GENOTYPES AND POPULATION GENETICS OF *CRYPTOCOCCUS NEOFORMANS* AND *CRYPTOCOCCUS GATTII* SPECIES COMPLEXES IN EUROPE AND THE MEDITERRANEAN AREA

Massimo Cogliati[1], Marie Desnos-Ollivier[2], Ilka McCormick-Smith[3], Volker Rickerts[3], Kennio Ferreira-Paim[4,5], Wieland Meyer[4], Teun Boekhout[6,7], Ferry Hagen[6], Bart Theelen[6], João Inácio[8], Beatriz Alonso[8,9], Maria F. Colom[10], Luciana Trilles[11], Maria Teresa Montagna[12], Antonella De Donno[13], Serdar Susever[14], Cagri Ergin[15], Aristea Velegraki[16], Mohamed S. Ellabib[17], Simona Nardoni[18], Cristina Macci[19], Laura Trovato[20], Ludovico Dipineto[21], Sevim Akcaglar[22], Emilija Mlinaric-Missoni[23], Sebastien Bertout[24], Ana C.F. Vencá[25], Ana C. Sampaio[26], Giuseppe Criseo[27], Stéphane Ranque[28], Nilgün Çerikçioğlu[29], Anna Marchese[30], Luigi Vezzulli[31], Macit Ilkit[32], Vincenzo Pasquale[33], Itzhack Polacheck[34], and Shawn R. Lockhart[35]

1. Dip. Scienze Biomediche per la Salute, Università degli Studi di Milano, Milano, Italy
2. Institut Pasteur, Molecular Mycology Unit, National Reference Center for Invasive Mycoses & Antifungal, CNRS UMR2000, Paris, France
3. Robert-Koch Institute, Berlin, Germany
4. Molecular Mycology Research Laboratory, Center for Infectious Diseases, Faculty of Medicine and Health, Sydney Medical School, Westmead Clinical School, Marie Bashier Institute for Emerging Infectious Diseases and Biosecurity, University of Sydney, Westmead Hospital, (Research and Educational Network) Westmead Institute for Medical Research, Westmead, NSW, Australia
5. Department of Microbiology, Federal University of Triangulo Mineiro, Uberaba, Minas Gerais, Brazil
6. Westerdijk Fungal Biodiversity Institute, Utrecht, The Netherlands
7. Institute of Biodiversity and Ecosystem Dynamic (IBED), University of Amsterdam, Amsterdam, The Netherlands
8. School of Pharmacy and Biomolecular Sciences, University of Brighton, Brighton, UK
9. Instituto de Investigación Sanitaria Gregorio Marañón (IisGM), Hospital Gegorio Marañón, Madrid, Spain
10. Universidad Miguel Hernández, Alicante, Spain
11. FIOCRUZ, Fundaçao Oswaldo Cruz, Rio de Janeiro, Brazil
12. Università degli Studi di Bari "Aldo Moro", Bari, Italy
13. DiSTeBA, Università del Salento, Lecce, Italy
14. Cyprus Near East University, Nicosia, Cyprus
15. Pamukkale University, Denizli, Turkey
16. Medical School, National and Kapodistrian University of Athens, Athens, Greece
17. Medical College University of Tripoli, Tripoli, Libya
18. Università di Pisa, Pisa, Italy
19. National Research Council, Research Institute on Terrestrial Ecosystems (IRET) Pisa, Italy
20. Università degli Studi di Catania, Catania, Italy
21. University of Napoli Federico II, Napoli, Italy
22. Uludag University, School of Medicine, Bursa, Turkey
23. Croatian National Institute of Public Health, Zagreb, Croatia
24. Unité Mixte Internationale "Recherches Translationnelles sur l'infection à VIH et les Maladies Infectieuses", Université de Montpellier, Montpellier, France
25. Instituto de Higiene e Medicina Tropical, Lisbon, Portugal
26. Universidade de Trás-os-Montes e Alto Douro, CITAB, Quinta dos Prados, Vila Real, Portugal

27. Department of Chemical, Biological, Pharmaceutical and Environmental Sciences, University of Messina, Italy
28. Aix-Marseille University, IRD, APHM, SSA, VITROME, IHU-Méditerranée Infection, Marseille, France
29. Marmara University,School of Medicine, Istanbul, Turkey
30. Sezione di Microbiologia del DISC, Università di Genova-IRCCS Policlinico San Martino Genova, Genova, Italy
31. Dipartimento di Scienze della Terra, dell'Ambiente e della Vita (DISTAV), Università di Genova, Genova, Italy
32. University of Çukurova Sarıçam, Adana, Turkey
33. Università degli Studi di Napoli Parthenope, Napoli, Italy
34. Hadassah-Hebrew University Medical Center, Jerusalem, Israel
35. Centers for Disease Control and Prevention , Atlanta, USA

Corresponding author:

Massimo Cogliati
Lab. Micologia Medica
Dip. Scienze Biomediche per la Salute
Università degli Studi di Milano
Via Pascal 36, 20133 Milano, Italy
Phone: +39 0250315144
Fax: +39 0250315146
E-mail: massimo.cogliati@unimi.it

Running title:
Genotypes of *Cryptococcus* in Europe

Keywords:
Cryptococcus neoformans, *Cryptococcus gattii*, MLST, population genetics, Europe

Conflicts of interests:

All authors declare no conflicts of interests.

ABSTRACT

A total of 476 European isolates (310 *Cryptococcus neoformans* var. *grubii*, 150 *C. neoformans* var. *neoformans*, and 16 *C. gattii* species complex) from both clinical and environmental sources were analyzed by multi-locus sequence typing. Phylogenetic and population genetic analyses were performed. Sequence analysis identified 74 sequence types among *C. neoformans* var. *neoformans* (VNIV), 65 among *C. neoformans* var. *grubii* (56 VNI, 8 VNII, 1 VNB), and 5 among the *C. gattii* species complex (4 VGI and 1 VGIV) isolates. ST23 was the most frequent genotype (22%) among VNI isolates which were mostly grouped in a large clonal cluster including 50% of isolates. Among VNIV isolates, a predominant genotype was not identified. A high percentage of autochthonous STs were identified in both VNI (71%) and VNIV (96%) group of isolates. The 16 European *C. gattii* species complex isolates analyzed in the present study originated all from the environment and all belonged to a large cluster endemic in the Mediterranean area. Population genetic analysis confirmed that VNI group of isolates were characterized by low variability and clonal expansion while VNIV by a higher variability and a number of recombination events. However, when VNI and VNIV environmental isolates were compared, they showed a similar population structure with a high percentage of shared mutations and the absence of fixed mutations. Also linkage disequilibrium analysis reveals differences between clinical and environmental isolates showing a key role of *PLB1* allele combinations in host infection as well as the key role of *LAC1* allele combinations for survival of the fungus in the environment. The present study shows that genetic comparison of clinical and environmental isolates represents a first step to understand the genetic characteristics that cause the shift of some genotypes from a saprophytic to a parasitic life style.

INTRODUCTION

The encapsulated basidiomycetes yeasts belonging to *Cryptococcus neoformans* and *Cryptococcus gattii* species complexes are important pathogens causing cryptococcosis, a fungal infection affecting primarily immunocompromised hosts. Cryptococcosis is acquired from the environment by exposure of the host to sources contaminated by the fungus such as soil, plant materials, and bird excreta. Dehydrated blastospores and basidiospores produced by the fungus are able to reach pulmonary alveoli and may cause a systemic disease by hematogenous dissemination. Cryptococcal yeasts are also able to cross the blood-brain barrier and cause meningitis and menigo-encephalitis that leads to host death unless treated (Kwon-Chung et al. 2014). Both *C. neoformans* and *C. gattii* have been recognized as species complexes each containing several lineages that are object of investigation for their differences in genetic, pathogenic, epidemiological, and ecological characteristics as well as their clinical manifestations. At present, five major molecular types have been recognized within the *C. neoformans* species complex: VNI, VNII, and VNB, all presenting the capsular antigen A (serotype A) and classified as *C. neoformans* var. *grubii*; VNIV presenting the capsular antigen D (serotype D) and classified as *C. neoformans* var. *neoformans*; and VNIII which represents diploid and aneuploid hybrids between the two varieties (AD hybrids). *Cryptococcus gattii* species complex includes at least four molecular types, VGI (serotype B), VGII (serotype B), VGIII (serotype B or C) and VGIV (majority of serotype C) (Meyer et al. 2003, Meyer et al. 2009). Since molecular typing has been performed for only a small percentage of isolates (about 10 %) of the more than 68,000 isolates reported in the literature (Cogliati 2013), it is likely that additional lineages will be found as more strains are typed and their relationship with existing lineages are analyzed. Hence the names "*C. neoformans* species complex" and "*C. gattii* species complex" have been proposed to indicate that both species contain genetically diverse lineages but it is premature to draw boundaries between

lineages to warrant each lineage as separate species. The status of "species complex" was provisional for the sake of nomenclatural stability until the vast majority of cryptococcal populations is analyzed and the character of each lineage becomes readily identifiable (Kwon-Chung et al. 2017). The proposed 'species complex' was in response to an earlier proposal for a revision of cryptococcal taxonomy which was based on the currently identified major lineages (Hagen et al. 2015, Hagen et al. 2017). The proposed revised system classifies, the VNI, VNII and VNB molecular types as *C. neoformans*, VNIV as *C. deneoformans*, VNIII as *C. neoformans* x *C. deneoformans* hybrids; VGI as *C. gattii*, VGII as *C. deteurogattii*, VGIII (serotype B and C) as *C. bacillisporus*, a cluster of VGIII serotype C as *C. decagattii*, and VGIV as *C. tetragattii*. Although this newly proposed classification system has a merit for the progress in the taxonomy of these pathogens within the genus *Cryptococcus*, it is premature as it is already encountering some difficulties revealed by the most recent molecular studies. VNII, VNB, VGIII, and VGIV lineages have been understudied in the past and only recently some results indicate them to be highly heterogeneous. The recent study on VNI, VNII, and VNB *C. neoformans* var. *grubii* isolates by whole genome sequencing showed that VNB could be further separated in two lineages VNBI and VNBII (Rhodes et al. 2017). Similar findings have been obtained for VGIII and VGIV *C. gattii* isolates, showing that these also include different lineages (Cogliati et al. 2012a, Firacative et al. 2016, Nyazika et al. 2016).

In Europe, as well as in most of the high-income countries, the number of cases of cryptococcosis has decreased during the last two decades since the introduction of high active anti-retroviral therapy, which has drastically reduced the number of AIDS patients (ecdc.europa.eu/en/publications-data/hivaids-surveillance-europe-2017-2016-data). On the other hand, the population of patients with other immunocompromising conditions, such as those with transplantation (http://www.transplant-observatory.org/count2), hematological malignancies (Miranda-Filho et al. 2018) or autoimmune diseases

(www.niaid.nih.gov/sites/default/files/adccfinal.pdf), are on the rise. Cases of cryptococcosis in these patients are often difficult to diagnose due to the rarity and the different clinical manifestation of the disease (Pappas 2013, George et al. 2018, Henao-Martìnez et al. 2018, Shuai et al. 2018, Wang et al. 2018, Hamdam et al. 2018, Morse et al. 2018, Fiorentino et al. 2018, Dioverti et al. 2018, Yamada et al. 2018, Swam and Gottlieb 2018, Posteraro et al. 2018). Furthermore, cryptococcosis can also occur in immunocompetent hosts, especially when the infection is caused by *C. gattii* species complex VGI or VGII, making the diagnosis more complicated (Chen et al. 2014). Early diagnosis is critical for effective therapy and patient management while delayed diagnosis could hamper the therapeutic trajectory and increase the risk of death.

In Europe, most cases of cryptococcosis are caused by isolates of the *C. neoformans* species complex with VNI being the most prevalent molecular type, followed by VNIV, and VNIII hybrids (Viviani et al. 2006, Cogliati 2013, Sanchini et al. 2014, Desnos-Ollivier et al. 2015). A higher prevalence of VNIV, and consequently of VNIII hybrids, than in other continents is a peculiar characteristic of *Cryptococcus* epidemiology in Europe (Cogliati 2013). A recent niche modeling analysis suggested that VNIV isolates are well adapted to European climatic conditions characterized by areas with sub-continental climate, which provides the optimal ecological niche for this pathogen (Cogliati et al. 2017). Furthermore, VNIII isolates resulting from hybridization of VNI and VNIV genotypes could play an important role in the evolution of the *C. neoformans* species complex since it represents a mechanism of generating new genetic variability and therefore potentially highly virulent strains (Sun et al. 2009, Shahid et al. 2008, Cogliati et al. 2012b).

Cases of cryptococcosis due to *C. gattii* species complex isolates were also reported in Europe, with VGI being the most prevalent molecular type (Hagen et al. 2012, Cogliati 2013). In addition, some VGII isolates were the cause of infection in cases of cryptococcosis most

probably acquired outside the continent (Hagen et al. 2010, Hagen et al. 2012). In contrast, most infections due to VGI isolates were autochthonous (Hagen et al. 2012) as confirmed by recent European environmental surveys that showed the presence of VGI isolates in live trees and soil (Chowdhary et al. 2012, Colom et al. 2012, Linares et al. 2015, Cogliati et al. 2016a, Montagna et al. 2018). Monitoring the distribution of genotypes of *C. neoformans* and *C. gatti* species complexes in the European and Mediterranean environment is crucial to evaluate the evolution of this pathogen which seems capable to rapidly colonize new habitats. This is illustrated by a VGI and VGII outbreak, which was first detected in 1999 in Vancouver Island (Bartlett et al. 2012) and is now spreading along the Northern Pacific Coast of North America (Engelthaler et al. 2014). The outbreak is mainly due to variants of the VGII genotype (VGIIa, VGIIb, and VGIIc) that were likely imported from South America (Hagen et al. 2013, Billmyre et al. 2014, Engelthaler et al. 2014, Souto et al. 2016, Roe et al. 2018) and has adapted to the temperate climate of the Vancouver Bay (Mak et al. 2010). The ability of VGI and VGII isolates to potentially behave as primary pathogens infecting healthy patients has alarmed public health authorities.

A recent niche modeling study showed that Europe and the non-European Mediterranean area (Northern Africa, Israel, Lebanon, Syria and Turkey) present optimal bioclimatic conditions for the survival in the environment of both *C. neoformans* and *C. gattii* species complexes with several areas where VNI, VNIV, and VGI can coexist (Cogliati et al. 2017). Therefore there is a potential risk for the emergence of new virulent genotypes or the establishment of virulent genotypes from other continents. In 2013 the Screen (**S**urvey of **CR**yptococcosis through a **E**uropean **E**pidemiological **N**etwork) Project, including about 200 European researchers, was created with the purpose of monitoring the epidemiology of *Cryptococcus* and cryptococcosis in Europe through data and isolate collection from clinical and veterinary cases as well as from environmental surveys

(sites.google.com/view/screenprojectcryptococcus). The present study reports the first population genetic analysis carried out integrating all multi-locus sequence typing (MLST) data available for European isolates comparing the different genetic populations in both clinical and environmental isolates.

MATERIALS AND METHODS

Isolates

In the present study 476 isolates of *C. neoformans* and *C. gattii* species complexes from clinical (n=275), veterinary (n=16), and environmental (n=185) sources were investigated. Isolates were from Italy (n=159), France (n=118), Germany (n=96), Turkey (n=41), Greece (n=30), Spain (n=16), Cyprus (n=4), Denmark (n=4), Belgium (n=4), Libya (n=3), and Portugal (n=1) (Table 1). Environmental isolates were recovered mainly from arboreal samples (n=176), eight were from pigeon droppings and one from dust (Cogliati et al. 2016a, Cogliati et al. 2016b). Most of the isolates were mating type α and included 36 mating type **a** isolates (three VNI and 33 VNIV)(Table 1).

MLST analysis

All isolates were genotyped using the MLST scheme standardized by the "ISHAM Working Group for Genotyping of *C. neoformans* and *C. gattii*" (Meyer et al. 2009). Six housekeeping loci (*CAP59*, *GPD1*, *LAC1*, *PLB1*, *SOD1*, and *URA5*) and the IGS1 rDNA intergenic region were amplified and sequenced. Sequences were then compared with those present in the *Cryptococcus* MLST database (www.mycologylab.org) in order to assign an allele number. The combination of the seven-allele numbers of each isolate was also searched in the MLST database and a sequence type (ST) was assigned. Allele types and STs, which were not present

in the MLST database, were sent to the curator of the database in order to assign a new allele or ST code. Most of the environmental isolates (176 isolates) were sequenced in the present study (Supplementary material, Table S1), whereas sequences of clinical isolates from Italy, France, Germany and data from Belgium and Denmark were obtained from previously reported results (Hagen et al. 2012, Cogliati et al. 2013, Danesi et al. 2014, Sanchini et al. 2014, Desnos-Ollivier et al. 2015, Cogliati et al. 2016b). The sequences of each allele type are available at the *Cryptococcus* MLST database and, furthermore, sequences of the new allele types were deposited in GenBank (www.ncbi.nlm.nih.gov). Accession numbers are reported in the supplementary material, Table S1.

Phylogenetic analysis

The concatenated sequences of the seven MLST loci of each isolate were aligned using Clustal Omega (www.ebi.ac.uk). The alignment file was then imported in MEGA software (www.megasoftware.com) for phylogenetic analysis. The maximum likelihood algorithm was used to infer a phylogenetic tree for both *C. neoformans* and *C. gattii* species complex isolates. Additional MLST sequence data from a set of VGI isolates of the *C. gattii* species complex previously reported in Europe were included in the phylogenetic analysis (Hagen et al. 2012). STs of European VGII and VGIII isolates were not included in the phylogenetic tree since they were already analyzed and discussed by Hagen et al. (2012). Statistical analysis was performed by bootstrap analysis with 1000 replicates. The same alignment files were imported into Network 4 software (www.fluxus-engineering.com) to construct a median-joining network for VNI and VNIV isolates. In addition, allele type and ST data were imported in Phyloviz software (www.philoviz.net) to construct a minimum spanning tree, using the goeBurst algorithm, for *C. neoformans* var. *grubii*, *C. neoformans* var. *neoformans* and the *C. gattii* species complex.

Population genetics analysis

To perform population genetics analysis isolates were stratified by different categories: by molecular type (VNI, VNII, VNIV), by source (clinical and environmental/veterinary, ENV), by molecular type and source (clinical VNI, ENV VNI, clinical VNIV, and ENV VNIV), and by country (France, Germany, Greece, Italy, Spain, and Turkey).

Intra- and inter-population genetic variability was estimated by several parameters: number of haplotypes (h), number of segregating sites (S), haplotype diversity (Hd), nucleotide diversity (π), and average number of nucleotide differences (k). The neutrality tests Tajima's D, Fu & Li's D, and Fu & Li's F, were also calculated to verify if populations were at equilibrium or selection mechanisms were occurring. Minimum number of recombination events for each locus in each population was also estimated. In addition, number of fixed mutations, number of shared mutations, and the number of monomorphic mutations in one population and polymorphic in the other, were calculated to estimate genetic divergence between populations. Inter-population comparison of nucleotide diversity was also performed. Finally, a map showing the most probable routes of spreading of European genotypes was inferred comparing the VNI populations of the countries with a number of isolates ≥ 10. Statistical analysis was performed calculating F_{ST} index to verify genetic exchange between VNI populations of different countries. The genotypes which were present in more than one country, were shown in the map as well as the statistical support to justify the spreading from a country to another. All parameters and tests were calculated using DNAsp v5.10.1 software (www.ub.edu/dnasp). Linkage disequilibrium test between pairs of loci was performed by Arlequin software v3.0 (cmpg.unibe.ch/software/arlequin3) to evaluate the level of linkage between loci in a population.

RESULTS

Most of the STs are exclusive of Europe and ST23 is the most common

MLST analysis identified 74 STs among *C. neoformans* var. *neoformans* (VNIV), 65 STs among *C. neoformans* var. *grubii* (56 VNI, 8 VNII, 1 VNB), and 5 STs amongst the *C. gattii* species complex (4 VGI and 1 VGIV) isolates. Whatever the origin, ST23 was the most common genotype (22%) among the VNI isolates followed by ST63 (10%), whereas the other STs showed a lower frequency (0.3-7%) (Table 2). In contrast, among the VNIV isolates, a prevalent genotype was not identified, with the highest frequency observed for ST499 and ST160 (10% and 8%, respectively), followed by ST35 (7%), ST116 (6%), and ST112 (5%). Most of the VNIV STs (58/77, 75%) were represented by only one isolate. ST160 and ST499 were the most frequent genotypes (20% and 26%, respectively) among VNIV ENV isolates, whereas among clinical isolates they showed only 3% and 0% frequency, respectively (Table 3). In addition, VNI isolates included nine STs that were present only among ENV isolates, and 37 STs were present only among clinical isolates. Amongst VNIV isolates 16 STs were exclusive to the ENV group, and 57 to the clinical group. Furthermore, comparing European STs results with those globally reported in the literature we found that 71% of VNI and 96% of VNIV STs analyzed in the present study were exclusive to Europe. Among the 13 VNII isolates eight STs were identified: ST40, ST41, ST42, ST43, ST96, ST106, ST107, and ST490. Finally, the 16 *C. gattii* species complex isolates analyzed in the present study were all from the environment, four isolates from Italy belonged to genotype ST156 (VGI) and ST197 (VGI), six from Spain to ST194 (VGI) and ST367 (VGIV), and six from Greece to ST204 (VGI).

Most of VNI STs are included in a large clonal cluster

Figure 1A shows the four clonal clusters (CCs) identified by goeBurst algorithm among VNI isolates. The main cluster was CC23 including 30 STs and 98 isolates (50%), followed by CC5 with 9 STs and 29 isolates (15%), CC69 with 4 STs and 25 isolates (13%), and CC75 with 3 STs and 16 isolates (8%). The remaining 128 isolates representing 10 STs were not clustered.

Among VNIV isolates only four CCs were identified (Figure 1B): CC135 including 5 STs and 14 isolates (9%), CC160 with 3 STs and 14 isolates (9%), CC499 including one ST and 16 isolates (10%), and CC121 with 5 STs and 14 isolates (9%). Seven CCs including 2 STs each were also identified, whereas others 49 STs were not clustered. In addition, the minimum spanning tree showed that 16% (9/56) of the STs in VNI isolates included both clinical and ENV isolates whereas only 5% (4/77) were present in VNIV isolates ($p<0.05$) (Tables 2 and 3, Figures 1A and 1B).

With respect to the geographical origin, 11 VNI STs were represented by isolates from different countries, including ST63 distributed in six countries, ST58 in five countries, and ST5, ST23, and ST69 in four countries (Figure 1A). On the contrary, among VNIV isolates only two STs (ST160 and ST135) were distributed in three different countries (Figure 1B).

The 15 *C. gattii* VGI isolates clustered in two CCs: CC156 including isolates from Italy and Spain belonging to ST156, ST197, and ST194, and a second cluster including isolates from Greece, which all belonged to ST204 (Figure 1C).

Some STs do not fit in the four major clusters VNI, VNII, VNB, and VNIV

Maximum likelihood phylogenetic reconstruction, performed including the 139 STs of the *C. neoformans* species complex of the present study, identified four major molecular types VNI, VNII, VNB, and VNIV, which were supported by more than 90% of trees produced by bootstrap analysis. However, some STs presented an ambiguous topology either among VNIV (ST496, ST506, ST507, ST508, ST516 and ST520), VNI (ST232, and ST488) or VNII (ST42 and ST107) (Figure 2).

When the four VGI STs identified in the present study were compared to those previously reported in Europe, they were grouped in a large cluster including STs mostly present exclusively in Europe. In this group only ST56 and ST58 included isolates from other continents. Isolates belonging to ST225, ST232, and ST57 belonged to other clusters and were probably imported from other continents. The VGIV isolate belonging to ST367 was unique and was different from any other global isolate so far reported (Figure 3).

Network analysis showed that the VNII group of isolates is not a homogenous group but includes genotypes with high genetic diversity that has some genetic relationship both with VNI and VNB groups (Figure 4). In addition, comparison of network reconstruction performed for VNI and VNIV isolates showed the different genetic evolution of these two groups. VNI isolates displayed numerous large clonal clusters and few dispersed genotypes, in contrast VNIV isolates showed a high number of dispersed genotypes and few small CCs (Figure 4).

High Hd in both VNI and VNIV group of isolates

Table 4 shows the genetic indexes calculated for the different group of isolates investigated in the present study. Comparison of VNI, VNII, and VNIV isolates showed that VNIV and VNI have the highest Hd values which was lower in VNII isolates. In contrast, the π value in VNII isolates was higher than that found in VNI with a k of 12.4 *versus* 8.6, respectively. Neutrality test carried out with three different algorithms showed that only VNIV isolates are under selective pressure. Independent analysis of each of the seven MLST loci showed that there was no polymorphism of *LAC1* in VNII isolates, and a low Hd was also observed for the *GPD1* locus in VNIV and the *SOD1* locus in VNI isolates. The IGS1 locus was the most variable in the three group of isolates with VNIV showing the highest value followed by VNII and VNI (Supplementary material Figure S1A). When the *C. neoformans* species complex isolates were grouped in two categories, clinical and ENV isolates, no differences in Hd, π, and neutrality test were found. A similar result was confirmed by single loci analysis, except for the *GPD1* locus, which showed a low Hd in the ENV isolates (Supplementary material Figure S1B). Stratification of VNI and VNIV by clinical and ENV isolates revealed differences in genetic parameters. The group including clinical VNIV isolates showed the highest Hd, whereas the lowest value was observed in the group including ENV VNI isolates. The other two categories (clinical VNI and ENV VNIV) had a similar Hd intermediate value. A high π value (0.01048) was observed in ENV VNIV isolates, but lower values were calculated for ENV VNI and clinical VNIV isolates, and a very low value (0.00195) was found for the clinical VNI group. Neutrality tests showed that selective

pressure is occurring in all groups of isolates except the clinical VNI group. Differences in Hd were also observed analyzing single loci (Supplementary material Figure S1C). The *GDP1* locus showed a very low Hd values in the ENV VNI group compared to the other ones. The greatest inter-population difference in Hd values was found for the IGS1 locus with the two VNIV group of isolates (clinical and ENV) showing a value ranging from 0.7 to 0.8 and VNI isolates showing a value lower than 0.3. A similar difference, but to a minor extent, was observed in the *URA5* locus where the calculated values were 0.8 for the VNI groups and 0.4-0.6 for VNIV. In both *CAP59* and *SOD1* loci Hd values were decreasing in the following order: ENV VNIV>clinical VNIV>ENV VNI>clinical VNI. In contrast, all groups showed a similar Hd in *LAC1* and *PLB1* loci.

Stratification of isolates by country revealed that isolates from France had the highest Hd and π values, whereas those from Greece had the lowest Hd value but a high π value. The other groups analyzed (Italy, Germany, Spain, and Turkey) showed similar Hd and π values ranging from 0.84 to 0.95 and from 0.026 to 0.041, respectively. Selective pressure was observed in Italy, France, Turkey, and Spain but not in Germany and Greece.

Linkage disequilibrium of PLB1 in VNI isolates and LAC1 in clinical VNIV isolates

Linkage disequilibrium test between pairs of loci performed for the VNI, VNII, and VNIV isolates, showed that in both VNI and VNIV isolates most of loci were linked, but some differences were observed. In the VNI group the *PLB1* locus was linked to each of the other loci, whereas this was not true in the VNIV group. By contrast the *LAC1* and *SOD1* loci showed a high percentage of linkage in VNIV isolates, but not in the VNI group. Pairwise comparison of loci in VNII isolates revealed that most of the loci were not linked to each other (Figure 5).

Stratified analysis by clinical and ENV of VNI isolates revealed that linkage disequilibrium pattern of the ENV group was identical to that of the whole VNI group of isolates, in contrast the clinical VNI group showed mostly non-linked loci, except for *PLB1.* Amongst VNIV isolates,

the clinical and ENV groups showed a difference in the *LAC1* linkage pattern, which was only present in the former group of isolates but not in the latter one (Figure 5).

Environmental isolates are the most recombinants in both VNI and VNIV groups

The minimum number of recombination events, calculated for each group of isolates, is reported in Table 5. Recombination was observed in VNIV isolates but not in VNI and VNII groups. The *LAC1* locus was the most recombinant one whereas no recombination events were found in the *PLB1* locus. Comparison of clinical and ENV isolates showed recombination in both groups with the *LAC1* locus having 21 recombination events in the ENV group *versus* 7 in the clinical one. Both clinical and ENV VNIV isolates had recombination at all loci except *PLB1*, whereas in the VNI isolates only the ENV group showed three recombinant loci (*LAC1*, *PLB1*, and *SOD1*). Finally, recombination events were observed in all countries, except in Spain.

Comparison of environmental VNI and VNIV isolates showed a high rate of shared mutations and no fixed mutations

The first row of graphics in Figure 6 shows the comparison between VNI, VNII, and VNIV isolates. VNIV *versus* VNI and VNII showed a high percentage of fixed mutations and a low percentage of shared mutations confirming the genetic distance between these group of isolates. Furthermore, in both comparisons VNIV showed the highest percentage of polymorphic sites compared to the other two group of isolates. Genetic comparison between VNI and VNII showed a lower percentage of fixed mutations and a higher percentage of shared mutations compared to the previous two analyses, and that VNI was more polymorphic than VNII isolates. Comparing clinical *versus* ENV VNI isolates, and clinical *versus* ENV VNIV isolates, ENV isolates were more polymorphic than clinical ones, especially ENV VNI isolates, and no fixed mutations were observed. The last two comparisons showed the most interesting results. VNI *versus* VNIV clinical groups presented a high percentage of fixed mutations and a lower percentage of shared mutations as expected, with clinical VNIV being more polymorphic than VNI. On the contrary, comparisons between ENV

isolates surprisingly did not present any fixed mutations and a high percentage of shared mutations, with VNI being more polymorphic than VNIV.

Inter-population comparison of nucleotide diversity is shown in Figure 7. Clinical and ENV isolates presented an overlapping nucleotide diversity curve (k = 196 *versus* 173, respectively) (Figure 7A and Table 4). When these two group of isolates were compared by considering only VNI isolates, the difference became evident (k = 6 *versus* 24, respectively) with clinical isolates presenting a lower nucleotide diversity in most of the loci, except *LAC1* and *PLB1*, compared to ENV ones (Figure 7B and Table 4). On the contrary, the nucleotide diversity in VNIV isolates was similar (k = 17 in clinical isolates *versus* 24 in environmental ones, Table 4) with low values of clinical VNIV only for *LAC1* and *SOD1* loci (Figure 7C).

Figure 7D shows a great difference in nucleotide diversity comparing VNI *versus* VNIV isolates (k = 9 and 21, respectively). This difference was not so evident when the two group of isolates (k = 23 in VNI and k = 24 in VNIV) where compared in the ENV group (Figure 7E and Table 4). In the clinical group nucleotide diversity was very low in both group of isolates (k = 6 in VNI and k =17 in VNIV) with the exception of higher values for *GPD1* and IGS1 loci in the VNIV group (Figure 7F and Table 4).

Genetic flow analysis supports the spreading in Europe of the most common VNI STs

Results of genetic flow analysis showed that, in Europe, the main genetic flows, which are supported by F_{ST} statistics, were identified between Italy, Germany, Spain, and Turkey. Analyzing the most frequent STs in Europe, that is ST23, ST63, and ST69, and ST5, all belonging to the VNI molecular type, it was possible to infer the main routes of spreading across the continent (Figure 8). ST23 was found in Italy, Germany, Turkey and France with a spreading direction South to North or *vice versa*. ST63 was the most spreading genotype, presenting an additional directional component towards or from Western Europe. ST69 was primarily confined in central Europe with main genetic flows between Italy and Germany. ST5 spreads mainly from Germany towards Southern and Western Europe.

DISCUSSION

The present study is the largest genetic population analysis carried out in Europe and the Mediterranean area integrating most of the available MLST datasets from clinical and environmental isolates reported in this geographical region. The analyzed collection of isolates included mainly isolates belonging to the *C. neoformans* species complex (VNI, VNII, VNIV and one VNB) and a few environmental *C. gattii* species complex isolates.

The results reported above show a high genetic diversity among the *C. neoformans* species complex isolates here investigated, confirming previous MLST studies carried out on clinical isolates in Italy, France, and Germany (Cogliati et al. 2013, Sanchini et al. 2014, Desnos Ollivier et al. 2015). The high haplotype diversity of VNI isolates is comparable to those calculated for other continents but higher than that observed in Asia (Khayhan et al. 2013, Ferreira-Paim et al. 2017). However, the present study includes more isolates from both clinical and environmental sources. In addition, a large number of clinical and environmental VNIV isolates were included in this study providing new insights about genetic differences and relationships with VNI isolates.

Our results show a major clonal cluster (CC23) being present in Europe including 50% of all VNI isolates with ST23 and ST63 being the most frequent genotypes. By contrast, fewer and smaller CCs are present amongst VNIV isolates with no prevalent genotypes. Interestingly, most of the VNI and VNIV STs (71% and 96%, respectively) seem unique to Europe and absent in other continents suggesting that these genotypes are not yet spreading outside the continent. On the other hand, some VNI STs belonging to CC5, CC69, and CC75 seem to be imported from other continents and are now spreading in Europe throughout Germany, France, and Italy.

Phylogenetic analysis showed that isolates from Europe and the Mediterranean area belonging to the *C. neoformans* species complex are grouped in four clusters corresponding to the molecular types VNI, VNII, VNB, and VNIV, and revealed some isolates that showed an unexpected branch topology suggesting that their genome contains genetic remains of ancestral groups of the *C. neoformans* species complex. These findings are supported by other studies which reported these kind of isolates (Cogliati et al. 2016b, Rhodes et al. 2017, Muñoz et al. 2018,). Further investigation of these isolates is needed to clarify their origin.

Network analysis, supported by haplotype diversity and recombination events values, confirmed the different genetic population structure of VNI and VNIV isolates, the former presenting a primarily clonal structure and the latter a recombinant structure. The same conclusions are reported by previous studies carried out with French clinical isolates and in a number of global VNIV isolates (Desnos-Ollivier et al. 2015, Cogliati et al. 2016b). In the present study genetic features of clinical and environmental isolates were compared for the first time providing new insights about genetic relationships and differences between VNI and VNIV isolates. The obtained results show that VNIV isolates present a higher genetic diversity compared to VNI, and that VNI and VNIV environmental isolates show a higher genetic diversity compared to their respective clinical isolates. These results suggest that in Europe VNIV strains has found favorable environmental conditions and have increased their genetic diversity most likely thanks to sexual recombination due to the presence of both mating type **a** and α in the environment (Cogliati et al. 2013, Cogliati et al. 2016a). In addition, recent studies showed that VNIV strains have a higher probability to be recovered in the areas of Europe characterized by a sub-continental climate due to tolerance of lower temperatures than VNI (Cogliati et al. 2017). Similarly, the higher genetic diversity of ENV VNI compared to clinical VNI could be explained with a residual ability of environmental isolates to sexually recombine. This is confirmed by the analysis of recombination events which were identified

only in the environmental VNI isolates. In addition, the presence of VNI mating type **a** and mating type α isolates in the European environment and their ability to mate were shown previously (Cogliati et al. 2016a).

Interesting conclusions were drawn by the analysis of linkage disequilibrium results and nucleotide diversity indexes comparing ENV and clinical groups in VNI and VNIV isolates. In the clinical VNIV isolates, the *LAC1* locus presents a full linkage with all other loci but this was not observed in ENV VNIV isolates. This means that only some alleles of the *LAC1* locus are represented amongst the clinical isolates, probably due to selective pressure. On the contrary, in ENV isolates, the *LAC1* locus presents a higher variability of combinations that could be explained by the necessity of the fungus to adapt to a wide range of substrates to which it is exposed in the environment. Nucleotide diversity values of the *LAC1* locus also support this hypothesis as π values of clinical VNIV are lower than those observed in ENV VNIV isolates. A recent study on the *LAC1* gene in a set of VNIV strains confirms these results, reporting an association between specific sequence polymorphisms in the *LAC1* gene and the level of melanin production (Samarashinge et al. 2018).

Linkage disequilibrium and nucleotide diversity analyses of VNI isolates identified a significant difference between the clinical and ENV groups. Clinical isolates do not show linkage between any locus, except for the *PLB1* locus, whereas most of the loci in ENV isolates, including the *PLB1* locus, are linked. Therefore, in both group of isolates, the *PLB1* locus showed a high level of linkage suggesting that evolution has selected a restricted set of *PLB1* alleles that have been conserved amongst the environmental isolates and that could play an important role in the parasitic phase of the fungal life cycle. Furthermore, the absence of linkage between the other loci observed in clinical VNI isolates suggests that a multitude of genotype combinations are able to cause infection conversely to what is observed in clinical VNIV isolates where linkage between loci limits the possible combinations. This is also

supported by the significant difference observed between VNI and VNIV in the percentage of STs containing both clinical and environmental isolates (16% *versus* 5%, respectively). These genetic differences between VNI and VNIV isolates could be the basis of their different ability to cause infections, with VNI being the most prevalent molecular type causing cryptococcosis globally (Cogliati 2013). Inter-population comparisons revealed that the major source of polymorphism is the environment for both VNI and VNIV isolates, having 93% and 53% of polymorphic sites, respectively, observed in ENV isolates. This suggests that environmental stress is more effective to generate new genotypes than host condition stress. Our results also showed that, although clinical VNI and clinical VNIV isolates are genetically divergent with 54% fixed mutations, no fixed mutation has been observed by comparing environmental isolates of VNI and VNIV which presented 25% shared mutations. Therefore, ENV VNI and VNIV isolates are less genetically divergent than clinical ones. At present, genetic population studies on *C. neoformans* species complex reporting a high genetic divergence between VNI and VNIV isolates are mostly based on clinical isolates, taking in scarce consideration of environmental isolates. By contrast, the present study carried out a more detailed genetic population analysis of the two group of isolates due to the abundance of VNIV isolates in Europe and a wide environmental survey recently carried out on the continent (Cogliati et al. 2016a) that allowed several environmental isolates to be collected.

Genetic flow analysis of the most prevalent VNI genotypes in Europe showed that four main STs are spreading in Europe (ST23, ST63, ST69, and ST5). Germany and Italy seem to be the main fulcrum of diffusion of both endemic (ST23 and ST63) and imported genotypes of the *C. neoformans* species complex genotypes (ST69 and ST5). More detailed investigations are needed to understand the mechanisms by which this pathogen has crossed continents to and from Europe.

Finally, the 15 environmental *C. gattii* VGI and the unique VGIV isolates investigated in the present study provided further evidence of the endemic presence of these fungi in the Mediterranean area, confirming the results previously reported by Hagen et al. (2012) which identified a cluster containing most of the VGI isolates reported from Europe. Furthermore, our study observed that in Europe ST156 is the most frequent VGI genotype belonging to a wider CC present in Italy and Spain, whereas another CC (CC204) seems to be located in Greece.

Although this study provides new insights about molecular epidemiology and population genetics of *C. neoformans* and *C. gattii* species complexes, it is difficult to give an exhaustive picture of the situation in Europe due to the following limits: a) thousands of *Cryptococcus* isolates have been isolated in Europe but at present only 6% of isolates have been genotyped by MLST (Cogliati 2013); b) the sequenced isolates were reported from only a few European and Mediterranean countries but for most of the other European countries data are lacking; c) about 16% of European isolates are AD hybrids and little is known about their genotypes; d) environmental isolates represent a small percentage (10%) of the total number of isolates available (Cogliati 2013). These limits make this study a preliminary study and should be considered a basis for future investigations. Future efforts to overcome these limitations should be carried out in order to elucidate the real epidemiology of *Cryptococcus* and cryptococcosis in Europe and in the Mediterranean area.

In conclusion, by genetic investigation of environmental populations of *Cryptococcus neoformans* and *Cryptococcus gattii* species complexes, our study has revealed that clinical and environmental isolates in this collection of 467 isolates have different genotypic features. This approach is a first step to understanding the genetic characteristics that cause the shift of some genotypes from a saprophytic to a parasitic life style.

ACKNOWLEDGMENTS

We thank Dr. Kwon Chung and Prof. Dromer for contribution in the critical reading of the manuscript.

The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

W Meyer and S Lockhart were supported by an NH&MRC grant # APP1031943.

AC Sampaio was supported by Fundação para a Ciência e Tecnologia and FEDER/COMPETE/POCI provided to CITAB under the projects POCI-01-0145-FEDER-006958 and UID/AGR/04033/2019.

A Velegraki was supported by SARG K.A.70/4/5905 and 70/3/6915, National and Kapodistrian University of Athens and the Bodosakis Foundation, Athens, Greece.

REFERENCES

Bartlett KH, Cheng PY, Duncan C, Galanis E, Hoang L, Kidd S, Lee MK, Lester S, MacDougall L, Mak S, Morshed M, Taylor M, Kronstad J. A decade of experience: *Cryptococcus gattii* in British Columbia. Mycopathologia. 2012;173:311-9.

Billmyre RB, Croll D, Li W, Mieczkowski P, Carter DA, Cuomo CA, Kronstad JW, Heitman J. Highly recombinant VGII *Cryptococcus gattii* population develops clonal outbreak clusters through both sexual macroevolution and asexual microevolution. MBio. 2014;5:e01494-14.

Chen SC, Meyer W, Sorrell TC. *Cryptococcus gattii* infections. Clin Microbiol Rev. 2014;27:980-1024.

Chowdhary A, Randhawa HS, Boekhout T, Hagen F, Klaassen CH, Meis JF. Temperate climate niche for *Cryptococcus gattii* in Northern Europe. Emerg Infect Dis. 2012;18:172-4.

Cogliati M, Chandrashekar N, Esposto MC, Chandramuki A, Petrini B, Viviani MA. *Cryptococcus gattii* serotype-C strains isolated in Bangalore, Karnataka, India. Mycoses. 2012a;55:262-8.

Cogliati M, Barchiesi F, Spreghini E, Tortorano AM. Heterozygosis and pathogenicity of *Cryptococcus neoformans* AD-hybrid isolates. Mycopathologia. 2012b;173:347-5.

Cogliati M, D'Amicis R, Zani A, Montagna MT, Caggiano G, De Giglio O, Balbino S, De Donno A, Serio F, Susever S, Ergin C, Velegraki A, Ellabib MS, Nardoni S, Macci C, Oliveri S, Trovato L, Dipineto L, Rickerts V, McCormick-Smith I, Akcaglar S, Tore O, Mlinaric-Missoni E, Bertout S, Mallié M, Martins Mda L, Vencà AC, Vieira ML, Sampaio AC, Pereira C, Criseo G, Romeo O, Ranque S, Al-Yasiri MH, Kaya M, Cerikcioglu N, Marchese A, Vezzulli L, Ilkit M, Desnos-Ollivier M, Pasquale V, Korem M, Polacheck I, Scopa A, Meyer W, Ferreira-Paim K, Hagen F, Theelen B, Boekhout T, Lockhart SR, Tintelnot K, Tortorano AM, Dromer F, Varma A, Kwon-Chung KJ, Inácio J, Alonso B, Colom MF. Environmental distribution of *Cryptococcus neoformans* and *C. gattii* around the Mediterranean basin. FEMS Yeast Res. 2016a;16. pii: fow045.

Cogliati M, Puccianti E, Montagna MT, De Donno A, Susever S, Ergin C, Velegraki A, Ellabib MS, Nardoni S, Macci C, Trovato L, Dipineto L, Rickerts V, Akcaglar S, Mlinaric-Missoni E, Bertout S, Vencà ACF, Sampaio AC, Criseo G, Ranque S, Çerikçioğlu N, Marchese A, Vezzulli L, Ilkit M, Desnos-Ollivier M, Pasquale V, Polacheck I, Scopa A, Meyer W, Ferreira-Paim K, Hagen F, Boekhout T, Dromer F, Varma A, Kwon-Chung KJ, Inácio J, Colom MF. Fundamental niche prediction of the pathogenic yeasts *Cryptococcus neoformans* and *Cryptococcus gattii* in Europe. Environ Microbiol. 2017;19:4318-4325.

Cogliati M, Zamfirova RR, Tortorano AM, Viviani MA; Fimua Cryptococcosis Network. Molecular epidemiology of Italian clinical *Cryptococcus neoformans* var. *grubii* isolates. Med Mycol. 2013;51:499-506.

Cogliati M, Zani A, Rickerts V, McCormick I, Desnos-Ollivier M, Velegraki A, Escandon P, Ichikawa T, Ikeda R, Bienvenu AL, Tintelnot K, Tore O, Akcaglar S, Lockhart S, Tortorano AM, Varma A. Multilocus sequence typing analysis reveals that *Cryptococcus neoformans* var. *neoformans* is a recombinant population. Fungal Genet Biol. 2016b;87:22-9.

Cogliati M. Global Molecular Epidemiology of *Cryptococcus neoformans* and *Cryptococcus gattii*: An Atlas of the Molecular Types. Scientifica (Cairo). 2013;2013:675213.

Colom MF, Hagen F, Gonzalez A, Mellado A, Morera N, Linares C, García DF, Peñataro JS, Boekhout T, Sánchez M. *Ceratonia siliqua* (carob) trees as natural habitat and source of infection by *Cryptococcus gattii* in the Mediterranean environment. Med Mycol. 2012;50:67-73.

Danesi P, Firacative C, Cogliati M, Otranto D, Capelli G, Meyer W. Multilocus sequence typing (MLST) and M13 PCR fingerprinting revealed heterogeneity amongst *Cryptococcus* species obtained from Italian veterinary isolates. FEMS Yeast Res. 2014;14:897-909.

Desnos-Ollivier M, Patel S, Raoux-Barbot D, Heitman J, Dromer F; French Cryptococcosis Study Group. Cryptococcosis serotypes impact outcome and provide evidence of *Cryptococcus neoformans* speciation. MBio. 2015;6:e00311.

Dioverti MV, Parikh SA, Osmon DR, Habermann TM, Tande AJ. *Cryptococcus neoformans* infections in patients with lymphoproliferative neoplasms. Leuk Lymphoma. 2018: doi: 10.1080/10428194.2018.

Engelthaler DM, Hicks ND, Gillece JD, Roe CC, Schupp JM, Driebe EM, Gilgado F, Carriconde F, Trilles L, Firacative C, Ngamskulrungroj P, Castañeda E, Lazera Mdos S, Melhem MS, Pérez-Bercoff A, Huttley G, Sorrell TC, Voelz K, May RC, Fisher MC, Thompson GR 3rd, Lockhart SR, Keim P, Meyer W. *Cryptococcus gattii* in North American Pacific Northwest: whole-population genome analysis provides insights into species evolution and dispersal. mBio. 2014;5:e01464-14.

Ferreira-Paim K, Andrade-Silva L, Fonseca FM, Ferreira TB, Mora DJ, Andrade-Silva J, Khan A, Dao A, Reis EC, Almeida MT, Maltos A, Junior VR, Trilles L, Rickerts V, Chindamporn A, Sykes JE, Cogliati M, Nielsen K, Boekhout T, Fisher M, Kwon-Chung J, Engelthaler DM, Lazéra M, Meyer W, Silva-Vergara ML. MLST-Based Population genetic analysis in a global context reveals clonality amongst *Cryptococcus neoformans* var. *grubii* VNI isolates from HIV patients in Southeastern Brazil. PLoS Negl Trop Dis. 2017;11:e0005223.

Firacative C, Roe CC, Malik R, Ferreira-Paim K, Escandón P, Sykes JE, Castañón-Olivares LR, Contreras-Peres C, Samayoa B, Sorrell TC, Castañeda E, Lockhart SR, Engelthaler DM, Meyer W. MLST and Whole-genome-based population analysis of *Cryptococcus gattii* VGIII links clinical, veterinary and environmental strains, and reveals divergent serotype specific sub-populations and distant ancestors. PLoS Negl Trop Dis. 2016;10:e0004861.

Florentino ISM, Júnior JBD, Cantilino A. Neurocryptococcosis in immunocompetent patient with psychosis. Clin Schizophr Relat Psychoses. 2018. doi: 10.3371/CSRP.FLJU.061518.

George IA, Spec A, Powderly WG, Santos CAQ. Comparative Epidemiology and Outcomes of Human Immunodeficiency virus (HIV), Non-HIV non-transplant, and solid organ transplant associated cryptococcosis: a population-based Study. Clin Infect Dis. 2018;66:608-611.

Hagen F, Ceresini PC, Polacheck I, Ma H, van Nieuwerburgh F, Gabaldón T, Kagan S, Pursall ER, Hoogveld HL, van Iersel LJ, Klau GW, Kelk SM, Stougie L, Bartlett KH, Voelz K, Pryszcz LP, Castañeda E, Lazera M, Meyer W, Deforce D, Meis JF, May RC, Klaassen CH, Boekhout T. Ancient dispersal of the human fungal pathogen *Cryptococcus gattii* from the Amazon rainforest. PLoS One. 2013;8:e71148.

Hagen F, Colom MF, Swinne D, Tintelnot K, Iatta R, Montagna MT, Torres-Rodriguez JM, Cogliati M, Velegraki A, Burggraaf A, Kamermans A, Sweere JM, Meis JF, Klaassen CH, Boekhout T. Autochthonous and dormant *Cryptococcus gattii* infections in Europe. Emerg Infect Dis. 2012;18:1618-24.

Hagen F, Khayhan K, Theelen B, Kolecka A, Polacheck I, Sionov E, Falk R, Parnmen S, Lumbsch HT, Boekhout T. Recognition of seven species in the *Cryptococcus gattii/ Cryptococcus neoformans* species complex. Fungal Genet Biol. 2015;78:16-48.

Hagen F, Lumbsch HT, Arsic-Arsenijevic V, Badali H, Bertout S, Billmyre RB, Bragulat MR, Cabañes FJ, Carbia M, Chakrabarti A, Chaturvedi S, Chaturvedi V, Chen M, Chowdhary A, Colom MF, Cornely OA, Crous PW, Cuétara MS, Diaz MR, Espinel-Ingroff A, Fakhim H, Falk R, Fang W, Herkert PF, Ferrer Rodríguez C, Fraser JA, Gené J, Guarro J, Idnurm A, Illnait-Zaragozi MT, Khan Z, Khayhan K, Kolecka A, Kurtzman CP, Lagrou K, Liao W, Linares C, Meis JF, Nielsen K, Nyazika TK, Pan W, Pekmezovic M, Polacheck I, Posteraro B, de Queiroz Telles F Filho, Romeo O, Sánchez M, Sampaio A, Sanguinetti M, Sriburee P, Sugita T, Taj-Aldeen SJ, Takashima M, Taylor JW, Theelen B, Tomazin R, Verweij PE, Wahyuningsih R, Wang P, Boekhout T. Importance of resolving fungal nomenclature: the case of multiple pathogenic species in the *Cryptococcus* genus. mSphere. 2017;2. pii: e00238-17.

Hagen F, van Assen S, Luijckx GJ, Boekhout T, Kampinga GA. Activated dormant *Cryptococcus gattii* infection in a Dutch tourist who visited Vancouver Island (Canada): a molecular epidemiological approach. Med Mycol. 2010;48:528-31.

Hamdan N, Billon Grand R, Moreau J, Thines L. Cryptococcal meningitis in an immunocompetent patient with obstructive hydrocephalus: A case report. Neurochirurgie. 2018;64:324-326.

Henao-Martínez AF, Chastain DB, Franco-Paredes C. Treatment of cryptococcosis in non-HIV immunocompromised patients. Curr Opin Infect Dis. 2018;31:278-285.

Khayhan K, Hagen F, Pan W, Simwami S, Fisher MC, Wahyuningsih R, Chakrabarti A, Chowdhary A, Ikeda R, Taj-Aldeen SJ, Khan Z, Ip M, Imran D, Sjam R, Sriburee P, Liao W, Chaicumpar K, Vuddhakul V, Meyer W, Trilles L, van Iersel LJ, Meis JF, Klaassen CH, Boekhout T. Geographically structured populations of *Cryptococcus neoformans* variety *grubii* in Asia correlate with HIV status and show a clonal population structure. PLoS One. 2013;8:e72222.

Kwon-Chung KJ, Bennett JE, Wickes BL, Meyer W, Cuomo CA, Wollenburg KR, Bicanic TA, Castañeda E, Chang YC, Chen J, Cogliati M, Dromer F, Ellis D, Filler SG, Fisher MC, Harrison TS, Holland SM, Kohno S, Kronstad JW, Lazera M, Levitz SM, Lionakis MS, May RC, Ngamskulrongroj P, Pappas PG, Perfect JR, Rickerts V, Sorrell TC, Walsh TJ, Williamson PR, Xu J, Zelazny AM, Casadevall A. The case for adopting the "species complex" nomenclature for the etiologic agents of cryptococcosis. mSphere. 2017;2. pii: e00357-16.

Kwon-Chung KJ, Fraser JA, Doering TL, Wang Z, Janbon G, Idnurm A, Bahn YS. *Cryptococcus neoformans* and *Cryptococcus gattii*, the etiologic agents of cryptococcosis. Cold Spring Harb Perspect Med. 2014;4:a019760.

Linares C, Colom MF, Torreblanca M, Esteban V, Romera Á, Hagen F. Environmental sampling of *Ceratonia siliqua* (carob) trees in Spain reveals the presence of the rare *Cryptococcus gattii* genotype AFLP7/VGIV. Rev Iberoam Micol. 2015;32:269-72.

Mak S, Klinkenberg B, Bartlett K, Fyfe M. Ecological niche modeling of *Cryptococcus gattii* in British Columbia, Canada. Environ Health Perspect. 2010;118:653-8.

Meyer W, Aanensen DM, Boekhout T, Cogliati M, Diaz MR, Esposto MC, Fisher M, Gilgado F, Hagen F, Kaocharoen S, Litvintseva AP, Mitchell TG, Simwami SP, Trilles L, Viviani MA, Kwon-Chung J. Consensus multi-locus sequence typing scheme for *Cryptococcus neoformans* and *Cryptococcus gattii*. Med Mycol. 2009;47:561-70.

Meyer W, Castañeda A, Jackson S, Huynh M, Castañeda E; IberoAmerican Cryptococcal Study Group. Molecular typing of IberoAmerican *Cryptococcus neoformans* isolates. Emerg Infect Dis. 2003;9:189-95.

Miranda-Filho A, Piñeros M, Ferlay J, Soerjomataram I, Monnereau A, Bray F. Epidemiological patterns of leukaemia in 184 countries: a population-based study. Lancet Haematol. 2018;5:e14-e24.

Montagna MT, De Donno A, Caggiano G, Serio F, De Giglio O, Bagordo F, D'Amicis R, Lockhart SR, Cogliati M. Molecular characterization of *Cryptococcus neoformans* and *Cryptococcus gattii* from environmental sources and genetic comparison with clinical isolates in Apulia, Italy. Environ Res. 2018;160:347-352.

Morse JC, Gelbard A. Laryngeal cryptococcoma resulting in airway compromise in an immunocompetent patient: A case report. Laryngoscope. 2018. doi: 10.1002/lary.27456.

Muñoz M, Camargo M, Ramírez JD. Estimating the intra-taxa diversity, population genetic structure, and evolutionary pathways of *Cryptococcus neoformans* and *Cryptococcus gattii.* Front Genet. 2018;9:148.

Nyazika TK, Hagen F, Meis JF, Robertson VJ. *Cryptococcus tetragattii* as a major cause of cryptococcal meningitis among HIV-infected individuals in Harare, Zimbabwe. J Infect. 2016;72:745-752.

Pappas PG. Cryptococcal infections in non-HIV-infected patients. Trans Am Clin Climatol Assoc. 2013;124:61-79.

Posteraro P, Vacca K, Celi M, De Angelis G, Vella A, Sanguinetti M, Posteraro B. Fatal fulminant cryptococcemia complicating sarcoidosis: Is it to be expected? Med Mycol Case Rep. 201;22:42-44.

Rhodes J, Desjardins CA, Sykes SM, Beale MA, Vanhove M, Sakthikumar S, Chen Y, Gujja S, Saif S, Chowdhary A, Lawson DJ, Ponzio V, Colombo AL, Meyer W, Engelthaler DM, Hagen F, Illnait-Zaragozi MT, Alanio A, Vreulink JM, Heitman J, Perfect JR, Litvintseva AP, Bicanic T, Harrison TS, Fisher MC, Cuomo CA. Tracing genetic exchange and biogeography of *Cryptococcus neoformans* var. *grubii* at the global population level. Genetics. 2017;207:327-346.

Roe CC, Bowers J, Oltean H, DeBess E, Dufresne PJ, McBurney S, Overy DP, Wanke B, Lysen C, Chiller T, Meyer W, Thompson GR 3rd, Lockhart SR, Hepp CM, Engelthaler DM. Dating the *Cryptococcus gattii* dispersal to the North American Pacific Northwest. mSphere. 2018;3. pii: e00499-17.

Samarasinghe H, Aceituno-Caicedo D, Cogliati M, Kwon-Chung KJ, Rickerts V, Velegraki A, Ackaglar S, Xu J. Genetic factors and genotype-environment interactions contribute to variation in melanin production in the fungal pathogen *Cryptococcus neoformans*. Sci Rep. 2018;8:9824.

Sanchini A, Smith IM, Sedlacek L, Schwarz R, Tintelnot K, Rickerts V. Molecular typing of clinical *Cryptococcus neoformans* isolates collected in Germany from 2004 to 2010. Med Microbiol Immunol. 2014;203:333-40.

Shahid M, Han S, Yoell H, Xu J. Fitness distribution and transgressive segregation across 40 environments in a hybrid progeny population of the human-pathogenic yeast *Cryptococcus neoformans.* Genome. 2008;51:272-81.

Shuai SY, Xiong L, He XL, Yu F, Xia Q, Zhou Q. A unique case report of endobronchial cryptococcosis and review of the literature. Respir Med Case Rep. 2018;25:247-252.

Sohail A, Smibert OC, Snell G, Paraskeva M, Jenney A. Cryptococcal infection in lung transplant recipients: A 5-year retrospective review at an Australian transplant center. Transpl Infect Dis. 2018:e12976.

Souto AC, Bonfietti LX, Ferreira-Paim K, Trilles L, Martins M, Ribeiro-Alves M, Pham CD, Martins L, Dos Santos W, Chang M, Brito-Santos F, Santos DC, Fortes S, Lockhart SR, Wanke B, Melhem MS, Lazéra MS, Meyer W. Population genetic analysis reveals a high genetic diversity in the Brazilian *Cryptococcus gattii* VGII population and shifts the global origin from the Amazon rainforest to the semi-arid desert in the Northeast of Brazil. PLoS Negl Trop Dis. 2016;10:e0004885.

Sun S, Xu J. Chromosomal rearrangements between serotype A and D strains in *Cryptococcus neoformans.* PLoS One. 2009;4:e5524.

Swan CD, Gottlieb T. *Cryptococcus neoformans* empyema in a patient receiving ibrutinib for diffuse large B-cell lymphoma and a review of the literature. BMJ Case Rep. 2018;2018. pii: bcr-2018-224786.

Taimur S. Yeast infections in solid organ transplantation. Infect Dis Clin North Am. 2018;32:651-666.

Viviani MA, Cogliati M, Esposto MC, Lemmer K, Tintelnot K, Colom Valiente MF, Swinne D, Velegraki A, Velho R; European Confederation of Medical Mycology (ECMM) Cryptococcosis Working Group. Molecular analysis of 311 *Cryptococcus neoformans* isolates from a 30-month ECMM survey of cryptococcosis in Europe. FEMS Yeast Res. 2006;6:614-9.

Wang Q, Wang Z, Hao Y, Li W, Xin T, Chen M, Gao P. Coinfection with *Cryptococcus* and *Aspergillus* in an immunocompetent adult: A case report. Medicine (Baltimore). 2018;97:e12612.

Yamada W, Yamada H, Murata K, Kosugi H, Asano Y, Mochizuki K, Ishida K. Case of cryptococcal choroiditis in adult with T-cell leukemia/lymphoma. J Infect Chemother. 2018. pii: S1341-321X(18)30190-9.

Figure legends

Figure 1. Minimum spanning tree showing the main clonal clusters (CC) linked by a single locus difference. Isolates have been identified according to the source and the country of isolation. Each circle represents a unique sequence type (ST), and the circle size is proportional to the number of isolates within each ST.

Figure 2. Maximum likelihood phylogenetic tree inferred using the concatenated sequences of the seven MLST loci (*CAP59, GPD1, LAC1, PLB1, SOD1, URA5*, and the IGS1 region) of the *Cryptococcus neoformans* species complex sequence types (ST) investigated in the present study. Bootstrap values, based on 1,000 replicates, are reported at each branch node. Red dots indicate STs including clinical isolates, green dots environmental isolates, and blue dots veterinary isolates. STs highlighted in light yellow include mating type **a** isolates. Black arrows indicate STs with an ambiguous topology.

Figure 3. Maximum likelihood phylogenetic tree inferred using the concatenated sequences of the seven MLST loci (*CAP59, GPD1, LAC1, PLB1, SOD1, URA5*, and the IGS1 region) of the *Cryptococcus gattii* species complex sequence types (ST) investigated in the present study (marked in light red). STs are compared with STs from European VGI isolates previously reported by other authors (Hagen et al. 2010) who analyzed also further VGII and VGIII isolates not included in the current analysis. Bootstrap values, based on 1,000 replicates, are reported at each branch node. Black squares indicate European autochthonous STs.

Figure 4. Network analysis showing genetic relationships between VNI, VNII and VNB isolates (right-side panel) and between VNI and VNIV isolates (left-side panels). Each circle represents

a sequence type, the size of the circles is proportional to the numbers of isolates, and the length of each branch is proportional to genetic distance.

Figure 5. Linkage disequilibrium between pair of loci calculated by Arlequin v3.0. Each graphic represents the results obtained in a different group of isolates. Red squares indicate that two loci are linked whereas green squares indicate no linkage.

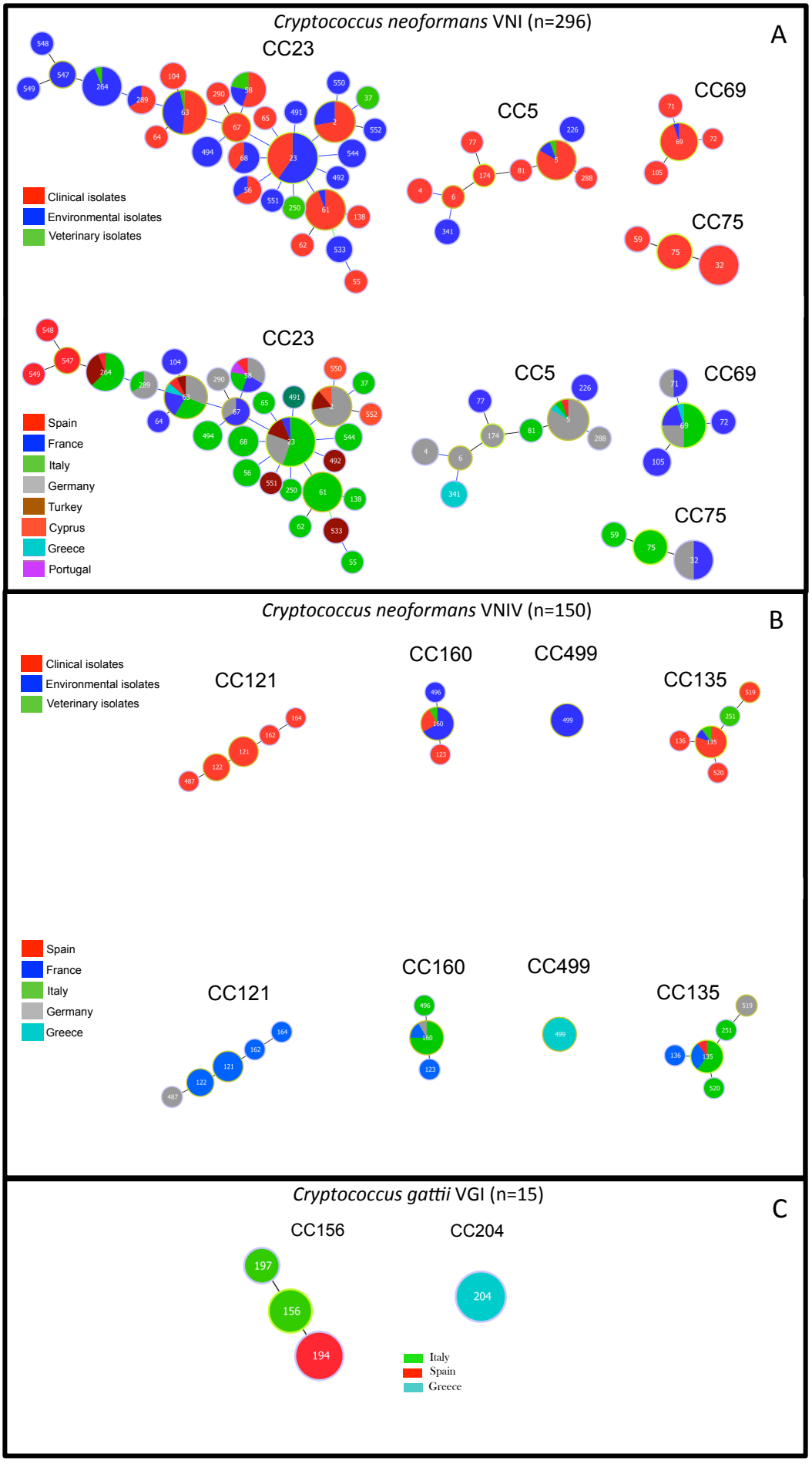
Figure 6. Inter-population genetic analysis inferred by DNAsp v5.10.1. In each comparison four parameters are represented with different colors: percentage of fixed mutations, percentage of shared mutations, percentage of polymorphic sites in population 1 and monomorphic in population 2, and percentage of polymorphic sites in population 2 and monomorphic in population 1.

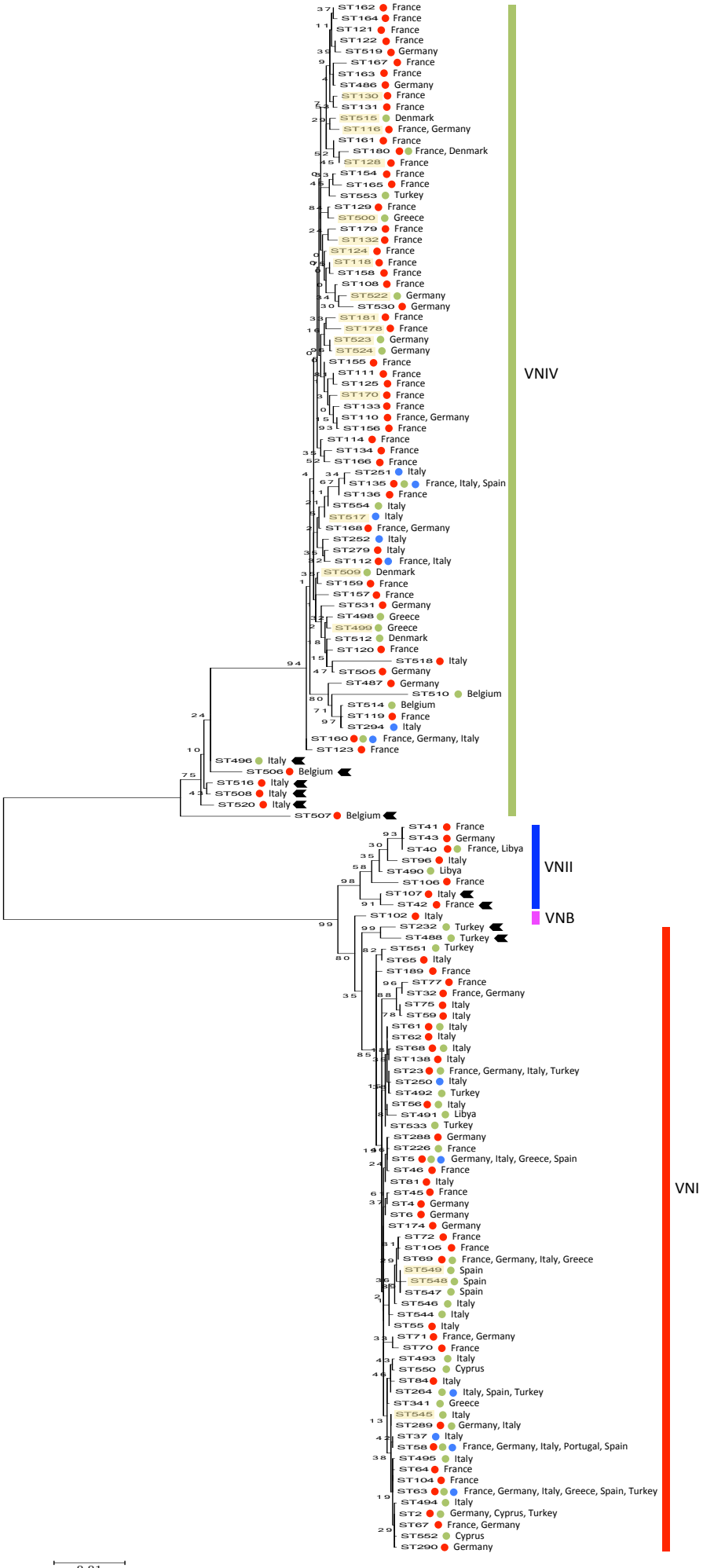
Figure 7. Variation of nucleotide diversity along the seven-loci concatenated sequences in VNI and VNIV clinical and environmental/veterinary (ENV) isolates. Portion of sequence covered by each of the seven loci is indicated on X ax. **A**. Comparison between clinical and ENV isolates. **B**. Comparison between clinical VNI and ENV VNI isolates. **C**. Comparison between clinical VNIV and ENV VNIV isolates. **D**. Comparison between VNI and VNIV isolates. **E**. Comparison between ENV VNI and ENV VNIV isolates. **F**. Comparison between clinical VNI and clinical VNIV isolates.

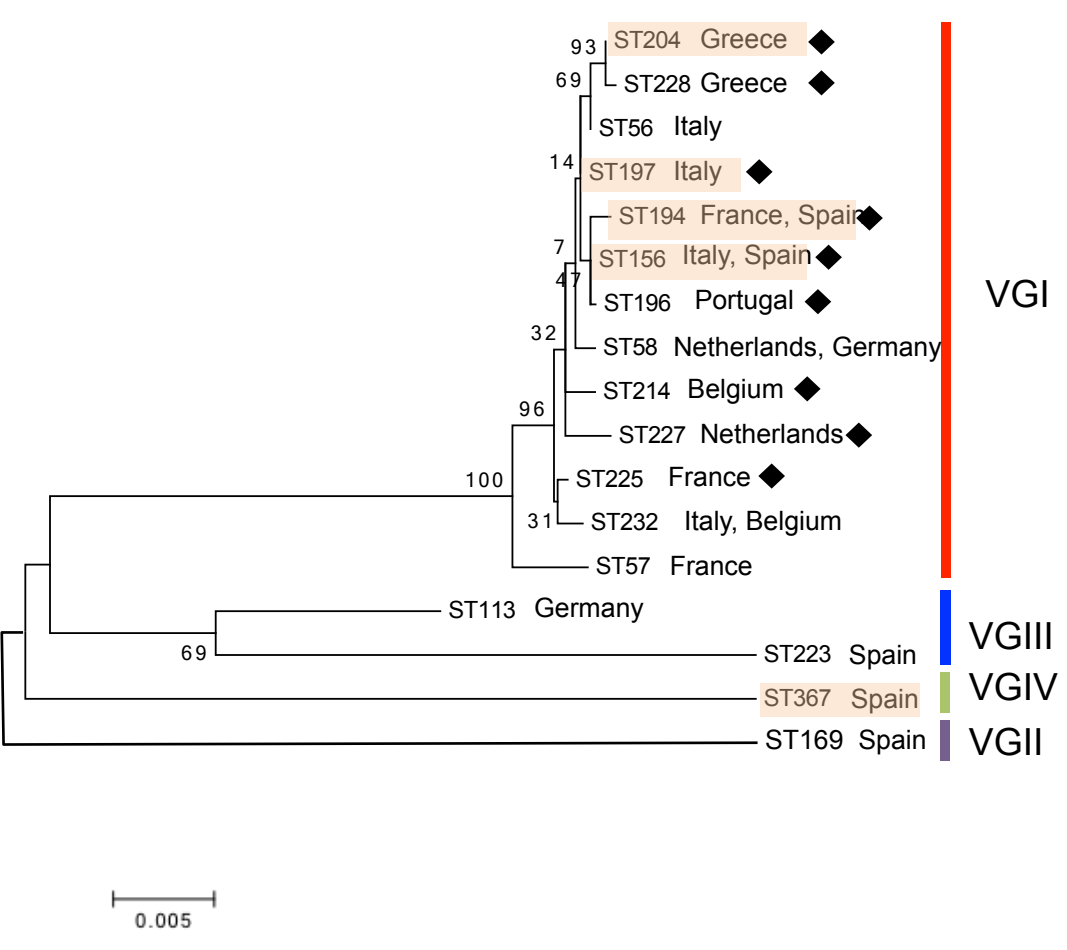
Figure 8. Maps representing the main genetic flux directions between European and Mediterranean countries. Pairwise F_{ST} test was calculated by DNAsp v5.10.1 and is reported on each arrow. Asterisk indicates that the genetic flux is supported by statistical analysis ($p<0.05$). The most probable spreading directions for ST5, ST23, ST63, and ST69 (the most

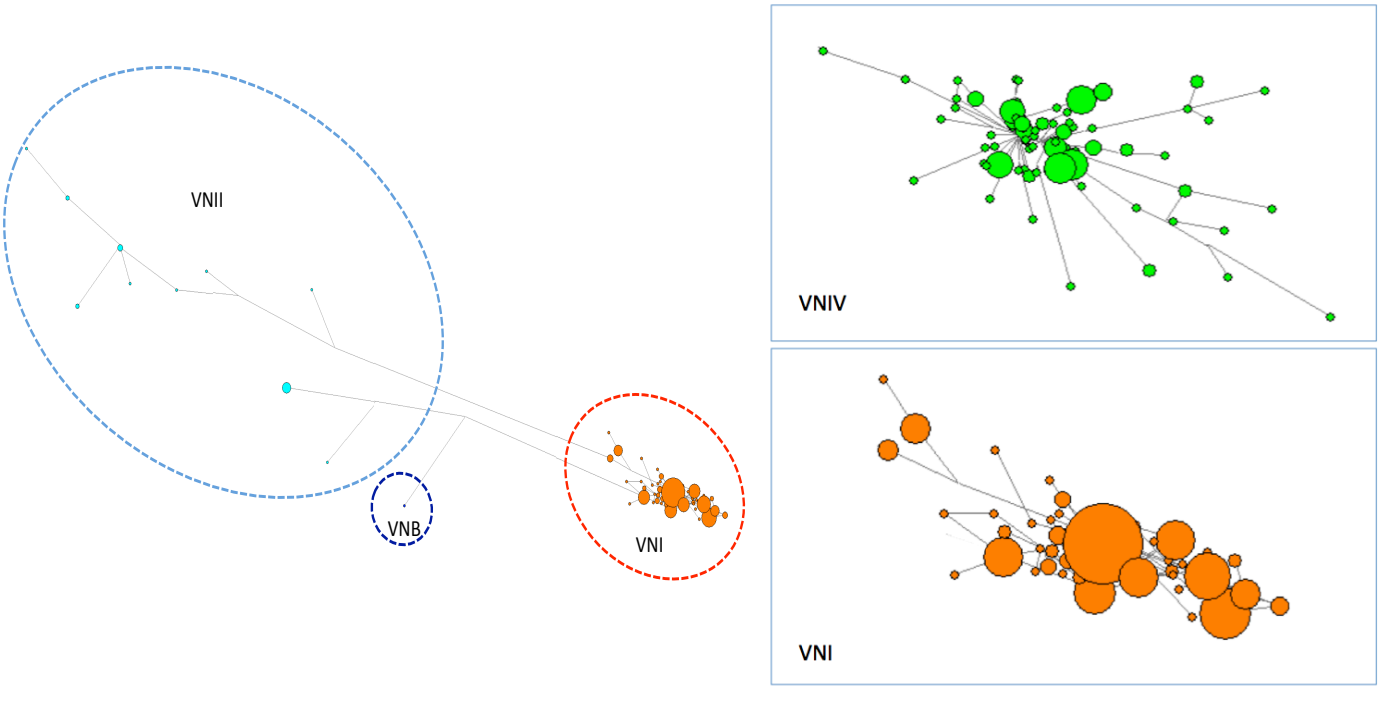
prevalent genotypes in Europe) are indicated as black arrows. Light blue arrows represent alternative spreading directions not supported by statistical analysis. Green circles represent environmental isolates and red circles the clinical ones. The number in each circle is the number of isolates.

Figure S1. Haplotype diversity values calculated for each locus. **A**. Comparison of VNI, VNII, and VNIV isolates. **B**. Comparison of clinical and environmental/veterinary (ENV) whole isolates. **C**. Comparison of clinical and ENV VNI, and clinical and ENV VNIV isolates.

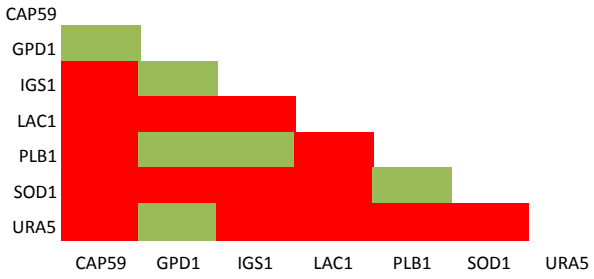

A
Cryptococcus neoformans VNI (n=296)
CC23
CC5
CC69
CC75
Clinical isolates
Environmental isolates
Veterinary isolates
Spain
France
Italy
Germany
Turkey
Cyprus
Greece
Portugal
B
Cryptococcus neoformans VNIV (n=150)
CC121
CC160
CC499
CC135
C
Cryptococcus gattii VGI (n=15)
CC156
CC204
Italy
Spain
Greece


ST162 France
ST164 France
ST121 France
ST122 France
ST519 Germany
ST167 France
ST163 France
ST486 Germany
ST130 France
ST131 France
ST515 Denmark
ST116 France, Germany
ST161 France
ST180 France, Denmark
ST128 France
ST154 France
ST165 France
ST553 Turkey
ST129 France
ST500 Greece
ST179 France
ST132 France
ST124 France
ST118 France
ST158 France
ST108 France
ST522 Germany
ST530 Germany
ST181 France
ST178 France
ST523 Germany
ST524 Germany
ST155 France
ST111 France
ST125 France
ST170 France
ST133 France
ST110 France, Germany
ST156 France
ST114 France
ST134 France
ST166 France
ST251 Italy
ST135 France, Italy, Spain
ST136 France
ST554 Italy
ST517 Italy
ST168 France, Germany
ST252 Italy
ST279 Italy
ST112 France, Italy
ST509 Denmark
ST159 France
ST157 France
ST531 Germany
ST498 Greece
ST499 Greece
ST512 Denmark
ST120 France
ST518 Italy
ST505 Germany
ST487 Germany
ST510 Belgium
ST514 Belgium
ST119 France
ST294 Italy
ST160 France, Germany, Italy
ST123 France
ST496 Italy
ST506 Belgium
ST516 Italy
ST508 Italy
ST520 Italy
ST507 Belgium
VNIV
ST41 France
ST43 Germany
ST40 France, Libya
ST96 Italy
ST490 Libya
ST106 France
ST107 Italy
ST42 France
VNII
ST102 Italy
VNB
ST232 Turkey
ST488 Turkey
ST551 Turkey
ST65 Italy
ST189 France
ST77 France
ST32 France, Germany
ST75 Italy
ST59 Italy
ST61 Italy
ST62 Italy
ST68 Italy
ST138 Italy
ST23 France, Germany, Italy, Turkey
ST250 Italy
ST492 Turkey
ST56 Italy
ST491 Libya
ST533 Turkey
ST288 Germany
ST226 France
ST5 Germany, Italy, Greece, Spain
ST46 France
ST81 Italy
ST45 France
ST4 Germany
ST6 Germany
ST174 Germany
ST72 France
ST105 France
ST69 France, Germany, Italy, Greece
ST549 Spain
ST548 Spain
ST547 Spain
ST546 Italy
ST544 Italy
ST55 Italy
ST71 France, Germany
ST70 France
ST493 Italy
ST550 Cyprus
ST84 Italy
ST264 Italy, Spain, Turkey
ST341 Greece
ST545 Italy
ST289 Germany, Italy
ST37 Italy
ST58 France, Germany, Italy, Portugal, Spain
ST495 Italy
ST64 France
ST104 France
ST63 France, Germany, Italy, Greece, Spain, Turkey
ST494 Italy
ST2 Germany, Cyprus, Turkey
ST67 France, Germany
ST552 Cyprus
ST290 Germany
VNI
0.01

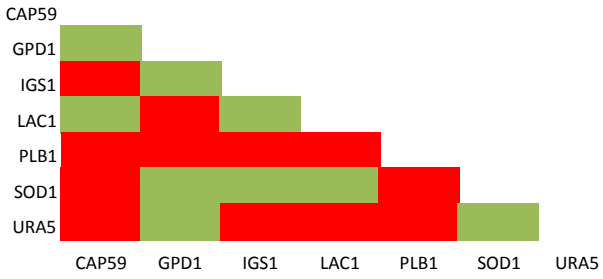

93
69
14
7
47
32
96
100
31
69
ST204 Greece
ST228 Greece
ST56 Italy
ST197 Italy
ST194 France, Spain
ST156 Italy, Spain
ST196 Portugal
ST58 Netherlands, Germany
ST214 Belgium
ST227 Netherlands
ST225 France
ST232 Italy, Belgium
ST57 France
ST113 Germany
ST223 Spain
ST367 Spain
ST169 Spain
VGI
VGIII
VGIV
VGII
0.005


VNII
VNB
VNI
VNIV
VNI

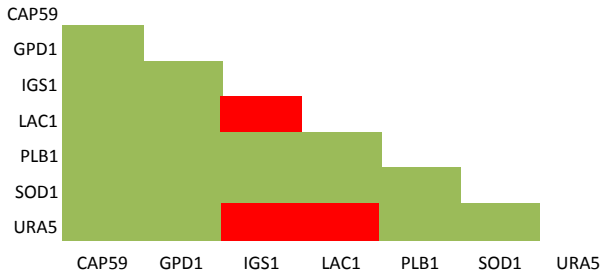
VNIV (n=150)



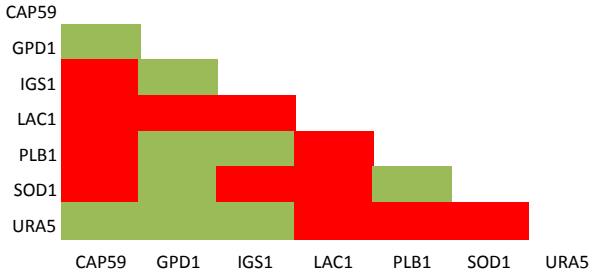
VNI (n=296)



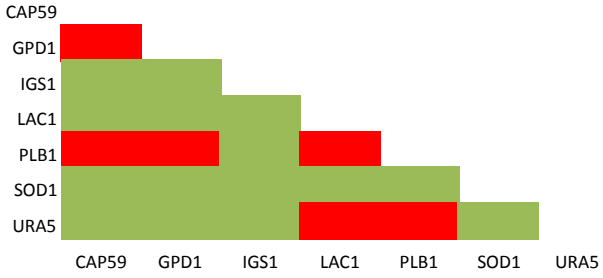
VNII (n=13)



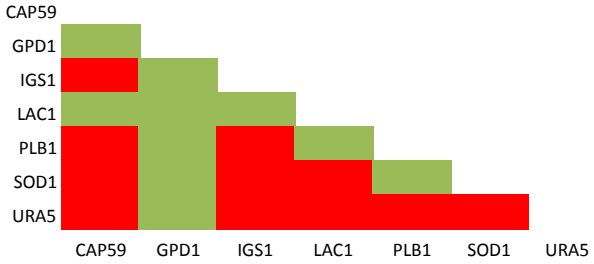
CLINICAL VNIV (n=99)



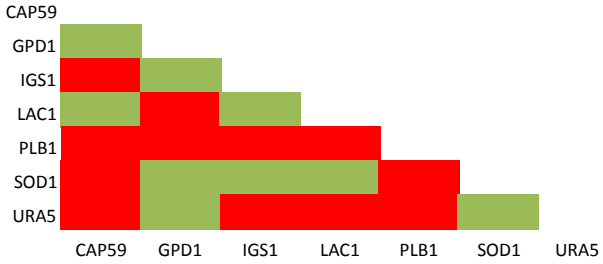
CLINICAL VNI (n=164)

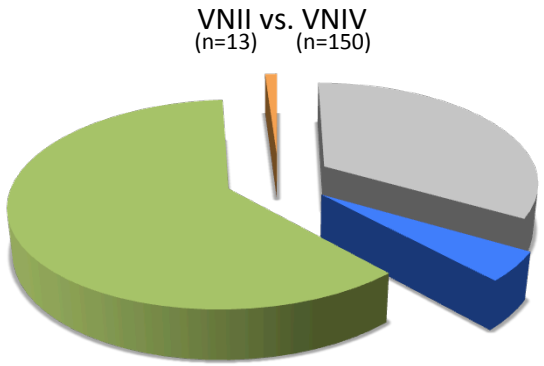


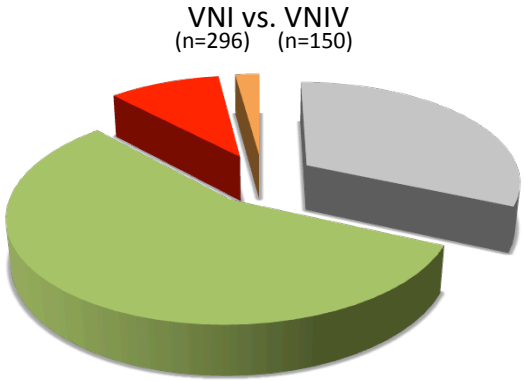
ENVIRONMENTAL VNIV (n=51)

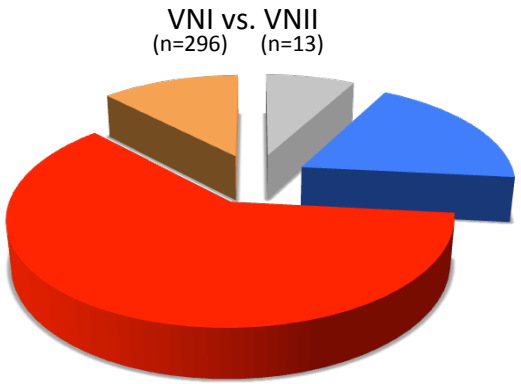


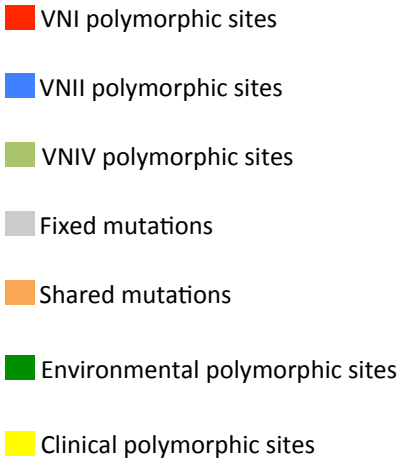
ENVIRONMENTAL VNI (n=132)

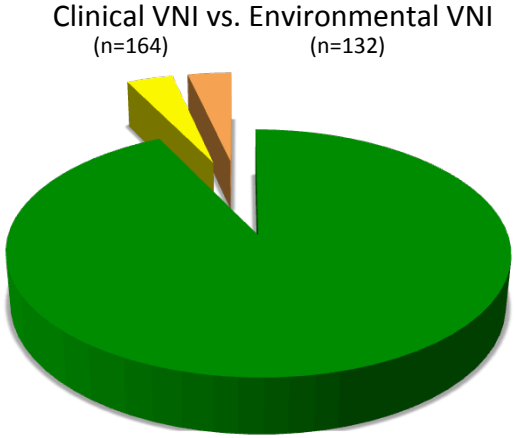


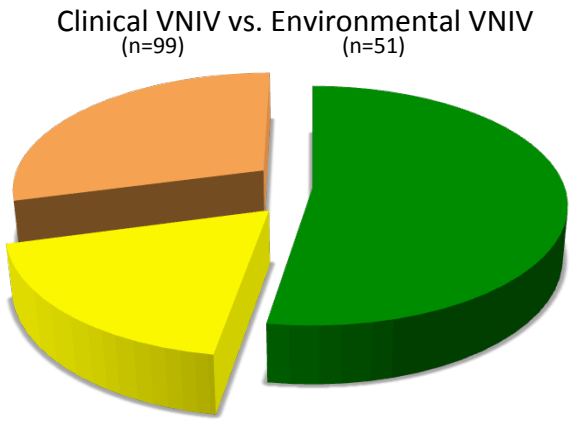

VNII vs. VNIV
(n=13) (n=150)

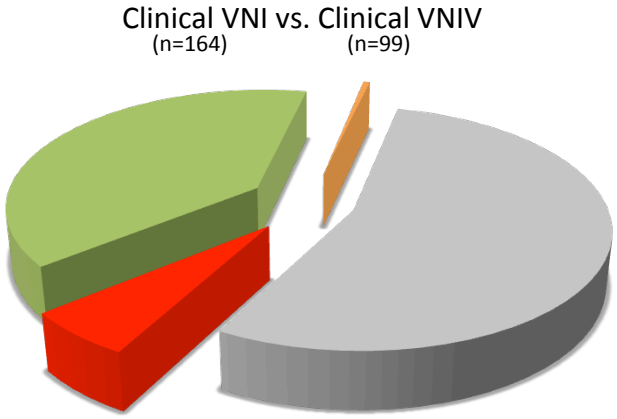

VNI vs. VNIV
(n=296) (n=150)

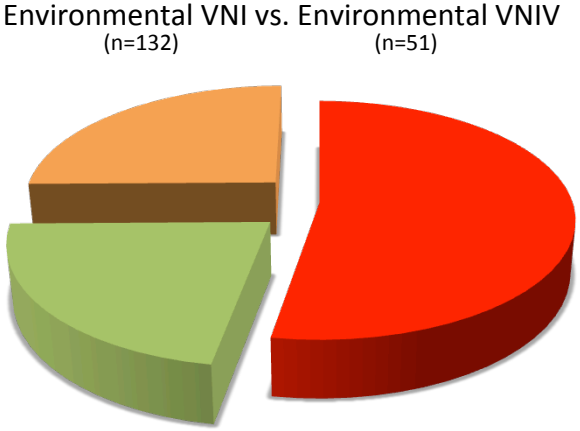

VNI vs. VNII
(n=296) (n=13)

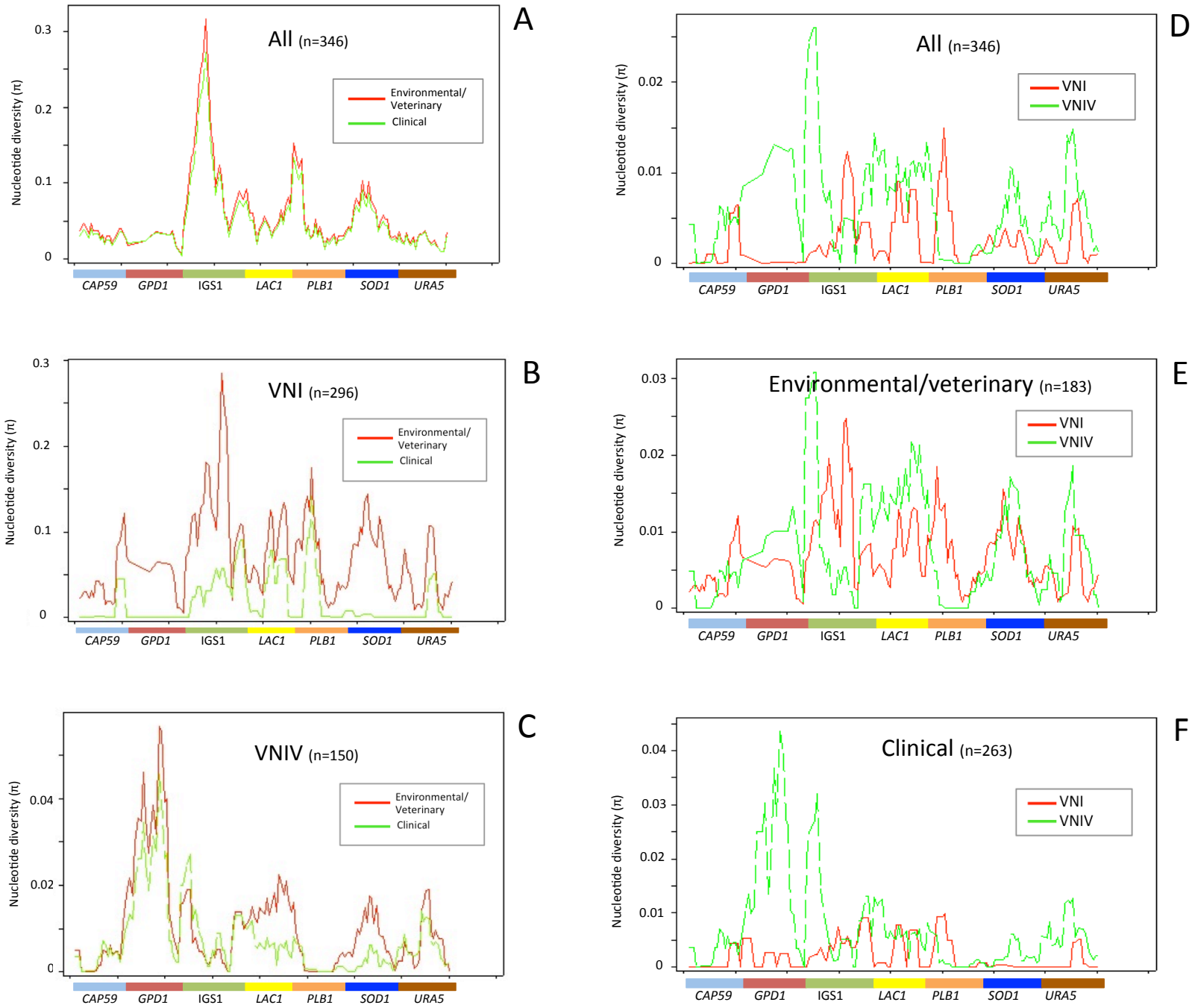

VNI polymorphic sites
VNII polymorphic sites
VNIV polymorphic sites
Fixed mutations
Shared mutations
Environmental polymorphic sites
Clinical polymorphic sites

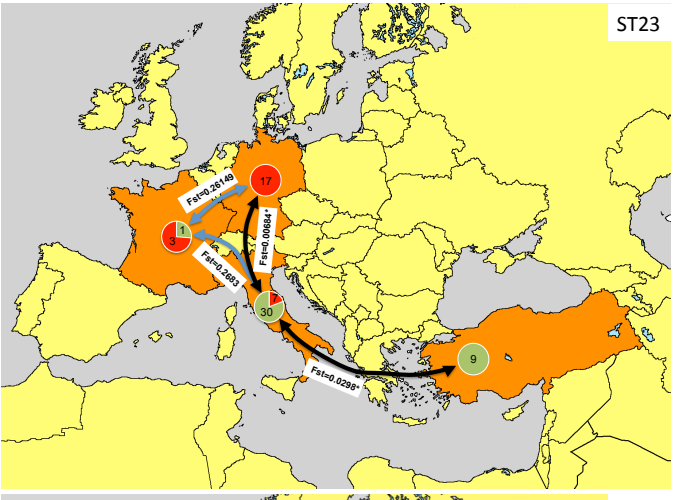

Clinical VNI vs. Environmental VNI
(n=164) (n=132)

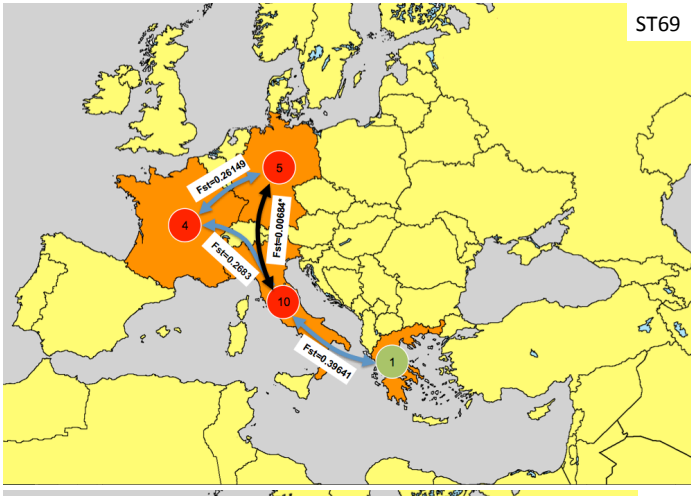

Clinical VNIV vs. Environmental VNIV
(n=99) (n=51)

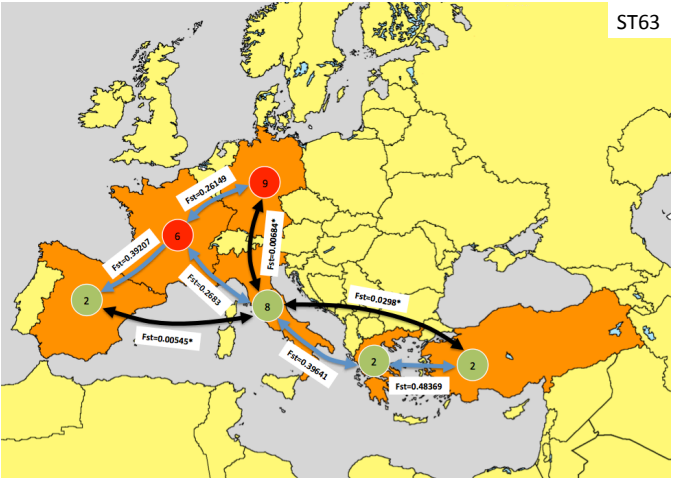

Clinical VNI vs. Clinical VNIV
(n=164) (n=99)


Environmental VNI vs. Environmental VNIV
(n=132) (n=51)


A
All (n=346)
Environmental/ Veterinary
Clinical
Nucleotide diversity (π)
0.3
0.2
0.1
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5
B
VNI (n=296)
Environmental/ Veterinary
Clinical
Nucleotide diversity (π)
0.3
0.2
0.1
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5
C
VNIV (n=150)
Environmental/ Veterinary
Clinical
Nucleotide diversity (π)
0.04
0.02
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5
D
All (n=346)
VNI
VNIV
Nucleotide diversity (π)
0.02
0.01
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5
E
Environmental/veterinary (n=183)
VNI
VNIV
Nucleotide diversity (π)
0.03
0.02
0.01
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5
F
Clinical (n=263)
VNI
VNIV
Nucleotide diversity (π)
0.04
0.03
0.02
0.01
0
CAP59 GPD1 IGS1 LAC1 PLB1 SOD1 URA5


ST23
Fst=0.26149
Fst=0.00684*
Fst=0.2683
Fst=0.0298*
17
1
3
30
9


ST69
Fst=0.26149
Fst=0.00684*
Fst=0.2683
Fst=0.39641
5
4
10
1


ST63
Fst=0.26149
Fst=0.00684*
Fst=0.39207
Fst=0.2683
Fst=0.0298*
Fst=0.00545*
Fst=0.39641
Fst=0.48369
9
6
2
8
2
2

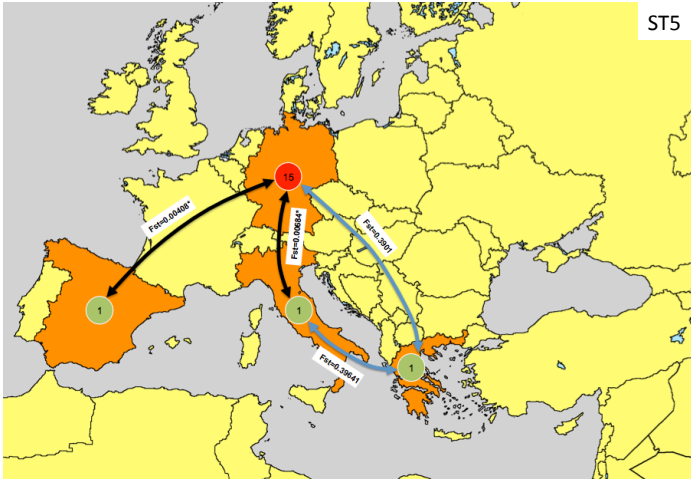

ST5
Fst=0.00408*
Fst=0.00684*
Fst=0.3901
Fst=0.39641
15
1
1
1

Table 1. Geographical origin, source, and molecular type of the 476 *C. neoformans* and *C. gattii* species complex isolates investigated in the present study

		Molecular types						
		VNI	VNII	VNB	VNIV	VGI	VGIV	Total
Belgium	C				2			2
	E				2			2
Cyprus	C							0
	E	4						4
Denmark	C							0
	E				4			4
France	C	38	7		70			115
	E	3						3
Germany	C	77	1		15			93
	E				3			3
Greece	C							0
	E	6			18	6		30
Italy	C	49	3	1	12			65
	E	71			19	4		94
Libya	C							0
	E	1	2					3
Portugal	C							0
	E	1						1
Spain	C							0
	E	9			1	5	1	16
Turkey	C							0
	E	37			4			41
Total	C	164	11	1	99	0	0	275
	E	132	2	0	51	15	1	201
Mating type	α	293	13	1	116	12	1	436
	a	3	0	0	33	3	0	40

C = Clinical isolates; E = Environmental isolates

Table 2. Sequence type frequency among clinical and environmental VNI isolates

Sequence type (ST)	No. of isolates (%)		
	Clinical (n=164)	Environmental (n=132)	Total (n=296)
23	27 (16)	40 (31)	67 (22)
63	15 (9)	14 (11)	29 (10)
69	19 (11)	1 (0.8)	20 (7)
2	13 (8)	5 (4)	18 (6)
5	15 (9)	3 (2)	18 (6)
61	16 (10)	1 (0.8)	17 (6)
264	-	16 (12)	16 (5)
232	-	13 (10)	13 (4)
32	10 (6)	-	10 (3)
58	5 (3)	4 (3)	9 (3)
68	2 (1)	3 (2)	5 (2)
75	5 (3)	-	5 (2)
46	4 (2)	-	4 (1)
494	-	4 (3)	4 (1)
544	-	3 (2)	3 (1)
45, 67	3 (2)	-	3 (1)
56, **289**	2 (1)	1 (0.8)	3 (1)
4, 71, **104**,**105**	2 (1)	-	2 (0.7)
226, **341**, **493**, **533**, **547**	-	2 (2)	2 (0.7)
6, 37, **55**, **59**, **62**, 64, **65**, **70**, **72**, 77, 81, **84**, **138**, 174, **250**, **288**, **290**, **488**, **491**, **492**, **495**, **545**, **546**, **548**, **549**, **550**, **551**, **552**	1 (0.6)	-	1 (0.3)

STs indicated in bold are exclusive of Europe

Table 3. Sequence type frequency among clinical and environmental VNIV isolates

Sequence type (ST)	No. of isolates (%)		
	Clinical (n=99)	Environmental (n=51)	Total (n=150)
499	-	16 (31)	16 (10)
160	3 (3)	9 (18)	12 (8)
135	8 (8)	2 (4)	10 (7)
116	9 (9)	-	9 (6)
121	7 (7)	-	7 (5)
112	4 (4)	2 (4)	6 (4)
114, **122**	4 (4)	-	4 (3)
553	-	4 (9)	4 (3)
180	2 (2)	1 (2)	3 (2)
110, **125**, 168	3 (3)	-	3 (2)
294	-	2 (4)	2 (1)
133, **134**, **165**, **508**	2 (2)	-	2 (1)
251, **252**, **496**, **498**, **500**, **517**, **522**, **523**, **524**, **554**	-	1 (2)	1 (0.7)
108, **111**, **118**, **119**, **120**, **123**, **124**, **128**, **129**, **130**, **131**, **132**, **136**, **154**, **155**, **156**, **157**, **158**, **159**, **161**, **162**, **163**, **164**, **166**, **167**, **170**, **178**, **179**, **181**, **279**, **486**, **487**, **505**, **506**, **507**, **509**, **510**, **512**, **514**, **515**, **516**, **518**, **519**, **520**, **530**, **531**	1 (1)	-	1 (0.7)

STs indicated in bold are exclusive of Europe

Table 4. Parameters values of population genetics calculated for the populations investigated in the present study

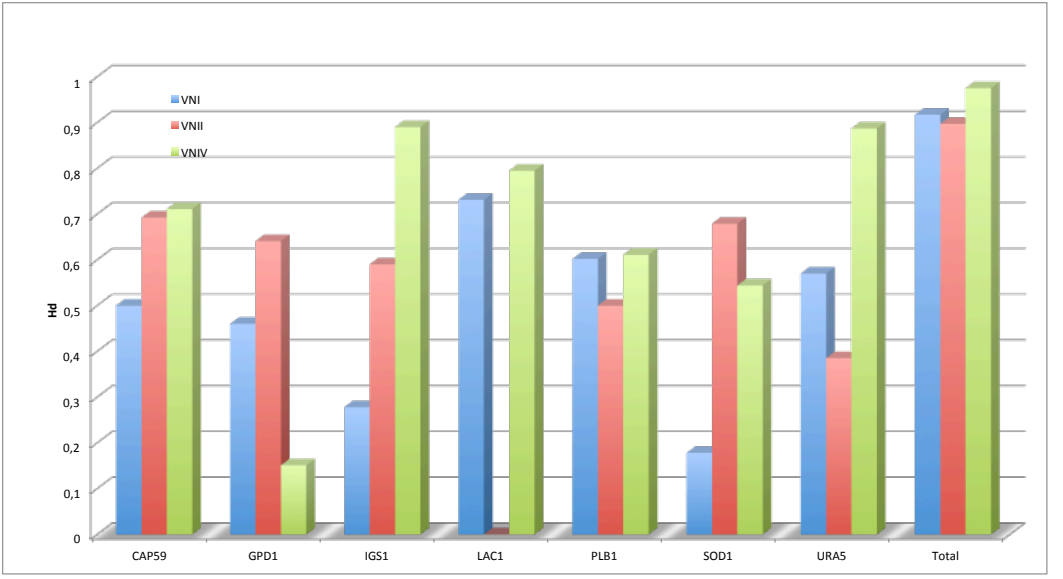
Population (no. of isolates)	Polymorphic sites (S)	No. of mutations (Eta)	No. of haplotypes	Haplotype diversity (Hd)	Average no. of nucleotide difference (K)	Nucleotide diversity (π)	TajimaD	FuLiD	FuLiF
VNI (n=296)	102	102	54	0.917	8.67526	0.00264	-1.0652	-1.6819	-1.6536
VNII (n=13)	43	43	8	0.897	12.41026	0.0037	0.3017	0.7634	0.7318
VNIV (n=150)	386	420	76	0.975	20.87302	0.00615	-2.3134**	-6.8658**	-5.6255**
CLINICAL (n=275)	640	675	99	0,969	195,99388	0,05824	3,3209**	0,2625	2,2016**
ENVIRONMENTAL (n=185)	602	642	55	0,921	172,85493	0,05015	1,9673*	0,5456	1,4837*
VNI CLINICAL (n=164)	43	43	31	0,919	6,21739	0,00195	0,0795	-0,3883	-0,2285
VNI ENVIRONMENTAL (n=132)	492	499	29	0,859	23,72921	0,00697	-2,3989**	2,4411**	0,2253
VNIV CLINICAL (n=99)	254	266	55	0,977	17,00417	0,00723	-1,5402	-3,0985*	-2,9078*
VNIV ENVIRONMENTAL (n=51)	415	450	25	0,909	24,39592	0,01048	-2,166*	-3,7735**	-3,7839**
FRANCE (n=118)	562	583	63	0,984	206,30268	0,06193	3,9762***	2,2139**	3,6125**
GERMANY (n=96)	557	574	26	0,901	136,09888	0,04178	1,5002	1,3588	1,7042*
GREECE (n=24)	544	569	11	0,770	187,37	0,05496	1,2337	0,8888	1,1779
ITALY (n=155)	556	578	43	0,928	134,18172	0,03963	1,5467	2,0203**	2,1465**
SPAIN (n=10)	491	493	8	0,956	85,53333	0,02593	-2,0918**	-2,4981**	-2,7097**
TURKEY (n=41)	538	552	11	0,840	103,38902	0,03157	-0,1907	1,6077**	1,1366

*p<0.05, **p<0.01, ***p<0.001

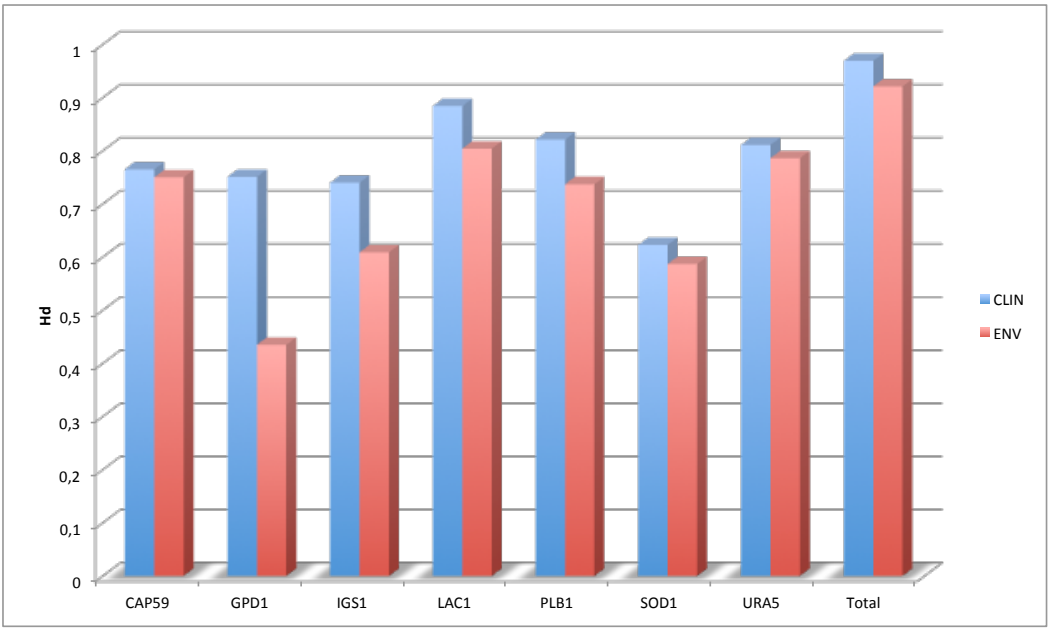
Table 5. Minimun number of recombination events calculated among populations investigated in the present study

Population	*CAP59*	*GPD1*	IGS1	*LAC1*	*PLB1*	*SOD1*	*URA5*	Total
VNI (n=296)	0	0	0	0	0	0	0	0
VNII (n=13)	0	0	0	0	0	0	0	0
VNIV (n=150)	4	2	6	17	0	3	5	37
CLINICAL (n=275)	5	3	9	7	1	3	5	33
ENVIRONMENTAL (n=185)	4	2	5	21	1	4	3	40
VNI CLINICAL (n=164)	0	0	0	0	0	0	0	0
VNI ENVIRONMENTAL (n=132)	0	0	0	1	1	3	0	5
VNIV CLINICAL (n=99)	4	0	5	1	0	3	5	18
VNIV ENVIRONMENTAL (n=51)	2	12	5	11	0	1	2	28
FRANCE (n=118)	3	3	6	2	1	2	3	20
GERMANY (n=96)	2	2	1	3	0	1	2	11
GREECE (n=24)	0	2	4	9	0	0	0	15
ITALY (n=155)	2	3	6	1	1	3	2	18
SPAIN (n=10)	0	0	0	0	0	0	0	0
TURKEY (n=41)	0	1	2	3	1	2	0	9

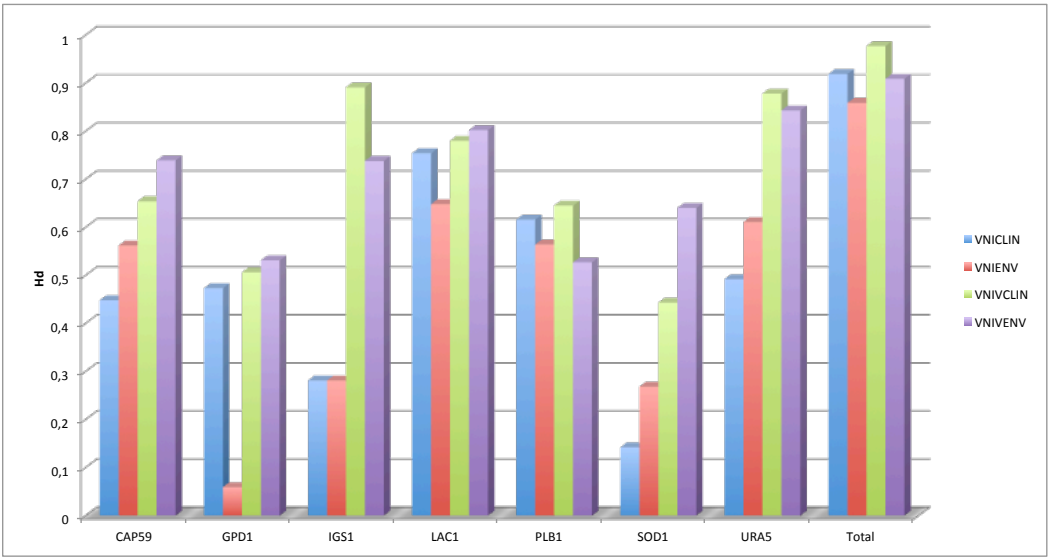
A


VNI
VNII
VNIV
Hd
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
CAP59
GPD1
IGS1
LAC1
PLB1
SOD1
URA5
Total

B


Hd
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
CAP59
GPD1
IGS1
LAC1
PLB1
SOD1
URA5
Total
CLIN
ENV

C


Hd
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
CAP59
GPD1
IGS1
LAC1
PLB1
SOD1
URA5
Total
VNICLIN
VNIENV
VNIVCLIN
VNIVENV