# Hunting for rare and undersampled phyla in Trunk-river lemonade

Moritz Buck
Uppsala University

Summer 2015

## 1 Introduction

The ARB database lists more than 50 phyla of Bacteria. However the four most represented phyla (Proteobacteria, Firmicutes, Bacteroidetes, and Actinobacteria) collect 80% of the total 16s sequences of this database, whereas the 80% rarest phyla only represent about 5% of all sequences. Many of these phyla represented in ribosomal gene databases do not have a single published genome, or isolate. This genetic space is chronically under-represented and under-studied and might contain a large variety of unknown and important biology [1].

Additionally the question is still open if and which of the many phyla are actually rare in nature. Some of these are few in the database simply because they are rare in nature. However, some of these can be highly distributed, and central to certain ecosystem or have not been observed in 16s censuses due to inherent biases in the methods. Famously any bacteria smaller than $0.2\mu m$ will not be observed in most studies as $0.2\mu m$ filters are often used to filter out cells for DNA extraction. Also primers used for 16s amplification can obscure major parts of a microbial ecology if the target 16s region contains miss-matches at unfortunate positions.

Not only are most of these organisms underrepresented in 16s databases, they are also very few, if any, isolates available for many of them, putting into forward the difficulty of cultivation of these phyla. Slow-growth, inability to attach, sensitivity to environmental parameter, are some but certainly not all of the reasons of this. But maybe also more human factors, like "uncoolness"

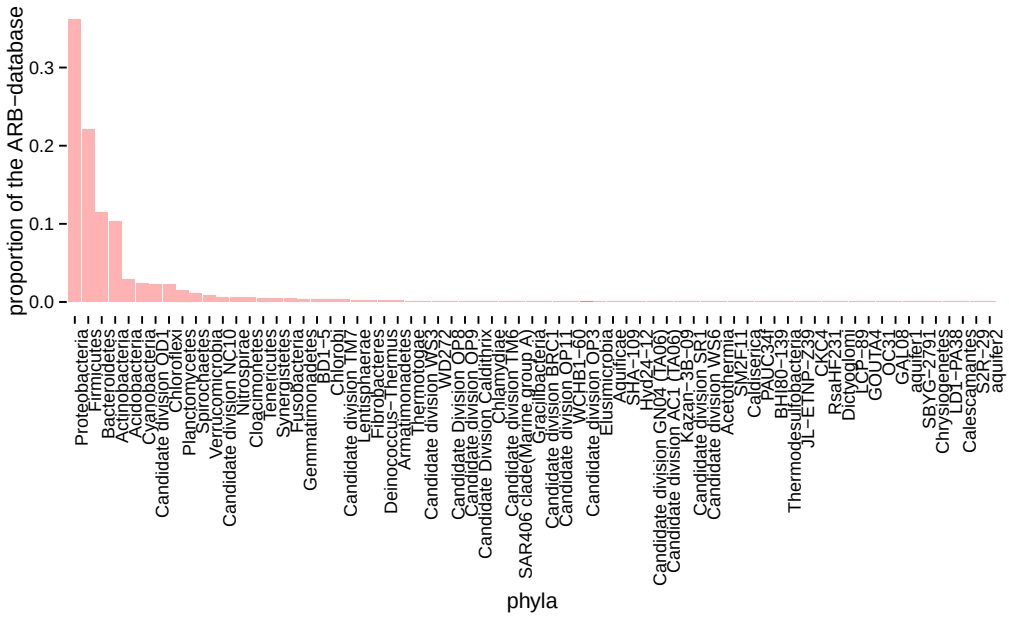
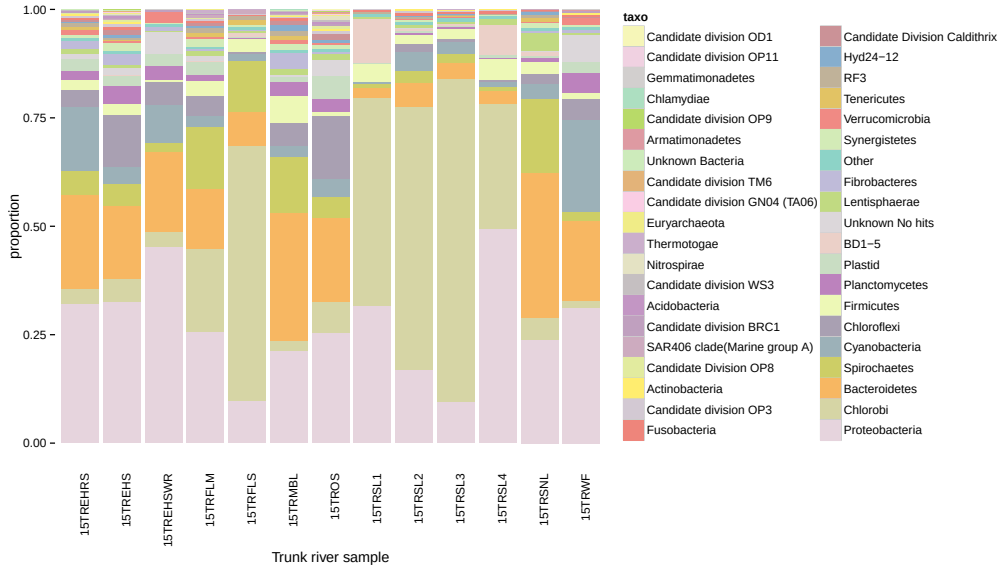



Figure 1: Proportion of sequences per phyla in the ARB database, total number of sequences $4.5 * 10^6$

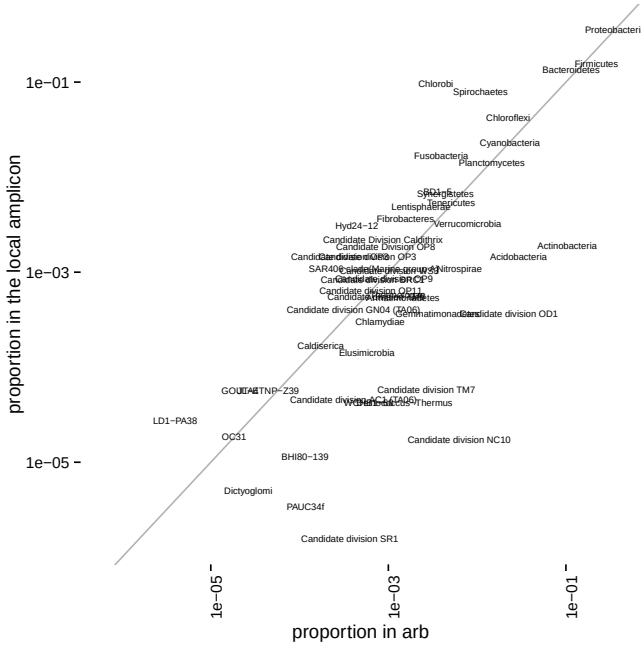
of cell colonies, or visibility of these, can influence the lack of isolates from rare taxa.

A surprising aspect of many of these rare phyla is that they can be found in many environments but only at low levels (corresponding closely to the levels in the databases, Fig. 2b), and also occasionally with higher abundance (see Chlorobi in Figure 2). So it could be concluded that many of the rare phyla are at least not rare due to the rarity of their niches.

In the course of this project I will attempt to detect and potentially isolate members of rare taxa, using a combination of cultivation and molecular techniques. Inocula from the Trunk-river site (Falmouth, MA) will be grown on plates and in natural liquid media. DNA will be extracted from these as well as from filters of water from "Trunk-river lemonade". This is then screened using phyla specific PCR-primers for a variety of taxa.




(a) Frequency of phylum in a 16s amplicon of diverse trunk-river samples




(b) Proportion of reads in the ARB database versus the proportion of reads in the Trunk-river amplicon dataset

Figure 2: Phyla in Trunk-river

# 2 Materials and Methods

## 2.1 SwampLTY

This Trunk-river inspired variant of the LTY culture media is composed, for 1 L:

- 10 mL 100x Salt Water Base
- 100 mg $NaSo_4$
- 1 g Trytone
- 1 mL trace element solution
- 0.1 g yeast extracted
- if making plates, add an additional 15g agar

Autoclave the initial medium and, after cooling down, additionally add:

- 50mg cyclohexamide
- 1mL multivitamin solution

## 2.2 Spiked swampL

A variant of the previous medium, replacing the typtone and yeast by dried, pulverized, autoclaved rotting sea-grass collected *in situ*, and complementing with a little additional Potassium and Nitrogen, for 1 L:

- 10 mL 100x Salt Water Base
- 100 mg $NaSo_4$
- $100\mu L$ $HPO_4$
- 1 mL $NH_4Cl$
- 1 g dried autoclaved rotting sea-grass
- if making plates, add an additional 15g agar

Autoclave the initial medium and, after cooling down, additionally add:

- 50 mg cyclohexamide
- 1 mL multivitamin solution

## 2.3 Filtered swamp media

The filtered swamp media is nothing else than water from Trunk-river filtered through a $0.2\mu m$ filter, with 1 g of dried autoclaved rotted sea-grass. For the swamp-plates dissolve 10 g of sterile agar in 200 mL of filtered water for 1 L of media in the microwave and mix the dissolved media into the media when appropriately cooled down, and pour plates.

## 2.4 DNA-extractions

The power-fecal (Probio) extraction kit is used to extract DNA from the $8\mu m$, $5\mu m$, $3\mu m$, $2\mu m$, $0.45\mu m$, $0.2\mu m$, and $0.1\mu m$ used for the filtration of the swamp water used for the medium.

Plated colonies are washed by a protocol inspired by [9]. 3 mL of SwLTY medium is are flushed repeatedly over the plates, which are then placed on a slow shaker for 30min to detach cells. Filled eppendorf tubes with supernatent, and pelleted the cells at 16000rpm for 15 min. 5mL of liquid media (or more dependent on the growth in the bottle) are pelleted for liquid culture at 16000rpm for 15min.

DNA is extracted from the pelleted liquid and plated colonies re-suspended in $50\mu L$ nuclease-free water. These were boiled in 95°C for 15min on a thermocycler.

## 2.5 PCRs

The Promega goTaq hotstart kit was used for PCR reactions. For each reaction $25\mu L$ of goTaq hotstart engine master mix was added to $3\mu L$ of extracted DNA, $20\mu L$ of nuclease-free water, and $1\mu L$ of forward- and reverse-primer (at xxxng/mL). They were run on the thermocycler for 30 cycles at various annealing temperatures (see Table 1). PCR products are verified on a 1% agarose gel stained with $1.5\mu L.L^{-1}$

| Phyla | FWD-primer | REV-primer | rpos | fpos | temp | citation |
|---|---|---|---|---|---|---|
| Chlorobi | GCTCAGGACGAACGYTG | TTAGCCCCAGTCACTAAG | 38 | 1476 | 51 | [2] |
| GN02 | CTGGCTCAGGGTGAAC | GTCAGTAGTCCTACCATG | 34 | 1467 | 51 | [2] |
| OD1[2] | universal forward | GACTACHVGGGTATCTAATCC | [NA] | 805 | 51 | [4] |
| Spirochetes | CACATTGGGACTGAGATAC | TACCTGTTAGTAACYGGCAGTAG | 312 | 1138 | 51 | [3] |
| SR1 | GATGAACGCTAGCGRAAYG | CTTAACCCCAGTCACTGATT | 46 | 1475 | 51 | [2] |
| Synergestes | TTTGATCCTGGCTCAGGA | GTTACGACTTCACCCYYCT | 39 | 1501 | 51 | [2] |
| Verrucomicrobia | TGGCGGCGTGGWTAAGA | TGCTACACCGWGAATTC | 37 | 673 | 51 | Scott |
| Bacteroidetes[1] | GGARCATGTGGTTTAATTCGATGAT | AGCTGACGACAACCATGCAG | 934 | 1060 | 54 | [5] |
| Acidobacteria | GCTCAGAATSAACGCTGG | universal reverse | 39 | [NA] | 54 | [6] |
| Chloroflexi | ATCAACCCCACCTTTGAC | GGCGGCGYGGCYTAATAC | 54 | 1464 | 54 | [2] |
| OP10 | universal forward | ASTACGGCCGCAAGGTTG | [NA] | 907 | 54 | [7] |
| Planctomycetes | GGCTGCAGTCGAGRATCT | TGTGTGAGCCCCCGTCAA | 325 | 920 | 54 | [8] |
| TM7 | GWAAAGAGTWGCGTAGGYGG | WTRCTTAACGCGTTAGCTTCGCT | 590 | 965 | 54 | NA |

Table 1: Details of PCR-primers used in this study.

# 3 Results

1 L of water from swamp-lemonade holes from Trunk-river (see future amazing paper) has been filtered trough serial $8\mu m$, $5\mu m$, $3\mu m$, $2\mu m$, $0.45\mu m$, $0.2\mu m$, and $0.1\mu m$. DNA extracted from these (12 samples, the $0.45\mu m$ and $0.1\mu m$ as triplicates), and media made with the water. Plates 1-5;22-24 are on SwLTY, 6-10;26 spiked swampL, 11-15;17-18;20-21 natural swamp plates. Each inoculated with $100\mu L$ of swamp lemonade. These are grown for a week in a 1% oxygen atmosphere for bottles with an ID smaller than 22, and at 3% for plates numbered 22 and larger. Liquid cultures 1-3 are SwLTY, 4-6 spiked swampL, 7-9 filtered swamp media, and 10 an unfiltered bottle of trunk lemonade. Bottles 1-2;4-5;7-8 are Inoculated with $100\mu L$ of swamp lemonade. Bottles 3,6,9 not inoculated (the water is filtered at $0.2\mu L$ so smaller bacteria are "the inoculum"). Bottle 10 is also not inoculated and just continues growing around. All bottles where at 1% oxygen. Bottles are left for their own devices for 2 weeks on the bench before DNA-extraction.

PCRs are done for all primers on all samples (546 total PCRs). PCR products are verified by gel and summarized in Table 2.

# 4 Discussion

As the table obviously shows, the number of hits is limited, and mostly to the positive controls. But all was not lost. The two "positive controls", OD1 and Bacto, have worked on a decent numbers of plates, proving that the DNA from plate washes and low-density (the bottles where low density mostly) liquid cultures was enough. Using the environmental sample as "DNA positive control", seams to work also as only phyla detected on filters have also been detected on cultures, but not nearly all. However the method needs a lot of optimization of primers, for examples the Chlorobi primer should have detected something as the original environment contain large numbers of them (see Figure 2a). This could be due to a variety of reasons, from badly designed primer, to the extraction protocol and annealing temperature optimization. For a larger study all these parameters as well as DNA-quantity should be optimized. Also adequate positive controls should be found for

[1] used as a positive control

[2] actually not OD1 specific, is a general primer that is not bad a detecting known OD1s, so basically used as a second positive control

| sample | OD1 | Verr | Syn | Bacto | TM7 | Plancto | media |
|---|---|---|---|---|---|---|---|
| $0.1\mu m$ | x | | x | x | | | filter |
| $0.1\mu m$ | x | | x | x | x | | filter |
| $0.1\mu m$ | x | | x | x | x | | filter |
| $0.2\mu m$ | x | | | x | x | x | filter |
| $0.45\mu m$ | x | x | x | x | x | x | filter |
| $0.45\mu m$ | x | x | x | x | x | x | filter |
| $0.45\mu m$ | x | x | x | x | x | x | filter |
| $2\mu m$ | x | | | x | | | filter |
| $2\mu m$ | x | | x | x | | x | filter |
| $3\mu m$ | x | | | x | | | filter |
| $5\mu m$ | | | | x | | | filter |
| $8\mu m$ | x | x | | x | x | x | filter |
| 1 | | | | x | | | plate |
| 2 | | | | x | | | plate |
| 3 | | | | x | | | plate |
| 4 | | | | x | | | plate |
| 5 | | | | x | | | plate |
| 6 | | | | x | | | plate |
| 7 | | | | x | | | plate |
| 8 | x | | | | | | plate |
| 9 | | | | x | | | plate |
| 10 | x | | | x | | | plate |
| 11 | | | | x | | | plate |
| B1 | x | | | | | | liquid |
| B2 | x | | | | | | liquid |
| B7 | x | | | | | | liquid |

Table 2: Successful PCRs. Primers and samples with no bands at all have been removed.

each primer, for obvious reasons. Also all primers should be tested against a variety of "other" DNA to check for non-specific binding.

An other line of improvement of this system is the time-scales. Many of the rarer taxa are suspected to be slower growing. Hence letting the cultures grow for longer time and finding ways of limiting invasion by copiotrophic generalists would be central. Also the downstream optimization once a culture containing a clade of interest is found, has to be addressed.

# References


[1] Christian Rinke, Patrick Schwientek, Alexander Sczyrba, Natalia N. Ivanova, Iain J. Anderson, Jan-Fang Cheng, Aaron Darling, Stephanie Malfatti, Brandon K. Swan, Esther A. Gies, Jeremy A. Dodsworth, Brian P. Hedlund, George Tsiamis, Stefan M. Sievert, Wen-Tso Liu10, Jonathan A. Eisen, Steven J. Hallam, Nikos C. Kyrpides, Ramunas Stepanauskas, Edward M. Rubin, Philip Hugenholtz, and Tanja Woyke, *Insights into the phylogeny and coding potential of microbial dark matter*, Nature, 2013.

[2] Anuj Camanocha and Floyd E. Dewhirst, *Host-associated bacterial taxa from Chlorobi, Chloroflexi, GN02, Synergistetes, SR1, TM7, and WPS-2 Phyla/candidate divisions*, Journal of Oral Microbiology, 2014.

[3] Mitsuo Sakamoto, Jose F. Siqueira, Jr., Isabela N. Rocas, and Yoshimi Benno, *Diversity of Spirochetes in Endodontic Infections*, Journal of Clinical Microbiology, 2009

[4] Herlemann DPR, Labrenz M, Jurgens K, Bertilsson S, Waniek JJ, Andersson AF. *Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea*, ISME Journal, 2011

[5] X. Guo, X. Xia , R. Tang , J. Zhou , H. Zhao and K. Wang, *Development of a real-time PCR method for Firmicutes and Bacteroidetes in faeces and its application to quantify intestinal population of obese and lean pigs*, Letters in Applied Microbiology, 2008

[6] Sang-Hoon Lee, Jae-Chang Cho, *Group-specific PCR primers for the phylum Acidobacteria designed based on the comparative analysis of 16S rRNA gene sequences*, Journal of Microbiological Methods, 2011

[7] Sergio E. Morales and William E. Holben, *Empirical Testing of 16S rRNA Gene PCR Primer Pairs Reveals Variance in Target Specificity and Efficacy Not Suggested by In Silico Analysis*, Applied and Environmental Microbiology, 2009

[8] Martin Mu hling, John Woolven-Allen, J Colin Murrell and Ian Joint, *Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities*, ISME Journal, 2008

[9] Bradley S. Stevenson, Stephanie A. Eichorst, John T. Wertz, Thomas M. Schmidt and John A. Breznak, *New Strategies for Cultivation and Detection of Previously Uncultured Microbes*, Applied and Environmental Microbiology, 2004

[10] Jantiya Isanapong, W Sealy Hambright, Austin G Willis, Atcha Boonmee, Stephen J Callister, Kristin E Burnum, Ljiljana Pasa-Tolic, Carrie D Nicora, John T Wertz, Thomas M Schmidt and Jorge LM Rodrigues, *Development of an ecophysiological model for Diplosphaera colotermitum TAV2, a termite hindgut Verrucomicrobium*, ISME Journal, 2013.