Check for updates

DATA NOTE

The Blastocystis Biobank: a curated, multi-host biological resource of cultures, nucleic acids, metabolites, and stool-derived aliquots with a controlled-access virtual catalogue for global research

[version 1; peer review: awaiting peer review]

Jaya Makkimane[1], Ana M. Figueiredo[2], Barbara Šoba[3], Carolina Allievi[4], Constance Denoyelle[5], David Carmena[6,7], Eleni Gentekaki[8], Eric Viscogliosi[5], Isabel Guadano-Procesi[9], Julia Walochnik[10], Laura Hurtado[11], Laura Lerma[12], María José Irisarri-Gutiérrez[12], Rens Zonneveld[13], Sabiha Sensöz[14], Vasana Jinatham[15], Yaseen Majid Salman Al-Adilee[16], Anastasios D. Tsaousis[1]

[1]School of Natural Sciences, University of Kent, Canterbury, England, UK
[2]Department of Biology & CESAM, University of Aveiro, Aveiro, Portugal
[3]Institute of Microbiology and Immunology, Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia
[4]Department of Veterinary Medicine and Animal Sciences, Università degli Studi di Milano, 26900 Lodi, Italy
[5]Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 9017 - CIIL - Centre d'Infection et d'Immunité de Lille, Lille, France
[6]Parasitology Reference and Research Laboratory, National Centre for Microbiology, Health Institute Carlos III, Majadahonda (Madrid), Spain
[7]Centre for Biomedical Research Network in Infectious Diseases (CIBERINFEC), Health Institute Carlos III, Madrid, Spain
[8]Department of Veterinary Medicine, University of Nicosia School of Veterinary Medicine, Nicosia, Cyprus
[9]Department of Clinical Sciences and Translational Medicine, Università di Roma Tor Vergata, Rome, Italy
[10]Institute of Specific Prophylaxis and Tropical Medicine, Medical University of Vienna, Vienna, Austria
[11]VALORIZA - Research Centre for Endogenous Resources Valorisation, Polytechnic Institute of Portalegre, Portalegre, Portugal
[12]Department of Preventive Medicine, Public Health and Microbiology, Faculty of Medicine, Autonomous University of Madrid, Madrid, Spain
[13]Clinical Parasitology Division, Department of Medical Microbiology & Infection Prevention, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands
[14]Department of Nutrition and Dietetics, Hitit University Faculty of Health Sciences, Çorum, Turkey
[15]Mae Fah Luang University, Mueang Chiang Rai, Chiang Rai, Thailand
[16]Department of Medical Microbiology, College of Medicine, Ninevah University, Mosul, Nineveh Governorate, Iraq

v1 **First published:** 17 Oct 2026, **6**:430
https://doi.org/10.12688/openreseurope.24809.1
Latest published: 17 Oct 2026, **6**:430
https://doi.org/10.12688/openreseurope.24809.1

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Abstract

Background

Blastocystis is among the most prevalent gut-associated microbial eukaryotes in humans and animals worldwide, yet reproducible cross-laboratory research remains hampered by the absence of widely accessible, well-annotated reference materials. Subtype (ST) diversity spanning at least 45 recognised STs (ST1–ST17, ST21, ST23–ST49), methodological heterogeneity in detection, culture, and nucleic-acid extraction, and the dispersal of valuable biological materials across individual laboratories represent ongoing obstacles to comparative analyses, assay benchmarking, and inter-laboratory reproducibility.

Methods

We established a physical collection and a linked virtual catalogue at the University of Kent. The collection holds cryopreserved xenic and axenic *Blastocystis* cultures spanning multiple STs and host categories, DNA extracts and, for a subset of accessions, matched RNA extracts, prepared either from culture or directly from faecal material, and stool-derived aliquots from both human and animal donors, at present predominantly of animal origin. Each deposit is assigned a persistent accession identifier linking all associated physical assets, subtype assignments with an explicit confidence level and mixed-colonisation flags, contamination screening outcomes, and quality-control metrics. The virtual catalogue provides a harmonised, versioned metadata framework for discovery and for controlled access under a material transfer agreement, with tiered visibility that accommodates ethical, permit-related and depositor constraints.

Data Available

Physical materials are available on request through the University of Kent Biobank Coordinator, subject to a material transfer agreement. The virtual catalogue and its metadata schema are versioned and citable. Contact details and access procedures are provided in this manuscript.

Reuse Potential

The resource supplies standardised, well-annotated reference materials for *Blastocystis* research. It is intended to support cross-laboratory comparison, assay benchmarking and inter-laboratory ring trials spanning a range of subtypes, hosts and methods.

Keywords
Blastocystis subtypes; inter-laboratory comparison; xenic and axenic cultures; reference material; metadata standardisation; material transfer agreement; One Health

This article is included in the COST Actions gateway.

This article is included in the Blastocystis collection.

Corresponding author: Anastasios D. Tsaousis (A.Tsaousis@kent.ac.uk)

Author roles: Makkimane J: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Validation, Visualization, Writing – Original Draft Preparation; **Figueiredo AM**: Writing – Review & Editing; **Šoba B**: Writing – Review & Editing; **Allievi C**: Writing – Review & Editing; **Denoyelle C**: Writing – Review & Editing; **Carmena D**: Writing – Review & Editing; **Gentekaki E**: Writing – Review & Editing; **Viscogliosi E**: Writing – Review & Editing; **Guadano-Procesi I**: Writing – Review & Editing; **Walochnik J**: Writing – Review & Editing; **Hurtado L**: Writing – Review & Editing; **Lerma L**: Writing – Review & Editing; **José Irisarri-Gutiérrez M**: Writing – Review & Editing; **Zonneveld R**: Writing – Review & Editing; **Sensöz S**: Writing – Review & Editing; **Jinatham V**: Writing – Review & Editing; **Majid Salman Al-Adilee Y**: Writing – Review & Editing; **D. Tsaousis A**: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This project has received funding from the European Union's Framework Programme for Research & Innovation as part of the COST Action CA21105, Blastocystis under One Health, as supported by the COST Association (European Cooperation in Science and Technology).
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2026 Makkimane J *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Makkimane J, Figueiredo AM, Šoba B *et al.* **The Blastocystis Biobank: a curated, multi-host biological resource of cultures, nucleic acids, metabolites, and stool-derived aliquots with a controlled-access virtual catalogue for global research [version 1; peer review: awaiting peer review]** Open Research Europe 2026, **6**:430
https://doi.org/10.12688/openreseurope.24809.1

First published: 17 Oct 2026, **6**:430 https://doi.org/10.12688/openreseurope.24809.1

Introduction

Blastocystis is one of the most frequently detected microbial eukaryotes in gut-associated samples from humans and animals. Reported prevalence varies markedly with the population sampled and with the detection method used, molecular assays consistently recovering higher rates than microscopy.[7,8] A metagenomic survey of 56,989 individuals across 32 countries found carriage among apparently healthy adults to be lowest in North America (6.6%) and considerably higher in Europe (22.1%), whereas PCR-based surveys in some low- and middle-income settings have reported far higher rates, reaching 100% in a paediatric cohort from the Senegal River Basin.[6,24] Despite this ubiquity, the organism remains unusually difficult to study in a reproducible and cross-comparable manner. A central limitation is the lack of widely accessible, well-annotated reference materials. Isolates are typically maintained within individual laboratories, described using non-standardised metadata frameworks, and processed through locally optimised workflows, which makes methodological variability difficult to disentangle from true biological difference.[3]

Blastocystis isolates recovered from mammalian and avian hosts are currently assigned to 45 subtypes (STs) (ST1–ST17, ST21 and ST23–ST49) defined on the small subunit ribosomal RNA (*SSU* rRNA) gene.[1,25] This framework was developed for, and is applied to, isolates of mammalian and avian origin; sequences recovered from other host groups are not currently accommodated within it. The count has risen steeply in recent years. The proximate driver is not simply wider geographic and host sampling but the routine availability of next-generation amplicon sequencing (NGS) and long-read platforms, which deliver the full-length or near full-length *SSU* rRNA sequences needed to validate a candidate subtype to the standard set by current nomenclatural practice.[23,25,26] Of these, 16 STs have been reported from humans (ST1–ST10, ST12, ST14, ST16, ST23, ST35 and ST41), with ST1–ST4 accounting for the majority of human records worldwide.[4,23,29,30] The remaining STs occur predominantly or exclusively in animals, which is one reason a One Health perspective is needed to describe the ecological range of the genus. Subtype identity carries biological information: STs differ in host range, in geographic distribution and in the level of intra-subtype genetic diversity they contain.[10] Reproducible associations between subtype and clinical outcome, by contrast, have not been established, and subtype should not at present be read as a proxy for pathogenic potential.[7,15,16] Subtype confirmation nonetheless requires sequencing, and a substantial fraction of the material reported in the literature carries no validated assignment. In addition, mixed-subtype colonisation is inconsistently reported. Mixed subtypes are resolved reliably only by NGS, which is performed routinely in comparatively few laboratories; Sanger sequencing of a PCR product typically recovers the predominant subtype alone and under-represents minor ones.[26,27] Culture-based studies, for their part, frequently omit passage history and whether a culture is xenic or axenic. That distinction is consequential: xenic culture, in which the organism is maintained together with an accompanying bacterial community, is achievable for many subtypes, whereas axenisation has succeeded for only a limited number of isolates, largely because growth appears to depend on interactions with those associated bacteria and no generally effective axenisation protocol is available.[28] Both omissions materially affect how downstream results can be interpreted.[3,28]

The challenge is compounded in One Health contexts. Animal-associated *Blastocystis* is critical for understanding host range, transmission pathways, and the ecological dimensions of colonisation. However, animal-derived material is typically more heterogeneous, collected under more variable conditions, and subject to additional ethical and permit-related constraints.[9,25] When useful materials do exist, they are rarely shared — partly because no established route for sharing exists, and partly because the wider community is often unaware that the material exists — and are seldom described in sufficient detail for others to assess fitness-for-purpose. This fragmentation limits cross-study comparison delays the development of community standards and prevents the field from exploiting the growing interest in *Blastocystis* as a potentially relevant member of the gut microbiome[11,12].

A physical biobank with a complementary virtual catalogue (defined in Methods) directly addresses these barriers. By accessioning materials under persistent identifiers, recording structured quality-control (QC) indicators, and providing a tiered-access framework that accommodates the legal and ethical complexity of human-derived, animal-derived and legacy materials, it becomes possible to treat these holdings as citable, referenceable resources rather than informal local collections. Human-derived material warrants a note: an established *Blastocystis* culture is itself a non-human organism, and in several jurisdictions such cultures fall outside the human-tissue governance that applies to the originating stool specimen. Requirements vary between countries and depositors remain responsible for confirming those applicable to them. The approach is consistent with wider moves towards research infrastructures aligned with the FAIR (Findable, Accessible, Interoperable, Reusable) principles[13,14] and with the multi-laboratory, multi-host analyses that One Health research requires. The collection is intended to serve three related purposes: to hold source material of human, animal and environmental origin; to make that material usable for One Health research questions; and to allow linkage across those categories within a single analytical framework.

Here, we describe the structure, content, and access model of the *Blastocystis* Biobank, which is physically and virtually based at the University of Kent**.** It is, as far as we are aware, the first international biobank dedicated to *Blastocystis* that couples curated biological materials to a harmonised virtual catalogue (Table 1) under a common metadata and quality-control framework. The physical collection holds material contributed by depositor laboratories worldwide: cryopreserved cultures spanning multiple STs and host categories, DNA extracts and, where the source material allowed, matched RNA extracts (RNA is not available for every accession), and stool-derived aliquots, all accessioned under a common metadata schema and curated against explicit quality-control (QC) indicators. The virtual catalogue provides the corresponding versioned, citable inventory with tiered visibility. It is hosted within BlastoDB (www.blastodb.com), the community *Blastocystis* database developed under the COST Action CA21105[21] "*Blastocystis* under One Health" (OneHealth*Blastocystis*), which already consolidates curated subtype data, genome and transcriptome resources, laboratory protocols, and a global research laboratory directory.[22] Taken together, the two components are intended to make cross-laboratory work in *Blastocystis* biology both reproducible and benchmarkable. This biobank is a primary deliverable of Working Group 2 (WG2) of the COST Action CA21105, which is mandated to establish and maintain both a biobank and a databank of *Blastocystis* STs and microbiome data, and to develop the associated material and data transfer frameworks.[21]

Methods

Physical biobank: scope, assets, and accessioning

The physical biobank is established at the School of Natural Sciences, University of Kent, and comprises four asset classes. Repository practices follow the ISBER recommendations for repositories,[19] within the wider framework of research biobanking practice.[18] Two terms recur throughout. The virtual catalogue is the digital, curated and versioned inventory of biobank holdings: it carries structured metadata on the identity, provenance, availability and access conditions of each deposit, and supports discovery and governance through a controlled interface. It is a description of the collection, not a duplicate of it. Quality-control (QC) indicators are the standardised metadata attributes and QC measurements recorded for each deposit that bear on its identity, integrity and fitness for downstream use; they are enumerated in the quality assurance section below and specified field by field in https://blastodb.com/biobank/.

1. **Frozen *Blastocystis* cultures**: axenic and xenic isolates cryopreserved in multiple vials per accession, spanning confirmed and provisional subtype assignments across human and animal host categories. Culture conditions (medium, passage number, axenic status, and recovery-test outcomes where performed) are documented per deposit.

2. **DNA extracts**: genomic and/or metagenomic DNA, linked either to a specific culture or directly to the source specimen (human, animal or environmental), with quality metrics (concentration, purity, total yield) recorded and the extraction method documented.

3. **RNA extracts**: total RNA or mRNA preparations where available, with RNA integrity metrics (RIN score or equivalent) recorded, given the sensitivity of RNA to degradation, and extraction and storage handling documented.

4. **Stool-derived aliquots**: biospecimens from human and animal donors, at present predominantly of animal origin, preserved under documented conditions (RNAlater, snap-freezing, or ethanol preservation as appropriate), retained for downstream molecular work, microbiome co-profiling, and assay benchmarking.

Each deposit is assigned a persistent biobank accession identifier (format: BBST-[YEAR]-[5-digit sequential number]) that functions as the primary key linking all associated physical assets, metadata, QC indicators and downstream outputs. Individual asset identifiers (Culture_ID, DNA_ID, RNA_ID, Stool_ID) provide item-level resolution within each accession, enabling one-to-many relationships (for example, one culture accession linked to multiple vials and multiple nucleic-acid extracts). Users can therefore request the specific asset they need (a single vial, a DNA extract, or a matched culture and nucleic-acid set) while the chain of custody remains intact and auditable.

Materials are stored under conditions appropriate to the asset type, including long-term cryogenic storage for cultures and stool aliquots, with continuous temperature monitoring and location tracking. Inventory is managed as aliquoted units to minimise freeze-thaw cycles. Inventory records, storage locations, aliquot counts, freeze–thaw events and quality-control (QC) outcomes are held in a controlled, access-restricted inventory spreadsheet, from which the catalogue is populated. Each asset carries an availability status (available, restricted, depleted, or under QC hold) that is reflected in the virtual catalogue (Table 1).

Table 1. Illustrative record structure for the *Blastocystis* Biobank virtual catalogue. The entries below are synthetic examples constructed to show the field structure and the range of permitted values; they do not correspond to specific real accessions or to material held by any named depositor.

Accession ID	Material Type	Subtype	Confidence	Host Species	Country	Year	Storage	DNA	RNA	MTA	Notes
BBST-2026-00001	Culture	ST1	Confirmed	Human	UK	2023	−80 °C	Yes	No	Not required	Axenic culture; passed viability check Feb 2026
BBST-2026-00002	Stool	ST2	Confirmed	Cattle	UK	2022	LN_2	Yes	No	Required - in place	Confirmed by SSU rRNA gene sequencing
BBST-2026-00003	DNA	ST3	Inferred	Dog	France	2024	−80 °C	Yes	Yes	Not required	Matched RNA available
BBST-2026-00004	DNA	ST4	Pending	Pig	Netherlands	2021	−80 °C	Yes	No	Required - pending	Legacy deposit; synthetic example, provenance not attributable to any depositor
BBST-2026-00005	Stool	ST5	Confirmed	Sheep	Australia	2023	LN_2	Yes	No	Required - in place	Zoonotic investigation cohort
BBST-2026-00006	RNA	Mixed	Confirmed	Human	UK	2024	−80 °C	Yes	Yes	Not required	Mixed ST1/ST3; paired with culture accession BBST-2026-00001

Metadata framework and controlled vocabulary

A core design principle is that every deposit be described using a controlled, extensible metadata schema rather than free-text fields.[20] The schema is implemented as the virtual catalogue database (described below and available as an extended data resource) and covers fields relevant to all asset types as well as asset-specific quality and provenance fields. Key metadata categories include host species, geographic origin (country and regional tier), collection date, sample preservation method, *Blastocystis* subtype with an explicit confidence level (see https://blastodb.com/biobank/), detection method, storage condition and location, MTA status, ethics reference, and linkage to any associated sequence accession in public repositories (GenBank/NCBI, ENA/EMBL-EBI, SRA).

All fields with discrete values are governed by controlled vocabulary lists to enable filtering and cross-accession comparison. Where subtype assignments are inferred rather than confirmed by *SSU* rRNA gene sequencing, this is flagged explicitly. Mixed-subtype colonisation and subtype-ambiguous deposits are recorded as such rather than collapsed onto the predominant subtype, so that users receive the assignment the underlying data actually support. Throughout this manuscript and in the catalogue vocabulary we use colonisation as the default term for the presence of *Blastocystis* in a host, and reserve infection for records to which a documented symptomatic clinical context is attached. This reflects the unresolved status of *Blastocystis* pathogenicity and avoids importing a clinical interpretation into a descriptive metadata field.[2,5,17]

In-house subtype confirmation. Where a deposit arrives without a validated assignment, or where the depositor assignment is based on a short fragment, subtyping is repeated in house. *Blastocystis* screening is performed using a direct PCR protocol targeting the *SSU* rRNA gene. The primer pair BhRDr (5′–GAG CTT TTT AAC TGC AAC AACG–3′) and RD5 (5´–ATC TGG TTG ATC CTG CCA GT–3′) are used to amplify a PCR product of ~600 bp.[31] DNA aliquots of *Blastocystis*-positive samples, whose *SSU* rRNA-PCR amplicons produced bands on agarose gels compatible with the presence of *Blastocystis* are sent for Sanger sequencing and confirmed through BLAST.

Quality assurance, curation, and data validation

Every deposit undergoes a curation step before being assigned active status in the catalogue. QC indicators recorded per asset include subtype assignment and confidence level, contamination screening status and quarantine or exclusion notes where applicable, culture-specific fields (axenic/xenic status, passage number where available, recovery testing outcomes), nucleic-acid QC metrics (minimum concentration, A260/280 and A260/230 purity ratios, total yield, RIN score or equivalent for RNA), and the processing and storage context, which bears directly on interpretation; particularly for stool-derived aliquots, where the preservation method constrains downstream RNA and metabolite work.

Biosafety and shipping. Cultures are handled and shipped in accordance with the biological-agent classification applicable at the depositing and receiving institutions. *Blastocystis* is not classified as a hazard group 3 or 4 agent, and material is processed under standard containment level 2 practice at the University of Kent; shipments are packed and labelled as UN3373 Biological Substance, Category B unless the receiving jurisdiction requires otherwise. Recipients are responsible for confirming the classification and import requirements that apply locally.

Recording these indicators lets the catalogue serve two different kinds of request: exploratory work, for which broader inclusion criteria are acceptable, and benchmarking or ring-trial studies, for which only confirmed, high-quality material is appropriate. Each record carries a completeness flag indicating whether the full curation workflow has been completed.

Current holdings. At the time of writing the collection comprises 634 accessions, corresponding to 149 cryopreserved cultures, 534 DNA extracts, 10 RNA extracts and 300 stool-derived aliquots, contributed by 4 depositor laboratories in 4 countries. Subtypes represented are ST1–9, ST13, ST15, ST24b, ST24c, ST25 and ST49 of which 323 accessions carry a sequencing-confirmed assignment and 311 are recorded as mixed. Host categories represented are human, pig, monkey, lemur, cockroach, baboon, orangutan, chicken, rat, quail, pheasant, goat, wild boar, red deer, sheep and cattle.

Virtual catalogue: structure, tiered access, and versioning

The virtual catalogue constitutes the discovery and governance layer of the biobank, hosted within BlastoDB (www.blastodb.com), the community *Blastocystis* database developed under the COST Action CA21105.[21] It serves as a central hub for *Blastocystis* research, providing access to curated ST data, complete and draft genome and transcriptome sequences, standardised laboratory protocols, and a global directory of active research groups.[22] Placing the catalogue within BlastoDB improves discoverability for the *Blastocystis* research community and keeps the collection alongside the other data resources built by the COST Action. It also supplies a curated, version-controlled inventory recording what exists, what is available for distribution, how each item was derived and which restrictions apply. The catalogue is implemented as a relational database (schema in Https://blastodb.com/biobank/) with a structured access interface.

Because certain deposits include materials subject to ethical approvals, permits, depositor agreements, and sensitive provenance fields (particularly for human-derived, animal-derived and legacy materials), the catalogue operates on a tiered visibility model:

Catalogue-level fields: openly visible for discovery and filtering (subtype, host species, country, asset type, availability status, QC completeness flag).

Restricted provenance fields: fine-grained site descriptors, collaborator-sensitive context, farm or site identifiers (released only on justified request).

Material access: physical samples available through a request-and-review process under a material transfer agreement.

The catalogue is maintained with record versioning and update timestamps, so that a study can cite the precise catalogue version it used, and so that physical materials can be linked to subsequent publications, sequence depositions and further QC outcomes. The virtual catalogue database schema is available in Table 2 and in the extended data repository at https://blastodb.com/biobank/.

Table 2. Core metadata fields in the *Blastocystis* Biobank virtual catalogue database schema.

Field Name	Asset Type(s)	Description	Controlled Vocabulary	Required
BioBank_ID	All	Persistent accession identifier (BBST-YYYY-#####)	No	Yes
Material_Type	All	Category of biological material	Culture; Stool Sample; DNA; RNA; Metabolite; Sequence Data	Yes
*Blastocystis*_Subtype	All	Subtype designation based on *SSU* rRNA classification	ST1–ST17; ST21; ST23–ST49; Mixed (two or more STs); Novel – pending designation; Unknown; Pending	Yes
Subtype_Confidence	All	Confidence level of subtype assignment	Confirmed – full-length *SSU* rRNA; Confirmed – partial *SSU* rRNA (Sanger); Confirmed – NGS amplicon; Inferred; Pending	Yes
Host_Species	All	Source host species	Scientific name plus NCBI Taxonomy ID (e.g. Homo sapiens, txid9606); "Environmental – no host" where applicable	Yes
Country_of_Origin	All	Country of sample collection	ISO 3166-1 alpha-3	Yes
Geographic_Region	All	Broad geographic region	UN M49 sub-region (e.g. Northern Europe; Western Asia; Eastern Africa)	Recommended
Collection_Date	All	Date of original sample collection	YYYY-MM-DD	Yes
Detection_Method	All	Method used to confirm *Blastocystis* positivity	Microscopy; Culture-based; Conventional PCR – *SSU* rRNA; qPCR; Sanger sequencing; NGS amplicon sequencing – *SSU* rRNA; Shotgun metagenomics	Yes
Storage_Condition	All	Long-term storage temperature and method	−196 °C (LN_2); −80 °C; −20 °C; RNAlater; Formalin-fixed	Yes

Table 2. *Continued*

Field Name	Asset Type(s)	Description	Controlled Vocabulary	Required
MTA_Status	All	Status of material transfer agreement	Not required; Required – in place; Required – pending; Under negotiation	Yes
Ethics_Ref	All (human/ animal)	Institutional ethics approval reference		Yes (if applicable)
Culture_Type	Culture	Axenic vs xenic culture status	Axenic; Xenic; Mixed culture; Unknown	Yes
Culture_Medium	Culture	Growth medium used	Modified Jones' medium; TSGYM-9; LYSGM; Robinson's medium; IMDM +10% horse serum; RPMI-1640;	Recommended
Passage_Number	Culture	Number of in vitro passages since isolation	Integer	Recommended
Viability_Status	Culture	Last confirmed viability status	Viable – confirmed; Viable – assumed; Non-viable; Unknown	Yes
DNA_Type	DNA	Type of DNA preparation	Genomic DNA; Metagenomic DNA; Whole genome amplified	Yes
Concentration_ng_uL	DNA, RNA	Nucleic acid concentration	Numeric (ng/µL)	Yes
Purity_A260_280	DNA, RNA	Purity ratio (A260/280)	Numeric	Recommended
RIN_Score	RNA	RNA Integrity Number	0–10	Recommended
Sequence_Accession	DNA, RNA, Sequence	Public repository accession number	Free text	Recommended
Clinical_Status	Stool	Clinical phenotype of source host	Asymptomatic colonisation; IBS-D/C/M; IBD; Acute gastroenteritis; Other (specify); Not recorded	Recommended
Sample_Preservation	Stool	Method of stool preservation	RNAlater; Snap-frozen; Ethanol (70%); Fresh; Formalin (accepted for morphology only – not suitable for culture or nucleic-acid work)	Yes
Bristol_Stool_Scale	Stool (human)	Stool consistency at collection	Type 1–7; Unknown	Recommended
Availability_Status	All	Current availability for distribution	Available; Restricted; Depleted; Under QC hold	Yes

Sustainability beyond the COST Action. The biobank is a WG 2 deliverable of COST Action CA21105, but it is not intended to end with it. The physical collection is hosted by the School of Natural Sciences at the University of Kent, which provides the laboratory space, cryogenic storage and monitored freezer capacity; the virtual catalogue is hosted within BlastoDB, which the Kent group maintains. Recurrent running costs (liquid nitrogen, consumables, curation time) are met from institutional resources supplemented by the cost-recovery model described below, and the Action consortium is pursuing longer-term support for a part-time curator. Should hosting at Kent become untenable, the material transfer and depositor agreements provide for transfer of the collection to another consortium laboratory or to an established public culture collection rather than for its disposal, and the catalogue records, being versioned and independently deposited, would survive such a transfer.

Access pathway

Access to catalogue metadata and/or physical materials is provided through a structured request process managed by the University of Kent Biobank Coordinator. Requests are evaluated for scientific rationale, biosafety, and logistical feasibility, and for compliance with any restrictions specific to the requested accessions. Catalogue metadata are provided free of charge. Physical materials are supplied on a not-for-profit, cost-recovery basis: requesters meet shipping, packaging, customs and permit costs, and a nominal handling fee may be levied towards consumables and the upkeep of the physical repository. No charge is made for the biological material itself. Approved releases are documented in the catalogue record, giving an auditable chain of custody. Users are required to cite the biobank accession identifier(s) and the catalogue version used, and to observe any depositor restrictions attached to the accession.

Ethics and permits

All material was collected under the approvals held by the originating laboratory as follows:

- All procedures for the collection of faecal samples from live animals were accomplished following good clinical practices in the respect of animal welfare according to current legislation. The faecal sampling conducted with the approval of the Institutional Animal Care and Use Committee of Università degli Studi di Milano ("Organismo Preposto al Benessere degli Animali," Protocol number OPBA_ 90_2023).

- Sampled animals were legally hunted under Portuguese and EU (RD 173/99) legislation. All the hunters held hunting licenses, and no animals were hunted for the project's sake, but only as part of the annual hunting activities permitted by the abovementioned legislation. Professional personnel collected the faecal samples from hunter-harvested wild boars and red deer during the regular hunting seasons.

- Most animal samples were collected non-invasively (directly from the ground), so they did not require permits.

Depositors are required to confirm at accession that the material was obtained with the relevant institutional ethical approval, informed consent (for human-derived specimens) and animal-use or wildlife-sampling permits, and to supply the corresponding reference, which is recorded in the Ethics_Ref field and released to recipients on request. A clear explanation of the research objectives prior to enrolment was provided to the volunteers, and written informed consent was obtained from each adult included in this study for the collection and use of human-derived specimens. (Consent form provided in the supplementary material).

The University of Kent does not re-consent or re-permit deposited material; onward use by recipients is governed by the material transfer agreement and by the recipient's own institutional approvals.

Ethics and consent

Ethical approval and informed consent were obtained for human-derived and animal-derived materials as follows:

Human-derived specimens. Written informed consent was obtained from each adult donor prior to collection and biobanking of faecal specimens. A clear explanation of research objectives was provided at enrolment. Specific ethics reference numbers are recorded in the catalogue Ethics_Ref field and released to recipients upon request.

Animal-derived specimens. The Institutional Animal Care and Use Committee of Università degli Studi di Milano ("Organismo Preposto al Benessere degli Animali," Protocol OPBA_90_2023) approved all sampling. Faecal sampling of live animals was conducted according to good clinical practices in respect of animal welfare per current legislation.

Wildlife-derived specimens. Faecal samples from hunter-harvested wild boars and red deer were collected in accordance with Portuguese and EU legislation (RD 173/99). All hunters held valid hunting licenses. Sampling was non-invasive, conducted directly from the ground, and required no additional permits.

The University of Kent does not re-consent or re-permit deposited material. Onward use by recipients is governed by the material transfer agreement and the recipient's own institutional approvals. The consent form used for human specimen collection is available in the supplementary material and in the extended data repository at https://blastodb.com/biobank/

Software availability

BlastoDB (www.blastodb.com), the community *Blastocystis* database that hosts the virtual catalogue, is freely accessible online; catalogue browsing is open, while restricted provenance fields and material requests require registration. No new standalone software was developed for this study: the catalogue is implemented as a relational schema and access

interface within the existing BlastoDB platform. The metadata schema itself (https://blastodb.com/biobank/) is a static data resource and requires no specific software. The schema will additionally be deposited in a general-purpose repository under a CC BY licence, with a version-specific DOI, so that it can be cited independently of this article. Data sharing is not applicable to this article as no data were created or analysed in this research.

Data availability statement

The *Blastocystis* Biobank virtual catalogue is publicly accessible via the BlastoDB platform at www.blastodb.com. Physical materials comprising cryopreserved cultures, nucleic acid extracts, and stool-derived aliquots are available from the University of Kent *Blastocystis* Biobank Coordinator (Blastocystis.cost@gmail.com) under controlled access.

Access to physical materials reflects three foundational requirements. First, human-derived specimens were collected under informed consent agreements specifying use for approved research, with ethics approvals that vary by jurisdiction and donor country. Second, animal-derived materials are subject to wildlife protection legislation, institutional animal care protocols, and in certain jurisdictions import/export permit requirements. Third, the biobank operates under material transfer agreements ensuring appropriate use, traceability of distribution, and recognition of contributing laboratories.

Researchers wishing to obtain materials should contact the *Blastocystis* Biobank Coordinator (Blastocystis.cost@gmail.com) providing the following: a description of the proposed research with scientific rationale, confirmation of institutional ethics approval or exemption status, evidence of adequate biosafety facilities (BSL-2 minimum for culture work), and confirmation that the requesting institution can comply with material transfer agreement terms. Applicants then submit supporting documentation including ethics approvals, institutional biosafety certification, and any required import/export permits, followed by signature of the standard material transfer agreement.

Recipients agree to cite biobank accession identifiers and catalogue version in publications, observe any depositor-specific restrictions attached to accessions, refrain from transferring material to third parties without explicit prior approval, and provide feedback to the Biobank Coordinator on material use and research outcomes. The catalogue's core metadata—subtype assignments with confidence level, host species, geographic origin, asset types, availability status, and quality control completeness flags—are openly discoverable at www.blastodb.com. This metadata supports research planning and enables identification of suitable materials before submission of a formal access request.

Restricted provenance fields including specific collection site details, collaborator information, and farm identifiers are released only on justified request, protecting depositor confidentiality. The virtual catalogue database schema is provided in Table 2 under Creative Commons Attribution 4.0 licence and is publicly available for reuse without restriction, provided this article is cited. Supporting materials including the consent form used for human specimen collection are available at https://blastodb.com/biobank/

The biobank is hosted at the School of Natural Sciences, University of Kent, and is integrated with BlastoDB, the COST Action CA21105 community Blastocystis database. BlastoDB automatically tracks versioning with timestamps, enabling researchers to cite the specific catalogue state used in their analyses. Material transfer and depositor agreements specify transfer of the collection to another consortium laboratory should Kent hosting become unavailable, ensuring long-term preservation and continued access. The virtual catalogue records, being versioned and independently managed within BlastoDB, would remain permanently accessible following any such transfer.

Acknowledgments

The authors thank all contributing laboratory partners who have deposited or committed to depositing materials in the biobank. We are grateful to colleagues who provided feedback on the metadata schema and access model during the development of this resource.

References

1. Stensvold CR, Suresh GK, Tan KS, *et al.*: **Terminology for *Blastocystis* subtypes – a consensus.** *Trends Parasitol.* 2007; **23**(3): 93–96.
PubMed Abstract | Publisher Full Text

2. Tan KS: **New insights on classification, identification, and clinical relevance of *Blastocystis* spp.** *Clin. Microbiol. Rev.* 2008; **21**(4): 639–665.
PubMed Abstract | Publisher Full Text | Free Full Text

3. Clark CG, van der Giezen M, Alfellani MA, *et al.*: **Recent developments in *Blastocystis* research.** *Adv. Parasitol.* 2013; **82**: 1–32.
Publisher Full Text

4. Alfellani MA, Stensvold CR, Vidal-Lapiedra A, *et al.*: **Variable geographic distribution of *Blastocystis* subtypes and its potential implications.** *Acta Trop.* 2013; **126**(1): 11–18.
PubMed Abstract | Publisher Full Text

5. Stensvold CR, Clark CG: **Current status of *Blastocystis*: a personal view.** *Parasitol. Int.* 2016; **65**: 763–771.
Publisher Full Text

6. El Safadi D, Gaayeb L, Meloni D, *et al.*: **Children of Senegal River Basin show the highest prevalence of *Blastocystis* sp. ever observed worldwide.** *BMC Infect. Dis.* 2014; **14**: 164.
PubMed Abstract | Publisher Full Text | Free Full Text

7. Beghini F, Pasolli E, Truong DT, *et al.*: **Large-scale comparative metagenomics of *Blastocystis*, a common member of the human gut microbiome.** *ISME J.* 2017; **11**(12): 2848–2863.
PubMed Abstract | Publisher Full Text | Free Full Text

8. Andersen LO, Bonde I, Nielsen HB, *et al.*: **A retrospective metagenomics approach to studying *Blastocystis*.** *FEMS Microbiol. Ecol.* 2015; **91**: 1–9.
Publisher Full Text

9. Cian A, El Safadi D, Osman M, *et al.*: **Molecular epidemiology of *Blastocystis* sp. in various animal groups from two French zoos and evaluation of potential zoonotic risk.** *PLoS One.* 2017; **12**(1): e0169659.
PubMed Abstract | Publisher Full Text | Free Full Text

10. Stensvold CR, Alfellani M, Clark CG: **Levels of genetic diversity vary dramatically between *Blastocystis* subtypes.** *Infect. Genet. Evol.* 2012; **12**(2): 263–273.
PubMed Abstract | Publisher Full Text

11. Nieves-Ramírez ME, Partida-Rodríguez O, Laforest-Lapointe I, *et al.*: **Asymptomatic intestinal colonization with protist *Blastocystis* is strongly associated with distinct microbiome ecological patterns.** *mSystems.* 2018; **3**(3): e00007–e00018.
Publisher Full Text

12. Nourrisson C, Scanzi J, Pereira B, *et al.*: ***Blastocystis* is associated with decrease of fecal microbiota protective bacteria: comparative analysis between patients with irritable bowel syndrome and control subjects.** *PLoS One.* 2014; **9**(11): e111868.
PubMed Abstract | Publisher Full Text | Free Full Text

13. Wilkinson MD, Dumontier M, Aalbersberg IJ, *et al.*: **The FAIR Guiding Principles for scientific data management and stewardship.** *Sci. Data.* 2016; **3**: 160018.
PubMed Abstract | Publisher Full Text | Free Full Text

14. Roche DG, Kruuk LE, Lanfear R, *et al.*: **Public data archiving in ecology and evolution: how well are we doing?.** *PLoS Biol.* 2015; **13**(11): e1002295.
PubMed Abstract | Publisher Full Text | Free Full Text

15. Skotarczak B: **Genetic diversity and pathogenicity of *Blastocystis*.** *Ann. Agric. Environ. Med.* 2018; **25**(3): 411–416.
Publisher Full Text

16. Scanlan PD: ***Blastocystis*: past pitfalls and future perspectives.** *Trends Parasitol.* 2012; **28**(8): 327–334.
Publisher Full Text

17. Andersen LO, Stensvold CR: ***Blastocystis* in health and disease: are we moving from a clinical to a public health perspective?.** *J. Clin. Microbiol.* 2016; **54**(3): 524–528.
PubMed Abstract | Publisher Full Text | Free Full Text

18. Hewitt RE: **Biobanking: the foundation of personalized medicine.** *Curr. Opin. Oncol.* 2011; **23**(1): 112–119.
Publisher Full Text

19. Snapes E: *ISBER Best Practices: Recommendations for Repositories.*5th edVancouver: International Society for Biological and Environmental Repositories; 2023.
Reference Source

20. Yilmaz P, Kottmann R, Field D, *et al.*: **Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIxS) specifications.** *Nat. Biotechnol.* 2011; **29**(5): 415–420.
PubMed Abstract | Publisher Full Text | Free Full Text

21. Tsaousis AD, Gentekaki E, Stensvold CR: **Advancing research on *Blastocystis* through a One Health approach.** *Open Res Eur.* 2024; **4**: 145.
PubMed Abstract | Publisher Full Text | Free Full Text

22. **BlastoDB: The *Blastocystis* Database.** [cited 2026 May].
Reference Source

23. Santin M, Figueiredo A, Molokin A, *et al.*: **Division of *Blastocystis* ST10 into three new subtypes: ST42–ST44.** *J. Eukaryot. Microbiol.* 2024; **71**(1): e12998.
PubMed Abstract | Publisher Full Text

24. Piperni E, Nguyen LH, Manghi P, *et al.*: **Intestinal *Blastocystis* is linked to healthier diets and more favorable cardiometabolic outcomes in 56,989 individuals from 32 countries.** *Cell.* 2024; **187**(17): 4554–4570.e18.
PubMed Abstract | Publisher Full Text

25. Figueiredo AM, Santin M, Maloney JG, *et al.*: ***Blastocystis* in free-ranging wild ruminant species across the Iberian Peninsula.** *Vet. Res.* 2025; **56**: 145.
PubMed Abstract | Publisher Full Text | Free Full Text

26. Maloney JG, Molokin A, Santin M: **Next generation amplicon sequencing improves detection of *Blastocystis* mixed subtype infections.** *Infect. Genet. Evol.* 2019; **73**: 119–125.
PubMed Abstract | Publisher Full Text

27. Rojas-Velázquez L, Maloney JG, Molokin A, *et al.*: **Use of next-generation amplicon sequencing to study *Blastocystis* genetic diversity in a rural human population from Mexico.** *Parasit. Vectors.* 2019; **12**(1): 566.
PubMed Abstract | Publisher Full Text | Free Full Text

28. Shaw D, Denoyelle C, Tan KSW, *et al.*: **Comprehensive tools for culturing *Blastocystis*: a standardized resource for research and diagnostics.** *Curr Protoc.* 2025; **5**(8): e70175.
PubMed Abstract | Publisher Full Text | Free Full Text

29. Akdur-Öztürk E, Al-Adilee YMS, Edwards W, *et al.*: ***Blastocystis* across humans, animals and the environment in rural Türkiye, and relationships with the human intestinal microbiome.** *Front. Microbiol.* 2025; **16**: 1665966.
PubMed Abstract | Publisher Full Text | Free Full Text

30. Maloney JG, Molokin A, Seguí R, *et al.*: **Identification and molecular characterization of four new *Blastocystis* subtypes designated ST35–ST38.** *Microorganisms.* 2023; **11**(1): 46.
PubMed Abstract | Publisher Full Text | Free Full Text

31. Scicluna SM, Tawari B, Clark CG: **DNA barcoding of *Blastocystis*.** *Protist.* 2006; **157**: 77–85.
Publisher Full Text