SHORT REPORT   **Open Access**

DNA metabarcoding of forensic mycological samples

Saverio Giampaoli[1], Elisabetta De Vittori[2], Filippo Barni[3], Anna Anselmo[3], Teresa Rinaldi[4], Marina Baldi[5], Kevin Charles Miranda[6], Arnold Liao[6], Daniel Brami[6], Giovanni Vanni Frajese[1*] and Andrea Berti[3]

Abstract

Background: DNA metabarcoding and massive parallel sequencing are valuable molecular tools for the characterization of environmental samples. In forensic sciences, the analysis of the sample's fungal population can be highly informative for the estimation of post-mortem interval, the ascertainment of deposition time, the identification of the cause of death, or the location of buried corpses. Unfortunately, metabarcoding data analysis often requires strong bioinformatic capabilities that are not widely available in forensic laboratories.

Results: The present paper describes the adoption of a user-friendly cloud-based application for the identification of fungi in typical forensic samples. The samples have also been analyzed through the QIIME pipeline, obtaining a relevant data concordance on top genus classification results (88%).

Conclusions: The availability of a user-friendly application that can be run without command line activities will increase the popularity of metabarcoding fungal analysis in forensic samples.

Keywords: NGS, Fungi, Data analysis, Metabarcoding, Forensic science

Background

The analysis of biological components in a specific place can be crucial for the determination of the events that took place at the crime scene. In particular, the identification of non-human biological traces can better clarify transfer events, storage conditions, and the environmental origins of items (Iyengar and Hadi 2014). Environmental samples are characterized by the presence of different biological species that can be identified by DNA sequencing and analysis. The non-human DNA pool present in those samples reflects the biodiversity of a specific milieu. In particular, environmental fungi represent an important group of organisms strictly correlated with decomposition and that require specific living conditions. Specific fungal populations characterize well-defined ecological niches and can be informative about the conservation conditions of items that can be found

at the crime scene (Dash and Das 2018). In addition, forensic analysis of the fungal population can be highly informative for the identification of human activities connected to specific jobs, the estimation of post-mortem interval, the ascertainment of deposition time, the identification of the cause of death, and the location of buried corpses (Hawksworth and Wiltshire 2011; Di Piazza et al. 2018). The identification of mycological profiles has been presented also as a highly informative tool and probative evidence in crime reconstruction with real caseworks (Wiltshire and Hawksworth 2014).

In the last few years, the analysis of biological matrices through metabarcoding and high-throughput sequencing has changed classical approaches to biodiversity estimation. New protocols have been developed in basic science fields, and they are now becoming available to applied biomedical areas like public health and forensic science (Giampaoli et al. 2017; Littlefair and Clare 2016; Young et al. 2017; Bianchi et al. 2018). The efficacy of metabarcoding analysis is strictly correlated with the availability of DNA reference libraries and robust

* Correspondence: giovanni.frajese@uniroma4.it
[1]Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Piazza Lauro De Bosis, 6, 00135 Rome, Italy
Full list of author information is available at the end of the article

bioinformatic tools (Taberlet et al. 2012). For prokaryotic analysis, the sequencing of the 16S ribosomal RNA gene (rRNA) is considered a valuable approach unless the researchers want to study the intraspecific differences. Several 16S rRNA gene databases (like Greengenes, SILVA, RDP) are available and incorporated in software packages for bacteria metabarcoding (DeSantis et al. 2007; Glöckner et al. 2017; Cole et al. 2014).

For eukaryotes, other databases have been suggested, in general trying to focus on the most informative DNA sequence regions for each kingdom. For *Fungi*, a specific database, named UNITE (http://unite.ut.ee), is based on the internal transcribed spacer (ITS) region and is available and ready to be used with several software packages like QIIME or mothur (Kõljalg et al. 2013). Unfortunately, QIIME and mothur are software packages for bioinformatics data processing that require informatic expertise and experience in command line operations (Oulas et al. 2015). This constrain strongly reduces the expansion of fungal metabarcoding in applied medical fields where bioinformatic resources are not easily available.

Recently, certain cloud pipelines have been generated (e.g., PipeCraft) that try to meet higher flexibility and ease-of-use standards (Anslan et al. 2017). Nevertheless, additional simplification of workflows is required for next-generation sequencing data analysis, and user-friendly applications are strongly desirable in order to support standardization and incorporation of fungal metabarcoding into routine protocols.

The present paper describes a new user-friendly application for the analysis of fungal metabarcoding data. This application is the adaptation of a previous cloud-based prokaryotic tool and can represent a useful and rapid instrument for the analysis of fungal biodiversity in forensic samples. The aim of the paper is not to perform a validation of a bioinformatic algorithm/pipeline, but the description of the versatility of existing informatic tools with a specific application with the UNITE database. Hopefully, the informatic steps adopted for the generation of the new application could be implemented with other curated sequence reference DNA datasets, like those for plants, protozoa, or animals.

Methods

Samples and DNA extraction

A total of 16 samples were considered for this study: four food samples with mold, eleven environmental (indoor and outdoor) samples, and one commercially available fungal sample, brewer's yeast, as a positive control (Table 1). Samples 7, 12, 18, and 22 are from a kitchen in an apartment. Samples 11, 34, 35, 53, and 54 are from different rooms in a country house. Sample 29 is from an office building. Sample 51 is from a carpentry

Table 1 A total of 16 samples were selected for the study. The table shows the sample ID and the sample description. The samples have been chosen from indoor and outdoor environments as an example of items contaminated by fungi that can characterize specific milieus. Sample 25 was included as a positive control, due to the fact that brewer's yeast is mainly composed of *Saccharomyces*

Sample ID	Description
7	Swab of spoiled tomato with mold
11	Swab of bath mat with black mold
12	Swab of green tomato with mold
18	Spoiled bread with mold
22	Swab of spoiled pear with mold
23	Swab of the surface of a car washing station
25	Brewer's yeast
26	Wet soil
29	Swab of air conditioning system with dust and mold
30	Soil from horse racetrack
31	Soil from plant pot
34	Swab of a chair seat with mold
35	Swab of a clothes brush with mold
51	Swab of the floor of a woodworking
53	Swab of a laundry basket with mold (internal surface)
54	Swab of a laundry basket with mold (external surface)

building. Samples 23, 26, 30, and 31 are from outdoor environments. All samples were collected using ForensiX swabs (Prionics AG, Zurich, Switzerland) with the exception of spoiled bread, brewer's yeast, and soil specimens that were directly processed starting from tiny amounts (less than 250 mg).

These samples have been chosen mainly as an example of items characterizing specific milieus and containing fungi. Vegetables with mold are largely present in a specific working area (food storages, agriculture, kitchens), while other swabs or items can characterize specific indoor environments of possible crime scenes. They are not strictly correlated to human death, but mainly to environmental characterization.

DNA was purified according to Giampaoli et al. (2012) with minor modifications: in particular, after physical disaggregation of the samples, an incubation (30 min at 30°C) with Lyticase (Sigma-Aldrich, St. Louis, MO, USA) at a final concentration of 2.5 U/μl was performed.

Library preparation and NGS sequencing

NGS samples were prepared according to the "16S Metagenomic Sequencing Library Preparation" guide (Part# 15044223 rev. A; Illumina, San Diego, CA, USA), with slight modifications. In particular, the following primers for fungal ITS region 1 (ITS1) and containing Illumina overhang adapters were used: ITS1-F (Gardes

and Bruns 1993) and ITS2 (Jumpponen and Jones 2009) (expected amplicon size 320 bp) (Table 2). A limited number of samples were amplified also with primers ITS3 and ITS4 (White et al. 1990) (expected amplicon size 520 bp) (Table 2). The annealing condition for the amplicon PCR was identified through gradient PCR on genomic DNA of *Candida albicans* and agarose gel electrophoresis visualization: the annealing at 57°C was selected in order to guarantee an adequate yield while also starting from a very low amount of sample.

Libraries were quantified through Qubit Fluorometric assay (Thermo Fisher Scientific, Waltham, MA, USA) and library size validated through agarose gel electrophoresis. After library normalization and pooling (according to the "16S Metagenomic Sequencing Library Preparation" guide), a 2 × 251 sequencing run was performed on a MiSeq FGx desktop sequencer (Illumina) in research use only (RUO) mode. The resulting sequencing data was demultiplexed using the bcl2fastq software (Illumina) before uploading of fastq data files to Illumina's cloud platform: BaseSpace Sequence Hub (BSSH; basespace.illumina.com).

Bioinformatics

The BSSH application (app) named "16S Metagenomics v1.0.1" performs taxonomic classification of 16s rRNA sequences using a naive Bayesian classifier as described by (Wang et al. 2007; Illumina 2014). Briefly, the naive Bayesian classifier searches different kmer lengths against the corresponding database and applies statistical tests to find significant hits. Here, the application was modified by the substitution of the default 16S rRNA database (Green-Genes) with the eukaryotic ITS database (UNITE) (Kõljalg et al. 2013; Abarenkov et al. 2010). To perform database substitution, custom scripts were used to combine the UNITE reference dynamic files (sh_taxonomy_qiime_ver7_dynamic_31.01.2016.txt and sh_refs_qiime_ver7_dynamic_31.01.2016.fasta; the default threshold value of the dynamic file is 98.5%) into a single tab-delimited file. The new application with the ITS database has been named "ITS Metagenomics v1.0.0." The previously uploaded fastq

files were then processed by the "ITS Metagenomics v1.0.0" app.

The same dataset was also analyzed with QIIME pipeline version 1.9.1 for comparative purposes (Caporaso et al. 2010). The read data in fastq format were subjected to default quality processing and removal of singleton reads. The taxonomic classification of the quality-processed reads was based on the open reference clustering of sequences into operational taxonomical units (OTUs), using the UCLUST method (Edgar 2010). In the open-reference OTU picking process, reads were clustered against a reference sequence collection, and all the reads which did not hit the reference sequence collection were subsequently clustered de novo. The read clusters were BLAST assigned to taxonomies using the UNITE reference database (release 31.01.2016).

Results and discussion

In order to make available a user-friendly application for fungal metabarcoding to the forensic scientific community, the "ITS Metagenomics" tool was added to the bioinformatic apps available on the Illumina cloud platform BSSH. The fastq.gz files generated by the MiSeq FGx desktop sequencer were directly uploaded to BSSH and, after the identification of reads as paired-ends, analyzed by the "ITS Metagenomics" tool. The application generates single-sample reports and an aggregate summary. For each sample, the report file summarizes classification statistics at different taxonomic levels (kingdom, phylum, class, order, family, genus, species) and shows the relative abundance of the classification results through an interactive sunburst chart together with bar and pie charts. The report, with exception of the interactive sunburst chart, is available also in a pdf file. The application generates a csv file containing the classification results for each single sample.

The aggregate summary presents the beta-diversity analysis as pdf or html files, together with the csv files containing the classification results at all taxonomic levels: an interactive principal coordinate analysis chart (PCoA) and a hierarchical clustering dendrogram show the differences in the distribution of taxonomic classifications between samples.

The availability of report files containing beta-diversity analysis is particularly useful in forensic science. One of the main questions for forensic scientists is about the similarity of two or more samples. The PCoA chart visualizes similarities or dissimilarities of data and helps the identification of subsets of specimens, allowing the forensic scientist to speculate if two samples are strictly correlated (e.g., because they have been collected in similar environments) or not. The "ITS Metagenomics" tool can generate a PCoA chart at all taxonomic levels (kingdom, phylum, class, order, family, genus, species).

Table 2 Nucleotide sequence of the primers used for library preparation

Primer	Sequence (from 5′ to 3′)
ITS1-F	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCTTGGTCATTTAG AGGAAGTAA
ITS2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTGCGTTCTTC ATCGATGC
ITS3	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCATCGATGAA GAACGCAGC
ITS4	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCCTCCGCTTAT TGATATGC

The application version described in the present paper was built on the 2016 release of the UNITE database, but future updates will be scheduled. The database can support the analysis of both ITS1 and ITS2 regions. The selection of the sequencing primers for fungal metabarcoding should reflect technical limits and taxonomic needs. According to some papers, the ITS2 region should be less affected by amplification bias, but at the same time, the larger size and the variability in length can represent a limit for several sequencing platforms (Turenne et al. 1999; Op De Beeck 2014). Indeed, shorter fragments are much more suitable for metabarcoding assays. Considering that the aim of the present work was not the identification of the best barcoding sequence in fungi and that a genus-level classification can be acceptable for the definition of a microbiological signature in many environmental samples, the adoption of primers ITS1-F and ITS2 was considered. In addition, few samples were also sequenced with primers ITS3 and ITS4, obtaining comparable outputs (data not shown).
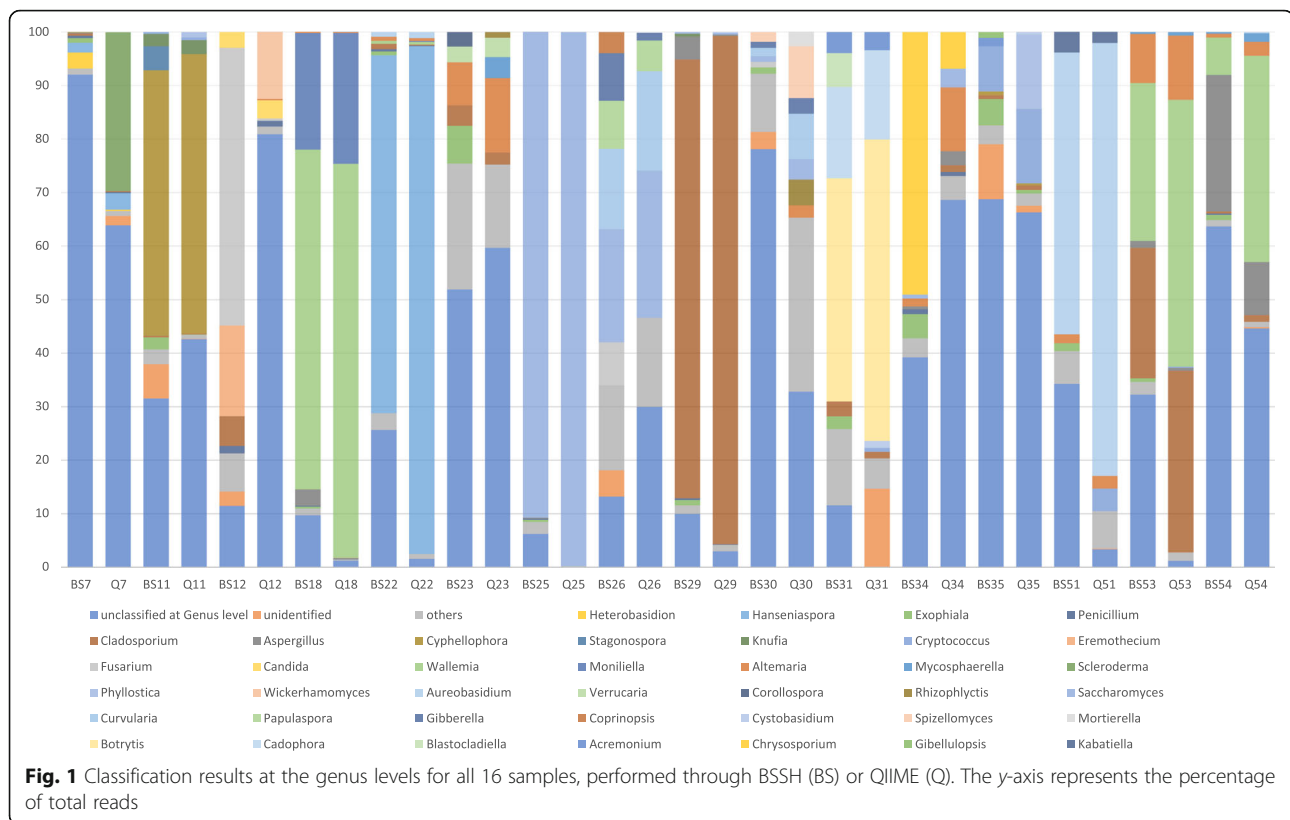
Considering that fungal contamination can be linked to item storage circumstances, microclimatic conditions, and environmental ecosystems, specific samples from typical indoor and outdoor locations were collected for this technical evaluation. In particular, samples from furniture, food, soil, indoor surfaces, and positive control were collected according to usual forensic procedures (Table 1): samples from soil and indoor or outdoor surfaces are usually collected for crime reconstruction, while food traces can be present as contaminants on clothes worn by victims or suspects. For all samples, it has been possible to generate libraries with the selected primers for the ITS1 region, and sequence files were analyzed on BSSH: the minimum number of reads passing quality filtering per sample was approximately 60,000 while the maximum was close to 2,000,000 reads (Table 3). The high variability in the number of reads was probably the consequence of library quantification through gel electrophoresis: nonetheless, the number of reads obtained appears sufficient for metagenomic surveys of the selected samples or for the estimation of predominant taxa (Ni et al. 2013). The number of species identified ranged from 208 to 900, while the number of genera ranges from 127 to 566.

The metabarcoding results showed the presence of fungal genera that typically colonize specific environments or surfaces (Fig. 1). For example, *Wallemia sebi*, largely found on sample 18 (spoiled bread with mold), is a xerophilic fungus commonly found on highly sugared or salted foods, such as jams, bread, or cakes, and for this reason can be considered a possible marker for places where food is stored (Zajc and Gunde-Cimerman 2018). Sample 22 (spoiled pear) was characterized by the predominance of the genus *Hanseniaspora*, an organism associated with fruits and blossoms of different fruit trees (Vadkertiová et al. 2012). Sample 51 (floor of a woodworking) was characterized by the predominance of the genus *Aureobasidium*: this fungus colonizes leaf surfaces of several trees and various wood surfaces that

Table 3 Statistics of taxon identification obtained through BSSH and QIIME (in brackets). The Shannon Species Diversity Index is reported as a measure of species diversity in each sample. The QIIME analysis shows a reduced number of species and genera identified when compared to BSSH

Sample ID	Number of reads PF	% reads classified to genus	Shannon Species Diversity	Number of species identified	Number of genera identified
54	74,821	36.32%	1.071	208 (88)	127 (61)
51	116,158	65.75%	1.464	324 (155)	252 (106)
11	1,275,602	68.47%	1.397	582 (242)	421 (170)
34	226,684	60.75%	0.562	370 (152)	290 (106)
22	71,584	74.34%	1.137	294 (116)	226 (83)
7	393,827	7.97%	0.417	390 (148)	313 (91)
35	1,202,353	31.22%	1.191	596 (240)	448 (166)
53	158,877	67.75%	1.867	439 (192)	293 (125)
18	206,093	90.26%	1.113	216 (119)	172 (83)
29	594,710	90.02%	1.043	524 (246)	380 (168)
23	59,588	48.09%	1.845	511 (272)	339 (178)
31	221,512	88.41%	2.169	611 (262)	429 (168)
30	307,746	21.89%	1.258	548 (194)	382 (138)
26	569,935	86.75%	2.908	900 (394)	566 (256)
12	150,823	88.50%	1.432	208 (66)	184 (49)
25	1,992,943	93.77%	0.644	473 (155)	384 (113)

Fig. 1 Classification results at the genus levels for all 16 samples, performed through BSSH (BS) or QIIME (Q). The y-axis represents the percentage of total reads

can be easily found in a woodworking area (Andrews et al. 2002; van Nieuwenhuijzen et al. 2016).

In order to compare the new tool with other available software tools, the sequence files of each sample were also analyzed with the QIIME pipeline, and the results were compared with those obtained with the BSSH "ITS Metagenomics" app. The comparison of two different bioinformatic software tools can raise concerns about computational irreproducibility and database usage (Di Tommaso et al. 2017). In this situation, the two software adopted the same database release, but differences in the classifier can lead to variations in unclassified reads and in correct assignment at the taxonomic level. While this problem cannot be easily resolved, it is reasonable to expect that two different bioinformatic tools can lead to similar results especially in a sample with a reduced number of unclassified reads.

The QIIME analysis showed a reduced number of species (ranging from 394 to 88) and genera identified (from 256 to 61; see Table 3). Figure 1 represents the classification results at the genus level for the 16 samples of the study, both through QIIME and through BSSH analysis. The genus-level classification is widely adopted for the comparison of environmental matrices; in addition, the analysis at a species level can be inaccurate in a single genetic locus approach (Thomas et al. 2012; Moussa et al. 2017; Rotimi et al. 2018). About the

samples described in the present paper, when not considering the unclassified and unidentified reads, in 11 samples (69%), the most abundant genus identified was the same with both software tools. The concordance percentage was higher (88%) when considering the concordance of the first genus or at least of two genera on the three most abundant. The positive control was correctly analyzed with both bioinformatic tools: *Saccharomyces* was identified as the predominant genus both with BSSH (90% of assigned reads) and QIIME (99%). Only the analysis of sample 12 (green tomato with mold) and sample 35 (clothes brush with mold) showed a reduced concordance. It is important to notice that the QIIME analysis of sample 12 showed more than 80% of unclassified reads at the genus level, while with BSSH, this datum is 11.5%: this output has probably affected the correct investigation of the sample diversity. Similar considerations can be done for sample 35 where approximately 70% of reads were unclassified with both approaches. Nonetheless, a more in-depth examination of the identified genera in sample 35 underlined many concordances between the two analyses, showing the presence of *Cryptococcus*, *Exophiala*, *Cyphellophora*, and *Cladosporium* amongst the six most representative taxa in both situations (Table 4). While the presence of a large quantity of unclassified or unidentified reads does not automatically mean that the analysis is

Table 4 Top genus classification results for sample 35, obtained through BSSH and QIIME. The genera *Cryptococcus*, *Exophiala*, *Cyphellophora*, and *Cladosporium* are present in both the analysis but with differences in the relative abundance

QIIME genus classification	% reads	BaseSpace genus classification	% reads
Unclassified	66.33	Unclassified	68.78
Phyllostica	13.91	Unidentified	10.25
Cryptococcus	13.88	*Cryptococcus*	8.41
Unidentified	1.19	*Exophiala*	4.91
Cladosporium	0.86	*Acremonium*	1.58
Exophiala	0.65	*Gibellulopsis*	1.11
Cystobasidium	0.47	*Cyphellophora*	0.72
Cyphellophora	0.38	*Cladosporium*	0.71
Others	2.33	Others	3.53

compromised, it looks clear that it enhances uncertainty, suggesting additional investigations: the exclusion from the present study of those samples with a high level of unclassified reads leads to an increased concordance between the two approaches. Sample 7 (a swab of a spoiled tomato) presented the largest percentage of unclassified reads at the genus level. The sample was collected from a vegetable largely decomposed. It is possible that the swab was highly contaminated from tomato cells: in this situation, the PCR amplification had reduced stringency due to the large presence of plant DNA: the NGS was able to clean up the datum, working only on the sequences with correspondence in the UNITE database.

Conclusions

At the moment, many different software solutions are available for NGS metabarcoding (or marker gene metagenomics) data analysis. This is a large and rapidly moving field where specific informatic programming capabilities are strongly required. Commonly used tools for metabarcoding data analysis and denoising include QIIME, Mothur, SILVAngs, MEGAN, and AmpliconNoise (Oulas et al. 2015). QIIME was initially implemented for use of 454 pyrosequencing datasets, but it has been then modified to accept the Illumina fastq file format. Unfortunately, this open-source software package is mostly implemented with the PYTHON language, not easily accessible to current forensic scientists. For this reason, user-friendly applications are strongly desirable in order to support metabarcoding studies in forensics. The BSSH application named "ITS Metagenomics" can represent an interesting tool for rapid analysis of fungal biodiversity in several environmental samples, including forensic caseworks. When compared to the QIIME classification, the results look similar: differences can be ascribed to variances in the two informatic pipelines or caused by the large presence of unclassified reads in specific samples (Zajc and Gunde-Cimerman 2018). This latter point could be solved in the future,

through the adoption of new releases of the UNITE reference database. Interestingly, the "ITS Metagenomics" app seems able to identify a larger number of taxa per sample. In our opinion, this first release of the application can represent an important starting point for data analysis in forensics, and informatic steps adopted for the implementation of UNITE database in the app can be suitable for other reference databases. In addition, its presence on the BSSH cloud is fully compatible with the implementation of automatic analysis of data generated from Illumina sequencers like the MiSeq system.

The importance of bioinformatic applications for NGS in forensic science has been recently underlined (Allwood et al. 2020a; Allwood et al. 2020b; Giampaoli et al. 2018). The analysis of biological material, even from trace amounts in dust, has clear advantages for specific forensic applications, in particular, when we need to discern sample origin or geolocation (Allwood et al. 2020a). Recently, several authors focused their studies on fungal succession during mammalian cadaver decomposition, underlining the value of performing mycological analysis with NGS technology in the burial context (Fu et al. 2019; Procopio et al. 2020). Interestingly, metabarcoding applied to diatom DNA can support forensic discrimination of drowning incidents in postmortem examinations (Liu et al. 2020). In addition, NGS metabarcoding applied also to other taxa can be useful in linking specific samples to ecological habitats (Fløjgaard et al. 2019).

As a final consideration, the correct interpretation of metabarcoding data in legal disputations should be standardized by the definition of international guidelines, developed after a large empirical interlaboratory evaluation. This point will be crucial for the diffusion of the metabarcoding NGS analysis in forensic science.

Availability of data and materials
Please contact the corresponding author for data requests.

Declarations

Ethics approval and consent to participate
According to the national laws and international guidelines, the research activities described in the paper do not require an approval from an ethics committee. The Institutional Review Board of the University of Rome "Foro Italico" (CAR-IRB, http://www.uniroma4.it/?q=node/4997) confirmed the exemption from requiring ethics approval.

Consent for publication
This study does not include publishing of personal data.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Movement, Human and Health Sciences, University of Rome "Foro Italico", Piazza Lauro De Bosis, 6, 00135 Rome, Italy. [2]Central Laboratory of National DNA Database, Department of Penitentiary Administration, Ministry of Justice, via del Casale di San Basilio 168, 00156 Rome, Italy. [3]Carabinieri, Reparto Investigazioni Scientifiche di Roma, Viale di Tor di Quinto, 119, 00191 Rome, Italy. [4]Department of Biology and Biotechnology, Sapienza Università di Roma, Rome, Italy. [5]Legal Genetics - Forensic Consulting, Via F. Cesi, 21, 00193 Rome, Italy. [6]Illumina, Inc., 5200 Illumina Way, San Diego, CA 92122, USA.

References

Abarenkov K, Henrik Nilsson R, Larsson KH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, Pennanen T, Sen R, Taylor AF, Tedersoo L, Ursing BM, Vrålstad T, Liimatainen K, Peintner U, Kõljalg U (2010) The UNITE database for molecular identification of fungi--recent updates and future perspectives. New Phytol 186(2):281–285. https://doi.org/10.1111/j.1469-8137.2009.03160.x

Allwood JS, Fierer N, Dunn RR (2020a) The future of environmental DNA in forensic science. Appl Environ Microbiol 86(2):e01504–e01519

Allwood JS, Fierer N, Dunn RR, Breen M, Reich BJ, Laber EB, Clifton J, Grantham NS, Faith SA (2020b) Use of standardized bioinformatics for the analysis of fungal DNA signatures applied to sample provenance. Forensic Sci Int 310: 110250. https://doi.org/10.1016/j.forsciint.2020.110250

Andrews JH, Spear RN, Nordheim EV (2002) Population biology of Aureobasidium pullulans on apple leaf surfaces. Can J Microbiol 48(6):500–513. https://doi.org/10.1139/w02-044

Anslan S, Bahram M, Hiiesalu I, Tedersoo L (2017) PipeCraft: flexible open-source toolkit for bioinformatics analysis of custom high-throughput amplicon sequencing data. Mol Ecol Resour 17(6):e234–e240. https://doi.org/10.1111/1755-0998.12692

Banchi E, Ametrano CG, Stanković D, Verardo P, Moretti O, Gabrielli F, Lazzarin S, Borney MF, Tassan F, Tretiach M, Pallavicini A, Muggia L (2018) DNA metabarcoding uncovers fungal diversity of mixed airborne samples in Italy. PLoS One 13(3):e0194489. https://doi.org/10.1371/journal.pone.0194489

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7(5):335–336. https://doi.org/10.1038/nmeth.f.303

Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, Brown CT, Porras-Alfaro A, Kuske CR, Tiedje JM (2014) Ribosomal database project: data and tools for high throughput rRNA analysis. Nucleic Acids Res 42(D1):D633–D642. https://doi.org/10.1093/nar/gkt1244

Dash HR, Das S (2018) Microbial degradation of forensic samples of biological origin: potential threat to human DNA typing. Mol Biotechnol 60(2):141–153. https://doi.org/10.1007/s12033-017-0052-5

DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL (2007) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol 72(7): 5069–5072

Di Piazza S, Zotti M, Barranco R, Cecchi G, Greco G, Ventura F (2018) Post-mortem fungal colonization pattern during 6 weeks: two case studies. Forensic Sci Int 289:e18–e23. https://doi.org/10.1016/j.forsciint.2018.05.037

Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E, Notredame C (2017) Nextflow enables reproducible computational workflows. Nat Biotechnol 35(4):316–319. https://doi.org/10.1038/nbt.3820

Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Fløjgaard C, Frøslev TG, Brunbjerg AK, Bruun HH, Moeslund J, Hansen AJ, Ejrnæs R (2019) Predicting provenance of forensic soil samples: linking soil to ecological habitats by metabarcoding and supervised classification. PLoS 14(7):e0202844. https://doi.org/10.1371/journal.pone.0202844

Fu X, Guo J, Finkelbergs D, He J, Zha L, Guo Y, Cai J (2019) Fungal succession during mammalian cadaver decomposition and potential forensic implications. Sci Rep 9(1):12907. https://doi.org/10.1038/s41598-019-49361-0

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes--application to the identification of mycorrhizae and rusts. Mol Ecol 2(2):113–118. https://doi.org/10.1111/j.1365-294X.1993.tb00005.x

Giampaoli S, Alessandrini F, Frajese GV, Guglielmi G, Tagliabracci A, Berti A (2018) Environmental microbiology: perspectives for legal and occupational medicine. Leg Med (Tokyo) 35:34–43. https://doi.org/10.1016/j.legalmed.2018.09.014

Giampaoli S, Berti A, Valeriani F, Gianfranceschi G, Piccolella A, Buggiotti L, Rapone C, Valentini A, Ripani L, Romano Spica V (2012) Molecular identification of vaginal fluid by microbial signature. Forensic Sci Int Genet 6(5):559–564. https://doi.org/10.1016/j.fsigen.2012.01.005

Giampaoli S, DeVittori E, Valeriani F, Berti A, Romano Spica V (2017) Informativeness of NGS analysis for vaginal fluid identification. J Forensic Sci 62(1):192–196. https://doi.org/10.1111/1556-4029.13222

Glöckner FO, Yilmaz P, Quast C, Gerken J, Beccati A, Ciuprina A, Bruns G, Yarza P, Peplies J, Westram R, Ludwig W (2017) 25 years of serving the community with ribosomal RNA gene reference databases and tools. J Biotechnol 261: 169–176. https://doi.org/10.1016/j.jbiotec.2017.06.1198

Hawksworth DL, Wiltshire PE (2011) Forensic mycology: the use of fungi in criminal investigations. Forensic Sci Int 206(1-3):1–11. https://doi.org/10.1016/j.forsciint.2010.06.012

Illumina. MiSeq reporter metagenomics workflow reference guide. Part#15042317Rev.D December 2014. https://support.illumina.com/content/dam/illumina-support/documents/documentation/software_documentation/miseqreporter/miseq-reporter-metagenomics-workflow-reference-guide-15042317-d.pdf Accessed 20 March 2020

Iyengar A, Hadi S (2014) Use of non-human DNA analysis in forensic science: a mini review. Med Sci Law 54(1):41–50. https://doi.org/10.1177/0025802413487522

Jumpponen A, Jones KL (2009) Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate Quercus macrocarpa phyllosphere. New Phytol 184(2):438–448. https://doi.org/10.1111/j.1469-8137.2009.02990.x

Kõljalg U, Nilsson RH, Abarenkov K, Tedersoo L, Taylor AF, Bahram M, Bates ST, Bruns TD, Bengtsson-Palme J, Callaghan TM, Douglas B, Drenkhan T, Eberhardt U, Dueñas M, Grebenc T, Griffith GW, Hartmann M, Kirk PM, Kohout P, Larsson E, Lindahl BD, Lücking R, Martín MP, Matheny PB, Nguyen NH, Niskanen T, Oja J, Peay KG, Peintner U, Peterson M, Põldmaa K, Saag L, Saar I, Schüßler A, Scott JA, Senés C, Smith ME, Suija A, Taylor DL, Telleria MT, Weiss M, Larsson KH (2013) Towards a unified paradigm for sequence-based identification of fungi. Mol Ecol 22(21):5271–5277. https://doi.org/10.1111/mec.12481

Littlefair JE, Clare EL (2016) Barcoding the food chain: from sanger to high-throughput sequencing. Genome 59(11):946–958. https://doi.org/10.1139/gen-2016-0028

Liu M, Zhao Y, Sun Y, Wu P, Zhou S, Ren L (2020) Diatom DNA barcodes for forensic discrimination of drowning incidents. FEMS Microbiol Lett 367(17):fnaa145

Moussa TAA, Al-Zahrani HS, Almaghrabi OA, Abdelmoneim TS, Fuller MP (2017) Comparative metagenomics approaches to characterize the soil fungal communities of western coastal region, Saudi Arabia. PLoS One 12(9): e0185096. https://doi.org/10.1371/journal.pone.0185096

Ni J, Yan Q, Yu Y (2013) How much metagenomic sequencing is enough to achieve a given goal? Sci Rep 3(1):1968. https://doi.org/10.1038/srep01968

Op De Beeck M, Lievens B, Busschaert P, Declerck S, Vangronsveld J, Colpaert JV (2014) Comparison and validation of some ITS primer pairs useful for fungal metabarcoding studies. PLoS One 9(6):e97629. https://doi.org/10.1371/journal.pone.0097629

Oulas A, Pavloudi C, Polymenakou P, Pavlopoulos GA, Papanikolaou N, Kotoulas G, Arvanitidis C, Iliopoulos I (2015) Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. Bioinform Biol Insights 9:75–88. https://doi.org/10.4137/BBI.S12462

Procopio N, Ghignone S, Voyron S, Chiapello M, Williams A, Chamberlain A, Mello A, Buckley M (2020) Soil fungal communities investigated by metabarcoding within simulated forensic burial contexts. Front Microbiol 11:1686. https://doi.org/10.3389/fmicb.2020.01686

Rotimi AM, Pierneef R, Reva ON (2018) Selection of marker genes for genetic barcoding of microorganisms and binning of metagenomic reads by Barcoder software tools. BMC Bioinformatics 19(1):309. https://doi.org/10.1186/s12859-018-2320-1

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol 21(8):2045–2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x

Thomas T, Gilbert J, Meyer F (2012) Metagenomics - a guide from sampling to data analysis. Microb Inform Exp 2(1):3. https://doi.org/10.1186/2042-5783-2-3

Turenne CY, Sanche SE, Hoban DJ, Karlowsky JA, Kabani AM (1999) Rapid identification of fungi by using the ITS2 genetic region and an automated fluorescent capillary electrophoresis system. J Clin Microbiol 37(6):1846–1851. https://doi.org/10.1128/JCM.37.6.1846-1851.1999

Vadkertiová R, Molnárová J, Vránová D, Sláviková E (2012) Yeasts and yeast-like organisms associated with fruits and blossoms of different fruit trees. Can J Microbiol 58(12):1344–1352. https://doi.org/10.1139/cjm-2012-0468

van Nieuwenhuijzen EJ, Houbraken JA, Meijer M, Adan OC, Samson RA (2016) Aureobasidium melanogenum: a native of dark biofinishes on oil treated wood. Antonie Van Leeuwenhoek 109(5):661–683. https://doi.org/10.1007/s10482-016-0668-7

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73(16):5261–5267. https://doi.org/10.1128/AEM.00062-07

White TJ, Bruns TD, Lee SB, Taylor JW (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ (eds) PCR protocols: a guide to methods and applications. Academic Press, Cambridge (USA), pp 315–322

Wiltshire PE, Hawksworth DL, Webb JA, Edwards KJ (2014) Palynology and mycology provide separate classes of probative evidence from the same forensic samples: a rape case from southern England. Forensic Sci Int 244: 186–195. https://doi.org/10.1016/j.forsciint.2014.08.017

Young JM, Austin JJ, Weyrich LS (2017) Soil DNA metabarcoding and high-throughput sequencing as a forensic tool: considerations, potential limitations and recommendations. FEMS Microbiol Ecol 93(2):fiw207

Zajc J, Gunde-Cimerman N (2018) The genus Wallemia—from contamination of food to health threat. Microorganisms 6(2):46. https://doi.org/10.3390/microorganisms6020046

Publisher's Note