MOLECULAR PROFILING AND PHYLOGENETIC ANALYSIS OF THE *ERG11* GENE OF FLUCONAZOLE RESISTANT STRAINS OF *CANDIDA* SPECIES ISOLATED FROM HUMANS AND DOGS OF REPRODUCTIVE AGE

BY

ODIBA, AROME SOLOMON
(PG/Ph.D/14/76654)

DEPARTMENT OF BIOCHEMISTRY
FACULTY OF BIOLOGICAL SCIENCES
UNIVERSITY OF NIGERIA
NSUKKA

OCTOBER, 2017

TITLE PAGE

MOLECULAR PROFILING AND PHYLOGENETIC ANALYSIS OF THE *ERG11* GENE OF FLUCONAZOLE RESISTANT STRAINS OF CANDIDA SPECIES ISOLATED FROM HUMANS AND DOGS OF REPRODUCTIVE AGE

A DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF DOCTOR OF PHOLOSOPHY (Ph.D) IN MOLECULAR BIOLOGY AND GENETICS, UNIVERSITY OF NIGERIA, NSUKKA

BY

ODIBA, AROME SOLOMON

(PG/Ph.D/14/76654)

DEPARTMENT OF BIOCHEMISTRY
FACULTY OF BIOLOGICAL SCIENCES
UNIVERSITY OF NIGERIA
NSUKKA

SUPERVISORS: PROF. B. C. NWANGUMA
PROF. I. N. E. ONWURAH

OCTOBER, 2017

CERTIFICATION

Odiba, Arome Solomon, a postgraduate student of the Department of Biochemistry, University of Nigeria Nsukka, with Registration Number, PG/Ph.D/14/76654, has satisfactorily completed the dissertation in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy (Ph.D) in Molecular Biology and Genetics. The work embodied in this dissertation is original and has not been submitted in part or full for any other diploma or degree of this or any other university to the best of our knowledge.

…… ……

Prof. B. C. Nwanguma (Supervisor)

Prof. I. N. E. Onwurah (Supervisor)

…… ……

Prof. F. C Chilaka (Head of Department)

External Examiner

DEDICATION

This work is dedicated to all researchers and scientists who are working relentlessly at understanding the language of God hidden in the DNA.

ACKNOWLEDGEMENTS

All the knowledge I have acquired is through the grace of God. He provides life, wisdom and understanding for every step, and I am full of thanks to Him.

I am grateful to my supervisor, Prof. B. C. Nwanguma who never forced anything on me; rather he created the environment necessary for me to thrive in the field of my passion; Molecular Biology and Genetics. To my second supervisor, Prof. I. N. E Onwurah, who sent me to where I could get all the appropriate help I needed, I say I am grateful. My thanks go to the current Head of Department, Prof. F. C. Chilaka, who was prompt in validating all the approvals I needed for this work. I am very grateful to my lecturers and the entire staff of this department for contributing to my life. My loads of appreciation also goes to Dr. A. Mgbeahuruike who inspired the conception of this research. With all sense of gratitude, I say thanks to Prof. I. M. Ezeonu of the Department of Microbiology, University of Nigeria, Nsukka, for her immense contribution to the molecular phase of this work. My sincere thanks in a similar trend goes to Mr. N. N. Ntun of the Department of Microbiology, University of Nigeria, Nsukka, who guided me through the microbiology phase of this work. To Mr. Durojaye Olanrewaju, a friend, and a colleague who shared his bioinformatics skills with me, I am grateful. To Dr. Parker Elijah Joshua, your effort has been outstanding and you are an immediate help in time of need. My appreciation extends to the entire staff of the Library Department and other student related offices that provided all the necessary resources for the successful completion of this program. To the entire family of Odiba, James Alidu, my God given family, and to everyone in my life, I am grateful.

ABSTRACT

The clinical resistance of *Candida* species to antifungal medications, particularly fluconazole, is rising. A major mechanism responsible for this resistance is the alteration in the nucleotide sequence of the gene that codes for the *ERG11* protein, a key enzyme in the ergosterol synthesis pathway, which remains the target of fluconazole. This study investigated the distribution of different species of *Candida* in human and dog vaginal swabs, the susceptibility of the *Candida* species to fluconazole, the profile of the *ERG11* gene, and the phylogenetic relationship of the *Candida* species based on the nucleotide sequence of the *ERG11* gene. A total of 57 human samples and 7 dog samples were screened for the presence of *Candida* species. Twenty-eight (28) of the human samples were positive (+ve) to yeast growth. A total of 37 *Candida* isolates were obtained from the 28 human specimens that were positive to yeast growth. Of the 37 isolates, 13 (35%) were *C. albicans*, 4 (9%) *C. glabrata*, 4 (9%) *C. krusei*, 2 (6%) C. *tropicalis*, and 14 (38%) other *Candida* species. Of the 28 human specimens that were positive to *Candida* growth, 21 had single species, 5 had two different species and 2 had three different species. Four different species of *Candida*, including *C. albicans, C. tropicalis, C. krusei,* and *C. glabrata* were identified. The antifungal susceptibility test revealed that 33 (89.2%) of the *Candida* species were susceptible (≥19mm) to 25 µg fluconazole. The most fluconazole-resistant isolate was *C. glabrata*, while the most fluconazole-susceptible isolate was a *C. albicans*. Both isolates were obtained from humans. The four (4) isolates whose *ERG11* genes were sequenced, were the most resistant, *C. glabrata*, the susceptible dose-dependent, *C. albicans*, the most susceptible, *C. albicans* and the dog isolate, *C. krusei*. The nucleotide sequence lengths of the *ERG11* gene of these isolates varied from 1431 bases (*Can Iso-001*) to 1668 bases (*Can Iso-029*). Similarly, the molecular weight varied from 56183.10Da (*Can Iso-001*) to 65139.25Da (*Can Iso-029*), while the isoelectric point varied from 8.88 pI (*Can Iso-029*) to 9.60 pI (*Can Iso-001*). The predicted half-life ($t_{1/2}$) of these proteins in mammalian cells was 100 hours, and the instability coefficients were 35.70, 36.98, 44.50 and 37.69, for *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029, respectively. The grand average of hydrophobicity (GRAVY) of these *Candida ERG11* proteins were 0.195, 0.478, 0.195 and 0.576, respectively. The four *Candida ERG11* proteins were predicted to be localized in the plasma membrane. The first 12 amino acids in the multiple sequence alignment (MSA) make-up the first major conserved domain, while the second major conserved regions are in positions 331-337, 332-338, 330-336 and 330-336, for *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029, respectively. The total residues of alpha helix, beta pleated sheets and turns for these proteins varied. The most fluconazole-resistant isolate (*Can Iso*-001) had the highest percentage of α-helix in its *ERG11* protein, while the most fluconazole-susceptible isolate (*Can Iso-028*) had the lowest percentage of α-helix in its *ERG11* protein. There was a low-level similarity between the *Can Iso-017* and *Can Iso-028 ERG11* tertiary structural models. The Tertiary protein structure of the *Can Iso-001 ERG11* was not similar to the other isolates. *Can Iso-001 ERG11* protein had no single fluconazole-binding site, while *Can Iso-017, Can Iso-028* and *Can Iso-029 ERG11* proteins had unique binding sites to which the drug can effectively bind. Similarly, *Can Iso-001 ERG11* protein possesses 10 antigenicity sites, while the *Can Iso-017, Can Iso-028* and *Can Iso-029 ERG11* proteins, possess 16, 21 and 15 antigenicity sites, respectively. The amino acids; 2-8: ETVIDGI was identified as a disease-causing region common to all, in addition to other disease-causing regions that were peculiar to each of the isolates. The phylogenetic analysis based on the *ERG11* gene showed that, the four isolates were closely related. However, the most resistant isolate (*Can Iso*-001) and the dog isolate (*Can Iso*-029), seemed to have originated from a common ancestor, implying an even closer evolutionary relatedness.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABC	ATP-binding cassette
AgNPs	silver nanoparticles
BLAST	Basic Local Alignment Search Tool
BLASTn	Nucleotide Basic Local Alignment Search Tool
Can Iso	Candida Isolate
CDR1	Candida drug resistance 1
CDR2	Candida drug resistance 2
CDR1p	Candida drug resistance 1 protein
CDR2p	Candida drug resistance 2 protein
CFSSP	Chou & Fasman Secondary Structure Prediction server
CLSI	Clinical and Laboratory Standards Institute
DHVS	Dog High Vaginal Swab
EMBOSS	European Molecular Biology Open Software Suite
EMEA	European Medicines Agency
ERG11	ERGosterol 11
ERG11-F	Ergosterol 11 Forward Primer
ERG11p	ERGosterol 11 protein
FDA	Food and Drug Administration
GRAVY	Grand Average of Hydropathicity
HHVS	Human High Vaginal Swab
MCL	Maximum Composite Likelihood
MDR1	Multi-drug resistance 1
MEGA7	Molecular Evolutionary Genetics Analysis 7
MFS	Major Facilitator Superfamily

MSA	Multiple Sequence Alignment
NAC	Non-albicans Candida
NCBI	National Center for Biotechnology Information
NIH	National Institutes of Health
SDA	Sabouraud Dextrose Agar
SDD	Susceptible Dose-Dependent
TAC1	Transcriptional activator of CDR genes 1
UPC2	Uptake control 2
VVC	vulvovaginal candidiasis
5-FC	5-Fluorocytosine

CHAPTER ONE
INTRODUCTION

The incidence and prevalence of invasive fungal infections have increased since the 1980s, especially in the large population of immunocompromised patients (Espinel-Ingroff *et al.,* 2009). *Candida* species are important human fungal pathogens that cause both mucosal and deep tissue infections. *Candida* species belong to the normal microbiota of an individual's mucosal oral cavity, gastrointestinal tract, the vagina and other endo-mucosal surfaces (Shao *et al.,* 2007), and are responsible for various clinical manifestations from mucocutaneous overgrowth to bloodstream infections (Eggimann *et al.,* 2003). These yeasts are commensal in healthy humans, and, possibly could cause systemic infections in immunocompromised situations. More than 17 different *Candida* species are known to be aetiological agents of human infection. However, more than 90% of invasive *Candida* infections are caused by *Candida albicans*, *Candida glabrata*, *Candida parapsilosis*, *Candida tropicalis* and *Candida krusei* (Ortega *et al.,* 2011; Pfaller *et al.,* 2015). *Candida albicans* and other non-*albicans Candida* (NAC) species such, as *C. glabrata, C. parapsilosis, C. tropicalis,* and *C. krusei* are capable of causing superficial oral, vaginal mucosa, disseminated bloodstream and deep-tissue infections. However, species involvement varies by infection site and by geographical location (Sharifzadeh *et al.,* 2013; Cleveland *et al.,* 2015; Klingspor *et al.,* 2015). *C. glabrata* is the most common NAC species found to be the causative agent in vulvovaginal candidiasis (VVC) (Vermitsky *et al.,* 2008; Mahmoudi Rad *et al.,* 2012). *C. parapsilosis* is well known for its threat to the pediatric population, as it is responsible for 17–50% of all fungemia in infants and neonates (Krcmery *et al.,* 1999). *C. parapsilosis* is also second to *C. albicans* in incidence as a cause of *Candida* endocarditis with mortality rates between 42% and 65% (Garzoni *et al.,* 2007). *C. tropicalis* infections are commonly associated with malignancy, with some studies reporting higher prevalence among patients with hematologic diseases such as acute myeloid leukemia (Tang *et al.,* 2015; Cornely *et al.,* 2015). The mortality rate ranges from 30% to 70%, with the highest rates commonly observed among the elderly (Cornely *et al.,* 2015; Wang *et al.,* 2015). *C. krusei* is the fourth most common NAC species associated with invasive candidiasis and candidemia, accounting for approximately 2.7% of NAC species isolated in clinical studies (Pfaller *et al.,* 2014). The pathogenicity of *Candida* species is attributed to certain virulence factors, such as adherence to mucosal surfaces, ability to evade host defenses, biofilm

formation (on host tissue and on medical devices) and the production of tissue-damaging hydrolytic enzymes such as proteases, phospholipases and haemolysin (Verstrepen and Klis, 2006). Biofilms are biological communities of *Candida* with a high degree of organization residing in carbohydrate polymers, in which microorganisms form structured, coordinated and functional colonies. These biological communities are embedded in a highly organized self-created extracellular matrix, primarily composed of structural carbohydrate polymers. Currently, an increase in the number of *Candida* species that are resistant to antifungal drugs is recognized worldwide (Ingham *et al.,* 2012). The increase in resistant strains necessitates the search for new molecular targets in the organism, as well as new antifungal agents to replace existing ones. There are numerous classes of compounds used to treat *Candida* infections. The polyenes, azoles, echinocandins, nucleoside analogs, and allylamines are used with varying efficacy, depending on the type and site of infection, as well as the susceptibility of the *Candida* species to the agent (Pfaller *et al.,* 2015; Pappas *et al.,* 2015). The most commonly prescribed antifungal agents used for most *C. albicans* infections is fluconazole; a member of the azole class of antifungals (Pfaller *et al.,* 2010). Azoles inhibit lanosterol 14 alpha-demethylase, encoded by the *ERG11* gene, which is an enzyme involved in the biosynthesis of the fungal-specific membrane sterol; ergosterol (Lortholary *et al.,* 2011; Fothergill *et al.,* 2014). Azole antifungals have long provided effective treatment for *Candida* infections, however, recent epidemiological studies indicated that intrinsic azole resistance in some *Candida* species, including the onset of high-level azole resistance is a problem of critical importance in clinical settings (Pfaller *et al.,* 2015; Shields *et al.,* 2015). While extensive studies to elucidate the molecular machineries of high-level azole resistance in *C. albicans* has uncovered the role of ergosterol biosynthesis gene (*ERG11* gene) mutation and drug efflux pump upregulation as key mediators of azole resistance, there are other factors at play that contribute significantly to such resistance. From previous studies, there exist clear mutations in the *ERG11* gene that are found to influence azole resistance in clinical isolates among *Candida* species (Pfaller *et al.,* 2015; Shields *et al.,* 2015). As azole resistance continues to emerge in these species, a better understanding of the important differences among resistance machineries employed by these species is needed in order to circumvent this crucial clinical problem. Azole antifungals such as fluconazole are often the preferred treatment for many *Candida* infections due to their efficiency, lower costs, limited toxicity and availability in oral administration. There is an extensive documentation of intrinsic and developed resistance to azole antifungals among

numerous *Candida* species at different degrees, and in different geographical locations (Vermitsky *et al.,* 2008; Sharifzadeh *et al.,* 2013; Cleveland *et al.,* 2015; Klingspor *et al.,* 2015). As the frequency of azole resistant *Candida* isolates in the clinical setting increases, it is essential to elucidate the machineries of such resistance in order to both preserve and improve upon the azole class of antifungals for the treatment of *Candida* infections. Many studies have documented the ability of *Candida* to develop high-level resistance to azole antifungals (Oxman *et al.,* 2010; Lortholary *et al.,* 2011), therefore, a clear understanding of molecular machineries driving the intrinsic and onset of high-level azole resistance is necessary.

1.1 Epidemiology of *Candida* Infection

Numerous *Candida* species are commensals and colonize the skin and mucosal surfaces of humans. Critically ill or otherwise immunocompromised patients are more predisposed to developing both superficial and life-threatening *Candida* infections (Hasan *et al.,* 2009). *C. albicans* is the predominant cause of invasive fungal infections (Horn *et al.,* 2009), and represents a serious public health challenge with increasing medical and economic importance. This is due to the high mortality rates and increased costs of care (Lai *et al.,* 2012). Although *C. albicans* is the most prevalent species involved in invasive fungal infections, the incidence of infections due to non-*albicans* species is increasing. In a recent study, it was found that 28.3% of patients exhibited invasive fungal infection, with *C. albicans* as the most frequently isolated (58%), followed by *C. tropicalis* (17%) and *C. glabrata* (15%) (Yapar, 2014). In Europe, an analysis showed that more than half of the cases of Candidaemia were caused by *C. albicans*, and the incidence rates for non-*albicans* Candidaemia infections were 14% each for *C. glabrata* and *C. parapsilosis*, 7% for *C. tropicalis* and 2% for *C. krusei* (Tortorano *et al.,* 2006). Changes in the epidemiology have also been observed in Latin America. For instance, in Chile, the prevalence of *C. albicans* has changed, and a progressive increase of non-*albicans* infection has been observed; *C. parapsilosis* is the most frequent species, followed by *C. tropicalis* and *C. glabrata*. According to the Brazilian Network Candidaemia study, *C. albicans* accounted for 40.9% of cases in Brazil, followed by *C. tropicalis* (20.9%), *C. parapsilosis* (20.5%) and *C. glabrata* (4.9%) (Nucci *et al.,* 2010). For Nigeria, however, such a cohort of study is yet to be executed and reported.

1.2 Pathogenicity of *Candida* Species

Candida species are considered important pathogens due to their versatility and ability to survive in various anatomical sites. *Candida* species are commensal eukaryotic opportunistic pathogens that reside on the mucosa of the gastrointestinal tract, mouth, oesophagus, vagina and other mucosal linings in an asymptomatic manner. While *Candida* species can infect different anatomical sites of the human host, there are indications that immune protection is site-specific. Cutaneous candidiasis and vaginal candidiasis are more likely to be connected with a phagocytic response involving neutrophils and mononuclear phagocytes (Vidigal and Svidzinski, 2009). It can however, become one of the most significant causes of death if not treated effectively (Wisplinghoff *et al.*, 2006; Vincent *et al.*, 2009). Infection caused by *Candida* is called candidiasis or candidosis, with a wide spectrum of clinical manifestations. It can be classified as superficial (as with cutaneous and mucosal infections), deep, widespread and of high severity (as is the case of invasive candidiasis). In years past, fungi emerged as major causes of nosocomial infections, mainly affecting immunocompromised patients or those who were hospitalized for long periods as a result of serious underlying diseases (Vidigal and Svidzinski, 2009). Most people usually have a single strain of *Candida* in different places in the body for a long period, while a comparatively lower number of individuals have more than one strain or species at the same time, as frequently observed among hospitalized patients (Kojic and Darouiche, 2004; Klotz *et al.*, 2007). Virulence in *C. albicans* comprises of host recognition, which enables the pathogen to bind to host cells and proteins, with degradative enzymes playing special roles in their virulence. Extracellular hydrolytic enzymes appear to play an important part in adherence, tissue penetration, invasion and the damage of host tissues (Silva *et al.*, 2009). *Candida* pathogenicity is aided by a number of virulence factors, the most important of which are those for adherence to host tissues and medical devices, biofilm formation and secretion of hydrolytic enzymes (e.g. proteases, phospholipases and haemolysins). The primary mechanism in the fungal colonization of human tissues is adherence to host surfaces; a process controlled and induced by numerous cell-signaling cascades in both the fungus and its immediate surrounding environment. The initial attachment of *Candida* cells is facilitated by non-specific factors (electrostatic forces and hydrophobicity) and promoted by specific adhesins that are present on the surface of fungal cells, and can identify ligands such as proteins (including fibrinogen and fibronectin). Adhesins can unambiguously bind to amino acids and sugars on the surface of other cells (Verstrepen and Klis, 2006). The presence of biofilm

matrix restricts the penetration of drugs, through the formation of a diffusion barrier, causing clinical problems of concern, by increasing resistance to antifungal therapy, since only the superficial layers are in contact with lethal doses of the drug (Kojic and Darouiche, 2004). Recent evidence suggests that many of the diseases produced by *C. albicans* are associated with biofilm growth (Ramage and Lo´pez-Ribot, 2005). Biofilms can thrive on any moist biotic or abiotic surface as a form of protection for their proliferation and defense against antifungal treatment, as well as to withstanding competitive pressure from other organisms. This strategy also encourages symbiotic relationships, and allows survival in hostile environments (Davey and O'toole, 2000; Ramage and Lopez-Rib, 2005). In lung infections, the association between *C. albicans* and *Pseudomonas aeruginosa* is an instance of an antagonistic interaction between bacteria and fungi, where *P. aeruginosa* kills yeast hyphae and biofilms of *C. albicans* (Morales *et al.,* 2010). Generally, the biofilm matrix comprises carbohydrates, proteins, phosphorus and hexosamines; though, environmental circumstances such as medium composition, pH, oxygen concentration and the strain can affect biofilm formation along with matrix composition. For instance, *C. parapsilosis* biofilms contain large amounts of carbohydrates, and the protein content is lower in comparison with the biofilms of *C. glabrata* and *C. tropicalis* (Silva *et al.,* 2009). According to the US National Institutes of Health (NIH), biofilms are the most common form of microbial growth in nature, and cause the majority of infections in humans (Nett *et al.,* 2010). Several studies have shown that a relationship exists between an increase in the activity of extracellular hydrolytic enzymes and an increase in the pathogenic capacity of the yeasts, leading to clinical signs of severe candidiasis (Ingham *et al.,* 2012). The roles of these fungal extracellular lipases include the digestion of lipids for nutrient acquisition, adhesion to host cells and tissues, unspecific initiation of inflammatory processes and self-defense through lysing of any competing microflora (Verstrepen and Klis, 2006). Iron, an inorganic element, is also essential for the survival of microorganisms, including yeasts (*Candida*), and the capacity to obtain this element is contributory to the on-set of an infectious process (Dongari-Bagtzoglou *et al.,* 2009). Biofilms are hard to diagnose and treat, and have the possibility to serve as infectious reservoirs for a variety of microorganisms that include bacteria (such as *Staphylococcus epidermidis, Staphylococcus aureus* and *Enterococcus* species) and fungi (Klotz *et al.,* 2007; Harriott and Noverr, 2011). Biofilms however, thus far, have not been demonstrated in the gastrointestinal tract (Harriott and Noverr, 2011). *C. albicans* reversibly transforms from unicellular yeast cells to either pseudohyphae or hyphae (filamentous growth

form); a morphogenesis phenomenon. This phenomenon has been observed in *C. albicans* and *C. dubliniensis* (Bruder-Nascimento *et al.,* 2010). The growth of hyphae, a virulence mechanism, plays an important role in tissue invasion and resistance to phagocytosis (Jayatilake *et al.,* 2006). The morphological transformation from the yeast to the mycelial form (dimorphic switching) is induced by many environmental factors, such as serum, high temperatures (37^0C) and neutral pH (Yi *et al.,* 2011). Genetic analyses show that both yeast cells and hyphae are crucial for biofilm formation, suggesting that each cell type has a unique role in this process (Douglas, 2003).

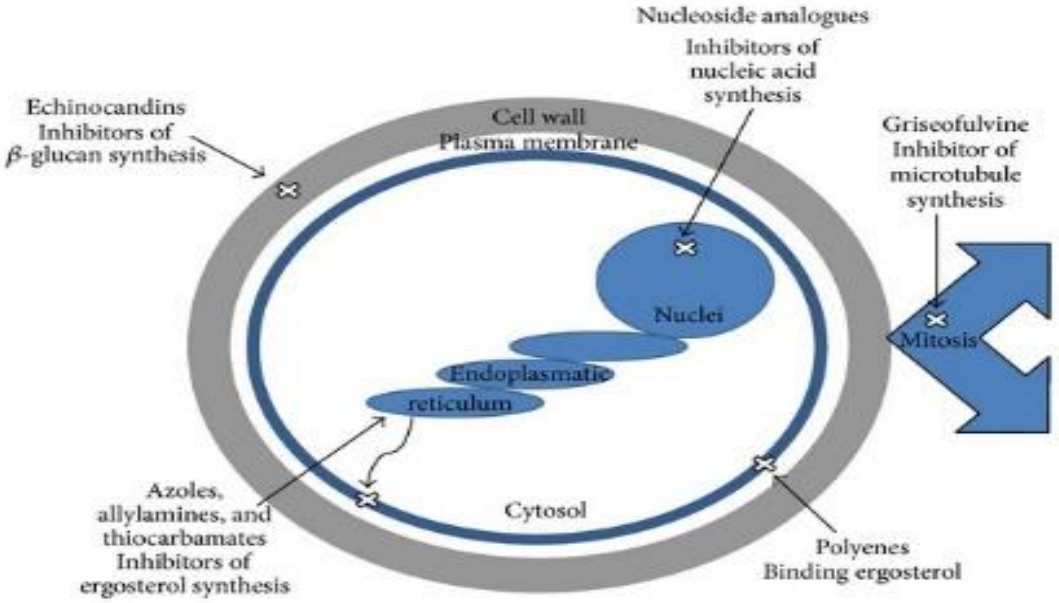
1.3.1 *Candida* Biofilms and Conventional Antifungals

Different antifungal classes utilize a different means to inhibit the growth of fungal pathogens (Pfaller *et al*., 2012). The molecular mechanisms of antifungal resistance are categorized as either primary or secondary, and are connected to intrinsic or acquired qualities of the fungal pathogen. This encompasses either interference with the antifungal machineries of the corresponding drug or a reduction in the drug levels. Resistance also surfaces when environmental influences lead to the colonization or replacement of a susceptible species with a resistant species. The antifungal properties of polyene and azole antifungals could be attributed to their actions on the fungal cell membrane, whereas echinocandins act by disrupting the fungal cell wall (Pfaller *et al*., 2012). The ability of *Candida* to form drug-resistant biofilms is a vital factor in its influence in human disease. The development of biofilms causes clinical complications of concern because they elevate the resistance to antifungal therapies, and the mechanism of biofilm resistance to antimicrobial agents is currently not completely known. A notable hypothesis to account for this resistance is that, the formation of a diffusion barrier, through the presence of the matrix, restricts the infiltration of drugs (Nett *et al.,* 2011); therefore, only the most superficial layers are in direct contact with lethal doses of antimicrobials. Numerous molecular tools of resistance to antifungal agents in *C. albicans* have been described. In particular, these include the increased efflux of antifungal agents as a result of the overexpression of the efflux genes, *CDR1*, *CDR2* and *MDR1*. The *CDR1* and *CDR2* are in the family of ABC (ATP-binding cassette) membrane transport proteins (Sardi *et al.,* 2011). *CDR1*, *CDR2* and other genes are often co-regulated, and are overexpressed at the same time (Staib *et al.,* 1999). Amino acid substitutions in the enzyme *ERG11*p (lanosterol 14-α-demethylase), encoded by the gene *ERG11* is also another possible factor that could be responsible (Flowers *et al*., 2015). Due to the resulting increased fungal infections, two triazoles (voriconazole and posaconazole) and three echinocandins (anidulafungin, caspofungin and micafungin) have been developed and

approved to treat and prevent these infections (Mattiuzzi and Giles, 2005) (Fig. 1). Among these three classes of antifungal agents currently in clinical use, only amphotericin B and the echinocandins, e.g. caspofungin, have validated reliable *in vitro* activity against *C. albicans* biofilms (Kuhn *et al.,* 2002). Despite the success of these two agents, *Candida* biofilm-related infections are awfully difficult to eradicate. The combined use of echinocandins with other drugs that have antifungal activity is becoming an important alternative form of therapy in mycoses resulting from fungi that are resistant to standard antifungal monotherapy in biofilm-associated diseases. In *C. albicans* biofilms, only a small subcategory of yeast cells are described to be highly resistant to amphotericin B (LaFleur *et al.,* 2006).

1.3.2 *Candida* Biofilm and New Antifungal Strategies

The rising occurrence of drug-resistant pathogens and the side-effects associated with existing antifungal substances has attracted attention in the direction of the antimicrobial action of natural products. The small number of medications available for fungal treatment, most of which are fungistatic, and the developing resistance to antifungal agents encourage the search for more suitable substitute treatments (Sardi *et al.,* 2011). Plants are found to be good options for obtaining a wide diversity of medications used in medicine due to their easy accessibility and application to various pathologies (Sardi *et al.,* 2011). Plants therefore have proven to be an excellent source for substances that are useable in the formulation of new antifungal agents (Holetz *et al.,* 2002). The extracts of some Romanian medicinal plants such as *Artemisia absinthium, Arnica montana* and *Urtica dioica*, have significant antimicrobial activities that are preferentially directed against fungi (*C. albicans*) and bacteria (*S. aureus*) (Stanciuc *et al.,* 2011). Humans have also been a source of antifungal agents. In a study by Rossignol *et al*. (2011), the 18-amino acid cationic tryptophan-rich ApoEdpL-W peptide, derived from human ApoE apolipoprotein was studied, and showed antifungal activity against pathogenic yeasts of the *Candida* genus, with the exception of *C. glabrata*. ApoEdpL-W proved to be active against planktonic cells and early-stage biofilms but with less activity against mature biofilms, possibly because of its attraction for extracellular matrix b-glucans (Stanciuc *et al.,* 2011).



Antifungal class	Mode of action	Drugs
Azoles	Inhibitors of lanosterol 14-α-demethylase	Miconazole Econazole Clotrimazole Ketoconazole Fluconazole Itraconazole Voriconazole Posaconazole
Echinocandins	Inhibitors of (1,3)-β-D-glucan synthase	Caspofungin Micafungin Anidulafungin
Polyenes	Binding ergosterol	Nystatin Amphotericin B
Nucleoside analogues	Inhibitor of DNA/RNA synthesis	Flucytosine
Allylamines	Inhibitors of squalene-epoxidase	Terbinafine Amorolfine Naftifine
Thiocarbamates	Inhibitors of squalene-epoxidase	Tolnaftate Tolciclate
Antibiotic	Interaction with β-tubulin	Griseofulvin

Figure 1: Primary targets and mode of action of several antifungal agents (Source; Spampinato and Leonardi, 2013)

Furthermore, ApoEdpL-W partly prevents the development of biofilms on medical devices. In a study by Mandal *et al*. (2011), *in vitro* growth inhibition of *C. tropicalis* and interrupted biofilm development in a concentration-dependent manner with purified Tn-AFP1 (the peptide derived from *Trapa natans*) was identified. The study also confirmed the downregulation of *MDR1* and *ERG11* gene expression. It has however, been documented as well that, usnic acid has inhibitory and fungicidal activity against biofilms of *C. parapsilosis* and *C. orthoparapsilosis* (Pires *et al.*, 2011).

Other studies have been done using natural products to assess interference with *C. albicans* biofilms (Taweechaisupapong *et al.*, 2010; Coleman *et al.*, 2010; Furletti *et al.*, 2011). A different antifungal strategy with high prospect is the use of silver nanoparticles (AgNPs) (Percival *et al*., 2005). Silver has antimicrobial activity and has a viable application in medicine, characterized by a well-tolerated tissue response and a low toxicity profile. Silver has been proven to inhibit the multiplication of microorganisms by interfering with microbial DNA replication within bacteria and fungi (Percival *et al.,* 2005), thereby hindering the production of biofilms. Silver ions also cause protein denaturation and cell death as a result of their reaction with nucleophilic amino acid residues in proteins, and their attachment to amino, imidazole, thiol, phosphate and carboxyl groups of membrane proteins or enzymes (Mastrolorenzo *et al.*, 2000; Percival *et al.*, 2005). Other promising approaches such as the use of nanoparticles, antibodies and photodynamic inactivation for treating fungal infections are being studied (Percival *et al.*, 2005).

1.4 Antifungal Drugs Classes in use for the Clinical Treatment of *Candida* Infections

The antifungal drugs used in clinical treatments are a few group of compounds presently available to treat mucosal or systemic infections with *Candida* spp. (Kathiravan *et al.*, 2012). The common classes of drugs used in clinical antifungal treatments are discussed.

1.4.1 Polyenes: Binding to Ergosterol

Polyenes, including nystatin and amphotericin B (both isolated from *Streptomyces spp*.) bind to ergosterol and interrupt the major lipidic constituent of the fungal cell membrane, resulting in the production of aqueous pores (Figure 1). As a result, the cellular permeability is changed and leads to the leakage of cytosolic constituents and consequently, the death of the fungi (Sanglard and Odds, 2002). One of the earliest examples of the polyenes is amphotericin B, considered to be the

reference drug for the treatment of most systemic fungal infections. Amphotericin B acts across all *Candida* spp., some species of *Aspergillus*, *Blastomyces dermatitidis*, and other fungi. It is not absorbed by the gastrointestinal tract and must be administered intravenously. Notwithstanding the more than 30 years of the clinical use of this drug, some resistance to amphotericin B has been developed (Sanglard and Odds, 2002). The main problem connected with the prophylactic use of conventional amphotericin B has always been its well-known side-effects and toxicity (Laniado-Labor´ın and Cabrales-Vargas, 2009). Resistance to amphotericin B has a tendency to be species dependent. A significant proportion of *C. krusei* and *C. glabrata* isolates resistant to amphotericin B have been reported (Kontoyiannis and Lewis, 2002; Pappas *et al.,* 2004; Krogh-Madsen *et al.,* 2006). *C. krusei* and *C. glabrata* are usually considered to be Susceptible Dose-Dependent (SDD) to amphotericin B, even though they show high MICs to polyenes (Kontoyiannis and Lewis, 2002; Pappas *et al.,* 2004; Krogh-Madsen *et al.,* 2006). As a result, higher than normal doses of amphotericin B have been recommended by the Infectious Diseases Society of America for treating candidemia to which *C. glabrata* and *C. krusei* are the causal organisms. During acquired resistance to polyenes, a decrease or lack of ergosterol content in cell membranes usually is observed. In effect, membranes of polyene-resistant *Candida* isolates have relatively low ergosterol composition when compared to those of polyene-susceptible isolates. These deficiencies are probably as a result of the loss of function (mutations) in the *ERG3* or *ERG6* genes which encode some of the enzymes that are associated with ergosterol biosynthesis (Espinel-Ingroff, 2009). On the whole, alterations in the steroids of the cell membrane, their phospholipid profile, their defense against oxidative damage, and mutations associated with genes involved in ergosterol biosynthesis (especially in *ERG11*and *ERG6*) are connected to resistance in *Candida* spp. In a study by Vandeputte *et al*. (2008), it was identified that a specific non-sense mutation (encoding a stop codon) in *ERG6* that resulted in a decrease in the ergosterol content of a clinical isolate of *C. glabrata* resistant to amphotericin B is key to resistance. Similarly, Hull *et al*. (2012) identified two mutations (T121V and T121I) in *ERG2* (substitution of the threonine in the 121st position by valine or isoleucine), in two isolates of *C. glabrata* with reduced sensitivity to amphotericin B.

1.4.2 Echinocandins: Inhibitors of the Glucan Synthesis

The echinocandins (anidulafungin, caspofungin and micafungin) are lipopeptidic antifungal compounds that inhibit the synthesis of fungal wall by noncompetitive blockage of the (1, 3)-β-

Dglucan synthase (Figure 1). Caspofungin, anidulafungin and micafungin, are proven to be effective in treating candidiasis (Chandrasekar and Sobel, 2006). This enzyme inhibition produces the formation of fungal cell walls with compromised structural integrity, finally leading to cell vulnerability and osmotic lysis (Grover, 2010). All these agents display concentration-dependent fungicidal activity against most species of *Candida* (Cappelletty and Eiselstein-McKitrick, 2007), and are approved by FDA, for the treatment of esophageal and invasive candidiasis, including candidemia (Ostrosky-Zeichner, *et al.*, 2005). Echinocandins are actually the most recent class of antifungal agent to be introduced in clinical practice for the management of fungal infections, especially those caused by *Candida* sp (Cappelletty and Eiselstein-Mckitrick, 2007). They inhibit β-(1, 3)-D-glucansynthetase, which is responsible for fungal cell wall production. This effectiveness notwithstanding, reports of echinocandin resistance in patients with infections due to *C. albicans, C. tropicalis, C. krusei and C. glabrata,* are on the increase. For instance, resistance in *C. glabrata* rose from 4.9% to 12.3% between 2001 and 2010 in a study at Duke University hospital Durham, North Carolina (Shields *et al*., 2015). Even more, emergence of co-resistance to both echinocandins and azoles in clinical isolates of *C. glabrata* has also been reported (Garcia-Effron *et al.,* 2008). In addition, intrinsic echinocandin resistance of *C. orthopsilosis, C. parapsilosis,* C. *guilliermondii* and *C. metapsilosis,* has been described (Garcia-Effron *et al.,* 2008). Advanced resistance has also been attributed to point mutations in the *FKS1* and/or *FKS2* genes, which encode the (1, 3)-β-D-glucan synthase complex (Balashov *et al.,* 2006). These mutations in *FKS1* was shown not to alter substrate binding, however, it lowered V_{max} values (Garcia-Effron *et al.,* 2009). The decrease in sensitivity to echinocandins is discovered to be associated with mutations in the *FKS1*p and *FKS2*p subunits of the β-1,3-glucansynthetase complex, which is required for the production of β-1,3 glucan, an essential constituent of the *Candida* cell wall (Desnos-Ollivier *et al.,* 2008). Precisely, mutations have been found in two regions; the hot spot 1 and hot spot 2 (composed of 9 and 8 amino acids, respectively), which are found in both genes. These mutations in *FKS1* and *FKS2* result in the inability of the echinocandins to block the production of 1, 3-β- glucan. Mutations in hot spot 1 of *FKS1* and *FKS2* are the most predominant among a variety of fungal species resistant to this group of drugs. Findings by Zimbeck *et al*. (2010) support associations between mutations in the hot spot1 of both *FKS1* and *FKS2* in *C. glabrata* resistant to echinocandins (Naicker *et al.,* 2016). In one of the discovered mutations in the hotspot 1 region of the FKS2 gene, serine was replaced by phenylalanine at

position 663 (S663F). Other mutations, such as the substitution of arginine with lysine at amino acid position 1377 (R1377K) was identified in the FKS2 hotspot 2 region. Shields *et al*. (2015) found *Candida* spp. FKS mutations in 5% of sequenced isolates and 2% of overall isolates. It is observed that FKS mutations are uncommon among non-*C. glabrata* species, even with prior echinocandin exposure (Shields *et al.*, 2015).

1.4.3 Nucleoside Analogues: The Inhibitors of DNA/RNA Synthesis

Flucytosine (5-fluorocytosine or 5-FC) is a pyrimidine analogue capable of being transported into fungal cells by cytosine permeases. It is then deaminated to 5-fluorouracil, and further phosphorylated to 5-fluorodeoxyuridine monophosphate. This fluorinated nucleotide inhibits thymidylate synthase and thus interferes with DNA synthesis (Vermes *et al.,* 2000). Furthermore, the 5-fluorodeoxyuridine monophosphate can be phosphorylated and incorporated into RNA, thus affecting RNA and protein synthesis (Onishi *et al.,* 2000). These final molecules behave as antimetabolites and impede the normal biosynthesis of nucleic acids and nucleotides vital for fungal growth. This drug is prescribed for infections caused by *Candida* spp., *Cryptococcus neoformans, Aspergillus and Torulopsis* spp. (Vermes *et al.,* 2000). Primary resistance to flucytosine is actually low (<2%). Secondary resistance depends on the inactivation of different enzymes of the pyrimidine pathway as described in the following:

i. Decreased Intracellular Drug Accumulation: under this mechanism, uptake of the drug is affected by point mutations in the *FCY2* gene that codes for the cytosine permease.
ii. Counteraction of the Effect of the Drug: under this mechanism, acquired resistance to flucytosine also results from point mutations in the *FCY1* gene that codes for the cytosine deaminase or *FUR1* gene that codes for the uracil phosphoribosyl transferase. These enzymes catalyze the transformation of 5-fluorocytosine to 5-fluorouracil, and 5-fluorouracil to 5- fluorouridine monophosphate, respectively. The most often acquired resistance to flucytosine resides in the point mutations in the *FUR1* gene. Several point mutations have been described in *C. albicans, C. glabrata,* and *C. lusitaniae* (Vandeputte *et al.,* 2011).

Mutations that lead to a reduction or termination in the drug's import, or its intracellular conversion are often accountable for resistance to pyrimidine analogs. Point mutation that results in the

substitution of arginine with cysteine at the 101st position of *FUR1* has been discovered to be associated with resistance to 5-FC in *C. albicans* (Vermes *et al.,* 2000). Also, mutation T26C, which results in the amino acid change M19T of the *FCY1* gene coding for cytosine deaminase, and responsible for the conversion of 5-FC to 5-fluorouridine monophosphate, has been found to be associated with a *C. lusitane* isolate resistance to 5-FC (Vermes *et al.,* 2000).

1.4.4 Azoles: The Inhibitors of the Lanosterol 14-α-Demethylase

Currently, the broadest group of antifungal drugs in clinical use is the azole family, which disrupt the cell membrane by inhibiting the activity of the enzyme lanosterol 14-α-demethylase (Pfaller *et al.,* 2014), involved in the biosynthesis of ergosterol (Figure 1). Ergosterol (in fungi), a compound analogous to cholesterol (in animal cells), is the largest sterol constituent of the fungal cell membrane. Ergosterol and cholesterol have sufficient structural differences, and most antifungal agents targeted to ergosterol binding or biosynthesis do not cross-react with host cells. The azole family includes the imidazoles (econazole, miconazole, ketoconazole and clotrimazole), the triazoles (fluconazole, itraconazole, and the latest agents; voriconazole and posaconazole) (Hof, 2006). Voriconazole is a second-generation, synthetic triazole derivative of fluconazole, while posaconazole is a hydroxylated analogue of itraconazole. Many azoles are effective both for topical use as well as for the treatment and prophylaxis of invasive fungal infections (Ribas, 2016). These agents have been approved by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMEA) for clinical use. The azoles have a broad spectrum of action, and are the most commonly used class of drugs for the treatment of diseases caused by fungi, particularly *Candida* spp. In the ergosterol biosynthetic pathway, acetic acid is converted into ergosterol using a number of enzymes, similar to the processes involved in the biosynthesis of cholesterol in mammals. Fluconazole, itraconazole, voriconazole and posaconazole interfere with a step in the ergosterol biosynthetic pathway through the inhibition of the enzyme lanosterol 14-α-demethylase (*ERG11*p or *14DM*). The conversion of lanosterol into ergosterol is therefore prevented, thereby increasing both the permeability and the progressive instability of the fungal cell (Vandeputte *et al.,* 2005; Dos-Santos-Silva *et al.,* 2016). Fluconazole and itraconazole have an extensive use for chemoprophylaxis and treatment of systemic fungal infections for up to a decade now due to their favorable oral bioavailability and safety profiles (Livermore, 2004). However, fluconazole clinical resistance has been shown to rise in patients with vaginal

candidiasis and candidemia (Redding *et al.*, 2003; Skiest *et al.*, 2007). Intrinsically decreased susceptibility to fluconazole has been immensely reported for non-*albicans* species like *C. krusei*, *C. lusitaniae* and *C. glabrata* (Vazquez *et al.*, 2001; Pfaller *et al.*, 2015), and there are numerous molecular mechanisms that could lead to azole resistance (No¨el, 2012)

i. Decreased Intracellular Drug Accumulation. This mechanism for reducing the intracellular concentration of azole depends on an upregulation of two principal families of efflux pumps; *Candida* drug resistance 1 and 2 proteins (*CDR1*p and *CDR2*p). There is another pump belonging to the major facilitator superfamily (MFS) of transporters and is encoded by the *MDR1* gene in *C. albicans*. It is a secondary transporter that uses proton gradient as a source of energy, and specific for fluconazole (Coste *et al.*, 2006; Cannon *et al.*, 2009; Spampinato and Leonardi, 2013).
ii. Reduced affinity for the drug. The target of azole antifungals is the lanosterol 14-α-demethylase encoded by the *ERG11* gene. Several point mutations have been characterized and associated with *ERG11* gene, believed to be a mechanism to prevent the action of the azole.
iii. Counteraction of the Drug Effect. Two molecular tools contribute to the counterbalancing of the drug effects. The first involves an upregulation of the *ERG11* gene, leading to an intracellular increase of the target protein (*ERG11*p) (No¨el, 2009). The second machinery, involves the modification of the late steps of the biosynthesis of ergosterol via *ERG3* inactivation, which leads to the total inactivation of the C5 sterol desaturase. Consequently, toxic 14-*α*-methylated sterols are no longer accumulated, and the yeast produce cell membranes lacking ergosterol, though other sterols may be present. This second mechanism, though very infrequent, has been recognized in several clinical isolates of *C. albicans* (Martel *et al.*, 2010).

1.5 Molecular Machineries Responsible for *Candida* spp Resistance to Azoles

Antifungal resistance could be based on different molecular mechanisms, namely:
(i) Decreased drug intracellular accumulation
(ii) Reduced affinity for the drug, and

(iii) Counteraction of the drug effect.

The mechanism of resistance is dependent on the mode of action of the antifungal compounds. Numerous studies have been carried out to elucidate the molecular machineries accountable for the onset of the resistance of clinical isolates of *Candida* spp to azoles. Deficiency in DNA repair may account for the accelerated development of various genetic variations accountable for the resistance. The contribution of DNA repair in fungal pathogens, especially to the emergence of antifungal resistance, is yet to be discovered in depth (Healey *et al.,* 2016). Healey *et al*. (2016) established that a mutator phenotype resulting from a mismatch repair defect is prevalent in clinical isolates of *C. glabrata*. This mechanism enhances the gain of resistance to multiple antifungal drugs, which to some extent explains the elevated rates of azole and multi-drug resistance associated with *Candida*. One such mechanism involves the efflux pumps that export the antifungal agent from the intracellular environment to the extracellular environment, in order to reduce the intracellular concentration of drugs (Holmes *et al.,* 2006). Studies on *Candida* with fluconazole have demonstrated that this drug is actively transported into the extracellular environment by fungal cells, via an energy-dependent manner, and the overexpression of genes that encode membrane transport proteins is responsible for most of the antifungal efflux (Holmes *et al.,* 2006). The two notable families of efflux membrane transporters distinguished in yeasts according to the energy source used for this extrusion of substrates are the *CDR1* and *CDR2* (*Candida* drug resistance 1 and 2) proteins (Marie and White, 2009). The genes *CDR1* and *CDR2* encode the ABC-type transport proteins which act as transmembrane efflux pumps, using the hydrolysis of ATP to transport substrates across the membrane. The expression of *CDR1* and *CDR2* is regulated by *TAC1* (Transcriptional activator of *CDR* genes 1). Hyperactivation of the *TAC1* transcription factor is manifest in the gain-of-function mutations that consequently promote the overexpression of *CDR1* and *CDR2* (Coste *et al.,* 2004). In addition to *CDR1* and *CDR2*, the *MDR1* gene (Multi-drug resistance 1; which encodes a permease protein of the *MFS* (Major facilitator superfamily) is directly involved in fluconazole resistance. This gene acts as a membrane carrier using a proton electrochemical gradient for the transport of substrates. The expression of *MDR1* is controlled by at least three transcription factors, including the *MRR1* transcription factor (multi-drug resistance regulator 1). The hyperactivation of *MRR1* leads to the overexpression of *MDR1* (Schubert *et al.,* 2008). Changes associated with ergosterol biosynthesis covers another mechanism of azole resistance. The principle of this mechanism is that it

circumvents the inactivation of the sterol enzyme Δ5, 6-desaturase which is encoded by the gene *ERG3*. It acts upstream of 14-α-demethylase in the ergosterol biosynthesis pathway, converting 14-α-metilfecoesterol into 14-α-methyl-3, 6-diol (Chau *et al.,* 2005). Thus, isolates with such modifications in this enzymatic step will definitely show a selective advantage when subjected to azoles (Akins, 2005). Some other transporter genes (*CgCDR1*, *CgCDR2* and *CgSNQ2*), have been reported to be upregulated in azole-resistant *C. glabrata* (Torelli *et al.,* 2008). *CgCDR2* was formerly named *PDH1*, and *CgSNQ2* is another ABC transporter. This observation has been extended to other species; *C. dubliniensis* (*CdCDR1* and *CdCDR2*), and *C. tropicalis* (*CDR1*-homologue) isolates (Lamping *et al.,* 2009; Paul and Moye-Rowley, 2014). In *C. glabrata*, *CgCDR1*, *CgCDR2*, and *CgSNQ2* genes are all regulated by the *CgPDR1* transcription factor (Vermitsky *et al.,* 2006). Mutations in the *ERG11* gene of *Candida* spp. are predominantly also involved in resistance of *Candida* to azoles. *ERG11* is positioned on chromosome 5 and shows significant variation in size, from 1,569 bp to 2,669 bp, depending on the species and strain. It could also me more or less, depending on the prevailing situations that are particular to the particular strain. Mutations in *ERG11* are largely responsible for resistance to azoles, manifested through the reduction in the binding affinity of the drug to the target enzyme (Dos-Santos-Silva *et al.,* 2016). Series of studies have tried comparing *ERG11*gene sequences from isolates of different *Candida* spp. that are susceptible as well as resistant to azoles, with *C. albicans* being the most studied species in this respect (De-Almeida *et al.,* 2015). Vandeputte *et al*. (2005) studied isolates of *C. tropicalis* that were resistant to fluconazole, and discovered a missense mutation (Y132F) responsible for resistance previously reported in *C. albicans* by Chau *et al*. (2004). De-Almeida *et al*. (2015) also investigated *ERG11* mutations in clinical isolates of *C. albicans*, *C. glabrata* and *C. tropicalis* that had been previously evaluated by fluconazole-susceptibility tests, and identified 14 different missense mutations, five of which were not previously described, including a new L321F mutation identified in an isolate of *C. albicans*. Dos-Santos-Silva *et al*. (2016) identified three new synonymous mutations in the *ERG11* gene in the isolates of *C. glabrata* (C108G, C423T and A1581G) in addition to two new nonsynonymous mutations in the isolates of *C. krusei*: -A497C (Y166S) and G1570A (G524R). An elevated *ERG11* gene expression could also results in resistance to antifungal agents due to an elevated concentration of 14-α-demethylase in the intracellular environment, requiring larger amounts of antifungals to inhibit the enzyme activity. This mechanism has been revealed in various isolates of *C. albicans* resistant to fluconazole as

studied by Xu *et al.* (2015). *ERG11* gene expression is regulated by the *UPC2* (uptake control 2) transcription factor. Gain-of-function mutations in the *UPC2* gene leads to hyperactivity. Such hyperactivity leads to the over-activation of *ERG11* gene expression. Over-expression of *ERG11*, in turn, significantly reduces the effect of the antifungal agent in cells, thereby decreasing the cell's sensitivity and susceptibility to the drug (Heilmann *et al.*, 2010).

1.6 Species-Specific Azole Antifungal Resistance Machineries

1.6.1 Azole Antifungal Resistance Machineries in *Candida albicans*

Resistance to azole antifungals in *Candida albicans* has been the most extensively studied. One amongst the many possible mechanisms of resistance identified in this species is the occurrence of point mutations in *ERG11*. Previous studies have recognized amino acid substitutions to account for the decreased fluconazole susceptibility. In a recent study in which 63 fluconazole-resistant *C. albicans* clinical isolates were examined for mutations within their *ERG11* alleles, 55 were found to carry at least one mutation that resulted in amino acid substitutions, with 9 of such predicted amino acid substitutions found to be novel (Flowers *et al.*, 2015). Molecular modeling of the substitutions that resulted in decreased fluconazole susceptibility showed that the mutations clustered in either the predicted catalytic site, the fungus-specific external loop, or on the proximal surface potentially interacting with the loop or near the heme. Furthermore, a study involving site-directed mutagenesis of wild-type *ERG11* in 23 *C. albicans* clinical isolates, showed nine of these mutations to result in increased fluconazole resistance (Xiang *et al.*, 2013). Five (5) of the amino acid substitutions were predicted to be at or near the active site of the *ERG11*p. Another notable mechanism of fluconazole resistance in *C. albicans* is the increased expression of *ERG11* gene as a result of the mutations in the gene encoding the zinc-cluster transcriptional regulator *UPC2*p. Initially, *UPC2* involvement in fluconazole resistance in *C. albicans* was demonstrated when *C. albicans* strains were shown to be highly susceptible to azoles as those over-expressing *UPC2* had increased fluconazole resistance (MacPherson *et al.*, 2005). Subsequent studies also noted a matched set of fluconazole-susceptible and resistant *C. albicans* clinical isolates in which fluconazole resistance was not linked with overexpression of drug efflux pumps (Dunkel *et al.*, 2008; Heilmann *et al.*, 2010; Hoot *et al.*, 2011). Three additional matched sets of *ERG11*-overexpressing clinical *C. albicans* isolates have been described with no sequence differences in *UPC2* between the susceptible and resistant isolates, indicating that other possible machineries of

ERG11 upregulation exist. A study involving 63 fluconazole-resistant *C. albicans* clinical isolates showed 47 of these isolates to overexpress *ERG11* by at least 2-fold (Flowers *et al.*, 2015). Twenty-nine (29) of these *ERG11*-overexpressing isolates had a missense mutation in *UPC2*, and eight (8) single amino acid substitutions elucidated from their *UPC2* alleles. Seven (7) of these alleles were also found to be associated with increased *ERG11* expression, increased ergosterol production, and decreased fluconazole susceptibility. Two other established machineries of fluconazole resistance in *C. albicans* involve the overexpression of drug efflux pumps *MDR1*p and *CDR1*p/*Cdr2*p. *TAC1* (transcriptional activator of *CDR1* genes) is a zinc-cluster transcription factor whose regulon is hallmarked by the ATP-binding cassette (ABC) transporter-encoding genes *CDR1* and *CDR2* (Coste *et al.,* 2004). Activating the expression of *TAC1* regulon is via the binding of *TAC1* to the DRE (drug response element) present in the promoter regions of *TAC1*-regulated genes (Coste *et al.,* 2004; Liu *et al.,* 2007; Coste *et al.,* 2007; Selmecki *et al.*, 2008). *MRR1*, multidrug resistance regulator 1, is likewise, identified by comparing the transcriptomes of a set of matched isolates in which the fluconazole-resistant isolates overexpressed *MDR1*. Disruption of *MRR1* in these resistant isolates leads to a decrease in fluconazole MIC, while introduction of each of the mutant alleles separately into a wildtype fluconazole-susceptible background in the native *MRR1* locus confers fluconazole resistance to constructed strain (Dunkel *et al.,* 2008). A less common mechanism of azole resistance in *C. albicans* is the inactivation of the *ERG3* gene, which encodes the ergosterol biosynthesis enzyme sterol Δ5, 6 desaturase. The *ERG3*p catalyzes one of the final steps in the pathway and also converts toxic 14-α-methylated sterol intermediates into the non-toxic sterol 14-α-methylergosta-8, 24(28)-dien-3β,6α-diol, which therefore, averts such toxic sterols from being synthesized (Chau *et al.*, 2005; Martel *et al.*, 2010). Another phenomenon that plays a role in azole resistance in *C. albicans* as demonstrated by genome hybridization is aneuploidy, involving Chr5 (Selmecki *et al.,* 2006). Similarly, loss of heterozygosity (LOH) has also been shown to occur in azole-resistant *C. albicans* (Coste *et al.*, 2006). Investigation of *TAC1* in a matched set of azole-susceptible and resistant *C. albicans* isolates showed that the susceptible isolate harbored two wildtype alleles of *TAC1*, while the resistant isolate contained only one of those alleles (Coste *et al.*, 2004).

1.6.2 Azole Antifungal Resistance Machineries in *Candida parapsilosis*

Since azole resistance has been extensively studied in *C. albicans*, efforts to reveal machineries of azole resistance in *C. parapsilosis* have involved examining orthologous genes, and has yielded

mixed results. For instance, a study of a sequence of six isolates from a particular patient found that single SNP in *MRR1* exists in the two fluconazole-resistant isolates (Zhang *et al.*, 2015; Souza *et al.*, 2015). In addition to the established association between *CDR1* and *ERG11* as well as fluconazole resistance in *C. parapsilosis*, this causal link has not been definitively proven. In previous efforts to identify potential machineries of azole resistance on a genome-wide scale in *C. parapsilosis*, fluconazole-, voriconazole-, and posaconazole-resistant strains were developed experimentally by serial passage in liquid culture containing either of these drugs (Silva *et al.*, 2009). It was observed that the fluconazole- and voriconazole-resistant strains were cross-resistant to both fluconazole and voriconazole, and possessed similar transcriptional profiles (Branco *et al.*, 2015). It is also remarkable to note that *ERG11* was not differentially expressed in these strains (Branco *et al.*, 2015). Among the genes that were differentially expressed in fluconazole- and voriconazole-resistant *C. parapsilosis* strains, the stress response gene *GRP2*, *MDR1* and *MRR1* are common (Branco *et al.*, 2015). Past studies involving laboratory strains of *C. parapsilosis*, in which previously determined gain-of-function alleles of *CpMRR1* were introduced into the native locus of the strains containing *MRR1*p with a G583R amino acid substitution from a fluconazole-resistant *C. parapsilosis* isolate, led to greater resistance to fluconazole and voriconazole when compared to strains harboring the wildtype allele (Branco *et al.*, 2015). Likewise, studies have revealed that while differential expression of efflux pumps is commonly found in azole-resistant *C. parapsilosis* isolates, the resistant phenotype is not exclusively due to their overexpression, but instead multifactorial, and involves *ERG11* mutation and/or overexpression as well (Berkow *et al.*, 2015).

1.6.3 Azole Antifungal Resistance Machineries in *Candida tropicalis*

A relatively little information is known about the machineries of azole resistance in *C. tropicalis*. The average *ERG11* expression level is shown to be 4-fold higher among fluconazole-resistant isolates than susceptible isolates (Jiang *et al.*, 2012). Furthermore, *ERG11* expression is even higher among a subset of fluconazole-resistant isolates that are also resistant to itraconazole and voriconazole (Choi *et al.*, 2016). The sequence of the *UPC2* gene of *C. tropicalis* had revealed numerous heterozygous and homozygous mutations in resistant strains (MacPherson *et al*., 2005). Many of these mutations however, have been observed in fluconazole-susceptible isolates that are not found to overexpress *ERG11*, therefore, characterization of their impact on the regulatory function of *UPC2* is required. Molecular characterization of azole-resistant clinical *C. tropicalis*

isolates has also showed alterations in the ergosterol biosynthetic pathway, similar to that found in the *C. albicans* (Vandeputte *et al.*, 2005; Eddouzi *et al.*, 2013; Jiang *et al.*, 2012; Choi *et al.*, 2016). Some notable *ERG11* mutations consist of a deletion of 132 nucleotides resulting in a D275V amino acid substitution and the loss of 44 amino acids near the N-terminus of the *ERG11*p (Eddouzi *et al.*, 2013). In addition, *ERG11* mutation resulting in decreased fluconazole susceptibility due to the amino acid substitution Y132F that has been well characterized in *C. albicans* has also been recently observed in a fluconazole-resistant *C. tropicalis* isolate (Barchiesi *et al.*, 2000). It has also been previously observed that all isolates with reduced susceptibility to fluconazole demonstrated increased expression of both *C. tropicalis MDR1* and a gene with high homology to *C. albicans CDR1*. The role of efflux pump overexpression in azole resistance among clinical *C. tropicalis* isolates has not been so clearly defined (Jiang *et al.*, 2012). Meanwhile, expression of both *MDR1* and *CDR1* has been observed to be significantly higher among both fluconazole-susceptible and fluconazole-resistant isolates (Wang *et al.*, 2015).

1.6.4 Azole Antifungal Resistance Machineries in *Candida krusei*

Though *C. krusei* is intrinsically resistant to fluconazole, the precise mechanism is yet to be completely understood. Several studies have attributed *C. krusei*'s innate azole resistance to its efflux pump activity, through the ATP-binding cassette transporter *ABC1*p, leading to reduced drug accumulation, in combination with reduced azole affinity for *ERG11*p (Lamping *et al.*, 2009). Alterations in the cell membrane affecting the membrane fluidity possibly may also be implicated in azole resistance, since there is evidence to suggest that intracellular azole accumulation occurs through one or possibly both machineries of passive and facilitated diffusion (Kolaczkowska and Kolaczkowski, 2016). The discovery of a trisomy in the *ERG11*-containing chromosome in a *C. krusei* strain suggests that aneuploidy possibly may also play a part in the behaviors of this species, though the effects as it relates to azole resistance are not yet fully uncovered (Lamping *et al.*, 2009). Analysis of itraconazole-resistant *C. krusei* isolates showed that reduced intracellular content of the drug (and not altered affinity for the drug target) possibly drives resistance to itraconazole (Tavakoli *et al.*, 2010). More recently, it has been recommended that overexpression of genes coding for *ERG11*p and the efflux pump *ABC2*p may possibly play a role with itraconazole resistance (Tavakoli *et al.*, 2010). Resistance to voriconazole has also appeared, and research supports a theory where overexpression of the genes encoding *ABC2* (the efflux pump) and *ERG11* impart more transient resistance properties. Meanwhile, increased expression of

*ABC1*p and point mutations in *ERG11* prevail as time progresses to yield a stably resistant pathogen in the prolonged presence of voriconazole. *ERG11*p amino acid substitutions have also been shown in azole-resistant *C. krusei* (Dos-Santos-Silva *et al.,* 2016). Newer antifungal agents such as posaconazole and isavuconazole have shown good activity against *C. krusei* (Rybak *et al.,* 2015), and reports of resistance against these agents are still relatively sparse (Pfaller *et al.,* 2014; Pfaller *et al.,* 2015). The machineries of resistance in *C. krusei* against these agents remains a strong area to be investigated.

1.6.5 Azole Antifungal Resistance Machineries in *Candida glabrata*

C. glabrata is a unique species among the *Candida* species outlined so far as it is a haploid yeast. The onset of azole resistance in clinical isolates of *C. glabrata* has been particularly linked to the incidence of mutations in the zinc cluster transcription factor *PDR1* (Vermitsky and Edlind, 2004), leading to differential expression of the downstream targets. Nearly all clinical isolates examined in previous studies have been found to have *PDR1* mutations, with such mutations located in the inhibitory domain, middle homology region, activating domain, and the xenobiotic binding region (Healey *et al.,* 2016). The activating mutations display discrete expression patterns of the downstream effector genes, with the exclusion of increased expression of *CDR1* and *PUP1*, and no relationship has been establish between location of the mutation and altered gene expression (Caudle *et al.,* 2011). Amid the genes whose pleiotropic drug response element (PDRE) is directly bound by *PDR1* (Paul and Moye-Rowley, 2014), only three; the ABC transporters *CDR1*, *PDH1* (*CDR2*), and *SNQ2* (Torelli *et al.,* 2008), have been connected directly to azole resistance. Contemporary work has shown increased expression of four MFS transporters in clotrimazole resistant isolates compared to clotrimazole susceptible clinical isolates. Interruption of one of these, TPO3, discreetly increased susceptibility to clotrimazole and fluconazole (Coste et al., 2006; Cannon et al., 2009; Spampinato and Leonardi, 2013). This result recommends that MFS transporters possibly will have minor roles in azole resistance in *C. glabrata*. Remarkably, *ERG11* does not seem to play an important role in clinical azole resistance in *C. glabrata* (Sanguinetti *et al.,* 2005), as opposed to the findings in the other earlier discussed species. However, according to a finding, exceptional elevated expression of *ERG11* has been detected in only two clinical isolates of *C. glabrata* (Redding *et al.,* 2003). The upregulation in one isolate was observed to be as a result of the duplication of the entire chromosome containing *ERG11*, and the phenotype was lost with subsequent passaging in azole-free media. A particular resistant clinical isolate of *C. glabrata* has

demonstrated a total absence of nonfunctional 14-α-sterol demethylase as a result of a missense mutation in *ERG11*, leading to the complete absence of ergosterol in the cell membrane (Hull *et al.*, 2012). *C. glabrata* possesses the capacity of growing with altered cell membrane sterols, which permits for the evasion of azole treatment. *C. glabrata* is able to take up exogenous sterols (Nakayama *et al.*, 2000), when the ergosterol biosynthesis pathway is obstructed (Bard *et al.*, 2005). *AUS1*p has also been recognized as the sterol transporter responsible for tolerance to azoles in the presence of exogenous sterols (Nakayama *et al.*, 2007). Comparatively, *C. albicans* has lately been shown to take up sterols under aerobic conditions while, *C. glabrata* is more substantial in its capacity to take up sterols and does so in both aerobic and anaerobic conditions. In the presence of serum and fluconazole, it improves on the uptake under aerobic conditions (Zavrel *et al.*, 2013). Azole resistance in *C. glabrata* has also been ascribed to the development of petite mutants (cells that have lost mitochondrial function, resulting in respiratory deficiency) (Brun *et al.*, 2003). Petite mutants can be produced in the laboratory by treatment with azoles or ethidium bromide (Ferrari *et al.*, 2011), but not common among clinical isolates. Azole resistance in petite mutants has been credited to the upregulation of the ABC transporters *CDR1*, *CDR2*, and *SNQ2* (Ferrari *et al.*, 2011), exhibiting altered sterol profiles with an inconsistent quantity of ergosterol and very little of ergosterol intermediates; however, no changes in the sequence of *ERG11* or its expression have been detected so far (Brun *et al.*, 2003).

1.7 Justification for Study

The resistance of *Candida* species to drugs has become a major problem in clinical practice as many *Candida* species are becoming increasingly resistant to first-line and second-line antifungal medications. One of the most potent class of drugs against *Candida* species is a group of compounds called the azoles, of which fluconazole is particularly proven to be outstanding in treating *Candida* infections. Unfortunately, many *Candida* species have devised a way of subverting the potency of this drug. A major mechanism responsible for this resistance is the mutation in the DNA nucleotide sequence of the gene that codes for the protein (enzyme), *ERG11*p; the target of fluconazole. The mutation alters the final structure and conformation of the enzyme lanosterol 14-α-demethylase, to prevent the binding of fluconazole, thereby, subverting the inhibitory action of the drug. Profiling the DNA changes in the *ERG11* gene, alongside modelling studies, will provide more understanding to the molecular mechanism of this resistance, and the knowledge gained will provide more insight in drug design

1.8 Aim of the Study

1.8.1 Aim of the Study

The aim of this study was to investigate the distribution of different species of *Candida* in human and dog vaginal swabs, the susceptibility of the *Candida* species to fluconazole, a first-line antifungal drug, the profile of the *ERG11* gene, and the phylogenetic relationship of the *Candida* species based on the nucleotide sequence of the *ERG11* gene.

1.8.2 Specific Objectives of the Study

The aim was achieved through the following specific objectives:

i. Screen collected samples for *Candida* species using conventional screening methods
ii. Identify *Candida* species using differential/selective media (Chromogenic Agar)
iii. Carry-out antifungal (fluconazole) susceptibility studies on the *Candida* species isolated and identified
iv. Profile the *ERG11* gene of the different *Candida* isolates of interest, by direct gene sequencing

v. Carry-out a genetic sequence comparison of the *ERG11* gene across nucleotide databases for nucleotide homology and variations, using the nucleotide sequence of the *Candida* species of interest

vi. Carry-out protein modelling studies on the *ERG11* gene sequence to ascertain the relationship between the biological roles of the predicted proteins and the structural basis for drug design.

vii. Investigate the phylogenetic (evolutionary) relationship between the different *Candida* species using the *ERG11* gene sequence

CHAPTER TWO
MATERIALS AND METHODS

2.1 Materials

2.1.1 *Candida* Species

Different species of *Candida* were isolated from samples collected from both humans and animals.

2.1.2 Human Samples

Collection of human samples was approved by the authority of the Medical Center, University of Nigeria, Nsukka from women of reproductive age, between the ages of 20 and 35 years, who provided informed consent, and confirmed by a qualified physician to show signs and symptoms of vulvovaginal candidosis.

2.1.3 Animal Samples

Collection of samples from female dogs of reproductive age was approved by the authority of the Veterinary Teaching Hospital, University of Nigeria, Nsukka.

2.1.4 Chemicals/Reagents

Chemicals/Reagents	Manufacturer
Sabouraud Dextrose Agar (SDA)	Sigma Aldrich, Germany
Mueller-Hinton agar	Sigma Aldrich, Germany
Fungal/Bacterial DNA Miniprep Kit (Catalog No. D6005)	ZYMO RESEARCH, USA
Beta-mercaptoethanol	Sigma Aldrich, Germany
ABI V3.1 Big dye gene sequencing kit	Thermo Fisher Scientific
ExoSAP PCR kit	Thermo Fisher Scientific

2.1.5 Equipment/Instruments

Equipment	Manufacturer
Glass wares	Pyrex, England
Autoclaved	Rodwell, Germany

Light microscope	Olympus, China
Water bath	Chikkpas, England
Refrigeration	Thermocool, Germany
Micro-centrifuge	Labaclence, England
Thermocycler	Epperndorf, Germany
ABI3500XL analysers	Thermo Fisher Scientific

2.2 Methods

2.2.1 Experimental Design

High Vaginal Swabs were collected from both women and dogs of reproductive age, and screened for *Candida* species. The isolates were further characterized using a differential medium (chromogenic agar). The isolates were subjected to susceptibility testing with 25 μg fluconazole. The ERG11 gene of the isolates of interest from the susceptibility testing were sequence for possible nucleotide variations. Sequences obtained were analyzed using a series of bioinformatics tools, for sequence homology, predicted protein characteristics, and phylogenetic (evolutionary) relationship between the *Candida* species based on the *ERG11* gene sequence.

2.2.2 Collection of Samples

Human samples were collected by licensed laboratory technologists in the Medical Centre, University of Nigeria, by directly inserting a swab collection stick into the vagina of the women in the direction of the cervix. Similarly, samples were collected from female dogs by qualified veterinary doctors in the veterinary teaching hospital, University of Nigeria, by directly inserting a specially tilted swab collection stick into the vagina in the direction of the cervix.

2.2.3 Identification and Characterization of *Candida* Species

The collected samples were screened for yeast (*Candida* species) using conventional screening methods that include growth on Sabouraud Dextrose Agar (SDA) and microscopic examinations for morphology. The yeast isolates were then characterized for specific *Candida* species using differential/selective (Chromogenic Agar) media.

2.2.3.1 Conventional Screening on Nutrient Media

The culture medium used in isolating the clinical *Candida* species was Sabouraud Dextrose Agar (SDA) (40 g/l dextrose, 10 g/l peptone, 20 g/l agar, pH 5.6). SDA was prepared according to the manufacturer's instructions. An amount, 65 g of the SDA was dispensed in one litre of purified (distilled) water. Heat with frequent agitation was applied and boiled for one minute to completely dissolve the medium. It was then autoclaved at 121° C for 15 minutes. The medium was thereafter cooled to 50°C, and poured into petri dishes for culture, and Bijou bottles for slants. For processing of specimen, samples were inoculated on the SDA-prepared Petri dishes by direct streaking, using inoculating loop. Plates were incubated at 30°C in an inverted position (agar side up), and monitored for growth for 24 – 48 hours.

2.2.3.2 Identification of Isolates using Microscopy

Isolates were phenotypically identified based on their microscopic and macroscopic characteristics as described by classical methodologies, including colony formation, and production of germinative tubes (Almeida *et al.,* 2013). Identification of *Candida* isolates was done using the germ tube test. The clinical samples were incubated in normal saline for 3 hours at 37°C. Samples were observed on slides under a light microscope and viewed for the, short, slender, tube-like structures (germ tube). Presence of germ tube indicated the presence of *Candida* spp.

2.2.3.3 Characterization of Isolates using Chromogenic Media

Isolated yeasts (*Candida spp.*) were grown on Chromogenic *Candida* Agar (CHROMagar®; Paris, France). The identity of all the isolates were confirmed at the species level using the CHROMagar *Candida*®. One liter of the CHROMagar *Candida*® was prepared according to the manufacturer's manual. To prepare 1L of the media, 47.7 g of the agar powder was slowly dispersed in 1L of purified (distilled) water and stirred until the agar was well thickened. It was then heated and brought to boiling (100°C) while swirling (stirring) regularly. The media was cooled in a water bath to 50°C, while swirled gently. The media was poured into sterile petri dishes. It was left to solidify and dry. The dishes were protected from light and dehydration by storing in the dark under refrigeration (2/8°C) before use. For inoculation, the agar plates were brought to a temperature of 25^0C before inoculation, by streaking the sample onto the plates. The plates were incubated in aerobic conditions at 37°C for 48 hours.

2.2.4 Antifungal Drug Susceptibility Testing

The susceptibility of the *Candida* isolates to the selected antifungal agent (fluconazole) was carried out on culture plates, following the methods of the Clinical and Laboratory Standards Institute (CLSI) document M27-A316 (Fothergill, 2012). Suspension of inoculums was prepared in 5 ml of sterile saline (0.85%) and the turbidity adjusted to 0.5 McFarland standards. Within 15 minutes of adjusting the turbidity, each isolate was inoculated onto a dried surface of prepared Mueller-Hinton agar plates using sterile cotton swabs. Antimicrobial disks containing 25 µg of fluconazole were dispensed onto the surface of the inoculated agar plate. Each disk was pressed down to ensure its complete contact with the agar surface. The plates were incubated at 37ºC and examined after 24 hours and 48 hours of incubation. The zones of inhibition were measured in millimeters and the results were interpreted using validated CLSI interpretive breakpoints for *in vitro* susceptibility testing of fluconazole. *Candida* species were reported as susceptible "S" (zone diameter ≥ 19 mm); susceptible dose dependent "SDD" (15 to 18 mm) and resistant "R" (≤ 14 mm).

2.2.5 Isolation of DNA

ZYMO RESEARCH Fungal/Bacterial DNA Miniprep Kit with Catalog No. D6005 was used for the isolation of the *Candida* DNA. Four (4) isolates, comprising three (3) human isolates from the susceptibility testing classifications (Resistant, SDD and Susceptible) and one (1) dog isolate were the isolates of interest. For each of the 4 isolates of interest, beta-mercaptoethanol (user supplied) was added to the Genomic Lysis Buffer to a final dilution of 500 µl per 100 ml. Each of the *Candida* samples of interest was resuspended in 200 µl of water. An amount, 100 mg (wet weight) of the *Candida* cells was added into a ZR BashingBead™ Lysis Tubes. Thereafter, 750 µl of Lysis Solution was added. This was secured in a bead beater fitted with a 2 ml tube holder assembly and processed at 10,000 x *G* for 6 minutes. The ZR BashingBead™ Lysis Tube was centrifuged at 10,000 x *G* for 1 minute. In the following step, 400 µl of the supernatant was transfer to a Zymo-Spin™ IV Spin Filter in a collection tube and centrifuged at 7,000 x *G* for 1 minute. In a follow-up step, 1,200 µl of Genomic Lysis Buffer was then added to the filtrate in the collection tube. In the next step, 800 µl of the mixture was then transferred to a Zymo-Spin™ IIC Column in a collection tube and centrifuged at 10,000 x *g* for 1 minute. The flow-through (effluent) from the collection tube was discarded, and the step repeated. Afterwards, 200 µl of DNA Pre-Wash Buffer was then added to the Zymo-Spin™ IIC Column in a new collection tube and centrifuged at 10,000 x *g* for 1 minute. In a further step, 500 µl of g-DNA Wash Buffer was then added to the Zymo-

Spin™ IIC Column and centrifuged at 10,000 x *g* for 1 minute. The Zymo-Spin™ IIC Column was transferred to a clean 1.5 ml micro-centrifuge tube and 100 µl DNA Elution Buffer added directly to the column matrix. It was centrifuged at 10,000 x *G* for 30 seconds to elute the DNA. The ultra-pure DNA was stored at -20^{0}C, ready for further use.

2.2.6 PCR Amplification

The reaction solution consisted of 2.5 ml of PCR reaction buffer (50 mM KCl, 1.5 mM MgCl2, 10 mM Tris-HCl (pH 8.3), 10 ng of target genomic DNA, 200 mM of dNTPs, 1 mM of each primer, 0.5 U Taq polymerase, and sterilized distilled water to a final volume of 25 ml. PCR was carried out by using primers that span the *ERG11* open reading frame: 5'-GTT GAA ACT GTC ATT GAT GG (forward) and 5'-TCA GAA CAC TGA ATC GAA AG (reverse). The PCR was performed in a 25 well thermocycler (Epperndorf, Germany). The PCR products were resolved using 2% agarose gel electrophoresis to assess their quality and integrity. The amplification program for all reactions was as follows: initial denaturation at 94ºC for 5 min, 30 denaturation cycles at 94ºC for 30 s, annealing at 50ºC for 40 s, extension at 72ºC for 50 s, followed by final extension at 72ºC for 10 min.

2.2.7 Sequencing of the *ERG11* gene

PCR products were cleaned using ExoSAP in accordance with the following protocol. The Exo/SAP master mix was prepared by adding 50.0 µl Exonuclease I (NEB M0293) 20U/ul and 200.0 µl Shrimp Alkaline Phosphatase (NEB M0371) 1U/ul to a 0.6 ml micro-centrifuge tube. This was followed by mixing 10.0 µl of PCR Mixture and 2.5 µl Exo/SAP Mix. It was thereafter mixed well and incubated at 37°C for 30 min. The reaction was stopped by heating the mixture at 95°C for 5 min. Gene sequencing was done with the ABI V3.1 Big dye kit following the instructions stipulated in the manufacturer's manual. The labelled products were cleaned with the Zymo Seq clean-up kit in accordance with the manufacturer's instruction. Thereafter, the cleaned products were injected on the ABI3500XL analysers with a 50cm array, using POP7.

2.2.8 Bioinformatics Studies

The sequence analysis of the *ERG11* gene was carried out using the nucleotide sequence of the *Candida* isolates of interest. The phylogenetic (evolutionary) relationship between the *Candida* isolates was determined using the *ERG11* gene and other related gene families. The sequences

were assembled and aligned using the sequences obtained, and compared to *ERG11* sequences in the nucleotide databases using Basic Local Alignment Tool (BLASTn) accessible through the National Center for Biotechnology Information (NCBI) server. Genetic comparison of the *Candida ERG11* Gene sequences was carried out using the NCBI-BLAST alignment tool (Arnaud *et al.,* 2006). A comprehensive comparison of the evolutionary (phylogenetic) relationship of the *ERG11* Gene sequences of the *Candida* species was done by retrieving sequences of high and topmost homology at the NCBI server (Kumar *et al.,* 2016). Protein modelling studies on the *ERG11* gene sequence, to comprehend the relationship between the biological roles of the predicted protein, as well as the structural basis for drug design was also carried out. Proteomic analysis focused on obtaining the amino acids sequences (Primary protein structure), Protein secondary, tertiary and quaternary structure of lanosterol 14-α- demethylase. Analyses such as antigenic sites and ligand binding sites, for more efficient drug design were also carried out.

2.7.1 Multiple Sequence Alignment

The genetic comparison of the *ERG11* gene of *Candida* species was done using Multiple Sequence Alignment (MSA) for matching the nucleotide and predicted protein sequences, using the NCBI-BLAST alignment tool (Arnaud *et al.,* 2006).

2.7.2 Translation

The gene sequences were translated using the Biopython library (Stafford, 2014). Biopython includes modules for reading and writing different sequence file formats and multiple sequence alignments, dealing with 3D macro molecular structures, interacting with common tools such as BLAST, ClustalW and EMBOSS, accessing key online databases, as well as providing numerical methods for statistical learning (Stafford, 2014).

2.7.3 Physiological–Biochemical Characterization

The Expasy Protparam server was used for the physicochemical characterization and to know the molecular weight, theoretical isoelectric point (pI), total number of negative and positive residues, aliphatic index, extinction coefficient, instability index, and grand average hydropathicity (GRAVY) of the predicted protein (Gasteiger *et al.,* 2005; Yu *et al.,* 2010; Artimo *et al.,* 2012).

2.7.4 Subcellular Localization

Subcellular localization of the protein was identified by Psort (http://psort.nibb.aC.jp/form2.html) (Gardy *et al.,* 2004; Wang *et al.,* 2005).

2.7.5 Secondary Structure Prediction

The Chou and Fasman Secondary Structure Prediction (CFSSP) server (Gardy *et al.,* 2004; Wang *et al.,* 2005) was used to analyze the secondary structure of the *ERG11* gene amino acid sequence (Fasman, 2012; Kumar, 2013; Kelley *et al.,* 2015). Secondary structural properties of the protein includes alpha helix, beta pleated sheets, and beta turns.

2.7.6 Identification and Comparison of Conserved Domain

The conserved domain was identified using the Molecular Evolutionary Genetics Analysis 7 (MEGA7) offline bioinformatics tool (Yang *et al.,* 2015).

2.7.7 3D Structure Prediction

Phyre2, a web-based software (Yang *et al.,* 2015), was used to predict the 3D structure of the proteins (Liu *et al.,* 2016).

2.7.8 Active Site Prediction

The Active Site Prediction was done using 3DLigandSite, an automated software that is used to predict ligand binding sites (Wass *et al.,* 2010).

2.7.9 Antigenicity Site Prediction

The antigenicity site was predicted using the Immunomedicine group antigenicity prediction online server (Yao *et al.,* 2012). This program predicts those segments from within a protein sequence that are likely to be antigenic; capable of eliciting an antibody response. Predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes (Linding *et al.,* 2003). Segments are only reported if they have a minimum size of 8 residues. The reported accuracy of method is about 75%.

2.7.10 Disease Causing Region Prediction

The Disease Causing Region prediction was done using the globplot online server (Linding *et al.,* 2003). This web service looks for order/globularity or disorder tendency in the query protein based

on a running sum of the propensity for an amino acid to be in ordered or disordered state by searching domain databases and known disorders in proteins.

2.7.12 Phylogeny

Sequence retrieval of existing *Candida ERG11* gene sequences for phylogenetic analysis was done on NCBI server (Wheeler, 2007). The construction of a phylogenetic tree to depict the evolutional relationship between species and strains of *Candida* using the *ERG11* gene sequence was carried out using Molecular Evolutionary Genetic Analysis Software 7 (Mega7) (Kumar *et al.*, 2016).

2.8 Statistical Analysis

The statistical tools used are integrated algorithms in the analytical software employed in this study, including Clustal 2.1 and test of Maximum Likelihood based on the Tamura-Nei model, for the phylogenetic analysis.

CHAPTER THREE

RESULTS

3.1 Results of the Screening Experiment

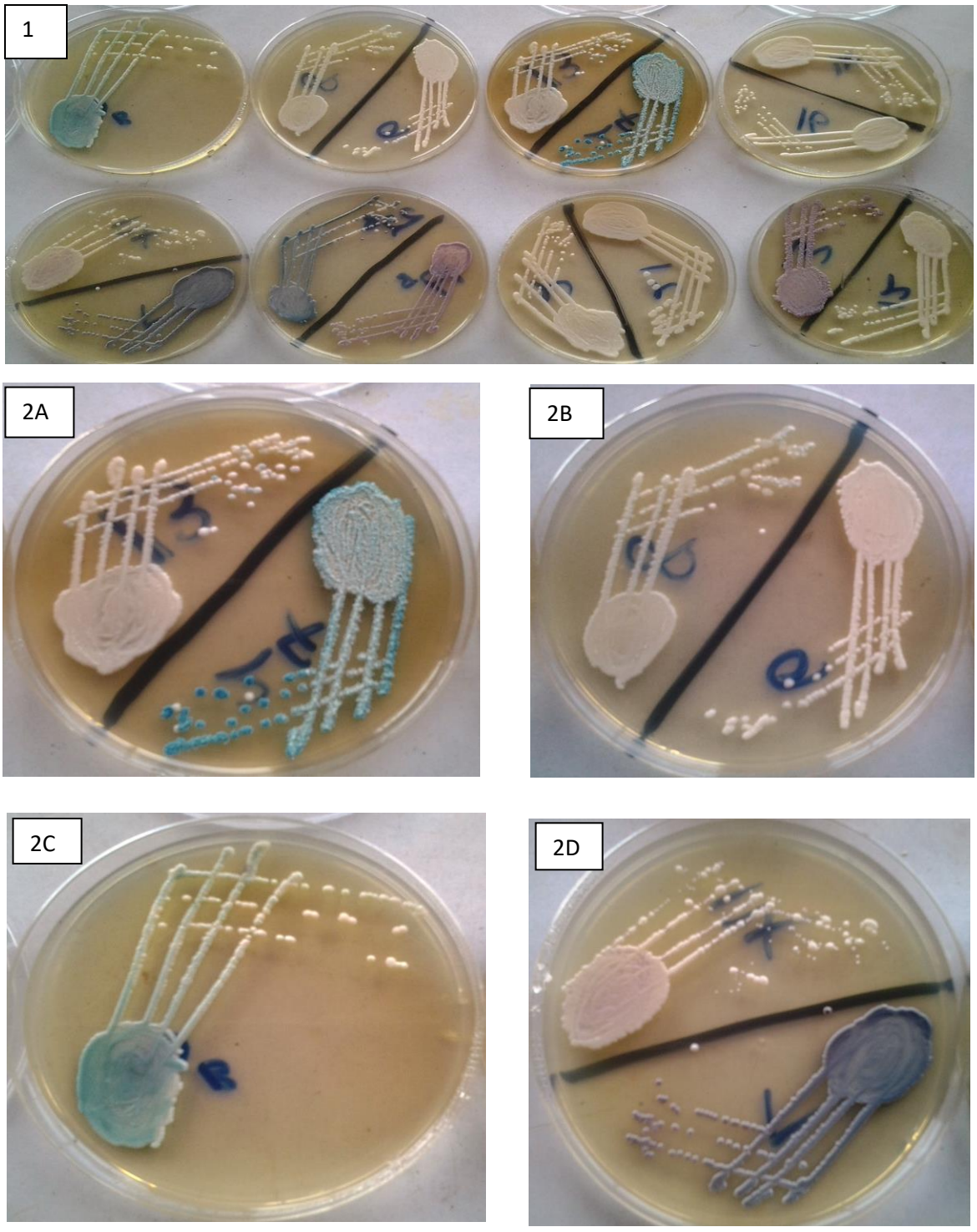
A total of 57 human samples (HHVS - Human High Vaginal Swab) and 7 dog samples (DHVS-Dog High Vaginal Swab) were screened for *Candida* species. Twenty eight (49%) of the human samples were positive (+ve) to growth and to germ-tube test. The result is presented in Table 1. Correspondingly, 29 (51%) of the samples were negative (-ve) to growth on the media, and germ-tube test. For the dog samples, 4 (57%) samples were positive to growth and to germ-tube test. Complimentarily, 3 (43%) samples were negative to -ve to growth on the media and germ-tube test. Of the 4 Dog samples, however, after considering the growth characteristics, only one dog sample, DHVS-001 was judge suitable for the next phase of the experiment.

Table 1: Prevalence of *Candida* Species in the Samples Collected

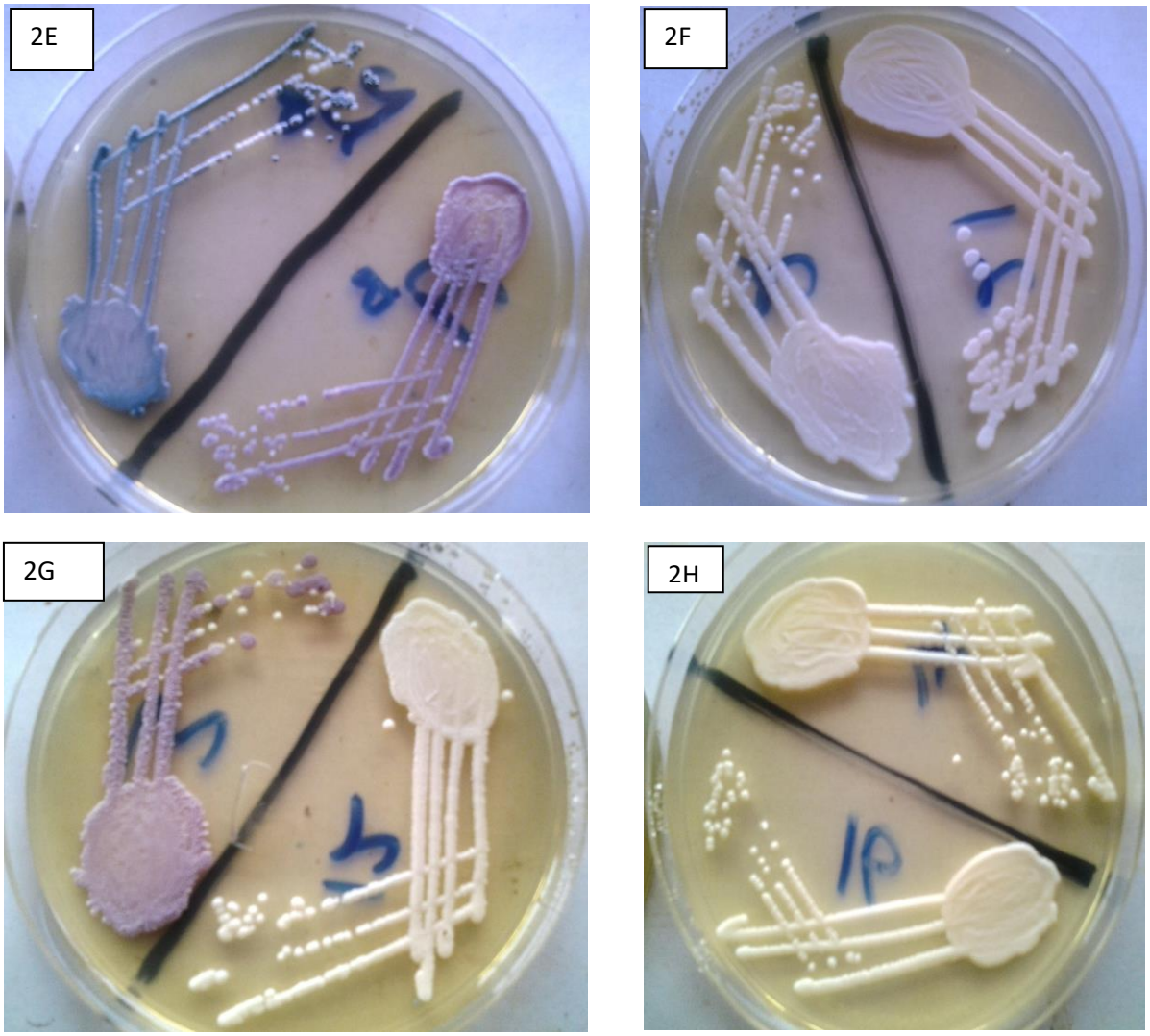
	Frequency of +ve indications	**Percentage of +ve indications**	**Frequency of –ve indications**	**Percentage of –ve indications**
Human samples	28	49%	29	51%
Dog samples	4	57%	3	43%

3.2 Differential (Chromogenic) Characterization of Yeast Species Isolated

The result revealed the presence of different species of *Candida*, with *C. albicans C. tropicalis, C. krusei,* and *C. glabrata* identified by varying colours as shown on Plates 1 and 2 (2A – 2H). The growth has the following microorganism typical colony colour appearance interpretations. *C. albicans* → green; *C. tropicalis* → metallic blue; *C. krusei* → pink - fuzzy; *C. glabrata* → mauve-brown; other species → white to mauve. Some showed mixed growth while others showed the growth of a single isolate. Twenty one samples had single species, 5 samples had two different (mixed) species (HHVS-006, HHVS-019, HHVS-038, HHVS-041, HHVS-047) and 2 samples had three different (mixed) species (HHVS-029, HHVS-039). The result also revealed a total of 13 *C. albicans*, 4 *C. glabrata*, 5 C. *krusei*, 2 *C. tropicalis*, and 14 other *Candida* species; making a total of 38 isolates. The isolate from the dog is a *C. krusei*. The summary of findings are presented on Table 2.



Plates 1 and 2: Differential characterization of yeast on *Candida*-Specific chromogenic media. Plates 1 and 2A – 2D represents both the single and mixed isolates from the samples



Plates 2: Differential characterization of yeast on *Candida*-Specific chromogenic media. Plates 2E – 2H represents both single and mixed isolates from the samples

Table 2: *Candida* species isolated from the samples

S/No	High Vaginal Sample (HVS)	Yeast Microscopic Characterization	Isolates	*Candida* Species
1	HHVS-001	+VE	*Can Iso-001*	*C. glabrata*
2	HHVS-002	+VE	*Can Iso-002*	*C. krusei*
3	HHVS-006	+VE	*Can Iso-003A*	*C. glabrata*
			Can Iso-003B	*C. krusei*
4	HHVS-007	+VE	*Can Iso-004*	other species
5	HHVS-010	+VE	*Can Iso-005*	other species
6	HHVS-011	+VE	*Can Iso-006*	other species
7	HHVS-017	+VE	*Can Iso-007*	*C. tropicalis*
8	HHVS-019	+VE	*Can Iso-008A*	other species
			Can Iso-008B	*C. albicans*
9	HHVS-020	+VE	*Can Iso-009*	*C. albicans*
10	HHVS-021	+VE	*Can Iso-010*	*C. albicans*
11	HHVS-025	+VE	*Can Iso-011*	*C. albicans*
12	HHVS-027	+VE	*Can Iso-012*	other species
13	HHVS-028	+VE	*Can Iso-013*	*C. albicans*
14	HHVS-029	+VE	*Can Iso-014A*	*C. krusei*
			Can Iso-014B	*C. glabrata*
			Can Iso-014c	other species
15	HHVS-033	+VE	*Can Iso-015*	*C. albicans*
16	HHVS-035	+VE	*Can Iso-016*	other species
17	HHVS-036	+VE	*Can Iso-017*	*C. albicans*
18	HHVS-037	+VE	*Can Iso-018*	other species
19	HHVS-038	+VE	*Can Iso-019A*	*C. krusei*
			Can Iso-019B	*C. tropicalis*
20	HHVS-039	+VE	*Can Iso-020A*	other species
			Can Iso-020B	*C. albicans*
			Can Iso-020C	*C. glabrata*
21	HHVS-040	+VE	*Can Iso-021*	other species
22	HHVS-041	+VE	*Can Iso-022A*	other species
			Can Iso-022B	*C. albicans*
23	HHVS-044	+VE	*Can Iso-023*	*C. albicans*
24	HHVS-049	+VE	*Can Iso-024A*	*C. albicans*
			Can Iso-024B	other species
25	HHVS-050	+VE	*Can Iso-025*	other species
26	HHVS-052	+VE	*Can Iso-026*	*C. albicans*
27	HHVS-053	+VE	*Can Iso-027*	other species
28	HHVS-055	+VE	*Can Iso-028*	*C. albicans*
29	DHVS-001	+VE	*Can Iso-029*	*C. krusei*

Albicans = 13, *glabrata* = 4, *krusei* = 5, tropicalis = 2, others = 14; **Total**= 38 single isolates.

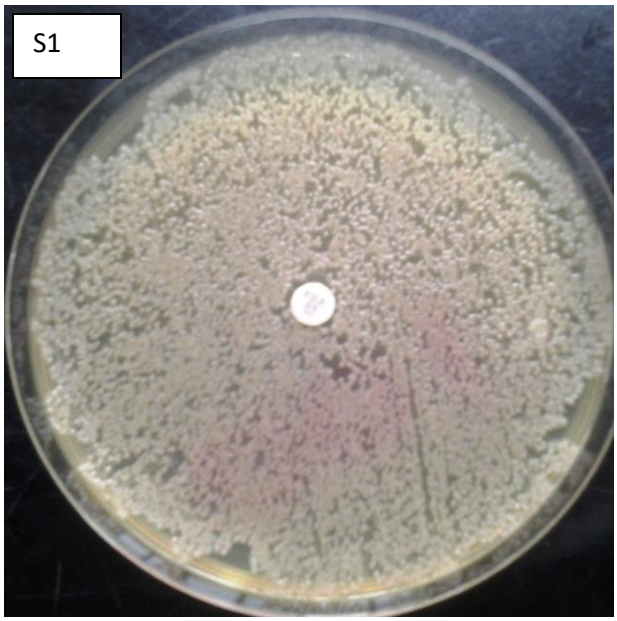
Table 2: *Candida* species isolated from the samples

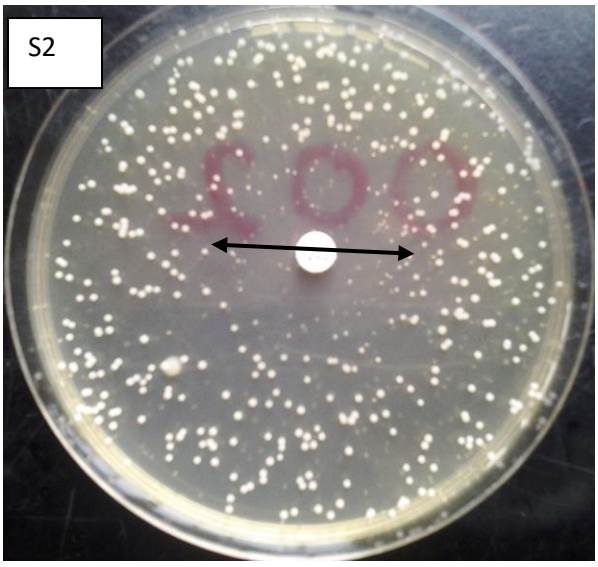
S/No	High Vaginal Sample (HVS)	Yeast Microscopic Characterization	Isolates	*Candida* Species
1	HHVS-001	+VE	*Can Iso-001*	*C. glabrata*
2	HHVS-002	+VE	*Can Iso-002*	*C. krusei*
3	HHVS-006	+VE	*Can Iso-003A*	*C. glabrata*
			Can Iso-003B	*C. krusei*
4	HHVS-007	+VE	*Can Iso-004*	other species
5	HHVS-010	+VE	*Can Iso-005*	other species
6	HHVS-011	+VE	*Can Iso-006*	other species
7	HHVS-017	+VE	*Can Iso-007*	*C. tropicalis*
8	HHVS-019	+VE	*Can Iso-008A*	other species
			Can Iso-008B	*C. albicans*
9	HHVS-020	+VE	*Can Iso-009*	*C. albicans*
10	HHVS-021	+VE	*Can Iso-010*	*C. albicans*
11	HHVS-025	+VE	*Can Iso-011*	*C. albicans*
12	HHVS-027	+VE	*Can Iso-012*	other species
13	HHVS-028	+VE	*Can Iso-013*	*C. albicans*
14	HHVS-029	+VE	*Can Iso-014A*	*C. krusei*
			Can Iso-014B	*C. glabrata*
			Can Iso-014c	other species
15	HHVS-033	+VE	*Can Iso-015*	*C. albicans*
16	HHVS-035	+VE	*Can Iso-016*	other species
17	HHVS-036	+VE	*Can Iso-017*	*C. albicans*
18	HHVS-037	+VE	*Can Iso-018*	other species
19	HHVS-038	+VE	*Can Iso-019A*	*C. krusei*
			Can Iso-019B	*C. tropicalis*
20	HHVS-039	+VE	*Can Iso-020A*	other species
			Can Iso-020B	*C. albicans*
			Can Iso-020C	*C. glabrata*
21	HHVS-040	+VE	*Can Iso-021*	other species
22	HHVS-041	+VE	*Can Iso-022A*	other species
			Can Iso-022B	*C. albicans*
23	HHVS-044	+VE	*Can Iso-023*	*C. albicans*
24	HHVS-049	+VE	*Can Iso-024A*	*C. albicans*
			Can Iso-024B	other species
25	HHVS-050	+VE	*Can Iso-025*	other species
26	HHVS-052	+VE	*Can Iso-026*	*C. albicans*
27	HHVS-053	+VE	*Can Iso-027*	other species
28	HHVS-055	+VE	*Can Iso-028*	*C. albicans*
29	DHVS-001	+VE	*Can Iso-029*	*C. krusei*

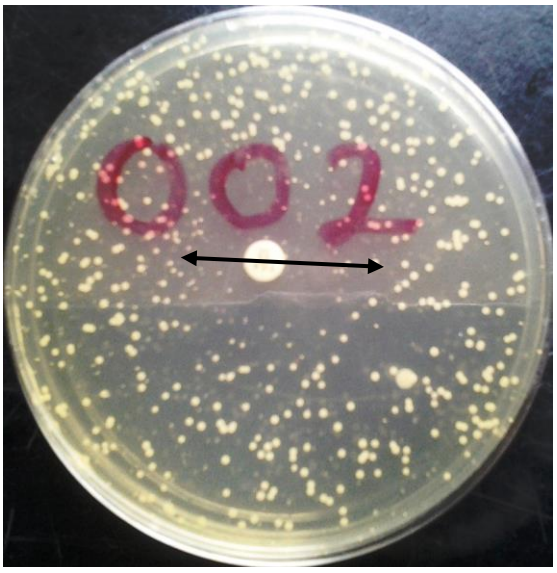
Albicans = 13, *glabrata* = 4, *krusei* = 5, tropicalis = 2, others = 14; **Total**= 38 single isolates.

3.3 Susceptibility of *Candida* Isolates to 25 µg Fluconazole

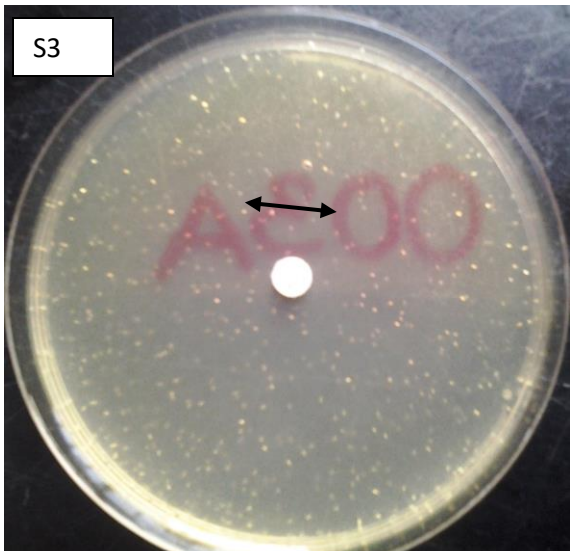
The antifungal susceptibility testing of the human samples revealed that 33 of the *Candida* isolates were susceptible (≥19mm) to 25 µg fluconazole, 2 were resistant (≤14mm), 2 were susceptible dose dependent (15 – 18mm), and the isolate from dog was susceptible. Of the 13 species of *C. albicans* tested, only one was resistant to fluconazole, one was susceptible dose dependent, while the remaining 11 were susceptible. HHVS-001(*Can Iso-001- C. glabrata*) was the most resistant (0.00mm). HHVS-055 (*Can Iso-028 - C. albicans*) was the most susceptible. The susceptibility plates (top and bottom views) are presented in plates S1 – S38 (where S refers to susceptibility). Table 3 shows that there was no significant difference in the value of zone of inhibition of the *Candida* Isolates to 25 µg fluconazole at 24 hours (M = 25.22, SD = 7.07) and at 48 hours (M= 26.87, SD=7.64) conditions; t(74)= -0.976, p = 0.332. The 24 hours schedule was therefore used.

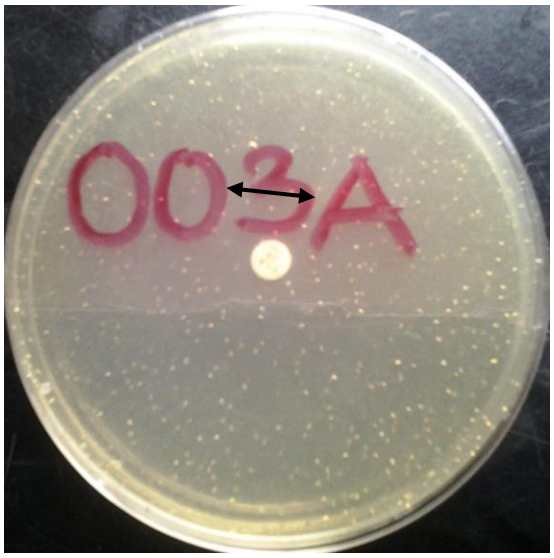


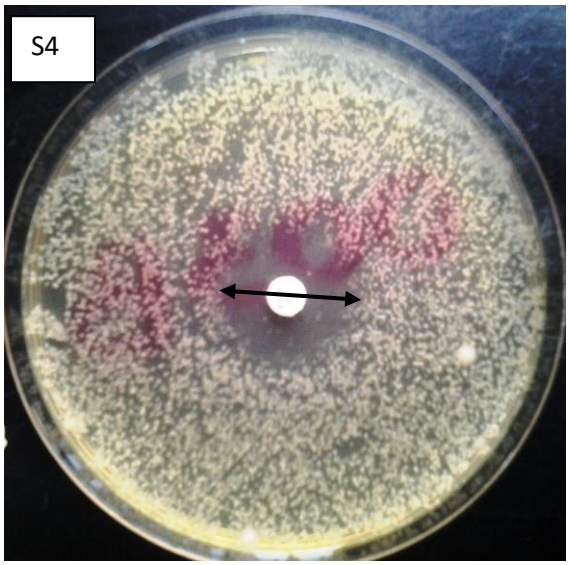


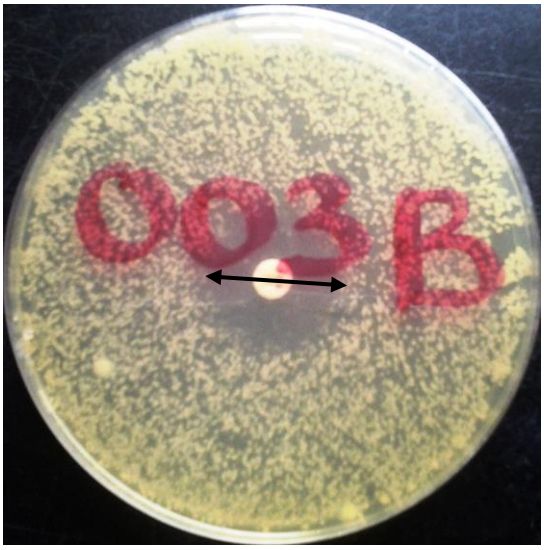


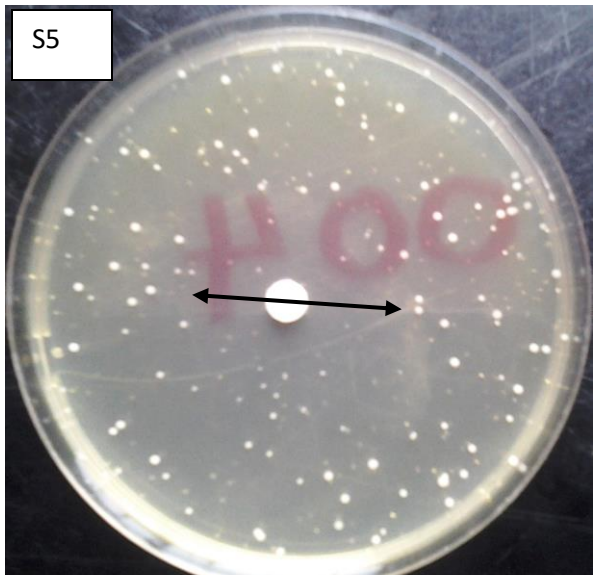
Plates S1 – S38: Susceptibility of *Candida* isolates to 25µg fluconazole

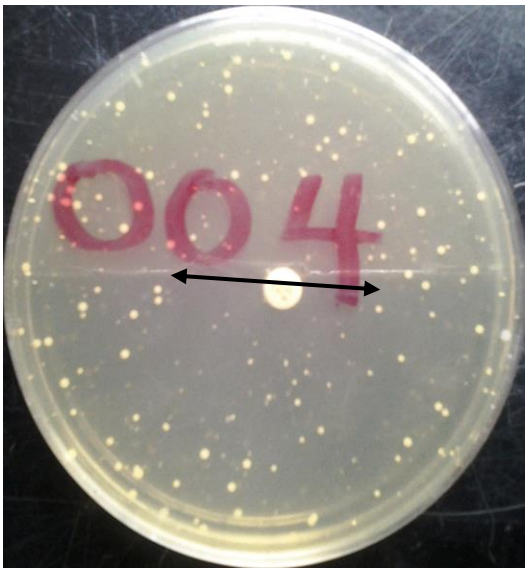

S3
003A

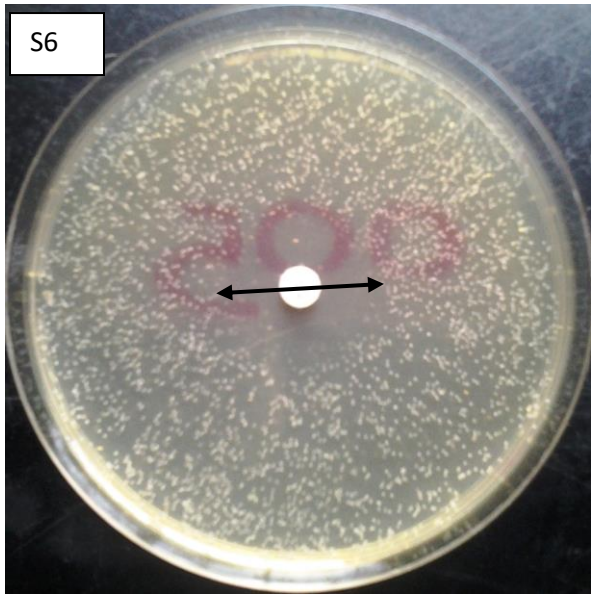

003A

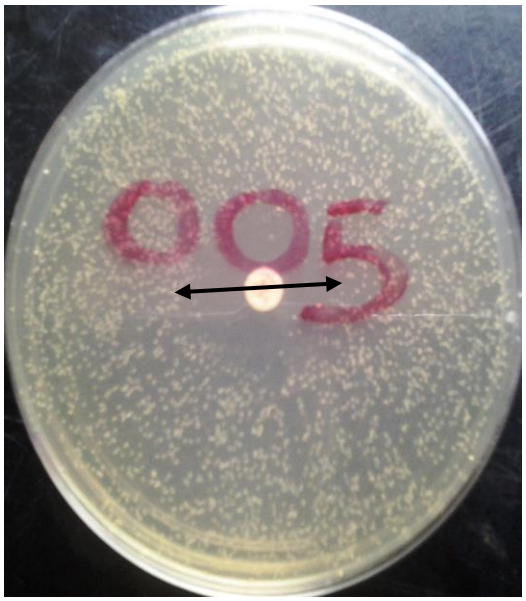

S4

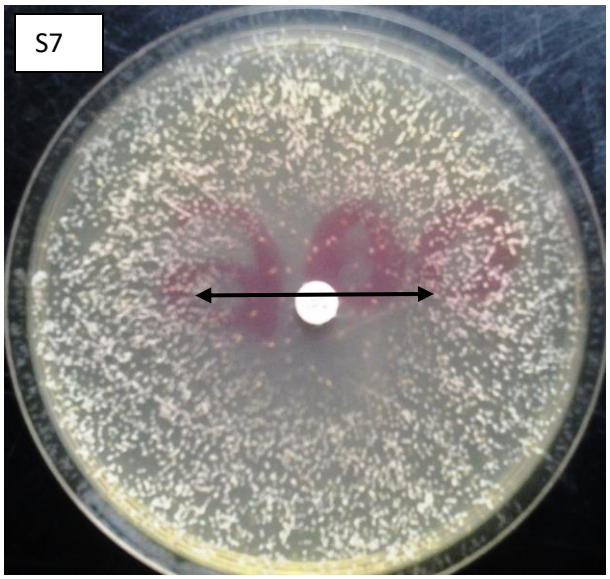

003B

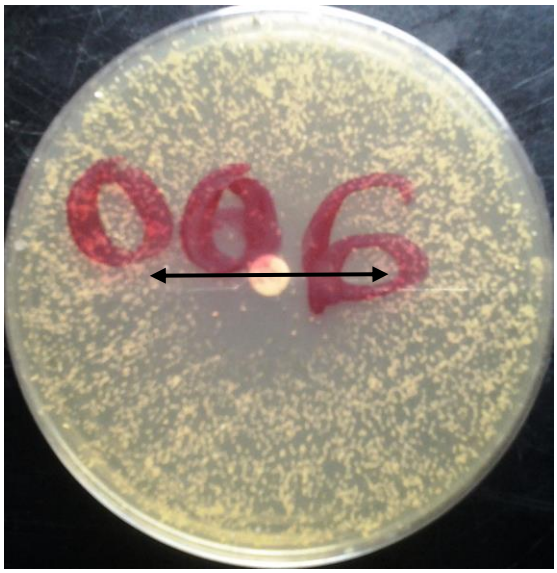

S5

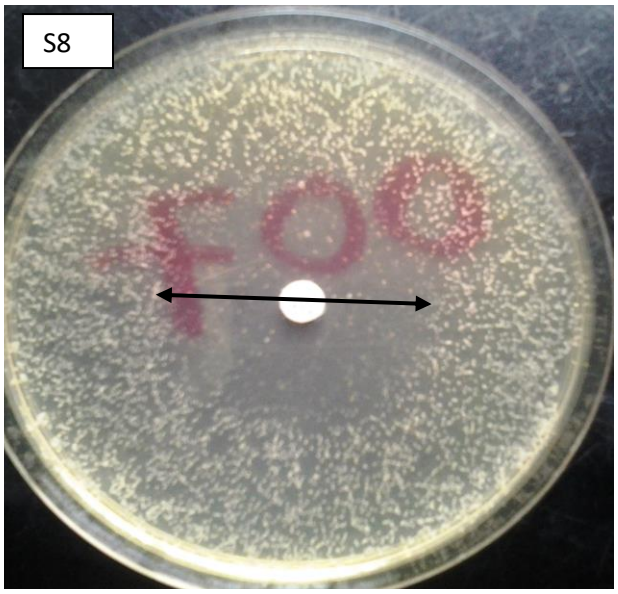

004

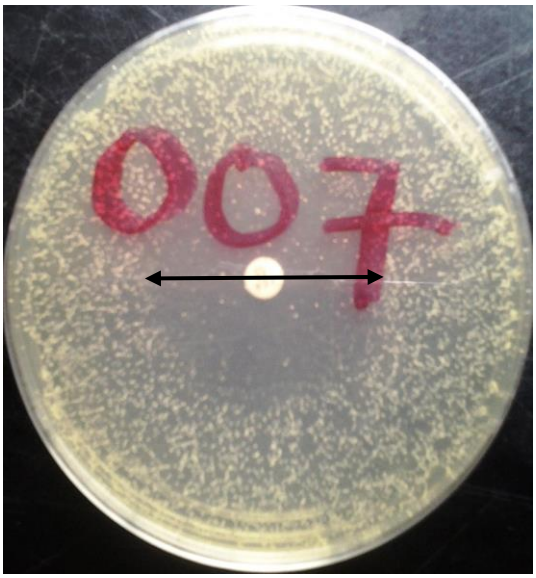

S6

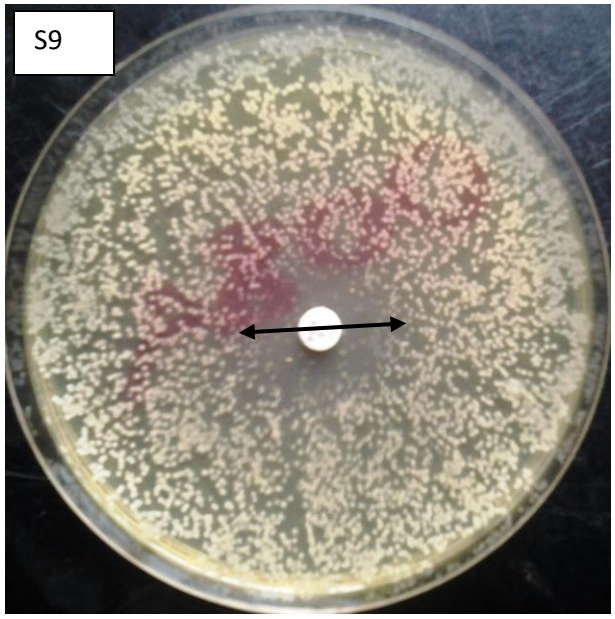

005

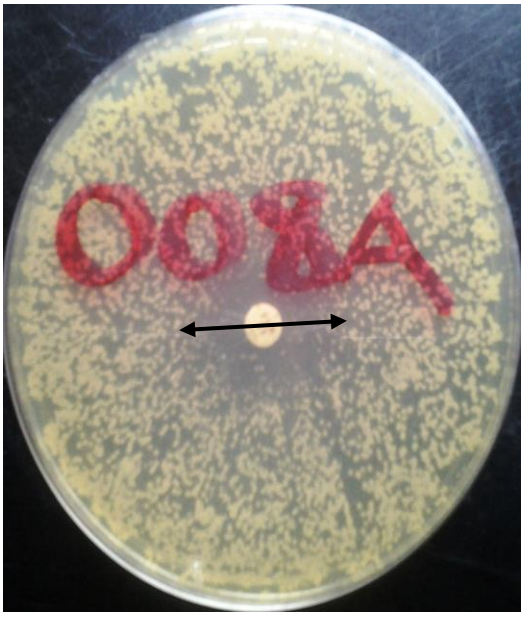

S7

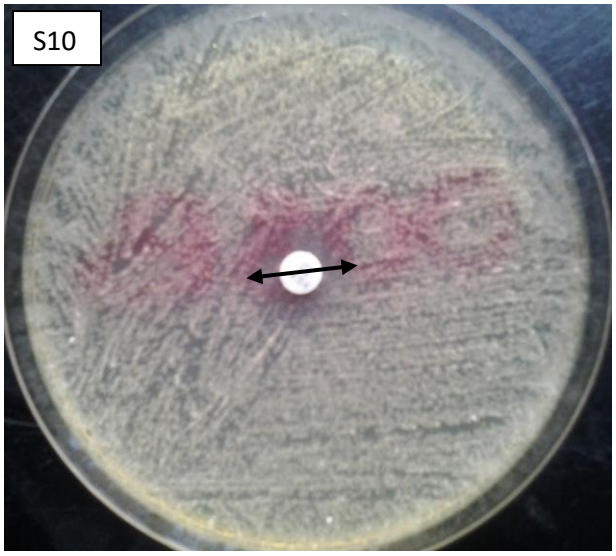

006

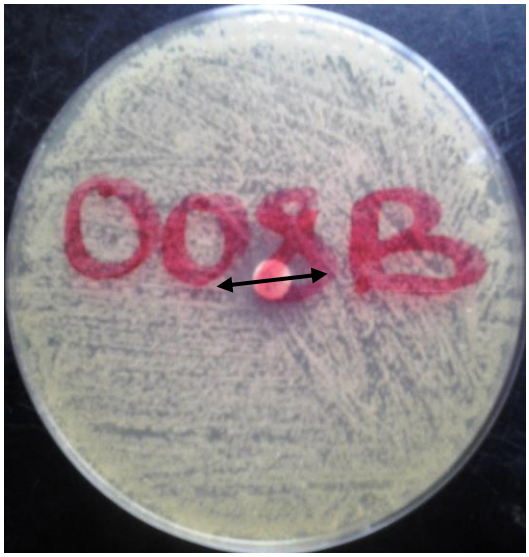

S8
100

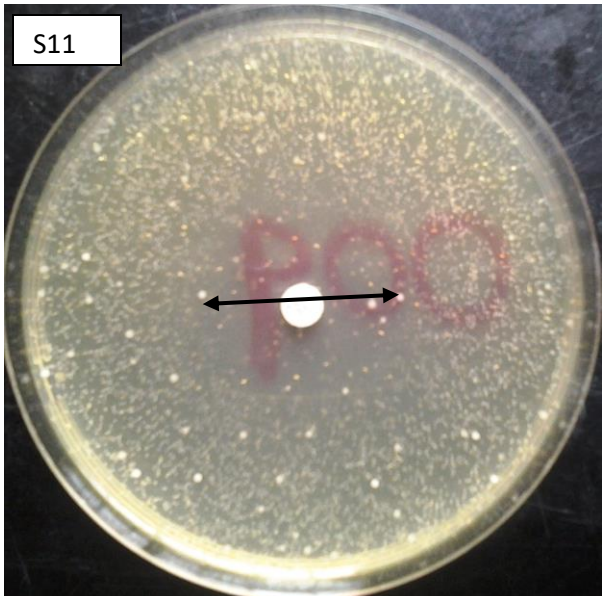

007

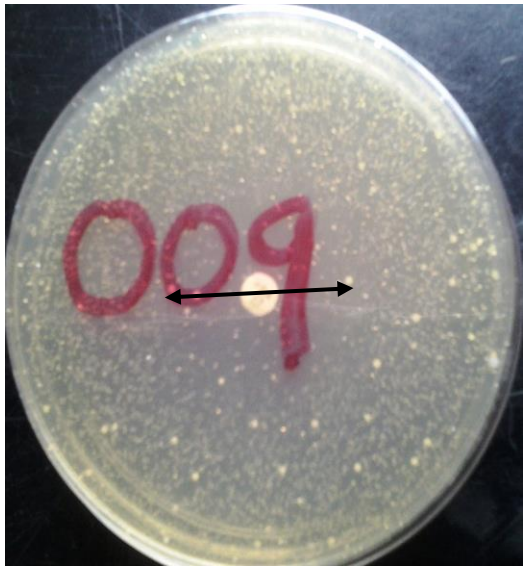

S9

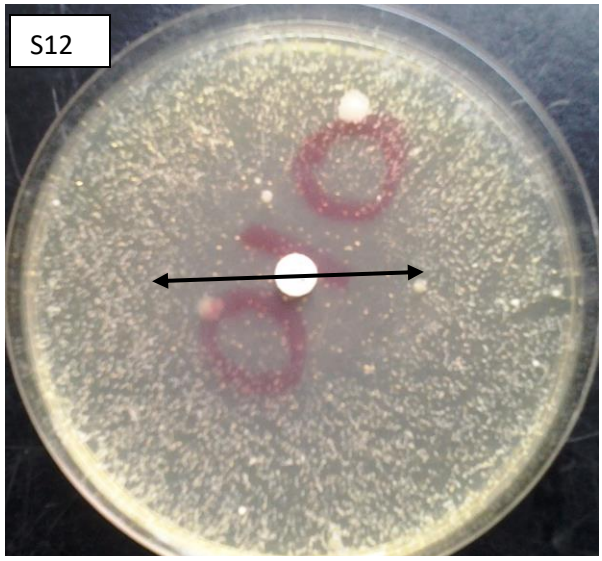

008A

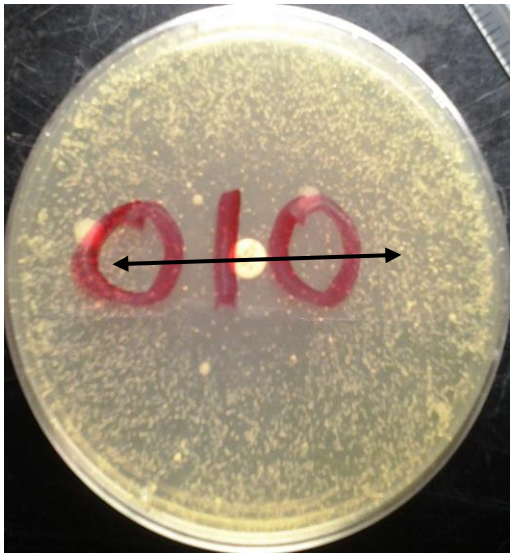

S10

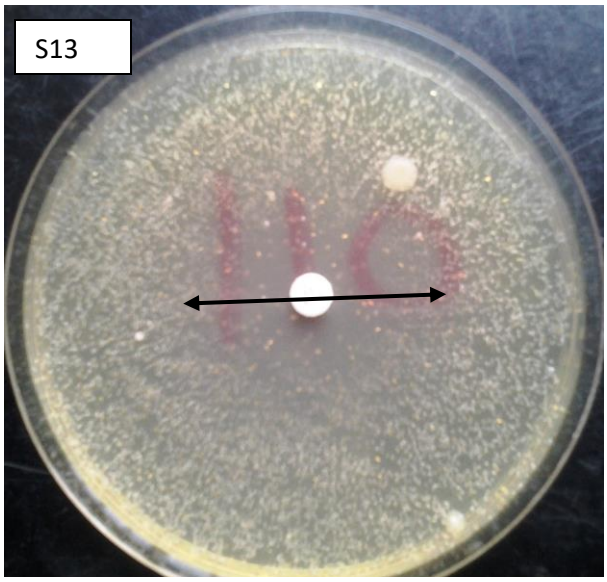

008B

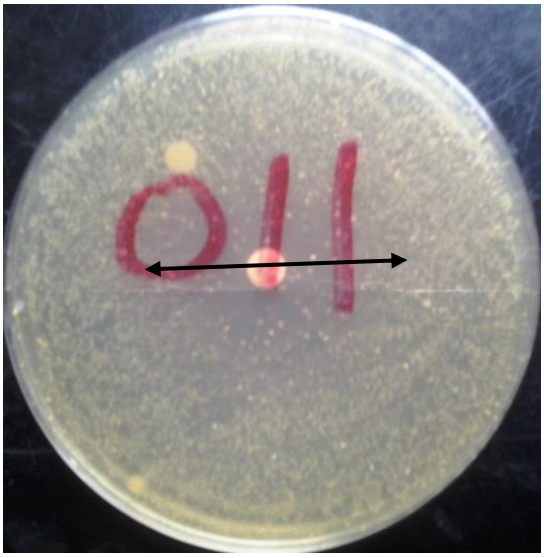

S11

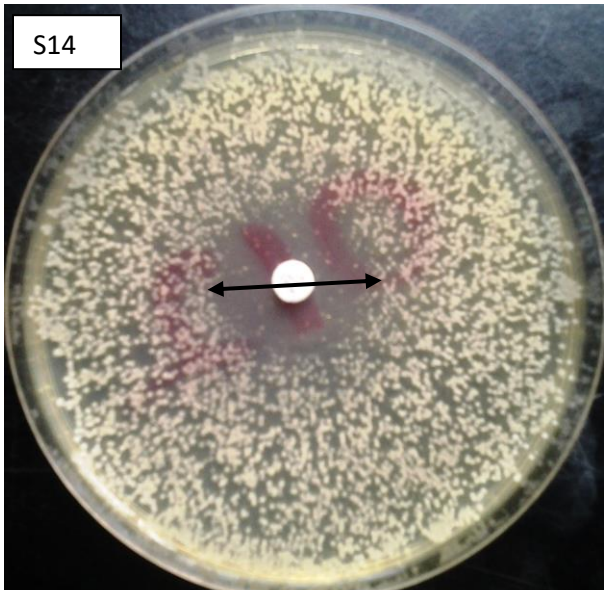


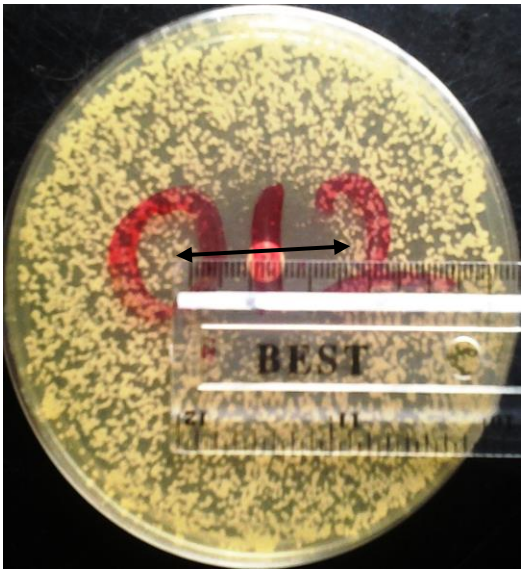

S12

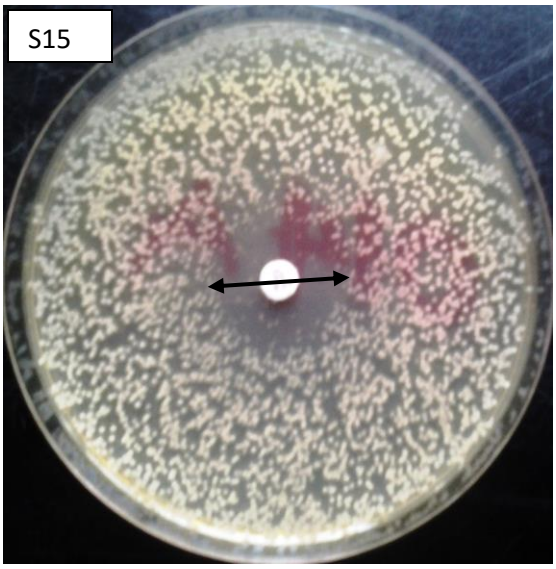


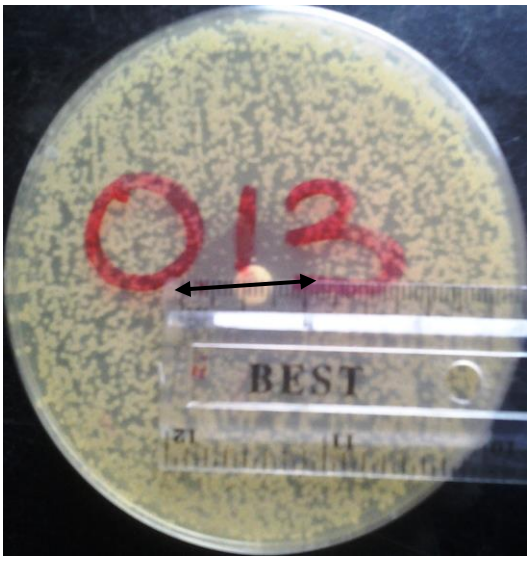

S13

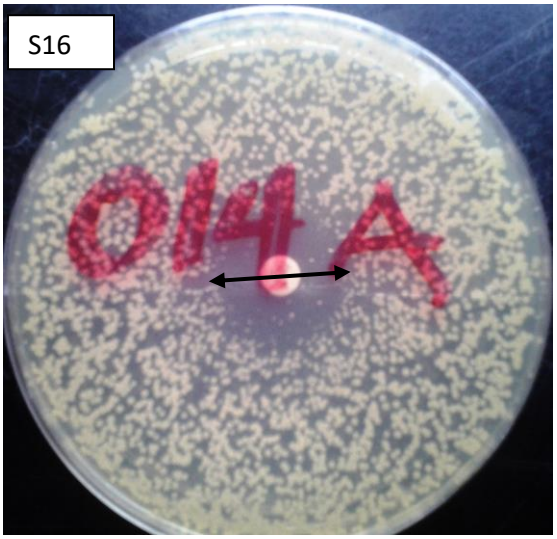


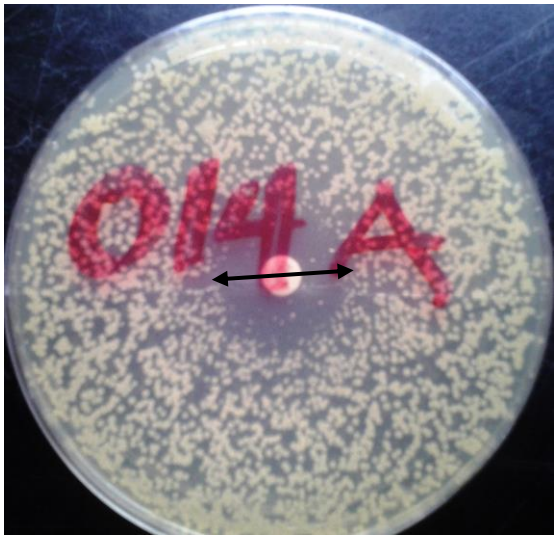

S14

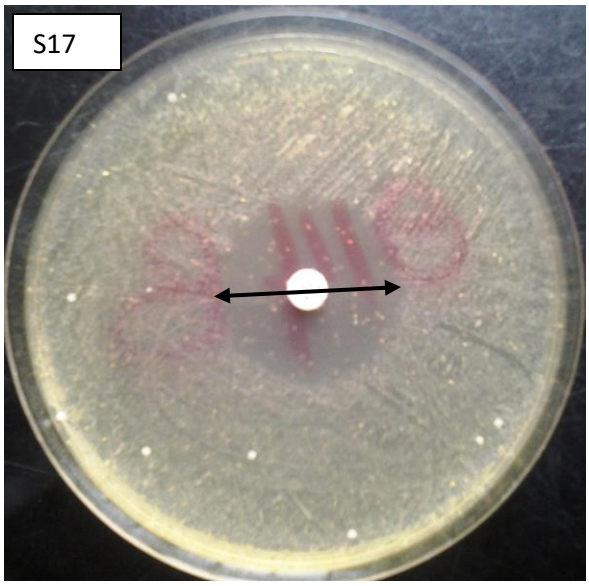

BEST

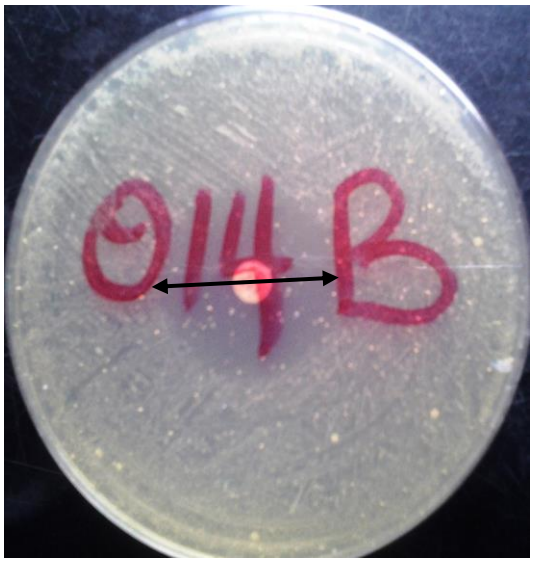

S15


013
BEST

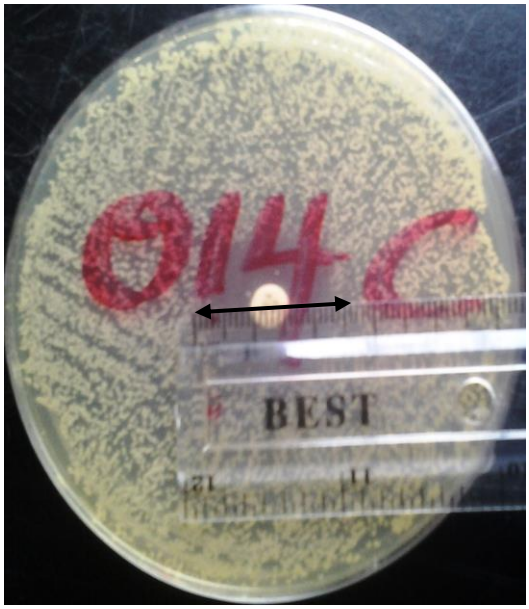

S16
014A

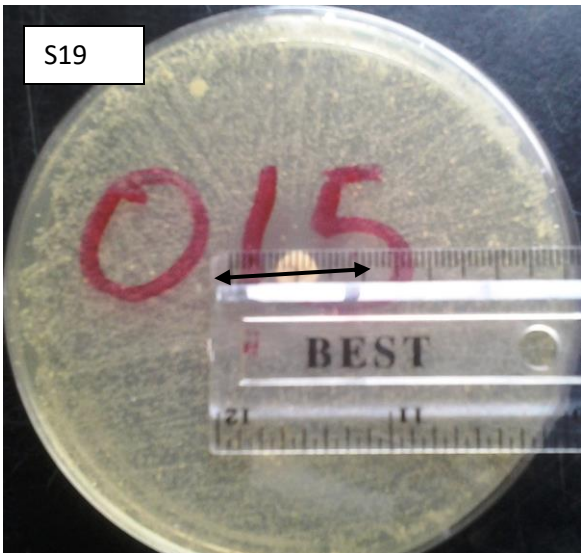

014A

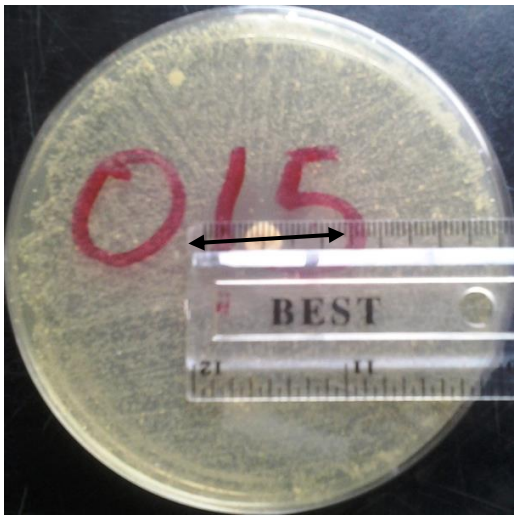

S17

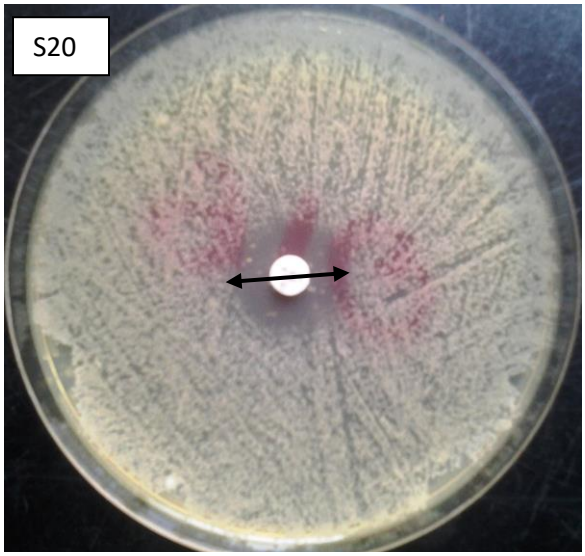

014B

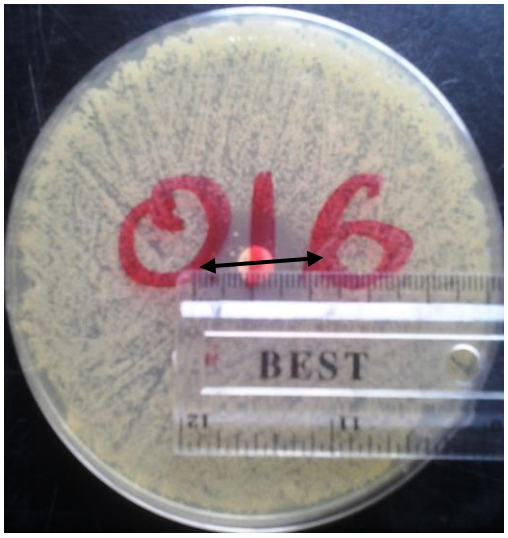

S18
014C
BEST

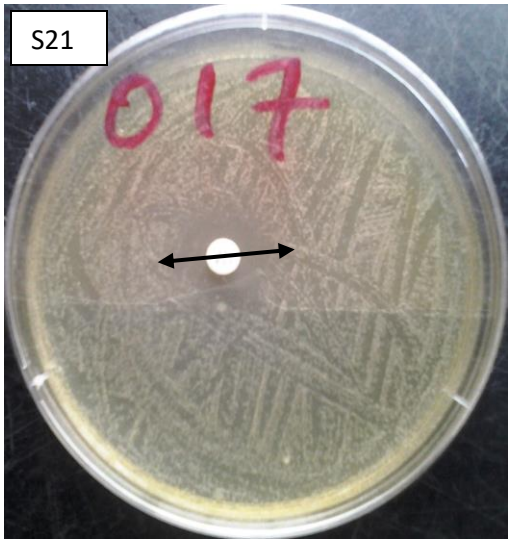

014C
BEST

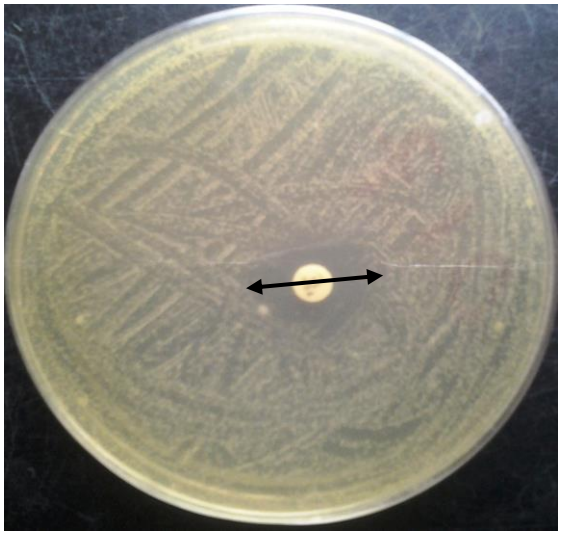

S19
015
BEST

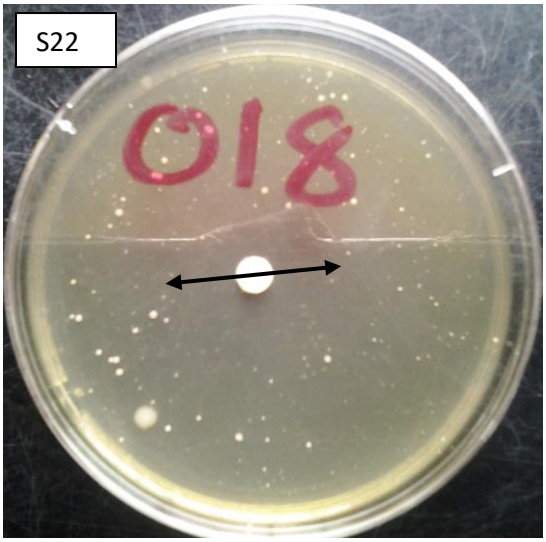

015
BEST

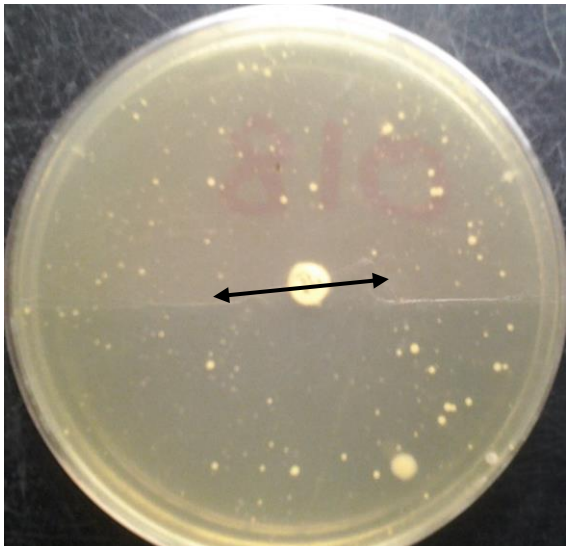

S20

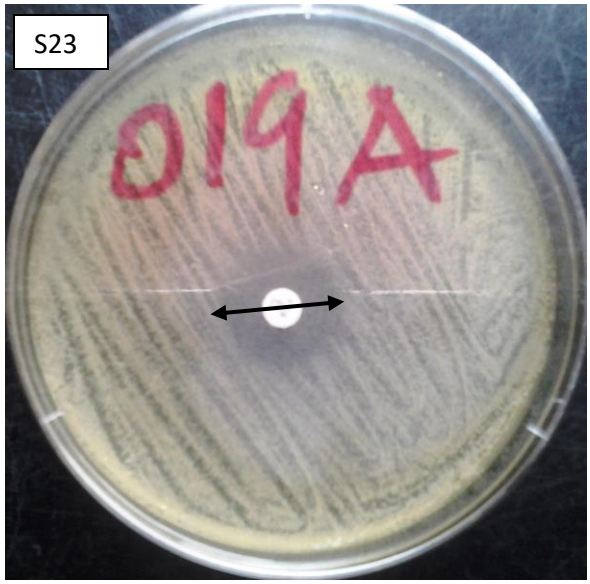

016
BEST

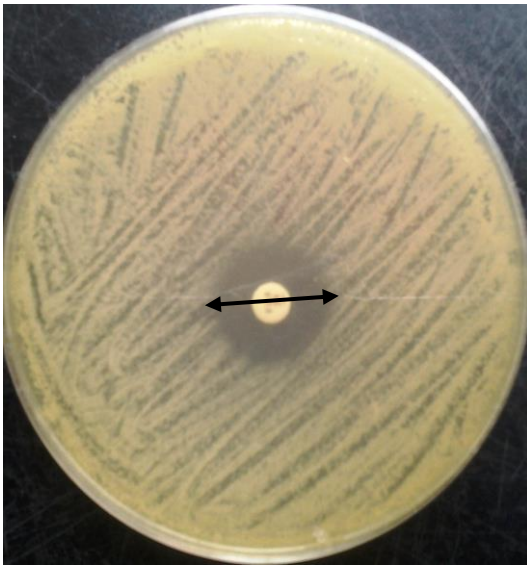

S21
017

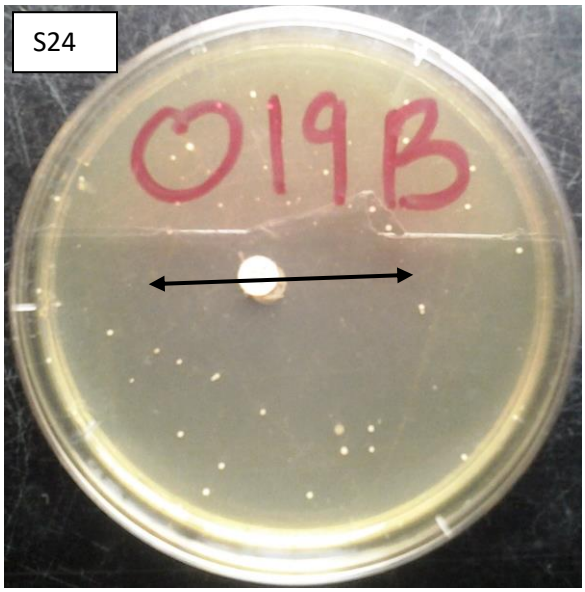


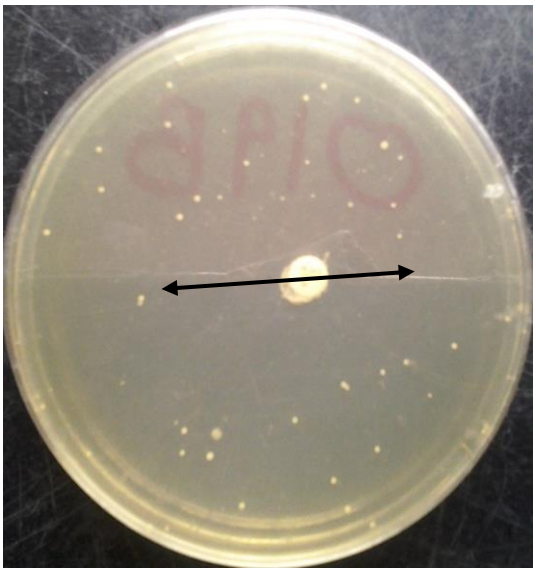

S22
018

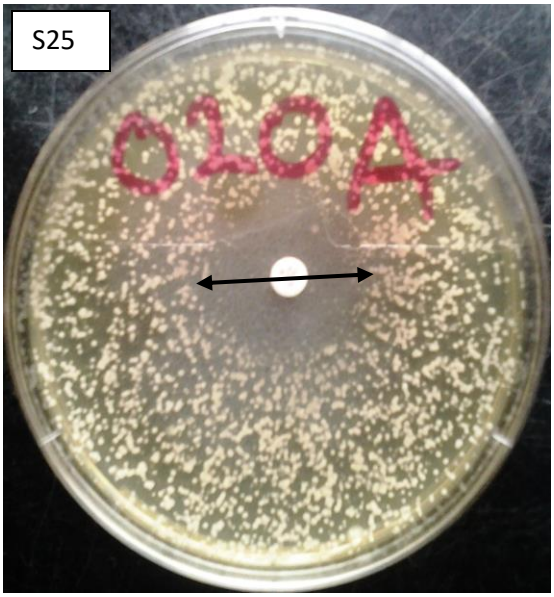


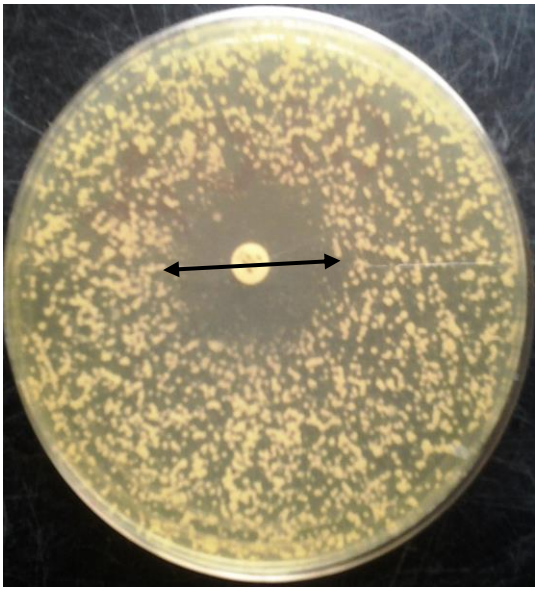

S23
019A



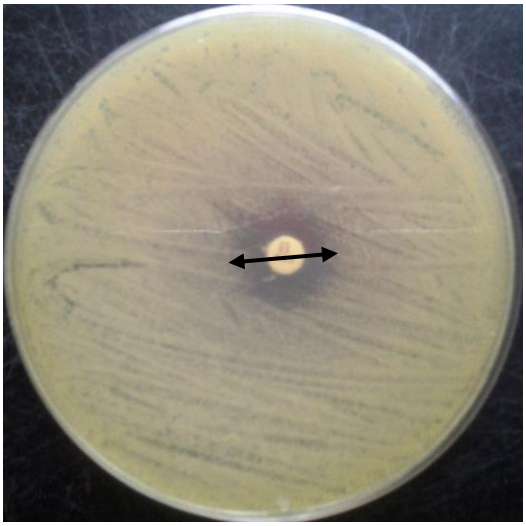

S24
019B

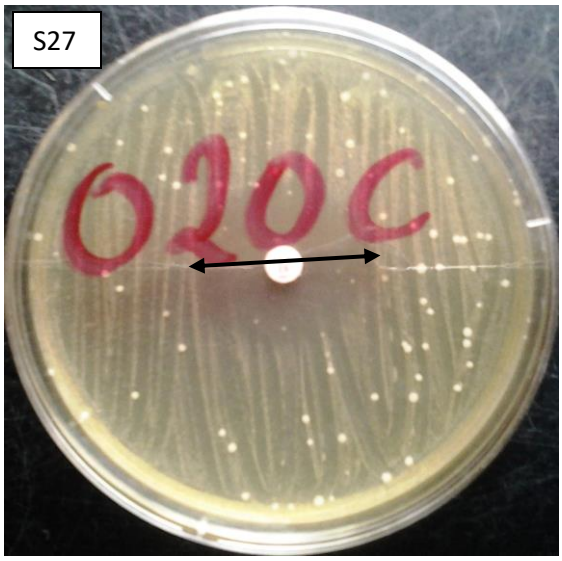


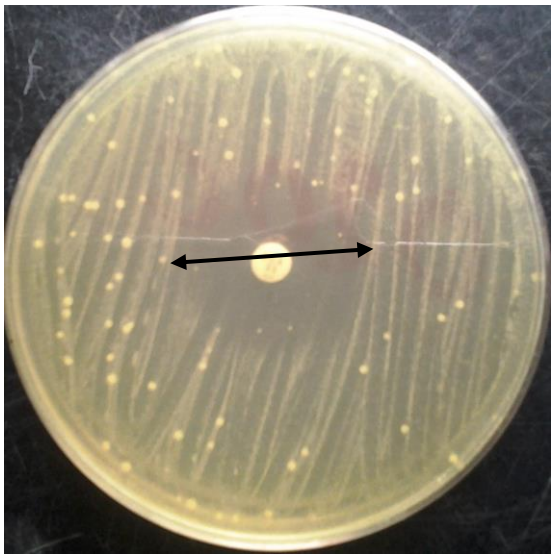

S25
010A

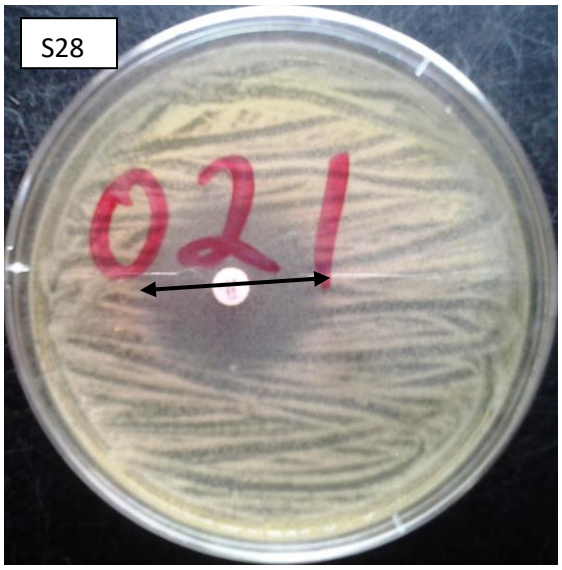


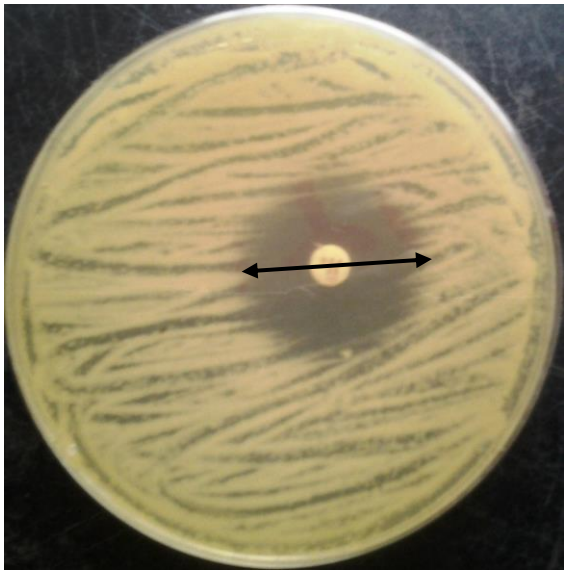

S26
020B



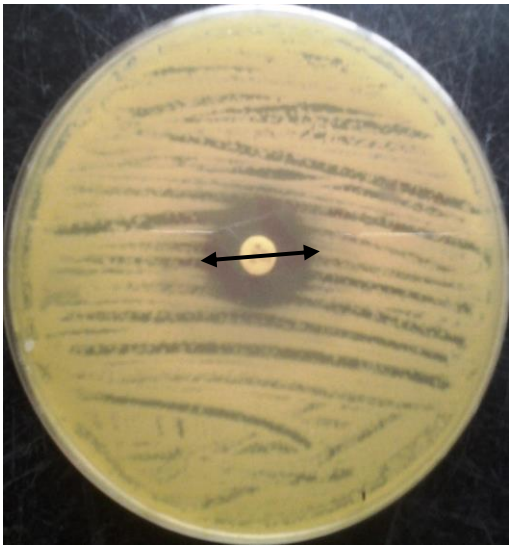

S27
020C

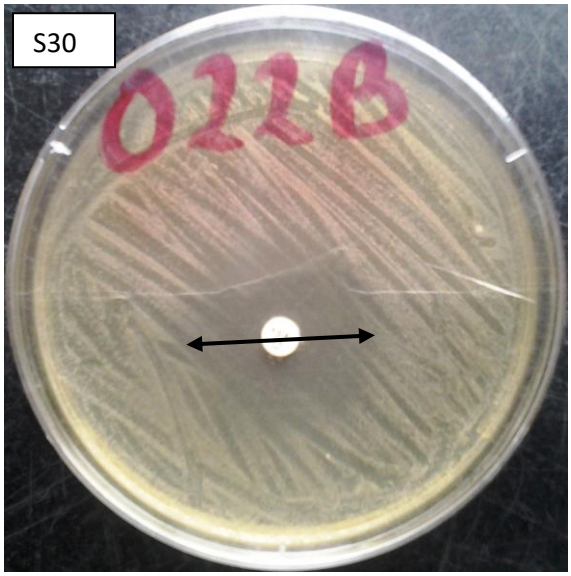


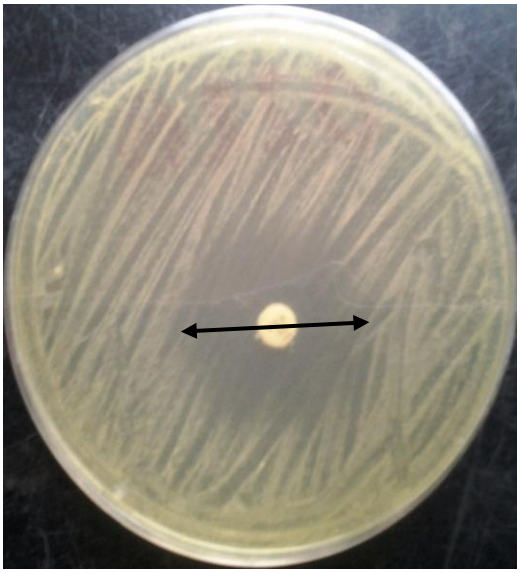

S28
021

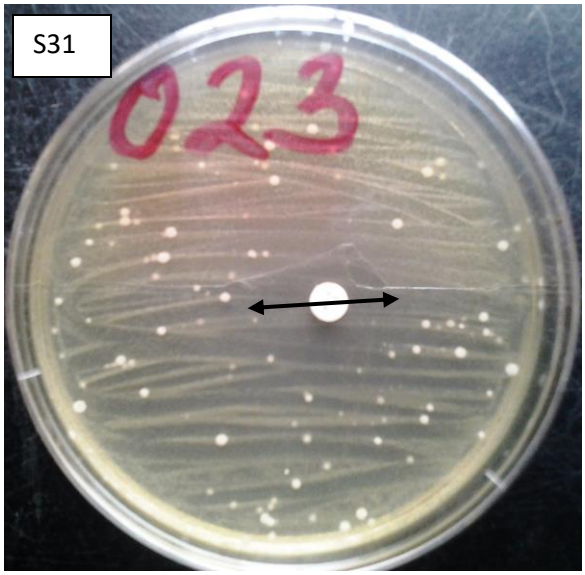


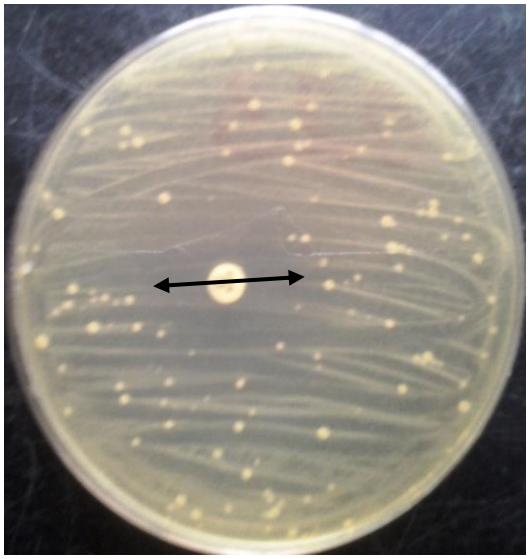

S29
022A

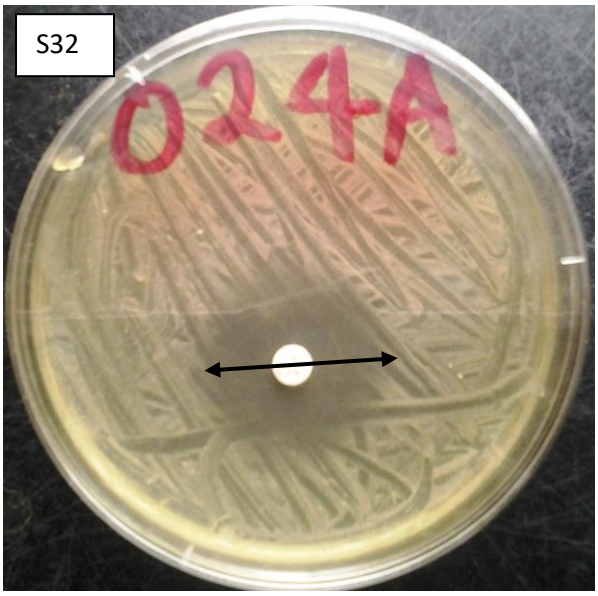


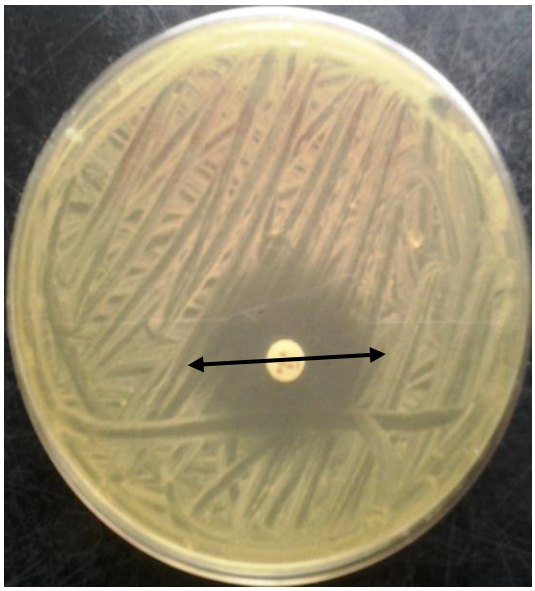

S30
022B

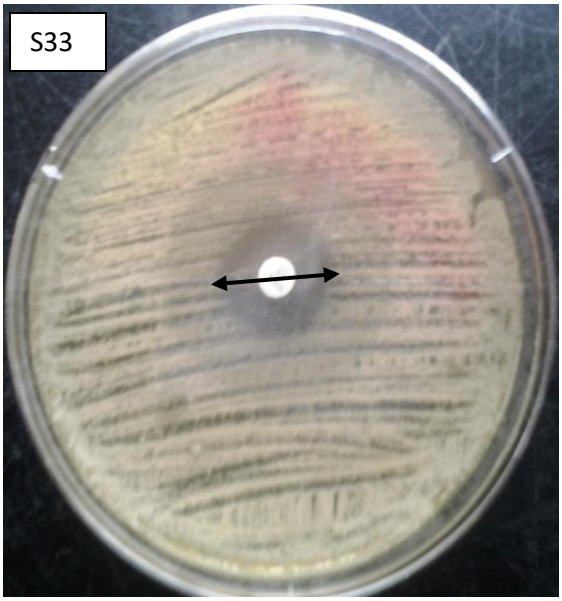


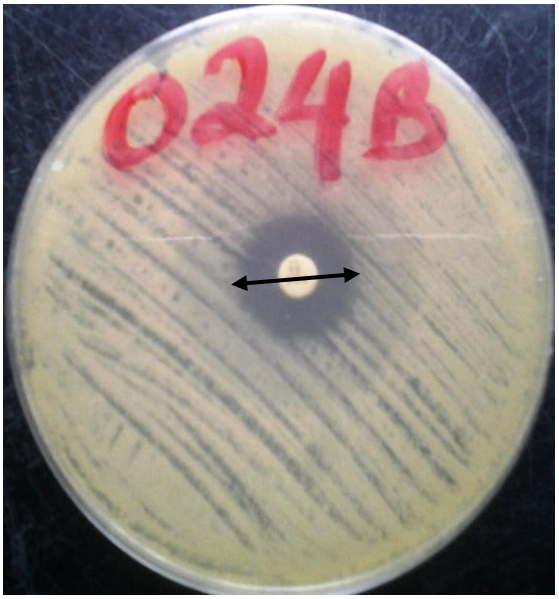

S31
023

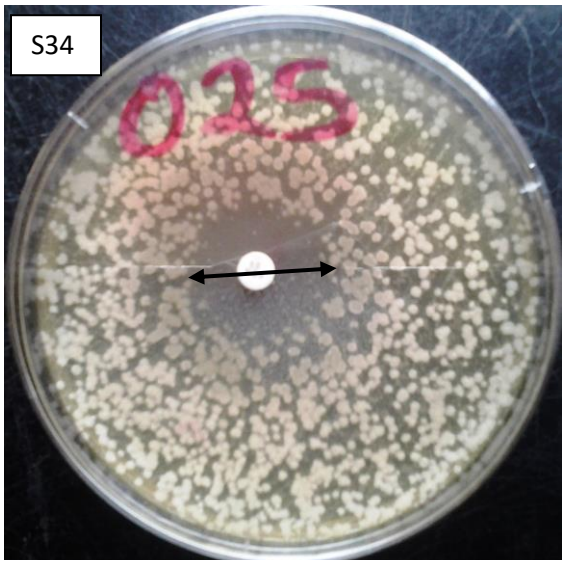


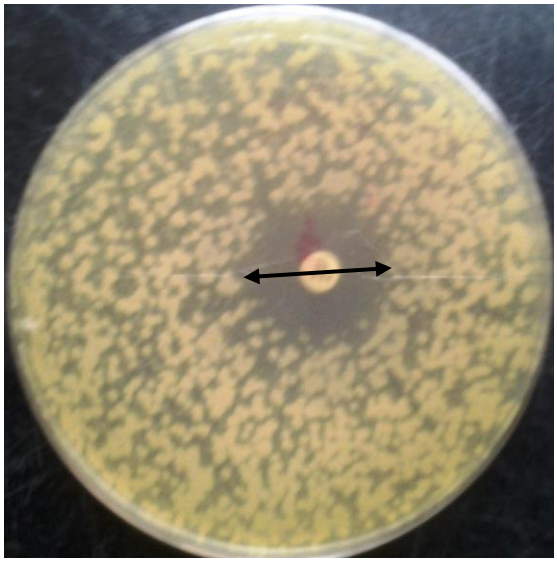

S32
024A

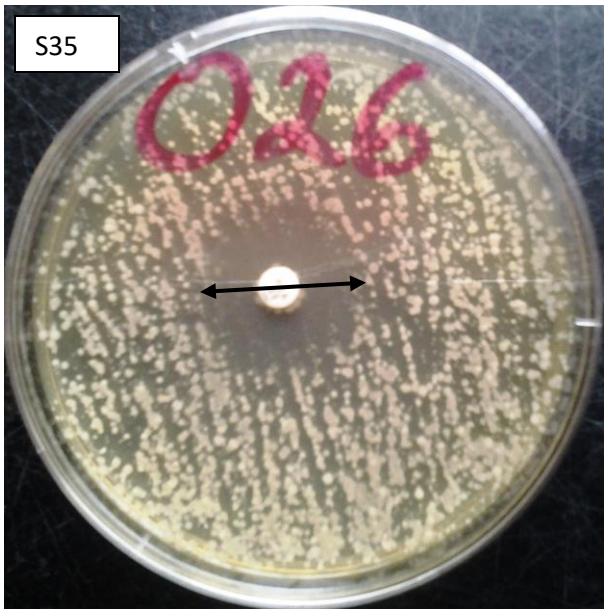


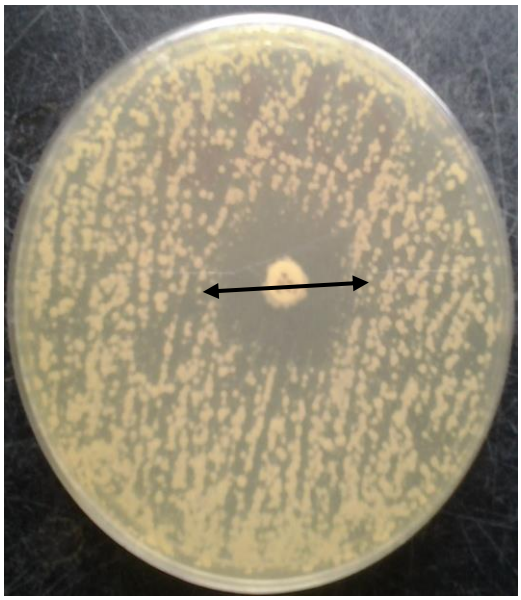

S33

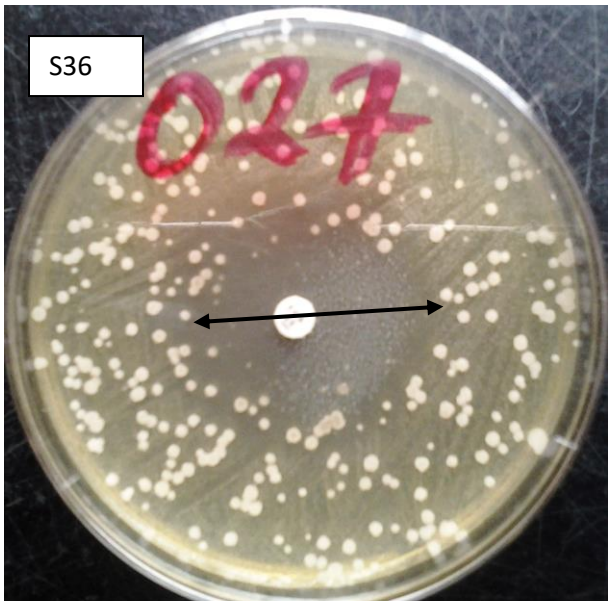

024B

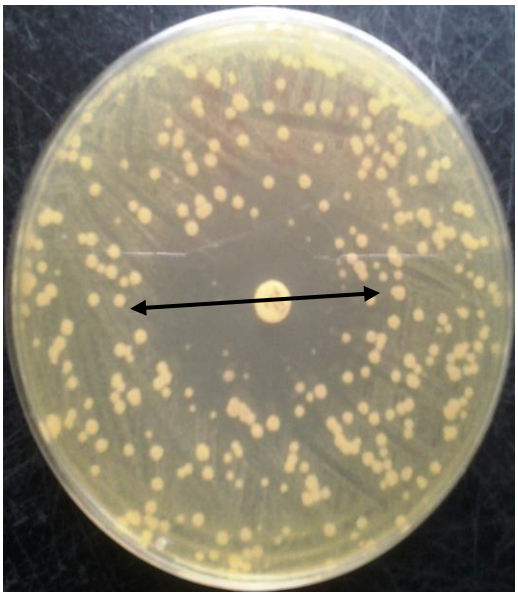

S34
025




S35
026




S36
027



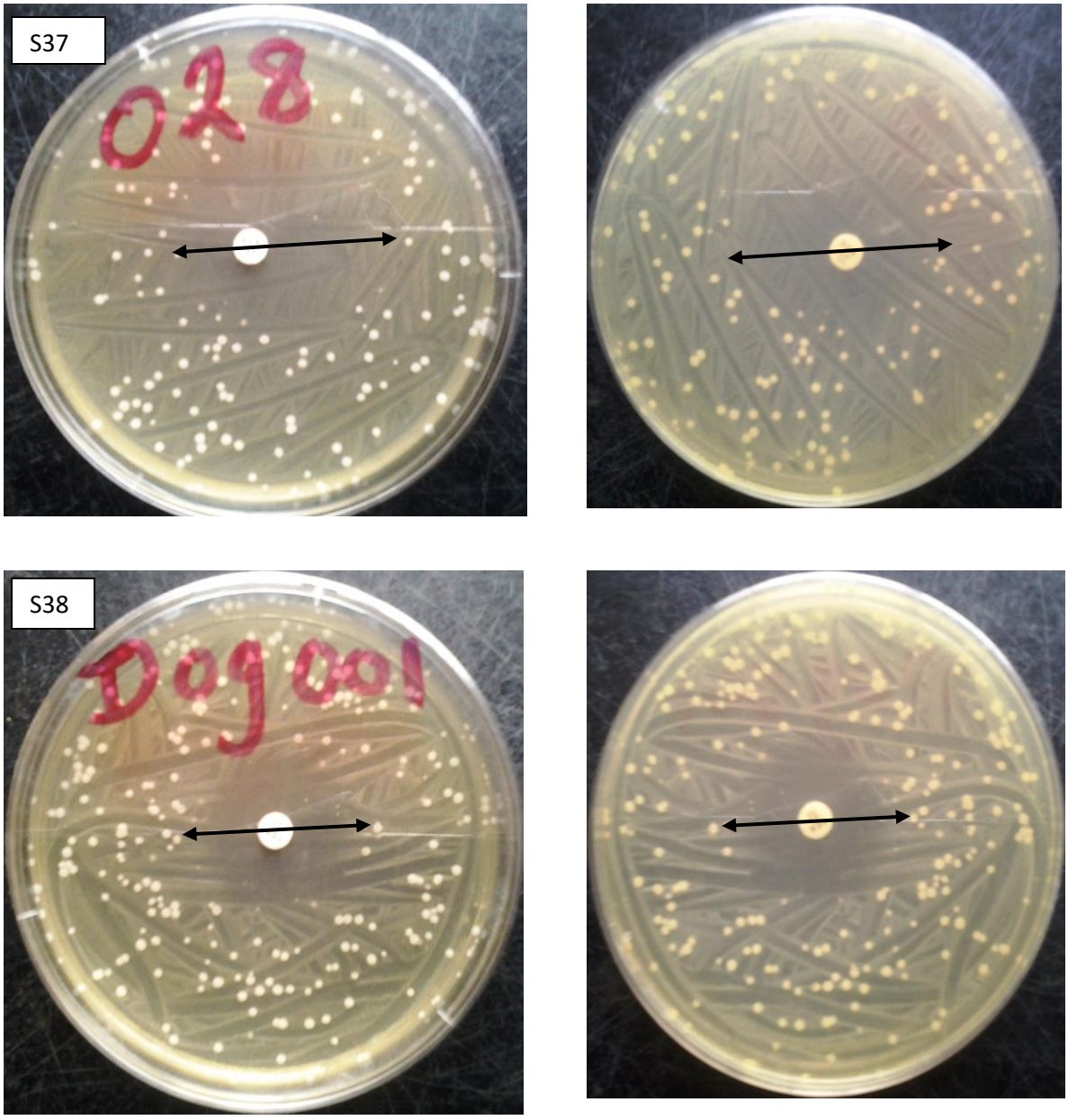


Table 3: Zone of Inhibition of the *Candida* Isolates to 25µg Fluconazole at 24 hours and 48 hours

S/N	*Candida* Isolate	24 HOURS	48 HOURS
1	***Can Iso-001***	**0.00**	**0.00**
2	*Can Iso-002*	28.33	30.00
3	*Can Iso-003A*	30.00	36.00
4	*Can Iso-003B*	24.00	24.00
5	*Can Iso-004*	33.33	33.33
6	*Can Iso-005*	23.33	26.33
7	*Can Iso-006*	29.33	30.67
8	*Can Iso-007*	39.33	38.67
9	*Can Iso-008A*	21.00	22.33
10	*Can Iso-008B*	14.00	15.33
11	*Can Iso-009*	29.00	29.00
12	*Can Iso-010*	33.33	30.67
13	*Can Iso-011*	26.67	26.67
14	*Can Iso-012*	25.33	26.67
15	*Can Iso-013*	20.67	22.33
16	*Can Iso-014A*	22.67	24.67
17	*Can Iso-014B*	30.00	30.33
18	*Can Iso-014c*	20.33	24.67
19	*Can Iso-015*	21.00	22.33
20	*Can Iso-016*	18.67	20.33
21	***Can Iso-017***	**17.67**	**20.33**
22	*Can Iso-018*	24.00	26.00
23	*Can Iso-019A*	19.33	19.33
24	*Can Iso-019B*	28.00	25.67
25	*Can Iso-020A*	25.00	21.67
26	*Can Iso-020B*	19.00	22.67
27	*Can Iso-001C*	31.00	29.33
28	*Can Iso-021*	26.67	32.00
29	*Can Iso-022A*	19.67	21.33
30	*Can Iso-022B*	30.67	38.00
31	*Can Iso-023*	28.00	33.00
32	*Can Iso-024A*	29.00	31.67
33	*Can Iso-024B*	20.00	21.00
34	*Can Iso-025*	23.67	24.67
35	*Can Iso-026*	25.67	25.67
36	*Can Iso-027*	33.33	33.00
37	***Can Iso-028***	**38.00**	**42.00**
38	***Can Iso-029 (Dog)***	**29.67**	**39.67**

Susceptible = 33 isolates
Susceptible Dose-Dependent (S-DD) ** (15 – 18mm) = 2 isolates
Resistant (≤14mm) = 2 isolates
Dog (susceptible)

3.4 Nucleotide Sequence of the *ERG11* Gene of the *Candida* Isolates of Interest

Four (4) isolates were selected for gene sequencing. The most resistant (*Can Iso-001*), a dose-dependent susceptible (*Can Iso-17*), the most susceptible (*Can Iso-028*) and the dog isolate (*Can Iso-029*). The nucleotide sequences showed the nucleotide compositions of each of the *ERG11* genes sequenced to vary. It also shows the nucleotide count. The sequence in FASTA format begins with a single-line description, followed by lines of sequence data. The definition line (defline) is distinguished from the sequence data by a ">" symbol at the beginning. The word following the ">" symbol is the identifier of the sequence. In this case, the identifier was assigned since the sequence is yet to be uploaded on NCBI. The isolates were then assigned a new nomenclature as follows: *Can Iso-001* as 01_*ERG11*-F, *Can Iso-17* as 17_*ERG11*-F, *Can Iso-028* as 28_*ERG11*-F and *Can Iso-029* as Dog_*ERG11*-F. The 01_*ERG11*-F had 1431 bases, 17_*ERG11*-F had 1590 bases, 28_*ERG11*-F had I578 bases, and Dog_*ERG11*-F had 1668 bases. The sequences are presented as FASTA 1, FASTA 2, FASTA 3, and FASTA 4.

FASTA 1: *ERG11* Gene Nucleotide Sequence for >01_*ERG11*-F

>01_*ERG11*-F

GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTTGACATATTGTATACGACATATCAGTA
TATTATTATTGGTTCCTCTTGTAATCTACTGATTTGTTGTTTTTTCTTTGATTCAGAAATAATT
AATCTCCTCATGACAGTTTGACATCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTT
ATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAA
AATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGAT
GTTTCTGTTGAAGAAGCTTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATG
ATTGTCCAAATTCTAGATTAATGGAACAAAAAAAACTTGCTAAATTTGCTTTGACTACTGATTC
ATTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGT
TTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTA
TTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGC
TCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCT
TTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAA
TTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGAT
TCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGT
ATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTG
AAAAACCTCATTTACAAGATGTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGTG
ATTGAATGATTGACTTATGAGATTACAAAATTACATCAGTCATACACTATAGGAACTCTCAGAT
GCATATGCATTACATTCTATTTTTAGAAGTACTACCATTAGGATTCTTGAACATTATATGGTCC
AAGGTCATATGTTAAGCTTCTCAGGTAATGCTCATACTAGTGAAAGATATTTTGATAACCCTGA
AGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCT
TCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCAT
TTAGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTT
AACTACTTTTGTTTATAACTTAA

Sequence length: 1431b

FASTA 2: *ERG11* Gene Nucleotide sequence for >17_*ERG11*-F

>17_*ERG11*-F

GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTACCAACAGATCAGTATA
TTATTAGTGGTACCTTTGTATACTACTGATCATGTGTTATTTATATTCATTAAGAAAAGATAGA
TCTCCATTATTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTT
ATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAA
AATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGAT
GTTTCTGTTGAAGAAGCTTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATG
ATTGTCCAAATTCTAGATTAATGGAACAAAAAAAACTTGCTAAATTTGCTTTGACTACTGATTC
ATTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGT
TTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTA
TTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGC
TCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCT
TTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAGAAAT
TAAACTGAGAAGAGAACGTGGTGATATTGATCCAATCGTGATTTAATTGATTCCTTATTGATTC
ATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTAT
TCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAA
AAACCTCATTTACAAGATGTATTTATCAGAAGTTGTTGAATTATTGAAAAAAAAAGGTGTGATT
TGAATGATTTGACTTATGAAGATTTACAAAATTACCATCAGTCATACACTATAGGAACTCTCAG
ATGCATATGCCATTACATTCTATTTTTAGAAGTACTACCATTAGATTCCTGAACATTATATTGA
TCCAAGGTCATAATGGATTAAGCTTCTTCCAGGTTATGTGCCTCATACTAGTGAAAGATATTTT
GATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTT
CATTTAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACC
TTATTTACCATTTGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTG
GGAACCATTTTAACTACTTTTGTTTATAACTTAAGATGGACTATTGATGGTTATAAAGTGCCTG
ACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAG
AGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGA

Sequence length: 1590b

FASTA 3: *ERG11* Gene Nucleotide sequence for >28_*ERG11*-F

>28_*ERG11*-F

GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGATGACCATATTATTACACCATATCAGTT
ATTATTAATGGTACCTTTGTATCTACTTATCTGTTGTTATTTATATTCTTCAGAAATGATTAAT
CTCCTCATGGCTTTATTCGACTTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTA
TGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAA
ATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATG
TTTCTGCTGAAGAAGCTTATAAACATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGA
TTGTCCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCA
TTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTT
TCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTAT
TTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCT
CAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTT
TACCTCATTATTGGAGACGTGATGCTGCTCAAAGAAAATCTCTGCTACTTATATGAAAGAAATT
AAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTC
ATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTAT
TCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAA
AAACCTCATTTACAAGATGTATTTATCAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGAT
TTGAATGATTTGACTTATGAGATTACAAAATACCATCAGTCAATACACTATACGAACTCTCAGA
TGCATATGCCATACTCTATTTTTAGAAGTACTATCATAGATCCTGAATCAATTATTGATCCAAA
GTCATTATGTTTAGTTCTTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTT
CTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATT
TGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTA
ACTACTTTTGTTTATAATTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATA
GTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTAT
GTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGA

Sequence length: I578b

FASTA 4: *ERG11* Gene Nucleotide sequence for >Dog_*ERG11*-F

>Dog_*ERG11*-F

GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTTTGATAATTGTATACACCATATCAGTA
TATTATTATTGGTTCCTCTTGTAATCTACTTATCTGTTGTTTTTTCTTTCTTCAGAAATGATTA
ATCTCCTTCATTGACATTATTCGACTTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAAC
CTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGG
GAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCT
GATGTTTCTGCTGAAGAAGCTTATAAACATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTT
ATGATTGTCCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGA
TTCATTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAA
AGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTA
CTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATT
TGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTA
CCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAG
AAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATT
GATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATT
GGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAG
GTGAAAAACCTCATTTACAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAAGAAAAAGG
TGTGATTTGAATGATTTGACTTATGAAGATTACAAAATACCATCAGTCATACACTATATGAACT
CTCAGAATGCATATGCCATACATCCTATTTTAGAAAAGTACTATCATAGATTCCTGAATCAATT
ATTTGATCCAAAGTCATTTATGGTTTAGTTTTCTCCCGTTATGCTCATACTAGTGAAAGATATT
TTGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGT
TTCATTTAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCA
CCTTATTTACCATTTGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAAT
TAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACTATTGATGGTTATAAAGTGCC
TGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAA
AGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTC
ATTTTGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAGA
ATAG

Sequence length: 1668b

3.5 Amino Acid Sequence (Primary Protein Structure) of the Translated *ERG11* Gene Nucleotide Sequence

Amino acid Sequence (Primary Protein Structure) translated from the *ERG11* gene nucleotide sequence of the various isolate showed the amino acid composition to vary. In this case, the identifier was assigned since the sequence is yet to be uploaded on NCBI. **01_*ERG11-F*** had 476 amino acids, **17_*ERG11-F*** had 529 amino acids, **28_*ERG11-F*** had 525 amino acids, and **Dog_*ERG11-F*** had 555 amino acids. The amino acid sequences are presented as FASTA 5, FASTA 6, FASTA 7, and FASTA 8.

FASTA 5: Amino Acid Sequence of Translated *ERG11* Gene Nucleotide Sequence for 01_*ERG11-F*

>01_*ERG11-F*

VETVIDGINYFLLTYCIRHISILLLVPLVIYYFVVFSLIQKKLISSSQFDILGLVLQLHM

VNNLMNFSNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNLFSMLNYLMFLLKKLISIILLQ

FSVQGLFMIVQILDDWNKKNLLNLLLLLIHLKDMFLRLEKKFFIILLLMKVSNNKKKLMG

LPMLLKLNQKLLFSLLQDLYLVMKKEEFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIG

DVMLLKRKSLLLIIKKLNNEENVVILIQIVIILIPYYFIQLIKMVVKKLIKKLLIFFLVF

LWVVNILLLLLLLGSCYIIVKNLIYKMYLSRSCCIIERKRCDDMIDLLDYKITSVIHYRN

SQMHMHYILFLEVLPLGFLNIIWSKVICCASQVMLILVKDILITLKILIQLDGILLLPKL

ILFHLTLLMKLIMGLGKFLKGFLHLIYHLVVVDIDVLGNNLLMFNWEPFFLLLFIT

Sequence length: 484

FASTA 6: Amino Acid Sequence of Translated ERG11 Gene Nucleotide Sequence for 17_*ERG11-F*

>17_*ERG11-F*

VETVIDGINYFLSLSVPTDQYIISGTFVYYYSCVIYIHHEKIDLHYCFIGFLGLVLQLHM

VNNLMNFSNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNLFSMLNYLMFLLKKLISIILLQ

FSVQGLFMIVQILDDWNKKNLLNLLLLLIHLKDMFLRLEKKFFIILLLMKVSNNKKKLMG

LPMLLKLNQKLLFSLLQDLYLVMKKEEFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIG

DVMLLKRKSLLLIIRNNTEKRTWWYYSNRDLIDSLLIHSTYKDGVKMTDQEIANLLIGIL

MGGQHTSASTSAWFLLHLGEKPHLQDVFIRSCCIIEKKRCDLNDLTYEDLQNYHQSYTIG

TLRCICHYILFLEVLPLDSSTLYYSKVIMDDASSRLCASYYYKIFFFPPRFFSNNMGYCC

CQSSFCFIILFFFSSLWVWESFFRGFFTLFTIWWWWTTMYWGTICLCSIGNHFNYFCLLL

KMDYYWLLSAAPPLLFNGGFTYYTSRNHLGKKRNLYVLIKRQLSFDSVF

Sequence length: 537

FASTA 7: Amino Acid Sequence of Translated ERG11 Gene Nucleotide Sequence for 28_*ERG11*-F

>28_*ERG11-F*

VETVIDGINYFLMTILLHHISYYYWYLCIYLSVVIYILQKKLISSWLYSTSLVWFCSFIW

STTLLIFRIMSSKVWWCIFIYVIRENYDGLFRSKRSSICFQCCIIICFCCRSLLTFNYSS

FRYRGYLLLSKFQINGTKKICCICFDYYFIIKICSSDDRRNFELFCYYYKFQIERKNSWG

CQCYENSTRNYYFHCFKIFIWWWNEKNFFPFICSTIFFFRRRFYPYYFCFPPFTFTSLLE

TTCCSKKISATYMKEIKLRRERGDIDPNRDLIDSLLIHSTYKDGVKMTDQEIANLLIGIL

MGGQHTSASTSAWFLLHLGEKPHLQDVFIRSCCIIERKRWWFEEFDLLDYKIPSVNTLYE

LSDAYAILYFFKYYHRSSINYYSKVIMFSSSRLCSYYYKIFFFPPRFFSNNMGYCCCQSS

FCFIILFFFSSLWVWESFFRGFFTLFTIWWWWTMMYWGTICLCSIRNHFNYFCLLFKMDY

YWLLSAAPPLLFNGGFTYYTSRNHLGKKRNLYVLIKRQLSFDSVF

Sequence length: 534

FASTA 8: Amino Acid Sequence of Translated ERG11 Gene Nucleotide Sequence for Dog_*ERG11*-F

>dog_*ERG11-F*

VETVIDGINYFLFDNCIHHISILLLVPLVIYLSVVFSFFRNDDSPSLTLFDFLGLVLQLH

MVNNLMNFSNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNLFSMLNYLMFLLKKLINIILL

QFSVQGLFMIVQIPDDWNKKNLLNLLLLLIHLKDMFLRLEKKFFIILLLMKVSNNKKKLM

GLPMLLKLNQKLLFSLLQDLYLVMKKEEFLTVHLLNYILIIIKVLPLLILFSLIYLYLII

GDVMLLKRKSLLLIIKKLNNEENVVILIQIVIILIPYYFIQLIKMVVKKLIKKLLIFFLV

FLWVVNILLLLLLLGSCYIIVKNLIYRCYLSRSCCIIEKKKVVFEEFDLLRLQNTISHTL

YELSECICHTSYFRKVLSSIPESIIISKVIYGLVFSRYAHTSERYFDNPEDFDPTRWDTA

AAKANSVSFNSSDEVDYGFGKVSKGVSSPYLPFGGGRHRCIGEQFAYVQLGTILTTFVYN

LRWTIDGYKVPDPDYSSMVVLPTEPAEIIWEKRETCMFFFNGNFLSIQCSDCFHFVTTLD

DHIYTYTYKYMIHIE

Sequence length: 565

3.6 Physical and Chemical Characteristics/ Subcellular Localization of the Proteins

The amino acid sequence of the predicted *Candida ERG11* protein for 01*_ERG11*-F, 17*_ERG11*-F, 28*_ERG11*-F and Dog*_ERG11*-F contains 476, 529, 525, 555 amino acid residues, respectively, with a molecular weight of 56183.10Da, 62897.16Da, 64159.15Da, 65139.25Da and a predicted pI (pH at which protein carries no net electric charge) of 9.60, 9.03, 9.03 and 8.88, respectively. The estimated half-live of the proteins in mammalian cells is 100 hours, and the instability coefficients (stability in the test tube) are 35.70, 36.98, 44.50 and 37.69, respectively. 01*_ERG11*-F, 17*_ERG11*-F and Dog*_ERG11*-F are stable enzymes while the 28*_ERG11*-F is an unstable enzyme. The grand average of hydrophobicity (GRAVY) of the *Candida ERG11* protein is 0.195, 0.478, 0.195 and 0.576, respectively. All the *Candida ERG11* protein were predicted to be localized in the plasma membrane.

3.7 Comparative Analysis of the *ERG11* gene Nucleotides

Multiple Sequence Alignment (MSA) was used to compare the nucleotide sequence variations in the *ERG11* gene of the isolates. The sequence alignment in Figures 1a, 1b, 1c and 1d are targeted at revealing the *Candida ERG11* protein conserved domains and variable sites in both human and the dog infecting species. The graphical annotation of the MSA of the nucleotide sequences is shown in Figure 2. The red bars denotes significant homology. The Percentage Identity Matrix - created by Clustal 2.1, and the sequences significant alignments measured as Max score, Total score, Percentage Query cover, E value, and percentage Ident, are shown in Tables 4 and 5.

```
CLUSTAL O(1.2.4) multiple sequence alignment


Sample_17_ErgII-F       GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTA--CCAACAGATC    58
Sample_01_ErgII-F       GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTTGACATATTGTATACGACATATC    60
Sample_28_ErgII-F       GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGATGACCATATTATTACACCATATC    60
Sample_Dog_ErgII-F      GTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTTTGATAATTGTATACACCATATC    60
                        ************************************              *  ** ***

Sample_17_ErgII-F       AGTATATTATTAGTGGTACCTTTGTAT-ACTACTGATCATGTGTTATTTATATTCATTAA    117
Sample_01_ErgII-F       AGTATATTATTATTGGTTCCTCTTGTAATCTACTGATTTGTTG-TTTTTTCTTTGATTCA    119
Sample_28_ErgII-F       AGTTATTATTAATGGTACCT---TTGTATCTACTTATCTGTTG-TTATTTATATTCTTCA    116
Sample_Dog_ErgII-F      AGTATATTATTATTGGTTCCTCTTGTAATCTACTTATCTGTTG-TTTTTT-CTTTCTTCA    118
                        ***   *  * *  *   *          ***** **    ** *  **    *  ** *

Sample_17_ErgII-F       GAAAAGATAGATC--TCCATTATTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTC    175
Sample_01_ErgII-F       GAAATAATTAATCTCCTC----ATGACAGTTTGACATCCTTGGTTTGGTTCTGCAGCTTC    175
Sample_28_ErgII-F       GAAATGATTAAT--CTCCTCATGGCTTTATTCGACTTCCTTGGTTTGGTTCTGCAGCTTC    174
Sample_Dog_ErgII-F      GAAATGATTAATCTCCTTCATTGACATTATTCGACTTCCTTGGTTTGGTTCTGCAGCTTC    178
                        ****  **  **                * *   ************************

Sample_17_ErgII-F       ATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT    235
Sample_01_ErgII-F       ATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT    235
Sample_28_ErgII-F       ATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT    234
Sample_Dog_ErgII-F      ATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT    238
                        ************************************************************

Sample_17_ErgII-F       TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATT    295
Sample_01_ErgII-F       TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATT    295
Sample_28_ErgII-F       TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATT    294
Sample_Dog_ErgII-F      TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATT    298
                        ************************************************************

Sample_17_ErgII-F       TGTTTTCAATGCTAAATTATCTGATGTTTCTGTTGAAGAAGCTTATAAGCATTTAACTAC    355
Sample_01_ErgII-F       TGTTTTCAATGCTAAATTATCTGATGTTTCTGTTGAAGAAGCTTATAAGCATTTAACTAC    355
Sample_28_ErgII-F       TGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGCTTATAAACATTTAACTAC    354
Sample_Dog_ErgII-F      TGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGCTTATAAACATTTAACTAC    358
                        ******************************** *************** ***********

Sample_17_ErgII-F       TCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATGGAACAAAA    415
Sample_01_ErgII-F       TCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATGGAACAAAA    415
Sample_28_ErgII-F       TCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATGGAACAAAA    414
Sample_Dog_ErgII-F      TCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATGGAACAAAA    418
                        ****************************************** *****************
```

Figure 1a: Multiple sequence alignment of the nucleotide sequences (1 to ~400) of the *ERG11* gene of the isolates

```
Sample_17_ErgII-F       AAAACTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAG    475
Sample_01_ErgII-F       AAAACTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAG    475
Sample_28_ErgII-F       AAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAG    474
Sample_Dog_ErgII-F      AAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAG    478
                        **** **************************************************

Sample_17_ErgII-F       AGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCA    535
Sample_01_ErgII-F       AGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCA    535
Sample_28_ErgII-F       AGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCA    534
Sample_Dog_ErgII-F      AGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCA    538
                        ************************************************************

Sample_17_ErgII-F       TGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATC    595
Sample_01_ErgII-F       TGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATC    595
Sample_28_ErgII-F       TGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATC    594
Sample_Dog_ErgII-F      TGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATC    598
                        ************************************************************

Sample_17_ErgII-F       TTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGA    655
Sample_01_ErgII-F       TTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGA    655
Sample_28_ErgII-F       TTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGA    654
Sample_Dog_ErgII-F      TTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGA    658
                        ********************************************** **********

Sample_17_ErgII-F       TTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTA    715
Sample_01_ErgII-F       TTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTA    715
Sample_28_ErgII-F       TTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTA    714
Sample_Dog_ErgII-F      TTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTA    718
                        ************************************************************

Sample_17_ErgII-F       TTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAGAA-ATTAAACT    774
Sample_01_ErgII-F       TTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACT    775
Sample_28_ErgII-F       TTGGAGACGTGATGCTGCTCAAA-GAAAATCTCTGCTACTTATATGAAAGAAATTAAACT    773
Sample_Dog_ErgII-F      TTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACT    778
                        ********************** ********************** * ********

Sample_17_ErgII-F       GAGAAGAGAACGTGGTGATATTGATC-CAATCGTGATTTAATTGATTCCTTATTGATTCA    833
Sample_01_ErgII-F       GAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCA    835
Sample_28_ErgII-F       GAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCA    833
Sample_Dog_ErgII-F      GAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCA    838
                        *************************  ********************************
```

Figure 1b: Multiple sequence alignment of the nucleotide sequences (~470 to ~800) of the *ERG11* gene of the isolates

```
Sample_17_ErgII-F       TTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGG    893
Sample_01_ErgII-F       TTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGG    895
Sample_28_ErgII-F       TTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGG    893
Sample_Dog_ErgII-F      TTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGG    898
                        ************************************************************

Sample_17_ErgII-F       TATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTT    953
Sample_01_ErgII-F       TATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTT    955
Sample_28_ErgII-F       TATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTT    953
Sample_Dog_ErgII-F      TATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTT    958
                        ************************************************************

Sample_17_ErgII-F       AGGTGAAAAACCTCATTTACAAGATGTATTTATCAGAAGTTGTTGA-ATTATTGAAAAAA    1012
Sample_01_ErgII-F       AGGTGAAAAACCTCATTTACAAGATGTATTTATCAAGAAGTTGTTGAATTATTGAAAGAA    1015
Sample_28_ErgII-F       AGGTGAAAAACCTCATTTACAAGATGTATTTATCA-GAAGTTGTTGAATTATTGAAAGAA    1012
Sample_Dog_ErgII-F      AGGTGAAAAACCTCATTTACAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAAGA    1018
                        ********************     *********  *  *  *   **********  *

Sample_17_ErgII-F       AAAGGTGT-GATTTGAATGATTTGACTTATGAAGATTTACAAAATTACCATCAGTCATAC    1071
Sample_01_ErgII-F       AAAGGTGTGATTGAA---TGATTGACTTAT--GAGATTACA-AAATTACATCAGTCATAC    1069
Sample_28_ErgII-F       AAAGGTGGTGATTTGAATGATTTGACTTAT--GAGATTACAAAATACCATCAGTCAATAC    1070
Sample_Dog_ErgII-F      AAAAGGTGTGATTTGAATGATTTGACTTAT--GAAGATTACAAAATACCATCAGTCATAC    1076
                        *** *      *         *********       *    **           ****

Sample_17_ErgII-F       ACTATAGGAACTCTCAGATGCATATGCCATTACATTCTATTTTTA-GAAGTACTACCATT    1130
Sample_01_ErgII-F       ACTATAGGAACTCTCAGATG-CATATGCATTACATTCTATTTTTA-GAAGTACTACCATT    1127
Sample_28_ErgII-F       ACTATACGAACTCTCAGAT-GCATATGCCATACTCTATTTTTA---GAAGTACTATCATA    1126
Sample_Dog_ErgII-F      ACTATATGAACTCTCAGAATGCATATGCCATACATCCTATTTTAGAAAAGTACTATCATA    1136
                        ****** ***********        *  ***    * ***     ******** ***

Sample_17_ErgII-F       AGATTCCTGAACATTATATTGATCCAAGGTCATAATGGATTAAGCTTCTTCCAGGTTATG    1190
Sample_01_ErgII-F       AGGATTCTTGAACATTATATGGTCC----AAGGTCATAT----GTTAAGCTTCTCAGGTA    1179
Sample_28_ErgII-F       GATC-CTGAATCAATTAT-TGATCCA--AAGTCATTATG----TTTAGTTCTTCCAGGTT    1178
Sample_Dog_ErgII-F      GATTCCTGAATCAATTATTTGATCCAAAGTCATTTATGG----TTTAGTTTTCTCCCGTT    1192
                                *    ** ***                    *           *

Sample_17_ErgII-F       TGCCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGG    1250
Sample_01_ErgII-F       ATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGG    1239
Sample_28_ErgII-F       ATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGG    1238
Sample_Dog_ErgII-F      ATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGG    1252
                          ************************************************************
```

Figure 1c: Multiple sequence alignment of the nucleotide sequences (~800 to ~1200) of the *ERG11* gene of the isolates

```
Sample_17_ErgII-F       ATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTATG    1310
Sample_01_ErgII-F       ATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTATG    1299
Sample_28_ErgII-F       ATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTATG    1298
Sample_Dog_ErgII-F      ATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTATG    1312
                        ************************************************************

Sample_17_ErgII-F       GGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGAC    1370
Sample_01_ErgII-F       GGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAGAC    1359
Sample_28_ErgII-F       GGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGAC    1358
Sample_Dog_ErgII-F      GGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGAC    1372
                        ********************************************* ************

Sample_17_ErgII-F       ATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTG    1430
Sample_01_ErgII-F       ATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTG    1419
Sample_28_ErgII-F       ATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTG    1418
Sample_Dog_ErgII-F      ATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTG    1432
                        ********************************** ***********************

Sample_17_ErgII-F       TTTATAACTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAA    1490
Sample_01_ErgII-F       TTTATAACTTAA------------------------------------------------    1431
Sample_28_ErgII-F       TTTATAATTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAA    1478
Sample_Dog_ErgII-F      TTTATAATTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAA    1492
                        ******* ****

Sample_17_ErgII-F       TGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGT    1550
Sample_01_ErgII-F       ------------------------------------------------------------    1431
Sample_28_ErgII-F       TGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGT    1538
Sample_Dog_ErgII-F      TGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGT    1552


Sample_17_ErgII-F       TTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGA--------------------    1590
Sample_01_ErgII-F       ------------------------------------------------------------    1431
Sample_28_ErgII-F       TTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGA--------------------    1578
Sample_Dog_ErgII-F      TTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTGTTACTT    1612


Sample_17_ErgII-F       ------------------------------------------------------        1590
Sample_01_ErgII-F       ------------------------------------------------------        1431
Sample_28_ErgII-F       ------------------------------------------------------        1578
Sample_Dog_ErgII-F      AGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAG       1668
```

Figure 1d: Multiple sequence alignment of the nucleotide sequences (~1300 to end) of the *ERG11* gene of the isolates

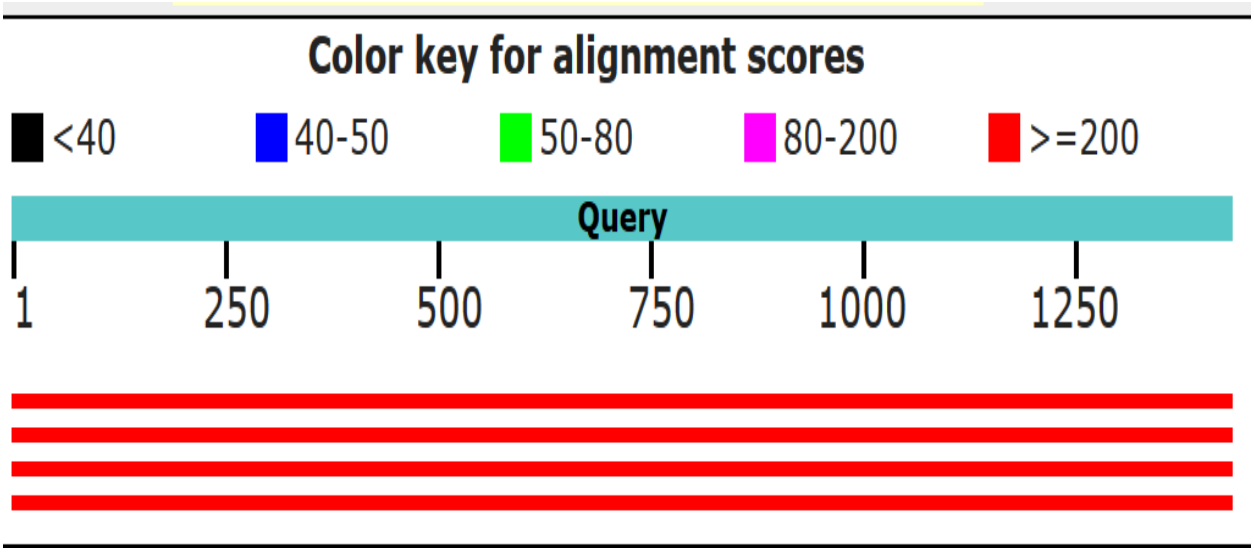


Figure 2: Representation of the multiple sequence alignment of the nucleotide sequences of the *ERG11* gene of the isolates

Table 4: Percentage nucleotide Identity Matrix of the four isolates

	01_*ERG11*-F	017_*ERG11*-F	28_*ERG11*-F	Dog_*ERG11*-F
01_*ERG11*-F	100.00%	91.56%	92.02%	93.15%
17_*ERG11*-F	91.56%	100.00%	91.04%	91.47%
28_*ERG11*-F	91.02%	92.04%	100.00%	94.55%
Dog_*ERG11*-F	93.15%	91.47%	94.55%	100.00%

Table 5: Sequences significance nucleotide alignments, measured by max score, total score, percentage query cover, E value, and percentage Ident

Description	Max score	Total score	Query cover	E value	Ident	Query Accession
01_*ERG11*-F ↕ **17**_*ERG11*-F	2264	2426	100%	0.0	96%	Query_91717
01_*ERG11*-F ↕ **28**_*ERG11*-F	2242	2408	100%	0.0	95%	Query_91718
01_ErgII-F ↕ Dog_*ERG11*-F	2273	2456	100%	0.0	96%	Query_91719
17_*ERG11*-F ↕ **28**_*ERG11*-F	2515	2647	100%	0.0	96%	Query_122320
17_*ERG11*-F ↕ **Dog**_*ERG11*-F	2473	2604	100%	0.0	95%	Query_122321
28_*ERG11*-F ↕ **Dog**_*ERG11*-F	2599	2733	100%	0.0	97%	Query_28305

3.8 Comparative Analysis of the Amino Acid Sequences of the Translated *ERG11* Gene Nucleotide Sequences

Amino acid Multiple Sequence Alignment (MSA) was used to compare the amino acid sequence variations of the *ERG11* gene of the isolates. The sequence alignment shown in Figure 3 are targeted at revealing the *Candida ERG11* protein conserved domains and variable sites in both the human- and the dog-infecting *Candida* species. The amino acid MSA of the four isolates of *Candida* in this study revealed numerous variable sites with two major conserved domains. The first 12 amino acids in the sequences make up the first major conserved domain. Due to the differential length of each of the *Candida ERG11* protein sequences used for this study, the position of the second conserved domain varies in each of the sequences. The amino acid sequence making up the second major conserved region in 01_*ERG11*-F, 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F are in positions 331-337, 332-338, 330-336 and 330-336, respectively. The graphical annotation of the MSA of the amino acid sequence of the *ERG11* gene of the isolates is shown in Figure 4. Red bars express significant homology, while black bars represent non-significant homology. The Percentage Identity Matrix - created by Clustal 2.1, and the amino acid sequences significant alignments measured as Max score, Total score, Percentage Query cover, E value, and percentage Ident are shown in Tables 6 and 7.

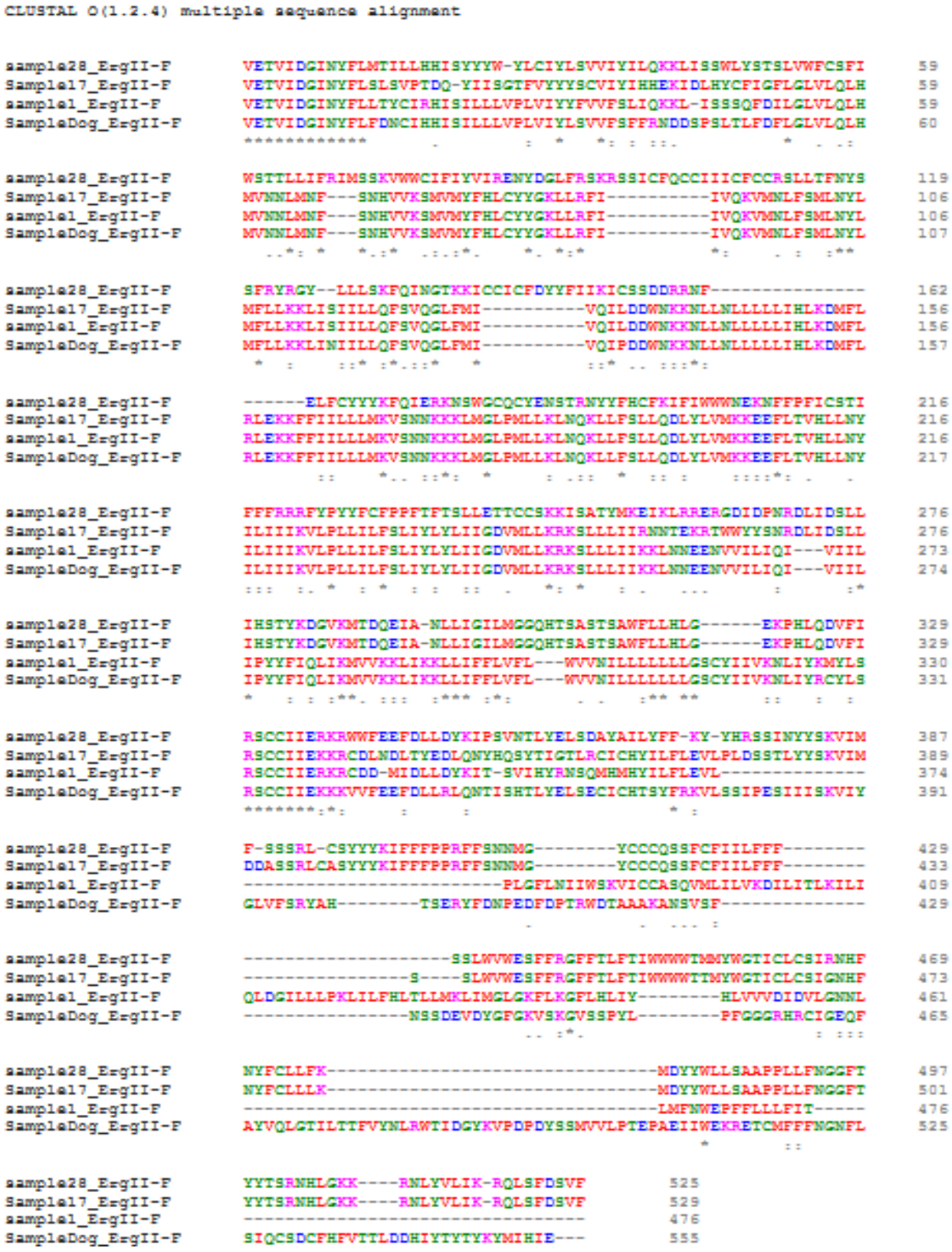


Figure 3: Multiple sequence alignment of the translated amino acid sequences (1 to end) of the *ERG11* gene of the isolates

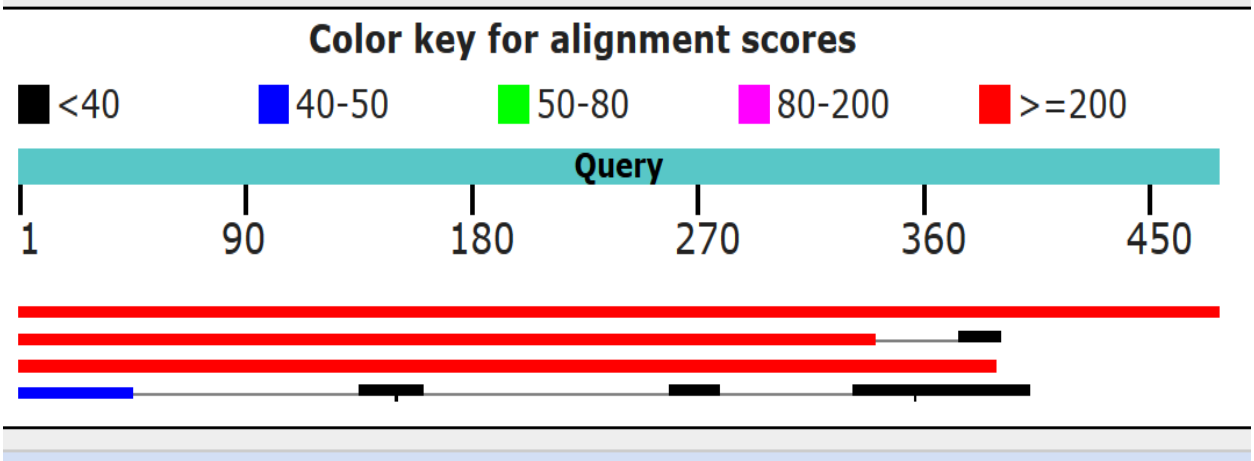


Figure 4: Representation of the multiple sequence alignment of the Amino acid sequence of the *ERG11* gene of the isolates

Table 6: Amino acid alignment Percentage Identity Matrix of the four isolates

	01_*ERG11*-F	17_*ERG11*-F	28_*ERG11*-F	Dog_*EG11*-F
01_*ERG11*-F	100.00%	61.34%	24.32%	76.68%
17_*ERG11*-F	61.34%	100.00%	52.20%	54.82%
28_*ERG11*-F	24.32%	52.20%	100.00%	23.78%
Dog_*ERG11*-F	76.68%	54.82%	23.78%	100.00%

Table 7: Sequences significance amino acid alignments measured by max score, total score, percentage query cover, E value, and percentage Ident

Description	Max score	Total score	Query cover	E value	Ident	Query Accession
01_*ERG11*-F ↕ **17**_*ERG11*-F	385	385	81%	2e-133	65%	Query_150085
01_*ERG11*-F ↕ **28**_*ERG11*-F	43.5	155	34%	3e-08	61%	Query_150086
01_ErgII-F ↕ **Dog**_*ERG11*-F	510	530	75%	0.0	93%	Query_150087
17_*ERG11*-F ↕ **28**_*ERG11*-F	427	541	58%	2e-149	83%	Query_66352
17_*ERG11*-F ↕ **Dog**_*ERG11*-F	358	358	74%	5e-122	64%	Query_66353
28_*ERG11*-F ↕ **Dog**_*ERG11*-F	44.7	98.6	14%	5e-09	62%	Query_164643

3.9 Amino Acid Composition of the Translated *ERG11* Proteins

All the translated *ERG11* amino acid sequences have varying composition in amino acids as shown in Figure 5. The result showed a low concentration of amino acids with hydrophilic side chains in all species of the *Candida ERG11* gene sequenced. These amino acids include Arginine, Asparagine, Aspartate, Glutamine, Glutamate, Histidine, Lysine and Serine.

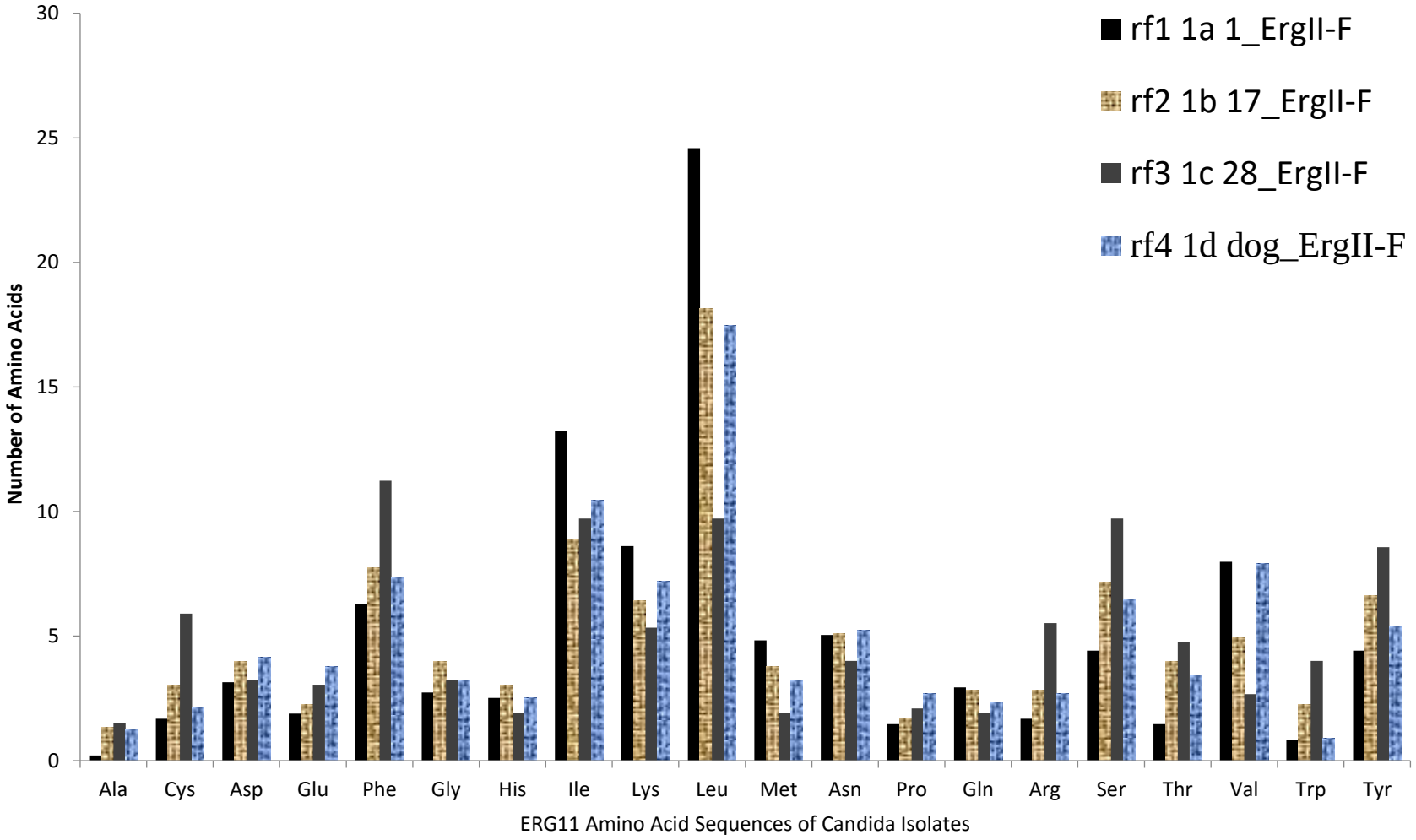
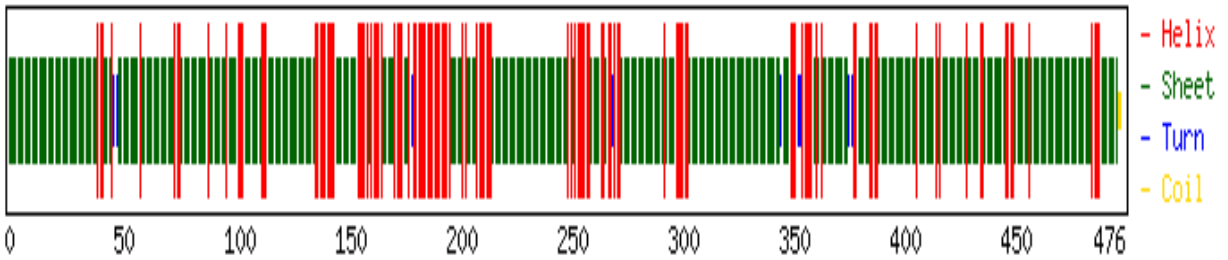


Figure 5: Comparison of the amino acid composition of the *ERG11* proteins sequences

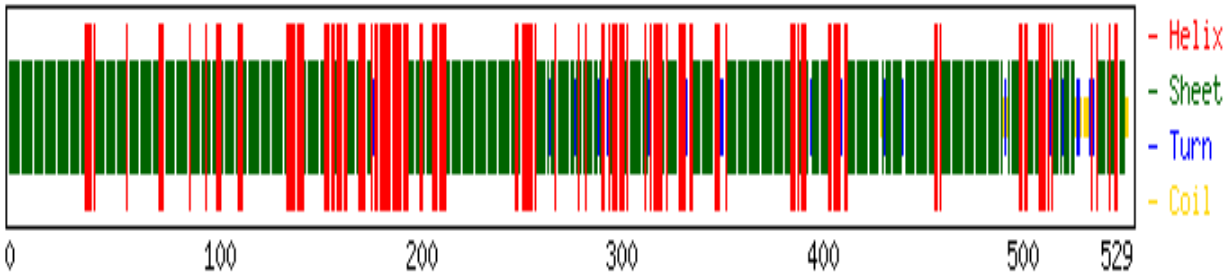
3.10 Secondary Structure of *ERG11* Gene Protein of the Isolates Sequenced

In the result of the secondary structure prediction shown in Figures 6, 7, 8 and 9, the total residues (expressed in percentage) of the *ERG11* gene alpha helix, beta pleated sheets, and turns vary. Regions of α-helices, β-sheets, turns and coils are shown in red, green, blue and yellow, respectively.



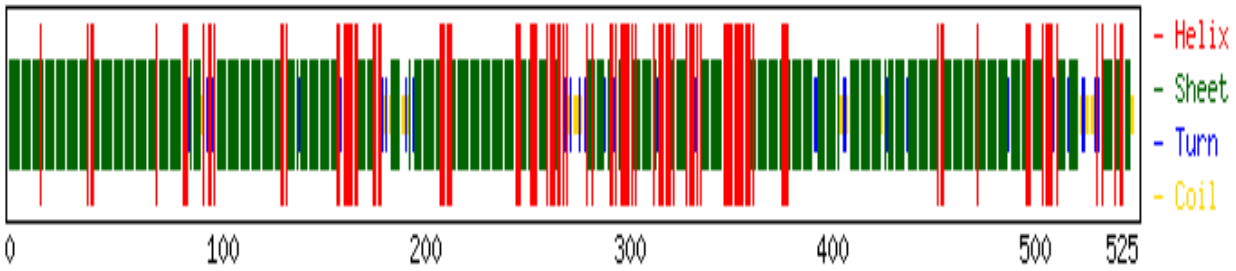
Total residues:
Helix: 441, Sheet: 442, Turn: 25.
Percentage (%):
Helix: 92.6, Sheet: 92.9, Turn: 5.3.

Figure 6: Predicted protein secondary structure of *ERG11* gene for 01_*ERG11*-F



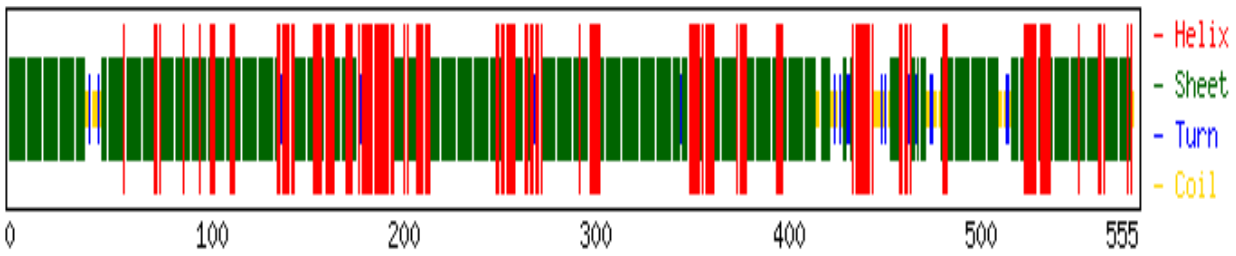
Total residues:
Helix: 385, Sheet: 452, Turn: 43.
Percentage (%):
Helix: 72.8, Sheet: 85.4, Turn: 9.1.

Figure 7: Predicted protein secondary structure of *ERG11* gene for 17_*ERG11*-F



Total residues:
Helix: 347, Sheet: 427, Turn: 48.
Percentage (%):
Helix: 66.1, Sheet: 81.3, Turn: 9.1.

Figure 8: Predicted protein secondary structure of *ERG11* gene for 28_*ERG11*-F



Total residues:
Helix: 439, Sheet: 457, Turn: 42.
Percentage (%):
Helix: 79.1, Sheet: 82.3, Turn: 7.6.

Figure 9: Predicted protein secondary structure of *ERG11* gene for Dog_*ERG11*-F

3.11 Tertiary Structure of the *ERG11* Gene Protein of the Four Isolates Sequenced

The tertiary structures were predicted by PHYRE2 using the amino acid residues in each *Candida ERG11* gene sequence. Modelling was done with a 100% confidence by the single highest scoring template. The results shown in Figures 10, 11, 12 and 13 indicated a low level of similarity between the 17_*ERG11*-F and 28_*ERG11*-F structural models, while 01_*ERG11*-F show a significant similarity to none. The models differ widely in complexity and size.

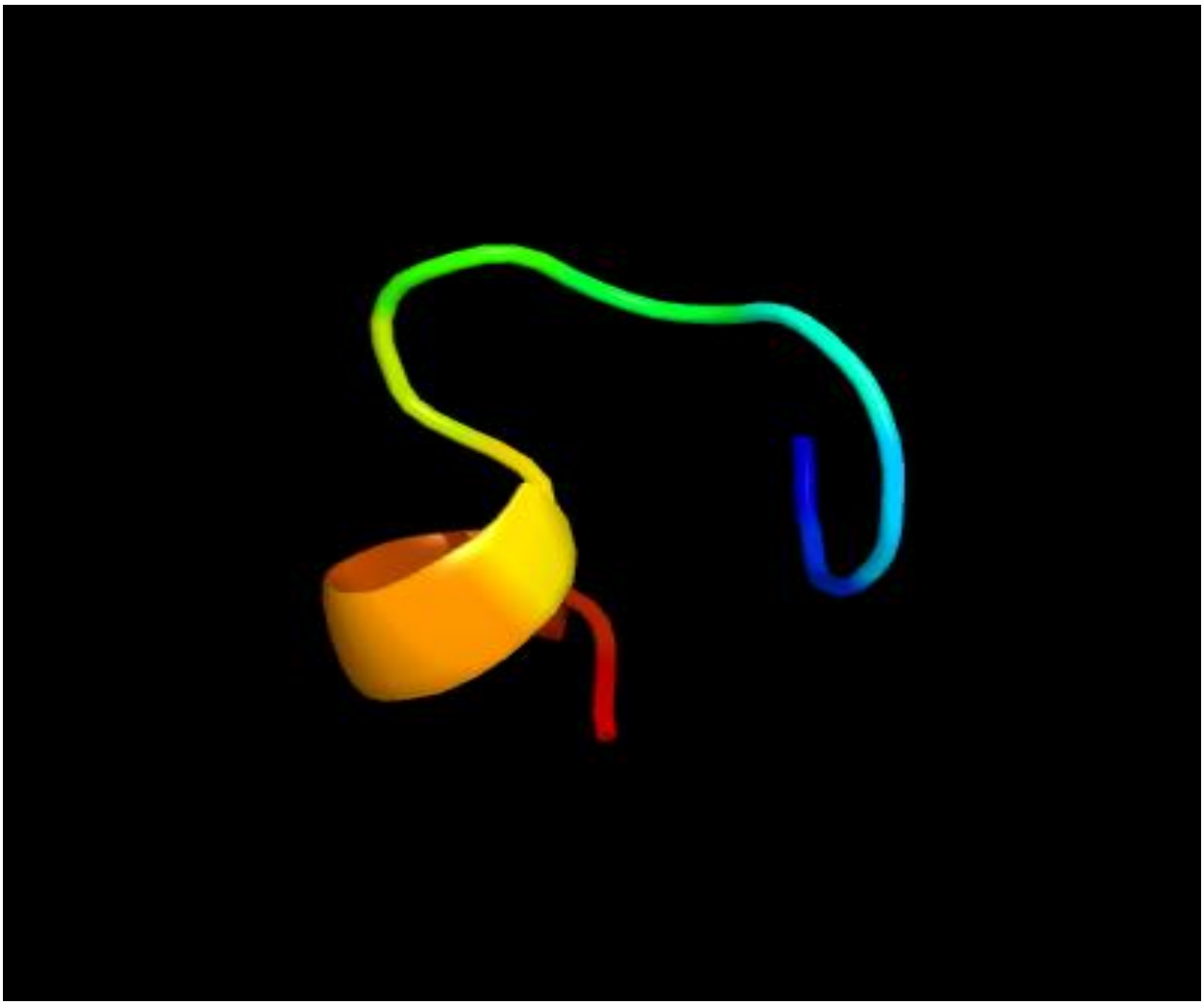


Figure 10: Tertiary Structure of the *ERG11* gene protein of 01_*ERG11*-F

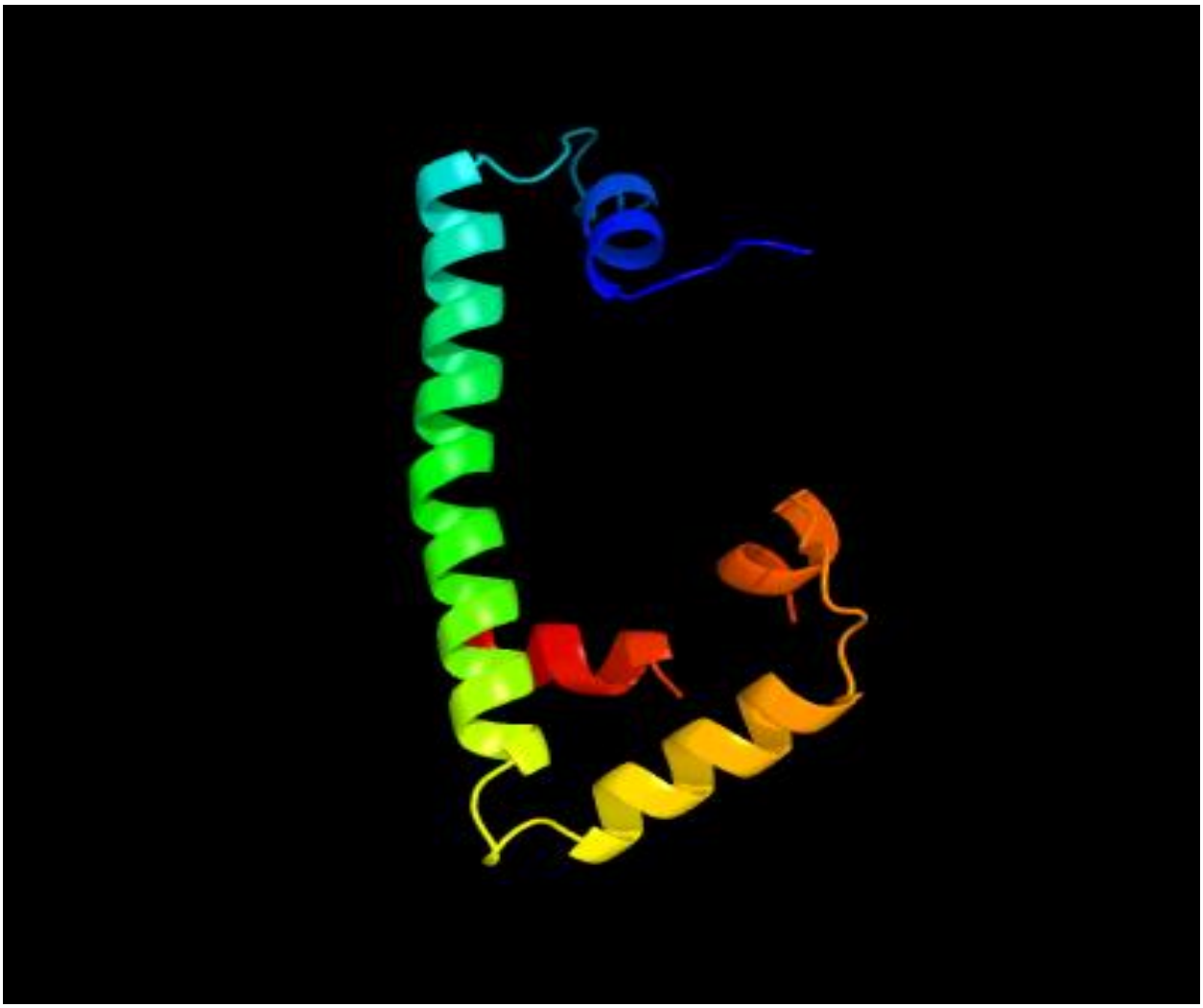


Figure 11: Tertiary structure of the *ERG11* gene protein of 17_*ERG11*-F

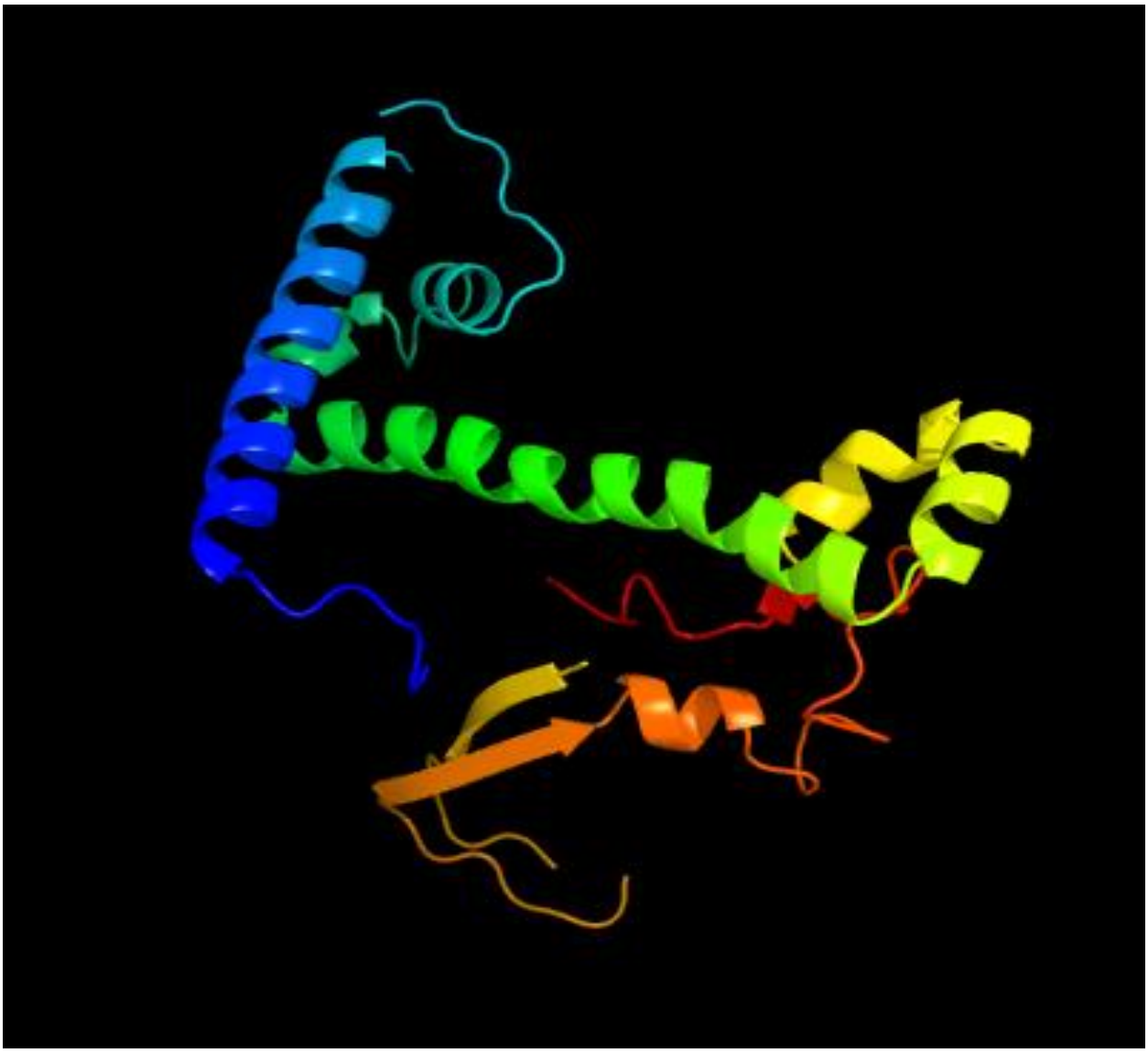


Figure 12: Tertiary structure of the *ERG11* gene protein of 28_*ERG11*-F

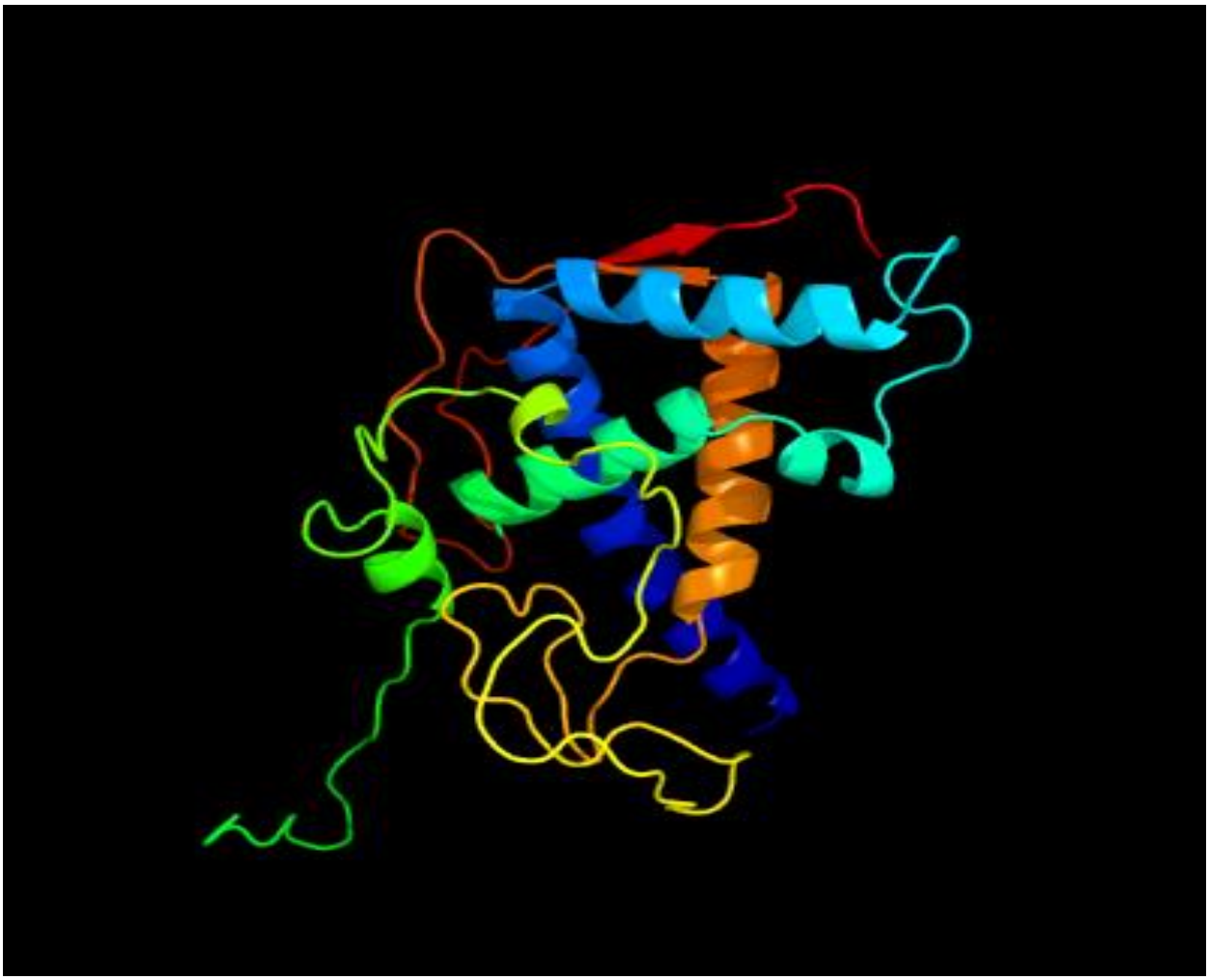
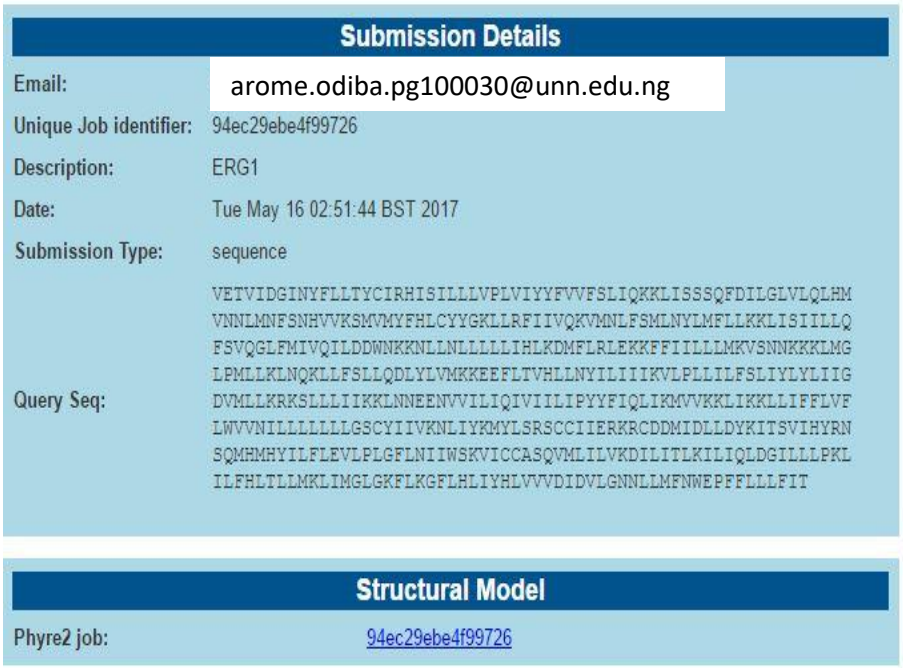


Figure 13: Tertiary structure of the *ERG11* gene protein of Dog_*ERG11*-F

3.12 *ERG11* Protein Ligand Binding Sites for the Four Isolates Sequenced

The 3D ligand automated software predicted the active sites of *Candida ERG11* protein with metallic heterogens to which drugs can bind. The ligand/drug binding sites of the *ERG11* proteins, and the associated amino acid sequences are shown in Figures 14, 15, 16, and 17. These sites have been reported with the observation that 01_*ERG11*-F has no single drug binding site. 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F however, exhibit unique binding sites to which drug can effectively bind.



Submission Details

Email: arome.odiba.pg100030@unn.edu.ng

Unique Job identifier: 94ec29ebe4f99726

Description: ERG1

Date: Tue May 16 02:51:44 BST 2017

Submission Type: sequence

Query Seq:

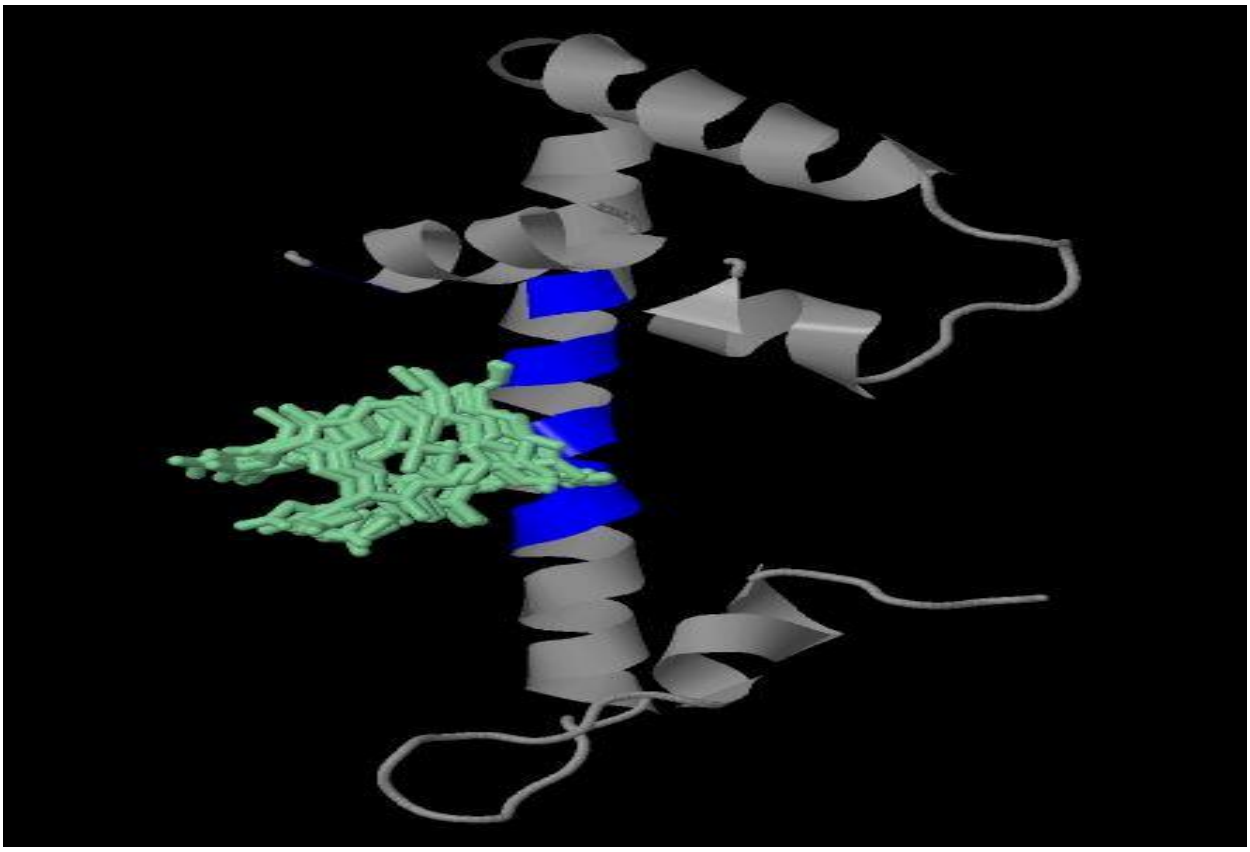
```
VETVIDGINYFLLTYCIRHISILLLVPLVIYYFVVFSLIQKKLISSSQFDILGLVLQLHM
VNNLMNFSNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNLFSMLNYLMFLLKKLISIILLQ
FSVQGLFMIVQILDDWNKKNLLNLLLLLIHLKDMFLRLEKKFFIILLLMKVSNNKKKLMG
LPMLLKLNQKLLFSLLQDLYLVMKKEEFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIG
DVMLLKRKSLLLIIKKLNNEENVVILIQIVIILIPYYFIQLIKMVVKKLIKKLLIFFLVF
LWVVNILLLLLLLGSCYIIVKNLIYKMYLSRSCCIIERKRCDDMIDLLDYKITSVIHYRN
SQMHMHYILFLEVLPLGFLNIIWSKVICCASQVMLILVKDILITLKILIQLDGILLLPKL
ILFHLTLLMKLIMGLGKFLKGFLHLIYHLVVVDIDVLGNNLLMFNWEPFFLLLFIT
```

Structural Model

Phyre2 job: 94ec29ebe4f99726

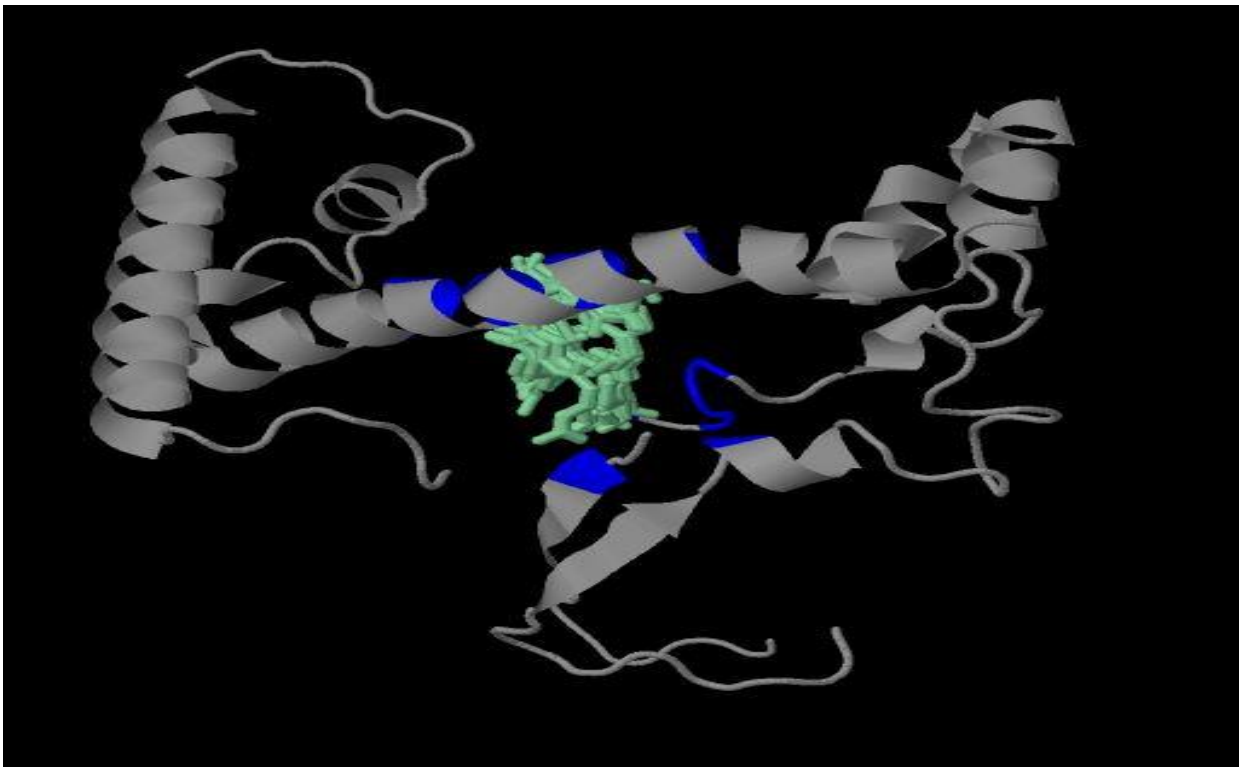
No prediction made as there were insufficient homologous structures with ligands bound identified

Figure 14: *ERG11* Protein ligand binding sites of 01*_ERG11*-F



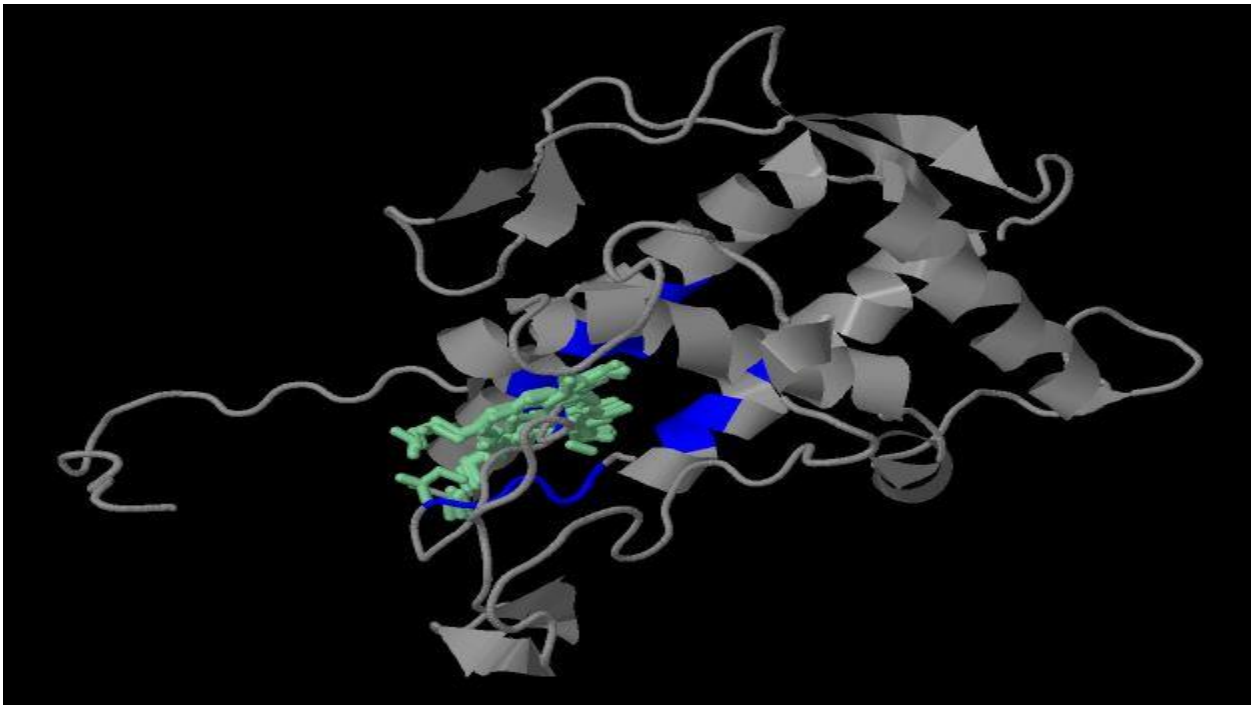
Residue	Amino acid
298	GLY
299	ILE
300	LEU
302	GLY
303	GLY
306	THR
307	SER
310	THR
362	LEU

Figure 15: *ERG11* Protein ligand binding sites of 17*_ERG11*-F



Residue	Amino acid
298	GLY
299	ILE
302	GLY
303	GLY
306	THR
307	SER
310	THR
347	LEU
371	PHE
400	ILE
401	PHE
402	PHE
403	PHE
405	PRO
406	ARG

Figure 16: *ERG11* Protein ligand binding sites of 28_*ERG11-F*

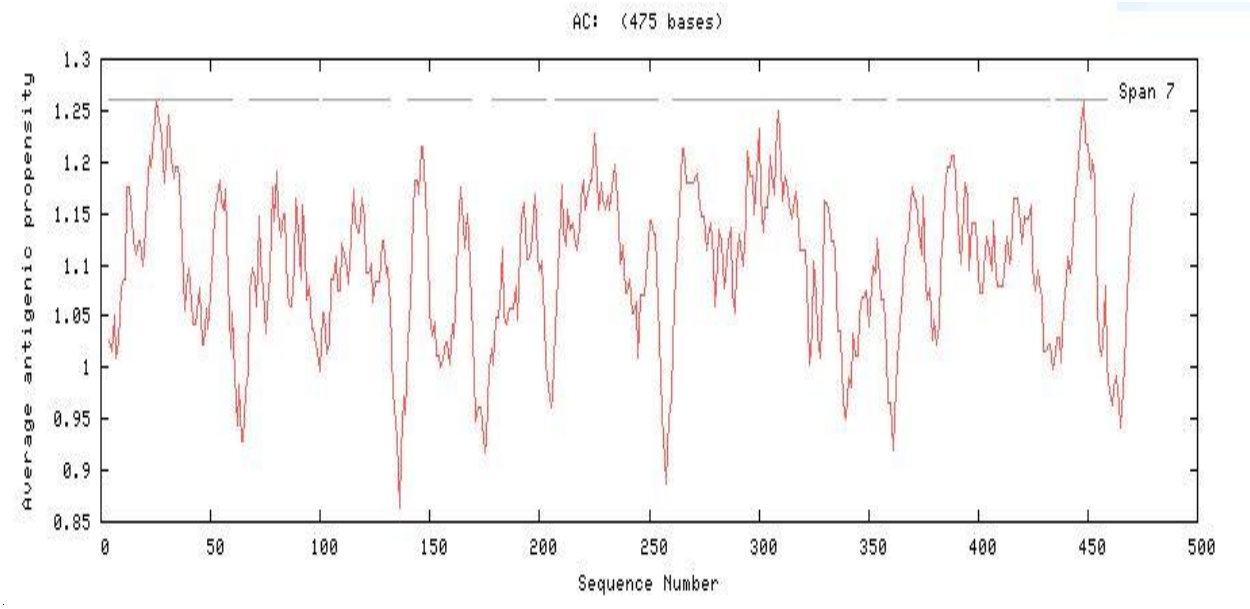


Residue	Amino acid
311	LEU
314	LEU
315	GLY
318	TYR
319	ILE
322	LYS
377	LEU
452	PRO
458	HIS
459	ARG
460	CYS
461	ILE
465	PHE
466	ALA
470	LEU

Figure 17: *ERG11* Protein ligand binding sites of Dog_*ERG11*-F

3.12 Antigenic Sites of the *ERG11* Protein of the Four Isolates Sequenced

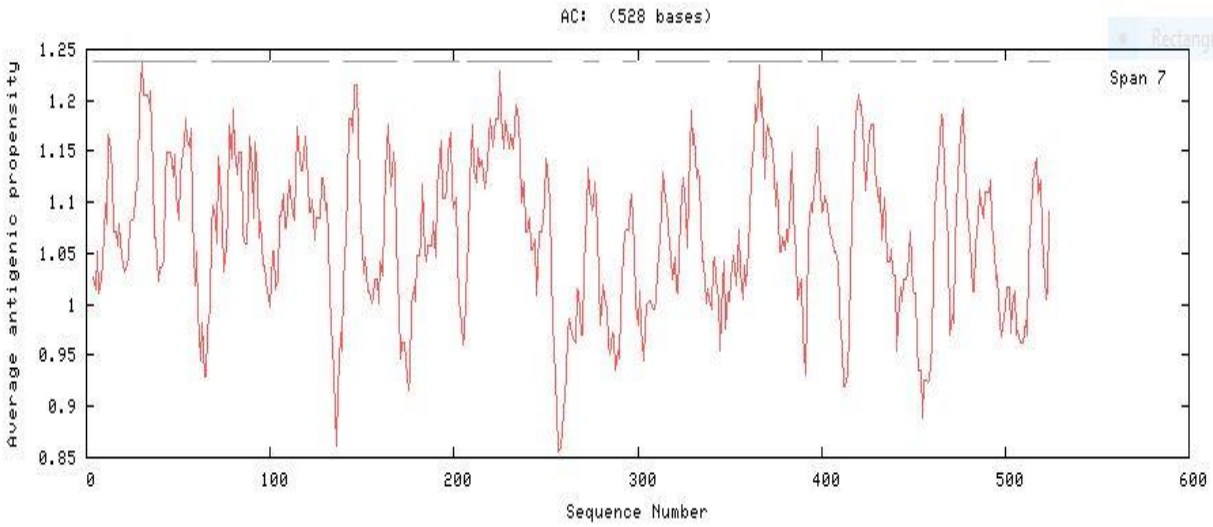
The chromatogram of the antigenic sites of the *ERG11* protein of 01_*ERG11*-F, 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F are shown in Figures 18, 19, 20 and 21. Several antigenic sites (determinants) were predicted for each of the isolates. The amino acids in the peaks also accompany the chromatogram. The 01_*ERG11*-F possesses 10 antigenicity sites (the lowest when compared to others), while 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F, possess 16, 21 and 15 antigenicity sites, respectively.



There are 10 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	4	VIDGINYFLLTYCIRHISILLLVPLVIYYFVVFSLIQKKLISSSQFDILGLVLQLHM	60
2	68	SNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNL	99
3	101	SMLNYLMFLLKKLISIILLQFSVQGLFMIVQI	132
4	140	NLLNLLLLLIHLKDMFLRLEKKFFIILLLM	169
5	178	LMGLPMLLKLNQKLLFSLLQDLYLVM	203
6	207	EFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIGDVMLLKRKSLLLII	254
7	261	ENVVILIQIVIILIPYYFIQLIKMVVKKLIKKLLIFFLVFLWVVNILLLLLLLGSCYIIVKNLIYKMYLSRSCCIIE	337
8	343	DMIDLLDYKITSVIHY	358
9	363	MHMHYILFLEVLPLGFLNIIWSKVICCASQVMLILVKDILITLKILIQLDGILLLPKLILFHLTLLMKLIM	433
10	435	LGKFLKGFLHLIYHLVVVDIDVLGN	459

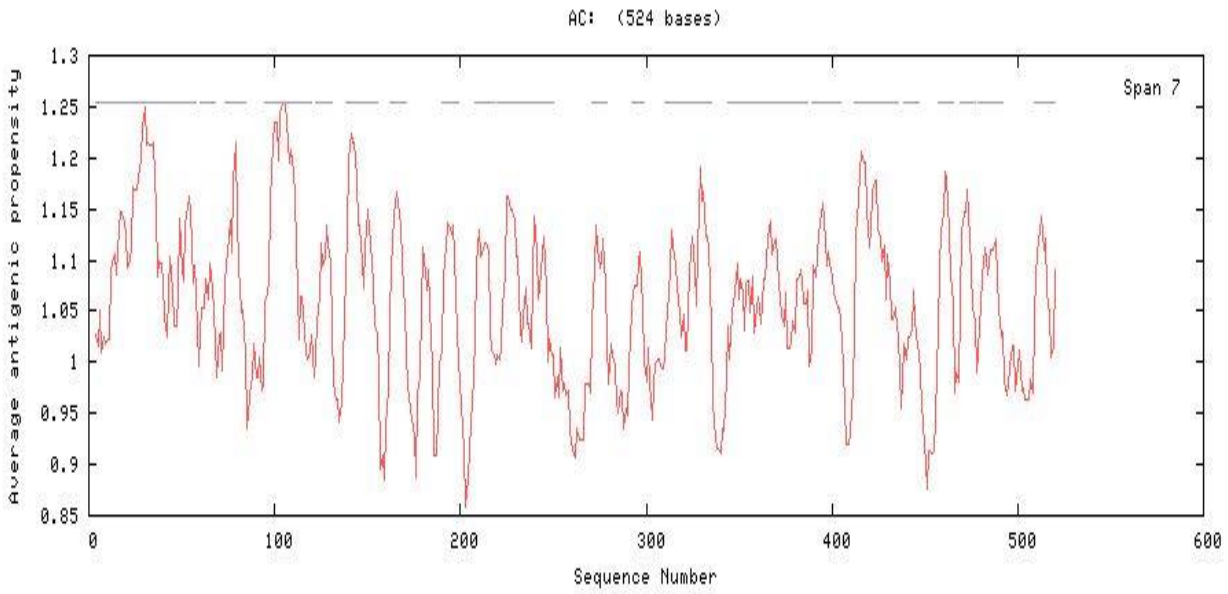
Figure 18: *ERG11* protein antigenic sites chromatogram of 01_*ERG11*-F



There are 16 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	4	VIDGINYFLSLSVPTDQYIISGTFVYYYSCVIYIHHEKIDLHYCFIGFLGLVLQLHM	60
2	68	SNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNL	99
3	101	SMLNYLMFLLKKLISIILLQFSVQGLFMIVQI	132
4	140	NLLNLLLLLIHLKDMFLRLEKKFFIILLLM	169
5	178	LMGLPMLLKLNQKLLFSLLQDLYLVM	203
6	207	EFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIGDVMLLKRKSLLLI	253
7	271	LIDSLLIHS	279
8	292	IANLLIGI	299
9	310	TSAWFLLHLGEKPHLQDVFIRSCCIIEKKR	339
10	349	DLQNYHQSYTIGTLRCICHYILFLEVLPLDSSTLYYSKVIM	389
11	393	SSRLCASYYYKIFFFPP	409
12	416	MGYCCCQSSFCFIILFFFSSLWVWE	440
13	443	FRGFFTLFT	451
14	461	WGTICLCSI	469
15	473	FNYFCLLLKMDYYWLLSAAPPLLF	496
16	513	RNLYVLIKRQLS	524

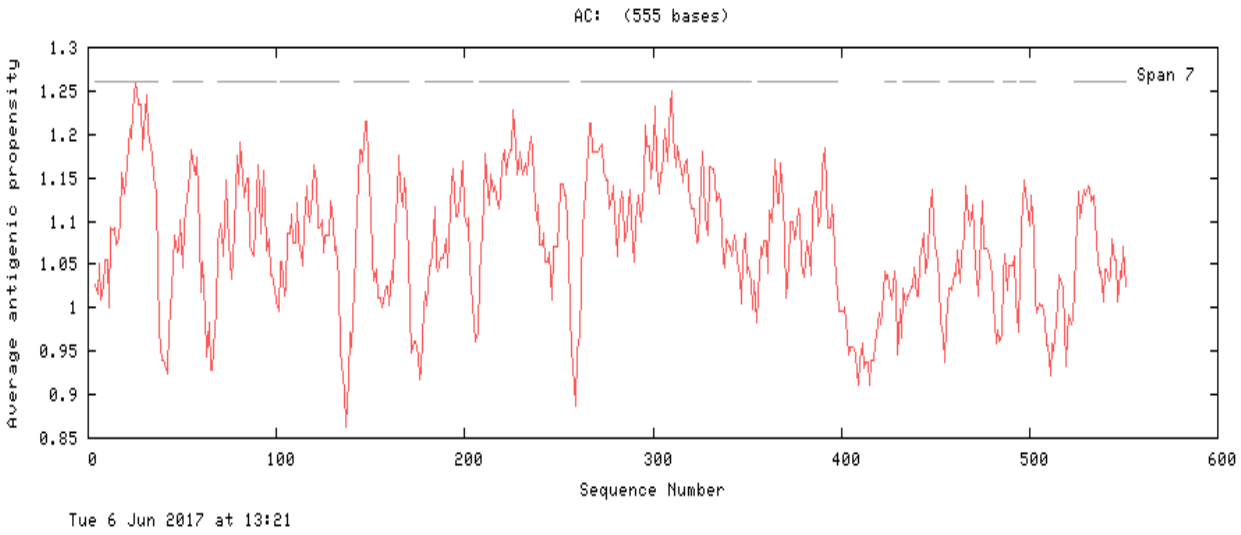
Figure 19: *ERG11* protein antigenic sites chromatogram of 17*_ERG11*-F



There are 21 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	4	VIDGINYFLMTILLHHISYYYWYLCIYLSVVIYILQKKLISSWLYSTSLVWFCSF	58
2	60	WSTTLLIFR	68
3	73	KVWWCIFIYVIR	84
4	95	RSSICFQCCIIICFCCRSLLTFNYSS	120
5	122	RYRGYLLLSK	131
6	138	KKICCICFDYYFIIKICS	155
7	162	FELFCYYYKF	171
8	190	NYYFHCFKIF	199
9	208	FFPFICSTIFF	218
10	220	RRRFYPYYFCFPPFTFTSLLETTCCSKKISA	250
11	271	LIDSLLIHS	279
12	292	IANLLIGI	299
13	310	TSAWFLLHLGEKPHLQDVFIRSCCII	335
14	344	EFDLLDYKIPSVNTLYELSDAYAILYFFKYYHRSSINYYSKVIM	387
15	389	SSSRLCSYYYKIFFFPP	405
16	412	MGYCCCQSSFCFIILFFFSSLWVWE	436
17	439	FRGFFTLFT	447
18	457	WGTICLCSI	465
19	469	FNYFCLLFK	477
20	479	DYYWLLSAAPPLLF	492
21	509	RNLYVLIKRQLS	520

Figure 20: *ERG11* protein antigenic sites chromatogram of 28_*ERG11*-F



There are 15 antigenic determinants in your sequence:

n	Start Position	Sequence	End Position
1	4	VIDGINYFLFDNCIHHISILLLVPLVIYLSVVFS	37
2	45	PSLTLFDFLGLVLQLHM	61
3	69	SNHVVKSMVMYFHLCYYGKLLRFIIVQKVMNL	100
4	102	SMLNYLMFLLKKLINIILLQFSVQGLFMIVQI	133
5	141	NLLNLLLLLIHLKDMFLRLEKKFFIILLLM	170
6	179	LMGLPMLLKLNQKLLFSLLQDLYLVM	204
7	208	EFLTVHLLNYILIIIKVLPLLILFSLIYLYLIIGDVMLLKRKSLLLII	255
8	262	ENVVILIQIVIILIPYYFIQLIKMVVKKLIKKLLIFFLVFLWVVNILLLLLLLGSCYIIVKNLIYRCYLSRSCCIIEKKKVVFEEFDLLRL	352
9	356	ISHTLYELSECICHTSYFRKVLSSIPESIIISKVIYGLVFSRY	398
10	423	KANSVSF	429
11	433	DEVDYGFGKVSKGVSSPYLP	452
12	457	RHRCIGEQFAYVQLGTILTTFVYNL	481
13	486	DGYKVPDP	493
14	495	YSSMVVLPT	503
15	524	FLSIQCSDCFHFVTTLDDHIYTYTYKYM	551

Contact Pedro Reche

Figure 21: *ERG11* protein antigenic sites chromatogram of Dog_*ERG11*-F

3.14 The Disordered/Disease Causing Region of the *ERG11* Protein of the Four Isolates Sequenced

The disease causing regions of the *ERG11* protein of 01_*ERG11*-F, 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F are shown in figures 22, 23, 24 and 25. Each figure indicates the amino acid composition of the *ERG11* protein disease-causing region for each isolate. The disordered/disease causing region for the isolates are painted blue on the X-axis. The amino acid sequence 2-8: ETVIDGI is common to all.

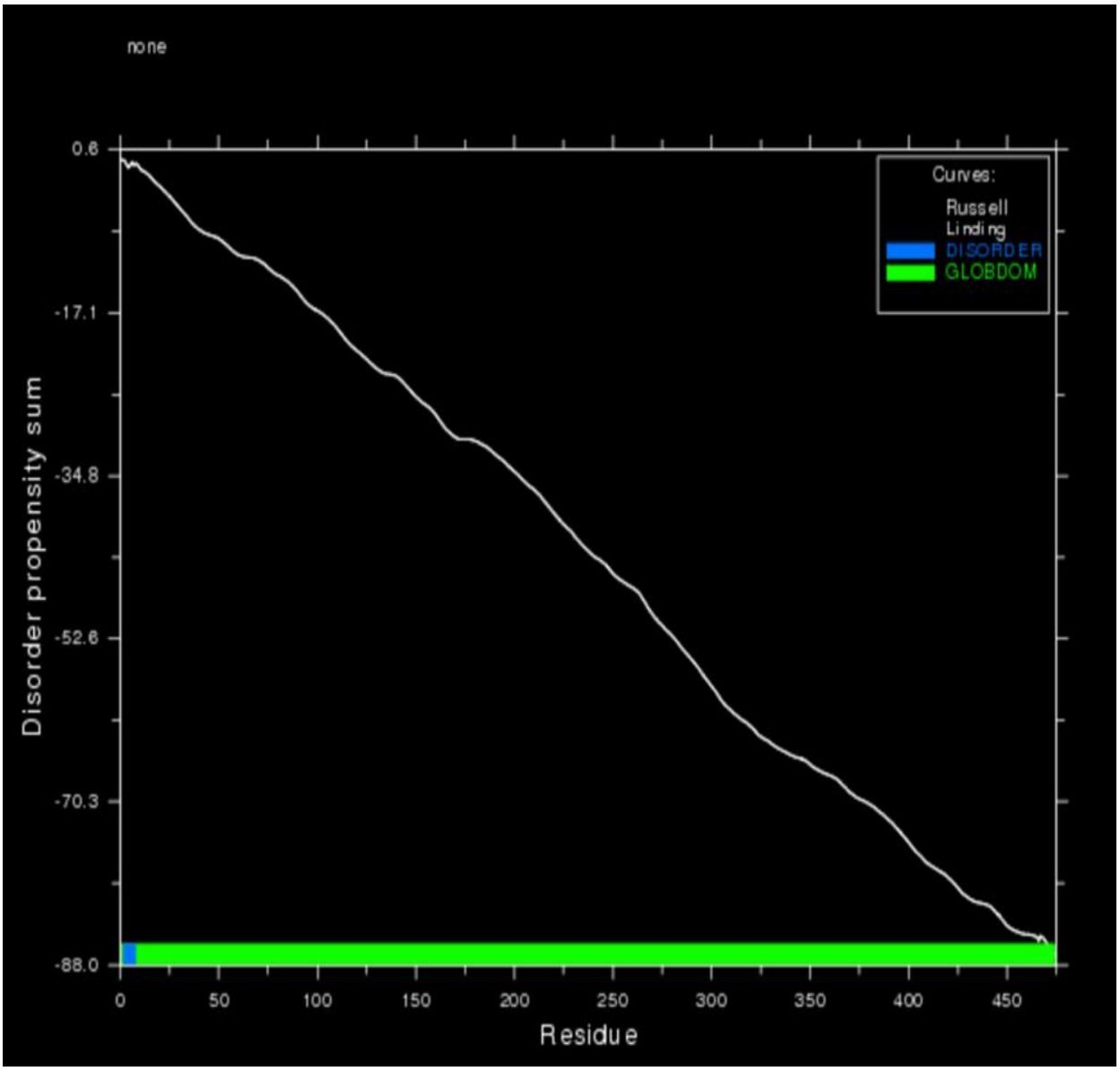


Figure 22: The disordered/disease causing region of 01_*ERG11*-F. The amino acids in this region painted in blue ranges from; 2-8: ETVIDGI

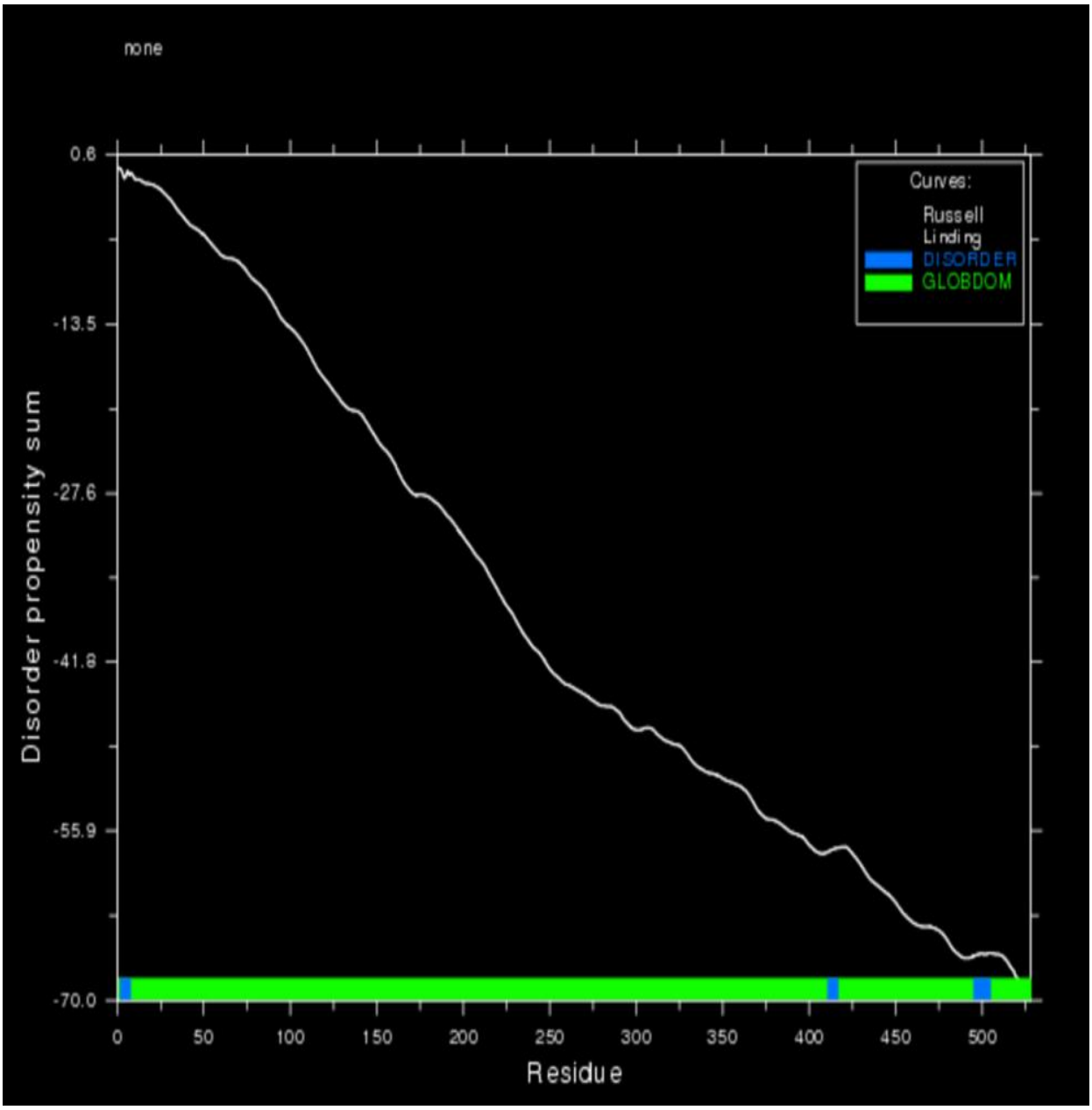


Figure 23: The disordered/disease causing region of 17_*ERG11*-F. The amino acids in this region painted in blue ranges from;
2-8: ETVIDGI
411-417: FFSNNMG
495-505: LFNGGFTYYTS

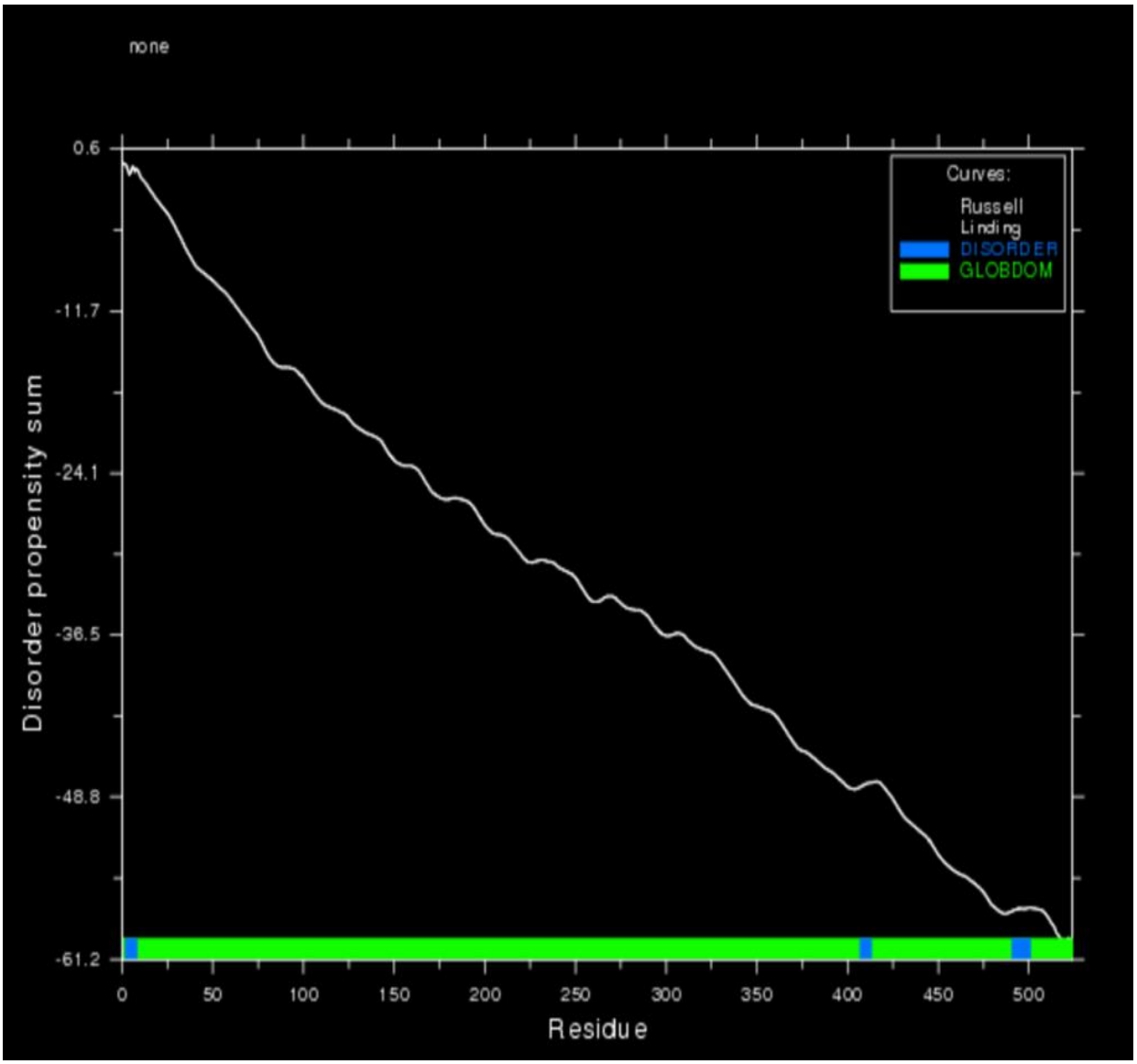


Figure 24: The disordered/disease causing region of 28_*ERG11*-F. The amino acids in this region painted in blue ranges from;
2-8: ETVIDGI
407-413: FFSNNMG
491-501: LFNGGFTYYTS

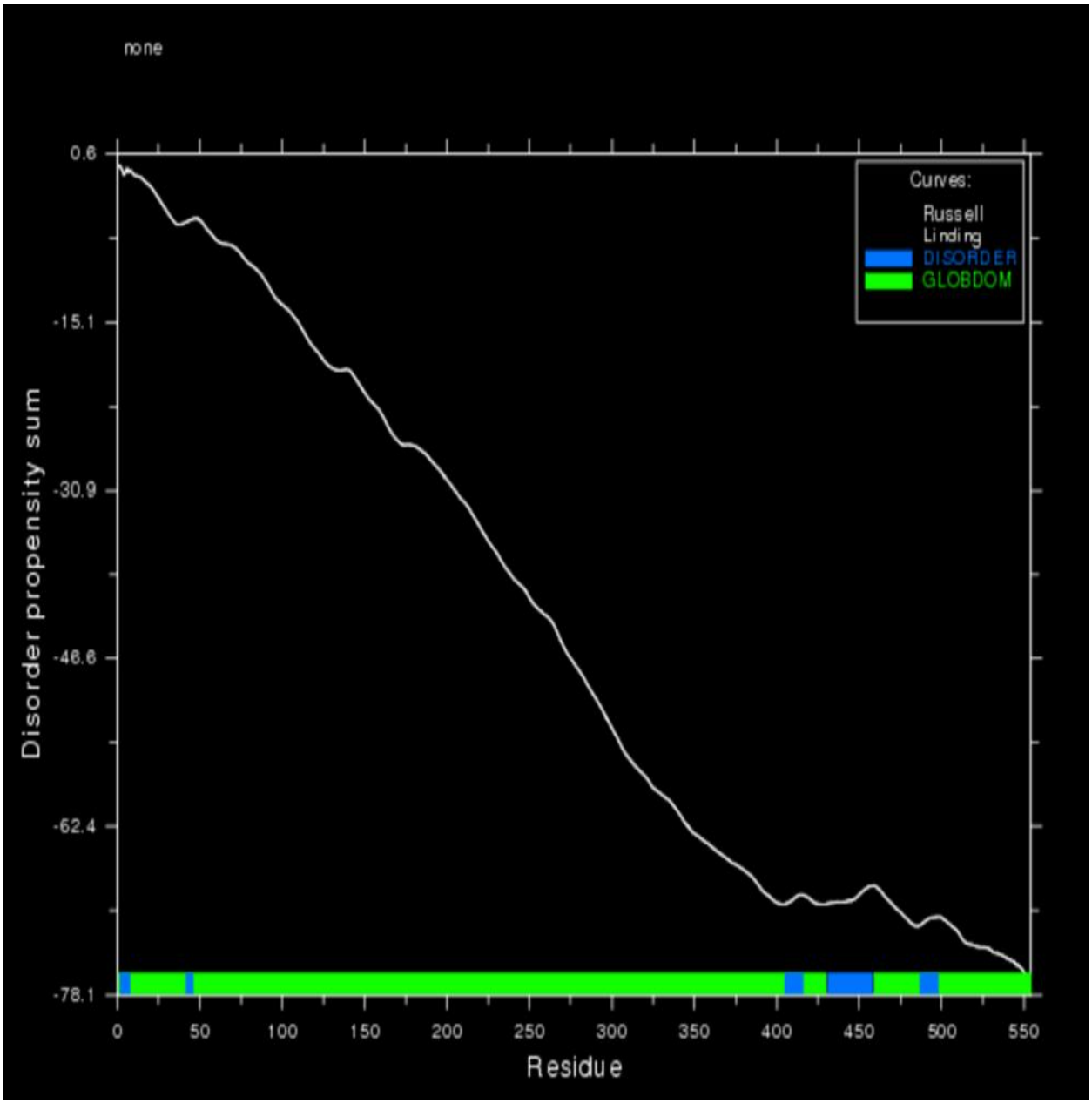


Figure 25: The disordered/disease causing region of Dog_*ERG11*-F. The amino acids in this region painted in blue ranges from;

2-8: ETVIDGI
42-46: DDSPS
405-416: YFDNPE DFDPTR
431-458: SSDEVDYGFG KVSKGVSSPY LPFGGGRH
487-498: GYKVPDPDYSSM

3.15 Phylogeny

The percentage of trees in which the associated taxa clustered together is shown next to the branches as shown in figure 26 and 27. Isolates 01_*ERG11*-F, 17_*ERG11*-F, 28_*ERG11*-F and Dog_*ERG11*-F are closely related organisms but more specifically, 01_*ERG11*-F and Dog_*ERG11*-F (Dog) are shown to have originated from a common ancestor.

Figure 26: Condensed Phylogenetic tree of *Candida ERG11*

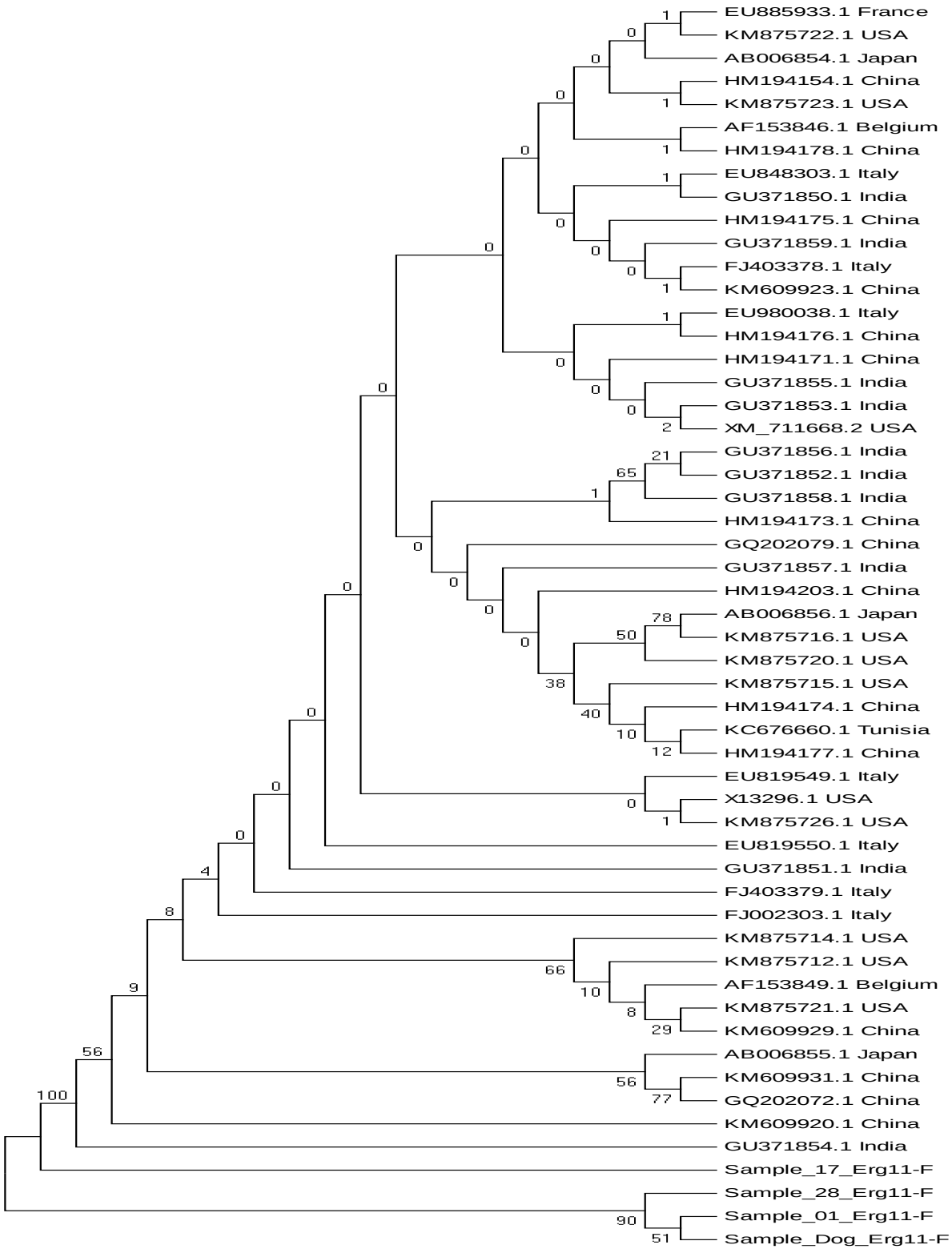


Figure 27: Bootstrap consensus phylogenetic tree of *Candida*

CHAPTER FOUR

DISCUSSION

The resistance of *Candida* species to drugs is a major problem in clinical practice as many *Candida* species are becoming increasingly resistant to first-line and second-line antifungal medications. One of the mechanisms responsible for this resistance is the alteration in the DNA sequence of the gene that codes for the protein (enzyme), *ERG11* which is the target of fluconazole. This study investigated the distribution of different species of *Candida* in human and dog vaginal swabs, the susceptibility of the *Candida* species to fluconazole, a first-line antifungal drug, the profile of the *ERG11* gene, and the phylogenetic relationship of the *Candida* species based on the nucleotide sequence of the *ERG11* gene.

A total of 57 human samples (HHVS- Human High Vaginal Swab) and 7 dog samples (DHVS- Dog High Vaginal Swab) were screened for the presence of *Candida* species. Twenty-eight (28) of the human samples were positive (+ve) to yeast growth as shown in Table 1. A total of 37 *Candida* isolates were obtained from the 28 human specimens that were positive to yeast growth as represented in Plates 1 and 2. The different strains of *Candida* isolated showed different morphological characteristics. Four different species of *Candida*, including *C. albicans, C. tropicalis, C. krusei*, and *C. glabrata* were identified in this study (Table 2). Some samples showed mixed growth, while others showed single isolate growth. Of the 28 human specimens that were positive to *Candida* growth, 21 had single species, 5 had two species and 2 had three species. Of the 37 isolates, 13 (35%) were *C. albicans*, 4 (9%) *C. glabrata*, 4 (9%) *C. krusei*, 2 (6%) C. *tropicalis*, and 14 (38%) other *Candida* species. Sundar *et al.* (2017) recently reported a distribution of 56% *C. albicans*, 20% *C. tropicalis*, 14% *C. glabrata* and 10% *C. krusei*. In this study, *C. albicans* was present in about 57% of the mixed cultures. This is similar to the findings of Yang *et al.* (2003), who reported a 60% occurrence of *C. albicans* in mixed cultures of *Candida*. The result confirms that other *Candida* species are haboured in the vagina of humans and dogs apart from the common *C. albicans, C. krusei, C. glabrata,* and *C. tropicalis*. Fears have been expressed that some of the uncommon *Candida* species such as *C. dubliniensis* could replace *C. albicans* as the predominant species under favorable conditions. *C. dubliniensis* which has been

reported in Singapore has caused great concern with its inducible fluconazole resistance (Tan *et al.,* 2002).

The antifungal susceptibility tests revealed that 89.2% of the *Candida* species were susceptible (≥19mm) to 25 µg fluconazole as shown in plates S1 – S38. This observation implies that fluconazole remains a very potent antifungal agent for *Candida* infection in this region. Overall, only 5.4% of the isolates from humans were resistant to fluconazole (≤14mm). This low percentage of fluconazole resistance suggests that there may be no pressing need for new antifungal drugs against *Candida* in the region of this study. The only isolate from dog was also susceptible to the 25 µg fluconazole. This implies that fluconazole still supports the treatment of *Candida* infections in dogs. One of the remarkable findings of this research is the observation of a very resistant strain (a *C. glabrata*), isolated from humans, with a zone of inhibition of 0.00mm. This strain may have undergone a series of mutations as well as evolutionary adaptations, such that it circumvents all the possible mechanisms of action of fluconazole (Stephanie *et al.,* 2012). This finding implies that fluconazole will be the wrong option for treating any patient infected with this strain. The most susceptible isolate was a *C. albicans* species isolated from humans. This study also showed 95.8% of the non-*albicans* species to be susceptible to fluconazole. This suggests that the non-*albicans* species are not necessarily a resistance threat to fluconazole.

The four (4) isolates whose *ERG11* genes were sequenced were *Can Iso*-001 (a *C. glabrata*; the most resistant), *Can Iso*-17 (a *C. albicans*; a susceptible dose-dependent), *Can Iso*-028 (a *C. albicans*; the most susceptible) and *Can Iso*-029 (a *C. krusei*; the dog sample). The nucleotide sequence composition of the *ERG11* gene of the various isolates varied considerably as reflected in the nucleotide counts. The most fluconazole-resistant isolate (*C. glabrata*) had 1431 bases. This is shorter than the sequence length of 1602 bases reported by Silva *et al.* (2009). The susceptible dose-dependent isolate (*C. albicans*) had 1590 bases, and the most susceptible (*C. albicans*) had I578 bases. For *C. albicans* species, the most commonly reported sequence length is 1587 bases (Ying *et al.,* 2013; Flowers *et al.,* 2015; Hu *et al.,* 2015; Rossini *et al.,* 2015; Fakhim *et al.,* 2016), however, Wang *et al.* (2009) reported a sequence length of 1370 bases. The isolate from dog, a *C. krusei* had 1668 bases. Silva *et al.* (2016) reported a sequence length of 1587 bases in *C. krusei*, while, Healey *et al.* (2016) reported a much higher nucleotide sequence length of 3907 bases for

C. krusei. Nucleotide sequence lengths of the same gene of even the same species have been observed to vary, and this could be attributed to the molecular adaptation in the particular strains (Zimbeck *et al*., 2010). The differences observed in the sequences of the strains of the same species could be attributed to mutations or single nucleotide polymorphisms (Forche *et al.,* 2004; Gomez *et al.,* 2008). The resistance may also be due to changes resulting from other counter mechanisms against antifungal agents (Sanglard and White, 2006; Coste *et al.,* 2007).

The primary structure of the *ERG11* proteins could have effects on electrostatic properties, protein packing, local and global structure, stability, interactions, activity and abundance. Thus, the observed variations in the *ERG11* gene may manifest in different ways as observed in the widely different sequence compositions. Many deleterious variations, such as large deletions, protein truncations, amphigoric amino acid insertions, deletions and indels, are easy to explain, while the minor sequence alterations, most often resulting in amino acid substitutions are more difficult to explain. Variations in primary, secondary and tertiary structures of *ERG11* proteins are also more difficult to explain than those of the DNA sequence itself (Kircher *et al.,* 2014; Kucukkal *et al.,* 2014; Perniola and Musco 2014). The amino acid sequence (primary structure) translated from the *ERG11* nucleotide sequence of the four (4) isolates show the amino acid composition to vary considerably (Wang *et al.,* 2009; Xiang *et al.,* 2013; Silva *et al.,* 2016). *Can Iso*-001 had 476 amino acids, *Can Iso*-17 had 529 amino acids, *Can Iso*-028 had 525 amino acids, and the dog isolate, *Can Iso*-029, had 555 amino acids. The corresponding proteins of the *Candida ERG11* gene have molecular weights of 56,183.10Da, 62,897.16Da, 64,159.15Da, and 65,139.25Da, respectively. These are proteins with high molecular weights. Debnath and Addya (2014), reported a molecular weight of 60675.4Da for the *ERG11* protein of *Candida*, which is similar to the findings of this study. The observed differences are due to the changes in nucleotide sequences that may have resulted from mutation. The *ERG11* proteins from this study had a predicted pI (isoelectric point) of 9.60, 9.03, 9.03 and 8.88, for *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029, respectively. In a solution with a pH that is above the pI of these *ERG11* proteins, the surface of the protein is predominantly negatively charged. Likewise, in a solution with a pH that is below the pI of these *ERG11* proteins, the surface of the protein is predominantly positively charged. This knowledge would be found useful in the purification of these enzymes for further kinetic studies. The pI of most proteins is in the pH range of 4 to 7. In a study carried out by Debnath and

Addya (2014), the pI for *ERG11* protein was 6.69. The *ERG11* proteins in this study, however, had higher pI values. These enzymes are alkaline in nature and are predicted to be localized in the plasma membrane. The predicted half-live ($t_{1/2}$) of these proteins in mammalian cells is 100 hours, and the instability coefficients are 35.70, 36.98, 44.50 and 37.69, for *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029, respectively. A protein with instability index smaller than 40 is predicted as stable (Fang *et al.,* 2011). This instability index also provides an estimate of the stability of a protein in a test tube. The implication of this finding is that in some cases, at some point in the life-cycle of the *Candida*, the fungi may get ruptured and this protein gets into the host's system. The protein, however, needs to be cleared off the system, and by 100 hours, the amount of this disease-causing protein will be halved (Benson *et al.,* 2009). This information could be important in pharmaceutical drug designs, and dosage studies. Based on this, all the *ERG11* proteins of the *Candida* species that were sequenced for study, could be said to be stable, except the most susceptible (*Can Iso*-028). This lower stability of the most susceptible isolate could explain its high susceptibility to 25 μg fluconazole. The grand average of hydrophobicity (GRAVY) of these *Candida ERG11* proteins are 0.195, 0.478, 0.195 and 0.576, for *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029, respectively. The more positive the value, the more hydrophobic the amino acids located in that region of the protein are (Jaspard *et al.,* 2012). Therefore, in employing water as a therapeutic option, infection with *Can Iso*-17 and the *Can Iso*-029 are more easily treated than infections with *Can Iso*-001 and *Can Iso*-28.

The multiple sequence alignment of both the nucleotide and amino acid sequence of the four *Candida* isolates sequenced in this study revealed numerous variable sites, with two major conserved domains (Fig. 1a, 1b, 1c, 1d, 2, 3 and 4). The major principle of comparative genomics is that common features of two or more organisms will often be conserved within the DNA (Koonin, and Galperin, 2003). There have been concerns as to whether to use nucleotide sequences or amino acid sequences in the analysis of conserved domains. However, due to the redundancy of the genetic code the more reliable option is the amino acid multiple sequence alignment (Yang *et al.,* 2015). This does not in any way negate the credibility of nucleotide multiple sequence alignment. The Percentage Identity Matrix (PIM) between the four isolates in the nucleotide multiple sequence alignment were all above 90%, and specifically in the range of 91.02% to 94.55% (Tables 4 and 5). However, the PIM of the four isolates in the amino acid multiple

sequence alignment were all below 80%, and specifically in the range of 23.78% to 76.68% (Table 6 and 7). All the multiple sequence alignment E-values are at the significant levels (0.0 for nucleotide multiple sequence alignment, and a range of 0.0 to 5e-09 for amino acid multiple sequence alignment), as E-values below 0.05 are considered significant. The first 12 amino acids in the sequences (as shown by the amino acid multiple sequence alignment) make up the first major conserved domain. Due to the varying length of the *Candida ERG11* gene sequences of the four isolates, the position of the second conserved domain varies in each of the sequences. The amino acid sequence making up the second major conserved region in *Can Iso*-001, *Can Iso*-017, *Can Iso*-028 and *Can Iso*-029 are in positions 331-337, 332-338, 330-336 and 330-336, respectively (fig. 3). These regions are potential target sites for therapeutic agents (Maurice *et al*., 2014). The subcellular localization prediction confirmed their *Candida ERG11* proteins to be a membrane enzyme which catalyzes the ergosterol synthetic pathway (Fothergill *et al*., 2014). All the *ERG11* proteins had varying composition of amino acids as shown in Fig. 5. The result shows a low concentration of amino acids with polar side chains in all four species of *Candida ERG11* protein. These amino acids include Arginine, Asparagine, Aspartate, Glutamine, Glutamate, Histidine, Lysine and Serine.

The secondary structure prediction of the *ERG11* protein showed a high percentage of α-helix and β-sheet, contributing to the stability and conservation of this enzyme (Fig. 6, 7, 8, and 9). There is a relationship between the α-helix content of a protein and its stability. A protein with a higher percentage of α-helix appears to be more stable (Deller *et al.*, 2016). This stability reflects in the susceptibility to the action of drugs. In this study, the most resistant strain of *Candida* had the highest percentage of α-helix, while the most susceptible strain had the lowest percentage of α-helix (Fig. 6, 7, 8, and 9). From the findings of this study, there appears to also be a relationship between the level of β-sheets in a protein and its susceptibility to fluconazole, as the most resistant strain of *Candida* had the highest percentage of β-sheets, while the most susceptible strain had the lowest percentage of β-sheets. The tertiary structure comparison of the most resistant, the SDD, the most resistant and the dog isolate gave an insight into the different levels of drug resistance observed between them (Fig. 10, 11, 12 and 13). Apart from the most resistant species, whose *ERG11* protein had a strikingly different tertiary structure, all the other three *ERG11* proteins had tertiary structures that predictively explain their susceptibility to fluconazole (Shi *et al*., 2012).

Ligand/drug binding sites were detected in the *ERG11* protein of three of the four isolates sequenced. The most resistant isolate was the only exception, which had no single drug binding site (fig. 14, 15, 16, and 17). The absence of a drug-binding site in the *ERG11* protein of the resistant isolate may explain why the strain is completely resistant to fluconazole.

Antigenicity of a protein refers to its ability to elicit a specific immune response, and the activation of the immune system by the invading antigen is faster when an organism with a wider range of antigenicity sites is involved (Soria-Guerra *et al.*, 2015). The most resistant strain had 10 antigenicity sites, which appears to be the lowest compared to others. The other 3 isolates sequenced, had 16 (the SDD), 21 (the most resistant) and 15 (the isolate from dog) antigenicity sites, respectively (Fig. 18, 19, 20 and 21). Thus, the lower number of antigenic sites observed in the most resistant strain may be a survival strategy to circumvent detection by the host's immune system (Soria-Guerra *et al.*, 2015). These regions can also be regarded as ideal targets for drugs, which could be considered during the design as well as in the development of vaccines (Lorenzo and Fenton, 2013).

The experiment to identify the disease-causing regions of the *ERG11* protein revealed the presence of disease-causing regions in the *ERG11* protein of all the four strains of the *Candida* sequenced. The amino acid sequence of the regions is displayed in Fig. 22, 23, 24 and 25. Amongst the varying disease-causing regions in each of the proteins, it was observed that all these *ERG11* proteins have a common disease-causing region. This ranges within the amino acid residues 2-8: ETVIDGI. However, the isolates have other disease-causing regions that are peculiar to each of the strains (Fig. 22, 23, 24 and 25). These disease-causing regions can also serve as a potential target for drugs.

One of the major objectives of this study was to use phylogenetic analysis of the *ERG11* to determine the evolutionary relatedness of the isolates to each other, as well as their relatedness to *Candida* species from some other parts of the world. The result of this investigation revealed that the four isolates were closely related (Fig. 26 and 27). However, the most resistant strain (*Can Iso*-001) and the dog isolate (*Can Iso*-029), seemed to have originated from a common ancestor, implying an even closer evolutionary relatedness. A further revelation of the phylogenetic analysis is that, the isolates obtained in this study are more closely related to strains isolated in India and

China, than those isolated in USA, Belgium, Italy and France (Tamura *et al.*, 2013; Kumar *et al.*, 2016). The closer relationship between the Nigerian isolates to isolates from India and China (both in Asia) may be attributed to the closer similarity between the climates of Africa and Asia. This finding also suggests that, other drugs apart from fluconazole, used in the treatment of *Candida* infection in India and China may also be effective against *Candida* in Nigeria.

Conclusion

This study investigated the distribution of different species of *Candida* in human and dog vaginal swabs, the susceptibility of the *Candida* species to fluconazole, a first-line antifungal drug, the profile of the *ERG11* gene, and the phylogenetic relationship of the *Candida* species based on the nucleotide sequence of the *ERG11* gene. Different species of *Candida* were isolated, with *C. albicans* being the most predominant species. The behaviour of the *Candida* strains whose *ERG11* gene profile were investigated are shown to be strongly connected to the variations in their respective nucleotide and amino acid sequences. These variations could be attributed to mutations, to circumvent the action of fluconazole, as well as evolutionary changes overtime. The susceptibility experiment showed that most of the *Candida* strains are susceptible to 25 μg fluconazole, however, one of the isolates was fully resistant to 25 μg fluconazole. The four isolates whose *ERG11* genes were sequenced varied in their nucleotide and amino acid sequences, and this was confirmed by multiple sequence alignment, as well as in their physicochemical properties. The predicted primary, secondary and tertiary structure of the *ERG11* protein all varied, with a particularly marked difference in the most resistant isolate. All the predicted *ERG11* proteins varied in their ligand-binding sites, antigenic sites, and disease causing regions. The phylogenetic analysis based on the *ERG11 gene* showed that, the four isolates sequenced are closely related in origin, but more specifically, the most resistant isolate and the isolate from the dog are shown to have originated from a common ancestor. The information obtained from this study can be gainfully employed in the design of new drugs and vaccines, which could relief the key challenges originating from *Candida* in clinical practice.

Recommendations for Further Studies

- Apart from the *ERG11* gene profiling, studies on the *ABC* transporter genes should also be done alongside *ERG11* gene profiling experiments in subsequent studies, since they could also be contributory factors to the response of the *Candida* to fluconazole.
- A study is needed to determine the prevalence of the *C. glabrata* because of its absolute resistance to fluconazole.
- Further laboratory and clinical experiments on the design of potential anti-fungal drugs and vaccines that will translate through clinical trial phases are needed. The key knowledge revealed in the conserved domain should be effectively harnessed since it could serve as a root target for all forms of *Candida* infections, irrespective of species.

REFERENCES

Akins, R. A. (2005). An update on antifungal targets and mechanisms of resistance in *Candida albicans*. *Medical Mycology*, **43**(4): 285-318.

Almeida, A. A. D., Mesquita, C. S. S., Svidzinski, T. I. E. and Oliveira, K. M. P. D. (2013). Antifungal susceptibility and distribution of *Candida* spp. isolates from the University Hospital in the municipality of Dourados, State of Mato Grosso do Sul, Brazil. *Revista da Sociedade Brasileira de Medicina Tropical*, **46**(3): 335-339.

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research*, **25**(17): 3389-3402.

Altschul, S. F., Wootton, J. C., Gertz, E. M., Agarwala, R., Morgulis, A., Schäffer, A. A. and Yu, Y. K. (2005). Protein database searches using compositionally adjusted substitution matrices. *The FEBS journal*, **272**(20): 5101-5109.

Arnaud, M. B., Costanzo, M. C., Skrzypek, M. S., Shah, P., Binkley, G., Lane, C. and Sherlock, G. (2006). Sequence resources at the *Candida* genome database. *Nucleic acids research*, *35*(suppl_1), D452-D456.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E. and Grosdidier, A. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic acids research*, **40**(W1): W597-W603.

Barchiesi, F., Calabrese, D., Sanglard, D., Di Francesco, L. F., Caselli, F., Giannini, D. and Scalise, G. (2000). Experimental induction of fluconazole resistance in *Candida tropicalis* ATCC 750. *Antimicrobial Agents and Chemotherapy*, **44**(6): 1578-1584.

Bard, M., Sturm, A. M., Pierson, C. A., Brown, S., Rogers, K. M., Nabinger, S. and Hazen, K. C. (2005). Sterol uptake in *Candida glabrata*: rescue of sterol auxotrophic strains. *Diagnostic Microbiology and Infectious Disease*, **52**(4): 285-293.

Berkow, E. L., Manigaba, K., Parker, J. E., Barker, K. S., Kelly, S. L. and Rogers, P. D. (2015). Multidrug transporters and alterations in sterol biosynthesis contribute to azole antifungal resistance in *Candida parapsilosis*. *Antimicrobial Agents and Chemotherapy*, **59**(10): 5942-5950.

Branco, J., Silva, A. P., Silva, R. M., Silva-Dias, A., Pina-Vaz, C., Butler, G. and Miranda, I. M. (2015). Fluconazole and voriconazole resistance in *Candida parapsilosis* is conferred by gain-of-function mutations in MRR1 transcription factor gene. *Antimicrobial Agents and Chemotherapy*, **59**(10): 6629-6633.

Bruder-Nascimento, A., Camargo, C. H., Sugizaki, M. F., Sadatsune, T., Montelli, A. C., Mondelli, A. L. and Bagagli, E. (2010). Species distribution and susceptibility profile of *Candida* species in a Brazilian public tertiary hospital. *BMC Research Notes*, **3**(1): 1-5.

Brun, S., Aubry, C., Lima, O., Filmon, R., Bergès, T., Chabasse, D. and Bouchara, J. P. (2003). Relationships between respiration and susceptibility to azole antifungals in *Candida glabrata*. *Antimicrobial Agents and Chemotherapy*, **47**(3): 847-853.

Cannon, R. D., Lamping, E., Holmes, A. R., Niimi, K., Baret, P. V., Keniya, M. V. and Monk, B. C. (2009). Efflux-mediated antifungal drug resistance. *Clinical Microbiology Reviews*, **22**(2): 291-321.

Cappelletty, D. and Eiselstein-McKitrick, K. (2007). The echinocandins. Pharmacotherapy: The *Journal of Human Pharmacology and Drug Therapy*, **27**(3): 369-388.

Caudle K. E., Barker K. S., Wiederhold N. P., Xu L., Homayouni, R. and Rogers P. D. (2011). Genome-wide expression profile analysis of the *Candida glabrata* Pdr1 regulon. *Eukaryot. Cell*, **10**, 373–383.

Chandrasekar, P. H. and Sobel, J. D. (2006). Micafungin: A new echinocandin. *Clinical Infectious Diseases*, **42**(8): 1171-1178.

Chau, A. S., Mendrick, C. A., Sabatelli, F. J., Loebenberg, D. and McNicholas, P. M. (2004). Application of real-time quantitative PCR to molecular analysis of *Candida albicans* strains exhibiting reduced susceptibility to azoles. *Antimicrobial Agents and Chemotherapy*, **48**(6): 2124-2131.

Cheng, J. W., Liao, K., Kudinha, T., Yu, S. Y., Xiao, M., Wang, H. and Xu, Y. C. (2017). Molecular epidemiology and azole resistance mechanism study of *Candida* guilliermondii from a Chinese surveillance system. *Scientific Reports*, **7**(1):907-914.

Choi, M. J., Won, E. J., Shin, J. H., Kim, S. H., Lee, W. G., Kim, M. N. and Im, Y. J. (2016). Resistance mechanisms and clinical features of fluconazole-nonsusceptible *Candida* tropicalis isolates compared with fluconazole-less-susceptible isolates. *Antimicrobial Agents and Chemotherapy*, **60**(6): 3653-3661.

Cleveland, A. A., Harrison, L. H., Farley, M. M., Hollick, R., Stein, B., Chiller, T. M. and Park, B. J. (2015). Declining incidence of candidemia and the shifting epidemiology of *Candida* resistance in two US metropolitan areas, 2008–2013: Results from population-based surveillance. *PLoS One*, **10**(3): e0120452.

Coleman, J. J., Okoli, I., Tegos, G. P., Holson, E. B., Wagner, F. F., Hamblin, M. R. and Mylonakis, E. (2010). Characterization of plant-derived saponin natural products against *Candida albicans*. *ACS Chemical Biology*, **5**(3): 321-332.

Cornely, O. A., Gachot, B., Akan, H., Bassetti, M., Uzun, O., Kibbler, C. and Ameye, L. (2015). Epidemiology and outcome of fungemia in a cancer Cohort of the Infectious Diseases Group (IDG) of the European Organization for Research and Treatment of Cancer (EORTC 65031). *Clinical Infectious Diseases*, **61**(3): 324-331.

Coste, A. T., Karababa, M., Ischer, F., Bille, J. and Sanglard, D. (2004). TAC1, transcriptional activator of CDR genes, is a new transcription factor involved in the regulation of *Candida albicans* ABC transporters CDR1 and CDR2. *Eukaryotic Cell*, **3**(6): 1639-1652.

Coste, A., Selmecki, A., Forche, A., Diogo, D., Bougnoux, M. E., d'Enfert, C. and Sanglard, D. (2007). Genotypic evolution of azole resistance mechanisms in sequential *Candida albicans* isolates. *Eukaryotic Cell*, **6**(10): 1889-1904.

Coste, A., Turner, V., Ischer, F., Morschhäuser, J., Forche, A., Selmecki, A. and Sanglard, D. (2006). A mutation in Tac1p, a transcription factor regulating CDR1 and CDR2, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in *Candida albicans*. *Genetics*, **172**(4): 2139-2156.

Davey, M. E. and O'toole, G. A. (2000). Microbial biofilms: From ecology to molecular genetics. *Microbiology and Molecular Biology Reviews*, **64**(4): 847-867.

De-Almeida, A. A., Nakamura, S. S., Fiorini, A., Grisolia, A. B., Svidzinski, T. I. E. and de Oliveira, K. M. P. (2015). *Revista Iberoamericana de Micología*, **32**(3): 153-158.

Deller, M. C., Kong, L. and Rupp, B. (2016). Protein stability: a crystallographer's perspective. *Acta Crystallographica Section F: Structural Biology Communications*, **72**(2): 72-95.

Desnos-Ollivier, M., Ragon, M., Robert, V., Raoux, D., Gantier, J. C. and Dromer, F. (2008). Debaryomyces hansenii (*Candida* famata): a rare human fungal pathogen often misidentified as Pichia guilliermondii (*Candida* guilliermondii). *Journal of clinical microbiology*, **46**(10): 3237-3242.

Dongari-Bagtzoglou, A., Dwivedi, P., Ioannidou, E., Shaqman, M., Hull, D. and Burleson, J. (2009). Oral *Candida* infection and colonization in solid organ transplant recipients. *Molecular Oral Microbiology*, **24**(3): 249-254.

Dos-Santos-Silva, D. B., Grisolia, A. B. and de Oliveira, K. M. P. (2016). Genetic determinants of antifungal resistance in *Candida* species. *African Journal of Biotechnology*, **15**(40): 2259-2264.

Douglas, L. J. (2003). *Candida* biofilms and their role in infection. *Trends in Microbiology*, **11**(1): 30-36.

Dunkel, N., Blaß, J., Rogers, P. D. and Morschhäuser, J. (2008). Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of

MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. *Molecular Microbiology*, **69**(4): 827-840.

Eddouzi, J., Parker, J. E., Vale-Silva, L. A., Coste, A., Ischer, F., Kelly, S. and Sanglard, D. (2013). Molecular mechanisms of drug resistance in clinical *Candida* species isolated from Tunisian hospitals. *Antimicrobial Agents and Chemotherapy*, **57**(7): 3182-3193.

Eggimann, P., Garbino, J. and Pittet, D. (2003). Epidemiology of *Candida* species infections in critically ill non-immunosuppressed patients. *The Lancet Infectious Diseases,* **3**(11): 685-702.

Espinel-Ingroff, A., Canton, E., Peman, J., Rinaldi, M. G. and Fothergill, A. W. (2009). Comparison of 24-hour and 48-hour voriconazole MICs as determined by the Clinical and Laboratory Standards Institute broth microdilution method (M27-A3 document) in three laboratories: results obtained with 2,162 clinical isolates of *Candida* spp. and other yeasts. *Journal of Clinical Microbiology*, **47**(9): 2766-2771.

Fang, Q., Kani, K., Faca, V. M., Zhang, W., Zhang, Q., Jain, A. and Mallick, P. (2011). Impact of protein stability, cellular localization, and abundance on proteomic detection of tumor-derived proteins in plasma. *PloS one,* **6**(7): e23090.

Fasman GD, editor. Prediction of protein structure and the principles of protein conformation. Springer Science and Business Media; 2012 Dec 6. Springer US

Ferrari, S., Ischer, F., Calabrese, D., Posteraro, B., Sanguinetti, M., Fadda, G. and Sanglard, D. (2009). Gain of function mutations in CgPDR1 of *Candida glabrata* not only mediate antifungal resistance but also enhance virulence. *PLoS Pathogens,* **5**(1): e1000268.

Flowers, S. A., Colón, B., Whaley, S. G., Schuler, M. A. and Rogers, P. D. (2015). Contribution of clinically derived mutations in ERG11 to azole resistance in *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **59**(1): 450-460.

Fothergill, A. W. (2012). Antifungal susceptibility testing: clinical laboratory and standards institute (CLSI) methods. In *Interactions of Yeasts, Moulds, and Antifungal Agents* (pp. 65-74). Humana Press.

Fothergill, A. W., Sutton, D. A., McCarthy, D. I. and Wiederhold, N. P. (2014). The impact of new antifungal breakpoints on antifungal resistance in *Candida* species. *Journal of Clinical Microbiology*, **52**(3):994-997.

Furletti, V. F., Teixeira, I. P., Obando-Pereda, G., Mardegan, R. C., Sartoratto, A., Figueira, G. M. and Höfling, J. F. (2011). Action of *Coriandrum sativum* L. essential oil upon oral *Candida albicans* biofilm formation. *Evidence-Based Complementary and Alternative Medicine*, 2011: 1-9

Garcia-Effron, G., Katiyar, S. K., Park, S., Edlind, T. D. and Perlin, D. S. (2008). A naturally occurring proline-to-alanine amino acid change in Fks1p in *Candida parapsilosis*, *Candida orthopsilosis*, and *Candida metapsilosis* accounts for reduced echinocandin susceptibility. *Antimicrobial Agents and Chemotherapy*, **52**(7): 2305-2312.

Garcia-Effron, G., Park, S. and Perlin, D. S. (2009). Correlating echinocandin MIC and kinetic inhibition of fks1 mutant glucan synthases for *Candida albicans*: implications for interpretive breakpoints. *Antimicrobial Agents and Chemotherapy*, **53**(1): 112-122.

Gardy, J. L., Laird, M. R., Chen, F., Rey, S., Walsh, C. J., Ester, M. and Brinkman, F. S. (2004). PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. *Bioinformatics*, **21**(5): 617-623.

Garzoni, C., Nobre, V. A. and Garbino, J. (2007). *Candida parapsilosis* endocarditis: A comparative review of the literature. *European Journal of Clinical Microbiology and Infectious Diseases*, **26**(12): 915-926.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S. E., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005). Protein identification and analysis tools on the ExPASy server (pp. 571-607). Humana Press.

Gomez-Lopez, A., Alastruey-Izquierdo, A., Rodriguez, D., Almirante, B., Pahissa, A., Rodriguez-Tudela, J. L. and Barcelona Candidemia Project Study Group. (2008). Prevalence and susceptibility profile of *Candida* metapsilosis and *Candida* orthopsilosis: results from population-based surveillance of candidemia in Spain. *Antimicrobial agents and chemotherapy*, **52**(4): 1506-1509.

Gregory, T. R. (2008). Understanding evolutionary trees. *Evolution: Education and Outreach*, **1**(2): 121-137.

Grossman, N. T., Pham, C. D., Cleveland, A. A. and Lockhart, S. R. (2015). Molecular mechanisms of fluconazole resistance in *Candida parapsilosis* isolates from a US surveillance system. *Antimicrobial Agents and Chemotherapy*, **59**(2): 1030-1037.

Grover, N. D. (2010). Echinocandins: A ray of hope in antifungal drug therapy. *Indian Journal of Pharmacology*, **42**(1): 9–11

Harriott, M. M. and Noverr, M. C. (2011). Importance of *Candida*–bacterial polymicrobial biofilms in disease. *Trends in Microbiology*, **19**(11): 557-563.

Hasan, F., Xess, I., Wang, X., Jain, N. and Fries, B. C. (2009). Biofilm formation in clinical *Candida* isolates and its association with virulence. *Microbes and Infection*, **11**(8): 753-761.

Healey, K. R., Zhao, Y., Perez, W. B., Lockhart, S. R., Sobel, J. D., Farmakiotis, D. and Jimenez-Ortigosa, C. (2016). Prevalent mutator genotype identified in fungal pathogen *Candida glabrata* promotes multi-drug resistance. *Nature Communications*, **7**, 11117–11128.

Heilmann, C. J., Schneider, S., Barker, K. S., Rogers, P. D. and Morschhäuser, J. (2010). An A643T mutation in the transcription factor Upc2p causes constitutive ERG11 upregulation and increased fluconazole resistance in *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **54**(1): 353-359.

Heilmann, C. J., Schneider, S., Barker, K. S., Rogers, P. D. and Morschhäuser, J. (2010). An A643T mutation in the transcription factor Upc2p causes constitutive ERG11 upregulation and increased fluconazole resistance in *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **54**(1): 353-359.

Hof, H. (2006). A new, broad-spectrum azole antifungal: Posaconazole–mechanisms of action and resistance, spectrum of activity. *Mycoses*, **49**(s1): 2 - 6.

Holetz, F. B., Pessini, G. L., Sanches, N. R., Cortez, D. A. G., Nakamura, C. V. and Dias Filho, B. P. (2002). Screening of some plants used in the Brazilian folk medicine for the treatment of infectious diseases. *Memórias do Instituto Oswaldo Cruz*, **97**(7): 1027-1031.

Holmes, A. R., Tsao, S., Ong, S. W., Lamping, E., Niimi, K., Monk, B. C. and Cannon, R. D. (2006). Heterozygosity and functional allelic variation in the *Candida albicans* efflux pump genes CDR1 and CDR2. *Molecular Microbiology*, **62**(1): 170-186.

Hoot, S. J., Smith, A. R., Brown, R. P. and White, T. C. (2011). An A643V amino acid substitution in Upc2p contributes to azole resistance in well-characterized clinical isolates of *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **55**(2): 940-942.

Horn, D. L., Neofytos, D., Anaissie, E. J., Fishman, J. A., Steinbach, W. J., Olyaei, A. J. and Webster, K. M. (2009). Epidemiology and outcomes of candidemia in 2019 patients: Data from the prospective antifungal therapy alliance registry. *Clinical Infectious Diseases*, **48**(12): 1695-1703.

Hull, C. M., Bader, O., Parker, J. E., Weig, M., Gross, U., Warrilow, A. G. and Kelly, S. L. (2012). Two clinical isolates of *Candida glabrata* exhibiting reduced sensitivity to amphotericin B both harbor mutations in ERG2. *Antimicrobial Agents and Chemotherapy*, **56**(12): 6417-6421.

Hull, C. M., Parker, J. E., Bader, O., Weig, M., Gross, U., Warrilow, A. G. and Kelly, S. L. (2012). Facultative sterol uptake in an ergosterol-deficient clinical isolate of *Candida glabrata* harboring a missense mutation in ERG11 and exhibiting cross-resistance to azoles and amphotericin B. *Antimicrobial Agents and Chemotherapy*, **56**(8): 4223-4232.

Ingham, C. J., Boonstra, S., Levels, S., De Lange, M., Meis, J. F. and Schneeberger, P. M. (2012). Rapid susceptibility testing and microcolony analysis of *Candida* spp. cultured and imaged on porous aluminum oxide. *PLoS One*, **7**(3): e33818.

Jaspard, E., Macherel, D. and Hunault, G. (2012). Computational and statistical analyses of amino acid usage and physico-chemical properties of the twelve late embryogenesis abundant protein classes. *PLoS One*, **7**(5): e36968.

Jayatilake, J. A. M. S., Samaranayake, Y. H., Cheung, L. K. and Samaranayake, L. P. (2006). Quantitative evaluation of tissue invasion by wild type, hyphal and SAP mutants of *Candida albicans*, and non-*albicans Candida* species in reconstituted human oral epithelium. *Journal of Oral Pathology and Medicine*, **35**(8): 484-491.

Jiang, C., Dong, D., Yu, B., Cai, G., Wang, X., Ji, Y. and Peng, Y. (2012). Mechanisms of azole resistance in 52 clinical isolates of *Candida tropicalis* in China. *Journal of Antimicrobial Chemotherapy*, **68**(4): 778-785.

Kathiravan, M. K., Salake, A. B., Chothe, A. S., Dudhe, P. B., Watode, R. P., Mukta, M. S. and Gadhwe, S. (2012). The biology and chemistry of antifungal agents: A review. *Bioorganic and Medicinal Chemistry*, **20**(19): 5678-5698.

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. and Sternberg, M. J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nature protocols*, **10**(6): 845-858.

Klingspor, L., Tortorano, A. M., Peman, J., Willinger, B., Hamal, P., Sendid, B. and Ruhnke, M. (2015). Invasive *Candida* infections in surgical patients in intensive care units: A prospective, multicentre survey initiated by the European Confederation of Medical Mycology (ECMM)(2006–2008). *Clinical Microbiology and Infection*, **21**(1): 87-e1.

Klotz, S. A., Gaur, N. K., De Armond, R., Sheppard, D., Khardori, N., Edwards Jr, J. E. and El-Azizi, M. (2007). *Candida albicans* Als proteins mediate aggregation with bacteria and yeasts. *Medical Mycology*, **45**(4): 363-370.

Kojic, E. M. and Darouiche, R. O. (2004). *Candida* infections of medical devices. *Clinical Microbiology Reviews*, **17**(2): 255-267.

Kołaczkowska, A. and Kołaczkowski, M. (2016). Drug resistance mechanisms and their regulation in non-*albicans Candida* species. *Journal of Antimicrobial Chemotherapy*, **71**(6): 1438-1450.

Kontoyiannis D. P. and Lewis, R. E. (2002). Antifungal drug resistance of pathogenic fungi. *The Lancet*, **359**(9312): 1135– 1144.

Koonin, E. V. and Galperin, M. Y. (2003). Principles and methods of sequence analysis. In *Sequence—Evolution—Function* (pp. 111-192). Springer US.

Krcmery Jr, V., Mrazova, M., Kunova, A., Grey, E., Mardiak, J., Jurga, L. and West, D. (1999). Nosocomial Candidaemias due to species other than *Candida albicans* in cancer patients. *Supportive Care in Cancer*, **7**(6): 428-431.

Krogh-Madsen, M., Arendrup, M. C., Heslet, L. and Knudsen, J. D. (2006). Amphotericin B and caspofungin resistance in *Candida glabrata* isolates recovered from a critically ill patient. *Clinical Infectious Diseases*, **42**(7): 938-944.

Kuhn, D. M., George, T., Chandra, J., Mukherjee, P. K. and Ghannoum, M. A. (2002). Antifungal susceptibility of *Candida* biofilms: unique efficacy of amphotericin B lipid formulations and echinocandins. *Antimicrobial Agents and Chemotherapy*, **46**(6): 1773-1780.

Kumar, S., Stecher, G. and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, **33**(7): 1870-1874.

Kumar, T. A. (2013). CFSSP: Chou and Fasman secondary structure prediction server. *Wide Spectrum*, **1**(9):15-9.

LaFleur, M. D., Kumamoto, C. A. and Lewis, K. (2006). *Candida albicans* biofilms produce antifungal-tolerant persister cells. *Antimicrobial Agents and Chemotherapy*, **50**(11): 3839-3846.

Lai, C. C., Wang, C. Y., Liu, W. L., Huang, Y. T. and Hsueh, P. R. (2012). Time to positivity of blood cultures of different *Candida* species causing fungaemia. *Journal of Medical Microbiology*, **61**(5): 701-704.

Lamping, E., Ranchod, A., Nakamura, K., Tyndall, J. D., Niimi, K., Holmes, A. R. and Cannon, R. D. (2009). Abc1p is a multidrug efflux transporter that tips the balance in favor of innate azole resistance in *Candida krusei*. *Antimicrobial Agents and Chemotherapy*, **53**(2): 354-369.

Laniado-Labor´ın R. and Cabrales-Vargas M. N. (2009). Amphotericin B: side effects and toxicity. *Revista Iberoamericana de Micologia*, **26**(4): 223–227.

Linding, R., Russell, R. B., Neduva, V. and Gibson, T. J. (2003). GlobPlot: exploring protein sequences for globularity and disorder. *Nucleic acids research*, **31**(13): 3701-3708.

Liu, T. T., Znaidi, S., Barker, K. S., Xu, L., Homayouni, R., Saidane, S. and Rogers, P. D. (2007). Genome-wide expression and location analyses of the *Candida albicans* Tac1p regulon. *Eukaryotic Cell*, **6**(11): 2122-2138.

Liu, X., Zhao, H., Cao, W., Liu, Y., Zhang, C., Lan, X. and Ma, X. (2016). Bioinformatic prediction of the antigenic epitopes of recombinant ferritin of Echinococcus granulosus. *Molecular medicine reports*, **13**(1): 888-894.

Livermore, D. M. (2004). The need for new antibiotics. *Clinical Microbiology and Infection*, **10**(s4): 1-9.

Lorenzo, M. M. G. and Fenton, M. J. (2013). Immunobiology of influenza vaccines. *CHEST Journal*, **143**(2): 502-510.

Lortholary, O., Desnos-Ollivier, M., Sitbon, K., Fontanet, A., Bretagne, S. and Dromer, F. (2011). Recent exposure to caspofungin or fluconazole influences the epidemiology of candidemia: A prospective multicenter study involving 2,441 patients. *Antimicrobial Agents and Chemotherapy*, **55**(2): 532-538.

MacPherson, S., Akache, B., Weber, S., De Deken, X., Raymond, M. and Turcotte, B. (2005). *Candida albicans* zinc cluster protein Upc2p confers resistance to antifungal drugs and is an activator of ergosterol biosynthetic genes. *Antimicrobial Agents and Chemotherapy*, **49**(5): 1745-1752.

Mahmoudi Rad, M., Zafarghandi, A. S., Amel Zabihi, M., Tavallaee, M. and Mirdamadi, Y. (2012). Identification of *Candida* species associated with vulvovaginal candidiasis by multiplex PCR. *Infectious Diseases in Obstetrics and Gynecology*, 2012:1-5.

Mandal, S. M., Migliolo, L., Franco, O. L. and Ghosh, A. K. (2011). Identification of an antifungal peptide from Trapa natans fruits with inhibitory effects on *Candida tropicalis* biofilm formation. *Peptides*, **32**(8): 1741-1747.

Marie, C. and White, T. C. (2009). Genetic basis of antifungal drug resistance. *Current Fungal Infection Reports*, **3**(3): 163-169.

Martel, C. M., Parker, J. E., Bader, O., Weig, M., Gross, U., Warrilow, A. G. and Kelly, S. L. (2010). Identification and characterization of four azole-resistant erg3 mutants of *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **54**(11): 4527-4533.

Mastrolorenzo, A., Scozzafava, A. and Supuran, C. T. (2000). Antifungal activity of Ag (I) and Zn (II) complexes of aminobenzolamide (5-sulfanilylamido-1, 3, 4-thiadiazole-2-sulfonamide) derivatives. *Journal of Enzyme Inhibition*, **15**(6): 517-531.

Masur, H., Brooks, J. T., Benson, C. A., Holmes, K. K., Pau, A. K. and Kaplan, J. E. (2014). Prevention and treatment of opportunistic infections in HIV-infected adults and adolescents: Updated Guidelines from the Centers for Disease Control and Prevention, National Institutes of Health, and HIV Medicine Association of the Infectious Diseases Society of America. *Clinical infectious diseases*, **58**(9): 1308-1311.

Mattiuzzi, G. and Giles, F. J. (2005). Management of intracranial fungal infections in patients with haematological malignancies. *British Journal of Haematology*, **131**(3): 287-300.

Maurice, D. H., Ke, H., Ahmad, F., Wang, Y., Chung, J. and Manganiello, V. C. (2014). Advances in targeting cyclic nucleotide phosphodiesterases. *Nature reviews Drug discovery*, **13**(4): 290-314.

Morales, D. K., Jacobs, N. J., Rajamani, S., Krishnamurthy, M., Cubillos-Ruiz, J. R. and Hogan, D. A. (2010). Antifungal mechanisms by which a novel Pseudomonas aeruginosa phenazine toxin kills *Candida albicans* in biofilms. *Molecular Microbiology*, **78**(6): 1379-1392.

Morris, M. I. and Villmann, M. (2006). Echinocandins in the management of invasive fungal infections, Part 2. *American Journal of Health-System Pharmacy*, **63**(19): 1813-1820.

Naicker, S. D., Magobo, R. E., Zulu, T. G., Maphanga, T. G., Luthuli, N., Lowman, W. and Govender, N. P. (2016). Two echinocandin-resistant *Candida glabrata* FKS mutants from South Africa. *Medical Mycology Case Reports*, **11**, 24-26.

Nakayama, H., Izuta, M., Nakayama, N., Arisawa, M. and Aoki, Y. (2000). Depletion of the squalene synthase (ERG9) gene does not impair growth of *Candida glabrata* in mice. *Antimicrobial Agents and Chemotherapy*, **44**(9): 2411-2418.

Nakayama, H., Tanabe, K., Bard, M., Hodgson, W., Wu, S., Takemori, D. and Chibana, H. (2007). The *Candida glabrata* putative sterol transporter gene CgAUS1 protects cells against azoles in the presence of serum. *Journal of Antimicrobial Chemotherapy*, **60**(6): 1264-1272.

Nett, J., Lincoln, L., Marchillo, K., Massey, R., Holoyda, K., Hoff, B. and Andes, D. (2007). Putative role of β-1, 3 glucans in *Candida albicans* biofilm resistance. *Antimicrobial Agents and Chemotherapy*, **51**(2): 510-520.

Noël, T. (2012). The cellular and molecular defense mechanisms of the *Candida* yeasts against azole antifungal drugs. *Journal de Mycologie Médicale/Journal of Medical Mycology*, **22**(2): 173-178.

Nucci, M., Queiroz-Telles, F., Tobón, A. M., Restrepo, A. and Colombo, A. L. (2010). Epidemiology of opportunistic fungal infections in Latin America. *Clinical Infectious Diseases*, **51**(5): 561-570.

Oliveira, L. F. D., Jorge, A. O. C. and Santos, S. S. F. D. (2006). In vitro minocycline activity on superinfecting microorganisms isolated from chronic periodontitis patients. *Brazilian oral research*, **20**(3): 202-206.

Onishi, J., Meinz, M., Thompson, J., Curotto, J., Dreikorn, S., Rosenbach, M. and Cabello, A. (2000). Discovery of novel antifungal (1, 3)-β-D-glucan synthase inhibitors. *Antimicrobial Agents and Chemotherapy*, **44**(2): 368-377.

Ortega, M., Marco, F., Soriano, A., Almela, M., Martínez, J. A., López, J. and Mensa, J. (2011). *Candida* species bloodstream infection: epidemiology and outcome in a single institution from 1991 to 2008. *Journal of Hospital Infection*, **77**(2): 157-161.

Ostrosky-Zeichner, L., Kontoyiannis, D., Raffalli, J., Mullane, K. M., Vazquez, J., Anaissie, E. J., and Lau, W. (2005). International, open-label, noncomparative, clinical trial of micafungin alone and in combination for treatment of newly diagnosed and refractory candidemia. *European Journal of Clinical Microbiology and Infectious Diseases*, **24**(10): 654-661.

Oxman, D. A., Chow, J. K., Frendl, G., Hadley, S., Hershkovitz, S., Ireland, P. and Golan, Y. (2010). Candidaemia associated with decreased in vitro fluconazole susceptibility: Is *Candida* speciation predictive of the susceptibility pattern? *Journal of Antimicrobial Chemotherapy*, **65**(7): 1460-1465.

Pappas, P. G., Kauffman, C. A. andes, D. R., Clancy, C. J., Marr, K. A., Ostrosky-Zeichner, L., and Zaoutis, T. E. (2015). Clinical practice guideline for the management of candidiasis: 2016 update by the Infectious Diseases Society of America. *Clinical Infectious Diseases*, **62**(4): e1-e50.

Pappas, P. G., Rex, J. H., Sobel, J. D., Filler, S. G., Dismukes, W. E., Walsh, T. J. and Edwards, J. E. (2004). Guidelines for treatment of candidiasis. *Clinical Infectious Diseases*, **38**(2): 161-189.

Paul, S. and Moye-Rowley, W. S. (2014). Multidrug resistance in fungi: Regulation of transporter-encoding gene expression. *Frontiers in physiology*, 5:143

Percival, S. L., Bowler, P. G. and Russell, D. (2005). Bacterial resistance to silver in wound care. *Journal of Hospital Infection*, **60**(1): 1-7.

Pfaller, M. A. andes, D. R., Diekema, D. J., Horn, D. L., Reboli, A. C., Rotstein, C. and Azie, N. E. (2014). Epidemiology and outcomes of invasive candidiasis due to non-*albicans* species of *Candida* in 2,496 patients: data from the Prospective Antifungal Therapy (PATH) registry 2004–2008. *PLoS One*, **9**(7): e101510.

Pfaller, M. A., Castanheira, M., Lockhart, S. R., Ahlquist, A. M., Messer, S. A. and Jones, R. N. (2012). Frequency of decreased susceptibility and resistance to echinocandins among fluconazole-resistant bloodstream isolates of *Candida glabrata*. *Journal of Clinical Microbiology*, **50**(4): 1199-1203.

Pfaller, M. A., Diekema, D. J., Gibbs, D. L., Newell, V. A., Ellis, D. and Tullio, V. (2010). Results from the ARTEMIS DISK Global Antifungal Surveillance Study, 1997 to 2007: A 10.5-year analysis of susceptibilities of *Candida* species to fluconazole and voriconazole as determined by CLSI standardized disk diffusion. *Journal of Clinical Microbiology*, **48**(4): 1366-1377.

Pfaller, M. A., Rhomberg, P. R., Messer, S. A., Jones, R. N. and Castanheira, M. (2015). Isavuconazole, micafungin, and 8 comparator antifungal agents' susceptibility profiles for common and uncommon opportunistic fungi collected in 2013: temporal analysis of antifungal drug resistance using CLSI species-specific clinical breakpoints and proposed epidemiological cutoff values. *Diagnostic Microbiology and Infectious Disease*, **82**(4): 303-313.

Pires, R. H., Santos, J. M. D., Zaia, J. E., Martins, C. H. G. and Mendes-Giannini, M. J. S. (2011). *Candida parapsilosis* complex water isolates from a haemodialysis unit: Biofilm production and *in vitro* evaluation of the use of clinical antifungals. *Memórias do Instituto Oswaldo Cruz*, **106**(6): 646-654.

Rabideau, A. E. and Pentelute, B. L. (2015). A d-amino acid at the N-terminus of a protein abrogates its degradation by the N-end rule pathway. *ACS central science*, **1**(8): 423-430.

Ramage, G. and López-Rib, J. L. (2005). Techniques for antifungal susceptibility testing of *Candida albicans* biofilms. *Antifungal Agents: Methods and Protocols*, **118**:71-79.

Redding, S. W., Kirkpatrick, W. R., Saville, S., Coco, B. J., White, W., Fothergill, A. and Lopez-Ribot, J. (2003). Multiple patterns of resistance to fluconazole in *Candida glabrata* isolates from a patient with oropharyngeal candidiasis receiving head and neck radiation. *Journal of Clinical Microbiology*, **41**(2): 619-622.

Ribas, R., Daniela, A., Spolti, P., Del Ponte, E. M., Donato, K. Z., Schrekker, H., & Fuentefria, A. M. (2016). Is the emergence of fungal resistance to medical triazoles related to their use in the agroecosystems? A mini review. Brazilian Journal of Microbiology, **47**(4): 793-799.

Rossignol, T., Kelly, B., Dobson, C. and d'Enfert, C. (2011). Endocytosis-mediated vacuolar accumulation of the human ApoE apolipoprotein-derived ApoEdpL-W antimicrobial peptide contributes to its antifungal activity in *Candida albicans*. *Antimicrobial Agents and Chemotherapy*, **55**(10): 4670-4681.

Rybak, J. M., Marx, K. R., Nishimoto, A. T. and Rogers, P. D. (2015). Isavuconazole: Pharmacology, pharmacodynamics, and current clinical experience with a new triazole antifungal agent. Pharmacotherapy: *The Journal of Human Pharmacology and Drug Therapy*, **35**(11): 1037-1051.

Sanglard, D. and Odds, F. C. (2002). Resistance of *Candida* species to antifungal agents: molecular mechanisms and clinical consequences. *The Lancet Infectious Diseases*, **2**(2): 73-85.

Sanguinetti, M., Posteraro, B., Fiori, B., Ranno, S., Torelli, R. and Fadda, G. (2005). Mechanisms of azole resistance in clinical isolates of *Candida glabrata* collected during a hospital survey of antifungal resistance. *Antimicrobial Agents and Chemotherapy*, **49**(2): 668-679.

Sardi, J. C. O., Almeida, A. M. F. and Giannini, M. J. S. M. (2011). New antimicrobial therapies used against fungi present in subgingival sites—A brief review. *Archives of Oral Biology*, **56**(10): 951-959.

Schubert, S., Rogers, P. D. and Morschhäuser, J. (2008). Gain-of-function mutations in the transcription factor MRR1 are responsible for overexpression of the MDR1 efflux pump in fluconazole-resistant *Candida dubliniensis* strains. *Antimicrobial Agents and Chemotherapy*, **52**(12): 4274-4280.

Selmecki, A., Forche, A. and Berman, J. (2006). Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science*, **313**(5785): 367-370.

Senthilkumar, B., Sailo, S., Guruswami, G. and Nachimuthu, S. (2012). Prot-Prop: J-tool to predict the subcellular location of proteins based on physiochemical characterization. *Interdisciplinary sciences, computational life sciences*, **4**(4): 296.

Shao, L. C., Sheng, C. Q. and Zhang, W. N. (2007). Recent advances in the study of antifungal lead compounds with new chemical scaffolds. *Acta Pharmaceutica Sinica*, **42**(11): 1129-1136.

Sharifzadeh, A., Khosravi, A. R., Shokri, H., Jamnani, F. A., Hajiabdolbaghi, M. and Tamami, I. A. (2013). Oral microflora and their relation to risk factors in HIV+ patients with oropharyngeal candidiasis. *Journal de Mycologie Médicale/Journal of Medical Mycology*, **23**(2): 105-112.

Shi, Z., Sellers, J. and Moult, J. (2012). Protein stability and in vivo concentration of missense mutations in phenylalanine hydroxylase. *Proteins: Structure, Function, and Bioinformatics*, **80**(1): 61-70.

Shields, R. K., Nguyen, M. H., Press, E. G., Cumbie, R., Driscoll, E., Pasculle, A. W. and Clancy, C. J. (2015). Rate of FKS mutations among consecutive *Candida* isolates causing bloodstream infection. *Antimicrobial Agents and Chemotherapy*, **59**(12): 7465-7470.

Silva, S., Henriques, M., Martins, A., Oliveira, R., Williams, D. and Azeredo, J. (2009). Biofilms of non-*Candida albicans Candida* species: quantification, structure and matrix composition. *Sabouraudia*, **47**(7): 681-689.

Singh, N. (2001). Changing spectrum of invasive candidiasis and its therapeutic implications. *Clinical Microbiology and Infection*, **7**(2): 1-7.

Skiest, D. J., Vazquez, J. A., Anstead, G. M., Graybill, J. R., Reynes, J., Ward, D. and Isaacs, R. (2007). Posaconazole for the treatment of azole-refractory oropharyngeal and esophageal candidiasis in subjects with HIV infection. *Clinical Infectious Diseases*, **44**(4): 607-614.

Soria-Guerra, R. E., Nieto-Gomez, R., Govea-Alonso, D. O. and Rosales-Mendoza, S. (2015). An overview of bioinformatics tools for epitope prediction: implications on vaccine development. *Journal of biomedical informatics*, **53**, 405-414.

Souza, A. C. R., Fuchs, B. B., Pinhati, H. M., Siqueira, R. A., Hagen, F., Meis, J. F. and Colombo, A. L. (2015). *Candida parapsilosis* resistance to fluconazole: Molecular mechanisms and *in vivo* impact in infected Galleria mellonella larvae. *Antimicrobial Agents and Chemotherapy*, **59**(10): 6581-6587.

Spampinato, C. and Leonardi, D. (2013). *Candida* infections, causes, targets, and resistance mechanisms: traditional and alternative antifungal agents. *BioMed research international*, 2013: 1-13

Stafford, R. L., Zimmerman, E. S., Hallam, T. J. and Sato, A. K. (2014). A General Sequence Processing and Analysis Program for Protein Engineering. *Journal of chemical information and modeling*, **54**(10): 3020-3032.

Staib, P., Kretschmar, M., Nichterlein, T., Köhler, G., Michel, S., Hof, H. and Morschhäuser, J. (1999). Host-induced, stage-specific virulence gene activation in *Candida albicans* during infection. *Molecular Microbiology*, **32**(3): 533-546.

Stanciuc, A. M., Gaspar, A., Moldovan, L., Saviuc, C., Popa, M. and Măruţescu, L. (2010). In vitro antimicrobial activity of Romanian medicinal plants hydroalcoholic extracts on planktonic and adhered cells. *Roumanian Archives of Microbiology and Immunology*, **70**(1): 11-14.

Tang, J. L., Kung, H. C., Lei, W. C., Yao, M., Wu, U. I., Hsu, S. C. and Chou, W. C. (2015). High incidences of invasive fungal infections in acute myeloid leukemia patients receiving induction chemotherapy without systemic antifungal prophylaxis: A prospective observational study in Taiwan. *PloS one*, **10**(6): e0128410.

Tavakoli, M., Zaini, F., Kordbacheh, M., Safara, M., Raoofian, R. and Heidari, M. (2010). Upregulation of the ERG11 gene in *Candida krusei* by azoles. *Daru: journal of Faculty of Pharmacy, Tehran University of Medical Sciences*, **18**(4): 270 - 276.

Taweechaisupapong, S., Singhara, S., Lertsatitthanakorn, P. and Khunkitti, W. (2010). Antimicrobial effects *of Boesenbergia pandurata* and *Piper sarmentosum* leaf extracts on planktonic cells and biofilm of oral pathogens. *Pakistan Journal of Pharmaceutical Sciences*, **23**(2).

Torelli, R., Posteraro, B., Ferrari, S., La Sorda, M., Fadda, G., Sanglard, D. and Sanguinetti, M. (2008). The ATP-binding cassette transporter–encoding gene CgSNQ2 is contributing to the CgPDR1-dependent azole resistance of *Candida glabrata*. *Molecular Microbiology*, **68**(1): 186-201.

Tortorano, A. M., Kibbler, C., Peman, J., Bernhardt, H., Klingspor, L. and Grillot, R. (2006). Candidaemia in Europe: epidemiology and resistance. *International Journal of Antimicrobial Agents*, **27**(5): 359-366.

Uversky, V. N. (2013). A decade and a half of protein intrinsic disorder: biology still waits for physics. *Protein Science*, **22**(6): 693-724.

Vandeputte, P., Ferrari, S. and Coste, A. T. (2011). Antifungal resistance and new strategies to control fungal infections. *International Journal of Microbiology*, 2012.

Vandeputte, P., Larcher, G., Bergès, T., Renier, G., Chabasse, D. and Bouchara, J. P. (2005). Mechanisms of azole resistance in a clinical isolate of *Candida tropicalis*. *Antimicrobial Agents and Chemotherapy*, **49**(11): 4608-4615.

Vandeputte, P., Tronchin, G., Larcher, G., Ernoult, E., Bergès, T., Chabasse, D. and Bouchara, J. P. (2008). A nonsense mutation in the ERG6 gene leads to reduced susceptibility to polyenes in a clinical isolate of *Candida glabrata*. *Antimicrobial Agents and Chemotherapy*, **52**(10): 3701-3709.

Varland, S., Osberg, C. and Arnesen, T. (2015). N-terminal modifications of cellular proteins: The enzymes involved, their substrate specificities and biological effects. *Proteomics*, **15**(14): 2385-2401.

Varshavsky, A. (2011). The N-end rule pathway and regulation by proteolysis. *Protein science*, **20**(8): 1298-1345.

Vazquez, J. A., Peng, G., Sobel, J. D., Steele-Moore, L., Schuman, P. and Holloway, W. (2001). Evolution of antifungal susceptibility among *candida* species isolates recovered from human immunodeficiency virus—infected women receiving fluconazole prophylaxis. *Clinical Infectious Diseases*, **33**(7), 1069-1075.

Vermes, A., Guchelaar, H. J. and Dankert, J. (2000). Flucytosine: A review of its pharmacology, clinical indications, pharmacokinetics, toxicity and drug interactions. *Journal of Antimicrobial Chemotherapy*, **46**(2): 171-179.

Vermitsky, J. P. and Edlind, T. D. (2004). Azole resistance in *Candida glabrata*: coordinate upregulation of multidrug transporters and evidence for a Pdr1-like transcription factor. *Antimicrobial Agents and Chemotherapy*, **48**(10): 3773-3781.

Vermitsky, J. P., Earhart, K. D., Smith, W. L., Homayouni, R., Edlind, T. D. and Rogers, P. D. (2006). Pdr1 regulates multidrug resistance in *Candida glabrata*: gene disruption and genome-wide expression studies. *Molecular Microbiology*, **61**(3): 704-722.

Vermitsky, J. P., Self, M. J., Chadwick, S. G., Trama, J. P., Adelson, M. E., Mordechai, E. and Gygax, S. E. (2008). Survey of vaginal-flora *Candida* species isolates from women of

different age groups by use of species-specific PCR detection. *Journal of Clinical Microbiology*, **46**(4): 1501-1503.

Verstrepen, K. J. and Klis, F. M. (2006). Flocculation, adhesion and biofilm formation in yeasts. *Molecular Microbiology*, **60**(1): 5-15.

Vidigal, P. G. and Svidzinski, T. I. E. (2009). Yeasts in the urinary and respiratory tracts: Is it a fungal infection or not? *Jornal Brasileiro de Patologia e Medicina Laboratorial*, **45**(1): 55-64.

Vincent, J. L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C. D. and Reinhart, K. (2009). International study of the prevalence and outcomes of infection in intensive care units. *Jama*, **302**(21): 2323-2329.

Wang, E., Farmakiotis, D., Yang, D., McCue, D. A., Kantarjian, H. M., Kontoyiannis, D. P. and Mathisen, M. S. (2015). The ever-evolving landscape of Candidaemia in patients with acute leukaemia: Non-susceptibility to caspofungin and multidrug resistance are associated with increased mortality. *Journal of Antimicrobial Chemotherapy*, **70**(8): 2362-2368.

Wang, J., Sung, W. K., Krishnan, A. and Li, K. B. (2005). Protein subcellular localization prediction for Gram-negative bacteria using amino acid subalphabets and a combination of multiple support vector machines. *BMC Bioinformatics*, **6**(1): 174.

Wass, M. N., Kelley, L. A. and Sternberg, M. J. (2010). 3DLigandSite: predicting ligand-binding sites using similar structures. *Nucleic acids research*, **38**(suppl_2): W469-W473.

Whaley, S. G., Berkow, E. L., Rybak, J. M., Nishimoto, A. T., Barker, K. S. and Rogers, P. D. (2017). Azole antifungal resistance in *Candida albicans* and emerging non-*albicans Candida* species. *Frontiers in microbiology*, **7**, 2173.

Wheeler, D. L., Barrett, T., Benson, D. A., Bryant, S. H., Canese, K., Chetvernin, V. and Feolo, M. (2007). Database resources of the national center for biotechnology information. *Nucleic acids research*, **36**(suppl_1): D13-D21.

Wisplinghoff, H., Seifert, H., Wenzel, R. P. and Edmond, M. B. (2006). Inflammatory response and clinical course of adult patients with nosocomial bloodstream infections caused by *Candida* spp. *Clinical Microbiology and Infection*, **12**(2): 170-177.

Xiang, M. J., Liu, J. Y., Ni, P. H., Wang, S., Shi, C., Wei, B. and Ge, H. L. (2013). Erg11 mutations associated with azole resistance in clinical isolates of *Candida albicans*. *FEMS Yeast Research*, **13**(4): 386-393.

Xu, Y., Sheng, F., Zhao, J., Chen, L. and Li, C. (2015). ERG11 mutations and expression of resistance genes in fluconazole-resistant *Candida albicans* isolates. *Archives of Microbiology*, **197**(9): 1087-1093.

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J. and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nature methods*, **12**(1): 7-8.

Yao, B., Zhang, L., Liang, S. and Zhang, C. (2012). SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. PloS one, **7**(9): e45152.

Yapar, N. (2014). Epidemiology and risk factors for invasive candidiasis. *Therapeutics and Clinical Risk Management*, **10**; 95–105.

Yi, S., Sahni, N., Daniels, K. J., Lu, K. L., Srikantha, T., Huang, G. and Soll, D. R. (2011). Alternative mating type configurations (a/α versus a/a or α/α) of *Candida albicans* result in alternative biofilms regulated by different pathways. *PLoS Biology*, **9**(8): e1001117.

Yu, N. Y., Wagner, J. R., Laird, M. R., Melli, G., Rey, S., Lo, R. and Brinkman, F. S. (2010). PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics*, **26**(13): 1608-1615.

Zavrel, M., Hoot, S. J. and White, T. C. (2013). Comparison of sterol import under aerobic and anaerobic conditions in three fungal species, *Candida albicans*, *Candida glabrata*, and *Saccharomyces cerevisiae*. *Eukaryotic Cell*, **12**(5): 725-738.

Zhang, L., Xiao, M., Watts, M. R., Wang, H., Fan, X., Kong, F. and Xu, Y. C. (2015). Development of fluconazole resistance in a series of *Candida parapsilosis* isolates from a persistent candidemia patient with prolonged antifungal therapy. *BMC Infectious Diseases*, **15**(1): 340.

Zhang, Z., Schwartz, S., Wagner, L. and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *Journal of Computational Biology*, **7**(1-2): 203-214.

Zimbeck, A. J., Iqbal, N., Ahlquist, A. M., Farley, M. M., Harrison, L. H., Chiller, T. and Lockhart, S. R. (2010). FKS mutations and elevated echinocandin MIC values among *Candida glabrata* isolates from US population-based surveillance. *Antimicrobial Agents and Chemotherapy*, **54**(12): 5042-5047.

APPENDICES

Appendix 1: Sequences Retrieved from NCBI for Phylogenic Analysis

>KM875721.1 USA

ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875714.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTAATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>AB006854.1 Japan
AGGGAATTCAATCGTTATTCTTTCCATATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAA
GAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAG
TGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATAT
TCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATG
GTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATT
AGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGAT
GTTTCTGCTGAAGAAGCTTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTC
CAAATTCTAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATA
TGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAA
ACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTAT
TTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTT
TACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAG
AAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTG
ATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGC
TAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTA
CATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAG
GTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGA
AACTCTTAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAA
ACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGAT
ATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTC
ATTTAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTA
CCATTTGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAA
CTACTTTTGTTTATAACTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAAT
GGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAAC
GGCAACTTTCTTTCGATTCAGTGTTCTGATTG

>AF153849.1 Belgium
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGCTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194171.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>AF153846.1 Belgium
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGTTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875712.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TTTGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACY
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194175.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGSGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM609931.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTCATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTAGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAGGAGAAACTTGTATGTTTTAA

>KM875715.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGACCGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>GQ202079.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGCGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>GQ202072.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTCATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTAGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>EU819550.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>EU819549.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>AB006856.1 Japan
AGGGAATTCAATCGTTATTCTTTCCATATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAA
GAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAG
TGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATAT
TCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATG
GTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATT
AGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGAT
GTTTCTGCTGAAGAAGCTTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTTTGATTGTC
CAAATTCTAGATTAATGGAACAAAGAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATA
TGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAA
ACTCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTAT
TTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTT
TACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAG
AAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTG
ATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGC
TAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTA
CATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAG
GTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGA
AACTCTTAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAA
ACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGAT
ATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTC
ATTTAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTA
CCATTTGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAA
CTACTTTTGTTTATAACTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAAT
GGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAAC
GGCAACTTTCTTTCGATTCAGTGTTCTGATTG

>HM194174.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGACCGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCYAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>AB006855.1 Japan
AGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTG
TTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTC
ATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGT
CAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAG
GGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGT
TTCTGCTGAAGAAGCTTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTCATGATTGTCCA
AATTCTAGATTAATGGAACAAAAAAAACTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATG
TTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAAC
TCATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTT
GGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTA
CCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAA
AATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGAT
TTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTA
ATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACA
TTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGT
GGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAA
CTCTTAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAAC
CAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATAT
TTTGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCAT
TTAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACC
ATTTGGTGGTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACT
ACTTTTGTTTATAACTTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGG
TGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGG
CAACTTTCTTTCGATTCAGTGTTCTGATTG

>HM194173.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGCGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCYAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875720.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAGAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGACCGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875716.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTTTGATTGTCCAAATTCCAGATTAATG
GAACAAAGAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194178.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTYATTTATCAAGAAGTTGTTGAATTRTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTYAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTAYGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAY
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>EU980038.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAAYTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>EU848303.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTYAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAACATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM609929.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCG
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTCCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAGAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTGATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGGAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTAGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTCTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194176.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAWGC
TTATAARCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCACGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCYAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KC676660.1 Tunisia
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGACCGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTGTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194177.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTYGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGAWGC
TTATAARCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCTAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAATTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGACCGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTYATTTATCAAGAAGTTGTTGAATTRTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCYAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTGGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194154.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTYAATGCTAAATTATCTGATGTTTCTGCTGAAGAWGC
TTATAARCATTTAACTACTCCAGTTTTCGGTAMAGGGGTTATTTATGATTGTCCAAATTCYAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAAYTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTRTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTYAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTRGGAACCATTTTAACTACTTTTGTTTATAAY
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>EU885933.1 France
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTYAATGCTAAATTATCTGATGTTTCTGCTGAAGAWGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAMAGGGGTTATTTATGATTGTCCAAATTCYAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATRTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAAYTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTRTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTYAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTRGGAACCATTTTAACTACTTTTGTTTATAAY
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>GU371855.1 India
AATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATA
TTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGT
TGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTA
GGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAG
TGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATC
ATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGT
CCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGCTTATAAACATT
TAACTACTCCAGTTTTCGGTACAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATGGAACAAAAAAA
ATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTTTG
AATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAA
CTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGA
CCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAAT
TTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAA
TTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTC
AACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGT
GGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAG
ATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGA
AGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATATGCCATTACAT
TCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCCAAAAGGTCATT
ATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGATCC
AACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTAT
GGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACATAGATGTA
TTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGAC
TATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAA
ATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTG
ATTGTTTTCATTTTGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATAT
AGAATAGAAATT

>GU371859.1 India
TCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATATTACTTGTC
TTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGAT
GGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTT
AGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTG
CAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATG
TTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGT
TTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCA
GATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGA
GAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTAT
GAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACC
GTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTA
CCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGA
ACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAA
TGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCT
TGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGA
AAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAA
CTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTAT
ATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGA
AGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTG
ATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACATAGATGTATT
GGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACTATTGATGG
TTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAA
GAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTGTTACTTAG
TTGGATTAACATATATACACATATACATACAAATATATG

>GU371857.1 India
TTTCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCAT
ATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAAC
TGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTG
TTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGG
TTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT
TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAAT
TATCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGT
CCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCC
TAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTG
CCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGA
ATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAA
TTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAAC
TGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGAT
GGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTC
TACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAAT
TATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACT
ATTAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGA
AACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTG
ATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCT
GATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACA
TAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGA
CTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATT
TGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTT
TGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAG

>GU371854.1 India
TCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATATTACTTGTC
TTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGAT
GGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTT
AGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTG
CAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATG
TTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGT
TTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCA
GATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGA
GAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTAT
GAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACC
GTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTA
CCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGA
ACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAA
TGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCT
TGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGA
AAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAA
CTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTAT
ATTGTTCCAAAAGGTCATTATGTTTTAGTTTTTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGA
AGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTG
ATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACATAGATGTATT
GGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACTATTGATGG
TTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAA
GAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTGTTACTTAG
TTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAGAAATT

>GU371851.1 India
TTTCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCAT
ATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAAC
TGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTG
TTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGG
TTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT
TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAAT
TATCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGT
CCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCC
TAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTG
CCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGA
ATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAA
TTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAAC
TGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGAT
GGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTC
TACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAAT
TATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACT
ATTAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGA
AACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTG
ATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCT
GATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACA
TAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGA
CTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATT
TGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTT
TGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAGAAATT

>X13296.1 USA
ATCTTACTTCTTTCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTA
TTCTTTCCATATTACTTGTCTTCTTTTTATTATATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTA
TTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGG
GTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTG
GATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATG
GTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTT
AATGCTAAATTATCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTAT
TTATGATTGTCCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAA
GATATGTTCCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACT
CATGGGGTTGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGA
AATGAGAAGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAA
GAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAAC
TTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATA
CTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAA
GTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGT
CAATAACACTATTAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAA
GAATCCCTGAAACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAA
AGATATTTTGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATT
TAACTCTTCTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTG
GTGGTAGACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGC
AGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATT
GTTTTCATTTTGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAGAAA
TTA

>GU371858.1 India
ATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATATTACTTGTCTTCTTTTTAT
TATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGATGGCATTAATT
ATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAA
TATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATA
TGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGA
AAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAA
GATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATGGA
ACAAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAGAAGAAATTT
TGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTATGAAAACTCAA
CCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACCGTTCATTTGC
TCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTACCTCATTATT
GGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGAT
ATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCA
AGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGT
TACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGT
GATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAAT
GCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCCAA
AAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAAGATTTTGAT
CCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTGATTATGGGTT
TGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAAACATAGATGTATTGGGGAACAAT
TTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACTATTGATGGTTATAAAGTG
CCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTG
TATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTGTTACTTAGTTGGATTAAC
ATATATACACATATACATACAAATATATG

>GU371856.1 India
TCTTTCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCC
ATATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAA
ACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATT
TGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTT
GGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTA
TTTTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAA
ATTATCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATT
GTCCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTT
CCTAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGT
TGCCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAA
GAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCT
AATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAA
ACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAG
ATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCT
TCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGA
ATTATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACA
CTATTAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCT
GAAACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTT
TGATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTT
CTGATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAAA
CATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATG
GACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCA
TTTGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCAT
TTTGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAG

>GU371852.1 India
TCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATATTACTTGTC
TTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAACTGTCATTGAT
GGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTTTACAACTT
AGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTG
CAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATG
TTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGT
TTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCA
GATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGA
GAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAATGTTAT
GAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAATTTTTGACC
GTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATTTACCTTTA
CCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACTGAGAAGAGA
ACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGGTGTGAAAA
TGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCT
TGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGA
AAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAA
CTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTAT
ATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGA
AGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATGAAGTTG
ATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAAACATAGATGTATT
GGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACTATTGATGG
TTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTGGGAAAAAA
GAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTGTTACTTAG
TTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAGAAATT

>GU371850.1 India
TCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCATAT
TACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAACTG
TCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTGTT
TACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGGTT
TGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTT
CATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTA
TCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCC
AAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTA
AGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCC
AATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGAAT
TTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAATT
TACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAACTG
AGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGATGG
TGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTCTA
CTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTA
TTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTAT
TAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAA
CCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGAT
AACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTATTTCATTTAACTCTTCTGA
TGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACATA
GATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGACT
ATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATTTG
GGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTTTG
TTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAG

>GU371853.1 India
TTTCTTTCAATCTTTAATAAATCAATTTTTATATATAAATAGACAAAGAAAGGGAATTCAATCGTTATTCTTTCCAT
ATTACTTGTCTTCTTTTTATTATATATATAAGTTTCTTTTCAAGAAGATCATAACTCAATATGGCTATTGTTGAAAC
TGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCAGTATATTATTAGGGGTTCCATTTG
TTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAGAGCTCCATTAGTGTTTTATTGGATTCCTTGG
TTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAATTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATT
TTCATTTATGTTATTAGGGAAAATTATGACGGTTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAAT
TATCTGATGTTTCTGCTGAAGATGCTTATAAACATTTAACTACTCCAGTTCTCGGTAAAGGGGTTATTTATGATTGT
CCAAATTCCAGATTAATGGAACAAAAAAAATTTGCTAAATTTGCTCTGACTACTGATTCATTTAAAAGATATGTTCC
TAAGATTAGAGAAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTG
CCAATGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGAAGA
ATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTGTTTTCCCTAA
TTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTATATGAAAGAAATTAAAC
TGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTATTGATTCATTCAACTTATAAAGAT
GGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTATTCTTATGGGTGGTCAACATACTTCTGCTTC
TACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACCTCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAAT
TATTGAAAGAAAAAGGTGGTGATTTGAATGATTTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACT
ATTAAGGAAACTCTCAGAATGCATATGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGA
AACCAATTATATTGTTCCAAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTG
ATAACCCTGAAGATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCT
GATGAAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAGACA
TAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAATTTAAGATGGA
CTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTGAACCAGCAGAAATCATT
TGGGAAAAAAGAGAAACTTGTATGTTTTAATAAAACGGCAACTTTCTTTCGATTCAGTGTTCTGATTGTTTTCATTT
TGTTACTTAGTTGGATTAACATATATACACATATACATACAAATATATGATACATATAGAATAGAAATT

>XM_711668.2 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>FJ403378.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTCAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTRTTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM609923.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAGTTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAGGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM609920.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGTTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875723.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGAGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875722.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGAGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>HM194203.1 China
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTYAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCYAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>FJ403379.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGAAGC
TTATAAGCATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTTAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>FJ002303.1 Italy
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCTAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGATTATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAC
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

>KM875726.1 USA
ATGGCTATTGTTGAAACTGTCATTGATGGCATTAATTATTTTTTGTCCCTTAGTGTTACACAACAGATCA
GTATATTATTAGGGGTTCCATTTGTTTACAACTTAGTATGGCAATATTTATATTCATTAAGAAAAGATAG
AGCTCCATTAGTGTTTTATTGGATTCCTTGGTTTGGTTCTGCAGCTTCATATGGTCAACAACCTTATGAA
TTTTTCGAATCATGTCGTCAAAAGTATGGTGATGTATTTTCATTTATGTTATTAGGGAAAATTATGACGG
TTTATTTAGGTCCAAAAGGTCATGAATTTGTTTTTAATGCTAAATTATCTGATGTTTCTGCTGAAGATGC
TTATAAACATTTAACTACTCCAGTTTTCGGTAAAGGGGTTATTTATGATTGTCCAAATTCCAGATTAATG
GAACAAAAAAAATTTGCTAAATTTGCTTTGACTACTGATTCATTTAAAAGATATGTTCCTAAGATTAGAG
AAGAAATTTTGAATTATTTTGTTACTGATGAAAGTTTCAAATTGAAAGAAAAAACTCATGGGGTTGCCAA
TGTTATGAAAACTCAACCAGAAATTACTATTTTCACTGCTTCAAGATCTTTATTTGGTGATGAAATGAGA
AGAATTTTTGACCGTTCATTTGCTCAACTATATTCTGATTTAGATAAAGGTTTTACCCCTATTAATTTTG
TTTTCCCTAATTTACCTTTACCTCATTATTGGAGACGTGATGCTGCTCAAAAGAAAATCTCTGCTACTTA
TATGAAAGAAATTAAACTGAGAAGAGAACGTGGTGATATTGATCCAAATCGTGATTTAATTGATTCCTTA
TTGATTCATTCAACTTATAAAGATGGTGTGAAAATGACTGATCAAGAAATTGCKAATCTTTTAATTGGTA
TTCTTATGGGTGGTCAACATACTTCTGCTTCTACTTCTGCTTGGTTCTTGTTACATTTAGGTGAAAAACC
TCATTTACAAGATGTTATTTATCAAGAAGTTGTTGAATTATTGAAAGAAAAAGGTGGTGATTTGAATGAT
TTGACTTATGAAGATTTACAAAAATTACCATCAGTCAATAACACTATTAAGGAAACTCTCAGAATGCATA
TGCCATTACATTCTATTTTTAGAAAAGTTACTAACCCATTAAGAATCCCTGAAACCAATTATATTGTTCC
AAAAGGTCATTATGTTTTAGTTTCTCCAGGTTATGCTCATACTAGTGAAAGATATTTTGATAACCCTGAA
GATTTTGATCCAACTAGATGGGATACTGCTGCTGCCAAAGCTAATTCTGTTTCATTTAACTCTTCTGATG
AAGTTGAATATGGGTTTGGGAAAGTTTCTAAAGGGGTTTCTTCACCTTATTTACCATTTGGTGGTGGTAG
ACATAGATGTATTGGGGAACAATTTGCTTATGTTCAATTAGGAACCATTTTAACTACTTTTGTTTATAAT
TTAAGATGGACTATTGATGGTTATAAAGTGCCTGACCCTGATTATAGTTCAATGGTGGTTTTACCTACTG
AACCAGCAGAAATCATTTGGGAAAAAAGAGAAACTTGTATGTTTTAA

Appendix 2: Identification and Characterization of Yeast Samples using Conventional methods

S/N	Sample	Media Growth/ Yeast Microscopic Characterization (germ tube test)
1	HHVS-001	+VE (+ve to Germ tube test)
2	HHVS-002	+VE (+ve to Germ tube test)
3	HHVS-003	Growth Absent (-VE)
4	HHVS-004	-VE
5	HHVS-005	-VE
6	HHVS-006	+VE (+ve to Germ tube test)
7	HHVS-007	+VE (+ve to Germ tube test)
8	HHVS-008	-VE
9	HHVS-009	-VE
10	HHVS-010	+VE (+ve to Germ tube test)
11	HHVS-011	+VE (+ve to Germ tube test)
12	HHVS-012	-VE
13	HHVS-013	-VE
14	HHVS-014	-VE
15	HHVS-015	-VE
16	HHVS-016	-VE
17	HHVS-017	+VE (+ve to Germ tube test)
18	HHVS-018	-VE
19	HHVS-019	+VE (+ve to Germ tube test)
20	HHVS-020	+VE (+ve to Germ tube test)
21	HHVS-021	+VE (+ve to Germ tube test)
22	HHVS-022	-VE
23	HHVS-023	-VE
24	HHVS-024	-VE
25	HHVS-025	+VE (+ve to Germ tube test)
26	HHVS-026	-VE
27	HHVS-027	+VE (+ve to Germ tube test)
28	HHVS-028	+VE (+ve to Germ tube test)
29	HHVS-029	+VE (+ve to Germ tube test)
30	HHVS-030	-VE
31	HHVS-031	-VE
32	HHVS-032	-VE
33	HHVS-033	+VE (+ve to Germ tube test)
34	HHVS-034	-VE
35	HHVS-035	+VE (+ve to Germ tube test)
36	HHVS-036	+VE (+ve to Germ tube test)
37	HHVS-037	+VE (+ve to Germ tube test)
38	HHVS-038	+VE (+ve to Germ tube test)
39	HHVS-039	+VE (+ve to Germ tube test)
40	HHVS-040	+VE (+ve to Germ tube test)

Appendix 2 Cont'd

S/N	Sample	Media Growth/ Yeast Microscopic Characterization (germ tube test)
41	HHVS-041	+VE (+ve to Germ tube test)
42	HHVS-042	-VE
43	HHVS-043	-VE
44	HHVS-044	+VE (+ve to Germ tube test)
45	HHVS-045	-VE
46	HHVS-046	-VE
47	HHVS-047	-VE
48	HHVS-048	-VE
49	HHVS-049	+VE (+ve to Germ tube test)
50	HHVS-050	+VE (+ve to Germ tube test)
51	HHVS-051	-VE
52	HHVS-052	+VE (+ve to Germ tube test)
53	HHVS-053	+VE (+ve to Germ tube test)
54	HHVS-054	-VE
55	HHVS-055	+VE (+ve to Germ tube test)
56	HHVS-056	-VE
57	HHVS-058	-VE
58	DHVS-001	+VE (+ve to Germ tube test)
59	DHVS-002	-VE
60	DHVS-003	-VE
61	DHVS-004	-VE
62	DHVS-005	+VE (+ve to Germ tube test)
63	DHVS-006	+VE (+ve to Germ tube test)
64	DHVS-007	+VE (+ve to Germ tube test)

HHVS - Human High Vaginal Swab; (28 present, 29 absent) total = 57
DHVS - Dog High Vaginal Swab; Dogs 4 present, 3 absent, total = 7