

10-1-2011

SPECIES AND GENOTYPE DIVERSITIES OF YEASTS IN THE CLINICAL AND NATURAL ENVIRONMENTS IN HAMILTON

Harinad Babu Maganti
McMaster University, maganthb1@gmail.com

Follow this and additional works at: <http://digitalcommons.mcmaster.ca/opensdissertations>

 Part of the [Environmental Microbiology and Microbial Ecology Commons](#), and the [Pathogenic Microbiology Commons](#)

Recommended Citation

Maganti, Harinad Babu, "SPECIES AND GENOTYPE DIVERSITIES OF YEASTS IN THE CLINICAL AND NATURAL ENVIRONMENTS IN HAMILTON" (2011). *Open Access Dissertations and Theses*. Paper 5589.

This Thesis is brought to you for free and open access by the Open Dissertations and Theses at DigitalCommons@McMaster. It has been accepted for inclusion in Open Access Dissertations and Theses by an authorized administrator of DigitalCommons@McMaster. For more information, please contact scom@mcmaster.ca.

**SPECIES AND GENOTYPE DIVERSITIES OF YEASTS IN THE CLINICAL AND
NATURAL ENVIRONMENTS IN HAMILTON**

By

HARINAD B. MAGANTI, B.Sc.

A Thesis

Submitted to the School of Graduate Studies in

Partial Fulfillment of the Requirements

for the Degree

Master of Science

McMaster University

©Copyright by Harinad B. Maganti, June 2011

McMaster University MASTER OF SCIENCE (2011) Hamilton, Ontario (Biology)

TITLE: Species and genotype diversities of yeasts in the clinical and natural environments in Hamilton

AUTHOR: Harinad B. Maganti B.Sc (McMaster University)

SUPERVISOR: Professor Jianping Xu

NUMBER OF PAGES: 108

ABSTRACT

In Canada the incidence of yeast infections have increased over the past decade, which in turn has resulted in the increased mortality and morbidity rates among the immuno-compromised patients. Yeasts are ubiquitous in nature and constitute a healthy portion of human skin and gut flora. Factors such as the urban settings and food have been previous shown to influence the yeast flora people harbour. This makes us believe that to effectively tackle the rising yeast infections in Canada we need to not only conduct epidemiological yeast studies in clinical settings but should also understand the diversity and distribution of them in the urban environment. This thesis constitutes of an epidemiological fungemia study and an urban environmental yeast profiling study conducted in the city of Hamilton.

In the first chapter of the thesis I discuss the results of the epidemiological candidemia study. We noticed that over the past decade the mean age of the population with candidemia in hospitals within Hamilton has increased by 10 years. DNA fingerprinting analysis suggested that 33% of the blood stream *Candida* isolates from January 2005 to February 2009 belonged to 18 clusters, some of which were shared between wards and hospitals. we found that for each of the four species, strains isolated closer to each other temporally were overall genetically more similar to each other as well, which suggested that nosocomial sources likely caused repeated candidemia infections. The study is the first of its sort in Canada and the results of this chapter are expected to aid infection control practitioners in the Hamilton hospitals and make the stay of patients in hospitals safer.

In the second chapter, we discuss the diversity and distribution of yeasts prevalent on trees in and around Hamilton. We identified a total of 88 environmental yeasts belonging to 20 species (based on ITS sequence data). The yeast populations were highly heterogeneous in both species and genotype composition. Among the 14 tree species sampled, yeasts were frequently found on cedar, cottonwood and basswood. Interestingly all the *Candida parapsilosis* strains were found from pine tree only. Some of the potential environmental factors shaping the distribution of yeast populations in Hamilton are discussed.

Preface:

Each chapter in this thesis has been written as an individual manuscript. Chapter 1 has been accepted by the journal of Medical Mycology, while chapter 2 has been submitted to the journal FEMS Yeast Research and we are currently addressing the reviewers' comments. Written preparation of the manuscripts, result analyses and interpretation of both chapters was done by Jianping Xu and Harinad Maganti. Deborah Yamamura provided the clinical blood stream isolates analysed in the first chapter. She also helped provide access to patient databases and supervised the patient chart data entry and chart analysis. David Bartfai helped in the sampling of yeast in the second chapter. He also provided the ITS sequences for most of the 88 yeast strains. He scored the PCR fingerprints independently of Harinad Maganti to ensure consistency.

Acknowledgements

This work would not have been possible without the valuable contributions of many individuals. Considering I have only acknowledged the funding agencies within the body of my chapters, I would like to mention the names of individuals who deserve particular recognition.

Primarily, I would like to mention the names of my graduate colleagues Amanda Wilson, Aaron Vogan, Jane Shen, Iqbal Satiadi and Hermina Ghenu who were all very helpful in relation to both my data analysis and in helping me deal with the insanity one comes across after long hours of lab work. Both these contributions played a significant role to the timely completion of my thesis.

Secondly, I would like to convey my sincere gratitude to the microbiology department staff members of Hamilton General Hospital for their help in the stocking, storage and transfer of blood stream *Candida* isolates to our research facility in the department of Biology at McMaster University.

Sreenivas Maganti, Subbamma Maganti (my parents) and Ram Maganti (my brother) were of inordinate value to my success. Former and current post doctorates: Dr Porter, Dr Haerty and Dr James played a significant role in helping me learn, how to use certain bioinformatic tools, statistical tools and lab techniques.

Finally and most crucially, I acknowledge the excellent tutelage provided by my supervisors Dr Xu and Dr Yamamura. I have been very fortunate to work with Dr Xu whose unique cognitive construct, time management skills, enthusiasm and sportsmanship have inspired me thoroughly during my stay as a graduate student at McMaster. He made pursuing research fun and has helped me learn the process of writing scientific manuscripts as well as provided me the rare opportunity of being a corresponding author on one of the manuscripts. I will always be indebted to him, for all the mentoring that he has provided me since undergrad that helped me make appropriate career choices. He has played a pivotal role in ensuring that the standard of my experience was set extraordinarily high.

Harinad B Maganti B. Sc.

McMaster University

June 2011

Table of Contents

Abstract	iii
Acknowledgements	iv
Table of Contents	v
List of Tables	x
List of Figures	xi
Background	
Ubiquitous nature of fungi	1
Environmental factors known to shape yeast distribution	1
Need for urban environmental studies	2
Nosocomial infections & their threat to Canada	2
Phylogenetic position of the genus <i>Candida</i>	3
Candidiasis disease classification: Invasive vs non-invasive	4
Epidemiology & limitations of current studies	4
Nosocomial transmission of <i>Candida</i> species	5
Scope of this thesis	7
References	8
Chapter 1: Prevalent nosocomial clusters among causative agents for candidemia in Hamilton, Canada	
Title Page	11
Abstract	12
Introduction	13
Methods	16

Strains	16
DNA extraction and genotyping	17
Data analysis	18
Discriminatory index	18
Test for temporal clustering of strains	18
Mantel Test	19
Results	20
Epidemiology	21
Strain clusters	21
Temporal genotype clustering of strains	22
Discussions	23
Acknowledgements	28
References	29
Appendix	34

Chapter 2: Ecological structuring of yeasts associated with trees around Hamilton, Ontario, Canada.

Title Page	62
Abstract	63
Introduction	64
Materials and methods	66
Sample collection	66
Sample processing	67

DNA-PCR amplification	67
DNA sequencing	67
Species identification	68
ITS Phylogeny construction	68
PCR Fingerprinting	68
PCR Fingerprinting data analysis	69
Phylogram construction and discriminatory index	69
Species richness comparison	70
Host tree preference analysis	70
Difference between deciduous and coniferous trees	70
Results	71
Species diversity of yeast populations at different geographical sites	71
Fingerprinting results	71
Discriminatory power of the markers	72
Yeast species richness based on geographical sites	73
Host tree – Yeast species correlation	73
Discussion	73
Acknowledgements	77
References	78
Appendix	83
Concluding statements and future directions	107

List of Tables

Table 1.1 Summary information for patient and yeast species demographics of the strains causing candidemia in Hamilton, Ontario, Canada from 2005 to 2009

Table 1.2 The discriminatory index (DI) of the PCR fingerprint pattern using three primers M13, (GACA)₄ and OPA3 across the four species: *C. albicans*, *C. tropicalis*, *C. glabrata*, and *C. parapsilosis*.

Table 1.3 Comparison of the mean pairwise genetic distance between strains isolated at different time intervals (those <90 d apart vs. those >90 d apart) within each of the four *Candida* species analyzed here.

Table 1.4 Nosocomial clusters observed in the clinical isolates causing candidemia in Hamilton, Ontario, Canada from 2005 to 2009.

Table 1.5: Mantel comparison of the pairwise genetic distance between strains isolated at different time intervals within each of the four *Candida* species.

Table 2.1 Summary information relating to the total number samples collected at each sampling site and percentage of them that turned out to be yeast.

Table 2.2 Summary information of yeast species and tree species identified at each sampling site and the number of yeast isolates for each species at each site.

Table 2.3 Summary of positive yeast samples and the total number of samples taken from each tree as well as the Z-score correlation between them.

Table 2.4 Fisher's exact test analysis showing tree type preference

List of Figures

Figure 1.1 Genetic relationships among independent strains of *C. albicans* (CA) in this study.

Figure 1.2 Genetic relationships among independent strains of *C. tropicalis* (CT) analyzed in this study

Figure 1.3 Genetic relationships among strains of *C. glabrata* (CG) analyzed in this study

Figure 1.4 Genetic relationships among strains of *C. parapsilosis* (CP) analyzed in this study

Figure 1.5 Banding patterns of independent strains present in nosocomial clusters using M13 and (GACA)₄ primers. Lanes 1 to 16 correspond to strains CA64, CA84 (a duplicate of CA64), CA85, CA86; CP92, CP94, CP102, CP105, CG123, CG124, CG131, CG133, CT115, CT116, CT121, CT122) respectively. Strains CP92, CP94 and CP102 and CP105 were obtained from the same patient sequentially at one week intervals.

Figure 1.6 The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 86 strains of *Candida albicans* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

Figure 1.7 The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 10 strains of *Candida glabrata* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

Figure 1.8 The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 11 strains of *Candida tropicalis* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

Figure 1.9 The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 17 strains of *Candida parapsilosis* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

Figure 2.1 Environmental sampling in the greater Hamilton region. The sites A-H sampled between May 26, 2008 and September 28, 2008 are indicated in the map.

Figure 2.2 ITS tree of all the 88 environmental yeast strains sampled. Each strain in the phylogram was named in the following manner: Strain identification number _ Species abbreviation _ tree species isolated from _ geographical site isolated from. * denotes reference sequences for each respective yeast species. The names of the tree species have been abbreviated in the following manner: Ash (As), Basswood (Bw), Birch (Bi), Cedar (Cd), Cherry (Ch), Chestnut (Cn), Cotton wood (Ct), Hawthorn (Ht), Maple (Mp), Oak (Ok), Pine (Pi), Sycamore (Sy), Raspberry (Rb) and Beech (Be). Detritus and bird droppings have been abbreviated as Di and Bd.

Figure 2.3 Phylogram of 9 *C.parapsilosis* (Cp) strains. 8 genotypes were identified and none of the genotypes were present at more than one site. All the *C.parapsilosis* strains were identified to grow only on Pine (Pi) trees. Each strain in the phylogram was named in the following manner: Strain identification number _ Species abbreviation _ tree species isolated from _ geographical site isolated from.

Figure 2.4 Phylogram of 9 *K.aerobia* (Ka) strains. 8 genotypes were identified and none of the genotypes were present at more than one site. 8 of the *K.aerobia* strains were identified to grow on Ash trees (As) while 1 strain was identified to grow on Sycamore (Sy). Each strain in the phylogram, was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Figure 2.5 Phylogram of 10 *K.lactis* (Kl) strains. 9 genotypes were identified and none of the genotypes were present at more than one site. *K.lactis* strains were identified to grow on many trees namely, Sycamore (Sy), Ash (As), Beech (Be), Cotton wood (Ct), Maple (Mp) and Cedar (Cd). *K.lactis* was also identified to grow on the shrub Raspberry (Rb). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Figure 2.6 Phylogram of 10 *T.delbrueckii* (Tm) strains. 10 genotypes were identified and none of the genotypes were present at more than one site. *T.delbrueckii* strains were identified to grow on many trees namely, Chest nut (Cn), Beech (Be), Cotton wood (Ct), Basswood (Bd) and Sycamore (Sy). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Figure 2.7 Phylogram of 17 *S.paradoxus* (Sp) strains. 17 genotypes were identified and none of the genotypes were present at more than one site. *S.paradoxus* strains were identified to grow on many trees namely, Cedar (Cd), Basswood (Bd), Oak (Ok), Cotton wood (Ct), Maple (Mp) and Beech (Be). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Figure 2.8 Phylogram of 7 *D.hanseniii* (Dh) strains. 7 genotypes were identified and none of the genotypes were present at more than one site. *D.hanseniii* strains were identified to grow on many trees namely, Ash (As), Hawthorn (Ht), Cotton wood (Ct), Basswood (Bd) and Maple (Mp). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Background

Ubiquitous nature of fungi

Yeasts are eukaryotic microorganisms belonging to the kingdom of Fungi. As of 2006 researchers had described ~1500 yeast species which are estimated to constitute about 1% of species present in the kingdom of Fungi (Kurtzman and Fell, 2006; Kurtzman and Piškur, 2006). Please note that considering all yeasts are fungi, I use the words yeast and fungi interchangeable in this thesis document. Yeasts are ubiquitous in nature. They are commonly found on sugar rich materials such skin of grapes, apples, peaches, exudates from plants and rotting wood. They have also been found in marine environments. In addition yeasts have also been identified to play an integral role in the beetle's ability to utilize wood as a part of its diet. The ability to degrade complex materials (such as lipids, hemicelluloses and xylan) is believed to have partly contributed to their broad range distribution (Oyeka and Ugwu, 2002; Sláviková and Vadkertiová, 2003; Suh *et al.*, 2005).

Environmental factors known to shape yeast distribution

Many of the yeast profiling studies conducted around the world have been done in non-urban plantation or forested settings. A study conducted in the north-western area of India and southern state of Tamil Nadu concluded that the trees *Azadirachta indica*, *Cassia fistula*, *Mimusops elengi*, *Manilkara hexandra* and *Syzygium cumini* demonstrated a high rate of recovery for *Cryptococcus gattii* and *Cryptococcus neoformans* both of which are opportunistic fungal pathogens that cause human infections (Randhawa *et al.*, 2008). They also observed soil to be an important ecological niche for yeasts, along with both salt water and fresh water environments. These findings helped to create a more comprehensive picture of the dispersal of both these human pathogenic species and demonstrated the ability of the yeasts to spread along

waterways and forestry activities (Randhawa *et al.*, 2008). Other factors impacting yeast distribution include pH, electrical conductivity of the soil, and rainfall (Vishniac, 2006).

Need for urban environmental studies

The ubiquitous nature of yeasts makes humans to always be in close contact with them. In humans yeast species such as *Candida albicans*, *Rhodotorula rubra*, *Trichosporon cutaneum* have all been previously identified to constitute a significant portion of their skin flora (Oyeka and Ugwu, 2002). In addition to skin, species like *C. albicans* are also commonly found in the human gut and mouth (Martini, 1992; Bass *et al.*, 2007; Fauci *et al.*, 2008; Kutty and Philip, 2008). Even though yeasts are commonly commensals, they are known to cause invasive fungal infections among the immuno-compromised (e.g. HIV infected) and immuno-deficient patients (e.g. neonates and old aged) (Fauci *et al.*, 2008). The fact that humans (irrespective of whether they are immuno-compromised or deficient) are in close contact with many opportunistic yeasts, some of which are regularly consumed along with natural food products (fruits, milk) makes us believe that if we are to effectively tackle invasive yeast infections, it is important to gain a comprehensive understanding of yeast distributions in not only clinical settings but also in urban environmental settings.

Nosocomial fungal infections and their threat to Canada

Nosocomial invasive fungal infections are a threat to hospitalized patients worldwide. Infections in hospital patients are defined as nosocomial if they appear within the first 48 hours or more after hospital admission or within 30 days after discharge. Recent studies have shown the incidence of nosocomial fungal infections to be on the rise over the past decade (Asmundsdottir *et al.*, 2002; 2005; 2008). Since many human pathogenic fungi are ubiquitous in nature, controlling the infections caused by them has often been very tough in clinical settings

and costs millions of dollars (Yamamura *et al.*, 1999; Asmundsdottir *et al.*, 2002; Chowdhary *et al.*, 2003; Asmundsdottir *et al.*, 2005; Asmundsdottir *et al.*, 2008; Rotstein, 2008).

Considering the aging population of Canada, this means that over the next few years more people will be at risk to invasive fungal infections and as such it is very important to study and control the rise of these infections. Apart from age, pharmaceutical agents (ex: antibacterials, parenteral glucocorticoids, cytotoxics), indwelling catheters (ex: intravascular, urinary), hyperalimentation fluids, respirators and neutropenia have all been previously noticed as aetiological factors triggering nosocomial invasive fungal infections (Asmundsdottir *et al.*, 2002; 2005; 2008; Fauci *et al.*, 2008).

Phylogenetic position of the genus *Candida*

The genus *Candida* was first taxonomically described in 1923 by the botanist Christine Marie Berkhout (Gantner and Verlag, 2006). Since then this genus has been identified to contain some of the most clinically prevalent fungal pathogens responsible for causing nosocomial fungemia among the immuno-compromised and immuno-depressed. Though this polyphyletic genus possesses over 50 human pathogenic fungal species, traditionally researchers have observed *C. albicans*, *C. parapsilosis*, *C. glabrata*, *C. metapsilosis*, *C. orthopsilosis*, *C. krusei*, and *C. guilliermondii* to be the most clinically prevalent *Candida* species with *C. albicans* causing the bulk of infections (Fauci *et al.*, 2008; Miceli *et al.*, 2011). Historically, the genus *Candida* has been known as the “taxonomic pit” into which many asexual fungi have been deposited. Its members are biologically diverse including yeasts with ascomycetous and basidiomycetous affinities. Typically, under a microscope *Candida* species have been noticed to be between 4-6 μm in diameter and some can exist in four morphological forms: blastospores,

pseudohyphae, hyphae and a unicellular vegetative stage (McCullough *et al.*, 1996; Miceli *et al.*, 2011).

Candidiasis disease classification: Invasive vs non-invasive

The infection caused by *Candida* species is termed as candidiasis. *Candida* infections are broadly characterized into two categories: non-invasive candidiasis and invasive candidiasis. Non-invasive candidiasis is defined as the infection of the mucosal and cutaneous layers (Fauci *et al.*, 2008). A common example of mucosal infection is oral candidiasis (commonly called as thrush) while an example of the cutaneous layer would be the infection of sub-mammary areas. Together these two types of non-invasive infections are termed as muco-cutaneous infections. Non-invasive infections are more common than invasive infections; they are often easy to treat and are very seldom lethal. The penetration of yeast through the muco-cutaneous layers leading to infections in the underlying surfaces such as blood, urinary tract and gastro-intestinal (GI) tract symbolize invasive candidiasis. Invasive candidiasis is often hard to treat and if left untreated, easily spreads through the body (with blood acting as the carrier) often causing multi-organ system failure (Fauci *et al.*, 2008). The medical term used to define the presence of yeast belonging to the genus *Candida* in blood is candidemia and it has been on the rise around the world (Fauci *et al.*, 2008; Asmundsdottir *et al.*, 2005; 2008).

Epidemiology & limitations of current studies

In Canada the incidence of infection by *Candida* has increased from 1.2 to 5.1/100000 over the past decade (Yamamura *et al.*, 1999; Maganti *et al.*, 2011). One of the few places in the world where extensive candidemia data has been recorded and analyzed is Iceland. A 20 year cross-sectional study conducted there had shown hospital acquired candidemia to rise by four

folds in between 1980-1999 (Asmundsdottir *et al.*, 2005). In 2008 Asmundsdottir *et al.* (2008) showed that the risk of acquiring candidemia varied among the different wards and that it was highest in the intensive care unit (ICU) wards. They also showed that candidemia in one third of the hospital admitted patients of Iceland between 1991 and 2006 was triggered by nosocomial clusters [closely related isolates ($\geq 90\%$ identical fingerprinting pattern) from ≥ 2 patients in the same ward or at the same hospital within a period of 90 days] (Asmundsdottir *et al.*, 2002; 2005; 2008). Though candidemia related studies have been conducted in Canada, most have only focused on the etiological factors causing candidemia, antifungal resistance and its impact on species composition (Yamamura *et al.*, 1999; Rotsein, 2008). There has been very little information published on the relative contributions of nosocomial clusters and nosocomial transmission to candidemia in Canada as well as in most other parts of the world (Table 1).

Nosocomial transmission of *Candida* species

Nosocomial transmission is defined as the transmission of an infection causing pathogen from one part of hospital to another. Some of the previously noticed means of transfer were medical infusates, health personnel (white coats and nails), blood culture bottles and milk products (Ruiz-Diez *et al.*, 1997; Asmundsdottir *et al.*, 2002; 2005; 2008; Miceli *et al.*, 2011). Some *Candida* species (*C. kefyr*, *C. lusitaniae*) have been identified to be prevalent in soil, raising speculations that they (*Candida* species) can potentially be transferred via shoes, brooms and mops. This versatility found in the adaptation of *Candida* species is facilitated by their tight adhering nature on surfaces which in turn aids them in causing infections. The fact that *Candida* species are present on both animate and inanimate objects in diverse habitats has made clinicians often wonder whether the sources of hospital acquired cases of candidiasis are actually nosocomial in nature or not; if they are truly indigenous or are actually exogenous to hospitals.

Many infection control practitioners around the world have reported the problem of not being able to differentiate between exogenous environmental *Candida* strains and indigenous hospital *Candida* strains. The major obstacle/cause for this is the lack of enough molecular based urban environmental profiling studies that report intra/inter species variation/diversity of *Candida* populations present in the natural environments. Although researchers in the past have often reported changes in *Candida* population dynamics (for example, many studies have reported the rise of non-albican *Candida* species) very few of them have explored the intra species variation present within the different *Candida* populations (clinical versus non-clinical) and the potential factors shaping this change (Miceli *et al.*, 2011). The lack of such information has restricted infection control practitioners from accurately identifying whether the candidemia infections acquired by the admitted patients are actually being caused by the strains prevalent in the hospital or are by commensal flora present on their body (that they bring with them to the hospitals). This in turn prevents them from laying down fool-proof infection control guidelines. This is particularly applicable to the city of Hamilton, where the last candidemia based study was conducted in the early 1990's. Furthermore it is important to note that there have been no environmental profiling studies looking at the urban environmental yeast flora and the factors shaping the yeast community (Yamamura *et al.*, 1999).

Scope of this thesis

Since the last retrospective candidemia study involving Hamilton in Canada was conducted in the early 1990's and the fact that there haven't been any molecular fingerprinting studies attempted previously in Canada to deduce the impact of nosocomial clusters in causing candidemia, the first objective of this project is to study this in the Hamilton Health Sciences

(HHS) hospitals located in the city of Hamilton which has a sizable population of 550,000. The HHS hospitals are reported to annually serve over 3 million people and we believe the conclusions unearthed through this objective would in the long run save the HHS hospitals millions of dollars and make the stay of patients admitted in the hospitals safer. It will also provide vital information that could be applicable in other areas of nosocomial septicemia.

The second focus of the thesis was to study the prevalence and diversity of yeasts on trees in the urban environments of Hamilton. We are interested in identifying how many human pathogenic yeasts are associated with trees and how large the environmental repertoire of the various *Candida* species is. We are also interested in studying the dispersal pattern of yeast populations in their natural environments. If there is a large reservoir of *Candida* species we are interested in comparing their (environmental strains) genotypes to those of the invasive *Candida* samples as this will help answer the question of whether the source of infections in hospitals are exogenous or endogenous to the hospitals.

References

- Ásmundsdóttir LR, Erlendsdóttir H, Gottfredsson M. (2005) Improving survival of patients with candidaemia: analysis of prognostic factors from a long-term, nationwide study in Iceland. *Scand J Infect Dis* 37: 111–120.
- Ásmundsdóttir LR, Erlendsdóttir H, Gottfredsson M. (2002) Increasing incidence of candidemia: results from a 20-year nationwide study in Iceland. *J Clin Microbiol* 40: 3489–3492.
- Ásmundsdóttir LR, Erlendsdóttir H, Haraldsson, G, *et al.* (2008) Molecular epidemiology of candidemia: evidence of clusters of smoldering nosocomial infections. *Clin Infect Dis* 47: 17–24.
- Bass D, Howe A, Brown N, Barton H, Demidova M, Michelle H, Li L, Sanders H, Watkinson SC, Willcock S, Richards TA. (2007) Yeast forms dominate fungal diversity in the deep oceans. *Proceedings. Biological Sciences/The Royal Society* 274 (1629): 3069–3077.
- Chowdhary A, Becker K, Fegeler W, *et al.* (2003) An outbreak of candidemia due to *Candida tropicalis* in a neonatal intensive care unit. *Mycoses*; 46: 287–292.
- Clark TA, Slavinski SA, Morgan J, *et al.* (2004) Epidemiologic and molecular characterization of an outbreak of *Candida parapsilosis* bloodstream infections in a community hospital. *J Clin Microbiol* 42: 4468–4472.
- Fauci, Braunwald, Kasper, Hauser, (2008). *Harrison's Principles of internal medicine (Voll)*. Toronto: McGraw Hill.
- Gantner ARG, Verlag KG.(2006) International Code of Botanical Nomenclature (Vienna Code). *Regnum Vegetabile* 146.
- Kurtzman, Fell (2006). *Yeast Systematics and Phylogeny - Implications of Molecular Identification Methods for Studies in Ecology*, The Yeast Handbook, Springer.

- Kurtzman, Piškur, (2006). Taxonomy and phylogenetic diversity among the yeasts. Berlin: Springer.
- Kutty, SN, Philip R. (2008) Marine yeasts—a review. *Yeast* 25(7): 465.
- Maganti H, Yamamura D, Xu J (2011) Prevalent nosocomial clusters among causative agents for candidemia in Hamilton, Canada. *Med Mycol* DOI:10.3109/13693786.2010.547880.
- Martini A. (1992) Biodiversity and conservation of yeasts. *Biodiversity and Conservation* 1: 324–333.
- McCullough MJ, Ross BC, Reade PC. (1996) *Candida albicans*: a review of its history, taxonomy, epidemiology, virulence attributes, and methods of strain differentiation. *Int. J. Oral Maxillofac. Surg* 25: 136-144.
- Miceli HM, Diaz AJ, Lee AS. (2011) Emerging opportunistic yeast infections. *Lancet Infect Dis* 11: 142-151.
- Oyeka, CA, Ugwu LO. (2002) Fungal flora of human toe webs. *Mycoses* 45(11-12): 488–491.
- Randhawa HS, Kowshik T, Chowdhary A, Preeti Sinha K, Khan ZU, Sun S, Xu J. (2008) The expanding host tree species spectrum of *Cryptococcus gattii* and *Cryptococcus neoformans* and their isolations from surrounding soil in India. *Med Mycol* 46(8): 823-833.
- Rotstein C. (2008) Invasive candidiasis in the ICU: prophylaxis versus preemptive treatment. *Curr Infect Dis Rep* 10(6): 454-458.
- Ruiz-Diez B, Martinez V, Alvarez M, Rodriguez-Tudela JL, Martinez-Suarez JV. (1997) Molecular tracking of *Candida albicans* in a neonatal intensive care unit: long-term colonizations versus catheter-related infections. *J Clin Microbiol* 35: 3032–3036.

- Sláviková E, Vadkertiová R. (2003) The diversity of yeasts in the agricultural soil. *Journal of Basic Microbiology* 43(5): 430–436.
- Smith PB, Morgan J, Benjamin JD, *et al.* (2007) Excess costs of hospital care associated with neonatal candidemia. *Pediatr Infect Dis J* 26: 197–200.
- Suh SO, McHugh JV, Pollock DD, Blackwell M. (2005) The beetle gut: a hyperdiverse source of novel yeasts. *Mycological Research* 109 (3): 261–265.
- Vishniac HS. (2006) A Multivariate Analysis of Soil Yeasts Isolated from a Latitudinal Gradient. *Microb Ecol* 50 (2006): 90-103.
- Yamamura D, Rotstein C, Nicolle L, Ioannou S, and The fungal disease registry of The Canadian Infectious Disease Society. (1999) Candidemia at selected Canadian sites: results from the Fungal Disease Registry, 1992-1994. *CMAJ* 160:493–499.

Prevalent nosocomial clusters among causative agents for candidemia in Hamilton, Canada

Harinad Maganti¹, Deborah Yamamura^{2,3}, Jianping Xu^{1,2}

¹Department of Biology, McMaster University, 1280 Main St W, Hamilton, ON, L8S 4K1, Canada;

²Institute of Infectious Diseases Research, Michael G. DeGroot School of Medicine, McMaster University, 1280 Main St W, Hamilton, ON, L8S4L8, Canada,

³Hamilton Regional Laboratory Medicine Program, 237 Barton Street East, Hamilton, ON, L8L 2X2. Canada

Short Title: Candidemia nosocomial clusters

Corresponding author: Dr. Jianping Xu, Department of Biology, McMaster University, 1280 Main St. West, Hamilton, Ontario, L8S 4K1, Canada. Phone: 1-905-525-9140 ext. 27934. Fax: 1-905-522-6066. Email: jpxu@mcmaster.ca

Abstract

We investigated the prevalence of nosocomial clusters among the causative agents for candidemia, the bloodstream infection caused by strains of the yeast genus *Candida*, in Hamilton, Ontario, Canada during a period from January 2005 to February 2009. In Canada, the incidence of candidemia over the years varied from 1.2-5.1 cases/100,000 and *Candida* species were the third most common agents of bloodstream infections in intensive care units. However, the relative contributions of nosocomial transmission to candidemia remain poorly understood. In this study, we genotyped 134 isolates from 125 unrelated patients with candidemia. Among the 134 isolates, there were 87 *C. albicans*, 20 *C. parapsilosis*, 11 *C. glabrata*, 15 *C. tropicalis*, and one *C. krusei*. Our PCR fingerprinting analyses using three primers of highly polymorphic sites identified a total of 99 genotypes with 18 genotypes shared by 44 independent isolates. Nine pairs of isolates were obtained from the same patients at the same time and each pair had an identical fingerprint. Interestingly, all strains belonging to each of the shared genotypes were isolated from patients within 3-months stay in the Hamilton hospitals. Both inter- and intra-ward clusters were found, including one that contained strains from intensive care units in two hospitals. Our results indicated that 33% of the patients with candidemia were infected by nosocomial clusters and suggested that measures should be taken in hospitals to prevent nosocomial acquisition of *Candida* infections.

Introduction

The fungal genus *Candida* includes over 200 species. These organisms are ubiquitous in nature and are commonly found on inanimate objects, foods, plants, and animals. Many *Candida* species are normal commensals of humans, frequently inhabiting the oral mucosal surface, the gastrointestinal tract, the urogenital tract, and the skin (Rentz *et al.*, 1998; Gudlaugsson *et al.*, 2003; Hajjeh *et al.*, 2004). However, many of these organisms are capable of causing infections, primarily in patients with compromised immunity (Rentz *et al.*, 1998; Hajjeh *et al.*, 2004). Indeed, it has been long known that *Candida* infection, or candidiasis, is very common among debilitated patients (Banerjee *et al.*, 1991; Asmundsdottir *et al.*, 2002; Asmundsdottir *et al.*, 2005; Asmundsdottir *et al.*, 2008). Candidiasis is broadly classified into muco-cutaneous candidiasis and invasive candidiasis. Candidemia, the focus of this study, is a common form of invasive candidiasis and is defined as the presence of yeast belonging to genus *Candida* in blood. The common risk factors for invasive candidiasis include HIV infection, immuno-suppressive therapy, recent surgery, presence of central line, prolonged stay in the intensive care unit, and for small children, a very low birth weight (Banerjee *et al.*, 1991; Rentz *et al.*, 1998; Asmundsdottir *et al.*, 2002; Gudlaugsson *et al.*, 2003; Hajjeh *et al.*, 2004; Asmundsdottir *et al.*, 2005; Asmundsdottir *et al.*, 2008). Over the past 10 years, the incidence of candidemia in hospital intensive care units (ICU) has increased significantly in some regions of the world, and nosocomial infections have been hypothesized as a major contributor (Rentz *et al.*, 1998; Gudlaugsson *et al.*, 2003; Hajjeh *et al.*, 2004).

Strains causing nosocomial infections could be from two sources. In the first, these strains could originate from endogenous strains brought in to the hospital by patients themselves. In the second, the strains could originate from nosocomial transmission from the hospital

environment. Several studies have implicated nosocomial transmission from a common source as a cause of candidemia, with potential sources including contaminated infusates, the hands of health-care workers, hospital food, and medical devices (Strausbaugh *et al.*, 1994; Pfaller *et al.*, 1995; Marco *et al.*, 1999; Asmundsdottir *et al.*, 2002; Asmundsdottir *et al.*, 2008; Van Asbeck *et al.*, 2007). If the strains were from an endogenous source and patients were unrelated, these strains should be genetically different from each other without any spatial and temporal clustering. In contrast, if the strains were from a common source from within the hospital environment, clusters of strains with similar or identical genotypes should be found and these strains should show spatial and temporal clustering.

With the development of modern molecular techniques for genotyping, it's now possible to examine the relationships among strains and to identify whether strains are from the same or difference sources (Pfaller *et al.*, 1996; Finkelstein *et al.*, 2003). Indeed, with the aid of discriminatory molecular markers, genotyping has enabled investigators to identify and cluster strains as well as to set apart these clustered strains from one another based on temporal and geographical factors (Thanos *et al.*, 1996; Pujol *et al.*, 1997; Soll, 2000; Xu *et al.*, 2000; Robles *et al.*, 2004). In this study, a nosocomial cluster is defined as the isolation of strains sharing identical genotypes at the examined markers in samples obtained from ≥ 2 patients in the same hospital within a period of 90 days (Gudlaugsson *et al.*, 2003).

A previous retrospective study conducted by Yamamura *et al.* (1999) demonstrated that between 1992 and 1994, 88.4% of the patients with candidemia across Canada were adults. They found that the majority of the clinical isolates (69%) belonged to one species *C. albicans* and patients > 60 years old who stayed in the ICU were at the highest risk. However, no genotypic information was available for that set of strains. In a recent epidemiological study performed in

Iceland, Ásmundsdóttir *et al.* (2008) demonstrated that 33% of candidemia could be attributed to nosocomial clusters. Their conclusion was based on DNA fingerprinting through PCR using arbitrary primers. PCR fingerprinting has been observed to have a high discriminatory power for related and unrelated isolates of clinically relevant *Candida* species (Thanos *et al.*, 1996; Pujol *et al.*, 1997; Soll, 2000; Xu *et al.*, 2000; Robles *et al.*, 2004). Ásmundsdóttir *et al.* (2008) concluded that the risk of candidemia in hospitals varied among wards, with the highest being in ICUs. Since the last retrospective study on candidemia in Canada was performed in the early 1990s and there is little information on the relative roles of nosocomial clusters in causing candidemia in Canada, we sought to investigate the prevalence of nosocomial clusters in Hamilton, Ontario, Canada.

We investigated whether nosocomial clusters for strains causing candidemia were present in three tertiary care hospitals (part of the Hamilton Health Sciences, HHS) in Hamilton, Ontario, Canada, and if so, their prevalence within and among wards as well as within and among hospitals. We were also interested in the potential changes in the distribution of *Candida* species causing candidemia over the 10 years since the last epidemiological survey was conducted in the 1990s in this region (Yamamura *et al.*, 1999). Several recent studies (Rentz *et al.*, 1998; Asmundsdottir *et al.*, 2002; Gudlaugsson *et al.*, 2003; Hajjeh *et al.*, 2004; Asmundsdottir *et al.*, 2005) have shown a growing incidence of non-*C. albicans* *Candida* species causing candidemia. Whether a similar shift occurred in this region is unknown. Information from this study should help us identify the importance of nosocomial transmission and help develop strategies to prevent the increasing incidence of nosocomial candidemia.

Materials and Methods

Strains

Candida isolates were obtained from patients with candidemia admitted to Hamilton Health Sciences (HHS) hospitals from January 1, 2005 to February 28, 2009. Blood cultures were incubated using the BacTAlert system (bioMérieux Canada Inc, St. Laurent, Quebec). Isolates were identified to the species level with the use of standard morphological and physiological tests in clinical microbiology labs. Specifically, strains were identified to the species level based on their germ tube formation ability, a set of macroscopic and microscopic features on bile oxgall agar, and by physiological characteristics determined using the Vitek Legacy YBC (2005-2007), Vitek 2 YST card (2007-present), and the API 20C AUX systems as required. The Vitek systems and the API 20 kits were purchased all from bioMérieux Canada Inc (St. Laurent, Quebec).

The clinical history for each of these patients was retrieved retrospectively using a standardized form by two of the investigators (HM, DY). Patient's age, sex, the ward location within the admitted hospital, underlying diseases, treatment regimes for invasive candidiasis, dates of the first *Candida*-positive blood culture as well as subsequent dates with positive blood cultures were all recorded. Following traditional guidelines, an episode of candidemia was defined as ≥ 1 blood culture positive for *Candida* species and episodes that happened 30 d apart or were caused by different species of *Candida* were considered as separate cases of candidemia (Asmundsdottir *et al.*, 2008).

DNA Extraction and Genotyping

Isolates were stored at -80°C at the Microbiology Laboratory of Hamilton General Hospital. In total, there were 161 isolations of *Candida* strains from blood cultures during the study period, of which 152 were stored. Among the 152 stored isolates, 134 were viable and used for subsequent genotypic analyses. To obtain DNA for genotyping through PCR fingerprinting, all isolates were first streaked onto yeast extract–peptone-dextrose agar (Fisher Scientific, Fair Lawn, New Jersey) and incubated 48 h at 33°C. Genomic DNA was extracted from the actively growing yeast cells using protoplasting buffer and lysing buffer, followed by Iso-propanol precipitation as described in Xu *et al.* (Xu *et al.*, 2000). The extracted DNA was stored at -20°C and diluted to working concentration (~50-100 ng/ml) before being amplified by PCR using the following three separate primers: (GACA)₄ (5'-GACAGACAGACAGACA-3'), OPA3 (5'-AGTCAGCCAC-3'), or M13 (5'-GTAAAACGACGGCCAGT-3') (Asmundsdottir *et al.*, 2002; Asmundsdottir *et al.*, 2005; Asmundsdottir *et al.*, 2008). Each standard PCR contained 8 µl of Ready-to-Go-PCR mix (Promega, Madison, Wisconsin), 4 µl of working concentration of template DNA, and 4 µl of 0.10 µM of the desired primer. The amplified products were subjected to gel electrophoresis on 1% agarose gels with 1 x Tris-Borate-EDTA buffer (pH 8) for 3 h at 100V. Ethidium bromide was added to the gels prior to loading the PCR products and upon completion of the electrophoresis, the gels were photographed digitally under ultraviolet light using a Chemi-Imager (Alpha InnovTech Corporation, San Leandro, California). Representative strains from each of the analyzed species were tested multiple times for PCR and gel electrophoresis to check for reproducibility of the PCR-fingerprinting patterns.

Data Analysis

All the amplified DNA bands for all samples were scored. Similar to previous studies, a position tolerance setting of 2% was used (Soll, 2000; Asmundsdottir *et al.*, 2002). Strains with identical fingerprint patterns for all analyzed primers were considered to have the same genotype and belong to the same clone. To ensure consistency and minimize scoring errors, the PCR fingerprinting patterns were scored two times, each time by two different people. These genotypes were then compared to patient history to identify nosocomial clusters. Here, we define a nosocomial cluster as a group of isolates with an identical genotype obtained from ≥ 2 patients in the same hospital system within a period of 90 d. Our definition is more stringent than those used in previous studies, which employed 95% of PCR fingerprint identity as a cutoff for nosocomial clusters (Soll, 2000; Asmundsdottir *et al.*, 2002). The phylogenetic analysis program PAUP* (Swofford, 2003) was used to construct a phenogram illustrating strain relationships and identifying clusters of strains.

Discriminatory index

Following PCR amplification and DNA band scoring, phylograms were constructed independently for the fingerprinting pattern of each primer using PAUP* (Swofford, 2003). The discriminatory index (DI) for each primer for each species was calculated using the Simpson's Diversity Index, following previous studies (Xu *et al.*, 2000; Gudlaugsson *et al.*, 2003).

Test for temporal clustering of strains

Temporal clustering of strains was assessed via two additional tests that examined the relationship between pairwise genetic distances of strains and their times of isolation. These tests were conducted for each of the four *Candida* species where multiple strains were isolated. Here,

the pairwise genetic distance between each pair of strains was calculated based on band sharing using the program PAUP* (Swofford, 2003), the temporal difference between their times of isolation for each pair of strains (i.e. number of days separating the isolation time of the two strains) was derived from the retrieved patients' records.

Using the obtained genetic distance data and the temporal time difference data, the first test examined Pearson's correlation co-efficient between these two variables. This analysis was conducted separately for each species. In the second test, pairs of isolates were divided into two groups, one group consisting of strain pairs that were obtained less than 90 d from each other and the other group including strains separated by greater than 90 d of each other. The mean genetic distances and standard deviations were calculated for each of the two groups for each *Candida* species. A t-test was conducted to determine the statistical significance of the observed difference between the two groups using the software package SPSS version 14 (SPSS Inc, Chicago, Illinois).

Mantel Test

The Mantel test was conducted to study the relationship between the pairwise strain genetic distance and their time differences in date of isolation. In the Mantel test, the first column contained pairwise genetic distance between strain pairs, while the second column contained the difference in the time of isolation between the strain pairs (in days). The distance matrixes were generated for each of the columns and the R and P values were calculated using GenAlEx 6.41 (Peakall and Smouse, 2006). This was done independently for each of the 4 *Candida* species. The null hypothesis of the test was that there is no correlation between the strain variability and date of isolation, among each of the *Candida* species.

Results

Based on our retrospective data, 161 isolates associated with candidemia were recorded, of which, 29 were duplicates (i.e., isolated from the same body site in the same patient at the same time). Thus, there were 132 episodes of candidiasis at HHS during that period and 132 independent isolates (Table 1.1). Among these 132 episodes, 89 were caused by *C. albicans*, 19 by *C. parapsilosis*, 12 by *C. tropicalis*, 10 by *C. glabrata*, and 2 by *C. krusei* (Table 1.1). *C. albicans* was the most prevalent species causing candidemia in the HHS hospital environment, accounting for 67% of the total yeast isolates (Table 1.1).

Among the 161 isolates, 134 isolates were successfully revived and further analyzed. Nine of the 134 isolates were duplicates. All nine pairs showed the same PCR fingerprinting profile. In order to minimize the effects of duplicate isolates on our analyses, only one representative from each of the 9 strain pairs were included in the calculations of discriminatory index and nosocomial clusters. Among these 125 isolates representing independent episodes of candidemia, there were 86 *C. albicans*, 16 *C. parapsilosis*, 10 *C. glabrata*, 12 *C. tropicalis*, and one *C. krusei*.

Of the three primers used, M13 had the highest discriminatory power (DI) in identifying genotypes of *C. albicans*. However, OPA3 possessed the highest DI for identifying the genotypes of the three other species that contained multiple strains (Table 1.2). When DNA fingerprint results from all three primers were analyzed together, their strain genotype DIs for *C. albicans*, *C. tropicalis*, *C. parapsilosis* and *C. glabrata* were 0.92, 0.9, 0.83 and 0.83, respectively (Table 1.2). The average DI including all the markers was 0.87 for the four species analyzed here. Our PCR fingerprinting results were consistent with species identifications of the

125 isolates based on morphological and physiological characters. Clear differences in the banding patterns were observed among strains belonging to the five *Candida* species.

Epidemiology:

Of the 132 patients with candidemia, 63 were male and 69 were female. The mean age was 62 years (range 6 months to 87 years) with a higher age in females than males (68 vs 55 years). The mean age of candidemia patients increased by 10 years compared to that from the previous survey (mean age 52 years) from this region (Yamamura *et al.*, 1999). The most prevalent risk factors for the current study population were receipt of antimicrobial therapy (90%) and presence of a central line (68.2%).

Strain clusters:

Among the 86 strains of *C. albicans*, a total of 69 genotypes were identified. Nine of these genotypes were each shared by two or more strains while the remaining 60 genotypes were represented by one strain each (Figure 1.1). Among the nine shared genotypes, the most common was shared by five strains, two were shared by four strains each, one by three strains, and five by two strains each. Of these nine clusters, three were found in the ICU of one hospital. In addition, two were simultaneously present in the ICU of two hospitals. Strains CA05 and CA33 were found at hospital A, while CA34 and CA43 were found at hospital B (Figure 1.1). *C. albicans* clusters were observed to exist at all hospital sites.

Among the 12 independent *C. tropicalis* strains analyzed, ten different genotypes were found. Two of these genotypes were shared by two strains each (Figure 1.2). The clusters of *C. tropicalis* were all from one hospital. For *C. glabrata*, seven genotypes were observed among the 10 strains (Figure 1.3). Of the seven observed genotypes, three were shared by two strains each

(Figure 1.3). All three clusters were present in one hospital, with one cluster found in the ICU ward. Finally, amongst the 18 *C. parapsilosis* strains analyzed here, 12 genotypes were identified (Figure 1.4). Of the 12 genotypes, two were shared by two strains each and two were shared by three strains each (Figure 1.4). These three clusters of *C. parapsilosis* were found in two different hospitals. It should be noted that these 18 strains in Figure 1.4 were from 15 patients with four strains (CP92, CP94, CP102 and CP105) all from one patient but they were isolated one week apart from each other during a four-week span. Strains CP92 and CP94 were from the first two weeks and strains CP102 and CP105 were from the later two weeks. The differential clustering of strains CP92, CP94, CP102 and CP105 from one patient foreshadowed a single case of *C. parapsilosis* strain replacement and represented two independent episodes of candidemia (Figures 1.4 and 1.5).

Temporal genotypic clustering of strains

The mean difference between pairs of strains in their banding patterns was 16% for *C. albicans*, 44% for *C. glabrata*, 28% for *C. parapsilosis* and 32% for *C. tropicalis* (Table 1.3). These results suggest abundant diversity within each of the species and that the observed intra- and inter-ward clusters within hospitals and the inter-hospital strain clusters were neither due to the lack of polymorphisms for the markers nor due to the lack of diversity in the population samples (Table 1.3). The positive correlation observed between pairwise genetic distance and the pairwise temporal time differences across the four-year study period for all the four *Candida* species is consistent with the hypothesis that the hospital environment plays a significant role in candidemia in Hamilton, Ontario, Canada (Figures 1.6-1.9). This was also supported by the Mantel test (Table 1.5).

Discussion

In this study, *C. albicans* was found to be the most prevalent species isolated from patients with candidemia in Hamilton, Ontario, followed by *C. parapsilosis*, *C. tropicalis* and *C. glabrata* (Table 1.1). This result was largely consistent with several recent studies that showed an increasing proportion of non-*C. albicans* *Candida* species causing candidemia, likely due to increased azole treatments that select for intrinsically more resistant species (Lewis and Klepser, 1999; Ásmundsdóttir, 2002; Gottfredsson *et al.*, 2003; Gudlaugsson *et al.*, 2003; Ásmundsdóttir *et al.*, 2008; Leroy *et al.*, 2009). However, none of the patients in our study had prior history of candidemia and very few had significant exposure to azoles. The lack of a prolonged selective pressure likely contributed to the slightly higher prevalence of *C. albicans* in our patient population than in most other samples from other countries/regions. Our observations were very similar to the recent studies conducted by Laupland *et al.* (2005) in Calgary, Alberta, Canada and by St. Germain *et al.* (2008) in Quebec, Canada.

The relative proportions of different yeast species causing candidemia as revealed here were also similar to those observed by Yamamura *et al.* (1999) for the 1992-1994 surveyed samples in Canada. In that study, *C. albicans* accounted for 68.9% of all cases of candidemia in Canada, followed by *C. parapsilosis* (10.4%), *C. glabrata* (8.2%), and *C. tropicalis* (6.5%). In our study, the proportions of *C. albicans*, *C. parapsilosis*, *C. glabrata*, and *C. tropicalis* were 67.4%, 14.3%, 7.5% and 9.1% respectively. Different from the previous retrospective study conducted by Yamamura *et al.* (1999), the mean age of patients with candidemia in our study increased by about 10 years over the last decade. Since studies by Laupland *et al.* (2005) and St. Germain *et al.* (2008) conducted elsewhere in Canada also indicated a gradual increase in the mean population age when compared to Yamamura *et al.* (1999), this trend likely reflected the

relative increase of the negative health effects suffered for seniors over that of the general population in Canada. For example, a large number of the patients admitted to hospitals in this study were due to fractures (especially women) or solid tumors that often needed surgeries. Such surgeries were often accompanied by pain-relieving drugs and the application of broad-spectrum antibiotics, both of which could favor the growth of opportunistic fungal pathogens such as *Candida* (Prentice, 2009).

Previous studies conducted in Iceland focused on intra-ward strain clusters and showed that nosocomial transmission played a major role in causing candidemia in wards such as the ICU (Ásmundsdóttir *et al.*, 2002; Ásmundsdóttir *et al.*, 2005; Ásmundsdóttir *et al.*, 2008). In this study, we broadened our scope and investigated both inter- and intra-ward clusters as supposed to only intra-ward clusters. Our analyses identified a likely strain transfer event between wards, even between hospitals. Another difference between this study and that of Ásmundsdóttir *et al.* (2008) is the definition of a “nosocomial cluster” where they used 95% genotype similarity as the cutoff for a strain cluster while we used 100% genotype similarity. If we relaxed our criterion to 95% genotype similarity, over 50% of the strains in our study would be in nosocomial clusters (Figures 1.6-1.9).

To demonstrate that these clusters were not due to the lack of polymorphisms among the markers analyzed or due to the lack of diversity in the analyzed populations, we calculated the discriminatory power of the markers. Discriminatory power of a marker is the ability of the marker to differentiate between any two strains in a given population sample. A low discriminatory power would result in an over-estimate of the number of strain clusters due to the grouping of different strains in the same cluster. In this study, the discriminatory power of the combined three markers was very high, close to 90% for each of the species (Table 1.2). This

result and the retrospectively collected patient data helped ensure that the nosocomial clusters identified here were robust.

We observed a single *C. albicans* cluster containing clinical isolates from patients admitted in the ICU wards of two hospitals located about 10 km from each other. While strains CA5 and CA33 were found in the ICU wards of hospital A, strains CA34 and CA43 were found in the ICU wards of hospital B (Figure 1.1). There are two possibilities for this observation. The first is the lack of genetic diversity between the exogenous strains of the analyzed *C. albicans* population and that the clustering was purely by chance. However, the pairwise analysis of the 125 strains from different patients indicated that there was a high diversity among the strains of all four species of *Candida*, including *C. albicans*. The second hypothesis for the shared genotype between the two hospitals within 45 d was due to inter-hospital transmission. A possible route of such a transmission between hospitals could be through healthcare workers (doctors and nurses). Indeed, healthcare workers have been reported as likely agents of transmission for pathogens in several previous studies (Strausbaugh *et al.*, 1994; Pfaller, 1995; Reagan *et al.*, 1995; Pfaller *et al.*, 2001; Chowdhary *et al.*, 2003).

The observed genotype sharing across all four *Candida* species suggests that hospital environments (including healthcare workers) contributed to candidemia in Hamilton, Ontario. This hypothesis is further supported by our correlation analyses between genetic distance and temporal time difference, allowing us to determine the overall statistical significance of such clustering in the whole samples. The positive correlation observed between these two variables suggested that nosocomial infections contributed significantly to the observed candidemia in Hamilton hospitals (Figures 1.6-1.9). We found that for each of the four species, strains isolated closer to each other temporally were overall genetically more similar to each other as well

(Figures 1.6-1.9). This was further supported by our Mantel analysis we were observed a positive R value meaning a positive correlation exists (Table 1.5). The P value of < 0.01 which meant the positive correlation was significant and that the null hypothesis was rejected.

Indeed, our observations suggest that nosocomial sources likely caused repeated infections via strain replacement in one patient, who was affected by *C. parapsilosis*. In this patient, four blood culture isolates, CP92, CP94, CP102 and CP105 (Figure 1.4) of *C. parapsilosis* were obtained, each one-week apart from each other. The blood isolates CP92 and CP94, taken in the first 2 weeks clustered separately from CP102 and CP105, the ones that were isolated in the last 2 weeks (Figure 1.4). We hypothesized this might be due to two possible reasons: one was strain replacement and the second was mutation and microevolution. Microevolution has been observed to occur in *Candida*, especially in the presence of stress (Pfaller *et al.*, 1998; Xu *et al.*, 2000). In our study, this patient was treated with caspofungin and fluconazole between the second and third strain isolations. However, microevolution typically refers to small changes within a clonal framework. In the specific case here, our fingerprinting analysis identified seven band differences between the two sets of strains when amplified with the M13 core primer and two band differences when amplified with the (GACA)₄ primer (Figure 1.5). This magnitude of difference in the banding patterns suggested that strain replacement, not microevolution was likely responsible for the observed genotype change. Interestingly, the patient was a 6-month old baby with a low birth weight. The reduced immunity at this age accompanied by the low birth weight condition might have played a significant role in acquiring the new strain in such a short period of time (Ruiz-Diez *et al.*, 1997; Gottfredsson *et al.*, 2003; Clark *et al.*, 2004; Pronovost *et al.*, 2006; Smith *et al.*, 2007).

Previous studies have shown that the presence of a central line, solid tumor, diabetes, surgery and antibiotic therapy are major risk factors for candidemia. For example, a central line has been noted to be present in about 85% of the patients infected with *C. parapsilosis* (Sanchez *et al.*, 1993; Pronovost *et al.*, 2006). In our study, among the 19 patients with candidemia caused by *C. parapsilosis*, all but one had a central line. *Candida parapsilosis* is known to produce biofilms over catheter surfaces leading to repeated infection of different patients by the same strain (Sanchez *et al.*, 1993). Similarly, diabetic patients are known to have compromised leukocyte function and immunity, leading to multiple abnormalities in host responses to microbial invasions. Urinary tract infections continue to be problematic in these patients because glucose in the urine provides an enriched medium. This could further complicate the issue if patients suffer from poor bladder emptying. Infections could ascend from the bladder to the kidney. Apart from people with diabetes, those with poor wound healing during post surgery periods are often at a high risk of developing candidemia from the endogenous and exogenous microflora (Zaoutis *et al.*, 2005).

In summary, we found that 33% of the patients sampled here from 2005-2009 in Hamilton, Ontario were affected by strains sharing a genotype with at least another strain from unrelated hosts within a three-month period. Of the 18 clonal clusters, 5 were found in one hospital and 13 in another. The average age of intra-cluster patients with candidemia was 66 years (Tables 1.1 and 1.4), about four years older than that of the mean age for all patients in this study. Our analyses suggest that the hospital environment could be a significant source of strains for candidemia and that measures should be taken in hospitals to reduce such infections.

Acknowledgements

This study was supported by a grant from the Natural Science and Engineering Research Council (NSERC) of Canada and by an internal grant from the Institute of Infectious Diseases Research in the Michael G. DeGroot School of Medicine of McMaster University.

References

- Ásmundsdóttir LR, Erlendsdóttir H, Gottfredsson M. (2005) Improving survival of patients with candidaemia: analysis of prognostic factors from a long-term, nationwide study in Iceland. *Scand J Infect Dis* 37: 111–120.
- Ásmundsdóttir LR, Erlendsdóttir H, Gottfredsson M. (2002) Increasing incidence of candidemia: results from a 20-year nationwide study in Iceland. *J Clin Microbiol* 40: 3489–3492.
- Ásmundsdóttir LR, Erlendsdóttir H, Haraldsson, G, *et al.* (2008) Molecular epidemiology of candidemia: evidence of clusters of smoldering nosocomial infections. *Clin Infect Dis* 47: 17–24.
- Banerjee SN, Emori TG, Culver DH, *et al.* (1991) Secular trends in nosocomial primary bloodstream infections in the United States, 1980–1989. National Nosocomial Infections Surveillance System. *Am J Med* 91: 86S–89S.
- Chowdhary A, Becker K, Fegeler W, *et al.* (2003) An outbreak of candidemia due to *Candida tropicalis* in a neonatal intensive care unit. *Mycoses* 46: 287–292.
- Clark TA, Slavinski SA, Morgan J, *et al.* (2004) Epidemiologic and molecular characterization of an outbreak of *Candida parapsilosis* bloodstream infections in a community hospital. *J Clin Microbiol* 42: 4468–4472.
- Finkelstein R, Reinhertz G, Hashman N, Merzbach D. (2003) Outbreak of *Candida tropicalis* fungemia in a neonatal intensive care unit. *Mycoses* 14: 587–590.
- Gottfredsson M, Vredenburg JJ, Xu J, Schell WA, Perfect JR. (2003) Candidemia in women with breast carcinoma treated with high-dose chemotherapy and autologous bone marrow transplantation. *Cancer* 98: 24–30.
- Gudlaugsson O, Gillespie S, Lee K, *et al.* (2003) Attributable mortality of nosocomial candidemia, revisited. *Clin Infect Dis* 37: 1172–1177.

- Hajjeh AR, Sofair NA, Harrison LH, *et al.* (2004) Incidence of bloodstream infections due to *Candida* species and in vitro susceptibilities of isolates collected from 1998 to 2000 in a population-based active surveillance program. *J Clin Microbiol* 42: 1519–1527.
- Laupland KB, Gregson DB, Church DL, Ross T, Elsayed S. (2005) Invasive *Candida* species infections: a 5 year population based assessment. *J Antimicrob Chemother* 56: 532–537.
- Leroy O, Gangneux J-P, Montravers P, *et al.* (2009) Epidemiology management, and risk factors for death of invasive *Candida* infections in critical care: A multicenter, prospective, observational study in France (2005-2006). *Crit Care Med* 37: 1612–1618.
- Lewis RE, & Klepser M. (1999) The changing face of nosocomial candidemia: epidemiology, resistance, and drug therapy. *Am J Health Syst Pharm* 56: 525–533.
- Marco F, Lockhart SR, Pfaller MA, *et al.* (1999) Elucidating the origins of nosocomial infections with *Candida albicans* by DNA fingerprinting with the complex probe Ca3. *J Clin Microbiol* 37: 2817–2828.
- Peakall, R., Smouse, P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6: 288-295.
- Pfaller MA, Diekema DJ, Jones RN, *et al.* (2001) International surveillance of bloodstream infections due to *Candida* species: frequency of occurrence and in vitro susceptibilities to fluconazole, ravuconazole, and voriconazole of isolates collected from 1997 through 1999 in the SENTRY antimicrobial surveillance program. *J Clin Microbiol* 39: 3254–3259.
- Pfaller MA, Lockhart SR, Pujol C, *et al.* (1998) Hospital specificity, region specificity, and fluconazole resistance of *Candida albicans* bloodstream isolates. *J Clin Microbiol*; 36: 1518–1529.

- Pfaller MA. (1995) Epidemiology of candidiasis. *Journal of Hospital Infection* 30: 329–338.
- Pfaller MA. (1996) Nosocomial candidiasis: emerging species, reservoirs, and modes of transmission. *Clin Infect Dis* 22(2): S89–S94.
- Prentice AG. (2009) Oral and gastrointestinal Candidosis: Prophylaxis during immunosuppressive therapy. *Mycoses* 32: 42-46.
- Pronovost P, Needham D, Berenholtz S, *et al.* (2006) An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*; 355: 2725–2732.
- Pujol C, Joly S, Lockhart SR, *et al.* (1997) Parity among the randomly amplified polymorphic DNA method, multilocus enzyme electrophoresis, and Southern blot hybridization with the moderately repetitive DNA probe Ca3 for fingerprinting *Candida albicans*. *J Clin Microbiol* 35: 2348–2358.
- Reagan DR, Pfaller MA, Hollis RJ, Wenzel RP. (1995) Evidence of nosocomial spread of *Candida albicans* causing bloodstream infection in a neonatal intensive care unit. *Diagn Microbiol Infect Dis* 21: 191–194.
- Rentz AM, Halpern MT, Bowden R. (1998) The impact of candidemia on length of hospital stay, outcome, and overall cost of illness. *Clin Infect Dis* 27: 781–788.
- Robles JC, Koreen L, Park S, Perlin DS. (2004) Multilocus sequence typing is a reliable alternative method to DNA fingerprinting for discriminating among strains of *Candida albicans*. *J Clin Microbiol* 42: 2480–2488.
- Ruiz-Diez B, Martinez V, Alvarez M, Rodriguez-Tudela JL, Martinez-Suarez JV. (1997) Molecular tracking of *Candida albicans* in a neonatal intensive care unit: long-term colonizations versus catheter-related infections. *J Clin Microbiol* 35: 3032–3036.

- Sanchez V, Vazquez JA, Barth-Jones D, *et al.* (1993) Nosocomial acquisition of *Candida parapsilosis*: an epidemiologic study. *Am J Med* 94: 577–582.
- Smith PB, Morgan J, Benjamin JD, *et al.* (2007) Excess costs of hospital care associated with neonatal candidemia. *Pediatr Infect Dis J* 26: 197–200.
- Soll DR. (2000) The ins and outs of DNA fingerprinting the infectious fungi. *Clin Microbiol Rev* 13: 332–370.
- St. Germain, Laverdière M, Pelletier R, *et al.* (2008) Epidemiology and antifungal susceptibility of bloodstream *Candida* isolates in Quebec: Report on 453 cases between 2003 and 2005. *Can J Infect Dis Med Microbiol* 19: 55–62.
- Strausbaugh LJ, Sewell DL, Ward TT, *et al.* (1994) High frequency of yeast carriage on hands of hospital personnel. *J Clin Microbiol* 32: 2299–2300.
- Swofford, D.L. (2003). PAUP*: Phylogenetic analysis using parsimony (and other methods), version 4.0b 10. Sinauer Associates, Sunderland, Massachusetts.
- Thanos M, Schonian G, Meyer W, *et al.* (1996) Rapid identification of *Candida* species by DNA fingerprinting with PCR. *J Clin Microbiol* 34: 615–621.
- Van Asbeck EC, Huang YC, Markham AN, Clemons KV, Stevens DA. (2007) *Candida parapsilosis* fungemia in neonates: genotyping results suggest healthcare workers hands as source, and review of published studies. *Mycopathologia*; 164: 287–293.
- Xu J, Ramos AR, Vilgalys R, Mitchell TG. (2000) Clonal and spontaneous origins of fluconazole resistance in *Candida albicans*. *J Clin Microbiol* 38: 1214–1220.
- Yamamura D, Rotstein C, Nicolle L, Ioannou S, and The fungal disease registry of The Canadian Infectious Disease Society. (1999) Candidemia at selected Canadian sites: results from the Fungal Disease Registry, 1992-1994. *CMAJ* 160: 493–499.

Zaoutis TE, Argon J, Chu J, *et al.* (2005) The epidemiology and attributable outcomes of candidemia in adults and children hospitalized in the United States: a propensity analysis.

Clin Infect Dis 41: 1232–1239.

Appendix

Table 1.1: Summary information for patient and yeast species demographics of the strains causing candidemia in Hamilton, Ontario, Canada from 2005 to 2009

<i>Candida</i> species or patient age	Total number of patients with candidemia			Number of isolates genotyped in this study (Number of duplicates)
	Male patients	Female patients	Total number of isolates obtained from these patients (Number of duplicates)	
Mean age of patients	55	68	62	62
<i>C. albicans</i>	39	50	100 (11)	87 (1)
<i>C. glabrata</i>	5	5	14 (4)	11 (1)
<i>C. parapsilosis</i>	11	8	27 (8)	20 (4)
<i>C. tropicalis</i>	8	4	18 (6)	15 (3)
<i>C. krusei</i>	0	2	2 (0)	1 (0)
Total	63	69	161 (29)	134 (9)

Table 1.2: The discriminatory index (DI) of the PCR fingerprint pattern using three primers M13, (GACA)₄ and OPA3 across the four species: *C. albicans*, *C. tropicalis*, *C. glabrata*, and *C. parapsilosis*.

Species	PCR fingerprinting primers			Mean DI ¹ per species
	M13	(GACA) ₄	OPA3	
<i>C. albicans</i>	0.95	0.90	0.91	0.92
<i>C. tropicalis</i>	0.90	0.88	0.92	0.90
<i>C. parapsilosis</i>	0.83	0.80	0.87	0.83
<i>C. glabrata</i>	0.77	0.79	0.93	0.83
Average DI per Primer	0.86	0.84	0.91	0.87

¹DI = Discriminatory index

Table 1.3: Comparison of the mean pairwise genetic distance between strains isolated at different time intervals (those <90 d apart vs. those >90 d apart) within each of the four *Candida* species analyzed here.

Species	All pair-wise strain comparisons		Strains pairs within 90 d of each other		T value
	N	Mean distance \pm SD ¹	N	Mean distance \pm SD	
<i>C. albicans</i>	3740	0.17 \pm 0.0623	1421	0.0999 \pm 0.0317	162.092***
<i>C. glabrata</i>	55	0.44 \pm 0.2368	17	0.0989 \pm 0.0686	13.668***
<i>C. parapsilosis</i>	190	0.28 \pm 0.1399	100	0.1729 \pm 0.0792	27.71***
<i>C. tropicalis</i>	103	0.32 \pm 0.2391	49	0.1105 \pm 0.0574	13.546***

¹, SD, standard deviation; ***, p<0.001.

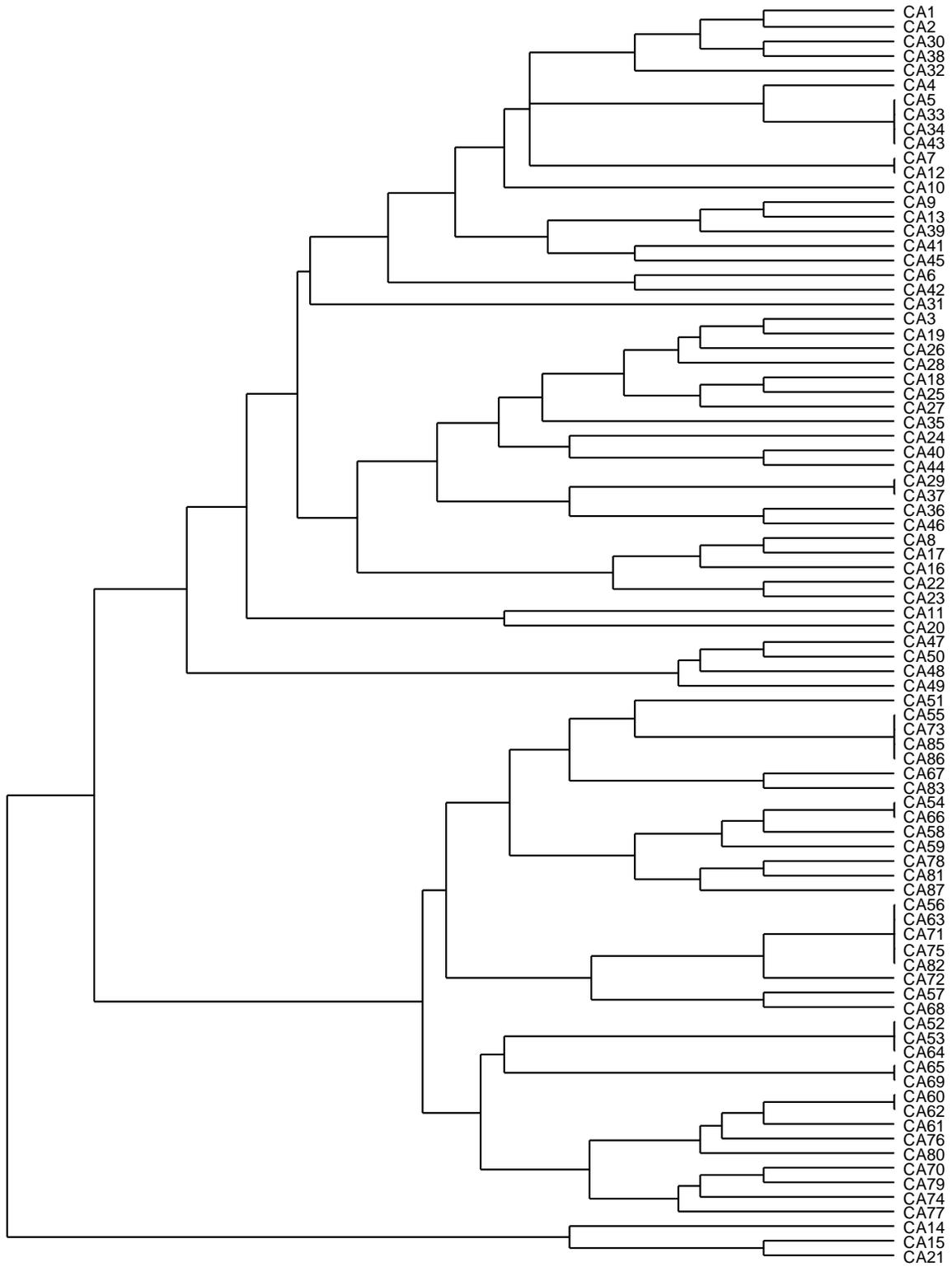
Table 1.4: Nosocomial clusters observed in the clinical isolates causing candidemia in Hamilton, Ontario, Canada from 2005 to 2009.

	No. isolates in clusters	Mean age of patients with strains in clusters (years)	No. Strains in clusters (number of clusters in parenthesis)			
			<i>C. albicans</i>	<i>C. tropicalis</i>	<i>C. glabrata</i>	<i>C. parapsilosis</i>
All patients	44 (33%)	66	26 (9)	4 (2)	6 (3)	8 (4)
Male patients	26	65	18	3	2	3
Female patients	18	67	8	1	4	5

Table 1.5: Mantel test results of the pairwise genetic distance between strains isolated at different time intervals within each of the four *Candida* species.

Species	R Value	P Value
<i>Candida albicans</i>	0.05	0.01
<i>Candida tropicalis</i>	0.05	0.01
<i>Candida glabrata</i>	0.12	0.01
<i>Candida parapsilosis</i>	0.07	0.01

UPGMA



— 0.01 changes

Figure 1.1: Genetic relationships among independent strains of *C. albicans* (CA) in this study

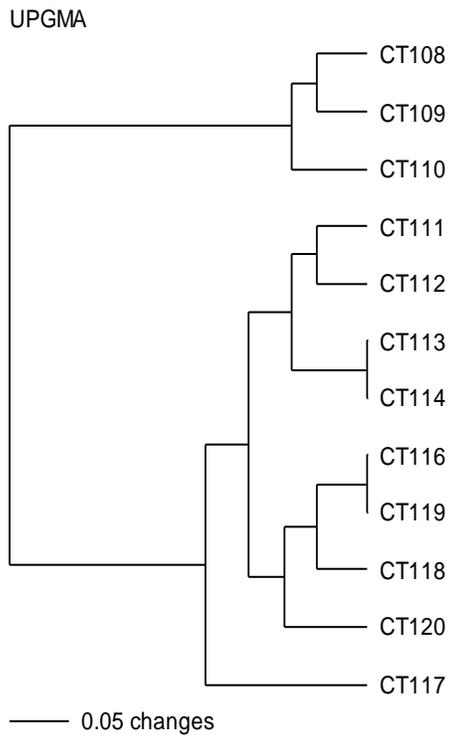


Figure 1.2: Genetic relationships among independent strains of *C. tropicalis* (CT) analyzed in this study

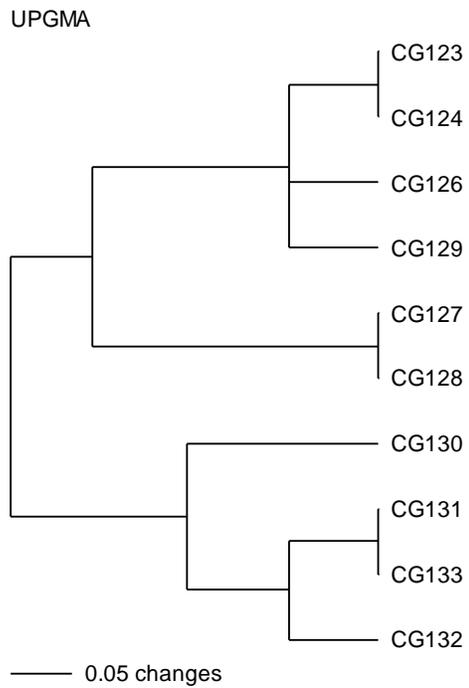


Figure 1.3: Genetic relationships among strains of *C. glabrata* (CG) analyzed in this study

UPGMA

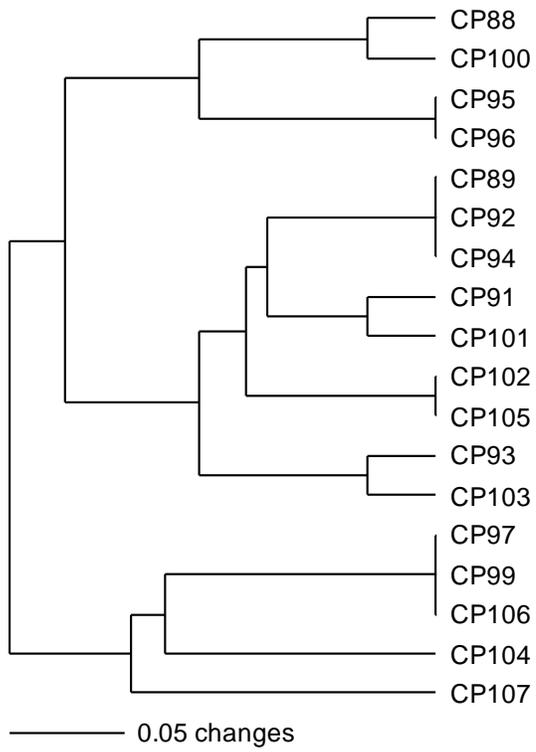


Figure 1.4: Genetic relationships among strains of *C. parapsilosis* (CP) analyzed in this study

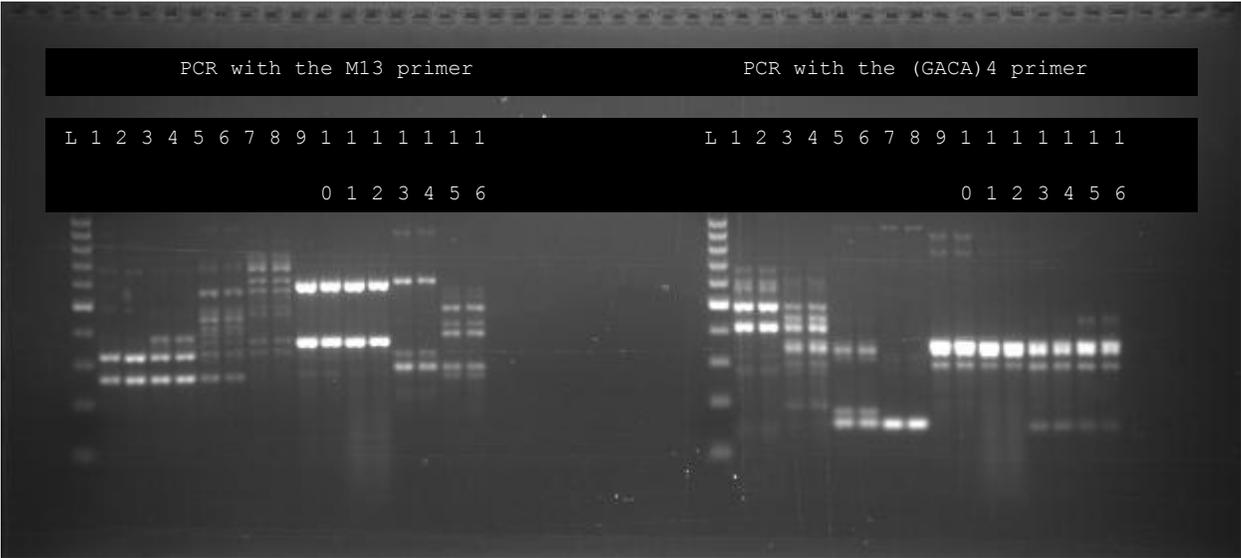


Figure 1.5: Banding patterns of independent strains present in nosocomial clusters using M13 and (GACA)₄ primers. Lanes 1 to 16 correspond to strains CA64, CA84 (a duplicate of CA64), CA85, CA86; CP92, CP94, CP102, CP105, CG123, CG124, CG131, CG133, CT115, CT116, CT121, CT122) respectively. Strains CP92, CP94 and CP102 and CP105 were obtained from the same patient sequentially at one week intervals.

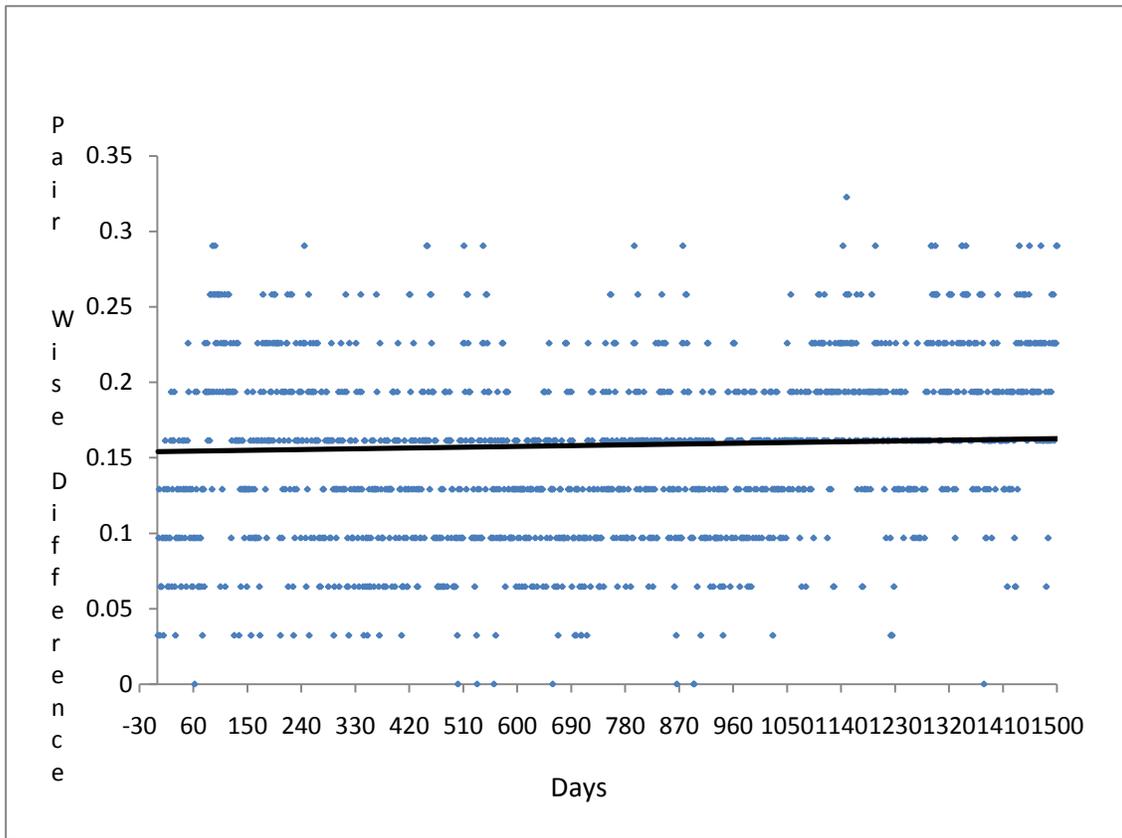


Figure 1.6: The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 86 strains of *Candida albicans* isolated from patients

with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

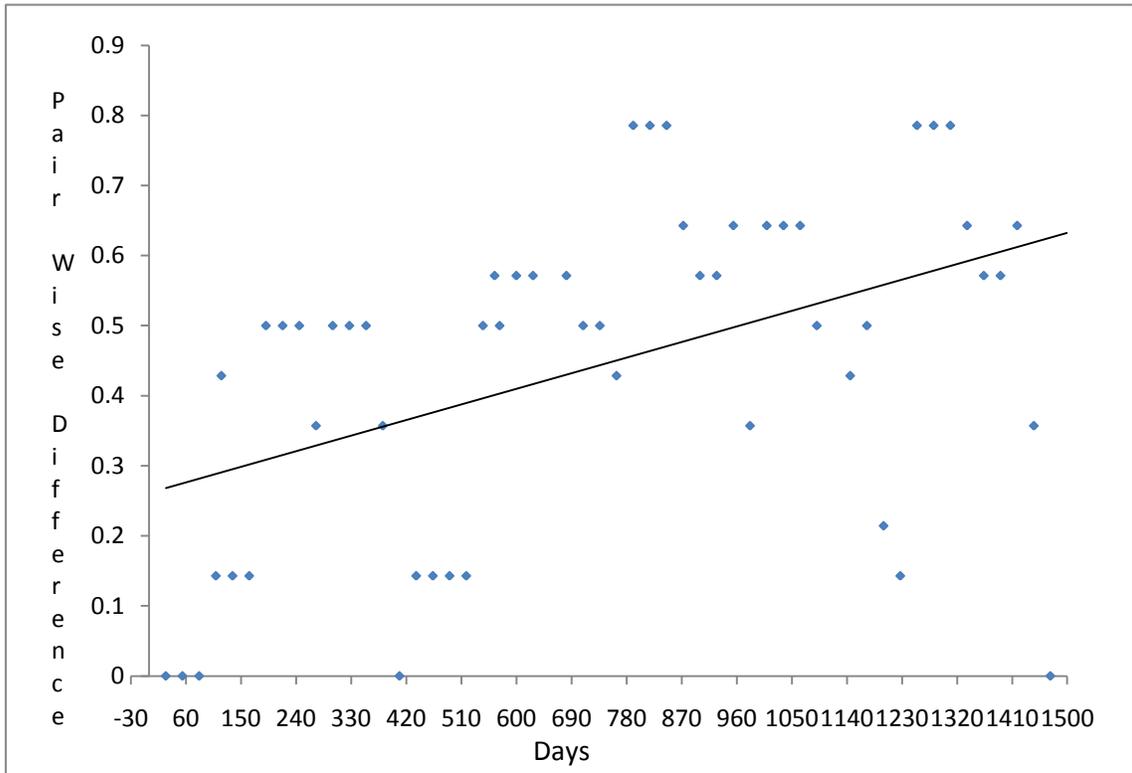


Figure 1.7: The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 10 strains of *Candida glabrata* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

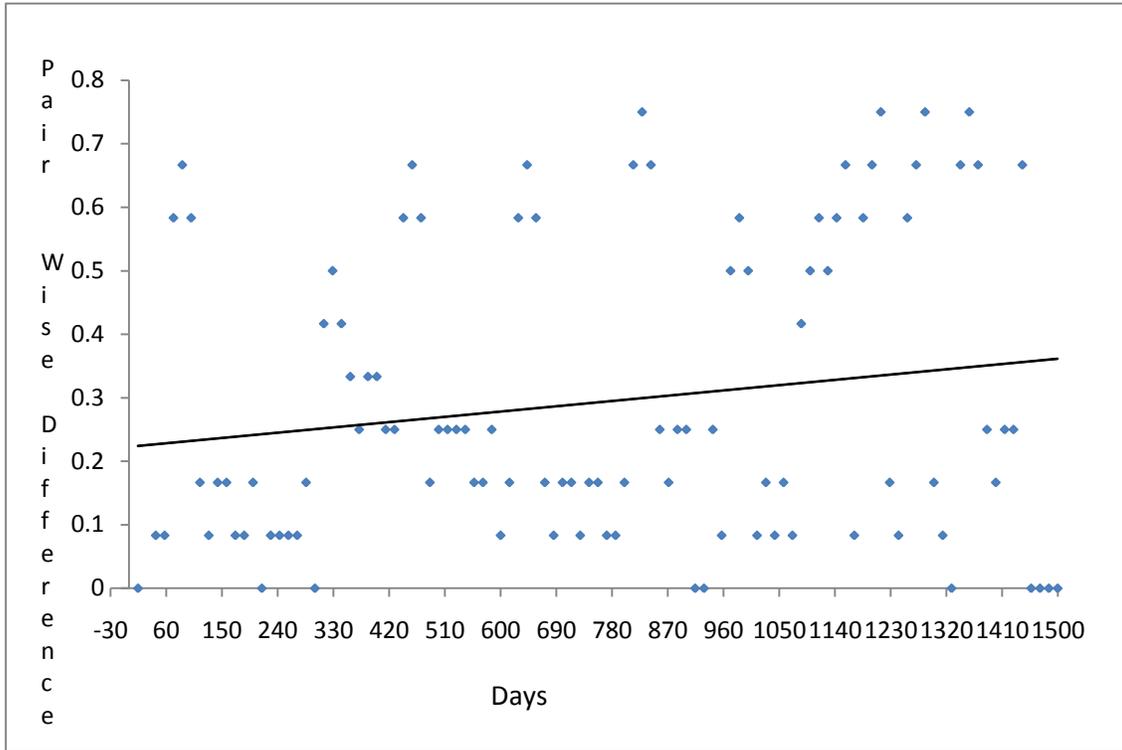


Figure 1.8: The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 11 strains of *Candida tropicalis* isolated from patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

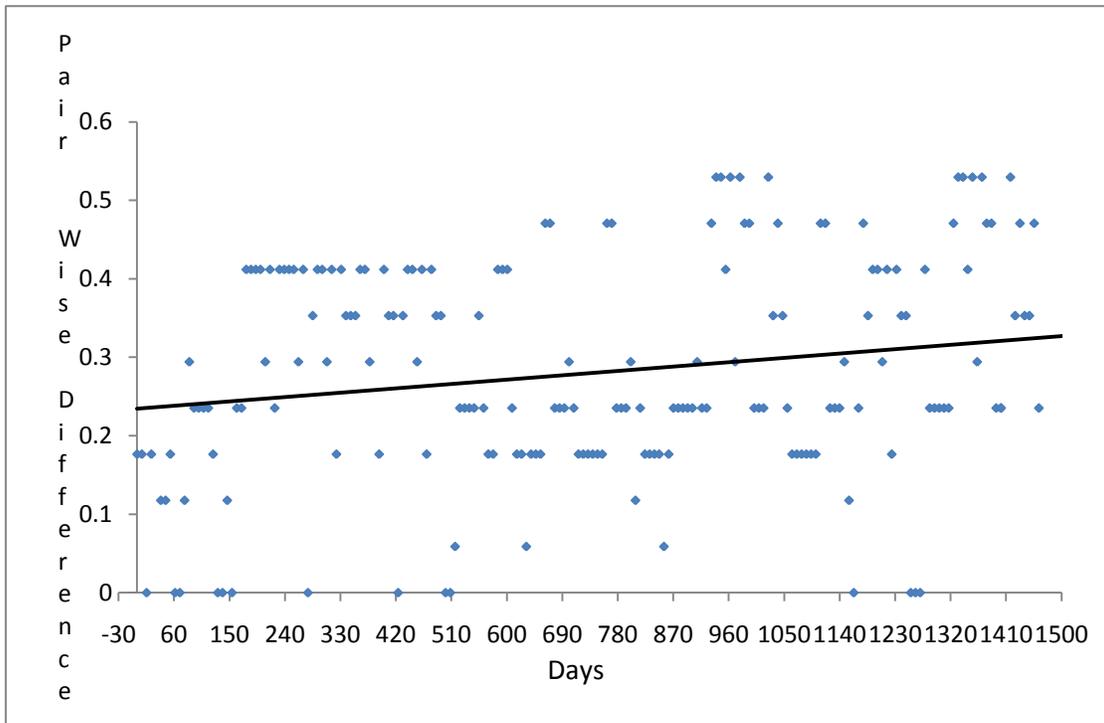


Figure 1.9: The relationship between pairwise strain genetic distance (Y-axis) and temporal time difference (X-axis, in days) for 17 strains of *Candida parapsilosis* isolated from

patients with Candidemia between 2005 and 2009 in Hamilton, Ontario, Canada. A significant positive correlation was found ($p < 0.05$).

Ecological structuring of yeasts associated with trees around Hamilton, Ontario, Canada

Harinad Maganti¹ and David Bartfai^{1,2} , Jianping Xu¹

¹Department of Biology, McMaster University, 1280 Main St W, Hamilton, ON, L8S 4K1, Canada;

Current Address:

²Department of Cell & Systems Biology, University of Toronto, 3359 Mississauga Rd North, ON, L5L 1C6, Canada;

Short Title: Yeast diversity among city trees

Keywords: Yeast diversity, ITS, yeast-tree association, microgeographic structuring

Accession numbers range: JF916480 to JF916567

Corresponding author: Harinad Maganti, Department of Biology, McMaster University, 1280 Main St. West, Hamilton, Ontario, L8S 4K1, Canada. Phone: 1-905-525-9140 ext. 27996. Fax: 1-905-522-6066. Email: maganthb@mcmaster.ca

Abstract

This study seeks to determine the distribution and diversity of yeasts in and around the Hamilton area in Canada. In light of the increasing number of fungal infections along with rising morbidity and mortality rates, especially among the immunocompromised, understanding the diversity and distribution of yeasts in natural environments close to human habitations has become an increasingly relevant topic. In this study, we analyzed 1110 samples obtained from the hollows of trees, shrubs and avian droppings at 8 geographical sites in and around Hamilton, Ontario, Canada. A total of 88 positive yeast strains were isolated and identified belonging to 20 yeast species. Despite the relative proximity of the sampling sites, our DNA fingerprinting results showed that the yeast populations were highly heterogenous in nature. Among the 14 tree species sampled, Cedar, Cottonwood and Basswood hollows had relatively high yeast colonization rates. Interestingly, *Candida parapsilosis* was isolated almost exclusively from Pine trees only. Our results are consistent with microgeographic and ecological differentiation of yeast species in an urban environment.

Introduction

Yeasts are known to inhabit a wide range of ecological niches. For example, yeasts in the Ascomycota and Basidiomycotas phyla have been isolated from a wide variety of surfaces such as those of leaves, rotting wood, soil and marine environments (Middlehoven, 1996; Fraser *et al.*, 2006). In addition, yeasts have been isolated from beetle guts and have been found to play an integral role in the beetle's ability to utilize wood as part of its diet (Tanahashi *et al.*, 2010). This ability of yeasts to degrade and draw nutrients from certain complex organic molecules is speculated to play an important role in the distribution of these microscopic eukaryotes. Yeasts have been shown to possess different abilities to degrade and utilize nutrients such as lipids, hemicelluloses, xylan, and ferulic, hydroxycinnamic, gallic, and tannic acids (Middlehoven, 1997; Bhadra *et al.*, 2008). The variable degrading abilities of yeasts could contribute to differences of yeast species compositions found among yeast populations in soil and rotting wood (Middlehoven, 1997; Inacio *et al.*, 2002; Slavikova *et al.*, 2007).

Population studies of yeasts have shown a surprising amount of variation in yeast species composition among neighbouring regions. In soil populations of yeasts, it was found that variations in factors such as pH, nitrogen concentration and water saturation could impact the quantity and composition of yeast populations (Yarwood *et al.*, 2009). It has been shown that while most yeasts are able to tolerate a wide range of environmental conditions, barriers to their spread over even relatively short distances do exist. For example, the *Cryptococcus gattii* outbreak on Vancouver Island, British Columbia which began in 1999 has yet to establish a stable and comparable reservoir on the mainland or any other neighbouring islands (Fraser *et al.*, 2006; MacDougall *et al.*, 2007; Byrnes and Heitman, 2009). Given the ability of *C. gattii* to spread via airborne or waterborne spores as well as many anthropomorphic routes, this failure to

colonize nearby landmasses may be due to mechanisms other than that of its ability to disperse over natural barriers (Frasier *et al.*, 2006; Kidd *et al.*, 2007).

The failure to migrate and colonize new niches has been observed within established populations of microorganisms, especially those competing for similar nutrients. In these cases, established populations may be better adapted to the particular niche or climate than new migrants and can outcompete and prevent the spread of new populations to those niches (Lachance *et al.*, 2003). For example, population dynamics studies of *Aureobasidium pullulans* have shown that resident populations of this yeast are highly resistant to the immigration of cells, even if they are from the same species (Woody *et al.*, 2007). In most geographic areas, fungal (including yeast) populations often contained not only a high degree of diversity, but also that these communities are rarely homogeneous and frequently contain very different species compositions (Nakase, 2000; Inacio *et al.*, 2002). Even for strains belonging to the same species from a single ecological niche, a measurable genetic variability could still be found, often with their degrees of genetic relationship correlated to geographic distances (Johnson *et al.*, 2004; Koufopanou *et al.*, 2006). However, most current studies have focused on natural environments far from human habitations. In this study we attempted to survey the diversity of yeasts in and around an urban area in Hamilton, Ontario, Canada. We were specifically interested in not only the yeast species distribution patterns but also the potential genetic variability within individual species to determine whether frequent anthropogenic activities could obscure the potential genetic differentiations among the different sites. Yeasts were isolated from 8 sites within and around Hamilton and their species designation was identified based on sequences at the internal transcribed spacer (ITS) region. In addition, DNA fingerprinting was used to determine their

genotypes. Yeast species diversity, distribution and their genetic variation were analyzed and compared among the geographic sites and host trees.

Materials and Methods:

Sample collection:

Yeast samples were collected between May 26, 2008 and September 28, 2008 from 8 sites within and around Hamilton: Site A (43.2143, -79.8220), Site B (43.2565, -79.86715), Site C (43.27793, -79.97966), Site D (43.24747, -79.7489), Site E (43.26191, -79.9193), Site F (43.26286, -79.92069), Site G (43.26910, -79.9031) and Site H (43.22990, -79.9760) (Fig. 1). These eight sites were all within a 30km region and the majority were within 10 km radius from site E. To describe Site A and Site G were sport complexes, Site B was a park area located in Hamilton downtown. Site C was a waterfall area in a conservation area with aging trees, while Site D was a harbor area. Site E (also the center of this study) was McMaster university campus. Site F and Site H were conservation areas.

Samples were located within roughly 500 meters of the center of each of the sampling locations. Within each area, all tree hollows were sampled, with each hollow sampled three times. Selected avian droppings were also swabbed. A total of 370 different surfaces were swabbed (tree hollows and avian droppings together), representing 1110 samples in all. For each sample, a sterile cotton swab was first soaked in sterile 0.5% saline solution immediately before the swab was applied to the wood debris. Each swab was placed in a 2000 μ L tube containing 500 μ L of the same saline solution. Whenever possible and necessary, a section of branch with multiple leaves was also obtained and labelled to allow for the host tree species identification soon after returning to the lab. All samples were taken from live trees only.

Sample processing:

Upon returning to the lab, the samples were enriched with 500µL of YEPD (yeast extract – peptone-dextrose) liquid agar supplemented with 50mg/mL of the antibiotic chloramphenicol. Enriched samples were incubated at 30°C for 48 hours. Samples were then vortexed and plated onto solid YEPD medium and incubated at 30°C for 48 hours. Any colonies displaying yeast-like morphology were isolated and re-streaked onto YEPD agar and incubated at 30°C for 48 hours. Single colonies were then suspended in 500 µL of 20% Glycerol and frozen at -80°C. Frozen samples were revived by plating 10 µL on YEPD agar and incubated for 48 hours at 30°C for DNA extraction and genotyping.

DNA PCR Amplification:

DNA was extracted as shown in Xu *et al.* (2000) from 88 purified yeast-like isolates. The ITS gene fragment of each isolate was amplified using primers ITS1 (5' TCCGTAGGTGAACCTGCGG 3') and ITS4 (5' TCCTCCGCTTATTGATATGC 3') and the following PCR cycle protocol; DNA was denatured at 95^o C for 4 minutes followed by 40 cycles at 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 1 minute and a final cycle at 72°C for 6.5 minutes.

DNA sequencing

Amplified ITS DNA was then cleaned for sequencing by adding an equal volume of MicroClean© to 10-15µL of PCR product, vortexed and incubated at 20°C for 5 minutes. Samples were then centrifuged at 13000 RPM for 5 minutes and visible supernatant was added to 15µL of double distilled H₂O and incubated at 20°C for 5 minutes. Samples were then centrifuged a second time at 13000 RPM for 5 minutes and the supernatant transferred to a new

container. Presence of purified DNA product was tested via electrophoresis on 1.0% agarose gel with 1 X TBE buffer (Tris-Borate-EDTA [pH 8]) stained with ~1 µL of Ethidium Bromide for 1h at 100V. The gels were then photographed under ultraviolet light (Chemi-imager; Alpha InnovTech). Sequencing was carried out by the MOBIX laboratory of McMaster University Life Sciences building and the Life Sciences Core Laboratory Centre of Cornell University.

Species identification

The ITS sequences were compared to archived ITS sequences on the NCBI database using the BLAST algorithm to determine the yeast species. Species were identified based upon the best fit with a nucleotide identity of at least 97%. Species that were found to be present in more than one geographical site were subjected to PCR fingerprinting, to observe the similarity in the genetic constitution of the strains.

ITS Phylogeny construction

All the ITS sequences of the yeast strains were aligned by ClustalW, along with reference sequences for the individual species previously deposited in GenBank. The aligned sequences were used to construct a Maximum-Parsimony tree (Tamura, Peterson, Stecher *et al*, 2011). The bootstrap support for individual branches were generated with 1000 replicate randomized dataset.

PCR Fingerprinting

All yeast isolates were genotyped by PCR fingerprinting using 3 arbitrary primers (GACA)₄, OPA3, and M13, each separately (Rentz *et al.*, 1998; Gudlaugsson *et al.*, 2003; Hajjeh *et al.*, 2003; Maganti *et al.*, 2011). Each standard PCR tube contained 8 µl of Ready-to-

Go-PCR mix (Amersham Biosciences), 4 μ l of working concentration of DNA, and 4 μ l of 0.10 μ M of desired primer. The amplifications were subjected to gel electrophoresis on 1.0% agarose gels with 1 X TBE buffer for 3h at 100V. The agarose gels were stained with Ethidium Bromide prior to loading the samples. On completion of gel electrophoresis the gels were photographed digitally under ultraviolet light (Chemi-imager; Alpha Innovtech). Representative strains from each of the analyzed species were tested multiple times for PCR and gel electrophoresis to check for reproducibility of the PCR-fingerprinting patterns.

PCR Fingerprinting data analysis

The DNA bands on the PCR fingerprinting gels were scored manually. A position tolerance setting of 2% was used (Soll, 2000; Ásmundsdóttir *et al.*, 2008). Strains with identical genotypes were expected to show identical fingerprint patterns for all the 3 primers used. To ensure consistency and minimize errors, all the finger printing results were scored twice by two individuals (HM and DB) and this was done for each of the 3 primers.

Phylogram construction and discriminatory index

The scored fingerprinting results of the amplified DNA were used to construct phylograms to determine the relationships among the strains using the phylogenetic program PAUP (Swofford, 2003). This was done independently for each of the three primers as well as for the combined dataset. The Discriminatory Index for each of the three primers was calculated using the Simpsons Diversity Index as done elsewhere (Xu *et al.*, 2000; Gudlaugsson *et al.*, 2003).

Species richness comparison

The species richness for each site was determined as a ratio of the number of unique yeast species found at each site.

Host tree preference analysis

The percentage of positive yeast samples obtained from a tree species was calculated by dividing the total number of positive yeast isolates obtained from that tree by the total number of swabs taken from it. This was termed as ‘Tree Positive Percentage’ (TPP) and was done for each of the individual tree hosts. The mean and standard deviation for all these values was calculated. We refer to the overall mean as the ‘Mean Positive Percentage’ (MPP) and the standard deviation as SD.

The significance of any observed yeast - tree preference was evaluated through Z score analysis. The null hypothesis of the test is that the percentage of positive yeast samples obtained from all the trees is similar to each other and close to the mean. A Z value of > 1.64 ($Q > 0.95$) indicates that the specific tree has a significantly higher rate than the mean. The Z score was calculated by subtracting MPP from TPP and then dividing this number with SD.

Difference between deciduous and coniferous trees

The statistical significance of yeast preference to grow on deciduous trees over coniferous trees or vice versa was examined using the Fisher’s exact test. The null hypothesis of the test was that the yeasts show no preference between deciduous trees or coniferous trees.

Results:

Species diversity of yeast populations at different geographical sites.

A total of 370 tree hollows were sampled, of which 88 were positive for yeast. Of the 88 positive yeast samples, thirteen were from the Site A, five were from Site B, thirteen were from Site C, 21 were from Site D, nineteen were from Site E, nine were from Site F, seven were from Site G and one were from Site H (Table 2.1). Based on the ITS sequences, the 88 yeast samples belonged to 20 yeast species (Figure 2.2).

Six of the 20 yeast species (*Candida parapsilosis*, *Kazachstania aerobia*, *Kluyveromyces lactis*, *Torulaspora delbrueckii*, *Debaryomyces hansenii* and *Saccharomyces paradoxus*) containing 62 strains were identified to be present at more than one site. 14 of the species were each found at only one of the sampled sites (Figure 2.2; Table 2.2).

Fingerprinting results:

Our combined PCR fingerprinting results identified that the nine strains of *Candida parapsilosis* belonged to eight distinct genotypes. Seven of these genotypes were present at site D, while one was present at site G. None of the eight genotypes overlapped between the two sites (Figure 2.3).

The nine strains of *Kazachstania aerobia* belonged to eight unique genotypes. Of the eight genotypes, two were found at site D while the remaining six were found at site E. None of the eight genotypes overlapped between the two different sites (Figure 2.4).

The ten strains of *Kluyveromyces lactis* belonged to nine genotypes. Two of the nine genotypes were present at Site D, while four and three genotypes were found at site E and site F respectively. None of the nine genotypes overlapped across the three different sites (Figure 2.5).

The ten strains of *Torulaspota delbrueckii* each had a different genotype. Of the ten genotypes one was found at site D, five at site G, three at site E, and one at site C. None of the ten genotypes overlapped across the four different sites (Figure 2.6).

The seventeen strains of *Saccharomyces paradoxus* were identified to belong to seventeen genotypes. Of the seventeen genotypes, one was found at site D, two at site F, thirteen at site A and one at site E. None of the seventeen genotypes overlapped across the four different sites (Figure 2.7).

The seven strains of *Debaryomyces hansenii* belonged to seven genotypes. Of the seven genotypes, one was found at site E and site C each, while five genotypes were found at site D. None of the seven genotypes overlapped across the three different sites (Figure 2.8).

Discriminatory power of the markers

The discriminatory powers of the three PCR fingerprinting primers were similar. For *Candida parapsilosis*, the M13 PCR fingerprinting primer had the highest power of 0.87 followed by OPA3 and (GACA)₄ with values of 0.86 and 0.85. The combined discriminatory power for the markers was identified to be 0.86. Among *Kazachstania aerobia* strains, the OPA3 primer had the highest discriminatory power of 0.87, followed by (GACA)₄ and M13 with powers of 0.86 and 0.85. The combined discriminatory power for the three markers was 0.86. As for *Kluyveromyces lactis* strains, M13 had the highest discriminatory power of 0.90 followed by OPA3 and (GACA)₄ with values of 0.86 and 0.85. Their combined discriminatory power for the three markers was 0.87.

Among the *Torulaspota delbrueckii* strains, the marker OPA3 had the highest discriminatory power of 0.91 followed by (GACA)₄ and M13 with powers of 0.90 and 0.89. The

combined discriminatory power for the three markers was identified to be 0.90. Among *Saccharomyces paradoxus* strains the marker M13 had the highest discriminatory power of 0.96 followed by (GACA)₄ and OPA3 with powers of 0.95 and 0.91. The combined discriminatory power for the three markers was identified to be 0.94. Amid the *Debaryomyces hansenii* strains the marker (GACA)₄ had the highest discriminatory power of 0.87 followed by M13 and OPA3 with powers of 0.86 and 0.85. The combined discriminatory power for the three markers was identified to be 0.86.

Yeast species richness based on geographic sites

Site C was identified to be the most species rich site, containing strains belonging to nine species. This was followed by site D (eight species), site E (seven species), site F (four species), site G (two species) and lastly sites A, B and H with each containing strains belonging to only one species each (Table 2.2).

Host tree – yeast species correlation:

Host tree species also appeared to have an impact on the number and type of yeast isolates collected from each site. A significantly high percentage of samples swabbed from White Cedar, Basswood and Cotton wood turned out to be yeast (Table 2.3). Over all the probability of finding yeast species on Conifers was significantly higher than Deciduous trees (Table 2.4). Interestingly, all the *C. parapsilosis* isolated strains were found to be present on Pine trees only.

Discussion

In this study we isolated and identified 20 yeast species from 370 tree hollows across the 8 sampling sites. Six of these 20 yeast species were distributed at more than one site (Figure 2.2;

Table 2.2). The genotype analysis based on PCR fingerprinting showed that no genotype was shared between any of the sites (Figures 2.4-2.8). Our results are consistent with microgeographic and ecological structuring among the yeasts in and around an urban environment. The high amount of heterogeneity found prevalent among the yeast populations of Hamilton was unexpected due to the relative proximity of the sampled areas (Figure 2.1), and the availability of vectors (e.g. birds, rodents and humans) to shuttle these yeasts between the areas. In addition, sites C, G, F and D were all connected by waterways. While the detailed mechanisms are unknown, differences in population history could have contributed to the observed spatial heterogeneity. For example, established yeast populations have been previously identified to monopolize the nutrients of a particular site thus blocking the founding of new yeast populations (Vishniac *et al.*, 2006).

In terms of species richness, we observed the species richness to vary between the different sites. There was no observed correlation between yeast species richness and tree species richness. With the exception of site C which was observed to be the most species rich site for both yeasts and trees, the order of yeast species richness and tree species richness varied among the different sites (Table 2.2). Similar findings have been reported by other studies conducted elsewhere (Middlehoven, 1997, Inacio *et al.*, 2002, Slavikova *et al.*, 2007). However, the amount of decomposing matter present on the grounds at the sampled sites might be related to the rate of yeast isolation at the given site. Though we didn't quantify the decomposing organic matter, sites D and E contained visibly the largest amount of decomposing material and these two sites also contained the highest percentage of yeast isolation rate. In contrast, site H contained very little detritus and it had the lowest yeast isolation rate. Since previous studies have shown that increased detritus leads to increased carbon availability in the soil which in the past has shown to

lead to increased yeast and bacterial populations in soil (Gonzalez *et al.*, 1989; Wardle *et al.*, 1992; Wardle *et al.*, 2004; Höglberg & Read, 2006; Yarwood *et al.*, 2009), we speculate that these yeast species might have been carried onto the trees, from the soil by animals and anthropogenic factors, where they might have found it convenient to colonize the tree hollows.

The degradative abilities of yeasts likely also play a role in their distribution on trees (Middlehoven, 1997; Inacio *et al.*, 2002; Slavikova *et al.*, 2007). In this study, we observed that coniferous trees seemed more likely to harbour yeasts than deciduous ($P = 0.02$). A significantly high percentage of samples (compared to the mean) isolated from the detritus of cottonwood, Cedar and basswood had yeast (Table 2.3). Of these trees, Cedar is a softwood coniferous tree, while basswood and cottonwood are deciduous hardwood trees. This was unexpected because previous studies have shown no such bias between deciduous and coniferous trees (Sláviková *et al.*, 2007) (Table 2.4). Apart from the degradative abilities of yeast, another property that might have played a role in the distribution of yeast is, the pH of the tree barks which is known to vary between different trees.

Apart from the nature of trees (i.e. deciduous or coniferous), the innate resistance of fungi to anti-fungal chemicals present in the trees could also play an important role in deciding what species inhabits which tree. In this study all the *C. parapsilosis* strains were observed to colonize Pine trees only. *C. parapsilosis* has been previously observed to grow on the barks of injured pine trees (El-Tabey *et al.*, 1952). However, this has not been shown to be an exclusive association. These results can however be explained by the fact that Pine wood and pine needles contain a constituent named beta-pinene, which is known to possess anti-fungal properties that few yeasts in *Candida* were resistant (Uribe *et al.*, 1984; Krauze-Baranowska *et al.*, 2002;

Pozzatti *et al.*, 2010). This innate resistance of *C. parapsilosis* to this compound could potentially provide an advantage for *C. parapsilosis* over other yeasts to colonize the pine trees.

Environmental pollutants such as Poly Aromatic Hydrocarbons (PAHs) may also shape microbial flora in natural environments. In this study, site D is located downstream of two steel mills and a highway and high levels of PAHs have been detected in this area. The two species *C. parapsilosis* and *D. hansenni* known to be able to metabolize PAHs (MacGillivray *et al.*, 1993; Sofowote *et al.*, 2008) were found on the trees close to the steel mills and the highway. A similar trend was also observed in site C, site F and site G, all of which are known to be connected to site D by water ways. In these sites we observed the dominance of species belonging to *Candida*, *Pichia*, *Rhodotorula* on trees close to the water sources and likely contained high levels of PAHs. The common ability of these species is to be able to metabolize PAH's (MacGillivray *et al.*, 1993). A detailed test of the levels of PAH at these environments would be needed to demonstrate whether PAH contamination selects for certain yeast species in the natural environment. It is important to note that in this particular case many of the species in genera *Candida*, *Pichia* and *Rhodotorula* are capable of causing human and animal infections and if our speculation of the correlation between PAHs and human pathogenic yeast is true, then the rise in environmental PAHs could potentially lead to fungal outbreaks in the future. The fact that site F, which is about 900 meters away from McMaster children hospital (Figure 2.1) has been identified to possess strains belonging to genus *Pichia* and *Candida* both of which are known to contain species that cause nosocomial infections in clinical settings (Figure 2.1; Table 2.2) makes us believe further investigation should be made in this area and the necessary precautionary measures should be taken (Chakrabarti *et al.*, 2001; Maganti *et al.*, 2010).

In summary, we isolated 88 yeast strains and identified that they belonged to twenty yeast species. Six of the 20 species were identified to be present at more than one site. The fingerprinting results showed that despite the close proximity of the sampling sites and the presence of dispersal vectors, the populations of yeasts belonging to the same species were heterogeneous in nature. Our results suggested little evidence for clonal dispersal among the strains in the urban environments. While the probability of finding yeasts on Cedar, Basswood and Cottonwood was noticeably high, *C. parapsilosis* was observed only on Pine trees in our samples. Our results suggest that the substrate-utilization abilities of yeasts likely played a significant role in their distribution in the environment in and around Hamilton.

Acknowledgements

This study was supported by a grant from the Natural Science and Engineering Research Council (NSERC) of Canada.

References

- Ásmundsdóttir LR, Erlendsdóttir H, Haraldsson, G, Guo H, Xu J, Gottfredsson M (2008) Molecular epidemiology of candidemia: evidence of clusters of smoldering nosocomial infections. *Clin Infect Dis* 47:17-24.
- Bhadra B, Rao RS, Singh PK, Sarkar PK, Shivaji S (2008) Yeasts and yeast-like fungi associated with tree bark: diversity and identification of yeasts producing extracellular endoxylanases. *Curr Microbiol* 56(5): 489-494.
- Byrnes III JE and Heitman J (2009) *Cryptococcus gattii* outbreak expands into the Northwestern United States with fatal consequences. *F1000 Biology Reports* 1:62.
- Chakrabarti A, Sing K, Narang A, Singhi S, Batra R, Rao NLK, Ray P, Gopalan S, Das S, Gupta V, Gupta KA, Bose MS and McNeil MM (2001) Outbreak of *Pichia anomala* infection in the pediatric service of a tertiary-care center in northern India. *J Clin Microbiol* 39(5): 1702-1706.
- Sláviková E, Vadkertiova; R, and Vra;nova; D (2007) Yeasts colonizing the leaf surfaces. *J. Basic Microbio* 47: 344-350.
- El-Tabey Awad Shihata MA and Mrak ME (1952) Intestinal yeast floras of successive populations of *Drosophila*. *Evolution* 6(3): 325-332.
- Fraser AJ, Lim CMS, Diezmann S, Wenink CE, Arndt GC, Cox MG, Dietrich SF and Heitman J (2006) Yeast diversity sampling on the San Juan Islands reveals no evidence for the spread of the Vancouver Island *Cryptococcus gattii* outbreak to this locale. *FEMS Yeast Res* 6: 620-624.

- Gonzalez A, Martinez A, Almendros G & Grinbergs J (1989) A study of yeasts during the delignification and fungal transformation of wood into cattle feed in Chilean rain forest. *A Van Leeuw J Microb* 55:221-236.
- Gudlaugsson O, Gillespie S, Lee K, Vande Berq J, Hu J, Messer S, Herwaldt L, Pfaller M, Diekema D (2003) Attributable mortality of nosocomial candidemia, revisited. *Clin Infect Dis* 37:1172–1177.
- Högberg P and Read D (2006) Towards a more plant physiological perspective on soil ecology. *Trends Ecol Evol* 21: 548–554.
- Hajjeh, AR, Sofair NA, Harrison LH, Lyon MG, Arthington-Skaggs AB, Mirza AS, Phelan M, Morgan J, Lee-Yang W, Ciblak AM, Benjamin EL, Sanza TL, Huie S, Yeo FS, Brandt EM and Warnock WD (2004) Incidence of bloodstream infections due to *Candida* species and in vitro susceptibilities of isolates collected from 1998 to 2000 in a population-based active surveillance program. *J Clin Microbiol* 42: 1519-1527.
- Inacio J, Pereira P, De Carvalho M, Fonseca A, Amaral-Collac TM, and Spencer-Martins I (2002) Estimation and Diversity of Phylloplane Mycobiota on Selected Plants in a Mediterranean “Type Ecosystem in Portugal”. *Microb Ecol* 44: 344-353.
- Johnson LJ, Koufopanou, V, Goddard M R, Hetherington R, Schafer SM, and Burt A (2004) Population genetics of the wild yeast *Saccharomyces paradoxus*. *Genetics* 166: 43–52.
- Kidd SE, Bach JP, Hingston OA, Mak S, Chow Y, MacDougall L, Kronstad WJ, and Bartlett HK (2007) *Cryptococcus gattii* Dispersal Mechanisms, British Columbia, Canada. *Emerg Infect Dis* 13: 51-57.

- Koufopanou V, Hughes J, Bell G, and Burt A (2006) The spatial scale of genetic differentiation in a model organism: the wild yeast *Saccaromyces paradoxus*. *Philos Trans R Soc Lond B Biol Sci* 361: 1941-1946.
- Krauze-Baranowska M, Mardarowicz M, Wiwart M, Pობlocka L, and Dynowska M (2002) Antifungal activity of the essential oils from some species of the genus *Pinus*. *Z Naturforsch C* 57(5-6): 478-482.
- Lachance MA, Bowles MJ, and Starmer TW (2003) Geography and niche occupancy as determinants of yeast biodiversity: the yeast insect morning glory ecosystem of Kipuka Puauulu, Hawaii. *FEMS Yeast Res* 4: 105-111.
- MacDougall L, Fyfe M, Romney M, Starr M, and Galanis E. (2011) Risk factors for *Cryptococcus gatii* infection, British Columbia, Canada. *Emerg Infect Dis* 17(2): 193-199;
- MacGillivray RA, Shiaris PM. (1993) Biotransformation of polycyclic aromatic hydrocarbons by yeasts isolated from costal sediments. *Applied AEM* 59(5): 1613-1618.
- Maganti H, Yamamura D, Xu J (2011) Prevalent nosocomial clusters among causative agents for candidemia in Hamilton, Canada. *Med Mycol*: DOI:10.3109/13693786.2010.547880.
- Middlehoven JW (1997) Identity and biodegradative abilities of yeasts isolated from plants growing in an arid climate. *A Van Leeuw J Microb* 72(2): 81-89.
- Nakase T (2000) Expanding world of ballistosporous yeasts: Distribution in the phyllosphere, systematic and phylogeny. *J Gen Appl Microbiol* 46: 189-216.
- Pozzatti P, Loreto SE, Nunes Mario AD, Rossato L, Santurio MJ, Alves HS (2010) Activities of essential oils in the inhibition of *Candida albicans* and *Candida dubliniensis* germ tube formation. *Med Mycol* 20: 185-189.

- Rentz AM, Halpern MT, Bowden R (1998) The impact of candidemia on length of hospital stay, outcome, and overall cost of illness. *Clin Infect Dis* 27:781–788.
- Sofowote UM, McCarry BE, Marvin CH (2008) Source apportionment of PAH in Hamilton Harbour suspended sediments: Comparison of two factor analysis methods. *Environ Sci Technol* 42: 6007–6014.
- Soll DR (2000) The ins and outs of DNA fingerprinting the infectious fungi. *Clin Microbiol Rev* 13: 332–370.
- Swofford DL (2003) PAUP*: Phylogenetic analysis using parsimony (and other methods), version 4.0b 10. Sinauer Associates, Sunderland, Massachusetts.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, and Kumar S (2011) MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Molecular Biology and Evolution* (submitted).
- Tanahashi M, Kubota K, Matsushita N and Togashi K (2010) Discovery of mycangia and the associated xylose-fermenting yeasts in stag beetles (Coleoptera: Lucanidae). *Naturwissenschaften* 97(3):311-317.
- Uribe S, Ramirez J and Pena A (1985) Effects of beta-pinene on yeast membrane functions. *J Bacteriol* 161(3):1195-1200.
- Vishniac and Helen S (2006) A Multivariate Analysis of Soil Yeasts Isolated from a Latitudinal Gradient. *Microb Ecol* 50: 90-103.
- Wardle DA, Bardgett RD, Klironomos JN, Setälä H, Van der Putten WH and Wall DH (2004) Ecological linkages between aboveground and belowground biota. *Science* 304:1629–1633.
- Wardle DA (1992) A comparative assessment of factors which influence microbial biomass carbon and nitrogen levels in soil. *Biol Rev* 67: 321–358.

Woody ST, Ives RA, Nordheim VE, and Andrews HJ (2007) Dispersal density dependence and population dynamics of a fungal microbe on leaf surface. *Ecology* 88: 1513-1524.

Xu J and Mitchell TG (2003) Geographic differences in human oral yeast flora. *Clin Infect Dis* 36: 221-224.

Yarwood AS, Myrold DD and Hogberg NM (2009) Termination of belowground C allocation by trees alters soil fungal and bacterial communities in a boreal forest. *FEMS Microbiol Ecol* 70: 151–162.

Appendix

Table 2.1: Summary information relating to the total number samples collected at each sampling site and percentage of them that contained yeast.

Location	Total Samples (370 tree hollows x 3)	Samples with yeast	Percentage of Positive samples
Site A	120	13	10.83
Site B	120	5	4.17
Site C	120	13	10.83
Site D	90	21	23.33
Site E	120	19	15.83
Site F	360	9	2.50
Site G	105	7	6.67
Site H	75	1	1.33
Total	1110	88	7.93
Average	138.75	11	
Standard Deviation	90.98	6.85	

Table 2.2: Summary information of yeast species and tree species identified at each sampling site and the number of yeast isolates for each species at each site.

Site Name	Yeast Species Identified [no. of isolates]	Tree Species Identified (need latin name for these tree species)
Site A	<i>S. paradoxus</i> [13]	Cedar, Bass wood, Oak, Maple, Cotton wood
Site B	<i>R.mucilaginosa</i> [5],	Cherry
Site C	<i>D.hansenii</i> [1], <i>P. guilliermondii</i> [4], <i>S.martinae</i> [1], <i>T.delbrueckii</i> [1] <i>C.psychrophila</i> [1], <i>H.opuntiae</i> [1], <i>F.inconspicuum</i> [1], <i>R.toruloides</i> [1], <i>S.citricola</i> [2]	Hawthorn, Wall nut, Beech, Ash, Bass wood, Birch, Sycamore, Cotton wood
Site D	<i>K.lactis</i> [2], <i>T.delbrueckii</i> [1], <i>C.parapsilosis</i> [7], <i>K.aerobia</i> [3], <i>D.pseudopolymo</i> [1], <i>D.hansenii</i> [5], <i>P.parasitica</i> [1], <i>S. paradoxus</i> [1]	Pine, Cotton wood, Bass wood, Ash, Maple, Raspberry
Site E	<i>D.hansenii</i> [1], <i>K.lactis</i> [5], <i>T.delbrueckii</i> [3], <i>K. aerobia</i> [6], <i>R.mucilaginosa</i> [1], <i>L.kluyveri</i> [2], <i>S. paradoxus</i> [1]	Ash, Sycamore, Cotton wood, Chestnut, Beech
Site F	<i>P.hampshirensis</i> [3], <i>C.tropicalis</i> [1], <i>K.lactis</i> [3], <i>S.paradoxus</i> [2]	Cotton wood, Cedar, Bass wood, Sycamore, Maple
Site G	<i>C.parapsilosis</i> [2], <i>T.delbrueckii</i> [5],	Pine, Beech, Cotton wood, Chestnut
Site H	<i>C. lignohabitans</i> [1]	Cottonwood

Table 2.3: Summary of positive yeast samples and the total number of samples taken from each tree as well as the Z -score correlation between them.

Sample Source	Latin names	No.of yeast samples	No.of samples taken	Total Samples	TPP	Z Score	Q > 0.95
Ash	<i>Zanthoxylum americanum</i>	13	54	162	8.02	-0.19	-
Basswood	<i>Tilia americana</i>	7	11	33	19.44	1.67	+++
Birch	<i>Betula populifolia</i>	1	18	54	1.85	-1.06	-
Cedar	<i>Thuja occidentalis</i>	11	17	51	21.57	1.72	+++
Cherry	<i>Prunus avium</i>	5	15	45	11.11	0.25	-
Chestnut	<i>Aesculus hippocastanum</i>	2	18	54	3.7	-0.80	-
Cotton Wood	<i>Populus deltoides</i>	13	20	60	20.63	1.74	+++
Hawthorn	<i>Crataegus punctata</i>	2	20	60	3.33	-0.85	-
Maple	<i>Acer nigrum</i>	4	36	108	3.7	-0.80	-
Oak	<i>Quercus velutina</i>	2	20	60	3.33	-0.85	-
Pine	<i>Pinus ponderosa</i>	9	27	81	11.11	0.25	-
Sycamore	<i>Platanus occidentalis</i>	6	12	36	13.89	1.03	-
Detritus		2	33	99	2.02	-1.04	-
Avian droppings		3	43	129	2.32	-0.99	-
Raspberry	<i>Rubus occidentalis</i>	1	4	12	11.11	0.15	-
Beech	<i>Fagus grandifolia</i>	6	18	54	11.11	0.25	-
Walnut	<i>Juglans nigra</i>	1	4	12	11.11	0.15	-
TOTAL		88	370	1110	100		
Mean		5.18	21.76	65.29	9.38 ^Δ		
⁺ SD					7.09		

*TPP = Tree Positive Percentage; ^ΔMPP = Mean positive Percentage; ⁺SD = Standard deviation; +++ = significantly higher than the mean (Q > 0.95 .i.e. reject null hypothesis); - = not significantly higher than the mean (Q < 0.95 .i.e. accept null hypothesis).

Table 2.4: Fisher's exact test analysis showing tree type preference

	Deciduous trees	Coniferous trees	Total
Samples without yeast	676	112	788
Samples with yeasts	62	20	82
Total	738	132	870

*P = 0.0223



Figure 2.1: Environmental sampling in the greater Hamilton region. The sites A-H sampled between May 26, 2008 and September 28, 2008 are indicated in the map.

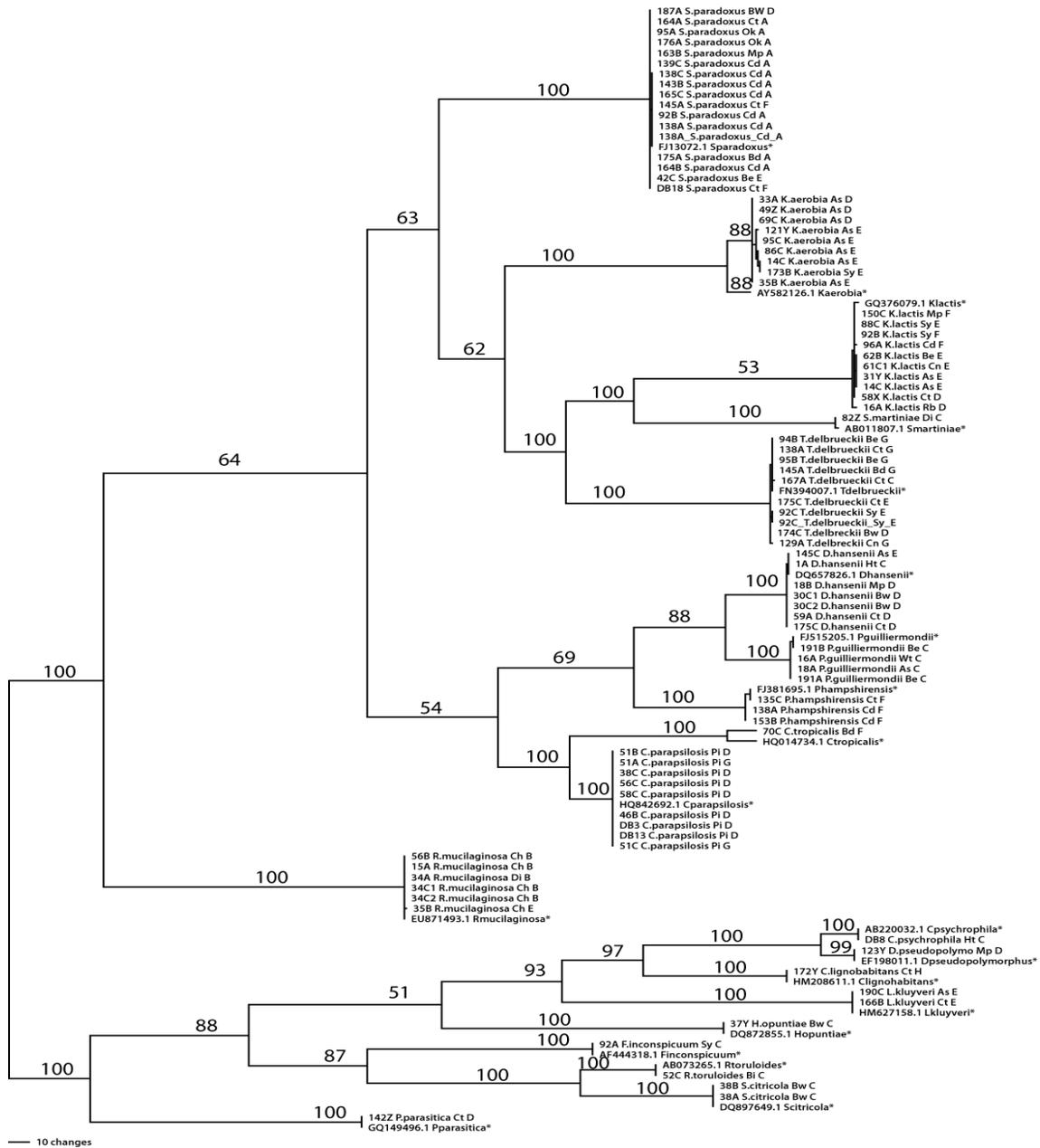


Figure 2.2: ITS tree of all the 88 environmental yeast strains sampled. Each strain in the phylogram was named in the following manner: Strain identification number _ Species abbreviation _ tree species isolated from _ geographical site isolated from. * denotes reference sequences for each respective yeast species. The names of the tree species have been abbreviated in the following manner: Ash (As), Basswood (Bw), Birch (Bi), Cedar (Cd), Cherry (Ch), Chestnut (Cn), Cotton wood (Ct), Hawthorn (Ht), Maple (Mp), Oak (Ok), Pine (Pi), Sycamore (Sy), Raspberry (Rb) and Beech (Be). Detritus and bird droppings have been abbreviated as Di and Bd.

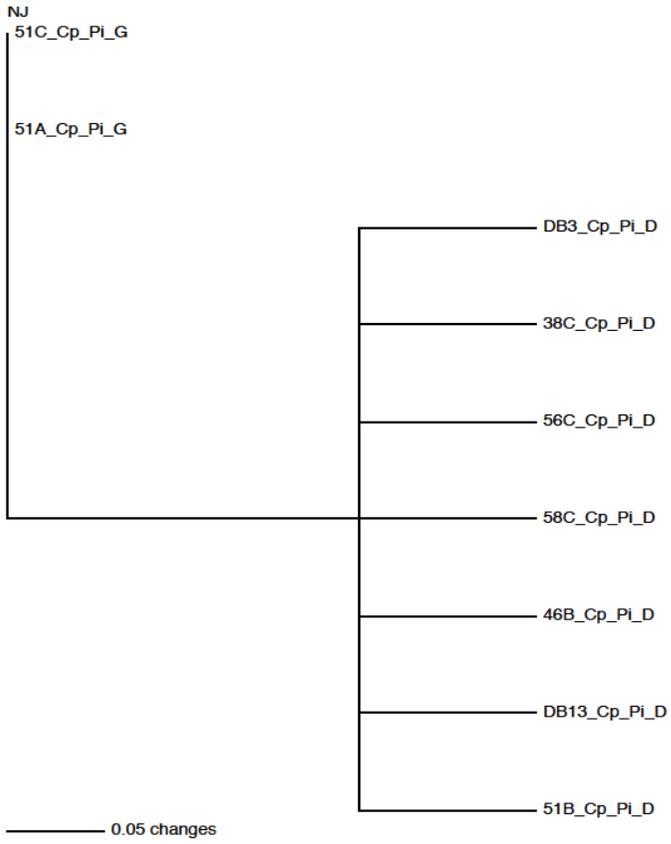


Figure 2.3: Phylogram of 9 *C.parapsilosis* (Cp) strains. 8 genotypes were identified and none of the genotypes were present at more than one site. All the *C.parapsilosis* strains were identified to grow only on Pine (Pi) trees. Each strain in the phylogram was named in the following manner: Strain identification number _ Species abbreviation _ tree species isolated from _ geographical site isolated from.

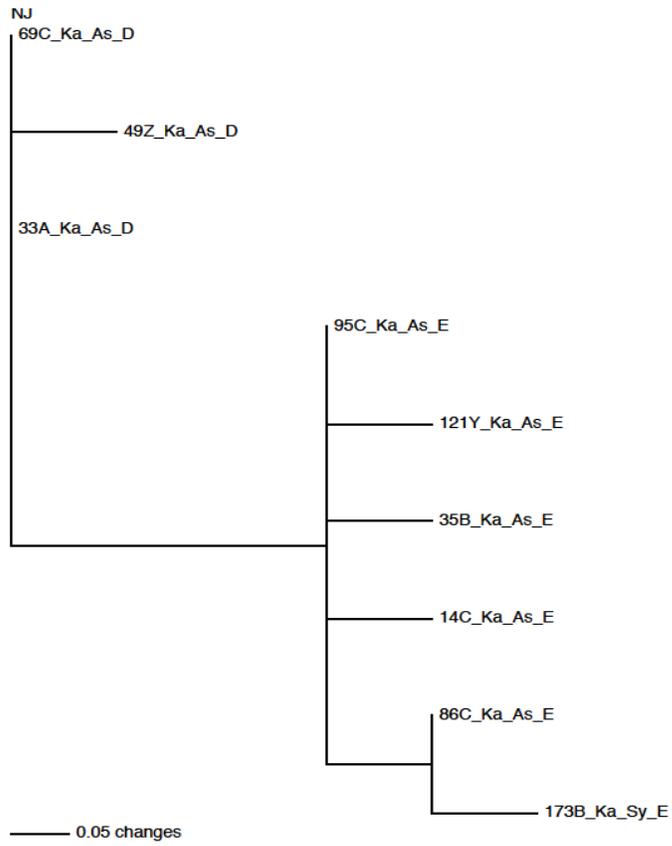


Figure 2.4: Phylogram of 9 *K.aerobia* (Ka) strains. 8 genotypes were identified and none of the genotypes were present at more than one site. 8 of the *K.aerobia* strains were identified to grow on Ash trees (As) while 1 strain was identified to grow on Sycamore (Sy). Each strain in the phylogram, was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

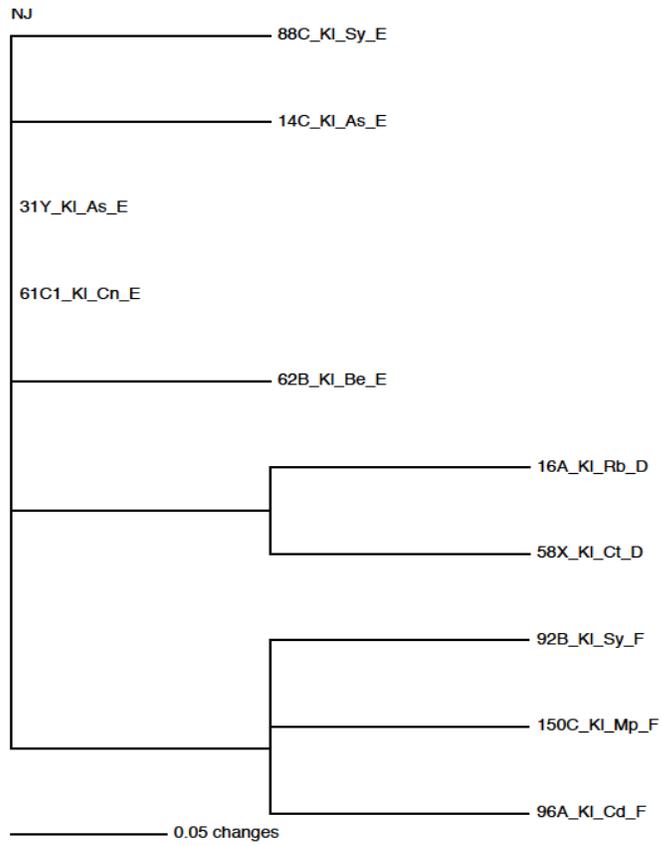


Figure 2.5: Phylogram of 10 *K.lactis* (Kl) strains. 9 genotypes were identified and none of the genotypes were present at more than one site. *K.lactis* strains were identified to grow on many trees namely, Sycamore (Sy), Ash (As), Beech (Be), Cotton wood (Ct), Maple (Mp) and Cedar (Cd). *K.lactis* was also identified to grow on the shrub Raspberry (Rb). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

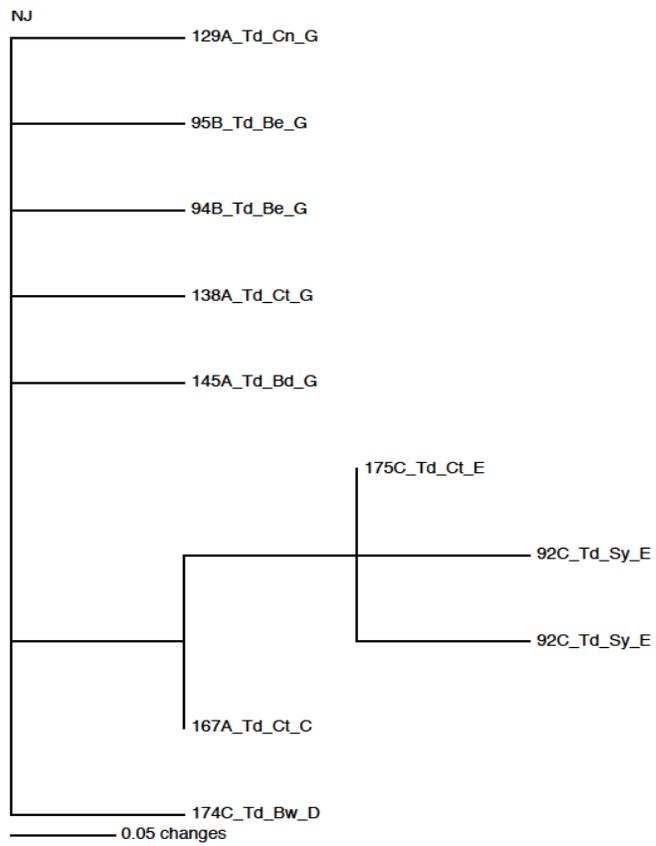


Figure 2.6: Phylogram of 10 *T.delbrueckii* (Tm) strains. 10 genotypes were identified and none of the genotypes were present at more than one site. *T.delbrueckii* strains were identified to grow on many trees namely, Chest nut (Cn), Beech (Be), Cotton wood (Ct), Basswood (Bd) and Sycamore (Sy). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

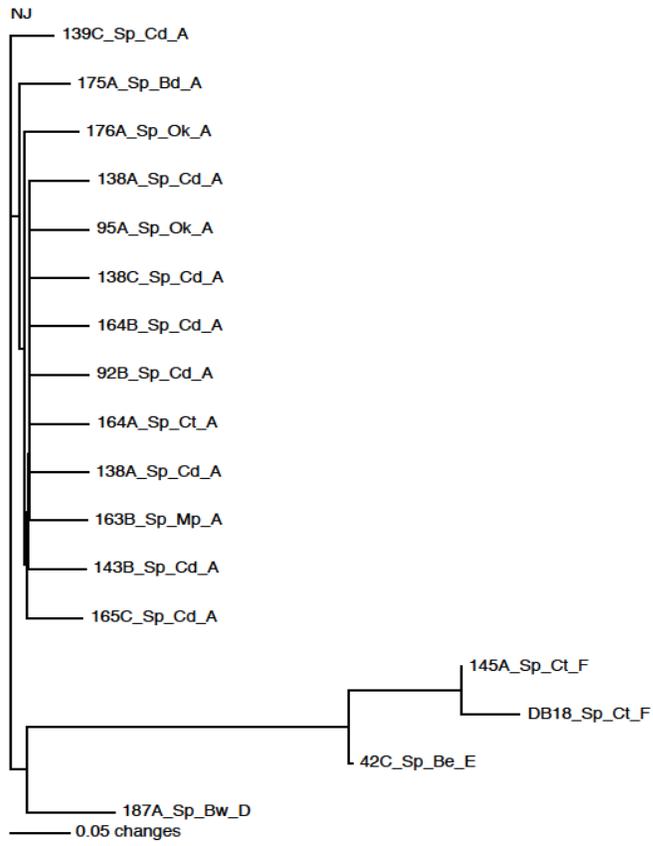


Figure 2.7: Phylogram of 17 *S.paradoxus* (Sp) strains. 17 genotypes were identified and none of the genotypes were present at more than one site. *S.paradoxus* strains were identified to grow on many trees namely, Cedar (Cd), Basswood (Bd), Oak (Ok), Cotton wood (Ct), Maple (Mp) and Beech (Be). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

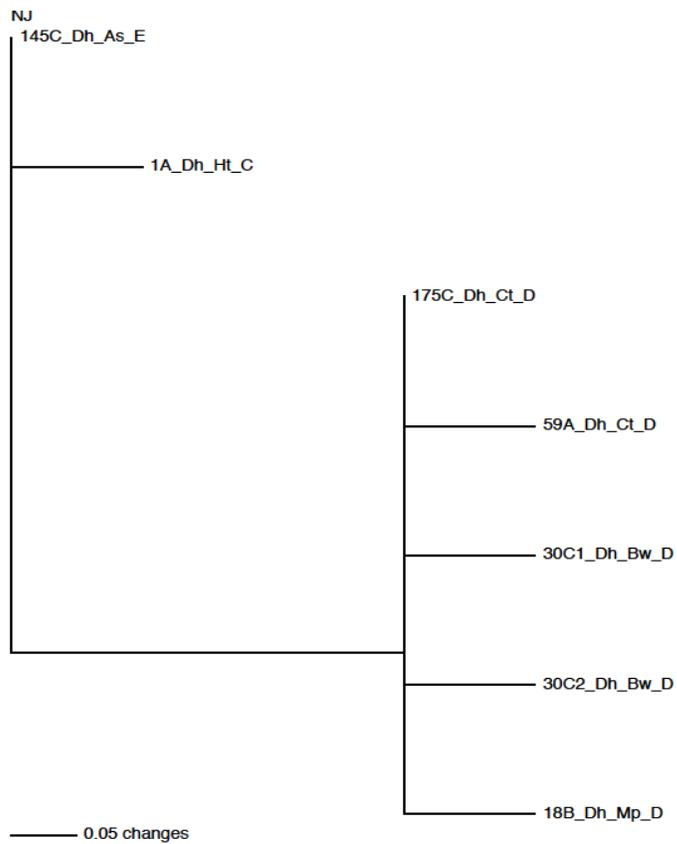


Figure 2.8: Phylogram of 7 *D.hansenii* (Dh) strains. 7 genotypes were identified and none of the genotypes were present at more than one site. *D.hansenii* strains were identified to grow on many trees namely, Ash (As), Hawthorn (Ht), Cotton wood (Ct), Basswood (Bd) and Maple (Mp). Each strain in the phylogram was named in the following manner: Strain identification number _Species abbreviation _ Tree species isolated from _ Geographical site isolated from.

Concluding statements and Future Directions:

Our PCR fingerprinting work showed nosocomial clusters to play an important role in triggering candidemia in the HHS hospitals. We identified that candidemia in 33% of the patients admitted between January 2005 to February 2009 was caused by strains in nosocomial clusters. A total of 18 clusters were identified, with 13 of them found in one hospital and 5 found in another. Furthermore we found that for each of the four species with multiple strains in our sample, strains isolated closer to each other temporally were overall genetically more similar to each other as well, which suggested that nosocomial sources likely caused repeated candidemia infections.

With regards to environmental profiling we obtained 88 yeast isolates across the 8 geographical sites. The 88 yeast isolates belonged to 20 species. Of the 20 species, 6 are opportunistic pathogens, including *C. tropicalis* and *C. parapsilosis*. Even though the sampling sites were within the same municipality and were connected by waterways, human and bird movement, the yeast populations were heterogeneous. Among the 14 tree species sampled, cedar, cottonwood and basswood hollows had relatively high yeast colonization rates. The lack of a large repertoire of *Candida* species in and around the region of Hamilton city prevented us from understanding the role played by various environmental *Candida* species in causing infections among patients in clinical settings. We are currently collecting mucosal and blood stream invasive samples from patients. In the future we intend to compare the genotypes of the mucosal and invasive *Candida* samples. This would help us understand the proportion of the clinical infections that might be caused by indigenous strains prevalent within the hospital environments versus the exogenous strains brought in by the patients themselves.

Apart from genotype studies, one should also try to test the antifungal Minimum Inhibitory Concentration (MIC) of the mucosal and the environmental *Candida* strains we obtained in this project and compare these values to those causing invasive infections. The mucosal and invasive sample MIC values should be correlated and studied in the context of antifungal treatment strategies the patients had undergone in the recent past. This will not only allow researchers to learn the variation in the level of anti-fungal resistance present among the commensal *Candida* human flora and those present in natural environments, but would also help clinicians to learn the effect of their anti-fungal treatments. This could potentially provide valuable information for developing better treatment strategies.