

Usefulness of Microbiome for Forensic Geolocation

Subjects: **Biochemistry & Molecular Biology**

Contributor: Christian Haarkötter

Forensic microbiomics is a promising tool for crime investigation. Geolocation connects an individual to a certain place or location by microbiota.

forensic microbiology

forensic science

geolocation

microbiome

microbiota

1. Introduction

The microbiome is not a novel concept, given that the term was developed during the late 1980s by Whipps et al. ^[1] to refer to a group of microorganisms living in a defined area. The common factor uniting the various fungi and bacteria in a particular location is the location itself. Today, the term has evolved into two different concepts: 'microbiota' is a term used for a group of microorganisms or viruses that are centred and interact in a certain area, and 'microbiome' is a term for the genomic study of a community of microorganisms ^[2]. However, the microbiome has been defined as a certain microbial community that lives in a defined area with certain physical and chemical properties (it includes the microorganisms and their environment), and the microbiota includes an assembly of microorganisms that belong to different kingdoms, including their microbial structures, metabolic reagents or products and mobile or relic DNA/RNA elements. Thus, the original definition as stated by Whipps appears to be the most accurate ^[3].

Various methodologies and strategies have been developed to describe and classify microorganisms. Prior to 1960, the methodology was mostly based on morphology, metabolic requirements or pathogenicity. In 1960, a numerical taxonomy was introduced into bacterial systematics with the mol% guanine–cytosine content of DNA as a quantitative measurement. Therefore, no more than 2–3% of the variation in guanine–cytosine content was expected in the same species of microorganisms. Chemotaxonomy, the description of new species based on the study of the composition of cell walls or bacterial cytochromes, became common from 1960 to 1980; however, it was supplanted by the arrival of 16S ribosome DNA or rDNA (see **Figure 1**) gene sequencing during the mid-1990s. This approach implied that strains with less than 98.7% sequence similarity were a new species. Given that 16S rDNA is easily isolated, ubiquitous and constrained (constraints are mechanisms that limit or restrict adaptative evolution), it is commonly studied; that is the reason why it is the most common approach in literature ^[4]. Most recently, the introduction of high-throughput technologies, commonly known as next generation sequencing (NGS), allowed whole genome sequencing, in which new species are defined by the comparison between two chromosomes ^[5].

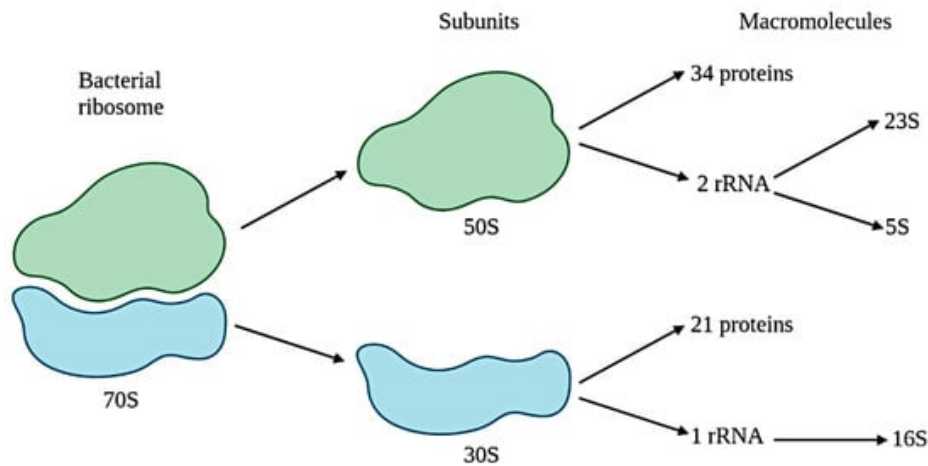


Figure 1. Bacterial ribosome and 16S rRNA.

Currently, there is no official or recognised system for the classification of bacteria; however, the most commonly used system is the polyphasic approach, which includes phenotypic, chemotaxonomic, genotypic and phylogenetic data [6]. Microbiologists use Linnaeus's binomial naming system to designate microorganisms, with *Proteobacteria* divided into seven orders: *Chromatiales*, *Thiotrichales*, *Legionellales*, *Pseudomonadales*, *Vibrionales*, *Enterobacteriales* and *Pasteurellales*. Each order includes several genera, and each genus a variety of species; for example, the family *Enterobacteriaceae* (from the order *Enterobacteriales*) includes the genera *Enterobacter*, *Escherichia*, *Klebsiella*, *Proteus*, *Salmonella*, *Serratia*, *Shigella* and *Yersinia* [7].

Given the importance of the microbiota, the Human Genome Project led to the Human Microbiome Project, whose main objectives are creating a draft database of the human-associated microbiome by 16S rRNA sequencing, studying individuals who represent specific clusters and analysing global human microbiome diversity [8]. As an example of the large variety of microbiota that live in the human body, the most common genera in stomach, small and large intestine, oral cavity, male and female urogenital system and skin are shown in **Figure 2**.

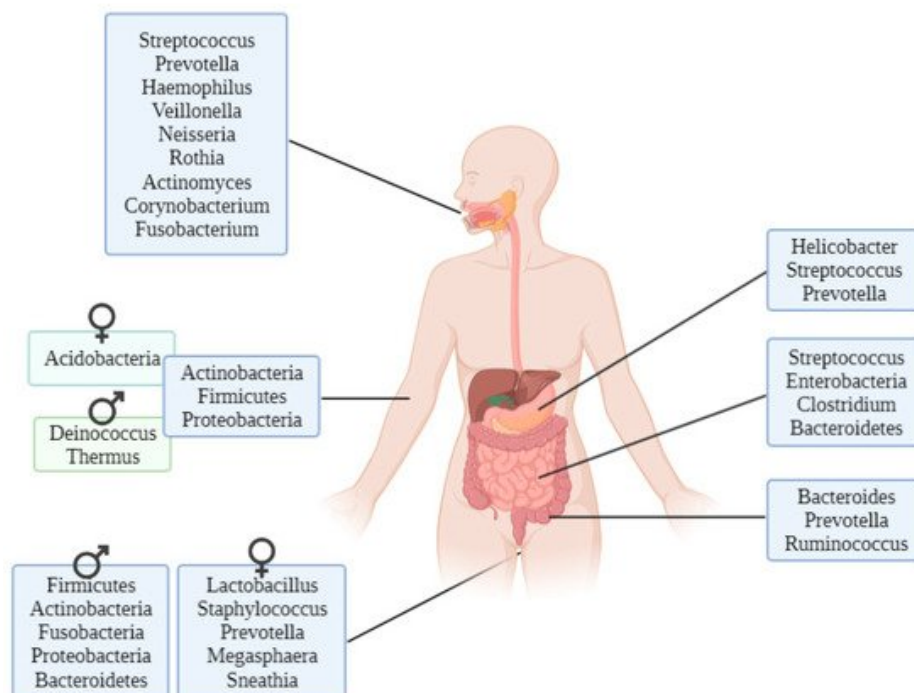


Figure 2. Most common bacteria genera in the different parts of the human body [9].

Forensic microbiology is a fairly new field in the forensic sciences, and it has been developing since the terrorist attacks in the United States in 2001 due to the fear of a possible biological attack. Forensic microbiologists were concerned with developing tools to identify bioweapons and those who use them [10]. Since then, and thanks to major developments in sequencing technologies, its applications are growing rapidly [11]. There are currently three main areas of interest in forensic science [12]:

Identification. The microbiome has the potential to identify an individual in the population based on their characteristic microbes. It appears to be possible to identify the items a person has touched, and therefore to define biogeographical patterns in the items.

Post-Mortem. Interval Estimation. Research shows that there are distinctive microorganisms that can be sequenced at various time points and body locations during decomposition.

Geolocation. Microbiota differ in composition across geographical locations due to climate, rainfall, altitude, soil and energy sources in the environment; thus, the knowledge of specific bacteria composing a certain area would could link a person or item to a certain place.

2. Forensic Microbiome as a Tool for Geolocation

“Every contact leaves a trace” is probably the most important axiom in forensic science, given that it was first established by Locard during the early 20th century. This statement has been applied by forensic scientists since then in all forensic fields, and it can also be applied to microbiome studies [13]. If a certain place contains a characteristic microbiota that is different from other locations, we can analyse a person’s microbiome and possibly establish where they have been, which is precisely the main principle of microbiome geolocation.

Several studies have been performed to characterise the urban and transit microbiome, demonstrating that certain areas of a city contain unique microbiome profiles [14]. Along these lines, the Earth Microbiome Project, EMP (<http://www.earthmicrobiome.org>) must be mentioned. It was created in 2010 to sample the whole planet’s microbial communities with the aim of understanding the biogeographic variations and principles that govern microbial communities by using standardised protocols and environmental descriptors in an open science model [15]. The various samples and their connection by similarity (containing similar types of microbial communities) are shown in **Figure 3**.

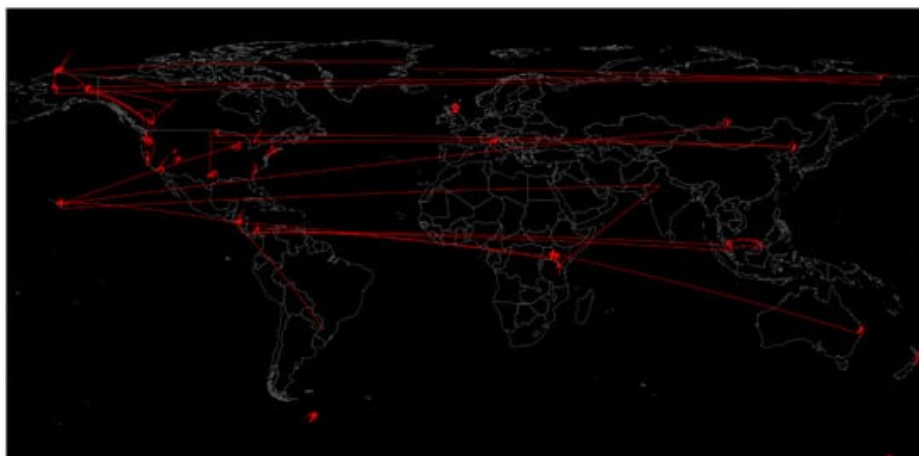


Figure 3. Soil microbiome samples collected by the Earth Microbiome Project by similarity [15].

2.1. Soil and Surface Microbiome

The literature has demonstrated that a whole city's microbiome can be analysed by swab sampling of subway stations, public parks and waterways. Certain species have been found to be linked to certain areas of the city, with a degree of fluctuation observed in some genera during the day. However, an important issue was also discovered: many samples did not match any known organism [16], which calls attention to the importance of projects such as the Earth Microbiome Project. A combined effort to study the urban metagenome can be found in the Metagenomics and Metadesign of the Subways and Urban Biomes, (MetaSUB) International Consortium, which was created with the aim of helping with city planning, public health and architectural design matters [17]. Moreover, in 60 cities across a three-year longitudinal study, it was established that there is geographic variation among microbial communities in type and density [18]; thus, it is possible to create a map of the various microbiota that can be found in specific cities. Interestingly, it has been observed that a relationship can be established between a geographic metagenome and organisms' diversity, acting as a type of 'molecular echo' [19]. This molecular echo could be useful information for future correlations between the microbiome and forensic entomology. Recent advances in city microbiome studies suggest that certain species are especially useful for geolocation, given that some of them are invariably present in every studied city, thus, some genera was particular to each location [20].

2.2. In Vivo Microbiome

There are genera of microorganisms that allow researchers to assess a person's geographical origin. For example, *Helicobacter pylori* extracted from gastric mucosa has been used to determine the geographical origin of unidentified Asian cadavers, resulting in three different clusters: East Asian, Western and Southeast Asian [21]. Furthermore, studies focusing on the relationship between microbiota and diseases such as obesity have found differences between Colombians, Americans, Europeans, Japanese and South Koreans and their relative disposition to increased body mass index [22]. These differences have also been found in studies conducted to evaluate the relationship between the microbiome and infectious diseases such as *Plasmodium falciparum* infection, finding again geographical differences among people in their stool microbiota [23]. Other studies performed with human hair microbiota have found differences between samples from California and Maryland, and interestingly, scalp hair resulted in better prediction of geolocation than pubic hair [24].

Firmicutes and *Bacteroidetes* appear to have a certain pattern depending on the latitude. In a study conducted with healthy individuals' gut microbiota, it was found that the *Firmicutes* and *Bacteroidetes* proportion differs with latitude: the proportion of *Firmicutes* is much higher in the Northern Hemisphere than in the Southern Hemisphere [25]. The explanation of the differences in microbiota remains unclear, although there are three proposed models: host genes, the environment itself or host plasticity.

2.3. Machine Learning and Geolocation

Machine learning automates computers to make predictions based on data. Machine learning has been used in biomedical research, cancer diagnosis and with the human microbiome to predict categorical or numerical values by classification and regression, respectively [26]. The program itself learns from each classification it makes, so the next classification contemplates the previous ones. There are numerous machine learning techniques available for the classification of the human microbiota [27], and random forest is one of them. It is the most commonly used technique in microbiome forensics [28]. A random forest algorithm is a combination of tree predictors (a tree is a type of flux diagram in which every internal node is an attribute, the branch is a decision rule and every leaf a result). Each tree has the same distribution, and its values depend on a random vector sampled independently [29]. Roughly, a random forest works as follows (see **Figure 4**): a data set is introduced into the algorithm, which generates the statistically best decision trees for the given variables, and the algorithm is trained so it can learn from its successes and mistakes (as any other machine learning based algorithm). Then, a problem sample is given so the algorithm makes decisions with the various trees generated, ultimately giving a category result (for example, a country) based on a majority vote of the tree results.

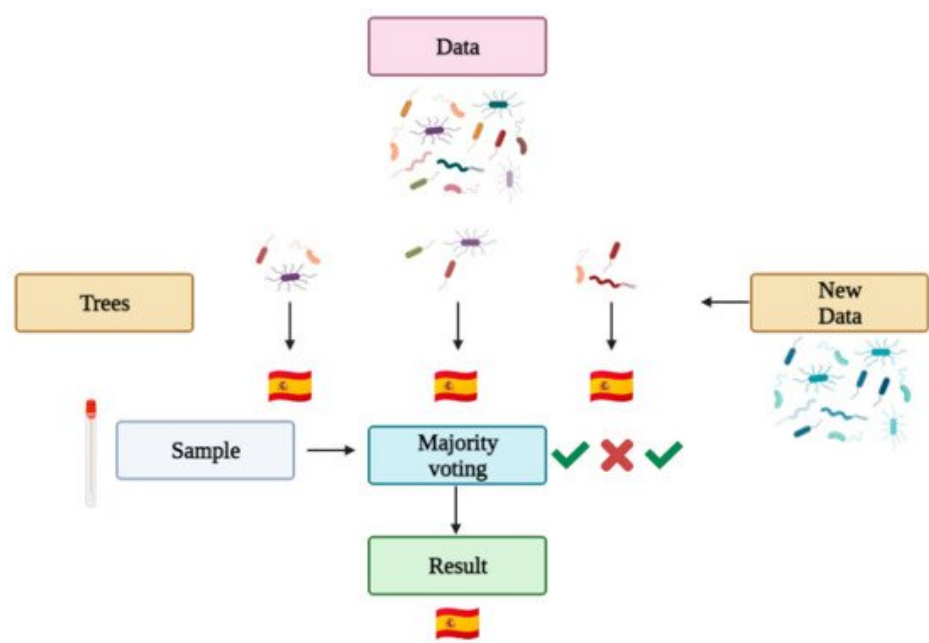


Figure 4. Random forest prediction in microbiome Forensics.

Although random forest is an accurate and unbiased predictor that needs no rigid statistical assumptions of the target variable, it has some disadvantages: greater computational intensity with the increase in calibration data, high sensitivity of predictions to the quality of the input data and variations in obtained model interpretation [30]. Several algorithms have been developed to make machine learning more accessible to forensic scientists. An example applied to microbiome geolocation is DeepSpace, which is based on deep neural network classifiers (algorithms of machine learning that assimilate data representation when they recognise, for example, a human face in a pixel image), which could correctly classify dust from different countries with a 90% accuracy just by using fungi data [31].

2.4. Protocols

Several protocols for sampling, DNA extraction and amplification are available; however, given that swabs are a reliable technique and the DNA extraction methodology is crucial, a reduction in host DNA is recommended [32].

2.4.1. Sampling

The Earth Microbiome Project has designed a protocol for collaborators who want to contribute samples. The protocol depends on the specific sample type, and is summarised in Table 1.

Table 1. Sampling protocol from Earth Microbiome Project [33].

Sampling	Samples Should Be Collected Fresh and Then Frozen without Using Any Buffer or Solution.	
	Soil	Swabs
Procedure	Split fresh sample into 2 mL tubes (10) with, at least, 200 mg biomass and store at -80 or -20 °C.	Take 10 replicate swabs with no buffers or solutions and store in -80 or -20 °C
Shipping	Samples should be shipped with dry ice in an extruded polystyrene foam container or similar.	

2.4.2. DNA Extraction

For good-quality environmental samples there are several commercial platforms for microbiome studies, all of which are magnetic beads based: KingFisher Flex Purification System (ThermoFisher Scientific, Waltham, MA, USA); epMotion 5075

TMX (Eppendorf, Hamburg, Germany); and Tecan Freedom EVO Nucleic Acid Purification (Tecan, Morrisville, NC, USA) [34]. The platforms have been tested with a variety of samples, including faeces, oral, skin, soil and water. The various commercial DNA extraction kits available are shown in **Table 2**. A special strategy has been developed for low-template microbiome samples, for as few as 50-500 cells, called KatharoSeq. It is based on Mo Bio PowerSoil and the QIAGEN Ultra Clean kit [35]. Other kits not designed for the microbiome have been validated for forensic microbiome workflows [36]; however, they present the challenge of not eliminating the non-bacterial DNA present in samples.

Table 2. Commercial kits for microbiome DNA extraction.

Commercial Kit	Principle	Format	Time	Automation
MagMAX Microbiome Ultra Nucleic Acid Isolation Kit (ThermoFisher Scientific) [37]	Magnetic beads	100 reactions	~60 min	KingFisher™ Duo Prime, Flex and Presto
Invitrogen PureLink Microbiome DNA Purification Kit (ThermoFisher Scientific) [38]	Spin column	100 reactions	120 min	-
QIAamp DNA Microbiome Kit (QIAGEN) [39]	Silica columns	50 reactions	~180 min	-
MO BIO's PowerMag® Soil DNA Isolation Kit (QIAGEN) [40]	Magnetic beads	4 × 96 or 32 × 12	60–120 min	epMotion®

2.4.3. Sequencing

References

The 16S amplification protocol was designed for prokaryotes, bacteria and archaea, given that it is an excellent phylogenetic marker, and it provides insight into both communities and individual microbial taxa. The protocol's ability to relate trends of species to hosts or environments has been proven. The polymerase chain reaction primers were developed for the V4 region of 16S rRNA [41].

3. Challenges and Limitations

1. Kambours, M.E.; Velegraki, A.; Pafinos, G.P.; Zerva, E. Introduction: The Microbiome as a Concept: Vogue or Necessity? In Microbiomes; Kambours, M.E., Velegraki, A., Eds.; Academic Press: Cambridge, MA, USA, 2020; pp. 1–4. ISBN 978-0-12-816664-2.

2. Berg, G.; Rybakova, D.; Fischer, D.; Cernava, T.; Vergès, M.-C.C.; Charles, T.; Chen, X.; Cocolin, L.; Eversole, K.; Corai, G.H., et al. Microbiome definition re-visited: Old concepts and new challenges. *Microbiome* 2020, 8, 103.

3. Fox, G.E.; Magrum, J.; Balch, W.E.; Wolfe, R.S.; Woese, C.R. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. *Proc Natl Acad Sci USA* 1977, 74, 4537–4541.

4. Jan-da, J.M. Taxonomic Classification of Bacteria. In *Practical Handbook of Microbiology*; Green, L.H., Goldman, E., Eds.; CRC Press: Boca Raton, FL, USA, 2021; pp. 161–167. ISBN 9780367567637.

5. Schleifer, K.H. Classification of Bacteria and Archaea: Past, present and future. *Syst Appl Microbiol* 2009, 32, 533–542.

6. Wiley, J.M.; Sherwood, L.M.; Woolverton, C.J. *Prescott's Principles of Microbiology*; McGraw-Hill: New York, NY, USA, 2009; ISBN 978-0-07-337523-6.

7. Turnbaugh, P.J.; Ley, R.E.; Hamady, M.; Fraser-Liggett, C.M.; Knight, R.; Gordon, J.I. The Human Microbiome Project. *Nature* 2007, 449, 804–810.

8. Marchesi, J.R. *The Human Microbiota and Microbiome*; Advances in Molecular and Cellular Microbiology, CABI: Wallingford, UK, 2014; ISBN 9781780640495.

9. Budowle, B.; Schutzer, S.E.; Einseln, A.; Kelley, L.C.; Walsh, A.C.; Smith, J.A.L.; Marrone, B.L.; Robertson, J.; Campos, J. Building Microbial Forensics as a Response to Bioterrorism. *Science* 2003, 301, 1852–1853.

10. Bur-cham, Z.M.; Jordan, H.R. History, current, and future use of microorganisms as physical evidence. In *Forensic Microbiology*; Carter, D.O., Tomberlin, J.K., Benbow, M.E., Metcalf, J.L., Eds.; Wiley: Chichester, UK, 2017; pp. 25–56. ISBN 9781119062554.
11. Clarke, T.H.; Gomez, A.; Singh, H.; Nelson, K.E.; Brinkac, L.M. Integrating the microbiome as a resource in the forensics toolkit. *Forensic Sci. Int. Genet.* 2017, 30, 141–147.
12. National Institute of Justice. The Forensic Microbiome: The Invisible Traces We Leave Behind. Available online: <https://nij.ojp.gov/topics/articles/forensic-microbiome-invisible-traces-we-leave-behind> (accessed on 25 October 2021).
13. Robinson, J.M.; Pasternak, Z.; Mason, C.E.; Elhaik, E. Forensic Applications of Microbiomics: A Review. *Front. Microbiol.* 2021, 11, 3455.
14. Thompson, L.R.; Sanders, J.G.; McDonald, D.; Amir, A.; Ladau, J.; Locey, K.J.; Prill, R.J.; Tripathi, A.; Gibbons, S.M.; Ackermann, G.; et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* 2017, 551, 457–463.
15. Afshinnkoo, E.; Meydan, C.; Chowdhury, S.; Jaroudi, D.; Boyer, C.; Bernstein, N.; Maritz, J.M.; Reeves, D.; Gandara, J.; Chhangawala, S.; et al. Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics. *Cell Syst.* 2015, 1, 72–87.
16. The MetaSUB International Consortium; Mason, C. The Metagenomics and Metadesign of the Subways and Urban Biomes (MetaSUB) International Consortium inaugural meeting report. *Microbiome* 2016, 4, 24.
17. Danko, D.; Bezdan, D.; Afshinnkoo, E.; Ahsanuddin, S.; Bhattacharya, C.; Butler, D.J.; Chng, K.R.; DeFilippis, F.; Hecht, J.; Kahles, A.; et al. Global Genetic Cartography of Urban Meta genomes and Anti-Microbial Resistance. *bioRxiv* 2019, 724526.
18. Rosenfeld, J.A.; Reeves, D.; Brugler, M.R.; Narechania, A.; Simon, S.; Durrett, R.; Foox, J.; Shianna, K.; Schatz, M.; Gandara, J.; et al. Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nat. Commun.* 2016, 7, 10164.
19. Walker, A.R.; Datta, S. Identification of city specific important bacterial signature for the MetaSUB CAMDA challenge microbiome data. *Biol. Direct* 2019, 14, 11.
20. Habtom, H.; Pasternak, Z.; Matan, O.; Azulay, C.; Gafny, R.; Jurkevitch, E. Applying microbial biogeography in soil forensics. *Forensic Sci. Int. Genet.* 2019, 38, 195–203.
21. Escobar, J.S.; Klotz, B.; Valdes, B.E.; Agudelo, G.M. The gut microbiota of Colombians differs from that of Americans, Europeans and Asians. *BMC Microbiol.* 2014, 14, 311.
22. Yooseph, S.; Kirkness, E.F.; Tran, T.M.; Harkins, D.M.; Jones, M.B.; Torralba, M.G.; O'Connell, E.; Nutman, T.B.; Doumbo, S.; Doumbo, O.K.; et al. Stool microbiota composition is associated with the prospective risk of *Plasmodium falciparum* infection. *BMC Genom.* 2015, 16, 631.
23. Brinkac, L.; Clarke, T.H.; Singh, H.; Greco, C.; Gomez, A.; Torralba, M.G.; Frank, B.; Nelson, K.E. Spatial and Environmental Variation of the Human Hair Microbiota. *Sci. Rep.* 2018, 8, 9017.
24. Suzuki, T.A.; Worobey, M. Geographical variation of human gut microbial composition. *Biol. Lett.* 2014, 10, 20131037.
25. Metcalf, J.L.; Xu, Z.Z.; Bouslimani, A.; Dorrestein, P.; Carter, D.O.; Knight, R. Microbiome Tools for Forensic Science. *Trends Biotechnol.* 2017, 35, 814–823.

26. Knights, D.; Costello, E.K.; Knight, R. Supervised classification of human microbiota. *FEMS Microbiol. Rev.* 2011, 35, 343–359.
27. Pavlov, Y.L. *Random Forests*; De Gruyter: Berlin, Germany, 2019; Volume 45, ISBN 9783110941975.
28. Young, J.M.; Weyrich, L.S.; Breen, J.; Macdonald, L.; Cooper, A. Predicting the origin of soil evidence: High throughput eukaryote sequencing and MIR spectroscopy applied to a crime scene scenario. *Forensic Sci. Int.* 2015, 251, 22–31.
29. Hengl, T.; Nussbaum, M.; Wright, M.N.; Heuvelink, G.B.M.; Gräler, B. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. *PeerJ* 2018, 6, e5518.
30. Grantham, N.S.; Reich, B.J.; Laber, E.B.; Pacifici, K.; Dunn, R.R.; Fierer, N.; Gebert, M.; Allwood, J.S.; Faith, S.A. Global forensic geolocation with deep neural networks. *J. R. Stat. Soc. Ser. C Appl. Stat.* 2020, 69, 909–929.
31. Bjerre, R.D.; Hugerth, L.W.; Boulund, F.; Seifert, M.; Johansen, J.D.; Engstrand, L. Effects of sampling strategy and DNA extraction on human skin microbiome investigations. *Sci. Rep.* 2019, 9, 17287.
32. Thompson, L.; Ackermann, G.; Humphrey, G.; Gilbert, J.; Jansson, J.; Knight, R. EMP Sample Submission Guide v1. protocols.io 2018. Available online: <https://www.protocols.io/view/emp-sample-submission-guide-pfqdjmw> (accessed on 25 October 2021).
33. Marotz, C.; Amir, A.; Humphrey, G.; Gaffney, J.; Gogul, G.; Knight, R. DNA extraction for streamlined metagenomics of diverse environmental samples. *Biotechniques* 2017, 62, 290–293.
34. Minich, J.J.; Zhu, Q.; Janssen, S.; Hendrickson, R.; Amir, A.; Vetter, R.; Hyde, J.; Doty, M.M.; Stillwell, K.; Benardini, J.; et al. KatharoSeq Enables High-Throughput Microbiome Analysis from Low-Biomass Samples. *mSystems* 2018, 3, e00218-17.
35. Alessandrini, F.; Brenciani, A.; Fioriti, S.; Melchionda, F.; Mingoia, M.; Morroni, G.; Tagliabracci, A. Validation of a universal DNA extraction method for human and microbial DNA analysis. *Forensic Sci. Int. Genet. Suppl. Ser.* 2019, 7, 256–258.
36. Applied Biosystems Mag MAX™ Microbiome Ultra Nucleic Acid Isolation Kit. Available online: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FFLSG%2Fmanuals%2FMAN0018070_MagMAXMicrobiomeNuclAcidIsolatKit_SoilSalivaUrine_Automated_U (accessed on 25 October 2021).
37. Invitrogen Pure Link™ Microbiome DNA Purification Kit. Available online: https://www.thermofisher.com/document-connect/document-connect.html?url=https://assets.thermofisher.com/TFS-Assets%2FFLSG%2Fmanuals%2FMAN0014331_PureLinkMicrobiome_Soil_UG.pdf (accessed on 25 October 2021).
38. QIAGEN QIAamp DNA Microbiome Kit. Available online: <https://www.qiagen.com/us/products/discovery-and-translational-research/dna-rna-purification/dna-purification/microbial-dna/qiaamp-dna-microbiome-kit/> (accessed on 25 October 2021).
39. QIAGENMO-BIO's Power Mag Soil DNA Isolation Kit EPH and Book. Available online: <https://www.qiagen.com/us/resources/resourcedetail?id=faaba49d-9d10-48cf-a461-01d5302fea49&lang=en> (accessed on 25 October 2021).

40. Caporaso, J.G.; Lauber, C.L.; Walters, W.A.; Berg-Lyons, D.; Lozupone, C.A.; Turnbaugh, P.J.; Fierer, N.; Knight, R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* 2011, 108 (Suppl. 1), 4516–4522.
41. Amaral-Zettler, L.A.; McCliment, E.A.; Ducklow, H.W.; Huse, S.M. A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes. *PLoS ONE* 2009, 4, e6372.
42. European Network of Forensic Science Institutes ENFSI Guideline for Evaluative Reporting in Forensic Science. Available online: https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (accessed on 25 October 2021).
43. Neckovic, A.; Van Oorschot, R.A.H.; Szkuta, B.; Durdle, A. Challenges in Human Skin Microbial Profiling for Forensic Science: A Review. *Genes* 2020, 11, 1015.

Retrieved from <https://encyclopedia.pub/entry/history/show/40839>