PYROSEQUENCING REVEALS SHARED BACTERIAL TAXA ACROSS CONTINENTS

<http://dx.doi.org/10.4322/apa.2014.098>

Tauê Posada Pereira, Anthony Diego Barboza, Afnan Khalil Ahmad Suleiman,
Antônio Batista Pereira, Luiz Fernando Wurdig Roesch*

Universidade Federal do Pampa - UNIPAMPA, Campus São Gabriel,
Avenida Antônio Trilha, 1847, CEP 97300-000, São Gabriel, RS, Brazil

*e-mail: luizroesch@unipampa.edu.br

Abstract: Soil bacterial community structures differ markedly between environments but apparently there is a degree of prokaryote endemism across broad geographic scales. With this work we aimed to analyze the membership of the soil bacterial communities from soil samples collected across a continental scale and test whether contrasting environments present a subset of the soil microbial community that is shared between environments. Bacterial community composition from soils collected at the Keller Peninsula, Antarctica and from the Brazilian Pampa was assessed by pyrosequencing of the 16S rRNA gene fragments. The sequences obtained were processed to calculate the Good's coverage, to perform a Principal Coordinate Analysis (PCoA), to calculate Jaccard's and Sorensen's beta diversity dissimilarity indexes, and to perform a network-based analysis for the visualization of the OTUs that are either unique or shared by specific groups of soil samples. The overall composition of the bacterial community differed whether or not the abundance of taxa was considered. However, it was possible to detect the presence of 7.56% bacterial taxa (1,940 OTUs) shared between biomes. Our results showed that although organisms in environments at the extremes of temperature are presumably under strong selective pressures contrasting soils can harbor at least a fraction of shared bacterial taxa.

Keywords: Antarctica, Brazilian Pampa, Biodiversity, Culture-independent Technique

Introduction

Soil bacterial community structures differ markedly between systems, and even across landscapes that appear macroscopically homogeneous. Whole community studies, and studies on the genotypes of individual taxa have revealed that, at high levels of taxonomic resolution, there is a degree of prokaryote endemism apparent across broad geographic scales (Fulthorpe *et al.*, 2008). The mechanisms involved in this endemism might be related with the small size and high abundance of microorganisms as well as other aspects like probability of chance dispersal via an accidental vector, wind, water, among others (Green & Bohannan, 2006). One of the major questions involving the soil microbiome is whether there is a subset of the soil microbial community that is shared between most soils. Here we used pyrosequencing of 16S rRNA genes to compare the bacterial communities of soil samples from ice-free areas of Keller Peninsula and from the Brazilian Pampa. Our

goal was to obtain a detailed baseline description of the soil bacterial communities from soil samples collected across a continental scale in order to address the following question: Is there a common diversity structure within soil bacterial communities from contrasting environments under strong selective pressures?

Materials and Methods

Soil samples were collected from three distinct environments across two continents, including eight samples from the Brazilian Pampa and eight samples from Antarctica. From the Brazilian Pampa, we collected four soil samples from a well-preserved gallery forest (30° 24' 09.3" S; 53° 52' 59.1" W), and four samples from grassland under severe degradation (30° 24' 08.9" S; 50° 53' 05.9" W). From Antarctica, we collected eight soil samples from ice-free

areas in the northwest side of the Keller Peninsula, King George Island (62° 03' 51.1" S; 58° 24' 47.5" W). Soil was collected removing the plant cover and taking cores of 5 cm diameter and 5 cm depth. All soil samples were stored on ice upon collection and transported to the laboratory for DNA extraction.

Bacterial community composition was assessed by pyrosequencing of the 16S rRNA gene fragments (V1–V2 hypervariable regions) using 454 GS FLX Titanium (Lib-L) chemistry for unidirectional sequencing of the amplicon libraries. The DNA extraction, 16S rRNA amplification and pyrosequencing were performed following the procedures adopted by Roesch *et al.* (2012).

The raw sequences obtained were processed using QIIME (Caporaso *et al.*, 2010) with default parameters. Briefly, the multiplexed reads were first filtered for quality and assigned to the starting soil samples. The filtering criteria included a perfect match to the sequence barcode and primer, 200 bp in length, no more than two undetermined bases, and a least 60% match to a previously determined 16S rRNA gene sequence (Hamady *et al.*, 2008). The fasta-output file was used for building a table with the Operational Taxonomic Unit (OTU) abundance of each sample and the taxonomic assignments for each OTU. To do this, the sequences were clustered into OTUs based on the relatedness of the sequences (97% similarity) and a representative sequence from each OTU was selected. These representative sequences were subjected to the RDP naïve Bayesian rRNA Classifier (Wang *et al.*, 2007), which attaches complete taxonomic information from domain to species to each sequence in the database with 80% taxonomy confidence and an e-value of 0.001. The representative set of sequences was also used to align the sequences against a reference database and to build a phylogenetic tree necessary for downstream measurements.

The dataset was used to: a) calculate the Good's coverage (Good, 1953) in order to verify how well an environment was sampled and indicate the percentage of individuals sampled in a microbial community; b) perform a Principal Coordinate Analysis (PCoA) used for the overall comparison for significant differences among the bacterial communities by finding clusters of samples that reflect

the similarity of the biological communities; c) calculate Jaccard's and Sorensen's beta diversity dissimilarity indexes; d) perform a network-based analysis for the visualization of the OTUs that are either unique or shared by specific groups of soil samples. All raw sequences were submitted to the NCBI Sequence Read Archive (SRA) under the accession numbers SRA048722.2 (Antarctica) and SRA013204.1 (Brazilian Pampa).

Results

A total of 240,000 sequences were obtained from the sixteen soil samples. This number of sequences represented on average 96% coverage for soils from Antarctica and 89% coverage for the soils from the Brazilian Pampa. These values are indicative that the number of OTUs identified was sufficient to reasonably describe the communities.

The dataset was used to find clusters of similar bacterial communities by applying a Jackknifed Principal Coordinate Analysis (PCoA) (Figure 1). The weighted and unweighted PCoA analyses showed that the soil bacterial communities clustered in two distinctive groups using both weighted distance metric, which accounts for changes in the relative abundance of taxons, and unweighted distance metric, which accounts for presence/absence of taxons. Also, a pairwise dissimilarity calculation using Jaccard's and Sorensen's metrics showed 92% and 86% of dissimilarity between the bacterial communities from Antarctica and the Brazilian Pampa, respectively. Thus, the overall composition of the bacterial community differed whether or not the abundance of taxa was considered. However, when we mapped soil microbial community composition and structure and displayed microbial partitioning across soil samples in a network, it was possible to detect the presence of 7.56% bacterial taxa (1,940 OTUs) shared between both biomes (Figure 2). Across all shared OTUs, 1.9% was not able to be classified below the domain level. Of the classifiable sequences, 15 phyla were identified in both continents. They were Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, Firmicutes, Gemmatimonadetes, Nitrospira, OD1, OP10, Planctomycetes, Proteobacteria, SR1, TM7 and Verrucomicrobia.

Figure 1. Jackknifed PCoA showing clusters of similar bacterial communities with unweighted UniFrac distance metric, which accounts for presence/absence of taxa and weighted UniFrac distance metric, which accounts for changes in the relative abundance of taxa. The clusters were generated using a subset of 1000 sequences from each environment for 100 replicate trials. The positions of the points are the average for the jackknife replicates and are displayed with a network around the points representing the statistical dispersion in each axis.

Figure 2. Network based analysis showing the clusters of bacterial communities and the co-occurrences of taxonomic unities in soil samples from ice-free areas of Antarctica and Brazilian Pampa. Small circles are representative sequences from each taxonomic unity clustered based on the relatedness of the sequences (97% similarity). Large circles represent the environment where the bacterial community came from. Each environment is connected with a taxonomic unity through edges. The nodes connected by two edges represent a taxonomic unity shared between environments. The nodes connected by only one edge represent a taxonomic unity found specifically in one environment.

Discussion

The data illustrates the amount of similarity/dissimilarity found between randomly chosen soils, separated by up to 3,700 km. The main finding of this study was that less than 8% of the bacterial OTUs, as defined at the 97% similarity level, were detected with pyrosequencing in all soils. Large population sizes and short generation times are claimed to be the major mechanism resulting in high bacterial dispersal rates (Finlay & Clarke, 1999). For instance, most soils, including agricultural ones, seem to be dominated by only a small number of phyla, namely Proteobacteria, Acidobacteria, Actinobacteria, Bacteroidetes and Firmicutes (Janssen, 2006; Roesch *et al.*, 2007; Fierer *et al.*, 2009; Lauber *et al.*, 2009; Uroz *et al.*, 2010). Community similarity declines with increasing geographic distance if species tend to be locally adapted or if they are dispersal limited (Bell, 2010; Nekola & White, 1999). Although organisms in environments at the extremes of temperature are presumably under strong selective pressures, our results showed that the soils in ice-free areas of Keller Peninsula are dominated by these same phyla. This indicates that at least at a coarse level of taxonomic resolution it is possible to find some degree

of ubiquity in soil samples. Some taxonomic units have a wide distribution.

Conclusion

The strategy applied here allowed us to detect different patterns in bacterial community structure in soil samples across a continental scale. Although organisms in environments at the extremes of temperature are presumably under strong selective pressures, our results showed that contrasting soils can harbor at least a fraction of shared bacterial taxa.

Acknowledgements

This work integrates the National Institute of Science and Technology Antarctic Environmental Research (INCT-APA) that receives scientific and financial support from the National Council for Research and Development (CNPq process: n° 574018/2008-5) and Carlos Chagas Research Support Foundation of the State of Rio de Janeiro (FAPERJ n° E-16/170.023/2008). The authors also acknowledge the support of the Brazilian Ministries of Science, Technology and Innovation (MCTI), of Environment (MMA) and Inter-Ministry Commission for Sea Resources (CIRM).

References

- Bell, T. (2010). Experimental tests of the bacterial distance-decay relationship. *The ISME Journal*, 4(11): 1357-65.
- Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peña, A.G.; Goodrich, J.K.; Gordon, J.I.; Huttley, G.A.; Kelley, S.T.; Knights, D.; Koenig, J.E.; Ley, R.E.; Lozupone, C.A.; McDonald, D.; Muegge, B.D.; Pirrung, M.; Reeder, J.; Sevinsky, J.R.; Turnbaugh, P.J.; Walters, W.A.; Widmann, J.; Yatsunenko, T.; Zaneveld, J. & Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7: 335-336.
- Fierer, N.; Strickland, M.; Liptzin, D.; Bradford, M. & Cleveland, C. (2009). Global patterns in belowground communities. *Ecology Letters*, 12(11): 1238-1249
- Finlay, B.J. & Clarke, K.J. (1999). Ubiquitous dispersal of microbial species. *Nature*, 400: 828.
- Fulthorpe, R.R.; Roesch, L.F.W.; Riva, A. & Triplett, E.W. (2008). Distantly sampled soils carry few species in common. *The ISME Journal* 2: 901-10.
- Good, I.J. (1953). The population frequencies of species and the estimation of the population parameters. *Biometrika*, 40(3): 237-64.
- Green, J.L. & Bohannan, B. (2006). Spatial scaling of microbial biodiversity. *Trends in Ecology and Evolution*, 21: 501-507.
- Hamady, M.; Walker, J.J.; Harris, J.K.; Gold, N.J. & Knight, R. (2008). Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods*, 5(3): 235-7.

-
- Janssen, P. (2006). Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. *Applied and Environmental Microbiology*, 72(3): 1719-28.
- Lauber, C.; Hamady, M.; Knight, R. & Fierer, N. (2009). Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. *Applied and Environmental Microbiology*, 75(15):5111-20.
- Nekola J.C. & White P.S. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, 26(4): 867-78.
- Roesch, L.F.W.; Fulthorpe, R.R.; Riva, A.; Casella, G.; Hadwin, A.K.M.; Kent, A.D.; Daroub, S.H.; Camargo, F.A.; Farmerie, W.G. & Triplett, E.W. (2007). Pyrosequencing enumerates and contrasts soil microbial diversity. *The ISME Journal* 1(4): 283-90.
- Roesch, L.F.W.; Fulthorpe, R.R.; Pereira, A.B.; Pereira, C.K.; Lemos, L.N.; Barbosa, D.A.; Suleiman, A.K.A.; Gerber, A.L.; Pereira, M.G.; Loss, A. & Costa, E.M. (2012). Soil bacterial community abundance and diversity in ice-free areas of Keller Peninsula, Antarctica. *Applied Soil Ecology*, 61: 7-15.
- Uroz, S.; Buee, M.; Murat, C; Frey-Klett, P. & Martin, F. (2010). Pyrosequencing reveals a contrasted bacterial diversity between oak rhizosphere and surrounding soil. *Environmental Microbiology Reports*, 2(2): 281-8.
- Wang, Q.; Garrity, G.M.; Tiedje, J.M. & Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental Microbiology*, 73(16): 5261-7.