Identifying Biogenic Pyrite Using Random Forests™ Analysis of In Situ Laser Ablation-ICPMS Trace Element Data

D.D. Gregory[1], M.J. Cracknell[2], M. Figueroa[1], R.R. Large[2], S. Kuhn[2], A.S. Stepanov[2], N. Fox[2], M.J. Baker[2], V.V. Maslennikov[3], I.A. Belousov[2], and T.W. Lyons[1], [1]Department of Earth Sciences, University of California, Riverside CA 92521, USA, [2]CODES ARC Centre of Excellence in Ore Deposits, Private Bag 79, University of Tasmania, Hobart Tasmania, Australia, 7001, [3]456301 Institute of Mineralogy, Ural Division, Russian Academy of Sciences, Miass, Russia.

Introduction: Identifying the presence of life in the deep past and on other planets is one of the important problems facing science today. Obtaining trace or body fossils is unlikely when sampling planets other than Earth due to the level of evolution required to form fossils, paucity of the fossil records, and difficulty in collecting samples. Therefore, we must turn to geochemical techniques to identify chemical conditions produced by, or conducive for life. Traditional bulk sample geochemical techniques can also be frustrated by overprints of secondary metamorphic or hydrothermal fluids or weathering effects obscuring the original geochemical signatures produced during early deposition. To avoid these complications we focus on the analyses of sedimentary pyrite which can be preserved to mid-green schist facies[1]. Sedimentary pyrite is inextricably related to life through biological sulfate reduction. Additionally, pyrite is able to accommodate a wide suite of elements that reflect the trace element composition of the fluids from which it formed and thus the availability of a number of bioessential elements[2,3].

Methods: In this study we present a database of 3830 laser ablation ICP-MS pyrite trace element analyses from 84 different sedimentary formations that formed via biotic reduction of sulfate and hydrothermal settings unrelated to biological activity. These samples were selected to represent as wide a range of variation in geologic time and to represent as many different hydrothermal systems as possible to ensure that pyrite from no known hydrothermal system can be consistently mistaken for biogenic pyrite.

The resultant data set had 12 different trace elements per analysis resulting in a total of over 45,000 data points making traditional interpretation of data plots difficult if not impossible. To overcome this impediment we used Random Forests™, a supervised classification algorithm that is effective at classifying large geological datasets[4]. Random Forests™ induces a large number (500) random decision tree classifiers to cast a vote for the class label of each data point. The winning class (the one with the majority of votes) is the predicted classification.

The data was split into two groups. The first group consisting of 3071 analyses from 56 locations to train the classifier and test the random forest. The other group, consisting of 759 analyses and 28 different locations, was utilized as a blind test of the Random Forest™ classifier.

Group 1 was further subdivided into training and testing groups. Sixty analyses from each different hydrothermal setting and sedimentary pyrite were used to train the random forest and the remaining analyses were used to test it.

Results: The results are very encouraging. The Random forest was able to correctly predict the sedimentary pyrite from the other pyrite types 94% of the time and seafloor hydrothermal pyrtite that may be related to biology between 84% and 88% of the time. To further validate the method we used the trained classifier to predict the classification of the blind test data, which was from locations that are unrelated to those of the training data. These results were similarly encouraging with sedimentary pyrite being correctly identified 81% of the time.

Conclusions: These results suggest that Random Forests™ analysis of pyrite trace element content is a viable method to test whether a given pyrite was formed via a hydrothermal pathway or via biological sulfate reduction. Further development of this technique will make it possible to test for the presence of life on Mars by analyzing pyrite in rocks collected from the surface or direct analyse by future rovers regardless of metamorphic grade of the rocks sampled.

References: [1] Large, R.R. et al., (2007) *Econ. Geol.*, 102, 1233–1267. [2] Gregory, D.D. et al., (2015) *Econ. Geol.*, 110, 1389–1410. [3] Gregory, D.D. et al., (2015) *GCA*, 149, 223-250. [4] Cracknell, M.J., et al. (2014) Aust. J. Earth Sci., 61, 287-304.