

# **Human Microbiota of the Argentine Population- A pilot study**

Belén Carbonetto<sup>1</sup>, Mónica Carolina Fabbro<sup>1</sup>, Mariela Sciara<sup>2</sup>, Analía Seravalle<sup>2</sup>,  
Guadalupe Méjico<sup>2</sup>, Santiago Revale<sup>1</sup>, María Soledad Romero<sup>1</sup>, Bianca Brun<sup>1</sup>, Marcelo  
Fay<sup>2</sup>, Fabián Fay<sup>2</sup>, Martin Vazquez<sup>1†</sup>.

5

1- Instituto de Agrobiotecnología de Rosario (INDEAR), Ocampo 210 bis (2000), Predio  
CCT Rosario, Santa Fe, Argentina.

2 CIBIC – Roca 746 (2000), Rosario, Santa Fe, Argentina

(†) Corresponding author: Martín P. Vázquez, Ph.D. Instituto de Agrobiotecnología de  
Rosario (INDEAR), Ocampo 210 bis, 2000, Predio CCT Rosario, Santa Fe, Argentina.  
Tel/Fax: +54-341-4861100. E-mail: martin.vazquez@indear.com

Keywords: human microbiome, dysbiosis, South America, healthy microbiota, amplicon  
sequencing

14

## **Background**

The human microbiota is the collection of microorganisms living in or on the human body.  
An imbalance or dysbiosis in these microbial communities can be associated with disease  
(Pham and Lawley, 2014, Petersen and Round, 2014, Zaura et al., 2014). Moreover, when  
the microbiota of the same bodysites are compared between different healthy individuals,  
specific microbial community features are apparent (Relman, 2015, Li et al., 2012, Zhou et

al., 2013). In addition differing selective pressures are found at distinct body sites leading to different patterns in microbial community structure and composition (Yatsunenko et al., 2012, Oh et al., 2014, Stearns et al., 2011). Because of the existence of these natural variations, a comprehensive characterization of the healthy microbiota is critical for predicting alterations related to disease. This characterization should be based on a broad population of healthy humans over time, geography and cultural traditions (Yatsunenko et al., 2012, Shetty et al., 2013, Ross et al., 2015, Leung et al., 2015). The study of healthy individuals representing different ages, cultural traditions and ethnic origins will enable to understand how the associated microbiota varies between populations and respond to different lifestyles. It is important to address these natural variations in order to later detect other kind of variations that are related to disease.

During the last decade researchers from around the world have characterized and defined geographical differences in the composition of microbiota in healthy adult humans. The most important projects so far have been the Human Microbiome Project (HMP) (Peterson et al., 2009), MetaHIT (Arumugam et al., 2011, Qin et al., 2010) and the American Gut Project . Despite the big effort conducted, these major projects are based on American and European populations and have not covered all ethnic groups and different socio-economic, geographic and cultural settings. Some projects have been conducted to fill this gap (Yatsunenko et al., 2012, Ross et al., 2015, Clemente et al., 2015, Leung et al., 2015). Nevertheless there are still no records on human microbiota of urban middle-income populations in South America. Here we characterized the microbiota of six different body habitats: palatine tonsils, saliva, buccal mucosa, throat, anterior nares and gut from samples of 20 healthy middle-income male and females between the ages of 20 and 50 years living

in Rosario City, Argentina. Genomic DNA was purified from each sample and the 16S rRNA gene V1-V3 region was amplified. Amplicons were sequenced on a 454 GS-FLX+ Titanium platform and produced a total of 323,110 filtered reads. We compared microbial community composition and structure data with the same data derived from the Human Microbiome Project. Results showed that microbiota of buccal mucosa, palatine tonsils and gut microbiota differed between the Argentine and American populations. We made publically available the first 16S rDNA profile dataset of human body microbiota of an Argentine urban cohort. Our results enhance the idea that the generation of more and larger local datasets of healthy individuals is needed in order to analyze different microbiota dysbiosis related to disease in the Argentine middle-income urban population.

## **Methods:**

### Sampling

Samples were collected in a non-invasive manner from palatine tonsils, saliva, buccal mucosa, throat, anterior nares and gastrointestinal tract (stool) of healthy men and women between 20 and 50 years old, who live in Rosario city in the central region of Argentina. Subjects donated blood to examine the presence of viral markers and metadata was collected by medical examination in order to select healthy individuals (Supp. Table1). Samples were collected in a single visit and sent immediately to the lab where they were processed.

### Ethics, consent and permissions

The study was approved by the Institutional Ethic's Committee from the Hospital Italiano Garibaldi in Rosario, Argentina. Moreover, a consent form with information about the study, including the rights of the subject and the risks and benefits involved in participating in the study was signed by each individual.

#### DNA extraction and sequencing

Total genomic DNA was extracted from 200mg of each fresh stool sample using QIAamp (Qiagen, CA) DNA Stool Mini Kit following manufacturer's instructions. Palatine tonsils, saliva, buccal mucosa, throat and anterior nares swabs were first resuspended in 200µl steril saline solution. Genome DNA was extracted from this solution using QIAamp (Qiagen, CA) DNA Mini Kit following manufacturer's instructions.

For the construction of pyrotag libraries the V1-V3 hyper variable regions of the 16s rRNA gene was amplified using the 27F (5'AGAGTTTGATCCTGGCTCAG3') and 534R (5'ATTACCGCGGCTGCTGG3') tagged primers (2012). Samples were amplified using two rounds of PCR, the first to amplify the 16S rRNA gene (30 cycles) and a second short round to add barcodes for sample identification (10 cycles), following the procedures detailed in Rascovan et al. 2013 (Rascovan et al., 2013). Duplicated reactions were performed in both rounds of PCR to reduce amplification biases and then pooled. All amplicons were cleaned using Ampure DNA capture beads (Agencourt- Beckman Coulter, Inc.) and pooled in equimolar concentrations before sequencing on a Genome Sequencer FLX (454-Roche Applied Sciences) using 454 GS FLX+ chemistry according to the manufacturer's instructions.

#### Amplicon sequence processing, OTU classification and taxonomic assignment

Our dataset and the HMP dataset (<http://hmpdacc.org/HMQCP/>) were processed using the QIIME v1.8 analysis pipeline (Caporaso et al., 2010b). For comparative purposes the same number of individuals (N=20) was selected randomly from the HMP dataset. First a random selection of 1000 reads per sample was done using a custom-made script. Sequences were then clustered into Operational Taxonomic Units (OTUs) using the pick\_otus.py script with the Uclust method at 97% sequence similarity (Edgar, 2010). OTU representative sequences were aligned using PyNast algorithm with QIIME default parameters (Caporaso et al., 2010a). Phylogenetic trees containing the aligned sequences were then produced using FastTree (Price et al., 2009). Richness alpha diversity metrics and rarefaction curves were calculated by sub-sampling the OTU tables at different depths and counting the resulting number of phylotypes using 10 iterations per sample. Phylogeny-based beta diversity distances between OTUs were calculated using weighted Unifrac (Lozupone and Knight, 2005, Lozupone et al., 2007). Taxonomic classification of sequences was done with Ribosomal Database Project (RDP) Classifier using the Greengenes V13.5 database and a 50% confidence threshold (Wang et al., 2007).

## Numerical analyses

Unifrac phylogenetic pairwise distances among samples were visualized with principal coordinates analysis. Analysis of similarity statistics was calculated to test a-priori sampling groups. Mann-Whitney non-parametric tests were performed to elucidate differences in taxa abundances. All calculations were carried out with R packages 'BiodiversityR' and 'Vegan'.

## Results

### Argentine human microbiota overview

Results showed that OTU richness differed between body habitats. Saliva presented the higher richness and the nose was the less diverse site (Figure 1A). Beta diversity also revealed differences in community structure between habitats (Figure 1B). The PCoA visualization revealed a separation of the data in three main groups: oropharyngeal region, nose and gut (ANOSIM  $p < 0.05$ ).

### Comparison with the HMP dataset

We observed that the microbiota of the Argentine and American populations differed in composition and community structure. Beta diversity results based on Weighted Unifrac distances showed that the microbiota of the buccal mucosa, the palatine tonsils and the gastrointestinal tract differed between Argentine and American individuals (Supp. Figure 1, ANOSIM  $p < 0.05$ ). Moreover, the taxonomic composition of these body habitats was also different between populations. We observed differences in the abundance of the most predominant taxa (Figure 2). For example the abundance of Bacteroidaceae family was higher in the American population gut microbiota, while Ruminococcaceae, Lachnospiraceae, Rikenellaceae and Prevotellaceae were more abundant in the Argentine gut microbiota. It is known that the variation in the levels of three main taxa in the gut microbiota can define enterotypes (Arumugam et al., 2011). The higher relative abundance of three genera defines the enterotypes: Bacteroides defines enterotype 1, Prevotella defines enterotype 2 and of Ruminococcus defines enterotype 3. Although we found differences in Bacteroidaceae and Ruminococcaceae abundances (Figure 2, Mann-Whitney  $p < 0.05$ ) and

131 Bacteroides, Ruminococcus and Prevotella abundances (Mann-Withney  $p < 0.05$ , data not  
 132 shown) between populations, both populations can be assigned to enterotype 1 (Supp.  
 133 Figure 2). This is, to our knowledge the first time gut microbiota of middle-income  
 134 populations of North and South America are compared. Other author's did find  
 135 differences in Prevotella and Bacteroides abundances between southamerican and US  
 136 populations, but they compared ameridian individuals living in little villages near the  
 137 Amazonas with western citizens living in metropolitan areas (Yatsunenko et al., 2012).  
 138 Southamerican ameridians presented higher Prevotella abundance than US individuals. As  
 139 expected the authors hypothesized that differences in diet were mainly explaining  
 140 differences in the gut microbiota. Our results encourage this hypothesis. We observed  
 141 differences in the relative abundances of the most important taxa but the  
 142 Bacteroides/Prevotella trade-off (enterotype 1) was the same for the argentine and US  
 143 populations. The diet of people in the main metropolitan areas of Argentina is more similar  
 144 to the one of individuals in the main cities of the US than it is from individuals in the  
 145 mentioned villages near the Amazonas, which is mainly based on maize and native roots  
 146 rather than processed food. Our initial findings support this idea that diet is more relevant  
 147 when defining gut enterotypes than it is geographic location or ethnics. Nevertheless, the  
 148 difference in taxa abundances between our dataset and the HMP dataset reflects the  
 149 differences in cultural traditions in the Argentine and US cities. As regards the buccal  
 150 mucosa microbiota we observed that Streptococcaceae family was more abundant in the  
 151 Argentine population while Gemellaceae family was higher in the US individuals (Figure 2,  
 152 Mann-Withney  $p < 0.05$ ). The abundance of Streptococcaceae family was also higher in the  
 153 palatine tonsils of argentine individuals. We hypothesize that the differences in the  
 154 observed oral/pharyngeal profiles may be due to differences in the use of antibiotic

treatments in both populations. Long-term impacts in oral microbiota due to antibiotic administration has been reported, moreover antibiotic resistance were observed in these cases (Jakobsson et al., 2010). Although this is just a hypothesis that needs to be proven our results encourage the idea that the human microbiota ecosystem has multiple states of equilibrium and that these variations are present between and within healthy populations. Then, these multiple states related to different lifestyles, location, ethnics, cultural tradition, age, gender, etc. need to be studied in depth in order to further understand alterations on the equilibrium due to disease.

## **Conclusions**

Here we present the first dataset based on human microbiota samples of an urban middle-income population in South America. We characterized the microbiota of six different body habitats: palatine tonsils, saliva, buccal mucosa, throat, anterior nares and gut from samples of healthy individual living in a metropolitan area in Argentina. Our initial findings revealed differences in the structure and composition of the microbial communities compared to the US urban population.

By sharing our data, we want to actively encourage its reuse for comparison purposes. This will hopefully result in novel biological insights on the variability of the microbiota of healthy individuals across populations worldwide. Moreover, the understanding of the human microbiota ecosystem in a health-associated state will help to answer questions related to the role of the microbiota in disease.

## **Data Access**



Datasets supporting the results of this study are available in the NCBI SRA database under the BioProject with accession number: PRJNA293521 (<http://www.ncbi.nlm.nih.gov/bioproject/PRJNA293521/>).

# **Conflict or interest statement**

The authors declare that they have no conflict of interest.

# **Authors' contributions**

BC and MCF analyzed the data and performed the numerical and statistical analyses. BC wrote the manuscript. SR processed the sequence raw data. MF and MS designed the protocol and criteria for healthy individual selection. MS wrote the protocol for recruiting the individuals, supervised the recruiting process, and collected the metadata for the individuals. GM performed DNA preparation from human samples. MSR and BB performed amplicon library preparation and sequencing. FF directed the project and supervised the study. MV directed the project, conceived the study, supervised the bioinformatics analysis and supervised the manuscript.

# **Funding**

Funding for this work was provided by Agencia Nacional de Promoción Científica y Tecnológica in Argentina.

# **References**

*American Gut* [Online]. Available: <http://americangut.org/>.  
2012. A framework for human microbiome research. *Nature*, 486, 215-221.  
ARUMUGAM, M., RAES, J., PELLETIER, E., LE PASLIER, D., YAMADA, T., MENDE, D. R.,  
FERNANDES, G. R., TAP, J., BRULS, T., BATTO, J.-M., BERTALAN, M., BORRUEL,

N., CASELLAS, F., FERNANDEZ, L., GAUTIER, L., HANSEN, T., HATTORI, M.,  
HAYASHI, T., KLEEREBEZEM, M., KUROKAWA, K., LECLERC, M., LEVENEZ, F.,  
MANICHANH, C., NIELSEN, H. B., NIELSEN, T., PONS, N., POULAIN, J., QIN, J.,  
SICHERITZ-PONTEN, T., TIMS, S., TORRENTS, D., UGARTE, E., ZOETENDAL, E.  
G., WANG, J., GUARNER, F., PEDERSEN, O., DE VOS, W. M., BRUNAK, S., DORE, J.,  
WEISSENBAACH, J., EHRLICH, S. D. & BORK, P. 2011. Enterotypes of the human gut  
microbiome. *Nature*, 473, 174-180.

CAPORASO, J. G., BITTINGER, K., BUSHMAN, F. D., DESANTIS, T. Z., ANDERSEN, G. L. &  
KNIGHT, R. 2010a. PyNAST: a flexible tool for aligning sequences to a template  
alignment. *Bioinformatics*, 26, 266-267.

CAPORASO, J. G., KUCZYNSKI, J., STOMBAUGH, J., BITTINGER, K., BUSHMAN, F. D.,  
COSTELLO, E. K., FIERER, N., PENA, 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. 2010b. QIIME allows analysis of  
high-throughput community sequencing data. *Nat Meth*, 7, 335-336.

CLEMENTE, J. C., PEHRSSON, E. C., BLASER, M. J., SANDHU, K., GAO, Z., WANG, B.,  
MAGRIS, M., HIDALGO, G., CONTRERAS, M., NOYA-ALARCÓN, Ó., LANDER, O.,  
MCDONALD, J., COX, M., WALTER, J., OH, P. L., RUIZ, J. F., RODRIGUEZ, S., SHEN,  
N., SONG, S. J., METCALF, J., KNIGHT, R., DANTAS, G. & DOMINGUEZ-BELLO, M.  
G. 2015. The microbiome of uncontacted Amerindians. *Science Advances*, 1.

EDGAR, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*,  
26, 2460-2461.

JAKOBSSON, H. E., JERNBERG, C., ANDERSSON, A. F., SJÖLUND-KARLSSON, M.,  
JANSSON, J. K. & ENGSTRAND, L. 2010. Short-Term Antibiotic Treatment Has  
Differing Long-Term Impacts on the Human Throat and Gut Microbiome. *PLoS ONE*, 5,  
e9836.

LEUNG, M. H. Y., WILKINS, D. & LEE, P. K. H. 2015. Insights into the pan-microbiome: skin  
microbial communities of Chinese individuals differ from other racial groups. *Scientific  
Reports*, 5, 11845.

LI, K., BIHAN, M., YOOSEPH, S. & METHÉ, B. A. 2012. Analyses of the Microbial Diversity  
across the Human Microbiome. *PLoS ONE*, 7, e32118.

LOZUPONE, C. & KNIGHT, R. 2005. UniFrac: a New Phylogenetic Method for Comparing  
Microbial Communities. *Applied and Environmental Microbiology*, 71, 8228-8235.

LOZUPONE, C. A., HAMADY, M., KELLEY, S. T. & KNIGHT, R. 2007. Quantitative and  
Qualitative  $\beta$  Diversity Measures Lead to Different Insights into Factors That Structure  
Microbial Communities. *Applied and Environmental Microbiology*, 73, 1576-1585.

OH, J., BYRD, A. L., DEMING, C., CONLAN, S., PROGRAM, N. C. S., KONG, H. H. & SEGRE,  
J. A. 2014. Biogeography and individuality shape function in the human skin metagenome.  
*Nature*, 514, 59-64.

PETERSEN, C. & ROUND, J. L. 2014. Defining dysbiosis and its influence on host immunity and  
disease. *Cell Microbiol*, 16, 1024-33.

PETERSON, J., GARGES, S., GIOVANNI, M., MCINNES, P., WANG, L., SCHLOSS, J. A.,  
BONAZZI, V., MCEWEN, J. E., WETTERSTRAND, K. A., DEAL, C., BAKER, C. C., DI  
FRANCESCO, V., HOWCROFT, T. K., KARP, R. W., LUNSFORD, R. D.,  
WELLINGTON, C. R., BELACHEW, T., WRIGHT, M., GIBLIN, C., DAVID, H., MILLS,  
M., SALOMON, R., MULLINS, C., AKOLKAR, B., BEGG, L., DAVIS, C.,  
GRANDISON, L., HUMBLE, M., KHALSA, J., LITTLE, A. R., PEAVY, H., PONTZER,  
C., PORTNOY, M., SAYRE, M. H., STARKE-REED, P., ZAKHARI, S., READ, J.,

WATSON, B. & GUYER, M. 2009. The NIH Human Microbiome Project. *Genome Res*, 19, 2317-23.

PHAM, T. A. N. & LAWLEY, T. D. 2014. Emerging insights on intestinal dysbiosis during bacterial infections. *Current Opinion in Microbiology*, 17, 67-74.

PRICE, M. N., DEHAL, P. S. & ARKIN, A. P. 2009. FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. *Molecular Biology and Evolution*, 26, 1641-1650.

QIN, J., LI, R., RAES, J., ARUMUGAM, M., BURGDORF, K. S., MANICHANH, C., NIELSEN, T., PONS, N., LEVENEZ, F., YAMADA, T., MENDE, D. R., LI, J., XU, J., LI, S., LI, D., CAO, J., WANG, B., LIANG, H., ZHENG, H., XIE, Y., TAP, J., LEPAGE, P., BERTALAN, M., BATTO, J.-M., HANSEN, T., LE PASLIER, D., LINNEBERG, A., NIELSEN, H. B., PELLETIER, E., RENAULT, P., SICHERITZ-PONTEN, T., TURNER, K., ZHU, H., YU, C., LI, S., JIAN, M., ZHOU, Y., LI, Y., ZHANG, X., LI, S., QIN, N., YANG, H., WANG, J., BRUNAK, S., DORE, J., GUARNER, F., KRISTIANSEN, K., PEDERSEN, O., PARKHILL, J., WEISSENBAACH, J., BORK, P., EHRLICH, S. D. & WANG, J. 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464, 59-65.

RASCOVAN, N., CARBONETTO, B., REVALE, S., REINERT, M., ALVAREZ, R., GODEAS, A., COLOMBO, R., AGUILAR, M., NOVAS, M., IANNONE, L., ZELADA, A., PARDO, A., SCHRAUF, G., MENTABERRY, A. & VAZQUEZ, M. 2013. The PAMPA datasets: a metagenomic survey of microbial communities in Argentinean pampean soils. *Microbiome*, 1, 21.

RELMAN, D. A. 2015. The human microbiome and the future practice of medicine. *JAMA*, 314, 1127-1128.

ROSS, M., MUZNY, D., MCCORMICK, J., GIBBS, R., FISHER-HOCH, S. & PETROSINO, J. 2015. 16S gut community of the Cameron County Hispanic Cohort. *Microbiome*, 3, 1-11.

SHETTY, S., MARATHE, N. & SHOUCHE, Y. 2013. Opportunities and challenges for gut microbiome studies in the Indian population. *Microbiome*, 1, 24.

STEARNS, J. C., LYNCH, M. D. J., SENADHEERA, D. B., TENENBAUM, H. C., GOLDBERG, M. B., CVITKOVITCH, D. G., CROITORU, K., MORENO-HAGELSIEB, G. & NEUFELD, J. D. 2011. Bacterial biogeography of the human digestive tract. *Scientific Reports*, 1, 170.

WANG, Q., GARRITY, G. M., TIEDJE, J. M. & COLE, J. R. 2007. Naïve Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, 73, 5261-5267.

YATSUNENKO, T., REY, F. E., MANARY, M. J., TREHAN, I., DOMINGUEZ-BELLO, M. G., CONTRERAS, M., MAGRIS, M., HIDALGO, G., BALDASSANO, R. N., ANOKHIN, A. P., HEATH, A. C., WARNER, B., REEDER, J., KUCZYNSKI, J., CAPORASO, J. G., LOZUPONE, C. A., LAUBER, C., CLEMENTE, J. C., KNIGHTS, D., KNIGHT, R. & GORDON, J. I. 2012. Human gut microbiome viewed across age and geography. *Nature*, 486, 222-227.

ZAURA, E., NICU, E. A., KROM, B. P. & KEIJSER, B. J. F. 2014. Acquiring and maintaining a normal oral microbiome: current perspective. *Frontiers in Cellular and Infection Microbiology*, 4, 85.

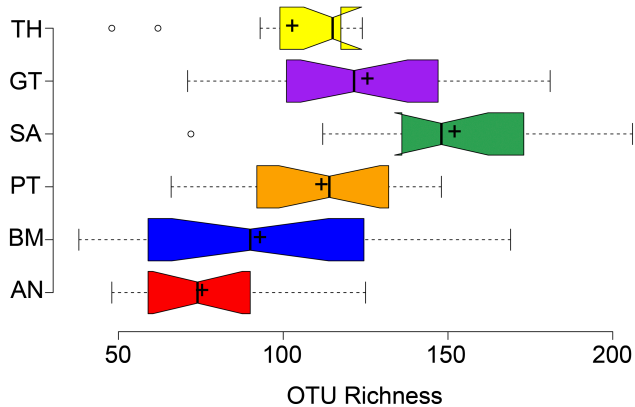
ZHOU, Y., GAO, H., MIHINDUKULASURIYA, K., LA ROSA, P., WYLIE, K., VISHNIVETSKAYA, T., PODAR, M., WARNER, B., TARR, P., NELSON, D., FORTENBERRY, J., HOLLAND, M., BURR, S., SHANNON, W., SODERGREN, E. & WEINSTOCK, G. 2013. Biogeography of the ecosystems of the healthy human body. *Genome Biology*, 14, R1.

## Figure Legends

**Figure 1: Argentine human microbiota diversity analysis.** (A) Alpha diversity analysis based on OTU richness BM stands for buccal mucosa, TH for throat, PT for palatine tonsils, AN for anterior nares, SA for saliva and GT for gut. (B).Beta diversity analysis based on Weighted unifracs pairwise distances, each color represents a body site as in A.

**Figure 2: Taxonomic profiles.** Most abundant bacterial families in the Argentine and American (HMP) human microbiota. ARG refers to the Argentina population and US to the American population. BM stands for buccal mucosa, TH for throat, PT for palatine tonsils, AN for anterior nares, SA for saliva and GT for gut.

A



B

