Enabling tools for *Toxoplasma* glycobiology

Elisabet Gas-Pascual[1,2,*], H. Travis Ichikawa[1,5,*], M. Osman Sheikh[3], M. Isabella Serji[1,6], Bowen Deng[1,2], Msano Mandalasi[1,2], Giulia Bandini[4], John Samuelson[4], Lance Wells[1,3], and Christopher M. West[1,2,7]

[1]Department of Biochemistry & Molecular Biology, [2]Center for Tropical and Emerging Global Diseases, [3]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602 USA

[4]Department of Molecular and Cell Biology, Henry M. Goldman School of Dental Medicine, Boston University, 72 East Concord Street, Boston, MA 02118 USA

[5] New Materials Institute, University of Georgia, Athens, GA 30602 USA
[6] George Washington University School of Public Health, Washington, D.C. 20009 USA

[7] to whom requests for information or materials should be sent at Dept. of Biochemistry & Molecular Biology, 120 E. Green St., Davison Life Sciences – B129, University of Georgia, Athens GA 30602 USA, telephone +1-706-542-4259, email westcm@uga.edu

* These authors contributed equally to this project

Keywords: CRISPR/Cas, glycomics, *Toxoplasma gondii*, glycosyltransferase, glycobiology, mass spectrometry, parasitology

Abbreviations: AAL, *Aleuria aurantia* agglutinin or lectin; Dol-P-sugar, a conjugate of Glc or Man to dolichol phosphate; EIC, extracted ion chromatogram; FBS, fetal bovine serum; GPI, glycophosphatidylinositol; GT, glycosyltransferase; HAc, acetic acid; HdH, a disaccharide consisting of a hexose and a deoxy-hexose residue; HFF, human foreskin fibroblasts; hTERT, HFF immortalized with human telomerase reverse transcriptase; H6N2, an oligosaccharide consisting of 6 hexose (Hex) and 2 N-acetylhexosamine (HexNAc) residues; NHEJ, non-homologous end-joining; nLC-MS/MS, nano-liquid chromatography coupled to multi-dimensional mass spectrometry; mAb, monoclonal antibody; pp-αGalNAcT, UDP-GalNAc-dependent polypeptide α-N-acetyl-D-galactosaminyltransferase; Tn antigen, a structure corresponding to αGalNAc-S/T; TSR, thrombospondin type 1 repeats

ABSTRACT

Infection by the protozoan parasite *Toxoplasma gondii* is a major health risk owing to its chronic nature, ability to reactivate to cause blindness and encephalitis, and high prevalence in human populations. Like nearly all eukaryotes, *Toxoplasma* glycosylates many of its proteins and lipids and assembles polysaccharides. Unlike most eukaryotes, *Toxoplasma* divides and differentiates in vacuoles within host cells. While their glycans resemble canonical models, they exhibit species-specific variations that have inhibited deeper investigations into their roles in parasite biology and virulence. The genome of *Toxoplasma* encodes a suite of likely glycogenes expected to assemble a range of *N*-glycans, *O*-glycans, a *C*-glycan, GPI-anchors, and polysaccharides, along with their requisite precursors and membrane transporters. To facilitate genetic approaches to investigate the roles of specific glycans, we mapped probable connections between 59 glycogenes, their enzyme products, and the glycans to which they contribute. We adapted a double-CRISPR/Cas9 strategy and a mass spectrometry-based glycomics workflow to test these relationships, and conducted infection studies in fibroblast monolayers to probe cellular functions. Through the validated disruption of 17 glycogenes, we also discovered novel Glc_{0-2}-Man_6-$GlcNAc_2$-type N-glycans, $GalNAc_2$- and Glc-Fuc-type O-glycans, and a nuclear O-Fuc type glycan. We describe the guide sequences, disruption constructs, and mutant strains, which are freely available to practitioners for application in the context of the relational map to investigate the roles of glycans in their favorite biological processes.

INTRODUCTION

Toxoplasma gondii is a worldwide, obligate intracellular apicomplexan parasite that can infect most nucleated cells of warm-blooded animals (Kim and Weiss, 2004). Up to 80% of some human populations are seropositive (Pappas et al., 2009. Toxoplasmosis, the disease caused by *Toxoplasma*, is associated with encephalitis and blindness in individuals whose parasites are reactivated such as in AIDS and other immune-suppressed patients (Pereira-Chioccola et al., 2009). *In utero* infections can cause mental retardation, blindness, and death (Luft and Remington, 1992). *Toxoplasma* is transmitted by digesting parasites from feline feces (as oocysts) or undercooked meat (as tissue cysts). Once in the host, parasites convert to the tachyzoite form that disseminates to peripheral tissues (e.g. brain, retina, and muscle). The resulting immune response and/or drugs can control tachyzoite replication, but the parasite survives by converting into slow growing bradyzoites that encyst. Cysts sporadically burst, and the parasites convert to tachyzoites, whose unchecked growth results in cell and tissue damage (Lambert and Barragan, 2010; Blader et al., 2015). Currently, no *Toxoplasma* vaccine exists, anti-toxoplasmosis drugs have severe side effects, and resistance is developing to these drugs (Katlama et al., 1996; Dannamann et al., 1992; Baatz et al., 2006; Aspinall et al., 2002; Weiss and Kim, 2000). As infected individuals remain infected for life, new anti-*Toxoplasma* drugs are needed, and ideal targets are those present in the parasite but not in the host.

Apicomplexan proteins were for a time considered to be minimally glycosylated, but lectin-, metabolic labeling-, genomic- and mass spectrometry-based studies over the past three decades have refuted the original dogma in *Toxoplasma*, *Plasmodium* spp. (agent of malaria), *Cryptosporidium* spp. (agent of cryptosporidiosis), and others (Rodrigues et al., 2015; Bushkin et al., 2010; Heimburg-Molinaro et al., 2013; Macedo et al., 2010; Sanz et al., 2013). While less diverse than for most other eukaryotes, the major known glycan classes, including *N*-glycans, *O*-glycans, *C*-mannose, GPI-anchors, glycolipids, and polysaccharides, are represented in *Toxoplasma*. *N*-glycans, *O*-glycans, or GPI-anchors have been described on hundreds of *Toxoplasma* proteins (Luo et al., 2011; Fauquenoy et al., 2008; Wang et al., 2016; Bandini et al., 2016). Their functions have been little studied, but glycans have been implicated in invasion, O_2-sensing, cyst wall formation, cyst persistence, and nutrient storage, each of which may be important for virulence (Luo et al., 2011; Debierre-Grockiego et al., 2010; Tomita et al, 2017; Rahman et al., 2016; Caffaro et al., 2013; Uboldi et al., 2015). Many glycosyltransferase genes have low fitness scores when disrupted in a genome-wide survey (Sidik et al., 2016), indicating unknown but vital biological functions for the parasite (Fig. 1).

Glycosylation associated drugs are beginning to make an impact in human health, and there is reason to anticipate success in combating pathogens with their distinctive glycan synthetic pathways. Unique glycoconjugates are prime candidates for vaccine development and as markers for diagnostic monitoring. But before the full potential of these strategies can be realized, more structure-function studies are needed to guide selection of the most promising targets. Progress in exploiting these strategies has been inhibited by the uniqueness of the glycomes of *Toxoplasma* and other parasites, and lack of information about the genetic basis of its non-template driven mechanisms of assembly.

Further understanding of the roles of glycans will benefit from genetic approaches to manipulate their biosynthesis. Sufficient knowledge currently exists to allow informed predictions of glycogenes associated with glycan assembly, and mass spectrometry-based technologies for interpreting glycomic consequences of glycogene perturbations are available. This report describes the generation of a collection of CRISPR/Cas9-associated guide DNAs that are suitable for disrupting glycogenes, using an approach that is expected to be suitable for strains regardless of their NHEJ (Ku80) status, and presents the results of trials on the type 1 RH strain. A streamlined mass spectrometry workflow was implemented to benchmark the new studies, which led to some revisions of existing models and new structures including novel N-glycans, a novel mucin-type O-glycan, and a novel nuclear O-Fuc. Finally, the glycogene mutants were screened in a cell culture growth assay for information about their physiological significance. It is expected that the availability of these findings and associated resources will help stimulate interrogation of the role of glycosylation in future functional studies.

EXPERIMENTAL PROCEDURES

Glycogene prediction. To search for candidate glycosyltransferases (GTs) that contribute to the *Toxoplasma* glycome, we expanded approaches previously implemented in *Dictyostelium discoideum* (West et al., 2005) and *D. purpureum* (Sucgang et al., 2011). The predicted proteome (8460 proteins) from Toxodb.org was analyzed by i) the SUPERFAMILY server, which assigns protein domains at the SCOP 'superfamily' level using hidden Markov models; ii) dbCAN, an automated Carbohydrate-active enzyme ANnotation database which utilizes a CAZyme signature domain-based annotation based on a CDD (conserved domain database) search, literature curation, and a hidden Markov model; iii) the Pfam database; and iv) Bernard Henrissat (Marseille) who supervises the CAZy database (Lombard et al., 2014). This yielded 39 GT-like coding regions. No gene candidates for DNA glucosylation as characterized in trypanosomatids (Bullard et al., 2014) are apparent in apicomplexans. The search was expanded to include genes predicted to contribute to the synthesis of sugar nucleotide precursors and their transport into the secretory pathway, genes that contribute to N-glycan processing, and genes associated with processing of starch, yielding 28 additional glycogenes.

Parasites and cell culture

Human foreskin fibroblasts (HFF, ATCC SCRC-1041) and HFF immortalized with human telomerase reverse transcriptase (hTERT, ATCC CRL-4001) were maintained in Complete Medium, which consisted of Dulbecco's Modified Eagle Medium (DMEM, Corning Inc., 15-1013-CV) supplemented with 10% (v/v) fetal bovine serum (FBS, Corning, 35-015-CV), 2 mM L-glutamine (Corning, 25-005-CI), and 100 IU/ml penicillin/100 µg/ml streptomycin (Corning, 30-002-CI). *Toxoplasma gondii* strains RH and Pru, and their derivative strains, were routinely maintained on hTERT cells in Complete Medium containing 1% (v/v) FBS in a humidified incubator with 5% CO_2 at 37°C.

Intracellular parasites were prepared by forcing well-infected fibroblast cells through a 27-gauge needle attached to 10 ml syringe 5 times for hTERT cells and 3 times for HFF cells and collected by centrifugation at 2000 × *g* for 8 min at 4°C.

Parasite freezer stocks were prepared from a T25 flask culture of hTERT or HFF monolayer that were 30 to 50% lysed. Scraped cells were collected and pelleted as described above and resuspended in 1-1.5 ml ice-cold DMEM:FBS (1:1). After incubating 10 min on ice, an equal volume of ice-cold DMEM:DMSO (Sigma, D2650) in a 4:1 ratio was added dropwise to the cell suspension. Cell suspensions were slowly frozen in 1-ml cryo-vials at -80°C and transferred to liquid N_2.

Dual-guide plasmid construction and DHFR cassette amplification

Plasmid 2 and Plasmid 3 were derived from pU6-Universal (Sidik et al., 2016; Addgene plasmid #52694) by the introduction of a novel XhoI site either upstream of the U6 promoter, or between the guide RNA scaffold and the NsiI site, respectively. Plasmid 2 (p2) was constructed by replacing the PvuI-PciI fragment of pU6-Universal with a new PvuI-PciI fragment generated by PCR of pU6-Universal with Plasmid 2 FOR and Plasmid 2 REV (which contains an XhoI site) (Table S3), using pU6-universal as the template. Similarly, Plasmid 3 (p3) was constructed by replacing the HindIII-NsiI fragment with the HindIII-NsiI fragment of a PCR amplicon generated using Plasmid 3 FOR and Plasmid 3 XhoI (which contains an XhoI site). All enzymatic manipulations of plasmid DNA were performed as described by the manufacturer, and all PCR amplification reactions were performed with Q5 DNA polymerase (NEB, M0491L). p2 and p3 were prepared in *dam⁻/dcm⁻* Competent *E. coli* (NEB, C2925I). Guide sequences were selected using the Eukaryotic Pathogen CRISPR guide RNA/DNA Design Tool (EuPaGDT; Ranzinger et al., 2017; http://grna.ctegd.uga.edu) based on the calculated efficiency score against GT1 genome (and/ or Me49 genome) in the latest available version of ToxoDB (http://toxodb.org/toxo/). Guide sequences for each gene were cloned into p2 and p3, as described in Addgene protocol with minor modifications. Briefly, annealed synthetic "top" and "bottom" ssDNA guide oligonucleotides flanked by BsaI compatible 4-nt overhangs were directionally cloned and insertion was confirmed by PCR reactions using the "top" strand guide oligonucleotide and a downstream primer, Plasmid 1 REV. Dual-guide plasmids (pDG) were constructed

by transferring the XhoI-NsiI fragment of p2 into p3 that was double digested with XhoI and NsiI (Fig. 2). Plasmids were prepared in transformed Top10 cells (Thermo Fisher Scientific) using a ZR Plasmid Miniprep-Classic (Zymo Research, D4016) kit. Finally, the inserted guide sequences were confirmed by sequencing the pDG plasmid with primers gRNA For and gRNA Rev (Table S3). All available plasmids are referred to in Table S1 and defined in Table S4.

The DHFR expression cassette amplicon was prepared using primers (DHFR-Fw or DHFR-For LoxP2 and DHFR-Rv) (Table S3) and purified using a DNA Clean & Concentrator-5 kit (Zymo Research, D4003). For transfection, pDG and the DHFR expression amplicon were EtOH-precipitated, rinsed with 70% EtOH, and resuspended in Cytomix Buffer (120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4, 25 mM HEPES-NaOH, 2 mM EDTA, 5 mM $MgCl_2$, final pH 7.6).

Transfection and cloning

Transfection was performed as described (Striepen and Soldati, 2007). Briefly, 10 million freshly prepared intracellular parasites were electroporated with 1 µg DHFR expression amplicon ± 10 µg pDG, or no DNA, in Cytomix Buffer plus 100 mM each of ATP and glutathione in a final vol of 400 µl in a 2 mm-wide cuvette (Bulldog BIO, 12358-346), at a setting of 1.5 kV, 25 µF in a Gene Pulser Xcell (Bio-Rad). After 10 min of incubation at room temperature, treated parasites were diluted into 5 ml of Complete Medium with 10% FBS and inoculated onto confluent HFFs in a T25 flask. After standard incubation overnight, the medium was replaced with 5 ml of Complete Medium with 10% FBS containing 1 µM pyrimethamine (Sigma, 46703) for selection. If the majority of HFF cells started lysing before the mock-transfected parasites died, the parasites were transferred to another T25 flask with hTERT monolayer for further selection in Complete Medium with 1% FBS containing 1 µM pyrimethamine until no more visibly viable mock-transfected parasites were observed.

Pyrimethamine-resistant clones were isolated by limiting dilution on confluent HFF in Complete Medium with 1% FBS containing 1 µM pyrimethamine. Six serial two-fold dilutions starting with 25 parasites per well were prepared in octuplet in a 96-well plate and grown for 6 to 10 d. Single clones were expanded in hTERT monolayers in T25 flasks for genomic DNA isolation and freezer stock preparation.

Analysis of *Toxoplasma gondii* glycogene loci

Genomic DNA was prepared from extracellular parasites using a Quick DNA Miniprep Plus kit (Zymo Research, D4069), eluted in 100 µl of DNA Elution Buffer and stored at 4°C in the DNA Elution Buffer. Q5 DNA polymerase was used to PCR amplify 10-20 ng of genomic DNA with primer sets flanking expected DSB sites (Table S2) and near the ends of DHFR amplicon (Table S3) to characterize the glycogene loci disruption.

Plaque assay

Plaque assays were performed on 100% confluent, tightly packed HFF monolayers under 3 ml Complete Medium, 1% FBS without pyrimethamine in a 6-well plate. Freshly lysed-out intracellular parasites were inoculated at 100 cells/well, and the medium was replaced after 3 h by aspiration. On d 6, cultures were aspirated and rinsed 3× with 3 ml PBS (Corning, 21-040-CV) before fixation with -20°C 100% methanol for 5 min followed by air drying. The monolayer was stained with 1 ml 2% (w/v) crystal violet, 0.8% ammonium oxalate (w/v) in 25% ethanol for 5 min at room temperature, rinsed 3× with PBS, and air dried. Mutant and control (RH) strains were always compared on the same plate. Monolayers were scanned by Odyssey CLx Imaging System (Li-Cor) and plaque areas were measured using ImageJ software (National Institute of Health, USA). Mutant plaque areas were referenced to control plaque areas from the same plate, and all control plaque areas were normalized to a value of one for presentation. Data were presented and statistically analyzed using GraphPad Prism version 6.

Preparation of parasites

The type 1 RH strain (w/t) and mutant strains were grown on confluent hTERTs in Complete Medium supplemented with 1% FBS. The medium was collected after 36-48 h, and spontaneously lysed-out

tachyzoites were harvested by centrifugation (8 min, 2000 × *g*, room temperature). The pellets were syringe-lysed (through a 27G needle) three times, filtered through a 3 µm Nuclepore® filter (Whatman) followed by 2 rinses of DMEM to remove host cell debris. Pellets were resuspended and centrifuged 3× with ice-cold PBS, frozen in liquid N_2, and stored at -80°C. 125 mg wet weight of cells (~5 × 10^9 cells) were typically recovered from 14 T175 flasks.

Preparation of delipidated total cell pellets and Folch-partition

Cell pellets (~5 × 10^9 cells) were delipidated, and total protein powder was prepared as described (Aoki et al., 2007). Briefly, cell pellets were disrupted by Dounce homogenization in a mixture of ice-cold MeOH and H_2O. The resulting suspension was transferred to a screw-top 8-ml glass tube and adjusted with water and $CHCl_3$ to a final ratio of $CHCl_3$:MeOH:H_2O (4:8:3;v:v:v). The extract was incubated for 3 h at room temperature with end-over-end agitation. The insoluble proteinaceous material was collected by centrifugation and re-extracted three times in chloroform/MeOH (2:1;v:v) and twice in $CHCl_3$:MeOH:H_2O (4:8:3;v:v:v). The final pellet was dried under a stream of N_2, washed with acetone, dried, weighed (typically ~25-30 mg), and stored at −20°C until analysis. The clarified lipid extract was subjected to Folch-partitioning (Folch et al., 1957). The resulting aqueous fraction (GIPLs) and organic fraction (other lipids) were dried under a stream of nitrogen and stored separately at -20°C.

Enrichment in GPI-anchored proteins and GPI-glycan core release

GPI-anchored proteins were enriched from delipidated pellets by partitioning in butan-1-ol (n-BuOH, 99%; Alfa Aesar, Tewksbury, MA) as described (Almeida et al., 1994; Koeller et al., 2014). Briefly, 20 mg of protein powder was extracted 3× with 4 ml H_2O saturated with n-BuOH (~9%) for 4 h at room temperature. The n-BuOH-saturated water fractions were pooled and extracted with an equal vol of n-BuOH. After centrifugation for 15 min at 3500 × *g* at room temperature, upper butanolic and lower aqueous phases were separately dried and stored at -20°C. Butanolic fractions, enriched in GIPLs, were saved for future analysis. Aqueous fractions, containing GPI-anchored proteins, were treated with 100 µl 0.5 M NH_4OH (in H_2O) for 6 h at 4°C to release fatty acids. The sample was then dried under a N_2-stream, redissolved in H_2O, and dried again twice. The glycan core was released by treatment with 50 µl 48% aqueous HF at 4°C for 12 h (Masuishi et al., 2016), resuspended in 500 µl 0.05% trifluoroacetic acid (ThermoFisher Scientific, 28904) in H_2O, filtered through a C_{18}-SepPak to capture protein and lipid, and dried under vacuum. The sample was subjected to N-acetylation by addition of 10 µl of pyridine (Sigma, 360570) and 50 µl of acetic anhydride (Sigma, 539996) in 0.5 ml of MeOH, incubation for 30 min at room temperature (Kozulic et al., 1979), and drying under a stream of N_2.

Preparation of *N*-linked glycans

N-linked glycans were prepared as described (Aoki et al, 2007). 2 to 5 mg of protein powder was resuspended in 450 µl of trypsin digestion buffer (20 mM NH_4HCO_3, natural pH, 1 M urea) with brief bath sonication. 25 µl each of 20 ng/µl trypsin (Sigma, T8003) and 10 µg/µl chymotrypsin (Sigma, C4129) was added at a final ratio ~1:20 protease:protein and incubated for 18 h at 37°C. The reaction was boiled for 5 min, cooled and centrifuged to remove insoluble material. The supernatant was adjusted to 5% (v/v) acetic acid (HAc) and loaded onto a Sep-Pak C_{18} cartridge (Sep-Pak® Vac 1cc C_{18} cartridge, Waters, Milford, MA), that had been pre-equilibrated by washing with >3 ml of 100% acetonitrile and >3 ml of 5% HAc, and washed with 10 column vol of 5% HAc. Glycopeptides were eluted stepwise, as described (Alvarez-Manilla et al., 2010), with 2 vol 20% 2-propanol in 5% HAc, 2 vol 40% 2-propanol in 5% HAc and 3 vol 100% 2-propanol. The 2-propanol fractions were pooled, evaporated to dryness by vacuum centrifugation, resuspended in 50 µl of 20 mM Na phosphate buffer (pH 7.5), and digested with PNGaseF (NEB, P0705S) for 18 h at 37°C. Reaction mixtures were evaporated to dryness, resuspended in 500 µl 5% HAc and loaded onto a Sep-Pak C_{18} cartridge column. The column run-through and an additional wash with 3 column vols of 5% HAc, containing released oligosaccharides, were collected together and evaporated to dryness. Core α3-linked Fuc, which interferes with PNGaseF action, has not been detected on *Toxoplasma N*-glycans consistent the absence of candidate α3FucT encoding genes in the genome.

For samples to be treated with α-mannosidase, the PNGaseF digestion was performed instead in 25 mM NH_4HCO_3 (pH 7.8) and taken to dryness. The sample was resuspended in either 50 µl of ~8 U/ml of Jack Bean α-mannosidase (Prozyme, GKX-5010) in 20 mM NaAc (pH 5.0), 0.4 mM Zn^{2+}, or of ~5 mU/ml of *Aspergillus satoi* α-mannosidase (Prozyme, GKX-5009, in 100 mM NaAc (pH 5.0), and digested for 18 h at 37°C. The reaction mixture was evaporated to dryness, resuspended in 500 µl 5% HAc, purified through a Sep-Pak C_{18} cartridge, and again evaporated to dryness.

Preparation of *O*-linked glycans

O-linked glycans were obtained by reductive β-elimination as described previously (Aoki et al., 2008) with minor modifications. Briefly, ~5 mg of protein powder was resuspended with bath sonication in 500 µl of 50 mM NaOH (diluted from 50% stock; Fisher Scientific, SS254-500) containing 1 M $NaBH_4$ (Sigma, 213462-25G), and incubated for 18 h at 45°C in an 8-ml glass tube sealed with a Teflon-lined screw top. The reaction mixture was then neutralized with 10% HAc and desalted by applying to a 1-ml bed of Dowex 50W-X8 (H^+ form; Sigma, 217506) and washing with 5 bed volumes of 5% HAc. The flow-through fraction, containing released oligosaccharide alditols, was evaporated to dryness. Borate was removed as an azeotrope with MeOH by adding 0.5 ml of 10% HAc in MeOH, drying under a nitrogen stream at 42°C, and repeating three additional times. To remove residual peptide and reagent contaminants, the dried material was resuspended in 500 µl of 5% HAc and loaded onto a C_{18}-SepPak as above. Released oligosaccharide alditols were recovered by collecting the column run-through and an additional 2 ml of wash with 5% HAc. The run-through and wash were combined and evaporated to dryness.

Mass spectrometry of permethylated glycans – direct infusion

The released glycans from the original 5-20 mg of protein powder were permethylated, as described (Anumula and Taylor, 1992), and the final CH_2Cl_2 phase was dried down under N_2 and redissolved in 100 µl of MeOH. 10 µl of the glycan sample was combined with 35 µl of MS Buffer (1 mM NaOH in 50% MeOH) and 5 µl of ^{13}C-permethylated isomaltopentaose (DP5), as an internal standard at a final concentration of 0.2 pmol/µl. Samples were directly infused (1 µl/min) into an Orbitrap mass spectrometer (LTQ Orbitrap XL or Orbitrap Elite, ThermoFisher Scientific) using nanospray ionization in positive ion mode, with the capillary temperature set to 200°C. The Full FTMS was collected and the Top 10 most intense peaks were selected for MS^2 fragmentation using collision-induced dissociation (CID) at 40% collision energy. Once an ion was fragmented, the *m*/*z* was automatically put on an 'exclusion list' to prevent re-analysis in the run. MS^2 product fragments matching a preselected list (Table S5) were selected for MS^3 CID fragmentation at the same collision energy.

Mass spectrometry of permethylated glycans – nLC-MS/MS

10 µL of permethylated glycans from above were combined with 2 µL of an internal standard mix [^{13}C-permethylated isomaltose series (DP4, DP5, DP6, and DP7) at a final concentration of 0.2 pmol/µl each] and 38 µL of LC-MS Buffer A (1 mM LiAc, 0.02% HAc). 5 µl was injected into a PepMap Acclaim analytical C18 column (75 µm × 15 cm, 2 µm pore size) maintained at 60°C in an Ultimate 3000 RSLC (ThermoFisher Scientific/Dionex), which was coupled to a ThermoScientific Velos Pro Dual-Pressure Linear Ion Trap mass spectrometer. For *N*-Glycans, after equilibrating the column in 99% LC-MS Buffer A and 1.5-min ramp up to 45% LC-MS Buffer B [80% (v/v) acetonitrile, 1 mM LiAc, 0.02% HAc], separation was achieved using a linear gradient from 45% to 70% Buffer B over 150 min at a flow rate of 300 nl/min. For *O*-Glycans, after equilibrating the column in 99% LC-MS Buffer A and a 1.5-min ramp to 30% Buffer B, separation was achieved using a linear gradient from 30% to 70% Buffer B over 150 mins at a flow rate of 300 nl/min. The column was regenerated after each run by ramping to 99% Buffer B over 6 min and maintaining at 99% Buffer B for 14 min, before a 1-min ramp back down to 99% Buffer A. The effluent was introduced into the mass spectrometer by nanospray ionization in positive ion mode via a stainless-steel emitter with spray voltage set to 1.8 kV and capillary temperature set at 210°C. The MS method consisted of first collecting a Full ITMS (MS^1) survey scan, followed by MS^2 fragmentation of the Top 3 most intense peaks using CID at 42% collision energy and an isolation window of 2 *m*/*z*. Dynamic

exclusion parameters were set to exclude ions for fragmentation for 15 s if they were detected and fragmented 5 times within 15 s.

Glycan annotation

Structural assignments for the glycans detected at the reported *m*/*z* values were based on the compositions predicted by the exact mass (±0.1 amu) of the intact molecule, the presence of diagnostic MS^2 and MS^3 fragment ions that report specific glycan features, and the limitations imposed on structural diversity by known glycan biosynthetic pathways. The GRITS toolbox and its GELATO annotation tool (Ranzinger et al., 2017) were used to help identify glycan compositions using customized glycan databases and with the following tolerance settings: MS^2 accuracy 500 ppm, max number of adducts (charge states) = 3.

Western blotting and immunofluorescence analysis.

Whole cell pellets (3×10^6 cells/lane) were solubilized in Laemmli sample buffer containing 50 mM dithiothreitol, subjected to SDS-PAGE on 4–12% preformed gels in MES buffer (NuPAGE Novex, Invitrogen), and Western blotted onto nitrocellulose membranes using an iBlot system (Invitrogen). Blots were blocked with 5 mg/ml bovine serum albumin in PBS. For Tn antigen, blots were probed with anti-Tn mAbs 1E3 and 5F4 (Steentoft et al., 2011), which were validated for specificity by comparison of reactivity with wild-type and β3GalT (Core 1)-KO mouse colon samples generously provided by Lijun Xia (OMRF), and a 1:10,000-fold dilution of Alexa-680-labeled rabbit anti-mouse IgG secondary antibody (Invitrogen). For *O*-Fuc, blots were probed with biotinylated *Aleuria aurantia* lectin (AAL) and ExtrAvidin HRP (Sigma) and imaged on an ImageQuant LAS4000 imager (GE Healthcare) as previously described (Bandini et al., 2016). Blots were imaged on a Li-Cor Odyssey infrared scanner and processed in Photoshop, with contrast maintained at γ=1.

Parasites were analyzed by immunofluorescence with AAL as previously described (Bandini et al., 2016).

RESULTS AND DICUSSION

Predicted glycogenes. Our algorithms predicted 38 GT-like genes expected to contribute directly to the assembly of *N*-, *O*- and *C*-glycans on proteins, GPI anchors and other glycolipids, starch, and a wall polysaccharide. An additional five genes are likely to be involved with starch assembly and disassembly, one in GPI processing, three in *N*-glycan processing, 16 in the assembly of sugar nucleotide precursors from simple sugars, and three in transport of sugar nucleotides into the secretory pathway. Table 1 summarizes characteristics of the candidates, organized according to the class of glycoconjugates for which they are expected to contribute, along with findings from the literature and studies described below that provide support for the predicted functions. All but eight of the genes are correlated graphically with known or predicted glycosylation pathways in Figs. 4-8.

For an initial indication of the cellular roles of the glycogenes, we analyzed data from a published genome-wide CRISPR gRNA-based screen for genes that contribute to fitness in a fibroblast monolayer infection setting (Sidik et al. 2016). The change of representation of guide DNA sequences in the population after three rounds of invasion and growth was translated into a fitness score based on a $\log_2$ scale. An analysis of the 68 glycogenes in Table 1 revealed a range of fitness values (Fig. 1) that roughly superimposed on the range of fitness values of all genes. 19 had fitness values <-2.5, suggesting that they may be important for viability. This group was enriched for predicted functions in GPI-anchor assembly, and deficient in genes associated with *O*-glycosylation in the secretory pathway and glycosylation in the nucleocytoplasmic compartment.

Glycogene disruption strategy. To associate glycogenes with assembly of the glycome and cellular activities, we investigated consequences of glycogene disruption. As summarized in Fig. 2, we implemented a variation of the double CRISPR/Cas9 approach described by Shen et al. (2014), in which two gRNAs are employed with the intent of replacing most of the coding region with the pyrimethamine resistance marker in natively Ku80-positive parasite strains (Long et al., 2016). gRNA sequences were chosen using the

Eukaryotic Pathogen CRISPR guide RNA/DNA Design Tool (Peng and Tarleton, 2015). We modified a previously constructed transient expression plasmid designed to express a single gRNA sequence and the Cas9 endonuclease to express 2 gRNAs as shown in Fig. 2E. Transient expression of the plasmid, together with a universal DHFR amplicon prepared by a simple PCR reaction, results in two double stranded breaks that can be repaired via non-homologous end joining (NHEJ) by the insertion of an amplicon encoding DHFR cassette. Treatment with pyrimethamine yields parasites that have incorporated the amplicon, many of which have insertions at the Cas9 cleavage sites. Disruption was confirmed by PCR experiments utilizing oligonucleotide primer pairs that flank both gRNA target sites, and primer pairs that bridge from outside the expected deletion and either end of the DHFR cassette, as outlined in Fig. 2F. This approach has the benefit of using a universal DHFR amplicon and does not stress the cells due to the absence of NHEJ (Ku80), which is commonly used for molecular biology studies of *Toxoplasma*. In addition, this method can in principle be applied to any wild-type clinical isolate regardless of the family type. We applied that double-CRISPR/Cas9 disruption method to selected genes representing each of the major glycan classes but whose fitness scores (Fig. 1) did not suggest inviability, as summarized in Table 1.

Of the 68 glycogene candidates, dual gRNA plasmids were constructed for 30, and single gRNA plasmids were generated for 4 others (summarized in Table S1). Of the 30 for which pDG constructs were available, 17 were used to modify the target gene, and 8 were unsuccessful, as determined by PCR studies (Table S1). Five genes were not tested, usually because mutants became available from other studies during the course of this project. Of the eight unsuccessful trials, each was independently repeated, with the same outcome, and three were re-tried with new pairs of gRNAs, with the same negative outcome (Table S1). This was unexpected based on fitness scores. Of the 17 successful disruptions, ten were full deletions engaging both gRNA sites, and seven resulted in effects at only one of the gRNA sites. Fifteen of these were examined glycomically to address their function and to assess whether the genetic disruption resulted in a loss of function (see below). They were also examined for cell level effects using a plaque assay in the next section.

Each transfection generated a spectrum of genetic lesions, which were evaluated by PCR trials involving primers that hybridized with gene-specific flanking sequences and DHFR specific sequences. Clones were classified based on evidence for a double cut, in which the intervening coding region was deleted and replaced with the DHFR cassette and/or other DNA, and single cut at either the 5'- or 3'-gRNA site and insertion of the DHFR and/or other DNA, or only an improper repair with no introduced DNA, as summarized for each gene in Table S1. Most if not all clones analyzed for individual genes showed evidence consistent with incorporation of the DHFR resistance cassette at one or both of the double stranded break sites. An expanded analysis of the results from three representative gene loci (*galE*, *spy* and *pofut2*) is presented in Figures S2-S4.

Assessment of cellular consequences. To examine cellular effects of disrupting a glycogene, we employed a 6-d monolayer infection model, in which spreading infection from an initial single cell generates a lytic plaque, whose area is a measure of successful cycles of parasite invasion, proliferation and egress. As shown in Fig. 3, nine of the 13 strains examined exhibited significant deficiency in this setting. These included mutations predicted to affect *N*- or *O*-glycosylation in the secretory pathway, *O*-glycosylation in the nucleocytoplasm, or sugar-nucleotide precursor supply, as detailed below. The effects of nine of the disruptions reasonably correlated with fitness scores from Sidik at al. (Fig. 1), but diverged for four genes. The effects of three disruptions were more severe than observed in the fitness assay. Though this might be attributable to unknown spurious genetic changes, it is notable that these three genes are associated with protein fucosylation: synthesis of GDP-Fuc or its utilization by the Spy or POFUT2 glycosyltransferases (see below). Thus the variations might be due to the different parasite activities assayed as, in contrast to the plaque assay, the fitness assay did not test parasite egress from host cells. Though these correlations strongly implicate the importance of different glycosylation pathways for overall parasite success in cell culture, follow-up studies beyond the scope of this resource will be required to confirm causality for individual glycogenes.

Assessment of glycomic consequences. In order to assess the biochemical consequences of glycogene disruption, we adapted established methods to evaluate the parasite glycome. Since the parasite grows within human host cells, which are themselves reared in the presence of bovine serum, measures are required to minimize contamination with human and bovine glycans. Thus we allowed tachyzoites to spontaneously lyse out of fibroblast monolayers in low (1%) serum medium, and washed the parasites by centrifugation with serum-free medium. Since glycans are found in glycolipids and glycoproteins, and as polysaccharides, which vary in their chemical characteristics, we adapted a classical scheme of Folch extraction to generate fractions enriched in glycolipids, GPI-anchored glycoproteins, and other glycoproteins, and then applied chemical or enzymatic methods to release glycans from these fractions to be analyzed by mass spectrometry (Fig. S5). The various chemical linkages of glycans to protein amino acids necessitate specialized methods for their release. In general, *N*-glycans are released enzymatically using PNGaseF, *O*-glycans are released chemically using NaOH to induce β-elimination, and GPI-anchor glycans are released using aqueous HF that cleaves phosphodiester linkages. For monosaccharide glycans that are too short to recover by current methods, we Western-blotted the carrier glycoproteins with antibodies (Abs) and lectins; alternatively, glycans might be analyzed as glycopeptides by MS after trypsinization.

***N*-glycans.** *N*-glycans are present on dozens if not hundreds of proteins that traverse the *Toxoplasma* secretory pathway (Fauquenoy et al., 2011; Fauquenoy et al., 2008; Luo et al., 2011), though the density of *N*-sequons to which they are linked is less than in most eukaryotes (Bushkin et al., 2010). We identified genes expected to assemble a $Glc_3Man_5GlcNAc_2$-P- moiety on the P-Dol precursor (Table 1), in accord with previous studies (Samuelson et al., 2005). *N*-glycans are typically extensively processed after transfer to the protein, but only enzymes that remove the Glc residues were detected in *Toxoplasma* (Table 1). Notably, the processing α-mannosidases and UDP-Glc-dependent α-glucosyltransferase, found in most eukaryotes, are absent. Together with the absence of other genes likely due to secondary loss, *N*-glycans are not expected to have a role in chaperone assisted protein folding and quality control (Banerjee et al., 2007; Samuelson and Robbins, 2015). Studies indicate that *N*-glycosylation is important for parasite motility and invasion (Fauquenoy et al., 2011; Luk et al., 2008), consistent with the low fitness scores (<-2.5) for 4 of the 5 enzymes that assemble the $Man_5GlcNAc_2$- core (Fig. 4A).

nLC-MS(n)-based glycomic profiling was employed to interrogate consequences of *N*-glycogene disruptions. Analysis of unfractionated tachyzoites identified three prominent *N*-glycan-like species: H8N2, H7N2 and H6N2 (Fig. 4B), and other minor species including H9N2, H5N2 and H4N2 (Fig. S6A), where H indicates the number of hexose and N the number of HexNAc residues. H8N2 was initially expected to represent $Glc_3Man_5GlcNAc_2$, with H7N2 and H6N2 being processed forms lacking one or two Glc residues. The absence of complex, hybrid and paucimannose *N*-glycans confirmed minimal contamination from host or serum glycans (Fig. S6B). However, the H6N2 species elutes at the same time both in RH and in a mutant disrupted in *alg5* (Fig. 4), the enzyme that generates the Dol-P-Glc precursor required for glucosylation. H7N2 and H8N2 were absent from *alg5*Δ cells, suggesting that they represented H6N2 containing one or two Glc caps, respectively. To address the identity of the Hex residues, the samples were treated with jack bean α-mannosidase, which cleaves α-1-2/3/6-linked mannoses. H6N2 from both RH (parental) and *alg5*Δ cells, which are likely identical owing to co-elution of both of their anomers from the C18 column, were both converted to H1N2 (Fig. 4B, C), indicating that the extra Hex residue is an αMan. Furthermore, α2-mannosidase converted the H6N2 to H3N2, indicating that the αMan is 2-linked. After treatment with either α-mannosidase, H8N2 appeared to convert to H7N2 (similar elution time as the original H7N2), and H7N2 to convert to H6N2 (different elution time compared to original H6N2), which suggests that the extra α2Man occurs in both of these species. This conclusion is based on the assumption that jack bean α-mannosidase, at the enzyme concentration used, can only access the α6Man residue after the α3-arm is partially trimmed, as can only occur in the absence of the αGlc caps.

In this new model, we propose that *Toxoplasma* assembles an unusual $Glc_3Man_6GlcNAc_2$-linked precursor (Fig. 4A), and that the 6th Man residue is α2-linked to the α6Man, though linkage to the α3Man or to the non-reducing end α2Man cannot be excluded. The responsible αManT is unknown, but might be

Alg11, which applies the other two α2Man before the precursor flips from the cytoplasm to the lumen of the rER, or might be attached after flipping by the GPI-anchored pigB, which also applies an α2Man to an α6Man. This model is consistent with $Hex_9GlcNAc_2$ being the largest and most abundant lipid-linked oligosaccharide in *Toxoplasma* (Garénaux et al., 2008). After transfer to protein by the oligosaccharyltransferase (Shams-Eldin et al., 2005), partial trimming down to H6N2 is consistent with evidence for α-glucosidase-I and α-glucosidase-II genes (Fig. 4A, Table 1A). The minor H9N2, which was inconsistently observed, is evidently converted to H8N2 by jack bean α-mannosidase, and the minor H5N2 and H4N2 species were eliminated by treatment with jack bean α-mannosidase, indicating that they are canonical $Man_5GlcNAc_2$ and $Man_4GlcNAc_2$ species that might have resulted from processing by lysosomal hydrolases, as described in *Dictyostelium* (Hykollari et al., 2014). We detected no effect of *alg10*Δ, expected to assemble the 3rd, α2-linked, Glc, consistent with its efficient removal by α-glucosidase-1 (Fig. 4A), as expected. Our analysis of the putative *alg8*-mutant yielded a reduced level of H8N2 relative to H7N2 and H6N2 and a small plaque phenotype (Fig. 3). This difference suggested the genetic lesion, which consisted of an insertion of DHFR in a poorly conserved region of the 5'-end of the protein coding region, is likely a hypomorphic rather than a null allele. We found no evidence for host derived *N*-glycans (Garénaux et al., 2008), but this could be the result of our use of spontaneously lysed extracellular parasites, which may have turned over host-derived glycans.

Thus, *Toxoplasma* proteins appear to be decorated with just three major species of *N*-glycans, each with an unexpected and unusual structure: $Glc_2Man_6GlcNAc_2$, $Glc_1Man_6GlcNAc_2$, and $Man_6GlcNAc_2$ (Fig. 4A). *N*-glycans of *Cryptosporidium*, a related apicomplexan, include just two forms, $Hex_6HexNAc_2$ and $Hex_5HexNAc_2$, each of which lacks an extension on the short mannose arm and so likely represents $Glc_1Man_5GlcNAc_2$ and $Man_5GlcNAc_2$, respectively (Haserick et al., 2017). Failure to trim the Glc residues is consistent with the absence of the quality control mechanism, in which they participate in other eukaryotes (Bushkin et al., 2010; Samuelson and Robbins, 2015). Attempts to genetically disrupt assembly of the $Man_6GlcNAc_2$- portion of the *N*-glycan precursor were unsuccessful, consistent with essentiality of its highly conserved $Man_5GlcNAc_2$ core. Disruption of assembly of the terminal αGlc residues, which are retained on many *N*-glycans, was deleterious for the parasite. Thus the αGlc termini contribute functionality to *Toxoplasma N*-glycans, which can in the future be investigated using the respective pDG plasmids described here.

***O*-glycan glycogenes associated with the secretory pathway.** *O*-glycans are also found on dozens if not hundreds of glycoproteins that are processed via the parasite rER and Golgi based mainly on lectin capture studies (Luo et al., 2011; Wang et al., 2016). The genome encodes five pp-αGalNAcTs (Table 1C and Fig. 5A) that initiate mucin-type *O*-glycan assembly (Stwora-Wojczyk et al., 2004), i.e., glycans that are linked via αGalNAc-residues to Thr/Ser residues that are often clustered in hydrophilic sequence repeats. In addition, two genes (Table 1C, Fig. 5B) are predicted to assemble a second type of *O*-glycan (Glcβ1,3Fuc1α-) involved in quality control of thrombospondin type 1 repeat-domains (TSR) and protein-protein interactions in animals (Vasudevan and Haltiwanger, 2014). Four Golgi-associated GTs of unknown function are also predicted (Table 1E and Fig 5D). With fitness scores ranging from -1.9 to 1.4 (Table 1), the contributions of these genes to tachyzoite biology is predicted to be modulatory rather than essential for viability. Mutational analyses of the pp-αGalNAcTs and a sugar nucleotide transporter that may selectively impact *O*-glycan assembly, indicate functional roles in the integrity of semi-dormant bradyzoite cyst walls (Tomita et al., 2017), which contains mucin type glycoproteins including Cst1 (Tomita et al., 2013), and tissue cyst persistence in mice (Caffaro et al., 2013).

Little is known about *Toxoplasma O*-glycans, so we initiated a global *O*-glycomic analysis of material released from total tachyzoite protein powder by classical reductive β-elimination (Fig. S7A). Remarkably, only two *O*-glycan types could be confidently assigned, with compositions of N2 (HexNAc2) and HdH (Hex-deoxyHex). The assignments were verified by exact mass and diagnostic MS^2 and MS^3 fragment ions. These ions did not appear to be host derived because they are very much less abundant than other host *O*-glycans, including HN, H2N, H3N, and sialylated derivatives (Fig. S7B), which were not detected in the *Toxoplasma* sample (Fig. S7A). The N2 and HdH ions were detected at higher levels than the $Man_6GlcNAc_2$

N-glycan that is inefficiently released by conditions of β-elimination. Monosaccharide modifications cannot be confidently assigned by this method. These findings indicate a limited repertoire of *O*-glycosylation at the proliferative tachyzoite stage and document the selectivity of the method for parasite glycans.

The N2 species is potentially representative of a mucin-type *O*-glycan and, based on studies using the lectins jacalin, *Vicia villosa* agglutinin, *Dolichos biflorus* agglutinin, and *Helix pomatia* agglutinin, expression of a GalNAc-GalNAc- structure was previously proposed (Fig. 6A). Analysis of the *O*-glycomes of pp-αGalNAcT2 (*gant2*Δ) and pp-αGalNAcT3 (*gant3*Δ) mutants showed that N2 depended on Gant2, with no effect observed on the expression of the HdH species (Fig. 5F). N2 was detected in *gant3*Δ extracts, though its apparent abundance relative to HdH ion was consistently reduced. Furthermore, N2 was also dependent on the expression of *galE* (see below), the epimerase that generates the sugar nucleotide UDP-GalNAc precursor for the pp-αGalNAcTs from UDP-GlcNAc (Fig. 5G). These results indicate that the reducing end N residue is GalNAc, but do not reveal the identity of the second N. The involvement of two genes predicted to encode Golgi HexNAcTs, TGGT1_266310 and TGGT1_207750 (Table 1E), have been excluded based on analysis of their *O*-glycans (Fig. S7 and data not shown). Thus tachyzoites express only one major mucin-type *O*-linked oligosaccharide, a disaccharide that may correspond to the previously predicted core 5 structure (GalNAcα1,3GalNAcα1-Thr). This prediction was based on lectin studies on bradyzoites (Tomita et al., 2017), which matched findings from metabolic labeling of a disaccharide on the GRA2 glycoprotein (Zinecker et al., 1998) that can be interpreted as GalNAc-GalNAc- based on more recently available information that the jacalin lectin can recognize non-reducing terminal GalNAc as well as Gal (Tachibana et al., 2006).

To investigate the potential expression of non-extended αGalNAc, referred to as Tn-antigen, we utilized anti-Tn mAbs developed from studies on human cancer (Steentoft et al., 2011). As shown in Fig. 5I, using two validated (see Methods) anti-Tn mAbs, we observed that both type 1 (RH) and type 2 (Pru) strains expressed at least 6 proteins (A-F) carrying the Tn-antigen in the M_r range of 20,000-60,000. Proteins in this M_r range have also been detected with various GalNAc-preferring lectins (Tomita et al., 2017; Wang et al., 2016). mAb recognition was dependent on both Gant2 and Gant3, as well as, as expected, GalE. Since only one pp-αGalNAcT is required to express Tn-antigen, these findings support the model of Tomita et al. (2017), consistent with precedent in animals, that Gant2 is a priming pp-αGalNAcT, and Gant3 is a follow-on pp-αGalNAcT that adds to the number of structures. Since N2 is more dependent on Gant2 than on GAnt3, N2 is likely to be assembled on the amplified number of Tn-antigen structures generated by Gant3.

The HdH ion is a candidate for the Glcβ1,3Fucα1- disaccharide that has been reported to be linked to Ser/Thr residues at consensus positions of TSR domains in animal glycoproteins (Vasudevan et al, 2015), and of *Plasmodium* CSP and TRAP proteins that also have the consensus peptide motif (Swearingen et al., 2016). MS^2 and MS^3 fragmentation indicated that the order of the sugars is Hex-dHex- (data not shown). Disruption of the *Plasmodium* homolog of *pofut2*, which catalyzes addition of αFuc to Ser/Thr, prevents assembly of the glycan (Lopaticki et al., 2017). Similarly, our *pofut2*Δ strain also blocked accumulation of HdH (Fig. 5H), demonstrating by analogy that the core deoxyHex residue is αFuc. Two recent preliminary reports confirm this result in *Toxoplasma* (Bandini et al., 2018; Khurana et al., 2018). The genome harbors a CAZy GT31 gene (TGGT1_239752 or *glt2*; Table 1) with homology to animal B3GlcT, which catalyzes addition of the βGlc cap at the 3-position of Fuc. Our *glt2*Δ strain also failed to express the disaccharide (Fig. 5H), confirming its role in its assembly, and indicating conservation of the Glcβ1,3Fucα1-disaccharide in *Toxoplasma* (Fig. 5B). Both assembly mutants exhibit a small plaque phenotype (Fig. 3), suggesting conservation of a role in quality control of folding of the several GRA proteins that bear the consensus motif in their TSR-repeats (Vasudevan et al., 2015).

In this first investigation of its *O*-glycome by mass spectrometry, *Toxoplasma* tachyzoites are found to exhibit a limited repertoire of *O*-linked oligosaccharides, just two disaccharide species that can be detected at the whole cell level. The precursor for the GalNAc-GalNAc- species, Tn-antigen (GalNAc-), also accumulates at detectable levels, and *O*-Fuc was also recently detected on TSR5 of MIC2 in a glycopeptide analysis (Bandini et al. 2018). Phylogenetic analyses of the genes known to assemble *Toxoplasma O*-glycans, including the Tn antigen, Glcβ1,3Fucα1-, and *C*-Man described below, are found in many other

apicomplexans but more related to their metazoan counterparts than genes found in other protists (unpublished). This is good evidence for the horizontal transfer of *O*- and *C*-glycosylation mechanisms from the genome of a host cell, within which apicomplexans are obligate residents.

GPI-anchor and GIPL- glycogenes. Numerous parasite proteins are linked to the cell surface via glycophosphatidylinositol (GPI) anchors. The glycan portion of *Toxoplasma* GPI anchors conforms to the conventional Man_3GlcNH_2inositol core formed by most eukaryotes but are reportedly modified at the core Man by a 3-linked βGalNAc residue (Azzouz et al., 2006). Some GPI anchor precursors also remain free as GIPLs, and are distinguished by an additional αGlc 4-linked to the βGalNAc (Fig. 6A). In addition, there is strain variation with regard to the presence of the phosphoethanolamine and αGlc termini in GIPLs (Niehus et al., 2014). Besides serving a structural role, GPI anchors and possibly GIPLs are recognized by host cell TLR2 and probably galectins as co-receptors (Debierre-Grockiego et al., 2010). Studies targeting either the GPI anchor or the galectin implicate their participation in the infection of host cells. The novel Glcα1,4GalNAcβ1- branch has been the subject of vaccine candidate and antibody-based diagnostics (Anish et al., 2014).

We developed a new MS-based method to analyze GPI anchor glycans released from an enriched pool of GPI-anchored proteins, by saponification and HF to break the phosphodiester linkages at both ends of the glycan. Unexpectedly, our analysis of strain RH tachyzoites yielded only the linear Man_3GnIn core lacking a side chain (Fig. 6B).

The high degree of conservation of the enzymes that assemble the canonical Man_3GnIn backbone enabled confident prediction of the enzyme coding genes (Table 1B). The phosphatidylinositol synthase (*pis*) that assembles the initial precursor from CDP-diacylglycerol and myoinositol has been described (Seron et al., 2000). *pis*, the *pigA* GlcNAcT, and the *pigM*, *pigV*, and *pigB* ManTs have very low fitness values, -3.1 to -5.1 (Fig. 1), indicating essentiality of the complete backbone. Our attempts to disrupt the final αManT (*pigB*) of the pathway were unsuccessful (Fig. 6A), consistent with the fitness data. With respect to the GTs that mediate the formation of the disaccharide side chain, the *Toxoplasma* genome encodes two βHexNAcT-like GTs that are predicted to be Golgi-associated type 2 membrane proteins, TGGT1_266310, TGGT1_207750, and one predicted αHexT that is a type 2 membrane protein, TGGT1_266320. These candidates, which have modest fitness values, were each successfully disrupted in our RH strain, but their effects on the assembly of the GPI-anchor glycan branch could not be evaluated owing to its absence in our studies (Fig. 6B).

In this first examination of *Toxoplasma* GPI-anchor glycans by mass spectrometry, we confirmed the core backbone, but did not detect a previously described unique glycan feature. Nevertheless, we have enumerated a plausible genetic basis for the assembly of the glycan part of the anchor and GIPLs with their branch chains, and developed tools for their analysis in future studies in other strains and under other conditions.

Other glycogenes associated with the secretory pathway. In addition to *N*- and *O*-glycosylation of secretory proteins, and their C-terminal glycolipidation, *Toxoplasma* also modifies a C-atom of the side chain of certain Trp residues with αMan (Fig. 5C). Originally found in association with a multi-Trp sequence motif in animal TSRs, this monosaccharide modification and the motif are conserved on CSP and TRAP proteins in *Plasmodium*, a related apicomplexan (Vasudevan et al., 2015). Both *Plasmodium* and *Toxoplasma* genomes encode a high scoring homolog of Dpy19 (Table 1D) that mediates formation of this linkage from Dol-P-Man in worms, and the *Toxoplasma* enzyme was recently shown to harbor the same activity (Hoppe et al., 2018). Using our double CRISPR plasmid we achieved a single insertion near the C-terminus of the predicted protein (Fig. 5C), but it is unclear whether this affects its C-αManT activity. Its -2.4 fitness score is consistent with importance for growth, potentially related to quality control of secretion of the proteins carrying TSRs, possibly in concert with Glc-Fuc- modifications (Scherbackova et al., 2017). Since there is no method known to release C-Man, monitoring the outcome will depend on glycopeptide analyses beyond the scope of this project.

The oocyst wall, formed during the sexual cycle in the gut of the cat host, contains disulfide rich

proteins and a β3glucan polysaccharide (Fig. 5E), and the genome encodes a predicted β3glucan synthase (Bushkin et al., 2012; Samuelson et al., 2013). Its disruption had no effect on proliferation of tachyzoites (Fig. 3), as expected because the polymer is assembled at another stage of the life cycle. Deletion of β3glucan synthase also has no effect on bradyzoite and tissue cyst formation *in vitro* (Bushkin et al., 2012). A previously described hint for the presence of chitin in the bradyzoite cyst wall is not supported by evidence for a chitin synthase-like sequence in the genome.

In addition to the free GIPLs, there is evidence for a class of glycosphingolipid-like molecules that can be metabolically labeled with tritiated Ser and Gal (Azzouz et al., 2002). They are so classified based on the resistance of the radiolabeled material to saponification and their sensitivity to known inhibitors of sphingolipid synthesis. The tritiated Gal was evidently incorporated as GalNAc, which might be mediated by one of the two predicted βHexNAcTs found in the genome (Table 1E).

Our GT predictions yield four unidentified genes (Table 1E, Fig. 5D), and our inventory of known glycans yields four linkages in need of GTs (N2, GPI-anchor, and glycosphingolipid). Future studies are needed to make the potential connections, and this summary does not exclude the possibility of lower abundance glycans or novel GT sequences yet to be discovered.

Cytoplasmic glycosylation glycogenes. Proteins of the nucleocytoplasmic compartment of eukaryotic cells are also potentially subject to glycosylation (West and Hart, 2017). Because these involve different GT genes and often require detection by different methods, this class of glycosylation events is often overlooked.

A subunit of the SCF class of E3 ubiquitin ligases, Skp1, has been shown to be modified by an *O*-linked pentasaccharide (Fig. 7A). A homolog of this glycan was originally discovered in the social amoeba *Dictyostelium* by metabolic labeling with [^{3}H]fucose, and subsequently defined by mass spectrometry and NMR (Sheikh et al., 2017). The glycan is linked in alkali-resistant fashion to a hydroxyproline, which therefore must be analyzed as tryptic glycopeptides, and may be restricted to only a single protein and so is not abundant. The *Toxoplasma* genome contains homologs of two of the GT genes required to assemble the *Dictyostelium* glycan (Rahman et al., 2016), which led to the discovery of a related pentasaccharide in *Toxoplasma* that is assembled by the two related and two unrelated GTs (Rahman et al., 2017). The Skp1 GT genes are not essential, but their disruption by conventional double crossover homologous recombination each generated small plaques, consistent with a role for the glycan in O_2-sensing originating from the O_2-dependence of the prolyl hydroxylase that forms the Hyp residue on which it is assembled (Xu et al., 2012; West and Blader, 2015). pDG plasmids and mutants for the first and final GTs in the pathway, *gnt1* and *gat1* (Table 1G), are available.

Recently, *Toxoplasma* nuclei were shown to be enriched in a novel Fuc modification that could be detected cytologically with the lectin *Aleuria aurantia* (AAL), which glycoproteomic analysis showed consisted of αFuc- modifications of Ser/Thr residues (Fig. 7B) often in clusters (Bandini et al., 2016). Our analysis predicted four unknown cytoplasmic (or nuclear) GTs, one of which was homologous to the *O*-βGlcNAcT (OGT or Secret agent) that modifies Ser/Thr residues of metazoan and higher plant nucleocytoplasmic proteins with *O*-βGlcNAc. Disruption of this gene, more closely related to its paralog *spindly* (*spy*) in *Arabidopsis*, abolishes AAL labeling of nuclei (Fig. 7C) and of extracts analyzed by Western blotting (Fig. 7D), indicating that Spy is an αFucT (or OFT) of nucleocytoplasmic proteins. This finding parallels that of a recent report showing *Arabidopsis* Spy is a *bona fide* OFT (Zentella et al., 2017). Despite the similarity of the Spy and OGT (Sec) sequences, a phylogenetic analysis of their evolution indicates that they are related by an ancient gene duplication possibly in prokaryotes, and diverged their donor sugar substrates at least by the time they appeared in eukaryotes (West et al., 2004). Spy appears to contribute functionality to the parasite according to the small plaque size phenotype (Fig. 3), though it did not contribute to fitness in the high throughput assay.

The cytoplasm is also the site of accumulation of amylopectin (Fig. 7E), a polymer of 4-linked αGlc with periodically spaced 6-linked α4glucan chains (Guérardel et al., 2005). Amylopectin assembles into semicrystalline floridean starch granules that are highly induced during bradyzoite differentiation. The *Toxoplasma* genome contains sequences for UDP-Glc dependent amylopectin assembly and disassembly

that relate more to fungal and animal glycogen metabolism than to the ADP-Glc dependent pathways in plants (Coppin et al., 2005). The branched polymer is assembled by a starch synthase and potentially two branching enzymes (Fig. 7E), and disassembled by a debranching enzyme, starch phosphorylase (Sugi et al., 2017) (Table 1I), and exoglycosidases (not listed). A gene with sequence similarity to glycogenin, an α4GlcT that primes the synthesis of the related polymer glycogen in yeast and animals, is the Skp1 glycosyltransferase Gat1, which might have been the evolutionary precursor of glycogenin (Mandalasi et al., 2018). The starch assembly genes do not evidently contribute much to tachyzoite fitness, consistent with the induction of starch accumulation in bradyzoites and oocysts. Thus this pathway was not targeted in this study.

The genome encodes a gene with an α,α-trehalose-6-P synthase (TPS)-like domain, and a domain with resemblance to trehalose-6-phosphatase (TPP), glycohydrolase GH77 (α4-glucanotransferase), and CBM20 starch binding domains (Table 1I). Genes for trehalose formation and turnover are widely distributed in eukaryotes (Avonce et al., 2006), but the absence of a candidate for trehalase, or evidence for trehalose in *Toxoplasma*, suggests a potential role for this gene in starch metabolism. The function of this gene, which evidently contributes to fitness (-2. 9), remains unassigned.

There are two reports of a glycolipid, characterized by a dihexosyl-moiety, possibly including a di-Gal epitope, on ceramide, diacylglycerol and acylalkylglycerol lipids (Botté et al., 2008; Maréchal et al., 2002). This is reminiscent of galactolipid type structures associated with plastids in plants, so they might potentially be associated with the apicoplast, the relict plastid that persists in *Toxoplasma*. The enzymatic basis for the addition of these sugars remains to be determined.

Three GT-like sequences predicted to reside in the cytoplasmic compartment (Table 1H, Fig. 7F) remain to be assigned. pDG plasmids (as well as mutants for TGGT1_288390 and TGGT1_217692) are available for their investigation, and disruption of the β3GlcNAcT-like exhibits a small plaque phenotype (Fig. 3). The detection of their presumptive glycan products might require specialized approaches to be detected, if they follow the example of the Skp1 glycan.

Sugar precursor glycogenes. Almost all glycosyltransferases rely on a sugar nucleotide or a Dol-P-sugar as the high energy sugar phosphate donor for the transfer of the sugar to the acceptor molecule. The formation of these donors from metabolic precursors tends to be well conserved. Fig. 8A-C summarize the metabolic pathways proposed to generate these compounds based on genomic analyses and experimental reports, and the genes are summarized in Table 1J. Based on information linking specific precursors to glycan types (Fig. 8D), genetic disruption of a precursor pathway offers an alternative approach to investigate the functional significance of a glycosylation pathway with multiple redundant GT genes.

Current information suggests that all sugars normally derive from Glc or Gal (Fig. 8B). Glc enters the cell via the plasma membrane associated Glc transporter TgGT1 (Zineker et al., 1998), and the α-anomer is converted to Glc-6-PO_4 (Glc6P) by hexokinase (Saito et al., 2002). Glc6P can be the source of all glycosylation precursors, and is also the C-source for central carbon metabolism via the classical Emden–Meyerhof glycolytic pathway and the pentose phosphate shunt. Alternatively, precursors for glucosylation and galactosylation may derive from the α-anomer of Gal via a salvage pathway (see below), as TgGT1 also transports Gal, Man and GlcNAc (Blume et al., 2009). Finally, glutamine can contribute to sugar precursor assembly via Glc generated by gluconeogenesis (Blume et al., 2015).

The lipid linked *N*-glycan donor contains GlcNAc, Man and Glc, whose precursors derive from each of the three main branches of precursor assembly (Fig. 8A-C). GlcNAc derives from UDP-GlcNAc which is assembled from Glc6P using the classical hexosamine biosynthetic pathway involving 5 enzymes (Fig. 8A). The highly negative fitness score (-2.5 to -5.2) of several of the enzymes implies essentiality of the pathway, consistent with the importance of *N*-glycosylation, so the significance of the less negative fitness score (-1.4) of the middle enzyme GNA is unclear. The ability to radiolabel glycans with tritiated $GlcNH_2$ (Azzouz et al., 2006; Azzouz et al, 2002), which presumably requires a transporter and a 6-hexokinase to enter the pathway, is however not explained by current knowledge. Owing to the likely necessity of these enzymes for viability (Cova et al., 2018), UDP-GlcNAc formation was not targeted in this study.

Man derives from GDP-Man, which is formed via three conventional enzymatic transformations from

Glc6P (Fig. 8C). The strongly negative fitness scores (-2.7 to -5.0) also suggest essentiality, consistent with the fitness scores for the *N*-glycan ManTs. The very negative fitness score for *pmi* (-5.0) indicates that Glc6P is the only C-source for this pathway. Thus GDP-Man biosynthesis was not targeted in this study. The ability of parasites to be radiolabeled with tritiated Man is consistent with the activities of the Glc transporter and hexokinase to utilize this sugar in addition to Glc.

The GlcNAc and Man residues of *N*-glycans are incorporated while the lipid-linked precursor is oriented toward the cytoplasm and derive from sugar nucleotide donors (Aebi et al., 2010). In contrast, the Glc residues are assembled after the precursor is flipped to the luminal face by a scramblase whose identity is unknown in any eukaryote. As is typical, the 3 αGlcT's, Alg6, Alg8 and Alg10, utilize Dol-P-Glc as donors. Dol-P-Glc is formed from UDP-Glc by Dol-P-Glc synthase (Alg5), whose -3.4 fitness score suggests importance of this modification, because there is no evidence that other GTs utilize this precursor. *alg5* was successfully disrupted, which resulted in the absence of the Glc modifications and small plaque sizes as expected (see *N*-glycans above).

UDP-Glc is in turn assembled from Glc6P via two traditional enzymatic transformations (Fig. 8B). Conversion to Glc1P is expected to be mediated by a phosphoglucomutase (Imada et al., 2010), but a double-KO of the two candidate genes does not affect viability (Saha et al., 2017). Therefore, this function might also be provided by phospho-N-acetylglucosamine mutase (*pagm*) or phosphomannomutase (*pmm*), as these enzymes are promiscuous in other parasites (Bandini et al., 2012). There is a single UDP-sugar pyrophosphorylase (*usp*) candidate whose score (-0.9) suggests that it also does not contribute much to fitness. Since this varies from the negative fitness score for Dol-P-Glc synthase, the UDP-GlcNAc pyrophosphorylase may be promiscuous as in *Arabidopsis* (Yang et al., 2010) allowing for redundancy. Nevertheless we have been unsuccessful in disrupting this gene using our double CRISPR approach.

Alternatively, Gal may be an alternate source of UDP-Glc via the so-called Isselbacher pathway found in plants and other eukaryotic parasites (Robo-Rojas, 2016). The Glc transporter TgGT1 also transports Gal (Blume et al., 2009) and *Toxoplasma* has predicted galactokinase (GalK; fitness -2.2) and UDP-Glc/UDP-Gal epimerase (GalE; fitness +1.7) enzymes. There is no good candidate for the Gal1P uridyltransferase of the classical Leloir pathway, suggesting that USP, the pyrophosphorylase that forms UDP-Glc from Glc1P, may promiscuously perform this role as observed in a trypanosome and plants (Damerow et al., 2010; Yang and Bar-Peled, 2010). The importance of this pathway was investigated by disrupting *galE*. This deletion resulted in a rather strong plaque defect (Fig. 3) despite a neutral fitness score, consistent with disruption of the Skp1 GalTs. However, GalE is also involved in the formation of UDP-GalNAc from UDP-GlcNAc as demonstrated by effects on mucin-type *O*-glycosylation (Fig. 5G). A selective effect on UDP-Gal formation could be investigated by disrupting GalK, for which a double CRISPR plasmid was generated. Thus *Toxoplasma* may utilize redundancy in its phosphoglycomutases and UDP-sugar pyrophosphorylases, as observed in other eukaryotes, to ensure access to precursors within varied host cells. Alternatively, other genes may yet to be identified, or a critical intermediate might derive directly from the host cell.

Assembly of the GPI anchor backbone, and *C*-Man of TSR domains, depend on Dol-P-Man. The Dol-P-Man synthase gene was not targeted owing to its very negative fitness score (-4.3) and the likely essentiality of the GPI anchor glycan (Fig. 6).

Fuc is known to be selectively applied to TSR domains in the secretory pathway, and to Skp1 and over 60 other proteins in the nucleus and cytoplasm. *Toxoplasma* lacks good candidate genes for the salvage pathway for GDP-Fuc assembly but does possess the two genes for the so-called conversion pathway. The second enzyme gene, the epimerase (or GDP-Fuc synthase), was modified by an insertion at near the 5'-end of its coding region yielded a modest small plaque phenotype (Fig. 3). However, the Glc-Fuc- structure was expressed normally, emphasizing the value of the double CRISPR method and the importance of biochemically confirming consequences of genetic modifications.

Altogether, the metabolic pathways summarized in Fig. 8 that generate activated sugars in *Toxoplasma* are thus far consistent with the gene repertoire and biochemical studies. However, it is still not clear how parasites can be labeled with free tritiated Man and $GlcNH_2$ and how tritiated Gal can label GalNAc (68). Nevertheless, the gene assignments resemble those of the agent for malaria, the related apicomplexan *Plasmodium* (Cova et al., 2015). Two major differences are that i) *Toxoplasma* also possesses enzymes for

the utilization of Gal, consistent with the ability to metabolically label the parasite with tritiated Gal (Azzouz et al., 2002), and good candidate genes for GalK and GalE, and ii) *Toxoplasma* has a Dol-P-Glc synthase to glucosylate its *N*-glycans within the rER lumen.

The sugar nucleotides described above are synthesized in the cytoplasm. Their access to GTs within the secretory pathway is afforded by specific transmembrane sugar nucleotide transporters (Fig. 8E), of which three are predicted in *Toxoplasma* (Table 1K). Nst1 has been examined biochemically and shown genetically to be important for the availability of UDP-GalNAc for *O*-glycosylation and for UDP-GlcNAc (Caffaro et al., 2013). Recent studies demonstrated a role for Nst2 in the transport of GDP-Fuc (Bandini et al., 2018). Nst3, which has a very negative fitness score (-4.9) has yet to be studied and may contribute to the transport of UDP-Gal and/or UDP-Glc, but biochemical studies are required to evaluate this speculation.

CONCLUSIONS

The *Toxoplasma* genome encodes at least 37 predicted GTs, and one phospho-GT, that modify a variety of proteins and lipids or generate polysaccharides. Twenty-six appear to be associated with the secretory pathway and 11 with cytoplasmic glycosylation. An additional 22 genes appear to be involved in generation of glycosylation precursors and their transport into the secretory pathway. This enzyme gene enumeration creates a new opportunity to explore gene-glycan-cell function relationships.

The range and abundance of *Toxoplasma* glycan classes, as profiled by mass spectrometry and including findings from radiolabeling, mAb and lectin studies, is more than initially appreciated but relatively limited compared to other protozoans such as trypanosomatid parasites and the amoebozoan *Dictyostelium* (Rodrigues et al., 2015; Sucgang et al., 2011; Feasley et al., 2015; Schiller et al., 2012). Furthermore, there is notable lack of known anionic modifications of *Toxoplasma* glycans that are prominent elsewhere in the form of anionic sugars or substitutions with PO_4 or SO_4. These differences may be a consequence of the homogenous environment experienced by this parasite, because when it is not residing within the parasitophorous vacuole of its host cell, it is enclosed within a protective cell wall.

N-glycosylation is remarkably simple, with incomplete Glc-trimming and little evidence for further remodeling of the novel $Man_6GlcNAc_2$- precursor, consistent with the absence of *N*-glycan mediated quality control mechanisms in the parasite rER (Bushkin et al., 2010; Samuelson and Robbins, 2015). *O*-glycans of the *Toxoplasma* secretory pathway are also simple, consisting mainly (if not only) of GalNAc-GalNAc- and GalNAc- in the mucin class, and Glc-Fuc- and Fuc- in the TSR class. *Toxoplasma* pp αGalNAcTs are more related to animal pp αGalNAcTs (CAZy GT27) than typical protist pp αGlcNAcTs (CAZy GT60) from which they evolved (West et al., 2004), suggesting that they derived from a host by horizontal gene transfer. This interpretation is consistent with the uniquely apicomplexan presence of the TSR glycosylation genes, and the evident absence of more diverse *O*-glycosylation, though the presence of four uncharacterized GT-like genes allows for the possibility of additional complexity at lower abundance or at other stages of the parasite life cycle.

Distinct *O*-glycosylation pathways occur in the parasite nucleus and cytoplasm. Over 60 different nucleocytoplasmic proteins have been shown to be modified by *O*-Fuc (Bandini et al., 2016) in a pattern that is reminiscent of *O*-GlcNAc modifications that occurs on thousands of animal nucleocytoplasmic proteins. This parallel is emphasized by the paralogous relationship, presented here, of their enzyme genes, which is certain to expose new mechanisms of protein regulation as has unfolded from studies on *O*-GlcNAc. The cytoplasm also harbors 5 additional GT activities that assemble a pentasaccharide on a single protein (West and Blader, 2015), Skp1, which contrasts with the role of the *O*-Fuc (Spy) GT to add a single sugar to many proteins. The cytoplasm rather than the apicoplast plastid is the locus of the enzymes that process amylopectin (starch), and additional unknown glycans are suggested by 3 other unknown cytoplasmic GT-like genes.

Altogether, glycosylation has many clear roles in the parasite that remain to be explored. The double-CRISPR approach is designed to facilitate these studies by its applicability to all wild-type strains of *Toxoplasma* (owing to NHEJ DNA repair), but the validated guide DNAs are readily adaptable to use in Ku80Δ strains lacking NHEJ, by using PCR to append 45-bp gene-specific homology arms to the ends of the DHFR amplicon. Use of Ku80Δ strains may limit promiscuous insertions of the DHFR expression

cassette in locations other than that targeted by the pair of guide RNAs. All guide sequences, CRISPR/Cas9 plasmids, and disrupted strains as summarized in Tables S1 and S4 are available to the community; individuals may contact C.M.W. It is anticipated that these strains and tools will be useful for others to investigate contributions of glycosylation to their favorite cellular or protein functions.

SUPPLEMENTAL INFORMATION

Tables S1-S5, and Figures S1-S7.

AUTHOR CONTRIBUTIONS

HTI and MIS generated the plasmids and most of the gene edited cell lines, and validated the target gene loci. BD and MM created and characterized additional gene edited lines. EGP prepared the samples for glycomic studies, performed independent genomic confirmation, interpreted the MS glycomic data, and correlated the data. MOS performed the MS glycomic studies. GB and JS characterized the spyΔ strain and assisted in glycogene identifications. LW and CMW designed the study, CMW predicted and organized the glycogene inventory, and wrote the manuscript, which was edited by EGP, MOS, GB, JS, and approved by all authors.

ACKNOWLEDGMENTS

We thank Bernard Henrissat (CNRS, Marseilles) and Kazi Rahman (Georgia) for their assistance in glycogene identifications. We are grateful to Drs. Tadakimi Tomita and Louis Weiss for the pp GalNAcT mutants, and to Drs. Ulla Mandel and Henrik Clausen for the gift of anti-Tn mAbs. This work was funded by a grant from the NIH Common Fund to Promote the Glycosciences (5R21 AI123161), and received additional support from NIH R01-037539 and NIH R01-084383.

REFERENCES

Aebi M, Bernasconi R, Clerc S, Molinari M (2010) N-glycan structures: recognition and processing in the ER. T*rends Biochem Sci* **35**:74-82.

Almeida IC, Ferguson MA, Schenkman S, Travassos LR (1994) Lytic anti-alpha-galactosyl antibodies from patients with chronic Chagas' disease recognize novel O-linked oligosaccharides on mucin-like glycosyl-phosphatidylinositol-anchored glycoproteins of *Trypanosoma cruzi*. *Biochem J* **304**:793-802.

Alvarez-Manilla G, Warren NL, Atwood J, Orlando R, Dalton S, Pierce M (2010) Glycoproteomic analysis of embryonic stem cells: identification of potential glycobiomarkers using lectin affinity chromatography of glycopeptides. *J Proteome Res* **9**:2062-2075.

Anish C, Schumann B, Pereira CL, Seeberger PH (2014) Chemical biology approaches to designing defined carbohydrate vaccines. *Chem Biol* **21**:38-50.

Anumula, K. R. and P. B. Taylor (1992) A comprehensive procedure for preparation of partially methylated alditol acetates from glycoprotein carbohydrates. *Anal Biochem* **203**:101-108.

Aoki K, Perlman M, Lim JM, Cantu R, Wells L, Tiemeyer M (2007) Dynamic developmental elaboration of N-linked glycan complexity in the *Drosophila melanogaster* embryo. *J Biol Chem* **282**:9127-9142.

Aoki K, Porterfield M, Lee SS, Dong B, Nguyen K, McGlamry KH, Tiemeyer M (2008) The diversity of O-linked glycans expressed during *Drosophila melanogaster* development reflects stage- and tissue-specific requirements for cell signaling. *J Biol Chem* **283**:30385-30400.

Aspinall TV, Joynson DH, Guy E, Hyde JE, Sims PF (2002) The molecular basis of sulfonamide resistance in *Toxoplasma gondii* and implications for the clinical management of toxoplasmosis. *J Infect Dis* **185**:1637-1643.

Avonce N, Mendoza-Vargas A, Morett E, Iturriaga G (2006) Insights on the evolution of trehalose biosynthesis. *BMC Evolutionary Biology* **6**:109.

Azzouz N, Rauscher B, Gerold P, Cesbron-Delauw MF, Dubremetz JF, Schwarz RT (2002) Evidence for de novo sphingolipid biosynthesis in *Toxoplasma gondii*. *Int J Parasitol* **32**:677-684.

Azzouz N, Shams-Eldin H, Niehus S, Debierre-Grockiego F, Bieker U, Schmidt J, Mercier C, Delauw MF, Dubremetz JF, Smith TK, Schwarz RT (2006) *Toxoplasma gondii* grown in human cells uses GalNAc-

containing glycosylphosphatidylinositol precursors to anchor surface antigens while the immunogenic Glc-GalNAc-containing precursors remain free at the parasite cell surface. *Int J Biochem Cell Biol* **38**:1914-1925.

Baatz H, Mirshahi A, Puchta J, Gumbel H, Hattenbach LO (2006) Reactivation of *Toxoplasma* retinochoroiditis under atovaquone therapy in an immunocompetent patient. *Ocul Immunol Inflamm* **14**:185-187.

Bandini G, Mariño K, Güther ML, Wernimont AK, Kuettel S, Qiu W, Afzal S, Kelner A, Hui R, Ferguson MA (2012) Phosphoglucomutase is absent in *Trypanosoma brucei* and redundantly substituted by phosphomanno- mutase and phospho-N-acetylglucosamine mutase. *Mol Microbiol* **85**:513–534.

Bandini G, Haserick JR, Motari E, Ouologuem DT, Lourido S, Roos DS, Costello CE, Robbins PW, Samuelson J (2016) *O*-fucosylated glycoproteins form assemblies in close proximity to the nuclear pore complexes of *Toxoplasma gondii*. *PNAS* **113**:11567-11752.

Bandini G, Leon DR, Hoppe CM, Zhang Y, Agop-Nersesian C, Shears MJ, Mahal LK, Routier FH, Costello CE, Samuelson J (2018) *O*-fucosylation of thrombospondin-like repeats is required for processing of MIC2 and for efficient host cell invasion by *Toxoplasma gondii* tachyzoites. bioRxiv doi: https://doi.org/10.1101/382515

Banerjee S, Vishwanath P, Cui J, Kelleher DJ, Gilmore R, Robbins PW, Samuelson J (2007) The evolution of N-glycan-dependent endoplasmic reticulum quality control factors for glycoprotein folding and degradation. *PNAS* **104**:11676-11681.

Blader IJ, Coleman BI, Chen CT, Gubbels MJ (2015) Lytic cycle of *Toxoplasma gondii*: 15 Years Later. Annu Rev Microbiol **69**:463-85.

Blume M, Rodriguez-Contreras D, Landfear S, Fleige T, Soldati-Favre D, Lucius R, Gupta N (2009) Host-derived glucose and its transporter in the obligate intracellular pathogen *Toxoplasma gondii* are dispensable by glutaminolysis. *PNAS* **106**:12998-13003.

Blume M, Nitzsche R, Sternberg U, Gerlic M, Masters SL, Gupta N, McConville MJ (2015) A *Toxoplasma gondii* gluconeogenic enzyme contributes to robust central carbon metabolism and is essential for replication and virulence. *Cell Host Microbe* **18**:210-220.

Botté C, Saïdani N, Mondragon R, Mondragón M, Isaac G, Mui E, McLeod R, Dubremetz JF, Vial H, Welti R, Cesbron-Delauw MF, Mercier C, Maréchal E (2008) Subcellular localization and dynamics of a digalactolipid-like epitope in *Toxoplasma gondii*. *J Lipid Res* **49**:746-762.

Bullard W, Lopes da Rosa-Spiegler J, Liu S, Wang Y, Sabatini R (2014) Identification of the glucosyltransferase that converts hydroxymethyluracil to base J in the trypanosomatid genome. *J Biol Chem* **289**:20273-20282.

Bushkin GG, Ratner DM, Cui J, Banerjee S, Duraisingh MT, Jennings CV, Dvorin JD, Gubbels MJ, Robertson SD, Steffen M, O'Keefe BR, Robbins PW, Samuelson J (2010) Suggestive evidence for Darwinian Selection against asparagine-linked glycans of *Plasmodium falciparum* and *Toxoplasma gondii*. *Eukaryot Cell* **9**:228-241.

Bushkin GG, Motari E, Magnelli P, Gubbels MJ, Dubey JP, Miska KB, Bullitt E, Costello CE, Robbins PW, Samuelson J (2012) β-1,3-glucan, which can be targeted by drugs, forms a trabecular scaffold in the oocyst walls of *Toxoplasma* and *Eimeria*. *mBio* **3**:e00258.

Caffaro CE, Koshy AA, Liu L, Zeiner GM, Hirschberg CB, Boothroyd JC (2013) A nucleotide sugar transporter involved in glycosylation of the *Toxoplasma* tissue cyst wall is required for efficient persistence of bradyzoites. *PLoS Pathog* **9**:e1003331.

Coppin A, Varré JS, Lienard L, Dauvillée D, Guérardel Y, Soyer-Gobillard MO, Buléon A, Ball S, Tomavo S (2005) Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. *J Mol Evol* **60**:257-267.

Cova M, Rodrigues JA, Smith TK, Izquierdo L (2015) Sugar activation and glycosylation in *Plasmodium*. *Malar J* **14**:427.

Cova M, López-Gutiérrez B, Artigas-Jerónimo S, González-Díaz A, Bandini G, Maere S, Carretero-Paulet L, Izquierdo L (2018) The Apicomplexa-specific glucosamine-6-phosphate N-acetyltransferase gene family encodes a key enzyme for glycoconjugate synthesis with potential as therapeutic target. *Sci Rep*

8:4005.

Damerow S, Lamerz AC, Haselhort T, Führing J, Zarnovican P, von Itztein M, Routier FH (2010) *Leishmania* UDP-sugar pyrophosphorylase: the missing link in galactose salvage? *J Biol Chem* **285**:878–887.

Dannemann B, McCutchan JA, Israelski D, Antoniskis D, Leport C, Luft B, Nussbaum J, Clumeck N, Morlat P, Chiu J, Vilde JL, Orellana M, Feigal D, Bartok A, Heseltine P, Leedom MD, Remington J (1992) Treatment of toxoplasmic encephalitis in patients with AIDS. A randomized trial comparing pyrimethamine plus clindamycin to pyrimethamine plus sulfadiazine. The California Collaborative Treatment Group. *Ann Intern Med* **116**:33-43.

Debierre-Grockiego F, Niehus S, Coddeville B, Elass E, Poirier F, Weingart R, Schmidt RR, Mazurier J, Guérardel Y, Schwarz RT (2010) Binding of *Toxoplasma gondii* glycosylphosphatidylinositols to galectin-3 is required for their recognition by macrophages. *J Biol Chem* **285**:32744-32750.

Fauquenoy S, Morelle W, Hovasse A, Bednarczyk A, Slomianny C, Schaeffer C, Van Dorsselaer A, Tomavo S (2008) Proteomics and glycomics analyses of N-glycosylated structures involved in *Toxoplasma gondii*—host cell interactions. *Mol Cell Proteomics* **7**:891–910.

Fauquenoy S, Hovasse A, Sloves PJ, Morelle W, Dilezitoko Alayi T, Slomianny C, Werkmeister E, Schaeffer C, Van Dorsselaer A, Tomavo S (2011) Unusual N-glycan structures required for trafficking *Toxoplasma gondii* GAP50 to the inner membrane complex regulate host cell entry through parasite motility. *Mol Cell Proteomics* **10**:M111.008953.

Feasley CL, van der Wel H, West CM (2015) Evolutionary diversity of social amoebae N-glycomes may support interspecific autonomy. *Glycoconjugate J* **32**:345–359.

Folch J, Lees M, Stanley GHS (1957) A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem* **226**:497-509.

Garénaux E, Shams-Eldin H, Chirat F, Bieker U, Schmidt J, Michalski JC, Cacan R, Guérardel Y, Schwarz RT (2008) The dual origin of *Toxoplasma gondii* N-glycans. *Biochemistry* **47**:12270-12276.

Guérardel Y, Leleu D, Coppin A, Liénard L, Slomianny C, Strecker G, Ball S, Tomavo S (2005) Amylopectin biogenesis and characterization in the protozoan parasite *Toxoplasma gondii*, the intracellular development of which is restricted in the HepG2 cell line. *Microbes Infect* **7**:41-48.

Haserick JR, Leon DR, Samuelson J, Costello CE (2017) Asparagine-linked glycans of *Cryptosporidium parvum* contain a single long arm, are barely processed in the endoplasmic reticulum (ER) or Golgi, and show a strong bias for sites with threonine. *Mol Cell Proteomics* **16**(4 suppl 1):S42-S53.

Heimburg-Molinaro J, Priest JW, Live D, Boons GJ, Song X, Cummings RD, Mead JR (2013) Microarray analysis of the human antibody response to synthetic *Cryptosporidium* glycopeptides. *Int J Parasit* **43**:901-907.

Hoppe CM, Albuquerque-Wendt A, Bandini G, Leon DR, Shcherbakova A, Buettner FFR, Izquierdo L, Costello CE, Bakker H, Routier FH (2018) Apicomplexan C-mannosyltransferases modify thrombospondin Type I-containing adhesins of the TRAP family. *Glycobiology* **28**:333-343.

Hykollari A, Dragosits M, Rendić D, Wilson IB, Paschinger K (2014) N-glycomic profiling of a glucosidase II mutant of *Dictyostelium discoideum* by "off-line" liquid chromatography and mass spectrometry. *Electrophoresis* **35**:2116-2129.

Imada M, Kawashima S, Kanehisa M, Takeuchi T, Asai T (2010) Characterization of alpha-phosphoglucomutase isozymes from *Toxoplasma gondii*. *Parasitol Int* **59**:206-210.

Katlama C, De Wit S, O'Doherty E, Van Glabeke M, Clumeck N (1996) Pyrimethamine-clindamycin vs. pyrimethamine-sulfadiazine as acute and long-term therapy for toxoplasmic encephalitis in patients with AIDS. *Clin Infect Dis* **22**:268-275.

Kim K, Weiss LM (2004) *Toxoplasma gondii*: the model apicomplexan. *Int J Parasitol* **34**:423-432.

Koeller CM, van der Wel H, Feasley CL, Abreu F, da Rocha JD, Montalvao F, Fampa P, Dos Reis FC, Atella GC, Souto-Padron T, West CM, Heise N (2014) Golgi UDP-GlcNAc:polypeptide O-alpha-N-Acetyl-D-glucosaminyltransferase 2 (TcOGNT2) regulates trypomastigote production and function in *Trypanosoma cruzi*. *Eukaryot Cell* **13**:1312-1327.

Kozulic B, Ries B, Mildner P (1979) N-acetylation of amino sugar methyl glycosides for gas-liquid

chromatographic analysis. *Anal Biochem* **94**:36-39.
Khurana S, Coffey MJ, John A, Ubold ADi, Huynh M-H, Stewart RJ, Carruthers VB, Tonkin CJ, Goddard-Borger ED, Scott N (2018) POFUT2-mediated *O*-glycosylation of MIC2 is dispensable for *Toxoplasma gondii* tachyzoites. *bioRxiv* doi: https://doi.org/10.1101/391623
Lambert H, Barragan A (2010) Modelling parasite dissemination: host cell subversion and immune evasion by *Toxoplasma gondii*. *Cell Microbiol* **12**:292-300.
Lobo-Rojas ÁE, González-Marcano EB, Valera-Vera EA, Acosta HR, Quiñones WA, Burchmore RJ, Concepción JL, Cáceres AJ (2016) *Trypanosoma cruzi* contains two galactokinases; molecular and biochemical characterization. *Parasitol Int* **65**(5 Pt A):472-482.
Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* **42**(Database issue):D490-495.
Long S, Wang Q, Sibley LD (2016) Analysis of Noncanonical Calcium-Dependent Protein Kinases in *Toxoplasma gondii* by Targeted Gene Deletion Using CRISPR/Cas9. Infect Immun **84**:1262-1273.
Lopaticki S, Yang ASP, John A, Scott NE, Lingford JP, O'Neill MT, Erickson SM, McKenzie NC, Jennison C, Whitehead LW, Douglas DN, Kneteman NM, Goddard-Borger ED, Boddey JA (2017) Protein O-fucosylation in *Plasmodium falciparum* ensures efficient infection of mosquito and vertebrate hosts. *Nat Commun* **8**:561
Luft BJ, Remington JS (1992) Toxoplasmic encephalitis in AIDS. *Clin Infect Dis* **15**:211-222.
Luk FC, Johnson TM, Beckers CJ (2008) N-linked glycosylation of proteins in the protozoan parasite *Toxoplasma gondii*. *Mol Biochem Parasitol* **157**:169-78.
Luo Q, Upadhya R, Zhang H, Madrid-Aliste C, Nieves E, Kim K, Angeletti RH, Weiss LM (2011) Analysis of the glycoproteome of *Toxoplasma gondii* using lectin affinity chromatography and tandem mass spectrometry. *Microbes Infect* 13:1199-210.
Macedo CS, Schwarz RT, Todeschini AR, Previato JO, Mendonça-Previato L (2010) Overlooked post-translational modifications of proteins in *Plasmodium falciparum*: *N*- and *O*-glycosylation -- a review. *Mem Inst Oswaldo Cruz* **105**:949-956.
Mandalasi M, Kim HW, Rahman K, Zhao P, Daniel N, Sheikh MO, van der Wel H, Thieker D, Ichikawa TH, Glushka J, Wells L, Wood ZA, West CM (2018)A homolog of glycogenin controls growth of *Toxoplasma gondii* via glycosylation of an E3 ubiquitin ligase. In preparation.
Maréchal E, Azzouz N, de Macedo CS, Block MA, Feagin JE, Schwarz RT, Joyard J (2002) Synthesis of chloroplast galactolipids in apicomplexan parasites. *Eukaryot Cell* **1**:653-656.
Masuishi Y, Kimura Y, Arakawa N, Hirano H (2016) Identification of glycosylphosphatidylinositol-anchored proteins and omega-sites using TiO_2-based affinity purification followed by hydrogen fluoride treatment. *J Proteomics* **139**:77-83.
Niehus S, Smith TK, Azzouz N, Campos MA, Dubremetz JF, Gazzinelli RT, Schwarz RT, Debierre-Grockiego F (2014) Virulent and avirulent strains of *Toxoplasma gondii* which differ in their glycosylphosphatidylinositol content induce similar biological functions in macrophages. *PLoS One* **9**:e85386.
Pappas G, Roussosa N, Falagasa ME (2009) Toxoplasmosis snapshots: Global status of *Toxoplasma gondii* seroprevalence and implications for pregnancy and congenital toxoplasmosis. *Int J Parasit* **39**:1385-1394.
Peng D, Tarleton R (2015) EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. *Microb Genom* **1**:e000033.
Pereira-Chioccola VL, Vidal JE, Su C (2009) *Toxoplasma gondii* infection and cerebral toxoplasmosis in HIV-infected patients. *Future Microbiol* **4**:1363-1379.
Rahman K, Zhao P, Mandalasi M, van der Wel H, Wells L, Blader IJ, West CM (2016) The E3-ubiquitin ligase adaptor protein Skp1 is glycosylated by an evolutionarily conserved pathway that regulates protist growth and development. J *Biol Chem* **291**:4268-4280.
Rahman K, Mandalasi M, Zhao P, Sheikh MO, Taujale R, Kim HW, van der Wel H, Matta K, Kannan N, Glushka JN, Wells L, West CM (2017) Characterization of a cytoplasmic glucosyltransferase that extends the core trisaccharide of the *Toxoplasma* Skp1 E3 ubiquitin ligase subunit. *J Biol Chem*

292:18644-18659.
Ranzinger R, Porterfield M, Weatherly B, Khan S, Arpinar S, Al-Jadda K, Myoung KT (2017) GRITS-toolbox. http://www.grits-toolbox.org/.
Rodrigues JA, Acosta-Serrano A, Aebi M, Ferguson MA, Routier FH, Schiller I, Soares S, Spencer D, Titz A, Wilson IB, Izquierdo L (2015) Parasite Glycobiology: A Bittersweet Symphony. PLoS Pathog **11**:e1005169.
Saha S, Coleman BI, Dubey R, Blader IJ, Gubbels MJ (2017) Two Phosphoglucomutase Paralogs Facilitate Ionophore-Triggered Secretion of the *Toxoplasma* Micronemes. *mSphere* **2**. pii: e00521-17.
Saito T, Maeda T, Nakazawa M, Takeuchi T, Nozaki T, Asai T (2002) Characterisation of hexokinase in *Toxoplasma gondii* tachyzoites. *Int J Parasitol* **32**:961-967.
Samuelson J, Banerjee S, Magnelli P, Cui J, Kelleher DJ, Gilmore R, Robbins PW (2005) The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. *PNAS* **102**:1548-1553.
Samuelson J, Bushkin GG, Chatterjee A, Robbins PW (2013) Strategies to discover the structural components of cyst and oocyst walls. *Eukaryot Cell* **12**:1578-1587.
Samuelson J, Robbins PW (2015) Effects of N-glycan precursor length diversity on quality control of protein folding and on protein glycosylation. *Semin Cell Dev Biol* **41**:121-128.
Sanz S, Bandini G, Ospina D, Bernabeu M, Mariño K, Fernández-Becerra C, Izquierdo L (2013) Biosynthesis of GDP-fucose and other sugar nucleotides in the blood stages of *Plasmodium falciparum*. *J Biol Chem* **288**:16506-16517.
Schiller B, Hykollari A, Yan S, Paschinger K, Wilson IB (2012) Complicated N-linked glycans in simple organisms. *Biol Chem* **393**:661-673.
Seron K, Dzierszinski F, Tomavo S (2000) Molecular cloning, functional complementation in *Saccharomyces cerevisiae* and enzymatic properties of phosphatidylinositol synthase from the protozoan parasite *Toxoplasma gondii*. *Eur J Biochem* **267**:6571–6579.
Shams-Eldin H, Blaschke T, Anhlan D, Niehus S, Müller J, Azzouz N, Schwarz RT (2005) High-level expression of the *Toxoplasma gondii* STT3 gene is required for suppression of the yeast STT3 gene mutation. *Mol Biochem Parasitol* **143**:6-11.
Shcherbakova A, Tiemann B, Buettner FF, Bakker H (2017) Distinct C-mannosylation of netrin receptor thrombospondin type 1 repeats by mammalian DPY19L1 and DPY19L3. *PNAS* **114**:2574-2579.
Sheikh MO, Thieker D, Chalmers G, Schafer CM, Ishihara M, Azadi P, Woods RJ, Glushka JN, Bendiak B, Prestegard JH, West CM (2017) O_2 sensing-associated glycosylation exposes the F-box-combining site of the *Dictyostelium* Skp1 subunit in E3 ubiquitin ligases. *J Biol Chem* **292**:18897-18915.
Shen B, Brown KM, Lee TD, Sibley LD (2014) Efficient gene disruption in diverse strains of *Toxoplasma gondii* using CRISPR/CAS9. *mBio* **5**:e01114-14.
Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, Nasamu AS, Thiru P, Saeij JPJ, Carruthers VB, Niles JC, Lourido S (2016) A genome-wide CRISPR screen in *Toxoplasma* identifies essential apicomplexan genes. Cell **167**:1423–1435.
Steentoft C, Vakhrushev SY, Vester-Christensen MB, Schjoldager KT, Kong Y, Bennett EP, Mandel U, Wandall H, Levery SB, Clausen H. (2011) Mining the O-glycoproteome using zinc-finger nuclease-glycoengineered SimpleCell lines. Nat Methods **8**:977-982.
Striepen B, Soldati D (2007) Genetic manipulation of *Toxoplasma gondii*. In: *Toxoplasma gondii*. The model apicomplexan: perspectives and methods. Eds: Weiss LM, Kim K. Elsevier. pp. 391-418.
Stwora-Wojczyk MM, Dzierszinski F, Roos DS, Spitalnik SL, Wojczyk BS (2004) Functional characterization of a novel *Toxoplasma gondii* glycosyltransferase: UDP-N-acetyl-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-T3. *Arch Biochem Biophys* **426**:231-240.
Sucgang R, Kuo A, Tian X, Salerno W, Parikh A, Feasley CL, Dalin E, Tu H, Huang E, Barry K, Lindquist E, Shapiro H, Bruce D, Schmutz J, Salamov A, Fey P, Gaudet P, Anjard C, Babu MM, Basu S, Bushmanova Y, van der Wel H, Katoh-Kurasawa M, Dinh C, Coutinho PM, Saito T, Elias M, Schaap P, Kay RR, Henrissat B, Eichinger L, Rivero F, Putnam NH, West CM, Loomis WF, Chisholm RL, Shaulsky G, Strassmann JE, Queller DC, Kuspa A, Grigoriev IV (2011) Comparative genomics of the

social amoebae *Dictyostelium discoideum* and *Dictyostelium purpureum*. *Genome Biol* **12**:R20.
Sugi T, Tu V, Ma Y, Tomita T, Weiss LM (2017) *Toxoplasma gondii* requires glycogen phosphorylase for balancing amylopectin storage and for efficient production of brain cysts. *mBio* **8**:e01289-17.
Swearingen KE, Lindner SE, Shi L, Shears MJ, Harupa A, Hopp CS, Vaughan AM, Springer TA, Moritz RL, Kappe SH, Sinnis P (2016) Interrogating the *Plasmodium* sporozoite surface: Identification of surface-exposed proteins and demonstration of glycosylation on CSP and TRAP by mass spectrometry-based proteomics. *PLoS Pathog* **12**:e1005606.
Tachibana K, Nakamura S, Wang H, Iwasaki H, Tachibana K, Maebara K, Cheng L, Hirabayashi J, Narimatsu H (2006) Elucidation of binding specificity of Jacalin toward O-glycosylated peptides: quantitative analysis by frontal affinity chromatography. *Glycobiology* **16**:46–53.
Tomita T, Bzik DJ, Ma YF, Fox BA, Markillie LM, Taylor RC, Kim K, Weiss LM (2013) The *Toxoplasma gondii* cyst wall protein CST1 is critical for cyst wall integrity and promotes bradyzoite persistence. *PLoS Pathog* **9**:e1003823.
Tomita T, Sugi T, Yakubu R, Tu V, Ma Y, Weiss LM (2017) Making home sweet and sturdy: *Toxoplasma gondii* ppGalNAc-Ts glycosylate in hierarchical order and confer cyst wall rigidity. *mBio* **8**:e02048-16.
Uboldi AD, McCoy JM, Blume M, Gerlic M, Ferguson DJ, Dagley LF, Beahan CT, Stapleton DI, Gooley PR, Bacic A, Masters SL, Webb AI, McConville MJ, Tonkin CJ (2015) Regulation of starch stores by a Ca(2+)-dependent protein kinase Is essential for viable cyst development in *Toxoplasma gondii*. *Cell Host Microbe* **18**:670-681.
Varki A, Cummings RD, Aebi M, Packer NH, Seeberger PH, Esko JD, Stanley P, Hart G, Darvill A, Kinoshita T, Prestegard JJ, Schnaar RL, Freeze HH, Marth JD, Bertozzi CR, Etzler ME, Frank M, Vliegenthart JF, Lütteke T, Perez S, Bolton E, Rudd P, Paulson J, Kanehisa M, Toukach P, Aoki-Kinoshita KF, Dell A, Narimatsu H, York W, Taniguchi N, Kornfeld S (2015) Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**:1323-1324.
Vasudevan D, Haltiwanger RS (2014) Novel roles for O-linked glycans in protein folding. *Glycoconj J* **31**:417-426.
Vasudevan D, Takeuchi H, Johar SS, Majerus E, Haltiwanger RS (2015) Peters plus syndrome mutations disrupt a noncanonical ER quality-control mechanism. *Curr Biol* **25**:286-295.
Wang K, Peng ED, Huang AS, Xia D, Vermont SJ, Lentini G, Lebrun M, Wastling JM, Bradley PJ (2016) Identification of novel O-Linked glycosylated *Toxoplasma* proteins by *Vicia villosa* lectin Chromatography. *PLoS One* **11**:e0150561.
Weiss LM, Kim K (2000) The development and biology of bradyzoites of *Toxoplasma gondii*. *Front Biosci* **5**:D391-405.
West CM, Van Der Wel H, Sassi S, Gaucher EA (2004) Cytoplasmic glycosylation of protein-hydroxyproline and its relationship to other glycosylation pathways. *Biochim Biophys Acta* **1673**:29-44.
West CM, van der Wel H. Coutinho PM, Henrissat B (2005) Glycosyltransferase genomics in *Dictyostelium discoideum*. In: *Dictyostelium* Genomics. Eds.: W.F. Loomis & A. Kuspa (Horizon Scientific Press, Norfolk, UK) pp. 235-264.
West CM, Blader IJ (2015) Oxygen sensing by protozoans: how they catch their breath. *Curr Opin Microbiol* **26**:41-47.
West CM, Hart GW (2017) In: Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, editors. Essentials of Glycobiology [Internet]. 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Ch. 18.
Xu Y, Brown KM, Wang ZA, van der Wel H, Teygong C, Zhang D, Blader IJ, West CM (2012) The Skp1 protein from *Toxoplasma* is modified by a cytoplasmic prolyl 4-hydroxylase associated with oxygen sensing in the social amoeba *Dictyostelium*. *J Biol Chem* **287**:25098-25110.
Yang T, Bar-Peled M (2010) Identification of novel UDP-sugar pyrophosphorylase with broad substrate specificity in *Trypanosoma cruzi*. *Biochem J* **429**:533–543.
Yang T, Echols M, Martin A, Bar-Peled M (2010) Identification and characterization of a strict and a promiscuous N-acetylglucosamine-1-P uridylyltransferase in *Arabidopsis*. *Biochem J* **430**:275-284.

Zentella R, Sui N, Barnhill B, Hsieh WP, Hu J, Shabanowitz J, Boyce M, Olszewski NE, Zhou P, Hunt DF, Sun TP (2017) The Arabidopsis O-fucosyltransferase SPINDLY activates nuclear growth repressor DELLA. *Nat Chem Biol* **13**:479-485.

Zinecker CF, Striepen B, Tomavo S, Dubremetz JF, Schwarz RT (1998) The dense granule antigen, GRA2 of *Toxoplasma gondii* is a glycoprotein containing O-linked oligosaccharides. *Mol Biochem Parasitol* **97**:241-246.

Table 1. *Toxoplasma* glycogenes and their characteristics. Summary of documented and predicted genes contributing directly and indirectly to the parasite glycome. Genes are organized according to glycan type to which they are expected to contribute. Characteristics and evidence regarding their functions from this project and others are summarized.

Glycan Class	ToxoDB ID	Gene name[1]	Description	CAZy[2]	Ref.	Fitness score[3]	gRNA plasmid[4]	Mutants available	Other mutants[5]	Glycome data	Plaque phenotype	Figure
A. *N*-linked												
lipid-linked precursor	TGGT1_244520	*alg7*	UDP-GlcNAc GlcNAc-P-T			-5.05						4A
	TGGT1_207070	*alg14*	UDP-GlcNAc β4GlcNAcT	GT1		-2.63						4A
	TGGT1_230590	*alg1*	GDP-Man β4ManT	GT33		-2.51	2	no				4A
	TGGT1_227790	*alg2*	GDP-Man α3/6ManT	GT4		-1.06	2	no				4A
	TGGT1_246982	*alg11*	GDP-Man α2ManT	GT4		-2.7	2	n.a.				4A
	TGGT1_314730	*alg6*	DPG α3GlcT1	GT57		-0.97	2	no				4A
	TGGT1_262030	*alg8*	DPG α3GlcT2	GT57		-1.94	2	yes		MS	small	4A
	TGGT1_263690	*alg10*	DPG α2GlcT1	GT59		-0.84	2	yes		MS	n.a.	4A
protein-associated	TGGT1_231430	*stt3*	oligosaccharyl transferase catalytic subunit	GT66	a	-3.45						4B
	TGGT1_242020	gluI	α-glucosidase-I	GH63		-2.55						4B
	TGGT1_253030	gluIIA	α-glucosidase-II, α-subunit	GH31		-0.92						4B
	TGGT1_253400	gluIIB	α-glucosidase-II, β-subunit	none		0.37						4B
B. GPI-anchor precursor												
	TGGT1_207710	*pis*	phosphatidylinositol synthase		b	-3.57						6
	TGGT1_241860	*pigA*	PI α6GlcNAcT catalytic subunit	GT4		-5.14	1,1	n.a.				6
	TGGT1_258110	*pigL*	GlcNAc-phosphatidylinositol deacetylase			-4.37						6
	TGGT1_310100	*pigM*	DPM α4ManT catalytic	GT50		-4.66						6
	TGGT1_270220	*pigV*	DPM α6ManT, smp3	GT76		-3.05						6
	TGGT1_321660	*pigB*	DPM α2ManT	GT22		-3.27	2	no				6
C. *O*-linked (secretory)												
mucin-type	TGGT1_259530	gant1	pp αGalNAcT1	GT27	c	1.15	1	n.a.	d			5A
	TGGT1_258770	gant2	pp αGalNAcT2	GT27	c	-0.07	2	no	d	lectin		5A
	TGGT1_318730	gant3	pp αGalNAcT3	GT27	c	-0.03	2	no	d	lectin		5A
	TGGT1_256080	gant4	pp αGalNAcT4	GT27	c	1.37	2	n.a.				5A
	TGGT1_278518	gant5	pp αGalNAcT5	GT27	c	-1.14	1,1	n.a.				5A
thrombospondin-type	TGGT1_273550	pofut2	POFUT2; GDP-Fuc:Ser/Thr O-FucT	GT65	e,o	-0.34	2	yes	e	MS	small	5B
	TGGT1_239752	glt2	UDP-Glc:Fuc β3GlcT	GT31		-1.36	2	yes		MS	small	5B
D. *C*-linked (secretory)												
	TGGT1_280400	dpy19	C-αManT	GT98		-2.37	2	no				5C
E. unknown GTs (secretory)												
	TGGT1_266310		βHexNAcT, peripheral	GT13		-1.89	2	yes		MS	small	5D
	TGGT1_266320		αHexT, peripheral (retaining)	GT32		-0.99	2	yes		MS	normal	5D
	TGGT1_207750		βHexNAcT, peripheral	GT17		-0.47	2	yes		MS	normal	5D
	TGGT1_238040A		LPS- or pp βGlc/XylT-like (inverting)	GT90		-0.53	2	no				5D
F. Polysaccharide (secretory)												
	TGGT1_278110		β3glucan synthase component	GT48	f	0.86	2	yes	f		normal	5E
G. *O*-linked (cytoplasmic)												
	TGGT1_315885	*gnt1*	Skp1 pp-Hyp αGlcNAcT	GT60	g	-0.58	2	yes	g	MS	small	7A
	TGGT1_260650	*pgtA*	Skp1 β3GalT/α2FucT	GT2/GT74	g	-4.15			g	MS	small	7A
	TGGT1_205060	*glt1*	Skp1 α3GlcT	GT32	h	0.96			h	MS	small	7A
	TGGT1_310400	*gat1*	Skp1 α3GalT	GT8	i	2.45	2	yes	i	MS	small	7A
	TGGT1_273500	*spy*	pp-Ser/Thr αFucT (inverting)	GT41		-0.11	2	yes		lectin	small	7B
H. unknown GTs (cytoplasmic)												
	TGGT1_288390		β3GlcNAcT-like	GT2		-1.93	2	yes			small	7F
	TGGT1_462962		β4GalT-like	GT7C		-0.95	1	n.a.				7F
	TGGT1_217692		retaining GT	GT47		0.45	2	yes			normal	7F
I. Cytoplasmic polysaccharide												
	TGGT1_222800	ssy	UDP-Glc: starch synthase	GT5/GT1		-0.7						7E
	TGGT1_209960	sbe1	α4glucan branching enzyme-like	GT1/GT4		-0.75						7E
	TGGT1_316520	sbe2	α4glucan branching enzyme-like			0.2						7E
	TGGT1_271210	sde1	debranching enz.; α4-glucanotransferase-like			0.27						
	TGGT1_310670	aph1	α4glucan phosphorylase	GT35	j	1.01			j			
	TGGT1_226910	amy1	amylo-α6glucosidase	GH133		-0.24						
	TGGT1_297720		α,α-trehalose-P synthase/phosphatase-like	GT20/GH77		-2.89						
J. Donor formation												
UDP-Glc	TGGT1_265450	hk1	hexokinase		k	-1.37						8B
	TGGT1_318580	*pgm2*	phosphoglycomutase 2		l	0.25			l			8B
	TGGT1_285980	*pgm1*	phosphoglycomutase 1 (A/B)		l	1.2/1.5			l			8B
	TGGT1_218200	*usp*	UDP-sugar pyrophosphorylase			-0.92	2	no				8B
Dol-P-Glc	TGGT1_216540	*alg5*	Dolichyl-phosphate βGlc synthase, catalytic	GT2		-3.36	2	yes		MS	small	8B
UDP-Gal	TGGT1_229780	*galK*	galactokinase			-2.18	2	n.a.				8B
	TGGT1_225880	*galE*	UDP-Glc 4-epimerase			1.73	2	yes		MS	small	8B
UDP-GlcNAc	TGGT1_283780	*pgi*	Glc-6-PO4 isomerase			-2.52						8A
	TGGT1_231350	*gfat*	glutamine fructose-6-PO4 isomerase			-4.57						8A
	TGGT1_243600	*gna*	GlcNH2-6-PO4 N-acetyltransferase		m	1.41						8A
	TGGT1_264650	*pagm*	phospho-N-acetylglucosamine mutase			-2.94						8A
	TGGT1_264780	*uap*	UDP-GlcNAc pyrophosphorylase			-5.22						8A
GDP-Man	TGGT1_260140	*pmi*	phosphomannose isomerase			-4.96						8C
	TGGT1_239710	*pmm*	phosphomannomutase			-2.66						8C
	TGGT1_257960	*mpg*	GDP-Man pyrophosphorylase			-4.6						8C
Dol-P-Man	TGGT1_277970	*dpm1*	Dolichyl-phosphate αMan synthase	GT2		-4.37						8C
GDP-Fuc	TGGT1_238940	*gmd*	GDP-Man dehydratase			-0.53						8C
	TGGT1_286220	*fx*	GDP-Man epimerase/reductase; GDP-Fuc synth.			0.53	2	yes		MS	small	8C
K. Xporter, sugar nucleotide												
	TGGT1_267380	*nst1*	UDP-GalNAc/GlcNAc transporter		n	0.97	2	n.a.	n	lectin		8E
	TGGT1_267730	*nst2*	GDP-Fuc transporter		e	1.23	2	n.a.	e			8E
	TGGT1_254580	*nst3*	pred. UDP-Gal/Glc transporter			-4.9						8E

FIGURE LEGENDS

Fig. 1. Fitness of glycogenes. Fitness data for glycogenes for invasion of and growth in human fibroblasts were extracted from Sidik et al. (2016) and graphed according to rank order of fitness value for separate categories of glycosylation as indicated at the bottom. Fitness scores are based on a log2 scale and values below -2.5 (red dashed line) have increasing likelihood of being essential. Filled data points represent genes for which mutants are available: green- from this study, blue- from other studies. Red fill refers to unsuccessful attempts from this study, and unfilled circles indicate no mutants attempted (but dual guide plasmids, or pDGs, available, in black) or known (in gray). Genes names correspond to listings in Table 1.

Fig. 2. Plasmids and gene replacement. The previously described pU6-Universal plasmid **(A)**, designed for single CRISPR/Cas9 editing, contains a BsaI site for convenient introduction of a synthetic duplex guide DNA **(B)**. pU6 was modified by the addition of a XhoI site at the 5'-end of its guide DNA cassette to generate p2 **(C)**, or its 3'end to generate p3 **(D)**. Thus the guide2 cassette of p2 could be conveniently excised from and re-ligated into p3 to generate the dual guide plasmid pDG **(E)**. **(F)** Gene replacement strategy. Co-transient transfection of pDG, expressing the gRNAs and Cas9, with the DHFR amplicon, results in the excision of the intervening genomic DNA (a) and ligation of the DHFR amplicon by non-homologous end joining (b). The desired replacement clones are selected for in the presence of pyrimethamine (c).

Fig. 3. Phenotype analysis. A clone of each strain was analyzed using a plaque assay on confluent human foreskin fibroblast (HFF) monolayers, and plaque areas were measured 6 d after clonal infection. (**A)** Image comparing plaques from RH and *alg8*Δ strains. (**B)** Dot plots from pairwise comparisons of mutant and parental (RH) strains. Gene names are color-coded to indicate their glycosylation pathway association (red, *N*-glycans; green, other secretory pathway glycans; purple, cytoplasmic glycans; black, precursor assembly), and Fitness scores from Fig. 1 are indicated in parentheses. To facilitate comparisons between trials, the mean areas of RH plaques were normalized to one and the respective mutant plaque areas scaled accordingly (only one representative RH experiment is graphed for clarity). Mean values ± 2 standard deviations (95% confidence interval) are shown with red bars. Statistical significance is shown at the top. Data were pooled from 3 independent trials, which each independently conformed to the same trend.

Fig. 4. *N*-glycan glycogenes and nanoLC-MS analysis of parasite *N*-glycans. (**A**) The *N*-glycan precursor is assembled on dolichol-pyrophosphate, one sugar at a time from right to left. Genes that contribute to biosynthetic steps are depicted. The status of gene editing is color-coded according to the key. The complete $Glc_3Man_6GlcNAc_2$ glycan is transferred *en bloc* to protein Asn side chains in an NxS/T motif by *stt3* and variably trimmed by α-glucosidases-I and –II (Table 1A). Glycans are represented according to the Consortium for Functional Glycomics system of nomenclature (Varki et al., 2015). (**B, C**) *N*-glycans were prepared by treating tryptic digests of total parasite protein powders with PNGaseF. The released glycans were permethylated and subjected to nLC on a C18 column followed by MS(n) in a linear ion-trap MS. Extracted ion chromatograms (EIC) for all detected *N*-glycan-like, lithiated compositions (H1N2 through H8N2) species are shown before and after digestion with jack bean α-mannosidase (α1-2/3/6) or *Aspergillus satoi* (α1-2) mannosidase to probe the identity of the hexose (H) species. The two peaks observed for each species are α- and β-anomers. Minor related *N*-glycan-like species are shown in Fig. S6A. (**B**) RH (parental). (**C**) *alg5*Δ-mutant, which is unable to form Dol-P-Glc, the Glc donor for terminal α-glucosylation.

Fig. 5. *O*-glycan, *C*-mannose and polysaccharide glycogenes in the secretory pathway and their glycomic analysis. (A) Mucin-type *O*-glycans. Tn antigen (GalNAcα-Ser/Thr) and the N2 disaccharide, consisting of a HexNAc of uncertain identity (potentially GalNAc) linked to αGalNAc, inferred from MS glycomic and lectin labeling studies. Five pp GalNAcTs are predicted to catalyze formation of the Tn

antigen, and T2 and T3 are confirmed. T2 is required for accumulation of N2, but the enzyme responsible for the second sugar is unknown, except that genes in green in panel D are not required. (**B**) TSR-type Glc-Fuc– disaccharide and its documented GT genes (see Fig. 7C). (**C**) GT for *C*-mannosylation. (**D**) Four genes encoding predicted sugar nucleotide dependent GT that appear to be type 2 membrane proteins in the secretory pathway. The glycans to which they contribute are yet to be identified. (**E**) GT for assembly of the oocyst wall β3-glucan, whose priming mechanism is unknown. (**F-H**) *O*-glycans recovered reductive β-elimination, permethylated, and subjected to nLC on a C18 column followed by MS(n) in a linear ion-trap MS as in Fig. 4. EICs for N2 (drawn as GalNAc-GalNAc-) and HdH (drawn as Glc-Fuc-) species (lithiated adducts ± 70 ppm), and the DP4 standard, are shown. (**F**) Parental strains (Type 2 Pru) and two derivative mutant strains (pp-αGalNAcT2 and pp-αGalNAcT3). (**G**) The parental strain (Type 1 RH) and the derivative *galE*Δ strain, and an extract of host hTERT cells. (**H**) RH and two derivative mutants predicted to encode the sequential addition of αFuc (*POFUT2*Δ) and βGlc (*β3GlcT*Δ). (**I**) Western blot analysis of total RH strain tachyzoites for the Tn-antigen, probed with mAb 5F4. Similar results were obtained with mAb 1E3, except that mAb-reactive proteins were differentially emphasized.

Fig. 6. GPIL and GPI-anchor biosynthesis glycogenes. (**A**) Free GIPLs and GPI-anchors share a related glycan linked to a diacylglycerol on the right. Glycan assembly proceeds stepwise from right to left along the upper arm, followed by the lower arm (Azzouz et al., 2006), by the enzymes predicted. The GPI-anchor precursor receives only the lower arm β4GalNAc, whereas the GPIL also receives an α4Glc (dashed line). Phosphoethanolamine is then conjugated to the terminal α2Man, and the resulting NH_2-group of the GPI precursor is conjugated to the polypeptide C-terminus by a multisubunit transamidase. Potential fatty acyl modifications are not shown. (**B**) The glycan region of GPI-anchored proteins enriched from RH strain tachyzoites was released after saponification by HF, which cleaves phosphodiester bonds, subjected to re-N-acetylation and permethylation, and analyzed by MS(n). *m/z* values corresponding to glycans containing one or two of the sugars of the predicted lower arm are indicated in gray, but were not detected in this study.

Fig. 7. Cytoplasmic GT glycogenes. (**A**) Assembly of the hydroxyproline-linked pentasaccharide on the $E3^{SCF}$Ub-ligase subunit Skp1. (**B**) Formation of *O*-Fuc on Ser/Thr residues of nucleocytoplasmic proteins. (**C**) AAL (green) binding to 4 nuclei of parental RH strain parasites (white arrow, upper left), housed within the same parasitophorous vacuole within a host fibroblast, is lost in the six *spy*Δ cells shown (lower left). To the right, IMC3 (red) labeling localizes to the parasite inner membrane complex whose organization varies according to the cell cycle. (**D**) AAL labeling of Western blots of parental strain whole cell extracts is lost in *spy*Δ cells. Subjecting the blotted membrane to conditions of β-elimination removed the majority of labeling, consistent with its *O*-linkage. (**E**) Genes associated with amylopectin assembly. Genes associated with its disassembly are in Table 1I. (**F**) Other genes predicted to encode sugar nucleotide-dependent cytoplasmic (or nuclear) GTs.

Fig. 8. Precursor assembly pathways and transporters. Predicted and validated glycogenes associated with formation and transport of sugar nucleotides, and formation of dolichol-linked monosaccharides, that are utilized as sugar donors by glycosyltransferases. (**A**) UDP-GlcNAc and UDP-GalNAc are derived from Glc via Glc6P. (**B**) UDP-Glc, UDP-Gal, and Dol-P-Glc can derive from either Glc or Gal. (**C**) GDP-Man, Dol-P-Man and GDP-Fuc also ultimately derive from Glc. (**D**) Destination glycans for the precursors boxed in orange in panels A-C. (**E**) Transporters transfer sugar nucleotides from the cytoplasm where they are synthesized into the lumina of vesicles containing the glycosyltransferases. The scramblases that re-orient Dol-P-Glc and Dol-P-Man from the cytoplasmic to the luminal face of the rER are unknown. (**F**) Definition of symbols and color coding of the status of gene analyses.

Figure 1

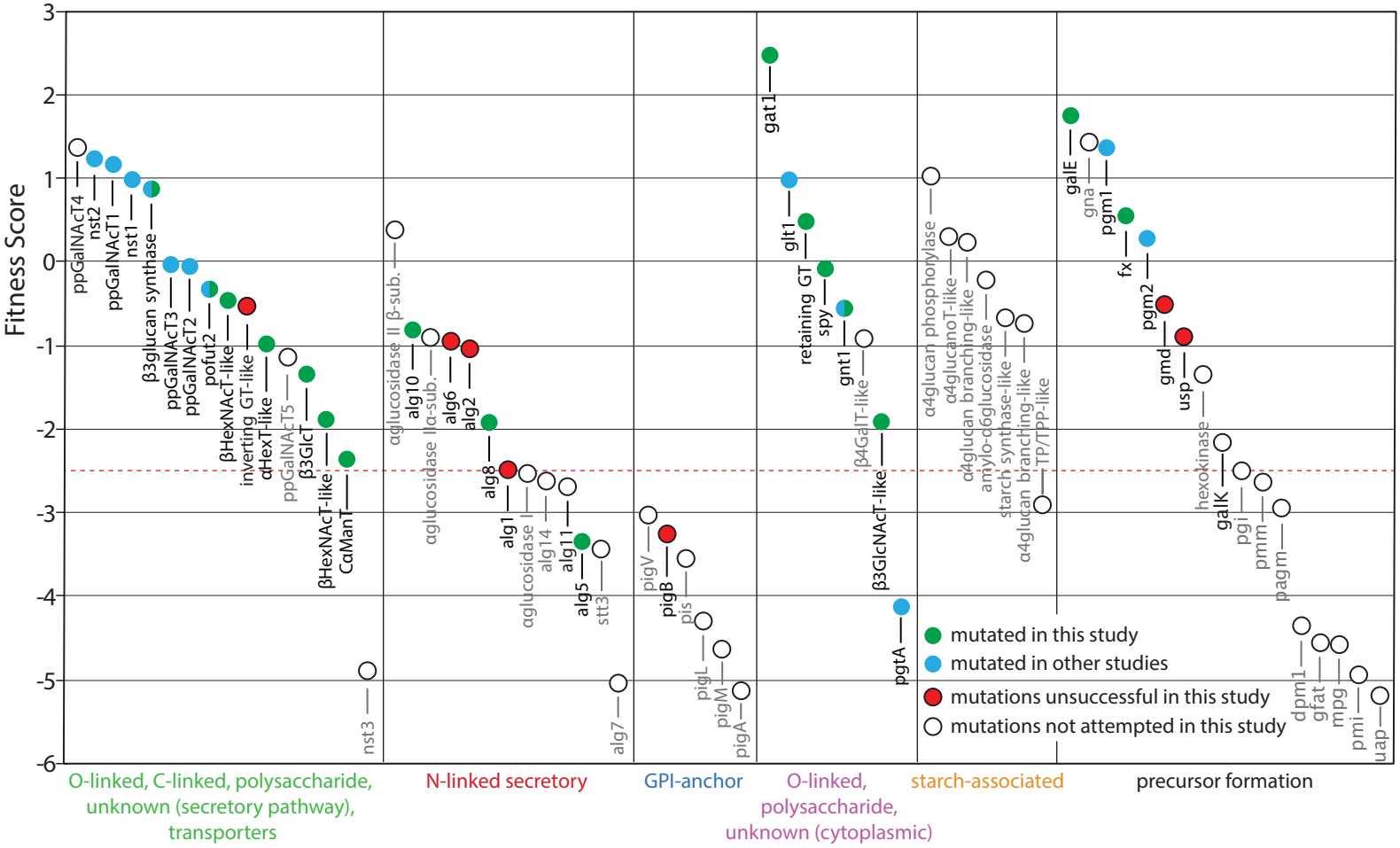

Fitness Score
3
2
1
0
-1
-2
-3
-4
-5
-6
ppGalNAcT4
nst2
ppGalNAcT1
nst1
β3glucan synthase
ppGalNAcT3
ppGalNAcT2
pofut2
βHexNAcT-like
inverting GT-like
αHexT-like
ppGalNAcT5
β3GlcT
βHexNAcT-like
CdManT
nst3
αglucosidase II β-sub.
alg10
αglucosidase IIα-sub.
alg6
alg2
alg8
alg1
αglucosidase I
alg14
alg11
alg5
stt3
alg7
pigV
pigB
pis
pigL
pigM
pigA
gat1
glt1
retaining GT
spy
gnt1
β4GalT-like
β3GlcNAcT-like
pgtA
α4glucan phosphorylase
α4glucanoT-like
α4glucan branching-like
amylo-α6glucosidase
starch synthase-like
α4glucan branching-like
TP/TPP-like
galE
gna
pgm1
fx
pgm2
gmd
usp
hexokinase
galK
pgi
pmm
pagm
dpm1
gfat
mpg
pmi
uap
mutated in this study
mutated in other studies
mutations unsuccessful in this study
mutations not attempted in this study
O-linked, C-linked, polysaccharide, unknown (secretory pathway), transporters
N-linked secretory
GPI-anchor
O-linked, polysaccharide, unknown (cytoplasmic)
starch-associated
precursor formation

Figure 2

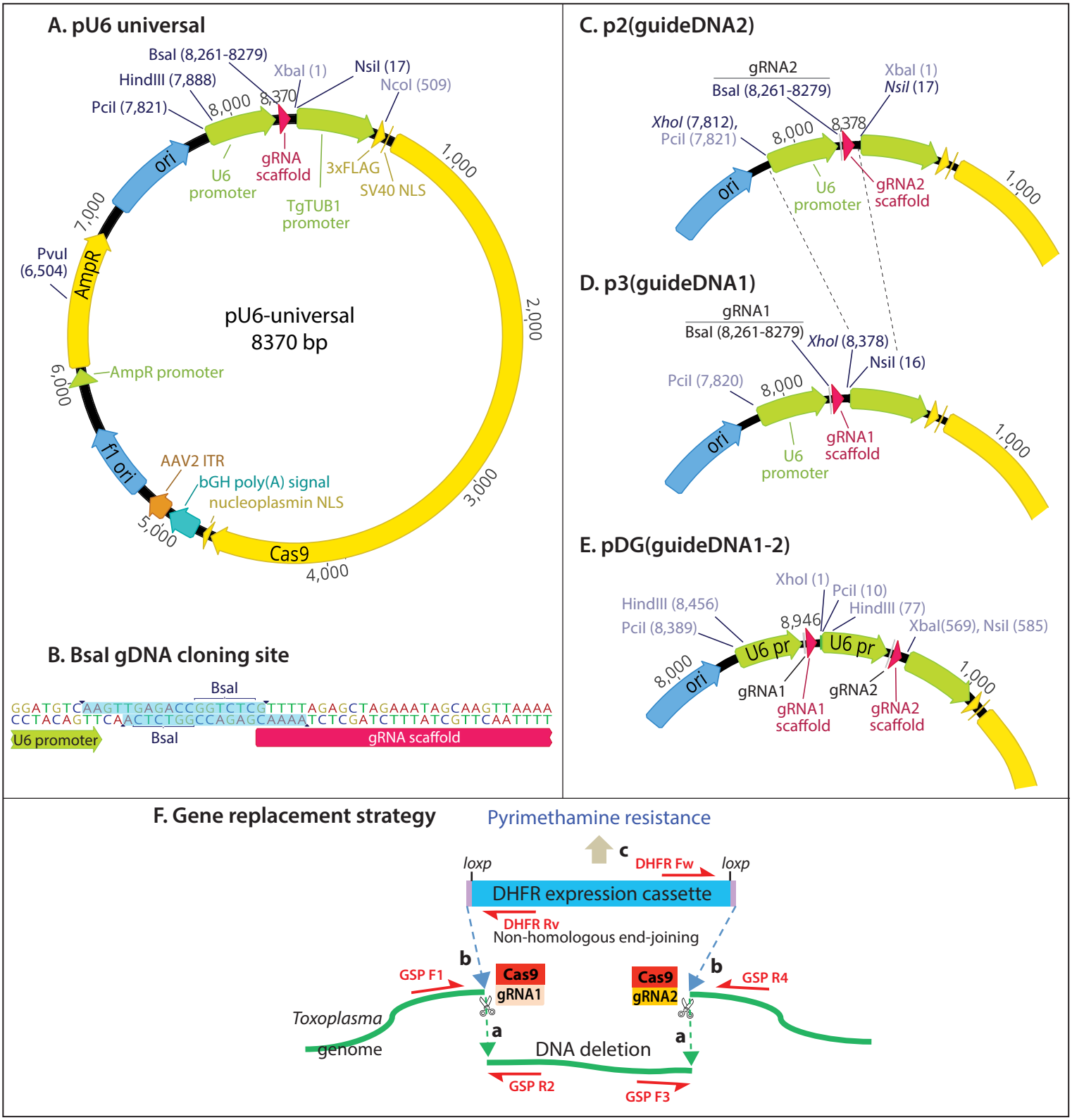

A. pU6 universal
BsaI (8,261-8279)
HindIII (7,888)
PciI (7,821)
XbaI (1)
NsiI (17)
NcoI (509)
8,000
8,370
ori
U6 promoter
gRNA scaffold
3xFLAG
SV40 NLS
TgTUB1 promoter
1,000
7,000
PvuI (6,504)
AmpR
pU6-universal
8370 bp
2,000
AmpR promoter
6,000
f1 ori
AAV2 ITR
bGH poly(A) signal
nucleoplasmin NLS
3,000
5,000
Cas9
4,000
B. BsaI gDNA cloning site
BsaI
GGATGTCAAGTTGAGACCGGTCTCGTTTTAGAGCTAGAAATAGCAAGTTAAAA
CCTACAGTTCAACTCTGGCCAGAGCAAAATCTCGATCTTTATCGTTCAATTTT
U6 promoter
BsaI
gRNA scaffold
C. p2(guideDNA2)
gRNA2
BsaI (8,261-8279)
XbaI (1)
NsiI (17)
XhoI (7,812),
PciI (7,821)
8,000
8,378
ori
U6 promoter
gRNA2 scaffold
1,000
D. p3(guideDNA1)
gRNA1
BsaI (8,261-8279)
XhoI (8,378)
NsiI (16)
PciI (7,820)
8,000
ori
U6 promoter
gRNA1 scaffold
1,000
E. pDG(guideDNA1-2)
XhoI (1)
PciI (10)
HindIII (8,456)
HindIII (77)
PciI (8,389)
8,946
XbaI(569), NsiI (585)
8,000
ori
U6 pr
U6 pr
gRNA1
gRNA2
gRNA1 scaffold
gRNA2 scaffold
1,000
F. Gene replacement strategy
Pyrimethamine resistance
c
loxp
DHFR Fw
loxp
DHFR expression cassette
DHFR Rv
Non-homologous end-joining
b
b
GSP F1
Cas9
gRNA1
Cas9
gRNA2
GSP R4
Toxoplasma
genome
a
a
DNA deletion
GSP R2
GSP F3

Figure 3

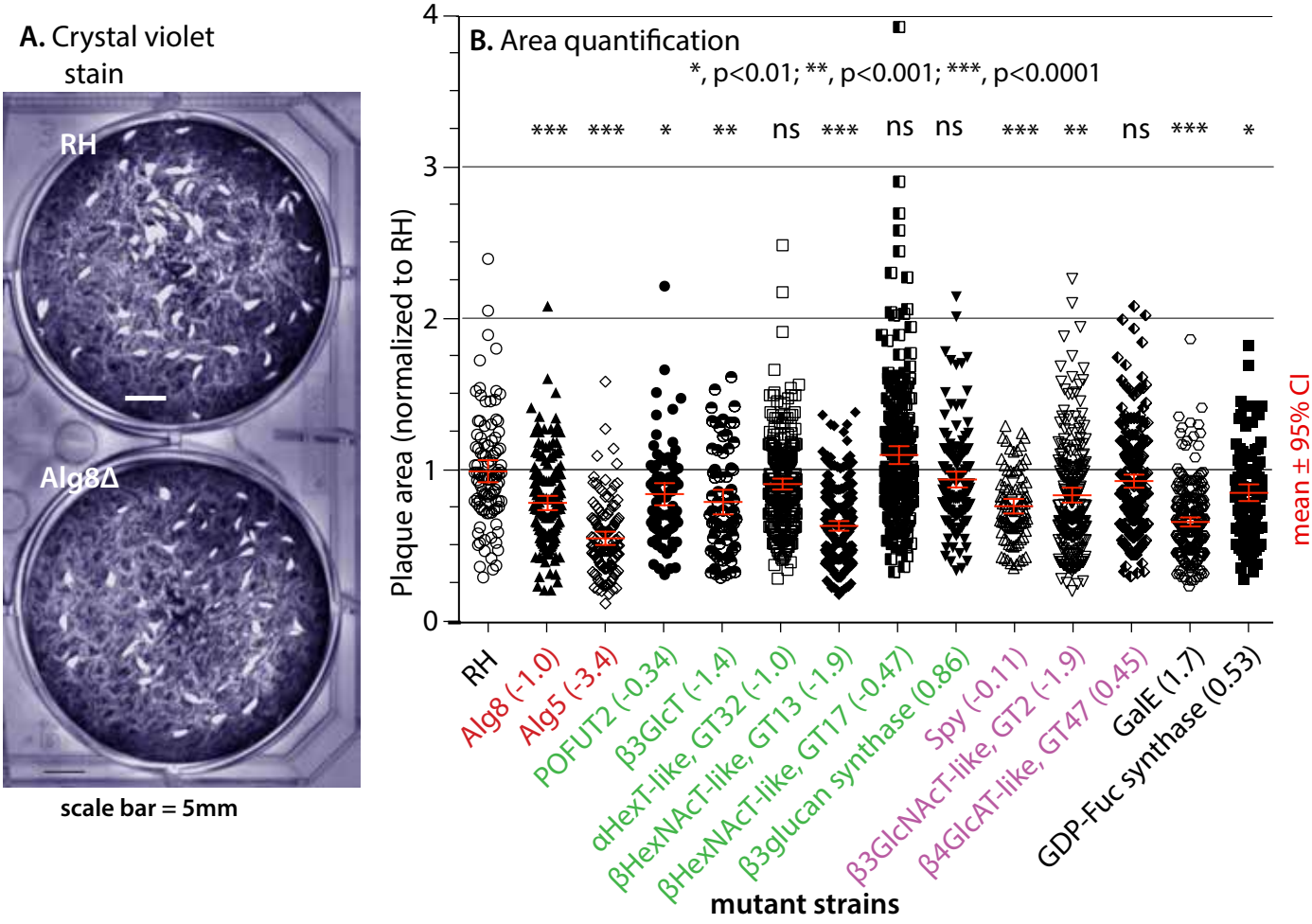

A. Crystal violet stain
RH
Alg8Δ
scale bar = 5mm
B. Area quantification
*, p<0.01; **, p<0.001; ***, p<0.0001
*** *** * ** ns *** ns ns *** ** ns *** *
Plaque area (normalized to RH)
0
1
2
3
4
mean ± 95% CI
RH
Alg8 (-1.0)
Alg5 (-3.4)
POFUT2 (-0.34)
β3GlcT (-1.4)
αHexT-like, GT32 (-1.0)
βHexNAcT-like, GT13 (-1.9)
βHexNAcT-like, GT17 (-0.47)
β3glucan synthase (0.86)
Spy (-0.11)
β3GlcNAcT-like, GT2 (-1.9)
β4GlcAT-like, GT47 (0.45)
GalE (1.7)
GDP-Fuc synthase (0.53)
mutant strains

Figure 4

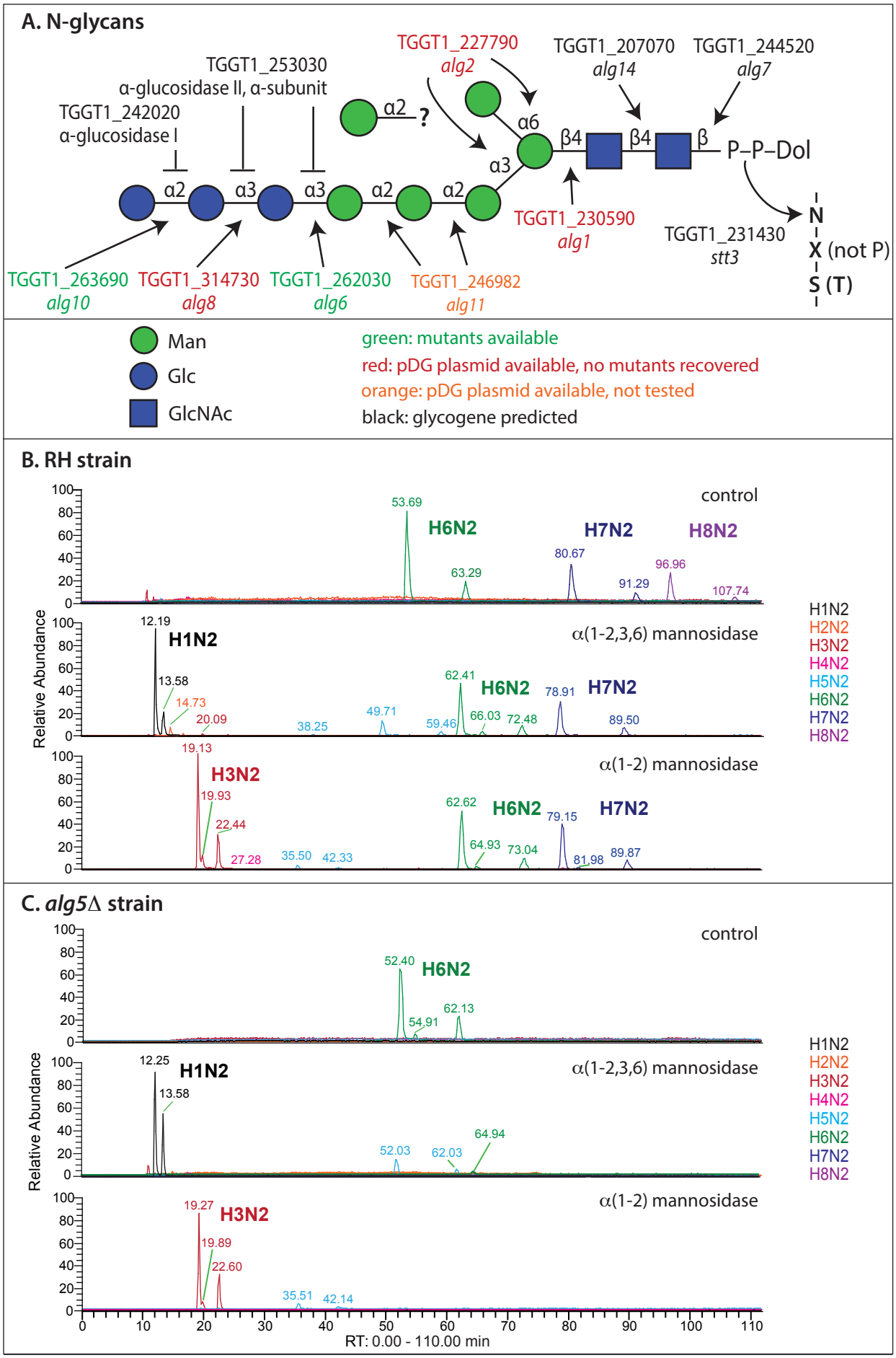

A. N-glycans
TGGT1_227790
alg2
TGGT1_207070
alg14
TGGT1_244520
alg7
TGGT1_253030
α-glucosidase II, α-subunit
TGGT1_242020
α-glucosidase I
α2
?
α6
β4
β4
β
P–P–Dol
α3
α2
α3
α3
α2
α2
TGGT1_230590
alg1
TGGT1_231430
stt3
N
X (not P)
S (T)
TGGT1_263690
alg10
TGGT1_314730
alg8
TGGT1_262030
alg6
TGGT1_246982
alg11
Man
Glc
GlcNAc
green: mutants available
red: pDG plasmid available, no mutants recovered
orange: pDG plasmid available, not tested
black: glycogene predicted
B. RH strain
control
53.69
H6N2
63.29
H7N2
80.67
H8N2
96.96
91.29
107.74
α(1-2,3,6) mannosidase
12.19
H1N2
13.58
14.73
20.09
38.25
49.71
59.46
62.41
H6N2
66.03
72.48
78.91
H7N2
89.50
α(1-2) mannosidase
19.13
H3N2
19.93
22.44
27.28
35.50
42.33
62.62
H6N2
64.93
73.04
79.15
H7N2
81.98
89.87
Relative Abundance
H1N2
H2N2
H3N2
H4N2
H5N2
H6N2
H7N2
H8N2
C. alg5Δ strain
control
52.40
H6N2
54.91
62.13
α(1-2,3,6) mannosidase
12.25
H1N2
13.58
52.03
62.03
64.94
α(1-2) mannosidase
19.27
H3N2
19.89
22.60
35.51
42.14
RT: 0.00 - 110.00 min

Figure 5

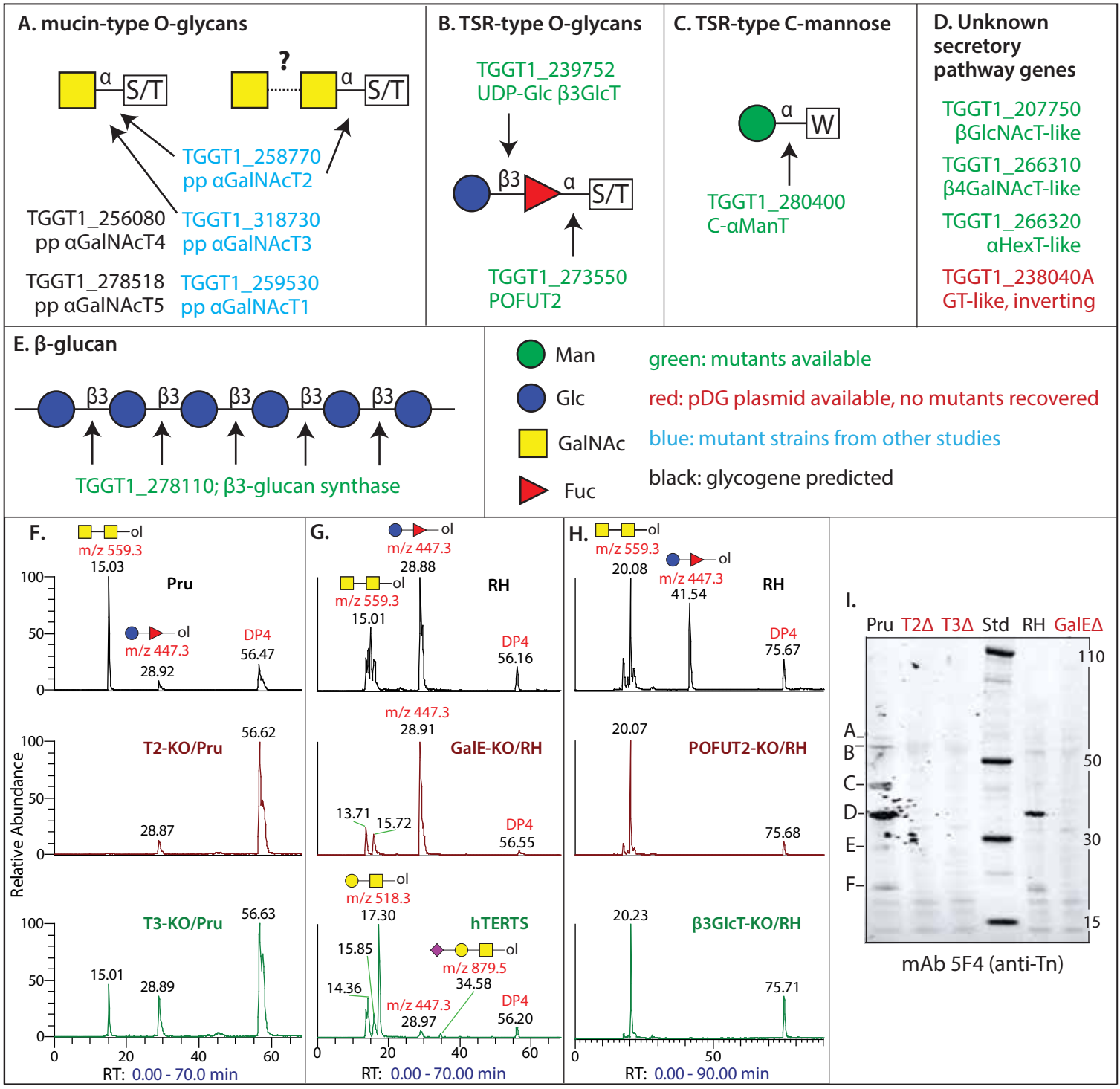

A. mucin-type O-glycans
α
S/T
?
TGGT1_258770
pp αGalNAcT2
TGGT1_256080
pp αGalNAcT4
TGGT1_318730
pp αGalNAcT3
TGGT1_278518
pp αGalNAcT5
TGGT1_259530
pp αGalNAcT1
B. TSR-type O-glycans
TGGT1_239752
UDP-Glc β3GlcT
β3
TGGT1_273550
POFUT2
C. TSR-type C-mannose
W
TGGT1_280400
C-αManT
D. Unknown secretory pathway genes
TGGT1_207750
βGlcNAcT-like
TGGT1_266310
β4GalNAcT-like
TGGT1_266320
αHexT-like
TGGT1_238040A
GT-like, inverting
E. β-glucan
TGGT1_278110; β3-glucan synthase
Man
Glc
GalNAc
Fuc
green: mutants available
red: pDG plasmid available, no mutants recovered
blue: mutant strains from other studies
black: glycogene predicted
F.
m/z 559.3
15.03
Pru
m/z 447.3
28.92
DP4
56.47
T2-KO/Pru
56.62
28.87
T3-KO/Pru
56.63
15.01
28.89
Relative Abundance
RT: 0.00 - 70.0 min
G.
m/z 447.3
28.88
RH
m/z 559.3
15.01
DP4
56.16
m/z 447.3
28.91
GalE-KO/RH
13.71
15.72
DP4
56.55
m/z 518.3
17.30
hTERTS
15.85
m/z 879.5
34.58
14.36
m/z 447.3
28.97
DP4
56.20
RT: 0.00 - 70.00 min
H.
m/z 559.3
20.08
m/z 447.3
41.54
RH
DP4
75.67
20.07
POFUT2-KO/RH
75.68
20.23
β3GlcT-KO/RH
75.71
RT: 0.00 - 90.00 min
I.
Pru T2Δ T3Δ Std RH GalEΔ
110
50
30
15
A
B
C
D
E
F
mAb 5F4 (anti-Tn)

Figure 6

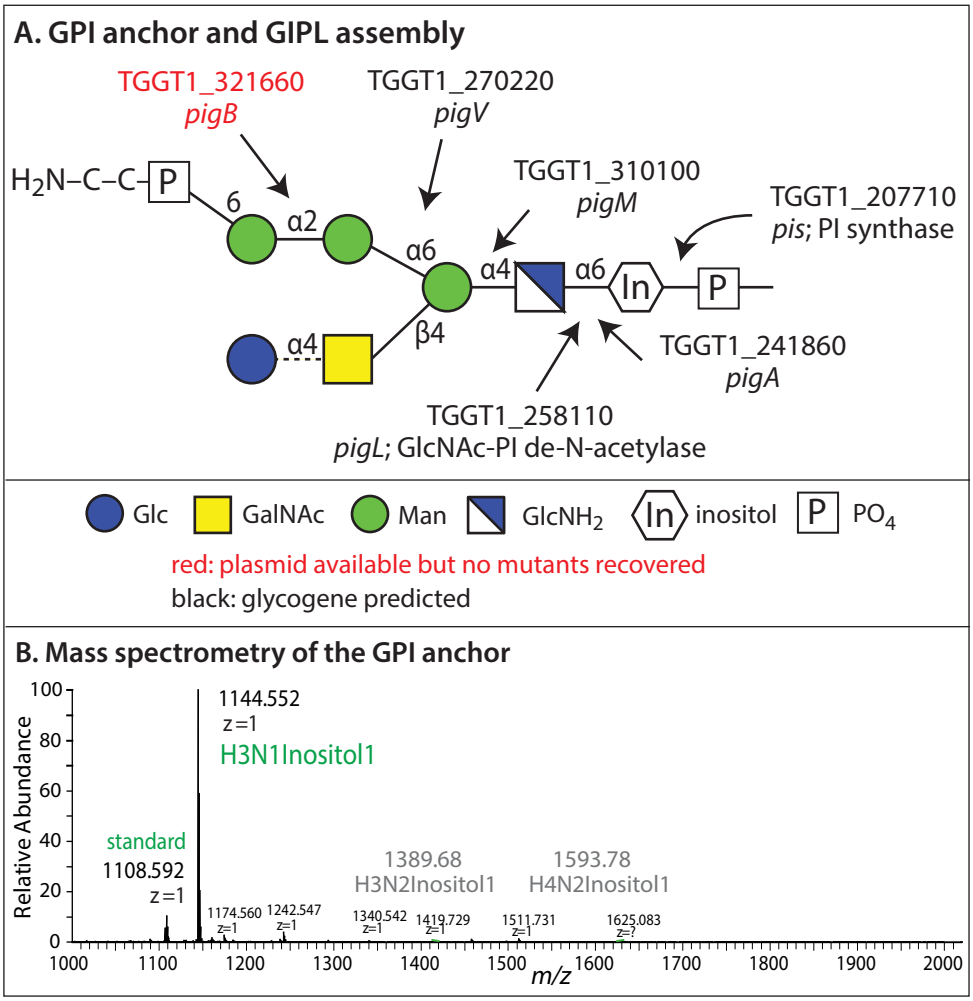

A. GPI anchor and GIPL assembly
TGGT1_321660
pigB
TGGT1_270220
pigV
TGGT1_310100
pigM
TGGT1_207710
pis; PI synthase
H2N–C–C–P
6
α2
α6
α4
α6
In
P
α4
β4
TGGT1_241860
pigA
TGGT1_258110
pigL; GlcNAc-PI de-N-acetylase
Glc
GalNAc
Man
GlcNH2
In inositol
P PO4
red: plasmid available but no mutants recovered
black: glycogene predicted
B. Mass spectrometry of the GPI anchor
Relative Abundance
1144.552
z=1
H3N1Inositol1
standard
1108.592
z=1
1389.68
H3N2Inositol1
1593.78
H4N2Inositol1
1174.560 z=1
1242.547 z=1
1340.542 z=1
1419.729 z=1
1511.731 z=1
1625.083 z=?
m/z

Figure 7

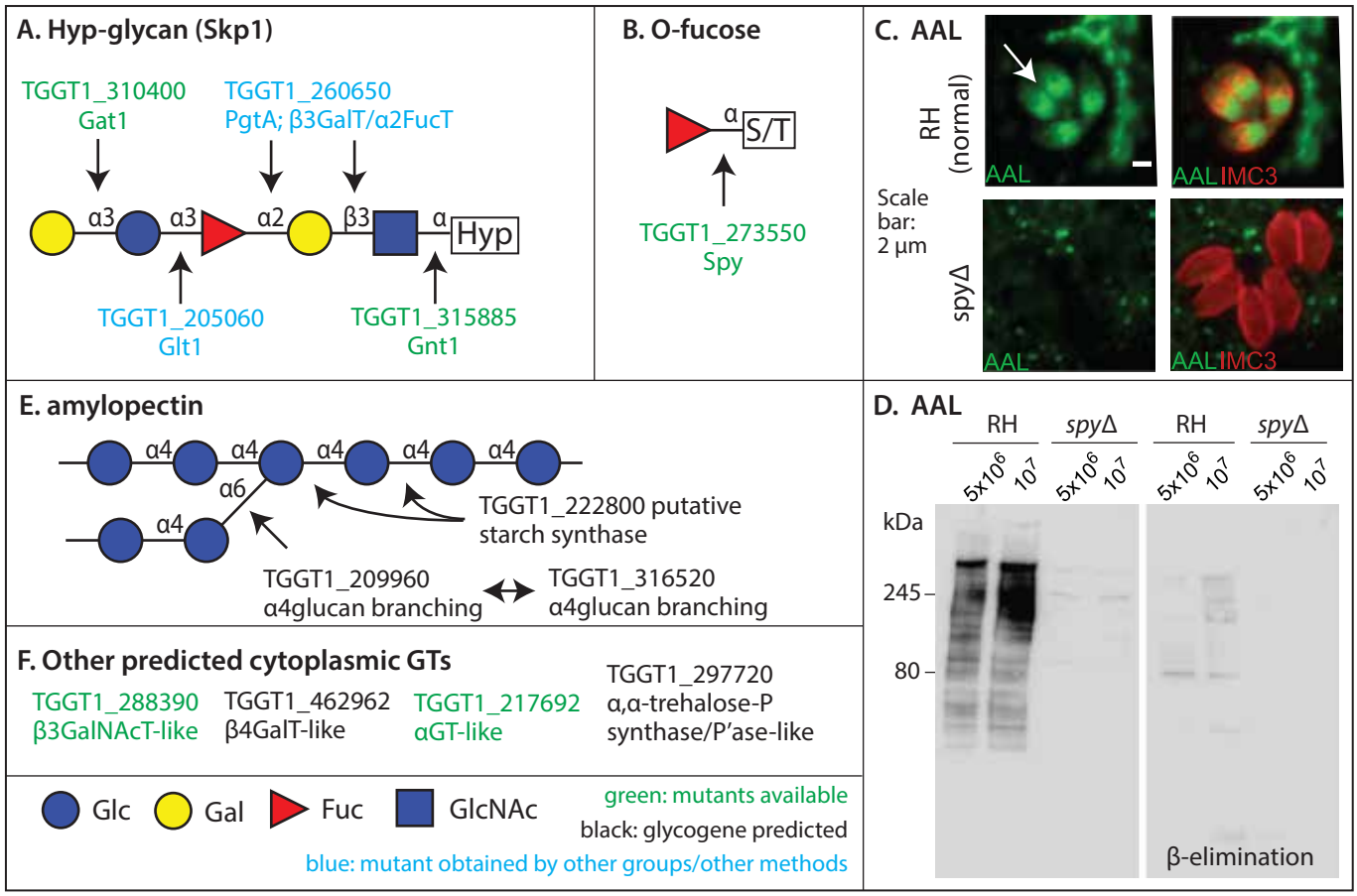
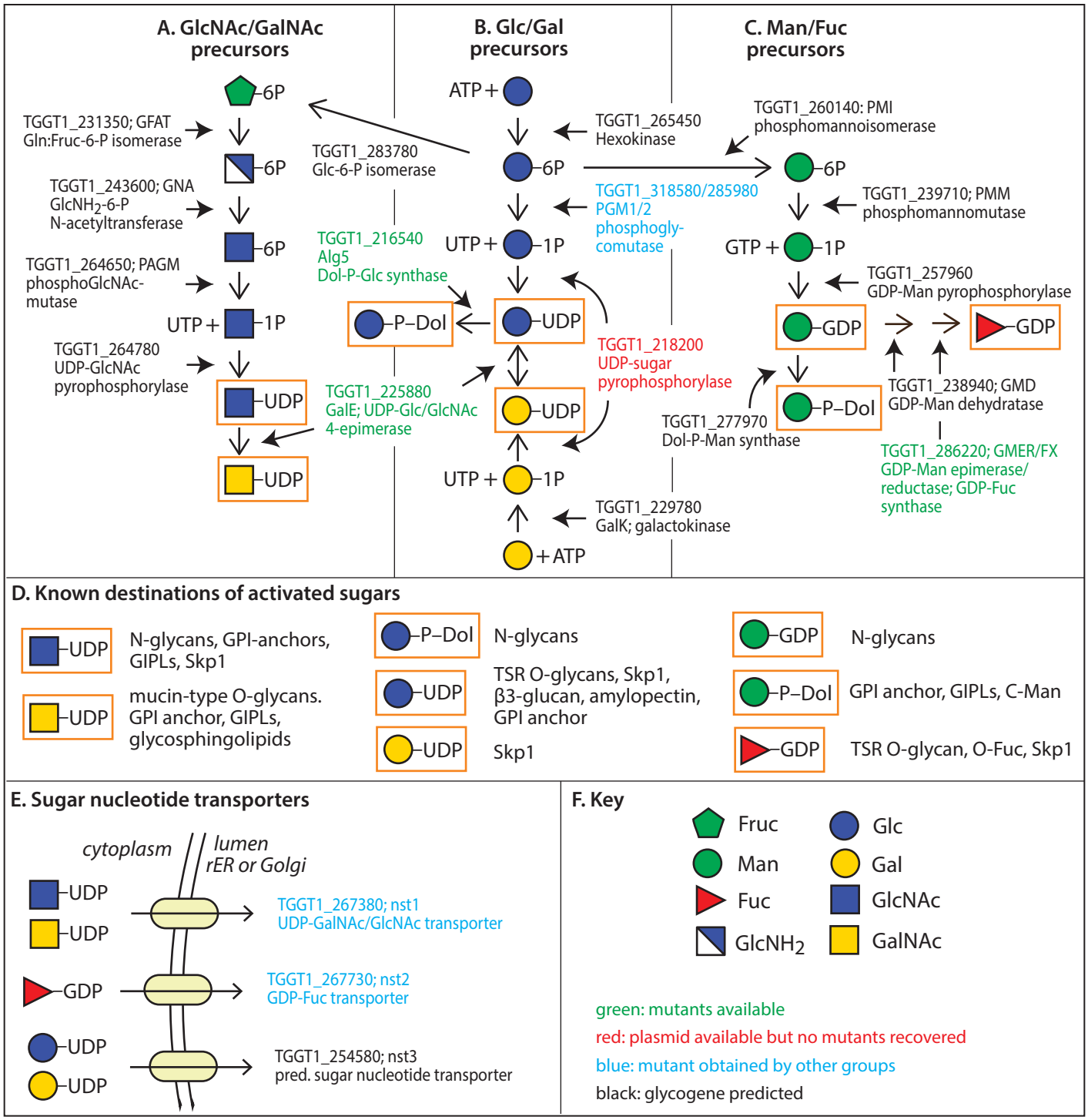

A. Hyp-glycan (Skp1)
TGGT1_310400
Gat1
TGGT1_260650
PgtA; β3GalT/α2FucT
α3
α3
α2
β3
α
Hyp
TGGT1_205060
Glt1
TGGT1_315885
Gnt1
B. O-fucose
α
S/T
TGGT1_273550
Spy
C. AAL
RH
(normal)
spyΔ
Scale
bar:
2 µm
AAL
AAL IMC3
AAL
AAL IMC3
E. amylopectin
α4
α4
α4
α4
α4
α6
α4
TGGT1_222800 putative
starch synthase
TGGT1_209960
α4glucan branching
TGGT1_316520
α4glucan branching
D. AAL
RH
spyΔ
RH
spyΔ
5x10^6
10^7
5x10^6
10^7
5x10^6
10^7
5x10^6
10^7
kDa
245
80
β-elimination
F. Other predicted cytoplasmic GTs
TGGT1_288390
β3GalNAcT-like
TGGT1_462962
β4GalT-like
TGGT1_217692
αGT-like
TGGT1_297720
α,α-trehalose-P
synthase/P'ase-like
Glc
Gal
Fuc
GlcNAc
green: mutants available
black: glycogene predicted
blue: mutant obtained by other groups/other methods

Figure 8


A. GlcNAc/GalNAc precursors
B. Glc/Gal precursors
C. Man/Fuc precursors
6P
TGGT1_231350; GFAT
Gln:Fruc-6-P isomerase
6P
TGGT1_243600; GNA
GlcNH2-6-P
N-acetyltransferase
6P
TGGT1_264650; PAGM
phosphoGlcNAc-
mutase
UTP +
1P
TGGT1_264780
UDP-GlcNAc
pyrophosphorylase
UDP
UDP
TGGT1_283780
Glc-6-P isomerase
TGGT1_216540
Alg5
Dol-P-Glc synthase
P–Dol
TGGT1_225880
GalE; UDP-Glc/GlcNAc
4-epimerase
ATP +
TGGT1_265450
Hexokinase
6P
TGGT1_318580/285980
PGM1/2
phosphogly-
comutase
UTP +
1P
UDP
TGGT1_218200
UDP-sugar
pyrophosphorylase
UDP
UTP +
1P
TGGT1_229780
GalK; galactokinase
+ ATP
TGGT1_260140: PMI
phosphomannoisomerase
6P
TGGT1_239710; PMM
phosphomannomutase
GTP +
1P
TGGT1_257960
GDP-Man pyrophosphorylase
GDP
GDP
P–Dol
TGGT1_277970
Dol-P-Man synthase
TGGT1_238940; GMD
GDP-Man dehydratase
TGGT1_286220; GMER/FX
GDP-Man epimerase/
reductase; GDP-Fuc
synthase
D. Known destinations of activated sugars
UDP
N-glycans, GPI-anchors, GIPLs, Skp1
UDP
mucin-type O-glycans. GPI anchor, GIPLs, glycosphingolipids
P–Dol
N-glycans
UDP
TSR O-glycans, Skp1, β3-glucan, amylopectin, GPI anchor
UDP
Skp1
GDP
N-glycans
P–Dol
GPI anchor, GIPLs, C-Man
GDP
TSR O-glycan, O-Fuc, Skp1
E. Sugar nucleotide transporters
cytoplasm
lumen
rER or Golgi
UDP
UDP
TGGT1_267380; nst1
UDP-GalNAc/GlcNAc transporter
GDP
TGGT1_267730; nst2
GDP-Fuc transporter
UDP
UDP
TGGT1_254580; nst3
pred. sugar nucleotide transporter
F. Key
Fruc
Glc
Man
Gal
Fuc
GlcNAc
GlcNH2
GalNAc
green: mutants available
red: plasmid available but no mutants recovered
blue: mutant obtained by other groups
black: glycogene predicted