Genomic reconstruction of an azole-resistant *Candida parapsilosis* outbreak and the creation of a multilocus sequence typing scheme

Running title: Genomic epidemiology and sequence typing of *Candida parapsilosis*

Phillip Brassington[1], Frank-Rainer Klefisch[2], Barbara Graf[3], Roland Pfüller[4], Oliver Kurzai[5,6], Grit Walther[5], Amelia E. Barber[1]

Address correspondence to: Amelia E. Barber (amelia.barber@uni-jena.de) or Grit Walther (grit.walther@leibniz-hki.de)

[1] Institute for Microbiology. Friedrich Schiller University Jena. Jena, Germany.
[2] Paulinen Hospital. Berlin, Germany.
[3] Labor Berlin. Berlin, Germany.
[4] MVZ Hauptstadtlabor, Labservice GmbH. Berlin, Germany.
[5] National Reference Center for Invasive Fungal Infections (NRZMyk). Leibniz Institute for Natural Product Research and Infection Biology-Hans Knöll Institute. Jena, Germany.
[6] Institute for Hygiene and Microbiology. University of Würzburg. Würzburg, Germany.

Abstract

Background

Fluconazole-resistant *Candida parapsilosis* has emerged as a significant healthcare-associated pathogen with a propensity to spread patient-to-patient and cause nosocomial outbreaks, similar to *Candida auris*. This study investigates a prolonged outbreak of fluconazole-resistant *C. parapsilosis* across multiple years and healthcare centers in Germany.

Methods

We employed whole-genome sequencing of isolates from the outbreak, other regions within Germany, and compared them with isolates from a global distribution to understand the molecular epidemiology of this outbreak. Additionally, we used the genomic dataset of 258 samples to identify loci with high discriminatory power to establish the first multi-locus sequence typing (MLST) strategy for *C. parapsilosis*.

Findings

A clonal, azole-resistant strain of *C. parapsilosis* was observed causing invasive infections over multiple years and in multiple hospitals within the outbreak city. Including this outbreak clone, we identified three distinct *ERG11* Y132F azole-resistant lineages in Germany, marking the first description of this azole-resistance in the country and its endemic status. Using the novel MLST strategy, isolates were categorized into 31 sequence types, proving the utility of the typing scheme for genetic epidemiology and outbreak investigations as a rapid alternative to whole genome sequencing.

Interpretation

Temporal and genomic reconstruction of the outbreak indicated that transfer of patients between healthcare facilities was likely responsible for the persistent reimportation of the drug-resistant clone and subsequent person-to-person transmission. This research underscores the importance of monitoring of *C. parapsilosis* epidemiology, not only in Germany but globally. The emergence of azole-resistant lineages necessitates continuous surveillance and rigorous infection control measures. By combining large-scale genomic epidemiology and introducing a novel typing method, our study offers valuable insights into the management of emerging healthcare-associated pathogens, with direct implications for public health and clinical practice.

Funding

German Federal Ministry for Education and Research, German Research Foundation, German Ministry of Health

Research in Context

Evidence before this study

The epidemiology of candidemia has undergone dramatic changes in recent years. New pathogenic species, such as *Candida auris,* have emerged, and existing species like *Candida parapsilosis* have increased in prominence. There has also been a worrying increase in drug resistance among *Candida* species. Moreover, numerous drug-resistant outbreaks of *C. parapsilosis* have been reported worldwide and are challenging to control due to their prolonged and intermittent nature. The evidence sources considered for this study were journal articles found through the search terms “Candida parapsilosis”, “outbreak”, “azole resistance”, and/or “fluconazole” in PubMed and Google Scholar.

Added value of this study

This study adds to the existing evidence by utilizing whole genome sequencing in conjunction with hospital records to analyze a prolonged outbreak of clonal, azole-resistant *C. parapsilosis* that occurred across multiple years and medical centers. This study demonstrates that patient transfers can result in the reimportation of outbreak clones, posting a significant challenge for infection control. We also reveal that the outbreak clone is closely related to drug-resistant isolates from other continents, highlighting the global spread of drug-resistant *C. parapsilosis*. Furthermore, the study addresses the need for rapid strain differentiation in outbreak settings by establishing and validating a set of four loci for sequence-based typing, which provide a highly discriminatory tool for epidemiologic investigations.

Implications of all the available evidence

This study underscores the global challenge of azole-resistant *C. parapsilosis* and its importance as the causative agent of nosocomial outbreaks. Clinicians should be aware of the evolving epidemiology of *C. parapsilosis* and the prevalence of drug-resistant strains, emphasizing the importance of appropriate antifungal stewardship and infection control measures. The study emphasizes the challenges caused by inter-hospital transmission and their role in persistent outbreaks, highlighting the need for robust surveillance and coordination among healthcare facilities. Furthermore, the establishment of a sequence-based typing scheme is a valuable tool for rapid assessment of outbreak clones, which can aid in outbreak tracking and containment efforts.

Introduction

During the last decade, the epidemiology of candidemia has seen dramatic changes due to the emergence of new pathogenic species like *Candida auris,* and dramatic increases in the prominence of existing species such as *Candida parapsilosis* [1–4]. *C. parapsilosis* is now the second to fourth leading cause of systemic *Candida* infections and has gone from being historically associated with pediatric infection to a frequent cause of catheter-associated bloodstream infections in adults [2]. A worrying accompaniment to the changing species epidemiology is a corresponding increase in drug resistance in *Candida* spp, including *C. parapsilosis*. Previously, isolates of *C. parapsilosis* were overwhelmingly azole susceptible. Now, there are more than 36 studies from 15 countries documenting local fluconazole resistance rates of >10% in *C. parapsilosis* and an exponential increase in reports of fluconazole-resistant *C. parapsilosis* since 2019 [4]. Fluconazole-resistant invasive disease results in a three-fold increase in mortality compared to infection with a susceptible strain (16% for susceptible to 50% resistant infections) [5]. The most dominant fluconazole resistance mechanism seen clinically is the Y132F amino acid substitution in the azole target gene, lanosterol14-alpha-demethylase (*ERG11*). This mutation has been reported in every country where fluconazole-resistant *C. parapsilosis* has been reported. Other resistance mutations, including additional substitutions in *ERG11* (e.g. K128N, K134R, G458S), as well as gain-of-function mutations in the transcription factors *MRR1* and *TAC1* have also been observed clinically, but are less common [6–8].

Alongside the rapid increase in fluconazole resistance, there has been an alarming number of drug-resistant outbreaks of *C. parapsilosis* reported in the last ~5 years [5,9–18]. This organism is commonly found on the skin, and, as a result, it can spread within healthcare facilities in a patient-to-patient manner through contaminated medical devices, environmental surfaces, or the hands of healthcare workers [10,14,19]. Outbreaks of *C. parapsilosis* can also persist despite strict infection control strategies and may show gaps in the occurrence of outbreak strains [4,20]. Recognition of nosocomial outbreaks and implementation of successful infection control measures requires a rapid and reproducible method for discriminating closely related isolates. However, there is currently no gold standard for strain typing in *C. parapsilosis*. Currently, the most widely used method is a species-specific set of microsatellite markers [21]. However, microsatellite typing requires specialized equipment, is prone to PCR artifacts, and requires specialized knowledge for interpretation [22,23]. Moreover, microsatellite-based typing results are not as easily shared and compared across centers. Whole genome sequencing (WGS) has exceptional resolution for discriminating between closely related isolates and is incredibly valuable for outbreak investigation, as has been demonstrated for *C. auris* [24–26]. However, WGS is not feasible for many settings due to its cost, time to result, and limitations in trained personnel to analyze the data. As demonstration of this, to our knowledge there is no large-scale, genome-level analysis of a *C. parapsilosis* outbreak, despite their increasing frequency.

In this study we report an azole-resistant outbreak caused by Y132F *C. parapsilosis* that occurred in Berlin, Germany across several years and medical centers. We use a large-scale genomic dataset

130 composed of 258 *C. parapsilosis* isolates to investigate outbreak dynamics, confirm the high degree
131 of relatedness among outbreak isolates, and compare them to isolates from around the world. Finally,
132 we leverage this genome collection to identify and validate four markers for Sanger sequence-based
133 strain typing, epidemiology, and rapid differentiation of isolates in future outbreaks.
134
135 **Results**
136 Clinical case description and outbreak reconstruction
137 We report a prolonged outbreak of *Candida parapsilosis* that occurred across multiple healthcare
138 institutions across Berlin, Germany during the period of 2018 - 2022 (Figure 1a). Cases were initially
139 detected by physicians at Hospital #1 (name anonymized for privacy), who recognized an increase in
140 isolates with a distinct antifungal susceptibility profile in their facility. The isolates were resistant to
141 fluconazole (FLC^R) and voriconazole (VOR^R), but susceptible to other triazoles. After a period of
142 approximately five months, during which no FLC^R and VOR^R infections were reported, infection
143 isolates with this susceptibility profile were once again observed at Hospital #1. The outbreak was
144 officially declared at that time (October 2019), and was formally investigated including environmental
145 surveillance. As *C. parapsilosis* is known to colonize the ear, 21 ICU patients at Hospital #1 were
146 screened with swabs of both ears, but none were positive for *C. parapsilosis*. To investigate possible
147 transfer by staff, sixteen bedside stethoscopes were screened. Only one sample was positive, but the
148 resulting isolate of *C. parapsilosis* was azole susceptible. In addition, opened and sealed hand creams
149 used by staff were screened, but the pathogen was not detected. Hand sampling of staff was not
150 performed as this is discouraged in Germany due to stigmatization.
151
152 To prevent further cases, emphasis was placed on hygiene measures and particularly hand
153 disinfection. Following implementation of the infection control measures, there was a nine month
154 period without FLC^R and VOR^R strains reported. However, additional cases were detected in
155 September 2021 and the start-stop pattern continued through 2022 (Fig 1a). Consultations with other
156 hospitals and clinical laboratories in Berlin as well as the National Reference Center for Invasive
157 Fungal Infections (NRZMyk) identified *C. parapsilosis* infections with this same antifungal
158 susceptibility profile at five other institutions around Berlin. Subsequent epidemiological and genomic
159 investigations revealed a total of 33 outbreak cases that are analyzed in the rest of this manuscript.
160 The majority of outbreak cases were first reported as wound infections (39%), followed by infections
161 associated with central venous catheters (29%) (Table 1). Outbreak isolates were all resistant to
162 fluconazole with a median minimum in inhibitory concentration (MIC) of 64 mg/l and voriconazole
163 (median MIC 1 mg/l), but susceptible to itraconazole (median MIC 0.125 mg/l) and posaconazole
164 (median MIC 0.03 mg/l) (Table 1).
165
166 To understand the transmission dynamics of the outbreak, we reconstructed the temporal and spatial
167 relationship of the outbreak using admission and discharge records. The reconstruction of inpatient
168 admissions and transfers between medical centers in Berlin revealed that even though the largest

fraction of outbreak cases (n = 20) were diagnosed at Hospital #1, the majority of these patients had been directly transferred to this facility from a major surgical facility (indicated as Hospital #0 in Figure 1b). Supporting nosocomial transmission, the majority of outbreak cases (31/33; 94%) had verifiable, temporal overlap with other cases, i.e. inpatient stays at a given hospital whose dates overlapped with >one other outbreak patient (Figure 1b). Despite lacking specific ward information for all outbreak cases, seven patient clusters involving a total of 22 patients had demonstrated spatial and temporal overlap, with outbreak patients confirmed to have been on the same ward at the same time as other outbreak cases, strongly supporting patient-to-patient spread (Figure 1b). In summary, we report a multiyear, sporadic outbreak of azole-resistant *C. parapsilosis* in Germany that was not limited to any single hospital but involved a network of hospitals in Berlin.

Genomic epidemiology of *C. parapsilosis* confirms the clonal nature of Berlin outbreak strains and reveals the presence of the outbreak strain on additional continents.

To define the genomic epidemiology of *C. parapsilosis* in Germany and examine the genomic relationship between the putative outbreak cases from the hospitals in Berlin, we performed whole genome sequencing (WGS) on 68 isolates (33 outbreak isolates; 33 non-outbreak isolates) collected by the German National Reference Centre for Invasive Fungal Infections (NRZMyk) between 2016 and 2022. Sequenced isolates originated from 29 medical centers in 16 cities. 64% (n = 44) of the sequenced samples originated from Berlin, including samples suspected to be part of a local outbreak and 14 non-linked isolates from different medical centers. Publicly available sequence data for an additional 190 samples was also included to facilitate comparison of the German sequenced isolates with the global *C. parapsilosis* population, bringing the total number of samples in the dataset to 258 genomes. WGS samples analyzed had a mean depth of coverage of 146x (range 10-248x) after data quality control and mapping (Supplemental File 1).

When built into a whole genome phylogeny, samples from Germany were distributed throughout the tree, with no phylogenetic group being unique to a specific country or continent (Figure 2). The close genetic relationship between the outbreak strains was evident by their tight grouping and short branch lengths in the phylogeny, supporting the nosocomial transmission that was inferred from the spatial and temporal overlap in cases (Figure 1b). The quasi-clonal relationship between outbreak isolates was further confirmed by pairwise single nucleotide variant (SNV) analysis, where isolates in the Berlin outbreak clade differed by an average of 36 SNVs (Figure 3). In contrast, the average number of SNVs between non-outbreak isolates from other locations in Germany was significantly higher at 2,079 SNVs ($p < 0.001$ by student's t-test). Further, the number of SNVs observed between duplicate sequencing of the same freezer stock of an isolate strain was 15 SNVs, indicating that the number of pairwise SNV differences between repeat sequencing of the same sample was similar to the genetic differentiation observed between different outbreak isolates. Strikingly, the public sequence data revealed samples from Turkey, South Korea, and Kuwait that were closely related to the outbreak strains from Berlin (Figure 2). These isolates were also reported to be fluconazole resistant [27–29],

208 highlighting the global spread of azole-resistant *C. parapsilosis*. These tightly related global strains
209 differed from the Berlin outbreak strains by an average of 96 SNVs (range 18-191), compared to the
210 mean 36 SNVs that the Berlin cluster differed from each other on average. Altogether, whole genome
211 sequencing confirmed the close genetic relationship of the outbreak strains and its presence of this
212 lineage in multiple other countries.
213
214 *ERG11* Y132F is the primary azole resistance mechanism among *C. parapsilosis* strains from
215 Germany
216 Of the 68 German isolates newly described and sequenced in this study, 65% were resistant to one
217 or more triazole (n = 44) (Supplemental File 1). *ERG11* Y132F was the dominant resistant
218 mechanism, with 36 resistant isolates possessing this mutation (n = 31 outbreak; n = 5 non-outbreak).
219 Interestingly, the non-outbreak Y132F strains were in two distantly related groups of samples based
220 on the phylogeny (Figure 2; Supplemental Figure 1). Each of these Y132F-containing phylogenetic
221 groups contained isolates from geographically distant regions of Germany, suggesting that there are
222 at least 3 genomically-distinct lineages of Y132F *C. parapsilosis* circulating widely throughout the
223 country.
224
225 We also observed three closely related but unlinked samples from different cities in Germany that
226 were pan-azole resistant and carried a G458S substitution in *ERG11* (Figure 2; Supplemental Figure
227 1) that has been linked to clinical pan-azole resistance previously [12,18,30,31]. Also included in our dataset
228 were eight isolates that were resistant to one more azoles, but did not contain any resistance-
229 associated polymorphisms in *ERG11*. However, these isolates contained sequence polymorphisms
230 in several genes that have been linked to drug resistance, including the efflux pumps *MDR1* and
231 *CDR1B*, the multidrug resistance gene regulator *MRR1,* and the transcriptional activator *TAC1*
232 (Supplemental Table 1). While a casual role for most of these polymorphisms in drug resistance have
233 not been experimentally established, one resistant isolate harbored a L518F substitution in *TAC1*
234 experimentally proven to confer azole resistance [8]. Among our newly described isolates and the public
235 genomes we observed additional amino acid polymorphisms in *ERG11* present in susceptible strains,
236 including M178T, N283Y, and both resistant and susceptible strains R398I [32–34]. Together, we find
237 that *ERG11* Y132F is the dominant resistant mechanism among both outbreak and non-outbreak
238 fluconazole-resistant strains in Germany and, surprisingly, discover that our outbreak strains are
239 closely related to resistant samples previously described from East and West Asia.
240
241 Identification of markers for multilocus typing in *C. parapsilosis*
242 As outbreaks of *C. parapsilosis* are being increasingly reported, there is an urgent need for a rapid
243 and reproducible tool capable of differentiating isolates. Motivated by this, we set out to identify a
244 minimal set of loci that could be used for Sanger sequence-based typing of both general epidemiology
245 and outbreak detection. Based on the whole genome phylogeny, we identified 14 distinct genetic
246 clusters of *C. parapsilosis* in our dataset of 258 genomes (Figure 4a, b). The clusters contained an

247 average of 17 samples (range 2-105) and were numbered to be concordant with the five major clades
248 of *C. parapsilosis* described in [35].
249
250 To identify loci with high discriminatory power, we performed a genome-wide screen of genetic variant
251 profiles first using full-length gene coding sequences, followed by 750 bp regions of CDS with the
252 high discriminatory power, and finally assessing these loci in combination. Using this approach, we
253 identified four 750 bp loci in CPAR2_101470, CPAR2_108720, CPAR2_212940, and CPAR2_808110
254 (Table 2) that were capable of differentiating isolates into 31 unique sequence types with a Hunter's
255 Discriminatory Power of 0·87. As *C. parapsilosis* is a diploid organism, an appreciable fraction of the
256 discriminatory power comes from accounting for whether observed sequence polymorphisms were
257 heterozygous or homozygous. When nucleotide zygosity was not considered, we were only able to
258 identify 22 sequence types and the discriminatory power was reduced to 0·81. The four selected loci
259 showed a high degree of agreement with the whole genome phylogeny. Our typing scheme was able
260 to correctly identify samples to their whole genome-defined genetic cluster with 97·6% sensitivity and
261 99·9% specificity and a phylogeny created from loci sequence largely replicated the structure of the
262 whole genome tree (Figure 4c). To confirm that the selected loci were present as a single copy and
263 the discriminatory power was not being inflated by the presence of multiple copies in the genome, we
264 assessed them for genome-wide copy number changes. Using three different copy number variation
265 softwares, we did not detect any CNV changes in these loci, either additional copies or loss of the
266 region (Supplemental Figure 2). Based on orthology with *Candida albicans* and *Saccharomyces*
267 *cerevisiae*, the identified loci have putative functions in the cytoskeleton, cell wall organization, and
268 ion transport (Table 2).
269
270 Laboratory validation of *in silico*-identified markers
271 To validate our *in silico* typing, we designed primers to PCR amplify the four identified loci (Table 3).
272 Using a representative 14 samples from our dataset, including four outbreak strains expected to have
273 identical sequence type profiles, we amplified and Sanger sequenced the four loci. In each sample,
274 we detected the expected variants from the *in silico* analysis by Sanger sequencing and did not
275 discover any variants that were not in the whole genome data (Supplemental Table 2). Resultantly,
276 we could correctly assign samples to the nine unique sequence types predicted from the *in silico*
277 typing and the six genetic clusters assigned to them by the whole genome phylogeny. As expected,
278 the four outbreak samples had identical sequences at all four loci and were assigned to the same
279 sequence type. An important note is that Sanger sequencing providers may not report dual peaks
280 caused by heterozygous sites using IUPAC nucleotide codes. As this was the case for us, this was
281 resolved by manually recoding positions in the chromatograms where equal sized dual peaks were
282 recorded (for example, a double peak reporting an A and a G were recoded to the corresponding
283 IUPAC base of R). To facilitate the typing of novel isolates by future researchers and clinical labs,
284 FASTA alignments containing nucleotide sequence from each sequence type of each of the four
285 makers are provided in Supplemental File 2.

286

287 Demonstration of the utility of the typing scheme for rapidly identifying outbreak clones

288 To demonstrate the value of our typing scheme for the rapid identification of clones within an outbreak
289 setting, we tested our typing loci on a set of seven novel isolates that were not part of the dataset
290 used to create the typing scheme. These isolates were collected from three hospitals in Berlin and
291 were suspected to be additional outbreak cases based on their antifungal susceptibility profile (FLC^R,
292 VOR^R, POS^S, ITR^S) and their isolation from hospitals with previous outbreak cases. PCR amplification
293 and Sanger sequencing of the four typing loci revealed that their sequence type was identical to the
294 outbreak isolates analyzed by WGS, and strongly suggesting that they were additional outbreak cases
295 (Supplemental Table 2). Taken together, we selected and validated four loci that can be used for
296 sequence-based typing of *C. parapsilosis* and demonstrated their value in rapidly identifying closely
297 related samples.

298

299 **Discussion**

300 In this work, we use genomic epidemiology coupled with patient tracking to investigate an outbreak
301 of azole-resistant *C. parapsilosis* that occurred across several years and medical centers in Berlin,
302 Germany. *C. parapsilosis* is an emerging pathogen and our report joins the increasing list of
303 nosocomial outbreaks reported for this pathogen. Similar to *Candida auris*, this pathogen is often
304 drug resistant and has the propensity to spread patient-to-patient. Due to slow outbreak kinetics and
305 the low case density, the outbreak we describe might have gone unnoticed if not for the atypical
306 susceptibility pattern that was detected by clinicians and clinical microbiologists. Based on the
307 increasing number of outbreaks for *C. parapsilosis* around the globe and the need to rapidly
308 discriminate between isolates, we also establish and validate a set of four loci for sequence-based
309 strain typing.

310

311 The reconstructed transmission dynamics suggest that persistent outbreaks of *C. parapsilosis* may
312 result from the reintroduction of outbreak strains from other hospitals, underlining the myriad
313 complex challenges of outbreak control and the need for heightened infection control measures. In
314 our case, patient transfers were a common occurrence among cases and facilitated inter-hospital
315 transmission and subsequent intra-hospital transmission. Unfortunately, the high-resolution
316 temporal and spatial monitoring of patients is lacking for other outbreak descriptions, so it is unclear
317 the degree to which this has driven other reported *C. parapsilosis* outbreaks. However, the pattern
318 of prolonged but tightly related isolates causing disease across several years and multiple medical
319 centers matches other outbreak reports from France, Spain, China, and Brazil [15,17,31,36], suggesting
320 that reintroductions and/or patient transfers could have also contributed to the prolonged nature of
321 these outbreaks. One limitation to our study is that while patients without infection were screened
322 for colonization during the first wave of cases at Hospital #1, a comprehensive screening of all the
323 involved hospitals and throughout the extended period of the outbreak was not performed. This
324 obscures an understanding of the relationship between colonization and subsequent infection

during the outbreak and a mechanistic understanding of how *C. parapsilosis* is transmitted within healthcare facilities.

The finding of azole-resistant strains from Turkey, South Korea, and Kuwait that were very closely related to the strains spreading within healthcare facilities in Berlin highlights the global challenge of antifungal resistance and the emergence of resistant *Candida* spp that are easily transmitted from patient-to-patient. Whether this drug-resistant lineage has been present in each of these regions for a longer period or was recently imported is an open question. We also cannot deduce whether it was originally imported to Germany from any of these sites. The observation of highly related, drug resistant clones observed on multiple continents mirrors the behavior of *Candida auris*. Nearly identical clades of this pathogen have been reported on multiple continents and continue to spread globally [37,38], underscoring the fact that *C. parapsilosis* has the potential to become a problem of similar scale. Additionally, three phylogenetically distinct Y132F lineages of fluconazole-resistant *C. parapsilosis* were discovered in Germany during the study period, each with multiple isolates reported from distinct geographic areas. This demonstrats that fluconazole-resistant *C. parapsilosis* is endemic in Germany, mirroring what has been recently reported in other countries [4], and highlights the changing epidemiology of the pathogen that clinicians should be aware of.

Microsatellite typing in *C. parapsilosis* is typically performed using four loci and provides a discriminatory power between 0·63 [39] and 0·99 [21]. The four loci MLST established in this study achieved a discriminatory power of 0·87, within the range of what is achieved using microsatellite markers. Importantly, the MLST methodology could unambiguously compare samples from >9 years and around the world, a task not possible using existing microsatellite strategies based on gel electrophoresis profiles. Moreover, it recapitulated whole genome-based phylogenetic relationships with ~97% accuracy, demonstrating its use not only in outbreak tracking and epidemiology but in phylogenetic contexts. Taken together, our study provides a high resolution description of a fluconazole-resistant *C. parapsilosis* outbreak, identifies hospital transfers as recurring source of pathogen introductions, and defines a sequence-based typing scheme for epidemiology and rapid identification of potential outbreaks.

Methods

Outbreak cases and sample collection

The 68 novel strains of *C. parapsilosis* sequenced in this study were sent to the German National Reference Centre for Invasive Fungal Infections (NRZMyk) by German healthcare facilities and laboratories for species confirmation and susceptibility testing between 2014 and 2022 (Table 1). For tracking outbreak dynamics, admission and discharge records were obtained for all outbreak patients beginning January 2018, corresponding to one year before the first outbreak case, and extending to December 2022. Patient information was obtained from the sample submission form of the NRZMyk. Samples were de-identified via the assignment of an NRZ ID (e.g. NRZ-XXXX) which cannot be linked

364 to a patient, hospital, or region. The Ethics Committee of the University Hospital Jena approved this
365 study under registration number 2024-3255-Daten.
366
367 Strains submitted to the NRZMyk from outside Berlin were selected for whole genome sequencing
368 (WGS) if they displayed resistance to one or more medical azole. For cities with azole resistant
369 samples selected for sequencing, an additional number of azole susceptible strains were included
370 from this city. In total 44 azole resistant strains (FLC^R, VOR^R, POS^S, or ITR^R) and 25 fully azole
371 susceptible strains were analyzed. Strains were identified as *C. parapsilosis* prior to WGS by Sanger
372 sequencing of the ITS locus. This locus was amplified with the forward primer V9G (5′-
373 TTACGTCCCTGCCCTTTGTA-3′) and the reverse primer LS266 (5′-
374 GCATTCCCAAACAACTCGACTC-3). The resulting sequences were compared against
375 *C. parapsilosis* sequences in the NCBI database using BLAST.
376
377 Antifungal susceptibility testing
378 *In vitro* antifungal susceptibility testing was performed by broth microdilution technique following the
379 European Committee on Antimicrobial Susceptibility Testing (EUCAST) standard methodology for the
380 following antifungals: fluconazole (FLC; Pfizer Inc., Peapack, NJ, USA), itraconazole (ITR,
381 Chemicalpoint, Deisenhofen, Germany), posaconazole (POS; MSD, Rahway, NJ, USA), and
382 voriconazole (VOR; Pfizer Inc., Peapack, NJ, USA). Minimum inhibitory concentration (MICs) were
383 assessed with a nephelometer (Labsystems Nepheloskan Ascent Microplate Reader Type 750) after
384 24 h of incubation at 35°C. The endpoint of growth was as a 50% inhibition in comparison with the
385 drug-free control. Reference strain *C. parapsilosis* ATCC 22019 was used as a quality control.
386
387
388 DNA extraction and genome sequencing
389 For extraction of genomic DNA, single colonies of each putative strain were isolated and grown
390 overnight on YPD plates at 35°C. Fungal material was transferred to a tube containing acid-washed
391 glass beads and lysis buffer (50 mM Tris, 50 mM sodium EDTA, 3% [wt/vol] sodium dodecyl sulfate
392 [SDS] [pH 8]). The samples were homogenized using a vortex adapter, followed by incubation in a
393 thermomixer. Next, cellular debris was pelleted by centrifugation, and the supernatant was transferred
394 to a new tube. RNAse treatment was performed for 1h at 37°C. Subsequently, 3M sodium acetate
395 was added to the tube, samples were incubated at -20°C, then centrifuged and the supernatant
396 transferred to a new tube. DNA was then precipitated using EtOH and the resulting pellet washed
397 twice. DNA extracts that did not meet the requirement of purity were purified using the Monarch
398 genomic DNA purification kit (New England BioLabs Inc., Frankfurt a. M., Germany). Library
399 preparation and 2x250bp paired end sequencing was performed by GeneWiz using the Illumina
400 MiSeq platform.
401
402 Genome sequencing, quality control, alignment, and variant identification

403 Raw sequence reads were trimmed for quality using TrimGalore v0.6.6 in paired end mode using the
404 default error rate threshold of 0.1. Sequence data quality was assessed before and after trimming
405 using FastQC v0.11.9. Mapping to the CDC317 reference genome (FungiDB, release 50) was
406 performed using bwa-mem2 v2.2.1 [40]. Duplicates were marked using the MarkDuplicates function of
407 Picard v2.24.0. Alignment and genome coverage statistics were assessed using
408 CollectAlignmentMetrics and CollectWgsMetrics from the Picard tools (v2.24.0). Only samples with a
409 mean genome coverage of 10x and >90% of reads mapping to the CDC317 reference genome were
410 included in downstream analyses. Genome coverage and fraction of mapped reads for each sample
411 are listed in Supplemental File 1. Variant calling was performed using Freebayes v1.3.2-dirty. Variants
412 were filtered using vcffilter v0.2 and the following filter string: "QUAL > 30 & DP > 5 & SAF > 0 & SAR
413 > 0 & RPR > 1 & RPL > 1". Variants were annotated using snpEFF v5.1 [41] and a custom database
414 made from the CDC317 reference genome files and using the alternative yeast nuclear and yeast
415 mitochondrial codon tables. For defining features within upstream and downstream regions, a cutoff
416 of 500 bp was used.
417
418 SNV-based phylogeny and identification of phylogenetic clusters
419 For the phylogeny and SNV distance-based analyses, only variant positions with non-zero coverage
420 in all samples and SNVs were included. Variant positions with zero coverage, as well as all insertion
421 or deletion variants were removed from the vcf files. The resulting multi-sample VCF was converted
422 to a phylip alignment with a length of 38,567 positions. Pairwise SNV distances between isolates were
423 calculated using SNP Dist [42]. A maximum-likelihood (ML) phylogeny was inferred by IQ-TREE 2
424 v2.2.0.3 [43] using the GTR+ASC model to account for ascertainment bias as constant positions were
425 excluded. The ModelFinder module [44] of IQ-TREE was used to confirm that GTR was also the best
426 fitting model based on Akaike and Corrected Akaike Information Criterion. The *ERG11* amino acid
427 annotations on the phylogeny are reported relative to the *C. parapsilosis* reference genome CDC317,
428 except for cDNA position 395, corresponding to amino acid position 132. CDC317 is azole resistant
429 and contains a Y132F substitution relative to the EUCAST control strain ATCC 22019. Given this,
430 samples reporting no variant at position 395 of the CDS were re-coded to a Y132F amino acid
431 substitution and samples recording a 395T>A change relative to CDC317 were reported as not having
432 an amino acid substitution at position 132. The annotated phylogenies were created using ggtree [45].
433
434 Identification of phylogenetic clusters was performed using TreeCluster v1.0.3 [46] using a cut threshold
435 of 0.700. The identified clusters were labeled in accordance with the five major clades described in
436 [35]. As TreeCluster calculated sufficient intraclade genetic distance to further subdivide clades 2 and
437 4, clusters belonging to these clades were assigned the IDs of 2a, 2b and 4a-g.
438
439 Identification of discriminatory loci for sequence typing
440 Variant profiles of complete gene coding regions were obtained from vcf files and the number of
441 distinct profiles present among the isolates calculated by hierarchical clustering analysis. Protein

coding regions with no discriminatory power were discarded and the top 212 CDS were further evaluated for their discriminatory power in 750 bp sliding windows with a step size of 50 bp. Windows including intronic sequence were excluded for downstream selection.

To ensure that potential sequence typing loci were present in the genomes of all samples as a single copy, loci were screened for the presence of the presence of copy number variations (CNVs) using Control-FREEC v11.6 [47] , DELLY v1.1.6 [48], and CNVator v0.4.1 [49]. Multiple tools utilizing different approaches for detecting copy number changes, including tools utilizing read depth, GC normalization, and split read mapping, were used as many CNV calling software suffer from false positives. DELLY and CNVator were run with default parameters. As FREEC uses GC normalization which varies by organism, 0.25 was used for the minExpectedGC and 0.45 for maxExpectedGC. Loci that fell within CNVs detected by two or more callers were excluded from further consideration. Several samples from the publicly available data displayed prominent chromosome-end bias [50]. As this overrepresented coverage at the chromosome ends reduces confidence for CNV calling, these samples were excluded from CNV analysis.

In silico strain typing

Locus sequence was extracted *in silico* from the vcf files using the consensus function of BCFtools v1.9 [51] and aligned using MUSCLE v3.8 via the R package msa. As BCFtools is not able to convert heterozygous sites to their corresponding IUPAC base, this was performed after alignment using an in-house R script. Concatenated alignments from the four loci were built into a maximum-likelihood phylogeny using RaxML (v2.0.10) using the GTR substitution model and 100 bootstrap replicates. The resulting phylogeny was visualized using ggtree [45].

Hunter's discriminatory power was calculated using the following equation [52]:

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{s} x_j(x_j - 1)$$

PCR-amplification and Sanger sequencing of sequence typing markers

Primers for the selected loci were designed using NCBI Primer Blast and are listed in Table 3, along with the predicted fragment size. PCR was performed using the following cycling parameters: 4 min at 95°C, 40 cycles each consisting of a 45 sec denaturation at 95°C followed by 45 sec of annealing at 57°C for CPAR2_808110 and CPAR2_212940 and 59°C for CPAR2_101470 and CPAR2_108720, and an extension at 72°C for 2 min, and a final 10 min at 72°C. Sanger sequencing was performed by LGC (Berlin, Germany). As our Sanger sequence data was not reported with IUPAC codes representing heterozygous sites, these were manually corrected by manual review of the sequencing chromatogram. We observed cases where the CPAR2_101470 sequencing reaction failed to cover

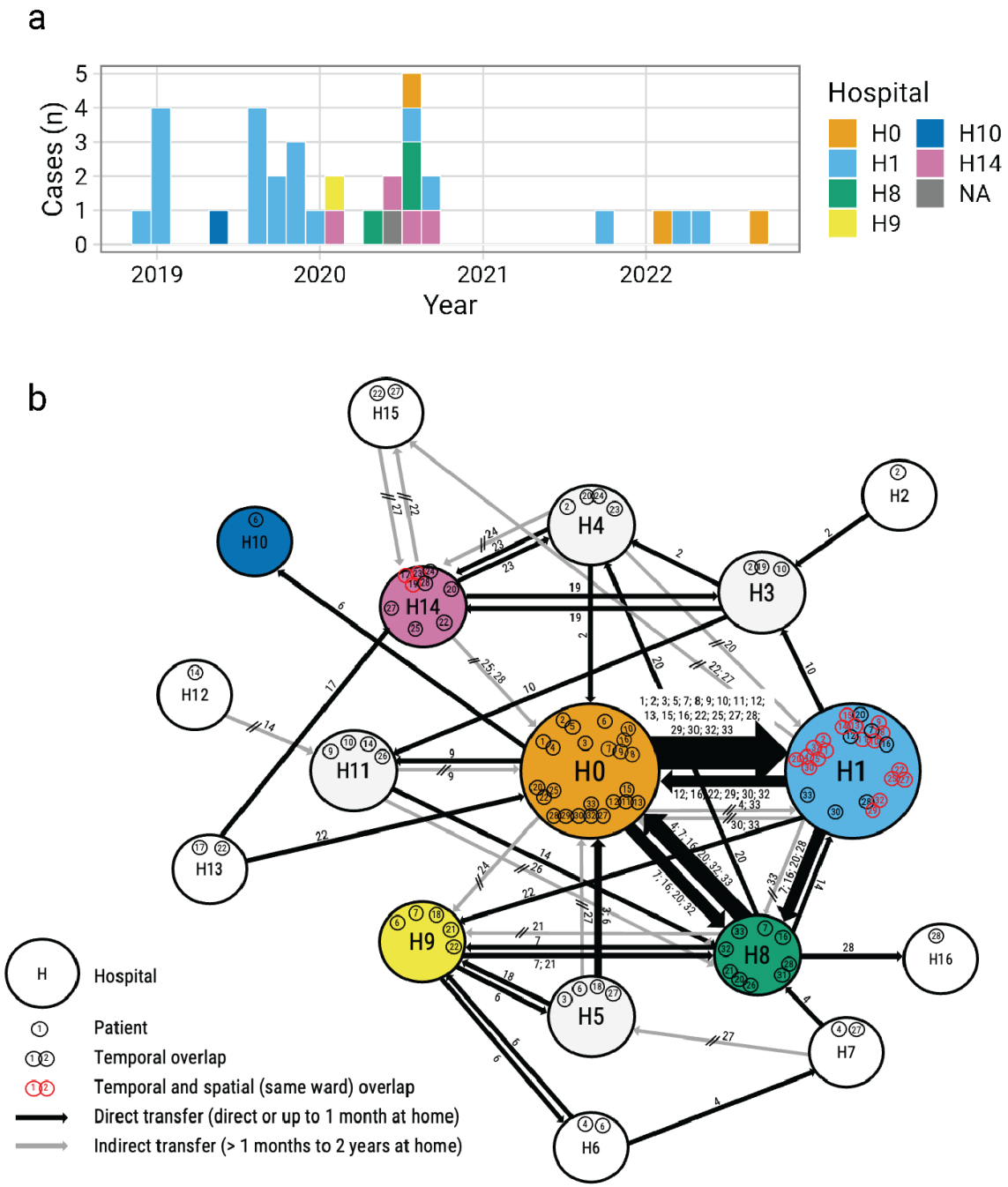
478 the entire amplicon when using the F primer, missing the final variant which we resolved by performing
479 a second Sanger reaction using the R strand. This affects the differentiation of clusters 2b and 2d.
480
481 Data availability
482 Raw FASTQ files for novel isolates sequenced in this study were uploaded to the NCBI Sequence
483 Read Archive and are publicly available under BioProject PRJNA996760. Accession numbers for
484 the publicly available sequence data used in this study are listed in Supplemental File 1. Genbank
485 records for the Sanger sequencing of the MLST loci performed in this study are listed in
486 Supplemental Table 2.
487
488 Acknowledgements
489 We are grateful to Peggy Karanatsiou of the University of Jena and Philipp Hupel, Carmen Karkowski,
490 and Christiane Weigel of the NRZMyk for technical assistance. We also thank MSD and Pfizer for
491 providing the drugs for antifungal susceptibility testing free of charge.
492
493 Funding
494 This project was supported by The Federal Ministry for Education and Science (Bundesministerium
495 für Bildung und Forschung) within the framework of InfectControl 2020 (Projects FINAR 2.0, grant
496 03ZZ0834). The work of the German National Reference Center for Invasive Fungal Infections is
497 supported by the Robert Koch Institute from funds provided by the German Ministry of Health (grant
498 1369-240). AEB is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research
499 358 Foundation) under Germany's Excellence Strategy – EXC 20151 – Project-ID 390813860.
500
501 Authors' contributions
502 AEB and GW conceptualized the study. PB, FRK, BG, RP, OK, GW and AEB contributed to
503 investigation, methodology, and formal analysis. PB, GW, and AEB were responsible for visualization.
504 AEB, PB, and GW were responsible for writing the original draft. All authors participated in reviewing
505 and editing the manuscript.
506
507 Conflicts of Interest
508 The authors have no conflicts of interest to report.
509
510 **References**
511 1. Mesini, A. *et al.* Changing epidemiology of candidaemia: Increase in fluconazole-resistant
512 Candida parapsilosis. *Mycoses* **63**, 361–368 (2020).
513 2. Tóth, R. *et al.* Candida parapsilosis: From genes to the bedside. *Clin. Microbiol. Rev.* **32**, 1–38
514 (2019).
515 3. Rhodes, J. & Fisher, M. C. Global epidemiology of emerging Candida auris. *Curr. Opin.*
516 *Microbiol.* **52**, 84–89 (2019).
517 4. Daneshnia, F. *et al.* Worldwide emergence of fluconazole-resistant Candida parapsilosis:

518 current framework and future research roadmap. *Lancet Microbe* **4**, e470–e480 (2023).
519 5. Arastehfar, A. *et al.* Clonal Candidemia Outbreak by Candida parapsilosis Carrying Y132F in
520 Turkey: Evolution of a Persisting Challenge. *Front. Cell. Infect. Microbiol.* **11**, (2021).
521 6. Branco, J. *et al.* Fluconazole and Voriconazole Resistance in Candida parapsilosis Is Conferred
522 by Gain-of-Function Mutations in MRR1 Transcription Factor Gene. *Antimicrob. Agents*
523 *Chemother.* **59**, 6629–6633 (2015).
524 7. Branco, J. *et al.* Clinical azole cross-resistance in Candida parapsilosis is related to a novel
525 MRR1 gain-of-function mutation. *Clin. Microbiol. Infect.* **28**, 1655.e5-1655.e8 (2022).
526 8. Daneshnia, F. *et al.* Determinants of fluconazole resistance and echinocandin tolerance in C.
527 parapsilosis isolates causing a large clonal candidemia outbreak among COVID-19 patients in a
528 Brazilian ICU. *Emerg. Microbes Infect.* **11**, 2264–2274 (2022).
529 9. Pinhati, H. M. S. *et al.* Outbreak of candidemia caused by fluconazole resistant Candida
530 parapsilosis strains in an intensive care unit. *BMC Infect. Dis.* **16**, 4–9 (2016).
531 10. Qi, L. *et al.* Nosocomial outbreak of Candida parapsilosis sensu stricto fungaemia in a neonatal
532 intensive care unit in China. *J. Hosp. Infect.* **100**, e246–e252 (2018).
533 11. Thomaz, D. Y. *et al.* An Azole-Resistant Candida parapsilosis Outbreak: Clonal Persistence in
534 the Intensive Care Unit of a Brazilian Teaching Hospital. *Front. Microbiol.* **9**, (2018).
535 12. Arastehfar, A. *et al.* First report of candidemia clonal outbreak caused by emerging fluconazole-
536 resistant Candida parapsilosis isolates harboring Y132F and/or Y132F+K143R in Turkey.
537 *Antimicrob. Agents Chemother.* (2020) doi:10.1128/AAC.01001-20.
538 13. Corzo-Leon, D. E., Peacock, M., Rodriguez-Zulueta, P., Salazar-Tamayo, G. J. & MacCallum,
539 D. M. General hospital outbreak of invasive candidiasis due to azole-resistant Candida
540 parapsilosis associated with an Erg11 Y132F mutation. *Med. Mycol.* 1–8 (2020)
541 doi:10.1093/mmy/myaa098.
542 14. Danilo Y. Thomaz, João N. de Almeida Jr. *, O. N. E. S., Gilda M.B. Del Negro, Gabrielle O. M.
543 H. Carvalho, V. M. F. G., Maria Emilia B. de Souza, Amir Arastehfar, Carlos H. Camargo, A. L.,
544 Motta, Flávia Rossi, David S. Perlin, Maristela P. Freire, Edson Abdala, G., & Benard.
545 Environmental clonal spread of azole-resistant Candida parapsilosis with Erg11-Y132F mutation
546 causing a large candidemia outbreak in a Brazilian Cancer Referral Center. *J. Fungi* (2021).
547 15. Fekkar, A. *et al.* Hospital Outbreak of Fluconazole-Resistant Candida parapsilosis: Arguments
548 for Clonal Transmission and Long-Term Persistence. *Antimicrob. Agents Chemother.* **65**,
549 e02036-20 (2021).
550 16. Guinea, J. *et al.* Whole genome sequencing confirms Candida albicans and Candida
551 parapsilosis microsatellite sporadic and persistent clones causing outbreaks of candidemia in
552 neonates. *Med. Mycol.* **60**, myab068 (2022).
553 17. Thomaz, D. Y. *et al.* A Brazilian Inter-Hospital Candidemia Outbreak Caused by Fluconazole-
554 Resistant Candida parapsilosis in the COVID-19 Era. *J. Fungi* **8**, 100 (2022).
555 18. Trevijano-Contador, N. *et al.* Global emergence of resistance to fluconazole and voriconazole in
556 Candida parapsilosis in tertiary hospitals in Spain during the COVID-19 pandemic. *Open Forum*
557 *Infect. Dis.* ofac605 (2022) doi:10.1093/ofid/ofac605.
558 19. Hernández-Castro, R. *et al.* Outbreak of Candida parapsilosis in a neonatal intensive care unit:
559 A health care workers source. *Eur. J. Pediatr.* **169**, 783–787 (2010).
560 20. Presente, S. *et al.* Hospital Clonal Outbreak of Fluconazole-Resistant Candida parapsilosis
561 Harboring the Y132F ERG11p Substitution in a French Intensive Care Unit. *Antimicrob. Agents*
562 *Chemother.* **67**, e01130-22 (2023).
563 21. Sabino, R. *et al.* New polymorphic microsatellite markers able to distinguish among Candida
564 parapsilosis sensu stricto isolates. *J. Clin. Microbiol.* **48**, 1677–1682 (2010).
565 22. Garcia-Hermoso, D., Desnos-Ollivier, M. & Bretagne, S. Typing Candida Species Using
566 Microsatellite Length Polymorphism and Multilocus Sequence Typing. in *Candida Species:*
567 *Methods and Protocols* (eds. Calderone, R. & Cihlar, R.) 199–214 (Springer, 2016).

568 doi:10.1007/978-1-4939-3052-4_15.
569 23. de Valk, H. A., Meis, J. F. G. M. & Klaassen, C. H. W. Microsatellite based typing of Aspergillus
570 fumigatus: Strengths, pitfalls and solutions. *J. Microbiol. Methods* **69**, 268–272 (2007).
571 24. Hamprecht, A. *et al.* Candida auris in Germany and Previous Exposure to Foreign Healthcare.
572 *Emerg. Infect. Dis.* **25**, 1763–1765 (2019).
573 25. Lockhart, S. R. *et al.* Simultaneous emergence of multidrug-resistant candida auris on 3
574 continents confirmed by whole-genome sequencing and epidemiological analyses. *Clin. Infect.*
575 *Dis.* **64**, 134–140 (2017).
576 26. Rhodes, J. *et al.* Genomic epidemiology of the UK outbreak of the emerging human fungal
577 pathogen Candida auris article. *Emerg. Microbes Infect.* **7**, (2018).
578 27. Asadzadeh, M., Dashti, M., Ahmad, S., Alfouzan, W. & Alameer, A. Whole-Genome and
579 Targeted-Amplicon Sequencing of Fluconazole-Susceptible and -Resistant Candida
580 parapsilosis Isolates from Kuwait Reveals a Previously Undescribed N1132D Polymorphism in
581 CDR1. *Antimicrob. Agents Chemother.* **65**, 10.1128/aac.01633-20 (2021).
582 28. Arastehfar, A. *et al.* Candidemia Among Coronavirus Disease 2019 Patients in Turkey Admitted
583 to Intensive Care Units: A Retrospective Multicenter Study. *Open Forum Infect. Dis.* **9**, ofac078
584 (2022).
585 29. Kim, T. Y. *et al.* Evolution of Fluconazole Resistance Mechanisms and Clonal Types of Candida
586 parapsilosis Isolates from a Tertiary Care Hospital in South Korea. *Antimicrob. Agents*
587 *Chemother.* **66**, e00889-22 (2022).
588 30. Demirci-Duarte, S., Arikan-Akdagli, S. & Gülmez, D. Species distribution, azole resistance and
589 related molecular mechanisms in invasive Candida parapsilosis complex isolates: Increase in
590 fluconazole resistance in 21 years. *Mycoses* **64**, 823–830 (2021).
591 31. Díaz-García, J. *et al.* Evidence of Fluconazole-Resistant Candida parapsilosis Genotypes
592 Spreading across Hospitals Located in Madrid, Spain and Harboring the Y132F ERG11p
593 Substitution. *Antimicrob. Agents Chemother.* **66**, e00710-22 (2022).
594 32. Grossman, N. T., Pham, C. D., Cleveland, A. A. & Lockhart, S. R. Molecular Mechanisms of
595 Fluconazole Resistance in Candida parapsilosis Isolates from a U.S. Surveillance System.
596 *Antimicrob. Agents Chemother.* **59**, 1030–1037 (2015).
597 33. Berkow, E. L. *et al.* Multidrug Transporters and Alterations in Sterol Biosynthesis Contribute to
598 Azole Antifungal Resistance in Candida parapsilosis. *Antimicrob. Agents Chemother.* **59**, 5942–
599 5950 (2015).
600 34. Asadzadeh, M., Ahmad, S., Al-Sweih, N. & Khan, Z. Epidemiology and Molecular Basis of
601 Resistance to Fluconazole Among Clinical *Candida parapsilosis* Isolates in Kuwait. *Microb.*
602 *Drug Resist.* **23**, 966–972 (2017).
603 35. Bergin, S. A. *et al.* Systematic Analysis of Copy Number Variations in the Pathogenic Yeast
604 Candida parapsilosis Identifies a Gene Amplification in RTA3 That is Associated with Drug
605 Resistance. *mBio* **13**, e01777-22 (2022).
606 36. Wang, H. *et al.* Investigation of an unrecognized large-scale outbreak of Candida parapsilosis
607 sensu stricto fungaemia in a tertiary-care hospital in China. *Sci. Rep.* **6**, 27099 (2016).
608 37. Chow, N. A. *et al.* Multiple introductions and subsequent transmission of multidrug-resistant
609 Candida auris in the USA: a molecular epidemiological survey. *Lancet Infect. Dis.* **3099**, 1–8
610 (2018).
611 38. Chow, N. A. *et al.* Tracing the evolutionary history and global expansion of Candida auris using
612 population genomic analyses. *mBio* **11**, 2020.01.06.896548 (2020).
613 39. Romeo, O. *et al.* Microsatellite-based genotyping of Candida parapsilosis sensu stricto isolates
614 reveals dominance and persistence of a particular epidemiological clone among neonatal
615 intensive care unit patients. *Infect. Genet. Evol.* **13**, 105–108 (2013).
616 40. Vasimuddin, Md., Misra, S., Li, H. & Aluru, S. Efficient Architecture-Aware Acceleration of BWA-
617 MEM for Multicore Systems. in *2019 IEEE International Parallel and Distributed Processing*

Symposium (IPDPS) 314–324 (2019). doi:10.1109/IPDPS.2019.00041.
41. Cingolani, P. *et al.* A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)* **6**, 80–92 (2012).
42. Seemann, T. Pairwise SNP distance matrix from a FASTA sequence alignment. https://github.com/tseemann/snp-dists (2018).
43. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
44. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
45. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).
46. Balaban, M., Moshiri, N., Mai, U., Jia, X. & Mirarab, S. TreeCluster: Clustering biological sequences using phylogenetic trees. *PLOS ONE* **14**, e0221068 (2019).
47. Boeva, V. *et al.* Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics* **28**, 423–425 (2012).
48. Rausch, T. *et al.* DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012).
49. Abyzov, A., Urban, A. E., Snyder, M. & Gerstein, M. CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* **21**, 974–984 (2011).
50. Abbey, D. A. *et al.* YMAP: a pipeline for visualization of copy number variation and loss of heterozygosity in eukaryotic pathogens. *Genome Med.* **6**, 100 (2014).
51. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
52. Hunter, P. R. Reproducibility and indices of discriminatory power of microbial typing methods. *J. Clin. Microbiol.* **28**, 1903–1905 (1990).

646 **Figures**



647

648 **Figure 1. Timeline and hospital network dynamics of a *Candida parapsilosis* outbreak in**
649 **Berlin, Germany.** (a) Timeline histogram of the 33 outbreak cases. Dates are taken from the date
650 of first positive culture. Color indicates the inpatient hospital at the time of isolation. A single case
651 was identified in an ambulatory patient, which is indicated with “NA” for the inpatient hospital. (b)
652 Inpatient stays of the outbreak cases represented as a network, where nodes (large circles)
653 represent hospitals in the Berlin area and edges (lines) represent direct or indirect patient transfers
654 between two institutions. Edge thickness depicts the relative number of transfers between two
655 nodes. Individual patient case numbers are indicated within the node and along the edge. For the
656 network, admission and discharge data beginning 12 months prior to the first detected outbreak and
657 extending until the end of 2022 was utilized.

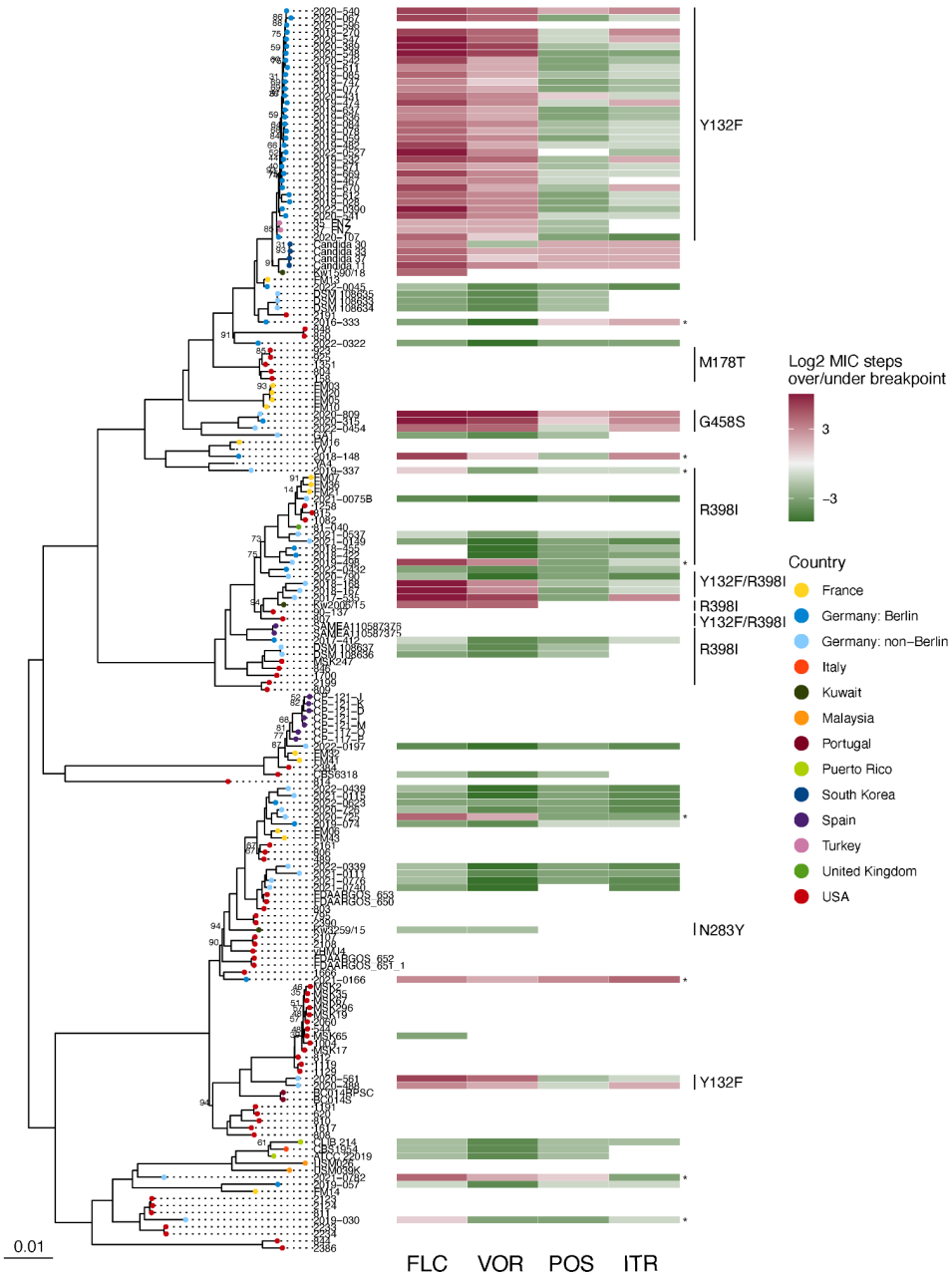


Figure 2. Whole genome phylogeny of *Candida parapsilosis* isolates from a global distribution and their azole susceptibility profiles and *ERG11* polymorphisms. Maximum-likelihood phylogeny constructed from SNV data from 38,567 variable positions. Tree is shown rooted at the midpoint. To increase tree readability, samples from the same sequencing study that showed little genetic differentiation were randomly downsampled to a maximum of 4 samples per sequencing project - with the exception of the samples from the samples described for the first time in our study. Branches are supported by ultrafast bootstrap values of >0.95 unless otherwise indicated. Where available, azole susceptibility data relative to EUCAST breakpoints and *ERG11* polymorphisms are indicated. Strains that were resistant to one or more azole but where no polymorphisms in *ERG11* were detected are marked with an asterisk.

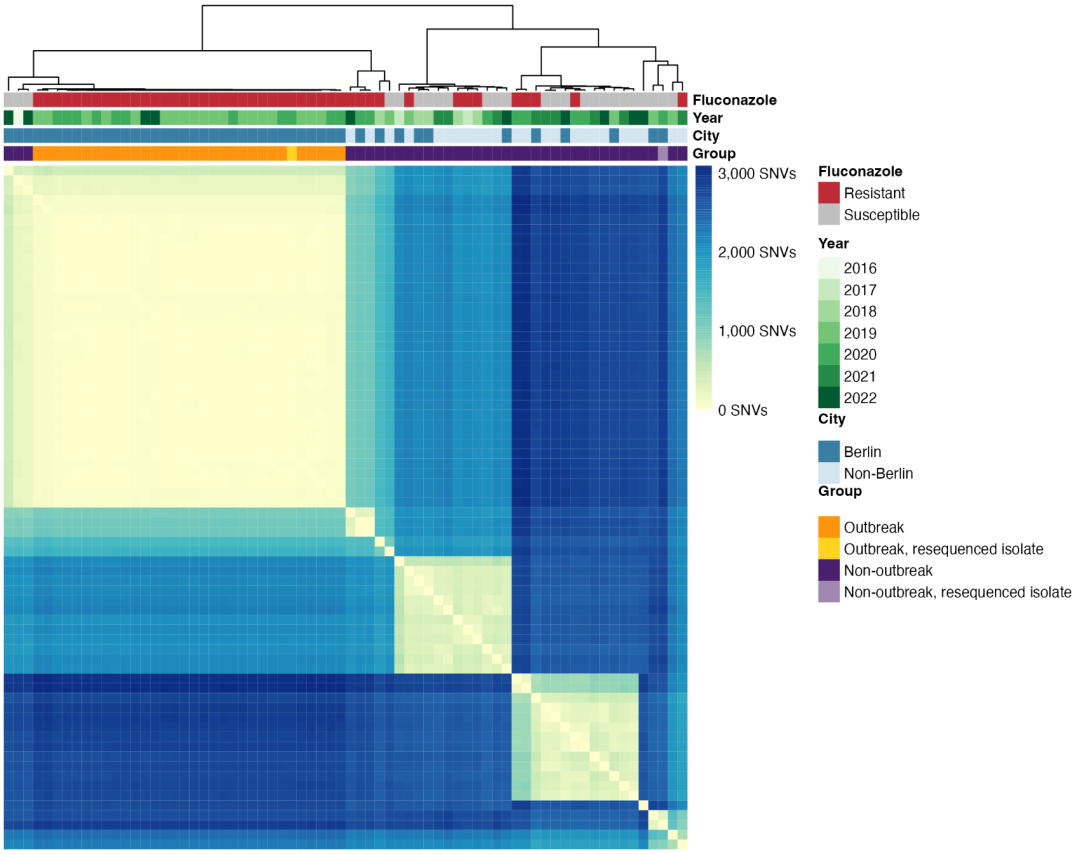


Figure 3. SNV matrix of the German isolates sequenced in this study.
Heatmap showing the pairwise SNV differences between the novel German isolates described in this study. Metadata bands indicate (from top to bottom): fluconazole susceptibility, year of sample isolation, and whether the sample is from the city of Berlin (including both outbreak and non-outbreak strains) or another city in Germany.

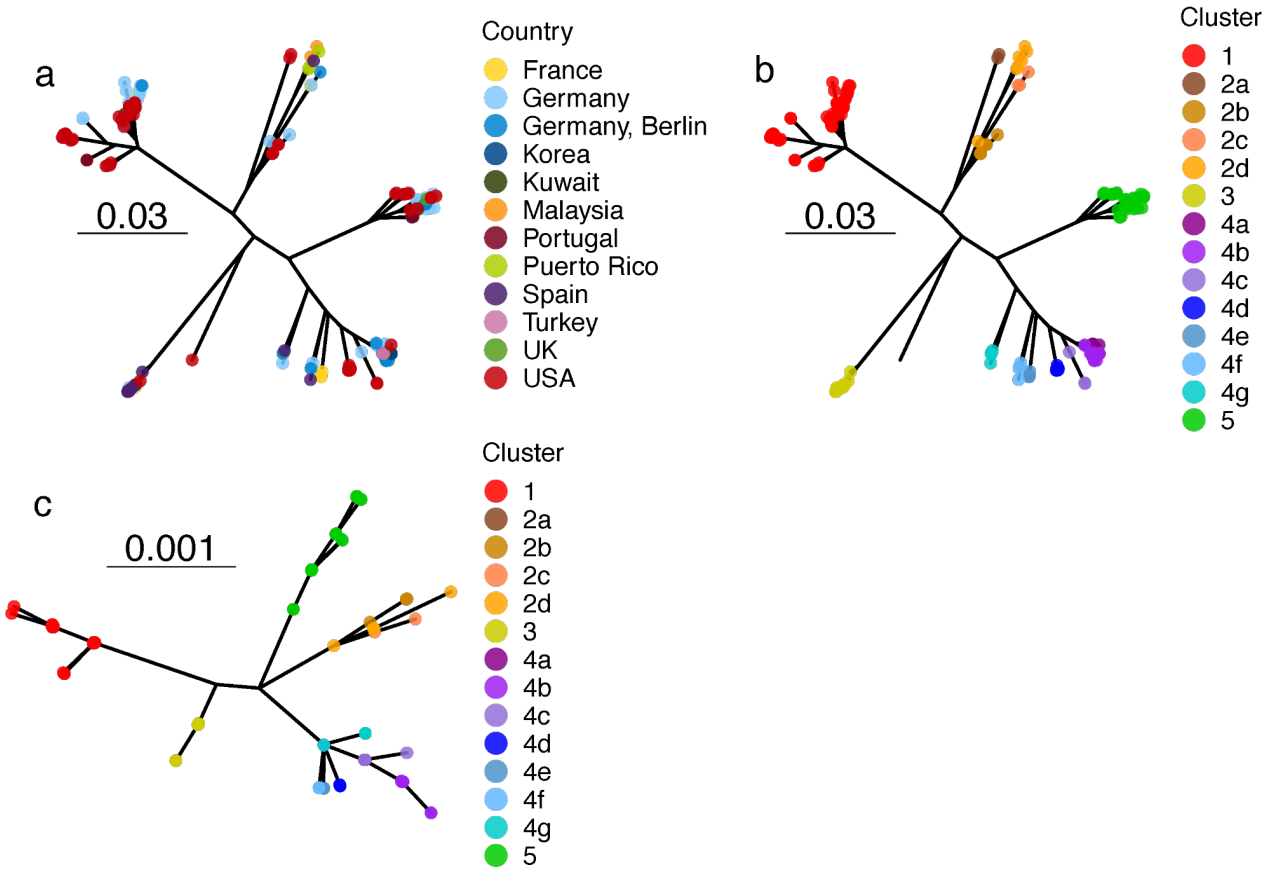


Figure 4. **Comparison of whole genome and four loci phylogenies of *C.parapsilosis***. (a-b) Maximum-likelihood whole genome SNV phylogeny with tips colored by country of origin (a) or genetic cluster based on whole-genome data (b). (c) Maximum-likelihood phylogeny constructed from ~700 bp of sequence from CPAR2_101470, CPAR2_108720, CPAR2_212940, and CPAR2_808110. Tip color indicates the genetic cluster samples were assigned based on the whole genome phylogeny in panel b.

683 **Table 1. Patient characteristics, isolate sources, and MIC data for the newly-described and sequenced**
684 **German *C. parapsilosis* isolates in this study.**
685

	Outbreak (n = 33)	Non-outbreak (n = 38)
Patient		
Sex F	15 (45.5%)	-
Age (years)	Mean 68 (range 33 - 83)	-
Source		
Wound	12 (38.7%)	5 (13.2%)
Central venous catheter	9 (29.0%)	7 (18.4%)
Urine	5 (16.1%)	0 (0.0%)
Blood	2 (6.5%)	17 (44.7%)
Other	3 (9.7%)	9 (23.7%)
MIC (mg/l)		Median (Min - Max)
Fluconazole	64 (16 - >64)	1 (0.25 - >64)
Voriconazole	1 (0.25 - 4)	0.03 (≤0.016 - 8)
Posaconazole	0.03 (≤0.016 - 0.25)	≤0.016 (≤0.016 - 0.5)
Itraconazole	0.125 (≤0.016 - 0.5)	0.125 (≤0.016 - 1)
Isavuconazole	≤0.016 (≤0.016 - 0.125)	≤0.016 (≤0.016 - 2)
Anidulafungin	1 (0.5 - 4)	1 (0.5 - 2)
Amphotericin B	0.5 (0.25 - 1)	0.5 (0.06 - 1)

686

687 **Table 2. Loci used for *Candida parapsilosis* sequence typing.**

688

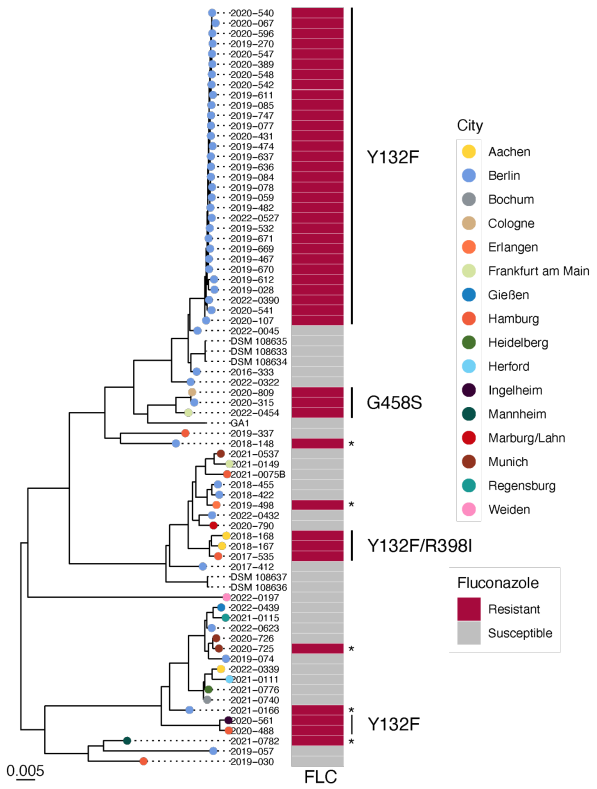
Gene	Chromosome	Position	Predicted Function	*C. albicans* Ortholog
CPAR2_101470	Contig005569	298860-299861	Structural constituent of cytoskeleton Mitotic spindle orientation Nuclear migration	orf19.3976 (*JNM1*)
CPAR2_108720	Contig005569	1835814-1833166	Cellular response to heat Fungal-type cell wall organization	orf19.3038 (*TPS2*)
CPAR2_212940	Contig005809	2776693-2775623	Nucleotide binding Oxidoreductase activity	orf19.2036 (*DDD1*)
CPAR2_808110	Contig005807	1878755-1876554	Ferric-chelate reductase activity Copper ion import Iron ion transport Plasma membrane	orf19.1932 (*CFL4*)

689 Gene IDs, genomic coordinates, and descriptions of the four target genes selected for sequence typing.
690 Rightmost column includes the orthologous ORF in *Candida albicans*.
691
692

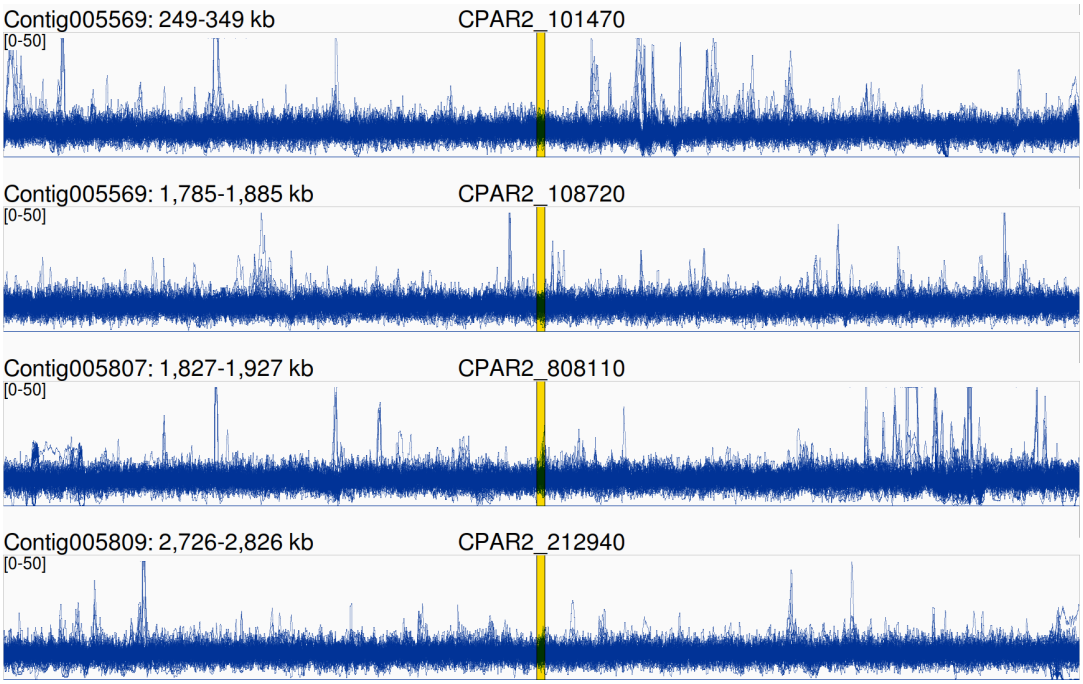
693 **Table 3. Oligo sequences for amplification and sequencing of *Candida parapsilosis* typing loci.**
694

Oligo ID	Seq?	Sequence	Gene position	Amplicon size
CPAR2_101470_F	X	GTGTGTCAACAACGGAGCAT	(-15) - 680	695 bp
CPAR2_101470_R		CATTTTTGGGGACTGGCACA		
CPAR2_108720_F	X	GGAAACCAATCTTGTCAGCACC	1891 - 2592	701 bp
CPAR2_108720_R		ACATCCATGCGGTATCTCCG		
CPAR2_212940_F		TAGTATACAAATGGTACACAATCGC	(-7) - 818	825 bp
CPAR2_212940_R	X	TCTGGATGTGGACGTGGTTT		
CPAR2_808110_F		AGTTGAGTCGTTATCTGAGGTTAG	265 - 1022	757 bp
CPAR2_808110_R	X	TTGCTGGTACTTGGGTGC		

695 Table of forward and reverse primers used for the amplification of the four loci selected during this study. The
696 oligo used for subsequent Sanger sequencing is indicated in the ‘Seq’ column. Gene positions are relative to
697 the first coding base in the reference genome of CDC317.
698



Supplemental Figure 1. Whole genome phylogeny of *Candida parapsilosis* isolates from Germany, their fluconazole susceptibility, and *ERG11* polymorphisms. Maximum-likelihood phylogeny constructed from SNV data from 38,567 variable positions. Tree is shown rooted at the midpoint. Tip colors indicate the city where the isolate originated from. Fluconazole resistant isolates where no polymorphisms in *ERG11* were detected are marked with an asterisk.



Supplemental Figure 2. Genomic coverage at the four sequence typing loci plus 50 bp of up- and downstream sequence. Overlaid line traces of normalized read depth for 250 samples included in the copy number variation screen (see Methods). The region used for sequence typing is indicated in the yellow box. Coverage depth was calculated as reads per million (RPM).

712 **Supplemental Table 1. Genetic variants observed in resistance-associated genes among**
713 **azole-resistant isolates lacking ERG11 Y132 substitution.**

	NRZ-2016-333	NRZ-2018-148	NRZ-2019-030	NRZ-2019-337	NRZ-2019-498	NRZ-2020-725	NRZ-2021-0166	NRZ-2021-0782
CPAR2_304370 (CDR1b)	P70fs	A1214V		R224*				
CPAR2_301760 (MDR1)			W104*					
CPAR2_603010 (MDR1b)		G247R		L236F				
CPAR2_207540 (MDR1)		F463I						T307K
CPAR2_303510 (TAC1)		Y792fs		L518F				
CPAR2_807270 (MRR1)					G583E		S629G	N725Y
CPAR2_207280 (UPC2)							N455D	
CPAR2_212210 (ALD5)								T18N
CPAR2_401350 (DAG7)							P77fs	
CPAR2_405010 (ERG6)								C300Y

714 Abbreviations in addition to standard one-letter amino acid codes: fs = frameshift; * = nonsense
715 mutation (premature stop codon).

716 **Supplemental Table 2. *C. parapsilosis* strains used for laboratory validation of sequence**
717 **typing markers as well as their genetic cluster and sequence type identity and GenBank**
718 **accession number for the Sanger data generated from each sample.**

***C. parapsilosis* strain ID**	**Sequence type**	**GenBank accession**			
		CPAR2_101470	**CPAR2_108720**	**CPAR2_212940**	**CPAR2_808110**
NRZ-2016-033	ST 16 (0-0-0-0)	OR413203	OR413216	OR413229	OR413242
NRZ-2018-148	ST 24 (0-0-1-4)	OR413204	OR413217	OR413230	OR413243
NRZ-2018-167	ST 28 (3-0-3-10)	OR413205	OR413218	OR413231	OR413244
NRZ-2019-030	ST 08 (1-0-2-7)	OR413206	OR413219	OR413232	OR413245
NRZ-2019-074	ST 01 (4-6-4-7)	OR413207	OR413220	OR413233	OR413246
NRZ-2019-077	ST 18 (0-0-0-1)	OR413208	OR413221	OR413234	OR413247
NRZ-2019-337	ST 22 (0-0-0-4)	OR413209	OR413222	OR413235	OR413248
NRZ-2019-498	ST 27 (3-4-3-8)	OR413210	OR413223	OR413236	OR413249
NRZ-2019-532	ST 18 (0-0-0-1)	OR413211	OR413224	OR413237	OR413250
NRZ-2019-747	ST 18 (0-0-0-1)	OR413212	OR413225	OR413238	OR413251
NRZ-2020-315	ST 23 (0-0-0-5)	OR413213	OR413226	OR413239	OR413252
NRZ-2020-540	ST 18 (0-0-0-1)	OR413214	OR413227	OR413240	OR413253
NRZ-2020-726	ST 01 (4-6-4-7)	OR413215	OR413228	OR413241	OR413254
NRZ-2021-0591	ST 18 (0-0-0-1)	OR987460	OR987469	OR987478	OR987451
NRZ-2022-0323	ST 18 (0-0-0-1)	OR987461	OR987470	OR987479	OR987452
NRZ-2022-0798	ST 18 (0-0-0-1)	OR987462	OR987471	OR987480	OR987453
NRZ-2023-0851	ST 18 (0-0-0-1)	OR987454	OR987463	OR987472	OR987445
NRZ-2023-0852	ST 18 (0-0-0-1)	OR987455	OR987464	OR987473	OR987446
NRZ-2023-0854	ST 18 (0-0-0-1)	OR987456	OR987465	OR987474	OR987447
NRZ-2023-0850	ST 18 (0-0-0-1)	OR987457	OR987466	OR987475	OR987448
NRZ-2023-0853	ST 18 (0-0-0-1)	OR987458	OR987467	OR987476	OR987449
NRZ-2023-0855	ST 18 (0-0-0-1)	OR987459	OR987468	OR987477	OR987450

719

720

721 **Supplemental File 1. Accession numbers for the publicly available genomes analyzed in this**
722 **study and strain information for the outbreak strains described in this study.**

723

724 **Supplemental File 2. FASTA alignments of the four typing loci containing representative**
725 **sequence for each of the 31 sequence types.**