Identification of genetically distinct petroleum tribes in the Middle Magdalena Valley, Colombia

William Thompson-Butler, Kenneth E. Peters, Leslie B. Magoon, Allegra Hosford Scheirer, J. Michael Moldowan, Vladimir Orlando Blanco, Roman Eugenio Gonzalez, Stephan A. Graham, John E. Zumberge, and David A. Wavrek

ABSTRACT

Six distinct oil tribes were established using multivariate statistical analysis of source-related biomarker and isotopic ratios for 73 crude oil samples from the Middle Magdalena Valley (MMV), Colombia. These six tribes show a systematic distribution by both basin location and reservoir rock age and may originate from different source rocks or different organofacies of the same source rock. Biomarker and isotopic data further differentiate the tribes with respect to source rock depositional environment, lithology, organic matter type, and thermal maturity. The thermal maturity and reservoir interval for the northernmost tribe 5 suggest a middle Cretaceous Tablazo Formation source rock. In contrast, tribes 1 through 4 are likely derived from the primary regional source rock, the Upper Cretaceous La Luna Formation. However, we observe regional differences in bulk properties, thermal maturity, terrigenous input, and oxicity between the four La Luna–derived oil tribes. In addition, tribe 3 appears to result from end-member mixing between tribes 2 and 4. Finally, the southernmost tribe 6 is the only oil with terrigenous character. Diamondoid analysis shows the presence of significant secondary cracking in the tribe 6 oil with low levels of cracking present in oil samples from tribes 1 and 2 in the central MMV. This suggests a more deeply buried nonmarine source along the western flank of the Andean Eastern Cordillera. The integration of chemometric, biomarker, and diamondoid analyses have improved our understanding of the MMV petroleum system and advocate for the presence of three or more source rock intervals within the basin.

AUTHORS

WILLIAM THOMPSON-BUTLER ~ *Department of Geological Sciences, Stanford University, Stanford, California; rthompsonbutler@gmail.com*

William Thompson-Butler is a Ph.D. candidate in geological sciences at Stanford University School of Earth, Energy and Environmental Sciences. His research focus is basin and petroleum system modeling and sedimentary geology. He received his B.A. in environmental science and public policy and earth and planetary sciences at Harvard University. His professional experience includes industry internships with ExxonMobil Development Company, BHP Regional Exploration, and Chevron Energy Technology Company.

KENNETH E. PETERS ~ *Schlumberger, Mill Valley, California; Department of Geological Sciences, Stanford University, Stanford, California; kpeters2@slb.com*

Kenneth E. Peters is a Schlumberger geochemistry advisor. He has 40 years with Chevron, Mobil, ExxonMobil, the US Geological Survey (USGS), and Schlumberger and published more than 165 geoscience books and papers. Ken is adjunct professor in the Stanford Basin and Petroleum System Modeling Industrial Affiliates Program. He holds the 2009 Alfred Treibs, 2013 AAPG Honorary Member, 2016 European Association of Geoscientists and Engineers Alfred Wegener medals, a 2017 AAPG Heritage of the Petroleum Geologist honoree, and the 2019 Sidney Powers Memorial Award.

LESLIE B. MAGOON ~ *Department of Geological Sciences, Stanford University, Stanford, California; lmagoon@stanford.edu*

Leslie B. Magoon is currently an adjunct professor in basin and petroleum system modeling at Stanford University. He worked for Shell Oil Company and the USGS. Since 1981, he popularized the petroleum system through talks, courses, and AAPG Memoir 60, *The Petroleum System—From Source to Trap*, for which he and his coeditor received the R. H. Dott, Sr. Award in 1996.

ALLEGRA HOSFORD SCHEIRER ~
Department of Geological Sciences, Stanford University, Stanford, California; allegras@ stanford.edu

Allegra Hosford Scheirer is a research scientist at Stanford University where she codirects the Basin and Petroleum System Modeling Industrial Affiliates Program. Prior to Stanford, Allegra was a research geophysicist at USGS. She is the editor of USGS Professional Paper 1713 and a past associate editor of the *Journal of Geophysical Research*. Allegra's degrees are from the Massachusetts Institute of Technology–Woods Hole Oceanographic Institution Joint Program in Oceanography (Ph.D.) and Brown University (geology and physics/math).

J. MICHAEL MOLDOWAN ~ *Biomarker Technologies, Inc., Rohnert Park, California; jmmoldowan@biomarker-inc.com*

J. Michael (Mike) Moldowan attained a Ph.D. in chemistry in 1972 from the University of Michigan. He became the eventual leader of the Chevron biomarker team from 1974 to 1993, which established the application of biological marker technology to petroleum exploration. He joined the Department of Geological Sciences at Stanford University as a professor (research) in 1993. In 2012, he founded the laboratory Biomarker Technologies, Inc., which offers unique applications of advanced geochemical technologies to industry. Moldowan has published more than 100 papers on geochemistry, including 2 editions of *The Biomarker Guide* in 1993 and 2005. In 2011, Moldowan received the Treibs Medal, considered the ultimate career award for organic geochemistry.

VLADIMIR ORLANDO BLANCO ~ *Instituto Colombiano del Petróleo, Ecopetrol, Bucaramanga, Colombia; vladimir.blanco@ ecopetrol.com.co*

Vladimir Orlando Blanco is a senior researcher and the current head of the organic geochemistry and petroleum charge team at Ecopetrol–Instituto Colombiano del Petróleo, where he provides technical assistance to exploration and production projects. He holds a B.Sc. (2004) in geology from Universidad Industrial de Santander and an M.Sc. (2012)

INTRODUCTION

The Middle Magdalena Valley (MMV) is an intermontane basin and petroleum province in northwestern Colombia that, along with the Eastern Cordillera and Llanos Basin, acted as a regional sedimentary basin from the Triassic into the late Neogene (Figure 1). The MMV is a productive petroleum province, accounting for nearly 3 billion bbl of oil and 3 tcf (~85 billion m^3) of gas from more than 40 fields in over a century of exploration (IHS Markit, 2010). The distribution of fields, in addition to the estimated ultimate recovery of oil equivalent, are shown in Figure 2 and described in more detail in Table 1. Many of the early discoveries were the larger accumulations in the central and central-eastern part of the study area (Morales and Colombian Petroleum Industry, 1958; Dickey, 1992). However, as exploration progressed, large fields in the southern part of the MMV, as well as several smaller accumulations in the northern MMV, have contributed to reserves. Despite this long history of petroleum exploration, the complex burial and variable exhumation history associated with the Andean orogeny add uncertainty to the evaluation of petroleum systems in the MMV by impacting the distribution, quality, and thermal maturity of the thick Cretaceous succession of organic-rich source rocks.

From the Campanian to the Maastrichtian, the axis of sediment deposition in Colombia migrated eastward toward the present-day position of the western foothills of the Eastern Cordillera as the Central Cordillera was gradually uplifted (Villamil, 1999). Restricted marine conditions during this interval resulted in deposition of organic-rich carbonate and siliciclastic sediments in a broad foreland basin within the MMV (Cooper et al., 1995). The Upper Cretaceous La Luna Formation, a lateral equivalent to the primary source rock for the Maracaibo Basin in Venezuela, has been well documented by source rock analysis and oil–source rock extract correlation as the primary regional source rock within the MMV (e.g., Zumberge, 1980, 1984; Rangel et al., 2000; Ramón et al., 2001). However, other source rocks could also contribute to charge in the basin. Specifically, deposition of the Aptian–Albian Tablazo Formation was coincident with a period of ocean anoxia and local deepening that resulted in organic matter preservation (Villamil, 1998). A regional study of Colombian oils by Rangel et al. (2017) identified two petroleum families in the MMV and suggested that one originated from the calcareous facies of the La Luna Formation, whereas the other originated from siliciclastic facies of the Tablazo Formation. Figure 3 shows a log suite of Rock-Eval T_{max} (the temperature corresponding to the maximum production of hydrocarbon products by kerogen cracking during pyrolysis), resistivity, total organic carbon (*TOC*), and hydrogen index (*HI*) data for a location in the central MMV and illustrates the source potential of the Cretaceous interval. Whereas

the La Luna and Tablazo Formations show high *TOC* and resistivity values, elevated thermal maturity and resulting depletion of *HI* in the Tablazo Formation raise uncertainty regarding the timing of a Tablazo-sourced petroleum system in the central MMV. However, *HI* values for the Tablazo Formation typically range from 100 to 600 mg HC/g *TOC* with intervals in the less deeply buried northern MMV often in the 300–500 mg HC/g *TOC* range (Sarmiento, 2011).

Plate convergence along the western margin of South America acted as the primary driver for the Paleogene evolution of the MMV. The oblique accretion of the Panama arc and the western terranes of the Central and Western Cordilleras drove the eastward propagation of deformation from the Central Cordillera toward the Eastern Cordillera (Restrepo-Pace et al., 2004; Parra et al., 2009). During this period, the environment of deposition in the MMV transitioned to nonmarine with intervals of fluvial and alluvial fan sandstones denuded from exhumed intervals in the flanking cordilleras (Reyes-Harker et al., 2015). Most oil and gas fields in the MMV produce from the Paleogene fluvial-sandstone reservoirs of the Colorado, Mugrosa, Esmeraldas, and La Paz Formations (Table 1). However, new play concepts targeting Cretaceous siliciclastic reservoirs are being explored (e.g., Prince et al., 2011). Oil geochemistry and oil–oil correlation in the basin can be used as tools to identify regional patterns in geochemical character and as a starting point to address unanswered questions about petroleum systems in the MMV pertinent to these new exploration campaigns.

The objectives of this study are to (1) build a large oil geochemical database across the MMV for oil–oil correlation; (2) create a regional classification of genetic petroleum "tribes" using chemometric analysis; and (3) infer the age, lithology, mixing, and depositional environment of the source rocks in the MMV using geochemical and diamondoid data for each tribe.

METHODOLOGY

A total of 106 oil samples were gathered for this study. Analyses for 84 of the oils were provided by GeoMark Research. Data included API gravity, sulfur, vanadium, nickel, stable carbon isotopes of saturate and aromatic fractions, and quantitative gas chromatography–mass spectrometry (GCMS) of saturate and aromatic fractions. An additional 22 oil samples provided by Petroleum Systems International (PSI) and Ecopetrol were analyzed at Biomarker Technologies, Inc. (BTI) for quantitative GCMS analysis of diamondoids and saturates as well as gas chromatography tandem mass spectrometry (GCMSMS) analysis. A subset of four samples were selected and processed by both laboratories to aid in the comparison of chemometric classifications

in petroleum geochemistry from Newcastle University. His interest fields include source rock prediction, crude oil characterization, and petroleum system modeling.

ROMAN EUGENIO GONZALEZ ~ *Exploration Vice-Presidency, Onshore Division, Ecopetrol, Bogotá, Colombia; roman. gonzalez@ecopetrol.com.co*

Roman Eugenio Gonzalez received a B.Sc. in geology from the Colombian Industrial University of Santander in 2001. He has worked in the Colombian oil industry since then and joined Ecopetrol in 2009. He is currently working as an exploration geologist for the onshore team in Colombia.

STEPHAN A. GRAHAM ~ *Department of Geological Sciences, Stanford University, Stanford, California; sagraham@ stanford.edu*

Stephan A. Graham is the C. Naramore Dean and W. J. and M. L. Crook Professor for the School of Earth, Energy and Environmental Sciences at Stanford University. His research foci include the tectonics of sedimentary basins, basin and petroleum systems modeling, and deep-water siliciclastic depositional systems. He is a recipient of AAPG's Sproule Award and the Grover E. Murray Memorial Distinguished Educator Award as well as the SEPM's Pettijohn Medal.

JOHN E. ZUMBERGE ~ *GeoMark Research Ltd., Houston, Texas; jzumberge@ geomarkresearch.com*

John E. Zumberge has been vice president and cofounder of GeoMark Research in Houston since 1991. He was the manager of geochemical and geological research for Cities Service–Occidental and director of geochemical services for Core Laboratories. He obtained a B.S. in chemistry from the University of Michigan and a Ph.D. in organic geochemistry from the University of Arizona.

DAVID A. WAVREK ~ *Petroleum Systems International, Inc., Salt Lake City, Utah; dwavrek@petroleumsystems.com*

David A. Wavrek is president of Petroleum Systems International, an energy services company. He actively participates in service projects for the geologic community, including technical services for the USGS, and mentoring

between the two data sets that otherwise might be obscured by differences in analytical techniques.

Laboratory Analysis

GeoMark Research completed laboratory analysis of the 84-sample set (see the appendix of Peters et al., 2007, for a description of analytical technique). Analysis of the second data set of oil samples provided by Ecopetrol and PSI was performed at BTI as follows. Small aliquots of each sample were measured and spiked using internal standards for diamondoid, aromatic, and saturate analysis. The saturate and aromatic hydrocarbons for each sample were then separated using liquid chromatography by elution of a light-hydrocarbon solvent and dichloromethane through an activated silica gel column. To avoid interference of normal alkane (n-alkane or paraffin) peaks during chromatography, an additional column elution through a zeolite ZSM-5 gel was performed on the saturate fraction. The GCMS of the prepared fractions used an Agilent GC-MSD for quantitative diamondoids and saturate biomarkers and an Agilent TQ System for GCMSMS. The analysis of diamondoids was performed using the saturate fraction of each oil sample. Each oil sample included a deuterated diamondoid internal standard added prior to the silica gel chromatography separation (He et al., 2012).

Sample Screening

The following samples were excluded from the chemometric analysis: (1) source rock extracts, (2) heavily biodegraded samples with a Peters–Moldowan (PM) biodegradation rank of 5 or more (Peters and Moldowan, 1993), and (3) highly mature samples in which biomarkers are low or absent (<10-ppm steranes). The training set excluded source rock extracts because of the difficulty in finding a discrete rock extract that reflects the composition of oil generated from the full section of thermally mature source rock. Low PM rank values indicate light to moderate levels of biodegradation and potential degradation of n-alkanes and isoprenoids. While useful, these compounds are not required for the chemometric categorization. However, as biodegradation continues, more important source-related parameters, including steranes and terpanes, are at risk. As a result, a PM rank of 5 is used as a threshold to ensure no significant degradation (Figure 4) of the sterane and terpane classes (Peters et al., 2007). Of the initial 84 samples in the data set from GeoMark Research, 29 samples were removed because of biodegradation during the screening process, resulting in 55 samples for chemometric analysis. Of the 22 samples provided by PSI and Ecopetrol, 4 were removed because of biodegradation during the screening process, resulting in 18 samples for the chemometric
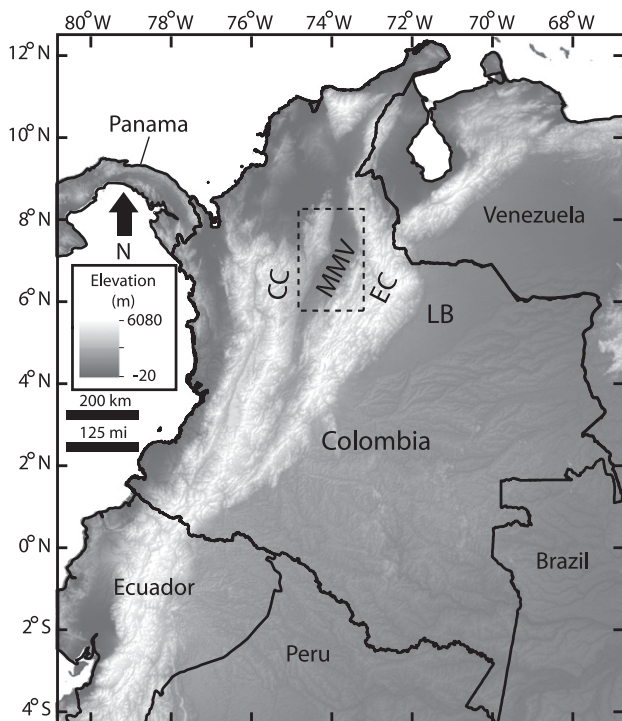
Figure 1. Map of Colombia showing elevation as well as the location of the Middle Magdalena Valley (MMV), Central Cordillera (CC), Eastern Cordillera (EC), and Llanos Basin (LB). Dashed box indicates location of MMV map used in Figures 2, 8, and 19.

analysis. Table 5 in Appendix 1 provides well location and data source information for the 73 samples used. Although there are more biodegraded samples in the central and southern portions of the MMV where reservoirs are shallower, biodegradation is ubiquitous in the basin (Ramón et al., 1997; Ramón and Dzou, 1999).

Chemometric Analysis

Chemometrics provide methods for the statistical analysis of multivariate chemical data. One benefit of chemometric analysis is the use of multidimensional space for the exploration, visualization, and categorization of multivariable data. Early applications of chemometrics in petroleum geochemistry by Christie et al. (1984), Peters et al. (1986), and Zumberge (1987) paved the way for regional-scale chemometrics studies (Peters et al., 2007, 2016). In this study, we used the chemometric software program Pirouette® from Infometrix, Inc. to perform hierarchical cluster analysis (HCA) and principal component analysis (PCA) to identify genetic affinities among oil samples from the MMV. In

addition, alternating least squares of concentration (ALS-C) is employed to identify end-members and deconvolute mixtures for a subset of the oil samples. The HCA in Pirouette was performed using autoscale preprocessing, Euclidean metric distance, and incremental linkage. The use of autoscale preprocessing normalizes each parameter value to the standard deviation for that parameter, resulting in an equal weight for each parameter during the chemometric analysis (Peters et al., 2008a).

Chemometric data analysis was used to identify distinct petroleum tribes in the MMV. In this analysis, tribe refers to a group of crude oil samples that are broadly similar with regard to their geochemical characteristics but may originate from different source rocks or organofacies of the same source rock and show systematic distributions by map location and reservoir age (e.g., Peters et al., 2016). The HCA and
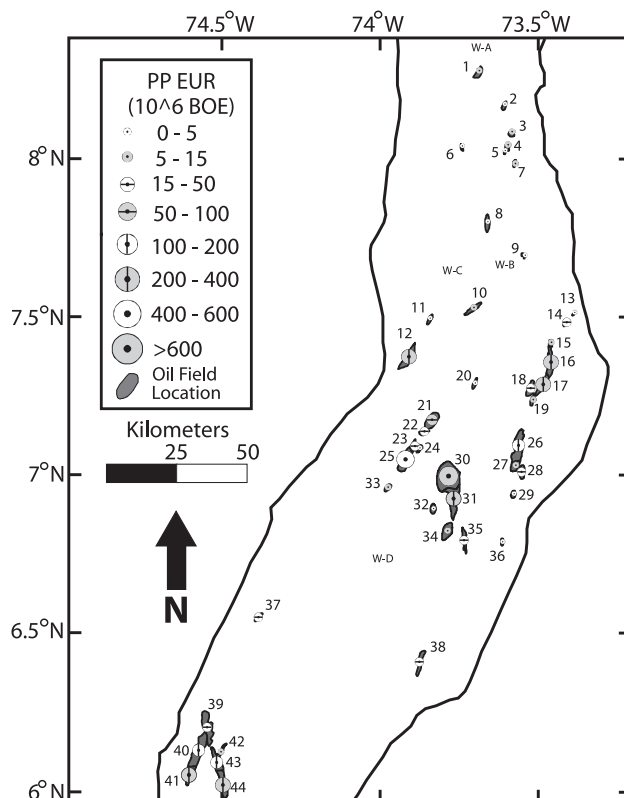
Figure 2. Map of the Middle Magdalena Valley showing the location of oil fields in addition to four wells with fluid samples used in this study (W-A to W-D). Circle sizes indicate the proven probable estimated ultimate recovery in 10[6] BOE for each field (internal Ecopetrol reporting). Numeric labels are used as map identifiers referenced in Table 1. Location of the study area is shown by the dashed box in Figure 1. EUR = estimated ultimate recovery; PP = proven probable.

Table 1. Map Identification Corresponding to Figure 2, Proven Probable Estimated Ultimate Recovery in 10^6 BOE from Internal Ecopetrol Reporting, Gas-to-Oil Ratio, Primary Reservoir Interval(s), and Number of Samples per Reservoir for Each Tribe for 44 Fields and 4 Wells in the Middle Magdalena Valley

Map ID	Field Name	PP EUR, 10^6 BOE	GOR, CF/bbl	Primary Reservoir	Tribe Distribution by Field					
					Tribe 1	Tribe 2	Tribe 3	Tribe 4	Tribe 5	Tribe 6
1	Buturama-Olivo	8.8	1000	La Luna, Rosablanca	–	–	–	–	Ros (1), CBG (3)	–
2	Totumal	1.2	318	La Luna, Rosablanca	–	–	–	LaL (1)	–	–
3	Los Angeles	9.5	88	Lisama	–	–	–	–	–	–
4	Tisquirama	11.3	182	Lisama	–	–	Lis (2)	–	–	–
5	Caiman	0.2	0	Lisama	–	–	–	Lis (2), Umi (1)	–	–
6	Learija	0.7	200	Lisama	–	–	–	Umi (1)	–	–
7	San Roque	5.2	200	Lisama	–	–	Mug (1)	–	–	–
8	Santa Lucia	4.4	235	Esmeraldas	–	–	Umi (1)	–	–	–
9	Cachira-Pavas	0.5	20	Mugrosa	–	–	–	Mug (1)	–	–
10	Cristalina	8.0	367	Esmeraldas	–	–	Lis (1)	–	–	–
11	Garzas	2.6	200	Esmeraldas	–	–	–	–	–	–
12	Cantagallo-Yarigui	270.8	500	La Paz	–	–	Mug (1), LaP (1), Esm-LaP (1)	Lis (1)	–	
13	LaTigra	0.1	8600	Lisama	Lis (1)		Unk (1)	–	–	–
14	Bonanza	30.0	1200	Esmeraldas	–	Mug (1)	–	–	–	–
15	La Liebre	8.0	857	Esmeraldas	–	–	–	–	–	–
16	Provincia-Santos	391.7	4444	Lisama	–	LaP (3), Esm (2)	–	–	–	–
17	Payoa	211.6	9249	Lisama	–	LaP (2)	LaP (4), Urk (2)	–	–	–
18	La Salina-Aguas Claras	47.5	2143	La Paz, Mugrosa	Esm (1)	Esm (1), Esm-LaP (1)	Esm (2), LaP (1), Esm-LaP (1)	–	–	–
19	Corazon	11.9	3889	Esmeraldas	–	LaP (3)	Lis (1)	–	–	–
20	Sogamoso	1.5	160	Esmeraldas	–	–	–	–	–	–
21	Llanito	52.0	568	Mugrosa		–	–	–	–	–
22	Gala	16.6	225	Mugrosa	–	–	–	–	–	–
23	Galan	34.8	531	Mugrosa	Col (2)	–	–	–	–	–
24	San Silvestre	1.8	750	Colorado	–	–	–	–	–	–
25	Casabe	455.7	588	Mugrosa	Mug (1), Col (1)	–	–	–	–	–
26	Lisama	100.2	1513	Lisama	Mug (1)	Lis (1)	–	–	–	–
27	Tesoro	12.2	2133	Colorado	–	Mug (1)	–	–	–	–
28	Nutria	18.0	1200	Mugrosa	–	Mug (1)	–	–	–	–
29	Peroles	2.8	2500	Mugrosa	–	Mug (1)	–	–	–	–
30	La Cira	749.0	967	Mugrosa	Mug (3)	–	–	–	–	–
31	Infantas	326.2	634	Mugrosa	Mug (1), Col (1)	–	–	–	–	–
32	Tenerife	0.3	400	Mugrosa	–	–	–	–	–	–
33	Penas Blancas	11.6	345	Mugrosa	–	–	–	–	–	–
34	Aguas Blancas	10.3	150	Mugrosa	Mug (1)	–	–	–	–	–
35	Colorado	16.7	4000	Mugrosa	Mug (3)	–	–	–	–	–
36	San Luis	1.0	444	Mugrosa	–	–	–	–	–	–
37	Chicala	45.0	0	Mugrosa	–	–	–	–	–	–
38	Opon	33.0	121,212	La Paz	–	–	–	–	–	LaP (1)
39	Moriche	45.8	111	Mugrosa	–	–	–	–	–	–
40	Nare-Nare Norte	165.8	219	Mugrosa	–	–	–	–	–	–

(continued)

Table 1. Continued

Map ID	Field Name	PP EUR, 10^6 BOE	GOR, CF/bbl	Primary Reservoir	Tribe Distribution by Field					
					Tribe 1	Tribe 2	Tribe 3	Tribe 4	Tribe 5	Tribe 6
41	Corcorna-Teca-Nare Sur	262.5	300	Mugrosa	–	–	Mug (1)	–	–	–
42	Ermirtano-Caipal	7.8	240	Mugrosa	–	–	–	–	–	–
43	Palagua	148.5	254	Mugrosa	–	–	–	–	–	–
44	Velasquez	217.5	523	Mugrosa	Col (2)	–	–	–	–	–
Well A	Norean	–	–	Tablazo	–	–	–	–	Tab (1)	–
Well B	Punta Piedra	–	–	Simiti	–	–	–	–	Sim (1)	–
Well C	Las Lajas	–	–	–	–	–	–	Unk (1)	–	–
Well D	Baracma	–	–	–	Unk (1)	–	–	–	–	–
–	Total	3763	–	–	20 samples	17 samples	21 samples	8 samples	6 samples	1 sample

The "Tribe Distribution by Field" columns show reservoir (number of samples) for every field and tribe.

Abbreviations: – = no samples were found and processed; CBG = Calcareous Basal Group; Col = Colorado; Esm = Esmeraldas; EUR = estimated ultimate recovery; GOR = gas-to-oil ratio; ID = identifier; LaL = La Luna; LaP = La Paz; Lis = Lisama; Mug = Mugrosa; PP = proven probable; Ros = Rosablanca; Sim = Simiti; Tab = Tablazo; Umi = Umir; Unk = unknown.

PCA, performed initially on the 55 samples provided by GeoMark Research, identify 5 geochemically distinct petroleum tribes in the MMV (HCA-1 in Figure 5). However, the 55 samples from GeoMark Research cannot be directly combined in HCA or PCA with the 18 oil samples processed at BTI because sample preparation, instrumentation, and column performance during separation can vary significantly between laboratories and impact chemometric categorization. As a result, a second HCA dendrogram, which we call HCA-2, was built for the 18 samples processed at BTI (Figure 6), and this identifies 4 geochemically distinct petroleum tribes. Of the second sample set of 18 samples, 4 were selected and sent to GeoMark Research for their standard analytical suite. After processing at GeoMark Research, these four samples (indicated by an asterisk in Figures 5 and 6 and Table 5 in Appendix 1) were incorporated into HCA-1 to help enable a more direct comparison between the tribes identified in HCA-1 and HCA-2.

Source-Related Geochemical Parameters

The chemometric analysis of the data provided by GeoMark uses 20 source-related biomarker and isotopic ratios (see Table 2). Previous geochemical studies show that certain parameters are affected by secondary processes, such as biodegradation and extreme thermal maturation. Excluding these parameters allows the chemometric study to indicate genetic relationships among oil and bitumen samples.

The selected ratios show no evidence of significant alteration by secondary processes among the training set samples. However, minor interference because of biodegradation or high thermal maturation is possible. The selected parameters are described in Peters et al. (2005) and include 13 terpane ratios, 5 sterane ratios, and 2 stable carbon isotope ratios. The terpane ratios are C_{19}/C_{23}, C_{22}/C_{21}, C_{24}/C_{23}, and C_{26}/C_{25} tricyclic terpanes; C_{24} tetracyclic terpane/C_{23} tricyclic terpane; C_{27} tetracyclic terpane/C_{27} tricyclic terpane ($C_{27}T/C_{27}$); 28,30-bisnorhopane/hopane (C_{28}/H); C_{29} 30-norhopane/hopane; C_{30} diahopane/hopane ($C_{30}X/H$); oleanane/hopane (Ol/H); C_{31} homohopane 22R/hopane; C_{27} 18α-trisnorneohopane/17α-trisnorhopane (Ts/Tm); and $C_{35}22S/C_{34}22S$ hopanes ($C_{35}S/C_{34}S$). The sterane ratios include steranes/hopanes (sterane and hopanes used for this ratio described in detail in the appendix of Peters et al., 2007), the diasterane ratio (rearranged steranes to regular steranes) and %C_{27}, %C_{28}, and %C_{29} (based on 5α,14β,17β steranes from mass-to-charge ratio [m/z] 218). The stable carbon isotope ratios include $\delta^{13}C_{aromatics}$ and $\delta^{13}C_{saturates}$. Table 2 includes average values and standard deviations by tribe for each of the biomarker and isotopic ratios included in the chemometric analysis of the data from GeoMark Research. Other parameters in the table are API gravity, weight percent sulfur, saturate/aromatic ratio, the triaromatic steroid cracking ratio (TAS3[CR]; modified from Mackenzie et al., 1981, as described in Peters et al., 2005), dibenzothiophene/phenanthrene (DBT/P) ratio (Hughes
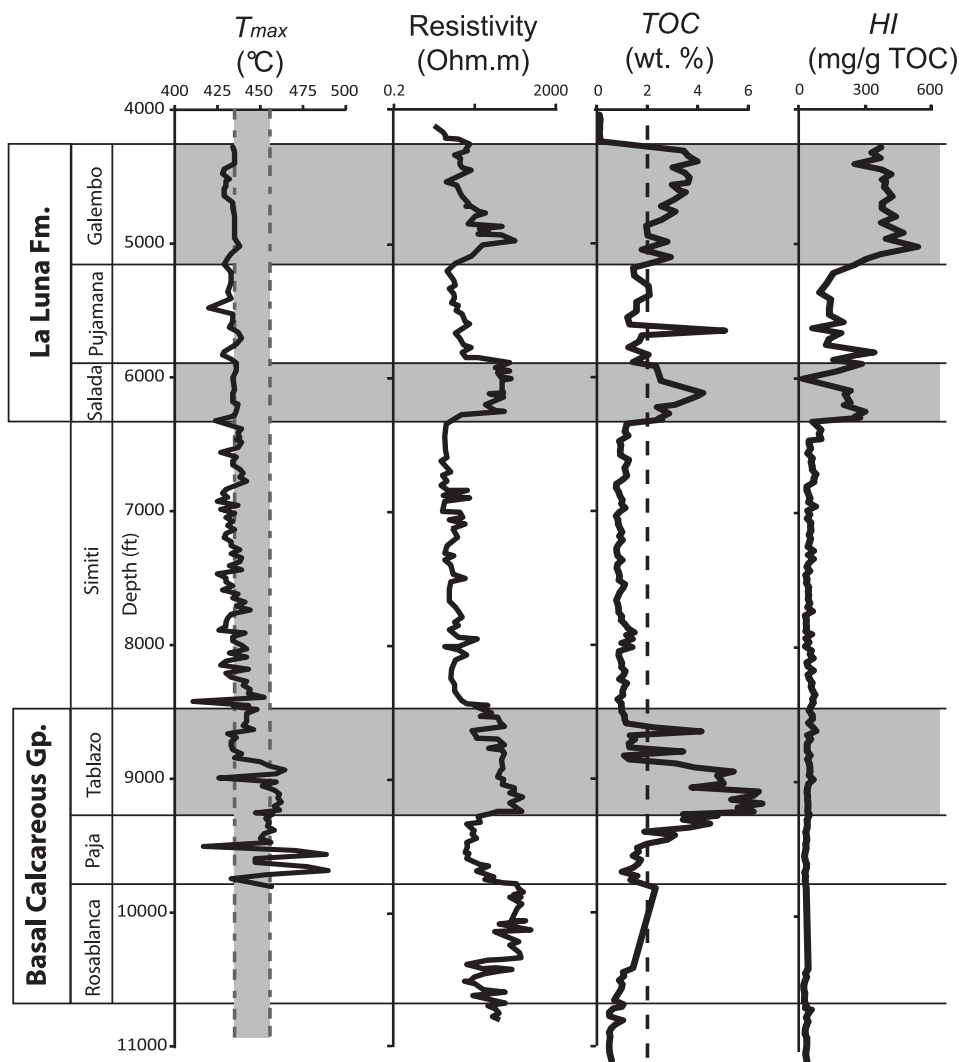
Figure 3. Logs including Rock-Eval T_{max} (the temperature corresponding to the maximum production of hydrocarbon products by kerogen cracking during pyrolysis), resistivity, total organic carbon (*TOC*), and hydrogen index (*HI*) for the Cretaceous section of a well in the central Middle Magdalena Valley. A *TOC* of 2 wt. % and T_{max} range of 435°C to 455°C are demarcated by dashed lines. Fm. = Formation; Gp. = Group.

et al., 1995), vanadium/nickel (V/Ni) ratio (Lewan, 1984), pristane/phytane (Pr/Ph) ratio, the C_{29}-Ts/C_{29}-hopane (C_{29}D/H) ratio, and the C_{26}/C_{28} 20S triaromatic steroid (C_{26}/C_{28} STAS) ratio. Figure 7 gives an *m/z* 191 chromatogram for six oils from the MMV and illustrates differences in the numbered saturate peaks (Table 3) between the tribes identified in the chemometric analysis.

The second chemometric analysis of oils performed at BTI included all of the parameters used in the first analysis with the exception of the stable carbon isotope ratios and C_{27}T/C_{27}. In addition, 11 parameters from the more extensive GCMS and GCMSMS analyses were included (see Table 4 and Appendix 2 for parameter values and definitions).

These parameters were 3β-methylhopane index, $\alpha\alpha\alpha C_{29}$(20R)/total $\alpha\alpha\alpha(C_{27} + C_{28} + C_{29})$, total C_{29} steranes/total ($C_{27} + C_{28} + C_{29}$), extended tricyclics ratio ($C_{28} + C_{29}$ tricyclics/18α-trisnorneohopane), homohopane index (HHI), C_{25}–highly branched isoprenoids (C_{25}-HBI), 2/3-methylhopane index, 21-norcholestane ratio (21-NCR), 24-nordiacholestane ratio (24-NDR), bicadinane index, and diahopane index. The addition of 11 parameters helped bolster the chemometric analysis, leading to a more robust geochemical categorization. Table 4 includes average values and standard deviations by tribe for each of the biomarker ratios included in the chemometric analysis of the data set processed at BTI. Other parameters included in the table include
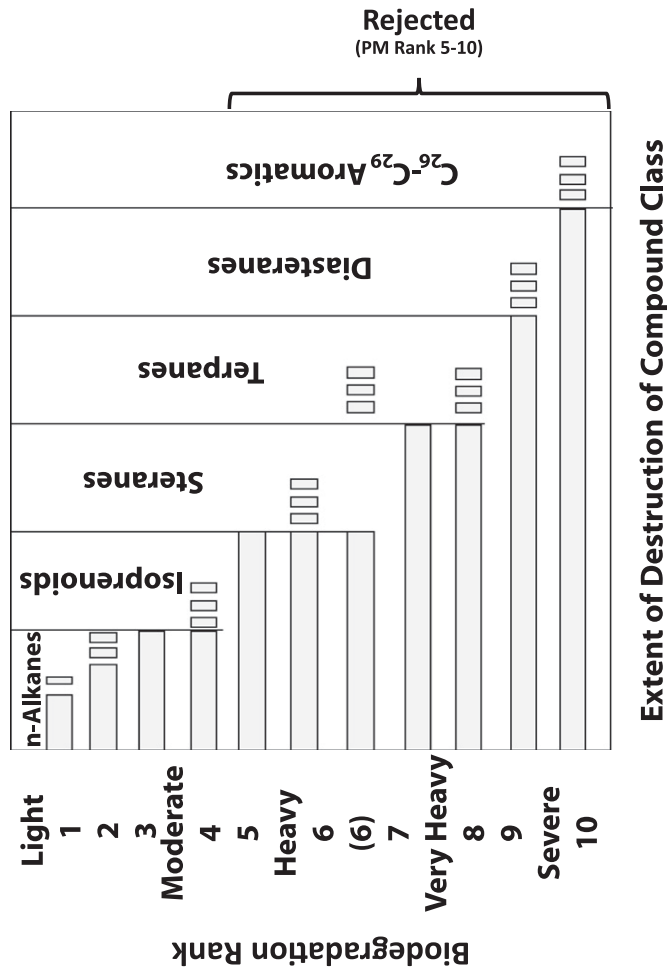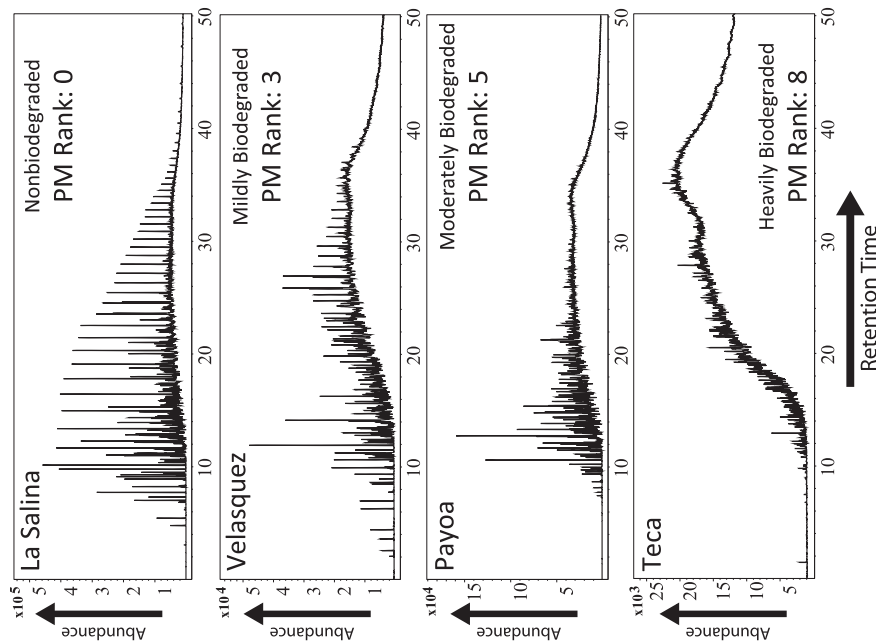
Figure 4. (Left) Oil samples from the La Salina, Velasquez, Payoa, and Teca fields illustrate the impact of various levels of biodegradation in the Middle Magdalena Valley. The Teca and Payoa oil samples were excluded from the chemometric analysis because of the potential impact of biodegradation on steranes, terpanes, and diasteranes used for classification. (Right) Peters–Moldowan (PM) biodegradation rank scale (modified from Peters and Moldowan, 1993) used to screen and select oil for the chemometric analysis. Oil samples with a PM rank of 5 or greater were rejected from the chemometric analysis.
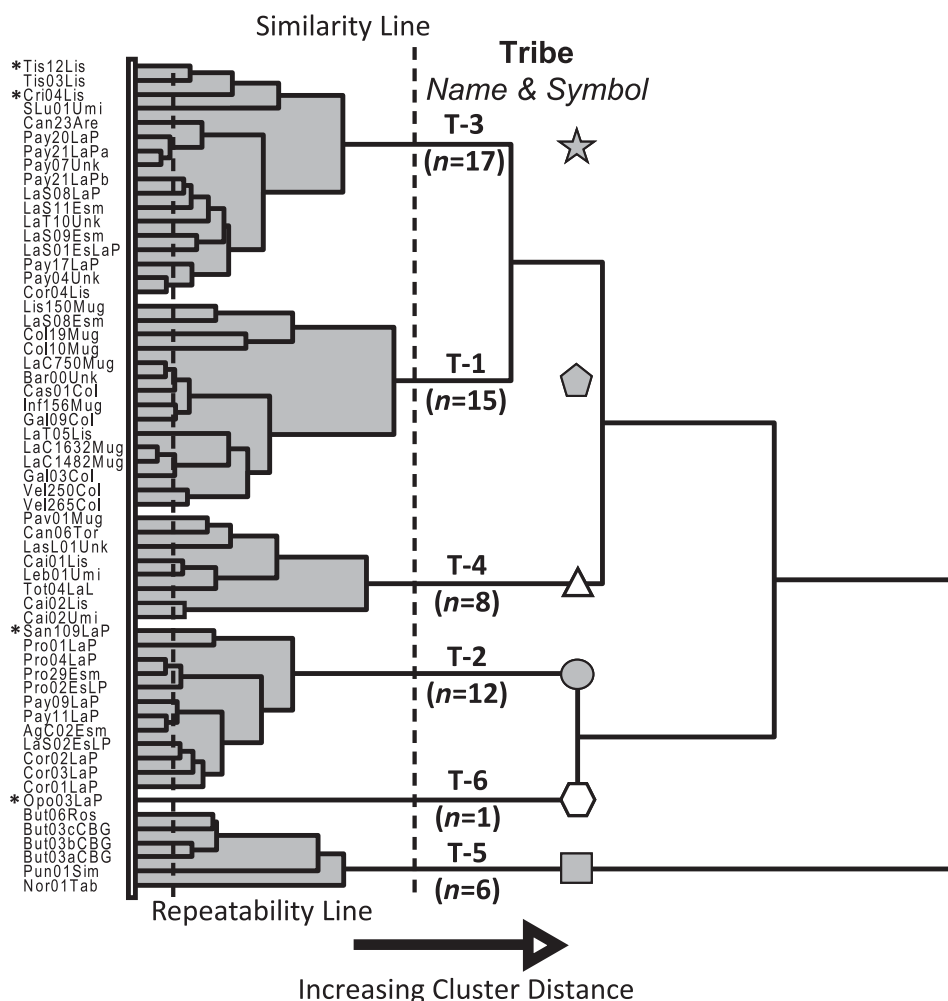
Figure 5. Hierarchical cluster analysis (HCA) dendrogram for the oil database provided by GeoMark Research; we designate this classification as HCA-1. This HCA-1 was built using 20 biomarker and isotopic ratios shown in Table 2. It identifies six total tribes in the Middle Magdalena Valley with "T" indicating the tribe name for each cluster and "n" indicating the number of samples per tribe. The map symbology for each tribe is given to the right of each cluster and the cluster distance for differentiating tribes is indicated by the dashed similarity line. Samples marked by an asterisk were processed by both Biomarker Technologies, Inc. and GeoMark Research and are included in both HCA-1 and HCA-2. Analytical repeatability, indicated by the repeatability line, is based on the Pay20LaP and Pay21LaPa samples in tribe 3 as well as the Pay09LaP and Pay11LaP samples in tribe 2. Samples with a cluster distance that is greater than the repeatability line are considered geochemically distinct.

gammacerane/$C_{31}R$ as well as parameters used in the diamondoid analysis, including 1- + 2-methyl adamantane, 3- + 4-methyl diamantane, and C_{29}-20R stigmastane.

RESULTS AND DISCUSSION

Tribe Map and Reservoir Distributions

Chemometric data analysis within the MMV classifies six distinct petroleum tribes: three major tribes identified in both HCAs (tribes 1, 2, and 3), two

smaller tribes unique to HCA-1 (tribes 4 and 5), and a single-sample tribe (tribe 6) that is distinct in both HCAs. Preliminary observations show that these tribes, identified by chemometric categorization, vary systematically by location and primary reservoir interval (Figure 8). Trends in geographic distribution and sample reservoir interval among tribes identified in both HCA-1 and HCA-2 are similar for the two data sets. Figures 9 and 10 give the reservoir distribution for samples from each tribe for HCA-1 and HCA-2, respectively. Tribes 1 and 2 are generally located in Oligocene and Eocene reservoirs, respectively, in the central part of the study area, with tribe 2 localized
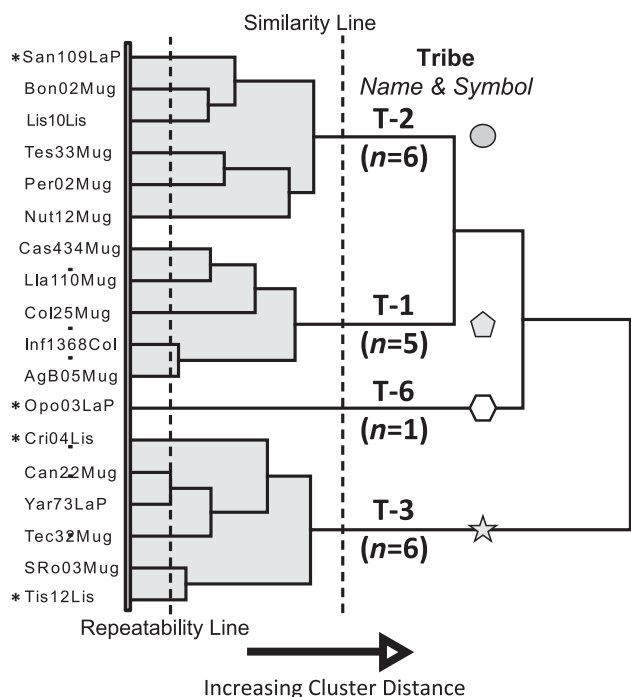
Figure 6. Hierarchical cluster analysis (HCA) dendrogram for the oil database processed at Biomarker Technologies, Inc. (BTI); we designate this classification as HCA-2. This HCA-2 was built using 28 biomarker parameters shown in Table 3. It identifies four tribes in the Middle Magdalena Valley (MMV) with "T" indicating the tribe name for each cluster and "*n*" indicating the number of samples per tribe. The map symbology for each tribe is given to the right of each cluster and the cluster distance for differentiating tribes is indicated by the dashed similarity line. Samples marked by an asterisk were processed by both BTI and GeoMark Research and are included in both HCA-1 and HCA-2. The repeatability line is based on the Can22Mug and Yar73LaP samples, both from the Cantagallo–Yarigui field in the Western MMV.

along the eastern edge of the basin near the Eastern Cordillera thrust front. Tribe 5 is located in Lower to middle Cretaceous reservoirs in the northern part of the study area. Tribe 4 is also located in the northern part of the study area and is found primarily in Upper Cretaceous to lower Paleogene reservoir intervals. The HCA-2 data set has no samples in the northern part of the study area from Cretaceous to lower Paleogene reservoirs, and, as a result, tribes 4 and 5 are not represented. Tribe 3 occurs between tribes 4 and 2, both regionally and in terms of reservoir interval. Finally, tribe 6 is a single sample from the La Paz Formation in the Opon field located in the southern part of the MMV near the western flank of the Eastern Cordillera.

Although broadly similar, each tribe shows different ranges of values for selected biomarker and isotopic ratios as well as bulk properties, such as API gravity, sulfur content, saturate/aromatic ratio, and weight percentage $<C_{15}$. These differences can be used to infer changes in source rock depositional environment, lithology, or thermal maturity within the source rocks of the MMV. Tribes 5 and 6 are the most distinct tribes based on PCA. Whereas stable carbon isotope ratios for the saturate and aromatic fractions of the oil samples in tribes 1 through 5 indicate marine organic matter input, the tribe 6 oil is the only sample indicating terrigenous organic matter input (Figure 11) (Sofer, 1984). The primary inferred source rocks in the MMV, the La Luna and Tablazo Formations, are marine calcareous units with little terrigenous input (Zumberge, 1984; Rangel et al., 2000) Whereas tribes 1 through 5 likely originated from these Cretaceous marine source rocks within the MMV, tribe 6 might represent either extra-basinal charge from thrust-stacked intervals of more proximal and terrigenous source units in the Eastern Cordillera or an extension of the estuarine Upper Cretaceous Umir Formation (Rangel et al., 2002). Tribe 5 differs from tribes 1 through 4 by a combination of lithologic and maturity-driven indicators discussed in more detail below. Although geochemical and diamondoid data for each tribe allows inferences on the age, lithology, mixing, and depositional environment of the source rocks in the MMV, an extensive oil–source rock extract study is needed to more explicitly establish correlations.

Alternating Least Squares to Deconvolute Mixtures

Mixture analysis employing alternating least squares (ALS) in Pirouette requires compound concentrations (Peters et al., 2008b). The ALS-C analysis for 48 biomarkers from the GeoMark database was used to deconvolute three MMV petroleum tribes within the black oil molecular weight range. The oil samples comprising petroleum tribe 3 fall between tribes 2 and 4 in terms of geographic location and reservoir rock interval (Figures 8, 9). In addition, for most biomarker parameters used in this study, values for tribe 3 fall between those of tribes 2 and 4 (Table 2). Finally, the linear trend from tribe 2 through tribe 3 and on to tribe 4 in principal component space is similar to distributions observed in other mixture analysis studies (Peters et al., 2008b) and again

Table 2. Biomarker and Isotopic Ratios Used in the Chemometric Analysis of the Data Set from GeoMark Research

Geochemical Variable	Middle Magdalena Valley Oil Tribes					
	Tribe 1 (15)	Tribe 2 (12)	Tribe 3 (17)	Tribe 4 (8)	Tribe 5 (6)	Tribe 6 (1)
API gravity	27.65 ± 10.07	33.88 ± 4.12	28.15 ± 5.61	19.13 ± 6.30	29.18 ± 7.55	−
Wt. % sulfur	0.74 ± 0.28	0.70 ± 0.15	1.13 ± 0.36	1.49 ± 0.17	0.71 ± 0.35	0.08
Sat/Aro	1.51 ± 0.68	1.77 ± 0.25	1.18 ± 0.30	0.89 ± 0.10	1.42 ± 0.31	5.42
C_{19}/C_{23}*	0.05 ± 0.02	0.07 ± 0.02	0.03 ± 0.01	0.02 ± 0.00	0.09 ± 0.02	0.30
C_{22}/C_{21}*	0.30 ± 0.03	0.38 ± 0.03	0.39 ± 0.04	0.42 ± 0.05	0.60 ± 0.05	0.28
C_{24}/C_{23}*	0.96 ± 0.08	0.85 ± 0.04	0.84 ± 0.08	0.69 ± 0.06	0.58 ± 0.08	1.05
C_{26}/C_{25}*	0.76 ± 0.04	0.80 ± 0.03	0.78 ± 0.03	0.72 ± 0.04	0.76 ± 0.07	0.85
Tet/C_{23}*	0.16 ± 0.02	0.25 ± 0.03	0.16 ± 0.02	0.12 ± 0.01	0.23 ± 0.01	0.26
$C_{27}T/C_{27}$*	0.36 ± 0.07	0.18 ± 0.05	0.37 ± 0.08	0.38 ± 0.08	0.04 ± 0.02	0.17
C_{28}/H*	0.10 ± 0.03	0.06 ± 0.02	0.13 ± 0.01	0.14 ± 0.03	0.05 ± 0.01	0.05
C_{29}/H*	0.64 ± 0.03	0.59 ± 0.02	0.64 ± 0.03	0.75 ± 0.03	0.56 ± 0.04	0.54
$C_{30}\times/H$*	0.05 ± 0.02	0.05 ± 0.01	0.03 ± 0.01	0.01 ± 0.00	0.10 ± 0.04	0.34
Ol/H*	0.05 ± 0.01	0.08 ± 0.01	0.04 ± 0.01	0.02 ± 0.01	0.05 ± 0.01	0.16
$C_{31}R/H$*	0.30 ± 0.03	0.29 ± 0.03	0.39 ± 0.04	0.42 ± 0.02	0.34 ± 0.02	0.17
$Ga/C_{31}R$	0.59 ± 0.15	0.38 ± 0.10	0.43 ± 0.16	0.42 ± 0.05	0.50 ± 0.11	0.23
$C_{35}S/C_{34}S$*	0.68 ± 0.07	0.71 ± 0.07	0.87 ± 0.07	0.89 ± 0.03	0.98 ± 0.17	0.98
Ster/Hop*	0.73 ± 0.21	0.39 ± 0.05	0.58 ± 0.08	0.78 ± 0.07	1.53 ± 0.27	0.82
Rearr/Reg*	0.88 ± 0.40	0.67 ± 0.14	0.51 ± 0.13	0.44 ± 0.05	1.44 ± 0.37	2.33
$\%C_{27}$*	37.96 ± 1.19	35.67 ± 1.15	36.84 ± 0.82	40.51 ± 1.00	38.33 ± 1.04	31.34
$\%C_{28}$*	35.55 ± 1.37	34.95 ± 0.55	35.10 ± 0.95	33.21 ± 1.06	30.08 ± 1.06	36.78
$\%C_{29}$*	26.47 ± 1.06	29.38 ± 1.16	28.06 ± 1.17	26.30 ± 0.80	31.57 ± 0.63	31.88
Ts/Tm*	0.72 ± 0.22	0.70 ± 0.05	0.54 ± 0.15	0.31 ± 0.06	4.96 ± 1.48	1.98
$\delta^{13}C_{saturates}$*	−27.48 ± 0.17	−27.63 ± 0.11	−27.78 ± 0.25	−28.12 ± 0.08	−26.67 ± 0.20	−27.26
$\delta^{13}C_{aromatics}$*	−27.04 ± 0.17	−27.06 ± 0.08	−27.23 ± 0.13	−27.54 ± 0.06	−26.08 ± 0.26	−26.04
Pr/Ph	1.20 ± 0.26	1.57 ± 0.17	1.01 ± 0.24	0.73 ± 0.05	0.75 ± 0.07	2.14
DBT/P	0.58 ± 0.28	0.56 ± 0.08	0.89 ± 0.19	1.09 ± 0.20	0.78 ± 0.39	0.09
$C_{29}D/H$	0.22 ± 0.04	0.27 ± 0.02	0.20 ± 0.04	0.12 ± 0.02	0.96 ± 0.20	0.60
TAS3(CR)	0.19 ± 0.11	0.14 ± 0.02	0.13 ± 0.03	0.08 ± 0.02	0.40 ± 0.05	0.76
$C_{26}/C_{28}STAS$	0.35 ± 0.10	0.27 ± 0.02	0.23 ± 0.04	0.12 ± 0.05	0.46 ± 0.14	0.72
V/(Ni + V)	0.74 ± 0.08	0.76 ± 0.10	0.81 ± 0.11	0.88 ± 0.03	0.64 ± 0.03	−

Values are listed as the mean plus or minus the standard deviation for each tribe. The number of samples measured for each tribe is given in parentheses in the column heading for that tribe.

Abbreviations: − = not enough sample to process; $C_{26}/C_{28}STAS = C_{26}/C_{28}$ 20S triaromatic steroid ratio; $C_{29}D/H = C_{29}$-18α-trisnorneohopane (Ts)/C_{29}-hopane; DBT/P = dibenzothiophene/phenanthrene; Ga = gammacerane; OI/H = oleanane/hopane; Pr/Ph = pristane/phytane; Rearr/Reg = diasterane ratio; Sat/Aro = saturates/aromatics; Ster/Hop = sterane/hopane ratio; TAS3(CR) = triaromatic steroid cracking ratio; Ts/Tm = C_{27} 18α-trisnorneohopane/17α-trisnorhopane.

*Parameters used in hierarchical cluster analysis 1 (Figure 5).

suggests a mixing trend (Figure 12). Although these qualitative observations suggest a mixture, ALS-C analysis provides a robust measure of the relative amounts of two or more oil types in a mixture without requiring the identification of pure end-member oil prior to the analysis. This methodology is built on the observation that the concentrations of compounds in oil mixtures within the black oil molecular weight range vary as linear functions of the amount of each

oil type, thus allowing quantitative mixture deconvolution (Peters et al., 2008b). Conditions for this analysis included specification that the sum of all concentrations remains constant and that the amount of each compound must be nonnegative.

The results of the ALS-C analysis support the mixture hypothesis. The tribe 3 samples plot as mixtures of end-member contributions from tribes 2 and 4 for the biomarker parameters selected within
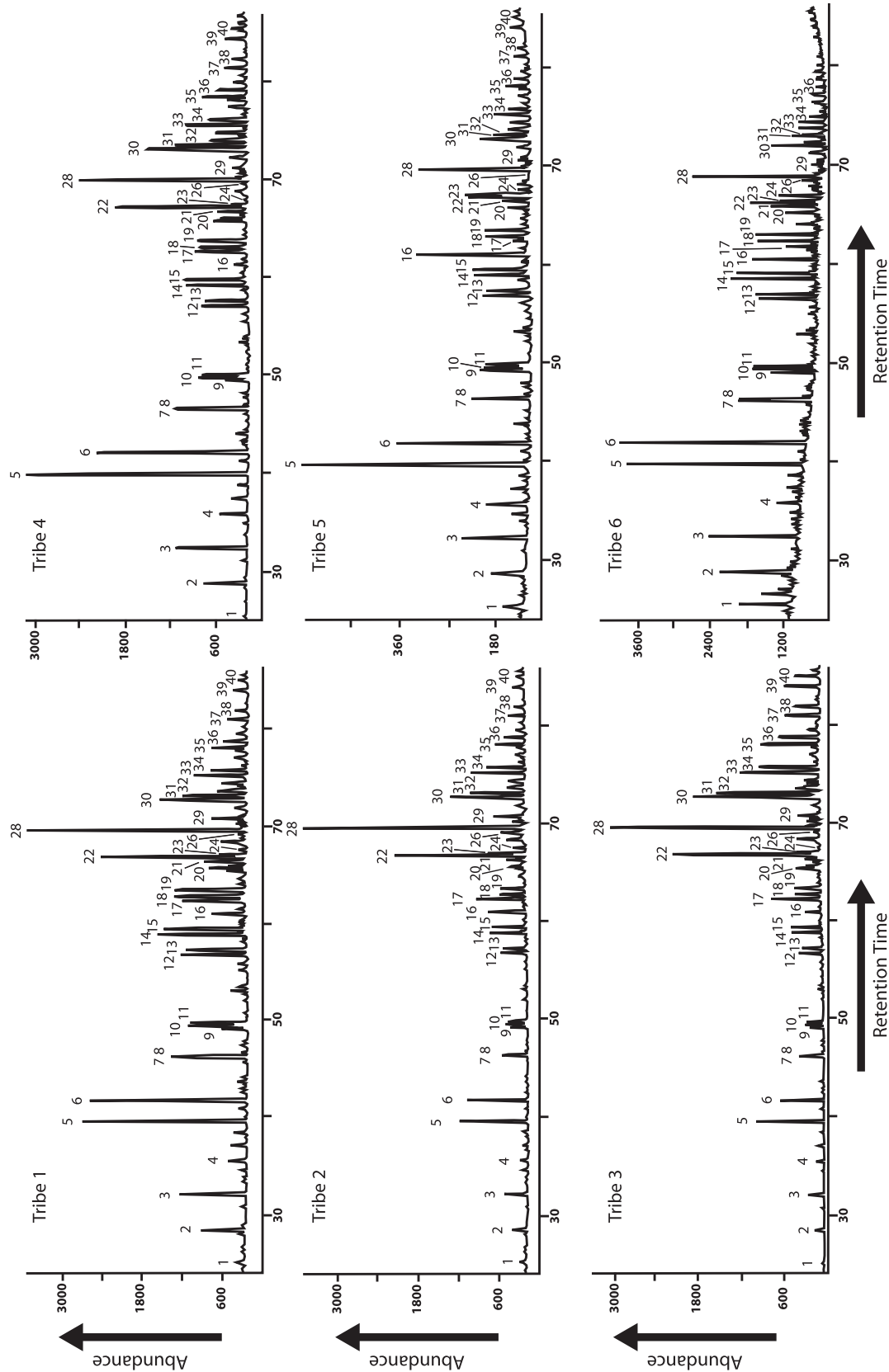
Figure 7. Mass-to-charge ratio 191 chromatograms from a representative oil from each tribe identified in the chemometric analysis. The selected saturate compounds corresponding to each numbered peak are given in Table 3. Variations in abundance between peaks from different tribes reflect differences in source rock maturity, lithology, and depositional setting, among other characteristics. Taken together, these differences are used to identify the petroleum tribes in the Middle Magdalena Valley.

Table 3. The Full Compound Names Corresponding to the Select Labeled Saturate Peaks from the Mass-to-Charge Ratio 191 Chromatograms in Figure 7

Peak No.	Compound
1	C_{19} tricyclic diterpane
2	C_{20} tricyclic diterpane
3	C_{21} tricyclic diterpane
4	C_{22} tricyclic terpane
5	C_{23} tricyclic terpane
6	C_{24} tricyclic terpane
7	C_{25} tricyclic terpane S
8	C_{25} tricyclic terpane R
9	C_{24} tetracyclic terpane
10	C_{26} tricyclic terpane S
11	C_{26} tricyclic terpane R
12	C_{28} extended tricyclic terpane S
13	C_{28} extended tricyclic terpane R
14	C_{29} extended tricyclic terpane S
15	C_{29} extended tricyclic terpane R
16	Ts 18α-trisnorhopane
17	Tm 17α-trisnorhopane
18	C_{30} extended tricyclic terpane S
19	C_{30} extended tricyclic terpane R
20	C_{28} 17α, 18α, 2lβ-bisnorhopane
21	C_{29} 25-norhopane
22	C_{29} Tm 17α, 21β-norhopane
23	C_{29} Ts 18α-norneohopane
24	C_{30} 17α-diahopane
26	18α-oleanane
28	C_{30} 17α-hopane
29	C_{30} moretane
30	C_{31} 17α-homohopane 22S
31	C_{31} 17α-homohopane 22R
32	Gammacerane
33	C_{32} 17α-bishomohopane 22S
34	C_{32} 17α-bishomohopane 22R
35	C_{33} 17α-trishomohopane 22S
36	C_{33} 17α-trishomohopane 22R
37	C_{34} 17α-extended hopane 22S
38	C_{34} 17α-extended hopane 22R
39	C_{35} 17α-extended hopane 22S
40	C_{35} 17α-extended hopane 22R

the black oil molecular weight range (Figure 13). Figure 13 orders the samples geographically from southeast to northwest. The tribe 2 contribution generally appears more dominant than the tribe 4 contribution in most tribe 3 samples. However, tribe 3 samples in the north and northwestern MMV have a more significant contribution from a tribe 4 source than in the central and eastern portions of the basin.

As Peters et al. (2008b) acknowledged, biomarker concentrations can be impacted by reservoir processes, such as thermal cracking and biodegradation. Tribes 2, 3, and 4 have similarly low levels of biodegradation, reducing the likelihood of significant alteration by postdepositional processes. However, the present-day burial depth increases for the samples, from deeper burial in the area of tribe 2 to shallower burial in the area of tribe 4. The increased present-day burial depth in the area of tribe 2 allows the possibility of input from a more deeply buried source mixing with tribe 2 samples but falls outside of the purview of this mixture analysis because such a source likely would not be in the black oil molecular weight range.

Inferred Source Rock Depositional Environment, Lithology, and Age

Several biomarker ratios in Tables 2 and 4 are used to describe the source rock depositional environment and lithology of each petroleum tribe. Used in combination, these parameters can help identify petroleum generated from source rock deposited under more reducing or anoxic conditions as well as those likely deposited in more oxic environments. These categorizations are not specific enough to establish the more rigorous classifications for oxygenation levels defined by Tyson and Pearson (1991). However, they identify differences in relative oxicity that are important in breaking out the tribes in the MMV. In addition, the biomarker ratios can suggest source rock lithology by indicating varying degrees of terrigenous input as well as clay-rich or clay-poor sources. To simplify these lithologic classifications, we use the qualitative categories of shale, marl, and carbonate while acknowledging that the source rocks of the MMV exhibit a broader array of facies. In addition, whereas the mean values of a parameter for a tribe might suggest a specific classification, values within that tribe can vary as a result of regional facies changes or secondary processes, such as thermal maturity or biodegradation.

Bulk properties like weight percentage of sulfur, API gravity, and the ratio of saturates to aromatics

Table 4. Biomarker and Isotopic Ratios Used in the Chemometric Analysis of the Data Set from Ecopetrol and Petroleum Systems International in Addition to Diamondoid Values

Geochemical Variable	Middle Magdalena Valley Oil Tribes			
	Tribe 1 (5)	Tribe 2 (6)	Tribe 3 (6)	Tribe 6 (1)
$C_{19}/(C_{19} + C_{23})$*	0.11 ± 0.03	0.15 ± 0.04	0.04 = 0.01	0.48
$C_{22}/(C_{22} + C_{21})$*	0.26 ± 0.01	0.30 ± 0.02	0.35 = 0.02	0.23
C_{24}/C_{23}*	1.02 ± 0.06	0.88 ± 0.05	0.78 ± 0.07	1.10
C_{26}/C_{25}*	0.87 ± 0.03	0.90 ± 0.01	0.83 ± 0.03	0.92
Tet/C_{23}*	0.34 ± 0.06	0.45 ± 0.13	0.28 = 0.02	0.48
C_{28}/H*	0.13 ± 0.04	0.07 ± 0.03	0.17 = 0.03	0.03
C_{29}/H*	0.04 ± 0.01	0.03 ± 0.01	0.06 ± 0.01	0.02
C_{31}R/H*	0.29 ± 0.03	0.28 ± 0.02	0.44 ± 0.03	0.20
C_{35}S/C_{34}S*	0.88 ± 0.09	0.81 ± 0.11	1.16 = 0.03	0.57
Ga/C_{31}R	0.10 ± 0.01	0.07 ± 0.01	0.09 ± 0.01	0.11
C_{26}/Ts*	1.77 ± 0.36	1.11 ± 0.37	2.29 = 0.29	1.34
%C_{27}*	0.28 ± 0.01	0.27 ± 0.01	0.28 ± 0.01	0.25
%C_{28}*	0.42 ± 0.01	0.41 ± 0.01	0.41 ± 0.01	0.41
%C_{29}*	0.30 ± 0.01	0.32 ± 0.02	0.30 = 0.01	0.34
Rearr/Reg*	0.29 ± 0.02	0.32 ± 0.04	0.21 = 0.02	0.49
Ts/Tm*	0.47 ± 0.06	0.44 ± 0.03	0.31 ± 0.05	0.73
ETR*	4.67 ± 0.86	2.69 ± 0.67	5.92 = 0.66	2.93
HHI*	0.04 ± 0.01	0.06 ± 0.01	0.06 ± 0.02	0.01
C_{25}-HBI*	50.06 ± 8.74	53.54 ± 10.17	74.76 = 10.11	21.05
OI-index*	0.05 ± 0.01	0.10 ± 0.01	0.03 = 0.01	0.20
2/3-Methylhopane*	0.72 ± 0.02	0.64 ± 0.05	0.64 = 0.04	0.83
3-Methylhopane ratio*	0.006 ± 0.001	0.004 ± 0.000	0.002 = 0.001	0.02
3β-Methylhopane index*	0.009 ± 0.001	0.012 ± 0.003	0.013 = 0.002	0.01
$\alpha\alpha\alpha C_{29}$20R/total $\alpha\alpha\alpha$(C27-C29)*	0.24 ± 0.02	0.33 ± 0.05	0.26 = 0.02	0.36
Total C_{29}/total (C$_{27}$-C$_{29}$)*	0.27 ± 0.02	0.31 ± 0.02	0.29 = 0.01	0.38
24-NDR*	0.19 ± 0.01	0.19 ± 0.01	0.19 ± 0.01	0.22
Bicadinane index*	0.19 ± 0.08	0.22 ± 0.12	0.03 = 0.01	0.83
21-NCR*	0.22 ± 0.02	0.23 ± 0.02	0.17 = 0.01	0.29
Diahopane index*	0.05 ± 0.01	0.04 ± 0.01	0.02 ± 0.00	0.25
1- + 2-Methyl adamantane	41.55 ± 8.51	56.19 ± 5.52	6.36 = 2.41	276.15
3- + 4-Methyl diamantane	2.07 ± 0.56	2.09 ± 0.40	0.44 = 0.18	10.25
Stigmastane 20R	20.52 ± 5.47	15.72 ± 4.93	31.21 = 6.01	1.04

Values are listed as the mean plus or minus the standard deviation for each tribe. The number of samples measured for each tribe is given in parentheses in the column heading for that tribe.

Abbreviations: NCR = norcholestane ratio; NDR = nordiacholestane ratio; HBI = highly branched isoprenoid; ETR = extended tricyclics ratio; Ga = gammacerane; HHI = homohopane index; OI = oleanane; Rearr = rearranged steranes; Reg = regular steranes; Tet = C_{24} tetracyclic terpane; Tm = 17α-trisnorhopane; Ts = 18α-trisnorneohopane.

*Parameters used in the hierarchical cluster analysis 2 (Figure 6).

can help differentiate lithology and depositional environments in oils. The high sulfur contents for samples in tribe 4 (~1.49 wt. %) relative to tribes 1 and 2 (~0.74 and 0.70 wt. %, respectively) are indicative of more reducing conditions (Orr, 2001). The tribe 6 oil has the lowest sulfur content (0.08 wt. %), again

reflective of a nonmarine organic matter type. According to classifications by Orr (2001), tribes 2, 3, 4, and 5 correspond with medium sulfur kerogen type, whereas tribe 1 has lower sulfur content (Figure 14). Meanwhile, the higher values of the saturate/aromatic ratio in tribes 1 and 2 (~1.51 and ~1.77,
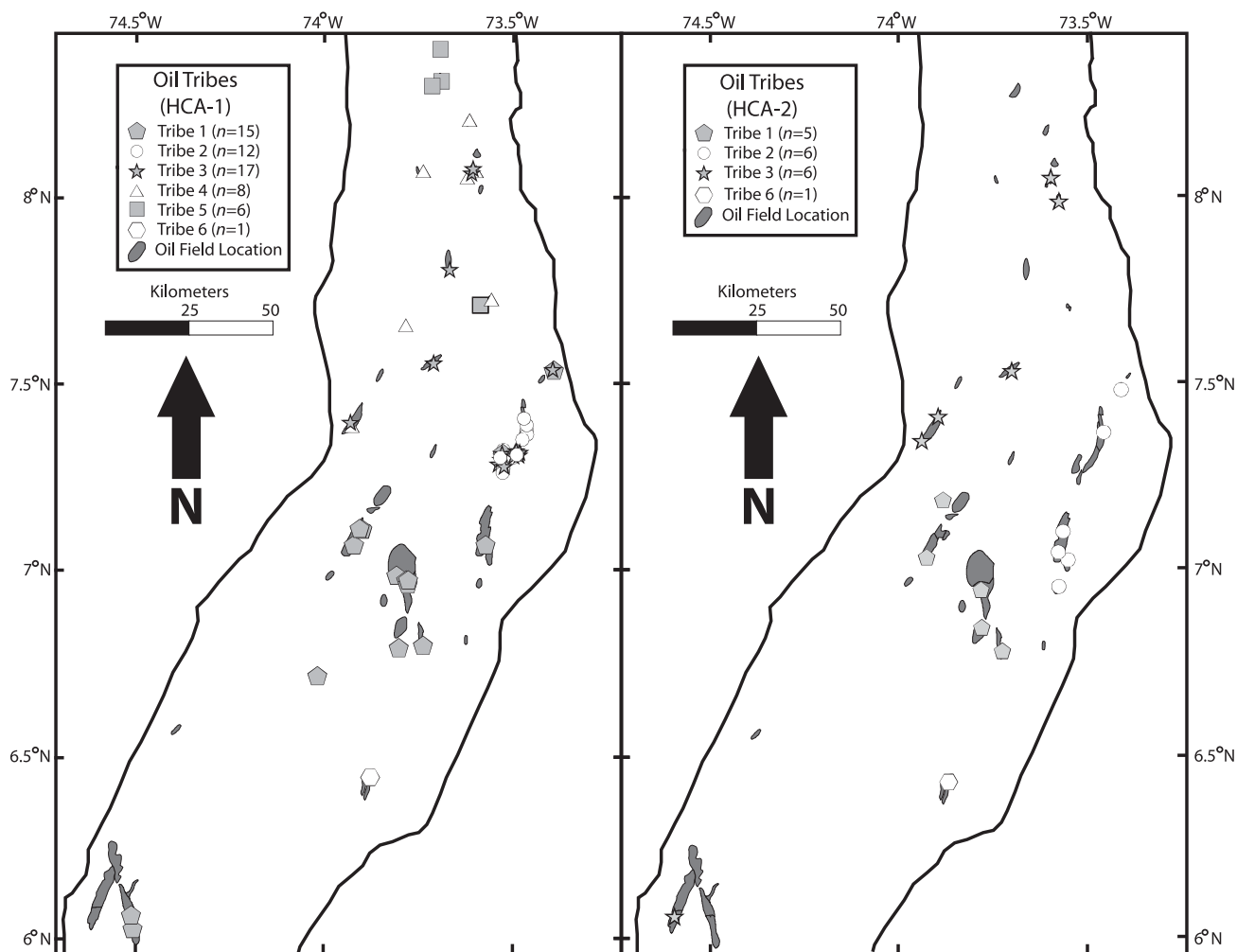
Figure 8. (Left) Map of the Middle Magdalena Valley (MMV) showing the locations of oil samples used in hierarchical cluster analysis (HCA)-1, with symbols indicating the petroleum tribe of each sample. Gray polygons are the locations of the primary oil fields in the MMV. For oil field name and size, refer to Figure 2 and Table 1. (Right) Map of the MMV showing the locations of oil samples used in HCA-2, with symbols indicating the petroleum tribe of each sample. Tribes 4 and 5 are not represented here, likely because no samples from Cretaceous to early Paleogene reservoirs were included in the data set for HCA-2. Location of the study area is shown by the dashed box in Figure 1. For both maps, "*n*" indicates the number of samples from each tribe in the figure.

respectively) are more likely to originate from shale, whereas the lower ratio (~0.89) for tribe 4 suggests a carbonate source (Waples and Curiale, 1999). However, the impact of thermal maturity on bulk parameters has been noted by several authors (Orr, 2001; Peters et al., 2016) and thus limits the use of these parameters.

The analysis of a database of more than 500 crude oils performed by GeoMark Research demonstrated that C_{22}/C_{21} and C_{24}/C_{23} tricyclic terpane ratios could be a useful tool to determine source rock lithology (Peters et al., 2005). High values of C_{24}/C_{23} tricyclic terpanes in addition to low values of C_{22}/C_{21} tricyclic terpanes for tribe 1, tribe 2, and tribe 6

indicate shale source lithology, whereas the inverse relationships for tribes 4 and 5 indicate marl to carbonate lithology (Figures 15 and 7). The DBT/P and Pr/Ph ratios corroborate this classification. Hughes et al. (1995) found that high values of Pr/Ph (>1) and low values of DBT/P (<1) are typical of a marine shale source rock, whereas the inverse is typical of a marine carbonate to marl source rock. High DBT/P and low Pr/Ph for tribes 4 and 5 indicate, tentatively, a marine marl to carbonate source, whereas the low DBT/P and high Pr/Ph ratios for tribes 1, 2, and 6 indicate a shale source (Table 2). In both cases, tribe 3 shows values between tribes 2 and 4, further supporting a mixture hypothesis for tribe 3.
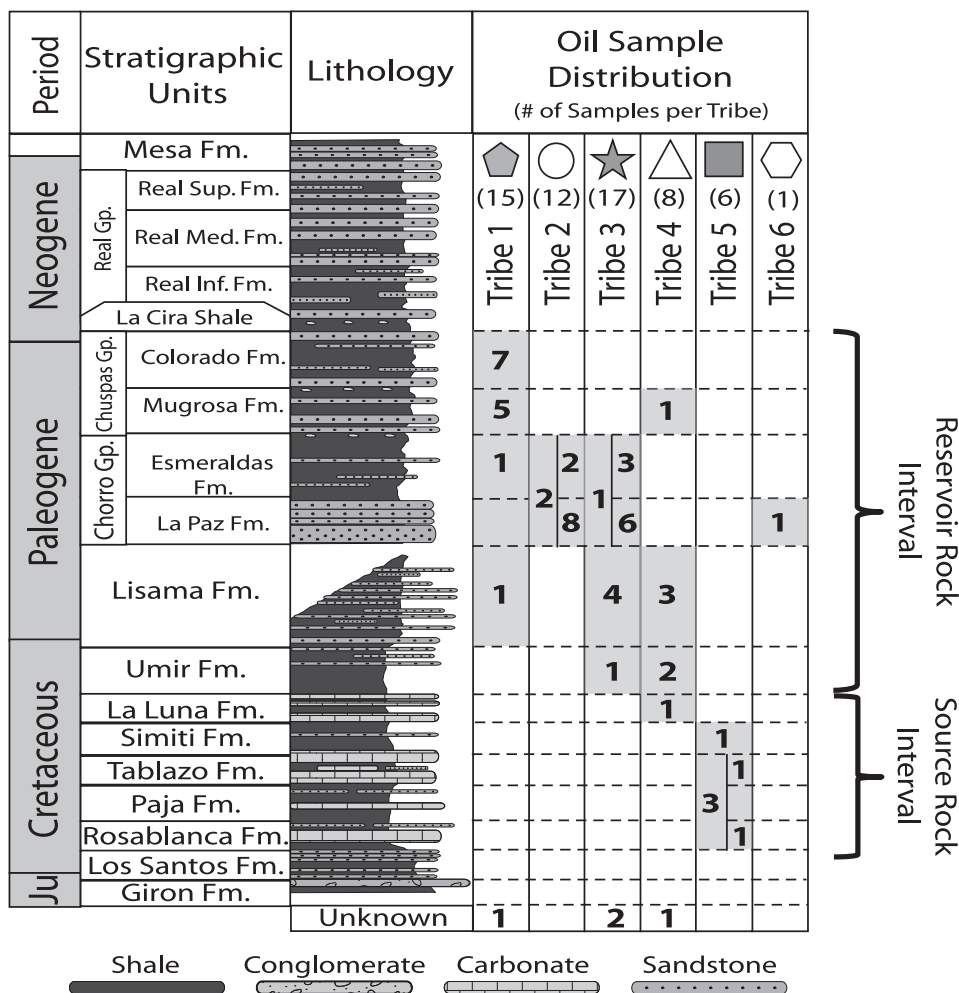
Figure 9. The reservoir rock interval for the samples used in hierarchical cluster analysis 1. Each row is a reservoir rock interval, and the number of samples from that reservoir for each tribe is shown. Samples listed on the boundary between reservoirs indicate that multiple reservoir rocks were recorded for the sample (i.e., "Esmeraldas–La Paz"). Three samples from tribe 5 were listed as "basal Calcareous group," and the Tablazo, Paja, and Rosablanca Formations (Fm.) were not specified. Finally, if a reservoir interval was not specified for a sample, the sample is listed as "unknown." Note that the reservoir distributions vary systemically by tribe. Gp. = Group; Inf. = lower; Ju = Jurassic; Med. = middle; Sup. = upper.

Elevated C_{28}/H and $C_{35}S/C_{34}S$ in tribes 3 and 4 are typical of petroleum generated from source rock deposited under reducing to anoxic conditions (Katz and Elrod, 1983; Peters and Moldowan, 1991), whereas lower values for tribes 1 and 2 indicate relatively more oxicity. Tribes 5 and 6 show high C_{35} hopanes but low C_{28}/H. However, high levels of thermal maturity may impact these parameters. The porphyrin V/Ni ratio is resistant to higher levels of thermal maturity (Lewan, 1984). In a study of oil samples from various Venezuelan and Spanish basins, Galarraga et al. (2008) observed that V/(Ni + V) ratios above 0.75 suggest marine anoxic deposition from a calcareous shale or limestone source rock. The high values of V/(Ni + V)

in tribes 4 and 3 (~0.88 and ~0.81, respectively) are consistent with this interpretation, whereas the lower tribe 5 value (~0.64) indicates source rock deposition under dysoxic–oxic conditions (Galarraga et al., 2008), with tribe 1 and 2 values (~0.74 and ~0.76, respectively) falling in between.

Finally, Zumberge (1987) observed that higher ratios of oleanane to hopane and C_{19} to C_{23} tricyclic terpanes indicate more abundant higher-plant input to a source rock. The sole tribe 6 oil originated from terrigenous organic matter (Figure 12) and shows much higher Ol/H and C_{19}/C_{23} ratios than tribes 1 through 5 (Table 2). Although the marine tribes 1 through 5 all show low higher-plant input, tribe 2 has relatively higher values than tribes 1, 3, 4, and 5, suggesting local
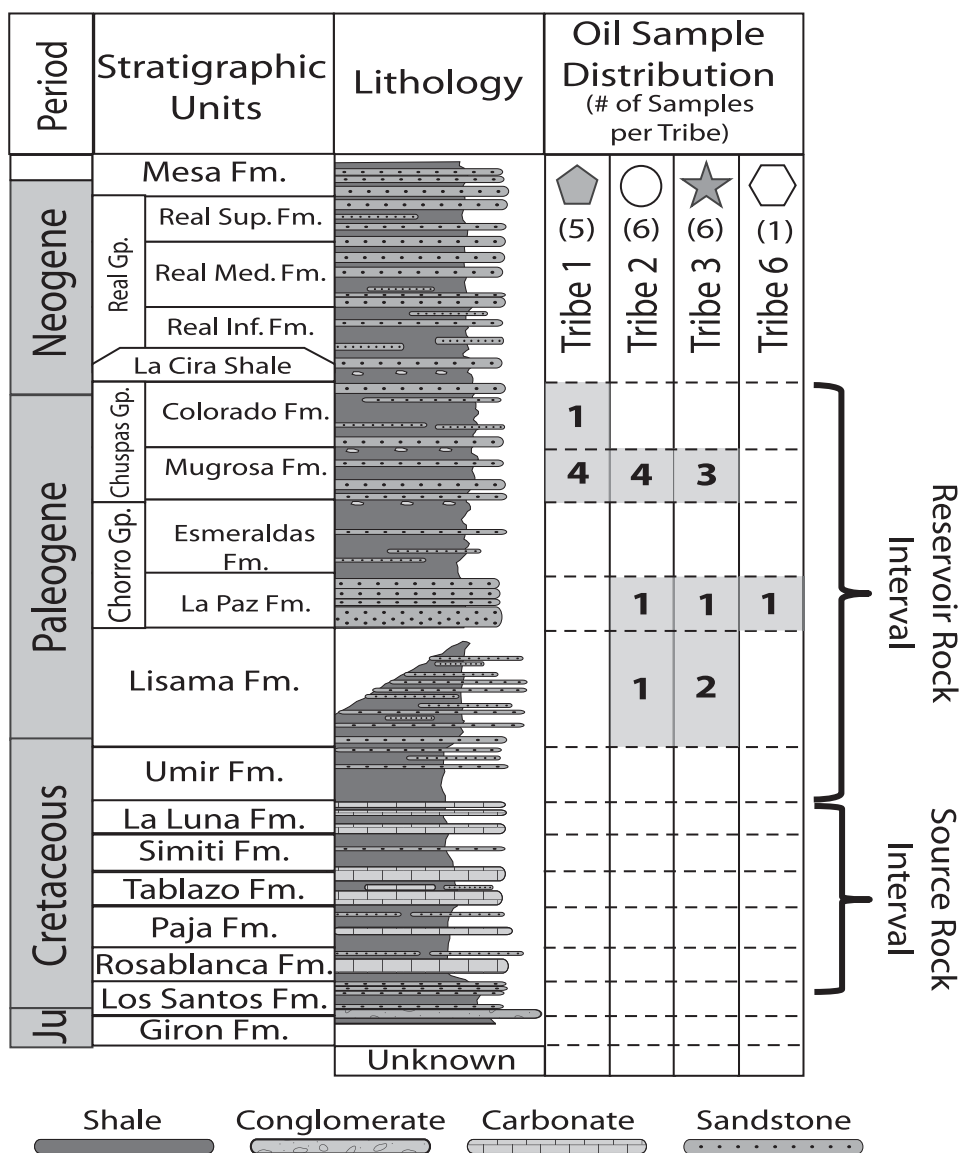
Figure 10. The reservoir rock interval for the samples used in hierarchical cluster analysis (HCA) 2. Each row is a reservoir rock interval, and the number of samples from that reservoir for each tribe is shown. No samples from tribe 4 or tribe 5 were identified in HCA-2. Fm. = Formation; Gp. = Group; Inf. = lower; Ju = Jurassic; Med. = middle; Sup. = upper.

enrichment in terrigenous organic matter. Oleanane and bicadinane index values from the GCMSMS data set (Table 4) support these trends because tribe 6 has significantly higher values than tribes 1 through 3. The Ol/H ratio, oleanane index, and bicadinane index can also be used to describe the source rock age for an oil sample (Moldowan et al., 1994). As a result, although elevated values of these three parameters in tribe 6 would suggest a younger age relative to tribes 1 through 5, it is more likely an indication of a terrigenous source when considering the stable carbon isotope fractions for tribe 6 (Figure 11).

Select biomarker ratios, such as C_{24} nordiacholestane ratio (24-NDR) and C_{25}–highly branched isoprenoids (C_{25}-HBI) (Table 4), are also useful for describing the source rock age of an oil sample. The C_{24} nordiacholestane is used in distinguishing Paleogene from Cretaceous oil because 24-NDR likely originates from diatoms that became increasingly prevalent beginning in the Jurassic (Holba et al., 1998). The Opon oil (tribe 6) has the highest value for 24-NDR, suggesting a younger age than tribes 1, 2, and 3. However, C_{25}-HBI is also reported to increase in concentration during the Cretaceous, but tribe 6 has the lowest
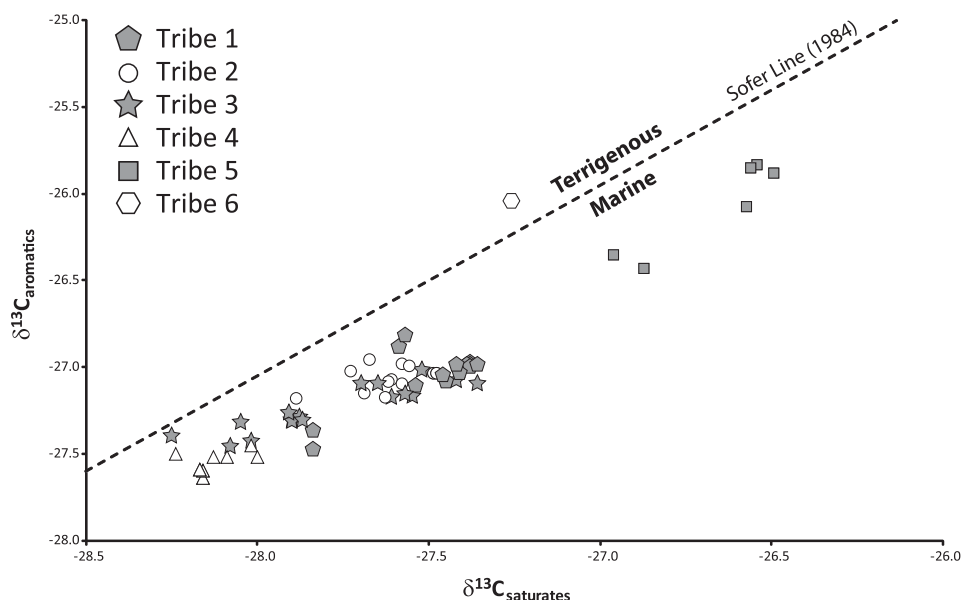
Figure 11. The Sofer Line, established by Sofer (1984), differentiates oils with marine organic matter input and terrigenous organic matter input by cross-plotting the stable carbon isotope fractions for saturates and aromatics. Tribes 5 and 6 are distinct from tribes 1 through 4. Tribe 6 is the only sample with terrigenous organic matter input. Tribes 1 and 2 are similar, whereas tribe 3 is distributed between tribes 2 and 4.

C_{25}-HBI value of any tribe, contrary to the 24-NDR results (Sinninghe Damsté et al., 1989). One possible explanation is that C_{25}-HBI is also impacted by depositional environment. Lower values in this case could be a result of a terrigenous tribe 6 oil versus marine oil in tribes 1 through 4. In summary, the trends for age parameters in this study are unclear either because the impact of age was affected by source-related differences or because the difference in age was not dramatic enough to cause significant differences in parameter values.

Source Rock Maturity from Oil Composition

The extent of source rock thermal maturation differs between the petroleum tribes and is also a key factor in relating the tribes to the basin and petroleum system history of the MMV. As with the biomarker ratios used previously, thermal maturity parameters can be impacted or overprinted by differences in depositional environment, kerogen type, lithology, and other factors. As a result, using bulk parameters and a variety of biomarker indicators for assessing thermal maturity is critical for a thorough interpretation.

Consistent with high thermal maturity, tribe 6 samples show very high saturate/aromatic ratios in addition to lower sulfur content (e.g., Radke, 1988; Orr, 2001). Tribe 4, however, has low saturate/aromatic

ratios and API gravity relative to the other tribes. This may be indicative of early oil window maturity (Peters et al., 2016). However, the probable terrigenous origin of tribe 6 and calcareous marl to carbonate origin for tribe 4 likely impact these values. Supporting these interpretations with other biomarker parameters is necessary for a more conclusive maturity assessment. The C_{27} Ts/Tm and C_{27} diasterane to regular sterane ratios are commonly used to compare thermal maturities of oil samples (Seifert and Moldowan, 1978; Peters et al., 1990). Elevated values for tribes 5 and 6 indicate late oil window maturity, whereas low values for tribes 3 and 4 indicate early oil window maturity (Table 2). In addition, tribe 1 samples from the Colorado and Velasquez fields show locally elevated values. The contrast of the low Ts/Tm values of tribes 3 and 4 with the high Ts/Tm values of tribes 5 and 6 are particularly evident in the *m/z* 191 chromatograms (Figure 7). Although low Ts/Tm and diasterane to regular sterane ratios are consistent with the inferred clay-poor source rock of tribe 4 (Moldowan et al., 1986), the values are also consistent with early oil window maturity. Adding confidence to this assessment, triaromatic steroid ratios, such as the TAS3(CR) and C_{26}/C_{28} STAS ratio, repeat the trends observed for Ts/Tm and the C_{27} diasterane to sterane ratio (Figure 16). The TAS3(CR) and C_{26}/C_{28} STAS ratios are good indicators for thermal maturity at elevated
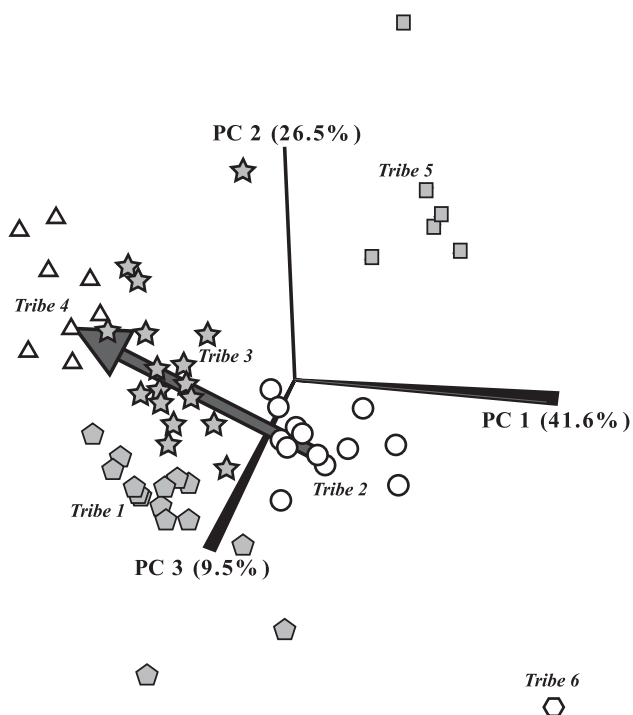
Figure 12. Principal component analysis of tribes 1 through 6 in which multivariable GeoMark data are condensed into three primary axes of variation for visualization in three-dimensional space. The amount of variation encapsulated in each principal component axis is given as a percent in parentheses. The linear trend of samples from tribe 2 to tribe 3 to tribe 4 suggests that mixing may be occurring. PC = principal component axis.

maturities still within the range of the oil window and help show the elevated maturity of tribe 6 compared with tribe 5 (Beach et al., 1989; Peters et al., 2005).

In the central part of the MMV, where many of the largest petroleum fields in the basin occur, bulk property and biomarker analysis consistently indicate early-to-peak oil maturity for both tribes 1 and 2. Samples from tribe 1 in the Colorado and Velasquez oil fields are notable exceptions and show higher maturity values than the other tribe 1 samples. These fields apparently contain oil from source rocks similar to the other tribe 1 samples but were likely buried more deeply because of local faulting in the basin (Caballero et al., 2013). Tribe 4 oil samples, in the northern part of the MMV, are less mature than the tribe 1 and 2 oils, reflecting early oil window maturity. Like the source-related biomarker ratios, the tribe 3 maturity ratios generally plot between those for tribes 2 and 4, supporting mixed accumulations. Despite proximity to many of the tribe 4 oil samples, the tribe 5 oil samples

show late oil window maturity indicative of a deeper source rock within the Cretaceous. Finally, the tribe 6 sample also shows high maturity, possibly indicating migration from more deeply buried organic-rich units in the western flank of the Eastern Cordillera.

Diamondoid Analysis for Cracking from Deep Sources

Multiple late-window oil tribes within the MMV suggest that local contributions from thermally cracked, deep-source intervals may exist. In particular, the organic-rich Aptian–Albian Tablazo Formation (Figure 3) is deeply buried in portions of the central and eastern MMV exhibiting late-window or overmature values based on T_{max} and vitrinite reflectance data (Ramón et al., 1997; Sarmiento, 2011). Many biomarker maturity parameters reach their endpoint prior to the end of the oil window and the onset of secondary cracking and are, therefore, ineffective indicators for highly mature oil (Farrimond et al., 1998). Diamondoids, however, show high thermal stability compared with other compounds. As a result, diamondoids become progressively more concentrated during thermal maturation in a manner that can be used to predict the degree of cracking and identify oil mixtures composed of low- and high-maturity components (Dahl et al., 1999). As a nearby example of this methodology, diamondoid analysis in Colombia's Llanos Basin helped to unravel the complex charge history associated with progressive orogeny of the Eastern Cordillera as well as mixed-maturity oil tribes (Springer et al., 2010; Moldowan et al., 2015).

By comparing concentrations of the biomarker C_{29} stigmastane 20R and the methyldiamantane diamondoid, Dahl et al. (1999) were able to distinguish low-mature, mixed-mature, and postmature or cracked oil. Stigmastane is depleted at approximately the same time that diamondoid concentrations begin to rise, and, as a result, the transition from moderately mature oil to cracked and then heavily cracked oil can be identified. However, plots of stigmastane versus 3- + 4-methyl diamantane from the MMV show few cracked oil samples (Figure 17). Low concentrations of stigmastane and high methyldiamantanes for tribe 6 indicate higher maturity and low to moderate levels of cracking. Tribe 3 shows low-maturity oil with a consistent baseline. Tribes 1 and 2 appear more mature than tribe 3. In addition, whereas tribe 3 plots along a
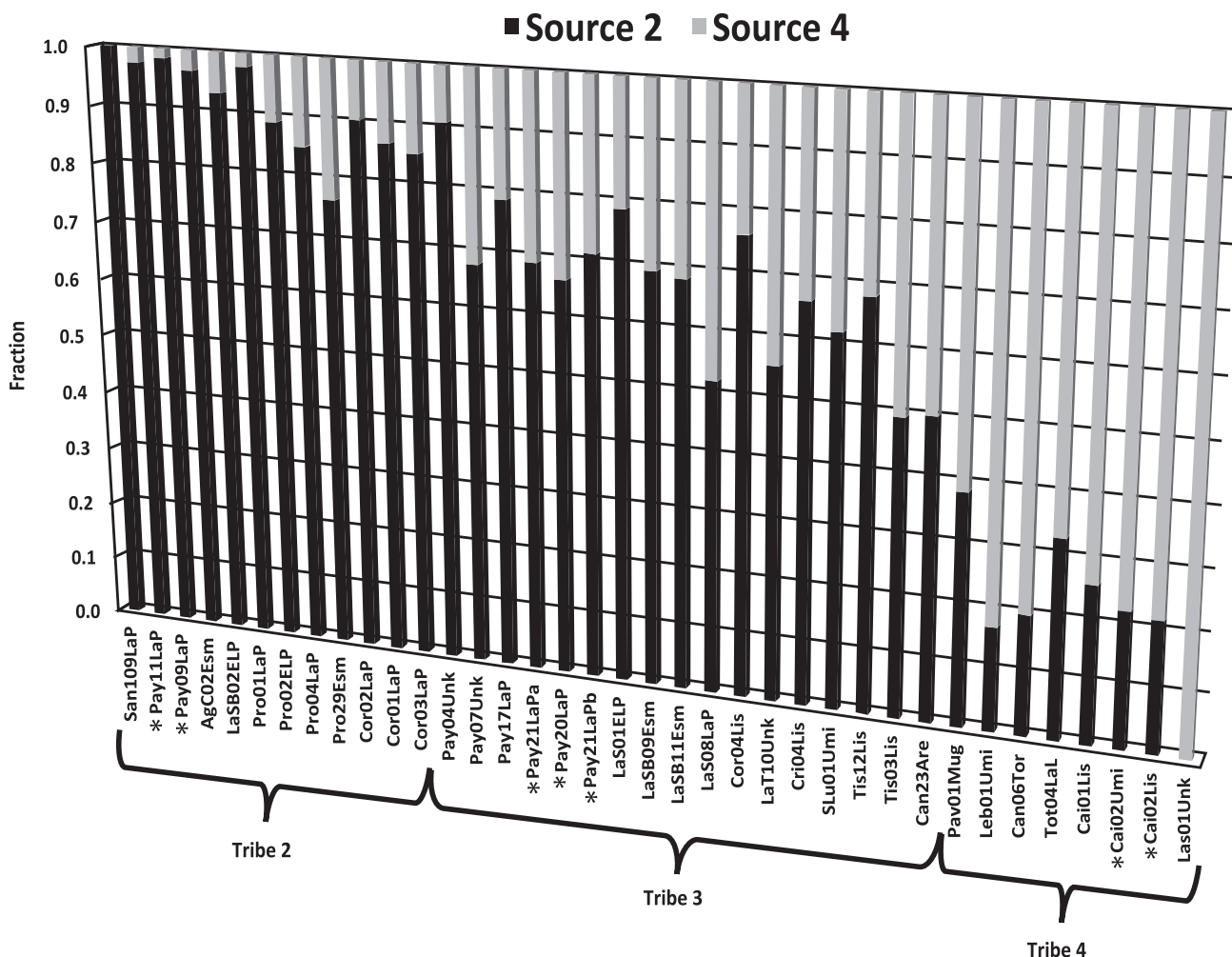
Figure 13. The results of alternating least squares analysis of compound concentrations from 48 biomarker compounds on tribes 2, 3, and 4 from the GeoMark data set. The analysis independently identifies tribes 2 and 4 as end-member tribes with tribe 3 a mixture of the two. Samples with asterisks are taken from the same or proximal wells at similar reservoir intervals and indicate the high level of repeatability in the analysis.

clearly defined diamondoid baseline, no obvious baseline exists for tribes 1 and 2, likely indicating that these two tribes have some contribution from a source rock with low levels of cracking, such as the organic-rich and more deeply buried Tablazo Formation.

Finally, taking into consideration that biodegradation can impact the concentrations of stigmastane, concentrations of the more-resistant methyl adamantanes versus methyl diamantanes can be employed to reduce the impact for samples with low to moderate biodegradation, such as tribe 1. Figure 18 shows elevated thermal maturity and cracking in tribe 6 compared with tribes 1 through 3, with tribe 3 also indicating lower maturity than tribes 1 and 2. These observations are in line with biomarker-based interpretations.

Inferences of Tribe Origin from Chemometrics

Without source rock extract analyses, direct oil–source correlation could not be conducted. However, geochemical characteristics of the oil samples allowed inferences on the source rock and depositional environment for each tribe.

Tribes 1 and 2

Tribes 1 and 2 have many similar geochemical characteristics and account for approximately three-quarters of the oil discovered in the MMV (Table 1). Tribe 1 contains 20 oil samples from Paleogene reservoirs, with 17 of these samples produced from
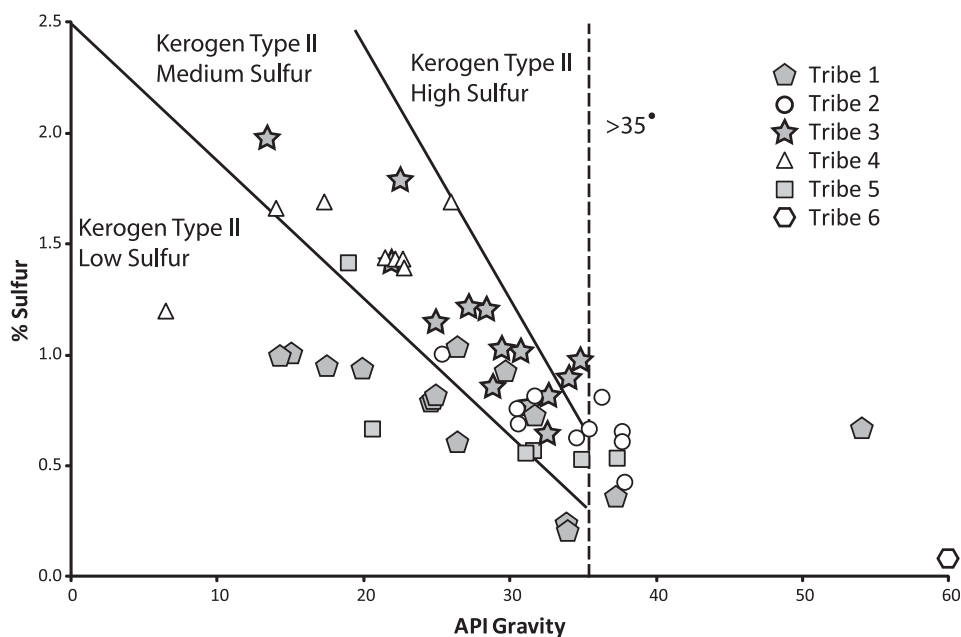
Figure 14. Cross-plot of API gravity versus weight percent of sulfur by tribe in the Middle Magdalena Valley. Low-sulfur, medium-sulfur, and high-sulfur thresholds are based on Orr (2001) for type II kerogen. Tribe 4 tends to show the highest sulfur and lowest API gravity values on average, whereas tribe 2, with the exception of the tribe 6 sample, shows the reverse. Differences in maturity and biodegradation impact the API gravity and weight percent sulfur, contributing to the variation between samples as well as tribes.

the Oligocene Colorado and Mugrosa Formations. In contrast, to the east, tribe 2 consists of 17 samples, 12 of which are from accumulations in the Eocene Esmeraldas and La Paz Formations. Although both tribes originated from clay-rich marine shale, the most notable differences relate to terrigenous organic matter input and depositional conditions. Higher values of Pr/Ph and saturate/aromatics and lower values of DBT/p, sulfur content, and C_{28}/H all indicate that tribe 2 originates from a relatively more clay-rich interval that was deposited under relatively less reducing conditions than that of tribes 1, 3, 4, and 5 (Tables 2, 4). Tribe 2 oil indicates a more terrigenous input than any tribe except tribe 6, as evidenced by elevated Ol/H, oleanane index, and bicadinane index (Tables 2, 4).

Although the source rocks for tribe 2 oil were more proximal to the primary source of terrigenous organic matter to the east, other factors may contribute to this trend. Specifically, the deeper present-day burial of the Cretaceous section in the area of tribe 2 (Figure 19), adjacent to the western flank of the Eastern Cordillera, relative to tribe 1 could impact the stratigraphic intervals that entered the oil generative window. Input from the more transitional marine units, overlying the La Luna Formation, could result in the more clay-rich and oxic biomarker

signatures observed in tribe 2. An oil–source rock extract comparison (Rangel et al., 2002) identified oil generated from the latest Cretaceous Umir Formation along the eastern flank of the MMV near many of the oil samples identified as tribe 2. Contributions from this estuarine–coastal plain unit, in combination with local variations in oxicity in the La Luna Formation at the time of deposition, may help to explain the geochemical character of tribe 2. Finally, enrichment in 3- + 4-methyl diamantine in the samples from both tribe 1 and tribe 2 suggest that the deeply buried Lower Cretaceous Tablazo Formation may be contributing gas or condensate locally (Jimenez et al., 2015) and that multiple charge events may have occurred.

Ramón and Dzou (1999) supported the hypothesis of multiple charge events, observing that oil in the central MMV shows an n-alkane and isoprenoid distribution normally found in moderately degraded oil while also containing a high abundance of gasoline-range hydrocarbons (Figure 20). The presence of these short-chain n-alkanes in a moderately biodegraded oil is rare because these hydrocarbons are some of the most rapidly biodegraded by bacteria. The explanation proposed by Ramón and Dzou (1999) is a sequence of events comprising initial oil migration into the reservoir,
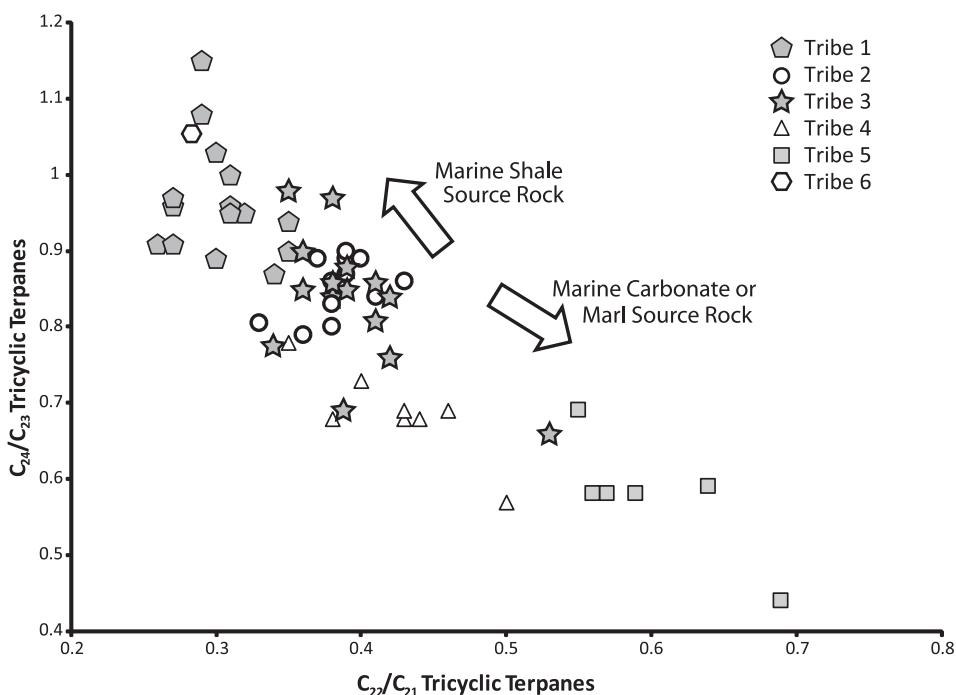
Figure 15. The C_{22}/C_{21} and C_{24}/C_{23} tricyclic terpane ratios are useful tools for differentiating marine carbonate, marl, and marine shale source rock (Aquino Neto et al., 1983; Peters et al., 2005). Tricyclic terpanes from the Middle Magdalena Valley show a marine marl or carbonate source for tribes 4 and 5 and a more siliciclastic marine shale source for tribes 1, 2, 3, and 6.

biodegradation, and finally a secondary pulse of gas or condensate into the reservoir as the section progresses through a later stage of catagenesis (Tissot and Welte, 1978). The diamondoid and biomarker data for tribes 1 and 2 offer further evidence for this secondary pulse proposal because both biomarkers supporting early-to-peak maturity oil (Figure 16) and elevated 3- + 4-methyl diamantanes supporting contributions from secondary cracking (Figure 17) are found in the samples.

Among tribe 1 oil, samples from the Colorado and Velasquez fields are somewhat distinct. The Velasquez oil is from a southern trend of large, often biodegraded fields. The Velasquez field appears to have migrated from a more deeply buried section of the La Luna to the east, explaining the higher maturity compared with most tribe 1 samples. The Colorado field oil also has elevated maturity in addition to high Pr/Ph, C_{24}/C_{23} tricyclic terpanes, and low V/Ni, more characteristic of the depositional environment described for tribe 2. The Colorado field is one of the easternmost tribe 1 fields, with oil likely migrating updip from a more deeply buried section of the La Luna to the east. Proximity to the tribe 2 trend in addition to deeper burial of the Cretaceous section

could help explain why the Colorado field oil differs slightly from the general tribe 1 trends.

Tribes 3, 4, and 5

Tribe 4 contains eight samples, including six samples from Upper Cretaceous and lower Paleogene reservoirs. Tribe 5 contains six samples from Lower Cretaceous reservoirs (Figure 9). Numerous parameters, including saturates/aromatics, sulfur content, DBT/P, Pr/Ph, and $C_{35}S/C_{34}S$ and C_{26}/C_{25} tricyclic terpanes, suggest marine carbonate or marl source rock for tribe 4. Although several of these parameters, specifically Pr/Ph, DBT/P, and tricyclic terpanes, also indicate marine carbonate or marl source rock for tribe 5, many differences exist between tribes 4 and 5. Tribe 4 has the highest V/Ni and lowest $C_{30}X/H$ values (Figure 7) of tribes 1 through 5, indicating the most reducing conditions at the time of deposition. Meanwhile, tribe 5 has characteristics of a depositional environment that is relatively more oxic.

Furthermore, thermal maturity indicators, including Ts/Tm, $C_{29}D/H$, TAS3(CR), and C_{27} diasterane to regular sterane ratio, show late oil window maturity for tribe 5 and early oil window maturity for tribe 4, despite their proximity in the northern part of
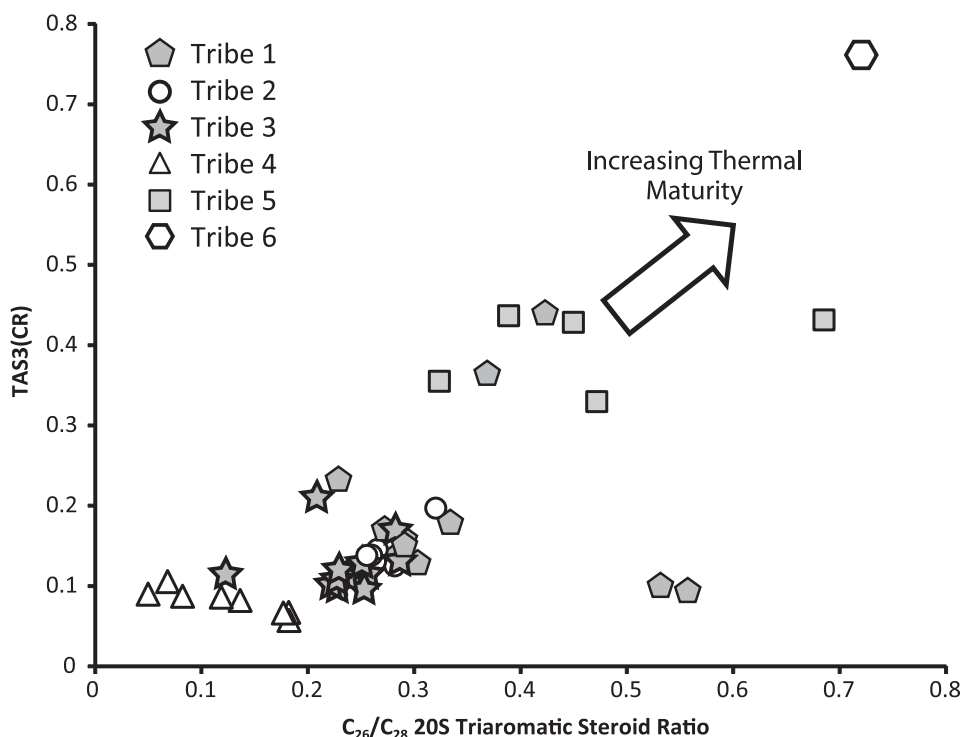
Figure 16. The triaromatic steroid cracking ratio, TAS3(CR), and the C_{26}/C_{28} 20S triaromatic steroid ratio are good indicators for thermal maturity, with higher values indicating greater thermal maturity. These triaromatic steroids are able to resolve higher levels of maturity than the C_{27} 18α-trisnorneohopane/17α-trisnorhopane and C_{27} diasterane to regular sterane ratios and, as a result, are able to resolve that the tribe 6 oil is more thermally mature than the tribe 5 oil. Tribe 4 oil is early oil window maturity, with tribes 1 and 2 showing early to middle oil window maturities.

the MMV. The middle Cretaceous reservoir rock interval, stratigraphically beneath the La Luna Formation, and high maturity of the tribe 5 samples, in addition to favorable *TOC* (>2 wt. %) and *HI* (>300 mg HC/g *TOC*) in the Tablazo Formation (Sarmiento, 2011; Jimenez et al., 2015; US Energy Information Administration, 2015), suggest it may act as the source rock for tribe 5. Despite regional proximity with tribe 5, the stark contrast in thermal maturity makes Tablazo contributions to tribe 4 seem unlikely. Rather, the Upper Cretaceous La Luna Formation is likely the source rock for tribe 4. Previous authors (Ramón et al., 2001) suggested a possible facies change in the La Luna Formation from a marine, carbonate-rich facies in the north to a marine, clastic-rich facies to the southeast. The results of this study support this hypothesis because the northern tribe 4 has biomarker signatures most in favor of a carbonate-rich source rock (Table 2).

The geographic and stratigraphic distribution of tribe 3, as well as the ALS-C analysis of biomarkers, support mixing of tribe 2 and tribe 4 end members.

The tribe 3 samples more closely resemble tribe 4 samples in the north and tribe 2 samples in the southeast, as would be expected in a mixing scenario. Therefore, evidence suggests that in the black oil molecular weight range, tribe 4 likely originates from a locally calcareous facies of the La Luna Formation in the northern MMV, whereas tribe 2 originates largely from a more oxic and clay-rich facies of the La Luna Formation in the central and eastern MMV and may include overlying transitional marine units, such as the Umir Formation. Tribe 3 is the result of local mixing between tribes 2 and 4.

Tribe 6

Tribe 6 consists of a single sample from the Opon field. The Sofer diagram (Figure 11) suggests the tribe 6 oil originated primarily from terrigenous organic matter. Highly elevated values of oleanane and bicadinane indices relative to the other tribes support this observation of a terrigenous source (Table 4). Low $C_{35}S/C_{34}S$, C_{28}/H,
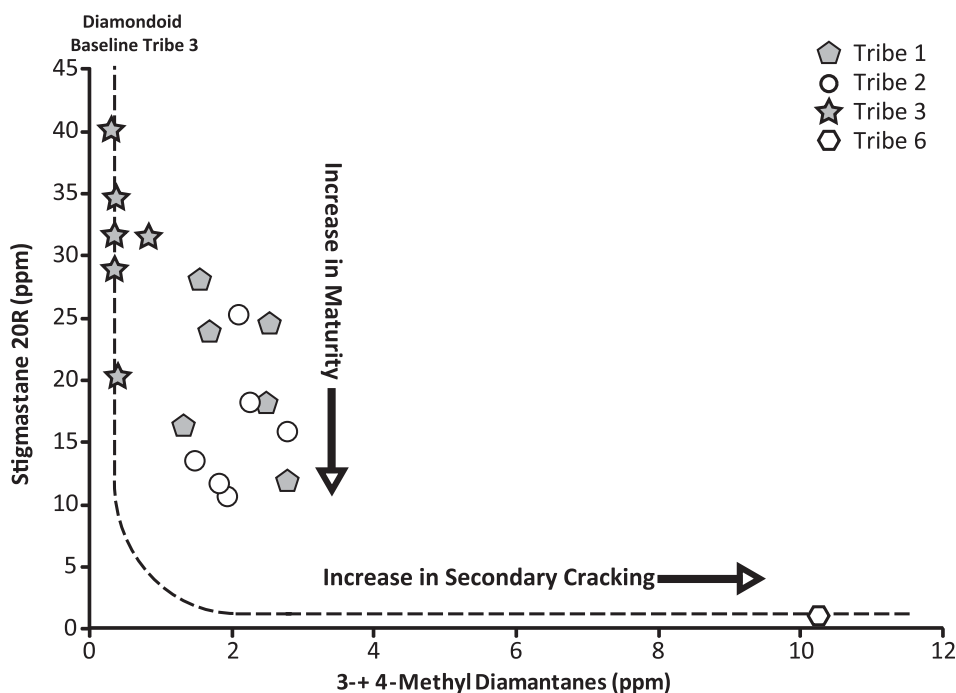
Figure 17. The 3- + 4-methyl diamantane concentration plotted against the biomarker C_{29} stigmastane 20R can help identify low-maturity, high-maturity, or cracked oils in addition to mixed oils with charge from a deeper source with secondary cracking (Dahl et al., 1999). In the Middle Magdalena Valley, the tribe 6 sample from the Opon field shows secondary cracking. Tribe 3 shows the lowest maturity in addition to a distinct diamondoid baseline. Tribes 1 and 2, however, show increased maturity and no clear diaomondoid baseline. The samples from tribes 1 and 2 may show enrichment in 3- + 4-methyl diamantanes, indicating mixing with a deeper source that has experienced low levels of secondary cracking. The more deeply buried middle Cretaceous Tablazo Formation is a possible candidate for the secondary source in these instances.
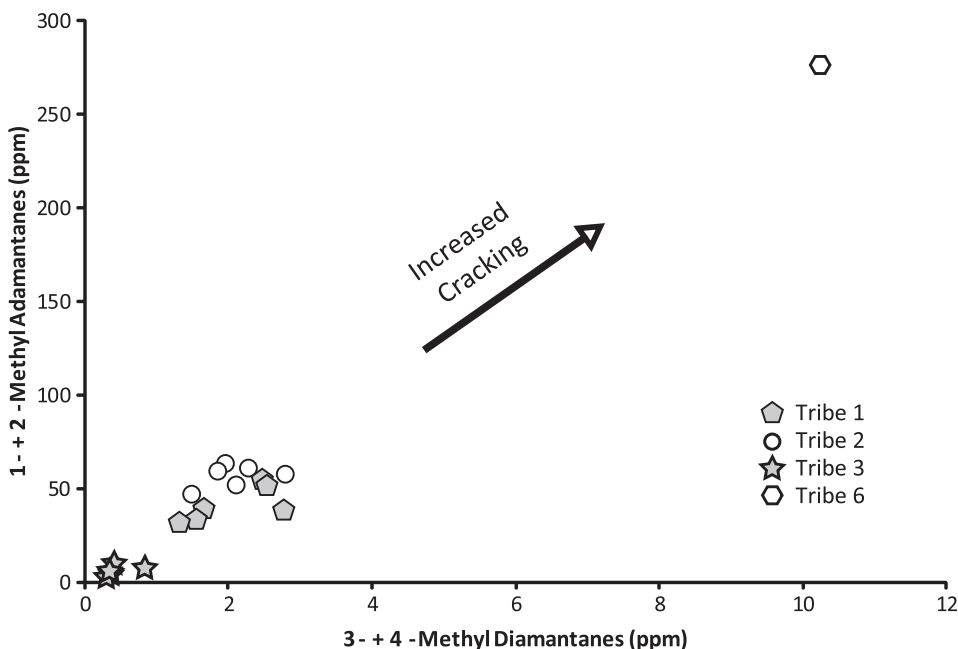
Figure 18. Increased concentrations of 3- + 4-methyl diamantanes and 1- + 2-methyl adamantanes indicate an increase in secondary cracking. Tribe 6 oil shows significantly greater secondary cracking than tribes 1, 2, and 3, corroborating Figure 17. Tribe 3 shows the lowest thermal maturity levels, whereas tribes 1 and 2 are marginally more mature. Again, the higher diamondoid concentrations observed in samples from tribes 1 and 2 may indicate a contribution from a deeper source.

Figure 19. Map of the Middle Magdalena Valley (MMV) contouring the depth to the top of the Cretaceous source rock package, marked by the Eocene unconformity, in total vertical depth subsea (TVDSS) from more than 140 wells. The top of the Cretaceous generally deepens from both north to south and west to east in the MMV. Location of the study area is shown by the dashed box in Figure 1.

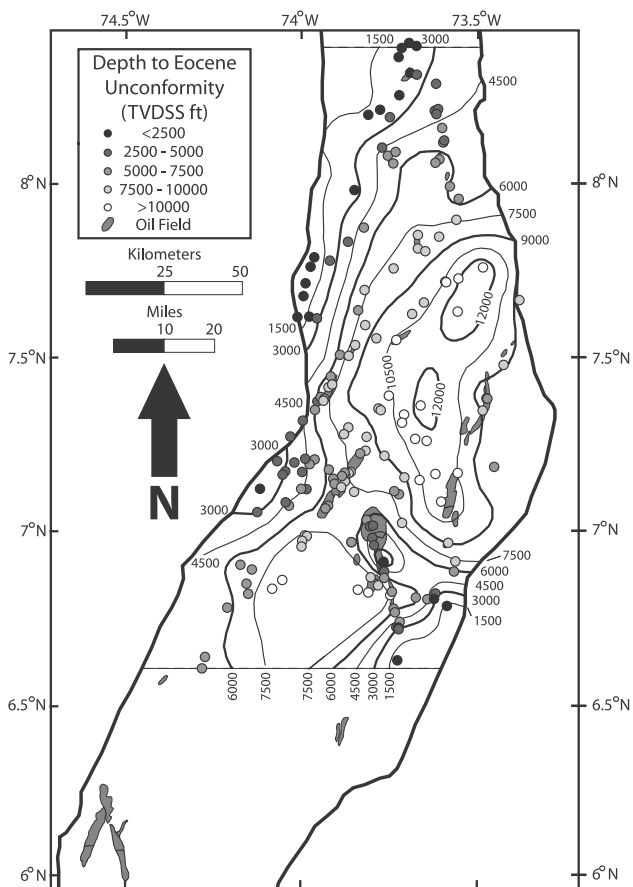and high $C_{30}X/H$ also indicate more oxic deposition (Table 2). Finally, diamondoids and maturity parameters indicate that tribe 6 is highly mature and cracked (Figures 17, 18). The location of the Opon field proximal to the thrust front along the western margin of the Eastern Cordillera suggests that this oil migrated from a thrust-stacked terrigenous source rock, possibly an extension of the estuarine Upper Cretaceous Umir Formation.

CONCLUSIONS

- Chemometric analysis of source-related biomarker and isotopic parameters from 73 oil samples identifies 6 genetically distinct petroleum tribes in the MMV.

- The tribes vary systematically in the MMV both by geographic location and by reservoir rock interval. In addition, multivariate statistical analyses of separate data sets provide the same tribe classification despite being processed by different laboratories and built using different sets of biomarker and isotopic ratios.

- Tribes 4 and 5 are located in the northern MMV and found in Upper Cretaceous and Lower Cretaceous reservoirs, respectively. In contrast, the remainder of the tribes were recovered from overlying Paleogene reservoir rocks (Figure 9). Bulk properties and biomarkers such as DBT/P, Pr/Ph, and tricyclic terpanes indicate a marine carbonate or marl source for both units. The middle Cretaceous reservoir rock interval (stratigraphically beneath the La Luna Formation) and late oil window maturity of the tribe 5 samples suggest a Lower Cretaceous Tablazo Formation source. In contrast, the Upper Cretaceous reservoir interval, early oil window maturity, and indication of a more anoxic depositional environment suggest a La Luna Formation source for tribe 4.

- Bulk properties and biomarker parameters indicate that tribes 1 and 2 derived from the most clay-rich source rocks of the marine-sourced tribes 1 through 5, with tribe 2 being the most clay rich and oxic and the least reducing. In addition, of the marine tribes, tribe 2 has the greatest terrigenous input. This is likely a reflection of increased proximity to the Eastern Cordillera as well as greater burial of the Cretaceous interval, allowing possible contributions from overlying transitional units, such as the Umir Formation.

- Tribe 3 is distributed, both geographically and in terms of reservoir rock interval, between tribes 2 and 4. Using ALS of compound concentrations from 48 biomarkers compounds, tribe 3 was identified as a mixture of tribes 2 and 4 within the molecular weight range of black oil.

- Tribe 6, from the Opon field along the western flank of the Eastern Cordillera, is the sole terrigenous oil identified in this study. In addition, diamondoid analysis of the Opon oil indicates high thermal maturity and moderate cracking. Taken together, these factors suggest updip migration from a more deeply buried terrigenous source in the Eastern Cordillera foothills.

- Biomarker thermal maturity parameters of the tribe 1 through 3 oil samples suggest early-to-peak
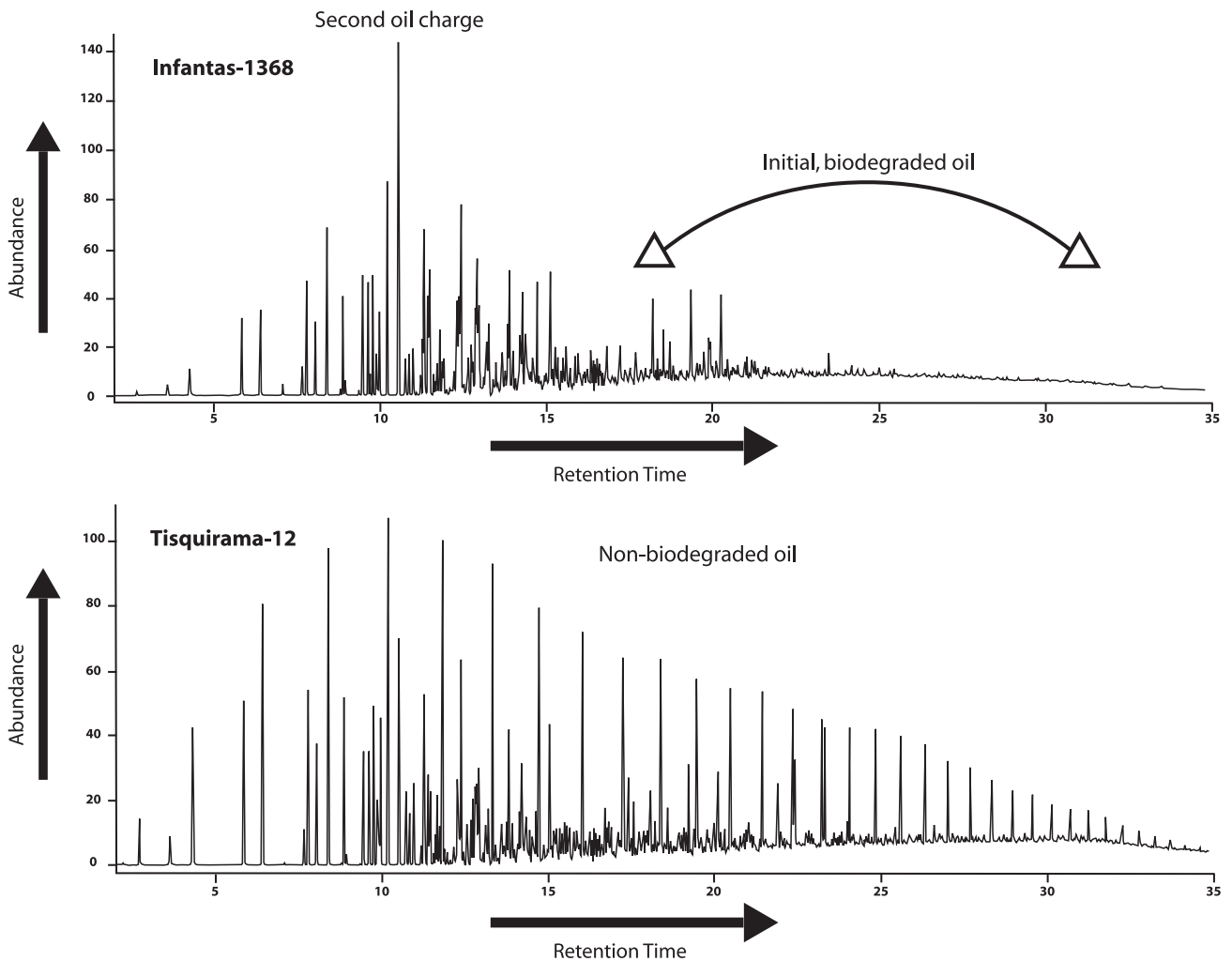
Figure 20. Gas chromatography–flame ionization detector fingerprints for two oil samples, Tisquirama-12 and Infantas-1368, from the Middle Magdalena Valley. Whereas the Tisquirama-12 oil illustrates a typical, nonbiodegraded fingerprint, the Infantas-1368 oil shows an n-alkane and isoprenoid distribution indicative of a moderately biodegraded oil. However, the Infantas-1368 oil also shows a high abundance of gasoline-range hydrocarbons. This combination suggests that an initial oil charge occurred and was biodegraded before a secondary pulse of gas or condensate was emplaced in the reservoir (Ramón and Dzou, 1999).

oil window maturities. However, the tribe 1 and 2 oil samples show evidence of contribution from a more deeply buried source as evidenced by 3- + 4-methyl diamantanes enriched beyond the established diamondoid baseline. The pattern of biodegradation in several of the oils in the MMV further suggest the presence of a secondary gas or condensate pulse (Ramón and Dzou, 1999). The Tablazo Formation is in the condensate-to-gas generative window in much of the central MMV and is the most likely candidate for a more deeply buried, high thermal maturity source.

APPENDIX 1

Table 5. Metadata for Each Oil Sample Used in the Chemometric Study

Tribe	Well Name	Field	Reservoir	Sample Top		Sample Base		Data Source	Latitude	Longitude
				ft	m	ft	m			
1	Lisama-150	Lisama	Mugrosa	7976	2431	8503	2592	GeoMark	7.0400	−73.5589
1	Salina B-08	La Salina	Esmeraldas	3532	1077	3554	1083	GeoMark	7.2830	−73.5127
1	Colorado-19	Colorado	Mugrosa	8600	2621	−	−	GeoMark	6.7673	−73.7892
1	Colorado-10	Colorado	Mugrosa	8600	2621	−	−	GeoMark	6.7733	−73.7253
1	La Cira-750	La Cira	Mugrosa	3295	1004	3662	1116	GeoMark	6.9587	−73.7957
1	Unknown	Baracma	Unknown	−	−	−	−	GeoMark	6.7500	−74.0000
1	Casabe-1	Casabe	Colorado	2802	854	3543	1080	GeoMark	7.0403	−73.9093
1	Infantas-156	Infantas	Mugrosa	2359	719	2656	810	GeoMark	6.9363	−73.7649
1	Galan-9	Galan	Colorado	2650	808	−	−	GeoMark	7.0843	−73.8931
1	La Tigra-5	La Tigra	Lisama	−	−	−	−	GeoMark	7.5043	−73.3767
1	La Cira-1632	La Cira	Mugrosa	2200	671	−	−	GeoMark	6.9860	−73.8010
1	La Cira-1482	La Cira	Mugrosa	1900	579	−	−	GeoMark	6.9471	−73.7700
1	Galan-3	Galan	Colorado	2500	762	−	−	GeoMark	7.0813	−73.8863
1	Velasquez-250	Velasquez	Colorado	5300	1615	6400	1951	GeoMark	6.0558	−74.5013
1	Velasquez-265	Velasquez	Colorado	6156	1876	7150	2179	GeoMark	6.0202	−74.4966
1	Casabe-434	Casabe	Mugrosa	−	−	−	−	PSI	7.0240	−73.9239
1	x	Llanito	Mugrosa	−	−	−	−	Ecopetrol	x	x
1	x	Colorado	Mugrosa	−	−	−	−	Ecopetrol	x	x
1	Infantas-1368	Infantas	Colorado	−	−	−	−	PSI	6.9369	−73.7783
1	AguasBlancas-5	Aguas Blancas	Mugrosa	−	−	−	−	PSI	6.8355	−73.7761
2*	Santos-109	Santos	La Paz	−	−	−	−	PSI	7.3584	−73.4496
2	Suerte-1	Provincia	La Paz	2675	815	−	−	GeoMark	7.3755	−73.4565
2	Santos-4	Provincia	La Paz	8500	2591	−	−	GeoMark	7.3453	−73.4497
2	Santos-29	Provincia	Esmeraldas	−	−	−	−	GeoMark	7.3329	−73.4482
2	Conde-2	Provincia	Esmeraldas–La Paz	−	−	−	−	GeoMark	7.3198	−73.4613
2	Payoa-9	Payoa	La Paz	9308	2837	10,090	3075	GeoMark	7.2788	−73.4767
2	Payoa-11	Payoa	La Paz	9452	2881	9657	2943	GeoMark	7.2730	−73.4844
2	AguasClaras-2	Aguas Claras	Esmeraldas	4232	1290	4875	1486	GeoMark	7.2927	−73.5115
2	Salina B-02	La Salina	Esmeraldas–La Paz	1584	483	2022	616	GeoMark	7.2717	−73.5206
2	Corazon-2	Corazon	La Paz	7816	2382	7846	2391	GeoMark	7.2317	−73.5143
2	Corazon-3	Corazon	La Paz	7564	2306	7640	2329	GeoMark	7.2438	−73.5116
2	Corazon-1	Corazon	La Paz	6500	1981	−	−	GeoMark	7.2376	−73.5124
2	Bonanza-2	Bonanza	Mugrosa	−	−	−	−	PSI	7.4725	−73.4032
2	x	Lisama	Lisama	−	−	−	−	Ecopetrol	x	x
2	Tesoro-33	Tesoro	Mugrosa	−	−	−	−	PSI	7.0373	−73.5721
2	Peroles-2	Peroles	Mugrosa	−	−	−	−	PSI	6.9449	−73.5697
2	Nutria-12	Nutria	Mugrosa	−	−	−	−	PSI	7.0170	−73.5452
3*	x	Tisquirama	Lisama	x	x	−	−	Ecopetrol	x	x
3	Tisquirama-3	Tisquirama	Lisama	8000	2438	−	−	GeoMark	8.0259	−73.5918
3*	Cristalina-4	Cristalina	Lisama	−	−	−	−	PSI	7.5244	−73.6969
3	Santa Lucia-1	Santa Lucia	Umir	9570	2917	9582	2921	GeoMark	7.7734	−73.6537

(*continued*)

Table 5. Continued

Tribe	Well Name	Field	Reservoir	Sample Top ft	Sample Top m	Sample Base ft	Sample Base m	Data Source	Latitude	Longitude
3	Cantagallo-23	Cantagallo	Arenas Cantagallo	6872	2095	7290	2222	GeoMark	7.3639	−73.9182
3	Payoa-20	Payoa	La Paz	10,690	3258	10,847	3306	GeoMark	7.2807	73.4703
3	Payoa-21	Payoa	La Paz	10,504	3202	10,572	3222	GeoMark	7.2915	−73.4776
3	Payoa-7	Payoa	Unknown	10,262	3128	10,342	3152	GeoMark	7.2874	−73.4687
3	Payoa-21	Payoa	La Paz	10,368	3160	10,860	3310	GeoMark	7.2915	−73.4776
3	Salina-08	La Salina	La Paz	3416	1041	3432	1046	GeoMark	7.2684	−73.5177
3	Salina B-11	La Salina	Esmeraldas	2744	836	3287	1002	GeoMark	7.2849	−73.5152
3	La Tigra-10	La Tigra	Unknown	2000	610	−	−	GeoMark	7.5075	−73.3797
3	Salina B-09	La Salina	Esmeraldas	2528	771	2617	798	GeoMark	7.2774	−73.5145
3	Salina-1	La Salina	Esmeraldas–La Paz	1370	418	2230	680	GeoMark	7.2566	−73.5258
3	Payoa-17	Payoa	La Paz	7730	2356	8263	2519	GeoMark	7.2850	−73.4807
3	Payoa-4	Payoa	Unknown	7065	2153	7724	2354	GeoMark	7.2783	−73.4842
3	Corazon-4	Corazon	Lisama	8112	2473	−	−	GeoMark	7.2502	−73.5098
3	Cantagallo-22	Cantagallo	Mugrosa	−	−	−	−	PSI	7.3368	−73.9375
3	x	Yarigui	La Paz	−	−	−	−	Ecopetrol	x	x
3	x	Teca	Mugrosa	−	−	−	−	Ecopetrol	x	x
3	x	San Roque	Mugrosa	x	x	x	x	Ecopetrol	x	x
4	Pavas-1	Pavas	Mugrosa	11,725	3574	−	−	GeoMark	7.6928	−73.5430
4	Cantagallo-6	Cantagallo	Toro Shale	5500	1676	6035	1839	GeoMark	7.3561	−73.9144
4	Las Lajas-1	Las Lajas	Unknown	9920	3024	10032	3058	GeoMark	7.6247	−73.7702
4	Caiman-1	Caiman	Lisama	5300	1615	−	−	GeoMark	8.0175	−73.6035
4	Lebrija-1	Lebrija	Umir	9731	2966	−	−	GeoMark	8.0379	−73.7389
4	Totumal-4	Totumal	La Luna	7305	2227	8649	2636	GeoMark	8.1703	−73.6009
4	Caiman-2	Caiman	Lisama	7614	2321	9017	2748	GeoMark	8.0336	−73.5917
4	Caiman-2	Caiman	Umir	11,061	3371	12,004	3659	GeoMark	8.0336	−73.5917
5	Buturama-6	Buturama-Olivo	Rosa Blanca	5890	1795	−	−	GeoMark	8.2597	−73.7003
5	Buturama-3	Buturama-Olivo	Calcareous Basal Gr.	6170	1881	6357	1938	GeoMark	8.2728	−73.6751
5	Buturama-3	Buturama-Olivo	Calcareous Basal Gr.	6170	1881	6357	1938	GeoMark	8.2728	−73.6751
5	Buturama-3	Buturama-Olivo	Calcareous Basal Gr.	6170	1881	6357	1938	GeoMark	8.2728	−73.6751
5	Puntapiedra-1	Punta Piedra	Simiti	11,546	3519	12,172	3710	GeoMark	7.6772	−73.5724
5	Norean-1	Norean	Tablazo	5668	1728	5710	1740	GeoMark	8.3584	−73.6784
6*	x	Opon	La Paz	−	−	−	−	Ecopetrol	x	x

Tribe, well name, field, reservoir, sample interval top and base, data source, latitude, and longitude for each sample used in the chemometric analysis. For certain samples with restricted access, the well name, sample reservoir interval, latitude, and longitude are omitted. Asterisks next to tribe number indicate samples processed both at Biomarker Technologies, Inc. and GeoMark Research. Sample depth information was not available for many of the samples.

Abbreviations: − = data unavailable; x = data access restricted; Gr. = Group; PSI = Petroleum Systems International.

APPENDIX 2

By using both the GCMS and GCMSMS at BTI, 11 additional biomarker ratios and indices were incorporated into HCA-2. Compound-based definitions for several of these indices and ratios are included here: 3β-methylhopane index = 3β-methylhopane (αβ)/(3β-methylhopane [αβ] + C30 hopane); HHI = (C35 αβ-hopane [22S] + C35 αβ-hopane [22R])/(C35 αβ-hopane [22S] + C35 αβ-hopane [22R] + C34 αβ-hopane [22S] + C34 αβ-hopane [22R] + C33 αβ-hopane [22S] + C33 αβ-hopane [22R] + C32 αβ-hopane [22S] + C32 αβ-hopane [22R] + C31 αβ-hopane [22S] + C31 αβ-hopane [22R]); 2/3-methylhopane index = (C31 2β-methylhopane [αβ] + C31 2α-methylhopane [αβ])/(C31 2β-methylhopane [αβ] + C31 2α-methylhopane + 3β-methylhopane [αβ]); 21-NCR = C26 21-nor ααα + αββ/(C26 21-nor ααα + αββ + C26 ααα-27-nor-20S + C26 αββ-27-nor-20R + C26 αββ-27-nor-20S + C26 ααα-27-nor-20R); 24-NDR = (C26 13β 17α-24-nordia-20S + C26 13β 17α-24-nordia-20R)/ (C26 13β 17α-24-nordia-20S + C26 13β 17α-24-nordia-20R

+ C26 13β 17α-27-nordia-20S + C26 13β 17α-27-nordia-20R); bicadinane index = bicadinane-T/(bicadinane-T + C30 hopane); and diahopane index = oleanane-α/(oleanane-α + oleanane-β).

REFERENCES CITED

Aquino Neto, F. R., J. M. Trendel, A. Restle, J. Connan, and P. A. Albrecht, 1983, Occurrence and formation of tricyclic and tetracyclic terpanes in sediments and petroleums, in M. Bjoroy, C. Albrecht, C. Cornford, K. De Groot, G. Eglinton, E. Galimov, D. Leythaeuser, R. Pelet, J. Rullkotter, and G. Speers, eds., Advances in organic geochemistry: New York, John Wiley & Sons, p. 659–676.

Beach, F., T. M. Peakman, G. D. Abbott, R. Sleeman, and J. R. Maxwell, 1989, Laboratory thermal alteration of triaromatic steroid hydrocarbons: Organic Geochemistry, v. 14, no. 1, p. 109–111, doi:10.1016/0146-6380(89)90024-7.

Caballero, V., M. Parra, A. Mora, C. Lopez, L. E. Rojas, and I. Quintero, 2013, Factors controlling selective abandonment and reactivation in thick-skin orogens: A case study in the Magdalena Valley, Colombia, in M. Nemčok, A. Mora, and J. W. Cosgrove, eds., Thick-skin-dominated orogens: From initial inversion to full accretion: Geological Society, London, Special Publications 2013, v. 377, p. 343–367, doi:10.1144/SP377.4.

Christie, O. H. J., K. Esbensen, T. Meyer, and S. Wold, 1984, Aspects of pattern recognition in organic geochemistry: Organic Geochemistry, v. 6, p. 885–891, doi:10.1016/0146-6380(84)90111-6.

Cooper, M. A., F. T. Addison, R. Alvarez, M. Coral, R. H. Graham, A. B. Hayward, S. Howe, et al., 1995, Basin development and tectonic history of the Llanos Basin, Eastern Cordillera and Middle Magdalena Valley, Colombia: AAPG Bulletin, v. 79, no. 10, p. 1421–1443.

Dahl, J. E., J. M. Moldowan, K. E. Peters, G. E. Claypool, M. A. Rooney, G. E. Michael, M. R. Mello, and M. L. Kohnen, 1999, Diamondoid hydrocarbons as indicators of natural oil cracking: Nature, v. 399, no. 6731, p. 54–57, doi:10.1038/19953.

Dickey, P., 1992, La Cira-Infantas field, Colombia, Middle Magdalena Basin, in E. A. Beaumont and N. H. Foster, eds., Structural traps VII: AAPG Treatise of Petroleum Geology—Atlas of Oil and Gas Fields, p. 323–347.

Farrimond, P., A. Taylor, and N. Telnaes, 1998, Biomarker maturity parameters: The role of generation and thermal degradation: Organic Geochemistry, v. 29, no. 5–7, p. 1181–1197, doi:10.1016/S0146-6380(98)00079-5.

Galarraga, F., K. Reategui, A. Martinez, M. Martinez, J. F. Llamas, and G. Marquez, 2008, V/Ni ratio as a parameter in palaeoenvironmental characterization of nonmature medium-crude oils from several Latin American basins: Journal of Petroleum Science Engineering, v. 61, no. 1, p. 9–14, doi:10.1016/j.petrol.2007.10.001.

He, M., J. M. Moldowan, A. Nemchenko-Rovenskaya, and K. E. Peters, 2012, Oil families and their inferred source rocks in the Barents Sea and northern Timan-Pechora Basin, Russia: AAPG Bulletin, v. 96, no. 6, p. 1121–1146, doi:10.1306/10181111043.

Holba, A. G., L. I. P. Dzou, W. D. Masterson, W. B. Hughes, B. J. Huizinga, M. S. Singletary, J. M. Moldowan, M. R. Mello, and E. Tegelaar, 1998, Application of 24-norcholestanes for constraining source age of petroleum: Organic Geochemistry, v. 29, no. 5–7, p. 1269–1283, doi:10.1016/S0146-6380(98)00184-3.

Hughes, W. B., A. Holba, and L. Dzou, 1995, The ratios of dibenzothiophene to phenanthrene and pristane to phytane as indicators of depositional environment and lithology of petroleum source rocks: Geochimica et Cosmochimica Acta, v. 59, no. 17, p. 3581–3598, doi:10.1016/0016-7037(95)00225-O.

IHS Markit, 2010, Oil and gas, accessed December 1, 2015, https://www.ihs.com/industry/oil-gas.html.

Jimenez, M., C. Rojas, and Y. Valderrama, 2015, Initial analysis of hydrocarbon potential in the Tablazo Formation; Center and northern areas of Middle Magdalena Valley, MMV, Colombia: Unconventional Resources Technology Conference, San Antonio, Texas, July 20–22, 2015, 6 p., doi:10.15530/urtec-2015-2153136.

Katz, B. J., and L. W. Elrod, 1983, Organic geochemistry of DSDP Site 467, offshore California, Middle Miocene to Lower Pliocene strata: Geochimica et Cosmochimica Acta, v. 47, no. 3, p. 389–396, doi:10.1016/0016-7037(83)90261-2.

Lewan, M. D., 1984, Factors controlling the proportionality of vanadium to nickel in crude oils: Geochimica et Cosmochimica Acta, v. 48, no. 11, p. 2231–2238, doi:10.1016/0016-7037(84)90219-9.

Mackenzie, A. S., C. F. Hoffmann, and J. R. Maxwell, 1981, Molecular parameters of maturation in the Toarcian shales, Paris Basin, France—III. Changes in aromatic steroid hydrocarbons: Geochimica et Cosmochimica Acta, v. 45, no. 8, p. 1345–1355, doi:10.1016/0016-7037(81)90227-1.

Moldowan, J. M., J. Dahl, B. J. Huizinga, F. J. Fago, L. J. Hickey, T. M. Peakman, and D. W. Taylor, 1994, The molecular fossil record of oleanane and its relation to angiosperms: Science, v. 265, no. 5173, p. 768–771, doi:10.1126/science.265.5173.768.

Moldowan, J. M., S. Moldowan, V. Blanco-Velandia, Y. Blanco-Velandia, C. Orejuela-Parra, G. Bott, and J. Dahl, 2015, Llanos Basin: Unraveling its complex petroleum systems with advanced geochemical technologies: AAPG Search and Discovery article 10776, accessed February 15, 2016, http://www.searchanddiscovery.com/pdfz/documents/2015/10776moldowan/ndx_moldowan.pdf.html.

Moldowan, J. M., P. Sundararaman, and M. Schoell, 1986, Sensitivity of biomarker properties to depositional environment and/or source input in the Lower Toarcian of SW-Germany: Organic Geochemistry, v. 10, no. 4–6, p. 915–926, doi:10.1016/S0146-6380(86)80029-8.

Morales, L. G., and Colombian Petroleum Industry, 1958, General geology and oil occurrences of Middle Magdalena Valley, Colombia, in L. G. Weeks, ed., AAPG Habitat of Oil, p. 641–695.

Orr, W. L., 2001, Evaluating kerogen sulfur content from crude oil properties—Cooperative Monterey organic geochemistry study, *in* C. M. Isaacs and J. Rullkotter, eds., The Monterey Formation—From rocks to molecules: New York, Columbia University Press, p. 348–367.

Parra, M., A. Mora, E. R. Sobel, M. R. Strecker, and R. Gonzalez, 2009, Episodic orogenic front migration in the northern Andes: Constraints from low-temperature thermochronology in the Eastern Cordillera, Colombia: Tectonics, v. 28, no. 4, TC4004, 27 p., doi:10.1029/2008TC002423.

Peters, K. E., F. D. Hostettler, T. D. Lorenson, and R. J. Rosenbauer, 2008a, Families of Miocene Monterey crude oil, seep, and tarball samples, coastal California: AAPG Bulletin, v. 92, no. 9, p. 1131–1152, doi:10.1306/04180807113.

Peters, K. E., and J. M. Moldowan, 1991, Effects of source, thermal maturity, and biodegradation on the distribution and isomerization of homohopanes in petroleum: Organic Geochemistry, v. 17, no. 1, p. 47–61, doi:10.1016/0146-6380(91)90039-M.

Peters, K. E., and J. M. Moldowan, 1993, The biomarker guide—Interpreting molecular fossils in petroleum and ancient sediments: New Jersey, Prentice Hall, 363 p.

Peters, K. E., J. M. Moldowan, M. Schoell, and W. B. Hempkins, 1986, Petroleum isotopic and biomarker composition related to source rock organic matter and depositional environment: Organic Geochemistry, v. 10, no. 1–3, p. 17–27, doi:10.1016/0146-6380(86)90006-9.

Peters, K. E., J. M. Moldowan, and P. Sundararaman, 1990, Effects of hydrous pyrolysis on biomarker thermal maturity parameters, Monterey phosphatic and siliceous members: Organic Geochemistry, v. 15, no. 3, p. 249–265, doi:10.1016/0146-6380(90)90003-I.

Peters, K. E., L. S. Ramos, J. E. Zumberge, Z. C. Valin, and K. E. Bird, 2008b, De-convoluting mixed crude oil in Prudhoe Bay field, North Slope, Alaska: Organic Geochemistry, v. 39, no. 6, p. 623–645, doi:10.1016/j.orggeochem.2008.03.001.

Peters, K. E., L. S. Ramos, J. E. Zumberge, Z. C. Valin, C. R. Scotese, and D. L. Gautier, 2007, Circum-Arctic petroleum systems identified using decision-tree chemometrics: AAPG Bulletin, v. 91, no. 6, p. 877–913, doi:10.1306/12290606097.

Peters, K. E., C. C. Walters, and J. M. Moldowan, 2005, The biomarker guide: Cambridge, United Kingdom, Cambridge University Press, 1155 p.

Peters, K. E., T. L. Wright, L. S. Ramos, J. E. Zumberge, and L. B. Magoon, 2016, Chemometric recognition of genetically distinct oil families in the Los Angeles basin, California: AAPG Bulletin, v. 100, no. 1, p. 115–135, doi:10.1306/08031515068.

Prince, M., R. Acevedo, C. Sanchez, and D. Rojas, 2011, Upper Cretaceous in the Middle Magdalena Valley, Colombia, South America: A new exploratory target in an old mature basin: AAPG Search and Discovery article 10296, accessed November 3, 2015, http://www.search-anddiscovery.com/documents/2011/10296prince/.

Radke, M., 1988, Application of aromatic compounds as maturity indicators in source rocks and crude oils: Marine and Petroleum Geology, v. 5, no. 3, p. 224–236, doi:10.1016/0264-8172(88)90003-7.

Ramón, J. C., and L. I. Dzou, 1999, Petroleum geochemistry of Middle Magdalena Valley, Colombia: Organic Geochemistry, v. 30, no. 4, p. 249–266, doi:10.1016/S0146-6380(99)00005-4.

Ramón, J. C., L. I. Dzou, and B. Giraldo, 1997, Geochemical evaluation of the Middle Magdalena Basin, Colombia: Ciencia, Tecnologia y Futuro, v. 1, no. 3, p. 47–66.

Ramón, J. C., L. I. Dzou, W. B. Hughes, and A. G. Holba, 2001, Evolution of the Cretaceous organic facies in Colombia: Implications for oil composition: Journal of South American Earth Sciences, v. 14, no. 1, p. 31–50, doi:10.1016/S0895-9811(01)00010-4.

Rangel, A., J. M. Moldowan, C. Nino, P. Parra, and B. N. Giraldo, 2002, Umir Formation: Organic geochemical and stratigraphic assessment as cosource for Middle Magdalena Basin oil, Colombia: AAPG Bulletin, v. 86, no. 12, p. 2069–2087.

Rangel, A., J. F. Osorno, J. C. Ramirez, J. De Bedout, J. L. Gonzalez, and J. M. Pabon, 2017, Geochemical assessment of the Colombian oils based on bulk petroleum properties and biomarker parameters: Marine and Petroleum Geology, v. 86, p. 1291–1309, doi:10.1016/j.marpetgeo.2017.07.010.

Rangel, A., P. Parra, and C. Nino, 2000, The La Luna formation: Chemostratigraphy and organic facies in the Middle Magdalena Basin: Organic Geochemistry, v. 31, no. 12, p. 1267–1284, doi:10.1016/S0146-6380(00)00127-3.

Restrepo-Pace, P. A., F. Colmenares, C. Higuera, and M. Mayorga, 2004, A fold-and-thrust belt along the western flank of the Eastern Cordillera of Colombia – Style, kinematics, and timing constraints derived from seismic data and detailed surface mapping, *in* K. R. McClay, ed., Thrust tectonics and hydrocarbon systems: AAPG Memoir 82, p. 598–613.

Reyes-Harker, A., C. F. Ruiz-Valdivieso, A. Mora, J. C. Ramirez-Arias, G. Rodriguez, F. de la Parra, V. Caballero, et al., 2015, Cenozoic paleogeography of the Andean foreland and retroarc hinterland of Colombia: AAPG Bulletin, v. 99, no. 8, p. 1407–1453, doi:10.1306/06181411110.

Sarmiento, L. F., 2011, Petroleum geology of Colombia, Middle Magdalena basin: Medellin, Colombia, Agencia Nacional de Hidrocarburos, 193 p.

Seifert, W. K., and J. M. Moldowan, 1978, Applications of steranes, terpanes, and monoaromatics to the maturation, migration and source of crude oils: Geochimica et Cosmochimica Acta, v. 42, no. 1, p. 77–95, doi:10.1016/0016-7037(78)90219-3.

Sinninghe Damsté, J. S., E. R. Koert, A. C. Kock-van Dalen, J. W. de Leeuw, and P. A. Schenck, 1989, Characterization of highly branched isoprenoid thiophenes occurring in sediments and immature crude oils: Organic Geochemistry, v. 14, no. 5, p. 555–567, doi:10.1016/0146-6380(89)90035-1.

Sofer, Z., 1984, Stable carbon isotope compositions of crude oils: Application to source depositional environments and petroleum alteration: AAPG Bulletin, v. 68, no. 1, p. 31–49.

Springer, M. V., D. F. Garcia, F. T. T. Goncalves, L. Landau, and D. A. Azevedo, 2010, Diamondoid and biomarker characterization of oils from the Llanos Orientales Basin, Colombia: Organic Geochemistry, v. 41, no. 9, p. 1013–1018, doi:10.1016/j.orggeochem.2010.03.002.

Tissot, B. P., and D. H. Welte, 1978, Diagenesis, catagenesis and metagenesis of organic matter, in B. P. Tissot and D. H. Welte, eds., Petroleum formation and occurrence: Berlin, Springer, p. 69–73, doi:10.1007/978-3-642-96446-6_6

Tyson, R. V., and T. H. Pearson, 1991, Modern and ancient continental shelf anoxia: An overview, in R. V. Tyson and T. H. Pearson, eds., Modern and ancient continental shelf anoxia: Geological Society, London, Special Publications 1991, v. 58, p. 1–24, doi:10.1144/GSL.SP.1991.058.01.01

US Energy Information Administration, 2015, Technically recoverable shale oil and shale gas resources: Northern South America: Washington, DC, US Department of Energy, 30 p.

Villamil, T., 1998, Chronology, relative sea-level history and a new sequence stratigraphic model for basinal cretaceous facies of Colombia: Tulsa, Oklahoma, SEPM Special Publication 58, p. 161–216, doi:10.2110/pec.98.58.0161

Villamil, T., 1999, Campanian-Miocene tectonostratigraphy, depocenter evolution and basin development of Colombia and western Venezuela: Palaeogeography, Palaeoclimatology, Palaeoecology, v. 153, no. 1–4, p. 239–275, doi:10.1016/S0031-0182(99)00075-9.

Waples, D. W., and J. A. Curiale, 1999, Oil-oil and oil-source rock correlations, in E. A. Beaumont and N. H. Foster, eds., Treatise of Petroleum Geology, Handbook of petroleum geology: Exploring for oil and gas traps: AAPG Treatise Handbook 3, p. 1–71.

Zumberge, J. E. 1980, Oil-oil and oil-source rock correlations of bacterially degraded oils and Cretaceous outcrops from Colombia, South America (abs.): International Geological Congress 26, Paris, July 7–17, 1980, p. 806.

Zumberge, J. E., 1984, Source rocks of the La Luna Formation (Upper Cretaceous) in the Middle Magdalena Valley, Colombia, in J. C. Palacas, ed., Petroleum geochemistry and source rock potential of carbonate rocks: AAPG Studies in Geology 18, p. 127–133.

Zumberge, J. E., 1987, Prediction of source rock characteristics based on terpane biomarkers in crude oils: A multivariate statistical approach: Geochimica et Cosmochimica Acta, v. 51, no. 6, p. 1625–1637, doi:10.1016/0016-7037(87)90343-7.