XLIX CONGRESSO NAZIONALE AMCLI

26 febbraio - 1 marzo 2022 | Palacongressi Rimini

Aspetti tecnici e procedurali per una buona sperimentazione sul microbioma

Technical and procedural aspects for a good experimentation on the microbiome

SESSIONE 6-MICROBIOTA COME ORGANO-SISTEMA: DAL LABORATORIO ALLA CLINICA

Simona Panelli

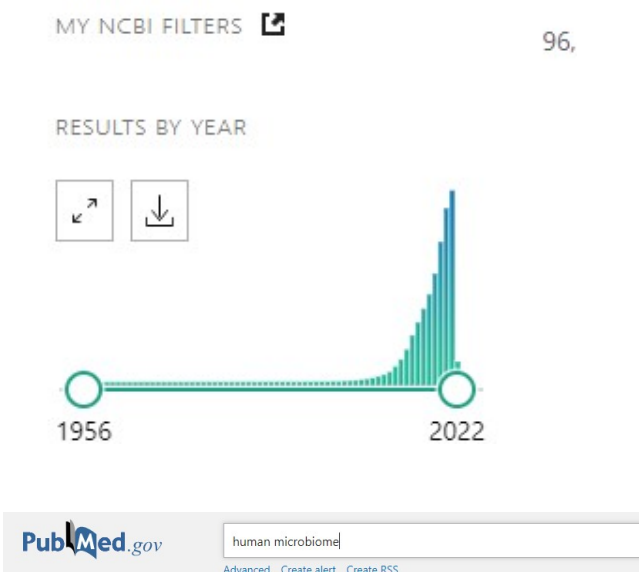
Centro di Ricerca Pediatrica «Romeo ed Enrica Invernizzi», Dipartimento e Scienze Biomediche e Cliniche «L. Sacco», Università di Milano

Università degli Studi di Milano, ***simona.panelli1@unimi.it***

In the last decades a plethora of publications have used DNA- and RNA-based profiling methods to interrogate microbial communities in locations ranging from ice columns in the remote arctic to the human body.

The human microbiota comprises the microorganisms, with their genes and metabolites, colonizing (in varying compositions) different body niches, such as mouth, genitals, skin, airways and the gut

Disruption of the balanced interaction of microbes with the host (dysbiosis) is associated with a plethora of pathologies: hence the great medical interest in clarifying community composition and functions to open perspectives of applications in various areas, notably therapeutics



The gastrointestinal tract is the most studied microbial ecosystem within the human host, due to its large volume, relevance to health and disease, and high diversity:

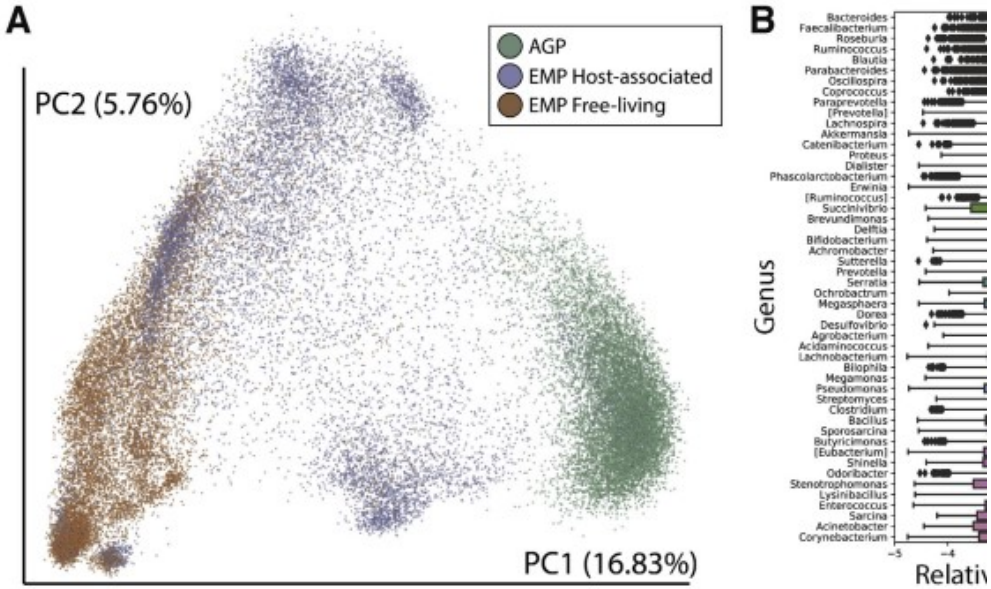


Figure 1. Intersubject variability of the gut microbiome. (*A*) A principal coordinates plot of unweighted computed using the Earth Microbiome Project (EMP) data set[24] and the fecal samples from the America

Clinical Gastroenterology and Hep

Microbiome 101: Studying, Analyzing, and Interpreting G Microbiome Data for Clinicians

Celeste Allaband,[*] Daniel McDonald,[‡] Yoshiki Vázquez-Baeza,[‡] Jeremiah J. Minich,[§] Anupriya Tripathi,[‡] David A. Brenner,[¶] Rohit Loomba,[¶,] Larry Smarr,[#,**,‡‡] William J. Sandborn,[¶,§§] Bernd Schnabl,[¶,#,§§] Pieter Dorres Amir Zarrinpar,[¶,#,§§] and Rob Knight[‡,#,**,¶¶]

The diversity detectable in human stool (AGP, American Gut Project) is comparable with the diversity spanned by completely different kinds of environments in the Earth Microbiome Project (EMP)

XLIX CONGRESSO NAZIONALE AMCLI

26 febbraio - 1 marzo 2022 | Palacongressi Rimini

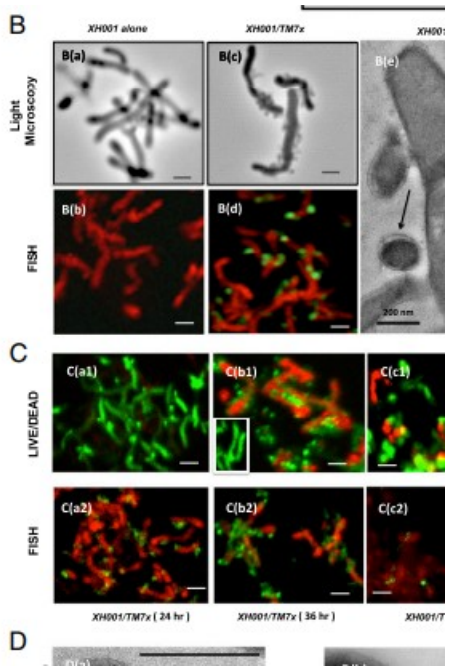
… And new pieces in this puzzle are continuously discovered….

PNAS

Cultivation of a human-associated TM7 phylo reveals a reduced genome and epibiotic parasitic lifestyle

Xuesong He[a,1], Jeffrey S. McLean[b,c,1,2], Anna Edlund[a,c], Shibu Yooseph[c], Adam P. Hall[c], Su-Yang Liu Pieter C. Dorrestein[e], Eduardo Esquenazi[f], Ryan C. Hunter[g], Genhong Cheng[d], Karen E. Nelson[c], Re and Wenyuan Shi[a,2]

[a]Section of Oral Biology, School of Dentistry, and [d]Department of Microbiology, Immunology, and Molecular Genetics, Universit CA 90095; [b]Department of Periodontics, University of Washington, Seattle, WA 98195; [c]Microbial and Environmental Genomics San Diego, CA 92037; [e]Departments of Chemistry and Biochemistry and Pharmacology, Skaggs School of Pharmacy & Pharmace



Significance

TM7 is one of the most enigmatic bacterial phyla uncultivated candidate phyla referred to as "mic matter," and it has potential pathogenic associ revealed molecular insights into its uncultivability genicity, as well its unique epibiotic and parasi phases. These novel discoveries shed significant li biological, ecological, and medical importance of T as providing useful information for culturing othe

Cell Reports

Acquisition and Adaptation of Ultra-small Pa

Journal of Microbiology (2020) Vol. 58, No. 5, pp. 377–386
DOI 10.1007/s12275-020-9365-3

Full-repertoire comparison of the microscopic objects compo human gut microbiome with sequenced and cultured comm

Edmond Kuete Yimagou[1], Jean-Pierre Baudoin[1], Rita Abou Abdallah[1], Fabrizio Di Pinto[1], Jacques Yaacoub Bou Khalil[2*], and Didier Raoult[1*]

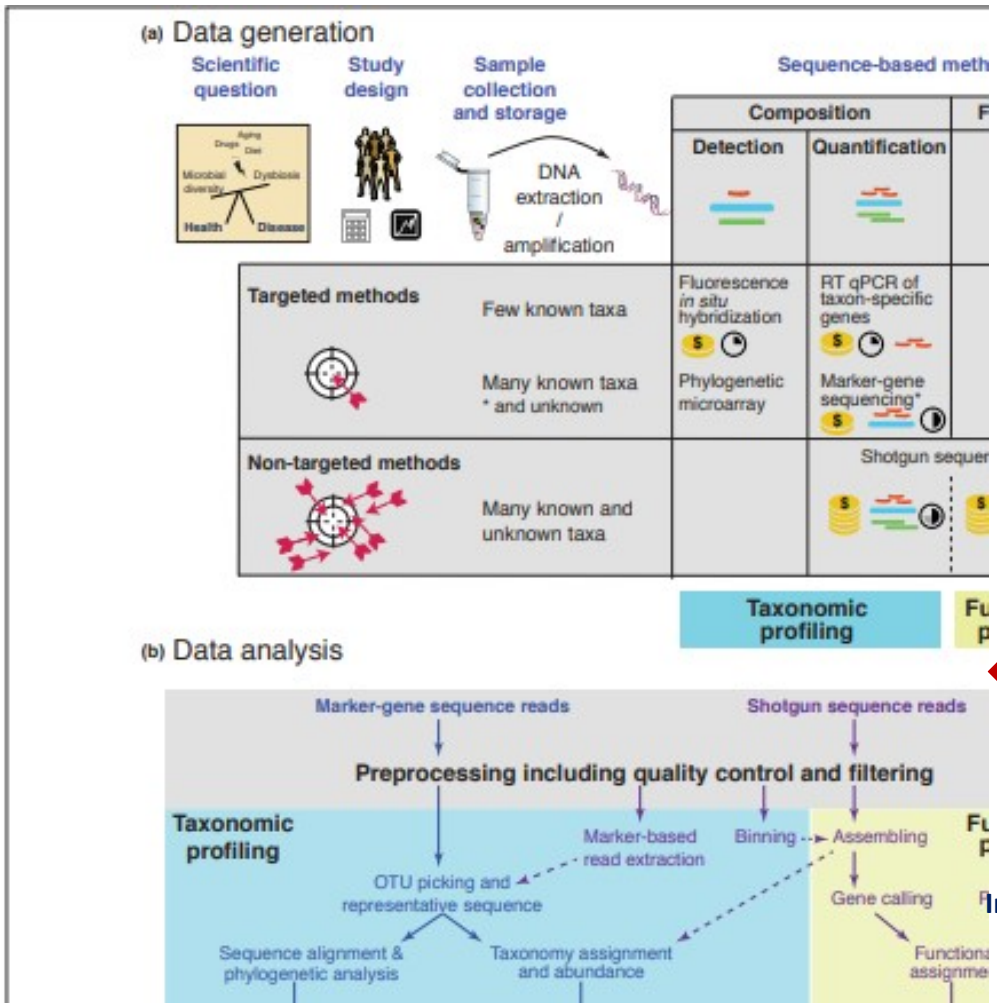
[1]Aix Marseille Univ., IRD, AP-HM, MEPHI, VITROME,

components remains a challenging objec 2012; Hugon *et al.*, 2013, 2017). The human sists of the microorganisms living in or or (Whitman *et al.*, 1998; Lloyd-Price *et al.* ganisms are especially abundant in the hu composition of the gut microbiome varies conditions, such as age, geographical origin

Conducting a robust experiment is not trivial in microbiota research:

«A metagenomics study includes multiple consecutive steps and requires careful attention to the study design, the choice of methods, and rigor of study»

«Unlike [other technologies] that have already converged toward standardized methods and procedures, leading to reproducibility and comparability between studies, metagenomics has not yet reached such level of maturity»



Overview of strategies for interrogating the microbiota.

Amplicon (taxonomic profiling)
Vs
Shotgun (taxonomic + functional profiling)
metagenomics

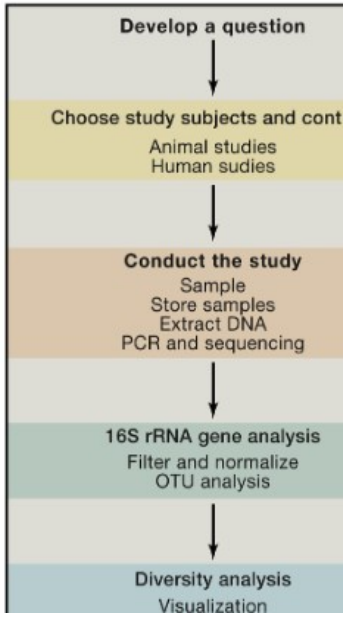
In this presentation, I will focus on amplicon metagenomics, or Marker gene analysis

Therefore, what should we worry about when planning a study of amplicon metagenomics?

1-Study design and scientific questions

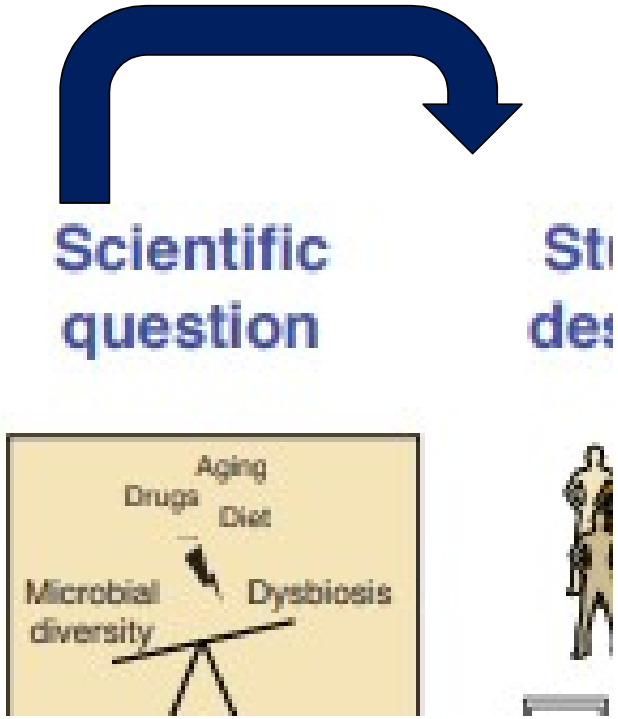
2-Data generation: considerations on the experimental flow

3-Data analysis



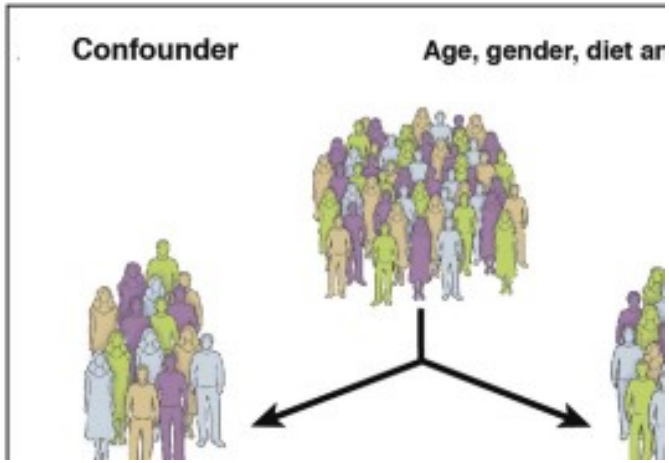
1-Study design and scientific questions

A reliable microbiome study depends on defining clear scientific questions and objectives that drive the design and the choice of the experimental and computational strategies



Typical scientific questions, such as those addressed in case-control and longitudinal or intervention studies, can all be addresed in microbiome studies

HOWEVER: many confounding factors, if not controlled, can obscure patterns in microbiome data:



In case-control studies potential confounders are: heterogeneity of clinical phenotypes among patients, age, body mass index, pregnancy, gender, diet, medications, other lifestyle factors

Longitudinal studies control these confounding factors

50
AMCLI
associazione
microbiologi
clinici italiani
1970 – 2020

Key points:

1-CASE-CONTROL STUDIES:

CASES: try to recruit a «diseased» population as homogeneous as possible: the heterogeneity of clinical phenotypes dilutes the statistical estimates of effect sizes of the microbiome.

Example. in a study examining the role of the microbiota in the development of pneumonia: focus on the most common type rather than trying to profile all the different etiological types.

HOW TO ACHIEVE THIS?

Specify well-defined inclusion and exclusion criteria, e.g., based on respiratory cultures

CONTROLS: their choice is crucial, and often challenging.

They must:

-be matched with cases for all those factors impacting the microbiota (sex, age, use of drugs…), to avoid confounding factors

-have a phenotype in clear contrast with that of cases. This does not necessarily imply that they must be «healthy»

(If faisible, multiple control groups recruited on the basis of a variety of criteria and methods could offer more insights into the heterogeneous effects of the microbiome)

2-LONGITUDINAL STUDIES:

Individuals are used as their own controls (e.g., by collecting samples before and after a treatment)

TIMING AND FREQUENCY OF SAMPLE COLLECTION are often determined/influenced by:

-monetary resources

-invasiveness of the sampling procedure

-subject compliance to study protocol

ALL STUDIES: EFFECT SIZE

The number of samples required for a microbiome study depends on the effect size:

an adequate number of subjects must be recruited to ensure that the effect expected from the exposure or intervention of interest can be detected.

<u>Evaluate the effect size with the help of:</u>

-statistical analysis and tools, also available online
-a pilot study
- scientific literature, looking at similar previous studies

COMMENTARI

Microbial Diversity in Clinical Microbiome Studies: Sample Size and

Gastroenterology 2020;158:1524–1528

README.rst

E-vident: elucidating sampling effort for microb analysis studies

FINAL, IMPORTANT REMARK: METADATA

Organize for an accurate collection and curation of metadata about all aspects of each sample, from clinical variables to sample processing: document everything and be consistent!

Record the maximum information about subjects, samples, and experimental procedures (that must be kept absolutely constant throughout the study) for use during downstream statistical and bioinformatics analysis

TABLE 2

Experimental and computational considerations to design, conduct and reproduce a microbiome study

Stage	Challenging factors to consider	Recommendations and risk mitigation
Study design	Scientific question(s)	• Clearly define the objective(s) of the study before deciding on the study design and statistica
	Effect size and target population	• Evaluate the effect size with a pilot study (or from scientific literature) • Estimate the study size based on anticipated effect size and similar previous studies (power a • Focus on specific aspects with well-defined inclusion and exclusion criteria to reduce heterog
	Control group(s)	• Appropriate control group should have a clinical phenotype distinct from the one under study criteria to avoid confounding factors • In longitudinal studies, individuals can be used as their own control by collecting baseline samp • Multiple control groups selected on various criteria can provide more insights
	Timing and frequency of sample collection	• Cross-sectional sampling from patients is appropriate for diagnostic microbiome signatures an microbiome types. It has the advantage that more individuals can be analyzed • Repeated sampling of the same subject over time should be chosen to get insight into temporal view of microbial community changes. It is also appropriate for monitoring disease severity or res suitable to discriminate between dysbiosis associated with disease and preclinical states and sh resilience. For repeated sampling, sampling frequency should be similar between subjects
	Additional measurements	• Select factors with continuous rather than categorical measurements as surrogates of disease

2-Data generation. Considerations on the experimental flow

MANY CHALLENGING FACTORS TO CONSIDER:

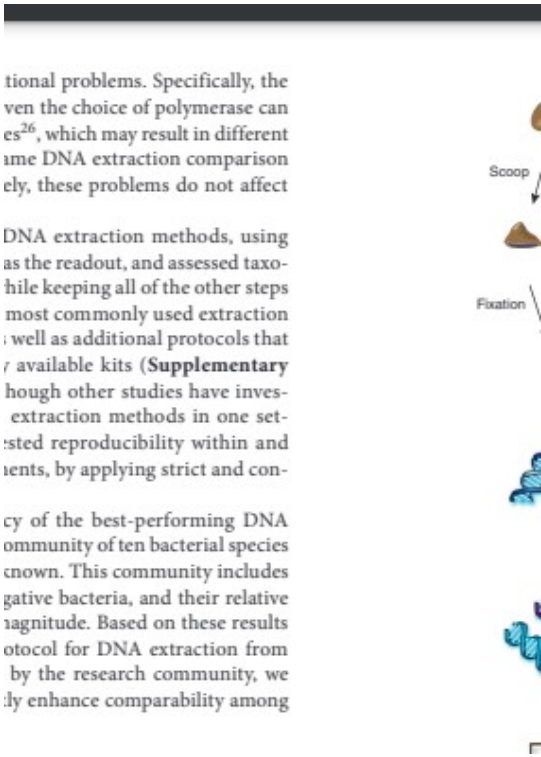
in metagenomics, the process of data generation is complex and consists of multiple steps, each of which is subject to technical variability

nature biotechnology

ANALYSIS

Towards standards for human fecal sample processing in metagenomic studies

er Nature. All rights reserved.



tional problems. Specifically, the
ven the choice of polymerase can
es[26], which may result in different
ıme DNA extraction comparison
ely, these problems do not affect

DNA extraction methods, using
as the readout, and assessed taxo-
hile keeping all of the other steps
most commonly used extraction
well as additional protocols that
available kits (**Supplementary**
hough other studies have inves-
extraction methods in one set-
sted reproducibility within and
ıents, by applying strict and con-

cy of the best-performing DNA
ommunity of ten bacterial species
cnown. This community includes
gative bacteria, and their relative
ıagnitude. Based on these results
otocol for DNA extraction from
by the research community, we
ly enhance comparability among

Scoop

Fixation

1-SAMPLE COLLECTION, TRANSIT AND STORAGE

Important variables in microbiota studies outcomes because they impact the DNA yield and quality->THEREFORE, WHATEVER THE CONDITIONS USED, IT IS MOST IMPORTANT TO KEEP THEM CONSTANT ACROSS SAMPLES AND AVOID FREEZE-THAWS

The most widely accepted protocols include immediate freezing in dry ice or liquid nitrogen, followed by storage at -80°C. Before freezing, the sample can be homigenized to generate multiple, identical aliquots

Immediate freezing is not always feasible, e.g., for stool samples collected at home. However, several studies underline that the effects of short-term storage conditions on community diversity and structure are generally small. Here, refrigeration can bridge the time between collection and processing (few days at +4°C: minimum effect)

Another possibility is the use of preservatives/stabilizing solutions for storage at +4°C or RT

ANYWAY: the most important factor that affects the composition of the microbial community is the number of freeze-thaw cycles

OXFORD FEMS JOURNALS investing in science

FEMS Microb

doi: 10.1093/fe
Advance Acce
Review Article

REVIEW ARTICLE

Practical considerations for large-microbiome studies

Doris Vandeputte[1,2,3], Raul Y. Tito[1,2,3], Rianne V
and Jeroen Raes[1,2,†]

[1]KU Leuven, Department of Microbiology and Immunology, Rega Ins

Table 2. Overview of reported significant effects of different storage conditions on alpha diversi
non Diversity index (SDI)), beta diversity measurements (e.g. unifrac, Bray–Curtis dissimilarity
immediate freezing (at −20°C or −80°C).

	Alpha-diversity	Beta-diversity
No buffer (4°C)	Richness ≈ (72 h[b], 24 h[a]) SDI ≈ (24 h[a])	Individual > storage, with BC (72 h[a,b])
No buffer (RT)	Richness ≈ (72 h[b]; <2 h[m], 24 h[a], 96 h[a]) Richness ❖ (<96 h[a]) Evenness ↘ (72 h[b]) SDI ↘ (72 h[b,p]) SDI ≈ (72 h[f], <2 h[m], 24 h[a], 96 h[a]) SDI ❖ (<96 h[a])	Individual > storage, with weighted unifrac (<2 h[m], <96 h[a,o]), unweighted unifrac and BC (<96 h[a,o]) Increased BC, (un)weighted unifrac and Jaccard over time from 4 h RT storage onwards[p]
TE-buffer (RT)	Richness ≈ (72 h[b]) Richness ❖ (<96 h[a]) SDI ❖ (<96 h[a])	Individual > storage, with unweighted unifrac, weighted unifrac, BC (<96 h[a])
RNAlater (RT)	Richness ≈ (24 h[q], 72 h[b], 96 h[a]) Richness ❖ (<96 h[a]) Evenness ↘ (72 h[b]) SDI ↘ (72 h[f]) SDI ≈ (96 h[a]) SDI ❖ (<96 h[a])	Individual > storage, with unweighted unifrac (24 h[q],96 h[a,o]), weighted unifrac (24 h[q],96 h[a,o]), BC (24 h[q], 72 h[b], 96 h[a,o])
Ethanol 70% (RT)	Richness ❖ (<96 h[a]) SDI ≈ (1–8 weeks[k]) SDI ❖ (<96 h[a])	Individual > storage, with unweighted unifrac, weighted unifrac, BC (<96 h[a])
Ethanol 95% (RT)	Richness ≈ (24 h[q], 96 h[a]) SDI ≈ (24 h[q], 96 h[a])	Individual > storage, with unweighted unifrac (24 h[q], 96 h[a], 1–8 weeks[k]), weighted unifrac (24 h[q], 96 h[a], 1–8 weeks[k]), BC (24 h[q], 72 h[b], 96 h
Ethanol 95% + lyophilisation (RT)	Richness ≈ (24 h[q])	Individual > storage, with unweighted unifrac (24 h[q]), weighted unifrac (24 h[q]), BC (24 h[q], 72 h
RTTVs (RT)	Richness ≈ (72 h[b], 1/28 days[j]) SDI ≈ (1–2 weeks[r]) SDI ↗ (1/28 days[j])	Individual > storage, with BC dissimilarity (1/28 d[j]), unweighted unifrac (1–8 weeks[k]), weighted unifrac (1–8 weeks[k]), spearman distance (1–2 weeks[r]),

2-DNA EXTRACTION

The most important source of technical variation and bias: different DNA extraction protocols result in different diversity profiles (and ecological indexes) because of differential sensitivity to common mechanical or chemical lysis methods of microbial cells within the community

Combination of chemical and mechanical (zirconia beads) lysis is proposed to capture the most accurate community composition and to efficiently extract DNA from Gram-positive bacteria

Extraction protocols MUST be applied consistently within a study. The use of commercial kits increases the possibility for consistency between independent studies

Contamination issues are particularly important, especially when working with «extreme» environments with low microbial biomass (eg., meconia): run a blank extraction control through DNA extraction and PCR

In the lab, if possible, spatially separate pre- and post-PCR steps

3-CHOICE OF PRIMERS: WHICH V-REGIONS SHOULD I TARGET?

-Rely on previous studies

-Consider that different couples of universal 16S primers may be biased toward (or against) certain bacterial taxa, thus giving artefactual over- (or under-representations) of them.

-For example, the 27F/338R primer sets (targeting the V1–V3 regions) is biased against the amplification of *Bifidobacteria*.

-Other primer sets poorly resolve specific taxa. For example, *Enterobacteriaceae*, *Clostridiaceae*, *Peptostreptococcaceae* are poorly resolved by primers on V3-V4

4-DNA AMPLIFICATION

PCR amplification could lead to biases when DNA template concentration is low and/or PCR cycle number is high.

Some hints to minimize bias:

-Use a standard (relatively high) DNA concentration across samples
-Pool multiple PCR for each sample
-Use proof-reading polymerases
-Use longer annealing times to reduce chimera formation
-Use PCR enhancers to improve yields

OTHER HINTS: THE IMPORTANCE OF CONTROLS

1-include negative controls at different sample processing steps:

-Blank extraction control: DNA extraction and all subsequent steps performed on water to evidence the «kitome»
-Blank PCR control: PCR and library generation performed without template addition
-Blank sequencing control: sterile swab opened in the sequencing lab undergoes full sequencing control

(results on low biomass samples are potentially indistinguishable from contamination!)

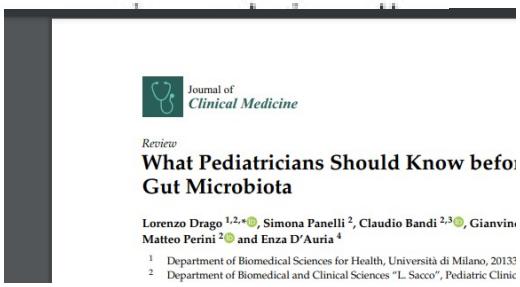
2-include positive controls. Use mock communities to check for potential introduction of bias or distortion during sample processing

XLIX CONGRESSO NAZIONALE AMCLI

26 febbraio - 1 marzo 2022 | Palacongressi Rimini

A step-by-step workflow for a m

1. Identify a well-designed target
2. Clearly define primary endpoints of m
3. Choosing timing and frequency of sample collect
4. Use the right storage and samplin



Journal of *Clinical Medicine*

Review

What Pediatricians Should Know befo
Gut Microbiota

Lorenzo Drago [1,2,*], Simona Panelli [2], Claudio Bandi [2,3], Gianvin
Matteo Perini [2] and Enza D'Auria [4]

[1] Department of Biomedical Sciences for Health, Università di Milano, 20133
[2] Department of Biomedical and Clinical Sciences "L. Sacco", Pediatric Clinic

3-Data analysis: Mothur, QIIME & Co

1650 www.drugdiscoverytoday.com

TABLE 3

Overview of computational methods and recommendations for the context of applicability for metagenomic data analysis

	Methods	Description	Advantages	Limitations	Recommendations and/or mit
Taxonomic profiling: to identify taxa and estimate their relative abundances in microbiome samples					
Marker gene sequencing analysis	Step 1: operational taxonomic unit (OTU) picking – cluster sequence reads into taxonomic units (phylotypes) based on sequence identity Preferred marker genes: 16S for bacteria and archaea and internal transcribed spacer (ITS) for fungi				
	De novo	Used to cluster sequences without any external reference sequences	• No need for *a priori* knowledge	• Computationally intensive: large memory and long execution times	• Not suitable for very large da exhaustive *de novo* clustering is • Greedy algorithms are more • First group input sequences i bins, and then cluster within ea
	Closed-reference	Used to cluster sequences with external reference sequence database or collection	• Fast and scalable (parallelization)	• Unmatched sequences discarded	• Not recommended for enviro samples because they contain that mostly lack reference genc
	Open-reference	Hybrid approach: initially clusters all sequences using external reference sequence database or collection and then applies *de novo* method for sequence reads that do not match any reference	• More reads can be mapped • Fast		• Most suitable approach to us • Not possible to use it when c different regions of the 16S rRN
	Step 2: OTU taxonomic assignment – assign OTUs to taxa				
	Numerous algorithms available	Closed-reference: transfer taxonomy of the reference sequence *De novo* and open-reference: search for the closest match to an OTU, from which a taxonomic lineage is inferred using reference databases of marker genes		• Many OTUs will lack a complete taxonomy label, resulting from a lack of confidence or from a matching to a branch in the phylogeny that lacks taxonomic information • Genus and/or species information might be biased toward reference strains that can be cultured	• OTU quality filtering based o • OTUs of similarity >97% cons species for bacteria and archae • Rarefy (i.e., randomly select e sequences from each sample fc OTU table to mitigate differenc • Important to report database names can be different from or • Use Biological Observation M for OTU tables
	Step 3: Taxonomy collapsing – assess abundances at higher taxonomic level				

	taxonomic information	
Downstream analysis and visualization		
Alpha-diversity	Describe microbiome taxa or functional diversity within a sample	• Visualize proportions c taxonomic ranks using p
Beta-diversity	Describe microbiome taxa or functional diversity between samples	
Principal coordinate analysis (PCoA)	Reduce microbiome data dimensionality to ease data visualization and interpretation	• Use principal coordina separation and bacterial
Clustering or classification – supervised	Build a classifier to discriminate sample groups and train predictive model for classification	• Identify taxa- and/or fu predictors of subject clas regression-based approa
Clustering or classification – unsupervised	Cluster samples based on taxa or function-associated abundances without prior knowledge	• Use distance metric suc a clustering algorithm • Perform clustering in s robustness, because the distance metric and algo • Use resampling metho in a cluster • Heatmap visualization
Statistical modeling	Develop statistical model(s) to test scientific hypotheses	• Define key dependent primary study objectives • Check and, if necessary, experimental and technica • Transform data for norm sequence read counts to p or use nonparametric stat • Possibly drop rare taxa, w the analysis • Avoid categorization (e.g variables, but rather keep • Once numerically encod categorical variables • Impute missing data for

Times they're changing

OXFORD

Human Molecular Genetics, 2018, Vol. 27,

doi: 10.1093/hmg/ddy177
Advance Access Publication Date: 14 May 2018
Invited Review

INVITED REVIEW

Long reads: their purpose and place

Martin O. Pollard[1,2,*], Deepti Gurdasani[1,2], Alexander J. Mentzer
Tarryn Porter[1,2] and Manjinder S. Sandhu[1,2]

[1]Human Genetics - Wellcome Sanger Institute, Hinxton, Cambridge CB10 1HH, UK, [2]Univers
Department of Medicine, Addenbrookes Hospital, Box 157, Hills Road, Cambridge CB2 0QQ, U

PLOS ONE

RESEARCH ARTICLE

Comparing bioinformatic pipelines f
microbial 16S rRNA amplicon seque

Andrei Prodan[1]*, Valentina Tremaroli[2], Harald Brolin[2], Aeilko H. Zw
Max Nieuwdorp[1], Evgeni Levin[1,4]

1 Department of Experimental Vascular Medicine, Amsterdam University Medical C
Netherlands, 2 Wallenberg Laboratory for Cardiovascular and Metabolic Research,

Third generation platform (Pac Bio) allow sequencing of the entire 16S gene, with an enhanced taxonomic resolution

OTUs vs ASVs (amplicon sequence variant): while OTU approaches attempt to collaps similar sequences into a consensus sequence, thus minimizing the influence of any sequencing errors within the pool of reads, the ASV tries to reconstruct each exact sequence present in a sample

50
AMCLI
associazione
microbiologi
clinici italiani
1970 – 2020

Thank you