21, p. 295-312.

Readman, J.W., Mantoura, R.F.C., Llewellyn, C.A., Preston, M.R., and Reeves, A.D., 1986, The use of pollutant and biogenic markers as source discriminants of organic inputs to estuarine sediments: International Journal of Environmental Analytical Chemistry, v. 27, p. 29-54.

Shaw, P.M., and Johns, R.B., 1986, Organic geochemical studies of a recent Inner Great Barrier Reef sediment--II. Factor analysis of sedimentary organic materials in input source determinations: Organic Geochemistry, v. 9, p. 237-244.

Simoneit, B.R.T., and Mazurek, M.A., 1982, Organic matter of the troposphere---II. Natural background of biogenic lipid matter in aerosols over the rural western United States: Atmospheric Environment, v. 16, p. 2139-2159.

Soiffer, B., 1985, Troubled waters: SF Focus, August 1985, p. 147-150.

Spies, R.B, Rice, D.W., Jr., Montagna, P.A., and Ireland, R.R., 1985, Reproductive success, xenobiotic contaminants and hepatic mixed-function oxidase (MFO) activity in *Platichthys stellatus* populations from San Francisco Bay: Marine Environmental Research, v. 17, p. 117-121.

Spies, R.B., Andresen, B.D., and Rice, D.W., Jr., 1987, Benzthiazoles in estuarine sediments as indicators of street runoff: Nature, v. 327, p. 697-699.

Thomson, E.A., Luoma, S.N., Johansson, C.E., and Cain, D.J., 1984, Comparison of sediments and organisms in identifying sources of biologically available trace metal contamination: Water Research, v. 18, p. 755-765.

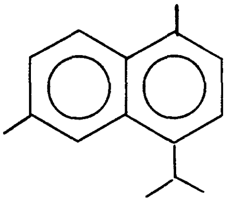
Venkatesan, M.I., Brenner, S., Ruth, E., Bonilla, J., and Kaplan, I.R., 1980, Hydrocarbons in age-dated sediment cores from two basins in the Southern California Bight: Geochimica et Cosmochimica Acta, v. 44, p. 789-802.

Wade, T.L., and Quinn, J.Q., 1979, Geochemical distribution of hydrocarbons in sediments from mid-Narragansett Bay, Rhode Island: Organic Geochemistry, v. 1, p. 157-167.

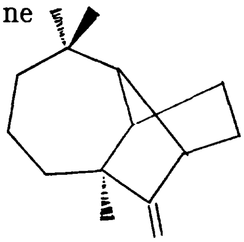
Zumberge, J.E., 1987, Prediction of source rock characteristics based on terpane biomarkers in crude oils: a multivariate statistical approach: Geochimica et Cosmochimica Acta, v. 51. p. 1625-1637.

APPENDIX. Structures cited in the text.

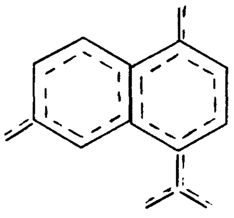
I. Cadalene



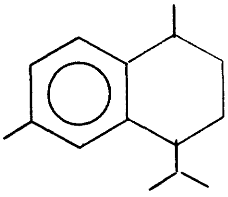
II. Longifolene



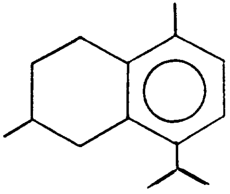
III. Cadinenes (2 double bonds)



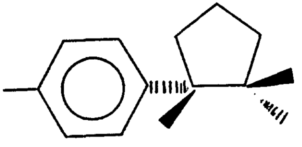
IV. Calamenene



V. 5,6,7,8-Dihydrocadalene



VI. Cuparene



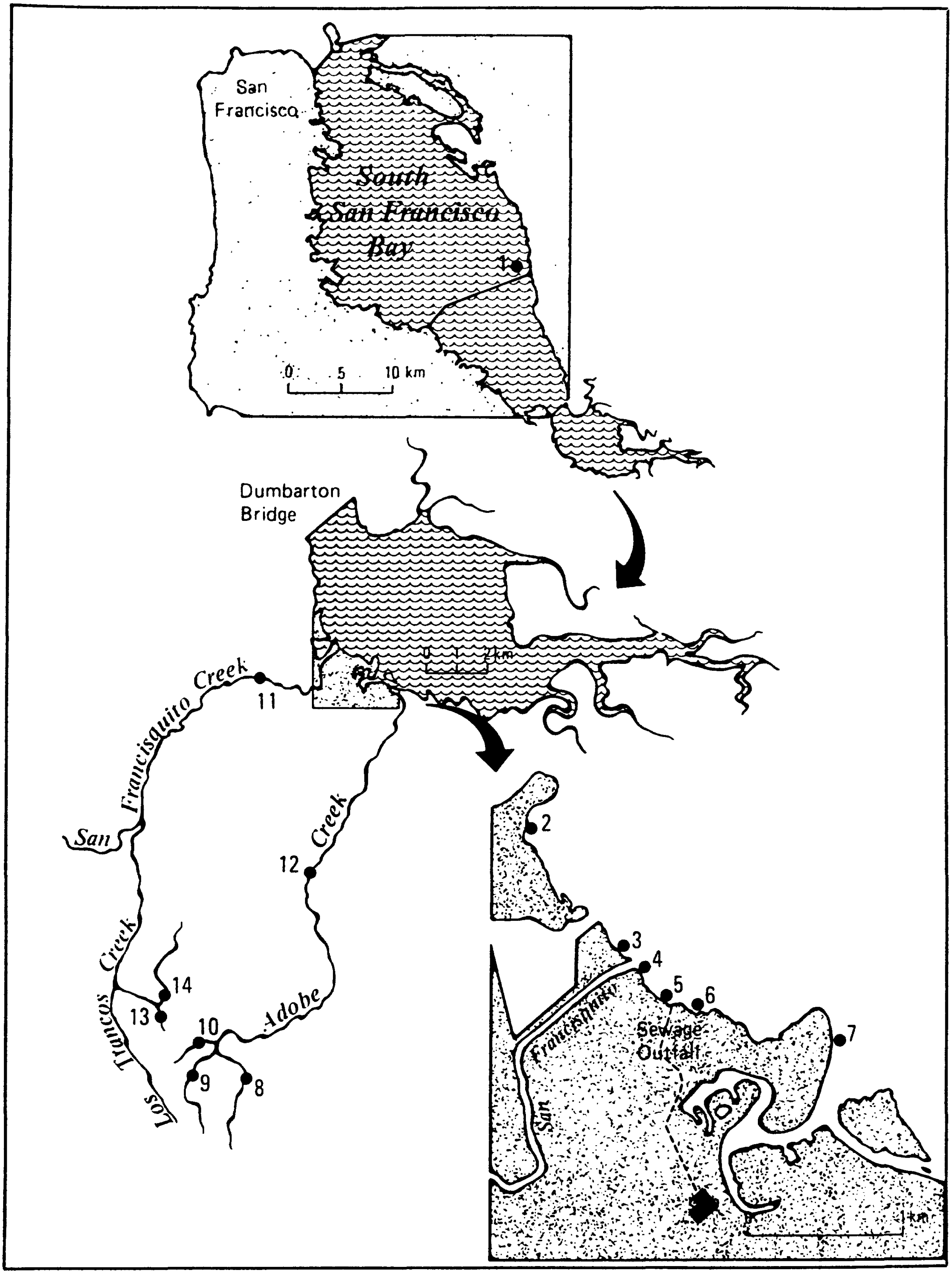


Figure 1. Location of sampling stations.

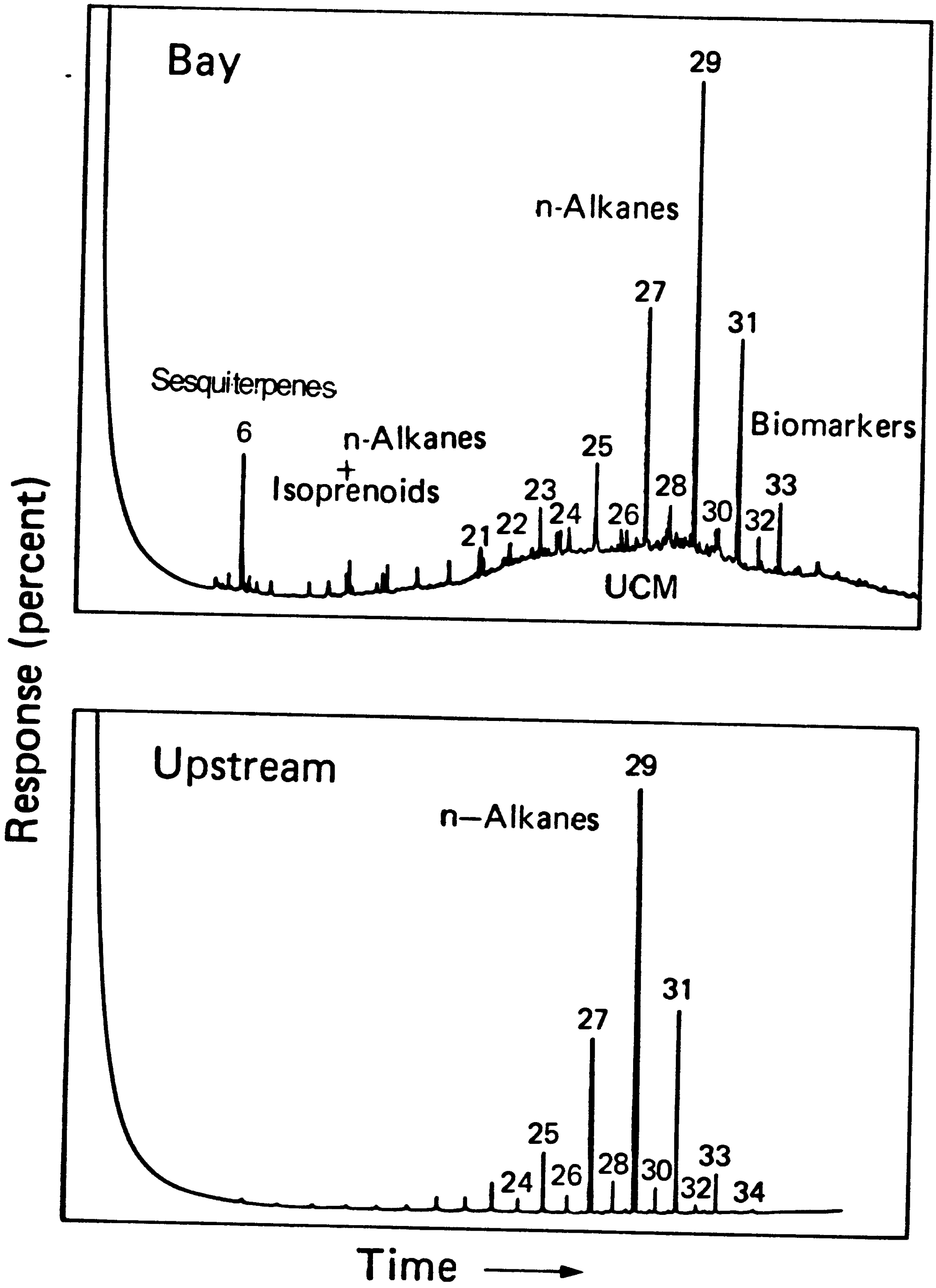
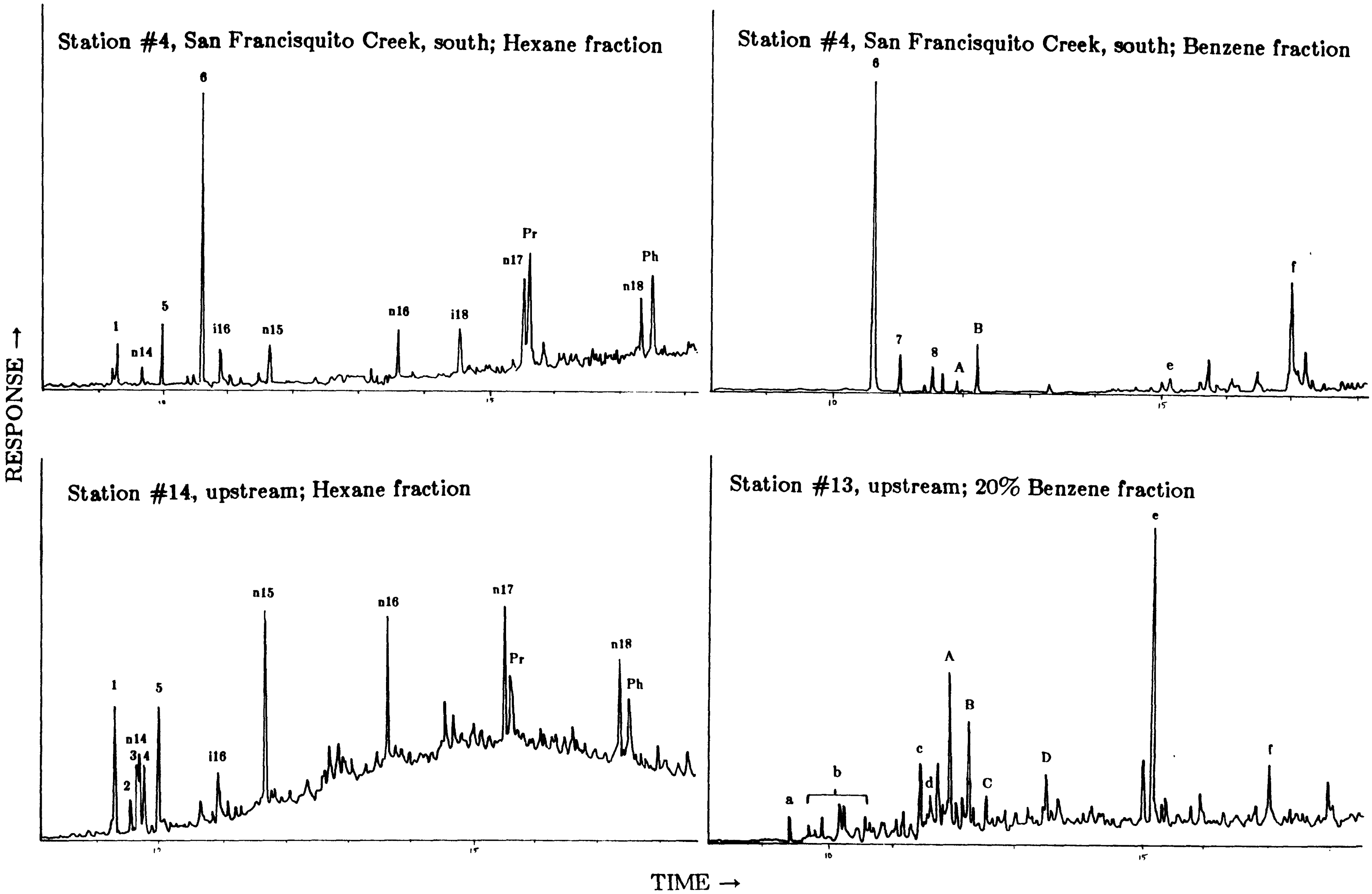


Figure 2. Chromatography of hexane fractions. Numbers 21-34 are n-alkanes; 6 is a sesquiterpene as shown in Table 2, part 2.


Station #4, San Francisquito Creek, south; Hexane fraction

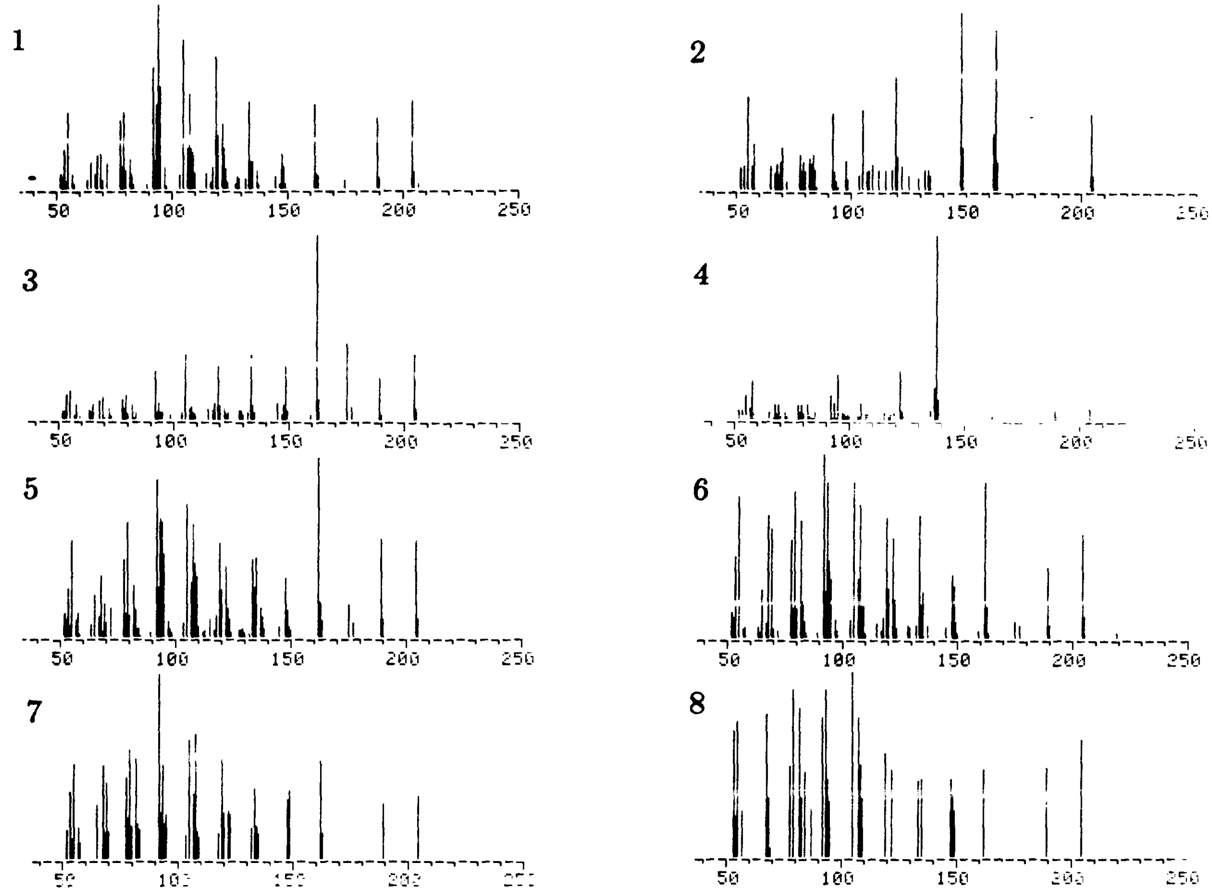
Station #4, San Francisquito Creek, south; Benzene fraction

Station #14, upstream; Hexane fraction

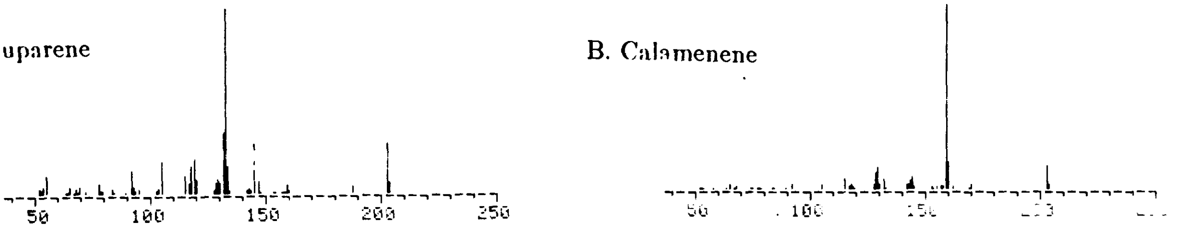
Station #13, upstream; 20% Benzene fraction

Figure 3. Sesquiterpenoids from chromatograms of representative hexane and benzene fractions

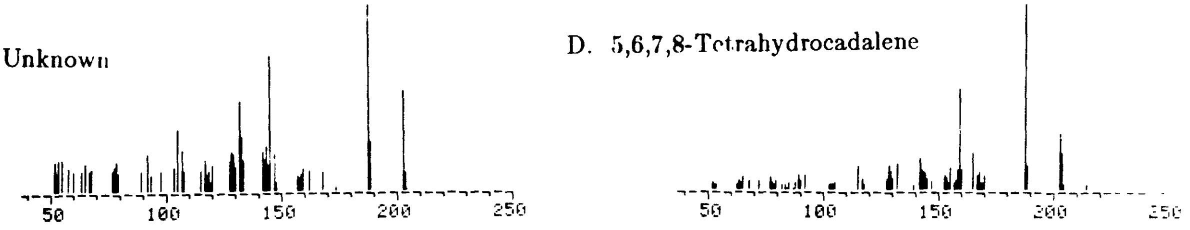
1. Molecular weight 204, $C_{15}H_{24}$ (2 rings, 2 double bonds or 3 rings, 1 double bond)



2. Molecular weight 202, $C_{15}H_{22}$ (2 rings, 1 fully aromatized)


A. Cuparene

B. Calamenene


C. Unknown

D. 5,6,7,8-Tetrahydrocadalene

3. Molecular weight 198, $C_{15}H_{18}$ (2 rings, both aromatic)

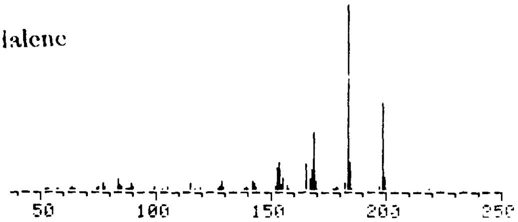

c. Cadalene

Figure 4. Mass spectra of sesquiterpenoids

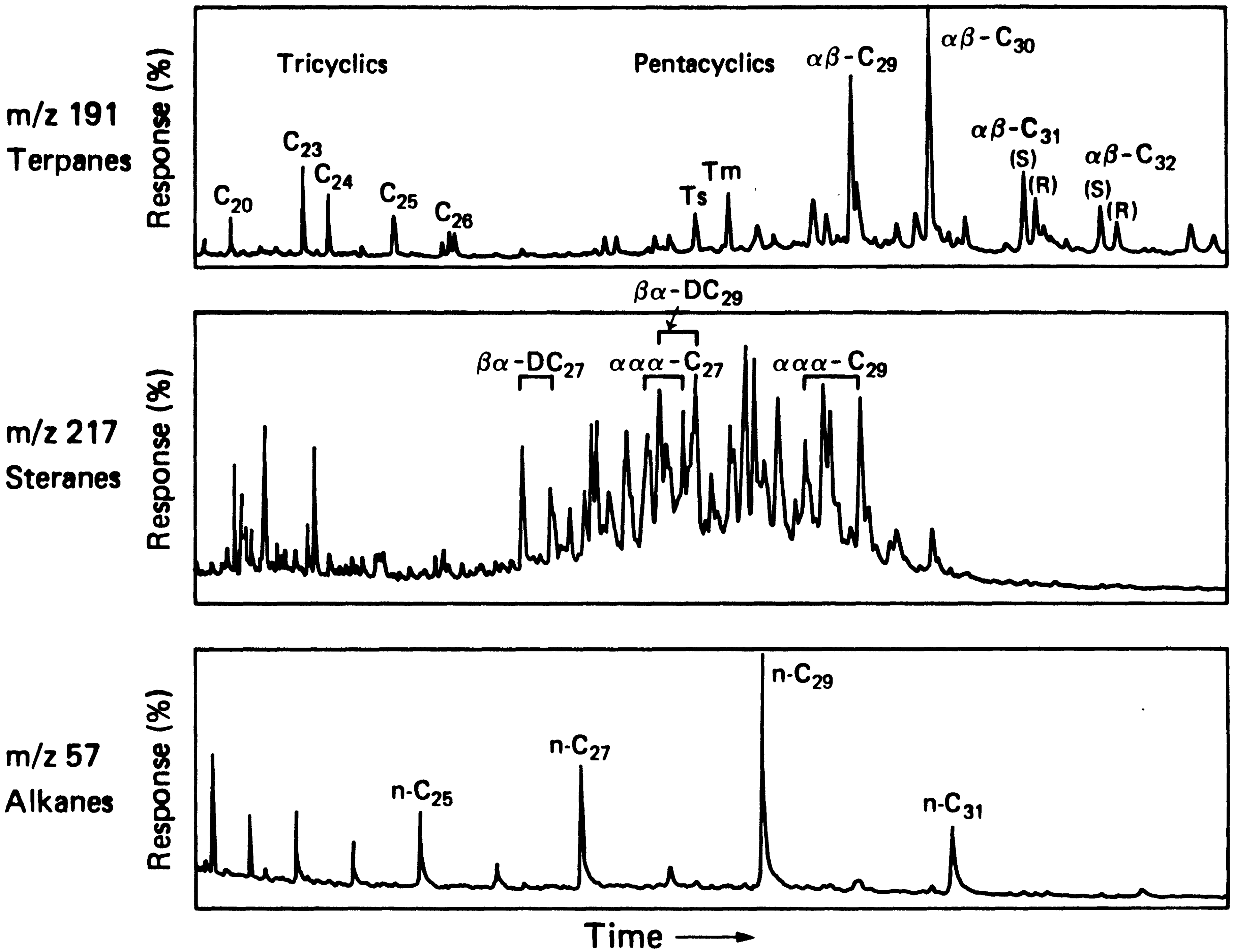


Figure 5. Chromatography--GC/MS SIM of biomarkers. For Terpanes, $\alpha\beta$ refers to compounds in the $17\alpha(H),21\beta(H)$-hopane series; for Steranes, $\alpha\alpha\alpha$ and $\beta\alpha D$ refer to compounds in the $5\alpha(H),14\alpha(H),17\alpha(H)$-cholestane and $13\beta(H),17\alpha(H)$-diacholestane series respectively.

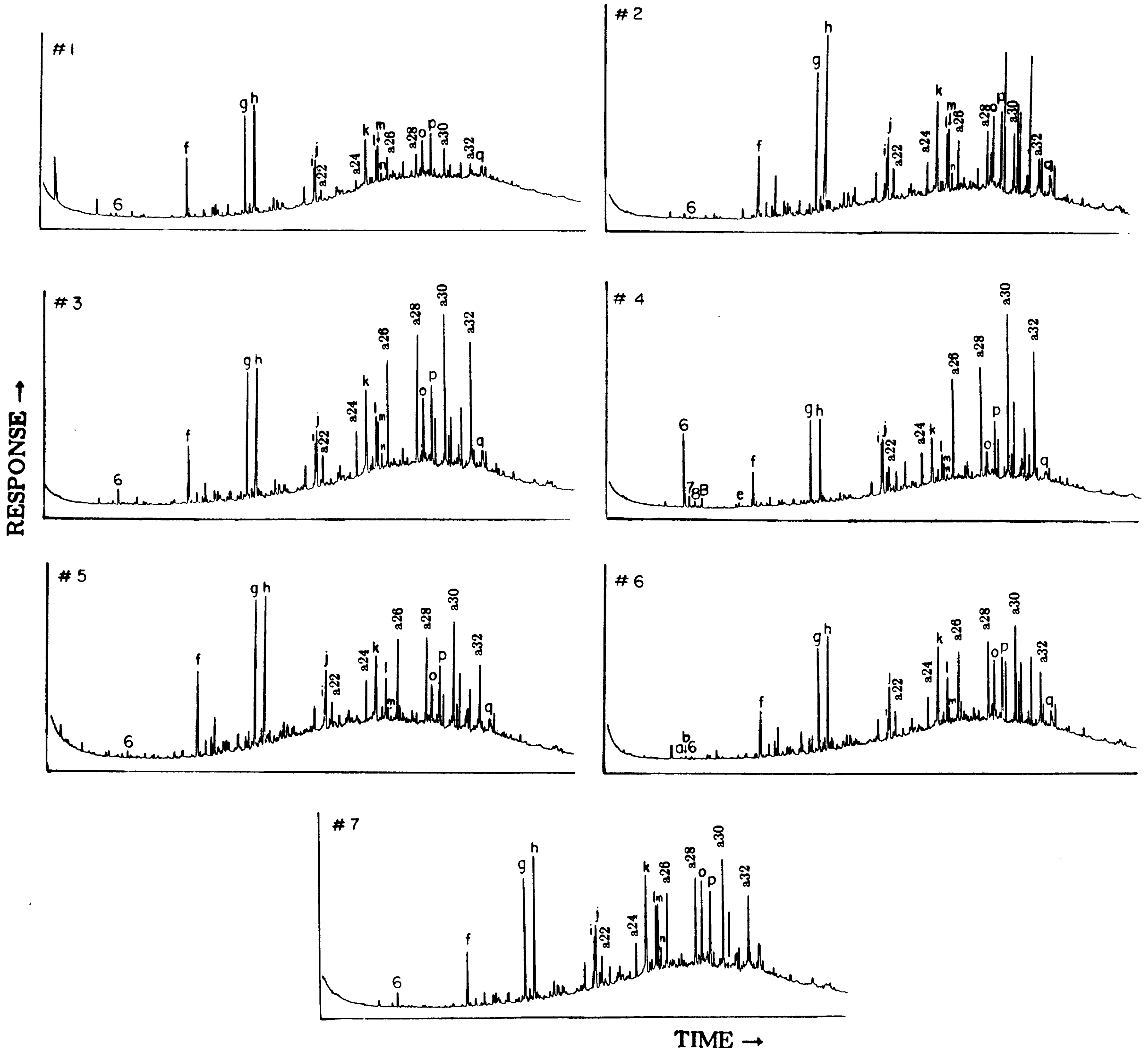
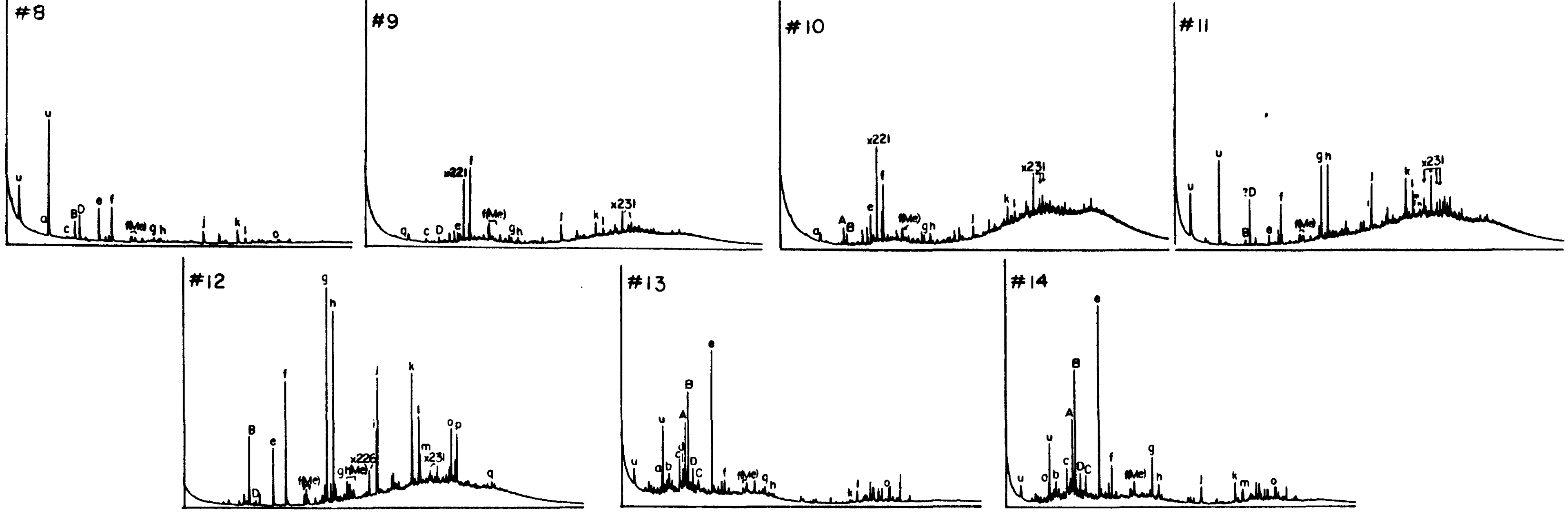


Figure 6. Chromatography of benzene fractions of Bay sediments

a. 20% benzene fractions



b. 60% benzene fractions

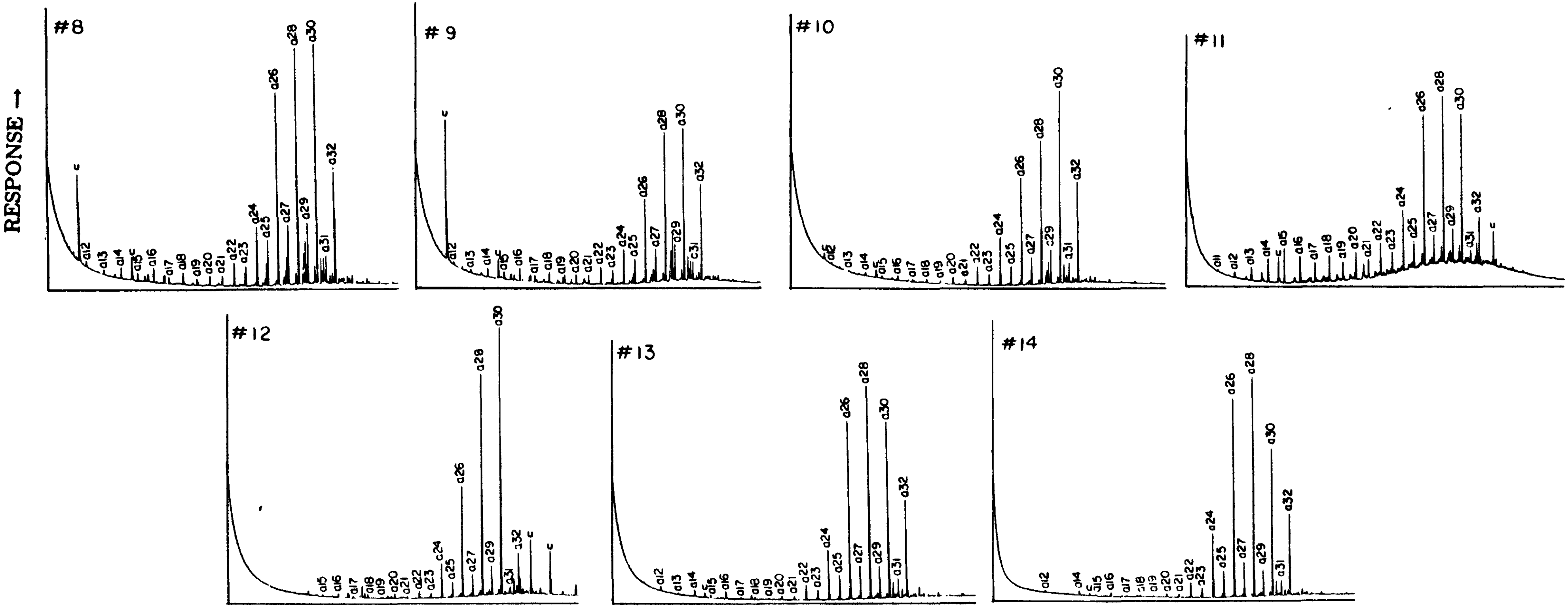


Figure 7. Chromatography of benzene fractions of upstream sediments

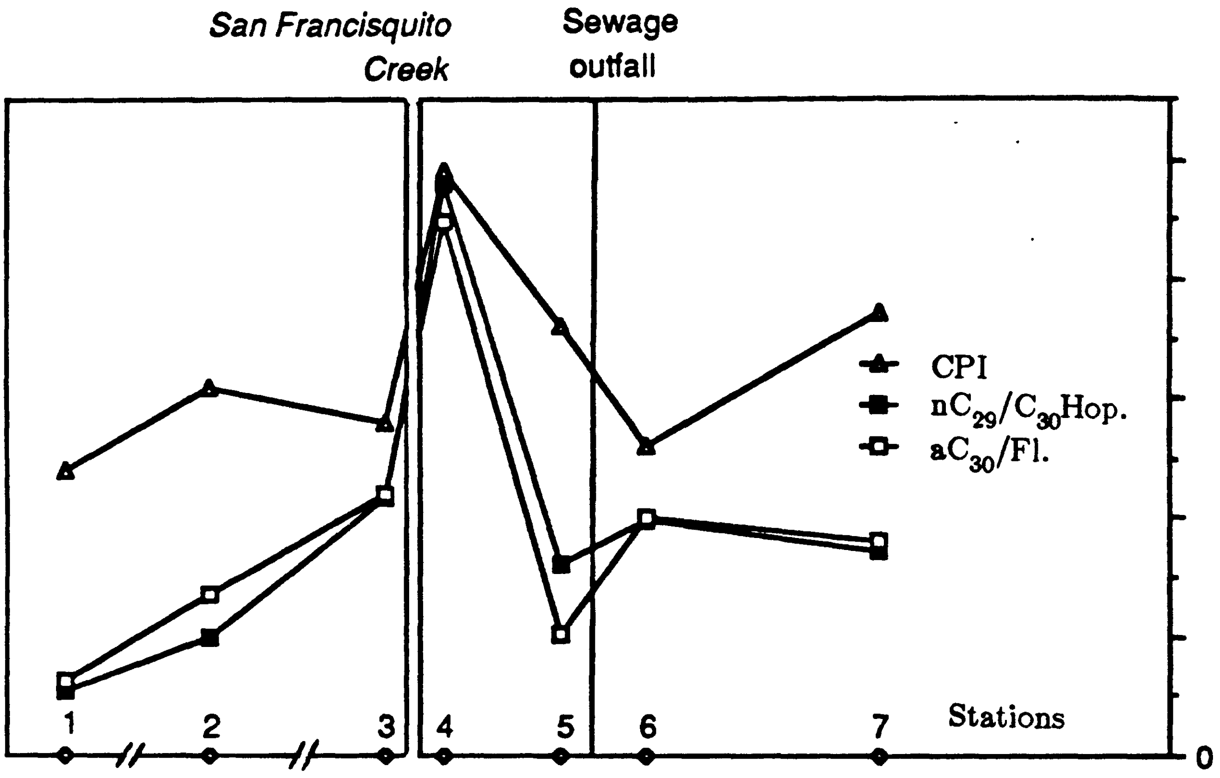


Figure 8. Results of parameter comparisons

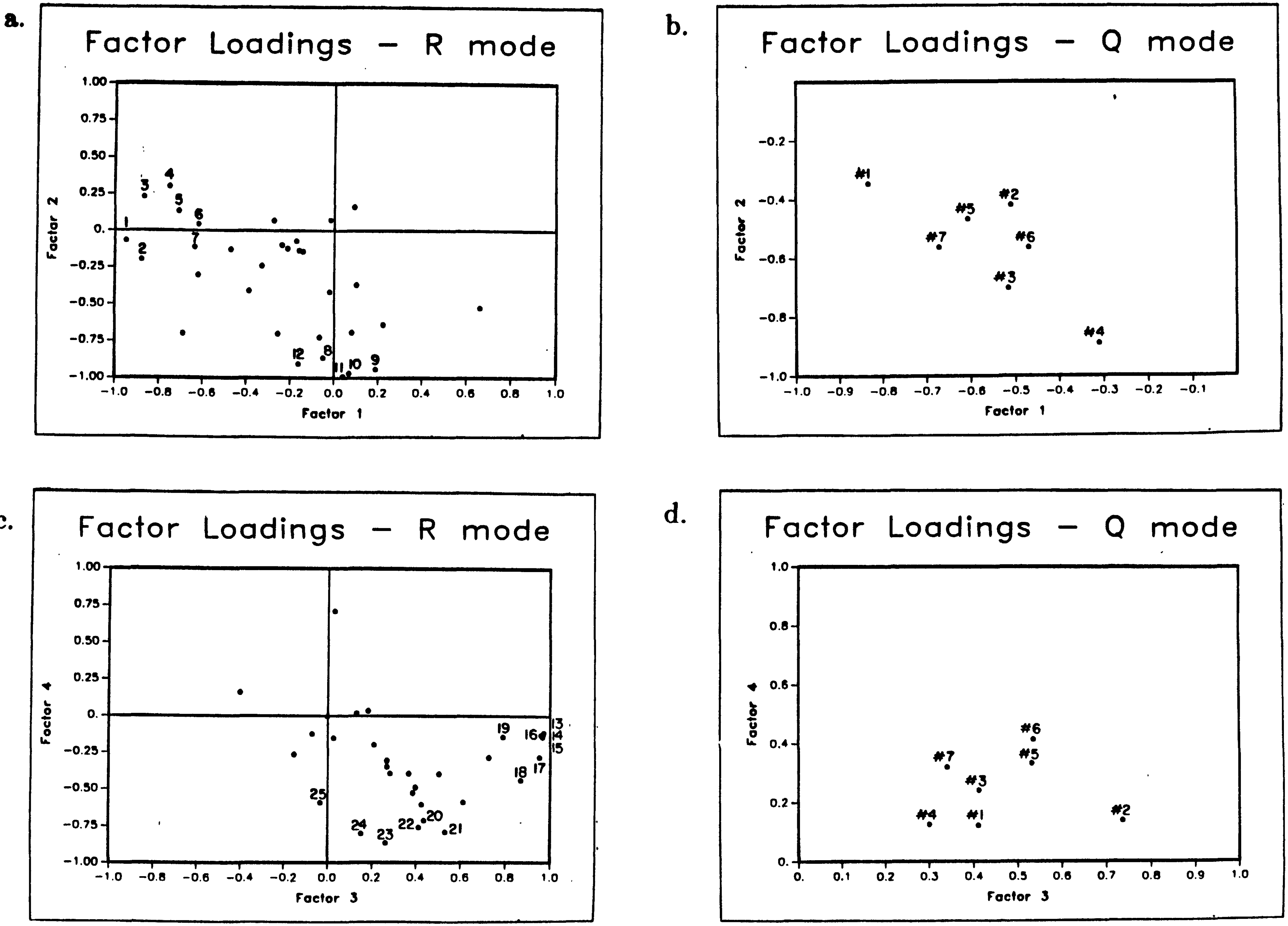


Figure 9. Multivariate analysis factor loadings

Table 1. Geochemical Analyses

Station	Location	EOM (μg/g)	Fractions, Weight(μg/g(%EOM)) Hexane	Benzene*	Methanol	TOC	[Pyrene] (ppb)	[ΣPAH] (ppb)	CPI	$\frac{n\text{-}C_{29}}{C_{30}\text{-}\alpha\beta\text{hopane}}$	$\frac{a\text{-}C_{30}}{\text{Fluoranthene}}$
South Bay Samples:											
1	San Mateo Bridge	160	31(19%)	48(30%)	78(48%)	0.90	236	1437	4.8	2.1	0.25
2	University Ave. site	405	80(20%)	62(15%)	194(48%)	1.17	531	2371	6.2	4.0	0.55
3	100 m. N. of San Francisquito Cr.	222	71(32%)	59(27%)	81(36%)	1.13	283	1841	5.6	8.7	0.88
4	100 m. S. of San Francisquito Cr.	212	54(25%)	52(25%)	92(43%)	0.60	135	997	9.8	19.2	1.79
5	100 m. N. of sewage outfall	456	132(29%)	89(19%)	176(39%)	1.04	236	1402	7.2	6.5	0.41
6	200 m. S. of sewage outfall	262	97(37%)	66(25%)	87(33%)	1.18	225	1646	5.2	7.9	0.81
7	1 km. S. of sewage outfall	197	67(34%)	49(25%)	68(35%)	0.93	268	1987	7.4	6.9	0.73
Upstream Samples:											
8	Adobe Creek, left fork	26	1.9(8%)	2.8(11%)	20(78%)	0.29	1.6	43	11.7	>400	34
9	Adobe Creek, middle fork	28	6.9(24%)	7.2(25%)	13(46%)	0.23	1.8	45	8.4	34	40
10	Adobe Creek, Bunny Cr. fork	59	16(27%)	18(31%)	17(29%)	0.31	3.2	110	10.6	89	72
11	San Francisquito Cr. by Univ. Ave.	76	21(28%)	19(25%)	25(33%)	--	18	122	3.1	3.0	1.4
12	Adobe Creek, Los Altos	358	126(35%)	105(29%)	71(20%)	0.36	493	2897	5.8	19	1.5
13	Los Trancos Creek, Foothill Park	82	8.5(10%)	19(23%)	39(48%)	--	nd	50	11.8	>500	22
14	Los Trancos Creek, Foothill Park	47	10.1(22%)	18(38%)	16(34%)	--	3.6	78	12.5	750	34

EOM = Extractable Organic Material
μg/g = μg of fraction/g of dry sediment
*Benzene: South Bay samples = fraction B; Upstream samples = Σ20,40,60&100% benzene eluates
TOC = Total Organic Carbon
CPI = Defined in text

Table 2. Identification of fraction constituents.

Peak	Compound
1. Polycyclic aromatic hydrocarbons (PAH's)	
a	Biphenyl
b	C_2-naphthalene
c	3-Methylbiphenyl
d	2-Methylbiphenyl
e	Cadalene
f	Phenanthrene
g	Fluoranthene
h	Pyrene
i	Benzanthracene
j	Chrysene
k	Benzo(b) and (k) fluoranthenes
l	Benzo(e)pyrene
m	Benzo(a)pyrene
n	Perylene
o	Indeno(1,2,3-cd)pyrene
p	Benzo(ghi)perylene
q	Coronene
2. Aliphatic and partially aromatized hydrocarbons	
n#	n-alkane of # carbon chain length
i#	isoprenoid of # carbon atoms
Pr	Pristane (i19)
Ph	Phytane (i20)
1-8	Major sesquiterpenoid hydrocarbon isomers, MW 204
A-D	Major sesquiterpenoid hydrocarbon isomers, MW 202
	A. Cuparene
	B. Calamanene
	C. Unknown
	D. 5,6,7,8-Dihydrocadalene
3. Aldehydes	
a#	n-aldehyde of # carbon chain length
4. Unknowns	
x#	Unknown, # = base peak on MS
u	Unknown, not relevant

Table 3. Compound Identifications for Factor Analysis

1. Benzo(a)Pyrene
2. Perylene
3. 2-Me Phenanthrene
4. Unknown
5. Phenanthrene
6. Pyrene
7. Fluoranthene
8. n-C_{24} Aldehyde
9. n-C_{26} Aldehyde
10. n-C_{28} Aldehyde
11. n-C_{30} Aldehyde
12. n-C_{32} Aldehyde

13-20. Unknowns

21. Benzo(b)Fluoranthene
22. Benzo(e)Pyrene
23. Indeno(1,2,3-cd)Pyrene
24. Benzo(ghi)Perylene
25. Coronene