The potential for detecting life as do don't know it by fractal complexity analysis

Journal:	*International Journal of Astrobiology*
Manuscript ID:	IJA-AR-12-0371.R1
Manuscript Type:	Research Article
Date Submitted by the Author:	07-May-2013
Complete List of Authors:	Azúa-Bustos, Armando; Pontificia Universidad Catolica de Chile, Departamento de Genetica Molecular y Microbiologia Vega Martínez, Cristian; Pontificia Universidad Catolica de Chile, Departmento de Astronomía y Astrofísica
Keyword:	Life, fractal, complexity

The potential for detecting "life as we don't know it" by fractal complexity analysis

Armando Azua-Bustos[1,2*] Cristian Vega-Martínez[3]

[1]Pontificia Universidad Católica de Chile, Faculty of Biological Sciences, Department of Molecular Genetics and Microbiology, Santiago, Chile. Centro de Estudios Generales, Universidad de los Andes, Santiago, Chile.[3] Instituto de Astrofísica de La Plata (CCT La Plata, CONICET - UNLP). [*]To whom correspondence should be addressed.

E-mail: ajazua@uc.cl

Running title: detecting life by Fractal Complexity

Abstract

Finding life in the Universe entirely different to the one evolved on Earth elsewhere is probable. This is a significant constraint for life-detecting instruments that were sent and may be sent elsewhere in the solar system, as how could we detect life as “we don’t know it”? How could we detect something that we have no prior knowledge of its composition or how it looks like? Here we argue that disregarding the type of lifeform that could be envisioned, all must share in common the attribute of being entities that decrease their internal entropy at the expense of free energy obtained from its surroundings. As entropy quantifies the degree of disorder in a system, any envisioned lifeform must have a higher degree of order than its supporting environment. Here we show that by using fractal mathematics analysis alone, one can readily quantify the degree of entropy difference (and thus, their structural complexity) of living processes (lichen growths and plant growing patterns in this case) as distinct entities separate from its similar abiotic surroundings. This approach may allow the possible detection of unknown forms of life based on nothing more than entropy differentials of complementary data sets. Future explorations in the solar system, like Mars or Titan, may incorporate this concept in their mission planning in order to detect potential endemic lifeforms.

Introduction

There is a chance of finding life "as we don't know it" elsewhere in the Universe. These lifeforms may be entirely different from life that evolved on Earth; not using water as a solvent, DNA/RNA as informational molecules, etc., thus reflecting an independent origin of life (Davies *et al*. 2009). This is a significant constraint not only for life-detecting instruments sent, and to be sent to different bodies of the solar system, but also for life-signature detecting techniques to be applied onto more distant exoplanets.

Life evolves against entropy, keeping the information gained and increasing it (Avery 2003). As Erwin Schrödinger stated (1945), "life feeds on negative entropy". Thus, as lifeforms become more complex, their entropy decreases (Crutcheld & Young 1989). Death causes the opposite effect; that is, the quantitative symmetry in the long term between the entropy of the lifeform and its surrounding environment.

Although a universally accepted definition for complexity has not been reached so far it may yet be stated that life is a complex adaptive system. As such, it contains interdependent constituents that interact nonlinearly, possessing a structure that spans several scales (Baranger 2011).

Thus, one should be able to detect entropy differences (and thus levels of complexity) between lifeforms and the environments where they thrive (Kleidon 2010). It then may be expected that extraterrestrial lifeforms will also have lower entropy states in comparison with similar abiotic phenomena found in their environments. In fact, Lovelock proposed that in order to find signs of life, "one must look for a reduction or a reversal of entropy" (Lovelock 1979). On Earth, even the simplest of microorganisms show a high degree of complexity (Passalacqua *et al*. 2009).

One mathematical tool that allows an objective quantification of complexity is fractal geometry analysis. Fractal geometry analysis can readily analyze the complexity of different structures, in particular of those spanning several scales. It has been previously shown that fractal geometry better approaches the complexity of many life related phenomena at different scales, from tree distribution in forests to neural activity patterns (Losa 2009). This is estimated by calculating the fractal dimension "D", a statistical quantity that indicates the degree of completeness in which a fractal structure appears to fill a data space as finer and finer scales are analyzed. In general, an object may be called a "fractal" if its D value exceeds its topological dimension. Consequently, different D values identify different levels of complexity.

Fractal geometry analysis may then be applied to the examination of interesting candidates elsewhere in the Universe to check whether a higher order of complexity, compared with environmental features of its system, may be indicative of unknown lifeforms, and potentially, also of abiotic processes altered by life processes. This approach has the advantage of requiring no prior information of the potential lifeform to be analyzed. For the analysis of such candidates, a proof of concept needs to be developed, along a proper analytical technique, both of which we describe here.

Material and Methods

Model systems

We chose two different biotic-abiotic model pairs for our analyses, which in addition covered two different orders of magnitude.

At the microscopic level, we analyzed lichen covered rocks in the high Andes Mountains near Santiago, Chile. Here the aim was to compare the structure of lichen growths with the structure of the rock were it develops, in order to see if we could detect a higher complexity associated to the lichen growth in relation to its abiotic surroundings.

At the macroscopic level, we analyzed *Tillandsia* shrub growing patterns on sand dunes of the Atacama Desert. In these dunes, the Bromeliad *Tillandsia landbeckii* self-organize in characteristic growth bands of shrubs in order to maximize the interception of fogs (Borthagaray *et al*. 2010), creating banding patterns that from above, appear very similar to geologically banding structures. Here the aim was to compare *Tillandsia* banding patterns with geological banding patterns of similar inter-band distances, in order to see if we could detect a higher complexity associated to the *Tillandsia* banding in comparison of that of similar bandings of geological origin.

Images

High resolution photographs of lichen covered rocks were taken at the Andes Mountains near Santiago, Chile (FIG. 1A). For the analysis of lichens and rock surfaces, we analyzed 20 bare rocks images (FIG 2A) and 110 lichen images (FIG 2B). The lower number of bare rocks surfaces analyzed was due to the scarcity of unequivocally naturally bare rock surfaces immediately adjacent to lichen colonized areas.

Geological bandings patterns in the vicinity of areas covered by *Tillandsia* bands were obtained by downloading 30 images (FIG. 2C) from Google Earth. Care was taken to consider that the inter-band distance of geological bandings was about 10 meters, which is the mean inter-band distance measured for *Tillandsia* banding growths. 33 *Tillandsia* bandings images (FIG. 2D) were also downloaded from Google Earth. *Tillandsia* covered sites, which are found near the City of Iquique at the Atacama Desert of Chile, were visited as to confirm their biological origin and independency of geological processes (FIG. 1B). This was also confirmed for the banding patterns of geological nature.

In addition, we selected 24 banding patterns images (FIG. 2E) taken by the HiRise camera onboard of the Mars Reconnaissance Orbiter at Meridiani Planum on Mars. These banding patterns are thought to have formed through the accumulation of sediments transported by flowing water. The number of images analyzed in this case corresponded to images with similar inter-band distances to the bandings of Earthly origin.

In all cases the images were cropped in order to ensure the analysis of just the phenomena of interest (FIG. 2).

Development of FrAn, a Fractal Analysis tool

The box-counting method is one of many methods developed for fractal analysis (Soille and Rivet, 1996). It works by covering a set of data (an image in this case) with "boxes" (squares) and then evaluating how many "boxes" are needed to re-describe the data set completely in a metric space. Repeating this measurement with boxes of decreasing sizes results in a logarithmical function of box size (x-axis) and number of boxes needed to cover the pixels image dataset (y-axis) (FIG. 3). The slope of this function is referred as the *box dimension*, which is a considered to be a good approximation of the fractal dimension (Soille and Rivet, 1996).

The software we used for our analyses, "FrAn" (for **Fr**actal **An**alyzer), was created using as a reference a previous program (HarFa), a software that was compiled to perform harmonic and wavelet analysis of digitized images and calculations of their fractal parameters based on the box counting technique (Harfa, 2010). Using this software as an example, we created a new algorithm that also uses the box counting technique, incorporating a new and critical parameter for our aims, the fractal excess (FE) (defined in detail below).

Image data analysis

Using FrAn, we first estimated the fractal dimension of the sets images of lichens-colonized and uncolonized Andean rock surfaces: the images were first trimmed in order to consider unequivocally fully colonized and fully uncolonized selections (FIG. 2). FrAn then transforms these images into grey scale versions, and then to black and white images (FIG. 4). The grey scale to black and white transformation is performed for all possible threshold (T) values (256 in total), where $\bar{X} - 1.5\sigma < T < \bar{X} + 1.5\sigma$, where is the mean intensity of all pixels and σ the standard deviation of all pixel intensities. The importance of the thresholding process may be understood if it is considered that the strength of the signal (the brightness and contrast of the image as determined by the intensity value for each pixel) has a direct effect on the amount of information that can be extracted from any particular image (FIG. 5).In practice, this means that every sample image was analyzed at least seven times in the range of -1.5σ and 1.5σ. In addition, to every image analyzed, a "null model" was created, which corresponded to the same image data set, but randomized (FIG. 2F).

Then, the Minkowski–Bouligand dimension (or box dimension) was calculated; $\ln N(\varepsilon) = D_{box} \ln (1/\varepsilon) + K$, where ε is the box length, N is the number of different boxes of size ε, and K is an arbitrary constant. This was done for all image analyzed and their null counterparts.

After these analyses the computation of the Fractal Excess (FE) was then calculated. The FE is defined as the difference between the fractal spectrum of the sample image fractal spectrum and its null model. Therefore, the incorporation of a null model for each of the samples being analyzed guarantees a true estimation of the complexity of the analyzed images.

Thus, the final output of the image data analysis is a curve of the entire spectra of FE values for each threshold situation. As a result, a low FE value indicates a high fractal dimension and correspondingly, a higher complexity of the sample being analyzed. In turn, high FE values indicate a low fractal dimension, and thus, a lower complexity of the sample.

All analysis were performed with a Hewlett Packard cluster composed by 62 Intel Titanium II processors and a total of 48 GB of RAM available at the Center for bioinformatics of the Faculty of Biological Sciences of the Pontificia Universidad Católica de Chile. Using this processing power, mean analysis time for each image selection was about 20 minutes. FrAn is available on request.

Results

Figure 6 shows that the Fractal Excess spectrum curve for all lichen images analyzed is unique, and different to the Fractal Excess spectrum curve for all bare rock surface images. Bare rock samples show higher FE values than lichen FE values, the latter being lower and closer to the x axis.

Larger distances between both curves are observed at the positive range of the spectra, although appreciable distances can also be seen at the negative range of the. FE minima also reflect these differences, with bare rock surface being 0.26 and Lichen being 0.18 (Figure 7). At both ends of the FE spectra maxima also follow the same pattern, with lichen maxima being lower than bare rock surface values. This result provides good evidence that biotic processes do have a higher complexity than the environment where they occur.

Repeating these analyses in another order of scale show a similar trend. We processed images of banding vegetation patterns of *Tillandsia* shrubs growing on the Atacama Desert, comparing them with geological structures of similar appearance, inter-band distance and size in the surrounding area.

As can be seen in Figure 8, the FE curve profile of vegetation patterns is readily separated from similarly looking geological banding structures. Similarly to the results obtained with lichen growth analysis, *Tillandsia* growing patterns show a curve with lower values than similar abiotic phenomena, thus revealing the higher complexity associated to living phenomena. As for FE minima, they also reflect these differences, with geological banding patterns being 0.28 and *Tillandsia* banding patterns being 0.20 (FIG. 7).

Interestingly, when similar geological banding patterns found in Meridiani Planum on Mars were analyzed, its curve showed a distinctive profile, between those from geological and biological banding patterns on Earth (FIG. 8).

Discussion

As a proof of concept of our approach, we focused on two extreme environments as models, which in addition represent two scales of size (FIG. 1).

First, we analyzed lichens growing on rocks at high altitude on the Andes Mountains of Chile. Lichens are important drivers of biogeochemistry and the first lifeforms to colonize barren environments (Cornelissen *et al*. 2007). Seen from the point of view of an alien civilization, lichens may not be at first recognized as living entities, a hypothetical scenery that emulate a situation in which humans explore an extraterrestrial environment. Hence, we set to determine whether the complexity of lichens could be differentiated from the similarly looking rock background where they grow. By using FrAn we show that indeed, the complexity structure of bare rocks and lichens differ and are readily separable (FIG. 6). Lichen samples have lower FE values than Bare rock samples, the former being closer to the x axis, thus reflecting their higher complexity and entropy. Minima values from both curves reflect this same difference, with the value of lichen minima being much lower than that of bare rock minima (FIG. 7).

In addition, we found that thresholding is useful as a "dial" for extracting the maximum of information for the analyzed images. For the case of the lichen model, thresholding in the range of 0 and 1.5 gives the better resolution.

Thus, although we have the experimental advantage of "knowing" that lichens are living entities, just by analyzing high definition images alone, and with the aid of fractal geometry, we could readily differentiate them as distinct higher entropy entities from its similar abiotic surroundings. By analyzing seasonal time series, further spatial and temporal patterns changes in size and or structure (tridimensional for example) could be incorporated as well.

In order to test our hypothesis at a larger scale, we applied our analysis to banded vegetation patterns of *Tillandsia* species growing on the Atacama Desert, comparing them with geological banding structures of similar appearance found in the area. Similarly to the case of lichen analysis, *Tillandsia* banding growths also showed lower values than those of similar geological banding patterns, thus confirming their higher complexity. Analog to the case of the lichen model, curve minima also reflect these differences, with *Tillandsia* values being lower than geological values (FIG. 8). In this case, thresholding in the range of -1.5 and -0.5 gave a better resolution.

The *Tillandsia* banding pattern analysis results show that even from afar, life can be distinguished as a distinctive entropy process, distinct from similar abiotic processes.

The curve behavior of Mars banding patterns with lower FE values than of Earth abiotic banding patterns but higher than biotic banding patterns remains puzzling for now. As the shapes of the curves of Earth and Mars banding patterns are different, but have similar minima, one may speculate if this could this be either reflecting differences in the geological processes that gave rise to both structures or (more interestingly) some limited influence of life processes in the genesis of Mars strata?

It is important to note that our analysis allows the estimation of the complexity of datasets being analyzed, (images in this case) but as in practice the input data for the FrAn software are pixel intensity numerical values, any type of data can be fitted as well. Thus, future examination of extraterrestrial features could be fractally analyzed with our method at different informational layers; high resolution images, thermal images, elemental and isotopic compositions, pH, etc. In this way, complementary fractal data values could be then overlaid in a matrix to calculate a Lifeform Probability Index (LPI) for extant or past life related phenomena based on their entropy measurements analysis.

Finally, a number of applications may be envisioned for the approach presented here; rovers on distant planets could be "taught" to use the LPI with the aid of artificial intelligence as a way to tag features that human experience intuitively classify as "interesting". Potential fossil structures could be compared with this method, in order to appraise its true biological origin, both on Earth or elsewhere in the Solar System. Radio signals from potential extraterrestrial civilizations could as well be analyzed in search of fractal patterns, as use of fractal mathematics may be considered as a marker of advanced technological development.

Acknowledgements

Armando Azua-Bustos receives a Scholarship for graduate studies from CONICYT Chile, and additional financial support from: the Millennium Institute for Fundamental and Applied Biology (MIFAB), Chile, from FONDECYT project nº 1110597 and from the AngelicvM Foundation. Cristian Vega received financial support from the Center for Bioinformatics, Pontificia Universidad Católica de Chile.

We also thanks Carlos González-Silva, from the CENIMA, Iquique, for his help in field sampling, and Cristián Arenas-Fajardo, for his help with statistical analysis of our data.

References

1.- Avery, J. (2003) in Information Theory and Evolution. 217 pp. World Scientific Pub Co Inc.

2.- Baranger, M. Chaos, Complexity, and Entropy. A physics talk for non-physicists. MIT-CTP-3112. Available online at:

http://www.necsi.edu/faculty/baranger.html

3.- Borthagaray, A.I., Fuentes, M.A. & Marquet, P.A. (2010). Vegetation pattern formation in a fog-dependent ecosystem. *J Theor Biol* **265**, 18-26.

4.- Burlando, B. (1993). The fractal geometry of evolution. *J Theor Biol* **163**,161-172.

5.- Cornelissen, J.H., Lang, S.I., Soudzilovskaia, N.A., & During, H.J. (2007). Comparative cryptogam ecology: a review of bryophyte and lichen traits that drive biogeochemistry. Ann Bot **99**, 987-1001 .

6.- Crutcheld, J.P. & Young, K. (1989). Inferring statistical complexity. Phys Rev Lett **63**, 105-108.

7.- Davies, P.C., Benner, S.A., Cleland, C.E., Lineweaver, C.H., McKay, C.P., & Wolfe-Simon, F. (2009). Signatures of a shadow biosphere. Astrobiology **9**,241-249.

8.- Guarino, V., Guaccio, A., Netti, P.A. & Ambrosio, L. (2010). Image processing and fractal box counting: user-assisted method for multi-scale porous scaffold characterization. J Mater Sci Mater Med **21**, 3109-3118.

9.- HarFa. 2010. Available on line at: http://www.fch.vutbr.cz/lectures/imagesci/

10.- Kleidon, A. (2010). Life, hierarchy, and the thermodynamic machinery of planet Earth. Phys Life Rev **7**, 424-460.

11.- Losa, G.A. (2009). The fractal geometry of life. Riv Biol **102**, 29- 59.

12.- Lovelock, J. (1979) in GAIA – A New Look at Life on Earth. Oxford University Press, 176 pp.

13.- Passalacqua, K.D., Varadarajan, A., Ondov, B.D., Okou, D.T., Zwick, N.E., & Bergman, N.H. (2009). Structure and complexity of a bacterial transcriptome. J Bacteriol **191**, 3203-3211.

14.- Rodríguez-Pascua, M.A, De Vicente, G., Calvo, J.P., Pérez-López, R. (2003). Similarities between recent seismic activity and paleoseismites during the late miocene in the external Betic Chain (Spain): relationship by ‘b’ value and the fractal dimension. J Struct Geol 25: 749-763.

15.- Schrödinger, E. (1945) in What is Life–the Physical Aspect of the Living Cell. The Macmillan Company, 91 pp .

16.- Soille, P. and Rivet, Jean-P. (1996). On the Validity of Fractal Dimension Measurements in Image Analysis. J Vis Comm Ima Rep 7:217-229.

Figure legends

Figure 1.- Sites used as models. A.- Lichen colonized rocks at the Andes Mountains near Santiago. B- *Tillandsia* banding growths on sand dunes at the Atacama Desert near Iquique, Chile.

Figure 2.- Cropped images samples. A, bare rock surface. B, lichen colonized rock surface. C, Atacama Desert geological bands. D, *Tillandsia* banding growths. E, Mars geological bands. F, the same image of E, after being randomized (null model).

Figure 3.- A, The Box-counting technique applied in a theoretical fractal set (E). B, Plotting the box number (N(e)) that includes one part of the fractal set (E) at the very least, versus the size of the box (e). D, the fractal dimension, is the slope of the straight fitted line. Modified from Rodríguez-Pascua *et al*., 2003.

Figure 4.- Image data transformation process. High definition color images were taken of colonized and uncolonized rocks, and areas of the images where only-

colonized and only-uncolonized surfaces were cut and stored. These subsets were then transformed to gray scale and then to black and white images, this last format being the input of the fractal analysis.

Figure 5.- Threshold importance on subsequent data analysis. As the threshold of the black and white image has a direct impact on the amount of information than can reconstruct the original color image, all threshold values between 0 and 256 were taken into account for subsequent fractal analysis.

Figure 6.- Fractal Excess (FE) curves of lichen colonized and bare rock surfaces. Shaded areas for each curve correspond to the standard error of the means.

Figure 7.- Fractal excess minimum statistical analysis. Fractal excess minima of the two models shown were separately analyzed by t student test and one-way ANOVA (Tukey's Multiple Comparison post test) respectively. Means with different letters are significantly different ($P < 0.05$; a >b).

Figure 8.- Fractal Excess curves of different banding processes on Earth and on Mars. Shaded areas for each curve correspond to the standard error.

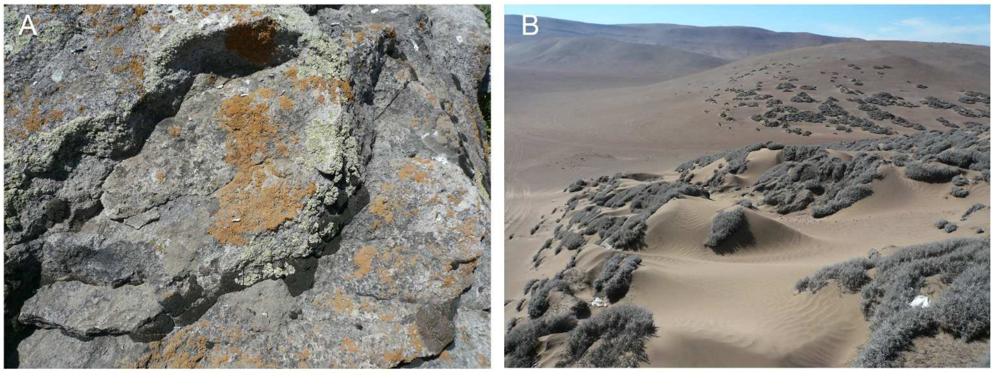

Figure 1.- Sites used as models. A.- Lichen colonized rocks at the Andes Mountains near Santiago. B- Tillandsia banding growths on sand dunes at the Atacama Desert near Iquique, Chile.
453x172mm (199 x 199 DPI)

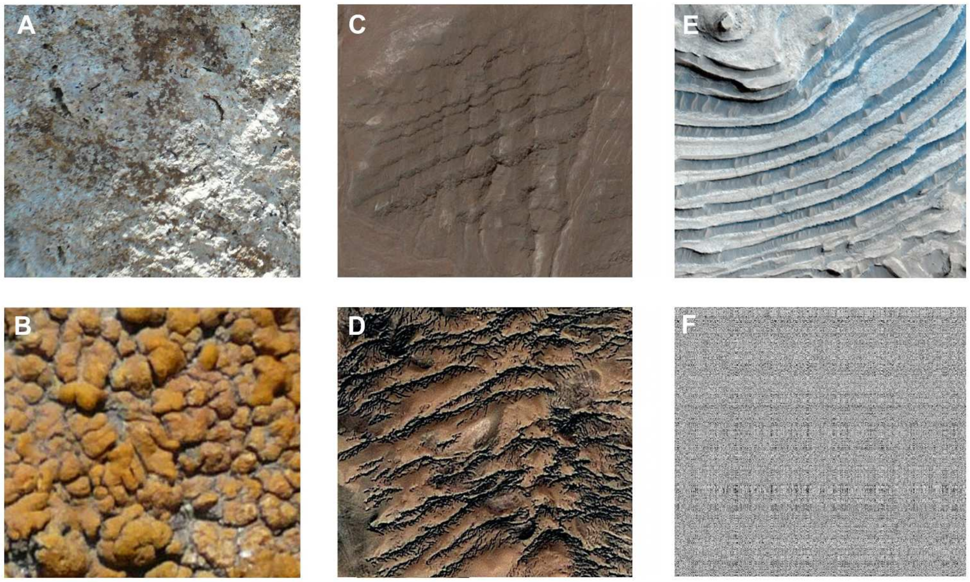

Figure 2.- Cropped images samples. A, bare rock surface. B, lichen colonized rock surface. C, Atacama Desert geological bands. D, Tillandsia banding growths. E, Mars geological bands. F, the same image of E, after being randomized (null model).
382x231mm (200 x 200 DPI)

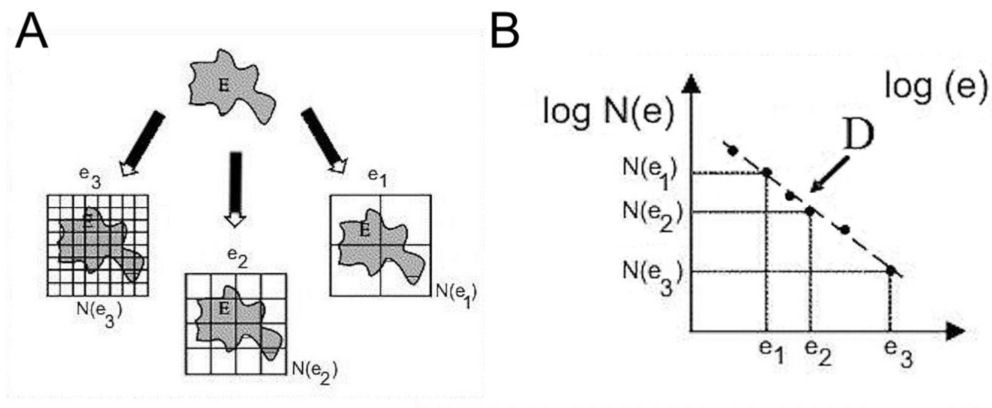

Figure 3.- A, The Box-counting technique applied in a theoretical fractal set (E). B, Plotting the box number (N(e)) that includes one part of the fractal set (E) at the very least, versus the size of the box (e). D, the fractal dimension, is the slope of the straight fitted line. Modified from Rodríguez-Pascua et al., 2003.
169x71mm (200 x 200 DPI)

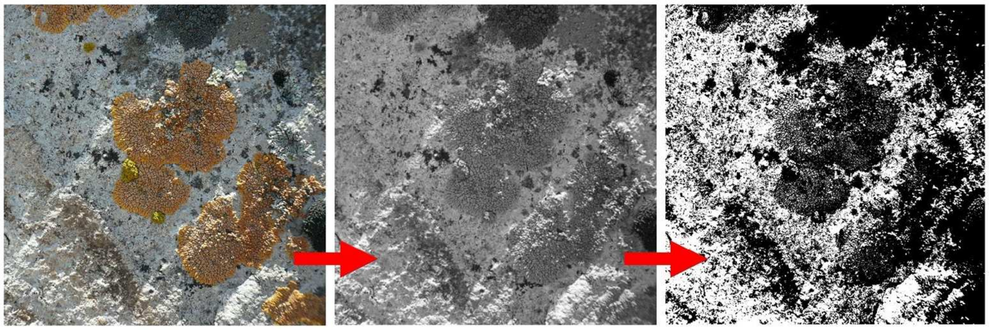

Figure 4.- Image data transformation process. High definition color images were taken of colonized and uncolonized rocks, and areas of the images where only-colonized and only-uncolonized surfaces were cut and stored. These subsets were then transformed to gray scale and then to black and white images, this last format being the input of the fractal analysis.
201x68mm (200 x 200 DPI)

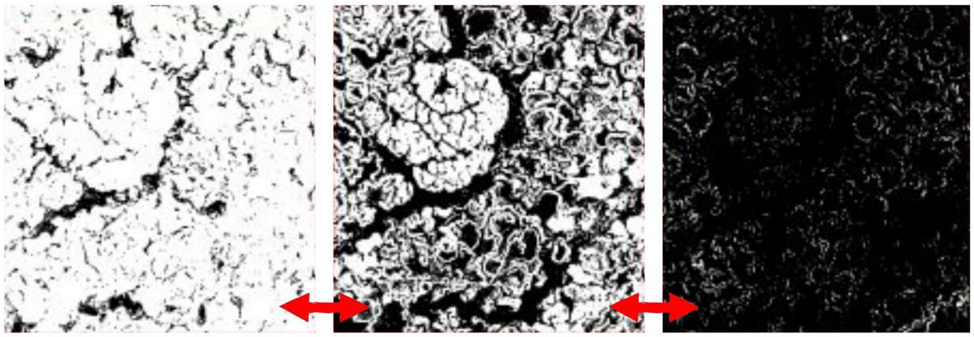

Figure 5.- Threshold importance on subsequent data analysis. As the threshold of the black and white image has a direct impact on the amount of information than can reconstruct the original color image, all threshold values between 0 and 256 were taken into account for subsequent fractal analysis.
252x91mm (200 x 200 DPI)

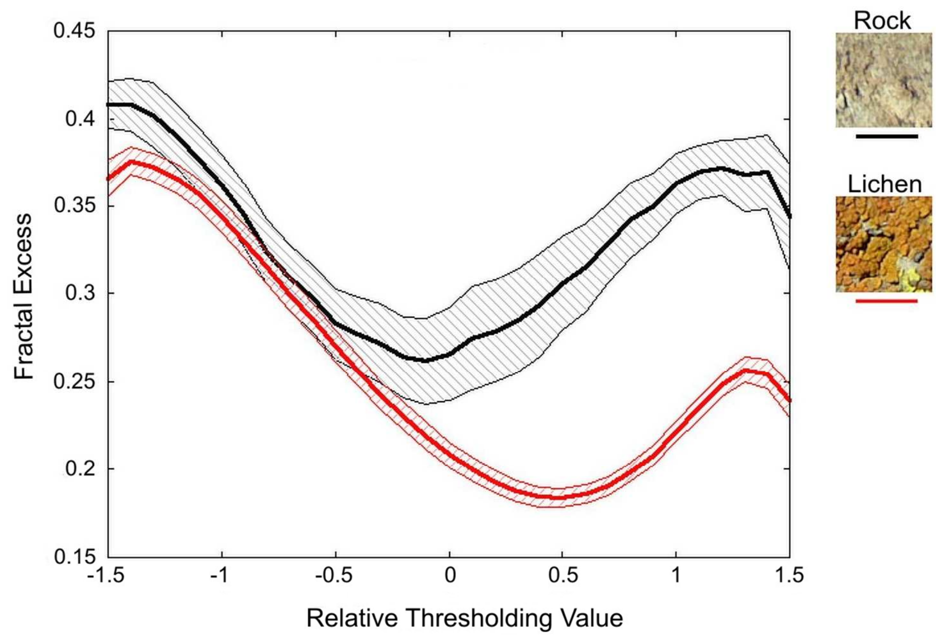

Figure 6.- Fractal Excess (FE) curves of lichen colonized and bare rock surfaces. Shaded areas for each curve correspond to the standard error of the means.
192x126mm (199 x 199 DPI)

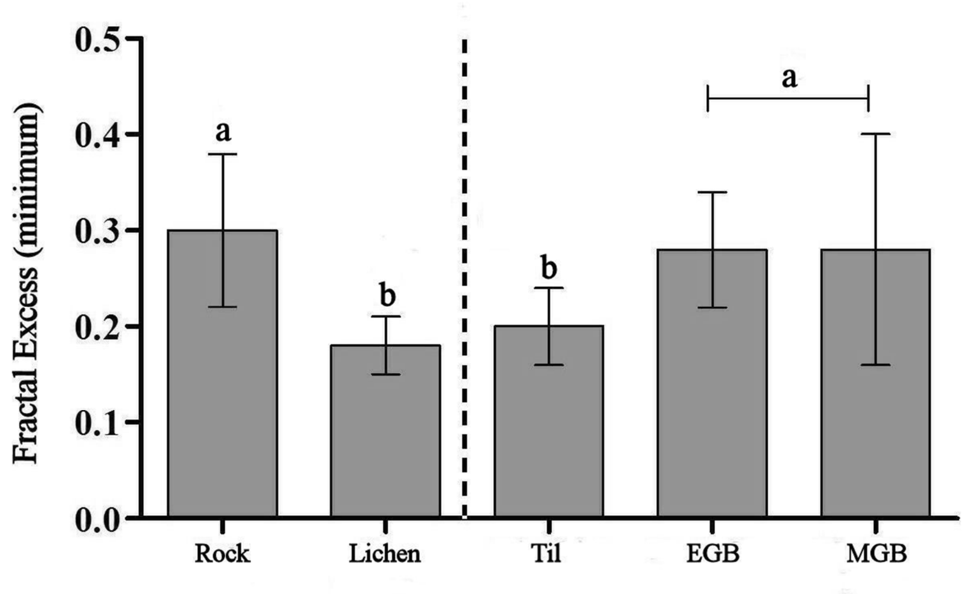

Figure 7.- Fractal excess minimum statistical analysis. Fractal excess minima of the two models shown were separately analyzed by t student test and one-way ANOVA (Tukey's Multiple Comparison post test) respectively. Means with different letters are significantly different ($P < 0.05$; a >b).
105x65mm (300 x 300 DPI)

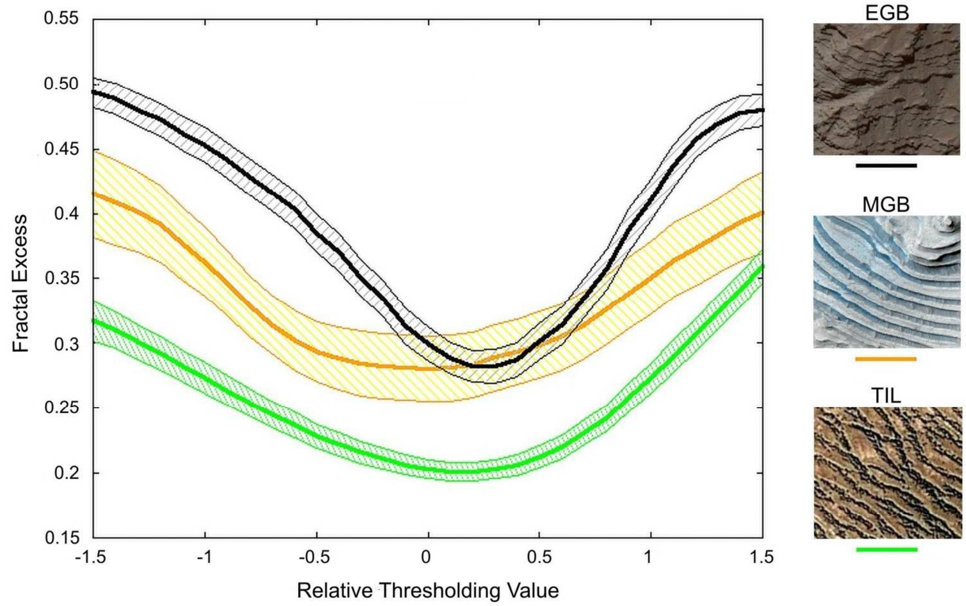

Figure 8.- Fractal Excess curves of different banding processes on Earth and on Mars. Shaded areas for each curve correspond to the standard error.
233x148mm (199 x 199 DPI)