Caracterización por secuenciación de genoma completo de aislados clínicos de *Candida tropicalis* provenientes de un hospital de tercer nivel ubicado en Guayaquil

Almeida Aguilar, Stephania Susana

Facultad Hábitat, Infraestructura y Creatividad

Maestría en Biología computacional

Trabajo de titulación, previo a la obtención del título de Magister en Biología Computacional

Ing. Francisco Javier Flores, Ph.D.

25 de abril de 2025

Dedicatoria

A mi esposo Sunito Lee, por su amor, paciencia, apoyo incondicional y acompañarme en cada paso de este camino, siendo la razón para seguir adelante en los momentos más difíciles.

A mis padres René y Rocío, mis hermanos Alejandro y Elizabeth, por nunca dejar de ser el apoyo constante y la tranquilidad que solo el calor de una familia brinda.

A mis tutores, por saber guiarme en el camino y ser fuente de inspiración para superar mis propios límites.

Y, especialmente, a Dios por ser la razón de mi existir.

Stephania Susana Almeida Aguilar

Agradecimientos

En primer lugar, agradezco a Dios por darme la salud y permitirme llegar hasta donde estoy.

A mi esposo por su trabajo, sacrificio, paciencia y apoyo incondicional y nunca dejarme sola ante cualquier adversidad.

A mis padres y hermanos, por nunca dejar que me rinda y siempre ante todo darme su amor y apoyo.

A toda mi familia por su apoyo, confianza y cariño.

A Ing. Francisco Flores PhD, por todos los conocimientos impartidos. Gracias por su apoyo y confianza.

A Derly Andrade Molina PhD, por su aporte de conocimientos y respaldo en este proyecto.

Stephania Susana Almeida Aguilar

Índice de Contenidos

Dedicatoria **2**
Agradecimientos **3**
Índice de Contenidos **4**
Índice de Figuras **6**
Índice de Tablas **6**
Índice de Anexos **6**
Resumen **7**
Abstract **8**
Capítulo I: Introducción **9**
1.1 Antecedentes 9
1.2 Justificación del Problema 10
1.3 Objetivos de la Investigación 11
1.3.1 Objetivo General 11
1.3.2 Objetivos Específicos 12
Capítulo II: Revisión de la Literatura **13**
2.1 *Candida tropicalis* 13
2.1.1 *Relevancia Clínica* 13
2.1.2 *Características Genéticas* 14
2.1.3 *Resistencia a los Antifúngicos* 14
2.2 Caracterización Genética 16
2.2.1 *Secuenciación de Genoma Completo* 16
2.2.2 *MLST* 17
2.2.3 *Genes de Resistencia* 17
Capítulo III: Metodología **19**
3.1 Fuente de las Secuencias 19
3.2 Análisis de Calidad 19

3.3 Identificación de las Secuencias.....19
3.4 Análisis de Secuenciación de Genoma Completo 20
3.5 MLST 20
3.6 Evaluación de Genes Asociados a Resistencia Antifúngica 21
Capítulo IV: Resultados 22
4.1 Identificación de las Secuencias.....22
4.2 Perfiles MLST.....23
4.3 Genes asociados a la resistencia.....24
Capítulo V: Discusión 26
Capítulo VI: Conclusiones 28
Capítulo VII: Referencias 29
Capítulo VIII: Anexos.....38

Índice de Figuras

Figura 1. Mecanismos de resistencia a antifúngicos descritos en Candida tropicalis 15

Figura 2. Arbol de filogenia inferida usando la region ITS 22

Índice de Tablas

Tabla 1. Genes involucrados con mecanismos de resistencia antifúngica en Candida tropicalis 18

Tabla 2. Perfiles alélicos y complejo clonal de muestras de C. tropicalis.....23

Tabla 3. Mutaciones en genes asociados a resistencia en Candida tropicalis. 24

Índice de Anexos

Anexo A. Metada de las muestras analizadas 38

Anexo B. Script personalizado de Python.....39

Anexo C. Mutaciones previamente descritas 41

Resumen

Candida tropicalis es una especie patógena oportunista causante de infecciones nosocomiales. En los últimos años se ha presentado un aumento en los casos de candidiasis invasiva provocada por esta levadura. El objetivo de este estudio fue caracterizar por secuenciación de genoma completo (WGS) aislados clínicos de *C. tropicalis* provenientes de un hospital de tercer nivel ubicado en Guayaquil. Se identificaron las muestras mediante análisis filogenético de la región conservada ITS, usando la base de datos PubMLST se generaron perfiles alélicos para cada una de ellas y se detectaron mutaciones en los genes asociados a resistencia antifúngica. Las 11 muestras analizadas fueron identificadas como *C. tropicalis*, cuyos perfiles alélicos mostraron una alta diversidad genética al agruparse en 7 complejos clonales. Las mutaciones detectadas en los genes ERG11, ERG3, FKS2, CDR1 y TAC1 no coincidieron con aquellas descritas en la literatura como asociadas a la resistencia antifúngica en *C. tropicalis*.

Abstract

Candida tropicalis is an opportunistic pathogen that causes nosocomial infections. In recent years, the number of invasive candidiasis caused by this yeast has been increasing. This study aims to characterize clinical *C. tropicalis* isolates from a tertiary level hospital in Guayaquil by whole genome sequencing. These samples were identified by phylogenetic analysis of the conservative region ITS, MLST profiles for each one were generated using the PubMLST database and mutations in genes associated with antifungal resistance were detected. 11 samples analyzed were identified as *C. tropicalis*, where their allelic profiles showed high genetic diversity by grouping up in 7 clonal complexes. Mutations found in the genes ERG11, ERG3, FKS2, CDR1 and TAC1 were not consistent with those previously described in the literature as associated with antifungal resistance in *C. tropicalis.*

Capítulo I: Introducción

1.1 Antecedentes

Candida tropicalis es una levadura patógena oportunista, siendo la especie más común de candidiasis en el subcontinente índico (Lima et al., 2022), y frecuentemente aislada en América Latina y el sudeste asiático (Keighley et al., 2022). Estudios recientes sugieren que existe un aumento en la resistencia a múltiples fármacos en los aislados de origen clínico de este patógeno (Favarello et al., 2021). En el Ecuador, esta es la especie non-*Candida albicans* (NCA) más identificada en muestras clínicas (Acosta-Mosquera et al., 2024; Yánez, 2020).

La caracterización genética de levaduras oportunistas comprende el estudio de diversidad genética, estructura poblacional, epidemiología y resistencia a los antifúngicos (Q. Wang et al., 2020). En este sentido, el método mejor descrito para el estudio de diversidad alélica en *C. tropicalis* es la tipificación de secuencias multi locus (MLST, por sus siglas en inglés), permitiendo comparar poblaciones de diferentes orígenes geográficos utilizando bases de datos (Wu et al., 2019). Por otro lado, en el ámbito de la farmacorresistencia se analizan mutaciones específicas en el genoma asociadas al desarrollo de resistencia (Lima et al., 2022).

A nivel mundial, la mayoría de los estudios realizados en *C. tropicalis* se han enfocado por separado en la diversidad alélica por MLST (Al-Obaid et al., 2017; Dougue et al., 2022; Tavanti et al., 2005; Q. Wang et al., 2020) y en el análisis de mutaciones en genes relacionados con resistencia (Arastehfar et al., 2020; Megri et al., 2020; Spruijtenburg et al., 2023), siendo escasos los estudios que combinen ambos enfoques

(Keighley et al., 2022). Sin embargo, la información para América Latina es limitada (Magri et al., 2013), y particularmente en Ecuador no existen investigaciones previas.

Con el presente estudio se pretende aportar nuevos conocimientos sobre la diversidad alélica y la existencia de mutaciones en los genes relacionados con resistencia en cepas locales de *C. tropicalis* de origen clínico.

1.2 Justificación del Problema

C. tropicalis es una especie patógena oportunista causante de infecciones nosocomiales, con potencial de invadir el pulmón, el tracto urogenital e intraabdominal, entre otros, en pacientes enfermos (D. Wang et al., 2021; Marcos-Zambrano et al., 2014a). En los últimos años se ha presentado un aumento en los casos de candidiasis invasiva provocada por este patógeno frente a una disminución de *C. albicans* (O'Brien et al., 2021; Zuza-Alves et al., 2017); y en comparación a esta última, conlleva a un mayor tiempo de hospitalización y una tasa de mortalidad más elevada (Arastehfar et al., 2020).

El informe sobre patógenos prioritarios de origen fúngico emitido por la Organización Mundial de la Salud (2022) señala que la candidiasis invasiva asociada a *C. tropicalis* presenta una mortalidad del 55% al 60% en adultos y del 26% al 40% en pacientes pediátricos. De ahí la necesidad de realizar más estudios a nivel global permitiendo contrarrestar la falta de información de este patógeno (Keighley et al., 2024).

La mayoría de brotes de candidiasis provocados por *C. tropicalis* ocurren en UCI, atribuyendo este hecho al riesgo inherente de contaminación cruzada en el uso de catéteres, y al estado inmunodeprimido de los pacientes (Barac et al., 2020). De manera similar, en Guayaquil, la especie predominante en pacientes hospitalizados es *C. tropicalis*, incluso por encima de *C. albicans* (Acosta-Mosquera et al., 2024).

Evidencias recientes apuntan a un aumento en la resistencia a azoles en *C. tropicalis*, llevando al tratamiento con equinocandinas, que puede desencadenar en la generación de cepas resistentes a múltiples fármacos (Lima et al., 2022). Como estrategia de prevención se plantea el estudio de los marcadores asociados a la resistencia antifúngica mediante la secuenciación de genoma completo (Keighley et al., 2022), debido a que se desconoce el efecto de las mutaciones presentes en los diferentes genes de esta especie (Xu, 2021).

Por todo lo expuesto anteriormente, resulta crucial estudiar a profundidad la diversidad alélica y las mutaciones asociadas a resistencia antifúngica a través del análisis de secuenciación de genoma completo de cepas nativas de *C. tropicalis*, con el objetivo de mejorar la comprensión de este patógeno en el ámbito clínico.

1.3 Objetivos de la Investigación

1.3.1 Objetivo General

Caracterizar por secuenciación de genoma completo (WGS) aislados clínicos de *Candida tropicalis* provenientes de un hospital de tercer nivel ubicado en Guayaquil.

1.3.2 Objetivos Específicos

- Identificar los aislados pertenecientes a la especie *C. tropicalis* utilizando marcadores moleculares.
- Definir las características genéticas de *C. tropicalis* mediante herramientas bioinformáticas.
- Detectar genes o patrones moleculares asociados a resistencia antifúngica y virulencia en las muestras de *C. tropicalis*.

Capítulo II: Revisión de la Literatura

2.1 *Candida tropicalis*

Candida tropicalis es una levadura diploide miembro del reino *Fungi*, división *Ascomycota*, orden *Saccharomycetales* (Ann-Chai et al., 2010), comúnmente hallado en la naturaleza, capaz de colonizar con frecuencia la piel humana, la cavidad oral y otros tejidos (Singh et al., 2020).

Esta especie fue descrita por primera vez en 1912 por Catellani como Oidi*um tropicale* o *Endomyces tropicalis*, aislada de pacientes con bronquitis fúngica. Es responsable de provocar infecciones nosocomiales, productora de biopelículas y altamente adherente a células epiteliales y endoteliales (Marcos-Zambrano et al., 2014b)

2.1.1 *Relevancia Clínica*

Junto a otras especies de *Candida*, *C. tropicalis* es el agente causal de la candidiasis (Zuza-Alves et al., 2017), una infección oportunista, al invadir sistemáticamente presenta una mayor tasa de mortalidad independientemente de la edad del paciente, pabellón de tratamiento o cepa del patógeno (Fan et al., 2017).

Entre los factores de virulencia de esta especie se destacan la secreción de enzimas extracelulares, capaces de alterar la composición del tejido hospedero; la formación de biopelículas, que afectan la acción de la respuesta inmune; y la resistencia a antifúngicos (Sasani et al., 2021).

En Ecuador, se ha reportado infecciones en el torrente sanguíneo provocadas por este patógeno, denominadas candidemias; en las provincias de Pichincha y Manabí

ha sido asilado en muestras de mujeres en estado de gestación (Betancourt y Carrera, 2012; Chila Santana et al., 2022).

2.1.2 *Características Genéticas*

El tamaño total del genoma completo es de aproximadamente 14,6 Mb, con un contenido de guanina citosina del 33%, se han descrito 6258 genes codificantes (Butler et al., 2009), una gran parte de estos están asociados a factores de virulencia como la adhesión al tejido del hospedero, formación de biopelículas y producción de enzimas proteasas, fosfolipasas y hemolíticas (Xu, 2021).

C tropicalis pertenece al clado CUG-Ser1, un grupo de especies que codifica serina con el codón CUG en lugar de leucina (Krassowski et al., 2018). Registra reordenamientos cromósomicos, tales como inversión y traslocación, que provocan una acumulación de mutaciones y divergencia genética entre los cromosomas homólogos (Xu, 2021).

Las características genéticas de *C. tropicalis* son muy similares a las de *C. albicans*, aparentemente debido a la reproducción clonal predominante en estas especies. Sin embargo, los eventos de recombinación ocurren con frecuencia suficiente para generar poblaciones con características similares (Zuza-Alves et al., 2017)

2.1.3 *Resistencia a los Antifúngicos*

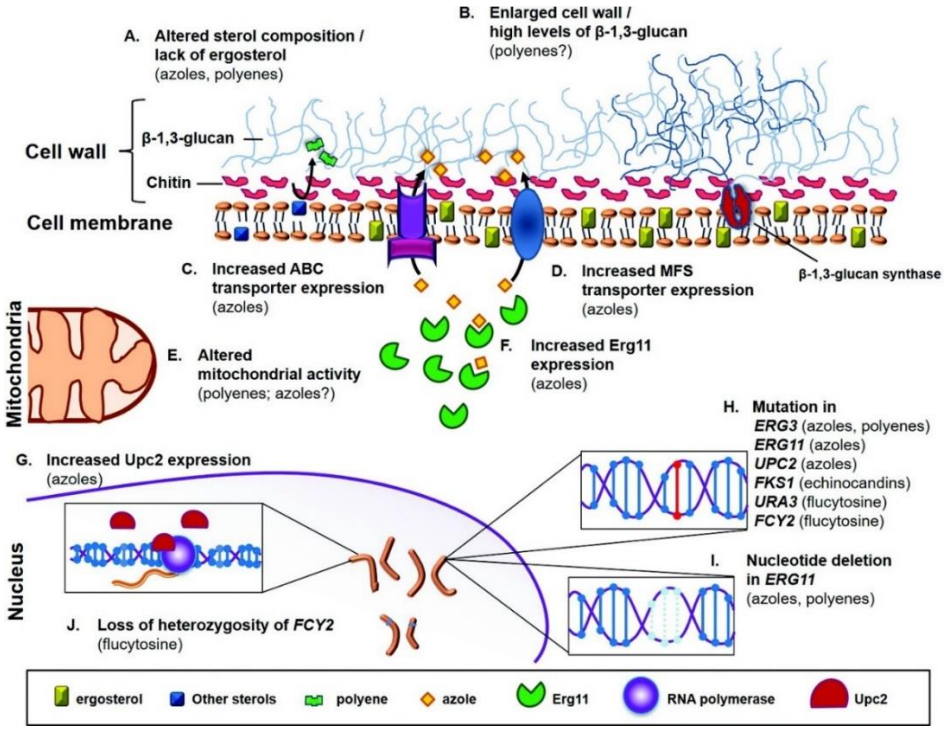
La resistencia a los antifúngicos se define como la no susceptibilidad de un hongo frente a un determinado agente antimicrobiano, es evaluada mediante ensayos in vitro de susceptibilidad clasificándose en resistencia intrínseca o resistencia adquirida (Kanafani y Perfect, 2008). Estudios recientes reportan el aumento de casos de

resistencia a fármacos en *C. tropicalis*, lo que supone un grave problema terapéutico (Zuza-Alves et al., 2017).

Existen diversos mecanismos de resistencia en *C. tropicalis*, como la alteración en la interacción entre antifúngico y molécula objetivo, el aumento o disminución de la cantidad de molécula objetivo, la expulsión selectiva de fármacos de la célula, entre otros. Generalmente, estos procesos están mediados por mutaciones en genes específicos o por alteraciones en la expresión génica (Oliveira et al., 2020). Un resumen de los mecanismos de resistencia que presenta este patógeno se puede observar en la Figura 1.

Figura 1

Mecanismos de resistencia a antifúngicos descritos en Candida tropicalis



Nota. (A) Producción de esteroles alternativos debido a la inactivación de Erg3. (B) Los β-glucano de la pared celular inhiben la penetración del antifúngico. (C) Expresión

aumentada de bombas de eflujo tipo casete de unión a ATP (C) y de la superfamilia mayor de facilitadores (MFS) reduce el nivel intracelular de azoles. (E) La actividad mitocondrial alterada reduce la acumulación intracelular de especies reactivas de oxígeno. (F) La actividad de los azoles es sobrepasada por la expresión incrementada de Erg11. (G) Una alta expresión del factor de transcripción Upc2 induce el aumento de la expresión de ERG11. (H) Mutaciones y (I) deleciones relacionadas con la reducción de la afinidad al fármaco. (J) Mutación y pérdida de heterocigocidad en FCY2 disminuye el influjo de flucitosina a la célula. Reproducido de “The yeast, the antifungal, and the wardrobe: a journey into antifungal resistance mechanisms of *Candida tropicalis*” (p. 380), por Oliveira et al., 2020, *Canadian journal of microbiology*, 66(6).

2.2 Caracterización Genética

Es el proceso de identificación y diferenciación de recursos genéticos, especialmente la detección de variantes producto de las diferencias en la secuencia genética, en genes específicos o en factores modificantes (Vicente et al. 2006).

2.2.1 *Secuenciación de Genoma Completo*

La secuenciación de genoma completo (WGS, por sus siglas en inglés) es el proceso de obtención de la totalidad de la secuencia del genoma de un organismo a partir de millones de fragmentos de ADN, abarcando secuencias codificantes y no codificantes, variaciones en el número de copias, ADN circulante y reordenamientos estructurales genómicos (deleciones, traslocaciones, etc.) (Beedanagari y John, 2014).

Los beneficios de WGS incluyen la identificación de especies, serotipos, genotipos, genes de resistencia y virulencia, anotación del genoma y la filogenia. A nivel de investigación, está consolidándose como el método de referencia para la

caracterización de patógenos, ofreciendo ventajas considerables para comprender la evolución y la diseminación de la resistencia a los antimicrobianos (Calero-Cáceres et al., 2023).

2.2.2 *MLST*

MLST es una técnica en biología molecular que permite la tipificación de cepas basada en las secuencias de ADN de genes conservados o "*house-keeping*". Pequeñas variaciones en las secuencias de estos genes permiten diferenciar las distintas cepas de una misma especie comparándolas con bases de datos (Marcos-Zambrano et al., 2014a).

La base de datos más utilizada es PubMLST dado que cuenta con información de más de 100 especies y géneros microbianos diferentes, su sistema analiza gen por gen los genomas microbianos, donde cada secuencia depositada es anotada y curada para identificar y catalogar sistemáticamente su variación (Jolley et al., 2018).

En *C. tropicalis* el esquema de MLST fue desarrollado y establecido por Tavanti et al. (2005), basado en los genes ICL1, SAP2, SAP4, MDR1, XYR1 y ZWF1a, capaz de diferenciar entre cepas de este hongo con un poder discriminatorio mayor al 99%; convirtiéndolo en el método más utilizado debido a su alta reproducibilidad y mayor discriminación (Q. Wang et al., 2020).

2.2.3 *Genes de Resistencia*

Los genes de resistencia son uno de los mecanismos que utilizan los hongos para contrarrestar el efecto de los antifúngicos (Navarro, 2019). En *C. tropicalis* se han descrito varios genes asociados a la resistencia a antifúngicos de uso común, junto con la función biológica que cumplen (Tabla 1).

Tabla 1

Genes involucrados con mecanismos de resistencia antifúngica en Candida tropicalis

Fármaco	Gen	Producto génico	Función biológica
Azoles	ERG11	Erg11p	Vía de biosíntesis del ergosterol
	ERG3	Erg3p	
	MDR1	Mdr1p	Transporte dependiente de la energía
	CDR1	Cdr1p	
Anfotericina B	ERG3	Erg3p	Vía de biosíntesis del ergosterol
	ERG6	Erg6p	
	ERG11	Erg11p	
Equinocandinas	FKS1	Fks1p	Acción catalítica
	FKS2		
	FKS3		
	RHO1	Rho1p	Regulación de la biosíntesis de -1,3-D-glucano y otros procesos celulares

Nota. Adaptado de “An Update on *Candida tropicalis* Based on Basic and Clinical Approaches” (p.15), por Zuza-Alves et al., 2017, *Frontiers in Microbiology*, 8.

El aumento en los casos de resistencia antifúngica a azoles en *C. tropicalis* se asocia con la sobreexpresión en las bombas de eflujo y mutaciones en los genes ERG11 y/o ERG3. Por ejemplo, la mutación A395T/W en ERG11 produce el cambio proteico Y132F provocando una menor afinidad entre los azoles y su molécula diana. Además, mutaciones sin sentido pueden reducir el efecto antifúngico de los azoles, equinocandinas y polienos (Oliveira, et al., 2020).

Capítulo III: **Metodología**

3.1 Fuente de las Secuencias

Los archivos fastq fueron proporcionados por la Dra. Derly Andrade-Molina de la Universidad Espíritu Santo (UEES), provenientes de muestras clínicas de un hospital terciario en la ciudad de Guayaquil (Acosta-Mosquera et al, 2024), secuenciados mediante tecnología Illumina® de extremos emparejados (*paired-end*). La metadata de las muestras se detalla en el Anexo A.

3.2 Análisis de Calidad

Las secuencias crudas fueron evaluadas en el programa FastQC v0.11.9 para comprobar contenido de GC, presencia de adaptadores, secuencias duplicadas o sobre representadas y calidad de secuenciación. Los adaptadores se removieron de los archivos fastq pareados utilizando Trimmomatic v0.39 con los parámetros predeterminados (Bolger et al., 2014).

3.3 Identificación de las Secuencias

Las secuencias libres de adaptadores se ensamblaron de *novo* usando el programa MEGAHIT v1.2.9 (Li et al., 2015; Li et al., 2016). La calidad de los ensamblajes se evaluó utilizando QUAST v5.3.0 (Mikheenko et al., 2023).

La identificación de la región del espaciador transcrito interno (ITS, por sus siglas en inglés) se realizó mediante BLAST+ (Camacho et al., 2009), comparando las muestras y otras especies de *Candida* obtenidas de Candida genome database (http://www.candidagenome.org/download/sequence/) con una base de datos local generada a partir de secuencias del género *Candida* disponibles en Nucleotide de NCBI

(Sayers et al., 2024). Las regiones obtenidas fueron alineadas para generar un árbol de filogenia basado en el método de máxima verosimilitud en Mega v11.0.13 (Tamura et al., 2021).

3.4 Análisis de Secuenciación de Genoma Completo

Los archivos fastq libres de adaptadores fueron mapeados al genoma de referencia *C. tropicalis* cepa MYA-3404 ensamblaje ASM633v3 (número de acceso GCF_000006335.3) utilizando el programa BWA-MEM 2 v2.2.1 con sus parámetros por defecto (Vasimuddin et al., 2019). Los archivos SAM fueron convertidos a BAM, los cuales se ordenaron e indexaron usando SAMtools v1.21 (Danecek et al., 2021), las lecturas duplicadas se marcaron con Picard MarkDuplicates v3.0.

Los polimorfismos de un solo nucleótido (SNPs, por sus siglas en inglés) se infirieron en FreeBayes con los parámetros min alt 0.3 y min depth 10 (Garrison y Marth, 2012). Los archivos VCF fueron normalizados utilizando el programa vt (Tan et al., 2015) y anotados en SnpEff v5.2 (Cingolani et al., 2012).

3.5 MLST

El perfil MLST se determinó utilizando secuencias consenso que fueron generadas en el programa BCFtools (Danecek et al., 2021) a partir de los archivos VCF normalizados. Cada muestra fue subida a la base de datos de tipificación de *Candida tropicalis* en PubMLST (https://pubmlst.org/bigsdb?db=pubmlst_ctropicalis_seqdef) de donde se obtuvo los perfiles alélicos y se analizó si estos correspondían a una secuencia tipo (ST) conocida. Además, se verificó si se podían agrupar las muestras en base a sus perfiles.

3.6 Evaluación de Genes Asociados a Resistencia Antifúngica

Basado en estudios previos de *C. tropicalis* se seleccionaron los genes ERG11, ERG3, FKS2, CDR1 y TAC1 asociados a resistencia antifúngica. Estos fueron extraídos del genoma de referencia mediante bedtools (Quinlan y Hall, 2010), a partir de las coordenadas descritas en el estudio realizado por Keighley et al. (2022), generando una base de datos local que permitió determinar mediante BLAST+ (Camacho et al., 2009) las coordenadas de cada gen en las muestras.

Los genes asociados a la resistencia presentes en el genoma de referencia fueron comparados, mediante un script personalizado de Python (Anexo B), con aquellos obtenidos de las muestras para investigar sustituciones en los aminoácidos previamente asociados a resistencia antifúngica en este patógeno.

Capítulo IV: Resultados

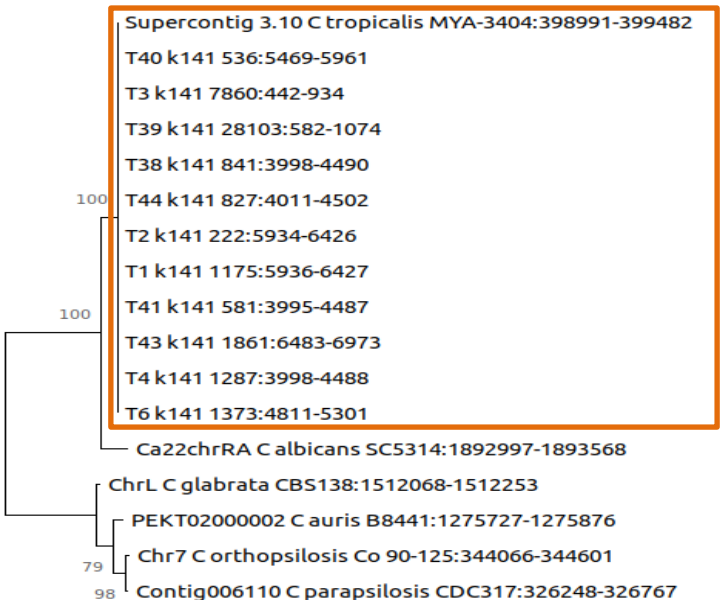
4.1 Identificación de las Secuencias

Se analizaron un total de 11 archivos fastq emparejados, los cuales pasaron los controles de calidad, por tanto fueron ensamblados y evaluados en QUAST, presentando una longitud total aproximada de 14,6 Mb, N50 superior a 60000 pb, y porcentaje de cobertura respecto a la referencia mayor al 96%, indicadores de un buen ensamblaje.

A partir de las secuencias ensambladas, se utilizó la región ITS para confirmar la identidad de las muestras mediante el alineamiento con varias especies del género Candida usando el algoritmo MUSCLE. Las relaciones filogenéticas mostraron que las 11 muestras formaron un grupo monofilético con la secuencia de referencia de *C. tropicalis*, confirmando así la identidad de las mismas (Figura 2).

Figura 2

Arbol de filogenia inferida usando la region ITS en especies del genero Candida, basado en el metodo de maxima verosimilitud con el modelo Tamura-Nei. Los valores en las ramas fueron evaluados mediante bootstrap con 1000 replicas.



4.2 Perfiles MLST

El perfil MLST de las muestras generado a partir de la base de datos PubMLST se detalla en la Tabla 2. Ninguna de estas coincidió con un ST descrito anteriormente, razón por la cual, se agruparon en complejos clonales aquellas que presentaron el mismo alelo por lo menos en 5 loci distintos al compararse entre ellas. De las 11 muestras, se identificó dos (T2 y T4) con el mismo perfil alélico en todos los loci estudiados.

Tabla 2

Perfiles alélicos y complejo clonal de muestras de C. tropicalis

Muestra	ICL1	MDR1	SAPT2	SAPT4	ZWF1	XYR1	Complejo clonal
T1	3	90	3	93	6	52	1
T2	1	175	3	17	3	54	2
T38	1	180	9	10	9	200	3
T39	1	209	18	1	1	76	4
T3	1	117	3	7	1	2	5
T40	1	209	18	1	1	5	4
T41	1	209	18	1	1	107	4
T43	1	9	12	23	9	92	6
T44	1	180	3	7	3	58	7
T4	1	175	3	17	3	54	2
T6	1	180	3	7	3	92	7

El locus XYR1 presentó el mayor grado de polimorfismo (10 alelos únicos), seguido de los loci MDR1 y SAPT4 (6 alelos únicos), con ICL1 mostrando la menor cantidad de polimorfismos (2 alelos únicos).

4.3 Genes asociados a la resistencia

Se identificaron un total de 28 mutaciones no sinónimas en los cinco genes analizados, descritas en la Tabla 3.

Tabla 3.

Mutaciones detectadas en las muestras en genes asociados a resistencia en Candida tropicalis.

Gen	Sustitución aminoácidos	Muestras portadoras
ERG11	S137H	T3
ERG3	E123K	T2, T38, T43, T44, T4, T6
FKS2	T72I	T2
	P132S	T43
	H1291L	T44
	F1438L	T2, T38, T44, T6
CDR1	M15V	T38
	P60R	T2, T44, T4, T6
	K332N	T1
	D355N	T1, T39, T40, T41, T4
	F367Y	T40, T41
	A377T	T2, T44, T4, T6
	F623L	T2, T4, T6
	T644I	T43
	V660D	T39, T40, T41
	A758S	T39, T40, T41
	D966G	T1, T43
	R1064Q	T3
	A1066V	T3
	V1412M	T39, T40, T41, T44, T6

TAC1	R201G	T1, T2, T38, T44, T4, T6
	M336R	T6
	T530S	T1, T2, T43, T44, T4, T6
	Q544*	T1, T2, T43, T44, T4, T6
	I549F	T1, T2, T43, T44, T4, T6
	E631G	T2
	D790N	T1, T2
	V917E	T43

Las sustituciones R201G, T530S, I549F, y Q544* del gen TAC1; y la sustitución E123K del gen ERG3, fueron halladas en 6 de las 11 muestras analizadas. Además, las muestras T2, T6 y T44 presentaron un mayor número de mutaciones.

Capítulo V: **Discusión**

Este es el primer estudio en Ecuador que caracteriza muestras clínicas de *Candida tropicalis* mediante secuenciación de genoma completo. Según Hilt y Ferrieri (2022), esta técnica permite la identificación y tipificación de microorganismos y posibles mecanismos de resistencia a antimicrobianos.

La identificación de las muestras se realizó a partir de la extracción de la región ITS del genoma completo comparada frente a un panel de especies del género *Candida*, encontrando que todas las secuencias analizadas pertenecen a la especie *C. tropicalis*; Ramírez y De la Torre (2008) mencionan que la región ITS permite identificar organismos muy cercanos a nivel filogenético.

Los perfiles de MLST obtenidos en las muestras de esta investigación indican una alta diversidad genética en *C. tropicalis*, donde ninguna coincidió con secuencias tipo al compararla con la base de datos PubMLST, similar a lo descrito por Keighley et al. (2022) y Al-Obaid et al. (2017) quienes encontraron 53 ST nuevos en 61 aislados, y 59 ST nuevos en 63 aislados, respectivamente. Esto pone de manifiesto que *C. tropicalis* es una especie aún no descrita en su totalidad y por tanto es necesario continuar con su estudio para alimentar la base de datos con nueva información.

Se encontró 10 ST diferentes en las 11 muestras estudiadas con un 90,1% de diversidad alélica, a diferencia de lo publicado por Dougue et al. (2022), quienes reportaron apenas 53 ST en 103 muestras de 3 países diferentes, con un 51,46% de diversidad alélica. Esta proporción podría estar afectada por el tamaño de la muestra, pero resalta la variabilidad intraespecífica de este patógeno, incluso en un mismo hospital.

Al relacionar entre el complejo clonal asignado, el tipo de muestra y la hospitalización en unidad de cuidados intensivos (UCI), no se evidenció una correlación notable entre estas variables en las muestras de este estudio. Keighley et al. (2022) y Dougue et al. (2022) resaltan la importancia de disponer y analizar la información relacionada con el paciente (metadata) de los aislados para evaluar el impacto clínico de los marcadores moleculares encontrados mediante secuenciación de genoma completo.

Las mutaciones detectadas en este estudio no se han reportado previamente en la literatura como factores directos asociados a la resistencia en *C. tropicalis*. Sin embargo, algunas de estas variantes (Anexo C) coinciden con lo reportado por Keighley et al. (2022), donde se encontraron tanto en cepas resistentes como susceptibles; sugiriendo que la resistencia podría tener origen multifactorial (Oliveira et al., 2020).

Las sustituciones Y132F y Y257H/N en el gen ERG11, asociadas a la resistencia en cepas de Latinoamérica (Spruijtenburg et al., 2023), no fueron detectadas en ninguna de las muestras analizadas en este estudio; así como la mutación K143R, reportada por Xisto et al. (2017) en muestras clínicas de Brasil resistentes a azoles, que tampoco fue hallada.

El presente trabajo proporciona, por primera vez en cepas de *C. tropicalis* en Ecuador, la diversidad alélica mediante perfiles de MLST y la identificación de mutaciones en genes asociados a resistencia, sentando las bases para futuras investigaciones que amplíen estos análisis al incluir un mayor número de muestras de diversas ubicaciones del país.

Capítulo VI: Conclusiones

- Se identificaron 11 muestras como *C. tropicalis* a partir del análisis de filogenia para la región conservada ITS

- La caracterización genética mediante perfiles MLST evidenció una alta diversidad alélica agrupando a las muestras en siete complejos clonales

- Las mutaciones detectadas en los genes ERG11, ERG3, FKS2, CDR1 y TAC1 no coincidieron con aquellas descritas en la literatura como asociadas a la resistencia antifúngica en *C. tropicalis*.

Capítulo VII: Referencias

Acosta-Mosquera, Y., Tapia, J. C., Armas-González, R., Cáceres-Valdiviezo, M. J., Fernández-Cadena, J. C. y Andrade-Molina, D. (2024). Prevalence and Species Distribution of Candida Clinical Isolates in a Tertiary Care Hospital in Ecuador Tested from January 2019 to February 2020. Journal of fungi (Basel, Switzerland), 304.

Al-Obaid, K., Asadzadeh, M., Ahmad, S. y Khan, Z. (2017). Population structure and molecular genetic characterization of clinical Candida tropicalis isolates from a tertiary-care hospital in Kuwait reveal infections with unique strains. PloS one, e0182292.

Ann-Chai, L. Y., Denning, D. W. y Warn, P. (2010). Candida tropicalis in human disease. Critical reviews in microbiology, 282-298.

Arastehfar, A., Daneshnia, F., Hafez, A., Khodavaisy, S., Najafzadeh, M. J., Charsizadeh, A., Zarrinfar, H., Salehi, M., Shahrabadi, Z. Z., Sasani, E., Zomorodian, K., Pan, W., Hagen, F., Ilkit, M., Kostrzewa, M. y Boekhout, T. (2020). Antifungal susceptibility, genotyping, resistance mechanism, and clinical profile of Candida tropicalis blood isolates. Medical Mycology, 766-773.

Barac, A., Cevik, M., Colovic, N., Lekovic, D., Stevanovic, G., Micic, J. y Rubino, S. (2020). Investigation of a healthcare-associated Candida tropicalis candidiasis cluster in a haematology unit and a systematic review of nosocomial outbreaks. Mycoses, 326-333.

Beedanagari, S. y John, K. (2014). Next Generation Sequencing. En P. Wexler, Encyclopedia of toxicology (págs. 501-503). Elsevier.

Betancourt, E. y Carrera, V. (2012). Prevalencia de candidiasis vaginal en mujeres embarazadas de Quito-Ecuador: identificación de especies utilizando dos medios de cultivo. *Revista médica Vozandes*, 113-118.

Bolger, A. M., Lohse, M. y Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics (Oxford, England), 2114-2120.

Butler, G., Rasmussen, M. D., Lin, M. F., Santos, M. A., Sakthikumar, S., Munro, C. A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J. L., Agrafioti, I., Arnaud, M. B., Bates, S., Brown, A. J., Brunke, S., Costanzo, M. C., Fitzpatrick, D. A., de Groot, P. W., Harris, … Cuomo, C. A. (2009). Evolution of pathogenicity and sexual reproduction in eight Candida genomes. Nature, 657-662.

Calero-Cáceres, W. O.-G., Gomes-Dias, C. A., Bastidas-Caldes, C., Ramírez, M. S. y Harries, A. D. (2023). Whole-genome sequencing for surveillance of antimicrobial resistance in Ecuador: present and future implications. Revista panamericana de salud publica = Pan American journal of public health, e8.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. y Madden, T. L. (2009). BLAST+: architecture and applications. BMC bioinformatics, 421.

Chila Santana, L. A., Anzules Guerra, J. B., Milian Hernández, E. J. y Izaguirre Bordelois, M. (2022). Perfil clínico-microbiológico de la Candidiasis Vulvovaginal en mujeres embarazadas. *Higía de la Salud*.

Cingolani, P., Platts, A., Wang, leL., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X. y Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly, 80-92.

Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M. y Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience, giab008.

Dougue, A. N., El-Kholy, M. A., Giuffrè, L., Galeano, G., D Aleo, F., Kountchou, C. L., Nangwat, C., Dzoyem, J. P., Giosa, D., Pernice, I., Shawky, S. M., Ngouana, T. K., Boyom, F. F. y Romeo, O. (2022). Multilocus sequence typing (MLST) analysis reveals many novel genotypes and a high level of genetic diversity in Candida tropicalis isolates from Italy and Africa. Mycoses, 989-1000.

Fan, X., Xiao, M., Liao, K., Kudinha, T., Wang, H., Zhang, L., Hou, X., Kong, F. y Xu, Y. C. (2017). Notable Increasing Trend in Azole Non-susceptible Candida tropicalis Causing Invasive Candidiasis in China (August 2009 to July 2014): Molecular Epidemiology and Clinical Azole Consumption. Frontiers in microbiology, 464.

Favarello, L. M., Nucci, M., Queiroz-Telles, F., Guimarães, T., Salles, M. J., Sukiennik, T. C. T., da Matta, D. A., Melo, A. S. A. y Colombo, A. L. (2021). Trends towards lower azole susceptibility among 200 Candida tropicalis bloodstream isolates from Brazilian medical centres. Journal of global antimicrobial resistance, 199-201.

Garrison, E. y Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv preprint, 3907.

Hilt, E. E. y Ferrieri, P. (2022). Next Generation and Other Sequencing Technologies in Diagnostic Microbiology and Infectious Diseases. *Genes*, 1566.

Jolley, K. A., Bray, J. E. y Maiden, M. C. (2018). Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome open research, 124.

Kanafani, Z. A. y Perfect, J. R. (2008). Antimicrobial resistance: resistance to antifungal agents: mechanisms and clinical impact. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 120-128.

Keighley, C., Gall, M., van Hal, S. J., Halliday, C. L., Chai, L. Y. A., Chew, K. L., Biswas, C., Slavin, M. A., Meyer, W., Sintchenko, V. y Chen, S. C. A. (2022). Whole Genome Sequencing Shows Genetic Diversity, as Well as Clonal Complex and Gene Polymorphisms Associated with Fluconazole Non-Susceptible Isolates of Candida tropicalis. Journal of fungi (Basel, Switzerland), 896.

Keighley, C., Kim, H. Y., Kidd, S., Chen, S. C., Alastruey, A., Dao, A., Bongomin, F., Chiller, T., Wahyuningsih, R., Forastiero, A., Al-Nuseirat, A., Beyer, P., Gigante, V., Beardsley, J., Sati, H., Morrissey, C. O. y Alffenaar, J. W. (2024). Candida tropicalis-A systematic review to inform the World Health Organization of a fungal priority pathogens list. Medical mycology, myae040.

Krassowski, T., Coughlan, A. Y., Shen, X. X., Zhou, X., Kominek, J., Opulente, D. A., Riley, R., Grigoriev, I. V., Maheshwari, N., Shields, D. C., Kurtzman, C. P., Hittinger, C. T., Rokas, A. y Wolfe, K. H. (2018). Evolutionary instability of CUG-Leu in the genetic code of budding yeasts. Nature communications, 1887.

Li, D., Liu, C. M., Luo, R., Sadakane, K. y Lam, T. W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics (Oxford, England), 1674-1676.

Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., Yamashita, H. y Lam, T. W. (2016). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. Methods (San Diego, Calif.), 3-11.

Lima, R., Ribeiro, F. C., Colombo, A. L. y de Almeida, J. N. (2022). The emerging threat antifungal-resistant Candida tropicalis in humans, animals, and environment. Frontiers in fungal biology, 957021.

Magri, M. M., Gomes-Gouvêa, M. S., de Freitas, V. L., Motta, A. L., Moretti, M. L. y Shikanai-Yasuda, M. A. (2013). Multilocus sequence typing of Candida tropicalis shows the presence of different clonal clusters and fluconazole susceptibility profiles in sequential isolates from candidemia patients in Sao Paulo, Brazil. Journal of clinical microbiology, 268-277.

Marcos-Zambrano, L. J., Escribano, P., Bouza, E. y Guinea, J. (2014). Aplicación de las técnicas de tipificación molecular al estudio de brotes hospitalarios de candidemia. Revista Iberoamericana de Micología, 97-103.

Marcos-Zambrano, L. J., Escribano, P., Bouza, E. y Guinea, J. (2014). Production of biofilm by Candida and non-Candida spp. isolates causing fungemia: comparison of biomass production and metabolic activity and development of cut-off points. International journal of medical microbiology : IJMM, 1192-1198.

Megri, Y., Arastehfar, A., Boekhout, T., Daneshnia, F., Hörtnagl, C., Sartori, B., Hafez, A., Pan, W., Lass-Flörl, C. y Hamrioui, B. (2020). Candida tropicalis is the most prevalent yeast species causing candidemia in Algeria: the urgent need for antifungal stewardship and infection control measures. Antimicrobial resistance and infection control, 50.

Mikheenko, A., Saveliev, V., Hirsch, P. y Gurevich, A. (2023). WebQUAST: online evaluation of genome assemblies. Nucleic acids research, W601-W606.

Navarro, P. (2019). Mecanismos de resistencia a azoles en especies de Candida no-albicans. Universitat Rovira i Virgili.

O'Brien, C. E., Oliveira-Pacheco, J., Ó Cinnéide, E., Haase, M. A. B., Hittinger, C. T., Rogers, T. R., Zaragoza, O., Bond, U., & Butler, G. (2021). Population genomics of the pathogenic yeast Candida tropicalis identifies hybrid isolates in environmental samples. PLoS pathogens, e1009138.

Oliveira, J. S., Pereira, V. S., Castelo-Branco, D. S., Cordeiro, R. A., Sidrim, J. J., Brilhante, R. S. y Rocha, M. F. (2020). The yeast, the antifungal, and the wardrobe: a journey into antifungal resistance mechanisms of Candida tropicalis. Canadian journal of microbiology, 377-388.

Quinlan, A. R. y Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics (Oxford, England), 841-842.

Ramirez, J. F. y de la Torre, A. (2008). Evaluación de los cebadores TS3 e ITS4 para la detección de infecciones por Candida spp. en muestras de flujo vaginal y humor acuoso. *Infectio*, 103-110.

Sasani, E., Yadegari, M. H., Khodavaisy, S., Rezaie, S., Salehi, M. y Getso, M. I. (2021). Virulence Factors and Azole-Resistant Mechanism of Candida Tropicalis Isolated from Candidemia. Mycopathologia, 847-856.

Sayers, E. W., Beck, J., Bolton, E. E., Brister, J. R., Chan, J., Comeau, D. C., Connor, R., DiCuccio, M., Farrell, C. M., Feldgarden, M., Fine, A. M., Funk, K., Hatcher, E., Hoeppner, M., Kane, M., Kannan, S., Katz, K. S., Kelly, C., Klimke, W., … Sherry, S. T. (2024). Database resources of the National Center for Biotechnology Information. Nucleic acids research, D33-D43.

Singh, D. K., Tóth, R. y Gácser, A. (2020). Mechanisms of Pathogenic Candida Species to Evade the Host Complement Attack. Frontiers in cellular and infection microbiology, 94.

Spruijtenburg, B., Spruijtenburg, B., Baqueiro, C. C. S. Z., Colombo, A. L., Meijer, E. F. J., de Almeida, J. N., Jr, Berrio, I., Fernández, N. B., Chaves, G. M., Meis, J. F., de Groot, T. y On Behalf Of The Latin American Group For Investigating Candida Tropicalis Resistance (2023). Short Tandem Repeat Genotyping and Antifungal Susceptibility Testing of Latin American Candida tropicalis Isolates. Journal of fungi (Basel, Switzerland), 207.

Tamura, K., Stecher, G. y Sudhir Kumar. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution, 3022-3027.

Tan, A., Abecasis, G. R. y Kang, H. M. (2015). Unified representation of genetic variants. Bioinformatics (Oxford, England), 2202-2204.

Tavanti, A., Davidson, A. D., Johnson, E. M., Maiden, M. C., Shaw, D. J., Gow, N. A. y Odds, F. C. (2005). Multilocus sequence typing for differentiation of strains of Candida tropicalis. Journal of clinical microbiology, 5593-5600.

Vasimuddin, M., Misra, S., Li, H. y Aluru, S. (2019). Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS), 314-324.

Vicente, M., Guzmán, F., Engels, J. y Rao, R. (2006). Genetic characterization and its use in decision-making for the conservation of crop germplasm.

Wang, D., An, N., Yang, Y., Yang, X., Fan, Y. y Feng, J. (2021). Candida tropicalis distribution and drug resistance is correlated with ERG11 and UPC2 expression. Antimicrobial resistance and infection control, 54.

Wang, Q., Li, C., Tang, D. y Tang, K. (2020). Molecular epidemiology of Candida tropicalis isolated from urogenital tract infections. MicrobiologyOpen, e1121.

Wu, J. Y., Zhou, D. Y., Zhang, Y., Mi, F. y Xu, J. (2019). Analyses of the Global Multilocus Genotypes of the Human Pathogenic Yeast Candida tropicalis. Frontiers in microbiology, 900.

Xisto, M. I., Caramalho, R. D., Rocha, D. A., Ferreira-Pereira, A., Sartori, B., Barreto-Bergter, E., Junqueira, M. L., Lass-Flörl, C. y Lackner, M. (2017). Pan-azole-resistant Candida tropicalis carrying homozygous erg11 mutations at position K143R: a new emerging superbug? *The Journal of antimicrobial chemotherapy*, 988-992.

Xu, J. (2021). Is Natural Population of Candida tropicalis Sexual, Parasexual, and/or Asexual? Frontiers in cellular and infection microbiology, 751676.

Yánez, D. D. (2020). Caracterización molecular de las especies de Candida almacenadas en el laboratorio de micología clínica de la Carrera de Bioquímica Clínica, Facultad de Medicina de la Pontificia Universidad Católica del Ecuador en la ciudad de Quito correspondientes al p. Quito: Pontificia Universidad Católica del Ecuador.

Zuza-Alves, D. L., Silva-Rocha, W. P. y Chaves, G. M. (2017). An Update on Candida tropicalis Based on Basic and Clinical Approaches. Frontiers in microbiology, 1927.

Capítulo VIII: Anexos

Anexo A. Metada de las muestras analizadas

Tabla A.1

Información clínica de los aislados

Código muestra	Código laboratorio	Origen	Fecha	UCI (Si / No)
T38	DCT38	Urocultivo	24-02-2020	Sí
T40	DCT40	Urocultivo	24-02-2020	No
T39	DCT39	Urocultivo	24-02-2020	No
T41	DCT41	LB	24-02-2020	Sí
T44	DCT44	Secreción de herida	24-02-2020	No
T43	DCT43	Urocultivo	24-02-2020	Sí
T3	DCT3	Orina	11-01-2020	No
T4	DCT4	Aspirado traqueal	19-02-2019	No
T2	MCT2	Aspirado traqueal	30-01-2019	Sí
T6	DCT6	Hemocultivo	21-03-2019	No
T1	KCT1	Hemocultivo	11-02-2019	Sí

Anexo B. Script personalizado de Python

Figura B. 1

Script de Python para identificar variantes nucleotídicas y aminoacídicas comparando la secuencia de referencia con la de las muestras.

```python
from Bio import SeqIO
from Bio.Seq import Seq
from Bio.Data import CodonTable
from Bio.Align import substitution_matrices as mat
import itertools
import csv
import argparse
from collections import Counter

iupac_codes = {
    'A': {'A'}, 'C': {'C'}, 'G': {'G'}, 'T': {'T'}, 'U': {'T'},
    'R': {'A', 'G'}, 'Y': {'C', 'T'}, 'S': {'G', 'C'}, 'W': {'A', 'T'},
    'K': {'G', 'T'}, 'M': {'A', 'C'}, 'B': {'C', 'G', 'T'}, 'D': {'A', 'G', 'T'},
    'H': {'A', 'C', 'T'}, 'V': {'A', 'C', 'G'}, 'N': {'A', 'C', 'G', 'T'}
}

blosum80 = mat.load('BLOSUM80')

def read_alignment(file_path):
    """Reads a FASTA alignment file and returns a dictionary of sequences."""
    sequences = {}
    for record in SeqIO.parse(file_path, "fasta"):
        sequences[record.id] = str(record.seq)
    return sequences

def resolve_iupac(seq_list):
    """Resolves IUPAC ambiguous codes by choosing the most frequent or, if tied, the most common biologically relevant nucleotide."""
    resolved_seq = []
    for bases in zip(*seq_list):
        base_counts = Counter()
        for base in bases:
            base_counts.update(iupac_codes.get(base, {base}))  # Expand ambiguous bases
        most_common_base = base_counts.most_common(1)[0][0]  # Select most frequent base
        resolved_seq.append(most_common_base)
    return "".join(resolved_seq)

def detect_nucleotide_changes(sequences):
    """Finds nucleotide differences across aligned sequences."""
    sequence_ids = list(sequences.keys())
    ref_seq = resolve_iupac([sequences[sid] for sid in sequence_ids])
    alignment_length = len(ref_seq)

    changes = []
    for i in range(alignment_length):
        column = [sequences[sid][i] for sid in sequence_ids]
        unique_bases = set(column)
        if len(unique_bases) > 1:
            changes.append((i + 1, column))

    return changes

def translate_sequences(sequences):
    """Translates nucleotide sequences to amino acid sequences, handling ambiguous codons safely."""
    standard_table = CodonTable.unambiguous_dna_by_id[1]
    aa_sequences = {}

    for seq_id, sequence in sequences.items():
        resolved_sequence = resolve_iupac([sequence])  # Ensure no ambiguous bases
        codons = [resolved_sequence[i:i+3] for i in range(0, len(resolved_sequence), 3)]

        aa_seq = []
        for codon in codons:
            if len(codon) == 3 and all(base in "ACGT" for base in codon):  # Check for valid codon
                aa_seq.append(str(Seq(codon).translate(table=standard_table, to_stop=False)))
            else:
                aa_seq.append("X")  # Replace invalid codons with "X"

        aa_sequences[seq_id] = "".join(aa_seq)

    return aa_sequences

def detect_amino_acid_changes(aa_sequences):
    """Finds amino acid differences using BLOSUM80 for scoring."""
    sequence_ids = list(aa_sequences.keys())
    ref_aa_seq = aa_sequences[sequence_ids[0]]
    alignment_length = len(ref_aa_seq)

    aa_changes = []
    for i in range(alignment_length):
        column = [aa_sequences[sid][i] for sid in sequence_ids]
        unique_aas = set(column)

        if len(unique_aas) > 1:
            scores = {}
            for aa1 in unique_aas:
                for aa2 in unique_aas:
                    if aa1 != aa2:
                        scores[(aa1, aa2)] = blosum80.get((aa1, aa2), blosum80.get((aa2, aa1), -4))

            aa_changes.append((i + 1, column, scores))

    return aa_changes

```

```
def write_tsv(output_file, nucleotide_changes, amino_acid_changes):
    """Writes detected changes to a TSV file."""
    with open(output_file, "w", newline="") as f:
        writer = csv.writer(f, delimiter="\t")
        writer.writerow(["Position", "Nucleotide Variants"])
        for pos, bases in nucleotide_changes:
            writer.writerow([pos] + bases)

        writer.writerow([])
        writer.writerow(["Position", "Amino Acid Variants", "BLOSUM80 Scores"])
        for pos, aas, scores in amino_acid_changes:
            score_str = ", ".join([f"{pair}: {score}" for pair, score in scores.items()])
            writer.writerow([pos] + aas + [score_str])

def main():
    parser = argparse.ArgumentParser(description="Detect nucleotide and amino acid changes from an alignment file.")
    parser.add_argument("fasta_file", help="Input FASTA alignment file")
    parser.add_argument("output_file", help="Output TSV file")
    args = parser.parse_args()

    sequences = read_alignment(args.fasta_file)
    nucleotide_changes = detect_nucleotide_changes(sequences)
    aa_sequences = translate_sequences(sequences)
    amino_acid_changes = detect_amino_acid_changes(aa_sequences)

    write_tsv(args.output_file, nucleotide_changes, amino_acid_changes)
    print(f"Results saved to {args.output_file}")

if __name__ == "__main__":
    main()
```

Anexo C. Mutaciones previamente descritas

Tabla C.1

Sustituciones aminoacídicas identificadas en este estudio y reportadas previamente por Keighley et al.

Gen	Sustitución aminoácidos
FKS2	F1438L
CDR1	M15V
	P60R
	K332N
	D355N
	F367Y
	A377T
	F623L
	A758S
	D966G
	V1412M
TAC1	D790N