Biomarker Fingerprinting: Application and Limitation for Correlation and Source Identification of Oils and Petroleum Products

Zhendi Wang, M. Fingas, C. Yang and B. Hollebone*

Emergencies Science and Technology Division
Environmental Technology Centre, Environment Canada
335 River Road, Ottawa, Ontario, Canada K1A 0H3
Tel: (613) 990-1597, Fax: (613) 991-9485
E-mail: zhendi.wang@ec.gc.ca

INTRODUCTION

Biological markers or biomarkers are complex molecules derived from formerly living organisms. **Figure 1** shows molecular structures of representative cyclic biomarker compounds. Biomarkers are useful because they retain all or most of original carbon skeleton of the original natural product and this structural similarity reveals more information about their origins than other compounds.

Biomarker fingerprinting has historically been used by petroleum geochemists (1) in characterization of oils in terms of the type(s) of precursor organic matter in the source rock, correlation of oils with each other and oils with their source rock, determination of depositional conditions, assessment of thermal maturity of oil during burial, and evaluation of migration and the degree of biodegradation.

The conversion of a vast number of the precursor biochemical compounds from living organisms into biomarkers creates a vast suite of compounds in crude oils that have distinct structures. Further, due to the wide variety of geological conditions and ages under which oil has formed, every crude oil exhibit an essentially unique biomarker "fingerprint". Therefore, chemical analysis of source-characteristic and environmentally-persistent biomarkers generates information of great importance in determining the source of spilled oil, differentiating oils, monitoring the degradation process and weathering state of oils under a wide variety of conditions. In the past decade, use of biomarker fingerprinting techniques to study spilled oils has greatly increased, and biomarker parameters have been playing a prominent role in almost all oil spill work.

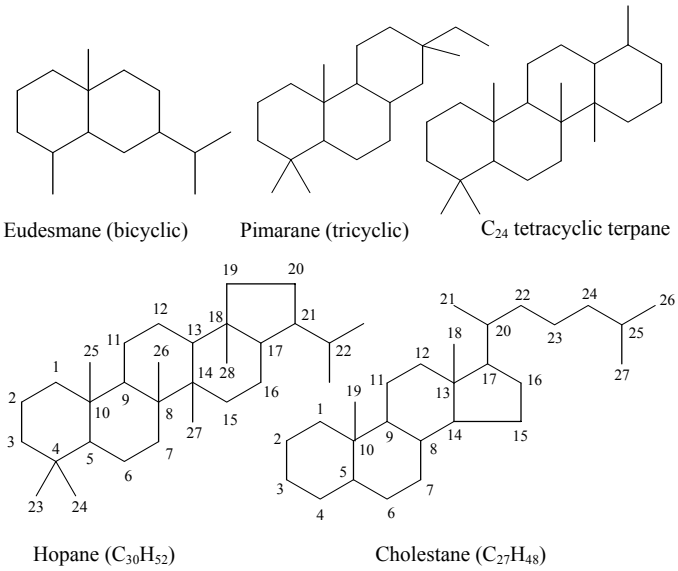

Figure 1 Molecular structures of representative cyclic biomarker compounds.

EXPERIMENTAL

Approximately 0.4 gram of oils or petroleum products were weighed, dissolved in hexane and made up to the final volume of 5.00 mL. A 200 µL of the oil solutions containing ~ 16 mg of oil was spiked with appropriate surrogates (100 µL 200 ppm of o-terphenyl and 100 µL of mixture of deuterated acenaphthene, phenanthrene, benz[a]anthracene, and perylene, 10 ppm each), and then transferred into a 3-g silica gel microcolumns, which was topped with about 1-cm anhydrous granular sodium sulfate and had been pre-conditioned using 20 mL of hexane, for sample cleanup and fractionation.

Hexane (12 mL) and 50% benznene in hexane (v/v, 15 mL) were used to elute the saturated and aromatic hydrocarbons, respectively. For each sample, half of the hexane fraction (labeled F1) was used for analysis of aliphatics, n–alkanes, and biomarker terpane and sterane compounds; half of the 50% benzene fraction (labeled F2) was used for analysis of alkylated homologous PAHs and other EPA priority unsubstituted PAHs; the remaining halves of the hexane fraction and 50% benzene fraction were combined into a fraction (labeled F3) and used for the determination of the total GC-detectable petroleum hydrocarbons (TPH) and the unresolved complex mixture of hydrocarbons (UCM). These three fractions were concentrated under a stream of nitrogen to appropriate volumes (~0.4 mL), spiked with appropriate internal standards (50 µL of 200 ppm 5-α-androstane and 50 µL of 20 ppm C_{30}-ββ-hopane, 50 µL of 10 ppm terphenyl-d_{14}, and 50 µL of 200 ppm 5-α-androstane for F1, F2, and F3 respectively), and then adjusted to an accurate pre-injection volume of 0.50 mL for GC-FID and GC-MS analyses (2, 3).

Analyses for *n*-alkane distribution (n-C_8 through n-C_{41}, pristane and phytane) and TPH were performed on a Hewlett-Packard (HP) 5890 gas chromatograph equipped with a flame-ionization detector (FID) and an HP 7683 autosampler. Analyses of target PAH compounds (including 5 alkylated PAH homologous groups and other EPA priority PAHs) and biomarker terpanes and steranes were performed on an HP 6890 GC equipped with a HP 5973 mass selective detector (MSD). System control and data acquisition were achieved with an HP G1701 BA MSD ChemStation. The individual biomarker compounds were determined using GC-MS in the selected ion monitoring mode (SIM).

RESULTS AND DISCUSSION

Biomarker Distributions

Generally, oil biomarkers of interest to environmental forensic investigation can be categorized into two classes: acyclic or aliphatic biomarkers, and cyclic biomarkers.

The terpanes in petroleum include sesqui- (C_{15}), di- (C_{20}), sester- (C_{25}), and triterpanes (C_{30}). Terpanes are found in nearly all oils. These bacterial terpanes include several homologous series, including bicyclic, tricyclic, tetracyclc, and pentacyclic (e.g., hopanes) compounds. Hopanes with the 17 αβ-configuration in the range of C_{27} to C_{35} are characteristic of petroleum because of their greater thermodynamic stability compared to other epimeric (ββ and βα) series. The steranes are a class of biomarkers containing 21 to 30 carbons that are derived from sterols, and they include regular steranes, rearranged diasteranes, and mono- and tri-aromatic steranes. Among them, the regular C_{27}-C_{28}-C_{29} homologous sterane series are the most common and useful steranes because they are highly specific for correlation.

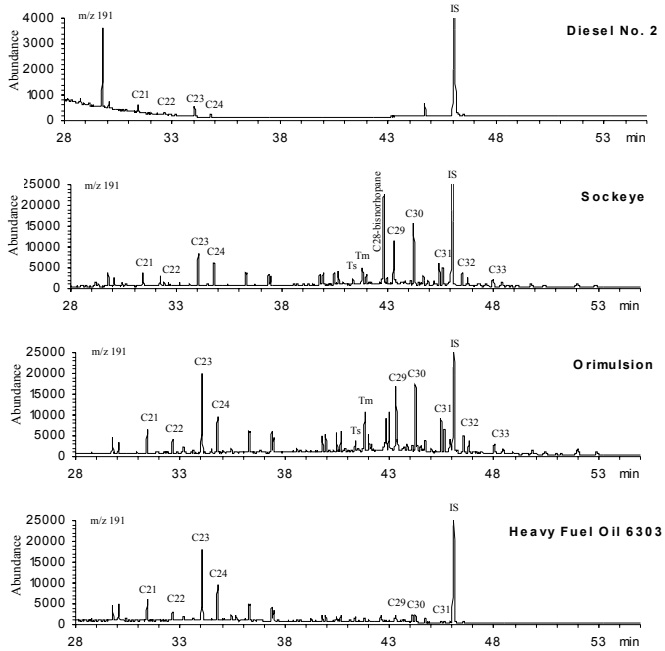

Figure 2 Distribution of biomarker terpane compounds (at m/z 191) for Diesel No. 2, Sockeye oil, Orimulsion, and HFO 6303, illustrating the differences in the relative distribution of terpanes between oils and oil products.

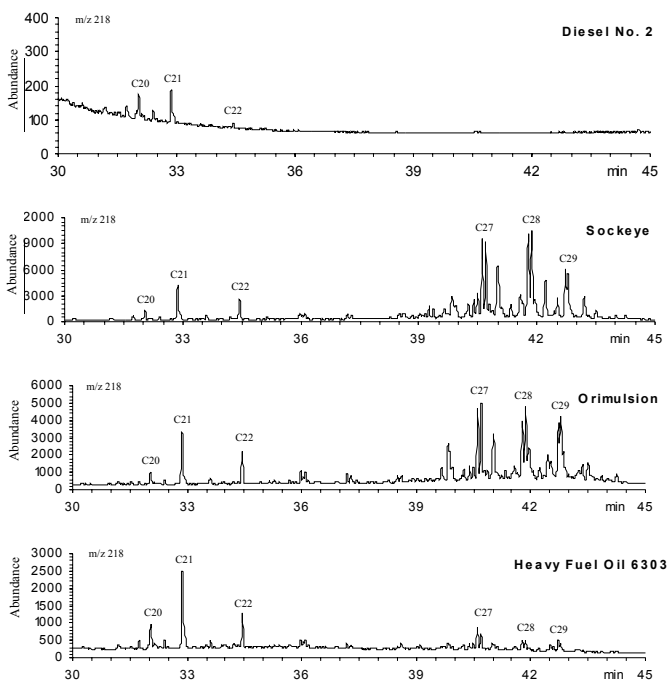

Figure 3 Distribution of biomarker sterane compounds (at m/z 218) for Diesel No. 2, Sockeye oil, Orimulsion, and HFO 6303, illustrating the differences in the relative distribution of steranes between oils and oil products.

Characterization of these biomarker compounds are achieved by using GC-MS at m/z 191 for tricyclic, tetracyclc and pentacyclic terpanes, at m/z 123 for bicyclic sesquiterpanes, at m/z 217 and 218 for steranes, at m/z 217/259 for diasteranes, at m/z 253 for mono-aromatic steranes, and at m/z 231 for tri-aromatic steranes. For many oils, the GC-MS chromatograms of terpanes (m/z 191) are characterized by the terpane distribution in a wide range from C_{20} to C_{30} often with C_{23} and C_{24} tricyclic terpanes and C_{29} αβ- and C_{30} αβ-pentacyclic hopanes being prominent. As for steranes (at m/z 217 and 218), the dominance of C_{27}, C_{28}, and C_{29} 20S/20R homologues, particularly the epimers of αββ-steranes, among the C_{20} to C_{30} steranes is often apparent.

The distribution patterns of biomarkers are, in general, different from oil to oil and from oil to refined products. **Figures 2** and **3** show GC-MS chromatograms at m/z 191 and 218 for Sockeye (California), Orimulsion-400 (Venezuela), HFO 6303, and Diesel No.2 (Ontario). The differences in the relative distribution of terpanes and steranes between Sockeye oil and Orimulsion are very apparent. For Sockeye, C_{28}-bisnorhopane, C_{29} and C_{30} αβ hopane are the most abundant with the concentration of C_{28} being even higher than C_{29} and C_{30} hopane. For Orimulsion, C_{23} terpane is the most abundant and the concentration of C_{29} is lower than C_{30} hopane. For HFO 6303, C_{23} terpane is the most abundant, but nearly no homohopanes of C_{31} to C_{35} were detected. Different from most Bunker C type oils, the concentrations of both terpanes and steranes are quite low in HFO 6303. As for refined products, only trace of C_{20} to C_{24} terpanes and C_{20} to C_{22} steranes were detected in Diesel No. 2. In contrast, most lube oils contain very high quantity of biomarkers. Obviously, refining processes have removed or concentrated high molecular mass biomarkers from the corresponding crude oil feed stocks.

Biomarker distribution in weathered oil

Biomarkers are source-characteristic and highly degradation-resistant. Therefore, for severely weathered oils, they may exhibit completely different GC-FID chromatograms and n-alkane profiles and isoprenoid distribution from their source oil, but their biomarker distribution patterns may be unaltered. Thus the fingerprinting of terpane and sterane biomarkers provides us with a powerful tool for tracking the source of the long-term weathered oil. Characterization of many long-term spilled oils demonstrated that n-alkane and isoprenoid compounds in these severely weathered oil samples can be completely lost, but, the profiles of their GC-MS fingerprints at m/z 191 and 217/218 can be nearly identical. Furthermore, the computed diagnostic ratios of a series of target pairs of biomarker compounds can be also nearly identical, clearly indicating that these samples were from the same source.

Biomarker distribution in petroleum products with similar chromatographic profiles

On some other cases where two oils may exhibit similar or even nearly identical n-alkane and isoprenoid distributions, however, their biomarker distribution may be markedly different. Thus, the successful forensic investigation will require to perform detailed analysis and comparison of not only the concentrations but also the diagnostic ratios of biomarkers among similar product types. **Figure 4** shows the GC-FID chromatograms of three unknown oil samples received form Montreal on March of 2001 for product identification and differentiation. Three samples show nearly identical GC chromatographic profiles and distribution patters. The relative ratios of low abundant hydrocarbons n-C_{17}/n-C_{18}, n-C_{17}/pristane and n-C_{18}/phytane are also very similar. All the GC trace features suggest that the oils are hydraulic fluid type products. The questions which must be answered at this stage are: do these three oil samples really come from the same source or not? The GC-MS analysis for biomarker characterization (**Figure 5**) indicate that samples #1 and

#2 are nearly identical in distribution patterns and profiles of terpanes and steranes, but sample #3 shows markedly different distribution pattern of biomarkers from sample #1 and #2. The concentrations of C_{23} and C_{24}, and the sum of C_{31} through C_{35} homohopanes in sample #3 are significantly lower and higher than the corresponding compounds in samples #1 and #2, respectively. The biomarker analysis results, in combination with other GC analysis results, clearly demonstrated that sample #1 and #2 are identical and from the same source, and the sample #3 is different from sample #1 and #2 and does not come from the same source as samples #1 and #2 do.

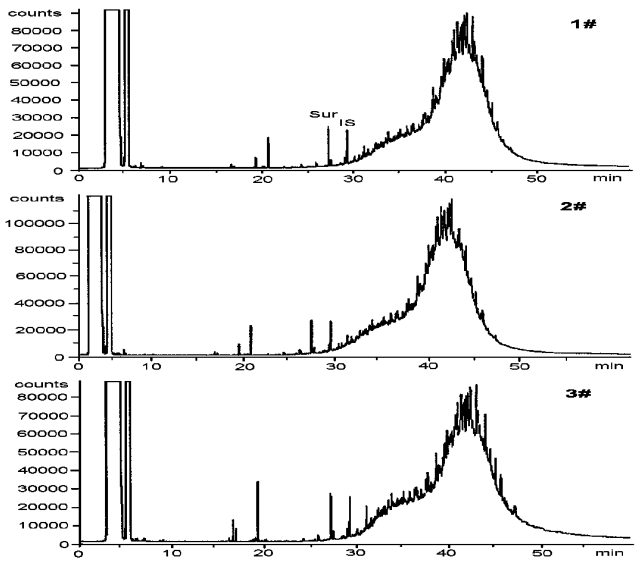

Figure 4 GC-FID chromatograms of Fraction 3 for *n*-alkane and TPH analysis of three unknown oil samples. These three samples show very similar GC chromatographic profiles and distribution patters.

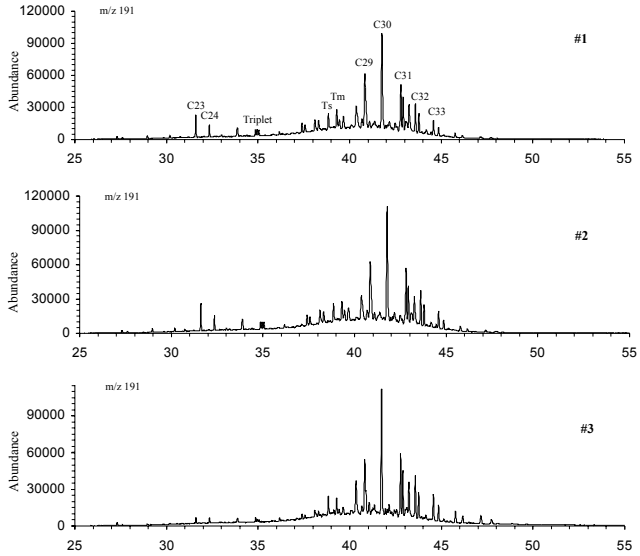

Figure 5 Distribution of biomarker terpane compounds (m/z 191) for the three unknown oil samples. Samples #1 and #2 are nearly identical in distribution patterns and profiles of terpanes, but sample #3 shows markedly different distribution pattern of biomarkers.

Sesquiterpane and Diamondoid Biomarkers in Oils and Lighter Petroleum Products

Refining processes have removed most high-molecular weight biomarkers from the corresponding crude oil feed stocks. Therefore the high boiling point pentacyclic triterpanes and steranes are generally absent in lighter petroleum products jet fuels and most diesels. However, several bicyclic sesquiterpanes including eudesmane and drimane (C_{15}), which boil in diesel range and are ubiquitous in sediments and crude oils, are still present in diesels and certain jet fuels with significant abundances after refining processes. Examination of GC-MS chromatograms of these bicyclic biomarkers at m/z 123, 179, 193, and 207 provide a comparable and highly diagnostic means of comparison for diesel type products (1, 4).

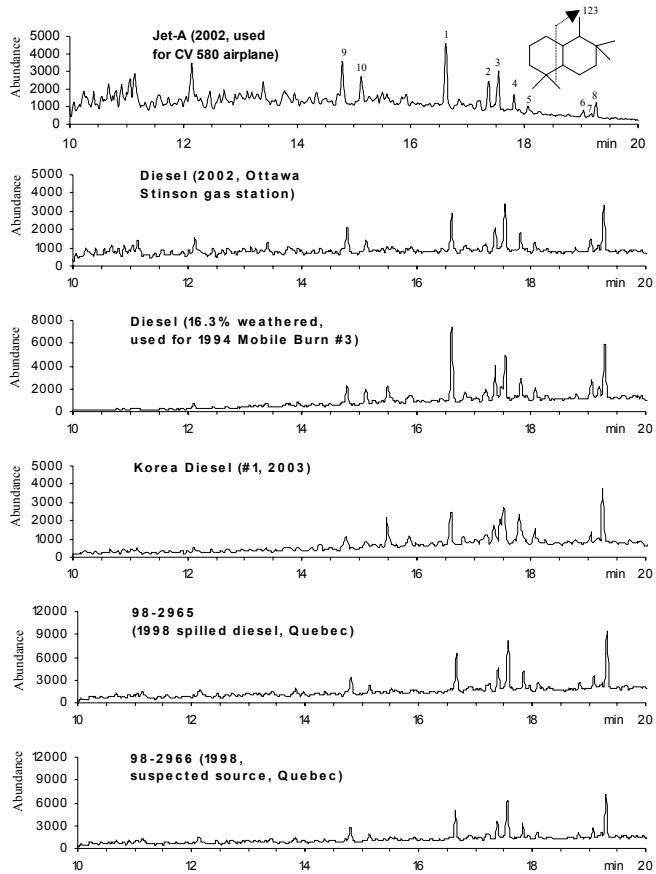

Figure 6 GC-MS chromatograms at m/z 123 for sesquiterpane analysis of Jet A (2002), Diesel No.2 (2002, from Ottawa Stinson Gas Station), Diesel No.2 (for 94 Mobile Burn, 16.3% weathered), a Korea Diesel (#1, 2003, from Korea), 1998-spilled-diesel (from Quebec) and 1998-suspected-source diesel (from Quebec). The different distributions of the sesquiterpanes demonstrates the differences between diesels and between diesel and jet fuel.

Another promising group of low boiling cyclic biomarkers of interest for environmental forensic investigators are "diamondoid" hydrocarbons (the collective term of adamantane (C_{10}), diamantane (C_{14}), and their alkyl homologous series). Diamondoids are rigid, three-dimensionally fused cyclohexane-ring alkanes that have a diamond-like (cage-like) structure. The diamond structure endows these molecules with an unusually high thermal stability and high resistance to biodegradation. Adamantanes and diamantanes are examined at m/z 135, 136, 149, 163, 177, and 191, and at m/z 187,

188, 201, 215 and 229, respectively. The sesquiterpane and diamondoid biomarkers are useful and promising for source correlation and differentiation of refined products, in particular for the weathered refined products, because of their stability and resistance to biodegradation

Figure 6 shows the GC-MS chromatograms at m/z 123 for sesquiterpane analysis of a jet fuel, Diesel No.2 (2002, from an Ottawa Gas Station)), Diesel No.2 (for 94 Mobile Burn, 16.3% weathered), Diesel No.2 (2003, from Korea), 1998-spilled-diesel (from Quebec) and 1998-suspected-source diesel (from Quebec). The different distributions of the sesquiterpanes demonstrates the differences between diesels and between diesel and jet fuel. **Figure 6** also clearly shows the spilled-diesel had nearly identical GC chromatogram with the suspected-source diesel (from Quebec). These similarities, in combined with other GC analysis results (such as hydrocarbon groups, n-alkanes and PAHs), argued strongly that the spilled diesel was from the suspected source.

Unique Biomarker Compounds

Biomarker terpanes and steranes are common constituents of crude oils. However, a few "specific" biomarker compounds including several geologically rare acyclic alkanes are found to exist only in certain oils and, therefore, can be used as unique markers to provide an interpretational advantage in fingerprinting sources of spilled oils and to provide additional diagnostic information on the types of organic matter that give rise to the crude oil. For example, the geologically rare acyclic alkane botryococcane ($C_{34}H_{70}$) was used to identify a new class of Australian non-marine crude oils. The presence of botryococcane indicates that the source rock contains remains of the algae *Botryococcus braunii*.

Other "specific" pentacyclic terpanes include C_{30} 17α (H)-diahopane (it may be related to bacterial hopanoid precursors that have undergone oxidation and rearrangement by clay-mediated acidic catalysis), β-carotane ($C_{40}H_{78}$, highly specific for lacustrine deposition, highly abundant in Green River Shale), gammacerane ($C_{30}H_{52}$, it has been tentatively suggested as a marker for hypersaline episodes of source rock deposition), lupanes and bisnorlupanes (they are believed to indicate terrestrial organic matter input), and bicadinanes (highly specific for resinous input from higher plants).

Diagnostic Ratios of Biomarkers

The diagnostic parameters of biomarkers most used in oil spill fingerprinting include: tricyclic C23/C24, C29 αβ/C30 αβ hopane, oleanane/C30 αβ hopane, homohopane (C31 to C35) distribution, Ts/Tm & Ts/(Ts + Tm), triplet ratio (C24 tertracyclic terpane/C26 tricyclic (S)/C26 tricyclic (R) sterane), regular sterane distribution (C27-C28-C29 ααα and αββ steranes (20S + 20R), C26-C27-C28 triaromatic steranes (TA) and C27-C28-C29 monoaromatic steranes (MA) distribution, and others.

The biomarker fingerprinting results strongly suggest a "*basic rule*" in the environmental forensic investigations: a negative correlation of biomarkers is strong evidence for lake of relationship between samples, and a positive correlation of biomarkers is not necessarily "proof" that samples are related because some oils from different sources can show similar characteristics of biomarkers. In order to reliably and defensively to correlate or differentiate samples, the "multi-parameter approach" must be initiated, that is, analyses of more than one suite of analytes must be performed.

Biodegradation of Biomarkers

It has well recognized that terpane and sterane compounds are very resistant to biodegradation. The studies of biodegradation of nine Alaska oils and oil products and eight Canadian oils by a defined bacterial consortium incubated under freshwater and cold/marine conditions in our lab showed that no changes of biomarkers after incubation, despite extensive saturate and aromatic losses. Also, the ratios of selected paired biomarkers remained constant. Therefore, biomarkers can, in many cases, be and have been used as conserved internal references for estimation of oil weathering percentages.

However, it does not mean cyclic biomarkers can not be biodegraded. Actually, they are still biodegradable in severe weathering conditions. In a recent study on long-term fate and persistence of the spilled Metula oil in a marine marsh environment (5), we found that in highly degraded asphalt pavement samples, even the most refractory biomarker compounds showed various degrees of biodegradation. The degree of biodegradation of biomarkers was not only molecular mass and size dependent, but also stereo-isomer dependent. The biomarkers were generally degraded in the declining order of importance as: diasterane > C_{27} steranes > tricyclic terpanes > pentacyclic terpanes > norhopanes ~ C_{29} αββ-steranes. The degradation of steranes was in the order of $C_{27} > C_{28} > C_{29}$ with the stereochemical degradation sequence 20R ααα steranes > (20R + 20S) αββ steranes > 20S ααα steranes. C29-18α(H), 21β(H)-30 norneohopane and C29 αββ 20R and 20S stigmastanes appeared to be the most biodegradation-resistant terpane and sterane compounds, respectively, among the studied target biomarkers.

REFERENCES

1. Peters, K. E., Moldowan, J. W., *The Biomarker Guide: Interpreting Molecular Fossils in Petroleum and Ancient Sediments*, Prentice Hall, New Jersey, **1993**.
2. Wang, Z. D., Fingas, M., Li, K., *J. Chromatogr. Sci.*, **1994**, 32, 361-366 (Part I) and 367-382 (Part II).
3. Wang, Z. D., Fingas, M., Sergy, G., *Environ. Sci. Technol.*, **1994**, 28, 1733-1746.
4. Stout, S. A., Uhler, A. D., McCarthy, K. J., Emsbo-Mattingly, S., in: *Introduction to Environmental Forensics* (B. L. Murphy and R. D. Morrison, eds.), Academic Press, London, **2002**, Chapter 6.
5. Wang, Z. D., Fingas, M., Owens, E. H., Sigouin, L., Brown, C.E., *J. Chromatogr.*, **2001**, 926, 275-190.