Pretended category: Research article

Title: What does 16S rRNA gene-targeted next generation sequencing contribute to the study of infective endocarditis in valve tissue?

Running title: 16S rRNA gene-targeted NGS in infective endocarditis.

Authors: Paula Santibáñez^a, Aránzazu Portillo^a, Sonia Santibáñez^a, Lara García-Álvarez^a, María de Toro^b, José A. Oteo^{a#}.

^a Center for Rickettsiosis and Arthropod-Borne Diseases, Infectious Diseases Department, San Pedro University Hospital -Center for Biomedical Research from La Rioja (CIBIR), Logroño, Spain.

^b Genomics and Bioinformatics Core Facility, CIBIR, Logroño, Spain.

#Address correspondence to: José A. Oteo, jaoteo@riojasalud.es

Hospital Universitario San Pedro-CIBIR. Infectious Diseases Department. Center of Rickettsiosis and Arthropod-borne diseases.
C/ Piqueras, 98 26006 – Logroño (La Rioja) SPAIN.

Phone: 00 34 941 298993. Fax: 00 34 941 298667. email address: jaoteo@riojasalud.es

Abstract

Infective endocarditis (IE) is a severe and life-threatening disease. Identification of infectious etiology is essential for establishing the appropriate antimicrobial treatment and decreasing mortality. The aim of this study was to explore potential utility of metagenomics for improving microbiological diagnosis of IE. In this work, next-generation sequencing (NGS) of V3-V4 region of the 16S rRNA gene was performed in 27 heart-valve tissues (18 natives, 5 intravascular devices, and 4 prosthetics) of patients diagnosed by IE. Initial microbiological diagnosis, blood culture (BC) and/or PCR, was

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

compared with NGS-based diagnosis. Metagenomics matched with conventional techniques diagnosis in 24/27 cases (88.9%). The same bacterial family was assigned to 24 cases, the same genus to 23 cases, and the same specie for 13 cases. In 22 of them, the etiological agent was represented by percentages >99% of the reads and in two by ~70%. *Staphylococcus aureus* was detected in a previously undiagnosed patient, making the microbiological diagnosis possible in one more sample than with previously used techniques. The remaining two patients showed no coincidence between traditional and NGS microbiological diagnoses. Minority records verified mixed infections in four cases and suggested confectons in two cases, supported by clinical data. In conclusion: 16S rRNA gene-targeted NGS allowed to diagnose one case of IE without microbiological entity based on traditional techniques. However, the application of metagenomics to the study of IE in resected heart valves provides no benefits in comparison with BC and/or PCR. More studies are needed before implementation of NGS for the diagnosis of IE.

Keywords: Infective endocarditis, metagenomics, 16S rRNA gene, NGS, heart valves, Spain.

Text

Introduction

Infective endocarditis (IE) is defined as an infection of a native or prosthetic cardiac valve, endocardial surface, or indwelling cardiac device (1). Despite trends towards earlier diagnosis, pharmacotherapy, and surgical intervention, IE remains a major medical concern with an unchanged incidence (1.7-10 cases/100,000 inhabitants) and

mortality (13%-25% in-hospital, approaching 40% within the first year) in the last two decades (2, 3).

Identification of the causative agent/s is essential for optimal patient management and guides treatment duration and antibiotic choice. Microbiological diagnosis mainly relies on cultured-based methods (4, 5). However, blood cultures are negative in 2.5%-31% of IE cases (3, 6). Causes of blood culture negative endocarditis (BCNE) include sub-optimal specimen collection, antibiotic treatment prior to blood culture collection, and/or infection due to fastidious microorganisms or intra-cellular bacteria (e.g. deficient streptococci, members of the HACEK group, *Propionibacterium acnes*, *Candida* sp., *Brucella* spp., *Legionella* spp., *Listeria monocytogenes*, *Mycoplasma* spp., mycobacteria, *Coxiella burnetii*, *Bartonella* spp., *Tropheryma whipplei*, etc.) (2, 7). These limitations are overcome by molecular biology techniques. Thus, specific and broad range 16S rRNA PCR assays from blood and/or valvular biopsies have proven to be useful for etiologic diagnosis (6, 8, 9). However, heart valves are not usually available, and they are only after surgery, which may delay an early diagnosis.

In the last two decades, next-generation sequencing (NGS) technology has been extensively used in research studies, particularly focusing on the human microbiota and its association with health and disease (10). Continuous improvements of NGS technology as, well as the reduction of costs, have effectively transformed biomedical research (11-13). Nowadays, clinical laboratories evaluate metagenomics as an infectious diseases diagnostic tool for identification of microorganisms directly from patient samples (14).

The aim of this work was to evaluate the contribution of targeted 16S rRNA gene-NGS analysis in heart valve tissues to the study of IE.

Materials and Methods

Samples

DNA from resected heart valves extracted using QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) and stored at -80°C, were retrospectively selected from the CRETAV collection (CIBIR, La Rioja, Spain). They corresponded to 18 native heart valves, five intravascular devices, and four prosthetic heart valves from 27 patients diagnosed with IE according to the modified Duke criteria in our hospital (Hospital Universitario San Pedro, La Rioja, Spain) from 2009 to 2017 (Table 1). Clinical data and results of blood culture (BC) and 16S ribosomal RNA (rRNA) gene PCR (15) and Sanger sequencing from the heart valve tissues were available in all cases (Table 1).

Approval of the regional ethics committee was obtained (Comité Ético de Investigación Clínica-Consejería de Sanidad de La Rioja, Ref. CEICLAR PI-19). Informed consent was obtained from all participants. All procedures were in accordance with the ethical standards of the research committee and with the 1964 Helsinki declaration and its later amendments.

DNA quantification and quality determination

DNA was quantified with a Qubit 3.0 fluorometer (Thermo Scientific) using Qubit dsDNA HS (High Sensitivity) assay kit. The quality of DNA was assessed with the Fragment Analyzer by Agilent formerly Advanced Analytical (AATI) using Genomic DNA 50kb kit. A total of 12.5 ng DNA per sample were added.

Samples were manipulated under sterile conditions in a Class II biosafety cabinet using cycles of UV light prior and between uses to prevent contamination. Sterile single-use instruments were used. DNA extraction, preparation of PCR master mix, and amplification were performed in separate rooms to prevent contamination.

16S rRNA gene amplification, library preparation and sequencing

Primers targeting the hypervariable V3 and V4 regions of 16S rRNA gene were used (16). Amplified regions were purified and indexed with Nextera XT Index kit (Illumina). The library quality was assessed on a Qubit 3.0 Fluorometer (Thermo Scientific) and Fragment Analyzer (AATI) using dsDNA reagent (35-5000bp) kit. Paired-end 300 bp sequences were obtained on an Illumina MiSeq platform.

Sequence processing and analysis

Quality controls of raw reads were carried out with FastQC software <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (17), and trimmed with the Trimmomatic software (18). The V3-V4 amplified region (550-580 bp) was reconstructed through paired reads according the Quantitative Insights Into Microbial Ecology (QIIME) protocol (v1.9.1) (19). Operational Taxonomic Units (OTUs) were defined as sequences with at least 97% similarity versus Greengenes database (20) using UClust clustering algorithm http://drive5.com/usearch/manual/uclust_algo.html (21) and following the open-reference method described by QIIME (22). OTUs with <0.01% relative abundance of the total read counts on a per-sample basis were removed (spurious and chimeric reads). Each OTU's sequences were refined with BLAST tool against GenBank https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome# (23). The causative pathogen was attributed to the microorganism with which the most amplicon reads matched. Low-abundant taxa were also considered relevant when a diagnosis with mixed infections matched epidemiologic, clinical and microbiologic results of patients.

Results

The 16S rDNA gene amplicon sequencing successfully revealed bacterial pathogens in all samples of the cohort (Table 1). A total of 8,733,978 reads [average counts per

sample 323,481; minimum: 67,043 and maximum: 672,563] distributed among 118 OTUs were obtained. The metagenomic analysis of bacteria in hearth valve tissues revealed the same microbiologic diagnosis than conventional techniques (BC and/or PCR) in 24 out of 27 cases. However, the accuracy of the 16S-NGS test did not enable us to give all these results at the level of species. With this technology, it was possible to assign the same bacterial family as with traditional techniques for 24 cases, the same genus for 23 cases, and the same species for 13 cases. For 22 out of these 24 cases that showed consistent results among NGS data and conventional methods, the bacterial taxa with maximum representation were found in a proportion of reads that ranged 99.1-99.9%. However, for two cases (patient IDs #3137 and #3507), around 70% of the reads confirmed previous results (*S. mutans* and *E. faecalis*, respectively) whereas reads at nearly 30% corresponded to *E. faecalis* and *H. parainfluenzae*, respectively. For all these 24 patients, NGS results allowed us to retrospectively confirm the previous diagnosis. In addition, the 16S-NGS test data assessment, led us to corroborate mixed infections within the clinical and epidemiologic context when considering relatively low abundant microorganisms. Thus, for patient ID #1649, whose initial microbiologic diagnosis included *Brucella melitensis* (by BC and PCR) and *Coxiella burnetii* (by indirect immunofluorescence assays and specific PCR) (24), metagenomic results showed *Brucellaceae* as the most relative abundant taxa (99.7%) and, interestingly, *C. burnetii* was also detected at low relative abundance (<1%). For patient ID #1990, conventional methods (BC and PCR) led to an IE diagnosis due to *E. faecalis*, and probably to *S. epidermidis* since this bacterium was obtained from the culture of the tip of the catheter. According to the 16S rRNA gene profiling analysis, 99.6% of reads mapped to *E. faecalis*, and members of the genus *Staphylococcus* were also observed (<1%). For patient ID #2882, an initial diagnosis with IE by *Staphylococcus aureus* was

given according to BC and PCR methods. Furthermore, culture of the resected hearth valve tissue yielded *Escherichia coli*, and bacteremia by *E. faecalis* was detected by BC after surgery. The 16S-NGS work revealed that the majority of the reads (99.6%) corresponded to *S. aureus*, and *E. faecalis* was also present in a lower proportion of reads (<1%). In this case, no sequenced reads mapping to *E. coli* were obtained. Patient ID #3106 was diagnosed with IE by viridans group streptococci based on traditional techniques (BC yielded *Streptococcus intermedius* and *Streptococcus anginosus* was detected by PCR). According to the metagenomic analysis of this specimen, the percentage of sequenced reads observed matching *S. anginosus* reached 99.7%, and other streptococci including *Streptococcus sanguinis*, *Streptococcus cristatus* and *Streptococcus mutans* were found with low relative abundance (<1%).

In addition, two cases were reclassified as mixed infections since sequence reads of uncultured microorganisms were detected and they were subsequently assessed in the clinical-epidemiologic context. NGS test results from patient ID #2730 showed high relative abundance of *S. agalactiae* (99.9%) and low relative abundance of *Coxiellaceae* (<1%). These data corresponded to a patient diagnosed with IE by *S. agalactiae* by BC and PCR, with *C. burnetii* phase II IgG titer of 400 and phase I IgG not detected, with negative PCR results for *C. burnetii*, and positive PCRs for bacteria within the *Coxiellaceae* family. For the other case (patient ID #2969), diagnosed with IE by *T. whipplei* based on PCR results, the NGS analysis showed *T. whipplei* as the most abundant bacteria (99.8%) and <1 % of reads were assigned to *Coxiellaceae*, suggesting a mixed infection supported by serological criteria (*C. burnetii* phase I IgG titer $\geq 1,600$ and phase II IgG titer of 800).

For the sample in which no microbiologic diagnosis had been achieved either by BC or by PCR (patient ID #2208), the analysis of metagenomics sequences evidenced that

most of them (95.1% of relative abundance) corresponded to *S. aureus*. Thus, with our NGS approach we were able to detect the causative agent of IE in one more sample than with the remaining methods.

In contrast, the testing by V3-V4 16S metagenomics yielded taxonomic predictions that differed from those obtained with conventional methods for two cases. Nevertheless, according to our NGS data, the clinical entities found with conventional techniques were also detected in these two samples, although at low relative amounts (patient IDs #2799 and #3042).

Metagenomic results corresponding to the remaining 16 patients not detailed above were consistent with those obtained by traditional techniques, and NGS technology did not provide additional information considered relevant for diagnosis.

Detailed information about data that support the diagnosis of the studied patients is showed in Table 1.

189 Table 1: Contributions of NGS analysis from hearth valve tissue specimens to the diagnosis of infective endocarditis (IE) for the 27 patients of
190 this study.

Patient ID	Initial microbiologic diagnosis before NGS	Studied tissues/device	Identification technique/s for diagnosis	Bacterial taxa with the highest relative abundance and minority findings suspicious of being clinically important by 16S NGS
#1649	<i>Brucella melitensis</i> and <i>Coxiella burnetii</i> ¹	IVD	BC, PCR IFA ¹ , <i>htpAB</i> PCR ¹	<i>Brucellaceae</i> (99.7%) <i>C. burnetii</i> (<1%)
#1990	<i>Enterococcus faecalis</i> (and probably <i>Staphylococcus epidermidis</i> ²)	Native aortic	BC, PCR CTC ²	<i>E. faecalis</i> (99.6%) <i>Staphylococcus</i> spp. (<1%)
#2103	<i>C. burnetii</i>	IVD	PCR	<i>C. burnetii</i> (99.5%)
#2155	<i>Streptococcus bovis</i> group	IVD	BC, PCR	<i>Streptococcus</i> spp. (99.1%)
#2208	No etiology	Native mitral	BC, PCR	<i>Staphylococcus aureus</i> (95.1%)
#2502	<i>S. aureus</i> ³	IVD	PCR, WEC ³	<i>Staphylococcus</i> spp. (99.9%)
#2578	<i>S. bovis</i> group	Native aortic	BC, PCR	<i>Streptococcus</i> spp. (99.1%)
#2622	<i>S. epidermidis</i>	Prosthetic aortic	BC, PCR	<i>Staphylococcus</i> spp. (99.8%)

#2730	<i>Streptococcus agalactiae</i>	Native mitral	BC, PCR	<i>S. agalactiae</i> (99.9%) <i>C. burnetii</i> (<1%)
#2799	<i>E. faecalis</i>	Native mitral	BC, PCR	<i>Streptococcus</i> spp. (26.4%) <i>E. faecalis</i> (15.9%)*
#2882	<i>S. aureus</i> / <i>Escherichia coli</i> ⁴ and bacteremia by <i>E. faecalis</i> ⁵	Native aortic	BC, PCR VC ⁴ BC after surgery ⁵	<i>S. aureus</i> (99.6%) <i>E. faecalis</i> (<1%)
#2969	<i>Tropheryma whipplei</i>	Native aortic	PCR	<i>T. whipplei</i> (99.8%) <i>C. burnetii</i> (<1%)
#2994	<i>Streptococcus milleri</i>	Prosthetic mitral	BC	<i>Streptococcus</i> spp. (99.6%)
#3039	<i>E. faecalis</i>	Native mitral	BC, PCR	<i>E. faecalis</i> (99.9%)
#3041	<i>E. faecalis</i>	Native aortic	BC, PCR	<i>E. faecalis</i> (99.8%)
#3042	<i>Haemophilus parainfluenzae</i>	IVD	BC	<i>Streptococcus</i> spp. (99.2%) <i>H. parainfluenzae</i> (<1%)
#3106	<i>Streptococcus anginosus</i> group (<i>Streptococcus intermedius</i> by	Native mitral	BC, PCR	<i>S. anginosus</i> (99.7%) <i>Streptococcus</i> spp. (<1%)

	BC and <i>S. anginosus</i> by PCR)			
#3112	<i>E. faecalis</i>	Native mitral	BC, PCR	<i>E. faecalis</i> (99.9%)
#3137	<i>Streptococcus mutans</i>	Native mitral	BC, PCR	<i>S. mutans</i> (69.5%) <i>E. faecalis</i> (28.6%)
#3186	<i>Streptococcus mitis</i>	Native aortic	BC, PCR	<i>Streptococcus</i> spp. (99.8%)
#3226	<i>Streptococcus sanguinis</i>	Native mitral	BC, PCR	<i>Streptococcus</i> spp. (99.8%)
#3248	<i>S. mitis</i>	Native mitral	BC, PCR	<i>Streptococcus</i> spp. (99.9%)
#3461	<i>H. parainfluenzae</i>	Native mitral	BC, PCR	<i>H. parainfluenzae</i> (99.5%)
#3471	<i>Streptococcus oralis</i>	Native mitral	BC, PCR	<i>Streptococcus</i> spp. (99.9%)
#3507	<i>E. faecalis</i>	Prosthetic aortic	BC, PCR	<i>E. faecalis</i> (68.8%) <i>H. parainfluenzae</i> (29.7%)
#3511	<i>E. faecalis</i>	Prosthetic aortic	BC, PCR	<i>E. faecalis</i> (99.7%)
#3515	<i>Staphylococcus lugdunensis</i>	Native aortic	BC, PCR	<i>Staphylococcus</i> spp. (99.9%)

191 ID: identification number; IVD: intravascular device; BC: blood culture; PCR: polymerase chain reaction (targeting 16S rRNA gene from hearth
192 valve tissues); ¹, ², ³, ⁴, and ⁵: indicates the sample in which the microorganism was detected; IFA: immunofluorescence assay; *htpAB* PCR: PCR
193 targeting the *htpAB* gene for *C. burnetii*; CTC: catheter tip culture; VC: hearth valve culture; WEC: wound exudate culture. *In this sample, a
194 total of 11 bacterial taxa were detected at relative abundance >1%.

Discussion

IE is still a severe disease with high morbidity and prolonged hospital stay as well as very high mortality during admission and during the 1-year follow-up (3). Therefore, techniques to reliably guide the correct antimicrobial treatment in order to get the sterilization of the affected tissues and decrease the mortality are needed (8). NGS has recently emerged as a comprehensive method for exploring causative agents of infectious diseases without prior culture. However, reports about metagenomic analysis for pathogen identification in heart valve tissues are scarce, and each of them consists of very few patients (25- 34). Here, we used 16S rRNA targeted-NGS from heart valve tissues as an approach to the diagnosis of IE in a cohort of 27 patients.

Herein, metagenomics allowed the microbiological diagnosis (*S. aureus*) in patient #2208, in which the causative agent had not been detected either by PCR or by HC. In addition, the same microbiological diagnosis was obtained using NGS or routine techniques for 24 patients (88.9%), although at higher taxonomical level for 11 of them. Taxonomic assignment of sequence reads below the genus level is a challenge in metagenomics data analysis. Only the combination of multiple hypervariable regions or the nearly complete sequence of 16S rRNA gene gives accurate measures of taxonomic diversity (35). Third-generation sequencing provides long-read sequences, but high base-calling error rates (36). Besides, consensus in current NGS protocols is essential since microbiome studies are potentially biased at every methodological stages, from sampling to bioinformatic analysis (37, 38).

The interpretation of the metagenomic results for IE cases has been based on considering the bacteria represented by the highest percentage of reads as the causative agent (26, 30, 32). In our series, this was possible for 22 patients with values from 99.1% to 99.9% of the reads and in 1 patient (patient ID #2208) with 95.1% of the

reads. However, we found two cases (patient IDs #3137 and #3507) in which the most represented bacteria matched to the initial causative agent, but the proportion reached 70% of the reads, and the remaining ones corresponded to *E. faecalis* and *H. parainfluenzae*, respectively. The biological significance of reads at around 30% of relative abundance is currently unknown. Moreover, in patient ID #2799 the highest proportion of reads was as low as 26.4% and mapped to *Streptococcus* spp., and other 10 bacterial taxa were detected at relative abundance >1% (Supplementary data).

One main advantage of 16S rRNA-NGS technology is its capacity to classify all bacteria from a sample without intermediate culturing steps (14). Minority findings constitute a concern for 16S rRNA gene NGS projects. The cutoff value indicating how many reads of a microorganism in a sample are no relevant for the analysis has not been established for microbiome studies. It will require a vast experience to establish which are spurious reagent contaminants, sample processing contaminants, cross-contamination in multiplexed libraries, etc. or true infections or coinfections. Regarding to the relatively low abundant microorganisms, we have described four cases of previously characterized mixed infections (patient IDs #1649, #1990, #2882 and #3106) and two cases reclassified as mixed infections without clinical suspicion, but in agreement with serological results (patient IDs #2969 and #2730). However, there were two cases in which the microorganism formerly considered as causative agent by HC and/or PCR was barely represented in metagenomic results (patients IDs #2799 and #3042). In order to understand this, it is important to take into consideration that, as was previously mentioned, each step of NGS analysis influences the relative abundances observed (36-38). However, it cannot be discarded that these small percentages are inherent failures of the technique. The concordance of minority percentages with clinical data gave value to the diagnosis of mixed infectious in the six cases mentioned

above (patient IDs #1649, #1990, #2882, #3106 #2969 and #2730), in contrast to patient ID #3042, in which no clinical or epidemiological data available supported streptococcal infection, and patient ID #2799, in which 11 bacterial taxa were found. Addressing laboratory contamination is an urgent task and it is important to scrutinize NGS data with an understanding of its potential for false positive results. Bacterial identification using 16S rRNA-NGS may be biased because of unequal amplification of certain species, and it is influenced by several factors, such as the region(s) sequenced, amplification efficiency, sequencing technology, and bioinformatics workflow(s). Including spiking-in mock microorganisms that provide comparable results across research groups and time, as well as positive and negative controls in order to assure the quality of the NGS results, is recommended (36, 38).

NGS technology has been suggested as an essential supplement to culture-based methods for the diagnosis of IE, particularly for unculturable or difficult-to-culture microorganisms (25, 26 31, 32). However, in our three BCNE cases, the microbiological diagnosis had already been achieved by 16S rRNA PCR (patient IDs #2103, #2502 and #2969). Therefore, the application of NGS techniques may not always be valuable, even more considering that is more expensive and time-consuming, and requires equipment hardly affordable by most clinical laboratories and personnel trained in bioinformatics analysis.

Anyway, this is so far one of the largest series published, and the concordance of our results with the previous microbiological diagnoses in almost all patients highlights the importance of this work.

Conclusions

Results of 16S rRNA-NGS are mostly consistent with those of BC and/or conventional PCR but do not improve the diagnosis of IE cases. Metagenomics may be helpful to IE patients after valve replacement surgery, especially when conventional tests fail to yield a diagnosis. Moreover, minority findings supported by clinical data could suggest mixed infections not previously suspected, although more efforts should be made in order to understand them. Hence, further studies are required to validate the clinical usefulness of this method.

Acknowledgments

To María Bea, laboratory technician from the Genomics and Bioinformatics Core Facility (CIBIR, La Rioja, Spain).

Partial results were presented at VII Congreso SEICAV (Sociedad Española de Infecciones Cardiovasculares), Sevilla, 2018.

Transparency declaration

The authors have no conflicts of interest to disclose. No specific external funding was available for this study.

References

1. Cahill TJ, Prendergast BD. Infective endocarditis. Lancet 2016; 387:882-93.
2. Murdoch DR, Corey GR, Hoen B, Miró JM, Fowler VG Jr, Bayer AS, Karchmer AW, Olaison L, Pappas PA, Moreillon P, Chambers ST, Chu VH, Falcó V, Holland DJ, Jones P, Klein JL, Raymond NJ, Read KM, Tripodi MF, Utili R, Wang A, Woods CW, Cabell CH; International Collaboration on Endocarditis-Prospective Cohort Study (ICE-PCS) Investigators. 2009. Clinical presentation, etiology, and

outcome of infective endocarditis in the 21st century: the International
Collaboration on Endocarditis-Pro prospective Cohort Study. *Arch Intern Med* 169:
463–473.

3. Muñoz P, Kestler M, De Alarcon A, Miro JM, Bermejo J, Rodríguez-Abella H,
Fariñas MC, Cobo Belaustegui M, Mestres C, Llinares P, Goenaga M, Navas E,
Oteo JA, Tarabini P, Bouza E; Spanish Collaboration on Endocarditis-Grupo de
Apoyo al Manejo de la Endocarditis Infecciosa en España (GAMES) (see
Acknowledgment). 2015. Current epidemiology and outcome of infective
endocarditis: a multicenter, prospective. Cohort Study *Med (Baltimore)* 94:e1816.

4. Baddour LM, Wilson WR, Bayer AS, Fowler VG Jr, Tleyjeh IM, Rybak MJ, Barsic
B, Lockhart PB, Gewitz MH, Levison ME, Bolger AF, Steckelberg JM, Baltimore
RS, Fink AM, O'Gara P, Taubert KA; American Heart Association Committee on
Rheumatic Fever, Endocarditis, and Kawasaki Disease of the Council on
Cardiovascular Disease in the Young, Council on Clinical Cardiology, Council on
Cardiovascular Surgery and Anesthesia, and Stroke Council. 2015. Infective
Endocarditis in Adults: Diagnosis, Antimicrobial Therapy, and Management of
Complications: A Scientific Statement for Healthcare Professionals From the
American Heart Association. *Circulation* 13; 132(15):1435-1486.

5. Habib G, Lancellotti P, Antunes MJ, Bongiorni MG, Casalta JP, Del Zotti F,
Dulgheru R, El Khoury G, Erba PA, Iung B, Miro JM, Mulder BJ, Plonska-
Gosciniak E, Price S, Roos-Hesselink J, Snygg-Martin U, Thuny F, Tornos Mas P,
Vilacosta I, Zamorano JL; ESC Scientific Document Group. 2015. 2015 ESC
Guidelines for the management of infective endocarditis: The Task Force for the
Management of Infective Endocarditis of the European Society of Cardiology
(ESC). Endorsed by: European Association for Cardio-Thoracic Surgery (EACTS),

the European Association of Nuclear Medicine (EANM). Eur Heart J 21;
36(44):3075-3128.

6. Fournier PE, Gourié F, Casalta JP, Lepidi H, Chaudet H, Thuny F, Collart F,
Habib G, Raoult D. 2017. Blood culture-negative endocarditis: Improving the
diagnostic yield using new diagnostic tools. Medicine (Baltimore) 96(47): e8392.

7. Tattévin P, Watt G, Revest M, Arvieux C, Fournier PE. Update on blood culture-
negative endocarditis. Med Mal Infect. 2015; 45(1-2): 1-8.

8. Marín M, Muñoz P, Sánchez M, Del Rosal M, Alcalá L, Rodríguez-Créixems M,
Bouza E; Group for the Management of Infective Endocarditis of the Gregorio
Marañón Hospital (GAME). 2007. Molecular diagnosis of infective endocarditis by
real-time broad-range polymerase chain reaction (PCR) and sequencing directly
from heart valve tissue. Medicine (Baltimore) 86(4):195-202.

9. García-Álvarez L, Sanz MM, Marín M, Fariñas M, Montejo M, Goikoetxea J,
Rodríguez García R, de Alarcón A, Almela M, Fernández-Hidalgo N, Alonso Socas
Mdel M, Goenaga MÁ, Navas E, Vicioso L, Oteo JA; Spanish Collaboration on
Endocarditis-Grupo de Apoyo al Manejo de la Endocarditis Infecciosa en España
(GAMES). 2016. *Tropheryma whippelii* endocarditis in Spain: Case reports of 17
prospective cases. Medicine (Baltimore) 95(26):e4058.

10. Sirisinha S. 2016. The potential impact of gut microbiota on your health: Current
status and future challenges. Asian Pac J Allergy Immunol 34(4):249-264

11. Villanueva-Millán MJ, Pérez-Matute P, Oteo JA. 2015. Gut microbiota: a key
player in health and disease. A review focused on obesity. J Physiol Biochem
71(3): 509-525.

12. Portillo A, Palomar AM, de Toro M, Santibáñez S, Santibáñez P, Oteo JA. 2019. Exploring the bacteriome in anthropophilic ticks: To investigate the vectors for diagnosis. *PLoS One* 19; 14 (3):e0213384
13. Wilson MR, Sample HA, Zorn KC, Arevalo S, Yu G, Neuhaus J, Federman S, Stryke D, Briggs B, Langelier C, Berger A, Douglas V, Josephson SA, Chow FC, Fulton BD, DeRisi JL, Gelfand JM, Naccache SN, Bender J, Dien Bard J, Murkey J, Carlson M, Vespa PM, Vijayan T, Allyn PR, Campeau S, Humphries RM, Klausner JD, Ganzon CD, Memar F, Ocampo NA, Zimmermann LL, Cohen SH, Polage CR, DeBiasi RL, Haller B, Dallas R, Maron G, Hayden R, Messacar K, Dominguez SR, Miller S, Chiu CY. 2019. Clinical Metagenomic Sequencing for Diagnosis of Meningitis and Encephalitis. *N Engl J Med* 13; 380(24):2327-2340.
14. Miller S, Chiu C, Rodino KG, Miller MB. 2020. Point-Counterpoint: Should we be performing metagenomic next-generation sequencing for infectious disease diagnosis in the clinical laboratory? *J Clin Microbiol* 24; 58(3):e01739-19.
15. Weisburg WG, Barns SM, Pelletier DA, Lane DJ. 1991. 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* 173(2): 697-703.
16. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glöckner FO. 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 41:e1.
17. Babraham Bioinformatics—FastQC A quality control tool for high throughput sequence data. [http:// www.bioinformatics.babraham.ac.uk/projects/fastqc/](http://www.bioinformatics.babraham.ac.uk/projects/fastqc/)
18. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* 30:2114–2120.
19. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D,

Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336.

20. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimerachecked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol.* 72:5069–5072.

21. http://drive5.com/usearch/manual/uclust_algo.html.

22. Rideout JR, He Y, Navas-Molina JA, Walters WA, Ursell LK, Gibbons SM, Chase J, McDonald D, Gonzalez A, Robbins-Pianka A, Clemente JC, Gilbert JA, Huse SM, Zhou HW, Knight R, Caporaso JG. 2014. Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *Peer J* 2:e545.

23. https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome 20. metagenomeS

24. Oteo JA, Pérez-Cortés S, Santibáñez P, Gutiérrez E, Portillo A, Blanco JR, de Alarcón A. 2012. Q fever endocarditis associated with a cardiovascular implantable electronic device. *Clin Microbiol Infect* 18(11): E482-484.

25. Imai A, Gotoh K, Asano Y, Yamada N, Motooka D, Fukushima M, Kanzaki M, Ohtani T, Sakata Y, Nishi H, Toda K, Sawa Y, Komuro I, Horii T, Iida T, Nakamura S, Takashima S. 2014. Comprehensive metagenomic approach for detecting causative microorganisms in culture-negative infective endocarditis. *Int J Cardiol* 15; 172 (2): e288-289.

26. Fukui Y, Aoki K, Okuma S, Sato T, Ishii Y, Tateda K. 2015. Metagenomic analysis for detecting pathogens in culture-negative infective endocarditis. *J Infect Chemother*. 21(12): 882-884.
27. Oberbach A, Schlichting N, Feder S, Lehmann S, Kullnick Y, Buschmann T, Blumert C, Horn F, Neuhaus J, Neujahr R, Bagaev E, Hagl C, Pichlmaier M, Rodloff AC, Gräber S, Kirsch K, Sandri M, Kumbhari V, Behzadi A, Behzadi A, Correia JC, Mohr FW, Friedrich M. 2017. New insights into valve-related intramural and intracellular bacterial diversity in infective endocarditis. *PLoS One* 14; 12(4): e0175569.
28. Cheng J, Hu H, Kang Y, Chen W, Fang W, Wang K, Zhang Q, Fu A, Zhou S, Cheng C, Cao Q, Wang F, Lee S, Zhou Z. 2018. Identification of pathogens in culture-negative infective endocarditis cases by metagenomic analysis. *Ann Clin Microbiol Antimicrob* 20; 17(1):43.
29. Chan WS, Au CH, Leung HC, Ho DN, Li D, Chan TL, Lam TW, Ma ES, Tang BS. 2019. Potential utility of metagenomic sequencing for improving etiologic diagnosis of infective endocarditis. *Future Cardiol* 15(6):411-424.
30. Choutko V, Lazarevic V, Gaïa N, Girard M, Renzi G, Leo S, Keller PM, Huber C, Schrenzel J. 2019. Rare Case of Community-Acquired Endocarditis Caused by *Neisseria meningitidis* Assessed by Clinical Metagenomics. *Front Cardiovasc Med* 6; 6:112.
31. Cheng J, Hu H, Fang W, Shi D, Liang C, Sun Y, Gao G, Wang H, Zhang Q, Wang L, Wu H, Hu L, Chen L, Zhang J, Lee S, Wang F, Zhou Z. 2019. Detection of pathogens from resected heart valves of patients with infective endocarditis by next-generation sequencing. *Int J Infect Dis* 83: 148-153.

32. Kolb M, Lazarevic V, Emonet S, Calmy A, Girard M, Gaïa N, Charretier Y, Cherkaoui A, Keller P, Huber C, Schrenzel J. 2019. Next-Generation Sequencing for the Diagnosis of Challenging Culture-Negative Endocarditis. *Front Med (Lausanne)*. 20; 6:203.
33. Million M, Gaudin M, Melenotte C, Chasson L, Edouard S, Verdonk C, Prudent E, Amphoux B, Meresse S, Dorent R, Lepidi H, La Scola B, Gorvel JP, Desnues C, Raoult D. 2020. Metagenomic Analysis of Microdissected Valvular Tissue for Etiologic Diagnosis of Blood Culture Negative Endocarditis. *Clin Infect Dis* 23; 70(11):2405-2412.
34. Ai JW, Liu H, Li HX, Ling QX, Ai YQ, Sun SJ, Wang X, Zhang BY, Zheng JM, Jin JL, Zhang WH. 2020. Precise diagnosis of *Neisseria macacae* infective endocarditis assisted by nanopore sequencing. *Emerg Microbes Infect* 9(1):1864-1868.
35. Yarza P, Yilmaz P, Priesse E, Glöckner FO, Ludwig W, Schleifer KH, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12(9): 635-645
36. Boers SA, Jansen R, Hays JP. 2019. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. *Eur J Clin Microbiol Infect Dis* 38:1059–1070.
37. Pollock J, Glendinning L, Wisedchanwet T, Watson M. 2018. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. *Appl Environ Microbiol* 19; 84(7): e02627-17.

- 439 38. Forbes JD, Knox NC, Peterson CL, Reimer AR. 2018. Highlighting Clinical
440 Metagenomics for Enhanced Diagnostic Decision-making: A Step Towards Wider
441 Implementation. Comput Struct Biotechnol J 27; 16: 108-120.

Supplementary data

Bacteria	IE agent	Clinical importance	Reference
<i>Abiotrophia defectiva</i>	<1%	Normal flora of the oral cavity, the urogenital and the intestinal tracts. It has been involved in ocular, osteoarticular, arthroplasty and sinus infections, otitis, cerebral abscesses, iatrogenic meningitis and pancreatic abscesses.	Birlutiu et al., 2017
<i>Actinomyces</i> spp.	30 cases since 1939	Infections are usually limited to cervicofacial, thoracic, and abdominopelvic regions.	Phichaphop et al., 2020
<i>Bergeyella zoohelcum</i>	rare	It is a well-known zoonotic pathogen causing cellulitis, leg abscess, tenosynovitis, septicemia, pneumonia and meningitis.	Pan et al., 2020
<i>Capnocytophaga sputigena</i>	yes	Normal flora of the oral cavity. Has occasionally been associated with periodontitis, sinusitis, conjunctivitis, pleural empyemas, subphrenic abscesses, osteomyelitis and septicemia.	Torru et al., 1995
<i>Corynebacterium</i> spp.	yes	Normal skin flora. Contaminate blood cultures and cause opportunistic severe infections.	Rasmussen et al., 2020
<i>Enterococcus faecalis</i>	10% (Enterococci)	Part of the human gastrointestinal flora. Cause urinary, intra-abdominal, wound and pelvic infections and bacteremia. It is the third most common cause of IE.	Dahl et al., 2013
<i>Haemophilus parainfluenzae</i>	1.2 - 3% (HACEK group)	Frequently colonize the oropharynx and nasopharynx and may also be part of the genital tract microbiota. Related to infections of the respiratory tract although brain abscess, surgical site infections, soft tissue infections, prosthetic joint infections, and hepatic and biliary tract infections have also been documented. Most common cause of	Revest et al., 2016

		HACEK endocarditis.	
<i>Lautropia mirabilis</i>	no	Commonly found in the human oral cavity and the upper respiratory tract. Described in HIV-positive children and in the sputum of one cystic fibrosis patient.	Rossmann et al., 1998; Voronina et al., 2018
<i>Neisseria elongata</i>	very rare	Normal oropharyngeal bacterial flora. Associated with septicemia and osteomyelitis.	Youssef et al., 2019
<i>Rothia aeria</i>	9 cases	Colonizes the human oral cavity and upper gastrointestinal tract. Sporadically related as cause of neck abscess, shoulder joint infection, lung infection and neonatal sepsis.	Kim et al., 2014
<i>Streptococcus</i> spp.	30%	Streptococci are found in almost every location in the human body and are the dominant species in the human oral cavity and upper respiratory tract. They are the second leading cause of IE.	Kim et al., 2018

Supplementary References

1. Birlutiu V, Birlutiu RM. 2017. Endocarditis due to *Abiotrophia defectiva*, a biofilm-related infection associated with the presence of fixed braces. *Medicine* 96:e8756.
2. Phichaphop C, Apiwattanakul N, Wanitkun S, Boonsathorn S. 2020. Bacterial Endocarditis Caused by *Actinomyces oris*: First Reported Case and Literature Review. *J Investig Med High Impact Case Rep* 8:2324709620910645.
3. Pan H, Li W, Sun E, Zhang Y. 2020. Characterization and whole genome sequencing of a novel strain of *Bergeyella cardium* related to infective endocarditis. *BMC Microbiol* 12; 20(1):32.
4. Torrús D, González JA, Boix V, Portilla J. 1995. Endocarditis aórtica por *Capnocytophaga sputigena* [Aortic endocarditis due to *Capnocytophaga sputigena*]. *Med Clin (Barc)* 17; 105(3):119. Spanish.
5. Rasmussen M, Mohlin AW, Nilson B. 2020. From contamination to infective endocarditis-a population-based retrospective study of *Corynebacterium* isolated from blood cultures. *Eur J Clin Microbiol Infect Dis* 39(1):113-119.
6. Dahl A, Bruun NE. 2013. *Enterococcus faecalis* infective endocarditis: focus on clinical aspects. *Expert Rev Cardiovasc Ther* 11(9):1247-1257.
7. Revest M, Egmann G, Cattoir V, Tattevin P. 2016. HACEK endocarditis: state-of-the-art. *Expert Rev Anti Infect Ther* 14(5):523-530.
8. Rossmann SN, Wilson PH, Hicks J, Carter B, Cron SG, Simon C, Flaitz CM, Demmler GJ, Shearer WT, Kline MW. 1998. Isolation of *Lautropia*

468 *mirabilis* from oral cavities of human immunodeficiency virus-infected
469 children. J Clin Microbiol 36(6):1756-1760.

470 9. Voronina OL, Kunda MS, Ryzhova NN, Aksenova EI, Sharapova NE,
471 Semenov AN, Amelina EL, Chuchalin AG, Gintsburg AL. 2018. On
472 Burkholderiales order microorganisms and cystic fibrosis in Russia.
473 BMC Genomics 19(Suppl 3):74.

474 10. Youssef D, Marroush TS, Levine MT, Sharma M. 2019. Endocarditis due
475 to *Neisseria elongata*: A case report and review of the literature. Germs
476 2; 9(4):188-192.

477 11. Kim UJ, Won EJ, Kim JE, Jang MO, Kang SJ, Jang HC, Park KH, Jung
478 SI, Shin JH. 2014. *Rothia aeria* infective endocarditis: a first case in
479 Korea and literature review. Ann Lab Med 34(4):317-320.

480 12. Kim SL, Gordon SM, Shrestha NK. 2018. Distribution of streptococcal
481 groups causing infective endocarditis: a descriptive study. Diagn
482 Microbiol Infect Dis 91(3):269-272.