1 **The consensus N_{glyco}-X-S/T motif and a previously unknown N_{glyco}-N -linked glycosylation**
2 **are necessary for growth and pathogenicity of *Phytophthora***

3 Running title: N-glycosylation is necessary for *Phytophthora*

4 Can Zhang¹, Shanshan Chen¹, Fan Zhang¹, Tongshan Cui¹, Zhaolin Xue¹, Weizhen Wang¹, Borui
5 Zhang¹, Xili Liu^{1,2*}

6 ¹ *Department of Plant Pathology, China Agricultural University, Beijing 100193, China*

7 ² *State Key Laboratory of Crop Stress Biology for Arid Areas, Northwest A&F University, Yangling*
8 *712100, China*

9 *Correspondence author: Xili Liu, Email: seedling@cau.edu.cn.

Abstract:

Asparagine (Asn, N) -linked glycosylation within the glycosylation motif (N_{glyco}-X-S/T; X ≠ P) is a ubiquitously distributed post-translational modification that participates in diverse eukaryotic cellular processes. However, little is known about the characteristic features and roles of N-glycosylation in oomycetes. In this work, it found that 2.5 µg/ml tunicamycin (N-glycosylation inhibitor) completely inhibited *Phytophthora sojae* growth, suggesting that N-glycosylation is necessary for oomycete development. We conducted a glycoproteomic analysis of *P. sojae* to identify and map all N-glycosylated proteins and to quantify differentially expressed glycoproteins associated with mycelia, asexual cysts, and sexual oospores. A total of 355 N-glycosylated proteins were found, containing 496 glycosites that likely participate in glycan degradation, carbon metabolism, glycolysis, or other central metabolic pathways. To verify the glycoproteomic results and further examine the function of N-glycosylation in *P. sojae*, two proteins were selected for PNGase F deglycosylation assays and CRISPR/Cas9-mediated site-directed mutagenesis, including a GPI transamidase protein (GPI16) up-regulated in cysts, with the consensus N_{glyco}-X-S/T motif at Asn 94, and a heat shock protein 70 (HSP70) up-regulated in cysts and oospores with a previously unknown N_{glyco}-N motif at Asn 270. We demonstrated that the GPI16 and HSP70 are both N-glycosylated proteins, confirming that the N_{glyco}-N motif is a target site for asparagine - oligosaccharide N-glycosidic linkage. Glycosite mutations of Asn 94 in the GPI16 led to impaired cyst germination and pathogenicity, while HSP70 mutants exhibited decreased cyst germination and oospore production. This work describes an integrated map of oomycete N-glycoproteomes and advances our understanding of N-glycosylation in oomycetes. Moreover, we confirm that the consensus N_{glyco}-X-S/T and the N_{glyco}-N -linked glycosites are both essential for the growth of *Phytophthora sojae*, indicating that there are multiple N-glycosylation motifs in oomycetes.

Keywords:

N-glycosylation, oomycetes, mycelium, cyst, oospore, N_{glyco}-N motif, site-directed mutagenesis

Introduction

Asparagine (Asn, N) -linked glycosylation is among the most abundant post-translational modifications (PTMs) in a given genome, and is widely distributed across the tree of life (1, 2). This type of modification participates in the regulation of diverse biological processes and reportedly plays a major role in the development of complex multicellular organisms (3, 4). In humans, aberrant glycosylation has been implicated in a number of diseases, including neurodegenerative disorders and cancer (2, 5). Overall, more than 50% of proteins are estimated to be modified by covalent attachment of sugar molecules at some point in their life cycle, higher even than the 30% proportion of proteins predicted to undergo phosphorylation in organisms (6). Due to the complexity of this modification, relatively few large-scale studies of glycosylation have been conducted to date, though development in scope and throughput has increased in recent years. These studies have primarily focused on mammals and lower model organisms such as *Arabidopsis thaliana* and *Saccharomyces cerevisiae* (5, 7, 8). Thus, in order to determine the contributions of N-glycosylation to reproductive fitness and pathogenicity among microorganisms, it is necessary to map N-glycoproteomes in more organisms as well as in more developmental stages of those organisms.

N-glycosylation in the endoplasmic reticulum (ER) is a conserved, two-phase process in eukaryotic cells. First, the assembly of the lipid-linked oligosaccharide (LLO) initiates at the cytoplasmic side of the ER membrane and terminates in the lumen (9, 10). Then, the oligosaccharyltransferase (OST) catalyzes the *en bloc* transfer of the LLO to the asparagine side chain of the acceptor polypeptides in the ER lumen (9, 10). The OST selects polypeptides carrying an N_{glyco}-X-S/T (X ≠ P) motif and generates the N-glycosidic linkage between the side chain amide of asparagine and the LLO (9-12). Moreover, other conserved N-glycosylation motifs have been identified through liquid chromatography - tandem mass spectrometry (LC-MS/MS). For example, the N_{glyco}-X-C motif in mice (8), N_{glyco}-G and N_{glyco}-X-C motifs in yeast (5), as well as the N_{glyco}-A, N_{glyco}-N and N_{glyco}-D motifs in *F. graminearum* (13). However, the verification of these findings and further functional characterization of these N-glycosylation motifs has yet to be conducted.

N-glycosylation affects protein folding either directly or indirectly (14-16). After the protein is correctly folded, three glucose residues are removed and the glycoprotein is exported to the Golgi apparatus for further glycan processing (16, 17). If the N-glycosylation of a polypeptide is disrupted, the protein cannot properly fold into its native conformation, thus indicating the necessity for N-glycosylation in maintaining protein functionality during growth and reproductive development in different mammals and plants species (18-20). Two distinct processes have been described that help eukaryotic cells cope with this problem: ER-associated degradation (ERAD) and the unfolded protein response (UPR) (21). The ERAD system eliminates misfolded proteins via degradation in the cytosol; and the transcriptional induction of UPR target genes allows the

74 cell to adjust the capacity of protein folding in ER (21-23).

75 Oomycetes classified in the kingdom Stramenopila, are eukaryotic fungal-like microbes (24),
 76 many of which, especially *Phytophthora* spp., are important plant pathogens that cause tens of
 77 billions of dollars in annual losses to agriculture and aquaculture, globally (25, 26). *Phytophthora*
 78 *sojae* is a widely distributed, destructive oomycete plant pathogen that has been developed as a
 79 model for oomycete biology (25, 27). Like most *Phytophthora* species, *P. sojae* reproduces both
 80 sexually and asexually. During a disease cycle, the asexual life cycle (mycelium and cyst) is
 81 mainly responsible for rapid propagation and dispersal. In contrast, thick-walled oospores
 82 produced during sexual reproduction are essential for the pathogen's long-term survival by
 83 introducing greater diversity through recombination and prevention of genetic bottleneck within
 84 populations (28-30). Previous research, such as an iTRAQ-based proteomic comparison of
 85 mycelia and germinating cysts in *P. capsici* identified several candidate proteins involved in fatty
 86 acid degradation for energy production, thereby distinguishing the pre-infection stage from the
 87 mycelial stage (31). Proteomic analyses of mature oospores and non-mating hyphae identified
 88 proteins enriched in oospores were β -glucanases potentially involved in digesting the oospore wall
 89 during germination (32).

90 Given the relevance of the asexual and sexual phases in the epidemiology of *P. sojae*, we
 91 conducted a comprehensive glycoproteomic analysis of the N-glycosylated proteins associated
 92 with somatic growth, asexual cysts, and sexual oospores, in order to identify and map the full
 93 range of N-glycosylated proteins in this oomycete. In addition, glycoproteomic analysis allowed
 94 us to quantify differentially expressed glycoproteins during individual reproductive stages in *P.*
 95 *sojae*. Interestingly, we identified four previously unknown, conserved N-glycosylation motifs in
 96 addition to the consensus N_{glyco}-X-S/T motif. Furthermore, this work confirmed the
 97 N-glycosylation of two highly differentially expressed glycoproteins, including a GPI
 98 transamidase protein (GPI16) carrying the canonical consensus N_{glyco}-X-S/T target sequence and a
 99 heat shock protein 70 (HSP70) carrying a novel N_{glyco}-N motif, both of which were shown to be
 100 critical for full reproductive capability, and hence pathogenicity in *P. sojae*. To our knowledge, this
 101 is the first study to systematically characterize glycoproteins in oomycetes, and to report on
 102 function of the N_{glyco}-N glycosylation motif.

103 **Experimental Procedures**

104 **Effect of N-glycosylation inhibitor tunicamycin on *P. sojae***

105 Tunicamycin is a potent inhibitor of protein N-glycosylation that has frequently been used for
 106 experimental chemical inhibition of protein glycosylation in cells (33). Cultures of *P. sojae* strain
 107 PS6 were cultivated as described previously (34). The effect of 2.5 μ g/ml and 5 μ g/ml
 108 tunicamycin on *P. sojae* mycelial growth and cyst germination was evaluated using protocols from
 109 a previously published study (35). The colony diameters were measured perpendicularly after 6

110 days dark-incubation, and the germination rate of the cysts was assessed under a light microscope.
111 Tunicamycin was added to the 10% V8 liquid medium following 4 days of incubation of *P. sojae*
112 in the dark at 25°C. Oospore production was quantified 5 days later.

113 **Preparation of *P. sojae* mycelia, cysts, and oospores**

114 Mycelia were cultured in 100 ml of 10% V8 liquid medium for 4 days at 25°C. Mycelia were then
115 washed with autoclaved, distilled water, then extracted by filtration. Asexual zoospores were
116 obtained following established methods (36). One minute of vigorous shaking of *P. sojae* cultures
117 was used to induce the zoospores to encyst. After harvesting, mycelia and cysts were stored at
118 -80 °C until further use.

119 Since *P. sojae* is a homothallic species, oospores can be collected from individual colonies after
120 two weeks of dark-incubation on V8 agar at 25°C. The separation of oospores was carried out
121 following the techniques described by Niu (29), with some modifications. Briefly, a Polytron
122 (speed setting 7, Brinkman Instruments, Westbury, NY) was used to homogenize the mating
123 tissues collected from five dishes into 50 ml of sterile deionized water, using five intervals of 2
124 min each at 4°C. Suspensions of homogenized tissue were then centrifuged at 650 g for 5 min and
125 the supernatant was removed. After filtration through a 100 µm nylon mesh, the hyphae were
126 digested for 1 h in a 20 ml lyase cocktail (0.3 g lyase and 0.12 g cellulase), then washed three
127 times with sucrose solution and deionized water. Finally, the oospore suspension was centrifuged
128 at 4000 X g for 10 min to concentrate the oospores. Oospores were then flash frozen in liquid
129 nitrogen and stored at -80°C.

130 **Protein extraction and trypsin digestion**

131 The mycelia, cysts, and oospores were thoroughly ground to a fine powder in liquid nitrogen. The
132 powder was sonicated three times on ice using a high intensity ultrasonic processor (Scientz,
133 Ningbo, China) in lysis buffer (10 mM dithiothreitol, 1% protease inhibitor cocktail, and 1%
134 dephosphorylase inhibitor). An equal volume of Tris-saturated phenol (pH 8.0) was then added,
135 and the upper phenol phase was precipitated by the addition of at least four volumes of ammonium
136 sulfate-saturated methanol. After centrifugation at 4°C for 10 min, the remaining precipitate was
137 washed once with ice-cold methanol, followed by three washes in ice-cold acetone. The protein
138 was redissolved in 8 M urea and the protein concentration was determined using a BCA kit
139 (Beyotime Institute, Shanghai, China) according to the manufacturer's instructions.

140 Prior to digestion, the protein solution was reduced with 5 mM dithiothreitol for 30 min at 56 °C,
141 then alkylated with 11 mM iodoacetamide for 15 min at room temperature in darkness. The
142 protein sample was diluted by adding 100 mM triethylammonium bicarbonate (TEAB) to adjust
143 the urea concentration to lower than 2M. After this preparation, trypsin was added at a 1:50
144 trypsin-to-protein mass ratio for the first overnight digestion and at a 1:100 trypsin-to-protein mass
145 ratio for a second 4 h digestion.

146 **Tandem mass tag (TMT) labeling, high-pressure liquid chromatography (HPLC)** 147 **fractionation, and N-glycopeptide enrichment**

148 After trypsin digestion, the peptides were desalted using a Strata X C18 SPE column
149 (Phenomenex Torrance, USA) and vacuum-dried. Peptides were reconstituted in 0.5 M TEAB and
150 processed following the instruction manual for the TMT kit (Thermo Fisher Scientific, Waltham,
151 USA). The tryptic peptides were fractionated by high pH reverse-phase HPLC using a Thermo
152 Betasil C18 column (5 μ m particles, 10 mm ID, 250 mm length). The hydrophilic interaction
153 liquid chromatography (HILIC) enrichment of N-glycopeptides were performed according to
154 previous reports with minor modifications (37, 38). The peptide fractions were first dissolved in
155 the enrichment loading buffer (80% acetonitrile (ACN); 1% trifluoroacetic acid (TFA)) and then
156 pipetted into a HILIC tip (Jingjie PTM BioLab, Hangzhou, China). After centrifugation at 4000 X
157 g for 15 min, the HILIC tip was washed with 40 μ l of loading buffer three times. The enriched
158 N-glycopeptides were eluted with 10% ACN and lyophilized until dry. For deglycosylation, 200
159 units of PNGase F (NEB) was added into 50 μ l of 50 mM NH_4HCO_3 in H_2^{18}O (Aladdin, Shanghai,
160 Beijing) and incubated at 37°C overnight.

161 **Liquid chromatography separation and mass spectrometric analysis**

162 Digested peptides were separated using an Acclaim PepMap RSLC C18 capillary reversed-phase
163 analytical column (Thermo) with a 60 min 4-85% ACN/water gradient containing 0.1% formic
164 acid (FA) at a constant flow rate of 700 nL/min with an EASY-nLC 1000 UPLC system (Thermo).
165 The eluted peptides were further ionized and sprayed into the nanospray-ionization (NSI) source
166 followed by tandem mass spectrometry (MS/MS) with a Q ExactiveTM Plus instrument (Thermo)
167 coupled online to UPLC. The electrospray voltage applied was 2.0 kV. The m/z scan range was
168 350 to 1800 for full scan, and intact peptides were detected in the Orbitrap at a resolution of
169 70,000. Peptides were then selected for MS/MS using an NCE setting of 28 and the fragments
170 were detected in the Orbitrap at a resolution of 17,500. Automatic gain control (AGC) was set at
171 5E4.

172 **Database processing and bioinformatic analysis**

173 The fragmentation data were analyzed using the MaxQuant software package (version 1.5.2.8,
174 <http://www.maxquant.org/>) (39) to match ions against the *P. sojae* P6497 reference database
175 containing 25721 sequences (downloaded from the Uniprot database, <https://www.uniprot.org/>).
176 Trypsin/P was specified as the cleavage enzyme, allowing up to 2 missing cleavages. Mass error
177 was set to 5 ppm for precursor ions and 0.02 Da for fragment ions. Carbamidomethylation on Cys
178 (+57.0215 Da) was specified as a fixed modification, while oxidation on Met (+15.9949 Da), and
179 deamidation (^{18}O) on Asn (+2.9882 Da) were specified as variable modifications. The false
180 discovery rate (FDR) thresholds for protein, peptide, and modification sites were specified at 0.01.

181 Soft MoMo (motif-x algorithm, <http://meme-suite.org/tools/momo>) analysis was used to model the

amino acid sequence composition at specific positions in glycosylated 21-mers (10 amino acids upstream and downstream of the site) (40). Secondary structure analysis was performed using NetSurfP (<http://www.cbs.dtu.dk/services/NetSurfP/>) with *p* values calculated by Wilcoxon test (41). Gene Ontology (GO) terms for the identified proteins were assessed by querying the UniProt-GO annotation database (<http://www.ebi.ac.uk/GOA>) (42). The protein domain functional annotations were assigned using InterProScan, as implemented in the InterPro domain database (<http://www.ebi.ac.uk/interpro/>) (43). Soft Wolf PSORT version 2.0 (<http://www.genscript.com/psort/wolfsort.html>) was used for protein subcellular localization prediction (44).

Analysis of differentially expressed modified proteins

The ratios of the TMT reporter ion intensities in MS/MS spectra (*m/z*) from raw data sets were used to calculate fold changes between samples. Whole-proteome analysis was also performed to identify differentially expressed N-glycosylated proteins between the mycelial, cyst, and oospore growth and reproductive stages of *P. sojae*, as described above (except for the enrichment step). The N-glycosylation data were then normalized to the proteome data. In general, the differentially expressed modified and non-modified proteins were defined as having a fold-change ≥ 1.50 or ≤ 0.67 , and a *p* value < 0.05 .

Plasmid construction for overexpression and site-directed mutagenesis

To further validate the N-glycosylation modifications identified by proteomic analysis, a GPI transamidase glycoprotein (GPI16) with a consensus N_{glyco}-X-S/T motif that was highly expressed in the cysts, and a heat shock protein 70 (HSP70) with a N_{glyco}-N motif highly expressed in oospores were cloned. The sequences of the targeted candidate proteins were obtained by searching the *P. sojae* genome, v3.0 (<https://genome.jgi.doe.gov/Physo3/Physo3.home.html>). The full-length genes were amplified using EasyPfu DNA Polymerase (TransGen Biotech, China) and confirmed by Sanger sequencing (Primers in Table S1). Substitution mutations were introduced into the genes encoding the two candidate proteins by oligonucleotide-directed mutagenesis (Table S1) (45). The oligonucleotides used to alter the potential N-glycosylation sites were designed to replace the codon for asparagine (Asn, N) with that of alanine (Ala, A). After sequencing, no changes were detected other than the desired nucleotide substitutions in the cDNAs of mutant variants.

For protein overexpression to validate the identified glycoproteins, the target genes with or without glycosite mutation were cloned into the pTOR::FLAG vector using the In-fusion HD Cloning kit (Clontech, Beijing, China) (Table S1). The CRISPR/Cas9 system was used for excision and replacement of the wild-type target proteins with glycosite-mutagenized variants in order to better understand the functions of these N-glycosylation sites in *P. sojae* growth and reproduction. Constructs encoding two single guide RNAs (sgRNAs) (sg214 for GPI16: ACTCGGCGTCGTGTTTCTCG; and sg603 for HSP70: CGTCTGGTAGTCCTCCTCCG) were

cloned separately into plasmid PYF515. To avoid the sgRNA-mediated Cas9 cleavage of the homology directed repair donor, the sgRNA binding residues in the donor plasmid were mutated without changing the encoded amino acids, producing the two donor plasmids. Flanking sequences, 1,000 bp upstream and downstream of the mutation site, were amplified using GPI16-F1/R1, GPI16-F2/R2, and HSP70-F1/R1, HSP70-F2/R2 primers (Table S1) and ligated into pBS-SK⁺ using the In-fusion HD Cloning kit (Clontech, Beijing, China), respectively, to create the basic donor plasmid (46).

Transformation of *P. sojae*

The transformation of *P. sojae* was accomplished by following protocols described in Fang and Tyler (47). To acquire the overexpression mutants, the pTOR::FLAG plasmids carrying each glycoprotein were respectively transformed into *P. sojae*. PYF515-sgRNA was co-transformed into *P. sojae* PS6 together with its donor plasmid for the site-directed mutagenesis of the two target glycoproteins. The mutants were initially screened in pea broth medium amended with 30 µg/ml geneticin (Amresco, Solon, United States) and then on V8 agar plates containing 50 µg/ml geneticin. The hyphae of the mutants were then collected for genomic DNA (gDNA) extraction. For verification of the mutants, the amplification products of the full-length target candidate genes were confirmed by sequencing in all mutants.

The validation of target glycoproteins

Harvested hyphae were ground to a powder using a pestle and mortar, and the total protein was extracted using a Protein Extraction kit (Invent, Beijing, China). Cell debris was removed by centrifugation for 10 min at 13000 x g at 4°C. Anti-FLAG M2 beads (Sigma-Aldrich, Shanghai, China) were added to the lysate, followed by incubation for 5 h at 4°C with gentle shaking. Proteins binding to the beads were eluted into elution buffer (7 M urea, 2 M thiourea, 65 mM dithiothreitol and 200 mM PMSF). PNGase F (NEB, Beijing, China) was used for the deglycosylation analysis with an overnight incubation at 37°C.

Proteins were separated in SDS-PAGE gels and the separated proteins were transferred to polyvinylidene difluoride membranes. The membranes were then blocked using PBSTM buffer (phosphate-buffered saline with 0.1% Tween 20 and 5% non-fat milk) for 30 min at room temperature with 60 rpm shaking. Anti-FLAG antibodies (1:5000; Sigma-Aldrich, Shanghai, China) were then added to the PBSTM. After addition of the antibody, the membranes were incubated at room temperature for 3 h then washed three times (10 min each) with PBST buffer. The membranes were then incubated with an HRP-conjugated goat anti-mouse IgG antibody (CWBIO, Beijing, China) at a dilution of 1:5000 in PBSTM at room temperature for 1 h with 60 rpm shaking, and followed by three washes (5 min each) with PBST. The signals were detected by Azure c600 imager (Azure Biosystems, USA).

Phenotypic analysis of the site-directed mutants

254 Mycelial growth of *P. sojae* mutant strains was assessed by measuring the colony diameter in two
 255 perpendicular directions after 6 days of incubation in the dark on 10% V8 agar. Sporangia and
 256 zoospores were obtained using previously published protocols (48). Zoospore encystment was
 257 measured by pipetting 100 μ l of zoospore suspension onto a glass plate and counting the number
 258 of encysted zoospores after 4 h incubation at 25°C and 80% humidity. Oospores of *P. sojae* were
 259 collected after two weeks of incubation in the dark on V8 agar at 25°C. The 4-day-old etiolated
 260 soybean seedlings (cv. Ribenqing, grown in the dark at 25°C and 80% humidity) were used for the
 261 *in planta* pathogenicity assays (36). The seedlings were spot inoculated at 2 cm below the
 262 beginning of the root zone with 10 μ l zoospores suspension (1×10^4 zoospores/ml), and incubated
 263 at 25°C for 2 days in the dark. The lengths of the disease lesions above and below the point of
 264 inoculation on the hypocotyl were then measured. Each treatment contained at least 10 seedlings
 265 and the entire experiment was conducted at least twice.

266 RNA extraction and Quantitative real time PCR assays

267 Quantitative real time PCR (qRT-PCR) was used to assess the relative expression levels of the
 268 genes in the ER-associated degradation (ERAD) pathway and the unfolded protein response (UPR)
 269 pathway in the glycosite mutant strains and wild-type *P. sojae*. The biological samples were
 270 prepared according protocols established in the above section. Total RNA was first extracted using
 271 the SV Total RNA Isolation kit (Promega, Beijing, China) before cDNA was synthesized using the
 272 PrimeScript RT reagent kit with gDNA Eraser (Takara, Beijing, China) following the protocols
 273 provided by the manufacturer. The qPCR reactions were performed using the SYBR Premix
 274 Dimer Eraser kit (Takara, Beijing, China). The relative quantities of the PCR products were
 275 calculated using the $2^{-\Delta\Delta C_t}$ method, with the *RPS* and *RPL13a* genes used as internal references to
 276 normalize the transcript levels. The primers used in the qRT-PCR experiments are listed in Table
 277 S1. All treatments were represented by at least three replicates and the entire experiment
 278 conducted three times.

279 Results

280 N-glycosylation inhibitor tunicamycin significantly inhibited growth of *P. sojae*

281 The N-glycosylation inhibitor tunicamycin is used extensively for manipulation or prevention of
 282 protein N-glycosylation in cells (33), although its effects on oomycetes have not been reported.
 283 We first examined the effect of tunicamycin on mycelial growth, cyst germination, and oospore
 284 production in *P. sojae* and found that 2.5 μ g/ml tunicamycin completely inhibited the growth of *P.*
 285 *sojae* (Fig. 1A). The rate of cyst germination dropped from 81.5% in the untreated control to 10.8%
 286 and 7.6% following the addition of 2.5 μ g/ml or 5 μ g/ml tunicamycin, respectively (Fig. 1B). The
 287 WT *P. sojae* exhibited a slight decrease in oospore production under exposure to 2.5 μ g/ml
 288 tunicamycin, but a significant decrease under treatment with 5 μ g/ml tunicamycin ($P < 0.001$, Fig.
 289 1C). These results indicated that N-glycosylation likely plays an essential role in the growth of *P.*

290 *sojae*.

291 **A large set of glycosylation sites were identified across a range of proteins in *P. sojae***

292 In order to obtain a comprehensive view of N-glycosylation patterns in *P. sojae*, we conducted
293 tandem mass spectrometry on the enriched N-glycopeptides from mycelia, cysts, and oospores. A
294 total of 474 N-glycopeptides were identified with high confidence, corresponding to 355 proteins
295 and 496 sites, using Maxquant software to search our MS/MS fragments against the *P. sojae*
296 P6497 reference proteome. The false discovery rate (FDR) for both glycopeptides and
297 glycosylation sites was adjusted to <1%. Among the identified N-glycosylation sites and proteins,
298 we quantified the glycosylation level, or number of modified sites per protein for 341 sites across
299 239 proteins in the three different developmental stages (Fig. S1A). To our knowledge, this dataset
300 represents the first oomycete quantitative N-glycoproteome to date.

301 The mass accuracy distribution of the identified glycopeptides (Fig. S1B) showed that all of the
302 peptides had a mass error of < 5.5 ppm, supporting the high confidence of their identification. A
303 majority of the N-glycosylated proteins contained between one and three post-translational
304 modification (PTM) sites. Specifically, 71.27% of the 355 proteins had one N-glycosylation site,
305 whereas 20.00% and 7.32% of the PTM proteins had two and three N-glycosylation sites,
306 respectively (Fig. S1C). The number of modified proteins with four or more PTM sites was
307 significantly lower. The alpha-mannosidase protein (accession No: G4Z5K4) had the highest
308 number (six) of N-glycosylation sites, while the tyrosinase protein (accession No: G5A886) had
309 five sites. We thus found and identified a large number of N-glycosylated proteins, indicating that
310 N-glycosylation is widely present in *Phytophthora*. These glycoproteins showed differences in the
311 number of PTM sites, suggesting that N-glycosylation may not affect each glycoprotein equally.

312 **N-glycosylation preferentially located on five conserved N-glycosylation motifs**

313 To identify sequence motifs that were generally associated with N-glycosylation sites, we
314 examined a 21-amino acid sequence window surrounding the central N_{glyco} residues of each site
315 using the web-based Motif-X program with a significance threshold of $P < 0.000001$. Previous
316 research showed that the consensus sequence for N-glycosylation motif is N_{glyco}-X-S/T (X≠P)
317 (10-13). In this study, we found that most of the 248 N-glycosylation sites matched the
318 N_{glyco}-X-S/T motif (Fig. 2A). Interestingly, four other conserved motifs were identified among the
319 regions surrounding N-glycosylation peptides, including N_{glyco}-A, N_{glyco}-N, N_{glyco}-D, and N_{glyco}-S
320 (Fig. 2A). As indicated by the heat map of the frequencies of different amino acids that were
321 adjacent to the N_{glyco} residues (Fig. 2B), N-glycosylated sequences were non-randomly distributed
322 among residues, with higher frequencies associated with glycine (G) at the -1, -6 site and lysine (K)
323 at the -7, -8 site. Three nonpolar residues, including proline (P), leucine (L), and phenylalanine (F),
324 more frequently occupied the +1 or +2 site (the cut-off value for the log₁₀ transformed Fisher's
325 exact test P value was set at 1).

To further understand the effects of N-glycosylation on protein structure, secondary structural predictions were performed using the NetSurfP algorithm on both the modified subset and un-modified Asn residues. We observed that N-glycosylation was more likely to occur on the coil structures (P value = 3.72×10^{-5}). In contrast, N-glycosylated residues were less likely to be located on beta-strands (P value = 5.25×10^{-4}), and no significant differences were found between modified and un-modified Asn residues on alpha-strands (P value = 1.32×10^{-1}) (Fig. 2C, left). The surface accessibility of N-glycosylation sites was significantly higher than that of un-modified Asn residues (P value = 1.53×10^{-11}), indicating that N-glycosylated residues may be preferentially located on the outside of the modified proteins (Fig. 2C, right). This result may also explain the predominant co-localization with hydrophilic amino acids near the N_{glyco} motifs.

***P. sojae* glycoproteins belonged to a wide range of biological processes**

To clarify the functions of these N-glycosylation proteins in *P. sojae*, all of the identified N-glycosylated proteins were subjected to subcellular localization prediction, GO functional classification, and KEGG pathway analyses. Predictions of subcellular localization suggested that among the identified glycoproteins in mycelia, cysts, and oospores, 29.58% were targeted for extracellular secretion, 20.28% were localized in mitochondria, 19.72% were distributed in the cytoplasm, and 14.37% were targeted to the plasma membrane (Fig. 3A). Moreover, most of the glycoproteins matched the N_{glyco}-X-S/T motif were more likely to be extracellular (Fig. 3B). However, the glycoproteins carrying four other conserved N-glycosylation motifs, including N_{glyco}-A, N_{glyco}-N, N_{glyco}-D, and N_{glyco}-S, preferred to localize in mitochondria, with a higher percentage of 28.93% than in other localizations (Fig. 3C).

Previous studies have demonstrated that post-translational N-glycosylation is widely distributed throughout many important central metabolic pathways (49). Thus, to better understand the general functions of glycoproteins identified in *P. sojae*, the GO and KEGG database were used to explore the specific metabolic pathways associated with these N-glycosylated proteins. Within the three GO categories (biological process, molecular function, and cellular component), N-glycosylated proteins were assigned to a wide range of biological process sub-categories, including carboxylic acid metabolic process, aerobic respiration, nucleoside metabolic process, single-organism carbohydrate catabolic process, and membrane lipid metabolic process (Fig. 3D). KEGG analysis revealed that the N-glycosylated proteins were involved in glycan degradation (2.96%), biosynthesis of antibiotics (12.32%), carbon metabolism (8.87%), glycolysis/gluconeogenesis (5.42%), ribosome (5.42%), TCA cycle (3.45%), and protein processing in the endoplasmic reticulum (6.40%) (Fig. 3E). These results suggest that *P. sojae* may directly regulate important metabolic pathways through N-glycosylation, especially for carbon metabolism, aerobic respiration, and nucleoside metabolic processes, as well as other functional proteins.

Changes in N-glycosylation profiles across three developmental stages

We analyzed the differentially expressed N-glycosylated proteins (DEPs) with normalization of the proteome data in the three samples of *P. sojiae*. Among the 239 glycoproteins, 89 of them, containing 106 glycosites, showed differences in abundance between cysts and mycelia (Fig. S2A). Moreover, 77 DEPs were common to both oospore comparison groups (Oo vs Cy; Oo vs My), while only 32 or 24 DEPs were shared in cyst comparisons (Cy vs My; Oo vs Cy) or mycelium comparisons (Cy vs My; Oo vs My), respectively (Fig. S2B). These results suggested that in *P. sojiae*, the diversity, or the number of different DEPs, was greater during oospore life stages. Based on GO, KEGG, and domain pathway enrichment analysis, N-glycosylated proteins associated with iron binding and protein transport were highly enriched in oospores (Fig. S3). However, a higher proportion of these proteins in the Cy vs My comparison could not be functionally clustered, reflecting the fact that many of the proteins associated with cyst development have yet to be characterized in *P. sojiae* and other related organisms.

Two significantly differentially expressed glycoproteins with different N-glycosylation motifs were selected to first validate the glycoproteomic data, and second to examine the contribution of glycosylation to protein function in reproductive development and pathogenicity. These proteins included a GPI transamidase component protein (GPTI, accession No: G4Z0S6) enriched in the cyst stage and carrying a consensus N_{glyco}-X-S/T motif, and a heat shock protein 70 (HSP70, accession No: G5A5V6) carrying a novel N_{glyco}-N motif that was up-regulated in the cyst and oospore stage (Table 1). The mass spectra used for N-glycosylation identification of these three glycopeptides are shown in Fig. S4. The GPTI glycoprotein only had one glycosite at Asn 94, while the HSP70 protein had two glycosites: one at Asn 270 and the other at Asn 305 (Table 1). The expression level of GPTI-Asn 94 was 2.00-fold higher in cysts than in mycelia and 5.56-fold higher than in oospores, respectively. Similarly, the expression of HSP70-Asn 270 was 1.38- and 1.20-fold higher in cyst and oospores than in mycelia, respectively. Since HSP70-Asn 305 showed no significant up-regulation in either cyst or oospore stage, this Asn residue glycosite was not included in further experiments (Table 1).

Two candidate glycoproteins carrying different conserved glycosylation motif were verified

To verify N-glycosylation of the two candidate glycoproteins, the proteins were checked by immunoblot following treatment with peptide-N-glycosidase F (PNGase F), which catalyzes the removal of N-linked oligosaccharide chains from glycoproteins (50). The results showed that the apparent molecular weight of GPI16 and HSP70 were clearly lower after PNGase F deglycosylation, indicating that both proteins were originally glycosylated (Fig. 4A and Fig. 4C). To further characterize the N-glycosylation sites, GPI16 protein was mutagenized by point mutation at glycosite Asn 94 (GPI16^{N94A}) and HSP70 was given a point mutation at glycosite Asn 270 (HSP70^{N270A}), then checked again by immunoblotting. As expected, GPI16^{N94A} and HSP70^{N270A} could be separated from GPI16 and HSP70, respectively; and the molecular weight became smaller compared to each parental wild-type protein (Fig. 4B and Fig. 4D). This result

demonstrated that the Asn 94 consensus N_{glyco}-X-S/T motif and Asn 270 N_{glyco}-N motif are both glycosites, and more than one conserved glycosylation motif was found in *P. sojae*.

The N-glycosylation motifs of GPI16 and HSP70 were further analysed in different oomycetes species, including four *Phytophthora* species, and two *Pythium* species. The N-glycosylation motifs of these two target proteins, no matter the N_{glyco}-X-S/T or the N_{glyco}-N motif were both conserved in *Phytophthora* species, as indicated by PSURA for *P. ramorum*, PTPG for *P. infestans*, PPTG for *P. parasitica*, and PHYCA for *P. capsici* (Fig. 4E and Fig. 4F). In addition, only the N_{glyco}-N motif of HSP70 was conserved among *Pythium* species, while the motif of GPI16 did not appear in *Pythium* species at the corresponding position (Fig. 4E and Fig. 4F). The conservation of glycosylation sites in oomycetes, especially in *Phytophthora*, suggest they might share similar important functions in oomycetes.

N-glycosylation of the target glycoproteins is necessary for the growth of *P. sojae*

In order to investigate the function of significantly differentially expressed N-glycosylation site(s) in the two candidate glycoproteins, the CRISPR/Cas9 system was used to replace the Asn residue target site with an Ala residue. We characterized and compared the phenotypes of the wild-type PS6 (WT), a *P. sojae* control transformed with the empty vector (EV), and *P. sojae* isolates containing mutations that prevent Asn target site glycosylation for each of the glycoproteins. No effect was found on mycelial growth and oospore production on GPI16 mutants (Fig. 5A and Fig. 5C), whereas GPI16 mutations significantly affected the germination rate of cysts and pathogenicity of the mutant isolates against soybean ($P < 0.001$, Fig. 5B and Fig. 5D). The cyst germination rate was 82.7% and 78.5% in the WT and EV isolates, respectively, but only 14.8 and 19.2% in GPI16 homozygous mutants M16 and M29, respectively (Fig. 5B). Additionally, etiolated soybean seedlings inoculated with WT and EV zoospores exhibited a lesion length between 66.8-70.0 mm, while the mutants produced significantly shorter lesions of 48.6-49.6 mm in the mutants (Fig. 5D).

The HSP70 mutants showed similar colony diameter as the WT and EV isolates (Fig. 5E). The cyst germination rate was significantly decreased in the HSP70 homozygous mutants compared to WT and EV isolates ($P < 0.01$; Fig. 5F). We observed a significant drop in the average number of oospores from 189.9 to 118.2 and 139.8 in each cm² zone among the mutants M12 and M15, respectively (Fig. 5G). Moreover, pathogenicity was also affected in the HSP70 homozygous mutations ($P < 0.001$, Fig. 5H). In brief, these results of site-directed Asn glycosite mutagenesis were consistent with the relative expression assays indicating low expression of glycosites during developmental stages where mutagenesis had little effect, and high expression in stages that were severely impaired by abolishing glycosylation. This finding further supports the necessity for N-glycosylation of these target proteins for normal growth of *P. sojae*, for both the N_{glyco}-X-S/T sequence and the N_{glyco}-N motif.

Elimination of N-glycosylation induces expression of the ERAD and UPR pathway genes

The inhibition of protein N-glycosylation leads to the accumulation of unfolded proteins in the ER (ER stress), and induces expression of the unfolded protein response (UPR) and ER-associated degradation (ERAD) pathway target genes to adapt to the decreased stability of glycoproteins (21-23). We then sought to determine whether UPR and ERAD pathway gene expression was modulated by these mutagenized DEPs. Using qRT-PCR to measure the transcript levels of major ER chaperones and foldases, including *CNE1-2* (accession No: G4YMY1 and G4YVS5), encoding the *P. sojae* homolog of calnexin; *PDII-3* (accession No: G4YXR9, G4YKG9, G5A3Y0), encoding protein disulfide isomerase; and *ERAD1-5* (accession No: G4Z2L8, G4ZMP2, G4Z0Q7, G4YIZ4, G4ZYS3), encoding ER-associated degradation proteins in WT, EV, and glycosite mutant strains. The expression of two calnexin genes *CNE1-2* was not affected in both GPI16 mutant M16 and HSP70 mutant M12 (Fig. 6A and Fig. 6B). The mRNA levels of *PDII* and *PDII3* were approximately 1.67- and 1.89-fold higher in the M12 mutant compared with WT, respectively (Fig. 6C and Fig. 6E). In the GPI16 mutant, the expression levels of five *ERAD* genes were all up-regulated, and *ERAD1* appeared 3.03-fold up higher than WT (Fig. 6F-6J). These results indicated that abnormal growth of *P. sojae* caused by elimination of N-glycosylation in these target glycoproteins was potentially related to aberrant protein folding, with different UPR and ERAD genes are playing roles in this process.

Discussion

Currently, N-glycoproteome data is very limited in phytopathogenic fungi and oomycetes, with the only fungal N-glycoproteome, that of *F. graminearum*, reported in 2016 (13). In the present study, we identified 496 N-glycosylation sites and 355 glycosylated proteins in three developmental stages of *P. sojae*. Similarly, 425 and 516 high-confidence sites were previously mapped in fission and budding yeast (8), while 406 proteins containing 774 sites have been identified in *F. graminearum* (13). According to these results, the proportion of N-glycosylated proteins among the total proteins appears lower in microorganisms than the proportion reported in mammals, using the same cut-off scores for identifying PTM peptides (38). Overall, this study extends the number of known sites and target motifs, and advances the current knowledge surrounding the oomycete N-glycoproteome, providing insight for further exploration of the potential physiological functions of these glycosylated proteins in reproduction and pathogenicity.

Most of the 496 N-glycosylation sites matched with the consensus motif for N-glycosylation N_{glyco}-X-S/T (X ≠ P) as shown in Fig. 2A. This finding is in concordance with the fact that the underlying core “N-glycosylation machinery” consists of few proteins, which are highly conserved in eukaryotic species (11, 51, 52). However, the motif patterns, such as N_{glyco}-A, N_{glyco}-N, and N_{glyco}-D in *P. sojae* (Fig. 2A) differed from glycosylation sites in other species (38, 53). The N_{glyco}-N motif has only been previously reported in *F. graminearum* (13), suggesting that these sites may not be conserved among higher organisms. The N_{glyco}-G motif has been reported in

several species including mammals (8), plants (5), and fungi (5, 13), and the N_{glyco}-X-C motif was found in mammals (5, 8). However, these two motifs are absent in oomycetes, indicating an evolutionary divergence in the specificity of N-glycosylation. Overall, these results suggest there are more than one conserved glycosylation motifs in organisms.

N-glycosylation adheres to very rigid topological and sequence constraints, and N-glycosylated proteins have been reported to be mainly located in the cell exterior or in intracellular compartments such as the ER and Golgi (5, 38, 53). The results of this study were consistent with previously reported properties and distributions of N-glycoproteins, with most glycoproteins located in the extracellular matrix or plasma membrane (Fig. 3A), thus supporting the confidence of our N-glycoprotein identifications. However, in contrast to previous studies, we found that the second most common subcellular compartment for glycoproteins was mitochondria in oomycetes (Fig. 3A), especially in the glycoproteins carrying N_{glyco}-A, N_{glyco}-N, N_{glyco}-D, and N_{glyco}-S motifs (Fig. 3C), which has only been reported in *F. graminearum* (13), thus reflecting a novel and potentially unique role of N-glycosylation in phytopathogenic microorganisms. In addition, some glycoproteins were localized to the cytoplasm based on curated annotation data (Fig. 3A-3C). An alternative type of N-glycosylation has been identified in bacteria that takes place in the cytoplasm (54-57). This novel cytosolic N-glycosylation is performed by an enzyme structurally different from OST, but which displays the same acceptor site specificity for N_{glyco}-X-S/T (11). In light of this finding, we speculated that more kinds of OSTs could be identified in oomycetes due to the unique subcellular localization of glycoproteins.

GO and KEGG pathway analyses revealed that N-glycosylated proteins expressed during the stages examined in this work were involved in 'carbon metabolism', 'TCA cycle', and 'fatty acid metabolism' among others, confirming *P. sojae* glycoproteins belonged to a wide range of biological processes. The calcium ion binding and cation binding functional categories were enriched for the most highly up-regulated N-glycoproteins in the Cy vs My comparison group; this finding indicates that ion flux and homeostasis are the most quantitatively distinct processes between cysts and the mycelia. Calcium is necessary for signaling and function in eukaryotic flagella, and silencing of *Pibzpl1*, a regulator of calcium signaling, adversely affected the cyst germination rate and the ability to infect plants in other *Phytophthora* species (58, 59). Notably, KEGG pathway analysis of the differentially expressed N-glycoproteins in the Oo vs My and Oo vs Cy comparisons found that 'protein processing in endoplasmic reticulum', 'other glycan degradation', and 'sphingolipid metabolism' pathways were all enriched for glycoproteins in the oospore stage. N-glycosylation is conducted in the endoplasmic reticulum (10, 11); this result is consistent with the differential increase in expression of glycosylated proteins in oospores compared to the other two stages, indicating a necessary demand for glycoproteins during sexual reproduction.

Deglycosylation analyses and site-directed mutagenesis of the identified glycosites were

performed in the current work to validate the results of our N-glycoproteome annotation, and to further interrogate the significant differentially expressed glycoproteins for their role in development and reproduction of *P. sojae*. This study demonstrated that GPI16 and HSP70 are both N-glycosylated at Asn residues in conserved motifs in *P. sojae*, further validating the N-glycoproteome data and confirming the presence and functionality of the N_{glyco}-N motif in oomycetes. We also found that both the N_{glyco}-X-S/T and the N_{glyco}-N motifs are conserved at the genus level, suggesting that these OST recognition sequences have potential for development as fungicide target sites for the control of *Phytophthora*.

The glycosites of these two proteins were mutagenized using CRISPR/Cas9 to ablate their capacity for post-translational glycosylation. Among these, the GPI16 protein participates in GPI anchoring, a ubiquitous mode of post-translational modification in eukaryotic organisms (60). Attachment of GPI anchors to proteins is mediated by the GPI transamidase enzyme complex (60-64). N-glycosylation was necessary for the expression of functional GPI16, as indicated by the decreases in the rate of cyst germination and pathogenicity among glycosite mutants (Fig. 5). The GPI transamidase subunits GPI8, GPI16, and GAA1 form three-component core sub-complexes in *Saccharomyces cerevisiae*, which is conserved in many species (65). Structural modeling suggests that GPI16 may present the substrate protein to GPI8p, and deletion of *GPI16* gene in *Plasmodium berghei* reduces GPIs and GPI-anchored proteins in parasite membranes (66). Further work is needed to determine whether N-glycosylation of GPI16 affects the biosynthesis of GPIs in *Phytophthora*. The second glycosite-mutagenized protein, HSP70 was shown to contribute to cyst germination, oospore production and pathogenicity (Fig. 5). Hsp70 is, by far, a more evolutionarily conserved protein than GPI16 (67, 68), which is consistent with the similarity analysis in this study showing conservation of the N-glycosylation motif between *Phytophthora* and *Pythium* species.

Many lines of evidence have shown that glycoproteins are more stable than their corresponding unglycosylated counterparts, regardless of whether or not the glycosylation is accompanied by major structural changes (69). One proposed explanation is that glycosylation reduces disorder in unfolded proteins, thus increasing the reversibility of the unfolding process and resulting in overall entropic effects that reinforce protein stability. Alternatively, N-linked glycans may be necessary to stabilize some secondary structures that are necessary for protein function or that maintain folded protein conformation (70, 71). To further explore the downstream effects of N-glycosylation on gene expression during different reproductive stages in *Phytophthora*, we measured the relative expression levels of 14 UPR and ERAD target genes in the glycosylation knockout mutant strains. The mRNA levels of UPR pathway genes *PDII* and were up-regulated in the HSP70 mutants, while the five ERAD genes were all highly expressed in the GPI16 mutants (Fig. 6). Similarly, previous work revealed that in *Trichoderma reesei*, abolishing N-glycosylation of CBH1 led to the induction UPR target gene expression, with 5-, 4-, and 2-fold up-regulation of

547 *CNE1*, *BIP1*, and *PDII*, respectively (72). These results thus demonstrated that elimination of
548 glycosylation sites can lead to a range of different effects on the folding and stabilization of
549 glycoproteins in oomycetes, though further exploration is needed to characterize the full extent
550 and underlying mechanisms of these effects.

551 The developmental stages of plant pathogens have been shown to characteristically differ at the
552 transcriptional, post-transcriptional, translational, and post-translational levels (31, 73-75). In
553 particular, N-glycosylation is a post-translational modification critical for fundamental processes
554 underlying growth and reproduction (28-30). In this study, we used PTM-omic approaches to
555 survey changes in N-glycosylation events in the mycelia, cysts, and oospores of *Phytophthora*
556 *sojae*. Our results indicate that N-glycosylation of proteins occurs extensively throughout the life
557 stages of *Phytophthora* examined in this study. This study is the first, to our knowledge, to
558 confirm the presence of multiple N-glycosylation motifs in oomycetes. This work also
559 demonstrated the crucial role of N-glycosylation in *P. sojae*, and will ultimately provide a
560 large-scale proteomic and proteome-wide glycosylation dataset for *Phytophthora* oomycetes
561 which can be progressively interrogated for insights into the function of glycoproteins in specific
562 processes as well as in macroscopic developmental phenotypes in oomycetes. An increasingly
563 comprehensive understanding of the physiological functions of PTMs in *P. sojae* will aid in the
564 improvement of management and prevention strategies for *Phytophthora*-associated diseases, and
565 may be relevant to homologous PTM-mediated developmental processes in higher organisms.

566 **Acknowledgments**

567 This work was funded by the National Natural Science Foundation of China (31672052, 31801761)
568 and China Postdoctoral Science Foundation (2019T120160). The authors acknowledge Jingjie
569 PTM BioLabs, Inc. for technical assistance with peptide and protein identification.

570 **Data availability**

571 Mass spectrometry data have been deposited to the ProteomeXchange Consortium via PRIDE
572 with the data set identifier PXD015294.

573 **Declaration of Interest Statement**

574 The authors have declared that no competing interest exist.

References

1. Nagashima, Y., von Schaewen, A., and Koiwa, H. (2018) Function of N-glycosylation in plants. *Plant Sci.* **274**, 70–79.
2. Ohtsubo, K., and Marth, J.D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* **126**, 855–67.
3. Roth, J. (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem. Rev.* **102**, 285–303.
4. Spiro, R.G. (2002) Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* **12**, 43R–56R.
5. Zielinska, D.F., Gnad, F., Schropp, K., Wiśniewski, J.R., and Mann, M. (2012) Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell.* **46**, 542–548.
6. Van den Steen, P., Rudd, P.M., Dwek, R.A., and Opdenakker, G. (1998) Concepts and principles of O-linked glycosylation. *Crit. Rev. Biochem. Mol. Biol.* **33**, 151–208.
7. Wollscheid, B., Bausch-Fluck, D., Henderson, C., O'Brien, R., Bibel, M., Schiess, R., Aebersold, R., and Watts, J.D. (2009) Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins. *Nat. Biotechnol.* **27**, 378–386.
8. Zielinska, D.F., Gnad, F., Wiśniewski, J.R., and Mann, M. (2010) Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* **141**, 897–907.
9. Schwarz, F., Aeby, M. (2011) Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struc. Biol.* **21**, 576–582.
10. Kelleher, D.J., and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* **16**, 47–62.
11. Kelleher, D.J., Banerjee, S., Cura, A.J., Samuelson, J., and Gilmore, R. (2007) Dolichol-linked oligosaccharide selection by the oligosaccharyltransferase in protist and fungal organisms. *J. Cell Biol.* **177**, 29–37.
12. Poljak, K., Selevsek, N., Ngwa, E., Grossmann, J., Losfeld, M.E., and Aeby, M. (2018) Quantitative profiling of N-linked glycosylation machinery in yeast *Saccharomyces cerevisiae*. *Mol. Cell Proteomics* **17**, 18–30.
13. Yu, L.Y., He, H.B., Hu, Z.F., and Ma, Z.J. (2016) Comprehensive quantification of N-glycoproteome in *Fusarium graminearum* reveals intensive glycosylation changes against fungicide. *J. Proteomics.* **142**, 82–90.
14. Hanson, S.R., Culyba, E.K., Hsu, T.L., Wong, C.H., Kelly, J.W., and Powers, E.T. (2009) The core trisaccharide of an N-linked glycoprotein intrinsically accelerates folding and enhances stability. *Proc. Natl. Acad. Sci. USA* **106**, 3131–3136.
15. Shental-Bechor, D., and Levy, Y. (2008) Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. USA* **105**, 8256–8261.
16. Culyba, E.K., Price, J.L., Hanson, S.R., Dhar, A., Wong, C.H., Gruebele, M., Powers, W.T., and Kelly, J.W. (2011) Protein native-state stabilization by placing aromatic side chains in N-glycosylated reverse turns. *Science* **331**, 571–575.
17. Aeby, M. (2013) N-linked protein glycosylation in the ER. *BBA-Mol. Cell Res.* **1833**, 2430–2437.
18. Wheatley, M., and Hawtin, S.R. (1999) Glycosylation of G-protein-coupled receptors for hormones central to

614 normal reproductive functioning: its occurrence and role. *Hum. Reprod. Update* **5**, 356–364.

615 19. Tan, H., Liang, W., Hu, J., and Zhang, D. (2012) MTR1 encodes a secretory fasciclin glycoprotein required for
616 male reproductive development in rice. *Dev. Cell* **22**, 1127–1137.

617 20. Marada, S., Navarro, G., Truong, A., Stewart, D.P., Arensdorf, A.M., Nachtergaele, S., Angelats, E., Opferman,
618 J.T., Rohatgi, R., McCormick, P.J., and Ogden, S.K. (2015) Functional divergence in the role of N-linked
619 glycosylation in smoothened signaling. *PLoS Genet.* **11**, e1005473.

620 21. Travers, K.J., Patil, C.K., Wodicka, L., Lockhart, D.J., Weissman, J.S., and Walter, P. (2000) Functional and
621 genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated
622 degradation. *Cell* **101**, 249–258.

623 22. Zattas, D., and Hochstrasser, M. (2015) Ubiquitin-dependent protein degradation at the yeast endoplasmic
624 reticulum and nuclear envelope. *Crit. Rev. Biochem. Mol. Biol.* **50**, 1–17.

625 23. Chapman, R., Sidrauski, C., and Walter, P. (1998) Intracellular signaling from the endoplasmic reticulum to the
626 nucleus. *Annu. Rev. Cell Dev. Biol.* **14**, 459–485.

627 24. Baldauf, S.L. (2003) The deep roots of eukaryotes. *Science* **300**, 1703–1706.

628 25. Kamoun, S., Furzer, O., Jones, J.D., Judelson, H.S., Ali, G.S., Dalio, R.J., Roy, S.G., Schena, L., Zambounis,
629 A., Panabières, F., Cahill, D., Ruocco, M., Figueiredo, A., Chen, X.R., Hulvey, J., et al. (2015) The top 10
630 oomycete pathogens in molecular plant pathology. *Mol. Plant Pathol.* **16**, 413–434.

631 26. Pennisi, E. (2010) Armed and dangerous. *Science* **327**, 804–805.

632 27. Tyler, B.M. (2007) *Phytophthora sojae*: root rot pathogen of soybean and model oomycete. *Mol. Plant Pathol.*
633 **8**, 1–8.

634 28. Cvitanich, C., and Judelson, H.S. (2003) A gene expressed during sexual and asexual sporulation in
635 *Phytophthora infestans* is a member of the Puf family of translational regulators. *Eukaryot Cell* **2**, 465–473.

636 29. Niu, X.F. (2010) Gene regulatory machinery and proteomics of sexual reproduction in *Phytophthora infestans*.
637 Diss UC Riverside.

638 30. Prakob, W., and Judelson, H.S. (2007) Gene expression during oosporogenesis in heterothallic and homothallic
639 *Phytophthora*. *Fungal Genet. Biol.* **44**, 726–739.

640 31. Pang, Z.L., Srivastava, V., Liu, X.L., and Bulone, V. (2017) Quantitative proteomics links metabolic pathways
641 to specific developmental stages of the plant-pathogenic oomycete *Phytophthora capsici*. *Mol. Plant Pathol.* **18**,
642 378–390.

643 32. Niu, X., Ah-Fong, A.M., Lopez, L.A., and Judelson, H.S. (2018) Transcriptomic and proteomic analysis reveals
644 wall-associated and glucan-degrading proteins with potential roles in *Phytophthora infestans* sexual spore
645 development. *PLoS One* **13**, e0198186.

646 33. Xiao, H.P., Smeeckens, J.M., and Wu, R.H. (2016) Quantification of tunicamycin-induced protein expression
647 and N-glycosylation changes in yeast. *Analyst* **141**, 3737–3745.

648 34. Cai, M., Miao, J.Q., Song, X., Lin, D., Bi, Y., Chen, L., Liu, X.L., and Tyler, B.M. (2016) C239S mutation in
649 the β -Tubulin of *Phytophthora sojae* confers resistance to zoxamide. *Front. Microbiol.* **7**, 762.

650 35. Miao, J.Q., Dong, X., Lin, D., Wang, Q.S., Liu, P.F., Chen, F.R., Du, Y.X., and Liu, X.L. (2016) Activity of the
651 novel fungicide oxathiapiprolin against plant-pathogenic oomycetes. *Pest Manag. Sci.* **72**, 1572–1577.

652 36. Li, A.N., Zhang, M., Wang, Y.L., Li, D.L., Liu, X.Y., Tao, K., Ye, W.W., Wang, Y.C. (2014) PsMPK1, an

SLT2-type mitogen-activated protein kinase, is required for hyphal growth, zoosporeogenesis, cell wall integrity, and pathogenicity in *Phytophthora sojae*. *Fungal Genet. Biol.* **65**, 14–24.

37. Zhu, J., Wang, F.J., Chen, R., Cheng, K., Xu, B., Guo, Z.M., Liang, X.M., Ye, M.L., and Zou, H.F. (2012) Centrifugation assisted microreactor enables facile integration of trypsin digestion, hydrophilic interaction chromatography enrichment, and on-column deglycosylation for rapid and sensitive N-glycoproteome analysis. *Anal. Chem.* **84**, 5146–5153.

38. Zhu J, Sun Z, Cheng K, Chen R, Ye ML, Xu B, Sun, D.G., Wang, L.M., Liu, J., Wang, F.J., and Zou, H.F. (2014) Comprehensive mapping of protein N-glycosylation in human liver by combining hydrophilic interaction chromatography and hydrazide chemistry. *J Proteome Res.* **13**, 1713–1721.

39. Cox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V., and Mann, M. (2009) A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **4**, 698.

40. Schwartz, D., and Gygi, S.P. (2005) An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* **23**, 1391–1398.

41. Klausen, M.S., Jespersen, M.C., Nielsen, H., Jensen, K.K., Jurtz, V.I., Soenderby, C.K., Sommer, M.O.A., Nielsen, O.W.M., Petersen, B., and Marcatili, P. (2019) NetSurfP-2.0: Improved prediction of protein structural features by integrated deep learning. *Proteins* **87**, 520–527.

42. Gene Ontology Consortium. (2014) The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, 258–261.

43. Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., and Lopez, R. (2005) InterProScan: protein domains identifier. *Nucleic Acids Res.* **33**, 116–120.

44. Horton, P., Park, K.J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007) WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* **35**, W585–W587.

45. Fukushima, Y., Oka, Y., Saitoh, T., Katagiri, H., Asano, T., Matsushashi, N., Takata, K., Breda, E.V., Yazaki, Y., and Sugano, K. (1995) Structural and functional analysis of the canine histamine H2 receptor by site-directed mutagenesis: N-glycosylation is not vital for its action. *Biochem. J.* **310**, 553–558.

46. Fang, Y.F., Cui, L.K., Gu, B., Arredondo, F., and Tyler, B.M. (2017) Efficient genome editing in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Curr. Protoc. Microbiol.* **44**:21A–1.

47. Fang, Y.F., and Tyler, B.M. (2016) Efficient disruption and replacement of an effector gene in the oomycete *Phytophthora sojae* using CRISPR/Cas9. *Mol. Plant Pathol.* **17**, 127–139.

48. Miao, J.Q., Chi, Y.D., Lin, D., Tyler, B.M., and Liu, X.L. (2018) Mutations in ORP1 conferring oxathiapiprolin resistance confirmed by genome editing using CRISPR/Cas9 in *Phytophthora capsici* and *P. sojae*. *Phytopathology* **108**, 1412–1419.

49. Ruiz-May, E., Hucko, S., Howe, K.J., Zhang, S., Sherwood, R.W., Thannhauser, T.W., and Rose, J.K.C. (2014) A comparative study of lectin affinity based plant N-glycoproteome profiling using tomato fruit as a model. *Mol. Cell Proteomics* **13**, 566–579.

50. Norris, G.E., Stillman, T.J., Anderson, B.F., and Baker, E.N. (1994) The three-dimensional structure of PNGase F, a glycosyl asparaginase from *Flavobacterium meningosepticum*. *Structure* **2**, 1049–1059.

51. Woods, R.J., Edge, C.J., and Dwek, R.A. (1994) Protein surface oligosaccharides and protein function. *Nat. Struct. Biol.* **1**, 499–501.

692 52. Mohorko, E., Glockshuber, R., and Aeby, M. (2011) Oligosaccharyltransferase: the central enzyme of N-linked
693 protein glycosylation. *J. Inherit. Metab. Dis.* **34**, 869–878.

694 53. Kaji, H., Shikanai, T., Sasaki-Sawa, A., Wen, H., Fujita, M., Suzuki, Y., Sugahara, D., Sawaki, H., Yamauchi,
695 Y., Shinkawa, T., Taoka, M., Takahashi, N., Isobe, T., and Narimatsu, H. (2012) Large-scale identification of
696 N-glycosylated proteins of mouse tissues and construction of a glycoprotein database, *J. Proteome Res.* **11**, 4553–
697 4566.

698 54. Choi, K.J., Grass, S., Paek, S., Geme III, J.W.S., and Yeo, H.J. (2010) The *Actinobacillus pleuropneumoniae*
699 HMW1C-like glycosyltransferase mediates N-linked glycosylation of the *Haemophilus influenzae* HMW1 adhesin.
700 *PLoS One* **5**, e15888.

701 55. Grass, S., Buscher, A.Z., Swords, W.E., Apicella, M.A., Barenkamp, S.J., and Ozchlewski, N. (2003) The
702 *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1C and
703 phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. *Mol. Microbiol.* **48**, 737–751.

704 56. Grass, S., Lichti, C.F., Townsend, R.R., Gross, J., and Geme III, J.W.S. (2010) The *Haemophilus influenzae*
705 HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin.
706 *PLoS Pathog.* **6**, e1000919.

707 57. Gross, J., Grass, S., Davis, A.E., Gilmore-Erdmann, P., Townsend, R.R., and Geme, J.W.S. (2008) The
708 *Haemophilus influenzae* HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. *J.*
709 *Biol. Chem.* **283**, 26010–26015.

710 58. Blanco, F.A., and Judelson, H.S. (2005) A bZIP transcription factor from *Phytophthora* interacts with a protein
711 kinase and is required for zoospore motility and plant infection. *Mol. Microbiol.* **56**, 638–648.

712 59. Warburton, A.J., and Deacon, J.W. (1998) Transmembrane Ca^{2+} fluxes associated with zoospore encystment
713 and cyst germination by the phytopathogen *Phytophthora parasitica*. *Fungal Genet. Biol.* **25**, 54–62.

714 60. McConville, M.J., and Menon, A.K. (2000) Recent developments in the cell biology and biochemistry of
715 glycosylphosphatidylinositol lipids. *Mol. Membr. Biol.* **17**:1–16.

716 61. Ohishi, K., Inoue, N., Maeda, Y., Takeda, J., Riezman, H., and Kinoshita, T. (2000) Gaa1p and gpi8p are
717 components of a glycosylphosphatidylinositol (GPI) transamidase that mediates attachment of GPI to proteins. *Mol.*
718 *Biol. Cell.* **11**, 1523–1533.

719 62. Udenfriend, S., and Kodukula, K. (1995) How glycosyl-phosphatidylinositol-anchored membrane proteins are
720 made. *Annu. Rev. Biochem.* **64**, 563–591.

721 63. Tiede, A., Bastisch, I., Schubert, J., Orlean, P., and Schmidt, R.E. (1999) Biosynthesis of glycosylphosphatidyl
722 -inositols in mammals and unicellular microbes. *Biol. Chem.* **380**, 503–524.

723 64. Ikezawa, H. (2002) Glycosylphosphatidylinositol (GPI)-anchored proteins. *Biol. Pharm. Bull.* **25**, 409–417.

724 65. Meyer, U., Benghezal, M., Imhof, I., and Conzelmann, A. (2000) Active site determination of Gpi8p, a
725 caspase-related enzyme required for glycosylphosphatidylinositol anchor addition to proteins. *Biochemistry* **39**,
726 3461–3471.

727 66. Liu, Q., Zhao, Y., Zheng, L., Zhu, X., Cui, L., and Cao, Y. (2018) The GPI transamidase complex subunit
728 PbGPI16 of *Plasmodium berghei* is important for inducing experimental cerebral malaria. *Infect. Immun.* AI–
729 00929.

730 67. Gupta, R.S., and Singh, B. (1994) Phylogenetic analysis of 70 Kd heat shock protein sequences suggest a

731 chimeric origin for the eukaryotic cell nucleus. *Curr. Biol.* **4**, 1104–1114.

732 68. Hunt, C., and Morimoto, R.I. (1985) Conserved features of eucaryotic hsp70 genes revealed by comparison
733 with the nucleotide sequence of human hsp70. *Proc. Natl. Acad. Sci. USA* **82**, 6455–6459.

734 69. Imperiali, B., and O'Connor, S.E. (1999) Effect of N-linked glycosylation on glycopeptide and glycoprotein
735 structure. *Curr. Opin. Chem. Biol.* **3**, 643–649.

736 70. DeKoster, G.T., and Robertson, A.D. (1997) Thermodynamics of unfolding for Kazal-type serine protease
737 inhibitors: entropic stabilization of ovomucoid first domain by glycosylation. *Biochemistry* **36**, 2323–2331.

738 71. Wang, C., Eufemi, M., Turano, C., and Giartosio, A. (1996) Influence of the carbohydrate moiety on the
739 stability of glycoproteins. *Biochemistry* **35**, 7299–7307.

740 72. Qi, F.F., Zhang, W.X., Zhang, F.J., Chen, G.J., and Liu, W.F. (2014) Deciphering the effect of the different
741 N-glycosylation sites on the secretion, activity, and stability of cellobiohydrolase I from *Trichoderma reesei*. *Appl.*
742 *Environ. Microbiol.* **80**, 3962–3971.

743 73. Srivastava, V., Rezinciuc, S., and Bulone, V. (2018) Quantitative proteomic analysis of four developmental
744 stages of *Saprolegnia parasitica*. *Front. Microbiol.* **8**, 2658.

745 74. Fan, B., Li, Y.L., Li, L., Peng, X.J., Bu, C., Wu, X.Q., Borriss, R. (2017) Malonylome analysis of
746 rhizobacterium *Bacillus amyloliquefaciens* FZB42 reveals involvement of lysine malonylation in polyketide
747 synthesis and plant-bacteria interactions. *J. Proteomics*. **154**, 1–12.

748 75. Okanishi, H., Kim, K., Masui, R., and Kuramitsu, S. (2014) Lysine propionylation is a prevalent
749 post-translational modification in *Thermus thermophilus*. *Mol. Cell Proteomics* **13**, 2382–238.

750

751 TABLE LEGENDS

752 **Table 1 Annotation and recognition sequences of two candidate glycoproteins with increased**
753 **abundance in cysts and oospores.**

754 FIGURE LEGENDS

755 **Fig. 1 Effect of N-glycosylation inhibitor tunicamycin on the growth of *P. sojae*.** 2.5 µg/ml and
756 5 µg/ml of tunicamycin was added to medium, and changes in *P. sojae* colony size, cyst
757 germination, and oospore production were recorded. (A) The colony diameters of *P. sojae* cultured
758 on V8 medium for five days. (B) Cysts were treated with tunicamycin germination was
759 determined after 8-12 hours. (C) Tunicamycin was added to the 10% V8 liquid medium following
760 4 days of incubation of *P. sojae* in the dark at 25°C. Oospore production was determined after 5
761 days of incubation. Bars represent the mean ± standard error (SD) from at least three biological
762 replicates, and three repeated experiments. An asterisk indicates significant differences based on
763 unpaired Student's *t* test with the *P* values marked (**P* < 0.05; ***P* < 0.01, ****P* < 0.001).

764 **Fig. 2 Properties of N-glycosylated peptides in *P. sojae*.** (A) Motif analysis of N-glycosylation
765 peptides in *P. sojae*. Sequence models comprising amino acids in specific positions of
766 modified-21-mers (10 amino acids upstream and downstream of the site) of all of the protein
767 sequences. All of the database protein sequences were used as the background database parameter.
768 Both the foreground and background histograms are shown on the left; (B) Heatmap showing the
769 frequency of different amino acid residues around glycosylated Asn residues in *P. sojae*; (C)
770 Distribution of glycosylated Asn and non-glycosylated Asn in protein secondary structures. The
771 mean secondary structure probabilities of modified Asn residues were compared against the mean
772 secondary structure probabilities of the dataset containing non-modified Asn residues from all
773 N-glycosylated proteins identified in this study. *P* values were calculated using non-paired Wilcox
774 tests.

775 **Fig. 3 Subcellular prediction and functional classifications of N-glycosylation proteins in *P.***
776 ***sojae*.** (A) Subcellular localization of the identified N-glycosylation proteins represented as a pie
777 chart; (B) Subcellular localization of the N-glycosylation proteins carrying N_{glyco}-X-S/T motif
778 represented as a pie chart; (C) Subcellular localization of the N-glycosylation proteins carrying
779 N_{glyco}-A, N_{glyco}-N, N_{glyco}-D, and N_{glyco}-S motifs represented as a pie chart; (D) GO-based
780 enrichment analysis of identified proteins in *P. sojae*. A two-tailed Fisher's exact test was
781 employed to test the enrichment of the identified proteins in each category against all of the
782 database proteins. GOs with corrected *P* values < 0.05 were considered significantly enriched; (E)
783 KEGG pathway-based enrichment analysis of N-glycosylated proteins in *P. sojae*. Analysis of the
784 enriched, identified proteins compared to all KEGG database proteins was used to determine
785 enriched pathways via two-tailed Fisher's exact tests. Pathways with corrected *P* values < 0.05
786 were considered significantly enriched.

Fig. 4 The GPI transamidase (GPI16) with a N_{glyco}-X-S/T motif and heat shock protein 70 (HSP70) with a N_{glyco}-N motif are both N-glycosylated in *P. sojae*. Total protein extracts were treated with PNGase F or buffer only before immunoblotting; Anti-FLAG antibodies were used for GPI16 (A) and HSP70 (B). Deglycosylation analysis was conducted for the GPI16^{N94A} and HSP70^{N270A} mutants, each carrying a point mutation at the Asn 94 for GPI16 (C) and Asn 94 for HSP70 (D) respective glycosylation sites. The conservation of glycosylation site on the two target proteins GPI16 (E) and HSP70 (F) were analyzed in oomycetes, as indicated by PHYSODRAFT for *P. sojae*, PSURA for *P. ramorum*, PTPG for *P. infestans*, PPTG for *P. parasitica*, and PHYCA for *P. capisici*, EPrPVT for *Pythium vexans*, and PIW for *Py. iwayamai*, respectively.

Fig. 5 The characteristics of *P. sojae* mutants compared to the wild-type controls. The parental wild-type isolate was PS6; EV was the control isolate of *P. sojae* transformed with empty vector; M16 and M29 were mutants of GPI16 with a point mutation at the Asn 94 (A-D) glycosylation site; M12 and M15 mutant lines carried an HSP70 variant with a point mutation in the Asn 270 glycosylation site (E-H). The mycelial growth (A and E), rate of cyst germination (B and F), the oospore production (C and G), and lesion length on soybean seedlings (D and H) in mutant, WT or control isolates. Bars represent the mean $\pm$ standard error (SD) from three biological replicates, and three repeated experiments. An asterisk indicates significant differences based on unpaired Student's *t* test with the *P* values marked (**P* < 0.05; ***P* < 0.01, ****P* < 0.001).

Fig. 6 Transcript levels of five unfolded protein response (UPR) and five ER-associated degradation (ERAD) target genes of *P. sojae* in GPI16 and HSP70 mutants. The WT parental strain of *P. sojae* was PS6; EV was the *P. sojae* isolate transformed with an empty vector; M16 was the glycosite mutant of GPI16 with a point mutation at Asn 94 (A-E); M12 was the glycosite mutant of HSP70 with a point mutation at Asn 270 (F-J). (A-B) two major ER chaperones and foldase genes in the UPR pathway; (C-E) four protein disulfide isomerase genes in the UPR pathway; (F-J) five genes in the ERAD pathway. Columns and bars indicate means $\pm$ standard deviation (SD) from three biological replicates, and three repeated experiments. An asterisk indicates significant differences based on unpaired Student's *t* test with the *P* values marked (**P* < 0.05; ***P* < 0.01, ****P* < 0.001).

815

Table 1. Annotation and recognition sequences of two candidate glycoproteins with increased abundance in cysts and oospores.

Protein accession ^a	Protein description	Protein Name ^b	Position ^c	Modified sequence ^d	Cy vs My ^e	Oo vs Cy	Oo vs My
G4Z0S6	GPI transamidase component	GPI16	94	YGSAQIFN(1)R	2.00	0.18	0.35
G5A5V6	Heat shock 70 kDa protein	HSP70	270 305	AETTISN(0.997)N(0.003)R N(1)DLEQFIYR	1.38 1.13	0.87 0.57	1.20 0.65

^a The accession number of candidate proteins in JGI database.

^b The abbreviated name corresponding to each glycoprotein used in this study.

^c Position means the N-glycosylation site of each candidate protein.

^d The number after the amino acid Asn (N) indicates the localization probability of each N-glycosylation site.

^e The comparison ratio of N-glycosylation levels in the two comparison groups.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6