Clinical metagenomics of bone and joint infections: a proof of concept study

Etienne Ruppé[1], Vladimir Lazarevic[1], Myriam Girard[1], William Mouton[2], Tristan Ferry[3], Frédéric Laurent[2], Jacques Schrenzel[1,4]

1. Genomic Research Laboratory, Service of Infectious Diseases, Geneva University Hospitals, rue Gabrielle-Perret-Gentil 4, 1205 Geneva, Switzerland.

2. Centre International de Recherche en Infectiologie, INSERM U1111, Pathogenesis of staphylococcal infections, University of Lyon 1, Lyon, France; Department of Clinical Microbiology, Northern Hospital Group, Hospices Civils de Lyon, Lyon, France.

3. Centre International de Recherche en Infectiologie, INSERM U1111, Pathogenesis of staphylococcal infections, University of Lyon 1, Lyon, France; Infectious Diseases Department, Northern Hospital Group, Hospices Civils de Lyon, Lyon, France.

4. Bacteriology Laboratory, Service of Laboratory Medicine, Department of Genetics and Laboratory Medicine, Geneva University Hospitals, 4 rue Gabrielle-Perret-Gentil, 1205 Geneva, Switzerland.

Key words: next-generation sequencing, metagenomics, bone and joint infections, bio-informatics, antibiotic resistance genes.

Running title: Metagenomics in bone and joint infections

***Corresponding author**

Dr Etienne RUPPE (PharmD, PhD)

Genomic Research Laboratory

Service of Infectious Diseases

Geneva University Hospitals

4 rue Gabrielle-Perret-Gentil

CH-1205 Geneva

Switzerland

Phone: +41(0)22 372 93 16

Fax: +41(0)22 372 73 04

etienne.ruppe@hcuge.ch

Article's main point: We applied metagenomic sequencing to 24 bone and joint infection samples, and showed that it was a promising tool to complement, but not yet to replace conventional methods in order to detect the bacterial pathogens and their antibiotic susceptibility patterns.

Abstract (250 words)

Background. Bone and joint infections (BJI) are severe infections that require a tailored and protracted antibiotic treatment. The diagnostic of BJI relies on the culture of surgical specimens, yet some bacteria would not grow because of extreme oxygen sensitivity or fastidious growth. Hence, metagenomic sequencing could potentially address those limitations. In this study, we assessed the performances of metagenomic sequencing of BJI samples for the identification of pathogens and the prediction of antibiotic susceptibility.

Methods. A total of 179 samples were considered. The DNA was extracted with a kit aiming to decrease the amount of human DNA (Molzym), and sequenced on an Illumina HiSeq2500 in 2x250 paired-end reads. The taxonomy was obtained by MetaPhlAn2, the bacterial reads assembled with MetaSPAdes and the antibiotic resistance determinants (ARDs) identified using a database made of Resfinder+ARDs from functional metagenomic studies.

Results. We could sequence the DNA from 24 out of 179 samples. For monomicrobial samples (n=8), the presence of the pathogen was confirmed by metagenomics in all cases. For polymicrobial samples (n=16), 32/55 bacteria (58.2%) were found at the species level (41/55 [74.5%] at the genus level). Conversely, a total of 273 bacteria not found in culture were identified, 182 being possible pathogens undetected in culture and 91 contaminants. A correct antibiotic susceptibility could be inferred in 94.1% cases for monomicrobial samples and in 76.5% cases in polymicrobial samples.

Conclusions. When sufficient amounts of DNA can be extracted from samples, we found that clinical metagenomics is a potential tool to support conventional culture.

Introduction

Bone and joint infections (BJI) are severe infections that affect a growing number of patients [1]. Along with the surgical intervention, the microbiological diagnosis is a keystone of the management of BJI in *(i)* identifying the bacteria causing the infection and (*ii*) assessing their susceptibility to antibiotics. Currently, this is achieved by culturing surgical samples on various media and conditions, together with a long time of incubation to recover fastidiously-growing bacteria that can be involved in BJI. Still, some bacteria would not grow under these conditions because of extreme oxygen sensitivity, a prior antibiotic intake or metabolic issues (quiescent bacteria in chronic infections). Consequently, the antibiotic treatment may not span all the bacteria involved in the infection, which can favor the relapse and the need for a new surgery.

Clinical metagenomics refers to the concept of sequencing all the DNA (i.e. all the genomes) present in a clinical sample with the purpose of recovering pathogens and inferring their antibiotic susceptibility pattern [2]. This new, culture-independent method takes advantages of the thrilling development of next-generation sequencing (NGS) technologies since the mid-2000s. These sequencers typically yield thousands to millions of short reads (sequences of size ranging from 100 bp to a few kbp), which virtually enables to recover the sequences of all the genes present in the sample, yet in a disorganized fashion. Substantial bio-informatics efforts are thereby needed to re-construct and re-order the original sequences in genomes, and are referred to as the assembly process. Hence, various information such as the taxonomic identification of the present species, antibiotic resistance determinants (ARDs), mutations (as compared to a reference genome or sequence), single nucleotide polymorphisms (for clonality assessment) and virulence genes can be found.

Clinical metagenomics is an emerging field in medicine. So far, a few attempts to use metagenomics on clinical samples have been performed (on urines [3,4], cerebrospinal fluid or brain biopsy [5,6], blood [7] and skin granuloma [8]) likely because of the high price of metagenomics and the complexity of the management of sequence data for clinical microbiologists. To the best of our knowledge, metagenomics has never been applied to BJI samples.

As for the inference of antibiotic susceptibility testing from the genomic information, a few studies focusing on *Escherichia coli*, *Klebsiella pneumoniae*, *Mycobacterium tuberculosis* and *Staphylococcus aureus* have constantly showed excellent correlations between the analysis of the

genomic content of antibiotic resistance determinants (ARDs) and the phenotype [9–15] while performances were not as good for *Pseudomonas aeruginosa* [16]. Furthermore in metagenomic data, the possible presence of multiple pathogens poses the issue of linking ARDs to their original host in order to infer its antibiotic susceptibility pattern [3]. So far no method has been proposed to address this question.

Applying metagenomics in the context of BJI is thus seducing in that 1) there is no limit in the number of species and ARDs that could be detected (as opposed to PCR-based methods which detect targeted bacteria), 2) unculturable bacteria, fastidious growers (such as *Propionibacterium* sp.) or bacteria altered by prior antibiotic use would be recovered, and 3) the antibiotic susceptibility inference would benefit from both the detection of ARDs (such as *mecA*, *qnr, dfr, erm, etc.*) and the identification of mutations leading to resistance to key antibiotics used in BJI. Here, our main objective is to assess the performances of clinical metagenomics in BJI in terms of pathogen identification and inference of AST, as compared with conventional microbiology (gold standard).

Material and methods

Samples

We initially included 179 per-operative samples recovered from 47 patients (range 1-8 samples per patient). All but 2 (swabs) were solid specimens. The quantity of material for each non-swab sample (n=177) was macroscopically estimated: less than 1 mL (n=100), from 1 to 10 mL (n=60) and more than 10 mL (n=17). The samples were collected from September 2015 to January 2016 in the orthopedic departments of the CRIOAc (Regional Reference Center for Complex Osteo-Orticular Infections), Lyon, France (https://www.crioac-lyon.fr) and stored at –80ºC until shipment in dry ice to the Genomic Research Laboratory in Geneva on April 13, 2016. The samples had previously been cultured (see Supplementary Methods for the detailed protocol): a single bacteria or yeast was recovered for 104 out of 179 samples (58.1%), the remaining yielding 2 (24/179, 13.4%), 3 (26/179, 14.5%), 4 (14/179, 7.8%) or 5 (11/179, 6.1%) bacteria and yeasts. The exploitation of the collection used in this study was approved by the Ethical Committee of the Lyons University Hospital (September 25, 2014).

DNA manipulations

The detailed protocol for DNA manipulations can be found in the Supplementary Methods. Briefly, the DNA from samples was extracted by the Ultra-Deep Microbiome Prep kit (Molzym, Bremen, Germany) according to the manufacturer's instructions (Version 2.0) for tissue samples. The concentration of bacterial and human DNA was determined by qPCR experiments as described previously [17]. About 3 ng of DNA were sent to Fasteris (Plan-les-Ouates, Switzerland) for DNA purification and subsequent sequencing in Rapid Run mode for 2x250+8 cycles on an Illumina HiSeq 2500 instrument (with a HiSeq Rapid Flow Cell v2).

Bioinformatic methods

The pipeline of read processing is displayed in Figure 1, and detailed in the Supplementary methods. Briefly, quality-filtered reads were processed with MetaPhlAn2 to get the taxonomic profile of the microbial community [18,19]. The bacterial reads assembled using metaSPAdes [20]. The identification of ARDs was achieved in mapping the total quality-filtered reads onto a database made of the ARDs from the ResFinder database [21] and ARDs from functional metagenomic studies [22–24]. To get the depth of sequencing of the bacterial species and of the ARDs in samples, and the single nucleotide variants (SNVs), we separately mapped the reads against the contigs assigned to one given species and against the ARDs identified in this sample The same pipeline was applied to all samples after downsizing to 1M reads.

Results

DNA extraction

We first extracted the samples for which the quantity of material exceeded or was equal to 1 mL (n=77), and the two swabs. We recovered more than 1 pg bacterial DNA mostly for samples that had grown at >100 CFUs (Supplementary Figure 1, panel A), while the concentration of human DNA did not seem to correlate with bacterial load (Supplementary Figure 1, panel B). Accordingly, the remaining samples, that had grown at least 100 CFUs (n=25), were submitted to extraction. In total, the DNA of 104 samples was extracted, from which 24 met the requirement to be sequenced (*i.e.* contained at least 1 pg/µL bacterial DNA, and less than 99% human DNA, Supplementary Table 1). Among the 24 samples and all throughout the manuscript, we will refer as monomicrobial (n=8) and

polymicrobial (n=16) samples those which respectively yielded one and more than one bacterial species in culture.

Bioinformatics

After trimming, we obtained a mean number of 10,046,084 reads per sample (range 4,128,425-14,549,687, Supplementary Table). With the Kraken classifier, the mean rate of classified reads (as bacteria, archea or virus) was 27.9% (range 1.8-85.7, Supplementary Table 1). Of note, the classification rate was correlated to the proportion of bacterial DNA as found by qPCR (Pearson's correlation test, $p<0.001$, Supplementary Figure 2). The assembly of the classified reads with metaSPAdes yielded a mean number of contigs of 10,444 (range 3,087-18,513, Supplementary Table 1), for a mean total number of base pairs of 8.3M (range 2.9M-16.5M, Supplementary Table 1). Of note, the total number of base pairs of contigs was higher in polymicrobial samples than in the monomicrobial ones (respectively 9.7M vs. 5.5M, t test $p <0.05$, Supplementary Figure 3). The mean size of the contigs was 805 bp (median 369 bp, maximum 445,300 bp, Supplementary Figure 4).

Identification of the pathogens

In monomicrobial samples (n=8, Table 2), 8/8 (100%) of the pathogens identified by culture were found by MetaPhlAn2, mostly at very high abundances (over 94.6%) at the exception of sample 46 in which *Streptococcus anginosus* was only found at a 2.2% abundance (Figure 2). In polymicrobial samples (n=16, Table 2), 55 bacterial species were found in culture 32 of which (58.2%) were found by MetaPhlAn2. At the genus level the match rate increased to 41/55 (74.5%). The presence of all bacteria found by culture in a given sample was confirmed by MetaPhlAn2 for 11/24 (45.8%) samples at the species level, including 3/16 (18.8%) for polymicrobial infections. At the genus level, 15/24 (62.5%) samples were in agreement with cultures, including 7/16 (43.8%) samples with polymicrobial infections.

Identification of other bacteria and possible contaminants

Apart from the bacteria that were found in culture (n=63) in the 24 positive samples, a total of 273 bacteria, not found in culture, were identified by MetaPhlAn2 (Figures 2 and 3). One (*Propionibacterium acnes*) was found in 20/24 samples (supplementary Figures 5 and 6). Moreover,

the abundance of *P. acnes* in samples was negatively correlated to their total DNA concentration sample (Supplementary Figure 5), supporting that *P. acnes* was a contaminant in this study [25,26]. For other species, such correlation could not be tested because of their low occurrence in samples. We identified 66 likely contaminants (Figure 3, Supplementary Table 2), some being commonly found in culture (such as *Micrococcus luteus*) or as reagents contaminants [27]. Others were unexpected such as *Borrelia sp.* (samples 103, 104, 108 and 110) or *Rickettsia japonica* (sample 117). Still, the taxonomic assignment of the contigs did not confirm the presence of those species and manual blastn of reads against the NCBI nr database supported that they were likely *in silico* contaminants (data not shown) [27,28]. Besides, we identified 25 species that could be due to a misclassification of reads to closely related bacteria, such as in samples 184 (where *Corynebacterium striatum* was found in culture, and some metagenomic reads were identified as *Dermabacter* sp. and *Corynebacterium pyruviciproducens*), samples from patient C (19, 103, 104 where *Acinetobacter baumannii* and *Achromobacter xylosoxidans* were found in culture, and some reads were identified as from other *Acinetobacter* spp., *Achromobacter* spp., or *Advenella kashmirensis*, a bacterium close to *Achromobacter*) (Supplementary Table 2). Hence, a total of 182 bacteria not recovered in culture and not acknowledged as contaminants were identified in metagenomic sequencing. For one sample that was monomicrobial in culture (sample 46, that yielded *S. anginosus*), 38 other species were identified by metagenomics. Interestingly, these species appeared to be commonly found in the oropharyngeal microbiota, which was consistent with the site of the infection (mandible). In polymicrobial samples such as samples 4 and 140 (patient B), 90 and 158 (patient H), 108 and 181 (patient I), metagenomic sequencing identified several more anaerobic bacteria (range 3-40, see Supplementary Table 2, in consistence with the sporadic isolation of such bacteria in the routine culture of these of samples. In both samples 4 and 140 from patient B, the most abundant species was *Propionibacterium propionicum* (respective abundances of 71.5% and 43.2%) that was not found in culture. Arguments in contradiction with *P. propionicum* being a contaminant in these samples are that the species found in other samples was *P. acnes*, and that the abundance of *P. propionicum* was high (supplementary Figure 6) whereas the abundance of *P. acnes* was low in the samples where it was identified.

Identification of clones within species

Based on this assumption that in case of multiple clones within one species, the SNVs would be homogeneously distributed along the contigs (see Supplementary methods and Supplementary Figure 7), we found polyclonal populations for 29 of the 74 (39.2%) bacterial species that were tested. Among the bacteria that were found in culture and that were tested (n=32), 8 (25%) displayed a polyclonal population: *Morganella morganii* (samples 4 and 140), *Streptococcus agalactiae* (samples 103 and 117), *Staphylococcus aureus* (samples 28 and 110), *S. anginosus* (sample 158) and *Pseudomonas aeruginosa* (sample 128). Moreover, we observed that for *M. morganii* (samples 4 and 140), no mutations on the topoisomerases of M. morganii were found in the sample 140, while in sample 4 the Ser83Ile and Ser84Ile were found the in GyrA and ParC, respectively. This suggests that one population of *M. morganii* was susceptible to fluoroquinolones and the other was not. In culture though, only the fluoroquinolone resistant clone was found.

Antibiotic resistance determinants, linkage with the host and inference of antibiotic susceptibility

A total of 151 ARDs (61 unique) were identified from the 24 samples (range 2-22, Table 2). The most frequent ARD families were beta-lactamases (n=30), Tet(M) (n=26), Erm (n=18) and Dfr (n=16). For monomicrobial samples, we assumed that the ARDs identified by metagenomics were expressed by the bacterium that was recovered in culture. Considering together (*i*) the antibiotic class the ARDs usually confer resistance to, (*ii*) the antibiotic susceptibility of wild-type species and (*iii*) the analysis of the sequence of specific genes (*gyrA*, *parC*, *rpoB*), we could infer a *in silico* susceptibility in agreement with the phenotypic susceptibility in 94.1% (111/118) cases (Figure 4 and supplementary tables, a case being defined as the susceptibility testing of one antibiotic for one sample). Of note, the six major errors (overprediction of resistance as compared to culture) originated from sample 46 where anaerobic bacteria and likely associated ARDs were found in metagenomic sequencing but not in culture. For polymicrobial samples, as we could not rely on the depth of sequencing of ARDs and bacterial contigs to infer some connections (Supplementary Figure 8), we separately considered the ARDs and the bacteria found in the sample (Supplementary Table 2). Accordingly, we inferred a correct susceptibility in 76.5% (192/251) cases. Very major errors mostly occurred because some bacteria with specific resistance patterns were not detected in sequencing (Supplementary Table 2). Conversely and along with the observations with monomicrobial samples, most major errors occurred

because some bacteria and ARDs were found in sequencing but not in culture. Of note, the prediction of susceptibility to fluoroquinolones was correct in 100% (24/24) samples

Influence of downsizing the samples to 1M reads

We ran the same pipeline analysis onto the 24 samples downsized at 1M reads. We observed that the taxonomic distribution did not apparently change for the most abundant species (Supplementary Figure 9), but the mean genome coverage of the main pathogen was lower in the downsized group than in the full-reads group (3.9% vs. 8.9%, Student paired test $p<0.001$, Supplementary Figure 10). Also, only 86 ARDs were found after downsizing while 151 were detected before (Student paired test $p<0.001$, Supplementary Figure 9). Of note, the impact of downsizing was observed in both monomicrobial and polymicrobial samples (Supplementary Figures 11 and 12).

Discussion

The main result of this study is that we showed that metagenomic sequencing could be a potential tool in the diagnostic of BJI. Indeed for monomicrobial infections, the pathogen was identified in 100% (8/8) samples and the antibiotic susceptibility prediction was successful in 94.1% (111/128) cases. In case of polymicrobial samples, the high abundance of several bacteria (mostly anaerobes) did occasionally prevent from the correct identification of the pathogens and their antibiotic susceptibility profiles. Accordingly, our findings support that currently, metagenomic sequencing of BJI samples could not replace conventional methods based on culture due to the limitations encountered when several bacterial populations are present in the samples, but rather be performed in support.

Interestingly, metagenomic sequencing yielded in some ways more information than culture. First, metagenomic sequencing identified many more bacterial species than culture. Besides likely contaminants, some bacteria were probably true positive that were not detected by culture and may not have been targeted by the selected antibiotic regimen. Second, we could observe within species at least two clonal populations, which could differ in their susceptibility to antibiotics as we observed for fluoroquinolones in *M. morganii*. In all, using metagenomic data could help to tailor the antibiotic regimen for the treatment of BJI, and the added-value of clinical metagenomics in BJI should now be assessed.

However, there are several hindrances to the application of metagenomic sequencing to BJI samples. First, we could only sequence 24 out of 179 samples, due to a low amount of bacterial DNA that could be recovered from the samples. This is the main limitation of this study as it reduced the diversity of clinical situations that we could address. Nonetheless, the samples from this study have been frozen and thawed, which decays bacteria and releases DNA. As the DNA extraction method we used eliminates free DNA after lysing eukaryotic cells, it is likely that we could have sequenced more samples if they would not have been frozen. This said, recovering enough bacterial DNA (in terms of quantity and proportion with respect to human DNA) remains challenging. Also, the high cost of NGS currently prevents its routine application. We tested the impact of a lower depth of sequencing and showed that despite the taxonomic profiles of the bacterial populations were similar, the inference of antibiotic susceptibility was less accurate due to a lower recovery of genes involved in antibiotic resistance. Our results suggest that clinical metagenomics should indeed benefit from the highest depth of sequencing. Another limitation of the study is that we did not concomitantly sequence a negative control to identify the putative contaminants that would originate from the sample process. Some contaminants have been identified in studies using 16S rDNA amplifications [27], but some of them are also met as BJI pathogens (e.g. *P. acnes*). A solution to this issue would be to include a negative control for every run or at least when new batches of kits are used, and to subtract from the clinical samples the bacteria found in the negative control based on their abundance [29].

Besides, our observations suggest that clinical metagenomics will soon require, as for clinical microbiology, a specific expertise combining clinical, biological and bioinformatic skills in order to infer clinically relevant results from metagenomic data. In this perspective, the development of clinical metagenomics will need the definition of quality standards, e.g. what is the sufficient genome coverage for a given bacterium to consider that its antibiotic susceptibility profile can be likely inferred. In the long term, algorithms should be built to provide clinicians with clear data and robust algorithms to support clinical decisions.

In conclusion, we showed that metagenomic sequencing of BJI samples was a potential tool to support conventional methods. In this perspective, its main limitations (DNA extraction, cost and data management) should be tackled, and the clinical benefit provided by clinical metagenomics should now be assessed in a prospective fashion.

References

1. Grammatico-Guillon L, Baron S, Gettner S, et al. Bone and joint infections in hospitalized patients in France, 2008: clinical and economic outcomes. J. Hosp. Infect. **2012**; 82:40–48.

2. Ruppé E, Baud D, Schicklin S, Guigon G, Schrenzel J. Clinical metagenomics for the management of hospital- and healthcare-acquired pneumonia. Future Microbiol. **2015**; Ahead of print.

3. Hasman H, Saputra D, Sicheritz-Ponten T, et al. Rapid whole-genome sequencing for detection and characterization of microorganisms directly from clinical samples. J. Clin. Microbiol. **2014**; 52:139–146.

4. Schmidt K, Mwaigwisya S, Crossman LC, et al. Identification of bacterial pathogens and antimicrobial resistance directly from clinical urines by nanopore-based metagenomic sequencing. J. Antimicrob. Chemother. **2016**;

5. Wilson MR, Naccache SN, Samayoa E, et al. Actionable diagnosis of neuroleptospirosis by next-generation sequencing. N. Engl. J. Med. **2014**; 370:2408–2417.

6. Frémond M-L, Pérot P, Muth E, et al. Next-Generation Sequencing for Diagnosis and Tailored Therapy: A Case Report of Astrovirus-Associated Progressive Encephalitis. J. Pediatr. Infect. Dis. Soc. **2015**; 4:e53–57.

7. Gyarmati P, Kjellander C, Aust C, Song Y, Öhrmalm L, Giske CG. Metagenomic analysis of bloodstream infections in patients with acute leukemia and therapy-induced neutropenia. Sci. Rep. **2016**; 6:23532.

8. Bodemer C, Sauvage V, Mahlaoui N, et al. Live rubella virus vaccine long-term persistence as an antigenic trigger of cutaneous granulomas in patients with primary immunodeficiency. Clin. Microbiol. Infect. Off. Publ. Eur. Soc. Clin. Microbiol. Infect. Dis. **2014**; 20:O656–663.

9. Köser CU, Holden MTG, Ellington MJ, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. N. Engl. J. Med. **2012**; 366:2267–2275.

10. Köser CU, Bryant JM, Becq J, et al. Whole-genome sequencing for rapid susceptibility testing of M. tuberculosis. N. Engl. J. Med. **2013**; 369:290–292.

11. Zankari E, Hasman H, Kaas RS, et al. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. J. Antimicrob. Chemother. **2013**; 68:771–777.

12. Gordon NC, Price JR, Cole K, et al. Prediction of Staphylococcus aureus antimicrobial resistance by whole-genome sequencing. J. Clin. Microbiol. **2014**; 52:1182–1191.

13. Stoesser N, Batty EM, Eyre DW, et al. Predicting antimicrobial susceptibilities for Escherichia coli and Klebsiella pneumoniae isolates using whole genomic sequence data. J. Antimicrob. Chemother. **2013**; 68:2234–2244.

14. Tyson GH, McDermott PF, Li C, et al. WGS accurately predicts antimicrobial resistance in Escherichia coli. J. Antimicrob. Chemother. **2015**;

15. Bradley P, Gordon NC, Walker TM, et al. Rapid antibiotic-resistance predictions from genome sequence data for Staphylococcus aureus and Mycobacterium tuberculosis. Nat. Commun. **2015**; 6:10063.

16. Kos VN, Déraspe M, McLaughlin RE, et al. The resistome of Pseudomonas aeruginosa in relationship to phenotypic susceptibility. Antimicrob. Agents Chemother. **2015**; 59:427–436.

17. Lazarevic V, Gaïa N, Girard M, François P, Schrenzel J. Comparison of DNA extraction methods in analysis of salivary bacterial communities. PloS One **2013**; 8:e67699.

18. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat. Methods **2012**; 9:811–814.

19. Truong DT, Franzosa EA, Tickle TL, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat. Methods **2015**; 12:902–903.

20. Nurk S, Meleshko D, Korobeynikov A, Pevzner P. metaSPAdes: a new versatile de novo metagenomics assembler. ArXiv160403071 Q-Bio **2016**; Available at: http://arxiv.org/abs/1604.03071. Accessed 13 July 2016.

21. Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. J. Antimicrob. Chemother. **2012**; 67:2640–2644.

22. Sommer MOA, Dantas G, Church GM. Functional characterization of the antibiotic resistance reservoir in the human microflora. Science **2009**; 325:1128–1131.

23. Moore AM, Patel S, Forsberg KJ, et al. Pediatric fecal microbiota harbor diverse and novel antibiotic resistance genes. PloS One **2013**; 8:e78822.

24. Pehrsson EC, Tsukayama P, Patel S, et al. Interconnected microbiomes and resistomes in low-income human habitats. Nature **2016**; 533:212–216.

25. Lazarevic V, Gaïa N, Emonet S, et al. Challenges in the culture-independent analysis of oral and respiratory samples from intubated patients. Front. Cell. Infect. Microbiol. **2014**; 4:65.

26. Willner D, Daly J, Whiley D, Grimwood K, Wainwright CE, Hugenholtz P. Comparison of DNA extraction methods for microbial community profiling with an application to pediatric bronchoalveolar lavage samples. PloS One **2012**; 7:e34605.

27. Salter SJ, Cox MJ, Turek EM, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. **2014**; 12:87.

28. Gonzalez A, Vázquez-Baeza Y, Pettengill JB, Ottesen A, McDonald D, Knight R. Avoiding Pandemic Fears in the Subway and Conquering the Platypus. mSystems **2016**; 1:e00050–16.

29. Lazarevic V, Gaïa N, Girard M, Schrenzel J. Decontamination of 16S rRNA gene amplicon sequence datasets based on bacterial load assessment by qPCR. BMC Microbiol. **2016**; 16:73.

30. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. **2014**; 15:R46.

31. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinforma. Oxf. Engl. **2013**; 29:1072–1075.

Funding source

None.

Conflict of interest

All authors declare that they have no conflict of interest.

Acknowledgements

None.

Tables

Table 1: Characteristics of the 14 patients for whom 24 samples were sequenced. ASA: American Society of Anesthesiologists.

Patient	Samples	Age	Gender	ASA score	Body mass index	Post-operative infection (type of surgery)	Delay between surgery and infection	Body site	Material involved
A	2, 66, 28	51	M	2	30.4	Yes (material)	<1 month	Ankle	Osteosynthesis
B	4, 140	50	F	2	39.8	No	NA	Clavicle	None
C	19, 103, 104	54	M	2	24.1	Yes (material)	<1 month	Toe	Osteosynthesis
D	110	66	M	2	29.4	Yes (material)	Between 1 and 3 months	Tibia	Osteosynthesis
E	42	61	F	3	50.6	Yes (material)	<1 month	Knee	Total knee prothesis
F	46	63	M	2	18.0	Yes (material)	<1 month	Mandible	Osteosynthesis
G	59, 117, 136	69	M	2	25.5	Yes (bone resection)	NA	Tibia	None
H	90, 158	64	F	2	21.2	No	NA	Sacrum	None
I	108, 181	86	F	2	26.7	Yes (material)	Between 1 and 3 months	Knee	Total knee prothesis
J	121, 172	50	F	1	24.2	No	NA	Tibia	None
K	128	86	F	2	30.1	No	>3 months	Knee	Osteosynthesis
L	171	51	M	1	25.6	Yes (material)	>3 months	Tibia	Osteosynthesis
M	178	87	F	3	26.1	Yes (material)	<1 month	Knee	Total knee prothesis
N	184	60	M	3	34.3	No	NA	Greater trochanter and ischium	None

Table 2: Description of the 24 samples sequenced in this study. WT: wild-type. NA: not assembled.

Patient	Sample number	Monomicrobial or polymicrobial	Culture (proportion in %)	Species identified in metagenomic sequencing (≥0.1% abundance)	Antibiotic resistance genes (Resfinder)	Antibiotic resistance genes (functional metagenomic studies)	GyrA	ParC	RpoB
A	2	Polymicrobial	Staphylococcus aureus (29.4), Klebsiella pneumoniae (5.9), Klebsiella oxytoca (5.9), Peptoniphilus asaccharolyticus (29.4), Finegoldia magna (29.4)	Staphylococcus aureus (99.6)	blaZ , norA	None	WT (*S. aureus*)	WT (*S. aureus*)	D320N (*S. aureus*)
B	4	Polymicrobial	Morganella morganii (99.9), Streptococcus anginosus (0.1)	Morganella morganii (5.6), Propionibacterium propionicum (71.5), Bacteroides fragilis (14.2), Prevotella bivia (4.1), Atopobium rimae (3.0), Parvimonas unclassified (1.0), Parvimonas micra (0.2), Prevotella buccalis (0.1)	aadA1 , aph(3')-Ia , blaDHA-1 , catA1 , catA2 , cepA , cfxA3 , dfrA1 , erm(A) , erm(B) , strA , sul2 , tet(D) , tet(M) , tet(Q)	dfr, dfr, van, tet(B)			
C	19	Polymicrobial	Acinetobacter baumanii complex (90.8), Streptococcus agalactiae (9.2)	Acinetobacter baumanii complex (22.1), Streptococcus agalactiae (73.1), Finegoldia magna (1.7), Acinetobacter_pittii_calcoaceticus nosocomialis (0.9), Corynebacterium resistens (0.9), Helcococcus kunzii (0.7), Advenella kashmirensis (0.3), Propionibacterium acnes (0.2), Achromobacter unclassified (0.1), Achromobacter piechaudii (0.1)	blaADC-25 , blaOXA-328 , erm(B) , tet(M)	None	WT (A. baumannii, S. agalactiae)	WT (A. baumannii, S. agalactiae)	WT (S. agalactiae)
A	28	Polymicrobial	Staphylococcus aureus (33.3), Klebsiella pneumoniae (16.7), Klebsiella oxytoca (16.7), Finegoldia magna (16.7), Peptostreptococcus anaerobius (16.7)	Staphylococcus aureus (90.3), Finegoldia magna (4.3), Peptoniphilus harei (3.7), Propionibacterium acnes (0.9)	ant(6)-Ia , aph(3')-III , norA	tet(O)	WT (*S. aureus*)	WT (*S. aureus*)	D320N (*S. aureus*)
E	42	Polymicrobial	Staphylococcus epidermidis (50.0), Streptococcus mitis/oralis (50.0)	Staphylococcus epidermidis (99.1), Propionibacterium acnes (0.4)	aac(6')-aph(2") , aph(3')-Ia , blaZ , erm(C) , fosB , mecA	blaTEM	S84Y (S. epidermidis)	S80F, D84Y (S. epidermidis)	T700S, D837E (S. epidermidis)
F	46	Monomicrobial	Streptococcus anginosus (100.0)	Streptococcus anginosus (2.2), Mogibacterium_sp CM50 (34.5), Olsenella uli (13.8), Atopobium_sp_oral_taxon 199 (13.0), Peptostreptococcus stomatis (6.9), Parvimonas unclassified (6.7), Fretibacterium fastidiosum (6.3), Pseudoramibacter alactolyticus (5.6), Slackia unclassified (3.2), Slackia exigua (2.2), Alloprevotella tannerae (2.0), Prevotella oris (0.8), Eubacterium infirmum (0.5), Treponema maltophilum (0.4), Parvimonas micra (0.3), Prevotella baroniae (0.2), Treponema socranskii (0.2), Bacteroidetes_bacterium_oral_taxon 272 (0.2), Dialister invisus (0.2), Prevotella denticola (0.2), Tannerella forsythia (0.1), Porphyromonas uenonis (0.1), Treponema vincentii (0.1), Treponema denticola (0.1), Peptoniphilus lacrimalis (0.1)	cfxA3 , lsa(C) , tet(M)	dfr, dfr, dfr, tet(M), van	C96Y	WT	D492A
G	59	Monomicrobial	Streptococcus agalactiae (100.0)	Streptococcus agalactiae (99.9)	erm(B) , tet(M)	None	WT	WT	WT
A	66	Polymicrobial	Staphylococcus aureus (27.8), Klebsiella pneumoniae (11.1), Klebsiella oxytoca (5.6), Finegoldia magna (27.8), Peptoniphilus asaccharolyticus (27.8)	Staphylococcus aureus (99.9)	blaZ , norA	None	WT (*S. aureus*)	WT (*S. aureus*)	D320N (*S. aureus*)
H	90	Polymicrobial	Streptococcus anginosus (82.6), Enterococcus faecalis (2.5), Bacteroides fragilis (8.3), Clostridium ramosum (4.1), Clostridium clostridioforme (2.5)	Streptococcus anginosus (0.6), Olsenella_sp_oral_taxon 809 (55.9), Pseudoramibacter alactolyticus (9.6), Atopobium_sp_oral_taxon 199 (8.4), Eggerthella unclassified (6.6), Slackia unclassified (5.8), Mogibacterium_sp CM50 (5.3), Slackia exigua (1.8), Peptoniphilus_sp_oral_taxon 375 (1.6), Anaerococcus lactolyticus (1.0), Peptoniphilus harei (0.6), Finegoldia magna (0.5), Parvimonas unclassified (0.4), Peptoniphilus lacrimalis (0.4), Olsenella uli (0.3), Eggerthella lenta (0.3), Parvimonas micra (0.1), Porphyromonas asaccharolytica (0.1), Actinomyces europaeus (0.1), Actinomyces turicensis (0.1), Coriobacteriaceae_bacterium BV3Ac1 (0.1), Porphyromonas somerae (0.1), Subdoligranulum unclassified (0.1)	ant(6)-Ia , aph(3')-III , erm(A) , erm(B) , erm(X) , strA , sul2 , tet(32) , tet(M) , tet(W)	dfr, tet(O)			
C	103	Polymicrobial	Acinetobacter baumanii complex (90.8), Streptococcus agalactiae (9.2)	Acinetobacter baumanii complex (66.9), Streptococcus agalactiae (26.9), Acinetobacter_pittii_calcoaceticus nosocomialis (2.7), Finegoldia magna (1.5), Corynebacterium resistens (0.4), Staphylococcus simulans (0.4), Propionibacterium acnes (0.3), Staphylococcus lugdunensis (0.2), Achromobacter unclassified (0.1), Helcococcus kunzii (0.1), Bordetella unclassified (0.1), Advenella kashmirensis (0.1)	blaADC-25 , blaOXA-328 , erm(B) , tet(M)	None	WT (A. baumannii, S. agalactiae)	WT (A. baumannii, S. agalactiae)	WT (S. agalactiae)
C	104	Polymicrobial	Acinetobacter baumanii complex (90.8), Streptococcus agalactiae (9.2), Achromobacter xylosoxidans (0.0)	Acinetobacter baumanii complex (70.1), Streptococcus agalactiae (24.2), Acinetobacter_pittii_calcoaceticus nosocomialis (3.9), Corynebacterium resistens (1.1), Propionibacterium acnes (0.2), Achromobacter unclassified (0.1), Bordetella unclassified (0.1), Enhydrobacter aerosaccus (0.1)	blaADC-25 , blaOXA-328 , tet(M)	None	WT (A. baumannii, S. agalactiae)	WT (A. baumannii, S. agalactiae)	WT (S. agalactiae)
I	108	Polymicrobial	Enterococcus faecalis (1.0), Peptoniphilus asaccharolyticus (99.0)	Peptoniphilus harei (98.7), Propionibacterium acnes (0.7), Streptococcus agalactiae (0.1), Deinococcus unclassified (0.1), Acinetobacter unclassified (0.1)	None	tet(M), tet(M)	NA	NA	NA
D	110	Monomicrobial	Staphylococcus aureus (100.0)	Staphylococcus aureus (98.2), Propionibacterium acnes (1.7)	blaZ , norA	None	WT	WT	WT
G	117	Monomicrobial	Streptococcus agalactiae (100.0)	Streptococcus agalactiae (99.9), Rickettsia japonica (0.1)	erm(B) , tet(M)	None	WT	WT	WT
J	121	Monomicrobial	Staphylococcus aureus (100.0), ,	Staphylococcus aureus (99.6), Propionibacterium acnes (0.4)	blaZ , norA	None	WT	NA	WT
K	128	Polymicrobial	Proteus mirabilis (NA), Klebsiella oxytoca (NA), Pseudomonas aeruginosa (MA)	Klebsiella oxytoca (74.2), Pseudomonas aeruginosa (0.8), Klebsiella unclassified (23.5), Rothia mucilaginosa (0.3), Pseudomonas unclassified (0.3), Propionibacterium acnes (0.2)	aac(3)-IIa , aac(6')Ib-cr , blaCTX-M-11 , blaOXA-1 , blaOXY-2-8 , dfrA14 , fosA , oqxA , oqxB , QnrB1 ,	Putative beta-lactamase, blaTEM, vanB	T83I (*K. oxytoca*)	S80I (*K. oxytoca*)	
G	136	Monomicrobial	Streptococcus agalactiae (100.0)	Streptococcus agalactiae (99.8), Propionibacterium acnes (0.1)	erm(B) , tet(M) , tet(M)	None	WT	WT	WT
B	140	Polymicrobial	Morganella morganii (76.9), Streptococcus anginosus (0.1), Prevotella bivia (7.7), Bifidobacterium (7.7), Peptoniphilus (7.7)	Morganella morganii (14.6), Prevotella bivia (6.6), Propionibacterium propionicum (43.2), Bacteroides fragilis (24.4), Atopobium rimae (8.3), Parvimonas unclassified (1.2), Dermabacter_sp HFH0086 (1.0), Parvimonas micra (0.3), Anaeroglobus geminatus (0.1), Proteus mirabilis (0.1)	aadA1 , aph(3')-Ia , blaMOR , blaTEM-1A , catA1 , catA2 , cepA , cfxA3 , dfrA1 , dfrA14 , erm(B) , QnrS1 , strA , sul1 , sul2 , tet(D) , tet(M) , tet(Q)	dfr, dfr, tet(B), van			
H	158	Polymicrobial	Enterococcus faecalis (42.6), Streptococcus anginosus (42.6), Corynebacterium coyleae (14.9)	Enterococcus faecalis (3.8), Streptococcus anginosus (1.3), Olsenella_sp_oral_taxon 809 (69.8), Mogibacterium_sp CM50 (6.5), Slackia unclassified (3.8), Peptoniphilus harei (1.6), Atopobium_sp_oral_taxon 199 (1.5), Finegoldia magna (1.4), Pseudoramibacter alactolyticus (1.2), Anaerococcus lactolyticus (1.2), Peptoniphilus_sp_oral_taxon 375 (1.1), Slackia exigua (1.0), Porphyromonas asaccharolytica (0.8), Actinomyces turicensis (0.8), Subdoligranulum unclassified (0.7), Peptoniphilus lacrimalis (0.7), Eggerthella unclassified (0.4), Olsenella uli (0.4), Anaerococcus vaginalis (0.2), Lachnospiraceae_bacterium_5_1 57FAA (0.2), Dialister invisus (0.2), Clostridium clostridioforme (0.2), Actinomyces europaeus (0.2), Porphyromonas somerae (0.2), Clostridiales_bacterium BV3C26 (0.1), Anaerococcus obesiensis (0.1), Parvimonas unclassified (0.1), Helcococcus kunzii (0.1), Prevotella timonensis (0.1), Facklamia hominis (0.1), Eggerthella lenta (0.1)	ant(6)-Ia , aph(3')-III , cmx , erm(A) , erm(B) , erm(X) , lsa(A) , strA , strB , tet(M)	cfxA, dfr, dfr, dfr, tet(O), tet(O)			
L	171	Polymicrobial	Staphylococcus aureus (40.0), Staphylococcus lugdunensis (20.0), Anaerococcus vaginalis (40.0)	Staphylococcus aureus (97.8), Anaerococcus vaginalis (1.6), Propionibacterium acnes (0.5), Bartonella unclassified (0.1)	blaZ , norA	None	WT (*S. aureus*)	WT (*S. aureus*)	WT (*S. aureus*)
J	172	Monomicrobial	Staphylococcus aureus (100.0), ,	Staphylococcus aureus (94.6), Micrococcus luteus (3.1), Propionibacterium acnes (2.2)	norA	blaTEM	WT	NA	WT
M	178	Monomicrobial	Staphylococcus epidermidis (100.0)	Staphylococcus epidermidis (99.7), Propionibacterium acnes (0.2)	aac(6')-aph(2") , blaZ , erm(A) , fusB , mecA , spc , vat(B) , vga(A) , vga(B)	None	S84F	S80Y	WT
I	181	Polymicrobial	Enterococcus faecalis (4.5), Staphylococcus carnosus (4.5), Staphylococcus lugdunensis (45.5), Propionimicrobium , 45.5, Anaerococcus vaginalis (0.0)	Enterococcus faecalis (62.5), Propionimicrobium (1.2), Anaerococcus vaginalis (0.5), Peptoniphilus harei (20.7), Actinomyces neuii (7.5), Finegoldia magna (7.1), Anaerococcus obesiensis (0.2), Propionibacterium acnes (0.1)	aph(3')-III , lsa(A)	dfr, blaTEM, tet(M)	WT (*E. faecalis*)	WT (*E. faecalis*)	WT (*E. faecalis*)
N	184	Polymicrobial	Escherichia coli (27.3), Enterococcus faecalis (45.5), Corynebacterium striatum (27.3)	Enterococcus faecalis (2.3), Corynebacterium striatum (4.7), Finegoldia magna (53.6), Dermabacter_sp HFH0086 (13.7), Peptoniphilus harei (10.3), Varibaculum cambriense (8.6), Anaerococcus vaginalis (2.3), Propionibacterium acnes (2.0), Anaerococcus obesiensis (0.9), Escherichia_unclassified (0.5), Corynebacterium pyruviciproducens (0.1)	erm(A) , erm(X)	tet(M), tet(O)	S83I (*F. magna*)		

Figures

Figure 1: Bioinformatic analysis performed in this study. ARDs: antibiotic resistance determinants. Fastq: format for the files that embeds the read sequences and their per-base quality score. FMG: functional metagenomics; ARDs: antibiotic resistance determinants; ORF: open reading frame.

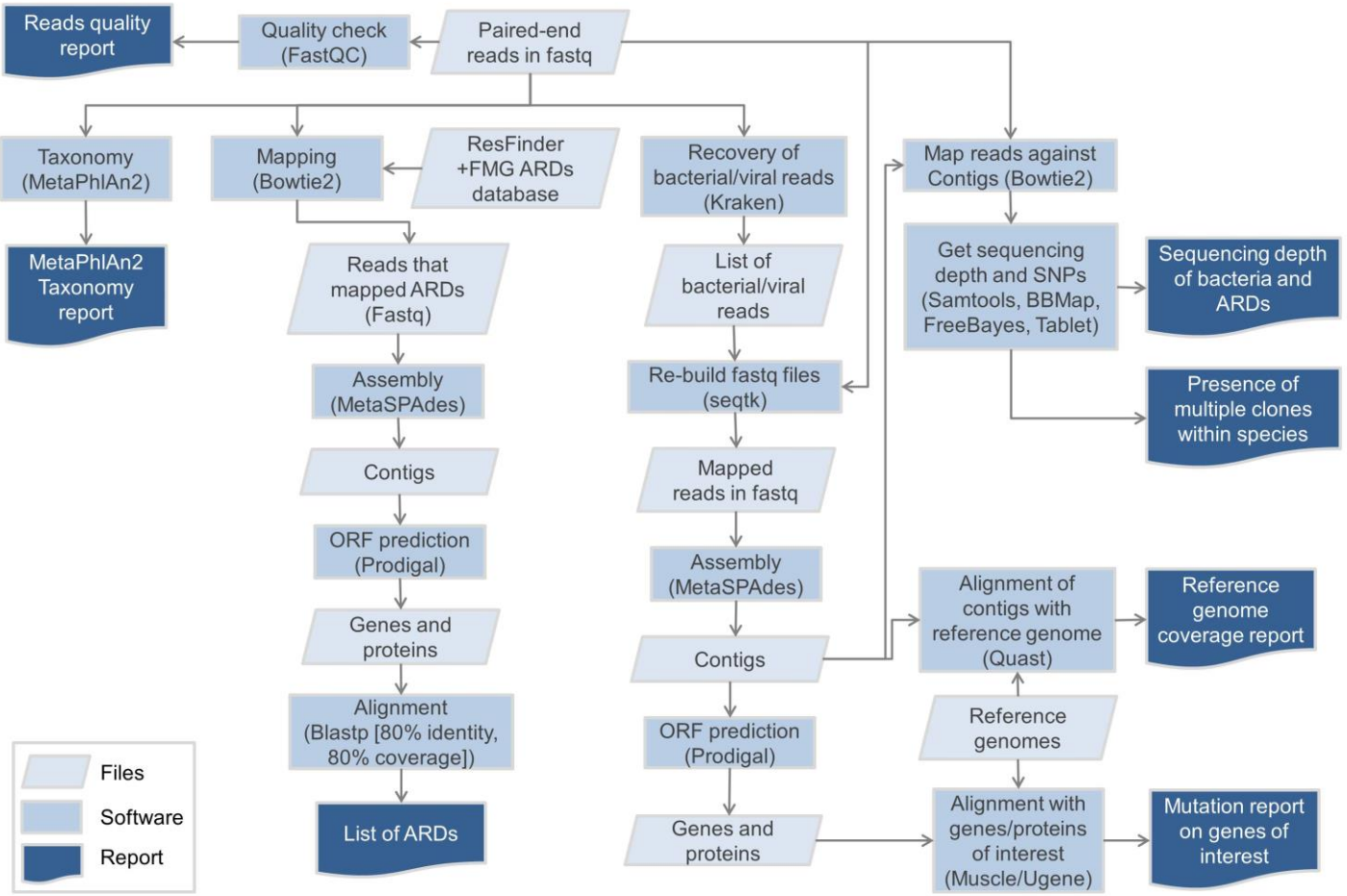


Figure 2: Proportions of the species recovered in culture and from reads (using MetaPhlAn2 [19]).

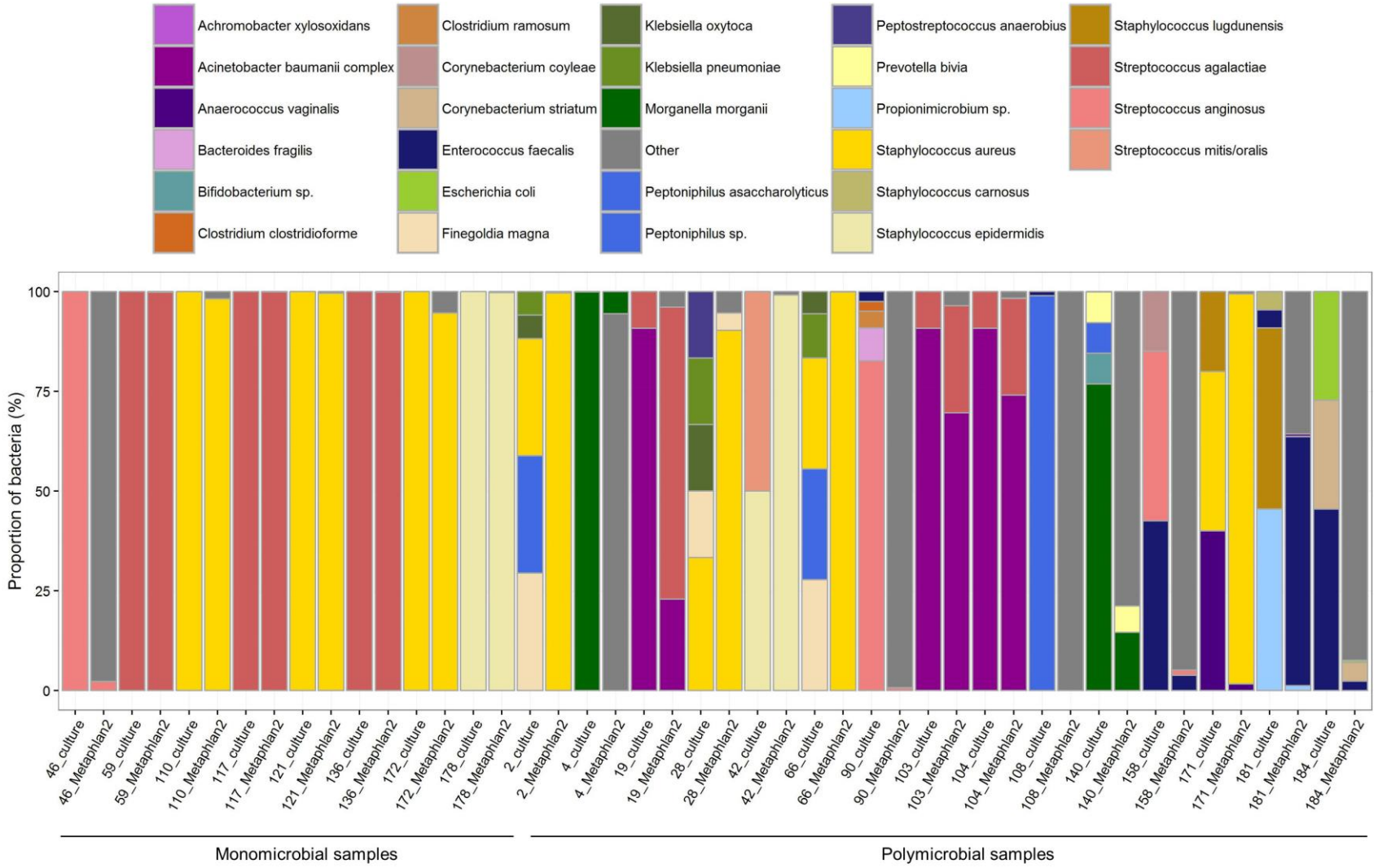


Figure 3: Distribution of the number of species found both in culture and metagenomic sequencing, only in culture and only metagenomic sequencing (in this case, putative pathogenic species and contaminants/misclassifications are depicted apart). All species identified by MetaPhlAn2 were considered (not only those above 0.1% abundance).

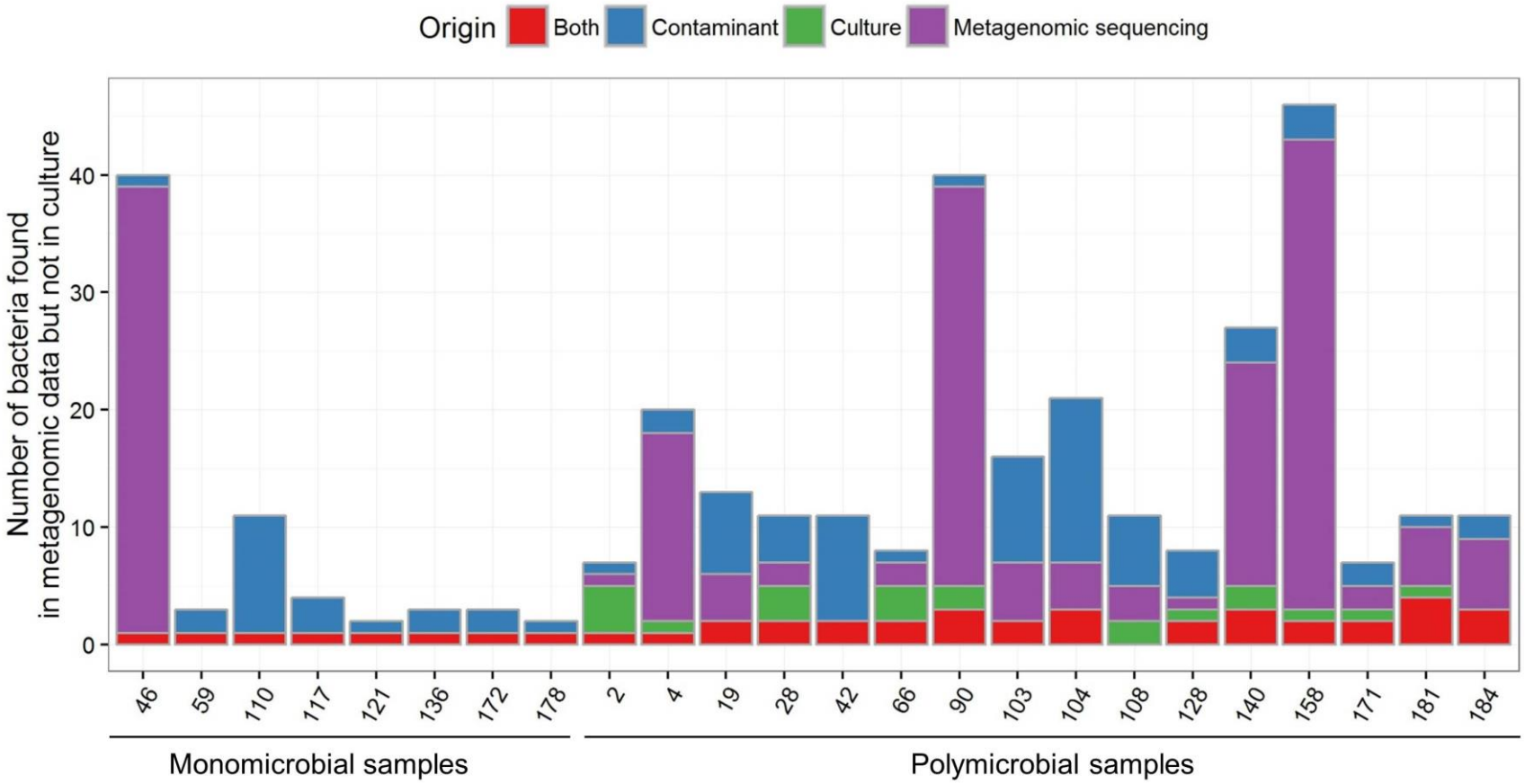
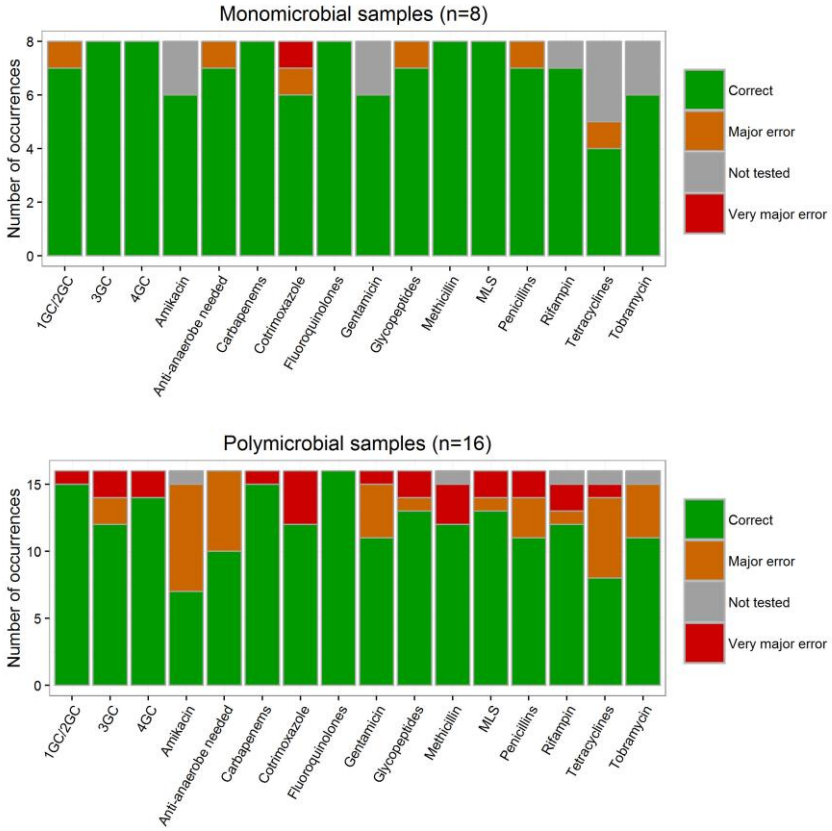
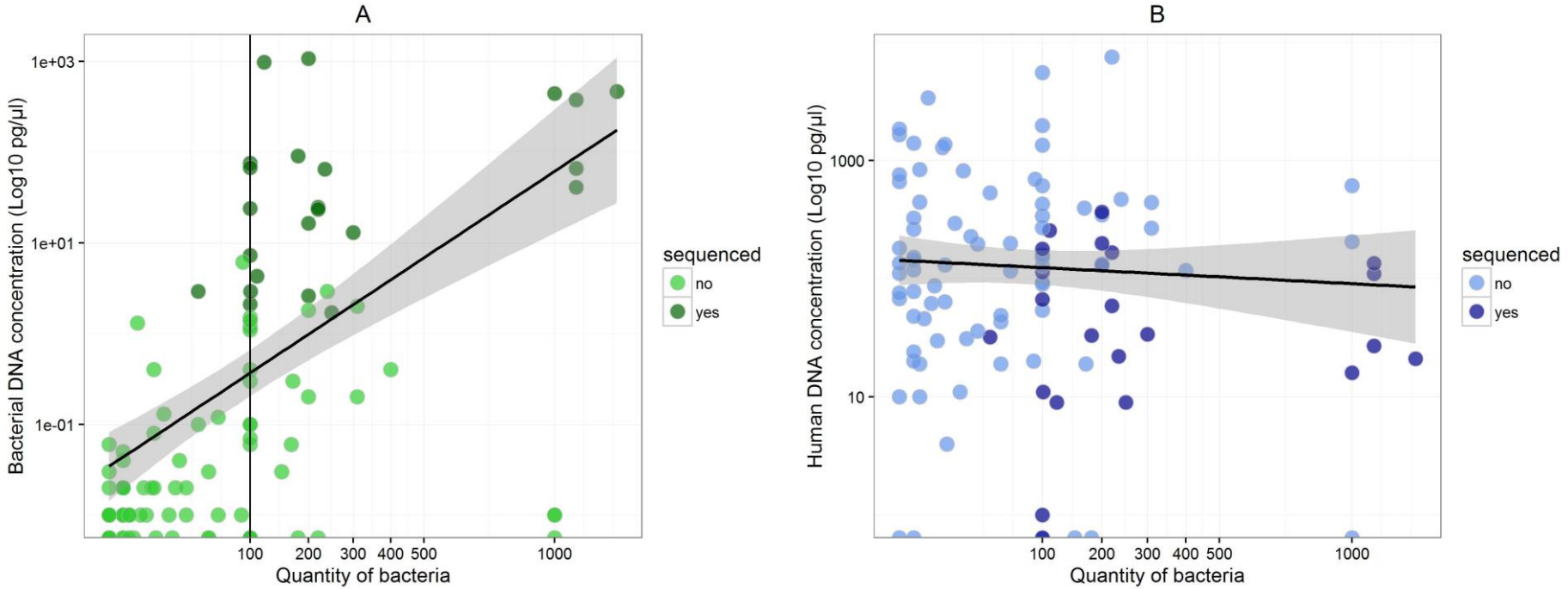


Figure 4: Antibiotic susceptibility inference from metagenomic data compared to culture and conventional antibiotic susceptibility testing (gold standard). 1GC/2GC, 3GC, 4GC: 1st, 2nd, 3rd and 4th generation cephalosporins, respectively. MLS: macrolides, lincosamides, streptogramines. Correct: metagenomic result consistent with the result given by conventional methods. Very major error: metagenomic data did not predict antibiotic resistance while at least one bacteria identified by conventional methods was resistant to this antibiotic. Major error: metagenomic data predicted antibiotic resistance while all the bacteria identified by conventional methods were susceptible. Not tested: no molecule from the antibiotic class was tested with conventional methods.

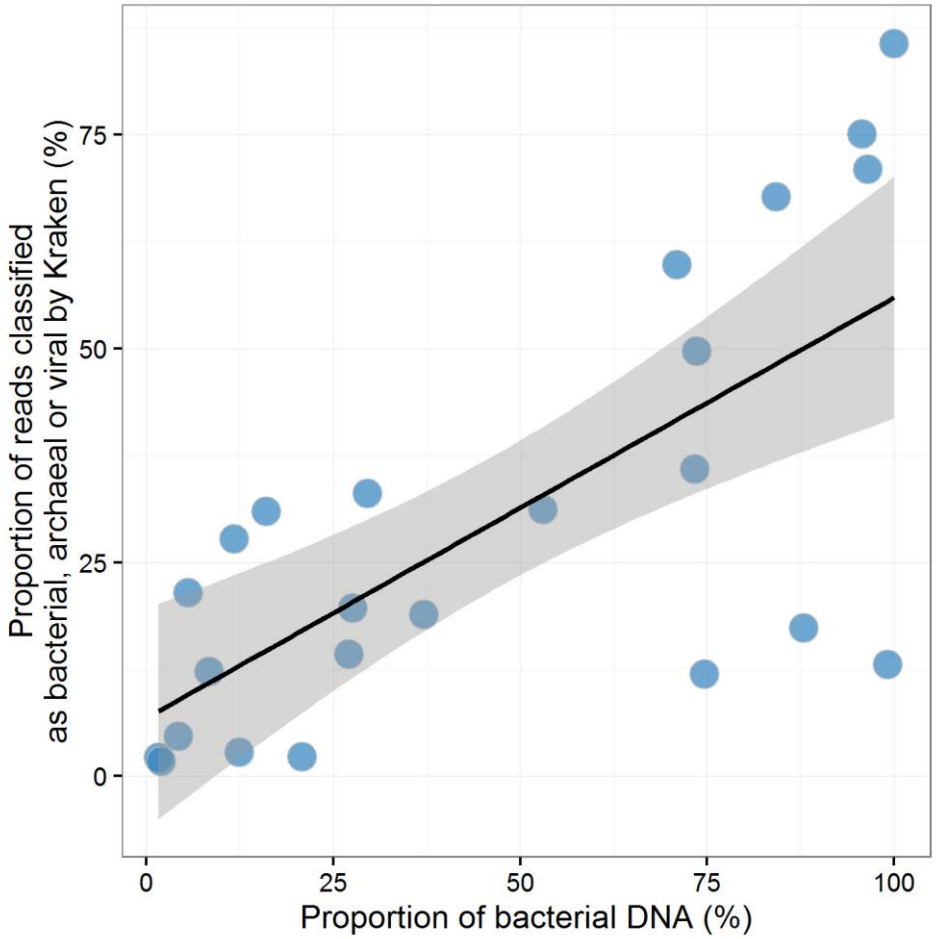


Supplementary Figures

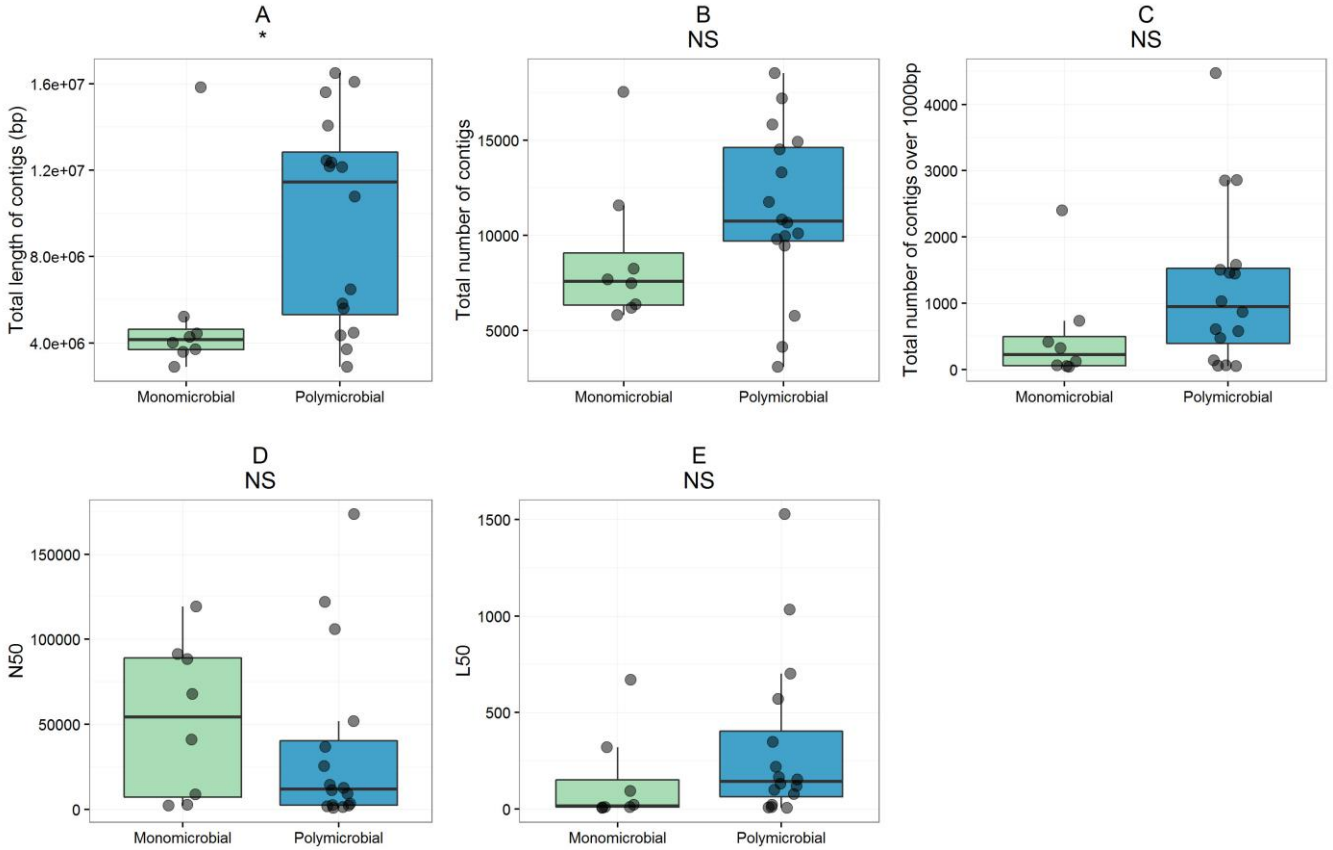
Supplementary Figure 1: Total quantity of bacteria obtained in culture and concentrations of bacterial (A) and human DNA (B) for 102 samples for which DNA was extracted. DNA concentrations in DNA extracts were determined by qPCR (see methods). The shaded grey area depicts the 95% confidence interval around the linear regression line. The X-axis is square-root transformed for visibility purposes. One sequenced sample is missing (sample 128) because the bacterial concentrations were missing. Eventually, 19 and 7 samples with null bacterial and human DNA concentrations are not shown in the panel A and B, respectively.



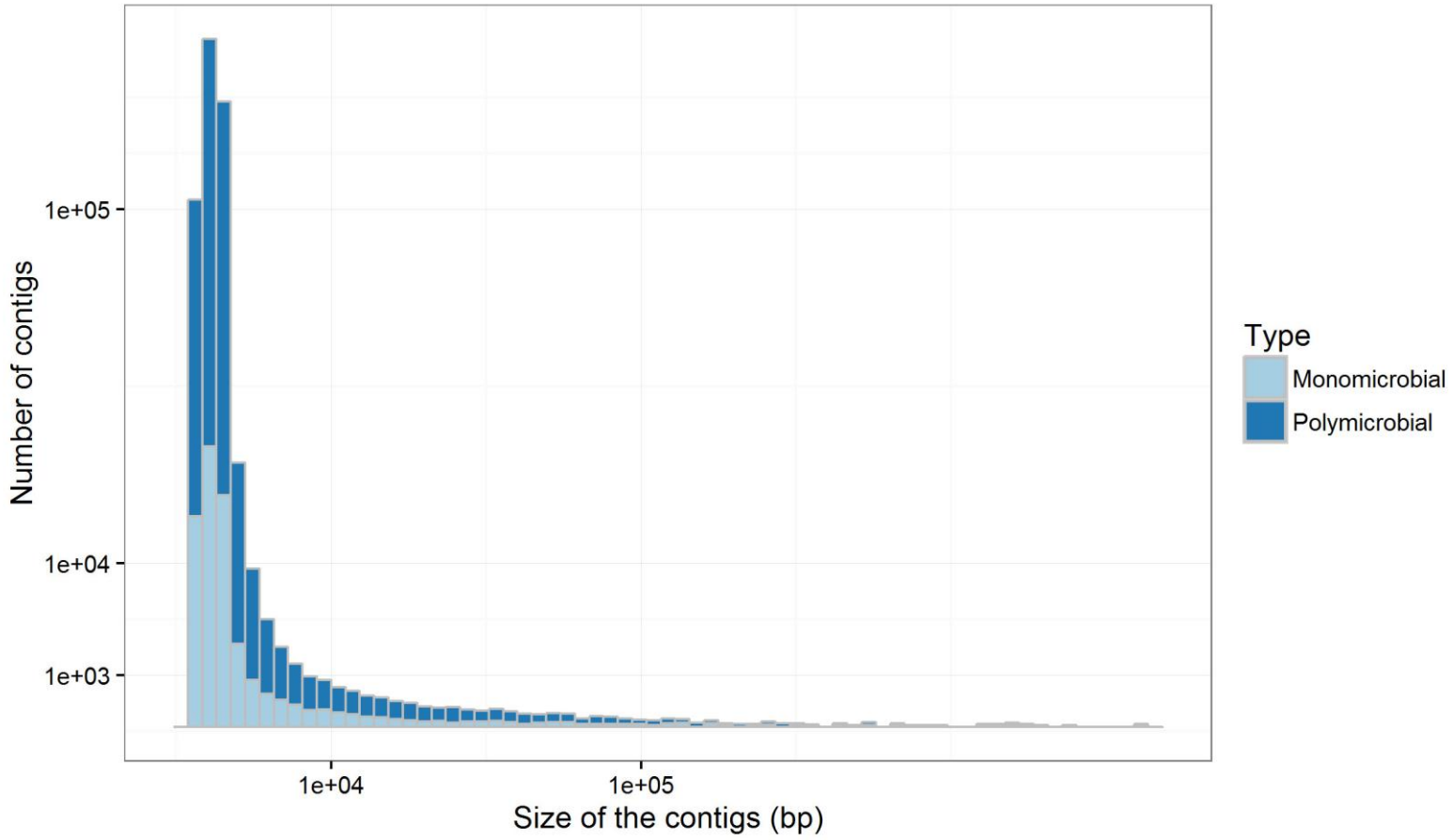
Supplementary Figure 2: Proportion of bacterial DNA as determined by qPCR (the proportion being calculated as the percentage of bacterial DNA on the total DNA [bacterial and human]) on the X-axis, and proportion of the quality-filtered reads classified as bacterial, archaeal or viral by the Kraken classifier. The shaded grey area depicts the 95% confidence interval around the linear regression line.



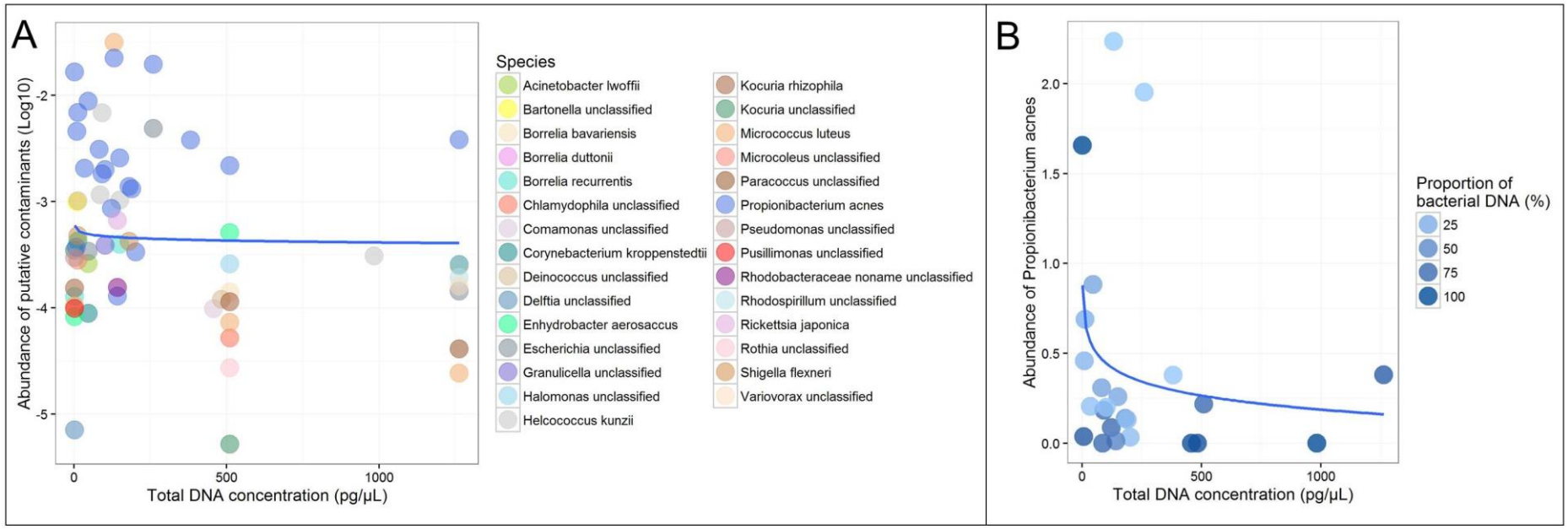
Supplementary Figure 3: Boxplots of the the length of summed contigs (A), the total number of contigs (B), the total number of contigs exceeding 1000bp (C), the N50 (D), the L50 (E) according to the number of bacteria recovered in culture (monomicrobial or polymicrobial). The contigs were obtained by the assembly by MetaSPAdes [20] of reads classified by Kraken [30]. The parameters showed on this figure were obtained by Quast [31]. *: $p<0.05$; NS: not significant. The boxplot limits represents (from bottom to roof) the 25th, 50th (median) and 75th percentiles.



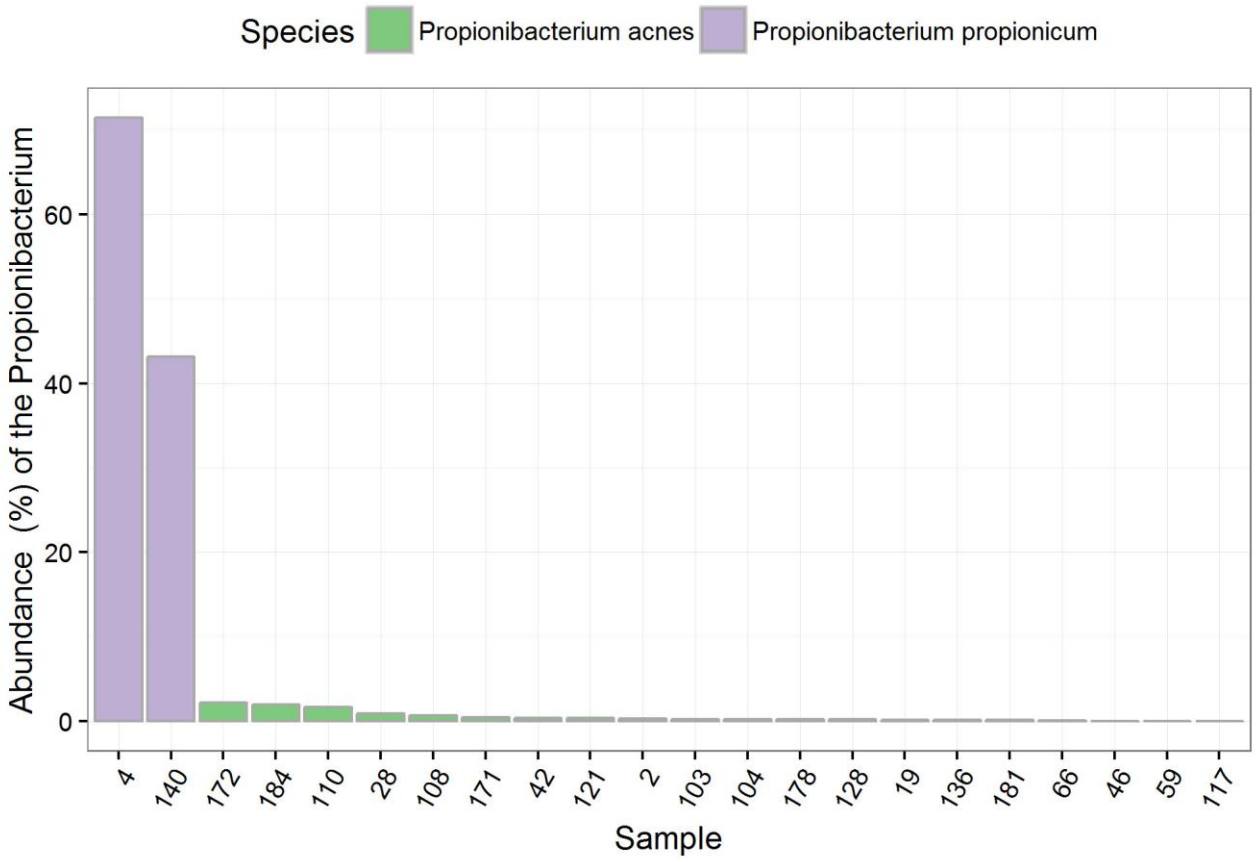
Supplementary Figure 4: Distribution of the size of the contigs for all the samples (n=24). Both Y-axis and X-axis are square-root transformed.



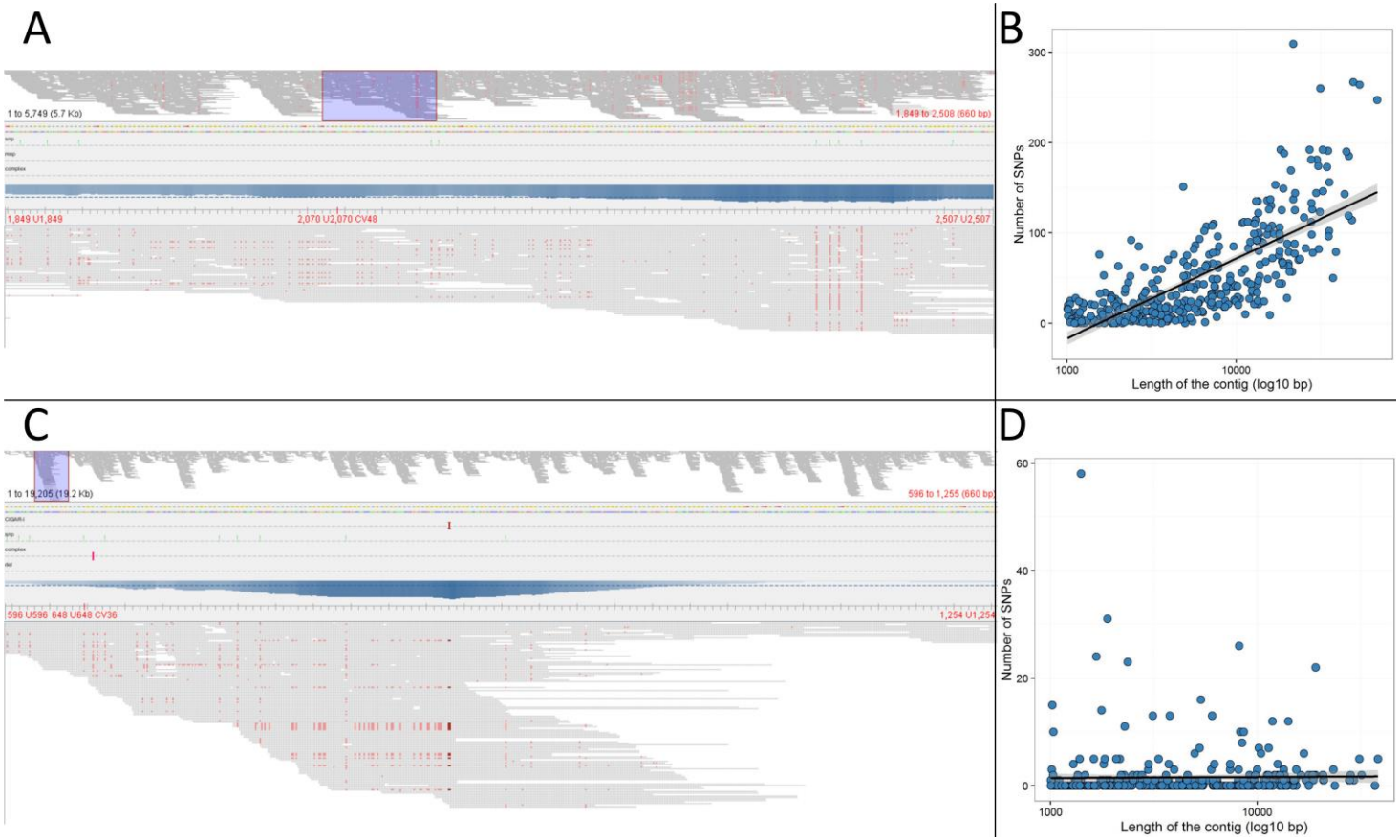
Supplementary Figure 5: (A) Dot plot of the abundance of the species (Log10) that were considered as contaminants in this study (see supplementary Table 1) along with the total DNA (sum of bacterial and human DNA) concentrations (pg/µL). The blue line depicts the linear regression between the abundances of the species (Log10) and the Log10 of the DNA concentrations. (B) Dot plot of the abundance of *Propionibacterium acnes* (%) along with the total DNA (sum of bacterial and human DNA) concentrations (pg/µL). Samples with an abundance of 0 are also showed. The blue line depicts the linear regression between the abundances of the species and the Log10 of the DNA concentrations.



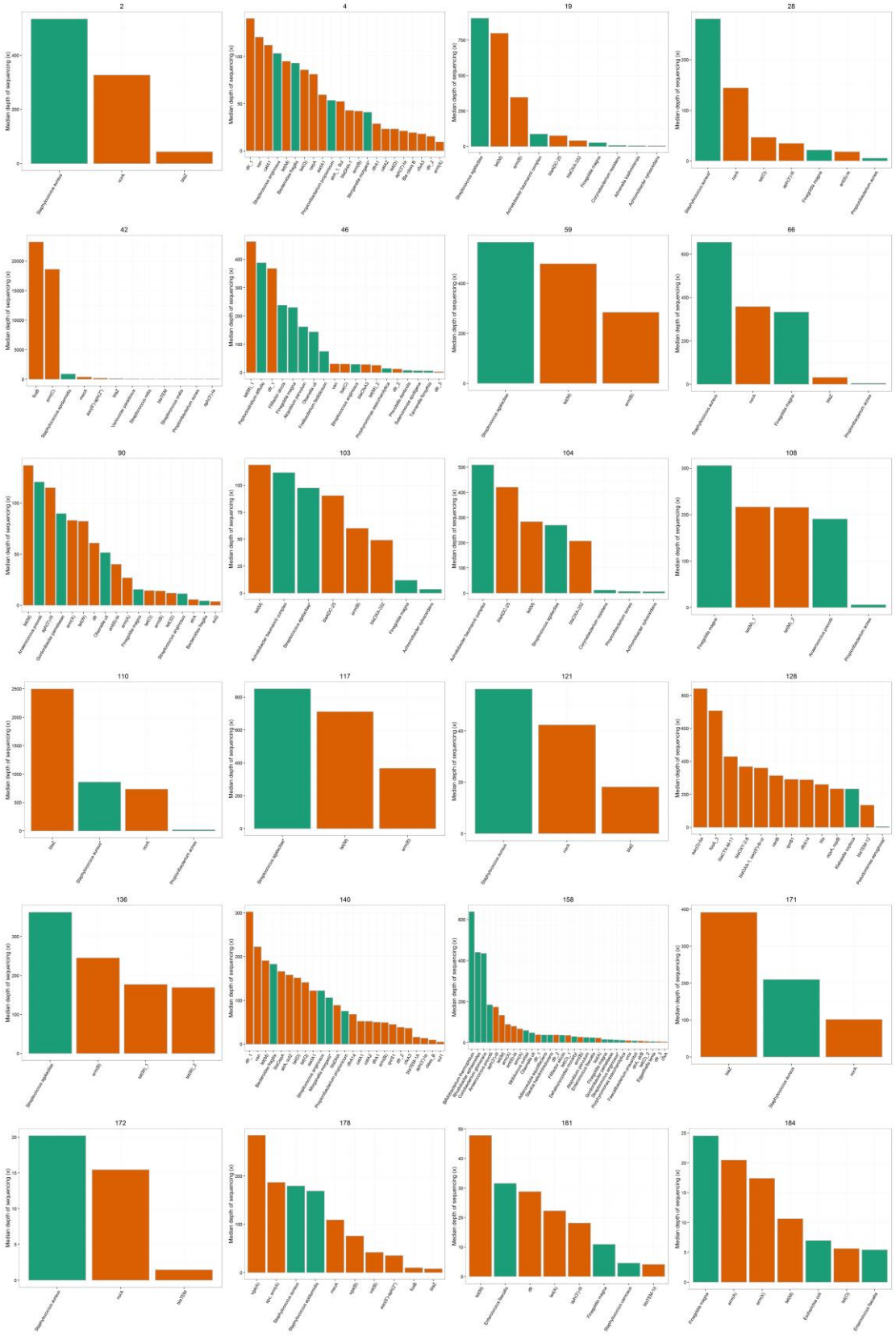
Supplementary Figure 6: Distribution of the abundances (%) of *Propionibacterium propionicum* (purple) and *Propionibacterium acnes* (green) in the 22 samples where they were identified.



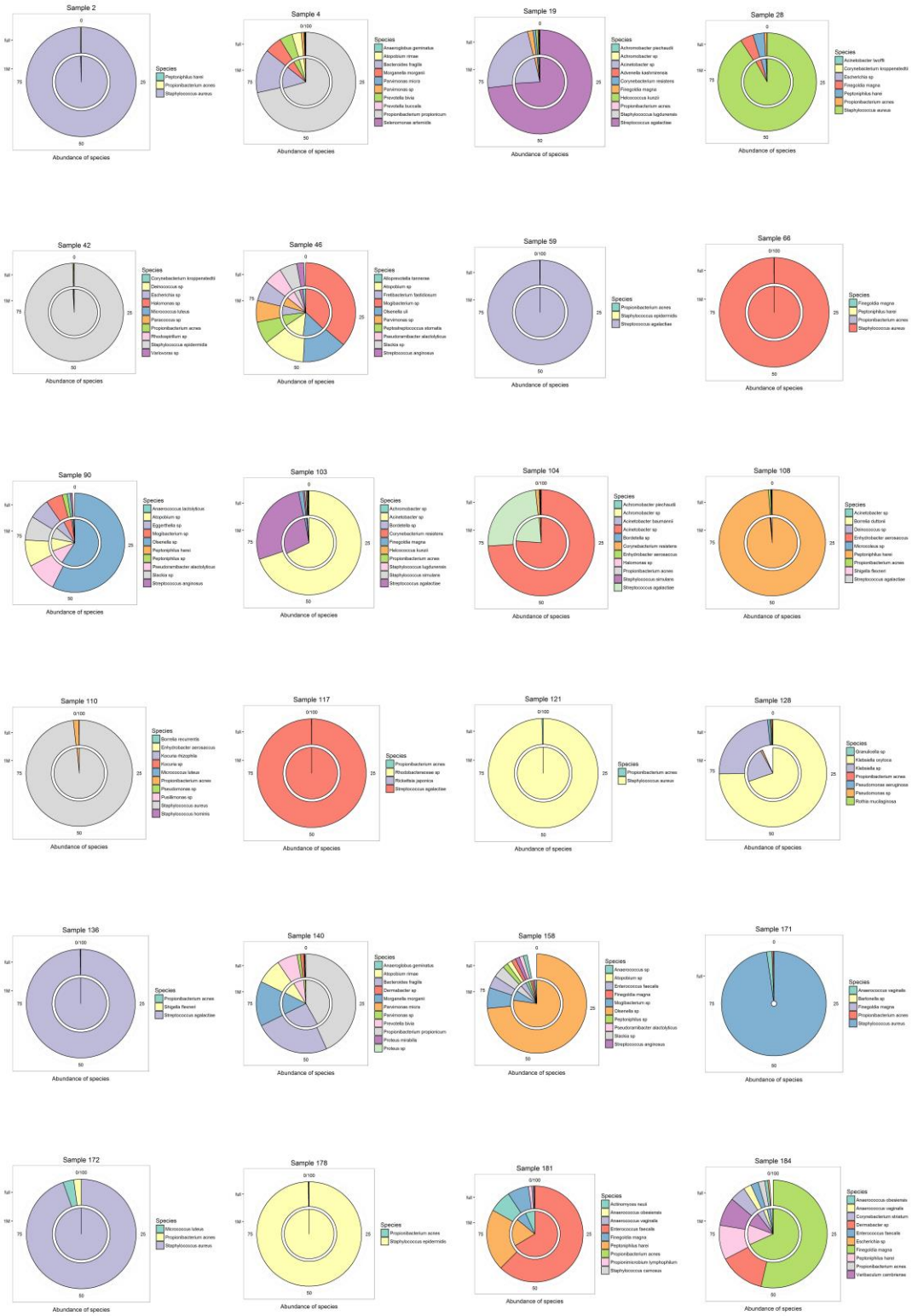
Supplementary Figure 7: Examples of the assessment of a polyclonal population within species. A: The Tablet view of a 5,749 bp contig from *Morganella morganii* in sample 4. B: The dot-plot of the number of SNVs and the size of the contigs from *Morganella morganii* in sample 4 (Pearson's correlation test p<0.001). C: The Tablet view of a 19,205 bp contig from *Staphylococcus aureus* in sample 121. In this specie, the SNVs are concentrated in some regions (such as the one showed) and not homogeneously scattered. D: The dot-plot of the number of SNVs and the size of the contigs from *Staphylococcus aureus* in sample 121 (Pearson's correlation test p=0.4).



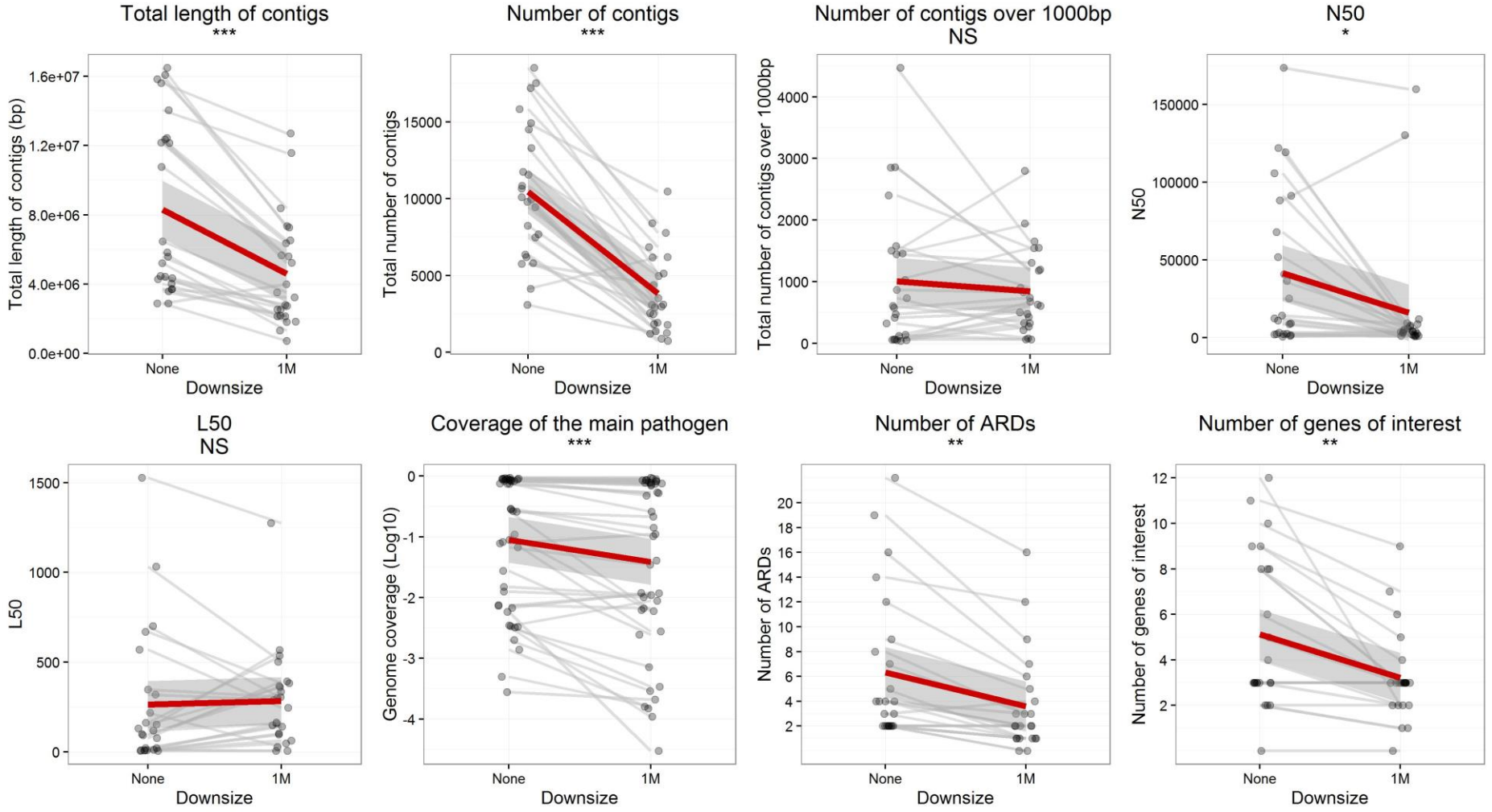
Supplementary Figure 8: Bar plot of the median depth of sequencing (expressed in ×) of the contigs of the bacterial species (in green) and depth of sequencing of the ARDs (in orange) found in the sample.



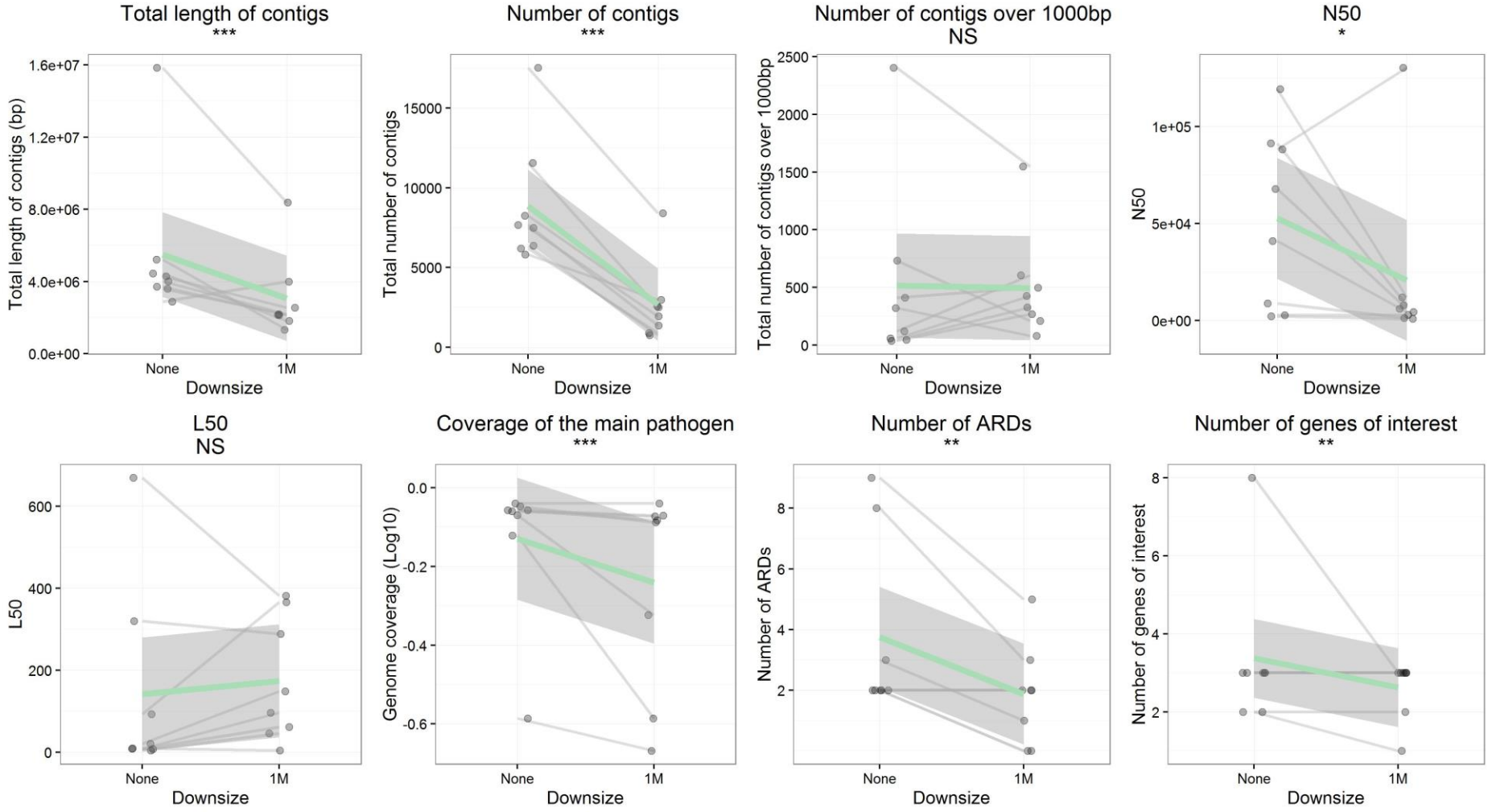
Supplementary Figure 9: Influence of the downsizing to 1M reads on the taxonomic classification of reads by MetaPhlAn2. The outer and the inner circles represent the distribution of the main species with no and 1M downsize, respectively.



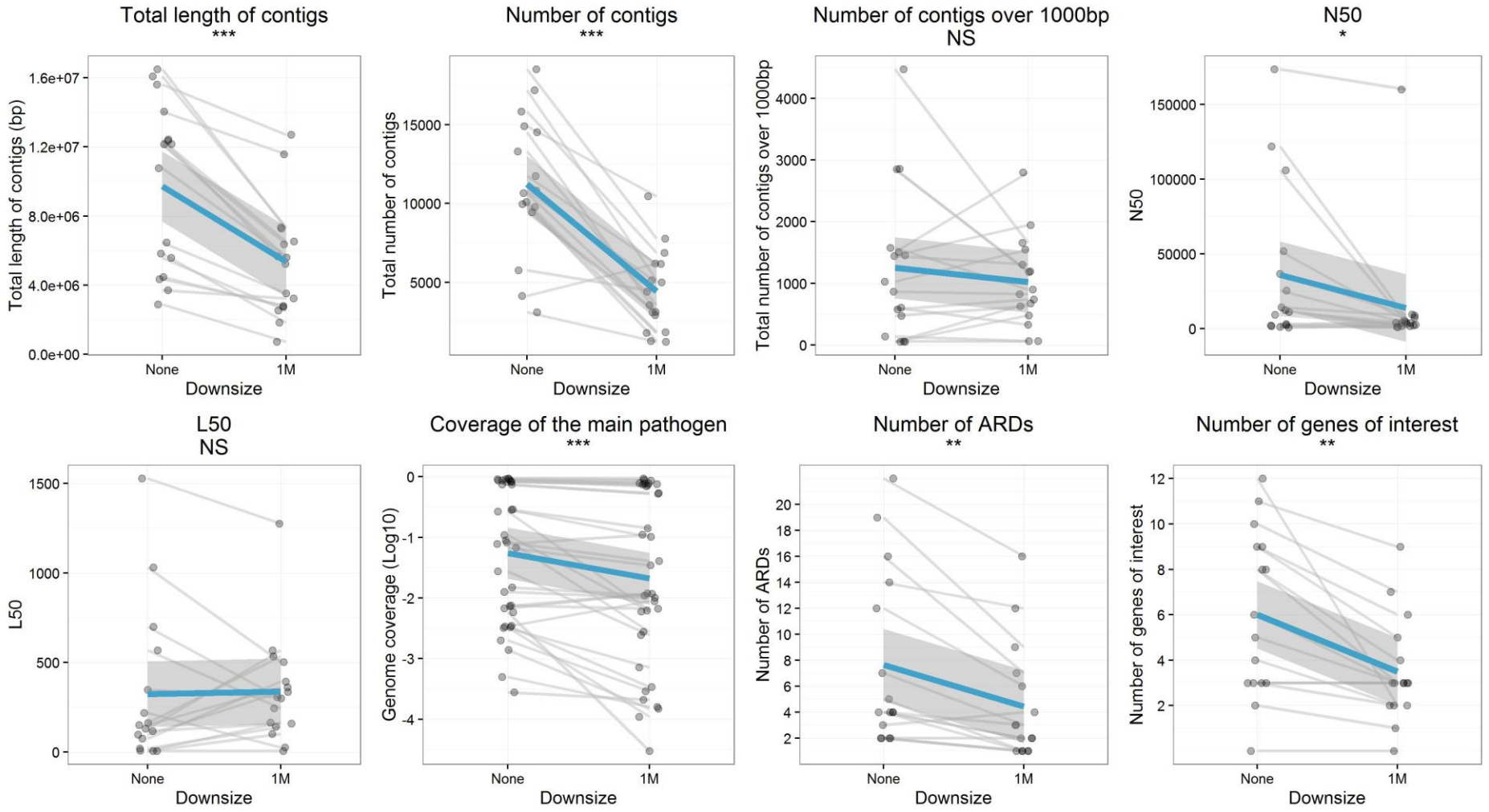
Supplementary Figure 10: Effect of downsizing to 1M reads on the assembly performances for all samples (n=24). Paired t-tests were performed. ARDs: antibiotic resistance determinants. The genes of interest included *gyrA*, *parC* and *rpoB* (with sizes over 80% of the reference genes). ***: p<0.001; **:p<0.01; NS: not significant. The shaded grey area depicts the 95% confidence interval around the linear regression line.



Supplementary Figure 11: Effect of downsizing to 1M reads on the assembly performances for monomicrobial samples (n=8). Paired t-tests were performed. ARDs: antibiotic resistance determinants. The genes of interest included *gyrA*, *parC* and *rpoB* (with sizes over 80% of the reference genes). ***: p<0.001; **:p<0.01; NS: not significant. The shaded grey area depicts the 95% confidence interval around the linear regression line.



Supplementary Figure 12: Effect of downsizing to 1M reads on the assembly performances for polymicrobial samples (n=16). Paired t-tests were performed. ARDs: antibiotic resistance determinants. The genes of interest included *gyrA*, *parC* and *rpoB* (with sizes over 80% of the reference genes). ***: p<0.001; **:p<0.01; NS: not significant. The shaded grey area depicts the 95% confidence interval around the linear regression line.



Supplementary Tables legend

Supplementary Table 1: Summary of the sequencing and assembly results for the 24 samples. N50 is the median contig size of the metagenomic assembly. L50 is the number of contigs that accounts for more than 50% of the metagenomic assembly.

Supplementary Table 2: Inference of antibiotic susceptibility from metagenomic data and antibiotic susceptibility of the bacteria found in culture.