

PS Using Hierarchical Cluster Analysis to Improve Facies Definitions in Permian Mudrocks (Wolfcamp and Lower Leonard), Midland Basin, Texas*

Robert W. Baumgardner, Jr.¹ and Harry D. Rowe¹

Search and Discovery Article #41986 (2017)**

Posted February 6, 2017

*Adapted from poster presentation given at AAPG 2016 Annual Convention and Exhibition, Calgary, Alberta, Canada, June 19-22, 2016

**Datapages © 2017 Serial rights given by author. For all other rights contact author directly.

¹Bureau of Economic Geology, University of Texas at Austin, Austin, Texas, United States (robert.baumgardner@beg.utexas.edu)

Abstract

Mudrocks are notoriously difficult to describe. XRF analysis with calibrated hand-held instruments gives quantitative elemental results that provide insight into mineralogical composition of these fine-grained, dark rocks, if XRF results are supplemented by mineralogical analysis (XRD). When properly interpreted, XRF data can improve facies definitions, but understanding the abundance of elemental data (20+ major and trace elements for every data point) is problematic. The object of this study was an 852-ft continuous core through a sequence of interbedded basinal hemipelagic and sediment gravity flow deposits. Facies delineation based on core description was refined using iterative hierarchical cluster analysis, a technique that treats the rock as a whole, rather than analyzing individual elements or element ratios (e.g. Ca, Al, Si/Al, Si/Ti). Elements were interpreted as proxies for productivity (Ni, Zn, V), reducing conditions (Mo, U), detrital deposition (Si, Al, Ti, Zr, Rb), carbonate deposition (Ca, Mg, Sr), phosphate enrichment (P, Y), and sulfur enrichment (S). The most significant cluster-defining elements were determined by applying analysis of variance and a partitioning index to elements in each cluster. This approach produced chemofacies (e.g. high-detrital siliceous mudrock) that cannot be ascertained as rapidly or as quantitatively by other methods, delineated a previously-unrecognized geochemical boundary between the Wolfcamp and lower Leonard (sulfur-enriched mudrocks below vs. high-redox/high productivity mudrocks above), and revealed sub-meter-scale cyclicity of chemofacies that is not otherwise apparent. Calibrated XRF data subjected to cluster analysis provide finely detailed, core-based, geochemical ‘ground truth’ that is not available by any other means. This technique is a valuable supplement to traditional description of lithofacies based on depositional features seen in core.

References Cited

Algeo, T.J., R. Hannigan, H. Rowe, M. Brookfield, A. Baud, L. Krystyn, and B.B. Ellwood, 2007, Sequencing events across the Permian-Triassic boundary, Guryul Ravine (Kashmir, India): *Palaeogeography, Palaeoclimatology, Palaeoecology*, v. 252, p. 328–346.

Calvert, S.E., and T.F. Pedersen, T.F., 2007, Elemental proxies for palaeoclimatic and palaeoceanographic variability in marine sediments: Interpretation and Application: Chapter 14, *Developments in Marine Geology*, v. 1, p. 567–644.

Phillips II, N.D., 1991, Refined subsidence analysis as a means to constrain late Cenozoic fault movement, Ventura Basin, California: MS Thesis, The University of Texas at Austin, 121 p.

Templ, M., P. Filzmoser, and C. Reimann, 2008, Cluster analysis applied to regional geochemical data: Problems and possibilities: *Applied Geochemistry*, v. 23, p. 2198–2213.

Wahlman, G.P., and D.R. Tasker, 2013, Lower Permian (Wolfcampian) carbonate shelf-margin and slope facies, Central Basin Platform and Hueco Mountains, Permian Basin, West Texas, USA: in Verwer, K., Playton, T.E., and Harris, P.M., eds., *Deposits, architecture and controls of carbonate margin, slope and basinal settings: SEPM Special Publication 13*, p. 305–333.

Using Hierarchical Cluster Analysis to Improve Facies Definitions in Permian Mudrocks (Wolfcamp and Lower Leonard), Midland Basin, Texas; Part I: Background

Problem statement and objectives

Mudrocks are notoriously difficult to describe. XRF analysis with calibrated hand-held instruments gives quantitative elemental results that provide insight into mineralogical composition of these fine-grained, dark-colored rocks, when supplemented by mineralogical analysis (XRD). XRF analysis generates a wealth of elemental information that can form the basis for better facies definitions. But properly interpreting the abundance of elemental data is problematic.

Application

Core from lower Leonard/upper Wolfcamp in southern Midland Basin was described and analyzed for mineralogical content (XRD) and elemental content (XRF). Initial facies definitions were based on visual examination supplemented by major element (Ca, Si) percentages. Then, facies analysis was refined by iterative hierarchical cluster analysis (HCA) of XRF data.

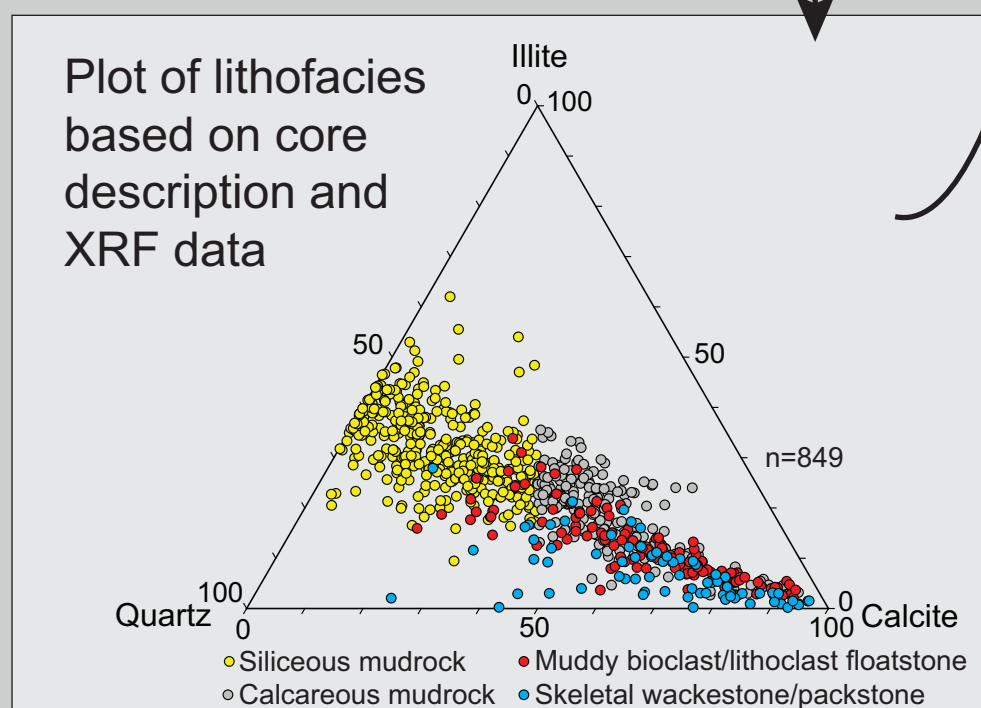
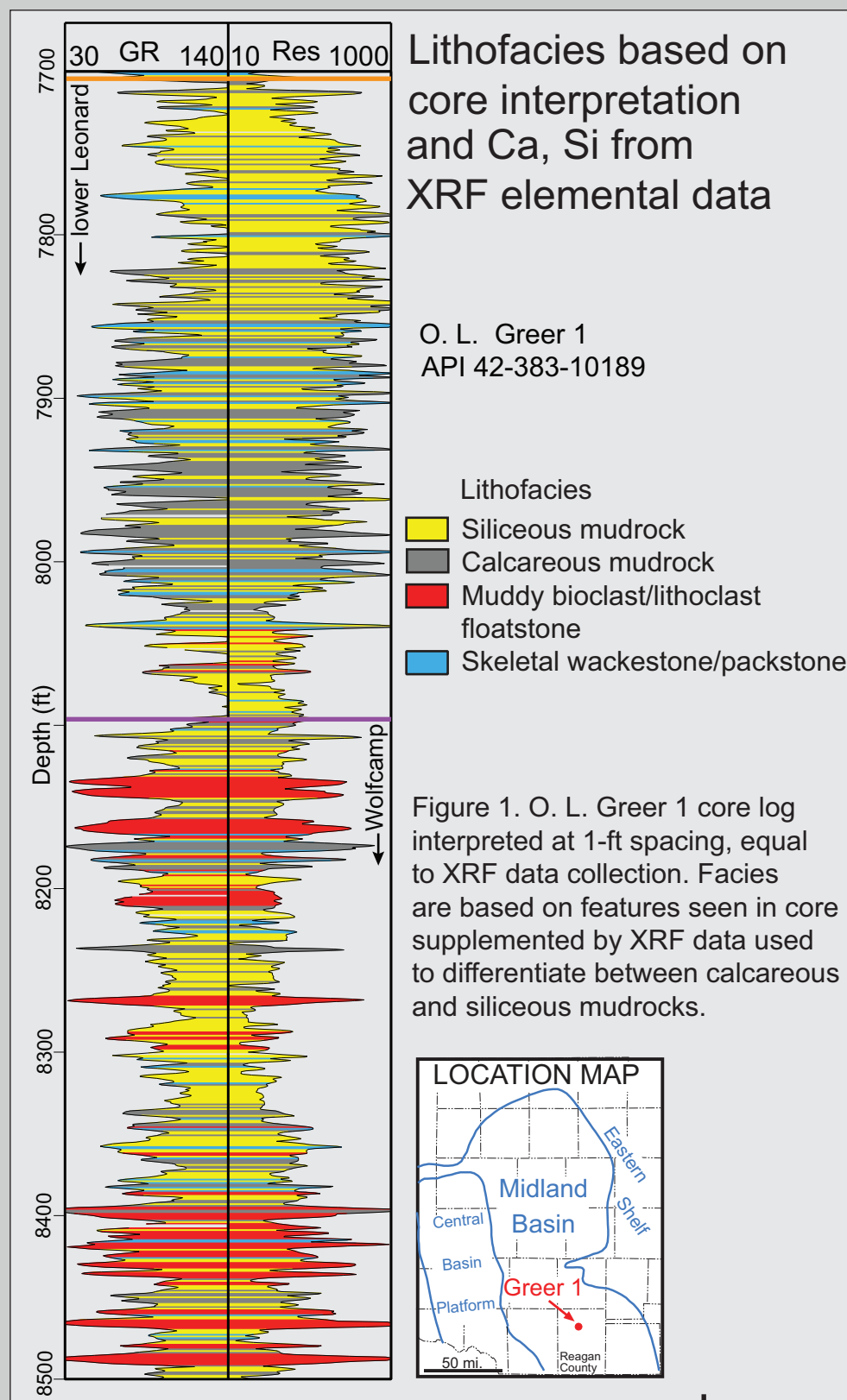


Figure 2. Facies interpreted on basis of core description and XRF data. Quartz, illite, and calcite based on Si, Al, and Ca measured by XRF (see equations below). But to exploit the full range of elemental data provided by XRF scanning, further analysis is required.

Equations used to calculate mineralogy from elemental data

- 1) Calcite (%) = $\frac{\text{Ca}_{\text{meas}} \times 100}{40}$. [40 = molar wt of Ca. Assumes all Ca is in calcite. Confirmed by XRD data and TIC_{meas} vs Ca plot (not shown)].
- 2) Clay minerals (%) = $\frac{\text{Al}_{\text{meas}} \times 100}{k1}$. [K/Al ratio indicates most K is in illite. XRD data show most clay is illite.] (Algeo et al., 2007).
- 3) Quartz (%) = $\frac{\text{SiO}_2(\text{meas}) - (\text{Al}_{\text{meas}}/27 \times k2 \times 60.1)}{100}$. [Assumes all Si not in illite is in quartz.] (Algeo et al., 2007).

where: k1 = average concentration of Al in illite, the dominant clay mineral.
k2 = 1.26 = Si/Al molar ratio in illite.
27 = molar wt of aluminum, 60.1 = molar wt SiO₂

Robert W. Baumgardner, Jr., and Harry D. Rowe
Bureau of Economic Geology
The University of Texas at Austin

Stratigraphic plots of single elements from XRF scans

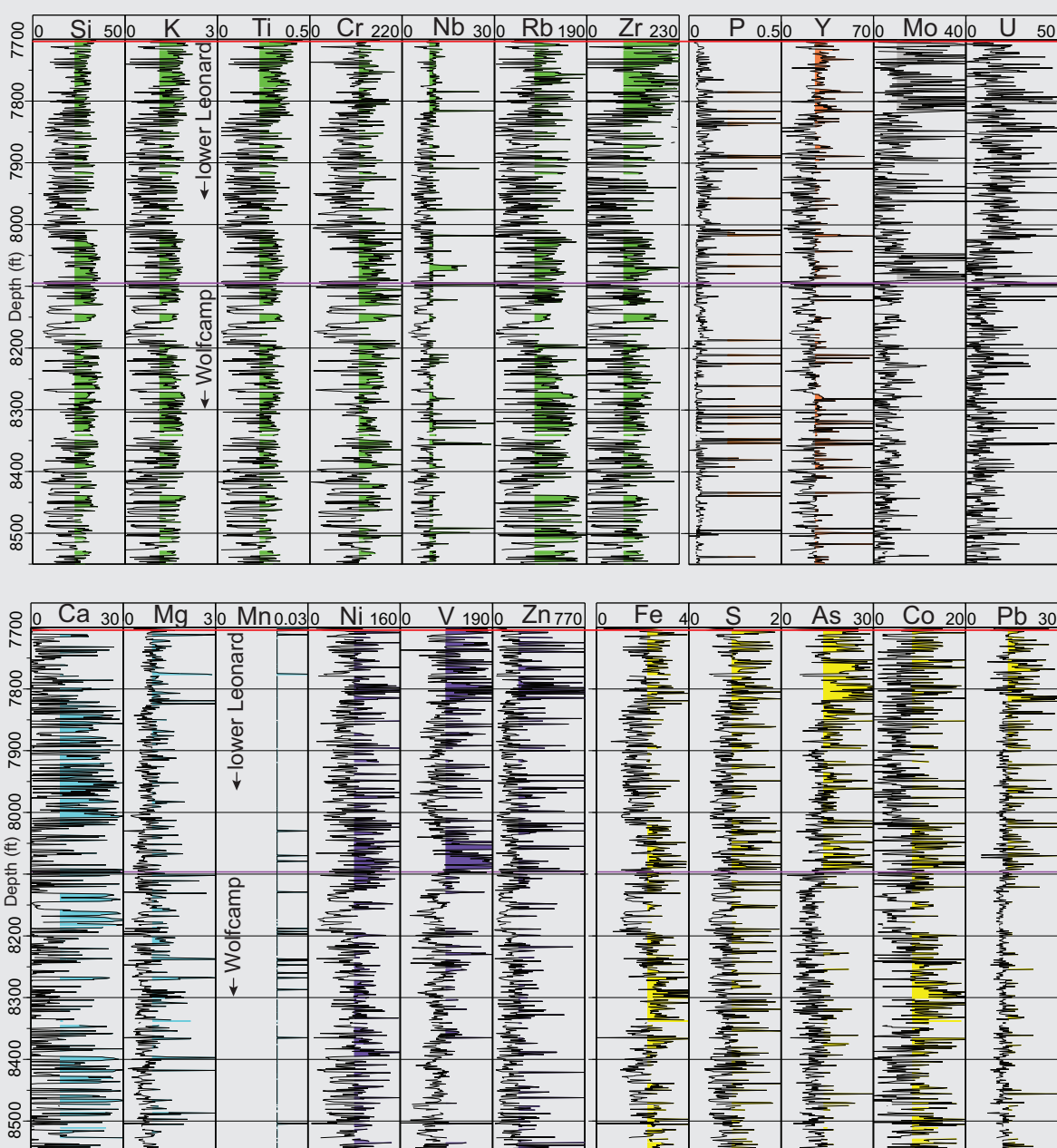


Figure 3. The sheer abundance of XRF data is a challenge to interpret. More than 20 major and trace elements are measured. At this stage, cluster analysis can be used in an exploratory approach to the data to sort them into meaningful groups. Cluster analysis is properly used as a tool of discovery, revealing useful associations and structures in the data.

"The principal aim of cluster analysis is to partition multivariate observations into a number of meaningful multivariate homogeneous groups...A good outcome of cluster analysis will result in a number of clusters where the observations within a cluster are as similar as possible while the differences between the clusters are as large as possible."

Templ et al., 2008

Partition Index values show relative enrichment of elements in each cluster

cluster 63 n = 2	cluster 13 n = 32	cluster 25 n = 55	cluster 14 n = 133	cluster 19 n = 160	cluster 35 n = 24	cluster 12 n = 133	cluster 11 n = 83	cluster 37 n = 81	cluster 10 n = 125
element	element	element	element	element	element	element	element	element	element
	Mo	Mo	Mo						
	6.59	Zr	2.24						
	3.08	V	1.69	Zr	1.44				
	2.31	Zn	1.67	Cr	1.43				
U	26.88	As	2.24	As	1.63	%Si	1.42		
Nb	10.49	Co	2.17	%Ti	1.62	%K	1.39		
Pb	5.74	Ni	1.79	%S	1.59	%Ti	1.35	Rb	1.80
Co	4.09	Zr	1.77	%K	1.38	Rb	1.33	%K	1.53
Ni	3.52	%S	1.63	%Si	1.28	V	1.29	Zr	1.49
Zn	3.11	Pb	1.52	Rb	1.28	Co	1.29	Co	1.42
Y	2.50	Rb	1.52	Nb	1.25	Ni	1.25	%Ti	1.35
As	1.94	%K	1.49	Y	1.23	%Si	1.22	%Si	1.33
%S	1.64	%Ti	1.44	Co	1.21	Zn	1.22	Ni	1.18
Cr	1.51	%Fe	1.35	Pb	1.18	Nb	1.20	Nb	1.18
%Si	1.41	%Si	1.34	Ni	1.15	Mo	1.14	Cr	1.18
%K	1.35	Cr	1.32	Cr	1.14	%S	1.07	%Fe	1.13
V	1.33	Nb	1.12	%Fe	1.12	%Fe	1.03	Y	1.09
%Ti	1.27	Y	1.00	%Mg	1.05	Pb	1.00	Zn	1.02
%Mn	1.00	%Mn	1.00	%Mn	1.00	%Mn	1.00	%Mn	1.00
%Fe	0.96	%Mg	0.82	U	0.98	As	0.99	%S	0.98
%Mg	0.73	U	0.69	%P	0.24	U	0.75	Pb	0.95
%P	0.29	%P	0.34	%Ca	0.19	%Mg	0.69	V	0.80
%Ca	0.05	%Ca	0.05	%Ca	0.09	%Ca	0.09	%Ca	0.09
Mo	0.00			%Ca	0.09	%Ca	0.09	%Ca	0.09
Rb	0.00			%Ca	0.09	%Ca	0.09	%Ca	0.09
Zr	0.00			%Ca	0.09	%Ca	0.09	%Ca	0.09

Table 1. Partition Index (PI) values (Phillips, 1991) for clusters defined by iterative hierarchical cluster analysis (HCA). Elements are ranked from highest (most-enriched) to lowest (most-depleted) relative to their average value. Colors indicate elements that are interpreted as indicators of the same geochemical or depositional factor (e.g., blue = presence of carbonate).

HIERARCHICAL CLUSTER ANALYSIS

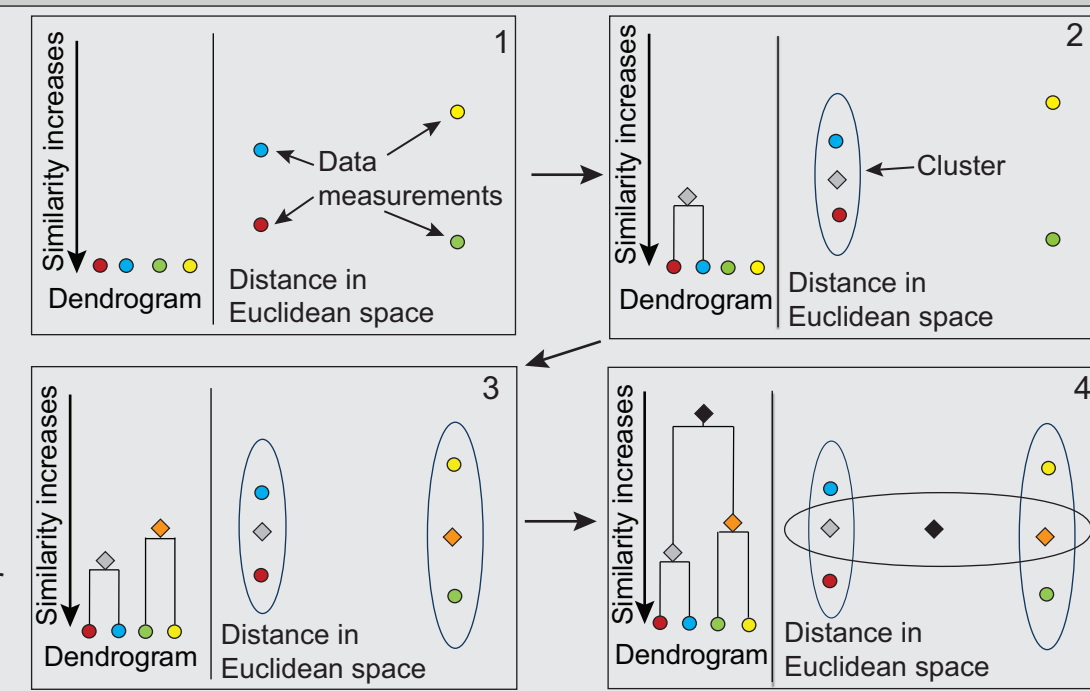
Explanation of clustering technique

Figure 4. Hierarchical cluster analysis gathers elemental data into nearest-neighbor groups of elements.

Data measurements are clustered based on proximity in 2-D Euclidean space, a.k.a. similarity.

A hierarchy of similarity is built as data measurements are grouped into clusters.

All uncorrelated elemental data (22 major and trace elements) were used to define clusters.



Clustering dendrogram

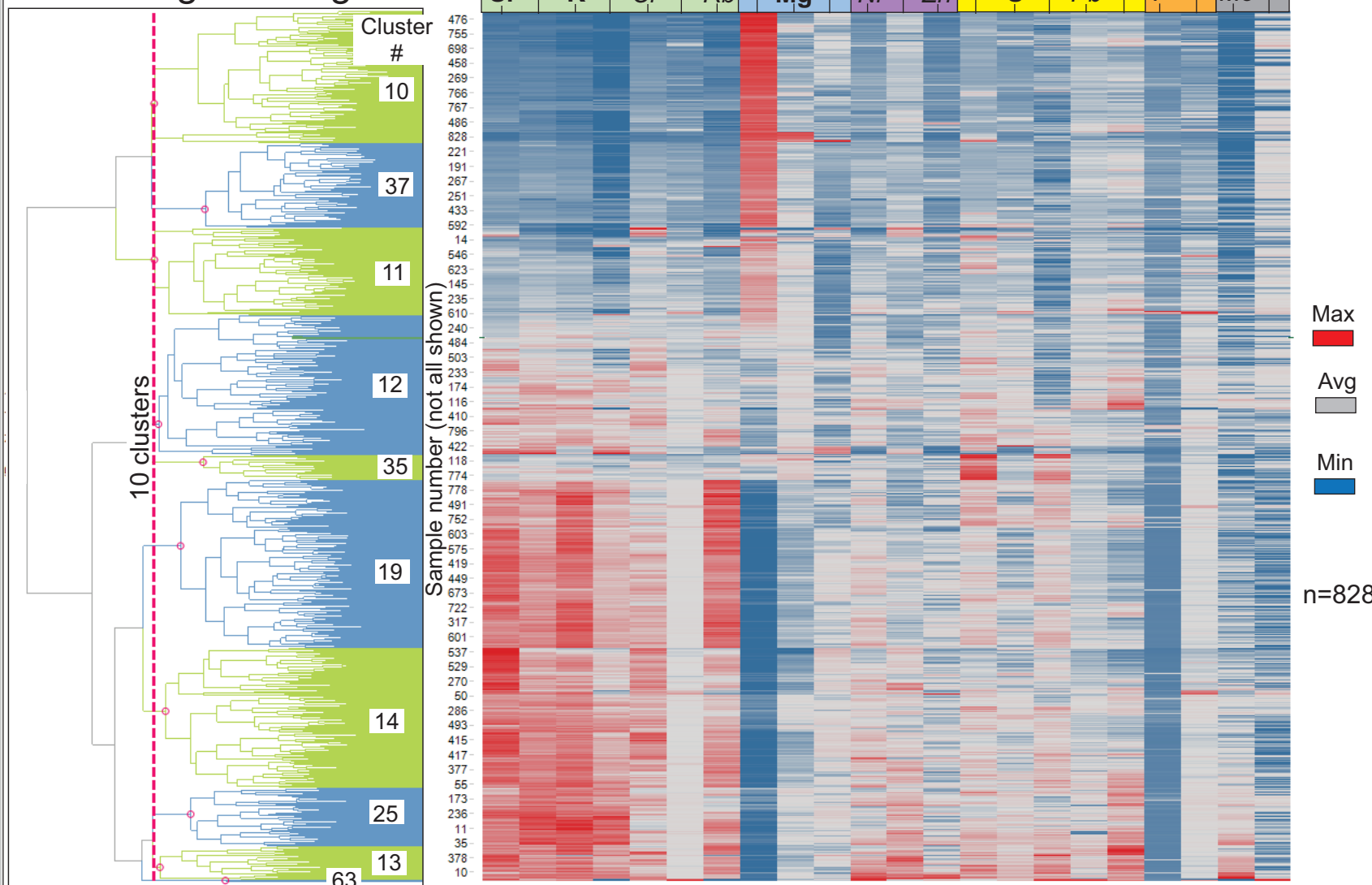


Figure 5. Cluster analysis is an iterative process. The number of clusters is, in part, a judgment call by the analyst, based on knowledge of the rocks and level of detail needed to delineate facies based on the clusters. In this case, 10 clusters were chosen to differentiate between different levels of detrital enrichment (red area at lower left).

Cluster analysis "distills" elemental data from XRF into a manageable number of clusters, which can then be analyzed in terms of elemental abundance and association and further interpreted as facies.

Plot of lithofacies classified by initial cluster analysis

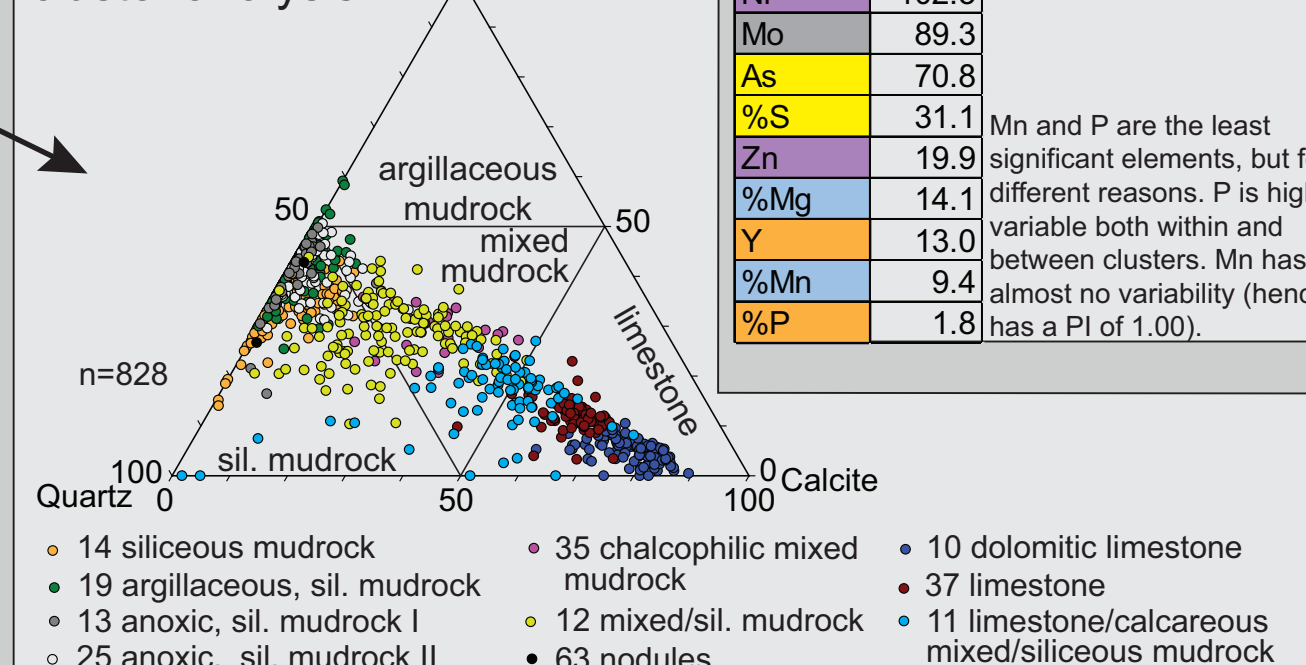


Figure 6. Plot of preliminary clusters shows that clustering provides basis for identifying facies that are not visible in core (e.g., clusters 13, 14, 19, 25, compared to Fig. 2). However, clustering "smears" data between recognized rock types (limestones, mixed mudrocks, and siliceous mudrocks), if cluster is based on elements not represented by apices of ternary diagram (e.g., P in cluster 11). One solution is to separate data into major rock types before clustering, herein referred to as "categorized hierarchical clustering analysis".

See main poster for explanation.

Using Hierarchical Cluster Analysis to Improve Facies Definitions in Permian Mudrocks (Wolfcamp and Lower Leonard), Midland Basin, Texas; Part II: Categorized HCA

Problem statement and objectives

Calibrated X-ray fluorescence (XRF) scanning of core generates large amounts of elemental data. Determining which elements are most important for characterizing rocks can be daunting. One method in use is hierarchical clustering analysis (HCA), which clusters geologic data without regard to lithofacies. In contrast, using “categorized” HCA, the analyst subdivides the data into categories—major rock types (such as limestone, and siliceous and mixed mudrocks)—beforehand. With presorted data, categorized HCA avoids grouping disparate rock types into clusters based on similar amounts of minor rock constituents, which “smears” the distinction between recognized rock types. The goal of this work is to develop a systematic approach to analysis of XRF data that efficiently incorporates geochemical data into standard core description and lithofacies delineation.

Pros and cons of HCA

Advantages of cluster analysis:

- 1) Uses quantitative (elemental XRF) data
- 2) Treats elemental data as an assemblage, i.e., like a rock
- 3) With use of partitioning index and analysis of variance, can determine which elements are most important for defining clusters

Disadvantages of cluster analysis:

- 1) Results are dependent on the current data set, are not directly transferable to other data sets (other cores or basins)
- 2) A few anomalously high values for a single element can give appearance of a cluster dominated by that element
- 3) Requires repetition to achieve best results

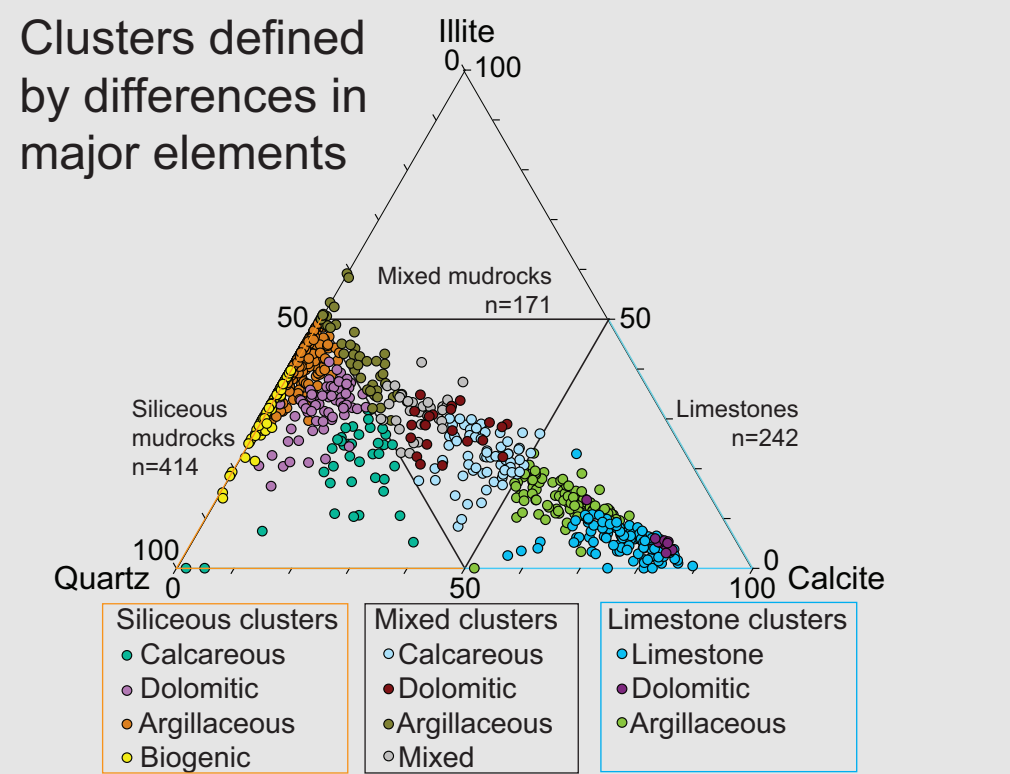
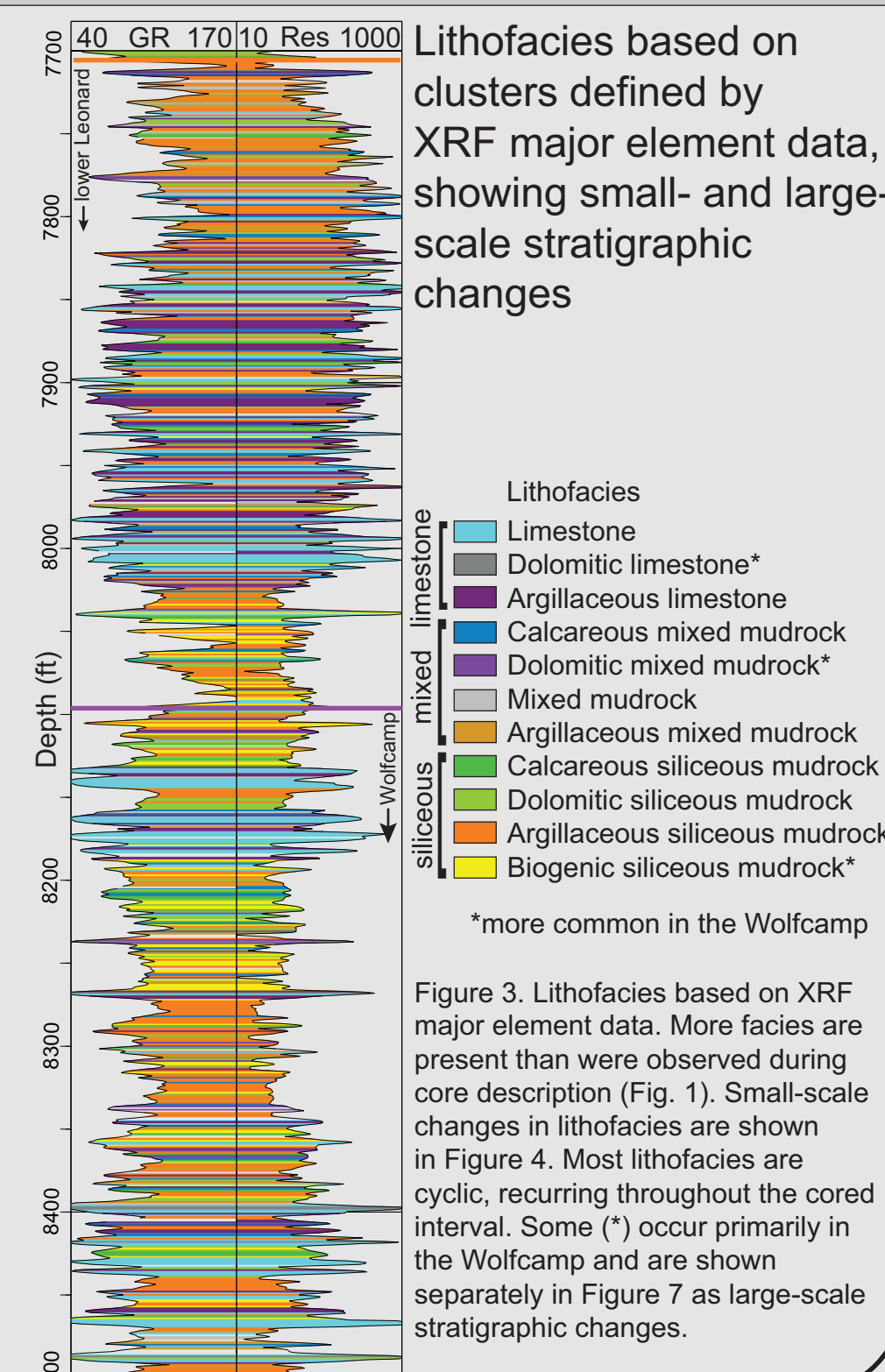
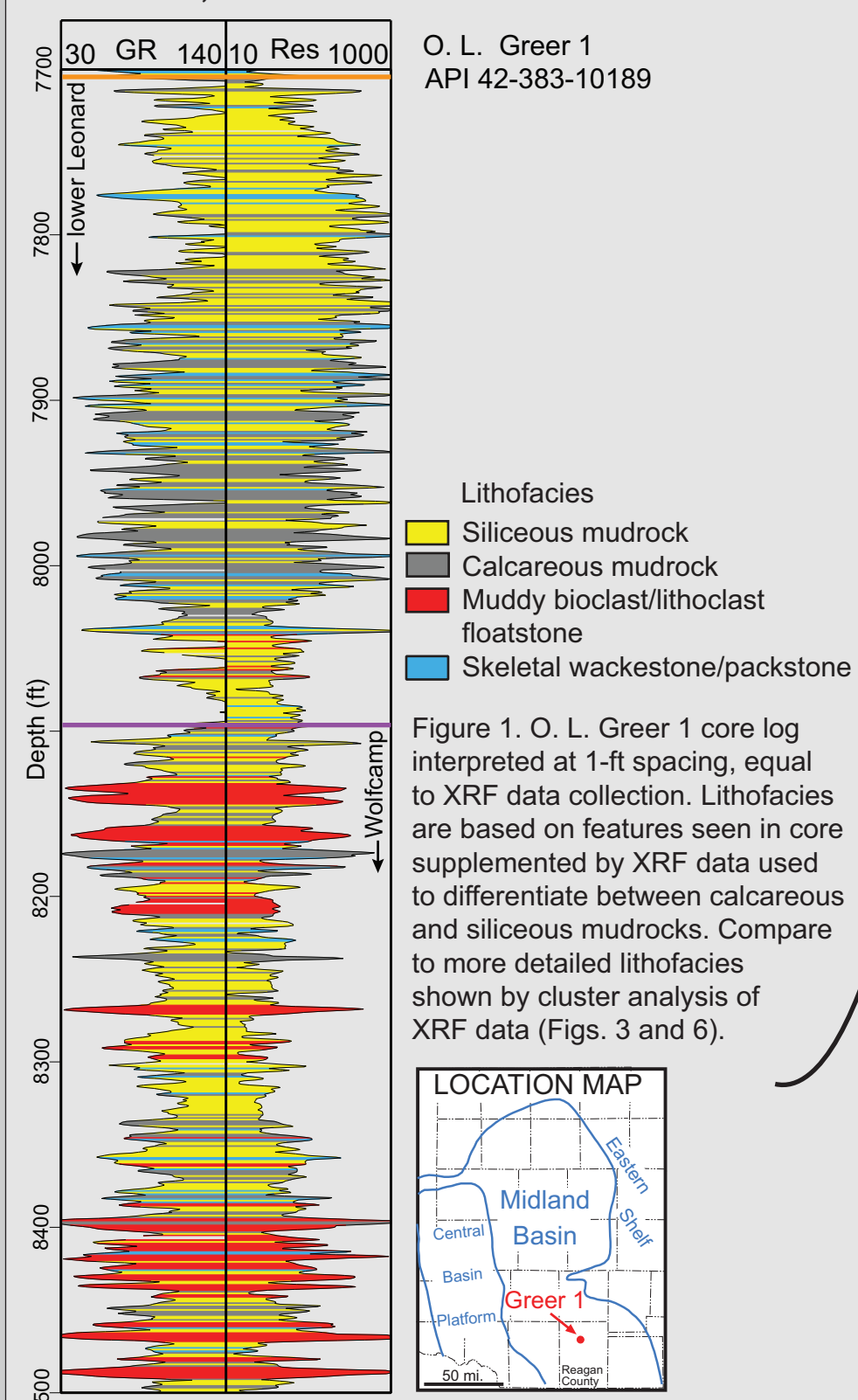


Figure 2. Ternary diagram of calibrated XRF data. Quartz, illite, and calcite based on Si, Al, and Ca from XRF data (see background poster for details). Categorized HCA was applied to data already classified into major rock types (limestone, mixed and siliceous mudrocks), yielding clusters interpreted from abundance of major elements. Presence of dolomite and biogenic silica requires confirmation with XRD and/or thin-section study.

Lithofacies based on core interpretation and Ca, Si from XRF elemental data



Categorized HCA defined lithofacies based on geochemical differences not seen in core

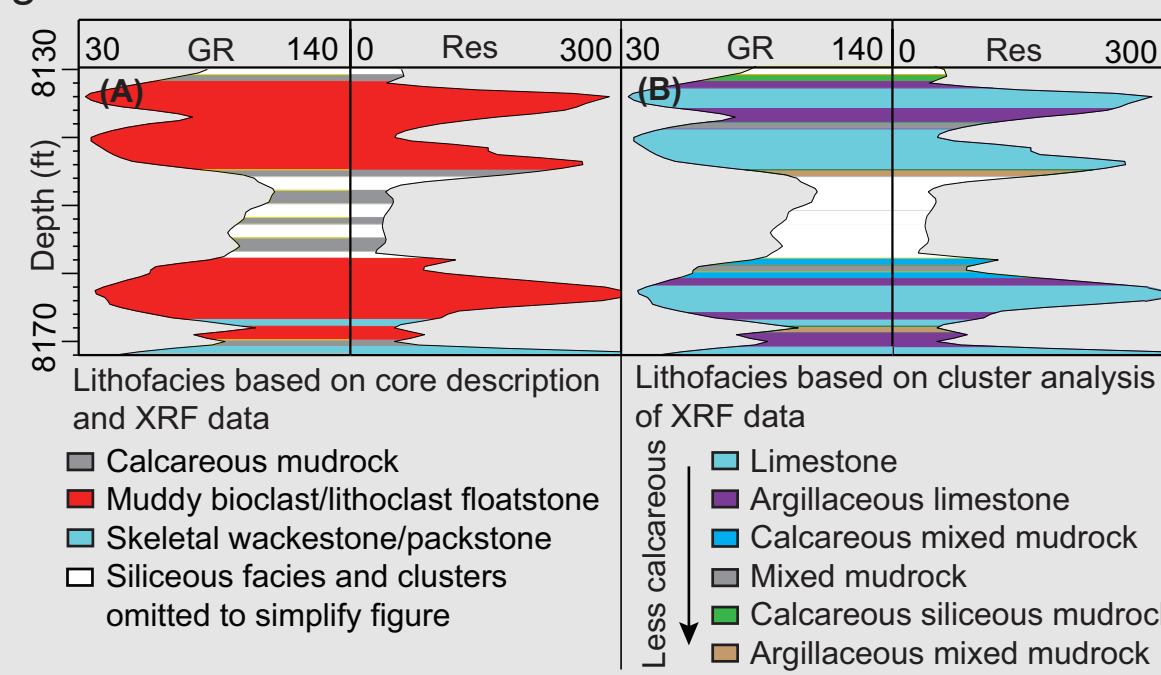


Figure 4. (A) Carbonate-rich lithofacies defined by core description supplemented by XRF data. (B) Lithofacies based on cluster analysis of XRF data. Cluster-based lithofacies reflect cyclic changes in geochemistry, which are not visible in core. Muddy bioclast/lithoclast floatstones (A) are commonly composed of layers of limestone (B) sandwiched between layers of less calcareous, more argillaceous and siliceous mudrock.

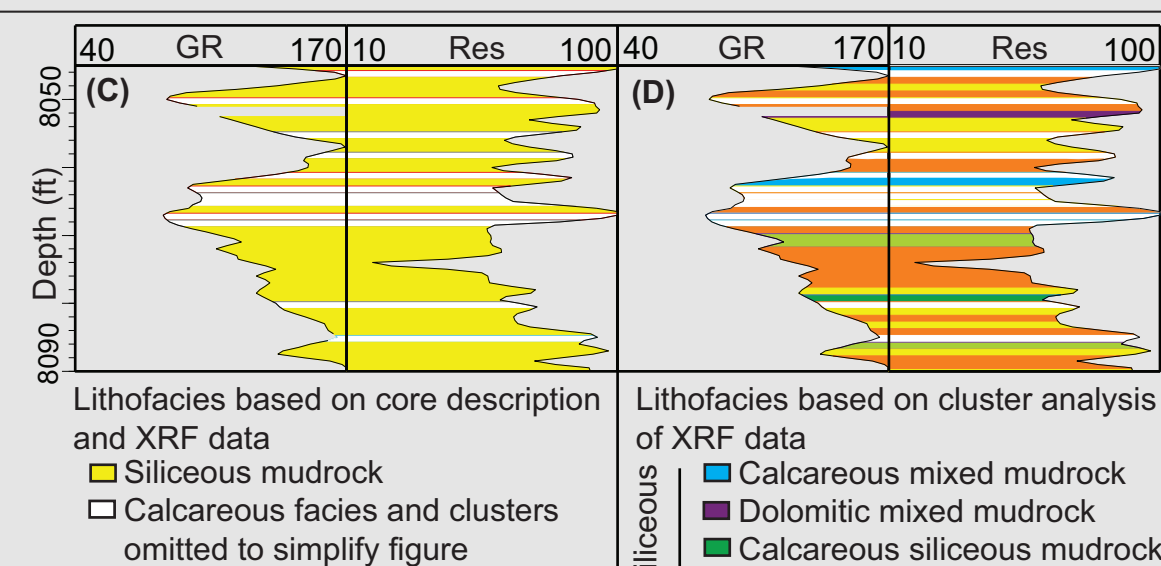


Figure 5. Ternary diagram of calibrated XRF data. Quartz, illite, and calcite are based on Si, Al, and Ca data from XRF data (see background poster for details). Clusters are defined by abundance of trace elements. Most clusters with enriched traces are siliceous. Most clusters with depleted traces are limestones, indicating that carbonate deposition interrupts anoxia, detrital input, and accumulation of organic matter. Stratigraphic differences shown in Figures 6 and 7.

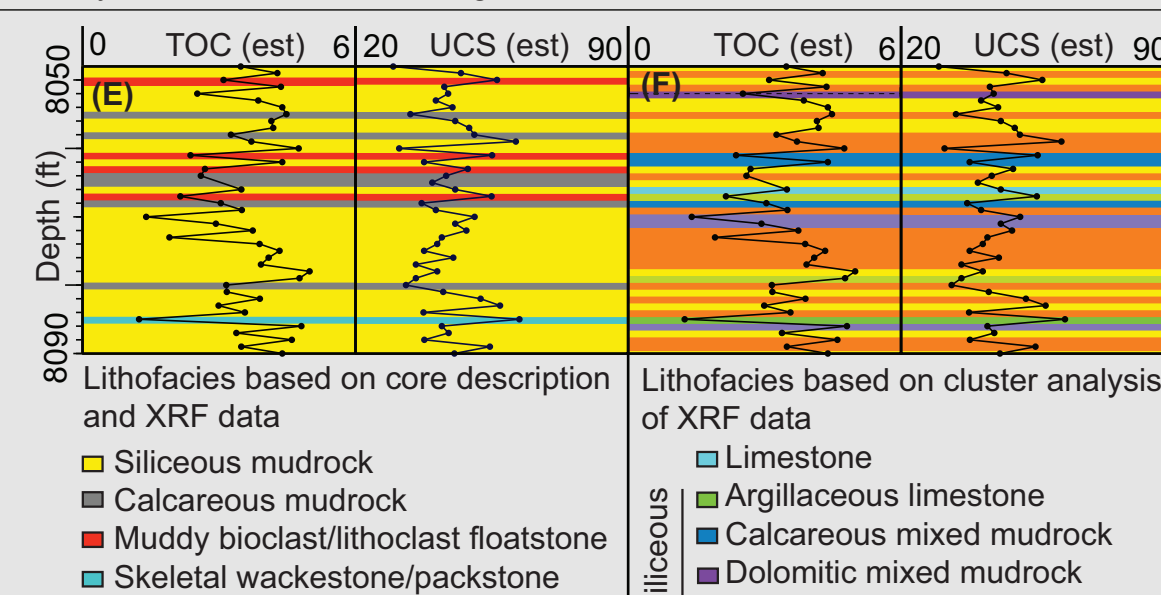


Figure 6. Cluster names are based on interpretations of trace-element abundance. Most clusters are cyclic, showing no change in abundance stratigraphically. Stratigraphic changes are interpreted to reflect changes in ocean water chemistry and sediment source over time. Clusters that show stratigraphic changes (+) are displayed separately in Figure 7.

Trace elements interpreted as indicators of depositional environment

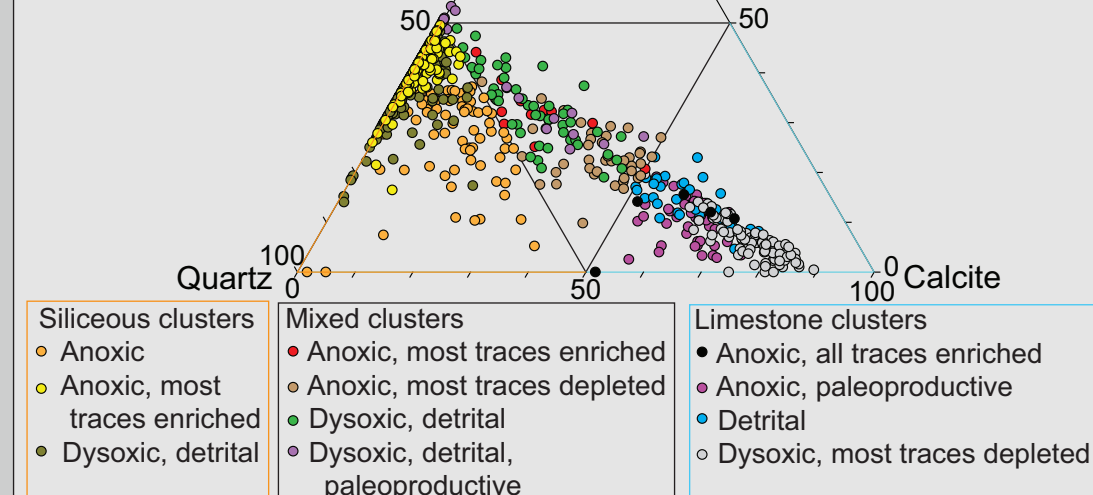
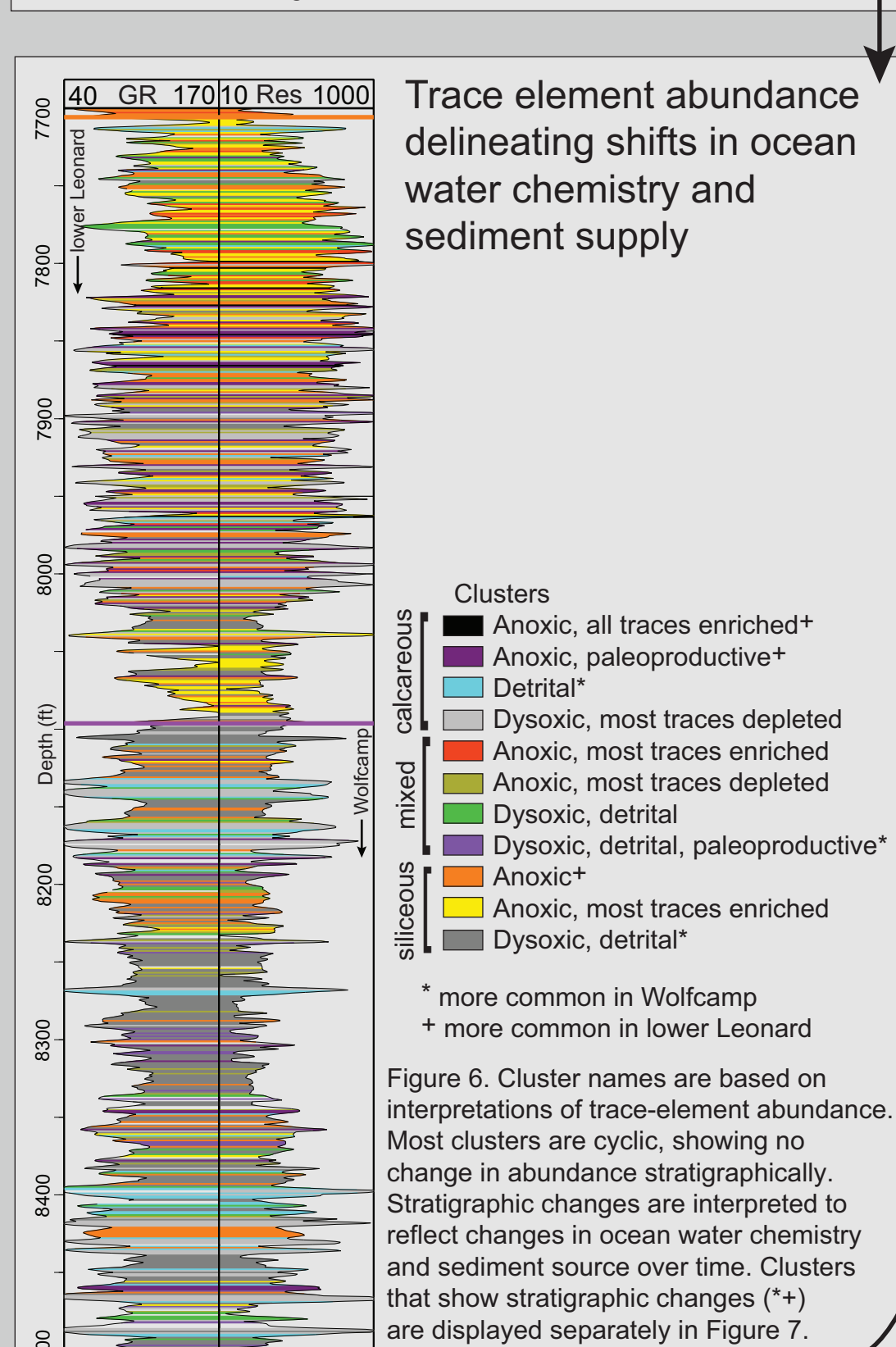


Figure 7. Some lithofacies and clusters are more common in the Wolfcamp than in the lower Leonard and vice versa. Lithofacies and clusters more common in the Wolfcamp tend to be more dolomitic, with more biogenic silica-enriched in detrital proxies but depleted in proxies for anoxia. Clusters more common in the lower Leonard are enriched in proxies for paleoproductivity and anoxia.



Categorized HCA delineating large-scale geochemical, stratigraphic changes in some lithofacies and clusters

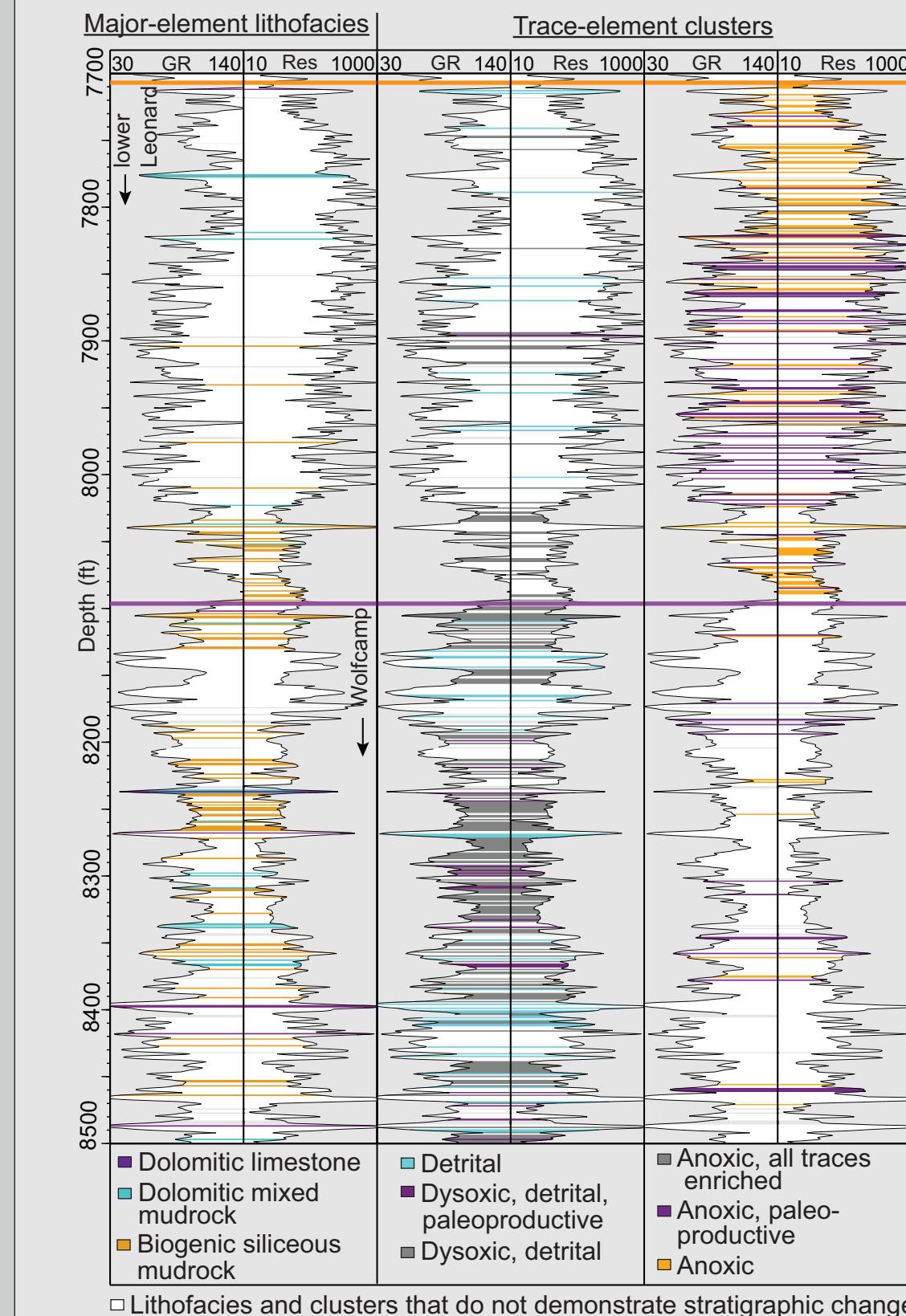


Figure 9. Some lithofacies and clusters are more common in the Wolfcamp than in the lower Leonard and vice versa. Lithofacies and clusters more common in the Wolfcamp tend to be more dolomitic, with more biogenic silica-enriched in detrital proxies but depleted in proxies for anoxia. Clusters more common in the lower Leonard are enriched in proxies for paleoproductivity and anoxia.

- Confirm interpretations of lithofacies and clusters with analysis of thin sections and XRD mineralogical data.
 - Incorporate confirmed lithofacies and trace-element clusters into revised lithofacies names/descriptions.
- ## Conclusions
- Categorized HCA should be done **before** core description, in order to incorporate geochemical data into lithofacies descriptions and to guide collection of more expensive data, such as XRD, TOC, Rock-Eval, and thin sections.
 - Categorized HCA is best used as an initial survey tool to find large-scale geochemical changes and, if desired, select small-scale changes for further study.
 - Categorized HCA revealed sub-meter-scale geochemical changes, e.g., argillaceous and biogenic siliceous mudrock lithofacies correlated with changes in estimated rock strength and TOC.
 - Categorized HCA detected large-scale geochemical changes—e.g., from late Wolfcampian to early Leonardian time detrital proxies decrease while proxies for anoxia and paleoproductivity increase—that coincide with relative rise in sea level, suggesting eustatic influence on anoxia, paleoproductivity, and detrital input in the deep basin.

References

- Algeo, T. J., Hannigan, R., Rowe, H., Brookfield, M., Baud, A., Krystyn, L., and Ellwood, B. B., 2007, Sequencing events across the Permian-Triassic boundary, Guryul Ravine (Kashmir, India): Palaeogeography, Palaeoclimatology, Palaeoecology, v. 252, p. 328–346.
- Calvert, S. E. and Pedersen, T. F., 2007, Elemental proxies for palaeoclimatic and palaeoceanographic variability in marine sediments: Interpretation and Application. Chapter 14, Developments in Marine Geology, v. 1, p. 567–644.
- Phillips II, N. D., 1991, Refined subsidence analysis as a means to constrain late Cenozoic fault movement, Ventura Basin, California. MS Thesis, The University of Texas at Austin, 121 p.
- Templ, M., Filzmoser, P., and Reimann, C., 2008, Cluster analysis applied to regional geochemical data: Problems and possibilities: Applied Geochemistry, v. 23, p. 2198–2213.
- Wahlman, G. P. and Tasker, D. R., 2013, Lower Permian (Wolfcampian) carbonate shelf-margin and slope facies, Central Basin Platform and Hueco Mountains, Permian Basin, West Texas, USA, in Verwer, K., Playton, T. E., and Harris, P.M., eds., Deposits, architecture and controls of carbonate margin, slope and basinal settings: SEPM Special Publication 13, p. 305–333.

Acknowledgments

The authors gratefully acknowledge support of MSRL members: Anadarko, Apache, BHP, BP, Cenovus, Centrica, Chesapeake, Chevron, Cima, Cimarex, Concho, ConocoPhillips, Cypress, Devon, Encana, ENI, EOG, EXCO, Exxon-Mobil, FEI, Hess, Husky, IMP, Kerogen, Marathon, Murphy, Newfield, Oxy, Penn Virginia, Penn West, Pioneer, QEP, Samson, Shell, Statoil, Talisman, Texas American Resources, The Unconventionals, US EneerCorp, Valence, and YPF.