A genetic engineering toolbox for the lignocellulolytic anaerobic gut fungus *Neocallimastix frontalis*

Casey A. Hooker[1,2], Radwa Hanafy[2], Ethan T. Hillman[1†], Javier Muñoz[1,3], and Kevin V. Solomon[1,2*]

[1]Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, IN, 47907 USA

[2]Department of Chemical and Biomolecular Engineering, University of Delaware, Newark, DE, 19716 USA

[3]Department of Biomedical Engineering, Purdue University, West Lafayette, IN, 47907 USA

[†]Present Address: Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, 48109 USA

*****Correspondence: kvs@udel.edu

Abstract

Anaerobic fungi are powerful platforms for biotechnology that remain unexploited due to a lack of genetic tools. These gut fungi encode the largest number of lignocellulolytic carbohydrate active enzymes (CAZymes) in the fungal kingdom, making them attractive for applications in renewable energy and sustainability. However, efforts to genetically modify anaerobic fungi have remained limited due to inefficient methods for DNA uptake and a lack of characterized genetic parts. We demonstrate that anaerobic fungi are naturally competent for DNA and leverage this to develop a nascent genetic toolbox informed by recently acquired genomes for transient transformation of anaerobic fungi. We validate multiple selectable markers (HygR and Neo), an anaerobic reporter protein (iRFP702), enolase and TEF1A promoters, TEF1A terminator, and a nuclear localization tag for protein compartmentalization. This work establishes novel methods to reliably transform the anaerobic fungus *Neocallimastix frontalis*, thereby paving the way for strain development and various synthetic biology applications.

Keywords: anaerobic gut fungi, transformation, parts, protein localization, gene expression, anaerobic reporter

Introduction

Anaerobic gut fungi (AGF; phylum Neocallimastigomycota), native to the digestive tracts of ruminant and hindgut animals, are primary colonizers of plant biomass that readily degrade crude lignocellulose.[1,2] Gut fungi thrive on a variety of complex plant materials ranging from crude C3 and C4 grasses, which differ in how carbon is fixed during photosynthesis, to untreated agricultural and forestry residues.[3] The ability of gut fungi to efficiently hydrolyze these feedstocks is proposed to be driven by the high diversity and abundance of carbohydrate active enzymes (CAZymes) encoded by their genomes.[4,5] The plethora of CAZymes, combined with the use of extracellular multienzyme lignocellulose-processing complexes or cellulosomes and the physical penetration of lignocellulosic biomass contribute to high lignocellulolytic activity.[5] Beyond applications in renewable energy technologies, anaerobic fungi may also be leveraged to identify new natural products as well as intermediates for drop-in biofuels. Multiple analyses of sequenced gut fungal genomes reveal an abundance of biosynthetic gene clusters encoding genes that produce uncharacterized natural products.[6–8] Together, anaerobic fungi offer a wealth of biotechnological potential that is only beginning to be realized.[9]

Efforts to characterize and tap into this biotechnological potential have revealed many proteins from anaerobic fungi that offer superior properties that are desirable for industrial processes.[10–12] These industrially relevant properties include thermostability, resistance to protease degradation, and preferential cofactor requirements, among others.[10,13] Additional efforts reveal signaling and/or transport proteins from gut fungi that may be leveraged as orthogonal systems in model organisms.[14–17] However, heterologous expression of many anaerobic fungal proteins remains challenging due to: 1) differences in host reducing conditions, which inhibits proper folding[18]; 2) a compositional bias towards amino acids encoded by AT-rich codons[13], which can present a metabolic burden that is not fully addressed by codon optimization[13,19], and 3) glycosylation and other post-translational modifications required for function that are not easily replicated in model hosts.[20] Consequently, developing genetic tools to manipulate AGF proteins in their native context will enable the discovery of novel anaerobic fungal genes and metabolic products of biotechnological importance.

Developing genetic tools and techniques to domesticate non-model organisms is an attractive strategy for building robust microbial cell factories.[21,22] These genetic toolkits at a minimum require efficient transformation schemes for introduction of nucleic acids, selectable markers and/or reporters to help validate transformation, and identification of constitutive promoters to drive expression of these genes[22]. With this preliminary set of genetic tools, additional parts such as terminator sequences, localization tags, replication sequences[23], and centromeric sequences[24] can be identified to enable stable expression, replication, and partitioning of genetic constructs *in vivo*.[25–27] Development of proper genetic tools enabling strain engineering efforts in anaerobic fungi has been hampered due to the fragmented nature of currently available genomes.[4,28,29] This high fragmentation is caused by shearing of the genome during lysis of the thick, chitin rich cell walls[28], extremely high genomic AT content that is challenging to assemble[4,30], and large homopolymeric sequences that impede genome assembly and thus annotation.[4,30] Beyond genome assembly and annotation, additional challenges are introduced by the complex lifestyle and life cycle of gut

fungi.[2,31] Anaerobic fungal life cycles begin with a motile flagellated zoospore stage that subsequently attaches and encysts into plant biomass to initiate a non-motile phase characterized by sporangia and rhizomycelia.[2] This complex alternating lifecycle coupled with an obligate anaerobic nature hinder the use of traditional transformation methods such as *Agrobacterium tumefaciens* mediated transformation [32,33] (ATMT), and protoplast-mediated transformation (PT), that are commonly used in aerobic fungi. ATMT does not occur above 29 °C, well below the growth temperature of anaerobic fungi at 39 °C.[32] Moreover, the thick chitin and glucan rich cell walls of anaerobic fungi significantly impede the efficacy of lytic enzymes commonly used in protoplast-mediated transformations. Despite these challenges, they are not insurmountable and there have been limited reports of transient anaerobic fungal transformation.

The feasibility of anaerobic fungal nucleic acid transformation was first demonstrated with harsh biolistic bombardment of fungal biomass to introduce and transiently express a β-glucuronidase (GUS) reporter gene.[34] While promising results were achieved, this delivery method resulted in high rates of cell mortality and low transformation efficiencies (~4 transformants/μg DNA), and relied on isolation of fungal zoospores in preparing cultures for transformation. Thus, gentler, higher throughput methods are needed to reliably transform and domesticate anaerobic fungi. More recently, it was demonstrated that germinating zoospores naturally uptake foreign, short RNA sequences (~21 bases).[35] This natural competency was exploited to deliver small interfering RNA (siRNA) targeting the D-lactate dehydrogenase (*ldh*) gene, significantly reducing gene expression and thus production of D-lactate. While delivery of siRNAs via natural competency enabled higher cell survival rates, this protocol is limited to AGF species that can produce a high concentration of zoospores. However, zoospore quantity is greatly affected by culture age, anaerobic fungal species, and cultivation conditions. Thus, more efficient transformation strategies, particularly those that enable uptake of larger nucleic acids and those that are not limited to a specific lifecycle stage are greatly needed to improve transformation of gut fungi

In this work we establish transformation of anaerobic fungi with deoxyribonucleic acids (DNA) that exploits the natural competency of their cells at multiple lifecycle stages. This approach removes the need to enrich and selectively transform fungal zoospores enabling higher-throughput transformation methods for this unique phylum of fungi. Using a novel AGF isolate from giraffe, *Neocallimastix frontalis* var *giraffae* (Gf-ma), as a model system, we leverage our transformation protocol to evaluate the functionality of two promoter sequences, multiple reporter constructs, and several selectable markers. We demonstrate that coupling these selectable markers to anaerobic fluorophores enables screening of promoter activity and monitoring of protein expression profiles *in vivo*. We also develop flow cytometry protocols to screen and quantify transformation efficiency of these genetic constructs in fungal cells. Lastly, we have identified and verified functionality of a nuclear localization signal sequence to enable targeting of proteins to anaerobic fungal nuclei, which will be essential for future strain development efforts via gene editing technologies. The methods and genetic tools developed here provide a framework for future synthetic biology and functional genomic studies to leverage anaerobic fungi for their diverse biotechnological potential.

Results

Anaerobic fungi naturally take up DNA

As the first step for transformation, we needed to establish robust and efficient methods for DNA uptake. Prior biolistic methods were low efficiency with high cell mortality[34], while natural competency based methods were only established for short RNA sequences.[35] We first assessed whether anaerobic fungi were naturally competent for DNA and whether this ability was independent of life cycle stage (motile and vegetative). We generated short 50 nt dsDNA probes via primer annealing that we subsequently labeled with Cy3 fluorescent dye and mixed with young cultures (~18 h) of Gf-ma. Cy3 labeled DNA was rapidly taken up by multiple life stages of anaerobic fungi and observed in motile zoospores, vegetative sporangia, sporangiophores, and to a limited extent in rhizomycelia immediately after mixing (Figure 1). Within 60 minutes the Cy3-labeled DNA localized to the nucleus, suggesting that natural competence is a potential delivery vehicle for heterologous gene expression. More importantly, however, these results establish that anaerobic fungi are naturally competent for DNA at all lifecycle stages, obviating the need for time-consuming enrichment of zoospores as in previous methods.[34,35]

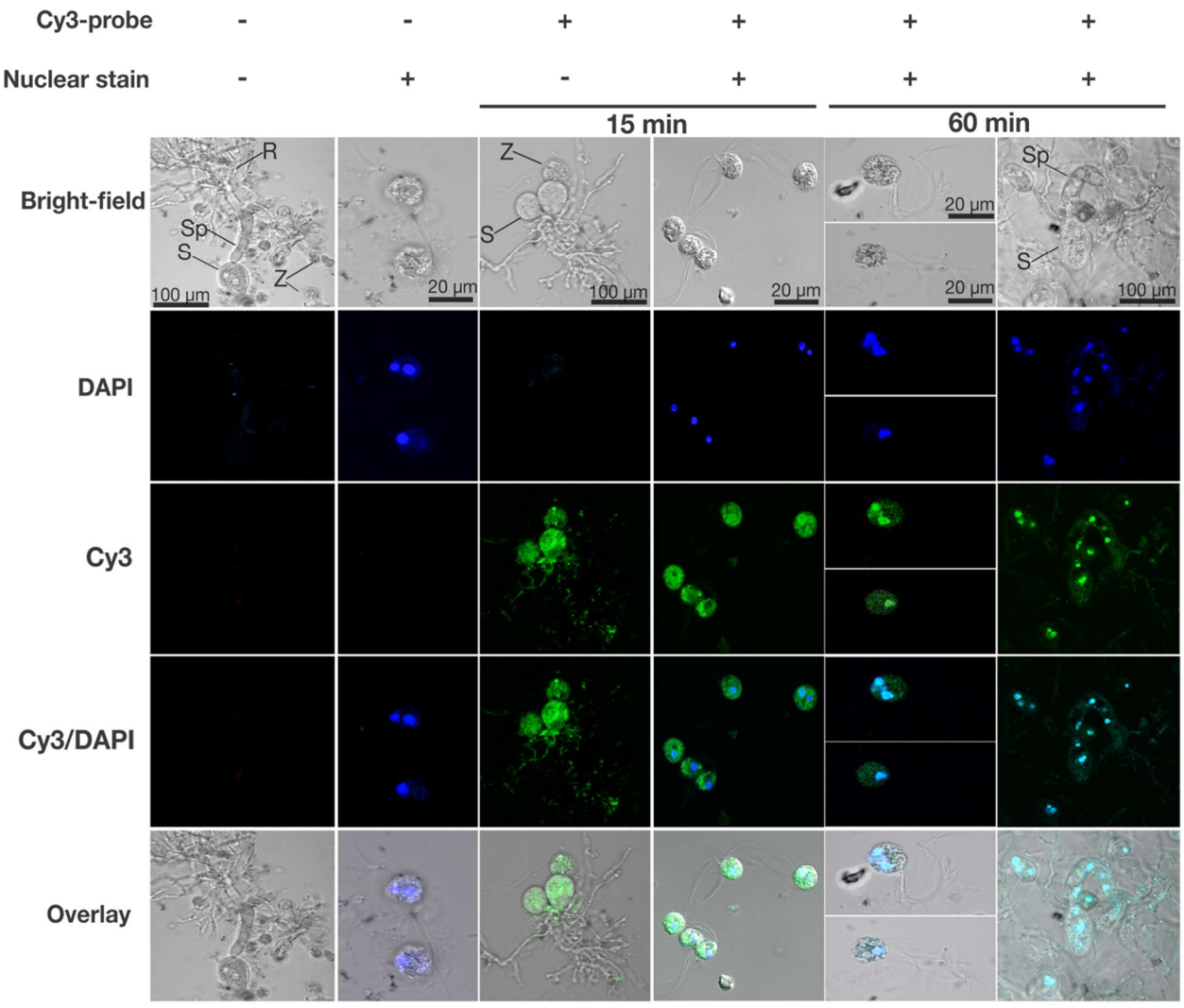


Figure 1: *Neocallimastix frontalis* var *giraffae* is competent for DNA, which localizes to the nucleus
Bright-field micrographs of fungal rhizomycelia (R), sporangia (S), sporangiophores (Sp), and zoospores (Z). Counterstaining Gf-Ma cells with DAPI reveal fungal nuclei. Cy3 channel visualization of fungal cells reveal sporangia and zoospores readily take up Cy3 labelled nucleic acid probes. Merging of DAPI and Cy3 channels show nucleic acid probes localize to nuclei of motile zoospores and vegetative sporangia.

Transformation efficiency via natural competency was then established via flow cytometry (Figure 2). The arithmetic mean of Cy3 fluorescence as measured with a phycoerythrin (PE) filter set increased 60% in transformed cells relative to untransformed controls (Figure 2; transformed population = $3.0 \pm 1.9 \times 10^4$ AU; untransformed population = $1.9 \pm 0.9 \times 10^4$ AU). This increase in fluorescence was statistically significant (unpaired t-test p-val < 2.2×10^{-16}), suggesting significant uptake of DNA within the assayed cells. To calculate a minimum transformation efficiency, we counted the fraction of cells within the transformed population whose fluorescence was at least 2 standard deviations greater than the mean of the untransformed control group (i.e., statistically significant with $p \leq 0.05$). At least 16.5% of all observed cells were transformed with short nucleic acid probes (>40 transformants/µg DNA). However, these estimates are a conservative estimate or lower bound for transformation efficiency. To prevent clogging of the instrument only smaller zoospores (< 40 µm) were analyzed, excluding transformed sporangia and sporangiophores, which are too large to be run through a flow cytometer (Figure 1). Nonetheless, the flow cytometry independently confirmed DNA uptake via natural competency that could facilitate heterologous gene expression.

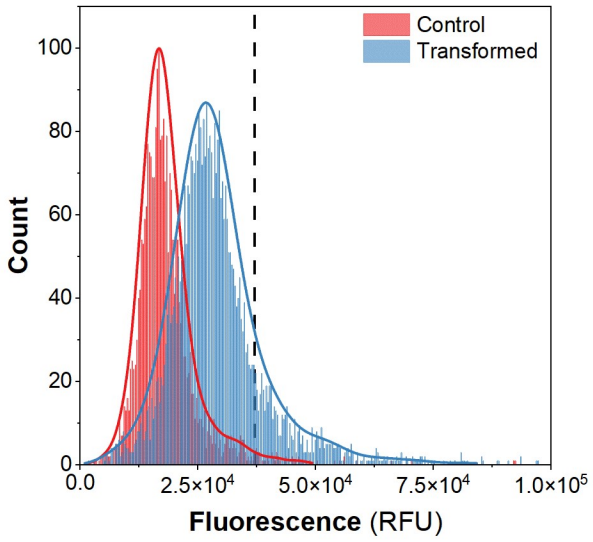


Figure 2: Flow cytometry of fungal zoospores for tracking DNA uptake efficiency
Cy3 fluorescence of gated events measured on the PE channel show transformed zoospores have higher mean fluorescence than untransformed controls. Mean fluorescence + 2 standard deviations of the control (dashed line) is used as the threshold to define a transformant. (See Table S5 for additional statistics of flow cytometry analysis and Figure S9 for collected events and gates used).

Selection systems for anaerobic fungi

As proof of concept for heterologous gene expression, we next identified potential selectable markers for anaerobic fungi and minimal inhibitory concentrations (MIC) of frequently used antibiotics for selection. Fungal growth was tracked by pressure over a multiple log concentration range (Figure 3).[36] Hygromycin, geneticin (G418), neomycin, puromycin, and blasticidin strongly inhibited AGF growth with varying levels of sensitivity (Table 1). AGF were weakly sensitive to neomycin with concentrations in excess of 200 µg/ml needed to knock down growth (Figure 3C). This result is not unexpected as most fungi are not susceptible to aminoglycosides except at high concentrations.[37] Puromycin and blasticidin were highly lethal to AGF with no growth under the concentrations tested (MIC < 3 µg/ml; Figure 3D-E). However, AGF were moderately sensitive to hygromycin and G418 with MICs of 5 and 10 µg/ml, respectively. Conversely, AGF are resistant to bleomycin with no growth inhibition at concentrations of up to 50 µg/ml bleomycin (Figure 3F). Subsequent genome analysis of all sequenced AGF confirmed the presence of a conserved bleomycin resistance gene (lactoglutathione lyase- *lgl*), which forms part of the glyoxyl detoxification pathway.[38,39] This pathway is responsible for protecting hosts from oxidative stress[40,41], which would be evolutionarily advantageous for anaerobic fungi in their native environment. Thus, this resistance gene may serve as a convenient counter-selectable marker for future genome engineering efforts. Given their clear and reproducible MICs, relatively low cost, and high sensitivity, hygromycin and G418 selection were used to establish heterologous expression in anaerobic fungi.

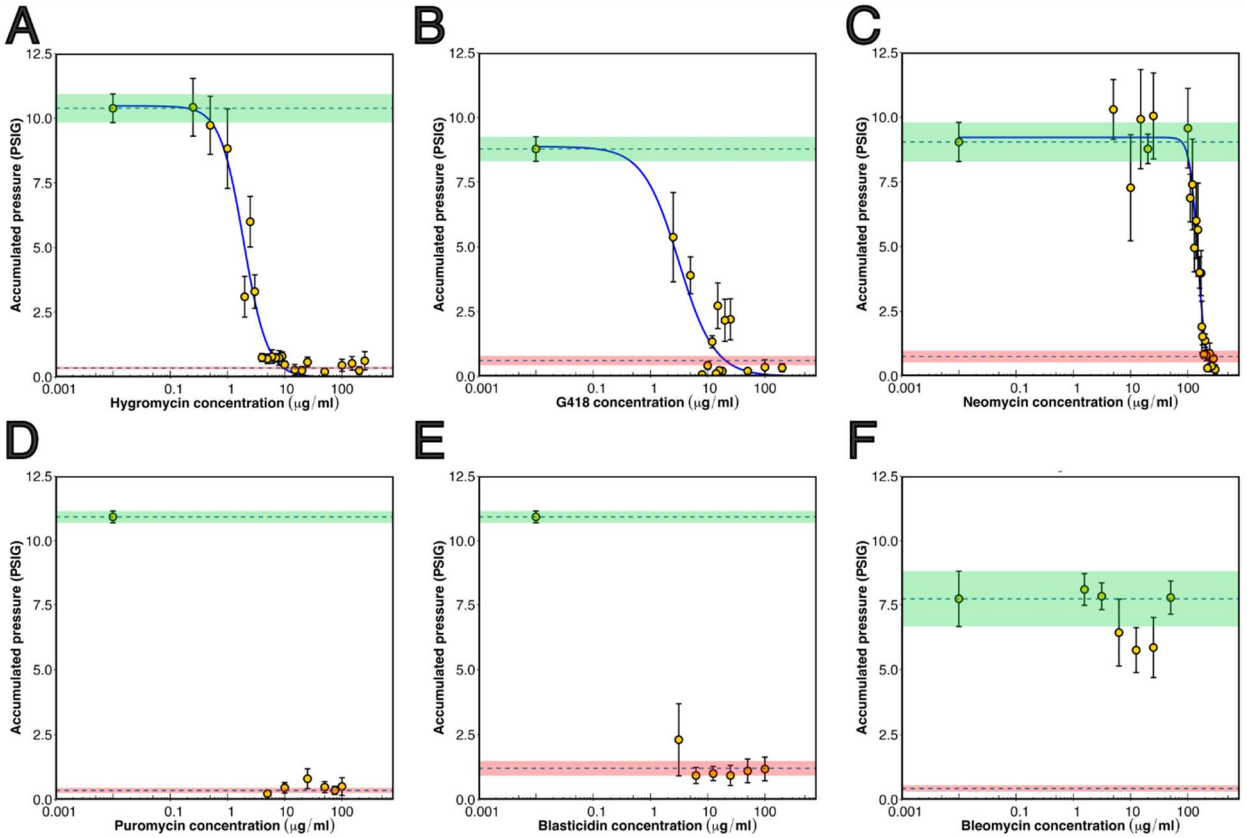


Figure 3: *Neocallimastix frontalis* var *giraffae* is susceptible to hygromycin (A), G418 (B) Neomycin (C), Puromycin (D), Blasticidin (E), but not Bleomycin (F). N ≥ 4 biological replicates. Error bars

represent standard error of the mean. Dashed lines represent the average pressure of uninoculated tubes (lower line) or positive control (upper line) while the red ribbon represents the standard error of the uninoculated controls and the green ribbons represents the standard error of the positive control.

Table 1: Minimal Inhibitory Concentrations of Tested Antibiotics

Antibiotic	Mechanism	MIC (µg/ml)
Blasticidin-S	Inhibits translation	<3
Bleomycin	Induced DNA breaks	Not detected
Geneticin (G418)	Inhibits translation	10
Hygromycin	Inhibits translation	5
Neomycin	Inhibits translation	200
Puromycin	Inhibits translation	<3

Genetic parts for heterologous expression

To develop a robust heterologous expression system in anaerobic fungi, we identified constitutive promoters and transcriptional terminators from conserved non-coding upstream and downstream sequences flanking highly conserved and highly expressed essential enolase (ENOL) and translation elongation factor 1 alpha (TEF1A) genes. These two genes were constitutively expressed on seven different carbon sources in evolutionarily related *Neocallimastix californiae* and *Piromyces finnis* isolates,[4] and due to their essential nature in carbohydrate metabolism and amino acid biosynthesis hypothesized to express at similarly constitutive levels in Gf-Ma. This analysis provided 2 putatively strong constitutive promoters, P_{ENOL} and P_{TEF1A}, and 1 terminator, T_{TEF} terminator (see Supplementary text for details), which we used to construct several heterologous gene cassettes with codon harmonized selectable markers (Hyg – *hph*; G418 – *neo;* Table 2). Genes were codon harmonized (see Supplementary text for details) by selecting AGF codons with relative abundances (Table S1) that matched those of the gene in their native hosts to ensure high expression.

We then tested the ability of these genetic constructs to rescue fungal growth in the presence of hygromycin or G418. As these genetic constructs lacked Autonomously Replicating Sequences (ARS) and centromeric sequences (CEN) for replication and partitioning to daughter cells[42], we dosed in ~2 µg of linearized plasmid daily (Figure 4A). Control studies with labelled plasmid confirmed that larger plasmid DNA was also taken up and localized to the nucleus (Figure 4B and Figure S5), potentially enabling gene expression. We transformed our cells with unlabeled DNA and grew them in non-selective media for 72 h to express the selectable marker before subculturing in selective media (Figure 4A). Cells without construct showed negligible growth (pressure accumulation) when grown in the presence of antibiotic (Figures 4C and D). Cells transformed with P_{ENOL}-*hph* cassettes were able to grow to similar levels above the construct-free negative control in selective media ($p = 0.005$; Figure 4C, Table 2) with significant levels of biomass generation. However, this growth was only ~60% of positive (non-selective) control levels ($p = 0.001$) as measured by pressure evolution (Figure 4C and Figure S6A). These effects are only observed with regular dosing of exogenous DNA (data not shown); that is, provided DNA constructs do not persist in the absence of stabilizing sequences such as an ARS. Similarly, the P_{TEF1A}-*neo* construct was able to rescue ~75% of

growth relative to positive (non-selective) controls (Figure 4D). Promoter-less control constructs (pUCM6-AGF-124) do not rescue growth in selective media and are indistinguishable from untransformed negative controls (Figure S7). Antibiotic resistance phenotypes are dependent on our cloned upstream regions validating that they contain constitutive promoters that are active in anaerobic fungi and can drive heterologous gene expression.

To facilitate the development of gene editing tools and future approaches for more stable transformations, we identified nuclear localization sequences (NLS) that could be used to traffic Cas endonucleases and/or other enzymes to the nucleus. Analysis of the leader sequence of nuclear-localized histone H2B across multiple fungal lineages revealed strong conservation in the first 40 amino acids of anerobic fungal histone proteins (Figure 5A, and Dataset SD5)[5,39,43–45]. This region was distinct from that of homologs in other fungal lineages forming an AGF-specific putative NLS. The AGF consensus sequence (Figure 5B, and Dataset SD6) was used to flank an anaerobic fluorescent reporter protein, iRFP702[46,47], which we codon-optimized for anaerobic fungi. We also generated control reporter constructs that lacked NLS sequences. All reporter constructs were also fused to hygromycin selection cassettes to enable enrichment of transformed cells in selective media. Cells transformed with reporter plasmids all exhibited increased far-red fluorescence as visualized by confocal microscopy and strong fungal growth under selective conditions (Figure 5C; Figure S6A). Expression was observed in all fungal life cycle stages examined. While NLS-free constructs showed more diffuse fluorescence, constructs flanked by NLS sequences displayed punctate or localized fluorescence that coincided with stained nuclei. In other words, addition of the NLS sequence did not disrupt fluorescence of the iRFP fluorophore and, more importantly, localized iRFP to fungal nuclei. To quantify this phenomenon, transformed cells were analyzed via flow cytometry. The arithmetic mean fluorescence of transformed cells lacking the NLS sequence was 3.50 times as much as that of untransformed cells (Figure 6). NLS-tagged transformants exhibited 4.55 times the fluorescence of untransformed controls, potentially due to the localized or concentrated nature of protein expression (Figure 6). In summary, we successfully identified and validated anaerobic fluorescent reporters and NLS sequences for AGF establishing powerful new tools to track protein expression and localize proteins to specific compartments.

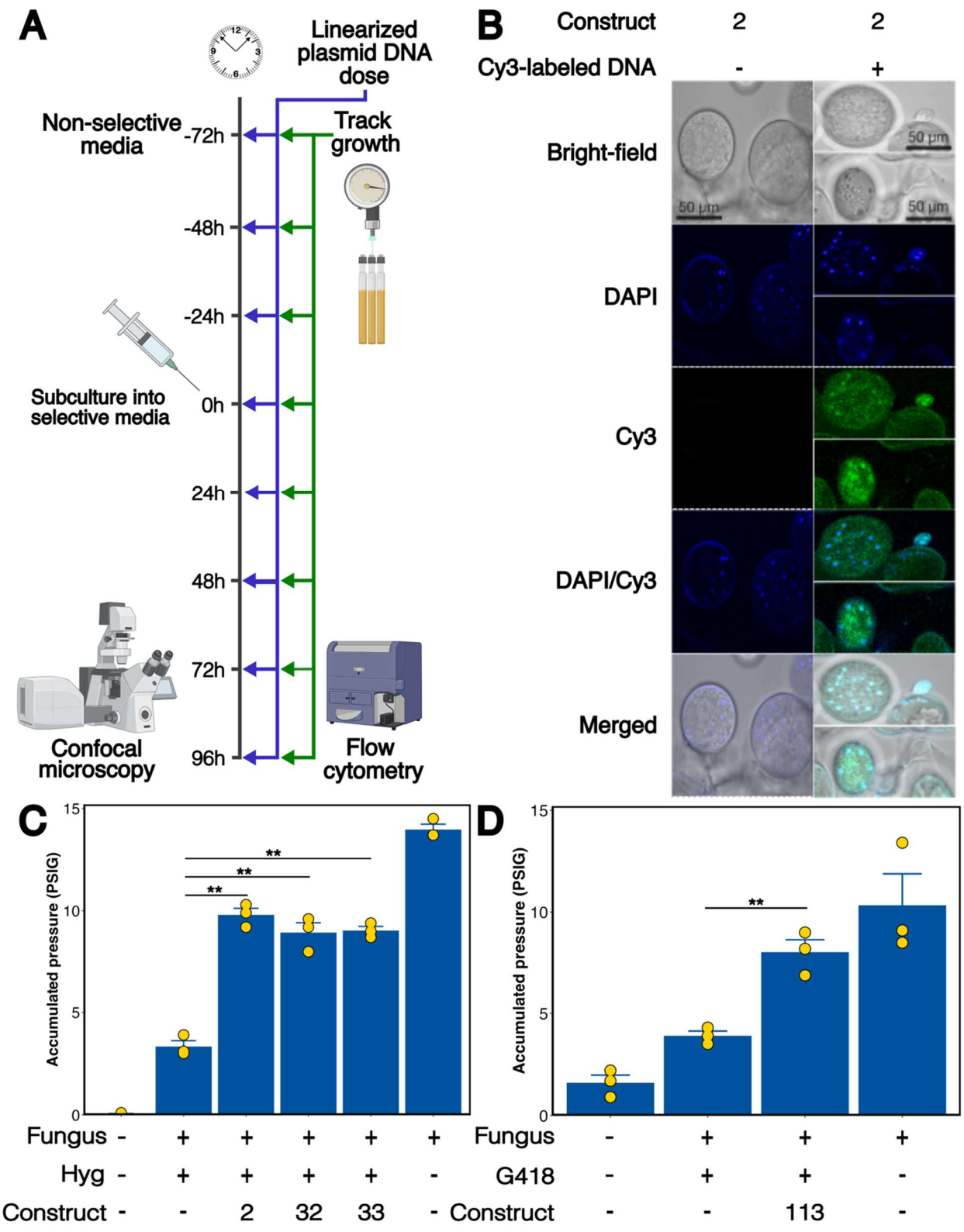


Figure 4: Dosing linearized plasmids into *Neocallimastix frontalis* var *giraffae* rescues fungal growth under selective conditions. A) Plasmid dosing scheme B) Cy3 labelled pUCM6-AGF-2 harboring P_{ENOL}-*hph* is taken up and localizes to fungal nuclei C) Hygromycin resistance markers can be heterologously

expressed and rescue growth in selective media relative to untransformed control via P_{ENOL} [Constructs express: 2 – Hyg^R, 32 – Hyg^R + a nuclear localized fluorescent reporter, 33 – Hyg^R + fluorescent reporter] D) *neo* resistance marker is heterologously expressed from P_{TEF1A} in construct 113 to rescue growth on G418. See Table 2 for a full description of each construct and Figure S6 for corresponding growth curves. ** p-value ≤ 0.0050

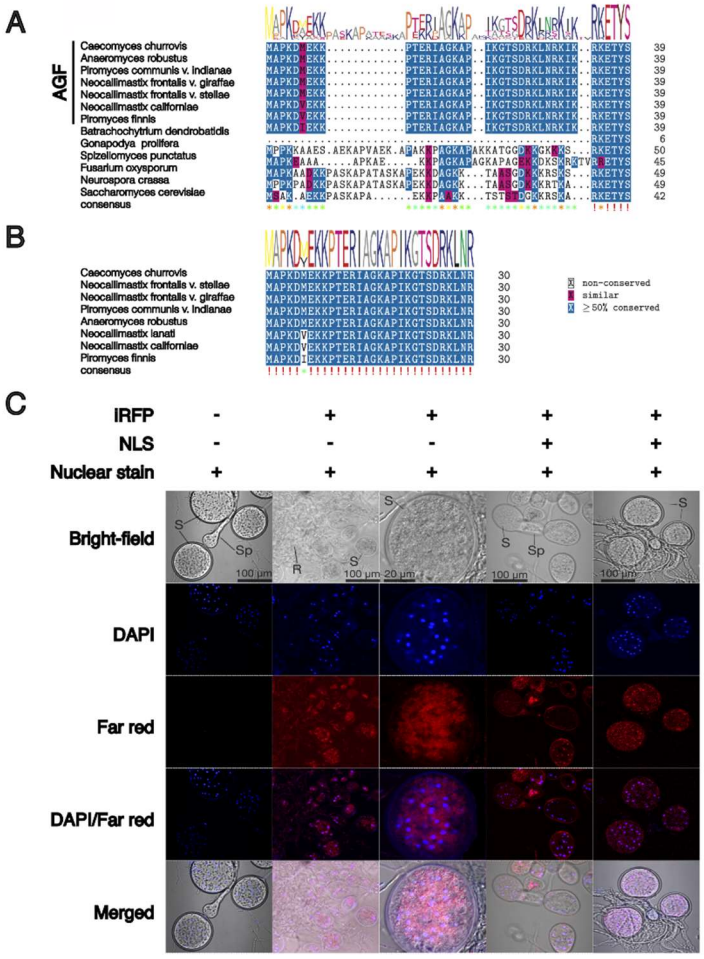


Figure 5: NLS sequence from Histone H2B localizes anaerobic fluorescent reporters to the nuclei of fungal cells. A) Anaerobic fungi have unique 5' leader sequences on histone H2B that are distinct from leader sequences of model fungal lineages. B) Consensus sequence of conserved anaerobic fungi Histone 2B leader sequence or putative NLS. C) Confocal microscopy of fungal cells transformed with constructs containing hygromycin resistance and iRFP enable selection and visualization of iRFP expression. See Table 2 for full descriptions of these constructs.

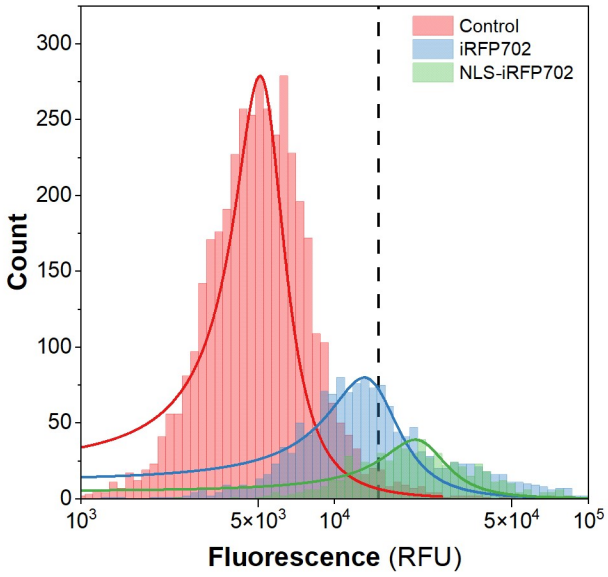


Figure 6: Flow cytometry analysis of fungal cells transformed with iRFP cassettes Fluorescence for gated events of fungal zoospores transformed with iRFP containing constructs have a higher mean fluorescence (2.0 x 10^5) than untransformed controls (5.8 x 10^4). Similarly, fluorescence for gated events of zoospores transformed with NLS-tagged iRFP also show higher mean fluorescence (2.6 x 10^5) than untransformed controls. Table 2 contains full descriptions of the constructs used to generate these plots, and Table S5 provides additional details on collection of flow cytometry data. Raw flow cytometry data and gates used provided in Figure S10

Discussion

Anaerobic fungi offer a wealth of biotechnological potential such as superior enzymes to hydrolyze plant biomass[4], or previously unidentified natural products [6,8], however efforts to domesticate and engineer these organisms has significantly lagged behind that of many emerging model organisms. Methods and tools to transform model fungal lineages relying on protoplasting or *Agrobacterium* mediated protocols[32] are not amenable to anaerobic fungi due to the thick cell wall and differences in cultivation temperature. In this work we establish natural competency as a robust scheme to introduce DNA into fungal cells (Figure 1). Natural competency has been reported in model fungal lineages as a strategy by which nucleic acids can be reliably introduced. While natural competency for RNA has been reported and leveraged for RNAi silencing in anaerobic fungi[35], our work demonstrates that DNA competency can be similarly exploited for heterologous expression. This approach is less lethal than previous biolistic methods[34], the only previously demonstrated method for DNA transformation of anaerobic fungi, and is at least 10 times more efficient than that of previously reported biolistic methods (~40 transformants/µg vs < 4 transformants/µg via biolistic methods).[34]. Moreover, this approach is more accessible to labs, removing the need for specialized equipment such as an anaerobic chamber to manipulate the fungi. All that is needed are sealed Hungate tubes to cultivate anaerobic fungi and syringes to inoculate and introduce DNA. Further, our transformation scheme also expands on past natural competency transformation methods by transforming all lifecycle stages instead of selectively transforming anaerobic fungal zoospores.[35] By transforming all lifecycle stages

instead of enriching only fungal zoospores we are able to perform transformations approximately 72 hours faster than currently reported methods, significantly reducing the Design Build Test Learn cycle. Additionally, we established the first methods for performing flow cytometry of anaerobic fungal zoospores (Figure 2 and Figure 6) which not only enables rapid screening of transformants, but also establishes a procedure for estimating transformation efficiency. In summary, our scheme for transforming anaerobic fungi overcomes past bottlenecks of high cell mortality, high cost, and low throughput.

In this work we also establish a suite of genetic parts to enable genetic engineering of gut fungi including promoters, resistance markers, nuclear localization sequences, and fluorescent reporters. To date the only functional genetic parts identified for anaerobic fungi were the enolase promoter of *Neocallimastix frontalis* and the enzymatic reporter β-glucoronidase[34]. However, as this was developed before community repositories such as Addgene and the source isolate lost, no genetic parts are currently available for these organisms. To address this lack of parts, we identified and validated functionality of two promoter sequences (Figures S1-S2, Figure 4C-D) and demonstrated that they reliably drive expression of two antibiotic resistance markers, *hph* and *neo*. Moreover, we identified and validated functionality of an anaerobic fungal NLS sequence (Figure 5A-B). This sequence was anaerobic fungal specific and, similarly to those found on other histone proteins in other species, did not display a classical monopartite or bipartite structure.[48] Anaerobic fungal genomes are especially GC-poor leading to a depletion of amino acids encoded by GC-rich codons such as arginine, which are common in NLSs.[48,49] Our identified NLS contained no GC-rich codon repeats and was ~10 residues shorter than NLSs from later diverging fungi. Thus, anaerobic fungi may have evolved unique nuclear importins that do not rely as strongly on canonical arginine residues for function. Despite this, the identified NLS retains a net positive charge (pI ~ 10.1) under homeostatic conditions, like canonical NLS sequences, suggesting some conserved mechanism for nuclear transport.

We also validate the use of the anaerobic reporter iRFP in anaerobic fungi to visualize protein trafficking within anaerobic fungi (Figure 5C) and leveraged this expression to develop flow cytometry protocols to rapidly screen transformed populations (Figure 6). Transformants were 3-4x brighter than control cells validating heterologous expression in specific cellular compartments. However, the low brightness suggested low transformation or expression. While iRFP was selected in this study due to the low background fluorescence of anaerobic fungi within its fluorescence spectra, it is a dim fluorophore with a brightness (quantum yield × molar extinction) up to an order of magnitude lower than other anaerobic fluorophores such as iLOV[46] and UnaG.[50] Thus, future studies should expand this palette of reporters to improve the sensitivity of these tools and more robustly quantify transformation efficiency.

In this study, we have developed simple and reliable methodologies for the transformation of the emerging anaerobic fungal chassis. We demonstrate that these species are naturally competent for DNA during multiple stages of their life cycle and can heterologous express and traffic a number of enzymes and reporters to various cellular compartments. While only transient, this transformation can be sustained through periodic dosing of heterologous DNA directly to cell culture – no equipment or sample preparation

is required. More importantly, these methods are potentially enabling for future strain development efforts such as those that use CRISPR technology. In many fungi and higher eukaryotes, transient expression is optimal for gene editing due to toxicity arising from off-target events with high stable expression of Cas9.[51–53] Use and further development of these tools unlocks new avenues of research in this class of lignocellulolytic fungi, accelerating the development of these species for biomanufacturing,[12,54,55] drug discovery[6,7] biogas,[56,57] and agriculture.[9]

Methods:

Anaerobic fungal isolate, media and growth conditions:

Neocallimastix frontalis var *giraffae* (Gf-ma) was isolated from the feces of a giraffe (*Giraffa reticulata*) from the Indianapolis Zoo and maintained on a semi-defined medium B supplemented with chloramphenicol at a final concentration of 35 µg/ml as previously described.[12] All cultures were grown in 14 ml Hungate tubes at 39 °C unless otherwise specified. Briefly, media was prepared by microwaving all components except glucose and cysteine for approximately 13 minutes prior to sparging with CO_2 (Keen Gas, Newark, DE, USA) for approximately 1 hour to remove oxygen present in the media. Subsequently, glucose (5 g/l) (Fisher Scientific, Waltham, MA, USA) and L-Cysteine hydrochloride (1.25 g/l) (Fisher Scientific, Waltham, MA, USA) were added prior to dispensing media into individual tubes. Specifically, anaerobic liquid media was dispensed to 14 ml Hungate tubes or 100 ml serum bottles under 100% CO_2. Anaerobic tubes and bottles were then sealed and maintained under 100% CO_2 headspace prior to autoclaving with an exposure time of 20 minutes at 120 °C. Sterile anaerobic media was stored at 4 ºC until it was ready for use. All media was prewarmed to 39 ºC one hour before inoculation. To assess the phylogenetic affiliation of strain Gf-ma in relation to the other *Neocallimastix* species as well as other anaerobic fungal genera, four copies of the D1/D2 of nuc 28S rDNA genes sequences were extracted from the genome.[39] The obtained sequences were aligned to anaerobic fungal reference 28S rDNA sequence dataset downloaded from the National Center for Biotechnology Information (NCBI) GenBank nr database using Clustal W using default parameters and manually refined in MEGA7.[58] The generated alignment was then used to construct a maximum likelihood phylogenetic tree in FastTree using the gtr model[59] with *Chytriomyces* sp. WB235A as the outgroup. Bootstrap values were calculated based on 100 replicates (Figure S8).

Transformations of Neocallimastix. frontalis var giraffae with labeled nucleic acid probes:

To assess the natural competency of strain Gf-ma to exogenous dsDNA*,* a 50 bp long ssDNA (5'-ATCTGTACCGGCTAATGCGAATCGATAGCTACCGATTGCGATTCGGATTC-3') was purchased from Integrated DNA Technologies, Inc.(Coralville, IA, USA) and labeled with Cyanine3 (Cy3) fluorescent dye at the 5′ end to enable visual detection of DNA uptake. Cy3 labeling was conducted using *Label* IT™ nucleic acid labeling kit, Cy3 from Mirus Bio (Madison, WI, USA) following the manufacturer's instructions. Then we created a double stranded DNA probe by annealing the Cy3-labeled ssDNA oligomers with the unlabeled complementary strand by incubating them at 95 °C followed by a gradual cooling down at room

temperature for 45 min to permit hybridization. Young cultures of strain Gf-ma (18 h, grown on medium B) were mixed with Cy3-labeled dsDNA at a final concentration of 20 nM, and incubated at 39 °C for 90 min. Samples were taken every 15 min., counterstained with 4′,6 diamidino-2-phenylindole (DAPI; 10 µg/mL) and examined using confocal microscopy (see *Confocal Microscopy and Image processing*) for nuclear visualization and detection of DNA uptake by different lifecycle stages. The Cy3 kit was also used to label a heterologous construct harboring hygromycin resistance (See Figure 4B) following the same workflow.

Kill curve assays

To determine the minimal inhibitory concentration of antibiotics on *Neocallimastix*. *frontalis* var *giraffae*, 14 ml Hungate tubes containing 9 ml of media B and chloramphenicol (see *Anaerobic fungal isolate, media and growth conditions*) received either hygromycin B, neomycin, puromycin dihydrochloride, blasticidin s hydrochloride, or G418. All antibiotic solutions were diluted in water to achieve 100X solutions at each concentration range to be tested. 100 µl of 100X working solution was then added to each tube prior to inoculation with 1 ml of actively growing *Neocallimastix frontalis* var *giraffae* culture. In order to overcome biological noise, all test conditions as well as positive and negative controls were carried out in biological quadruplicate at a minimum. Additionally, all Hungate tubes were zeroed out to atmospheric pressure immediately after inoculation. Growth of *Neocallimastix frontalis* var *giraffae* was determined by pressure accumulation [36] in the sealed Hungate tubes via a pressure transducer (Automation Products Group, Logan, UT, USA) daily over a period of 7 days. Total accumulated pressure was then determined for all test conditions. Growth curves of pressure accumulation versus antibiotic concentration were generated in R with the ggplot2 package.[60] The pracma package was used to fit Hill equations to the data.[61]

Transformations of Neocallimastix. frontalis var giraffae with reporter constructs on linearized plasmids

Constructs tested in this report were first transformed into chemically competent *E. coli* strain DH5α. Single colonies were inoculated into 5 ml of Lauria-Bertani (LB) broth containing 100 µg/ml ampicillin and cultured at 37 ºC at 250 RPM. After approximately 16-18 hours, these 5 ml liquid cultures were used to inoculate 50 ml LB cultures containing 100 ug/ml Ampicillin in 25 0ml baffled flasks. The 50 ml *E. coli* cultures were incubated at 37 ºC at 250 RPM until the optical density reached approximately 2.5 absorbance units. Liquid cultures were subsequently centrifuged at 10,000 RCF for 10 minutes at 4 ºC prior to midiprepping with a GeneJet midiprep kit (Fisher Scientific, Waltham, MA, USA). Midiprepped plasmid DNA was linearized with Fast Digest BamHI or EcoRI in Fast Digest buffer. DNA digest reactions were heat inactivated for 5 minutes at 80 ºC, according to the manufacturer's instructions. Sterile anaerobic water (1.25 g/l cysteine hydrochloride, 1.5 g/l PIPES) was prepared in sterile 100 ml serum bottles and sparged with CO_2 for five minutes to remove oxygen. Sterile anaerobic water was subsequently prewarmed to 39 ºC and used to dilute digested DNA to a final volume of 300 µl. 100 µl of this digested DNA was added to each Hungate tube so that all constructs were tested in biological triplicate. After addition of digested plasmid DNA, Hungate tubes were zeroed to atmospheric pressure. Since these genetic constructs are unstable due to the lack of Autonomously Replicating Sequences (ARS) and centromeric sequences (CEN), we

dosed in ~2 µg of linearized plasmid daily before and after antibiotics selection and fluorophore detection (Figure 4A).

Flow cytometry of Neocallimastix frontalis var *giraffae cultures*

Fungal zoospores were collected from liquid cultures by opening Hungate tubes, filtering through 8 layers of cheesecloth directly into sterile 50ml Falcon tubes to remove large fungal biomass and rhizomycelia while also enriching for fungal zoospores. This filtered fermentation liquor was then centrifuged at 4,500 RCF for 10 minutes at 39 ºC to pellet fungal zoospores. The supernatant was gently decanted and the resulting pellet was resuspended in 1 ml of sterile anaerobic water prewarmed to 39 ºC (See *Transformations of Neocallimastix. frontalis* var *giraffae with reporter constructs on linearized plasmids*). The concentrated zoospores were then carefully pipetted through BD FACS tubes with 35 µm mesh size (BD, Franklin Lakes, NJ, USA) to eliminate any clumps or cell debris. Samples were immediately analyzed using a NovoCyte flow cytometer. For flow cytometry analysis on Cy3 transformed cultures 50,000 ungated events were collected with an opening of 21.1 µm and a flow rate of 104 µl per minute. A PE (R-Phycoerythrin) laser with 488 nm excitation and 660/20 nm detection was used for all Cy3 transformed fungal zoospores. Gating was performed by running media only controls to determine which events were caused by debris in fungal media. For flow cytometry on cultures transformed with iRFP containing constructs, the Alexa-Fluor 700 channel was used with 640 nm excitation and 695/40 nm detection with the same flow rate and opening size. Gating of events was performed through the same procedure as Cy3 transformed zoospores. Fluorescence intensity on PE-H or Alexa Fluor 700-H was compared to untransformed controls, and transformation efficiency was calculated by determining the number of gated events having fluorescence signal at least 2 standard deviations above the mean fluorescence intensity of untransformed controls. For all analyses, the arithmetic mean fluorescence was used when calculating transformation efficiency. Data collection and analysis were performed with NovoExpress 1.5.0.

Confocal microscopy and image processing

Fungal biomass was harvested from liquid cultures and stained with 4,6′-diamidino-2-phenylindole (DAPI; 10 µg/mL) to be scanned using confocal laser scanning microscopy (CLSM). Images were captured using a Zeiss LSM 980/Airyscan 2 (Carl Zeiss, Inc., Gottingen, Germany), equipped with 40X and 60X oil immersion objective lens. Fungal cells transformed with Cy3 were examined using a Cy3 (Cyanine-3) channel applying the following wavelengths: 534 nm for excitation and 694 nm for detection. While fungal cells transformed with iRFP were examined using Alexa-Fluor 700 channel applying the following wavelengths: 640 nm for excitation and 755 nm for detection. Nuclear visualization and localization were detected via a DAPI channel applying the following wavelengths: 410 nm for excitation and 605 nm for detection. Image acquisition and processing were handled using an AxioCam and Fiji software[62].

Promoter and terminator alignments

DNA sequences for regions upstream of Neocallimastigomycota enolase genes were downloaded from Mycocosm[39] and aligned in RStudio Version 4.2.0.[63] DNA sequences were aligned with the multiple

sequence alignment (MSA) package version 1.28.0 within the Bioconductor suite version 3.15.[64,65] Sequences were aligned using ClustalW with the default substitution matrix, the neighbor joining method, a gap open penalty of 15, and gap extension penalty of 6.6 with 3 iterations.[64] For all promoter alignments 5' untranslated sequences, if annotated, were included in the alignment. Terminator sequence alignments were created using the same alignment parameters, however 3' untranslated sequences, if present in the genome annotation, were excluded. All protein IDs used to create these alignments are presented in Tables S1 and S2). Complete multiple sequence alignments for these analyses are available as Datasets SD1-SD6.

Identification of a Nuclear Localization Signal Sequence

To identify a putative Nuclear Localization Signal (NLS) sequence in anaerobic fungi, amino acid sequences of histone H2B, a protein that is targeted to the nuclei of eukaryotic cells, were collected from Mycocosm.[39] This preliminary list included both species of fungi in the Neocallimastigomycota lineage as well as model fungi (Figure 4A). Protein IDs collected from Mycocosm[39] from the corresponding genomes are listed in Table S3. These sequences were aligned using the MSA R package within the Bioconductor suite.[64,65] The histone H2B multiple sequence alignment files are available as Supplemental Dataset SD5 which includes both anaerobic fungal and model fungal lineages, and Dataset SD6 which only includes gut fungal H2B sequences. The first 30 amino acids of histone H2B (protein ID 728204) for *Neocallimastix frontalis* var *giraffae* were selected as an NLS and cloned on the 5' and 3' ends of an AGF codon harmonized variant of iRFP702. 5X glycine linkers were used between NLS sequences and the iRFP702 gene. These NLS sequences were additionally codon harmonized to further promote expression when transformed into anaerobic fungi. pUCM6-AGF-32, which contains the NLS tagged iRFP fluorophore, is described in more detail in Table 2.

Table 2: Description of constructs used in the present work.

Construct ID	Plasmid Name	Relevant genoytpe	Size (bp)
2	pUCM6-AGF-2	ColE1(pBR322) ori, *bla*(Amp^r), P_{ENOL}-HygR(AGF-harmonized)-T_{TEF1A}	5465
32	pUCM6-AGF-32	ColE1(pBR322) ori, *bla*(Amp^r), P_{ENOL}-HygR(AGF-harmonized)-T_{TEF1A}-P_{ENOL}-NLS-iRFP702(AGF-harmonized)-NLS-T_{TEF1A}	7862
33	pUCM6-AGF-33	ColE1(pBR322) ori, *bla*(Amp^r), P_{ENOL}-HygR(AGF-harmonized)-T_{TEF1A}-P_{ENOL}-iRFP702(AGF-harmonized)-T_{TEF1A}	7628
79	pUCM6-AGF-79	ColE1(pBR322) ori, *bla*(Amp^r), P_{TEF1A}-HygR(AGF-harmonized)-T_{TEF1A}-P_{ENOL}-iRFP702(AGF-harmonized)-T_{TEF1A}	7633
113	pUCM6-AGF-113	ColE1(pBR322) ori, *bla*(Amp^r), P_{TEF1A}-Neo(AGF-optimized)-T_{TEF1A}	5252
124	pUCM6-AGF-124	ColE1(pBR322) ori, *bla*(Amp^r), HygR(AGF-harmonized)-T_{TEF1A}	5497

Supporting information

- Supplementary Info: Description of how anaerobic fungal parts were developed, information about sequences used in bioinformatic analyses, flow cytometry data, fungal growth curves, and phylogenetic analysis
- Dataset SD1: SD1_Enolase_1000_MSA.txt - alignment of anaerobic fungal enolase leader sequences used to identify an enolase promoter
- Dataset SD2: SD2_TEF1A_1000_Promoter_MSA.txt- alignment of Translation Elongation Factor 1 Alpha (TEF1A) leader sequences used to identify a TEF1A promoter for anaerobic fungi
- Dataset SD3: SD3_TEF1A_500_Terminator.txt - alignment of TEF1A downstream sequences used to identify a TEF1A terminator for anaerobic fungi
- Dataset SD4: SD4_Gfma_TEF1A_500_Terminator.txt - alignment of *Neocallimastix frontalis* TEF1A promoter sequences used to identify a *Neocallimastix frontalis* TEF1A promoter
- Dataset SD5: SD5_NLS_all_MSA.txt - alignment of fungal histone H2B sequences used to assess Nuclear Localization Signal sequence conservation across multiple fungal lineages
- Dataset SD6: SD6_Neocallimastigomycota_H2B.txt- alignment of anaerobic fungal histone H2B sequences to assess Nuclear Localization Signal sequence conservation across all Neocallimastigomycota

Conflicts of Interest

The authors declare no competing financial interest.

Author contributions

K.V.S conceived the study. EH performed strain isolation. C.H., E.H. J.M, and R.H. performed genomic analyses. CH, EH, and RH designed constructs generated in this study. C.H. and R.H. performed transformations in this study. R.H. performed confocal microscopy and image processing. C.H. performed flow cytometry experiments and analyses. C.H. and R.H. and K.V.S wrote the manuscript.

Acknowledgments

CH was supported by a National Science Foundation Graduate Research Fellowship under DGE-1333468. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Number DE-SC0022018/DE-SC0022206 and the National Science Foundation under Grant No. IOS-2128272. A portion of this research was performed under the Facilities Integrating Collaborations for User Science (FICUS) program (proposal doi: 10.46936/fics.proj.2018.50397/60000049) and used resources at the DOE Joint Genome Institute (https://ror.org/04xm1d337) and the Environmental Molecular Sciences Laboratory (https://ror.org/04rc0xn13), which are DOE Office of Science User Facilities operated under Contract Nos. DE-AC02-05CH11231 (JGI) and DE-AC05-76RL01830 (EMSL). The LSM980 confocal microscope was acquired with a shared instrumentation grant (1S10RR027273-01) from the National Institutes of Health (NIH). Microscopy access was supported by grants from the NIH-NIGMS (P20 GM103446), the NIGMS (P20 GM139760) and the State of Delaware. This publication was made possible by the Delaware INBRE program, supported by a grant from the National Institute of General Medical Sciences – NIGMS (P20 GM103446) from the National Institutes of Health and the State of Delaware. Last, we would like to express our gratitude and pay our respects to the late Dr. Richard West for his considerable help in the Flow Cytometry Core facility.

References

(1) Gruninger, R. J.; Puniya, A. K.; Callaghan, T. M.; Edwards, J. E.; Youssef, N.; Dagar, S. S.; Fliegerova, K.; Griffith, G. W.; Forster, R.; Tsang, A.; McAllister, T.; Elshahed, M. S. Anaerobic Fungi (Phylum Neocallimastigomycota): Advances in Understanding Their Taxonomy, Life Cycle, Ecology, Role and Biotechnological Potential. *FEMS Microbiol. Ecol.* **2014**, *90* (1), 1–17. https://doi.org/10.1111/1574-6941.12383.

(2) Orpin, C. Studies on the Rumen Flagellate Neocallimastix Frontalis. *Microbiology* **1975**, *91* (2), 249–262. https://doi.org/10.1099/00221287-91-2-249.

(3) Hooker, C. A.; Hillman, E. T.; Overton, J. C.; Ortiz-Velez, A.; Schacht, M.; Hunnicutt, A.; Mosier, N. S.; Solomon, K. V. Hydrolysis of Untreated Lignocellulosic Feedstock Is Independent of S-Lignin Composition in Newly Classified Anaerobic Fungal Isolate, Piromyces Sp. UH3-1. *Biotechnol. Biofuels* **2018**, *11* (1), 293. https://doi.org/10.1186/s13068-018-1292-8.

(4) Solomon, K. V.; Haitjema, C. H.; Henske, J. K.; Gilmore, S. P.; Borges-Rivera, D.; Lipzen, A.; Brewer, H. M.; Purvine, S. O.; Wright, A. T.; Theodorou, M. K.; Grigoriev, I. V.; Regev, A.; Thompson, D. A.; O'Malley, M. A. Early-Branching Gut Fungi Possess a Large, Comprehensive Array of Biomass-Degrading Enzymes. *Science* **2016**, *351* (6278), 1192. https://doi.org/10.1126/science.aad1431.

(5) Haitjema, C. H.; Gilmore, S. P.; Henske, J. K.; Solomon, K. V.; de Groot, R.; Kuo, A.; Mondo, S. J.; Salamov, A. A.; LaButti, K.; Zhao, Z.; Chiniquy, J.; Barry, K.; Brewer, H. M.; Purvine, S. O.; Wright, A. T.; Hainaut, M.; Boxma, B.; van Alen, T.; Hackstein, J. H. P.; Henrissat, B.; Baker, S. E.; Grigoriev, I. V.; O'Malley, M. A. A Parts List for Fungal Cellulosomes Revealed by Comparative Genomics. *Nat. Microbiol.* **2017**, *2* (8), 17087. https://doi.org/10.1038/nmicrobiol.2017.87.

(6) Swift, C.L.; Louie, K.B.; Bowen, B. P.; Olson, H. M.; Purvine, S. O.; Salamov, A.; Mondo, S. J.; Solomon, K. V.; Wright, A. T.; Northen, T. R.; Grigoriev, I. V.; Keller, N. P.; O'Malley, M. A. Anaerobic Gut Fungi Are an Untapped Reservoir of Natural Products. *Proc. Natl. Acad. Sci.* **2021**, *118* (18), e2019855118. https://doi.org/10.1073/pnas.2019855118.

(7) Hillman, E. T.; Readnour, L. R.; Solomon, K. V. Exploiting the Natural Product Potential of Fungi with Integrated-Omics and Synthetic Biology Approaches. *Curr. Opin. Syst. Biol.* **2017**, *5*, 50–56. https://doi.org/10.1016/j.coisb.2017.07.010.

(8) Brown, J. L.; Swift, C. L.; Mondo, S. J.; Seppala, S.; Salamov, A.; Singan, V.; Henrissat, B.; Drula, E.; Henske, J. K.; Lee, S. Co-cultivation of the Anaerobic Fungus Caecomyces Churrovis with Methanobacterium Bryantii Enhances Transcription of Carbohydrate Binding Modules, Dockerins, and Pyruvate Formate Lyases on Specific Substrates. *Biotechnol. Biofuels* **2021**, *14* (1), 1–16. https://doi.org/10.1186/s13068-021-02083-w.

(9) Hooker, C. A.; Lee, K. Z.; Solomon, K. V. Leveraging Anaerobic Fungi for Biotechnology. *Curr. Opin. Biotechnol.* **2019**, *59*, 103–110. https://doi.org/10.1016/j.copbio.2019.03.013.

(10) Li, X.-L.; Ljungdahl, L. G.; Ximenes, E. A.; Chen, H.; Felix, C. R.; Cotta, M. A.; Dien, B. S. Properties of a Recombinant β-Glucosidase from Polycentric Anaerobic Fungus Orpinomyces PC-2 and Its Application for Cellulose Hydrolysis. *Appl. Biochem. Biotechnol.* **2004**, *113* (1), 233–250. https://doi.org/10.1385/abab:113:1-3:233.

(11) Seppälä, S.; Wilken, St. E.; Knop, D.; Solomon, K. V.; O'Malley, M. A. The Importance of Sourcing Enzymes from Non-Conventional Fungi for Metabolic Engineering and Biomass Breakdown. *Metab. Eng.* **2017**, *44*, 45–59. https://doi.org/10.1016/j.ymben.2017.09.008.

(12) Hillman, E. T.; Li, M.; Hooker, C. A.; Englaender, J. A.; Wheeldon, I.; Solomon, K. V. Hydrolysis of Lignocellulose by Anaerobic Fungi Produces Free Sugars and Organic Acids for Two-Stage Fine Chemical Production with Kluyveromyces Marxianus. *Biotechnol. Prog.* **2021**, *37* (5), e3172. https://doi.org/10.1002/btpr.3172.

(13) Hillman, E. T.; Frazier, E. M.; Shank, E. K.; Ortiz-Velez, A. N.; Englaender, J. A.; Solomon, K. V. Anaerobic Fungal Mevalonate Pathway Genomic Biases Lead to Heterologous Toxicity Underpredicted by Codon Adaptation Indices. *Microorganisms* **2021**, *9* (9), 1986. https://doi.org/10.3390/microorganisms9091986.

(14) Podolsky, I. A.; Seppälä, S.; Xu, H.; Jin, Y.-S.; O'Malley, M. A. A SWEET Surprise: Anaerobic Fungal Sugar Transporters and Chimeras Enhance Sugar Uptake in Yeast. *Metab. Eng.* **2021**, *66*, 137–147. https://doi.org/10.1016/j.ymben.2021.04.009.

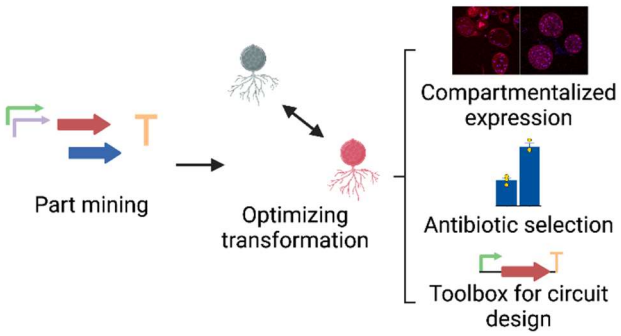
(15) Podolsky, I. A.; Schauer, E. E.; Seppälä, S.; O'Malley, M. A. Identification of Novel Membrane Proteins for Improved Lignocellulose Conversion. *Curr. Opin. Biotechnol.* **2022**, *73*, 198–204. https://doi.org/10.1016/j.copbio.2021.08.010.

(16) Seppälä, S.; Yoo, J. I.; Yur, D.; O'Malley, M. A. Heterologous Transporters from Anaerobic Fungi Bolster Fluoride Tolerance in Saccharomyces Cerevisiae. *Metab. Eng. Commun.* **2019**, *9*, e00091. https://doi.org/10.1016/j.mec.2019.e00091.
(17) Yoo, J. I.; Seppälä, S.; O'Malley, M. A. Engineered Fluoride Sensitivity Enables Biocontainment and Selection of Genetically-Modified Yeasts. *Nat. Commun.* **2020**, *11* (1), 1–9. https://doi.org/10.1038/s41467-020-19271-1.
(18) Gasser, B.; Saloheimo, M.; Rinas, U.; Dragosits, M.; Rodríguez-Carmona, E.; Baumann, K.; Giuliani, M.; Parrilli, E.; Branduardi, P.; Lang, C.; Porro, D.; Ferrer, P.; Tutino, M. L.; Mattanovich, D.; Villaverde, A. Protein Folding and Conformational Stress in Microbial Cells Producing Recombinant Proteins: A Host Comparative Overview. *Microb. Cell Factories* **2008**, *7* (1), 11. https://doi.org/10.1186/1475-2859-7-11.
(19) Seppälä, S.; Lankiewicz, T. S.; Saxena, M.; Henske, J. K.; Salamov, A. A.; Grigoriev, I. V.; O'Malley, M. A. Genomic and Proteomic Biases Inform Metabolic Engineering Strategies for Anaerobic Fungi. *Metab. Eng. Commun.* **2020**, *10*, e00107. https://doi.org/10.1016/j.mec.2019.e00107.
(20) O'Malley, M. A.; Theodorou, M. K.; Kaiser, C. A. Evaluating Expression and Catalytic Activity of Anaerobic Fungal Fibrolytic Enzymes Native to Piromyces Sp E2 in Saccharomyces Cerevisiae. *Environ. Prog. Sustain. Energy* **2012**, *31* (1), 37–46. https://doi.org/10.1002/ep.10614.
(21) Riley, L. A.; Guss, A. M. Approaches to Genetic Tool Development for Rapid Domestication of Non-Model Microorganisms. *Biotechnol. Biofuels* **2021**, *14* (1), 30. https://doi.org/10.1186/s13068-020-01872-z.
(22) Fatma, Z.; Schultz, C.; Zhao, H. Recent Advances in Domesticating Non-model Microorganisms. *Biotechnol. Prog.* **2020**, e3008. https://doi.org/10.1002/btpr.3008.
(23) Nakamura, Y.; Nishi, T.; Noguchi, R.; Ito, Y.; Watanabe, T.; Nishiyama, T.; Aikawa, S.; Hasunuma, T.; Ishii, J.; Okubo, Y. A Stable, Autonomously Replicating Plasmid Vector Containing Pichia Pastoris Centromeric DNA. *Appl. Environ. Microbiol.* **2018**, *84* (15), e02882-17. https://doi.org/10.1128/AEM.02882-17.
(24) Cao, M.; Seetharam, A. S.; Severin, A. J.; Shao, Z. Rapid Isolation of Centromeres from Scheffersomyces Stipitis. *ACS Synth. Biol.* **2017**, *6* (11), 2028–2034. https://doi.org/10.1021/acssynbio.7b00166.
(25) Teh, M. Y.; Ooi, K. H.; Danny Teo, S. X.; Bin Mansoor, M. E.; Shaun Lim, W. Z.; Tan, M. H. An Expanded Synthetic Biology Toolkit for Gene Expression Control in Acetobacteraceae. *ACS Synth. Biol.* **2019**, *8* (4), 708–723. https://doi.org/10.1021/acssynbio.8b00168.
(26) Crozet, P.; Navarro, F. J.; Willmund, F.; Mehrshahi, P.; Bakowski, K.; Lauersen, K. J.; Pérez-Pérez, M.-E.; Auroy, P.; Gorchs Rovira, A.; Sauret-Gueto, S. Birth of a Photosynthetic Chassis: A MoClo Toolkit Enabling Synthetic Biology in the Microalga Chlamydomonas Reinhardtii. *ACS Synth. Biol.* **2018**, *7* (9), 2074–2086. https://doi.org/10.1021/acssynbio.8b00251.
(27) Mózsik, L.; Pohl, C.; Meyer, V.; Bovenberg, R. A.; Nygård, Y.; Driessen, A. J. Modular Synthetic Biology Toolkit for Filamentous Fungi. *ACS Synth. Biol.* **2021**, *10* (11), 2850–2861. https://doi.org/10.1021/acssynbio.1c00260.
(28) Solomon, K. V.; Henske, J. K.; Theodorou, M. K.; O'Malley, M. A. Robust and Effective Methodologies for Cryopreservation and DNA Extraction from Anaerobic Gut Fungi. *Anaerobe* **2016**, *38*, 39–46. https://doi.org/10.1016/j.anaerobe.2015.11.008.
(29) Youssef, N. H.; Couger, M. B.; Struchtemeyer, C. G.; Liggenstoffer, A. S.; Prade, R. A.; Najar, F. Z.; Atiyeh, H. K.; Wilkins, M. R.; Elshahed, M. S. The Genome of the Anaerobic Fungus Orpinomyces Sp. Strain C1A Reveals the Unique Evolutionary History of a Remarkable Plant Biomass Degrader. *Appl. Environ. Microbiol.* **2013**, *79* (15), 4620. https://doi.org/10.1128/AEM.00821-13.
(30) Chen, Y.-C.; Liu, T.; Yu, C.-H.; Chiang, T.-Y.; Hwang, C.-C. Effects of GC Bias in Next-Generation-Sequencing Data on De Novo Genome Assembly. *PLOS ONE* **2013**, *8* (4), e62856. https://doi.org/10.1371/journal.pone.0062856.
(31) Orpin, C. The Rumen Flagellate Piromonas Communis: Its Life-History and Invasion of Plant Material in the Rumen. *Microbiology* **1977**, *99* (1), 107–117. https://doi.org/10.1099/00221287-99-1-107.
(32) Michielse, C. B.; Hooykaas, P. J.; van den Hondel, C. A.; Ram, A. F. Agrobacterium-Mediated Transformation as a Tool for Functional Genomics in Fungi. *Curr. Genet.* **2005**, *48* (1), 1–17. https://doi.org/10.1007/s00294-005-0578-0.
(33) Turgeon, B. G.; Condon, B.; Liu, J.; Zhang, N. Protoplast Transformation of Filamentous Fungi. In *Molecular and cell biology methods for fungi*; Springer, 2010; pp 3–19.

(34) Durand, R.; Rascle, C.; Fischer, M.; Fèvre, M. Transient Expression of the B-Glucuronidase Gene after Biolistic Transformation of the Anaerobic Fungus Neocallimastix Frontalis. *Curr. Genet.* **1997**, *31*, 158–161. https://doi.org/10.1007/s002940050190.
(35) Calkins, S. S.; Elledge, N. C.; Mueller, K. E.; Marek, S. M.; Couger, M. B.; Elshahed, M. S.; Youssef, N. H. Development of an RNA Interference (RNAi) Gene Knockdown Protocol in the Anaerobic Gut Fungus Pecoramyces Ruminantium Strain C1A. *PeerJ* **2018**, *6*, e4276–e4276. https://doi.org/10.7717/peerj.4276.
(36) Theodorou, M. K.; Davies, D. R.; Nielsen, B. B.; Lawrence, M. I.; Trinci, A. P. Determination of Growth of Anaerobic Fungi on Soluble and Cellulosic Substrates Using a Pressure Transducer. *Microbiology* **1995**, *141* (3), 671–678. https://doi.org/10.1099/13500872-141-3-671.
(37) Lee, H. B.; Kim, Y.; Kim, J. C.; Choi, G. J.; Park, S.-H.; Kim, C.-J.; Jung, H. S. Activity of Some Aminoglycoside Antibiotics against True Fungi, Phytophthora and Pythium Species. *J. Appl. Microbiol.* **2005**, *99* (4), 836–843. https://doi.org/10.1111/j.1365-2672.2005.02684.x.
(38) Chakraborty, S.; Gogoi, M.; Chakravortty, D. Lactoylglutathione Lyase, a Critical Enzyme in Methylglyoxal Detoxification, Contributes to Survival of Salmonella in the Nutrient Rich Environment. *Virulence* **2015**, *6* (1), 50–65. https://doi.org/10.4161/21505594.2014.983791.
(39) Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otillar, R.; Riley, R.; Salamov, A.; Zhao, X.; Korzeniewski, F.; Smirnova, T.; Nordberg, H.; Dubchak, I.; Shabalov, I. MycoCosm Portal: Gearing up for 1000 Fungal Genomes. *Nucleic Acids Res.* **2014**, *42* (D1), D699–D704. https://doi.org/10.1093/nar/gkt1183.
(40) Korithoski, B.; Lévesque, C. M.; Cvitkovitch, D. G. Involvement of the Detoxifying Enzyme Lactoylglutathione Lyase in Streptococcus Mutans Aciduricity. *J. Bacteriol.* **2007**, *189* (21), 7586–7592. https://doi.org/10.1128/JB.00754-07.
(41) Xu, D.; Liu, X.; Guo, C.; Zhao, J. Methylglyoxal Detoxification by an Aldo-Keto Reductase in the Cyanobacterium Synechococcus Sp. PCC 7002. *Microbiology* **2006**, *152* (7), 2013–2021. https://doi.org/10.1099/mic.0.28870-0.
(42) Varoquaux, N.; Liachko, I.; Ay, F.; Burton, J. N.; Shendure, J.; Dunham, M. J.; Vert, J.-P.; Noble, W. S. Accurate Identification of Centromere Locations in Yeast Genomes Using Hi-C. *Nucleic Acids Res.* **2015**, *43* (11), 5331–5339. https://doi.org/10.1093/nar/gkv424.
(43) Goffeau A.; Barrell B. G.; Bussey H.; Davis R. W.; Dujon B.; Feldmann H.; Galibert F.; Hoheisel J. D.; Jacq C.; Johnston M.; Louis E. J.; Mewes H. W.; Murakami Y.; Philippsen P.; Tettelin H.; Oliver S. G. Life with 6000 Genes. *Science* **1996**, *274* (5287), 546–567. https://doi.org/10.1126/science.274.5287.546.
(44) Russ C.; Lang B. F.; Chen Z.; Gujja S.; Shea T.; Zeng Q.; Young S.; Cuomo C. A.; Nusbaum C. Genome Sequence of Spizellomyces Punctatus. *Genome Announc.* **2016**, *4* (4), e00849-16. https://doi.org/10.1128/genomeA.00849-16.
(45) Baker, S. E.; Schackwitz, W.; Lipzen, A.; Martin J.; Haridas, S.; LaButti, K.; Grigoriev, I. V.; Simmons, B. A.; McCluskey, K. Draft Genome Sequence of Neurospora Crassa Strain FGSC 73. *Genome Announc.* **2015**, *3* (2), e00074-15. https://doi.org/10.1128/genomeA.00074-15.
(46) Shcherbakova, D. M.; Verkhusha, V. V. Near-Infrared Fluorescent Proteins for Multicolor in Vivo Imaging. *Nat. Methods* **2013**, *10* (8), 751–754. https://doi.org/10.1038/nmeth.2521.
(47) Lecoq, J.; Schnitzer, M. J. An Infrared Fluorescent Protein for Deeper Imaging. *Nat. Biotechnol.* **2011**, *29* (8), 715–716. https://doi.org/10.1038/nbt.1941.
(48) Lu, J.; Wu, T.; Zhang, B.; Liu, S.; Song, W.; Qiao, J.; Ruan, H. Types of Nuclear Localization Signals and Mechanisms of Protein Import into the Nucleus. *Cell Commun. Signal.* **2021**, *19* (1), 60. https://doi.org/10.1186/s12964-021-00741-y.
(49) Wilken, St. E.; Seppälä, S.; Lankiewicz, T. S.; Saxena, M.; Henske, J. K.; Salamov, A. A.; Grigoriev, I. V.; O'Malley, M. A. Genomic and Proteomic Biases Inform Metabolic Engineering Strategies for Anaerobic Fungi. *Metab. Eng. Commun.* **2020**, *10*, e00107. https://doi.org/10.1016/j.mec.2019.e00107.
(50) Shitashima, Y.; Shimozawa, T.; Asahi, T.; Miyawaki, A. A Dual-Ligand-Modulable Fluorescent Protein Based on UnaG and Calmodulin. *Biochem. Biophys. Res. Commun.* **2018**, *496* (3), 872–879. https://doi.org/10.1016/j.bbrc.2018.01.134.
(51) Wilson, F. M.; Harrison, R. J. CRISPR/Cas9 Mediated Editing of the Quorn Fungus Fusarium Venenatum A3/5 by Transient Expression of Cas9 and SgRNAs Targeting Endogenous Marker Gene PKS12. *Fungal Biol. Biotechnol.* **2021**, *8* (1), 15. https://doi.org/10.1186/s40694-021-00121-8.

(52) Foster, A. J.; Martin-Urdiroz, M.; Yan, X.; Wright, H. S.; Soanes, D. M.; Talbot, N. J. CRISPR-Cas9 Ribonucleoprotein-Mediated Co-Editing and Counterselection in the Rice Blast Fungus. *Sci. Rep.* **2018**, *8* (1), 14355. https://doi.org/10.1038/s41598-018-32702-w.
(53) Jiang, W.; Brueggeman, A. J.; Horken, K. M.; Plucinak, T. M.; Weeks, D. P. Successful Transient Expression of Cas9 and Single Guide RNA Genes in Chlamydomonas Reinhardtii. *Eukaryot. Cell* **2014**, *13* (11), 1465–1469. https://doi.org/10.1128/EC.00213-14.
(54) Henske, J. K.; Wilken, St. E.; Solomon, K. V.; Smallwood, C. R.; Shutthanandan, V.; Evans, J. E.; Theodorou, M. K.; O'Malley, M. A. Metabolic Characterization of Anaerobic Fungi Provides a Path Forward for Bioprocessing of Crude Lignocellulose. *Biotechnol. Bioeng.* **2018**, *115* (4), 874–884. https://doi.org/10.1002/bit.26515.
(55) Ranganathan, A.; Smith, O. P.; Youssef, N. H.; Struchtemeyer, C. G.; Atiyeh, H. K.; Elshahed, M. S. Utilizing Anaerobic Fungi for Two-Stage Sugar Extraction and Biofuel Production from Lignocellulosic Biomass. *Front. Microbiol.* **2017**, *8*. https://doi.org/10.3389/fmicb.2017.00635.
(56) Yıldırım, E.; Ince, O.; Aydin, S.; Ince, B. Improvement of Biogas Potential of Anaerobic Digesters Using Rumen Fungi. *Renew. Energy* **2017**, *109*, 346–353. https://doi.org/10.1016/j.renene.2017.03.021.
(57) Procházka, J.; Mrázek, J.; Štrosová, L.; Fliegerová, K.; Zábranská, J.; Dohányos, M. Enhanced Biogas Yield from Energy Crops with Rumen Anaerobic Fungi. *Eng. Life Sci.* **2012**, *12* (3), 343–351. https://doi.org/10.1002/elsc.201100076.
(58) Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35* (6), 1547–1549. https://doi.org/10.1093/molbev/msy096.
(59) Price, M. N.; Dehal, P. S.; Arkin, A. P. FastTree: Computing Large Minimum Evolution Trees with Profiles Instead of a Distance Matrix. *Mol. Biol. Evol.* **2009**, *26* (7), 1641–1650.
(60) Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis*; Springer-Verlag New York, 2016.
(61) Borchers, H. *pracma: Practical Numerical Math Functions*. https://CRAN.R-project.org/package=pracma (accessed 2022-08-04).
(62) Schindelin, J.; Arganda-Carreras, I.; Frise, E.; Kaynig, V.; Longair, M.; Pietzsch, T.; Preibisch, S.; Rueden, C.; Saalfeld, S.; Schmid, B.; Tinevez, J.-Y.; White, D. J.; Hartenstein, V.; Eliceiri, K.; Tomancak, P.; Cardona, A. Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* **2012**, *9* (7), 676–682. https://doi.org/10.1038/nmeth.2019.
(63) R Core Team. *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/ (accessed 2022-07-28).
(64) Bodenhofer, U.; Bonatesta, E.; Horejš-Kainrath, C.; Hochreiter, S. Msa: An R Package for Multiple Sequence Alignment. *Bioinformatics* **2015**, *31* (24), 3997–3999. https://doi.org/10.1093/bioinformatics/btv494.
(65) Gentleman, R. C.; Carey, V. J.; Bates, D. M.; Bolstad, B.; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J. Bioconductor: Open Software Development for Computational Biology and Bioinformatics. *Genome Biol.* **2004**, *5* (10), 1–16. https://doi.org/10.1186/gb-2004-5-10-r80.

711 **For Table of Contents Only**



712