Migration and standing genetic variation in vaginal and rectal yeast populations in recurrent vulvovaginal candidiasis

Abdul-Rahman Adamu Bukari[a], Rebekah J. Kukurudz-Gorowski[a], Alexia de Graaf[a*], Devin Habon[a*], Beamlak Manyaz[a*], Yana Syvolos[a], Aruni Sumanarathne[a], Vanessa Poliquin[b], Aleeza Gerstein[a,b,#]

[a]Department of Microbiology, Faculty of Science, University of Manitoba, Winnipeg, Manitoba, Canada

[b]Obstetrics, Gynecology and Reproductive Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

[c]Department of Statistics, Faculty of Science, University of Manitoba, Winnipeg, Manitoba, Canada

Running head: Diversity in recurrent vulvovaginal candidiasis

* Alexia de Graaf, Devin Habon, and Beamlak Manyaz contributed equally to this work and are listed alphabetically

Author Correspondence to Aleeza Gerstein, aleeza.gerstein@umanitoba.ca

Abstract

Vulvovaginal candidiasis is one of the most common vaginal and fungal infections. The majority of infections are successfully treated with antifungal drugs. However, ~8% of cases lead to chronic recurrent vulvovaginal candidiasis ("RVVC"), and approximately half of RVVC cases are idiopathic. Previous research has generally found closely-related isolates within vaginal and rectal populations and between subsequent infections. However, their coarse methods preclude assessing the fine-scale relationships among closely related isolates and measuring standing genetic variation, a fundamental property of populations with implications for evolutionary potential. To address this gap, we isolated 12 vaginal and 12 rectal yeast isolates during symptomatic relapse from four individuals with a history of RVVC. Three participants had *Candida albicans* infections, while the fourth had *Nakaseomyces glabratus*. All isolates were whole-genome sequenced and phenotyped. The isolates were placed into the global phylogenies, which included constructing an updated *N. glabratus* tree containing over 500 isolates. Multiple analyses were consistent with frequent migration between sites. Although there are extremely few comparables, *C. albicans* population nucleotide diversity was similar to most commensal oral and rectal populations, while *N. glabratus* was similar to some bloodstream infections, yet higher than others. Diversity was largely driven by single nucleotide changes; no aneuploidies were found, and although loss-of-heterozygosity tracts were common in the populations, only a single region on chr1L varied among isolates from one participant. There was very little phenotypic diversity for drug response or growth and no consistent difference between isolates from different sites for invasive growth. Combined, this study provides baseline measurements and describes analysis techniques to quantify within-population diversity. We highlight a critical need for comparable studies that use the same sampling effort, sequencing, and analysis methods to understand the interplay between selection, drift, and migration in shaping fungal microbial communities in this and other important contexts.

Author Summary

Recurrent vaginal yeast infections are relatively common, and we do not understand why some people experience these chronic infections when many others will have a single infection that is successfully treated and cleared. Many open questions remain about the basic biology of the yeast populations involved. We quantified diversity using modern sequencing technology within vaginal and rectal yeast populations from four individuals with a history of recurrent yeast infections experiencing symptoms. Three participants had a *Candida albicans* infection (the most common causative species), while the fourth had a *Nakaseomyces glabratus* infection (the second most common and increasingly implicated). We found that vaginal and rectal isolates were closely related, indicating the same population is present at the two sites. We found, surprisingly, that diversity was similar to the yeast populations found at other body sites in healthy people. Our study highlights a critical need for additional studies to be done following the same methods in different contexts to understand better the fungal microbial populations that reside in our bodies.

Introduction

Vulvovaginal candidiasis (VVC, colloquially "yeast infection") is a mucosal infection of the lower female reproductive tract. It is very common, affecting approximately 75% of women at least once in their lives [1–3]. The disease burden of VVC results in global annual treatment costs of ~ 1.8 billion USD [4,5], with a loss in productivity in high-income countries of ~ 14 billion USD [4]. Treatment involves topical or oral antifungal medication, which is effective at symptom abatement in most cases. However, ~ 8% of women experience recurrence (RVVC), defined as three or more symptomatic episodes a year [4,5]. Even with weekly drug treatment, most people with RVVC experience symptomatic recurrence within 8-9 months, compared to 2 months in an untreated control group [6]. Many diverse factors are known to trigger recurrence or predispose individuals to RVVC [7,8], and references within], including known human genetic variants, such as mannose-binding lectin deficiency [9,10] and TLR2 Pro631His polymorphism [11]. However, approximately half of all people with RVVC have no identifiable risk factors [12], signifying the need for additional studies on the biological basis of this chronic condition.

Candida albicans is responsible for 50-90% of VVC and RVVC cases (collectively, R/VVC [13–17]). *Nakaseomyces glabratus* (formerly *Candida glabrata*) is the second most prevalent cause of R/VVC, globally attributed to ~ 8% of cases [18–20]. Here, we collectively describe these species using the colloquial term "yeast," which reflects a shared morphology, rather than the historical and still common "*Candida* spp." given our current understanding of their phylogenetic relationships and recent official renaming of *N. glabratus* [21,22]. To be consistent with clinical practice, we will continue to use the R/VVC abbreviations while noting that "candidiasis" does not reflect the updated genera names. Treatment recommendation for R/VVC differs by species, as many isolates from *N. glabratus* (and other non-*albicans* pathogenic yeast species) have intrinsic resistance to the common azole antifungals [23]. Other species-specific differences suggest there may be clinical and phylogenetic reasons not to lump these species together. For example, Pra1, whose expression in *C. albicans* sequesters zinc, promoting neutrophil recruitment and consequently leading to inflammatory vaginitis in an RVVC setting, is absent in *N. glabratus* [24,25].

Understanding the etiology of R/VVC is complicated in part since yeast exists as a common commensal member of the vaginal microbiota without causing symptomatic VVC. Identification of vaginal yeast in the absence of symptoms does not warrant treatment. Approximately 70% of healthy women were found to have vaginal yeast at a single time point through an amplicon-based study [26], which is much higher than the 20-30% asymptomatic commensal colonization estimate frequently cited from culture-based methods (e.g., [27]).

Furthermore, self-diagnosis due to symptoms has also been shown to be problematic, as other gynecological infections and dermatologic conditions have overlapping symptoms [28]. Genetic typing and phylogenetic strain analyses in *C. albicans* have repeatedly found no strict phylogenetic differentiation between commensal and pathogenic strains. However, some phylogenetic clades are potentially over-represented by commensal strains and those causing superficial disease [29,30].

Relapse in RVVC, i.e., return of symptoms after treatment, could theoretically be due to either incomplete eradication of the vaginal yeast population between symptomatic infections or complete vaginal eradication followed by re-colonization from an endogenous or exogenous population [31]. There are decades of studies that have sought to understand the etiology of RVVC, as the answer has potential implications for improving treatment and reducing or eliminating symptom recurrence. Re-colonization from a sexual partner has been proposed as a possible mechanism. Yet, recurrence also occurs in non-sexually active individuals [32], and a review of eight randomized clinical trials found that treating asymptomatic sexual partners did not influence or cure recurrence rates [33]. The GI tract has also been suggested as a potential endogenous source population. A study in 1979 that treated RVVC patients with oral nystatin to reduce the resident GI population found that it did not decrease the time to recurrence [34]. Furthermore, studies examining yeast colonization of the GI tract through feces or rectal swabs during an active infection find that not all participants are culture-positive [34–40]. However, this does not necessarily preclude that a small GI population is present in all individuals (below the culture detection limit in some), which could act as an endogenous reintroduction source under the right host conditions.

Examining the diversity of strains at different body sites and among recurrent infections can potentially differentiate between relapse scenarios. If vaginal isolates are closely related to GI tract/rectal isolates but less diverse, that would be consistent with reintroduction. If vaginal isolates sampled at different time points consistently have the same genotype that is not present in the GI tract/rectal isolates, this would be consistent with incomplete eradication. Looking at the level of diversity within a single time point during an active symptomatic infection ("standing genetic variation") at different body sites can also help us to understand the adaptive potential and migration dynamics of the yeast populations within the body.

The earliest studies examined within-population diversity by looking at 'resistogram' or phenotypic profiles of strains, by typing strains based on their ability to grow in a panel of diverse chemical challenges [40–42]. This was followed by studies using many different methods of genomic profiling to assess relatedness among strains, such as rDNA restriction

109 fragment length [43], DNA/PCR fingerprinting [35,44–46], CHEF gel analysis [47], microsatellite
110 loci [48–51], and randomly amplified polymorphic DNA (RAPD) [46,52–54]. A higher genomic
111 resolution and better ability to accurately map phylogenetic relationships among divergent
112 strains come from recent studies using multilocus sequence typing (MLST) [17,55,56].
113 Considering this body of existing work, the overall picture does not change drastically with the
114 method employed. The overarching results are consistent with a role in the maintenance of
115 genotypes (i.e., the same or very similar genotype is recovered from different body sites and
116 within the vagina between successive symptomatic episodes), alongside an occasional
117 opportunity for novel genotypes to arise (i.e. reinfection, sometimes referred to as strain
118 turnover). However, previous studies do not necessarily rule out a role for standing genetic
119 variation to have been consistently present yet undetected, i.e., relapse masquerading as
120 reinfection, as typically only one or a small number of isolates were examined at a given time.
121 Only a single study sequenced two vaginal isolates from different time points through whole
122 genome sequencing (WGS) [57] , and no previous studies have used WGS to compare multiple
123 isolates from the same time point. WGS is required to accurately assess relatedness, diversity,
124 and the potential source of genetic novelty. WGS removes the need to rely on a single or small
125 number of markers, which could potentially over or underinflate the true level of diversity. As an
126 example, Sitterlé et al. showed that while MLST revealed occasional differences among *C.*
127 *albicans* oral isolates collected from three healthy individuals, whole-genome sequencing
128 revealed that the three examined isolates were closely related in each case [58].
129 Within-species standing genetic variation of yeasts has only been quantified in a handful
130 of contexts. Two studies in *C. albicans* that whole genome sequenced 3-6 oral and rectal
131 isolates from six healthy individuals found that although isolates were closely related in most
132 cases, they differed by numerous single nucleotide polymorphisms, primarily resulting from
133 short-range loss-of-heterozygosity tracts [30,58]. Two individuals were, however, simultaneously
134 colonized with oral isolates from different clades [30]. A single *N. glabratus* study sequenced up
135 to ten isolates from nine patients with bloodstream candidemia [59]; a pairwise SNP analysis
136 was also consistent with closely related isolates. Variation in RVVC has yet to be quantified
137 through WGS; hence, it is unknown whether it is similar to commensal populations. It also
138 needs to be determined how many isolates need to be sequenced to capture population
139 diversity accurately. It is generally thought a large sample size is required to obtain an accurate
140 picture of the genetic variation [60]. Unfortunately, larger samples result in extra sequencing and
141 computational costs. Past studies have shown that population genetics analysis can be
142 conducted with small sample sizes when using many single nucleotide polymorphisms (e.g.

143 from WGS or RADseq) [61]. Consequently, simulation and empirical analyses have shown that
144 the sample size can be reduced to four to six individuals when using a large number of SNPs
145 [60,62,63]. These studies were conducted in sexually reproducing organisms; it remains
146 unexplored how many yeast isolates are required to access the population level of diversity.
147 Here we build on previous work by using WGS paired with high-throughput phenotyping
148 to quantify standing genetic variation in vaginal and rectal isolates during symptomatic RVVC.
149 We obtained isolates from four individuals with a history of RVVC in Winnipeg, Canada. Three
150 participants had *C. albicans* infections, and one had *N. glabratus*. We collectively refer to these
151 isolates as the "THRIVE-yeast" isolates, following the name of our local umbrella research
152 program that studies The Host-microbial Relationships and Immune function in different Vaginal
153 Environments ("THRIVE," http://www.mthrive.ca). From all individuals, we found a complete
154 phylogenetic overlap of vaginal and rectal isolates consistent with high levels of migration. We
155 found no evidence diversity in populations from the two sites was different; levels of standing
156 genetic variation were similar to what has been observed among commensal oral and rectal
157 populations in *C. albicans* and some bloodstream infections in *N. glabratus*. This suggests that
158 despite frequent population bottlenecks caused by treatment, diversity can be maintained or
159 rapidly restored.

Results

THRIVE-yeast isolates in the global species phylogenies

Participants with a history of RVVC were recruited from a monthly clinic specializing in recurrent yeast infections in Winnipeg, Canada. A high number of vaginal and rectal isolates were obtained from swabs taken from four of seventeen participants who were enrolled and screened intermittently between January 2020 and November 2022. Isolates from one participant were *N. glabratus* (YST6), while *C. albicans* was isolated from the other three (TVY4, TVY10, and YST7). We haphazardly isolated 12 vaginal and 12 rectal yeast isolates from each participant. All *C. albicans* isolates were MAT-heterozygous diploids (a/α), while the *N. glabratus* isolates were all MTL1a. We collectively refer to these 93 isolates as the "THRIVE-yeast" isolates.

The phylogenetic relationship of the THRIVE-yeast isolates was evaluated in the context of available whole-genome sequenced (WGS) isolates from each species. We exhaustively searched NCBI for available *N. glabratus* sequences and, in doing so, built the largest *N. glabratus* tree to date. The 526 isolates we found were overwhelmingly blood isolates (n=430); we found only a single vaginal isolate constructed a new global *N. glabratus* phylogeny from 526 previously sequenced isolates. The YST6 isolates are monophyletic and cluster with 68 bloodstream isolates and three isolates of unknown provenance from the United States, Canada, and Australia (Fig. 1A). Unfortunately, over 80% of the sequenced isolates are bloodstream infection isolates, with only one previous isolate annotated as originating from the vagina, rectum, or stool (Table S1), precluding further analysis. The globally published *C. albicans* phylogeny comprises 182 isolates from a wide breadth of geographic and anatomical sites [29]. Isolates from TVY4, TVY10 and YST7 all form monophyletic groups and fall within a subgroup that contains fifteen additional isolates in clade 1, the most common clade (Fig. 1B). TVY4 and TVY10 isolates are beside each other, without other closely related isolates. The YST7 isolates are most closely related to three vaginal isolates (one each from Brazil, Morocco, and China) and one oral isolate from Niger. Five of the remaining twelve isolates in the clade 1 subgroup were also isolated from the vagina. This is a statistical enrichment for vaginal isolates compared to the rest of the isolates in clade 1 (THRIVE-yeast isolates from each participant were counted as a single isolate; Fisher exact test comparing 11 vaginal isolates out of 19 in the subgroup to 5 vaginal isolates out of 24 in the rest of clade 1, $p = 0.011$). If we discount the 35 predominantly vaginal isolates in clade 13, which is now recognized as likely a separate species (*C. africanus)* [64,65], clade 1 as a whole is also statistically overrepresented for vaginal isolates compared to the *C. albicans* tree in general (16 vaginal isolates in clade 1 out of 42 total isolates, compared to 18 vaginal isolates out of 105 total isolates; Fisher exact test, $p = 0.009$).

194 Thus, although the sequenced vaginal isolates are located in six different clades, they are over-
195 represented in clade 1 relative to a neutral expectation that vaginal isolates are equally likely to
196 be found anywhere in the existing tree.
197 The *N. glabratus* phylogeny is composed of long internal branches with only a small
198 amount of variation within the clades; clades defined through MLST are mainly consistent with
199 WGS [66,67]. An *in silico* MLST analysis of our WGS data found that the YST6 isolates cluster
200 with 71 isolates: 67 ST16, two ST60, and two ST187. ST60 is distinguished by a single
201 polymorphism at position 298 in *LEU2,* while ST60 is distinguished by a single polymorphism at
202 position 250 in *NMT1*. *C. albicans* diploid sequence types (DST) do not map as well to WGS
203 clades. Each of the clade 1 subgroup isolates has a different DST [29], and the differences are
204 mainly registered in 39 positions across the genes.
205
206 **Vaginal and rectal isolates are closely related and phylogenetically overlapping.**
207 Following the observed monophyly among isolates from the same participant, we next
208 assessed the relatedness of vaginal and rectal isolates. We first generated phylogenies for each
209 individual using RAxML [68], as previously done for closely related bacteria strains [69]. The
210 vaginal and rectal isolates from all four participants are phylogenetically overlapping (Fig. 2A).
211 Few branches within the four phylogenies had bootstrap support exceeding 80%, yet well-
212 supported clusters in all participants included isolates from both body sites. For each population,
213 we also conducted a local principal component analysis (PCA) that examines the regions of
214 high genomic heterogeneity within populations. From all participants, the regions of high
215 differences were generally distributed throughout the genome, and the PCA failed to segregate
216 vaginal and rectal isolates (Fig. S1).
217
218 **Pairwise differences in single nucleotide polymorphisms among isolates per participant**
219 Although isolates were closely related, WGS data picked up SNP variation among all
220 pairs of participant isolates. To test whether within-population vaginal diversity was lower than
221 within-population rectal diversity, we compared pairwise SNP differences among vaginal
222 isolates to the pairwise SNP differences among rectal isolates. To test whether there was a
223 signal of divergence between sites, we also compared single-site differences to pairwise
224 differences between isolates across sites. The average pairwise SNP differences between
225 vaginal isolates were very similar to average pairwise SNP differences between rectal isolates
226 and between isolates from different sites (Fig. 2B). The only significant difference was in TVY4,
227 where the average SNP differences among rectal isolates were significantly lower than the

228 vaginal isolates (YST7: $F_{2, 250}$ = 2.164, P = 0.117; YST6: $F_{2, 250}$ = 0.785, P =0.457; TVY4: $F_{2, 250}$ =
229 8.599, p =0.000244; TVY10: $F_{2, 250}$ = 1.66, P = 0.192; Tukey's HSD Test for multiple
230 comparisons; p = 0.0001). The distributions were fairly normal, as expected for isolates with low
231 population structure, except that TVY10 isolates showed a bimodal distribution of pairwise SNP
232 differences in both the vaginal and rectal sites.
233
234 **Minimum number of isolates for estimating standing genetic variation within participants**
235 A major goal of our work was to compare diversity within RVVC populations to diversity
236 observed in other contexts to make inferences about the evolutionary process based on the
237 observed degree of standing genetic variation. However, the small number of comparable
238 studies all sequenced different numbers of isolates, and nucleotide diversity (π) will decrease
239 with an increased number of samples taken from a population. To quantify the scale of this
240 effect of changing the number of isolates, we conducted a bootstrap analysis using our 12
241 vaginal isolates from each individual. We repeatedly resampled 3-10 isolates and recalculated
242 diversity. There was subtle variation in nucleotide diversity among the three participants with *C.*
243 *albicans* infection. In all cases, the shape of the diversity curve with the number of isolates was
244 very similar---an elbow was observed around n = 6 (Fig. 3). Nucleotide Diversity in YST6 (*N.*
245 *glabratus*) was two orders of magnitude lower, and there was not a consistent change in
246 diversity with the number of isolates.
247
248 **THRIVE-yeast isolates share similar diversity as isolates from commensal and other**
249 **disease settings**
250 We downloaded the fastq files from two previous *C. albicans* studies on commensal
251 populations [30,58] and used our pipeline to calculate the average nucleotide diversity for each.
252 For all populations, including our own, where necessary we downsampled the number of
253 isolates to three, consistent with the lowest number of isolates sampled from the commensal
254 populations. The average nucleotide diversity was very similar across most populations (Fig.
255 4A). The exception was populations from two participants previously shown to have isolates
256 from different phylogenetic clades. The YST6 vaginal isolate population was compared to
257 reanalyzed fastq data from nine different BSI populations from Badrane et al., which all had
258 between 9-10 isolates [59]. The average nucleotide diversity from our population was similar to
259 the diversity from four participants yet much higher than the average nucleotide diversity from
260 the other five (Fig. 4B).
261

262 **Little variation in copy number or loss of heterozygosity within populations**

263 We examined copy number variation (CNV) and loss-of-heterozygosity (LOH) events
264 among THRIVE-yeast isolates and their closest relatives using Y_{MAP} [70]. No CNVs were
265 identified in any YST6 isolates (Fig. 5A). A single ~50 kb CNV on the right arm of chr3 was
266 identified in all YST7 isolates (Fig. 5B). This CNV is also present in the closest relative to the
267 YST7 isolates, vaginal strain 9518, but is absent in the next two closely related strains that are
268 also vaginal in origin (B116 and M17). As Y_{MAP} visualizations are based on averages across
269 5000 bp sliding windows, we examined the region in finer detail. Coverage was measured from
270 the BAM files to compute the depth at each position in that region. Mapped coverage was
271 inconsistent with the profile of a typical CNV; the majority of the region had only a slightly
272 elevated copy number relative to the rest of the genome (Fig. 6C). Two small (< 200 bp) regions
273 spiked up to ~6-fold and ~14-fold coverage, the first internal to *ALS6* and the second to *ALS7*
274 (Fig. 6C). A third region comprised of elevated coverage maps to another gene with close
275 homology to other genes in the genome, *CYP5*, a putative peptidyl-prolyl cis-trans isomerase
276 [71]. The region identified in Y_{MAP} is thus likely to primarily reflect an error in mapping rather
277 than a true CNV with potential biological effects. No other CNVs or aneuploidies were identified
278 in the other THRIVE-yeast isolates.

279 LOH analysis was consistent with the phylogenetic analysis and diversity metrics.
280 Generally, all isolates from the same participant shared an LOH profile. Among the THRIVE-
281 yeast *C. albicans* isolates compared to the SC5314 reference, nearly all chromosomes had at
282 least one LOH event extending to the telomere (Fig. 6B, Fig. S2). No large interstitial LOH
283 events were identified. TVY4 isolates share a complicated LOH region on the left arm of chr2
284 that flips back and forth from the two haplotypes, as well as small LOH regions on both the left
285 and right arms of chr4. YST7 shares LOH blocks on the right arm of chr2 from the centromere to
286 the telomere, on the left arm of chr5, and on the left arm of chrR with the vaginal isolates it is
287 most closely related to (i.e., 9518, B116, M17) but not Niger8, an oral isolate. The exception of
288 wholly shared LOH blocks among participant isolates was in TVY10, where a single LOH event
289 on the left arm of chr1 was present in twelve TVY10 isolates that cluster together on the
290 phylogeny.

291 In addition to the characteristic LOH regions within populations, some LOH regions were
292 also shared between isolates from different participants. All isolates have an LOH block on the
293 left arm of chr3 and the right arm of chr6. TVY10 and TVY4 isolates share a ~300 kb region on
294 the right arm of chr1. A LOH region on chr7 beside the ancestral chr7 right arm LOH block is
295 also present in both TVY10 and TVY4, though it is larger in TVY10. TVY10 and YST7 (but not

296 TVY4) both have LOH blocks on the left arm of chr5, though the allelic profile between them is
297 different.
298
299 **Phenotypic variation**
300 We quantified within-population phenotypic variation in parallel to genotypic variation.
301 The average growth rate for YST6 *N. glabratus* isolates was higher than the *C. albicans*
302 populations in both Roswell Park Memorial Institute (RPMI) medium (Fig. 7A) and vaginal
303 simulative medium (VSM, Fig. 7B). Growth rates were either the same between vaginal and
304 rectal isolates or the rectal isolates were higher when grown in either medium (Welch Two
305 Sample t-test, Table S3, rectal isolates higher in TVY4 grown in RPMI: t = 2.50, df = 16.1, p-
306 value = 0.023; TVY4 grown in VSM: t = -2.36, df = 18.1, p-value = 0.030; YST7 grown in VSM: t
307 = -3.63, df = 18.8, p-value = 0.002). Consistent with a visual inspection, single statistical rectal
308 outliers with increased growth rate relative to other isolates were seen in YST7 and TVY4 grown
309 in RPMI, and YST6 and TVY10 grown in VSM (Fig. 7, Rosner's test for outliers). No additional
310 outliers were identified in the other populations.
311 We then compared drug resistance and drug tolerance from vaginal and rectal isolates
312 for two common drugs prescribed in our local clinic, fluconazole (FLC) and boric acid (BA). In
313 addition to drug resistance, drug tolerance, the ability of drug-susceptible populations to grow
314 slowly in the presence of high levels of fungistatic drugs, has also emerged as a trait that varies
315 among different fungal species and isolates [25,26]. Tolerance may be implicated in the
316 propensity to cause fungal disease in other infection contexts [27] but has not previously been
317 examined in the context of R/VVC. Drug response was assessed through disk diffusion assays
318 done at pH 4.2. The site of isolation was only significant for YST6 BA resistance (vaginal
319 isolates were slightly more tolerant than rectal isolates; t-test, t = 2.77, df = 18.2, p-value =
320 0.012, Table S2, Fig. 8). A visual examination of the data also indicated that the majority of
321 isolates were very similar to each other. Formal outlier statistical tests that grouped vaginal and
322 rectal isolates were broadly consistent with the qualitative visual assessment, identifying only
323 two outlier isolates for tolerance (both rectal isolates in BA, one from YST6, one from YST).
324 There was considerable variation among participants and isolates for invasive growth
325 (Fig. 8, Table S2). However, there was no difference between YST6 or TVY10 vaginal and
326 rectal isolates. In contrast, YST7 vaginal isolates had higher invasive growth than rectal isolates
327 (W = 115.5, P = 0.0002), and TVY4 rectal isolates exhibited higher invasive growth than vaginal
328 isolates (W = 372, p-value = 0.032). The overall picture is thus that invasive growth seems to
329 vary more between participants than between sites of isolation and that statistical differences

330 between sites are likely due to neutral processes rather than selection for invasive growth in the
331 vaginal environment.

332 **Discussion**

333 The biological basis of RVVC needs to be better understood. To study the vaginal yeast
334 populations implicated in disease, we quantified the diversity of 12 vaginal and 12 rectal isolates
335 from four people with a history of RVVC who were experiencing symptomatic recurrence at the
336 time of sampling. In each case, the isolates formed monophyletic groups, and the vaginal and
337 rectal isolates were phylogenetically overlapping, consistent with frequent migration between
338 the two sites. The level of diversity within the three *C. albicans* populations was generally similar
339 to the small number of populations from previous studies on the diversity of commensal oral and
340 rectal, while the *N. glabratus* population was similar to or higher than populations from BSI
341 infections. Phenotypically, there was minimal diversity in drug responses or growth ability in
342 clinically relevant medium and no consistent difference between isolates from different isolation
343 sites for invasive growth.

344 The observed phylogenetic overlap of vaginal and rectal isolates indicates that the yeast
345 populations in these anatomical sites share a high degree of genetic similarity, suggesting a
346 common ancestral source and a high level of ongoing migration (at least during symptomatic
347 recurrence). Our phenotypic analyses are consistent with the genetic results, with generally low
348 variation and non-statistically different phenotypes between the vaginal and rectal isolate
349 populations. We primarily identified species-level differences in growth rate and drug response
350 phenotypes, which was not unexpected, given the well-established intrinsic differences among
351 isolates from *C. albicans* and *N. glabratus* [72–74]. This is also in concordance with previous
352 studies that found high genetic similarity using coarse sequencing methods between vaginal
353 and rectal isolates [52,75,76] as well as similar expression of virulence factors [77]. The
354 proportion of participants we screened with a history of RVVC who were positive for both
355 vaginal and rectal yeast was similar to previous culture-based studies (rectal colonization in
356 tandem with vaginal colonization between 45-86% [35–38,78]).

357 We found no evidence for selection acting differently at the two sites. This could be due
358 to similar selective pressures in both environments (leading to selection for the same traits) or
359 could indicate that selection cannot overcome migration, potentially coupled with drift due to low
360 effective population size. The isolates come from individuals with a long history of repeatedly
361 taking antifungal drugs, and the yeast population sizes likely undergo many bottleneck cycles
362 over time, functionally reducing the efficacy of selection. Nevertheless, we did observe 100s of
363 SNP differences between all isolate pairs, indicating the presence of what has been termed
364 microvariation, and could be of the magnitude sufficient for adaptation. These results

365 underscore the need for further investigations to understand the relationship between
366 populations at different body sites and over time in individuals with RVVC.
367 The observed average nucleotide diversity for the three RVVC *C. albicans* populations
368 was similar to what we calculated from previously sequenced commensal oral and rectal
369 isolates from seven healthy individuals, yet lower than two commensal oral populations that
370 were previously known to have isolates from different clades [30,58]. The only *N. glabratus*
371 study we could find that examined intra-population diversity was nine individuals with
372 bloodstream infections. The RVVC *N. glabratus* population had similar diversity to four BSI
373 populations yet was much higher than the other five [59]. It was recently shown that *N. glabratus*
374 isolates causing bloodstream infections were closely related to isolates from commensal sites in
375 the same individuals [79]. Overall, nucleotide diversity in RVVC populations was higher than we
376 had anticipated. Of note, we chose isolates for sequencing blind to the phenotypic data to
377 provide an unbiased estimate of diversity. Although it is tempting in experimental work when
378 resources are limited to always pick the most diverse isolates for characterization, it makes it
379 more difficult to compare results among studies [63]. The level of diversity we uncovered
380 suggests that standing genetic variation may be maintained within the vaginal yeast population,
381 particularly since the available comparable WGS studies purposefully selected isolates that
382 maximized phenotypical differences (i.e., [30,59]). Interestingly, the presence of diversity is
383 consistent with one of the earliest genetic studies, which used DNA fingerprinting to examine
384 diversity at a single time point in up to 14 vaginal isolates from six RVVC populations [80].
385 As whole genome sequencing becomes ever more routine, we hope additional studies
386 will examine intra-population diversity. Our work establishes the importance of sample size in
387 comparing diversity among studies in *C. albicans*. To our knowledge, our intra-population WGS
388 characterization of up to 12 isolates at two body sites is the largest number of isolates examined
389 to date in any context for *C. albicans* or *N. glabratus*. Previous work in other taxonomic groups
390 suggested that a sample size of four to eight isolates is sufficient to accurately estimate
391 population genetics parameters such as expected heterozygosity, observed heterozygosity and
392 pairwise genetic differentiation (i.e., F_{ST}) [60,62,63]. To conduct our analysis in the context of
393 microbial fungi, we subsampled from the 12 isolates and conducted a bootstrap analysis of
394 nucleotide diversity. From our results, we suggest that other studies interested in calculating
395 diversity should sequence six isolates per population, given the tradeoff between precision
396 (more isolates will always be better) and cost. Our results also highlight that if the goal is to
397 compare among studies, it is vital to ensure the same number of isolates are used for diversity
398 calculations. Through the use of pixy to calculate nucleotide diversity, results among studies

399 should be more comparable than other methods, as pixy has been shown to better deal with
400 bias that results from the implicit assumption that missing genotypes, including those at sites not
401 represented in the VCF, are homozygous for the reference allele [81]. We hope that our work
402 provides a baseline genetic diversity that will be used in the future to contextualize what
403 happens to the vaginal yeast population between symptomatic episodes and to generally act as
404 a launching point in studies comparing genetic diversity among different body sites and
405 research aims.
406 The use of the WGS enables us to highlight that isolates from the same population
407 assigned the same MLST can still differ by 100s of SNPs. Furthermore, our work also highlights
408 that ST groups are not always monophyletic. For example, we found two non-ST16 *N. glabratus*
409 isolates are nested within the ST16 clade. Hence, conclusions based on MLST should place
410 their results into a phylogenetic framework to assess relatedness rather than relying on ST
411 designation alone, and studies relying on MLST should ensure nuance in making statements
412 about the relatedness of isolates. Although not always feasible, WGS removes the need to rely
413 on a single or small number of markers, which could over or underinflate the actual level of
414 diversity, as this demonstrates.
415 Whole-genome sequencing enabled us to quantify differences among isolates in SNPs
416 as well as chromosomal aneuploidy, CNVs, and LOH. Non-SNP mutations have a higher
417 mutation rate than the per-base pair substitution rate [82–90]. Aneuploidy has been commonly
418 observed in yeast populations evolving to drugs in the lab [90,91] and among clinical isolates
419 [89,92]. Variation in LOH has similarly been found to be prevalent among closely related
420 isolates from other contexts in *C. albicans*, e.g., serial isolates from HIV patients with
421 oropharyngeal candidiasis [93], among oral commensal isolates from healthy individuals [58],
422 among commensal isolates from the digestive tract of healthy individuals [94], among
423 populations evolved in the presence of *C. elegans* [95], and among experimentally evolved *S.*
424 *cerevisiae* populations in many environments (e.g., rich medium, beer wort, wine must, high
425 salt, minimal medium, [96,97]). Given that RVVC individuals take drugs repeatedly [98], we
426 expected instances of aneuploidies, CNVs and LOH among the isolates in this study.
427 Interestingly, we did not observe any aneuploidy or well-supported CNV within our isolates. We
428 did, however, see fairly extensive variation in LOH when comparing among populations, even
429 though they were all from the same subclade, highlighting inter-population diversification despite
430 being in the same clade. Each of the populations had unique LOH tracks as well as shared LOH
431 tracks. The only population with variation in LOH was TVY10, where only some of the isolates
432 carried an LOH on chr1L. Future work will determine whether there is evidence for any unique

433 LOH signatures common among RVVC vaginal isolates, possibly pinpointing regions under
434 selection.
435 All three *C. albicans* RVVC populations were in a subgroup in the global phylogeny in
436 clade 1. This subgroup is enriched for vaginal isolates compared to the entire tree and even
437 compared to the other part of clade 1. Compared to different clades, a greater proportion of
438 clade 1 isolates have previously been noted to be significantly associated with superficial
439 infections [99], including in the context of VVC [17,55,56,100]. Although vaginal isolates can be
440 found throughout the phylogenetic tree, there may be something unique to the common
441 ancestor of the clade 1 subgroup that makes them more amenable to colonizing and invading
442 epithelial surfaces in general and hence able to cause vaginal diseases [101,102]. Most
443 phenotypes of potential clinical interest have previously been found to vary among isolates
444 within the same clade, precluding clear phenotype × clade associations (e.g., [103,104]).
445 However, it may be that more fine-scale phylogenetic resolution is required to tease apart
446 relationships; if only a subgroup of clade 1 isolates is enriched for a particular phenotype, that
447 might not be seen if all clade 1 isolates are grouped together. Sala *et al.* recently found that
448 VVC isolates induced greater fungal shedding from epithelial cells and differently stimulated
449 epithelial signaling pathways compared to isolates from healthy women [105]. This is the
450 clearest *in vitro* assay able to differentiate VVC and healthy isolates phenotypically, and hence,
451 a strong target for a GWAS analysis to potentially pinpoint the genetic basis of this seemingly
452 important trait. It will be of great interest in the future to determine whether there is a genotypic
453 association between common variants in the subgroup of clade 1 isolates (including isolates
454 from other body sites) and their interaction with vaginal epithelial cells.
455 Significantly less work has been done to examine *N. glabratus* in the context of R/VVC
456 compared to *C. albicans*, despite the increasing incidence of *N. glabratus* globally as an
457 etiological agent of R/VVC [18,106]. We found only a single vaginal isolate out of 526 total
458 isolates that have been whole genome sequenced. This sharply contrasts with *C. albicans*,
459 where 35 % of all sequenced isolates in the current phylogeny have been annotated as vaginal
460 origin [29]. With only two independent vaginal isolates (since the RVVC isolates from our study
461 are not independent, we only count them as one), this extremely low number precludes making
462 any conclusions. There are also only two vaginal isolates in the *N. glabratus* MLST database
463 [107] out of over 2000 isolates. Interestingly, they all belong to ST16, which is also the ST of the
464 RVVC *N. glabratus* population we analyzed. All three are from North America; ST16 is the most
465 common group in the US [79] and is seldom found in other areas such as Australia and Asia
466 [108–111]. Although there is ongoing debate about the utility of renaming *N. glabratus* away

467 from *C. glabrata* [112], the closest common ancestor is from over 100 million years ago [113].
468 *N. glabratus* exhibits different pathogenesis and immunological responses compared to *C.*
469 *albicans* [18,114,115] and differs in the approach to inactivating non-beneficial nutrients, such
470 as excess copper and zinc [116–118]. Given the long evolutionary divergence and biological
471 differences, it may be that lumping these species together under the colloquial vaginal "yeast
472 infection" downplays differences that could have disease implications [115].
473 We have conducted the most extensive study to date that employed whole genome
474 sequencing, modern methods for calculating and comparing diversity, and high throughput
475 phenotypic analyses to compare vaginal and rectal isolates from symptomatic RVVC infections.
476 We find no evidence that rectal isolates are different than vaginal isolates, which is inconsistent
477 with the hypothesis that the GI tract is a source population for vaginal reinfection. We observed
478 a near-identical average nucleotide diversity between our populations and some populations
479 from commensal (*C. albicans*) and BSI (*N. glabratus*) settings. It remains unanswered how the
480 nucleotide diversity is kept at such elevated levels despite repeated use of antifungals by
481 individuals with RVVC. Combined, this study provides relevant information on the baseline
482 diversity of RVVC vaginal and rectal populations. It emphasizes the need to investigate the role
483 of other body sites in shaping fungal microbial communities across various contexts.

484 **Methods**

485

486 **Clinical isolates**

487 Clinical samples were collected from four non-pregnant female participants between the
488 ages of 18-50 years with a history of recurrent vulvovaginal candidiasis (RVVC). Seventeen
489 consenting female participants who attended a clinic for individuals with a history of RVVC in
490 Winnipeg, Canada, were sampled at the clinic for possible inclusion in this study. Vaginal and
491 rectal swabs were acquired from all participants. Swabs were kept at -4 ℃ and on ice during
492 transport and processed within 5 h of acquisition. Swabs were then agitated for ~30 s in 1 mL
493 PBS, and a dilution series (1, 1:10, 1:100) was conducted. 100 µL from each of the three
494 dilutions was spread onto SDA and chromogenic *Candida* agar plates. Plates were incubated
495 for 48 h at 30 °C, at which time 12 vaginal and 12 rectal isolates were haphazardly isolated
496 from dilution plates that had clear margins between colonies. Colonies were suspended in 1 mL
497 of 20% glycerol and kept at -70 °C. We collectively refer to all colonies collected in this study as
498 "THRIVE-yeast" isolates. Of the seventeen participants, vaginal colonies were acquired from
499 ten. Of those, six were also positive for rectal isolates; we recovered less than ten rectal
500 colonies from the two participants who were rectal swab positive, leaving us with samples from
501 four participants for this project. This study has been approved by the University of Manitoba
502 Biomedical Research Ethics Board (HS24769 (B2021:026)) and Shared Health (SH2021:038).

503 **DNA extraction and sequencing**

504 Genomic DNA was extracted from 24 single colonies (12 each from the vagina and
505 rectum) from one individual with an *N. glabratus* infection (YST6) and three individuals with *C.*
506 *albicans* infections (YST7, TVY4, TVY10). We followed a standard phenol-chloroform protocol
507 as previously described [91]. DNA quality and concentration were assessed on a Thermo
508 Scientific™ NanoDrop 2000 and Qubit® 2.0 Fluorometer, respectively. Genomic DNA was sent
509 to either Microbial Genome Sequencing Center ("MIGS," Pittsburgh, USA; YST6 and YST7) or
510 SeqCoast Genomics (New Hampshire Ave., USA; TVY4 and TVY10) for sequencing. At MIGS,
511 sample libraries were prepared using the Illumina DNA Prep kit and IDT 10bp UDI indices and
512 sequenced on NextSeq 2000 using a 300-cycle flow cell kit, producing 2×151bp reads. The bcl-
513 convert v3.9.3 software was used to assess read quality, demultiplex and trim adapter
514 sequences. At SeqCoast Genomics, samples were prepared using an Illumina DNA Prep
515 tagmentation kit and unique dual indexes. Sequencing was performed on the Illumina
516 NextSeq2000 platform using a 300-cycle flow cell kit to produce 2×150bp paired reads.

517 DRAGEN v3.10.11 was used to assess read quality, demultiplex and trim adapter sequences.
518 The reads have been deposited at the National Center for Biotechnology Information (NCBI)
519 Sequence Read Archive under BioProject ID PRJNA991137.
520 In addition to the 93 genomes we sequenced, we downloaded an additional 182 *C.*
521 *albicans* FASTQ files from the NCBI Sequence Read Archive database [119] from BioProject
522 Accession PRJNA432884 [29] and 526 *N. glabratus* FASTQ files from 19 different projects on
523 the SRA database (assessed on February 05, 2022), including 99 *N. glabratus* FASTQ files
524 from SRA PRJNA361477 [67] and PRJNA669061 [66] (see Table S1 and Table S3).
525
526 **Variant calling**
527 The sequence reads were trimmed with Trimmomatic (v0.39) [120] with standard
528 parameters (LEADING: 10, TRAILING: 3, SLIDINGWINDOW:4:15, MINLEN: 31, TOPHRED33,
529 following [121]). Quality was assessed with FASTQC
530 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and MultiQC [122]. *C. albicans*
531 trimmed paired-end reads were mapped using bwa-mem [123] to the SC5314 haplotype A
532 reference genome (A22-s07-m01-r160) downloaded from the Candida Genome Database [124].
533 The resulting SAM file was coordinate-sorted and converted to a Binary Alignment Map (BAM)
534 file using samtools v1.9 [125]. *N. glabratus* isolates were mapped to the CBS138 reference
535 genome (GCA000002545v2) downloaded from the Ensembl Genome Database [126].
536 Alignment quality was assessed with CollectAlignmentSummaryMetrics from Picard v2.26.3
537 (http://broadinstitute.github.io/picard) and consolidated across all samples with MultiQC [122].
538 All files had a >95% mapping quality. BAM files were further processed with Picard by adding a
539 read group annotation so that samples with the same Bioproject ID had the same read group,
540 removing duplicate PCR amplicons, and fixing mate pairs. Base quality scores for the *C.*
541 *albicans* aligned reads were recalibrated with known single-nucleotide polymorphisms obtained
542 from the Candida Genome Database website
543 (http://www.candidagenome.org/download/gff/C_albicans_SC5314/Assembly22/A22_Jones_PM
544 ID_15123810_Polymorphisms.vcf; downloaded on July 29, 2020) [127] using the
545 BaseRecalibrator and ApplyBQSR from the Genome Analysis Toolkit 4.2.4.0. The average
546 coverage for each isolate was estimated using samtools v1.9 [125].
547 The GATK Best Practices were adapted for variant calling. In sequence,
548 HaplotypeCaller, CombineGVCFs, GenotypeVCFs, VariantFiltration, and SelectVariants [128–
549 130] were used to identify single nucleotide variants (SNPs) among all sequenced isolates in
550 diploid and haploid mode for *C. albicans* and *N. glabratus* respectively. The results SNP table

551 was hard filtered using the suggested parameters and to match [29] (QualByDepth <2.0,
552 FisherStrand >60.0, root mean square mapping quality <30.0, MappingQualityRankSumTest <
553 −12.5, ReadPosRankSumTest < −8.0). We excluded variants that were called in known
554 repetitive regions of the genome, as these are likely to reflect sequencing misalignments rather
555 than true variants, i.e., the subtelomeric regions (15kb from the start and end of each
556 chromosome), the centromeres, and the major repeat sequence regions, (Table S4, start and
557 stop positions from candidagenome.org).
558
559 **Phylogeny construction**
560 Phylogenetic trees were constructed for *C. albicans* and *N. glabratus* isolates. For each
561 species, the multi-sample VCF file consisting of only genomic DNA SNPs was converted to a
562 FASTA alignment using a publicly available Python script that creates an alignment matrix for
563 phylogenetic analysis (vcf2phylip.py v2.8, downloaded from
564 https://github.com/edgardomortiz/vcf2phylip) [131]. For heterozygous SNPs in *C. albicans,* the
565 consensus sequence is preferentially made based on the reference (haplotype A) base.
566 Ambiguous bases are written following IUPAC nucleotide ambiguity codes in the matrix. Ploidy
567 is not specified, as the most parsimonious ploidy is detected by the script. The FASTA
568 alignment was parsed in FastTree (2.1.11) [132] in the double precision mode to construct an
569 approximate maximum-likelihood phylogenetic tree using the general time reversible model and
570 the -gamma option to rescale the branch lengths. Compared to the time-consuming ML-based
571 phylogeny predictors such as RAxML [68], FastTree has been found to produce large dataset
572 trees with similar accuracy within a significantly shorter amount of time [133]. The phylogeny
573 was visualized and annotated with the Interactive Tree Of Life (iTOL, v5) [134]. Isolates from *C.*
574 *albicans* clade 13, which is believed to be a closely related but different species termed *C.*
575 *africanus* [64,65,135] was used to root the phylogeny. The *N. glabratus* phylogeny was rooted at
576 the midpoint. While constructing the phylogeny with FastTree with this large data set yielded a
577 phylogeny comparable to *C. albicans* whole-genome phylogenies previously generated by other
578 groups using RAxML [29,136], an attempt to use FastTree to resolve the relatedness of the
579 THRIVE-yeast isolates yielded poor resolution, as most of the SNP sites were ignored during
580 the tree construction. Hence to generate a more-refined phylogeny depicting the relatedness
581 among the isolates from each participant, we constructed a maximum likelihood phylogeny with
582 RAxML v8.2.12 [68] following standard practice using the GTR+G model with 20ML inferences

583 on the alignment, inferring bootstrap replicate trees, applying MRE-based bootstrapping test,
584 and drawing support values using TBE (Transfer Bootstrap Expectation) on the best-scoring
585 tree.
586
587 **Pairwise SNP differences between isolates from the same site in a participant**
588 We also looked at the variation in SNPs among isolates from the same participant within
589 a site. Using RTG tools with the vcfsplit option [137], the VCF files for each isolate were
590 extracted from the multi-sample VCF file. Using "bcftools query -f" the VCF files were converted
591 to bed files. A custom R script was used to do a pairwise comparison of the isolates from a site
592 to determine differences in SNP positions. An ANOVA test was performed to compare pairwise
593 SNP differences in the rectal and vaginal sites in each participant.
594
595 ***In silico* MLST typing**
596 As clades in *N. glabratus* are commonly identified and named through MLST analysis
597 ([66,67]), we conducted *in silico* MLST analyses of the YST6 isolates as well as the cluster in
598 which the YST6 isolated are associated with using stringMLST [138] which uses a predefined
599 MLST *N. glabratus* allele library (FKS, LEU2, NMT1, TRP1, UGP1, URA3) in the PubMLST
600 database [139]. The diploid sequence types (DSTs) of the global *C. albicans* isolates were
601 determined in a previous study using a PCR-based method [29]. *In silico* identification of DST
602 from the WGS data was error-prone and unable to replicate known DSTs.
603
604 **Determining segregating genomic blocks within isolates from the same participant**
605 We conducted a principal component analysis of windows of genomic regions that
606 differed among isolates from the same participant. Briefly, used the "templated script" from the
607 R package lostruct (local PCA/population structure, v.0.0.0.9) [140] with run parameters -t
608 (window type): bp, -s (window size): 5000, -npc (number of pcs): 2 and -m (number of MDS
609 coordinates): 2. To check for possible segregation of vaginal and rectal isolates, principal
610 component analysis plots were generated based on the extreme heterogeneous genomic
611 windows across the entire genome.
612
613 **Relationship between average nucleotide diversity (π) and number of samples**
614 We assessed the potential limitation of the sample size using average pairwise diversity
615 differences between all possible isolate pair estimates from Pixy [81] in a randomized
616 bootstrapping framework. We estimated the π for different numbers of samples (n = 2,3,4,6,10)

617 from the same population. Each number (n) was randomly selected from the vaginal samples,
618 and the π was estimated with Pixy [81]. Additionally, n = 12 was generated by randomly
619 selecting 6 samples each from the rectal and vaginal isolates and combined. All random
620 selections were bootstrapped 50 times, and the standard deviation was estimated. Finally, the
621 number of nucleotide diversity of the total number of isolates per participant (i.e. n=23 for
622 TVY10, TV4, YST7 and n=24 for YST6) was also included. We estimated the nucleotide
623 diversity within (π) using Pixy (v1.2.6.beta1) [81]. Briefly, variants were called using
624 GenotypeGVCFs with --all-sites option activated. Vcftools (v0.1.16) [141] was used to filter the
625 variants (with --max-meanDP 500, --min-meanDP 20, --max-missing 0.8). Indels and
626 mitochondrial DNA were excluded (--remove-indels, --not-chr). The nucleotide diversity was
627 then plotted against the number of isolates to assess the value of *n* where near π saturation
628 was achieved.
629
630 **Nucleotide diversity from published isolates from other studies**
631 We compared our nucleotide diversities to instances where multiple isolates have been
632 retrieved in either commensal or disease settings. For *C. albicans*, 3-4 isolates from 6 and 3
633 individuals respectively from Andersen *et al.* [30] and Sitterlé *et al*. [58] were assessed. All the
634 isolates are from commensal settings in oral and rectal sites. For *N. grabratus*, nine to ten isolates
635 from blood cultures in each of the nine patients [59] were compared to the 12 YST6 isolates.
636
637 ***In silico* mating-type locus detection**
638 To determine mating type-like locus in the *C. albicans* isolates (YST7, TVY4 and
639 TVY10), the reads were aligned to both haplotypes and consensus sequences of the MAT locus
640 on chromosome 5 (MATa1 and MATa2 for hapA and MATα1 and MATα2) were extracted. A
641 BLAST search was then conducted to confirm the locus. Similarly, the mating type-like locus of
642 *N. glabratus* (YST6) isolates was determined by determining the consensus sequence for MTL1
643 (MTLalpha1 and MTLalpha2) and MTL3 on chromosome B, and MTL2 on chromosome E and
644 confirming the MTL by a BLAST search.
645
646 **Loss of heterozygosity and copy number analyses**
647 We conducted genome-wide loss of heterozygosity (LOH) and copy number variant
648 (CNV) analyses using the web-based yeast analysis pipeline (Y_{MAP}) [70]. Mitochondrial DNA is
649 by default excluded from the pipeline. Paired-end reads of the *N. glabratus* isolates were
650 uploaded and analyzed to the CBS138 reference genome (CGD: s05-m01-r09). As *N. glabratus*

651 is haploid, the ploidy was set as 1, which by default excludes LOH analyses. For the *C. albicans*
652 isolates, paired-end read data for each isolate was uploaded and analyzed against the SC5314
653 A22-s02-m09-r10 reference genome. Ploidy was left at the default value (two, i.e. diploid), and
654 correction was enabled for GC-content bias and chromosome-end bias. THRIVE-yeast isolates
655 were compared against the SC5314 haplotype map to distinguish between ancestral and newly
656 evolved LOH signatures and determine which allele was retained. CNV profiles were compared
657 to the reference SC5134 isolate as well as closely related isolates in the phylogeny. The
658 genomic elements within observed CNV regions were identified using the "gene/sequence
659 resources" section of the Candida Genome Database (CGD, https://candidagenome.org).
660
661 **Growth Rate Assay**
662 Two separate growth rate assays were conducted to measure growth in RPMI (10.4 %
663 w/v RPMI powder, 1.5 % w/v dextrose, 1.73 % w/v 3-(N-morpholino) propanesulfonic acid
664 (MOPS), adjusted to pH 7 with NaOH tablets) and vaginal simulative medium ("VSM", following
665 [29]: 1.16 % v/v 5 mM NaCl, 3.6 % v/v 0.5 M KOH, 0.0128 % v/v 99% glycerol, 20 % v/v 0.01 M
666 $Ca(OH)_2$, 1.34 % v/v 0.5 M Urea, 6.6 % v/v 0.5 M glucose, 0.67% w/v solid YNB, 0.85 % v/v 2 M
667 acetic acid, 0.192 % v/v lactic acid, adjusted to pH 4.2 with NaOH tablets). For each experiment,
668 5 µL of frozen glycerol stock from all THRIVE-yeast isolates was inoculated in duplicate into 500
669 µL of RPMI or VSM and incubated for 48 h at 37 °C with agitation at 250 rpm. Cultures were
670 then standardized to an optical density (OD) of 0.01 A600 in RPMI or VSM, and 200 µL was
671 transferred into a 96-well round bottom plate and sealed with a Breathe-Easier sealing
672 membrane (Electron Microscopy Sciences, PA, United States). OD_{600} readings were taken by
673 the Epoch plate reader (Biotek) every 15 minutes, with continuous shaking at 37 °C for 48 h.
674 From each well, the maximal growth rate was calculated as the spline with the highest slope
675 using a custom R script written by Dr. Richard Fitzjohn (https://github.com/acgerstein/THRIVE-
676 variation/scripts_real/growthrate_generic.R). The average growth rate between two technical
677 replicates for each isolate in each growth medium was used for visualization and statistical
678 analysis. Statistical outliers were determined through Rosner's test of outliers available through
679 the rosterTest function in the EnvStats R package [142]. For each population, we started with a
680 k value of one (i.e., testing for a single outlier). If that was significant, we increased k by one
681 until no additional outliers were identified.
682

683 **Drug resistance and tolerance**

684 Disk diffusion assays were carried out to measure resistance and tolerance to
685 fluconazole and boric acid at pH 4.2. Standard CLSI M44 guidelines for antifungal disk diffusion
686 susceptibility testing [143] were generally followed except for pH adjustment as appropriate. To
687 initiate each experiment, frozen stock of all THRIVE-yeast was streaked onto Sabouraud
688 dextrose agar (SDA) plates and incubated for 72 h at room temperature. Isolates were then
689 subcultured by streaking one colony onto a fresh SDA plate and incubating at 37 ºC for 24 h.
690 Colonies from each isolate were then suspended in 200 µL of 0.85% saline solution and
691 standardized to an OD_{600} of 0.01 in 1 mL of saline solution. Within 15 minutes of
692 standardization, 100 µL of the culture was spread evenly using autoclaved glass beads on
693 Mueller-Hinton (MH) plates that had been adjusted to a pH of 4.2. Plates were left to dry for 20
694 minutes before placing a single 5 mg boric acid or 25 mg fluconazole disk in the center of the
695 plate. Plates were incubated upside down at 37 ºC. After 48 h, photographs of each plate were
696 taken on a lightbox. Previous work demonstrated consistent drug resistance values at 24 h and
697 48 h. The entire experiment was conducted twice for each isolate × drug × pH combination, with
698 two technical replicates for each of the two biological replicates.

699 The 48 h images were processed in ImageJ as previously described [72]. They were
700 cropped to a uniform size, the colours were inverted, and brightness and contrast were adjusted
701 to maximize the contrast between the white disk and the black background. The adjusted
702 images were run through the diskImageR package [144] for drug resistance (RAD_{20}) and
703 tolerance (FOG_{20}) quantification. Briefly, diskImageR calculates resistance as RAD_{20}, as the
704 radius of the zone of inhibition where growth is reduced by 20% relative to growth on the
705 margins of the plate where there is no drug, and tolerance as FoG_{20}, the fraction of realized
706 growth between RAD_{20} and the disk.

707 A Welch two-sample t-test that did not assume equal variance was used for each
708 participant × drug combination to compare vaginal and rectal isolates. Statistical outliers were
709 determined through Rosner's test of outliers for growth rates. All statistical analysis was done at
710 a type I error rate of 0.05.

711

712 **Invasive growth assay**

713 To examine invasive growth, we revised the methods from [145]. Freezer stock from
714 THRIVE-yeast isolates were streaked onto 20 mL yeast peptone dextrose (YPD) plates (2% w/v
715 peptone, 2% w/v yeast extract, 1.8% w/v agar, 1% w/v glucose, 0.00016% w/v adenine
716 sulphate, 0.00008% w/v uridine, 0.1% v/v of chloramphenicol and ampicillin) and grown for 72 h

717 at room temperature. A single colony was then randomly chosen from each isolate and
718 inoculated into 200 µL YPD. If no single colonies were available, a similar amount of culture
719 from the colony lawn was used. Cultures were standardized to OD_{600} 0.01 in 1 mL of liquid YPD
720 media, then 2 µL of standardized culture was spotted onto the surface of a 20 mL solid YPD
721 plate in a hexagonal pattern for a total of 7 spots per plate (i.e., spotted at each vertex and in
722 the center). Plates were incubated for 96 h at 37 ºC. The surface growth was washed off using
723 distilled water and a photograph was taken in a dark room on a lightbox. Two biological
724 replicates were performed for each isolate.
725 The qualitative amount of invasive growth for each isolate was determined by visual
726 examination of the post-wash photographs. To develop a five-point scale, two different people
727 independently went through the post-wash pictures from YST6 and YST7 and selected two to
728 six representative pictures that fit into five levels of invasive growth (scored as 1-5). The
729 independent selections were then compared, and one image from each person was chosen as
730 the most representative for each level of the scale. Using these as a reference, each isolate was
731 then categorized into the five levels of the scale (0 - no growth/pipette tip indent, 0.25 - pinprick
732 growth, 0.5 - circular growth evident, 0.75 - circular growth with pinprick, 1 - dense growth
733 throughout). The maximum score between the two bio-replicates of each isolate was used for
734 statistical analysis, though the same statistical conclusions were obtained if the mean score was
735 used instead. For each participant, a Wilcoxon rank sum test was used to compare vaginal and
736 rectal isolates.
737
738 **Data availability**
739 FASTQ files generated for this project have been deposited at the National Center for
740 Biotechnology Information (NCBI) Sequence Read Archive under BioProject ID PRJNA991137.
741 All phenotypic data and code required to reproduce figures and statistical analyses are available
742 at https://github.com/acgerstein/THRIVE_yeast-VR. Large files (e.g., BAM, VCF) are available
743 upon request.

Acknowledgements & Funding

We are extremely thankful to the THRIVE-yeast participants. We are grateful to Dr. Alicia Berard, Dr. Adam Burgener, Kenzie Birse and other THRIVE investigators for sharing their knowledge and resources, which were integral in the initiation of this study and thank other members of the THRIVE-yeast research team including study clinicians, laboratory and administrative staff for their many contributions. The study was funded through the Manitoba Medical Service Foundation, a University of Manitoba Faculty of Science Collaborative grant, the Canadian Institute for Advanced Research (CIFAR), and the National Science and Engineering Rearch Council of Canada. ACG acknowledges the support of the CIFAR Azrieli Global Scholars Program and start-up funding from the University of Manitoba. A-RAB and AS were supported by EvoFunPath (NSERC CREATE) fellowships. DAH was funded by a University of Manitoba Research Award and the University of Manitoba Department of Microbiology Fellowship for Education Purposes. AdG was supported by a University of Manitoba Science Students' Association and the Faculty of Science Undergraduate Student Research Award (USRA). BM was supported by a Faculty of Science USRA. Some swabs and Amies buffer tubes were donated by COPAN Italia.

References:

1. Benedict K, Singleton AL, Jackson BR, Molinari NAM. Survey of incidence, lifetime prevalence, and treatment of self-reported vulvovaginal candidiasis, United States, 2020. BMC Womens Health. 2022;22: 147.

2. Yano J, Sobel JD, Nyirjesy P, Sobel R, Williams VL, Yu Q, et al. Current patient perspectives of vulvovaginal candidiasis: incidence, symptoms, management and post-treatment outcomes. BMC Womens Health. 2019;19: 48.

3. Rathod SD, Buffler PA. Highly-cited estimates of the cumulative incidence and recurrence of vulvovaginal candidiasis are inadequately documented. BMC Womens Health. 2014;14: 43.

4. Denning DW, Kneale M, Sobel JD, Rautemaa-Richardson R. Global burden of recurrent vulvovaginal candidiasis: a systematic review. Lancet Infect Dis. 2018;18: e339–e347.

5. Foxman B, Barlow R, D'Arcy H, Gillespie B, Sobel JD. Candida vaginitis: self-reported incidence and associated costs. Sex Transm Dis. 2000;27: 230–235.

6. Donders G, Sziller IO, Paavonen J, Hay P, de Seta F, Bohbot JM, et al. Management of recurrent vulvovaginal candidosis: Narrative review of the literature and European expert panel opinion. Front Cell Infect Microbiol. 2022;12: 934353.

7. Rosati D, Bruno M, Jaeger M, Ten Oever J, Netea MG. Recurrent vulvovaginal candidiasis: An immunological perspective. Microorganisms. 2020;8. doi:10.3390/microorganisms8020144

8. Fidel PL Jr. History and update on host defense against vaginal candidiasis. Am J Reprod Immunol. 2007;57: 2–12.

9. Babula O, Lazdane G, Kroica J, Ledger WJ, Witkin SS. Relation between recurrent vulvovaginal candidiasis, vaginal concentrations of mannose-binding lectin, and a mannose-binding lectin gene polymorphism in Latvian women. Clin Infect Dis. 2003;37: 733–737.

10. Wojitani MDKH, de Aguiar LM, Baracat EC, Linhares IM. Association between mannose-binding lectin and interleukin-1 receptor antagonist gene polymorphisms and recurrent vulvovaginal candidiasis. Arch Gynecol Obstet. 2012;285: 149–153.

11. Rosentul DC, Delsing CE, Jaeger M, Plantinga TS, Oosting M, Costantini I, et al. Gene polymorphisms in pattern recognition receptors and susceptibility to idiopathic recurrent vulvovaginal candidiasis. Front Microbiol. 2014;5: 483.

12. Sobel JD. Management of patients with recurrent vulvovaginal candidiasis. Drugs. 2003;63: 1059–1066.

13. Zhang J-Y, Liu J-H, Liu F-D, Xia Y-H, Wang J, Liu X, et al. Vulvovaginal candidiasis: species distribution, fluconazole resistance and drug efflux pump gene overexpression. Mycoses. 2014;57: 584–591.

14. Shi X-Y, Yang Y-P, Zhang Y, Li W, Wang J-D, Huang W-M, et al. Molecular identification and antifungal susceptibility of 186 Candida isolates from vulvovaginal candidiasis in

799 southern China. J Med Microbiol. 2015;64: 390–393.

800 15. Guzel AB, Ilkit M, Akar T, Burgut R, Demir SC. Evaluation of risk factors in patients with
801 vulvovaginal candidiasis and the value of chromID Candida agar versus CHROMagar
802 Candida for recovery and presumptive identification of vaginal yeast species. Med Mycol.
803 2011;49: 16–25.

804 16. van Schalkwyk J, Yudin MH, INFECTIOUS DISEASE COMMITTEE. Vulvovaginitis:
805 screening for and management of trichomoniasis, vulvovaginal candidiasis, and bacterial
806 vaginosis. J Obstet Gynaecol Can. 2015;37: 266–274.

807 17. Song N, Kan S, Pang Q, Mei H, Zheng H, Li D, et al. A prospective study on vulvovaginal
808 candidiasis: multicentre molecular epidemiology of pathogenic yeasts in China. J Eur Acad
809 Dermatol Venereol. 2022;36: 566–572.

810 18. Kennedy MA, Sobel JD. Vulvovaginal candidiasis caused by non-albicans Candida species:
811 New insights. Curr Infect Dis Rep. 2010;12: 465–470.

812 19. Parazzini F, Di Cintio E, Chiantera V, Guaschino S. Determinants of different *Candida*
813 species infections of the genital tract in women. Eur J Obstet Gynecol Reprod Biol.
814 2000;93: 141–145.

815 20. Richter SS, Galask RP, Messer SA, Hollis RJ, Diekema DJ, Pfaller MA. Antifungal
816 susceptibilities of *Candida* species causing vulvovaginitis and epidemiology of recurrent
817 cases. J Clin Microbiol. 2005;43: 2155–2162.

818 21. Borman AM, Johnson EM. Name changes for fungi of medical importance, 2018 to 2019. J
819 Clin Microbiol. 2021;59. doi:10.1128/JCM.01811-20

820 22. Kidd SE, Abdolrasouli A, Hagen F. Fungal nomenclature: Managing change is the name of
821 the game. Open Forum Infect Dis. 2023;10: ofac559.

822 23. Pfaller MA, Rhomberg PR, Messer SA, Jones RN, Castanheira M. Isavuconazole,
823 micafungin, and 8 comparator antifungal agents' susceptibility profiles for common and
824 uncommon opportunistic fungi collected in 2013: temporal analysis of antifungal drug
825 resistance using CLSI species-specific clinical breakpoints and proposed epidemiological
826 cutoff values. Diagn Microbiol Infect Dis. 2015;82: 303–313.

827 24. Citiulo F, Jacobsen ID, Miramón P, Schild L, Brunke S, Zipfel P, et al. *Candida albicans*
828 scavenges host zinc via Pra1 during endothelial invasion. PLoS Pathog. 2012;8: e1002777.

829 25. Roselletti E, Pericolini E, Nore A, Takacs P, Kozma B, Sala A, et al. Zinc prevents vaginal
830 candidiasis by inhibiting expression of an inflammatory fungal protein. Sci Transl Med.
831 2023;15: eadi3363.

832 26. Drell T, Lillsaar T, Tummeleht L, Simm J, Aaspõllu A, Väin E, et al. Characterization of the
833 vaginal micro- and mycobiome in asymptomatic reproductive-age Estonian women. PLoS
834 One. 2013;8: e54379.

835 27. Beigi RH, Meyn LA, Moore DM, Krohn MA, Hillier SL. Vaginal yeast colonization in
836 nonpregnant women: A longitudinal study. Obstetrics & Gynecology. 2004;104: 926.

837 28. Bullarbo M, Andersch B, Samuelson E, Lindgren A, Kondori N, Baltzer IM. Self-diagnosis of

838 vulvovaginal candidiasis is poor - A comparison of diagnostic methods introducing β-glucan
839 as a complement. Bioenergetics. 2017;06. doi:10.4172/2161-038x.1000202

840 29. Ropars J, Maufrais C, Diogo D, Marcet-Houben M, Perin A, Sertour N, et al. Gene flow
841 contributes to diversification of the major fungal pathogen *Candida albicans*. Nat Commun.
842 2018;9: 2253.

843 30. Anderson FM, Visser ND, Amses KR, Hodgins-Davis A, Weber AM, Metzner KM, et al.
844 *Candida albicans* selection for human commensalism results in substantial within-host
845 diversity without decreasing fitness for invasive disease. PLoS Biol. 2023;21: e3001822.

846 31. Tasić S, Tasić N, Tasić A, Mitrović S. Recurrent genital candidosis of women; consequence
847 of reinfection of relapse. Med Biol. 2002;9. Available: http://dx.doi.org/

848 32. Lanchares JL, Hernández ML. Recurrent vaginal candidiasis changes in etiopathogenical
849 patterns. Int J Gynaecol Obstet. 2000;71 Suppl 1: S29-35.

850 33. Giraldo PC, Rodrigues HM, Gosson AM, Amaral R, Eleutério J, Gonçalves AKS.
851 Vulvovaginitis and the treatment of asymptomatic partners: a systematic review and meta-
852 analysis. J Bras Doenças Sex Transm. 2014;26: 15–20.

853 34. Milne JD, Warnock DW. Effect of simultaneous oral and vaginal treatment on the rate of
854 cure and relapse in vaginal candidosis. Br J Vener Dis. 1979;55: 362–365.

855 35. El-Din SS, Reynolds MT, Ashbee HR, Barton RC, Evans EG. An investigation into the
856 pathogenesis of vulvo-vaginal candidosis. Sex Transm Infect. 2001;77: 179–183.

857 36. Fong IW. The rectal carriage of yeast in patients with vaginal candidiasis. Clin Invest Med.
858 1994;17: 426–431.

859 37. Sobel JD. Recurrent vulvovaginal candidiasis. N Engl J Med. 1986;315: 1455–1458.

860 38. Mårdh P-A, Novikova N, Stukalova E. Colonisation of extragenital sites by Candidain
861 women with recurrent vulvovaginal candiosis. BJOG. 2003;110: 934–937.

862 39. Spinillo A, Nicola S, Colonna L, Marangoni E, Cavanna C, Michelone G. Frequency and
863 significance of drug resistance in vulvovaginal candidiasis. Gynecol Obstet Invest. 1994;38:
864 130–133.

865 40. O'Connor MI, Sobel JD. Epidemiology of recurrent vulvovaginal candidiasis: identification
866 and strain differentiation of *Candida albicans*. J Infect Dis. 1986;154: 358–363.

867 41. Warnock DW, Speller DC, Day JK, Farrell AJ. Resistogram method for differentiation of
868 strains of *Candida albicans*. J Appl Bacteriol. 1979;46: 571–578.

869 42. Odds FC, Abbott AB. A simple system for the presumptive identification of *Candida*
870 *albicans* and differentiation of strains within the species. Sabouraudia. 1980;18: 301–317.

871 43. Stein GE, Sheridan VL, Magee BB, Magee PT. Use of rDNA restriction fragment length
872 polymorphisms to differentiate strains of *Candida albicans* in women with vulvovaginal
873 candidiasis. Diagn Microbiol Infect Dis. 1991;14: 459–464.

874 44. Lockhart SR, Reed BD, Pierson CL, Soll DR. Most frequent scenario for recurrent Candida

875 vaginitis is strain maintenance with “substrain shuffling”: demonstration by sequential DNA
876 fingerprinting with probes Ca3, C1, and CARE2. J Clin Microbiol. 1996;34: 767–777.

877 45. Soll DR, Galask R, Isley S, Rao TV, Stone D, Hicks J, et al. Switching of *Candida albicans*
878 during successive episodes of recurrent vaginitis. J Clin Microbiol. 1989;27: 681–690.

879 46. Mendling W, Pinto De Andrade M, Gutschmidt J, Gantenberg R, Presber W, Schonian G.
880 Strain specificity of yeasts isolated from different locations of women suffering from vaginal
881 candidosis, and their partners Vergleich der Stammspezifitat von Hefepilzen verschiedener
882 Lokalisation bei Frauen mit Vaginalcandidosen. Mycoses. 2000;43: 387–392.

883 47. Vazquez JA, Sobel JD, Demitriou R, Vaishampayan J, Lynch M, Zervos MJ. Karyotyping of
884 *Candida albicans* isolates obtained longitudinally in women with recurrent vulvovaginal
885 candidiasis. J Infect Dis. 1994;170: 1566–1569.

886 48. Amouri I, Sellami H, Abbes S, Hadrich I, Mahfoudh N, Makni H, et al. Microsatellite analysis
887 of Candida isolates from recurrent vulvovaginal candidiasis. J Med Microbiol. 2012;61:
888 1091–1096.

889 49. Sampaio P, Gusmão L, Correia A, Alves C, Rodrigues AG, Pina-Vaz C, et al. New
890 microsatellite multiplex PCR for *Candida albicans* strain typing reveals microevolutionary
891 changes. J Clin Microbiol. 2005;43: 3869–3876.

892 50. Güzel AB, Döğen A, Aydın M, Serin A, Serin MS, Kalkancı A, et al. Genotyping reveals no
893 link between *Candida albicans* genotype and vaginitis severity in Turkish women.
894 Mycopathologia. 2013;175: 287–294.

895 51. Hong N, Lei Y, Chen H, Chen X, Tsui KM, Hu D, et al. Genotyping and drug resistance
896 profile of clinical isolates of *Candida albicans* from vulvovaginal candidiasis in the Eastern
897 China. Mycopathologia. 2022;187: 217–224.

898 52. Araújo Paulo de Medeiros M, Vieira de Melo AP, Gonçalves SS, Milan EP, Chaves GM.
899 Genetic relatedness among vaginal and anal isolates of *Candida albicans* from women with
900 vulvovaginal candidiasis in north-east Brazil. J Med Microbiol. 2014;63: 1436–1445.

901 53. Chong PP, Abdul Hadi SR, Lee YL, Phan CL, Tan BC, Ng KP, et al. Genotyping and drug
902 resistance profile of *Candida* spp. in recurrent and one-off vaginitis, and high association of
903 non-albicans species with non-pregnant status. Infect Genet Evol. 2007;7: 449–456.

904 54. Faria-Gonçalves P, Gaspar C, Oliveira AS, Palmeira-de-Oliveira R, Gonçalves T, Martinez-
905 de-Oliveira J, et al. Evaluation of overtime phenotypic variation of yeasts in chronic
906 vulvovaginal candidosis cases. Med Mycol. 2021;59: 1166–1173.

907 55. Tian J-Y, Yang Y-G, Chen S, Teng Y, Li X-Z. Genetic diversity and molecular epidemiology
908 of *Candida albicans* from vulvovaginal candidiasis patients. Infect Genet Evol. 2021;92:
909 104893.

910 56. Zhu Y, Fang C, Shi Y, Shan Y, Liu X, Liang Y, et al. *Candida albicans* multilocus sequence
911 typing clade I contributes to the clinical phenotype of vulvovaginal candidiasis patients.
912 Front Med. 2022;9: 837536.

913 57. Bradford LL, Chibucos MC, Ma B, Bruno V, Ravel J. Vaginal *Candida* spp. genomes from

914 women with vulvovaginal candidiasis. Pathog Dis. 2017;75: ftx061

915 58. Sitterlé E, Maufrais C, Sertour N, Palayret M, d'Enfert C, Bougnoux M-E. Within-host
916 genomic diversity of *Candida albicans* in healthy carriers. Sci Rep. 2019;9: 2563.

917 59. Badrane H, Cheng S, Dupont CL, Hao B, Driscoll E, Morder K, et al. Genotypic diversity
918 and unrecognized antifungal resistance among populations of *Candida glabrata* from
919 positive blood cultures. Nat Commun. 2023;14: 5918.

920 60. Willing E-M, Dreyer C, van Oosterhout C. Estimates of genetic differentiation measured by
921 F(ST) do not necessarily require large sample sizes when using many SNP markers. PLoS
922 One. 2012;7: e42649.

923 61. Jeffries DL, Copp GH, Lawson Handley L, Olsén KH, Sayer CD, Hänfling B. Comparing
924 RADseq and microsatellites to infer complex phylogeographic patterns, an empirical
925 perspective in the Crucian carp, Carassius carassius, L. Mol Ecol. 2016;25: 2997–3018.

926 62. Nazareno AG, Bemmels JB, Dick CW, Lohmann LG. Minimum sample sizes for population
927 genomics: an empirical study from an Amazonian plant species. Mol Ecol Resour. 2017;17:
928 1136–1147.

929 63. Qu W-M, Liang N, Wu Z-K, Zhao Y-G, Chu D. Minimum sample sizes for invasion
930 genomics: Empirical investigation in an invasive whitefly. Ecol Evol. 2020;10: 38–49.

931 64. Romeo O, Tietz H-J, Criseo G. *Candida africana*: Is it a fungal pathogen? Curr Fungal
932 Infect Rep. 2013;7: 192–197.

933 65. Mixão V, Saus E, Boekhout T, Gabaldón T. Extreme diversification driven by parallel events
934 of massive loss of heterozygosity in the hybrid lineage of *Candida albicans*. Genetics.
935 2021;217. doi:10.1093/genetics/iyaa004

936 66. Helmstetter N, Chybowska AD, Delaney C, Da Silva Dantas A, Gifford H, Wacker T, et al.
937 Population genetics and microevolution of clinical *Candida glabrata* reveals recombinant
938 sequence types and hyper-variation within mitochondrial genomes, virulence genes, and
939 drug targets. Genetics. 2022. doi:10.1093/genetics/iyac031

940 67. Carreté L, Ksiezopolska E, Pegueroles C, Gómez-Molero E, Saus E, Iraola-Guzmán S, et
941 al. Patterns of genomic variation in the opportunistic pathogen *Candida glabrata* suggest
942 the existence of mating and a secondary association with humans. Curr Biol. 2018;28: 15-
943 27.e7.

944 68. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
945 phylogenies. Bioinformatics. 2014;30: 1312–1313.

946 69. Matsutani M, Wakinaka T, Watanabe J, Tokuoka M, Ohnishi A. Comparative genomics of
947 closely related *Tetragenococcus halophilus* strains elucidate the diversity and
948 microevolution of CRISPR elements. Front Microbiol. 2021;12: 687985.

949 70. Abbey DA, Funt J, Lurie-Weinberger MN, Thompson DA, Regev A, Myers CL, et al. YMAP:
950 a pipeline for visualization of copy number variation and loss of heterozygosity in eukaryotic
951 pathogens. Genome Med. 2014;6: 100.

952 71. Pemberton TJ. Identification and comparative analysis of sixteen fungal peptidyl-prolyl

953 cis/trans isomerase repertoires. BMC Genomics. 2006;7: 244.

954 72. Salama OE, Gerstein AC. Differential response of species morphologies and isolates to
955 fluconazole and boric acid. Antimicrob Agents Chemother. 2022;66: e0240621.

956 73. Bhattacharjee P. Epidemiology and antifungal susceptibility of *Candida* species in a tertiary
957 care hospital, Kolkata, India. Curr Med Mycol. 2016;2: 20–27.

958 74. Hashemi SE, Shokohi T, Abastabar M, Aslani N, Ghadamzadeh M, Haghani I. Species
959 distribution and susceptibility profiles of *Candida* species isolated from vulvovaginal
960 candidiasis, emergence of C. lusitaniae. Curr Med Mycol. 2019;5: 26–34.

961 75. Sampaio P, Gusmão L, Alves C, Pina-Vaz C, Amorim A, Pais C. Highly polymorphic
962 microsatellite for identification of *Candida albicans* strains. J Clin Microbiol. 2003;41: 552–
963 557.

964 76. Shi W-M, Mei X-Y, Gao F, Huo K-K, Shen L-L, Qin H-H, et al. Analysis of genital *Candida*
965 *albicans* infection by rapid microsatellite markers genotyping. Chin Med J . 2007;120: 975.

966 77. Araújo Paulo de Medeiros M, Vieira de Melo AP, Maia de Sousa AM, Silva-Rocha WP,
967 Pipolo Milan E, Maranhão Chaves G. Characterization of virulence factors of vaginal and
968 anal isolates of *Candida albicans* sequentially obtained from patients with vulvovaginal
969 candidiasis in north-east Brazil. J Mycol Med. 2017;27: 567–572.

970 78. Spinillo A, Carratta L, Pizzoli G, Lombardi G, Cavanna C, Michelone G, et al. Recurrent
971 vaginal candidiasis. Results of a cohort study of sexual transmission and intestinal
972 reservoir. J Reprod Med. 1992;37: 343–347.

973 79. Lott TJ, Frade JP, Lyon GM, Iqbal N, Lockhart SR. Bloodstream and non-invasive isolates
974 of *Candida glabrata* have similar population structures and fluconazole susceptibilities. Med
975 Mycol. 2012;50: 136–142.

976 80. Lockhart SR, Fritch JJ, Meier AS, Schröppel K, Srikantha T, Galask R, et al. Colonizing
977 populations of *Candida albicans* are clonal in origin but undergo microevolution through C1
978 fragment reorganization as demonstrated by DNA fingerprinting and C1 sequencing. J Clin
979 Microbiol. 1995;33: 1501–1509.

980 81. Korunes KL, Samuk K. pixy: Unbiased estimation of nucleotide diversity and divergence in
981 the presence of missing data. Mol Ecol Resour. 2021;21: 1359–1368.

982 82. Dutta A, Lin G, Pankajam AV, Chakraborty P, Bhat N, Steinmetz LM, et al. Genome
983 Dynamics of hybrid *Saccharomyces cerevisiae* during vegetative and meiotic divisions. G3 .
984 2017;7: 3669–3679.

985 83. Dutta A, Dutreux F, Schacherer J. Loss of heterozygosity results in rapid but variable
986 genome homogenization across yeast genetic backgrounds. Elife. 2021;10: e70339

987 84. Sui Y, Qi L, Wu J-K, Wen X-P, Tang X-X, Ma Z-J, et al. Genome-wide mapping of
988 spontaneous genetic alterations in diploid yeast cells. Proc Natl Acad Sci U S A. 2020;117:
989 28191–28200.

990 85. Sharp NP, Sandell L, James CG, Otto SP. The genome-wide rate and spectrum of
991 spontaneous mutations differ between haploid and diploid yeast. Proc Natl Acad Sci U S A.

992 2018;115: E5046–E5055.

993 86. Zhu YO, Siegal ML, Hall DW, Petrov DA. Precise estimates of mutation rate and spectrum
994 in yeast. Proc Natl Acad Sci U S A. 2014;111: E2310-8.

995 87. Smukowski Heil C. Loss of heterozygosity and its importance in evolution. J Mol Evol.
996 2023;91: 369–377.

997 88. Lynch M, Sung W, Morris K, Coffey N, Landry CR, Dopman EB, et al. A genome-wide view
998 of the spectrum of spontaneous mutations in yeast. Proc Natl Acad Sci U S A. 2008;105:
999 9272–9277.

1000 89. Selmecki A, Gerami-Nejad M, Paulson C, Forche A, Berman J. An isochromosome confers
1001 drug resistance *in vivo* by amplification of two genes, ERG11 and TAC1. Mol Microbiol.
1002 2008;68: 624–641.

1003 90. Todd RT, Selmecki A. Expandable and reversible copy number amplification drives rapid
1004 adaptation to antifungal drugs. Elife. 2020;9: 58349.

1005 91. Kukurudz RJ, Chapel M, Wonitowy Q, Adamu Bukari A-R, Sidney B, Sierhuis R, et al.
1006 Acquisition of cross-azole tolerance and aneuploidy in *Candida albicans* strains evolved to
1007 posaconazole. G3 . 2022;12: jkac156.

1008 92. Yang F, Scopel EFC, Li H, Sun L-L, Kawar N, Cao Y-B, et al. Antifungal Tolerance and
1009 Resistance Emerge at Distinct Drug Concentrations and Rely upon Different Aneuploid
1010 Chromosomes. MBio. 2023;14: e0022723.

1011 93. Ford CB, Funt JM, Abbey D, Issi L, Guiducci C, Martinez DA, et al. The evolution of drug
1012 resistance in clinical isolates of *Candida albicans*. eLife. 2015;4: e00662.

1013 94. Diogo D, Bouchier C, d'Enfert C, Bougnoux M-E. Loss of heterozygosity in commensal
1014 isolates of the asexual diploid yeast *Candida albicans*. Fungal Genet Biol. 2009;46: 159–
1015 168.

1016 95. Smith AC, Morran LT, Hickman MA. Host defense mechanisms induce genome instability
1017 leading to rapid evolution in an opportunistic fungal pathogen. Infect Immun. 2022;90:
1018 e0032821.

1019 96. James TY, Michelotti LA, Glasco AD, Clemons RA, Powers RA, James ES, et al.
1020 Adaptation by loss of heterozygosity in *Saccharomyces cerevisiae* clones under divergent
1021 selection. Genetics. 2019;213: 665–683.

1022 97. Ene IV, Farrer RA, Hirakawa MP, Agwamba K, Cuomo CA, Bennett RJ. Global analysis of
1023 mutations driving microevolution of a heterozygous diploid fungal pathogen. Proc Natl Acad
1024 Sci U S A. 2018;115: E8688–E8697.

1025 98. Pappas PG, Kauffman CA, Andes D, Benjamin DK Jr, Calandra TF, Edwards JE Jr, et al.
1026 Clinical practice guidelines for the management of candidiasis: 2009 update by the
1027 Infectious Diseases Society of America. Clin Infect Dis. 2009;48: 503–535.

1028 99. Odds FC, Bougnoux M-E, Shaw DJ, Bain JM, Davidson AD, Diogo D, et al. Molecular
1029 phylogenetics of *Candida albicans*. Eukaryot Cell. 2007;6: 1041–1052.

1030 100. Ge S-H, Xie J, Xu J, Li J, Li D-M, Zong L-L, et al. Prevalence of specific and
1031 phylogenetically closely related genotypes in the population of *Candida albicans* associated
1032 with genital candidiasis in China. Fungal Genet Biol. 2012;49: 86–93.

1033 101. Schmid J, Herd S, Hunter PR, Cannon RD, Yasin MSM, Samad S, et al. Evidence for a
1034 general-purpose genotype in *Candida albicans*, highly prevalent in multiple geographical
1035 regions, patient types and types of infection. Microbiology. 1999;145 (Pt 9): 2405–2413.

1036 102. Giblin L, Edelmann A, Zhang N, von Maltzahn NB, Cleland SB, Sullivan PA, et al. A DNA
1037 polymorphism specific to *Candida albicans* strains exceptionally successful as human
1038 pathogens. Gene. 2001;272: 157–164.

1039 103. Cravener MV, Do E, May G, Zarnowski R, Andes DR, McManus CJ, et al. Reinforcement
1040 amid genetic diversity in the *Candida albicans* biofilm regulatory network. PLoS Pathog.
1041 2023;19: e1011109.

1042 104. MacCallum DM, Castillo L, Nather K, Munro CA, Brown AJP, Gow NAR, et al. Property
1043 differences among the four major *Candida albicans* strain clades. Eukaryot Cell. 2009;8:
1044 373–387.

1045 105. Sala A, Ardizzoni A, Spaggiari L, Vaidya N, van der Schaaf J, Rizzato C, et al. A new
1046 phenotype in *Candida*-epithelial cell interaction distinguishes colonization- versus
1047 vulvovaginal candidiasis-associated strains. MBio. 2023; e0010723.

1048 106. Makanjuola O, Bongomin F, Fayemiwo SA. An update on the roles of non-albicans
1049 *Candida* species in vulvovaginitis. J Fungi (Basel). 2018;4: 121.

1050 107. Dodgson AR, Pujol C, Denning DW, Soll DR, Fox AJ. Multilocus sequence typing of
1051 *Candida glabrata* reveals geographically enriched clades. J Clin Microbiol. 2003;41: 5709–
1052 5717.

1053 108. Chen P-Y, Huang Y-S, Chuang Y-C, Wang J-T, Sheng W-H, Chen Y-C, et al. Implication of
1054 genotypes for prognosis of *Candida glabrata* bloodstream infections. J Antimicrob
1055 Chemother. 2024;22: dkae200.

1056 109. Won EJ, Choi MJ, Kim M-N, Yong D, Lee WG, Uh Y, et al. Fluconazole-resistant *Candida*
1057 *glabrata* bloodstream isolates, South Korea, 2008-2018. Emerg Infect Dis. 2021;27: 779–
1058 788.

1059 110. Khalifa HO, Arai T, Majima H, Watanabe A, Kamei K. Genetic basis of azole and
1060 echinocandin resistance in clinical *Candida glabrata* in Japan. Antimicrob Agents
1061 Chemother. 2020;64: 10-128.

1062 111. Chew KL, Achik R, Osman NH, Octavia S, Teo JWP. Genomic epidemiology of human
1063 candidaemia isolates in a tertiary hospital. Microb Genom. 2023;9: 001047.

1064 112. de Hoog S, Walsh TJ, Ahmed SA, Alastruey-Izquierdo A, Alexander BD, Arendrup MC, et
1065 al. A conceptual framework for nomenclatural stability and validity of medically important
1066 fungi: a proposed global consensus guideline for fungal name changes supported by ABP,
1067 ASM, CLSI, ECMM, ESCMID-EFISG, EUCAST-AFST, FDLC, IDSA, ISHAM, MMSA, and
1068 MSGERC. J Clin Microbiol. 2023;61: e0087323.

1069 113. Gabaldón T, Carreté L. The birth of a deadly yeast: tracing the evolutionary emergence of
1070 virulence traits in *Candida glabrata*. FEMS Yeast Res. 2016;16: fov110.

1071 114. Galocha M, Pais P, Cavalheiro M, Pereira D, Viana R, Teixeira MC. Divergent Approaches
1072 to Virulence in *C. albicans* and *C. glabrata*: Two Sides of the Same Coin. Int J Mol Sci.
1073 2019;20: 2345.

1074 115. Brunke S, Hube B. Two unlike cousins: *Candida albicans* and *C. glabrata* infection
1075 strategies. Cell Microbiol. 2013;15: 701–708.

1076 116. Thorvaldsen JL, Sewell AK, McCowen CL, Winge DR. Regulation of metallothionein genes
1077 by the ACE1 and AMT1 transcription factors. J Biol Chem. 1993;268: 12512–12518.

1078 117. Weissman Z, Berdicevsky I, Cavari BZ, Kornitzer D. The high copper tolerance of *Candida*
1079 *albicans* is mediated by a P-type ATPase. Proc Natl Acad Sci U S A. 2000;97: 3520–3525.

1080 118. Riedelberger M, Penninger P, Tscherner M, Hadriga B, Brunnhofer C, Jenull S, et al. Type
1081 I interferons ameliorate zinc intoxication of *Candida glabrata* by macrophages and promote
1082 fungal immune evasion. iScience. 2020;23: 101121.

1083 119. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, et al. Database
1084 resources of the national center for biotechnology information. Nucleic Acids Res. 2022;50:
1085 D20–D26.

1086 120. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence
1087 data. Bioinformatics. 2014;30: 2114–2120.

1088 121. Todd RT, Wikoff TD, Forche A, Selmecki A. Genome plasticity in *Candida albicans* is
1089 driven by long repeat sequences. Elife. 2019;8: e45954.

1090 122. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for
1091 multiple tools and samples in a single report. Bioinformatics. 2016;32: 3047–3048.

1092 123. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
1093 arXiv [q-bio.GN]. 2013. Available: http://arxiv.org/abs/1303.3997

1094 124. Skrzypek MS, Binkley J, Binkley G, Miyasato SR, Simison M, Sherlock G. The Candida
1095 Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and
1096 visualization of high throughput sequencing data. Nucleic Acids Res. 2017;45: D592–D596.

1097 125. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence
1098 alignment/map format and SAMtools. Bioinformatics. 2009;25: 2078–2079.

1099 126. Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, et al. Ensembl genomes
1100 2022: an expanding genome resource for non-vertebrates. Nucleic Acids Res. 2022;50:
1101 D996–D1003.

1102 127. Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, et al. The diploid
1103 genome sequence of *Candida albicans*. Proc Natl Acad Sci U S A. 2004;101: 7329–7334.

1104 128. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for
1105 variation discovery and genotyping using next-generation DNA sequencing data. Nat
1106 Genet. 2011;43: 491–498.

1107 129. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et
1108 al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best
1109 practices pipeline. Curr Protoc Bioinformatics. 2013;43: 11.10.1-11.10.33.

1110 130. Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, et
1111 al. Scaling accurate genetic variant discovery to tens of thousands of samples. bioRxiv.
1112 2018. p. 201178. doi:10.1101/201178

1113 131. Ortiz EM. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for
1114 phylogenetic analysis. 2019. doi:10.5281/zenodo.2540861

1115 132. Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for
1116 large alignments. PLoS One. 2010;5: e9490.

1117 133. Liu K, Linder CR, Warnow T. RAxML and FastTree: comparing two methods for large-scale
1118 maximum likelihood phylogeny estimation. PLoS One. 2011;6: e27731.

1119 134. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree
1120 display and annotation. Nucleic Acids Res. 2021;49: W293–W296.

1121 135. Mixão V, Gabaldón T. Genomic evidence for a hybrid origin of the yeast opportunistic
1122 pathogen Candida albicans. BMC Biol. 2020;18: 48.

1123 136. Szarvas J, Rebelo AR, Bortolaia V, Leekitcharoenphon P, Schrøder Hansen D, Nielsen
1124 HL, et al. Danish whole-genome-sequenced *Candida albicans* and *Candida glabrata*
1125 samples fit into globally prevalent clades. J Fungi (Basel). 2021;7. doi:10.3390/jof7110962

1126 137. Cleary JG, Braithwaite R, Gaastra K, Hilbush BS, Inglis S, Irvine SA, et al. Comparing
1127 variant call files for performance benchmarking of next-generation sequencing variant
1128 calling pipelines. bioRxiv. 2015. p. 023754. doi:10.1101/023754

1129 138. Gupta A, Jordan IK, Rishishwar L. stringMLST: a fast k-mer based tool for multilocus
1130 sequence typing. Bioinformatics. 2017;33: 119–121.

1131 139. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb
1132 software, the Pubmlst.org website and their applications. Wellcome Open Res. 2018;3:
1133 124.

1134 140. Li H, Ralph P. Local PCA shows how the effect of population structure differs along the
1135 genome. Genetics. 2019;211: 289–304.

1136 141. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call
1137 format and VCFtools. Bioinformatics. 2011;27: 2156–2158.

1138 142. Millard SP. EnvStats: An R Package for environmental statistics. Springer New York; 2013.

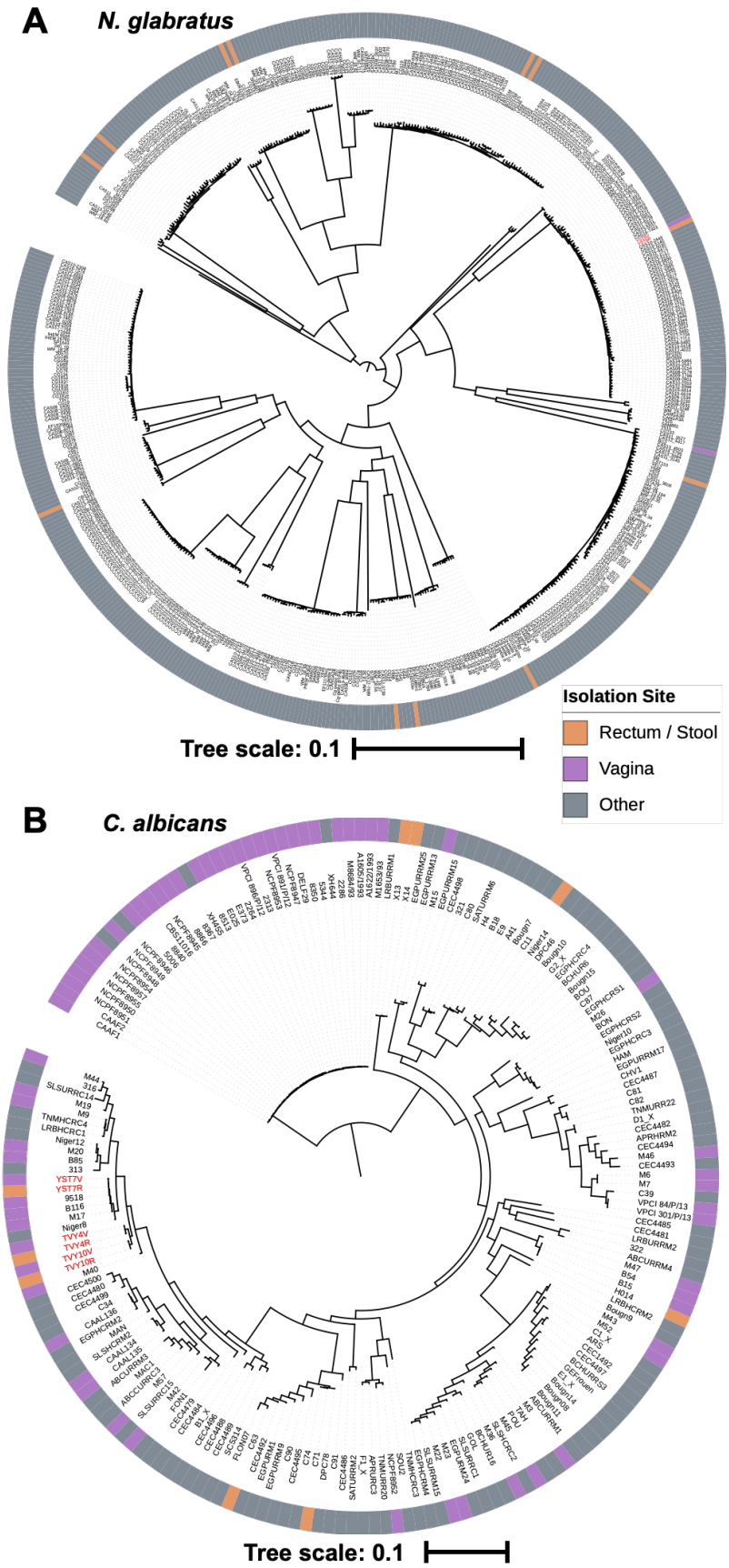
1139 143. NCCLS. Reference method for antifungal disk diffusion susceptibility testing of yeasts;
1140 approved guideline. NCCLS document M44-A National Committee for Clinical Laboratory
1141 Standards. Wayne.

1142 144. Gerstein AC, Rosenberg A, Hecht I, Berman J. diskImageR: quantification of resistance
1143 and tolerance to antimicrobial drugs using disk diffusion assays. Microbiology. 2016;162:
1144 1059–1068.

1145 145. Zupan J, Raspor P. Quantitative agar-invasion assay. J Microbiol Methods. 2008;73: 100–
1146 104.

1147 **Figures**

1148


A
N. glabratus
Tree scale: 0.1
Isolation Site
Rectum / Stool
Vagina
Other
B
C. albicans
Tree scale: 0.1

1149

1150 **Fig. 1** Approximate maximum likelihood phylogenies of (A) *N. glabratus* (n = 526 isolates and
1151 eight THRIVE-yeast isolates) and (B) *C. albicans* (n = 182 isolates and 24 THRIVE-yeast
1152 isolates). THRIVE-yeast isolates are indicated with red labels. The YST6 *N. glabratus* isolates
1153 fall in a cluster of predominantly ST16 isolates (ST16 shown by blue branches, black indicates
1154 not ST16). TVY4, TVY10, and YST7 *C. albicans* isolates fall in clade 1 (blue branches). The
1155 physiological sites of isolation are indicated as a colour strip around the phylogeny, indicating
1156 vaginal (purple), rectal (orange), and all other sites (grey). The *N. glabratus* phylogeny was
1157 rooted at the midpoint, and the *C. albicans* tree was rooted by clade 13 (grey labels).

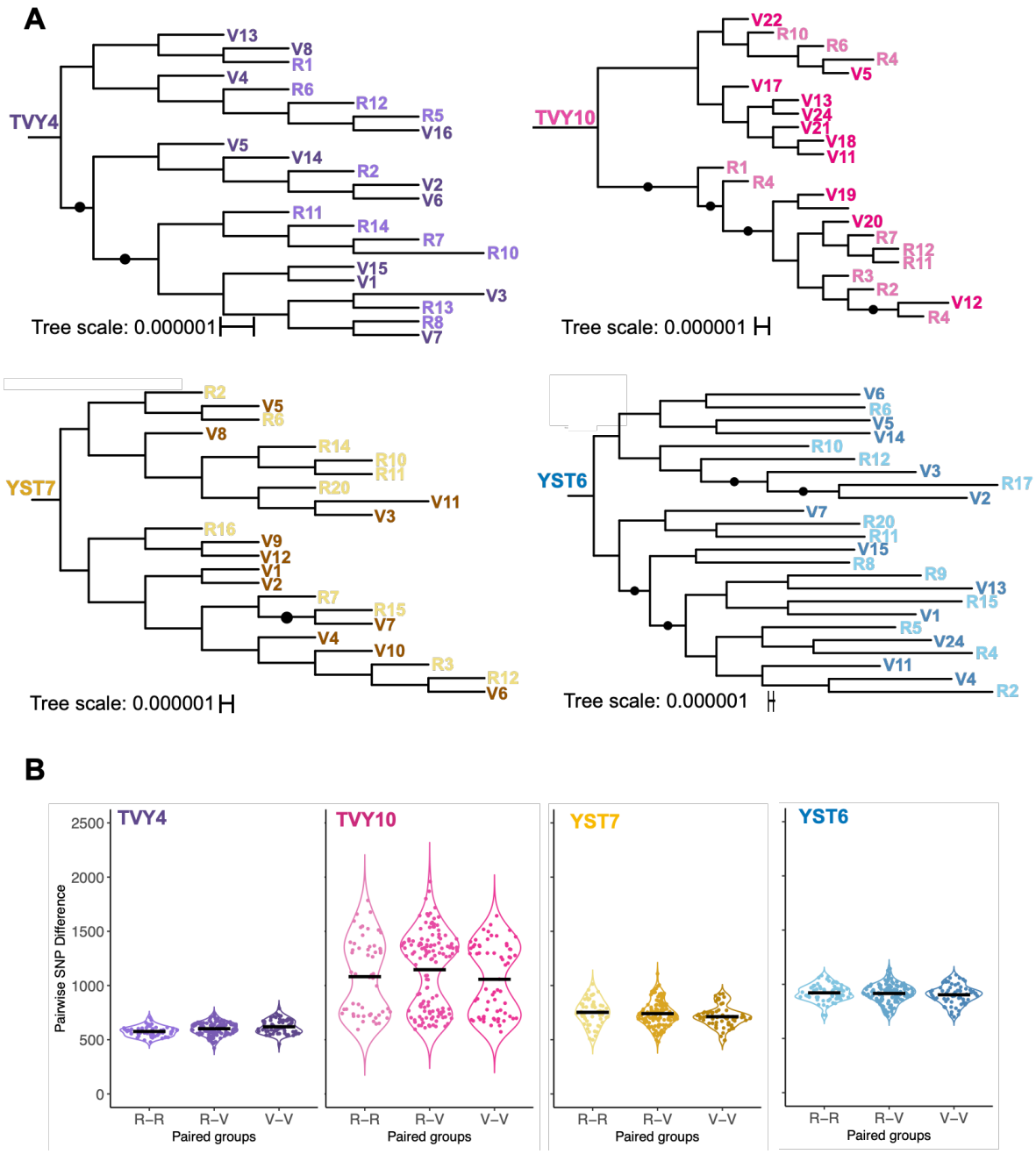


Fig. 2 Within participant phylogenetic and single nucleotide position (SNP) analyses. (A) The fine-scale phylogenetic structure among THRIVE-yeast isolates. Vaginal (V) and rectal (R) isolates were acquired during symptomatic infection from four participants with a history of RVVC (YST6: *N. glabratus*; TVY4, TVY10, YST7: *C. albicans*). Reads were aligned to their respective reference genomes (SC5314 and CBS138), and variants were called using a custom variant calling pipeline adopted from the GATK best practices. Mitochondrial variants, as well as variants in repeat regions, were excluded. Phylogenetic trees were generated with RAxML;

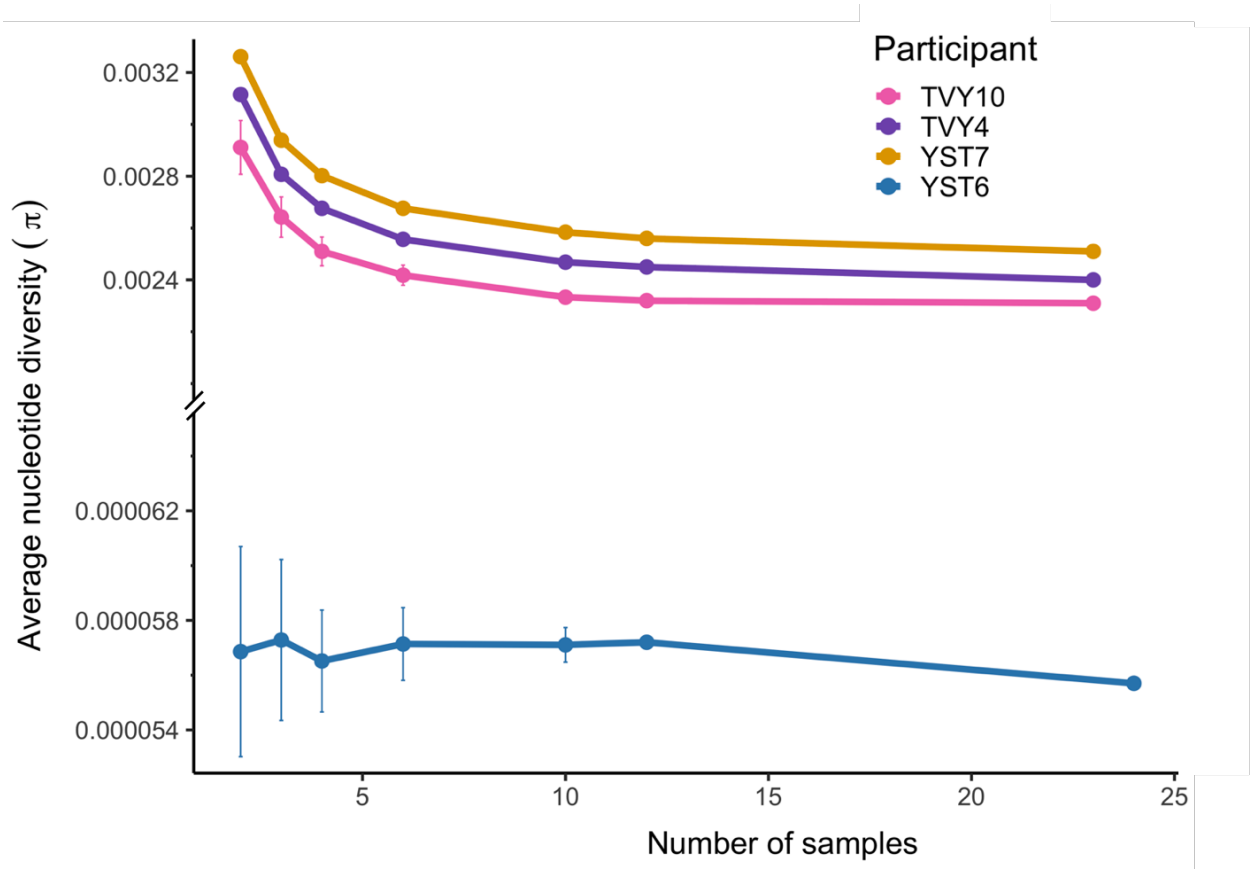
1167 black circles indicate branches with bootstrap support ≥ 0.8. (B) Within-participate pairwise

1168 comparison of single nucleotide positions between isolates from the indicated sites (V-V:

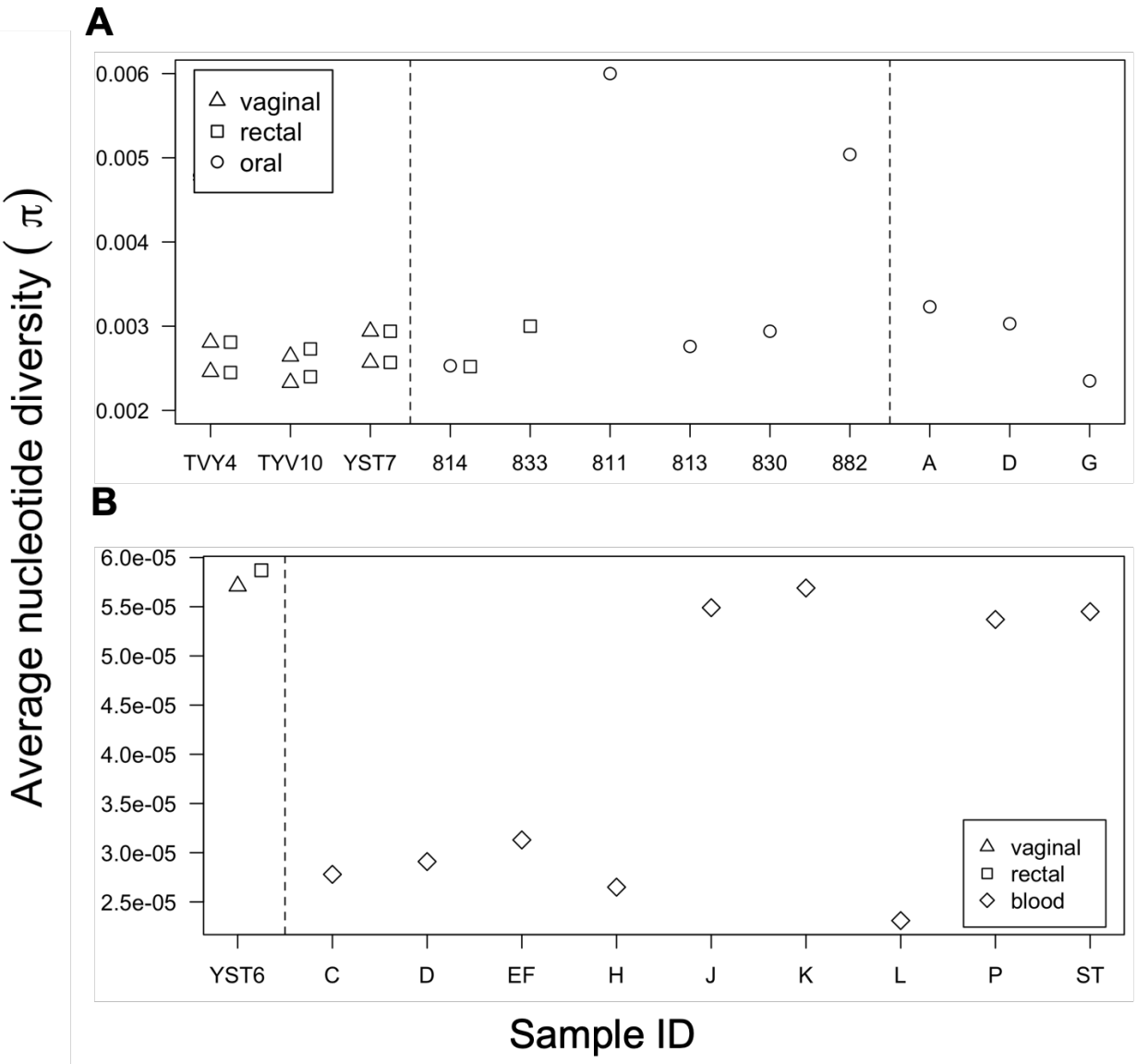
1169 vaginal-vaginal, R-V: rectal-vaginal, R-R: rectal-rectal).

1170

1171



1172
1173
1174 **Fig. 3** Relationship between sample size and average nucleotide diversity. π was calculated for
1175 different numbers of samples from each individual. For n= 2, 3, 4, 6, or 10, the given number of
1176 samples was randomly selected from the vaginal isolate set. For n = 12, the datasets were
1177 generated by randomly choosing 6 samples from each of the rectal and vaginal isolates sets.
1178 The bootstrap analysis was done 50 times for each sample size, and the mean and standard
1179 deviation among data sets were calculated. For n = 23 (*C. albicans*, TVY10, TVY4, YST7) or 24
1180 (*N. glabratus,* YST6), the nucleotide diversity of all samples was calculated.
1181



1182
1183
1184 **Fig. 4** Comparison of the average nucleotide diversity between our samples and isolates from
1185 other studies. (A) Comparison of pi in THRIVE *C. albicans* isolates from TVY4, TVY10, and
1186 YST7 to commensal isolates from two previously published studies. For accurate comparison,
1187 three randomly chosen samples were subsampled from each site in the THRIVE-yeast isolates.
1188 (B) Comparison of pi in THRIVE *N. glabratus* (YST6) isolates to bloodstream infection isolates
1189 from 10 individuals in a previous study.
1190

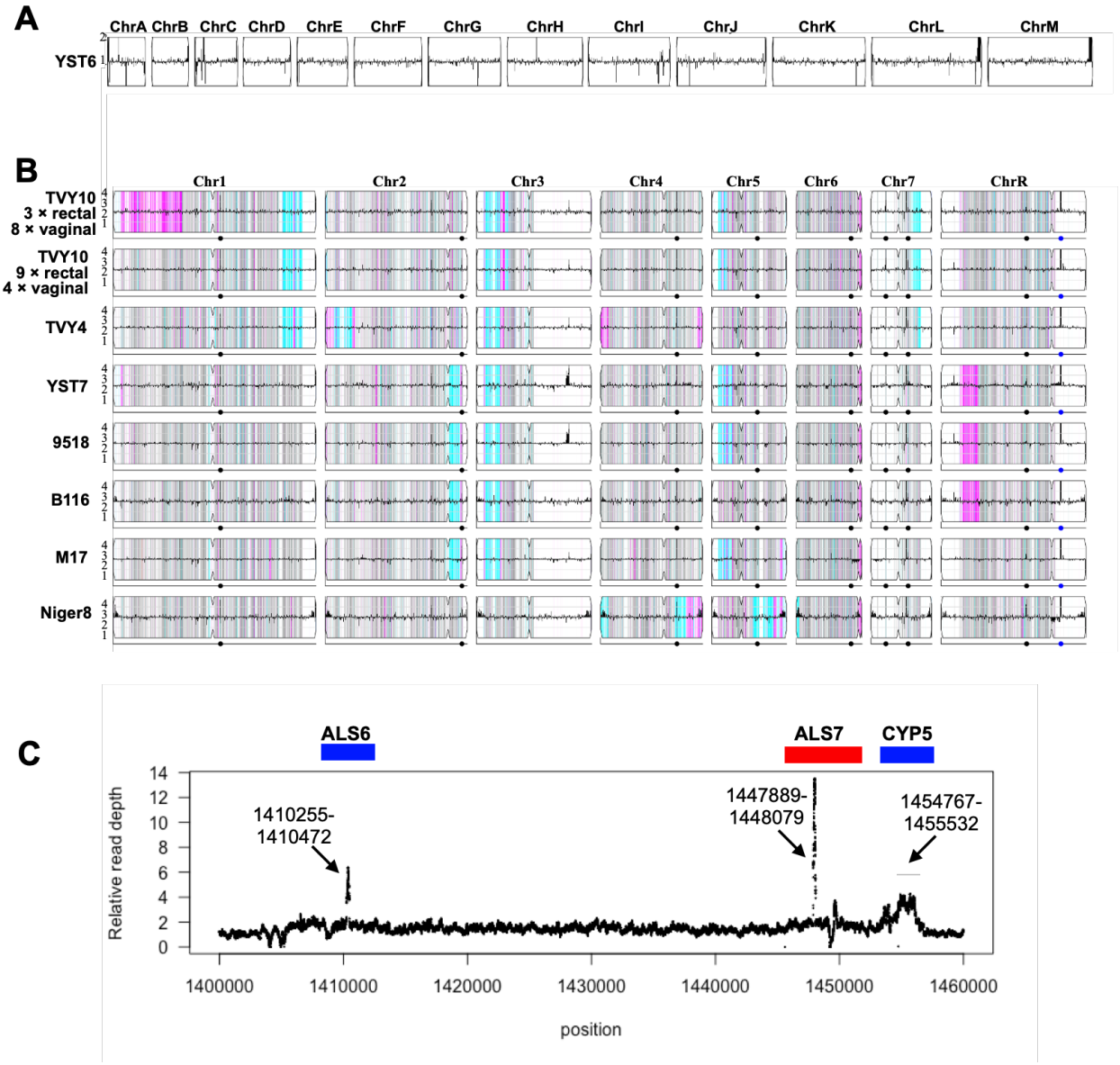


Fig. 5 CNV and LOH profiles of THRIVE-yeast. (A) Representative trace of a YST6 isolate (*N. glabratus*) and (B) TVY10 isolates, representative traces of TVY4 and YST7 isolates, and the four isolates that are closely related to YST7 (three vaginal isolates and Niger8, oral isolate). The relative copy number compared to the reference genomes is shown as the horizontal black line with the scale on the y-axis. For *C. albicans* isolates, the density of heterozygous SNPs is shown as the vertical lines spanning the height of each chromosome box, with the intensity representing the number of SNPs in 5 kb bins. Heterozygous SNPs are gray, homozygous SNPs are coloured based on the retained SC5314 homolog: cyan for “AA” and magenta for “BB.” White indicates an ancestral LOH in SC5314. For each chromosome, the centromere is indicated by an indentation in the box. The dots on the bottom line below each box indicate the

1202 positions of major repeat sequences. (C) Fine-scale coverage mapping of the putative CNV on
1203 chr3. Shown is one representative trace from YST7 R2, all isolates have a similar pattern. Gene
1204 positions above the figure are approximated.
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217

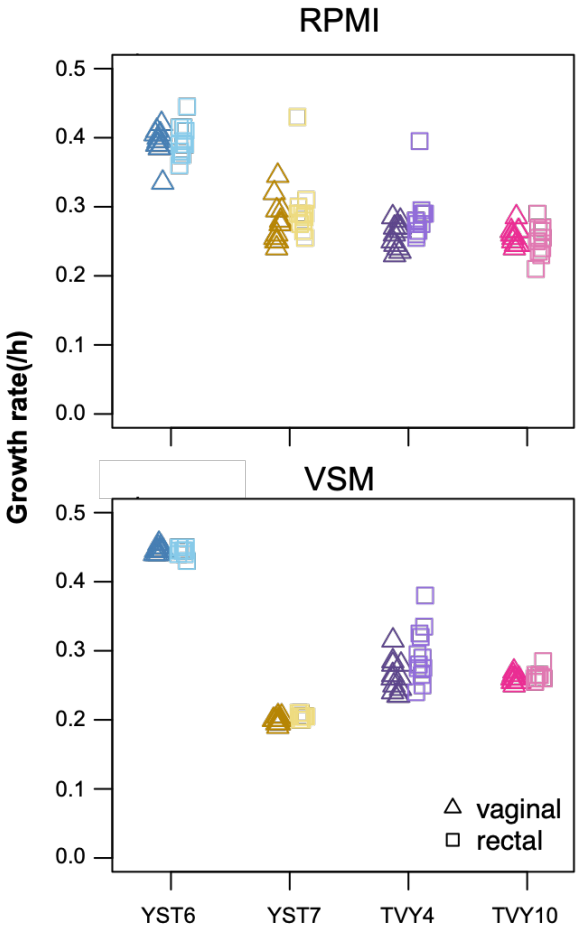
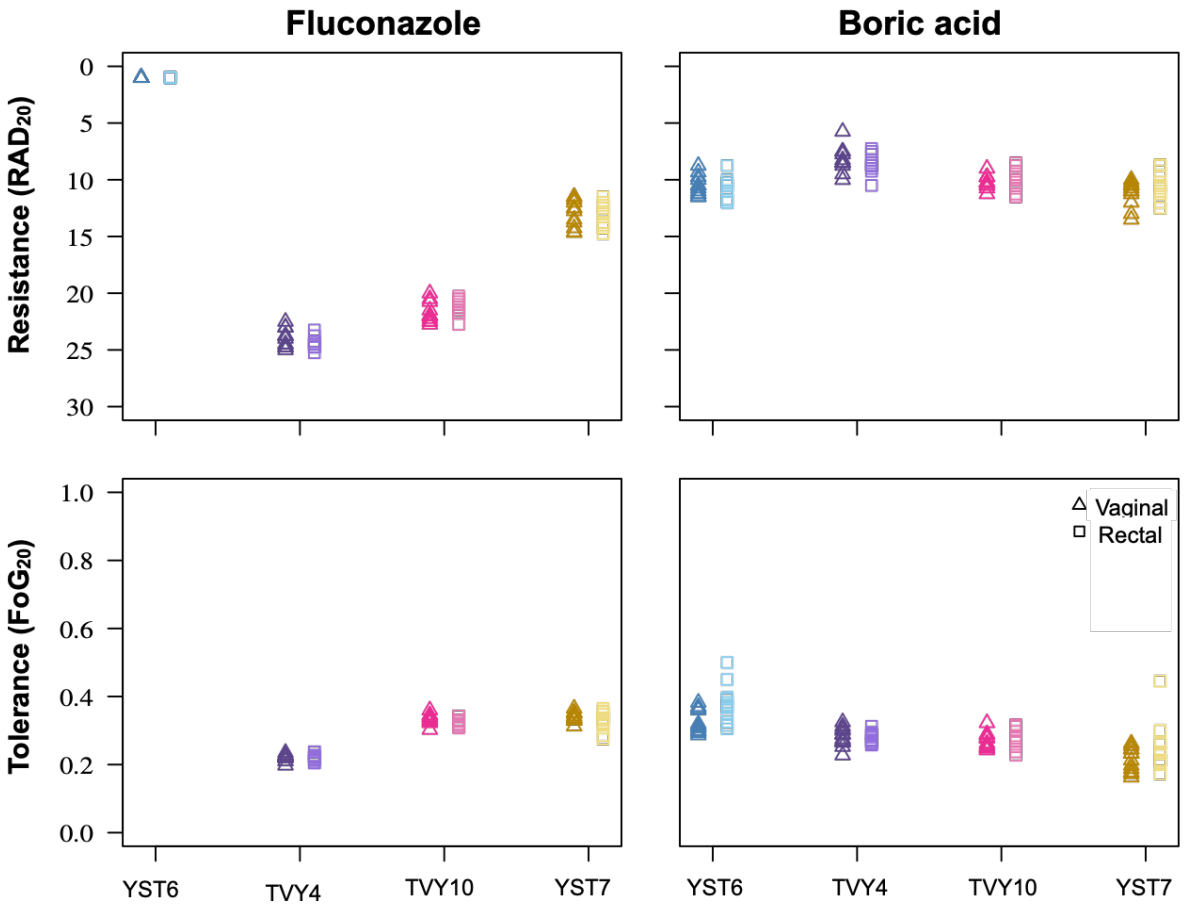


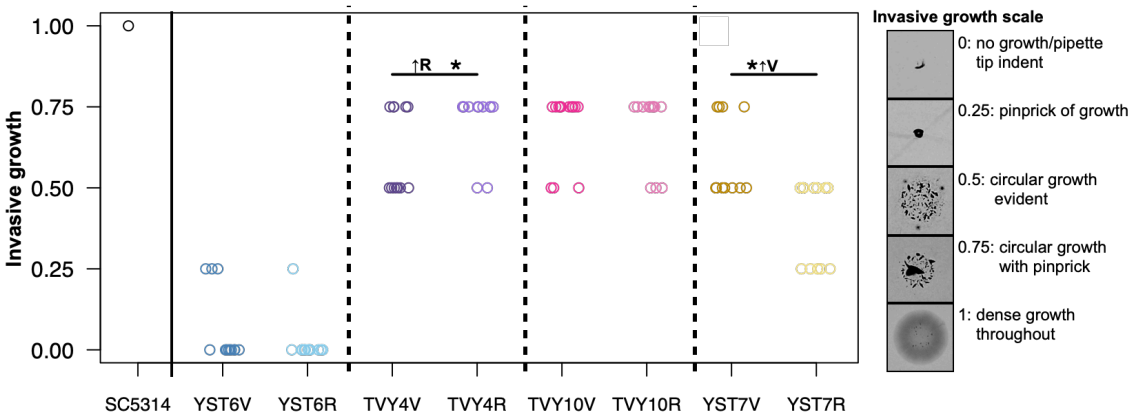
Fig. 6 Growth rate was measured from 12 vaginal and 12 rectal isolates from each population. Optical density was recorded every 15 minutes in a plate reader with constant shaking and incubation at 37 ℃. Each point represents the mean of two technical replicates for each of the two biological isolates, 24 isolates were measured for each group. The growth rate was calculated as the spline with the highest slope using a custom R script.



1224

1225 **Fig. 7** Drug response phenotypes from disk diffusion assays. Drug resistance (top panels) and
1226 drug tolerance (bottom panels) were measured from disk diffusion assays for fluconazole (FLC)
1227 and boric acid (BA). Drug response was measured on pH 4 Mueller-Hinton plates using the R
1228 package *diskImageR* [144], which computationally measures response parameters from images
1229 of disk diffusion assays. Each point represents the mean of four replicates (two technical
1230 replicates for two biological replicates).

1231



1232

1233 **Fig. 8** Invasive growth was qualitatively scored after growth on YPD plates for 96 h. Each point
1234 indicates the maximum score between two bio-replicates for each isolate.
1235