Reviewed Preprint
v1 • June 4, 2026
Not revised

Reviewed Preprint
v2 • July 10, 2026
Revised by authors

A genetic toolkit for stable transgenesis in the anaerobic gut parasite *Blastocystis* ST7-B

M Rey Toleco[1,‡], Kevin SW Tan[2], Mark van der Giezen[1,3,4,5]

[1]Department of Chemistry, Bioscience, and Environmental Engineering, University of Stavanger, Stavanger, Norway • [2]Laboratory of Molecular and Cellular Parasitology, Healthy Longevity Translational Research Programme and Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore • [3]Research Department, Stavanger University Hospital, Stavanger, Norway • [4]Department of Biosciences, University of Exeter, Exeter, United Kingdom • [5]Natural Resources Institute, University of Greenwich, London, United Kingdom

For correspondence:
mitchellrey.toleco@uis.no
mark.vandergiezen@uis.no

‡ Present address: Evolution of Protist Symbioses Laboratory, Institute of Parasitology, Biology Centre, Czech Academy of Sciences, České Budějovice, Czech Republic

Competing interests: No competing interests declared

Funding: See page 26

Reviewing editor: Wendy S Garrett, Harvard T.H. Chan School of Public Health, United States

eLife Assessment

This study makes a **valuable** contribution by broadening the range of eukaryotic model systems and establishing Blastocystis, the most prevalent microeukaryote in the human gut, tractable for reverse-genetics investigations. The presented imaging data are **convincing** and informative, although confirmation by molecular methods would further strengthen the study. The work should interest readers studying host-microbe interactions in the human gut, as well as those developing new systems for eukaryotic research.

https://doi.org/10.7554/eLife.111816.2.sa3

Abstract

Blastocystis is among the most prevalent microbial eukaryotes in the human gut, yet it has remained largely inaccessible to functional genetics. Here, we report a combinatorial toolkit for *Blastocystis* ST7-B that enables stable transgene maintenance under antibiotic selection and recovery of colony-derived transgenic lines. Guided by a proteomics-informed candidate screen, we identified endogenous promoter–terminator pairs and benchmarked their activity using NanoLuc luciferase (Nluc), defining near-background, weak, moderate, and robust expression tiers. We optimised square-wave electroporation and establish conditions that balance DNA delivery with culture viability, providing a practical operating regime for routine transfection. Using resazurin-based viability assays alongside culture outgrowth validation, we identified puromycin and trimethoprim as the most reliable selectable systems. A three-stage workflow combining liquid enrichment, solid-phase selection, and liquid culture expansion supports recovery of colony-derived transgenic lines that can be cryopreserved and revived with retained growth, antibiotic resistance, and reporter expression. Finally, bicistronic constructs incorporating a codon-optimised P2A peptide supported selection-linked expression of anaerobic-compatible reporters (UnaG, smURFP, and SNAP-tag®). Results showed reporter-dependent performance consistent with constraints such as chromophore availability and substrate permeability. Together, this toolkit makes *Blastocystis* ST7-B markedly more amenable to genetic engineering.

Introduction

Blastocystis is an anaerobic microbial eukaryote that colonises the intestinal lumen of its host under extremely low oxygen tension (Stensvold and van der Giezen, 2018). Genetically diverse and morphologically variable, it colonises a wide range of hosts, including humans (Tan, 2008). It is also one of the most prevalent eukaryotic microorganisms in the human gut and is estimated

to colonise approximately one to two billion people globally (Scanlan and Stensvold, 2013). Yet mechanistic understanding remains limited, particularly with respect to its metabolic physiology and interaction with its host.

At least 44 distinct subtypes (STs) have been identified based on small subunit ribosomal RNA gene sequences, with ST1 to ST9 found in humans (Koehler et al., 2024; Šejnohová et al., 2024). Among these, ST1 to ST4 are the most frequently detected in human microbiomes (Alfellani et al., 2013). Recently, the presence of *Blastocystis* in the gut was reported to correlate with favourable cardiometabolic profiles and lower BMI values, and to be negatively associated with diseases linked to altered gut ecology (Piperni et al., 2024). However, such generalisations obscure important subtype-specific differences that shape its biology. In animal models, *Blastocystis* ST7-B has been associated with shifts in microbiome composition that include reduced abundance of beneficial taxa, whereas *Blastocystis* ST4-WR1 has been linked to the opposite trend (Yason et al., 2019; Deng and Tan, 2022). In epithelial cell culture, several ST7 isolates trigger tight junction disruption, whereas the ST4 isolates tested did not (Wu et al., 2014). These paradoxical subtype-specific effects complicate any anempt to classify *Blastocystis* as strictly commensal or pathogenic (Deng and Tan, 2025).

Blastocystis metabolism is equally unconventional. It encodes both the classical mitochondrial pyruvate dehydrogenase complex and the anaerobic pyruvate:ferredoxin oxidoreductase (PFO), consistent with the potential for dual modes of pyruvate decarboxylation (Stechmann et al., 2008). It encodes an iron-only [FeFe] hydrogenase, typically linked to molecular hydrogen generation, though hydrogen production has not been detected under the conditions tested in *Blastocystis* (Stechmann et al., 2008). Its tricarboxylic acid (TCA) cycle is incomplete, missing key enzymes such as citrate synthase and isocitrate dehydrogenase, and may function in malate dismutation rather than canonical respiration (Stechmann et al., 2008; Gentekaki et al., 2017). Intriguingly, *Blastocystis* expresses alternative oxidase (AOX), a mitochondrial enzyme that uses oxygen as a terminal electron acceptor, although its physiological relevance remains unresolved (Tsaousis et al., 2018). Together, these features reflect the hybrid and evolutionarily distinct nature of its mitochondria.

A further hallmark of *Blastocystis* physiology is mitochondrial glycolysis, where multiple glycolytic enzymes localise to the mitochondria (Abrahamian et al., 2017; Bártulos et al., 2018; Pyrihová et al., 2024). By relocating core carbon flux across the cytosol–mitochondria boundary, this organisation implies rewired redox and ATP economics and makes *Blastocystis* a useful experimental system for testing how metabolism and organelle function co-evolve under anaerobic constraint. A link between mitochondrial glycolysis and serine biosynthesis has been proposed in oomycetes (Abrahamian et al., 2017), but does not apply to *Blastocystis*, which does not encode mitochondrial versions of the key enzymes required for serine biosynthesis. In parallel, comparative genomics supports a trajectory of reductive evolution, including a streamlined genome and coordinated loss of peroxisomes and flagella, consistent with long-term adaptation to an anaerobic gut niche (Gentekaki et al., 2017; Záhonová et al., 2023).

Though *Blastocystis* occupies a unique position at the intersection of metabolic specialisation, reductive evolution, and global gut colonisation, it remains refractory to molecular-level experimental manipulation. Without functional genetics, mechanistic insight across cellular, metabolic, and ecological dimensions remains speculative. A major step toward genetic manipulation in *Blastocystis* ST7-B was achieved with the development of a protocol for transient transfection (Li et al., 2019). This approach employed a species-specific promoter derived from the upstream region of the legumain gene, paired with its native 3' UTR containing a conserved motif downstream of the polyadenylation signal essential for expression (Klimeš et al., 2014). Among the regulatory elements tested, this was the only combination shown to reliably drive transgene expression. Furthermore, the system was constrained by low transfection efficiency, dependence on a single validated promoter–terminator pair, and the absence of effective drug-selectable markers at the time. These limitations prevented sustained expression, precluded clonal propagation, and restricted functional studies that rely on persistent gene activity. Addressing these gaps will require further optimisation, including identification of additional endogenous

regulatory elements, development of robust selection strategies, and establishment of methods for stable transgene maintenance and clonal line generation. Because transfection has been demonstrated in *Blastocystis* ST7-B, it provides a practical entry point for systematic tool development.

To address these gaps, we developed a genetic toolkit for *Blastocystis* ST7-B. It combines an optimised electroporation workflow for DNA delivery, a screened set of endogenous promoters and terminators for expression control, validated antibiotic markers for selection, and reporter readouts based on P2A-linked constructs and anaerobic-compatible fluorescent proteins.

This combinatorial toolkit establishes *Blastocystis* ST7-B as a genetically accessible system for studying the regulatory logic of its unorthodox cellular physiology and probing organellar evolution under anaerobic constraint. By enabling sustained transgene expression under drug selection, reporter-based assays, and colony-derived transgenic line generation, it provides an entry point for mechanistic experiments on colonisation-relevant traits and subtype-linked host interaction readouts. Beyond its application to *Blastocystis*, the toolkit offers a transferable framework for molecular tool development in other non-model microbial eukaryotes. By prioritising endogenous regulatory elements, modular construct architecture, and empirical optimisation, this work offers an actionable blueprint for bringing molecular genetics to lineages that remain inaccessible to molecular investigation.

Methods

Cell culture

Axenic cultures of *Blastocystis* ST7-B were maintained in pre-reduced Iscove's Modified Dulbecco's Medium (IMDM) with stable glutamine in 25 mM HEPES (L0191; Biowest) supplemented with 10% (v/v) heat-inactivated horse serum (Gibco). Cultures were incubated at 37 °C under anaerobic conditions in 2.5 L anaerobic jars with an anaerobic gas pack (Oxoid) for 3 to 4 days, or until a marked increase in cell density was evident, typically accompanied by a colour change in the medium.

Promoter and terminator selection

Candidate promoter and terminator sources were selected using the omics resources available for *Blastocystis* at the inception of this study. Because no genome-wide promoter map, transcription start site dataset, or experimentally defined regulatory annotation was available for *Blastocystis* ST7-B, we used the published abundance-ranked proteomic dataset for *Blastocystis* ST4-WR1 as a practical starting point for candidate discovery (Armengaud et al., 2017). Specifically, abundant *Blastocystis* ST4-WR1 proteins were used to identify their corresponding annotated genes, after which clear homologous genes were sought in the *Blastocystis* ST7-B genome. We reasoned that ST7-B homologs of highly abundant ST4-WR1 proteins would provide a rational starting set of loci whose endogenous promoter and downstream regulatory regions were more likely to support detectable transgene expression. This rationale was supported by the highly skewed *Blastocystis* ST4-WR1 proteome, in which 193 proteins contribute approximately 50% of the detected proteome and the 13 most abundant proteins account for approximately 10% (Armengaud et al., 2017). This strategy was intended to enrich for candidate regulatory elements for toolkit development, rather than to randomly sample predicted genes or systematically survey the full range of promoter strengths in the *Blastocystis* ST7-B genome.

The 1,000 most abundant *Blastocystis* ST4-WR1 proteins were cross-referenced with the accompanying proteogenomic dataset to identify loci with experimentally supported C-terminal peptide evidence, providing stronger support for the annotated gene-end boundaries used in candidate selection (Armengaud et al., 2017). Corresponding homologs in the *Blastocystis* ST7-B genome annotation, based on Denoeud et al. (2011), were identified using NCBI BLAST searches (Altschul et al., 1990; Camacho et al., 2009) and retained if they showed ≥60% query coverage and ≥75% identity. For each retained *Blastocystis* ST7-B homolog, the putative promoter region was defined as the intergenic sequence upstream of the start codon, extending to the annotated 3'

boundary of the neighbouring upstream gene, as displayed in the NCBI Genome Browser. Where intergenic space was limited, candidate promoter regions sometimes extended into the annotated coding sequence of the neighbouring gene.

Alternative promoter lengths were therefore selected according to locus-specific genomic constraints and evaluated individually. Candidate terminator regions were defined operationally as the native 500 bp sequence immediately downstream of the stop codon and were kept constant across promoter-length variants from the same locus to keep the cloning and screening design tractable.

Cloning

All cloning was performed using the in-house protocol (Toleco and van der Giezen, 2025) adapted from the NEBuilder HiFi DNA Assembly kit (NEB). Plasmids were sent to Microsynth AG (Germany) for Sanger sequencing. For large-scale plasmid preparation, maxi- or giga-preps were carried out using commercial kits (Qiagen) or the Miraprep method (Pronobis et al., 2016). Selection marker genes (*Ecdhfr*, *Hsdhfr*, and *pac*) and anaerobic fluorescent reporter genes (*UnaG*, *smURFP*, and *SNAP-tag*®) were codon-optimized for *Blastocystis* ST7-B and synthesized by Invitrogen GeneArt Gene Synthesis Services (Thermo Fisher Scientific; Supplemental Data 1). All oligonucleotides were designed using Benchling (*https://benchling.com*) and ordered from Thermo Fisher Scientific (Supplemental Table 1).

The constructs used in this study were derived from the pXS2-$P_{Legumain}$ vector described by Li et al. (2019), which adapted a heterologous expression-vector backbone for transient plasmid-based expression in *Blastocystis* ST7-B. Here, the same molecular backbone was used as a plasmid scaffold carrying *Blastocystis*-derived regulatory elements and transgenes.

Estimation of IC_{50}

IC_{50} estimation using resazurin-based viability assay was done as described by Mirza et al. (2011). Antibiotics tested were puromycin (Thermo Fisher Scientific), trimethoprim (MedChemExpress), and WR99210 (MedChemExpress); an additional lot of WR99210 was provided by Jacobus Pharmaceutical (NJ, USA) as a generous gift.

Transfection protocol

The basic transfection protocol was adapted from Li et al. (2019) with modifications. *Blastocystis* ST7-B cells, cultured for 2–3 days, were harvested by centrifugation at 1,000 x g for 5 minutes at room temperature. The pellet was washed with 5 mL of pre-reduced, incomplete cytomix buffer (10 mM K_2HPO_4/KH_2PO_4, pH 7.6, 120 mM KCl, 0.15 mM $CaCl_2$, 25 mM HEPES, 2 mM EGTA, and 5 mM $MgCl_2$) and gently vortexed. Cells were pelleted again under the same conditions, and the supernatant was removed by decanting or gentle vacuum suction, leaving approximately 0.5 mL of buffer above the cell pellet. One millilitre of complete cytomix buffer (pre-reduced incomplete cytomix buffer supplemented with 2 mM ATP and 5 mM glutathione, prepared immediately before use) was added, and the pellet was gently resuspended. The resulting approximately 1.5 mL pooled cell suspension was used for total viable cell counting using a hemacytometer.

After counting, the volume corresponding to 5 x 10^7 cells was transferred to each electroporation reaction and combined with 25 µg of plasmid DNA. The total electroporation volume was adjusted to 500 µL with complete cytomix buffer. After electroporation (GenePulse Xcell™, Bio-Rad), the contents were transferred into a 2 mL microcentrifuge tube using a sterile 1.5 mL Pasteur pipene. The electroporation cuvene was rinsed with 1 mL of complete IMDM medium, and the rinse was pooled with the transfected cells. Cells were centrifuged at 1,000 x g for 5 minutes at room temperature. The supernatant was discarded, and 1.8 mL of complete IMDM was added. The pellet was resuspended by gentle pipening or flicking of the tube. Further electroporation details are described in the main text. Cultures were incubated under standard conditions inside an anaerobic box (Oxoid) with an anaerobic gas pack (Oxoid), as described above.

Nluc assay

Promoter–terminator activity was assessed 16–18 h after electroporation using a transient NanoLuc assay adapted from Li et al. (2019), with modifications. This early time point was selected to capture reporter output within the transient-expression window after DNA delivery, before prolonged plasmid loss, differential outgrowth, or culture-level changes could dominate the readout. Because the constructs did not contain a selectable marker, the measured NanoLuc signal reflects early transient reporter output rather than promoter strength independent of DNA uptake, early plasmid retention, or post-transfection recovery. Cells were harvested 16–18 hours post-transfection by centrifugation at 10,000 x g for 2 minutes at 4 °C. The supernatant was discarded, and the pellet was washed with 500 µL sterile PBS. After a second centrifugation under the same conditions, the pellet was resuspended in 350 µL PBS, and 200–300 mg of sterile zirconia beads (Ø 0.1 mm; Biospec) were added. Tubes were placed on ice, and the TissueLyser LT (Qiagen) adapter was flash-frozen in liquid nitrogen for no longer than 2 seconds. Lysis was carried out using the TissueLyser LT at 50 Hz for two cycles of 2 minutes each, with cooling on ice between cycles. The lysate was clarified by centrifugation at 17,000 x g for 10 minutes at 4 °C, and the supernatant was transferred to a fresh 1.5 mL microcentrifuge tube to serve as crude lysate for luciferase assays. For each reaction, 100 µL of crude lysate was mixed with 50 µL Nano-Glo® Luciferase Assay Reagent (Promega) in a luminescence-compatible 96-well plate, protected from direct light, and incubated at room temperature for 10 minutes. Luminescence was measured using a SpectraMax iD5 plate reader (Molecular Devices) following 30 seconds of orbital shaking. All assays were performed in three independent electroporation runs and analysed in technical duplicates; mock-transfected (no plasmid DNA) cells were included as a negative control.

Selection strategy and clonal propagation

The selection strategy was carried out in phases: liquid phase and solid phase selections followed by re-establishment of the liquid culture. In the liquid phase selection, transfected cells were allowed to recover and propagate until the medium turned yellow without drug treatment, usually within 2 days post-transfection. Cells were then pelleted by centrifugation at 1,500 x g for 5 minutes at room temperature. The supernatant was discarded, and the pellet was resuspended in fresh, pre-reduced IMDM supplemented with antibiotic. The antibiotic concentrations and the graded selection strategy are described in detail in the main text.

Once cells had adapted to high-dose antibiotic concentrations, they were subjected to solid-phase selection. Cultures were diluted with pre-reduced PBS or IMDM to a final density of 1,000 cells/mL. A 200 to 300 µL aliquot of the diluted cell suspension was spread onto pre-reduced IMDM plates containing 1% agar and 0.1% sodium thioglycolate using a sterile cell spreader (Tan et al., 1996a; Tan et al., 1996b; Tan et al., 2000). Plates were incubated overnight, agar-side down, to allow cells to senle. The following day, plates were inverted and incubated for 10–14 days until visible colonies appeared. Individual colonies were picked and transferred to 2 mL microcentrifuge tubes (with perforated lids for gas exchange), containing 1.8 mL of fresh, pre-reduced IMDM and cultured until the medium turned yellow. These clonal lines were subjected to the same liquid phase selection regime and ultimately maintained at a high drug dose referred to as the maintenance concentration.

All incubations were carried out at 37 °C under anaerobic conditions as described above. For contamination control, media were supplemented with an antibiotic cocktail (ABC) composed of penicillin G (100 U/mL), streptomycin (100 µg/mL), levofloxacin (25 µg/mL), and polymyxin B (100 U/mL).

Confocal Laser Scanning Microscopy

Sample preparation for confocal laser scanning microscopy (CLSM) was adapted from Pyrihová et al. (2024), with minor modifications. All imaging experiments were carried out on circular No. 1.5 glass coverslips (Marienfeld) pre-coated with poly-D-lysine (Gibco).

For UnaG visualisation, washed transgenic *Blastocystis* ST7-B cells were treated with 50 µM bilirubin (BR; Thermo Fisher Scientific) and incubated anaerobically at 37 °C for 1 h. Cells were washed twice with PBS before mounting. Wild-type (WT) cells processed in parallel served as controls.

For smURFP imaging, prepared transgenic cells were permeabilised with 0.1% saponin (Thermo Fisher Scientific) for 10 min and washed twice with pre-reduced PBS. Permeabilised cells were then incubated with 50 µM biliverdin (BV; Merck) under anaerobic conditions at 37 °C for 2 h, washed twice with PBS, and mounted for imaging. WT cells treated identically were used as controls.

For SNAP-tag® imaging, cells were incubated with SNAP-Cell® TMR-Star (NEB) at 0, 3, and 5 µM concentrations for 30 min at 37 °C under anaerobic conditions. Cells were washed twice with PBS prior to mounting. The same SNAP-tag®–expressing *Blastocystis* ST7-B line incubated without substrate was used as a negative control.

All imaging protocols used the same wash regime to remove excess substrate, and cells were counterstained with Hoechst 33342 (Molecular Probes) to visualize DNA. Cells were mounted in ProLong Gold Antifade Mountant (Molecular Probes), sealed with colourless nail polish, and cured overnight prior to image acquisition.

Confocal image stacks were acquired on a Leica TCS SP8 laser-scanning confocal microscope equipped with a 63x/1.40 NA oil-immersion objective (Leica Microsystems). Z-stacks were collected using system-optimized sampling compliant with the Nyquist–Shannon sampling theorem and LAS X LIGHTNING deconvolution-optimized senings (Leica Microsystems), except during SNAP-tag® imaging due to technical issues with our microscope. Laser power, pinhole diameter and other imaging senings were kept constant within each fluorescent reporter system to allow unbiased visual comparison of fluorescence intensities across conditions. For presentation, a subset of optical sections (10 consecutive slices) was collapsed into a maximum-intensity Z-projection. The linear dynamic range for each signal channel was kept constant across treatments.

Single transmined-light slices were background-corrected using the rolling-ball algorithm (radius 60 px; light background; sliding paraboloid and smoothing enabled) to bener visualize cellular morphology relative to the fluorescence signals. The background-corrected transmined-light images were then overlaid with the corresponding fluorescence channels in the final merged images, and a cell outline was manually traced to approximate cellular boundaries. All image processing was performed in Fiji (ImageJ distribution, version 1.54p).

Statistical analysis

Data distributions and variance panerns were examined, and the assumptions for parametric testing were not met. Consequently, group differences were evaluated with the Kruskal–Wallis rank-sum test. When omnibus results were significant, post-hoc pairwise comparisons were performed using Dunn's test with P-values adjusted for multiple comparisons using the Benjamini–Hochberg method. Unless noted otherwise, tests were two-sided with $\alpha = 0.05$. Dose–response curves were estimated via nonlinear regression of sigmoidal dose–response models. Analyses were conducted in R (version 4.5.1) and GraphPad Prism (10.6.1 (892)).

Results

Endogenous promoter–terminator pairs support graded transgene expression in *Blastocystis* ST7-B

Transgene expression in *Blastocystis* ST7-B has so far relied on a single validated endogenous promoter–terminator pair from the legumain locus (LeguP1; Li et al., 2019). To expand the available regulatory parts, we screened 23 NanoLuc reporter constructs containing putative endogenous promoter–terminator pairs from 11 of 14 candidate loci; three loci could not be cloned after two independent anempts and are indicated in Table 1. Each construct paired a candidate upstream promoter region with the corresponding downstream terminator region from

the same locus, defined here as the native 500 bp sequence immediately downstream of the stop codon. Where multiple promoter lengths were tested for the same locus, the terminator fragment was kept constant (Table 1; Figure 1A).

This design allowed construct-level benchmarking of paired promoter–terminator modules, but did not test promoter strength or terminator activity independently. Because the constructs lacked a selectable marker and were assayed 16–18 h after electroporation, NanoLuc activity should be interpreted as early transient reporter output rather than absolute promoter strength. Reporter output across three independent transfections per construct spanned a broad range, from near-background to values approaching the LeguP1 reference, with some individual replicates exceeding the reference measured in parallel. Robust signals were obtained with SARFGBPpro-600, 60SRPL3pro-1000, and EamApro-1023, while 60SRPL3pro-500, EamApro-1245, AATpro-1000, and SARFGBPpro-1325 showed moderate activity (Figure 2A). These constructs therefore expand the set of endogenous regulatory elements that support detectable to strong transgene expression in *Blastocystis* ST7-B.

Most of the remaining constructs produced weak reporter output relative to LeguP1, and several remained near-background, including AALpro-1000, AALpro-1500, CBMCPpro-378, and CBMCPpro-572. For AAL and CBMCP, increasing promoter length did not improve activity under the conditions tested. Candidate selection had been guided by *Blastocystis* ST4-WR1 protein abundance and the presence of experimentally supported C-terminal peptide, but these features did not reliably predict regulatory performance in *Blastocystis* ST7-B. For example, although AAL is highly abundant in *Blastocystis* ST4-WR1, both AAL promoter fragments drove only near-background expression in *Blastocystis* ST7-B (Figure 2A).

P2A supports selection-linked co-expression in *Blastocystis* ST7-B

To add a bicistronic expression option in *Blastocystis* ST7-B, we inserted a codon-optimised P2A sequence into a LeguP1-driven cassene linking a selectable marker and UnaG reporter (Figure 1B). The P2A peptide is expected to promote ribosomal skipping during translation, allowing two separate polypeptides to be produced from a single open reading frame. UnaG was used because GFP-family fluorophores require oxygen for chromophore maturation and are therefore unsuitable for anaerobic systems. Before drug selection, fluorescence microscopy of bulk-transfected cultures detected a subset of UnaG-positive cells, indicating expression from the bicistronic construct (Figure 3A). After puromycin selection, UnaG-positive cells were retained and enriched in the surviving population. These results are consistent with functional P2A-linked co-expression of both a selectable marker and a reporter protein in *Blastocystis* ST7-B. However, protein-level separation was not directly tested, and the extent of any residual uncleaved fusion product remains unresolved.

Square-wave electroporation improves DNA delivery and culture recovery in *Blastocystis* ST7-B

Applying the time-constant electroporation protocol of Li et al. (2019) at the recommended 370 V frequently caused arcing and loss of viability in *Blastocystis* ST7-B. Under these conditions, cells darkened, senled at the bonom of the tube, failed to recover, and cultures showed no colour change after 3 to 5 days of incubation. Lowering the voltage to 350 V eliminated arcing, but transfection outcomes remained inconsistent across replicates, indicating that reducing voltage alone did not resolve the limited reproducibility. We therefore replaced time-constant operation, which produces exponential voltage decay, with square-wave electroporation, which maintains constant voltage, and optimised the senings empirically using the LeguP1-Nluc reporter construct.

Voltage was first tested across 150 to 400 V using a single 20 ms pulse in 4 mm cuvenes, while cell input, DNA amount, and transfection volume were held constant at 5 x 10^7 cells, 25 µg DNA, and 500 µL, respectively (Figure 2B). At 150 to 200 V, corresponding to 0.375 to 0.50 kV/cm, Nluc signals were low. Reporter output increased sharply at 250 V and peaked at 300 V, corresponding to 0.625 and 0.75 kV/cm, respectively, whereas signal dropped markedly at 400 V even in replicates

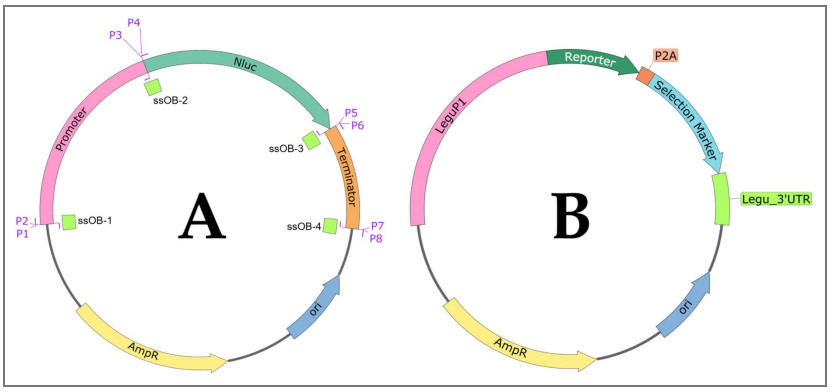

Figure 1.

(**A**) Schematic of the plasmid architecture used to assay promoters and terminators using Nluc as a reporter. Coloured blocks/arrows denote interchangeable modules; P1-P8 (purple) represent the necessary primers to PCR amplify the DNA fragments for assembly and green fragments represent the ssOBs needed to complete the DNA assembly reaction. This configuration highlights the modularity of the approach. Plasmids were cloned and amplified in *E. coli* and then introduced into *Blastocystis* ST7-B, functioning in practice as an *E. coli* ↔ *Blastocystis* ST7-B shuttle vector. Robust expression and drug selection were obtained without adding a *Blastocystis*-specific or canonical eukaryotic origin of replication (e.g., 2μ in yeasts or SV40 ori in mammalian cells), indicating that sequences in the bacterial backbone were sufficient for plasmid maintenance in *Blastocystis* ST7-B. (**B**) A bicistronic construct model using P2A peptide to link fluorescent reporter and a selectable marker under a single promoter-terminator cassette. Reporters used were UnaG, smURFP, and SNAP-tag®; selection markers were *Ecdhfr* (trimethoprim) and *pac* (puromycin).

Table 1. Genes from which endogenous promoters and terminators were cloned to drive transgene expression in *Blastocystis* ST7-B.

Blastocystis **ST4-WR1**			*Blastocystis* **ST7-B**				
Abundance Ranking*	**NCBI protein accession No.**	**Protein Name†**	**ST7-B Gene Homolog**	**NCBI RefSeq**	**% query coverage**	**% Identity**	**Promoter Lengths Tested‡**
6	XP_014525968.1	aspartate ammonia-lyase; AAL	GSBLH_T000 03689001	NW_01317 1849.1	96.00	89.67	AALpro-1000; AALpro-1500
8	XP_014525496.1	small ADP ribosylation factor GTP-binding protein; SARFGBP	GSBLH_T000 04960001	NW_01317 1868.1	100.00	85.39	SARFGBPpro-600; SARFGBPpro-1325
9	XP_014529056.1	2-amino-3-ketobutyrate coenzyme A ligase; AKBCoAL	GSBLH_T000 00411001	NW_01317 1815.1	69.00	82.98	Cloning Unsuccessful
15	XP_014525225.1	60S ribosomal protein L3; 60SRPL3	GSBLH_T000 04213001	NW_01317 1865.1	100.00	83.22	60SRPL3pro-500; 60SRPL3pro-1000
21	XP_014529132.1	aspartate aminotransferase; AAT	GSBLH_T000 01883001	NW_01317 1821.1	89.00	85.14	AATpro-1000; AATpro-1500; AATpro-2438
32	XP_014527004.1	glycine dehydrogenase; GDH	GSBLH_T000 02695001	NW_01317 1827.1	90.00	76.32	GDHpro-1500
87	XP_014525499.1	glyoxysomal and mitochondrial malate dehydrogenases; G/M_MDH	GSBLH_T000 04946001	NW_01317 1868.1	96.00	82.59	G/M_MDHpro-1000; G/M_MDHpro-1500
124	XP_014529593.1	calcium binding mitochondrial carrier protein; CBMCP	GSBLH_T000 03387001	NW_01317 1838.1	100.00	81.97	CBMCPpro-378; CBMCPpro-572
129	XP_014528036.1	40S ribosomal protein S7; 40SRPS7	GSBLH_T000 05110001	NW_01317 1868.1	81.00	79.93	40SRPS7pro-1000; 40SRPS7pro-1500; 40SRPS7pro-2041
348	XP_014526323.1	mitochondrial phosphate transporter; MPT	GSBLH_T000 02675001	NW_01456 9325.1	100.00	80.42	MPTpro-903; MPTpro-1688
535	XP_014529047.1	cathepsin B1; CathB1	GSBLH_T000 05101001	XP_0128995 54.1	96	80.32	Cloning Unsuccessful
		cathepsin B2; CathB2	GSBLH_T000 05101001	XP_0128995 55.1	90	79.84	CathB2pro-500; CathB2pro-1000
595	XP_014525961.1	EamA-like transporter family; EamA	GSBLH_T000 04551001	NW_01317 1866.1	97	75.51	EamApro-1023; EamApro-1245
1000	XP_014526708.1	peptidase C1A family protein; PC1AFP	GSBLH_T000 01253001	NW_01317 1817.1	67	77.31	Cloning Unsuccessful

*Based on the *Blastocystis* ST4-WR1 proteome dataset (Armengaud et al., 2017).

†Protein names are from existing annotations in ST4-WR1 or ST7-B; given abbreviations are arbitrarily assigned for nomenclature convenience.

‡Promoter naming is based on the protein abbreviations used in this work followed by the number of bases upstream of the start codon. For example, AALpro-1000 means a genomic sequence of 1000 bp upstream of the start codon of aspartate ammonia-lyase (AAL) in *Blastocystis* ST7-B. Unsuccessful cloning indicates failure after two independent cloning attempts.

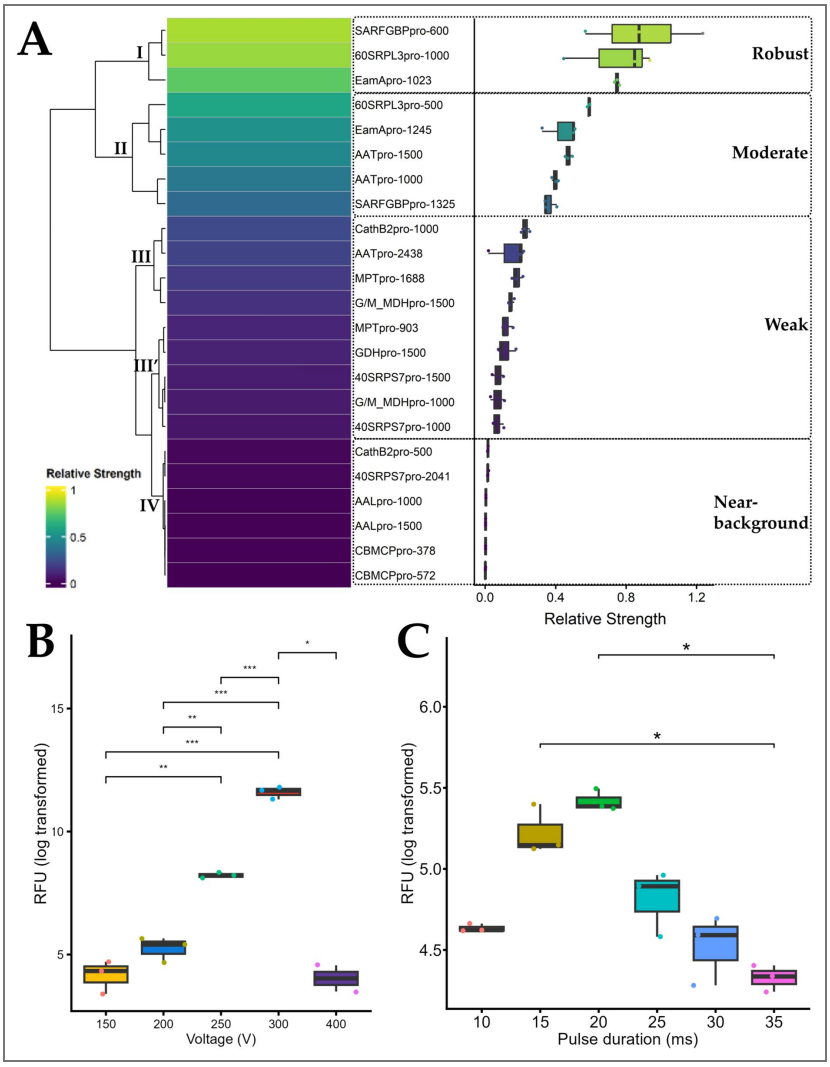

Figure 2.

(**A**) Heatmap with hierarchical clustering (left) and boxplots (right) of promoter–terminator activity in *Blastocystis* ST7-B. Activity was measured by Nluc luminescence. For each construct, raw relative light units (RLU; n = 3 independent replicates) were normalized to the LeguP1 reference construct (external standard) to obtain relative strength. Values are displayed on a 0–1 scale using min–max scaling after normalization; on this scale, 1.0 corresponds to the LeguP1 reference. Promoter fragment lengths are indicated by the numbers in the labels. Unsupervised clustering resolved into tiers: (I) robust, (II) moderate, (III and III') weak, and (IV) near-background. In the adjacent boxplots, points represent values from three independent electroporation runs, each assayed in technical duplicate. (**B**) Effect of square-wave voltage on Nluc luminescence (RLU; n = 3 independent replicates per condition) with pulse duration fixed at 20 ms. (**C**) Effect of single pulse duration on luminescence signal (RLU; n = 3 independent replicates per condition), with voltage fixed at 300 V. For **B** and **C**, electroporation runs were performed with a single pulse while plasmid amount and cell density were held constant. Boxplots show log-transformed raw RLU values (not normalized). Group differences were assessed by Kruskal–Wallis followed by Dunn's post hoc test (α = 0.05); brackets mark tested contrasts and *, **, *** indicate $P < 0.05$, 0.01, 0.001, respectively. Although the 20 ms pulse was not significantly different from 15, 25, or 30 ms, it was chosen for subsequent electroporation runs because it produced the narrowest interquartile range (least dispersion) while maintaining a high signal.

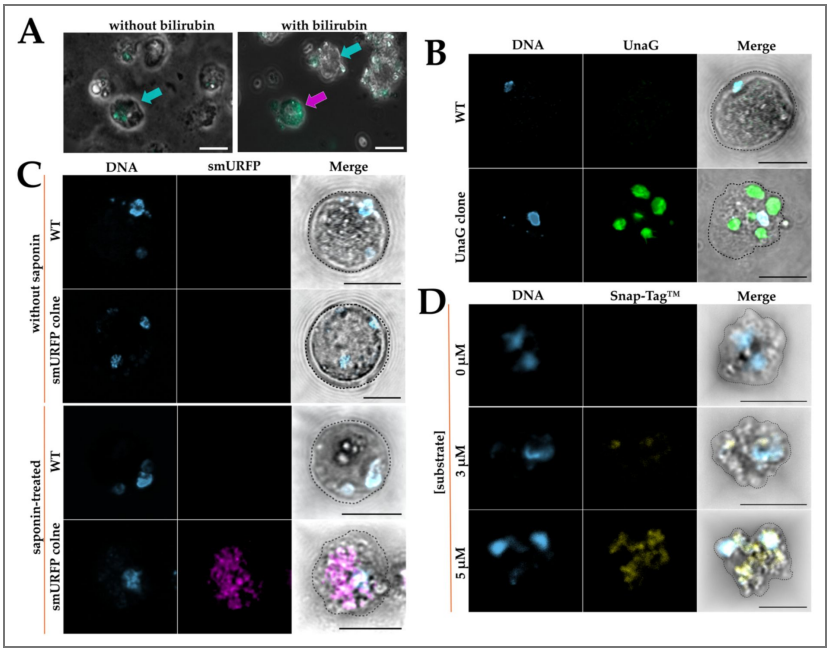

Figure 3.

(**A**) Fluorescence micrographs of *Blastocystis* ST7-B transfected with LeguP1:UnaG::P2A::pac:LeguTer and imaged 3 days post-transfection before antibiotic selection using an Olympus IX81 fluorescence microscope. Images show phase-contrast overlaid with the UnaG channel. In the absence of exogenous BR, only weak signal is detected, consistent with reported intrinsic autofluorescence in *Blastocystis* (yellow arrows; Nagel et al., 2015). After incubation with 1 mM BR, UnaG-positive cells show strong fluorescence (red arrow), whereas neighbouring non-transformed cells remain at background. Scale bar, 10 µm. (B) Representative single-cell confocal images showing compartmentalised UnaG fluorescence in *Blastocystis* ST7-B after incubation with 50 µM unconjugated BR. WT and UnaG-expressing cells were imaged under identical acquisition settings. For display purposes, the lower threshold of the UnaG channel in the transgenic clone was set to the maximum intensity observed in the WT control, so that only signal above the WT autofluorescence range is visible. (C) smURFP fluorescence in *Blastocystis* ST7-B is limited by chromophore access and is unmasked by saponin permeabilisation. Live cells expressing cytosolic smURFP were incubated with 50 µM BV and imaged by confocal microscopy either without permeabilisation or after treatment with 0.1% saponin. No signal was detected above background without saponin. After permeabilisation, smURFP fluorescence became strong and broadly distributed throughout the cytoplasmic compartment, consistent with the untargeted construct design. (D) SNAP-tag® labelling in *Blastocystis* ST7-B is dependent on substrate availability and concentration and yields broadly distributed fluorescence. Live clones expressing an untargeted cytosolic SNAP-tag® were incubated with increasing concentrations of a cell-permeant benzylguanine fluorophore substrate and imaged by confocal microscopy. Signal was at or near background at low substrate but increased progressively with substrate concentration and became broadly distributed throughout the cytoplasmic compartment, consistent with efficient covalent labelling in live cells. DNA was visualised using Hoechst 33342. Most cells contained two nuclei, and smaller Hoechst 33342-positive signals consistent with mitochondrial DNA were also observed in some instances. Because images for the different reporter systems were acquired under different imaging conditions, they are presented to demonstrate reporter detectability and labelling panern and should not be used for quantitative comparison of cell morphology across reporter systems. Scale bars in B, C, and D represent 5 µm.

without visible arcing. Both 250 and 300 V produced strong reporter output without arcing or obvious loss of viability, and cultures showed a colour change within 24 hours, with 300 V giving the strongest performance under matched conditions.

Pulse duration was then tested at fixed voltage by comparing single pulses from 10 to 35 ms at 300 V while holding all other parameters constant (Figure 2C). Reporter output increased from 10 to 20 ms and declined at longer pulse durations, with 20 ms consistently outperforming both shorter and longer pulses across replicates.

Because cuvene geometry, electroporation buffer, DNA amount, cell number, and reaction volume were unchanged across these experiments, the observed differences in reporter output and recovery were most consistent with effects of waveform, voltage, and pulse duration. Square-wave electroporation at 300 V for 20 ms in 4 mm cuvenes was therefore used for subsequent transfections.

Drug sensitivity profiling identifies candidate selection agents for *Blastocystis* ST7-B

To link transgene carriage to cell survival and enable recovery of stable *Blastocystis* ST7-B transformants, we evaluated candidate drugs to establish practical drug–marker pairs for post-transfection selection. Puromycin, trimethoprim, and WR99210 were prioritised because each has a well-established resistance cassene suitable for vector-based expression, namely *pac*, *E. coli* dhfr (*Ecdhfr*), and human DHFR (*Hsdhfr*), respectively (Vara et al., 1988; Asselbergs and Widmer, 1995; Fidock and Wellems, 1997; Remcho et al., 2020). WR99210 was additionally included because a putative DHFR–thymidylate synthase homolog is encoded in the *Blastocystis* ST7-B genome, providing a rationale to test whether this compound inhibits ST7-B growth.

Dose–response assays were performed using a resazurin-based viability readout to quantify drug sensitivity and define working concentration ranges (Figure 4). Cells were exposed to concentrations ranging from 0 to 1.5 μM for each compound, yielding IC_{50} values of 30 ± 6.5 nM for puromycin, 102 ± 24 nM for trimethoprim, and 27 ± 13 nM for WR99210. These values confirm that each compound reduces *Blastocystis* ST7-B viability under the conditions tested. These inhibitory profiles identified puromycin, trimethoprim, and WR99210 as candidate selection agents for further testing, with subsequent experiments supporting puromycin and trimethoprim as the most reliable systems under the conditions used here.

Recovery of colony-derived transgenic *Blastocystis* ST7-B lines through a staged selection regime

To recover *Blastocystis* ST7-B transformants after plasmid transfection, we developed a three-stage workflow comprising stepwise antibiotic selection in liquid culture, isolation of single colonies on solid medium, and re-establishment of individual colonies in liquid culture for expansion and verification. Complete growth inhibition (CGI) was defined as the lowest concentration at which no detectable growth occurred over 24 to 48 h of incubation. Growth and viability were monitored using simple culture phenotypes: healthy cultures shifted the medium towards yellow and formed dense pellets that resisted resuspension after addition of fresh medium, whereas non-viable cultures retained reddish pink medium and produced pellets that dispersed readily. In 1 x 10^7 WT cells, CGI occurred at 20.8x, 245x, and 250x the provisional IC_{50} for puromycin, trimethoprim, and WR99210, respectively (Figure 4). Direct transfer of transfected populations into CGI-level drug concentrations led to loss of viability, so selection pressure was increased gradually over several passages. Post-electroporation recovery and selection were performed in 2 mL round-bonom microcentrifuge tubes to conserve space and maintain high cell density. To support anaerobiosis while limiting contamination, each snap-cap lid was perforated with a single pinhole, and drug concentration was increased stepwise by pelleting cells, removing supernatant, and resuspending in fresh medium at the next concentration (Figure 5). Because axenisation is difficult in

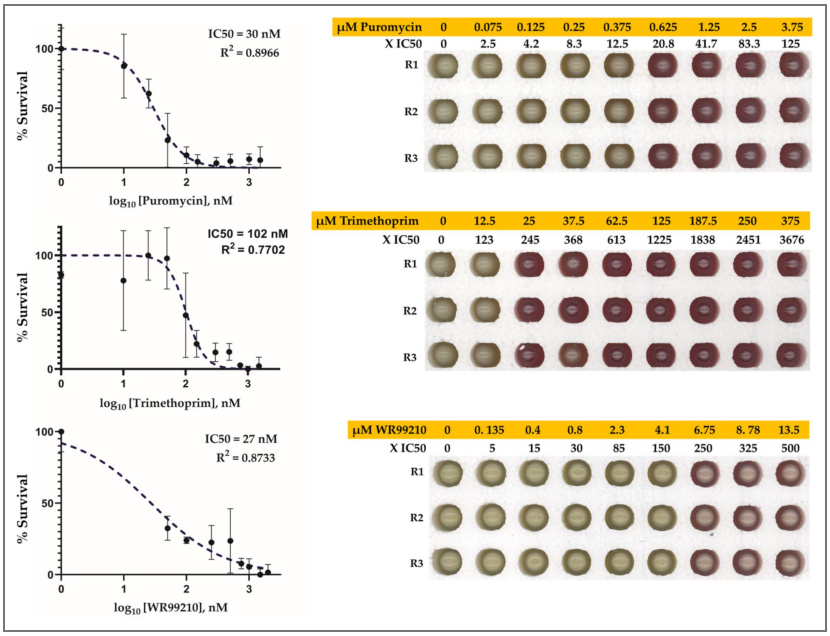

Figure 4. Antibiotic potency and selection-window determination in *Blastocystis* ST7-B.

Dose–response curves for puromycin, trimethoprim, and WR99210 were estimated from a resazurin-based viability assay (n = 3 independent replicates per drug per concentration). Points show mean ± SD, and the insets list the estimated IC_{50} values with R^2-values > 0.75 for all fined curves. The IC_{50} estimates represent the drug concentrations that reduced resazurin-based metabolic activity by 50% under the assay conditions. Right panels: small-culture complete growth inhibition assay using 1 x 10^7 WT *Blastocystis* ST7-B cells per culture, assayed in triplicate across a wide range of concentrations. Cultures were incubated for 2 days, and outgrowth was assessed using phenol red acidification of the medium as a culture-level readout, with yellow indicating growth and red indicating no detectable growth. The yellow-to-red transition was used to estimate the concentration required for complete growth inhibition and to guide the subsequent antibiotic selection strategy for *Blastocystis* ST7-B transformants. IC_{50} and CGI represent distinct assay endpoints: the former measures partial reduction in metabolic activity, whereas the laner identifies the concentration at which no detectable culture outgrowth occurs under the small-culture assay conditions.

Blastocystis (Chen et al., 1997), recovery media were supplemented with penicillin, streptomycin, levofloxacin, and polymyxin B, which suppressed bacterial contaminants without obvious effects on *Blastocystis* ST7-B growth or morphology during routine monitoring.

The solid-phase step yielded colony-derived lines consistent with outgrowth from individual viable units, although not all colonies resumed growth after transfer back into liquid culture, indicating strong density dependence during outgrowth. Recovery improved when plates were inoculated at low density, approximately 200 cells per plate, and incubated for up to 14 days before colony picking. Verification at this stage relied on phenotype and sustained growth under the specified sub-CGI drug concentration, as no growth was observed on plates supplemented at CGI levels, and mock-transformed controls plated in parallel also showed no growth. Across passages, puromycin and trimethoprim consistently discriminated transformed from non-transformed populations. By contrast, WR99210 showed marked source-dependent variability: an initial lot supported inhibition and selection at the expected levels, whereas an independently sourced lot, including freshly prepared stocks, failed to inhibit growth even at concentrations up to 1500ξ the earlier IC_{50} estimate, and low-passage WT controls no longer reproduced the previous response panern. ^{1}H and ^{13}C NMR spectra provided no evidence of WR99210 degradation (data not shown). WR99210 was therefore deprioritised, and puromycin and trimethoprim were used for routine selection in *Blastocystis* ST7-B under these conditions. To reduce repeated sampling of the same lineage during recovery and expansion, three to five electroporation runs were performed per construct, and at least one robust clone from each run was advanced. This workflow enabled recovery of colony-derived transformants that retained drug-resistant growth during extended passaging (>15) and after one year of cryogenic storage.

Anaerobic-compatible fluorescent proteins function in transgenic *Blastocystis* ST7-B lines

To functionally assess transgene expression in *Blastocystis* ST7-B and evaluate the suitability of anaerobic-compatible fluorescent reporters in this organism, three reporters, UnaG, smURFP, and SNAP-tag®, were expressed from bicistronic P2A vectors and examined by confocal microscopy (Kumagai et al., 2013; Rodriguez et al., 2016; Keppler et al., 2003). For each of these proteins, the addition of an external substrate is necessary to form the fluorescent holoenzyme. UnaG, smURFP, and SNAP-tag® require BR, BV, and SNAP-tag® substrates, respectively. Ultimately, each reporter places different demands on chromophore chemistry and substrate permeability. The P2A constructs combined codon-optimized reporters with selectable markers for puromycin or trimethoprim, and stable drug-resistant lines were isolated and expanded for functional characterization of transgene expression.

UnaG, a bilirubin-binding fluorescent protein originally isolated from the muscle of the Japanese eel (Kumagai et al., 2013), produced a bright green signal that rose clearly above WT autofluorescence when background was subtracted under identical senings. Under these conditions, the UnaG signal was readily detectable and visually robust. Efficient holoenzyme formation was observed without permeabilisation of the cells, which indicates that the reporter can become fluorescent *in vivo* (Figure 3B). The UnaG-derived fluorescence did not appear as a uniform cytosolic signal. Instead, it resolved into discrete, bright subcellular foci and larger compartment-like regions, with darker intervening areas that showed linle or no fluorescence. The *UnaG* coding sequence used in these constructs lacked any intentional targeting peptide or signal sequence, so the underlying plasmid design was expected to drive broadly ectopic expression throughout the cell.

smURFP (small ultra-red fluorescent protein) expression produced a clean far-red signal in *Blastocystis* ST7-B, but only when cells were permeabilised with 0.1% saponin prior to BV treatment (Figure 3C). Under identical laser power and detector senings, non-permeabilised smURFP-expressing cells remained at background levels that were indistinguishable from WT controls. After saponin treatment, smURFP fluorescence rose clearly above autofluorescence and was readily detectable, confirming that the reporter is expressed and can form a functional holoenzyme in *Blastocystis* ST7-B once the chromophore gains access to the cytoplasm. The

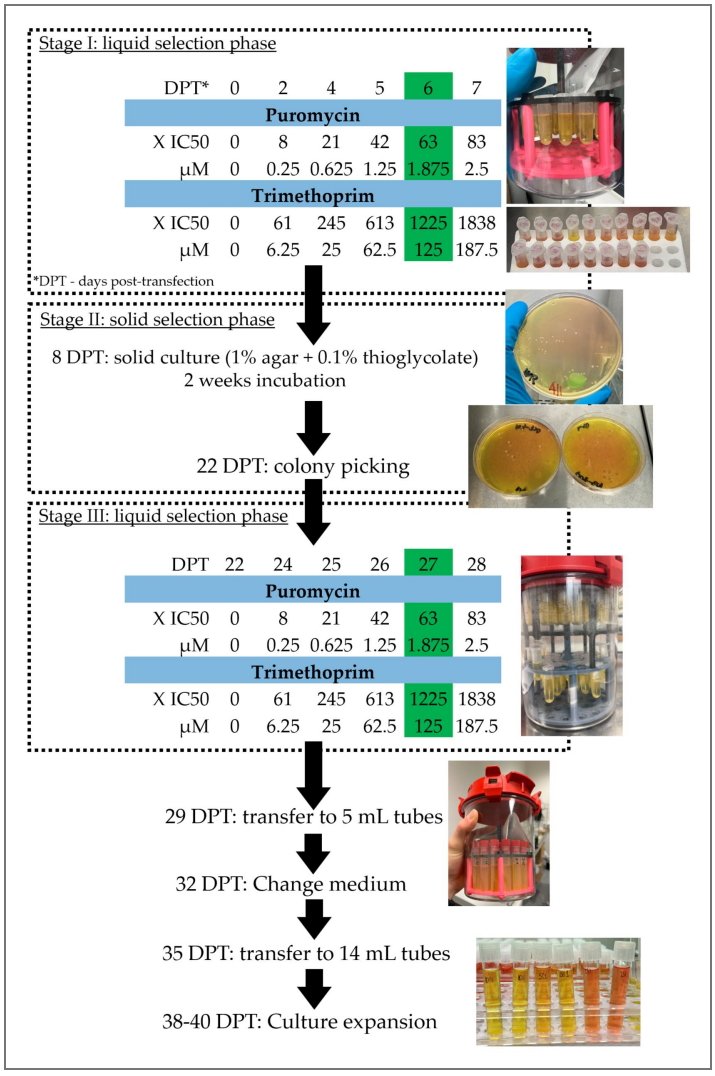

Figure 5. Stepwise selection and clonal propagation of *Blastocystis* ST7-B transformants with puromycin and trimethoprim monitored over days post-transfection (DPT).

Stage I: transfected cells are kept in 2 mL tubes inside anaerobic jars and exposed to increasing drug concentrations; values highlighted in green refer to the maintenance concentration. **Stage II**: enriched cultures are plated on solid medium incubated for 2 weeks, yielding well-isolated colonies that are picked at 22 DPT. At this stage, the drug concentrations are lowered to ensure that individual plated cells, which later form colonies, remain within their resistance capacity during early outgrowth on the plate. **Stage III**: colonies are re-established in liquid culture under the same stepwise antibiotic regime. Subsequently, volume expansion from 2 mL microcentrifuge tubes to 14 mL culture tubes proceeds through an intermediate 5 mL step to increase the success of culture establishment. The entire workflow takes up to 40 days.

smURFP signal appeared specific and free of obvious bleed-through into other imaging channels. Within permeabilised cells, the far-red signal was relatively global across the cytoplasm, without strong restriction to discrete subcellular compartments, consistent with the expected distribution of a broadly expressed cytosolic reporter. This panern matches the construct design, which did not include any intentional targeting peptide or signal sequence and was therefore anticipated to drive ectopic, global expression.

SNAP-tag® labelling was compatible with *Blastocystis* ST7-B and produced robust fluorescence. In cells expressing SNAP-tag®, addition of a cell-permeant SNAP substrate yielded a clear fluorescent signal across a range of substrate concentrations (Figure 3D). Signal intensity scaled with substrate dose over the tested range, indicating that labelling efficiency is substrate-limited under standard conditions. Fluorescence appeared broadly distributed throughout the cytoplasm rather than restricted to discrete organelles, which agrees with the construct design that lacks any explicit targeting sequence and is expected to drive ectopic, global expression. Efficient labelling was obtained without semi-permeabilisation, indicating that the substrate permeated the cell.

Discussion

This study establishes an integrated genetic toolkit for *Blastocystis* ST7-B that links construct design, DNA delivery, antibiotic selection, clonal recovery, and reporter validation into a single experimental pipeline. Its main advance lies not in any one component alone, but in showing that these elements function together to enable stable transgenesis under antibiotic selection and downstream cell biological analysis in this organism. The toolkit was intentionally designed as a combinatorial system of interchangeable genetic parts, allowing constructs to be tailored to the desired circuitry through promoter choice, P2A-based bicistronic design, and selection markers or oxygen-independent reporters. The previously reported ssOB cloning approach (Toleco and van der Giezen, 2025) further supports this design by enabling straightforward assembly and exchange of parts, giving the system a quasi-modular cloning property.

A key outcome of this study is the expansion of endogenous regulatory elements beyond the legumain module previously used for transgene expression in *Blastocystis* ST7-B (Li et al., 2019). This maners because a useful toolkit requires more than a single high-output cassene. Strong expression may be advantageous for selectable markers and bright reporters, but not for every cargo. Hence, weak or near-background promoter-terminator pairs may be useful for expressing toxic proteins or tightly regulated genes that would otherwise harm the host at high expression levels. The promoter-terminator combinations identified here therefore provide a practical expression range for future reporter design and multi-gene constructs. At the same time, the screen highlights an important constraint on part discovery across *Blastocystis* subtypes. Candidate nomination based on *Blastocystis* ST4-WR1 protein abundance and the presence of experimentally supported C-terminal peptide provided a reasonable starting point, but these features were not sufficient to predict activity in *Blastocystis* ST7-B, as illustrated by the poor performance of the AAL-derived fragments. Cross-subtype information can therefore guide candidate selection but cannot substitute for direct validation in the subtype being engineered. The promoter length series further indicates that regulatory output in *Blastocystis* ST7-B is strongly context-dependent. Longer fragments did not consistently improve activity, arguing against simple length-based design rules and instead pointing to a more compact promoter organisation in which relatively small sequence differences can alter initiation efficiency or introduce inhibitory context (Li et al., 2019). Although the present constructs were not designed to define the regulatory architecture of these promoter regions directly, they point to a clear next step in dissecting this logic, including transcription start site mapping and refinement of promoter boundaries. A similar principle likely applies to 3' regions, where locus-specific cleavage and yet unknown features may also contribute to the observed differences in expression.

A second important contribution of this study is the refinement of DNA delivery and recovery conditions for *Blastocystis* ST7-B. Electroporation in *Blastocystis* ST7-B appears to operate within a narrow window in which efficient DNA delivery must be balanced against cell survival and post-transfection recovery. Under the previously recommended time-constant senings at 370 V, arcing

was frequent and cultures often failed during recovery. By contrast, the optimised square-wave senings avoided arcing and preserved enough viable cells to support post-transfection outgrowth. This distinction maners because successful transfection in *Blastocystis* ST7-B is not defined solely by DNA entry, but also by whether a sufficient number of cells survive to re-establish growth, a practical constraint consistent with the long-recognised but still poorly understood density-dependent growth behaviour of *Blastocystis* in liquid culture (Ho et al., 1993). Selection and clonal recovery then complete the transition from transient uptake to stable antibiotic-selected line generation, and the present data indicate that this recovery phase is a major bonleneck. Immediate transfer into fully inhibitory drug concentrations caused culture collapse, whereas staged enrichment allowed resistant populations to emerge, indicating that stable antibiotic-selected line generation depends as much on managing post-transfection stress as on drug potency itself. Among the selectable systems tested, puromycin and trimethoprim proved the most reliable in practice. WR99210 was initially anractive because of its established use in other protists and the presence of a putative DHFR-TS homologue in the *Blastocystis* ST7-B genome (Fidock and Wellems, 1997; Remcho et al., 2020), but its inconsistent behaviour across sources argues against routine use in *Blastocystis*. This result is informative because selectable markers that perform well in one protist cannot be assumed to transfer cleanly to another, and in toolkit development practical robustness maners more than theoretical compatibility. This vulnerable recovery phase also places a premium on careful culture handling. Bacterial contamination can readily compromise recovering *Blastocystis* cultures, making contamination control a practical part of workflow feasibility rather than a minor technical detail.

The P2A bicistronic design adds useful flexibility to the system by enabling two coding sequences to be expressed from a single transcriptional unit, for example a selectable marker paired with a reporter. In a platform where fully validated regulatory parts remain limited, P2A offers a practical way to simplify construct architecture while expanding design options. P2A was selected because it is a well-characterised peptide with high reported separation efficiency in human cell lines, zebrafish embryos, and mice (Kim et al., 2011). It also has precedent across microbial eukaryotes, including the protist *Dictyostelium discoideum* (Zhu et al., 2023), the fungi *Aspergillus niger* (Schueåe and Meyer, 2017) and *Ustilago maydis* (Müntjes et al., 2020), and the apicomplexan parasites *Toxoplasma gondii* (Markus et al., 2019) and *Plasmodium falciparum* (Dans et al., 2024). However, P2A performance is context-dependent, and the evidence presented here is functional rather than biochemical. P2A-containing constructs support antibiotic-selected recovery and downstream reporter expression in *Blastocystis* ST7-B, but ribosomal skipping efficiency and any residual uncleaved product will require direct protein-level validation.

The reporter comparisons further show that fluorescent output in *Blastocystis* ST7-B is constrained by more than transgene expression alone, with each system exposing a different practical limitation. UnaG produced a clear signal in intact cells without oxygen-dependent chromophore maturation, which is a major advantage in an anaerobic protist (Kumagai et al., 2013; Kwon et al., 2020). However, UnaG fluorescence was compartmentalised despite the absence of an obvious targeting sequence, suggesting that signal distribution may be influenced by unconjugated BR availability or intracellular partitioning rather than reporter localization alone. This is plausible given the hydrophobic behaviour of unconjugated BR and its tendency to associate with lipid environments (Ostrow and Celic, 1984; Zucker et al., 1999; Bracken et al., 2024). Consistent with this, lipid-rich peripheral and intracellular structures have been reported in *Blastocystis* ST7-B, potentially providing favourable microenvironments for BR partitioning and contributing to the punctate UnaG fluorescence panern (Liao et al., 2023). An alternative possibility is that the observed signal panern reflects reporter sequestration or reporter-associated cellular stress. However, because UnaG-expressing lines were recovered, maintained under selection, and propagated through continued culture, toxicity remains a possible but untested explanation rather than a demonstrated mechanism. UnaG is therefore functional in *Blastocystis* ST7-B but appears less suitable for uniform whole-cell labelling or straightforward quantitative imaging across the cell. smURFP exposed a different constraint. Fluorescence was detectable only after saponin treatment, suggesting that limited chromophore access is likely the main barrier under the conditions tested (Rodriguez et al., 2016). This is consistent with the limited

membrane permeability expected for BV and with the barrier properties anributed to the *Blastocystis* cell surface (Lightner et al., 1996; Bulmer et al., 2008; Yoshikawa and Hayakawa, 1996; Yason and Tan, 2018; Zheng and Gallot, 2020). In its current form, smURFP therefore appears bener suited to permeabilised preparations than to straightforward live-cell imaging, although membrane-permeant bilins and newer monomeric variants may broaden its future utility (Hu et al., 2024; Maiti et al., 2023; Rodriguez et al., 2016). Among the reporters tested here, SNAP-tag® was the most adaptable in practice. Labelling worked in intact cells, signal scaled with substrate concentration, and the available chemistry provides access to a broad range of fluorophores and imaging modalities (Keppler et al., 2003; Gautier et al., 2008; Birke et al., 2022; Saimi and Chen, 2023; Porzberg et al., 2025). Although substrate-dependent labelling increases consumable cost, SNAP-tag® currently appears to offer the greatest practical versatility among the reporters evaluated here.

There are clear limitations to the present framework. The molecular maintenance state of the introduced constructs remains unresolved: episomal maintenance is a plausible working model, but genomic integration cannot be formally excluded. The constructs used here lack *Blastocystis* ST7-B homology arms, and comparative genomic analyses suggest that *Blastocystis* lacks canonical non-homologous end-joining components (Gentekaki et al., 2017), making targeted integration by standard repair routes unlikely but not impossible. Direct assays such as outward-facing PCR, plasmid rescue followed by full plasmid sequencing, FISH, or selection-withdrawal experiments will be required to distinguish episomal persistence from integration. Copy number, segregation, and long-term expression stability remain unresolved, and these variables are likely to contribute to clone-to-clone heterogeneity, particularly where more quantitative readouts are required. The toolkit also remains limited to transgene expression and reporter validation, with targeted integration, conditional control, and direct loss-of-function approaches still to be established.

Conclusions

Taken together, this work establishes a practical framework for stable transgenesis under antibiotic selection in *Blastocystis* ST7-B, substantially expanding the genetic tools available for this organism. By bringing endogenous regulatory part discovery, namely the identification of native promoter and terminator elements, DNA delivery, selection, clonal recovery, and reporter validation into a single pipeline, it creates a flexible foundation for routine transgene-based studies in this organism. In doing so, it shows that *Blastocystis* ST7-B can be genetically modified and that stable, colony-derived transgenic lines can be generated, opening the way to more direct molecular analysis of its cell biology. More broadly, the strategy taken here provides a practical blueprint for developing genetic tools in other less understood microbial eukaryotes where genetic access remains limited.

Data availability

Proteomic dataset used in this study was generated and published by Armengaud et al., 2017.

Supplemental data and table

Supplemental Data 1. Gene sequences of reporter genes, selection markers, Kozak, and P2A linker

>JQ437370.1 NanoLuc luciferase ATGGTCTTCACACTCGAAGATTTCGTTGGGGACTGGCGACAGACAGCC GGCTAC AACCTGGACCAAGTCCTTGAACAGGGAGGTGTGTCCAGTTTGTTTCAGAATCTC GGGGTGTCCGT AACTCCGATCCAAAGGATTGTCCTGAGCGGTGAAAATGGGCTG AAGATCGACATCCATGTCATCATCCCGTA TGAAGGTCTGAGCGGCGACCAAATG GGCCAGATCGAAAAAATTTTTAAGGTGGTGTACCCTGTGGATGATC ATCACTTTA AGGTGATCCTGCACTATGGCACACTGGTAATCGACGGGGTTACGCCGAACATGA TCGACTATT TCGGACGGCCGTATGAAGGCATCGCCGTGTTCGACGGCAAAAAGA TCACTGTAACAGGGACCCTGTGGAA CGGCAACAAAATTATCGACGAGCGCCTG ATCAACCCCGACGGCTCCCTGCTGTTCCGAGTAACCATCAACGG AGTGACCGGC TGGCGGCTGTGCGAACGCATTCTGGCGTAA

>smURFP *Blastocystis* ST7-B codon-optimized ATGGCTAAGACGTCGGAGCAGCGCGTGAACATTGCGAC GCTGCTGACGGAAAA CAAGAAAAAGATCGTGGATAAGGCGTCGCAGGATCTGTGGCGACGCCATCCGG AT CTGATTGCGCCTGGCGGAATCGCGTTCTCGCAGCGCGATCGAGCGCTGTGCCT GCGCGATTACGGATGGTT CCTGCACCTGATCACGTTCTGCCTGCTGGCTGGCGAT AAGGGACCGATCGAGTCGATTGGACTGATCTCGAT CCGCGAGATGTATAACTCG CTGGGAGTGCCGGTGCCTGCTATGATGGAATCGATCCGCTGCCTGAAAGAGG CG TCGCTGTCGCTGCTGGATGAGGAAGATGCGAACGAGACTGCGCCTTACTTCGAT TACATCATCAAGGCT ATGTCGTAA

>SNAP-tag *Blastocystis* ST7-B codon-optimized ATGGATAAGGATTGCGAGATGAAGCGCACGACGCTGG ATTCGCCTCTGGGAAAG CTGGAACTGTCGGGATGCGAGCAGGGACTGCACCGCATCATCTTTCTCGGAAAG GGAACGTCGGCTGCGGATGCGGTGGAAGTGCCTGCGCCTGCTGCTGTGCTCGGA GGACCTGAGCCTCTGAT GCAGGCGACGGCGTGGCTGAACGCGTACTTCCATCAG CCTGAGGCGATCGAGGAATTCCCGGTGCCTGCTC TGCACCATCCGGTGTTCCAG CAAGAGTCGTTCACGCGACAGGTGCTGTGGAAGCTGCTGAAGGTGGTGAAG TT CGGAGAGGTGATCTCGTACTCGCACCTGGCTGCGCTGGCTGGAAACCCTGCGGC TACGGCTGCGGTGAA AACGGCTCTGTCGGGAAACCCGGTGCCGATCCTGATTCC GTGCCACCGCGTGGTGCAGGGCGATCTGGATG TTGGAGGATACGAAGGCGGACT GGCGGTGAAAGAGTGGCTGCTGGCGCACGAGGGACACCGCCTCGGAA AGCCTG GACTGGGATAA

>UnaG *Blastocystis* ST7-B codon-optimized ATGGTCGAGAAGTTCGTTGGAACGTGGAAGATCGCGGATT CGCACAACTTCGGA GAGTACCTGAAGGCGATCGGAGCGCCTAAAGAACTGTCTGATGGCGGAGATGCT AC GACGCCGACGCTGTACATTTCGCAAAAGGATGGCGATAAGATGACCGTGAA GATCGAGAACGGACCGCCG ACCTTCCTGGATACGCAAGTGAAGTTCAAGCTGG GAGAAGAGTTCGATGAGTTCCCGTCGGATCGCCGAAA GGGCGTGAAGTCGGTG GTTAACCTCGTGGGAGAGAAGCTGGTGTACGTGCAGAAGTGGGACGGAAAAGA AACGACCTACGTGCGCGAGATCAAGGATGGAAAGCTGGTGGTTACGCTGACGAT GGGAGATGTGGTGGCT GTGCGATCGTATCGCCGAGCGACTGAATAA

>*E. coli* dhfr *Blastocystis* ST7-B codon-optimized ATGGAAAACGCTATGCCGTGGAACCTGCCTGCGGAT CTGGCGTGGTTCAAGCGC AACACGCTGAACAAGCCGGTGATCATGGGACGCCACACGTGGGAGTCGATCG G ACGCCCTCTGCCTGGACGCAAGAACATCATCCTGTCGTCGCAGCCTGGAACGGA TGACCGCGTGACGTGG GTGAAGTCGGTGGATGAGGCGATTGCGGCTTGCGGAG ATGTGCCCGAGATCATGGTGATCGGAGGCGGAC GCGTGTACGAGCAGTTCCTGC CGAAGGCGCAGAAGCTGTACCTGACGCACATCGATGCGGAGGTGGAAGG CGAT ACGCACTTCCCGGATTACGAGCCGGATGATTGGGAGAGCGTGTTCTCGGAGTTC CACGATGCGGATG CGCAGAACTCGCACTCGTACTGCTTCGAGATCCTGGAACGC CGATAA

>human DHFR *Blastocystis* ST7-B codon-optimized ATGCACGGATCGCTGAACTGCATCGTGGCGGTGTC GCAGAACATGGGAATCGGA AAGAACGGCGATTACCCGTGGCCTCCGCTGCGCAACGAGTTCCGCTACTTCC AG CGCATGACGACGACGTCGAGCGTGGAAGGCAAGCAGAACCTGGTGATCATGGG AAAGAAAACGTGGT TCTCGATCCCCGAGAAGAATCGCCCTCTGAAGGGACGCA TCAACCTGGTGCTGTCGCGAGAGCTGAAAGAA CCGCCTCAGGGTGCGCACTTCC TGTCGCGCTCGCTGGATGATGCGCTGAAGCTGACGGAACAGCCCGAGCT GGCG AACAAGGTGGATATGGTGTGGATCGTTGGAGGATCGTCGGTGTACAAAGAAGCT ATGAACCATCCG GGACACCTGAAGCTGTTCGTGACGCGCATCATGCAGGATTTC GAGTCGGATACGTTCTTCCCCGAGATCGAC CTGGAAAAGTACAAGCTGCTGCCT GAGTACCCTGGCGTGCTGTCGGATGTGCAAGAGGAAAAGGGAATCA AGTACAA GTTCGAGGTGTACGAGAAGAACGACTAA

>pac *Blastocystis* ST7-B codon-optimized ATGACCGAGTACAAGCCTACTGTGCGACTGGCGACGCGAGAT GATGTTCCTAGA GCTGTGCGAACGCTGGCGGCTGCGTTTGCTGATTATCCTGCTACGCGCCACACGG TGGA TCCTGATCGCCATATTGAACGCGTGACGGAACTGCAAGAGCTGTTCCTGA

CGCGCGTGGGACTCGATATCGGAAAAGTGTGGGTTGCCGATGATGGCGCTGCTG TGGCTGTGTGGACTACG CCTGAATCTGTGGAAGCGGGTGCTGTGTTCGCGGAGA TTGGACCTAGAATGGCGGAACTGTCTGGATCTAG ACTGGCGGCGCAGCAGCAGA TGGAAGGACTTTTGGCGCCGCACCGACCTAAAGAACCTGCTTGGTTTCTG GCGA CCGTGGGAGTGTCTCCGGATCACCAAGGCAAAGGATTGGGATCTGCTGTGGTGC TGCCTGGCGTTG AAGCTGCTGAACGAGCTGGTGTTCCGGCTTTCTTGGAAACGTC GGCGCCTCGCAACCTGCCTTTCTATGAAC GCCTGGGATTCACCGTGACGGCGGA TGTGGAAGTTCCTGAAGGACCTCGAACGTGGTGCATGACGCGCAA GCCTGGTGC TTAA

>P2A sequence *Blastocystis* ST7-B codon-optimized GGCAGCGGAGCGACGAACTTCTCGCTGCTGAAGC AGGCTGGCGACGTGGAAGA GAACCCTGGACCG

Primer/ssOB	Sequence (5' to 3')
pMRT_AATpro-1000	
MRT-75	TCATGGATCGAATTGCTTCCATC
MRT-76	TGTATAGAATCAATAGAAATCTATATGAAACAGGAG
MRT-77	GCCATTTGAATTCATGGTTGTTTCAATTAC
MRT-78	TATTTGGACTCGCGCGCG
ssOB_18	TTTCATATAGATTTCTATTGATTCTATACAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_20	CACGACGTTGTAAAACGACGGCCAGTGCCATCATGGATCGAATTGCTTCCATCAGACGTA
ssOB_21	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGCCATTTGAATTCATGGTTGTTTCAATTAC
ssOB_22	CAACACGTCCGTCGCGCGCGAGTCCAAATAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_AATpro-1500	
MRT-74	GGGTTGACATAGGGGATTCTTGTC
MRT-76	TGTATAGAATCAATAGAAATCTATATGAAACAGGAG
MRT-77	GCCATTTGAATTCATGGTTGTTTCAATTAC
MRT-78	TATTTGGACTCGCGCGCG
ssOB_18	TTTCATATAGATTTCTATTGATTCTATACAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_19	CACGACGTTGTAAAACGACGGCCAGTGCCAGGGTTGACATAGGGGATTCTTGTCTGGAAG
ssOB_21	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGCCATTTGAATTCATGGTTGTTTCAATTAC
ssOB_22	CAACACGTCCGTCGCGCGCGAGTCCAAATAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_AATpro-2438	
MRT-73	TAAGCGAATAAAATGATATGTATTAATACCTACCATAG
MRT-76	TGTATAGAATCAATAGAAATCTATATGAAACAGGAG
MRT-77	GCCATTTGAATTCATGGTTGTTTCAATTAC
MRT-78	TATTTGGACTCGCGCGCG
ssOB_17	CACGACGTTGTAAAACGACGGCCAGTGCCATAAGCGAATAAAATGATATGTATTAATACC
ssOB_18	TTTCATATAGATTTCTATTGATTCTATACAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_21	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGCCATTTGAATTCATGGTTGTTTCAATTAC
ssOB_22	CAACACGTCCGTCGCGCGCGAGTCCAAATAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_40SRPS7pro-1000	
MRT-64	ACTTCCTTTTACCCGATTTAGCATTGAAG
MRT-65	GGTCTACACAACGAATTAGCGAC
MRT-66	GAAAGTCGATACTTGAGGGTGGG
MRT-67	ACAAAGTAAAATCGGTACTTTTCGTAGATG
ssOB_1	CACGACGTTGTAAAACGACGGCCAGTGCCAACTTCCTTTTACCCGATTTAGCATTGAAGC
ssOB_2	TCGAATAGTCGCTAATTCGTTGTGTAGACCATGGTCTTCACACTCGAAGATTTCGTTGGG

Supplemental Table 1. Oligonucleotide sets used for expression vector cloning.

In cases where only the promoter fragment was changed, most oligonucleotides were reused; however, complete sets are shown here for clarity and completeness. Primers used for PCR amplification of DNA fragments are designated MRT-*xxx*, whereas single-stranded oligonucleotide bridges used for DNA assembly are termed ssOB-*yyy*. Oligonucleotides are grouped by plasmid, named pMRT-*zzz*.

ssOB_5	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGAAAGTCGATACTTGAGGGTGGGTTATACT
ssOB_6	CATCTACGAAAAGTACCGATTTTACTTTGTAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_40SRPS7pro-1500	
MRT-63	AGATAGAAATCCAAATTCTGAGCACAGC
MRT-65	GGTCTACACAACGAATTAGCGAC
MRT-66	GAAAGTCGATACTTGAGGGTGGG
MRT-67	ACAAAGTAAAATCGGTACTTTTCGTAGATG
ssOB_2	TCGAATAGTCGCTAATTCGTTGTGTAGACCATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_3	CACGACGTTGTAAAACGACGGCCAGTGCCAAGATAGAAATCCAAATTCTGAGCACAGCGC
ssOB_5	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGAAAGTCGATACTTGAGGGTGGGTTATACT
ssOB_6	CATCTACGAAAAGTACCGATTTTACTTTGTAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_40SRPS7pro-2041	
MRT-62	CGAAGCCAAGAAAAATGGAACCAAAATC
MRT-65	GGTCTACACAACGAATTAGCGAC
MRT-66	GAAAGTCGATACTTGAGGGTGGG
MRT-67	ACAAAGTAAAATCGGTACTTTTCGTAGATG
ssOB_2	TCGAATAGTCGCTAATTCGTTGTGTAGACCATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_4	CACGACGTTGTAAAACGACGGCCAGTGCCACGAAGCCAAGAAAAATGGAACCAAAATCAC
ssOB_5	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGAAAGTCGATACTTGAGGGTGGGTTATACT
ssOB_6	CATCTACGAAAAGTACCGATTTTACTTTGTAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_60SRPL3pro-500	
MRT-69	AATGACAAAATTCACGCCTTCAACCAG
MRT-70	GCTGTTTGATTAGGAAAAAAGTGAGG
MRT-71	GAGAGTCGAACGATCGATTGGGAAAT
MRT-72	TTTTAAATTGTTTTTTGTGTCCTTCATGC
ssOB_7	CACGACGTTGTAAAACGACGGCCAGTGCCAAATGACAAAATTCACGCCTTCAACCAGAAC
ssOB_8	TTTTCCTCACTTTTTTCCTAATCAAACAGCATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_10	AGCATGAAGGACACAAAAAACAATTTAAAAAATTCGTAATCATGGTCATAGCTGTTTCCT
ssOB_11	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGAGAGTCGAACGATCGATTGGGAAATAATT
pMRT_60SRPL3pro-1000	
MRT-68	ACAGACAGCTAGCTAGGTCATTCA
MRT-70	GCTGTTTGATTAGGAAAAAAGTGAGG
MRT-71	GAGAGTCGAACGATCGATTGGGAAAT
MRT-72	TTTTAAATTGTTTTTTGTGTCCTTCATGC
ssOB_8	TTTTCCTCACTTTTTTCCTAATCAAACAGCATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_9	CACGACGTTGTAAAACGACGGCCAGTGCCAACAGACAGCTAGCTAGGTCATTCAGAATGG
ssOB_10	AGCATGAAGGACACAAAAAACAATTTAAAAAATTCGTAATCATGGTCATAGCTGTTTCCT
ssOB_11	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGAGAGTCGAACGATCGATTGGGAAATAATT
pMRT_AALpro-1000	
MRT-80	CACGCAAACGCACCCAGC
MRT-81	TTTCAAGATAAGGGCAAAGAAAAAATATGATTTTC
MRT-82	AGAGAAGAGAAGGGATTGGGAAG
MRT-83	GAAATCGAGTTGAGGCTTCAGG
ssOB_12	CACGACGTTGTAAAACGACGGCCAGTGCCACACGCAAACGCACCCAGCACTGTTGCACGA

Supplemental Table 1. (continued)

ssOB_13	TCATATTTTTTCTTTGCCCTTATCTTGAAAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_15	TGGCGGCTGTGCGAACGCATTCTGGCGTAAAGAGAAGAGAAGGGATTGGGAAGAGTGCTG
ssOB_16	CTCCAAGTCCTGAAGCCTCAACTCGATTTCAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_AALpro-1500	
MRT-79	TCGATATTCTTCCCACATGTGATTTTC
MRT-81	TTTCAAGATAAGGGCAAAGAAAAAATATGATTTTC
MRT-82	AGAGAAGAGAAGGGATTGGGAAG
MRT-83	GAAATCGAGTTGAGGCTTCAGG
ssOB_13	TCATATTTTTTCTTTGCCCTTATCTTGAAAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_14	CACGACGTTGTAAAACGACGGCCAGTGCCATCGATATTCTTCCCACATGTGATTTTCTTT
ssOB_15	TGGCGGCTGTGCGAACGCATTCTGGCGTAAAGAGAAGAGAAGGGATTGGGAAGAGTGCTG
ssOB_16	CTCCAAGTCCTGAAGCCTCAACTCGATTTCAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_CathB2pro-500	
MRT-90	CAACGGTCCTATTGAGGTCGATTTC
MRT-91	ATTTACTTACTAAATATTTTCTGGGAAAGAGGATAAC
MRT-92	GGTGTTCTTCTGGTTGGTTTGTTTG
MRT-93	TCCTCCCTACAAGGAAAGCG
ssOB_33	CACGACGTTGTAAAACGACGGCCAGTGCCACAACGGTCCTATTGAGGTCGATTTCTCTGT
ssOB_34	TCTTTCCCAGAAAATATTTAGTAAGTAAATATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_36	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGGTGTTCTTCTGGTTGGTTTGTTTGTTTGA
ssOB_37	ACGTGAACATCGCTTTCCTTGTAGGGAGGAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_CathB2pro-1000	
MRT-89	GGTGATCTTCCTGAGTCTTTCGAC
MRT-91	ATTTACTTACTAAATATTTTCTGGGAAAGAGGATAAC
MRT-92	GGTGTTCTTCTGGTTGGTTTGTTTG
MRT-93	TCCTCCCTACAAGGAAAGCG
ssOB_34	TCTTTCCCAGAAAATATTTAGTAAGTAAATATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_35	CACGACGTTGTAAAACGACGGCCAGTGCCAGGTGATCTTCCTGAGTCTTTCGACCCCGTG
ssOB_36	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGGTGTTCTTCTGGTTGGTTTGTTTGTTTGA
ssOB_37	ACGTGAACATCGCTTTCCTTGTAGGGAGGAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_CBMCPpro-378	
MRT-85	GTTTGAGATGCTTGGGGCG
MRT-86	TCAAGTATAATTTCAAATCTCATTCATTTCTGAATC
MRT-87	ACACTGATTGGTGTTGTGACAATATTTATTTATTTATTTG
MRT-88	CACACGCGATAAAGACGAATCAATTC
ssOB_38	CACGACGTTGTAAAACGACGGCCAGTGCCAGTTTGAGATGCTTGGGGCGCTTATTGTGTA
ssOB_39	GAAATGAATGAGATTTGAAATTATACTTGAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_41	TGGCGGCTGTGCGAACGCATTCTGGCGTAAACACTGATTGGTGTTGTGACAATATTTATT
ssOB_42	GAACGAATTGATTCGTCTTTATCGCGTGTGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_CBMCPpro-572	
MRT-84	GTAGCAGATTCTTTGGTAGGTTCATC
MRT-86	TCAAGTATAATTTCAAATCTCATTCATTTCTGAATC
MRT-87	ACACTGATTGGTGTTGTGACAATATTTATTTATTTATTTG
MRT-88	CACACGCGATAAAGACGAATCAATTC

Supplemental Table 1. (continued)

ssOB_39	GAAATGAATGAGATTTGAAATTATACTTGAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_40	CACGACGTTGTAAAACGACGGCCAGTGCCAGTAGCAGATTCTTTGGTAGGTTCATCCTAG
ssOB_41	TGGCGGCTGTGCGAACGCATTCTGGCGTAAACACTGATTGGTGTTGTGACAATATTTATT
ssOB_42	GAACGAATTGATTCGTCTTTATCGCGTGTGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_EamApro-1023	
MRT-95	GTATGGAATGCGATGGAGTGTTC
MRT-96	TCTAGAGAATGAACTATTCTAGATTTCACCATAC
MRT-97	TGGTAGTGTTGCGTTTGTTAGAAAC
MRT-98	CACCAAATCCCAAATCTCTTTACGC
ssOB_43	CACGACGTTGTAAAACGACGGCCAGTGCCAGTATGGAATGCGATGGAGTGTTCTCTCTCT
ssOB_44	GGTGAAATCTAGAATAGTTCATTCTCTAGAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_46	TGGCGGCTGTGCGAACGCATTCTGGCGTAATGGTAGTGTTGCGTTTGTTAGAAACACGTT
ssOB_47	AAGCAGCGTAAAGAGATTTGGGATTTGGTGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_EamApro-1245	
MRT-94	GTAAGTTTCTTTCTAGAAGTAGAGTTGATCG
MRT-96	TCTAGAGAATGAACTATTCTAGATTTCACCATAC
MRT-97	TGGTAGTGTTGCGTTTGTTAGAAAC
MRT-98	CACCAAATCCCAAATCTCTTTACGC
ssOB_44	GGTGAAATCTAGAATAGTTCATTCTCTAGAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_45	CACGACGTTGTAAAACGACGGCCAGTGCCAGTAAGTTTCTTTCTAGAAGTAGAGTTGATC
ssOB_46	TGGCGGCTGTGCGAACGCATTCTGGCGTAATGGTAGTGTTGCGTTTGTTAGAAACACGTT
ssOB_47	AAGCAGCGTAAAGAGATTTGGGATTTGGTGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_GDHpro-1500	
MRT-99	GGCAACAATTTCTGCAAGATGAAG
MRT-101	CTATGTGATCAGAGCGGGACG
MRT-102	TCAACCAACCAACCAACCG
MRT-103	TGGGAGGAGCGAGGAGTAG
ssOB_49	TGTTGTCGGCGTCCCGCTCTGATCACATAGATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_50	CACGACGTTGTAAAACGACGGCCAGTGCCAGGCAACAATTTCTGCAAGATGAAGGAGGTG
ssOB_51	TGGCGGCTGTGCGAACGCATTCTGGCGTAATCAACCAACCAACCAACCGCTTCCTTCCTG
ssOB_52	CACATCAGAATCTACTCCTCGCTCCTCCCAAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_G/M_MDHpro-1000	
MRT-105	ATGACAACGAAATAAACGTAATGTAAGAGAAAAG
MRT-106	ATGAAAAGAAAATATTTTGCTAGATATCAATTTGTAGAAAC
MRT-107	ATCGTAATAGTTTTGGTTATGTTTTTTGAATGTTTG
MRT-108	GAGAGAACAGTACTGTGCTAAAAATCAAG
ssOB_53	CACGACGTTGTAAAACGACGGCCAGTGCCAATGACAACGAAATAAACGTAATGTAAGAGA
ssOB_54	TTGATATCTAGCAAAATATTTTCTTTTCATATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_56	TGGCGGCTGTGCGAACGCATTCTGGCGTAAATCGTAATAGTTTTGGTTATGTTTTTTGAA
ssOB_57	TCTTGATTTTTAGCACAGTACTGTTCTCTCAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_G/M_MDHpro-1500	
MRT-104	AGTGGGGAATGTCGGCACTTTTC
MRT-106	ATGAAAAGAAAATATTTTGCTAGATATCAATTTGTAGAAAC
MRT-107	ATCGTAATAGTTTTGGTTATGTTTTTTGAATGTTTG

Supplemental Table 1. (continued)

MRT-108	GAGAGAACAGTACTGTGCTAAAAATCAAG
ssOB_54	TTGATATCTAGCAAAATATTTTCTTTTCATATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_55	CACGACGTTGTAAAACGACGGCCAGTGCCAAGTGGGGAATGTCGGCACTTTTCTAAATGA
ssOB_56	TGGCGGCTGTGCGAACGCATTCTGGCGTAAATCGTAATAGTTTTGGTTATGTTTTTTGAA
ssOB_57	TCTTGATTTTTAGCACAGTACTGTTCTCTCAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_MPTpro-903	
MRT-110	TCGGGTAGATATCCGAAAACGAAG
MRT-111	TTGGAGAAATAAGTATGAGGGATCTAGG
MRT-112	GTGTTTTTCTTTTTGAATAATAACGTAACTAAATGA
MRT-113	CTAGTAATAACCTCAAATAGTCTTGTGCATTAC
ssOB_58	CACGACGTTGTAAAACGACGGCCAGTGCCATCGGGTAGATATCCGAAAACGAAGATTTAC
ssOB_59	TTCCTAGATCCCTCATACTTATTTCTCCAAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_61	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGTGTTTTTCTTTTTGAATAATAACGTAACT
ssOB_62	ATGCACAAGACTATTTGAGGTTATTACTAGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_MPTpro-1688	
MRT-109	GTGTGTTTGGAATGTCCGGAC
MRT-111	TTGGAGAAATAAGTATGAGGGATCTAGG
MRT-112	GTGTTTTTCTTTTTGAATAATAACGTAACTAAATGA
MRT-113	CTAGTAATAACCTCAAATAGTCTTGTGCATTAC
ssOB_60	TTCCTAGATCCCTCATACTTATTTCTCCAAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_59	TTCCTAGATCCCTCATACTTATTTCTCCAAATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_61	TGGCGGCTGTGCGAACGCATTCTGGCGTAAGTGTTTTTCTTTTTGAATAATAACGTAACT
ssOB_62	ATGCACAAGACTATTTGAGGTTATTACTAGAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_SARFGBPpro-600	
MRT-114	CGAACGATAAGGCTCGTATTGAG
MRT-116	CTATACAAAACTGTTAGATATAAGCTCTATATTGATATATTTG
MRT-117	AGTGATTTGAACTAATAGAAATAGCGTTTG
MRT-118	AGTCCACAGTCTCCACATTAAATCC
ssOB_64	CCCCAACGAAATCTTCGAGTGTGAAGACCATCTATACAAAACTGTTAGATATAAGCTCTAT
ssOB_65	CACGACGTTGTAAAACGACGGCCAGTGCCACGAACGATAAGGCTCGTATTGAGGATGCTA
ssOB_68	TGGCGGCTGTGCGAACGCATTCTGGCGTAAAGTGATTTGAACTAATAGAAATAGCGTTTG
ssOB_69	GCATTGGATTTAATGTGGAGACTGTGGACTAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_SARFGBPpro-600	
MRT-117	AGTGATTTGAACTAATAGAAATAGCGTTTG
MRT-118	AGTCCACAGTCTCCACATTAAATCC
MRT-123	TTCTCGATAGAATCTAGTTGAAAGAGAAAATTG
MRT-124	CTATACAAAACTGTTAGATATGAACTCTATATTGATATATTTG
ssOB_66	CACGACGTTGTAAAACGACGGCCAGTGCCATTCTCGATAGAATCTAGTTGAAAGAGAAAA
ssOB_67	ATAGAGTTCATATCTAACAGTTTTGTATAGATGGTCTTCACACTCGAAGATTTCGTTGGG
ssOB_68	TGGCGGCTGTGCGAACGCATTCTGGCGTAAAGTGATTTGAACTAATAGAAATAGCGTTTG
ssOB_69	GCATTGGATTTAATGTGGAGACTGTGGACTAATTCGTAATCATGGTCATAGCTGTTTCCT
pMRT_LeguP1:UnaG::P2A::pac:Legu3′UTR	
MRT-4	GCTAGCTTACAAATTTTTTATGGTATTATTTCTAC
MRT-5	GGATCCGCATTGAGTGTATATGTTTG

Supplemental Table 1. (continued)

MRT-159	ACTTTGTACAAAAAAGCAGGCTTCGC
MRT-160	TTCAGTCGCTCGGCGATACG
MRT-165	ATGACCGAGTACAAGCCTACTGTG
MRT-166	TTAAGCACCAGGCTTGCGC
ssOB_81	AATAATACCATAAAAAATTTGTAAGCTAGCACTTTGTACAAAAAAGCAGGCTTCGCCACC
ssOB_82	GCTGTGCGATCGTATCGCCGAGCGACTGAAGGCAGCGGAGCGACGAACTTCTCGCTGCTGAAGC AGGCTGGCGACGTGGA
ssOB_94	TGGTGCATGACGCGCAAGCCTGGTGCTTAAGGATCCGCATTGAGTGTATATGTTTGTTAT
ssOB_95	CAGTCGCACAGTAGGCTTGTACTCGGTCATCGGTCCAGGGTTCTCTTCCACGTCGCCAGCCTGCTT CAGCAGCGAG
pMRT_LeguP1:smURFP::P2A::Ecdhfr:Legu3'UTR	
MRT-4	GCTAGCTTACAAATTTTTTATGGTATTATTTCTAC
MRT-5	GGATCCGCATTGAGTGTATATGTTTG
MRT-159	ACTTTGTACAAAAAAGCAGGCTTCGC
MRT-161	ATGGAAAACGCTATGCCGTGGAAC
MRT-162	TTATCGGCGTTCCAGGATCTCGAAG
MRT-267	CGACATAGCCTTGATGATGTAATCGAAGTAAG
ssOB_81	AATAATACCATAAAAAATTTGTAAGCTAGCACTTTGTACAAAAAAGCAGGCTTCGCCACC
ssOB_85	CTGCTTCGAGATCCTGGAACGCCGATAAGGATCCGCATTGAGTGTATATGTTTGTTATAA
ssOB_86	AGGCAGGTTCCACGGCATAGCGTTTTCCATCGGTCCAGGGTTCTCTTCCACGTCGCCAG CCTGCTTCAGCAGCGAG
ssOB_165	TACTTCGATTACATCATCAAGGCTATGTCGGGCAGCGGAGCGACGAACTTCTCGCTGCTG
pMRT_LeguP1:SNAPtag::P2A::pac:Legu3'UTR	
MRT-4	GCTAGCTTACAAATTTTTTATGGTATTATTTCTAC
MRT-5	GGATCCGCATTGAGTGTATATGTTTG
MRT-159	ACTTTGTACAAAAAAGCAGGCTTCGC
MRT-165	ATGACCGAGTACAAGCCTACTGTG
MRT-166	TTAAGCACCAGGCTTGCGC
MRT-167	ACATGAACAGAATAATCCAGAACGAGGAAC
ssOB_81	AATAATACCATAAAAAATTTGTAAGCTAGCACTTTGTACAAAAAAGCAGGCTTCGCCACC
ssOB_94	TGGTGCATGACGCGCAAGCCTGGTGCTTAAGGATCCGCATTGAGTGTATATGTTTGTTAT
ssOB_95	CAGTCGCACAGTAGGCTTGTACTCGGTCATCGGTCCAGGGTTCTCTTCCACGTCGCCAGC CTGCTTCAGCAGCGAG
ssOB_167	GGACACCGCCTCGGAAAGCCTGGACTGGGAGGCAGCGGAGCGACGAACTTCTCGCTGCTG

Supplemental Table 1. (continued)

Acknowledgements

MRT was supported by Norwegian Research Council grant 301170 to MvdG. We thank Ms. Geok Choo Ng and Dr. Steven Santino Leonardi for generously sharing their wet-lab techniques for culturing *Blastocystis.* We thank Dr. Yohannes Seyoum for helping with the statistical analyses. We also thank Luz Aurora Martinez-Contreras, Daniel Simon Nagel, and Ayanara Mae Sadi Bukneberg for their help in routine culturing and maintenance of *Blastocystis* cultures.

Additional information

Author contributions

Conceptualisation and design of the research: MRT and MvdG Critical advice: MvdG and KSWT Experimental work and data analyses: MRT Writing and editing: MRT, MvdG, and KSWT All authors read and revised the paper.

Competing financial interests

The authors declare no competing interests.

Funding

Funder	Grant reference number	Author
Norges Forskningsråd (Forskningsrådet)	301170	M Rey Toleco Mark van der Giezen

Author ORCID iDs

Mark van der Giezen: https://orcid.org/0000-0002-1033-1335

References

Abrahamian M, Kagda M, Ah-Fong AMV, Judelson HS (2017) Rethinking the evolution of eukaryotic metabolism: novel cellular partitioning of enzymes in stramenopiles links serine biosynthesis to glycolysis in mitochondria. *BMC Evol Biol* **17**:241 https://doi.org/10.1186/s12862-017-1087-8 | PubMed

Alfellani MA, Stensvold CR, Vidal-Lapiedra A, Onuoha ESU, Fagbenro-Beyioku AF, Clark CG (2013) Variable geographic distribution of *Blastocystis* subtypes and its potential implications. *Acta Tropica* **126**:11-18 https://doi.org/10.1016/j.actatropica.2012.12.011 | PubMed

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**:403-410 https://doi.org/10.1016/S0022-2836(05)80360-2 | PubMed

Armengaud J, Pible O, Gaillard J-C, Cian A, Gantois N, Tan KSW, Chabé M, Viscogliosi E (2017) Proteogenomic insights into the intestinal parasite *Blastocystis* sp. subtype 4 isolate WR1. *Proteomics* **17**:1700211 https://doi.org/10.1002/pmic.201700211 | PubMed

Asselbergs FAM, Widmer R (1995) Use of the *Escherichia coli* chromosomal DHFR gene as selection marker in mammalian cells. *Journal of Biotechnology* **43**:133-138 https://doi.org/10.1016/0168-1656(95)00131-3 | PubMed

Bártulos CR, Rogers MB, Williams TA, Gentekaki E, Brinkmann H, Cerff R, Liaud M-F, Hehl AB, Yarlen NR, Gruber A, *et al.* (2018) Mitochondrial glycolysis in a major lineage of eukaryotes. *Genome Biol and Evol* **10**:2310-2325 https://doi.org/10.1093/gbe/evy164 | PubMed

Benchling (2025) Benchling. https://www.benchling.com/

Birke R, Ast J, Roosen DA, Lee J, Roßmann K, Huhn C, Mathes B, Lisurek M, Bushiri D, Sun H, *et al.* (2022) Sulfonated red and far-red rhodamines to visualize SNAP- and Halo-tagged cell surface proteins. *Org Biomol Chem* **20**:5967-5980 https://doi.org/10.1039/d1ob02216d | PubMed

Bracken ML, Fernandes MA, Mathura S (2024) Bilirubin and its crystal forms. *CrystEngComm* **26**:2136-2142 https://doi.org/10.1039/D4CE00123K

Bulmer AC, Blanchfield JT, Coombes JS, Toth I (2008) *In vitro* permeability and metabolic stability of bile pigments and the effects of hydrophilic and lipophilic modification of biliverdin. *Bioorg Med Chem* **16**:3616-3625 https://doi.org/10.1016/j.bmc.2008.02.008 | PubMed

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421 https://doi.org/10.1186/1471-2105-10-421 | PubMed

Chen XQ, Singh M, Ho LC, Tan SW, Ng GC, Moe KT, Yap EH (1997) Description of a *Blastocystis* species from *RaKus norvegicus*. *Parasitology Research* **83**:313-318 https://doi.org/10.1007/s004360050255 | PubMed

Deng L, Tan KSW (2025) From parasite to partner: unravelling the multifaceted role of *Blastocystis* in human health and disease. *The Lancet Microbe* **101155** https://doi.org/10.1016/j.lanmic.2025.101155 | PubMed

Deng L, Tan KSW (2022) Interactions between *Blastocystis* subtype ST4 and gut microbiota *in vitro*. *Parasites Vectors* **15**:80 https://doi.org/10.1186/s13071-022-05194-x | PubMed

Denoeud F, Roussel M, Noel B, Wawrzyniak I, Da Silva C, Diogon M, Viscogliosi E, Brochier-Armanet C, Couloux A, Poulain J, *et al.* (2011) Genome sequence of the stramenopile *Blastocystis*, a human anaerobic parasite. *Genome Biol* **12**:R29 https://doi.org/10.1186/gb-2011-12-3-r29 | PubMed

Fidock DA, Wellems TE (1997) Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil. *Proc Natl Acad Sci U S A* **94**:10931-10936 https://doi.org/10.1073/pnas.94.20.10931 | PubMed

Gautier A, Juillerat A, Heinis C, Corrêa IR, Kindermann M, Beaufils F, Johnsson K (2008) An engineered protein tag for multiprotein labeling in living cells. *Chem Biol* **15**:128-136 https://doi.org/10.1016/j.chembiol.2008.01.007 | PubMed

Gentekaki E, Curtis BA, Stairs CW, Klimeš V, Eliáš M, Salas-Leiva DE, Herman EK, Eme L, Arias MC, Henrissat B, *et al.* (2017) Extreme genome diversity in the hyper-prevalent parasitic eukaryote *Blastocystis*. *PLoS Biol* **15**:e2003769 https://doi.org/10.1371/journal.pbio.2003769 | PubMed

Ho LC, Singh M, Suresh G, Ng GC, Yap EH (1993) Axenic culture of *Blastocystis hominis* in Iscove's modified Dulbecco's medium. *Parasitol Res* **79**:614-616 https://doi.org/10.1007/bf00932249 | PubMed

Hu X, Xu Y, Yi J, Wang C, Zhu Z, Yue T, Zhang H, Wang X, Wu F, Xue L, *et al.* (2024) Using protein design and directed evolution to monomerize a bright near-infrared fluorescent protein. *ACS Synth Biol* **13**:1177-1190 https://doi.org/10.1021/acssynbio.3c00643 | PubMed

Keppler A, Gendreizig S, Gronemeyer T, Pick H, Vogel H, Johnsson K (2003) A general method for the covalent labeling of fusion proteins with small molecules *in vivo*. *Nat Biotechnol* **21**:86-89 https://doi.org/10.1038/nbt765 | PubMed

Kim JH, Lee S-R, Li L-H, Park H-J, Park J-H, Lee KY, Kim MK, Shin BA, Choi S-Y (2011) High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. *PLoS ONE* **6**:e18556 https://doi.org/10.1371/journal.pone.0018556 | PubMed

Klimeš V, Gentekaki E, Roger AJ, Eliáš M (2014) A large number of nuclear genes in the human parasite *Blastocystis* require mRNA polyadenylation to create functional termination codons. *Genome Biology and Evolution* **6**:1956-1961 https://doi.org/10.1093/gbe/evu146 | PubMed

Koehler AV, Herath HMPD, Hall RS, Wilcox S, Gasser RB (2024) Marked genetic diversity within *Blastocystis* in Australian wildlife revealed using a next generation sequencing–phylogenetic approach. *International Journal for Parasitology: Parasites and Wildlife* **23**:100902 https://doi.org/10.1016/j.ijppaw.2023.100902 | PubMed

Kumagai A, Ando R, Miyatake H, Greimel P, Kobayashi T, Hirabayashi Y, Shimogori T, Miyawaki A (2013) A bilirubin-inducible fluorescent protein from eel muscle. *Cell* **153**:1602-1611 https://doi.org/10.1016/j.cell.2013.05.038 | PubMed

Kwon J, Park J-S, Kang M, Choi S, Park J, Kim GT, Lee C, Cha S, Rhee H-W, Shim S-H (2020) Bright ligand-activatable fluorescent protein for high-quality multicolor live-cell super-resolution microscopy. *Nat Commun* **11**:273 https://doi.org/10.1038/s41467-019-14067-4 | PubMed

Li F-J, Tsaousis AD, Purton T, Chow VTK, He CY, Tan KSW (2019) Successful genetic transfection of the colonic protistan parasite *Blastocystis* for reliable expression of ectopic genes. *Sci Rep* **9**:3159 https://doi.org/10.1038/s41598-019-39094-5 | PubMed

Liao CC, Chen CH, Shin JW, Lin WC, Chen CC, Chu CT (2023) Lipid accumulation in *Blastocystis* increases cell damage in co-cultured cells. *Microorganisms* **11**:1582 https://doi.org/10.3390/microorganisms11061582 | PubMed

Lightner DA, Holmes DL, McDonagh AF (1996) On the acid dissociation constants of bilirubin and biliverdin: pKa values from 13C NMR spectroscopy. *Journal of Biological Chemistry* **271**:2397-2405 https://doi.org/10.1074/jbc.271.5.2397 | PubMed

Maiti A, Buffalo CZ, Saurabh S, Montecinos-Franjola F, Hachey JS, Conlon WJ, Tran GN, Hassan B, Walters KJ, Drobizhev M, *et al.* (2023) Structural and photophysical characterization of the small ultra-red fluorescent protein. *Nat Commun* **14**:4155 https://doi.org/10.1038/s41467-023-39776-9 | PubMed

Markus BM, Bell GW, Lorenzi HA, Lourido S (2019) Optimizing systems for Cas9 expression in *Toxoplasma gondii*. *mSphere* **4**:e00386-19 https://doi.org/10.1128/mSphere.00386-19 | PubMed

Mirza H, Teo JDW, Upcroft J, Tan KSW (2011) A rapid, high-throughput viability assay for *Blastocystis* spp. reveals metronidazole resistance and extensive subtype-dependent variations in drug susceptibilities. *Antimicrob Agents Chemother* **55**:637-648 https://doi.org/10.1128/AAC.00900-10 | PubMed

Müntjes K, Philipp M, Hüsemann L, Heucken N, Weidtkamp-Peters S, Schipper K, Zurbriggen MD, Feldbrügge M (2020) Establishing polycistronic expression in the model microorganism *Ustilago maydis*. *Front Microbiol* **11**:1384 https://doi.org/10.3389/fmicb.2020.01384 | PubMed

Nagel R, Gray C, Bielefeldt-Ohmann H, Traub RJ (2015) Features of *Blastocystis* spp. in xenic culture revealed by deconvolutional microscopy. *Parasitol Res* **114**:3237-3245 https://doi.org/10.1007/s00436-015-4540-x | PubMed

Ostrow JD, Celic L (1984) Bilirubin chemistry, ionization and solubilization by bile salts. *Hepatology* **4**:38S-45S https://doi.org/10.1002/hep.1840040807 | PubMed

Piperni E, Nguyen LH, Manghi P, Kim H, Pasolli E, Andreu-Sánchez S, Arrè A, Bermingham KM, Blanco-Míguez A, Manara S, *et al.* (2024) Intestinal *Blastocystis* is linked to healthier diets and more favorable cardiometabolic outcomes in 56,989 individuals from 32 countries. *Cell* **187**:4554-4570.e18. https://doi.org/10.1016/j.cell.2024.06.018 | PubMed

Porzberg N, Gries K, Johnsson K (2025) Exploiting covalent chemical labeling with self-labeling proteins. *Annual Review of Biochemistry* **94**:29-58 https://doi.org/10.1146/annurev-biochem-030222-121016 | PubMed

Pronobis MI, Deuitch N, Peifer M (2016) The miraprep: A protocol that uses a miniprep kit and provides maxiprep yields. *PLOS One* **11**:e0160509 https://doi.org/10.1371/journal.pone.0160509 | PubMed

Pyrihová E, King MS, King AC, Toleco MR, van der Giezen M, Kunji ER (2024) A mitochondrial carrier transports glycolytic intermediates to link cytosolic and mitochondrial glycolysis in the human gut parasite *Blastocystis*. *eLife* **13**:RP94187 https://doi.org/10.7554/eLife.94187 | PubMed

Remcho TP, Guggilapu SD, Cruz P, Nardone GA, Heffernan G, O'Connor RD, Bewley CA, Wellems TE, Lane KD (2020) Regioisomerization of antimalarial drug WR99210 explains the inactivity of a commercial stock. *Antimicrob Agents Chemother* **65**:e01385-20 https://doi.org/10.1128/AAC.01385-20 | PubMed

Rodriguez EA, Tran GN, Gross LA, Crisp JL, Shu X, Lin JY, Tsien RY (2016) A far-red fluorescent protein evolved from a cyanobacterial phycobiliprotein. *Nat Methods* **13**:763-769 https://doi.org/10.1038/nmeth.3935 | PubMed

Saimi D, Chen Z (2023) Chemical tags and beyond: Live-cell protein labeling technologies for modern optical imaging. *Smart Mol* **1**:e20230002 https://doi.org/10.1002/smo.20230002 | PubMed

Scanlan PD, Stensvold CR (2013) *Blastocystis*: gening to grips with our guileful guest. *Trends in Parasitology* **29**:523-529 https://doi.org/10.1016/j.pt.2013.08.006 | PubMed

Schueåe T, Meyer V (2017) Polycistronic gene expression in *Aspergillus niger*. *Microb Cell Fact* **16**:162 https://doi.org/10.1186/s12934-017-0780-z | PubMed

Šejnohová A, Koutenská M, Jirků M, Brožová K, Pavlíčková Z, Kadlecová O, Cinek O, Maloney JG, Santín M, Petrželková KJ, *et al.* (2024) A cross-sectional survey of *Blastocystis* sp. and *Dientamoeba fragilis* in non-human primates and their caregivers in Czech zoos. *One Health* **19**:100862 https://doi.org/10.1016/j.onehlt.2024.100862 | PubMed

Stechmann A, Hamblin K, Pérez-Brocal V, Gaston D, Richmond GS, van der Giezen M, Clark CG, Roger AJ (2008) Organelles in *Blastocystis* that blur the distinction between mitochondria and hydrogenosomes. *Current Biology* **18**:580-585 https://doi.org/10.1016/j.cub.2008.03.037 | PubMed

Stensvold CR, van der Giezen M (2018) Associations between gut microbiota and common luminal intestinal parasites. *Trends in Parasitology* **34**:369-377 https://doi.org/10.1016/j.pt.2018.02.004 | PubMed

Tan SW, Singh M, Thong KT, Ho LC, Moe KT, Chen XQ, Ng GC, Yap EH (1996a) Clonal growth of *Blastocystis hominis* in soft agar with sodium thioglycollate. *Parasitol Res* **82**:737-739 https://doi.org/10.1007/s004360050194 | PubMed

Tan SW, Singh M, Yap EH, Ho LC, Moe KT, Howe J, Ng GC (1996b) Colony formation of *Blastocystis hominis* in soft agar. *Parasitol Res* **82**:375-377 https://doi.org/10.1007/s004360050130 | PubMed

Tan KSW, Ng GC, Quek E, Howe J, Ramachandran NP, Yap EH, Singh M (2000) *Blastocystis hominis*: A simplified, high-efficiency method for clonal growth on solid agar. *Exp Parasitol* **96**:9-15 https://doi.org/10.1006/expr.2000.4544 | PubMed

Tan KSW (2008) New insights on classification, identification, and clinical relevance of *Blastocystis* spp. *Clin Microbiol Rev* **21**:639-665 https://doi.org/10.1128/CMR.00022-08 | PubMed

Toleco MR, van der Giezen M (2025) A simple and universal quasi-modular cloning system using NEBuilder® HiFi DNA assembly kit. *BioTechniques* **77**:257-262 https://doi.org/10.1080/07366205.2025.2542012 | PubMed

Tsaousis AD, Hamblin KA, Ellion CR, Young L, Rosell-Hidalgo A, Gourlay CW, Moore AL, van der Giezen M (2018) The human gut colonizer *Blastocystis* respires using complex II and alternative oxidase to buffer transient oxygen fluctuations in the gut. *Front Cell Infect Microbiol* **8**:371 https://doi.org/10.3389/fcimb.2018.00371 | PubMed

Vara JA, Pulido D, Lacalle RA, Jiménez A (1988) Two genes in *Streptomyces alboniger* puromycin biosynthesis pathway are closely linked. *Gene* **69**:135-140 https://doi.org/10.1016/0378-1119(88)90386-1 | PubMed

Wu Z, Mirza H, Tan KSW (2014) Intra-subtype variation in enteroadhesion accounts for differences in epithelial barrier disruption and is associated with metronidazole resistance in *Blastocystis* subtype-7. *PLoS Negl Trop Dis* **8**:e2885 https://doi.org/10.1371/journal.pntd.0002885 | PubMed

Yason JA, Liang YR, Png CW, Zhang Y, Tan KSW (2019) Interactions between a pathogenic *Blastocystis* subtype and gut microbiota: *in vitro* and *in vivo* studies. *Microbiome* **7**:30 https://doi.org/10.1186/s40168-019-0644-3 | PubMed

Yason JA, Tan KSW (2018) Membrane surface features of *Blastocystis* subtypes. *Genes* **9**:417 https://doi.org/10.3390/genes9080417 | PubMed

Yoshikawa H, Hayakawa A (1996) Freeze-fracture cytochemistry of membrane cholesterol in *Blastocystis hominis*. *International Journal for Parasitology* **26**:1111-1114 https://doi.org/10.1016/S0020-7519(96)80010-5

Záhonová K, Low RS, Warren CJ, Cantoni D, Herman EK, Yiangou L, Ribeiro CA, Phanprasert Y, Brown IR, Rueckert S, *et al.* (2023) Evolutionary analysis of cellular reduction and anaerobicity in the hyper-prevalent gut microbe *Blastocystis*. *Current Biology S* **0960982223006206** https://doi.org/10.1016/j.cub.2023.05.025 | PubMed

Zheng X, Gallot G (2020) Dynamics of cell membrane permeabilization by saponins using teraherå anenuated total reflection. *Biophys J* **119**:749-755 https://doi.org/10.1016/j.bpj.2020.05.040 | PubMed

Zhu X, Ricci-Tam C, Hager ER, Sgro AE (2023) Self-cleaving peptides for expression of multiple genes in *Dictyostelium discoideum*. *PLOS One* **18**:e0281211 https://doi.org/10.1371/journal.pone.0281211 | PubMed

Zucker SD, Goessling W, Hoppin AG (1999) Unconjugated bilirubin exhibits spontaneous diffusion through model lipid bilayers and native hepatocyte membranes. *J Biol Chem* **274**:10852-10862 https://doi.org/10.1074/jbc.274.16.10852 | PubMed

Armengaud J, Pible O, Gaillard JC, Cian A, Gantois N, Tan KSW, Chabé M, Viscogliosi E (2017) Proteogenomic of Blastocystis sp. subtype 4 isolate WR1. ProteomeXchange. ID PXD006588 https://proteomecentral.proteomexchange.org/?pxid=PXD006588

Peer reviews

Reviewer #1 (Public review):

Summary:

This paper presents a toolkit for the transformation of Blastocystis. The authors have screened a number of selectable agents, promoters and reporter genes and present their findings. This resource will be of immense use to those in Blastocystsis field, as well as those seeking to establish transformation tools in other species where such tools do not yet exist. Establishing new transformation tools is extremely challenging, and the authors have done an excellent job.

Strengths:

The authors have carried out a systematic screen of promoters, reporter genes and selectable agents. They have screened numerous for each, and all the data is presented. It is good to see when things did not work as well as when things did - so this data set is extremely useful indeed.

Weaknesses:

The findings are reported by reporter gene assay (microscopy). No evidence is given using genetics. The authors claim that the DNA is maintained episomally. However, could it be possible that there is integration? No PCRS/RT-PCRs are shown (although it can safely be assumed that the DNA/RNA is present where the transformation was successful), nor are any Western blots. These would have been useful to show that the P2A ribosomal skipping had occurred, and that proteins were expressed individually rather than as a polyprotein.

Comments on revised version.

The authors have revised their manuscript to clarify that molecular analyses have not yet occurred and have resolved the technical/publication issues with the figures. I look forward to seeing these tools used in future publications to answer important questions in Blastocystsis research.

https://doi.org/10.7554/eLife.111816.2.sa2

Reviewer #3 (Public review):

Summary:

The primary objective of this study was to establish a practical and functional framework for propagation of stable transgenic cell lines of Blastocystis, a common animal gut microeukaryote. Although the work focused on Blastocystis ST7-B, a subtype with relatively low prevalence in humans, this choice is justified by its association with more frequent negative health effects. Beyond their relevance to the medical field, the methodological advances described here have the potential to also expand cell biology studies of this anaerobic organism, including its unusual mitochondria and redox metabolism.

Strengths:

Prior to this work, genetic tools for Blastocystis were very limited, relying on a single strong promoter-terminator combination. The authors successfully expanded the available promoter set across a range of expression strengths by testing two dozen variants in luciferase-based assays. Critically, they developed an integrated workflow from a modular transgenic construct design to an expanded inventory of molecular components (promoters, reporters), optimized DNA delivery, stepwise antibiotic resistance-mediated clonal selection and propagation, and to reporter validation. The evaluation of several anaerobiosis-compatible labeling strategies for live (and fixed) cell optical imaging will be particularly useful, with the SNAP-tag system appearing especially promising for Blastocystis.

Weaknesses:

The presented data generally provide a solid support for the conclusions that the work reached, but clarification of reasoning and several inconsistencies, as well as amendments to visual presentation of the data would be highly beneficial, as detailed below.

(1) Episomal persistence of the construct:

The manuscript repeatedly assumes, including in its title, that constructs persist in Blastocystis in their episomal form, but no direct evidence is provided. Although this interpretation is plausible, it should be identified more clearly as provisional. Nuclear genomic integration (e.g., via NHEJ) remains a possible explanation unless supporting evidence or rationale is provided to exclude it. Testing whether the phenotype persists without drug-mediated selection in the generated transgenic cell lines would help strengthen the case for episomal maintenance.

(2) Promoters and terminators:

(2.1) There is a discrepancy between the claimed number of loci (14), from which promoters used to drive luciferase expression were derived, and those detailed as having been actually generated in Table 1 (11). This inconsistency should be corrected or explained, as it creates uncertainty around the accuracy of the dataset.

(2.2) Based on the presented evidence, constructs benchmarked in bioluminescence assays differed only in their promoter composition. Although terminator selection is mentioned in the Methods section, no additional details are provided; for instance, Table 1 and Figure 2 only list 23 promoters in total. Figure 2A likewise shows only promoter-dependent variation. If the terminator was held constant (LeguP1?), this should be stated explicitly. The authors may then consider revising the wording of having tested "23 promoter-terminator pairs" to better reflect that only promoters varied.

(2.3) Promoter benchmarking was done with a plasmid lacking a selection marker, so it is unclear how the maintenance of the luciferase construct was ensured. Without selection, the observed reporter intensity could reflect differential or stochastic plasmid retention rather than promoter strength alone. The luminescence assay was performed 16-18 hours after transfection, but the rationale for this particular timeframe should be explained. In this context, the authors should explicitly state whether the experiments shown in Fig.2A represent biological triplicates or technical triplicates from a single transfection.

(3) Figure 2:

(3.1) Several aspects of the current design may lead to ambiguity for the reader. The boxplots are colour-coded, but it is unclear whether the colours carry meaning or are purely decorative. Because the data are already spatially separated into bins, additional random colouring is redundant and may suggest distinctions that are not intended. In addition, the part A of Figure 2 is split into two panels with the scale for the left panel shown in the right panel and some of the boxplot colours falling in the range of the scale, but not in line with their counterparts in the left panel. Because the colour use is not consistent, it is difficult to tell whether the same scale should be applied to both panels or how it should be interpreted.

(3.2) The left panel of the part A uses a diverging blue-white-red colour scheme, which is most appropriate when the midpoint represents a meaningful central value such as zero. Because the values shown in this graph are only positive, a non-diverging 2-colour scale or a colour palette such as 'viridis' would make the plot easier to interpret.

(3.3) A black background should be avoided: 'B' and 'C' labels are invisible and it draws attention to a distracting design feature rather to the data themselves.

(4) Figure 3:

(4.1) Individual snapshots should be separated more clearly, either by using a white background or by adding visible borders to make the overall composition clearer. As currently displayed, some boundaries between fluorescent channels resemble image artifacts rather than intentional panel divisions.

(4.2) In the parts B-D, the legend should explain more clearly what each image shows and the figure itself would benefit from annotations. There seem to be three sub-panels in each 'condition' of part B (as well as C and D): while the middle and rightmost panel can be easily inferred to represent the fluorescent protein and bright-field image, what the leftmost panels represent is not specified. If DAPI was used to dye DNA, an explanation why mostly multiple labelled regions are visible should be provided.

(4.3) Cell morphology and appearance differ markedly between UnaG/smURFP and SNAP-tag images, which should be explained. A microscope issue is mentioned in the main text, but if that was the cause, the authors should consider replacing the images as the current distortions complicate interpretation.

Comments on revised version.

The revised version provides sufficient clarity and appropriate visual presentation. Some confusion evidently arose due to my misunderstanding, so I thank the authors for their comprehensive clarifications and patience.

https://doi.org/10.7554/eLife.111816.2.sa1

Author response:

The following is the authors' response to the original reviews.

In revising the manuscript, we have focused on three main priorities raised during review: (1) improving precision around evidential claims, particularly concerning vector maintenance and P2A-mediated protein separation; (2) substantially improving figure quality, accessibility, and legend clarity; and (3) correcting inconsistencies and expanding methodological detail where requested.

This study was intended as a foundational genetic toolkit and methodological framework for *Blastocystis* ST7-B, establishing practical workflows for DNA delivery, endogenous regulatory-element benchmarking, antibiotic-selected recovery, clonal propagation, and reporter-based analysis in a genetically challenging anaerobic microbial eukaryote. The central evidence presented is therefore functional in nature: reproducible transgene delivery, selectable recovery and propagation of colony-derived transgenic lines, and detectable reporter expression using multiple anaerobic-compatible reporter systems.

We agree with the reviewers that several additional experiments, including Western blot analysis of P2A-containing constructs, outward-facing PCR, plasmid rescue assays, and selection-withdrawal experiments, would further strengthen the mechanistic interpretation of the system and help distinguish episomal persistence from genomic integration. We have therefore revised the manuscript throughout to clearly separate what is directly demonstrated from what remains a plausible working interpretation or important future direction.

Importantly, the revised manuscript no longer presents episomal maintenance or complete P2A-mediated protein separation as demonstrated conclusions. Instead, these are now discussed explicitly as unresolved mechanistic questions requiring future molecular analysis. Nevertheless, the central methodological conclusion remains unchanged: stable selectable transgene expression, recovery of colony-derived transgenic lines, and reporter-positive *Blastocystis* ST7-B transformants can now be reproducibly obtained.

> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *This paper presents a toolkit for the transformation of Blastocystis. The authors have screened a number of selectable agents, promoters and reporter genes and present their findings. This resource will be of immense use to those in the Blastocystis field, as well as those seeking to establish transformation tools in other species where such tools do not yet exist. Establishing new transformation tools is extremely challenging, and the authors have done an excellent job.*
>
> *Strengths:*
>
> *The authors have carried out a systematic screen of promoters, reporter genes and selectable agents. They have screened numerous for each, and all the data is presented. It is good to see when things did not work as well as when things did, so this data set is extremely useful indeed.*
>
> *Weaknesses:*
>
> *The findings are reported by reporter gene assay (microscopy). No evidence is given using genetics. The authors claim that the DNA is maintained episomally. However, could it be possible that there is integration? No PCRS/RT-PCRs are shown (although it can safely be assumed that the DNA/RNA is present where the transformation was successful), nor are any Western blots. These would have been useful to show that the P2A ribosomal skipping had occurred, and that proteins were expressed individually rather than as a polyprotein.*

We thank the reviewer for the positive assessment of the manuscript and for recognising both the technical difficulty and broader utility of establishing genetic tools in *Blastocystis* and other experimentally challenging microbial eukaryotes. We also appreciate the reviewer's identification of the main evidential limitations in the original manuscript, particularly regarding vector maintenance and P2A-mediated protein separation.

First, regarding the question of vector topology and the interpretation of episomal maintenance.

We agree that the original manuscript presented episomal persistence too strongly relative to the evidence currently available. We have therefore revised the manuscript throughout to clarify that episomal maintenance should presently be regarded as a plausible working model rather than a directly demonstrated conclusion.

The transfection system used here was adapted from Li et al. (2019), including use of the pXS2-$P_{Legumain}$-derived plasmid framework. Importantly, the construct used in the present study does not contain the original *Trypanosoma brucei* tubulin-targeting region associated with homologous integration in the original pXS2 system. Complete plasmid sequencing confirmed that the constructs function here as heterologous expression plasmids carrying *Blastocystis* ST7-B regulatory elements and transgenes. While this does not demonstrate episomal persistence, it also means that genomic integration cannot be inferred from the historical pXS2 vector architecture alone.

We further note that comparative genomic analyses by Gentekaki et al. (2017) suggest that *Blastocystis* lacks components of the canonical non-homologous end-joining (NHEJ) machinery, implying that homologous recombination is likely to represent the principal route for double-stranded DNA repair. Because the constructs used here did not contain *Blastocystis* homology arms, there is currently no obvious mechanism favouring targeted homologous integration. Nevertheless, we fully agree that genomic integration cannot presently be excluded.

To reflect this appropriately, the revised manuscript now explicitly separates the demonstrated functional outcomes from unresolved mechanistic questions concerning vector maintenance. We also identify several future approaches that would help distinguish episomal persistence from genomic integration, including outward-facing PCR, plasmid rescue followed by full plasmid sequencing, Southern blotting, FISH, selection-withdrawal experiments, and long-read sequencing approaches.

We have revised the manuscript throughout to remove statements implying demonstrated episomal maintenance and now present episomal persistence only as a plausible working interpretation.

In the Methods section under Cloning, the following text has been added:

Lines 202–206: "The constructs used in this study were derived from the pXS2-$P_{Legumain}$ vector described by Li et al. (2019), which adapted a heterologous expression-vector backbone for transient plasmid-based expression in *Blastocystis* ST7-B. Here, the same molecular backbone was used as a plasmid scaffold carrying *Blastocystis*-derived regulatory elements and transgenes."

In the Discussion, the following text has been added/edited:

Lines 665–673: "The molecular maintenance state of the introduced constructs remains unresolved: episomal maintenance is a plausible working model, but genomic integration cannot be formally excluded. The constructs used here lack *Blastocystis* homology arms, and comparative genomic analyses suggest that *Blastocystis* lacks canonical non-homologous end-

joining components (Gentekaki et al., 2017), making targeted integration by standard repair routes unlikely but not impossible. Direct assays such as outward-facing PCR, plasmid rescue followed by full plasmid sequencing, FISH, or selection-withdrawal experiments will be required to distinguish episomal persistence from integration."

Second, regarding P2A-mediated protein separation.

We agree that Western blotting would provide the most direct biochemical assessment of P2A-mediated ribosomal skipping efficiency in *Blastocystis* ST7-B and would help determine the extent of any residual uncleaved fusion product. We have therefore revised the manuscript to avoid implying that complete protein-level separation was directly demonstrated.

The revised manuscript now states only what is directly supported by the current data: that P2A-containing bicistronic constructs supported antibiotic-selected recovery of transgenic lines together with detectable downstream reporter expression. The microscopy data therefore support functional downstream reporter expression, but do not by themselves exclude residual uncleaved fusion products.

We selected P2A because it is a compact and well-characterised peptide with high reported separation efficiency across multiple eukaryotic systems, including microbial eukaryotes. However, we agree that P2A performance can be context-dependent, and we now explicitly identify biochemical validation of P2A cleavage efficiency as an important future direction.

Importantly, these revisions do not alter the central methodological conclusion of the study, namely that selectable transgene expression, propagation of reporter-positive lines, and recovery of colony-derived *Blastocystis* ST7-B transformants can now be reproducibly achieved.

Text inserted in the Results:

Lines 394–396: "The P2A peptide is expected to promote ribosomal skipping during translation, allowing two separate polypeptides to be produced from a single open reading frame."

Lines 403–404: "However, protein-level separation was not directly tested, and the extent of any residual uncleaved fusion product remains unresolved."

Text inserted in the Discussion:

Lines 619–629: "P2A was selected because it is a well-characterised peptide with high reported separation efficiency in human cell lines, zebrafish embryos, and mice (Kim et al., 2011). It also has precedent across microbial eukaryotes, including the protest *Dictyostelium discoideum* (Zhu et al., 2023), the fungi *Aspergillus niger* (Schuetze and Meyer, 2017) and *Ustilago maydis* (Müntjes et al., 2020), and the apicomplexan parasites *Toxoplasma gondii* (Markus et al., 2019) and *Plasmodium falciparum* (Dans et al., 2024). However, P2A performance is context-dependent, and the evidence presented here is functional rather than biochemical. P2A-containing constructs support antibiotic-selected recovery and downstream reporter expression in *Blastocystis* ST7-B, but ribosomal skipping efficiency and any residual uncleaved product will require direct protein-level validation."

> ***Reviewer #1 (Recommendations for the authors):***
>
> *(1) Please could you show a Western blot to confirm if P2A has worked? It could be that the proteins are being expressed as a polyprotein.*

We agree that Western blotting would provide the most direct biochemical assessment of P2A-mediated ribosomal skipping efficiency in *Blastocystis* ST7-B and would help determine

the extent of any residual uncleaved fusion product. This is an important point, and we have revised the manuscript accordingly to avoid implying that complete protein-level separation was directly demonstrated.

The current study was designed as a first-generation functional genetic toolkit for *Blastocystis* ST7-B, focused primarily on establishing reproducible workflows for selectable transgene expression, reporter recovery, and propagation of transgenic lines in this experimentally challenging anaerobic microbial eukaryote. The toolkit is therefore validated here through functional outcomes, including antibiotic-selected survival, stable propagation through extended passaging (>15 passages) and cryopreservation, and detectable reporter fluorescence above wild-type autofluorescence.

P2A was selected because it is a compact and well-characterised peptide with high reported ribosomal skipping efficiency across multiple eukaryotic systems, including microbial eukaryotes, as discussed above. Nevertheless, we fully agree that direct biochemical validation would strengthen the mechanistic interpretation of the bicistronic system in *Blastocystis* ST7-B. We therefore now explicitly identify Western blot analysis, ideally using epitope-tagged upstream and downstream products, as an important future direction for quantitative assessment of P2A cleavage efficiency and any residual uncleaved fusion products.

Relevant manuscript revisions are described above under the general response to Reviewer 1.

> *(2) Something has gone wrong with figure formatting. Figure 2 is nearly illegible and I cannot read the text in section A. Sections B, C, and D have lost their labels and are fuzzy and surrounded by black. A similar issue affects Figure 3. Everything is just black with a few cells. It is illegible when printed.*

We thank the reviewer for highlighting these presentation issues and agree that the submitted figure quality significantly impaired readability and interpretation. The problems appear to have arisen primarily during manuscript compilation and export, particularly affecting image resolution, contrast, and panel labelling in the review PDF.

To address this, Figures 2 and 3 have been completely reformatted and replaced with revised high-resolution versions. We have also improved typography, panel separation, colour scaling, and legend clarity throughout. In response to additional reviewer suggestions, individual data points have now been added to Figures 2B and 2C to improve transparency and interpretability of the underlying data distributions.

Figures 2 and 3 have been replaced with fully revised high-resolution versions with improved panel labelling, accessibility, typography, and figure legends.

> *(3) The data from Figure 2B would be better placed in Table 1 with a column for robust/moderate/intermediate/weak/very weak. This would be much easier for the reader.*

We thank the reviewer for this helpful suggestion. We believe the comment refers to the promoter activity data shown in Figure 2A rather than the voltage optimisation data in Figure 2B. To improve readability and accessibility of these data, we have revised Figure 2A extensively to make the promoter activity tiers more legible and easier to interpret directly from the heat map and accompanying box plots.

We considered incorporating simplified activity classifications into Table 1. However, activity patterns were construct-specific rather than simply locus-specific. In several cases, multiple promoter fragments derived from the same locus produced substantially different reporter outputs, and activity did not scale monotonically with promoter fragment length. We

therefore felt that assigning a single categorical activity label at the locus level would oversimplify the dataset and reduce the construct-level resolution that is central to the toolkit value of the study.

Instead, we addressed the reviewer's concern by substantially improving the presentation and readability of Figure 2A, allowing readers to identify robust, moderate, intermediate, weak, and very weak expression constructs more directly while preserving the underlying construct-specific information.

Figure 2A has been revised to improve clarity, accessibility, and legibility of the promoter activity tiers, allowing construct-level expression classes to be interpreted more directly from the heat map and accompanying boxplots.

> *(4) How do you know if the constructs are maintained as episomes? Have you done an outward-facing PCR?*

We agree that direct molecular evidence distinguishing episomal persistence from genomic integration is currently lacking, and we appreciate the reviewer highlighting this important limitation. We have therefore revised the manuscript throughout to avoid presenting episomal maintenance as a demonstrated conclusion and now describe it only as a plausible working interpretation based on the current evidence and vector design.

We have not performed outward-facing PCR in the present study. As discussed in the general response above, we now explicitly identify outward-facing PCR, plasmid rescue followed by full plasmid sequencing, selection-withdrawal assays, FISH, and long-read sequencing approaches as important future directions for resolving the molecular maintenance state of the constructs.

The revised manuscript now clearly separates the demonstrated functional outcomes, including selectable transgene expression, recovery of colony-derived transgenic lines, and stable reporter-positive propagation under selection, from the unresolved mechanistic question of vector topology.

This issue has been addressed throughout the revised manuscript, including in the Methods and Discussion sections, where episomal maintenance is now presented as a plausible but unconfirmed interpretation rather than a demonstrated conclusion.

> *Minor Comments*
>
> *Line 66: is this one to two billion individuals with Blastocystis, or one to two billion Blastocystis cells per gut?*

The intended meaning was colonised individuals globally. We agree that the original phrasing was ambiguous and have corrected it for clarity.

Lines 66–67 revised to: "…microorganisms in the human gut, and is estimated to colonise approximately one to two billion people globally (Scanlan and Stensvold, 2013)."

> *Line 148: Supplier of IMDM?*

The supplier information was already present in the original manuscript as IMDM L0191 (Biowest).

No additional manuscript change required.

> *Line 157: Who annotated the dataset, the 2017 paper or the present study?*

The dataset annotation derives from Armengaud et al. (2017). We agree that the original wording was unclear and have revised this section substantially to improve clarity regarding

the rationale and workflow used for promoter and terminator candidate selection.

"The relevant Methods section has been extensively revised for clarity and expanded detail" (Lines 156–189).

> *Line 166: Who predicted the 3' UTR, the 2017 paper?*

This information derives from the NCBI annotation associated with the *Blastocystis* ST7-B genome based on Denoeud et al. (2011). This has now been clarified in the Methods section.

Clarified in revised Methods section.

> *Line 237: How long did it take in days?*

Approximately 2 days.

Line 270 revised to: "…turned yellow without drug treatment, usually within 2 days post-transfection."

> *Line 325: Typo, missing gap between Figure and 1A.*

Corrected in revised manuscript.

> ***Reviewer #2 (Public review):***
>
> *This manuscript presents a substantial technical advance for the genetic manipulation of Blastocystis by establishing an integrated workflow for stable episomal transgenesis, antibiotic selection, clonal recovery, and reporter-based imaging in the ST7-B subtype. The study is particularly valuable because it combines multiple previously fragmented approaches into a coherent and practically applicable toolkit, including endogenous regulatory elements, optimized electroporation conditions, selectable markers, and anaerobic compatible fluorescent reporters. This methodological work greatly expands the molecular toolbox and future studies focused on both basic and infection biology can now build on the ability to express and localize proteins in fixed as well as live cells.*
>
> *The microscopy data are convincing and clearly demonstrate functional reporter expression and successful recovery of stable transgenic lines. Nevertheless, because this is primarily a methodological paper, the study would be further strengthened by the inclusion of Western blot validation of reporter expression and bicistronic constructs. In particular, biochemical analysis of the P2A-containing constructs would help assess the efficiency of ribosomal skipping and exclude the possible presence of uncleaved fusion proteins, thereby providing stronger support for the interpretation of the imaging data and the functionality of the expression system.*

We thank the reviewer for this thoughtful and positive assessment of the manuscript and for recognising the value of integrating previously fragmented approaches into a coherent and practically usable genetic toolkit for *Blastocystis* ST7-B. We particularly appreciate the reviewer's recognition that the system expands the currently available molecular toolbox for both cell biological and infection-related studies in this experimentally challenging anaerobic microbial eukaryote.

We also appreciate the reviewer's comments regarding biochemical validation of the P2A-containing bicistronic constructs. We agree that Western blot analysis would strengthen the mechanistic interpretation of the reporter system by directly assessing ribosomal skipping efficiency and the possible presence of residual uncleaved fusion products. In response, we have revised the manuscript throughout to ensure that the conclusions remain appropriately evidence-based and do not imply that complete protein-level separation was directly demonstrated.

The revised manuscript now explicitly distinguishes the demonstrated functional outcomes, including selectable transgene expression, stable propagation of reporter-positive lines, and detectable downstream reporter expression, from unresolved mechanistic questions concerning P2A cleavage efficiency and vector maintenance state. We now also identify biochemical validation of P2A-mediated protein separation as an important future direction for further refinement of the system.

Relevant manuscript revisions addressing these points are described above under the response to Reviewer 1.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *The quality of images could be better. The figures lacked resolution — possibly a conversion artefact.*

We agree that the figure quality in the submitted review PDF significantly reduced readability and visual interpretation. The issues appear to have arisen primarily during manuscript compilation and export, particularly affecting image resolution, typography, panel labelling, and contrast rendering.

To address this, Figures 2 and 3 have been completely reformatted and replaced with revised high-resolution versions. We have also improved panel separation, typography, colour scaling, contrast settings, and figure legends to improve accessibility and interpretability both on screen and in print. In addition, the export workflow and file formatting have been updated to improve compatibility with journal production requirements and reduce the likelihood of compression-related rendering artefacts during manuscript compilation.

Figures 2 and 3 have been replaced with revised high-resolution versions with improved typography, panel labelling, contrast settings, and accessibility.

> ***Reviewer #3 (Public review):***
>
> *Summary:*
>
> *The primary objective of this study was to establish a practical and functional framework for the propagation of stable transgenic cell lines of Blastocystis, a common animal gut microeukaryote. Although the work focused on Blastocystis ST7-B, a subtype with relatively low prevalence in humans, this choice is justified by its association with more frequent negative health effects. Beyond their relevance to the medical field, the methodological advances described here have the potential to also expand cell biology studies of this anaerobic organism, including its unusual mitochondria and redox metabolism.*
>
> *Strengths:*
>
> *Prior to this work, genetic tools for Blastocystis were very limited, relying on a single strong promoter-terminator combination. The authors successfully expanded the available promoter set across a range of expression strengths by testing two dozen variants in luciferase-based assays. Critically, they developed an integrated workflow from a modular transgenic construct design, to an expanded inventory of molecular components (promoters, reporters), optimized DNA delivery, stepwise antibiotic resistance-mediated clonal selection and propagation, and to reporter validation. The evaluation of several anaerobiosis-compatible labeling strategies for live (and fixed) cell optical imaging will be particularly useful, with the SNAP-tag system appearing especially promising for Blastocystis.*
>
> *Weaknesses:*

> *The presented data generally provide solid support for the conclusions that the work reached, but clarification of reasoning and several inconsistencies, as well as amendments to the visual presentation of the data, would be highly beneficial, as detailed below.*
>
> *(1) Episomal persistence of the construct:*
>
> *The manuscript repeatedly assumes, including in its title, that constructs persist in Blastocystis in their episomal form, but no direct evidence is provided. Although this interpretation is plausible, it should be identified more clearly as provisional. Nuclear genomic integration (e.g., via NHEJ) remains a possible explanation unless supporting evidence or rationale is provided to exclude it. Testing whether the phenotype persists without drug-mediated selection in the generated transgenic cell lines would help strengthen the case for episomal maintenance.*

We thank the reviewer for this important point and agree that the original manuscript presented episomal persistence too strongly relative to the currently available evidence. In particular, we agree that the title and several sections of the manuscript implied a level of mechanistic certainty that was not directly demonstrated.

We have therefore revised the manuscript throughout to clarify that episomal maintenance should presently be regarded as a plausible working interpretation rather than a demonstrated conclusion. The revised text now explicitly distinguishes the demonstrated functional outcomes, including selectable transgene expression, recovery and propagation of colony-derived transgenic lines, and stable reporter-positive maintenance under selection, from the unresolved mechanistic question of vector topology.

As discussed in our response to Reviewer 1, the constructs used here do not contain *Blastocystis* homology arms, and comparative genomic analyses suggest that *Blastocystis* lacks canonical non-homologous end-joining components, making targeted integration by standard repair routes less strongly supported mechanistically, although genomic integration cannot presently be excluded.

We agree that selection-withdrawal experiments would provide useful additional evidence regarding construct persistence and have now explicitly identified such assays, together with outward-facing PCR, plasmid rescue, FISH, and long-read sequencing approaches, as important future directions for resolving the molecular maintenance state of the transgenes.

The manuscript has been revised throughout to remove wording implying demonstrated episomal maintenance. Episomal persistence is now discussed only as a plausible working interpretation pending direct molecular validation.

> *(2) Promoters and terminators:*
>
> *(2.1) There is a discrepancy between the claimed number of loci (14), from which promoters used to drive luciferase expression were derived, and those detailed as having been actually generated in Table 1 (11). This inconsistency should be corrected or explained, as it creates uncertainty around the accuracy of the dataset.*

We thank the reviewer for this careful reading and for identifying this inconsistency. We agree that the distinction between candidate loci and successfully generated promoter constructs was not sufficiently clear in the original manuscript and could create uncertainty regarding the dataset.

The original candidate set comprised 14 loci selected for promoter and terminator discovery. However, only 11 loci yielded successfully cloned and experimentally tested promoter constructs. The remaining three loci were retained in Table 1 for completeness and

transparency, as repeated cloning attempts were unsuccessful despite two independent efforts.

We have revised the manuscript to make this distinction explicit and to clarify that the reported NanoLuc benchmarking experiments were ultimately performed using constructs derived from 11 successfully cloned loci.

Lines 361–364: "To expand the available regulatory parts, we screened 23 NanoLuc reporter constructs containing putative endogenous promoter–terminator pairs from 11 of 14 candidate loci; three loci could not be cloned after two independent attempts and are indicated in Table 1."

> *(2.2) Based on the presented evidence, constructs benchmarked in bioluminescence assays differed only in their promoter composition. Although terminator selection is mentioned in the Methods section, no additional details are provided; for instance, Table 1 and Figure 2 only list 23 promoters in total. Figure 2A likewise shows only promoter-dependent variation. If the terminator was held constant (LeguP1?), this should be stated explicitly. The authors may then consider revising the wording of having tested "23 promoter-terminator pairs" to better reflect that only promoters varied.*

We thank the reviewer for the opportunity to clarify this point. We agree that the original presentation may have created the impression that promoter and terminator regions were independently varied and benchmarked, whereas the experimental design was primarily focused on construct-level comparison of endogenous regulatory modules.

As described in the Methods, each construct contained a candidate endogenous upstream promoter region together with the corresponding endogenous downstream terminator region derived from the same locus. For consistency and to keep the cloning and screening strategy experimentally tractable, a fixed 500 bp downstream terminator fragment was used for each locus rather than systematically varying terminator length or independently testing terminator activity.

We therefore retain the description "endogenous promoter–terminator pairs," since each construct contains both endogenous upstream and downstream regulatory regions from the same genomic locus. However, we agree that the assay was not designed to independently dissect promoter versus terminator contributions to reporter output. We have revised the manuscript accordingly to make this distinction explicit and avoid ambiguity regarding the scope of the benchmarking analysis.

Lines 365–368: "Each construct paired a candidate upstream promoter region with the corresponding downstream terminator region from the same locus, defined here as the native 500 bp sequence immediately downstream of the stop codon. Where multiple promoter lengths were tested for the same locus, the terminator fragment was kept constant (Table 1; Figure 1A)."

This design allowed construct-level benchmarking of paired promoter–terminator modules but did not test promoter strength or terminator activity independently.

> *(2.3) Promoter benchmarking was done with a plasmid lacking a selection marker, so it is unclear how the maintenance of the luciferase construct was ensured. Without selection, the observed reporter intensity could reflect differential or stochastic plasmid retention rather than promoter strength alone. The luminescence assay was performed 16-18 hours after transfection, but the rationale for this particular timeframe should be explained. In this context, the authors should explicitly state whether the experiments shown in Fig.2A represent biological triplicates or technical triplicates from a single transfection.*

We thank the reviewer for these important methodological points. We agree that the original manuscript did not sufficiently clarify the transient nature of the NanoLuc benchmarking assay or the rationale underlying the assay design and timing.

The promoter benchmarking assay was designed as an early transient-expression screen adapted from the NanoLuc-based workflow of Li et al. (2019), with modifications, rather than as a stable-maintenance assay. No selectable marker was included because the objective was to compare relative early reporter output across constructs shortly after DNA delivery, before prolonged culture effects became dominant.

The 16–18 h post-electroporation time point was selected based on the NanoLuc expression kinetics reported by Li et al. (2019) and empirical optimisation during assay development. This window allowed robust transient reporter detection while limiting confounding effects arising from prolonged plasmid loss, differential outgrowth, variable recovery, or later culture-level changes.

We agree that, in the absence of selection, the observed NanoLuc signal cannot be interpreted as an absolute measure of promoter strength independent of DNA uptake efficiency, early plasmid retention, or post-transfection recovery dynamics. We have therefore revised the manuscript to clarify that Figure 2A reports relative transient reporter output under standardized early post-transfection conditions rather than isolated promoter activity alone.

We now also explicitly state that the data shown in Figure 2A derive from three independent electroporation experiments per construct, each assayed in technical duplicate.

Lines 241–248: “Promoter–terminator activity was assessed 16–18 h after electroporation using a transient NanoLuc assay adapted from Li et al. (2019), with modifications. This early time point was selected to capture reporter output within the transient-expression window after DNA delivery, before prolonged plasmid loss, differential outgrowth, or culture-level changes could dominate the readout. Because the constructs did not contain a selectable marker, the measured NanoLuc signal reflects early transient reporter output rather than promoter strength independent of DNA uptake, early plasmid retention, or post-transfection recovery.”

Additional clarification added to Figure 2 legend stating that measurements derive from three independent electroporation experiments, each assayed in technical duplicate.

> *(3) Figure 2:*
>
> *(3.1) Several aspects of the current design may lead to ambiguity for the reader. The boxplots are colour-coded, but it is unclear whether the colours carry meaning or are purely decorative. Because the data are already spatially separated into bins, additional random colouring is redundant and may suggest distinctions that are not intended. In addition, part A of Figure 2 is split into two panels, with the scale for the left panel shown in the right panel and some of the boxplot colours falling in the range of the scale, but not in line with their counterparts in the left panel. Because the colour use is not consistent, it is difficult to tell whether the same scale should be applied to both panels or how it should be interpreted.*
>
> *(3.2) The left panel of part A uses a diverging blue-white-red colour scheme, which is most appropriate when the midpoint represents a meaningful central value such as zero. Because the values shown in this graph are only positive, a non-diverging 2-colour scale or a colour palette such as 'viridis' would make the plot easier to interpret.*
>
> *(3.3) A black background should be avoided: 'B' and 'C' labels are invisible, and it draws attention to a distracting design feature rather than the data themselves.*

We thank the reviewer for these detailed comments regarding figure design and visual interpretation. We agree that the original presentation of Figure 2 introduced unnecessary visual ambiguity through inconsistent colour usage, the use of a diverging colour scale for strictly positive values, and poor readability associated with the dark background and low-resolution export.

In response, Figure 2 has been extensively redesigned to improve clarity, accessibility, and interpretability. The previous blue–white–red diverging heatmap has been replaced with a sequential colour palette appropriate for positive-only expression data. Boxplot colouring has also been simplified and harmonised with the heatmap scheme to avoid implying unsupported categorical distinctions. In addition, panel organisation, typography, scaling, and legend structure have all been revised to improve readability and reduce ambiguity regarding interpretation of the plotted values.

We also agree that the black background distracted from the data presentation and impaired visibility of panel labels and image boundaries. The revised figures therefore use white backgrounds together with clearer panel separation and improved label visibility throughout.

Figure 2 has been completely reformatted using a sequential colour scale in panel A, simplified and harmonised boxplot colouring, larger typography, improved panel separation, revised legends, and white backgrounds throughout. Corrected high-resolution source figures have been provided.

> *(4) Figure 3:*
>
> *(4.1) Individual snapshots should be separated more clearly, either by using a white background or by adding visible borders to make the overall composition clearer. As currently displayed, some boundaries between fluorescent channels resemble image artifacts rather than intentional panel divisions.*

We thank the reviewer for this helpful comment regarding figure composition and panel separation. We agree that the original presentation made it difficult to distinguish intentional panel boundaries from imaging artefacts, particularly in the low-resolution review PDF generated during manuscript compilation.

To improve clarity, Figure 3 has been reformatted using white backgrounds, clearer panel spacing, and more explicit separation between individual snapshots and imaging channels. High-resolution source images have also been provided to ensure that fluorescence patterns, image boundaries, and panel organisation remain clearly interpretable both on screen and in print.

Figure 3 has been reformatted with improved panel separation, white backgrounds, clearer image boundaries, and revised high-resolution source figures.

> *(4.2) In parts B-D, the legend should explain more clearly what each image shows, and the figure itself would benefit from annotations. There seem to be three sub-panels in each 'condition' of part B (as well as C and D): while the middle and rightmost panel can be easily inferred to represent the fluorescent protein and bright-field image, what the leftmost panels represent is not specified. If DAPI was used to dye DNA, an explanation why mostly multiple labelled regions are visible should be provided.*

We thank the reviewer for these helpful suggestions regarding figure annotation and legend clarity. We agree that the original presentation did not sufficiently explain the composition of the imaging panels, particularly under the low-resolution conditions of the review PDF.

To improve interpretability, the revised Figure 3 now includes clearer panel organisation, improved annotations, and expanded figure legends explicitly identifying the individual

imaging channels and staining conditions shown in each subpanel. The leftmost panels in parts B–D are now more clearly identified in both the figure and legend, together with the corresponding fluorescence or staining conditions used in each experiment.

As mentioned in the Methods sections we used Hoechst 33342 to visualise DNA; but we agree that the Hoechst 33342-labelled structures required additional clarification. The revised legend section now explains that multiple Hoechst 33342-positive regions are commonly observed because *Blastocystis* cells can contain multiple nuclei depending on cell stage and subtype-specific morphology.

In addition, high-resolution source images have been provided to ensure that fluorescent signals, panel boundaries, and imaging features remain clearly interpretable both on screen and in print.

Figure 3 legends and annotations have been revised to clarify imaging channels, staining conditions, and panel organisation. The figure caption was also edited to include: “DNA was visualised using Hoechst 33342. Most cells contained two nuclei, and smaller Hoechst 33342-positive signals consistent with mitochondrial DNA were also observed in some instances.”

> *(4.3) Cell morphology and appearance differ markedly between UnaG/smURFP and SNAP-tag images, which should be explained. A microscope issue is mentioned in the main text, but if that was the cause, the authors should consider replacing the images, as the current distortions complicate interpretation.*

We thank the reviewer for this important observation and agree that the apparent morphological differences between the UnaG/smURFP and SNAP-tag panels required additional clarification.

The images shown for the different reporter systems were acquired under different imaging conditions and microscope configurations following an instrument-related issue during part of the imaging workflow, as noted in the Methods section. As a result, direct visual comparison of cell morphology between reporter systems is not appropriate. The primary purpose of these panels is instead to demonstrate reporter detectability, live-cell labelling capability, and the characteristic fluorescence patterns obtained with the different anaerobiosis-compatible reporter systems.

In particular, the SNAP-tag panels were included to demonstrate successful live-cell labelling without permeabilisation together with the expected increase in fluorescence signal at higher substrate concentrations, rather than to support quantitative comparison of cell morphology across imaging conditions.

We considered replacing the affected images. However, equivalent replacement datasets acquired under directly comparable conditions are not currently available. We have therefore retained the original images but revised the figure legend to clarify the intended interpretation and limitations of these panels explicitly.

Figure 3 legend revised to include:

“Because images for the different reporter systems were acquired under different imaging conditions, they are presented to demonstrate reporter detectability and labelling pattern and should not be used for quantitative comparison of cell morphology across reporter systems.”

> ***Reviewer #3 (Recommendations for the authors):***

> *The reader may find the current order confusing starting with construct design before testing which drug to use for selection. The narrative would work better if it started with antibiotic selection as the first logical step for generating stable cell lines.*

We thank the reviewer for this thoughtful suggestion regarding narrative structure and agree that multiple organisational strategies are possible for presenting a methodological workflow of this type.

We considered reorganising the Results section to begin with antibiotic selection and drug sensitivity profiling. However, we ultimately retained the overall structure because the manuscript is organised as a toolkit-development framework rather than as a strictly chronological experimental protocol. The Results therefore begin with regulatory-element discovery and construct design, which form the conceptual and experimental foundation of the toolkit, before progressing to DNA delivery optimisation, drug sensitivity profiling, clonal recovery, and reporter validation.

We felt that this structure most clearly reflects the dependency relationships within the system: regulatory elements are required before constructs can be assembled, constructs are required before electroporation conditions can be evaluated, and selectable constructs are required before stable selection and clonal recovery can be meaningfully assessed.

> *(2) The text states that the screen 'focused on the 1,000 most abundant proteins to establish a preliminary library capable of supporting varying levels of transcription.' Since the genome has ~6,000 protein-coding genes, the top 1,000 cover the most abundant proteins — not a wide expression range.*

We thank the reviewer for this important clarification. We agree that the original wording could incorrectly imply that the screen was intended to sample broadly across the full transcriptional range of the *Blastocystis* genome. This was not the case, and we have revised the manuscript accordingly.

Our strategy was instead designed to enrich for candidate loci with a higher prior likelihood of supporting detectable transgene expression. Because no genome-wide promoter map, transcription start site dataset, or experimentally validated regulatory annotation was available for *Blastocystis* ST7-B at the inception of this work, we used the abundance-ranked *Blastocystis* ST4-WR1 proteomic dataset of Armengaud et al. (2017) as a practical starting point for candidate discovery.

Importantly, the *Blastocystis* ST4-WR1 proteome is highly skewed, with 193 proteins contributing approximately 50% of the detected proteome and the 13 most abundant proteins contributing approximately 10% (Armengaud et al., 2017). We therefore selected the top 1,000 proteins not as a representation of the genome-wide expression range, but as a proteomics-guided enrichment strategy to identify loci more likely to contain active endogenous regulatory regions suitable for initial toolkit development.

We have revised the relevant Methods section substantially to clarify both the rationale and the workflow used for candidate selection, homolog identification, and promoter/terminator definition.

The Methods section (Lines 156–189) has been extensively revised to clarify the rationale underlying candidate regulatory-element selection. The revised text now explicitly states that the strategy was designed to enrich for likely active loci for toolkit development rather than to systematically survey the full range of promoter strengths across the *Blastocystis* genome.

Additional methodological detail has also been added regarding:

Use of the Armengaud et al. (2017) proteomic and proteogenomic datasets,

Homolog identification in *Blastocystis* ST7-B,

Locus selection criteria,

Promoter boundary definition,

And operational definition of candidate terminator regions.

> *(3) The Methods contain an inconsistency: cells were left in 0.5 mL, then 1 mL was added, but then only 0.5 mL is apparently used for transfection. What happened to the 1 mL?*

We thank the reviewer for identifying this ambiguity in the transfection workflow description. The apparent inconsistency arose because the protocol description moved from bulk cell resuspension to preparation of individual electroporation reactions without explicitly stating how the intermediate suspension was used.

After washing, approximately 0.5 mL of cytomix buffer remained above the pellet, and 1 mL of complete cytomix buffer was then added to generate an approximately 1.5 mL cell suspension. Cells were counted from this pooled suspension, after which the volume corresponding to 5×10^7 cells was transferred into each individual electroporation reaction. Following addition of DNA, each electroporation reaction was adjusted to a final volume of 500 µL with complete cytomix buffer. The remaining cell suspension was retained for additional transfections or control reactions.

We agree that the original wording could be misinterpreted and have revised the Methods section to clarify the sequential handling steps more explicitly.

Lines 225-229 revised to read: “The resulting approximately 1.5 mL pooled cell suspension was used for total viable cell counting using a hemacytometer.”

“After counting, the volume corresponding to 5×10^7 cells was transferred to each electroporation reaction and combined with 25 µg of plasmid DNA. The total electroporation volume was adjusted to 500 µL with complete cytomix buffer.”

> *(4) Figures 2 and 3 are too low-resolution for the font size used and for clearly viewing the microscopy images.*

We thank the reviewer for highlighting these readability issues. As noted in our responses above regarding Figures 2 and 3, the low-resolution appearance primarily resulted from manuscript compilation and PDF export artefacts affecting typography, image rendering, and panel clarity in the review version.

To address this, Figures 2 and 3 have been completely reformatted and replaced with revised high-resolution versions featuring improved typography, panel labelling, contrast, accessibility, and image clarity for both on-screen viewing and print reproduction.

Revised high-resolution versions of Figures 2 and 3 have been provided as described above. No additional manuscript changes were required beyond the figure revisions already outlined.

> *(5) Figure 4 is confusing because the left and right panels appear inconsistent, with much higher concentrations required for growth inhibition in the culture-based assay than the resazurin assay indicated. The rationale for the resazurin assay should be explained, and the complete growth inhibition (CGI) concentration should be highlighted in the right panel.*

We thank the reviewer for highlighting this potential source of confusion. We agree that the distinction between the two assay endpoints was not sufficiently emphasised in the original figure presentation and legend.

The apparent discrepancy arises because the two assays measure different biological endpoints under different assay conditions. The resazurin assay was used to estimate IC_{50} values, corresponding to the concentration at which metabolic activity was reduced by approximately 50% under the assay conditions. In contrast, the small-culture assay was designed to determine complete growth inhibition (CGI), defined operationally as the concentration at which no detectable culture outgrowth occurred after incubation, using phenol red acidification as a culture-level readout.

Because these assays measure partial metabolic inhibition versus complete suppression of detectable culture outgrowth, the corresponding concentration ranges are not expected to coincide directly. The higher concentrations observed in the right-hand panels therefore reflect the more stringent endpoint associated with complete growth inhibition rather than inconsistency between the assays.

We agree that this distinction should have been explained more clearly in the original manuscript. We have therefore substantially revised the Figure 4 legend to clarify the rationale underlying both assays, explicitly distinguish IC_{50} and CGI endpoints, and explain how the CGI values were used to guide subsequent antibiotic selection conditions for *Blastocystis* ST7-B transformants. The CGI transition range has also been made more visually explicit in the revised figure presentation.

Figure 4 caption revised to: “Antibiotic potency and selection-window determination in *Blastocystis* ST7-B. Dose–response curves for puromycin, trimethoprim, and WR99210 were estimated from a resazurin-based viability assay (n = 3 independent replicates per drug per concentration). Points show mean ± SD, and the insets list the estimated IC50 values with R^2-values > 0.75 for all fitted curves. The IC_{50} estimates represent the drug concentrations that reduced resazurin-based metabolic activity by 50% under the assay conditions.”

Right panels: “small-culture complete growth inhibition assay using 1×10^7 WT *Blastocystis* ST7-B cells per culture, assayed in triplicate across a wide range of concentrations. Cultures were incubated for 2 days, and outgrowth was assessed using phenol red acidification of the medium as a culture-level readout, with yellow indicating growth and red indicating no detectable growth. The yellow-to-red transition was used to estimate the concentration required for complete growth inhibition and to guide the subsequent antibiotic selection strategy for *Blastocystis* ST7-B transformants.”

“IC_{50} and CGI represent distinct assay endpoints: the former measures partial reduction in metabolic activity, whereas the latter identifies the concentration at which no detectable culture outgrowth occurs under the small-culture assay conditions.”

> *(6) In Figure 3B, the unexpected UnaG fluorescence pattern could be due to protein sequestration because the protein is mildly toxic to the cell. This should be discussed in addition to the reasons already provided.*

We thank the reviewer for this thoughtful suggestion and agree that protein sequestration or reporter-associated cellular stress represent plausible alternative interpretations of the observed UnaG fluorescence pattern.

We considered the possibility of UnaG-associated toxicity during interpretation of these data. However, under the conditions tested, we did not observe clear evidence of a substantial toxic effect: UnaG-expressing *Blastocystis* ST7-B cells could be recovered as stable lines, maintained under antibiotic selection, and propagated through continued culture. We

therefore felt that direct attribution of the observed fluorescence pattern to reporter toxicity would currently remain speculative.

At present, we consider the biochemical properties of the UnaG system itself to provide a more parsimonious explanation for the observed localisation pattern. In particular, unconjugated bilirubin is highly hydrophobic and would be expected to partition preferentially into lipid-rich cellular environments. This interpretation is consistent with the lipid-rich peripheral and intracellular structures previously reported in *Blastocystis* ST7-B (Liao et al., 2023).

We have therefore revised the Discussion to acknowledge that the observed UnaG fluorescence pattern may reflect a combination of reporter-specific biochemical behaviour, bilirubin partitioning, local intracellular environment, or possible sequestration phenomena. At the same time, we avoid assigning toxicity as a demonstrated mechanism in the absence of direct measurements of cell fitness, reporter abundance, or bilirubin distribution. Such experiments would be required to evaluate this possibility rigorously.

Lines 642-648: “Consistent with this, lipid-rich peripheral and intracellular structures have been reported in *Blastocystis* ST7-B, potentially providing favourable microenvironments for BR partitioning and contributing to the punctate UnaG fluorescence pattern (Liao et al., 2023). An alternative possibility is that the observed signal pattern reflects reporter sequestration or reporter-associated cellular stress. However, because UnaG-expressing lines were recovered, maintained under selection, and propagated through continued culture, toxicity remains a possible but untested explanation rather than a demonstrated mechanism.”

> *Minor Comments*
>
> *Figure 2: Parts B and C should also show individual datapoints for better reader assessment.*

We agree that inclusion of individual data points improves transparency and interpretability of the underlying data distributions.

Individual data points have now been overlaid on the boxplots in Figures 2B and 2C.

> *Figure 3A: Separate channels (fluorescence, bright-field, merge) should be shown rather than only the merge. The current overlay is difficult to interpret, especially for colour-blind readers.*

We appreciate the reviewer’s concern regarding accessibility and interpretability. We considered separating the fluorescence, bright-field, and merged channels for Figure 3A. However, this panel was intended primarily as an overview demonstrating reporter detectability within the bicistronic construct context, while the detailed fluorescence distribution is explored more extensively in the subsequent UnaG panels. We therefore retained the merged presentation for Figure 3A. Importantly, the image is not dependent on red–green discrimination, as it combines a greyscale bright-field background with a high-contrast green/cyan fluorescence signal that remains distinguishable through brightness and contrast differences. In addition, colour-blind-friendly lookup tables (LUTs) were used throughout the revised figure set.

To further improve accessibility, the original red annotation arrow has been replaced with a colour-blind-friendly annotation colour.

> *Briefly define system components (P2A, UnaG, smURFP, SNAP-tag) and add an abbreviation list.*

We agree that brief contextual definitions improve accessibility for readers less familiar with these reporter systems. Rather than adding a separate abbreviation list, we have added short

explanatory descriptions at the points where these components are first introduced in the manuscript.

Lines 394–396: “The P2A peptide is expected to promote ribosomal skipping during translation, allowing two separate polypeptides to be produced from a single open reading frame.” Line 515–516: “UnaG, a bilirubin-binding fluorescent protein originally isolated from the muscle of the Japanese eel (Kumagai et al., 2013)…” Line 527: “smURFP (small ultra-red fluorescent protein)…”

> *Abstract: “among the most prevalent microbial eukaryote” should be “eukaryotes”.*

Corrected in revised manuscript.

> *Conclusion (2nd sentence): unclear what “endogenous regulatory part discovery” means.*

We agree that this phrase required clarification. The intended meaning was the identification and benchmarking of native *Blastocystis* ST7-B promoter and terminator elements for construct design and toolkit development. We have clarified this directly in the revised Conclusion section.

Lines 682–683 revised to: “By bringing endogenous regulatory part discovery, namely the identification of native promoter and terminator elements, …”

> *Author contributions: “critical advise” should be “advice”.*

Corrected in revised manuscript.

Again, we thank the reviewers for their careful evaluation, constructive criticism, and thoughtful feedback on the manuscript. The review process has substantially strengthened the manuscript by helping us clarify the distinction between what is directly demonstrated experimentally and what remains mechanistically unresolved.

The central methodological conclusions of the study remain unchanged: the toolkit enables selectable transgene expression, recovery of colony-derived lines, and propagation of reporter-positive transgenic *Blastocystis* ST7-B lines, extending genetic accessibility in this organism substantially beyond the previous transient transfection framework.

At the same time, the revised manuscript now more explicitly acknowledges important unresolved mechanistic questions, including vector topology, P2A-mediated protein separation efficiency, and persistence in the absence of selection. These are now discussed transparently together with the future experimental approaches that will be required to address them directly.

We believe the revised manuscript now presents a clearer, more rigorous, and more accessible description of a practical genetic toolkit for *Blastocystis* ST7-B and hope that the revisions and clarifications satisfactorily address the reviewers’ concerns.

https://doi.org/10.7554/eLife.111816.2.sa0