Rapid genome sequencing for outbreak analysis of the emerging human fungal pathogen *Candida auris*

Johanna Rhodes[1], Alireza Abdolrasouli[2,3], Rhys A. Farrer[1], Christina A. Cuomo[4], David M. Aanensen[1,5], Darius Armstrong-James[2], Matthew C. Fisher[1], Silke Schelenz[2]

1 – Department of Infectious Disease Epidemiology, Imperial College London, London, United Kingdom
2 – National Heart and Lung Institute, Imperial College London, London, United Kingdom
3 – Department of Medical Microbiology, Charing Cross Hospital, Imperial College Healthcare NHS Trust, London, United Kingdom
4 – Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA
5 – The Centre for Genomic Pathogen Surveillance, Wellcome Trust Genome Campus, Cambridgeshire, UK

Corresponding author(s): S.Schelenz@rbht.nhs.uk and Johanna.Rhodes@imperial.ac.uk

Abstract

Background: *Candida auris* was first described in 2009, and has since caused nosocomial outbreaks, invasive infections and fungaemia across 11 countries in five continents. An outbreak of *C. auris* occurred in a specialised cardiothoracic London hospital between April 2015 and November 2016, which to date has been the largest outbreak reported worldwide, involving a total of 72 patients.
Methods: To understand the epidemiology of *C. auris* infection within this hospital, we sequenced the genomes of outbreak isolates using Oxford Nanopore Technologies and Illumina in order to type antifungal resistance alleles and to explore the outbreak within its local and global context.
Findings: Phylogenomic analysis placed the UK outbreak in the India/Pakistan clade, demonstrating an Asian origin. The outbreak showed similar diversity to that of the entire clade and limited local spatiotemporal clustering was observed. One isolate displayed resistance to both echinocandins and 5-flucytosine; the former was associated with a serine to tyrosine amino acid substitution in the gene *FKS1*, and the latter was associated with a phenylalanine to isoleucine substitution in the gene *FUR1*. These mutations are novel for this pathogen.
Interpretation: Multiple differential episodic selection of antifungal resistant genotypes has occurred within a genetically heterogenous population across this outbreak, creating a resilient pathogen and making it difficult to define local-scale patterns of transmission as well as implementing outbreak control measures.
Funding: Antimicrobial Research Collaborative, Imperial College London

Introduction

The emerging fungal pathogen *Candida auris* causes nosocomial invasive infections, predominantly in intensive care units (ICU). Since its first description in 2009 in Japan (1), reports of *C. auris* infections have been reported in several countries (2-10). *C. auris* demonstrates intrinsic multidrug resistant (MDR) phenotype (9), by exhibiting high level resistance to fluconazole and varying susceptibility to other azole drugs, amphotericin B and a newly introduced class of antifungals, echinocandins (9).

In 2016, we described the first large-scale *C. auris* outbreak (April 2015 to November 2016) occurring within a single specialist cardiothoracic hospital in London (6). Due to the high uncertainty as to the time and source of introduction of *C. auris* into the hospital, the rapid development of a molecular epidemiological toolkit was required. Outbreaks of other fungal pathogens have been previously investigated using short-read whole-genome sequencing (WGS), which provided sufficient information to discriminate between isolates and their phylogenetic relationships using single nucleotide polymorphism (SNP) analysis (11-13). Recently, the handheld, portable MinION sequencer, manufactured by Oxford Nanopore Technologies, UK (ONT) has made rapid WGS widely available in the field, and has been successfully used to analyse the molecular epidemiology of recent *Salmonella*, and Ebola and Zika viruses outbreaks (14-16).

Here, we describe the first use of the MinION nanopore sequencing technology to determine the genetic epidemiology of this fungal outbreak, both within the UK hospital and also within a global context, alongside reannotating the genome of *C. auris* and defining novel antifungal resistance alleles.

Methods

Fungal isolates

Twenty-eight *C. auris* isolates were studied, consisting of 26 clinical isolates from 22 patients and two isolates collected from the room of a patient known to be colonised with *C. auris* in the ICU (Table 1). The identification of *C. auris* was conducted by Matrix-assisted laser desorption/ionization Time of Flight (MALDI-TOF) mass spectrometry (Brucker Daltonics, Fremont, CA, USA) using a formic acid – acetonitrile extraction procedure. Scores were interpreted as >2.00 for species-level identification.

Library preparation and sequencing

Five clinical isolates (Table 1) representing a 146-day time frame were chosen for sequencing using the hand-held MinION sequencer (Oxford Nanopore Technologies, Oxford, UK).

Twenty-four isolates (Table 1), including the five clinical isolates sequenced using MinION, were chosen for Illumina sequencing using Nextera library preparation method as part of the MicrobesNG service (University of Birmingham, UK). These isolates represented a 155-day time frame, and included isolates from patients and the environment around infected patients. Isolates of *C. auris* for both MinION and Illumina sequencing were cultured as described in Supplementary Methods. All raw reads in this study have been submitted to the European Nucleotide Archive under the project accession PRJEB20230. Details of genome assembly and annotation are described in Supplementary Methods.

Alignment of Illumina reads and phylogenetic analysis

Raw Illumina reads were quality checked using FastQC (v0.11.3; Babraham Institute), and trimmed using Trimmomatic (v0.30) based on a Phred quality score of 15. Alignment of reads to the 16B25 reference genome and variant calling were carried out as described in Rhodes and colleagues (17).

Phylogenies for whole genome SNP data were constructed and visualised as described in Rhodes and colleagues (17). The rate of evolution (represented as the number of substitutions per day) along the tree topology was estimated using TempEst v1.5 (18), calibrated with sampling times. Root-to-tip regression was calculated and the root of the tree was selected to maximise R^2.

Mutation identification in *ERG11*, *FKS1* and *FUR1* genes

Orthologous sequences to *C. albicans ERG11* (SC5314) were extracted from each *C. auris* genome. Sequences were evaluated for amino acid substitutions to mutations within hot spot regions in *C. albicans* (19) as described in Lockhart and colleagues (9). Predicted *FKS1* and *FUR1* genes from the genome annotation were used to identify presence or absence of mutations in *C. auris* isolates.

Results

We sequenced 25 clinical *C. auris* isolates from a recently described outbreak (6), along with two environmental samples to better represent the overall genetic diversity within the impacted hospital. We also sequenced eight isolates derived from four patients taken days apart to establish possible within-patient diversity.

Rapid generation of outbreak-specific *C. auris* reference genome and Illumina sequencing

We assembled five high quality hybrid *de novo* reference genomes for *C. auris* using Illumina short-read sequences and MinION long-read sequences rapidly generated over 48 hours. Five isolates (15B5, 16B21, 16B25, 16B20 and 16B15a) were chosen to cover a range of dates (October 2015 to March 2016). Isolate 16B25 had the best overall assembly quality of 110 contigs, N_{50} = 396,317 bp and an estimated genome size of 12·3 Mb (Table 3 and Supplementary Table S1). 98·94% of the 16B25 assembly mapped to the *C. auris* genome B8441 assembled by Lockhart and colleagues (9).

We generated an average of 5·2 million Illumina reads passing quality control for 27 isolates recovered during the outbreak that mapped closely (average 95·5%) to our reference genome (Supplemental Material Table S2). The rapid availability of long reads from MinION sequencing demonstrates this technology is ideally used in an outbreak setting for providing high-quality contiguous assemblies. A total of 5,366 protein-coding genes, 4 rRNAs and 156 tRNAs were predicted using the genome annotation pipeline described in Supplementary Methods. Table 3 summarises the general features of 16B25, along with other pathogenic *Candida* genomes. The number of protein coding genes presented here is in line with the predicted number of genes in *C. lusitaniae* (20) (*n* = 5,941), the closest known relative of *C. auris*.

There are fewer protein coding genes, tRNAs and rRNAs predicted in this genome than previously reported for *C. auris* Ci 6684 in Chatterjee and colleagues (21), as shown in Table 3. Running our annotation pipeline on the B8441 isolate presented in Lockhart and colleagues (9) found similar numbers of protein coding genes, rRNAs and tRNAs (Table 3). Therefore, the different total numbers between 16B25 and Ci 6684 are likely due to the different annotation pipelines and not the quality of the reference assemblies. The number of protein coding genes identified in Chatterjee and colleagues is likely inflated due to over-prediction of short sequences, lack of filtering of repetitive sequences, and using only GenemarkS to predict the start of genes; our pipeline used additional criteria to achieve a predicted set of high-confidence genes.

Phylogenetic analysis reveals an Indian/Pakistani origin of *C. auris* outbreak

Phylogenetic analysis based on whole genome SNPs revealed the UK outbreak had an Indian/Pakistani origin (Figure 1). SNP calls for isolates from Venezuela, India, Pakistan, Japan and South Africa (9) were also included to add geographic context to the outbreak. The UK outbreak isolates were in the same clade as those from India and Pakistan (Figure 1a); on average, 240 SNPs separated UK outbreak isolates from isolates collected in India and Pakistan. There were no known patient travel links to India or Pakistan prior to admission into hospital, however. We found an average of 84 SNPs separating isolates within the UK outbreak; later isolates exhibited only 55 SNPs between them (October 2016), compared to earlier isolates (January 2016) that showed an average 130 SNPs separating them.

Fitting root-to-tip regression showed there was a linear relationship between sampling time (days) and the expected number of nucleotide substitutions along the tree, demonstrating clock-like evolution across the time-scale of the outbreak (Figure 2). The evolutionary rate of nuclear DNA (calculated from the slope of the regression) equated to 1.5204^{e-3} substitutions per site, comparable with nuclear DNA of other fungal species such as *Schizosaccharomyces pombe* beer strains (3.0^{e-3} (22)) and *Saccharyomyces cerevisiae* (5.7^{e-3} (23)). The time to the most recent common ancestor (TMRCA) was estimated to be early March 2015, one month prior to the first patient identified with a *C. auris* infection. One isolate, 16B22b, was identified as being less diverged than average for the sampling date; five isolates (16B15b, 16B26, 16B25, 16B22a and 16B24a) were identified as being more diverged than average for the sampling date given.

Clinical isolates of *C. auris* show multi-drug resistance

Overall, 14 isolates displayed MDR to two or more classes of antifungal drugs. Only five isolates displayed resistance to one drug, fluconazole. All isolates expressed elevated levels of resistance to fluconazole (MIC: >256 ug/ml), with varying levels of resistance to itraconazole (MIC: 0.03 ug/ml - >16 ug/ml), voriconazole

(MIC: 0.12 ug/ml - 8 ug/ml) and posaconazole (which has not been previously reported in *C. auris*). Four isolates also displayed resistance to amphotericin B (MIC: >2 ug/ml).

One isolate (16B15b) displayed elevated levels of resistance to all echinocandins (MIC: >8 ug/ml), but remained susceptible to all azole drugs, with the exception to fluconazole. 16B15b also displayed high levels of resistance to flucytosine (MIC >64 ug/ml), which was not seen in the other isolates; therefore, both mutations and associated resistance were unique to this outbreak isolate. This isolate belonged to a patient who received anidulafungin for 7 days for pancolitis, developing *C. auris* candidaemia 11 days afterwards, at which point isolate 16B15a was recovered. Treatment was switched to amphotericin and 5-flucytosine for 2 weeks. Six days after completing this treatment, a pan-resistant *C. auris* (16B15b) was recovered from the vascular tip. One non-synonymous SNP (nsSNP), causing a serine to tyrosine substitution (S652Y) was identified in the *C. auris FKS1* gene; a similar mutation (S645Y) in *FKS1* has been associated with echinocandin resistance in *C. albicans* (24). Another nsSNP caused a phenylalanine to isoleucine substitution (F211I) in *FUR1*, which has a role in 5-flucytosine resistance (25). Neither of these mutations have been reported previously.

Orthologous sequences to *C. albicans ERG11* were screened for substitutions that conferred known fluconazole resistance mutations (19). The outbreak isolates all had the Y132F substitution in *ERG11*, confirming an Indian/Pakistani origin. Lockhart and colleagues also found that these substitutions were strongly correlated with geographic clades (9).

Interpretation of typing results in relation to epidemiology of the outbreak

C. auris outbreak isolates grouped into two phylogenetic clusters (A and B; 26% and 74% respectively) (Figure 1b). Cluster A comprised seven isolates from 2016 that had on average 245 SNPs that distinguished it from Cluster B. On average 130 SNPs separated isolates within Cluster B. Cluster A was introduced into the hospital in early 2016 and formed the dominant outbreak strains towards the end of the outbreak, with only 55 SNPs separating those isolates. The phenotypic antifungal resistance varied among these isolates: all expressed high level fluconazole resistance (MIC: 128-256 ug/ml) and susceptibility to echinocandins (micafungin, caspofungin and anidulafungin MIC: 0.06-0.12 ug/ml) and 5-flucytosine (MIC 0.06 ug/ml), two clinical isolates expressed cross azole resistance (MIC: 4-8 ug/ml posaconazole, 8 ug/ml voriconazole, 2-16 ug/ml itraconazole). Phenotypic antifungal susceptibility profiling may therefore not be used reliably for addressing genetically indistinguishable strains during nosocomial transmission analysis.

Three patients with isolates in Cluster A were admitted to the ICU. Two patients acquired *C. auris* whilst staying in ICU, and another patient acquired *C. auris* on a surgical admission unit geographically placed next door to the high dependency unit in the same month, but not overlapping with the ICU patients. Transmission of *C. auris* between different units within this hospital likely occurred *via* the movement of *C. auris*-positive patients or contaminated equipment. However, because we did not sequence isolates from all patients during the outbreak alongside the heterogeneous nature of the founding population, we are currently unable to establish routes of transmission in more detail.

Analysis and interpretation of sequencing result in relation to epidemiology of individual patient transmission during the outbreak

Eight isolates within this study were sequential pairs of isolates from four separate patients (Table 1); we hypothesised that there may have been nosocomial horizontal transmission between patients and/or their surrounding environment, as suggested in previous studies (2,3,5,26), for the following pairs of isolates: 16B22a and 16B22b from patient A (isolated 12 days apart); 16I27a (MICs were not carried out for this isolate) and 16B27b from patient B (isolated 1 day later); 16B24a and 16B24b from patient C (isolated 5 days apart); and 16B15a and 16B15b from patient D (isolated 32 days apart).

In patient A, 16B22a (recovered from the axilla) showed resistance to fluconazole only (Table 2). The subsequent isolate from this patient, 16B22b (isolated from a central line tip) exhibited resistance to all azole drugs (Table 2). A large number of SNPs (277 SNPs) separated the two isolates (Figure 1), which also separated into different phylogenetic clusters (16B22a in Cluster B and 16B22b in Cluster A) suggesting independent acquisition of infection by this patient within the unit. The two isolates from patient B similarly differed by a large number of SNPs (164 SNPs), but were both placed in Cluster B. Given that 16I27a (from a body screen sample) was isolated one day prior to 16B27b (recovered from a positive blood culture), it is likely that this patients diversity represents an heterogenous infecting population. In patient C, 16B24b was isolated from a clinical pacing wire sampled five days after the initial isolation of 16B24a from sputum sample. These two isolates differ by over 400 SNPs, shown by their disparate position in the phylogeny (Figure 1) with 16B24a placed in Cluster A, and 16B24b placed in Cluster B. 16B15b showed raised MICs to all echinocandins and flucytosine, which was not displayed in 16B15a. These two isolates were separated by only 120 SNPs, and were phylogenetically placed in Cluster B (Figure 1b). Our results suggest that none of these patients were infected with a single, clonally propagating *C. auris* strain, and instead the hospital (and all patients) was seeded with a genetically heterogenous population.

Three isolates (16B25, 16B20,16I29a) clustered closely together with only 99 SNPs difference. 16I29a was recovered from the environment around a *C. auris*-positive patient (the isolate was not sequenced). The patient from which 16B20 was recovered was present in an adjacent side room at the same time, suggesting a potential transmission between these patients. However, it remains unclear how the organism may have been transferred between the two rooms and patients. The third isolate of this cluster (16B25) was recovered from a patient present in the same room where 16I29a was recovered. The patient was placed in this room 22 days after the previous *C. auris*-positive patient had left the room, and became positive within 14 days of being in this isolation room.

Isolates 15B5, 16B15b and 16B21 are phylogenetically related (separated by 95 SNPs; Figure 1b), and were recovered from patients sharing the same bay. The bay was initially populated with the patient from whom isolates 15B5 and 16B15b were recovered, and then saw the introduction of the patient from which 16B21 was recovered. Both examples of spatiotemporal clustering suggest that there may be environmental persistence of *C. auris* resulting in transmission to patients.

Discussion

C. auris is an MDR fungal pathogen, capable of causing invasive infections. Here we report WGS of *C. auris* infections from the largest outbreak to date, occuring between April 2015 and November 2016 in a London hospital ICU which spread to two other wards.

A gold standard reference genome for the outbreak was assembled using long MinION-generated reads, and Illumina short reads. Whilst the GC bias and base quality in >80% GC regions of the genome was similar in both Illumina and MinION sequencing, Illumina was more consistent across the whole genome (Figure S1). MinION reads displayed wide variation in base quality in >85% AT regions, ranging from a quality score of zero to 1·4, whereas Illumina reads ranged between quality scores of 0·75 and 0·85 for regions >85% AT. ONT has since released new chemistry that improves read quality, and therefore the variant calling, which will provide a competitive alternative to Illumina sequencing in both outbreak settings and routine research.

The speed of the MinION sequencer allowed rapid assembly and mapping of Illumina short reads to call high-confidence SNPs, which is of great importance in an outbreak (14,15). SNPs in *ERG11* correlated with known *C. albicans* hotspots (19) conferring resistance to the frontline drug fluconazole. All *C. auris* isolates in this study that exhibited high MICs to fluconazole (>250 ug/ml) contained amino acid substitutions known to cause resistance to fluconazole (19). Fourteen isolates were MDR, posing an important clinical challenge in the treatment of *C. auris* infections. Within this study we have also highlighted resistance to posaconazole, which

has not previously been reported in *C. auris*, alongside echinocandin and flucytosine resistance in one MDR isolate.

Echinocandin resistance is linked to mutations in the *FKS* genes in other *Candida* species (27), and our analysis identified the S645Y mutation in *FKS1* in an echinocandin-resistant isolate. Flucytosine resistance was observed in the same isolate, and was associated with the F211I mutation in *FUR1*. Our analysis suggests that systemic echinocandin and flucytosine treatment can rapidly select for resistant genotypes across the outbreak timescale. Rapid MinION sequencing of *C. auris* isolates would allow drug resistance mutation identification, providing time-sensitive information in a clinical setting.

Phylogenomic analysis showed weak support for monophyletic status of isolates within the hospital, suggesting multiple introductions of *C. auris* occurred. On average, only 240 SNPs separated the UK outbreak isolates from the Indian/Pakistani clade, clearly showing the UK outbreak had an Asian origin. When compared to other sequenced Asian *C. auris* isolates (Table S3) in Lockhart and colleagues (9), these SNP numbers suggest an anomalous amount of diversity within this outbreak. Future alignments will require clade-specific references due to the large evolutionary distances between South America/Africa and the Indian/Pakistani clades. Although the mode of introduction into the UK is unknown, temporal analysis of the outbreak isolates placed the most recent common ancestor as early March 2015, which correlates closely to the first confirmed infection within the hospital one month after this, suggesting a recent introduction into the UK.

Only one third of SNPs were shared between hospital environment isolates (bed and trolley of a confirmed *C. auris* infected patient). *C. auris* was also recovered from inanimate surfaces, suggesting a population of genotypes are capable of contaminating the hospital environment, causing onward infection of human hosts (9). Given SNP differences between isolates recovered from the same patient at different bodily locations (Patients A, B and C), either the outbreak diversity is due to multiple introductions, which is unlikely as all patients screened negative for *C. auris* upon admission, or a genetically heterogenous population seeded the hospital. Further, given the substantial number of SNPs separating the two isolates that were recovered from Patient C 24 hours apart, it appears clear that genetically disparate isolates are capable of infecting the same host at the same time. The genomic diversity of *C. auris* within this outbreak makes mapping local-scale transmission events difficult as genetic bottlenecks may result in rapid changes in allele frequencies within local spatiotemporal scales. Clearly, sequencing multiple isolates of *C. auris* from patients alongside those from other UK outbreak settings is needed to more finely resolve the population genomic structure of this pathogen in order to understand transmission dynamics.

This study represents the first use of ONT MinION sequencer on a human fungal pathogen. The association of nsSNPs in *FKS1* and *FUR1* with echinocandin and flucytosine resistance respectively are both clinically relevant and novel. Further investigation into these mutations is required to confirm these associations. Epidemiological analysis suggests that contact tracing was not sufficient to resolve fine-scale spatiotemporal processes across the outbreak due to the multiple differential episodic selection occurring across a genetically heterogenous *C. auris* population. Future research into *C. auris* should focus on sequencing many isolates from the same patient, from multiple body sites, in order to correctly establish the nature of persistence and transmission of *C. auris* within hospital environments; the genomic approaches underpinning this study will likely be cornerstones of future research into this increasingly important infectious disease.

Acknowledgements

We thank Oxford Nanopore Technologies for their generous contribution of flow cells and nanopore sequencing kits. We also wish to extend our thanks to Nicholas J. Loman (University of Birmingham) for providing insight into the initial conception of experiments, and to Anastasia Litvintseva, Nancy A. M. Chow and Kizee Etienne (Centers for Disease Control, USA) for the provision of additional sequence data. We also thank the sequencing team at Wellcome Trust Sanger Institute for sequencing four of the isolates presented here.

J.R. was supported by an Antimicrobial Research Collaborative (ARC) early career research fellowship (RSRO_54990).

R.A.F. was supported by an MIT/Wellcome Trust Fellowship.

M. C. F. was supported by the Natural Environmental Research Council.

D. A-J was supported by NERC and the Wellcome Trust.

C. A. C was supported by the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under grant number U19AI110818.

D. M. A. supported by the Wellcome Trust (Grant no. 099202)

Genome sequencing was provided by MicrobesNG (http://www.microbesng.uk), which is supported by the BBSRC (grant number BB/L024209/1)

Author contributions

Clinical and outbreak data analysis: S.S.

Collected isolates: S.S.

Conceived experiments: J.R., M.C.F., D.A-J., S.S.

DNA extractions: A.A., J.R.

MinION DNA sequencing: J.R.

Illumina sequencing: D.M.A. and MicrobesNG

Bioinformatic analysis: J.R.

Genome annotation: R.A.F., C.A.C.

Manuscript preparation: J.R., R.A.F., A.A, S.S.

Tables

Table 1: Clinical isolates of *C. auris* included in this study

Isolate ID	Sequencing technology	Isolate date	Site	Origin
15B5	MinION, Illumina	19/10/2015	Vascath site	Patient
15B6	Illumina	22/10/2015	Swan ganz tip	Patient
15B10	Illumina	28/12/2015	Groin swab	Patient
16B12	Illumina	04/01/2016	Axilla	Patient
16B13	Illumina	11/01/2016	Fem picco site	Patient
16I29b	Illumina	18/01/2016	Bed	Environment
16I29a	Illumina	18/01/2016	Bed window trolley	Environment
16B18	Illumina	01/02/2016	CVP tip arm	Patient
16B15a	MinION, Illumina	06/02/2016	Blood culture	Patient D
16I17	Illumina	08/02/2016	Screen	Patient
16B20	MinION, Illumina	16/02/2016	Blood culture	Patient
16B16	Illumina	21/02/2016	Urine catheter	Patient

16B21	MinION, Illumina	22/02/2016	Axilla swab	Patient
16B22a	Illumina	27/02/2016	Axilla swab	Patient A
16B24a	Illumina	07/03/2016	Sputum	Patient C
16B22b	Illumina	09/03/2016	CVP tip femoral	Patient A
16B15b	Illumina	09/03/2016	CVP tip	Patient D
16B24b	Illumina	12/03/2016	Pacing wire	Patient
16B25	MinION, Illumina	13/03/2016	ECMO site	Patient C
16I27a	Illumina	14/03/2016	Screen	Patient B
16B26	Illumina	14/03/2016	SCMO site	Patient
16B27b	Illumina	15/03/2016	Blood culture	Patient B
16I30	Illumina	21/03/2016	Screen	Patient
16I33	Illumina	16/06/2016	Groin	Patient
16B31	Illumina	16/10/2016	CVP line site	Patient
16B30	Illumina	17/10/2016	Groin	Patient
16I34	Illumina	23/10/2016	Axilla swab	Patient

Table 2: MICs at time of isolation of all isolates included in this study. MICs deemed above these breakpoint values (based on *Candida albicans* CSLI), and therefore resistant, are shaded in grey. ANI = Anidulafungin; MICA = micafungin; CAS = Caspofungin; 5FC = 5-fluorocytosine; POS = posaconazole; VOR = voriconazole; ITR = itraconazole; FLU = fluconazole; AmpB = Amphotericin B. N/A = No MICs carried out.

Isolate ID	ANI	MICA	CAS	5FC	POS	VOR	ITR	FLU	AmpB
15B5	0·12	0·12	0·25	0·12	8	8	16	256	2
15B6	0·06	0·06	0·06	0·06	0·015	0·25	0·06	256	1
15B10	0·12	0·12	0·5	0·12	8	8	16	256	2
16B12	0·12	0·12	0·5	0·25	8	8	16	256	2
16B13	0·12	0·12	0·25	0·06	0·03	0·5	0·06	256	1
16I29b	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
16I29a	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
16B18	0·25	0·25	1	0·06	0·12	2	0·12	256	1
16B15a	0·12	0·06	0·12	0·06	0·008	0·06	0·015	16	0·5
16I17	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
16B20	0·06	0·12	0·5	0·06	8	8	16	256	1
16B16	0·12	0·12	0·25	0·06	0·03	0·25	0·03	256	1
16B21	0·12	0·12	0·25	0·06	0·12	0·5	0·03	16	1
16B22a	0·12	0·12	0·25	0·06	0·008	0·25	0·03	128	0·5
16B24a	0·12	0·12	0·5	0·06	8	8	8	256	2
16B22b	0·12	0·12	0·25	0·06	4	8	2	256	1
16B24b	0·12	0·12	0·25	0·06	8	8	2	256	1
16B25	0·12	0·12	0·25	0·06	0·03	0·12	0·03	64	1
16I27a	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
16B26	0·12	0·12	0·25	0·06	8	8	16	256	1
16B27b	0·12	0·12	1	0·12	8	8	16	256	0·5
16I30	0·25	0·12	0·5	0·12	8	8	16	256	2
16B15b	8	8	8	64	0·015	0·25	0·03	8	1
16B30	0·12	0·06	0·12	0·06	0·15	0·5	0·06	128	1
16B31	0·12	0·06	0·12	0·06	0·03	0·5	0·06	256	1
16I33	0·06	0·06	0·06	0·06	0·06	0·06	0·03	128	0·5
16I34	0·06	0·06	0·06	0·06	0·008	0·25	0·03	256	0·5

Table 3: Summary of assembly and annotation statistics of the *C. auris* 16B25 genome, the *C. auris* B8441 reference (9)**, the *C. auris* Ci 6684 reference** (21)**, and other pathogenic *Candida* species reference genomes** (20)

Species	Genome size (Mb)	Number of chromosomes or scaffolds	GC content (%)	Number of protein coding genes	Average CDS size (bp)	Ploidy

C. auris (16B25)	12·3	110	45·13	5366	1564	Haploid
C. auris (B8441)	12·4	19	45·13	5439	1548	Haploid
C. auris (Ci 6684)	12·5	99	44·53	8358	1025	Haploid
C. albicans SC5314	14·3	7	33·5	6107	1468	Diploid
C. tropicalis MYA-3404	14·5	24	33·1	6163	1454	Diploid
C. parapsilosis	13·1	24	38·7	5733	1533	Diploid
C. lusitaniae	12·1	9	44·5	5941	1382	Haploid

Figure legends

Figure 1 – Phylogenetic analysis of *C. auris* isolates with bootstrap support (500 replicates) performed on WGS SNP data to generate unrooted maximum-likelihood phylogenies. Branches supported to 75% or higher unless otherwise stated. Branch lengths represent the average expected rates of substitution per site. **a)** Outbreak isolates from the UK (shown in blue) were combined with isolates from around the globe, including India (orange), Pakistan (red), Venezuela (pink), Japan (turquoise), and South Africa (green), to infer a possible geographical origin. Isolates with known mutations in the *ERG11* gene associated with resistance to fluconazole in *C. albicans* are shaded: Y132F as red, and F126L as blue **b)** Given the likely Indian/Pakistani origin of the outbreak isolates, phylogenetic analysis was repeated (as stated above), excluding isolates from South Africa, Venezuela and Japan to illustrate the UK outbreak. Isolates separating either into Cluster A (green) or B (purple) are depicted to infer likely introductions into the hospital.

Figure 2 – Root-to-tip regression analysis of all 28 *C. auris* outbreak isolates. Genetic distance is plotted against sampling time for the phylogeny of the *C. auris* outbreak. Each data point represents a tip on the phylogeny. The R^2 for the regression and the slope, reflecting the evolutionary rate (in substitutions per site per day) is also shown.

References

1. Satoh K, Makimura K, Hasumi Y, Nishiyama Y, Uchida K, Yamaguchi H. Candida auris sp. nov., a novel ascomycetous yeast isolated from the external ear canal of an inpatient in a Japanese hospital. Microbiology and Immunology. Blackwell Publishing Asia; 2009 Jan 1;53(1):41–4.

2. Chowdhary A, Sharma C, Duggal S, Agarwal K, Prakash A, Singh PK, et al. New clonal strain of Candida auris, Delhi, India. Emerging Infect Dis. 2013 Oct;19(10):1670–3.

3. Lee WG, Shin JH, Uh Y, Kang MG, Kim SH, Park KH, et al. First three reported cases of nosocomial fungemia caused by Candida auris. J Clin Microbiol. American Society for Microbiology; 2011 Sep;49(9):3139–42.

4. Magobo RE, Corcoran C, Seetharam S, Govender NP. Candida auris-associated candidemia, South Africa. Emerging Infect Dis. 2014 Jul;20(7):1250–1.

5. Calvo B, de Azevedo Melo AS, Perozo-Mena A, Hernandez M, Francisco EC, Hagen F, et al. First report of Candida auris in America: clinical and microbiological aspects of 18 episodes of candidemia. Journal of Infection. Elsevier Ltd; 2016 Jul 20;:1–24.

6. Schelenz S, Hagen F, Rhodes JL, Abdolrasouli A, Chowdhary A, Hall A, et al. First hospital outbreak of the globally emerging Candida auris in a European hospital. Antimicrobial Resistance & Infection Control. Antimicrobial Resistance & Infection Control; 2016 Oct 12;:1–7.

7. Ben-Ami R, Berman J, Novikov A, Bash E, Shachor-Meyouhas Y, Zakin S, et al. Multidrug-Resistant Candida haemulonii and C. auris, Tel Aviv, Israel. Emerging Infect Dis. 2017 Feb;23(1).

8. Schwartz IS, Hammond GW. First reported case of multidrug-resistant Candida auris in Canada. Canada Communicable Disease Report. 2017 Jul 10;43(7/8):150–3.

9. Lockhart SR, Etienne KA, Vallabhaneni S, Farooqi J, Chowdhary A, Govender NP, et al. Simultaneous emergence of multidrug resistant Candida auris on three continents confirmed by whole genome sequencing and epidemiological analyses. CLIN INFECT DIS. Oxford University Press; 2017;64(2):134–40.

10. Prakash A, Sharma C, Singh A, Kumar Singh P, Kumar A, Hagen F, et al. Evidence of genotypic diversity among Candida auris isolates by multilocus sequence typing, matrix-assisted laser desorption ionization time-of-flight mass spectrometry and amplified fragment length polymorphism. Clin Microbiol Infect. 2016 Mar;22(3):277.e1–9.

11. Litvintseva AP, Hurst S, Gade L, Frace MA, Hilsabeck R, Schupp JM, et al. Whole-genome analysis of Exserohilum rostratum from an outbreak of fungal meningitis and other infections. Land GA, editor. J Clin Microbiol. 2014 Sep;52(9):3216–22.

12. Etienne KA, Gillece J, Hilsabeck R, Schupp JM, Colman R, Lockhart SR, et al. Whole genome sequence typing to investigate the Apophysomyces outbreak following a tornado in Joplin, Missouri, 2011. Jordan IK, editor. PLoS ONE. 2012;7(11):e49989.

13. Engelthaler DM, Chiller T, Schupp JA, Colvin J, Beckstrom-Sternberg SM, Driebe EM, et al. Next-Generation Sequencing of Coccidioides immitis Isolated during Cluster Investigation. Emerging Infect Dis. Centers for Disease Control and Prevention; 2011 Feb 1;17(2):227–32.

14. Quick J, Ashton P, Calus S, Chatt C, Gossain S, Hawker J, et al. Rapid draft sequencing and real-time nanopore sequencing in a hospital outbreak of Salmonella. Genome Biol. Genome Biology; 2015 Jun 29;:1–14.

15. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable genome sequencing for Ebola surveillance. Nature. Nature Publishing Group; 2016 Feb 3;:1–17.

16. Faria NR, Sabino EC, Nunes MRT, Alcantara LCJ, Loman NJ, Pybus OG. Mobile real-time surveillance of Zika virus in Brazil. Genome Medicine. BioMed Central; 2016 Sep 29;8(1):97.

17. Rhodes J, Beale MA, Vanhove M, Jarvis JN, Kannambath S, Simpson JA, et al. A Population Genomics Approach to Assessing the Genetic Basis of Within-Host Microevolution Underlying Recurrent Cryptococcal Meningitis Infection. G3 (Bethesda). 2017 Apr 3;7(4):1165–76.

18. Rambaut A, Lam TT, Max Carvalho L, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. The Oxford University Press; 2016 Apr 10;2(1):vew007–7.

19. Morio F, Loge C, Besse B, Hennequin C, Le Pape P. Screening for amino acid substitutions in the Candida albicans Erg11 protein of azole-susceptible and azole-resistant clinical isolates: new substitutions and a review of the literature. Diagnostic Microbiology and Infectious Disease. Elsevier Inc; 2010 Apr 1;66(4):373–84.

20. Butler G, Rasmussen MD, Lin MF, Santos MAS, Sakthikumar S, Munro CA, et al. Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature. Nature Publishing Group; 2009 Jun 4;459(7247):657–62.

21. Chatterjee S, Alampalli SV, Nageshan RK, Chettiar ST, Joshi S, Tatu US. Draft genome of a commonly misdiagnosed multidrug resistant pathogen Candida auris. BMC Genomics. BMC Genomics; 2015 Sep 5;16(686):1–16.

22. Jeffares DC, Rallis C, Rieux A, Speed D, Převorovský M, Mourier T, et al. The genomic and phenotypic diversity of Schizosaccharomyces pombe. Nat Genet. Nature Research; 2015 Mar;47(3):235–41.

439 23. Liti G, Carter DM, Moses AM, Warringer J, Parts L, James SA, et al. Population genomics of domestic
440 and wild yeasts. Nature [Internet]. Nature Publishing Group; 2009 Mar 10;458(7236):337–41. Available
441 from: http://www.nature.com/doifinder/10.1038/nature07743

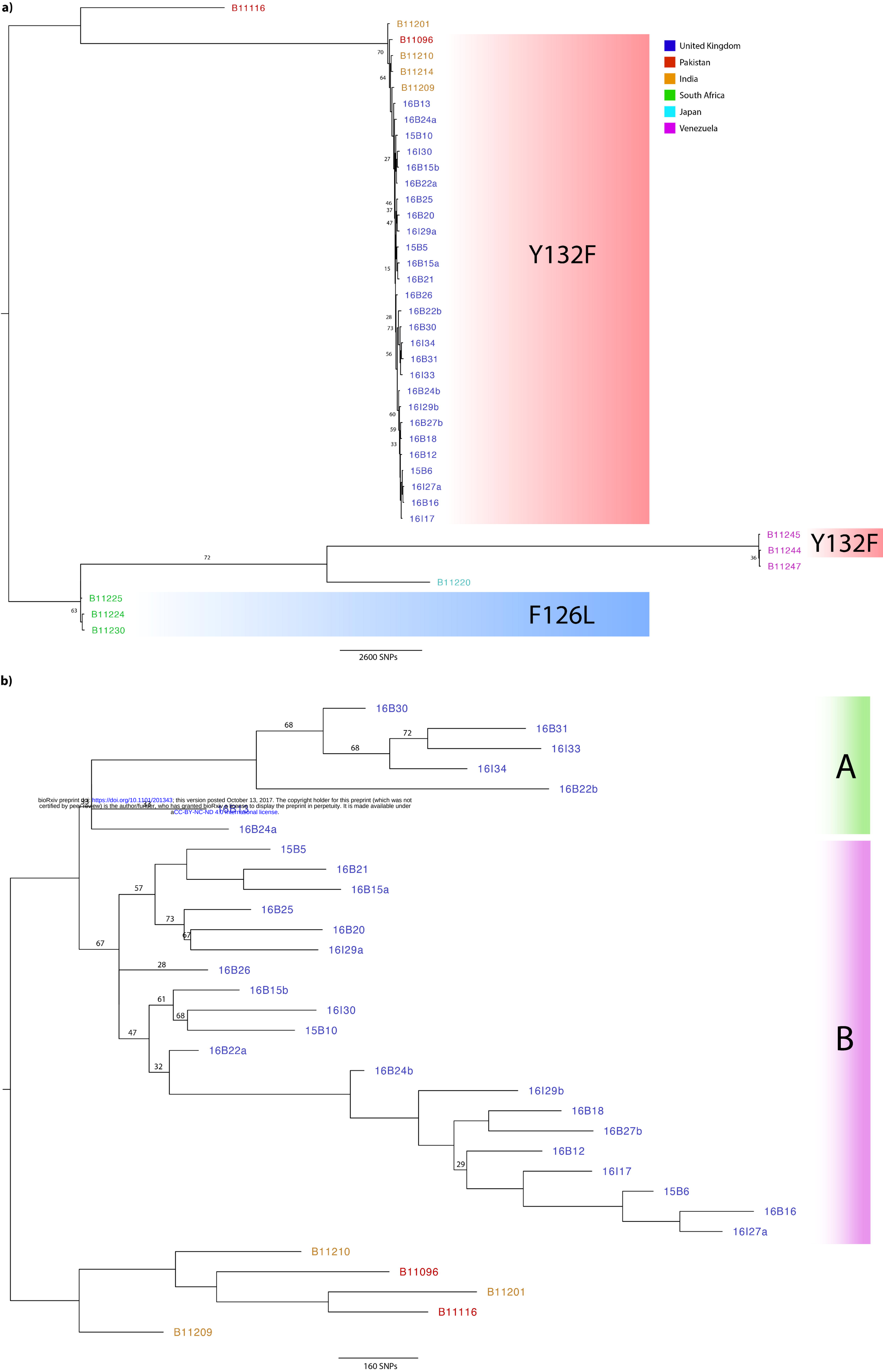
442 24. Balashov SV, Park S, Perlin DS. Assessing resistance to the echinocandin antifungal drug caspofungin
443 in Candida albicans by profiling mutations in FKS1. Antimicrob Agents Chemother. American Society
444 for Microbiology; 2006 Jun;50(6):2058–63.

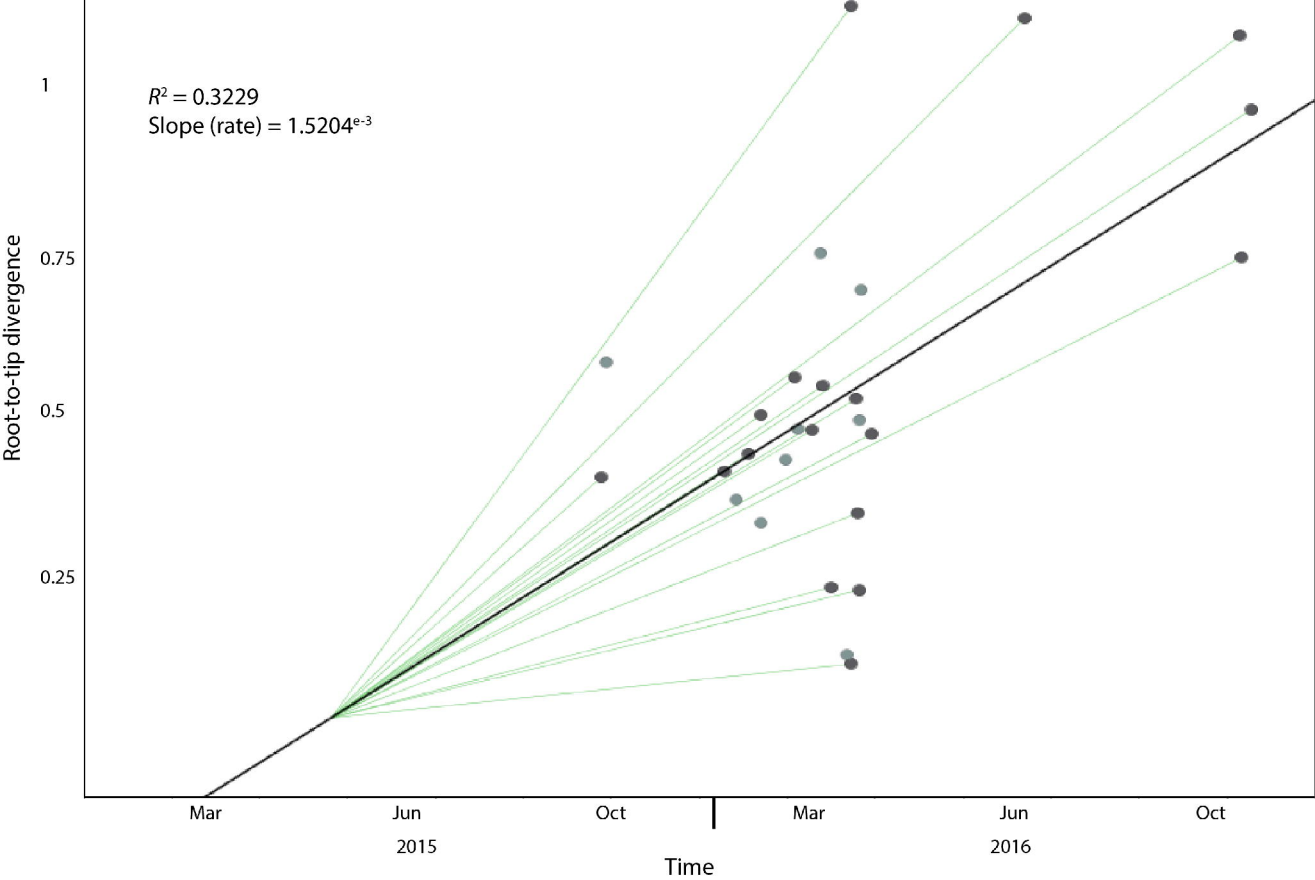
445 25. Dodgson AR, Dodgson KJ, Pujol C, Pfaller MA, Soll DR. Clade-specific flucytosine resistance is due to
446 a single nucleotide change in the FUR1 gene of Candida albicans. Antimicrob Agents Chemother.
447 American Society for Microbiology; 2004 Jun;48(6):2223–7.

448 26. Chowdhary A, Anil Kumar V, Sharma C, Prakash A, Agarwal K, Babu R, et al. Multidrug-resistant
449 endemic clonal strain of Candida auris in India. Eur J Clin Microbiol Infect Dis. Springer Berlin
450 Heidelberg; 2013 Dec 20;33(6):919–26.

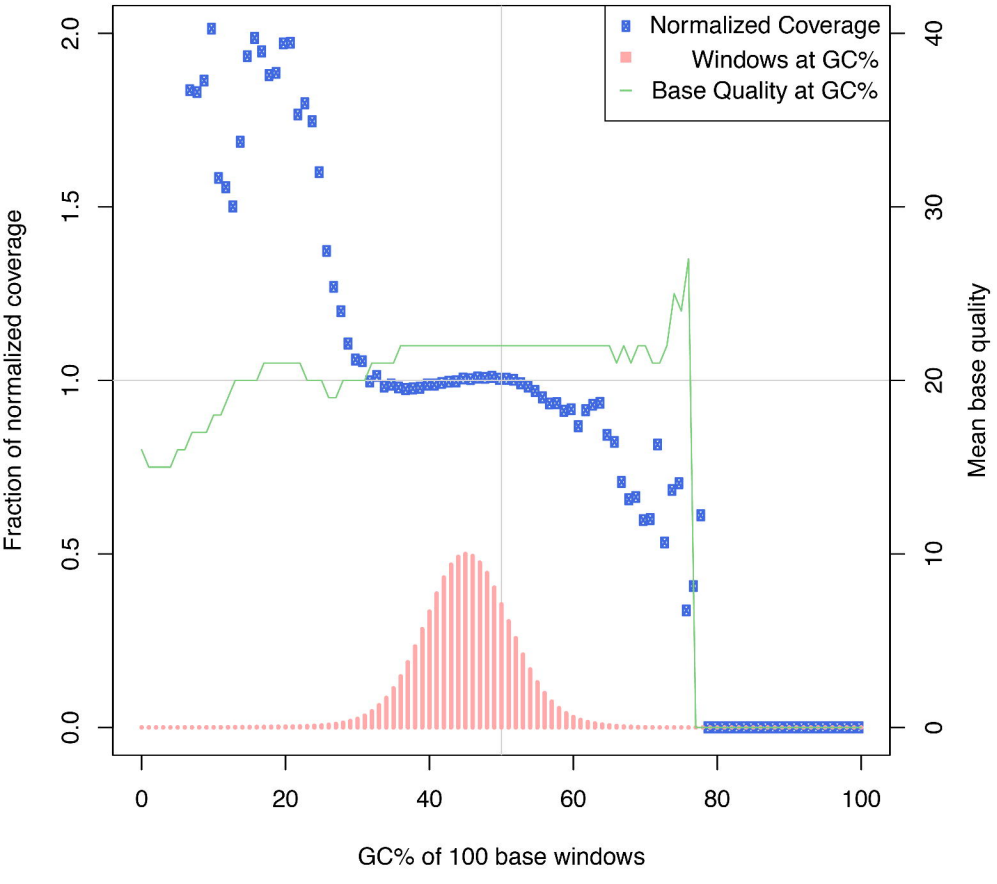
451 27. Garcia-Effron G, Lee S, Park S, Cleary JD, Perlin DS. Effect of Candida glabrata FKS1 and FKS2
452 mutations on echinocandin sensitivity and kinetics of 1,3-beta-D-glucan synthase: implication for the
453 existing susceptibility breakpoint. Antimicrob Agents Chemother. American Society for Microbiology;
454 2009 Sep;53(9):3690–9.

455

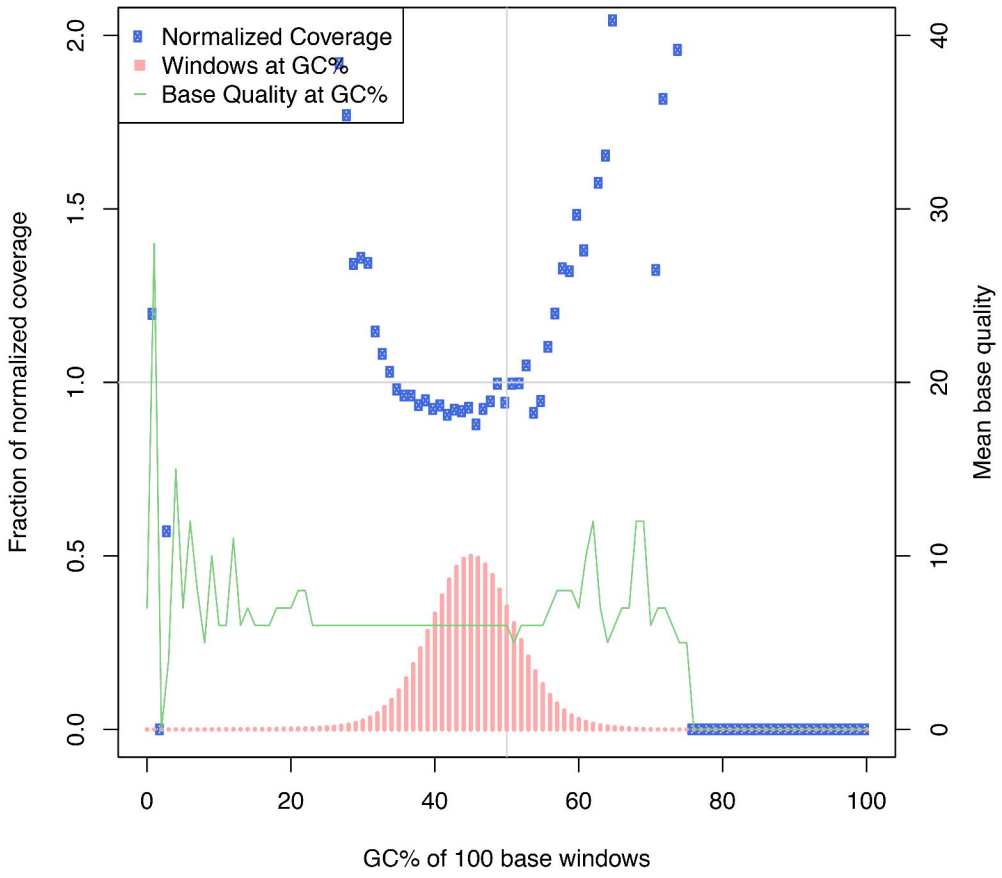



$R^2 = 0.3229$
Slope (rate) = 1.5204^{e-3}
1
0.75
0.5
0.25
Root-to-tip divergence
Mar
Jun
Oct
Mar
Jun
Oct
2015
2016
Time

a)


Illumina short reads aligned against hybrid reference
Normalized Coverage
Windows at GC%
Base Quality at GC%
Fraction of normalized coverage
Mean base quality
GC% of 100 base windows
0.0
0.5
1.0
1.5
2.0
0
10
20
30
40
0
20
40
60
80
100

b)


ONT long reads aligned against hybrid reference
Normalized Coverage
Windows at GC%
Base Quality at GC%
Fraction of normalized coverage
Mean base quality
GC% of 100 base windows
0.0
0.5
1.0
1.5
2.0
0
10
20
30
40
0
20
40
60
80
100