Genomic reconstruction of an azole-resistant *Candida parapsilosis* outbreak and the creation of a multilocus sequence typing scheme: a retrospective observational and genomic epidemiology study

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

Phillip Brassington, MSc[1,6]; Frank-Rainer Klefisch, Dr. med[2]; Barbara Graf, Dr. med[3]; Roland Pfüller, Dr. rer. nat. [4]; Oliver Kurzai, Univ. Prof. [5,6]; Grit Walther, Dr. rer. nat. [5, 8]; Amelia E. Barber, PhD[1, 7, 8]

Address correspondence to:

Amelia E. Barber (amelia.barber@uni-jena.de; Rosalind-Franklin-Str. 1, 07745 Jena, Germany)

[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.
[7] Cluster of Excellence Balance of the Microverse, Friedrich Schiller University Jena, Jena, Germany.
[8] Equal contribution

Summary

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 Berlin, Germany.

Methods

In this retrospective observational study, we used whole-genome sequencing of isolates from the outbreak in Berlin and 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 global samples to identify loci with high discriminatory power to establish a multi-locus sequence typing (MLST) strategy for *C. parapsilosis*.

Findings

A clonal, azole-resistant strain of *C. parapsilosis* was observed causing 33 cases of invasive infection from 2018-2022 in multiple hospitals within the outbreak city. Whole genome sequencing revealed that outbreak strains were separated by an average of 36 single nucleotide variants, while outbreak strains differed from outgroup samples from Berlin and other regions of Germany by an average of 2,112 variants. Temporal and genomic reconstruction of the outbreak cases 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. German outbreak strains were closely related to strains responsible for an outbreak in Canada and to others isolated in the Middle East and East Asia. Including the 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, a global collection of 386 isolates was categorized into 62 sequence types, with the outbreak strains all belonging to the same sequence type.

Interpretation

This study underscores the emergence of drug resistant fungal pathogens that can spread patient-to-patient within a healthcare system, but also around the globe. This highlights the importance of monitoring *C. parapsilosis* epidemiology globally and of continuous surveillance and rigorous infection control measures at the local scale. Through large-scale genomic epidemiology, our study offers a high-resolution view of how a drug-resistant clone behaved in a local healthcare system and how this clone fits into the global epidemiology of this pathogen. We also demonstrate the utility of the novel typing scheme for genetic epidemiology and outbreak investigations as a faster and less expensive alternative to whole genome sequencing.

Funding

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

Research in context

Evidence before this study

We searched PubMed and Google Scholar from database inception to Apr 25, 2024, using the search terms “Candida parapsilosis”, “outbreak”, “azole resistance”, and/or “fluconazole” in PubMed and Google Scholar. We applied no language or study type restrictions. 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 overwhelming majority of previous work has used microsatellite markers to infer genetic relationships among outbreaks strains, obscuring whether they are really clonal in nature, our understanding of the temporal and transmission dynamics of these outbreaks, and the genetic relationship between outbreak clones.

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 Sanger 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. While whole genome sequencing (WGS) is becoming more widely available, it is still not available in many settings due to cost, limitations in bioinformatic expertise, and the absence of standardized methodology and data interpretation. The establishment of a sequence-based typing scheme is a valuable tool for rapid assessment of samples, which can aid in outbreak tracking and containment efforts, and provide results more rapidly even in settings where WGS is available.

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* [1–3]. *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 [1]. 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 [3]. 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) [4]. The most dominant fluconazole resistance mechanism seen clinically is the Y132F amino acid substitution in the azole target gene, lanosterol 14-alpha-demethylase (*ERG11*).

Alongside the rapid increase in fluconazole resistance, there has been an alarming number of drug-resistant outbreaks of *C. parapsilosis* reported since 2018 [5–13]. 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 [5,9]. Outbreaks of *C. parapsilosis* can also persist despite strict infection control strategies and may show gaps in the occurrence of outbreak strains [3,14]. 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 [15]. However, microsatellite typing requires specialized equipment, is prone to PCR artifacts, and requires specialized knowledge for interpretation [16]. 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* [17,18]. However, WGS is not feasible for many settings due to its cost, time to result, and limitations in trained personnel to analyze the data. The aim of this study was to investigate a prolonged outbreak of azole-resistant *C. parapsilosis* using WGS to understand its epidemiology and to introduce a sequence typing strategy for rapid and inexpensive outbreak investigations.

Methods

<u>Study design, outbreak description, and isolate collection</u>

In this retrospective, national study, the genomes of *C. parapsilosis* isolates sent to the NRZMyk were sequenced. Samples were submitted by German healthcare facilities and laboratories for

susceptibility testing between 2016 and 2022. All potential outbreak samples were included in the study based on the following criteria: samples originated from Berlin, samples were resistant to fluconazole (FLC^R) and voriconazole (VOR^R), and were susceptible to posaconazole. Non-outbreak isolates were selected for inclusion in the study from the NRZMyk strain collection as follows: all strains that originated from outside Berlin and displayed resistance to one or more azole were included. For the azole resistant samples from outside Berlin, a matched number of azole susceptible strains from the city were randomly selected for inclusion from the NRZMyk collection.

Outbreak cases were initially detected in December 2018 and January 2019 by physicians at Hospital #1 (name anonymized for privacy) (Figure 1a), who noticed an increase in isolates with a distinct antifungal susceptibility profile. The isolates were resistant to fluconazole (FLC^R) and voriconazole (VOR^R), but susceptible to other triazoles. After a period of approximately five months, during which no FLC^R and VOR^R infections were reported at Hospital #1, infection isolates with this susceptibility profile were once again observed (Figure 1a). The outbreak was officially declared at that time (October 2019) and was formally investigated including environmental surveillance. As *C. parapsilosis* is known to colonize the ear, 21 ICU patients at Hospital #1 were screened with swabs of both ears, but none were positive for *C. parapsilosis*. To investigate possible transfer by staff, sixteen bedside stethoscopes were screened. Only one sample was positive, but the resulting *C. parapsilosis* isolate was azole susceptible. To prevent further cases, emphasis was placed on hygiene measures and hand disinfection. Following implementation of the infection control measures there was a nine-month period without FLC^R and VOR^R isolates reported. However, additional cases were detected in September 2021 and the start-stop pattern continued through 2022 (Figure 1a). Consultations with other hospitals and clinical laboratories in Berlin identified *C. parapsilosis* infections with the same antifungal susceptibility profile at five other institutions (H0, H8, H9, H10, H14). As outbreaks of *C. parapsilosis* are being increasingly reported, there is an urgent need for a rapid and reproducible tool capable of differentiating isolates. Motivated by this, we also set out to identify a minimal set of loci that could be used for Sanger sequence-based typing for both outbreak detection and epidemiology in our study. The Ethics Committee of the University Hospital Jena approved this study under registration number 2024-3255-Daten.

Procedures

Antifungal susceptibility testing was performed by the NRZMyk using the methodology described in appendix 1 p 2. For tracking outbreak dynamics, admission and discharge records were obtained for all outbreak patients beginning January 2018, corresponding to one year before the first detected outbreak case, and extending to December 2022 and the network of patient transfer manually assembled. These dates correspond to roughly 12 months prior to the first case and extending until there had been several months without new cases in 2022. Patient information, including date of birth, sex, underlying disease, and antifungal treatment history, was obtained from the intake form of the NRZMyk.The NRZMyk does not collect information on patient ethnicity or race.

Whole genome sequencing and computational analysis.

DNA extraction was performed as described in the appendix 1 p 2. Library preparation and paired end sequencing was performed by GeneWiz (Leipzig, Germany) or Novogene (Cambridge, UK) using the Illumina MiSeq platform and NovaSeq platforms, respectively (Illumina, San Diego, CA, USA). One outbreak sample was sequenced a second time using new outgrowth and DNA extraction from the freezer stock to assess the degree of technical variation between sequencing runs. Illumina data of global samples was obtained from the NCBI Sequence Read Archive and was included in all analyses except for those exclusively examining outbreak samples or comparisons between outbreak and other German samples. Sequence data quality control, processing, and identification of variants are described in appendix 1 p 2. The SNV-based phylogeny and all SNV distance-based analyses only considered variant positions with non-zero coverage in all 386 samples. Additional information about phylogeny construction, how genetic distances were calculated, and root-to-tip regression for assessing clock-like evolution are detailed in appendix 1 pp 2-3.

Sequence typing

The process by which loci were screened and identified for sequence typing and how the number of loci was selected is described in appendix 1 p 3. PCR amplification of typing loci was performed using the primers and cycling parameters described in Table 2. Our Sanger sequence provider did not report heterozygous sites with IUPAC codes and performed background subtraction, which in rare cases erroneously removed a heterozygous peak. These were resolved by requesting the unprocessed chromatograms and manually reviewing them to ensure that IUPAC bases were reported correctly. The samples used for validation were collected from three hospitals in Berlin and were suspected to be additional outbreak cases based on their antifungal susceptibility profile (FLC^R, VOR^R, POS^S) and their isolation from hospitals with previous outbreak cases.

Hunter's discriminatory power was calculated as follows, where N represents the number of strains tested and x_j the number of strains belonging to the jth type:

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

Statistical analysis

Statistical analyses were performed using R version 4.3.1 and significance was defined as $p < 0.05$. Mann-Whitney U test was used to determine the pairwise genetic difference between outbreak and outgroup strains. Spearman correlation was used to assess the relationship between isolate date and root-to-tip distance. ANOVA followed by Tukey's HDS test was used to test for pairwise SNV differences between the outbreak strains and the closely related strains from other countries.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

In this study, we included 38 outbreak isolates from 33 cases identified through epidemiological and genomic investigations. All isolates were fluconazole resistant with a median MIC of 64 mg/L (IQR 16 to >64) and voriconazole (median MIC 1 mg/l; IQR 1.25), but susceptible to posaconazole (Table 1). We also included 38 non-outbreak isolates from the NRZMyk, originating from unlinked hospitals in Berlin and outside Berlin displaying resistance to one or more azole (Table 1). Most outbreak cases were first reported as wound infections (39%; 12/33), followed by infections associated with central venous catheters (29%; 9/33) (Table 1). Antifungal treatment histories were available for 30 cases (appendix 1 p 5). 33% (10/30) of patients were treated with an azole within 60 days of *C. parapsilosis* isolation, while a larger proportion (53%; 16/30) were treated with an echinocandin within the previous 60 days (Table 1). The defined daily dose per 100 bed days of azole antifungals at Hospital #1 was similar to that of other care level three hospitals in the region and remained largely stable over the five-year study period (appendix 1 p 5).

Reconstructing the outbreak transmission dynamics revealed that even though the largest fraction of cases (n = 20) were diagnosed at Hospital #1, most of these patients had been directly transferred to this facility from a major surgical facility (Hospital #0 in Figure 1b). Supporting nosocomial transmission, 94% of outbreak cases (31/33) had verified temporal overlap with other cases. Ward information was available for 28 cases at hospitals H1 and H14, revealing five patient clusters of 21 patients with spatial and temporal overlap on the same ward at the same time as other outbreak cases, strongly supporting patient-to-patient spread (Figure 1c). Altogether, 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.

To examine the genomic relationship between the included isolates, we performed whole genome sequencing (WGS). 68% (51/75) of the sequenced samples originated from Berlin, including 38 isolates from the 33 cases identified to be part of the outbreak and 13 non-linked isolates from different medical centers in Berlin. Publicly available data for an additional 311 samples was included to facilitate comparison with the global *C. parapsilosis* population, bringing the total number to 386 genomes. WGS samples analyzed had a mean depth of coverage of 123x±80 (SD) (appendix 2).

When built into a phylogeny, we did not observe a strong population structure among the 386 samples and no lineage was unique to a particular country or continent (Figure 2). The close genetic relationship between the outbreak isolates was evident by their tight grouping and short branch lengths, supporting the nosocomial transmission that was inferred from the outbreak reconstruction.

266 The quasi-clonal relationship between outbreak isolates was further confirmed by pairwise single
267 nucleotide variant (SNV) analysis, where isolates in the outbreak clade differed by an average of
268 36±20 SNVs (SD) (Figure 3). In contrast, the mean number of SNVs between outbreak samples and
269 non-outbreak isolates from other locations in Germany was significantly higher at 2,112± 828 SNVs
270 SD ($p < 0.0001$, Mann-Whitney U test). Further, the number of SNVs observed between duplicate
271 sequencing of the same freezer stock of an isolate was 15, indicating that the number of SNV
272 differences between repeat sequencing of the same sample was similar to the genetic differentiation
273 between outbreak isolates. We did not detect an obvious spatial-genetic relationship among outbreak
274 isolates but did observe some evidence of clock-like evolution (appendix 1 p 7, Spearman $R = 0.36$,
275 $p = 0.063$).
276
277 Strikingly, samples from Turkey[19], Canada[13], South Korea[20], and Kuwait[21] were closely related to the
278 Berlin outbreak strains (Figure 2). The Canadian isolates were reported as part of a drug resistant
279 outbreak, revealing that this lineage has caused outbreaks on multiple continents. Although closely
280 related, the strains from each geographic region could still be distinguished by pairwise SNV
281 differences and were more similar to each other than to the other regions (appendix 1 p 8).
282 Interestingly, despite their close genetic relationship and all strains being fluconazole resistant,
283 multiple resistance mechanisms were present. The strains from Berlin and Turkey carried an *ERG11*
284 Y132F mutation, while strains from Canada were *ERG11* K143R (another mutation associated with
285 fluconazole resistance) and those from South Korea were *ERG11* wild-type. Altogether, whole
286 genome sequencing confirmed the close genetic relationship of the outbreak strains and the presence
287 of this lineage in multiple other countries.
288
289 *ERG11* Y132F was the dominant azole resistant mechanism across all German isolates, with 82% of
290 resistant isolates possessing this mutation (n = 38/38 outbreak; n = 5/13 non-outbreak). Interestingly,
291 the non-outbreak Y132F strains were in two distantly related groups of samples based on the
292 phylogeny (Figure 2; appendix 1 p 9). Each of these Y132F phylogroups contained isolates from
293 geographically distant regions, suggesting that there are at least 3 distinct lineages of Y132F *C.*
294 *parapsilosis* circulating widely in Germany. We also observed three closely related but unlinked
295 samples from different cities in Germany that were pan-azole resistant and carried a G458S
296 substitution in *ERG11* (Figure 2; appendix 1 p 9) that has been linked to pan-azole resistance
297 previously [7,22,23]. Also included in our dataset were six newly-sequenced isolates that were resistant
298 to one or more azoles, but did not contain any resistance-associated polymorphisms in *ERG11*. These
299 isolates contained sequence polymorphisms in several genes that have been linked to drug
300 resistance, including the efflux pumps *MDR1* and *CDR1B*, the multidrug resistance gene regulator
301 *MRR1,* and the transcriptional activator *TAC1* (appendix 1 pp 12-13). While a causal role for most of
302 these polymorphisms in drug resistance have not been experimentally established, one resistant
303 isolate harbored a L518F substitution in *TAC1* experimentally proven to confer azole resistance *in*
304 *vitro* [24].

305

306 Our analysis to select loci suitable for Sanger sequence-based typing identified four 750 bp loci in
307 CPAR2_101400, CPAR2_101470, CPAR2_108720, and CPAR2_808110 (Table 2). These loci the 386
308 isolates into 62 unique sequence types (appendix 3) with a Hunter's Discriminatory Power of 0·93.
309 Discriminatory power assumes unrelated strains, while our dataset contains multiple outbreaks and
310 sets of serial isolates. When the global dataset was pruned as described in appendix 1 p3, the
311 discriminatory power rose to 0·98. Conversely, *C. parapsilosis* is a diploid organism and an
312 appreciable fraction of the discriminatory power comes from whether sequence polymorphisms were
313 heterozygous or homozygous. When site zygosity was not considered, we were only able to identify
314 43 sequence types across the 386 samples and the discriminatory power was dropped to 0·90. The
315 four selected loci showed a high degree of agreement with the whole genome phylogeny and were
316 able to correctly identify samples to their whole genome-defined genetic cluster with 97·1% sensitivity
317 (0.951-0.982 95% CI) and 99·8% specificity (0.997-0.99 95% CI) (appendix 1 p 6). To confirm that
318 the selected loci were present as a single copy and the discriminatory power was not being inflated
319 by the presence of multiple copies in the genome, we assessed them for genome-wide copy number
320 changes. We did not detect any CNV changes at these loci, either additional copies or loss of the
321 region (appendix 1 p 10).
322

323 We validated our *in silico* typing using the primers in Table 2 and a representative 13 samples from
324 our dataset, including four outbreak strains expected to have identical sequence type profiles. In each
325 sample, we detected all anticipated variants by Sanger sequencing and did not discover any variants
326 that were not in the whole genome data (appendix 1 p 14). As expected, the four outbreak samples
327 had identical sequences at the four loci and were assigned to the same sequence type. To facilitate
328 the typing of novel isolates by researchers and clinical labs, FASTA alignments containing
329 representative allelic sequences of the four markers are provided as appendices 4-7.
330

331 Finally, we demonstrated the utility of the MLST scheme on seven isolates that were not part of the
332 dataset used to create the typing scheme, but were suspected to be additional outbreak cases based
333 on submitting hospital and antifungal susceptibility profiles (appendix 1 p 4). Their sequence type was
334 identical to the outbreak isolates (appendix 1 p 14), suggesting that they were outbreak cases 26-32
335 (Figure 1a), which was confirmed by WGS. Taken together, we identified and validated four loci that
336 can be used for sequence-based typing of *C. parapsilosis* and demonstrated their value in rapidly
337 identifying closely related samples.
338

339 **Discussion**

340 In this work, we use genomic epidemiology coupled with patient tracking to investigate an outbreak
341 of *ERG11* Y132F azole-resistant *C. parapsilosis* that occurred over several years and medical
342 centers in Berlin, Germany. We use a large-scale genomic dataset composed of 386 *C. parapsilosis*
343 isolates to investigate outbreak dynamics, confirm the high degree of relatedness among outbreak

344 isolates, and compare them to isolates from around the world—revealing that the outbreak clone is
345 closely related to samples from several other continents. Based on the increasing number of
346 outbreaks for *C. parapsilosis* and the need to rapidly discriminate between isolates, we also
347 establish and validate a set of four loci for Sanger sequence-based strain typing.
348
349 Due to slow outbreak kinetics and the low case density, the outbreak might have gone unnoticed if
350 not for the atypical susceptibility pattern that was detected by clinicians and clinical microbiologists.
351 The reconstructed transmission dynamics suggest that outbreaks of *C. parapsilosis* may result from
352 the reintroduction of outbreak strains from other hospitals, underlining the challenges of outbreak
353 control and the need for heightened infection control measures. In our case, patient transfers were a
354 common occurrence among cases and facilitated inter-hospital transmission and subsequent intra-
355 hospital transmission. Unfortunately, high-resolution temporal and spatial monitoring of patients is
356 lacking for other outbreaks, so it is unclear the degree to which this has driven other reported
357 *C. parapsilosis* outbreaks. However, the pattern of prolonged but tightly related isolates causing
358 disease across several years and multiple medical centers matches other outbreak reports [10,23,25,26],
359 suggesting that reintroductions and/or patient transfers could have also contributed to the prolonged
360 nature of these outbreaks. Though our study period concluded in 2022, given the prolonged and
361 sporadic nature, we believe it is unlikely that the outbreak has stopped for good.
362
363 One limitation to our study is that while patients without infection were screened for colonization
364 during the first wave of cases at Hospital #1, comprehensive screening of all the involved hospitals
365 and throughout the entire study period was not performed. This obscures an understanding of the
366 relationship between colonization and subsequent infection during the outbreak and a mechanistic
367 understanding of how *C. parapsilosis* is transmitted within healthcare facilities.
368
369 The finding of azole-resistant strains from different continents that were very closely related to the
370 strains spreading within healthcare facilities in Berlin highlights the global challenge of antifungal
371 resistant pathogens and the emergence of *Candida* spp that are easily transmitted patient-to-
372 patient. It is likely that drug resistance was acquired independently and recently in this lineage,
373 given the close genomic relationship but different resistance mechanisms in strains from Germany,
374 Canada, and South Korea. How this lineage acquired drug resistance and how it has spread across
375 continents are open questions, but we hypothesize that it may have enhanced properties that
376 facilitate its transmission and persistence. The observation of highly related clones observed on
377 multiple continents mirrors the behavior of *Candida auris*. Nearly identical clades of this pathogen
378 have been reported on multiple continents and continue to spread globally [27,28], underscoring the
379 fact that *C. parapsilosis* has the potential to become a problem of similar scale. Additionally, three
380 phylogenetically distinct Y132F lineages of fluconazole-resistant *C. parapsilosis* were discovered in
381 Germany during the study period, each with multiple isolates reported from distinct geographic
382 areas. We also identified six isolates that are resistant by an unknown mechanism, which will be an

interesting area to follow up on using genome-wide association approaches. Taken together these observations demonstrates that fluconazole-resistant *C. parapsilosis* is endemic in Germany, mirroring what has been recently been reported in other countries [3], and highlighting 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 [29] and 0·99 [15]. The four loci MLST established in this study achieved a discriminatory power of 0·93, within the range of what is achieved using microsatellites. Importantly, 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. Taken together, our study provides a high-resolution description of an azole-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.

Data sharing

Raw FASTQ files for novel isolates sequenced in this study were uploaded to the NCBI Sequence Read Archive and are publicly available under BioProject PRJNA996760. Accession numbers for the publicly available sequence data used in this study are listed in appendix 2. Genbank records for the Sanger sequencing of the MLST loci performed in this study are listed in appendix 1 p 14. FASTAs containing representative allele sequences for each of the four typing loci are available as appendices 4-7. Most data analysed in this study are included within the study or in the appendices. However, remaining datasets can be made available from the corresponding author on reasonable request.

Acknowledgements

We are grateful to Peggy Karanatsiou of the Friedrich Schiller University and Philipp Hupel, Carmen Karkowski, and Christiane Weigel of the NRZMyk for technical assistance. We thank Britta Schweickert of the Robert-Koch-Institute for contributing regional hospital data from the antibiotic consumption surveillance system. We also acknowledge MSD and Pfizer for providing the drugs for antifungal susceptibility testing to the NRZMyk free of charge.

Funding

This project was supported by The Federal Ministry for Education and Science (Bundesministerium für Bildung und Forschung) within the framework of InfectControl 2020 (Projects FINAR 2.0, grant 03ZZ0834). The work of the German National Reference Center for Invasive Fungal Infections is supported by the Robert Koch Institute from funds provided by the German Ministry of Health (grant 1369-240). AEB is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research 358 Foundation) under Germany's Excellence Strategy – EXC 20151 – Project-ID 390813860.

Contributors

AEB and GW conceptualized the study. PB, FRK, BG, RP, OK, GW and AEB contributed to investigation, methodology, and formal analysis. PB, GW, and AEB were responsible for visualization. AEB did the statistical analysis. AEB, PB, and GW wrote the original draft. All authors participated in reviewing and editing the manuscript. All authors had full access to all the data in the study and accept responsibility for the decision to submit for publication. AEB and GW verified the underlying data of the study.

Declaration of Interests

We declare no competing interests.

References

1. Tóth, R. *et al.* Candida parapsilosis: From genes to the bedside. *Clin. Microbiol. Rev.* **32**, 1–38 (2019).
2. Rhodes, J. & Fisher, M. C. Global epidemiology of emerging Candida auris. *Curr. Opin. Microbiol.* **52**, 84–89 (2019).
3. Daneshnia, F. *et al.* Worldwide emergence of fluconazole-resistant Candida parapsilosis: current framework and future research roadmap. *Lancet Microbe* **4**, e470–e480 (2023).
4. Arastehfar, A. *et al.* Clonal Candidemia Outbreak by Candida parapsilosis Carrying Y132F in Turkey: Evolution of a Persisting Challenge. *Front. Cell. Infect. Microbiol.* **11**, (2021).
5. Qi, L. *et al.* Nosocomial outbreak of Candida parapsilosis sensu stricto fungaemia in a neonatal intensive care unit in China. *J. Hosp. Infect.* **100**, e246–e252 (2018).
6. Thomaz, D. Y. *et al.* An Azole-Resistant Candida parapsilosis Outbreak: Clonal Persistence in the Intensive Care Unit of a Brazilian Teaching Hospital. *Front. Microbiol.* **9**, (2018).
7. Arastehfar, A. *et al.* First report of candidemia clonal outbreak caused by emerging fluconazole-resistant Candida parapsilosis isolates harboring Y132F and/or Y132F+K143R in Turkey. *Antimicrob. Agents Chemother.* (2020) doi:10.1128/AAC.01001-20.
8. Corzo-Leon, D. E., Peacock, M., Rodriguez-Zulueta, P., Salazar-Tamayo, G. J. & MacCallum, D. M. General hospital outbreak of invasive candidiasis due to azole-resistant Candida parapsilosis associated with an Erg11 Y132F mutation. *Med. Mycol.* 1–8 (2020) doi:10.1093/mmy/myaa098.
9. Danilo Y. Thomaz, João N. de Almeida Jr. *, O. N. E. S., Gilda M.B. Del Negro, Gabrielle O. M. H. Carvalho, V. M. F. G., Maria Emilia B. de Souza, Amir Arastehfar, Carlos H. Camargo, A. L., Motta, Flávia Rossi, David S. Perlin, Maristela P. Freire, Edson Abdala, G., & Benard. Environmental clonal spread of azole-resistant Candida parapsilosis with Erg11-Y132F mutation causing a large candidemia outbreak in a Brazilian Cancer Referral Center. *J. Fungi* (2021).
10. Fekkar, A. *et al.* Hospital Outbreak of Fluconazole-Resistant Candida parapsilosis: Arguments for Clonal Transmission and Long-Term Persistence. *Antimicrob. Agents Chemother.* **65**, e02036-20 (2021).
11. Guinea, J. *et al.* Whole genome sequencing confirms Candida albicans and Candida parapsilosis microsatellite sporadic and persistent clones causing outbreaks of candidemia in neonates. *Med. Mycol.* **60**, myab068 (2022).
12. Trevijano-Contador, N. *et al.* Global emergence of resistance to fluconazole and voriconazole in Candida parapsilosis in tertiary hospitals in Spain during the COVID-19 pandemic. *Open Forum Infect. Dis.* ofac605 (2022) doi:10.1093/ofid/ofac605.

13. McTaggart, L. R. *et al.* First Canadian report of transmission of fluconazole-resistant Candida parapsilosis within two hospital networks confirmed by genomic analysis. *J. Clin. Microbiol.* **0**, e01161-23 (2023).
14. Presente, S. *et al.* Hospital Clonal Outbreak of Fluconazole-Resistant Candida parapsilosis Harboring the Y132F ERG11p Substitution in a French Intensive Care Unit. *Antimicrob. Agents Chemother.* **67**, e01130-22 (2023).
15. Sabino, R. *et al.* New polymorphic microsatellite markers able to distinguish among Candida parapsilosis sensu stricto isolates. *J. Clin. Microbiol.* **48**, 1677–1682 (2010).
16. Garcia-Hermoso, D., Desnos-Ollivier, M. & Bretagne, S. Typing Candida Species Using Microsatellite Length Polymorphism and Multilocus Sequence Typing. in *Candida Species: Methods and Protocols* (eds. Calderone, R. & Cihlar, R.) 199–214 (Springer, New York, NY, 2016). doi:10.1007/978-1-4939-3052-4_15.
17. Hamprecht, A. *et al.* Candida auris in Germany and Previous Exposure to Foreign Healthcare. *Emerg. Infect. Dis.* **25**, 1763–1765 (2019).
18. Lockhart, S. R. *et al.* Simultaneous emergence of multidrug-resistant candida auris on 3 continents confirmed by whole-genome sequencing and epidemiological analyses. *Clin. Infect. Dis.* **64**, 134–140 (2017).
19. Arastehfar, A. *et al.* Candidemia Among Coronavirus Disease 2019 Patients in Turkey Admitted to Intensive Care Units: A Retrospective Multicenter Study. *Open Forum Infect. Dis.* **9**, ofac078 (2022).
20. Kim, T. Y. *et al.* Evolution of Fluconazole Resistance Mechanisms and Clonal Types of Candida parapsilosis Isolates from a Tertiary Care Hospital in South Korea. *Antimicrob. Agents Chemother.* **66**, e00889-22 (2022).
21. Asadzadeh, M., Dashti, M., Ahmad, S., Alfouzan, W. & Alameer, A. Whole-Genome and Targeted-Amplicon Sequencing of Fluconazole-Susceptible and -Resistant Candida parapsilosis Isolates from Kuwait Reveals a Previously Undescribed N1132D Polymorphism in CDR1. *Antimicrob. Agents Chemother.* **65**, 10.1128/aac.01633-20 (2021).
22. Demirci-Duarte, S., Arikan-Akdagli, S. & Gülmez, D. Species distribution, azole resistance and related molecular mechanisms in invasive Candida parapsilosis complex isolates: Increase in fluconazole resistance in 21 years. *Mycoses* **64**, 823–830 (2021).
23. Díaz-García, J. *et al.* Evidence of Fluconazole-Resistant Candida parapsilosis Genotypes Spreading across Hospitals Located in Madrid, Spain and Harboring the Y132F ERG11p Substitution. *Antimicrob. Agents Chemother.* **66**, e00710-22 (2022).
24. Daneshnia, F. *et al.* Determinants of fluconazole resistance and echinocandin tolerance in C. parapsilosis isolates causing a large clonal candidemia outbreak among COVID-19 patients in a Brazilian ICU. *Emerg. Microbes Infect.* **11**, 2264–2274 (2022).
25. Thomaz, D. Y. *et al.* A Brazilian Inter-Hospital Candidemia Outbreak Caused by Fluconazole-Resistant Candida parapsilosis in the COVID-19 Era. *J. Fungi* **8**, 100 (2022).
26. Wang, H. *et al.* Investigation of an unrecognized large-scale outbreak of Candida parapsilosis sensu stricto fungaemia in a tertiary-care hospital in China. *Sci. Rep.* **6**, 27099 (2016).
27. Chow, N. A. *et al.* Multiple introductions and subsequent transmission of multidrug-resistant Candida auris in the USA: a molecular epidemiological survey. *Lancet Infect. Dis.* **3099**, 1–8 (2018).
28. Chow, N. A. *et al.* Tracing the evolutionary history and global expansion of Candida auris using population genomic analyses. *mBio* **11**, 2020.01.06.896548 (2020).
29. Romeo, O. *et al.* Microsatellite-based genotyping of Candida parapsilosis sensu stricto isolates reveals dominance and persistence of a particular epidemiological clone among neonatal intensive care unit patients. *Infect. Genet. Evol.* **13**, 105–108 (2013).

Figures

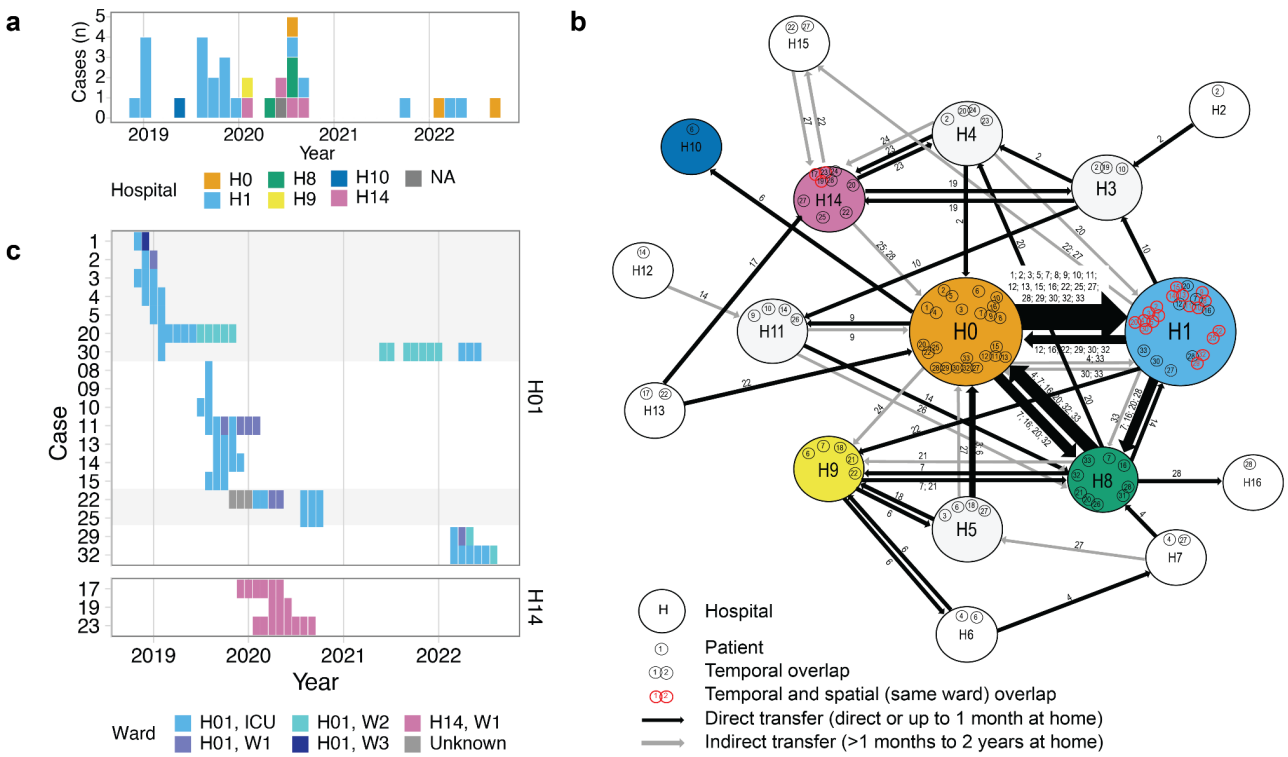

Figure 1. Timeline and hospital network dynamics of a *Candida parapsilosis* outbreak in Berlin, Germany. (a) Timeline histogram of the 33 outbreak cases. Dates are taken from the date of first positive culture. Color indicates the inpatient hospital at the time of isolation. A single case was identified in an ambulatory patient, which is indicated with “NA” for the inpatient hospital. (b) Inpatient stays of the outbreak cases represented as a network, where nodes (large circles) represent hospitals in the Berlin area and edges (lines) represent direct or indirect patient transfers between two institutions. Edge thickness depicts the relative number of transfers between two nodes. Individual patient case numbers are indicated within the node and along the edge. For the network, admission and discharge data beginning 12 months prior to the first detected outbreak and extending until the end of 2022 was utilized. (c) Timeline of 21 cases with temporal and spatial overlap. Alternating gray and white fills indicate five patient clusters where temporal and spatial (same ward) overlaps with other outbreak cases could be verified from inpatient records.

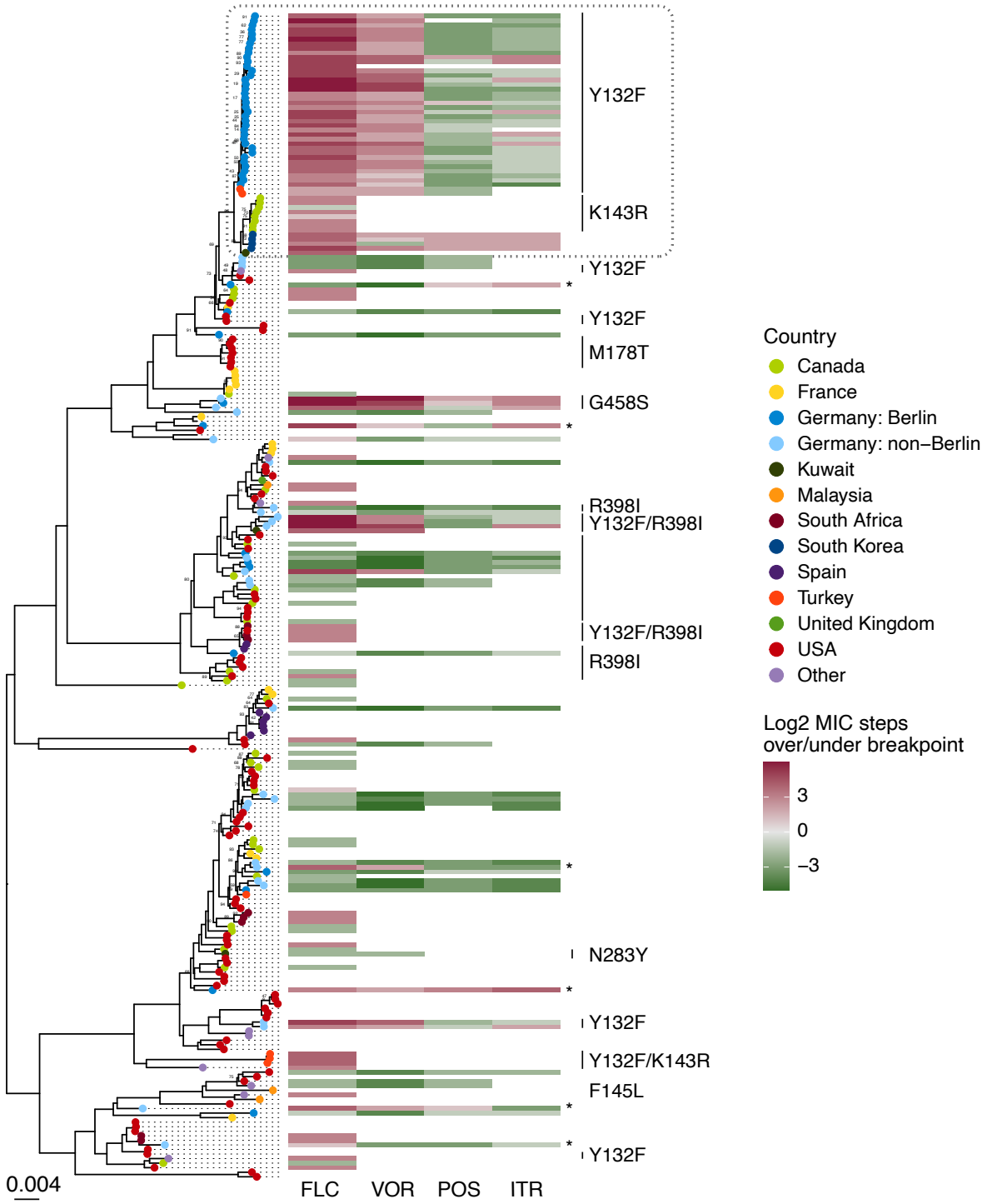

Figure 2. Whole genome phylogeny of *Candida parapsilosis* isolates from a global distribution and their azole susceptibility profiles and *ERG11* polymorphisms. Maximum-likelihood phylogeny constructed from SNV data from 51,494 variable positions. Tree is shown rooted at the midpoint. To increase tree readability, samples from the same sequencing study that showed little genetic differentiation were randomly downsampled to a maximum of 6 samples per sequencing project - with the exception of 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. The gray dashed box at the top of the phylogeny indicates the Berlin outbreak samples (medium blue) and the closely related samples from Turkey (red-orange), Canada (light green), South Korea (dark blue), and Kuwait (dark green) found in public data. Samples that were sequenced in our study and were resistant to one or more azole, but where no polymorphisms in *ERG11* were detected, are marked with an *.

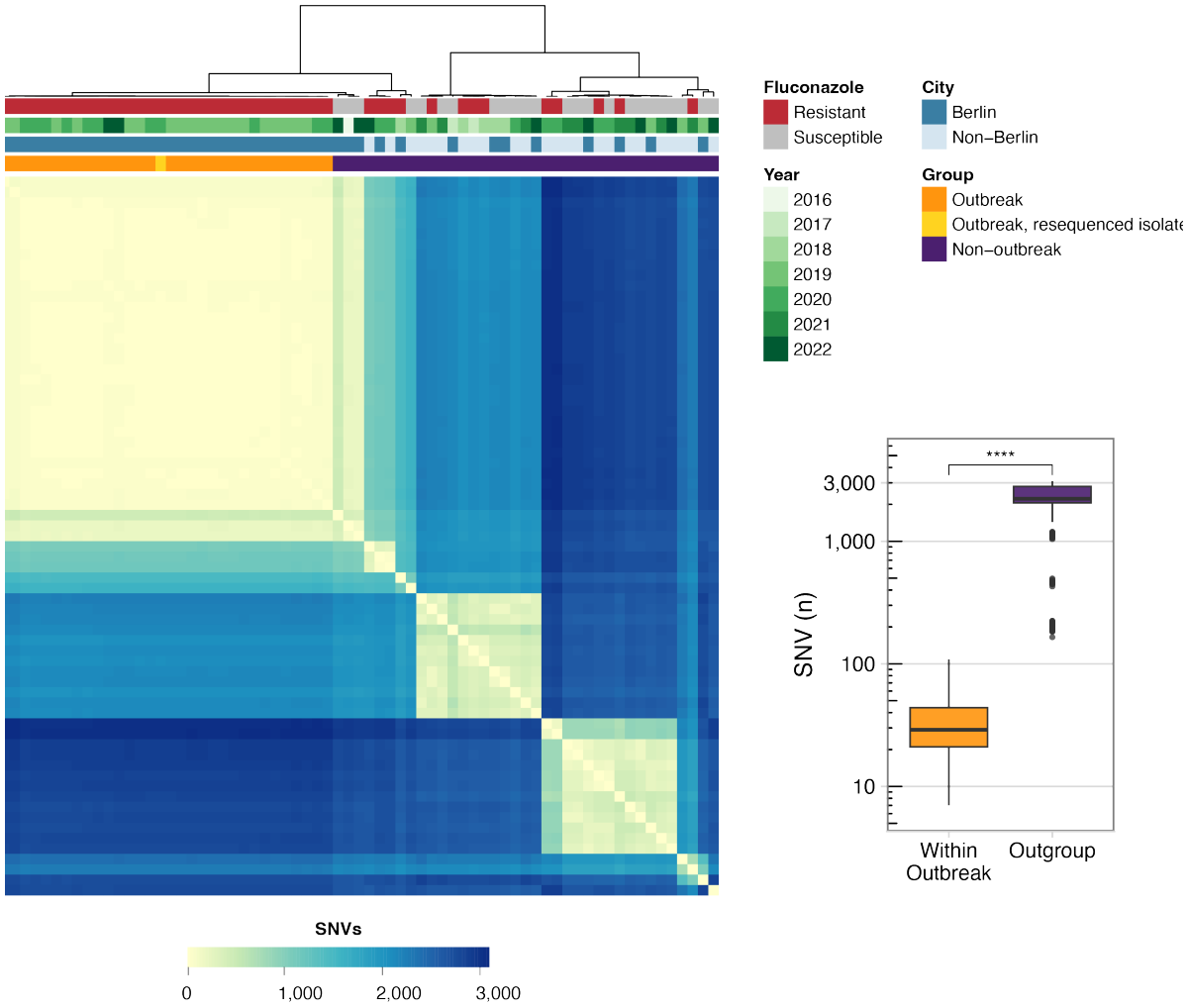

Figure 3. SNV matrix of the German isolates sequenced in this study.
Heatmap showing the pairwise SNV differences between the novel German isolates described in this study. Metadata bands indicate (from top to bottom): fluconazole susceptibility, year of sample isolation, and whether the sample is from the city of Berlin (including both outbreak and non-outbreak strains) or another city in Germany. Boxplot shows the number of pairwise SNV differences when the 33 outbreak strains are compared against each other compared to their pairwise differences with the other German isolates sequenced in this study (outgroup). Significance determined by Mann-Whitney U test, **** indicates $p < 0.0001$.