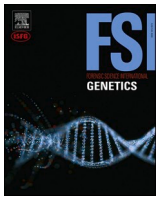
Forensic Science International: Genetics 56 (2022) 102627

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen



Integrating the human microbiome in the forensic toolkit: Current bottlenecks and future solutions

Celia Díez López [*], Athina Vidaki [1], Manfred Kayser [1]

Erasmus MC, University Medical Center Rotterdam, Department of Genetic Identification, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
Human microbiome
Microbial forensics
Forensic microbiome
Massively parallel sequencing
Forensic MPS
Forensic toolkit

ABSTRACT

Over the last few years, advances in massively parallel sequencing technologies (also referred to next generation sequencing) and bioinformatics analysis tools have boosted our knowledge on the human microbiome. Such insights have brought new perspectives and possibilities to apply human microbiome analysis in many areas, particularly in medicine. In the forensic field, the use of microbial DNA obtained from human materials is still in its infancy but has been suggested as a potential alternative in situations when other human (non-microbial) approaches present limitations. More specifically, DNA analysis of a wide variety of microorganisms that live in and on the human body offers promises to answer various forensically relevant questions, such as post-mortem interval estimation, individual identification, and tissue/body fluid identification, among others. However, human microbiome analysis currently faces significant challenges that need to be considered and overcome via future forensically oriented human microbiome research to provide the necessary solutions. In this perspective article, we discuss the most relevant biological, technical and data-related issues and propose future solutions that will pave the way towards the integration of human microbiome analysis in the forensic toolkit.

1. Introduction

A great variety of microorganisms live in and on the human body, including bacteria, archaea, lower and higher eukaryotes and viruses, collectively referred to as the human microbiota [1,2]. The term 'human microbiome' refers to these microorganisms, their genomes and the surrounding environmental conditions [2]. Human microbiome is also employed to refer to the collection of genes and genomes of the human microbiota [2]. Microorganisms are found in and on almost every human body part, where they have co-evolved with their hosts over a long period of time to form a complex mutualistic relationship [3–5]. The number of microorganisms living in and on the human body is of the same order as the number of human cells [6]. However, the genetic and genomic diversity of the human microbiota is estimated to surpass that of humans themselves by several orders of magnitude [1], providing a suitable prerequisite for human microbiome analysis in various areas of scientific research and applications such as in forensics.

The forensic interest in studying the human microbiome and its development through time can be graphically represented in a 'hype' cycle as shown in Fig. 1. In the mid 2000's, advances in massively parallel sequencing (MPS, also referred to as next generation sequencing, NGS) and bioinformatics analysis tools started to allow for population-level surveys of the human microbiome from different body parts. These include the Human Microbiome Project (HMP) Consortium in the USA [7,8] and the Metagenomics of the Human Intestinal Tract (MetaHIT) Project in the European Union [9]. The achievements of these large consortia boosted the knowledge on the human microbiome and led to the generation of large microbiome datasets, mostly relevant to the medical field. The public availability of such datasets triggered great interest in the interdisciplinary application of the human microbiome, including in the forensic field (innovation trigger, Fig. 1). Since then, more and more researchers from both within and outside the forensic community turned their attention to studying the microbiome with the aim to answer forensically relevant questions (peak of inflated expectations, Fig. 1), including individual identification [10], post-mortem interval estimation (PMI) [11] and tissue/body fluid identification [12,13], among others. Currently, although of great promise, forensic microbiome analysis faces various hurdles that need to be considered, further investigated and eventually solved before the human microbiome can fulfil its current promises to become an integral part of the

* Corresponding author.
E-mail addresses: c.diezlopez@erasmusmc.nl, celiadiezlopez@gmail.com (C. Díez López).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.fsigen.2021.102627
Received 5 July 2021; Received in revised form 12 October 2021; Accepted 27 October 2021
Available online 1 November 2021

forensic toolkit in the future (trough of disillusionment and slope of enlightenment, Fig. 1).

In this perspective paper, we provide an overview on the most important biological, technical and data-related issues current forensic human microbiome analysis is facing and propose future solutions how dedicated forensically-motivated and other related research can pave the way towards the integration of the human microbiome in the forensic toolkit.

2. Current state of forensic human microbiome research

Potential forensic microbiome applications have previously been investigated in areas where human (non-microbial) genetic solutions show limitations or fail for various reasons (Fig. 2), which will be summarized here.

2.1. Human individual identification

Since the human microbiome is shaped by numerous internal and external factors, which include but are not limited to host's genetics [14], sex [15], ancestry [16] and lifestyle [17], the microbiome of human individuals may contain a unique 'fingerprint' that can potentially distinguish between them and can thus be used for human individual identification purposes. This can be especially relevant in investigations where the recovered human DNA is not of sufficient quantity and/or quality to obtain a fully individualized DNA profile based on short tandem repeats (STRs); for example, when dealing with 'touched' samples that do not typically provide sufficient amounts of human DNA [18]. In this context, various forensically-motivated studies have linked objects (e.g. mobile phones, computer equipment) with their owners by comparing DNA similarities in the microbial composition of the 'touched' sample with the one obtained on the skin of the study individuals [10,19–22], reporting identification accuracies as high as 93% [10]. A few studies have also investigated the individualising potential of the human microbiome in other samples than skin, such as in saliva [23] and pubic hair [24,25]. Overall, microbiome-based human individual identification has been investigated so far on the basis of comparing individuals within the same study; however, for real forensic value, microbiome 'fingerprints' should be able to individualise any random person from the population, which requires appropriate statistical frameworks that are currently missing.

2.2. Post-mortem interval estimation

The succession of the human microbiome after death as well as associated soil microbial communities (in the case of outdoor crime scenes with bodies lying on the ground) has proven to be a suitable biomarker for the PMI estimation of human cadavers. Various studies have characterized the microbiome composition at different body sites during the decay process, including spleen, liver, brain, heart, blood, bones, gut, skin and oral cavity [11,26–32], with error rates as low as 1.7 days when analysing the first 25 days of decomposition [11]. However, some of these studies have reported confounding effects of the individual [31], sex [26] and abiotic factors (e.g. ambient temperature, solar irradiance) [27,31] on the microbiome succession, which can contribute to misleading PMI estimations and, hence, should be further investigated.

2.3. Geolocation inference

As previously mentioned, several factors can shape the human microbiome, some of which can be the basis for the inference of host's geolocation from the microbiome analysis of human traces left at crime scenes. These factors include geographical latitude [33,34], industrialization level of country of residence [35–38], and cultural and societal components [39]. Forensically oriented studies have shown that the

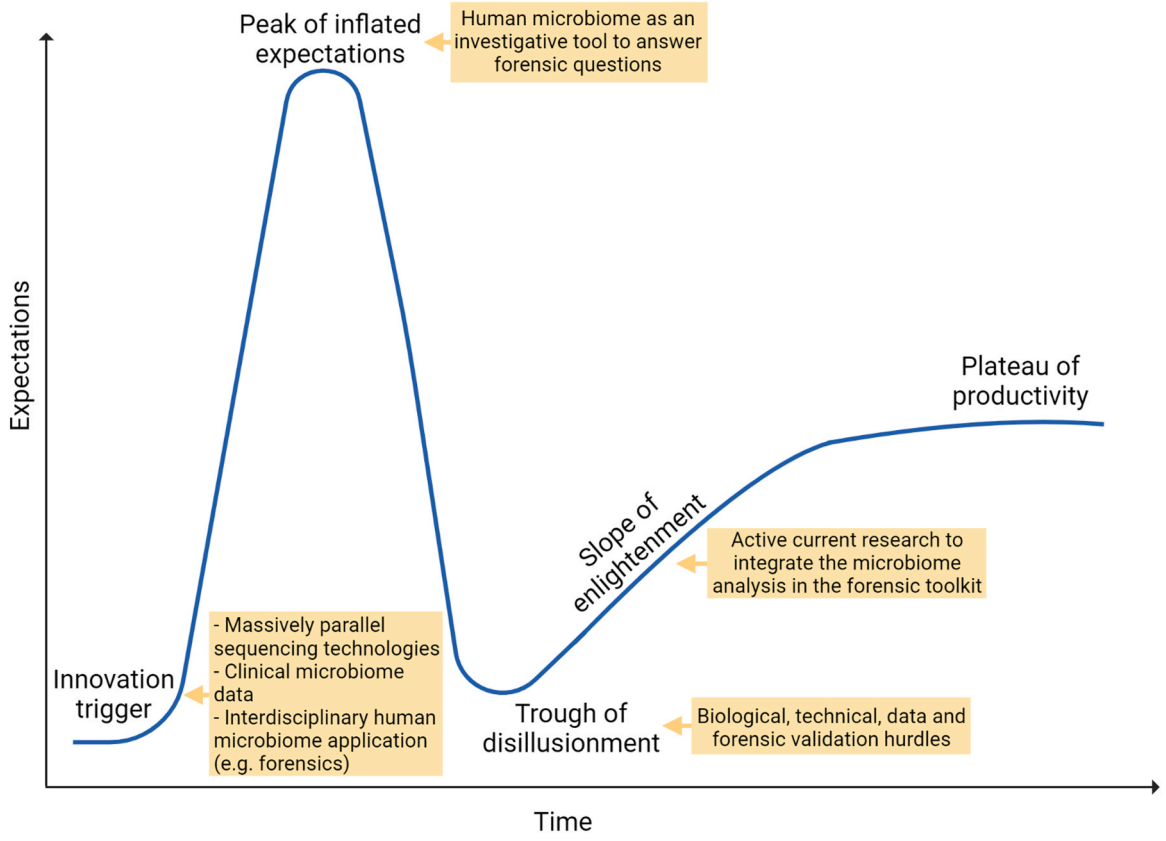


Fig. 1. The 'hype' cycle describing recent trends in forensic microbiome research in the last few years (figure created with BioRender).

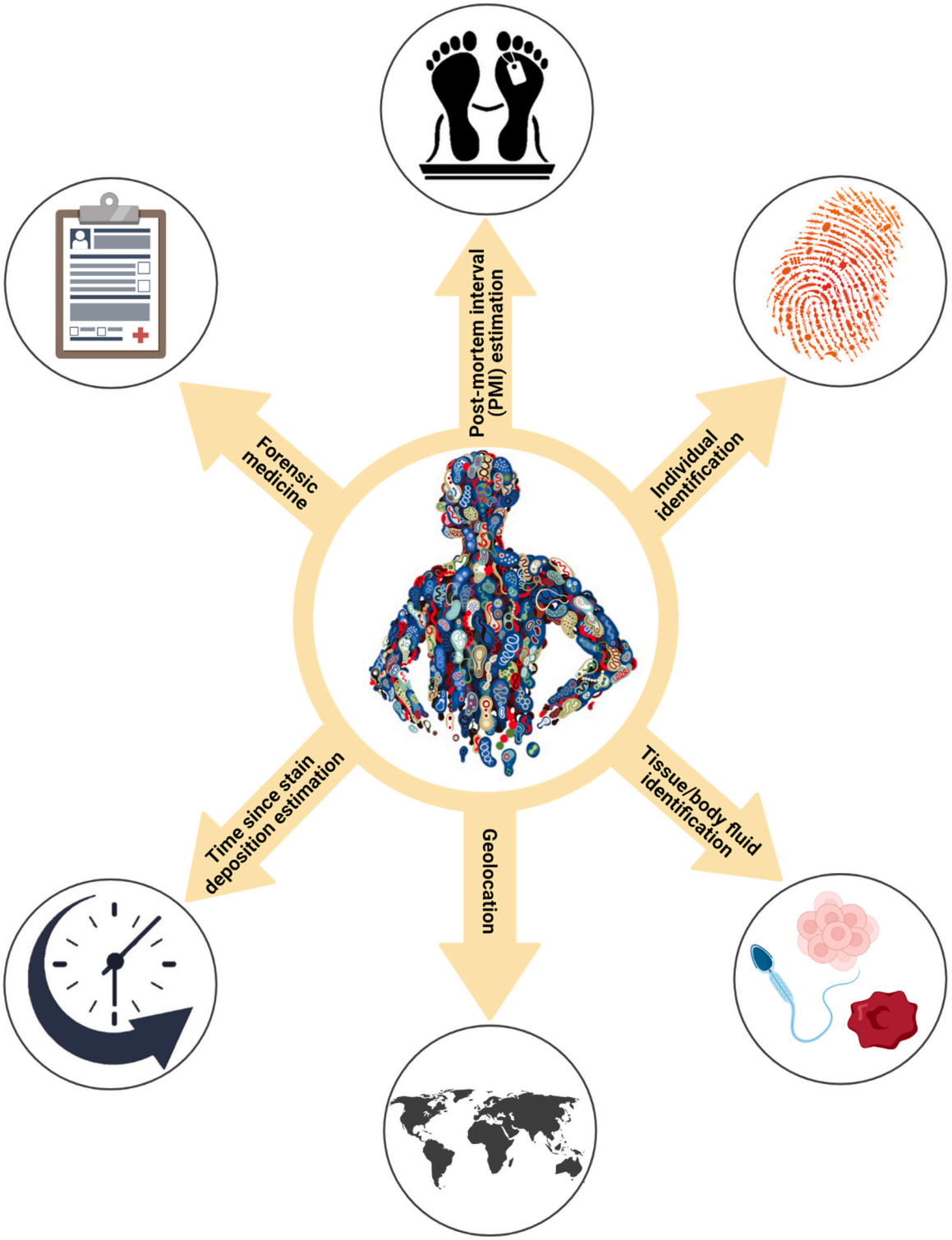


Fig. 2. State of the art of forensic microbiome research (figure created with BioRender).

microbial community composition of objects from owners living in the same city are more similar between them compared to objects from owners living in different cities [21,40,41], which can ultimately assign human biological samples to the city of origin via human microbiome analysis [42]. In a forensic casework context, the geographic emergence of the strains of the stomach bacteria *Helicobacter pylori* [43,44] was proved useful to identify the geographical origins of unidentified cadavers [45]. If successful, microbiome-based geolocation analysis shall be combined with bio-geographic ancestry (BGA) analysis based on human ancestry informative DNA markers to receive additional geographic information about the tested person.

2.4. Tissue/body fluid identification

Identifying the tissue/body fluid that a biological crime scene stain originated from can be informative in crime scene reconstruction and activity level reporting. For most forensically relevant human tissues/body fluids, the human microbiome is an outstanding biomarker based

on the differences of microbial communities across various sites of the human body [46,47], and due to the larger intra-individual differences among body sites compared to the inter-individual differences for the same body site [46,48]. Forensically motivated studies have mainly focused on the microbiome-based identification of vaginal fluid [12, 49–52] and less so in saliva [12,53], skin [12], and blood traces of different body site origins [13,54,55].

2.5. Time since stain deposition estimation

Information on the time when a stain was deposited at a crime scene can help police to assess alibis of known suspects and witnesses or provide an investigative lead to search for the right suspect. This forensic application had not been investigated in a microbiome context until recently when it has been shown that DNA- and RNA-based time dependent changes occur in the microbial composition of human biological traces [56,57]. Particularly, DNA profiling of commensal bacteria can estimate the time since deposition of an individual's saliva stain with an average error rate of 5 days when analysing the first 30 days after stain deposition [56]. However, further research is needed to better understand and model the microbial changes for this application.

2.6. Other forensic applications

Microbiome applications in forensic medicine include the identification of the manner of death, such as drowning [58,59], cardiovascular- and drug-related deaths [60]. On a different point, genotyping microbes responsible for sexually transmitted diseases can be useful to trace the source individual of the infection, which can be relevant in the legal and law enforcement context [61]. For instance, by this the infectious microorganism in the suspect (i.e. source of infection) and victim (i.e. infected individual) can be linked, which has proven useful in identifying the perpetrator in cases of child abuse [62,63].

3. Biological issues and potential solutions

The human microbiome presents several biological features suitable for forensic applications, such as ubiquity [6], response to changes in the environment [64,65] and capacity of being shed, deposited and exchanged [66,67]. However, other biological aspects of the human microbiome might be undesired depending on the forensic application in mind (Fig. 3) and therefore need to be further investigated and considered in future forensic research.

3.1. Inter- and intra-individual variation

Many human microbiome studies, including forensically motivated ones, have reported substantial inter-individual (between individuals) and intra-individual (within the same individuals over time) variation of the human microbiome, such as in the three most studied body sites in the forensic microbiome context; namely, vagina, skin and oral cavity, specifically saliva [19,20,49–51,53]. This inter- and intra-individual variation depends greatly on the body site of interest, with variation in skin and vaginal sites being higher than in oral sites as current studies suggest [68,69]. The variation also depends in the taxonomic level of analysis; this is, the hierarchical classification of living organisms, from highest to lowest: domain, kingdom, phylum, class, order, family, genus and species. Based on this, taxa prevalent between or within individuals diminish down to the genus, species and strain (group of related species) levels [70].

Inter-individual variation may be desirable depending on the forensic application as in principle it allows better discrimination between individual samples. While large inter-individual variation serves as basis for microbiome-based human individual identification, it is rather undesirable in forensic microbiome applications not aiming to differentiate individuals, such as for forensic tissue/body fluid identification. The majority of forensic microbiome research on forensic tissue/body fluid identification has focused on vaginal secretions, which is explained by the forensic relevance of this body fluid in the investigations of sexual assault cases. Previous studies mostly targeted

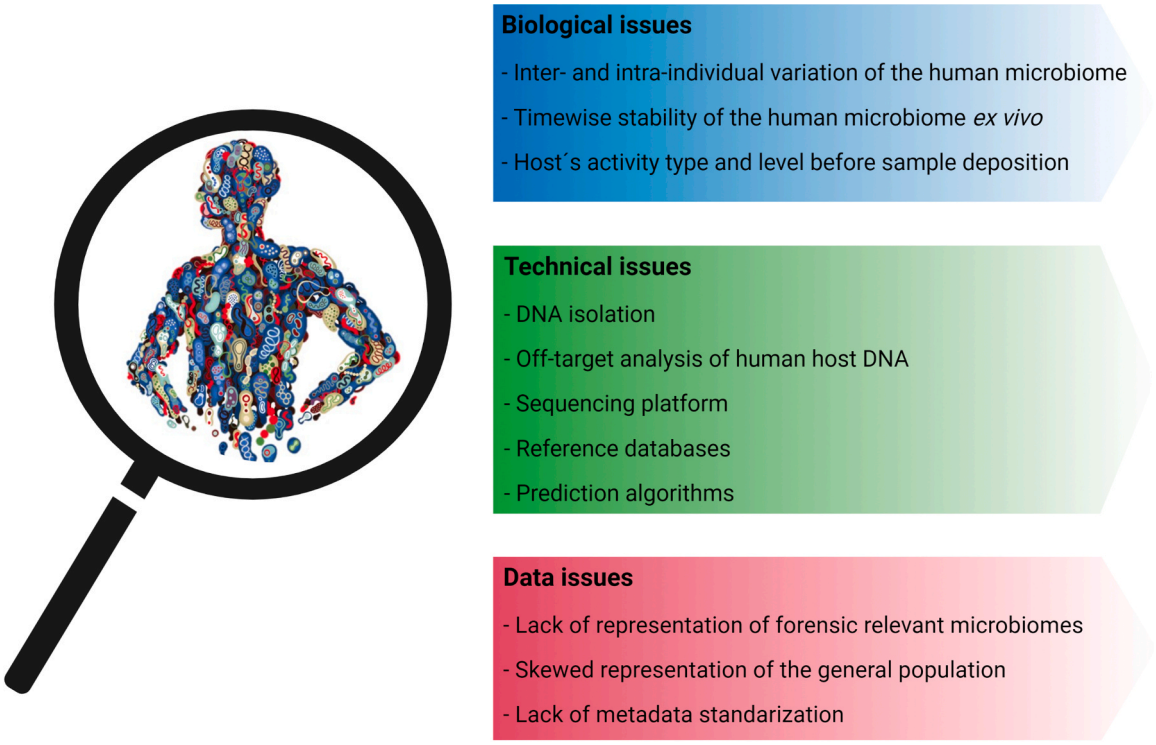


Fig. 3. Biological, technical and data-related issues of current forensic microbiome analysis (figure created with BioRender).

vaginal candidate bacteria, typically belonging to the *Lactobacillus* genus, via quantitative PCR (qPCR), capillary electrophoresis or microarray analysis [49–52,71]. However, between women, different *Lactobacillus* species are present in different abundances – from being the major microbial community constituent to being absent – leading to large inter-individual variation [72]. As a result, many of these forensic studies did not detect the targeted *Lactobacillus* species in all the analysed individuals [49–51], resulting in unreliable outcomes and limiting the generalization of these targeted approaches based on just a few candidate species. We recently showed that this limitation can be overcome by MPS approaches targeting the entire microbial community that is characteristic of a body site instead of restricting microbial DNA analysis to a small number of candidate species [12,13].

Intra-individual variation, however, is undesirable in all forensic applications, since it can hinder the comparison between a trace deposited at a crime scene at a time *x* and a reference sample collected at a later time *y*; with the time frame between *x* and *y* being days, weeks or, in some cases, even years depending on the circumstances. This can represent a great challenge for instance in skin microbiome-based human individual identification where transient microorganisms, resulting from environmental exposure, such as transfer from other skin sites from the individual itself (skin-to-skin contact), between individuals [19,66] or the built environment [21,41], can be confused with individualised microbial markers. Moreover, even though microorganisms preserve through time, their abundances change considerably due to internal and external factors [46,56,73]. For instance, saliva microbiome abundances show circadian rhythms [74], i.e. rhythmic fluctuations within the 24 h day-night cycle, which can influence the microbiome analysis results depending on the time during the day a sample is deposited at the crime scene or collected for reference purposes. Hence, intra-individual variation should be considered with caution in quantitative approaches relying on the absolute or relative abundance of microorganisms.

3.2. Timewise stability ex vivo

At crime scenes, human biological traces are exposed to various environmental factors for variable periods of time between sample deposition and sample collection. However, this ex vivo (outside the human body deposited as crime scene stains) timewise (in)stability of the human microbiome as well as its impact on investigation outcomes has been very little studied so far. Substantial instability of the human microbiome ex vivo might have a comparable effect to intra-individual variation, making the comparison between crime and reference samples challenging and potentially directing towards misleading results.

Additionally, the methodological approach and analytical technology (i.e., absolute vs. relative quantitative approaches, such as via qPCR vs. MPS) and the taxonomic level of data analysis also play decisive roles when analysing timewise (in)stability of the human microbiome ex vivo. A few studies have analysed the timewise (in)stability of the human microbiome in forensically relevant biological stains (venous blood, menstrual blood, saliva, skin, vaginal fluid, semen, pubic hair) exposed to indoor conditions (average room temperature of 20 °C) for variable periods of time (from 2 weeks up to 1.5 years) [22,24,57,75]. These studies employed targeted amplicon MPS and analysed whole microbiome profiles at the genus level [22,24,75], except one study that employed whole RNA sequencing and analysed whole RNA profiles at the domain and phylum levels [57]. All studies reported no significant variation in the microbial profiles stored for different time periods.

However, this apparent timewise stability of the human microbiome ex vivo can vanish when carefully selected candidate microbial species are analysed using absolute quantitative approaches. Recently, we demonstrated significant time-dependent changes of selected bacterial species in saliva stains exposed to indoor conditions up to 1 month [56]. Moreover, the detection of original members of the microbial community in a given stain might be surpassed by the detection of other organisms in the environment, especially in stains exposed to outdoor conditions (e.g. detection of surrounding vegetation) [57]. In forensic applications such as PMI estimation or time since stain deposition estimation, where some degree of timewise instability is desired, this environmental 'fingerprint' might limit the development of generalized approaches applicable to any case. One potential solution is to implement crime scene-specific approaches, as we have suggested recently [56]. Nevertheless, this area of research still needs to be further investigated in more detail including more tissues/body fluids and more environmental conditions.

3.3. Host's activity type and level

The type and level of human host's activities prior to sample deposition may or may not have an effect in the microbiome analysis results. In one of our studies, we reported that the potential transient microorganisms picked up from the environment during routine activities (e.g. type on the computer, touch door handles) does not have a negative effect on the correct identification of hand skin samples as part of tissue/body fluid identification [12]. This could be explained by the very distinct nature of the skin microbial communities compared to the oral and vaginal ones we analysed as part of our study [12]. However, in a different study, we observed that time between nose-blowing and nasal blood deposition can impact how similar or dissimilar a nasal blood stain microbiome is to the reference nasal mucosa or skin epithelium one [13]. Nose-blowing pushes out the mucus dragging out some nasal mucosa bacteria, while leaving exposed, for a time, other commensals in closer contact with the nasal epithelium. Hence, unknown nose-blowing activities prior to nasal blood deposition can provide a source of skin-misclassification of nasal blood samples collected at crime scenes [13]. On a different example, the saliva microbiome, which is highly exposed to fluctuated host's activities during the day, presents the highest live microbial load immediately upon waking, dropping after brushing teeth [76]. This highlights the need to investigate the impact of host's activity type and level prior to sample deposition on the microbiome and requires ways how to deal with it during analysis in an application-specific manner.

4. Technical issues and potential solutions

Biases in microbiome analysis outcomes can be introduced at every step in the analysis pipeline, from the very beginning during the nucleic acids isolation, via the initial amplification (in targeted amplicon MPS) and sequencing steps, to the final bioinformatics and statistical methods for data analysis. Such biases not only need to be considered through a suitable study design, but also need to be acknowledged by detailed and fair reports in scientific publications. The Achilles' heel of (forensic) microbiome research is with no doubt the lack of consensus in both experimental and analytical methods within and across different scientific communities. This lack of consensus makes the comparison of the results between studies difficult, if not impossible; hence, limiting meta-analysis of data from different studies and the establishment of an accurate overview of the forensic microbiome research's current stage and prospects. Also, this lack of standardization will eventually cause issues in acceptance of microbiome-based forensic evidence in court. Hereafter, we discuss the main technical considerations in microbiome analysis from a forensic standpoint (Fig. 3). In principle, many of the discussed aspects apply to any DNA-based human microbiome research, including both targeted amplicon sequencing and whole genome shotgun sequencing (WGS) metagenomics (Fig. 4), unless stated otherwise.

4.1. Microbial DNA isolation

The variation in microbiome analysis outcomes introduced by the DNA isolation method is currently under debate. Some studies have

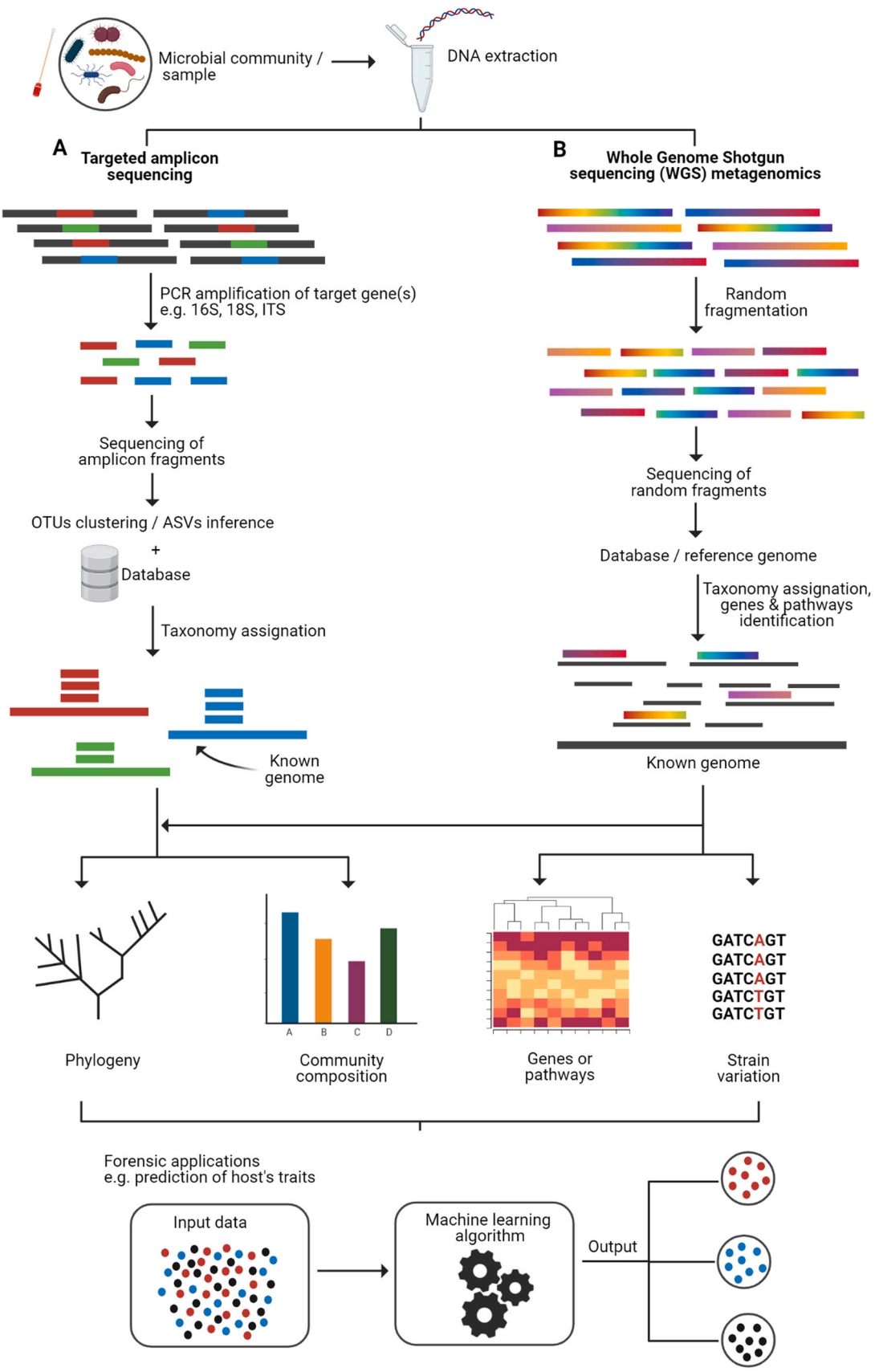


Fig. 4. Overview of currently available experimental and analytical tools for studying the human microbiome from obtaining microbial DNA of a sample to its forensic application. A. Massively parallel sequencing (MPS) of targeted amplicons. B. Whole Genome Shotgun sequencing (WGS) metagenomics (figure created with BioRender).

reported that different DNA isolation methodologies led to differences in the abundance of specific groups of bacteria [77–80]. The likely reason for this is that some bacterial cells are harder to lyse than others, such as Gram-positive bacteria compared to Gram-negative bacteria [81]. Moreover, the DNA isolation method of choice also affects the yield and quality of the obtained microbial DNA [79,82], as well as the presence of inhibitors that can affect downstream analysis steps [83]. However, this topic is not free of controversy since there are also studies reporting no significant effects of the DNA isolation method [84–86].

Microbial DNA isolation is based on 'harsh' extraction methods (e.g. mechanical, enzymatic) since, for instance, the bacterial cells are furnished with a wall of peptidoglycan matrix that opposes greater resistance compared to human cell membranes [87]. Therefore, microbial DNA isolation methods typically allow for the parallel extraction of human and microbial nucleic acids, meaning that human- and microbial-related forensic analysis can be carried out in the same extract, if material is available. From the vast variety of available microbial DNA isolation methods, it is recommended to use one that includes a bead-beating step, which has been linked with a more efficient DNA extraction from Gram-positive bacteria, achieving higher DNA yields and bacterial diversity [88–90]. Further, it is important to be aware that commercial DNA isolation kits themselves can introduce contamination, known as the 'kitome', which not only varies between different kits from the same and different manufacturers, but also between different batches of the same kit [91]. To this, it should be added that microorganisms are not only ubiquitous in the reagents but also in the lab environment, even in ultraclean laboratories [92]. This can be especially problematic in low-biomass samples (e.g. skin swabs, highly degraded samples), where contaminants can outnumber endogenous microorganisms within samples. Therefore, it is primordial to always include negative controls to monitor potential contamination. However, how to best account for the results obtained from these negative controls is still in question. It is not good scientific practice to simply remove the microbial taxa found in the negative controls from the data obtained from the study samples. It is unknown if their presence in the study samples goes back to the contamination picked up with the negative controls or is intrinsic to the study samples and may be enhanced by the contamination [93]. Thus, whenever possible, it is recommended to combine data from qPCR and sequencing to compare the absolute and relative abundances of potential contaminating taxa by using adequate statistical methods that assess whether or not they are real contaminants [94].

4.2. Off-target analysis of human host DNA

It is widely known that the amount of human host DNA poses a major challenge to WGS metagenomics analysis based on the generation of an overview of the gene composition in the samples, including both microbial and human host gene fractions. As a result, human host sequences could 'overwhelm' the microbial sequence fraction of interest. Off-target analysis of human host DNA has been generally overlooked also in targeted amplicon MPS analysis where it is a common practice to simply remove any sequencing reads that fall outside the average sequence length or are not classified as being of microbial origin [95]. In targeted amplicon MPS, a PCR amplification step of a region within the microbial gene(s) of interest, such as the widely used bacterial 16S ribosomal RNA (rRNA) gene, is performed prior to sequencing. The non-specific, co-amplification of human host DNA by 16S rRNA gene PCR primers can lead to partly masking the microbial fraction of interest in the downstream sequencing process. This can be notorious when analysing samples with a high human to microbial ratio. In other words, the lower the bacterial load relative to the human load, the more problematic these human genome-aligned reads are, as we previously reported when analysing human blood samples [13]. It is of high importance that the (forensic) research community is aware of this problem, especially since this issue has been reported when using 16 S rRNA gene PCR primers recommended by the HMP Consortium [13,96] that are widely used as reference. Recent research has shown that other 16S rRNA gene PCR primer pairs produce lower amounts of human sequence reads [95], though this should be further investigated in forensic-type samples with high human to microbial ratios, such as venous blood.

4.3. MPS platforms

The sequencing platform is an important choice in MPS-based microbiome analysis determining targeted fragment size, read length, sequence accuracy and cost [97]. Previous reports indicate that data are indeed reproducible across sequencing runs within and between different MPS technologies [98]. Even when obtained by different MPS platforms, results have proven to be robust as long as the datasets are generated following the same experimental protocol [99]. This has allowed researchers of forensic microbiome studies including us to reuse the microbiome data from large microbiome consortia, such as the HMP data, which were produced with the Roche 454 pyrosequencing platform that went out of business in 2013 [12,13,100].

The Illumina MiSeq is the preferred MPS platform in forensic microbiome studies [21,30,101], though we also opted for the alternative Ion Torrent PGM and S5 instruments (Thermo Fisher Scientific) in some of our studies [12,13]. MiSeq presents a high-throughput option that allows for high sequencing depth [102], is supposed to produce high-quality data, and allows for strict quality control parameters [99]. However, even though the Illumina MiSeq errors are reported to be around 0.01%, researchers have described errors as high as 10% and recommended the analysis of reads with complete overlap between the forward and reverse paired-end sequencing to correct for that [103]. Further, the long-read third-generation sequencing platforms, such as Oxford Nanopore MinION and Pacific Biosciences (PacBio) Sequel, though not commonly used in forensic microbiome as of yet, offer several advantages over the short-read second-generation sequencing platforms, like Illumina and Ion Torrent sequencing, as they provide full-length genes and increased resolution, which opens up new possibilities and applications options. Although these platforms are said to suffer from high error rates in the range of 5–15% [81], a study was able to reduce the observed error rate for the full-length of the bacterial 16S rRNA gene from 0.69% to 0.03% for PacBio data, which is comparable to error rates seen with the Illumina systems [104]. For both long-read sequencing solutions, the company providers are working on further reducing their current error rates, which will likely be achieved in the years to come.

4.4. Reference databases

In both, targeted MPS and WGS metagenomics, mistakes in the taxonomy assignment can occur due to sequencing errors and/or incorrect labelling [105]. Besides, commonly used microbiome databases are biased towards the presence of clinically relevant microorganisms for humans, with discrepancies observed in environmental microorganisms depending on the reference database used [106]. This is relevant for forensic applications dealing with samples exposed to diverse environmental conditions, where not only medically-relevant microorganisms are expected.

In targeted MPS, mistakes in the taxonomy assignment can occur due to errors during the PCR amplification step prior to sequencing [105]. Also, alignments by certain databases (i.e. Greengenes) are sometimes of poor quality, resulting in artificially inflated richness and diversity estimates [107]. It has been shown that the use of niche-specific databases (e.g. the expanded Human Oral Microbiome Database (eHOMD) [108]) leads to an increase in lower taxonomic assignments [109], which can result in higher taxonomic resolution. Nonetheless, care should be taken when using these niche-specific databases, since they might not contain all the microorganisms included in, for instance, commercial mock

communities that are commonly used to assess bias during PCR, library preparation, sequencing and initial analysis steps. An alternative would be to compare the mock community with a larger database or use an in-house mock sample that better reflects the microorganisms of interest.

Regarding WGS metagenomics, reference databases containing all domains of life provide the most suitable option rather than databases comprised of a single domain, since the later can lead to an unacceptably high number of false-positive assignations [110]. However, a WGS metagenomics database containing all reference genomes that may exist in a forensic sample does not (yet) exist. In consequence, researchers potentially fail to identify key microorganisms within their samples by use of WGS metagenomics data given the limitations of the available reference databases [111]. Researchers should be aware of the discrepancies across existing databases, which can have a substantial impact on the number and type of microorganisms identified in a sample (or an entire study) depending on the reference database of choice [111,112].

4.5. Prediction algorithms

Advances in MPS technologies coincided with improvements in bioinformatics data analysis methods such as machine learning (ML). Among other applications of forensic relevance, this potentially allows the prediction of the host's lifestyle or other traits from microbiome data [113], which can provide useful information to find unknown perpetrators, for instance together with human DNA-based prediction of appearance, bio-geographic ancestry and age [114]. Several authors including us have proven the promise of ML methods in forensic microbiome-based applications [12,13,24,30]. However, one major issue is the lack of generalizable prediction methods applicable to microbiome datasets other than the study-specific one. As an example, considering two studies with the same research question, the total number of microorganisms detected in study one is unlikely to be identified in study two due to biological and/or technical variation. This leads to incomparability issues of different studies so that a model based on data from study one cannot be applied to data from study two, because of missing data. Moreover, the total number of shared taxa between studies is expected to be even smaller when low taxonomic levels are analysed (e.g. species level). One way to solve this is to only select the taxa observed across the different studies for analysis, though this can reduce the prediction accuracy as it ignores informative taxa in each separate dataset.

Moreover, forensic microbiome studies should put more emphasis on the detailed and correct report of the ML pipeline used, e.g., describing reasons for the ML of choice, variation in the performance of different folds of cross-validation as well as between validation and test sets [115]. Ideally, these reports should also allow interpretation by non-experts in the field of ML. This can ultimately help other researchers to better assess the most appropriate ML method for a given application and to create awareness about the challenges and limitations encountered with different data types.

Additionally, if not used for intelligence purposes only, outcomes from the ML algorithms should be adjusted to the rigorous testing framework used in forensic investigations (e.g. likelihood ratios) in order to be accepted in court. To this, the statistical challenges intrinsic to mixed samples from different persons should be added. This is already highly complex when analysing mixed human DNA profiles of two or a few individuals and might be much more tedious in microbiome profiles consisting of dozens or hundreds of microorganisms, which is a pending task for the forensic microbiome research community.

5. Data issues and potential solutions

Current forensic microbiome research also faces data-related challenges that should be considered in future studies and eventually solved (Fig. 3). Most previous forensic microbiome studies suffered from (too) small sample size, partly caused by hurdles associated with data production that can be of ethical, administrative, availability and financial nature. A potential solution is to bypass the raw data production phase and directly make use of the large amounts of microbiome data deposited in public repositories, if such data are suitable to address the forensic study aim; e.g., the use of publicly available human microbiome data as training dataset in prediction modelling for tissue/body fluid identification like we did [12,13]. Over the recent years, publicly available human microbiome data have increased exponentially in number and size, especially those originating from large consortia, such as the HMP [8] and the American Gut Project (AGP) [116]. However, apart from the technical challenges discussed before, many other issues related to the data themselves must be considered too, such as the uneven microbiome research on different body sites based on their clinical relevance, which limits the application of public human microbiome data for non-medical purposes such as in forensics.

To illustrate this, we refer to the review by Proctor et al. [117] on the microbiome research carried out at the United States National Institutes of Health (NIH) during a 10 year period (2007–2016), which is a good proxy of the general human microbiome literature. According to this review, three quarters of the research has focused on just four body sites: gastrointestinal (GI) tract, urogenital tract (primarily vaginal), oral cavity and lung. From those, GI tract corresponds to 40% of the total research, which mainly analysed stool samples, while skin and nares microbiomes each represented only 3% of the total research output. The remaining body sites, tissues and systems include ear, eye, liver, blood, cardiovascular system and central nervous system. Hence, public microbiome data repositories typically lack data from forensically relevant tissues and body fluids, such as seminal fluid, venous blood or menstrual blood.

Additionally, samples included in forensic validation studies (e.g. mock casework samples) are prepared, stored and collected in very specific ways, that do not resemble those of the samples normally included in public repositories. As a result, in many cases, forensic microbiome researchers do not have any other option than produce raw microbiome data themselves, which requires availability of samples in large enough quantities as well as the resources for sample analysis. It is worth mentioning that the forensic community will enormously benefit from multi-collaborative efforts that allow studies of large sample size and the subsequent associated microbiome data deposition in public repositories, such as the Critical Assessment of Massive Data Analysis (CAMDA) of Metagenomics and Metadesign of Subways and Urban Biomes (MetaSUB) 'Forensic Challenge' [118].

Moreover, most forensic microbiome studies as well as the publicly available microbiome data from large consortia have focused on studying samples obtained from individuals living in the USA and Europe. Whether their findings can be generalized to other populations is still to be determined [119]. More recently, studies are emerging with study subjects from other countries with similar industrialization profiles and lifestyle including China, Japan and Israel, although there is a clear underrepresentation of populations in Africa, South America and other regions in Asia (e.g. India). These underrepresented regions potentially possess distinct genetic [120,121], ethnic [122,123], sociocultural [124,125] and lifestyle [126,127] backgrounds that shape their microbiome. Other factors, such as the donors' age, are also unevenly distributed across the samples described in the microbiome literature. From a forensic standpoint, it is important to have microbiome information from as many diverse human samples as possible in order to assess the generalization of any proposed investigative tool. Future research needs to establish whether such human population differences impact on the forensic microbiome applications.

Finally, although not specific only to human microbiome research, another aspect that we would like to point out is the high variability in metadata between studies, especially when publicly available microbiome data are used in future forensic research. Currently available

metadata is often misannotated, misleading and non-standardized, which makes sample reanalysis overly complex [128]. Importantly, metadata does not only refer to sample attributes, but also to sample preparation and processing [128]. The variability of metadata can limit the conclusions, since the observed patterns might not be attributable to a certain factor(s) [129,130] that should also be considered when selecting study subjects. In this context, we call attention to the necessity of validating metadata submitted to public repositories by providing immediate and informative feedback to the submitting researchers [128].

6. Future outlook

Although forensic microbiome analysis is still in its infancy, it has been demonstrated already by us and others that with further research and appropriate validation the use of the human microbiome for forensic purposes holds great promises. Once the current knowledge, technology and data issues are solved, the human microbiome can be applied as a very advantageous forensic tool, especially where other human (non-microbial) approaches present limitations. With this, we do not imply that other approaches should be replaced by microbiome analysis, but instead whenever possible to join forces with such other approaches aiming to answer challenging but relevant forensic questions. For instance, bacterial DNA seems to be more resistant to harsh environmental conditions (i.e. chemical and physical agents) compared to human DNA due to the circular nature of the bacterial DNA molecule and its localization within a cell furnished with a wall of peptidoglycan matrix [87]. We have previously shown that a human microbiome-based approach can perform better than human RNA-based methods for the tissue/body fluid identification of challenging samples, such as for skin traces deposited on 'touched' objects [12] and for body fluids of a complex nature like menstrual blood [13].

Further research is also needed on the impact the timespan between trace deposition at the crime scene and collection of a reference sample from a known suspect has on the microbiome outcome. Intra-individual variation in the microbial community between the two time points (e.g. based on change in individual's lifestyle habits, disease onset, circadian rhythms in the microbial community) could make the comparison between crime scene trace and reference samples difficult or unreliable. Also, the magnitude of the inter-individual variation on the application's generalizability should be further assessed and aimed to be solved for instance at the data or analysis level. Additionally, the timewise stability or instability of the microbiome ex vivo that may be desired depending on the forensic application should be further investigated.

Moreover, the forensic microbiome community should consider testing multi-omics approaches for answering relevant forensic questions. Studying 'what organisms are present' in a sample via DNA-based analysis is not always enough to capture the entire complexity of the human microbiome. It is known that microbial community assembly is primarily mediated by functional niches rather than a requirement for specific microorganisms [1,131]. Therefore, analysing 'what the microorganisms are doing' from metatranscriptomics [132], metaproteomics [133] and metabolomics [134] can help in adding new dimensions to address relevant forensic questions. Combining different microbial-derived omics data will also be important to avoid wrong conclusions. For example, some bacteria from different species perform the same function (e.g. lactic acid production in vagina), which based on metabolomics data would look the same, but based on DNA data might look different.

Lastly, the forensic microbiome community should overcome the current lack of experimental, analytical and (meta)data standardization. This will ultimately make forensic microbiome data findable, accessible, interoperable and reusable (FAIR) [135] and boost the forensic microbiome research and applications. In the USA, current multi-collaborative efforts such as the Microbiome Quality Control project (MBCQ) [136] and the International Human Microbiome Standards group (IHMS) [137] focus on the comprehensive evaluation of microbiome methods to promote the best practices through the field and increase the comparability of the results between studies. In the European Union, the Machine Learning for Microbiome 'ML4M' action aims to optimise and standardise the ML and statistical methods used in microbiome analysis and create publicly available benchmark datasets [138]. The forensic microbiome community would strongly benefit from standardised microbiome approaches to further optimise and validate them based on their specific needs for application in the forensic context.

Overall, we are optimistic that the current hurdles in forensic microbiome analysis outlined here will be overcome by further scientific and technical progress along the lines we proposed here, so that human microbiome analysis will become an integral part of the forensic toolkit in the not too distant future.

Conflict of interest

The authors declare not to have any conflicts of interest.

Acknowledgements

The authors would like to thank their colleagues from the forensic and non-forensic human microbiome communities for their previously published work and the datasets they made publicly available that led to the current achievements discussed here. The authors' work is supported by Erasmus MC, University Medical Center Rotterdam.

References

[1] P.J. Turnbaugh, R.E. Ley, M. Hamady, C.M. Fraser-Liggett, R. Knight, J.I. Gordon, The human microbiome project, Nature 449 (2007) 804–810, https://doi.org/10.1038/nature06244.
[2] J.R. Marchesi, J. Ravel, The vocabulary of microbiome research: a proposal, Microbiome 3 (2015) 31, https://doi.org/10.1186/s40168-015-0094-5.
[3] A.H. Moeller, Y. Li, E. Mpoudi Ngole, S. Ahuka-Mundeke, E.V. Lonsdorf, A. E. Pusey, M. Peeters, B.H. Hahn, H. Ochman, Rapid changes in the gut microbiome during human evolution, Proc. Natl. Acad. Sci. USA 111 (2014) 16431–16435, https://doi.org/10.1073/pnas.1419136111.
[4] J. Zaneveld, P.J. Turnbaugh, C. Lozupone, R.E. Ley, M. Hamady, J.I. Gordon, R. Knight, Host-bacterial coevolution and the search for new drug targets, Curr. Opin. Chem. Biol. 12 (2008) 109–114, https://doi.org/10.1016/j.cbpa.2008.01.015.
[5] R.E. Ley, M. Hamady, C. Lozupone, P.J. Turnbaugh, R.R. Ramey, J.S. Bircher, M. L. Schlegel, T.A. Tucker, M.D. Schrenzel, R. Knight, Evolution of mammals and their gut microbes, Science 320 (2008) 1647–1651.
[6] R. Sender, S. Fuchs, R. Milo, Revised estimates for the number of human and bacteria cells in the body, PLoS Biol. 14 (2016), e1002533, https://doi.org/10.1371/journal.pbio.1002533.
[7] The Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome, Nature 486 (2012) 207–214, https://doi.org/10.1038/nature11234.
[8] The Human Microbiome Project Consortium, A framework for human microbiome research, Nature 486 (2012) 215–221, https://doi.org/10.1038/nature11209.
[9] S.D. Ehrlich, The MetaHIT Consortium, MetaHIT: The European Union Project on metagenomics of the human intestinal tract. Metagenomics of the Human Body, Springer, 2011, pp. 307–316, https://doi.org/10.1007/978-1-4419-7089-3_15.
[10] J. Yang, T. Tsukimi, M. Yoshikawa, K. Suzuki, T. Takeda, M. Tomita, S. Fukuda, Cutibacterium acnes (Propionibacterium acnes) 16S rRNA genotyping of microbial samples from possessions contributes to owner identification, mSystems 4 (2019) e00594–19, https://doi.org/10.1128/mSystems.00594-19.
[11] A. Belk, Z.Z. Xu, D.O. Carter, A. Lynne, S. Bucheli, R. Knight, J.L. Metcalf, Microbiome data accurately predicts the postmortem interval using random forest regression models, Genes 9 (2018) 104, https://doi.org/10.3390/genes9020104.
[12] C. Díez López, A. Vidaki, A. Ralf, D. Montiel González, D. Radjabzadeh, R. Kraaij, A.G. Uitterlinden, C. Haas, O. Lao, M. Kayser, Novel taxonomy-independent deep learning microbiome approach allows for accurate classification of different forensically relevant human epithelial materials, Forensic Sci. Int. Genet. 41 (2019) 72–82, https://doi.org/10.1016/j.fsigen.2019.03.015.
[13] C. Díez López, D. Montiel González, C. Haas, A. Vidaki, M. Kayser, Microbiome-based body site of origin classification of forensically relevant blood traces, Forensic Sci. Int. Genet. 47 (2020), 102280, https://doi.org/10.1016/j.fsigen.2020.102280.
[14] R. Blekhman, J.K. Goodrich, K. Huang, Q. Sun, R. Bukowski, J.T. Bell, T. D. Spector, A. Keinan, R.E. Ley, D. Gevers, A.G. Clark, Host genetic variation impacts microbiome composition across human body sites, Genome Biol. 16 (2015) 191, https://doi.org/10.1186/s13059-015-0759-1.

[15] C. Dominianni, R. Sinha, J.J. Goedert, Z.H. Pei, L.Y. Yang, R.B. Hayes, J.Y. Ahn, Sex, body mass index, and dietary fiber intake influence the human gut microbiome, PLoS One 10 (2015), e0124599, https://doi.org/10.1371/journal.pone.0124599.
[16] J.M. Fettweis, J.P. Brooks, M.G. Serrano, N.U. Sheth, P.H. Girerd, D.J. Edwards, J. F. Strauss, K.K. Jefferson, G.A. Buck, V.M. Consortium, Differences in vaginal microbiome in African American women versus women of European ancestry, Microbiology 160 (2014) 2272–2282, https://doi.org/10.1099/mic.0.081034-0.
[17] V. Monda, I. Villano, A. Messina, A. Valenzano, T. Esposito, F. Moscatelli, A. Viggiano, G. Cibelli, S. Chieffi, M. Monda, G. Messina, Exercise modifies the gut microbiota with positive health effects, Oxid. Med. Cell. Longev. 2017 (2017) 3831972–3831978, https://doi.org/10.1155/2017/3831972.
[18] S. Aditya, A.K. Sharma, C.N. Bhattacharyya, K. Chaudhuri, Generating STR profile from "Touch DNA", J. Forensic Leg. Med. 18 (2011) 295–298, https://doi.org/10.1016/j.jflm.2011.05.007.
[19] D. Wilkins, M.H.Y. Leung, P.K.H. Lee, Microbiota fingerprints lose individually identifying features over time, Microbiome 5 (2017) 1, https://doi.org/10.1186/s40168-016-0209-7.
[20] J.F. Meadow, A.E. Altrichter, J.L. Green, Mobile phones carry the personal microbiome of their owners, PeerJ 2 (2014), e447, https://doi.org/10.7717/peerj.447.
[21] S. Lax, J.T. Hampton-Marcell, S.M. Gibbons, G.B. Colares, D. Smith, J.A. Eisen, J. A. Gilbert, Forensic analysis of the microbiome of phones and shoes, Microbiome 3 (2015) 21, https://doi.org/10.1186/s40168-015-0082-9.
[22] N. Fierer, C.L. Lauber, N. Zhou, D. McDonald, E.K. Costello, R. Knight, Forensic identification using skin bacterial communities, Proc. Natl. Acad. Sci. USA 107 (2010) 6477–6481, https://doi.org/10.1073/pnas.1000162107.
[23] S.L. Leake, M. Pagni, L. Falquet, F. Taroni, G. Greub, The salivary microbiome for differentiating individuals: proof of principle, Microbes Infect. 18 (2016) 399–405, https://doi.org/10.1016/j.micinf.2016.03.011.
[24] D. Williams, G. Gibson, Individualization of pubic hair bacterial communities and the effects of storage time and temperature, Forensic Sci. Int. Genet. 26 (2017) 9–20, https://doi.org/10.1016/j.fsigen.2016.09.006.
[25] S.R. Tridico, D.C. Murray, J. Addison, K.P. Kirkbride, M. Bunce, Metagenomic analyses of bacteria on human hairs: a qualitative assessment for applications in forensic science, Invest. Genet. 5 (2014) 16, https://doi.org/10.1186/s13323-014-0016-5.
[26] G.T. Javan, S.J. Finley, I. Can, J.E. Wilkinson, J.D. Hanson, A.M. Tarone, Human thanatomicrobiome succession and time since death, Sci. Rep. 6 (2016) 29598, https://doi.org/10.1038/srep29598.
[27] I. Can, G.T. Javan, A.E. Pozhitkov, P.A. Noble, Distinctive thanatomicrobiome signatures found in the blood and internal organs of humans, J. Microbiol. Methods 106 (2014) 1–7, https://doi.org/10.1016/j.mimet.2014.07.026.
[28] F.E. Damann, D.E. Williams, A.C. Layton, Potential use of bacterial community succession in decaying human bone for estimating postmortem interval, J. Forensic Sci. 60 (2015) 844–850, https://doi.org/10.1111/1556-4029.12744.
[29] K.A. Hauther, K.L. Cobaugh, L.M. Jantz, T.E. Sparer, J.M. DeBruyn, Estimating time since death from postmortem human gut microbial communities, J. Forensic Sci. 60 (2015) 1234–1240, https://doi.org/10.1111/1556-4029.12828.
[30] H.R. Johnson, D.D. Trinidad, S. Guzman, Z. Khan, J.V. Parziale, J.M. DeBruyn, N. H. Lents, A machine learning approach for using the postmortem skin microbiome to estimate the postmortem interval, PLoS One 11 (2016), e0167370, https://doi.org/10.1371/journal.pone.0167370.
[31] E.R. Hyde, D.P. Haarmann, J.F. Petrosino, A.M. Lynne, S.R. Bucheli, Initial insights into bacterial succession during human decomposition, Int. J. Leg. Med. 129 (2015) 661–671, https://doi.org/10.1007/s00414-014-1128-4.
[32] J.L. Metcalf, Z.Z. Xu, S. Weiss, S. Lax, W. Van Treuren, E.R. Hyde, S.J. Song, A. Amir, P. Larsen, N. Sangwan, Microbial community assembly and metabolic function during mammalian corpse decomposition, Science 351 (2016) 158–162, https://doi.org/10.1126/science.aad2646.
[33] T.A. Suzuki, M. Worobey, Geographical variation of human gut microbial composition, Biol. Lett. 10 (2014), 20131037, https://doi.org/10.1098/rsbl.2013.1037.
[34] J.S. Escobar, B. Klotz, B.E. Valdes, G.M. Agudelo, The gut microbiota of Colombians differs from that of Americans, Europeans and Asians, BMC Microbiol. 14 (2014) 311, https://doi.org/10.1186/s12866-014-0311-6.
[35] S. Ying, D.-N. Zeng, L. Chi, Y. Tan, C. Galzote, C. Cardona, S. Lax, J. Gilbert, Z.-X. Quan, The influence of age and gender on skin-associated microbial communities in urban and rural human populations, PLoS One 10 (2015), e0141842, https://doi.org/10.1371/journal.pone.0141842.
[36] A.V. Tyakht, E.S. Kostryukova, A.S. Popenko, M.S. Belenikin, A.V. Pavlenko, A. K. Larin, I.Y. Karpova, O.V. Selezneva, T.A. Semashko, E.A. Ospanova, Human gut microbiota community structures in urban and rural populations in Russia, Nat. Commun. 4 (2013) 1–9, https://doi.org/10.1038/ncomms3469.
[37] M.J. Blaser, M.G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, I. Estrada, Z. Gao, J.C. Clemente, E.K. Costello, R. Knight, Distinct cutaneous bacterial assemblages in a sampling of South American Amerindians and US residents, ISME J. 7 (2013) 85–95, https://doi.org/10.1038/ismej.2012.81.
[38] J. Li, D. Quinque, H.P. Horz, M.K. Li, M. Rzhetskaya, J.A. Raff, M.G. Hayes, M. Stoneking, Comparative analysis of the human saliva microbiome from different climate zones: Alaska, Germany, and Africa, BMC Microbiol. 14 (2014) 316, https://doi.org/10.1186/s12866-014-0316-1.
[39] G.E. Miller, P.A. Engen, P.M. Gillevet, M. Shaikh, M. Sikaroodi, C.B. Forsyth, E. Mutlu, A. Keshavarzian, Lower neighborhood socioeconomic status associated with reduced diversity of the colonic microbiota in healthy adults, PLoS One 11 (2016), e0148952, https://doi.org/10.1371/journal.pone.0148952.
[40] K.M. Hewitt, C.P. Gerba, S.L. Maxwell, S.T. Kelley, Office space bacterial abundance and diversity in three metropolitan areas, PLoS One 7 (2012), e37849, https://doi.org/10.1371/journal.pone.0037849.
[41] J. Chase, J. Fouquier, M. Zare, D.L. Sonderegger, R. Knight, S.T. Kelley, J. Siegel, J.G. Caporaso, Geography and location are the primary drivers of office microbiome composition, mSystems 1 (2016) e00022–16, https://doi.org/10.1128/mSystems.00022-16.
[42] L. Huang, C. Xu, W. Yang, R. Yu, A machine learning framework to determine geolocations from metagenomic profiling, Biol. Direct 15 (2020) 27, https://doi.org/10.1186/s13062-020-00278-z.
[43] S.L. McNulty, B.M. Mole, D. Dailidiene, I. Segal, R. Ally, R. Mistry, O. Secka, R. A. Adegbola, J.E. Thomas, E.M. Lenarcic, R.M. Peek Jr., D.E. Berg, M.H. Forsyth, Novel 180- and 480-base-pair insertions in African and African-American strains of Helicobacter pylori, J. Clin. Microbiol. 42 (2004) 5658–5663, https://doi.org/10.1128/jcm.42.12.5658-5663.2004.
[44] D. Kersulyte, A. Kalia, R.H. Gilman, M. Mendez, P. Herrera, L. Cabrera, B. Velapatiño, J. Balqui, F. Paredes Puente de la Vega, C.A. Rodriguez Ulloa, J. Cok, C.C. Hooper, G. Dailide, S. Tamma, D.E. Berg, Helicobacter pylori from Peruvian Amerindians: traces of human migrations in strains from remote Amazon, and genome sequence of an amerind strain, PLoS One 5 (2010), e15076, https://doi.org/10.1371/journal.pone.0015076.
[45] S. Nagasawa, H. Motani-Saitoh, H. Inoue, H. Iwase, Geographic diversity of Helicobacter pylori in cadavers: forensic estimation of geographical origin, Forensic Sci. Int. Genet. 229 (2013) 7–12, https://doi.org/10.1016/j.forsciint.2013.02.028.
[46] E.K. Costello, C.L. Lauber, M. Hamady, N. Fierer, J.I. Gordon, R. Knight, Bacterial community variation in human body habitats across space and time, Science 326 (2009) 1694–1697, https://doi.org/10.1126/science.1177486.
[47] C.A. Lozupone, M. Hamady, S.T. Kelley, R. Knight, Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities, Appl. Environ. Microbiol. 73 (2007) 73–85, https://doi.org/10.1128/aem.01996-06.
[48] J.G. Caporaso, C.L. Lauber, E.K. Costello, D. Berg-Lyons, A. Gonzalez, J. Stombaugh, D. Knights, P. Gajer, J. Ravel, N. Fierer, J.I. Gordon, R. Knight, Moving pictures of the human microbiome, Genome Biol. 12 (2011) R50, https://doi.org/10.1186/gb-2011-12-5-r50.
[49] T. Akutsu, H. Motani, K. Watanabe, H. Iwase, K. Sakurada, Detection of bacterial 16S ribosomal RNA genes for forensic identification of vaginal fluid, Leg. Med. 14 (2012) 160–162, https://doi.org/10.1016/j.legalmed.2012.01.005.
[50] C.C.G. Benschop, F.C.A. Quaak, M.E. Boon, T. Sijen, I. Kuiper, Vaginal microbial flora analysis by next generation sequencing and microarrays; can microbes indicate vaginal origin in a forensic context? Int. J. Leg. Med. 126 (2012) 303–310, https://doi.org/10.1007/s00414-011-0660-8.
[51] S. Giampaoli, A. Berti, F. Valeriani, G. Gianfranceschi, A. Piccolella, L. Buggiotti, C. Rapone, A. Valentini, L. Ripani, V. Romano Spica, Molecular identification of vaginal fluid by microbial signature, Forensic Sci. Int. Genet. 6 (2012) 559–564, https://doi.org/10.1016/j.fsigen.2012.01.005.
[52] S. Giampaoli, F. Alessandrini, A. Berti, L. Ripani, A. Choi, R. Crab, E. De Vittori, B. Egyed, C. Haas, H.Y. Lee, M. Korabecná, F. Noel, D. Podini, A. Tagliabracci, A. Valentini, V. Romano Spica, Forensic interlaboratory evaluation of the ForFLUID kit for vaginal fluids identification, J. Forensic Leg. Med. 21 (2014) 60–63, https://doi.org/10.1016/j.jflm.2013.10.016.
[53] H. Nakanishi, A. Kido, T. Ohmori, A. Takada, M. Hara, N. Adachi, K. Saito, A novel method for the identification of saliva by detecting oral streptococci using PCR, Forensic Sci. Int. Genet. 183 (2009) 20–23, https://doi.org/10.1016/j.forsciint.2008.10.003.
[54] A.E. Donaldson, M.C. Taylor, S.J. Cordiner, I.L. Lamont, Using oral microbial DNA analysis to identify expirated bloodspatter, Int. J. Leg. Med. 124 (2010) 569–576, https://doi.org/10.1007/s00414-010-0426-8.
[55] D.A. Power, S.J. Cordiner, J.A. Kieser, G.R. Tompkins, J. Horswell, PCR-based detection of salivary bacteria as a marker of expirated blood, Sci. Justice 50 (2010) 59–63, https://doi.org/10.1016/j.scijus.2009.04.006.
[56] C. Díez López, M. Kayser, A. Vidaki, Estimating the time since deposition of saliva stains with a targeted bacterial DNA approach: a proof-of-principle study, Front. Microbiol. 12 (2021), 647933, https://doi.org/10.3389/fmicb.2021.647933.
[57] A.P. Salzmann, N. Arora, G. Russo, S. Kreutzer, L. Snipen, C. Haas, Assessing time dependent changes in microbial composition of biological crime scene traces using microbial RNA markers, Forensic Sci. Int. Genet. 53 (2021), 102537, https://doi.org/10.1016/j.fsigen.2021.102537.
[58] A. Lucci, C.P. Campobasso, A. Cirnelli, G. Lorenzini, A promising microbiological test for the diagnosis of drowning, Forensic Sci. Int. Genet. 182 (2008) 20–26, https://doi.org/10.1016/j.forsciint.2008.09.004.
[59] M. Suto, N. Kato, S. Abe, M. Nakamura, R. Tsuchiya, K. Hiraiwa, PCR detection of bacterial genes provides evidence of death by drowning, Leg. Med. 11 (Suppl. 1) (2009) S354–S356, https://doi.org/10.1016/j.legalmed.2009.01.062.
[60] S.F. Kaszubinski, J.L. Pechal, K. Smiles, C.J. Schmidt, H.R. Jordan, M.H. Meek, M. E. Benbow, Dysbiosis in the dead: human postmortem microbiome beta-dispersion as an indicator of manner and cause of death, Front. Microbiol 11 (2020) 2212, https://doi.org/10.3389/fmicb.2020.555347.
[61] A. Gunn, S.J. Pitt, Microbes as forensic indicators, Trop. Biomed. 29 (2012) 311–330.
[62] C.M. Black, E.M. Driebe, L.A. Howard, N.N. Fajman, M.K. Sawyer, R.G. Girardet, R.L. Sautter, E. Greenwald, C.M. Beck-Sague, E.R. Unger, J.U. Igietseme, M. R. Hammerschlag, Multicenter study of nucleic acid amplification tests for detection of Chlamydia trachomatis and Neisseria gonorrhoeae in children being

evaluated for sexual abuse, Pedia Infect. Dis. J. 28 (2009) 608–613, https://doi.org/10.1097/INF.0b013e31819b592e.
[63] I.M.C. Martin, E. Foreman, V. Hall, A. Nesbitt, G. Forster, C.A. Ison, Non-cultural detection and molecular genotyping of Neisseria gonorrhoeae from a piece of clothing, J. Med. Microbiol. 56 (2007) 487–490, https://doi.org/10.1099/jmm.0.46956-0.
[64] R.E. Ley, M. Hamady, C. Lozupone, P.J. Turnbaugh, R.R. Ramey, J.S. Bircher, M. L. Schlegel, T.A. Tucker, M.D. Schrenzel, R. Knight, J.I. Gordon, Evolution of mammals and their gut microbes, Science 320 (2008) 1647–1651, https://doi.org/10.1126/science.1155725.
[65] J. Wu, B.A. Peters, C. Dominianni, Y.L. Zhang, Z.H. Pei, L.Y. Yang, Y.F. Ma, M. P. Purdue, E.J. Jacobs, S.M. Gapstur, H.L. Li, A.V. Alekseyenko, R.B. Hayes, J. Y. Ahn, Cigarette smoking and the oral microbiome in a large study of American adults, ISME J. 10 (2016) 2435–2446, https://doi.org/10.1038/ismej.2016.37.
[66] J.F. Meadow, A.C. Bateman, K.M. Herkert, T.K. O'Connor, J.L. Green, Significant changes in the skin microbiome mediated by the sport of roller derby, PeerJ 1 (2013), e53, https://doi.org/10.7717/peerj.53.
[67] J. Qian, D. Hospodsky, N. Yamamoto, W.W. Nazaroff, J. Peccia, Size-resolved emission rates of airborne bacteria and fungi in an occupied classroom, Indoor Air 22 (2012) 339–351, https://doi.org/10.1111/j.1600-0668.2012.00769.x.
[68] Y. Zhou, H. Gao, K.A. Mihindukulasuriya, P.S. La Rosa, K.M. Wylie, T. Vishnivetskaya, M. Podar, B. Warner, P.I. Tarr, D.E. Nelson, Biogeography of the ecosystems of the healthy human body, Genome Biol. 14 (2013) 1–18, https://doi.org/10.1186/gb-2013-14-1-r1.
[69] E.A. Franzosa, K. Huang, J.F. Meadow, D. Gevers, K.P. Lemon, B.J.M. Bohannan, C. Huttenhower, Identifying personal microbiomes using metagenomic codes, Proc. Natl. Acad. Sci. USA 112 (2015) E2930–E2938, https://doi.org/10.1073/pnas.1423854112.
[70] M.J. Blaser, Harnessing the power of the human microbiome, Proc. Natl. Acad. Sci. USA 107 (2010) 6125–6126, https://doi.org/10.1073/pnas.1002112107.
[71] R.I. Fleming, S. Harbison, The use of bacteria for the identification of vaginal secretions, Forensic Sci. Int. Genet. 4 (2010) 311–315, https://doi.org/10.1016/j.fsigen.2009.11.008.
[72] J. Ravel, P. Gajer, Z. Abdo, G.M. Schneider, S.S. Koenig, S.L. McCulle, S. Karlebach, R. Gorle, J. Russell, C.O. Tacket, R.M. Brotman, C.C. Davis, K. Ault, L. Peralta, L.J. Forney, Vaginal microbiome of reproductive-age women, Proc. Natl. Acad. Sci. USA 108 (Suppl. 1) (2011) S4680–S4687, https://doi.org/10.1073/pnas.1002611107.
[73] V. Lazarevic, K. Whiteson, D. Hernandez, P. François, J. Schrenzel, Study of inter- and intra-individual variations in the salivary microbiota, BMC Genom. 11 (2010) 523, https://doi.org/10.1186/1471-2164-11-523.
[74] L. Takayasu, W. Suda, K. Takanashi, E. Iioka, R. Kurokawa, C. Shindo, Y. Hattori, N. Yamashita, S. Nishijima, K. Oshima, M. Hattori, Circadian oscillations of microbial and functional composition in the human salivary microbiome, DNA Res. 24 (2017) 261–270, https://doi.org/10.1093/dnares/dsx001.
[75] A. Dobay, C. Haas, G. Fucile, N. Downey, H.G. Morrison, A. Kratzer, N. Arora, Microbiome-based body fluid identification of samples exposed to indoor conditions, Forensic Sci. Int. Genet. 40 (2019) 105–113, https://doi.org/10.1016/j.fsigen.2019.02.010.
[76] C. Marotz, J.T. Morton, P. Navarro, J. Coker, P. Belda-Ferre, R. Knight, K. Zengler, Quantifying live microbial load in human saliva samples over time reveals stable composition and dynamic load, mSystems 6 (2021) e01182–20, https://doi.org/10.1128/mSystems.01182-20.
[77] F. Teng, S.S. Darveekaran Nair, P. Zhu, S. Li, S. Huang, X. Li, J. Xu, F. Yang, Impact of DNA extraction method and targeted 16S-rRNA hypervariable region on oral microbiota profiling, Sci. Rep. 8 (2018) 16321, https://doi.org/10.1038/s41598-018-34294-x.
[78] L. Abusleme, B.-Y. Hong, A.K. Dupuy, L.D. Strausbaugh, P.I. Diaz, Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing, J. Oral Microbiol. 6 (2014) 23990, https://doi.org/10.3402/jom.v6.23990.
[79] R.D. Bjerre, L.W. Hugerth, F. Boulund, M. Seifert, J.D. Johansen, L. Engstrand, Effects of sampling strategy and DNA extraction on human skin microbiome investigations, Sci. Rep. 9 (2019) 17287, https://doi.org/10.1038/s41598-019-53599-z.
[80] P. Videnska, K. Smerkova, B. Zwinsova, V. Popovici, L. Micenkova, K. Sedlar, E. Budinska, Stool sampling and DNA isolation kits affect DNA quality and bacterial composition following 16S rRNA gene sequencing using MiSeq Illumina platform, Sci. Rep. 9 (2019) 13837, https://doi.org/10.1038/s41598-019-49520-3.
[81] J. Pollock, L. Glendinning, T. Wisedchanwet, M. Watson, The madness of microbiome: attempting to find consensus "best practice" for 16S microbiome studies, Appl. Environ. Microbiol. 84 (2018) e02627–17, https://doi.org/10.1128/AEM.02627-17.
[82] K. Gerasimidis, M. Bertz, C. Quince, K. Brunner, A. Bruce, E. Combet, S. Calus, N. Loman, U.Z. Ijaz, The effect of DNA extraction methodology on gut microbiota research applications, BMC Res. Notes 9 (2016) 1–10, https://doi.org/10.1186/s13104-016-2171-7.
[83] C. Schrader, A. Schielke, L. Ellerbroek, R. Johne, PCR inhibitors–occurrence, properties and removal, J. Appl. Microbiol. 113 (2012) 1014–1026, https://doi.org/10.1111/j.1365-2672.2012.05384.x.
[84] J. Rosenbaum, M. Usyk, Z. Chen, C.P. Zolnik, H.E. Jones, L. Waldron, J.B. Dowd, L.E. Thorpe, R.D. Burk, Evaluation of oral cavity DNA extraction methods on bacterial and fungal microbiota, Sci. Rep. 9 (2019) 1531, https://doi.org/10.1038/s41598-018-38049-6.
[85] Y. Lim, M. Totsika, M. Morrison, C. Punyadeera, The saliva microbiome profiles are minimally affected by collection method or DNA extraction protocols, Sci. Rep. 7 (2017) 8523, https://doi.org/10.1038/s41598-017-07885-3.
[86] L. Karstens, N.Y. Siddiqui, T. Zaza, A. Barstad, C.L. Amundsen, T.A. Sysoeva, Benchmarking DNA isolation kits used in analyses of the urinary microbiome, Sci. Rep. 11 (2021) 6186, https://doi.org/10.1038/s41598-021-85482-1.
[87] T. Pamela, G. D'Angiolella, P. Brun, I. Castagliuolo, S. Gino, L. Caenazzo, Skin microbiome analysis for forensic human identification: what do we know so far? Microorganisms 8 (2020) 873, https://doi.org/10.3390/microorganisms8060873.
[88] F. Guo, T. Zhang, Biases during DNA extraction of activated sludge samples revealed by high throughput sequencing, Appl. Microbiol. Biotechnol. 97 (2013) 4607–4616, https://doi.org/10.1007/s00253-012-4244-4.
[89] J. Maukonen, C. Simões, M. Saarela, The currently used commercial DNA-extraction methods give different results of clostridial and actinobacterial populations derived from human fecal samples, FEMS Microbiol. Ecol. 79 (2012) 697–708, https://doi.org/10.1111/j.1574-6941.2011.01257.x.
[90] B.E. Knudsen, L. Bergmark, P. Munk, O. Lukjancenko, A. Priemé, F.M. Aarestrup, S.J. Pamp, Impact of sample type and DNA isolation procedure on genomic inference of microbiome composition, mSystems 1 (2016) e00095–16, https://doi.org/10.1128/mSystems.00095-16.
[91] S.J. Salter, M.J. Cox, E.M. Turek, S.T. Calus, W.O. Cookson, M.F. Moffatt, P. Turner, J. Parkhill, N.J. Loman, A.W. Walker, Reagent and laboratory contamination can critically impact sequence-based microbiome analyses, BMC Biol. 12 (2014) 87, https://doi.org/10.1186/s12915-014-0087-z.
[92] L.S. Weyrich, A.G. Farrer, R. Eisenhofer, L.A. Arriola, J. Young, C.A. Selway, M. Handsley-Davis, C.J. Adler, J. Breen, A. Cooper, Laboratory contamination over time during low-biomass sample analysis, Mol. Ecol. Resour. 19 (2019) 982–996, https://doi.org/10.1111/1755-0998.13011.
[93] A. Glassing, S.E. Dowd, S. Galandiuk, B. Davis, R.J. Chiodini, Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples, Gut Pathog. 8 (2016) 1–12, https://doi.org/10.1186/s13099-016-0103-7.
[94] N.M. Davis, D.M. Proctor, S.P. Holmes, D.A. Relman, B.J. Callahan, Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data, Microbiome 6 (2018) 226, https://doi.org/10.1186/s40168-018-0605-2.
[95] S.P. Walker, M. Barrett, G. Hogan, Y. Flores Bueso, M.J. Claesson, M. Tangney, Non-specific amplification of human DNA is a major challenge for 16S rRNA gene sequence analysis, Sci. Rep. 10 (2020) 16356, https://doi.org/10.1038/s41598-020-73403-7.
[96] Jumpstart Consortium Human Microbiome Project Data Generation Working Group, 16S 454 Sequencing Protocol HMP Consortium, 2010.
[97] P. Ferretti, S. Farina, M. Cristofolini, G. Girolomoni, A. Tett, N. Segata, Experimental metagenomics and ribosomal profiling of the human skin microbiome, Exp. Dermatol. 26 (2017) 211–219, https://doi.org/10.1111/exd.13210.
[98] J.G. Caporaso, C.L. Lauber, W.A. Walters, D. Berg-Lyons, J. Huntley, N. Fierer, S. M. Owens, J. Betley, L. Fraser, M. Bauer, Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms, ISME J. 6 (2012) 1621–1624, https://doi.org/10.1038/ismej.2012.8.
[99] J. Tremblay, K. Singh, A. Fern, E. Kirton, S. He, T. Woyke, J. Lee, F. Chen, J. Dangl, S. Tringe, Primer and platform effects on 16S rRNA tag sequencing, Front. Microbiol. 6 (2015) 771, https://doi.org/10.3389/fmicb.2015.00771.
[100] J. Tackmann, N. Arora, T.S.B. Schmidt, J.F.M. Rodrigues, C. von Mering, Ecologically informed microbial biomarkers and accurate classification of mixed and unmixed samples in an extensive cross-study of human body sites, Microbiome 6 (2018) 192, https://doi.org/10.1186/s40168-018-0565-6.
[101] A.E. Woerner, N.M.M. Novroski, F.R. Wendt, A. Ambers, R. Wiley, S.E. Schmedes, B. Budowle, Forensic human identification with targeted microbiome markers using nearest neighbor classification, Forensic Sci. Int. Genet. 38 (2019) 130–139, https://doi.org/10.1016/j.fsigen.2018.10.003.
[102] R. D'Amore, U.Z. Ijaz, M. Schirmer, J.G. Kenny, R. Gregory, A.C. Darby, M. Shakya, M. Podar, C. Quince, N. Hall, A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling, BMC Genom. 17 (2016) 55, https://doi.org/10.1186/s12864-015-2194-9.
[103] J.J. Kozich, S.L. Westcott, N.T. Baxter, S.K. Highlander, P.D. Schloss, Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform, Appl. Environ. Microbiol. 79 (2013) 5112–5120, https://doi.org/10.1128/AEM.01043-13.
[104] P.D. Schloss, M.L. Jenior, C.C. Koumpouras, S.L. Westcott, S.K. Highlander, Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system, PeerJ 4 (2016), e1869, https://doi.org/10.7717/peerj.1869.
[105] K.E. Ashelford, N.A. Chuzhanova, J.C. Fry, A.J. Jones, A.J. Weightman, At least 1 in 20 16S rRNA sequence records currently held in public repositories is estimated to contain substantial anomalies, Appl. Environ. Microbiol. 71 (2005) 7724–7736, https://doi.org/10.1128/AEM.71.12.7724-7736.2005.
[106] J.J. Werner, O. Koren, P. Hugenholtz, T.Z. DeSantis, W.A. Walters, J.G. Caporaso, L.T. Angenent, R. Knight, R.E. Ley, Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys, ISME J. 6 (2012) 94–103, https://doi.org/10.1038/ismej.2011.82.
[107] P.D. Schloss, The effects of alignment quality, distance calculation method, sequence filtering, and region on the analysis of 16S rRNA gene-based studies, PLoS Comput. Biol. 6 (2010), e1000844, https://doi.org/10.1371/journal.pcbi.1000844.

[108] I.F. Escapa, T. Chen, Y. Huang, P. Gajare, F.E. Dewhirst, K.P. Lemon, New insights into human nostril microbiome from the expanded Human Oral Microbiome Database (eHOMD): a resource for the microbiome of the human aerodigestive tract, mSystems 3 (2018) e00187–18, https://doi.org/10.1128/mSystems.00187-18.

[109] J. Ritari, J. Salojärvi, L. Lahti, W.M. de Vos, Improved taxonomic assignment of human intestinal 16S rRNA sequences by a dedicated reference database, BMC Genom. 16 (2015) 1–10, https://doi.org/10.1186/s12864-015-2265-y.

[110] R. Marcelino, V. Holmes, E.C. Sorrell TC, The use of taxon-specific reference databases compromises metagenomic classification, BMC Genom. 21 (2020) 184, https://doi.org/10.1186/s12864-020-6592-2.

[111] C. Loeffler, A. Karlsberg, L.S. Martin, E. Eskin, D. Koslicki, S. Mangul, Improving the usability and comprehensiveness of microbial databases, BMC Biol. 18 (2020), 37, https://doi.org/10.1186/s12915-020-0756-z.

[112] F.P. Breitwieser, J. Lu, S.L. Salzberg, A review of methods and databases for metagenomic classification and assembly, Brief. Bioinform. 20 (2019) 1125–1136, https://doi.org/10.1093/bib/bbx120.

[113] G. Cammarota, G. Ianiro, A. Ahern, C. Carbone, A. Temko, M.J. Claesson, A. Gasbarrini, G. Tortora, Gut microbiome, big data and machine learning to promote precision medicine for cancer, Nat. Rev. Gastroenterol. Hepatol. 17 (2020) 635–648, https://doi.org/10.1038/s41575-020-0327-3.

[114] P.M. Schneider, B. Prainsack, M. Kayser, The use of forensic DNA phenotyping in predicting appearance and biogeographic ancestry, Dtsch. Arztebl. Int. 116 (2019) 873–880, https://doi.org/10.3238/arztebl.2019.0873.

[115] B.D. Topçuoğlu, N.A. Lesniak, M.T. Ruffin, J. Wiens, P.D. Schloss, A framework for effective application of machine learning to microbiome-based classification problems, mBio 11 (2020) e00434–20, https://doi.org/10.1128/mBio.00434-20.

[116] D. McDonald, E. Hyde, J.W. Debelius, J.T. Morton, A. Gonzalez, G. Ackermann, A. A. Aksenov, B. Behsaz, C. Brennan, Y. Chen, L. DeRight Goldasich, P. C. Dorrestein, R.R. Dunn, A.K. Fahimipour, J. Gaffney, J.A. Gilbert, G. Gogul, J. L. Green, P. Hugenholtz, G. Humphrey, C. Huttenhower, M.A. Jackson, S. Janssen, D.V. Jeste, L. Jiang, S.T. Kelley, D. Knights, T. Kosciolek, J. Ladau, J. Leach, C. Marotz, D. Meleshko, A.V. Melnik, J.L. Metcalf, H. Mohimani, E. Montassier, J. Navas-Molina, T.T. Nguyen, S. Peddada, P. Pevzner, K.S. Pollard, G. Rahnavard, A. Robbins-Pianka, N. Sangwan, J. Shorenstein, L. Smarr, S. J. Song, T. Spector, A.D. Swafford, V.G. Thackray, L.R. Thompson, A. Tripathi, Y. Vázquez-Baeza, A. Vrbanac, P. Wischmeyer, E. Wolfe, Q. Zhu, American Gut Consortium, R. Knight, American gut: an open platform for citizen science microbiome research, mSystems 3 (2018) e00031–18, https://doi.org/10.1128/mSystems.00031-18.

[117] L. Proctor, J. LoTempio, A. Marquitz, P. Daschner, D. Xi, R. Flores, L. Brown, R. Ranallo, P. Maruvada, K. Regan, R. Dwayne Lunsford, M. Reddy, L. Caler, Team NIHHMPA, A review of 10 years of human microbiome research activities at the US National Institutes of Health, Fiscal Years 2007–2016, Microbiome 7 (2019) 31, https://doi.org/10.1186/s40168-019-0620-y.

[118] R. Zhang, A.R. Walker, S. Datta, Unraveling city-specific signature and identifying sample origin locations for the data from CAMDA MetaSUB challenge, Biol. Direct 16 (2021) 1, https://doi.org/10.1186/s13062-020-00284-1.

[119] R. Brewster, F.B. Tamburini, E. Asiimwe, O. Oduaran, S. Hazelhurst, A.S. Bhatt, Surveying gut microbiome research in Africans: toward improved diversity and representation, Trends Microbiol. 27 (2019) 824–835, https://doi.org/10.1016/j.tim.2019.05.006.

[120] E.R. Davenport, D.A. Cusanovich, K. Michelini, L.B. Barreiro, C. Ober, Y. Gilad, Genome-wide association studies of the human gut microbiota, PLoS One 10 (2015), e0140301, https://doi.org/10.1371/journal.pone.0140301.

[121] J. Si, S. Lee, J.M. Park, J. Sung, G. Ko, Genetic associations and shared environmental effects on the skin microbiome of Korean twins, BMC Genom. 16 (2015) 992, https://doi.org/10.1186/s12864-015-2131-y.

[122] A.W. Brooks, S. Priya, R. Blekhman, S.R. Bordenstein, Gut microbiota diversity across ethnicities in the United States, PLoS Biol. 16 (2018), e2006842, https://doi.org/10.1371/journal.pbio.2006842.

[123] V.K. Gupta, S. Paul, C. Dutta, Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity, Front. Microbiol. 8 (2017) 1162, https://doi.org/10.3389/fmicb.2017.01162.

[124] A.M. Levin, A.R. Sitarik, S.L. Havstad, K.E. Fujimura, G. Wegienka, A.E. Cassidy-Bushrow, H. Kim, E.M. Zoratti, N.W. Lukacs, H.A. Boushey, Joint effects of pregnancy, sociocultural, and environmental factors on early life gut microbiome structure and diversity, Sci. Rep. 6 (2016) 1–16, https://doi.org/10.1038/srep31775.

[125] A. Lokmer, S. Aflalo, N. Amougou, S. Lafosse, A. Froment, F.E. Tabe, M. Poyet, M. Groussin, R. Said-Mohamed, L. Ségurel, Response of the human gut and saliva microbiome to urbanization in Cameroon, Sci. Rep. 10 (2020) 1–15, https://doi.org/10.1038/s41598-020-59849-9.

[126] J. Bai, Y. Hu, D.W. Bruner, Composition of gut microbiota and its association with body mass index and lifestyle factors in a cohort of 7–18 years old children from the American Gut Project, Pedia Obes. 14 (2019), e12480, https://doi.org/10.1111/ijpo.12480.

[127] P.A. Dimitriu, B. Iker, K. Malik, H. Leung, W.W. Mohn, G.G. Hillebrand, New insights into the intrinsic and extrinsic factors that shape the human skin microbiome, mBio 10 (2019) e00839–19, https://doi.org/10.1128/mBio.00839-19.

[128] P. Vangay, J. Burgin, A. Johnston, K.L. Beck, D.C. Berrios, K. Blumberg, S. Canon, P. Chain, J.-M. Chandonia, D. Christianson, S.V. Costes, J. Damerow, W. D. Duncan, J.P. Dundore-Arias, K. Fagnan, J.M. Galazka, S.M. Gibbons, D. Hays, J. Hervey, B. Hu, B.L. Hurwitz, P. Jaiswal, M.P. Joachimiak, L. Kinkel, J. Ladau, S. L. Martin, L.A. McCue, K. Miller, N. Mouncey, C. Mungall, E. Pafilis, T.B.K. Reddy, L. Richardson, S. Roux, L.M. Schriml, J.P. Shaffer, J.C. Sundaramurthi, L. R. Thompson, R.E. Timme, J. Zheng, E.M. Wood-Charlson, E.A. Eloe-Fadrosh, Microbiome metadata standards: report of the national microbiome data collaborative's workshop and follow-on activities, mSystems 6 (2021) e01194–20, https://doi.org/10.1128/mSystems.01194-20.

[129] E. Zaura, V.Y. Pappalardo, M.J. Buijs, C.M.C. Volgenant, B.W. Brandt, Optimizing the quality of clinical studies on oral microbiome: a practical guide for planning, performing, and reporting, Periodontology 2000 85 (2021) 210–236, https://doi.org/10.1111/prd.12359.

[130] N. Kachroo, D. Lange, K.L. Penniston, J. Stern, G. Tasian, P. Bajic, A.J. Wolfe, M. Suryavanshi, A. Ticinesi, T. Meschi, M. Monga, A.W. Miller, Standardization of microbiome studies for urolithiasis: an international consensus agreement, Nat. Rev. Urol. 18 (2021) 303–311, https://doi.org/10.1038/s41585-021-00450-8.

[131] A. Shafquat, R. Joice, S.L. Simmons, C. Huttenhower, Functional and phylogenetic assembly of microbial communities in the human microbiome, Trends Microbiol. 22 (2014) 261–266, https://doi.org/10.1016/j.tim.2014.01.011.

[132] S. Bashiardes, G. Zilberman-Schapira, E. Elinav, Use of metatranscriptomics in microbiome research, Bioinform. Biol. Insights 10 (2016) 19–25, https://doi.org/10.4137/BBI.S34610.

[133] M. Kleiner, Metaproteomics: much more than measuring gene expression in microbial communities, mSystems 4 (2019) e00115–e00119, https://doi.org/10.1128/mSystems.00115-19.

[134] S. Abubucker, N. Segata, J. Goll, A.M. Schubert, J. Izard, B.L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput. Biol. 8 (2012), e1002358, https://doi.org/10.1371/journal.pcbi.1002358.

[135] Laboratory LBN, National Microbiome Data Collaborative, 2021. ⟨https://microbiomedata.org/fair/⟩, (Accessed 11 June 2021).

[136] R. Sinha, G. Abu-Ali, E. Vogtmann, A.A. Fodor, B. Ren, A. Amir, E. Schwager, J. Crabtree, S. Ma, C.C. Abnet, R. Knight, O. White, C. Huttenhower, The Microbiome Quality Control Project C, Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium, Nat. Biotechnol. 35 (2017) 1077–1086, https://doi.org/10.1038/nbt.3981.

[137] P.I. Costea, G. Zeller, S. Sunagawa, E. Pelletier, A. Alberti, F. Levenez, M. Tramontano, M. Driessen, R. Hercog, F.-E. Jung, Towards standards for human fecal sample processing in metagenomic studies, Nat. Biotechnol. 35 (2017) 1069–1076, https://doi.org/10.1038/nbt.3960.

[138] Action ECiSaT, ML4Microbiome, 2019. ⟨https://www.ml4microbiome.eu/⟩, (Accessed 11 June 2021).