

Original Paper

Chemometric differentiation of oil families in the Mahu sag, Junggar Basin, NW China

Hang-Xin Cai^{a,d}, Jun Jin^b, Er-Ting Li^b, Zhong-Da Zhang^c, Shuang Yu^a, Chang-Chun Pan^{a,*}^a State Key Laboratory of Deep Earth Processes and Resources, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou, 510640, Guangdong, China^b Research Institute of Experiment and Testing, Xinjiang Oilfield Company, PetroChina, Karamay, 834000, Xinjiang, China^c Research Institute of Petroleum Exploration and Development, Working Stations for Postdoctors of Shengli Oilfield Company, SINOPEC, Dongying, 257000, Shandong, China^d University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Article history:

Received 26 February 2024

Received in revised form

4 January 2025

Accepted 26 May 2025

Available online 28 May 2025

Edited by Xi Zhang

Keywords:

Oil source assessment

Chemometric analysis

Carbon isotopes of individual *n*-alkanes

Biomarkers

Polynuclear aromatic hydrocarbons (PAH)

Junggar Basin

ABSTRACT

Most oil reservoirs that were found in the Junggar Basin are located in the Mahu sag and neighboring areas. Oil sources and classifications remain unresolved in this region. Oil source assessment can be partially inconsistent on the basis of different molecular and isotopic parameters. In the present study, classifications for the 92 studied oils from the Mahu sag and neighboring areas were performed using chemometric analysis, e.g., hierarchical cluster analysis (HCA) and principal component analysis (PCA) on the basis of integration of sixteen facies parameters. These parameters consist of isotope reversal index (RI), $\delta^{13}\text{C}$ of *n*-C₂₅, Ph/*n*-C₁₈, β -carotane/*n*-C₂₁, six terpene ratios of Ts/C₂₃ tricyclic terpanes, Ts/(C₂₈+C₂₉ tricyclic terpanes), C₂₉ Ts/C₂₃ tricyclic terpanes, C₂₉Ts/(C₂₈+C₂₉ tricyclic terpanes), C₃₀ di-hopane/C₂₃ tricyclic terpene and C₃₀ di-hopane/(C₂₈+C₂₉ tricyclic terpanes), and six ratios of polynuclear aromatic hydrocarbons (PAH) including trimethylnaphthalenes (TMNs)/(TMNs + phenanthrene (Phen)), tetramethylnaphthalenes (TeMNs)/(TeMNs + Phen), TMNs/(TMNs + methylphenanthrenes (MPs)), TeMNs/(TeMNs + MPs), TMNs/(TMNs + chrysene (Ch)) and TeMNs/(TeMNs + Ch). These sixteen parameters are mainly influenced by source facies and less influenced by maturity as demonstrated in the crossplots of these sixteen parameters versus concentrations of C₃₀ hopane. Oil classifications are more reliable and convenient using chemometric analysis (HCA and PCA) integrating the sixteen facies parameters, compared with using crossplots of two parameters or star charts of several parameters. The 92 oils are classified into three groups using HCA and PCA, i.e., Group I, II and III. Group I and II oils are derived from source rocks within the Lower Permian Fengcheng Formation (P_{1f}) and Middle Permian Lower Wuerhe Formation (P_{2w}), respectively. Group III oils are mixtures of Group I and II oils. Group I consists of fifty oils mainly located at the northeastern and central areas of the Mahu sag with only three oils at the southwestern area of the Mahu sag. Group II consists of fourteen oils at the southwestern area of the Mahu sag. Group III consists of twenty-eight oils located at the southwestern and central areas of the Mahu sag. Locations of Group I, II and III oils reflect the distributions of effective source rocks containing oil-prone Type I/II kerogen within the Fengcheng (P_{1f}) and Lower Wuerhe formations (P_{2w}). © 2025 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

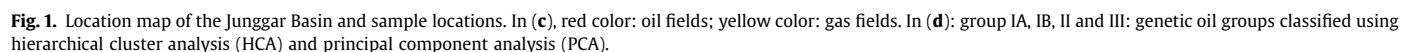
1. Introduction

The Junggar Basin is a major oil producing province of China. Most oil reservoirs are found in the northwestern areas of the basin (Fig. 1). To date, a huge amount of oil with total reserves proved in-place over 2×10^9 tonnes, but only a small amount of gas with the total reserves proved in-place about 13×10^9 m³, have been found in this region (e.g., Li et al., 2007; Wang et al., 2012;

* Corresponding author.

E-mail address: cpan@gig.ac.cn (C.-C. Pan).

Peer review under the responsibility of China University of Petroleum (Beijing).



Lower Wuerhe Formation (P_{2w}) in the northwestern Junggar Basin (Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993; Cao et al., 2005, 2006; Yu et al., 2017; Zhi et al., 2021; Wang et al., 2023).

Source rocks of the Fengcheng Formation (P_{1f}) were mainly deposited in alkaline saline lacustrine environment while source rocks within the Jiamuhe Formation (P_{1j}) and Lower Wuerhe Formation (P_{2w}) were deposited in fresh lacustrine environments (e.g., Zhang et al., 1993; Cao et al., 2015, 2020; Yu et al., 2017; Zhi et al., 2021). Source rocks within the Jiamuhe Formation (P_{1j}) contain mainly gas-prone Type III kerogen with Rock-Eval hydrogen index (HI) lower than 100 mg HC/g TOC (e.g. Cao et al., 2005; Wang et al., 2013). Source rocks within the Fengcheng Formation (P_{1f}) in northwestern border area of the basin contain oil-prone Type I/II kerogen (Jiang and Fan, 1983; Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993; Cao et al., 2005, 2020; Wang et al., 2013; Yu et al., 2017). Source rocks within the Lower Wuerhe Formation (P_{2w}) contain gas-prone Type III kerogen in the northern and central areas of the Mahu sag, but oil-prone Type I/II kerogen in the southern area of the Mahu sag and in the central area of the basin (Yang et al., 1985; Zhang et al., 1993; Cao et al., 2005, 2006; Wang et al., 2013; Yu et al., 2017). The distributions for the effective source rocks containing oil-prone Type I/II kerogen within the Lower Permian Fengcheng (P_{1f}) and Middle Permian Lower Wuerhe formations (P_{2w}) remain unclear due to too deep to drill although these two formations occur in the whole Mahu sag and the central and southern areas of the basin on the basis of seismic data (Zhang et al., 1993).

In the earlier studies (e.g., Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993), oils in the northwestern region of the basin were classified into two groups: Group I oils are derived from source rocks within the Lower Permian Fengcheng Formation (P_{1f}) while Group II oils are derive from source rocks within the Lower Permian Jiamuhe Formation (P_{1j}) below the Fengcheng Formation (P_{1f}) and the Middle Permian Lower Wuerhe Formation (P_{2w}) on the basis of molecular parameters. Group I oils and their source rocks within the Fengcheng Formation (P_{1f}) have higher Pr/*n*-C₁₇ and Ph/*n*-C₁₈ ratios and higher relative concentrations of gammacerane and carotanes but lower relative concentrations of Ts, C₂₉Ts and C₃₀ diahopane while Group II oils and their source rocks within the Jiamuhe Formation (P_{1j}) and Lower Wuerhe Formation (P_{2w}) are opposite (Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993). Later studies classified the oils into three groups: Group I and II oils are derived from source rocks within the Fengcheng Formation (P_{1f}) and Lower Wuerhe Formation (P_{2w}), respectively while Group III oils are the mixtures of Group I and II oils on the basis of $\delta^{13}\text{C}$ values of individual *n*-alkanes in the northwestern (Mahu sag) and central Junggar Basin (Yu et al., 2017; Pan et al., 2021; Zhang et al., 2022; Cai et al., 2023).

Source and maturity assessment remains a difficult task for oils which were derived from multiple source rocks with a wide range of maturities from the early to late oil generation window. Biomarker parameters are widely used in oil source correlation and assessments of oil and source rock maturities and oil biodegradation extent (e.g., Peters et al., 2005). Some biomarker parameters, such as Ts/(Ts + Tm), C₂₉Ts/(C₂₉ hopane + C₂₉Ts), C₃₀ diahopane/(C₃₀ diahopane + C₃₀ hopane) are both strongly influenced by source facies and maturity (e.g., Seifert and Moldowan, 1978, 1986; Peters et al., 2005). Zhang et al. (2022) presented a new approach to assess oil source on the basis of molecular facies parameters under maturity constraint using crossplots of one source parameter versus one maturity parameter. Concentrations of ΣC_{29} steranes and C₃₀ hopane were used as maturity indicators (Zhang et al., 2022). It is ideal to present a set of biomarker parameters that are mainly influenced by source facies but less influenced by maturity for source facies assessments for the studied oils.

Aromatic components are rarely used for oil source correlation. Requejo et al. (1996) suggested that the degree of alkylation of

polynuclear aromatic hydrocarbons (PAH) are influenced by kerogen composition and maturity of source rocks. For examples, carbonate-sourced oils contain higher relative amounts of PAH having three and four methyl groups while paralic-sourced oils contain higher relative amounts of PAH having one and two methyl groups (Requejo et al., 1996). It is possible to find some parameters of aromatic components to effectively distinguish oils that are derived from the two different source beds within the Fengcheng Formation (P_{1f}) and Lower Wuerhe Formation (P_{2w}), respectively because these two source beds formed at very different environments, i.e., alkaline saline lacustrine and fresh lacustrine, respectively (e.g., Zhang et al., 1993; Cao et al., 2020).

A large number of oil samples have been collected from the Mahu sag and systematically analyzed in the present study and previous studies (Zhang et al., 2022; Cai et al., 2023). Multivariate data are generally evaluated using chemometric analysis. Hierarchical cluster analysis (HCA) and principal component analysis (PCA) are widely employed for oil classification (e.g., Peters et al., 2005, 2013, 2016; Zhan et al., 2019; Murray and Peters, 2021; Wang et al., 2021; Zou et al., 2021). Oils in the studied region are derived from multiple sources and have a wide range of maturity. Routine biomarker parameters are influenced by both source facies and maturity. Source facies assessment and grouping for oils in the studied region remain unresolved and are partially inconsistent using different molecular and isotopic parameters. The main purposes for the present study are: (1) to select a set of source facies parameters for oil classification and source assessment, including parameters of isotopes, *n*-alkanes, isoprenoids, terpanes, steranes and polynuclear aromatic hydrocarbons (PAH) that are mainly influenced by source facies but less influenced by maturity, (2) to integrate the selected source facies parameters using hierarchical cluster analysis (HCA) and principle component analysis (PCA) for oil grouping, and (3) to demonstrate that HCA and PCA are more effective and reliable, and more convenience to perform compared with the method using a set of crossplots of source facies parameters for oil classification.

2. Geological setting

The Junggar Basin is located in the northern part of Xinjiang Uygur Autonomous Region of China, covering an area of about $130 \times 10^3 \text{ km}^2$ (Fig. 1). It is a composite stacked basin of the central landmass type, characteristic of the Precambrian crystalline and the lower Paleozoic bi-layer basement (Zhao, 1992a, 1992b). The generalized stratigraphy in the northwestern Junggar basin is shown in Fig. 2. The Middle-Upper Carboniferous consists mainly of volcanic and clastic sequences. In the Permian Period, the Junggar Basin (Fig. 1(b)) was divided into several sags and uplifts (Zhang et al., 1993). The Permian strata consist mainly of clastic rocks with some volcanic rocks in the lower part. At the end of the Permian and the beginning of the Triassic, the whole basin evolved into a unified lacustrine system. The Triassic and Jurassic strata consist of mainly sandstones, mudstones, carbonaceous mudstones and coals that were deposited in fluvial and shallow lacustrine environments. In the Cretaceous, the basement of the basin inclined to the south, leading to the Cretaceous and Cenozoic strata increasingly thicken southwards. The Cretaceous, Tertiary and Quaternary strata contain mainly alluvial and fluvial clastic sediments. In the Mahu sag and nearby areas, source rocks are located at the Carboniferous to Lower Permian Jiamuhe Formation (C-P_{1j}), Lower Permian Fengcheng Formation (P_{1f}) and Middle Permian Lower Wuerhe Formation (P_{2w}) while the reservoir rocks are located at nearly all formations from the Carboniferous to Cretaceous (Fig. 2, Zhang et al., 1993).

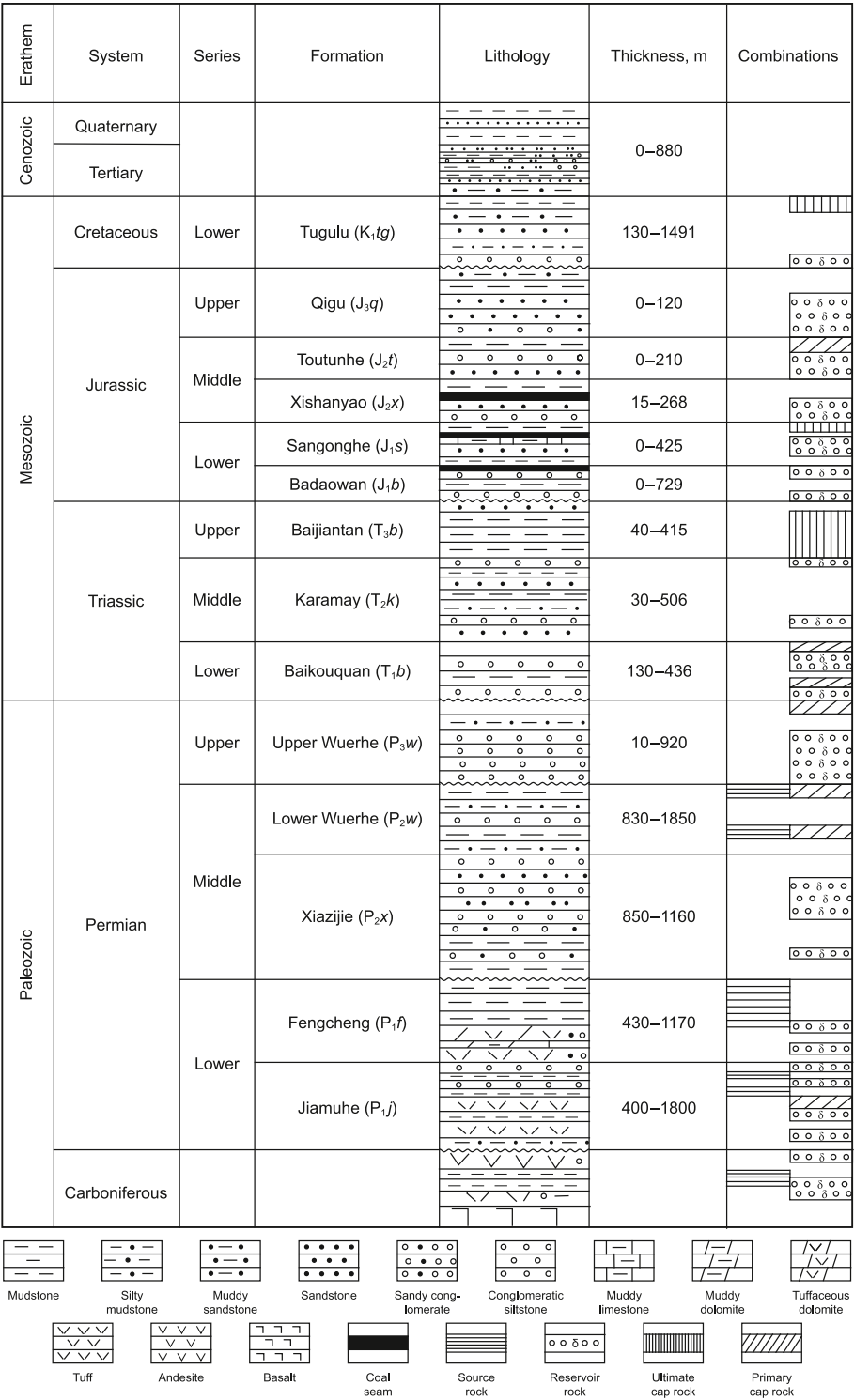


Fig. 2. Stratigraphic column of the northwestern Junggar Basin.

3. Samples and methods

3.1. Samples

Zhang et al. (2022) and Cai et al. (2023) performed quantitative GC, GC-MS and GC-IRMS analyses on 92 non-biodegraded oils

from the Mahu sag of the Junggar Basin (Fig. 1(d)), and reported the data for carbon isotopes, *n*-alkanes, isoprenoids, biomarkers, light hydrocarbons and diamondoids to document the sources and charging episodes for the studied oils. In the present study, quantitative GC-MS analyses on aromatic fractions of the 92 oils were performed.

3.2. Methods

3.2.1. Quantitative GC-MS analysis for aromatic components

Prior to oil fractionation, an internal standard dibenzothiophene- d_8 was added to the oils for quantification of aromatic compounds. Then, the samples were deasphalted using 40 $\times$ excess of *n*-hexane. The deasphalted samples were fractionated on a silica:alumina column using *n*-hexane, *n*-hexane:CH₂Cl₂ (2:1, v:v) and methanol:CH₂Cl₂ (9:1, v:v) to elute the saturated, aromatic and resin fractions, respectively.

Quantitative GC and GC-MS analyses for saturated fractions and Gas chromatography–isotope ratio mass spectrometry (GC-IRMS) analysis for individual *n*-alkanes were described in the previous studies (Zhang et al., 2022; Cai et al., 2023). Some parameters from the previous studies by Zhang et al. (2022) and Cai et al. (2023) were used in chemometric analysis (HCA and PCA) in the present study as notified in Table 1. In addition, we reorganized biomarker data from these two previous studies and presented six new terpane parameters of Ts/C₂₃ tricyclic terpanes, Ts/(C₂₈+C₂₉ tricyclic terpanes), C₂₉ Ts/C₂₃ tricyclic terpanes, C₂₉Ts/(C₂₈+C₂₉ tricyclic terpanes), C₃₀ diahopane/C₂₃ tricyclic terpane, C₃₀ diahopane/(C₂₈+C₂₉ tricyclic terpanes for chemometric analysis (HCA and PCA).

In the present study, aromatic fractions were analyzed using an Agilent 5975B MSD system interfaced to Agilent 6890N GC. The GC was fitted with a 60m $\times$ 0.25 mm i.d. column coated with 0.25 μ m film of HP-5MS. Helium was used as carrier gas. The oven temperature was programmed as follows: 80 $^{\circ}$ C for 1 min, raised from 80 to 310 $^{\circ}$ C at 3 $^{\circ}$ C/min, and then held at 310 $^{\circ}$ C for 16 min. Both selected ion monitoring (SIM) and full scan (*m/z* 50–550) detection modes were used. Aromatic compounds were quantified using the internal standard dibenzothiophene- d_8 .

3.2.2. Chemometric analysis

Sixteen isotopic and molecular parameters were selected for the 92 studied oils. Oil grouping on the basis of source facies was performed using hierarchical cluster analysis (HCA) and principal component analysis (PCA). For HCA, software IBM SPSS Statistics (version 26) was employed. HCA settings: preprocessing = auto-scale; distance = Squared Euclidean distance; linkage method = Ward's method; orientation = samples, and transforms = none. For PCA, Pirouette software version 4.5 (Infometrix, Inc.) was employed. PCA settings: preprocessing = autoscale, validation method = none, row = none, and transforms = none.

4. Results

4.1. Sixteen source facies parameters for oil grouping

The concentration of C₃₀ hopane and sixteen source facies parameters for oil grouping using hierarchical cluster analysis (HCA) and principal component analysis (PCA) are demonstrated in Table 1. For the 92 studied oils, two isotopic parameters, i.e., isotope reversal index (RI) and $\delta^{13}\text{C}$ of *n*-C₂₅, range from -0.09‰ to 2.94‰ and from -36.2‰ to -28.6‰ , respectively (Table 1, Fig. 3(a) and (b)). These two parameters have negative correlation with each other (Fig. 3(b)), and have no clear variation trends with decreasing concentration of C₃₀ hopane (increasing maturity) (Fig. 3(c) and (d)).

Gas chromatograms and mass chromatograms of *m/z* 191 and *m/z* 217 of oils M6070 (ON 28) and MH310 (ON 70), representatives of Group I and II oils, respectively on the basis of chemometric analysis (HCA and PCA) in this study are shown in Fig. 4. For the 92 studied oils, C₃₀ hopane concentrations are in the range of 18–3840 ppm (Table 1). The ratios of Ph/*n*-C₁₈ and β -carotane/*n*-

C₂₁ are in the ranges of 0.18–1.45 and 0.00–0.21, respectively (Table 1, Fig. 5(a)). Six terpane ratios of Ts/C₂₃ tricyclic terpanes, Ts/(C₂₈+C₂₉ tricyclic terpanes), C₂₉ Ts/C₂₃ tricyclic terpanes, C₂₉Ts/(C₂₈+C₂₉ tricyclic terpanes), C₃₀ diahopane/C₂₃ tricyclic terpane and C₃₀ diahopane/(C₂₈+C₂₉ tricyclic terpanes) are in the ranges of 0.019–0.234, 0.023–0.378, 0.002–0.305, 0.003–0.492, 0.003–0.180, 0.003–0.258, respectively (Table 1, Fig. 5(b)–(d)). All the eight saturated parameters show no clear variation trends with decreasing concentration of C₃₀ hopane (increasing maturity) (Figs. 6 and 7).

Six PAH ratios of TMNs/(TMNs + Phen), TeMNs/(TeMNs + Phen), TMNs/(TMNs + MPs), TeMNs/(TeMNs + MPs), TMNs/(TMNs + Ch) and TeMNs/(TeMNs + Ch) are in the ranges of 0.26–0.89, 0.12–0.85, 0.18–0.74, 0.07–0.69, 0.39–0.97, 0.25–0.94, respectively (Table 1, Fig. 8). These six PAH ratios also have no clear variation trends with decreasing concentration of C₃₀ hopane (increasing maturity) (Fig. 9).

4.2. Oil grouping from HCA and PCA

In the present study, the 92 studied oils from the whole Mahu sag were classified into three groups using hierarchical cluster analysis (HCA) and principal component analysis (PCA) integrating the sixteen source facies parameters. Group I consists of fifty oils, mainly located at the northeastern and central areas of the Mahu sag with only three oils (MH1601, MH702 and MH703) at the southwestern area of the Mahu sag (Table 1, Figs. 1(d) and 10). Group II consists of fourteen oils located at the southwestern area of the Mahu sag. Group III consists of twenty-eight oils located at the southwestern and central areas of the Mahu sag (Table 1, Figs. 10 and 11). Group I oils were derived from the source rocks within the Lower Permian Fengcheng Formation (P₁f) while Group II oils were derived from the source rocks within the Middle Permian Lower Wuerhe Formations (P₂w). Group III oils are mixtures of Group I and II oils as discussed later, in combination with the previous studies (Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993, 2022; Wang et al., 2013, 2023; Yu et al., 2017). Group I were further classified into Subgroup IA and IB, consisting of seventeen and thirty-three oils, respectively (Table 1, Figs. 10 and 11). The classification for Subgroup IA and IB oils reflects the subtle variation of source facies for source rocks within the Fengcheng Formation (P₁f). Subgroup IA are mainly located at the central area with only one oil (M15702) at the northeastern area of the Mahu sag (Table 1, Figs. 1(d) and 10). In the two-dimensional scores plot by PCA (Fig. 11), oil M15702 is separated from the other sixteen Subgroup IA oils but together with the Subgroup IB oils, demonstrating that source facies for this oil is possibly similar to those of Subgroup IB oils.

Principal components and percentage of variance and the loadings from the sixteen parameters for PC1 and PC2 in principal component analysis (PCA) are demonstrated in Tables 2 and 3, respectively. PC1 and PC2 occupy 66.8% and 8.6% of variance while the other fourteen principal components occupy 24.6% of variance in total (Table 2). The absolute values of the loading, i.e., the |loading| values, demonstrate the magnitudes of contributions from the selected sixteen parameters to principal components (Table 3). The higher value of the |loading| from a parameter mean the greater contribution from the parameter to the principal component (Zou et al., 2021). The |loading| values from the sixteen parameters for PC1 are in the range of 0.184–0.276, demonstrating similar contribution magnitudes from these sixteen parameters to PC1 (Table 3). In contrast, the |loading| values from the sixteen parameters for PC2 are in the range of 0.043–0.501, demonstrating that the contribution magnitudes from these sixteen parameters to PC2 vary substantially (Table 3). The three oil groups have major

Table 1Concentration of C₃₀ hopane and sixteen source facies parameters for oil grouping using HCA and PCA.

ON	Oils	St	Depth, m	GG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Northeastern																					
1	M0050	T _{1b2}	3367–3379	IB	110*	0.99*	−29.4*	0.85*	0.05*	0.080	0.078	0.041	0.040	0.035	0.034	0.64	0.54	0.48	0.38	0.80	0.74
2	M1380	T _{1b2}	3338–3353	IB	512*	0.70*	−28.9*	0.82*	0.07*	0.137	0.131	0.134	0.128	0.098	0.093	0.64	0.54	0.48	0.39	0.88	0.84
3	M1520	T _{1b2}	3142–3170	IB	280*	0.00*	−29.3*	1.17*	0.11*	0.061	0.046	0.032	0.024	0.049	0.037	0.65	0.48	0.42	0.26	0.88	0.78
4	M1540	T _{1b3}	3026–3037	IB	341*	0.61*	−28.8*	1.02*	0.07*	0.065	0.070	0.049	0.052	0.057	0.061	0.68	0.53	0.47	0.33	0.90	0.83
5	M1560	T _{1b3}	3152–3158	IB	1330*	0.88*	−30.2*	1.45*	0.21*	0.060	0.066	0.035	0.039	0.059	0.064	0.58	0.44	0.51	0.37	0.90	0.84
6	M15701	T _{1b3}	3023–3039	IB	250*	1.03*	−28.8*	0.86*	0.06*	0.061	0.063	0.041	0.042	0.037	0.038	0.66	0.51	0.43	0.29	0.90	0.83
7	M15702	T _{1b2}	3067–3072	IA	528*	0.81*	−29.1*	0.91*	0.13*	0.058	0.059	0.028	0.028	0.038	0.039	0.75	0.63	0.55	0.41	0.90	0.83
8	M150	T _{1b3}	3048–3056	IB	273*	0.44*	−28.8*	0.77*	0.05*	0.046	0.048	0.021	0.022	0.068	0.071	0.66	0.48	0.46	0.29	0.91	0.82
9	M1901	T _{1b3}	3464–3490	IB	1040*	0.69*	−28.8*	0.95*	0.09*	0.063	0.064	0.048	0.048	0.067	0.068	0.57	0.43	0.49	0.36	0.91	0.85
10	M1902	T _{1b2}	3522–3538	IB	223*	0.61*	−28.8*	0.90*	0.05*	0.085	0.097	0.062	0.070	0.035	0.040	0.77	0.68	0.58	0.47	0.87	0.81
11	M2170	P _{2w}	3999–4006	IB	399*	0.38*	−29.3*	1.20*	0.11*	0.102	0.110	0.064	0.069	0.038	0.041	0.68	0.55	0.47	0.34	0.86	0.78
12	M2180	P _{2w}	3940–3972	IB	35*	0.46*	−28.9*	1.13*	0.07*	0.095	0.101	0.043	0.046	0.031	0.033	0.68	0.59	0.41	0.32	0.85	0.80
13	M21901	T _{1b3}	3711–3714	IB	302*	0.34*	−29.4*	1.11*	0.04*	0.080	0.086	0.041	0.044	0.075	0.080	0.36	0.27	0.40	0.31	0.78	0.70
14	M21902	T _{1b}	3781–3789	IB	74*	0.40*	−28.9*	1.39*	0.10*	0.103	0.107	0.029	0.030	0.043	0.045	0.65	0.50	0.40	0.27	0.85	0.76
15	M2210	T _{1b1}	3703–3708	IB	236*	0.06*	−29.1*	1.33*	0.08*	0.083	0.084	0.035	0.036	0.028	0.029	0.71	0.62	0.52	0.42	0.86	0.80
16	M30	T _{1b}	3174–3185	IB	291*	0.06*	−28.6*	0.88*	0.07*	0.060	0.066	0.040	0.044	0.037	0.041	0.71	0.62	0.57	0.47	0.92	0.88
17	Ma200040	P _{2w}		IB	1120*	0.76*	−30.2*	1.10*	0.09*	0.074	0.068	0.056	0.052	0.069	0.065	0.61	0.38	0.48	0.26	0.90	0.78
18	Ma200100	P _{2w}		IB	400*	0.61*	−29.5*	1.32*	0.17*	0.058	0.060	0.029	0.030	0.051	0.053	0.68	0.50	0.49	0.31	0.92	0.85
19	Ma210210	P _{2w}		IB	314*	0.71*	−29.3*	0.92*	0.06*	0.066	0.065	0.042	0.041	0.045	0.044	0.58	0.37	0.37	0.20	0.88	0.76
20	MZ201	T _{1b1}	4274–4281	IB	26*	0.49*	−28.7*	0.86*	0.04*	0.091	0.099	0.027	0.029	0.035	0.038	0.73	0.68	0.53	0.48	0.86	0.83
21	MZ202	P _{2w}	4367–4370	IB	29*	0.87*	−29.1*	0.82*	0.04*	0.044	0.039	0.032	0.029	0.099	0.089	0.55	0.36	0.33	0.18	0.84	0.70
22	MZ50	T _{1b2}	4159–4164	IB	164*	−0.09*	−29.0*	0.98*	0.06*	0.062	0.047	0.013	0.010	0.097	0.073	0.50	0.40	0.33	0.24	0.76	0.68
Central																					
23	M180	T _{2k2}	3332–3334	IB	321*	1.08*	−29.7*	0.62*	0.06*	0.116	0.100	0.077	0.067	0.057	0.049	0.46	0.27	0.29	0.15	0.83	0.67
24	M6020	T _{1b2}	3846–3868	IB	295*	0.96*	−29.9*	0.89*	0.07*	0.029	0.031	0.012	0.013	0.066	0.071	0.69	0.56	0.55	0.41	0.88	0.80
25	M6040	T _{1b2}	3870–3890	IB	478*	0.71*	−29.9*	1.01*	0.05*	0.062	0.072	0.037	0.042	0.056	0.065	0.65	0.54	0.55	0.45	0.87	0.81
26	M60601	J _{1b}	2502–2510	IB	372*	0.92*	−30.1*	0.85*	0.04*	0.077	0.077	0.103	0.102	0.105	0.104	0.51	0.32	0.37	0.21	0.89	0.78
27	M60602	T _{1b2}	3745–3781	IB	270*	1.03*	−30.1*	0.70*	0.04*	0.069	0.076	0.027	0.030	0.038	0.042	0.69	0.58	0.54	0.42	0.87	0.80
28	M6070	P _{2w}	4100–4108	IB	1120*	0.64*	−29.7*	1.21*	0.09*	0.092	0.091	0.063	0.062	0.048	0.047	0.53	0.43	0.48	0.37	0.85	0.79
29	M6070d	P _{2w}	4100–4108	IB	940*	0.69*	−29.6*	1.13*	0.15*	0.085	0.084	0.058	0.057	0.048	0.048	0.51	0.44	0.46	0.38	0.84	0.79
30	MZ40	T _{3b}	3565–3617	IB	173*	1.16*	−30.2*	0.62*	0.04*	0.136	0.140	0.121	0.125	0.070	0.073	0.59	0.45	0.47	0.34	0.90	0.84
31	MH0110	P _{2w}	2676–2688	IB	1390*	0.78*	−30.3*	1.23*	0.17*	0.046	0.063	0.031	0.043	0.049	0.066	0.73	0.49	0.45	0.23	0.93	0.82
32	AH0130	T _{1b2}	3748–3769	IA	390*	0.73*	−29.9*	0.95*	0.10*	0.032	0.037	0.006	0.007	0.009	0.010	0.76	0.68	0.54	0.44	0.90	0.85
33	AH0150	T _{1b}	3399–3425	IA	250*	0.82*	−30.5*	0.86*	0.08*	0.030	0.035	0.002	0.003	0.012	0.014	0.73	0.69	0.55	0.50	0.84	0.81
34	AH110	T _{1b2}	2876–2890	IA	203*	1.13*	−30.4*	0.78*	0.10*	0.027	0.037	0.003	0.005	0.015	0.021	0.74	0.71	0.56	0.51	0.81	0.79
35	AH130	T _{1b2}	3138–3156	IA	209*	0.89*	−30.0*	0.86*	0.07*	0.021	0.027	0.005	0.007	0.007	0.010	0.68	0.62	0.52	0.45	0.83	0.79
36	AH130d	T _{1b2}	3138–3156	IA	241*	1.17*	−30.7*	0.89*	0.09*	0.029	0.035	0.004	0.005	0.009	0.010	0.69	0.63	0.52	0.45	0.85	0.81
37	AH140	T _{1b1}	3568–3578	IA	194*	0.98*	−29.7*	0.80*	0.07*	0.028	0.031	0.007	0.007	0.024	0.027	0.54	0.47	0.51	0.44	0.80	0.75
38	AH2010	T _{1b1}	3507–3527	IA	245*	0.81*	−30.1*	0.88*	0.09*	0.029	0.035	0.004	0.005	0.026	0.031	0.68	0.61	0.53	0.45	0.85	0.80
39	AH301	T _{3b3}	2812–2816	IA	239*	0.76*	−29.6*	0.95*	0.19*	0.029	0.034	0.008	0.009	0.010	0.012	0.89	0.85	0.69	0.62	0.92	0.89
40	AH302	T _{1b2}	3713–3720	IA	222*	0.81*	−29.7*	0.89*	0.19*	0.029	0.034	0.007	0.009	0.016	0.019	0.84	0.81	0.74	0.69	0.92	0.90
41	AH303	P _{2w}	3898–3968	IA	197*	0.68*	−29.6*	0.95*	0.16*	0.031	0.037	0.007	0.009	0.018	0.022	0.86	0.81	0.59	0.51	0.89	0.85
42	AH40	T _{1b3}	2854–2886	IA	463*	0.92*	−30.4*	0.95*	0.21*	0.023	0.028	0.005	0.006	0.017	0.021	0.69	0.60	0.50	0.40	0.84	0.78
43	AH60	T _{1b2}	3878–3909	IA	285*	0.87*	−30.1*	0.92*	0.16*	0.027	0.033	0.005	0.006	0.005	0.006	0.64	0.54	0.54	0.43	0.86	0.80
44	AH80	T _{1b2}	3791–3823	IA	239*	1.00*	−30.5*	0.72*	0.13*	0.019	0.023	0.006	0.007	0.021	0.025	0.65	0.55	0.47	0.37	0.72	0.64
45	AH0140	T _{1b1}	3117–3136	IA	182*	1.08*	−29.8*	0.64*	0.04*	0.033	0.042	0.005	0.006	0.004	0.005	0.67	0.58	0.51	0.41	0.84	0.77
46	AH1001	T _{2k2}	2954–2971	III	4008*	1.29*	−30.0*	0.75*	0.15*	0.048	0.055	0.023	0.026	0.016	0.018	0.61	0.43	0.46	0.29	0.94	0.88
47	AH1002	T _{1b2}	3342–3407	IA	248*	0.96*	−29.7*	0.75*	0.16*	0.027	0.031	0.009	0.010	0.003	0.003	0.54	0.46	0.52	0.44	0.83	0.78
48	AH1201	J _{1b}	2436–2440	III	387*	1.43*	−30.6*	0.72*	0.06*	0.055	0.063	0.015	0.018	0.023	0.026	0.55	0.35	0.41	0.24	0.92	0.84
49	AH1202	J _{1b}	2478–2482	III	443*	1.34*	−30.0*	0.78*	0.15*	0.034	0.038	0.014	0.015	0.012	0.013	0.52	0.36	0.45	0.30	0.88	0.79
50	AH1203	T _{2k2}	3156–3280	III	273*	1.30*	−30.2*	0.58*	0.05*	0.068	0.070	0.027	0.028	0.027	0.028	0.66	0.49	0.48	0.32	0.97	0.94
51	AH1204	T _{1b1}	3669–3708	IA	254*	1.49*	−30.7*	0.81*	0.11*	0.040	0.048	0.009	0.010	0.015	0.018	0.66	0.59	0.55	0.48	0.83	0.79
52	AH150	J _{1b1}	2379–2383	III	458*	1.75*	−31.4*	0.77*	0.11*	0.039	0.042	0.012	0.014	0.016	0.017	0.55	0.38	0.48	0.31	0.89	0.80
53	AH2040	T _{3b}	2392–2400	III	118*	1.44*	−29.4*	0.72*	0.03*	0.104	0.128	0.021	0.026	0.010	0.012	0.53	0.35	0.46	0.30	0.96	0.92

(continued on next page)

Table 1 (continued)

ON	Oils	St	Depth, m	GG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
54	AH50	T ₂ k ₂	3358–3361	III	189*	2.08*	−30.1*	0.54*	0.05*	0.062	0.061	0.026	0.026	0.005	0.005	0.61	0.41	0.46	0.27	0.94	0.87
55	AH50d	T ₂ k ₂	3358–3361	III	276*	1.23*	−29.9*	0.51*	0.11*	0.068	0.075	0.028	0.031	0.012	0.014	0.65	0.46	0.54	0.35	0.94	0.88
Southwestern																					
56	MH160	P ₁ f	4424–4437	IB	1880*	0.56*	−29.3*	1.25*	0.20*	0.036	0.053	0.032	0.047	0.056	0.082	0.70	0.46	0.39	0.19	0.94	0.84
57	MH0120	T ₁ b ₁	3453–3458	II	605*	2.53*	−34.6*	0.39*	0.04*	0.139	0.182	0.110	0.144	0.081	0.106	0.40	0.27	0.31	0.20	0.69	0.55
58	MH01301	T ₁ b	3521–3541	II	436*	2.67*	−34.8*	0.34*	0.02*	0.156	0.193	0.120	0.148	0.089	0.110	0.37	0.24	0.26	0.16	0.68	0.53
59	MH01302	P ₃ w	3642–3648	III	86*	2.33*	−32.8*	0.49*	0.03*	0.116	0.138	0.069	0.082	0.042	0.050	0.56	0.43	0.41	0.29	0.74	0.62
60	MH01302d	P ₃ w	3642–3648	III	77*	2.18*	−32.8*	0.48*	0.03*	0.133	0.166	0.069	0.085	0.048	0.060	0.57	0.43	0.44	0.30	0.74	0.61
61	MH01303	P ₂ w	3767–3782	II	661*	2.78*	−35.6*	0.29*	0.03*	0.155	0.203	0.136	0.178	0.097	0.127	0.34	0.18	0.23	0.11	0.70	0.49
62	MH0140	P ₃ w	3737–3844	II	210*	2.63*	−35.1*	0.27*	0.02*	0.149	0.179	0.080	0.097	0.123	0.148	0.32	0.14	0.20	0.08	0.74	0.50
63	MH0180	P ₃ w ₁	3591–3599	II	411*	2.87*	−36.1*	0.25*	0.01*	0.190	0.282	0.174	0.259	0.124	0.185	0.36	0.24	0.25	0.16	0.73	0.60
64	MH100	P ₂ w	3744–3748	II	979*	2.52*	−35.5*	0.42*	0.03*	0.234	0.378	0.305	0.492	0.141	0.228	0.42	0.19	0.29	0.12	0.83	0.62
65	MH110	P ₃ w	3422–3461	III	1030*	2.72*	−34.7*	0.51*	0.05*	0.071	0.108	0.090	0.137	0.071	0.109	0.48	0.24	0.30	0.13	0.86	0.67
66	MH140	P ₂ w	4163–4179	II	58*	2.47*	−35.5*	0.22*	0.00*	0.200	0.288	0.100	0.144	0.180	0.258	0.31	0.12	0.19	0.07	0.71	0.43
67	MH150	P ₃ w	3146–3154	III	875*	2.72*	−34.6*	0.59*	0.04*	0.112	0.173	0.105	0.162	0.083	0.129	0.47	0.27	0.30	0.15	0.82	0.66
68	MH210	T ₁ b ₂	3395–3398	II	435*	2.83*	−34.5*	0.43*	0.02*	0.201	0.268	0.140	0.187	0.100	0.133	0.49	0.46	0.40	0.38	0.68	0.65
69	MH220	P ₃ w ₁	3431–3465	II	264*	2.94*	−36.0*	0.18*	0.01*	0.222	0.186	0.197	0.166	0.153	0.128	0.26	0.12	0.18	0.08	0.65	0.41
70	MH310	P ₃ w ₂	3185–3190	II	1070*	2.66*	−35.7*	0.34*	0.02*	0.166	0.237	0.168	0.241	0.106	0.152	0.36	0.20	0.24	0.12	0.75	0.57
71	MH320	P ₃ w ₂	3184–3188	III	1190*	2.52*	−34.4*	0.64*	0.05*	0.068	0.096	0.089	0.126	0.070	0.098	0.53	0.32	0.35	0.18	0.87	0.74
72	MH0190	P ₃ w	3595–3611	III	1060*	1.60*	−31.4*	0.98*	0.15*	0.046	0.062	0.041	0.055	0.047	0.063	0.68	0.45	0.43	0.22	0.94	0.85
73	MH0300	P ₃ w ₂	3435–3506	III	466*	2.17*	−32.5*	0.59*	0.07*	0.056	0.072	0.063	0.081	0.040	0.052	0.532	0.377	0.406	0.266	0.849	0.750
74	MH130	J ₁ b	3013–3018	III	385*	1.31*	−31.7*	0.53*	0.03*	0.191	0.202	0.209	0.221	0.119	0.126	0.536	0.311	0.361	0.181	0.899	0.777
75	MH1801	P ₃ w ₂	3642–3682	III	212*	1.97*	−32.3*	0.52*	0.03*	0.074	0.086	0.070	0.082	0.040	0.047	0.562	0.435	0.553	0.425	0.956	0.929
76	MH1802	P ₃ w ₁	3741–3755	II	462*	2.81*	−35.9*	0.27*	0.01*	0.221	0.314	0.181	0.258	0.130	0.186	0.362	0.223	0.243	0.139	0.776	0.636
77	MH190	T ₁ b	3722–3726	III	18*	1.35*	−30.7*	0.30*	0.01*	0.066	0.085	0.023	0.029	0.033	0.042	0.627	0.571	0.498	0.440	0.387	0.333
78	MH101	P ₃ w	3420–3437	III	966*	1.56*	−31.8*	0.84*	0.12*	0.039	0.060	0.023	0.036	0.044	0.068	0.644	0.385	0.402	0.189	0.915	0.789
79	MH102	P ₂ w	3516–3531	III	1080*	1.49*	−31.5*	0.90*	0.16*	0.025	0.038	0.049	0.075	0.040	0.060	0.621	0.364	0.372	0.171	0.904	0.767
80	MH230	P ₃ w ₂	4193–4227	III	138*	2.04*	−32.0*	0.38*	0.02*	0.053	0.060	0.039	0.044	0.037	0.042	0.622	0.500	0.469	0.349	0.786	0.690
81	MH2401	T ₁ b	3055–3100	III	840*	2.13*	−33.4*	0.71*	0.05*	0.080	0.129	0.052	0.084	0.054	0.088	0.352	0.287	0.319	0.257	0.732	0.669
82	MH2402	P ₃ w ₂	3182–3205	II	751*	2.83*	−36.2*	0.26*	0.01*	0.234	0.336	0.185	0.265	0.143	0.205	0.302	0.166	0.182	0.093	0.701	0.519
83	MH270	P ₃ w ₂	3143–3171	III	810*	1.51*	−31.6*	0.76*	0.08*	0.040	0.047	0.055	0.065	0.028	0.033	0.567	0.380	0.391	0.231	0.862	0.745
84	MH290	P ₃ w ₂	3242–3245	III	1250*	2.24*	−32.5*	0.95*	0.10*	0.043	0.060	0.050	0.068	0.048	0.066	0.598	0.397	0.378	0.212	0.896	0.793
85	MH20	T ₃ b	2581–2612	III	1110*	2.36*	−33.6*	0.57*	0.09*	0.064	0.090	0.059	0.083	0.050	0.070	0.606	0.425	0.436	0.271	0.910	0.830
86	MH40	T ₁ b	3297–3335	III	3840*	2.32*	−33.4*	0.82*	0.18*	0.054	0.072	0.089	0.118	0.036	0.048	0.332	0.317	0.381	0.365	0.698	0.683
87	MH701	P ₃ w	3352–3442	III	1460*	1.45*	−31.3*	1.03*	0.11*	0.036	0.046	0.062	0.078	0.029	0.037	0.621	0.395	0.423	0.226	0.889	0.762
88	MH702	P ₁ f ₃	4012–4030	IB	929*	0.10*	−28.8*	1.24*	0.13*	0.026	0.036	0.012	0.016	0.030	0.041	0.738	0.473	0.475	0.224	0.928	0.805
89	MH702d	P ₁ f ₃	4012–4030	IB	861*	0.28*	−28.9*	1.33*	0.11*	0.024	0.034	0.027	0.038	0.031	0.044	0.720	0.478	0.435	0.215	0.920	0.805
90	MH801	P ₃ w ₁	3281–3286	III	565*	2.31*	−33.3*	0.47*	0.07*	0.067	0.075	0.063	0.071	0.057	0.064	0.493	0.253	0.259	0.108	0.864	0.689
91	MH802	P ₃ w	3342–3350	II	1490*	1.19*	−31.2*	0.91*	0.07*	0.220	0.261	0.222	0.263	0.102	0.121	0.292	0.128	0.346	0.160	0.478	0.247
92	MH803	P ₁ f	3442–3488	III	942*	1.01*	−30.5*	0.91*	0.12*	0.049	0.060	0.043	0.053	0.044	0.055	0.525	0.252	0.310	0.120	0.840	0.615

ON: oil number; St: strata as shown in Fig. 2; GG: genetic groups (IA: Subgroup IA; IB: Subgroup IB; II: Group II; III: Group III); 1: concentration of C₃₀ hopane in ppm; 2: Isotope reversal index (RI) in ‰; 3: $\delta^{13}\text{C}$ of *n*-C₂₅ in ‰; 4: Ph/*n*-C₁₈; 5: β -carotane/*n*-C₂₁; 6: Ts/C₂₃ tricyclic terpanes; 7: Ts/(C₂₈+C₂₉ tricyclic terpanes); 8: C₂₉ Ts/C₂₃ tricyclic terpanes; 9: C₂₉Ts/(C₂₈+C₂₉ tricyclic terpanes); 10: C₃₀ dihopane/C₂₃ tricyclic terpane; 11: C₃₀ dihopane/(C₂₈+C₂₉ tricyclic terpanes); 12: TMNs/(TMNs + Phen); 13: TeMNs/(TeMNs + Phen); 14: TMNs/(TMNs + MPs); 15: TeMNs/(TeMNs + MPs); 16: TMNs/(TMNs + Ch); 17: TeMNs/(TeMNs + Ch). *: data from Zhang et al. (2022); Cai et al. (2023).

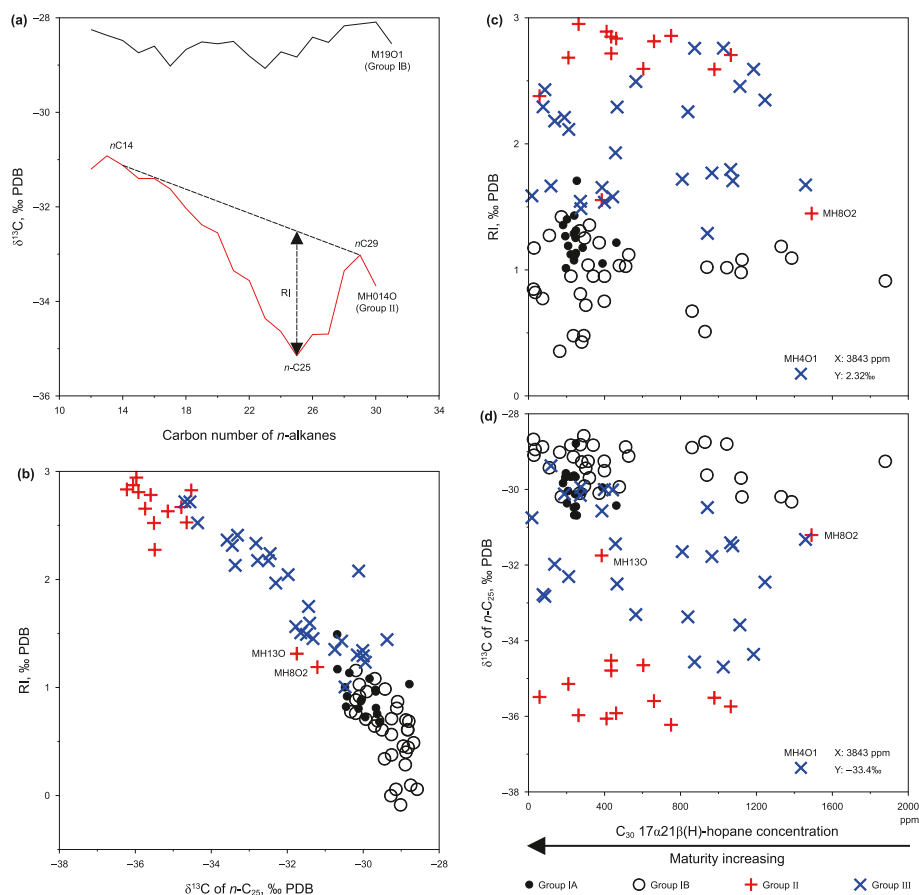


Fig. 3. Definition of reversal index (RI) (a), and crossplots of $\delta^{13}\text{C}$ values of *n*-C₂₅ versus parameter RI (b), concentration of C₃₀ hopane versus carbon isotopic parameter RI (c), and $\delta^{13}\text{C}$ value of *n*-C₂₅ (d) for the studied 92 oils.

differences in PC1 scores but no clear differences in PC2 scores (Fig. 11). PC1 scores demonstrate the source facies variations among the 92 studied oils. The geochemical interpretations for PC2 scores are unclear for us. In the two-dimensional scores plot by PCA (Fig. 11), Group I and II oils can be completely separated while Group III oils overlap slightly with Group II oils but significantly with Group I oils.

5. Discussion

5.1. Effectiveness of carbon isotopes of individual *n*-alkanes as source facies parameters

Compound specific isotopic analysis has been well established for nearly four decades (e.g., Hayes et al., 1987, 1990) and is a valuable method in oil-oil and oil-source rock correlation (Peters et al., 2005). Previous studies demonstrated that $\delta^{13}\text{C}$ values of individual *n*-alkanes and their variation trends with increasing carbon number vary significantly among oil groups from different source rocks in the Junggar Basin (Chen et al., 2016c; Yu et al., 2017; Pan et al., 2021; Zhang et al., 2022). Zhang et al. (2022) defined an isotopic parameter of reversal index (RI, Fig. 3(a)), and classified 52 oils from the central and southwestern Mahu sag into three groups on the basis of RI and $\delta^{13}\text{C}$ values of *n*-C₂₅, i.e., Group I, II and III that are derived from source rocks within the Lower Permian Fengcheng Formation (P₁f), Middle Permian Lower Wuerhe Formation (P₂w) and both, respectively. In the present study, the 92 oils from the whole Mahu sag were also classified

into three groups by HCA and PCA on the basis of sixteen parameters including RI and $\delta^{13}\text{C}$ values of *n*-C₂₅ (Table 1, Figs. 10 and 11). On the basis of new classification, RI values generally increase in the sequence of Group I < Group III < Group II oils while the $\delta^{13}\text{C}$ values of *n*-C₂₅ generally decrease in the sequence of Group I > Group III > Group II oils (Fig. 3b). However, there are some overlaps for these two parameters among the three oil groups. In particular, oils MH130 and MH802 of Group II have extraordinary lower RI values and higher $\delta^{13}\text{C}$ values of *n*-C₂₅ compared with the other Group II oils (Fig. 3(b)).

Previous studies demonstrated that oils in the Mahu sag have maturities from early to late oil generative window and over mature (Yu et al., 2017; Tao et al., 2021; Zhang et al., 2022; Cai et al., 2023). The studied oils have concentrations of biomarkers varying over three orders but yet similar low-maturity biomarker ratios, demonstrating oil mixing to various extents (Table 1, Figs. 3, 6, 7 and 9; Zhang et al., 2022; Cai et al., 2023). Biomarkers in these oils were predominantly derived from initially charging lower maturity oils. The initial oils were differentially diluted by later charging higher-maturity oils poor in biomarkers (Zhang et al., 2022; Cai et al., 2023). Previous studies demonstrated that concentrations of biomarker compounds (terpanes and steranes) decrease with maturity (e.g., Mackenzie et al., 1985; Farrimond et al., 1998; Wilhelms and Larter, 2004). Dahl et al. (1999) used the concentration of C₂₉ $\alpha\alpha\alpha$ 20R sterane to assess oil maturities at oil generation window prior to oil cracking. In the present study, concentration of C₃₀ hopane is used to indicate a bulk or averaged maturity for the studied oils (Table 1, Figs. 3, 6, 7 and 9). RI and $\delta^{13}\text{C}$

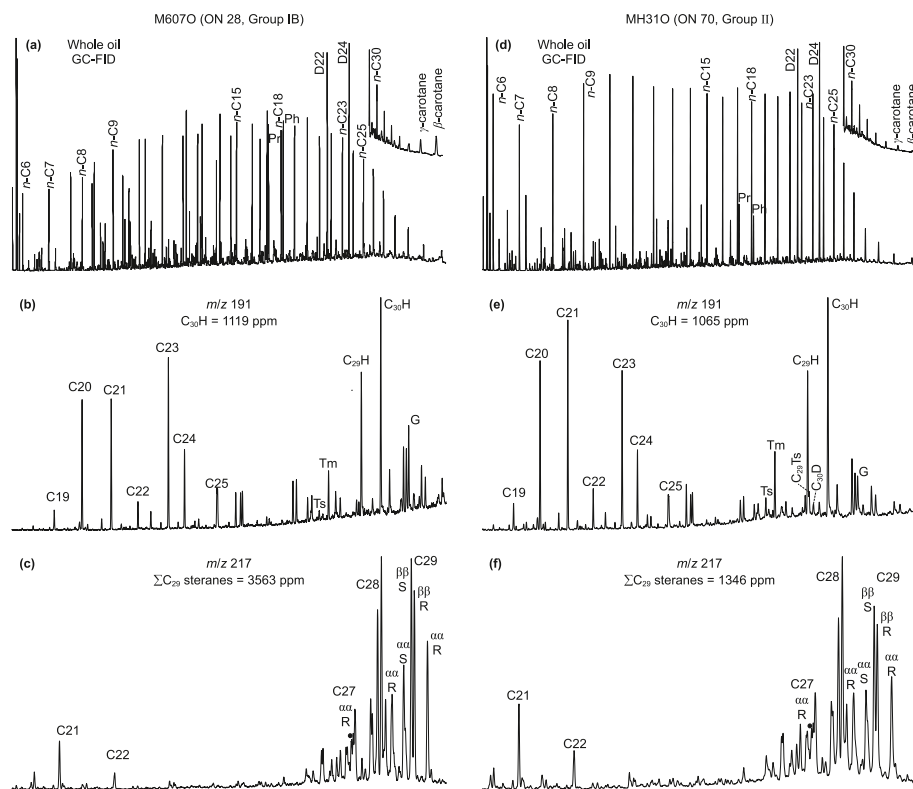


Fig. 4. Gas chromatograms and m/z 191 and m/z 217 mass chromatograms of selected oils M6070 and MH310 representatives of Group IB and II oils, respectively. in (a) and (d), D22: deuterated n -C₂₂; D24: deuterated n -C₂₄; in (b) and (e), C₃₀H: C₃₀ dihopane; C₁₉–C₂₅: C₁₉ to C₂₅ tricyclic terpanes; G: gammacerane.

values of n -C₂₅ have no clear variation trends with decreasing concentrations of C₃₀ hopane (increasing maturity), demonstrating that these two parameters are not influenced significantly by maturity (Fig. 3(c) and (d)).

5.2. Effectiveness of ratios of n -alkanes, isoprenoids and carotanes as source facies parameters

Previous studies demonstrated that ratios of Pr/ n -C₁₇, Ph/ n -C₁₈ and Pr/Ph are influenced by source facies, i.e., biological precursors of organic matter and depositional environments of source rocks (e.g., Connan, 1981; Palacas, 1984; Peters et al., 2005). These three ratios are also affected by maturity (e.g., Alexander et al., 1981; ten Haven et al., 1987). Several studies documented that higher relative concentration of β -carotane is associated primarily with anoxic, saline lacustrine, or highly restricted marine settings (e.g., Hall and Douglas, 1983; Jiang and Fowler, 1986; Fu et al., 1990; Irwin and Meyer, 1990; Peters et al., 2005). Ratios of Pr/ n -C₁₇, Ph/ n -C₁₈ and Pr/Ph and relative concentration of β -carotane were generally used for oil source correlation and classification of oils in the Junggar Basin (e.g., Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993; Pan and Yang, 2000; Pan et al., 2003, 2021; Cao et al., 2005, 2006, 2015, 2020; Wang et al., 2013; Chen et al., 2016a, 2016b; Tao et al., 2021).

For the three oil groups classified by HCA and PCA, Group I oils have higher relative concentrations of Pr, Ph and carotanes, and accordingly, lower relative concentrations of n -alkanes while Group II oils are opposite (Fig. 4(a) and (d)). The influence of source facies on ratios of Pr/ n -C₁₇, Ph/ n -C₁₈ and β -carotane/ n -C₂₁ is substantially greater compared with the influence of maturity for the

studied oils (Zhang et al., 2022). In the present study, two ratios of Ph/ n -C₁₈ and β -carotane/ n -C₂₁ were selected and integrated to the sixteen parameters for oil grouping using HCA and PCA. These two ratios decrease in the sequence of group I oils > group III oils > group II oils (Fig. 5(a)).

Group I and II oils can be sufficiently distinguished in the crossplots of concentration of C₃₀ hopane versus Ph/ n -C₁₈ and β -carotane/ n -C₂₁ ratios while Group III oils are reasonably separated from the other two groups in the crossplot of concentration of C₃₀ hopane versus Ph/ n -C₁₈ ratio but overlap heavily with Group I oils in the crossplot of concentration of C₃₀ hopane versus β -carotane/ n -C₂₁ ratio (Fig. 6). Ratios of Ph/ n -C₁₈ and β -carotane/ n -C₂₁ only have a very slightly decreasing trend with decreasing concentration of C₃₀ hopane (increasing maturity) (Fig. 6), reflecting that these two ratios are mainly influenced by source facies but only slightly influenced by maturity. At similar concentrations of C₃₀ hopane, ratios of Ph/ n -C₁₈ and β -carotane/ n -C₂₁ vary substantially greater for Group I oils than Group II oils (Fig. 6). This phenomenon can be mainly ascribed to deposition environments of source rocks for these two oil groups. Source rocks of the Fengcheng Formation (P₁f) were deposited in alkaline lacustrine environments with high salinity, strongly reducing conditions, and hydrothermal activity (Cao et al., 2015, 2020). Water geochemical condition (salinities, pH and Eh values and temperatures) and biological precursors for organic matter in sediments varied substantially at different stages of alkaline lacustrine evolution, leading to that ratios of Ph/ n -C₁₈ and β -carotane/ n -C₂₁ vary greatly for group I oils (Cao et al., 2015, 2020; Wang et al., 2022). In contrast, these two ratios vary in narrow ranges for Group II oils because their source rocks within the Lower Wuerhe Formation

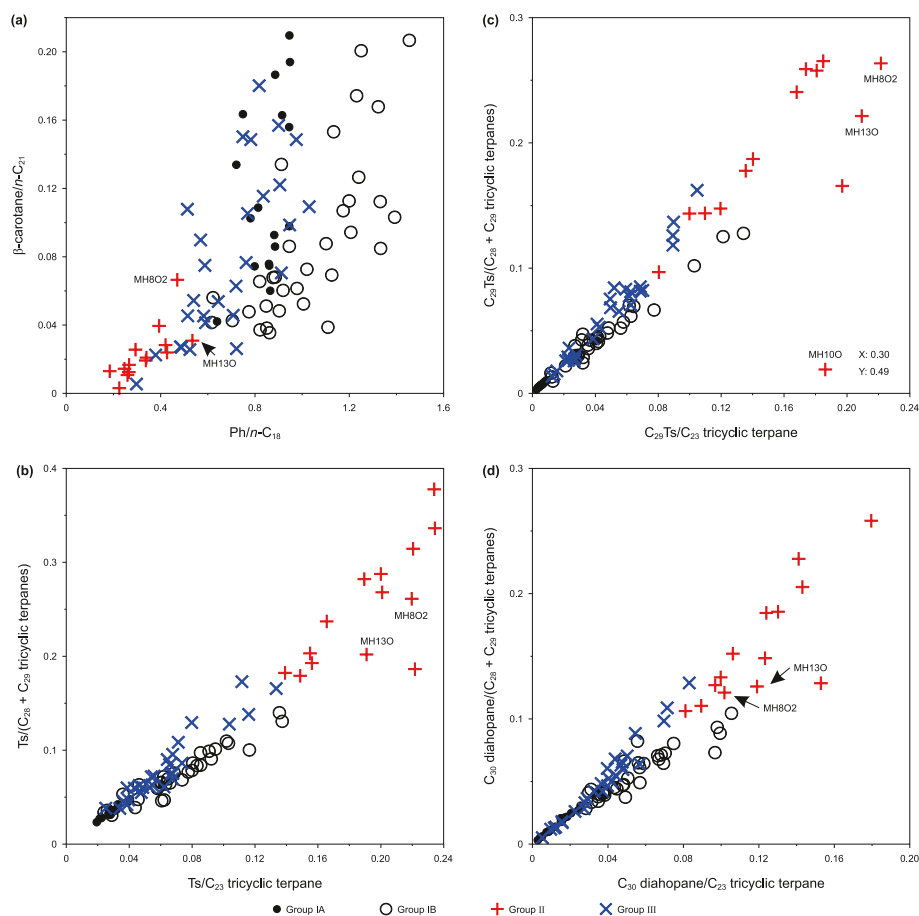


Fig. 5. Crossplots of Ph/*n*-C₁₈ ratio versus β-carotane/*n*-C₂₁ ratio (a), ratios of Ts/C₂₃ tricyclic terpane versus Ts/(C₂₈+C₂₉ tricyclic terpanes) (b), ratios of C₂₉Ts/C₂₃ tricyclic terpane versus C₂₉Ts/(C₂₈+C₂₉ tricyclic terpanes) (c) and ratios of C₃₀ diahopane/C₂₃ tricyclic terpane versus C₃₀ diahopane/(C₂₈+C₂₉ tricyclic terpanes) (d).

(P₂w) were deposited in a fresh lacustrine environment with relative stable water geochemical condition (e.g., Zhang et al., 1993; Wang et al., 2013).

5.3. Effectiveness of terpane parameters as source facies parameters

Previous studies suggested that oils and extracts from saline lacustrine and marine carbonate source rocks contain higher relative concentrations of tricyclic terpanes (e.g., Kruege et al., 1990; De Grande et al., 1993). Numerous studies demonstrated that the relative concentrations of Ts, C₂₉Ts and C₃₀ diahopane are influenced by lithology and deposition environment, and higher in oils and extracts from shale source rocks compared with carbonate source rocks (Seifert and Moldowan, 1978, 1986; Rullkötter et al., 1985; Moldowan et al., 1986, 1991; Farrimond and Telnæs, 1996). Previous studies on oils from the Junggar Basin demonstrated that Group I oils derived from the Fengcheng Formation (P₁f) have lower relative concentrations of Ts, C₂₉Ts and C₃₀ diahopane than Group II oils from other Permian source rocks (Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993; Wang et al., 2013, 2023; Yu et al., 2017). Zhang et al. (2022) further demonstrated that Group I oils have higher concentrations of tricyclic terpanes (e.g., C₂₀, C₂₁ and C₂₃ tricyclic terpanes) but lower concentrations of Ts, C₂₉Ts and C₃₀ diahopane than Group II oils at similar concentrations of ΣC₂₉ steranes and C₃₀ hopane. In the present study, six ratios of Ts/C₂₃ tricyclic terpane, Ts/(C₂₈ + C₂₉ tricyclic terpanes), C₂₉Ts/C₂₃ tricyclic terpane, C₂₉Ts/(C₂₈ + C₂₉ tricyclic

terpanes), C₃₀ diahopane/C₂₃ tricyclic terpane and C₃₀ diahopane/(C₂₈ + C₂₉ tricyclic terpanes) were selected for oil grouping for the studied 92 oils using HCA and PCA (Table 1). These six ratios are well positively correlated between each other (Fig. 5(b)–(d)). For the three oil groups classified using HCA, Group I oils have lower values of these six ratios than Group II oils (Figs. 5(b)–(d) and 7). Group I and II oils can be generally separated using these six ratios (Figs. 5(b)–(d) and 7). In particular, these two oil groups can be completely distinguished using the ratio of Ts/(C₂₈ + C₂₉ tricyclic terpanes) (Figs. 5(b) and 7(b)). In contrast, Group III oils heavily overlap with Group I oils on the basis of these six ratios (Figs. 5(b)–(d) and 7).

These six ratios do not show clear variation trends with decreasing concentration of C₃₀ hopane (increasing maturity), demonstrating that Ts, C₂₉Ts and C₃₀ diahopane have relatively similar thermal stabilities with C₂₃, C₂₈ and C₂₉ tricyclic terpanes (Fig. 7). A recent study by Bian et al. (2024) demonstrated that Ts, C₂₉Ts and C₃₀ diahopane have similar thermal stabilities of C₂₁–C₂₉ tricyclic terpanes on the basis of chemometric analysis on concentrations and ratios of terpanes and steranes for oils from the cratonic region of the Tarim Basin. Holba et al. (2001) used extended tricyclic terpane ratio ETR = (C₂₈ + C₂₉ tricyclic terpanes)/Ts to distinguish oils derived from source rocks within the Triassic–Jurassic strata. Holba et al. (2003) further demonstrated that oils from source rocks deposited in an oceanic upwelling zone may have very high ETR values. Peters et al. (2007) used the ratio of C₂₆ tricyclic terpanes/Ts to identify oil sources from Triassic, Jurassic and younger source rocks. Huang et al. (2017) performed

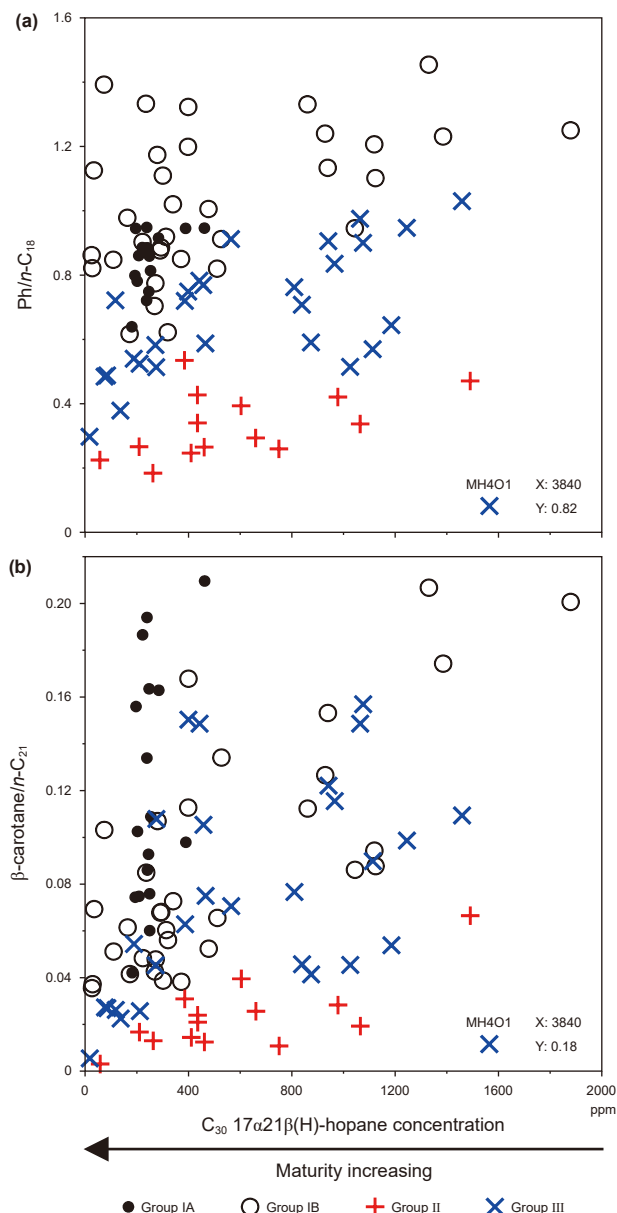


Fig. 6. Crossplots of concentration of C_{30} hopane versus $Ph/n-C_{18}$ ratio and β -carotane/ $n-C_{21}$ ratio.

an extensive study on ETR for oils and condensates from the Tazhong and Tabei uplifts, and concluded that these oils and condensates are derived from multiple source rocks within the Cambrian–Ordovician strata deposited under different environmental conditions. The two ratios of $Ts/(C_{28} + C_{29}$ tricyclic terpanes) and $C_{29}Ts/(C_{28} + C_{29}$ tricyclic terpanes) can be considered as modified ETR ratios.

5.4. Effectiveness of ratios of aromatic components as source facies parameters

Parameters of aromatic components were mainly used as maturity parameters, such as DNR, TMNr, TeMnr, MPI1 and MDR from dimethylnaphthalenes (DMNs), trimethylnaphthalenes (TMNs), tetramethylnaphthalenes (TeMNs), phenanthrene (P), methylphenanthrenes (MPs) and methyl dibenzothiophenes (MDBTs), respectively (Radke and Welte, 1983; Alexander et al., 1985; Radke et al., 1986; van Aarssen et al., 1999; Wang et al.,

2022). Only a limited previous studies documented some aromatic parameters for source facies (e.g., Hughes et al., 1995; Radke et al., 2000). Hughes et al. (1995) proposed ratio of DBT/Phen (dibenzothiophene/phenanthrene) while Radke et al. (2000) documented ratio of ADBT/ADBF ((DBT + MDBTs (methyl-dibenzothiophenes))/(DBF (dibenzofuran) + MDBFs (methyl-dibenzofurans))) as facies parameters. In the present study, the concentrations of DBT, ADBT, DBF and ADBF and ratios of DBT/Phen and ADBT/ADBF have no clear differences among Group I, II and III oils. Thus, these concentrations and ratios were not used for oil grouping in the present study.

Requejo et al. (1996) demonstrated that the degree of alkylation of polynuclear aromatic hydrocarbons (PAH) decreases in the sequence of carbonate-sourced oils > siliciclastic-sourced oils > paralic-sourced oils, i.e., carbonated-sourced oils contain higher amounts of PAH with three and four methyl groups while paralic-sourced oils contain higher amounts of PAH with one and two methyl groups, and suggested that this sequence can be ascribed to PAH precursor moieties in kerogens and maturities. In the present study, the concentrations of most PAH have no clear differences and heavily overlap among the three oil groups classified using HCA and PCA. However, the Group I oils have relatively higher concentrations of trimethylnaphthalenes (TMNs) and tetramethylnaphthalenes (TeMNs) but lower concentrations of phenanthrene (P), methylphenanthrenes (MPs) and chrysene (Ch) compared with Group II oils. Thus, six ratios of $TMNs/(TMNs + P)$, $TMNs/(TMNs + MPs)$, $TMNs/(TMNs + Ch)$, $TeMNs/(TeMNs + P)$, $TeMNs/(TeMNs + MPs)$ and $TeMNs/(TeMNs + Ch)$ were selected and integrated to the sixteen source facies parameters for oil grouping using HCA and PCA. These six ratios generally decrease in the sequence Group I > Group III > Group II for the three oil groups classified using HCA and PCA and are positively correlated to each other (Table 1, Fig. 8). They do not show any clear variation trends with decreasing concentration of C_{30} hopane (increasing maturity) (Fig. 9), suggesting that they are mainly influenced by source facies and not affected significantly by maturity. Aromatic components are thermally stable and do not decompose within the whole oil-generative window and even post mature (Hughes et al., 1995; Requejo et al., 1996). In addition, aromatic components are not influenced by biodegradation up to high level (Peters and Moldowan, 1993). Thus, the six aromatic facies ratios are applicable for oils at maturities of whole oil generative window and even post mature and for both non biodegraded and biodegraded oils up to higher levels (Peters and Moldowan, 1993; Hughes et al., 1995; Requejo et al., 1996).

5.5. Correlation among the sixteen source facies parameters

Pearson correlation coefficients (r) between each other for the sixteen source facies parameters are shown in Table 4. Among the 120 pairs in total, one pair has very strong negative linear correlation with r value of -0.94 , twenty pairs have very strong positive linear correlations with r values in the range of 0.80 – 0.97 , thirty-seven pairs have strong negative linear correlations with r values ranging from -0.79 to -0.60 , fifteen pairs have strong positive linear correlations with r values in the range of 0.60 – 0.77 , twenty-five pairs have moderate negative linear correlations with r values ranging from -0.58 to -0.40 , eighteen pairs have moderate positive linear correlations with r values in the range of 0.41 – 0.58 , and four pairs have weak positive linear correlations with r values in the range of 0.21 – 0.37 (Zou et al., 2021).

Isotopic compositions of n -alkanes inherit from those of bioprecursors (e.g., algae and bacteria species). Components for the eight saturated ratios are biomarkers. Concentrations of these saturated components are closely related to species of

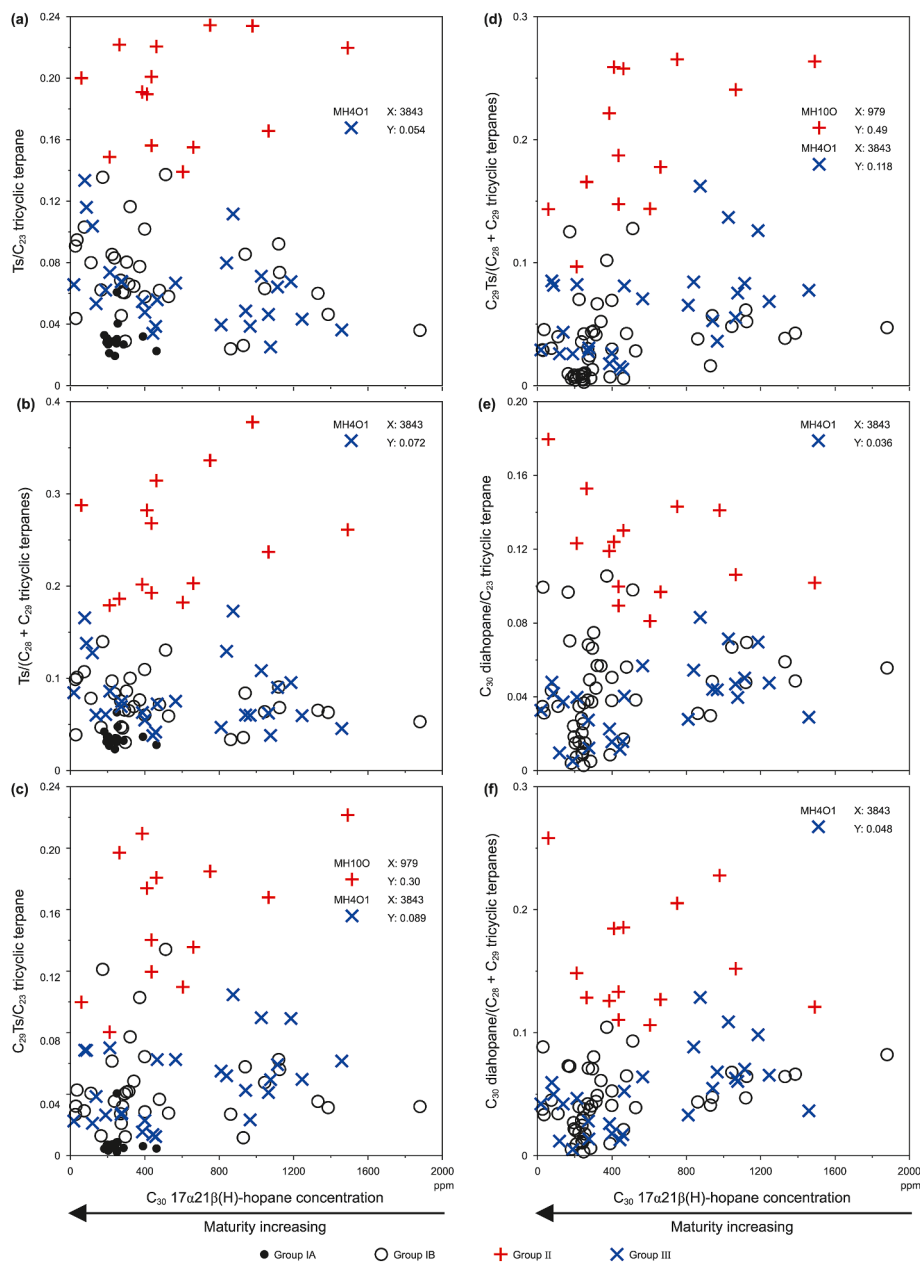


Fig. 7. Crossplots of concentration of C_{30} hopane versus ratios of Ts/C_{23} tricyclic terpane (a), $Ts/(C_{28} + C_{29})$ tricyclic terpanes (b), $C_{29}Ts/C_{23}$ tricyclic terpane (c), $C_{29}Ts/(C_{28} + C_{29})$ tricyclic terpanes (d), C_{30} diahopane/ C_{23} tricyclic terpane (e) and C_{30} diahopane/ $(C_{28} + C_{29})$ tricyclic terpanes (f).

bioprecursors. These eight ratios are influenced by deposition environment in different manner. Concentration of Ts , $C_{29}Ts$ and C_{30} diahopane are more sensitive to water pH value and mineral compositions of source rocks (acidic catalysis) while concentrations of Ph and β -carotane are more sensitive to water Eh value (reduced environment) (e.g., Moldowan et al., 1991; Peters and Moldowan, 1993; Peters et al., 2005). Wang et al. (2022) demonstrated that concentration of β -carotane can be influenced by lacustrine water temperature (hydrothermal activity). PAH components for the six aromatic ratios are non biomarkers. However, concentrations of these PAH components are closely related to kerogen compositions and deposition environment (Requejo et al., 1996). PAH components have complex origins and formation processes in natural system (Hughes et al., 1995; Requejo et al., 1996; Radke et al., 2000). The four pairs of weak linear

correlations with r values 0.21–0.37 are among the PAH ratios and between $TeMNs/(TeMNs + MPs)$ ratio versus ratios of $Ph/n-C_{18}$ and β -carotane/ $n-C_{21}$ (Table 4). The |loading| values from the sixteen parameters for PC1 in PCA vary in a narrow range of 0.184–0.276, demonstrating that all these parameters are effective source facies parameters for the studied oils (Table 3).

5.6. Distributions of effective oil prone source rocks and oil generation, migration and accumulation in the Mahu sag

The Lower Permian Fengcheng (P_1f) and Middle Permian Lower Wuerhe formations (P_2w) are distributed in the whole Mahu sag and the central and southern areas of the basin on the basis of seismic data (Fig. 1, Zhang et al., 1993). However, the distributions of the effective source rocks containing oil-prone

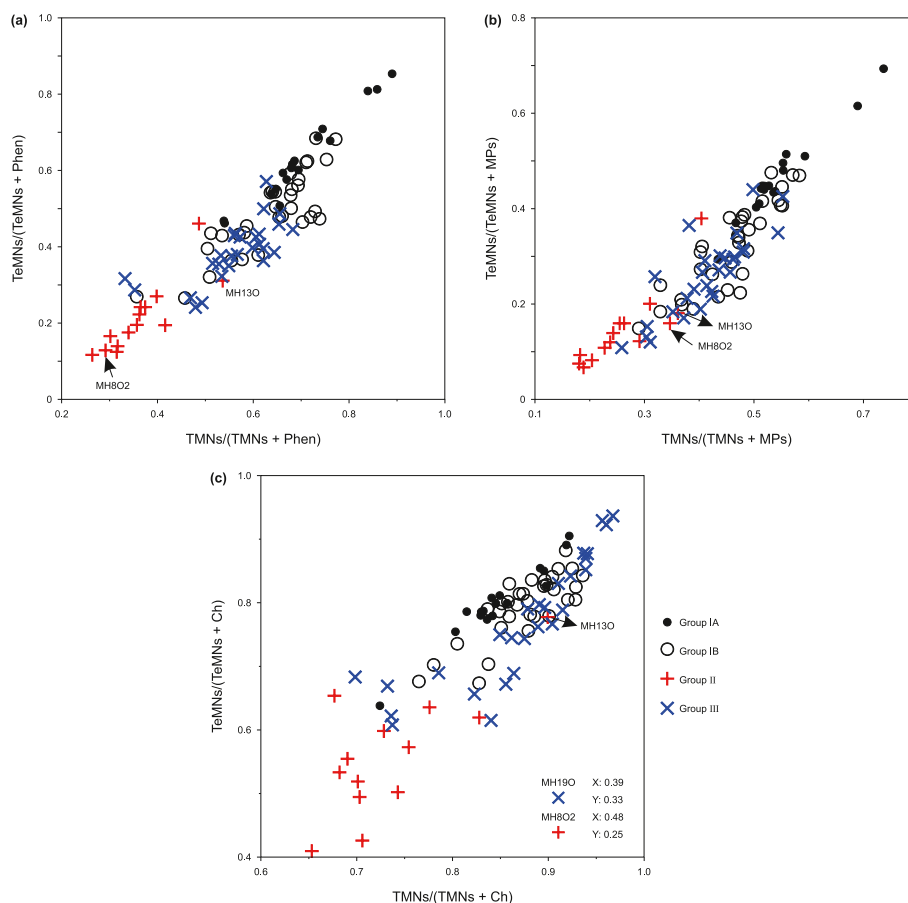


Fig. 8. Crossplots of ratios of $\text{TMNs}/(\text{TMNs} + \text{Phen})$ versus $\text{TeMNs}/(\text{TeMNs} + \text{Phen})$ (a), ratios of $\text{TMNs}/(\text{TMNs} + \text{MPs})$ versus $\text{TeMNs}/(\text{TeMNs} + \text{MPs})$ (b) and ratios of $\text{TMNs}/(\text{TMNs} + \text{Ch})$ versus $\text{TeMNs}/(\text{TeMNs} + \text{Ch})$ (c).

Type I/II kerogen within these two formations remain unclear due to too deep to drill. High quality source rocks containing oil-prone Type I/II kerogen within the Lower Permian Fengcheng Formation (P_1f) were drilled from the early 1980s in the border area of the northern section of K-W fault zone near Wuerhe area (Fig. 1(c), Jiang and Fan, 1983; Zhou et al., 1989; Yang et al., 1992; Wang et al., 2013; Cao et al., 2020). However, source rocks containing oil-prone Type I/II kerogen within the Middle Permian Lower Wuerhe Formation (P_2w) were not drilled by boreholes in the northern and central areas of the Mahu sag (e.g., Zhang et al., 1993; Wang et al., 2013). The Lower Wuerhe Formation (P_2w) drilled by the deep borehole AC1 in the central Mahu sag contain only source rocks with gas-prone Type III kerogen (Fig. 1(d) and (e), Yang et al., 1985). Fortunately, a column of dark mudstone core with 0.50 m in length were obtained from the Lower Wuerhe Formation (P_2w) at the depth about 4660 m from borehole JT1 drilled in 2013 in the border area of the southern Mahu sag (Fig. 1(c)). Source rocks of this core column contain oil-prone Type I/II kerogen, having TOC and Rock-Eval hydrogen indices (HI) up to 3.37% and 550 mg/g TOC, respectively (Yu et al., 2017).

For the 92 studied oils, 50 Group I oils are mainly located at the northeastern and central areas of the Mahu sag with only three oils (MH1601, MH702 and MH703) at the southwestern area of the Mahu sag (Table 1, Fig. 1(d)). 14 Group II oils are located at the southwestern area of the Mahu sag (Table 1, Fig. 1(d)). 28 Group III oils are located at the southwestern and central areas of the Mahu sag (Table 1, Fig. 1(d)). Group I and II were derived from the

source rocks within the Lower Permian Fengcheng (P_1f) and Middle Permian Lower Wuerhe formations (P_2w), respectively while Group III oils are mixtures of Group I and II on the basis of the sixteen source facies parameters of the present study in combination with the previous studies (Zhou et al., 1989; Yang et al., 1992; Zhang et al., 1993, 2022; Wang et al., 2013, 2023; Yu et al., 2017). On the basis of oil locations of the three groups, the effective oil prone source rocks with Type I/II kerogen of the Fengcheng Formation occur mainly in the northeastern and central areas of the Mahu sag with minor in the southwestern area of the Mahu sag while those of the Lower Wuerhe Formation occur only in the southwestern area of the Mahu sag in the studied region (Fig. 1(d)). The eight Group III oils in the central area of the Mahu sag are located at reservoir rocks within the strata from the Lower Jurassic Badaowan Formation (J_1b) to Middle Triassic Karamay Formation (T_2k) at burial intervals from 2436–2440 m to 3358–3361 m, which are generally younger and shallower compared with the reservoir rocks for the Group I oils in this region (Table 1, Fig. 1(d)). These Group III oils possibly consist of some Group II oil components that were generated from source rocks within the Middle Permian Lower Wuerhe Formation (P_2w) in the southern area of the Mahu sag and migrated along slope and faults to the younger and shallower reservoir rocks in the central area of the Mahu sag. During migration and accumulation to reservoirs, the Group II oil components were mixed with Group I oil components, resulting in the formation of these Group III oils in the central area of the

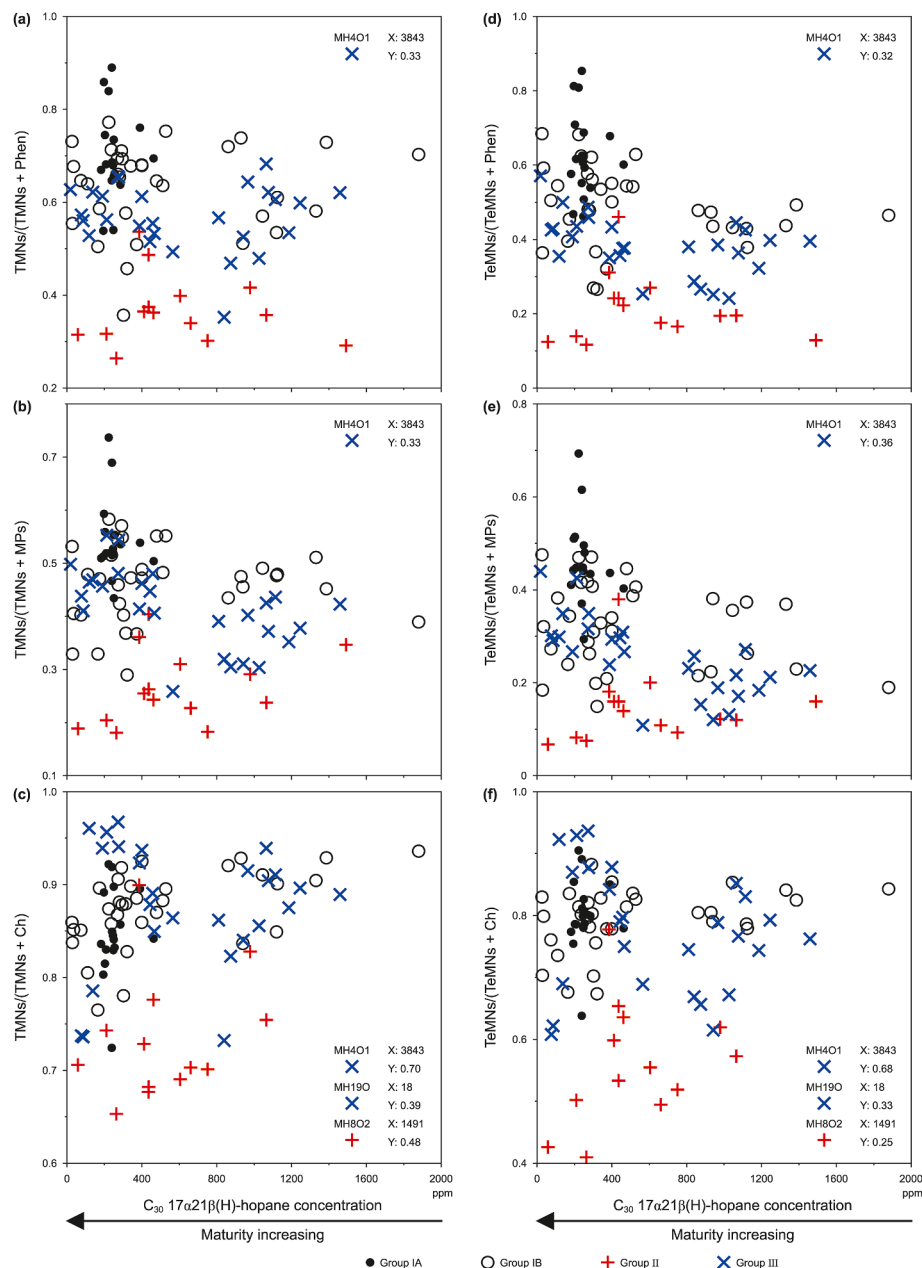


Fig. 9. Crossplots of concentration of C_{30} hopane versus ratios of TMNs/(TMNs + Phen) (a), TMNs/(TMNs + MPs) (b), TMNs/(TMNs + Ch) (c), TeMNs/(TeMNs + Phen) (d), TeMNs/(TeMNs + MPs) (e) and TeMNs/(TeMNs + Ch) (f).

Mahu sag (Fig. 1(d)). The strata from the Mesozoic to Cenozoic incline to the south because the Cretaceous and Cenozoic strata increasingly thicken southwards in the whole Junggar Basin (Zhang et al., 1993).

Source rocks within the Lower Permian Fengcheng Formation (P_{1f}) started oil generation from the beginning of Triassic Period and ended at the Middle Cretaceous Period while source rocks within the Middle Permian Lower Wuerhe Formation (P_{2w}) started oil generation from the Early Jurassic Period and continued to the present in the southern Mahu sag (e.g., Zhou et al., 1989; Zhang et al., 1993; Wang et al., 2013; Xiang et al., 2016; Tao et al., 2021). Source rocks of the Fengcheng Formation (P_{1f}) currently have maturities at the peak oil generation stage ($\%R_0$ 0.8–1.1) in the border area (fault zone) of the northern Mahu sag and post mature in the Mahu sag ($\%R_0 > 1.4$, up to 2.0 or even higher) while source rocks of the Lower Wuerhe Formation (P_{2w}) currently have

maturities at the late oil generative window to post mature in the Southwestern Mahu sag ($\%R_0$ 1.1–1.4) (e.g., Zhou et al., 1989; Zhang et al., 1993; Wang et al., 2013; Xiang et al., 2016; Tao et al., 2021). Oils from source rocks of the two formations migrated first via high-angle strike-slip faults, and then from the central area to the border area (fault zone) of the sag, or/and northwards from the southern area to the central area of the sag along the unconformities within the Permian, Triassic and Jurassic strata on the basis of geological setting of the Mahu sag (Fig. 1(e), Zhang et al., 1993; Kuang et al., 2013; Ablimit et al., 2016; Lei et al., 2017; Zhi et al., 2018; Tang et al., 2019; Tao et al., 2021). The distributions of Group I, II and III oils clearly demonstrate the locations and distributions of the effective source rocks containing oil prone Type I/II kerogen within the Fengcheng (P_{1f}) and Lower Wuerhe (P_{2w}) formations in the studied area. This result can be very useful for further oil exploration in this area.

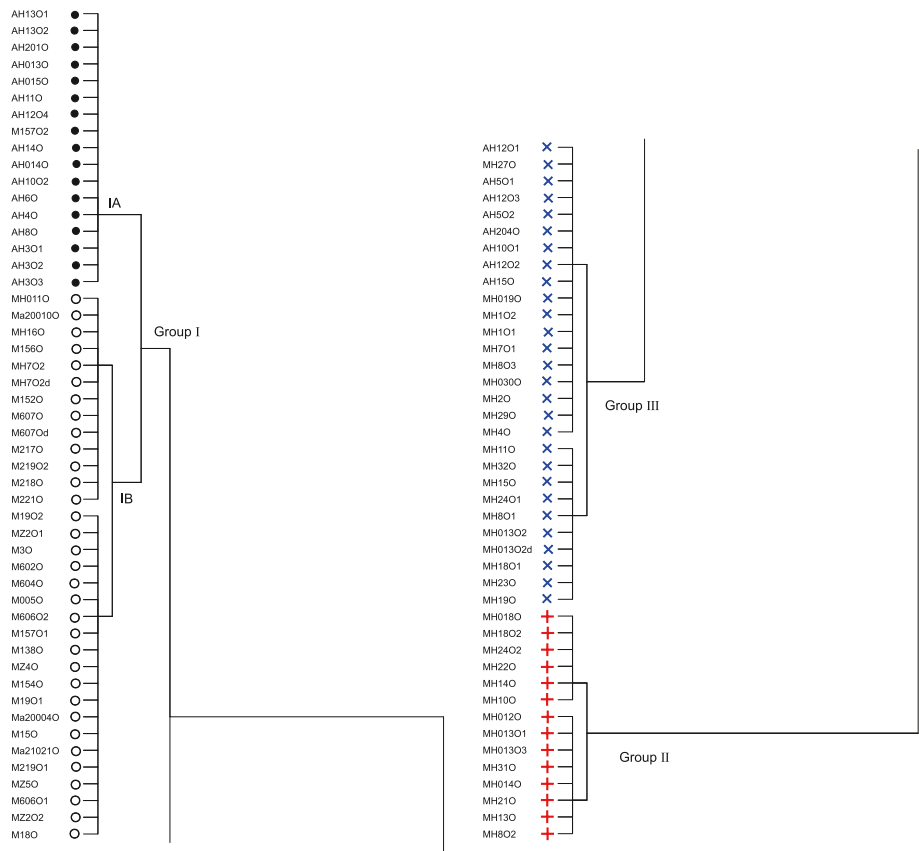


Fig. 10. Hierarchical cluster analysis dendrogram classifying the 92 studied oils on the basis of sixteen source facies parameters.

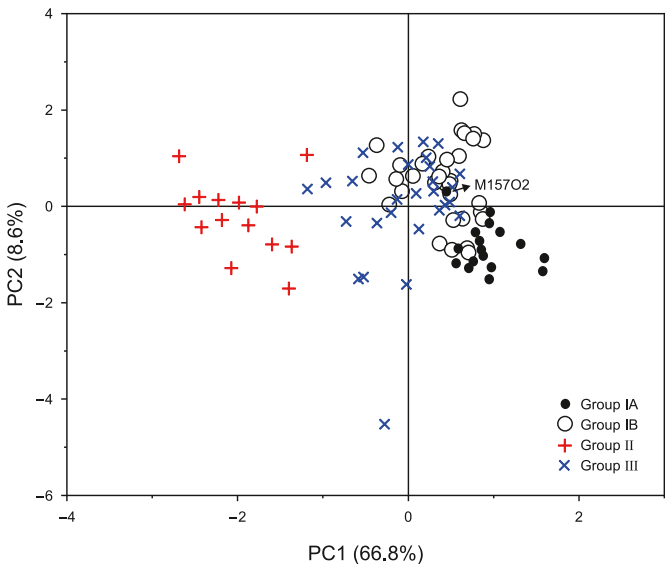


Fig. 11. Two-dimensional principle component scores plot for the 92 studied oils on the basis of sixteen source facies parameters.

5.7. Implications

Oil-source correlations are performed mainly on the basis of molecular and isotopic parameters (e.g., Mackenzie, 1984; Seifert and Moldowan, 1986; Peters et al., 2005). Oil classification and oil source correlation for a large number of oil samples are difficult tasks. In the present study, all the sixteen selected parameters vary

Table 2			
Principal components and percentage of variance.			
Principal components	Eigenvalue	Percentage of variance	Cumulative, %
PC1	10.68684	66.7927	66.793
PC2	1.38121	8.6326	75.425
PC3	1.22399	7.6499	83.075
PC4	0.96712	6.0445	89.120
PC5	0.68693	4.2933	93.413
PC6	0.33522	2.0951	95.508
PC7	0.22998	1.4374	96.945
PC8	0.16771	1.0482	97.994
PC9	0.14331	0.8957	98.889
PC10	0.0748	0.4675	99.357
PC11	0.05329	0.3331	99.690
PC12	0.03127	0.1954	99.885
PC13	0.01279	0.0799	99.965
PC14	0.00351	0.0219	99.987
PC15	0.00135	0.0085	99.996
PC16	7.09E-04	0.0044	100.000

consecutively and are partially inconsistent to each other. For example, Group II oils MH130 and MH802 are together with Group III oils and close to Group I oils in the crossplot of the two isotopic parameters (Fig. 3(b)). However, these two oils are together with the other Group II oils in the crossplots between each other of the six terpene parameters (Fig. 5(b)–(d)). The three oil groups classified from HCA and PCA partially overlap in all crossplots of one parameter versus another among the sixteen source facies parameters, in particular between Group III and Group I or II oils (Figs. 3, 5 and 8). It is difficult to identify which parameter is more reliable than the others. It is arbitrary or impossible to set some criteria values of some source facies

Table 3

Loadings from the sixteen parameters for PC1 and PC2 in PCA.

Parameters	PC1	PC2
RI in ‰	−0.234	−0.155
$\delta^{13}\text{C}$ of $n\text{-C}_{25}$ in ‰	0.261	0.098
Ph/ $n\text{-C}_{18}$	0.223	0.405
β -carotane/ $n\text{-C}_{21}$	0.193	0.252
Ts/ C_{23} tricyclic terpanes	−0.270	−0.043
Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes)	−0.276	−0.048
C_{29} Ts/ C_{23} tricyclic terpanes	−0.269	0.070
C_{29} Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes)	−0.267	0.066
C_{30} dihopane/ C_{23} tricyclic terpane	−0.264	0.177
C_{30} dihopane/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes)	−0.274	0.150
TMNs/(TMNs + Phen)	0.268	−0.054
TeMNs/(TeMNs + Phen)	0.258	−0.286
TMNs/(TMNs + MPs)	0.265	−0.261
TeMNs/(TeMNs + MPs)	0.228	−0.440
TMNs/(TMNs + Ch)	0.184	0.501
TeMNs/(TeMNs + Ch)	0.241	0.270

parameters to classify the 92 studied oils into two or three groups. HCA and PCA are suitable to resolve this complicated issue. Oil classification can be more reasonable and reliable on the basis of integration of a large set of source facies parameters using HCA and PCA compared with using a single crossplot of two parameters. Furthermore, it is convenient to classify the studied oils in groups by HCA and PCA integrating the data of all the selected facies parameters (Figs. 10 and 11). All the selected parameters generally increase in the sequence of Group I < Group III < Group II, or decrease in the sequence of Group I > Group III > Group II although there are some overlaps between each other of the three groups, demonstrating that these parameters are effective for oil classification (Figs. 3(b), 5 and 8). The six terpane ratios and six PAH ratios are rarely used for routine oil source correlation. This work is the first case study to utilize the six PAH ratios by HCA and PCA for oil classification and source facies assessment.

The sixteen source facies parameters are less influenced by maturity as demonstrated in the crossplots of these sixteen parameters versus concentration of C_{30} hopane (Figs. 3, 6, 7 and 9). Ratios of Ph/ $n\text{-C}_{18}$ and β -carotane/ $n\text{-C}_{21}$ are sensitive to oil biodegradation even at low extent (e.g., Peters and Moldowan,

1993). However, the twelve terpane and PAH ratios are not influenced by biodegradation up to medium extent (e.g., Peters and Moldowan, 1993). Thus, these terpane and PAH ratios can be widely used for oil source assessment using HCA and PCA in other regions of the Junggar Basin and elsewhere.

6. Conclusions

Sixteen facies parameters were selected and integrated to classify the 92 studied oils using HCA and PCA in the Mahu sag of the Junggar Basin. This classification is more reliable and convenient, compared with using crossplots of two parameters or star charts of several parameters. The sixteen parameters include isotope reversal index (RI), $\delta^{13}\text{C}$ of $n\text{-C}_{25}$, Ph/ $n\text{-C}_{18}$, β -carotane/ $n\text{-C}_{21}$, six terpane ratios of Ts/ C_{23} tricyclic terpanes, Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes), C_{29} Ts/ C_{23} tricyclic terpanes, C_{29} Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes), C_{30} dihopane/ C_{23} tricyclic terpane and C_{30} dihopane/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes), and six PAH ratios of TMNs/(TMNs + Phen), TeMNs/(TeMNs + Phen), TMNs/(TMNs + MPs), TeMNs/(TeMNs + MPs), TMNs/(TMNs + Ch) and TeMNs/(TeMNs + Ch). These sixteen parameters are mainly influenced by source facies and less influenced by maturity as demonstrated in the crossplots of these sixteen parameters versus concentration of C_{30} hopane. The 92 oils are classified into three groups, i.e., Group I, II and III that are derived from source rocks within the Lower Permian Fengcheng Formation (P_{1f}), Middle Permian Lower Wuerhe Formation (P_{2w}) and both, respectively. Group I consists of fifty oils mainly located at the northeastern and central areas of the Mahu sag with only three oils (MH1601, MH702 and MH703) at the southwestern area of the Mahu sag. Group II consists of fourteen oils at the southwestern area of the Mahu sag. Group III consists of twenty-eight oils located at the southwestern and central areas of the Mahu sag. Group I oils were further classified into Subgroup IA and IB, including seventeen and thirty-three oils, respectively. Subgroup IA oils are located at the central area of the Mahu sag while Subgroup IB oils occur widely in most areas of the Mahu sag. Classification of Subgroup IA and IB reflects subtle facies changes of source rocks within the Fengcheng Formation (P_{1f}). Locations of Group I, II and III oils indicate the distributions of effective source rocks containing oil-prone TypeI/II

Table 4Pearson correlation coefficients (r) between each other for the sixteen source facies parameters.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1.00	−0.94	−0.79	−0.45	0.55	0.63	0.58	0.63	0.47	0.58	−0.67	−0.64	−0.63	−0.52	−0.40	−0.53
2	−0.94	1.00	0.74	0.44	−0.65	−0.74	−0.68	−0.73	−0.63	−0.74	0.73	0.68	0.72	0.58	0.48	0.64
3	−0.79	0.74	1.00	0.66	−0.60	−0.63	−0.56	−0.56	−0.47	−0.53	0.60	0.50	0.49	0.35	0.54	0.60
4	−0.45	0.44	0.66	1.00	−0.61	−0.58	−0.49	−0.46	−0.53	−0.50	0.50	0.44	0.48	0.36	0.41	0.47
5	0.55	−0.65	−0.60	−0.61	1.00	0.96	0.90	0.86	0.82	0.82	−0.70	−0.62	−0.65	−0.52	−0.51	−0.63
6	0.63	−0.74	−0.63	−0.58	0.96	1.00	0.89	0.91	0.80	0.87	−0.70	−0.63	−0.66	−0.53	−0.51	−0.63
7	0.58	−0.68	−0.56	−0.49	0.90	0.89	1.00	0.97	0.81	0.82	−0.69	−0.66	−0.66	−0.58	−0.42	−0.58
8	0.63	−0.73	−0.56	−0.46	0.86	0.91	0.97	1.00	0.77	0.85	−0.67	−0.65	−0.65	−0.57	−0.41	−0.56
9	0.47	−0.63	−0.47	−0.53	0.82	0.80	0.81	0.77	1.00	0.95	−0.71	−0.70	−0.75	−0.66	−0.42	−0.62
10	0.58	−0.74	−0.53	−0.50	0.82	0.87	0.82	0.85	0.95	1.00	−0.70	−0.70	−0.76	−0.67	−0.42	−0.63
11	−0.67	0.73	0.60	0.50	−0.70	−0.70	−0.69	−0.67	−0.71	−0.70	1.00	0.92	0.84	0.70	0.58	0.72
12	−0.64	0.68	0.50	0.44	−0.62	−0.63	−0.66	−0.65	−0.70	−0.70	0.92	1.00	0.90	0.90	0.37	0.63
13	−0.63	0.72	0.49	0.48	−0.65	−0.66	−0.66	−0.65	−0.75	−0.76	0.84	0.90	1.00	0.93	0.44	0.69
14	−0.52	0.58	0.35	0.36	−0.52	−0.53	−0.58	−0.57	−0.66	−0.67	0.70	0.90	0.93	1.00	0.21	0.53
15	−0.40	0.48	0.54	0.41	−0.51	−0.51	−0.42	−0.41	−0.42	−0.42	0.58	0.37	0.44	0.21	1.00	0.91
16	−0.53	0.64	0.60	0.47	−0.63	−0.63	−0.58	−0.56	−0.62	−0.63	0.72	0.63	0.69	0.53	0.91	1.00

1: Isotope reversal index (RI) in ‰; 2: $\delta^{13}\text{C}$ of $n\text{-C}_{25}$ in ‰; 3: Ph/ $n\text{-C}_{18}$; 4: β -carotane/ $n\text{-C}_{21}$; 5: Ts/ C_{23} tricyclic terpanes; 6: Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes); 7: C_{29} Ts/ C_{23} tricyclic terpanes; 8: C_{29} Ts/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes); 9: C_{30} dihopane/ C_{23} tricyclic terpane; 10: C_{30} dihopane/($\text{C}_{28}+\text{C}_{29}$ tricyclic terpanes); 11: TMNs/(TMNs + Phen); 12: TeMNs/(TeMNs + Phen); 13: TMNs/(TMNs + MPs); 14: TeMNs/(TeMNs + MPs); 15: TMNs/(TMNs + Ch); 16: TeMNs/(TeMNs + Ch). $|r| > 0.80$: very strong positive or negative linear correlation; $0.80 > |r| > 0.60$: strong positive or negative linear correlation; $0.60 > |r| > 0.40$: moderate positive or negative linear correlation; $0.40 > |r| > 0.20$: weak positive or negative linear correlation; $|r| < 0.20$: no linear correlation (Zou et al., 2021).

kerogen within the Fengcheng (P_{1f}) and Lower Wuerhe formations (P_{2w}).

CRediT authorship contribution statement

Hang-Xin Cai: Software, Investigation, Data curation. **Jun Jin:** Resources, Investigation. **Er-Ting Li:** Resources, Investigation. **Zhong-Da Zhang:** Investigation, Data curation. **Shuang Yu:** Software, Investigation, Data curation. **Chang-Chun Pan:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was jointly funded by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA14010104) and the National S&T Major Project of China (Grant No. 2017ZX05008-002-030). We are very grateful to five anonymous reviewers for their helpful comments. We Thank the Editor Dr Hao Jie for editorial work. This is contribution No.IS-3655 from GIGCAS.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.petsci.2025.05.024>.

References

- Ablimit, I., Tang, Y., Cao, J., Chen, G., Chen, J., Tao, K., 2016. Accumulation mechanism and enrichment rules of the continuous hydrocarbon plays in the Lower Triassic Baikouquan Formation of the Mahu sag, Junggar Basin. *Natural Gas Geoscience* 27, 241–250 (in Chinese).
- Alexander, R., Kagi, R.L., Woodhouse, G.W., 1981. Geochemical correlation of windalia oil and extracts of winning group (cretaceous) potential source rock, barrow subbasin, western Australia. *AAPG Bull.* 65, 235–250. <https://doi.org/10.1306/2F9197B0-16CE-11D7-8645000102C1865D>.
- Alexander, R., Kagi, R.L., Rowland, S.J., Sheppard, P.N., Chirila, T.V., 1985. The effects of thermal maturity on distributions of dimethylnaphthalenes and trimethylnaphthalenes in some Ancient sediments and petroleum. *Geochem. Cosmochim. Acta* 49, 385–395. [https://doi.org/10.1016/0016-7037\(85\)90031-6](https://doi.org/10.1016/0016-7037(85)90031-6).
- Bian, Y., Zhou, C., Zhang, H., Xiao, Z., Zhan, Z., Yu, S., Pan, C., 2024. Chemometric analysis on concentrations and ratios of terpanes and steranes and implications for oils from the cratonic region of the Tarim Basin, NW China. *Petroleum Research* 9, 565–585. <https://doi.org/10.1016/j.ptlrs.2024.06.001>.
- Cai, H., Jin, J., Li, E., Zhang, Z., Gu, Y., Wang, Y., Yu, S., Pan, C., 2023. Unraveling gas charging and leakage for oil reservoirs in the Mahu sag of the Junggar Basin, NW China using concentrations and ratios of biomarkers, light hydrocarbons and diamondoids. *Org. Geochem.* 177, 104543. <https://doi.org/10.1016/j.orggeochem.2022.104543>.
- Cao, J., Zhang, Y., Hu, W., Yao, S., Wang, X., Zhang, Y., Tang, Y., 2005. The Permian hybrid petroleum system in the northwest margin of the Junggar Basin, NW China. *Mar. Petrol. Geol.* 22, 331–349. <https://doi.org/10.1016/j.marpetgeo.2005.01.005>.
- Cao, J., Yao, S.P., Jin, Z.J., Hu, W.X., Zhang, Y.J., 2006. Petroleum migration and mixing in the northwestern Junggar Basin (NW China): constraints from oil-bearing fluid inclusion analyses. *Org. Geochem.* 37, 827–846. <https://doi.org/10.1016/j.orggeochem.2006.02.003>.
- Cao, J., Lei, D., Li, Y., Tang, Y., Abulimit, I., Chang, Q., Wang, T., 2015. Ancient high-quality lacustrine source rocks discovered in the lower permian Fengcheng Formation, Junggar Basin. *Acta Petrol. Sin.* 36, 781–790. <https://doi.org/10.7623/syxb201507002> (in Chinese).
- Cao, J., Xia, L., Wang, T., Zhi, D., Tang, Y., Li, W., 2020. An alkaline lake in the Late Paleozoic Ice Age (LPIA): a review and new insights into paleoenvironment and petroleum geology. *Earth Sci. Rev.* 202, 103091. <https://doi.org/10.1016/j.earscirev.2020.103091>.
- Chen, G., Ablimit, Bai, L., Zhang, J., Bain, B., 2013. Petroleum accumulation field in the deep strata of the eastern slope area of the Mahu sag, Junggar Basin. *Journal of Southwest Petroleum University: Science & Technology Edition* 35 (6), 31–38 (in Chinese).
- Chen, G., An, Z., Abulimit, Li, X., Xu, Q., Zhang, L., 2014. Petroleum exploration prospects of carboniferous-permian in peripheral Mahu sag, Junggar Basin. *Xinjing Pet. Geol.* 35, 259–263. <https://doi.org/10.3863/j.issn.1674-5086.2013.06.004> (in Chinese).
- Chen, J., Wang, X., Deng, C., Liang, D., Zhang, Y., Zhao, Z., Ni, Y., Zhi, D., Yang, H., Wang, Y., 2016a. Geochemical Features of source rocks and crude oils in the Junggar Basin, Northwest China. *Acta Geol. Sin.* 90, 37–67. <https://doi.org/10.3969/j.issn.0001-5717.2016.01.003> (in Chinese).
- Chen, J., Wang, X., Deng, C., Liang, D., Zhang, Y., Zhao, Z., Ni, Y., Zhi, D., Yang, H., Wang, Y., 2016b. Oil and gas source, occurrence, and petroleum system in the Junggar Basin, Northwest China. *Acta Geol. Sin.* 90, 421–450. <https://doi.org/10.3969/j.issn.0001-5717.2016.03.002> (in Chinese).
- Chen, J., Wang, X., Deng, C., Zhao, Z., Ni, Y., Sun, Y., Yang, H., Wang, H., Liang, D., 2016c. Oil-source correlation of typical crude oils in the southern margin, Junggar Basin, Northwestern China. *Acta Pet. Sin.* 37 (2), 160–171. <https://doi.org/10.7623/syxb201602002> (in Chinese).
- Connan, J., 1981. Biological markers in crude oils. In: Mason, J.F. (Ed.), *Petroleum Geology in China*. Penn Well, Tulsa, OK, pp. 48–70.
- Dahl, J.E., Moldowan, J.M., Peters, K., Claypool, G., Rooney, M., Michael, G., Mellos, M., Kohnen, M., 1999. Diamondoid hydrocarbons as indicators of oil cracking. *Nature* 399, 54–56. <https://doi.org/10.1038/19953>.
- De Grande, S.M.B., Aquino Neto, F.R., Mello, M.R., 1993. Extended tricyclic terpanes in sediments and petroleum. *Org. Geochem.* 20, 1039–1047. [https://doi.org/10.1016/0146-6380\(93\)90112-0](https://doi.org/10.1016/0146-6380(93)90112-0).
- Farrimond, P., Telnæs, N., 1996. Three series of rearranged hopanes in Toarcian sediments (Northern Italy). *Org. Geochem.* 25, 165–177. [https://doi.org/10.1016/S0146-6380\(96\)00127-1](https://doi.org/10.1016/S0146-6380(96)00127-1).
- Farrimond, P., Taylor, A., Telnæs, N., 1998. Biomarker maturity parameters: the role of generation and thermal degradation. *Org. Geochem.* 29, 1181–1197. [https://doi.org/10.1016/S0146-6380\(98\)00079-5](https://doi.org/10.1016/S0146-6380(98)00079-5).
- Fu, J., Sheng, G., Xu, J., Eglinton, G., Gowar, A.P., Jia, R., Fan, S., Peng, P., 1990. Application of biological markers in the assessment of paleoenvironments of Chinese non-marine sediments. *Org. Geochem.* 16, 769–779. [https://doi.org/10.1016/0146-6380\(90\)90116-H](https://doi.org/10.1016/0146-6380(90)90116-H).
- Hall, P.H., Douglas, A.G., 1983. The distribution of cyclic alkanes in two lacustrine deposits. In: Bjorøy, M., Albrecht, P., Cornford, C., de Groot, K., Eglinton, G., Galimov, E., Leythaeuser, D., Pelet, R., Rullkötter, J., Speers, G. (Eds.), *Advances in Organic Geochemistry 1981*. John Wiley & Sons, pp. 576–587.
- Hayes, J.M., Takigiku, R., Ocampo, R., Callot, H.J., Albrecht, P., 1987. Isotopic compositions and probable origins of organic molecules of the Eocene Messel shale. *Nature* 329, 48–51. <https://doi.org/10.1038/329048a0>.
- Hayes, J.M., Freeman, K.H., Popp, B.N., Hoham, C.H., 1990. Compound-specific isotopic analyses: a novel tool for reconstruction of ancient biochemical processes. *Org. Geochem.* 16, 1115–1128. [https://doi.org/10.1016/0146-6380\(90\)90147-R](https://doi.org/10.1016/0146-6380(90)90147-R).
- Holba, A.G., Ellis, L., Dzou, L.L., Hallam, A., Masterson, W.D., Francu, J., Fincannon, A. L., 2001. Extended tricyclic terpanes as age discriminators between Triassic, early Jurassic and middle-late Jurassic oils. In: *20th International Meeting on Organic Geochemistry, Abstracts*, vol. 1, p. 464.
- Holba, A.G., Zumberge, J., Huizinga, B.J., Rosenstein, H., Dzou, L.I.P., 2003. Extended tricyclic terpanes as indicators of marine upwelling. In: *21st International Meeting on Organic Geochemistry. Book of Abstracts Part I*, p. 131.
- Huang, H., Zhang, S., Gu, Y., Su, J., 2017. Impacts of source input and secondary alteration on the extended tricyclic terpene ratio: a case study from Palaeozoic sourced oils and condensates in the Tarim Basin, NW China. *Org. Geochem.* 112, 158–169. <https://doi.org/10.1016/j.orggeochem.2017.07.012>.
- Hughes, W.B., Holba, A.G., Dzou, L.I.P., 1995. The ratios of dibenzothiophene to phenanthrene and pristane to phytane as indicators of depositional environment and lithology of petroleum source rocks. *Geochem. Cosmochim. Acta* 59, 3581–3598. [https://doi.org/10.1016/0016-7037\(95\)00225-0](https://doi.org/10.1016/0016-7037(95)00225-0).
- Irwin, H., Meyer, T., 1990. Lacustrine organic facies. A biomarker study using multivariate statistical analysis. *Org. Geochem.* 16, 176–210. [https://doi.org/10.1016/0146-6380\(90\)90040-7](https://doi.org/10.1016/0146-6380(90)90040-7).
- Jiang, Z., Fan, G., 1983. Organic geochemistry of source rocks within carboniferous Fengcheng Formation in Junggar Basin. *Xinjing Pet. Geol.* 3 (3), 74–91 (in Chinese).
- Jiang, Z., Fowler, M.G., 1986. Carotenoid-derived alkanes in oils from northwestern China. *Org. Geochem.* 10, 831–839. [https://doi.org/10.1016/S0146-6380\(86\)80020-1](https://doi.org/10.1016/S0146-6380(86)80020-1).
- Kruege, M.A., Hubert, J.F., Bensley, D.F., Crelling, J.C., Akes, R.J., Meriney, P.E., 1990. Organic geochemistry of a lower jurassic synrift lacustrine sequence, Hartford Basin, Connecticut, USA. *Org. Geochem.* 16, 689–701. [https://doi.org/10.1016/0146-6380\(90\)90110-L](https://doi.org/10.1016/0146-6380(90)90110-L).
- Kuang, L., Tang, Y., Lei, D., Wu, T., Qu, J., 2013. Exploration of Fan-controlled large-area lithological oil reservoirs of Triassic Baikouquan Formation in slope zone of Mahu depression in Junggar Basin. *China Petroleum Exploration* 19, 14–23 (in Chinese).
- Lei, D., Chen, G., Liu, H., Li, X., Abulimit, I., Tao, K., Cao, J., 2017. Study on the forming conditions and exploration fields of the Mahu Giant Oil (Gas) Province, Junggar

- Basin. *Acta Geologica Sinica* 91, 1604–1619. <https://doi.org/10.19762/j.cnki.dizhixuebao.2017.07.012> (in Chinese).
- Li, X., Zha, M., Wu, K., 2007. Geochemical feature of natural gas in Wu-Xia area, Junggar Basin. *Xinjing Pet. Geol.* 28 (4), 413–415 (in Chinese). <https://www.zgxpjg.com/CN/Y2007/V28/I4/413>.
- Liu, B., He, B., Huang, Z., Zhang, Y., Yin, Z., Guo, T., Wu, F., 2014. Sources and distribution patterns of natural gas of different genetic types at the northwestern margin of the Junggar Basin. *Nat. Gas. Ind.* 34 (9), 40–46. <https://doi.org/10.3787/j.issn.1000-0976.2014.09.006> (in Chinese).
- Mackenzie, A.S., 1984. Application of biomarkers in petroleum geochemistry. In: Brooks, J., Welte, D. (Eds.), *Advances in Petroleum Geochemistry*, vol. 1. Academic Press, London, pp. 115–214.
- Mackenzie, A.S., Rullkötter, J., Welte, D.H., Mankiewicz, P., 1985. Reconstruction of oil formation and accumulation in the North Slope, Alaska, using quantitative gas chromatography-mass spectrometry. In: Magoon, L.B., Claypool, G.E. (Eds.), *Alaska North Slope Oil/Source Rock Correlation Study*. American Association of Petroleum Geologists, Tulsa, OK, pp. 319–377.
- Moldowan, J.M., Sundaram, P., Schoell, M., 1986. Sensitivity of biomarker properties to deposition environment and/or source input in the Lower Toarcian of S. W. Germany. *Org. Geochem.* 10, 915–926. [https://doi.org/10.1016/S0146-6380\(86\)80029-8](https://doi.org/10.1016/S0146-6380(86)80029-8).
- Moldowan, J.M., Fago, F.J., Carlson, R.M.K., Young, D.C., van Duyne, G., Clardy, J., Schoell, M., Pillinger, C.T., Watt, D.S., 1991. Rearranged hopanes in sediments and petroleum. *Geochim. Cosmochim. Acta* 55, 3333–3353. [https://doi.org/10.1016/0016-7037\(91\)90492-N](https://doi.org/10.1016/0016-7037(91)90492-N).
- Murray, A.P., Peters, K.E., 2021. Quantifying multiple source rock contributions to petroleum fluids: bias in using compound ratios and neglecting the gas fraction. *AAPG Bull.* 105, 1661–1678. <https://doi.org/10.1306/03122120056>.
- Palacas, J.G., 1984. Carbonate rocks as source of petroleum: geological and geochemical characteristics and oil-source correlation. In: *Proceedings of the 11th World Petroleum Congress 1983*, vol. 2. John Wiley & Sons, Chichester, UK, pp. 31–43.
- Pan, C., Yang, J., 2000. Geochemical characteristics and implications of hydrocarbons in reservoir rocks of Junggar Basin, China. *Chem. Geol.* 167, 321–335. [https://doi.org/10.1016/S0009-2541\(99\)00236-3](https://doi.org/10.1016/S0009-2541(99)00236-3).
- Pan, C., Yang, J., Fu, J., Sheng, G., 2003. Molecular correlation of free oil and inclusion oil of reservoir rock in the Junggar Basin, China. *Org. Geochem.* 34, 357–374. [https://doi.org/10.1016/S0146-6380\(02\)00238-3](https://doi.org/10.1016/S0146-6380(02)00238-3).
- Pan, C., Wang, J., Yu, S., Xiang, B., Liao, J., Ren, J., Li, E., Zhang, X., 2021. Oil origins, mixing and biodegradation in southwestern Junggar Basin, NW China. *J. Petrol. Sci. Eng.* 196, 108017. <https://doi.org/10.1016/j.petrol.2020.108017>.
- Peters, K.E., Moldowan, J.M., 1993. *The biomarker guide*. In: *Interpreting Molecular Fossils in Petroleum and Ancient Sediments*. Prentice-Hall, Englewood Cliffs, NJ.
- Peters, K.E., Walters, C.C., Moldowan, J.M., 2005. *The Biomarker Guide*. In: *Biomarkers and Isotopes in Petroleum Exploration and Earth History*, 2. Cambridge University Press, UK, p. 1155. <https://doi.org/10.1017/CBO9781107326040>.
- Peters, K.E., Ramos, L.S., Zumberge, J.E., Valin, Z.C., Scotese, C.R., Gautier, D.L., 2007. Circum-Arctic petroleum systems identified using decision-tree chemometrics. *AAPG Bull.* 91, 877–913. <https://doi.org/10.1306/12290606097>.
- Peters, K.E., Coutrot, D., Nouvelle, X., Ramos, L.S., Rohrbach, B.G., Magoon, L.B., Zumberge, J.E., 2013. Chemometric differentiation of crude oil families in the San Joaquin Basin, California. *AAPG Bull.* 97, 103–143. <https://doi.org/10.1306/05231212018>.
- Peters, K.E., Wright, T.L., Ramos, L.S., Zumberge, J.E., Magoon, L.B., 2016. Chemometric recognition of genetically distinct oil families in the Los Angeles Basin, California. *AAPG Bull.* 100, 115–135. <https://doi.org/10.1306/08031515068>.
- Radke, M., Welte, D.H., 1983. The methylphenanthrene index (MPI): a maturity parameter based on aromatic hydrocarbons. In: Bjørøy, M. (Ed.), *Advances in Organic Geochemistry 1981*. Wiley, Chichester, pp. 504–512.
- Radke, M., Welte, D.H., Willsch, H., 1986. Maturity parameters based on aromatic hydrocarbons: influence of the organic matter type. *Org. Geochem.* 10 (1–3), 51–63. [https://doi.org/10.1016/0146-6380\(86\)90008-2](https://doi.org/10.1016/0146-6380(86)90008-2).
- Radke, M., Vriend, S.P., Ramanampisoa, L.R., 2000. Alkyldibenzofurans in terrestrial rocks: influence of organic facies and maturation. *Geochim. Cosmochim. Acta* 64, 275–286. [https://doi.org/10.1016/S0016-7037\(99\)00287-2](https://doi.org/10.1016/S0016-7037(99)00287-2).
- Requejo, A.G., Sassen, R., McDonald, T., Denoux, G., Kennicutt, M.C., Brooks, J.M., 1996. Polynuclear aromatic hydrocarbons (PAH) as indicators of the source and maturity of marine crude oils. *Org. Geochem.* 24, 1017–1033. [https://doi.org/10.1016/S0146-6380\(96\)00079-4](https://doi.org/10.1016/S0146-6380(96)00079-4).
- Rullkötter, J., Spiro, B., Nissenbaum, A., 1985. Biological marker characteristics of oils and asphalts from carbonate source rocks in a rapidly subsiding graben. Dead Sea, Israel. *Geochim. Cosmochim. Acta* 49, 141–157. [https://doi.org/10.1016/0016-7037\(85\)90286-8](https://doi.org/10.1016/0016-7037(85)90286-8).
- Seifert, W.K., Moldowan, J.M., 1978. Applications of steranes, terpanes, and monoaromatics to the maturation, migration, and source of crude oils. *Geochim. Cosmochim. Acta* 42, 77–95. [https://doi.org/10.1016/0016-7037\(78\)90219-3](https://doi.org/10.1016/0016-7037(78)90219-3).
- Seifert, W.K., Moldowan, J.M., 1986. Use of biological markers in petroleum exploration. In: Johns, R.B. (Ed.), *Biological Markers in the Sedimentary Record*. Elsevier, Oxford, pp. 261–290.
- Tang, Y., Guo, W., Wang, X., Bao, H., Wu, H., 2019. A new breakthrough in exploration of large conglomerate oil province in Mahu sag and its implication. *Xinjiang Petroleum Geology* 40, 127–137 (in Chinese).
- Tao, K., Cao, J., Hu, W., Zhi, D., Lei, D., Tang, Y., Qu, J., Kang, X., Magoon, L.B., 2021. Petroleum system for the continuous oil play in the lacustrine Lower Triassic, Junggar Basin, China. *AAPG Bull.* 105, 2349–2380. <https://doi.org/10.1306/07022119211>.
- ten Haven, H.L., de Leeuw, J.W., Rullkötter, J., Sinninghe, J.S., 1987. Restricted utility of the pristane/phytane ratio as a paleoenvironmental indicator. *Nature* 330, 641–643. <https://doi.org/10.1038/330641a0>.
- van Aarsen, B.G.K., Bastow, T.P., Alexander, R., Kagi, R.I., 1999. Distributions of methylated naphthalenes in crude oils: indicators of maturity, biodegradation and mixing. *Org. Geochem.* 30, 1213–1227. [https://doi.org/10.1016/S0146-6380\(99\)00097-2](https://doi.org/10.1016/S0146-6380(99)00097-2).
- Wang, Y., Luo, J., Gao, Q., Xiang, Y., Lei, L., 2012. Natural gas reserves growth prediction and exploration potential analysis in Junggar Basin. *Xinjing Pet. Geol.* 33, 614–616 (in Chinese).
- Wang, X., Zhi, D., Wang, Y., Chen, J., Qin, Z., Liu, D., Xiang, Y., Lan, W., Li, N., 2013. *Source Rocks and Oil-Gas Geochemistry in Junggar Basin*. Chinese Petroleum Industry Press (in Chinese).
- Wang, Q., Huang, H., Li, Z., Ma, Y., Zeng, J., Larter, S., 2021. Geochemical significance of β -carotene in lacustrine oils from the shahejie formation of the dongying depression, eastern China. *Org. Geochem.* 156, 104241. <https://doi.org/10.1016/j.orggeochem.2021.104241>.
- Wang, Q., Huang, H., He, C., Li, Z., Zheng, L., 2022. Methylation and demethylation of naphthalene homologs in highly thermal mature sediments. *Org. Geochem.* 163, 104343. <https://doi.org/10.1016/j.orggeochem.2021.104343>.
- Wang, Y., Cao, J., Tao, K., Xiao, W., Xiang, B., Li, E., Pan, C., 2022. Absence of β -carotene as proxies of hydrothermal activity in brackish lacustrine sediments. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 587, 110801. <https://doi.org/10.1016/j.palaeo.2021.110801>.
- Wang, D.Y., Li, M.J., Zhou, Y., Yang, L., Yang, Y.F., Li, E.T., Jin, J., Zou, X.L., Xu, B.D., 2023. Petroleum geochemistry and origin of shallow-buried saline lacustrine oils in the slope zone of the Mahu sag, Junggar Basin, NW China. *Pet. Sci.* 20, 3363–3378. <https://doi.org/10.1016/j.petsci.2023.08.024>.
- Wilhelms, A., Larter, S.R., 2004. Shaken but not always stirred. Impact of petroleum charge mixing on reservoir geochemistry. In: Cubitt, J.M., England, W.A., Larter, S.R. (Eds.), *Understanding Petroleum Reservoirs: towards an Integrated Reservoir Engineering and Geochemical Approach*, vol. 237. Geological Society Special Publications, pp. 27–35.
- Yang, B., Ma, Q., Gan, Z., 1985. Thermal evolution of organic matter and petroleum generation threshold of Aican-1 well. *Oil Gas Geol.* 6 (4), 379–385 (in Chinese).
- Xiang, B., Li, E., Gao, X., Wang, M., Wang, Y., Xu, H., Huang, P., Yu, Y., Liu, J., Zou, Y., Pan, C., 2016. Petroleum generation kinetics for Permian lacustrine source rocks in the Junggar Basin, NW China. *Organic Geochemistry* 98, 1–17. <https://doi.org/10.1016/j.orggeochem.2016.05.003>.
- Yang, B., Jiang, Z., Li, J., Wang, X., 1992. *Origins of petroleum in northwestern Junggar Basin*. In: Luo, B. (Ed.), *Petroleum Geology of the Junggar Basin*. Gansu Science and Technology Press, pp. 62–73 (in Chinese).
- Yu, S., Wang, X., Xiang, B., Ren, J., Li, E., Wang, J., Huang, P., Wang, G., Xu, H., Pan, C., 2017. Molecular and carbon isotopic geochemistry of crude oils and extracts from Permian source rocks in the northwestern and central Junggar Basin, China. *Org. Geochem.* 113, 27–42. <https://doi.org/10.1016/j.orggeochem.2017.07.013>.
- Zhan, Z., Lin, X., Zou, Y., Li, Z., Wang, D., Liu, C., Peng, P., 2019. Chemometric differentiation of crude oil families in the southern dongying depression, bohai bay basin, China. *Org. Geochem.* 127, 37–49. <https://doi.org/10.1016/j.orggeochem.2018.11.004>.
- Zhang, G., Wang, Z., Wu, M., Wu, Q., Yang, B., Yang, W., Yang, R., Fan, G., Zheng, D., Zhao, B., Peng, X., Yong, T., 1993. *Petroleum geology of Junggar Basin*. In: Zhai, G. (Ed.), *Petroleum Geology of China*, vol. 15. Chinese Petroleum Industry Press (in Chinese).
- Zhang, Z., Gu, Y., Jin, J., Li, E., Yu, S., Pan, C., 2022. Assessing source and maturity of oils in the Mahu sag, Junggar Basin: molecular concentrations, compositions and carbon isotopes. *Mar. Petrol. Geol.* 141, 105724. <https://doi.org/10.1016/j.marpetgeo.2022.105724>.
- Zhao, B., 1992a. Nature of basement of Junggar Basin. *Xinjing Pet. Geol.* 13 (2), 95–99 (in Chinese).
- Zhao, B., 1992b. Formation and evolution of Junggar Basin. *Xinjing Pet. Geol.* 13 (3), 191–196 (in Chinese).
- Zhi, D., Tang, Y., Zheng, M., Guo, W., Wu, T., Zou, Z., 2018. Discovery, distribution and exploration practice of large oil provinces of above-source conglomerate in Mahu sag. *Xinjiang Petroleum Geology* 39, 1–8 (in Chinese).
- Zhi, D., Tang, Y., He, W., Guo, X., Zheng, M., Huang, L., 2021. Orderly coexistence and accumulation models of conventional and unconventional hydrocarbons in lower permian Fengcheng Formation, Mahu sag, Junggar Basin. *Petrol. Explor. Dev.* 48, 1–14. [https://doi.org/10.1016/S1876-3804\(21\)60004-6](https://doi.org/10.1016/S1876-3804(21)60004-6).
- Zhou, Z., Sheng, G., Sheng, R., Ming, Y., Lin, M., Zhang, H., Song, M., 1989. *Petroleum Geochemistry of Junggar Basin*. Chinese Science Press, China, pp. 1–74 (in Chinese).
- Zou, Y., Wang, Y., Zhan, Z., 2021. *Petroleum Geochemometrics-Foundation and Applications*. Chinese Science Press, pp. 1–330 (in Chinese).