

## Effects of data categorization on paleocommunity analysis: A case study from the Pennsylvanian Finis Shale of Texas

Frank L. Forcino,<sup>1,2\*</sup> Emily S. Stafford,<sup>2</sup> Jared J. Warner,<sup>1</sup> Amelinda E. Webb,<sup>3</sup> Lindsey  
R. Leighton,<sup>2</sup> Chris L. Schneider,<sup>2,4</sup> Tova S. Michlin,<sup>5</sup> Lauren M. Palazzolo,<sup>1</sup> Jared R.  
Morrow,<sup>1</sup> and Stephen A. Schellenberg<sup>1</sup>

<sup>1</sup>San Diego State University, Department of Geological Sciences, San Diego, California 92182, USA; <sup>2</sup>University of Alberta, Earth and Atmospheric Sciences Department, Edmonton, Alberta T6G 2E3, Canada; <sup>3</sup>Yale University, Department of Geology and Geophysics, New Haven, Connecticut 06511, USA; <sup>4</sup>Alberta Geological Survey, Edmonton, Alberta T6B 2X3, Canada; <sup>5</sup>University of California, Davis, Department of Geology, Davis, California 95616, USA  
e-mail: forcino@ualberta.ca  
\*Corresponding author.

Keywords: paleoecology, sampling method, analytical method, ordination

### ABSTRACT

Paleocommunity research efforts have explored a multitude of faunal assemblages using a wide range of sampling and analytical methods to infer a paleoecological signal. Here, we derive six secondary datasets from a single stratigraphic series of faunal assemblages in the Finis Shale (Pennsylvanian) of Jacksboro, Texas, USA, using a variety of data categorization decisions (i.e., abundance versus calcified biomass, all taxa versus selected indicator taxa, and generic versus higher clade resolution). Biomass- and abundance-derived datasets were not significantly different in terms of evenness, Shannon's information index, or Simpson's diversity index. Using Bray-Curtis and nonmetric multidimensional scaling ordinations, with Sorenson and relative Sorenson distance measures, ordination axis scores of the six derived datasets were all significantly correlated with one another, suggesting little difference in their respective paleoecological signals. Three potential explanations for this consistent paleoecological signal, regardless of which data categorizations are employed, include: (1) the dominance of a few brachiopod taxa overwhelmingly influenced the community structure, (2) relatively constrained environmental conditions limited community variation, and (3) low variation in specimen size minimized potential differences among abundance and calcified biomass categorizations. We suggest that other datasets with greater diversities, greater evenness, or from a wider range of paleoenvironments might not show this consistency. Thus, to the degree possible and appropriate, paleoecological investigators should test the effects of these data categorization decisions on a paleoecological signal, regardless of the analytical method employed.

