(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
03 December 2020 (03.12.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/243675 A2

(51) International Patent Classification:
C12N 15/86 (2006.01)

(21) International Application Number:
PCT/US2020/035461

(22) International Filing Date:
30 May 2020 (30.05.2020)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/855,659 31 May 2019 (31.05.2019) US

(71) Applicant: **WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH** [US/US]; 455 Main Street, Cambridge, MA 02142 (US).

(72) Inventors: **WALDMAN, Ben**; 455 Main Street, Cambridge, MA 02142 (US). **LOURIDO, Sebastian**; 455 Main Street, Cambridge, MA 02142 (US).

(74) Agent: **BALICKY, Eric, M.** et al.; Hamilton, Brook, Smith & Reynolds, P.C., 530 Virginia Rd, P.O. Box 9133, Concord, MA 01742-9133 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:
— *of inventorship (Rule 4.17(iv))*

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: COMPOSITIONS AND METHODS FOR REGULATION OF CHRONIC TOXOPLASMA INFECTION

(57) Abstract: The present disclosure provides genetically altered protozoan parasites comprising a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. The genetically altered protozoan parasites can be utilized in vaccine compositions and in methods of treating apicomplexan parasite infection.

WO 2020/243675 A2

COMPOSITIONS AND METHODS FOR REGULATION OF CHRONIC TOXOPLASMA INFECTION

RELATED APPLICATION(S)

[0001] This application claims the benefit of U.S. Provisional Application No. 62/855,659, filed on May 31, 2019. The entire teachings of the above application are incorporated herein by reference.

GOVERNMENT SUPPORT

[0002] This invention was made with government support under Grant No. 1DP5OD017892 awarded by the National Institutes of Health. The government has certain rights in the invention.

BACKGROUND

[0003] *Toxoplasma gondii* chronically infects approximately a quarter of the world's population. *Toxoplasma* is a pathogen of humans and many domesticated animals. Recrudescence of latent infections can cause life-threatening disease in the immunocompromised and recurrent ocular lesions in the immunocompetent. Chronic infection is established when rapidly replicating tachyzoites differentiate into slow-growing bradyzoites, which form tissue cysts resistant to immune clearance and current therapeutics.

[0004] Treatments against *Toxoplasma* fail to clear chronic infections and are often poorly tolerated. Evidence suggests a high rate of congenital transmission in the absence of treatments. Thus, there is a need for new strategies to control *Toxoplasma* infection, particularly chronic infection.

SUMMARY OF THE INVENTION

[0005] In one aspect, the present invention provides genetically altered protozoan parasites comprising a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the parasite is an apicomplexan parasite, such as, e.g., a *Toxoplasma* parasite, a *Plasmodium* parasite, a *Hammondia* parasite, a *Neospora* parasite or a *Sarcocystis* parasite.

[0006] In some embodiments, the mutation is a deletion of all or a portion of the coding sequence of the *BFD1* gene. In some embodiments, the mutation is a deletion of the entire coding sequence of the *BFD1* gene.

[0007] In some embodiments, the mutation is a loss-of-function mutation. In some embodiments, the loss-of-function mutation is a null mutation.

[0008] In some embodiments, the mutation is a dominant negative mutation.

[0009] In some embodiments, the apicomplexan parasite is *Toxoplasma gondii*.

[0010] In some embodiments, the apicomplexan parasite is *Neospora caninum*.

[0011] In some embodiments, the apicomplexan parasite is *Sarcocystis neurona*.

[0012] In some embodiments, the apicomplexan parasite is *Hammondia hammondi*.

[0013] In some embodiments, the apicomplexan parasite is *Hammondia pardalis*.

[0014] In another aspect, the present invention provides compositions (e.g., vaccine compositions or pharmaceutical compositions) comprising (1) genetically altered protozoan parasites comprising a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite; and (2) a pharmaceutically-acceptable carrier.

[0015] In some embodiments, the composition is a vaccine composition.

[0016] In certain embodiments, the composition is a pharmaceutical composition.

[0017] In some embodiments, the vaccine composition further comprises an adjuvant.

[0018] In some embodiments, the vaccine composition comprises a live vaccine.

[0019] In some embodiments, the parasite expresses a heterologous antigen. For example, the heterologous antigen can be a cancer antigen. In some embodiments, the parasite expresses a therapeutic agent, such as, e.g., a peptide or a protein.

[0020] In another aspect, the present invention provides recombinant nucleic acid vectors comprising a nucleotide sequence encoding a BFD1 protein.

[0021] In some embodiments, the recombinant nucleic acid vector is an expression vector.

[0022] In a further aspect, the present invention provides host cells comprising recombinant nucleic acid vectors comprising a nucleotide sequence encoding a BFD1 protein.

[0023] In some embodiments, the host cell is *Toxoplasma gondii*.

[0024] In another aspect, the present invention provides methods of inducing an immune response to an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the apicomplexan parasite is *Toxoplasma gondii*.

[0025] In some embodiments, the subject is a human.

[0026] In some embodiments, the subject is a non-human mammal.

[0027] In some embodiments, the subject has an acute or chronic apicomplexan parasite infection.

[0028] In another aspect, the present invention provides methods of inhibiting (e.g., preventing or reducing likelihood of) a chronic apicomplexan parasite infection in a subject, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[0029] In another aspect, the present invention provides methods of treating a chronic infection by an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[0030] In an additional aspect, the present invention provides methods for inoculating a subject in need thereof with an apicomplexan parasite, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[0031] In a further aspect, the present invention provides methods of administering an antigen to a subject in need thereof comprising administering to the subject a composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite, and wherein the parasite comprises an antigen. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[0032] In some embodiments, the parasite is genetically altered to comprise the antigen.

[0033] In certain embodiments, the antigen is a cancer antigen.

[0034] The methods and compositions herein described can be used in pharmaceutical, medical, and veterinary applications, as well as fundamental scientific research and methodologies, as would be identifiable by a skilled person upon reading of the present disclosure. Other features and advantages of the invention will be understood by reference to the drawings, detailed description and examples that follow.

BRIEF DESCRIPTION OF THE DRAWINGS

[0035] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawings will be provided by the Office upon request and payment of the necessary fee.

[0036] The foregoing will be apparent from the following more particular description of example embodiments, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating embodiments.

[0037] FIGs. 1A-1H show that a forward genetic screen identifies a putative regulator of *T. gondii* differentiation. FIG. 1A. Construction of a differentiation reporter strain constitutively expressing RFP and Cas9. Following alkaline stress, the strain expresses mNeonGreen (mNG) under the regulation of a bradyzoite promoter, along with another marker of differentiation (SAG2Y). Images taken after 48 h of alkaline stress. Scale bar is 10 μm. FIG. 1B. Efficient gene disruption is mediated by constitutive Cas9 expression. Transfection and selection for a gRNA targeting the surface antigen SAG1 resulted in gene disruption in 98% of the resulting population. $n = 3$ biological replicates. 92–102 vacuoles were scored for each replicate. FIG. 1C. Percent of alkaline stressed reporter parasites expressing mNG, quantification by FACS. $n = 2$-4 biological replicates, Mean ± SD plotted. FIG. 1D. RNA-sequencing and differential expression (DE) analysis identifies 1311 genes significantly upregulated and 933 genes significantly downregulated in bradyzoites (adjusted $p < 0.001$). Parasites were allowed to invade host cells for 24 h, and then switched to either unstressed or stressed conditions. After 24 h of unstressed growth or 48 h of stressed growth, ~100,000 tachyzoites (unstressed, mNG^- parasites) or bradyzoites (stressed, mNG^+ parasites) were isolated by FACS. $n = 3$ independent experiments. FIG. 1E. Screening and analysis workflow. Libraries assembled from proteins containing nucleic acid-binding domains, and sub-divided based on differential expression (L1) or the presence of DNA-binding domains commonly found in transcription factors (L2). Libraries contain 5 gRNAs against each gene, control gRNAs against essential/non-essential genes, noncutting gRNAs, and gRNAs targeting the differentiation reporter. Following transfection, parasites were passaged four times in selective media, then split into alkaline (stressed) or standard culture conditions (unstressed). After 10 days, gRNAs were sequenced from bulk populations, FACS-sorted mNG^+ parasites from the stressed population, and from the input library. The $\log_2$ fold-changes from the input library to the final unstressed or stressed mNG^+ timepoints are

defined as the fitness or differentiation scores, respectively. Candidate genes will be depleted specifically from the mNG^+ stressed population (low differentiation score relative to their fitness score). FIGs. 1F and 1G show results for the screen with L1 (FIG. 1F) and L2 (FIG. 1G) show that in addition to the control (mNG), only TGME49_200385 (bradyzoite formation deficient 1; BFD1) displays a low differentiation score compared to its fitness score. FIG. 1H. Fitness and differentiation scores for individual gRNAs in L2.

[0038] FIGs. 2A-2I show that BFD1 is a nuclear factor necessary for differentiation. FIG. 2A. Neighbor-joining tree showing the phylogenetic relationship of the concatenated SANT domains from BFD1 and its closest homologs in other apicomplexans and humans. BFD1 forms a clade with human c-Myb (green), distinct from CDC5L-like sequences (blue). Bootstrap values for 1000 trials are displayed. FIG. 2B. Diagram of BFD1 and human c-Myb highlighting the SANT domains (green). The DNA-binding repeats of BFD1 are similar to the second and third repeats of c-Myb. FIG. 2C. Wild-type parasites transfected with a C-terminally Ty-tagged BFD1 cDNA, driven by the TUB1 promoter, show nuclear localization of the transgene. DNA stained with Hoechst (blue) and Ty1 immunostained with BB2 (red). Scale bar is 3 μm. FIG. 2D. Generation of a *BFD1* knockout through homologous recombination in ME49Δ*KU80*. The entire coding sequence of *BFD1* was replaced with an mNeonGreen-expression cassette using gRNAs against the 5′ and 3′ ends. The locus was fully sequenced to confirm complete deletion of *BFD1*. FIG. 2E. Plaque assay of parental ME49Δ*KU80* (WT) or ME49Δ*KU80*Δ*BFD1* (Δ*BFD1*) shows no defect in growth under unstressed conditions. Images taken 14 days post plaquing. FIG. 2F. *BFD1* parasites fail to differentiate. Differentiated parasites were stained with FITC-labeled *Dolichos biflorus* lectin (DBL). Images taken after 48 h of alkaline stress. Scale bar is 10 μm. FIGs. 2G-2I show quantification of differentiation in the parental strain and Δ*BFD1* parasites following 48 hours of alkaline stress (FIG. 2G), 48 hours of Compound 1 treatment (FIG. 2H) or occurring spontaneously (FIG. 2I). Mean ± SD of n = 3–4 biological replicates. % DBL positive vacuoles calculated from at least 100 vacuoles counted for each replicate.

[0039] FIGs. 3A-3J show that BFD1 is required for initiation of differentiation and expression of bradyzoite genes. FIG. 3A. Workflow for single-cell RNA-sequencing with Seq-Well. Parasites were allowed to invade host cells for 4 hours, and then shifted to unstressed or stressed conditions. After 24, 48, or 72 additional hours of growth, parasites were mechanically released and filtered, and approximately 12,000 parasites were loaded per Seq-Well array (1 array per strain under unstressed conditions, 2 arrays per strain under stressed conditions). FIG. 3B. Clustering of unstressed parasites of both genotypes after 72 h

of growth. Clustering performed by shared nearest neighbor (SNN) using principal components 1–6. Visualization by t-distributed stochastic neighbor embedding (t-SNE), each dot is a single cell. Arrow represents imputed direction of cell cycle. FIG. 3C. Six of seven clusters differentially express known cell cycle–regulated genes. Expression profiles for differentially expressed genes within each cluster in B) based on a microarray dataset of synchronized tachyzoites [1]. Expression profiles were scaled, mean-centered and averaged for each cluster. Differential expression determined by Wilcoxon rank test. Central cluster in gray does not have any differentially expressed genes. FIG. 3D. Proportion of cells in G1 or S/M clusters match the previously determined 60:40 ratio. FIG. 3E. The canonical stage-specific genes *SAG1* and *BAG1* show mutually exclusive expression. t-SNE projection of parasites from all timepoints, genotypes and growth conditions. Visualization downsampled to 500 cells per timepoint/genotype/growth condition (6,000 cells total). FIG 3F. t-SNE projection as in E), separated by sample of origin. WT and *BFD1* parasites cluster similarly under unstressed conditions, while stressed parasites cluster differently by genotype. Unstressed parasites of both genotypes are found primarily in *SAG1*$^{+}$ clusters. Under stressed conditions, only WT parasites are found in *BAG1*$^{+}$ clusters. FIG. 3G. Clustering of WT and *BFD1* parasites from unstressed and stressed cultures at 72 h timepoint. Clustering performed using principal components 1–18, visualized by t-SNE, and colored according to similarity to the tachyzoite clusters in B). FIG. 3H. Stressed *BFD1* parasites fail to upregulate canonical bradyzoite markers, and express tachyzoite markers weakly. Average scaled expression of the top 20 highly expressed up and downregulated genes identified by bulk RNA-sequencing. Each column is a single cell. Data downsampled to 1,000 cells per genotype/growth condition (4,000 cells total). FIG. 3I. Plotting cells in stressed, non-tachyzoite clusters according to the major cell-cycle varying principal components 1 and 3 reveals that only WT parasites are dividing in stressed conditions. Cells colored according to clustering in G). FIG. 3J. Under unstressed conditions, WT and *BFD1* replicate similarly as tachyzoites. Following alkaline stress induction, only WT parasites are able to differentiate into bradyzoites and continue dividing, while *BFD1* parasites seem to exit the cell cycle and stop replicating.

[0040] FIGs. 4A and 4B show that overexpression of BFD1 is sufficient to induce differentiation. FIG. 4A. Constructs and experimental workflow for transient overexpression of BFD1. cDNA of BFD1 (BFD1^{WT}) or BFD1 lacking its DNA-binding domain (BFD1$^{\Delta DBD}$) is regulated by the strong constitutive TUB1 promoter. WT or *BFD1* parasites were transfected with either BFD1^{WT} or BFD1$^{\Delta DBD}$ and fixed after 48 hours of growth under unstressed conditions. Parasites were immunostained for Ty1 (magenta) and for

differentiation with FITC-conjugated DBL (green). FIG. 4B. Transient overexpression of $BFD1^{WT}$ (WT), but not $BFD1^{\Delta DBD}$, is sufficient to induce differentiation in a majority of WT and *BFD1* parasites expressing the transgene. Ty^{+} vacuoles were identified and then scored for DBL positivity. Examples shown of vacuoles scored as DBL^{+}, $DBL^{+/-}$, and DBL^{-}. Scale bar is 10 μm. $n = 2$ independent replicates, 17-61 vacuoles counted per replicate. Mean ± SD plotted.

[0041] FIGs. 5A-5D show stage-specific RNA-sequencing of *Toxoplasma.* FIG. 5A. Sample FACS plots of the C16-B3 reporter strain at 24 or 48 h growth under unstressed or stressed conditions. 10,000 events per plot. FIG. 5B. Principal component analysis suggests the majority (98%) of variance observed in gene expression is due to growth condition. FIG. 5C. Comparing differential expression in this study to existing datasets shows good agreement among highly regulated genes, but only moderate correlation (spearman r^2 ~.54) overall. FIG. 5D. The majority of genes called as differentially expressed exclusively in the data set are more lowly expressed. Colors assigned by rank of baseMean expression as calculated by DESeq2.

[0042] FIG. 6 shows that integration rates in a non-laboratory-attenuated strain of *Toxoplasma* are low. Cas9-expressing ME49 parasites were selected for integration of a gRNA targeting SAG1. Plaquing efficiency post-selection was compared to pre-transfection viability rates to obtain a viability normalized integration rate. $n = 3$ independent experiments. Mean ± SD plotted.

[0043] FIG. 7 shows generation of a Δ*BFD1* reporter strain through Cas9-mediated frameshift. Transfection of single gRNA targeting the first exon of TGME49_200385 (*BFD1*) into the C16-B3 reporter strain allowed isolation of a strain with a frameshift mutation, resulting in a premature stop codon at amino acid 251.

[0044] FIG. 8 shows updated gene model and protein sequence of TGME49_200385. cDNA sequencing suggests the 5^{th}, 6^{th}, and 7^{th} exons as annotated on ToxoDB v. 42 are a single exon, resulting in a change of reading frame of the final exon. DNA-binding domains (SM00717) highlighted in blue.

[0045] FIG. 9 shows phylogenetic analysis of *Toxoplasma* SANT/myb-like domains. Neighbor-joining phylogenetic tree of SANT/Myb-like DNA-binding domains (SM00717) present in representative Apicomplexan genomes, along with human c-Myb and CDC5L. Alignment performed using ClustalW. Scale bar is substitutions per site.

[0046] FIG. 10 shows a schematic for generation of an early passage ME49ΔKu80. Two gRNAs were designed against both the 5′ and 3′ of the *KU80* locus. Transfection all four

gRNAs into an early passage ME49 strain followed by immediate subcloning allowed recovery of a clone that had deleted the intervening sequence.

[0047] FIGs. 11A-11C show gene detection is maximized and rRNA content minimized at the 72 hour timepoint. FIG. 11A. Distribution of UMIs across single cells from indicated samples and timepoints. Pre-processing quality control cutoffs required a minimum of 200 and maximum of 10,000 UMIs. FIG. 11B. As in FIG. 11A, but collapsed to unique genes detected. FIG. 11C. Proportion of UMIs corresponding to ribosomal genes. Pre-processing quality control cutoffs allowed a maximum of 40% rRNA reads.

[0048] FIGs. 12A-12B show TGME49_312330 and TGME49_208740 are novel cyst wall proteins robustly expressed earlier than canonical bradyzoite markers. FIG. 12A. Violin plots of expression of indicated genes in wildtype parasites after 24, 48 or 72 h growth under unstressed or stressed conditions. FIG. 12B. Endogenous tagging of TGME49_312330 and TGME49_208740 localizes both to the cyst wall. Pictures taken 72h post alkaline stress. Scale bar is 10 μm.

[0049] FIGs. 13A-13C show that Δ*BFD1* parasites do not overlap with BAG1$^+$ clusters. FIG. 13A. Mutually exclusive expression of the tachyzoite-specific SAG1 and the bradyzoite-specific BAG1 demonstrate strong separation of these two stages. Visualization by t-SNE, colored by normalized log-scale UMIs per cell. FIG. 13B. Coloring cells by genotype of origin shows distinct localization of wildtype and Δ*BFD1* parasites under stressed, but not unstressed, conditions. Notably, only wildtype parasites overlap with BAG1$^+$ clusters. FIG. 13C. *BFD1* is expressed more highly in BAG1$^+$ clusters.

[0050] FIGs. 14A-14D show that the majority of variance is driven by cell-cycle and stage-specific genes. FIG. 14A. Plotting the 18 principal components determined to be statistically significant by permutation analysis shows the first three explain the majority (66.4%) of variance seen at the 72 h timepoint. Statistical significance determined by JackStraw() as implemented by Seurat, p-value < .001. FIG. 14B. PC1 is driven by cell-cycle regulated genes, including previously known mid-S/M-specific markers of *Toxoplasma*, such as members of the ROP/RON and IMC genes families. Top 2000 cells for each loading, with expression shown for the top 30 gene loadings for each component. FIG. 14C. PC2 is driven by strongly tachyzoite-specific (e.g. SAG1, AMA1) or bradyzoite-specific (BAG1, LDH2, ENO1) genes. FIG. 14D. PC3 is driven cell-cycle regulated genes, particularly by genes associated with later events in the S/M phase, such as GAP45 and IMC8.

[0051] FIG. 15 shows that single-cell RNA-sequencing identifies markers specific to replicating bradyzoites. t-SNE visualization of selected markers specific to replicating bradyzoites. Colored by normalized log-scale UMIs per cell.

[0052] FIG. 16 shows stressed Δ*BFD1* parasites show aberrant morphology at later timepoints. Under alkaline stress conditions, Δ*BFD1* vacuoles frequently present disordered or bloated appearances, suggesting loss of viability. Pictures taken 72h post alkaline stress. GAP45 is a marker for the inner membrane complex. Scale bar is 10 μm.

[0053] FIGS. 17A-17D show BFD1 is a nuclear factor necessary for differentiation in cell culture. FIG. 17A. Generation of Δ*BFD1* and $BFD1^{WT}$ or $BFD1^{\Delta MYB}$ complemented parasites. To create Δ*BFD1* parasites, the endogenous coding sequence was replaced with a fluorescent cassette. The knockout was complemented with a wild-type (WT) (*BFD1WT*) or DNA-binding deficient ($BFD1^{\Delta MYB}$) Ty-tagged allele at the endogenous locus. FIG. 17B. Plaque assays of indicated strains grown under unstressed conditions for 14 days. Scale bar is 1 cm. FIG. 17C. Representative vacuoles after 48 h of alkaline stress. FITC-labeled *Dolichos biflorus* lectin (DBL) specifically stains differentiated vacuoles. Ty was stained with BB2 (magenta), and DNA was stained with Hoechst (blue). Scale bar is 10 mm. FIG. 17D. Quantification of differentiation in WT, knockout, and complemented parasites following 48 h of alkaline stress, 48 h of compound 1 treatment, or occurring spontaneously under unstressed conditions in the same time frame. The mean ± SD was plotted for n = 3–8 biological replicates, with percentage of DBL positive vacuoles calculated from at least 100 vacuoles per replicate. ****$p < 0.0001$, *$p < 0.05$, Student's one-tailed t test.

[0054] FIGS. 18A-E show that BFD1 is necessary for formation of brain cysts in mice. FIG. 18A. Timeline of mouse infections. Groups of CD-1 female mice were inoculated i.p. with 500 tachyzoites per animal from each strain or mock inoculated with PBS. Cyst formation was assayed in moribund animals starting 2 weeks post-infection and in all surviving animals at 5 weeks post-infection. FIG. 18B. Mean normalized weights of animals in each group. Graph represents mean ± SEM for all surviving animals at a given time point. Graphs are for n = 5 mock-inoculated mice and n = 15 for each parasite strain. FIG. 18C. Survival curve of animals in FIG. 18B. FIG. 18D. Representative cysts from WT and Δ*BFD1::*$BFD1^{WT}$-infected animals. The cyst wall was stained with DBL (green) and individual parasites with anti-CDPK1 (magenta). Scale bar is 20 mm. FIG. 18E. Cyst burden per animal, denoting those sacrificed before (open circles) or after (closed circles) 5 weeks of infection. Cysts per brain were estimated from counting four blinded replicates, with a limit

of detection of 56–71 cysts per brain, depending on the volume of the sample analyzed. Mean is plotted for each group. **p < 0.01, Student's one-tailed t test. See also FIGS. 19A-19E.

[0055] FIGS. 19A-E (related to FIGS. 18A-E) show virulence and brain cyst formation by Δ*BFD1* parasites in CBA/J Mice: FIGS. 19A, 19B. In groups of 5, female CBA/J mice were inoculated with 100 or 2,000 tachyzoites i.p. of WT, Δ*BFD1*, or Δ*BFD1::BFD1*^{WT} and surviving animals were sacrificed 2 weeks post-infection to assay brain cyst formation (FIG. 19A). Cyst burdens were estimated by counting 4 blinded samples from each animal. Mean ± SD is plotted with each dot representing an animal; ****p < 0.0001, Student's one-tailed t test (FIG. 19B). FIGS. 19C–19E. In groups of 5, female CBA/J mice were inoculated with 500 or 10,000 tachyzoites i.p of WT, Δ*BFD1*, or Δ*BFD1::BFD1*^{WT}. Starting at 3 weeks post-infection, brains were isolated from moribund animals, and at 5 weeks post-infection all surviving animals were sacrificed (FIG. 19C). Survival curve of animals infected with 10,000 (dotted lines) or 500 (solid lines) tachyzoites (FIG. 19D). Brain cyst burden of moribund or sacrificed animals, estimated by counting 4 blinded samples from each animal. Mean ± SD is plotted with each dot representing an animal (FIG. 19E).

DETAILED DESCRIPTION

[0056] A description of example embodiments follows.

[0057] Several aspects of the invention are described below, with reference to examples for illustrative purposes only. It should be understood that numerous specific details, relationships, and methods are set forth to provide a full understanding of the invention. One having ordinary skill in the relevant art, however, will readily recognize that the invention can be practiced without one or more of the specific details or practiced with other methods, protocols, reagents, cell lines and animals. The present invention is not limited by the illustrated ordering of acts or events, as some acts may occur in different orders and/or concurrently with other acts or events. Furthermore, not all illustrated acts, steps or events are required to implement a methodology in accordance with the present invention. Many of the techniques and procedures described, or referenced herein, are well understood and commonly employed using conventional methodology by those skilled in the art.

[0058] Unless otherwise defined, all terms of art, notations and other scientific terms or terminology used herein are intended to have the meanings commonly understood by those of skill in the art to which this invention pertains. In some cases, terms with commonly understood meanings are defined herein for clarity and/or for ready reference, and the inclusion of such definitions herein should not necessarily be construed to represent a

substantial difference over what is generally understood in the art. It will be further understood that terms, such as those defined in commonly-used dictionaries, should be interpreted as having a meaning that is consistent with their meaning in the context of the relevant art and/or as otherwise defined herein.

[0059] The terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. As used herein, the indefinite articles "a", "an" and "the" should be understood to include plural reference unless the context clearly indicates otherwise.

[0060] *Toxoplasma gondii* chronically infects approximately a quarter of the world's population. Recrudescence of latent infections can cause life-threatening disease in the immunocompromised and recurrent ocular lesions in the immunocompetent. Chronic infection is established when rapidly replicating tachyzoites differentiate into slow-growing bradyzoites, which form tissue cysts resistant to immune clearance and current therapeutics. Despite its central role in the *Toxoplasma* life cycle, the molecular basis of differentiation is not understood.

[0061] The length of infection is a critical parameter in the evolutionary fitness of infectious diseases. Pathogens can extend the period of infection by establishing a latent or chronic state, avoiding clearance through slow replication, altered immunogenicity, and a diminished impact on the host. Oftentimes, these reservoirs are resistant to chemotherapy due to decreased metabolic rates. Such persistent stages can recrudesce or contribute to transmission, and are important barriers to curing and eradicating infectious diseases.

[0062] The life cycles of single-celled protist parasites are marked by distinct developmental stages adapted to the specific stages of their complex life cycles. Within the phylum Apicomplexa, several examples of chronic stages play important roles in the life cycles of these pathogens. *Plasmodium vivax* hypnozoites in the liver are resistant to most antimalarial therapies, leading to long periods of latency, and complicating eradication efforts [2]. *Toxoplasma gondii* tachyzoites are capable of invading any nucleated cell in any warm-blooded animal, disseminating throughout the body and causing pathology through lysis of host cells. A proportion of tachyzoites differentiate into slower-growing bradyzoites, forming intracellular cysts with a tropism for brain and muscle tissue. Tissue cysts cannot be cleared by the immune system or by current therapies, and as a result up to ¼ of the world's population is thought to be chronically infected with *Toxoplasma* [3]. Tissue cysts are orally infectious to both the definitive felid host and to other intermediate hosts, providing flexibility in route of transmission. Approximately 2% of infections result in ocular lesions—

a leading cause of infectious blindness—with high rates of reactivation from chronic stages that persist after treatment. *Toxoplasma* infection is life threatening in immunocompromised individuals, and a majority of these cases result from recrudescent infections [4].

[0063] Major changes accompany differentiation of *Toxoplasma* from rapidly proliferating tachyzoites to cyst-forming bradyzoites. The parasitophorous vacuole *Toxoplasma* replicates within is modified into a heavily glycosylated cyst wall, containing many stage-specific protein products of unknown function [5-8]. Parasite metabolism changes drastically, relying on anaerobic glycolysis instead of aerobic respiration and accumulating cytoplasmic starch granules [9-11]. Underpinning these dramatic changes in lifestyle, between several hundred and several thousand genes have been identified as differentially regulated between tachyzoites and *in vitro* or *in vivo* bradyzoites [12-21]. Although differentiation can be induced *in vitro* through a variety of methods, including alkaline pH, heat shock, small molecules, and nutrient starvation, the molecular mechanisms driving bradyzoite differentiation remain poorly understood [22-25].

[0064] Attempts to enrich for mutants no longer able to differentiate have generated strains with decreased rates of stage conversion, though linking these phenotypes to inactivation of individual genes has proven challenging [26,27]. Deletion or chemical inhibition of histone modifying proteins has resulted in differentiation defects [17,28-31]. Interference with translational control through deletion of RNA-binding proteins or chemical modulation of eIF2α phosphorylation also impairs differentiation [32-35]. A single validated class of Apicomplexan transcription factors, the AP2 DNA-binding proteins (ApiAP2s), has been extensively investigated for potential regulators of differentiation. Knockouts of individual ApiAP2s modulate, but ultimately fail to completely ablate, bradyzoite differentiation, leading to the view that no master transcriptional regulator of this process exists in *Toxoplasma* [36-41].

[0065] Differentiation from tachyzoites to bradyzoites establishes chronic Toxoplasma infection; however, the molecular pathways regulating this transition have remained unclear, despite evidence that disparate inputs—heat shock, alkaline stress, and nutrient starvation—converge on a common transcriptional program.

[0066] Using bulk and single-cell RNA sequencing, differentiation is characterized in unprecedented detail. To functionally investigate this transition, over 200 putative DNA-binding proteins were screened, identifying a single factor, BFD1, as indispensable for differentiation. BFD1 knockout parasites are normal under standard conditions but fail to differentiate under all induction conditions tested. Overexpression of BFD1 is sufficient to

induce differentiation under standard conditions in both wildtype and knockout parasites, demonstrating its role as a master regulator of bradyzoite formation in *Toxoplasma*.

[0067] By profiling enriched differentiated populations, transcriptional differences between tachyzoites and bradyzoites were captured with greater sensitivity and dynamic range than achieved by previous datasets [15-18,20,21,42]. These changes likely reflect a combination of factors including altered replication, nutrient availability, and general stress responses, in addition to the bradyzoite differentiation program. The genetic handle afforded by BFD1 on differentiation will help disentangle the contributions of these variables.

[0068] A single family of DNA-binding proteins—the ApiAP2s—has been investigated for their role as *T. gondii* transcription factors and mediators of differentiation. While the phenotypes associated with many ApiAP2 mutants are striking, no single gene knockout has resulted in a complete block in differentiation, leading to the assumption that no single factor regulates bradyzoite development in *T. gondii* [36-41]. By screening a wider range of putative nucleic-acid binding proteins, especially those containing well-conserved DNA-binding motifs such as zinc finger and Myb-like domains, it was observed that inactivation of BFD1 completely ablates bradyzoite formation. This does not preclude important roles for ApiAP2 proteins as downstream mediators of the differentiation program. Toxoplasma encodes 13 other proteins containing SANT/Myb-like domains, suggesting the existence of a second extensive transcription factor family. Myb domain–containing proteins are widespread among eukaryotes, and have been implicated in the regulation of encystation in *Entamoeba* and *Giardia*, along with a wide variety of stress responses in plants [43-47]. In humans, c-Myb is thought to function as a pioneer transcription factor, binding to chromatin and recruiting histone acetyltransferases to commit cells to specific hematopoietic lineages [48,49]. Among apicomplexan parasites, a Myb domain–containing protein has been identified as important for erythrocytic growth of *P. falciparum* [50,51], suggesting that other family members will likely play important roles throughout the phylum.

[0069] Single-cell RNA-sequencing enables profiling of thousands of cells across asynchronous processes, and has been successfully used to examine commitment to sexual differentiation in *Plasmodium spp.* [52,53]. Implementing these approaches in Toxoplasma retained information about cell-cycle residency and timing of gene expression that is lost in bulk analyses. Single-cell sequencing allowed identification of novel markers specific to actively replicating bradyzoites and genes expressed during the earliest stages of differentiation. This detailed view of differentiation revealed that BFD1 knockout parasites progress normally throughout the tachyzoite cycle but completely fail to initiate bradyzoite

differentiation. *BFD1* therefore stands out from other genes known to influence differentiation for its complete essentiality during the process.

[0070] As a necessary and sufficient regulator of differentiation, BFD1 provides a focal point for the molecular mechanisms underlying differentiation. In *Plasmodium*, identification of AP2-G as the master transcriptional regulator of gametogenesis has permitted placement of multiple genes observed to affect sexual differentiation into a unified regulatory framework, and allowed directed investigation into their mechanisms of action [54-60]. The ability to induce synchronized sexual differentiation through conditional overexpression of AP2-G has allowed for finer temporal mapping of the gene expression changes accompanying gametogenesis, and application of a similar approach in Toxoplasma is now possible using BFD1[61].

[0071] The transcriptional profiling reveals that BFD1 is expressed in the 75th percentile in tachyzoites, and expression at this level or higher is corroborated by many other RNA-sequencing datasets [19-21,62-65]. BFD1 expression therefore does not appear to be stage specific despite a modest 1.5- to 3.6-fold upregulation in bradyzoites observed by bulk or single-cell RNA sequencing—the former below the cutoff for significance. These results suggest that regulation of BFD1 is post-transcriptional. Preferential translation of some transcripts under stress conditions has been reported, and mutations in RNA-binding proteins have resulted in severe defects that suggest an important role for translational control during differentiation [29,32]. Post-translational modifications, such as phosphorylation or acylation may provide additional layers of regulation by influencing BFD1 stability or function.

[0072] Mutations resulting in decreased rates of differentiation in cell culture generally display more profound defects in mice; however, the specificity of the BFD1 phenotype will help define the role of bradyzoites in pathogenesis and immunological memory during *Toxoplasma* infection [36,40,66,67]. The presence of BFD1 orthologs in other agriculturally significant parasites, such as *Neospora caninum* and *Sarcocystis neurona*, suggests ΔBFD1 parasites are an ideal attenuated vaccine strain—capable of proliferating robustly yet unable to enter a chronic state. Moreover, modulation of BFD1 holds clinical and biotechnological potential, since chronic infections represent a major barrier to both the treatment of *Toxoplasma* and its use in delivery of heterologous antigens and protein-based therapeutics.

[0073] The present invention is directed to compositions and methods for the treatment of apicomplexan parasite infection, such as *Toxoplasma* infection.

[0074] The present invention incorporates new approaches to investigate the tachyzoite to bradyzoite transition. A differentiation reporter strain compatible with Cas9-mediated

forward genetic screens was generated, and bulk and single-cell methods to characterize the transcriptional changes that underlie *Toxoplasma* differentiation were developed. The present disclosure provides a single factor necessary and sufficient for the initiation of the differentiation program. Characterization of this master regulator of bradyzoite formation represents an important step in the understanding of the establishment and maintenance of chronic *Toxoplasma* infection.

[0075] In one aspect, the present invention provides genetically altered protozoan parasites comprising a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. As such, attenuated parasites are provided. In some embodiments, the parasite is an apicomplexan parasite, such as e.g., a *Toxoplasma* parasite, a *Plasmodium* parasite, a *Hammondia* parasite, a *Neospora* parasite or a *Sarcocystis* parasite.

[0076] In some embodiments, the mutation is a deletion of all or a portion of the coding sequence of the *BFD1* gene (the sequence of the *BFD1* gene is provided in SEQ ID NO: 1). In some embodiments, the mutation is a deletion of the entire coding sequence of the *BFD1* gene. In other embodiments, the mutation is a deletion of a portion of the coding sequence of the *BFD1* gene. For example, the deletion can be a deletion of a portion of the *BFD1* gene comprising at least about 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more of the *BFD1* gene.

[0077] As described herein, the BFD1 protein contains two tandem SANT/myb-like DNA-binding domains (SMART accession 00717), flanked by large extensions lacking identifiable motifs. Collectively, the two tandem SANT/myb-like DNA-binding domains are referred to herein as the "BFD1 DNA binding domain." Individually, the SANT/myb-like DNA-binding domains are referred to herein as a "Myb-like DNA binding domain," "Myb-like domain" or "Myb domain." Accordingly, in some embodiments, the mutation is a deletion of a portion of the *BFD1* gene that encodes the BFD1 DNA binding domain (e.g., a deletion of a portion encoding both Myb-like DNA binding domains, a deletion of a portion encoding both Myb-like DNA binding domains along with the flanking extensions). In certain embodiments, the deletion encompasses a portion of the *BFD1* gene encoding amino acids 921–1019 of the BFD1 protein.

[0078] In other embodiments, the mutation is a deletion of a portion of the *BFD1* gene encoding one Myb-like DNA binding domain (e.g., a deletion of a portion encoding one of the Myb-like domains while leaving the other Myb-like domain intact, a deletion of a portion encoding one Myb-like domain and a portion of the second Myb-like domain). An example

of a mutation comprising a deletion of a portion of the *BFD1* gene encoding the BFD1 DNA binding domain is described in the Examples herein, and is referred to interchangeably as the ΔMYB or ΔDBD.

[0079] In further embodiments, the mutation can be one or more nucleotide base substitutions, one or more nucleotide base deletions, or one or more insertion (of one or more nucleotide bases or constructs (e.g., reporter genes)). In some embodiments, the insertion or deletion results in a frame-shift that changes the reading of subsequent codons, altering the entire amino acid sequence that follows the mutation. In some embodiments, the mutation can introduce a premature stop codon. In some embodiments, a nucleotide sequence of the *BFD1* gene comprises a nucleic acid sequence having at least 70% sequence identity to the sequence of SEQ ID NO: 1. In some embodiments, a nucleotide sequence of the *BFD1* gene comprises a nucleic acid sequence having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.

[0080] In some embodiments, the mutation is a loss-of-function mutation. As used herein, a "loss-of-function" mutation refers to a mutation that results in the altered gene product's loss of a specific biological function in which it is involved, such as, e.g., differentiation. Loss-of-function mutations can elicit a complete loss of function of the altered gene product (a "null mutation") or a partial loss of function. In certain embodiments, the loss-of-function mutation is a null mutation.

[0081] In some embodiments, the mutation is a dominant negative mutation. As used herein, a "dominant negative" mutation results in an altered gene product that acts antagonistically to the wild-type allele.

[0082] As used herein, a "coding sequence" of a gene consists of the nucleotide residues of the coding strand of the gene and the nucleotides of the non-coding strand of the gene which are homologous with or complementary to, respectively, the coding region of an mRNA molecule which is produced by transcription of the gene.

[0083] A "coding region" of an mRNA molecule also consists of the nucleotide residues of the mRNA molecule which are matched with an anti-codon region of a transfer RNA molecule during translation of the mRNA molecule or which encode a stop codon. The coding region may thus include nucleotide residues corresponding to amino acid residues which are not present in the mature protein encoded by the mRNA molecule (e.g., amino acid residues in a protein export signal sequence).

[0084] "Encoding" refers to the inherent property of specific sequences of nucleotides in a polynucleotide, such as a gene, a cDNA, or an mRNA, to serve as templates for synthesis of other polymers and macromolecules in biological processes having either a defined sequence of nucleotides (e.g., rRNA, tRNA and mRNA) or a defined sequence of amino acids and the biological properties resulting therefrom. Thus, a gene encodes a protein if transcription and translation of mRNA corresponding to that gene produces the protein in a cell or other biological system. Both the coding strand, the nucleotide sequence of which is identical to the mRNA sequence, and the non-coding strand, used as the template for transcription of a gene or cDNA, can be referred to as encoding the protein or other product of that gene or cDNA. Unless otherwise specified, a nucleotide sequence encoding an amino acid sequence includes all nucleotide sequences that are degenerate versions of each other and that encode the same amino acid sequence. Nucleotide sequences that encode proteins and RNA may include introns.

[0085] In some embodiments, the apicomplexan parasite is *Toxoplasma gondii*.

[0086] In some embodiments, the apicomplexan parasite is *Neospora caninum*.

[0087] In some embodiments, the apicomplexan parasite is *Sarcocystis neurona*.

[0088] In another aspect, the present invention provides compositions (e.g., vaccine compositions or pharmaceutical compositions) comprising (1) genetically altered protozoan parasites comprising a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite; and (2) a pharmaceutically-acceptable carrier.

[0089] As used herein, "vaccine composition" refers to a composition comprising a microbial immunogen (e.g., antigen or collection of antigens) that is capable of eliciting an adaptive immune response and/or immune memory against the microbe. The vaccine composition can further comprise a pharmaceutically acceptable carrier.

[0090] "Pharmaceutically acceptable" refers to those properties and/or substances which are acceptable to the patient from a pharmacological/toxicological point of view and to the manufacturing pharmaceutical chemist from a physical/chemical point of view regarding composition, formulation, stability, patient acceptance and bioavailability. "Pharmaceutically acceptable carrier" refers to a medium that does not interfere with the effectiveness of the biological activity of the active ingredient(s) and is not toxic to the host to which it is administered.

[0091] As used herein, the term "antigen" refers to any molecule capable of generating an immune response, such as a peptide, polypeptide, protein, polysaccharide, lipid, cell, cancer cell (such as a self-antigen associated with a cancer cell), live-attenuated pathogen (e.g.,

microbe), or heat-killed pathogen that has the potential to stimulate an immune response. Additionally, a "pathogen" refers to any organism capable of eliciting an immune response from a subject upon infection of the subject with pathogen. A nonlimiting example includes the pathogen *Toxoplasma gondii*. A given pathogen can be comprised of multiple antigens to which the subject's immune response may respond.

[0092] In some embodiments, the microbial immunogen comprises a live, genetically altered protozoan parasite that results in attenuation of the parasite. Thus, a vaccine composition refers to any composition that is administered to a subject with the goal of establishing an immune response and/or immune memory to a particular antigen or antigens. The vaccine compositions can comprise other substances designed to increase the ability of the vaccine to generate an immune response. For example, a live-attenuated vaccine can comprise the live-attenuated parasite plus an adjuvant.

[0093] Adjuvants can be any substance that enhances the immune response to the antigens in the vaccine composition. Non-limiting examples of adjuvants suitable for use in the present invention include Freund's adjuvant, incomplete Freund's adjuvant, saponin, surfactants such as hexadecylamine, octadecylamine, lysolecithin, demethyldioactadecyl ammonium bromide, N,N-dioctadecyl-N'-N-bis (2- hydroxyethylpropane diamine), methoxyhexa-decyl-glycerol, pluronic polyols, polyanions such as pyran, diethylaminoethyl (DEAE) dextran, dextran sulfate, polybrene, poly IC, polyacrylic acid, carbopol, ethylene maleic acid, aluminum hydroxide, and aluminum phosphate peptides, oil or hydrocarbon emulsions, and the like.

[0094] It is also contemplated that the vaccine compositions disclosed herein can be therapeutic or prophylactic. Thus, for example, the vaccine compositions disclosed herein can be used to prevent or reduce the likelihood of an infection such as, but not limited to, *Toxoplasma*. Alternatively, the vaccine compositions disclosed herein can be used therapeutically to treat one or more symptoms of an infection in an individual with a chronic infection.

[0095] The term "treating" as used within the context of the present invention is meant to include therapeutic treatment as well as prophylactic, or suppressive measures for the disease or disorder. Thus, for example, the term treatment includes the administration of an agent prior to or following the onset of a disease or disorder thereby preventing or removing all signs of the disease or disorder. As another example, administration of the agent after clinical manifestation of the disease to combat the symptoms of the disease comprises

"treatment" of the disease. This includes for instance, prevention of parasite propagation to uninfected cells of an organism.

[0096] Additionally, the present invention provides additional antigens in combination with the live, attenuated parasite in the vaccine compositions herein disclosed. The antigens provided in the mixture for vaccine compositions or immunization protocols can come from the same, different or unrelated targets. Thus, the antigens may be the same antigen, or the antigens may be heterologous antigens. For example, a vaccine composition can comprise a heterologous antigen, such as, e.g., a peptide of a protein of a target. In certain embodiments, the heterologous antigen is a peptide having a length of at least 5 amino acids (e.g., 5, 6, 7, 8, 9, 10, 11, 12, 15, 20, 25, 50, 75, 100 or more amino acids). The heterologous antigen can be, for example, a cancer antigen. In some embodiments, the parasite expresses the heterologous antigen.

[0097] The heterologous antigen can be an antigen found in various pathogens (e.g., viral, bacterial, fungal, protozoal, helminth) that may infect the subject. For example, the heterologous antigen may be derived from a different protozoan parasite (e.g., if the parasite is *T. gondii*, the antigen may be from a *Plasmodium* species). Alternatively, the heterologous antigen may be derived from a viral pathogen, such as a coronavirus (e.g., a SARS-CoV-2 virus). Examples of suitable antigens may include antigens derived from viral envelope or capsid proteins, proteins present at the cell surface (e.g., the cell surface of protozoan cells, bacterial cells, cancer cells), or secreted proteins. In some embodiments, the antigen is a coronavirus antigen, such as a spike protein or portion thereof (e.g., a SARS-CoV-2 spike protein or portion thereof, such as the S1 subunit or S2 subunit of the SARS-CoV-2 spike protein).

[0098] In some embodiments, the parasite that is genetically altered to have a mutation in the *BFD1* gene may be genetically altered to express the heterologous antigen. In specific embodiments, a coding sequence for the heterologous antigen can be inserted into the genome of the parasite at a location that is non-essential for parasite viability, e.g., into a non-essential gene. For example, the heterologous antigen can be inserted into a gene that is dispensable (see Example 2). In some embodiments, the coding sequence may be inserted into the BFD1 locus. The coding sequence can be operably linked to a promoter. The promoter can be, for example, the SAG1 or TUB1 promoter. In some embodiments, the heterologous antigen can be expressed as a fusion protein containing the heterologous antigen and all or a portion of an endogenous protein. In some embodiments, the endogenous protein, or portion thereof, is expressed at the cell surface (e.g., a SAG1 protein or a portion

thereof), and the resulting fusion protein is expressed at the cell surface (e.g., of the parasite). In other embodiments, the endogenous protein, or portion thereof, is a secreted protein, and the resulting fusion protein is secreted (e.g., by the parasite).

[0099] The antigen(s) can be coupled to a carrier protein. Non-limiting examples of suitable carrier proteins include albumin, ovalbumin, Pseudomonas exotoxin, tetanus toxin, ricin toxin, diphtheria toxin, cholera toxin, heat labile enterotoxin, keyhole lympet hemocyanin, epidermal growth factor, fibroblast growth factor, transferrin, platelet-derived growth factor, poly-L-lysine, poly-L-glutamine, mannose-6-phosphate, as well as various cell surface and membrane proteins, and the like.

[00100] Vaccine compositions can be formulated in aqueous solutions (carriers) such as water or alcohol, or in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer, including PBS. Vaccine compositions can also be prepared as solid form preparations which are intended to be converted, shortly before use, to liquid form preparations suitable for administration to a subject, for example, by constitution with a suitable vehicle, such as sterile water, saline solution, or alcohol, before use.

[00101] The vaccine compositions can also be formulated using sustained release vehicles (carriers) or depot preparations. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the vaccines may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt. Liposomes and emulsions can be used as delivery vehicles suitable for use with hydrophobic formulations. Sustained-release vehicles may, depending on their chemical nature, release the antigens over a range of several hours to several days to several weeks to several months.

[00102] The vaccine compositions may further include one or more antioxidants. Exemplary reducing agents include mercaptopropionyl glycine, N-acetylcysteine, β-mercaptoethylamine, glutathione, ascorbic acid and its salts, sulfite, or sodium metabisulfite, or similar species. In addition, antioxidants can also include natural antioxidants such as vitamin E, C, leutein, xanthine, beta carotene and minerals such as zinc and selenium.

[00103] Vaccine compositions may further incorporate additional substances to function as stabilizing agents, preservatives, buffers, wetting agents, emulsifying agents, dispersing agents, and monosaccharides, polysaccharides, and salts for varying the osmotic balance. The vaccine compositions can further comprise immunostimulatory molecules to enhance vaccine efficacy. Such molecules can potentiate the immune response, can induce

inflammation, and can be any lymphokine or cytokine. Nonlimiting examples of cytokines include interleukin (IL)-1, IL-2, IL-3, IL-4, IL-12, IL-13, granulocyte-macrophage colony stimulating factor (GMCSF), macrophage inflammatory factor, and the like. In some embodiments, the parasite may be genetically altered to express an immunostimulatory molecule such as any one or more of the aforementioned cytokines.

[00104] In some embodiments, the parasite utilized in a vaccine composition can express a therapeutic agent, such as e.g., a peptide or protein. In some embodiments, expression of a therapeutic agent or heterologous antigen in a tachyzoite can be under control of a promotor. In some embodiments, the promoter can be a constitutive promoter. In some embodiments, the promoter can be a promoter of a gene that encodes a protein that is expressed during the tachyzoite stage. For example, in some embodiments, the promoter can be the SAG1 or TUB1 promoter, in the case of *T. gondii*. Further, expression of a heterologous antigen by the parasite, e.g., a cancer antigen or antigen from a different pathogen can be under control of such promoters. A coding sequence of a heterologous polypeptide, e.g., a heterologous antigen or therapeutic agent, may be a codon optimized for expression by the parasite, e.g., *T. gondii*.

[00105] In the context of the present invention, the terms "peptide," "polypeptide" and "protein" are used herein. They refer to an amino acid chain, and include any post-translational modifications thereto (for example phosphorylation or glycosylation). Typically, the term protein is utilized when referring to a full-length product of a gene; and, the use of the term peptide and/or polypeptide is utilized when describing a fragment of a protein.

[00106] In some aspects, the present invention provides methods of treating cancer in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite, and wherein the parasite expresses a cancer antigen.

[00107] In another aspect, the present invention provides methods of inhibiting development or progression of cancer in a subject, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite, and wherein the parasite expresses a cancer antigen.

[00108] In another aspect, the present invention provides recombinant nucleic acid vectors comprising a nucleotide sequence encoding a BFD1 protein. In some embodiments, the nucleotide sequence encoding the amino acid sequence of a BFD1 protein comprises SEQ ID

NO:1. In some embodiments, a nucleotide sequence encoding the amino acid sequence of the BFD1 protein comprises a nucleic acid sequence having at least 70% sequence identity to the sequence of SEQ ID NO: 1. In some embodiments, a nucleotide sequence encoding the amino acid sequence of a BFD1 protein comprises a nucleic acid sequence having at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% sequence identity to the nucleic acid sequence of SEQ ID NO: 1.

[00109] In some embodiments, the amino acid sequence of a BFD1 protein comprises the sequence of SEQ ID NO:2.

[00110] In some embodiments, the recombinant nucleic acid vector is an expression vector. "Expression vector" refers to a vector comprising expression control sequences operatively linked to a nucleotide sequence to be expressed. An expression vector comprises sufficient cis-acting elements for expression; other elements for expression can be supplied by the host cell or in an *in vitro* expression system. Expression vectors include all those known in the art, such as cosmids, plasmids (e.g., naked or contained in liposomes) and viruses (e.g., lentiviruses, retroviruses, adenoviruses, and adeno-associated viruses) that incorporate the recombinant polynucleotide. The expression vectors can have cis-acting elements such as promoter sequences and non-promoter regulatory elements. As used herein, a “non-promoter regulatory element” refers to non-promoter sequence(s) of a nucleic acid molecule that are capable of increasing or decreasing the expression of specific genes within the recombinant vector. Such non-promoter regulatory elements include but are not limited to, e.g., enhancer elements, inducer elements, silencer elements, 5’ untranslated regions (UTRs), 3’UTRs, terminator elements, CAAT boxes, CCAAT boxes, Pribnow boxes, SECIS elements, polyadenylation signals, A-boxes, Z-boxes, C-boxes, E-boxes, G-boxes, and Cis-regulatory elements (CREs).

[00111] In some embodiments, the vector is a viral vector. Non-limiting examples of viral vectors that can be utilized by the present invention include DNA or RNA viral vectors including but not limited to retroviral vectors, herpes virus vectors, adenovirus vectors, lentivirus vectors, rabies virus vectors, lentiviral vectors, VSV vectors, vaccinia virus vectors, reovirus vectors, semliki forest virus, and sindbis virus vectors.

[00112] In some embodiments, the vector is a non-viral vector. Non-viral vectors can be plasmid DNA, liposome-DNA complexes (lipoplexes), and polymer-DNA complexes (polyplexes). Non-viral vectors can be plasmid RNA, liposome-RNA complexes (lipoplexes), and polymer-RNA complexes (polyplexes).

[00113] In some embodiments, the recombinant nucleic acid vector further comprises a selectable marker element. A "selectable marker element" is an element that confers a trait suitable for artificial selection. Examples of selectable marker elements useful in the present invention include, but are not limited to, Chloramphenicol acetyltransferase, Hypoxanthine-guanine phosphoribosyltransferase, bleomycin binding protein, beta-lactamase, neomycin resistance genes, mutant Fabl genes conferring triclosan resistance, URA3 elements, fluorescent gene products, affinity tags such as GST, His, CBP, MBP, and epitope tags such as Myc HA, FLAG. Selectable marker elements can be negative or positive selection markers.

[00114] In a further aspect, the present invention provides host cells comprising recombinant nucleic acid vectors comprising a nucleotide sequence encoding a BFD1 protein.

[00115] In some embodiments, the host cell is from an apicomplexan parasite. In some embodiments, the host cell is *Toxoplasma gondii*.

[00116] In another aspect, the present invention provides methods of inducing an immune response to an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the apicomplexan parasite is *Toxoplasma gondii*.

[00117] As used herein, the term, "subject," refers to an animal. Typically, the terms "subject" and "patient" may be used interchangeably herein in reference to a subject. As such, a "subject" includes an animal that is being treated for a disease, being immunized, or the recipient of a mixture of components as described here, such as a vaccine composition. The term "animal," includes, but is not limited to, mouse, rat, dog, guinea pig, cow, horse, sheep, chicken, cat, rabbit, pig, monkey, chimpanzee, and human. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal. In some embodiments, the subject is a bird.

[00118] The vaccine compositions described herein can be formulated for and administered by infusion or injection (intravenously, intraarterially, intramuscularly, intracutaneously, subcutaneously, intrathecally, intraduodenally, intraperitoneally, and the like). The vaccine compositions can also be administered intranasally, vaginally, rectally, orally, topically, buccally, transmucosally, or transdermally.

[00119] An effective antigen dosage to treat against apicomplexan parasite infection can be determined empirically, by means that are well established in the art. The effective dose

of the vaccine composition may depend on any number of variables, including without limitation, the size, height, weight, age, sex, overall health of the subject, the type of formulation, the mode or manner or administration, whether the parasite is active or attenuated, whether the patient is suffering from secondary infections, or other related conditions.

[00120] As an example, a suitable dose of genetically altered apicomplexan parasites, such as *Toxoplasma gondii*, per inoculation can be between about 1,000 to about 10 million tachyzoites, or between about 1,000 and 1 million tachyzoites, or between 5,000 and 50,000 tachyzoites, or between 10,000 and 50,000 tachyzoites. In some embodiments, a dose of at least about 1,000 genetically altered tachyzoites are administered to a subject, e.g., a human subject, per inoculation. In some embodiments, a dose of at between about 1,000 and 10,000 genetically altered tachyzoites are administered to a subject, e.g., a human subject. In some embodiments of the invention, a dose of at between about 10,000 and 100,000 genetically altered tachyzoites are administered to a subject, e.g., a human subject.

[00121] Vaccine regimens can also be based on the above-described factors. Vaccination can occur at any time during the lifetime of the subject, including development of the fetus through adulthood. Supplemental administrations, or boosters, may be required for full protection. To determine whether adequate immune protection has been achieved, seroconversion and antibody titers can be monitored in the patient following vaccination.

[00122] As used herein, "immune response," refers to an acquired and enhanced degree of protective immunity, preferably complete or sterile protection, against subsequent exposure to the parasites disclosed herein. In some embodiments, the protective immunity achieved after administration of the genetically altered protozoan parasites will not be complete but will reduce the severity of the infection symptoms after exposure to wild-type apicomplexan parasites (i.e., partial protection).

[00123] It is generally contemplated that inoculating a subject according to the methods described herein with genetically altered apicomplexan parasite will induce protective immunity against challenge with wildtype apicomplexan parasites of the same species. However, it is also contemplated by the present disclosure that immunization with one apicomplexan parasite can protect against challenge with another apicomplexan parasite of a different species, and eliciting cross-species protection in this manner is also within the scope of the invention.

[00124] As used herein, "immunization" or "vaccination" is intended for prophylactic or therapeutic immunization or vaccination. "Therapeutic vaccination" is meant for vaccination

of a patient with apicomplexan parasitic infection and/or for vaccination of a subject suffering from cancer, wherein the parasite expresses a cancer antigen.

[00125] Also, the disclosed methods can comprise the simultaneous or separate administration of multiple vaccine compositions. Thus, the present invention may further include the administration of a second, third, fourth, etc. antigen, wherein the second, third, fourth, etc. antigen is administered in a separate vaccine composition for administration at the same time as or 1, 2, 3, 4, 5, 6, 10, 14, 18, 21, 30, 60, 90, 120, 180, 360 days (or any number of days in between) after the first antigen.

[00126] In some embodiments, the parasite can be genetically altered to encode two or more antigens. In some embodiments, two or more antigens may be encoded as part of a single polypeptide. In some embodiments, two or more antigens may be encoded as separate polypeptides. In some embodiments, at least two of the antigens are from different proteins. In some embodiments, the two or more antigens include two or more different cancer antigens.

[00127] In some embodiments, the subject has an acute or chronic apicomplexan parasite infection.

[00128] In another aspect, the present invention provides methods of inhibiting or preventing a chronic apicomplexan parasite infection in a subject, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[00129] In another aspect, the present invention provides methods of treating a chronic infection by an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[00130] In another aspect, the present invention provides methods for inoculating a subject in need thereof with an apicomplexan parasite, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[00131] In another aspect, the present invention provides methods of administering an antigen to a subject in need thereof comprising administering to the subject a composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a *BFD1* gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite, and wherein the parasite comprises an antigen. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal.

[00132] In some embodiments, the parasite is genetically altered to comprise additional antigen(s).

[00133] In certain embodiments, the antigen is a cancer/tumor antigen. Non-limiting examples of possible cancer/tumor antigens that can be utilized in the present invention include alpha-fetoprotein (AFP), cancer antigen 125 (CA125), cancer antigen 15-3 (CA15-3), carbohydrate antigen 19-9 (CA19-9), carcinoembryonic antigen (CEA), HE4, chromogranin A (CgA), CD20, human chorionic gonadotropin (hCG or beta-hCG), lactate dehydrogenase, beta-2-microglobulin (B2M), calcitonin, neuron-specific enolase (NSE), programmed death ligand 1 (PD-L1), nuclear matrix protein 22, thyroglobulin, and prostate-specific antigen (PSA).

[00134] The present invention also provides for methods of generating genetically altered protozoan parasites comprising one or more mutations in the *BFD1* gene. As would be understood by those skilled in the art, routine molecular biology techniques, including but not limited to, gene knockouts, etc. can be utilized in generation of the genetically altered parasites described herein. For example, targeted mutation of a *BFD1* gene can be carried out in a lab setting by routine methods, such as homologous recombination techniques. In certain embodiments, the genetically altered protozoan parasites of the present invention are lab strains. Alternatively, the genetically altered protozoan parasites can be natural isolates.

[00135] In some embodiments, a targeted mutation of a *BFD1* gene may be generated using CRISPR methodology.

[00136] In some embodiments, the genetically altered protozoan parasite generated by the methods described herein is an apicomplexan parasite.

[00137] In certain embodiments, the genetically altered protozoan parasite generated by the methods described herein is *Toxoplasma gondii*.

[00138] In certain embodiments, the genetically altered protozoan parasite generated by the methods described herein is *Neospora caninum*.

[00139] In certain embodiments, the genetically altered protozoan parasite generated by the methods described herein is *Sarcocystis neurona*.

[00140] In certain embodiments, the genetically altered protozoan parasite generated by the methods described herein is *Hammondia hammondi*.

[00141] In certain embodiments, the genetically altered protozoan parasite generated by the methods described herein is *Hammondia pardalis*.

[00142] In a further aspect, the present invention provides for methods of identifying a candidate anti-parasitic compound comprising identifying a compound (small or large molecule compounds) that inhibits expression or activity of BFD1. The invention features methods for screening compounds that inhibit the expression or activity of the *BFD1* gene. In some embodiments, the screening assays involve contacting a proliferating tachyzoite with a test compound, and determining whether the test compound inhibits the tachyzoite from differentiation into a cyst-forming bradyzoite. In some embodiments, the screening assays involve contacting a proliferating tachyzoite with a test compound, and determining whether the test compound inhibits expression of the *BFD1* gene. In some embodiments, determining whether the test compound inhibits expression of the *BFD1* gene comprises measuring the level of mRNA encoding BFD1 and/or measuring the level of BFD1 protein in tachyzoites contacted with the test compound and comparing the level of BFD1 mRNA or protein with a suitable control, e.g., the level of BFD1 mRNA or protein found in tachyzoites not contacted with the test compound. In some embodiments, the screening assays involve contacting a proliferating tachyzoite with a test compound, and determining whether the test compound inhibits expression of one or more bradyzoite specific genes.

[00143] BFD1 may be necessary to maintain the differentiated state; therefore, interfering with its expression or activity (e.g., using compounds identified using screening methods described herein) may reactivate latent stages, which could subsequently be cleared with available anti-parasitic drugs (e.g., the subject could be treated with an anti-BRD1 compound and an anti-parasite drug).

[00144] In some embodiments, reporter assays can be utilized to screen for compounds that inhibit expression of a reporter gene operably linked to a BFD1-responsive promoter or promoter of any gene that is selectively expressed in the bradyzoite stage. The reporter gene may, for example, encode a reporter protein such as a fluorescent protein or an enzyme such as luciferase or beta-galactosidase, etc.

[00145] Compounds to be tested by the methods of the present invention include purified molecules, substantially purified molecules, molecules that are one or more components of a mixture of compounds, or a mixture of a compound with any other material. Test compounds can be organic or inorganic chemicals, or biomolecules, and all fragments, analogs,

homologs, conjugates, and derivatives thereof. Test compounds can be of natural or synthetic origin, and can be isolated or purified from their naturally occurring sources, or can be synthesized *de novo*. Test compounds can be defined in terms of structure or composition, or can be undefined. The compound can be an isolated product of unknown structure, a mixture of several known products, or an undefined composition comprising one or more compounds. Examples of undefined compositions include cell and tissue extracts, growth medium in which prokaryotic, eukaryotic, and archaebacterial cells have been cultured, fermentation broths, protein expression libraries, and the like. In some embodiments of this invention, the test compound could be a small peptide. Small peptides can be from 2 to about 10 amino acids in length, from about 10 to about 20 amino acids in length, from about 20 to about 30 amino acids in length, from about 30 to about 40 amino acids in length, from about 40 to about 50 amino acids in length, from about 50 to about 60 amino acids in length, from about 60 to about 70 amino acids in length, from about 70 to about 80 amino acids in length, from about 80 to about 90 amino acids in length, or from about 90 to about 100 amino acids in length. The peptides can contain naturally occurring amino acids, chemically modified amino acids and/or synthetic derivatives of amino acids. In another embodiment, the test compound is an antibody specific for the BFD1 protein. In some embodiments, the test compound is a small molecule. As used herein, a "small molecule" is a low molecular weight organic compound of about 50 daltons (D) up to about 1kD or about 2kD. In some embodiments, the small molecule can be < 900 daltons. Most drug active ingredients are small molecules.

[00146] In some embodiments, the screening assays involve contacting a proliferating tachyzoite with a test compound, and determining whether the test compound inhibits the tachyzoite from differentiation into a cyst-forming bradyzoite by molecular analysis of differentiation-related proteins or RNAs. If such differentiation-related proteins/RNAs show a reduction or elimination in expression, then the test compound is a candidate anti-parasitic compound. Analyses can be performed by methods understood in the art, such as, e.g., Western blot, Northern blot, and microarray analysis.

[00147] In another embodiment, candidate anti-parasitic compounds shown to affect differentiation of a tachyzoite can be further evaluated for their ability to reduce or inhibit progression of the tachyzoite into bradyzoites. In another embodiment, candidate anti-parasitic compounds shown to affect differentiation of a tachyzoite can be further evaluated for their ability to reduce or inhibit progression of acute infection to chronic infection in a subject.

[00148] SEQ ID NO: 1 - BFD1 genomic sequence:

ATGGAGGGAGCCAGTACTCAGCCAATGCCGCATTCGCAGCGGGAGCGTCAGGCT
CCTGTGCACGCCATCAGTGGCACGTGGGAAAACTGCGACCAGGCGGGTGAATAC
AAGGGGACTGCCTGTTGCGTCGTGGAGAGGCAAGTGTATTCTACGGAGACGCAC
GCAGTGGAAGGCTGCTGTTCCACGCGATGGAACGGAGTTAAGGAGGAAAAGGGC
GGTGGTGAAGTGTCCTCAAGAACTGCTTTAGCAGGGGTTGTCCATCTTTATACAA
ATTCAGACGGAAACGCATATACGCACATTTCTGACCAAGGAATGAGACAGGATG
AAGCTGTGGCGGGACAGGAGCAGCGTCACCTTAACTGTGAGGGGGAACACGAGA
AGAAAGGGAACGTGGCAGCAGGAGGTGGTACCCTCATGGTGACAAAATCGGAG
ATTGAACCGTGCGACGACTACTACTCCGTGCAGAGAGGACAGTCGTGCGCGGGA
GAGAGAGCGCCGCGTGACGGATGCTGTAGACTCCTTGATGGCAGCCGTGTCGAC
CCGGCAGAAGGAGGTTCGGAGGAAGATGAAAACTGTTATACACATGTGCTTCAG
AAGCATCTGGGTCAGATGCCGACGGCACCGTACCAAGGCGACGACCGCTTACAG
ACTGACTTACGACTGGGAGACTTCCAGTCCGGTCCTTCGGATGCCTTGTCCGGAC
ACCATGTAACACGGACACCAATTCTTTCGCCTCATGGCAGGTACTCTGAATACAT
CAGCGAACGTCTTGCGTGGCAGTATGCTGAGAGACCGGGGCATGGTATTGCTGC
GGAGAATACCGTTGTGATGCACAGCCATACTGGTGAGGCAACCGGATCATTGCG
AGCGGATGCGCCTTCGCAGTGGTCCACAGAGAGCCGCCTTCAGTTTCACGTCGGG
TCTCAATTTACCACCGAGAATCCGGAACTGTTCGCTGGCATTGTGGGATTAGACA
CAGAGCAATTCGACGCGAGAAATGCAGAAGCGCTGCACTGGAGACAACAGGGA
GAGGCTCGCTCAACGGAGACCGAATGCATCGCTGACATAACACCAGAGCGAGGC
AACCGGAGTACAATACGGCAATGGAATAAACCAGAAGCGGCGCGACAGGAATA
TCACCGTGCGTCAGgtgagagtggggctgcagtagtttgtctaacaggatgaacactaccagcacactagtcgcctgtaa
gactcgtgtgttttgtccggagtgtgattccaagatacaaggcgaccaagagctctttttgccagatgtgttgtgtccaacagCCTCT
GCTCGAGGGCCAGAGGAGACCAATATCATCACTCGGGGAGCACACTCCGACGAC
CAAAGTATAGACGCTGCTGCTCCAGGCTGCTGGGCTGCGCGCCATCTCACGTCTC
GACAGCATCCTAACCCGCGACCGAGGATGAAGGAGGAACATTGCGGTCGGGAAA
GGGAAGTCTTGTCTTCGGAGCAGCCCTCAGACTGTGGCGAAACACAGAAAACAC
CGgttagctgctagaactaaaaaatacatcaacatggctgatatccttcgagcgtacccacttagcctagcgtgaactgttgtgctcgc
gtggtaggcgaaaattggtgctgagcactgcccggcgagttgaaagaagtgacccgcaactcgcactcactaacactgcaccatggc
ggaatccgtttaccgattgtttccgtgttccacccacagcgtgatgacttcgctgtggagggagcctcatttcaattgctgcgtcgatgtgt
gtcctggttcctgtgcagGCGTCTCATTCACTTGTCTTGGATTCCAAGAGTCGAGCAGACGC
GAGTCAAACATCACCGTATACGCGTAACGACGAAATTCAAACGATTCGAGCATT
TGAGGAACAATGTCTTGAGGTGCAACTTCCTCAATCACGCTTAAGGCGTGGCATT

GACTATGAAGCTGGCATGTCACAGCTGGCTCCCACAACGCAAGAGTTGAAGGCC
AAAGTGTTGACAACGGGACTGTCCCCTTCGTGTCCAAACGCAGCCGTTCTTCGAG
CTAGTGTACCCTTGTCTATGGATACAACAGTTACGGTTCACCTTGCTGATGTCGA
GGATCCGAGTTCGAAGCGTCACACGCCTATCGAAAACTCTATTTCCAGCGGGCAT
TGTGATCCTCCGTGCTCGAGCCCGGGGGGAATGATGGAGCCCCGTGTATCCGTTT
CGGGTTTAGGCAGTCCTCACAACTGCGTGGAGGACAGAGAGATTGCGCGACACG
GAGACGGTATTGGAAGGGACATACTGGAAAGGCGTCCGCTACCATTCTCTCCTGC
TGCACTCCTGACGGGCTGCTCTAGCACGGCAGCCTCATCCGTTGTGTCTGGTGAC
GTGTTCACTTCGGTAGCAGCCTCTGCTCGCGCTAACGCGTGTAGATCGAGTACTG
ATACCCAGGAAGAGAGTGGCTGTGGCGACGCGCCGCGGTATGTACTACTGTTCC
CTGACGAACAAGACGAGAGATTAGGACATGCGCAAACCAGCATTGCCACGCCGT
CTGGATCAGGCCAATCAAAATGCATGTTTAAGGGCGACGGAAATGCAGAGGCCA
CATTCAGGGAGGAGGTAAAGAGTTCAGAACTTCGGACGCCCTCGAGGATACAAA
CACGGCGTTTGCTTCCCGGGGTACAGCTTCAAGGCATGGACTGTGACGGATGCGG
TGCCTCCGATCCCCAGAAAGGGAGACCTGCGCCAGAGACAGGTTTTTTGCCTGAG
ATTTGCCATTTCTCACCTCAGCATCCATGGCAACCGGGATCGGAAGTCAACCAGG
GGTACCGGGCGGGCGCTGACTACGGGACCTCTCGGGTACAGCAGAGTCTTGAAA
ATTGTTCATGGGAAGAGTCAGTGGATGAGGGTGAACAACCACGCACCACATCGT
CCTCCTCATACGGACAAGACACGCAGAAACGGGATTCCTTCCTCCAGTCAATCGA
TAACAGAATAGCAAGCGAACCGGTTGACCCCACAGGTACATGCGGATATAACTC
AACGCCTGAAACCTTTCGTAGTGGATGCATTATGCGTGATCAGATGCTCGGTGTC
AAGCCACATCTGTCTCATGTATCTTTAACGCGCGAGGGCAAGCATCGGGGACAG
CAAACGCACCTACGGAAAACTGGAGTGCCAACATGGAGCGCTGAGGAGGATGCG
AGTCTCGCAGAGTTGGTGTCTAGAAAGGGCTTCAAGTGGGCACTGATCTCTTCAC
AACTTACTGGCGCCTTTGGGATCCCACGGACTGgtaagcgtggatttgagtataccaagagagggtcat
ctgagcaagccaaacaccagatggcacgacggtacgagatttcccaagtgtgaaatcactgcttgtttttggtgattctcgactgtgcgt
gtgttctgattgtagGTAAACAATGTCGGGAACGATGGTTTAATCATGTGAATCCCGAGGT
TAAGAAAGGCGACTGGTCAGCCGAGGAGGATGCCATGATACTGATGCTTCAGAA
TGAACTGgtgagtggttgtatctcgggtggtcgatttgtttcacaatggcgttagacctctcaaaagtcagtgattgatgttgcggtg
ttgtccgaaaaagaaggttcttctctctgtgtttcgactctgcacgtgagtacccgtacaaggcaggctcacccagtatttttcctgctccat
gggctagtgcagacgagtgaattttcgtctgtctgcattcttatggaggtcacgaagatccacactgtgcaagcacgcgacgcgtgtga
gacttacgcagaaatgaggccgttgtttgccatgtttttgttcttgatgcagGGAAATCGCTGGGCGACTATCGCA
AAGAAGTTGCGGGGCCGCACAGAGAATGCCGTGAAGAACCGCTTTATTTCTCTC
AGCAATGCCCGCCTAGGTTATGGTCGTCCAAAGCGCGATGGCTCGAGCGCGGAC

TGCTTCAGTAACAGACGTACCGGTAGTGGGAAGTCGTCAGGCATAACAGGGATG
CCGAATCTATGTCAGTCTGTATGTTCCGCTGGCACAACGAAGAAGGACTCGTCTG
AATCGGGTAACCATTTCGTTATGTCCGTTGCCACGAAAGTTTTTGAATTTTCGGAC
GTGGCAGTGGATTCTGGGGTGAGCCGACCGAGACAGTGCACTGGCACGTCACCC
AGTTGTGGGCATCCGTCGGCAGGTGAGGGGGATCCATCTCATTTAAAGAACACT
GACGTTGTCGGCCGGGAACAACCCATACAGCGGAACAACAACGAATCAGGCAAA
GCTGCAGAGCAAACTGCGTTTTCAGGCGTGAAGACAGGGACACTATCGGTTTCA
CAGGATGCTGTCCCTGTCGGTCGTCTAGTCGTAGCAAGTGTAGGGCCTCAGCACA
TTCGTAGGAGCTTCCCGACTGACGAAACATTACCAAAATTCGCTGCTAAAGAGCA
TAACAACCAACAGCTGAACGACGAGAGAGAACATCTCGAACAAAGCAACTCCAC
GTCAGAGGGGAGCTTTCTGGCATCAGCTCATGAACATGCAGACATCGCGCGAAG
TGATCCTGATGAGGACACGCTGGAGCCCCATCAAAAACGACGACGCAAGCGATG
TGCAATAGCATACCAGGGCGAGGAACGTGGTGACAGCAACGGTCTTGATTCAAT
TGCGGATCGTGCAGAGCAAGCAGGGAACTTCCAGGCAATGAGGAAAGCTAACAC
GGACAACGGAAAAGTCGATTACTTGGAGCCTCATCGCTACGAAAAGTTGTCACC
ATGCGAGCAGGTGATCCAGCCGTCCTTGCGGCCAGCATGTGACCATCGTGGTGCA
CCCCAAAACTCAGTCGAGAGTGGTGAGCAGTCTCCCGATGCACAGCGACAGTTG
TGTAATCAAGGCTGTCGGACTTCCAACCGAACAGTACACAGCTCTGTTTACAGCA
ACGAAGTGGAGTCAAATGAGCTGCGCGGTGTGTTTCGGCTGGCGGAACAATCGC
TGCCTTCCCAGAGCGGGGATCCTGCGTGGTCGACAGCAGGATTTCAGTTGTCTAT
CCTGCCTCAGAAGGTGGAAGTACACTCCAGGAACAAATGCGATGGTCAAAATGT
GATGTATCGGTGTTCGCCCGGTTCTTTGCCGACTACCCATCAGCAGACAGTGTTT
CATTACGATAGAGATTCAAGCAGGTTCCCTTGCGCGGCCAAGCCTGCCGCAGCA
AGTGGAGCCCAGGGGACGATCGAGGAGAACGACGGATTAGTTAAAGAAGGACC
GAGCATGATTGTTACCGGGAGCAGTGTGGAAGTAGTACACTGCTGTTCAGTGTCC
CTTCGCAGGCGTGACCGGTCGTTACCAAGCGCTCAGTTGTGGACTTCACAAGAGA
CTGAGTCTGACACGAATCCATCACCGAACCAGCAGCATGAGAGTTGCCACCAAT
ACTGTAAGCGCCACGCTGCTTGGTGTGGGAAGACGGACCAGTTTTCCAAGCTGAC
TTCTTCGCATCAAGAGAACAGTAGTGGTAAAGATGCATGTCTTGTGTCTGTGTCT
CCAACAGTAACGCTAGATGATCTTCAGAAACAATCCCGAGGCACAGTTCTGTCA
GCGAAAGAAGAAATCGGCAAGCCGGAGACGTGGTCTCACGTTGTTGACAACACT
TACTCGAAGACAGACCACCAGCGTGCGTCGCTGTGTGCAGAGAATTCGTCAGGA
TGCGCCGAAGGAAGCACGGAACTCGTCAGATTCTCGGGAGGCTCTGTAAAATCA
GGGAGCAGTATGAGCGTCGACTGCGGAAATGGAAACCCAGATGACTGCCAGGAT

TGCAAAGCTGAGGAGATATGGCGAGGCGAGCAACGATATGCCGAAAGGGGCCA
CTCGGTTGAATCCAGGGGTGCTGGTTCAGTAGGTAGGAGCACGGATCTAACTATC
ACCGACTCTGGATCGATGCCGCTGTGTGCCAGCCCAATCGGTAGACCTCCTGCAG
ACAACGACACGCTGTTTTTAAGTGATGCCAGGTGCAACGTAGTGGCACAACTTAA
CCATCAAGATAACAGCAGGATTTCCCGATTAGCAAGCTGTGAAGAGGAATTCTT
GGCATGCGGAGGTGAGCGTCTGATTAATGCTTCCGGCGGTTTTAAGCCGGATGGT
GGATGCTTGTATCGCATGCAGCAAGCTGGTGCATGCAACACAAAACTTCACCGA
CCAGTGCATAGTTGCTCCACGATTGACTCGGAACAGCTTGAAGATCTTCCTTCTG
TGGAGAAGGCAGTTGGCGACCGAAGCTTCTCCAGCAAACGAAAAGGAGACATCC
CACCTTTCGCTGAATGGAAAAAAAATGACGAACTACGTGAGTTATACCGCGGCG
TGTCAGAAGCAGTTTCCCACGGTCAACCTGGCGACTGGAACGGCAGCTGGCCCG
GTATCTCAGGGCGCGCTCATCAAACGTCTTCGTGTTTCCCTGATAGGGTTAATGC
AAGCGACCGTAGAGAGCTGAATAGTTGGCGGCTCCACGTCTCCGCCGCCGCCGA
GCTAGGGTCAAGTCACATCTGGAACTCGCAATCATACGCTTCGGCCTCTGTGAGT
CGGGATAAACAACGGGAGCCACCTAAAAATGGGCTCACTGGCTGTGACGTGCCA
GAGTATCTGGGGACCAGTCAAAGTGCAGGACTACCTGCAGCAAACGCGCACGAA
CGCGGCAACTTCTATGGACATGATAGATGCCGACCACGTGAAGGAGAGCGGGTC
CGTTGGGTGGGCCTACAGCGAAATCGAAAACCCGAGGCATCGGTATCCTCTGGA
GCGAGCAACAGTGCGACAACAGCGAGACCGAAGGACAGCACGGAACCTGACGA
AGGCAACAGTGAAGGAGTGAGCACGAGGCGAAAAGACTCTGGTTCCACTGCAGC
GACTATCAGTCGGGCTGTTTCCCTGGGTATGGTGACACCGAGTGCTGCTTGTGAG
AATTCTTCTTCACTAACAGATACGTCTCCTCCTCTGAGCCATCGGCCCTCCTTCAG
CTTTACTCATTGCTGCGAGGAGACTTTGAGCAGGTGTAACTCCTCTAATTATTTGT
GTCCTCCAGCAACGTGTCATACGTCGGATGATGGTAGGTCTTTAGGGCCGTCGCG
AGAGGCACAAGCTCTCCGGTCATTGAGCCTAGCTTCCGGCTATGGTTACCCGGGG
ATACCTGCCGAGGCAACGTCCTTCTGGCAGGGCAGCTCACTGGAGCACTCAATA
ATGGAACCGCAGATGGTGCCATCCGATGATGAACTTCGATTGTGGGTCCATCCTC
GTGACGCAGCGAATTGGTCGCAGAGCACGTTGAAACCCGTTGCTGTCGTAAGCG
GGACCGACGCTGGTGACGATCAACATAAGACGCCAGAGAACTTGACTCCCGAAA
GCGGGCAGGCCCATCGCCGGGACGGCCACGACATGCAGCGTGTGCAACGGTGCG
ATGACGAGGGGGAGTGCCCCCCTACAACAGTCGAACTGACTTTTCCGCATTCACA
CTCTAGTGATGAAATGCAGGACTTGCCGTCGAAGGTACAAGGAAATTTTCTCTTG
CGAAGAGAGCTCTCGGACTCACTGCAACATGAGACTGCCGAAAGCGTTGCTGGT
TACGGGTGGATGCGTATACGGAATGCTGGAGATATTCCAAACTCGAAGGTACCTT

GTGCTTGGGAACAGTGCATGCCAGCATCGGAGAGAGAACGGGGAGTGAACGACC
ACATGAGTAGTGAAGCCTCACGAATGTCGAAGGCAGCAAGTTCAAGCTTCGTTC
CTAGCTCATGCACAGATGCGCCGGTGGTCAGAGTAGGTGAGGACACTACCAAGT
CGGTCTGTGAGGAGCAACAGCTGTGCGAAGGAGGTAATCGTGGTTCCTTGTCCCC
GGAAGCCACAGGCTTTGAGAGCTTGGGCCCACCGCTCCAGTTGTTGCTTGTTGAC
GGATACACTCCATTCGAGCCCGTAGTGGAAAAGGTCTCTCAAACAATGGAGCAG
ACTTTATTTCCTGTGCCAGGTCAGGAGACTGATACAAGAGATGAGGATGGACGA
TACAACTGgtaagaaaaacaatacactcaggcaaggccagggtgcgtgaaagagaagtgtatgctgtcccagtgtaggtagc
agtgctatttgtaggccttcgaaagaaagaaagtggaacagaaggctacttcgatgcggtagttccacagggtggcagggcagcccc
gccttgaggtgatctggcacgaagtcagtgttcaaatgctcaggcaccgattcaagccatatatccatatgtggccaacatagagaaaa
tggtgaatcaccacatttatttaagcacacataaaacagtcttcagtaactatcagtgtgcaagaggcggctataggcgtccggagccaa
gtgtaatggattcttgtccctcttcccacacggcagccgtgagatggtctggtgtacaatagccacggacttaaagctgtttccgggaga
cacagctggagactacacaacgttgtcaccgaagcctggtacgcgatgagtacacgcccgacctgtttcttgcctttctcctgtgttctca
gCGAGTGCTTACAAAACCGCCAGCCACCTTTGCATTCGGGGGGCTTGATG

[00149] SEQ ID NO: 2 - BFD1 amino acid sequence:

MEGASTQPMPHSQRERQAPVHAISGTWENCDQAGEYKGTACCVVERQVYSTETHA
VEGCCSTRWNGVKEEKGGGEVSSRTALAGVVHLYTNSDGNAYTHISDQGMRQDEA
VAGQEQRHLNCEGEHEKKGNVAAGGGTLMVTKSEIEPCDDYYSVQRGQSCAGERA
PRDGCCRLLDGSRVDPAEGGSEEDENCYTHVLQKHLGQMPTAPYQGDDRLQTDLRL
GDFQSGPSDALSGHHVTRTPILSPHGRYSEYISERLAWQYAERPGHGIAAENTVVMH
SHTGEATGSLRADAPSQWSTESRLQFHVGSQFTTENPELFAGIVGLDTEQFDARNAE
ALHWRQQGEARSTETECIADITPERGNRSTIRQWNKPEAARQEYHRASASARGPEET
NIITRGAHSDDQSIDAAAPGCWAARHLTSRQHPNPRPRMKEEHCGREREVLSSEQPS
DCGETQKTPASHSLVLDSKSRADASQTSPYTRNDEIQTIRAFEEQCLEVQLPQSRLRR
GIDYEAGMSQLAPTTQELKAKVLTTGLSPSCPNAAVLRASVPLSMDTTVTVHLADVE
DPSSKRHTPIENSISSGHCDPPCSSPGGMMEPRVSVSGLGSPHNCVEDREIARHGDGIG
RDILERRPLPFSPAALLTGCSSTAASSVVSGDVFTSVAASARANACRSSTDTQEESGC
GDAPRYVLLFPDEQDERLGHAQTSIATPSGSGQSKCMFKGDGNAEATFREEVKSSEL
RTPSRIQTRRLLPGVQLQGMDCDGCGASDPQKGRPAPETGFLPEICHFSPQHPWQPGS
EVNQGYRAGADYGTSRVQQSLENCSWEESVDEGEQPRTTSSSSYGQDTQKRDSFLQ
SIDNRIASEPVDPTGTCGYNSTPETFRSGCIMRDQMLGVKPHLSHVSLTREGKHRGQQ
THLRKTGVPTWSAEEDASLAELVSRKGFKWALISSQLTGAFGIPRTGKQCRERWFNH
VNPEVKKGDWSAEEDAMILMLQNELGNRWATIAKKLRGRTENAVKNRFISLSNARL

GYGRPKRDGSSADCFSNRRTGSGKSSGITGMPNLCQSVCSAGTTKKDSSESGNHFVM
SVATKVFEFSDVAVDSGVSRPRQCTGTSPSCGHPSAGEGDPSHLKNTDVVGREQPIQ
RNNNESGKAAEQTAFSGVKTGTLSVSQDAVPVGRLVVASVGPQHIRRSFPTDETLPK
FAAKEHNNQQLNDEREHLEQSNSTSEGSFLASAHEHADIARSDPDEDTLEPHQKRRR
KRCAIAYQGEERGDSNGLDSIADRAEQAGNFQAMRKANTDNGKVDYLEPHRYEKL
SPCEQVIQPSLRPACDHRGAPQNSVESGEQSPDAQRQLCNQGCRTSNRTVHSSVYSN
EVESNELRGVFRLAEQSLPSQSGDPAWSTAGFQLSILPQKVEVHSRNKCDGQNVMY
RCSPGSLPTTHQQTVFHYDRDSSRFPCAAKPAAASGAQGTIEENDGLVKEGPSMIVT
GSSVEVVHCCSVSLRRRDRSLPSAQLWTSQETESDTNPSPNQQHESCHQYCKRHAA
WCGKTDQFSKLTSSHQENSSGKDACLVSVSPTVTLDDLQKQSRGTVLSAKEEIGKPE
TWSHVVDNTYSKTDHQRASLCAENSSGCAEGSTELVRFSGGSVKSGSSMSVDCGNG
NPDDCQDCKAEEIWRGEQRYAERGHSVESRGAGSVGRSTDLTITDSGSMPLCASPIG
RPPADNDTLFLSDARCNVVAQLNHQDNSRISRLASCEEEFLACGGERLINASGGFKPD
GGCLYRMQQAGACNTKLHRPVHSCSTIDSEQLEDLPSVEKAVGDRSFSSKRKGDIPP
FAEWKKNDELRELYRGVSEAVSHGQPGDWNGSWPGISGRAHQTSSCFPDRVNASDR
RELNSWRLHVSAAAELGSSHIWNSQSYASASVSRDKQREPPKNGLTGCDVPEYLGTS
QSAGLPAANAHERGNFYGHDRCRPREGERVRWVGLQRNRKPEASVSSGASNSATTA
RPKDSTEPDEGNSEGVSTRRKDSGSTAATISRAVSLGMVTPSAACENSSSLTDTSPPLS
HRPSFSFTHCCEETLSRCNSSNYLCPPATCHTSDDGRSLGPSREAQALRSLSLASGYG
YPGIPAEATSFWQGSSLEHSIMEPQMVPSDDELRLWVHPRDAANWSQSTLKPVAVV
SGTDAGDDQHKTPENLTPESGQAHRRDGHDMQRVQRCDDEGECPPTTVELTFPHSH
SSDEMQDLPSKVQGNFLLRRELSDSLQHETAESVAGYGWMRIRNAGDIPNSKVPCA
WEQCMPASERERGVNDHMSSEASRMSKAASSSFVPSSCTDAPVVRVGEDTTKSVCE
EQQLCEGGNRGSLSPEATGFESLGPPLQLLLVDGYTPFEPVVEKVSQTMEQTLFPVPG
QETDTRDEDGRYNCECLQNRQPPLHSGGLM

[00150] The following examples are provided to describe the invention in more detail. They are intended to illustrate, not to limit, the invention.

Materials and Methods

[00151] Strains and cell culture

[00152] Human foreskin fibroblasts were maintained in DMEM (Gibco) supplemented with 3% inactivated fetal serum (IFS) and 10 μg/mL gentamicin, referred to as standard

media. If HFFs were to be used in bradyzoite experiments, host cells were maintained exclusively in DMEM supplemented with 10% IFS and 10 µg/mL gentamicin prior to infection. Alkaline stress media consists of RPMI 1640 (Sigma), supplemented with 1% IFS and 10 µg/mL gentamicin, and buffered with 50 mM HEPES adjusted to pH 8.1 with 10N NaOH.

[00153] Plasmids and primers

[00154] Oligos were ordered from IDT. All cloning was performed with Q5 2X master mix (NEB) unless otherwise noted. Primers and plasmids used or generated are found in Tables 1 and 2.

[00155] **Table 1. Plasmids utilized in present disclosure.**

Vector	Description	SEQ ID NO	Sequence
pU6_Library_DHFR	generic gRNA expression + pyrimethamine resistance. Library backbone.	3	ccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactat ggatgaacgaaatagacagatcgctgagataggtgcctcactgattaagcattggtaa ctgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaag gatctaggtgaagatccttttgataatctcatgaccaaaatcccttaacgtgagttttcgtt ccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctg cgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgcc ggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagatac caaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcac cgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagt cgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgg gctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaa ctgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaa ggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgaggga gcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgactt gagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagc aacgcggccttttacggttcctggccttttgctggccttttgctcaCATGgGATG AGACAAAGTGCGCGAGTTGAAATCGTCGTGGGGAC GATTTCACCGCGGCCACATGTTGGAGACACTGAGGG CACACGGGAAACGCGAAAGATTTCAAATTAACGTAC CCAAACGCGAAAGCTTGCGCAGCATACACTCGAAGC GAACATCCCGAACCATCGAGAGGCGGAGAGCGATA AGTCTTTCACGCTGCGAAGTGTTGCGACGGCTGCGC CGCTGCACTGTGAATTGGGCGCCAATATTGCATCCT AGGCCTGACGCGCCTCCTGCAGAACGCGAGACACTG GGATATGTAGAGCCAAGGGGGAAACCTTCGAACTCT CGAATGTCTTCTCTGACAAGAATCATATTTCCATCAG TTCTGTCAGATTTTCAAATGGCGACCTGCAGAGGCC TGCTTCCTCCCTGTGCGCTCTTCGAAGGGGCTTTCTG TCGCGCAGGGTCACCTCGTCCCCGAAGGGGGTGTTT GCCTTCTGGTAAATGGGGATGTCAAGTTAGAGACCG GTCTCAGTTTaAGAGCTAtgctgGAAAcagcaTAGCAAGT TtAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GCACCGAGTCGGTGCTTTTTTTTTCTTTTTCtctagaggtac CATGCA
pBAG1-mNG	pBAG1-mNG reporter	4	gagaaggggcgggccagagcgttcggaaaattatctgcaaagcccaggtcccgtat gatattcaaaaaagatgatggtgagcaagggcgaggaggataacatggcctctctccc agcgacacatgagttacacatctttggctccatcaacggtgtggactttgacatggtgg gtcagggcaccggcaatccaaatgatggttatgaggagttaaacctgaagtccaccaa

			gggtgacctccagttctccccctggattctggtccctcatatcgggtatggcttccatca gtacctgccctaccctgacgggatgtcgcctttccaggccgccatggtagatggctcc ggctaccaagtccatcgcacaatgcagtttgaagatggtgcctcccttactgttaactac cgctacacctacgagggaagccacatcaaaggagaggcccaggtgaaggggactg gtttccctgctgacggtcctgtgatgaccaactcgctgaccgctgcggactggtgcag gtcgaagaagacttaccccaacgacaaaaccatcatcagtacctttaagtggagttaca ccactggaaatggcaagcgctaccggagcactgcgcggaccacctacacctttgcca agccaatggcggctaactatctgaagaaccagccgatgtacgtgttccgtaagacgga gctcaagcactccaagaccgagctcaacttcaaggagtggcaaaaggcctttaccgat gtgatgggcatggacgagctgtacaagtaagcgcgcccagcccacagaagctgccc gtctctcgttttcctctcttttcggagggatcagggagagtgcctcgggtcggagagag ctgacgaggggtgccagagacccctgtgtcctttatcgaagaaaagggatgactctt catgtggcatttcacacagtctcacctcgccttgttttcttttgtcaatcagaacgaaagc gagttgcgggtgacgcagatgtgcgtgtatccactcggaatgcgttatcgttctgtatgc cgctagagtgctggactgttgctgtctgcccacgacagcagacaactttccttctatgca cttgcaggatggtgcagcgcaaacgacggagagaaaggagcaccctctcagtttccc tacgatgtgctgtcagtttcgactcttcaccgcgaacgattggcgatacgtctctgttgac ttgttaggctccgaccacgaagctcccttaactagataagccgcgacacctaagtgtac accatttgcagatcgataatctgcgaccgctgaatccgtccagatcagtaaaaccgcac cacctaagtgtaaaccttgtttaggtcgataaaatgctaccaaccccccacccacaatcg agccttgagcgtttctgcgcacgcgttggcctacgtgacttgctgatgcctgcctctggc cattcatgccagtcagtgcgcataaaaatgtggacacagtcggttgacaagtgttctgg caggctacagtgacaccgcggtggaggggatccactagttctactcgagggtcgac ggtatcgat
pSAG1-mNG	pSAG1-mNG reporter (used as template only)	5	ctcccagcgacacatgagttacacatctttggctccatcaacggtgtggactttgacatg gtgggtcagggcaccggcaatccaaatgatggttatgaggagttaaacctgaagtcca ccaagggtgacctccagttctccccctggattctggtccctcatatcgggtatggcttcc atcagtacctgccctaccctgacgggatgtcgcctttccaggccgccatggtagatgg ctccggctaccaagtccatcgcacaatgcagtttgaagatggtgcctcccttactgttaa ctaccgctacacctacgagggaagccacatcaaaggagaggcccaggtgaagggg actggtttccctgctgacggtcctgtgatgaccaactcgctgaccgctgcggactggtg caggtcgaagaagacttaccccaacgacaaaaccatcatcagtacctttaagtggagtt acaccactggaaatggcaagcgctaccggagcactgcgcggaccacctacacctttg ccaagccaatggcggctaactatctgaagaaccagccgatgtacgtgttccgtaagac ggagctcaagcactccaagaccgagctcaacttcaaggagtggcaaaaggcctttac cgatgtgatggGCATGGACGAGCTGTACAAGTAAgcgcgccca gcccacagaagctgcccgtctctcgttttcctctctttcggagggatcagggagagtg cctcgggtcggagagagctgacgagggggtgccagagaccccctgtgtcctttatcga agaaaagggatgactcttcatgtggcatttcacacagtctcacctcgccttgttttcttttg tcaatcagaacgaaagcgagttgcgggtgacgcagatgtgcgtgtatccactcggaat gcgttatcgttctgtatgccgctagagtgctggactgttgctgtctgcccacgacagcag acaactttccttctatgcacttgcaggatggtgcagcgcaaacgacggagagaaagga gcaccctctcagtttccctacgatgtgctgtcagtttcgactcttcaccgcgaacgattgg cgatacgtctctgttgacttgttaggctccgaccacgaagctcccttaactagataagcc gcgacacctaagtgtacaccatttgcagatcgataatctgcgaccgctgaatccgtcca gatcagtaaaaccgcaccacctaagtgtaaaccttgtttaggtcgataaaatgctaccaa cccccacccacaatcgagccttgagcgtttctgcgcacgcgttggcctacgtgacttgc tgatgcctgcctctggccattcatgccagtcagtgcgcataaaaatgtggacacagtcg gttgacaagtgttctggcaggctacagtgacaccgcggtggaggggatccactagtt ctactcgagggtcgacggtatcgat
pTUB1-BFD1_wt-Ty	Overexpression of BFD1	6	gtccaacccggtaagacacgacttatcgccactggcagcagccactggtaacaggatt agcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacg gctacactagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaa aaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttg tttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttt ctacggggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgaga ttatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaa agtatatatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatct cagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacg atacgggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgc

			tcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcag aagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctaga gtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtgg tgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgag ttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgt cagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctctt actgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattct gagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataatac cgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaa aactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcaccca actgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggc aaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactctt cctttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaa tgtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccac
pTUB1-BFD1_ΔDBD-Ty	Overexpression of BFD1ΔDBD	7	gcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctaca ctagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaaga gttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggtttttttgtttgca agcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacg gggtctgacgctcagtggaacgaaaactcacgttaagggattttggtcatgagattatc aaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtat atatgagtaaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagc gatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatac gggagggcttaccatctggccccagtgctgcaatgataccgcgagacccacgctcac cggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagt ggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaa gtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgtc acgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttac atgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcag aagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttact gtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgag aatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgc gccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaac tctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaact gatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaa atgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttcctt tttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgta tttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccac
p312330-Ty	Endogenous Ty tagging vector for 312330. Co-transfect with pCas9-CAT	8	gGCATGGACGAGCTGTACAAGTGAtcaccgttgtgctcacttctc aaatcgacaaaggaaacacacttcgtgcagcatgtgccccattataaagaaactgagtt gttccgctgtggcttgcaggtgtcacatccacaaaaaccggccgactctaaataggagt gtttcgcagcaagcagcgaaagtttatgactgggtccgaatctctgaacggatgtgtgg cggacctggctgatgttgatcgccgtcgacacacgcgccacatgggtcaatacacaa gacagctatcagttgttttagtcgaaccggttaacacaattcttgccccccccgaTGAct agaggtacCATGCATctagcatgtcattcgattttcacccccccgcgtagttcctgt gtgtcattcgttgtcgagacaactctgtcccgccccggtgctgttccatatgcgtgacttt cccgcaatttttttcagacttttcaggaaagacaggctccggaacgatctcgtccatgactg gtaaatccacgacaccgcaatggcccccagcacctctatctctcgtgccaggggacta acgttgtatgcgtctgcgtcttgtctttttgcattcgctttccaaaaaagagagccatccgtt cccccgcacattcaacgccgcgagtgcggtttttgtctttttgagtggtaggacgctttt catgcgcgaactacgtggacattaagttccattctcttttttcgacagcacgaaaccttgca ttcaaacccgcccgcggaagatccgatcttgctgctgttcgcagtcccagtagcgtcct gtcggccgcgccgtctctGttggtgggcagccgctacacctgttatCtgactgccgtg cgcgaaaatgacgccatttttgggaaaatcggggaacttcattctttaaaagtatgcgga ggtttccttttttcttctgttcgtttctttttctcgggtttgataaccgtgttcgatgtaagcacttt ccgtctctcctccgtgctttgttcgacatcgagaGcaggtgtgcagatccttcgcttgtc gatccggagacgcgtgtctcgtagaacctttcattttaccacacggcagtgcggagca ctgctctgagtgcagcagggacgggtgaagtttcgctttagtagtgcgtttctgctctac ggggcgttgtcgtgtctgggaag
p208740-Ty	Endogenous Ty	9	gagttgttccgctgtggcttgcaggtgtcacatccacaaaaaccggccgactctaaata ggagtgtttcgcagcaagcagcgaaagtttatgactgggtccgaatctctgaacggat

	tagging vector for 208740. Co-transfect with pCas9-CAT		gtgtggcggacctggctgatgttgatcgccgtcgacacacgcgccacatgggtcaata cacaagacagctatcagttgttttagtcgaaccggttaacacaattcttgcccccccgaT GActagaggtacCATGCATctagcatgtcattcgattttcacccccccgcgtagtt cctgtgtgtcattcgttgtcgagacaactctgtcccgccccggtgctgttccatatgcgtg actttcccgcaatttttttcagactttcaggaaagacaggctccggaacgatctcgtccat gactggtaaatccacgacaccgcaatggcccccagcacctctatctctcgtgccaggg gactaacgttgtatgcgtctgcgtcttgtcttttgcattcgctttccaaaaaagagagcca tccgttcccccgcacattcaacgccgcgagtgcggtttttgtctttttgagtggtaggac gcttttcatgcgcgaactacgtggacattaagttccattctcttttcgacagcacgaaac cttgcattcaaacccgcccgcggaagatccgatcttgctgctgttcgcagtcccagtag cgtcctgtcggccgcgccgtctctGttggtgggcagccgctacacctgttatCtgact gccgtgcgcgaaaatgacgccatttttgggaaaatcggggaacttcattctttaaaagta tgcggaggtttccttttcttctgttcgtttctttttctcgggtttgataaccgtgttcgatgtaa gcactttccgtctctcctccgtgctttgttcgacatcgagaGcaggtgtgcagatccttc gcttgtcgatccggagacgcgtgtctcgtagaacctttcatttttaccacacggcagtgc ggagcactgctctgagtgcagcagggacgggtgaagtttcgctttagtagtgcgtttct gctctacggggcgttgtcgtgtctgggaag
pU6-BFD1-DHFR	anti-BFD1 gRNA + pyrimethamine resistance, used to generated BFD1 frameshift	10	cggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagc ttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttga gcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaac gcggcctttttacggttcctggccttttgctggccttttgctcaCATGgGATGAG ACAAAGTGCGCGAGTTGAAATCGTCGTGGGGACGAT TTCACCGCGGCCACATGTTGGAGACACTGAGGGCAC ACGGGAAACGCGAAAGATTTCAAATTAACGTACCCA AACGCGAAAGCTTGCGCAGCATACACTCGAAGCGA ACATCCCGAACCATCGAGAGGCGGAGAGCGATAAG TCTTTCACGCTGCGAAGTGTTGCGACGGCTGCGCCG CTGCACTGTGAATTGGGCGCCAATATTGCATCCTAG GCCTGACGCGCCTCCTGCAGAACGCGAGACACTGGG ATATGTAGAGCCAAGGGGGAAACCTTCGAACTCTCG AATGTCTTCTCTGACAAGAATCATATTTCCATCAGTT CTGTCAGATTTTCAAATGGCGACCTGCAGAGGCCTG CTTCCTCCCTGTGCGCTCTTCGAAGGGGCTTTCTGTC GCGCAGGGTCACCTCGTCCCCGAAGGGGGTGTTTGC CTTCTGGTAAATGGGGATGTCAAGTTGTGTCCGGAC ACCATGTAACAgttttagagctagaaatagcaagttaaaataaggctagtcc gttatcaacttgaaaaagtggcaccgagtcggtgcTTTTTTTTTCTTTTTC t

[00156] Table 2. Primers utilized in present disclosure.

Oligo	Sequence	Description		SEQ ID NO
P1	AAGTTGACAAAATCCTCCTCCCTGGG	gRNA against exon 3 of HXGPRT		11
P2	AAAACCCAGGGAGGAGGATTTTGTCA	gRNA against exon 3 of HXGPRT		12
P3	AAGTTGGACATAGTGCTCGAAGAAGG	gRNA against exon 5 of HXGPRT		13
P4	AAAACCTTCTTCGAGCACTATGTCCA	gRNA against exon 5 of HXGPRT		14
P5	AAGTTCCACAGAACTTACTTCGGCGG	gRNA against exon 4 of		15

		HXGPRT		
P6	AAAACCGCCGAAGTAAGTTCTGTGGA	gRNA against exon 4 of HXGPRT		16
P7	ATGGCGTCCAAACCCATTGA	screening for HXGPRT deletions		17
P8	TCGTTGAAGTCGTAGCAGCA	screening for HXGPRT deletions		18
P9	acttccaatccaatttaatTATCCAGTTGCCCGGCTC	amplification of pBAG1		19
P10	ctcaccatCATCTTTTTTGAATATCATACGGGACC	amplification of pBAG1		20
P11	tggggatgtcaagttGACTGTGGGTTGAGTTACAAGg ttttagagctagaa	first gRNA upstream of Ku80		21
P12	ttctagctctaaaacCTTGTAACTCAACCCACAGTCaa cttgacatcccca	first gRNA upstream of Ku80		22
P13	tggggatgtcaagttGAACAGAGACATCATAGACGTg ttttagagctagaa	second gRNA upstream of Ku80		23
P14	ttctagctctaaaacACGTCTATGATGTCTCTGTTCaa cttgacatcccca	second gRNA upstream of Ku80		24
P15	tggggatgtcaagttGTTTTGTCAAAGACCGCCTGAgt tttagagctagaa	first gRNA downstream of Ku80		25
P16	ttctagctctaaaacTCAGGCGGTCTTTGACAAAACa acttgacatcccca	first gRNA downstream of Ku80		26
P17	tggggatgtcaagttGGTCTTTGACAAAACGGGAGgtt ttagagctagaa	second gRNA downsream of Ku80		27
P18	ttctagctctaaaacCTCCCGTTTTGTCAAAGACCaac ttgacatcccca	second gRNA downsream of Ku80		28
P19	CTGCAGAACGCGAGACACTG	for sequence verification of gRNA constructs		29
P20	ccgacggttcgatcctgagt	screening for Ku80 deletion		30
P21	ggactttccgaccagccctc	screening for Ku80 deletion		31
P22	GTACCGACTCTTCGCAAGCG	amplifies internal to Ku80, exon 3		32
P23	TACTATCGCGCCTCGTCACG	amplifies internal to Ku80, exon 3		33
P24	tggggatgtcaagttGTTGAGTCCAAGCAGAGCTCgtt ttagagctagaa	gRNA upstream of *BFD1*		34
P25	ttctagctctaaaacGAGCTCTGCTTGGACTCAACaac ttgacatcccca	gRNA upstream of *BFD1*		35

P26	tggggatgtcaagttGTGTAGAGTCGTGGAAGGAGgtt ttagagctagaa	gRNA downstream of *BFD1*		36
P27	ttctagctctaaaacCTCCTTCCACGACTCTACACaac ttgacatcccca	gRNA downstream of *BFD1*		37
P28	cgtcacccactcacatcgtgtgagttgagtccaagcagagGCTT TTACATCCGTTGcctt	amplification of pSAG1-mNG with homology to *BFD1*		38
P29	tcatactgccgttgcgcgctccactttcagcaccccactcTTACTT GTACAGCTCGTCCA	amplification of pSAG1-mNG with homology to *BFD1*		39
P30	acattaatgcgtgcgccgca	sequencing across *BFD1* locus		40
P31	tgcttcgggcaggcgactat	sequencing across *BFD1* locus		41
P32	atggagggagccagtactcag	*BFD1* cDNA amplification, forward		42
P33	CAATCGAGCGGGTCCTGGTTCGTGTGGACCT CcaTCAAGCCCCCCGAATGCAAAGGT	*BFD1* cDNA amplification, reverse		43
P34	ACCATAACCTAGGCGGGCATTTGTTGGCACTC CAGTTTTCCGT	*BFD1* dDBD cDNA amplification, reverse		44
P35	AATGCCCGCCTAGGTTATGGT	*BFD1* dDBD cDNA amplification, forward		45
P36	CTGAGTACTGGCTCCCTCCATTGTCGAAAAAG GGAATTCAAGaaaaaatgcc	TUB1 promoter		46
P37	ctcgaggtcgacggtatcgatattaattaaCCCCCCACTGC AAGCCCTACATTGACAAAATCCTCCTCCCCAT GCATGTCCCGCGTTCG	TUB1 promoter		47
P38	GAGGTCCACACGAACCAGGACCCGCTCGATT GAtgtaacagatggaagagggt	*BFD1* 3' UTR		48
P39	aaagggaacaaaagctggagctGCGGCCGCacttacCAA CTTCTCAACTCTGTCCTTGACCAATCCACCAtg actcgcaagcgtagcacg	*BFD1* 3' UTR		49
P40	TTAGACGAGCAGGTTTCTTGCCTAT	amplification of L1 from oligonucleotide pool		50
P41	AAGTAAGCTCGCGATGTAGACGTTT	amplification of L1 from oligonucleotide pool		51
P42	GCCGATTACACCGTTAAATAACCTG	amplification of L2 from oligonucleotide pool		52
P43	TGGCGTGACTATGTTCGGTTACTAC	amplification of L2 from		53

		oligonucleotide pool		
P44	TTCTGGTAAATGGGGATGTCAAGTT	to make compatible with Gibson assembly		54
P45	gctgTTTCcagcaTAGCTCTtAAAC	to make compatible with Gibson assembly		55
P46	GTAAATGGGGATGTCAAGTTGGACTTTGACATGGTGGGTCGTTTaAGAGCTAtgctgGAA	first gRNA against mNG		56
P47	TTCcagcaTAGCTCTtAAACGACCCACCATGTCAAAGTCCAACTTGACATCCCCATTTAC	first gRNA against mNG		57
P48	GTAAATGGGGATGTCAAGTTGGCACCGGCAATCCAAATGAGTTTaAGAGCTAtgctgGAA	second gRNA against mNG		58
P49	TTCcagcaTAGCTCTtAAACTCATTTGGATTGCCGGTGCCAACTTGACATCCCCATTTAC	second gRNA against mNG		59
P50	GTAAATGGGGATGTCAAGTTGGATTCTGGTCCCTCATATCGTTTaAGAGCTAtgctgGAA	third gRNA against mNG		60
P51	TTCcagcaTAGCTCTtAAACGATATGAGGGACCAGAATCCAACTTGACATCCCCATTTAC	third gRNA against mNG		61
P52	GTAAATGGGGATGTCAAGTTGGAAGCCATACCCGATATGAGTTTaAGAGCTAtgctgGAA	fourth gRNA against mNG		62
P53	TTCcagcaTAGCTCTtAAACTCATATCGGGTATGGCTTCCAACTTGACATCCCCATTTAC	fourth gRNA against mNG		63
P54	GTAAATGGGGATGTCAAGTTGCGGTAGTTAACAGTAAGGGGTTTaAGAGCTAtgctgGAA	fifth gRNA against mNG		64
P55	TTCcagcaTAGCTCTtAAACCCCTTACTGTTAACTACCGCAACTTGACATCCCCATTTAC	fifth gRNA against mNG		65
P56	AATGATACGGCGACCACCGAGATCTACACgaatgacacacaggaactacgcg	gRNA amplification + P5 adapater for Illumina sequencing		66
P57	CAAGCAGAAGACGGCATACGAGATTCGCCTTAgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	TAAGGCGA	67
P58	CAAGCAGAAGACGGCATACGAGATCTAGTACGgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	CGTACTAG	68
P59	CAAGCAGAAGACGGCATACGAGATTTCTGCCTgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	AGGCAGAA	69
P60	CAAGCAGAAGACGGCATACGAGATGCTCAGGAgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter	TCCTGAGC	70

		for Illumina sequencing		
P61	CAAGCAGAAGACGGCATACGAGATAGGAGTC Cgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	GGACTC CT	71
P62	CAAGCAGAAGACGGCATACGAGATCATGCCT Agattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	TAGGCAT G	72
P63	CAAGCAGAAGACGGCATACGAGATGTAGAGA Ggattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	CTCTCTA C	73
P64	CAAGCAGAAGACGGCATACGAGATCAGCCTC Ggattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	CGAGGC TG	74
P65	CAAGCAGAAGACGGCATACGAGATTGCCTCT Tgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	AAGAGG CA	75
P66	CAAGCAGAAGACGGCATACGAGATTCCTCTA Cgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	GTAGAG GA	76
P67	CAAGCAGAAGACGGCATACGAGATTCATGAG Cgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	GCTCATG A	77
P68	CAAGCAGAAGACGGCATACGAGATCCTGAGA Tgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	ATCTCAG G	78
P69	CAAGCAGAAGACGGCATACGAGATTAGCGAG Tgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	ACTCGCT A	79
P70	CAAGCAGAAGACGGCATACGAGATGTAGCTC Cgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	GGAGCT AC	80
P71	CAAGCAGAAGACGGCATACGAGATTACTACG Cgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	GCGTAG TA	81
P72	CAAGCAGAAGACGGCATACGAGATAGGCTCC Ggattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	CGGAGC CT	82
P73	CAAGCAGAAGACGGCATACGAGATGCAGCGT Agattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	TACGCTG C	83

P74	CAAGCAGAAGACGGCATACGAGATCTGCGCA Tgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	ATGCGC AG	84
P75	CAAGCAGAAGACGGCATACGAGATGAGCGCT Agattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	TAGCGCT C	85
P76	CAAGCAGAAGACGGCATACGAGATCGCTCAG Tgattttcaaatggcgacctgc	gRNA amplification + barcode + P7 adapter for Illumina sequencing	ACTGAG CG	86
P77	TTTTCAAGTTGATAACGGACTAGCCTTATTTAA ACTTGCTATGCTGTTTCCAGCATAGCTCTTAA AC	Custom sequencing primer		87
P78	GCGCGACAGAAAGCCCCTTCGAAGAGCGCAC AGGGAGGAAGCAGGCCTCTGCAGGTCGCCAT TTGAAAATC	Custom indexing primer		88
P79	AGACTTCCAGTCCGGTCCTT	200385 cDNA seq1		89
P80	TCCTCAATCACGCTTAAGGC	200385 cDNA seq2		90
P81	GTACAGCTTCAAGGCATGGA	200385 cDNA seq3		91
P82	GCGGATCGTGCAGAGCAAGC	200385 cDNA seq4		92
P83	GAGGCACAGTTCTGTCAGCG	200385 cDNA seq5		93
P84	ACTACGTGAGTTATACCGCG	200385 cDNA seq6		94
P85	CACAAGCTCTCCGGTCATTG	200385 cDNA seq7		95
P86	CATGCGTGCGTTACATTGTACCT	312330 forward primer		96
P87	TGACGTGCATGATTTTGTTGTGTCTGT	208740 forward primer		97
P88	gaccgtcagcagggaaacc	mNeonGreen reverse primer		98
P89	aagttGTGTCCGGACACCATGTAACAg	gRNA against exon 1 of BFD1		99
P90	aaaacTGTTACATGGTGTCCGGACACa	gRNA against exon 1 of BFD1		100

[00157] Strain generation

[00158] C16-B3. Starting with a robustly cyst-forming ME49 strain that constitutively expresses RFP (dsRed2.0) under control of the GRA1 promoter [68], the endogenous selectable marker *HXGPRT* was inactivated through transfection with three gRNAs targeting the third, fourth and fifth exons. These gRNA expression vectors were assembled by annealing oligos P1/P2, P3/P4, and P5/P6, ligating into BsaI (NEB) digested pU6-Universal[69], and sequence

verifying with P19[69]. Transfected parasites were selected with 300 μg/mL 6-thioxanthine and screened for large deletions with P7/P8[70]. This strain was made constitutively Cas9$^+$ by co-transfection with pCas9-CAT and pU6-Decoy as described previously[71]. The strain was further transfected with ScaI (NEB) linearized pBAG1-mNeonGreen containing the promoter of BAG1 (1.22 kb upstream of the coding sequence ATG), amplified with primers P9/10 driving expression of mNeonGreen and an HXGPRT resistance cassette, and selected for integration with 25 μg/ml mycophenolic acid and 50 μg/ml xanthine [16,72]. Note this plasmid contains two identical DHFR 3′ UTRs, and care had to be taken to avoid the loss of HXGPRT by recombination during growth in bacteria.

[00159] *BFD1*$^{\text{frameshift}}$. One gRNA was designed targeting the first exon of *BFD1*. Oligos P89/P90 were annealed, Gibson-assembled into pU6-Universal, and sequenced verified with P19, generating plasmid pU6-BFD1-DHFR. Bradyzoite reporter strain parasites were transfected with 50 μg of AseI (NEB) linearized pU6-BFD1-DHFR, and selected with 3 μM pyrimethamine in standard medium the next day. After stabilization of the population, parasites were subcloned into 96-well plates at 3 parasites per well. Clonal strains isolated from single plaques were screened and sequenced for polymorphisms at the targeted site.

[00160] ME49Δ*KU80*. Two gRNAs were designed targeting regions immediately upstream or downstream of the *KU80* coding sequence. Oligos P11/12, P13/14, P15/16, P17/18 were annealed, Gibson-assembled into pU6-Universal, and sequence verified with P19. An early passage ME49 strain was transfected with 25 μg of each plasmid, and immediately subcloned into 96-well plates at 20 or 40 parasites per well to account for loss of viability during transfection. Clonal strains isolated from single plaques were screened for deletion of *KU80* with P20/21, which amplifies a band of ~5.9 kb in wildtype parasites or ~500 bp if *KU80* is excised. A single mixed population was identified from 225 strains tested, and further subcloned to isolate ME49Δ*KU80*. Loss of *KU80* was also confirmed by complete sequencing of the locus and failure to amplify an internal fragment using P22/P23.

[00161] ME49Δ*KU80*Δ*BFD1*. Two gRNAs were designed targeting regions immediately upstream or downstream of *BFD1*. Oligos P24/25 and P26/27 were annealed, Gibson-assembled into pU6-Universal, and sequenced verified with P19. A repair template consisting of the SAG1 promoter driving expression of mNeonGreen was amplified from pSAG1-mNeonGreen using primers P28/29 with 40 bp of homology to regions flanking the targeted sites. ME49Δ*Ku80* was transfected with 50 μg of each gRNA and 10 μg of repair template. 5 days post-transfection parasites were sorted by green fluorescence and subcloned. Clonal

strains isolated from single plaques were further characterized by sequencing the locus using P30/31 to confirm complete deletion of *BFD1*.

[00162] Immunofluorescence assays

[00163] HFFs were grown on coverslips for 2–3 days before inoculation with *Toxoplasma*. Coverslips were fixed with 4% formaldehyde for 20 minutes, permeabilized with 1% Triton-X 100 for 8 minutes, and blocked (5% normal goat serum and 5% IFS in phosphate buffered saline (PBS)) for at least 15 minutes (FIGS. 1A-1H, 2A-2I, 4A-4B and 16). Alternatively, fixation was done using ice-cold 100% methanol for 2 minutes, without further permeabilization, followed by blocking as above (FIGs. 12A-12B). All primary and secondary antibody incubations were performed for 1 hour, with coverslips inverted on 50 µL of antibody dilutions in blocking buffer on Parafilm in a humidified chamber. Three washes with PBS were performed after each step. Coverslips were mounted on 5 µL of Prolong Diamond (ThermoFisher) and set for 30 minutes at 37°C or overnight at room temperature. DBL-488 (Vector Labs) was used at 1:500. Mouse-anti-Ty antibody (BB2) was used at 1:1000 [73]. Rabbit-anti-GAP45 was a gift from Dominique Soldati (University of Geneva) and was used at 1:1000[74]. Rabbit-anti-SAG2Y was used at 1:2000 [67]. Mouse-anti-SAG1 (DG52) was used at 1:500 [75]. Hoechst (Santa Cruz) was used at 1:2000. Secondary antibodies labeled with Alexa Fluor 488, 594 or 647 (ThermoFisher) were used at 1:1000.

[00164] Quantification of gene disruption

[00165] Cas9-expressing C16-B3 parasites were transfected with 50 µg of AseI (NEB) linearized pU6-SAG1-DHFR[71], encoding a gRNA targeting *SAG1*. Selection with 3 µM pyrimethamine in standard media was initiated the next day, and drug-resistant pools were inoculated onto coverslips two passages (five days) after transfection. Coverslips were fixed 24 h later with methanol, and stained for SAG1 to quantify KO rates, relative to an untransfected control. GAP45 was used as a counterstain. Knockout rates were quantified before each forward genetic screen to ensure Cas9 activity.

[00166] Endogenous tagging

[00167] To endogenously tag TGME49_312330 and TGME49_208740, ME49Δ*KU80* was co-transfected with 50 µg of pCas9-CAT and 50 µg of BsaI-linearized p312330-Ty or p208740-Ty. Selection with 3 µM pyrimethamine in standard media was initiated the next day. Parasites were subcloned in 96 well plates, and isolated clones screened for successful integration using primers P86/88 or P87/88, respectively, and validated by Sanger sequencing.

[00168] Overexpression vectors of *BFD1*^{WT} and *BFD1*$^{\Delta DBD}$

[00169] The sequence of *BFD1* was amplified from ME49 cDNA using primers P32/33. To amplify *BFD1* lacking the DNA binding domain (removing amino acids 921–1019), primers P32/34 and P35/P33 were used. *BFD1* fragments were Gibson assembled together with the TUB1 promoter (amplified with P36/37) and the native *BFD1* 3′ UTR (~1.1 kb amplified with P38/39), and sequence-verified by Sanger sequencing with oligos P79-P85.

[00170] Phylogenetic analysis of *BFD1*

[00171] Protein sequences containing SANT/Myb-like domains were obtained for representative apicomplexan genomes from EupathDB based on their annotation with SMART domain SM00717. Domains from human c-Myb and CDC5L were used for comparison. Individual domains were extracted from each sequence and aligned using ClustalW and the phylogenetic tree was generated by neighbor-joining (FIG. 9) [76]. Alignments were also prepared for the concatenated domain sequences for a subset of proteins, and Bootstrap values were calculated for 10,000 trials (FIG. 2A).

[00172] Library assembly

[00173] The gRNA oligonucleotide library was synthesized by Agilent and resuspended at 1 ng/μL in water. All library amplifications were done using iProof (Bio-Rad), using 2.5 ng of the oligonucleotide pool as template per 50 μL reaction. Sublibraries were amplified using primers P40/41 for library 1 and P42/43 for library 2, and subsequently amplified with primers P44/45 for cloning. Amplified libraries were Gibson assembled into gel-extracted (Zymo) BsaI-digested pU6_Library_DHFR, dialyzed against water, and electroporated into *E. cloni* (Lucigen). Coverage was assessed by dilution plating in comparison to a no-insert negative control. Libraries were maxiprepped (Zymo), and retransformed into chemically competent NEB 5-alpha (NEB) to improve yields. Both *E. cloni* and NEB 5-alpha libraries were sequenced to ensure diversity. Libraries were linearized with AseI, dialyzed 1 h against water, and divided into 50 μg aliquots. Guide RNAs against mNeonGreen were assembled separately by annealing primer pairs P46/47, P48/49, P50/51, P52/53, or P54/55 and Gibson assembling into gel-extracted, BsaI-digested pU6_Library_DHFR. Constructs were verified by sequencing with P19, and spiked into library aliquots at equimolar concentrations.

[00174] Forward Genetic Screening

[00175] C16-B3 reporter strain parasites were grown up in ~10 15-cm dishes per screen. 10 transfections were performed for each library as described previously, with 50 μg of library transfected into approximately 2.6 x 10^7 parasites in 400 μL cytomix for each reaction [69]. Transfections were pooled and split between four 15-cm dishes. Media was changed the

next day to standard media supplemented with 3 μM pyrimethamine and 10 μg/mL DNaseI (Sigma-Aldrich). At each passage of the screen, plates were scraped, parasites were mechanically released with a 27-gauge needle, and passed through a 3 μm filter. For the second passage of screen, all parasites were passed into 4 15-cm plates without counting. All subsequent passages were performed with at a multiplicity of infection (MOI) of 1 (6×10^6 parasites per plate). Plates lysed every 2–3 days under unstressed growth in standard media supplemented with 3 μM pyrimethamine. At fourth passage (14 days post-transfection), parasites were inoculated into seven 15-cm plates, and media was changed after 4 hours to standard media supplemented with 3 μM pyrimethamine (3 plates) or alkaline stress media (4 plates). Unstressed parasites were passaged at an MOI of 1 every 2–3 days into 1-2 15-cm plates, in standard media supplemented with 3 μM pyrimethamine. Parasites under stressed conditions did not lyse out and were not passaged for the duration of the experiment, and media was changed every 2 days to fresh alkaline stress media. At each passage of unstressed parasites, $1–4 \times 10^7$ parasites were frozen down. At 10 days post media change, stressed populations were scraped, parasites were mechanically released, passed through a 3 μm filter, and sorted based on green fluorescence. At final timepoints for stressed parasites, both bulk populations ($\sim 2 \times 10^5$ parasites) and mNG^+-sorted populations ($\sim 7 \times 10^5$) were frozen. DNA was isolated using the Qiagen Blood and Tissue kit, following the protocol for blood cells. Integrated gRNAs were amplified and barcoded using primers P56 and P57–76 in 50 μL reactions. Each reaction contained 200 ng or a maximum of 20 μL of template DNA. Amplicons were gel extracted (Zymo), eluted in water, and quantified using the QuBit dsDNA HS kit (ThermoFisher). Amplicons were pooled equally at a final concentration of 8 pM each and sequenced using a MiSeq v2 kit. Reads were 40 bp single-end and an 8 bp index. Custom sequencing primer P77 and custom indexing primer P78 were used. Guides were quantified using a custom Perl script. Guides not detected were assigned a pseudocount of 90% of the lowest detected gRNA in that sample. The phenotype or differentiation score for a gene was calculated by determining the mean $\log_2$ fold-change of all five gRNAs targeting that gene in the final sample compared to the input library. All analysis done in R (https://www_r-project_org/).

[00176] Stage-specific RNA-sequencing and analysis

[00177] Parasites were allowed to invade and replicate inside host cells for 24 h in standard media, and then switched to either standard or alkaline stress media. For FACS, parasites were mechanically released from host cells using a 27- followed by a 30-gauge syringe needle, and passed through a 3 μm filter. At 24 and 48 h post media change, $\sim 1 \times 10^5$

unstressed mNG^- or stressed mNG^+ parasites were sorted directly into TRIzol LS and frozen on dry ice. Sorting was done using a BD FACS Aria II, and visualization of events and gates using FCS Express 6. RNA was extracted by TRIzol-chloroform according to manufacturer's protocol, DNaseI digested, and TRIzol-chloroform extracted again. RNA quality was assessed by BioAnalyzer or Fragment analyzer. When possible, two samples were prepared per replicate and timepoint and treated as technical replicates in downstream processing. Libraries were generated using the SMARTseq low-input v4 kit, and sequenced on two lanes of a HiSeq 2000. Reads were 75 bp, paired-end. Alignment to the ToxoDB v. 36 assembly of the ME49 genome was done using STAR[77]. Differential expression analysis was done using the DESeq2 R package [78]. The cutoff for differential expression was an adjusted *p* value of 0.001 or lower.

[00178] Single-cell RNA-sequencing and analysis

[00179] Seq-Well was performed as previously described, with the following amendments to the protocol [79]. Single cell suspensions of *Toxoplasma* were prepared by syringe release of parasites from host cells with a 27 followed by a 30-gauge needle, followed by filtering through a 5 μm filter and counting on a haemocytometer. Approximately 12,000 parasites were loaded per array, with two arrays per strain and timepoint used for stressed samples, and one array per strain for unstressed. At the 48 h timepoint, one wildtype stressed and one Δ*BFD1* stressed array failed to seal correctly, resulting in only one array per strain and growth condition at this timepoint. Sequencing was done on two NovaSeq flowcells. Pre-processing, alignment to the ToxoDB v.41 assembly of the ME49 genome, and downstream processing done following the DropSeq Cookbook (http://mccarrolllab.org/dropseq/) [80]. An estimate of the number of single cells was made using plotCumulativeFractionOfReads (function implemented by the package "Dropbead" (https://github_com/rajewsky-lab/dropbead) that estimates the number of realistic cells sequenced, based on the fraction of cumulative reads assigned to each individual cell) from Dropbead in R with a maximum of 12,000 cells[81]. The corresponding cells were then further parsed and analyzed using the Seurat R package [82]. In the analysis of all timepoints, genotypes and growth conditions, cells were required to contain a minimum of 200 and a maximum of 10,000 non-rRNA mapping UMIs and have 40% or fewer total UMIs originating from rRNA. In the analysis of the final timepoint (72 h), cells were required to contain a minimum of 500 and a maximum of 5,000 non-rRNA mapping UMIs and have 10% or fewer total UMIs originating from rRNA. Cells were log-normalized and scaled to 10,000 UMIs, regressing out the number of UMIs detected. Variable genes were identified through outlier analysis of an average

expression/dispersion scatter plot. Principal component analysis was run using these variable genes. The number of principal components (PCs) chosen to use for clustering and t-SNE visualization was based on permutation analysis and visual inspection of standard deviations of PCs. Differential gene expression between clusters or groups of clusters was performed using the Wilcoxon rank sum test, with differentially expressed genes required to be expressed in at least 10 percent of cells in one of the compared groups, have a log-fold change of 0.5 or less, and an adjusted *p* value of 0.001 or less.

[00180] Example 1: Generation of a differentiation reporter compatible with Cas9-mediated screens

[00181] To screen for *Toxoplasma* mutants deficient in differentiation, a strain compatible with Cas9-mediated gene disruption and enrichment was developed for differentiated parasites [69] (FIG. 1A). Constitutive Cas9 expression was achieved as described previously [71]. Selection for a gRNA targeting the major tachyzoite surface antigen SAG1 resulted in 98% $SAG1^-$ parasites, confirming robust inactivation of genes in this background (FIG. 1B). To facilitate isolation of differentiated parasites, the reporter strain constitutively expresses RFP, and conditionally expresses the bright green fluorescent protein mNeonGreen (mNG) under the promoter of the canonical bradyzoite-specific gene *BAG1*. Growth of the reporter strain, referred to as C16-B3, under alkaline stress resulted in increasing proportions of parasites expressing mNeonGreen (FIGs. 1C and 5A).

[00182] To characterize transcriptomic differences between tachyzoites and bradyzoites, stage-specific bulk RNA-sequencing was performed using C16-B3. Gene expression of FACS-purified tachyzoites (mNG^-, 24 h unstressed growth) was compared to bradyzoites (mNG^+, 48 h stressed growth), and 1311 genes identified as upregulated and 933 genes as downregulated in bradyzoites (FIG. 1D). Principal component analysis shows 98% of variance is explained by growth condition, with minimal batch effects (FIG. 5B). Highly regulated genes show agreement with previous datasets, with the most highly upregulated genes including the canonical bradyzoite-specific genes bradyzoite antigen 1 (BAG1), lactate dehydrogenase 2 (LDH2), and enolase 1 (ENO1) (FIG. 5C). One of the most highly downregulated genes is the major tachyzoite surface antigen SAG1. Genes previously missed as differentially regulated tend to be more lowly expressed, suggesting enhanced sensitivity in the dataset (Fig. 5D).

[00183] Example 2: Forward genetic screening identifies a putative regulator of *T. gondii* differentiation

[00184] Genome-wide forward screens in *Toxoplasma* have been performed successfully, but performing these screens in a non-lab-attenuated strain presents additional challenges. In particular, the lower viability and integration rates observed suggest the largest number of genes that can be screened to be in the low hundreds (FIG. 6). Therefore, sets of candidate genes small enough to meet these technical limitations were curated. By combining differential expression analysis with domain annotation and gene ontology, two libraries of ~100 potential nucleic acid–binding proteins were assembled, targeting each gene with 5 guide RNAs (gRNAs). Library 1 (L1) consists of genes identified as differentially regulated in a preliminary RNA-seq experiment. Library 2 (L2) contains genes with DNA-binding domains commonly found in transcription factors, such as zinc finger and Myb-like domains. Across both libraries all 67 members of the ApiAP2 transcription factor family are targeted, along with 151 putative nucleic acid-binding proteins (Fig. 1E) [83,84]. As controls, each library additionally contains 10 genes known to be essential, 10 genes known to be dispensable, 10 non-cutting gRNAs, and 5 gRNAs against the mNG reporter itself. The effect inactivation of a gene has upon parasite fitness can be calculated by looking at changes in gRNA abundances over the course of an experiment. It is expected that gRNAs targeting highly essential genes to negatively affect parasite fitness and will be depleted quickly from the population, while gRNAs against non-essential genes or non-existent sequences should be retained. Guide against the mNG reporter should not affect fitness, but will mimic an inability to differentiate.

[00185] Following transfection of the libraries, parasites were passaged in selective media for four passages before being split between unstressed or stressed conditions for 10 days, passaging unstressed parasites as necessary (FIG. 1E). After 10 days, stressed mNG^+ parasites were isolated by FACS. Integrated gRNAs from the final unstressed passage or from stressed mNG^+ alkaline stress parasites were amplified, sequenced, and compared to the input library. The mean $\log_2$ fold-change of all gRNAs against a gene from input library to final unstressed sample or stressed, mNG^+ sample is referred to as a fitness or differentiation score, respectively. Candidate genes should be depleted specifically in the mNG^+ population (low differentiation score relative to their fitness score), as should guides against the mNG reporter. In L1, only control mNG gRNAs were depleted in the mNG^+ parasite population (FIG. 1F). However, in L2, gRNAs targeting a single other gene—TGME49_200385, renamed bradyzoite formation deficient 1 (*BFD1*)—were depleted along with control mNG gRNAs (FIGs. 1G–1H). Failure of a *BFD1* mutant to express the mNG reporter following alkaline stress was confirmed by transfecting a single *BFD1* targeting gRNA into C16-B3. A frameshifted clone was isolated with a single nucleotide insertion at the cut site (FIG. 7).

Under alkaline stress, the wildtype reporter strain shows robust mNG expression, while the frameshifted clone does not.

[00186] Example 3: BFD1 contains conserved DNA-binding domains and localizes to the nucleus

[00187] The sequence of the *BFD1* open reading frame was defined based on cDNA sequencing, which differed from the annotated gene model and encoded a protein of 2,415 amino acids (FIG. 8). BFD1 contains two tandem SANT/myb-like DNA-binding domains (SMART accession 00717), flanked by large extensions lacking identifiable motifs. The *Toxoplasma* genome encodes 14 proteins with SANT/myb-like domains. Phylogenetic analysis of these proteins reveals BFD1's two DNA-binding domains have homology to the R2 and R3 repeats of the human c-Myb protein, respectively (FIGs. 2A–2B and FIG. 9). BFD1 orthologs can be found in all Apicomplexan species that form tissue cysts, while different patterns of conservation are found for other SANT domain–containing proteins. Transiently expressed, full-length BFD1-Ty localized to the nucleus, and co-localized with a DNA stain (FIG. 2C).

[00188] Example 4: Loss of BFD1 blocks parasite differentiation

[00189] To provide a clean background for precise genetic manipulation, a low-passage, NHEJ-deficient ME49 strain was generated through deletion of *KU80* (FIG. 10). In this background, the entire coding sequence of *BFD1* was replaced with an mNG expression cassette (FIG. 2D). Deletion of *BFD1* caused no defect in tachyzoite growth as assayed by plaque formation (FIG 2E). *Dolichos biflorus* lectin (DBL) recognizes *N*-acetylgalactosamine on the bradyzoite-specific cyst wall protein CST1, providing a convenient way of distinguishing differentiating vacuoles [6]. The ability of a Δ*BFD1* strain to differentiate following 48 h of alkaline stress was tested. Wildtype vacuoles became robustly DBL^+ under alkaline stress conditions. By contrast, no Δ*BFD1* vacuoles developed DBL positivity (FIGs. 2F–2G). Δ*BFD1* parasites exhibited a similar defect when differentiation was induced by a small molecule known to induce differentiation in *Toxoplasma* (Compound 1)[24,25] or occurred spontaneously (FIGs. 2H–2I).

[00190] Example 5: Single-cell sequencing can be used to analyze Toxoplasma cell-cycle progression

[00191] To profile cell cycle progression and the asynchronous process of differentiation, the first single-cell RNA-sequencing of *T. gondii* was performed using Seq-Well [79]. Wildtype or Δ*BFD1* parasites were grown under unstressed or stressed conditions for 24, 48 or 72 h. Following downstream processing and alignment, 26,560 cells passed quality control cutoffs,

with an average of 1537 UMIs per cell representing an average of 685 genes per cell. (FIGs. 3A, 11A-11C). As cells from the 72 h timepoint were of the highest quality, unstressed parasites from this timepoint were clustered to examine the tachyzoite cell cycle. Seven clusters were identified, with six arranged in a circular pattern (FIG. 3B). 1,173 genes were identified as differentially expressed in one or more clusters. The average expression profile for markers from each cluster was determined using an existing dataset of synchronized tachyzoite gene expression [1]. The timing of marker expression indicates a progression through the cell cycle in a counter clock-wise direction through the various clusters (FIG. 3C). The proportion of cells identified as being in G1 or S/M matches the 60:40 ratio previously determined (FIG. 3D) [85].

[00192] Example 6: BFD1 is necessary for initiation of differentiation

[00193] Clustering cells from all timepoints, growth conditions and genotypes revealed a clear division between tachyzoite ($SAG1^+$) and bradyzoite ($BAG1^+$) containing clusters (FIG. 3E). Wildtype and Δ*BFD1* parasites respond to alkaline stress very differently over time (FIG. 3F). WT parasites quickly exit the tachyzoite cell cycle and begin progressing towards bradyzoite-containing clusters. By contrast, Δ*BFD1* parasites continue to replicate as tachyzoites until the 72 h timepoint, when they largely leave and cluster separately from both tachyzoites and bradyzoites. This suggests Δ*BFD1* parasites fail to initiate differentiation despite the presence of an inducing signal. Examination of the timing of expression of canonical bradyzoite markers in wildtype cells suggested BAG1 is not among the earliest genes induced (FIG. 12A). Endogenous tagging of two strong, early markers of differentiation (TGME49_312330, TGME49_208740) localized both to the cyst wall (FIG. 12B).

[00194] To understand the nature of the clusters parasites end up in, the clustering of unstressed and stressed parasites of both genotypes from the 72 h timepoint were analyzed (FIG. 3G). The canonical stage-specific genes *SAG1* and *BAG1* show high, mutually exclusive expression, demonstrating strong separation of tachyzoites and bradyzoites (FIG. 13A). Tachyzoites (high *SAG1*) are colored in red, while bradyzoites (high *BAG1*) are in blue, according to expression level of each marker. Clusters with minimal *BAG1* expression and weaker *SAG1* expression are in gray. Stressed wildtype parasites overlap strongly with BAG1-expressing clusters, while stressed Δ*BFD1* parasites either continue replicating as tachyzoites or move to the weakly $SAG1^+$ clusters (FIG. 13B). By contrast to bulk RNA-sequencing, in the single-cell data *BFD1* is identified as differentially regulated in bradyzoite-containing clusters (FIG. 13C).

[00195] Examining highly expressed, stage-specific genes identified by the bulk RNA-sequencing showed widespread defects in bradyzoite-specific gene expression by Δ*BFD1* parasites, including canonical markers such as LDH2 and ENO1 (FIG. 3H). Weaker expression of *SAG1* and other normally highly expressed genes in stressed Δ*BFD1* parasites made us wonder if these parasites were still actively dividing. Based on the cell-cycle regulation of the genes driving them, the first and third principal components explain the majority of variance due to the cell cycle (FIGs. 14A-14C). Wildtype stressed parasites plotted across these components show a clear cyclical pattern, indicating active bradyzoite replication (FIG. 3I). Many genes were identified as expressed specifically in replicating bradyzoites, including the previously identified bradyzoite rhoptry protein 1 (FIG. 15) [86]. 19% of bradyzoites are in clusters expressing S/M markers compared to 39.5% of tachyzoites, recapitulating the approximately two-fold slower replication previously observed [87]. Non-tachyzoite Δ*BFD1* stressed parasites show no such pattern, suggesting these parasites are no longer replicating (FIG. 3I). It has additionally been observed that aberrant morphologies in later timepoints of alkaline stressed Δ*BFD1* parasites (FIG. 16). These results suggest that BFD1 is necessary for parasites to initiate differentiation, and that after 72 hours under alkaline stress Δ*BFD1* parasites are dying due to their failure to respond to that stress appropriately (FIG. 3J).

[00196] Example 7: Overexpression of BFD1 is sufficient to induce differentiation in the absence of stress

[00197] As *BFD1* is necessary for differentiation to occur, it was investigated whether overexpression of *BFD1* would be sufficient to induce differentiation. Two constructs expressing epitope-tagged BFD1 under the TUB1 promoter–either full-length BFD1 ($BFD1^{WT}$) or a mutant protein with its DNA-binding domain removed ($BFD1^{\Delta DBD}$) were generated (FIG. 4A). Wildtype or Δ*BFD1* parasites were transfected with either $BFD1^{WT}$ or $BFD1^{\Delta DBD}$, and allowed to grow for 48 h under unstressed conditions. At 48 h, parasites were immunostained for BFD1 (Ty) and differentiation (DBL). Vacuoles containing Ty^{+} parasites were identified and scored for DBL positivity. Both constructs localized to the nucleus. Transient overexpression of $BFD1^{WT}$, but not $BFD1^{\Delta DBD}$, was sufficient to induce differentiation in over 60% of wildtype or Δ*BFD1* parasites, demonstrating that the inability of Δ*BFD1* parasites to differentiate is due to the specific absence of *BFD1* (FIG. 4B). The requirement of the DNA-binding domain of BFD1 suggests its activity as a *bona fide* transcription factor drives differentiation.

[00198] Example 8: Loss of BFD1 blocks parasite differentiation regardless of induction method

[00199] Δ*BFD1* parasites were complemented by introducing Ty-tagged cDNA copies of *BFD1* at the endogenous locus, either full-length (Δ*BFD1*::*BFD1*^{WT}-*Ty*) or with its Myb-like domains deleted (Δ*BFD1*::*BFD1*$^{\Delta MYB}$-*Ty*) (FIG. 17A). Deletion of *BFD1* and subsequent complementation caused no defect in tachyzoite growth as assayed by plaque formation. The complemented strains grew slightly faster, perhaps due to prolonged passaging in cell culture (FIG. 17B). Differentiating vacuoles can be identified using *Dolichos biflorus* lectin (DBL) staining, which recognizes N-acetylgalactosamine on the bradyzoite-specific cyst-wall protein CST1 (Tomita et al., 2013). Many cyst wall proteins, including CST1, are robustly detected in large proportions of vacuoles after only 24 h under alkaline stress, making DBL positivity an earlier and more robust marker for differentiating parasites than *BAG1* transcription which accumulates more slowly (see FIG. 1C). WT vacuoles became robustly DBL+ after 48 h under alkaline stress. By contrast, no Δ*BFD1* vacuoles developed DBL positivity under identical conditions.

[00200] Complementation with the WT but not the ΔMYB allele of *BFD1* restored differentiation and revealed nuclear localization of the transgenes (FIGS. 17C and 17D). Δ*BFD1* parasites also failed to differentiate spontaneously or when induced with the small molecule Compound 1 (FIG. 17D), and in both cases complementation with full-length *BFD1* restored differentiation to WT levels (Radke et al., 2006).

[00201] Example 9: Δ*BFD1* parasites fail to form brain cysts in mice

[00202] To determine whether BFD1 is necessary for the formation of tissue cysts in animals, acute virulence and chronic infection of mice were assessed. CD-1 female mice were infected by intraperitoneal injection with 500 tachyzoites of WT, Δ*BFD1*, or Δ*BFD1*::*BFD1*^{WT} parasites (FIG. 18A). Morbidity and mortality were comparable among the three strains (FIGS. 18B and 18C), indicating that BFD1 is dispensable for the acute symptoms of Toxoplasma infection. The marginal increase in Δ*BFD1*::*BFD1*^{WT} virulence may be attributed to the faster growth rate observed by plaque formation (FIG. 18D). Brain cysts from WT and Δ*BFD1*::*BFD1*^{WT} infections were morphologically identical, with a DBL positive cyst wall surrounding hundreds of bradyzoites (FIG. 18D). Starting at 2 weeks post-infection, cyst burden was measured by examining the brains of infected animals. Cyst numbers ranged from several hundred to several thousand per animal infected with WT or Δ*BFD1*::*BFD1*^{WT} parasites; however, cysts were never isolated from Δ*BFD1*-infected animals (FIG. 18E). Similar results were observed in experiments conducted with CBA/J

mice (FIG. 19). Taken together, these results reveal that loss of *BFD1* results in a specific and complete defect in tissue cyst formation during animal infections.

[00203] References

[00204] 1. Behnke, M. S. *et al.* Coordinated progression through two subtranscriptomes underlies the tachyzoite cycle of Toxoplasma gondii. *PLoS One* **5,** e12354 (2010).

[00205] 2. Baird, J. K. Resistance to Therapies for Infection by Plasmodium vivax. *Clinical microbiology reviews* **22,** 508–534 (2009).

[00206] 3. Montoya, J. G. & Liesenfeld, O. Toxoplasmosis. *The Lancet* **363,** 1965–1976 (2004).

[00207] 4. Porter, S. B. & Sande, M. A. Toxoplasmosis of the Central Nervous System in the Acquired Immunodeficiency Syndrome. *New England Journal of Medicine* **327,** 1643–1648 (1992).

[00208] 5. Tomita, T. *et al.* Making Home Sweet and Sturdy: Toxoplasma gondii ppGalNAc-Ts Glycosylate in Hierarchical Order and Confer Cyst Wall Rigidity. *MBio* **8,** e02048–16 (2017).

[00209] 6. Tomita, T. *et al.* The Toxoplasma gondii Cyst Wall Protein CST1 Is Critical for Cyst Wall Integrity and Promotes Bradyzoite Persistence. *PLOS Pathog* **9,** e1003823 (2013).

[00210] 7. Ferguson, D. J. P. Use of molecular and ultrastructural markers to evaluate stage conversion of Toxoplasma gondii in both the intermediate and definitive host. *International journal for parasitology* **34,** 347–360 (2004).

[00211] 8. Coppin, A. *et al.* 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 (2005).

[00212] 9. Denton, H., Roberts, C. W., Alexander, J., Thong, K.-W. & Coombs, G. H. Enzymes of energy metabolism in the bradyzoites and tachyzoites of Toxoplasma gondii. *FEMS Microbiology Letters* **137,** 103–108 (1996).

[00213] 10. Guérardel, Y. *et al.* Amylopectin biogenesis and characterization in the protozoan parasite Toxoplasma gondii, the intracellular development of which is restricted in the HepG2 cell line. *Microbes and Infection* **7,** 41–48 (2005).

[00214] 11. Sugi, T. *et al.* Toxoplasma gondii Requires Glycogen Phosphorylase for Balancing Amylopectin Storage and for Efficient Production of Brain Cysts. *MBio* **8,** e01289–17 (2017).

[00215] 12. Manger, I. D., Hehl, A., Parmley, S. & Sibley, L. D. Expressed sequence tag analysis of the bradyzoite stage of Toxoplasma gondii: identification of developmentally regulated genes. *Infection and Immunity* (1998).

[00216] 13. Yahiaoui, B. *et al.* Isolation and characterization of a subtractive library enriched for developmentally regulated transcripts expressed during encystation of Toxoplasma gondii. *Molecular and Biochemical Parasitology* **99,** 223–235 (1999).

[00217] 14. Cleary, M. D., Singh, U., Blader, I. J. & Brewer, J. L. Toxoplasma gondii asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryotic Cell* (2002).

[00218] 15. Radke, J. R. *et al.* The transcriptome of Toxoplasma gondii. *BMC Biology 2005 3:1* **3,** 1 (2005).

[00219] 16. Behnke, M. S., Radke, J. B., Smith, A. T., Sullivan, W. J., Jr & White, M. W. The transcription of bradyzoite genes in Toxoplasma gondii is controlled by autonomous promoter elements. *Molecular microbiology* **68,** 1502–1518 (2008).

[00220] 17. Naguleswaran, A., Elias, E. V., McClintick, J., Edenberg, H. J. & Sullivan, W. J. Toxoplasma gondii lysine acetyltransferase GCN5-A functions in the cellular response to alkaline stress and expression of cyst genes. *PLOS Pathog* **6,** e1001232 (2010).

[00221] 18. Lescault, P. J., Thompson, A. B., Patil, V., Lirussi, D. & Burton, A. Genomic data reveal Toxoplasma gondii differentiation mutants are also impaired with respect to switching into a novel extracellular tachyzoite state. *PLoS One* (2010).

[00222] 19. Buchholz, K. R. *et al.* Identification of tissue cyst wall components by transcriptome analysis of in vivo and in vitro Toxoplasma gondii bradyzoites. *Eukaryotic cell* **10,** 1637–1647 (2011).

[00223] 20. Pittman, K. J., Aliota, M. T. & Knoll, L. J. Dual transcriptional profiling of mice and Toxoplasma gondii during acute and chronic infection. *BMC Genomics 2014 15:1* **15,** 806 (2014).

[00224] 21. Ramakrishnan, C. *et al.* An experimental genetically attenuated live vaccine to prevent transmission of Toxoplasma gondii by cats. *Sci Rep* **9,** 55 (2019).

[00225] 22. Soete, M., Camus, D. & Dubrametz, J. F. Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of Toxoplasma gondii in vitro. *Experimental parasitology* (1994).

[00226] 23. Fox, B. A., Gigley, J. P. & Bzik, D. J. Toxoplasma gondii lacks the enzymes required for de novo arginine biosynthesis and arginine starvation triggers cyst formation. *International journal for parasitology* **34,** 323–331 (2004).

[00227] 24. Donald, R. G. K. *et al.* Toxoplasma gondii cyclic GMP-dependent kinase: chemotherapeutic targeting of an essential parasite protein kinase. *Eukaryotic cell* **1,** 317–328 (2002).

[00228] 25. Radke, J. R. *et al.* Changes in the Expression of Human Cell Division Autoantigen-1 Influence Toxoplasma gondii Growth and Development. *PLOS Pathog* **2,** e105 (2006).

[00229] 26. Singh, U., Brewer, J. L. & Boothroyd, J. C. Genetic analysis of tachyzoite to bradyzoite differentiation mutants in Toxoplasma gondii reveals a hierarchy of gene induction. *Molecular microbiology* (2002).

[00230] 27. Matrajt, M., Donald, R. & Singh, U. Identification and characterization of differentiation mutants in the protozoan parasite Toxoplasma gondii. *Molecular Microbiology* (2002).

[00231] 28. Saksouk, N. *et al.* Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite Toxoplasma gondii. *Molecular and cellular biology* **25,** 10301–10314 (2005).

[00232] 29. Narasimhan, J. *et al.* Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in Toxoplasma gondii. *J. Biol. Chem.* **283,** 16591–16601 (2008).

[00233] 30. Bougdour, A., Maubon, D., Baldacci, P. & Ortet, P. Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. *The Journal of Experimental Medicine* (2009).

[00234] 31. Maubon, D. *et al.* Activity of the histone deacetylase inhibitor FR235222 on Toxoplasma gondii: inhibition of stage conversion of the parasite cyst form and study of new derivative compounds. *Antimicrob. Agents Chemother.* **54,** 4843–4850 (2010).

[00235] 32. Gissot, M. *et al.* Toxoplasma gondii Alba Proteins Are Involved in Translational Control of Gene Expression. *Journal of Molecular Biology* **425,** 1287–1301 (2013).

[00236] 33. Sullivan, W. J., Narasimhan, J., Bhatti, M. M. & Wek, R. C. Parasite-specific eIF2 (eukaryotic initiation factor-2) kinase required for stress-induced translation control. *Biochemical Journal* **380,** 523–531 (2004).

[00237] 34. Konrad, C., Queener, S. F., Wek, R. C. & Sullivan, W. J. Inhibitors of eIF2α dephosphorylation slow replication and stabilize latency in Toxoplasma gondii. *Antimicrob. Agents Chemother.* **57,** 1815–1822 (2013).

[00238] 35. Holmes, M. J., Augusto, L. D. S., Zhang, M., WEK, R. C. & Sullivan, W. J., Jr. Translational Control in the Latency of Apicomplexan Parasites. *Trends in Parasitology* **33,** 947–960 (2017).

[00239] 36. Jeffers, V., Tampaki, Z., Kim, K. & Sullivan, W. J. A latent ability to persist: differentiation in Toxoplasma gondii. *Cell. Mol. Life Sci.* **75,** 2355–2373 (2018).

[00240] 37. Radke, J. B. *et al.* Transcriptional repression by ApiAP2 factors is central to chronic toxoplasmosis. *PLOS Pathog* **14,** e1007035 (2018).

[00241] 38. Radke, J. B. *et al.* ApiAP2 transcription factor restricts development of the Toxoplasma tissue cyst. *Proc. Natl. Acad. Sci. U.S.A.* **110,** 6871–6876 (2013).

[00242] 39. Walker, R. *et al.* The Toxoplasma nuclear factor TgAP2XI-4 controls bradyzoite gene expression and cyst formation. *Molecular microbiology* **87,** 641–655 (2013).

[00243] 40. Huang, S. *et al.* Toxoplasma gondii AP2IX-4 Regulates Gene Expression during Bradyzoite Development. *mSphere* **2,** e00054–17 (2017).

[00244] 41. Hong, D.-P., Radke, J. B. & White, M. W. Opposing Transcriptional Mechanisms Regulate Toxoplasma Development. *mSphere* **2,** e00347–16 (2017).

[00245] 42. Buchholz, K. R., Bowyer, P. W. & Boothroyd, J. C. Bradyzoite pseudokinase 1 is crucial for efficient oral infectivity of the Toxoplasma gondii tissue cyst. *Eukaryotic cell* **12,** 399–410 (2013).

[00246] 43. Ehrenkaufer, G. M., Hackney, J. A. & Singh, U. A developmentally regulated Myb domain protein regulates expression of a subset of stage-specific genes in Entamoeba histolytica. *Cellular Microbiology* **11,** 898–910 (2009).

[00247] 44. Sun, C.-H., Palm, D., McArthur, A. G., Svärd, S. G. & Gillin, F. D. A novel Myb-related protein involved in transcriptional activation of encystation genes in Giardia lamblia. *Molecular microbiology* **46,** 971–984 (2002).

[00248] 45. Huang, Y.-C. *et al.* Regulation of cyst wall protein promoters by Myb2 in Giardia lamblia. *J. Biol. Chem.* **283,** 31021–31029 (2008).

[00249] 46. Dubos, C. *et al.* MYB transcription factors in Arabidopsis. *Trends in Plant Science* **15,** 573–581 (2010).

[00250] 47. Ambawat, S., Sharma, P., Yadav, N. R. & Yadav, R. C. MYB transcription factor genes as regulators for plant responses: an overview. *Physiol Mol Biol Plants* **19,** 307–321 (2013).

[00251] 48. Sandberg, M. L. *et al.* c-Myb and p300 regulate hematopoietic stem cell proliferation and differentiation. *Dev. Cell* **8,** 153–166 (2005).

[00252] 49. Fuglerud, B. M., Ledsaak, M., Rogne, M., Eskeland, R. & Gabrielsen, O. S. The pioneer factor activity of c-Myb involves recruitment of p300 and induction of histone acetylation followed by acetylation-induced chromatin dissociation. *Epigenetics Chromatin* **11,** 35 (2018).

[00253] 50. Boschet, C. *et al.* Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 Plasmodium falciparum clones. *Molecular and Biochemical Parasitology* **138,** 159–163 (2004).

[00254] 51. Gissot, M., Briquet, S., Refour, P., Boschet, C. & Vaquero, C. PfMyb1, a Plasmodium falciparum Transcription Factor, is Required for Intra-erythrocytic Growth and Controls Key Genes for Cell Cycle Regulation. *Journal of Molecular Biology* **346,** 29–42 (2005).

[00255] 52. Poran, A. *et al.* Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. *Nature* **551,** 95–99 (2017).

[00256] 53. Bancells, C. *et al.* Revisiting the initial steps of sexual development in the malaria parasite Plasmodium falciparum. *Nature Microbiology 2018 4:1* **4,** 144–154 (2019).

[00257] 54. Eksi, S. *et al.* Plasmodium falciparum Gametocyte Development 1 (Pfgdv1) and Gametocytogenesis Early Gene Identification and Commitment to Sexual Development. *PLOS Pathog* **8,** e1002964 (2012).

[00258] 55. Ikadai, H. *et al.* Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis. *Proc. Natl. Acad. Sci. U.S.A.* **110,** E1676–84 (2013).

[00259] 56. Sinha, A. *et al.* A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* **507,** 253 (2014).

[00260] 57. Kafsack, B. F. C. *et al.* A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507,** 248 (2014).

[00261] 58. Brancucci, N. M. B. *et al.* Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host & Microbe* **16,** 165–176 (2014).

[00262] 59. Coleman, B. I. *et al.* A Plasmodium falciparum histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host & Microbe* **16,** 177–186 (2014).

[00263] 60. Filarsky, M. *et al.* GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. *Science ...* **359,** 1259–1263 (2018).

[00264] 61. Kent, R. S. *et al.* Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite Plasmodium berghei. *Nature Microbiology 2018 4:1* **3,** 1206–1213 (2018).

[00265] 62. Hassan, M. A., Vasquez, J. J., Guo-Liang, C., Meissner, M. & Nicolai Siegel, T. Comparative ribosome profiling uncovers a dominant role for translational control in Toxoplasma gondii. *BMC Genomics 2014 15:1* **18,** 961 (2017).

[00266] 63. Melo, M. B. *et al.* Transcriptional Analysis of Murine Macrophages Infected with Different Toxoplasma Strains Identifies Novel Regulation of Host Signaling Pathways. *PLOS Pathog* **9,** e1003779 (2013).

[00267] 64. Reid, A. J. *et al.* Comparative genomics of the apicomplexan parasites Toxoplasma gondii and Neospora caninum: Coccidia differing in host range and transmission strategy. *PLOS Pathog* **8,** e1002567 (2012).

[00268] 65. Swierzy, I. J. *et al.* Divergent co-transcriptomes of different host cells infected with Toxoplasma gondii reveal cell type-specific host-parasite interactions. *Sci Rep* **7,** 7229 (2017).

[00269] 66. Al-Anouti, F., Tomavo, S., Parmley, S. & Ananvoranich, S. The expression of lactate dehydrogenase is important for the cell cycle of Toxoplasma gondii. *J. Biol. Chem.* **279,** 52300–52311 (2004).

[00270] 67. Saeij, J., Arrizabalaga, G. & Boothroyd, J. C. A cluster of four surface antigen genes specifically expressed in bradyzoites, SAG2CDXY, plays an important role in Toxoplasma gondii persistence. *Infection and immunity* (2008).

[00271] 68. Oldenhove, G., Bouladoux, N., Wohlfert, E. A. & Hall, J. A. Decrease of Foxp3+ Treg cell number and acquisition of effector cell phenotype during lethal infection. *Immunity* (2009).

[00272] 69. Sidik, S. M., Hackett, C. G., Tran, F., Westwood, N. J. & Lourido, S. Efficient genome engineering of Toxoplasma gondii using CRISPR/Cas9. *PLoS One* (2014).

[00273] 70. Donald, R. G. K., Carter, D., Ullman, B. & Roos, D. S. Insertional tagging, cloning, and expression of the Toxoplasma gondii hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. *J. Biol. Chem.* **271,** 14010–14019 (1996).

[00274] 71. Sidik, S. M. *et al.* A Genome-wide CRISPR Screen in Toxoplasma Identifies Essential Apicomplexan Genes. *Cell* **166,** 1423–1435.e12 (2016).

[00275] 72. Bohne, W., Wirsing, A. & Gross, U. Bradyzoite-specific gene expression in Toxoplasma gondii requires minimal genomic elements. *Molecular and Biochemical Parasitology* Mar;85(1):89-98. (1997).

[00276] 73. Bastin, P., Bagherzadeh, Z., Matthews, K. R. & Gull, K. A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei. *Molecular and Biochemical Parasitology* **77,** 235–239 (1996).

[00277] 74. Plattner, F. *et al.* Toxoplasma profilin is essential for host cell invasion and TLR11-dependent induction of an interleukin-12 response. *Cell Host & Microbe* **3,** 77–87 (2008).

[00278] 75. Burg, J. L., Perelman, D., Kasper, L. H., Ware, P. L. & Boothroyd, J. C. Molecular analysis of the gene encoding the major surface antigen of Toxoplasma gondii. *The Journal of Immunology* **141,** 3584–3591 (1988).

[00279] 76. Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948 (2007).

[00280] 77. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29,** 15–21 (2013).

[00281] 78. Love, M., Anders, S. & Huber, W. Differential analysis of RNA-Seq data at the gene level using the DESeq2 package. (2013).

[00282] 79. Gierahn, T. M. *et al.* Seq-Well: portable, low-cost RNA sequencing of single cells at high throughput. *Nature* ... **14,** 395–398 (2017).

[00283] 80. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161,** 1202–1214 (2015).

[00284] 81. Alles, J. *et al.* Cell fixation and preservation for droplet-based single-cell transcriptomics. *BMC Biology 2005 3:1* **15,** 44 (2017).

[00285] 82. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36,** 411–420 (2018).

[00286] 83. Painter, H. J., Campbell, T. L. & Llinas, M. The Apicomplexan AP2 family: integral factors regulating Plasmodium development. *Molecular and Biochemical Parasitology* **176,** 1–7 (2011).

[00287] 84. Balaji, S., Babu, M. M., Iyer, L. M. & Aravind, L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic acids research* (2005).

[00288] 85. Radke, J. R. *et al.* Defining the cell cycle for the tachyzoite stage of Toxoplasma gondii. *Molecular and Biochemical Parasitology* **115,** 165–175 (2001).

[00289] 86. Schwarz, J. A., Fouts, A. E., Cummings, C. A., Ferguson, D. J. P. & Boothroyd, J. C. A novel rhoptry protein in Toxoplasma gondii bradyzoites and merozoites. *Molecular and Biochemical Parasitology* **144,** 159–166 (2005).

[00290] 87. Dzierszinski, F., Nishi, M., Ouko, L. & Roos, D. S. Dynamics of Toxoplasma gondii differentiation. *Eukaryotic cell* (2004).

CLAIMS

What is claimed is:

1. A genetically altered protozoan parasite comprising a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite.

2. The parasite of claim 1, wherein the mutation is a deletion of all or a portion of the coding sequence of the *BFD1* gene.

3. The parasite of claim 1 or 2, wherein the mutation is a deletion of the entire coding sequence of the *BFD1* gene.

4. The parasite of any one of claims 1-3, wherein the mutation is a loss-of-function mutation.

5. The parasite of claim 4, wherein the loss-of-function mutation is a null mutation.

6. The parasite of any one of claims 1-5, wherein the mutation is a dominant negative mutation.

7. The parasite of any one of claims 1-6, wherein the parasite is an apicomplexan parasite.

8. The parasite of claim 7, wherein the apicomplexan parasite is a *Toxoplasma* parasite, a *Plasmodium* parasite, a *Hammondia* parasite, a *Neospora* parasite or a *Sarcocystis* parasite.

9. The parasite of claim 7 or 8, wherein the apicomplexan parasite is *Toxoplasma gondii*.

10. The parasite of claim 7 or 8, wherein the apicomplexan parasite is *Neospora caninum*.

11. The parasite of claim 7 or 8, wherein the apicomplexan parasite is *Sarcocystis neurona*.

12. A vaccine composition comprising the parasite of any one of claims 1-11 and a pharmaceutically-acceptable carrier.

13. The vaccine composition of claim 12, further comprising an adjuvant.

14. The vaccine composition of claim 12 or 13, comprising a live vaccine.

15. The vaccine composition of any one of claims 12-14, wherein the parasite expresses a heterologous antigen.

16. The vaccine composition of claim 15, wherein the heterologous antigen is a cancer antigen.

17. The vaccine composition of any one of claims 12-16, wherein the parasite expresses a therapeutic agent.

18. The vaccine composition of claim 17, wherein the therapeutic agent is a peptide or protein.

19. A recombinant nucleic acid vector comprising a nucleotide sequence encoding a bradyzoite formation deficient 1 (BFD1) protein.

20. The nucleic acid vector of claim 19, wherein the vector is an expression vector.

21. A host cell comprising the nucleic acid vector of claim 19 or 20.

22. The host cell of claim 21, wherein the host cell is *Toxoplasma gondii*.

23. A method of inducing an immune response to an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite.

24. The method of claim 23, wherein the apicomplexan parasite is *Toxoplasma gondii*.

25. The method of claim 23 or 24, wherein the subject in need thereof is a human.

26. The method of claim 23 or 24, wherein the subject in need thereof is a non-human mammal.

27. The method of any one of claims 23-26, wherein the subject in need thereof has an acute or chronic apicomplexan parasite infection.

28. A method of inhibiting or preventing a chronic apicomplexan parasite infection in a subject, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite.

29. A method of treating a chronic infection by an apicomplexan parasite in a subject in need thereof, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite.

30. A method for inoculating a subject in need thereof with an apicomplexan parasite, comprising administering to the subject a vaccine composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite.

31. A method of administering an antigen to a subject in need thereof comprising administering to the subject a composition comprising a genetically altered protozoan parasite, wherein the parasite comprises a mutation in a bradyzoite formation deficient 1 (*BFD1*) gene, wherein the mutation inhibits differentiation of the parasite into a bradyzoite, and wherein the parasite comprises an antigen.

32. The method of claim 31, wherein the parasite is genetically altered to comprise the antigen.

33. The method of claim 31 or 32, wherein the antigen is a cancer antigen.

FIG. 1A

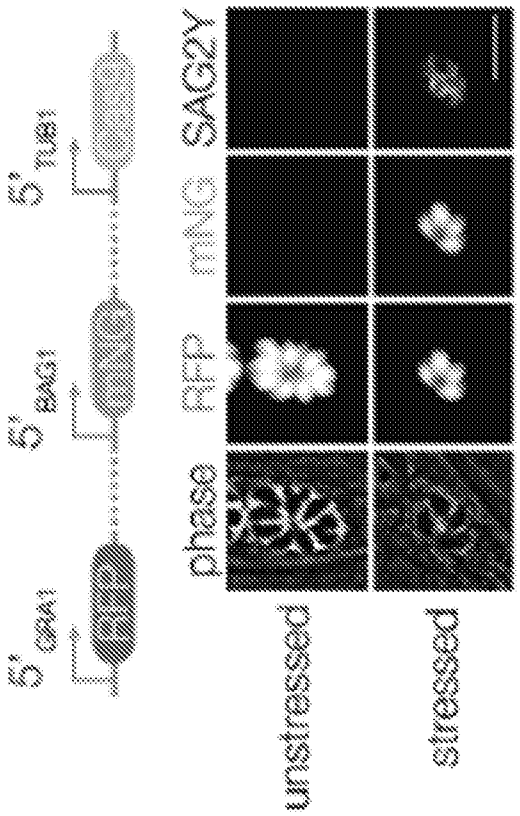

5' GPA1
RFP
5' BAG1
mNG
5' TUB1
phase
RFP
mNG
SAG2Y
unstressed
stressed

FIG. 1B

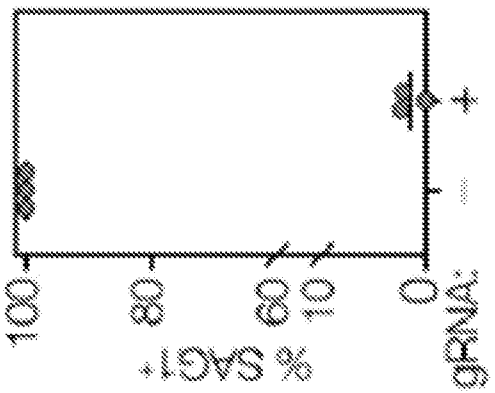

100
80
60
40
0
% SAG1+
-
+
gRNA:

FIG. 1D

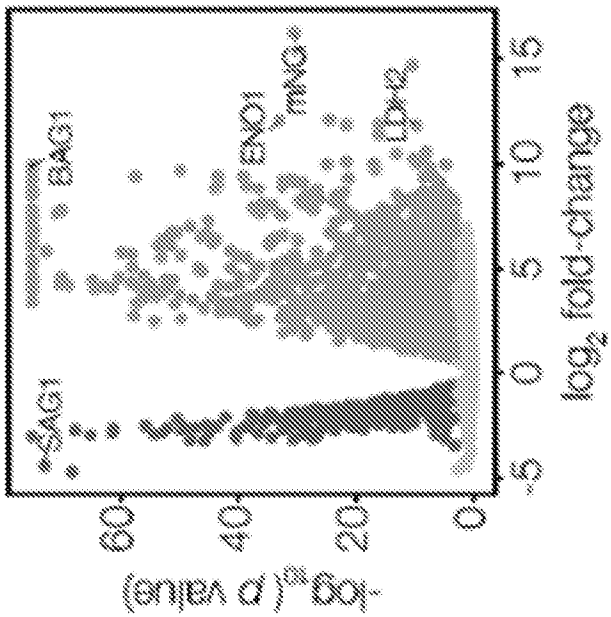

BAG1
ENO1
mNG+
Cpr2
ΔBAG1
-log10(p value)
log2 fold-change
60
40
20
0
-5
0
5
10
15

FIG. 1C

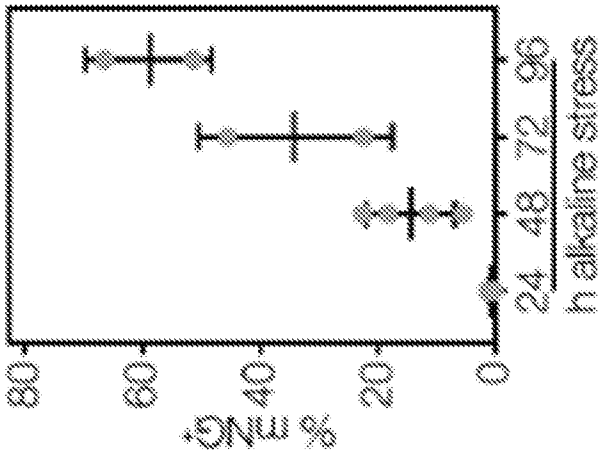

% mNG+
80
60
40
20
0
24
48
72
96
h alkaline stress

FIG. 1E

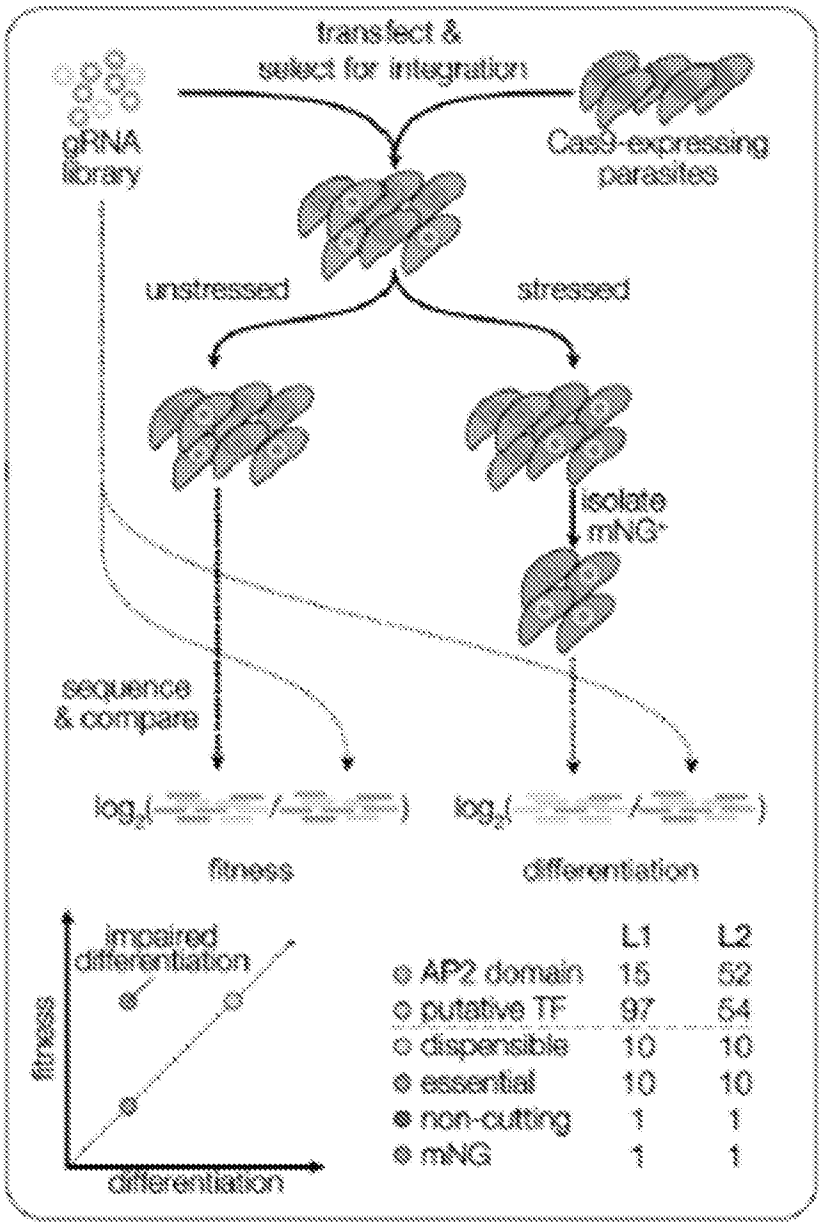

transfect &
select for integration
gRNA
library
Cas9-expressing
parasites
unstressed
stressed
isolate
mNG+
sequence
& compare
fitness
differentiation
impaired
differentiation
fitness
differentiation
L1 L2
AP2 domain 16 62
putative TF 97 64
dispensible 10 10
essential 10 10
non-cutting 1 1
mNG 1 1

FIG. 1G

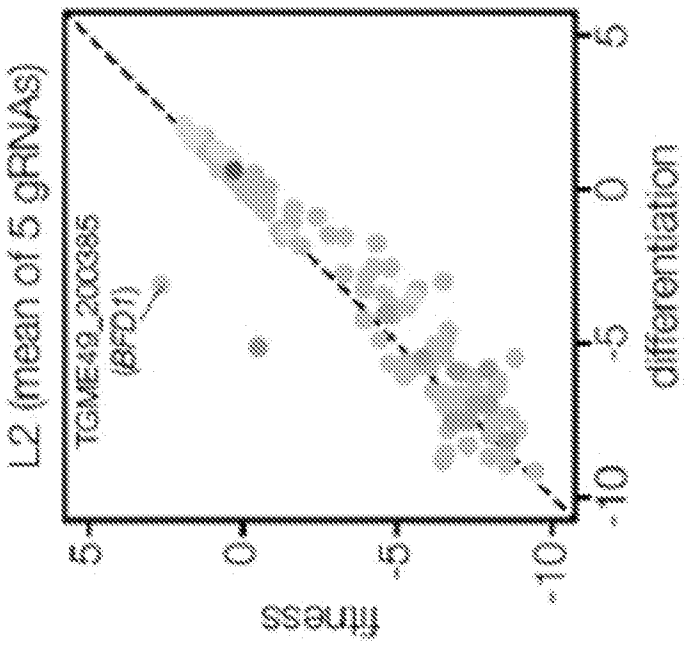

L2 (mean of 5 gRNAs)
fitness
differentiation
TGME49_200385
(BFD1)
5
0
-5
-10

FIG. 1F

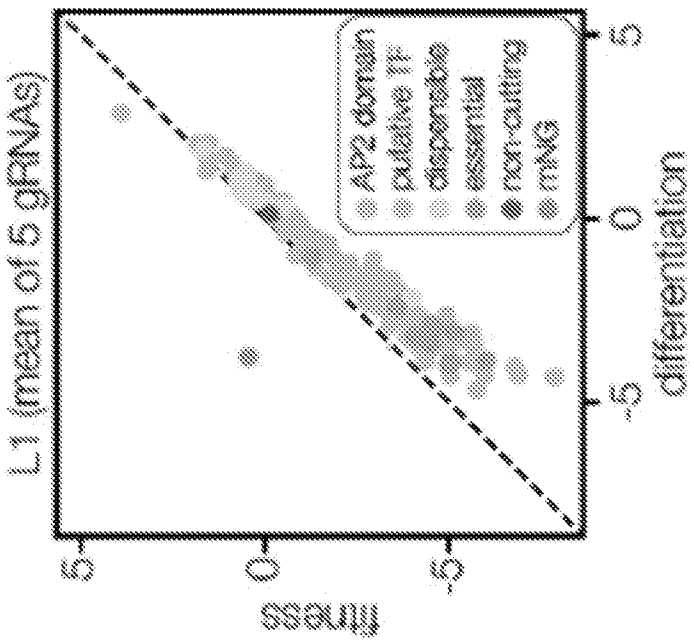

L1 (mean of 5 gRNAs)
fitness
differentiation
AP2 domain
putative TF
dispensable
essential
non-cutting
rhAG
5
0
-5

FIG. 1H

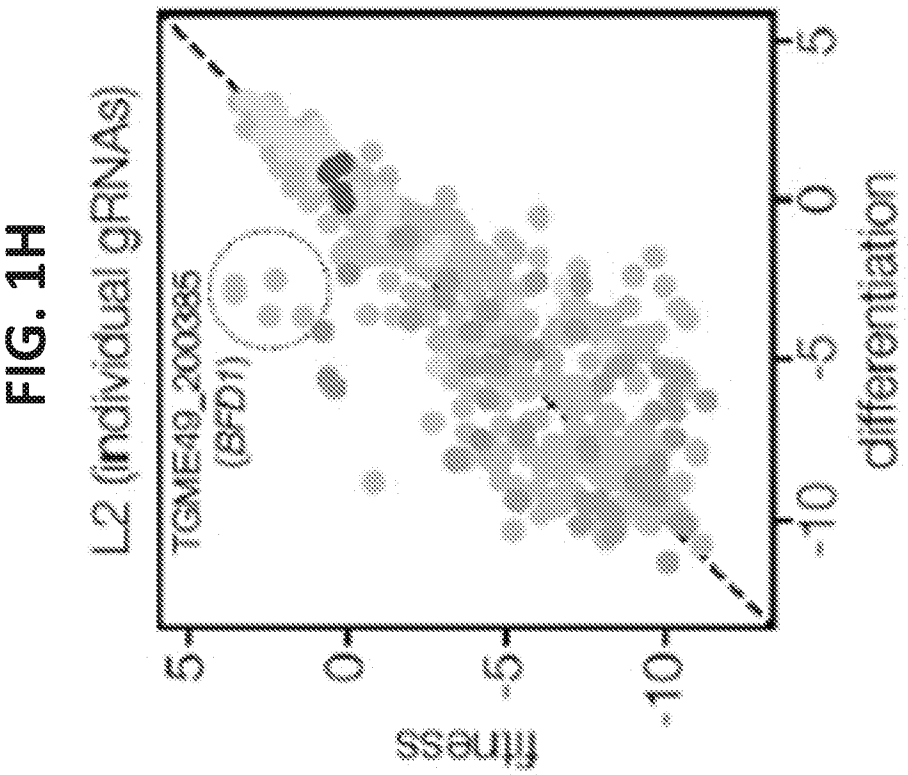

L2 (individual gRNAs)
fitness
differentiation
TGME49_200935
(GDI)
5
0
-5
-10

FIG. 2A

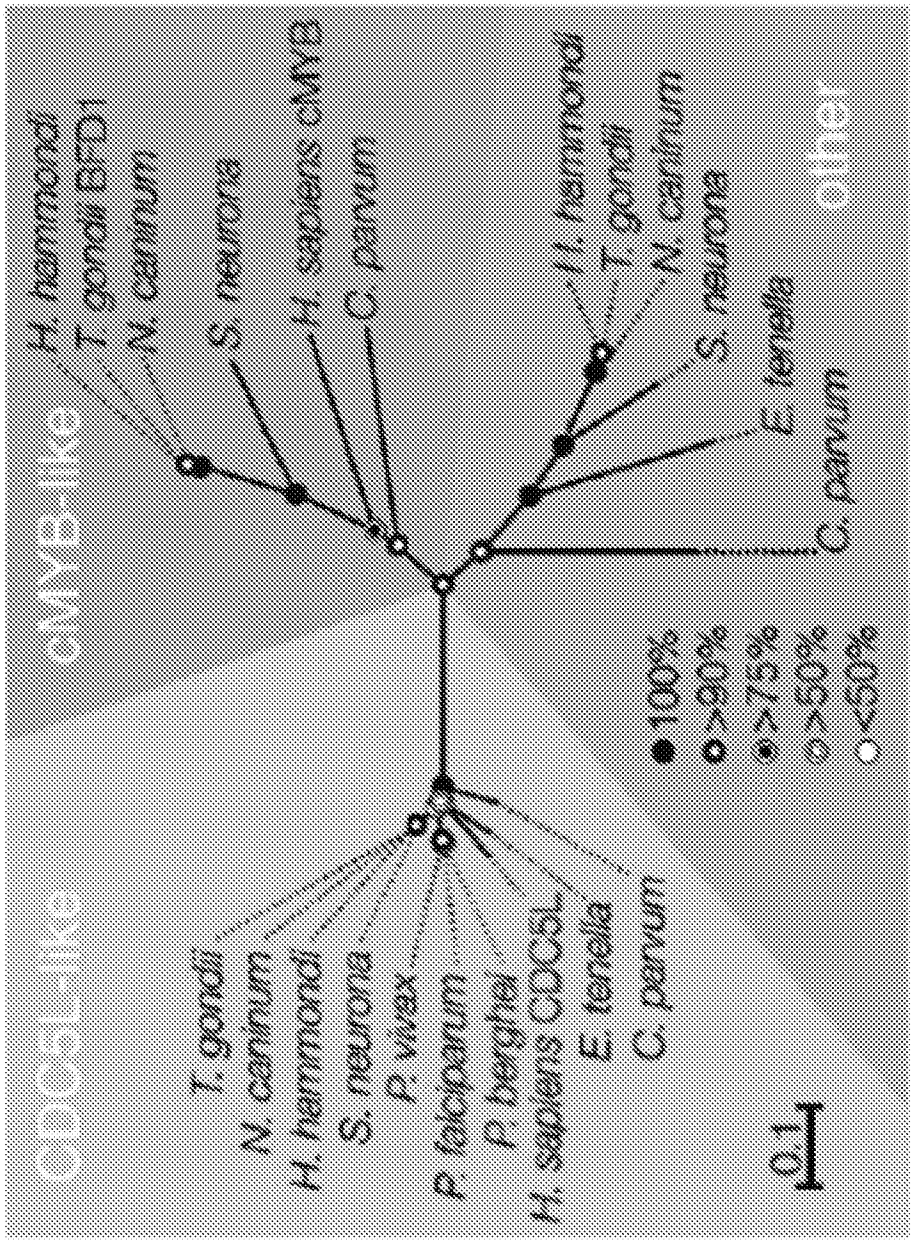

CDC6-like
cMYB-like
H. hammondi
T. gondii GT1
N. caninum
S. neurona
H. sapiens cMYB
C. parvum
H. hammondi
T. gondii
N. caninum
S. neurona
E. tenella
C. parvum
outlier
T. gondii
N. caninum
H. hammondi
S. neurona
P. vivax
P. falciparum
P. berghei
H. sapiens CDC6L
E. tenella
C. parvum
100%
>90%
>75%
>50%
<50%
0.1

FIG. 2B

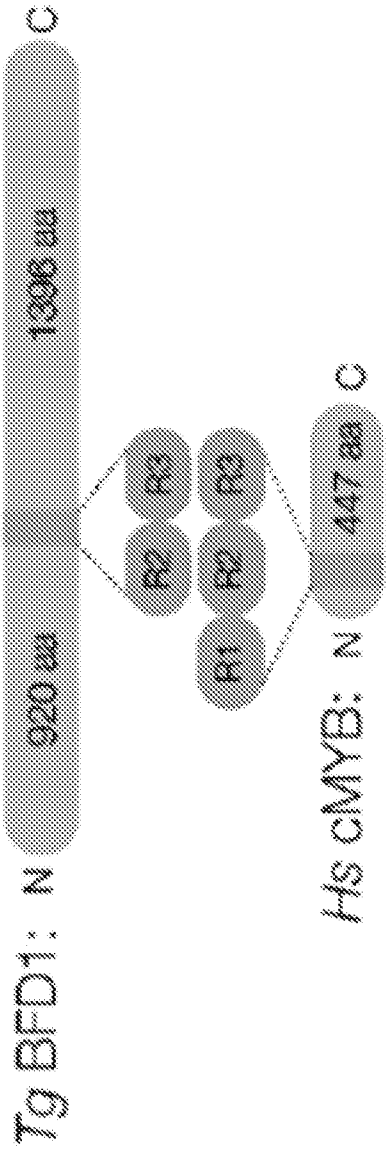
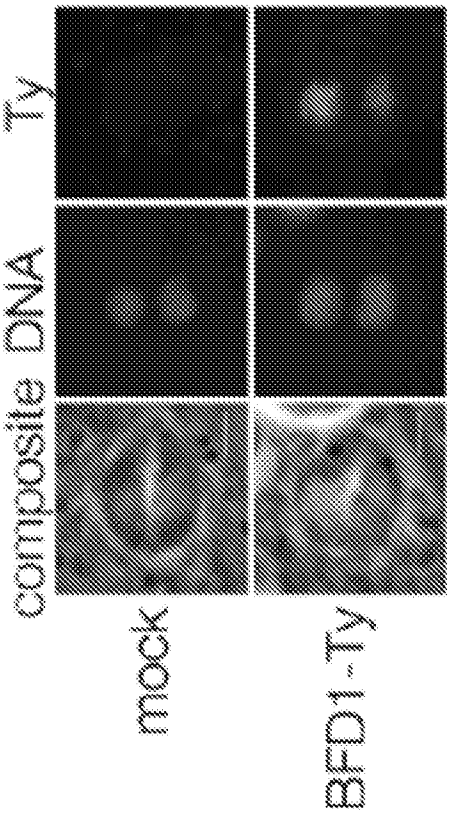

Tg BFD1: N
920 aa
1306 aa
C
R1
R2
R3
Hs cMYB: N
447 aa
C

FIG. 2C


composite DNA
Ty
mock
BFD1-Ty

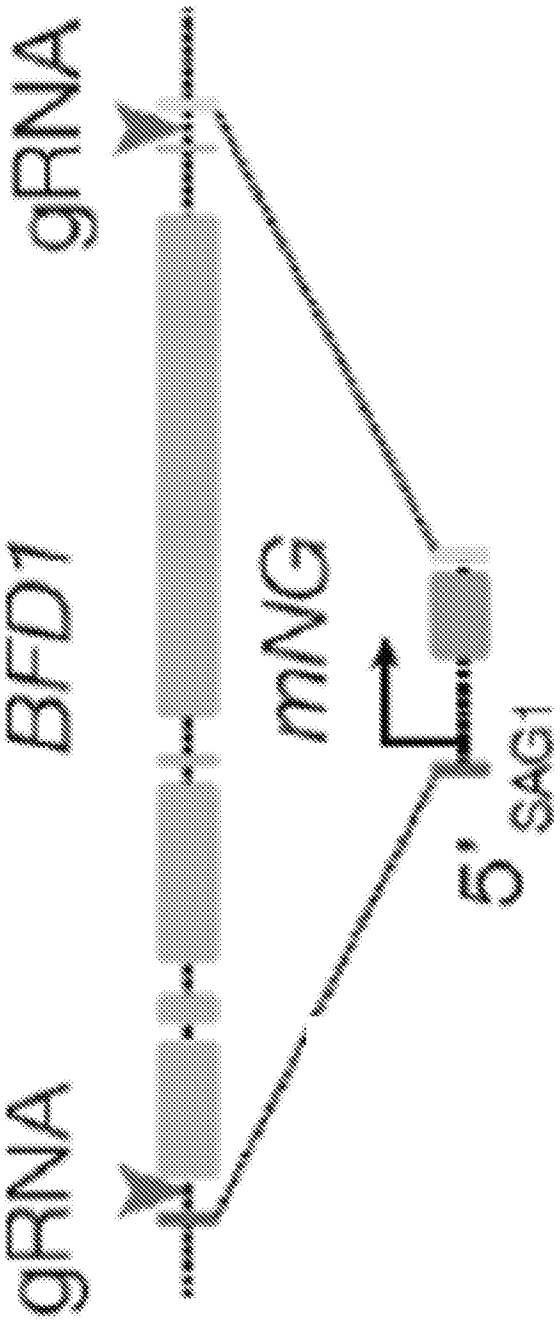


FIG. 2D

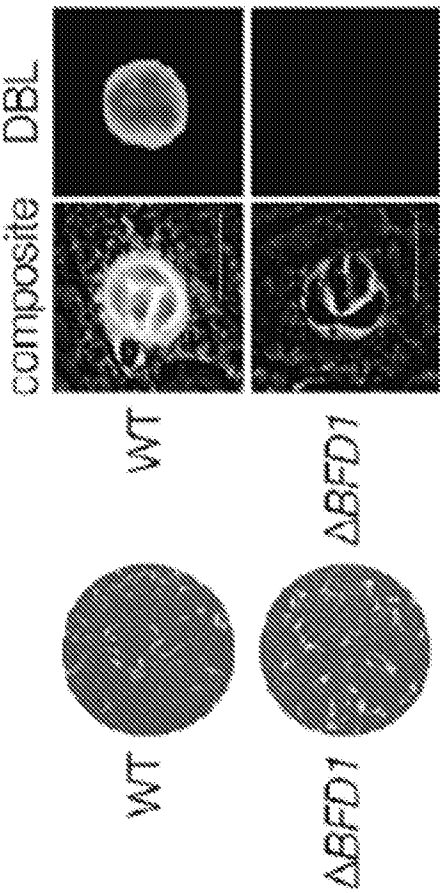


FIG. 2F

FIG. 2E

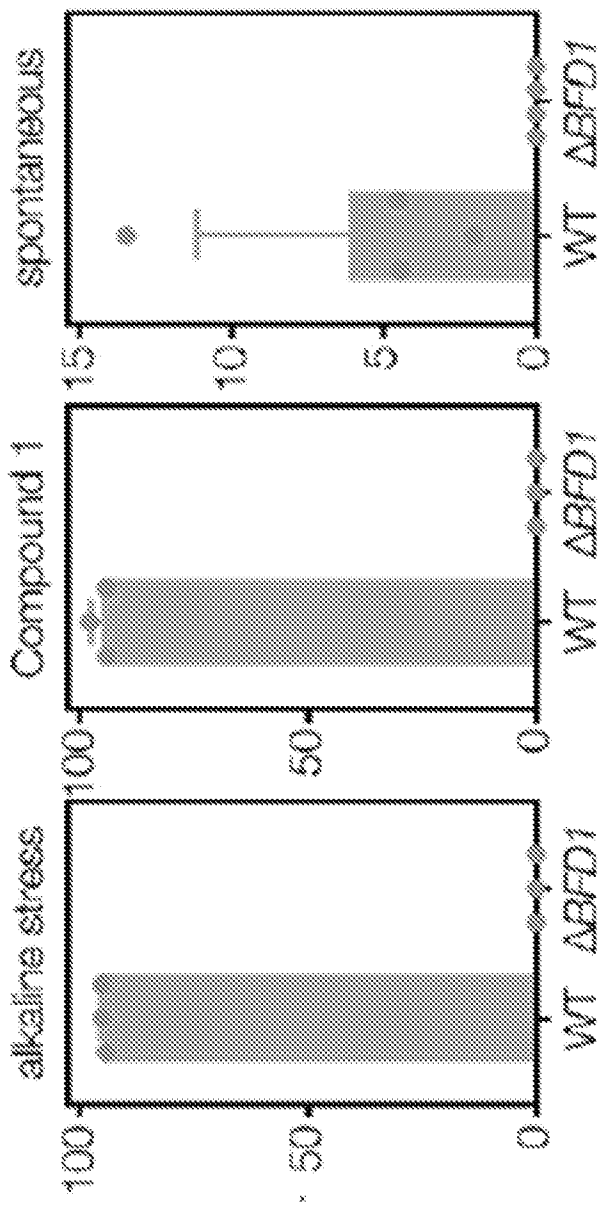


FIG. 2G

FIG. 2H

FIG. 2I

FIG. 3A

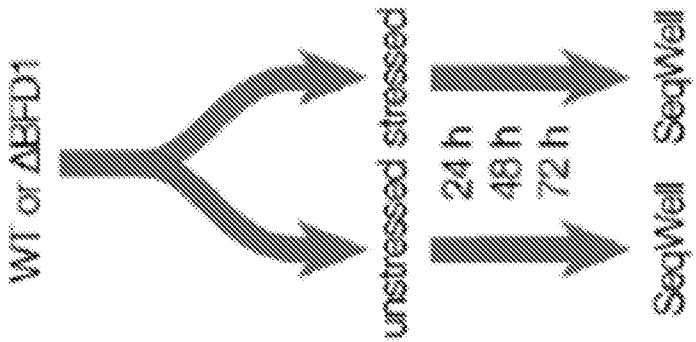

WT or ΔEPD1
unstressed
stressed
24 h
48 h
72 h
SeqWell
SeqWell

FIG. 3B

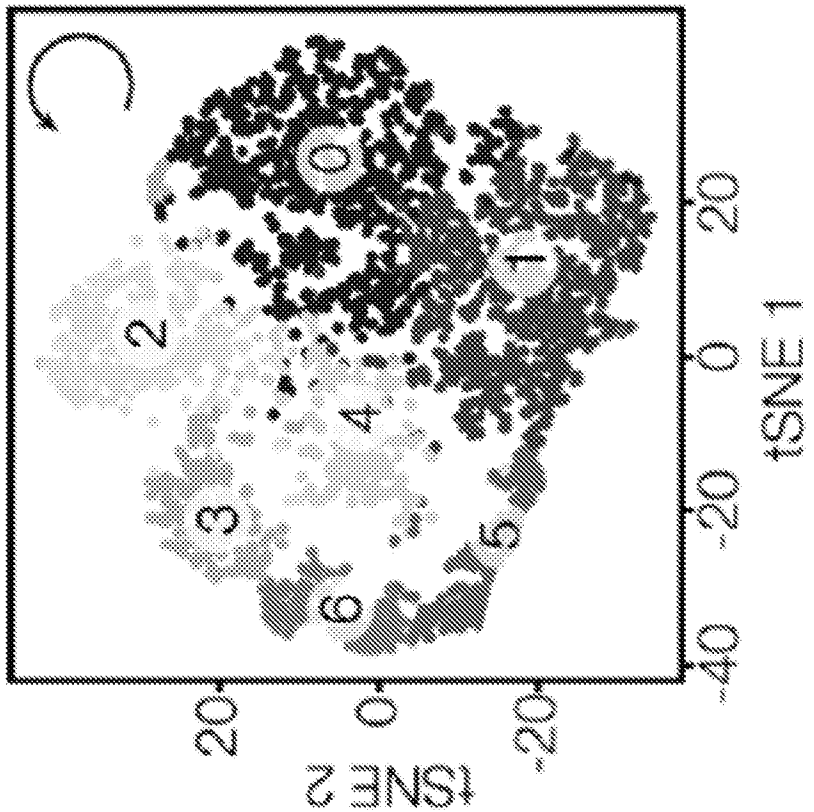

tSNE 2
20
0
-20
tSNE 1
20
0
-20
-40
0
1
2
3
4
5
6

FIG. 3D

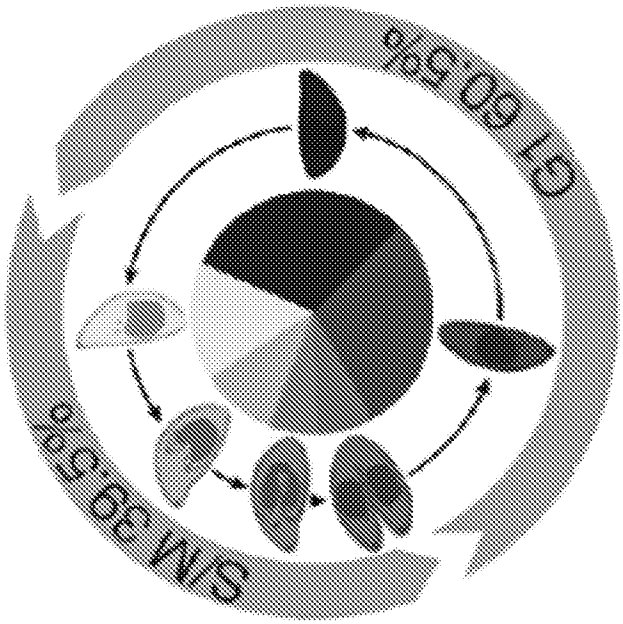

G1 60.5%
S/M 39.5%

FIG. 3C

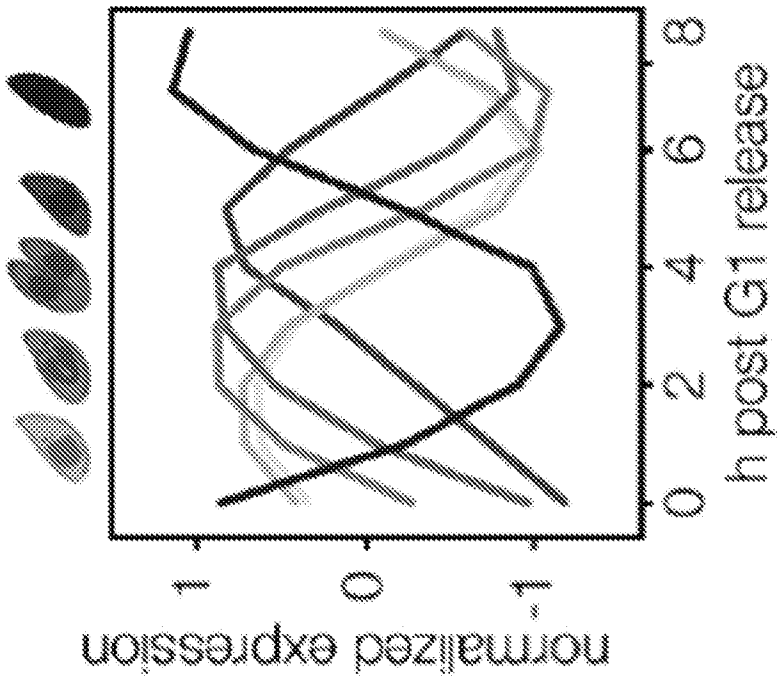

normalized expression
-1
0
1
0
2
4
6
8
h post G1 release

FIG. 3E

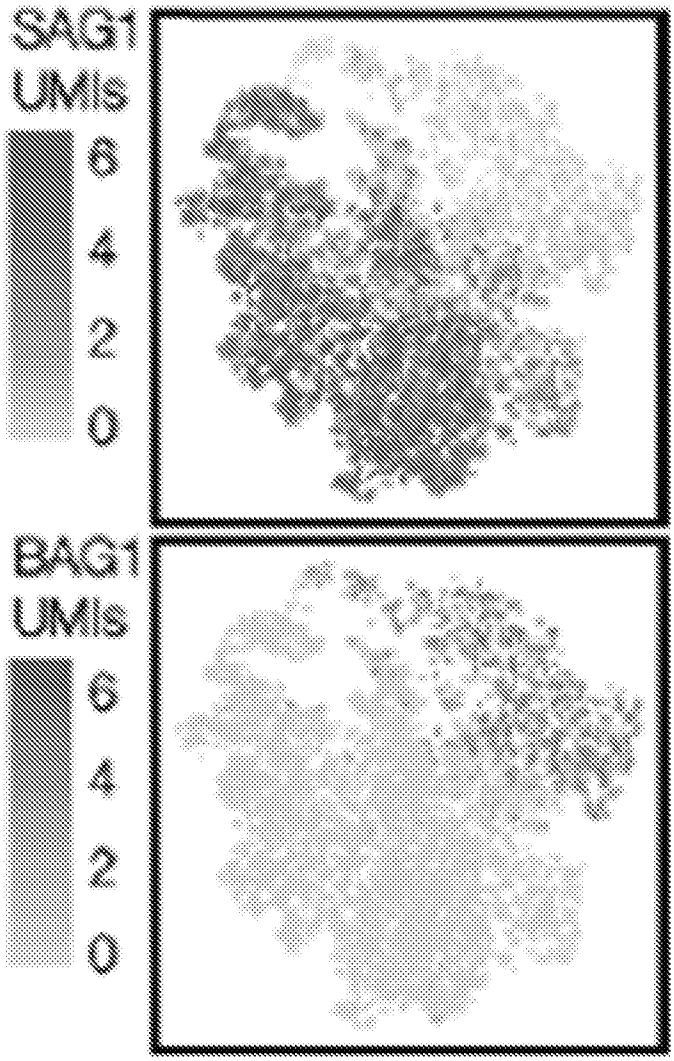


FIG. 3F

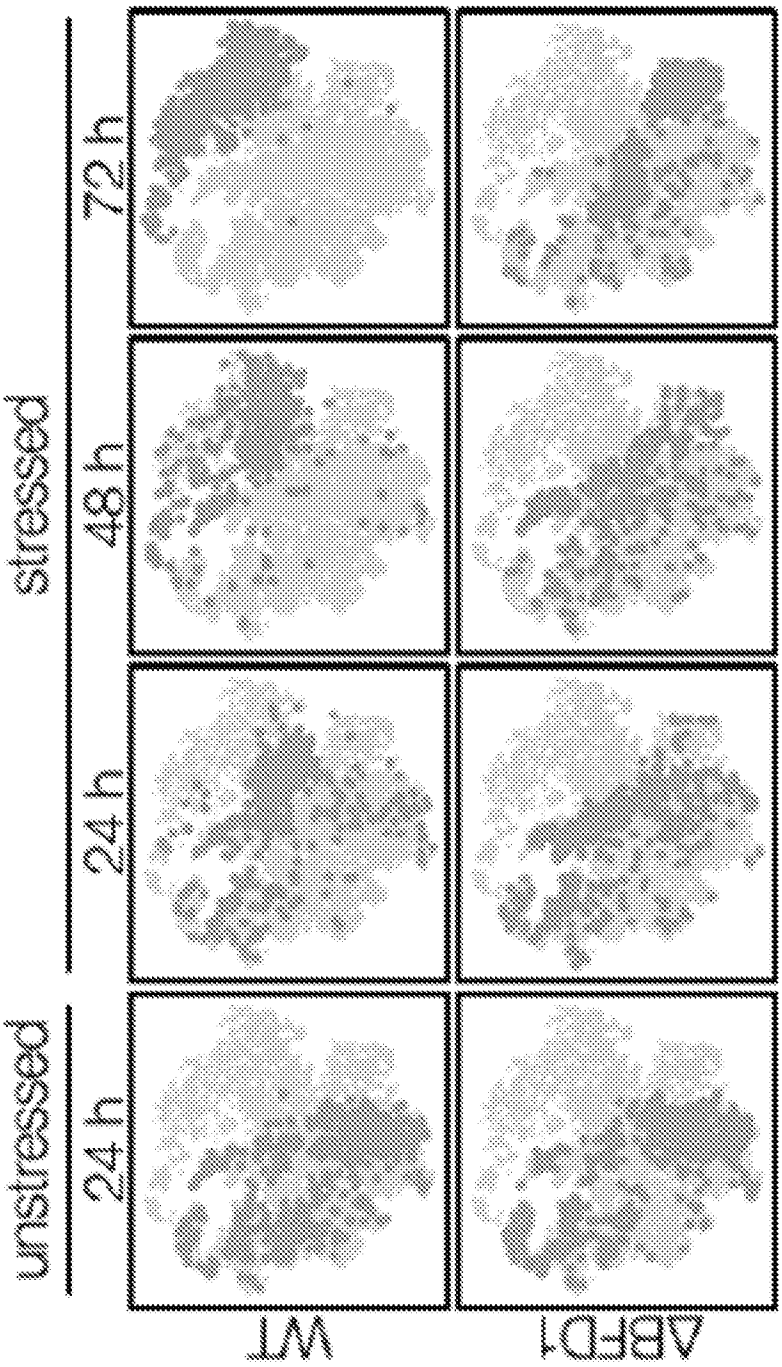


FIG. 3G

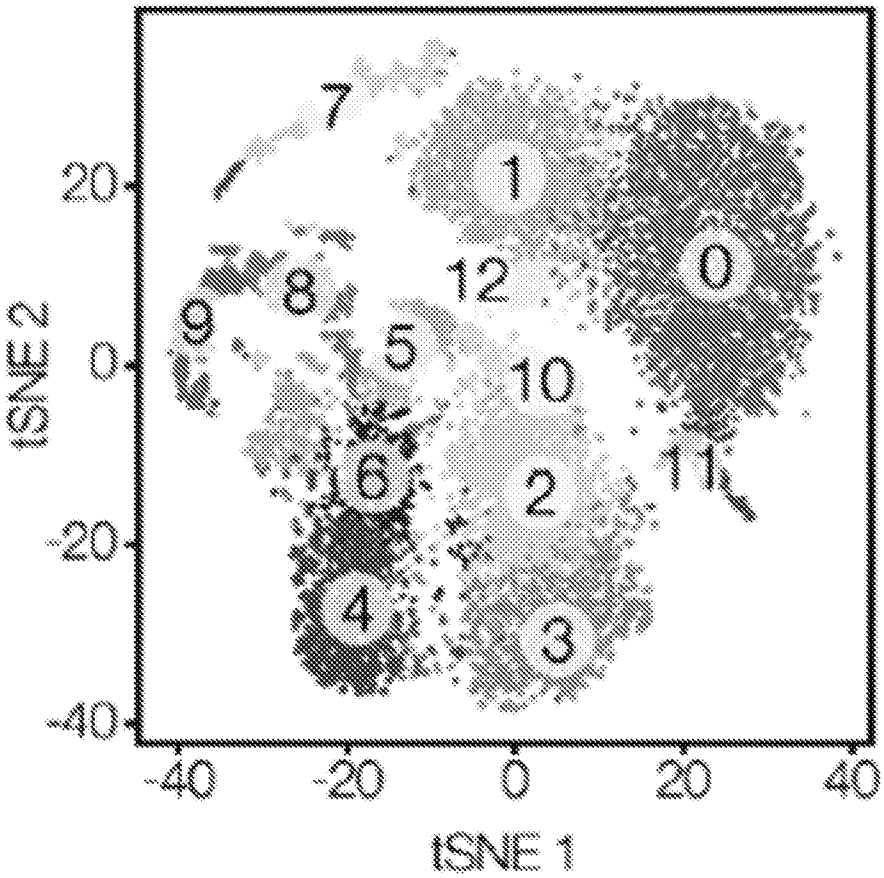


FIG. 3H

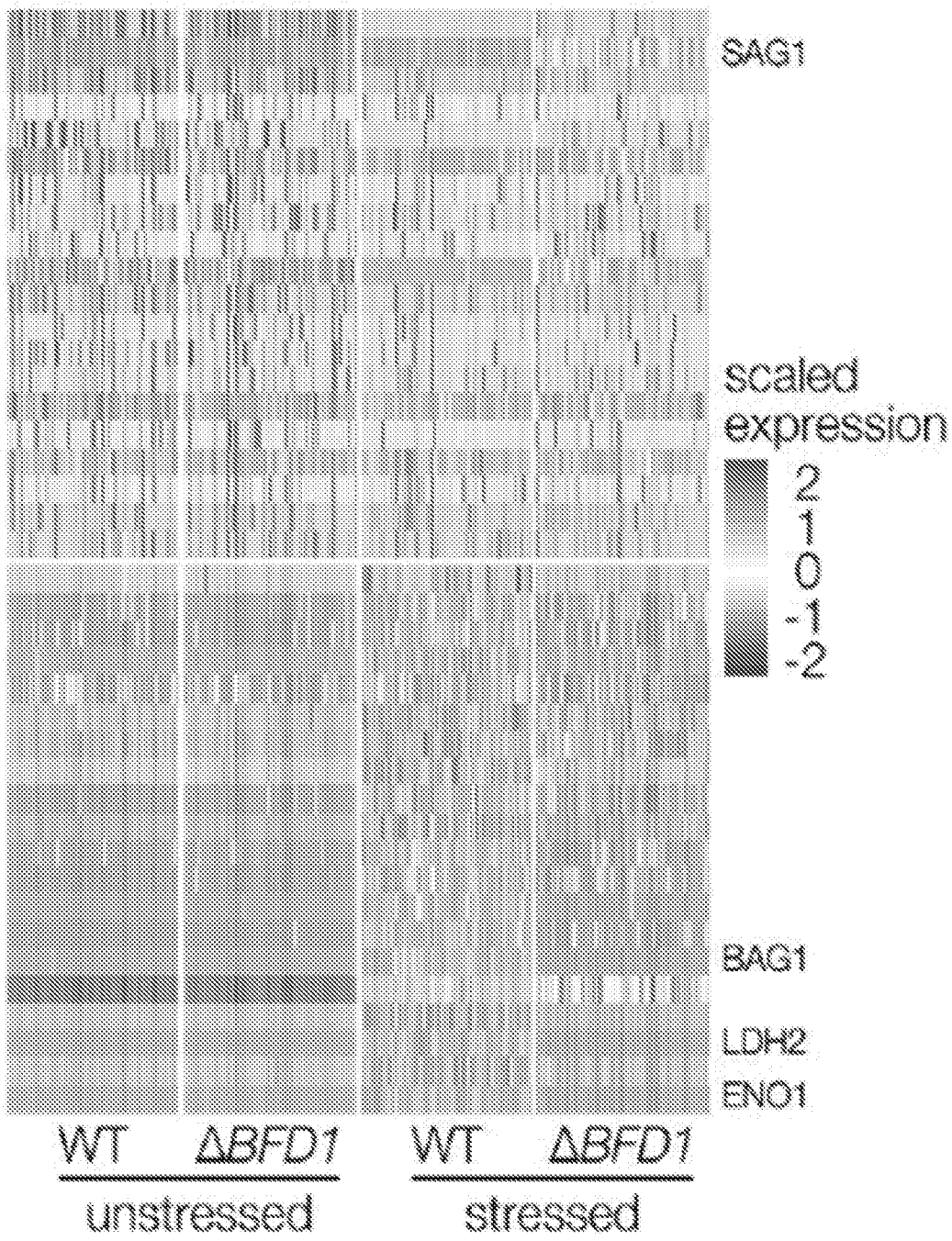
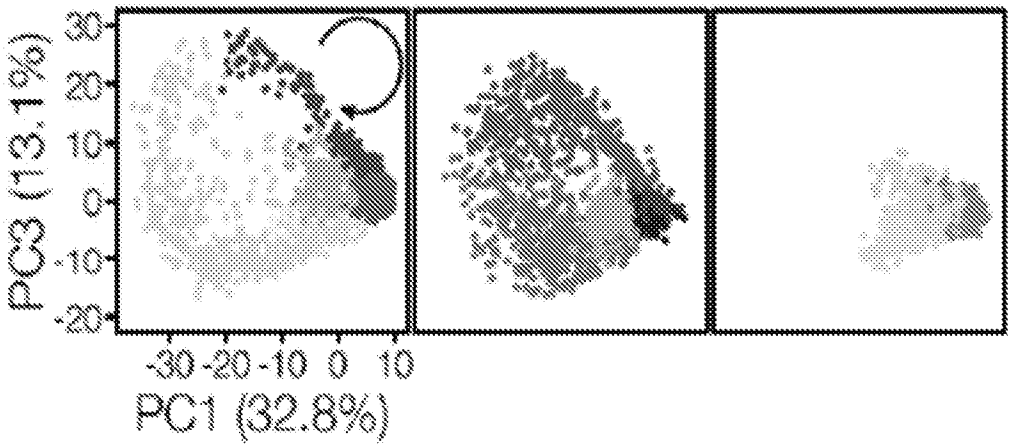

SAG1
scaled expression
2
1
0
-1
-2
BAG1
LDH2
ENO1
WT
ΔBFD1
WT
ΔBFD1
unstressed
stressed

FIG. 3I


PC3 (13.1%)
30
20
10
0
-10
-20
-30 -20 -10 0 10
PC1 (32.8%)

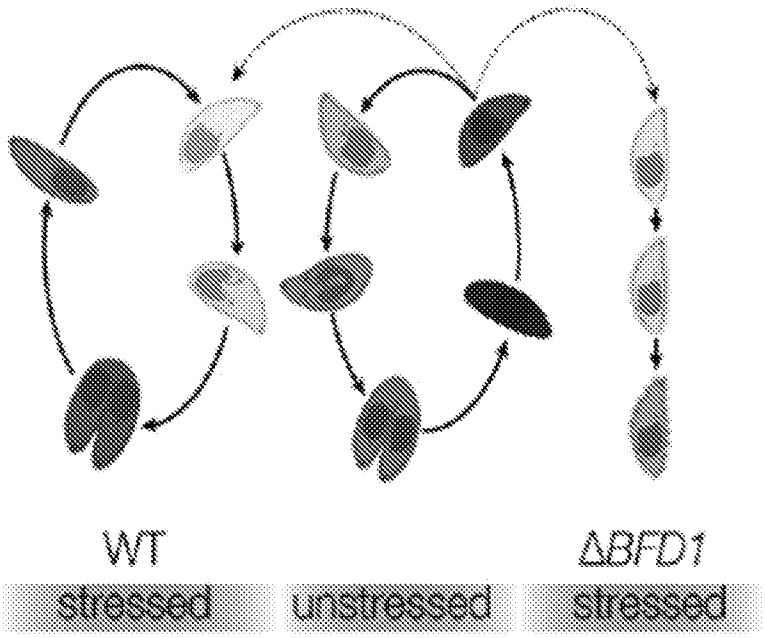

WT
ΔBFD1
stressed
unstressed
stressed

FIG. 3J

FIG. 4A

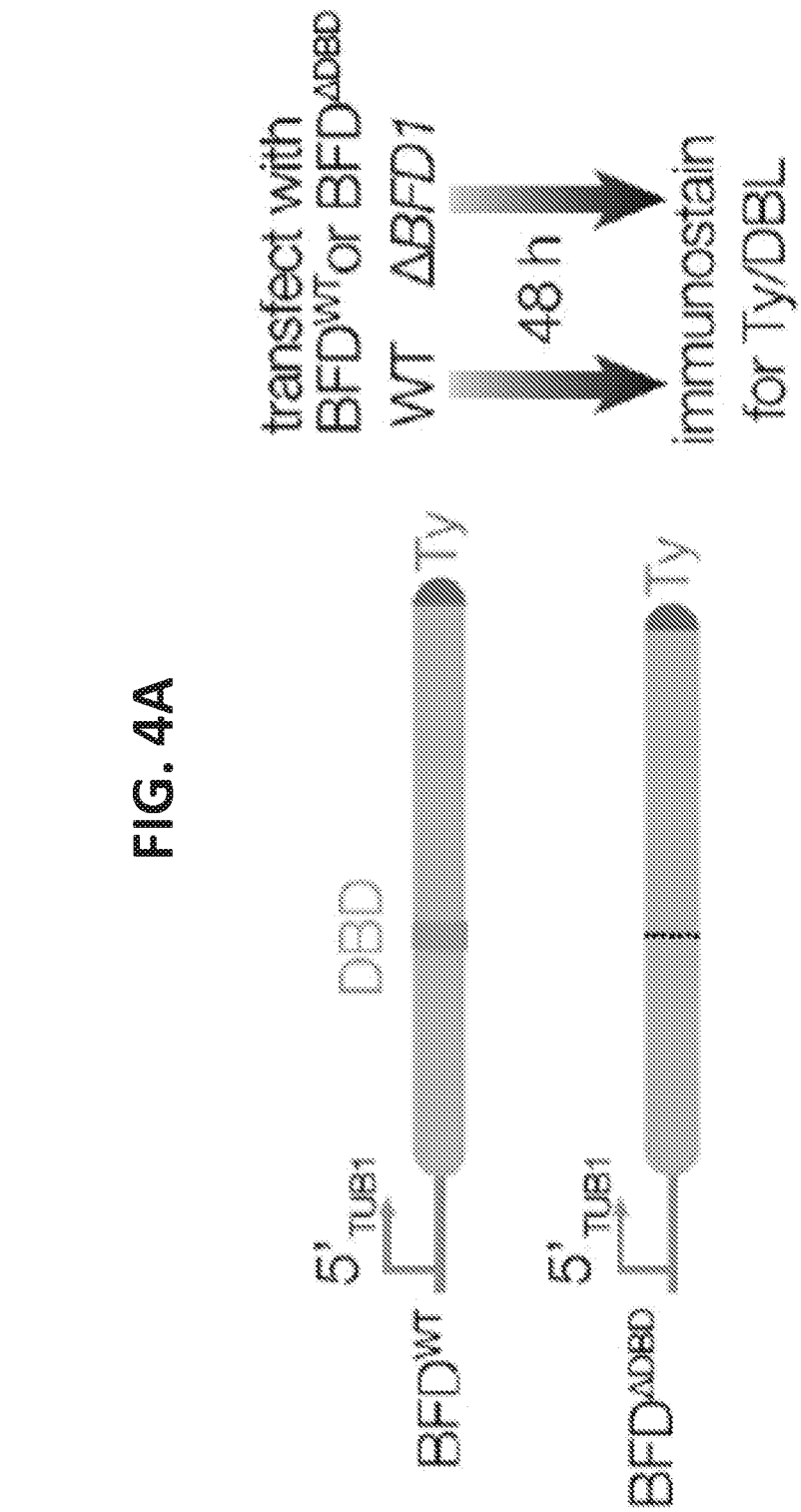

transfect with
BFD^WT or BFD^ΔDBD
WT ΔBFD1
48 h
immunostain
for Ty/DBL
5' TUB1
DBD
Ty
BFD^WT
5' TUB1
Ty
BFD^ΔDBD

FIG. 4B

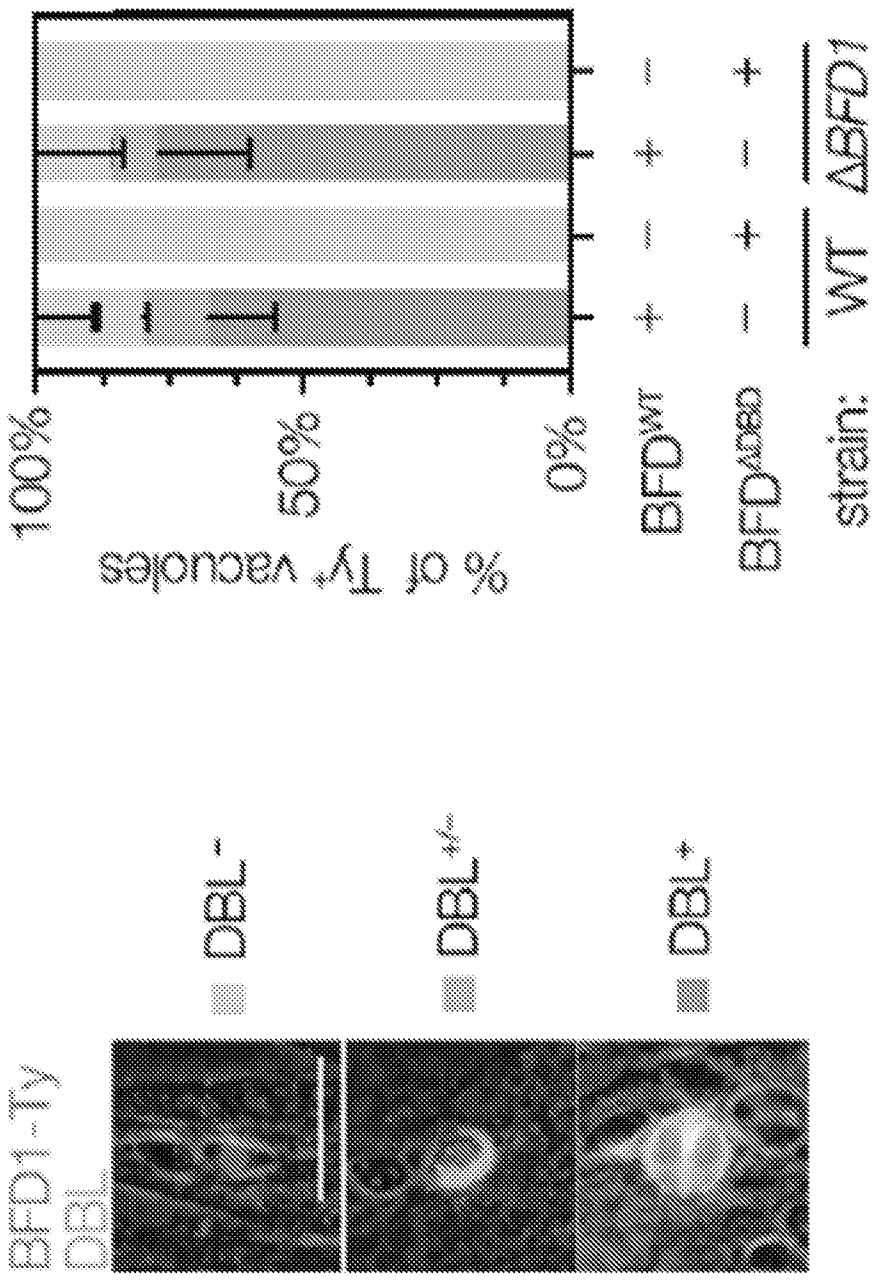

% of Ty+ vacuoles
100%
50%
0%
BFD1-Ty
DBL
DBL-
DBL+/-
DBL+
BFDWT
BFDΔdad
strain:
WT
ΔBFD1

FIG. 5A

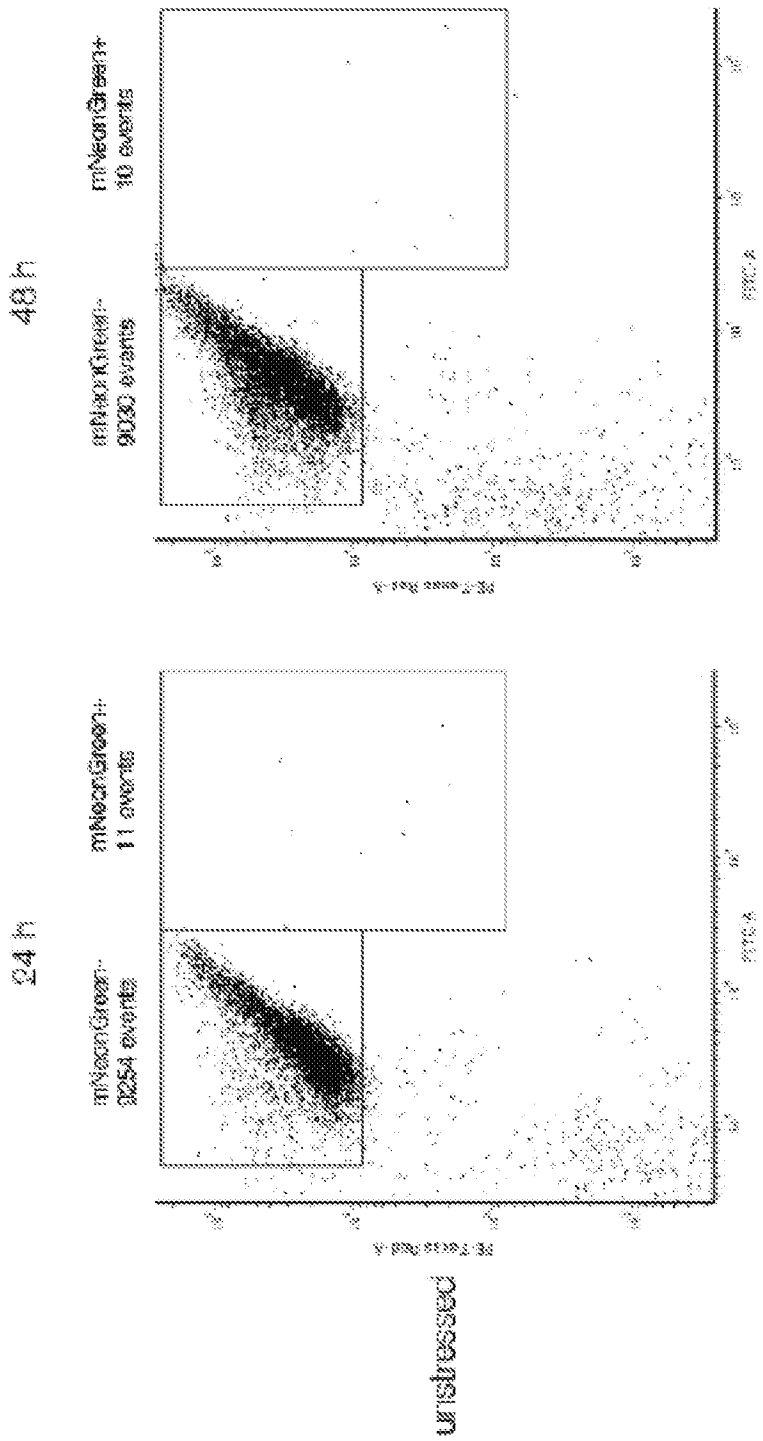

unstressed
24 h
mNeonGreen-
3254 events
mNeonGreen+
11 events
48 h
mNeonGreen-
9020 events
mNeonGreen+
10 events

FIG. 5A cont'd

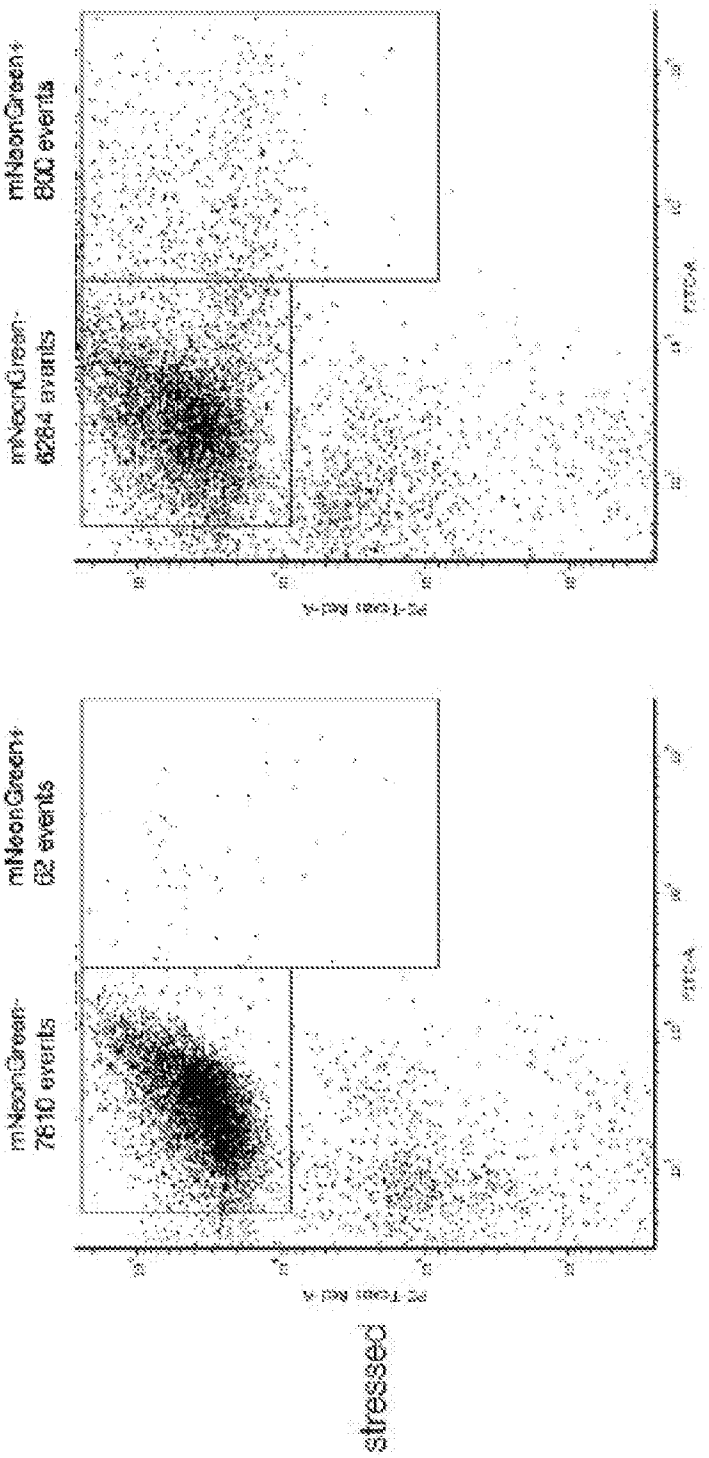

mNeonGreen+
7810 events
mNeonGreen+
62 events
mNeonGreen-
6284 events
mNeonGreen+
800 events
stressed

FIG. 5B

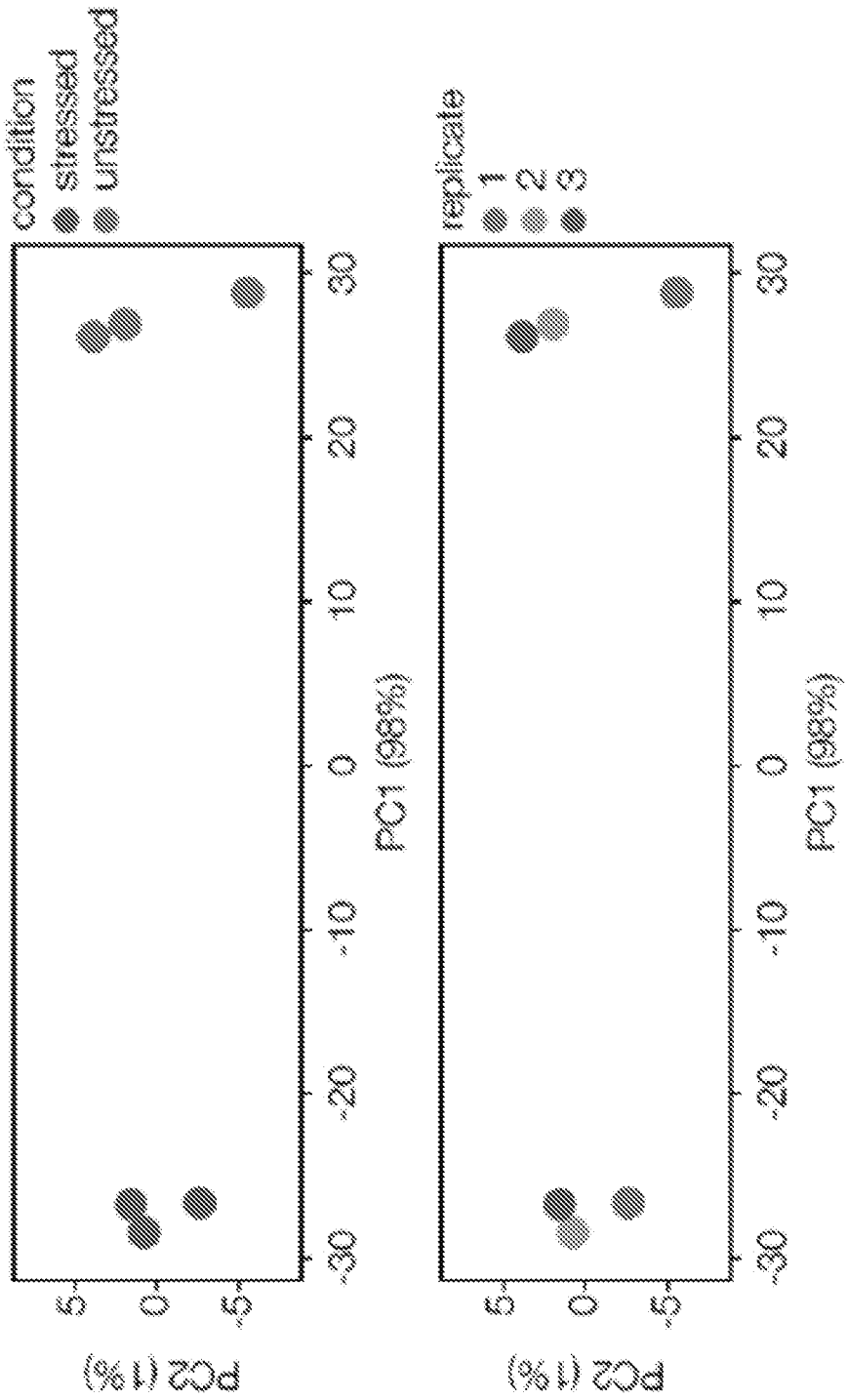

condition
stressed
unstressed
replicate
1
2
3
PC1 (98%)
PC2 (1%)
30
20
10
0
-10
-20
-30
5
0
-5

FIG. 5D

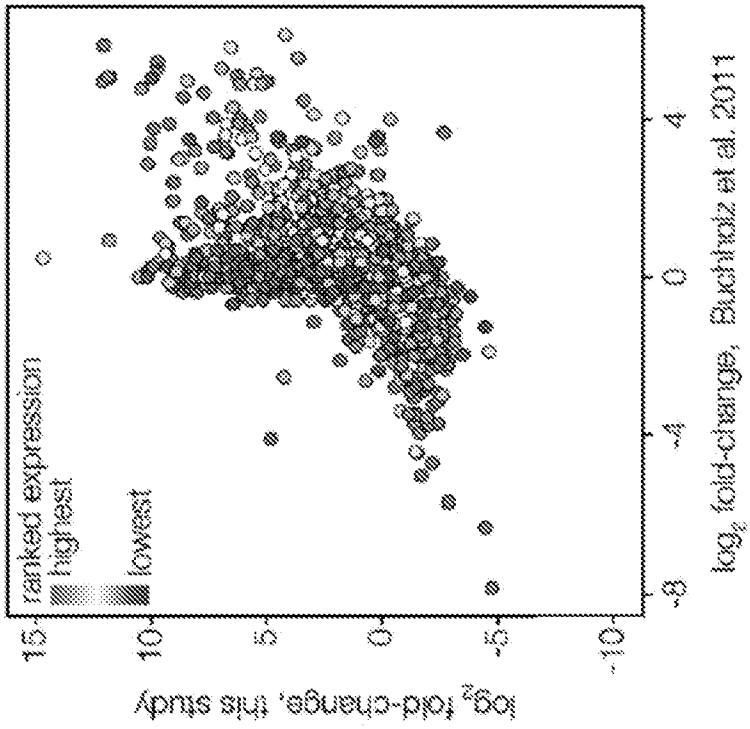

ranked expression
highest
lowest
15
10
5
0
-5
-10
log2 fold-change, this study
4
0
-4
-8
log2 fold-change, Buchholz et al. 2011

FIG. 5C

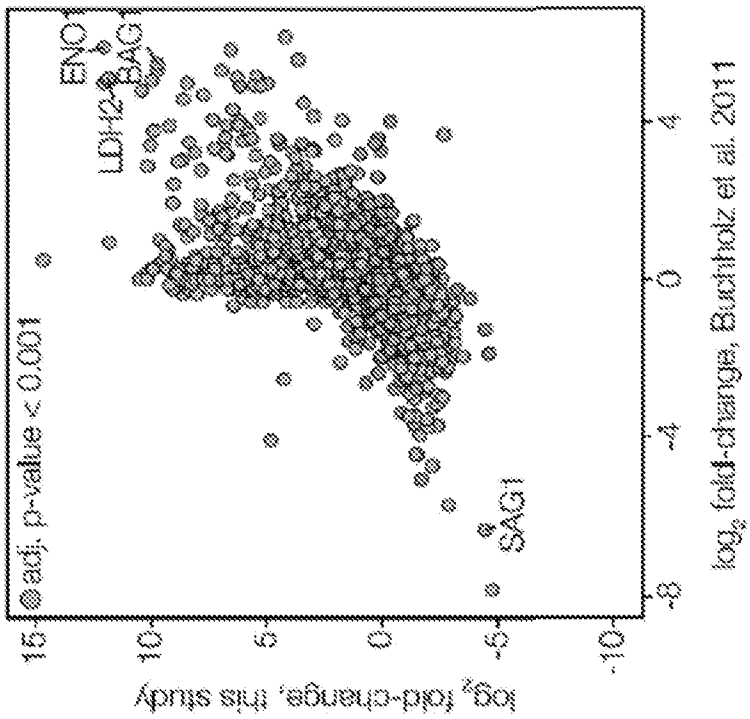

adj. p-value < 0.001
ENO1
LDH-12
BAG1
SAG1
15
10
5
0
-5
-10
log2 fold-change, this study
4
0
-4
-8
log2 fold-change, Buchholz et al. 2011

FIG. 6

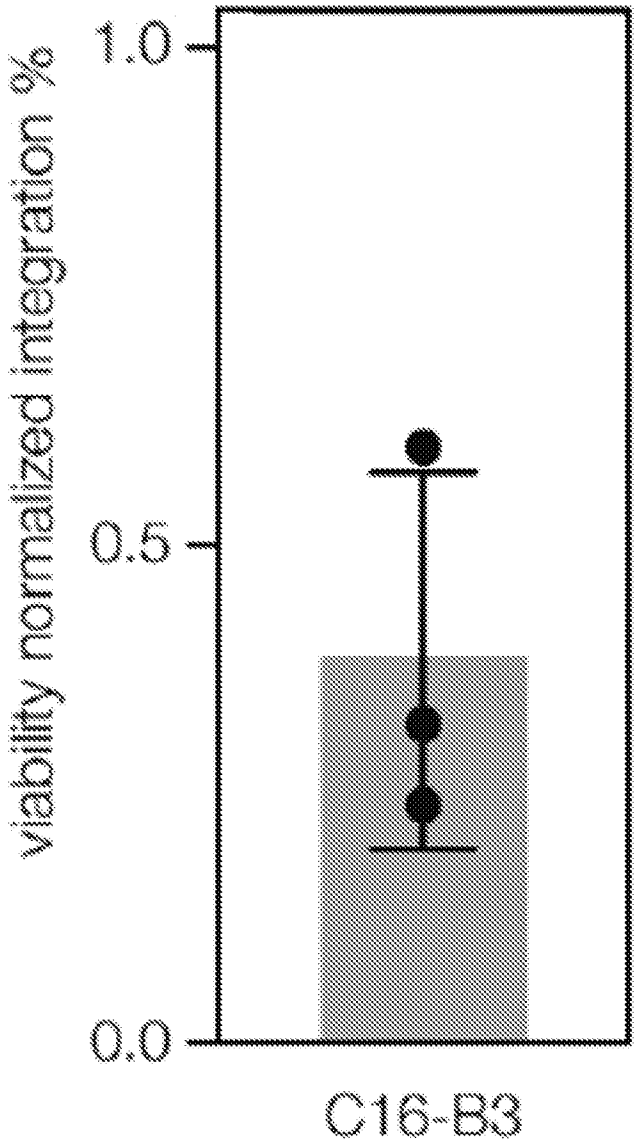


FIG. 7

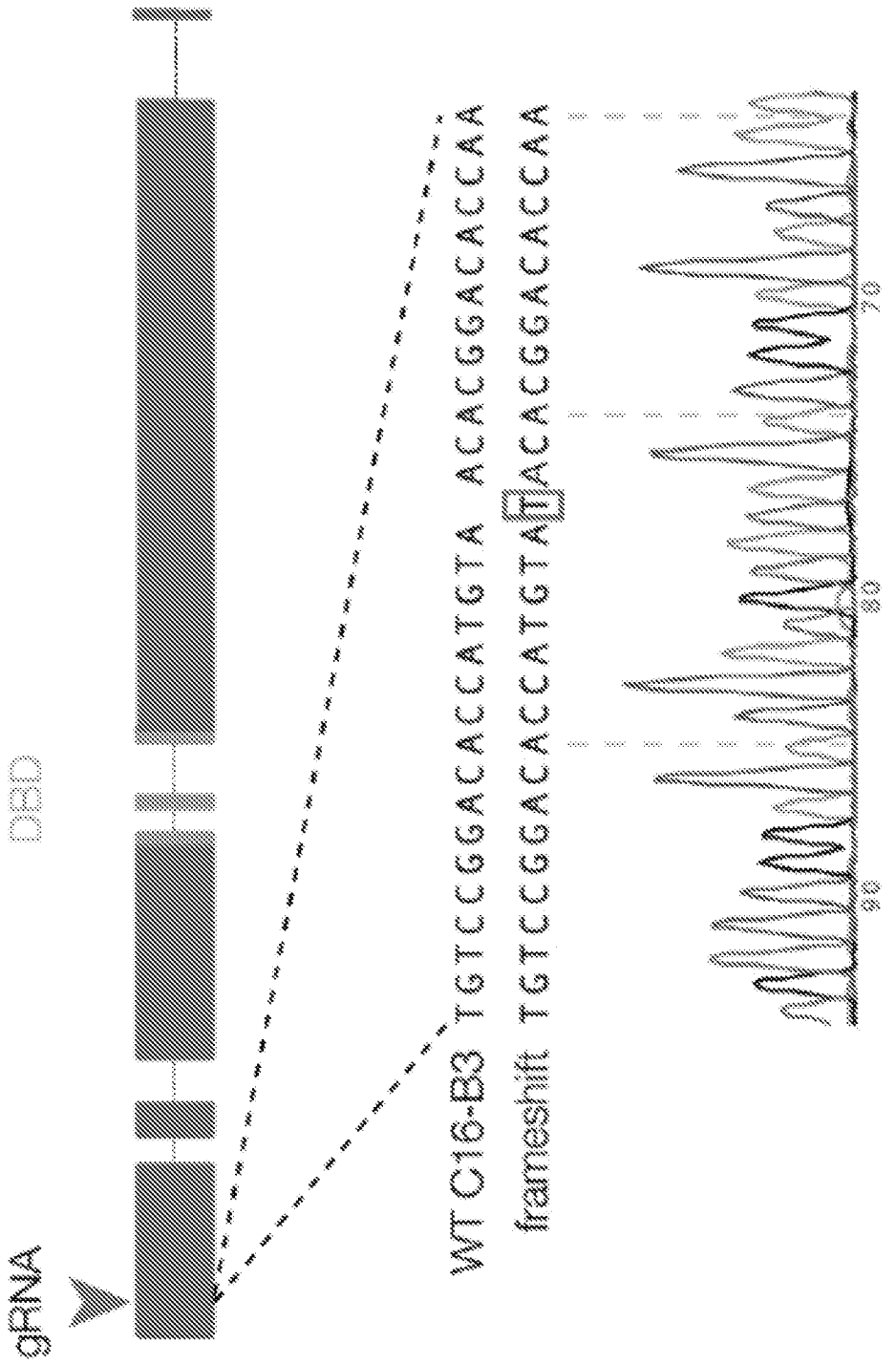
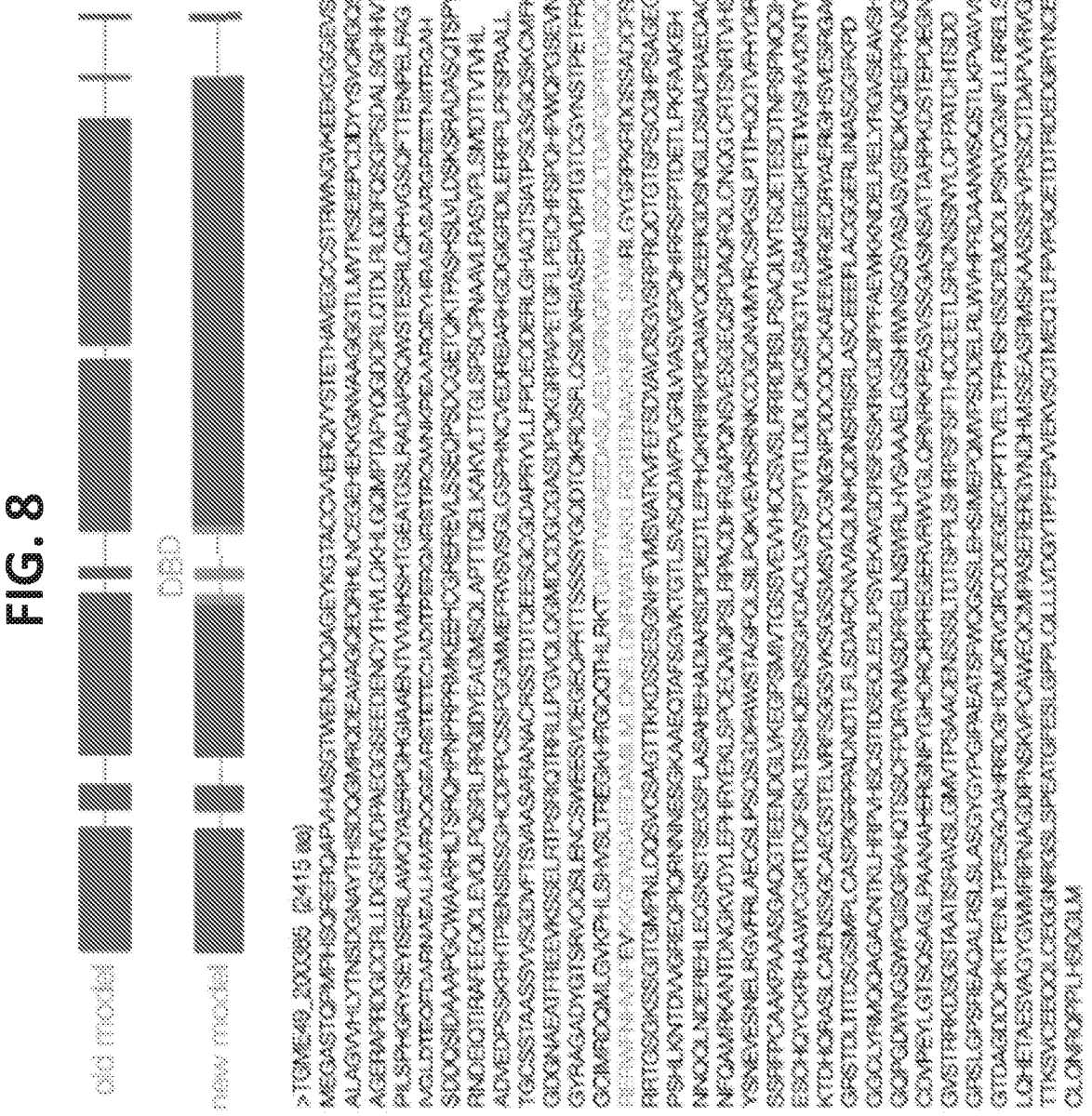

DBD
gRNA
WT C16-B3 TGTCCGGACACCATGTA ACACGGACACCAA
frameshift TGTCCGGACACCATGTATACACGGACACCAA
70
80
90

FIG. 8



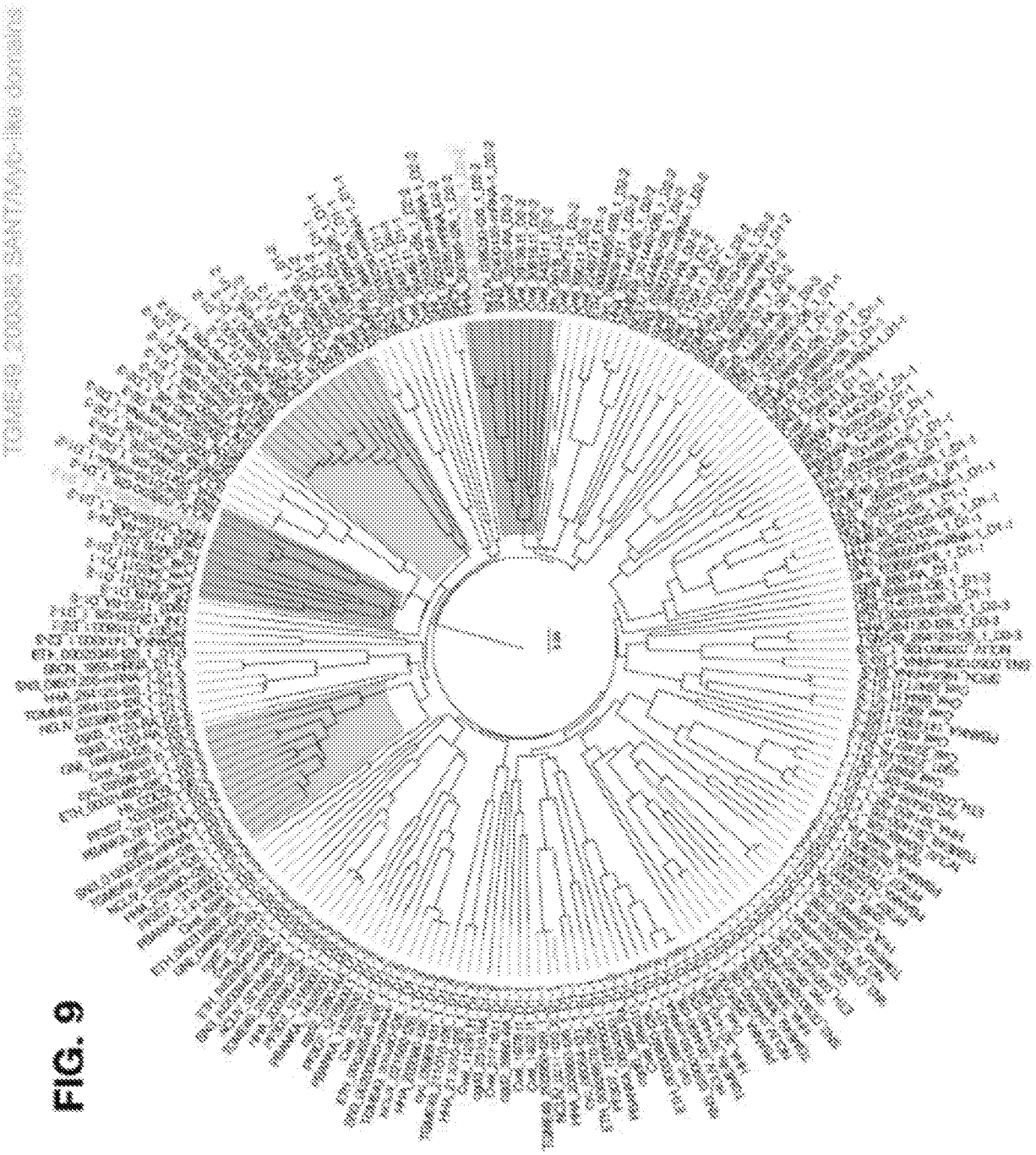


FIG. 9

FIG. 10

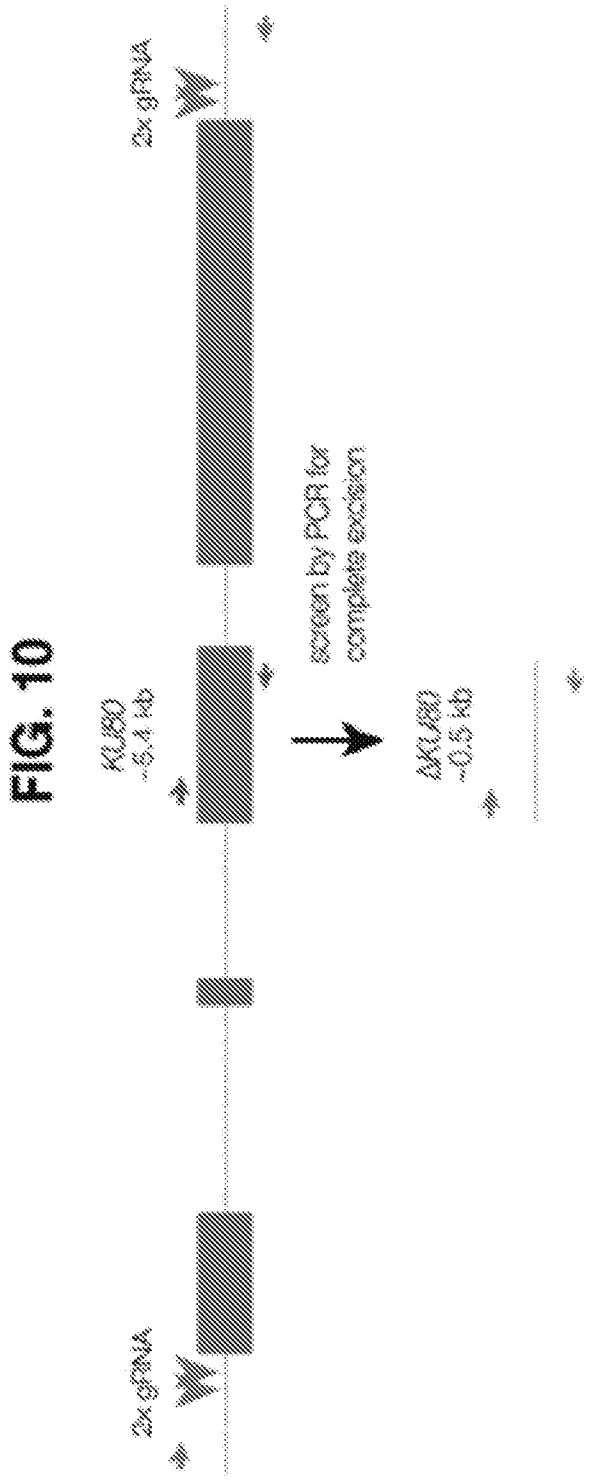

2x gRNA
KU80
~5.4 kb
screen by PCR for
complete excision
ΔKU80
~0.5 kb
2x gRNA

FIG. 11A

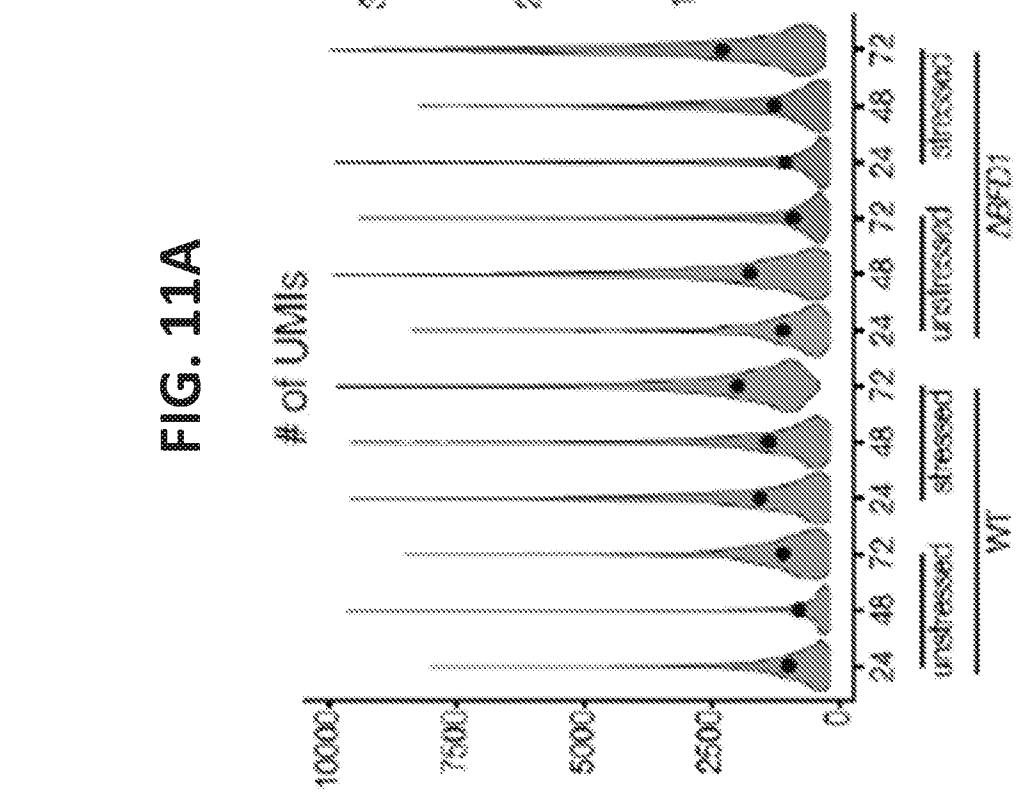


FIG. 11B

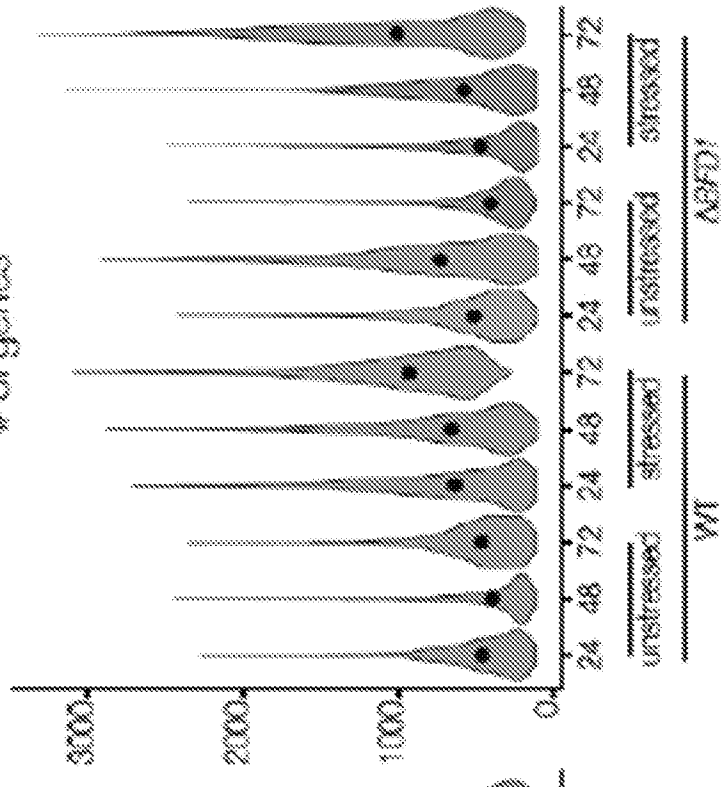


FIG. 11C

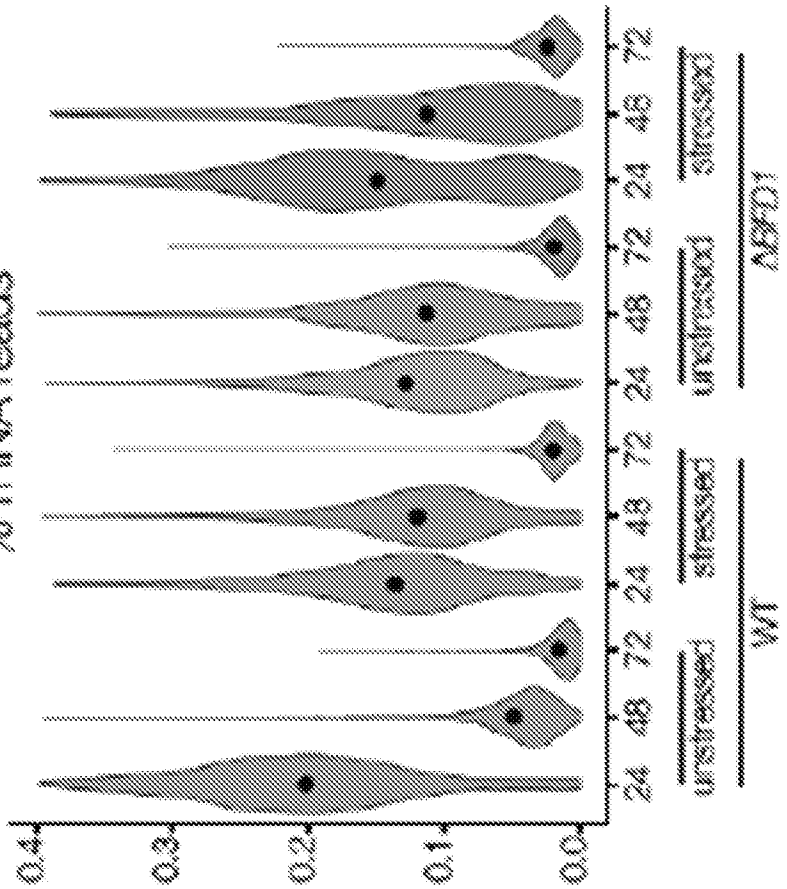

% rRNA reads
0.4
0.3
0.2
0.1
0.0
24
48
72
24
48
72
24
48
72
24
48
72
unstressed
stressed
unstressed
stressed
WT
ΔDFD1

FIG. 12A

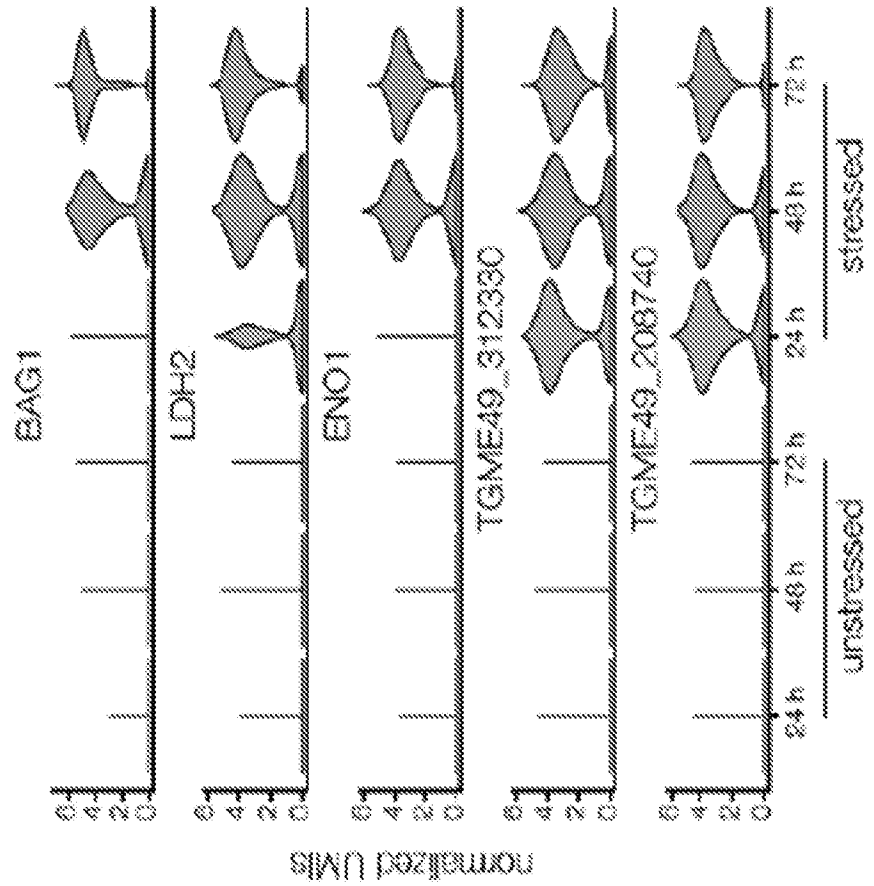

BAG1
LDH2
ENO1
TGME49_312330
TGME49_208740
normalized UMIs
0
2
4
6
24 h
48 h
72 h
unstressed
stressed

FIG. 12B

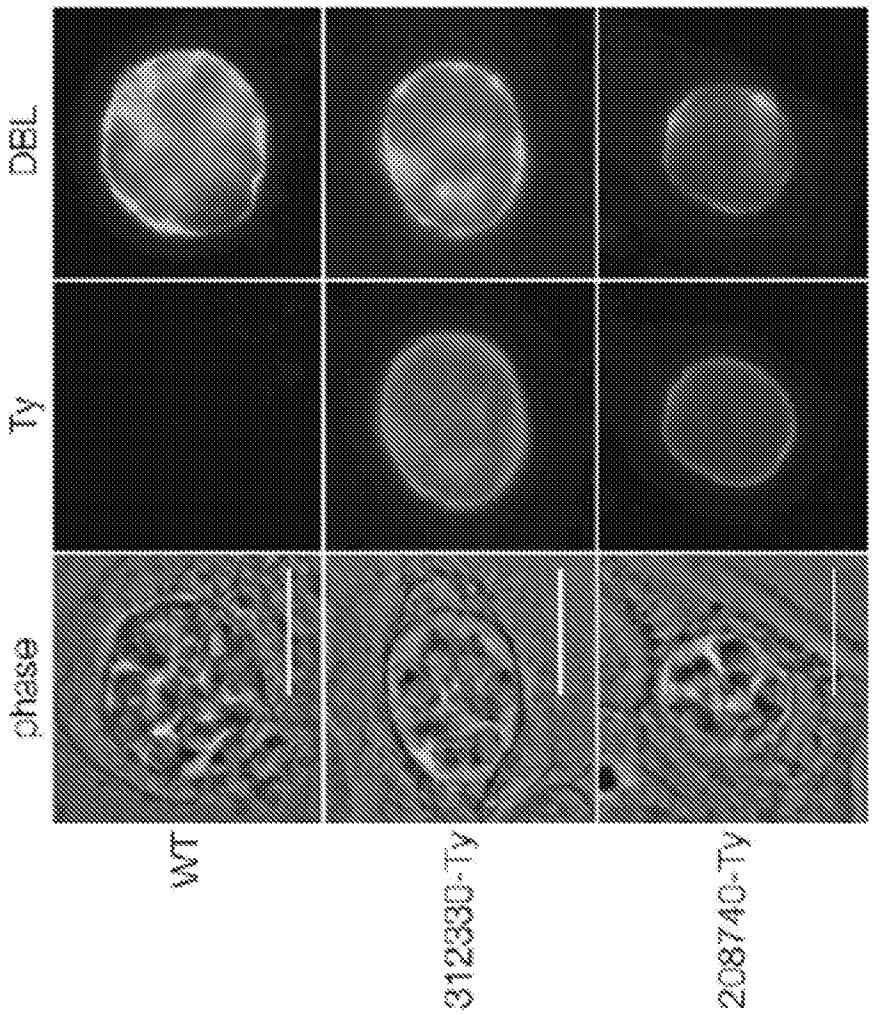


FIG. 13A

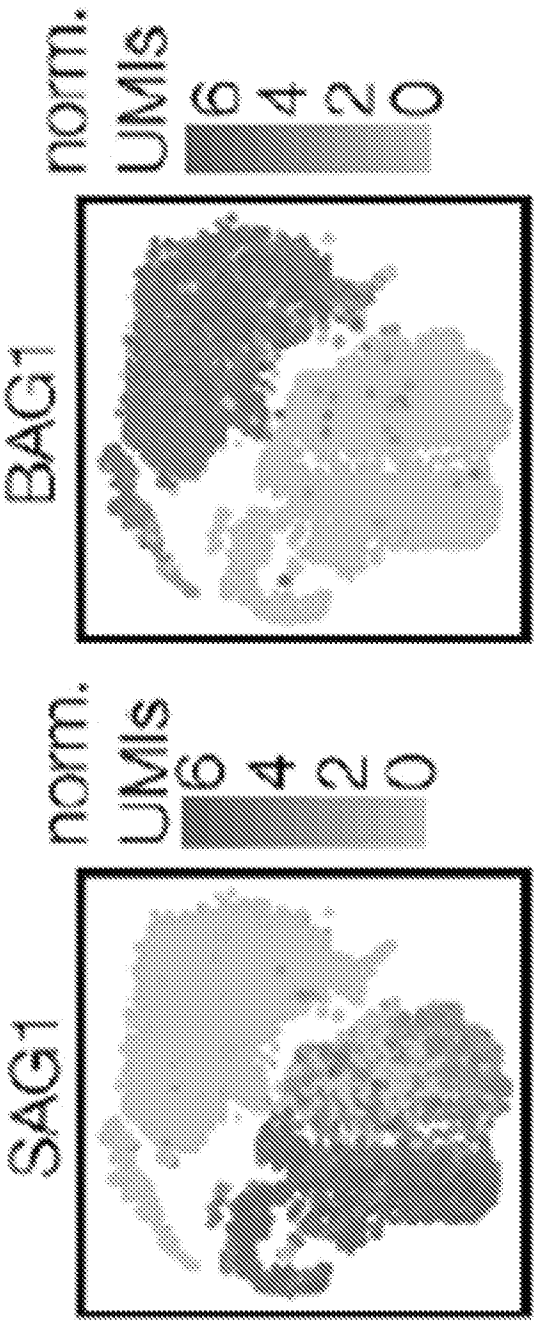


FIG. 13C

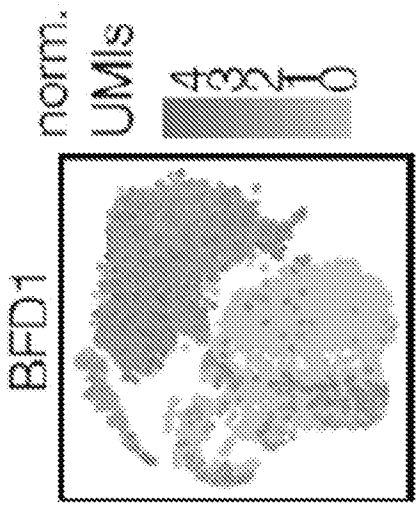


FIG. 13B

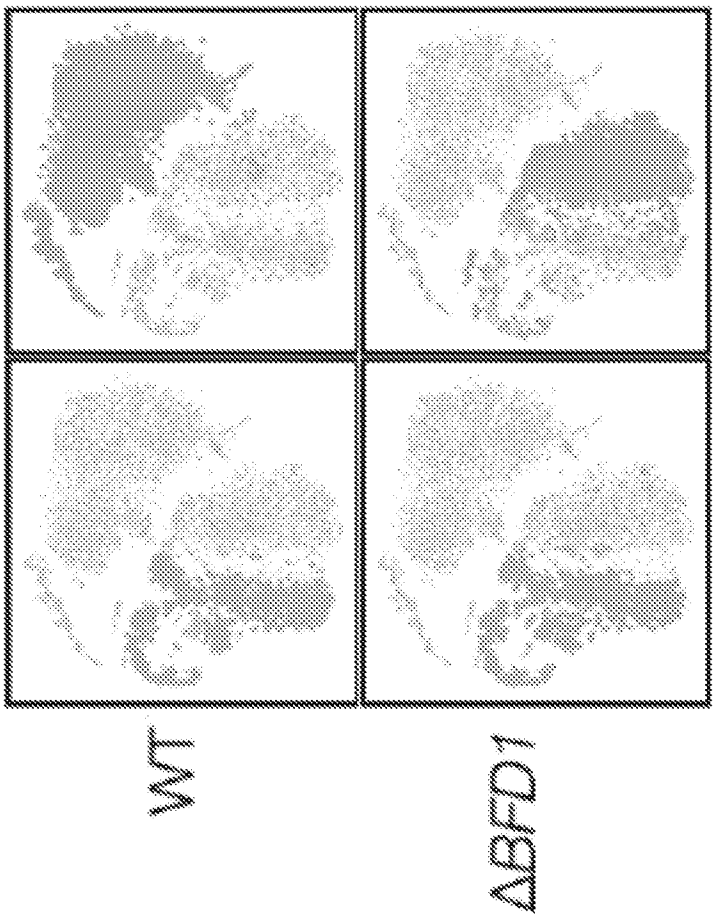


FIG. 14A

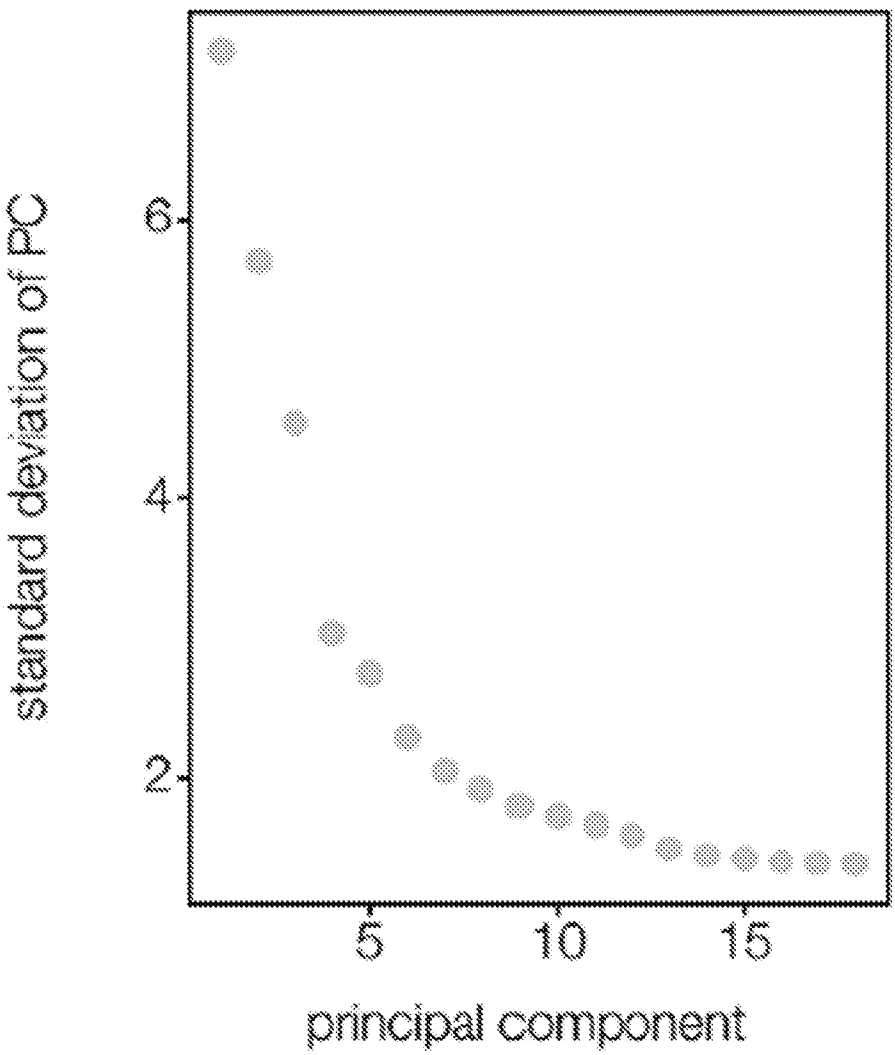


FIG. 14B

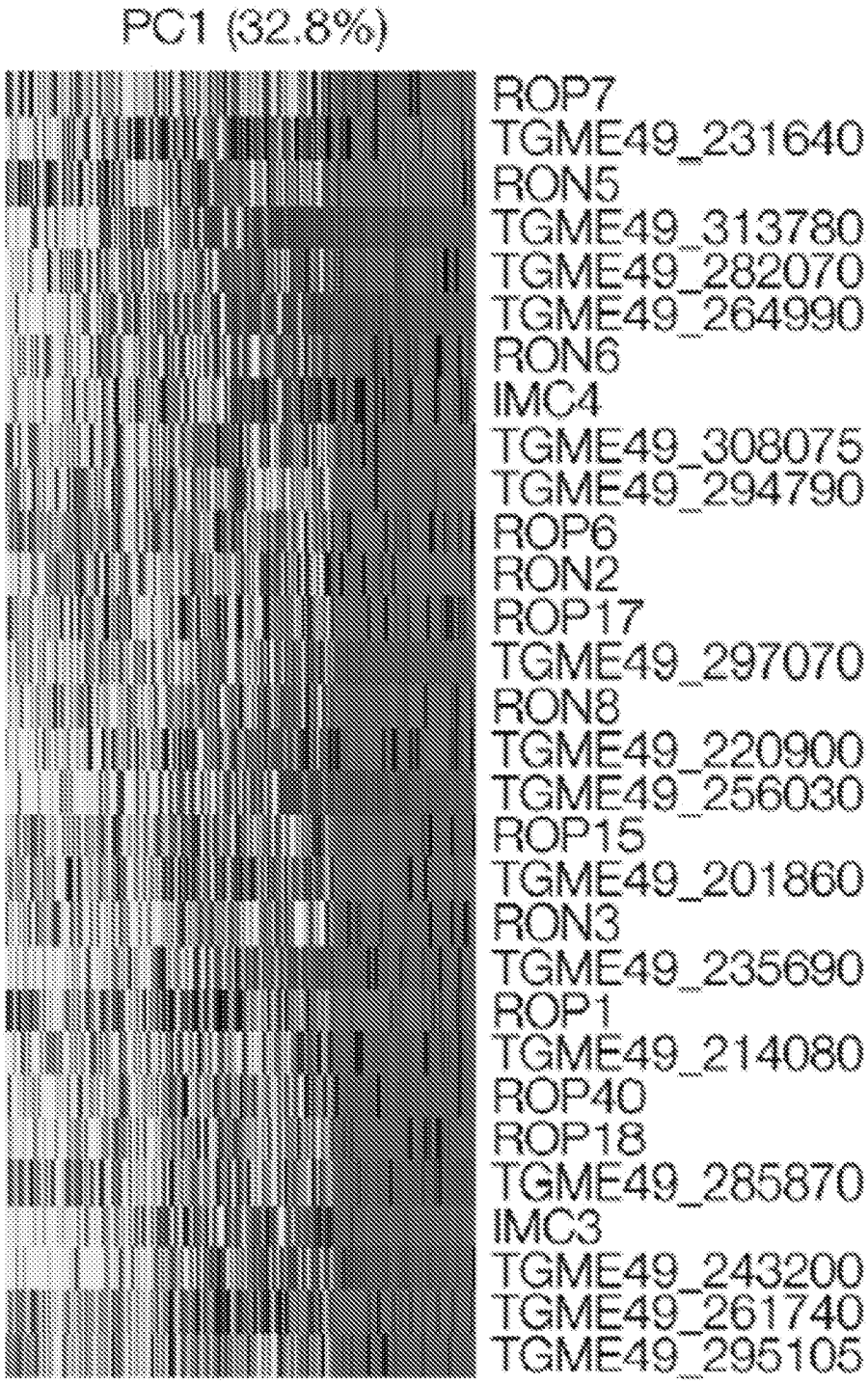

PC1 (32.8%)
ROP7
TGME49_231640
RON5
TGME49_313780
TGME49_282070
TGME49_264990
RON6
IMC4
TGME49_308075
TGME49_294790
ROP6
RON2
ROP17
TGME49_297070
RON8
TGME49_220900
TGME49_256030
ROP15
TGME49_201860
RON3
TGME49_235690
ROP1
TGME49_214080
ROP40
ROP18
TGME49_285870
IMC3
TGME49_243200
TGME49_261740
TGME49_295105

FIG. 14C

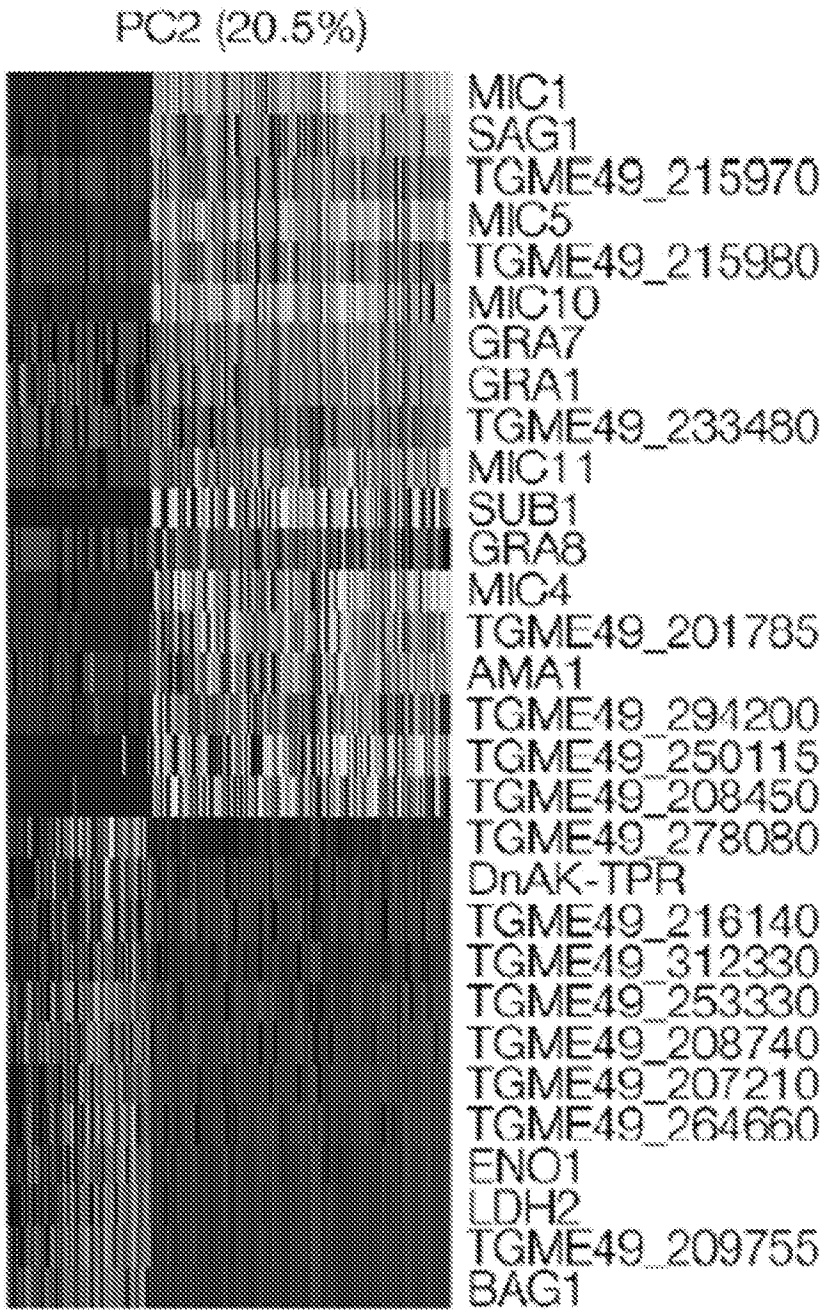

PC2 (20.5%)
MIC1
SAG1
TGME49_215970
MIC5
TGME49_215980
MIC10
GRA7
GRA1
TGME49_233480
MIC11
SUB1
GRA8
MIC4
TGME49_201785
AMA1
TGME49_291200
TGME49_250115
TGME49_208450
TGME49_278080
DnAK-TPR
TGME49_216140
TGME49_312330
TGME49_253330
TGME49_208740
TGME49_207210
TGME49_264660
ENO1
LDH2
TGME49_209755
BAG1

FIG. 14D

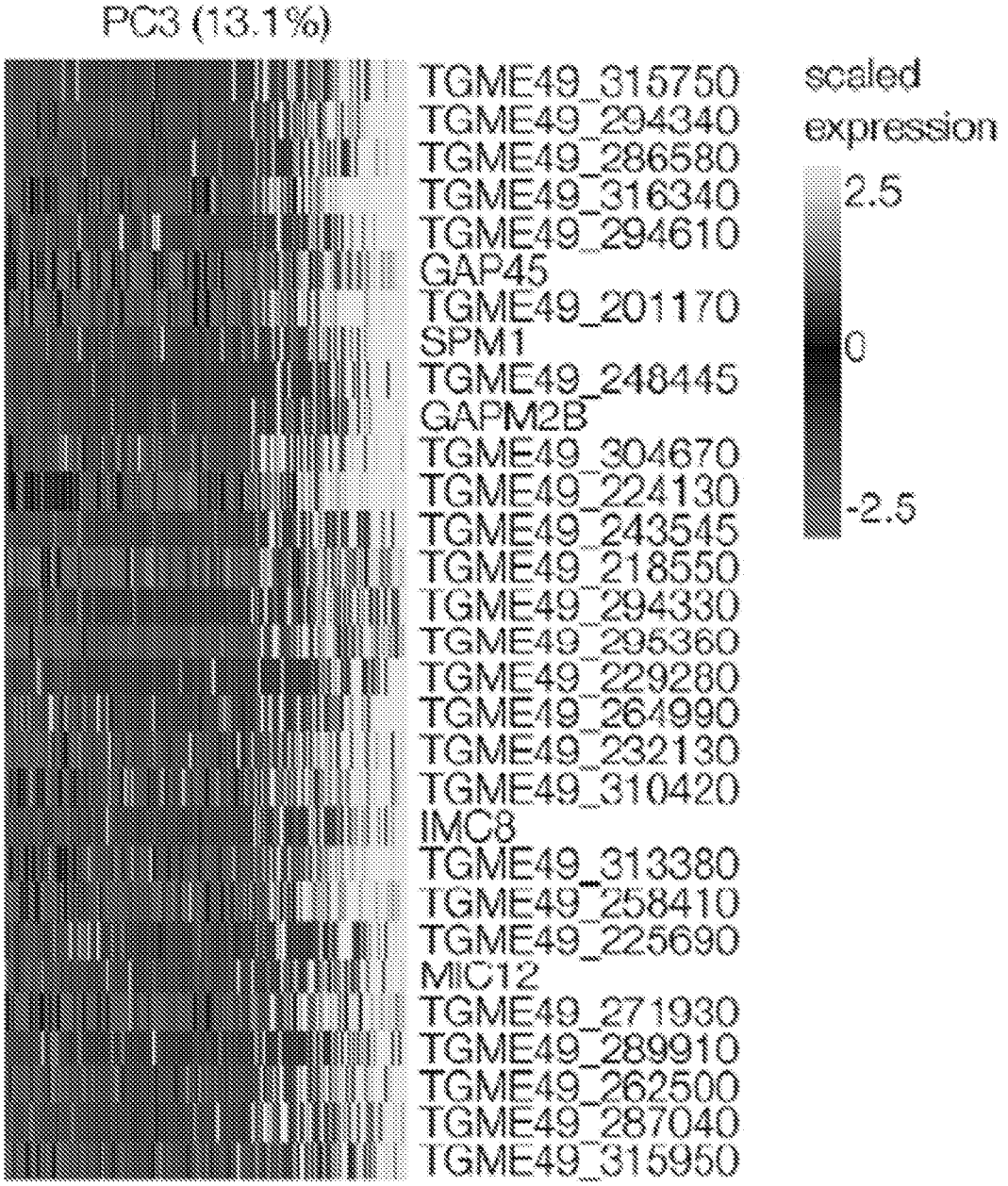

PC3 (13.1%)
TGME49_315750
TGME49_294340
TGME49_286580
TGME49_316340
TGME49_294610
GAP45
TGME49_201170
SPM1
TGME49_248445
GAPM2B
TGME49_304670
TGME49_224130
TGME49_243545
TGME49_218550
TGME49_294330
TGME49_295360
TGME49_229280
TGME49_264990
TGME49_232130
TGME49_310420
IMC8
TGME49_313380
TGME49_258410
TGME49_225690
MIC12
TGME49_271930
TGME49_289910
TGME49_262500
TGME49_287040
TGME49_315950
scaled expression
2.5
0
-2.5

FIG. 15

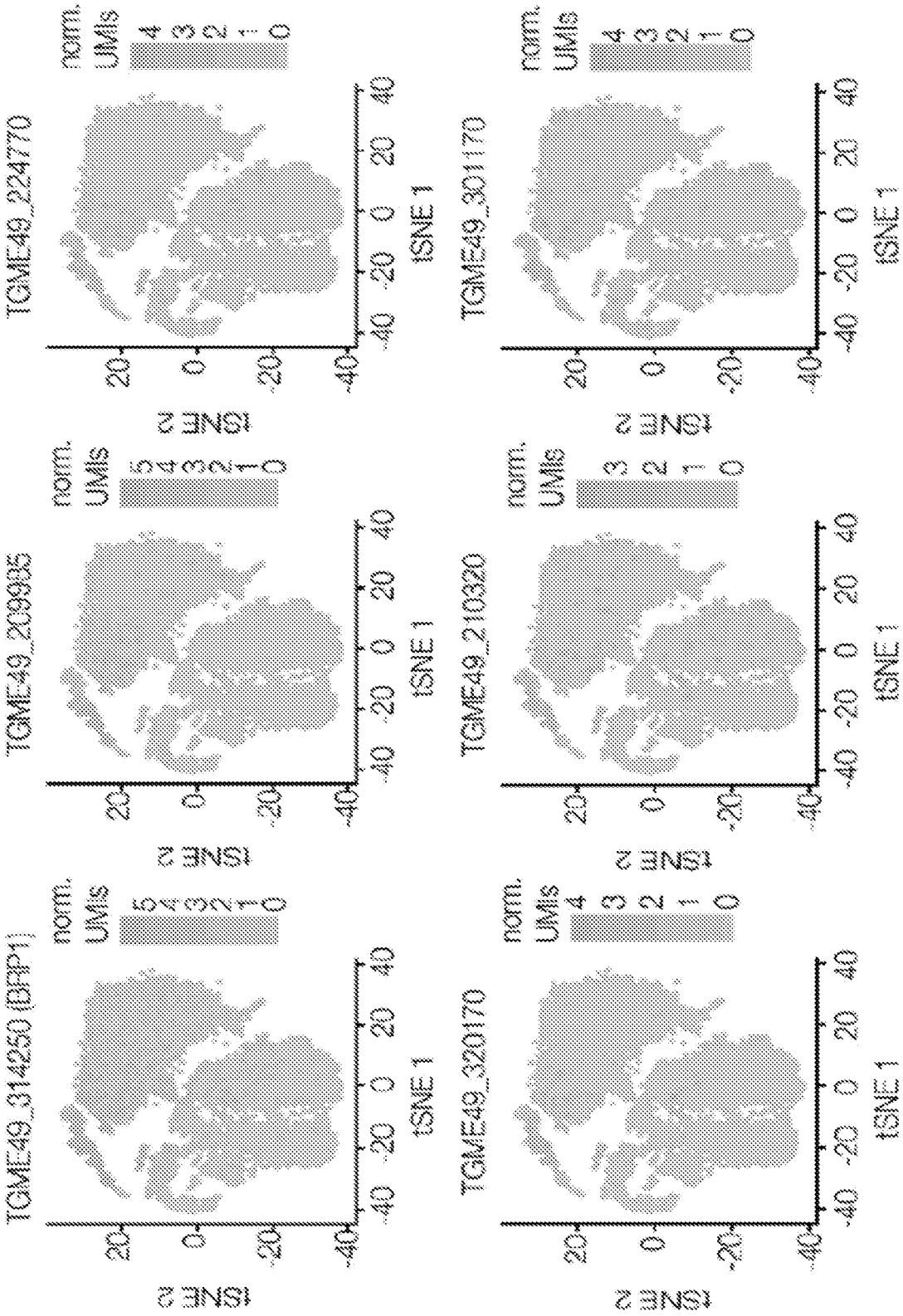

TGME49_314250 (BFP1)
TGME49_209935
TGME49_224770
TGME49_210320
TGME49_320170
TGME49_301170
norm. UMIs
tSNE 1
tSNE 2

FIG. 16

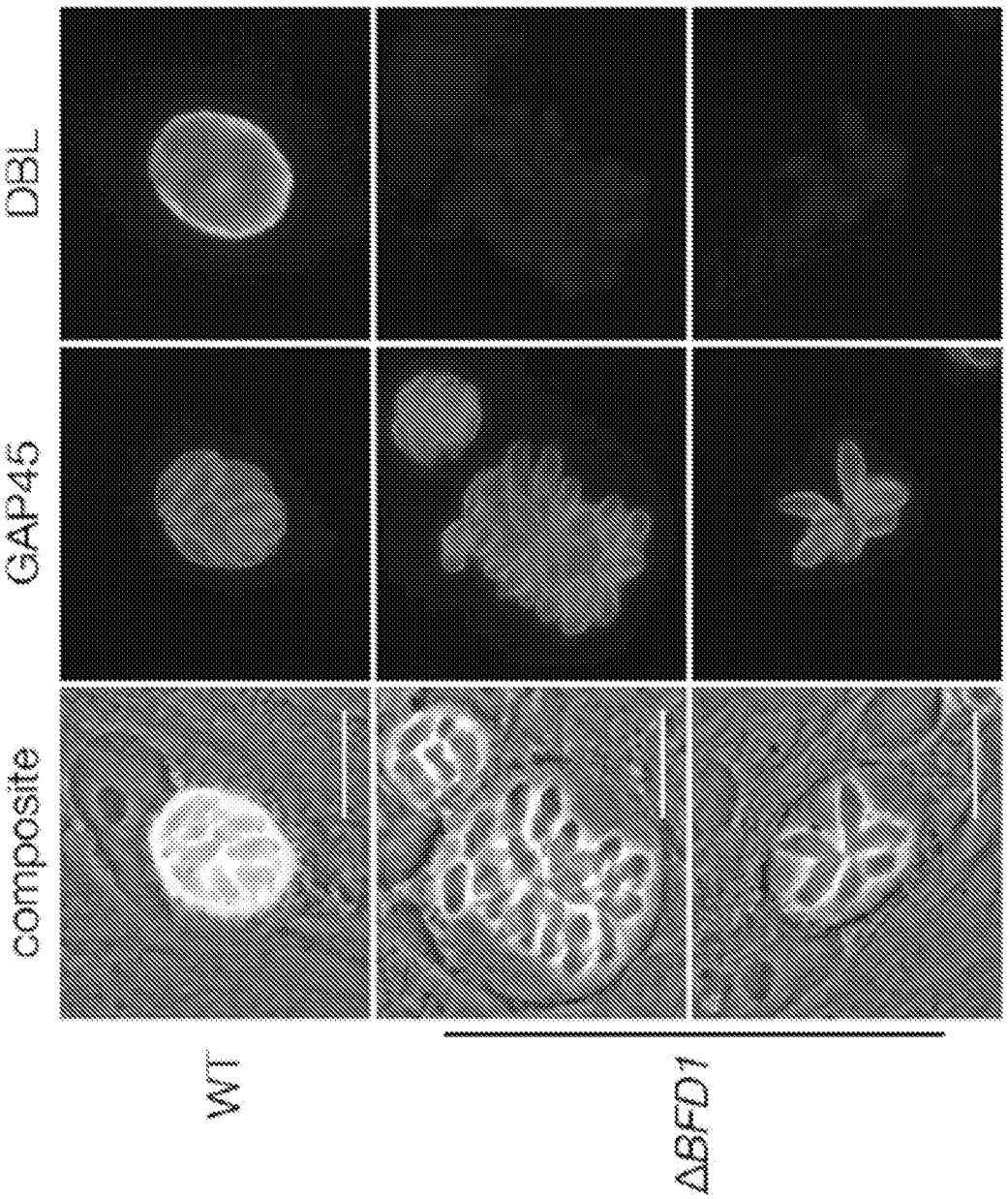
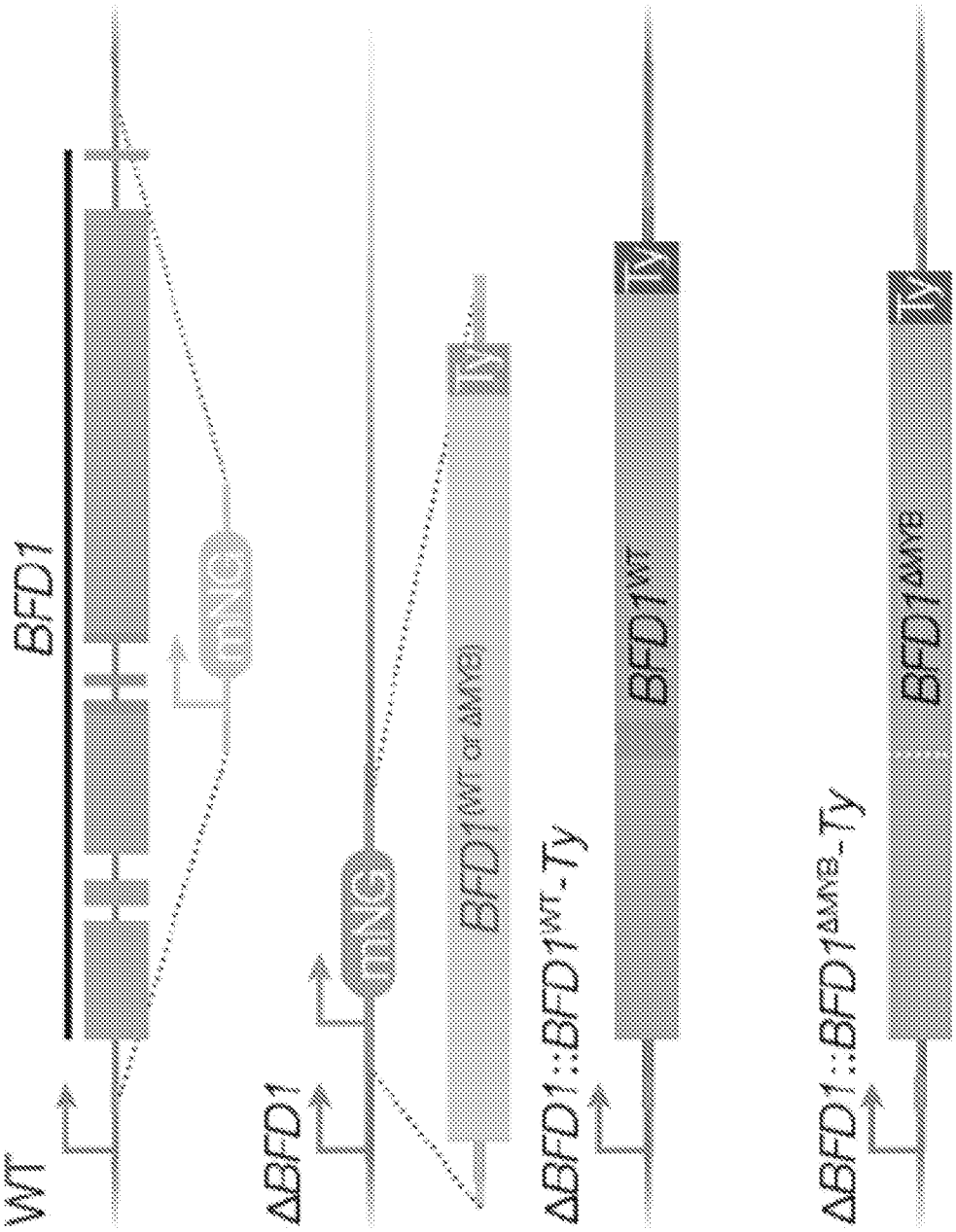


FIG. 17A


WT
BFD1
UTR
ΔBFD1
UTR
BFD1WT or ΔMYB
TY
ΔBFD1::BFD1WT-Ty
BFD1WT
TY
ΔBFD1::BFD1ΔMYB-Ty
BFD1ΔMYB
TY

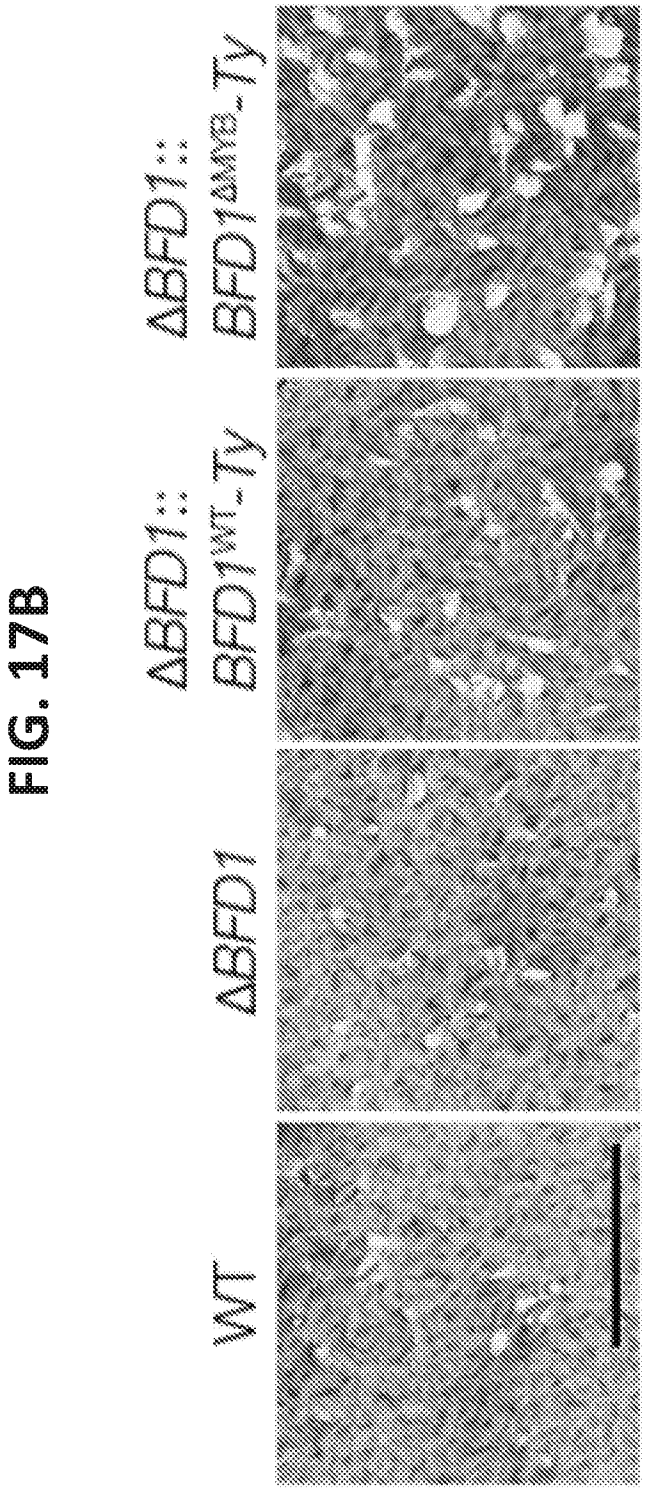
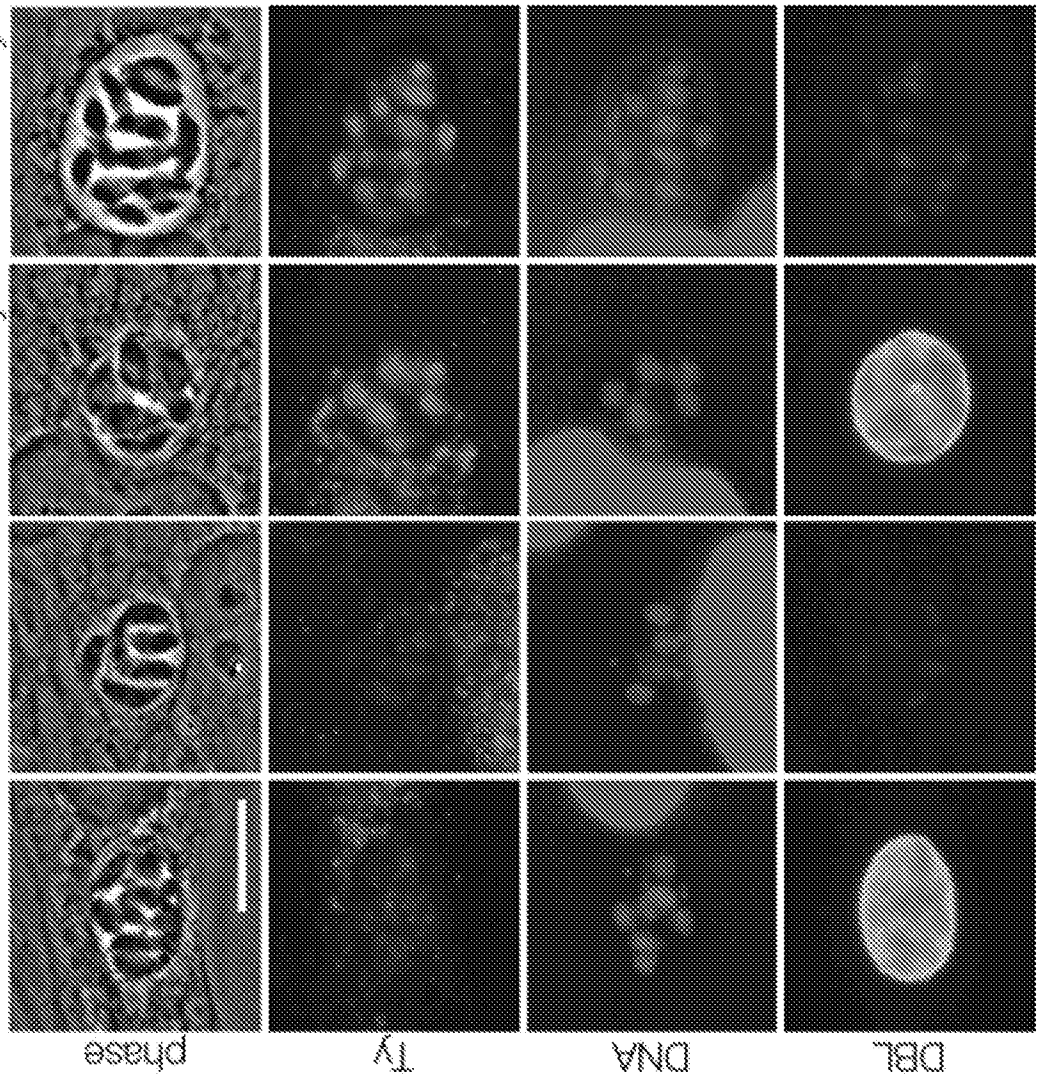


FIG. 17B

FIG. 17C



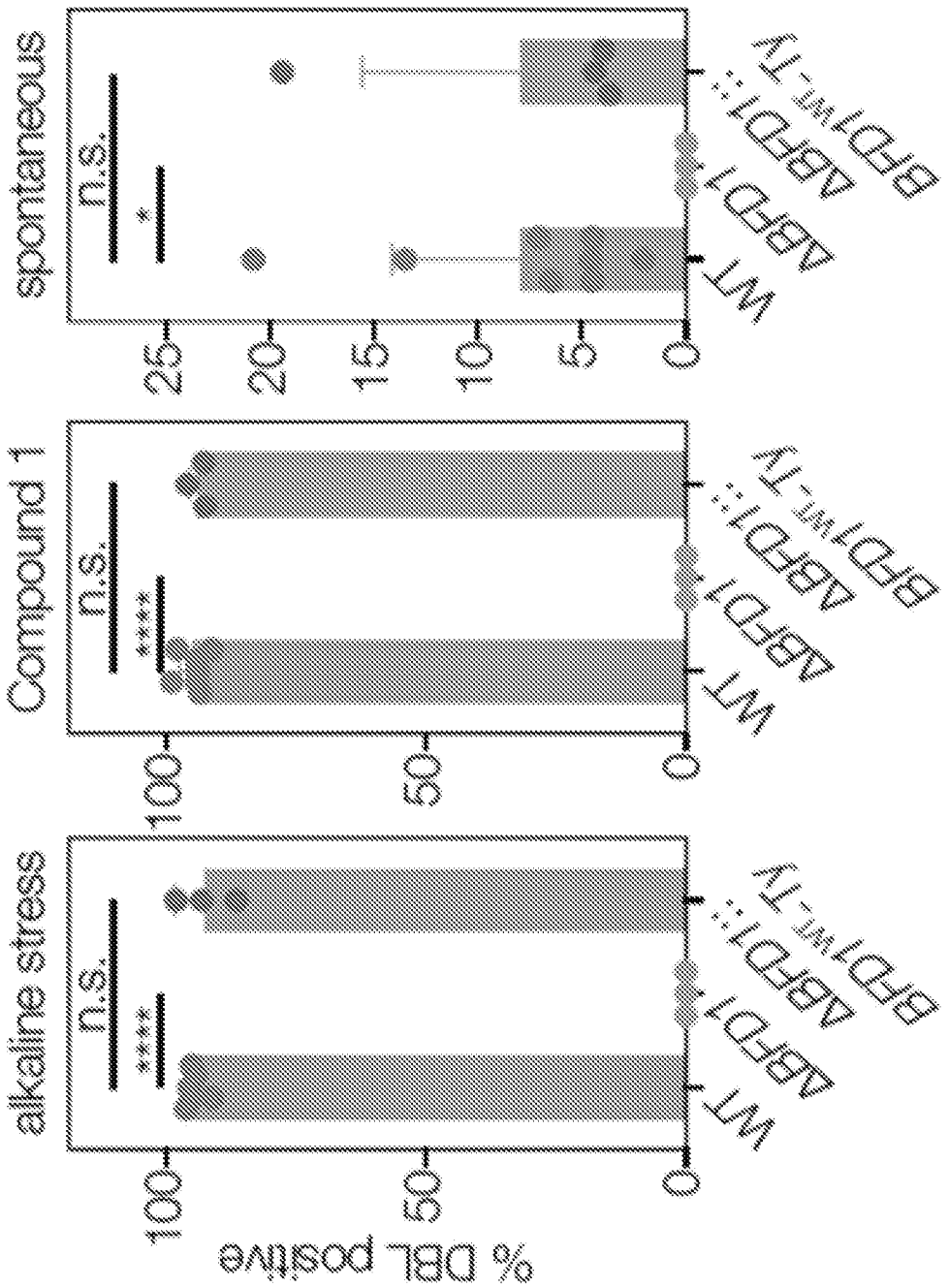

alkaline stress
Compound 1
spontaneous
% DBL positive
n.s.
****
n.s.
****
n.s.
*
100
50
0
100
50
0
25
20
15
10
5
0
WT
ΔBFD1
ΔBFD1::BFD1WT-TV
WT
ΔBFD1
ΔBFD1::BFD1WT-TV
WT
ΔBFD1
ΔBFD1::BFD1WT-TV

FIG. 17D

FIG. 18A

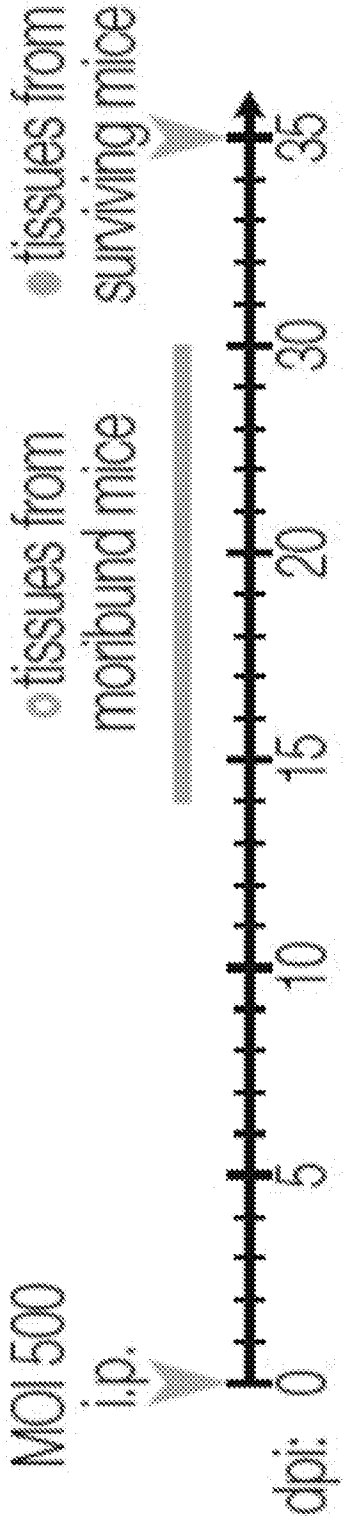
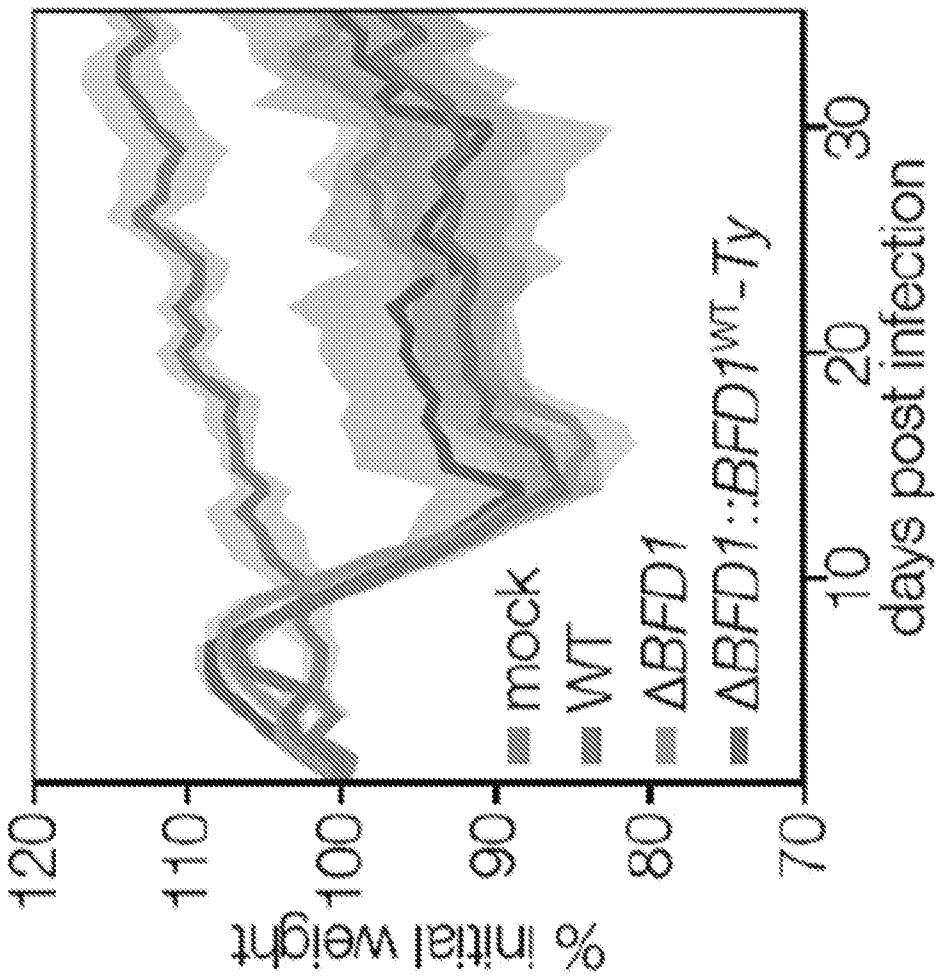

MOI 500
i.p.
tissues from moribund mice
tissues from surviving mice
dpi: 0
5
10
15
20
30
35

FIG. 18B


% initial weight
120
110
100
90
80
70
10
20
30
days post infection
mock
WT
ΔBFD1
ΔBFD1::BFD1WT_Ty

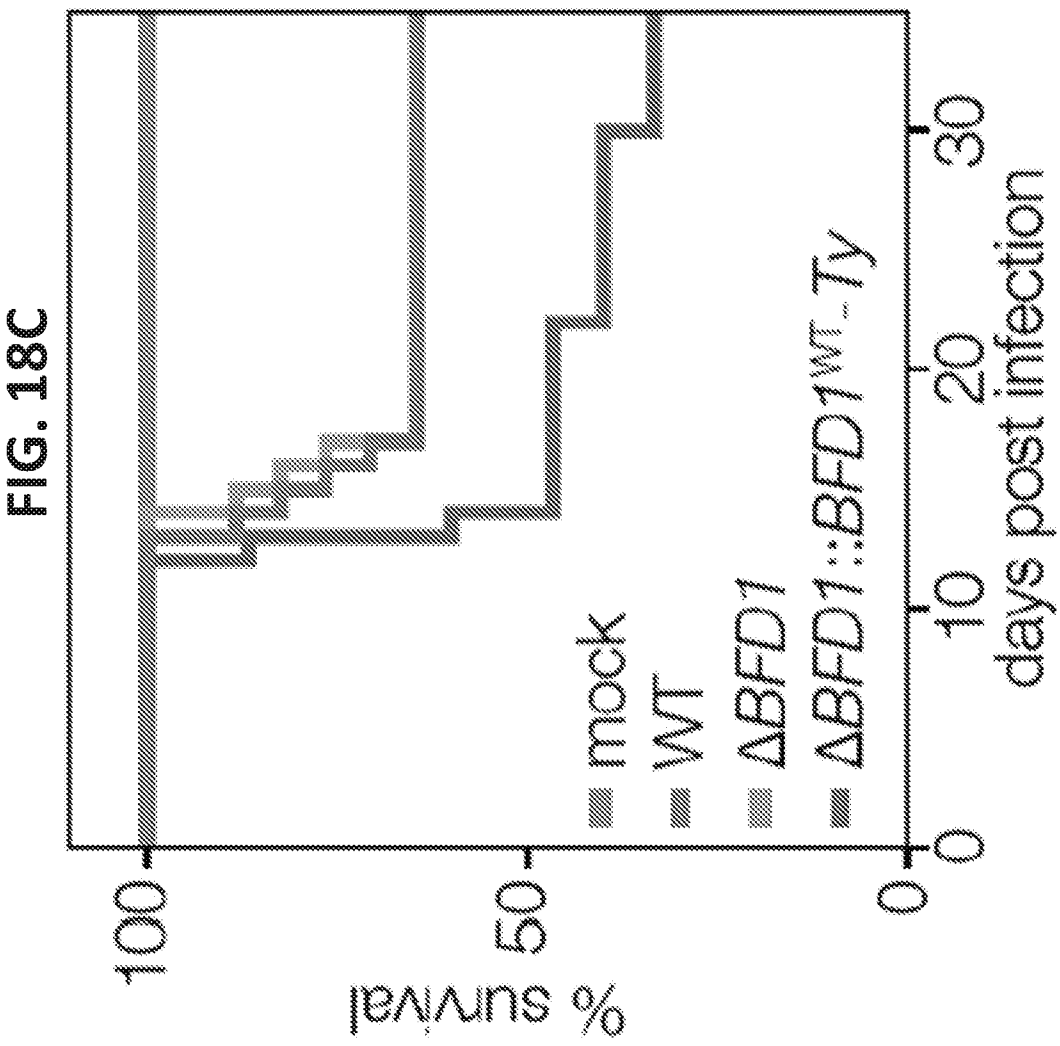

% survival
100
50
0
0
10
20
30
days post infection
mock
WT
ΔBFD1
ΔBFD1::BFD1WT-Ty

FIG. 18C

FIG. 18D

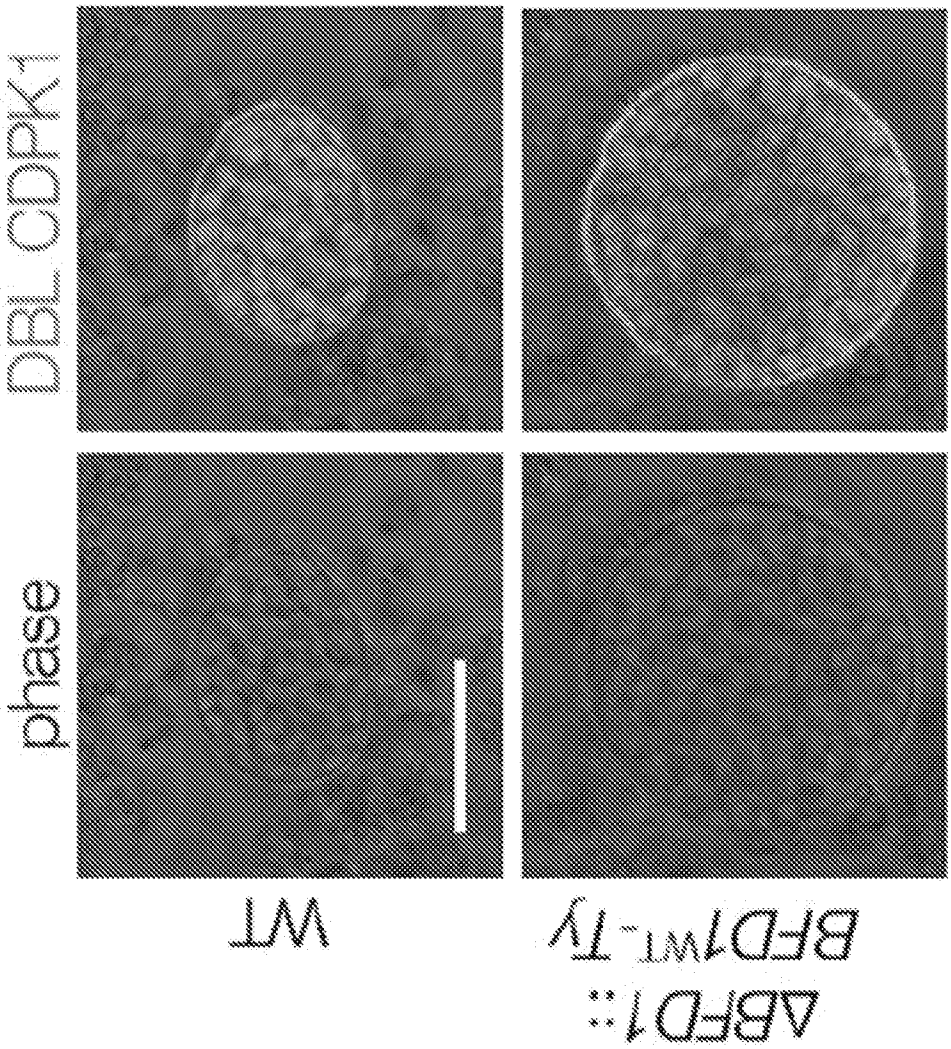


FIG. 18E

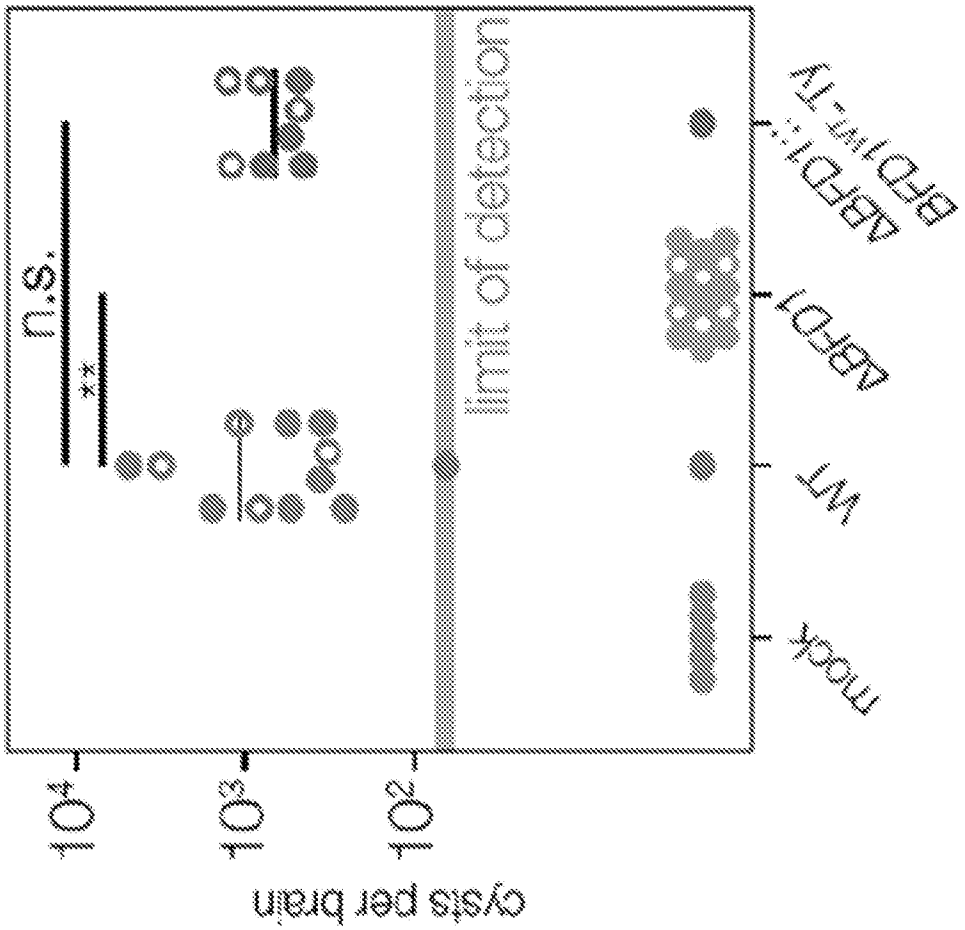

lysis per brain
10^4
10^3
10^2
limit of detection
n.s.
**
Mock
WT
ΔBRD1
ΔBRD1 + BRD1 WT

FIG. 19A

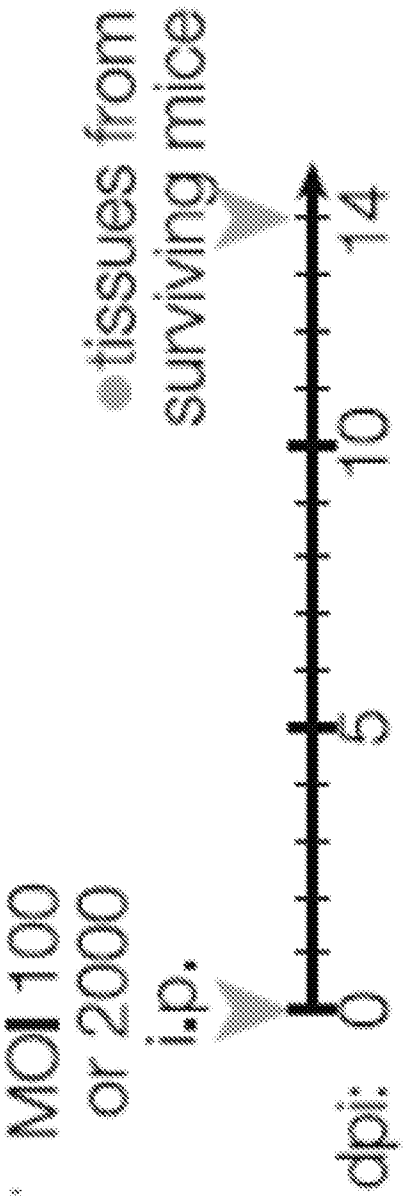
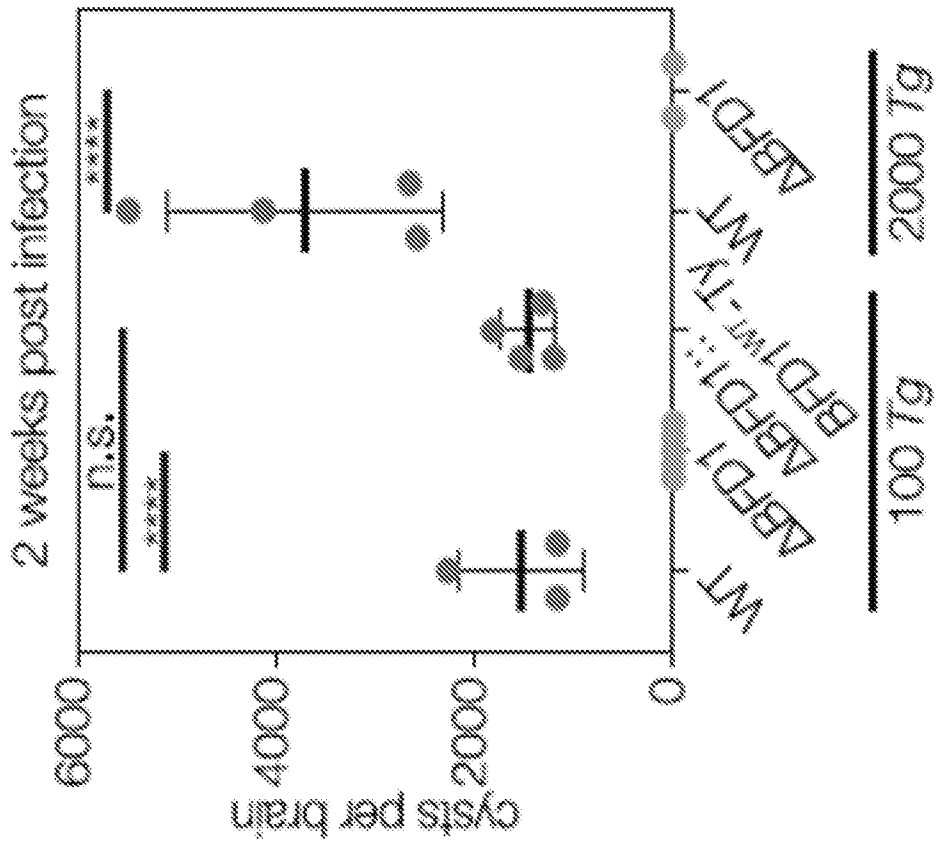


FIG. 19B


2 weeks post infection
cysts per brain
6000
4000
2000
0
n.s.
****
****
WT
ΔROP1
ΔROP1::ROP1HXG-TY
ΔROP1
WT
100 Tg
2000 Tg

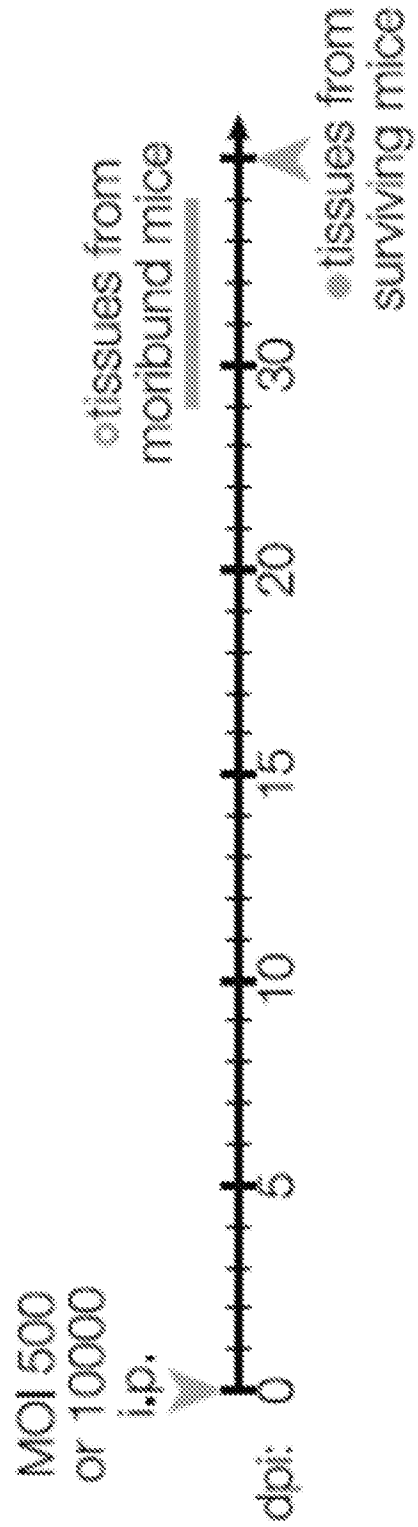

MOI 500
or 10000
i.p.
dpi:
0
5
10
15
20
30
tissues from
moribund mice
tissues from
surviving mice

FIG. 19C

FIG. 19E

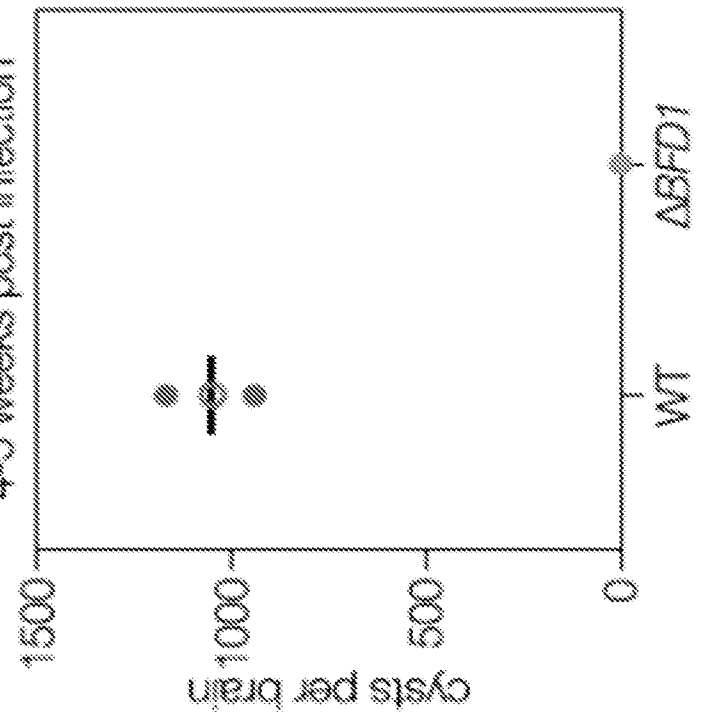


FIG. 19D