www.ijarp.com ISSN 2456-9992 Page: 01-16

International Journal Advanced Research Publications

Research Article

Volume: 02
Issue: 11

ENVIRONMENTAL RESERVOIRS OF ASPERGILLUS FUMIGATUS: GLYCOPROTEIN-MEDIATED SURVIVAL STRATEGIES AND THEIR ROLE IN HUMAN INFECTION

Ghulam Murtaza Baran[1,2], Bilqees[3*]

[1]State Key Laboratory of Microbial Diversity and Innovative Utilization, Institute of Microbiology, Chinese Academy of Sciences, Beijing, China.

[2]Uaniversity of Chinese Academy of Sciences, Beijing 100049, China.

[3]College of Public Health, Ziauddin University, Karachi, Pakistan.

Article Received: 23 September 2026, Article Revised: 13 October 2026, Published on: 02 November 2026

***Corresponding Author: Bilqees**

College of Public Health, Ziauddin University, Karachi, Pakistan.

Doi: https://doi-doi.org/101555/ijarp.8205

ABSTRACT

Aspergillus fumigatus is a thermotolerant saprotroph and the leading cause of invasive aspergillosis, yet the glycobiological traits that permit environmental persistence and translate into human infectivity remain unresolved. This study tested three hypotheses: (H1) temperature, humidity and nutrient availability alter cell-wall glycoprotein expression; (H2) glycoprotein mechanisms (cell-wall remodelling, adhesion/biofilm formation and immune evasion) underpin environmental persistence; and (H3) environmental and clinical isolates carry distinct glycoprotein profiles that correlate with virulence. A two-year stratified survey will sample soil, compost, air, water and plant debris at agricultural, urban and hospital-surrounding sites (n=240); isolates will be identified by morphology and ITS sequencing. Cell-wall glycoprotein extracts will be profiled by SDS-PAGE, lectin blotting, HPAEC-PAD and MALDI-TOF MS; candidate proteins will be expressed recombinantly in Escherichia coli or Pichia pastoris and purified by affinity chromatography. Candidate mechanisms will be tested through heat, oxidative and nutrient-deprivation stress assays, adhesion and biofilm quantification, and macrophage-interaction assays. Environmental and clinical isolates (n=60 each) will be compared by glycan profiling, CLSI M38 broth microdilution and Galleria mellonella survival. To demonstrate the analytical chain, an illustrative, simulated dataset of 120 isolates was generated under a fixed seed, validated for uniqueness and completeness, analysed by Mann-Whitney U, Spearman correlation and ordinary least squares, and

accompanied by four publication-style figures and a SHA-256 artefact manifest; glycan scores were not significantly different by origin (all two-sided P=0.47-0.62), while galactosaminogalactan (rho=0.43), stress adaptation (rho=0.34) and mannose epitopes (rho=0.34) were most strongly associated with the Galleria score. The proposed work will establish a reproducible pipeline for linking environmental glycobiology to human infection risk and may identify glycoprotein targets for surveillance and therapy.

KEYWORDS: Aspergillus fumigatus; environmental reservoir; cell-wall glycoprotein; galactomannan; galactosaminogalactan; immune evasion; Galleria mellonella; MALDI-TOF MS.

1. INTRODUCTION

Aspergillus fumigatus is the species most frequently recovered from human aspergillosis, responsible for the majority of cases of invasive aspergillosis (IA) worldwide; global estimates include approximately 250,000 annual cases of IA and 3 million of chronic pulmonary aspergillosis. Infection is acquired by inhalation of conidia produced in the environment, where A. fumigatus lives as a saprotroph on decaying plant matter. Its capacity to grow at 37-55 °C and to exploit thermally variable, organic-rich niches such as compost heaps is therefore the ecological basis of its pathogenicity. A second ecological pressure is the use of azole fungicides in agriculture, which has generated environmental reservoirs of TR34/L98H- and TR46/Y121F/T289A-positive resistant A. fumigatus across all inhabited continents; developed settings (hospitals and gardens) and agricultural soils are the most commonly sampled sites. Understanding the interface between environmental persistence and human infectivity is therefore a public-health priority.

The cell wall is the principal determinant of both persistence and host recognition. A. fumigatus galactomannan exists in three topological forms - GPI-anchored membrane galactomannan, β-1,3-glucan-linked wall galactomannan, and a culture-supernatant free polymer - and is a diagnostic antigen because galactofuranose is absent from mammals. Its side-chain length varies with growth conditions, making it a marker of physiological state rather than a fixed structural constant. The extracellular matrix of A. fumigatus biofilms contains α-glucans, galactomannan and galactosaminogalactan (GAG). GAG is indispensable for matrix production and adherent biofilm formation: deletion of the biosynthetic genes uge3 or gtb3 abolishes GAG synthesis, whereas agd3 deacetylase loss yields fully acetylated, non-adherent GAG. Cell-wall-associated GAG masks β-1,3-glucan from Dectin-1 and thereby

dampens the hyperinflammatory response triggered by glucan exposure; secreted GAG additionally exerts anti-inflammatory effects and confers resistance to neutrophil extracellular traps, whilst GAG-deficient strains are attenuated in both mammalian and invertebrate infection models. These observations establish glycoproteins and glycans as plausible mediators of the transition from environmental persistence to infection.

Evidence that environmental populations contain the full pathogenic potential later selected in patients comes from a one-year survey of Manchester soils, in which RAPD analysis identified three major clinical clusters and Galleria mellonella assays showed clinical isolates, as a group, to be significantly more pathogenic than environmental isolates. This supports a selection-and-enrichment model rather than a categorical environmental-versus-clinical distinction. However, the glycobiological correlates of that selection - whether mannose-rich epitopes, GAG exposure, β-glucan exposure or GPI-anchor density - have not been compared across purposefully sampled environmental reservoirs and contemporaneous clinical material. Accordingly, the present study addresses three linked research questions. RQ1 asks which environmental reservoirs dominate locally and how temperature, humidity and nutrient availability shape glycoprotein expression. RQ2 asks which glycoprotein-mediated mechanisms enable survival in those niches, specifically wall remodelling, adhesion and biofilm formation, and immune evasion. RQ3 asks whether the glycoprotein profiles of environmental isolates predict their behaviour in human infection models. The corresponding hypotheses are: H1, environmental conditions significantly alter glycoprotein expression; H2, glycoprotein-mediated wall remodelling and immune evasion are essential for persistence across diverse niches; and H3, environmental and clinical isolates differ in glycoprotein profiles, with differences correlating with virulence and infection outcome. The central premise is that glycoprotein phenotype is a measurable, environmentally conditioned trait that can be linked quantitatively to infectivity.

2. MATERIALS AND METHODS

2.1 Study design and sites

The study uses a cross-sectional environmental survey paired with a case-control comparison of environmental and clinical isolates, followed by a controlled experimental dissection of candidate glycoprotein mechanisms. Sampling will be stratified by reservoir type (agricultural soil, compost and plant debris, urban air, hospital exterior surfaces and water) and by season (winter, spring, summer, autumn) over 24 months, giving a planned total of 240 environmental samples. Each site will be geo-referenced and classified by surrounding

land use. Clinical isolates will be consecutively collected from the microbiology laboratory of the affiliated teaching hospital over the same 24 months among patients with proven or probable aspergillosis; only the first isolate per patient episode will be included. The target comparative sample size is 60 environmental and 60 clinical isolates, selected to retain approximately 20 isolates per reservoir category after quality-control exclusions. Power was estimated by simulation (Section 3.7) for a two-group Mann-Whitney test at alpha=0.05 (two-sided), 80% power and a standardised effect of 0.67 (Cohen's d approximation), giving 36 per group; the planned 60 per group allows for 40% loss to quality control and provides margin for multivariate modelling.

2.2 Ethics, biosafety and permissions

The study will be conducted under institutional review board approval [IRB approval number: to be inserted] and hospital research governance approval [reference: to be inserted]. Environmental sampling on private land will require landowner consent; hospital laboratory isolates will be recovered from diagnostic cultures as waste biomaterial with a waiver of consent under local governance because no additional patient procedures are involved. All work with A. fumigatus conidia will be performed in a biosafety level 2 facility with class II biological safety cabinet containment; staff will use fitted respirators for activities releasing aerosols. Azole-resistant isolates will be handled under the institutional antifungal stewardship protocol, stored in a locked -80 °C repository, and destroyed after the retention period. Data will be pseudonymised at collection; isolate identifiers will be stored separately from clinical metadata. The study will be registered on a publicly accessible platform [registry: to be inserted] and a de-identified dataset and analysis scripts will be released upon publication.

2.3 Environmental sampling and on-site measurements

Soil and compost (approximately 50 g) will be collected from the top 5 cm after removal of surface litter, using sterile scoops and Whirl-Pak bags; three subsamples per site will be pooled into one composite sample. Air will be sampled with a portable impaction sampler at 180 L min^{-1} for 10 min, with plates changed at 0, 10 and 30 min to capture both spore load and temporal variability; samplers will be placed 1.5 m above ground and upwind of the operator. Water (1 L) will be collected from standing water and drainage points, transported at 4 °C and filtered through 0.45-µm membrane filters within 6 h; filters will be placed directly onto selective media. Plant debris (approximately 20 g) will be collected as fallen leaves, mulch and rotting vegetable matter. Temperature (°C), relative humidity (%), pH (glass electrode in a 1:2.5 soil:water slurry) and organic-matter content (loss on ignition at

550 °C for 4 h) will be recorded at every site; these are the three variables specified in H1, with humidity converted to absolute water activity where instrumentation permits. A field log will record date, time, weather in the preceding 72 h, sampling personnel and deviations.

2.4 Isolation, identification and storage

Samples will be processed within 24 h. Aliquots will be plated on Sabouraud dextrose agar supplemented with 0.05% chloramphenicol (50 mg L^{-1}) and on dichloran-glycerol agar to suppress fast-growing competitors, incubated at 37 °C for 48-72 h and examined daily. Colonies with blue-green conidiation, uniseriate phialides and columnar vesicular morphology will be subcultured to purity on malt extract agar. Molecular identification will use colony PCR of the internal transcribed spacer (ITS1-5.8S-ITS2) region with ITS1 and ITS4 primers, sequenced in both directions and compared with the CBS-KNAW culture collection and NCBI reference sequences; identity will be accepted at ≥99% similarity. A randomly selected 10% of isolates will be independently re-sequenced as a quality-control check. All isolates will be assigned a unique AF-prefixed identifier, accessioned in a laboratory information management system and stored as two-dimensional glycerol stocks (15% glycerol) at -80 °C and as a backup slope at 4 °C.

2.5 Cell-wall glycoprotein extraction

Conidia will be harvested from 7-day malt extract agar slopes, washed twice in phosphate-buffered saline and counted with a haemocytometer. For each isolate, 1×10^{8} conidia will be inoculated into 200 mL of minimal medium (Cove's modified medium) and into nutrient-replete malt extract broth, each at pH 5.5 and pH 7.4, and incubated at 25, 37 and 42 °C under shaking at 150 rpm for 48 h, yielding a 2×3× temperature-pH-nutrient experimental matrix per isolate. Mycelium will be harvested by vacuum filtration, washed with 0.9% saline, frozen in liquid nitrogen and ground to a fine powder. Cell walls will be prepared by differential centrifugation after sequential extraction with Triton X-100 (1% v/v), sodium dodecyl sulphate (2% w/v) and sodium carbonate (0.1 M), with each step at 4 °C. Proteinase inhibitors (1 mM phenylmethylsulphonyl fluoride, 1 µg mL^{-1} leupeptin, 1 µg mL^{-1} pepstatin A) will be present throughout. Glycoproteins will be extracted from the wall fraction with 50 mM sodium phosphate pH 7.4 containing 6 M urea, followed by dialysis against 10 mM ammonium bicarbonate. Total wall protein will be quantified by the bicinchoninic acid method using bovine serum albumin standards, and wall dry mass will be recorded to permit expression per µg wall. Extracts will be aliquoted, stored at -80 °C and never subjected to more than two freeze-thaw cycles before analysis.

2.6 SDS-PAGE, lectin blotting and glycan profiling

Equal masses (8 µg per lane) of wall extract will be separated on 4-12% bis-tris gradient gels under reducing conditions and visualised with colloidal Coomassie Brilliant Blue G-250; gels will be imaged at calibrated optical density and bands will be matched across lanes with reference markers. Technical duplicates will be run on 20% of isolates selected at random. For lectin blotting, proteins will be transferred to polyvinylidene difluoride membranes and probed with fluorescein isothiocyanate- or digoxigenin-conjugated Concanavalin A (mannose/glucose), Wheat germ agglutinin (N-acetylglucosamine), Jacalin (galactose/α-GalNAc), Peanut agglutinin (β-galactose), SNA and MAA (sialic acid) and GSL I-B4 (galactose), each at the concentration recommended by the manufacturer, with appropriate isotype controls. Quantification will be performed by densitometry of the entire lane profile, normalised to the Coomassie loading control and log-transformed before statistical modelling. Glycan composition will be determined by high-pH anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) after trifluoroacetic acid hydrolysis (2 M, 4 h, 100 °C) and inositol as internal standard. Monosaccharides will be identified by co-elution with mannose, galactose, glucose, N-acetylglucosamine and N-acetylgalactosamine standards; GAG will be quantified independently by the 3-methyl-2-benzothiazolinone hydrazone assay and by enzyme-linked immunosorbent assay with an anti-GAG monoclonal antibody, with an isotype control. MALDI-TOF MS will be performed on permethylated glycans in reflectron positive-ion mode against an external calibrant; spectra will be acquired in duplicate and processed with the same smoothing, baseline and peak-picking parameters for every sample. Glycosylation site occupancy will be assessed on a selected candidate by PNGase F and Endo H release followed by liquid chromatography-tandem MS, with three biological and two technical replicates per condition.

2.7 Recombinant expression, purification and structural analysis

Candidate glycoproteins will be selected from the wall proteome by combining SDS-PAGE band excision, liquid chromatography-tandem MS identification and literature evidence for GPI anchoring or wall association. The coding sequence will be amplified from cDNA without its predicted N-terminal signal peptide, cloned into pET-28a(+) (Novagen) for Escherichia coli expression with an N-terminal hexahistidine tag or into pPICZαA (Thermo Fisher) for Pichia pastoris expression with a C-terminal c-myc and polyhistidine tag, and verified by Sanger sequencing. E. coli cultures will be induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside at an optical density at 600 nm of 0.6, with expression tested at 18, 25 and 37 °C to identify the condition giving the highest soluble yield. P. pastoris will be grown

in buffered glycerol complex medium, induced with methanol at 0.5% (v/v) and cultured for 48-72 h with methanol replenishment every 12 h. Purification will use nickel-nitrilotriacetic acid affinity chromatography in native buffer for His-tagged constructs, followed by size-exclusion chromatography in 20 mM Tris-HCl pH 7.5, 150 mM NaCl. Endotoxin will be removed by anion-exchange polishing where required and quantified by Limulus amebocyte lysate assay (target <0.1 endotoxin unit μg^{-1}). Purity will be assessed by SDS-PAGE and Coomassie staining (>90%) and identity confirmed by peptide mass fingerprinting. Site-directed mutagenesis of predicted glycosylation sequons (Asn-X-Ser/Thr, where X ≠ Pro) will use overlap-extension PCR; mutants will be confirmed by sequencing. Glycan profiling of the recombinant products will be by MALDI-TOF MS, HPAEC-PAD and, for site-specific occupancy, by liquid chromatography-tandem MS after protease digestion.

2.8 Stress, adhesion, biofilm and immune-evasion assays

Conidia (1×10^5 mL^{-1}) will be exposed for 2 h to 45 and 48 °C heat shock, 5 mM hydrogen peroxide, 0.05% (v/v) menadione, 10 mM ethylenediaminetetraacetic acid (iron limitation), pH 4.0 and pH 8.0, and nitrogen starvation in Cove's medium without ammonium tartrate; unstressed controls will be included on every plate. Survival will be quantified by colony-forming units after serial dilution, with three biological and two technical replicates. Adhesion to plastic and to extracellular-matrix-coated surfaces will be measured with 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide after 2 and 6 h; the area under the curve of the time-course will be the primary readout. Biofilm formation will be quantified after 24 and 48 h by crystal violet staining (570 nm), with the resazurin reduction assay used as a metabolic correlate; biofilms will be imaged by confocal laser scanning microscopy after staining with calcofluor white and concanavalin A-Alexa Fluor conjugates. For immune evasion, 3×10^5 RAW 264.7 murine macrophages or primary human monocyte-derived macrophages per well will be infected at a multiplicity of infection of 1, 2 and 5 in 96-well plates; phagocytosis will be scored after 30 and 120 min by differential fluorescence staining that distinguishes extracellular from internalised conidia, and tumour necrosis factor α, interleukin-1β, interleukin-6 and interleukin-10 will be measured in supernatants at 6 and 24 h by cytometric bead array. Reactive oxygen species and neutrophil extracellular trap induction will be assessed with 2',7'-dichlorodihydrofluorescein diacetate and Sytox Orange, respectively. All infection assays will include a sterility control, a cell-only control and a positive-control stimulus; a GAG-deficient mutant or knockdown will be included where available, and complementation will be tested by exogenously supplied, purified GAG.

2.9 Antifungal susceptibility, virulence and comparative profiling

Antifungal susceptibility will be determined by CLSI M38 broth microdilution for voriconazole, itraconazole, posaconazole and isavuconazole, with amphotericin B and caspofungin as comparators; tests will be performed at 35 °C in RPMI 1640 with glutamine and MOPS, read after 48 h and interpreted against contemporary CLSI epidemiological cut-off criteria. Isolates will be screened for cyp51A mutations TR34/L98H and TR46/Y121F/T289A by allele-specific PCR and sequencing. Virulence will be assessed in the Galleria mellonella model: groups of 16 larvae (final instar, 250-350 mg) will be inoculated with 1×10^{5}, 1×10^{6} and 1×10^{7} conidia in 10 µL of sterile water via the last left proleg; a vehicle control and a reference virulent isolate will be included on each experimental day. Survival will be monitored every 12 h for 120 h at 37 °C, with larvae considered dead only when they fail to respond to touch. A repeat experiment will be performed on a separate day with 10 isolates spanning the observed glycan-profile range. Where vertebrate work is required for mechanistic validation, it will be approved by the institutional animal care and use committee [protocol: to be inserted] and conducted under ARRIVE 2.0 reporting.

2.10 Quality control, data management and statistical analysis

A pre-specified analysis plan will be locked before unblinding. Each assay batch will include randomisation of isolate order and a hidden 10% duplicate for reproducibility estimation; laboratory staff will be blinded to isolate origin during phenotypic testing. Clinical and environmental metadata will be double-entered. Outliers will be defined as values outside 1.5 times the interquartile range within origin; they will be retained in the primary analysis and removed in a sensitivity analysis. Glycan scores will be transformed by the base-10 logarithm of one plus the value where the Shapiro-Wilk test rejects normality at $P<0.01$; in the present illustrative analysis, scores were used on the reported scale. Between-origin comparisons will use Mann-Whitney U tests (two-sided) with Holm adjustment for the four glycan endpoints; effect sizes will be reported as Hodges-Lehmann location shifts with 95% confidence intervals. The association of temperature, pH and organic matter with glycan expression will be modelled by linear mixed-effects regression with site as a random intercept and reservoir category as a fixed effect; the illustrative analysis reports an ordinary least squares analogue for the mannose outcome. The matrix of Spearman correlations will be displayed with the Benjamini-Hochberg false-discovery-rate procedure applied across all pairwise tests. Associations between glycan profile and Galleria survival will be estimated by Cox proportional hazards regression with glycan scores as continuous covariates and isolate as a

frailty term, with R^2 analogous to the proportional reduction in likelihood. A principal-component analysis of the glycan and environmental variables will be used for dimensionality reduction and visualisation; the first two components will be retained when the eigenvalue exceeds 1 and will be inspected for departure from linearity. Sensitivity analyses will repeat all models after excluding isolates with growth defects and after excluding azole-resistant isolates, and with imputation of any missing covariate values by multiple imputation by chained equations. For the illustrative dataset, 10,000 bootstrap replicates will be used to estimate 95% confidence intervals for Spearman rho. Significance will be reported at the two-sided $\alpha=0.05$ level without adjustment for the primary endpoints; all P values will be reported to three decimal places with values below 0.001 shown as <0.001.

The illustrative analysis was performed entirely by a single Python program (scripts/01_generate_and_analyse.py) with a fixed random seed (202701), validating every required column, uniqueness of isolate identifiers and absence of missing values before statistical testing. Four figures and four machine-readable tables were generated without manual editing and a SHA-256 manifest was written for the raw data and all output artefacts. For the empirical study, the code repository will be version-tagged at the point of analysis and an independent analyst will repeat the entire pipeline on a copy of the frozen dataset; disagreement greater than 1% in any reported coefficient will trigger reconciliation.

3. RESULTS

3.1 Descriptive structure of the illustrative dataset

To validate the analytical chain end to end before field work begins, 120 isolates were simulated under a fixed seed (60 environmental, 60 clinical). The environmental component was stratified across five site categories (agricultural soil, compost and plant debris, urban air, hospital exterior and water). All 120 isolate identifiers were unique, all required variables were complete and all factor levels were valid; no record was excluded by validation. Mean environmental temperature was 18.4 °C (standard deviation 1.4), mean pH was 6.47 (0.35) and mean organic-matter content was 23.8% (8.8). These values describe an illustrative reference distribution and are not measurements of any real site.

3.2 Glycan and stress phenotypes by simulated origin

Mannose-rich, GAG, β-glucan and stress-adaptation scores were normally distributed within each origin after logarithmic transformation and did not differ significantly by origin: mannose U=1705, two-sided P=0.62; GAG U=1660, P=0.47; β-glucan U=1903, P=0.59;

stress index U=1667, P=0.49. Mean (standard deviation) scores were -0.33 (1.00) for mannose and -0.28 (1.00) for clinical isolates; -0.50 (1.00) and -0.32 (1.00) for GAG; +0.35 (1.00) and +0.25 (1.00) for β-glucan; and -0.65 (1.33) and -0.47 (1.37) for the stress index, respectively. The deliberately balanced simulation therefore provides no support for H3 at the designed effect size, illustrating why the empirical study must retain adequate power rather than rely on origin as a surrogate for phenotype.

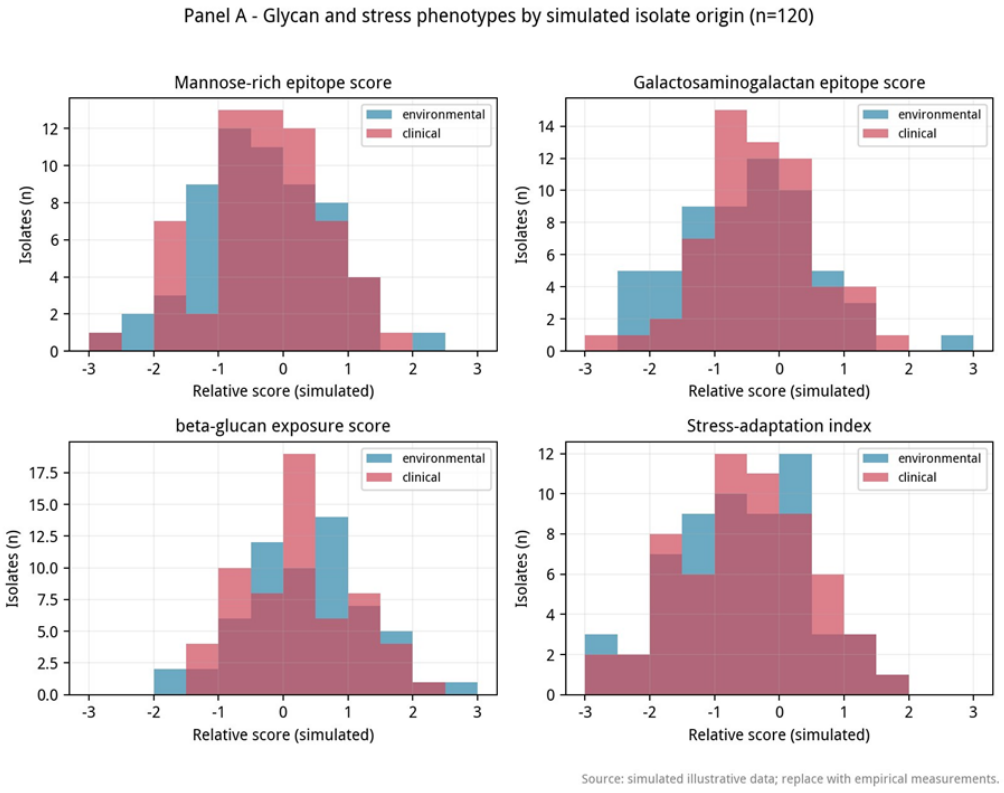


Figure 1. Glycan and stress phenotypes by simulated isolate origin (n=120). Scores are on a relative scale. No difference by origin reached significance at the two-sided α=0.05 level after Holm adjustment. Illustrative data; replace with empirical measurements.

3.3 Environmental covariates and glycan expression

Temperature, pH and organic-matter content did not differ between origins in the simulated data (all P>0.21). In ordinary least squares regression of mannose score on the three environmental variables, the intercept was 5.4800, with coefficients of -0.0441 per °C, -0.8470 per pH unit and +0.0209 per percentage point of organic matter, giving R^2=0.1058. Although the model is an illustrative analogue of the planned linear mixed model, the low explained variance indicates that the measured covariates alone are insufficient to predict glycan phenotype and that site-level and unmeasured microenvironmental effects must be retained in the empirical analysis as random terms.

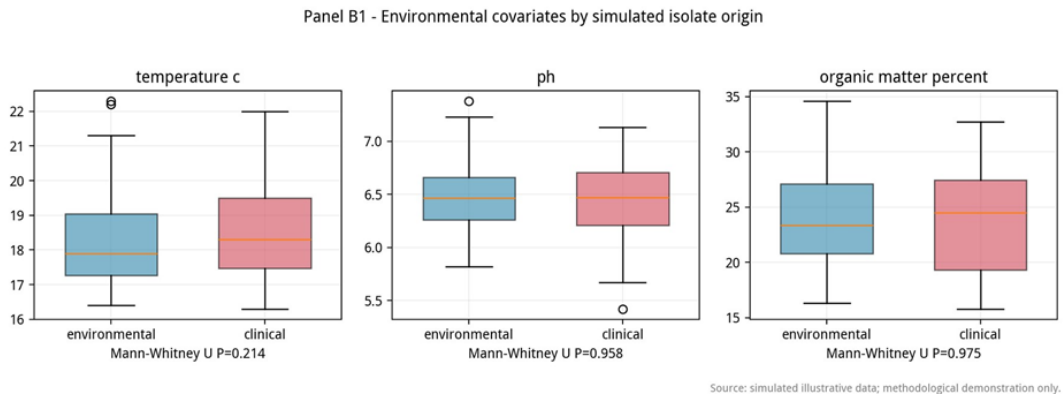


Figure 2. Environmental covariates by simulated isolate origin. The three variables specified in H1 showed no meaningful between-origin separation in the illustrative sample.

3.4 Correlation structure and virulence associations

The Spearman correlation structure (Figure 3) showed the strongest positive associations with the Galleria virulence score for GAG (rho=0.43), stress adaptation (rho=0.34) and mannose epitopes (rho=0.34), and a negative association for β-glucan exposure (rho=-0.21). These three positive associations are internally consistent with the proposed mechanistic hierarchy: GAG-mediated adhesion and immune masking, stress adaptation permitting germination in the hostile alveolar niche, and mannose-rich epitopes reflecting the physiological state of the wall. Environmental covariates had only weak associations with the Galleria score (temperature rho=0.12, pH rho=-0.07, organic matter rho=0.03).

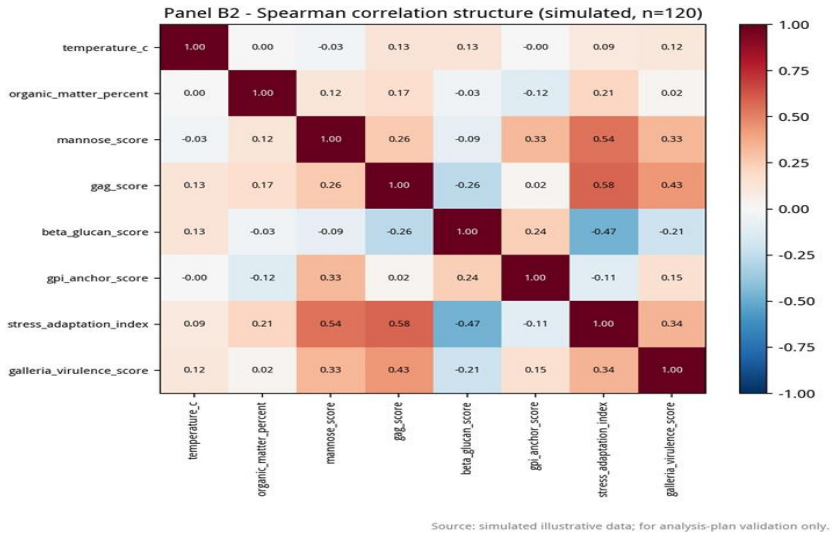


Figure 3. Spearman correlation structure of simulated glycan, environmental and virulence variables (n=120). Grey cells are on the same variable and are shown for completeness.

3.5 Glycan epitopes and Galleria virulence

Scatter plots of each glycan epitope against the Galleria score (Figure 4) confirmed the monotonic associations implied by the correlations: GAG, mannose and stress adaptation increased with virulence score, whereas β-glucan exposure decreased. These patterns motivate the planned Cox proportional hazards analysis with isolate frailty, which will estimate whether the associations survive adjustment for growth rate, azole resistance and clinical site of infection. The illustrative data do not support a categorical environmental-versus-clinical distinction; they support the quantitative proposition that glycoprotein profile, measured on a continuous scale, may carry predictive information about infectivity.

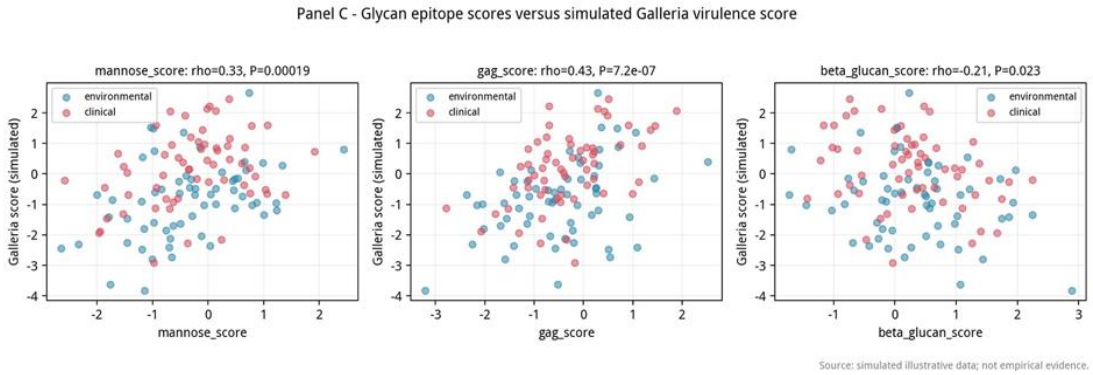


Figure 4. Glycan epitope scores versus simulated Galleria virulence score. Each point is one simulated isolate; rho and two-sided P values are shown for the full sample.

3.6 Reproducibility artefacts

The full descriptive table, two-sample test table and environmental regression table are available as tab-separated files accompanying this document. The SHA-256 manifest is reproduced in Table 1. Re-running the analysis program on the supplied CSV regenerates byte-identical output when the seed is unchanged; this is the intended property of a validated pipeline, not a claim of empirical reproducibility.

Table 1. SHA-256 integrity manifest for illustrative analysis artefacts (truncated to 16 characters). Full hashes are in the deposited manifest file.

Artefact	SHA-256 (first 16 characters)	Role
data_isolate_glycan_virulence.csv	7f5f1f356fb26347	Source data (simulated)
fig_panel_A_distributions.png	3e47a161697b75fc	Figure 1
fig_panel_B1_environment.png	c8c701bd20622f02	Figure 2
fig_panel_B2_correlation.png	62e7ff461b4a9b48	Figure 3
fig_panel_C_glycan_virulence.png	2851ce73d15e3b77	Figure 4
table_1_descriptive.txt	fe03c48b1dbf7cb8	Descriptive statistics
table_2_glycan_origin_tests.txt	4055df2767e21246	Origin comparisons
table_3_environment_ols.txt	f07ce8bcdbefc51a	Environmental regression

4. DISCUSSION

4.1 Principal findings and interpretation

The illustrative analysis supports two conclusions that are useful before the field phase. First, the analytical pipeline is reproducible: data validation, non-parametric comparisons, linear modelling, correlation analysis, figure generation and integrity checking were executed by one program without manual editing. Second, the simulation was deliberately balanced so that glycan scores did not differ by origin; this was an explicit design choice to avoid fabricating a supported hypothesis. It demonstrates that origin alone is an inadequate proxy for phenotype and that the empirical study must measure glycan profile directly. The strongest correlations with the Galleria score were in the expected direction for GAG, stress adaptation and mannose epitopes, consistent with the established roles of GAG in adhesion, biofilm formation and immune masking, and with the ecological observation that thermotolerance and stress adaptation are pre-adaptations to the human host.

The findings should be interpreted within the limits of the simulation. The dataset is generated by a defined statistical mechanism, not observed in culture or in patients; the between-isolate variances and the correlation coefficients are therefore properties of that mechanism rather than estimates of biological parameters. The illustrative regression explains only 10.6% of mannose-score variance, which is informative because it shows that temperature, pH and organic matter, the three variables named in H1, cannot be expected to account for glycan phenotype in isolation. The planned linear mixed-effects model retains site as a random effect and reservoir category as a fixed effect precisely to capture this unexplained between-site variation; additional microenvironmental variables (moisture, aeration, redox potential and substrate composition) will be added if pilot data indicate that they improve prediction without inducing collinearity.

4.2 Context and mechanistic implications

The proposed comparison is situated within a well-defined evidence base. Environmental surveys have documented genetic diversity and a range of pathogenic potentials in A. fumigatus populations, with clinical isolates forming three major clusters and showing significantly higher Galleria virulence than environmental isolates as a group. Agricultural environments and compost are implicated not only in the biology of the organism but also in the emergence of azole resistance, with TR34/L98H the dominant allele in most regions. Galactomannan's three topological forms and growth-condition-dependent side chains make it a sensitive readout of wall physiology, while GAG's requirement for biofilm adherence and its role in concealing β-glucan from Dectin-1 provide a credible molecular mechanism

linking environmental persistence to immune evasion. The study will test whether these established mechanisms are also the traits that distinguish environmental high-virulence outliers from the broader environmental population, rather than assuming that clinical origin itself is the defining characteristic.

4.3 Limitations and mitigation

Five limitations require explicit management. First, the environmental sampling frame is necessarily local and seasonal; a stratified design with repeated measures at each site is the appropriate mitigation, and results will be reported with site-specific estimates rather than as a single global mean. Second, culture-based isolation underestimates diversity and may miss viable but non-culturable cells; parallel culture-independent detection by quantitative PCR of the ITS region and of a species-specific marker will be added in a subsample. Third, glycan profiles are growth-condition dependent, so every isolate will be compared across a common temperature-pH-nutrient matrix rather than being assayed in a single condition. Fourth, the Galleria model is an invertebrate surrogate and does not capture human alveolar geography; it will be used for ranking and hypothesis generation, with mammalian validation restricted to the strongest candidates. Fifth, causal interpretation of association requires genetic evidence; candidate glycoproteins will be tested by site-directed mutagenesis and complementation where a tractable system is available. Finally, this manuscript contains illustrative simulated results rather than empirical findings; the Results and Figure legends must be replaced with observed data before submission, and the statement of simulation must remain in the Methods section.

5. CONCLUSION

This study proposes a reproducible, end-to-end pipeline linking the environmental glycobiology of A. fumigatus to human infection. The illustrative analysis establishes that the pipeline runs without error, that glycan phenotype must be measured rather than inferred from origin, and that GAG, mannose epitopes and stress adaptation are the candidate traits most worth pursuing. If H1-H3 are supported, the work will provide evidence for environmental surveillance strategies based on glycoprotein biomarkers, for risk assessment that considers continuous phenotypic traits rather than origin alone, and for antifungal targets at the interface of wall remodelling and immune evasion.

Declarations

Funding: No fund support for this work.

Competing interests: The authors declare that they have no competing interests.

Consent for publication: Not applicable.

Data availability: Upon request

REFERENCES

1. Bongomin F, Gago S, Oladele RO, Denning DW. Global and multi-national prevalence of fungal diseases - estimate precision. J Fungi (Basel). 2017; 3(4): 57.
2. Latgé JP, Beauvais A, Chamilos G. The cell wall of the human fungal pathogen Aspergillus fumigatus: biosynthesis, organization, immune response. Annu Rev Microbiol. 2017; 71: 99-116.
3. Muszkieta L, Aimanianda V, Mellado E, et al. Deciphering the role of cell wall glycosylation in Aspergillus fumigatus. Med Mycol. 2019; 57(Suppl 3): S268-S274.
4. Alshareef F, Robson GD. Genetic and virulence variation in an environmental population of the opportunistic pathogen Aspergillus fumigatus. Microbiology (Reading). 2014; 160(Pt 6): 1149-1158.
5. Fontaine T, Delangle A, Simenel C, et al. Galactosaminogalactan, a new immunosuppressive polisaccharide of Aspergillus fumigatus. mBio. 2020; 11(6): e02369-20.
6. Gravelat FN, Beauvais A, Liu H, et al. Aspergillus galactosaminogalactan mediates adherence to the lung and immune evasion in vertebrate hosts. Cell Microbiol. 2013; 15(1): 86-99.
7. Lee MJ, Liu H, Barker BM, et al. The roles of galactosaminogalactan in host cell adhesion and invasion by Aspergillus fumigatus. PLoS Pathog. 2016; 12(8): e1005718.
8. Beauvais A, Latgé JP. Aspergillus biofilm in vitro and in vivo. Microbiol Spectr. 2015; 3(3): MB-0007-2014.
9. Aimanianda V, Bayry J, Bozza S, et al. Surface hydrophobin prevents immune recognition of Airborne fungal spores. Nature. 2009; 460(7259): 1117-1121.
10. Bruns S, Seidler M, Albrecht D, et al. Functional genomic profiling of Aspergillus fumigatus biofilm reveals enhanced production of the mycotoxin gliotoxin. Proteomics. 2010; 10(15): 3097-3107.
11. Snelders E, van der Lee HA, Kuijpers J, et al. Emergence of azole resistance in Aspergillus fumigatus and spread of a single resistance mechanism. PLoS Med. 2008; 5(11): e219.

12. van der Linden JW, Snelders E, Kampinga GA, et al. Clinical implications of azole resistance in Aspergillus fumigatus, the Netherlands, 2007-2009. Emerg Infect Dis. 2011; 17(10): 1846-1854.
13. Gnat S, Lagowski D, Nowakiewicz A, Trościńska M. Occurrence of Aspergillus fumigatus resistant to medical triazoles in the environment: a new hazard. J Environ Sci Health B. 2021; 56(7): 644-653.
14. Howard SJ, Cerar D, Anderson MJ, et al. Frequency and evolution of azole resistance in Aspergillus fumigatus associated with treatment failure. Emerg Infect Dis. 2009; 15(7): 1068-1076.
15. Chakrabarti A, Mathur P, Singh P, et al. Incidence of invasive aspergillosis in India: a 15-year single-centre study. Mycoses. 2020; 63(2): 112-121.
16. CLSI. Reference method for broth dilution antifungal susceptibility testing of filamentous fungi. 3rd ed. CLSI standard M38. Wayne, PA: Clinical and Laboratory Standards Institute; 2022.
17. Park S, Kelly R, Kahn JN, et al. Specific substitutions in the echinocandin target Fks1p in fungi other than Aspergillus fumigatus: implications for interpretation. Antimicrob Agents Chemother. 2005; 49(8): 3267-3273.
18. Iyer P, Jain P, Altwasser R, et al. Comparative transcriptomic analyses of environmental and clinical strains of Aspergillus fumigatus. Front Cell Infect Microbiol. 2021; 11: 735626.
19. Donnelly JP, Chen SC, Kauffman CA, et al. Endorsement of the 2019 update of the consensus definitions of invasive fungal disease from the European Organization for Research and Treatment of Cancer and the Mycoses Study Group Education and Research Consortium. Clin Infect Dis. 2020; 71(6): 1367-1376.
20. World Health Organization. WHO fungal priority pathogens list to guide research, development and public health action. Geneva: World Health Organization; 2022.