



# Chemometrics reveals oil sources in the Fangzheng Fault Depression, NE China



Yao-Ping Wang<sup>a,b</sup>, Fan Zhang<sup>c</sup>, Yan-Rong Zou<sup>a,\*</sup>, Zhao-Wen Zhan<sup>a,b</sup>, Ping'an Peng<sup>a</sup>

<sup>a</sup> State Key Laboratory of Organic Geochemistry, Guangzhou Institute of Geochemistry, Chinese Academy of Sciences, Guangzhou 510640, PR China

<sup>b</sup> University of Chinese Academy of Sciences, Beijing 10039, PR China

<sup>c</sup> Exploration & Development Research Institute of Daqing Oilfield, PetroChina, Daqing 163712, PR China

## ARTICLE INFO

### Article history:

Received 1 June 2016

Received in revised form 22 September 2016

Accepted 23 September 2016

Available online 29 September 2016

### Keywords:

Fangzheng Fault Depression

Biomarkers

Oil–source rock correlation

Multidimensional scaling

Principal component analysis

## ABSTRACT

One hundred and twenty-five rock and oil samples collected from the Fangzheng Fault Depression were analyzed by GC–MS. Principal component analysis (PCA) and multidimensional scaling (MDS) based on nine biomarker ratios were used for oil–source rock correlation in this region. The oils are characterized by a predominance of low to middle molecular weight normal alkanes ( $n\text{-C}_{12}\text{-}n\text{-C}_{20}$  or  $n\text{-C}_{12}\text{-}n\text{-C}_{25}$ ), moderate to high Pr/Ph ratios (1.44–5.3), relatively low  $\text{C}_{27}/\text{C}_{29}$  regular steranes,  $\text{C}_{35}/\text{C}_{34}$  hopanes and gammacerane/ $\text{C}_{31}\text{R}$  ratios, all of which suggest relatively oxic conditions under fresh water depositional conditions with significant terrigenous organic matter input. PCA and MDS results show that the investigated oils were derived mainly from the Xinancun-Wuyun Formation and also illustrate the maturation and depositional conditions of the rocks and oils through an MDS plot. We show that MDS is a reliable, multi-parameter oil–source rock correlation method.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

The Fangzheng Fault Depression in northeast China (Fig. 1a), is located in the northern section of the Yishu Graben, and is controlled by two deep northwest-trending faults (Fig. 1b). The Yilan Fault Uplift and the Tanguan Fault Depression lie to the north of the Fangzheng Fault Depression, while the Shangzhi Fault Uplift, the Shengli Fault Depression, and the Shulan Fault Uplift lie to the south (Hu et al., 2010a). Fig. 1c shows the location of the Fangzheng Fault Depression (longitude:  $128^{\circ}15'\text{--}129^{\circ}15'$ ; latitude:  $45^{\circ}25'\text{--}46^{\circ}10'$ ), and the borehole distribution in the study area (Li et al., 2015) is shown in Fig. 1d.

The Fangzheng Fault Depression covers an area of about 1460 km<sup>2</sup> (Liu et al., 2014; Yang et al., 2014) and constitutes four sedimentary sequences on Paleozoic metamorphic basement (Dong et al., 2008; Hu et al., 2010b; Shao et al., 2013). From the bottom to the top these are: Muling (K<sub>2</sub>m), Wuyun (E<sub>1</sub>w), Xinancun (E<sub>2</sub>x), Dalianhe (E<sub>2</sub>d), Baoquanling (E<sub>3</sub>b) and Fujin (N<sub>1</sub>f) formations, with a total thickness of over 5000 m (Fig. 2). These are overlapped by Quaternary gravel and sandy clay having a thickness of 25–85 m. The Xinancun Formation and Wuyun Formation cannot be distinguished in the study area and are thus called the Xinancun-Wuyun Formation for convenience.

The Fangzheng Fault Depression is an important oil and gas-bearing depression and there has been considerable attention paid to this region in recent years. More than 20 wells have been drilled in the study area (Fig. 1d). Five commercial oil wells, two commercial gas wells and five wells with a low yield of oil were completed in the Xinancun-Wuyun Formation, Muling Formation and basement (Fu et al., 2014). However, the genetic relationships of the discovered crude oils are discussed only to a limited extent in the literature (He et al., 2011; Zhang, 2012) and oil–source correlations remain controversial.

The Xinancun-Wuyun Formation source rocks are very similar to those of the Muling Formation in terms of depositional conditions and geochemical character. This may explain why conventional analyses do not demonstrate clear genetic relationships between the oils and candidate source rocks. In addition, controversy over oil–source rock correlations in the study area may be the result of limited samples and the use of only a few biomarker parameters (He et al., 2011; Zhang, 2012).

In this paper, principal component analysis (PCA) was used to identify the affinities of oils and source rocks, and multidimensional scaling (MDS) based on multivariate statistical analysis was introduced for oil–source rock correlation. MDS has already been used to evaluate data from many disciplines (Shi et al., 2000; Lenz and Foran, 2010; Hollemeyer et al., 2012), including biomarker data (Revill et al., 1992; Zhou et al., 2015). Classical MDS generally uses the Euclidean distance to distinguish data sets,

\* Corresponding author. Fax: +86 20 85290706.

E-mail address: [zouyr@gig.ac.cn](mailto:zouyr@gig.ac.cn) (Y.-R. Zou).

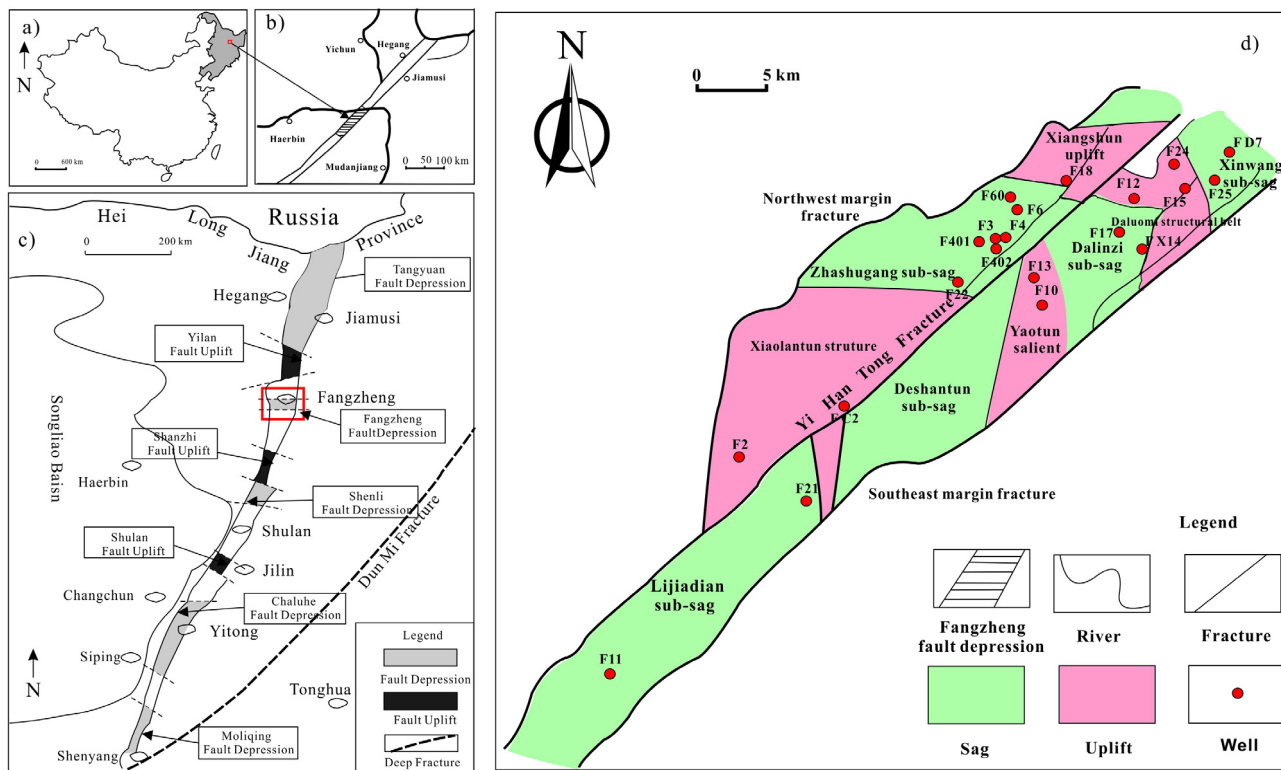


Fig. 1. Map showing location of: (a) NE China; (b) Tanlu Faults Belt; (c) the Fangzheng Fault Depression; and (d) wells in the study area (modified after Hu et al., 2010b; Li et al., 2015).

but this is not suitable as a dissimilarity measure for all types of data (McGee, 1968), depending on the study objectives. In this study, four distances were checked to find out the suitable distance measure for biomarker data. Then nine biomarker ratios frequently used for oil–source rock correlation were employed, projecting them onto a 2D plane using MDS dimension reduction in order to recognize genetic relationships between the oils and the source rocks.

## 2. Samples and methods

### 2.1. Samples

The boreholes were drilled with a water-based mud system. One hundred and thirteen core samples (mudstones) and 12 crude oil samples were collected in the present study (Table 1). All candidate source rocks from this depression were collected from 17 wells. One hundred and three core samples were collected from the Xinancun-Wuyun Formation, and 10 samples from the Muling Formation. All core samples were cleaned of surface impurities using a knife, washed with redistilled water, dried at 60 °C and then ground into powder. The 12 oils from DSTs were collected from eight different wells within the Xinancun-Wuyun and the Muling formations.

### 2.2. GC–MS analysis

In order to investigate possible oil–source genetic relationships, 113 source rock samples from the Xinancun-Wuyun and Muling Formations were selected, cleaned and then ground to powder. The powders were extracted for 72 h using Soxhlet apparatus and dichloromethane. The group compositions were separated using column chromatography for both oils and source rock

extracts. GC–MS analyses were performed on saturated hydrocarbon fractions using a Thermo Fisher Trace GC Ultra gas chromatograph equipped with an Agilent HP-5MS capillary column (60 m × 0.25 mm; 0.25 μm film thickness), coupled to a DSQ II mass spectrometer. Helium was used as the carrier gas. The oven temperature was initially kept at 60 °C for 1 min, and then programmed to 220 °C at a rate of 8 °C/min and then to 300 °C at a rate of 2 °C/min, and held isothermally for 25 min.

The mass spectrometer was operated in EI mode at 70 eV and the ion source temperature was 230 °C. The analysis was conducted using mode-combining selective ion monitoring (SIM) with full-scan detection in the scan range from 50 to 550 Da.

### 2.3. Computational methods

Principal component analysis is often used in petroleum geochemistry, but multidimensional scaling is less frequently used, so a brief introduction is given here.

#### 2.3.1. Brief explanation of multidimensional scaling

MDS is a statistical analysis method which represents similarity (or dissimilarity) measurements as distances between object points in a high-dimensional space. By means of MDS, the data based on multiple dimensions can be simplified to a 2D graph that displays correlations (Borg and Groenen, 2005). MDS adjusts the objects in low-dimensions in order to best depict distances in high-dimensional space and checks the goodness-of-fit between objects in low-dimensional space. Stress is a measure of goodness-of-fit, which is expressed as:

$$Stress = \sqrt{\frac{\sum (d_{ij} - d_{ij}^*)^2}{\sum d_{ij}^2}}$$

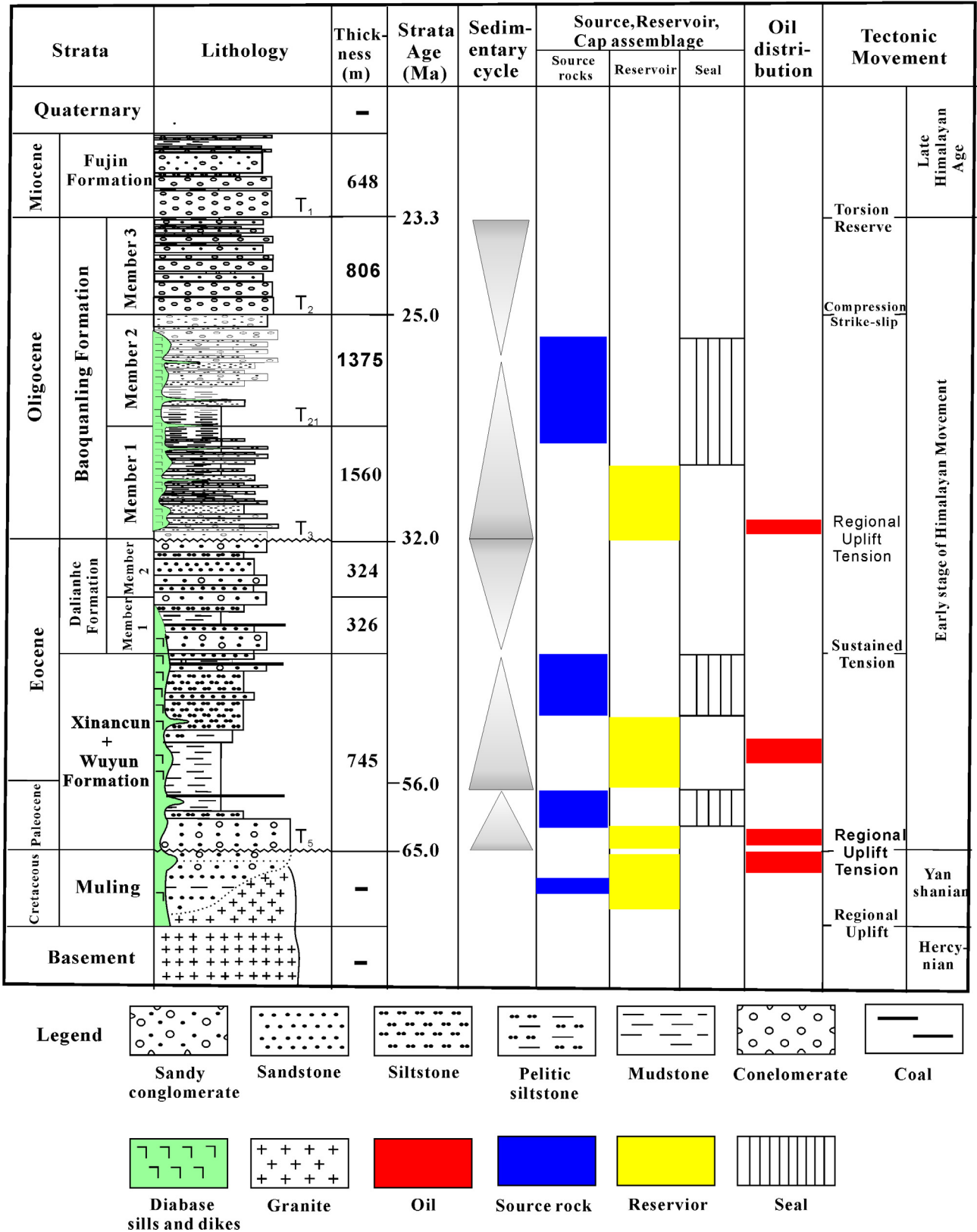


Fig. 2. General stratigraphic successions for the Fangzheng Fault Depression (modified after Fu et al., 2014; Chen et al., 2015).

Here  $d_{ij}$  is the distance between objects  $i$  and  $j$  in high-dimensional space and  $d^*_{ij}$  is the distance between objects  $i$  and  $j$  in low-dimensional space. Stress varies between 0 and 1, with values near 0 indicating a better fit. In Kruskal (1964), stress values were given

based on his experience (Table 2), and these are still used to indicate the goodness-of-fit of an MDS solution (Storti, 2016).

MDS is a graphical method to represent the similarity of data in low-dimensional space as simplified from high-dimensional space

**Table 1**  
Biomarker ratios of oils and source rocks collected from the Fangzheng Fault Depression.

No.	Strata	Lab no.	Well	Depth (m)	R1#	R2	R3	R4#	R5#	R6#	R7#	R8#	R9	R10	R11#	R12	R13#	R14#	R15	R16	R17	MDS results	
																						MDS-1	MDS-2
<i>Source rocks</i>																							
1	XW	XW-1	F3	2962	1.49	1.30	0.72	0.09	0.04	0.53	0.25	0.07	0.53	0.35	0.15	0.28	0.29	0.45	0.17	0.26	0.58	-0.071	-0.197
2	XW	XW-2	F3	2969	0.95	1.65	0.96	0.10	0.04	0.52	0.29	0.08	0.53	0.37	0.15	0.29	0.32	0.46	0.18	0.26	0.56	-0.106	-0.215
3	XW	XW-3	F4	2874	0.80	1.58	4.62	0.06	0.06	0.52	0.20	0.02	0.09	0.17	0.07	0.15	0.26	0.24	0.17	0.16	0.67	0.208	-0.264
4	XW	XW-4	F4	2886	0.86	0.36	0.32	0.12	0.05	0.48	0.26	0.08	0.47	0.27	0.15	0.23	0.67	0.49	0.31	0.23	0.46	-0.112	-0.185
5	XW	XW-5	F4	2900	1.26	1.32	0.63	0.04	2.11	1.09	0.12	0.03	0.06	0.19	0.13	0.15	0.68	0.30	0.35	0.15	0.50	0.283	-0.150
6	XW	XW-6	F4	2928	1.39	0.93	0.52	0.05	2.26	1.03	0.14	0.01	0.16	0.18	0.18	0.24	0.55	0.30	0.30	0.16	0.54	0.303	-0.198
7	XW	XW-7	F4	2986	1.14	0.76	0.52	0.33	0.08	0.44	0.46	0.04	0.50	0.39	0.21	0.38	1.03	0.69	0.38	0.25	0.37	-0.185	-0.080
8	XW	XW-8*	F4	3006	0.79	0.31	0.32	0.43	0.08	0.45	0.46	0.09	0.49	0.46	0.27	0.52	1.31	0.94	0.4	0.29	0.31	-0.280	-0.031
9	XW	XW-9	F4	3058	1.14	1.02	2.32	0.46	0.09	0.19	0.43	0.06	0.58	0.37	0.35	0.36	1.20	0.76	0.41	0.26	0.34	-0.265	-0.048
10	XW	XW-10*	F4	3094	0.43	0.42	0.31	0.25	0.09	0.55	0.35	0.04	0.51	0.41	0.20	0.42	1.09	0.71	0.39	0.25	0.36	-0.188	-0.139
11	XW	XW-11	F4	3110	2.50	1.74	0.40	0.08	0.13	0.74	0.14	0.01	0.38	0.19	0.11	0.28	0.63	0.48	0.30	0.23	0.47	0.152	-0.076
12	XW	XW-12	F4	3126	1.97	0.87	0.36	0.09	0.08	0.66	0.18	0.04	0.27	0.23	0.06	0.33	0.56	0.36	0.31	0.16	0.54	0.075	-0.074
13	XW	XW-13	F4	3129	0.77	1.15	3.73	0.11	0.07	0.62	0.26	0.02	0.48	0.27	0.10	0.34	0.85	0.66	0.34	0.26	0.40	-0.028	-0.234
14	XW	XW-14	F4	3140	3.75	1.68	0.31	0.08	0.13	0.69	0.21	0.02	0.31	0.21	0.07	0.37	0.50	0.35	0.27	0.19	0.54	0.147	-0.014
15	XW	XW-15	F4	3150	0.80	1.25	4.86	0.10	0.12	0.72	0.18	0.02	0.18	0.20	0.07	0.33	0.60	0.33	0.31	0.17	0.52	0.116	-0.190
16	XW	XW-16	F4	3160	2.77	1.22	0.31	0.18	0.20	0.62	0.21	0.04	0.27	0.23	0.09	0.31	0.63	0.47	0.30	0.22	0.48	0.005	-0.045
17	XW	XW-17	F4	3169	2.45	1.11	0.24	0.13	0.25	0.75	0.24	0.02	0.17	0.22	0.08	0.32	0.60	0.36	0.31	0.18	0.51	0.101	-0.064
18	XW	XW-18	F4	3175	2.36	0.95	0.36	0.22	0.14	0.62	0.25	0.04	0.43	0.32	0.10	0.38	0.54	0.53	0.36	0.23	0.41	-0.023	-0.044
19	XW	XW-19	F4	3176	2.87	1.53	0.51	0.11	0.21	0.92	0.16	0.02	0.18	0.24	0.05	0.38	0.41	0.33	0.23	0.19	0.58	0.175	-0.023
20	XW	XW-20	F4	3182	1.93	0.97	0.56	0.10	0.23	0.87	0.26	0.02	0.30	0.21	0.04	0.37	0.32	0.32	0.20	0.20	0.61	0.207	-0.081
21	XW	XW-21	F4	3189	2.83	1.11	0.31	0.12	0.15	0.68	0.20	0.03	0.27	0.23	0.06	0.36	0.61	0.37	0.31	0.19	0.50	0.083	-0.033
22	XW	XW-22	F4	3190	3.65	2.09	0.42	0.18	0.15	0.67	0.32	0.01	0.38	0.30	0.09	0.39	0.77	0.46	0.34	0.21	0.45	0.087	0.011
23	XW	XW-23	F4	3192	3.06	0.87	0.25	0.22	0.09	0.82	0.00	0.03	0.18	0.25	0.04	0.41	0.68	0.33	0.34	0.16	0.50	0.110	0.139
24	XW	XW-24	F4	3198	2.84	0.97	0.29	0.11	0.08	0.75	0.18	0.03	0.12	0.25	0.05	0.39	0.70	0.34	0.34	0.17	0.49	0.108	-0.042
25	XW	XW-25*	F4	3205	3.49	2.48	0.48	0.33	0.10	0.59	0.37	0.01	0.49	0.41	0.24	0.53	1.51	0.81	0.46	0.24	0.30	-0.162	0.100
26	XW	XW-26	F4	3205	2.00	1.04	0.49	0.08	3.51	0.68	0.23	0.03	0.25	0.20	0.08	0.38	0.55	0.38	0.29	0.20	0.52	0.233	-0.035
27	XW	XW-27*	F4	3217	1.56	0.28	0.18	0.55	0.24	0.50	0.28	0.03	0.51	0.52	0.16	0.52	0.47	0.38	0.26	0.20	0.54	-0.073	0.004
28	XW	XW-28*	F4	3221	1.91	0.92	0.33	0.42	0.10	0.51	0.43	0.07	0.55	0.44	0.26	0.55	1.37	0.98	0.41	0.29	0.30	-0.233	0.004
29	XW	XW-29*	F4	3226	1.76	0.38	0.20	0.52	0.23	0.50	0.26	0.06	0.55	0.51	0.17	0.48	0.55	0.38	0.29	0.20	0.52	-0.107	-0.003
30	XW	XW-30*	F4	3229	1.8	1.03	0.38	0.43	0.09	0.51	0.46	0.08	0.56	0.42	0.28	0.56	1.47	0.98	0.43	0.28	0.29	-0.250	0.007
31	XW	XW-31*	F4	3235	1.47	0.27	0.18	0.52	0.23	0.49	0.28	0.04	0.51	0.51	0.15	0.48	0.48	0.38	0.26	0.20	0.54	-0.083	-0.005
32	XW	XW-32*	F6	2760	0.62	0.50	0.78	0.38	0.06	0.55	0.43	0.04	0.52	0.44	0.99	0.48	1.05	0.86	0.36	0.30	0.34	-0.278	-0.130
33	XW	XW-33	F10	2860	1.71	1.68	0.43	0.16	0.05	0.77	0.21	0.02	0.42	0.26	0.07	0.37	0.60	0.51	0.28	0.24	0.47	0.052	-0.131
34	XW	XW-34	F10	3208	2.49	0.74	0.09	0.22	0.19	0.67	0.33	0.02	0.30	0.32	0.17	0.36	0.31	0.31	0.19	0.19	0.61	0.094	0.033
35	XW	XW-35*	F10	3216	3.71	0.88	0.15	0.27	0.06	0.82	0.29	0.08	0.44	0.41	0.04	0.52	0.81	0.50	0.35	0.22	0.43	-0.058	0.050
36	XW	XW-36*	F10	3217	4.41	0.84	0.14	0.29	0.06	0.71	0.32	0.04	0.37	0.47	0.05	0.59	0.62	0.42	0.31	0.20	0.49	0.009	0.043
37	XW	XW-37*	F10	3225	2.07	0.74	0.27	0.38	0.07	0.55	0.28	0.07	0.50	0.44	0.18	0.49	1.19	0.78	0.40	0.26	0.34	-0.186	-0.022
38	XW	XW-38*	F10	3233	2.15	1.05	0.28	0.41	0.08	0.57	0.29	0.09	0.52	0.43	0.19	0.45	1.02	0.69	0.38	0.26	0.37	-0.184	-0.007
39	XW	XW-39	F11	4294	2.81	1.45	0.38	0.19	0.07	0.90	0.23	0.04	0.43	0.29	0.05	0.44	0.72	0.38	0.34	0.18	0.48	0.042	-0.029
40	XW	XW-40	F11	4295	2.65	1.37	0.42	0.22	0.06	0.79	0.15	0.04	0.25	0.28	0.06	0.45	0.80	0.44	0.36	0.20	0.45	0.010	-0.031
41	XW	XW-41	F11	4296	2.90	1.57	0.39	0.20	0.05	0.81	0.16	0.07	0.24	0.29	0.06	0.45	0.87	0.44	0.38	0.19	0.43	-0.015	-0.003
42	XW	XW-42	F11	4296	0.79	0.98	1.16	0.28	0.07	0.64	0.40	0.01	0.51	0.33	0.15	0.30	0.69	0.49	0.32	0.23	0.46	-0.055	-0.191
43	XW	XW-43	F11	4296	2.91	1.62	0.41	0.21	0.06	0.79	0.17	0.05	0.24	0.28	0.06	0.43	0.84	0.36	0.38	0.16	0.45	0.024	-0.009
44	XW	XW-44	F11	4297	2.69	1.77	0.47	0.21	0.06	0.84	0.16	0.04	0.23	0.29	0.06	0.45	0.84	0.42	0.37	0.19	0.44	0.016	-0.022
45	XW	XW-45	F11	4298	3.14	1.84	0.41	0.18	0.05	0.84	0.16	0.04	0.24	0.29	0.06	0.46	0.86	0.41	0.38	0.18	0.44	0.039	0.001
46	XW	XW-46	F11	4299	0.64	1.55	4.19	0.16	0.06	0.83	0.19	0.03	0.23	0.28	0.06	0.44	0.73	0.34	0.35	0.17	0.48	0.019	-0.182
47	XW	XW-47	F11	4299	3.14	2.11	0.50	0.17	0.07	0.89	0.30	0.04	0.23	0.28	0.07	0.46	0.78	0.41	0.36	0.19	0.46	0.046	-0.006
48	XW	XW-48	F11	4300	2.71	2.12	0.54	0.19	0.08	0.84	0.14	0.04	0.28	0.30	0.07	0.46	0.77	0.38	0.36	0.18	0.46	0.052	-0.026
49	XW	XW-49	F11	4301	2.79	1.96	0.48	0.21	0.08	0.81	0.16	0.04	0.24	0.28	0.06	0.45	0.76	0.4	0.35	0.18	0.46	0.026	-0.026
50	XW	XW-50	F11	4301	2.74	1.92	0.51	0.24	0.07	0.78	0.17	0.05	0.23	0.28	0.07	0.43	0.79	0.38	0.36	0.17	0.46	0.002	-0.024
51	XW	XW-51	F11	4301	1.23	1.03	0.61	0.35	0.13	0.65	0.30	0.02	0.48	0.36	0.12	0.40	0.71	0.46	0.33	0.21	0.46	-0.034	-0.108
52	XW	XW-52	F11	4531	6.00	2.76	0.27	0.40	0.23	0.56	0.25	0.05	0.34	0.41	0.09	0.45	0.58	0.34	0.30	0.18	0.52	-0.013	0.094

53	XW	XW-53	F12	1815	4.06	0.87	0.17	0.08	0.02	1.34	0.35	0.01	0.43	0.21	0.02	0.46	0.94	0.49	0.39	0.20	0.41	0.230	0.063
54	XW	XW-54*	F12	1828	5.40	0.83	0.11	0.22	0.12	0.57	0.35	0.02	0.51	0.43	0.10	0.46	0.27	0.29	0.17	0.18	0.64	0.109	0.098
55	XW	XW-55	F12	1882	1.20	1.50	1.14	0.10	0.04	0.44	0.32	0.02	0.25	0.25	0.10	0.29	0.50	0.36	0.27	0.19	0.54	0.068	-0.205
56	XW	XW-56	F12	1956	1.39	1.28	1.01	0.16	0.04	0.46	0.30	0.01	0.31	0.28	0.12	0.30	0.63	0.41	0.31	0.20	0.49	0.031	-0.210
57	XW	XW-57	F13	3160	1.31	0.92	0.72	0.14	0.03	0.59	0.16	0.01	0.32	0.23	0.10	0.30	0.71	0.41	0.33	0.19	0.47	0.059	-0.227
58	XW	XW-58	F13	3210	1.71	0.95	0.43	0.14	0.04	0.63	0.22	0.01	0.25	0.19	0.08	0.34	0.55	0.32	0.29	0.17	0.54	0.140	-0.157
59	XW	XW-59	F13	3241	1.21	0.91	0.55	0.16	0.04	0.69	0.24	0.01	0.29	0.22	0.09	0.38	0.63	0.34	0.32	0.17	0.51	0.101	-0.193
60	XW	XW-60	F13	3340	2.21	1.18	0.51	0.16	0.05	0.72	0.19	0.01	0.26	0.22	0.07	0.4	0.56	0.33	0.29	0.18	0.53	0.152	-0.110
61	XW	XW-61	F13	3386	1.74	1.69	0.56	0.11	0.06	0.81	0.21	0.02	0.27	0.20	0.08	0.41	0.49	0.30	0.27	0.17	0.56	0.137	-0.130
62	XW	XW-62	F13	3410	2.15	1.45	0.57	0.12	0.04	0.71	0.23	0.01	0.25	0.22	0.08	0.37	0.52	0.32	0.28	0.17	0.54	0.145	-0.115
63	XW	XW-63	F13	3510	1.59	0.89	0.53	0.19	0.06	0.77	0.23	0.01	0.28	0.25	0.08	0.42	0.54	0.32	0.29	0.17	0.54	0.101	-0.127
64	XW	XW-64	F13	3630	1.37	0.64	0.5	0.18	0.06	0.83	0.24	0.01	0.26	0.23	0.07	0.46	0.44	0.29	0.26	0.17	0.58	0.162	-0.162
65	XW	XW-65	F13	3796	0.82	1.15	0.45	0.22	0.09	0.88	0.24	0.02	0.34	0.31	0.08	0.48	0.59	0.34	0.31	0.17	0.52	0.055	-0.172
66	XW	XW-66*	F13	3831	1.54	1.12	0.36	0.57	0.22	0.47	0.29	0.03	0.39	0.51	0.14	0.51	0.67	0.53	0.30	0.24	0.46	-0.108	-0.027
67	XW	XW-67*	F15	1839	3.06	0.61	0.15	0.19	0.11	0.00	0.31	0.03	0.38	0.44	0.15	0.4	0.27	0.29	0.17	0.19	0.64	0.037	0.206
68	XW	XW-68	F15	1854	1.22	1.50	0.96	0.11	0.05	0.51	0.27	0.02	0.17	0.17	0.11	0.23	0.57	0.41	0.29	0.21	0.50	0.028	-0.160
69	XW	XW-69	F15	1926	0.35	0.75	0.54	0.12	0.06	0.78	0.23	0.02	0.20	0.18	0.07	0.24	0.47	0.34	0.26	0.19	0.55	0.106	-0.234
70	XW	XW-70	F15	2020	1.50	1.18	0.40	0.14	0.08	0.50	0.25	0.02	0.22	0.21	0.09	0.24	0.35	0.38	0.20	0.22	0.58	0.066	-0.120
71	XW	XW-71	F15	2064	1.61	1.29	0.80	0.14	0.07	0.38	0.27	0.03	0.17	0.23	0.10	0.23	0.38	0.39	0.22	0.22	0.57	0.034	-0.116
72	XW	XW-72	F15	2096	1.35	1.47	0.73	0.22	0.09	0.86	0.20	0.03	0.34	0.31	0.08	0.47	0.61	0.34	0.31	0.17	0.51	0.036	-0.099
73	XW	XW-73	F15	2143	3.45	1.11	0.25	0.17	0.22	0.00	0.28	0.02	0.21	0.28	0.28	0.37	0.45	0.32	0.25	0.18	0.56	0.077	0.203
74	XW	XW-74	F15	2190	2.67	0.53	0.12	0.15	0.07	0.61	0.19	0.02	0.25	0.25	0.10	0.27	0.38	0.34	0.22	0.20	0.58	0.122	-0.051
75	XW	XW-75	F15	2273	2.28	1.15	0.37	0.10	0.08	0.53	0.24	0.02	0.16	0.19	0.10	0.26	0.42	0.36	0.23	0.20	0.56	0.118	-0.087
76	XW	XW-76	F15	2460	3.97	1.86	0.28	0.12	0.10	0.89	0.29	0.04	0.31	0.20	0.12	0.39	0.27	0.29	0.17	0.19	0.64	0.108	0.071
77	XW	XW-77	F15	2505	4.06	1.31	0.24	0.14	0.09	0.88	0.31	0.01	0.36	0.23	0.12	0.38	0.22	0.33	0.14	0.21	0.64	0.171	0.038
78	XW	XW-78	F15	2677	3.65	2.03	0.27	0.06	0.12	1.33	0.19	0.04	0.24	0.27	0.07	0.43	0.09	0.23	0.07	0.17	0.76	0.200	0.155
79	XW	XW-79	F15	2712	3.12	2.34	0.31	0.04	0.10	1.25	0.22	0.04	0.26	0.30	0.08	0.44	0.11	0.21	0.09	0.16	0.76	0.225	0.130
80	XW	XW-80	F16	2578	1.72	0.40	0.14	0.19	0.04	0.71	0.71	0.02	0.37	0.26	0.09	0.34	0.48	0.33	0.26	0.18	0.55	0.098	-0.153
81	XW	XW-81	F17	2444	1.16	1.24	0.62	0.18	0.04	0.74	0.27	0.02	0.19	0.25	0.07	0.41	0.71	0.29	0.35	0.15	0.50	0.070	-0.158
82	XW	XW-82	F17	2560	1.78	1.21	0.44	0.21	0.04	0.68	0.28	0.02	0.14	0.29	0.07	0.36	0.60	0.32	0.31	0.16	0.52	0.067	-0.099
83	XW	XW-83	F18	2324	0.73	1.39	0.85	0.11	0.06	0.39	0.28	0.01	0.20	0.30	0.14	0.24	0.54	0.33	0.29	0.17	0.53	0.054	-0.272
84	XW	XW-84	F18	2402	0.47	2.00	0.83	0.06	0.03	0.74	0.21	0.02	0.13	0.21	0.09	0.27	0.71	0.30	0.35	0.15	0.50	0.111	-0.269
85	XW	XW-85	F18	2421	1.09	1.46	0.63	0.06	0.02	0.76	0.20	0.01	0.12	0.19	0.08	0.31	0.70	0.31	0.35	0.15	0.50	0.153	-0.221
86	XW	XW-86	F21	3210	0.89	0.50	0.52	0.40	0.08	0.53	0.34	0.05	0.30	0.19	0.26	0.11	0.32	0.35	0.19	0.21	0.60	-0.148	-0.065
87	XW	XW-87	F21	3430	1.21	1.22	1.00	0.26	0.15	0.66	0.26	0.05	0.34	0.21	0.25	0.22	0.34	0.38	0.20	0.22	0.58	-0.058	-0.065
88	XW	XW-88	F21	3510	1.37	1.22	0.86	0.16	0.11	0.66	0.21	0.03	0.22	0.16	0.27	0.20	0.29	0.26	0.18	0.17	0.65	0.172	-0.071
89	XW	XW-89	F21	3740	1.26	0.94	0.90	0.38	0.04	0.63	0.17	0.05	0.32	0.24	0.04	0.36	0.73	0.39	0.34	0.18	0.47	-0.065	-0.109
90	XW	XW-90	F21	3786	1.43	1.32	2.06	0.37	0.06	0.63	0.21	0.06	0.27	0.23	0.08	0.31	0.73	0.51	0.33	0.23	0.45	-0.092	-0.090
91	XW	XW-91	F21	3816	1.11	0.36	0.38	0.33	0.06	0.61	0.19	0.07	0.26	0.19	0.06	0.33	0.69	0.42	0.33	0.20	0.47	-0.092	-0.121
92	XW	XW-92	F21	3850	1.20	0.46	0.47	0.36	0.05	0.62	0.20	0.06	0.27	0.24	0.07	0.34	0.82	0.53	0.35	0.23	0.43	-0.104	-0.103
93	XW	XW-93	F21	3925	0.71	0.21	0.22	0.35	0.06	0.8	0.19	0.06	0.28	0.24	0.06	0.37	0.65	0.45	0.31	0.21	0.48	-0.092	-0.143
94	XW	XW-94	F21	3972	0.83	0.32	0.34	0.39	0.05	0.72	0.23	0.05	0.30	0.25	0.06	0.36	0.62	0.46	0.30	0.22	0.48	-0.080	-0.125
95	XW	XW-95	F21	4022	1.03	0.34	0.28	0.31	0.06	0.70	0.22	0.05	0.28	0.22	0.05	0.36	0.66	0.39	0.32	0.19	0.49	-0.048	-0.115
96	XW	XW-96	F21	4126	0.91	0.37	0.36	0.50	0.06	0.70	0.25	0.07	0.38	0.35	0.09	0.40	0.73	0.54	0.32	0.24	0.44	-0.142	-0.104
97	XW	XW-97	F21	4158	1.00	0.68	0.85	0.55	0.06	0.65	0.28	0.08	0.41	0.37	0.09	0.38	0.65	0.61	0.29	0.27	0.44	-0.166	-0.105
98	XW	XW-98	F22	3086	1.58	0.87	0.44	0.10	0.14	0.76	0.11	0.04	0.23	0.23	0.08	0.19	0.61	0.50	0.29	0.24	0.47	0.000	-0.139
99	XW	XW-99	FD7	1466	5.29	4.22	0.36	0.15	0.07	0.87	0.31	0.01	0.35	0.31	0.14	0.43	0.61	0.38	0.30	0.19	0.50	0.150	0.028
100	XW	XW-100	FD7	1472	4.73	2.51	0.28	0.04	0.03	1.30	0.25	0.04	0.21	0.21	0.06	0.43	0.22	0.19	0.15	0.13	0.71	0.198	0.184
101	XW	XW-101	FX14	3442	5.00	3.27	0.33	0.12	0.10	0.67	0.29	0.01	0.50	0.64	0.06	0.60	0.21	0.23	0.15	0.16	0.69	0.278	0.048
102	XW	XW-102*	FX14	3442	3.19	0.87	0.16	0.27	0.11	0.60	0.42	0.03	0.36	0.65	0.07	0.66	0.58	0.37	0.30	0.19	0.51	0.027	0.027
103	XW	XW-103*	F402	3286	2.13	0.40	0.09	0.51	0.13	0.50	0.22	0.05	0.53	0.44	0.19	0.46	0.97	0.51	0.39	0.21	0.40	-0.132	-0.012
104	ML	ML-1	FC2	3036	0.68	0.36	0.35	0.3	0.15	0.40	0.71	0.01	0.24	0.18	0.49	0.16	0.79	0.68	0.32	0.28	0.40	-0.218	-0.240
105	ML	ML-2*	F4	3273	3.24	1.72	0.38	0.17	0.14	0.70	0.32	0.08	0.45	0.30	0.11	0.39	0.82	0.53	0.35	0.22	0.43	-0.051	0.014
106	ML	ML-3*	F4	3295	2.88	1.23	0.32	0.15	0.15	0.73	0.26	0.01	0.47	0.32	0.09	0.42	0.84	0.4	0.37	0.18	0.45	0.114	-0.029
107	ML	ML-4	F402	3349	0.88	0.13	0.11	0.42	0.15	0.51	0.46	0.04	0.41	0.32	0.19	0.21	2.08	0.75	0.54	0.20	0.26	-0.241	-0.065
108	ML	ML-5*	F402	3367	0.41	0.35	0.09	0.51	0.13	0.5	0.41	0.07	0.53	0.44	0.19	0.45	0.96	0.51	0.39	0.21	0.40	-0.216	-0.081
109	ML	ML-6*	F601	2788	1.80	0.44	0.19	0.63	0.18	0.39	0.65	0.04	0.61	0.53	0.13	0.51	0.44	0.33	0.25	0.19	0.57	-0.114	0.094
110	ML	ML-7*	F16</																				

Table 1 (continued)

No.	Strata	Lab no.	Well	Depth (m)	R1#	R2	R3	R4#	R5#	R6#	R7#	R8#	R9	R10	R11#	R12	R13#	R14#	R15	R16	R17	MDS results	
																						MDS-1	MDS-2
112	ML	ML-9	F10	3460	3.79	0.24	0.06	0.80	0.29	0.58	0.33	0.04	0.32	0.22	0.13	0.25	0.85	0.52	0.36	0.22	0.42	-0.107	0.079
113	ML	ML-10*	F17	3068	0.42	1.02	0.45	0.26	0.05	0.64	0.32	0.04	0.45	0.33	0.12	0.36	0.87	0.52	0.36	0.22	0.42	-0.119	-0.151
Crude oils																							
114	ML	ML-354	F6	3012	2.17	0.44	0.17	0.50	0.20	0.48	0.32	0.05	0.61	0.50	0.14	0.44	0.48	0.49	0.24	0.25	0.51	-0.103	0.014
115	ML	ML-427	F16	3211	2.67	0.30	0.12	0.59	0.13	0.51	0.28	0.07	0.49	0.50	0.16	0.53	0.65	0.45	0.31	0.22	0.47	-0.124	0.017
116	ML	ML-176	F12	1975	5.39	0.67	0.11	0.35	0.11	0.55	0.44	0.04	0.53	0.50	0.17	0.46	0.37	0.35	0.21	0.21	0.58	-0.021	0.111
117	ML	ML-198	F12	1975	3.87	0.59	0.17	0.32	0.13	0.65	(0.26)	0.04	0.56	0.55	0.20	0.46	0.34	0.37	0.20	0.22	0.59	0.003	0.068
118	ML	ML-389	F12	1975	4.16	0.64	0.15	0.24	0.14	0.59	0.27	0.03	0.46	0.53	0.12	0.40	0.24	0.26	0.16	0.17	0.67	0.069	0.100
119	XW	XW-916	F4	3214	2.50	0.37	0.15	0.55	0.23	0.51	(0.26)	0.03	0.51	0.51	0.16	0.50	0.51	0.37	0.27	0.19	0.53	-0.051	0.034
120	XW	XW-917	F4	3220	2.23	0.37	0.16	0.52	0.20	0.48	(0.26)	0.03	0.54	0.68	0.18	0.53	0.68	0.54	0.49	0.20	0.30	-0.167	0.038
121	XW	XW-1175	F6	2995	4.06	0.47	0.11	0.49	0.19	0.43	0.40	0.03	0.56	0.49	0.16	0.41	0.58	0.43	0.29	0.21	0.50	-0.062	0.077
122	XW	XW-157	F6	2995	2.06	0.34	0.16	0.79	0.18	0.47	0.32	0.04	0.52	0.53	0.13	0.46	0.55	0.41	0.28	0.21	0.51	-0.096	0.045
123	XW	XW-211	F402	3276	2.52	0.42	0.18	0.59	0.31	1.97	0.36	0.04	0.49	0.50	0.22	0.50	0.42	0.36	0.24	0.20	0.56	-0.072	0.128
124	XW	XW-212	F403	3328	1.44	0.18	0.13	0.48	0.11	0.63	0.43	0.06	0.57	0.49	0.21	0.46	0.62	0.45	0.30	0.22	0.48	-0.134	-0.043
125	XW	XW-390	F602	3129	2.27	0.37	0.16	0.38	0.23	0.54	0.30	0.04	0.38	0.45	0.10	0.33	0.19	0.40	0.12	0.25	0.63	-0.035	0.065

Note: XW = Xinancun + Wuyuan formations, ML = Muling Formation; "\*" samples were selected to oil-source rock correlation with normal cross plots; "#" biomarker parameters for PCA and MDS calculation; R1 = Pr/Ph; R2 = Pr/h-C<sub>17</sub>; R3 = Ph/h-C<sub>18</sub>; R4 = Ts/(Tm + Ts); R5 = Ol/H30; R6 = H29/H30; R7 = C<sub>25</sub>/C<sub>24</sub>; R8 = GA/C<sub>27</sub>; R9 = C<sub>29</sub>β/(αα + ββ); R10 = St/H; R11 = H<sub>32</sub> 22S/(22S + 22R); R12 = C<sub>29</sub> 20S/(20R + 20S); R13 = C<sub>29</sub> C<sub>27</sub>/C<sub>29</sub>; R14 = C<sub>28</sub>/C<sub>29</sub>; R15 = %C<sub>27</sub>; R16 = %C<sub>28</sub>; R17 = %C<sub>29</sub>. Pr = Pristane; Ph = Phytane; Ts = 18α(H),21β(H)-22,29,30-trinorhopane; Ol = Oleanane; H30 = 17α(H),21β(H)-hopane; H29 = 17α(H),21β(H)-30-norhopane; C<sub>25</sub> = 17α(H),21β(H)-30,31,32,33-tetrakisohopane; C<sub>24</sub> = 17α(H),21β(H)-30,31,32,33-tetrakisohopane; GA = Gammacerane; C<sub>27</sub>R = 22R-17α(H),21β(H)-30-homohopane; C<sub>29</sub> = C<sub>29</sub> 5α(H),14α(H),17α(H)-cholestane; C<sub>29</sub>β = C<sub>29</sub> 5α(H),14α(H),17β(H)-cholestane; St/H = sum of steranes/sum of hopanes; H32 = 17α(H),21β(H)-bishomohopane; R13 = C<sub>27</sub> regular steranes/C<sub>29</sub> regular steranes; R14 = C<sub>28</sub> regular steranes/C<sub>29</sub> regular steranes; R15 = C<sub>27</sub> regular steranes/sum of C<sub>27</sub>, C<sub>28</sub> and C<sub>29</sub> regular steranes; R16 = C<sub>28</sub> regular steranes/sum of C<sub>27</sub>, C<sub>28</sub> and C<sub>29</sub> regular steranes; R17 = C<sub>29</sub> regular steranes/sum of C<sub>27</sub>, C<sub>28</sub> and C<sub>29</sub> regular steranes; ( ) : missing data, replaced by mean value in brackets.

Table 2

MDS similarity/dissimilarity measures between two objects/samples (j and k).

No.	Dissimilarity	Formula
1	Euclidean distance	$[\sum_i (X_{ij} - X_{ik})^2]^{1/2}$
2	Chebyshev distance	$\max_i  X_{ij} - X_{ik} $
3	Chi-squared distance	$\left\{ \sum_i \left( \frac{1}{\sum_j X_{ij}} \right) [X_{ij} / (\sum_j X_{ij}) - X_{ik} / (\sum_j X_{ik})]^2 \right\}^{1/2}$
4	Bray-Curtis distance	$(\sum_i  X_{ij} - X_{ik} ) / [\sum_i (X_{ij} + X_{ik})]$

(Borg and Groenen, 2005) has been used widely in many scientific fields. Biomarker ratios are mainly used in the present study. There have been few papers on biomarker data using MDS, although there have been several papers using them with principal component analysis (e.g., Peters et al., 2008). Because the distance used and the goodness-of-fit reached for MDS calculation of the biomarker ratios data cannot be obtained from the literature, we performed various tests as discussed in the following.

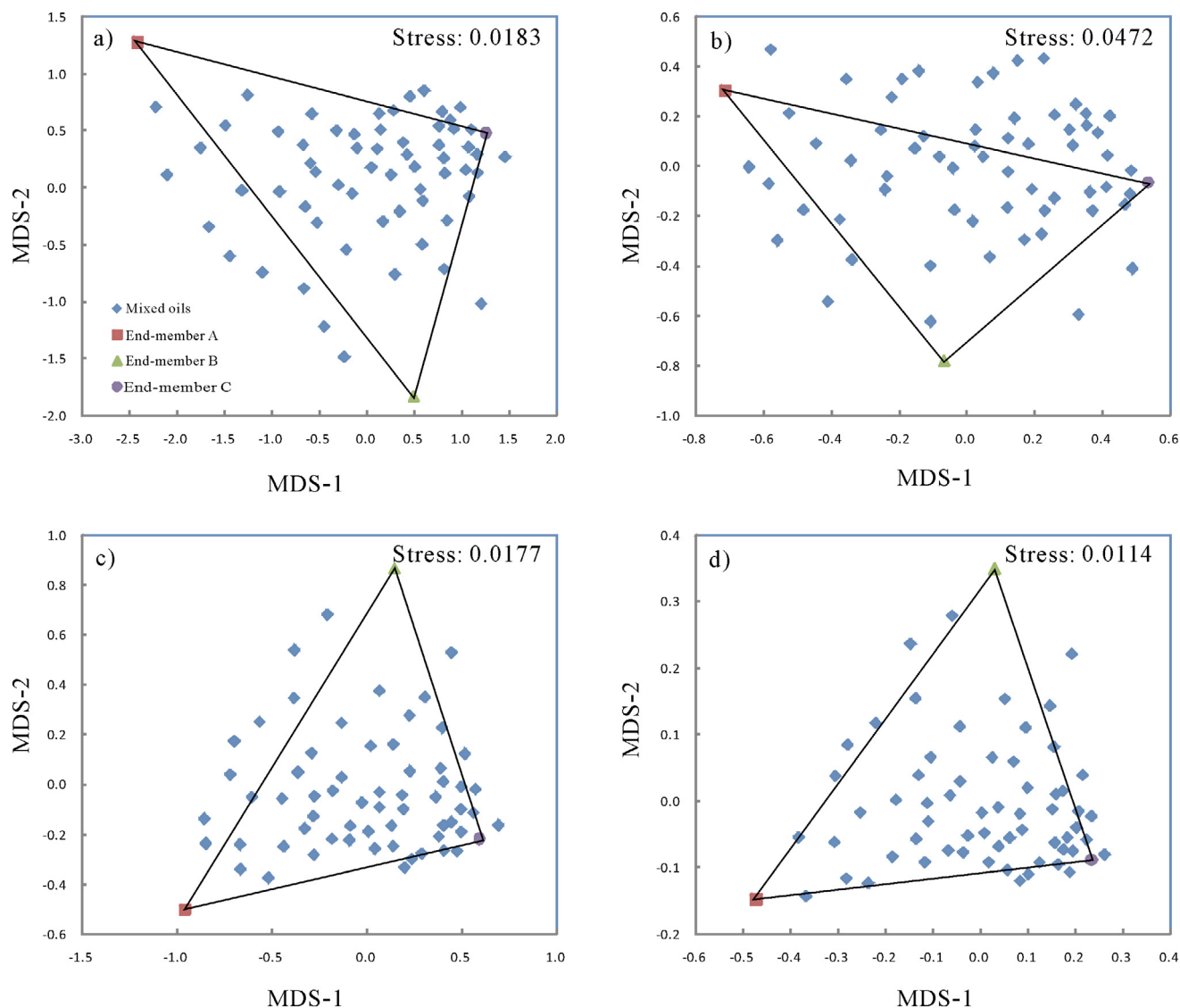
### 2.3.2. Similarity measure

The distance is used as the similarity/dissimilarity measure between samples/objects. Before distances are computed, the biomarker data are first normalized by the maximum-minimum range, i.e.  $X' = (X - X_{min}) / (X_{max} - X_{min})$ , in order to give equal weight to each individual biomarker ratio. The biomarker ratios of Zhan et al. (2016) were employed to identify which kind of distance is best for maintaining the original distribution of objects in high-dimensional space when mapped onto two dimensions. The data of Zhan et al. (2016) were obtained from laboratory mixtures of three end-member oils, demonstrating a triangular shape in high-dimensional space (biomarker concentrations). This shape can easily be distinguished in two dimensions. Twenty-four ratios were employed for each of 61 samples (Table 4 of Zhan et al., 2016) to choose the similarity measure. Four kinds of distances were selected (Euclidean distance, Chebyshev, Chi-squared and Bray-Curtis distance; Table 2) in order to check which distance measure is more suitable for biomarker ratios.

The Euclidean distance is frequently used in classical MDS and PCA. However, when used for the biomarker ratios it failed to distinguish the mixed oils (Fig. 3a). The Chebyshev distance presents the maximum difference of the ratio variants. The stress values noted in Fig. 3 clearly indicate that the Bray-Curtis distance is best. Our results are similar to these of Faith et al. (1987), who found that the Bray-Curtis measure is a robust distance, while Chi-squared distance and Euclidean distance are less robust. Therefore, we suggest that the Bray-Curtis distance can better reflect the genetic relationships between source rocks and the oils.

### 2.3.3. Examination of the MDS technique

Before the MDS technique is applied, it is necessary to test it for oil-oil and the oil-source rock correlation using known data sets. As Peters et al. (2007) demonstrated, there are four principal oil families on the Barrow Arch, North Slope, Alaska, including families 211, 212, 222 and 2321. Family 222 oils originated from the Triassic Shublik Formation and Family 212 oils are believed to be derived from the Cretaceous Hue Shale-GRZ; Family 211 is thought to be derived from a mixture of oils mainly from the Shublik Formation and Hue Shale-GRZ; while Family 2321 is from the basal (Hettangian-Aalenian) Kingak Shale source rock (Peters et al., 2008). Seventeen biomarker ratios and two stable carbon isotope ratios for 74 samples in Families 211, 212, 222 and 2321 were selected from Table A1 in Peters et al. (2008). The data were used to compute the Bray-Curtis distance and draw a 2D MDS diagram.



**Fig. 3.** Four kinds of distances used for MDS calculation were applied to biomarker data from Zhan et al. (2016); (1) Euclidean distance; (2) Chebyshev distance; (3) Chi-squared distance; (4) Bray-Curtis distance.

The results are shown in Fig. 4, in which the four oil families are clearly classified. Family 211 (mixed oils of Family 212 and 222) are obviously close to family 212, suggesting more contribution from family 212 oils. This is consistent with the conclusion of Peters et al. (2008).

In the Iranian sector of the Persian Gulf, there are two principal oil families (Mashhadi and Rabbani, 2015). Group I was interpreted to be generated from the Cretaceous Ahmadi member of the Sarvak Formation. Thirty-five samples and 13 biomarker ratios (data from Tables 4–6 and 8 of Mashhadi and Rabbani, 2015) were analyzed by MDS using the Bray-Curtis distance. Fig. 5 shows the MDS results, which distinguish Group I and Group II oil families. Group I was genetically related to Ahmadi member rocks (Mashhadi and Rabbani, 2015). In Fig. 5, two Ahmadi member rock samples are close to the Group I oils, supporting the conclusions of Mashhadi and Rabbani (2015). However, one sample of the Ahmadi member (Reshadat, CR-10-H3) is far from the Group I oils, in conflict with this interpretation. This conflict may be due to maturity of the rock sample. Biomarker parameters, including  $C_{32} \text{ 22S}/(22S + R)$  hopanes,  $C_{29} \beta\beta/(\alpha\alpha + \beta\beta)$  and  $C_{29} \text{ 20S}/(20R + 20S)$  steranes and

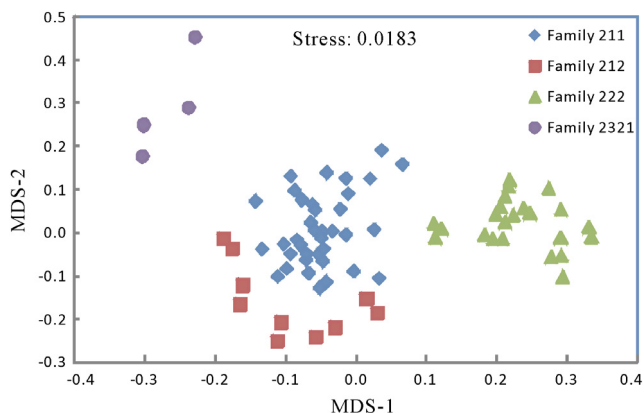
moretane/17 $\alpha$ -hopane (0.3, 0.28, 0.05 and 0.48, respectively), show that this sample has not reached the 'oil window'.

Based on an examination of the above published data, MDS is an effective method for oil–oil and the oil–source rock correlation using biomarker ratios. The genetic relationships between oils as well as oils and source rocks can be easily distinguished on a two-dimensional plot. In the following section, a case study using the Fangzheng Fault Depression samples reveals genetic relationships between source rocks and oils by MDS.

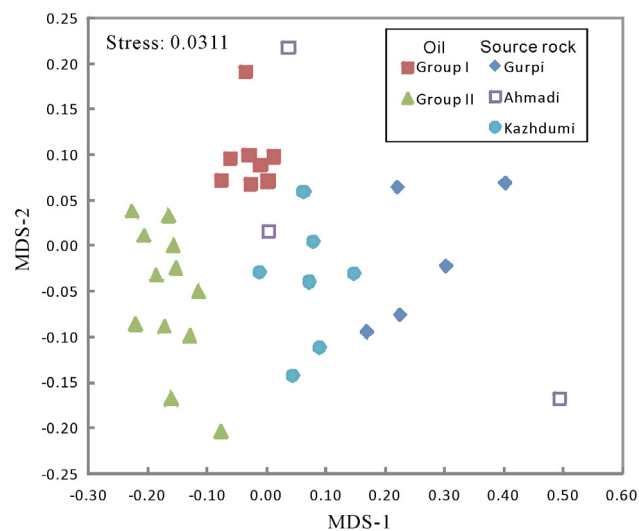
### 3. Results and discussion

#### 3.1. Source rock characteristics

Two main source rocks exist in the Fangzheng Fault Depression: The Xinancun-Wuyun Formation and the Muling Formation. Detailed source rock distribution and evaluation of the study area can be found in He et al. (2011), Li and Liu (2011), and Zhang (2012). The Xinancun-Wuyun source rocks are widespread, but only limited Muling strata have been found in the study area. Total



**Fig. 4.** MDS plot of data from Prudhoe Bay Field, North Slope, Alaska for use in oil–oil correlation (data from Peters et al., 2008). The 17 biomarker ratios and two stable carbon isotopic values were used for calculating MDS, including C<sub>19</sub>/C<sub>23</sub>, C<sub>22</sub>/C<sub>21</sub>, C<sub>24</sub>/C<sub>23</sub> and C<sub>26</sub>/C<sub>25</sub> tricyclic terpanes, C<sub>24</sub> tetracyclic terpane/C<sub>23</sub> tricyclic terpane (Tet/C<sub>23</sub>), C<sub>27</sub> tetracyclic terpane/C<sub>27</sub> tricyclic terpane [C<sub>27</sub>T/(Ts + Tm)], 28,30-bisnorhopane/hopane (C<sub>28</sub>/H), C<sub>29</sub> 30-norhopane/hopane (C<sub>29</sub>/H), C<sub>30</sub> diahopane/hopane(X/H), oleanane/hopane (OL/H), C<sub>31</sub> homohopane/hopane (C<sub>31</sub>R/H), gammacerane/C<sub>31</sub> homohopane 22R (GA/C<sub>31</sub>R), C<sub>35</sub> homohopane 22S/C<sub>34</sub> homohopane 22S (C<sub>35</sub>S/C<sub>34</sub>S), C<sub>26</sub> tricyclic terpane/trisnorhopane (C<sub>26</sub>T/Ts), steranes/hopanes (S/H), C<sub>27</sub>/C<sub>29</sub> regular steranes, C<sub>28</sub>/C<sub>29</sub> regular steranes, δ<sup>13</sup>C<sub>sat</sub> and δ<sup>13</sup>C<sub>aromatic</sub>.



**Fig. 5.** MDS plot for the Iranian sector of the Persian Gulf for use in oil–oil and oil–source rock correlation (data from Mashhadi and Rabbani, 2015). The biomarker ratios used for MDS are: Pr/Ph, Pr/*n*-C<sub>17</sub>, Ph/*n*-C<sub>18</sub>, C<sub>27</sub>/C<sub>29</sub> regular sterane, C<sub>28</sub>/C<sub>29</sub> regular sterane, C<sub>27</sub> diasterane/regular sterane, C<sub>29</sub>/C<sub>30</sub> hopane, C<sub>30</sub> diahopane/C<sub>30</sub> hopane, gammacerane/C<sub>30</sub> hopane, C<sub>19</sub>/C<sub>23</sub> tricyclic terpane, C<sub>29</sub> 20S/(S + R), C<sub>29</sub> αβ/(αα + αβ), moretanes/17α-hopane.

organic carbon (TOC) contents of the Xinancun-Wuyun Formation range from 0.19–10.36% with an average value of 1.42%. The source rocks of the Muling Formation have relatively higher TOC (0.53–10.43%) with an average value of 2.26%. The source rocks of the Xinancun-Wuyun and the Muling Formations have low to mature and mature organic matter, respectively.

All of the rock extracts of the Xinancun-Wuyun and the Muling formations show unimodal normal alkane distributions, typically dominated by short-chain (*n*-C<sub>15</sub> to *n*-C<sub>20</sub>) to longer-chain alkanes (*n*-C<sub>20</sub> to *n*-C<sub>25</sub>) (Fig. 6). Pristane/phytane is commonly used as an indicator of depositional conditions and source of organic matter

(Didyk et al., 1978). High Pr/Ph ratios (> 3) indicate terrigenous organic matter input under oxic conditions (Powell and McKirdy, 1973) and extremely low Pr/Ph ratios (< 0.8) are believed to indicate hypersaline or carbonate environments (Peters et al., 2005). Pr/Ph ratios of extracts from the Xinancun-Wuyun Formation range from 0.43–5.40 with an average of 2.30, which suggests deposition in oxic to anoxic conditions (Table 1) and significant terrigenous input. Samples of the Muling Formation have Pr/Ph ratios ranging from 0.41–3.24 with an average of 1.68 (Table 1), suggesting mainly terrigenous depositional conditions (Fig. 7).

The sterane distributions (*m/z* 217) for all samples are characterized by a predominance of C<sub>29</sub> relative to regular C<sub>27</sub> and C<sub>28</sub> steranes (Fig. 6). The source rock extracts of the Xinancun-Wuyun Formation have regular C<sub>27</sub>, C<sub>28</sub>, C<sub>29</sub> steranes in the range 17–43% (average 32%), 19–30% (average 23%) and 30–64 (average 45%), respectively (Table 1). The relative contents of regular C<sub>27</sub>, C<sub>28</sub>, and C<sub>29</sub> steranes from the Muling Formation range from 25–39% (average 35%), 18–22% (average 20%), 42–57% (average 45%), respectively. These values indicate major terrigenous organic matter input and less contribution of aquatic microbial organic matter, as supported by the diagram of regular C<sub>27</sub>, C<sub>28</sub>, C<sub>29</sub> steranes (Fig. 8).

Terpane distributions (*m/z* 191) for all samples are characterized by an abundant C<sub>30</sub> hopane (Fig. 6). The C<sub>35</sub>/C<sub>34</sub> hopane ratios for the Xinancun-Wuyun Formation range from 0.22–0.46 (average 0.34), indicating oxic depositional conditions. High gammacerane/C<sub>31</sub>R is usually associated with high salinity and reducing conditions (Fu et al., 1986). Gammacerane/C<sub>31</sub>R ratios for the samples are extremely low, ranging from 0.01–0.09, possibly suggesting a general absence of water column stratification. The samples from the Muling Formation also have gammacerane/C<sub>31</sub>R ratios in the range of 0.01–0.09, also reflecting a general absence of water column stratification. Therefore, we conclude that the organic matter from the Xinancun-Wuyun Formation was mainly derived from terrigenous input deposited under oxic conditions. Similarly, organic matter in the Muling Formation was also derived from major terrigenous input under normal water salinity.

Biomarker parameters were used to evaluate the thermal maturity of the 25 rock samples. The samples (18 from the Xinancun-Wuyun and seven from the Muling Formation) were selected to assess the level of thermal maturity and genetic relationships between oil–oil and oil–source rock by means of the cross-plots. Previous studies (He et al., 2011; Zhang, 2012) demonstrated that the depth of oil generative thresholds for the Xinancun-Wuyun and the Muling Formation are about 2000 m and 1716 m, respectively. The depths of all samples are listed in Table 1, which suggest that the investigated 25 samples are thermally mature. The ratios of C<sub>32</sub> hopane 22S/(22S + 22R) from the Xinancun-Wuyun and the Muling Formation samples are 0.55–0.61 and 0.54–0.58, respectively. With increasing maturity, C<sub>32</sub> hopane 22S/(22S + 22R) ratio rises from 0 to 0.6 and the equilibrium value is 0.57–0.62, indicating that rock samples of the Xinancun-Wuyun and the Muling Formation have at least entered the oil generation stage (Seifert and Moldowan, 1980). The level of thermal maturity of the Muling Formation samples is less than the Xinancun-Wuyun since it is shallower. The ratios of C<sub>29</sub> steranes 20S/(20R + 20S) and C<sub>29</sub> ββ/(αα + ββ) of the Xinancun-Wuyun and the Muling Formation are 0.38–0.66, 0.35–0.65 and 0.36–0.51 and 0.30–0.53 (Table 1), suggesting that the samples are mature (Peters et al., 2005).

### 3.2. Crude oils

The analyzed oils include a complete suite of acyclic isoprenoid (e.g., pristane and phytane) and low molecular-weight *n*-alkanes. Therefore, there is no evidence of significant biodegradation among the Xinancun-Wuyun and Muling oil samples (Fig. 9). In



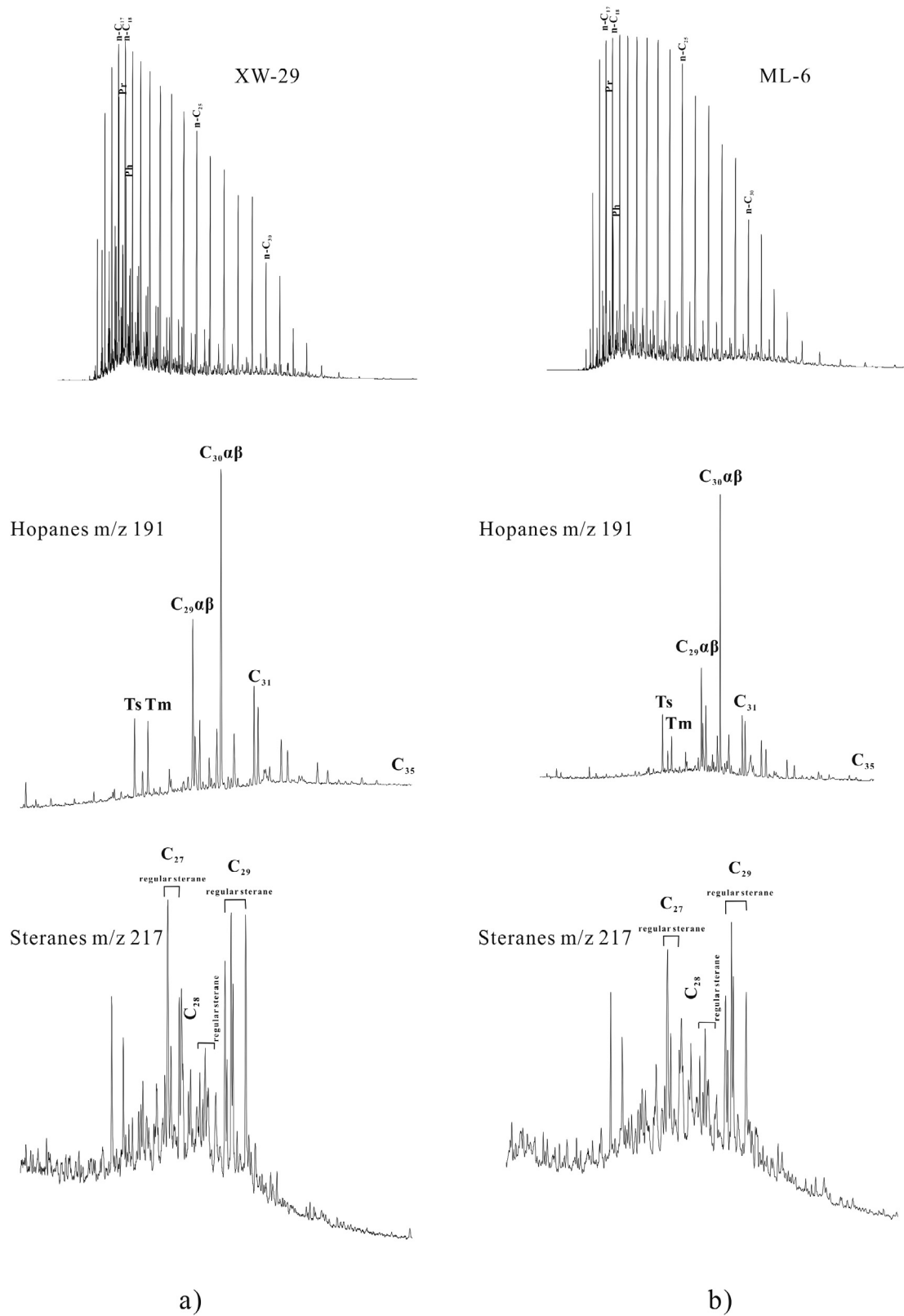


Fig. 6. GC-MS data for source rock samples from: (a) the Xinancun-Wuyun and (b) the Muling formations.

addition, very low values of  $Pr/n-C_{17}$  and  $Ph/n-C_{18}$  for all oils also indicate that these oil samples are not biodegraded. According to previous work (He et al., 2008; Liu et al., 2014), the Fangzheng Fault Depression contains mainly lacustrine facies. Gas

chromatograms of the saturated hydrocarbon fractions of oils from the Xinancun-Wuyun and the Muling Formations are presented in Fig. 9 and the geochemical parameters are in Table 1. Among 12 oil samples, the  $n$ -alkanes show a unimodal distribution with a pre-

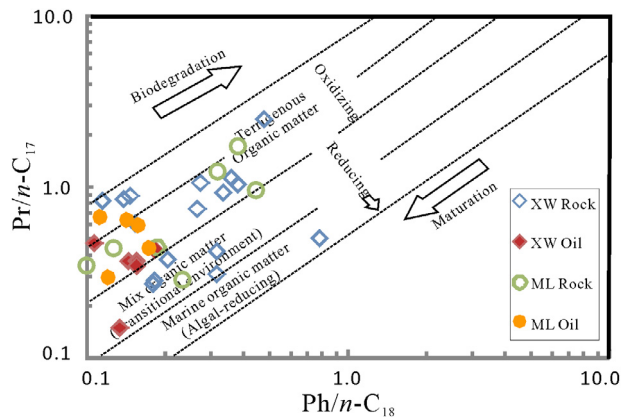


Fig. 7. Plot of  $Pr/n-C_{17}$  versus  $Ph/n-C_{18}$  ratios for the assessment of oil sources (after Connan and Cassou, 1980).

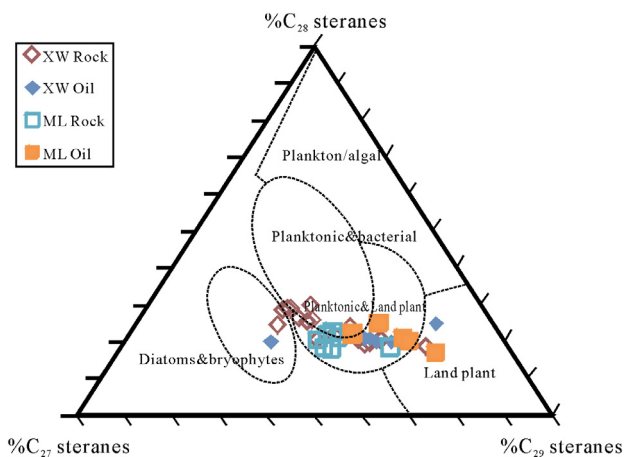


Fig. 8. Ternary diagram of regular steranes ( $C_{27}$ – $C_{29}$ ) showing the relationship between sterane compositions, source input and depositional conditions for the Fangzheng samples (modified after Huang and Meinschein, 1979).

dominance of low to middle molecular-weight compounds ( $n-C_{12}$ – $n-C_{20}$  or  $n-C_{12}$ – $n-C_{25}$ ). Significant amounts of acyclic isoprenoids occur in all 12 oil samples and their biomarker ratios (e.g., pristane/phytane) are listed in Table 1. As mentioned above,  $Pr/Ph$  is an effective parameter to assess redox conditions. In the studied oil samples,  $Pr/Ph$  ratios are in the range of 1.4–5.3, reflecting significant terrigenous input under oxic conditions (Peters et al., 2005).

The biomarker ratios for steranes ( $m/z$  217) in oils of the Xinancun-Wuyun Formation and the Muling Formation (Fig. 9) are presented in Table 1. Ternary plots of regular  $C_{27}$ ,  $C_{28}$  and  $C_{29}$  regular steranes are commonly used to distinguish different organic matter input, redox conditions of the depositional conditions and for oil–source rock correlation. Compared to  $C_{27}$  steranes [12–31%;  $C_{27}/(C_{27} + C_{28} + C_{29})$ ] and  $C_{28}$  steranes (17–25%), these oils have high proportions of  $C_{29}$  steranes (47–63%), suggesting a dominant contribution of terrigenous organic matter, as noted previously and as shown in Fig. 8. This is also consistent with the observation of low sterane/hopane ratios (0.1–0.22), indicating terrigenous or microbially reworked organic matter.

The  $m/z$  191 mass fragmentograms of the saturated hydrocarbon fraction of the oils are characterized by abundant  $C_{30}$  hopane. The  $C_{35}/C_{34}$  hopane ratios range from 0.27–0.43, indicating oxic depositional conditions, consistent with high  $Pr/Ph$  ratios. As discussed above, high gammacerane/ $C_{31}R$  is often related to high

salinity and reducing conditions. However, the Fangzheng oils have very low values of gammacerane/ $C_{31}R$  ratios in the range of 0.03–0.07. These low values are consistent with source rock deposited in fresh to brackish water conditions. On the basis of the biomarker analysis, including  $n$ -alkanes and isoprenoids, terpanes and steranes, the investigated oils in the Fangzheng Fault Depression are characterized by high  $Pr/Ph$ , high proportions of regular  $C_{29}$  steranes and low steranes/hopanes, all of which indicates that the organic matter was derived from a dominant contribution of terrigenous input deposited under oxic conditions.

The  $22S/(22S + 22R)$  ratios for  $C_{32}$  hopane, a parameter highly specific for immature to early oil generation (Peters et al., 2005), range from 0.51–0.62. This suggests that most of the oil samples have reached equilibrium (Seifert and Moldowan, 1980), which indicates the Fangzheng oils are mature. The ratios of  $C_{29}$  steranes  $20S/(20R + 20S)$  and  $C_{29}$   $\beta\beta/(\alpha\alpha + \beta\beta)$  are 0.33–0.50 and 0.45–0.68, respectively, indicating early-mature to peak oil window (Peters et al., 2005). A plot of  $Ph/n-C_{18}$  versus  $Pr/n-C_{17}$  ratios suggests the same interpretation and indicates that the investigated oils are at mature or peak oil window maturation (Fig. 7).

### 3.3. Oil–source rock correlation based on chemometrics

PCA is an effective tool to identify genetic affinities among large numbers of oils and source rocks (Peters et al., 2007, 2013). Biomarker parameters are generally used in chemometric oil–source rock correlations because they have source- and age-related meaning and are less altered by secondary processes. Compared to pentacyclic terpanes, the tricyclic terpanes are in very low abundance in the Fangzheng oils and source rocks (Figs. 6 and 9). Nine source-related biomarker ratios were selected for the chemometric oil–source rock correlations including PCA and MDS, while maturity-related parameters were excluded. The nine biomarker ratios employed are:  $Pr/Ph$ ,  $Ts/(Ts + Tm)$ ,  $OL/H_{30}$ ,  $H_{29}/H_{30}$ ,  $C_{35}/C_{34}$ , Gammacerane/ $C_{31}R$ ,  $S/H$ ,  $C_{27}/C_{29}$  and  $C_{28}/C_{29}$  (Table 1), which is similar to previous studies (Peters et al., 2007, 2013). In practice, sufficient sample size, usually over 30, is necessary; the sample number should be more than the variables, generally as much as 1.2 times the variates used in order to avoid unreliable results.

The biomarker ratios were preprocessed by range scaling in Pirouette® software (Infometrix, Inc.) before PCA was performed. The PCA results demonstrate the investigated crude oil samples can be divided into two groups (Fig. 10). Group A oils (116–118) correlate with the Xinancun-Wuyun Formation, while Group B oils (114, 115, 119–125) are associated with the source rocks of both the Xinancun-Wuyun Formation (1–103) and the Muling Formation (104–113).

MDS was also applied to oil–source rock correlation using in-house software. The biomarker parameters were the same as those used for the PCA. The stress of MDS in 2D space is computed as 0.0545 after 1000 iterations, suggesting that the goodness-of-fit is good to excellent (Table 3). The MDS goodness-of-fit (i.e., 94.5% explained) was a little better than 3D-PCA, in which 90.8% variance was explained (Fig. 10). Nonlinear MDS biplots (Greenacre and Primicero, 2013) are shown in Fig. 11. Fig. 11a exhibits the maturity direction using  $C_{29}$   $20S/(20R + 20S)$  ratio as indicator, and Fig. 11b shows the variation in depositional conditions when  $Pr/Ph$  ratio is used as the indicator, supporting the analysis of the MDS plot (Fig. 12).

Fig. 12 shows the MDS scenario of multi-parameter oil–source rock correlation for the Fangzheng Fault Depression. The Xinancun-Wuyun Formation (XW) rocks have a wide distribution range, while the Muling Formation (ML) source rocks are distributed in the left top of the figure and partially overlap the XW rocks, indicating the rocks have similar geochemical and depositional characteristics.

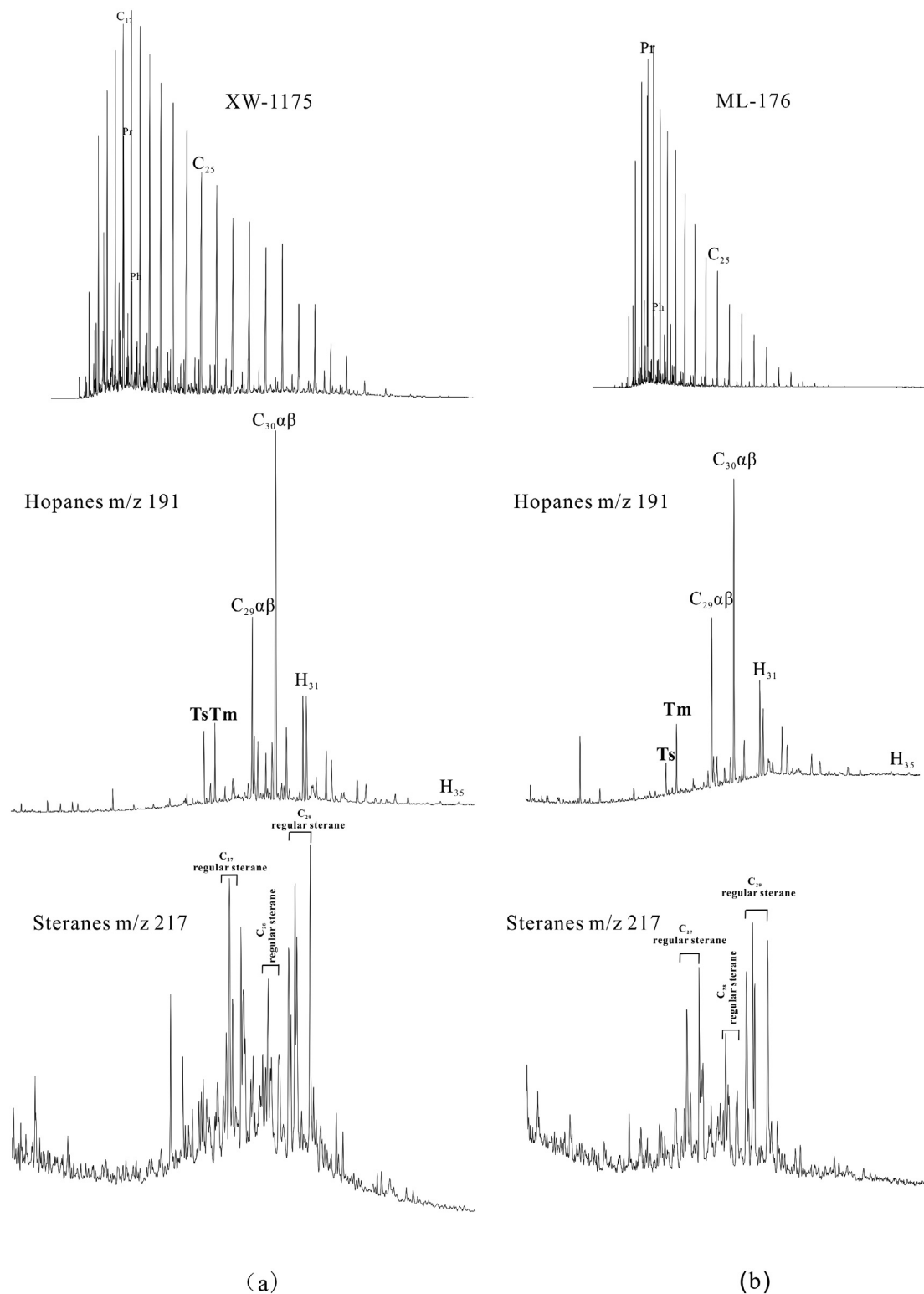
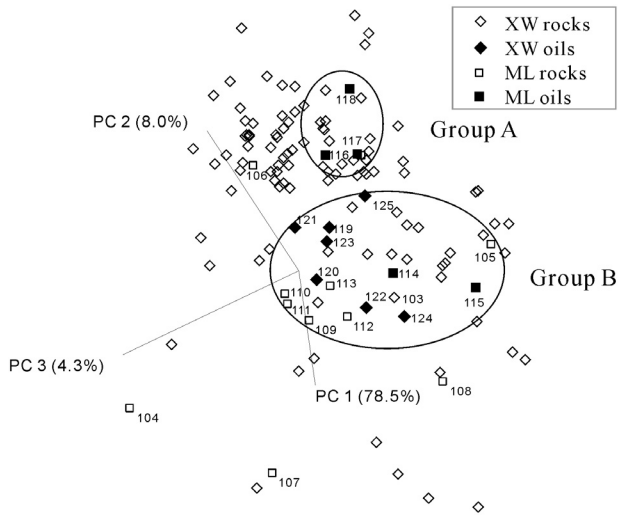


Fig. 9. TIC, terpene ( $m/z$  191) and sterane ( $m/z$  217) distributions of selected oils from the Fangzheng Fault Depression.

In the MDS plot, all of the oils plot in the mature rock area. The oils with relatively high Pr/Ph ratios from the Muling Formation (116–118), are closely related to the Xinancun-Wuyun Formation with relatively oxic depositional conditions. The other oils are associated with both XW and ML rocks, which is similar to Group B in PCA. However, the XW oils (119–125) were generated from

the Xinancun-Wuyun Formation, while the other two oils (114 and 115) could be a mixture generated from both formations, although they were discovered in the Muling Formation.

PCA and MDS provide an integrated diagram showing affinity between oils and source rocks in a visible style. Both PCA and MDS suggest that the oils from the Muling Formation were closely

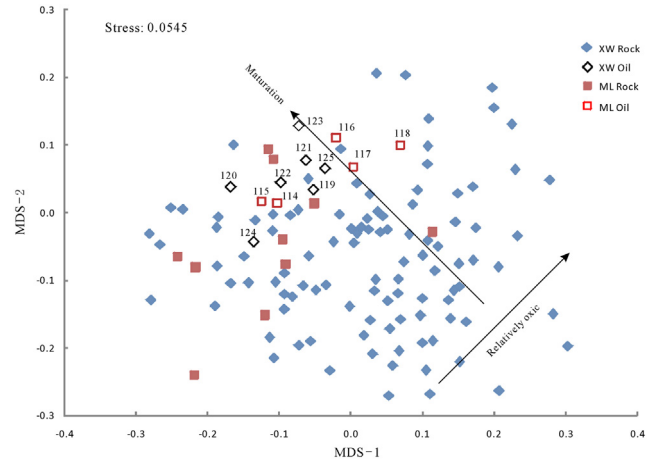


**Fig. 10.** Three-dimensional view of principal components analysis identifying oil genetics of the Fangzheng Fault Depression.

**Table 3**  
MDS stress and goodness-of-fit after Kruskal (1964) and Storti (2016).

Quality of configuration	Stress
Poor	> 0.20
Fair	0.10
Good	0.05
Excellent	0.025
Perfect	0.0

associated with source-rocks of the Xinancun-Wuyun Formation, which is likely the result of the huge thickness and wide distribution of the Xinancun-Wuyun Formation in the Depression. There is no evidence to show the oils were generated only from the Muling Formation.

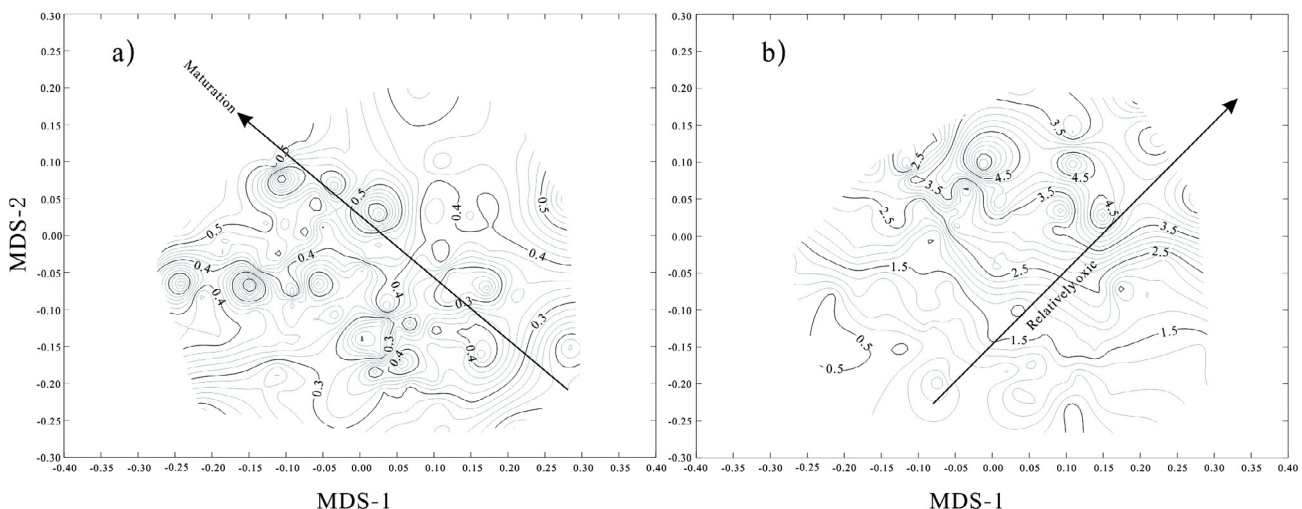


**Fig. 12.** MDS plot of the oil-source rock correlation for multiple biomarker parameters with variability in the maturation and depositional conditions indicated by arrows.

#### 4. Conclusions

The investigated oil fingerprints are characterized by a predominance of low- to mid-molecular weight compounds ( $n$ - $C_{12}$ - $n$ - $C_{20}$  or  $n$ - $C_{12}$ - $n$ - $C_{25}$ ), moderate to high Pr/Ph ratios (1.44–5.3), low  $C_{35}/C_{34}$  hopane ratios and relatively low  $C_{27}/C_{29}$  regular steranes. The oils found in the Fangzheng Fault Depression were mainly generated from source rocks of the Xinancun-Wuyun Formation with partial contribution from the Muling Formation based on MDS plots and geological evidence. The organic matter in the source rocks was mainly derived from terrigenous higher plants under relatively oxic conditions.

Chemometric methods (PCA and MDS) using multi-parameter biomarkers are effective tools for oil–oil and the oil–source rock correlation. The Bray-Curtis distance using MDS is a robust measure for the biomarker ratio data and the oil–source rock correlations. MDS plots can give useful information including the maturation and depositional conditions of the rocks and oils when combined in bi-plots.



**Fig. 11.** Nonlinear MDS bi-plots showing: (a)  $C_{29} S/(S + R)$  hopane contours and (b) the bi-plot of Pr/Ph ratio contours.

For the oil–source rock correlation, chemometric methods can be used to draw genetic relationships in an integrated diagram. However, the selected biomarker ratios must be based on compounds showing affinity, clear source- or age-related information, and resistance to secondary alteration. The available number of biomarker ratios depends heavily on the geochemical characteristics of the samples. Moreover, statistical rules must be followed: For instance, sample numbers should be > 30 and the number of variables used (in this case biomarker ratios) should not exceed the number of samples.

### Acknowledgements

We thank Drs. Ken Peters, J.A. Curiale and J.K. Volkman for the critical reviews, comments and extensive improvements to the grammar in our manuscript. This study was funded by the Natural Science Funding Council of China (Grant Nos. 41273059, 41173054), GIGCAS 135 project (Grant No. Y234021001) and Ear-marked Fund of the State Key Laboratory of Organic Geochemistry (Grant No. sklog2012A02). This is contribution No. IS-2286 from GIGCAS.

Associate Editor—Ken Peters

### References

- Borg, I., Groenen, P.J.F., 2005. *Modern Multidimensional Scaling: Theory and Applications*. Springer Science + Business Media, New York, pp. 1–14.
- Chen, D., Zhang, F., Chen, H., Dilek, Y., Yang, S., Meng, Q., Yang, C., 2015. Structural architecture and tectonic evolution of the Fangzheng sedimentary basin (NE China), and implications for the kinematics of the Tan-Lu fault zone. *Journal of Asian Earth Sciences* 106, 34–48.
- Connan, J., Cassou, A.M., 1980. Properties of gases and petroleum lipids derived from terrestrial kerogen at various maturation levels. *Geochimica et Cosmochimica Acta* 44, 1–23.
- Didyk, B.M., Simoneit, B.R.T., Brassell, S.C., Eglinton, G., 1978. Organic geochemical indicators of palaeoenvironmental conditions of sedimentation. *Nature* 272, 216–222.
- Dong, X., Liu, C., Zhao, R., Chen, X., Ma, L., 2008. Reservoir characteristics of Fang 4 well area of Xinancun Formation in Fangzheng Fault Depression. *Inner Mongolia Petrochemical* 17, 14–16 (in Chinese with English abstract).
- Faith, D.P., Minchin, P.R., Belbin, L., 1987. Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* 69, 57–68.
- Fu, G., Liu, T., Shi, J., Li, Y., Yang, L., 2014. Spatial matching relation between source rock and cap-rock and its control action on oil-gas accumulation in Fangzheng Fault Depression. *Lithologic Reservoirs* 5, 9–14 (in Chinese with English abstract).
- Fu, J., Sheng, G., Peng, P., Brassell, S.C., Eglinton, G., Jiang, J.G., 1986. Peculiarities of salt lake-sediments as potential source rocks in China. *Organic Geochemistry* 10, 119–126.
- Greenacre, M., Primicero, R., 2013. *Multidimensional Scaling Biplot, Multivariate Analysis of Ecological Data*. BBVA Foundation, Spain, pp. 138–148.
- He, X., Li, Y., Feng, Z., 2011. Source rock evaluation and oil–source rock correlation of Fangzheng Fault Depression. *Neijiang Science & Technology* 32 (1), 138–139 (in Chinese).
- He, X., Yang, J., Li, Y., 2008. Paleogene structural evolution and distribution characteristics of sedimentary facies in the Fangzheng Fault Depression of the Yishu Graben. *Geology in China* 5, 902–910 (in Chinese with English abstract).
- Hollemeier, K., Altmeyer, W., Heinze, E., Pitra, C., 2012. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry combined with multidimensional scaling, binary hierarchical cluster tree and selected diagnostic masses improves species identification of Neolithic keratin sequences from furs of the Tyrolean iceman Oetzi. *Rapid Communications in Mass Spectrometry* 26, 1735–1745.
- Hu, S., Guo, B., Lin, D., Yang, Y., Zhang, F., Zheng, J., 2010b. 3D seismic data fine structural interpretation and reservoir prediction in Fang-3 Well Block, Fangzheng Rift. *Oil Geophysical Prospecting* 45, 909–913 (in Chinese with English abstract).
- Hu, Y., Qi, R., Wei, Y., Wang, G., 2010a. Seismic evidence of strike-slip faults in the middle of Fangzheng Rift and favorable exploration areas. *Journal of Jilin University (Earth Science Edition)* 6, 1271–1277 (in Chinese with English abstract).
- Huang, W.Y., Meinschein, W.G., 1979. Sterols as ecological indicators. *Geochimica et Cosmochimica Acta* 43, 739–745.
- Kruskal, J.B., 1964. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29, 1–27.
- Lenz, E.J., Foran, D.R., 2010. Bacterial profiling of soil using genus-specific markers and multidimensional scaling. *Journal of Forensic Sciences* 55, 1437–1442.
- Li, C., Guo, W., Yu, J., 2015. Study on Palaeogene sedimentary facies of Fangzheng Fault Depression Daluomi-Xingwang area. *Jilin Geology* 34, 18–27 (in Chinese with English abstract).
- Li, X., Liu, H., 2011. Research on evaluating the source rock from Paleogene to Cretaceous in Fangzheng Fault Depression. *Inner Mongolia Petrochemical Industry* 8, 311–312 (in Chinese).
- Liu, Q., Guo, W., Jin, Z., 2014. The sedimentary facies of the second member of Baoquanling Formation around FZ4 exploratory well in the Fangzheng Fault Depression. *Geology and Resources* 6, 567–573 (in Chinese with English abstract).
- Mashhadi, Z.S., Rabbani, A.R., 2015. Organic geochemistry of crude oils and Cretaceous source rocks in the Iranian sector of the Persian Gulf: an oil–oil and oil–source rock correlation study. *International Journal of Coal Geology* 146, 118–144.
- McGee, V.C., 1968. Multidimensional scaling of n sets of similarity measures: a nonmetric individual differences approach. *Multivariate Behavioral Research* 3, 233–248.
- Peters, K.E., Coutrot, D., Nouvelle, X., Ramos, L.S., Rohrback, B.G., Magoon, L.B., Zumberge, J.E., 2013. Chemometric differentiation of crude oil families in the San Joaquin Basin, California. *American Association of Petroleum Geologists Bulletin* 97, 103–143.
- Peters, K.E., Scott Ramos, L., Zumberge, J.E., Valin, Z.C., Scotese, C.R., Gautier, D.L., 2007. Circum-Arctic petroleum systems identified using decision-tree chemometrics. *American Association of Petroleum Geologists Bulletin* 91, 877–913.
- Peters, K.E., Scott Ramos, L., Zumberge, J.E., Valin, Z.C., Bird, K.J., 2008. Deconvoluting mixed crude oil in Prudhoe Bay Field, North Slope, Alaska. *Organic Geochemistry* 39, 623–645.
- Peters, K.E., Walters, C.C., Moldowan, J.M., 2005. *The Biomarker Guide*. Cambridge University Press, Cambridge. 1155 pp.
- Powell, T.G., McKirdy, D.M., 1973. Relationship between ratio of pristane to phytane in crude oil composition and geological environment in Australia. *Nature* 243, 37–39.
- Reville, A.T., Carr, M.R., Rowland, S.J., 1992. Use of oxidative degradation followed by capillary gas chromatography–mass spectrometry and multi-dimensional scaling analysis to fingerprint unresolved complex mixtures of hydrocarbons. *Journal of Chromatography A* 589, 281–286.
- Seifert, W.K., Moldowan, J.M., 1980. The effect of thermal stress on source-rock quality as measured by hopane stereochemistry. *Physics and Chemistry of the Earth* 12, 229–237.
- Shi, L., Fan, Y., Lee, J.K., Waltham, M., Andrews, D.T., Scherf, U., Paull, K.D., Weinstein, J.N., 2000. Mining and visualizing large anticancer drug discovery databases. *Journal of Chemical Information and Computer Sciences* 40, 367–379.
- Shao, Z., Yang, J., Wang, H., Wu, G., Zhao, H., 2013. New understanding of interior structures and sediment–tectonic evolution of Fangzheng Fault Depression in eastern Heilongjiang Province. *Journal of Palaeogeography* 3, 339–350 (in Chinese with English abstract).
- Storti, D., 2016. Goodness of Fit. <[http://www.unesco.org/webworld/idams/advguide/Chapt8\\_1\\_1.htm](http://www.unesco.org/webworld/idams/advguide/Chapt8_1_1.htm)>.
- Yang, C., Ren, J., Zhang, Z., Zhang, J., 2014. Cenozoic structural characteristics and evolution in the Fangzheng Fault Depression, Northeast China. *Geotectonica et Metallogenia* 2, 388–397 (in Chinese with English abstract).
- Zhou, P., Chen, C., Ye, J., Shen, W., Xiong, X., Hu, P., Fang, H., Huang, C., Sun, Y., 2015. Combining molecular fingerprints with multidimensional scaling analyses to identify the source of spilled oil from highly similar suspected oils. *Marine Pollution Bulletin* 93, 121–129.
- Zhan, Z.-W., Zou, Y.-R., Shi, J.-T., Sun, J.-N., Peng, P., 2016. Unmixing of mixed oil using chemometrics. *Organic Geochemistry* 92, 1–15.
- Zhang, X., 2012. Hydrocarbon source rock evaluation and oil–source rock correlation in the fifteenth well of the Fangzheng Fault Depression. *Inner Mongolia Petrochemical Industry* 14, 127–128 (in Chinese).