BMC Genomics

RESEARCH

Open Access

Unraveling the ancient fungal DNA from the Iceman gut

Nikolay Oskolkov[1*†], Anna Sandionigi[2,3†], Anders Götherström[4], Fabiana Canini[5], Benedetta Turchetti[6], Laura Zucconi[5], Tanja Mimmo[7], Pietro Buzzini[6] and Luigimaria Borruso[7*]

Abstract

Background Fungal DNA is rarely reported in metagenomic studies of ancient samples. Although fungi are essential for their interactions with all kingdoms of life, limited information is available about ancient fungi. Here, we explore the possibility of the presence of ancient fungal species in the gut of Ötzi, the Iceman, a naturally mummified human found in the Tyrolean Alps (border between Italy and Austria).

Methods A robust bioinformatic pipeline has been developed to detect and authenticate fungal ancient DNA (aDNA) from muscle, stomach, small intestine, and large intestine samples.

Results We revealed the presence of ancient DNA associated with *Pseudogymnoascus* genus, with *P. destructans* and *P. verrucosus* as possible species, which were abundant in the stomach and small intestine and absent in the large intestine and muscle samples.

Conclusion We suggest that Ötzi may have consumed these fungi accidentally, likely in association with other elements of his diet, and they persisted in his gut after his death due to their adaptability to harsh and cold environments. This suggests the potential co-occurrence of ancient humans with opportunistic fungal species and proposes and validates a conservative bioinformatic approach for detecting and authenticating fungal aDNA in historical metagenomic samples.

Keywords Fungi, Ancient DNA (aDNA), *Pseudogymnoascus*, Iceman, Ancient metagenomics

†Nikolay Oskolkov and Anna Sandionigi contributed equally to this work.

*Correspondence:
Nikolay Oskolkov
nikolay.oskolkov@scilifelab.se
Luigimaria Borruso
luigimaria.borruso@unibz.it
[1] Department of Biology, Science for Life Laboratory, National Bioinformatics Infrastructure Sweden, Lund University, Lund, Sweden
[2] Department of Informatics, Systems and Communication, University of Milan-Bicocca, Milan, Italy
[3] Quantia Consulting Srl, Milan, Italy
[4] Centre for Palaeogenetics, Department of Archaeology and Classical Studies, Stockholm University, Stockholm, Sweden
[5] Department of Ecological and Biological Sciences, University of Tuscia, 01100 Viterbo, Italy
[6] Department of Agricultural, Food and Environmental Sciences, University of Perugia, Perugia, Italy
[7] Faculty of Agricultural, Environmental and Food Sciences, Free University of Bozen-Bolzano, Piazza Università 5, Bolzano, Italy

Background

Palaeobiological approaches based on ancient DNA (aDNA) have extensively been used to study the evolution of humans, animals, and plants [1–5]. Recent advances in sedimentary and environmental ancient metagenomics provided unprecedented time series resolution for understanding the evolution of flora and fauna without macroscopic remains such as bones or teeth [6–9]. Studies based on whole shotgun metagenomic sequencing leverage exciting information on bacterial species present in ancient samples [10–12].

Nevertheless, several challenges complicate ancient microbial metagenomics analysis, such as modern contamination, which makes authentication analysis of major importance [13, 14]. In addition, several different microbial taxa are typically present in metagenomic samples [15]. Thus, studying ancient microbiomes, especially when it comes to fungi, remains a challenge.

Fungal DNA is often underrepresented in metagenomic samples due to several limitations of sequencing methods [16]. According to the last estimates, the kingdom of fungi includes up to twelve million species [17] that play pivotal ecological roles in all ecosystems (e.g., in soil or dead material) as saprotrophs and in association (e.g., as symbionts or parasites) with plants, animals, and humans, including their gut microbiome [18, 19]. In the last few years, several studies have demonstrated that fungi are crucial components of the gut microbiome [20]. Nevertheless, the gut microbiome is dramatically susceptible to the environment; thus, distinguishing between non-resident fungal species ingested due to diet or, more generally, acquired by the environment and resident taxa is challenging [19, 21].

Here, we hypothesized the possible presence of ancient fungal species associated with the gut of the Iceman, nicknamed Ötzi, the naturally mummified human body found in 1991 in the Similaun glacier at 3,200 m asl in the Tyrolean Alps (border between Italy and Austria) [22] (Fig. 1).

Some works have already demonstrated the presence of aDNA fragments of a few bacterial species in the Iceman's gut [12, 23]; however, to the best of our knowledge, there is no information regarding the possible presence of ancient fungi. To test our hypothesis, we applied a robust bioinformatic pipeline to detect the endogenous genetic signals of fungal aDNA by using sequence data from the stomach, small intestine, and large intestine and considering muscle tissue and extraction blank as the negative controls (Fig. 1).

Methods

The Iceman dataset analyzed in this study was downloaded from the European Nucleotide Archive (ENA) under accession number ERP012908; all information about the samples and sequencing details can be found in Maixner et al. [12]. Raw sequences were analyzed using a bioinformatic pipeline consisting of three main phases: i) quality control, ii) taxonomic identification of the fungal taxa, and iii) authentication analysis of the detected fungi. All pipeline steps are illustrated in the flowchart of Fig. 2. The software BBDuk and BBMerge [24] were used to trim Illumina adapters and merge overlapping pair-end metagenomic reads. FastQC was used to inspect the quality metrics of the pre-processed reads [25]. More information about data pre-processing and quality control can be found in the Supplementary Information (S.M. 1 Quality Control subsection) .

Next, trimmed and merged reads were taxonomically classified via a *k*-mer-based approach with KrakenUniq [26] to get an overview of the microbial community in the samples and retrieve a fast taxonomic assignment for each read. Importantly, for unbiased and robust

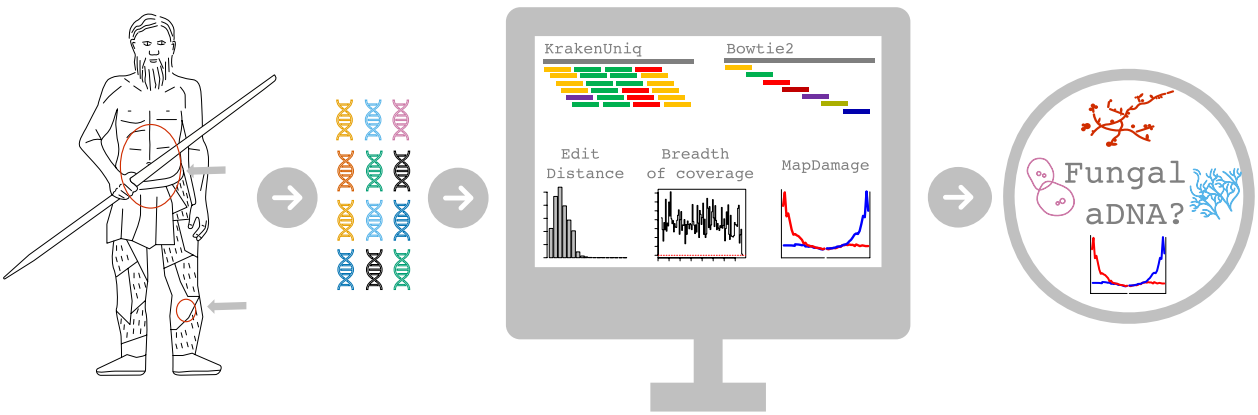

Fig. 1 Graphical abstract of workflow about the ancient fungal DNA detection in iceman sample

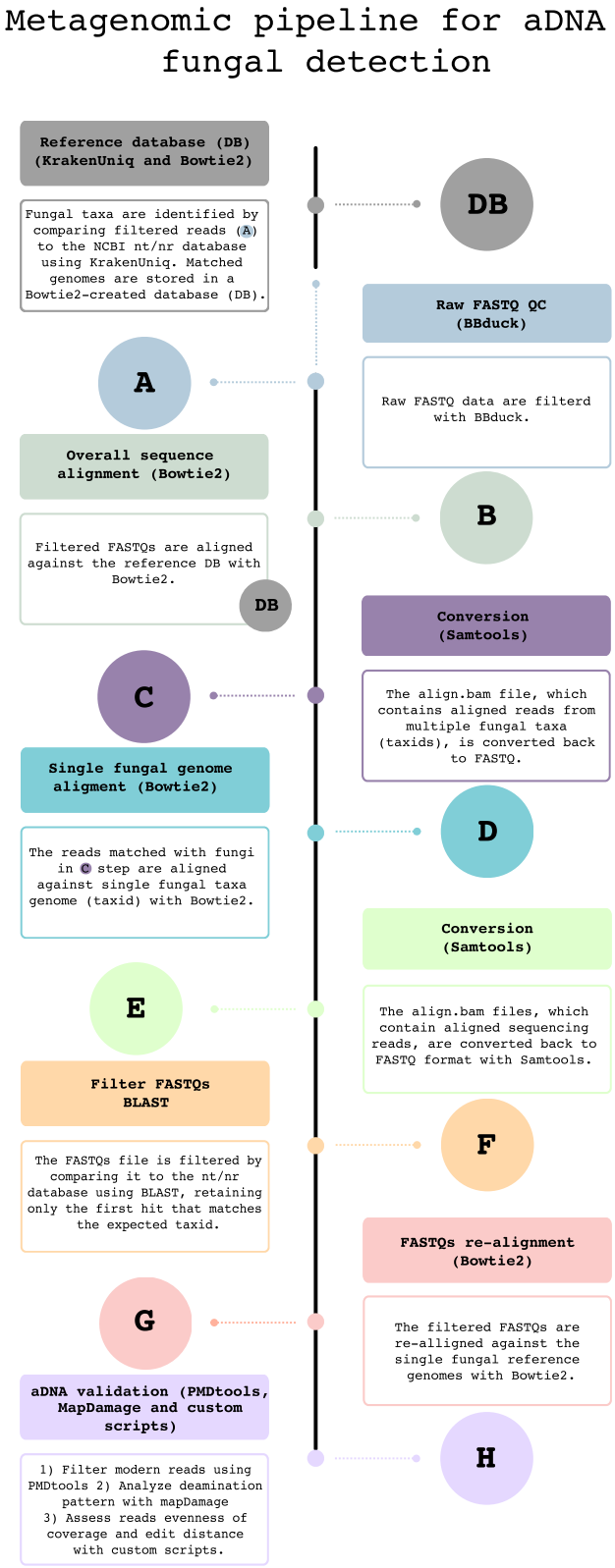

Fig. 2 Flowchart of the bioinformatic pipeline used to analyse the Iceman fungal community

taxonomic classification with KrakenUniq, we used one of the largest and most diverse nucleotide databases, which is NCBI GenBank (NT), accessed in December 2020, that included all microbial, vertebrate, non-vertebrate, and plant organisms, specifically 73,300,584 reference sequences corresponding to 9,556,710 taxa with 1,855,732 species. The KrakenUniq NCBI GenBank nucleotide (NT) database is publicly available through the SciLifeLab Figshare at https://doi.org/https://doi.org/10.17044/scilifelab.20205504. The organisms identified by KrakenUniq were conservatively filtered to minimize the number of false-positive discoveries, i.e. an organism was considered "detected" by KrakenUniq if it had at least 200 reads and at least 1000 unique *k*-mers assigned to it. The former quality metric implies filtering for depth of coverage, and the latter quality metric corresponds to the breadth of coverage, i.e. the reads have to uniformly cover a fungal reference genome for the fungus to be considered to be present in a sample. The KrakenUniq thresholds for depth and breadth of coverage used in this study were suggested in [14] and [26] as appropriate for filtering microbial community hits, see Supplementary Information (S.M. 1 K-mer based profiling with KrakenUniq subsection) and Figure S.M. 2 for more details. The organisms detected by KrakenUniq in the Iceman sample were visualized as Krona plots [27].

Further, we excluded the microorganisms found in the blank negative control samples from the final list of microbes detected by KrakenUniq. Hence, we did not proceed with validation and authentication analyses of these microbes.

Complementary to the *k*-mer-based classification with KrakenUniq, the pre-processed reads were also aligned via competitive mapping procedure to the indexed NCBI GenBank (NT) database with Bowtie2 [28] (the Bowtie2 index for the database is publicly available via SciLifeLab Figshare at https://doi.org/https://doi.org/10.17044/scilifelab.21070063), which allowed for a preliminary inspection of coverage and deamination signals with mapDamage [29]. Next, although the NCBI GenBank (NT) database is optimal in terms of its sensitivity for organism detection because it includes a variety of organisms, it often lacks good-quality reference genomes for some taxa. Therefore, alignments against a good-quality reference genome are typically required for a comprehensive follow-up of an interesting candidate. Thus, according to the taxonomic composition found with KrakenUniq, reference genomes of the most abundant fungal species were downloaded from the NCBI RefSeq resource, and the pre-processed metagenomic reads went through a second alignment round to the reference genomes by Bowtie2 [28]. These new good-quality alignments were used for computing a deamination profile with

mapDamage [29] and evenness of coverage with Samtools [30]. Only fungal candidates that demonstrated convincing deamination and evenness of coverage profiles were considered as likely present in a sample and of ancient origin. For more details about the alignment procedure, see Supplementary Information (S.M 1 Alignment subsection).

To further verify KrakenUniq and Bowtie2 findings, we also implemented another alignment iteration via BLAST [31]. For this purpose, the reads aligned to fungal references were extracted and used for BLAST analysis with the NCBI NR/NT database. Only sequences matching the first hit with the expected *taxid* were retained.

To follow up the most interesting fungal candidates, further sorting of damaged reads was performed with PMDtools [32]. To confirm the previously observed signals, we again evaluated the reads selected by BLAST and PMDtools regarding their deamination pattern with mapDamage [29] and evenness of coverage with Samtools [30]. For more details about the authentication procedure, see Supplementary Information (S.M. 1 Authentication analysis subsection).

Finally, Jellyfish [33] was utilized to extract and count the *k*-mers (with k = 5) of both the potential ancient fungal candidates (where *samtools consensus* was used for generating consensus fasta-files from bam-alignments) and all reference genome species belonging to their respective genera (i.e. *Pseudogymnoascus*) available in NCBI [33]. Additionally, four outgroup species (*Podosphaera leucotricha* PuE-3 k121 111, *Oidiodendron maius* Zn, *Xenosphaeropsis pyriputrescens* CBS 115176, and *Drepanopeziza brunnea* 'multigermtubi' MB m) selected from the same taxonomic phylum (Ascomicota) and class (Leotiomycetes) of *Pseudogymnoascus*, were included, along with two species from the same phylum (*Saturnispora hagleri* NRRL Y − 27,828 and *Saccharomyces cerevisiae* S288C) (Ascomicota).

The matrix of *k*-mer counts across the organisms produced by Jellyfish was normalized and log-transformed with log10 (x + 1), where x is the *k*-mer count, to make it more appropriate for cluster analysis [33]. Finally, cluster analysis based on Manhattan distance was performed with the "vegan" and "pvclust" R packages [34, 35], and the final hierarchical cluster dendrogram is presented in Fig. 5. Uncertainty of each cluster in the dendrogram was computed via bootstrap re-sampling procedure and is highlighted by red Approximately Unbiased (AU) and green Bootstrap Probability (BP) metrics. See Supplementary Information (S.M. 1 Cluster analysis subsection) for more details.

In summary, the bioinformatic pipeline applied in this work comprised multiple complementary detection, validation, and authentication steps to retrieve the most confident fungal signals. We focused only on highly reliable fungal hits, i.e., the ones detected by KrakenUniq, Bowtie2, and BLAST, and successfully validated and authenticated by Samtools, mapDamage, and PMDtools. All the scripts from the bioinformatic pipeline used in this study and the main output files, as well as computer simulation data and scripts used for testing KrakenUniq thresholds, can be found at https://github.com/NikolayOskolkov/Iceman_fungi.

Results and discussion

First of all, our bioinformatic pipeline was able to reproduce the main microbial findings of the Iceman reported by Maixner et al. [12], such as *Helicobacter pylori* and a few other bacteria belonging to *Clostridium* and *Pseudomonas* genera (S.M. 2 Table; S.M. 3–4 Figures and S.M. Figure 5–6). However, since this study focuses on the ancient fungal community of the Iceman samples, we will concentrate only on fungal hits and discuss their abundance and quality metrics (Figure S.M. 7).

The fungal species identified by KrakenUniq and validated with conservative depth and breadth filters are presented in Table 1. Two fungal genera were standing out as particularly abundant: *Rhodotorula* with the species *R. toruloides* (19,1057 reads and 14,868 unique *k*-mers assigned) and *Pseudogymnoascus* with *P. destructans* (1,298,527 reads and 473,667 unique *k*-mers assigned), *P. verrucosus* (437,040 reads and 405,670 unique *k*-mers assigned) and *P. pannorum* (787,723 reads and 405,670 unique *k*-mers assigned) (Table 1). The assigned unique *k*-mers for the two genera are far above our conservative detection limit of at least 1000 unique *k*-mers, which implies high confidence in discovering the two fungal genera.

Further, the competitive mapping procedure with Bowtie2 against the full NCBI GenBank nucleotide (NT) database confirmed that the number of reads mapped uniquely (i.e. with the mapping quality MAPQ above 1) was 2605 with *R. toruloides;* 472,871 reads with *P. destructans*; 170,837 with *P. pannorum* and 241,46 with *P. verrucosus* (Table 1). This indicates that mapped reads have a strong affinity towards a specific species since the reads were exposed via competitive mapping to millions of other reference sequences in the NCBI GenBank nucleotide (NT) database; however, they had much higher similarity with respect to the reported fungal species. Additionally, the breadth of coverage yielded the following results: 472,871 reads exhibited uniform coverage across the scaffolds of the representative genome of *P. destructans* (total length > 35.8 Mb), constituting 11.33% of the entire genome (Table 1). Furthermore, 241,461 reads encompassed 9.89% of the genome of *P. verrucosus* (total length > 30.17 Mb) (Table 1). In contrast,

Table 1 Detection, validation, and authentication metrics for fungal species discovered in the Iceman tissues. The table includes: taxID – taxonomic id according to NCBI nomenclature; Pers_Reads – fraction of reads covered by the clade rooted at this taxon; reads—number of reads covered by the clade rooted at this taxon; taxReads—number of reads assigned directly to this taxon; kmers – number of unique *k*-mers assigned to this taxon; dup—average number of times each unique *k*-mer has been seen and cov—coverage of the *k*-mers of the clade in the database

Taxonomy			KrakenUniq output							Reference genomes	Total reads mapped		Ancients reads (PMD3)	
Phylum	Genus	Species	taxID	Pers_Reads	reads	taxReads	kmers	dup	cov	GenBank assembly accession	n. reads	Genome covered	n. reads	Genome covered
Ascomycota	*Antarctomyces*	*A. pellizariae*	1,955,577	0.001575	6821	6821	1056	190	0.04621	GCA_010623925.1	4576	15.42%	82	/
	Aspergillus	*A. nidulans*	162,425	0.00148	6407	580	135,510	1.16	0.004575	GCF_000149205.2	237	0.03%	1	/
		A. niger	5061	0.00009031	391	222	1864	2.57	0.00008332	GCF_000002855.3	258	0.03%	69	/
	Botrytis	*B. cinerea*	40,559	0.0004633	2006	1557	1524	8.21	0.00003424	GCF_000143535.2	620	0.04%	0	/
	Pseudogymnoascus	*P. destructans*	655,981	0.2999	1,298,527	60,481	473,667	37.5	0.03685	GCF_001641265.1	472,871	11.33%	43,682	3.62%
		P. pannorum	79,858	0.1009	437,040	205	4830	1830	0.1145	GCA_001630605.1	170,837	0.69%	18,161	0.04%
		P. verrucosus	342,668	0.1819	787,723	569	405,670	21.7	0.03034	GCF_001662655.1	241,461	9.89%	29,647	3.02%
	Saccharomyces	*S. cerevisiae*	4932	0.000212	918	641	15,459	1.24	0.0006803	GCF_000146045.2	1577	0.09%	6	/
Basidiomycota	*Anthracocystis*	*A. flocculosa*	84,751	0.00434	18,791	253	2257	54.6	0.0001896	GCF_000417875.1	7	0.08%	0	/
	Apiotrichum	*A.mycotoxinovorans*	252,803	0.003578	15,492	9198	1413	69.8	0.00004321	GCA_013177335.1	414	0.02%	10	/
		A. porosum	105,984	0.00244	10,563	373	1111	54.2	0.00008571	GCF_003942205.1	522	0.04%	21	/
	Leucosporidium	*L. scottii*	5278	0.006938	30,038	5754	1718	185	0.001616	GCF_000320785.1	26,250	1.196%	6	/
	Malassezia	*M. globosa*	76,773	0.0001894	820	746	6336	1.88	0.0007186	GCF_000181695.1	620	0.06%	2	/
		M. restricta	76,775	0.002037	8820	8477	54,633	2.27	0.006595	GCF_003290485.1	1672	0.36%	2	/
	Phaffia	*P. rhodozyma*	264,483	0.001012	4382	652	1887	16.3	0.0001045	GCA_014706385.1	577	0.10%	10	/
	Phenoliferia	*P. glacialis*	418,497	0.003383	14,645	13,461	1056	72.5	0.2884	GCA_017654785.1	0	/	0	/
	Rhodotorula	*R. toruloides*	5286	0.04413	191,057	127,887	14,868	77.8	0.0007132	GCF_000320785.1	2605	0.23%	3	/
	Schizophyllum	*S. commune*	5334	0.002495	10,804	249	1185	53.5	0.00006565	GCF_000143185.1	12	0.01%	0	/
	Sporisorium	*S. reilianum*	72,558	0.002726	11,801	8808	1294	44.4	0.00005889	GCA_900162835.1	0	/	0	/

* Explanation of columns for KrakenUniq output: taxID – taxonomic id according to NCBI nomenclature, Pers_Reads – fraction of reads covered by the clade rooted at this taxon, reads—number of reads covered by the clade rooted at this taxon, taxReads—number of reads assigned directly to this taxon, kmers – number of unique *k*-mers assigned to this taxon, dup—average number of times each unique *k*-mer has been seen, cov—coverage of the *k*-mers of the clade in the database

P. pannorum and *R. toruloides* hits had breadth coverage < 1% of the reference genomes coverage; thus, we discarded both taxa from further analysis (Table 1). *P. destructans* and *P. verrucosus* had a moderate mismatch rate captured by their edit distance plots, i.e., most reads had one mismatch, which is not uncommon for damaged ancient reads (S.M. 8 Figures a and b) [36].

Deamination analysis in *P. destructans* and *P. verrucosus* showed a signal of deamination profiles with transition mutation frequencies reaching 4–5% at the ends of the reads (S.M. 9 Figures a and c). The distribution of reads mapped to *P. destructans* and *P. verrucosus* demonstrated a seemingly bimodal profile with a first peak at a read length of ~ 45 bp and a second at ~ 90 bp (S.M. 9 Figures b and d). The combination of relatively weak deamination, moderate mismatch rate, and bimodal read length distribution can imply that a complex mixture of different reads was mapped to the *P. destructans* and *P. verrucosus* representative genomes. We assume two reasons can explain the bimodal read length distribution and relatively weak overall read deamination: the presence of modern contamination and non-specific alignment to a homologous reference. First, DNA sequences from both ancient (damaged and short) and likely modern (not damaged and long) *P. destructans* and *P. verrucosus* can potentially be present in the Iceman samples [36]. Another explanation can be that ancient and modern reads from two or more species from the *Pseudogymnoascus* genus, whose reference genomes were not present in NCBI NT database, were mismapped to the *P. destructans* and *P. verrucosus* representative genomes [36].

Therefore, already at this stage, we can conclude that at least the *Pseudogymnoascus* genus was confidently detected in the Iceman data, while it is less certain that the exact two species, i.e. *P. destructans* and *P.verrucosus*, are present in the Ötzi's metagenomic samples. To exclude reads potentially coming from modern contamination, we filtered reads with a PMD score above 3, which implies a high confidence damage pattern [32]. Indeed, a typical authentication approach includes monitoring the fragmentation of DNA sequences and the deamination pattern of cytosine bases, which occur with a higher frequency at the extremities of the aDNA fragment [29]. In addition, we considered only reads shorter than 100 bp, which were more likely to be of ancient origin than longer reads [29]. This filtering step significantly improved the clarity of the deamination signal, where the frequency of transition mutations reached now ~ 20–30% at the ends of the reads (Fig. 3a and 3c) while high numbers of reads were still retained after the harsh filtering, i.e. 43,682 reads for *P. destructans* (3.02% of the genome covered), and 29,647 reads for *P. verrucosus* (3.62% of the genome covered) (Fig. 4 a-b-c-d). The edit distance in Figs. 4a and 4c demonstrated a relatively high mismatch rate, which can be expected since the PMD filtering prioritized most

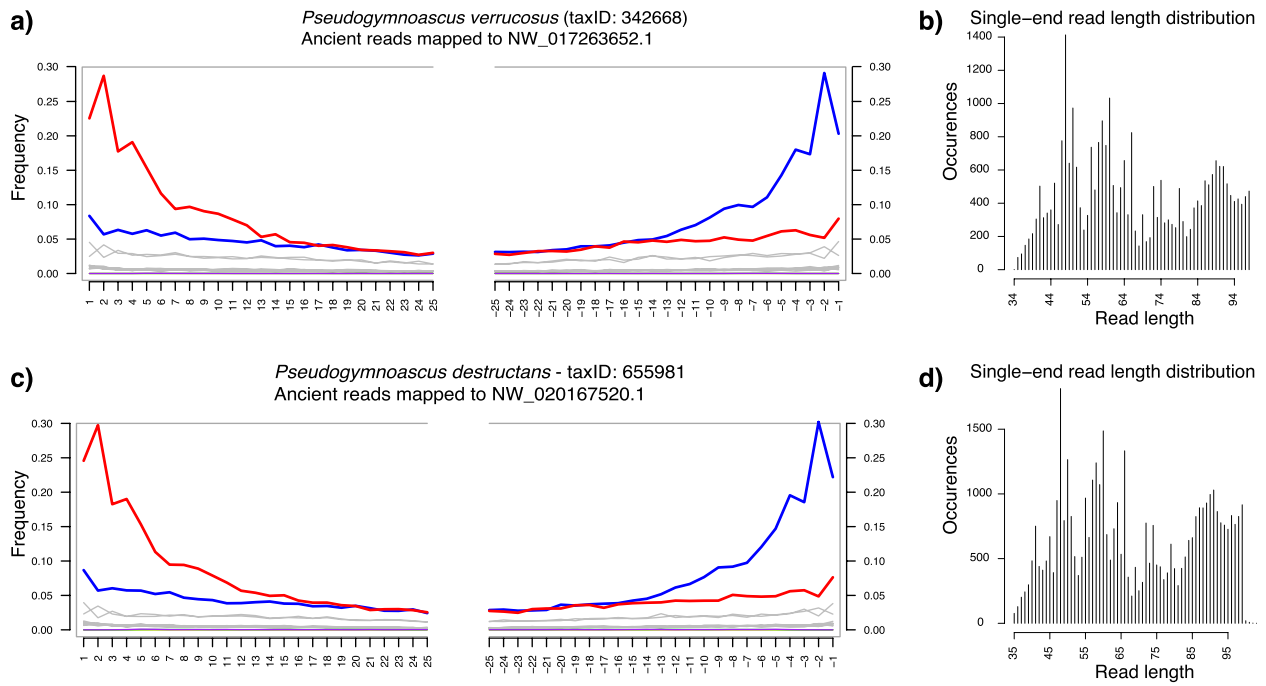

Fig. 3 Deamination pattern and read length distribution of *P. verrucosus* (**a**, **b**) and *P. destructans* (**c**, **d**) after filtering for read length (max. 100 bp long reads were selected) and damage (reads with PMD score > 3 were selected)

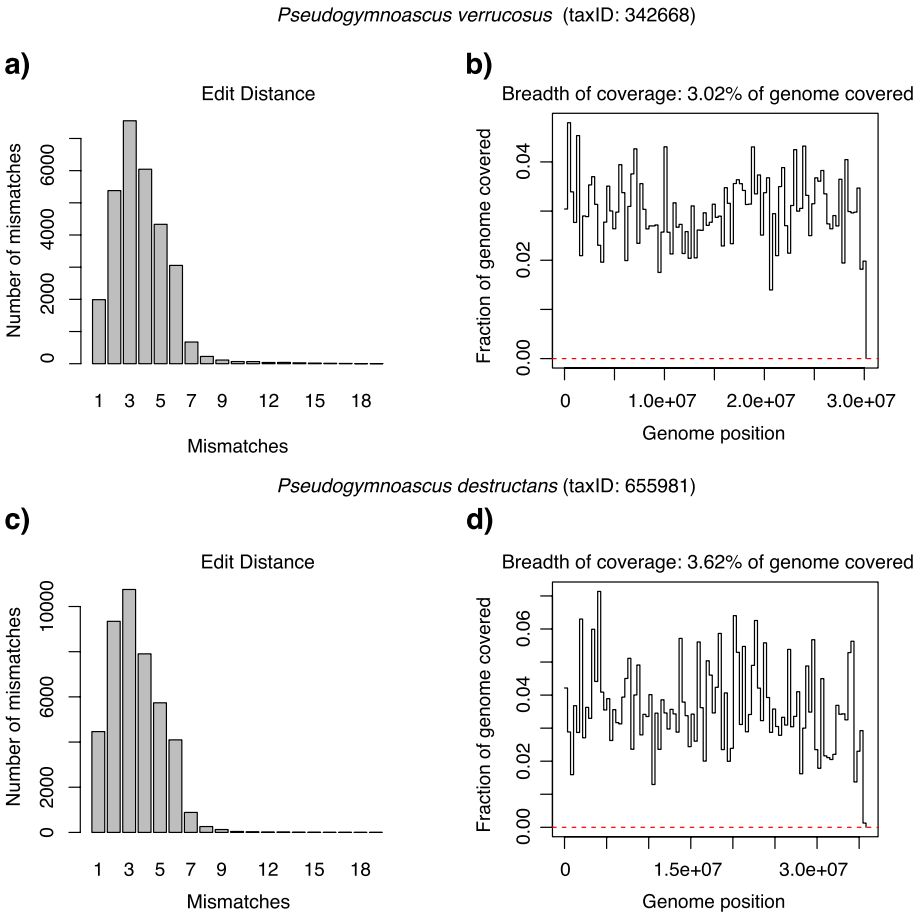

Fig. 4 Edit distance and evenness of coverage of *P. verrucosus* (**a** and **b**) and *P. destructans ancients* reads (PMD score > 3) (**c** and **d**)

deaminated/damaged reads with an increased mismatch rate toward the reference.

We proceeded with neighbor-joining phylogenetic/cluster analysis to further validate the presence and ancient origin of the fungal species found in this study (Fig. 5). The neighbor-joining phylogenetic relationships of the ancient reads aligned against the two potential fungal candidates were analyzed via an alignment-free approach based on *k*-mer statistics. Cluster analysis derived from *k*-mer frequency matrices of fungal species can provide valuable information regarding the taxonomic relationship among the different species [26]. Briefly, taxa that are grouped together are expected to have genetic characteristics that are more similar to those in other clusters. Both fungal candidates grouped together in the cluster with the other *Pseudogymnoascus* and not with any outgroup species (Fig. 5). This further confirms that we robustly identified *Pseudogymnoascus* genus in the Iceman data. However, the exact species is less confident. The two ancient candidates formed a separate sub-cluster, Fig. 5, which can be explained by a few reasons. Most likely, the two ancient *Pseudogymnoascus* fungi were clustering together and separately from the modern *Pseudogymnoascus* species due to the damage and missingness (i.e., only ~ 3% of their reference genomes were covered) in their genomic data (Fig. 4b, d). Another explanation can be their evolutionary diversity from the modern *Pseudogymnoascus* species.

After the different validation steps, we examined the distribution of the mapped reads of *P. destructans* and *P. verrucosus* along the different body sites of the Iceman. In both cases, the fungal reads were particularly abundant in the stomach and, to a lesser extent, in the small intestine compartments (Fig. 6) (Table 2). Fewer than five ancient reads were assigned to the two fungi in the large intestine, muscle, and extraction blank (Table 2). A few assigned/aligned reads typically reflect the noise level in metagenomic analysis and cannot serve as reliable evidence of microbial detection [14]. This confirms that *P. destructans* and *P. verrucosus* were unlikely from post-mortem environmental or laboratory contamination. Indeed, if the two fungal species in Ötzi's stomach

Fig. 5 Dendrogram based on Manhattan distance and built using the *k*-mer method. The dendrogram was constructed using all the genomes of the species belonging to Pseudogymnoascus available in NCBI, along with four outgroup species belonging to the same taxonomic phylum (*Ascomicota*) and class (*Leotiomycetes*) as *Pseudogymnoascus*, and two from the same phylum (*Ascomicota*). The two ancient candidates are highlighted in blue

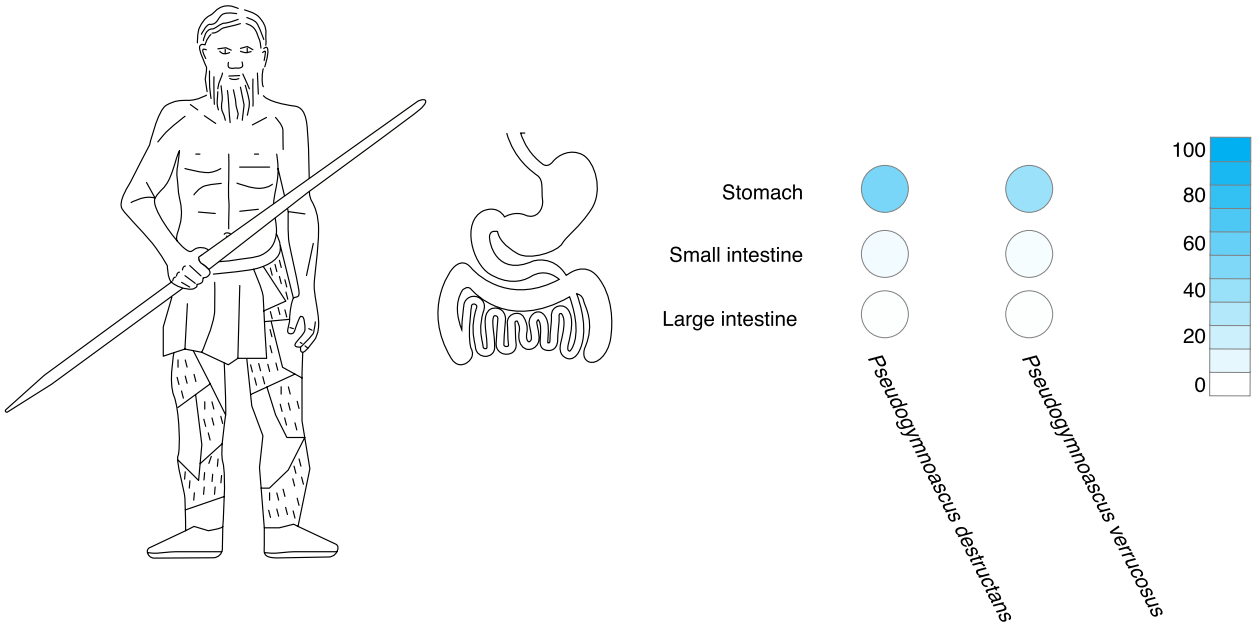

Fig. 6 Relative abundance of *P. verrucosus* and *P. destructans* in different Iceman samples

Table 2 Absolute number and frequency of the ancient reads across different body sites; n.d.: not detected

Body site	*P. verrucosus* (taxID: 342,668)		*P. destructans* (taxID: 655,981)	
	Number of ancient reads	Frequency of ancient reads	Number of ancient reads	Frequency of ancient reads
Stomach	28,969	3.6%	42,770	5.3%
Small intestine	677	38.9%	910	52.3%
Large intestine Upper part	n.d	/	n.d	/
Muscle tissue	n.d	/	n.d	/
Extraction Blank	n.d	/	n.d	/

reflected the local glacial environment, one would also expect to find them in the muscle sample, which was not the case in our analysis.

The presence of *Pseudogymnoascus* in the gut of the Iceman presents a complex puzzle. *Pseudogymnoascus* encompasses species that are psychrophilic or psychrotolerant saprotrophs, commonly found in diverse environments such as soil and associated with plant material[37]. Additionally, *P. destructans* can act as opportunistic pathogens, causing skin and respiratory tract infections [37, 38]. Palmer et al. suggested that *P. destructans* is likely a fungal pathogen of bats, the causative agent of white-nose syndrome, evolving alongside Eurasian bat species for millions of years [39]. Our proposed explanation is that *Pseudogymnoascus* in the Iceman's gut may reflect his dietary habits and the broader environmental conditions of his time. It is plausible that the Iceman could have accidentally ingested the fungal species through the diet. For example, the most likely explanation is that food items became contaminated before ingestion, so the fungus entered the Iceman's stomach. The absence of the *Pseudogymnoascus* fungus in other parts of Ötzi's body, such as muscle tissue, supports this hypothesis, suggesting that colonization from ice or permafrost is unlikely. After the iceman's death, the fungus found suitable survival and growth conditions for its proliferation. This hypothesis is also supported by the ecology of the *Pseudogymnoascus* species, which has a long-term persistence and can survive and grow at a temperature lower than −20 °C and in harsh environments [38, 40, 41]. Moreover, there is evidence that some *Pseudogymnoascus* species, such as *P. destructans,* can survive in the gastrointestinal tract [42].

Our study delved into historical metagenomic samples to unravel the presence of fungal aDNA. Despite the rarity of such findings in the literature, we successfully

identified the presence of the *Pseudogymnoascus* genus in the Iceman, providing potential insights into its co-occurrence with ancient humans. Furthermore, we developed a conservative bioinformatic approach that can be applied to other types of ancient sampling to detect and validate the possible presence of aDNA from filamentous fungi and yeasts.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-024-11123-2.

Supplementary Material 1.
Supplementary Material 2.
Supplementary Material 3.
Supplementary Material 4.
Supplementary Material 5.
Supplementary Material 6.
Supplementary Material 7.
Supplementary Material 8.
Supplementary Material 9.
Supplementary Material 10.

Acknowledgements
We are grateful to Manuela Dasser of Marameolab for the graphical illustrations. NO is financially supported by Knut and Alice Wallenberg Foundation as part of the National Bioinformatics Infrastructure Sweden at SciLifeLab. This work was supported by the Open Access Publishing Fund of the Free University of Bozen-Bolzano.

Authors' contributions
Conceptualization, L.B, N.O., P.B. and A.S.; methodology, L.B, N.O. and A.S; formal analysis, L.B, N.O., A.S and F.C..; investigation, N.O., A.S., A.G., F.C., B.T., L.Z., T.M., P.B. and L.B.; writing – original draft, L.B. and N.O.; writing – review & editing, N.O., A.S., A.G., F.C., B.T., L.Z., T.M., P.B. and L.B.; supervision L.B, N.O., and A.S; project administration, L.B.; funding acquisition, L.B. and N.O.

Funding
Not applicable.

Data availability
All the scripts from the bioinformatic pipeline used in this study and the main output files, as well as computer simulation related files and information, can be found at https://github.com/NikolayOskolkov/Iceman_fungi.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 21 October 2024 Accepted: 3 December 2024
Published online: 19 December 2024

References

1. Rasmussen M, Li Y, Lindgreen S, Pedersen JS, Albrechtsen A, Moltke I, et al. Ancient human genome sequence of an extinct Palaeo-Eskimo. Nature. 2010;463:757–62.
2. Van Der Valk T, Pečnerová P, Díez-del-Molino D, Bergström A, Oppenheimer J, Hartmann S, et al. Million-year-old DNA sheds light on the genomic history of mammoths. Nature. 2021;591:265–9.
3. Skoglund P, Ersmark E, Palkopoulou E, Dalén L. Ancient Wolf Genome Reveals an Early Divergence of Domestic Dog Ancestors and Admixture into High-Latitude Breeds. Curr Biol. 2015;25:1515–9.
4. Scott MF, Botigué LR, Brace S, Stevens CJ, Mullin VE, Stevenson A, et al. A 3,000-year-old Egyptian emmer wheat genome reveals dispersal and domestication history. Nat Plants. 2019;5:1120–8.
5. Fracasso I, Dinella A, Giammarchi F, Marinchel N, Kołaczek P, Lamentowicz M, et al. Climate and human impacts inferred from a 1500-year multi-proxy record of an alpine peatland in the South-Eastern Alps. Ecol Indic. 2022;145: 109737.
6. Weyrich LS, Duchene S, Soubrier J, Arriola L, Llamas B, Breen J, et al. Neanderthal behaviour, diet and disease inferred from ancient DNA in dental calculus. Nature. 2017;544:357–61.
7. Vernot B, Zavala EI, Gómez-Olivencia A, Jacobs Z, Slon V, Mafessoni F, et al. Unearthing Neanderthal population history using nuclear and mitochondrial DNA from cave sediments. *Science* 2021; 372: eabf1667.
8. Pedersen MW, De Sanctis B, Saremi NF, Sikora M, Puckett EE, Gu Z, et al. Environmental genomics of Late Pleistocene black bears and giant short-faced bears. Curr Biol. 2021;31:2728-2736.e8.
9. Wang Y, Pedersen MW, Alsos IG, De Sanctis B, Racimo F, Prohaska A, et al. Late Quaternary dynamics of Arctic biota from ancient environmental genomics. Nature. 2021;600:86–92.
10. Rasmussen S, Allentoft ME, Nielsen K, Orlando L, Sikora M, Sjögren K-G, et al. Early Divergent Strains of Yersinia pestis in Eurasia 5,000 Years Ago. Cell. 2015;163:571–82.
11. Sarhan MS, Lehmkuhl A, Straub R, Tett A, Wieland G, Francken M, et al. Ancient DNA diffuses from human bones to cave stones. *iScience*. 2021;24:103397.
12. Maixner F, Krause-Kyora B, Turaev D, Herbig A, Hoopmann MR, Hallows JL, et al. The 5300-year-old *Helicobacter pylori* genome of the Iceman. Science. 2016;351:162–5.
13. Key FM, Posth C, Krause J, Herbig A, Bos KI. Mining Metagenomic Data Sets for Ancient DNA: Recommended Protocols for Authentication. Trends Genet. 2017;33:508–20.
14. Pochon Z, Bergfeldt N, Kırdök E, Vicente M, Naidoo T, Van Der Valk T, et al. aMeta: an accurate and memory-efficient ancient metagenomic profiling workflow. Genome Biol. 2023;24:242.
15. Mande SS, Mohammed MH, Ghosh TS. Classification of metagenomic sequences: methods and challenges. Brief Bioinform. 2012;13:669–81.
16. De Vries RP, Tsang A, Grigoriev IV (eds). Fungal Genomics: Methods and Protocols. 2018. Springer New York, New York, NY.
17. Wu B, Hussain M, Zhang W, Stadler M, Liu X, Xiang M. Current insights into fungal species diversity and perspective on naming the environmental DNA sequences of fungi. Mycology. 2019;10:127–40.
18. Tedersoo L, Bahram M, Põlme S, Kõljalg U, Yorou NS, Wijesundera R, et al. Global diversity and geography of soil fungi. Science. 2014;346:1256688.
19. Borruso L, Checcucci A, Torti V, Correa F, Sandri C, Luise D, et al. I Like the Way You Eat It: Lemur (Indri indri) Gut Mycobiome and Geophagy. Microb Ecol. 2021;82:215–23.
20. Huseyin CE, O'Toole PW, Cotter PD, Scanlan PD. Forgotten fungi—the gut mycobiome in human health and disease. FEMS Microbiol Rev. 2017;41:479–511.
21. Lavrinienko A, Scholier T, Bates ST, Miller AN, Watts PC. Defining gut mycobiota for wild animals: a need for caution in assigning authentic resident fungal taxa. Anim Microbiome. 2021;3:75.
22. Müller W, Fricke H, Halliday AN, McCulloch MT, Wartho J-A. Origin and Migration of the Alpine Iceman. Science. 2003;302:862–6.
23. Mancabelli L, Turroni F, Ferrario C, Duranti S, Van Sinderen D, Ventura M. Ancient bacteria of the Ötzi's microbiome: a genomic tale from the Copper Age. Microbiome. 2017;5:5.
24. Bushnell B, Rood J, Singer E. BBMerge – Accurate paired shotgun read merging via overlap. PLoS ONE. 2017;12: e0185056.
25. Andrews S. FastQC: a quality control tool for high throughput sequence data. 2010. 2017.

26. Breitwieser FP, Baker DN, Salzberg SL. KrakenUniq: confident and fast metagenomics classification using unique k-mer counts. Genome Biol. 2018;19:198.
27. Ondov BD, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 2011;12:385.
28. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9.
29. Jónsson H, Ginolhac A, Schubert M, Johnson PLF, Orlando L. mapDamage2.0: fast approximate Bayesian estimates of ancient DNA damage parameters. *Bioinformatics* 2013; 29: 1682–1684.
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9.
31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
32. Skoglund P, Northoff BH, Shunkov MV, Derevianko AP, Pääbo S, Krause J, et al. Separating endogenous ancient DNA from modern day contamination in a Siberian Neandertal. Proc Natl Acad Sci. 2014;111:2229–34.
33. Marçais G, Kingsford C. A fast, lock-free approach for efficient parallel counting of occurrences of *k* -mers. Bioinformatics. 2011;27:764–70.
34. Jari Oksanen, Gavin L. Simpson, F. Guillaume Blanchet, Roeland Kindt, Pierre Legendre, Peter R. Minchin, R.B. O'Hara, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs, Helene Wagner, Matt Barbour, Michael Bedward, Ben Bolker, Daniel Borcard, Gustavo Carvalho, Michael Chirico, Miquel De Caceres, Sebastien Durand, Heloisa Beatriz Antoniazi Evangelista, Rich FitzJohn, Michael Friendly, Brendan Furneaux, Geoffrey Hannigan, Mark O. Hill, Leo Lahti, Dan McGlinn, Marie-Helene Ouellette, Eduardo Ribeiro Cunha, Tyler Smith, Adrian Stier, Cajo J.F. Ter Braak, James Weedon. vegan: Community Ecology Packag. Ordination methods, diversity analysis and other functions for community and vegetation ecologists.
35. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics. 2006;22:1540–2.
36. Der Sarkissian C, Velsko IM, Fotakis AK, Vågene ÅJ, Hübner A, Fellows Yates JA. Ancient Metagenomic Studies: Considerations for the Wider Scientific Community. *mSystems* 2021; 6: e01315–21.
37. Quandt CA, Haelewaters D. Phylogenetic Advances in Leotiomycetes, an Understudied Clade of Taxonomically and Ecologically Diverse Fungi. In: Zaragoza Ó, Casadevall A (eds). *Encyclopedia of Mycology*. 2021. Elsevier, Oxford, pp 284–294.
38. Hayes MA. The Geomyces Fungi: Ecology and Distribution. Bioscience. 2012;62:819–23.
39. Palmer JM, Drees KP, Foster JT, Lindner DL. Extreme sensitivity to ultraviolet light in the fungal pathogen causing white-nose syndrome of bats. Nat Commun. 2018;9:35.
40. Veselská et al. - 2020 - Comparative eco-physiology revealed extensive enzy.pdf.
41. Hoyt JR, Langwig KE, Okoniewski J, Frick WF, Stone WB, Kilpatrick AM. Long-Term Persistence of Pseudogymnoascus destructans, the Causative Agent of White-Nose Syndrome, in the Absence of Bats. EcoHealth. 2015;12:330–3.
42. Ballmann AE, Torkelson MR, Bohuski EA, Russell RE, Blehert DS. Dispersal hazards of pseudogymnoascus destructans by bats and human activity at hibernacula in summer. J Wildl Dis. 2017;53:725.

Publisher's Note