**Iraqi Journal of Biotechnology, 2024, Vol. 23, No. 1, 213-228**

# Phylogenetic Analysis of Pathogenic *Candida* Species Isolated from Iraqi Patients

**Zahraa A. Al-Ameri[1], Muna T. Al-Musawi[1]**

[1]Department of Biology, College of Science for Women, University of Baghdad.

**Received: September 4, 2023 / Accepted: October 31, 2023/ Published: September 23, 2024**

**Abstract:** Pathogenic species of Candida have extensively contributed to the raising mortality, morbidity and globally increase medical prices of healthcare-related infectious diseases. The current study was focused on isolation and identification of *Candida* spp. from different clinical specimens and their molecular phylogenetic analyses. The study was included clinical specimens (100) from different parts of the patient's body, such as (ear, mouth. vagina, skin, blood and urine) who attended to Gazi Al Hariri hospital, Baghdad teaching hospital in Medical City and Al-Yarmouk teaching hospital. The isolation of *Candida* spp. using selective media and vitek system. Then, a full inclusive tree, including the observed variant, was built by the neighbour-joining method and visualized as a circular cladogram using the iTOL suit. The findings of isolation and identification of *Candida* spp. in studied samples were recorded that the highest isolate was *C. parapsilosis* (27.7%) followed by *C glabrata* (22.2%), while the lowest isolates were *C. tropicalis* and *C. krusei* (5.50%) each one, *Candida* spp. also identified by ITS which results showed the presence of sizes ranging 550 - 625 bp depending on the difference in the lengths of the region (ITS 1 and ITS4) for yeast species, so they produce different sizes of DNA pieces. Sequencing reactions showed the accurate identity of the investigated samples and revealed that S1, S8, and S9 were homologous to *C. parapsilosis*, S2, S6, and S13 were similar to *C. albicans,* S3 was similar to *C. xylopsoci* while S4, S5, S7, S10, S11, and S12 were similar to *C. glabrata* (The presence of a total of 11 nucleic acid variants compared with the referring sequences of the *Candida* sequences was demonstrated. The identified variants were one deletion (g.492-494Adel) in S1, two substitutions (g.168A>G and g.233C>A) in S13, one substitution (g.83A>C) in S3, one substitution (g.116T>C) in S11, and four substitutions and one deletion (g.339C>A, g.353G>A, g.618G>A, g.356-359GGdel) in S4, S5, S7, S11, S12, one substitution in S7 (g.345C>G), one substitution (g.805T>C) in S4, S5, and S7. Meanwhile, the rest of the samples had shown a complete homology with the corresponding sequences and did not exhibit any detectable nucleic acid variations in comparison with the same reference sequences. This study suggests possible employment for these ribosomal amplicons to discriminate between the phylogenetic diversity among the other implemented tools.

**Keywords:** *Candida* spp., non-*albicans Candida*, molecular diagnosis, phylogenetic diversity

**Corresponding author**: (Email: zahraali1471997@gmail.com).

## Introduction

Pathogenic species of *Candida* have extensively contributed to the raising mortality, morbidity and globally increase medical prices of healthcare-related infectious diseases(1). Although *C. albicans* is still the most common cause of candidiasis, the widespread usage of certain antifungals in healthcare of human has

shifted the epidemiology of candidiasis towards non-*albicans Candida* spp., which are inherently less vulnerable to treatments (2). Systemic candidiasis, peritonitis and hepatosplenic candidiasis, are only a few of the invasive disorders that may be caused by *Candida* spp. There are more than 200 different species of *Candida*, although *C. tropicalis*, *C. parapsilosis*, *C. krusei*, *C. glabrata*, *C. albicans* and *C. dubliniensis*, are the most prevalent ones seen in medical settings. It's well knowledge that the rapid growth of certain strains of *Candida* has paved the way for the spread of other dangerous bacteria into human hosts (3). *Candida* species, on the other hand, are well-known for their biofilm-forming abilities, their morphological plasticity, and their capacity to rapidly transition between phenotypes. *Candida* spp. fungal infection rates have been observed to be on the rise (4). *Candida* is normal flora of the mouth and may be the root cause of several types of oral illnesses (5). However, impressive discrimination of multiple *Candida* species from one other in the oral cavity has been linked to an effective involvement of the internal spacer (ITS)-based ribosomal RNA (6). Distinguishing between infective *Candida* spp. is of great clinical importance because the kind of *Candida* spp. found may dictate the most effective course of therapy for oral infection (7). Molecular methods are recommended for determining the species of *Candida*. These approaches are costly, but their speed and precision cannot be disputed. Molecular techniques such as real-time polymerase chain reaction (real-time PCR), DNA fingerprinting, single-strand conformational polymorphism (SSCP), matrix-associated laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS), DNA sequencing, polymerase chain reaction (PCR), nested polymerase chain reaction (Nested PCR), random amplification of polymorphic DNA (RAPD), PCR-restriction length fragment polymorphisms (PCR-RFLP) and amplified fragment length polymorphism (AFLP), multiplex PCR techniques, are appropriate for identifying and diagnosing *Candida* spp. (8).

## Material and methods

### Sample collection

During the period of Dec. 2021 until February 2022, One hundred specimen collected(ear, mouth , vagina ,skin, blood and urine) from paitents attended to Gazi Al Hariri hospital, Al-Yarmouk teaching hospital and Baghdad teaching Hospital in Medical city.

### Isolation and Identic of *Candida* spp.

Chromagar medium (*Candida* agar) this a medium used for the purpose of detecting the species of the genus *Candida* and according to the colors that appear on the above medium (9). This medium is novel and utilized to differentiate *C. dubliniensis* from *C. albicans* , used to distinguish between types of *Candida* as show in figure(1) this medium was prepared according to the manufacturer's instructions, which involved dissolving the powdered medium of 47 g in 100 ml D.W., then letting it boil in a water bath, then pouring the medium into Petri dishes and holding it until use.

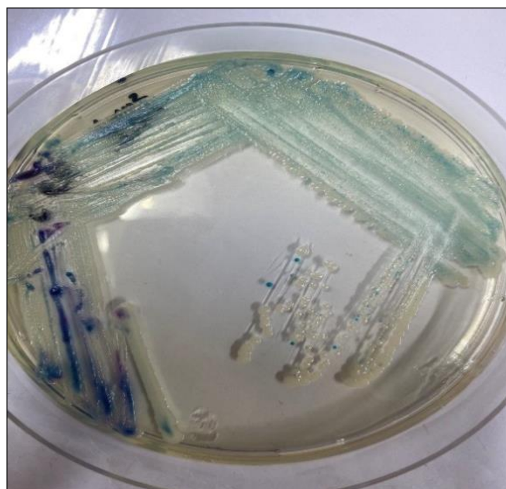


**Figure (1): The *Candida* spp. on Chromagar medium at 37° C after 24-48 hr. incubation. Mixed sample. blue to purple: *C. tropicals*, Green: *C. albicans*, white: *C. glabrata*.**

The Vitek-2 YST System was used for confirmation diagnostics. VITEK®2 ID cards are self-contained, disposable cards developed for use with the VITEK®2 family of equipment (10).

## DNA Extraction

Genomic DNA was isolated from Fungal growth according to the protocol of ABIOpure Extraction.

## Quantitation of DNA

Quantus Fluorometer was used to detect the concentration of extracted DNA in order to detect the goodness of samples for downstream applications. For 1 µl of DNA, 200µl of diluted Quantifluor Dye was mixed. After 5min incubation at room temperature, DNA concentration values were detected.

## Preparation of primer and reaction mixture

The primer under study was supplied by the Macrogen Company in a lyophilized form. The lyophilized primer in table (3-6) was dissolved in 300 µl nuclease-free water to give a final concentration of 100 pmol/ µl as a stock solution. A working solution of these primers was prepared by adding 10 µl of primer stock solution (stored at freezer -20 ºC) to 90 µl of nuclease-free water to obtain a working primer solution of 10 pmol/ µl as Table (1).

**Table(1):Primer sequence of ITS gene and their concentration.**

| Primer | Primer sequence (5 to 3) | Product size (bp) |
|---|---|---|
| ITS-F | 5`TCCGTAGGTGAACCTGCGG 3` | Variable |
| ITS-R | 3`TCCTCCGCTTATTGATATGC 5` | Variable |

**Table (2): PCR Master mix to detect its gene of *Candida* spp.**

| Components | Concentration | Volume(µl) |
|---|---|---|
| Master Mix | 2 x | 12.5 |
| F- primer | 10 µM | 1 |
| R- primer | 10 µM | 1 |
| Nuclease Free Water | | 7.5 |
| DNA | | 3 |
| Total volume | | 25 |

**Table (3): PCR program to detect ITS gene of *Candida* spp.**

| Step. | Temp. (°C) | Time (m: s) | Cycle |
|---|---|---|---|
| Initial Denaturation | 95 | 05:00 | 1 |
| Denaturation | 95 | 00:30 | 30 |
| Annealing | 55 | 00:30 | |
| Extension | 72 | 00:30 | |
| Final extension | 72 | 07:00 | 1 |
| Hold | 10 | 10:00 | 1 |

**Nucleic acids sequencing of PCR amplicons**

Commercial forward and reverse sequencing of the resolved PCR amplicons was performed according to the manufacturer's protocols (Macrogen Inc., Geumchen, Seoul, South Korea). To rule out the possibility that the variations and annotation are the result of sequencing or PCR objects, only clean chromatograms were analysed using ABI (Applied Biosystems) sequence files. The virtual locations and other features of the returned fragments of PCR were determined via comparing the observed sequences of nucleic acids of local specimens with the retrieved sequences of nucleic acids.

**Sequencing data interpretation**

Bio Edit Sequence Alignment Editor Software Version 7.1 (DNASTAR, Madison, WI, USA) was used to edit, align, and analyse the sequencing findings of the targeted samples' PCR products with their corresponding sequences in the reference database. PCR amplicons and their associated genomic locations were numbered to coincide with the variations found in each sample sequenced. PCR amplicons and their associated genomic locations for the detected nucleic acids were both catalogued. Each revealed variant within the sequences of fungi was marked via SnapGene Viewer ver. 4.0.4 (https://www.snapg ene.com).

**Comprehensive phylogenetic tree construction**

In this investigation, we used the neighbor-joining methodology outlined in (11) to build a complete tree tailored to our needs. With the use of the NCBI-BLASTn service (12), the detected variations were compared to their nearest homologous reference sequences. The observed variation was then included in a neighbor-joining tree, and the results were shown as a cladogram in the form of a sphere in the iTOL suit (13). A different colour was assigned to each species' sequence in the complete tree.

**Results and discussion**

**Identification of *Candida* Spp.**

In this study, all 100 isolates were tested primarily for colony characterization by culturing on the selective media, CHROMagar *Candida* and the percentage of each *Candida* spp. isolate as Figure (2):

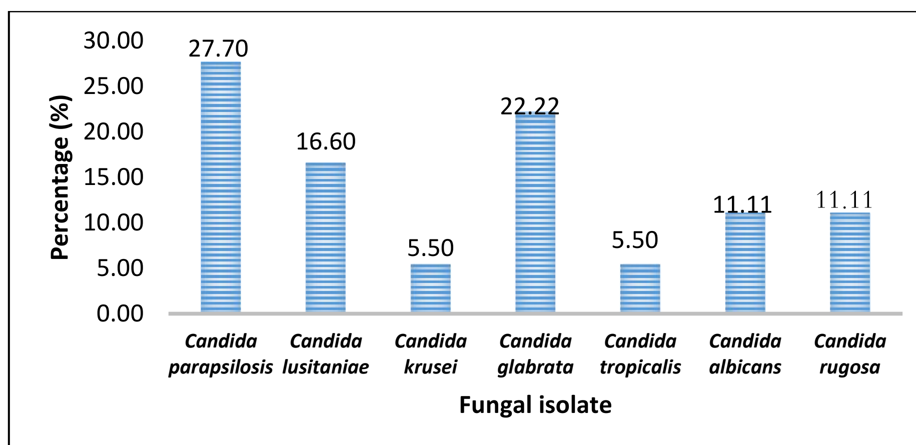



**Figure (2): Percentage of *Candida* spp. isolation in collected samples**

The isolation findings were agreed with (14) that revealed clinical samples were included five isolates from abscesses, four isolates from oral swabs, two isolates from otic secretion, two isolates from cutaneous secretion, five isolates from stool/rectal swabs, nine isolates from catheters, twelve isolates from blood, eighteen isolates from vaginal swabs, 45 isolates form sputum and 63 isolates from urine. Our data showed that *C. parapsilosis* (n = 5; 27.70%), not *C. albicans*, was the most often species obtained from blood samples. Fungemia is a growing problem, and *C. parapsilosis* complex species provide a potential explanation for this finding (15), and because it may colonise the skin, making it easier for healthcare workers to pass it to patients while handling intravenous catheters. *C. glabrata*, *C. tropicalis* and *C. albicans*, were shown to be the most common kinds of Candida in 11 clinical samples taken from hospitalised patients in Mexico City with suspected fungal infection (16). Using a sterile urinary cap, Othman *et al*. (17) collected 105 urine samples from patients with renal failure, and then they characterised the Candida spp. they found using a variety of diagnostic methods, *C. parapsilisos* and *C. albicans* at (20%) each one, *C.krusei* (27.27%), while *C. glabrata* (32.72%).

**Identification of *Candida* spp. by Vitek-2 system**

This system has been used in several previous researches due to its success in identifying and confirming biochemical tests. The Vitek-2 compact was used to identify all 100 clinical isolates of *Candida* spp. (*C. albicans* , *C. lusitaniae, C. rugose*, *C. parapsilosis* and *Cryptococcus neoformans*) and demonstrated that the results of this system matched those of traditional tests. The CHROMagar *Candida* culture medium was used to isolate the various species of Candida present, and those results may be used to infer which species are most likely to be at play. It has been employed the automated VíteK-2® and Etest® systems to screen for AFST because to their high repeatability and quick diagnostic testing with Candida spp. For antifungal susceptibility testing, the gold standard is the difficult and tedious to use CLSI standardized broth microdilution technique; nonetheless, both of these alternatives have merits.

**Molecular identification of *Candida* spp.**

The total (100) *Candida* isolates that were sequenced using ITS1 and ITS4 primers, distributed into seven species belonging to same genera ,The names of fungal genera were reported according to genus name from GenBank. The findings observed the presence of sizes ranging 550 - 625 bp depending on the difference in the lengths of the region (ITS 1 and ITS4) for yeast species, so they produce different sizes of DNA pieces that allow us to diagnose yeasts, for example, the polymerase chain reaction for a type *C. albicans* with a size of about 550 bp ,while it was for the type *C. parapsilosis* about 530 bp as in Figure (3).

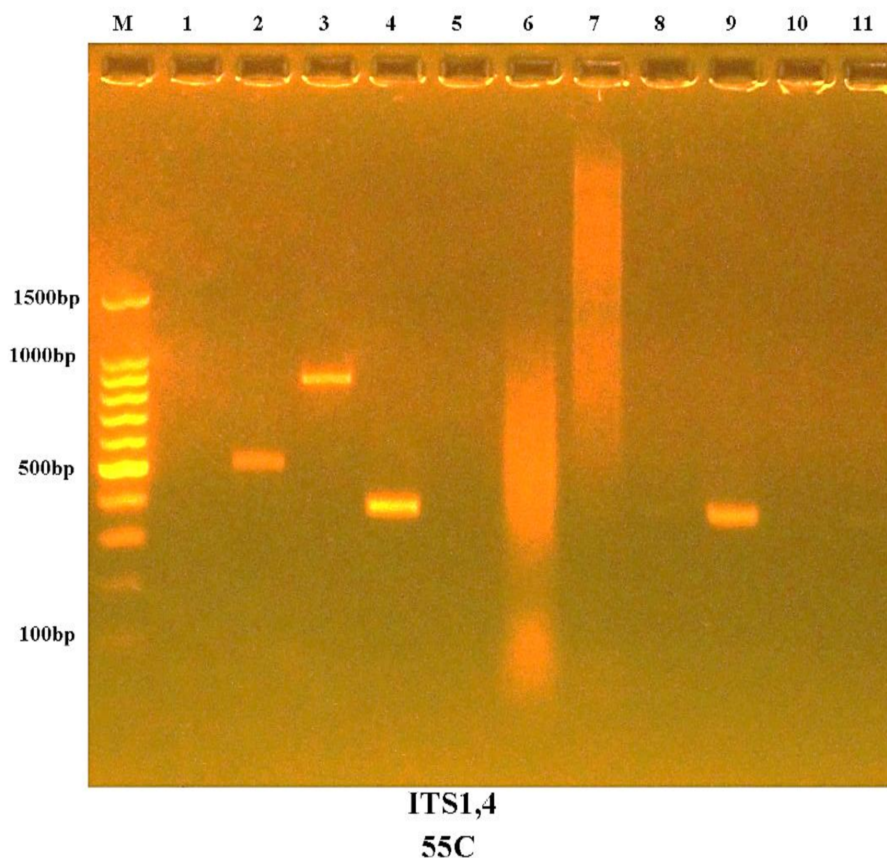



**Figure (3): Gel electrophoresis of extracted genomic DNA of (ITS) gene using (1.5%) agaros dyed with Eth.Br staining. M: 100bp marker ladder. For (1 hour) ,Lanes (1,2,3,4,5,6,7,8,9,10,11) positive result of products PCR.**

For the most prevalent *Candida* spp., the findings were consistent across all three approaches, as shown by comparison that conducted by (14). Sequences of target for ITS1 and ITS4 primers were trimmed using HardyCHROM®, MicroScan®(Omron Microscan Systems Inc, Renton, WA, USA), and PCR-RFLP (Restriction Fragment Length Polymorphisms). All clinical samples were able to successfully amplify fungal DNA; the most common fungi were members of the *C. albicans* complex, then

*C. glabrata* and *C. tropicalis*. The present results corresponded with those of a previous study (19) that used PCR to detect and identify some of the most frequent pathogenic species of *Candida* and assess the analytic value of these results in blood and bronchial lavage. According to analysis of sequence of the GenBank-deposited 18S-ITS1-5.8S-ITS2-28S section of the *Candida* spp. rDNA, two oligonucleotides, CandF and CandR, were developed. The designed oligonucleotides identified *C. dubliniensis*, *C. lusitaniae*/*Clavispora lusitaniae*, *C. guilliermondii*/ *Meyerozyma guilliermondii*, *C. krusei*/*Pichia kudriazevii*, *C. parapsilosis*, *C. tropicalis*, *C. glabrata* and *C. albicans*, using simplex PCR according to the size of amplicon, display a limit of detection of 10 pg/µL of DNA or $10^3$ yeasts/mL.

Habib *et al.* found Fifty-five clinical isolates from various sources, such as vaginal discharge, ear discharge, sputum, blood, urine, etc., were reanalyzed using the sequencing of ITS regions or 26S rRNA gene D1/D2 domains, Becton Dickinson Diagnostics (Phoenix system) as well as a Bruker Biotyper and a Vitek MS, which are two matrix-assisted laser desorption ionization-time of flight mass spectrometry analyzers. When ITS sequencing was performed on 38 isolates that had previously been classified as *C. famata* using the approach of Vitek 2, the vast majority (twenty-seven; 71.1%; isolates from total thirty-eight isolates) were found to be either *C. albicans* (7 isolates) or C. tropicalis (20 isolates), but not *C. famata*. Twenty isolates were tested, and seventeen were found to be *C. tropicalis* (or 85 percent). Patients with Vulvovaginal Candidiasis were used to analyse the phenotype and genotype of the Candida spp. responsible for the disease by Habib *et al.* (21). 100 patient samples and 30 control samples from females without the infection were identified. *Candida albicans* was the most public species, with 46.4%, then other species, include *C. krusei* (31%), *C. tropicalis* (18%), *C. parapsilosis* (7.2%) and *C. glabrata* (1.8%).

**Sequencing results**

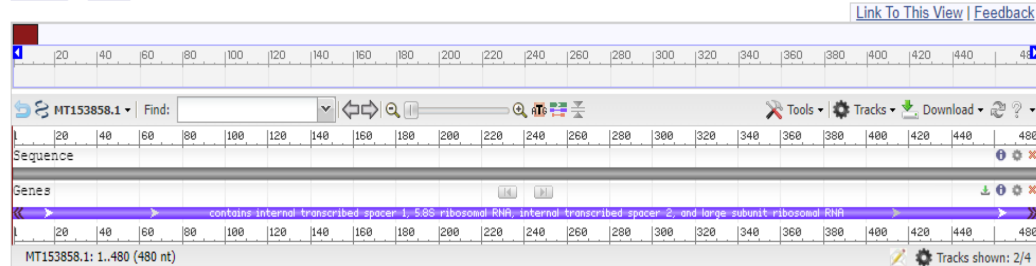
Within the targeted ribosomal loci, four fungal specimens were included in the present work. These specimens were analysed to see whether they might be used to amplify *Candida* ribosomal sequences. The variation in ribosomal sequences has the potential to be used to characterise *Candida* due to its resilience to varied genetic diversity, as demonstrated in many fungal species. After using NCBI blastn on these PCR amplicons, the sequencing reactions confirmed their true identification. About 98% to 99% sequence similarity was found between the sequenced specimens and three fungal reference target sequences (GenBank acc. MT153858.1, GenBank acc. MK886554.1, and GenBank acc. MF443854.1) when analysing the amplicons under investigation using the NCBI BLASTn engine. The locations and other features of the obtained fragments of PCR were determined by comparison of the observed and retrieved sequences of nucleic acids of the specimens under investigation (Figure 4). Sequencing reactions showed the presence of three species within *Candida* sequences, namely C. *parapsilosis* (Figure 4A), *C.lusitaniae* (Figure 4B), and *C. albicans* (Figure 4C).

*A) C. parapsilosis*

**Candida parapsilosis isolate 4Y139 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence**

GenBank: MT153858.1

GenBank FASTA

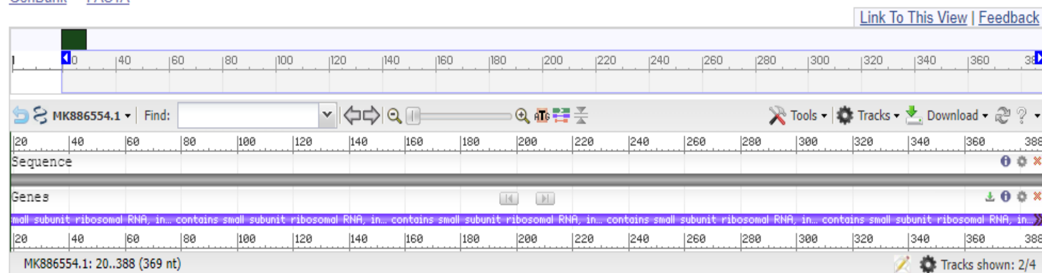



*B) Clavispora (Candida) lusitaniae*

**Clavispora lusitaniae strain RAS3 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence**

GenBank: MK886554.1

GenBank FASTA




*C) C. albicans*

**Candida albicans strain 10.120M internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence**

GenBank: MF443854.1

GenBank FASTA PopSet

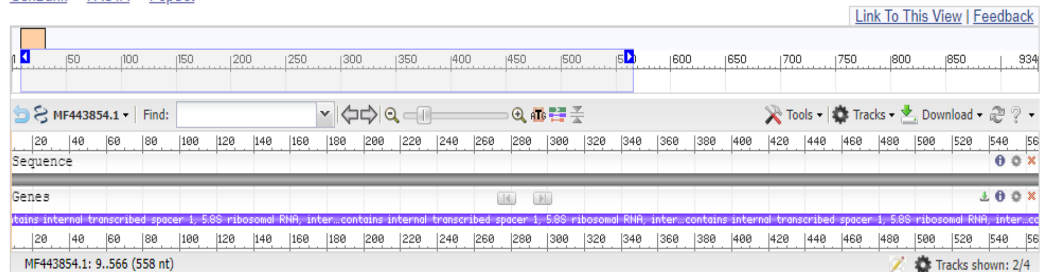



**Figure (4): The positioning of the retrieved rRNA amplicons that covered the internal transcribed spacer 1, 5.8S, internal transcribed spacer 2, and 18S rRNA within four species of *Candida* genomic sequences (GenBank acc. MT482736.1, GenBank acc. KM036428.1, GenBank acc. FR819718.1, and GenBank acc. KY792631.1). Targeted ribosomal amplicons were located within the Candida genome sequences, the sequences inside them were highlighted, and the amplicon lengths were calculated Table (4).**

**Table (4): The length and position of amplicons of PCR that are utilized for amplification of a portion of ribosomal sequences within three species of *Candida* genomic sequences (GenBank acc. MT153858.1, GenBank acc. MK886554.1, and GenBank acc. MF443854.1).**

| Amplicon | Reference locus sequences (5′ - 3′) | length |
|---|---|---|
| A) *C. parapsilosis* | GTGCTTAACTGCATTTTTTCTTACACATGTGTTTTTCTTTTTTTGA<br>AAACTTTGCTTTGGTAGGCCTTCTATATGGGGCCTGCCAGAGAT<br>TAAACTCAACCAAATTTTATTTAATGTCAACCGATTATTTAATA<br>GTCAAAACTTTCAACAACGGATCTCTTGGTTCTCGCATCGATGA<br>AGAACGCAGCGAAATGCGATAAGTAATATGAATTGCAGATATT<br>CGTGAATCATCGAATCTTTGAACGCACATTGCGCCCTTTGGTAT<br>TCCAAAGGGCATGCCTGTTTGAGCGTCATTTCTCCCTCAAACCC<br>TCGGGTTTGGTGTTGAGCGATACGCTGGGTTTGCTTGAAAGAAA<br>GGCGGAGTATAAACTAATGGATAGGTTTTTTCCACTCATTGGTA<br>CAAACTCCAAAACTTCTTCCAAATTCGACCTCAAATCAGGTAGG<br>ACTACCCGCTGAACTTAAGCATATCAAAGGCCGGAGGAA | 480 bp |
| B) *C. lusitaniae* | CGGAAGGATCATTAAAATAATACTTACACTTTGCATTTGCGAAC<br>AAAAAAAAGAACATTACACTTCTAATATATTTTTATCAAAACTT<br>TCAACAACGGATCTCTTGGTTCTCGCATCGATGAAGAACGCAGC<br>GAATTGCGATACGTAGTATGACTTGCAGACGTGAATCATCGAAT<br>CTTTGAACGCACATTGCGCCTCGAGGCATTCCTCGAGGCATGCC<br>TGTTTGAGCGTCGCATCCCCTCTAACCCCCGGTTAGGCGTTGCT<br>CCGAAATATCAACCGCGCTGTCAAACACGTTTACAGCACGACA<br>TTTCGCCCTCAAATCAGGTAGGACTACCCGCTGAACTTAAGCAT<br>ATCATAAAAGCGGAGGAA | 369 bp |
| C) *C.albicans* | GGTAGGGGGTGCCTGCGGAAGGATCATTACTGATTTGCTTAATT<br>GCACCACATGTGTTTTTCTTTGAAACAAACTTGCTTTGGCGGTG<br>GGCCCAGCCTGCCGCCAGAGGTCTAAACTTACAACCAATTTTTT<br>ATCAACTTGTCACACCAGATTATTACTTAATAGTCAAAACTTTC<br>AACAACGGATCTCTTGGTTCTCGCATCGATGAAGAACGCAGCG<br>AAATGCGATACGTAATATGAATTGCAGATATTCGTGAATCATCG<br>AATCTTTGAACGCACATTGCGCCCTCTGGTATTCCGGAGGGCAT<br>GCCTGTTTGAGCGTCGTTTCTCCCTCAAACCGCTGGGTTTGGTGT<br>TGAGCAATACGACTTGGGTTTGCTTGAAAGACGGTAGTGGTAA<br>GGCGGGATCGCTTTGACAATGGCTTAGGTCTAACCAAAAACATT<br>GCTTGCGGCGGTAACGTCCACCACGTATATCTTCAAACTTTGAC<br>CTCAAATCAGGTAGGACTACCCGCTGAACTTAAGCATATCAAA<br>GCGGGGGGGAGAAAAAAGGATCATTACTGATT | 558 bp |

When compared to the most comparable reference nucleic acid sequences of *C. parapsilosis*, *C. lusitaniae*, and *C. albicans*, the alignment findings of the amplified specimens indicated the existence of a total of seven nucleic acid alterations in various places of the analysed samples (Figure 5A - C). With regard to *C. parapsilosis* (Figure 5A), sequencing results identified one sample (S1) within this species, and detected the presence of two nucleic acid substitution mutations (g.291T>A and g.368T>A) in the S1 sample compared with the corresponding reference sequences (GenBank acc. MT153858.1). Concerning *C. lusitaniae* (Figure 5B), sequencing results identified one sample (S2) within this species, and showed the presence of only one nucleic acid substitution mutation (g.327A>G) in this sample, in comparison with the agreeing sequences of reference (GenBank acc. MK886554.1). Concerning *C. albicans* (Figure 5), sequencing results identified two samples (S3 and S4) within this species and showed the presence of one nucleic acid substitution mutation

(g.332T>C) in both S3 and S4 samples in addition three nucleic acid substitution mutations (g.341T>G, g.451A>T, and g.456T>C) in S4 sample in comparison with their corresponding reference sequences (GenBank acc. MF443854.1).

***A) C. parapsilosis***

```
         10        20        30        40        50        60        70        80        90       100
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
GTGCTTAACTGCATTTTTTCTTACACATGTGTTTTTCTTTTTTTGAAAACTTTGCTTTGGTAGGCCTTCTA
TATGGGGCCTGCCAGAGATTAAACTCAAC
S1  ....................................................................................................

        110       120       130       140       150       160       170       180       190       200
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
CAAATTTTATTTAATGTCAACCGATTATTTAATAGTCAAAACTTTCAACAACGGATCTCTTGGTTCTCGCA
TCGATGAAGAACGCAGCGAAATGCGATAA
S1  ....................................................................................................

        210       220       230       240       250       260       270       280       290       300
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
GTAATATGAATTGCAGATATTCGTGAATCATCGAATCTTTGAACGCACATTGCGCCCTTTGGTATTCCAA
AGGGCATGCCTGTTTGAGCGTCATTTCTCC
S1  ..........................................................................................A.........

        310       320       330       340       350       360       370       380       390       400
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
CTCAAACCCTCGGGTTTGGTGTTGAGCGATACGCTGGGTTTGCTTGAAAGAAAGGCGGAGTATAAACTA
ATGGATAGGTTTTTTCCACTCATTGGTACAA
S1  ..................................................................A.................................

        410       420       430       440       450       460       470       480
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
ACTCCAAAACTTCTTCCAAATTCGACCTCAAATCAGGTAGGACTACCCGCTGAACTTAAGCATATCAAAG
GCCGGAGGAA
S1  ................................................................................
```

***B) C. lusitaniae***

```
         10        20        30        40        50        60        70        80        90       100
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
CGGAAGGATCATTAAAATAATACTTACACTTTGCATTTGCGAACAAAAAAAAGAACATTACACTTCTAAT
ATATTTTTATCAAAACTTTCAACAACGGAT
S2  ....................................................................................................

        110       120       130       140       150       160       170       180       190       200
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
CTCTTGGTTCTCGCATCGATGAAGAACGCAGCGAATTGCGATACGTAGTATGACTTGCAGACGTGAATC
ATCGAATCTTTGAACGCACATTGCGCCTCGA
S2  ....................................................................................................

        210       220       230       240       250       260       270       280       290       300
    ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
ref.
GGCATTCCTCGAGGCATGCCTGTTTGAGCGTCGCATCCCCTCTAACCCCCGGTTAGGCGTTGCTCCGAA
ATATCAACCGCGCTGTCAAACACGTTTACAG
S2  ....................................................................................................

        310       320       330       340       350       360
    ....|....|....|....|....|....|....|....|....|....|....|....|....
ref.
CACGACATTTCGCCCTCAAATCAGGTAGGACTACCCGCTGAACTTAAGCATATCATAAAAGCGGAGGAA
S2  .........................G..................................
```

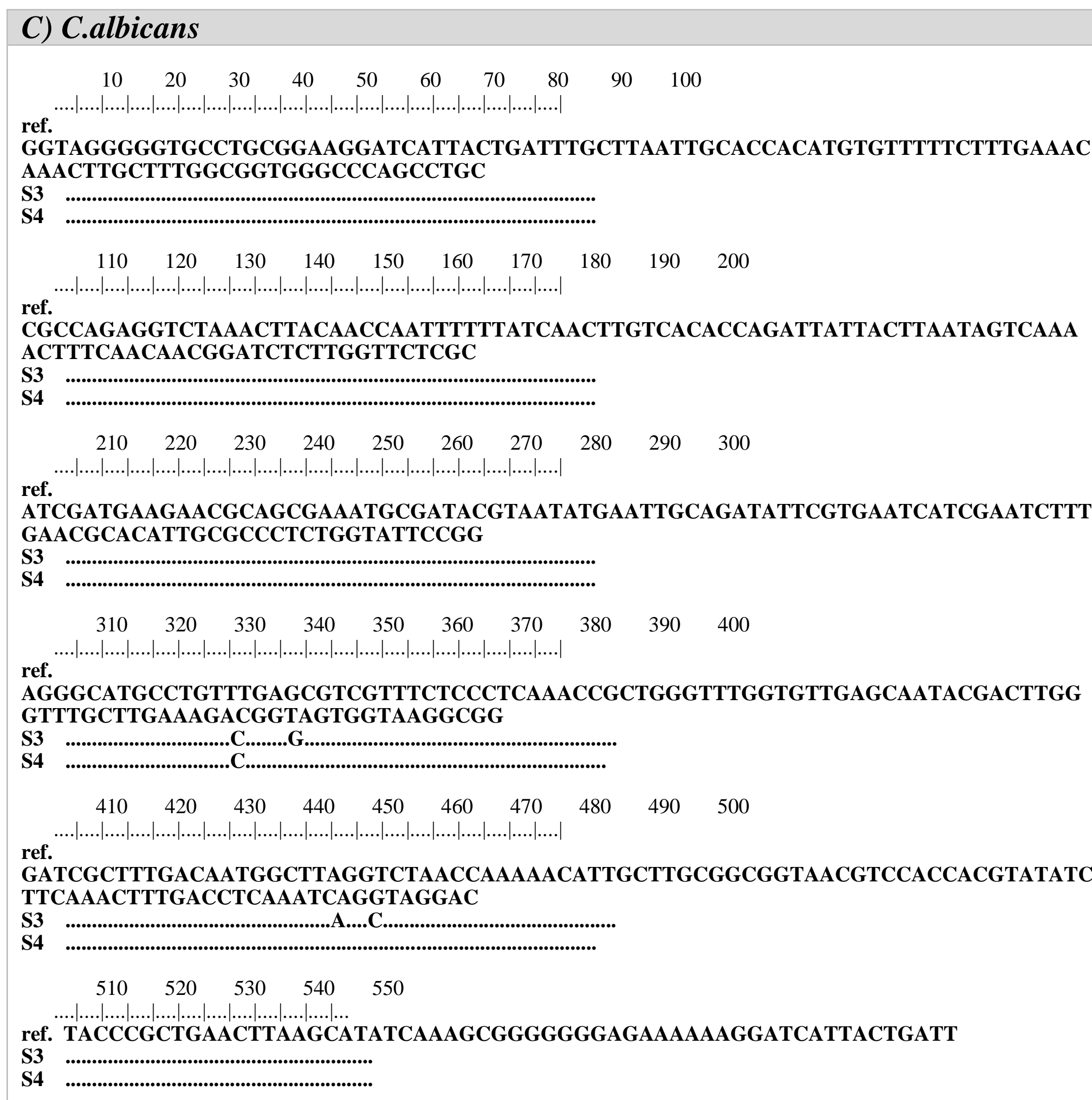



**Figure (5): Alignment of four *Candida* nucleic acid samples to reference sequences of rRNA amplicons from three different species (A): *C. parapsilosis* (B): *C. lusitaniae* (C): *C. albicans*. Sample numbers are denoted by the letter "S" followed by a number, and the sign "ref" represents the NCBI referencing sequence.**

In this work, it has been used the nucleic acid changes found in the ribosomal amplicons to construct a detailed phylogenetic tree. Samples S1–S13 were included in this phylogenetic tree, along with additional related *Candida* nucleic acid sequences. Our studied samples and their closest relatives' sequences were combined into the cladogram to form 4 distinct branches. These sequences were represented by four species of *Candida* sequences. These sequences were *C. parapsilosis, C. albicans, C. xylopsoci,* and *C. glabrata*. Each identified group was respectively incorporated within one phylogenetic clade without being interacted with the others. This finding demonstrated a strong ability of the employed rRNA to discriminate the investigated sequences into these four groups without including any noticeable

homology with other sequences from other species, and it also indicated a high degree of homology between the four indicated organisms and their corresponding species within each separate clade of the tree. This extensive tree had a total of 49 aligned nucleic acid sequences. A circular cladogram (Figure 6) were created to explain the two different depictions of the included *Candida* spp. Using the integrated *C. parapsilosis*, *C. albicans, C. xylopsoci,* and *C. glabrata* sequences, the analysed samples grouped into four phylogenetic clades in both versions of the produced cladogram. The ability of the used rRNA-based amplicons to classify the *Candida* sequences into these four phylogenetic distributions was the most fascinating finding among the fungal isolates we analysed. Within the clade of *C. parapsilosis*, thirteen sequences of the same species were incorporated alongside the investigated S1, S8, and S9 samples. All incorporated sequences exerted the same phylogenetic distributions within this clade. This observation indicated no possible phylogenetic role of the detected nucleic acid substitution of g.492-494Adel of the S1 sample in inducing any tilt within the clade of *C. parapsilosis* with regard to the wild-type (S8 and S9). The incorporated samples within this clade were suited in the immediate vicinity to various strains of *C. parapsilosis* deposited from Japan (GenBank acc. no. LC389898.1, LC389135.1, LC390091.1, LC389990.1, LC389986.1, LC390059.1,), Peru (GenBank acc. no. MZ540261.1), Austria (GenBank acc. no. AF455433.1), Kazakhstan (GenBank acc. no. MT482736.1), and China (GenBank acc. no. EF198017.1).

The S2, S6, and S13 samples were suited within the *C. albicans* major clade. Within this clade, these samples occupied close phylogenetic connections to an Egyptian strain (GenBank acc. no. KM036428.1). Despite the nucleic acid substitutions (g.168A>G and g.233C>A) observed in it, the altered S13 sample was incorporated alongside other wild-type samples evenly within the *C. albicans* major clade. This was due to the absence of any possible phylogenetic role of these mutations in causing any deviation within the currently generated tree. The S3 sample was incorporated within the clade of *C. xylopsoci* beside two strains deposited from Australia (GenBank acc. no. FR819718.1 and FM178339.1). As was observed in the other detected variations, no phylogenetic effect was identified from the nucleic acid substitution found in the S13 sample with respect to its positioning within the clade of *C. xylopsoci*. Within the last clade of *C. glabrata*, six of our investigated samples were incorporated (S4, S5, S7, S10, S11, and S12). Five of these samples (S4, S5, S7, S11, and S12) exerted variable numbers of mutations. These mutations collectively have induced a slight effect on the phylogenetic positioning of these samples with respect to the wild-type S10 sample. However, all these six samples were suited in the vicinity to variable strains isolated from Iran (GenBank acc. OM049826.1 and ON391956.1), and Kuwait (GenBank acc. HE993756.1).

Overall, it was inferred from the constructed tree that the detected nucleic acid variations showed a negligible evolutionary role in the variations observed in the fungal samples in comparison with the other investigated wild-type *Candida* sequences. The current observations of this comprehensive tree have confirmed sequencing reactions since it showed the actual neighbour-joining-based positioning in such observed variations. Interestingly, the multi-national origins of our investigated S1 and S13 samples could not be ignored.

Noteworthy, the utilization of the ITS1, 5.3 S, ITS2, and 18S rRNA sequences in this study has given further indication for the presence of accurate identification of these fungal organisms. Consequently, these observations are in line with each other to support our data on the divergence of these human-infecting pathogenic fungal sequences from the referred geographical sources. This rRNA - based comprehensive tree has provided comprehensive pieces of evidence about the high efficiency of such genetic fragments to identify this sort of phylogenetic distribution to discriminate between these four species without observing any cross-species phylogenetic similarity.

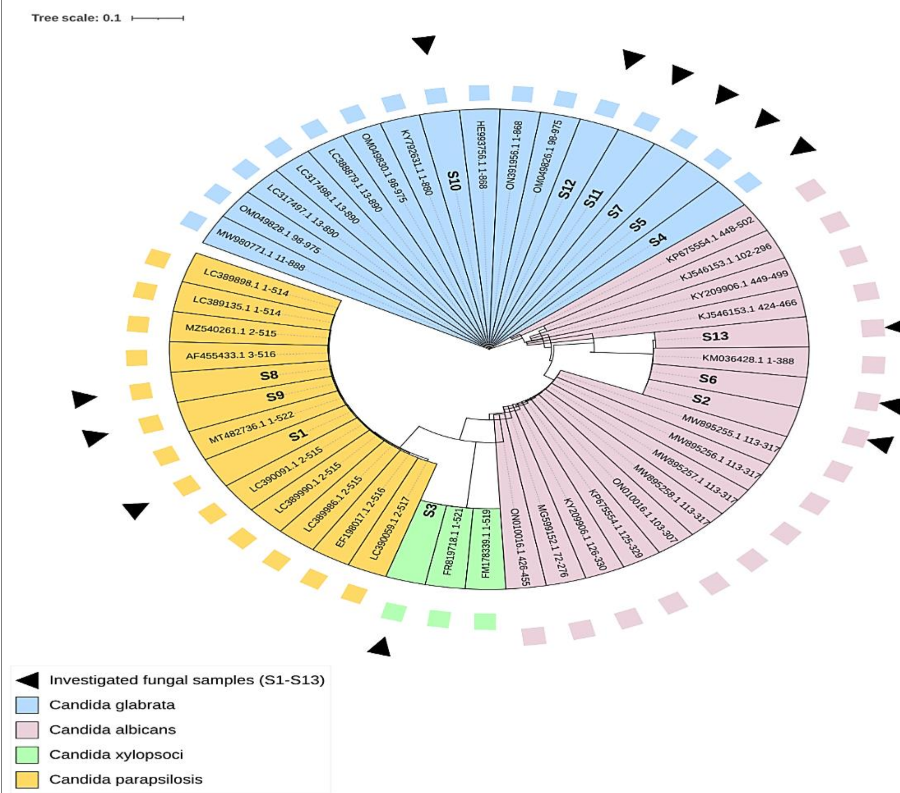



**Figure (6): A comprehensive circular cladogram phylogenetic tree of genetic variants of the rRNA fragment of four species of *Candida*. Fungi variations that were studied are shown by the black triangle. The numbers that were stated were the GenBank accession numbers for the respective species. The "0.1" at the top of the tree represents the relative degree of scale variation among the many taxonomic groups represented by the tree. The "S#" represents the sample codes that were used in the study.**

Eghtedar *et al*. (22) found that 88.9% of the isolates were part of clade I, whereas 11.10% were part of clade II. Subclade I comprised isolates from six different subclades and strains, including S7, H599, H40, H65, H60, and H40. Subclades were numbered from I through IV. H63 (*C. tropicalis*), H60 (*C. glabrata*), and H41 (*C. kefyr*) were found in HIV/AIDS patients, whereas H65 and H6 were C. albicans and H40 and H59 were isolates of *C. parapsilosis*. S7, which is acquired from HIV-negative people, was shown to be the strain of *C. guilliermondii*.

While Researchers Alrubayae *et al*.(23) identified pathogenic yeasts responsible for oral infections in children with cancer and used a phylogenetic tree to evaluate the genetic studied isolates relatedness. Analysis of sequence of regions of the internal transcribed spacer (ITS) allowed to identify *Candida dubliniensis, Candida glabrata*, and *Candida tropical*, as well as *Rhodotorula mucilaginosa* and *Kluyveromyces marxianus*. The ITS region of *R. mucilaginosa* was sequenced for the first time in Iraq, marking its official enrollment. The assembly was completed with the use of Maximum Likelihood (ML) analysis of the rDNA ITS1 and ITS2 sequences from individual isolates. Six reference strains' sequences were retrieved from the NCBI's public database: *Candida albicans*, *Candida tropicalis, Candida dubliniensis, Candida glabrata,* and *Rhodococcus mucilaginosus*. The findings indicated that the isolates might be divided into six distinct clades. The first clade consists of 12 *C. albicans* that are closely related to the reference strain KJ651874.1, while the second clade consists of *C. dubliniens* that is closely related to KY6731961.1. Similar to MT539196.1 and belonging to the third clade, *C. tropicalis*. Two *K. marxianus* strains, both belonging to the fourth clade and sharing characteristics with MN062966.1. The C. glabrata and R. mucilaginosa isolates belonged to clade 4, whereas the reference strains MK343435.1 and MT465994.1 belonged to clade 5. The 1100 replicates used to generate the bootstrap values.

## Conclusion

This study suggests possible employment for these ribosomal amplicons to discriminate between the phylogenetic diversity among the other implemented tools. These amplicons may be used effectively to identify the large range of biological variability present in *Candida* sequences. Additional information on the sequences of the discovered species of *Candida* may be gleaned by future investigation of these potential amplicons in a variety of clinical illnesses.

## Acknowledgements

We would like to thank the staff of the Gazi Al Hariri Hospital, Al-Yarmouk Teaching Hospital and Baghdad Teaching Hospital in Medical City.

## References

1. Al-Qursh, F.H. and Ali, I.A. (2023). Molecular detection of *Candida* spp. isolated from female patients infected with COVID-19 in baghdad city. Iraqi Journal of Biotechnology, 22(1): 238-244.
2. Ksiezopolska, E. and Gabaldón, T. (2018). Evolutionary emergence of drug resistance in *Candida* opportunistic pathogens. Genes, 9(9): 461-480.
3. Araújo, D.; Henriques, M. and Silva, S. (2017). Portrait of *Candida* species biofilm

regulatory network genes. Trends in Microbiology, 25(1): 62-75.

4. Aggarwal, N.; Bhateja, S.; Arora, G. and Yasmin, T. (2018). Candidiasis-The most common fungal infection of oral cavity. Biomedical Journal Science Technology Research, 8(1): 001-003.
5. Faraj, M.M.: AL-Jobori, K.M. and Risin, M.H..(2020). Evaluation of vitik2 system for clinical identification of *Candida* species and their antifungal exposure test. Iraqi Journal of Biotechnology, 19(1): 28-39.
6. Zahir, R. A. and Himratul,Aznita, W. H. (2013). Distribution of *Candida* in the oral cavity and its differentiation based on the internally transcribed spacer (ITS) regions of rDNA. Yeast, 30(1): 13-23.
7. Garcia-Cuesta, C.; Sarrion-Pérez, M. G. and Bagán, J. V. (2014). Current treatment of oral candidiasis: A literature review. Journal of Clinical and Experimental Dentistry, 6(5):576-590
8. Aslani, N.; Janbabaei, G.; Abastabar, M.; Meis, J. F.; Babaeian, M.; Khodavaisy, S.; and Badali, H. (2018). Identification of uncommon oral yeasts from cancer patients by MALDI-TOF mass spectrometry. BMC Infectious Diseases, 18(2): 1-11.
9. Khan, Z. U.; Ahmad, S.; Mokaddas, E. and Chandy, R. (2004). Tobacco agar, a new medium for differentiating *Candida dubliniensis* from *Candida albicans*. Journal of Clinical Microbiology, 42(10) :4796-4798.
10. Rennie, R. P.; Brosnikoff, C.; Turnbull, L.; Reller, L. B.; Mirrett, S.; Janda, W., *et al*. (2008). Multicenter evaluation of the Vitek 2 anaerobe and *Corynebacterium* identification card. Journal of Clinical Microbiology, 46(8): 2646-2651.
11. Hashoosh, Q.H. and AL-Araji, A. M. (2023). Molecular identification of *Tinea* spp. causing *Tinea* disease using ITS sequencing analysis. Iraqi Journal of Biotechnology, 22(1):220-228.
12. Zhang, Z.; Schwartz, S.; Wagner, L. and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. Journal Comput Biology 7(1-2):203-14.
13. Letunic, I. and Bork, P. (2019). Interactive Tree of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research, 47(1): 256-259.
14. Montes, K.; Ortiz, B.; Galindo, C.; Figueroa, I.; Braham, S. and Fontecha, G. (2019). Identification of *Candida* species from clinical samples in a Honduran Tertiary Hospital. Pathogens, 8(4): 237-250.
15. Van Asbeck, E. C.; Clemons, K. V. and Stevens, D. A. (2009). *Candida parapsilosis*: a review of its epidemiology, pathogenesis, clinical aspects, typing and antimicrobial susceptibility. Critical Reviews in Microbiology, 35(4): 283-309.
16. Camacho-Cardoso, J. L.; Martínez-Rivera, M. Á.; Manzano-Gayosso, P.; Méndez-Tovar, L. J.; López-Martínez, R. and Hernández-Hernández, F. (2017). Molecular detection of *Candida* species from hospitalized patient's specimens. Gac Med Mex, 15(3):581-589.
17. Othman, K. I.; Abdullah, S. M.; Ali, B. and Majid, M. (2018). Isolation and identification *Candida* spp from urine and antifungal susceptibility test. Iraqi Journal of Science, 7(1):1981-1988.
18. Clinical and Laboratory Standards Institute [CLSI]. (2021). M27-A3 reference method for broth dilution antifungal susceptibility testing of yeast; Approved Standard, 3rd Edn, 28(1) CLSI 6–12.
19. García-Salazar, E.; Acosta-Altamirano, G.; Betancourt-Cisneros, P.; Reyes-Montes, M. D. R.; Rosas-De-Paz, E.; Duarte-Escalante, E., *et al*. (2022). Detection and molecular identification of eight *candida* species in clinical samples by simplex PCR. Microorganisms, 10(2), 374-395.
20. Kim, T. H.; Kweon, O. J.; Kim, H. R. and Lee, M. K. (2016). Identification of uncommon *Candida* species using commercial identification systems. Journal of Microbiology and Biotechnology, 26(12): 2206-2213.
21. Habib, K.A.; Najee, E.N. and Abood, M.S. (2016). Identification of *candida* species isolated from vulvovaginal candidiasis patients by chromgen agar and PCR-RFLP Method. Baghdad Science Journal, 13(4): 291-297.
22. Nejad, E.; Almani, G.N.; Mohammadi, M. A. and Salari, S. (2020). Molecular identification of *Candida* isolates by real-

time PCR high resolution melting analysis and investigation of the genetic diversity of *Candida* species. Journal of Clinical Laboratory Analysis, 34(10): 23444-234869.
23. Alrubayae, I. M.; Al-laaeiby, A.; Minati, M. H. and ALibraheem, S. A. (2020). Determination of genetic relationships and pathogenicity of oral candidiasis etiological agents in pediatric malignant patients in Basrah province, Iraq. Systematic Reviews in Pharmacy, 11(12): 180-188.