nature communications

Article

https://doi.org/10.1038/s41467-026-71960-5

Wastewater intelligence predicts the emergence of clinically-relevant and drug-resistant *Candidozyma auris* at healthcare facilities

Received: 14 May 2025

Accepted: 3 April 2026

Published online: 18 April 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

The rapid evolution of antifungal resistance in *Candidozyma auris* (formerly *Candida auris*) presents significant challenges for conventional public health surveillance methods, particularly in detecting emergent and highly transmissible drug-resistant variants. Here, using wastewater-based epidemiology tools initially developed during the COVID-19 pandemic, we implement a high-resolution, facility-level early warning system to monitor *C. auris* infections and resistance patterns. Our evaluation across Southern Nevada demonstrates that upstream sewage monitoring at healthcare facilities provides significant sensitivity ($p < 0.001$) compared to wastewater treatment plant sampling. By combining amplicon sequencing and MALDI-TOF mass spectrometry, we identify clinically-relevant, resistance-associated variants in wastewater samples, while whole-genome sequencing reveals >90% genomic concordance between 443 wastewater-derived genomes and 2945 clinical isolates. We also detect previously unreported subclades and resistance mutations, including *FKS1* Phe635Leu and co-occurring *ERG11*/*FKS1* variants in wastewater samples up to nearly five months before their appearance in clinical settings. Transcriptomic profiling of drug-resistant isolates under antifungal and stress conditions identifies previously uncharacterized adaptation mechanisms, including differential regulation of ribosomal assembly pathways and cell cycle checkpoints. These findings highlight how wastewater intelligence can enhance traditional public health approaches for early detection and monitoring of *C. auris* outbreaks and antifungal resistance.

The global emergence of *Candidozyma auris* (formerly *Candida auris*) in healthcare settings represents a complex combination of rapid antifungal resistance development and environmental adaptability[1–7]. Since its identification in 2009, *C. auris* has rapidly evolved resistance to multiple antifungal drug classes, including azoles (e.g., fluconazole), echinocandins (e.g., micafungin), and polyenes (e.g., amphotericin B), leading to its designation as the only fungal organism classified as an urgent public health threat by the United States (U.S.) Centers for Disease Control and Prevention (CDC)[8]. In healthcare settings, this pathogen causes severe infections with mortality rates of up to 30–60% among infected patients[9]. The infections show a distinct demographic pattern, predominantly affecting older adults, especially those with compromised immune systems or those requiring indwelling medical devices, such as catheters or breathing tubes[10,11]. The pathogen's ability to develop resistance through multiple mechanisms, particularly mutations in ergosterol biosynthesis and

e-mail: daniel.gerrity@snwa.com; van.vo@unlv.edu; tin@wayne.edu; dhess@med.unr.edu; edwin.oh@unlv.edu

drug efflux pathways, has severely constrained treatment options and highlighted the need for development of alternative surveillance strategies[12,13].

The global dissemination of *C. auris* is characterized by six distinct geographical clades exhibiting unique resistance profiles and evolutionary trajectories[4,5,14]. Clade I isolates, which originated in South Asia, demonstrate highly prevalent rates of azole resistance (particularly to fluconazole) and enhanced transmissibility within healthcare settings. Clade II, first identified in East Asia, typically shows lower resistance rates to all three major antifungal classes while maintaining fitness in healthcare environments. Clade III strains, initially detected in Africa, exhibit heterogeneous susceptibility patterns, including reduced sensitivity to amphotericin B and moderate fluconazole resistance. Clade IV, emerging from South America, shows moderate levels of antifungal resistance, particularly to fluconazole and occasionally to echinocandins. Clade V, identified in Iran, represents a recently characterized lineage with high resistance to multiple antifungal classes. The newest addition, Clade VI, was first detected in Singapore and Bangladesh and shares the closest genetic relationships with Clade IV, but remains distinctly separated by at least 37,000 SNPs and demonstrates variable antifungal susceptibility patterns[14]. These clade-specific variations in resistance mechanisms and virulence factors underscore the complexity of tracking and responding to emerging antifungal resistance. Understanding these distinct evolutionary patterns has become increasingly relevant for developing precision-driven surveillance and treatment strategies, particularly as new variants continue to emerge.

Recent advances in high-resolution genomic epidemiology have revealed key molecular mechanisms underlying *C. auris* resistance, including clinically-significant mutations in *ERG11* (Tyr132-Phe) and *FKS1* (Ser639Phe) that confer resistance to azoles and echinocandins, respectively[13]. However, traditional clinical surveillance methods face inherent challenges in rapidly detecting and responding to new outbreaks and emerging resistance patterns. These limitations stem from the baseline timeframe for an infection to present clinically and undergo confirmatory testing, compounded by reliance on timely and active participation by healthcare facilities, which may be reluctant to engage in expanded surveillance. Such delays or outright refusals to participate create critical opportunities for resistant strains to disseminate within healthcare networks[15]. Recent efforts to leverage wastewater-based epidemiology (WBE) have demonstrated the feasibility of detecting *C. auris* DNA in wastewater treatment plants (WWTPs) across the U.S., establishing a promising foundation for community-level surveillance[16–20]. Despite these approaches to surveillance of bacterial, fungal, and viral pathogens[21–23], there is still a critical need to link wastewater findings to facility-specific and/or community-level transmission dynamics to further increase the actionability of wastewater intelligence data.

Here, we demonstrate that integrated wastewater intelligence provides early and actionable warning of emerging *C. auris* subclades during the 2021–2024 outbreak in Las Vegas, Nevada—the largest recorded outbreak in U.S. history to date[24,25]. Through genomic, transcriptomic, and phenotypic analyses, we establish that wastewater monitoring, particularly at healthcare facilities, enables detection of emerging subclades up to nearly five months before their appearance in clinical isolates. Our comparative analysis reveals a high 90% genomic concordance between wastewater and clinical isolates, while uncovering stress-response mechanisms and metabolic adaptations associated with antifungal resistance. These findings demonstrate the power of wastewater surveillance in tracking fungal pathogen evolution and resistance patterns, providing a robust framework for proactive infection monitoring and potentially guiding treatment strategies.

Results

Epidemiological and molecular surveillance reveals distinct patterns of *C. auris* transmission

Longitudinal clinical surveillance of hospitals and long-term care facilities in Southern Nevada from August 2021 to August 2024 revealed a significant outbreak pattern, with a marked increase in confirmed colonizations and clinical cases beginning in the spring of 2022 (Fig. 1A). For this study, we applied the CDC case definitions for *C. auris*, designating colonization cases as detections from screening or surveillance specimens and clinical cases as detections from diagnostic specimens collected in the context of care[26]. Demographic analysis showed distinct patterns in the affected population up to February 22, 2024, with a slight male predominance (59% male versus 39% female, 2% not reported)[27] (Fig. 1B). Age stratification demonstrated the outbreak's disproportionate impact on older adults, with the majority of cases occurring in individuals over 50 years of age. Peak incidence in colonization and clinical cases was observed in those over 70 years (46.7%), followed by the 60–69 age group (26.2%)[27], which is consistent with the previously documented vulnerability of older populations to *C. auris* infection[10]. Combined across colonization and clinical cases, counts by age group were: 50–59 (15.1%), 40–49 (6.8%), 30–39 (3.6%), 20–29 (1.3%), 10–19 (0.1%), and <10 (0.2%)[27].

To investigate whether wastewater surveillance could serve as a sensitive indicator of pathogen transmission within individual facilities, we conducted systematic wastewater monitoring at three major hospitals selected based on secure access to on-site sewer lines, and at their receiving municipal treatment plants (WWTP1 and WWTP4A) between August and December 2023 (Fig. 1C, D). Hospitals 1 and 2 discharge to WWTP1, and Hospital 3 discharges to WWTP4A[28]; municipal influent at both plants also includes surrounding residential, commercial, and industrial wastewater. This parallel sampling approach demonstrated that *C. auris* DNA concentrations in hospital wastewater were significantly correlated with the clinical case and colonization burden at each healthcare facility ($p < 0.001$). While *C. auris* DNA was successfully detected at both WWTPs (Fig. 1C), the hospital wastewater signal was significantly more consistent (detection frequency = 95.4% in hospital wastewater vs. 18.4% at WWTPs), and was significantly higher (Mann–Whitney $U = 15{,}341.5$, $p < 0.001$, Supplementary Figs. 1 and 2). In fact, the maximum concentrations in the hospital wastewater were nearly two orders of magnitude higher than the maximum concentrations at the community-scale WWTPs (~10^8 gc/L vs. ~10^6 gc/L). This dilution effect highlights the enhanced sensitivity and practical utility of facility-level surveillance for early outbreak detection. As of February 22, 2024, analysis of case distribution across the three hospitals revealed distinct spatial transmission patterns, with Hospital 1 experiencing the highest cumulative burden ($n = 769$ colonized cases), followed by Hospital 2 ($n = 203$ colonized cases) and Hospital 3 ($n = 95$ colonized cases)[29] (Fig. 1D).

Targeted amplicon sequencing reveals dynamic *C. auris* populations rather than persistent wastewater biofilms

A fundamental question raised by prior *C. auris* wastewater surveillance studies is whether detection of this biofilm-associated pathogen represents active transmission or potentially just a persistent biofilm within the premise plumbing and/or sewer system[16]. To distinguish between these possibilities, we designed and implemented an amplicon panel targeting 11 genes in *C. auris* implicated in antifungal resistance mechanisms: *CDR1*, *ERG3*, *ERG6*, *ERG11*, *FKS1*, *HSP90*, *MEC3*, *MLH1*, *MRR1A*, *TAC1B*, and *UPC2* (Supplementary Table 1). To check for species discrimination in complex wastewater matrices, we identified and incorporated unique sequence markers for both *C. auris* and *C. albicans*. This targeted approach enabled spatial and temporal monitoring of resistance-associated variants across the three hospitals (Fig. 2).

https://doi.org/10.1038/s41467-026-71960-5

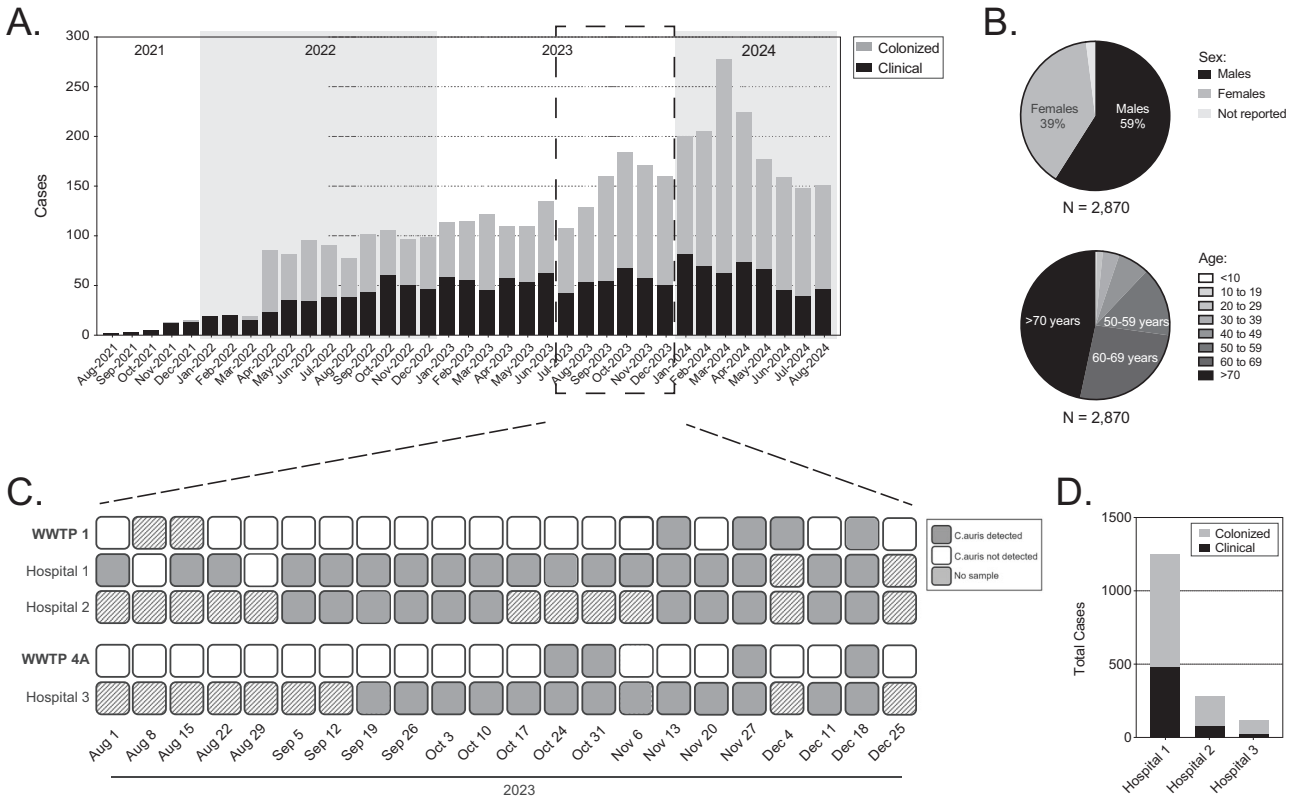

Fig. 1 | Characterization of *C. auris* outbreak dynamics through integrated clinical and wastewater surveillance approaches in Southern Nevada. **A** Monthly distribution of confirmed clinical cases (black bars) and colonization cases (gray bars) showing outbreak progression from August 2021 to August 2024, with a notable peak in late 2023/early 2024. Case data compiled from Southern Nevada Health District surveillance using CDC case definitions[26,27,29]. **B** Demographic analysis showing (top) sex distribution as a pie chart, with Males (black, 59%), Females (gray, 39%), and Not reported (light gray, 2%), and (bottom) age distribution as a pie chart, with age groups shown in grayscale gradient from light gray (<10 years) to black (>70 years) and intermediate shades for 10–19, 20–29, 30–39, 40–49, 50–59, and 60–69 (majority >60 years)[27]. **C** Wastewater surveillance results from three hospitals and two wastewater treatment plants (WWTP1 and 4A) between August and December 2023. Solid gray squares indicate *C. auris* detection, white squares indicate no detection, and hatched squares indicate no sample was collected on that date. **D** Total confirmed colonizations (gray bars) and clinical cases (black bars) at three major hospitals through February 22, 2024. Source data are included in the Source Data 1 file.

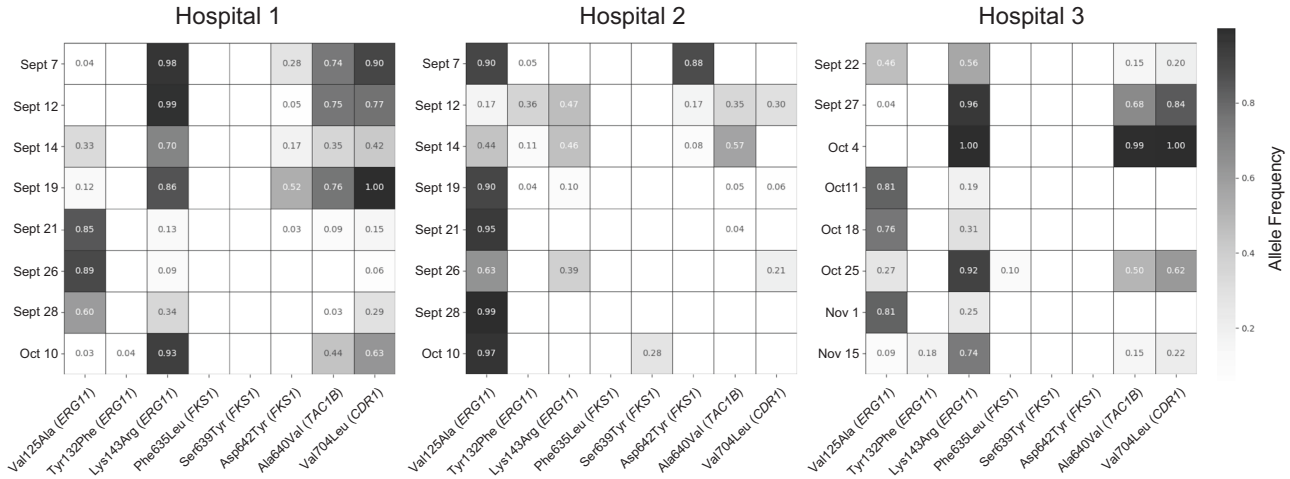

Fig. 2 | Spatial and temporal monitoring of resistance-associated variant frequencies at three hospitals between September 2023 and November 2023. Heatmaps show allele frequencies of mutations in genes associated with antifungal resistance: *ERG11* (azole resistance), *FKS1* (echinocandin resistance), *CDR1*, and *TAC1B* (multidrug efflux). Shading represents variant allele frequencies from non-detect (white) to >75% (black). Hospital 1 (left panel) and Hospital 2 (middle panel) show the emergence of *FKS1* Asp642Tyr independently. Hospital 2 (middle panel) displays rapid shifts in *ERG11* variants (Val125Ala/Phe126Leu and Lys143Arg) between September 7th and September 19th of 2023, and dynamic changes in multiple resistance markers over the sampling period. The temporal variation in mutation frequencies across all three facilities provides evidence for active transmission rather than stable biofilm colonization of wastewater systems. Source data are included in the Source Data 2 file.

https://doi.org/10.1038/s41467-026-71960-5

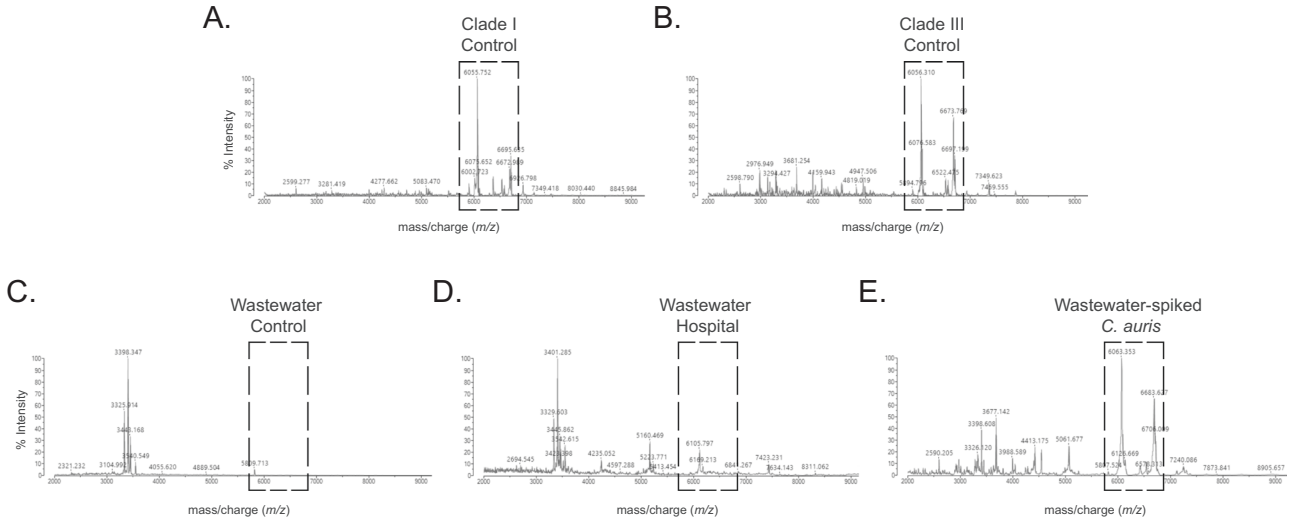

Fig. 3 | MALDI-TOF mass spectrometry for accurate *C. auris* detection in complex wastewater matrices. A Reference mass spectra from purified Clade I *C. auris* control isolate showing characteristic peaks in the 2000–9000 m/z range, with diagnostic signals at 6000–7000 m/z. **B** Reference mass spectra from Clade III *C. auris* control isolate displaying shared core spectral features with Clade I but exhibiting distinct peak patterns in the 6000–7000 m/z region, enabling clade discrimination. **C** Background mass spectra from control (non-spiked) wastewater showing low-intensity peaks around 3000 m/z, establishing the baseline signature of the wastewater matrix. **D** Analysis of non-spiked hospital wastewater samples following a 24–48 h pre-enrichment step, showing preservation of *C. auris*-specific spectral patterns, particularly in the 6000–7000 m/z range, validating detection of *C. auris* in complex wastewater matrices. **E** Mass spectra from wastewater samples spiked with known quantities of *C. auris*, demonstrating clear detection of pathogen-specific peaks against the background matrix.

High-resolution temporal analysis revealed rapid and independent fluctuations in variant frequencies that provided evidence for active pathogen transmission rather than static biofilm persistence. For example, in Hospital 2, we observed dramatic shifts in *ERG11* variants Val125Ala/Phe126Leu and Lys143Arg, with allele frequencies surging from <10% to >75% between September 7 and September 19, 2023 (Fig. 2). These abrupt changes occurred too rapidly to be explained by biofilm evolution alone. Supporting this interpretation, we detected the emergence of the *FKS1* Asp642Tyr mutation (associated with echinocandin resistance[30]) in Hospitals 1 and 2 at different frequencies and timepoints. While patient movement between facilities could not be definitively tracked, the temporal patterns and variant frequencies at each hospital could reflect independent evolution under similar selective pressures or circulation of distinct clones.

Validation of our detection approach included parallel monitoring of species-specific sequences from both *C. auris* and *C. albicans* (Supplementary Table 1). The dual-panel design achieved reliable species discrimination with a detection sensitivity of 3% variant frequency in complex wastewater matrices, enabling identification of minor variants before they achieved dominance. We consistently detected multiple resistance-associated mutations, including the clinically significant *ERG11* Val125Ala/Phe126Leu[31] and *CDR1* Val704Leu[32] variants, demonstrating robust performance in wastewater samples. In addition, the *CDR1*, *ERG11*, *FKS1*, and *TAC1B* variants exhibited distinct facility-specific patterns of emergence and progression. We also observed variants in *ERG3*, *ERG6*, *HSP90*, *MEC3*, *MLH1*, *MRR1A*, and *UPC2*; however, these were synonymous changes or polymorphisms of uncertain significance with no established role in resistance. These spatially resolved evolutionary trajectories showed facility-specific patterns at each hospital, with variant frequencies changing independently over time. The facility-specific nature of these changes, combined with their rapid temporal dynamics, provided strong evidence that wastewater detection captured active *C. auris* transmission rather than persistent biofilm populations. This result is further supported by subsequent genomic analysis showing concordance between these wastewater-detected mutations and clinical isolates from the same facilities (discussed in detail below).

MALDI-TOF mass spectrometry enables *C. auris* identification in complex wastewater matrices

To address the pressing need for accurate pathogen detection in healthcare settings, we developed and validated a MALDI-TOF mass spectrometry workflow for identifying *C. auris* in wastewater samples. Reference spectral libraries were built from locally circulating clades (I and III), which were the only clades documented in clinical genomes at the time hospital sampling began (Fig. 3A, B). These reference spectra revealed distinctive diagnostic peaks across the 2000–9000 m/z range, with robust signals at approximately 6000 m/z for both clades. Although sharing core spectral features, the clades exhibited unique peak patterns in the 6000–7000 m/z region, enabling clade differentiation. Control wastewater samples (confirmed negative for *C. auris* by qPCR) showed consistent background peaks at 3000 m/z, characteristic of the organic matrix (Fig. 3C).

Following a 24–48 h pre-enrichment step, hospital wastewater samples displayed *C. auris*-specific spectral patterns matching our reference isolates, with particularly strong signal preservation in the diagnostically crucial 6000–7000 m/z range, alongside the expected background signals at 3000 m/z (Fig. 3D). Systematic sensitivity analyses with wastewater samples spiked with defined quantities (~10^6 gc/L) of *C. auris* colonies (Fig. 3E) achieved a signal-to-noise ratio of 2:1 for diagnostic peaks, confirming robust detection of *C. auris* in complex wastewater matrices. The stability of these high molecular weight peaks demonstrated minimal matrix interference in critical mass ranges. This optimized MALDI-TOF protocol provides definitive identification of *C. auris* in environmental samples, complementing existing molecular methods. Our findings establish MALDI-TOF mass spectrometry as a powerful tool for real-time pathogen surveillance in healthcare and wastewater/environmental settings, enabling rapid implementation of infection control measures.

High genomic concordance between wastewater and clinical *C. auris* isolates enables early detection of emerging subclades

To evaluate whether wastewater surveillance could provide advance warning of emerging resistant variants, we conducted genomic surveillance across Southern Nevada from August 2021 to August 2024. This systematic molecular approach characterized 2977 patient-

https://doi.org/10.1038/s41467-026-71960-5

Table 1 | General description of clinical and wastewater *C.auris* isolates

Taxon breakdown	Clinical	Wastewater
Candidozyma auris	2945	557
Candida	6	1
Candida/Metschnikowiaceae	4	3
Candida albicans	2	0
Candida duobushaemulonis	1	0
Candida glabrata	4	4
Candida parapsilosis	6	4
Saccharomyces cerevisiae	1	0
NA	8	25
Total assemblies	2977	594
Clade breakdown for C. auris isolates	Clinical	Wastewater
Clade1	932	152
Clade2	0	0
Clade3	2012	401
Clade4	1	0
Top subgroups by clinical isolate count	Clinical (%)	Wastewater (%)
3.A.8.L.2	281 (9.5%)	1 (1.0%)
3.B.1.I	238 (8.1%)	12 (11.4%)
1.A.1.Z	193 (6.6%)	2 (1.9%)
3.A.8.C	177 (6.0%)	7 (6.7%)
3.A.8.A	169 (5.7%)	1 (1.0%)
1.B	168 (5.7%)	13 (12.4%)
3.A.8.Q	108 (3.7%)	0 (0.0%)
3.A.8.I	88 (3.0%)	10 (9.5%)
3.A.6.C	87 (3.0%)	1 (1.0%)
3.B.1.B	83 (2.8%)	5 (4.8%)
Top subgroups by wastewater isolate count	Clinical (%)	Wastewater (%)
1.B	168 (5.7%)	13 (12.4%)
3.B.1.I	238 (8.1%)	12 (11.4%)
3.A.8.N	7 (0.2%)	11 (10.5%)
3.A.8.I	88 (3.0%)	10 (9.5%)
3.A.8.C	177 (6.0%)	7 (6.7%)
3.A.8.X	45 (1.5%)	7 (6.7%)
3.B.1.B	83 (2.8%)	5 (4.8%)
1.A.1.E	26 (0.9%)	5 (4.8%)
3.A.21.B	10 (0.3%)	4 (3.8%)
1.A.1.A	55 (1.9%)	3 (2.9%)

Data represent counts and percentages from three hospitals in Southern Nevada, 2021–2024. Clinical isolates ($n = 2977$ assemblies) include both diagnostic and screening/colonization specimens; wastewater isolates ($n = 594$ assemblies) were recovered from hospital wastewater and municipal treatment plants.

derived isolates (diagnostic and screening/colonization specimens) from 45 healthcare facilities and wastewater isolates from hospital wastewater and WWTPs. Of the 2977 patient yeast colonies processed, 2945 were confirmed as *C. auris*, and all were included in clade/variant analyses (the remaining 32 were other yeasts). From wastewater, 594 yeast colonies were recovered; 557 were identified as *C. auris* by WGS, and 443 met quality control and coverage thresholds and were retained for downstream analyses (see "Methods" section). Initial phylogenetic analysis revealed structured populations dominated by Clade I (31.7%, 932/2945) and Clade III (68.3%, 2012/2945), with a single Clade IV isolate detected, suggesting multiple independent introductions in this geographic region (Table 1).

After removing duplicate same-day wastewater isolates (zero SNP differences) to ensure robust comparative analysis, we identified 105 unique wastewater isolates. These wastewater-derived sequences demonstrated high genomic concordance with the clinical population—90% matched clinical isolates within 3 SNPs and 48% showed exact matches (zero SNP differences) (Supplementary Table 2). Detailed phylogenetic analysis of clinical and wastewater samples within Clade I revealed multiple distinct subclades harboring clinically relevant resistance mutations (Supplementary Fig. 3). The *ERG11* Tyr132Phe mutation, conferring azole resistance, emerged consistently across evolving lineages, with particular prevalence in subclade 1.B (12.4% of wastewater isolates) (Table 1). While this high prevalence could represent repeated sampling of sustained shedding from one or more persistent colonization/case sources within the hospital, the lack of patient-level longitudinal data prevents definitive source attribution. Clinical isolates, in contrast, represent confirmed unique patient cases. Temporal analysis demonstrated the predictive power of wastewater surveillance across several lineages. While subclade 1.A.3 appeared in wastewater 13 days before detection of the first clinical isolates (Fig. 4A, C), the longest lead times were observed in Clade III isolates, which exhibited distinct evolutionary patterns characterized by *FKS1* mutations (Supplementary Fig. 4). Most significantly, subclade 3.A.8.X (6.7% of wastewater isolates) (Fig. 4B, D, and Table 1) was detected in wastewater nearly five months (143 days) before identification of the first clinical isolate—the longest advance warning period observed in our study. Similarly, subclade 3.A.8.Z appeared almost four months (112 days) before identification of the first clinical isolate (Fig. 4B, E).

Comparative analysis revealed distinct resistance mutation patterns between sample types. *FKS1* mutations showed significantly higher prevalence in wastewater isolates (16.2%) compared to clinical isolates (2.2%), with wastewater variants dominated by Asp642Tyr (9 isolates), Ser639Phe (3 isolates), and Ser639Tyr (4 isolates) (Table 2). Mutations in *ERG11* maintained similar frequencies between environments (wastewater: 30.5%; clinical: 30.4%), predominantly Lys143Arg and Tyr132Phe mutations (Table 3).

These findings illustrate how wastewater monitoring can uncover previously undetected subgroups and resistant lineages before their appearance in clinical settings. Wastewater detection preceded clinical identification by 13–143 days, depending on the subclade, though longer lead times may reflect clinical surveillance lag and persistent colonization sources. Combined with high genomic concordance with patient isolates, these results support facility-level wastewater surveillance as a valuable early warning tool for tracking the emergence and spread of *C. auris*.

Emergence of pan-resistant *C. auris* traced through paired wastewater and clinical surveillance

Building upon our observation of high genomic concordance between wastewater and clinical isolates, we next investigated whether culture-based wastewater surveillance could serve as an early indicator of emerging antifungal resistance. We conducted antifungal susceptibility testing (AST) across three major drug classes—azoles, echinocandins, and polyenes—using paired wastewater and clinical isolates (Fig. 5A–C). To establish baseline resistance profiles, AST was conducted on RPMI agar with the inclusion of quality control strains. *Candida parapsilosis* ATCC 22019, *Candida albicans* ATCC 90028, and *Candida krusei* ATCC 6258 yielded MIC values within established reference ranges, confirming assay performance across all drug classes (Fig. 5A and Supplementary Fig. 5). In addition, *C. auris* reference strains provided clade-specific benchmarks: ATCC MYA-5002 (Clade III) and CDC_CAU-27 (Clade III) displayed high-level fluconazole resistance (≥256 mg/L) while remaining susceptible to echinocandins (anidulafungin 0.064–0.094 mg/L, caspofungin 0.25–0.38 mg/L, micafungin 0.019–0.064 mg/L), whereas CDC_CAU-

https://doi.org/10.1038/s41467-026-71960-5

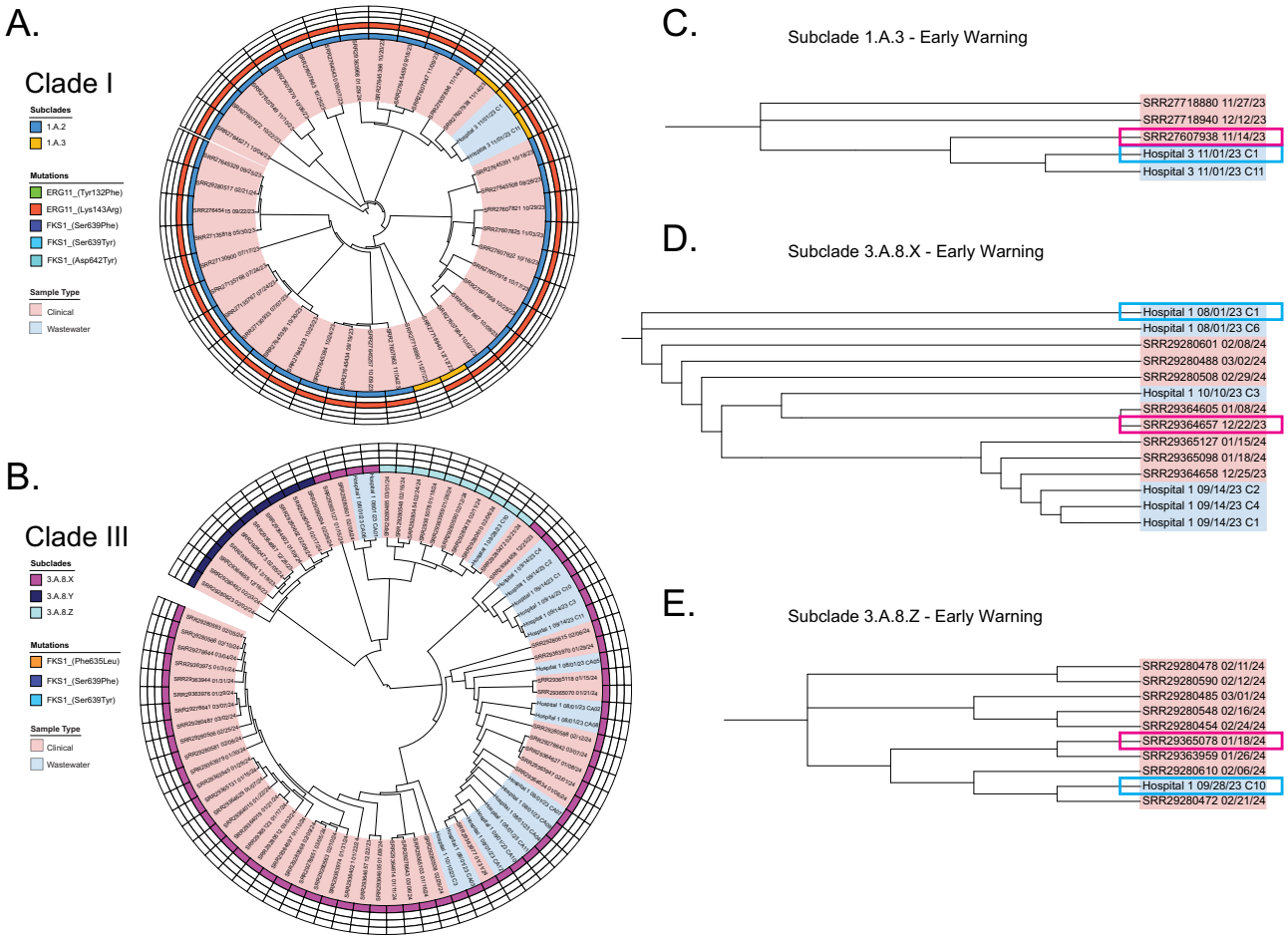

Fig. 4 | Phylogenetic analysis reveals early detection of resistance mutations in wastewater-derived *C. auris* isolates. A, **B** Maximum likelihood phylogenetic tree of Clade I and III isolates from Southern Nevada, showing major subclades and distribution of resistance mutations. Sample types indicated by shading: clinical isolates (pink/salmon shading), wastewater isolates (blue shading). Resistance mutations shown in outer rings by color: *ERG11*_Tyr132Phe (green), *ERG11*_Lys143Arg (orange), *FKS1*_Ser639Phe (dark blue), *FKS1*_Ser639Tyr (cyan), *FKS1*_Asp642Tyr (teal), *FKS1*_Phe635Leu (purple/orange for Clade III). Clade I subclades (**A**): 1.A.2 (blue) and 1.A.3 (orange/yellow). Clade III subclades (**B**): 3.A.8.X (purple), 3.A.8.Y (dark gray), and 3.A.8.Z (light blue/cyan). Detailed phylogenetic subtrees showing early detection of subclades in wastewater: **C** Subclade 1.A.3, **D** Subclade 3.A.8.X, and **E** Subclade 3.A.8.Z. Pink boxes with pink text highlight clinical isolates; blue boxes with blue text highlight wastewater isolates, demonstrating the temporal priority of wastewater detection. Source data are included in the Source Data 3 file.

09 (Clade I) demonstrated broad drug susceptibility, with consistently low echinocandin MIC breakpoints and amphotericin B values in the susceptible range (1–1.5 mg/L) (Fig. 5A, B and Supplementary Fig. 5). Alongside these benchmarks, we identified wastewater-derived isolates with high-level echinocandin resistance, demonstrating the capacity of wastewater testing to capture resistant phenotypes as they emerge in the hospital environment (Fig. 5A). Together, these internal and external controls provided robust anchors for interpreting the susceptibility profiles of wastewater and clinical isolates.

Longitudinal AST profiling of selected wastewater isolates revealed a striking stepwise evolution of drug resistance (Fig. 5B). The first signs of emerging echinocandin resistance appeared in Hospital 2 during early October 2023. While these isolates were predominantly Clade III, the phenotypic resistance patterns showed more complex relationships with genomic changes than simple one-to-one correspondence with known resistance mutations. Specifically, isolates C6 and C7 developed high-level resistance to echinocandins (anidulafungin and caspofungin ≥32 mg/L, micafungin 4–8 mg/L). This resistance pattern subsequently spread, with isolates from both Hospital 2 and Hospital 3 (C4, C8, C2) exhibiting echinocandin resistance (≥32 mg/L) by November-December 2023. Throughout this progression, amphotericin B (AMB) susceptibility remained stable (0.25–0.38 mg/L), suggesting highly specific selective pressure targeting echinocandin resistance mechanisms.

To validate the predictive potential of wastewater surveillance for resistance emergence, we performed paired AST analysis comparing wastewater isolates with their closest clinical matches. These pairs were identified through SNP-based phylogenetic analysis (Fig. 5C)[24,25]. To systematically evaluate phenotypic concordance, we first identified wastewater-clinical isolate pairs based on genomic similarity, focusing on those separated by two SNPs. This analysis revealed robust phenotypic concordance between paired isolates. For example, wastewater isolate Hospital 2_111623_C8 (separated by one SNP from clinical isolate SRR19738710) showed similar higher-level echinocandin resistance (MIC ≥ 32 mg/L). Similarly, Hospital 2_111623_C2 and its clinical counterpart SRR23109144 (two SNPs apart) both exhibited parallel resistance patterns across multiple drug classes. The strong correlation between resistance profiles in wastewater and clinical isolates supports wastewater surveillance as a relevant early warning system for emerging antifungal resistance. However, AST remains essential for definitively characterizing resistance patterns in newly detected isolates to enable truly real-time monitoring.

https://doi.org/10.1038/s41467-026-71960-5

Table 2 | *FKS1* mutations in clinical and wastewater *C. auris* isolates from Southern Nevada, 2021–2024

***FKS1* mutation comparisons**	**Clinical**	**Wastewater**
no *FKS1* mutation	2881	88
FKS1 (Asp642Tyr)	21	9
FKS1 (Leu640Val)	11	0
FKS1 (Ser639Tyr)	5	4
FKS1 (Arg641Ser)	5	0
FKS1 (Ser639Phe)	3	3
FKS1 (Arg1354His)	3	0
FKS1 (Leu638Phe)	3	0
FKS1 (Phe635Leu)	2	1
FKS1 (Ser639Pro)	2	0
FKS1 (Arg641Gly)	2	0
FKS1 (Ile301Phe)	1	0
FKS1 (Phe635Tyr)	1	0
FKS1 (Phe635del)	1	0
FKS1 (Ser1358Phe)	1	0
FKS1 (Ile1361Thr)	1	0
FKS1 (Met690Ile)	1	0
FKS1 (Leu686Phe)	1	0
***FKS1* mutation summary**	Clinical	Wastewater
FKS1 Mutation Total	64	17
Total C. auris	2945	105
% *FKS1* mutation	2.2%	16.2%

Values show counts of isolates harboring specific mutations.

Table 3 | *ERG11* mutations in clinical and wastewater *C. auris* isolates from Southern Nevada, 2021–2024

***ERG11* mutation comparisons**	**Clinical**	**Wastewater**
no *ERG11* mutation	2051	73
ERG11 (Lys143Arg)	746	19
ERG11 (Tyr132Phe)	148	13
***ERG11* mutation summary**	Clinical	Wastewater
ERG11 Mutation Total	894	32
Total C. auris	2945	105
% *ERG11* mutation	30.4%	30.5%

Values show counts of isolates harboring specific mutations.

Transcriptomic analysis reveals stress adaptation mechanisms in *C. auris*

To understand the molecular mechanisms enabling *C. auris* adaptation to environmental and therapeutic stresses, we performed transcriptomic profiling across multiple physiologically relevant conditions (Fig. 6A). We simultaneously exposed wastewater isolates to thermal stress (37 °C and 42 °C), antifungal pressure (fluconazole), and antibiotic exposure, generating high-depth RNA sequencing data (5-10 million reads per sample) to characterize stress-response transcriptomes.

We applied our consensus pathway analysis (CPA)[33,34] pipeline to systematically identify stress-activated pathways in *C. auris* (Fig. 6B). This computational framework integrated de novo transcriptome assembly with mapping to the *C. auris* B8441 reference genome, followed by BLAST-based gene annotation and pathway enrichment analysis. Differential expression analysis using DESeq2 revealed distinct stress-specific transcriptional signatures that we mapped to cellular pathways using stringent statistical criteria.

Analysis of Clade I isolates revealed a hierarchical organization of stress responses (Fig. 6C). The reference strain exhibited a canonical stress response, characterized by coordinated upregulation of ribosomal pathways, ribosome biogenesis, and carbon metabolism. Wastewater isolates 1 and 2 maintained this core response while activating additional survival mechanisms, including aminoacyl-tRNA biosynthesis pathways, suggesting enhanced protein quality control. Interestingly, wastewater isolate 3 displayed a fundamentally altered stress response, characterized by broad suppression of metabolic pathways while maintaining and expanding ribosomal and transcriptional responses. This isolate also uniquely upregulated DNA replication and cell cycle pathways, suggesting a shift from metabolic adaptation toward replication-driven survival strategies.

Clade III isolates demonstrated parallel evolutionary adaptations (Fig. 6D). While the reference strain and wastewater isolates 4 and 5 showed balanced activation of primary metabolic pathways, wastewater isolate 6 revealed complete rewiring of cellular stress responses. This isolated downregulated peroxisome function and fatty acid metabolism while strongly activating ribosomal pathways. It uniquely suppressed efferocytosis pathways and demonstrated fundamental restructuring of carbon metabolism through beta-alanine and propanoate pathways, indicating additional stress adaptation mechanisms.

These transcriptional profiles provide insights into *C. auris* stress adaptation, demonstrating that antifungal resistance emerges through complex reprogramming of cellular pathways rather than simple target modification. The extensive metabolic rewiring observed in wastewater isolates 3 and 6 suggests the evolution of adaptive mechanisms involving dynamic regulatory shifts in stress-response networks, energy metabolism, and cell wall remodeling pathways. These changes likely enhance cellular resilience, contributing to stable resistant phenotypes that persist in clinical settings. Furthermore, the identification of differentially expressed genes linked to oxidative stress resistance, membrane integrity, and metabolic plasticity reveals previously uncharacterized vulnerabilities that could serve as potential therapeutic targets for combating this highly adaptable fungal pathogen.

Discussion

Our integrated analysis of *Candidozyma auris* using wastewater-based epidemiology (WBE) provides a framework for tracking fungal pathogen evolution and resistance emergence. We observed advanced detection of 13–143 days compared with clinical surveillance, with longer lead times likely influenced by clinical screening frequency and sampling dynamics. Greater than 90% genomic concordance between wastewater and patient isolates indicates that wastewater signals largely reflect shedding from colonized or infected individuals rather than new resistance developing within the sewer environment. Because samples were collected upstream at hospital outflows with continuous-flow dynamics, conditions are less likely to support in situ selection, and the higher prevalence of *FKS1* mutations in wastewater is more plausibly explained by sampling bias from a subset of facilities contributing disproportionately to the dataset. This early warning capacity underscores how WBE can be used as a powerful tool for proactive resistance surveillance. These findings provide evidence for incorporating environmental/wastewater monitoring into global strategies addressing the rising threat of emerging pathogens and drug resistance[22,35]. The early detection of resistance mutations, including *FKS1* (Phe635Leu, Ser639Phe, Ser639Tyr, Asp642Tyr)[30,32,36] and *ERG11* (Tyr132Phe, Lys143Arg, and the V125Ala/F126Leu haplotype)[31] variants, provides a critical window for intervention before resistant strains establish dominance in healthcare settings. This advance warning enables implementation of targeted control measures, including enhanced screening, contact precautions, and antifungal stewardship[37]. However, the marked difference in *FKS1* mutation prevalence between wastewater (16.2%) and clinical isolates

https://doi.org/10.1038/s41467-026-71960-5

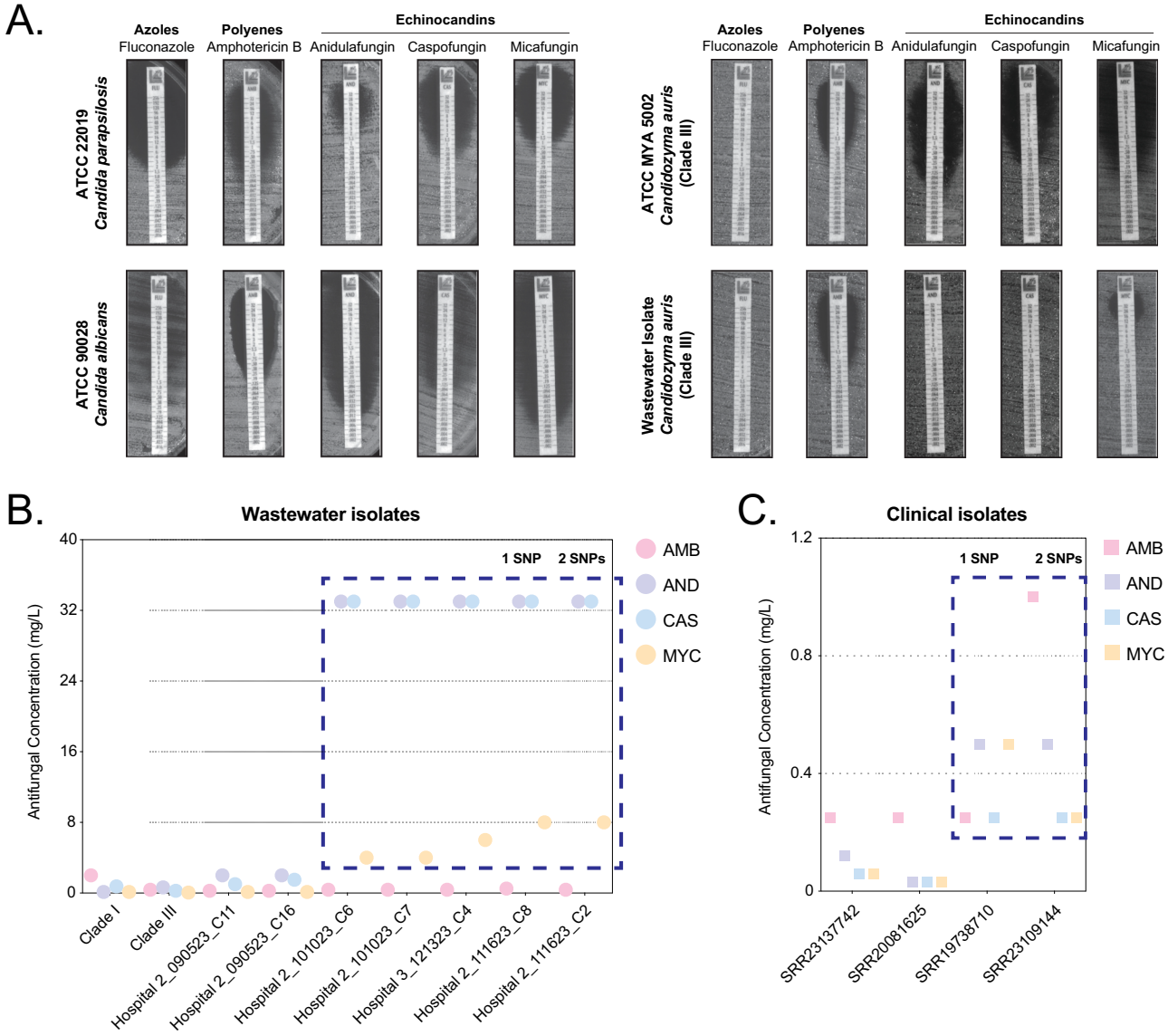

Fig. 5 | Temporal evolution of antifungal resistance profiles in *C. auris* isolates. **A** Representative antifungal susceptibility testing (AST) profiles using gradient diffusion strips showing drug responses against multiple antifungal classes. Reference and control strains shown: ATCC 22019 (*Candida parapsilosis*), ATCC 90028 (*Candida albicans*), and ATCC MYA-5002 (*Candidozyma auris*, Clade III), alongside a Hospital 2 wastewater-derived Clade III isolate showing elevated MICs consistent with echinocandin resistance. The gradient strips display concentration ranges from low (bottom) to high (top), with the elliptical zone of inhibition intersecting the strip at the MIC value. **B** Temporal progression of antifungal resistance in wastewater isolates from Hospital 2 and Hospital 3 collected September-December 2023. Wastewater isolates are designated by collection location, date (MMDDYY), and isolate number (e.g., Hospital 2_090523_C11 indicates Hospital 2, September 5, 2023, isolate C11). Drug concentrations shown on y-axis (mg/L); colors represent different antifungals: Amphotericin B (AMB, pink), Anidulafungin (AND, purple), Caspofungin (CAS, blue), and Micafungin (MYC, yellow). Early isolates (September 2023) show baseline resistance patterns, evolving to echinocandin resistance (≥ 32mg/L) in later isolates while maintaining amphotericin B susceptibility (0.25–0.38 mg/L). Blue dotted boxes highlight clinical isolates showing elevated MICs. **C** Phenotypic concordance between wastewater and clinical isolates. Blue dashed boxes highlight regions containing clinical isolates with elevated MICs (≤2 SNP difference). Two key paired examples demonstrate genetic-phenotypic relationships: Hospital 2_111623_C8 is separated by 1 SNP from clinical isolate SRR19738710, while Hospital 2_111623_C2 is separated by 2 SNPs from clinical isolate SRR23109144. Both pairs show similar resistance profiles, particularly for echinocandins and amphotericin B susceptibility patterns. Legend: Amphotericin B (AMB, pink), Anidulafungin (AND, purple), Caspofungin (CAS, blue), and Micafungin (MYC, yellow). Source data are included in the Source Data 4 file.

(2.2%) highlights the complex relationship between environmental detection and clinical outcomes, emphasizing the need for careful interpretation of wastewater surveillance data.

To better understand these wastewater findings, our longitudinal analysis reveals key insights into the stepwise evolution of echinocandin resistance. The progression from elevated anidulafungin MICs to pan-echinocandin resistance, while maintaining amphotericin B susceptibility, suggests evolutionary constraints on simultaneous resistance development[38]. These patterns likely reflect specific selective pressures within healthcare environments[39], possibly driven by antimicrobial usage patterns or unique wastewater system dynamics. Understanding these evolutionary constraints could inform therapeutic strategies that extend the efficacy of existing antifungal agents. Complementing these findings, our transcriptomic analyses uncover previously unrecognized dimensions of *C. auris* adaptation. The extensive metabolic rewiring observed in resistant isolates, including downregulation of peroxisome and fatty acid degradation pathways[40], challenges traditional resistance paradigms focused solely

https://doi.org/10.1038/s41467-026-71960-5

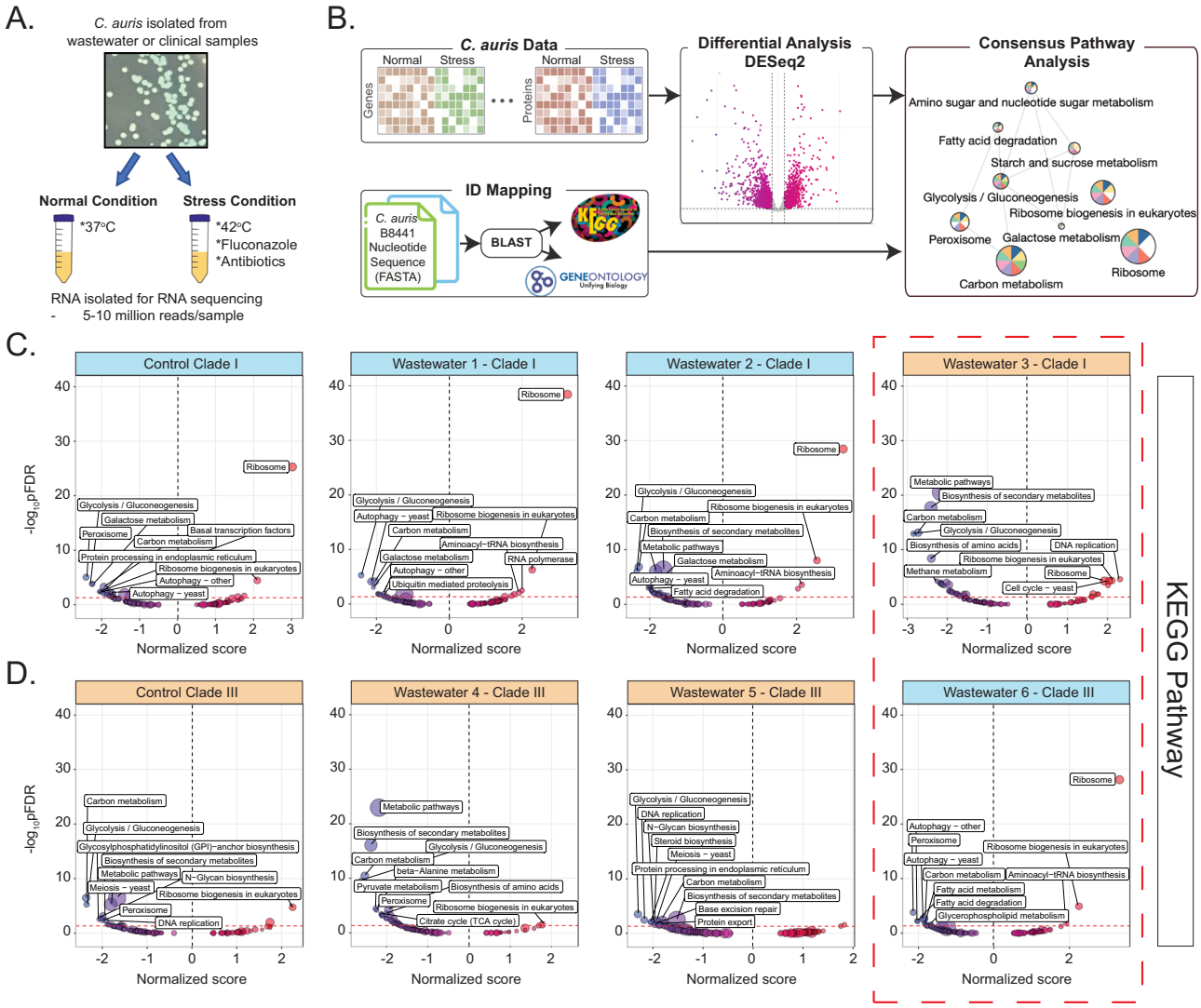

Fig. 6 | Comparative transcriptional profiling reveals distinct stress-response mechanisms in clinical and environmental *C. auris* isolates. A Experimental design for stress-response analysis. Wastewater and clinical *C. auris* isolates were exposed to multiple stressors (37°C, 42 °C, fluconazole, and antibiotics) with paired normal condition controls. RNA sequencing generated >5 million reads per sample. **B** Bioinformatic workflow for consensus pathway analysis. Heatmaps show transcriptomic (left, labeled "Genes") and proteomic (right, labeled "Proteins") expression profiles under normal and stress conditions, with colors representing expression levels: red (upregulated), green (downregulated), and black (no change). Gene identifiers were mapped from *C. auris* B8441 nucleotide sequences via BLAST to KEGG and Gene Ontology databases. Differential expression analysis was performed using DESeq2, followed by consensus pathway analysis (CPA) integrating transcriptomic data to identify significantly enriched KEGG pathways. **C** KEGG pathway enrichment analysis of Clade I isolates. Volcano-style plots show normalized pathway scores with -$\log_{10}$(pFDR) on the y-axis versus normalized enrichment score on the x-axis. Each labeled dot represents a KEGG pathway; pathways above the dashed horizontal line (pFDR <0.05) are significantly enriched. Panel backgrounds indicate transcriptional signature classification: blue denotes Clade I-like pathway profiles, orange denotes Clade III-like profiles (red dashed border highlights isolates whose signature diverges from their phylogenetic clade). Control Clade I and three wastewater isolates are shown. While wastewater isolates 1 and 2 maintain similar pathway activation patterns to the control (dominated by changes in ribosome biogenesis and carbon metabolism), wastewater isolate 3 shows distinct pathway utilization focused on DNA replication and cell cycle processes. **D** KEGG pathway analysis of Clade III isolates. Panel backgrounds indicate transcriptional signature classification as in (**C**): orange denotes Clade III-like pathway profiles, blue denotes Clade I-like profiles (red dashed border highlights isolates whose signature diverges from their phylogenetic clade). Control Clade III and three wastewater Clade III isolates are shown. While the control and wastewater isolates 4 and 5 maintain similar Clade III-like pathway activation patterns, wastewater isolate 6 exhibits a divergent stress response with significant downregulation of peroxisome and fatty acid degradation pathways and activation of ribosome biogenesis.

on drug target modifications. The engagement of additional pathways, particularly in carbon metabolism and cellular stress responses, reveals fundamental adaptation mechanisms that could serve as potential therapeutic targets.

Despite these promising results, the implementation of WBE faces several practical challenges that warrant attention. These include standardizing sampling protocols across diverse healthcare settings, understanding facility-specific influences on resistance patterns, and the inherent limitations of culture-based detection methods, as some cultures remain difficult to grow or isolate from complex wastewater matrices. This could create potential blind spots that may affect clinical decision-making. PCR-based detection methods may show cross-reactivity with other *Candida* species and require careful validation. Our findings raise important questions for future investigation: What drives the rapid emergence/evolution of resistance within healthcare facilities? How do these metabolic adaptations influence pathogen fitness and transmission? Can this surveillance approach be extended to other microbial pathogens, including both fungi and bacteria?

Addressing these questions requires integrating molecular biology, environmental science, and clinical epidemiology. The early detection of resistance variants, coupled with insights into previously undescribed adaptation mechanisms, demonstrates WBE's value as an important tool for pathogen surveillance. Importantly, the identification of previously undetected subclades and resistance mutations exclusively in wastewater highlights its value for uncovering cryptic resistance reservoirs.

Taken together, our study validates upstream (i.e., facility-specific) wastewater surveillance as a powerful approach for monitoring fungal pathogen evolution, providing a robust framework for integrating environmental and clinical data into public-health strategies. By advancing our understanding of *C. auris* resistance and adaptation mechanisms, we establish a foundation for proactive, scalable approaches to managing fungal outbreaks and combating the global challenge of antifungal resistance.

Methods

Data sources and ethical considerations

This study was reviewed by the University of Nevada, Las Vegas Institutional Review Board and determined to be exempt from human subjects research oversight, as wastewater surveillance does not involve identifiable human participants under federal regulations and university policy. Clinical *C. auris* isolates were collected and analyzed by the Nevada State Public Health Laboratory as part of routine public-health surveillance and outbreak investigation. All clinical isolates collected in Nevada are archived in the NCBI BioProject (Accession: PRJNA846332). Isolates were de-identified prior to genomic sequencing and analysis; accordingly, this work was not classified as human subjects research, and informed consent was not required.

Wastewater sample collection, processing, and PCR analysis

Raw wastewater samples were collected weekly from three major hospitals and their corresponding WWTPs in Southern Nevada between August and December 2023. A pilot sample from Hospital 1 was collected on July 5, 2023, to optimize sample collection, extraction, and culturing protocols prior to the start of routine weekly sampling; this sample passed quality control and is included in the dataset. Similar to our previous study[41], composite samples were placed on ice in the field and stored under refrigeration until processing (hold time <36 h). For each sampling event, approximately 100 mL of wastewater was pelleted by centrifugation at 2900 × *g* for 10 min at 4 °C. Total DNA was extracted from the pelleted solids using the DNeasy PowerSoil Pro Kit (Qiagen Catalog #47016). The pellet was resuspended in 800 μL of Solution CD1 and transferred to the PowerBead Pro tube. Nucleic acid extraction was then carried out according to the manufacturer's instructions. Hospital-specific sampling points were established within each facility's sewer network to capture wastewater flows before mixing with other municipal sewage streams[23,35]. At WWTPs, grab (due to practical limitations at Facility 4A) influent samples were collected[42,43]. This sampling strategy enabled comparison between facility-specific (i.e., hospital) signals and the more dilute pathogen concentrations typically observed at downstream treatment plants. Sample processing was conducted in dedicated biosafety facilities following standard protocols for handling potentially infectious materials. All sample collection points were georeferenced, and sample metadata (time, date, location, flow conditions) were recorded to maintain data quality and facilitate spatial-temporal analyses.

C. auris DNA in the samples was quantified by qPCR based on a CDC assay that targets the internal transcribed spacer 2 (ITS2) gene[16]. The sequences for the primers are: (1) Forward: 5′- CAG ACG TGA ATC ATC GAA TCT-3′, and (2) Reverse: 5′- TTT CGT GCA AGC TGT AAT TT-3′. The Probe sequence is: 5′-FAM-AAT CTT CGC GGT GGC GTT GCA TTC A-BHQ1-3′. A synthetic DNA gBlock Gene Fragment qPCR standard for *C. auris* was utilized at different concentrations to create a standard curve, as shown previously for other wastewater targets[22]. Primers, probes, and standards were acquired from Integrated DNA Technologies (IDT, Coralville, IA). Nucleic acid samples and standards were analyzed by qPCR using the Bio-Rad CFX Opus 96 instrument with the Millipore Sigma KiCqStart Probe qPCR ReadyMix (Catalog #KCQS04) as the reaction master mix. The final quantity in gene copies (gc) was divided by the equivalent sample volume (ESV), which was constant at 2 mL; after unit conversion, this yielded a final *C. auris* DNA concentration in gc/L, which represents the solids-associated signal. The limit of quantification (LoQ) for the *C. auris* qPCR assay was defined as 1 gene copy (gc) per reaction. Given the constant equivalent sample volume (ESV) of 2 mL used for analysis, the final reporting LoQ for wastewater was calculated to be 500 gc/L. All concentrations are reported in gc/L, and for logarithmic representation, this LoQ corresponds to 2.70 $\log_{10}$ gc/L (Supplementary Fig. 1). Differences in *C. auris* $\log_{10}$ gc/L concentrations between pooled hospital ($n = 130$) and pooled WWTP ($n = 125$) wastewater samples were compared using a two-sided Mann–Whitney *U*-test; detection frequency differences were assessed by Fisher's exact test (SciPy v1.17.0, Python).

Wastewater amplicon panel, library preparation, and bioinformatics

We designed an amplicon panel to detect antifungal resistance-associated genes in *C. auris*. The panel targeted eleven genes implicated in resistance mechanisms: *CDR1*, *ERG3*, *ERG6*, *ERG11*, *FKS1*, *HSP90*, *MEC3*, *MLH1*, *MRR1A*, *TAC1B*, and *UPC2*. The panels were designed using the *C. auris* strain B8441V2 reference genome (GCA_002759435.2). The first panel, generated using PrimalScheme version 1.4.1[44], comprised 345 primer pairs. The second panel, developed using Olivar version 1.1.1[45], contained 435 primer pairs and incorporated the Simulated Annealing Design using Dimer Likelihood Estimation (SADDLE) algorithm[46] to minimize primer dimer formation and non-specific amplification. To enable species-level discrimination, we identified unique sequence markers for both *C. auris* and *C. albicans* using unikmer version 0.19.2 and incorporated these sequences into the panel.

The primer panel was synthesized as oPools Oligo Pools by IDT. Library preparation consisted of two main steps. First, we performed targeted amplification using CleanPlex mPCR Mix (Paragon Genomics, Catalog #219011) with our custom primer pools. Second, we processed the amplified products using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina (Catalog #E7645L) and NEBNext® Multiplex Oligos for Illumina (Catalog #E7335L) from New England Biolabs. The final multiplexed libraries were sequenced on an Illumina NextSeq 1000 platform using a NextSeq 1000/2000 P1 flow cell with 300 cycles.

Raw sequencing reads were processed using fastp v0.23.4 to remove Illumina adapters, filtering low-quality reads (Phred quality score <20), and trimming of polyG tails[47]. The trimmed reads were aligned to a custom reference FASTA consisting of 11 genes in *C. auris* implicated in antifungal resistance (Supplementary Table 1) using BWA MEM v0.7.17-r1188 with default parameters[48]. Post-alignment sorting and indexing were performed using samtools v1.15.1[49]. Amplicon primers were trimmed using fgbio v2.4.0 TrimPrimers, and fgbio FilterBam was used to trim reads shorter than 40 base pairs. Variant calling was performed using ivar v1.4.3[50]. iVar's TSV output was converted to VCF for annotation with SnpEff version 5.2, using a GFF file generated from the concatenated reference[51]. Samples passing the following QC metrics were retained for downstream analysis: at least 50% of positions covered at ≥100× depth and a minimum read depth of 100 per SNP to balance sensitivity and specificity in variant detection.

Whole-genome sequencing of wastewater and clinical samples and bioinformatics

We performed genomic analysis of *C. auris* isolates from both clinical sources (collected between August 2021 and August 2024) and wastewater using two pipelines: the MycoSNP nextflow pipeline version 1.5[52,53] and TheiaEuk fungal genome specific workflow v2.0.0[24]. MycoSNP was used for reference-based alignment, variant calling, and downstream population-level analyses, while TheiaEuk provided species classification, subclade (clade) typing, and SNP-sharing analyses. For reference-based alignment, we used clade-specific genomes: *C. auris* B11205 (GenBank assembly GCA_016772135.1) for Clade I isolates and *C. auris* B11211 (GenBank assembly GCA_002775015.1) for Clade III isolates to align with CDC's MycoSNP practice and local public health genomics reports, ensuring comparability of tree topology with surveillance dataset[14,25]. To minimize false-positive variant calls, we first identified and masked repetitive elements (approximately 2% of the genome) using MUMmer (version 4.0) nucmer, followed by Bedtools makefasta[54,55]. Raw sequencing reads underwent quality control and preprocessing using FaQCs version 2.10 before alignment to the appropriate reference genome using BWA MEM version 0.7.17-r1188[48,56]. We processed the resulting alignments using Samtools version 1.15 to convert and sort sequence alignment map (SAM) files to binary alignment map (BAM) format[49]. Further refinement of alignments included marking duplicate reads with Picard version 2.26.10 MarkDuplicates and soft-clipping unaligned read regions using Picard CleanSam. For variant calling, we employed GATK HaplotypeCaller version 4.2.5.0, configured for haploid genome analysis. We applied stringent filtering criteria: quality by depth >2.0, Fisher strand bias <60.0, mapping quality >40.0, and read depth >30[57]. Additional quality control measures included filtering for minimum genotype quality (GQ) ≥ 50 and alternative allele frequency ≥ 80% in allele depth. Samples were kept for further analysis if they passed all quality control filters, including the requirement that they exhibit at least 70 percent average coverage across *C. auris* B8441V2 scaffolds 1–10 and 12–14 at a minimum of 30× depth. Variant annotation was performed using SnpEff and SnpSift version 5.2, with results exported to tab-separated format for downstream analysis[51,58]. To enable population-level analyses, we merged individual variant calls using GATK CombineGVCFs. We identified high-confidence SNPs showing ≤10% ambiguity across the sample set and concatenated them into a multi-fasta alignment using custom Python scripts[59]. Finally, we calculated SNP distance matrices using snp-dists and constructed maximum likelihood phylogenetic trees using RAxML-NG v.1.2.2. Tree visualization and annotation were performed using Interactive Tree Of Life[60].

Among wastewater genomes, 443 QC-passed assemblies were obtained; to avoid over representation of clonal replicates, identical same-facility, same-day genomes at 0-SNP distance were removed, yielding 105 non-redundant wastewater genomes for comparative analyses. For the clinical context, we selected nearest-neighbor clinical genomes (≤3 SNPs) to each wastewater genome and added representative clinical genomes to maintain lineage continuity across facilities and time. To keep trees interpretable, we collapsed 0-SNP clinical clusters and downsampled dense lineages stratified by facility and month. Tree inference used the high-quality core-SNP alignment described above; ERG11/FKS1 variants were annotated post hoc and were not used to build the trees.

MALDI-TOF mass spectrometry

Wastewater samples were concentrated by centrifugation, and approximately 10 mg of pelleted material was resuspended in 300 μL HPLC-grade deionized water. Cellular proteins were extracted using a modified ethanol-formic acid protocol. Briefly, samples were mixed with 900 μL pure ethanol, centrifuged (17,000 × *g*, 2 min, 25 °C), and the resulting pellet was washed with an additional ethanol step to ensure complete removal of contaminants. After air-drying, proteins were extracted using 40 μL of 70% formic acid followed by freeze-thaw treatment (−80 °C for 20 min, thawed 10 min at 25 °C). The extract was mixed with an equal volume of acetonitrile, centrifuged (17,000 × *g*, 2 min, 25 °C), and 70 μL of supernatant was collected. For MALDI-TOF MS analysis, 0.5 μL of protein extract was spotted in triplicate onto a MALDI target plate and overlaid with 0.5 μL of matrix solution (α-cyano-4-hydroxycinnamic acid, 100 mg/mL in 50% acetonitrile, 2.5% trifluoroacetic acid). Mass spectra were acquired using a Shimadzu MALDI-8020 mass spectrometer operated in positive linear ion mode over m/z 2000–20,000 Da; sample spots were interrogated using manual positioning, with 10 laser shots accumulated per profile and 100 profiles collected per spectrum at a laser power setting of 80–100, and data were acquired using SHIMADZU MALDI Solutions Data Acquisition software. The instrument was calibrated using DH5α competent cells (Thermo Fisher) as an external standard. All samples were analyzed in technical triplicate, and representative spectra are shown. Spectra were processed in the manufacturer's software using standard baseline subtraction and peak-picking settings. Species and clade assignments were made by multi-peak, library-based matching to published reference spectra generated from characterized Clade I and Clade III isolates[61].

Antifungal susceptibility tests

Minimum inhibitory concentrations (MICs) were determined using the gradient diffusion method with commercially available MIC test strips (Liofilchem). The inoculum was prepared by suspending isolated colonies from overnight cultures in sterile saline and adjusting the turbidity to a 0.5 McFarland standard. Sterile swabs were dipped into the suspension, blotted to remove excess fluid, and used to streak RPMI agar plates (Liofilchem) in three directions to ensure uniform coverage. After an initial inoculation, swabs were re-dipped and reapplied to achieve consistent density, and plates were allowed to dry completely before applying MIC test strips. Plates were incubated at 37 ± 1 °C for 24–48 h under aerobic conditions. MIC values were determined at the intersection of the inhibition ellipse with the test strip gradient markings. Discrete microcolonies within the inhibition zones were excluded from MIC determinations as per standard interpretative criteria. All susceptibility tests were performed in biological triplicate with appropriate quality control strains (*C. parapsilosis ATCC 22019, C. krusei ATCC 6258, and C. albicans ATCC 90028*). The following control *C. auris* isolates were used: ATCC MYA-5002, CDC_CAU-09 (AR-0389), and CDC_CAU27 (AR-1102).

RNA sequencing and data analysis

C. auris isolates were grown in YPD (Sigma Catalog #Y1375, 37 °C, overnight) or SSDB (Thomas Scientific Catalog #CHM01P620, 42 °C, 48 h). Cells were pelleted (2900 × *g*, 5 min), resuspended in YR Digestion Buffer (Zymo Catalog #R1001-1) with 25 U Zymolyase (Zymo E1004), and incubated at 37 °C for 60 min. An additional YR buffer and equal volume of ethanol (95–100%) were added, and RNA was purified according to the manufacturer's instructions for the RNeasy PowerMicrobiome Kit (Qiagen Catalog #26000-50), including DNase I treatment and elution in 54 μL RNase-free water.

Libraries were prepared using the Illumina Stranded mRNA Prep, Ligation Kit (Illumina Catalog #20040534) with IDT RNA UD Indexes Set A (Illumina Catalog #20040553). The library preparation followed the manufacturer's protocol. Briefly, RNA (1 μg) was incubated with RNA Purification Beads (RPBX) to capture mRNA, washed, eluted, and fragmented using Elute, Prime, Fragment High Mix (EPH3). First- and second-strand cDNA synthesis used First Strand Synthesis Mix (FSA +RVT) and Second Strand Marking Master Mix (SMM), followed by 1.8× bead cleanup with NucleoMag NGS Clean-up beads (Macherey Nagel Catalog #744970.500). Adenylation was performed with A-Tailing Mix

(ATL4), followed by adapter ligation with RNA Index Anchors and Ligation Mix (LIGX), and cleanup with a 0.8× bead ratio. Amplification was performed using Enhanced PCR Mix (EPM) and UD Index primers (14 cycles). Final libraries were cleaned with 1× beads and eluted in 17 μL RSB. Library quality was verified using the TapeStation system with High Sensitivity D1000 ScreenTape (Agilent Catalog #5067-5584), and concentrations were measured using the Qubit Flex Fluorometer with Qubit 1× dsDNA HS Assay Kit (Thermo Fisher Catalog #Q33231). Libraries were diluted to 2 nM, pooled, and sequenced on the NextSeq 1000 system using NextSeq 1000/2000 P2 Reagents (300 Cycles) v3 kit (Illumina Catalog #20046813).

We analyzed transcriptional profiles of wastewater-derived *C. auris* isolates using the nf-core/rnaseq pipeline version 3.13.2[53]. All analyses were performed using the *C. auris* B8441 reference genome (version s01-m03-r13) from the Candida Genome Database[62]. Quality control and read preprocessing involved multiple steps to ensure data integrity. First, we performed adapter removal and quality trimming using Trim Galore! version 0.6.7 with Cutadapt version 3.4, applying stringent quality score thresholds. To eliminate potential contamination, we screened reads using BBTools BBSplit version 38.90 for genomic DNA and SortMeRNA version 4.3.6 for ribosomal RNA sequences. For transcriptome analysis, we aligned processed reads to the reference genome using the STAR aligner version 2.7.9a, followed by transcript quantification with Salmon version 1.10.1[63]. The resulting alignments were sorted and indexed using Samtools version 1.17, with duplicate reads marked using Picard version 2.26.10 MarkDuplicates. We then assembled and quantified transcripts using StringTie version 2.2.1 to generate count matrices[64]. To validate transcriptome-derived genetic variants, we performed variant calling on RNA-seq alignments using Freebayes version 1.3.6. We applied stringent filtering criteria (allele frequency ≥0.9, sequencing depth ≥10) and compared these variants with our whole-genome sequencing results to confirm genotype consistency across different sequencing approaches. Differential expression analysis was conducted using the nf-core/differential abundance pipeline version 1.4.0 with DESeq2 version 1.34.0. We identified significantly differentially expressed genes using threshold criteria of adjusted p-value ≤ 0.05 and absolute $\log_2$ fold change ≥1. This analysis revealed distinct transcriptional signatures associated with stress responses and antifungal resistance, as detailed in our "Results" section.

Consensus pathway analysis (CPA) data analysis

To examine the biological pathways affected under different stress conditions, we developed a pathway analysis framework integrating both KEGG (release 111.0)[65] and Gene Ontology (GO, release 2024-09-08)[66] annotations. Given that neither database directly supported Candida Genome Database identifiers, we first established robust gene mapping protocols. For KEGG pathway mapping, we obtained nucleotide sequences directly from the KEGG database. GO gene sequences were sourced from the NCBI *C. auris* assembly ASM301371v2. We then mapped *C. auris* genes to their KEGG and GO counterparts using BLAST (version 2.5.0)[67,68], implementing a stringent similarity threshold of 95% to ensure high-confidence annotations. Pathway enrichment analysis was performed using our previously developed Consensus Pathway Analysis (CPA)[34,69] framework. We employed the *runGeneSetAnalysis()* function from the RCPA package (version 0.2.2) to execute Fast Gene Set Enrichment Analysis (FGSEA)[70] on differential expression results from DESeq2. Statistical significance was assessed using Benjamini–Hochberg False Discovery Rate[71] correction for multiple testing. To visualize pathway enrichment patterns, we generated volcano plots comparing normalized enrichment scores against $-\log_{10}$ adjusted p-values, revealing distinct stress-response signatures across different experimental conditions. The analysis was visualized using standard graphics and diagramming software (OmniGraffle and Adobe Illustrator).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw sequencing data generated in this study have been deposited in the NCBI BioProject database under the following accession numbers: PRJNA1279182 (whole-genome sequencing of wastewater isolates) [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1279182], PRJNA1279255 (wastewater amplicon sequencing) [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1279255], and PRJNA1279245 (RNA sequencing) [https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1279245]. Clinical isolate data collected by the Nevada State Public Health Laboratory are available through the existing NCBI BioProject PRJNA846332. Source data are provided with this paper.

Code availability

Bioinformatic workflows and source code supporting this study are available at https://github.com/M-Moshi/UNLV-NPM-Candida-auris. This repository includes analysis documentation, configuration files, software environment specifications, amplicon primer sequences, and reference files used in this work.

References

1. Satoh, K. et al. Candida auris sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. *Microbiol. Immunol.* **53**, 41–44 (2009).
2. O'Meara, T. R. Understanding pathogen emergence through the lens of Candida auris. Nat. Microbiol. 9, 1399–1400 (2024).
3. McCormick, T. S. & Ghannoum, M. Time to think antifungal resistance increased antifungal resistance exacerbates the burden of fungal infections including resistant dermatomycoses. *Pathog. Immun.* **8**, 158–176 (2023).
4. Chowdhary, A., Jain, K. & Chauhan, N. Candida auris genetics and emergence. *Annu. Rev. Microbiol.* **77**, 583–602 (2023).
5. Eix, E. F. & Nett, J. E. Candida auris: epidemiology and antifungal strategy. Annu. Rev. Med. 76, 57–67 (2025).
6. Chow, N. A. et al. Tracing the evolutionary history and global expansion of Candida auris using population genomic analyses. mBio 11, e03364-19 (2020).
7. Burrack, L. S., Todd, R. T., Soisangwan, N., Wiederhold, N. P. & Selmecki, A. Genomic Diversity across Candida auris Clinical Isolates Shapes Rapid Development of Antifungal Resistance In Vitro and In Vivo. *mBio* **13**, e0084222 (2022).
8. CDC. *Antibiotic Resistance Threats in the United States, 2019* (CDC, 2024).
9. Egger, N. B. et al. The rise of Candida auris: from unique traits to co-infection potential. *Micro. Cell* **9**, 141–144 (2022).
10. Benedict, K., Forsberg, K., Gold, J. A. W., Baggs, J. & Lyman, M. Candida auris–associated hospitalizations, United States, 2017-2022. *Emerg. Infect. Dis.* **29**, 1485–1487 (2023).
11. Soriano, A. et al. Invasive candidiasis: current clinical challenges and unmet needs in adult populations. *J. Antimicrob. Chemother.* **78**, 1569–1585 (2023).
12. Pappas, P. G., Lionakis, M. S., Arendrup, M. C., Ostrosky-Zeichner, L. & Kullberg, B. J. Invasive candidiasis. *Nat. Rev. Dis. Prim.* **4**, 18026 (2018).
13. Cortegiani, A. et al. Epidemiology, clinical characteristics, resistance, and treatment of infections by Candida auris. *J. Intensive Care* **6**, 69 (2018).
14. Suphavilai, C. et al. Detection and characterisation of a sixth Candida auris clade in Singapore: a genomic and phenotypic study. *Lancet Microbe* **5**, 100878 (2024).
15. Salam, M. A. et al. Antimicrobial resistance: a growing serious threat for global public health. Healthcare 11, 1946 (2023).

16. Barber, C. et al. Community-scale wastewater surveillance of Candida auris during an ongoing outbreak in Southern Nevada. *Environ. Sci. Technol.* **57**, 1755–1763 (2023).
17. Rossi, A. et al. Candida auris discovery through community wastewater surveillance during healthcare outbreak, Nevada, USA, 2022. *Emerg. Infect. Dis.* **29**, 422–425 (2023).
18. Babler, K. et al. Detection of the clinically persistent, pathogenic yeast spp. Candida auris from hospital and municipal wastewater in Miami-Dade County, Florida. *Sci. Total Environ.* **898**, 165459 (2023).
19. Zulli, A. et al. Prospective study of Candida auris nucleic acids in wastewater solids in 190 wastewater treatment plants in the United States suggests widespread occurrence. *mBio* **15**, e0090824 (2024).
20. Chavez, J. et al. Early Introductions of Candida auris Detected by Wastewater Surveillance, Utah, USA, 2022-2023. *Emerg. Infect. Dis.* **30**, 2107–2117 (2024).
21. Harrington, A. et al. Urban monitoring of antimicrobial resistance during a COVID-19 surge through wastewater surveillance. *Sci. Total Environ.* **853**, 158577 (2022).
22. Vo, V. et al. Identification and genome sequencing of an influenza H3N2 variant in wastewater from elementary schools during a surge of influenza A cases in Las Vegas, Nevada. *Sci. Total Environ.* **872**, 162058 (2023).
23. Vo, V. et al. Detection of the Omicron BA.1 variant of SARS-CoV-2 in wastewater from a Las Vegas Tourist Area. *JAMA Netw. Open* **6**, e230550 (2023).
24. Ambrosio, F. J. et al. TheiaEuk: a species-agnostic bioinformatics workflow for fungal genomic characterization. *Front. Public Health* **11**, 1198213 (2023).
25. Gorzalski, A. et al. The use of whole-genome sequencing and development of bioinformatics to monitor overlapping outbreaks of Candida auris in southern Nevada. *Front. Public Health* **11**, 1198189 (2023).
26. Centers for Disease Control and Prevention. *Candida auris* 2023 Case Definition. *National Notifiable Disease Surveillance System* (2023).
27. Hynes, M. 'Superbug' fungus cases hit record high in Southern Nevada. *Las Vegas Review-Journal*22 February (2024).
28. Crank, K. et al. Pathogen and indicator trends in southern Nevada wastewater during and after the COVID-19 pandemic. *Environ. Sci. Water Res. Technol.* **11**, 262–280 (2025).
29. Haas, G. Infections from superbug *C. auris* declining in Southern Nevada, state data shows. *8 News Now* 4 September (2024).
30. Hirayama, T. et al. Echinocandin resistance in Candida auris occurs in the murine gastrointestinal tract due to mutations. *Antimicrob. Agents Chemother.* **67**, e0124322 (2023).
31. Williamson, B. et al. Impact of Erg11 amino acid substitutions identified in Candida auris clade III isolates on triazole drug susceptibility. *Antimicrob. Agents Chemother.* **66**, e0162421 (2022).
32. Ben Abid, F. et al. Molecular characterization of Candida auris outbreak isolates in Qatar from patients with COVID-19 reveals the emergence of isolates resistant to three classes of antifungal drugs. *Clin. Microbiol. Infect.* **29**, 1083 e1081–1083.e1087 (2023).
33. Nguyen, H. et al. CCPA: cloud-based, self-learning modules for Consensus Pathway Analysis using GO, KEGG and reactome. Brief. Bioinform. 25 (Suppl. 1), bbae222 (2024).
34. Nguyen, H. et al. CPA: a web-based platform for consensus pathway analysis and interactive visualization. *Nucleic Acids Res.* **49**, W114–W124 (2021).
35. Harrington, A. et al. Environmental surveillance of flood control infrastructure impacted by unsheltered individuals leads to the detection of SARS-CoV-2 and novel mutations in the spike gene. *Environ. Sci. Technol. Lett.* **11**, 410–417 (2024).
36. Kordalewska, M. et al. Understanding echinocandin resistance in the emerging pathogen Candida auris. *Antimicrob. Agents Chemother.* **62**, e00238–18 (2018).
37. Barlam, T. F. et al. Implementing an antibiotic stewardship program: guidelines by the Infectious Diseases Society of America and the Society for Healthcare Epidemiology of America. *Clin. Infect. Dis.* **62**, e51–e77 (2016).
38. Arastehfar, A. et al. Echinocandin persistence directly impacts the evolution of resistance and survival of the pathogenic fungus Candida glabrata. *mBio* **15**, e0007224 (2024).
39. Larsson, D. G. J. & Flach, C. F. Antibiotic resistance in the environment. *Nat. Rev. Microbiol.* **20**, 257–269 (2022).
40. Huang, F. et al. Peroxisome disruption alters lipid metabolism and potentiates antitumor response with MAPK-targeted therapy in melanoma. J. Clin. Invest. 133, e166644 (2023).
41. Zhuang, X. et al. Early detection of emerging SARS-CoV-2 Variants from wastewater through genome sequencing and machine learning. *Nat. Commun.* **16**, 6272 (2025).
42. Zhuang, X. et al. Drug use patterns in wastewater and socioeconomic and demographic indicators. *JAMA Netw. Open* **7**, e2432682 (2024).
43. Vo, V. et al. Use of wastewater surveillance for early detection of Alpha and Epsilon SARS-CoV-2 variants of concern and estimation of overall COVID-19 infection burden. *Sci. Total Environ.* **835**, 155410 (2022).
44. Quick, J. et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **12**, 1261–1276 (2017).
45. Wang, M. X. et al. Olivar: towards automated variant aware primer design for multiplex tiled amplicon sequencing of pathogens. *Nat. Commun.* **15**, 6306 (2024).
46. Xie, N. G. et al. Designing highly multiplex PCR primer sets with Simulated Annealing Design using Dimer Likelihood Estimation (SADDLE). *Nat. Commun.* **13**, 1881 (2022).
47. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
48. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
49. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
50. Grubaugh, N. D. et al. An amplicon-based sequencing framework for accurately measuring intrahost virus diversity using PrimalSeq and iVar. *Genome Biol.* **20**, 8 (2019).
51. Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).
52. Bagal, U. R. et al. MycoSNP: a portable workflow for performing whole-genome sequencing analysis of Candida auris. *Methods Mol. Biol.* **2517**, 215–228 (2022).
53. Ewels, P. A. et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* **38**, 276–278 (2020).
54. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
55. Marcais, G. et al. MUMmer4: a fast and versatile genome alignment system. *PLoS Comput. Biol.* **14**, e1005944 (2018).
56. Lo, C. C. & Chain, P. S. Rapid evaluation and quality control of next generation sequencing data with FaQCs. *BMC Bioinform.* **15**, 366 (2014).
57. McKenna, A. et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. Cingolani, P. et al. Using Drosophila melanogaster as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front. Genet.* **3**, 35 (2012).

59. Broad Institute. broadinstitute/broad-fungalgroup: Broad-fungal Genomics group scripts. GitHub (2024).
60. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455 (2019).
61. Abdolrasouli, A. & Fraser, M. A. Candida auris identification and profiling by MALDI-ToF mass spectrometry. *Methods Mol. Biol.* **2517**, 21–32 (2022).
62. Skrzypek, M. S. et al. The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. *Nucleic Acids Re.s* **45**, D592–D596 (2017).
63. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
64. Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).
65. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
66. The Gene Ontology Consortium The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
67. Blast, G. PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**, 3389–3402 (1997).
68. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinform.* **10**, 1–9 (2009).
69. Nguyen, H. et al. RCPA: an open-source R package for data processing, differential analysis, consensus pathway analysis, and visualization. *Curr. Protoc.* **4**, e1036 (2024).
70. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at *bioRxiv* https://doi.org/10.1101/060012 (2021).
71. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289–300 (1995).

Acknowledgements

E.C.O., V.V., C.L.C., and M.M. are supported by NIH grants: MH109706 and AI194218. D.G., C.L., V.V., and E.C.O. were supported by Grant Number NH75OT000057-01-00 from the Centers for Disease Control and Prevention (CDC). T.N., Q.H.N., and H.N. were supported by NIH: R44GM152152 and U01CA274573. The project contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC. This work would not have been possible without the participation of the collaborating wastewater agencies in Southern Nevada. The first and corresponding authors have full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Author contributions

C.L.C., M.A.M. and Q.H.N. contributed equally to this work. Conceptualization: C.L.C. and E.C.O. Investigation: C.L.C., M.A.M., J.O., P.P., M.A., A.J.T., J.Y.I. and V.V. Formal analysis: C.L.C., M.A.M., Q.H.N., H.N., L.M., T.N., D.H. and E.C.O. Writing – original draft: C.L.C. and E.C.O. Writing – review and editing: all authors. Funding acquisition: E.C.O., V.V., D.G., C.L. and T.N. Supervision: E.K., C.L., H.Y.K., M.P., D.G., V.V., T.N., D.H. and E.C.O.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-026-71960-5.

Correspondence and requests for materials should be addressed to Daniel Gerrity, Van Vo, Tin Nguyen, David Hess or Edwin C. Oh.

Peer review information *Nature Communications* thanks Calabria de Araújo, Milena Kordalewska, and Marius Surleac for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2026

Ching-Lan Chang [1,2,13], Michael A. Moshi [1,2,13], Quang-Huy Nguyen [3,13], Junghun Oh[1], Ha Nguyen [4], Pratik Paranjape[2], Mohammed Abushanab[1], Austin J. Tang [1], Jose Yani Itorralba[1], Lauryn Massic[5], Eakalak Khan [6], Cassius Lockett[7], Horng-Yuan Kan[7], Mark Pandori[5], Daniel Gerrity [8] ✉, Van Vo [1] ✉, Tin Nguyen [9,10] ✉, David Hess [5] ✉ & Edwin C. Oh [1,2,11,12] ✉

[1]Laboratory of Neurogenetics and Precision Medicine, Center for Water Intelligence and Community Health, College of Sciences, University of Nevada Las Vegas, Las Vegas, NV, USA. [2]Neuroscience Interdisciplinary Ph.D. Program, University of Nevada Las Vegas, Las Vegas, NV, USA. [3]Department of Computer Science and Software Engineering, Auburn University, Auburn, AL, USA. [4]Department of Computer Science, Wayne State University, Detroit, MI, USA. [5]Nevada State Public Health Laboratory, Reno, NV, USA. [6]Civil and Environmental Engineering and Construction Department, University of Nevada Las Vegas, Las Vegas, NV, USA. [7]Southern Nevada Health District, Las Vegas, NV, USA. [8]Southern Nevada Water Authority, Las Vegas, NV, USA. [9]Department of Industrial and Systems Engineering, Wayne State University, Detroit, MI, USA. [10]Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit,

MI 48201, USA. [11]Department of Brain Health, Kirk Kerkorian School of Medicine at UNLV, Las Vegas, NV, USA. [12]Department of Internal Medicine, Kirk Kerkorian School of Medicine at UNLV, Las Vegas, NV, USA. [13]These authors contributed equally: Ching-Lan Chang, Michael A. Moshi, Quang-Huy Nguyen.
✉e-mail: daniel.gerrity@snwa.com; van.vo@unlv.edu; tin@wayne.edu; dhess@med.unr.edu; edwin.oh@unlv.edu