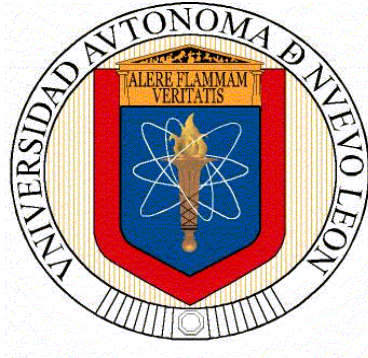


**UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN  
FACULTAD DE CIENCIAS BIOLÓGICAS**



**TESIS**

**CARACTERIZACIÓN FUNCIONAL DE TRES ELEMENTOS  
REGULADORES DE LA EXPRESIÓN GÉNICA QUE INCLUYE  
A UN PROMOTOR Y DOS SECUENCIAS 3'UTRS DE DOS  
GENES DE *Pichia pastoris* (*Komagataella* spp.).**

**POR**

**YANELIS ROBAINAS DEL PINO**


**COMO REQUISITO PARA OBTENER EL GRADO DE  
DOCTORADO EN CIENCIAS  
CON ORIENTACIÓN EN BIOTECNOLOGÍA**

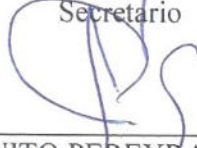
**2023**

CARACTERIZACIÓN FUNCIONAL DE TRES ELEMENTOS REGULADORES  
DE LA EXPRESIÓN GÉNICA QUE INCLUYE A UN PROMOTOR Y DOS  
SECUENCIAS 3'UTRS DE DOS GENES DE *Pichia pastoris* (*Komagataella* spp.)


Comité de tesis


  
DRA. MARTHA GUERRERO OLAZARÁN  
Presidente

  
DR. JOSÉ MARÍA VIADER SALVADÓ  
Secretario

  
DR. BENITO PEREYRA ALFÉREZ  
Vocal 1

  
DRA. MARÍA DEL SOCORRO FLORES GONZÁLEZ  
Vocal 2

  
DRA. SUSANA DE LA TORRE ZAVALA  
Vocal 3

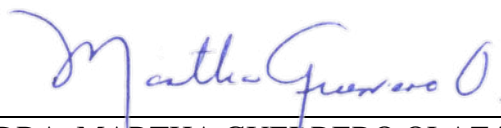
  
DRA. KATIUSHKA AREVALO NIÑO  
Subdirectora de Posgrado



SUBDIRECCIÓN  
DE POSGRADO

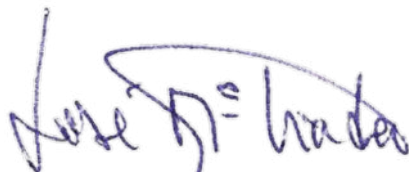
**CARACTERIZACIÓN FUNCIONAL DE TRES ELEMENTOS REGULADORES  
DE LA EXPRESIÓN GÉNICA QUE INCLUYE A UN PROMOTOR Y DOS  
SECUENCIAS 3'UTRS DE DOS GENES DE *Pichia pastoris* (*Komagataella spp.*).**

Dirección de tesis



---

DRA. MARTHA GUERRERO OLAZARÁN  
Director



---

DR. JOSÉ M. VLADER SALVADÓ  
Co-director

## **AVISO DE DERECHOS DE AUTOR**

**DERECHOS RESERVADOS©**

**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta Tesis está protegido, el uso de imágenes, fragmentos de vídeos y demás material contenido que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde se obtuvo mencionando al autor o autores.

## **FINANCIAMIENTO**

Agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada a título de la estudiante de doctorado Yanelis Robainas del Pino con número de matrícula 1940834 y número CVU 897751. Este trabajo fue financiado por el Programa de Apoyo a la Investigación Científica y Tecnológica, PAICYT, CN1599-21 y 604-CN-2022 y por el Fondo Sectorial de Investigación para la Educación SEP-CONACYT CB-2016-01-286093.

Parte del presente trabajo ha sido presentado en forma de cartel en los siguientes congresos:

Yanelis Robainas-del-Pino, Karla Fernández-Cano, José María Viader-Salvadó, Martha Guerrero-Olazarán “Identificación de la secuencia promotora y terminadora de un gen de *Pichia pastoris* (*Komagataella spp.*)” XVIII Congreso Nacional de Biotecnología y Bioingeniería, Sociedad Mexicana de Biotecnología y Bioingeniería. León, Guanajuato, México. 23-28 de junio de 2019.

Yanelis Robainas-del-Pino, Ana Lucía Herrera-Estala, José María Viader-Salvadó, Martha Guerrero-Olazarán. “Functional characterization of a novel non-methanol dependent *Pichia pastoris* promoter.” Pichia 2021. Modalidad Virtual. 24 de marzo-1 de abril de 2021.

Yanelis Robainas-del-Pino, Ana Lucía Herrera-Estala, José María Viader-Salvadó, Martha Guerrero-Olazarán “Caracterización funcional de la secuencia promotora y terminadora de un gen de *Pichia pastoris* (*Komagataella spp.*) no dependientes de metanol.” XIX Congreso Nacional de Biotecnología y Bioingeniería, Sociedad Mexicana de Biotecnología y Bioingeniería. Modalidad Virtual. 27 de septiembre-1 de octubre de 2021.

Yanelis Robainas-del-Pino, Ana Lucía Herrera-Estala, José María Viader-Salvadó, Martha Guerrero-Olazarán. “Functional characterization of the *Pichia pastoris* 1033 gene promoter and transcriptional terminator” Pichia 2022. Graz, Austria. 15-17 de noviembre de 2022.

Además, ha sido publicado en la revista “World Journal of Microbiology and Biotechnology”

Yanelis Robainas-del-Pino, José María Viader-Salvadó, Ana Lucía Herrera-Estala, Martha Guerrero-Olazarán. 2023. Functional characterization of the *Komagataella phaffii* 1033 gene promoter and transcriptional terminator. World Journal of Microbiology and Biotechnology. En impresión.

## AGRADECIMIENTOS

A mis padres y mi hermano por estar siempre ahí, por apoyarme incondicionalmente.

A Lire y Antonio que son mis segundos padres y siempre están pendientes en las buenas y en las malas.

A la Dra. Martha Guerrero, por aceptarme desde tan lejos y confiar en mí para ser parte del Laboratorio L-5, por enseñarme a modificar ADN y a crecerme ante las adversidades.

Al Dr. José María Viader por su paciencia y sus conocimientos que ayudaron a que este proyecto concluyera con éxito.

A todos los profesores del instituto que siempre estuvieron al tanto de mis avances, apoyándome y dándome ánimo para seguir, la Dra. Susana de La Torre, Dra. Socorro Flores, Dra. Miriam Elías, Dr. Benito Pereyra y Dra. Isela Quintero.

A mis compañeros de año que fueron mis cómplices y amigos, Nancy Pentón, Janeth Escudero, Daniel Sánchez, Myriam Ibarra y Fabiola García, juntos fuimos avanzando como equipo.

Al Dr. Antonio que me apoyó sobre todo en la recta final cuando ya van cediendo las fuerzas.

A mis compañeros del L-5, Abel Caballero, Ana Herrera, Karla Fernández, Alber Sánchez, Daniela Ordaz, Fabián Pérez, Rebeca Torres, David Gutiérrez y Mariana Sánchez. A todas las amigas del L-9 que fue mi segundo laboratorio.

A Nancy, Abel, Nancy Piña, Sofía y Abel Alejandro que fueron mi familia en Monterrey todos estos años.

A todos mis amigos que han sido incondicionales, en especial a Rafa que me ha levantado cuando los ánimos se pierden y a Judith que ha sido como una hermana.

A todos les estoy eternamente agradecida.

## DEDICATORIAS

*A Cami y Yovy que están siempre conmigo.*



# ÍNDICE

## Contenido

AVISO DE DERECHOS DE AUTOR.....	IV
FINANCIAMIENTO .....	V
AGRADECIMIENTOS .....	VII
DEDICATORIAS .....	VIII
ÍNDICE .....	IX
ÍNDICE DE TABLAS .....	XVI
ÍNDICE DE FIGURAS.....	XVIII
LISTA DE SÍMBOLOS Y ABREVIATURAS .....	XXV
RESUMEN.....	XXVIII
ABSTRACT.....	XXIX
1. INTRODUCCIÓN .....	1
2. ANTECEDENTES .....	4
2.1 Producción de proteínas recombinantes .....	4
2.2 Sistema hospedero <i>Pichia Pastoris</i> . Ventajas. ....	5
2.2.1 Sistema de expresión de <i>P. pastoris</i> . ....	6
2.2.1.1 Cepas de <i>P. pastoris</i> para la producción de proteínas heterólogas. ....	7
2.2.1.2 Vectores de expresión de proteínas heterólogas en <i>P. pastoris</i> . ....	8
2.2.1.3 Promotores utilizados en <i>P. pastoris</i> , inducibles y constitutivos.....	10
2.2.1.4 Necesidad de nuevos promotores en <i>Pichia pastoris</i> .....	13
2.2.1.5 Terminadores de la transcripción .....	14
2.2.1.6. Factores transcripcionales y herramientas para su predicción .....	15

2.2.1.7. Identificación de secuencias promotoras y terminadoras de la transcripción .....	16
2.2.1.8. Cuantificación de la expresión genética mediante RT-qPCR.....	18
2.2.1.9. Interacción promotor-terminador de la transcripción .....	19
2.2.1.10. El gen PAS_chr3_1033 (gen <i>1033</i> ).....	19
3. JUSTIFICACIÓN .....	21
4. HIPÓTESIS .....	22
5. OBJETIVOS .....	23
5.1. Objetivo general .....	23
5.2. Objetivos específicos.....	23
6. MATERIALES Y MÉTODOS .....	24
6.1. Lugar de trabajo.....	24
6.2. Material biológico y medios de cultivo .....	24
6.3. Materiales y equipos.....	25
6.4. Estrategia general experimental .....	26
6.5. Identificación de las secuencias del promotor y terminador de la transcripción del gen <i>1033</i> .....	27
6.5.1. Mapeo de los transcritos contra el genoma de referencia de <i>P. pastoris</i> .....	28
6.5.2. Identificación de motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas y determinación del núcleo del P <sub>1033</sub> .....	28
6.5.3. Identificación de secuencias consenso en terminadores de levaduras .....	29
6.6. Identificación de factores de transcripción que se unen a la secuencia del P <sub>1033</sub> ..	30
6.7. Diseño y síntesis de la secuencia sintética que incluyen las regiones de P <sub>1033</sub> y T <sub>1033</sub> .....	30
6.8. Construcción del vector de expresión pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> . .....	32

6.8.1. Transformación de <i>E. coli</i> TOP10 con los plásmidos pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII.....	33
6.8.2. Preparación y purificación de los vectores pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII .....	35
6.8.3. Digestión de pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII con las enzimas <i>XhoI</i> y <i>AatII</i> .....	35
6.8.3.1. Purificación de pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII digeridos con <i>XhoI</i> y <i>AatII</i> .....	36
6.8.4. Construcción y propagación del plásmido pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> .....	37
6.8.5. Construcción del vector pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	41
6.9. Construcción de las cepas recombinantes de <i>P. pastoris</i> KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	43
6.9.1. Extracción y purificación de los plásmidos pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	43
6.9.2. Digestión de pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> con la enzima <i>SaII</i> ...	44
6.9.3. Transformación de <i>P. pastoris</i> con pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	45
6.9.4. Caracterización genotípica de las cepas KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	47
6.10. Caracterización funcional de los elementos reguladores integrados en las cepas recombinantes construidas.....	49
6.10.1. Detección de transcritos por RT-PCR .....	49
6.10.1.1. Cultivo de las cepas KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	50
6.10.1.2. Preparación de RNA de las células de las cepas de <i>P. pastoris</i> construidas .....	50
6.10.1.3. Síntesis del DNA complementario (DNAc).....	52

6.10.1.4. Reacción de PCR para detección de transcritos específicos .....	53
6.10.2. Detección del producto proteico del gen reportero (fitasa) .....	54
6.10.2.1. Determinación de actividad volumétrica de fitasa extracelular .....	54
6.10.2.2. Detección de fitasa extracelular por SDS-PAGE.....	57
6.10.2.2.1. Ultrafiltración de sobrenadantes libres de células de los cultivos ...	57
6.7.2.2.2. Tratamiento para N-desglicosilación de las proteínas extracelulares de los concentrados de los cultivos de cepas recombinantes construidas. ....	58
6.10.2.2.3. Análisis de proteínas extracelulares por SDS-PAGE.....	59
6.11. Evaluación de la funcionalidad de la combinación de los elementos reguladores P <sub>1033</sub> -T <sub>AOX1</sub> y P <sub>1033</sub> -T <sub>1033</sub> en cepas unicopias cultivadas en dos fuentes de carbono .....	60
6.11.1. Identificación de clonas unicopia de las cepas KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	60
6.11.1.1. Cultivos en matraz agitado.....	61
6.11.1.2. Determinación de proteínas extracelulares de los cultivos .....	62
6.11.1.3. Determinación del rendimiento proteína extracelular por unidad de biomasa .....	63
6.11.2. Determinación de la dosis génica de las cepas unicopias identificadas .....	63
6.11.2.1. Cultivos de las clonas seleccionadas como unicopias .....	63
6.11.2.2. Preparación de DNA genómico de la identificación de las posibles clonas unicopias .....	64
6.11.2.3. Curvas para determinar la eficiencia de amplificación de los genes <i>FTEII</i> y <i>GAPDH</i> .....	65
6.11.3. Actividad transcripcional del P <sub>1033</sub> en combinación con T <sub>AOX1</sub> y T <sub>1033</sub> .....	66
6.11.3.1. Cultivos para cinéticas de crecimiento y de expresión de las cepas unicopias .....	66
6.11.3.2. Determinación de los niveles de transcritos de los genes <i>1033</i> , <i>FTEII</i> y <i>GAPDH</i> en células de las cepas recombinantes cultivadas en glucosa y glicerol	66

6.11.3.3. Determinación del rendimiento producto extracelular por unidad de biomasa .....	72
7. Resultados .....	75
7.1. Identificación de las secuencias del promotor y terminador de la transcripción del gen <i>1033</i> .....	75
7.1.1. Mapeo de los transcritos contra el genoma de referencia de <i>P. pastoris</i> .....	75
7.1.2. Identificación de motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas y determinación del núcleo del P <sub>1033</sub> .....	77
7.2. Identificación de factores de transcripción que se unen a la secuencia del P <sub>1033</sub> ..	77
7.3. Diseño y síntesis de la secuencia sintética que incluyen las regiones de P <sub>1033</sub> y T <sub>1033</sub> .....	79
7.4. Construcción de los vectores pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	82
7.4.1. Transformación de <i>E. coli</i> TOP10 con los plásmidos pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII .....	82
7.4.2. Preparación y purificación de los vectores pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII .....	82
7.4.3. Digestión de pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII con las enzimas <i>XhoI</i> y <i>AatII</i> .....	83
7.4.3.1 Purificación de pUCIDTP <sub>1033</sub> T <sub>1033</sub> y pGAHFTEII digeridos con <i>XhoI</i> y <i>AatII</i> .....	84
7.4.4. Construcción y propagación del plásmido pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> .....	85
7.4.5. Construcción y propagación del plásmido pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	89
7.5. Construcción de las cepas recombinantes de <i>P. pastoris</i> KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	91
7.5.1. Extracción y purificación de los plásmidos pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> y digestión con <i>Sal I</i> .....	91

7.5.2. Transformación de <i>P. pastoris</i> con pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	92
7.5.3. Caracterización genotípica de las cepas KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	92
7.6. Caracterización funcional de los elementos reguladores integrados en las cepas recombinantes construidas.....	95
7.6.1. Detección de transcritos por RT-PCR .....	95
7.6.2 Detección del producto proteico del gen reportero (fitasa) .....	98
7.6.2.1. Determinación de actividad volumétrica de fitasa extracelular .....	98
7.6.2.2. Detección de fitasa extracelular por SDS-PAGE.....	98
7.7. Evaluación de la funcionalidad de la combinación de los elementos reguladores P <sub>1033</sub> -T <sub>AOX1</sub> y P <sub>1033</sub> -T <sub>1033</sub> en cepas unicopias cultivadas en dos fuentes de carbono...	100
7.7.1. Determinación de proteínas extracelulares y rendimiento proteína/biomasa .....	100
7.7.2. Identificación de clonas unicopia de las cepas KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> .....	101
7.7.2.1. Cultivos y preparación de DNA genómico de las posibles clonas unicopias .....	102
7.7.2.2. Curvas para determinar la eficiencia de amplificación de los genes FTEII y GAPDH.....	104
7.7.3. Actividad transcripcional del P <sub>1033</sub> en combinación con T <sub>AOX1</sub> y T <sub>1033</sub> .....	106
7.7.3.1. Cinéticas de crecimiento de las cepas unicopias.....	106
7.7.3.2. Determinación de los niveles de transcritos de los genes <i>FTEII</i> , <i>1033</i> y <i>GAPDH</i> en células de las cepas recombinante cultivadas en glucosa y glicerol	107
7.7.3.3. Determinación del rendimiento producto extracelular por unidad de biomasa .....	115
8. DISCUSIÓN .....	118

9. CONCLUSIONES .....	127
10. PERSPECTIVAS .....	129
APÉNDICES.....	130
APÉNDICE A: Oligonucleótidos y sondas usados en este trabajo.....	130
APÉNDICE B: Preparación de medios, soluciones y reactivos.....	131
11. BIBLIOGRAFÍA .....	139

## ÍNDICE DE TABLAS

<b>Tabla 1.</b> Resumen de las principales ventajas de <i>P. pastoris</i> como sistema de expresión de proteínas recombinantes. ....	6
<b>Tabla 2.</b> Cepas comunes de la levadura metilotrófica <i>P. pastoris</i> . ....	8
<b>Tabla 3.</b> Vectores de expresión comerciales y marcadores de selección en <i>Pichia pastoris</i> . ....	9
<b>Tabla 4.</b> Promotores de <i>P. pastoris</i> identificados y caracterizados recientemente y los más usados para la producción de proteínas recombinantes. ....	11
<b>Tabla 5.</b> Configuración de la mezcla de reacción de los cortes enzimático de ambos plásmidos. ....	36
<b>Tabla 6.</b> Reacciones de ligación y controles negativos que se realizaron con los fragmentos purificados de inserto y vector. ....	38
<b>Tabla 7.</b> Mezcla de reacción de la digestión de pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> con las enzimas <i>XhoI</i> y <i>AatII</i> . ....	40
<b>Tabla 8.</b> Mezcla de reacción para la caracterización por PCR del plásmido pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> . ....	40
<b>Tabla 9.</b> Condiciones del termociclador para el ensayo de PCR. ....	40
<b>Tabla 10A.</b> Reacciones de digestión enzimática del plásmido pP <sub>1033</sub> -FTEII-T <sub>1033</sub> con <i>XhoI</i> y/o <i>Bsu36I</i> . ....	43
<b>Tabla 11.</b> Reacción de digestión de pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> . con la enzima <i>Sall</i> . ....	44
<b>Tabla 12.</b> Mezcla de reacción de PCR para verificar la integración del casete de expresión. ....	48
<b>Tabla 13.</b> Volumen y concentración final de los reactivos empleados en la reacción de retrotranscripción. ....	52
<b>Tabla 14.</b> Condiciones del termociclador para el ensayo de PCR para la detección de transcritos. ....	54
<b>Tabla 15.</b> Volúmenes para la preparación de los estándares para la curva de calibración. ....	56



<b>Tabla 16.</b> Reacciones de desnaturalización.....	59
<b>Tabla 17.</b> Reacciones de desglicosilación.....	59
<b>Tabla 18.</b> Proporciones de los reactivos que se utilizaron para preparar el gel de poliacrilamida.....	60
<b>Tabla 19.</b> Concentración de los estándares de BSA para la curva de calibración.....	62
<b>Tabla 20.</b> Coctel de reacción para síntesis de DNAc. ....	68
<b>Tabla 21.</b> Volúmenes de reacción para los ensayos de qPCR con sondas de hidrólisis o SYBR Green. ....	70
<b>Tabla 22.</b> Programa del termociclador para los ensayos de qPCR empleando SYBR Green. ....	71
<b>Tabla 23.</b> Programa del termociclador para los ensayos de qPCR empleando sondas de hidrólisis.....	71
<b>Tabla 24.</b> Volúmenes de las alícuotas tomadas de los cultivos a diferentes tiempos. ....	72
<b>Tabla 25.</b> Posición de las regiones no transcritas río arriba y río abajo del gen <i>1033</i> respecto al genoma de la cepa de <i>P. pastoris</i> CBS7435 a partir de los cinco transcriptomas analizados.....	75
<b>Tabla 26.</b> Posibles sitios de unión de factores de transcripción (TFBS) de la secuencia nucleotídica del P <sub>1033</sub> de <i>P. pastoris</i> con un p< 0.05 y matriz de puntuación utilizada para su identificación. ....	78
<b>Tabla 27.</b> Posibles factores de transcripción de <i>P. pastoris</i> que interaccionan con la secuencia nucleotídica del P <sub>1033</sub> . ....	78
<b>Tabla 28.</b> Resultados del análisis de las secuencias del “primer forward” PP1033 y el “primer reverse” FTE2 empleando el programa de Oligoanalyzer.....	81
<b>Tabla 29.</b> Resultados de la transformación de células de <i>Pichia pastoris</i> KM71 con el plásmido pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> . ....	92
<b>Tabla 30.</b> Preparación de DNA genómico para la determinación de dosis génica por qPCR. ....	104
<b>Tabla 31.</b> Cuantificación relativa de dosis génica por qPCR: Valores de ciclo umbral (Ct), media, coeficiente de variación (CV%) y ng de los genes <i>GAP</i> y <i>FTEII</i> de las clonas analizadas.....	106

## ÍNDICE DE FIGURAS

<b>Figura 1.</b> Diagrama de los vectores de <i>Pichia pastoris</i> pPink $\alpha$ -HC y pPICZ $\alpha$ -E.....	9
<b>Figura 2.</b> Estrategia general para la caracterización funcional de tres elementos reguladores de la expresión génica que incluye a un promotor y dos secuencias 3'UTRs de dos genes de <i>Pichia pastoris</i> . .....	27
<b>Figura 3.</b> Metodología empleada para identificar los factores de transcripción que se unen a la secuencia del P <sub>1033</sub> . .....	30
<b>Figura 4.</b> Estrategia y diseño de construcción de los vectores de expresión pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> . .....	33
<b>Figura 5.</b> Resultado del programa NEBicalculator. Relación molar inserto-vector para la reacción de ligación de acuerdo con los tamaños de inserto y vector y 50 ng de vector. ....	37
<b>Figura 6.</b> Plásmido de pP <sub>1033</sub> -FTEII-T <sub>1033</sub> conteniendo el casete de expresión que contiene los elementos reguladores y los sitios de corte. ....	41
<b>Figura 7.</b> Representación gráfica de los casetes de expresión integrados en el genoma de las levaduras transformadas con los plásmidos de pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> y pP <sub>1033</sub> -FTEII-T <sub>1033</sub> se muestran las regiones de hibridación de los oligonucleótidos y tamaño de los fragmentos esperados amplificados por PCR con cada par de oligonucleótidos.....	49
<b>Figura 8.</b> Esquema de las etapas en las reacciones para la síntesis de DNAc. ....	68
<b>Figura 9.</b> Esquema de la estructura del montaje para las reacciones de qPCR.....	70
<b>Figura 10.</b> Resultados del mapeo del análisis de los cinco transcriptomas visualizadas con el programa Integrative Genomics Viewer (IGV) identificando las regiones río arriba y río abajo del CDS del gen <i>I033</i> . Las subregiones donde no alinearon las lecturas de ARN se muestran con una flecha de doble punta. También se muestra la orientación del gen <i>I033</i> y los genes adyacentes respecto al genoma de la cepa <i>P. pastoris</i> CBS 7435, y las regiones 5'URT y 3'UTR del gen <i>I033</i> . nt: nucleótidos. ....	76
<b>Figura 11.</b> Representación esquemática de las regiones no transcritas río arriba y río abajo, 5'UTR, 3'UTR y CDS del gen <i>I033</i> indicando las posiciones respecto al genoma	

de <i>P. pastoris</i> CBS 7435 y el tamaño de las regiones del promotor y terminador del gen <i>1033</i> . nt: nucleótidos. ....	77
<b>Figura 12.</b> Representación esquemática de la secuencia P <sub>1033</sub> . Se muestran la región homopolimérica de 21 nucleótidos de desoxitimidina (Ts), la caja tipo TATA, el sitio de inicio de la transcripción (TSS) y el elemento tipo iniciador de mamíferos (INR). El inicio de la transcripción se indica con +1. Debajo de la representación lineal de la secuencia P <sub>1033</sub> se muestran en gris los factores de transcripción (TF) putativos identificados mediante el empleo consecutivo de la herramienta bioinformática TRAP, base de datos UniProt, y la técnica Best Reciprocal Blast Hit. ....	79
<b>Figura 13.</b> Secuencia sintética integrada en el plásmido PUCIDTP <sub>1033</sub> T <sub>1033</sub> que incluye la región promotora del gen P <sub>1033</sub> (rojo) desde el sitio <i>AatII</i> y la secuencia pre-pro del factor alfa (lila) con codones preferenciales para <i>P. pastoris</i> hasta el sitio <i>XhoI</i> , el sitio <i>NotI</i> seguido río abajo por la secuencia terminadora del gen P <sub>1033</sub> (azul), seguido de 500 nucleótidos del vector pPIC9 hasta el sitio <i>Bsu36I</i> . ....	80
<b>Figura 14.</b> Resultado del programa Primer3 v0.4.0 al introducir la secuencia molde a amplificar y el par de oligonucleótidos PP1033 y FTE2. ....	81
<b>Figura 15.</b> Resultado del programa PrimerBlast al introducir la secuencia de los oligonucleótidos PP1033 y FTE2, y la secuencia del casete de expresión. Además, se seleccionó el genoma de <i>P. pastoris</i> . ....	82
<b>Figura 16.</b> Electroforesis en gel de agarosa al 0.8 % en TBE 1X de preparaciones de DNA plasmídico. M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carriles 1 y 2: DNA plasmídico pUCIDTP <sub>1033</sub> T <sub>1033</sub> . Carriles 3 y 4: DNA plasmídico pGAHFTEII. ....	83
<b>Figura 17.</b> Electroforesis en gel de agarosa al 0.8 % (C) de la digestión enzimática de los DNA plasmídicos pPUCIDTP <sub>1033</sub> T <sub>1033</sub> (A) y pGAHFTEII (B). M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: DNA plasmídico pUCIDTP <sub>1033</sub> T <sub>1033</sub> ; Carril 2: DNA plasmídico pUCIDTP <sub>1033</sub> T <sub>1033</sub> digerido con <i>XhoI</i> ; Carril 3: DNA plasmídico pUCIDTP <sub>1033</sub> T <sub>1033</sub> digerido con <i>XhoI</i> y <i>AatII</i> ; Carril 4: DNA plasmídico pGAHFTEII; Carril 5: DNA plasmídico pGAHFTEII digerido con <i>XhoI</i> ; Carril 6: DNA plasmídico pGAHFTEII digerido con <i>XhoI</i> y <i>AatII</i> . ....	84

**Figura 18.** Electroforesis en gel de agarosa al 0.8 % en TBE 1X de los fragmentos purificados para llevar a cabo la ligación. M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: fragmento de DNA inserto de 423 pb; Carril 2: fragmento de DNA vector de 7,694 pb.....85

**Figura 19.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la ligación del fragmento AatII-XhoI del vector pGAHFTEII con el fragmento P<sub>1033</sub>-FTEII. M: marcador de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: vector e inserto con DNA T4 ligasa después de la incubación; Carril 2: vector e inserto (control negativo de ligación).....86

**Figura 20.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carriles 1-5: DNA plasmídico de cuatro colonias obtenidas de la transformación de E. coli TOP10 con el producto de ligación.....87

**Figura 21.** Electroforesis en gel de agarosa (B) al 0.8% en TBE 1X de la caracterización enzimática del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con XhoI y AatII (A). M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>; Carril 2: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> digerido con XhoI y AatII. ....88

**Figura 22.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con los iniciadores PP1033 y FTE2. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control (-) de la PCR. Carriles 2-4: Producto amplificado del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> de tres colonias de E. coli transformadas. ....88

**Figura 23.** Electroforesis en geles de agarosa al 0.8% del DNA plasmídico A) enviado por la compañía Gene Universal. M: Marcador de tamaño molecular DNA ladder 1kb de Promega. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> y B) obtenido de la extracción de DNA plasmídico de colonias transformadas de E. coli TOP10. M: Marcador de tamaño molecular DNA ladder 1kb de Promega. Carril 1-3: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> obtenido de E. coli TOP10 transformada.....89

**Figura 24.** Electroforesis en geles de agarosa (B) al 0.8% de la caracterización de pP<sub>1033</sub>-FTEII-T<sub>1033</sub> (A) con las enzimas de restricción XhoI y Bsu36I A. M: Marcador de tamaño

molecular DNA ladder 1kb de Promega. Carril 1: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> sin enzimas. Carril 2: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con *Xho*I y *Bsu*36I. Carril 3: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con *Xho*I. ....90

**Figura 25.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de preparaciones de DNA plasmídico a mediana escala. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). (A) pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> digerido con la enzima *Sal*I. (B) pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> digerido con la enzima *Sal*I.....91

**Figura 26.** Electroforesis en geles de agarosa al 0.8% en TBE 1X de DNA genómico aislado de cepas de *P. pastoris* KM71 transformadas con el vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. A) M: marcador de tamaño molecular 1 kb DNA Ladder. Carriles 1-5: DNA genómico de la placa A (colonias 1 a 5). Carriles 6-10: DNA genómico de la placa B (colonias 1 a 5). B) M: marcador de tamaño molecular 1 kb DNA Ladder. Carriles 1-5: DNA genómico de la placa 2 (colonias 1 a 5). Carriles 6-10: DNA genómico de la placa 3 (colonias 1 a 5). ...93

**Figura 27.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR con los oligonucleótidos PP1033 y FTE2 del DNA genómico aislado de colonias de la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: control negativo de la PCR. Carriles 3-6: Producto amplificado del DNA genómico de 4 colonias transformantes de *P. pastoris* KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. 94

**Figura 28.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR con los oligonucleótidos FTE1 y 3TH del DNA genómico aislado de las colonias de la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub> A). M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: DNA genómico *P. pastoris* de la placa A (1-5). Carriles 8-12: DNA genómico *P. pastoris* de la placa B (1-5). B) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1-5: DNA genómico *P. pastoris* de la placa 2 (1-5). Carriles 6-9: DNA genómico *P. pastoris* de la placa 3 (1-4). Carril 10: control positivo de PCR (pGAHFTEII). Carril 11: Control negativo de PCR. C) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: PCR de colonia de la placa A (6-10). PCR de colonia de la

placa B (6-10). D) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: PCR de colonia de la placa 2 (6-10). PCR de colonia de la placa 3 (7-10).....95

**Figura 29.** Electroforesis en gel de agarosa al 2.0% del producto amplificado por RT-PCR a partir de RNA de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: Marcador de tamaño molecular de 100 pb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control negativo de la reacción de PCR, coctel 1; Carril 2: control negativo de la retrotranscripción, coctel 1; Carril 3: producto amplificado del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (control positivo de la reacción); Carril 4: producto amplificado de la muestra con los oligonucleótidos FTE1 y FTE2, coctel 1; Carril 5: control negativo de la reacción de PCR, coctel 2; Carril 6: control negativo de la retrotranscripción, coctel 2; Carril 7: producto amplificado de la muestra con los oligonucleótidos 5qACT y 3qACT, coctel 2. ....96

**Figura 30.** Electroforesis en gel de agarosa al 0.8% en TBE 1X del RT-PCR a partir del RNA total a las 12 h de cultivo de células de *P. pastoris* KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. M1: marcador de tamaño molecular 1kb DNA ladder. Carril 1: control negativo PCR. Carril 2: control negativo RT. Carril 3: control positivo PCR (pP<sub>1033</sub>-FTEII-T<sub>1033</sub>). Carril 4: producto amplificado de la muestra con los oligonucleótidos FTE1 y FTE2. Carril 5: control negativo PCR ( $\beta$ -actina). Carril 6: control negativo RT. Carril 7: producto amplificado de la muestra con los oligonucleótidos 5qACT y 3qACT. M2: marcador de tamaño molecular de 100 pb. ....97

**Figura 31.** Gel de SDS-poliacrilamida de sobrenadantes concentrados (SC) de cultivo por ultrafiltración y tratados con o sin Endo H. Carril M: marcador de tamaño molecular. Carril 1: SC KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> tratado con Endo H. Carril 2: SC KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> sin Endo H. Carril 3: Endo H. Carril 4: SC KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> tratado con Endo H. Carril 5: SC KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> sin Endo H. ....99

**Figura 32.** Representación de diagrama de caja de la concentración de biomasa (A) la concentración de proteína extracelular (B) y el rendimiento de proteína/biomasa extracelular (C) para cultivos BMG de 24 h de 32 y 20 clones recombinantes (transformantes His+) de cada cepa (recuadro blanco, KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>; caja gris, KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>). Las líneas centrales muestran las medianas y los límites

de las cajas indican el 25 y el 75 percentil. Los bigotes se extienden 1.5 veces del rango intercuartil desde los percentiles 25 y 75 y los valores atípicos se representan con puntos.

..... 101

**Figura 33.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR de 3 clonas de cada construcción. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). (A) Caracterización con los oligonucleótidos PP1033 y FTE2. Carril 1: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: control negativo de la PCR. Carril 3: Clona 21, Carril 4: Clona 29, Carril 5: Clona 8. (B) Caracterización con los oligonucleótidos FTE1 y 3THIPO1. Carril 1: control negativo de la PCR Carril 2: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Carril 3: Clona 29, Carril 4: Clona 37, Carril 5: Clona 36. .... 103

**Figura 34.** Curvas de eficiencia de la amplificación por PCR de los genes *FTEII* (A) y *GAPDH* (B). .... 105

**Figura 35.** Cinética de crecimiento en medios BMGlc (●, ▲) y BMGly (■, ◆) para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (●, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▲, ◆). Los puntos representan la media de tres experimentos cinéticos independientes (coeficiente de variación inferior al 5%). .... 107

**Figura 36.** Curvas de eficiencia de la amplificación por PCR de los genes *FTEII* (A), *GAPDH* (B), *YPT1* (C) y *1033* (D). .... 109

**Figura 37.** Niveles de transcripción relativos del gen *FTEII* de para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas (p < 0,05) entre las dos cepas cultivadas en la misma fuente de carbono. .... 111

**Figura 38.** Niveles de transcripción relativos del gen *1033* para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. .... 112

**Figura 39.** Niveles de transcripción relativos del gen GAPDH de para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□ , ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas (p < 0,05) entre las dos cepas cultivadas en la misma fuente de carbono..... 113

**Figura 40.** Relación del nivel de transcripción de FTEII respecto a1033 en las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□ , ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos. Los datos se presentan como la media de tres cultivos independientes..... 114

**Figura 41.** Relación del nivel de transcripción de GAPDH /FTEII (A) y GAPDH /P<sub>1033</sub> (B) en las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□ , ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos. Los datos se presentan como la media de tres cultivos independientes. .... 115

**Figura 42.** Rendimiento de producto/biomasa extracelular (Yp/x) para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□ , ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▣, ▤) en los medios en glucosa BMGlc ((□, ▣) y glicerol BMGly ((■, ▤) a diferentes tiempos. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas (p < 0,05) entre las dos cepas cultivadas en la misma fuente de carbono..... 116



## LISTA DE SÍMBOLOS Y ABREVIATURAS

Los símbolos de elementos y compuestos químicos corresponden a los empleados por la Unión Internacional de Química Pura y Aplicada (*International Union of Pure and Applied Chemistry*, IUPAC). Los símbolos empleados en la cuantificación de magnitudes físicas, tanto en unidades básicas como derivadas, corresponden a los empleados por el Sistema Internacional de Unidades.

<i>AOX1</i>	Gen que codifica para la enzima alcohol oxidasa I
<i>AOX2</i>	Gen que codifica para la enzima alcohol oxidasa II
<i>ARG4</i>	Gen que codifica para la enzima argininosuccinato liasa
ASB	Albúmina sérica bovina
BMG	Medio mínimo amortiguado suplementado con glicerol 1 %
BMGly	Medio mínimo amortiguado suplementado con glicerol 30 mM
BMGlc	Medio mínimo amortiguado suplementado con glucosa 30 mM
Cat.	Número de catálogo
DNAc	DNA complementario
CDS	Secuencia codificante de un gen
CV	Coefficiente de variación
$C_t$	Ciclo umbral
DNA	Ácido desoxirribonucleico
dNTP	Desoxirribonucleótido trifosfato
DO <sub>600 nm</sub>	Densidad óptica a 600 nm
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ácido etilendiaminotetraacético
et al.	Y otros ( <i>et alter</i> )
FTEII	Fitasa termoestable II
<i>FTEII</i>	Gen que codifica para la fitasa termoestable II
g	Gravedades
gDCW	Gramos de peso celular seco ( <i>grams of dry cell weight</i> )
g/L	Gramos por litro
GAP	Gliceraldehído-3-fosfato deshidrogenasa

<i>GAP</i>	Gen que codifica para la enzima gliceraldehído-3-fosfato deshidrogenasa
h	Horas
h-1	Por hora
<i>HIS4</i>	Gen que codifica para la histidinol deshidrogenasa
<i>K. phaffii</i>	<i>Komagataella phaffii</i> (Nombre actual de <i>P. pastoris</i> )
kb	Kilobase o mil pares de bases
kDa	Kilodalton
KM71GAHFTEII	Cepa recombinante de <i>P. pastoris</i> productora y secretora de una fitasa termoestable regulada por el promotor GAP
KM71P <sub>1033</sub> -FTEII-T <sub>AOX1</sub>	Cepa recombinante de <i>P. pastoris</i> productora y secretora de una fitasa termoestable regulada por el promotor P <sub>1033</sub> y con el terminador T <sub>AOX1</sub>
KM71P <sub>1033</sub> -FTEII-T <sub>1033</sub>	Cepa recombinante de <i>P. pastoris</i> productora y secretora de una fitasa termoestable regulada por el promotor P <sub>1033</sub> y con el terminador P <sub>1033</sub>
KM71pPIC9	Cepa recombinante de <i>P. pastoris</i> transformada con el plásmido pPIC9
L	Litros
LB	Medio de cultivo Luria-Bertani
log	Logaritmo
M	Concentración molar
mg/mL	Miligramos por mililitro
min	Minutos
mM	Milimolar
Mut <sup>s</sup>	Fenotipo de utilización de metanol lento
<i>P. pastoris</i>	<i>Pichia pastoris</i>
<i>P<sub>AOX1</sub></i>	Promotor del gen que codifica para la enzima alcohol oxidasa I
<i>P<sub>GAP</sub></i>	Promotor del gen que codifica para la enzima gliceraldehído-3-fosfato deshidrogenasa
pb	Pares de bases
PCR	Reacción en cadena de la polimerasa
pH	Potencial de iones hidrógeno
qPCR	Reacción en cadena de la polimerasa cuantitativa
RDB	Medio de regeneración con dextrosa

RNA	Ácido ribonucleico
RNasa	Ribonucleasa
rpm	Revoluciones por minuto
RT	Retrotranscripción
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Dodecil sulfato de sodio
SDS-PAGE	Electroforesis en gel de poliacrilamida con dodecilsulfato de sodio
TE	Amortiguador de pH Tris-HCl-EDTA
Tris	Trizma o tris (hidroximetil) aminometano
TSNT	Tritón-SDS-NaCl-Tris-HCl
U	Unidades enzimáticas
UV	Radiación ultravioleta
X	Número de veces la concentración
$Y_{x/s}$	Rendimiento de biomasa por unidad de sustrato
$Y_{p/x}$	Rendimiento de producto por unidad de biomasa
YNB	Base nitrogenada de levadura
YPD	Medio de cultivo con extracto de levadura, peptona y dextrosa
<i>YPTI</i>	Gen que codifica para una proteína de unión a guanosín trifosfato, se empleó como gen normalizador
%	Porcentaje
$\mu$	Velocidad específica de crecimiento
$\mu\text{g/mL}$	Microgramos por mililitro
$\mu\text{L}$	Microlitros
$\mu\text{M}$	Micromolar

## RESUMEN

La levadura *Pichia pastoris* ha sido ampliamente usada en la producción de proteínas heterólogas por sus múltiples ventajas. Los promotores son elementos determinantes en un vector de expresión, sin embargo, un promotor fuerte no siempre es la opción más favorable para la producción de proteínas heterólogas, especialmente si el plegamiento correcto de la proteína y/o el procesamiento postraduccional es el paso limitante. El terminador transcripcional es otro elemento regulador que puede modificar los niveles de expresión del gen heterólogo ya que influye en la estabilidad del transcrito.

En este estudio se caracterizaron funcionalmente tres elementos reguladores que incluye un promotor y dos secuencias 3'UTRs de dos genes de *P. pastoris*. Se identificaron las regiones reguladoras del gen PAS\_chr3\_1033 (*1033*), el promotor ( $P_{1033}$ ) y el terminador transcripcional ( $T_{1033}$ ), un gen constitutivo débil no dependiente de metanol. Se diseñaron secuencias sintéticas que se integraron en dos casetes de expresión conteniendo la combinación promotor y terminador de la transcripción del gen *1033* ( $P_{1033}$ - $T_{1033}$ ) y la combinación de  $P_{1033}$  y región terminadora de la transcripción del gen *AOX1* ( $P_{1033}$ - $T_{AOX1}$ ), en ambos casos funcionalmente asociados a la secuencia pre-pro del factor alfa de *S. cerevisiae* y al gen reportero *FTEII* que codifica para una fitasa bacteriana. Los casetes de expresión conformaron dos vectores que se integraron en el genoma de *P. pastoris* para construir las cepas KM71/ $P_{1033}$ -*FTEII*- $T_{AOX1}$  (A) y KM71/ $P_{1033}$ -*FTEII*- $T_{1033}$  (H) y se determinó la funcionalidad de las regiones reguladoras por detección de transcritos (RT-PCR), detección del producto proteico por SDS-PAGE y actividad enzimática de fitasa. En clonas unicopias, se evaluó el impacto de las combinaciones de elementos reguladores en los niveles de transcripción de los genes heterólogo, *1033* y *GAPDH* en células cultivadas en glucosa o glicerol, y en el rendimiento producto extracelular/biomasa ( $Y_{p/x}$ ). La presencia de transcritos del gen reportero se evidenció por la detección en gel de agarosa de una banda de 536 pb producto de la RT-PCR. La actividad de fitasa extracelular a las 24 horas del cultivo en glicerol se detectó con niveles de  $0.23 \pm 0.01$  (cepa A) y de  $0.11 \pm 0.01$  (cepa H) U/mL, además se detectó por SDS-PAGE, en los sobrenadantes de los cultivos, una banda con una masa molecular aparente de 39 kDa que corresponde a la masa molecular de la fitasa recombinante producida de forma extracelular, confirmando la funcionalidad de la combinación de las regiones reguladoras evaluadas. El  $P_{1033}$  mostró una actividad transcripcional del 2-3% con respecto a  $P_{GAP}$  y es regulable por el crecimiento celular y la fuente de carbono. Las combinaciones de los elementos reguladores promotor-terminador de los genes estudiados contribuyeron a una actividad transcripcional diferenciada de los genes heterólogos y ésta fue dependiente de la fuente de carbono, lo que demuestra la influencia del terminador transcripcional en el casete de expresión sobre los niveles de transcritos de los genes heterólogos y endógenos. El par promotor-terminador y la fuente de carbono afectaron la ruta de traducción de genes heterólogos y/o secreción de proteínas. Además, los bajos niveles de transcripción del gen heterólogo en cultivos en glicerol están correlacionados con un aumento de la traducción y/o secreción de proteína, al menos para la proteína reportera utilizada en este trabajo. Este es el primer informe sobre la caracterización funcional del promotor y el terminador transcripcional del gen *1033* de *P. pastoris* lo cual amplía la caja de herramientas disponible para ajustar la expresión génica en *P. pastoris*.

## ABSTRACT

The *Pichia pastoris* yeast has been widely used in the production of heterologous proteins due to its multiple advantages. Promoters are key elements in an expression vector; however, a strong promoter is not always the most favorable choice for the production of heterologous proteins, especially if correct protein folding and/or post-translational processing is the limiting step. The transcriptional terminator is another regulatory element that can modify the expression levels of the heterologous gene since it influences on the transcript stability.

In this study, three regulatory elements were functionally characterized, including a promoter and two 3'UTRs sequences from two *P. pastoris* genes. The regulatory regions of the PAS\_chr3\_1033 (1033) gene, the promoter (P<sub>1033</sub>) and the transcriptional terminator (T<sub>1033</sub>), a weak housekeeping non-methanol dependent gene, were identified. Synthetic sequences were designed and that were integrated into two expression cassettes containing the combination promoter and transcription terminator of the 1033 gene (P<sub>1033</sub>-T<sub>1033</sub>) and the combination of P<sub>1033</sub> and the transcription terminator region of the *AOX1* gene (P<sub>1033</sub>-T<sub>AOX1</sub>), in both cases functionally associated to the pre-pro alpha factor sequence of *S. cerevisiae* and the *FTEII* reporter gene that codes for a bacterial phytase. Two vectors harboring the expression cassettes were integrated into the *P. pastoris* genome to build the strains KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (A) and KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (H). The functionality of the regulatory regions was determined by transcripts detection (RT-PCR), detection of the protein product by SDS-PAGE and phytase enzymatic activity. The impact of the combinations of regulatory elements on the transcription levels of the heterologous genes, 1033 and GAPDH genes in cells grown in glucose or glycerol, and on the extracellular product/biomass (Y<sub>p/x</sub>) yield was evaluated in unicopy clones.

The presence of reporter gene transcripts was evidenced by the detection of a 536 bp band in agarose gel as product of the RT-PCR. The extracellular phytase activity 24 hours after culture in glycerol was detected with levels of  $0.23 \pm 0.01$  (strain A) and  $0.11 \pm 0.01$  (strain H) U/mL, it was also detected by SDS-PAGE, in the cultures supernatants, a band with an apparent molecular mass of 39 kDa, that corresponds to the recombinant phytase molecular mass produced extracellularly, confirming the functionality of the regulatory regions combination evaluated. P<sub>1033</sub> showed 2-3% transcriptional activity relative to P<sub>GAP</sub> and is regulatable by cell growth and carbon source. The combinations of promoter-terminator regulatory elements of the studied genes contributed to a differentiated transcriptional activity of the heterologous gene, and this was dependent on the carbon source, which demonstrates the influence of the transcriptional terminator in the expression cassette on the transcripts levels of heterologous and endogenous genes. Promoter-terminator pair and carbon source affected the translation pathway of heterologous genes and/or protein secretion. In addition, low levels of transcription of the heterologous gene in glycerol cultures are correlated with increased protein translation and/or secretion, at least for the reporter protein used in this work. This is the first report on the functional characterization of the *P. pastoris* 1033 gene promoter and transcriptional terminator, which expands the toolbox available for tuning gene expression in *P. pastoris*.

# 1. INTRODUCCIÓN

La levadura metilotrófica *Pichia pastoris* (*Komagataella phaffii*) es un huésped ampliamente utilizado para producir proteínas heterólogas, debido a su eficiente sistema de secreción de proteínas que simplifica los pasos de purificación, así como a su capacidad para realizar modificaciones postraduccionales y para crecer hasta altas concentraciones de biomasa en medios de cultivo definidos de bajo costo. Además, puede producir altos niveles de proteínas heterólogas extracelulares en bioprocesos controlados. Normalmente, un casete de expresión que alberga una secuencia promotora, la secuencia codificante heteróloga y una secuencia terminadora de la transcripción se integra en el genoma de la levadura para expresar el gen heterólogo durante el crecimiento celular o en presencia de un inductor.

La transcripción del gen heterólogo es un paso crucial en la producción de proteínas recombinantes, normalmente se prefieren promotores fuertes y controlables para la producción eficiente de proteínas. Sin embargo, un promotor fuerte no siempre es la opción más favorable para la producción de proteínas heterólogas, especialmente cuando el plegamiento correcto de la proteína heteróloga y/o el procesamiento postraduccional en la vía secretora es el paso limitante para la obtención de altas cantidades del producto extracelular. Niveles altos de expresión génica heteróloga pueden sobrecargar la maquinaria postraduccional de la célula, provocando que una proporción significativa de la proteína recombinante se pliegue mal, no se procese, se localice mal y se degrade. Los promotores débiles se pueden usar para expresar genes que producen proteínas tóxicas para las células huésped y en aplicaciones de biología sintética que comúnmente requieren la coexpresión de múltiples genes.

Por otro lado, los niveles de producción de proteínas están determinados principalmente por la acción sinérgica del promotor y el terminador transcripcional. La función principal de los terminadores es la terminación del proceso de transcripción, pero la región 3' no traducida (3'UTR) contenida en la secuencia del terminador también modula la vida media del RNAm y las eficiencias transcripcional y traduccional, la exportación a través del

núcleo y la localización celular del RNAm. Por tanto, además del promotor, el terminador transcripcional es otro elemento regulador del casete de expresión que puede modificar los niveles de expresión del gen de interés.

En la extracción de datos de RNA-seq de dos cepas de *P. pastoris* KM71 cultivada en metanol, glucosa o glicerol, observamos que el gen PAS\_chr3\_1033 (*I033*) mostró niveles de transcripción bajos y casi constantes. Por tanto, el promotor del gen *I033* ( $P_{1033}$ ) debería ser débil y constitutivo. El gen *I033* (también referido como gen PP7435\_CHR3-0135 para el genoma de *K. phaffii* CBS 7435) codifica una proteína hipotética con una identidad de secuencia del 38% y 29% con las D-aminoácidos oxidasas de la levadura *Scheffersomyces stipitis* CBS 6054 (Genbank número de acceso XP\_001384459.2) y *K. phaffii* CBS 7435 (número de acceso de Genbank CCA39107.2), respectivamente.

Con el fin de caracterizar funcionalmente el arreglo de nuevos elementos reguladores que incluyen al promotor y terminador de la transcripción y determinar la influencia de la combinación de estos elementos reguladores en la expresión de un gen heterólogo en el sistema de *P. pastoris*, en este trabajo se identificaron y caracterizaron funcionalmente por primera vez el promotor ( $P_{1033}$ ) y el terminador transcripcional ( $T_{1033}$ ) del gen constitutivo *I033* con actividad transcripcional débil, no dependiente de metanol.

Para caracterizar completamente el  $P_{1033}$  junto con el  $T_{1033}$ , identificamos las secuencias  $P_{1033}$  y  $T_{1033}$ , junto con la secuencia 5'UTR, el sitio de inicio transcripcional (TSS), la secuencia 3'UTR y secuencias ricas en A-T río abajo del codón de paro para el gen *I033*. También se predijeron la secuencia de la caja tipo TATA, un elemento iniciador de la transcripción de tipo mamífero (INR), así como los posibles factores de transcripción (FT) para la secuencia de  $P_{1033}$  de *P. pastoris*. En este trabajo además se construyeron dos cepas de *P. pastoris* con dos combinaciones de elementos reguladores (pares  $P_{1033}$ - $T_{AOX1}$  y  $P_{1033}$ - $T_{1033}$ ;  $T_{AOX1}$ : terminador transcripcional del gen *AOX1*) en el casete de expresión que alberga el gen *FTEI*, que codifica para una fitasa bacteriana (Viader-Salvadó et al., 2010), como el gen reportero y asociado funcionalmente con la secuencia señal de secreción pre-pro del factor alfa de *S. cerevisiae*. Se seleccionaron clonas de cada construcción con una sola copia del casete de expresión integrado en el genoma de la levadura y se evaluó el impacto de las combinaciones de elementos reguladores en los niveles de transcripción

del gen *FTEII* heterólogo en células cultivadas en glucosa o glicerol como fuente de carbono. Además, como punto de referencia se analizaron los niveles de transcripción de los genes endógenos *1033* y *GAPDH*. También se determinaron los rendimientos de producto extracelular/biomasa ( $Y_{p/x}$ ) para correlacionarlo con los niveles de transcripción.

Los resultados de este estudio contribuyen con el estudio de nuevos elementos reguladores e indican que el  $P_{1033}$  en combinación con  $T_{AOX1}$  o  $T_{1033}$  podría usarse para producir proteínas tóxicas para el hospedero, co-expresar genes de la vía secretora y mantener un equilibrio metabólico en aplicaciones de ingeniería metabólica o biología sintética.

Este es el primer informe sobre la caracterización funcional del promotor y el terminador de la transcripción del gen *1033* de *P. pastoris* que amplía la caja de herramientas disponible para ajustar la expresión génica en esta levadura.



## 2. ANTECEDENTES

### 2.1 Producción de proteínas recombinantes

Las proteínas recombinantes se utilizan actualmente, en una amplia gama de aplicaciones en la industria y la medicina, incluyendo productos farmacéuticos, vacunas y productos agrícolas. Mediante el uso de la tecnología del DNA recombinante, las proteínas de interés se pueden producir en diferentes hospederos incluyendo bacterias, levaduras, hongos, células de mamíferos, plantas e insectos (Karbalaie 2020). El mercado de proteínas recombinantes está creciendo constantemente, se espera que los productos biofarmacéuticos alcancen un valor en el mercado de \$ 386.7 mil millones para fines de 2019 a una tasa de crecimiento anual compuesta (TCAC) del 10,6% (Highsmith 2015), mientras que, para las enzimas industriales, se prevé que alcancen los \$ 6,3 mil millones en 2021 a un TCAC del 4,7% (Chapman et al. 2018). La alta demanda y continua expansión del mercado proporciona incentivos para mejorar las plataformas de producción de proteínas, tanto para permitir la producción de nuevas proteínas como para reducir los costos de fabricación de los procesos existentes (García-Ortega et al. 2019).

La elección de un hospedero adecuado para la producción de una proteína recombinante es muy importante y probablemente uno de los pasos más críticos que deben abordarse en el comienzo del diseño del bioproceso (Matthews et al. 2017). Para seleccionar un sistema de expresión adecuado se deben tener en cuenta parámetros tales como costos, rendimiento, escala de tiempo de producción, capacidad de escalado y procesos posteriores, si la proteína requiere modificaciones postraduccionales, así como las propiedades y el uso del producto (Matthews et al. 2017).

En general, las bacterias proporcionan un excelente sistema de expresión para proteínas que no requieren síntesis en forma glicosilada o modificaciones postraduccionales y permiten procesos de producción rápidos y económicos. *Escherichia coli* ha sido el caballo de batalla para la producción de proteínas recombinantes, décadas de trabajo con este organismo han permitido contar con células bien caracterizadas y de rápido crecimiento, fáciles de cultivar y con una variedad de plásmidos que se utilizarán para ajustar los niveles de expresión de proteínas (Wiseman et al. 2020). Otros organismos

bacterianos como especies de *Bacillus*, *Pseudomonas fluorescens*, *Staphylococcus carnosus* y *Streptomyces lividans*, también son eficientemente utilizados como hospederos para producción de proteínas recombinantes (Adrio y Demain 2014; Westers y Quax 2004).

Los microorganismos eucariotas como la levadura emergen como un hospedero ideal pues combinan la capacidad eucariota de procesamiento de la proteína, como plegar, ensamblaje e introducción de modificaciones postraduccionales con importantes ventajas microbianas, como la capacidad de crecer rápidamente y altas densidades celulares en medios químicamente definidos. Además, las levaduras pueden secretar proteínas recombinantes en los medios extracelulares, facilitando enormemente la purificación (Porro et al. 2011; Porro y Mattanovich 2004). Las levaduras se implementaron por primera vez como una plataforma de síntesis de proteína recombinante en 1981 usando *Saccharomyces cerevisiae*. Desde entonces, otras levaduras se han convertido en alternativas que presentan algunas ventajas claras, incluida *Pichia pastoris*, así como otras levaduras, como *Hansenula polymorpha* o *Kluyveromyces lactis* (García-Ortega et al. 2019).

## **2.2 Sistema hospedero *Pichia pastoris*. Ventajas.**

La levadura *Pichia pastoris* reclasificada recientemente como *Komagataella phaffii* (Kurtzman 2009) es uno de los hospederos más importantes para la producción de proteínas recombinantes en la industria biotecnológica, principalmente relacionada con la producción farmacéutica (Adrio y Demain 2014). Este organismo metilotrófico se comenzó a utilizar como hospedero en 1985, sus principales ventajas se resumen en la Tabla 1. (García-Ortega et al. 2019).

Actualmente, se han expresado más de 5000 proteínas recombinantes en el sistema *P. pastoris* alcanzando niveles de expresión aproximadamente 80% de la proteína total secretada o hasta 30% de la proteína total celular (Potvin et al. 2012).

**Tabla 1.** Resumen de las principales ventajas de *P. pastoris* como sistema de expresión de proteínas recombinantes.

Ingeniería genética	Procesamiento de proteínas	Ingeniería de bioprocesos
Manipulación genética simple y estable	Capacidad eucariota para plegar, ensamblar y realizar modificaciones postraduccionales	Es considerado organismo GRAS carece de endotoxinas detectables
Numerosas herramientas disponibles para la manipulación genética. Incluido el sistema CRISPR / Cas9	Vía de procesamiento y secreción de proteínas similar a eucariotas superiores	Crecimiento no fermentativo en glucosa
Diferentes promotores versátiles y eficientes	Capacidad de secretar eficientemente las proteínas objetivo y bajos niveles de secreción de proteínas nativas	Crecimiento en medio químicamente definido hasta 120 g DCW L <sup>-1</sup>
Cepas reportadas de alto rendimiento y producción estable	Hiperglicosilación reducida y cepas construidas con glicosilación similar a humanos	Producción a gran escala bien establecida Procesamiento de purificación

(GRAS = generalmente reconocido como seguro. DCW = peso de células secas).

Aunque los sistemas de expresión basados en glucosa y glicerol también se aplican ampliamente para *P. pastoris* (Lee et al. 2017), la expresión inducida por metanol de proteínas recombinantes mediadas por promotores fuertemente inducibles tales como el promotor del gen de la alcohol oxidasa 1 (*AOX1*) puede considerarse como una de las características principales de *P. pastoris*. La preocupación por la toxicidad e inflamabilidad del metanol fomentó el desarrollo de nuevos sistemas de expresión independientes del metanol en *P. pastoris*, los enfoques recientes condujeron a la identificación de nuevos posibles promotores y estrategias de expresión para este hospedero (Ahmad et al. 2014; Vogl y Glieder 2013).

### 2.2.1 Sistema de expresión de *P. pastoris*.

El sistema de expresión de *P. pastoris* para la expresión génica recombinante está conformado por cepas modificadas de esta levadura y plásmidos recombinantes que integran en el genoma garantizando la estabilidad de la expresión. Este sistema cuenta con promotores fuertes como *pAOX1*, *pGAP* y *pFLD* y se utiliza mayormente como secuencia

señal el péptido pre-pro del factor  $\alpha$  de *S. cerevisiae*, además de marcadores de selección para seleccionar las transformantes (Akassou y Groleau 2019).

### **2.2.1.1 Cepas de *P. pastoris* para la producción de proteínas heterólogas.**

Las cepas de expresión de *Pichia pastoris* se derivan comúnmente de la cepa NRRL-Y 11430. Esta cepa parental, como mutante auxotrófico, es deficiente en uno o pocos genes que permiten la selección de transformantes (Cregg et al. 2000). Los genes *AOX1* y *AOX2* codifican para la alcohol oxidasa en la vía de oxidación de metanol, dependiendo de la eliminación de uno o dos de estos genes, hay tres tipos de cepas de *P. pastoris* que se denominan como  $Mut^+$ ,  $Mut^s$  y  $Mut^-$  (Theron et al. 2018; Cos et al. 2005). El fenotipo  $Mut^+$  incluye variantes que contienen ambas formas funcionales de genes *AOX*, el fenotipo  $Mut^s$  lo integran variantes que carecen del gen *AOX1* y el fenotipo  $Mut^-$  incluye variantes que tienen eliminado tanto *AOX1* como *AOX2*. En la cepa comercial MC100-3 (*his4 arg4 aox1* $\Delta$ : *SARG4 aox2* $\Delta$ : *Phis4*), se han eliminado los genes *AOX1* y *AOX2* del genoma; sin embargo, tiene la capacidad para inducir altos niveles de producción de proteínas bajo la regulación el promotor *AOX1*. Esta cepa demuestra el primer ejemplo de separar las fases de crecimiento y expresión de proteínas durante el mismo cultivo, por lo tanto, co-alimentarse con otro fuente de carbono, por ejemplo glucosa, además de cierta concentración de metanol que actúa solo como inductor (Looser et al. 2014).

Las cepas con una auxotrofia incluyen a las cepas JC254 (*ura3*), GS115 (*his4*) y GS190 (*arg4*); dos auxotrofías incluyen a las cepas GS200 (*arg4 his4*) y Jc227 (*ade1 arg4*); tres auxotrofías comprenden las cepas Jc300 (*ade1 arg4 his4*) y cuatro auxotrofías incluye Jc308 (*ade1 arg4 his4 ura3*). Además, hay tres tipos de mutantes deficientes en proteasa, que incluyen a las cepas SMD1163 (*his4 pep4 prb1*), SMD1165 (*his4 prb1*) y SMD1168 (*his4 pep4*), estas dos últimas son deficientes en proteasas para evitar la proteólisis de las proteínas heterólogas en los sistemas de producción de *P. pastoris* (Fickers 2014). En la Tabla 2 se muestran las cepas más comunes de la levadura metilotrónica *P. pastoris* (Baghban et al. 2019).

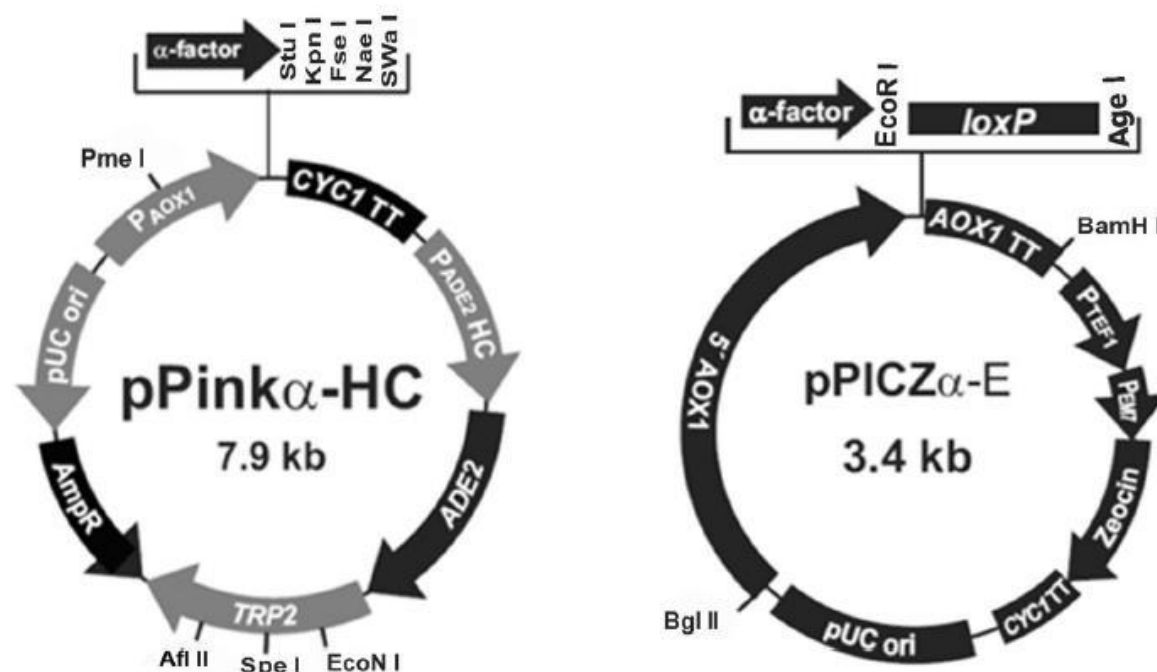
**Tabla 2.** Cepas comunes de la levadura metilotrófica *P. pastoris*.

Cepas de tipo salvaje	Y-11430 X-33	NRRLa * Daly and Hearn 2005
Cepa con diferente capacidad en el uso de metanol	Mut <sup>++</sup> ( <i>AOX1</i> <sup>+</sup> , <i>AOX2</i> <sup>+</sup> ) Mut <sup>s</sup> ( <i>AOX1</i> <sup>+</sup> , <i>AOX2</i> <sup>+</sup> ) Mut <sup>-</sup> ( <i>AOX1</i> <sup>-</sup> , <i>AOX2</i> )	Yin et al. 2007 Yin et al. 2007 Yin et al. 2007
Cepas con deficiencia en la actividad proteasa	Cepa SMD1163 ( <i>his4pep4prb1</i> ) Cepa SMD1165 ( <i>his4prb1</i> ) Cepa SMD1168 ( <i>his4pep4</i> )	Fickers 2014 Fickers 2014 Fickers 2014
Cepas con deficiencia de histidina deshidrogenasa	GS115 ( <i>his4</i> ) KM71 ( $\Delta$ aox1::SARG4 <i>his4 arg4</i> ) SMD1168 ( <i>His4, pep4</i> )	Vanz et al. 2012 Charoenrat et al. 2013 Fickers 2014

\*Laboratorios de Investigación Regionales del Norte, Peoria, IL.

### 2.2.1.2 Vectores de expresión de proteínas heterólogas en *P. pastoris*.

Los vectores estándar utilizados en genética de levadura son vectores lanzadera, que pueden replicarse en *E. coli* y mantenerse en *P. pastoris*. Los vectores de expresión de *P. pastoris* que están disponibles comercialmente tienen los siguientes elementos: un sitio múltiple de clonación (MCS); el promotor del gen de la alcohol oxidasa (*AOX1*); una secuencia de señal de secreción (SIG) [incluye *SUC2* (invertasa) y *PHO5* (fosfatasa ácida)], la proteína asesina *pGKL* y alfa-MF (factor de apareamiento  $\alpha$  de *S. cerevisiae*); un sitio de terminación de la transcripción (TT); *Amp<sup>r</sup>* (usado para la selección por ampicilina); *HIS4* (marcador para selección usando hidroxihistidinasa); y *ColB1*, un elemento de replicación para la proliferación de plásmidos en *E. coli* (Ahmad et al. 2014). La figura 1 muestra el diagrama de los vectores pPink $\alpha$ -HC y pPICZ $\alpha$ -E que contienen estos elementos generales. (Baghban et al. 2019).



**Figura 1.** Diagrama de los vectores de *Pichia pastoris* pPink $\alpha$ -HC y pPICZ $\alpha$ -E.

Los vectores más comunes usados para la expresión recombinante en *P. pastoris* se muestran en la Tabla 3 (Baghban et al. 2019).

**Tabla 3.** Vectores de expresión comerciales y marcadores de selección en *Pichia pastoris*.

Plásmido	Marcador de selección en levadura	Marcador de selección en bacteria	Referencia
DNA 2.0	Zeocina™, G418	Zeocina™, Ampicilina	Ahmad et al. 2014
Life Technologies™	Blasticidina, G418, Zeocina™, <i>HIS4</i>	Zeocina™, Ampicilina, Blasticidina	Ahmad et al. 2014
Life Technologies-PichiaPink™	<i>ADE2</i>	Ampicilina	Invitrogen, Carlsbad, CA
Life Technologies-pPICZ $\alpha$ -E Echo™	Zeocina™	Kanamicina	Invitrogen, Carlsbad, CA
BioGrammatics	Zeocina™, G418, Nurseotricina	Ampicilina	Ahmad et al. 2014
BioGrammatics GlycoSwitch®	Zeocina™, G418, Higromicina, <i>HIS4</i> , Nurseotricina	Ampicilina, Kanamicina, Zeocin™, Nurseotricina	Ahmad et al. 2014

### 2.2.1.3 Promotores utilizados en *P. pastoris*, inducibles y constitutivos.

Los promotores utilizados en levaduras para la expresión heteróloga de los genes se dividen en promotores constitutivos e inducibles, los constitutivos muestran altos niveles de expresión en casi todas las condiciones de cultivo, permitiendo una gran flexibilidad en la producción de la proteína. Por el contrario, los promotores inducibles necesitan condiciones más específicas y explícitas para la inducción, pero logran mayores niveles de producción de la proteína en comparación con los promotores constitutivos (Kamal et al. 2018).

El promotor fuerte e inducible con metanol más utilizado en *P. pastoris* para producir proteínas y enzimas heterólogas es el promotor de la enzima alcohol oxidasa 1 ( $P_{AOX1}$ ) conocido por inducir altos niveles de expresión (Vogl y Glieder 2013). Además, su homólogo débil, el  $P_{AOX2}$ , también ha sido utilizado. Otros promotores fuertes que se regulan con metanol ó metilamina reportados, son los promotores de la enzima formaldehído deshidrogenasa  $P_{FLDI}$  y de la dihidroxiacetona sintasa  $P_{DAS}$ , se reporta que éste último es incluso más fuerte que  $P_{AOX1}$ .

Sin embargo, el metanol es inflamable, peligroso para la salud y se requieren múltiples precauciones cuando se utiliza en escalas piloto e industrial; además de esto, el consumo de metanol genera grandes desventajas ya que puede causar procesos exotérmicos de elevadas temperaturas, dificultado el control del bioproceso y un incremento en la demanda de oxígeno durante la fase de lote alimentado (Prielhofer et al. 2013b; Ahmad et al. 2014).

El promotor constitutivo que se utiliza con frecuencia en la producción de proteínas heterólogas por tener casi los mismos niveles de expresión, que los que se tienen con  $P_{AOX1}$ , es el que controla la expresión basal del gliceraldehído 3-fosfato deshidrogenasa ( $P_{GAP}$ ) (Çalik et al. 2015); de hecho, su productividad se influencia controlando directamente la velocidad de crecimiento en la actividad óptima de  $P_{GAP}$  y disminuyendo los niveles disponibles de oxígeno (Rebnegger et al. 2014). Otros promotores constitutivos incluyen  $P_{TEF1}$  (factor de elongación de la traducción -1) (Ahn et al. 2007) y el promotor de enzima glicolítica 3 fosfoglicerato kinasa ( $P_{PGK1}$ ). También se ha descrito el promotor inducible con etanol isocitrato liasa ( $P_{ICL1}$ ) (Menendez et al. 2003).

Un estudio reciente identificó y caracterizó un nuevo promotor constitutivo fuerte  $P_{GCW14}$  el cual mostró alta actividad promotora bajo diferentes fuentes de carbono, incluso mayor que los promotores constitutivos más usados en *P. pastoris*  $P_{TEF1}$  y  $P_{GAP}$  (Liang et al. 2013). La producción de enzimas bajo el control del promotor  $P_{ADH3}$  ha sido comparado con los promotores  $P_{AOX1}$  y  $P_{GAP}$  de *P. pastoris*. A continuación, se muestran en la Tabla 4 los promotores de *P. pastoris* que se han identificado y caracterizado recientemente y los más usados en estos sistemas de expresión, ya sean inducibles o constitutivos para la fabricación de proteínas recombinantes (Kamal et al. 2018).

**Tabla 4.** Promotores de *P. pastoris* identificados y caracterizados recientemente y los más usados para la producción de proteínas recombinantes.

Nombre del gen	Gen producto/función	Niveles de expresión	Referencia
<b>Promotores inducibles</b>			
<i>AOX1</i>	Alcohol oxidasa 1	Fuerte (naturalmente ~ 5% del ARNm y ~ 30% de las proteínas totales)	Cereghino y Cregg 2000
<i>AOX2</i>	Alcohol oxidasa 2	~ 5–10% de $P_{AOX1}$	Cregg et al. 1989
<i>DAS</i>	Dihidroxiacetona sintasa	Fuerte (similar a $P_{AOX1}$ )	Küberl et al. 2011
<i>FLD1</i>	Formaldehído deshidrogenasa	Fuerte (similar a $P_{AOX1}$ )	Shen et al. 1998
<i>PEX8</i>	Proteína de la matriz peroxisomal	Débil	Cereghino y Cregg 2000
<i>THI1</i>	Proteína implicada en la síntesis del precursor de tiamina hidroximetilpirimidina	-	Ahn et al. 2007
<i>ENO1</i>	Enolasa	~ 20–70% de $P_{GAP}$	Stadlmayr et al. 2010
<i>ICL1</i>	Isocitrato liasa. Reprimido por glucosa, inducido en ausencia de glucosa / por adición de etanol	-	Menendez et al. 2003
<i>PHO89</i>	Putativo simulador de Na + / fosfato. Inducción en ausencia de fosfato	-	Ahn et al. 2009
<i>ADH1</i>	Alcohol deshidrogenasa. Reprimido en glucosa y metanol, inducido en glicerol y etanol	-	Cregg y Tolstorukov 2012
<i>GUT1</i>	Glicerol quinasa. Reprimido en metanol, inducido en glucosa, glicerol y etanol	-	Cregg y Tolstorukov 2012
<b>Promotores constitutivos</b>			
<i>GAP</i>	Gliceraldehído 3-fosfato deshidrogenasa	Fuerte (similar a $P_{AOX1}$ )	Waterham et al. 1997
<i>GPM1</i>	Fosfoglicerato mutasa	~ 15–40% de $P_{GAP}$	Stadlmayr et al. 2010
<i>HSP82</i>	Chaperona citoplasmática (familia Hsp90)	~ 10–40% de $P_{GAP}$	Stadlmayr et al. 2010
<i>ILV5</i>	Acetohidroxiácido isomeroeductasa	~ 15% de $P_{GAP}$	Vogl et al. 2016



<i>KAR2</i>	Chaperona residente de ER (también denominada Bip)	~ 10–70% de $P_{GAP}$	Stadlmayr et al. 2010
<i>KEX2</i>	Endopeptidasa involucrada en el procesamiento de proteínas secretadas	~ 10% de $P_{GAP}$	Vogl et al. 2016
<i>PET9</i>	Portador ADP / ATP de la membrana mitocondrial interna	~ 10–1700% de $P_{GAP}$	Stadlmayr et al. 2010
<i>PGK1</i>	Fosfoglicerato quinasa	~ 10% de $P_{GAP}$	Stadlmayr et al. 2010 Almeida et al. 2005
<i>SSA4</i>	Proteína de choque térmico	~ 10–25% de $P_{GAP}$	Stadlmayr et al. 2010
<i>TEF1</i>	Factor de elongación de traducción 1 alfa	Fuerte (similar a $P_{GAP}$ )	Stadlmayr et al. 2010 Vogl y Glieder 2013
<i>TP11</i>	Triosa fosfato isomerasa	~ 10–80% de $P_{GAP}$	Stadlmayr et al. 2010
<i>YPT1</i>	GTPasa implicada en la secreción	Débil	Sears et al. 1998 Cereghino y Cregg 2000
<i>GCW14</i>	Potencial expresión constitutiva de la proteína glicosil fosfatidil inositol (GPI) en glicerol, glucosa y metanol	-	Liang et al. 2013
<i>G1</i>	Transportador de glucosa de alta afinidad. Reprimido en glicerol, inducido por limitación de glucosa	-	Prielhofer et al. 2013
<i>G6</i>	Putativa aldehído deshidrogenasa. Reprimido en glicerol, inducido por limitación de glucosa	-	Prielhofer et al. 2013
<b>Promotores derivados de datos de la literatura</b>			
<i>GND1</i>	6-fosfogluconato deshidrogenasa (descarboxilación)	-	Almeida et al. 2005
<i>MCM1</i>	Factor de transcripción involucrado en la transcripción específica de tipo celular y la respuesta de feromonas	-	Almeida et al. 2005
<i>RPL1</i>	Proteína ribosómica de la subunidad grande (60S)	-	Almeida et al. 2005
<i>RAD2</i>	Endonucleasa de ADN monocatenario, reparación de daños en el ADN	-	Almeida et al. 2005
<i>RPS2</i>	Proteína ribosómica de la subunidad pequeña (40S)	-	Almeida et al. 2005
<i>RPS31</i>	Proteína de fusión que se escinde para producir una proteína ribosómica de la subunidad pequeña (40S) y ubiquitina	-	Almeida et al. 2005
<i>TKL1</i>	Transcetolasa	-	Ahn et al. 2007

#### 2.2.1.4 Necesidad de nuevos promotores en *Pichia pastoris*.

Se han realizado múltiples trabajos con el fin de localizar nuevos promotores que permitan la sobreexpresión del gen de interés y que puedan utilizarse para la producción de proteínas heterólogas. En un estudio de 26 promotores mediante microarreglos, se identificaron 3 de ellos con actividad constitutiva en las fuentes de carbono comparadas (glucosa, glicerol y metanol), además en medio de glucosa complejo, se identificaron algunos promotores que fueron más activos en la etapa temprana de crecimiento ( $P_{HSP82}$ ,  $P_{KAR2}$ ,  $P_{PET9}$ ,  $P_{RSP2}$ ,  $P_{RSP31}$ ,  $P_{TEF1}$  y  $P_{RSP1033}$ ) y en la etapa tardía de crecimiento ( $P_{PGK1}$ ,  $P_{ENO1}$ ,  $P_{TPI1}$  y  $P_{SSA4}$ ); en el caso de  $P_{KAR2}$  y  $P_{TEF1}$ , presentaron niveles de expresión iguales e incluso mayores que  $P_{GAP}$ . Además, con este estudio propusieron dos promotores altamente potenciales por su alta actividad transcripcional ( $P_{TEF1}$  y  $P_{THI11}$ ), donde éste último parece ser el primer promotor de *P. pastoris* regulable de manera independiente de los principales componentes nutricionales como la fuente de carbono y el nitrógeno (Stadlmayr et al. 2010).

En otro estudio reportado, se identificaron 6 nuevos promotores potenciales mediante microarreglos, que se activan en respuesta al agotamiento de la fuente de carbono, donde el promotor  $P_{G1}$ , fue el que mayor actividad presentó, y además superó la actividad promotora de  $P_{GAP}$  (por más del doble) en un cultivo de lote alimentado limitado con glucosa; el gen bajo su control se identificó como transportador de alta afinidad a glucosa ( $GTH1$ ) (Prielhofer et al. 2013).

Esto demuestra la oportunidad que existe en el sistema hospedero de *P. pastoris* para encontrar nuevos promotores fuertes y regulables.

El gen *1033* (también conocido como gen PP7435\_CHR3-0135 para el genoma de *K. phaffii* CBS 7435) codifica una proteína hipotética con una identidad de secuencia del 38% y 29% con las D-aminoácidos oxidasas de la levadura *Scheffersomyces stipitis* CBS 6054 (Genbank número de acceso XP\_001384459.2) y *K. phaffii* CBS 7435 (número de acceso de Genbank CCA39107.2), respectivamente.

### 2.2.1.5 Terminadores de la transcripción

Los terminadores o regiones 3'UTR, son el extremo 3' del RNA mensajero (ARNm) que no se traduce, contienen sitios de unión para muchas proteínas de unión a RNA (RBP) e integran una variedad de señales para la localización, degradación y traducción del RNAm (Mittal et al. 2014).

Un estudio que se llevó a cabo en levadura demostró que la estabilidad de las transcripciones no está correlacionada con la longitud de sus 3'UTR. Se discriminó sistemáticamente entre los fenotipos moleculares de la superposición de la codificación y los eventos de transcripción no codificantes de cada locus génico utilizando una nueva técnica de resolución de nucleótidos de todo el genoma para cuantificar las vidas medias de isoformas de transcripción de las regiones 3' en levadura. Los resultados revelaron diferencias generalizadas en la estabilidad entre las isoformas de cientos de genes en una sola condición, y esa variación de incluso un solo nucleótido en la región 3' no traducida (UTR) puede afectar la estabilidad del transcrito. Si bien se han informado casos anteriores de asociaciones negativas entre la longitud 3' UTR y la estabilidad de la transcripción, aquí, encontramos que las isoformas más cortas no son necesariamente más estables (Gupta et al. 2014).

Es importante tener en cuenta la secuencia 3' UTR a la hora de diseñar un vector de expresión porque se ha demostrado que influye en la estabilidad del transcrito, lo que redundará en la cantidad de la proteína recombinante que se desea producir.

Las regiones terminadoras no solo actúan en la terminación de la transcripción si no que pueden influir en los niveles de expresión de una proteína codificada río arriba del terminador (Yamanishi et al. 2013; Kathleen A. Curran et al. 2013). En general, en comparación con los estudios de promotores existen pocos estudios de terminadores para regular la expresión de proteínas de un gen de interés (Matsuyama 2019). Sin embargo, para *Saccharomyces cerevisiae*, en un estudio a gran escala, fue desarrollado y caracterizado un catálogo de 5302 terminadores (Yamanishi et al. 2013). Los terminadores resultantes en los niveles de expresión más altos y bajos fueron DIT1t y GIC1t, respectivamente, que diferían en más de 70 veces (Yamanishi et al. 2013; Ito et al. 2013). Estos resultados demuestran que al elegir una secuencia terminadora se pueden regular

los niveles de expresión de proteínas dentro de un amplio rango. Además, se demostró que en *S. cerevisiae*, los terminadores afectan la expresión de proteínas independientemente del promotor que esté río arriba y el gen reportero. Los estudios han demostrado que en *S. cerevisiae*, la elección del terminador influye en la abundancia de RNAm (Yamanishi et al. 2013; Ito et al. 2013), estabilidad del RNAm (Curran et al. 2013) e interacción postranscripcional con factores transcripcionales endógenos que actúan en trans (Ito et al. 2016).

En la levadura metilotrófica *Pichia pastoris* la aplicación de una estrategia de intercambio de terminadores podría mejorar la producción de proteínas recombinantes, sin embargo, solo existe un número limitado de terminadores evaluados en *P. pastoris*. En un estudio se utilizaron 15 terminadores endógenos de los genes de la vía de asimilación del metanol y 5 terminadores de *S. cerevisiae*. Estos 20 terminadores mostraron actividades comparables, con diferencias entre sus actividades de hasta 2 veces (Ito et al. 2020).

#### **2.2.1.6. Factores transcripcionales y herramientas para su predicción**

En eucariotas la regulación transcripcional está controlada fundamentalmente por la acción de factores de transcripción (FT) en la secuencia promotora de los genes. Los FT interactúan con elementos específicos del DNA, con otros FT y con la maquinaria de transcripción basal para regular la expresión de los genes a los que se unen, activándolos o reprimiéndolos (Pérez y López 2013). Los sitios de unión de factores de transcripción (SUFT) son elementos de DNA funcional que influyen en la actividad transcripcional (Priest et al. 2009).

Existen métodos para identificar y caracterizar sitios de unión de los FT, los cuales pueden clasificarse como métodos *in vivo* e *in vitro*. Las técnicas *in vivo* pueden revelar la unión de FT a sitios difíciles, especialmente cuando los sitios de unión están agrupados muy cerca (Rhee y Pugh 2011). Dentro de las técnicas *in vivo* están la Inmunoprecipitación de Cromatina acoplada a microarreglos (ChIP-chip) o secuenciación (ChIP-seq). Los métodos *in vivo* tienen algunos inconvenientes como la especificidad de la condición biológica para cada uno de los experimentos realizados que hace que no se capturen todos los posibles sitios de unión, además la detección de uniones indirectas o cooperativas para

un FT dado y la disponibilidad limitada de modelos de especificidad de interacciones de FT-DNA (Inukai et al. 2017).

Las técnicas de predicción de sitios de unión de FT *in vivo* e *in vitro* son costosas y toman mucho tiempo, por lo que se han desarrollado herramientas bioinformáticas que permiten la identificación de estos sitios de unión. Uno de los programas más usados es el de predicción de la afinidad de factores de transcripción (TRAP) que calcula la afinidad de los factores de transcripción por las secuencias de DNA sobre la base de un modelo biofísico. Este método ha demostrado ser útil para varias aplicaciones, incluida la determinación de los genes diana putativos de un factor dado (Thomas-Chollier et al. 2011). Esta herramienta posee un algoritmo que utiliza matrices de sitios de unión de factores de transcripción y un modelo fisicoquímico que permite una predicción con mayor seguridad de la interacción de las proteínas reguladoras con una secuencia DNA (Roeder et al. 2007). Los parámetros de TRAP fueron calibrados mediante información extraída de experimentos de CHIP-chip y microarreglos de unión a proteínas, además de utilizar las descripciones de los sitios de unión de la base de datos TRANSFAC (Wingender et al. 1996); y en comparación a métodos basados enteramente en tasas de unión determinadas experimentalmente, tiene una mayor capacidad predictiva que los métodos basados en resultados. El programa es lo suficientemente rápido para ser aplicado en colecciones completas de matrices y secuencias en cuestión de minutos (Roeder et al. 2007).

#### **2.2.1.7. Identificación de secuencias promotoras y terminadoras de la transcripción**

Se pueden identificar las secuencias de las regiones promotora y terminadora de un gen que no estén anotadas en genomas de referencia mediante análisis bioinformáticos. El mapeo de transcriptomas contra un genoma de referencia permite detectar zonas sin transcritos que coinciden con las secuencias 5' UTR y 3'UTR de un determinado gen. El análisis bioinformático para RNA-seq, se realiza utilizando la plataforma Galaxy y empieza con la filtración de secuencias con calidad aceptable para alinear fidedignamente al genoma o transcriptoma de referencia. Las secuencias alineadas deben ser cuantificadas, normalizadas y sometidas a un análisis estadístico para conocer los genes

que han sido expresados significativamente. Finalmente, se identifican los roles y funciones de los genes expresados (Peñazola y Mollinedo 2017).

A continuación, se realiza el control de calidad de las lecturas, el software más utilizado para ello es FastQ que permite obtener parámetros de calidad por nucleótido secuenciado (Chen et al. 2018). Después se realiza el mapeo genómico que consiste en alinear las secuencias (lecturas) obtenidas mediante secuenciación con el genoma (o transcriptoma) de referencia (Peñazola y Mollinedo 2017). Mapeando las secuencias obtenemos información de su localización en el genoma, la cual puede ser utilizada para identificar las secuencias de genes desconocidos o secuencias 5'UTR y/o 3' UTR. La primera consideración para elegir un software de alineamiento es el “splicing” alternativo. En organismos cuyos genomas poseen intrones es mejor utilizar alineadores como Tophat2 (Kim et al. 2013), Hisat2 (Kim et al. 2019), STAR (Dobin et al. 2013) o SOAP2 (Li et al. 2009).

La visualización de un genoma alineado nos permite comprobar los datos provenientes del análisis automático. Además de identificar genes de interés y zonas sin transcritos de forma manual. Existen diferentes tipos de software con determinadas ventajas pero uno de los más recomendados es el IGV (Integrative Genomics Viewer) el cual está disponible gratuitamente para su descarga desde <http://www.broadinstitute.org/igv>, bajo una licencia de código abierto.(Thorvaldsdóttir et al. 2013).

Una vez seleccionadas las secuencias promotora y terminadora mediante el mapeo de transcriptomas, se pueden utilizar otros programas bioinformáticos que permiten identificar motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas. El programa UTRscan (Pesole y Liuni 1999) de la plataforma ITBtools para identificar motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas. Los recursos de Internet para el análisis UTR que utiliza UTRscan son UTRdb y sitio UTR, que contienen datos de secuencias UTR de RNAm eucariotas que se actualizan frecuentemente (Pesole et al. 2002). La herramienta bioinformática YAPP permite predecir el núcleo del promotor en organismos eucariotas, incluyendo la caja TATA, sitio de inicio de la transcripción y otras secuencias características de la región

promotora, esta herramienta está disponible en línea <http://www.bioinformatics.org/yapp/cgibin/yapp.cgi>, 2018) (Mombeni et al. 2020).

#### **2.2.1.8. Cuantificación de la expresión genética mediante RT-qPCR**

La técnica de RT-qPCR se utiliza para cuantificar la expresión génica de determinados genes ya sea absoluta o relativa a genes normalizadores. Esta técnica consiste en realizar una retrotranscripción de RNA extraído de paquetes celulares de cultivos en condiciones específicas y en un momento determinado, ya sea en la fase exponencial o estacionaria, De la reacción de transcripción inversa se obtiene hebras monocatenarias (DNAc), las cuales se analizan por PCR en tiempo real mediante la interpretación de las señales fluorescentes que son emitidas durante la reacción en cadena de la polimerasa a través de una serie de ciclos, la variación de la señal fluorescente en cada ciclo es un valor medible por cada reacción y para cada muestra analizada (Schmittgen 2006).

En la cuantificación absoluta se cuantifica el amplicón de manera absoluta, es decir se determina el número de copias que se generan como resultado de la reacción en cadena de la polimerasa. Las curvas de calibración utilizadas en la cuantificación absoluta pueden basarse en concentraciones conocidas de moléculas estándar de DNA, por ejemplo, DNA plasmídico recombinante (DNArec), DNA genómico, productos de RT-PCR y oligonucleótidos grandes sintetizados comercialmente (Pfaffl y Hageleit 2001). Para una cuantificación absoluta es necesario conocer las cantidades específicas de los genes estándares porque estos datos servirán para relacionar entre la muestra y la curva estándar previamente realizada y así determinar el número de copias que se encuentran presentes en cada muestra (Pfaffl 2004).

La cuantificación relativa permite interpretar la cantidad en la que se expresa un gen específico normalizado frente a un control endógeno. Estos controles por lo general son genes de mantenimiento o house-keeping y poseen la característica de expresarse en cualquier condición metabólica del organismo, manteniendo una expresión constante (Schmittgen 2006). En la cuantificación relativa no se necesita de una curva estándar, aunque si es un requisito conocer la secuencia del gen de referencia, pero las unidades son

irrelevantes en la cuantificación porque lo único que se pretende en las investigaciones es entender los cambios fisiológicos o metabólicos a nivel de expresión génica. Esto quiere decir que se interpreta que un gen aumentó o disminuyó su expresión en relación a un gen que mantiene su expresión normal (Pfaffl et al. 2002).

#### **2.2.1.9. Interacción promotor-terminador de la transcripción**

En los sistemas de expresión utilizados para producir proteínas recombinantes se diseñan casetes de expresión en los cuales se coloca un terminador de la transcripción río abajo de la secuencia codificante para controlar principalmente la terminación de la transcripción, mientras que el promotor se coloca río arriba de la región de codificación para regular la actividad transcripcional del RNAm.

En eucariotas, se ha demostrado que el nivel de producción de proteínas de los genes heterólogos está determinado principalmente por la acción sinérgica del promotor y el terminador de la transcripción. Es por ello que, en ingeniería metabólica, la estrategia de intercambio de terminador o de intercambio de promotor se considera un buen método que produce efectos de sinergia sin competir con otras técnicas de modulación del nivel de expresión como optimización de codones (Matsuyama 2019). La función principal de los terminadores es la terminación del proceso de transcripción, pero la región 3' no traducida (3'UTR) contenida en la secuencia del terminador también modula la vida media del RNAm y las eficiencias transcripcional y traduccional, la exportación nuclear y la localización celular del ARNm (Kuersten y Goodwin 2003; Mayr 2019). Por tanto, además del promotor, el terminador transcripcional es otro elemento regulador del casete de expresión que puede modificar los niveles de expresión del gen de interés.

#### **2.2.1.10. El gen PAS\_chr3\_1033 (gen *I033*)**

En estudios previos de RNA-seq de cepas de *P. pastoris* KM71 realizados en nuestro grupo de trabajo, en los cuales se analizó la información de cinco transcriptomas de células de dos cepas KM71 recombinantes las cuales fueron crecidas en cultivos bajo diferentes



condiciones de pH, temperatura y fuente de carbono (metanol, glucosa o glicerol), se identificaron 8 genes de interés por las características en su expresión, entre ellos el gen PAS\_chr3\_1033 (gen *1033*), el cual mostró niveles de transcripción bajos y casi constantes a lo largo del cultivo y por lo tanto fue reportado como constitutivo con bajos niveles de expresión. El gen *1033* (también conocido como gen PP7435\_CHR3-0135 para el genoma de *K. phaffii* CBS 7435) codifica una proteína hipotética con una identidad de secuencia del 38% y 29% con las D-aminoácidos oxidasas de la levadura *Scheffersomyces stipitis* CBS 6054 (Genbank número de acceso XP\_001384459.2) y *K. phaffii* CBS 7435 (número de acceso de Genbank CCA39107.2), respectivamente.

En este trabajo se caracterizó funcionalmente el arreglo de elementos reguladores que incluyen al promotor y terminador de la transcripción del gen constitutivo *1033* de *Pichia pastoris* y al terminador de la transcripción del gen *AOX1* a través del desempeño de la expresión heteróloga de un gen reportero en dos fuentes de carbono.

### 3. JUSTIFICACIÓN

Las proteínas recombinantes son una alternativa de producir proteínas de interés para la industria biofarmacéutica y alimenticia, con altos niveles de producción y bajo costo de fabricación, lo cual tiene impacto social y económico.

La levadura *P. pastoris* ofrece una alternativa importante para la expresión de genes heterólogos. Los altos niveles de producción y secreción, la fácil manipulación genética y la posibilidad de desarrollar cultivos controlados a escala de laboratorio hacen de este microorganismo una opción viable para la producción de proteínas heterólogas a escala industrial.

Estudios con RNA-Seq de *P. pastoris* han permitido descubrir nuevos promotores, caracterizar las regiones reguladoras de estos promotores, e identificar factores de transcripción, lo cual permitirá contar con promotores alternos a pAOX1 que utiliza metanol para inducirse lo cual puede generar altos niveles de estrés celular, degradación de la proteína heteróloga y consecuentemente lisis celular.

Este estudio permitirá aportar conocimientos al caracterizar tres elementos reguladores que incluye un promotor y dos secuencias 3'UTRs de dos genes de *P. pastoris* y desarrollar sistemas de expresión alternos en este hospedero que se puedan utilizar en la producción de proteínas heterólogas para sustituir al metanol como fuente de carbono en los bioprocesos con *P. pastoris*.

## 4. HIPÓTESIS

La caracterización funcional del arreglo de elementos reguladores que incluyen al promotor y terminador de la transcripción del gen constitutivo *1033* de *Pichia pastoris* a través del desempeño de la expresión heteróloga de un gen reportero en diferentes fuentes de carbono permitirán determinar la influencia de la combinación del promotor y terminador de la transcripción del gen *1033* en la expresión de un gen heterólogo en este hospedero.

## 5. OBJETIVOS

### 5.1. Objetivo general

Caracterizar funcionalmente el arreglo de los elementos reguladores que incluyen al promotor y terminador de la transcripción del gen constitutivo *1033* de *Pichia pastoris* y al terminador de la transcripción del gen *AOXI* a través del desempeño de la expresión heteróloga de un gen reportero en dos fuentes de carbono diferentes.

### 5.2. Objetivos específicos

- Identificar y sintetizar la secuencia del promotor y terminador de la transcripción del gen *1033*.
- Diseñar y construir dos vectores conteniendo un casete de expresión conformado por el promotor del gen *1033* y su región terminadora de la transcripción o la región terminadora de la transcripción del gen *AOXI*, asociadas a un gen reportero (FTEII) y a una secuencia señal que permita la secreción del producto proteico.
- Construir dos cepas de *P. pastoris* conteniendo los casetes de expresión de cada uno de los vectores construidos, capaces de expresar y producir el producto proteico del gen reportero.
- Evaluar la funcionalidad de la combinación de las regiones reguladoras en la expresión génica de un gen reportero en cepas unicopias cultivadas en dos fuentes de carbono.

## 6. MATERIALES Y MÉTODOS

### 6.1. Lugar de trabajo

El trabajo experimental se realizó en el laboratorio de Biotecnología Molecular (L5) perteneciente al Instituto de Biotecnología de la Facultad de Ciencias Biológicas de la Universidad Autónoma de Nuevo León.

### 6.2. Material biológico y medios de cultivo

La cepa *Escherichia coli* TOP10 ( $F^-mcrA \Delta(mrr-hsdRMS-mcrBC) \phi80lacZ\Delta M15 \Delta lacX74 recA1 araD139 \Delta(ara-leu)7697 galU galK \lambda^-rpsL(Str^R) endA1 nupG$ ) se utilizó como hospedero de clonación y subclonación, se propagó a partir de conservas del cepario del laboratorio L5 del Instituto de Biotecnología, corroborando su viabilidad y que no tuviese el gen de resistencia a ampicilina, ni un plásmido que le confiriera esta resistencia.

*Pichia pastoris* KM71 (*his4, arg4, aox1* $\Delta::ARG4$ , Mut<sup>S</sup>) se adquirió de Thermo Fisher Scientific (Waltham, MA, EUA), fue utilizada como hospedero para la producción de fitasa FTEII bajo el control del promotor *1033* (P<sub>1033</sub>) y los terminadores de la transcripción de los genes *1033* y *AOX1* (T<sub>1033</sub> y T<sub>AOX1</sub>). Esta cepa se propagó a partir de conservas del cepario del laboratorio L5 del Instituto de Biotecnología, corroborando su viabilidad y que no contuviera una secuencia en su genoma que codifique para histidina.

El plásmido pUCIDTP<sub>1033</sub>T<sub>1033</sub> utilizado para la clonación se obtuvo de Integrated DNA Technologies, Inc. (Coralville, IA, EUA) El plásmido pGAHFTEII fue construido previamente en nuestro laboratorio (Herrera-Estala et al. 2022).

Se empleó el medio Luria Bertani (LB) para el cultivo y propagación de *E. coli*, el medio YPD, medio de Regeneración base de dextrosa (RDB) y medio mínimo amortiguado (BM) o adicionado con glicerol 1 % (p/v) (BMG) para el cultivo de *P. pastoris* se prepararon de acuerdo con el manual del Pichia Expression Kit (Thermo Fisher Scientific, Waltham, MA, EUA). Los medios BMGli y BMGlu se elaboraron a partir del medio BM adicionado con glicerol 30 mM (0,28 % [p/v]) o glucosa (0,54 % [p/v]) en lugar de glicerol al 1 % (p/v) empleado para el medio BMG. Todos los productos químicos utilizados

fueron de Sigma-Aldrich Co. (St. Louis, MO, EUA) o Productos Químicos Monterrey (Monterrey, Nuevo León, México).

### **6.3. Materiales y equipos**

La glucosidasa *Endo H*, las endonucleasas *AatII*, *XhoI*, *NotI* y *Bsu36I* y el Buffer CutSmart se adquirieron de New England Biolabs (Beverly, MA, EUA.). La endonucleasa *SalI* se adquirió de Clontech (Palo Alto, CA, EUA). La transcriptasa inversa M-MLV, la DNasa libre de RNasa RQ1, la ADN polimerasa GoTaq y el cebador oligo (dT) 15 fueron de Promega (Madison, WI, EUA), así como los reactivos Green Buffer, 1 kb DNA Ladder, dNTPs, MgCl<sub>2</sub>, agarosa, ligasa T4, buffer de ligación rápida T4, buffer de ligación T4, ribonucleasa RNase ONE. Los oligonucleótidos y sondas fueron adquiridos de Integrated DNA Technologies, Inc. (Coralville, IA, EUA); las secuencias se describen en el Apéndice B.

Los kits Wizard® Plus SV Minipreps DNA Purification System, PureYield™ Plasmid Midiprep System, Wizard® SV Gel and PCR Clean-Up System y SV Total RNA Isolation System se adquirieron de Promega (Madison, WI, EUA). Las unidades de filtración Amicon Ultra-0.5mL 3K y Amicon Ultra-4 10K se adquirieron de Merck Millipore (Millipore, MA, EUA). GelRed® Nucleic Acid Gel Stain fue adquirido de Biotium (Fremont, CA, EUA).

La acrilamida/bisacrilamida fue obtenidas de BioRad (Hercules, CA, Estados Unidos) y las celdas para espectrofotometría de BRAND GMBH (Wertheim, Alemania), RNAlater fue (Gaithersburg, MD, EUA).

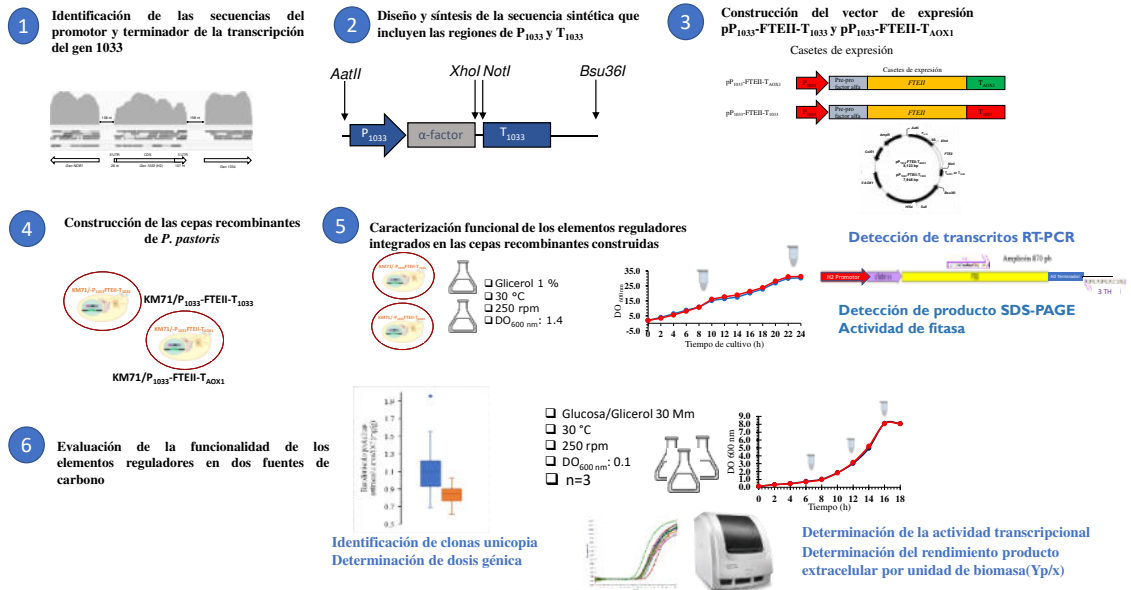
Las mezclas maestras para PCR cuantitativo en tiempo real utilizando sondas de ADN marcadas (qPCR ProbesMaster) y con tinción de ADN fluorescente verde SYBR® (qPCR SybrMaster), así como los kits SCRIPT cDNA Synthesis y gDNA Removal se obtuvieron de Jena Bioscience (Jena, Alemania).

Los equipos empleados para este trabajo se detallan en el texto.

#### 6.4. Estrategia general experimental

A partir de la información de cinco transcriptomas de dos cepas KM71 recombinantes de *P. pastoris* obtenidos en nuestro grupo de trabajo de cultivos a diferentes condiciones de pH, temperatura y fuente de carbono, se identificaron 8 genes entre ellos, el gen PAS\_chr3\_1033 (gen *I033*) el cual fue reportado como constitutivo con bajos niveles de expresión. Las regiones reguladoras de este gen no están anotadas en los genomas de *P. pastoris* reportados en bases de datos, por lo que se identificaron mediante un análisis bioinformático. Se diseñó una secuencia sintética que comprende las regiones promotora y terminadora de la transcripción del gen *I033*, la secuencia codificante de la región pre-pro del factor alfa de *Saccharomyces cerevisiae* y sitios de restricción para facilitar su clonación en un vector de expresión de *P. pastoris*. Se construyeron dos vectores de expresión, uno (pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>) que contiene un casete de expresión conformado por el promotor del gen *I033* (P<sub>1033</sub>) y la región 3'UTR del gen *AOX1* (T<sub>AOX1</sub>) y el segundo vector (pP<sub>1033</sub>-FTEII-T<sub>1033</sub>) que contiene un casete de expresión conformado por el P<sub>1033</sub> y la región 3'UTR del gen *I033* (T<sub>1033</sub>), ambos funcionalmente asociados a un gen reportero *FTEII*, secuencia que codifica para una fitasa termoestable (Viader-Salvadó et al. 2010) y a la secuencia codificante del péptido señal que permita la secreción del producto proteico. Se construyeron dos cepas de *P. pastoris* (KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>) integrando los casetes de expresión de los vectores previamente construidos y se determinó la funcionalidad de la combinación de las regiones reguladoras mediante la detección de transcritos por RT-PCR, la producción del producto proteico del gen reportero por SDS-PAGE y determinación de actividad volumétrica de fitasa. Se realizaron cultivos en BMG a pequeña escala de más de 20 clonas de ambas cepas a las cuales se les determinó el rendimiento de proteínas extracelulares/biomasa, y se seleccionaron las 3 clonas de menor rendimiento para encontrar una clona unicopia de cada cepa. Con las cepas unicopias seleccionadas, se realizaron cultivos en matraz empleando como fuente de carbono glucosa (BMGlc) o glicerol (BMGly) para evaluar la actividad transcripcional de la combinación de las regiones reguladoras determinando los niveles de expresión del gen reportero por RT-qPCR. La actividad transcripcional promovida por las regiones reguladoras en estudio se comparó con la presentada por los

genes endógenos *1033* y *GAPDH*. Además, se determinaron los niveles de producción de la proteína codificada por el gen reportero y los rendimientos proteína/biomasa (Fig. 2), con esta información se determinó la correlación entre actividad transcripcional y producción extracelular del producto proteico codificado por el gen reportero FTEII.



**Figura 2.** Estrategia general para la caracterización funcional de tres elementos reguladores de la expresión génica que incluye a un promotor y dos secuencias 3'UTRs de dos genes de *Pichia pastoris*.

## 6.5. Identificación de las secuencias del promotor y terminador de la transcripción del gen *1033*

Debido a que las regiones reguladoras del gen *1033* no están anotadas en los genomas de *P. pastoris* reportados en bases de datos, se analizaron 5 transcriptomas, determinados por RNA-seq, de la cepa KM71FTEII de *P. pastoris* recombinante cultivada a diferentes condiciones de cultivo, con el objetivo de identificar regiones donde no se mapeen transcritos río arriba y río abajo de la región codificante (CDS) del gen *1033*. Además, se realizó un análisis para identificar motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas y posibles factores transcripcionales de la región promotora del gen *1033*.

Los transcriptomas que se analizaron fueron los siguientes: transcriptoma KM71FTEII-Gli; KM71FTEII cultivada bajo condiciones limitantes de glicerol en fermentador,



transcriptoma KM71FTEII-Met (C0); KM71FTEII cultivada en batch de glicerol en fermentador, a las 0 h antes de iniciar la inducción con metanol, transcriptoma KM71FTEII-Met (CM 47-1 y CM 47-2); KM71FTEII cultivada en metanol en fermentador bajo condiciones de cultivo que favorecen altos niveles producción de la proteína heteróloga, a las 47 horas de inducción con metanol (replicados biológicos 1 y 2), transcriptoma KM71FTEII-Met (CP); KM71FTEII cultivada en metanol en fermentador bajo condiciones de cultivo que generan bajos niveles producción de la proteína heteróloga, a las 47 horas de inducción con metanol, de acuerdo con lo reportado previamente (Viader-Salvadó et al. 2013) y disponibles en NCBI Sequence Read Archive (SRA) con el número de acceso de BioProject PRJNA930494.

### **6.5.1. Mapeo de los transcritos contra el genoma de referencia de *P. pastoris***

El mapeo de cada uno de los transcriptomas de las cepas de *P. pastoris* contra la secuencia del genoma de referencia CBS7435 se realizó utilizando la plataforma bioinformática Galaxy versión 2.1.0 (<http://galaxyproject.org>). En esta plataforma se utilizó la herramienta bioinformática HISAT2 para alinear las secuencias y localizar zonas sin transcritos en las regiones río arriba y río abajo del CDS del gen *I033* empleando los parámetros estándar.

El programa Integrative Genomics Viewer (IGV) (Thorvaldsdóttir et al. 2013) se utilizó para visualizar los niveles de transcritos en las dos regiones analizadas con respecto al genoma de referencia. Como archivos de entrada se emplearon los formatos BAM y BAI generados por el programa HISAT2 de cada archivo FASTQ, así como del genoma de *P. pastoris*.

### **6.5.2. Identificación de motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas y determinación del núcleo del P<sub>1033</sub>**

El programa UTRscan de la plataforma ITBtools se utilizó para identificar motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas. Este programa busca secuencias de proteínas o nucleótidos (ADN, ARN, ARNt) para encontrar motivos UTR,

y determinar en una secuencia dada, motivos que caracterizan las regiones 3'UTR y 5'UTR. Dichos motivos están definidos en la base de datos UTRSite, una colección de patrones de secuencia funcional ubicados en las secuencias 5'UTR o 3'UTR. Cada entrada del sitio UTR se construye sobre la base de la información reportada en la literatura y revisada por científicos que trabajan experimentalmente en la caracterización funcional de elementos reguladores UTRs (Joy y Menon 2013).

La secuencia definida como del P<sub>1033</sub> también fue analizada con la herramienta bioinformática YAPP para predecir el núcleo del promotor en organismos eucariotas. Esta herramienta está disponible en línea <http://www.bioinformatics.org/yapp/cgi-bin/yapp.cgi> (Mombeni et al. 2020). Este programa busca los elementos de los promotores canónicos (cajas TATA, iniciadores, etc.) dentro de una secuencia de ADN, y combinaciones sinérgicas putativas, como un motivo Iniciador (INR) de tipo mamífero en la posición -2 en combinación con un motivo DPE en la posición +28. El enfoque se basa en el análisis estadístico de elementos del núcleo de promotores eucariota (Gershenzon y Ioshikhes 2005). La herramienta YAPP se utilizó también para predecir el nucleótido específico del sitio de inicio de la transcripción (TSS), a través de una combinación sinérgica de la caja similar a TATA y el elemento INR. El programa está diseñado para escanear regiones promotoras cortas, previamente identificadas o determinadas experimentalmente, en lugar de ser una herramienta predictiva para buscar grandes secuencias genómicas en busca de regiones promotoras.

### **6.5.3. Identificación de secuencias consenso en terminadores de levaduras**

Utilizando el programa BioEdit se buscaron en la secuencia del terminador del gen 1033, secuencias reportadas en la literatura como consenso en terminadores de levaduras, tales como el elemento de eficiencia, el elemento de posicionamiento y el sitio de poliadenilación (Poly A), al igual que regiones ricas en Ts que rodean el sitio Poly (A) (Curran et al. 2013). También se buscaron las secuencias que se refieren en la literatura como las más frecuentemente encontradas río abajo del codón de parada en los genes de levaduras (Van Helden et al. 2000).

## 6.6. Identificación de factores de transcripción que se unen a la secuencia del P<sub>1033</sub>

Los factores de transcripción que se unen a la secuencia del P<sub>1033</sub> se identificaron siguiendo el proceso que se muestra en la figura 3, utilizando la herramienta de predicción de afinidad de factores transcripcionales (TRAP) (Transcription factor Affinity Prediction), la matriz fungi y el modelo para promotores de levaduras, de la cual se seleccionaron las proteínas reguladoras con un valor de  $p < 0.05$ . Se localizaron las secuencias de los factores de transcripción de *S. cerevisiae* en la base de datos de Uniprot que se unen a los sitios de unión de factores de transcripción (SUFT) identificados y con ellas se determinaron las proteínas homólogas en *P. pastoris* mediante Best Reciprocal Blast Hit y se identificaron los sitios donde se unen a la secuencia promotora del gen *1033*, tal como se describió previamente (Hernández 2019).



**Figura 3.** Metodología empleada para identificar los factores de transcripción que se unen a la secuencia del P<sub>1033</sub>.

## 6.7. Diseño y síntesis de la secuencia sintética que incluyen las regiones de P<sub>1033</sub> y T<sub>1033</sub>

Se diseñó una secuencia sintética que incluye un sitio de restricción *AatII*, la secuencia del promotor del gen *1033*, la secuencia codificante de la región pre-pro del factor alfa de *S. cerevisiae* con codones preferenciales de *P. pastoris* (De Schutter et al. 2009) hasta el el *XhoI*, un sitio de corte de la enzimas *NotI*, 11 nucleótidos espaciadores, el terminador

de la transcripción del gen *1033*, y la secuencia desde el extremo 3' del terminador de *AOXI* hasta el sitio *Bsu36I* seguido de la secuencia de *HIS4* del vector pPIC9. Esta secuencia se sintetizó en la compañía Integrated DNA Technologies, Inc. (Coralville, IA, EUA) clonado en el vector pUCIDT-AMP para generar el vector pUCIDTP<sub>1033</sub>T<sub>1033</sub>.

Además, se diseñó un oligonucleótido 5' que hibrida en el promotor del gen *1033* empleando la herramienta primer3 v0.4.0 (<http://bioinfo.ut.ee/primer3>). Se utilizó como información de entrada la secuencia del sitio de restricción *AatII*, seguido de las secuencias del promotor del gen *1033*, la secuencia codificante de la región pre-pro del factor alfa, el sitio de restricción *XhoI* y la secuencia del gen *FTEII* que codifica para la fitasa. Para este diseño se utilizó como oligonucleótido 3' la secuencia del oligonucleótido FTE2 (Apéndice A) que hibrida en la secuencia del CDS del gen *FTEII*, diseñado previamente en nuestro laboratorio.

A continuación, se evaluó la secuencia del oligonucleótido diseñado (PP1033) para determinar su especificidad sobre el ADN blanco y la formación de estructuras secundarias. Primero se realizaron alineamientos con la herramienta Primer-BLAST (<https://www.sitio.nlm.nih.gov/tools/primer-blast/>) para evaluar la especificidad utilizando como información de entrada las secuencias de ambos oligonucleótidos y la secuencia del genoma de *P. pastoris*. Se tomaron en cuenta dos criterios: hibridación de ambos oligonucleótidos en la secuencia de interés directa y complementaria y que los oligonucleótidos no hibridaran en otra parte del genoma. Para evaluar la formación de estructuras secundarias, se realizaron análisis empleando la herramienta bioinformática OligoAnalyzer 3.1 (<http://www.idtADN.com/analyzer/applications/oligoanalyzer/>). Este programa permite analizar la formación de estructuras secundarias de tipo horquilla, homodímeros y heterodímeros, entre otras propiedades. Como información de entrada se emplearon las secuencias de ambos oligonucleótidos, se ajustaron las concentraciones de oligonucleótido a 0.5  $\mu$ M, Na 50 mM, Mg 1.5 mM y dNTPs 0.2 mM. Posteriormente se determinó el valor de  $\Delta G$ , el cual debe ser mayor a -10 kcal/mol para que no haya formación de estructuras secundarias.

Una vez evaluadas las propiedades del oligonucleótido diseñado, se mandó a sintetizar en la compañía Integrated DNA Technologies (Coralville, IA, EUA).

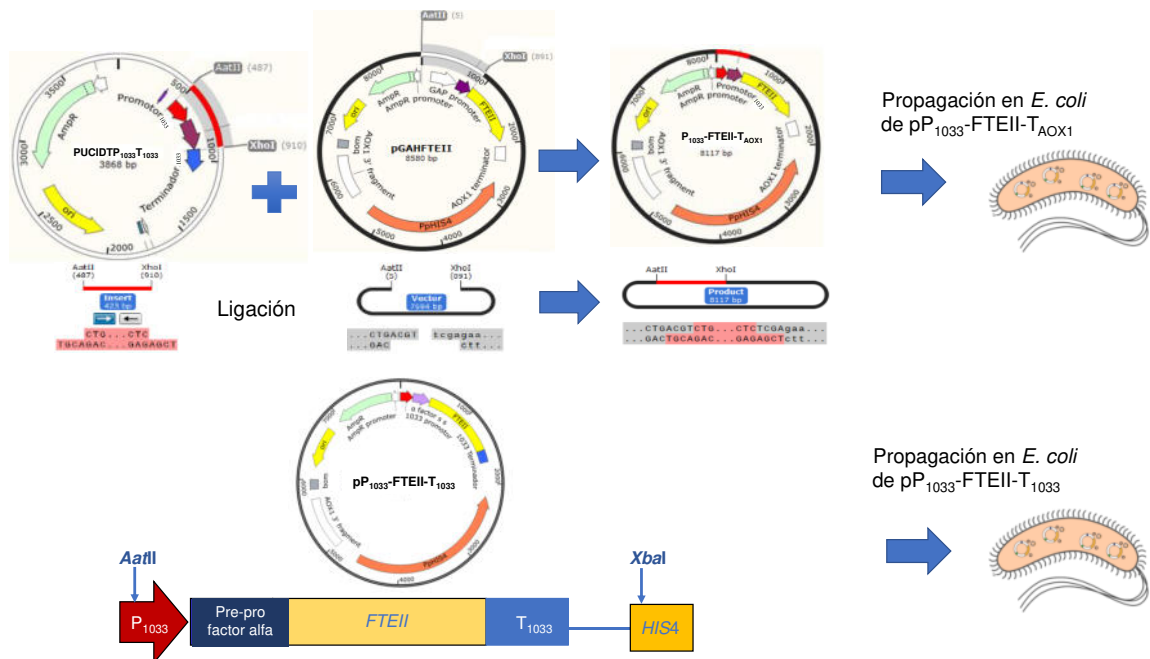
## 6.8. Construcción del vector de expresión pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>.

Se construyeron dos vectores, uno (pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>) que contiene un casete de expresión conformado por el P<sub>1033</sub> y la región 3'UTR del gen *AOX1*, el segundo vector (pP<sub>1033</sub>-FTEII-T<sub>1033</sub>) con un casete de expresión conformado por el P<sub>1033</sub> y la región 3'UTR del gen *1033*, ambos funcionalmente asociados a un gen reportero (*FTEII*), y a la secuencia nucleotídica que codifica para la secuencia señal pre-pro del factor alfa de *S. cerevisiae* que permitirá la secreción del producto proteico.

El vector pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> se construyó a partir del plásmido pUCIDTP<sub>1033</sub>T<sub>1033</sub> proporcionado por la compañía Integrated DNA Technologies y el vector pGAHFTEII obtenido en nuestro laboratorio (Herrera-Estala et al. 2022). La estrategia de clonación consistió en obtener el fragmento conteniendo *P*<sub>1033</sub> y la secuencia codificante de la región pre-pro del factor alpha de *S. cerevisiae* a partir de pUCIDTP<sub>1033</sub>T<sub>1033</sub> digerido con las enzimas con las enzimas *AatII* y *XhoI*, el cual fue posteriormente ligado al vector pGAHFTEII previamente digerido con las mismas enzimas para remover el promotor *GAP* y la secuencia codificante del péptido señal del vector receptor pGAHFTEII (Fig. 4).

Para el vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub> se diseñó un casete de expresión conteniendo los siguientes elementos: el sitio *AatII*, la secuencia promotora del gen *1033*, la secuencia codificante de la región pre-pro del factor alfa de *S. cerevisiae* el, gen reportero (*FTEII*), el terminador de la transcripción del gen *1033* y la secuencia desde el extremo 3' del terminador de *AOX1* hasta el sitio *XbaI* de la secuencia de *HIS4* del vector pPIC9. El casete de expresión con un tamaño total de 2140 pb se sintetizó en la compañía Gene Universal y fue proporcionado clonado en pPIC9 (Fig. 4).

Los plásmidos liofilizados enviados por la compañía (4 µg), se centrifugaron a 8000 rpm por 3 min a temperatura ambiente, se resuspendieron en 20 µL de agua MilliQ estéril, se homogenizaron mediante vórtex por 20 segundos. El ADN resuspendido se incubó a temperatura ambiente por 30 minutos y se centrifugó a 8000 rpm por 1 min a temperatura ambiente, finalmente se almacenó a -20°C.



**Figura 4.** Estrategia y diseño de construcción de los vectores de expresión pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>.

### 6.8.1. Transformación de *E. coli* TOP10 con los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII

La cepa TOP10 se utilizó para propagar los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII, para lo cual se comprobó que no estuviera portando algún plásmido incubándola por 16 horas en placas con medio LB (bactotripton 1%, extracto de levadura 0.5%, NaCl 1%) agar, para evaluar su viabilidad y con medio LB con ampicilina (100 µg/mL). La cepa *E. coli* TOP10 se activó inoculando 100 µL de células en 10 mL de medio LB bajo condiciones de esterilidad en una campana de flujo laminar (Labconco, EUA) y se incubó a 37 °C y 250 rpm, toda la noche, hasta alcanzar una DO<sub>600</sub> de 0.2 a 0.4. A continuación, se inocularon 100 µL en dos tubos con 5 mL de medio LB, se incubaron a 37 °C y 250 rpm hasta una DO<sub>600</sub> entre 0.2 y 0.4. El cultivo se dispensó en viales a razón de 2 mL, se centrifugó a 3000 rpm por 15 min a 4 °C, se desechó el sobrenadante y el pellet se incubó en hielo por 30 min. Estas operaciones se realizaron con el cuidado de mantener frío todo el material que está en contacto con las células durante todo el procedimiento. Se añadió CaCl<sub>2</sub> 0.1 M frío a razón de 1:10 del volumen inicial (200 µL), la mezcla se resuspendió

con golpes suaves en el fondo del vial y se incubó por 10 min en hielo. A continuación, se centrifugó a 3000 rpm por 10 min a 4 °C, se descartó el sobrenadante y se resuspendió con la mitad del volumen anterior de CaCl<sub>2</sub> frío (100 µL). Seguidamente se incubó a 4 °C toda la noche en una gradilla fría y al día siguiente se le adicionó el mismo volumen (100 µL) de glicerol frío, la mezcla se homogenizó con golpes suaves en el fondo del vial y las células calcio competentes se almacenaron a -70 °C.

La transformación de células de *E. coli* TOP10 calcio competentes se realizó con el plásmido pUCIDTP<sub>1033</sub>T<sub>1033</sub>, además se utilizó como plásmido control el pGAHFTEII (disponible en nuestro laboratorio) y como control negativo agua libre de nucleasas. Primeramente, a 100 µL de células calcio competentes se le agregaron 4.0 µL de cada plásmido y del control negativo. La suspensión celular se mezcló suavemente y se incubó en hielo por 30 minutos, luego se sometió a un choque térmico a 42°C por 45 segundos y se transfirió inmediatamente a hielo por 2 minutos. A continuación, se le agregaron 900 µL de medio SOC (Tryptona 2%, extracto de levadura 0.5%, NaCl 10 mM, KCl 2.5 mM MgCl<sub>2</sub> 10 mM, MgSO<sub>4</sub> 10 mM y glucosa 20 mM) y se incubó durante una hora a 37°C y 250 rpm. Por último, 100 µL de cada suspensión celular se sembraron en placas con medio LB con ampicilina (triptona 1%, extracto de levadura 0.5%, NaCl 1%, ampicilina 100 µg/mL y agar 15 g/L).

Para construir el vector pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> se utilizó además el vector construido previamente por el grupo de trabajo pGAHFTEII (Herrera-Estala 2019) que contiene la secuencia del gen reportero y una de las secuencias terminadoras (la región 3'UTR del gen *AOX1*) que se quieren evaluar. El vector pGAHFTEII se propagó en células de *E. coli* TOP10 y se conservó en medio LB con glicerol al 50 %.

A partir de las clonas transformadas con cada uno de los vectores que crecieron en las placas con medio LB agar con ampicilina (100 µg/mL) se picaron colonias con las cuales se inocularon 5 mL de medio LB con ampicilina a igual concentración contenidos en tubos de 15 mL y se incubaron a 37°C, 250 rpm por aproximadamente 22 horas para propagarlas y hacer conservas en medio LB con glicerol al 50 %.

### **6.8.2. Preparación y purificación de los vectores pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII**

Las cepas de *E. coli* TOP10 conteniendo los vectores pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII se activaron a partir de conservas en medio LB con glicerol al 50 %. Para ello se inocularon 100 µL de cada cepa en un tubo con 10 mL de medio LB con ampicilina (100 µg/mL), los cuales se incubaron a 37°C, 250 rpm por aproximadamente 5 horas hasta que alcanzaron una DO<sub>600</sub> de 0.2 a 0.4. El cultivo de cada cepa se dispensó en 2 viales de 2 mL y se centrifugó a 14000 g por 5 minutos. Posteriormente se desechó el sobrenadante, repitiendo este procedimiento hasta obtener un paquete celular de 5 mL de cultivo en cada vial.

La extracción y purificación de ambos plásmidos se realizó utilizando el sistema de purificación de ADN Wizard Plus SV Minipreps (Promega (Madison, WI, EUA), siguiendo el protocolo de uso de este producto. El ADN obtenido fue calentado a 65°C por 20 min para inactivar ADNsas y se analizó por electroforesis en gel de agarosa al 0.8% en buffer de tris-boratos-EDTA (Tris-HCl 90 mM, ácido bórico 89 mM, EDTA 2 mM) (TBE 1X). Los geles se prepararon pesando 800 mg de agarosa y disolviendo en 100 mL de TBE 1X estéril. La mezcla se calentó por 60 segundos en un microondas. La agarosa se colocó en un molde y se dejó solidificar a temperatura ambiente por 30 min. El gel fue montado en una cámara horizontal de electroforesis y se corrió durante 40 minutos con un voltaje de 100 V y una intensidad de corriente de 84 mA. Transcurrido este tiempo, el gel se desmontó de la cámara y se observó en un trans-iluminador 3UV de luz ultravioleta a 256 nm. El ADN plasmídico se cuantificó empleando el nanofotómetro P-300 (Implen, EUA).

### **6.8.3. Digestión de pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII con las enzimas *XhoI* y *AatII***

La digestión de los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII con las enzimas *AatII* y *XhoI* (New England Biolabs, Beverly, MA, EUA) se llevó a cabo para liberar el fragmento de ADN del *P<sub>1033</sub>* unido a la secuencia codificante del péptido señal y contenidos en pUCIDTP<sub>1033</sub>T<sub>1033</sub> y remover el promotor *GAP* y secuencia codificante del péptido señal del vector receptor pGAHFTEII. Los fragmentos liberados se utilizaron para construir el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> mediante una reacción de ligación y transformar células



electrocompetentes de *P. pastoris* KM71. Para la digestión, se emplearon 3 µg totales de ADN de cada plásmido y los reactivos empleados y mezcla de reacción se muestran en la Tabla 5.

**Tabla 5.** Configuración de la mezcla de reacción de los cortes enzimático de ambos plásmidos.

<b>pUCIDTP<sub>1033</sub>T<sub>1033</sub> (128 ng/µL)</b>		<b>pGAHFTEII (141 ng/µL)</b>	
<b>Reactivo</b>	<b>Volumen (µL)</b>	<b>Reactivo</b>	<b>Volumen (µL)</b>
Agua libre de nucleasas	58.57	Agua libre de nucleasas	58.73
Buffer CutSmart (10X)	10	Buffer CutSmart (10X)	10
ADN (3 µg)	23.43	ADN (3 µg)	21.27
Enzima <i>Xho</i> I (100 U/µL)	6	Enzima <i>Xho</i> I (100 U/µL)	6
Enzima <i>Aat</i> II (20 U/µL)	2	Enzima <i>Aat</i> II (20 U/µL)	2
Total	100	Total	100

Las reacciones se incubaron por 16 horas a 37 °C, pasado este tiempo se inactivaron las enzimas de restricción a 65 °C por 20 minutos. Se realizó una electroforesis a escala analítica en un gel de agarosa al 0.8% con 2 µL del producto resultante de cada digestión. El resto de la mezcla de cada reacción se sometió a una electroforesis preparativa para poder obtener los fragmentos de ADN deseados.

### **6.8.3.1. Purificación de pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII digeridos con *Xho*I y *Aat*II**

La purificación de los fragmentos a clonar se realizó a partir de geles de electroforesis a escala preparativa de las mezclas de reacción de los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII digeridos con las enzimas *Aat*II y *Xho*I. Previamente se sumergió la cámara de electroforesis, el peine, el molde y los adaptadores en SDS al 0.1 % y se dejaron reposar toda la noche, pasado este tiempo la cámara y accesorios se enjuagaron 3 veces con agua bidestilada estéril. Posteriormente se llenó la cámara hasta el límite máximo con solución tampón TBE 1X libre de nucleasas (estéril) y se realizó una electroforesis en agarosa 0.8 % en TBE 1X libre de nucleasas (estéril). Una vez finalizada la electroforesis, el gel se colocó sobre la base del transiluminador (previamente higienizada con SDS al 0.1%) con luz UV de 365 nm y se utilizó una navaja estéril para cortar los fragmentos de interés. Cada uno de los fragmentos obtenidos se colocaron en un tubo de microcentrífuga nuevo estéril previamente pesado y por diferencia de pesos se obtuvo el peso total de la agarosa

con ADN contenida en el tubo. Los fragmentos del gel de agarosa fueron purificados con el kit Wizard® SV Gel and PCR Clean-Up System (Promega, EUA) siguiendo las instrucciones del fabricante.

Las concentraciones de ADN de los fragmentos purificados se determinaron empleando el nanofotómetro P-300 de Implen y por electroforesis en gel de agarosa al 0.8% utilizando el marcador 1kb DNA Ladder (Promega, EUA). El ADN purificado del vector receptor y del inserto se conservó a 4°C.

#### 6.8.4. Construcción y propagación del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>

El plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> se obtuvo de una reacción de ligación con los fragmentos purificados de acuerdo con el protocolo de la sección previa. Primeramente, se utilizó el programa NEBcalculator (<https://nebiocalculator.neb.com/#!/ligation>) que calcula la relación molar inserto-vector que debe haber en la reacción de ligación, de acuerdo con los tamaños de inserto y vector y la cantidad de vector que se utilice. La relación molar 3:1 inserto-vector fue empleada para la reacción de ligación por lo que se utilizaron 8.25 ng de inserto por 50 ng de vector (Fig. 5).

Ligation Calculator

This tool will calculate the mass of insert required at several molar insert:vector ratios in the range needed for typical ligation reactions.

Insert DNA length  
423 bp

Vector DNA length  
7694 bp

Vector DNA mass  
50 ng

Required insert DNA mass

2.749 ng (1:1)
5.498 ng (2:1)
8.247 ng (3:1) ←
13.74 ng (5:1)
19.24 ng (7:1)

**Figura 5.** Resultado del programa NEBcalculator. Relación molar inserto-vector para la reacción de ligación de acuerdo con los tamaños de inserto y vector y 50 ng de vector.

Se prepararon 3 reacciones utilizando agua libre de nucleasas y buffer de ligación 10X de Promega (Madison, WI, EUA). La primera y la segunda reacción fueron controles negativos de la ligación: a la primera se le añadió el vector pGAHFTEII linearizado y la enzima DNA T4 ligasa, a la segunda se le añadió el vector pGAHFTEII linearizado e inserto sin la enzima ligasa y la tercera fue la reacción de ligación que contiene el inserto y vector pGAHFTEII linearizado en una relación molar 3:1 y la enzima DNA T4 ligasa (Tabla 6).

**Tabla 6.** Reacciones de ligación y controles negativos que se realizaron con los fragmentos purificados de inserto y vector.

**Reacción 1: (Vector + T4 ligasa)**

<b>Reactivo</b>	<b>Volumen (µL)</b>
Agua	3
Buffer ligación (10X)	1
Vector pGAHFTEII (50 ng)	5
T4 ADN ligasa	1
Total	10

**Reacción 2: (Vector + Inserto)**

<b>Reactivo</b>	<b>Volumen (µL)</b>
Agua	1
Buffer ligación (10X)	1
Vector pGAHFTEII (50 ng)	5
Inserto (8.25 ng)	3
Total	10

**Reacción 3: (Vector + Inserto + T4 ligasa 1:3)**

<b>Reactivo</b>	<b>Volumen (µL)</b>
Buffer ligación (10X)	1
Vector pGAHFTEII (50 ng)	5
Inserto (8.25 ng)	3
T4 ADN ligasa	1
Total	10

Las tres mezclas de reacción de 10  $\mu\text{L}$  se incubaron a 4  $^{\circ}\text{C}$  toda la noche y posteriormente se analizaron por electroforesis. Para ello se tomó 1  $\mu\text{L}$  de cada mezcla de reacción y el resto se empleó para la transformación de *E. coli* TOP10 calcio competentes.

Para realizar la transformación de células de *E. coli* TOP10 calcio competentes se siguió el procedimiento descrito en el acápite 6.8.1, se utilizó como plásmido control el pGAHFTEII y como control negativo agua libre de nucleasas. A continuación, 100  $\mu\text{L}$  de cada suspensión celular se sembraron en placas de LB agar con ampicilina (100  $\mu\text{g}/\text{mL}$ ). Para la selección de la cepa recombinante de *E. coli* portadora del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>, las placas se incubaron en una incubadora estática a 37 $^{\circ}\text{C}$  de 16-20 horas.

Las cepas transformadas con el producto de la reacción de ligación que crecieron en placas de LB con ampicilina se inocularon en medio LB líquido con ampicilina (100  $\mu\text{g}/\text{mL}$ ) y se incubaron a 37 $^{\circ}\text{C}$ , 250 rpm durante 14-18 horas para propagarlas y hacer conservas en medio LB con glicerol al 50 %.

Cuatro clonas de *E. coli* TOP10 que contenían el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> se activaron a partir de conservas en medio LB con glicerol al 50 %. Posteriormente se cultivaron en medio LB con ampicilina (100  $\mu\text{g}/\text{mL}$ ) y se obtuvieron paquetes celulares a los cuales se les extrajo el ADN plasmídico necesario para su caracterización por el método de lisis alcalina. La preparación obtenida se analizó por electroforesis en gel de agarosa al 0.8% y se almacenó a 4 $^{\circ}\text{C}$  hasta su posterior uso.

Para caracterizar el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>, el ADN plasmídico extraído de una de las clonas fue sometido a una reacción de digestión empleando las enzimas de restricción *AatII* y *XhoI* de New England Biolabs (Beverly, MA, EUA). Se preparó una mezcla de reacción tal como se presenta en la Tabla 7. La reacción de digestión enzimática se incubó a 37 $^{\circ}\text{C}$  por 16 horas, posteriormente se inactivaron las enzimas y se analizaron 3  $\mu\text{L}$  de la reacción en un gel de agarosa al 0.8%. La digestión con estas enzimas genera dos bandas, una de 423 pb y otra de 7694 pb para el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>.

**Tabla 7.** Mezcla de reacción de la digestión de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con las enzimas *XhoI* y *AatII*.

Reactivo	Volumen (μL)
Agua estéril libre de nucleasas	2.75
Buffer CutSmart NEB(10X)	1.0
Enzima <i>XhoI</i> (100 U/μL)	1.0
Enzima <i>AatII</i> (20 U/μL)	0.25
ADN (pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub> ) (92.5 ng/ μL)	5.0
Volumen total	10.0

El plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> se caracterizó también mediante una reacción de PCR con los oligonucleótidos PP1033 y FTE2 (Apéndice A). Para la PCR se preparó un coctel de 4 reacciones con los reactivos descritos en la Tabla 8. El coctel se dispensó en 4 viales de 0.2 mL en alícuotas de 24 μL, en tres de ellos se agregó 1 μL del ADN extraído de tres clonas transformadas y en el otro se añadió 1 μL de agua libre de nucleasas como control negativo (tabla 8). Posteriormente se ajustaron las condiciones en el termociclador (Tabla 9) y el producto amplificado se detectó en un gel de agarosa al 0.8%.

**Tabla 8.** Mezcla de reacción para la caracterización por PCR del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>.

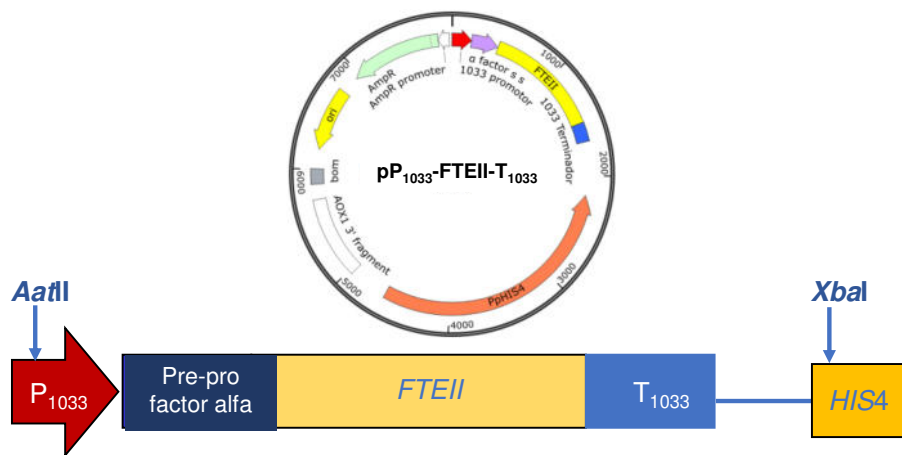
Reactivo	Concentración	Volumen (1X)	Volumen (4X)
		(μL)	(μL)
Agua estéril libre de nucleasas	-	13.3	53.2
Green Buffer	5X	5.0	20.0
dNTP's	10 mM	0.5	2.0
Oligonucleótido PP1033	5 μM	2.5	10.0
Oligonucleótido FTE2	5 μM	2.5	10.0
<i>GoTaq</i> DNA polimerasa	5 U/μL	0.2	0.8
DNA	100 ng/μL	1.0	-
Volumen final	-	25.0	96.0

**Tabla 9.** Condiciones del termociclador para el ensayo de PCR.

Etapa	Temperatura	Tiempo	Ciclos
	(°C)	(min)	
Desnaturalización inicial	95	2	1
Desnaturalización	95	1	
Hibridación	53	1	30
Extensión	72	2	
Extensión final	72	5	1
Conservación	4	∞	

### 6.8.5. Construcción del vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub>

Para la construcción del vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub> se diseñó un casete de expresión conteniendo los siguientes elementos: el sitio *Aat*II, la secuencia promotora del gen *1033*, la secuencia codificante de la región pre-pro del factor alfa de *S. cerevisiae*, el gen reportero (*FTEII*), el terminador el terminador de la transcripción del gen *1033* y la secuencia desde el extremo 3' del terminador de *AOXI* hasta el sitio *Xba*I de la secuencia de *HIS4* del vector pPIC9. El casete de expresión con un tamaño total de 2140 pb se sintetizó en la compañía Gene Universal y fue proporcionado clonado en pPIC9 (Fig. 6).



**Figura 6.** Plásmido de pP<sub>1033</sub>-FTEII-T<sub>1033</sub> conteniendo el casete de expresión que contiene los elementos reguladores y los sitios de corte.

Para preparar el plásmido liofilizado enviado por la compañía para su uso en la transformación de las células calcio competentes de *E. coli* TOP10, se realizó el siguiente procedimiento de acuerdo con el protocolo de la sección 6.8: Se centrifugó el plásmido liofilizado contenido en un tubo de 2.0 mL por 1 min a máxima velocidad y se resuspendió en 50  $\mu$ L de agua libre de nucleasas. Antes de la transformación se determinó su concentración por espectrofotometría y se realizó una electroforesis en geles de agarosa al 0.8%.

Después de verificar que el plásmido disuelto no estuviera degradado, se transformaron células de *E. coli* TOP10 de la siguiente manera: A 50  $\mu\text{L}$  de células de *E. coli* TOP10 competentes se le agregaron 2.0  $\mu\text{L}$  de plásmido.

La suspensión celular se mezcló suavemente, y se incubó en hielo por 30 minutos, luego se sometió a un choque térmico a 42°C por 45 segundos y se transfirió inmediatamente a hielo por 2 minutos. A continuación, se le agregaron 250  $\mu\text{L}$  de medio SOC y se incubó durante una hora a 37°C a 250 rpm. Por último, 100  $\mu\text{L}$  de la suspensión celular se sembraron en placas de LB con ampicilina (100  $\mu\text{g}/\text{mL}$ ). Para la selección de cepas recombinantes de *Escherichia coli* portadoras del plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub>, las placas se incubaron en una incubadora estática a 37°C de 16-24 horas.

De las cepas transformantes que crecieron en placas de LB con ampicilina, se seleccionaron tres colonias y se inocularon en medio LB líquido con ampicilina (100  $\mu\text{g}/\text{mL}$ ), se incubaron a 37°C, 250 rpm durante 14-18 horas y se extrajo el plásmido con el kit Wizard Plus SV Miniprep DNA Purification System (Promega) siguiendo las instrucciones del fabricante y de acuerdo con el protocolo descrito en la sección 6.5.2. Finalmente, el DNA fue analizado por electroforesis en geles de agarosa al 0.8% y cuantificado por espectrofotometría.

Para la caracterización del plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub> por digestión enzimática, se utilizó solo una de las muestras de DNA obtenidas. Se realizaron mezclas de reacción para 3 muestras (una para el control negativo y dos para la muestra seleccionada) empleando las enzimas de restricción *Xho*I y *Bsu*36I. Se realizó una mezcla de reacción para un coctel de 3 reacciones de acuerdo con la Tabla 10.

Se realizó un coctel de 54  $\mu\text{L}$  para 3 reacciones de 20  $\mu\text{L}$  cada una (Tabla 10 Sección 1), repartiendo 18  $\mu\text{L}$  del coctel para cada una de las tres reacciones (Tabla 10 Sección 2), cuyas diferencias de 2  $\mu\text{L}$  estuvo en el volumen de las enzimas añadidas a la reacción o agua libre de nucleasas en el caso del control negativo, reacción 1 (R1), a la reacción 2 (R2) se añadió 1  $\mu\text{L}$  de *Xho*I y 1  $\mu\text{L}$  de agua libre de nucleasas y a la reacción 3 (R3) se añadió 1  $\mu\text{L}$  de *Xho*I y 1  $\mu\text{L}$  de *Bsu*36I realizando las mezclas de reacción como se muestra en la Tabla 10B. Las reacciones se incubaron a 37°C por 1 h. Después del periodo de incubación, se analizaron 3  $\mu\text{L}$  cada reacción en un gel de agarosa al 0.8%.

**Tabla 10A.** Reacciones de digestión enzimática (Sección 1) y mezclas de reacción (Sección 2) del plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub> digerido con *Xho*I y/o *Bsu*36I.

<b>Sección 1</b>			
<b>Reactivo</b>	<b>Reacción 1</b>	<b>Reacción 2</b>	<b>Reacción 3</b>
	<b>Control negativo</b>	<b>Volumen (μL)</b>	<b>Volumen (μL)</b>
	<b>Volumen (μL)</b>		
Agua estéril libre de nucleasas	8.0	7.0	6.0
Buffer CutSmart (10X)	2.0	2.0	2.0
Enzima <i>Xho</i> I (20 U/μL)	-	1.0	1.0
Enzima <i>Bsu</i> 36I (10 U/μL)	-	-	1.0
ADN (pP <sub>1033</sub> -FTEII-T <sub>1033</sub> ) [96.0 ng/ μL]	10.0	10.0	10.0
Volumen total	20.0	20.0	20.0
<b>Sección 2</b>			
<b>Reactivo</b>	<b>1 reacción</b>	<b>3 reacciones</b>	
	<b>μL</b>	<b>μL</b>	
Agua estéril libre de nucleasas	6	18	
Buffer CutSmart (10X)	2	6	
Enzima <i>Xho</i> I (20 U/μL)	--		
Enzima <i>Bsu</i> 36I (10 U/μL)	--		
ADN (pP <sub>1033</sub> -FTEII-T <sub>1033</sub> ) (96.0 ng/ μL)	10	30	
Volumen total	18	54	

## 6.9. Construcción de las cepas recombinantes de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>

### 6.9.1. Extracción y purificación de los plásmidos pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>

A partir de dos clonas de la cepa de *E. coli* TOP10, transformadas con los plásmidos pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>, caracterizadas previamente por PCR y digestión enzimática y conservadas a -20 °C se tomaron 100 μL de células para inocular 10 mL de medio LB con ampicilina (100 μg/mL), y se incubó a 37°C y 250 rpm por aproximadamente 21 horas hasta que alcanzó DO<sub>600</sub> de 0.2 a 0.6. En un matraz de 500 mL con 50 mL de medio LB con ampicilina (100 μg/mL) se inocularon 10 mL del cultivo



reactivado, posteriormente se incubó a 37°C y 250 rpm por aproximadamente 5 horas hasta que se alcanzó una DO<sub>600</sub> de 0.2 a 0.6. Se dispensaron 30 mL de cultivo en dos tubos de centrífuga y se centrifugaron a 5000 g por 10 minutos. El sobrenadante se desechó y los paquetes celulares se almacenaron a 4 °C hasta el día siguiente.

La extracción y purificación de cada plásmido (pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>) se realizó utilizando el sistema de purificación de DNA PureYield™ Plasmid Midiprep System (Promega, EUA), siguiendo el protocolo de uso de este producto.

Las preparaciones de ambos plásmidos se analizaron en un nanoespectrofotómetro y por electroforesis en gel de agarosa al 0.8 % para determinar su calidad y concentración.

### 6.9.2. Digestión de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con la enzima *SalI*

Volúmenes de 550 µL (pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>) y 590 µL (pP<sub>1033</sub>-FTEII-T<sub>1033</sub>) de las preparaciones de los plásmidos purificados, se dispensaron en tubos de 0.6 mL y se concentraron en el concentrador CentriVap (LABCONCO, EUA) hasta un volumen de 10 µL.

Las alícuotas de los plásmidos pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con aproximadamente 40 µg y 12 µg de DNA plasmídico respectivamente se sometieron a la digestión con la enzima *SalI* (50 U/µL) (Takara, EUA número de catálogo 1080AH) para linealizar cada plásmido. La mezcla de reacción de la digestión con *SalI* se muestra en la Tabla 11. Las reacciones se incubaron a 37 °C por 16 horas.

**Tabla 11.** Reacción de digestión de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. con la enzima *SalI*.

Reactivo	pP <sub>1033</sub> -FTEII-T <sub>AOX1</sub>	pP <sub>1033</sub> -FTEII-T <sub>1033</sub>
	(4 µg/µL)	(1.2 µg/µL)
	Volumen (µL)	Volumen (µL)
DNA plasmídico (4 µg/µL)	10	10
Buffer H	5	4
Enzima <i>SalI</i> (50 U/µL)	3	2
Agua libre de nucleasas	32	34
Total	50	50

La enzima se inactivó a 65 °C por 20 minutos y cada producto resultante de la digestión fue analizado en un gel de agarosa al 0.8%. Los reactivos contaminantes de la reacción con la enzima *SaII* se eliminaron utilizando una unidad de ultrafiltración de Amicon® Ultra 3K – (3,000 NMWL límite nominal de peso molecular) 0.5 mL (Millipore, EUA). El volumen de la reacción de la digestión (50 µL) se colocó en una unidad de ultrafiltración insertada dentro de un tubo colector de 2.0 mL y posteriormente se añadieron 400 µL de agua estéril libre de nucleasas. Se centrifugó a 14000 g por 10 min a temperatura ambiente y al finalizar se descartó el sobrenadante. Se realizaron 3 lavados repitiendo el paso anterior y finalmente cuando quedaba un volumen aproximado de 100 µL, se recuperó la muestra colocando una unidad de filtración de manera invertida dentro de un tubo nuevo de 2.0 mL, se centrifugó a 1000 g (3000 rpm) por 2 min a temperatura ambiente y se calentó a 65°C por 20 min.

El volumen colectado fue de 100 µL de la preparación de cada plásmido se pasó a un vial de 0.6 mL estéril para concentrar la preparación de ADN en el concentrador CentriVap (LABCONCO, EUA) a un volumen final de 10 µL. Se utilizó el programa 3 del equipo que corresponde a una temperatura de 42°C a 1000 rpm y se mantuvo activo hasta que el volumen llegó a la marca de 10 µL que se hizo previamente en el vial. Se tomaron 0.5 µL de cada preparación, se diluyeron 1:10 con agua MilliQ estéril, y se analizaron en un gel de agarosa al 0.8%.

### **6.9.3. Transformación de *P. pastoris* con pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>**

Para la preparación de células electrocompetentes que se utilizaron en la transformación por electroporación, se activó la cepa de *Pichia pastoris* KM71 (disponible en nuestro laboratorio), sin transformar, lo cual se comprobó incubando por 72 horas en una placa con medio YPD (extracto de levadura 1%, peptona 2%, dextrosa 2%) para comprobar su viabilidad y en una placa con medio RBD His<sup>-</sup> (sorbitol 1 M, dextrosa 2%, YNB 1.34%, biotina 4 x 10<sup>-5</sup>%, aminoácidos 0.005%, agar 2%). Primeramente, se inocularon 100 µL de células conservadas a -20 °C en 10 mL de medio YPD y se incubó a 30 °C y 250 rpm por 20 horas hasta alcanzar una DO<sub>600</sub> entre 6 y 8. Se inocularon 10 mL de este cultivo en un matraz de 500 mL con 50 mL de medio YPD y se incubó a 30 °C y 250 rpm por 20

horas hasta alcanzar una  $DO_{600}$  entre 6 y 8. Este cultivo se dispensó en 2 tubos de centrífuga y se centrifugó a 4000 rpm, 4 °C por 10 minutos. El sobrenadante se desechó y los paquetes celulares se almacenaron a 4 °C hasta el día siguiente.

Cada paquete celular se resuspendió en 5 mL de medio YPD-HEPES, se depositaron en un matraz de 250 mL, y se añadieron 250  $\mu$ L de DTT (1 M). La mezcla de células y medio YPD-HEPES-DTT se incubó a 30 °C y 250 rpm por 15 minutos. Posteriormente se aforó a 40 mL con agua destilada fría estéril, se centrifugó a 2000 g y 4°C por 10 minutos, se desechó el sobrenadante y se resuspendió en 25 mL de agua fría y estéril. Seguidamente se centrifugó a 2000 g y 4°C por 10 minutos, se desechó el sobrenadante y se resuspendió en 2 mL de sorbitol 1 M frío y estéril. Finalmente se centrifugó a 5000 g y 4°C por 10 minutos, se desechó el sobrenadante, se resuspendió en 100  $\mu$ L de sorbitol 1 M frío y estéril y se dispensó en alícuotas de 40  $\mu$ L en viales de 2.0 mL, los cuales se conservaron 16 horas a 4 °C para transformar las células frescas sin congelar.

Para la transformación de las células electrocompetentes preparadas previamente se tomaron dos tubos con células (40  $\mu$ L) de *P. pastoris* KM71 preparadas y se mezclaron con 5  $\mu$ L (8  $\mu$ g) del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> o del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> respectivamente, ambos linealizados y desalados. El contenido de cada tubo se vertió en una cubeta de electroporación de 2 mm previamente enfriada en hielo durante 2 minutos. La electroporación se realizó en el sistema Gene Pulser® II. Electroporation System (BioRad, EUA). Primero, se conectó el tren de carga de la celda en el “Controller pulser” y a su vez, el “Controller pulser” fue conectado en el Gene Pulser II, y finalmente fue conectado a la fuente de electricidad. El equipo Gene Pulser II se encendió y se ajustaron los parámetros de resistencia (400  $\Omega$ ), capacitancia (25  $\mu$ F) y voltaje (1.5 kV). La celda con la mezcla de células y ADN se introdujo en el tren del electroporador y se desplazó hacia los electrodos del equipo. Una vez colocado en su posición, se realizó la descarga eléctrica a las células y se colocó la cubeta en hielo inmediatamente después de la descarga de voltaje. Se agregó 1 mL de sorbitol 1M frío y estéril y se homogenizó la celda por inversión 5 veces.

El contenido de cada celda se sembró por extensión en placa con medio YPD-Agar y en 5 placas con medio RBD His<sup>-</sup> aplicando volúmenes de 50, 100, 150, 200 y 500  $\mu$ L en cada placa. Todas las placas se incubaron a 30°C por 48-120 horas hasta la aparición

de colonias transformadas. Se seleccionaron inicialmente 20 colonias al azar, se inocularon en medio líquido YPD y se incubaron a 30°C en agitación a 250 rpm por 24 horas.

#### **6.9.4. Caracterización genotípica de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>**

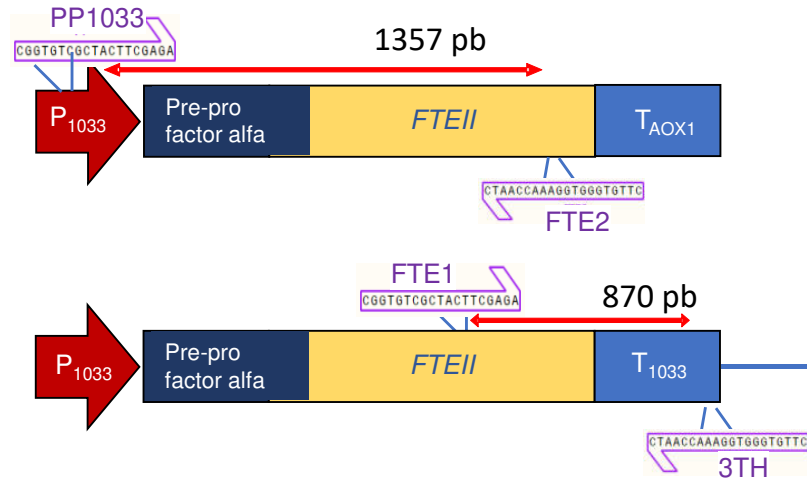
Para la caracterización genotípica se realizaron PCR de DNA genómico de las colonias recombinantes seleccionadas con parejas de oligonucleótidos dirigidos a los casetes de expresión integrados. De las colonias transformadas con ambos plásmidos se seleccionaron 4 al azar (una de cada placa), se inocularon en 5 mL de medio líquido YPD y se incubaron a 30°C y 250 rpm por 24 horas hasta DO<sub>600</sub> de 6 a 8.

Se tomaron 4 mL de cada cultivo para obtener los paquetes celulares realizando dos centrifugaciones sucesivas de 2 mL a 14000 rpm por 5 min y descartando los sobrenadantes. A partir de los paquetes celulares se obtuvieron preparaciones de DNA genómico, para ello se añadieron 200 µL de solución amortiguadora de lisis TSNT y se agitaron por inversión 10 veces. Seguidamente se agregaron 500 µL de fenol saturado con Tris-HCl y se agitaron nuevamente por inversión 5 veces. Posteriormente, se añadieron 100 µL de Sevag (cloroformo: alcohol isoamílico, 24:1) a cada tubo y se mezclaron en un vortex por 1 min. Después, se adicionaron 200 µL de TE 1X, los tubos se agitaron por inversión 5 veces y se centrifugaron a 14000 rpm durante 15 min a temperatura ambiente. La fase acuosa de los sobrenadantes obtenidos se transfirió a tubos nuevos de 2.0 mL, se añadió 1 mL de etanol al 100% frío a cada uno y se mezclaron por inversión 10 veces. Los tubos se centrifugaron a 14000 rpm por 10 min a 0°C y el sobrenadante resultante de cada tubo fue descartado. Las pastillas que quedaron en los tubos se lavaron con 1 mL de etanol 70% y se centrifugaron a 14000 rpm por 10 min a 0°C. A continuación, se descartó el sobrenadante de los tubos y se colocaron sobre papel absorbente para eliminar el exceso de etanol. Finalmente, el DNA se resuspendió en 100 µL de TE 1X, se analizó en geles de agarosa al 0.8%, y las preparaciones de DNA se calentaron a 65°C por 20 min para posteriormente almacenarlas a 4°C.

El ensayo de PCR se realizó con los oligonucleótidos específicos PP1033 y FTE2 (Apéndice A) para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y los oligonucleótidos FTE1 y 3TH (Apéndice A) para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>, utilizando una mezcla de reactivos para una reacción de 25  $\mu$ L o para un coctel de 6 reacciones como se describe en la Tabla 12 y considerando 1  $\mu$ L de una dilución 1:10 de DNA genómico extraído, de la misma manera se hicieron dos reacciones más, agregando agua libre de nucleasas en lugar de DNA como control negativo y DNA plasmídico como control positivo. Como control positivo de la reacción de PCR se utilizó el DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> para la caracterización de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> o el plásmido pGAHFTEII para el caso de la genotipificación de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>, como control negativo se utilizó agua estéril libre de nucleasas. Las condiciones del programa del termociclador fueron las mismas que se refieren en la tabla 9. Del producto de cada reacción se tomaron 3  $\mu$ L para ser analizados en un gel de agarosa al 0.8%. En la figura 7 se muestran los sitios de hibridación y los tamaños de los fragmentos amplificados de los casetes de expresión integrados en los genomas de las dos cepas de levaduras generadas.

**Tabla 12.** Mezcla de reacción de PCR para verificar la integración del casete de expresión.

Reactivo	Concentración	Volumen (1X)	Volumen (6X)
		$\mu$ L	$\mu$ L
Agua estéril libre de nucleasas	-	13.3	79.8
Green Buffer	5X	5.0	30.0
dNTP's	10 mM	0.5	3.0
Oligonucleótido 5'	5 $\mu$ M	2.5	15.0
Oligonucleótido 3'	5 $\mu$ M	2.5	15.0
GoTaq DNA polimerasa	5 U/ $\mu$ L	0.2	1.2
DNA	100 ng/ $\mu$ L	1.0	-
Volumen final	-	25.0	144.0



**Figura 7.** Representación gráfica de los casetes de expresión integrados en el genoma de las levaduras transformadas con los plásmidos de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> se muestran las regiones de hibridación de los oligonucleótidos y tamaño de los fragmentos esperados amplificados por PCR con cada par de oligonucleótidos.

Posteriormente, se realizó el análisis de 20 clonas más por PCR de colonia utilizando el siguiente procedimiento: se tomaron 9  $\mu$ L de medio de cultivo con células de *P. pastoris* transformadas (DO<sub>600</sub>: 6-8) y se colocaron en tubos de 0.2 mL y se añadió 1  $\mu$ L de liticasa (5 U/ $\mu$ L) para un total de 10  $\mu$ L. Se incubaron a 30°C por 10 min. Se utilizó 1  $\mu$ L de lisado para la PCR y se emplearon los mismos volúmenes y las mismas condiciones que se mencionaron previamente.

## 6.10. Caracterización funcional de los elementos reguladores integrados en las cepas recombinantes construidas.

### 6.10.1. Detección de transcritos por RT-PCR

Para evaluar la funcionalidad de los elementos reguladores integrados en ambas cepas recombinantes construidas, se realizó una detección de transcritos por RT-PCR de una de las clonas de cada cepa construida y que fue positiva en el análisis de genotipificación por la PCR. Con las clonas seleccionadas se realizaron cultivos en matraz para obtener muestras de células para la extracción de RNA y posterior ensayo de RT-PCR.

#### **6.10.1.1. Cultivo de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>**

Una clona PCR positiva para la integración de ambas cepas construidas, se reactivó inoculando 50 µL de las células conservadas a -20 °C en 5 mL de medio YPD y se incubaron a 30°C con agitación de 250 rpm hasta alcanzar un valor de DO<sub>600</sub> entre 6 y 8. A partir del cultivo reactivado se inocularon 25 mL de medio YPD para alcanzar una DO<sub>600</sub> inicial de 0.3. Posteriormente se incubó a 30 °C con agitación a 250 rpm hasta alcanzar un valor de DO<sub>600</sub> entre 6 y 8. Partiendo del cultivo anterior en YPD se calculó el volumen necesario a inocular para obtener una DO<sub>600</sub> inicial de 1.4 en 100 mL de medio BMG-CaCl<sub>2</sub>, se incubó a 30 °C con agitación a 250 rpm y se tomaron muestras en la fase exponencial tardía, a las 10 y 12 h del cultivo, para hacer extracción de RNA y determinar la presencia de transcritos del gen *FTEII* por RT-PCR.

Teniendo en cuenta la densidad celular (g/L) del cultivo a las 10 y 12 h, se determinó el volumen de muestra que se debía tomar para obtener un total de 50 mg de células en peso húmedo. Se tomaron 3 alícuotas a las 10 y 12 horas del cultivo conteniendo 50 mg de biomasa húmeda y se depositaron en tubos de 2 mL, los cuales se centrifugaron a 6,000 rpm, durante 10 minutos a 4 °C y se desechó el sobrenadante. Los paquetes celulares se resuspendieron en 500 µL (volumen equivalente a 5 veces el volumen del paquete celular) del reactivo RNeasy lysis solution<sup>®</sup> (Qiagen) y se incubaron toda la noche a 4 °C. Posteriormente se centrifugaron a 6000 rpm durante 10 minutos a 4 °C y se descartó el sobrenadante. Finalmente, las muestras fueron etiquetadas y almacenadas a -70 °C para su posterior uso.

#### **6.10.1.2. Preparación de RNA de las células de las cepas de *P. pastoris* construidas**

Las preparaciones de RNA a partir de las células de ambas cepas conservadas previamente, se obtuvieron utilizando el kit SV Total RNA Isolation System (Promega, EUA), siguiendo el protocolo del manual del fabricante con algunas modificaciones, trabajando en un ambiente libre de polvo u otros contaminantes y utilizando guantes, puntillas, tubos y agua libres de RNasa. Una de las preparaciones de 50 mg (peso húmedo) de células conservadas en RNeasy lysis solution<sup>®</sup> que corresponde a la muestra tomada a las 10 horas de cultivo de cada cepa, fue descongelada y se mezcló con 100 µL de una

solución de lisis preparada en el momento [0.1 M EDTA (pH 7.4), 1 M sorbitol, 0.1%  $\beta$ -mercaptoetanol y 50 unidades de liticasa]. La mezcla fue resuspendida mediante vortex a velocidad media durante 5 segundos, posteriormente se incubó durante 50 minutos a 30°C en un agitador termoestático (Thermomixer R, Eppendorf) con agitación a 300 rpm. A continuación, se agregaron 75  $\mu$ L de buffer de lisis (4 M tiocianato de guanidinio [TCG], 0.001 M Tris-HCl, 0.97%  $\beta$ -mercaptoetanol) y 250  $\mu$ L de buffer de dilución de RNA, se mezclaron por inversión cada uno de los tubos y se centrifugaron a 12000 rpm durante 10 minutos a 20 °C. El sobrenadante (color azul) se transfirió en un tubo nuevo de 2 mL, se añadieron 200  $\mu$ L de etanol 95%, se mezcló por pipeteo, y se transfirió la mezcla a un cartucho de filtración. Se centrifugó a 12,000 rpm durante 1 minuto a 20 °C y se descartó el líquido del tubo colector. Posteriormente se añadieron 600  $\mu$ L de solución de lavado de RNA y se centrifugó a 12000 rpm, durante 1 minuto, a 20 °C. Una vez terminada la centrifugación, se preparó la solución de incubación de DNasa, la cual se añadió al cartucho de filtración (40  $\mu$ L de Yellow Core Buffer, 10  $\mu$ L de  $MnCl_2$  0.09M y 10  $\mu$ L de DNasa I) y se incubó a temperatura ambiente durante 30 minutos. Luego, se agregaron 250  $\mu$ L de la solución para desactivar la DNasa y el cartucho se centrifugó a 12000 rpm durante 1 minuto a 20 °C. Nuevamente se añadieron 600  $\mu$ L de la solución de lavado de RNA, se centrifugó en las mismas condiciones y se descartó el líquido del tubo colector. Posteriormente, se añadieron 250  $\mu$ L de solución de lavado de RNA y se centrifugó a 12,000 rpm, durante 2 minutos a 20 °C. Por último, el cartucho de filtración se transfirió a un tubo de elución de 1.5 mL y se agregaron 100  $\mu$ L de agua libre de nucleasas dejándolo en reposo 1 minuto a temperatura ambiente. Se centrifugó a 12000 rpm, durante 1 minuto a 20 °C. El cartucho de filtración fue desechado y el contenido del tubo colector fue etiquetado y almacenado a -20 °C.

El RNA obtenido se cuantificó por espectrofotometría a 260 y 280 nm utilizando un equipo de nanoespectrofotometría (Nanophotometer IMPLN) y se analizó mediante electroforesis en gel de agarosa al 0.8%. Se detectó la presencia de las dos bandas de las subunidades de rRNA (18s y 28s) en base a su tamaño molecular comparando con la migración electroforética de los tamaños de un marcador molecular de 100 pb DNA Ladder de Promega ®.



Para eliminar cualquier contaminación con DNA del RNA obtenido, se realizó un tratamiento adicional con una enzima DNAasa, empleando el Kit de Promega RQ1 RNase-Free DNAase. El primer paso para realizar el tratamiento con DNasa consistió en mezclar 3 µg de RNA a tratar, agua libre de RNasas, 3 µL de buffer de reacción y 3 µL de enzima RQ1 RNase-Free DNase en tubos de 0.2 ml. La mezcla se incubó en un termociclador durante 1 hora a 37°C. Una vez terminada la incubación, se agregaron 3 µL de solución de paro de DNAasa y se incubó de nuevo la muestra en un termociclador durante 15 minutos a 65°C. Finalmente, se analizó con un gel de electroforesis al 0.8%, se cuantificó mediante espectrofotometría y se almacenó a -20 °C hasta su uso para la síntesis del DNA complementario o DNAc.

### 6.10.1.3. Síntesis del DNA complementario (DNAc)

La síntesis del DNAc se realizó mediante una reacción de transcripción inversa a partir de las preparaciones de RNA tratadas con DNAasa. La reacción se realizó con las cantidades de reactivo que se muestran en la Tabla 13, empleando la enzima transcriptasa inversa del virus Moloney de la Leucemia Murina (M-MLV, Promega, Num. Cat. M170A), buffer de la enzima M-LV 5X y oligonucleótido d(T)<sub>15</sub> (Num. Cat. C110A).

**Tabla 13.** Volumen y concentración final de los reactivos empleados en la reacción de retrotranscripción.

Reactivo	KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub>		KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub>	
	Concentración Stock	Volumen µL	Concentración Stock	Volumen µL
Oligonucleótido d(T) <sub>15</sub>	0.5 µg/µL	1.0	0.5 µg/µL	1.0
Solución de RNA	113.0 ng/µL	10.0	58.0 ng/µL	18.0
Agua tratada con DEPC	0.1%	8.5	0.1%	0.5
Buffer para M-MLV	5X	6.0	5X	6.0
dNTP's	10 mM	1.5	10 mM	1.5
DTT	100 mM	1.0	100 mM	1.0
M-MLV	200 U/µL	2.0	200 U/µL	2.0
Total		30		30

Se realizaron dos reacciones para cada RNA extraído, una donde se adicionó la enzima transcriptasa inversa y otra con todos los reactivos sin la enzima (control negativo). Para cada reacción de 10 µL de la muestra de RNA (aproximadamente 1 µg) se

agregó 1  $\mu\text{L}$  del oligonucleótido d(T)<sub>15</sub> (0.5  $\mu\text{g}/\mu\text{L}$ ). Se calentaron las mezclas de reacción a 70°C por 5 min e inmediatamente se pusieron en hielo, se agregaron el resto de los reactivos de la reacción de transcripción inversa y se incubaron a 42°C por 60 min en el termociclador, tiempo durante el cual se realizó la síntesis del DNAc. Finalmente, las reacciones se incubaron a 65°C por 15 min para desnaturalizar DNAasas y se almacenaron a -20 °C.

#### **6.10.1.4. Reacción de PCR para detección de transcritos específicos**

Para las reacciones de PCR se utilizó el DNAc sintetizado previamente y la enzima *GoTaq* DNA polimerasa (Promega, EUA) para un volumen de 25  $\mu\text{L}$  de reacción. Por cada muestra se realizaron 5 reacciones de PCR: una reacción con la muestra empleando oligonucleótidos específicos (FTE1 y FTE2) que hibridan en la secuencia del gen heterólogo *FTEII* (tamaño del amplicón: 536 pb/temperatura de hibridación: 50 °C), otra con la muestra empleando oligonucleótidos específicos que hibridan en la secuencia de un gen de referencia  $\beta$ -actina (tamaño del amplicón: 150 pb temperatura de hibridación: 56 °C), otra con la reacción negativa de la RT empleando oligonucleótidos específicos 5qACT y 3qACT (Apéndice A) para el gen de referencia  $\beta$ -actina, otra como control negativo de la reacción de PCR utilizando 1  $\mu\text{L}$  de agua libre de nucleasas empleando oligonucleótidos específicos FTE1 y FTE2 (Apéndice A) que hibridan en el gen heterólogo *FTEII* y otra como control positivo de la reacción de PCR utilizando DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y oligonucleótidos específicos que hibridan en el gen heterólogo *FTEII* (FTE1 y FTE2).

Las condiciones de temperatura y tiempo que se emplearon en el termociclador para llevar a cabo la reacción de PCR se muestran en la Tabla 14. Se tomó una alícuota de cada producto de PCR y se analizó por electroforesis en gel de agarosa al 2.0 %.

**Tabla 14.** Condiciones del termociclador para el ensayo de PCR para la detección de transcritos.

<b>Etapa</b>	<b>Temperatura (°C)</b>	<b>Tiempo (min)</b>	<b>Ciclos</b>
Desnaturalización inicial	95	2	1
Desnaturalización	95	1	
Hibridación oligos <i>FTEII</i>	50	1	30
Hibridación oligos <i><math>\beta</math>-actina</i>	56	1	
Extensión	72	2	
Extensión final	72	5	1
Conservación	4	$\infty$	

### **6.10.2. Detección del producto proteico del gen reportero (fitasa)**

#### **6.10.2.1. Determinación de actividad volumétrica de fitasa extracelular**

Para esta determinación se realizaron cultivos de 3 clonas de cada cepa en las mismas condiciones que se describen en el acápite 6.10.1.1, pero alimentando con glicerol a las 14 horas manteniendo la concentración del glicerol al 1% p/v en el cultivo y prolongando el cultivo hasta las 24 horas. A las 24 horas se tomaron alícuotas de 2 ml de cultivo, las cuales se centrifugaron a 5000 rpm y 4°C por 10 minutos para recuperar el medio de cultivo libre de células y se almacenaron a -20°C hasta su uso.

Los sobrenadantes libres de células se desalaron utilizando cromatografía de exclusión molecular, con columnas PD-10 Sephadex G-25. La columna se equilibró previamente con una solución amortiguadora (Tris-HCl 100 mM, NaCl 50 mM, CaCl<sub>2</sub> 5 mM y glicerol 2%) eluyendo por gravedad y adquiriendo los datos mediante un software (Biologic LP) acoplado a un cromatógrafo (Biorad, Biologic LP), hasta que la conductividad registrada fue igual a la conductividad del buffer de elución (medido mediante el detector de conductividad ubicado a la salida de la columna PD-10). Una vez que estuvo equilibrada la columna se eluyó el buffer hasta la maya superior de la columna y se aplicó 1 mL de la muestra (previamente filtrada a través de un filtro Millipore de 0.22  $\mu$ M) y se eluyó por gravedad hasta que el total de la muestra entró completamente en la resina de la columna, cuidando que no se quede sin líquido la parte superior de la maya de la columna PD-10. Después se agregaron a la columna 5 mL de solución amortiguadora y se eluyó por gravedad colectando en un tubo de 15 mL libre de fosfatos hasta que la

señal de conductividad comenzó a subir. La muestra desalada quedó diluida 2.5 veces respecto a su concentración inicial.

La determinación de la actividad de fitasa se llevó a cabo mediante el método descrito previamente (Guerrero-Olazarán et al. 2010). Ese método cuantifica la liberación del fosfato inorgánico por acción de la fitasa sobre el fitato de sodio por unidad de tiempo. Dicho método cuantifica la cantidad de fosfato liberado del fitato en la reacción catalizada por la fitasa empleando ácido ascórbico y molibdato de amonio. El ortofosfato liberado en la reacción enzimática de la fitasa forma un complejo de fosfomolibdato, que se cuantifica a 820 nm y se compara con estándares.

Para estos ensayos se empleó como sustrato fitato de sodio (sal dodecasódica del ácido hexafosfórico de inositol) (Sigma Aldrich Corporation, PM: 770.04). La actividad de fitasa fue medida en buffer de actividad con 100 mM de Tris-HCl pH 7.5 y 1 mM de  $\text{CaCl}_2$ . Esta solución fue precalentada a 37°C por 30 min antes de realizar la reacción enzimática.

La elaboración de la recta de calibración se realizó por triplicado utilizando una solución stock de fosfato de potasio monobásico 1.64 mM preparado con el buffer de actividad. Los estándares y muestras se incubaron a 37°C durante 30 minutos, luego se agregaron 600  $\mu\text{L}$  de TCA al 15% y 800  $\mu\text{L}$  de solución colorante recién preparado (ácido ascórbico 10%, molibdato de amonio 2.5%, ácido sulfúrico 6 N). La mezcla de reacción se incubó a 37°C por 45 min, y se midió la absorbancia a 820 nm calibrando a cero con una celda de poliestireno y buffer de actividad.

Se definió una unidad de fitasa como la cantidad de enzima que cataliza la liberación de 1  $\mu\text{mol}$  de fosfato en un minuto bajo las condiciones del ensayo (fitato 1.6 mM,  $\text{Ca}^{+2}$  1 mM, 37 C y pH 7.5).

Para determinar la actividad enzimática de fitasa, se realizaron curvas de calibración con fosfato monobásico de potasio (FMP) disuelto en agua ultrapura, utilizando materiales y recipientes lavados con jabón libre de fósforo y agua ultrapura. La Tabla 15 muestra los estándares utilizados para realizar las curvas de calibración del

ensayo, realizando triplicados de cada dilución. A cada preparación de 100 µL de dilución se le añadieron 600 µL de ácido tricloroacético (TCA) al 15% y 800 µL de colorante.

**Tabla 15.** Volúmenes para la preparación de los estándares para la curva de calibración.

<b>Estándar de fosfato (mM)</b>	<b>Volumen del estándar (µL)</b>	<b>Volumen de buffer de actividad (µL)</b>
0.0	0	100
0.57	35	65
1.02	62	38
1.64	100	0

Las muestras se midieron sin diluir y a 40 µL de las muestras se añadieron a 60 µL de solución de fitato de sodio 8.33 mM en buffer de actividad. Junto a cada corrida de reacciones se añadió un blanco de fitato añadiendo 40 µL de buffer de actividad en 60 µL de solución de fitato. Para cada muestra se incluyó un blanco de muestra añadiendo 60 µL de buffer y 40 µL de muestra. El método se llevó a cabo de la siguiente manera: se colocaron 40 µL de muestra por triplicado en tubos de 2 mL y 60 µL de buffer de actividad para el blanco de muestra, así como 40 µL de buffer para el blanco de fitato. Se incubaron por 10 minutos a 37 °C en un Thermomixer R (Eppendorf) para precalentar. Posteriormente se añadieron 60 µL de solución de fitato a cada muestra, 40 µL de dilución de muestra a los blancos de muestra y 60 µL de solución de fitato al blanco de fitato. La reacción enzimática se incubó por 30 minutos a 37 °C. Inmediatamente terminado el tiempo de incubación, se añadieron 600 µL de TCA 15% (para detener la reacción) a todos los tubos, incluyendo blancos. Después, se añadieron 800 µL de colorante a todos los tubos y se incubaron por 45 minutos a 37 °C. Finalizado el tiempo de incubación se midió su absorbancia a 595 nm.

El cálculo de la actividad volumétrica enzimática se realizó empleando la siguiente ecuación:

$$APL = APM - ABF - ABM$$

Donde APM es la absorbancia promedio de la mezcla de reacción enzimática, ABF es la absorbancia promedio del blanco de sustrato (fitato) y ABM es la absorbancia

promedio del blanco de muestra. Posteriormente, se calculó la concentración de fosfato libre (mM) en la mezcla de reacción realizando una interpolación matemática del valor de la absorbancia generada por el fosfato liberado en la mezcla de reacción (APL) en la recta de calibración, calculada previamente a través de la regresión lineal por el método de mínimos cuadrados de los valores de absorbancia y concentraciones de las soluciones estándares de fosfato.

$$x=(y-b)/m$$

Donde “x” es la concentración de fosfato libre de la muestra (mM); “y” es el valor de la absorbancia generada por el fosfato liberado en la mezcla de reacción (APL); “b” es el valor de la intersección con el eje de las ordenadas, y “m” es la pendiente de la recta de calibración. Finalmente se calculó la actividad volumétrica de la muestra (U/mL) mediante la ecuación:

$$AVf=(x/t) \cdot FD \cdot 2.5$$

Donde AVf es la actividad volumétrica de fitasa en la muestra, “x” es la concentración (mM) de fosfato libre en la mezcla de reacción, “t” es el tiempo de incubación de la reacción enzimática (30 min), FD es el factor de dilución de la muestra y 2.5 es un factor que proviene de emplear 40  $\mu$ L de muestra en un volumen final de reacción de 100  $\mu$ L ( $100/40 = 2.5$ ). Como criterio de calidad, para aceptar el valor de la actividad volumétrica final calculado se tomó en cuenta que el valor de APM estuviera por debajo del valor de absorbancia del estándar de mayor concentración (1.62 mM) o menor a 2.954 y que x/t estuviese comprendido entre 0.0050 y 0.083.

## **6.10.2.2. Detección de fitasa extracelular por SDS-PAGE**

### **6.10.2.2.1. Ultrafiltración de sobrenadantes libres de células de los cultivos**

Para la detección del producto proteico del gen reportero, la fitasa FTEII, se realizaron cultivos en BMG-CaCl<sub>2</sub>, de una de las colonias seleccionadas de cada cepa recombinante construida tal como se describe en la sección 6.10.2.1. De cada cultivo se tomaron 10 mL

de muestra a las 24 horas y se centrifugó a 5,000 rpm y 4 °C durante 10 min para obtener los sobrenadantes libres de células los cuales se concentraron y desalaron por ultrafiltración con dispositivos de ultrafiltración Amicon Ultra 10-kDa de 4 mL (Millipore, MA, EUA). Se añadieron 4 mL del sobrenadante al dispositivo de ultrafiltración y se centrifugaron a 4500 rpm y 4 °C durante 10 min.

A continuación, se desechó el efluente, se rellenó el dispositivo con la muestra y se centrifugó en las mismas condiciones hasta completar el volumen total de la muestra. Después se concentró hasta llegar a los 100 µL y se desechó el efluente. Para desalar se hicieron 3 lavados con 100 µL de agua ultrapura estéril y centrifugaciones repetidas a 4500 rpm y 4 °C durante 10 min desechando el efluente entre cada centrifugación. Se recuperaron los 100 µL del concentrado (concentrado 100 veces) diafiltrado del dispositivo de ultrafiltración con micropipetas y puntillas de 200 µL en tubos de 0.6 mL. Se determinó la concentración de proteínas del concentrado por la técnica de Bradford haciendo el cálculo aproximado para diluir la muestra en base al factor de concentración de manera que se pueda leer en la curva. Finalmente se tomaron alícuotas de 30 µL del concentrado y se almacenaron a -20°C.

#### **6.7.2.2.2. Tratamiento para N-desglicosilación de las proteínas extracelulares de los concentrados de los cultivos de cepas recombinantes construidas.**

Para desglicosilar las proteínas extracelulares de los sobrenadantes concentrados de los cultivos de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> se tomó una alícuota de cada sobrenadante concentrado equivalente a 16 µg de proteínas totales y se sometió al tratamiento de N-desglicosilación con la desglicosidasa Endo H (New England Biolabs). En cada tubo de reacción se agregó el volumen requerido del concentrado de proteína, se añadió 1 µL de buffer de desnaturalización 10X y se completó con agua ultrapura para tener un volumen final de 10 µL como se muestra en la Tabla 16.

**Tabla 16.** Reacciones de desnaturalización.

Reactivos	KM71/P <sub>1033</sub> - FTEII-T <sub>AOX1</sub> con Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>AOX1</sub> sin Endo H ( $\mu$ L)	Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>1033</sub> con Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>1033</sub> sin Endo H ( $\mu$ L)
Muestra concentrada	9	9	---	9	9
Buffer de desnaturalización (10X)	1	1		1	1
Agua ultrapura	---	---	9	---	---
Volumen Total	10	10	10	10	10

Se incubaron las reacciones a 100 °C durante 10 min y se les añadió 2  $\mu$ L de buffer de reacción G5 10X y el volumen de agua necesaria para completar un volumen final de 20  $\mu$ L considerando el volumen de 1  $\mu$ L de enzima EndoH que se agregó a algunas reacciones (Tabla 17).

**Tabla 17.** Reacciones de desglicosilación.

Reactivos	KM71/P <sub>1033</sub> - FTEII-T <sub>AOX1</sub> con Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>AOX1</sub> sin Endo H ( $\mu$ L)	Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>1033</sub> con Endo H ( $\mu$ L)	KM71/P <sub>1033</sub> - FTEII-T <sub>1033</sub> sin Endo H ( $\mu$ L)
Buffer de Reacción G5 (10X)	2	2	2	2	2
Agua ultrapura	7	8	7	7	8
Endo H (1000 U/ $\mu$ L)	1	---	1	1	---
Volumen Final	20	20	20	20	20

Las reacciones de desglicosilación se incubaron a 37 °C durante 1 h. Pasado ese tiempo se les añadió 20  $\mu$ L de buffer Sampler 2X y se incubaron a 95°C durante 5 min.

### 6.10.2.2.3. Análisis de proteínas extracelulares por SDS-PAGE

El gel para la electroforesis se preparó con 12% poliacrilamida antes de realizar la reacción de desglicosilación y se dejó en la cámara de electroforesis vertical con buffer de corrimiento a 4°C. En la Tabla 18 se muestran las proporciones de los reactivos que se utilizaron para preparar el gel de poliacrilamida.



**Tabla 18.** Proporciones de los reactivos que se utilizaron para preparar el gel de poliacrilamida.

<b>Solución Stock</b>	<b>Gel concentrador (<math>\mu\text{L}</math>)</b>	<b>Gel separador (<math>\mu\text{L}</math>)</b>
Agua destilada	3,050	3,100
Acrilamida/bisacrilamida 30%	600	3,200
Tris-HCl 2 M pH 6.8	250	---
Tris-HCl 2 M pH 8.8	---	1,500
SDS 10%	40	80
PSA 10%	10	50
TEMED	5	10
Volumen Final	3,955	7,940

A continuación, se preparó la solución colorante de Comassie (Azul brillante de Comassie 0.5% p/v, metanol 50% v/v, ácido acético 10% v/v, en agua ultrapura) y la solución decolorante (metanol 5% v/v, ácido acético 7% v/v, en agua ultrapura). Se aplicaron 5  $\mu\text{L}$  del marcador de masa molecular y 10  $\mu\text{L}$  de las muestras en el gel de poliacrilamida con micropipetas y puntillas de 10  $\mu\text{L}$  y se programó la corrida con los parámetros de corriente eléctrica de 20 mA durante 2 h en frío, para ello se colocó la cámara de electroforesis vertical con el gel dentro del refrigerador a 4°C y se inició la corrida. Concluida la corrida se tiñó el gel con una solución colorante de Comassie durante toda la noche. Al día siguiente se colocó el gel en una solución decolorante durante 1 ó 2 días hasta que se observaron bien diferenciadas las bandas azules y el fondo claro.

### **6.11. Evaluación de la funcionalidad de la combinación de los elementos reguladores P<sub>1033</sub>-T<sub>AOX1</sub> y P<sub>1033</sub>-T<sub>1033</sub> en cepas unicopias cultivadas en dos fuentes de carbono**

#### **6.11.1. Identificación de clonas unicopia de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>**

Para la identificación y selección de cepas unicopias de las cepas recombinantes construidas, se eligieron al azar 32 y 20 clonas de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> respectivamente, se realizaron cultivos en BMG-CaCl<sub>2</sub> y se determinó la concentración de proteína extracelular y rendimiento proteína extracelular/biomasa de muestras de cultivo a las 24 horas. Las cepas con rendimientos más bajos

fueron seleccionadas para determinar por qPCR la dosis génica de las cepas consideradas como unicopias.

#### **6.11.1.1. Cultivos en matraz agitado**

Para seleccionar las clonas unicopias de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> se realizaron cultivos de 32 y 20 clonas respectivamente de cada cepa en medio BMG. Primero se realizó un cultivo en lote con glicerol durante 14 h, seguido de un cultivo de lote alimentado con glicerol, realizando una única adición de glicerol para tener una concentración 1 % (p/v) a las 14 horas hasta un tiempo total de cultivo de 24 horas. Cada clona se activó inoculando 50 µL de células conservadas a 4 °C en 5 mL de medio YPD y se incubó a 30 °C con agitación a 250 rpm hasta alcanzar un valor de DO<sub>600</sub> entre 6 y 8. Se tomó el volumen necesario del cultivo anterior para una DO<sub>600</sub> inicial de 0.3, se inoculó en 10 mL de medio YPD y se incubó a 30 °C con agitación a 250 rpm hasta alcanzar un valor de DO<sub>600</sub> entre 6 y 8. Partiendo del cultivo anterior en YPD se calculó el volumen necesario para inocular en 25 mL de medio BMG-CaCl<sub>2</sub> (buffer de fosfatos 100 mM, glicerol 1%, YNB 1.34%, biotina 1X, CaCl<sub>2</sub> 0.1%) y obtener una DO<sub>600</sub> inicial de 1.4. Este volumen se centrifugó a 5,000 rpm por 10 min a 4 °C, se desechó el sobrenadante y se resuspendieron las células en igual volumen de medio BMG-CaCl<sub>2</sub>, el cual se incorporó al matraz de 250 mL que contenía el medio BMG-CaCl<sub>2</sub> para completar los 25 mL. Se midió la DO<sub>600</sub> inicial del cultivo, se tomó 1 mL de muestra para determinar proteínas totales y se incubó a 30 °C durante 14 h a 250 rpm. A las 14 h de incubación se midió la DO<sub>600</sub>, se tomó 1 mL de muestra para determinar proteínas totales, se añadieron 2.5 mL de glicerol al 10% y se incubó a las mismas condiciones por 10 h hasta un tiempo total de 24 h de cultivo. A las 24 h de cultivo, se midió la DO<sub>600</sub> y se tomó 1 mL de muestra para determinar concentración de proteínas totales y 2 mL para determinar concentración de fitasa extracelular.

### 6.11.1.2. Determinación de proteínas extracelulares de los cultivos

Para determinar la concentración de proteínas totales en el medio extracelular de los cultivos de KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> se utilizó el método de Bradford, utilizando una solución estándar de albúmina sérica bovina (BSA) a una concentración de 50 µg/mL y azul brillante de Coomassie G-250. Primero se hizo una curva de calibración realizando diluciones por triplicado de la solución estándar de BSA para un volumen total de 200 µL y agregando a cada reacción 800 µL de azul de Coomassie (Tabla 19), esta mezcla se homogenizó con ayuda de un vórtex por 10 segundos, se incubó 5 min a temperatura ambiente y se hicieron lecturas de absorbancia en un espectrofotómetro (Implen, Alemania) a 595 nm (Bradford 1976). Las muestras se analizaron añadiendo a 200 µL de los sobrenadantes del cultivo sin diluir y 800 µL de azul de Coomassie, siguiendo el mismo procedimiento que con las soluciones estándar. Los valores de absorbancia de la muestra se leyeron en el espectrofotómetro a 595 nm y se obtuvo el valor de concentración de la muestra en base a la última curva de calibración que se resguardó en el equipo utilizado.

**Tabla 19.** Concentración de los estándares de BSA para la curva de calibración.

<b>Concentración Estándar (µg/mL)</b>	<b>Volumen de BSA (50 µg/mL) (µL)</b>	<b>Agua destilada (µL)</b>
0.0	0	200
5.0	20	180
7.5	30	170
15.0	60	140
22.5	90	110
30.0	120	80
45.0	180	20

A todas las mezclas de BSA y agua se añadieron 800 µL de reactivo de Coomassie.

### **6.11.1.3. Determinación del rendimiento proteína extracelular por unidad de biomasa**

Para la determinación del rendimiento proteína extracelular por unidad de biomasa, se estimó la densidad o concentración celular en gramos de células seca por litro de cultivo de las muestras de 24 h de los cultivos de KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>. La biomasa en peso seco se obtuvo multiplicando el valor de la DO<sub>600</sub> de la muestra del cultivo a las 24 h por 0.23. El rendimiento de proteínas extracelulares por unidad de biomasa en peso seco (mg/g) se determinó dividiendo la concentración de proteínas extracelulares (mg/L) entre la densidad celular (g/L) a las 24 h de cultivo. Se seleccionaron las 3 clonas con menor rendimiento proteína/biomasa para realizar cultivos en matraz y determinar la dosis génica, con el objetivo de encontrar clonas con una copia del casete de expresión integrada en el genoma de la levadura.

### **6.11.2. Determinación de la dosis génica de las cepas unicopias identificadas**

La determinación de la dosis génica se realizó a partir de la curva de eficiencia obtenida por qPCR (protocolo descrito en la sección 6.11.2.3) de la secuencia de *FTEII* integrada en el genoma de una de las clonas identificada como unicopia por los valores de bajo rendimiento proteína extracelular/biomasa y teniendo como referencia la curva correspondiente al gen *GAPDH*. Para la obtención del DNA genómico y estimación de concentración se realizaron cultivos de las clonas seleccionadas, de acuerdo con el protocolo descrito en la sección 6.11.2.1 y 6.11.2.2. Para los ensayos de qPCR se usaron sondas para ambos genes evaluados. A partir de las curvas de eficiencia se determinó la dosis génica de las clonas consideradas como unicopias.

#### **6.11.2.1. Cultivos de las clonas seleccionadas como unicopias**

Tres clonas seleccionadas con menor rendimiento de proteínas extracelulares/biomasa se cultivaron a nivel de matraz para lo cual se inocularon 100 µL de cada clona en 10 mL de

medio YPD y se incubaron a 30 °C y 250 rpm hasta que alcanzaron una DO<sub>600</sub> entre 6 y 8. Se dispensó cada cultivo en viales de 2 mL, posteriormente se centrifugó a 6000 rpm y 4 °C por 10 minutos y se desechó el sobrenadante, se dispensó nuevamente 2 mL y se centrifugó en las mismas condiciones para obtener paquetes celulares de 4 mL de cultivo, los cuales se almacenaron a -20 °C para ser empleados en la preparación de DNA genómico.

#### **6.11.2.2. Preparación de DNA genómico de la identificación de las posibles clonas unicopias**

Partiendo de los paquetes celulares conservados a -20°C de los cultivos de las clonas de menor rendimiento de proteínas extracelulares/biomasa, se extrajo DNA genómico por el método de extracción con solventes orgánicos (TSNT) descrito previamente. Se purificó el DNA genómico obtenido, primero se realizó un tratamiento con la RNase One (Promega) para eliminar el RNA presente en las muestras, añadiendo a todos los tubos 1 µL de la enzima, los cuales se incubaron a 37 °C por 1 hora.

Posteriormente, se realizó una diafiltración de las muestras para eliminar nucleótidos contaminantes en los dispositivos de filtración Amicon Ultra 3-kDa de 0.5 mL (Millipore, MA, EUA). Se adicionaron 50 µL de cada muestra en los cartuchos acoplados dentro de tubos de 2.0 mL y se añadieron 400 µL de agua estéril libre de nucleasas a cada uno. Se centrifugaron a 12000 rpm por 10 min a temperatura ambiente y al sacarlos de la centrifuga se descartó el líquido de la parte inferior de los tubos. Siguiendo el mismo procedimiento, se hicieron 3 lavados adicionales. Finalmente, para recuperar el DNA genómico purificado, se colocaron los cartuchos de forma invertida en tubos nuevos de 2.0 mL, se centrifugaron a 3000 rpm durante 3 min y se descartaron los cartuchos. Los tubos con el ADN colectado se calentaron a 65 °C por 20 minutos. Se tomaron 3 µL de cada muestra, se analizaron en un gel de agarosa al 0.8% y se cuantificaron por espectrofotometría como se describió en otros apartados.

### 6.11.2.3. Curvas para determinar la eficiencia de amplificación de los genes *FTEII* y *GAPDH*

Los ensayos de qPCR se realizaron utilizando el kit qPCR ProbesMaster (Número de catálogo PCR-360S, Jena Bioscience), sondas para los genes *FTEII* y *GAPDH* y el termociclador Mx3005P QPCR System de Agilent Technologies.

Las rectas para determinar la eficiencia de amplificación de los genes *FTEII* y *GAPDH* se realizaron con 4 puntos con DNA genómico de la clona considerada como unicopia por su bajo rendimiento proteínas extracelulares/biomasa. Las curvas se realizaron con 25, 5, 1 y 0.2 ng de DNA genómico en la reacción final. Las reacciones de qPCR de cada punto de la curva se realizaron por duplicado en un volumen total de 20  $\mu$ L, para lo cual se añadieron 10  $\mu$ L de Master Mix, 2  $\mu$ L de agua libre de nucleasas, 4  $\mu$ L de DNA y 4  $\mu$ L de Prime Time para el gen *FTEII* o para el gen *GAPDH*. Para cada gen, se realizó un ensayo negativo de la reacción de qPCR con el objetivo de verificar la ausencia de DNA genómico contaminante en los reactivos utilizados.

Con las curvas de eficiencia de amplificación determinadas para *FTEII* y *GAPDH*, se analizaron las muestras de 3 clonas de ambas cepas que tenían bajo rendimiento proteínas extracelulares/biomasa. El análisis se realizó con 1 ng de DNA genómico en el ensayo de qPCR y se estimó la cantidad de DNAG del gen *FTEII* y *GAPDH* en el intervalo de linealidad de acuerdo con el valor de Ct obtenido y el valor de *FTEII* se normalizó respecto al obtenido para *GAPDH*. El valor normalizado de *FTEII* cercano a 1 fue considerado como el representativo para las clonas unicopias, es decir con una copia del casete de expresión integrada en el genoma de la levadura.

El cálculo de la eficiencia se determinó a partir de la recta formada por el logaritmo de la cantidad de DNAC (eje de ordenadas) y el valor de Ct para cada dilución (eje de las abscisas) y calculada a partir de la ecuación:

$$E = 10^{-1/m} - 1$$

Donde E es la eficiencia y m es la pendiente de la recta del gráfico de logaritmo de la cantidad de DNAC contra valor de Ct para cada dilución de ADN genómico.

### **6.11.3. Actividad transcripcional del P<sub>1033</sub> en combinación con T<sub>AOX1</sub> y T<sub>1033</sub>**

Para determinar la actividad transcripcional de los casetes de expresión conteniendo la combinación de las regiones reguladoras P<sub>1033</sub>- T<sub>AOX1</sub> y P<sub>1033</sub>- T<sub>AOX1</sub> integradas en los genomas de las cepas recombinantes construidas se realizaron cultivos de una clona unicopia de cada cepa construida para determinar cinéticas de crecimiento y de expresión. Además, se realizaron cinéticas de rendimiento producto extracelular/biomasa ( $Y_{p/x}$ ) para correlacionar con la actividad transcripcional de la cepa. Para las cinéticas de expresión se realizaron ensayos de RTqPCR de células a diferentes tiempos del cultivo de los genes *FTEII*, *1033* y *GAPDH* éste último como gen de referencia.

#### **6.11.3.1. Cultivos para cinéticas de crecimiento y de expresión de las cepas unicopias**

Las dos cepas unicopia seleccionadas (KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> clona C8 y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> clona C36) se cultivaron en matraces de 1L conteniendo 150 mL de medio BMGly-CaCl<sub>2</sub> y BMGlc-CaCl<sub>2</sub> (medios BM-CaCl<sub>2</sub> con 30 mM de glicerol [0.28 % p/v] o 30 mM de glucosa [0.54% p/v], respectivamente) durante 18 o 24 h a 30 °C y 250 rpm, iniciando con una DO<sub>600</sub> de 0.1 y sin alimentación adicional a lo largo del cultivo.

Se tomaron muestras cada 3 h hasta las 6 h de cultivo y después cada 6 h para medir el crecimiento celular mediante ensayo de turbidimetría a DO<sub>600</sub>. Las velocidades específicas de crecimiento ( $\mu$ ) se determinaron mediante la pendiente de la regresión lineal del logaritmo natural de la DO<sub>600</sub> frente al tiempo de la fase exponencial (3-12 h), y se compararon estadísticamente mediante una prueba t de Student con un nivel de significancia de 0.05.

#### **6.11.3.2. Determinación de los niveles de transcritos de los genes *1033*, *FTEII* y *GAPDH* en células de las cepas recombinantes cultivadas en glucosa y glicerol**

Las rectas para determinar la eficiencia de amplificación de los genes *FTEII*, *GAPDH*, *YPT1* y *1033* se realizaron con 5 puntos con DNA genómico de la clona unicopia (C8) de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Las curvas se realizaron con 25, 5, 1, 0.2 y 0.04 ng de DNA genómico en la reacción final. Las reacciones de qPCR de cada punto de la curva se realizaron por duplicado en un volumen total de 20  $\mu$ L, para lo cual se añadieron 10  $\mu$ L de Master Mix, 2  $\mu$ L de agua libre de nucleasas, 4  $\mu$ L de DNA y 4  $\mu$ L de Prime Time para

el gen *FTEII* o para el gen *GAPDH* o de los oligonucleótidos 5q1033 y 3q1033 (Apéndice A) para el gen *1033*, o bien 5qYPT1 y 3qYPT1 (Apéndice A) para el gen *YPT1*. Para cada gen, se realizó un ensayo negativo de la reacción de qPCR con el objetivo de verificar la ausencia de DNA genómico contaminante en los reactivos utilizados.

Los niveles de transcritos de los genes *1033*, *FTEII* y *GAPDH* se determinaron mediante ensayos de RT-qPCR de células de cultivos en glucosa o glicerol a una concentración de 30 mM, tal como se describe en el apartado 6.11.3.1. Las células fueron colectadas a las 0, 6, 12 y 18 horas de cultivo en BMGly-CaCl<sub>2</sub> y BMGlc-CaCl<sub>2</sub> de las dos cepas unicopia, KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> clona C8 y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> clona C36, colectando un volumen del cultivo necesario para obtener 2 mg de peso húmedo de células (Tabla 24).

Las muestras colectadas se centrifugaron para obtener paquetes celulares que se conservaron en 200 µL RNAlater por 16 horas a 4 °C. Posteriormente, los paquetes celulares se centrifugaron a 12000 rpm a 4 °C por 5 min para eliminar el excedente de la solución de RNAlater y se conservaron a -20 °C. El RNA se extrajo con el kit SV Total RNA Isolation System (Promega, EUA) como se describió previamente. La calidad del RNA obtenido se evaluó mediante la relación de absorbancia de 260 a 280 nm (NanoPhotometer Pearl, Implen GmbH, Munich, Alemania) y mediante electroforesis en gel de agarosa.

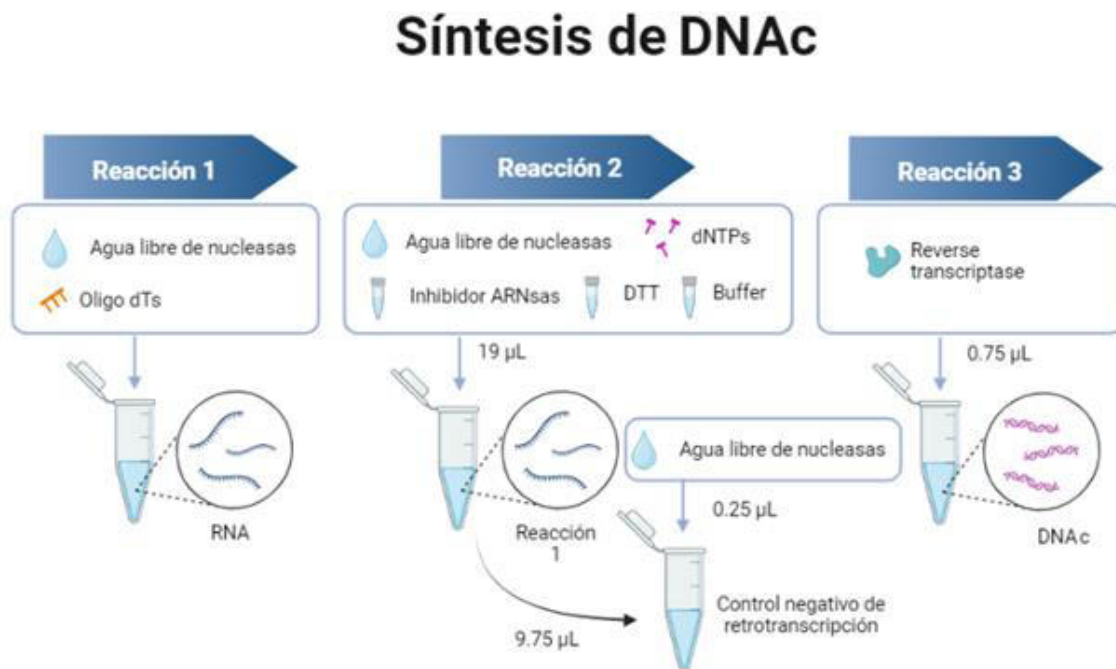
La síntesis del DNAc se realizó a partir de 1 µg de RNA total utilizando el kit SCRIPT cDNA Synthesis (Jena Bioscience GmbH, Jena, Alemania) de acuerdo con las instrucciones del fabricante, primero a cada volumen de RNA equivalente a 1 µg se le añadió 1 µL de Oligo dT y el volumen de agua libre de nucleasas necesario para completar reacciones de 20 µL, las cuales se incubaron a 70 °C por 5 minutos y se colocaron rápidamente en una hielera a -20 °C (Fig.8). A continuación, se preparó un coctel de reacción como se muestra en la Tabla 20 y se dispensaron 19 µL en cada una de las reacciones conservadas a -20 °C, se mezcló por pipeteo suavemente y sin sacar los viales de la hielera. De cada reacción se extrajeron 9.75 µL y se les añadió 0.25 µL de agua libre de nucleasas para utilizarlos como control negativo de retrotranscripción (Fig.8). Al resto del volumen de cada reacción se añadió 0.75 µL de la enzima retrotranscriptasa para completar un volumen total de 10 µL de cada reacción (Fig.8). Las reacciones y los



controles negativos se incubaron a 42 °C por 10 minutos y 50 °C por 60 minutos en un termociclador Mx3005P QPCR (Agilent Technologies, Santa Clara, CA, EUA), a continuación, se incubaron a 70 °C por 10 minutos para inactivar la enzima. Las muestras de DNAc y los controles negativos de retrotranscripción se almacenaron a – 20 °C.

**Tabla 20.** Coctel de reacción para síntesis de DNAc.

Reactivos	Volumen de reactivos para cada muestra de RNA (µL)
Agua ultrapura estéril	6
Buffer	8
dNTPs	2
DTT	2
Inhibidor de ARNsas	1
<b>Total</b>	<b>19 x número de muestras de RNA</b>



**Figura 8.** Esquema de las etapas en las reacciones para la síntesis de DNAc.

Cada DNAc específico de cada transcrito (genes *1033* y *FTEII*) de cada muestra se amplificó por duplicado utilizando el termociclador Mx3005P QPCR (Agilent

Technologies, Santa Clara, CA, E.U.A.). También se obtuvieron los niveles de transcritos del gen *YPT1* para emplearse como gen normalizador y los niveles de transcripción del gen *GADPH* como referencia de la actividad transcripcional. Para la determinación de los niveles de transcritos de *FTEII* y *GADPH* se empleó el qPCR ProbesMaster (Jena Bioscience Alemania) y el kit PrimeTime qPCR Probe Assays (Integrated DNA Technologies, Inc.) que incluye los iniciadores y sondas de hidrólisis específicos para *FTEII* y *GADPH*, mientras que para la determinación de los niveles de expresión de los genes endógenos *1033* y *YPT1* se utilizó el agente intercalante de DNA SYBR Green (qPCR SybrMaster, Jena Bioscience Alemania) con los oligonucleótidos 5q1033 y 3q1033 (Apéndice A) o bien 5qYPT1 y 3qYPT1 (Apéndice A). Para cada muestra se realizó un ensayo negativo de la reacción de retrotranscripción (RT) para verificar la ausencia de DNA genómico contaminante. En cada conjunto de reacciones o corrida de qPCR amplificando con una pareja de oligonucleótidos, se realizó una reacción negativa de qPCR utilizando agua libre de nucleasas estéril.

Las reacciones de qPCR se realizaron en un coctel ( de un volumen total según el número de muestras sujetas al ensayo) con los volúmenes de reacción que se muestran en la Tabla 21, mezclando en una primera etapa el master mix con agua libre de nucleasas (10 µL de master mix y 4 µL de agua libre de nucleasas para cada muestra), de esta mezcla se dispensaron 56 µL en viales de 0.2 mL para cada muestra a analizar, y a cada uno de estos viales se le añadieron 8 µL de DNAC a la dilución 1:125. Cada una de estas reacciones se dividió en 2 viales con 32 µL cada uno a los cuales se añadieron 8 µL de cada sonda para las reacciones con Probe Master o de cada mezcla de primers para las reacciones con Sybr Green, finalmente se distribuyeron 20 µL de cada mezcla en 2 tubos ópticos especiales de 8X y se cerraron con tapas ópticas especiales de 8X (Fig. 9). La Tabla 21 muestra los volúmenes utilizados en los ensayos de qPCR con sondas de hidrólisis y SYBR Green, en un volumen final de 20 µL por reacción, y de 10 µL por control negativo de qPCR (agua ultrapura estéril en lugar de DNAC en la reacción). Las reacciones se mezclaron en tubos de 0.65 mL a partir de los cuales se distribuyeron 20 µL de cada mezcla en 2 tubos ópticos especiales de 8X y se cerraron con tapas ópticas especiales de 8X. Se centrifugaron todos los tubos con las mezclas realizadas en una mini centrífuga durante 10 segundos y se observó que no tuvieran burbujas de aire dentro del líquido. En caso de tenerlas, se dio un

golpe en la parte baja del tubo y se centrifugaron otra vez por 10 segundos. Cuando ya no tuvieron burbujas se colocaron en el bloque de calentamiento del termociclador.

**Tabla 21.** Volúmenes de reacción para los ensayos de qPCR con sondas de hidrólisis o SYBR Green.

<b>SYBR Green</b>	<b>Volumen (μL)</b>	<b>Sondas de hidrólisis</b>	<b>Volumen (μL)</b>
qPCR SybrMaster	10	qPCR Probe Master	10
Agua ultrapura estéril	4	Agua ultrapura estéril	4
Oligonucleótido 5' (3 μM)	2	Prime Time	4
Oligonucleótido 3' (3 μM)	2	DNAc	2
DNAc	2		
<b>Total</b>	<b>20</b>		<b>20</b>



**Figura 9.** Esquema de la estructura del montaje para las reacciones de qPCR.

Los programas para los ensayos de qPCR utilizados para las reacciones con SYBR Green y con sondas de hidrólisis se muestran en las Tablas 22 y 23, respectivamente.

**Tabla 22.** Programa del termociclador para los ensayos de qPCR empleando SYBR Green.

Etapa	Ciclos	Temperatura (°C)	Duración (s)	Proceso
1	1	95	60	Desnaturalización
2	40	95	20	Desnaturalización
		60	20	Hibridación, elongación y almacenamiento de datos de fluorescencia
3	1	95	60	Desnaturalización
		55	30	Curva de fusión y almacenamiento de datos de fluorescencia
		95	30	

**Tabla 23.** Programa del termociclador para los ensayos de qPCR empleando sondas de hidrólisis.

Etapa	Ciclos	Temperatura (°C)	Duración (s)	Proceso
1	1	95	60	Desnaturalización
2	40	95	20	Desnaturalización
		60	20	Hibridación, elongación y almacenamiento de datos de fluorescencia

Se determinaron los valores del ciclo umbral (Ct) para cada reacción con el programa MxPro (Agilent Technologies) a un umbral de fluorescencia de 400. Para cada cultivo se consideraron tres réplicas biológicas. Cada muestra de RNA se retrotranscribió a DNAc y por cada DNAc se obtuvo un duplicado de amplificación de cada gen, los cuales se promediaron y se tomó como criterio de aceptación de la técnica de qPCR que el coeficiente de variación de estos duplicados de Ct fuera menor al 2 %.

Para el cálculo de los niveles de transcritos de los tres genes de interés (*FTEII*, *1033* y *GAPDH*), se realizaron rectas de calibración (logaritmo de la cantidad de DNA contra valor de Ct) de los cuatro genes (*FTEII*, *GAPDH*, *1033* y *YPT1*) empleando diluciones 1:5 de DNA genómico de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> clona C8 en el intervalo de 25.00 hasta 0.04 ng totales de DNA genómico (5 puntos). Con la pendiente (m) e intersección en el eje de ordenadas (b) de las rectas de calibración se determinó la eficiencia (E) de cada amplificación y la cantidad de DNA al valor de fluorescencia umbral (X<sub>t</sub>) según las siguientes ecuaciones:

$$E = 10^{(-\frac{1}{m})} - 1$$

$$X_t = 10^{-\frac{b}{m}}$$

El cálculo de los niveles de transcritos se realizó de forma relativa a los niveles de transcritos del gen *YPT1* según el método de  $2^{-\Delta Ct}$  con corrección de eficiencia y empleando la siguiente ecuación:

$$NE = \frac{X_{tGOI}(1 + E_{GOI})^{-C_{tGOI}}}{X_{tYPT1}(1 + E_{YPT1})^{-C_{tYPT1}}}$$

Donde NE es el nivel de transcritos del gen de interés (GOI: *FTEII*, 1033 y *GAPDH*) en una muestra relativos a los niveles de transcritos de *YPT1*.

### 6.11.3.3. Determinación del rendimiento producto extracelular por unidad de biomasa

A partir de los cultivos descritos en el punto 6.11.3.1 se tomaron muestras a las 0, 6, 12, 16 y 18 horas de cultivo para determinar la densidad celular por espectrofotometría a 600 nm y alícuotas para determinación de proteínas extracelulares y actividad volumétrica de fitasa. A partir de los valores de las  $DO_{600}$  se estimaron los volúmenes de las alícuotas necesario para obtener un total 8 mg de peso seco de cada muestra, teniendo en cuenta que una  $DO_{600}$  multiplicada por 0.23 equivale a 1mg de peso seco de células/ mL de cultivo. Los volúmenes de muestra tomados a cada tiempo para cada determinación se muestran en la Tabla 24.

**Tabla 24.** Volúmenes de las alícuotas tomadas de los cultivos a diferentes tiempos.

Tiempo de cultivo (h)	DO 600 nm	Volumen (mL) de muestra para obtener 2 mg de células en peso húmedo para preparación de RNA	Volumen (mL) de muestra para obtener 8 mg de células de peso seco para determinar proteínas y actividad de fitasa	Factor de concentración para determinar proteínas y actividad de fitasa
0	0.137	16.0	-	-
6	0.705	3.0	49.37	49.37
12	3.023	0.8	11.51	11.51
16	8.215	0.4	4.28	4.28
18	8.457	0.2	4.23	4.23

Las alícuotas de los sobrenadantes de los cultivos se diafiltraron por ultrafiltración con dispositivos de ultrafiltración Amicon Ultra 10-kDa de 4 mL (Millipore, MA, EUA). Se añadieron 4 mL del sobrenadante al dispositivo de ultrafiltración y se centrifugaron a 4500 rpm y 4°C durante 10 min, a continuación, se desechó el filtrado, se rellenó el dispositivo con la muestra y se centrifugó en las mismas condiciones hasta completar el volumen estimado según la Tabla 24. El sobrenadante se concentró hasta llegar a 1 mL desechando el efluente, para diafiltrar se hicieron 3 lavados con 3 mL de buffer de elución (Tris-Cl 100 mM pH 8.5, CaCl<sub>2</sub> 5 mM, NaCl 50 mM, Glicerol 2%) y centrifugaciones repetidas a 4500 rpm y 4 °C durante 10 min desechando el efluente entre cada centrifugación. El diafiltrado final se concentró hasta 1 mL y se recuperó del dispositivo de ultrafiltración.

Al diafiltrado concentrado se le determinaron proteínas por la técnica de Bradford tal como se describe en el acápite 6.11.1.2 y se dividió por el factor de concentración asociado a cada muestra en cada tiempo para calcular la concentración de proteínas en µg/mL de cultivo.

La actividad volumétrica de fitasa se determinó como se describe en el acápite 6.10.2.1.

La densidad celular o biomasa expresada en g/L se determinó a partir de la DO<sub>600</sub> determinada de muestras de los cultivos en glucosa o glicerol de las clonas de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> seleccionadas como unicopias, multiplicando por el factor de 0.23.

El rendimiento de producto por unidad de biomasa ( $Y_p/x$ ) se determinó dividiendo la actividad volumétrica de fitasa (U/L) del medio extracelular libre de células entre la densidad celular (g/L) a los 4 tiempos de cultivo analizados.

## **6.12. Análisis estadístico**

Se realizaron análisis estadísticos de comparación de medias mediante pruebas t de Student para muestras independientes y muestras pareadas sin asumir varianzas iguales con un nivel de significación de 0.05 para comparar los niveles de expresión relativa de cada gen con respecto a un gen normalizador en las dos cepas construidas y en las dos condiciones de cultivo, y también para comparar entre los tiempos de cultivo para cada

gen analizado. De igual manera se realizaron los análisis estadísticos de las tasas de crecimiento específico ( $\mu$ ) y los rendimientos extracelulares producto/biomasa ( $Y_{p/x}$ ) en los cultivos de ambas fuentes de carbono y para ambas cepas. Los cálculos se realizaron empleando el programa Microsoft Excel 365.

## 7. RESULTADOS

### 7.1. Identificación de las secuencias del promotor y terminador de la transcripción del gen *1033*

#### 7.1.1. Mapeo de los transcritos contra el genoma de referencia de *P. pastoris*

Para caracterizar las regiones del promotor ( $P_{1033}$ ) y el terminador transcripcional ( $T_{1033}$ ) del gen *1033*, se identificaron sus secuencias junto con la secuencia de la región 5'UTR, el sitio de inicio de la transcripción (TSS), y la secuencia de la región 3'UTR mediante el análisis de los cinco transcriptomas de RNA-seq de dos cepas de *P. pastoris* KM71 crecidas en metanol, glicerol o lactosa y localizando regiones río arriba y río abajo del CDS del gen *1033* donde no alinearon transcritos.

Las posiciones de inicio y final de las regiones no transcritas río arriba y río abajo del gen *1033* según las anotaciones del cromosoma 3 de la cepa CBS 7435 se describen en la Tabla 25.

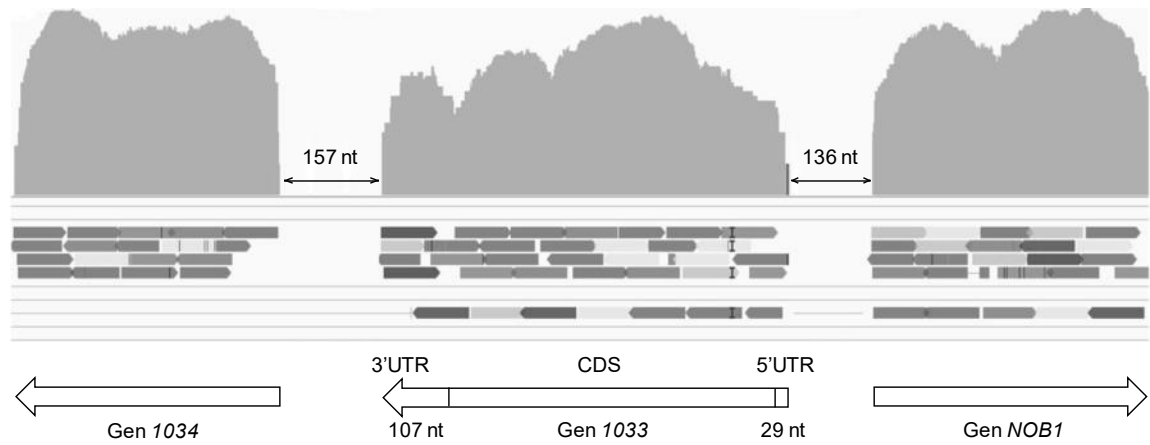
**Tabla 25.** Posición de las regiones no transcritas río arriba y río abajo del gen *1033* respecto al genoma de la cepa de *P. pastoris* CBS7435 a partir de los cinco transcriptomas analizados.

Transcriptomas	Región no transcrita río arriba del gen 1033		Región no transcrita río abajo del gen 1033	
	Inicio	Fin	Inicio	Fin
CM2M0	256738	256891	255319	255487
CM2M7	256754	256879	255341	255480
CM3M6	256744	256891	255338	255497
CP1M6	256747	256880	255325	255481
CP3M3	256756	256884	255318	255487
Promedio	256748	256885	255328	255486

La posición de inicio de la región no transcrita río arriba del gen *1033* tuvo un coeficiente de variación de 0.003% y la posición final de 0.002%. La longitud media de esta región fue de  $136 \pm 12$  nt. La posición de inicio de la región no transcrita río abajo del gen *1033*



tuvo un coeficiente de variación de 0.004% y la posición final de 0.003%. La longitud media de esta región fue de  $157 \pm 12$  nt (Fig. 10).



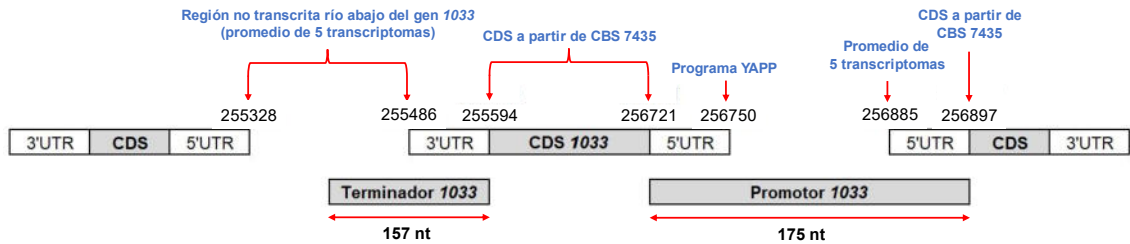
**Figura 10.** Resultados del mapeo del análisis de los cinco transcriptomas visualizadas con el programa Integrative Genomics Viewer (IGV) identificando las regiones río arriba y río abajo del CDS del gen *1033*. Las subregiones donde no alinearon las lecturas de ARN se muestran con una flecha de doble punta. También se muestra la orientación del gen *1033* y los genes adyacentes respecto al genoma de la cepa *P. pastoris* CBS 7435, y las regiones 5'UTR y 3'UTR del gen *1033*. nt: nucleótidos.

El inicio y fin de la región río arriba del CDS del gen *1033* sirvió para definir el inicio de las regiones 5'UTR del gen *1033* y del gen río arriba (*NOB1*), mientras que el inicio y fin de la región río abajo del CDS del gen *1033* sirvió para definir el final de la región 3'UTR del gen *1033* y el inicio de la región 5'UTR del gen río abajo (gen *1034*).

Debido a que la región río arriba del gen *1033* donde no mapearon transcritos (136 nucleótidos) abarcó el 80% de la región entre el CDS del gen *1033* y el CDS del gen río arriba (*NOB1*), se consideró usar como secuencia del promotor del gen *1033* (Fig. 11) a los 175 nucleótidos que abarcan toda la región entre los dos CDS. El sitio del inicio de la transcripción se identificó a 29 nucleótidos río arriba del codón de inicio de la traducción (Fig. 10).

El análisis transcriptómico mostró que las lecturas de RNA se alinearon hasta 107 nucleótidos río abajo del CDS del gen *1033* (Fig. 10), terminando con el triplete CAA que es un sitio característico de corte y poliadenilación de RNA para levaduras (Graber 2002). El sitio de corte y poliadenilación del RNA está flanqueado por dos secuencias ricas en A+T (ATATGTAT y ATTTATT) que están sobrerrepresentadas en secuencias río abajo de los CDS de levaduras (Van Helden et al. 2000). Como secuencia del terminador del

gen *1033*, se consideró la región 3'UTR del gen *1033* (107 nucleótidos desde dónde termina el CDS del gen *1033* hasta el sitio de corte y poliadenilación) más 50 nucleótidos adicionales, por ser lo recomendado en la literatura para terminadores de la transcripción de levaduras (Curran 2014), siendo un total de 157 nucleótidos (Fig. 11).



**Figura 11.** Representación esquemática de las regiones no transcritas río arriba y río abajo, 5'UTR, 3'UTR y CDS del gen *1033* indicando las posiciones respecto al genoma de *P. pastoris* CBS 7435 y el tamaño de las regiones del promotor y terminador del gen *1033*. nt: nucleótidos.

### 7.1.2. Identificación de motivos de secuencia característicos de regiones 5'UTR y 3'UTR de eucariotas y determinación del núcleo del P<sub>1033</sub>.

El análisis de la secuencia del P<sub>1033</sub> utilizando el programa UTRScan identificó un motivo de secuencia característico de regiones 5'UTR de eucariotas que es un SXL\_BS poli (dA-dT), el cual es una región común de muchos promotores constitutivos de levaduras. Con la herramienta bioinformática YAPP se predijo el núcleo del promotor determinando la posición de la caja tipo TATA (TAAAAAGA) y el elemento INR de tipo mamífero (CTAATCG), los cuales se ubicaron en -10 a -17 y +2 a + 8 nucleótidos del TSS, respectivamente (Fig. 12).

### 7.2. Identificación de factores de transcripción que se unen a la secuencia del P<sub>1033</sub>

Se identificaron 23 posibles sitios de unión de factores de transcripción (SUFT) en la secuencia del P<sub>1033</sub> (Tabla 26), se localizaron las secuencias de los factores de transcripción de *S. cerevisiae* en la base de datos de Uniprot que se unen a los SUFT identificados y con ellas se identificaron 6 proteínas homólogas en *P. pastoris* (Ecm22, Cat8, Rox1, Asg1, Yjl103c y Stp2) posibles factores de transcripción de *P. pastoris* que se unen a la secuencia del P<sub>1033</sub> y podrían participar en su regulación (Tabla 27). La

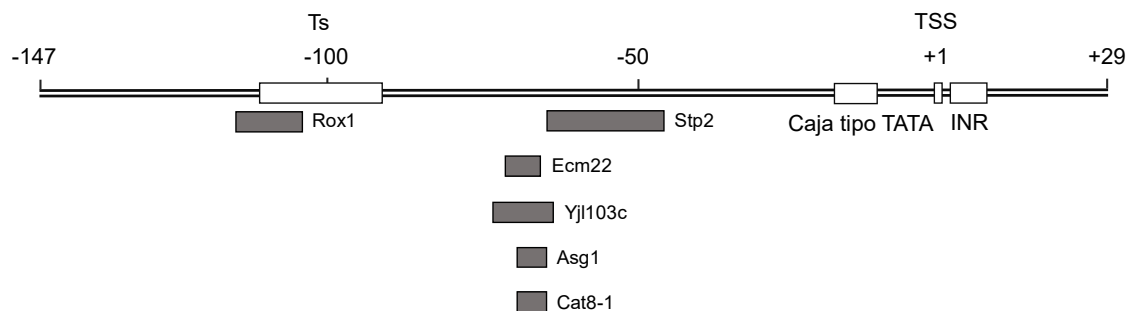
mayoría de los posibles factores de transcripción se unen en la región entre 76 y 88 nucleótidos río arriba de la caja tipo TATA del P<sub>1033</sub> (Fig. 12).

**Tabla 26.** Posibles sitios de unión de factores de transcripción (TFBS) de la secuencia nucleotídica del P<sub>1033</sub> de *P. pastoris* con un p < 0.05 y matriz de puntuación utilizada para su identificación.

<b>TFBS</b>	<b>p-value</b>	<b>ID de la matriz</b>
ECM22	0.0065	MA0292.1
F\$CAT8_01	0.0071	M01632
YER184C	0.0075	MA0424.1
F\$YER184C_01	0.0089	M01634
F\$ECM22_02	0.0120	M01639
F\$XBP1_Q2	0.0128	M00683
HSF1	0.0136	MA0319.1
LYS14	0.0144	MA0325.1
ROX1	0.0173	MA0371.1
F\$ECM22_01	0.0174	M01541
F\$YBR239C_01	0.0174	M01516
SIP4	0.0201	MA0380.1
YBR239C	0.0202	MA0420.1
ASG1	0.0245	MA0275.1
YJL103C	0.0252	MA0427.1
F\$LYS14_01	0.0255	M01576
F\$PDR1_01	0.0260	M01519
CAT8	0.0286	MA0280.1
STP2	0.0398	MA0395.1
F\$RDS1_01	0.0400	M01531
F\$SIP4_01	0.0402	M01517
F\$STP2_01	0.0437	M01560
F\$HSF_01	0.0476	M00029

**Tabla 27.** Posibles factores de transcripción de *P. pastoris* que interaccionan con la secuencia nucleotídica del P<sub>1033</sub>.

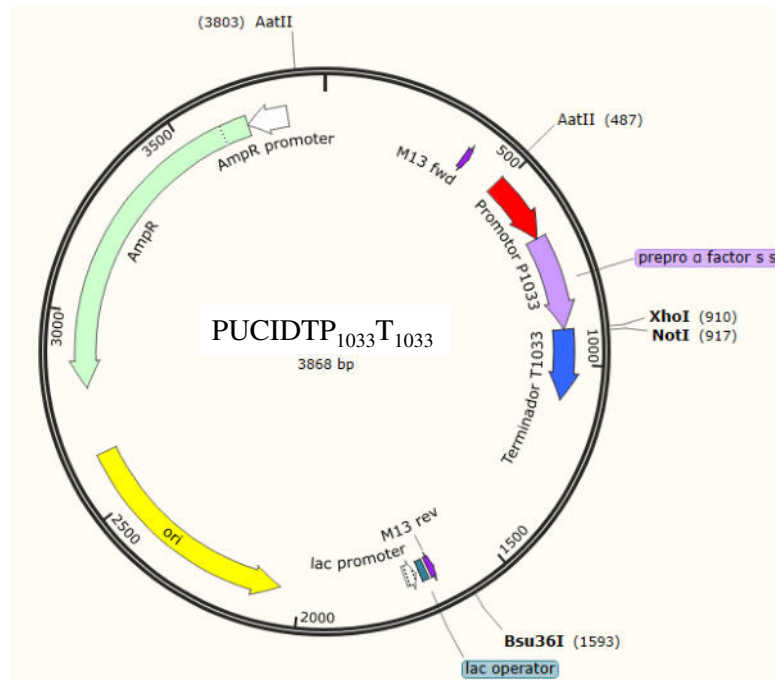
<b>Factor de transcripción</b>	<b>Clave UniProt</b>	<b>Clave GenBank en <i>P. pastoris</i></b>	<b>Clave PAS del factor de transcripción</b>
Ecm22	Q05958	XP_002490998	PAS_chr2-1_0114
Cat8-1	P39113	XP_002491690	PAS_chr2-1_0757
Rox1	P25042	XP_002490896	PAS_chr2-1_0017
Asg1	P40467	XP_002493608	PAS_chr4_0203
Yjl103C	P42950	XP_002489371	PAS_chr1-3_0036
Stp2	P38704	XP_002492619	PAS_chr3_0393



**Figura 12.** Representación esquemática de la secuencia  $P_{1033}$ . Se muestran la región homopolimérica de 21 nucleótidos de desoxitimidina (Ts), la caja tipo TATA, el sitio de inicio de la transcripción (TSS) y el elemento tipo iniciador de mamíferos (INR). El inicio de la transcripción se indica con +1. Debajo de la representación lineal de la secuencia  $P_{1033}$  se muestran en gris los factores de transcripción (TF) putativos identificados mediante el empleo consecutivo de la herramienta bioinformática TRAP, base de datos UniProt, y la técnica Best Reciprocal Blast Hit.

### 7.3. Diseño y síntesis de la secuencia sintética que incluyen las regiones de $P_{1033}$ y $T_{1033}$

Se diseñó una secuencia sintética que incluye la secuencia del promotor del gen *1033*, y la secuencia que codifica para la región pre-pro del factor alfa con codones preferenciales para *P. pastoris*, flanqueadas por los sitios de corte de las enzimas de restricción *AatII* y *XhoI*. A continuación, se añadieron 11 nucleótidos del vector pPIC9, seguido de la secuencia del terminador de la transcripción del gen *1033* y 500 nucleótidos del vector pPIC9, flanqueados por los sitios de corte de las enzimas de restricción *NotI* y *Bsu36I*. Esta secuencia con una longitud de 1116 pb (Fig. 13) fue sintetizada por la compañía Integrated DNA Technologies, Inc. (Coralville, IA, EUA) quien la proporcionó clonada en el vector pUCIDT-AMP, el cual posee la secuencia del gen de resistencia a ampicilina como marcador de selección en *E. coli*.



**Figura 13.** Secuencia sintética integrada en el plásmido PUCIDTP<sub>1033</sub>T<sub>1033</sub> que incluye la región promotora del gen P<sub>1033</sub> (rojo) desde el sitio *AatII* y la secuencia pre-pro del factor alfa (lila) con codones preferenciales para *P. pastoris* hasta el sitio *XhoI*, el sitio *NotI* seguido río abajo por la secuencia terminadora del gen P<sub>1033</sub> (azul), seguido de 500 nucleótidos del vector pPIC9 hasta el sitio *Bsu36I*.

Se diseñó un oligonucleótido 5' (PP1033) que hibrida en la secuencia del promotor del gen *1033* empleando la herramienta Primer3 v0.4.0 para utilizarlo en ensayos de PCR junto con el oligonucleótido FTE2 (disponible en el laboratorio) para caracterizar los vectores de expresión construidos y verificar la integración del casete de expresión en el genoma de *P. pastoris*. De las opciones de salida de este programa se seleccionó el oligonucleótido 5' (pP<sub>1033</sub>) con la secuencia CGGTGTCGCTACTTCGAGA, el cual tiene 19 pb y 57.89 % de GC.

El producto amplificado al utilizar los oligonucleótidos PP1033 y FTE2, que hibridan en el nucleótido 101 de la región promotora del gen *1033* y en el nucleótido 1457 de la secuencia del gen reportero respectivamente, tiene un tamaño de 1357 pb y corresponde a un fragmento de la secuencia del promotor P<sub>1033</sub>, secuencia pre-pro del factor alfa y un fragmento del CDS del gen *FTEII* (Fig. 14).

```

                start  len  tm  gc%  any  3'  seg
2 LEFT PRIMER    101  19  59.84  57.89  4.00  2.00  CGGTGTCGCTACTTCGAGA
  RIGHT PRIMER   1457  20  57.54  50.00  3.00  1.00  CTTGTGGGTGGAAACCAATC
  PRODUCT SIZE: 1357, PAIR ANY COMPL: 4.00, PAIR 3' COMPL: 2.00

```

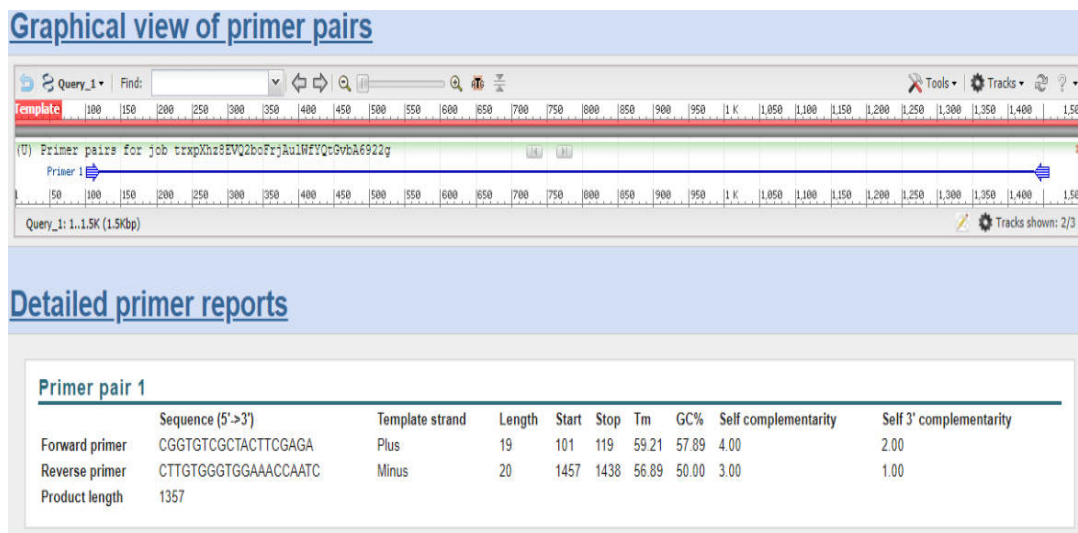
**Figura 14.** Resultado del programa Primer3 v0.4.0 al introducir la secuencia molde a amplificar y el par de oligonucleótidos PP1033 y FTE2.

La evaluación con la herramienta OligoAnalyzer 3.1 para para determinar la formación de estructuras secundaria al emplear el oligonucleótido PP1033 junto con el oligonucleótido FTE2 no mostró valores de  $\Delta G$  menores que -10 kcal/mol en ninguno de los oligonucleótidos (Tabla 28) por lo que se aseguró una baja probabilidad de formación de estructuras secundarias de tipo horquilla, homodímeros y heterodímeros. La temperatura de fusión ( $T_m$ ) de los oligonucleótidos fue de 63.9 y 61.8 °C para PP1033 y FTE2 respectivamente.

**Tabla 28.** Resultados del análisis de las secuencias del “primer forward” PP1033 y el “primer reverse” FTE2 empleando el programa de Oligoanalyzer.

Oligonucleótido	PP1033	FTE2
Secuencia	CGGTGTCGCTACTTCGAGA	CTTGTGGGTGGAAACCAATC
Horquilla (kcal/mol)	-0.81	-2.89
Homodímero (kcal/mol)	-6.76	-5.02
Heterodímero (kcal/mol)	-4.41	-4.41
Temperatura fusión (°C)	63.9	61.8

Para evaluar la especificidad de los oligonucleótidos PP1033 y FTE2, se realizaron alineamientos de los dos oligonucleótidos con la herramienta PrimerBlast buscando que no hibriden en alguna zona del genoma de *P. pastoris* (Fig. 15) para que no pudieran generar bandas inespecíficas en los ensayos de PCR. Estos oligonucleótidos solamente hibridaron en la región del casete de expresión integrado en el genoma de *P. pastoris* con un producto amplificado de 1357 pb.



**Figura 15.** Resultado del programa PrimerBlast al introducir la secuencia de los oligonucleótidos PP1033 y FTE2, y la secuencia del casete de expresión. Además, se seleccionó el genoma de *P. pastoris*.

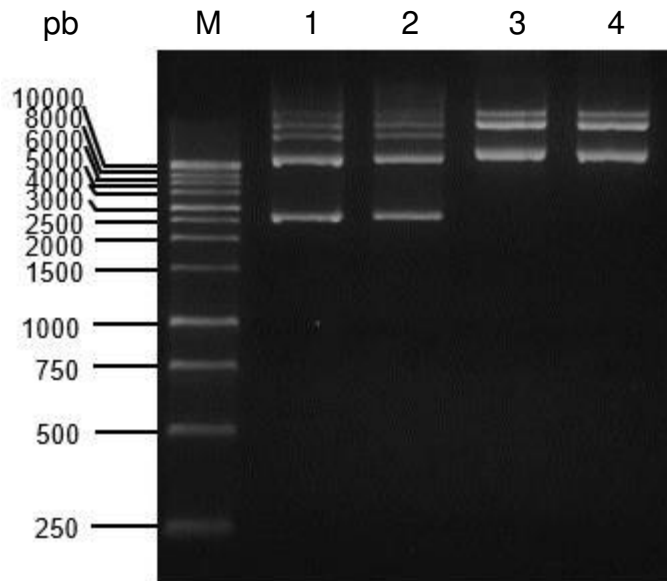
#### 7.4. Construcción de los vectores pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>

##### 7.4.1. Transformación de *E. coli* TOP10 con los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII

##### 7.4.2. Preparación y purificación de los vectores pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII

Los resultados de la preparación y purificación de los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII de células de *E. coli*., se muestran en la figura 16. El análisis en gel de agarosa para verificar la integridad de los DNA plasmídicos muestra las diversas conformaciones del DNA para cada plásmido, libre de RNA y sin señales de degradación, confirmando la integridad de ambos DNA plasmídicos aislados.

La cuantificación por espectrofotometría de ambos plásmidos resultó ser congruente con la intensidad de las bandas en la electroforesis. El DNA plasmídico pUCIDTP<sub>1033</sub>T<sub>1033</sub> presentó una concentración de 128 ng/μL y el DNA plasmídico pGAHFTEII obtuvo una concentración de 141 ng/μL. Para ambos plásmidos, la relación de absorbancia 260/280 nm estuvo entre 1.8 y 2.0.



**Figura 16.** Electroforesis en gel de agarosa al 0.8 % en TBE 1X de preparaciones de DNA plasmídico. M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carriles 1 y 2: DNA plasmídico pUCIDTP<sub>1033</sub>T<sub>1033</sub>. Carriles 3 y 4: DNA plasmídico pGAHFTEII.

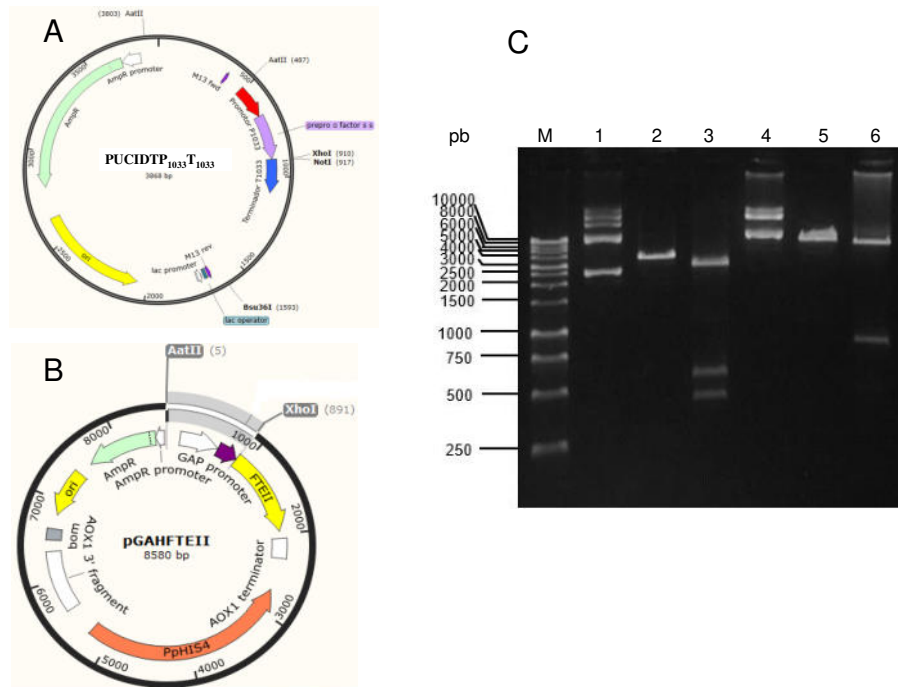
#### 7.4.3. Digestión de pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII con las enzimas *XhoI* y *AatII*

La digestión de los plásmidos pUCIDTP<sub>1033</sub>T<sub>1033</sub> (Fig.17A) y pGAHFTEII (Fig.17B) con las enzimas *XhoI* y *AatII* se realizó con el fin de obtener los fragmentos que integrarían el vector de expresión. Los resultados de las digestiones enzimáticas de ambos plásmidos se presentan en la figura 17, mostrando el análisis en gel de agarosa de los fragmentos generados del ensayo de digestión enzimática.

En el carril 2 se muestra una banda de 3868 pb correspondiente a pUCIDTP<sub>1033</sub>T<sub>1033</sub> linearizado digerido con *XhoI*, en el carril 3 se detectan 3 bandas de 2893, 552 y 423 pb al cortar el plásmido pUCIDTP<sub>1033</sub>T<sub>1033</sub> con las enzimas *XhoI* y *AatII*. La enzima *AatII* corta en el nucleótido 487 y en el 3804 de este plásmido generando dos bandas (Fig.17 C). El fragmento de 423 pb que va del nucleótido 488 al 910 es el que contiene la secuencia del promotor del gen P<sub>1033</sub> de *Pichia pastoris* de 178 pb y el factor alfa de *Saccharomyces cerevisiae* de 245 pb.

En el carril 5 se detecta el DNA plasmídico pGAHFTEII linearizado con *XhoI* (8580 pb) y en el carril 6 tratado con las enzimas *XhoI* y *AatII* mostrando las bandas esperadas de 7694 y 886 pb (Fig.17C).

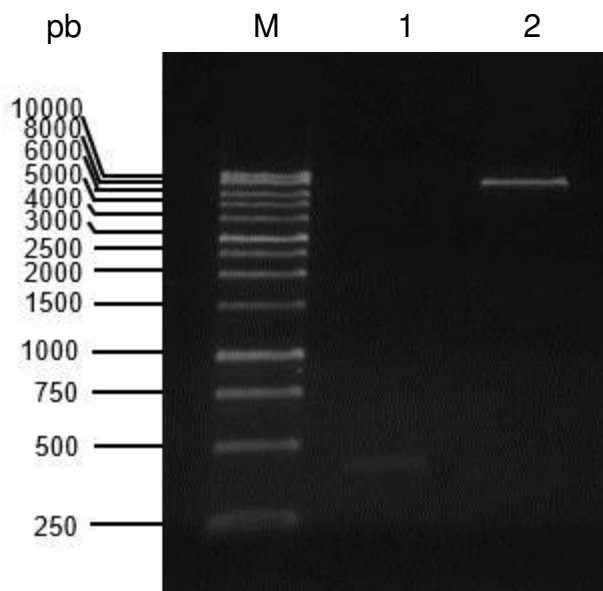




**Figura 17.** Electroforesis en gel de agarosa al 0.8 % (C) de la digestión enzimática de los DNA plasmídicos pUCIDTP<sub>1033</sub>T<sub>1033</sub> (A) y pGAHFTEII (B). M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: DNA plasmídico pUCIDTP<sub>1033</sub>T<sub>1033</sub>; Carril 2: DNA plasmídico pUCIDTP<sub>1033</sub>T<sub>1033</sub> digerido con *XhoI*; Carril 3: DNA plasmídico pUCIDTP<sub>1033</sub>T<sub>1033</sub> digerido con *XhoI* y *AatII*; Carril 4: DNA plasmídico pGAHFTEII; Carril 5: DNA plasmídico pGAHFTEII digerido con *XhoI*; Carril 6: DNA plasmídico pGAHFTEII digerido con *XhoI* y *AatII*.

#### 7.4.3.1 Purificación de pUCIDTP<sub>1033</sub>T<sub>1033</sub> y pGAHFTEII digeridos con *XhoI* y *AatII*

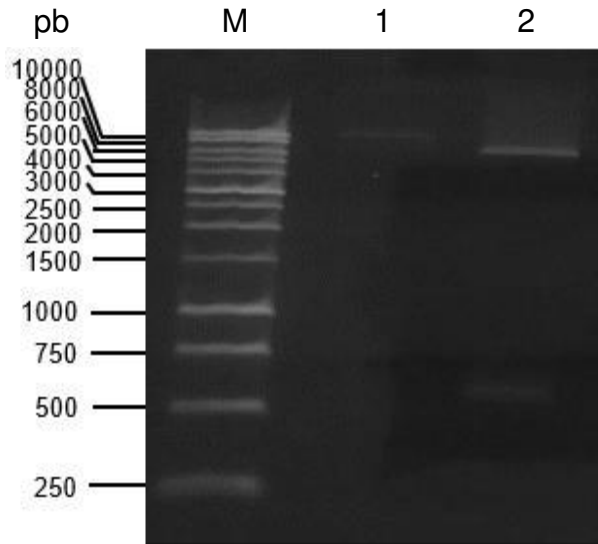
Para la obtención de los fragmentos de 423 pb liberado pUCIDTP<sub>1033</sub>T<sub>1033</sub> y el fragmento de 7694 pb liberado de pGAHFTEII, se realizó una digestión a escala preparativa con las enzimas *XhoI* y *AatII* de estos plásmidos. En la figura 18 se muestra la purificación de los fragmentos correspondientes al vector e inserto que se utilizaron en la reacción de ligación para la construcción del vector de expresión pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. De la purificación de la banda de 423 pb (carril 1) se obtuvieron 335 ng y de la banda de 7,694 pb (carril 2) 690 ng.



**Figura 18.** Electroforesis en gel de agarosa al 0.8 % en TBE 1X de los fragmentos purificados para llevar a cabo la ligación. M: Marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: fragmento de DNA inserto de 423 pb; Carril 2: fragmento de DNA vector de 7,694 pb.

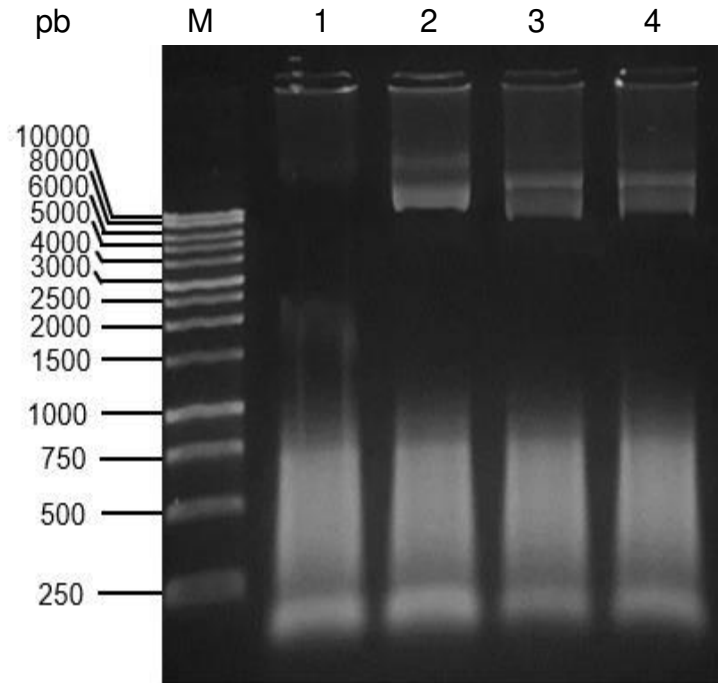
#### 7.4.4. Construcción y propagación del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>

Los resultados de la reacción de ligación de los fragmentos (vector-inserto) obtenidos de la digestión preparativa descritos en la sección 7. 4.3.1, se muestran en figura 19. En el carril 1 se muestra el producto de la ligación del vector y el inserto en el cual no se observaron bandas correspondientes al vector e inserto sin ligar, indicando que la ligación se llevó a cabo. En el carril 2 se muestra la reacción del control negativo de ligación (fragmentos a ligar sin ligasa) y en el cual no se detectan signos de degradación de los fragmentos utilizados.



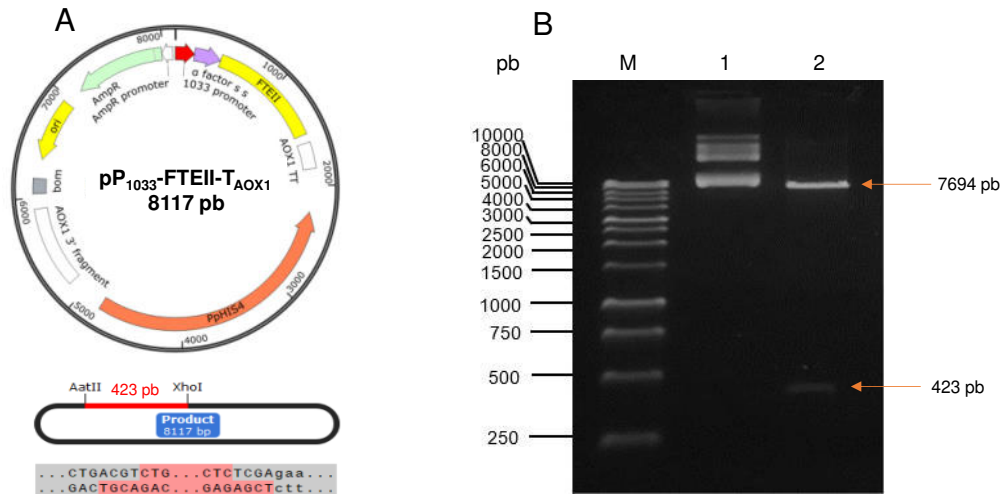
**Figura 19.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la ligación del fragmento AatII-*Xho*I del vector pGAHFTEII con el fragmento P<sub>1033</sub>-FTEII. M: marcador de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: vector e inserto con DNA T4 ligasa después de la incubación; Carril 2: vector e inserto (control negativo de ligación).

Con el producto de la reacción de ligación se transformaron células de *E. coli* TOP10, obteniendo más de 100 clonas que crecieron en las placas con ampicilina. A partir de cuatro de las colonias transformadas se preparó DNA plasmídico, cuyos resultados de análisis del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> obtenido se muestran en la figura 20. En el carril 1 se detectó la degradación del DNA plasmídico, en los carriles 2, 3 y 4 se detectaron bandas sin degradación con las distintas conformaciones características de los plásmidos, lo que confirmó la integridad del DNA plasmídico aislado. En todos los carriles se detectó la presencia de RNA degradado en las preparaciones.



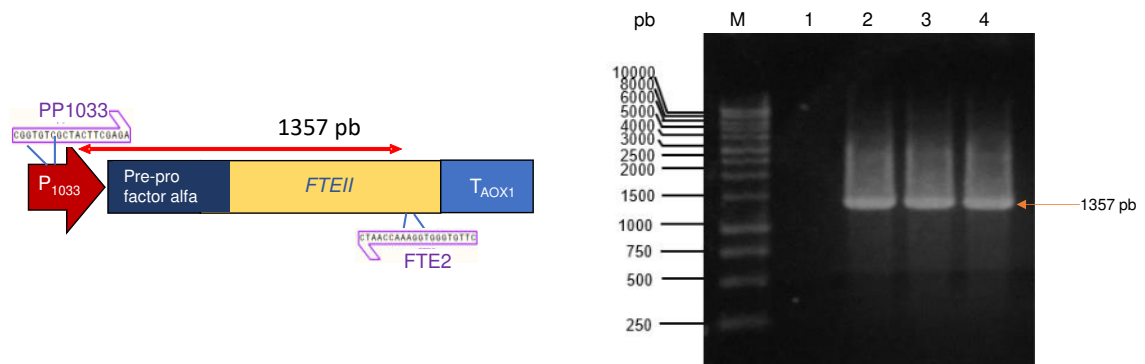
**Figura 20.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carriles 1-5: DNA plasmídico de cuatro colonias obtenidas de la transformación de E. coli TOP10 con el producto de ligación.

La caracterización enzimática del plásmido construido se muestra en la figura 21, en la cual se observa el análisis de la digestión del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con las enzimas *XhoI* y *AatII*. En el carril 1 se muestra el plásmido sin digerir y en el carril 2 se muestra la digestión de pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (previamente tratado con RNAsa) con las dos enzimas (*XhoI* y *AatII*), lo cual generó dos fragmentos, uno de 7,694 pb correspondiente al vector y otro de 423 pb correspondiente al inserto.



**Figura 21.** Electroforesis en gel de agarosa (B) al 0.8% en TBE 1X de la caracterización enzimática del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con *XhoI* y *AatII* (A). M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>; Carril 2: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> digerido con *XhoI* y *AatII*.

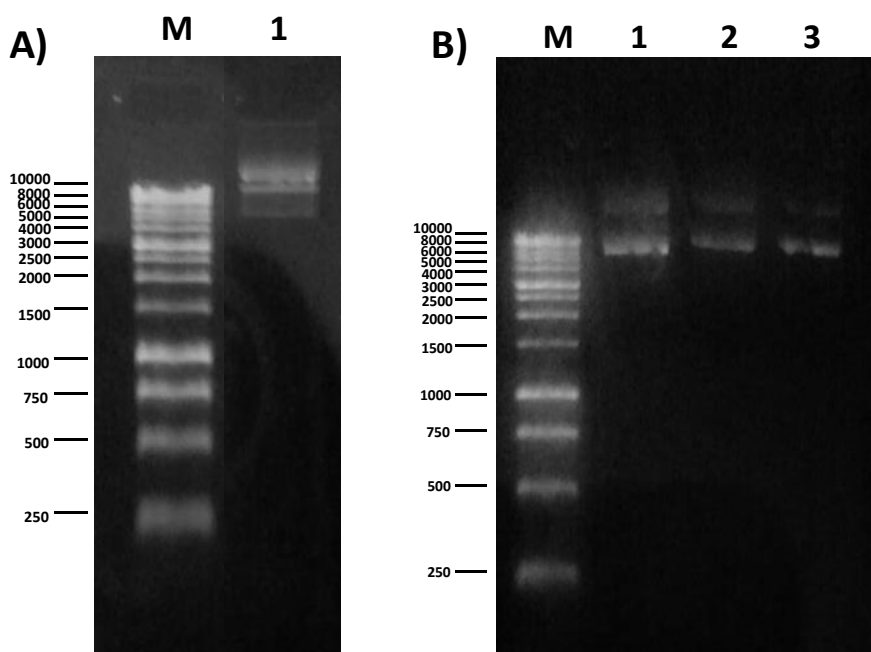
El plásmido pP1033-FTEII-T<sub>AOX1</sub> construido fue también caracterizado por PCR. En la figura 22 se muestra el resultado del PCR con los oligonucleótidos PP1033 y FTE2 del plásmido aislado de tres clonas de *E. coli* transformadas. El PCR del plásmido mostró la banda esperada de 1,357 pb que es el tamaño del amplicón de la secuencia construida con los oligonucleótidos utilizados.



**Figura 22.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> con los iniciadores PP1033 y FTE2. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control (-) de la PCR. Carriles 2-4: Producto amplificado del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> de tres colonias de *E. coli* trasformadas.

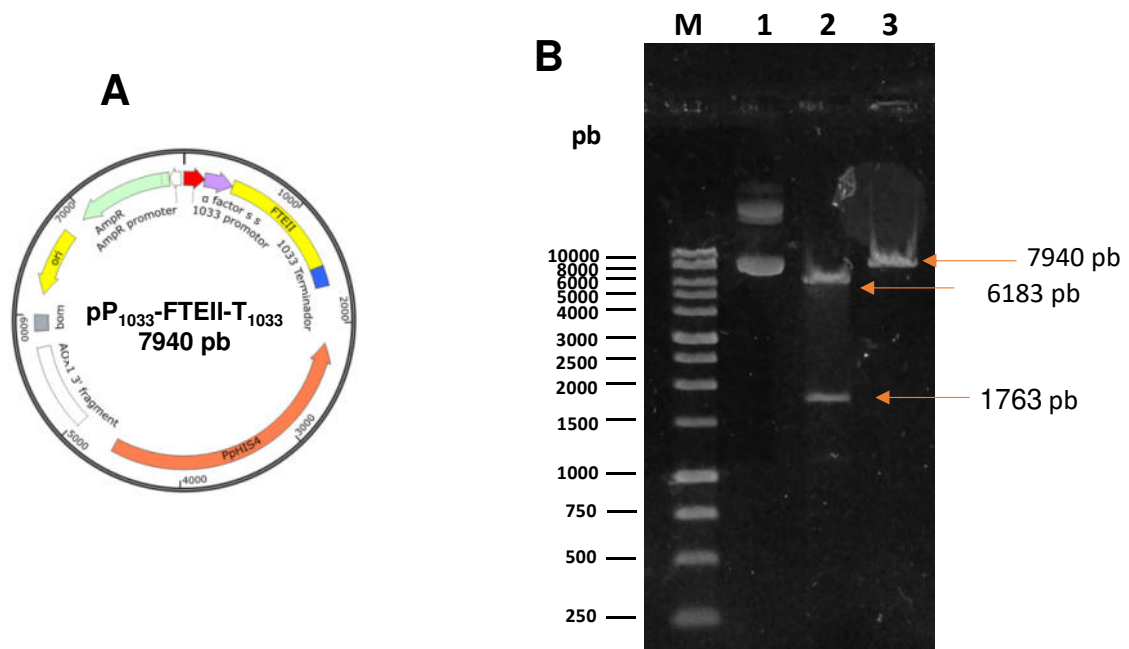
#### 7.4.5. Construcción y propagación del plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub>

El casete de expresión integrado en el vector pPIC9 y sintetizado de la compañía Gene Universal, se usó para transformar cepas de *Escherichia coli* TOP10, en donde el plásmido fue propagado y posteriormente purificado. En la figura 23A se muestra el análisis en gel de agarosa que comprueba la integridad del DNA obtenido de la compañía. En el carril 1 de la figura 23A, se observa que el DNA no presenta degradación y que posee buena calidad para su uso en la transformación de células de *E. coli* TOP10. Adicionalmente, el DNA presentó una concentración de 110 ng/μL y una relación de 260/280 de 1.9. En los tres carriles de la figura 23B, el gel de agarosa muestra la integridad del DNA del plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub> obtenido de la extracción de las colonias transformadas de *E. coli* TOP10. La concentración de DNA osciló entre 26.0 y 48.0 ng/μL y los valores de la relación 260/280 entre 1.811 y 1.870.



**Figura 23.** Electroforesis en geles de agarosa al 0.8% del DNA plasmídico A) enviado por la compañía Gene Universal. M: Marcador de tamaño molecular DNA ladder 1kb de Promega. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> y B) obtenido de la extracción de DNA plasmídico de colonias transformadas de *E. coli* TOP10. M: Marcador de tamaño molecular DNA ladder 1kb de Promega. Carril 1-3: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> obtenido de *E. coli* TOP10 transformada.

La caracterización enzimática pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con las enzimas *Xho*I y *Bsu*36I se muestra en la figura 24. En el carril 2, se observan las bandas del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> generadas de 6,183 y 1,763 pb. El fragmento más pequeño corresponde al gen *FTEII*, la región terminadora y una porción del gen *HIS4* del vector. El fragmento de 6,183 pb corresponde al resto del vector. En el carril 3 se observa una única banda de 7940 pb, que corresponde con el tamaño esperado del plásmido, por lo que el plásmido corresponde con las especificaciones dadas a la compañía y se utilizó para el resto de los experimentos.

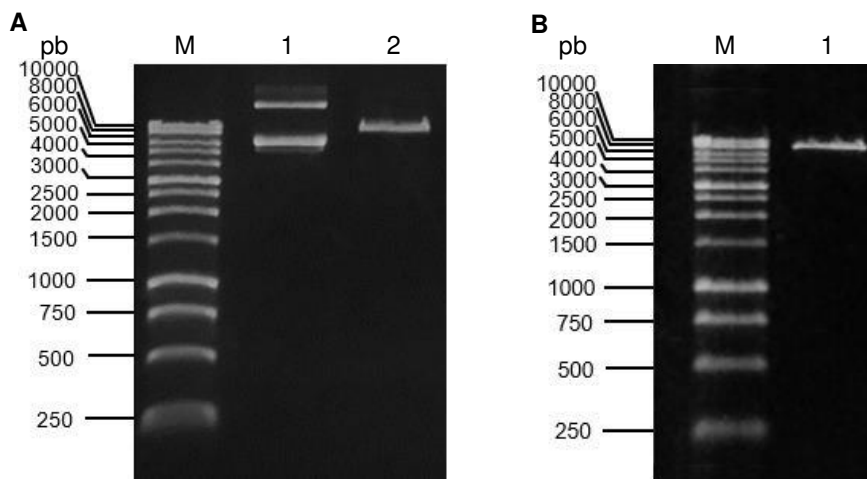


**Figura 24.** Electroforesis en gels de agarosa (B) al 0.8% de la caracterización de pP<sub>1033</sub>-FTEII-T<sub>1033</sub> (A) con las enzimas de restricción *Xho*I y *Bsu*36I. A. M: Marcador de tamaño molecular DNA ladder 1kb de Promega. Carril 1: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> sin enzimas. Carril 2: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con *Xho*I y *Bsu*36I. Carril 3: pP<sub>1033</sub>-FTEII-T<sub>1033</sub> con *Xho*I.

## 7.5. Construcción de las cepas recombinantes de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>

### 7.5.1. Extracción y purificación de los plásmidos pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> y digestión con *Sal* I.

Los plásmidos se obtuvieron de mediana escala y la digestión preparativa de los plásmidos pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> se realizó con la enzima *Sal*I para linealizarlos. En la figura 25 se muestra el análisis en gel de agarosa para verificar la integridad de ambos DNA plasmídicos y la correcta digestión con la enzima. En la figura 25A, carril 1 se observan las diversas conformaciones del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> sin presencia de RNA, confirmando la integridad del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> aislado. En el carril 2 se muestra el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> linearizado, con una banda bien definida de 8117 pb que corresponde al tamaño de este vector y demuestra la digestión completa del plásmido. En la figura 25B, se muestra el plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub> linearizado, la obtención de una banda bien definida de 7946 pb que corresponde al tamaño de este vector demostró el corte total del plásmido.



**Figura 25.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de preparaciones de DNA plasmídico a mediana escala. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). (A) pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> digerido con la enzima *Sal*I. (B) pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Carril 1: DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub> digerido con la enzima *Sal*I.



### 7.5.2. Transformación de *P. pastoris* con pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub>

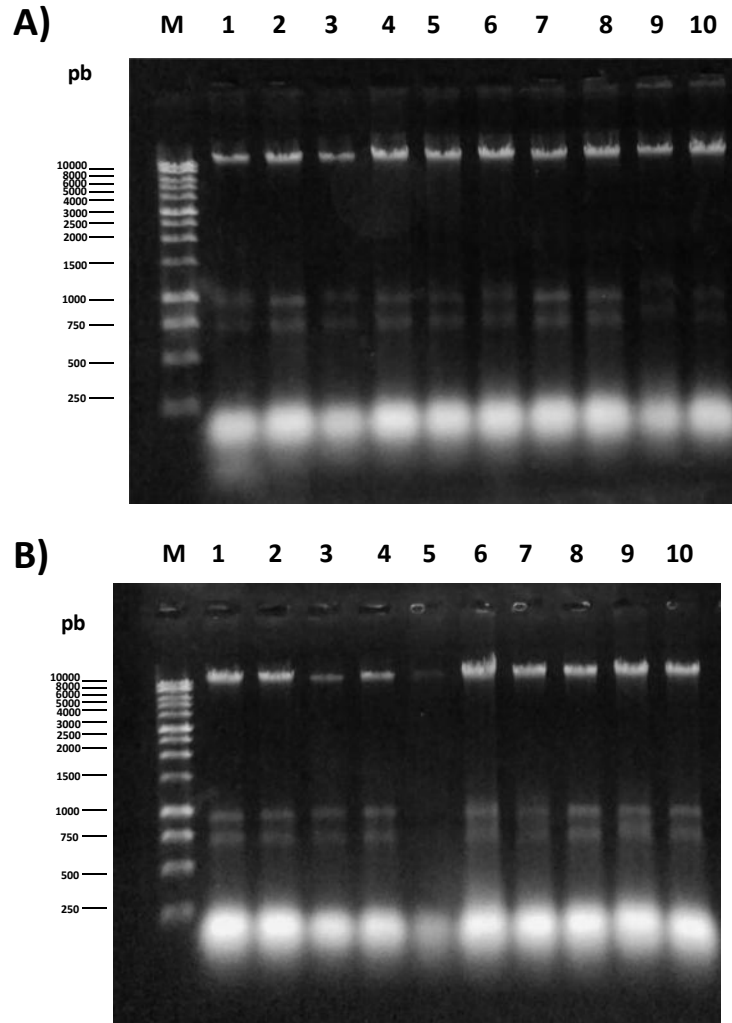
Con ambos plásmidos linearizados se transformaron células electrocompetentes de la cepa KM71 de *P. pastoris*, obteniéndose un total 1,990 clonas His<sup>+</sup> para el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (Tabla 29) y más de 1,000 clonas His<sup>+</sup> para el plásmido pP<sub>1033</sub>-FTEII-T<sub>1033</sub> distribuidas en cuatro placas de medio selectivo RDB His<sup>-</sup> (placas A, B y placas 2 y 3), todas las colonias con características morfológicas de *P. pastoris*.

**Tabla 29.** Resultados de la transformación de células de *Pichia pastoris* KM71 con el plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>.

Volumen de células sembrado (μL)	UFC
50	198
100	285
150	412
200	485
500	610

### 7.5.3. Caracterización genotípica de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>

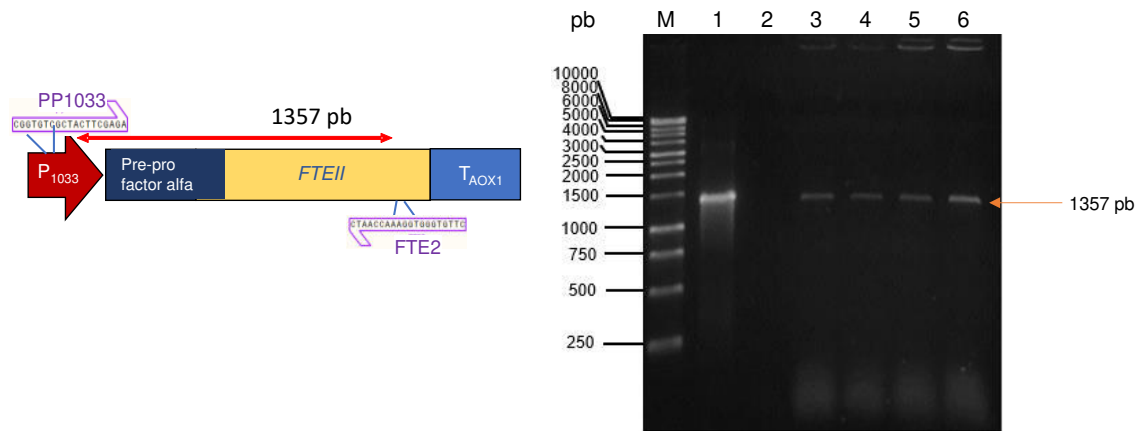
Para la caracterización genotípica se aisló el DNA genómico de 4 colonias His<sup>+</sup> de la cepa pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y 20 colonias de la cepa pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. El análisis en gel de agarosa para verificar la integridad de los DNA genómicos aislados de 20 colonias de *P. pastoris* cepa KM71 transformadas con el vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub> linealizado se muestra en la figura 26. La obtención de bandas intensas y definidas demostró la integridad de todos los DNA genómicos aislados.



**Figura 26.** Electroforesis en geles de agarosa al 0.8% en TBE 1X de DNA genómico aislado de cepas de *P. pastoris* KM71 transformadas con el vector pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. A) M: marcador de tamaño molecular 1 kb DNA Ladder. Carriles 1-5: DNA genómico de la placa A (colonias 1 a 5). Carriles 6-10: DNA genómico de la placa B (colonias 1 a 5). B) M: marcador de tamaño molecular 1 kb DNA Ladder. Carriles 1-5: DNA genómico de la placa 2 (colonias 1 a 5). Carriles 6-10: DNA genómico de la placa 3 (colonias 1 a 5).

A partir del DNA genómico se realizaron amplificaciones usando oligonucleótidos dirigidos hacia el casete de expresión integrado en los genomas de las cepas construidas. La caracterización genotípica de las colonias transformadas se llevó a cabo mediante la confirmación por PCR de la presencia del casete de expresión empleando la combinación de los oligonucleótidos específicos PP<sub>1033</sub> y FTE2 para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y los oligonucleótidos FTE1 y 3TH para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>.

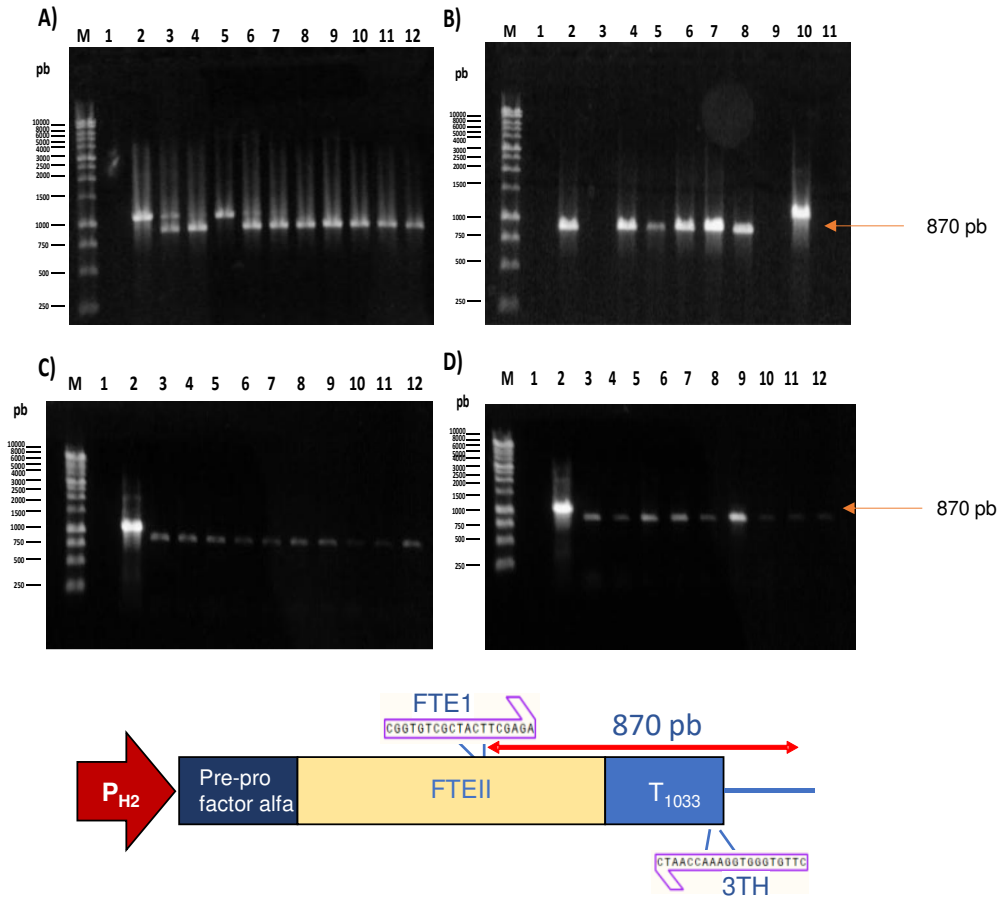
La figura 27 muestra el análisis en gel de agarosa de los productos del ensayo de PCR de DNA genómicos aislados de 4 colonias transformadas con el vector pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. En el gel se presentan bandas intensas y definidas con el tamaño esperado de 1,357 pb, lo que demostró la integración del casete de expresión, el vector pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>, en el genoma de *P. pastoris*.



**Figura 27.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR con los oligonucleótidos PP1033 y FTE2 del DNA genómico aislado de colonias de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: control negativo de la PCR. Carriles 3-6: Producto amplificado del DNA genómico de 4 colonias transformantes de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>.

En la figura 28 se muestra el análisis de electroforesis de los productos amplificados obtenidos en el ensayo de PCR de las clonas de las cepas KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. La presencia de una única banda de 870 pb confirmó la integración del casete de expresión en el genoma de la levadura.

De las 40 clonas analizadas, 33 dieron positivo al PCR, 3 clonas no presentaron la banda esperada y 4 presentaron una banda dando un resultado inconcluso.



**Figura 28.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR con los oligonucleótidos FTE1 y 3TH del DNA genómico aislado de las colonias de la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub> A). M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: DNA genómico *P. pastoris* de la placa A (1-5). Carriles 8-12: DNA genómico *P. pastoris* de la placa B (1-5). B) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1-5: DNA genómico *P. pastoris* de la placa 2 (1-5). Carriles 6-9: DNA genómico *P. pastoris* de la placa 3 (1-4). Carril 10: control positivo de PCR (pGAHFTEII). Carril 11: Control negativo de PCR. C) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: PCR de colonia de la placa A (6-10). PCR de colonia de la placa B (6-10). D) M: marcador de tamaño molecular 1kb DNA Ladder. Carril 1: control negativo de PCR. Carril 2: control positivo de PCR (pGAHFTEII). Carriles 3-7: PCR de colonia de la placa 2 (6-10). PCR de colonia de la placa 3 (7-10).

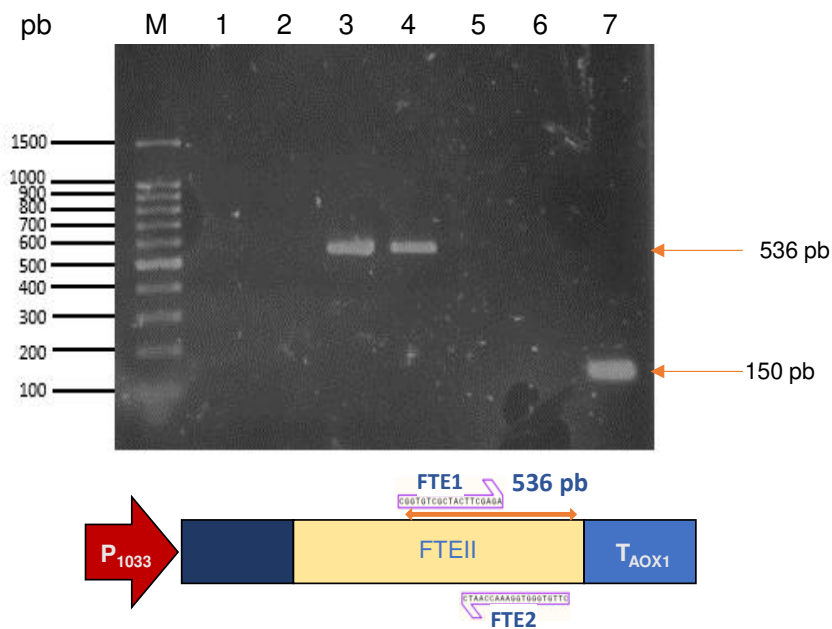
## 7.6. Caracterización funcional de los elementos reguladores integrados en las cepas recombinantes construidas.

### 7.6.1. Detección de transcritos por RT-PCR

La detección de transcritos del gen reportero por RT-PCR se empleó para determinar la funcionalidad de los elementos reguladores, promotor y terminador de la transcripción

integrados en el casete de expresión y a su vez en el genoma de *P. pastoris*. A partir de cultivos de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71P<sub>1033</sub>-FTEII-T<sub>1033</sub> se tomaron muestras en la fase exponencial de crecimiento a las cuales se les extrajo el RNA total que se utilizó para sintetizar DNAc y finalmente se amplificó por PCR.

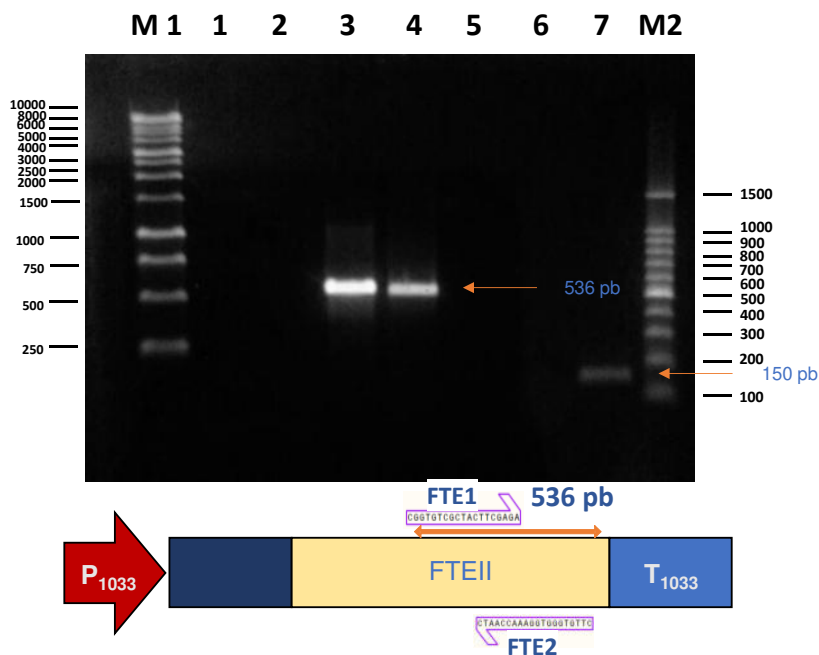
En la figura 29 se muestra el análisis en gel de agarosa para la detección del producto amplificado de la reacción de RT-PCR. Los controles negativos de la reacción de PCR y de la retrotranscripción (carriles 1 y 2) no muestran ninguna banda por lo que no hubo contaminación en estas reacciones. En los carriles 3 y 4 se detecta una banda entre 500 y 600 pb que corresponde con el producto amplificado esperado de 536 pb correspondiente a un fragmento de *FTEII* al emplear los oligonucleótidos FTE1 y FTE2 que hibridan en el nucleótido 922 y 1457 de la secuencia codificante del gen *FTEII*. En el carril 7 se observa una banda entre 100 y 150 pb, que corresponde con el producto amplificado esperado de 150 pb al emplear los oligonucleótidos 5qACT y 3qACT que hibridan en el CDS del gen de la  $\beta$ -actina.



**Figura 29.** Electroforesis en gel de agarosa al 2.0% del producto amplificado por RT-PCR a partir de RNA de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. M: Marcador de tamaño molecular de 100 pb DNA Ladder (Promega, Madison, WI, EUA). Carril 1: control negativo de la reacción de PCR, coctel 1; Carril 2: control negativo de la retrotranscripción, coctel 1; Carril 3: producto amplificado del DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (control positivo de la reacción); Carril 4: producto amplificado de la muestra con los oligonucleótidos FTE1 y FTE2, coctel 1; Carril 5: control negativo de la reacción de PCR, coctel 2; Carril 6: control negativo de la retrotranscripción, coctel 2; Carril 7: producto amplificado de la muestra con los oligonucleótidos 5qACT y 3qACT, coctel 2.

Este resultado demuestra que el P<sub>1033</sub> es funcional porque se evidencia la presencia de transcritos a las 10 horas de cultivo de una de las clonas transformadas que incorporó la secuencia del plásmido pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> en el genoma de *P. pastoris* KM71.

En la figura 30, se muestran el análisis en gel de agarosa de los productos amplificados por la PCR utilizando el DNAc sintetizado a partir de RNA total obtenido a las 12 h de cultivo de KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Los carriles 4 y 7 presentaron las bandas esperadas de 536 y 150 pb correspondientes a los transcritos de *FTEII* y  $\beta$ -actina respectivamente, demostrando que los elementos reguladores son funcionales, además la ausencia de amplificación en los controles negativos del RT en los carriles 2 y 6 indica que no hubo contaminación con DNA genómico, por lo que las bandas obtenidas provienen del DNAc sintetizado del RNA obtenido a partir de las células del cultivo.



**Figura 30.** Electroforesis en gel de agarosa al 0.8% en TBE 1X del RT-PCR a partir del RNA total a las 12 h de cultivo de células de *P. pastoris* KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. M1: marcador de tamaño molecular 1kb DNA ladder. Carril 1: control negativo PCR. Carril 2: control negativo RT. Carril 3: control positivo PCR (pP<sub>1033</sub>-FTEII-T<sub>1033</sub>). Carril 4: producto amplificado de la muestra con los oligonucleótidos FTE1 y FTE2. Carril 5: control negativo PCR ( $\beta$ -actina). Carril 6: control negativo RT. Carril 7: producto amplificado de la muestra con los oligonucleótidos 5qACT y 3qACT. M2: marcador de tamaño molecular de 100 pb.

## **7.6.2 Detección del producto proteico del gen reportero (fitasa)**

### **7.6.2.1. Determinación de actividad volumétrica de fitasa extracelular**

Los cultivos de 3 clonas de cada cepa construida (KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>) que se realizaron en las condiciones que se describen en el acápite 6.10.1.1, alimentando con glicerol al 1 % p/v a las 14 horas, alcanzaron una DO<sub>600</sub> en la fase de estacionaria a las 24 horas de cultivo de 7.70 a 8.31.

Los sobrenadantes libres de células de las muestras tomadas a las 24 horas de cultivo que se desalaron utilizando cromatografía de exclusión molecular (la muestra se diluye 2.5 veces al realizar el cambio de solución tampón) mostraron producción de FTEII en el medio de cultivo, pues se obtuvieron actividades volumétricas de fitasa extracelular de  $0.23 \pm 0.01$  U/mL para la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y  $0.11 \pm 0.01$  U/mL para la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Los valores obtenidos de las determinaciones de concentración de proteínas extracelulares a las 24 h de cultivo de las 3 clonas de las cepas KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub> fueron de  $9.06 \pm 0.03$  µg/mL y  $7.03 \pm 0.02$  µg/mL respectivamente. Estos resultados demuestran que ambas cepas tienen crecimiento celular y además producen la proteína recombinante y la secretan al medio de cultivo

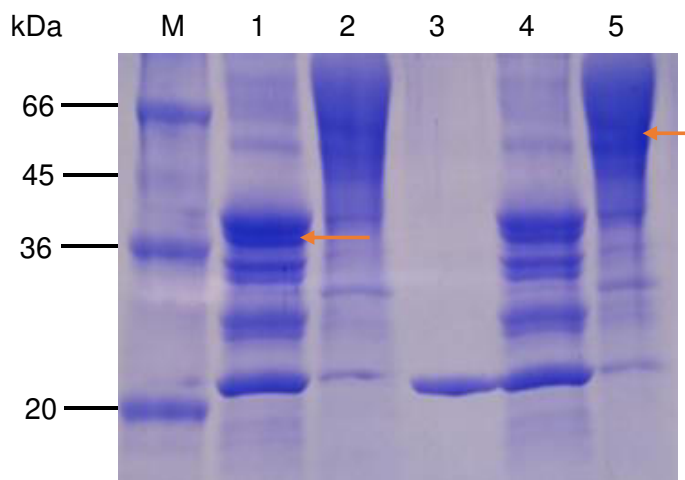
### **7.6.2.2. Detección de fitasa extracelular por SDS-PAGE**

La detección del producto proteico del gen reportero se empleó como evidencia de la actividad transcripcional y los procesos celulares que conllevan a la expresión de la proteína. Para la detección de la fitasa FTEII, se realizaron cultivos en BMG-CaCl<sub>2</sub>, de una colonia seleccionada de cada cepa recombinante construida, tal como se describe en la sección 6.10.2.1, tomando una muestra a las 24 horas de cada cultivo, las cuales se concentraron y desalaron por ultrafiltración. Cada muestra se concentró 100 veces. Estos cultivos mostraron DO<sub>600</sub> en la fase de estacionaria a las 24 horas de cultivo de 8.02 y 8.13 para las cepas KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub> respectivamente. Los sobrenadantes libres de células de las muestras tomadas a las 24 horas concentradas y desaladas tuvieron actividades volumétricas de fitasa extracelular de 0.25 U/mL para la

cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y 0.14 U/mL para la cepa KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Los valores obtenidos de las determinaciones de concentración de proteínas extracelulares a las 24 h de cultivo de las cepas KM71/pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/pP<sub>1033</sub>-FTEII-T<sub>1033</sub> fueron de 12.50 µg/mL y 8.52 µg/mL respectivamente.

En la figura 31 se muestran los resultados del análisis en SDS-poliacrilamida de muestras del sobrenadante de los cultivos de ambas cepas en los cuales se detectan el barrido característico de la fitasa FTEII, alrededor de 45 kDa (carriles 2 y 5). En la misma figura, en los carriles 1 y 4, se detecta una banda de 39 kDa, de las muestras tratadas con Endo H N-desglicosilasa. En estos carriles se aprecia un cambio en el patrón electroforético desapareciendo el barrido característico de la fitasa FTEII, alrededor de 45 kDa y en su lugar se detecta un patrón de bandas definidas, una de ellas a la altura de 39 kDa correspondiente a la masa molecular de FTEII desglicosilada. Este resultado correspondió a la masa molecular teórica de la fitasa FTEII en función de su secuencia de aminoácidos (Viader-Salvadó et al., 2010).

Estos resultados (detección de transcripción de *FTEII*, detección de fitasa FTEII por SDS-PAGE y actividad enzimática) confirmaron la funcionalidad correcta de las secuencias reguladoras y sus combinaciones (es decir, P<sub>1033</sub>-T<sub>AOX1</sub> y P<sub>1033</sub>-T<sub>1033</sub>).



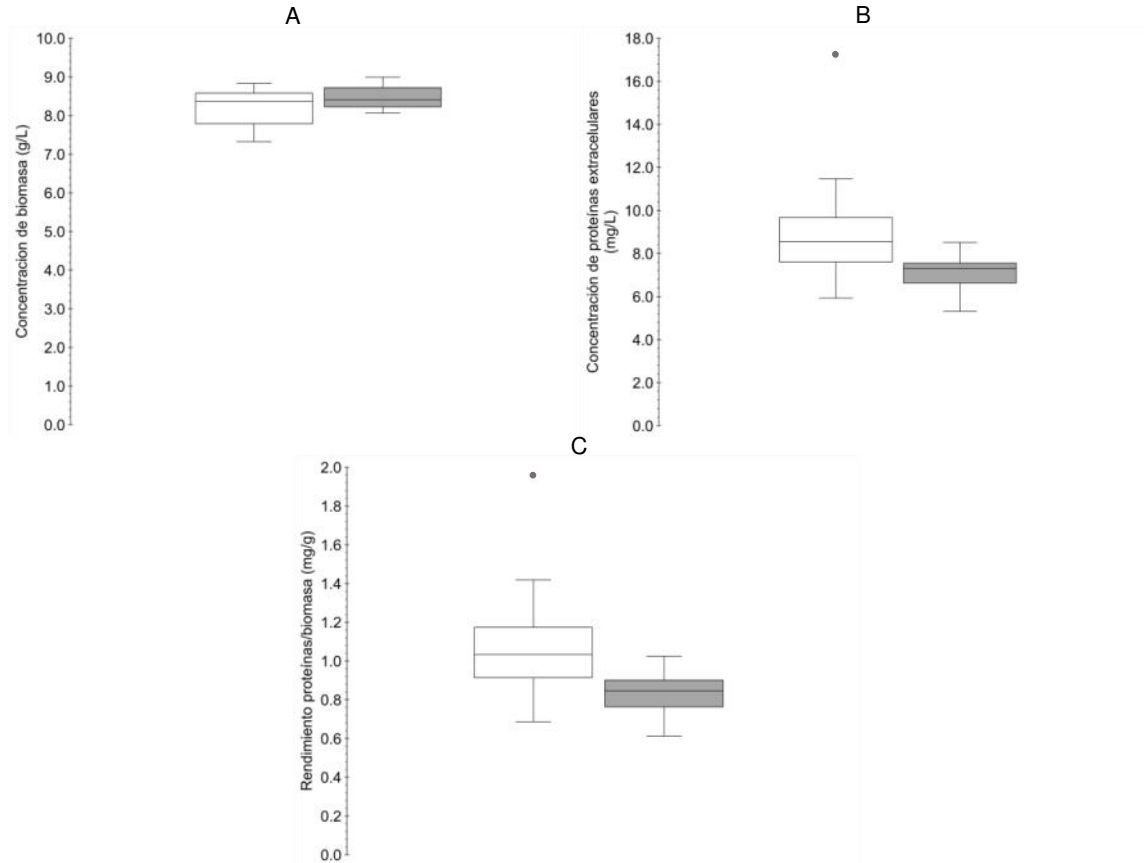
**Figura 31.** Gel de SDS-poliacrilamida de sobrenadantes concentrados (SC) de cultivo por ultrafiltración y tratados con o sin Endo H. Carril M: marcador de tamaño molecular. Carril 1: SC KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> tratado con Endo H. Carril 2: SC KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> sin Endo H. Carril 3: Endo H. Carril 4: SC KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> tratado con Endo H. Carril 5: SC KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> sin Endo H.



## **7.7. Evaluación de la funcionalidad de la combinación de los elementos reguladores P<sub>1033</sub>-T<sub>AOX1</sub> y P<sub>1033</sub>-T<sub>1033</sub> en cepas unicopias cultivadas en dos fuentes de carbono**

### **7.7.1. Determinación de proteínas extracelulares y rendimiento proteína/biomasa**

Se realizaron cultivos en matraz agitado con medio BMG de 32 clonas de *P. pastoris* KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y 20 clonas de KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> a los cuales se les determinó la densidad celular ó concentración de biomasa a las 24 horas del cultivo. Las densidades celulares, en gramos de células en peso seco por litro de cultivo, estuvieron comprendidas entre los valores de 7.33 y 8.83 para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> mostraron valores en el intervalo de 8.06 a 8.99 g/L (Fig. 32 A). Las 32 clonas de la cepa KM71P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> tamizadas mostraron concentraciones de proteínas extracelulares entre 5.93 y 17.20 µg/mL a las 24 horas de cultivo, por lo que la mayoría deben ser posibles clonas unicopia (con una sola copia del gen heterólogo *FTEII* insertada en el genoma de la levadura), solo una clona resultó ser un valor atípico o un punto fuera de la caja (clona 14) que la consideramos posible clona multicopia. Las 20 clonas de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> presentaron concentraciones en un intervalo entre 5.31 y 8.52 µg/mL a las 24 horas de cultivo, por lo que la mayoría deben ser posibles clonas unicopia del gen *FTEII* (Fig. 32 B). El rendimiento de proteínas extracelulares/biomasa para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> presentó valores entre 0.69 y 1.49 mg/g también con una clona que resultó ser un valor atípico (clona 14) y para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> el rendimiento de proteínas extracelulares/biomasa estuvo en un intervalo de 0.61 a 1.03 mg/g (Fig. 32 C).



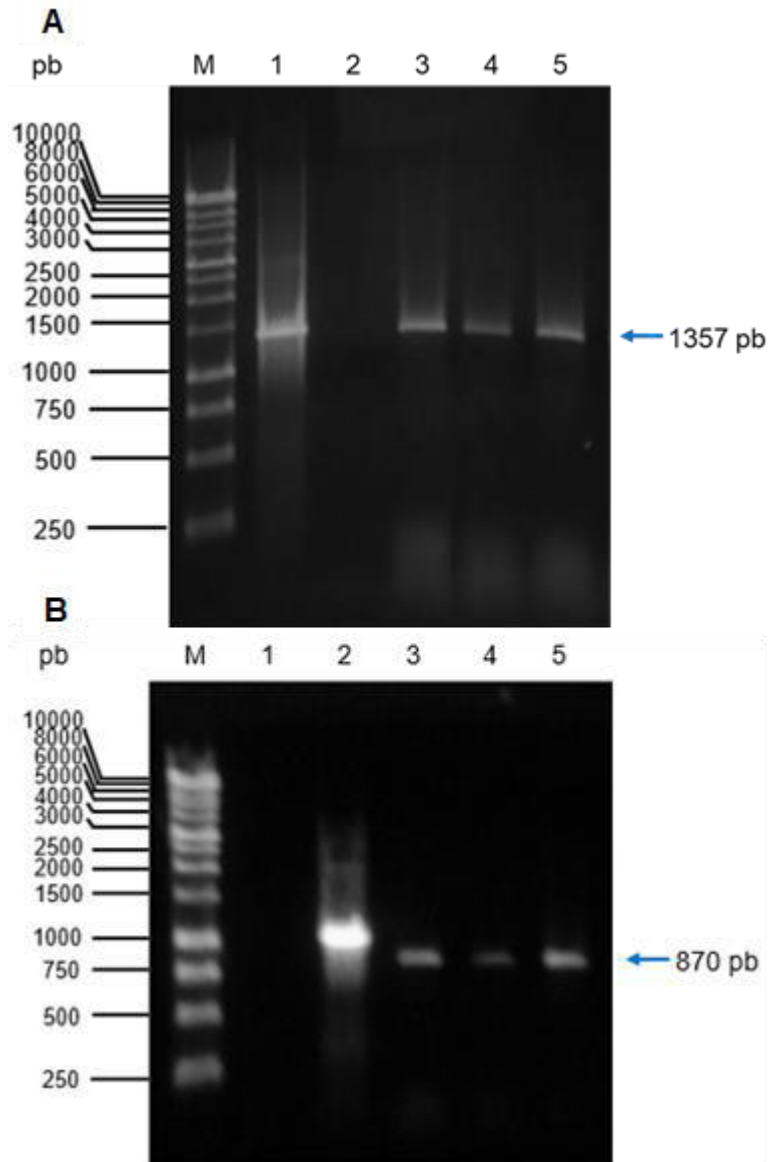
**Figura 32.** Representación de diagrama de caja de la concentración de biomasa (A) la concentración de proteína extracelular (B) y el rendimiento de proteína/biomasa extracelular (C) para cultivos BMG de 24 h de 32 y 20 clones recombinantes (transformantes His+) de cada cepa (recuadro blanco, KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>; caja gris, KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>). Las líneas centrales muestran las medianas y los límites de las cajas indican el 25 y el 75 percentil. Los bigotes se extienden 1.5 veces del rango intercuartil desde los percentiles 25 y 75 y los valores atípicos se representan con puntos.

### 7.7.2. Identificación de clonas unicopia de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>

Para la selección de la probable clona unicopia de cada cepa se consideraron las clonas que produjeron los rendimientos de producción de proteína/biomasa más bajos. Las clonas seleccionadas se cultivaron y se les aisló el DNA y luego se les realizó el análisis de qPCR del DNA de la cepa seleccionada de cada construcción y se confirmó la presencia de una copia del gen heterólogo en el genoma de la levadura.

### **7.7.2.1. Cultivos y preparación de DNA genómico de las posibles clonas unicopias**

Para la selección de la probable clona unicopia de cada cepa se consideraron las clonas que produjeron los rendimientos de proteína/biomasa más bajos, seleccionando 3 clonas de cada construcción, a las cuales se les determinó la integración del casete de expresión mediante PCR, obteniendo bandas intensas y definidas con un tamaño de 1357 pb para las 3 clonas de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>Aox1</sub> y 870 pb para la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>, lo que demostró la integración del casete de expresión en el genoma de ambas cepas de *P. pastoris*. (Fig. 33). Las 6 clonas seleccionadas se cultivaron en medio YPD hasta alcanzar DO a 600 nm de 6 a 8 y a partir de las células se extrajo ADN genómico por el método de TSNT y con posterior tratamiento con RNAsa y purificación del DNA genómico realizando lavados con agua libre de nucleasas en los dispositivos de filtración Amicon® Ultra-0.5, obteniendo preparaciones de buena calidad tal como se muestra en la Tabla 30.



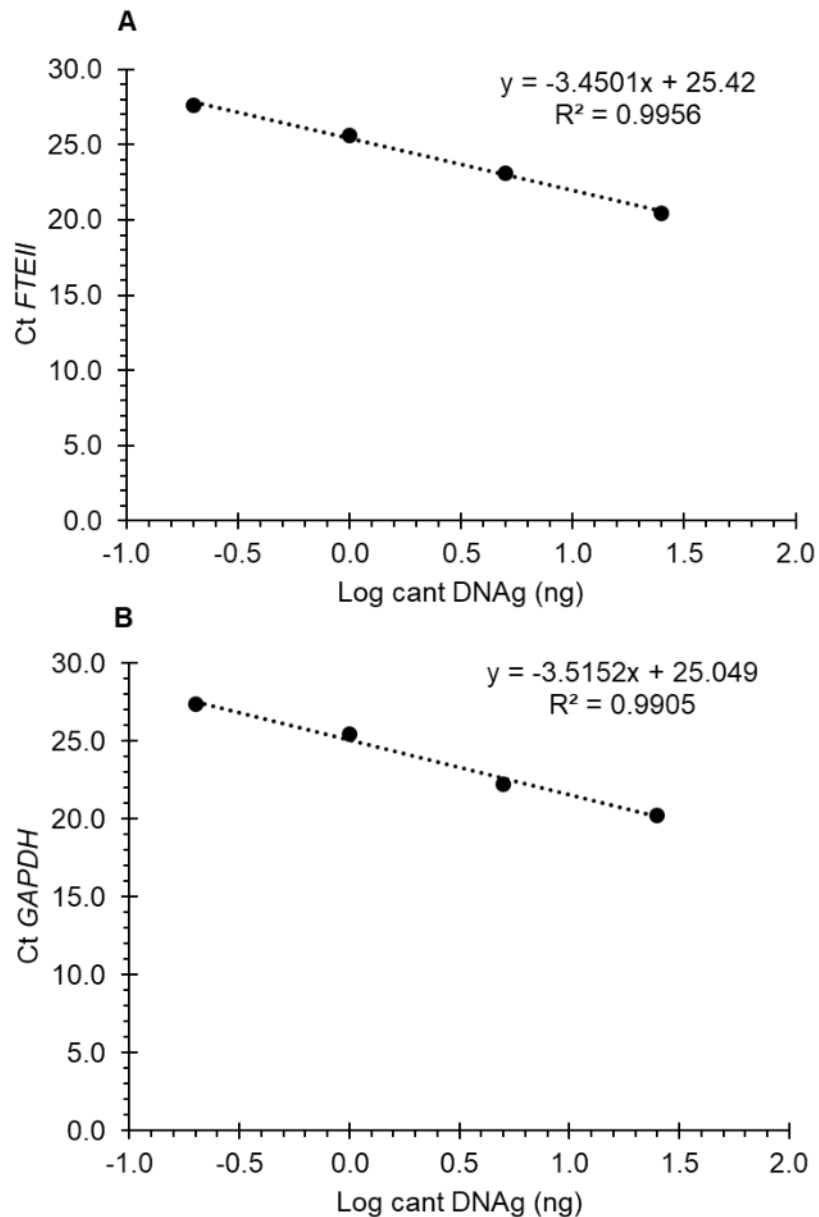
**Figura 33.** Electroforesis en gel de agarosa al 0.8% en TBE 1X de la caracterización por PCR de 3 clonas de cada construcción. M: marcador de tamaño molecular de 1kb DNA Ladder (Promega, Madison, WI, EUA). (A) Caracterización con los oligonucleótidos PP1033 y FTE2. Carril 1: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub>. Carril 2: control negativo de la PCR. Carril 3: Clona 21, Carril 4: Clona 29, Carril 5: Clona 8. (B) Caracterización con los oligonucleótidos FTE1 y 3THIPO1. Carril 1: control negativo de la PCR Carril 2: control positivo de PCR DNA plasmídico pP<sub>1033</sub>-FTEII-T<sub>1033</sub>. Carril 3: Clona 29, Carril 4: Clona 37, Carril 5: Clona 36.

**Tabla 30.** Preparación de DNA genómico para la determinación de dosis génica por qPCR.

<b>Preparaciones DNA genómico</b>	<b>Concentración ng/<math>\mu</math>L</b>	<b>Relación 260/280</b>	<b>Relación 260/230</b>
KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> Clona 21	1198	1.878	1.908
KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> Clona 29	1072	1.930	2.073
KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> Clona 8	1237	1.979	2.016
KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> Clona 29	1592	1.990	2.095
KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> Clona 36	1450	1.895	2.037
KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> Clona 37	1385	1.979	2.083

### **7.7.2.2. Curvas para determinar la eficiencia de amplificación de los genes FTEII y GAPDH**

Las curvas de eficiencia de los genes *FTEII* y *GAPDH* se realizaron graficando los valores de Ct, calculados con un umbral de fluorescencia de 400, frente al logaritmo en base 10 de la cantidad de DNA en el eje de las abscisas para cada gen. Las curvas de eficiencia de ambos genes mostraron linealidad en el intervalo de 0.2 a 25 ng de DNA genómico, con valores de pendientes de -3.4 y -3.5, un  $R^2$  mayor que 99 %. Los valores de eficiencia del qPCR obtenidos fueron de 94.9 y 92.5 % para *FTEII* y *GAPDH* respectivamente. Con las curvas de eficiencia y los valores de Ct obtenidos para cada muestra de DNA genómico de las clonas de bajo rendimiento proteínas/biomasa de ambas cepas se estimó la cantidad de DNA y con esta determinar la dosis génica mediante la cuantificación relativa del DNA del gen *FTEII* empleando el gen *GAPDH* como gen normalizador. En la figura 34A y 34B se muestra las curvas de eficiencia para los genes *FTEII* y *GAPDH* respectivamente.



**Figura 34.** Curvas de eficiencia de la amplificación por PCR de los genes *FTEII* (A) y *GAPDH* (B).

### 7.7.2.3 Determinación de la dosis génica de *FTEII* por qPCR

El número de copias o dosis génica del gen *FTEII* en las clonas seleccionadas como unicopias en base a los rendimientos de proteína extracelular/biomasa fue determinado de la cuantificación relativa del DNA del gen *FTEII* respecto al gen *GAPDH* en ambas cepas construidas. El análisis de qPCR para la cuantificación relativa del gen *FTEII* empleando

el gen *GAP* como normalizador del DNA de las clonas seleccionadas de cada construcción (KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> Clona 8 y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> Clona 36) muestra los valores de Ct obtenidos que fueron calculados con un umbral de fluorescencia de 400 para los genes *GAP* y *FTEII* (Tabla 31). Los valores de C<sub>t</sub> estuvieron comprendidos entre 24.39 y 26.06 para *GAP* y entre 24.58 y 26.24 para *FTEII*. Los coeficientes de variación de los valores de C<sub>t</sub> fueron menores del 1.3 % para *GAP* y 0.7 % para *FTEII*. Se cuantificaron los ng de DNA genómico de ambos genes con la media de Ct, la pendiente y el intercepto de la curva de eficiencia de cada gen, la dosis génica o el número de copias del gen *FTEII* de cada clona resultó de dividir los ng de DNA genómico de *FTEII* entre los ng de DNA genómico de *GAP* como gen normalizador (existe una única copia del gen *GAP* en el genoma de *Pichia pastoris*). El número de copias del gen *FTEII* en la clona 8 de la cepa KM71P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> recombinante fue de 0.98 y para la clona 36 de la cepa KM71P<sub>1033</sub>-FTEII-T<sub>1033</sub> fue de 1.05. Estos resultados indicaron que estas dos clonas tienen una sola copia del gen *FTEII* recombinante.

**Tabla 31.** Cuantificación relativa de dosis génica por qPCR: Valores de ciclo umbral (Ct), media, coeficiente de variación (CV%) y ng de los genes *GAP* y *FTEII* de las clonas analizadas.

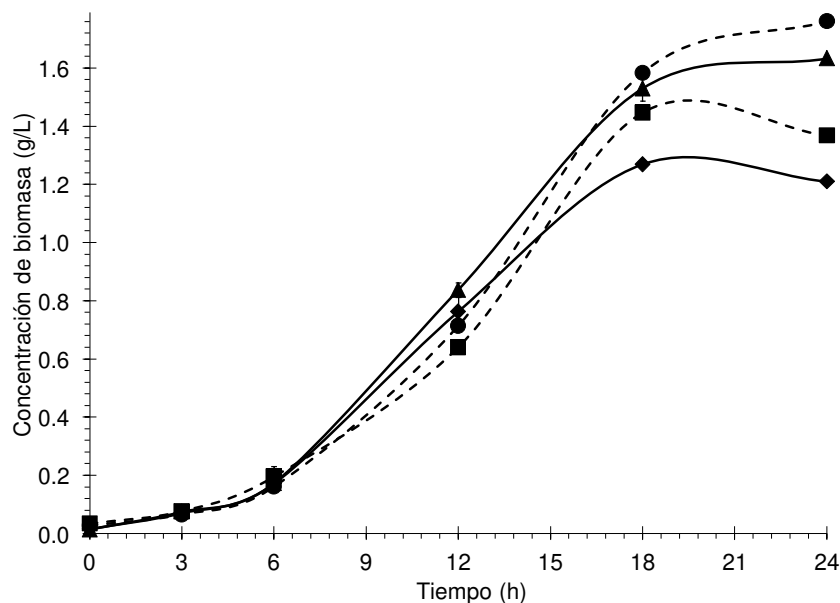
Clonas	FTEII					GAP					Número de copias
	C <sub>t1</sub>	C <sub>t2</sub>	Media	CV (%)	ng	C <sub>t1</sub>	C <sub>t2</sub>	Media	CV (%)	ng	
KM71/P <sub>1033</sub> -FTEII-T <sub>AOX1</sub> Clona 8	26.24	26.14	26.19	0.27	1.07	25.59	26.06	25.83	1.29	1.09	<b>0.99</b>
KM71/P <sub>1033</sub> -FTEII-T <sub>1033</sub> Clona 36	24.58	24.81	24.70	0.66	2.82	24.39	24.44	24.42	0.14	2.69	<b>1.05</b>

### 7.7.3. Actividad transcripcional del P<sub>1033</sub> en combinación con T<sub>AOX1</sub> y T<sub>1033</sub>

#### 7.7.3.1. Cinéticas de crecimiento de las cepas unicopias

La figura 35 muestra la cinética de crecimiento de ambas cepas cultivadas en glucosa y glicerol como fuente de carbono. El crecimiento celular incrementó exponencialmente de las 3 a las 12 horas del cultivo para las cuatro cinéticas, obteniendo valores de  $\mu$  en un intervalo de  $0.203 \pm 0.029$  a  $0.257 \pm 0.001$  h<sup>-1</sup> sin diferencias significativas entre los cultivos en glucosa y glicerol en una misma cepa o entre cepas crecidas en una misma

fuente de carbono. A pesar de que las cuatro cinéticas alcanzaron una meseta entre las 18 y 24 horas, los cultivos en glucosa obtuvieron 1.29 y 1.35 veces más concentración de biomasa a las 24 horas de cultivo que la obtenida en glicerol para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub>, respectivamente.



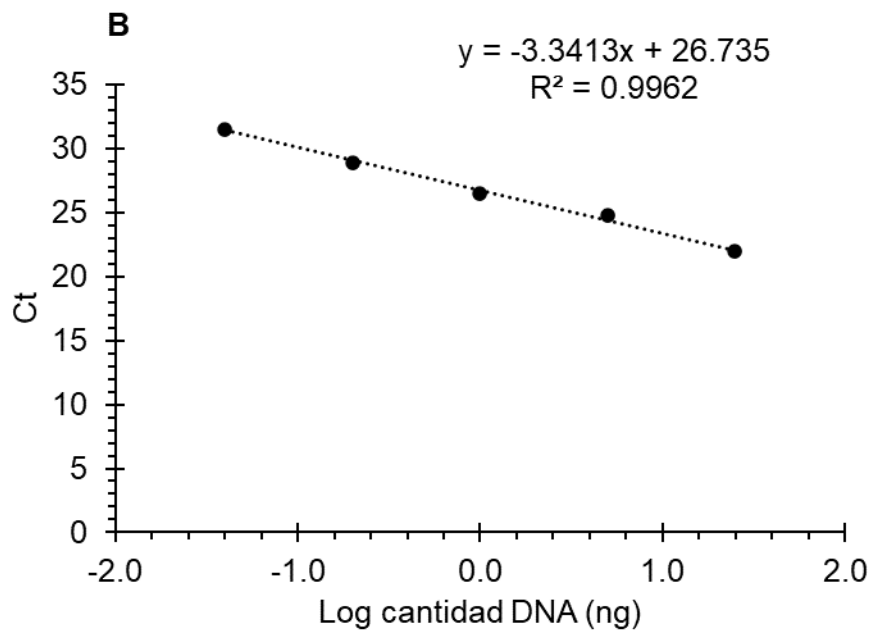
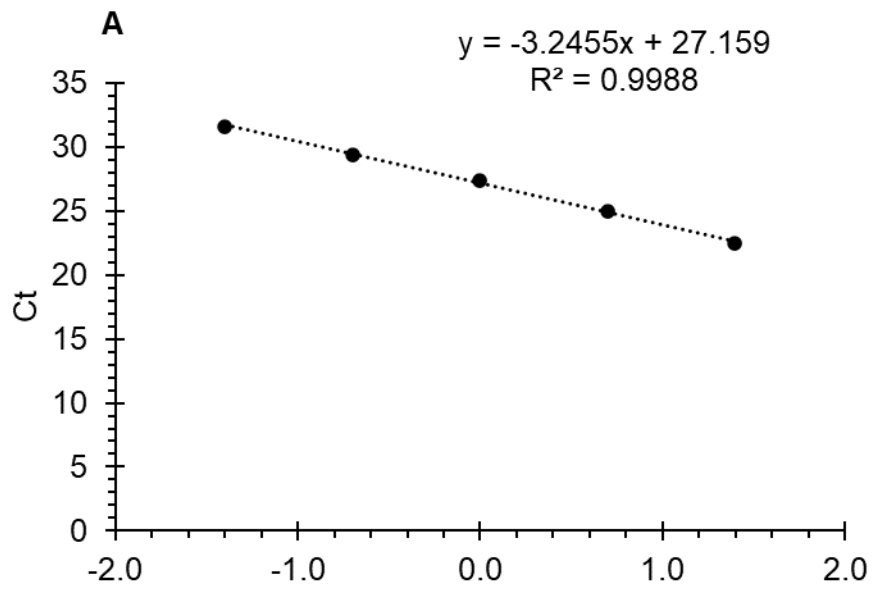
**Figura 35.** Cinética de crecimiento en medios BMGlc (●, ▲) y BMGly (■, ◆) para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (●, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▲, ◆). Los puntos representan la media de tres experimentos cinéticos independientes (coeficiente de variación inferior al 5%).

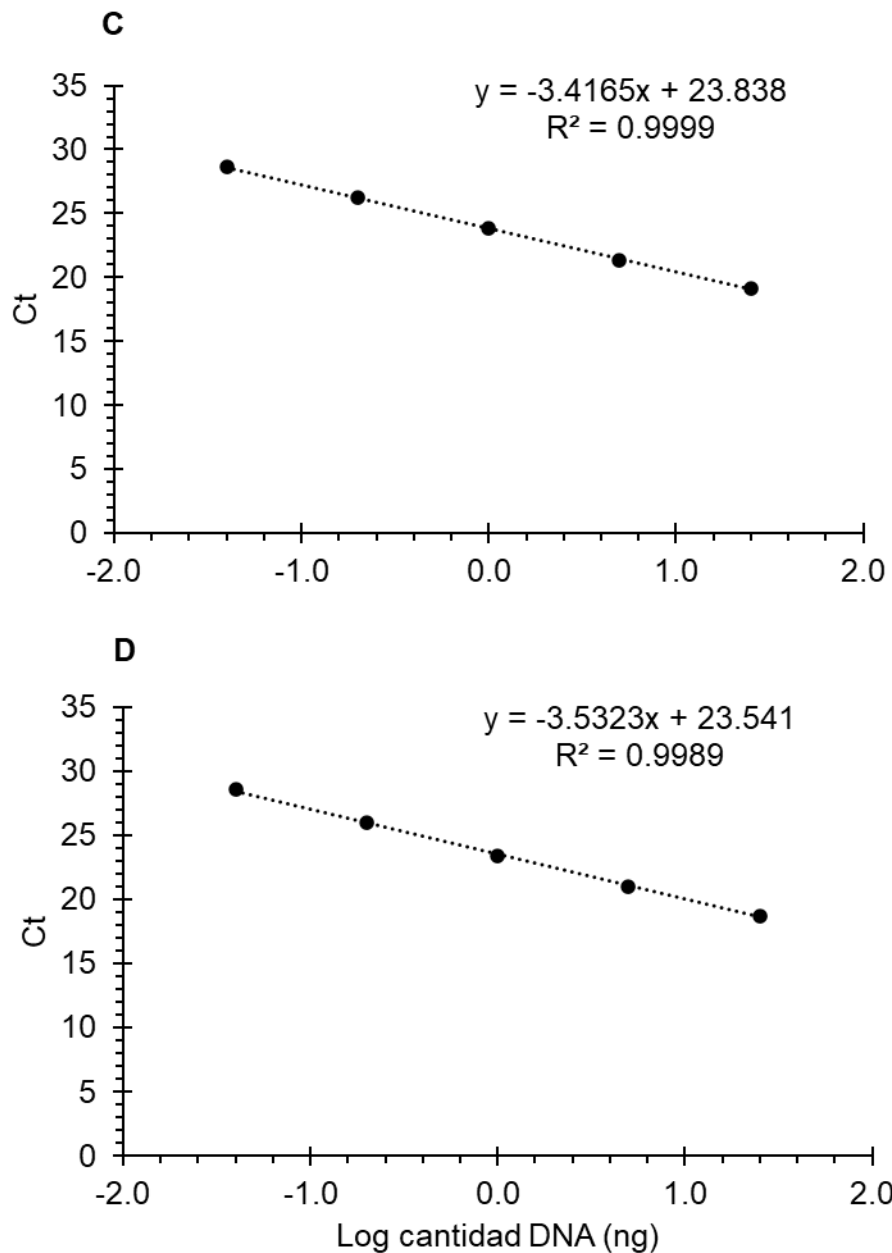
### 7.7.3.2. Determinación de los niveles de transcritos de los genes *FTEII*, *1033* y *GAPDH* en células de las cepas recombinante cultivadas en glucosa y glicerol

Se realizaron curvas de eficiencia de los genes *FTEII*, *GAPDH*, *1033* y *YPT1* graficando los valores de Ct, calculados con un umbral de fluorescencia de 400, frente al logaritmo en base 10 de la cantidad de DNA en el eje de las abscisas para cada gen. Las curvas de eficiencia de los 4 genes mostraron linealidad en el intervalo de 0.04 a 25 ng de DNA genómico, con valores de pendientes de -3.3 para los genes *FTEII* y *GAPDH*, -3.4 para *YPT1* y -3.5 para *1033*, y con un R<sup>2</sup> mayor que 99 %. Los valores de eficiencia del qPCR obtenidos fueron de 103.3, 99.2, 96.2 y 91.9 % para los genes *FTEII*, *GAPDH*, *YPT1* y



1033 respectivamente. En la figura 36 (A, B, C, D) se muestra las curvas de eficiencia para los genes *FTEII*, *GAPDH*, *YPT1* y *1033* respectivamente.





**Figura 36.** Curvas de eficiencia de la amplificación por PCR de los genes *FTEII* (A), *GAPDH* (B), *YPT1* (C) y *1033* (D).

Los niveles de transcritos de los genes *FTEII*, *1033* y *GAPDH* se midieron en células de ambas cepas recombinantes cultivadas en glucosa y glicerol a partir de paquetes celulares de 2 mg de biomasa, las preparaciones de RNA rindieron concentraciones en el rango de 50.4 a 213.2 ng/ $\mu$ L con buena calidad pues la relación de las lecturas de absorbancia a 260/280 nm estuvo en el intervalo de 1.814 a 2.023 y la relación de las lecturas de absorbancia a 260/230 nm estuvo entre 1.734 y 2.053.

Los RNA extraídos se sometieron a reacciones de retro transcripción para obtener DNAc, realizando a continuación una reacción de cuantificación por qPCR de un control negativo de retrotranscripción para cada DNAc sintetizado, en ninguno de los casos este control negativo dio valor de Ct por lo que los DNAc sintetizados no presentaban contaminación de DNA genómico de la levadura.

Las reacciones de qPCR se montaron según el acápite 6.11.3.2, poniendo el DNAc diluido mezclado con agua y mater mix en un tubo, del que se distribuye de manera homogénea a los tubos donde se va a realizar la cuantificación de cada gen en específico, esto garantizó mayor reproducibilidad en los resultados de la cuantificación de cada gen y en los duplicados de qPCR.

En las reacciones de qPCR se calcularon los valores de Ct con un umbral de 400 para cada gen, los cuales se encontraron dentro del intervalo de linealidad y mostraron valores de  $CV < 2\%$ , indicando una buena reproducibilidad de los duplicados analíticos en cada una de las muestras analizadas.

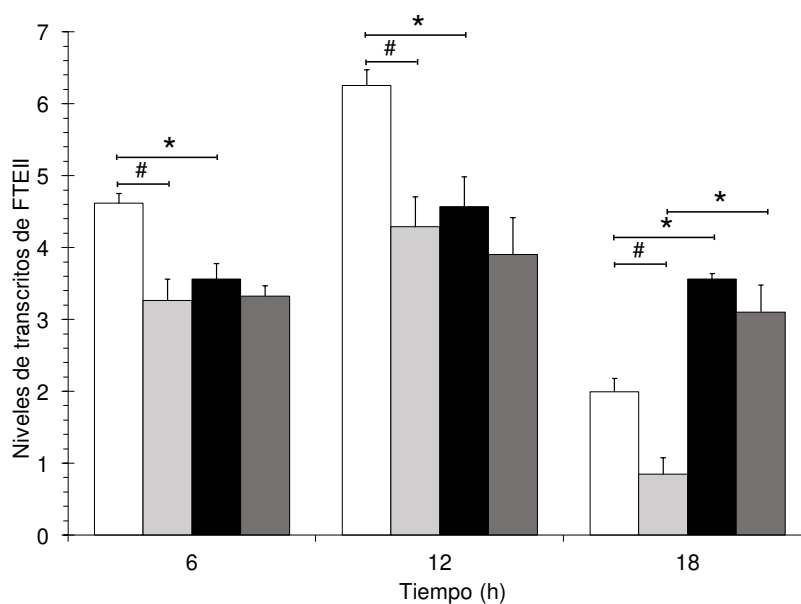
#### **7.7.3.2.1. Actividad transcripcional del gen *FTEII***

Los niveles de transcritos del gen *FTEII*, empleado como gen reportero, el cual es regulado por  $P_{1033}$ , fueron 1.4 veces más altos en las células de  $KM71/P_{1033}\text{-}FTEII\text{-}T_{AOX1}$  cultivadas en glucosa durante la fase exponencial que en las células  $KM71/P_{1033}\text{-}FTEII\text{-}T_{1033}$  (Fig. 37), mientras que en los cultivos en glicerol no se presentaron diferencias significativas en los niveles de transcritos obtenidos de las muestras celulares en todos los tiempos analizados entre ambas cepas (Fig. 37).

Por otro lado, el perfil de los niveles de transcritos de *FTEII* de la cepa  $KM71/P_{1033}\text{-}FTEII\text{-}T_{AOX1}$  crecida en glucosa estuvo ligada al crecimiento. Aunque  $KM71/P_{1033}\text{-}FTEII\text{-}T_{1033}$  crecida en glucosa también presentó un perfil de niveles de transcritos ligado al crecimiento, éste no fue estadísticamente significativo. En cambio, ambas cepas crecidas en glicerol mantuvieron niveles de transcritos de *FTEII* constante en la fase exponencial. Los niveles de transcritos de *FTEII* fueron 1.3 veces más altos en los cultivos en glucosa que en glicerol durante la fase exponencial en  $KM71/P_{1033}\text{-}FTEII\text{-}T_{AOX1}$ , mientras que en

las células de KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> no se mostraron diferencias significativas en los niveles de transcritos de *FTEII* crecidas en ambas fuentes de carbono.

El gen *FTEII* mostró una disminución de los niveles de transcritos más alto en glucosa que en glicerol durante la fase estacionaria, ya que los niveles de transcritos de *FTEII* fueron 1.8 y 3.6 veces más altos en las células crecidas en glicerol que en glucosa en KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> respectivamente durante la fase estacionaria cuando la fuente de carbono se ha agotado (Fig. 37).



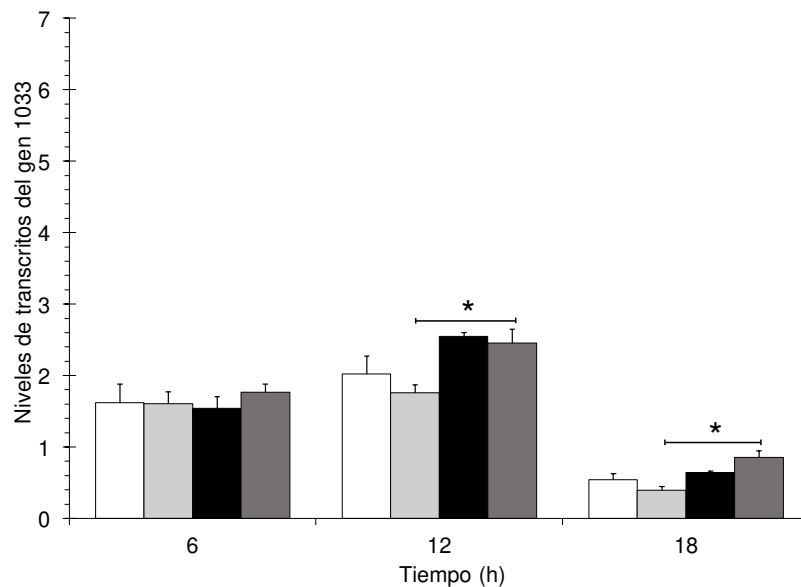
**Figura 37.** Niveles de transcripción relativos del gen *FTEII* de para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▤, ▥) en los medios en glucosa BMGlc ((□, ▤) y glicerol BMGly ((■, ▥) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas ( $p < 0,05$ ) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas ( $p < 0,05$ ) entre las dos cepas cultivadas en la misma fuente de carbono.

#### 7.7.3.2.2. Actividad transcripcional del gen *1033*

Los niveles de transcripción del gen *1033* endógeno no mostraron diferencias significativas entre las dos cepas cultivadas en glucosa o glicerol (Fig. 38).

El perfil del nivel del transcrito de *1033* se relacionó con el crecimiento celular en glicerol, ya que los niveles del transcrito aumentaron significativamente durante la fase de

crecimiento exponencial, 1.7 veces y 1.4 veces más entre las 6 y las 12 h de cultivo para KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71P<sub>1033</sub> -células FTEII-T<sub>1033</sub>, respectivamente (Fig. 38). En glucosa se mantiene la misma tendencia del perfil de niveles de transcritos de *1033* relacionado con el crecimiento celular que en glicerol entre las 6 y 12 horas, sin embargo, no fue estadísticamente significativo. Los niveles de transcritos del gen *1033* fueron hasta 1.4 veces más altos en las células crecidas en glicerol en comparación con las células cultivadas con glucosa a las 12 h de cultivo para ambas cepas, pero en KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> este incremento no fue estadísticamente significativo (Fig. 38). El gen *1033* se reguló a la baja en ambas cepas cuando la fuente de carbono se agotó en la fase estacionaria, esta disminución fue más alta en glucosa que en glicerol.



**Figura 38.** Niveles de transcripción relativos del gen 1033 para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▤, ▥) en los medios en glucosa BMGlc (□, ▤) y glicerol BMGly (■, ▥) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa.

### 7.7.3.2.3. Actividad transcripcional del gen *GAPDH*

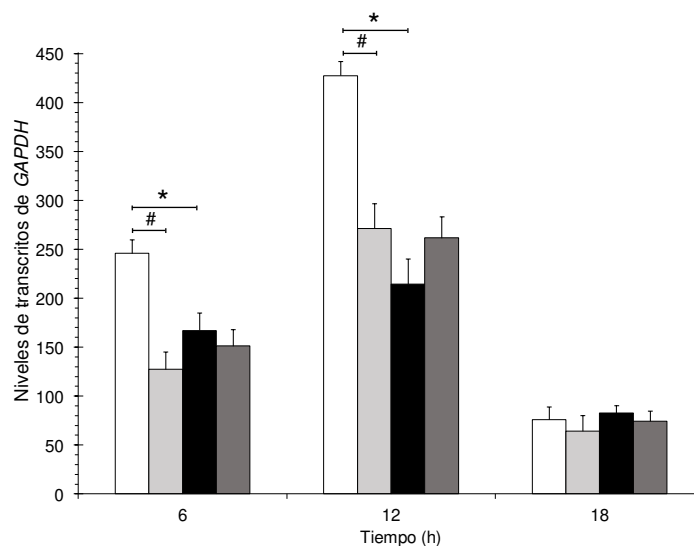
Los niveles de transcritos de *GAPDH* endógeno no mostraron diferencias significativas entre cepas cultivadas en glicerol como fuente de carbono, mientras que en los cultivos en glucosa los niveles de transcritos de *GAPDH* fueron hasta 1.8 veces más altos en

KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> que en la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> durante la fase exponencial de las 6 a las 12 h (Fig. 39).

El perfil de los niveles de transcritos de *GAPDH* se mantuvo ligada al crecimiento celular para ambas cepas y fuentes de carbono, con un aumento del nivel de transcripción durante la fase de crecimiento exponencial y una disminución cuando las células alcanzaron la fase estacionaria.

Los niveles de transcritos de *GAPDH* de las células de KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> cultivadas en glicerol no presentaron diferencias significativas entre las 6 y 12 horas del cultivo, lo cual podría asociarse con una velocidad específica de crecimiento casi constante durante la fase exponencial.

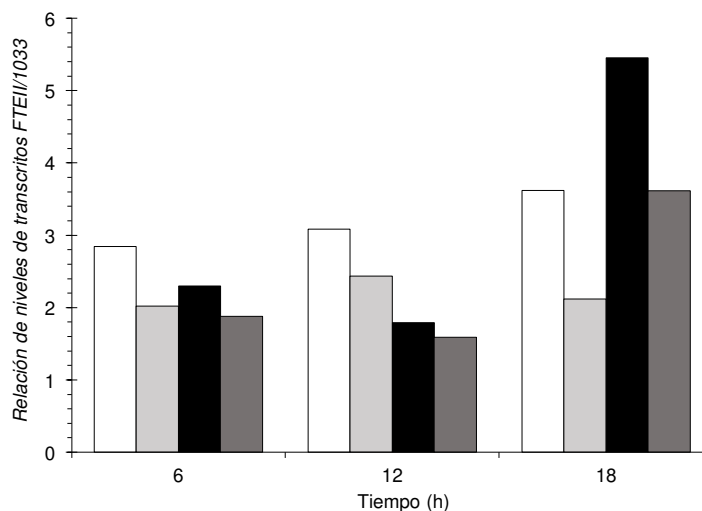
Los niveles de transcripción de *GAPDH* de la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> en la fase de crecimiento exponencial fueron 1.7 veces más altos para las células cultivadas con glucosa, en comparación con las células cultivadas con glicerol, mientras que los niveles de transcripción de *GAPDH* no mostraron diferencias significativas para las células KM71/ Células P<sub>1033</sub>-FTEII-T<sub>1033</sub> cultivadas con cualquiera de las dos fuentes de carbono (Fig.39).



**Figura 39.** Niveles de transcripción relativos del gen *GAPDH* de para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▤, ▥) en los medios en glucosa BMGlc ((□, ▤) y glicerol BMGly ((■, ▥) a diferentes tiempos de cultivo. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas (p < 0,05) entre las dos cepas cultivadas en la misma fuente de carbono.

#### 7.7.3.2.4. Comparación de la actividad transcripcional de los genes *1033*, *FTEII* y *GAPDH*

Los niveles de transcripción de *FTEII* fueron en promedio 2.8 y 2.1 veces más altos, y 2.0 y 2.2 veces más altos que los niveles de transcripción de *1033* para las células KM71/P<sub>1033</sub>-*FTEII*-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-*FTEII*-T<sub>1033</sub> cultivadas con glucosa. y para células KM71/P<sub>1033</sub>-*FTEII*-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-*FTEII*-T<sub>1033</sub> cultivadas en glicerol, respectivamente (Fig. 40). Durante la fase estacionaria, el incremento de la relación entre ambos genes se potencia en ambas fuentes de carbono (Fig. 40).

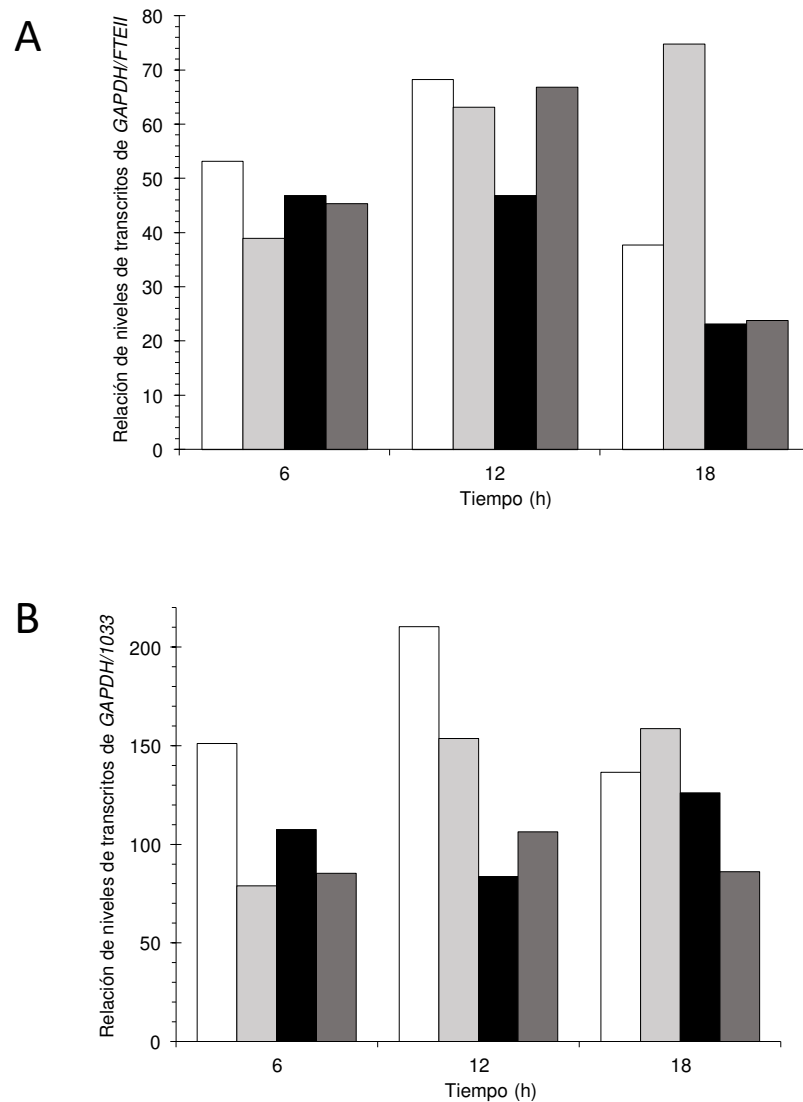


**Figura 40.** Relación del nivel de transcripción de *FTEII* respecto a *1033* en las cepas KM71/P<sub>1033</sub>-*FTEII*-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-*FTEII*-T<sub>1033</sub> (▤, ▥) en los medios en glucosa BMGlc ((□, ▤) y glicerol BMGly ((■, ▥) a diferentes tiempos. Los datos se presentan como la media de tres cultivos independientes.

Los niveles de transcritos de *GAPDH* fueron en promedio 50 y 60 veces más altos que *FTEII* (regulado por P<sub>1033</sub>) en los cultivos en glucosa, mientras que en los cultivos en glicerol fueron 34 y 38 veces más altos en KM71/P<sub>1033</sub>-*FTEII*-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-*FTEII*-T<sub>1033</sub> respectivamente para ambos casos (Fig. 41 A).

Los niveles de transcritos de *GAPDH* fueron en promedio 142 y 125 veces más altos que los niveles de transcritos del gen *1033* en las células de KM71/P<sub>1033</sub>-*FTEII*-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-*FTEII*-T<sub>1033</sub> en los cultivos en glucosa respectivamente. Mientras que, en los cultivos en glicerol, los niveles de transcritos de *GAPDH* fueron en promedio 106 y 93

veces más altos que los obtenidos para el gen *1033* en las células de KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> respectivamente a las 6 horas del cultivo (Fig. 41 B).



**Figura 41.** Relación del nivel de transcripción de *GAPDH* /FTEII (A) y *GAPDH* /P<sub>1033</sub> (B) en las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (▤, ▥) en los medios en glucosa BMGlc ((□, ▤) y glicerol BMGly ((■, ▥) a diferentes tiempos. Los datos se presentan como la media de tres cultivos independientes.

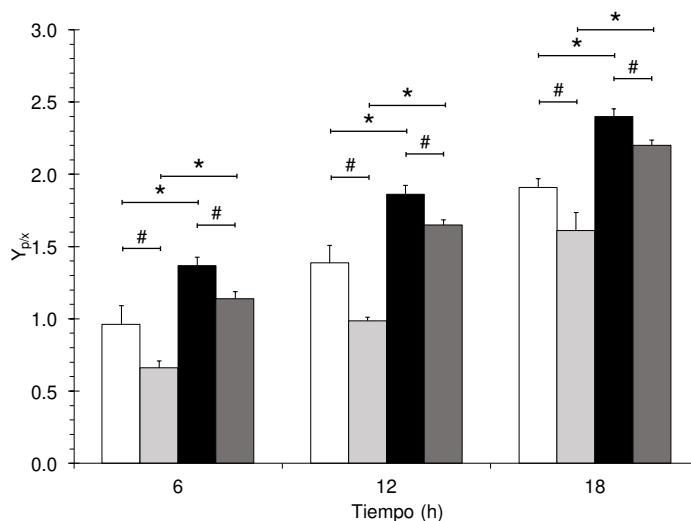
### 7.7.3.3. Determinación del rendimiento producto extracelular por unidad de biomasa

El rendimiento de la producción extracelular de fitasa,  $Y_{p/x}$ , fue en promedio 1.4 veces más alto en la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> que en la cepa KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> en los cultivos en glucosa (Fig. 42). Estos hallazgos son consistentes con los niveles de



transcritos 1.4 más altos de *FTEII* en KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> cultivada en glucosa respecto a los obtenidos en células KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> en las mismas condiciones de cultivo. La diferencia en el rendimiento Y<sub>p/x</sub> fue menos pronunciado (1.1) entre ambas cepas cultivadas en glicerol (Fig. 42), lo cual se relaciona con los niveles de transcritos sin diferencias significativas entre cepas en los cultivos en glicerol.

A pesar de que los niveles de transcritos de *FTEII* fueron más altos (1.4 veces) en los cultivos en glucosa que en glicerol en KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub>, el rendimiento producto/biomasa fue 1.3 veces más alto en los cultivos en glicerol que en glucosa (Fig. 42).



**Figura 42.** Rendimiento de producto/biomasa extracelular (Y<sub>p/x</sub>) para las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> (□, ■) y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> (■, ■) en los medios en glucosa BMGlc ((□, ■) y glicerol BMGly ((■, ■) a diferentes tiempos. Los datos se presentan como la media ± error estándar de tres cultivos independientes. \*Diferencias significativas (p < 0,05) entre células cultivadas con glicerol y glucosa de la misma cepa. #Diferencias significativas (p < 0,05) entre las dos cepas cultivadas en la misma fuente de carbono.

Por otro lado, a pesar de que no se presentaron diferencias significativas en los niveles de transcritos de *FTEII* en KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> entre los cultivos en glucosa y glicerol, el rendimiento producto/biomasa fue 1.7 veces más altos en glicerol que en glucosa.

Estos hallazgos indican que la ruta de traducción y secreción de proteínas tuvo un alto impacto en el Y<sub>p/x</sub> cuando se usó glicerol como fuente de carbono. El Y<sub>p/x</sub> más alto en ambas cepas ocurrió al final de la fase de crecimiento exponencial y al comienzo de la

fase de crecimiento estacionario, cuando los niveles de transcritos de *FTEII* disminuyeron significativamente en las dos fuentes de carbono evaluadas (Fig. 42).

## 8. DISCUSIÓN

Aunque un promotor fuerte inducible o constitutivo es la primera opción para expresar genes heterólogos en *P. pastoris*, recientemente ha aumentado el interés en los promotores débiles debido a su potencial para producir proteínas tóxicas para el hospedero (Ruth et al. 2010) o para co- expresar varios genes y mantener un equilibrio metabólico en el hospedero (Vogl et al. 2016; Jin et al. 2019; Dou et al. 2021). Como consecuencia de lo antes mencionado es importante identificar nuevos promotores para la producción de proteínas recombinantes.

Otra herramienta importante para regular los niveles de expresión de proteínas es el terminador de la transcripción. Las regiones terminadoras no solo actúan en la terminación de la transcripción si no que pueden influir en los niveles de expresión de una proteína codificada río arriba del terminador (Yamanishi et al. 2013; Curran et al. 2013). Sin embargo, se han realizado pocos estudios sobre terminadores transcripcionales que se puedan usar para ajustar la expresión génica, en comparación con los numerosos estudios de promotores (Ito et al. 2020).

Con el fin de caracterizar funcionalmente el arreglo de nuevos elementos reguladores que incluyen al promotor y terminador de la transcripción y determinar la influencia de la combinación de estos elementos reguladores en la expresión de un gen heterólogo en el sistema de *P. pastoris*, en este trabajo se identificaron y caracterizaron funcionalmente por primera vez el promotor (P<sub>1033</sub>) y el terminador transcripcional (T<sub>1033</sub>) del gen constitutivo *I033* (PAS\_chr3\_1033) con actividad transcripcional débil, no dependiente de metanol.

Se determinó la actividad del promotor en la transcripción del gen heterólogo y el efecto del terminador de la transcripción en la estabilidad del transcrito y se evaluó el impacto de los elementos reguladores en la traducción y la secreción de proteínas determinando el rendimiento de producto/biomasa extracelular (Y<sub>p/x</sub>).

Los resultados de este estudio indican que P<sub>1033</sub> tiene un 2-3% de actividad transcripcional con respecto a P<sub>GAP</sub> y puede ajustarse o regularse con la fase de crecimiento celular y la fuente de carbono. La combinación de los elementos reguladores promotor-terminador del

gen *1033* contribuyeron a una actividad transcripcional diferenciada de los genes heterólogos y endógenos y fue dependiente de la fuente de carbono, lo que demuestra la influencia del terminador transcripcional en el casete de expresión en los niveles de transcritos de los genes heterólogos y endógenos. Por otro lado, la combinación del par promotor-terminador y la fuente de carbono afectaron de forma diferenciada las vías de traducción y/o secreción de proteínas.

Las regiones 5'UTR y 3'UTR del gen *1033* no están anotados en las bases de datos de genomas disponibles de *P. pastoris*. Por esta razón, analizamos los datos de RNA-seq para ubicar subregiones río arriba y río abajo del CDS *1033* donde las lecturas de RNA no se alinearon con el genoma de referencia. En la región río arriba, la secuencia donde no se mapearon las lecturas de RNA tuvo una longitud de 136 nucleótidos, mientras que la región inter-CDS entre el gen *1033* y el gen adyacente río arriba está anotado con 175 nucleótidos. Dado que estas dos regiones son de tamaño similar, decidimos usar la secuencia de la región inter-CDS como la secuencia del promotor P<sub>1033</sub> a evaluar.

El P<sub>1033</sub> contiene una secuencia tipo TATA con dos cambios de nucleótidos de la secuencia consenso de la caja TATA (TATAWAW), que es característica de los promotores constitutivos (Donczew y Hahn 2018). Además, la secuencia P<sub>1033</sub> alberga un elemento INR de tipo mamífero río abajo de la caja tipo TATA. Las secuencias INR de tipo mamífero están presentes en, o adyacentes, al TSS de aproximadamente el 40% de los promotores de levadura (Yang et al. 2007). El supuesto TSS se ubicó en una secuencia de CA en la hebra codificante, es decir, dinucleótido de pirimidina-purina (PyPu) en los sitios -1 y +1 del TSS, con dos adeninas en una región de 7 nucleótidos inmediatamente río arriba de la secuencia de CA. Esta secuencia de nucleótidos TSS se ha descrito para especies de levadura que utilizan el modelo de exploración para el inicio de la transcripción (Lu y Lin 2021). La secuencia P<sub>1033</sub> contiene un tramo homopolimérico de 21 nucleótidos de desoxitimidina (Ts) en la cadena codificante. Las secuencias poli(dA:dT) son muy abundantes en los genomas eucarióticos (Segal y Widom 2009; Yang y Zhang 2018) y se consideran como un elemento característico del promotor río arriba en los promotores constitutivos de levadura que actúan bidireccionalmente para activar la transcripción en direcciones opuestas de dos genes adyacentes no relacionados

(Struhl 1985, 1986), como es el caso del gen *1033* y el gen adyacente río arriba, como es el caso del gen *NOB1*.

Estas secuencias albergan una estructura helicoidal más corta y un surco menor estrecho, que sirve para resistir la flexión requerida para la unión de histonas (Jansen y Verstrepen 2011). Las regiones Poli (dA:dT) podrían crear una “barrera” que evite la formación de nucleosomas en esa región, por lo que sólo se pueden formar en regiones adyacentes (Mavrich et al. 2008). Debido a su estructura inusual, los nucleosomas se estrechan en estas regiones lo que aumenta la accesibilidad del DNA para los factores transcripcionales (Segal y Widom 2009).

Las secuencias poli (dA:dT) se han utilizado para regular los promotores en levaduras, alterar la presencia y la longitud de las regiones poli (dA:dT) nativas pueden aumentar la accesibilidad y favorecer la unión de factores transcripcionales, regulando así la actividad del promotor (Lyer y Struhl 1995; Raveh-Sadka et al. 2012).

Un estudio en el promotor *AOX1* para analizar el efecto transcripcional de las regiones poli (dA: dT) alterando la presencia y longitud del elemento poli (dA:dT) nativo en diferentes sitios, indicó que la eliminación o el alargamiento de los trectos poli (dA:dT) nativos pueden alterar las actividades de las variantes que van desde ~ 0.25 a ~ 3.5 veces la actividad del promotor nativo, lo que demostró que este método es más eficaz que los métodos tradicionales para mejorar la actividad del promotor *AOX1* (Yang y Zhang 2018). Solo se identificaron 6 sitios de unión de factores de transcripción (SUFT) en el análisis de secuencia  $P_{1033}$ , lo cual es menos que los 11 y 21 SUFT que identificamos en las secuencias  $P_{AOX1}$  y  $P_{GAP}$ , respectivamente, realizando el mismo análisis *in silico* (datos no publicados). Entre los supuestos 6 factores de transcripción (FT) para  $P_{1033}$ , solo uno actúa sobre  $P_{1033}$  y  $P_{GAP}$  (Cat8-1).

Cat8 es un activador transcripcional con un motivo estructural de dedos de zinc necesario para la desrepresión de una variedad de genes bajo condiciones de crecimiento no fermentativo. En *S. cerevisiae*, Cat8 es uno de los FT más importantes que activa genes de la vía de gluconeogénesis (Schüller 2003; Turcotte et al. 2010). *Pichia pastoris* tiene dos homólogos putativos de Cat8 (Cat8-1 y Cat8-2) que están involucrados en la activación de la vía de asimilación del etanol (Barbay et al. 2021). Además, Cat8-1 es

necesario para activar genes en el ciclo del glioxilato que contribuyen a la vía de la gluconeogénesis.

La identificación de este SUFT en la secuencia P<sub>1033</sub> podría estar relacionada con la supuesta función de la proteína codificada por el gen *1033*, D-aminoácido oxidasa, que cataliza la oxidación de D-aminoácidos neutros a  $\alpha$ -cetoácidos, que pueden entrar en la gluconeogénesis, glucólisis y vías del ciclo de Krebs. Análisis de la región promotora del gen D-aminoácido oxidasa de *Rhodotorula gracilis* hasta 300 pb río arriba del codón de inicio ATG, no logró identificar elementos putativos que favorecieran la transcripción (Alonso et al. 1998). En la levadura metilotrófica *C. boidinii*, la D-aminoácido oxidasa participa en el crecimiento en D-Alanina como fuente de carbono, sin embargo, no han sido identificados elementos responsables de la transcripción en la región río arriba del gen D-aminoácido oxidasa en *C. boidinii* (Sakai et al. 1998; Yurimoto et al. 2001).

El análisis del transcriptoma río abajo del CDS del gen *1033* se usó para definir el extremo 3'UTR del gen *1033*, y en el cual se incluyeron 107 nucleótidos ubicados río abajo del final del CDS del gen *1033*, además consideramos incluir en la secuencia de la región 3'UTR, 50 nucleótidos adicionales para la secuencia terminadora de la transcripción del gen *1033* (T<sub>1033</sub>) a evaluar y siguiendo las sugerencias de la literatura para terminadores transcripcionales de levadura (Curran et al. 2013).

La construcción correcta de los vectores de expresión pP<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y pP<sub>1033</sub>-FTEII-T<sub>1033</sub> y de las cepas KM71/P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> y KM71/P<sub>1033</sub>-FTEII-T<sub>1033</sub> que albergan las secuencias de los dos elementos reguladores a evaluar (P<sub>1033</sub> y T<sub>1033</sub>) junto con T<sub>AOX1</sub> fueron confirmados por PCR.

Caracterizamos funcionalmente el P<sub>1033</sub> de *Pichia pastoris*, confirmamos el evento de expresión génica mediante la detección por RT-PCR de transcritos del gen *FTEII*, este resultado es similar al obtenido por otros autores que detectaron transcritos de *PhysC* bajo la regulación del promotor *AOXI* (Guerrero-olazarán et al. 2007). También se detectó la actividad volumétrica de fitasa extracelular en los cultivos de ambas cepas y corroboramos el tamaño molecular de la fitasa recombinante FTEII de los sobrenadantes de cultivo de la cepa KM71P<sub>1033</sub>-FTEII-T<sub>AOX1</sub> obteniendo una masa molecular aparente de 39 kDa que corresponde a la masa molecular teórica de esta fitasa y que coincide con lo obtenido en

investigaciones previas en nuestro grupo de trabajo en construcciones que tenían como gen reportero *FTEII* regulado por el promotor *AOXI* (Viader-Salvadó et al. 2010).

La débil actividad del promotor  $P_{1033}$  se observó inicialmente en los cultivos de clonas transformantes de las dos cepas construidas ya que el rango en el rendimiento de proteína/biomasa fue de 5 a 23 veces inferior a los obtenidos previamente para cultivos similares con clonas con arreglos  $P_{GAP}$ - $T_{AOXI}$  utilizando el mismo gen reportero (Herrera-Estala et al. 2022).

En trabajos anteriores (Viader-Salvadó et al. 2010; Herrera-Estala et al. 2022) se reportó que la integración del casete de expresión en el locus *HIS4* transformando por electroporación podría conducir a múltiples eventos de inserción en este locus. Algunos autores han encontrado una relación entre el número de copias del casete de expresión en el genoma de *P. pastoris* con la producción de la proteína recombinante, demostrando que en la medida que incrementa el número de copias del gen heterólogo se obtiene como resultado mayor rendimiento de la proteína heteróloga (Cereghino y Cregg, 2000; Nordén et al. 2011; Prielhofer et al. 2013a; Zhu et al. 2014; Looser et al. 2015). En un estudio donde la región promotora del gen metanol oxidasa ( $P_{MOX}$ ), de *Hansenula polymorpha*, se exploró para la expresión de proteínas heterólogas en cepas de *P. pastoris* utilizando las enzimas endoglucanasa 3 (CMC3) y endoglucanasa II (EgII) como proteínas reporteras, los resultados indicaron que el aumento de la dosis del gen aumentó significativamente la expresión de CMC3 y EgII bajo el control de  $P_{MOX}$  (Mombeni et al. 2020). Por esta razón, y para una mejor comprensión del efecto de los dos re-arreglos genéticos contenidos en el casete de expresión sobre los niveles de expresión, evaluamos la dosis génica de varios transformantes por qPCR y seleccionamos clonas de única copia para evitar el posible efecto del número de copias del gen sobre los niveles de transcritos y niveles de proteínas del gen heterólogo y para comparar con los niveles de transcripción de los genes endógenos *1033* y *GAPDH* en los cultivos de las clonas con los dos reordenamientos genéticos del casete de expresión.

Una proteína reportera intracelular como la proteína fluorescente verde, se ha utilizado generalmente para evaluar la fuerza de un promotor (Hartner et al. 2008; Qin et al. 2011; Prielhofer et al. 2013). Sin embargo, con este método, no se evalúa directamente la

actividad del promotor en la transcripción de genes, sino que se mide la transcripción junto con el proceso de traducción. Una de las mayores ventajas del sistema de expresión de *P. pastoris* es su eficiente sistema de secreción de proteínas. Por lo tanto, a menudo se prefiere este huésped para producir proteínas extracelularmente, lo que facilita el proceso de purificación. En consecuencia, decidimos evaluar la transcripción del gen heterólogo por RT-qPCR utilizando un casete de expresión conteniendo un gen reportero cuyo producto proteico se secreta en el medio de cultivo, ya que la secuencia codificante se asoció funcionalmente con una secuencia señal de secreción, y la combinación de dos reordenamientos genéticos del par promotor y terminador de la transcripción como elementos reguladores. Con este enfoque, y además de la actividad del promotor en la transcripción del gen heterólogo y el efecto del terminador de la transcripción en la estabilidad del transcrito, evaluamos el impacto de los elementos reguladores en la traducción y la secreción de proteínas determinando el rendimiento de producto/biomasa extracelular (Yp/x).

Las diferencias en la actividad transcripcional de P<sub>1033</sub> en *FTEII* entre las dos cepas cultivadas en glucosa no se observaron para el gen *1033* nativo. Por el contrario, los niveles de transcripción de *FTEII* y *1033* no mostraron diferencias significativas entre las dos cepas cultivadas en glicerol.

Estos resultados indican que los terminadores T<sub>AOX1</sub> y T<sub>1033</sub> no contribuyen a diferenciar los niveles de transcripción de *FTEII* en células cultivadas con glicerol, pero en glucosa, T<sub>AOX1</sub> aumenta los niveles de transcrito de *FTEII* con respecto a T<sub>1033</sub>, lo que muestra la influencia del terminador transcripcional en el casete de expresión sobre los niveles de transcritos del gen heterólogo. Estos hallazgos son similares a los descritos anteriormente (Ramakrishnan et al. 2020; Ito et al. 2020; Herrera-Estala et al. 2022), donde la actividad del T<sub>AOX1</sub> fue mayor que la de otros terminadores transcripcionales endógenos, indicando que se debía a la mayor estabilidad de la región 3'UTR y una mayor vida media del RNAm de T<sub>AOX1</sub>.

Nuestros hallazgos sugieren que la estabilidad del RNAm también podría depender de la fuente de carbono utilizada para el crecimiento celular.



El gen *1033* endógeno se expresó con bajos niveles de transcritos en comparación con el gen *GAPDH*, y su expresión fue modulada por el crecimiento celular y la fuente de carbono. Ambas cepas mostraron niveles de transcritos 1.4 veces más altos del gen *1033* en glicerol que en células cultivadas con glucosa durante la fase de crecimiento exponencial. Se observó una actividad transcripcional similar de  $P_{1033}$  en el gen *1033* para las dos cepas en las mismas fuentes de carbono.

En la fase estacionaria, los niveles del transcrito del gen *1033* disminuyeron, probablemente debido a la ausencia de la fuente de carbono. La disminución fue más pronunciada en los cultivos de glucosa que en los de glicerol. Aunque la expresión de los genes *FTEII* y *1033* fue dirigida por  $P_{1033}$  en ambas cepas, los niveles transcripcionales de *FTEII* fueron más altos que los del gen *1033*, siendo la proporción del nivel de transcritos de *FTEII* respecto a *1033* más alta en la cepa KM71/ $P_{1033}$ -*FTEII*- $T_{AOX1}$  en comparación con la cepa KM71/ $P_{1033}$ -*FTEII*- $T_{1033}$  en ambas fuentes de carbono (Fig. 38).

Estos hallazgos indican un efecto competitivo en la transcripción de *FTEII* y *1033* cuando se usó el mismo promotor dos veces en la misma cepa, lo que provocó una disminución en el nivel de transcripción de uno de los genes regulado por el mismo promotor, como se reportó previamente para genes heterólogos regulados por  $P_{GAP}$  y el gen *GAPDH* endógeno (Dou et al. 2021). Este fenómeno también se vio reforzado por la mayor estabilidad de los transcritos de *FTEII* que albergan la región 3'UTR del gen *AOX1* que los transcritos de *1033* que albergan la región 3'UTR del gen *1033*, fenómeno previamente descrito para otras regiones 3'UTR (Ito et al. 2020).

Los niveles de transcripción del gen *GAPDH* endógeno en la cepa KM71/ $P_{1033}$ -*FTEII*- $T_{AOX1}$  fueron diferentes según la fase de crecimiento y la fuente de carbono, con niveles de transcripción más altos en las células cultivadas con glucosa que en las células cultivadas con glicerol durante la fase exponencial. Esto es consistente con  $P_{GAP}$ , que tradicionalmente ha sido catalogado como un promotor constitutivo fuerte con mayor actividad a altas tasas de crecimiento específicas (Looser et al. 2015) y en cultivos de glucosa en comparación con cultivos de glicerol (Waterham et al. 1997).

Los niveles de transcritos de *GAPDH* fueron más altos en la cepa KM71/ $P_{1033}$ -*FTEII*- $T_{AOX1}$  que en la KM71/ $P_{1033}$ -*FTEII*- $T_{1033}$  cuando los cultivos crecieron en glucosa. Estos

resultados indican que la presencia de  $T_{AOX1}$  en el casete de expresión genera una mayor actividad transcripcional para la combinación endógena de promotor de *GAP*/terminador de *GAP* ( $P_{GAP}$ - $T_{GAP}$ ) en cultivos de glucosa, en comparación con el  $T_{1033}$  en el casete de expresión.

La actividad de los elementos reguladores en trans de  $P_{GAP}$  en los cultivos de glucosa probablemente se vio afectada por la combinación de promotor-terminador del casete de expresión. Por lo tanto, la regulación génica heteróloga impulsada por la combinación  $P_{1033}$ - $T_{AOX1}$  en el casete de expresión presentó una menor competencia con la regulación *GAPDH* endógena ( $P_{GAP}$ - $T_{GAP}$ ) en cultivos de glucosa, en comparación con la regulación génica heteróloga impulsada por la combinación  $P_{1033}$ - $T_{1033}$ . Por el contrario, no se observaron diferencias en los niveles de transcrito de *GAPDH* en células cultivadas con glicerol entre las dos cepas, lo que indica que la regulación impulsada por las dos combinaciones heterólogas de promotor-terminador compitió de manera similar con la regulación de *GAPDH*.

Por lo tanto, especulamos que los elementos reguladores (es decir, promotor y terminador transcripcional) del gen *1033* están regulados de manera similar a las del gen *GAPDH* en cultivos de glucosa, afectando la regulación de *GAPDH* cuando dos copias del par promotor-terminador (es decir, endógeno y heterólogo) del gen *1033* está presente en el genoma de la levadura, tal como ha sido reportado previamente en la regulación dependiente de  $P_{GAP}$  de genes heterólogos (Dou et al. 2021).

Aunque los niveles más altos de transcripción del gen heterólogo se alcanzaron en células KM71/ $P_{1033}$ -FTEII- $T_{AOX1}$  cultivadas en glucosa, los valores más altos de  $Y_{p/x}$  se obtuvieron en los cultivos de glicerol para las dos cepas. Estos resultados concuerdan con reportes previos que concluyeron que los niveles de transcripción por sí mismos no son suficientes para predecir los niveles de expresión de la proteína (Liu et al. 2016; Dou et al. 2021). Nuestros resultados indican que el par promotor-terminador junto con la fuente de carbono afectan las vías de traducción y/o la secreción de proteínas.

Nosotros encontramos resultados similares con una cepa  $P_{GAP}$ - $T_{AOX1}$  usando el mismo gen reportero (Herrera-Estala et al. 2022), donde las condiciones de cultivo que produjeron bajos niveles de transcripción del gen heterólogo también arrojaron valores altos de  $Y_{p/x}$ ,

que se correlacionó con una regulación positiva de los genes *KAR2* y *PSAI-1* de la vía de secreción en lugar del incremento en la transcripción de los genes heterólogos.

Además, una estrategia de cultivo de una fase discontinua de glicerol seguido de una fase discontinua alimentada con glucosa para un sistema de *P. pastoris* regulada por el promotor *GAP* solo aumentó 1.1 veces el valor de  $Y_{p/x}$ , en comparación con la estrategia glicerol-glicerol (Garcia-Ortega et al. 2013), aunque  $P_{GAP}$  muestra mayor actividad transcripcional en glucosa que en cultivos con glicerol (Waterham et al. 1997). En conjunto, estos datos apuntan a que bajos niveles de transcripción de genes heterólogos en cultivos de glicerol contribuyen en un aumento de la traducción y/o secreción de proteína, al menos para la proteína reportera utilizada en este trabajo.

Las combinaciones de los elementos reguladores de los genes *1033* y *AOX1* de *P. pastoris* (es decir, los pares  $P_{1033}-T_{AOX1}$  y  $P_{1033}-T_{1033}$ ) contribuyeron a una actividad transcripcional diferenciada de los genes heterólogos y endógenos y que fueron dependientes de la fuente de carbono, lo que demuestra la influencia del terminador transcripcional en el casete de expresión en los niveles de transcritos de los genes heterólogos y endógenos. El par promotor-terminador y la fuente de carbono afectaron las vías de traducción y/o secreción de proteínas. Además, los bajos niveles de transcritos de genes heterólogos de las células cultivadas en glicerol contribuyeron a un aumento de la traducción y/o la secreción de proteínas. El  $P_{1033}$  en combinación con  $T_{AOX1}$  o  $T_{1033}$  podría usarse para producir proteínas tóxicas para el hospedero, co-expresar genes de la vía secretora y mantener un equilibrio metabólico en aplicaciones de ingeniería metabólica o biología sintética.

Este es el primer informe sobre la caracterización funcional del promotor y el terminador de la transcripción del gen *1033* de *P. pastoris* que amplía la caja de herramientas disponible para ajustar la expresión génica en *P. pastoris*.

## 9. CONCLUSIONES

1. Se identificaron y sintetizaron la secuencia promotora y terminadora de la transcripción del gen *1033* (PAS\_chr3\_1033) y se construyeron dos cepas de *P. pastoris* conteniendo un casete de expresión conformado por el promotor y región terminadora de la transcripción del gen *1033* o la región terminadora de la transcripción del gen *AOXI*, asociadas a un gen reportero (*FTEII*) y a una secuencia señal que permitió la secreción del producto proteico.
2. Se caracterizaron funcionalmente por primera vez el promotor (P<sub>1033</sub>) y el terminador transcripcional (T<sub>1033</sub>) del gen constitutivo *1033* de *P. pastoris* con actividad transcripcional débil, no dependiente de metanol, ya que se detectaron transcritos del gen reportero, además de la actividad y detección de la proteína recombinante en ambas cepas y fuentes de carbono evaluadas.
3. La secuencia del promotor del P<sub>1033</sub> tiene 175 nucleótidos y abarca toda la región entre los dos CDS. El sitio del inicio de la transcripción se identificó a -29 nucleótidos río arriba del codón de inicio de la traducción, posee un motivo de secuencia característico de regiones 5'UTR de eucariotas que es un SXL\_BS poli (dA-dT), la caja tipo TATA (TAAAAAGA) y el elemento INR de tipo mamífero (CTAATCG), están a -10 a -17 y +2 a + 8 nucleótidos del TSS, respectivamente.
4. La secuencia T<sub>1033</sub> tiene 157 nucleótidos que incluyen la región 3'UTR del gen *1033* (107 nucleótidos desde dónde termina el CDS del gen *1033* hasta el sitio de corte y poliadenilación CAA) más 50 nucleótidos adicionales.
5. Se identificaron 6 sitios de unión de factores de transcripción (SUFT) que se unen a la región de P<sub>1033</sub>, siendo Cat8-1 el factor de transcripción (FT), que actúa sobre P<sub>1033</sub> y P<sub>GAP</sub>, más relevante ya que el SUFT podría estar relacionada con la supuesta función de la proteína codificada por el gen *1033*, D-aminoácido oxidasa, que cataliza la oxidación de D-aminoácidos neutros a  $\alpha$ -cetoácidos, que pueden entrar en la gluconeogénesis, glucólisis y vías del ciclo de Krebs.
6. El P<sub>1033</sub> tiene un 2 a 3% de actividad transcripcional con respecto a P<sub>GAP</sub> y puede ajustarse o regularse con la fase de crecimiento celular y la fuente de carbono.
7. La actividad transcripcional de los genes heterólogos y endógenos evaluados en este estudio fue influenciada por la combinación promotor-terminador

transcripcional en el casete de expresión, como fue el caso en el que la combinación  $P_{1033}$ - $T_{AOX1}$  en el caste de expresión presentó niveles de transcritos más altos del gen heterólogo que la combinación de  $P_{1033}$ - $T_{1033}$  en glucosa, y no así en glicerol

8. La combinación de los elementos reguladores promotor-terminador evaluados, junto con la fuente de carbono afectan de forma diferenciada las vías de traducción y/o la secreción de proteínas heterólogas y en general la expresión de genes endógenos del hospedero.
9. Los niveles de transcripción por sí mismos no son suficientes para predecir los niveles de producción de la proteína heteróloga, ya que entran en juego los procesos de traducción y secreción. Tal como se presentó en este estudio, que, aunque los niveles más altos de transcripción del gen heterólogo se alcanzaron en las células  $KM71/P_{1033}$ - $FTEII$ - $T_{AOX1}$  cultivadas en glucosa, los valores más altos de  $Y_p/x$  se obtuvieron en los cultivos de glicerol para las dos cepas.
10. El  $P_{1033}$  en combinación con  $T_{AOX1}$  o  $T_{1033}$  podría usarse para modular la producción de proteínas tóxicas para el hospedero, co-expresar genes de la vía secretora y mantener un equilibrio metabólico en aplicaciones de ingeniería metabólica o biología sintética.

## 10. PERSPECTIVAS

En este trabajo se logró caracterizar funcionalmente el promotor y el terminador transcripcional del gen *1033* de *P. pastoris*, lo cual amplía la caja de herramientas disponible para ajustar la expresión génica en esta levadura. La caracterización de los elementos reguladores del gen *1033* se llevó a cabo en cultivos a nivel de matraz de las dos cepas construidas, el siguiente paso sería evaluar funcionalmente estos elementos reguladores y la combinación promotor-terminador de ambas cepas a escala de bioproceso en biorreactor, donde se puede tener mejor control de las variables del cultivo.

Las cepas construidas integraron el casete de expresión en el locus *HIS4* del genoma de *P. pastoris*, se podrían construir dos cepas con los mismos casetes de expresión, pero integrándolos en el cromosoma 3 cerca de la secuencia nativa del gen *1033* para poder evaluar el efecto de los elementos reguladores en trans sobre la funcionalidad de estas combinaciones promotor-terminador en ese sitio de inserción.

La actividad transcripcional de los genes heterólogos y endógenos evaluados en este estudio fue influenciada por la combinación promotor-terminador transcripcional en el casete de expresión, se podrían evaluar otras combinaciones del promotor P<sub>1033</sub> con terminadores usualmente usados en los casetes de expresión de *P. pastoris*, como el terminador transcripcional del gen *GAP* que se expresa de manera constitutiva y con altos niveles de expresión.

## APÉNDICES

### APÉNDICE A: Oligonucleótidos y sondas usados en este trabajo

Oligonucleótido	Secuencia
5ACT	GCCCCAGAAGAGCACCCAGT
3ACT	ACCGGAAGCGTACAGGGACA
5PTFTEII	TCCAAATCACTGACGGTCCAGA
3PTFTEII	TCACCGTCTTGAGCAACGAACA
5PTGAPDH	TTCCATCTTCCACTGGTGCT
3PTGAPDH	TTGGGACACGGAAAGCCAAA
5q1033	CATTCCCCGTCCGTTGAGCA
3q1033	TCCAACGAATCCTGCGACCG
5qYPT1	GCGGTCAAGCTAGTGCTTCT
3qYPT1	TGCTGACAAAACGTAAAACCA
PP <sub>1033</sub>	CGGTGTCGCTACTTCGAGA
FTE2	CTTGTGGGTGGAAACCAA
FTE1	CGCTGACAAGAACGGTTACA
3TH	GGTGCCGAGGATGACGATGA
Sondas	Secuencia
pFTEII	TTTGGGTTTCGGTTTGGGTCCAGA
pGAPDH	CCAGAATTGAACGGTAAGCTGA

## **APÉNDICE B: Preparación de medios, soluciones y reactivos**

### **Extracción y análisis de ácidos nucleicos:**

#### **SDS 10% (Dodecil Sulfato de Sodio)**

Disolver 10 g de SDS y aforar a 100 mL con agua bidestilada. Calentar a 68°C para facilitar su disolución.

#### **Buffer Tris-EDTA (TE pH 8.0, 1X) (Tris-HCl 10 mM pH 8.0, EDTA 1 mM pH 8.0)**

Mezclar 2 mL de EDTA 0.5 mM pH 8.0 con 10 mL de Tris-Cl pH 8.0). Aforar a 1L con agua destilada. Esterilizar por autoclave y almacenar a temperatura ambiente

#### **EDTA 500 mM pH 8.0**

Disolver 186.1 g de EDTA en 800 mL de agua bidestilada, ajustar el pH a 8.0 con NaOH 10N (o lentejas de NaOH). Aforar a 1000 mL. Esterilizar por autoclave, almacenar a temperatura ambiente. Nota: El EDTA no se disolverá hasta que la solución alcance el pH 8.0

#### **Agarosa 0.8%**

Mezclar 0.8 g de agarosa con 100 mL de TBE 1X. Calentar de 1-2 min. Vaciar en el molde. Dejar solidificar.

#### **Agarosa 2.5%**

Mezclar 2.5 g de agarosa con 100 mL de TBE 1X. Calentar de 1-2 min. Vaciar en el molde. Dejar solidificar.

#### **Buffer TBE, 10X**

Solución Stock 10X. Disolver 108 g de Trizma base, 55 g de ácido bórico y 9.3 g de EDTA en 600 mL de agua bidestilada, aforar a 1000 mL con agua bidestilada y esterilizar en autoclave. Almacenar a temperatura ambiente.

#### **Buffer TBE, 1X**



Solución de trabajo 1X: Mezclar 100 mL de TBE 10X y agregar 900 mL de agua bidestilada (Volumen final 1 litro). Almacenar a temperatura ambiente.

**SEVAG (Cloroformo-Alcohol isoamílico 24:1)**

Mezclar 24 mL de cloroformo y 1 mL de alcohol isoamílico. No esterilizar.

**NaOH (10 N)**

Disolver 10 g de NaOH anhidro en agua ultrapura a 25 mL. Esterilizar por autoclave. Almacenar a temperatura ambiente hasta su uso.

**Solución amortiguadora de lisis (TSNT) (Triton 100 2%, SDS 1%, NaCl 100 mM, Tris-Cl 100 mM pH 8.8, EDTA 1 mM)**

Disolver 500  $\mu$ L de Triton 100 en 15 mL de agua destilada, posteriormente disolver 146.25 mg de NaCl, agregar 2.5 mL de SDS al 10%, agregar 500  $\mu$ L de Tris-Cl 500 mM pH 8 y 50  $\mu$ L de EDTA 500 mM. Aforar a 25 mL. Esterilizar en autoclave. Almacenar a temperatura ambiente.

**Tris-Cl 500 mM pH 8**

Disolver 6.05 g de Tris en 50 mL. Ajustar el pH a 8 con HCl. Aforar a 100 mL con agua bidestilada. Esterilizar por autoclave. Almacenar a 4°C o a temperatura ambiente.

**NaCl 100 mM**

Colocar 5 mL de NaCl 1 M y aforar a 50 mL en un matraz de aforación. Almacenar a temperatura ambiente.

**Gel Red Nucleic Acid 100 X**

Disolver 20  $\mu$ L de Gel Red Nucleic Acid 10000 X (Solución stock) en un volumen final de 2 mL de NaCl 100 mM.

**Transformación en bacterias:**

**Médo SOC (Tryptona 2%, Extracto de levadura 0.5%, NaCl 10 mM, KCl 10 mM, Mg<sup>2+</sup> 20 mM, Glucosa 20 mM)**

Disolver 2 g de triptona, 0.5 g de extracto de levadura, 1 mL de NaCl 1 M y 1 mL de KCl 1 M en 97 mL de agua bidestilada. Agitar hasta disolver. Autoclavear y enfriar a temperatura ambiente. Añadir 1 mL de stock de Mg<sup>2+</sup> 2 M y 1 mL de glucosa 2M, cada uno de estos 2 últimos hasta alcanzar la concentración final de 20 mM. Aforar a 100 mL con agua bidestilada estéril. Almacenar a 4°C o a temperatura ambiente

**NaCl 1 M (Medio SOC, También usado para tinción de ácidos nucleicos)**

Disolver 5.84 g de NaCl y aforar a 50 mL con agua bidestilada. Esterilizar por filtración. Almacenar a 4°C o a temperatura ambiente

**KCl 1 M (Medio SOC)**

Disolver 3.72 g de KCl y aforar a 50 mL con agua bidestilada. Esterilizar por filtración. Almacenar a 4°C o a temperatura ambiente

**Glucosa 2 M (Medio SOC)**

Disolver 36.04 g de D-glucosa en 50 mL de agua bidestilada estéril. Aforar a 100 mL con agua bidestilada estéril. Esterilizar por filtración. Almacenar a 4°C o a temperatura ambiente.

**Stock de Mg<sup>2+</sup> 2M (Medio SOC)**

Disolver 9.52 g de MgCl<sub>2</sub> y 24.65 g de MgSO<sub>4</sub>·7H<sub>2</sub>O en 100 mL de agua bidestilada estéril. Esterilizar por filtración.

**CaCl<sub>2</sub> 1 M**

Disolver 11.1 g de CaCl<sub>2</sub> en 80 mL de agua bidestilada. Aforar a 100 mL, esterilizar por autoclave. Almacenar a temperatura ambiente o en frío.

**X-Gal (50 mg/mL)**

Se empleó una solución preparada por Promega (Promega Num. Cat. V3941). Esta solución se empleó directamente del frasco botella. Esta solución es almacenada a -20°C hasta su uso.

**IPTG (10 mM)**

Disolver 238 mg de IPTG (Promega Num. Cat V395A) en 10 mL de agua bidestilada estéril. Esterilizar por filtración y almacenar a -20°C.

**Medio Luria-Bertani (Bactotripton 1%, Extracto de levadura 0.5%, NaCl 1%)**

Disolver 10 g de bactotripton, 5 g de extracto de levadura y 10 g de NaCl en 1 Litro de agua bidestilada. Ajustar el pH a 7.5 con NaOH 10 M. Cuando se requiera agar, adicionar a la mezcla anterior 15 g de agar. Esterilizar por autoclave. Enfriar, agregar la ampicilina (100 µg/mL), homogenizar y vaciar 4 mL en tubos corning capacidad 15 mL, o en cajas de Petri. Almacenar a 4°C o temperatura ambiente.

**Ampicilina (25 µg/µL)**

Disolver 250 mg en 10 mL de agua bidestilada previamente esterilizada. Esterilizar por filtración. Almacenar a -20°C hasta su uso.

**Preparación de levaduras electrocompetentes:****Sorbitol 1M**

Disolver 9.1 g de sorbitol en 50 mL de agua bidestilada. Esterilizar por autoclave. Almacenar a temperatura ambiente y/o a 4°C.

**HEPES 1 M**

Disolver 2.61 g de HEPES sal de sodio (4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid sodium salt) en 10 mL de agua bidestilada. Esterilizar por filtración. Almacenar a 4°C hasta su uso.

### **DTT 1M**

Disolver 1.54 g de DTT en 10 mL de agua bidestilada. Esterilizar por filtración. Almacenar a 4°C hasta su uso.

### **Transformación en levaduras:**

#### **YPD/ YPD agar. (Extracto de levadura 1%, peptona 2%, glucosa 1%)**

Disolver 10 g de extracto de levadura, 20 g de peptona y 20 g de glucosa en 1000 mL de agua bidestilada. Esterilizar en autoclave a 15 lb/pulg<sup>2</sup>. Almacenar a temperatura ambiente o a 4°C. La vida media de esta solución es de varios meses. Cuando se requiera YPD agar, la mezcla anterior y además 20 g de agar, mezclar bien y esterilizar por autoclave. Dejar enfriar y vertir cuidadosamente en placas Petri. Almacenar a 4°C o a temperatura ambiente.

#### **10 XD (Glucosa 20%)**

Disolver 20 g de glucosa en 100 mL de agua destilada, esterilizar por filtración. Almacenar a 4°C.

#### **Medio RDB y RDBH (Sorbitol 1M, agar 2%, glucosa 2%, YNB 1.34%, biotina 4 x 10<sup>-5</sup>%, aminoácidos 0.005%, ± histidina 0.004%)**

Disolver 93 g de sorbitol y 10 g de agar bacteriológico en 350 mL de agua bidestilada. Esterilizar en autoclave, enfriar y mantener a 60°C. Esterilizar en autoclave 44 mL de agua bidestilada, enfriar y agregar 50 mL de dextrosa 20%, 50 mL YNB 10X, 1 mL de biotina 500X, 5 mL de aminoácidos 100X y precalentar a 45°C. \*Agregar 5 mL de Histidina 100X a la mezcla precalentada, para el crecimiento de cepas his4 (prueba de viabilidad). Adicionar la mezcla precalentada a 45°C, al sorbitol y agar mantenidos a 60°C (345 mL de agua estéril más sorbitol y agar). Homogenizar, vaciar en cajas de Petri, y sellar una vez que solidifica el medio. Almacenar a 4°C o temperatura ambiente.

#### **Histidina 100X Histidina (0.4%)**

Disolver 400 mg de L-histidina en 100 mL de agua destilada y esterilizar por filtración. Almacenar a 4°C. La vida media de esta solución es de aproximadamente 1 año.

**Reactivación en *Pichia pastoris*:**

**Buffer de fosfatos 1M, pH 6**

Disolver 1.03 g de  $K_2HPO_4$  y 12.8 g de  $KH_2PO_4$  en 80 mL de agua destilada, ajustar el pH con  $H_3PO_4$  o con KOH, aforar a 100 mL y esterilizar 20 min a 15 lb/pulg<sup>2</sup> a 121°C. Almacenar a temperatura ambiente.

**500X B. (Biotina 0.02%)**

Disolver 20 mg de biotina en 100 mL de agua destilada, esterilizar por filtración y almacenar a 4 °C.

**Glicerol 10%**

Mezclar 100 mL de glicerol absoluto y aforar a 1000 mL de agua destilada. Esterilizar por autoclave. Almacenar a 4°C.

**10X YNB. (Base de levadura nitrogenada con sulfato de amonio y sin aminoácidos 13.4%)**

Disolver 134 g de YNB en 1000 mL de agua destilada, esterilizar por filtración. Almacenar a 4°C

**Aminoácidos (AA 100X) (0.5% de cada Aminoácido)**

Disolver 500 mg de los siguientes aminoácidos: ácido L-glutámico, L-Metionina, L-Lisina, L-Leucina e Isoleucina en 100 mL de agua. Esterilizar por filtración y almacenar a 4°C. La vida media de esta solución es aproximadamente 1 año.

**Medio mínimo amortiguado y con glicerol o BMG. (Fosfato de potasio 100 mM pH 6.0 YNB 1.34%, glicerol 1%, biotina 4 x 10<sup>-5</sup>%)**

Para 100 mL de BMG, mezclar 10 mL de buffer de fosfato de potasio 1M pH 6.0, 10 mL de YNB 10X, 0.2 mL de biotina 500X, 10 mL de glicerol 10% y 70 mL de agua destilada estéril. Distribuir en alícuotas y poner de 48-72 h a prueba de esterilidad (30°C), antes de su uso.

**Medio mínimo amortiguado y con metanol o BMM (Fosfato de potasio 100 mM pH 6.0 YNB 1.34%, biotina 4 x 10<sup>-5</sup>%)**

Para 100 mL de BMM, mezclar 10 mL de buffer de fosfato de potasio 1M pH 6.0, 10 mL de YNB 10X, 0.2 ml de B500X, y 80 mL de agua destilada estéril. Distribuir en alícuotas y poner de 48 -72 h a prueba de esterilidad (30°C), antes de su uso.

**Determinación de proteínas totales:**

**Azul brillante de Coomassie (Proteínas totales)**

Disolver 100 mg del colorante azul brillante de Coomassie G-250 en 50 mL de etanol al 95%. Agitar vigorosamente. Agregar 100 mL de ácido fosfórico al 85 %. Mezclar vigorosamente por 20-30 min. Llevar a 1 L con agua destilada, agitar por 20-30 min y filtrar con papel Whatman no. 1 para eliminar el colorante no disuelto. Guardar el reactivo en un frasco oscuro a 4°C.

**Solución de albúmina sérica bovina (1 µg/µL)**

Pesar 5 mg de ASB en un matraz de aforación de 5 mL y aforar con agua destilada.

**Solución de albúmina sérica bovina (50 µg/mL)**

Colocar 250 µL de ASB de 1 µg/µL y aforar a 5 mL con agua destilada.

**Electroforesis SDS-PAGE**

**Azul brillante de Coomassie (SDS-PAGE) (Azul brillante de Coomassie R250 0.05%, metanol 50%, ácido acético 10%)**

Disolver 0.05 g de azul brillante de Coomassie R250 en 30 mL de agua bidestilada. Agregar 50 mL de metanol y 10 mL de ácido acético. Aforar a 100 mL. Almacenar a 4°C en un frasco ámbar o cubierto de la luz.

**Buffer de electroforesis SDS-PAGE 5X**

Pesar 3.8 g de Trizma base, 18.0 g de glicina, 1.3 g de SDS y disolver cuidadosamente en 150 mL de agua bidestilada. Aforar a 250 mL con agua bidestilada. Guardar en refrigeración. Diluir a 1X con agua bidestilada justo antes de usarse.

**Solución Fijadora (metanol 50%, ácido acético 10%)**

Mezclar 250 mL de metanol, con 50 mL de ácido acético y aforar con agua destilada a 500 mL. Almacenar a temperatura ambiente.

**Solución decolorante (metanol 5%, ácido acético 7%)**

Mezclar 50 mL de metanol con 70 mL de ácido acético, y aforar con agua destilada a 1000 mL. Almacenar a temperatura ambiente.

**Buffer Tris-Cl/SDS pH 8.8 4X (Solución tampón concentrada para SDS-PAGE)**

Disolver 9.1 g de Trizma base en 30 mL de agua bidestilada. Ajustar el pH a 8.8 con HCl concentrado. Aforar a 50 mL con agua bidestilada. Agregar 0.2 g de SDS sólido y disolver. Almacenar a 4°C hasta un mes.

**Buffer Tris-Cl/SDS pH 6.8 4X (Solución tampón concentrada para SDS-PAGE)**

Disolver 0.605 g de Trizma base en 5 mL de agua. Ajustar el pH a 6.8 con HCl concentrado. Aforar a 10 mL con agua destilada. Agregar 0.04 g de SDS sólido y disolver. Almacenar a 4°C hasta un mes.

**Acrilamida-bisacrilamida**

Se empleó una solución preparada por BIORAD (BIORAD Num. Cat. 161-0154) que contiene los siguientes componentes: Acrilamida 28.5%, N,N'-metilen-bis-acrilamida 1.5% (Radio 19:1). Esta solución se empleó directamente de la botella. Esta solución es almacenada a 4°C hasta su uso sin desgasificar para geles SDS-PAGE 12 o 15%.

**Jugo azul (Azul de bromofenol 0.25%, xilencianol 0.25%, glicerol 30%)**

Mezclar 25 mg de azul de bromofenol y 25 mg de xilencianol con 3 mL de glicerol. Aforar con agua bidestilada a 10 mL. Almacenar a 4°C o a temperatura ambiente.

**Persulfato de amonio (10%)**

Disolver 200 mg de persulfato de amonio en 2 mL de agua bidestilada. Almacenar a 4°C. La vida media de esta solución es aproximadamente 2 semanas.

## 11. BIBLIOGRAFÍA

- Adrio JL, Demain AL. 2014. Microbial enzymes: Tools for biotechnological processes. *Biomolecules* 4: 117–139.
- Ahmad M, Hirz M, Pichler H, Schwab H. 2014. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Applied Microbiology and Biotechnology* 98: 5301–5317.
- Ahn J, Hong J, Lee H, Park M, Lee E, Kim Ch, Choi E, Jung J, Lee H. 2007. Translation elongation factor 1- $\alpha$  gene from *Pichia pastoris*: Molecular cloning, sequence, and use of its promoter. *Applied Microbiology and Biotechnology* 74: 601–608.
- Ahn J, Hong J, Park M, Lee H, Lee E, Kim Ch, Lee J, Choi ES, Jung JK, Lee H. 2009. Phosphate-responsive promoter of a *Pichia pastoris* sodium phosphate symporter. *Applied and Environmental Microbiology* 75: 3528–3534.
- Akassou M, Groleau D. 2019. Advances and challenges in the production of extracellular thermotolerant pullulanases by wild-type and recombinant microorganisms: a review. *Critical Reviews in Biotechnology* 39: 337–350.
- Alonso J, Barredo JL, Díez B, Mellado E, Salto F, Garcia JL, Cortes E. 1998. D-amino-acid oxidase gene from *Rhodotorula gracilis* (*Rhodospiridium toruloides*) ATCC 26217. *Microbiology* 144: 1095–1101.
- Baghban R, Farajnia S, Rajabibazl M, Ghasemi Y, Mafi A. 2019. Yeast expression systems : overview and recent advances. *Molecular Biotechnology* 61:365–384.
- Barbay D, Mačáková M, Sützl L, De S, Mattanovich D, Gasser B. 2021. Two homologs of the cat8 transcription factor are involved in the regulation of ethanol utilization in *Komagataella phaffii*. *Current Genetics* 67: 641–61.
- Bradford MM. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* 72: 248–254.
- Çalik P, Ata O, Günes H, Massahi A, Boy E, Keskin A, Öztürk S, Zerze GH, Özdamar TH. 2015. Recombinant protein production in *Pichia pastoris* under glyceraldehyde-3-phosphate dehydrogenase promoter: from carbon source metabolism to bioreactor operation parameters. *Biochemical Engineering Journal* 95: 20–36.
- Cereghino JL, Cregg JM. 2000. Heterologous Protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiology Reviews* 24: 45–66.
- Chapman J, Ismail AE, Dinu CZ. 2018. Industrial applications of enzymes: recent advances, techniques, and outlooks. *Catalysts* 8: 20–29.
- Charoenrat T, Khumruangsri N, Promdonkoy P, Rattanaphan N, Eurwilaichitr L,



- Tanapongpipat S, Roongsawang N. 2013. Improvement of recombinant endoglucanase produced in *Pichia pastoris* KM71 through the use of synthetic medium for inoculum and pH control of proteolysis. *Journal of Bioscience and Bioengineering* 116: 193–198.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34: i884–890.
- Cos O, Serrano A, Montesinos JL, Ferrer P, Cregg JM, Valero F. 2005. Combined effect of the methanol utilization (Mut) phenotype and gene dosage on recombinant protein production in *Pichia pastoris* fed-batch cultures. *Journal of Biotechnology* 117: 321–335.
- Cregg JM, Tolstorukov II. 2012. “*P. pastoris* ADH promoter and use thereof to direct expression of proteins. Washington, DC: U.S. Patent and Trademark Office Patent No.: 222,386.
- Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman CA. 1989. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. *Molecular and Cellular Biology* 9: 1316–1323.
- Cregg JM, Cereghino JL, Shi J, Higgins DR. 2000. Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology* 16: 23–52.
- Curran KA, Karim AS, Gupta A, Alper HS. 2013. Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic Engineering* 19: 88–97.
- Curran KA. 2014. “Optimizing gene expression in *Saccharomyces cerevisiae* for metabolic engineering applications.”
- Daehwan K, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biology* 14(R36).
- Daly R, Hearn Milton T. 2005. Expression of heterologous proteins in *Pichia pastoris*: a useful experimental tool in protein engineering and production. *Journal of Molecular Recognition* 18: 119–138.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski Ch, Jha S, Batut, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-Seq aligner. *Bioinformatics* 29: 15–21.
- Donczew R, Hahn S. 2018. Mechanistic differences in transcription initiation at TATA-Less and TATA-Containing promoters *Molecular and Cellular Biology* 38: 1–14.
- Dou W, Zhu Q, Zhang M, Jia Z, Guan W. 2021. Screening and evaluation of the strong endogenous promoters in *Pichia pastoris*. *Microbial Cell Factories* 20: 1–12.
- Fickers, P. 2014. *Pichia pastoris*: a workhorse for recombinant protein production. *Current research in Microbiology and Biotechnology* 2: 354-363.

- García-Ortega X, Cámara E, Ferrer P, Albiol J, Montesinos-Seguí JL, Valero F. 2019. Rational development of bioprocess engineering strategies for recombinant protein production in *Pichia pastoris* (*Komagataella phaffii*) using the methanol-free *GAP* promoter. Where do we stand? *New Biotechnology* 53: 24–34.
- García-Ortega X, Ferrer P, Montesinos-Seguí JL, Valero F. 2013. Fed-batch operational strategies for recombinant Fab production with *Pichia pastoris* using the constitutive *GAP* promoter. *Biochemical Engineering Journal* 79: 172–181.
- Gershenzon NI, Ioshikhes IP. 2005. Synergy of human pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics* 21: 1295–1300.
- Graber JH. 2002. Probabilistic prediction of *Saccharomyces cerevisiae* mRNA 3'-processing sites. *Nucleic Acids Research* 30: 1851–1858.
- Guerrero-Olazarán M, Rodríguez-Blanco L, Viader-Salvadó JM. 2007. Producción de una fitasa recombinante en *Pichia pastoris*. *Ciencia UANL* 10: 413–418.
- Gupta I, Clauder-Münster S, Klaus B, Järvelin AI, Aiyar RS, Benes V, Wilkening S, Huber W, Pelechano V, Steinmetz LM. 2014. Alternative polyadenylation diversifies post-transcriptional regulation by selective RNA–Protein interactions. *Molecular Systems Biology* 10: 719-730.
- Van Helden J, André B, Collado-Vides J. 2000. A web site for the computational analysis of yeast regulatory sequences. *Yeast* 16: 177–187.
- Hernández Mendoza, DA. 2019. Análisis de la expresión de genes que codifican para posibles factores de transcripción que podrían regular la expresión del gen *GAPDH* en *Pichia pastoris*. Universidad Autónoma de Nuevo León.
- Herrera-Estala, AL. 2019. Respuesta fisiológica de cepas de *Pichia pastoris* a condiciones de sobreproducción de proteínas heterólogas reguladas por el promotor constitutivo *GAP*. Universidad Autónoma de Nuevo León.
- Herrera-Estala, AL, Fuentes-Garibay JA, Guerrero-Olazarán M, Viader-Salvadó JM. 2022. Low specific growth rate and temperature in fed-batch cultures of a beta-propeller phytase producing *Pichia pastoris* strain under *GAP* promoter trigger increased *KAR2* and *PSA1-1* gene expression yielding enhanced extracellular productivity." *Journal of Biotechnology* 352(March): 59–67.
- Highsmith J. 2015. *Biological Therapeutic Drugs: Technologies and Global Markets*. BCC Res.
- Inukai S, Hong Kock K, Bulyk ML. 2017. Transcription factor–DNA binding: beyond binding site motifs. *Current Opinion in Genetics and Development* 43: 110–119.
- Ito Y, Yamanishia M, Ikeuchib A, Imamurab Ch, Tokuhirob K, Kitagawac T, Matsuyama T. 2013. Characterization of five terminator regions that increase the protein yield of a transgene in *Saccharomyces cerevisiae*. *Journal of Biotechnology* 168: 486–492.
- Ito Y, Kitagawac T, Yamanishi M, Katahira S, Izawa S, Irie K, Furutani-Seiki M, Matsuyama T. 2016. Enhancement of protein production via the strong DIT1

- terminator and two RNA-Binding proteins in *Saccharomyces cerevisiae*. *Scientific Reports* 6: 1–9.
- Ito Y, Terai G, Ishigami M, Hashiba N, Nakamura Y, Bamba T, Kumokita R, Hasunuma T, Asai K, Ishii J, Kondo A. 2020. Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. *Nucleic Acids Research* 48: 13000–13012.
- Jansen A, Verstrepen KJ. 2011. Nucleosome positioning in *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* 75: 301–320.
- Li-Qun J, Wei-Rong J, Zhang-Chi M, Qi S, Xue C, Zhi-Qiang L, Yu-Guo Z. 2019. Promoter engineering strategies for the overproduction of valuable metabolites in microbes. *Applied Microbiology and Biotechnology* 103: 8725–8736.
- Kamal S, Khan SU, Muhammad N, Shoaib M, Omar M, Pascal K, Marie Rose M, Sun FF. 2018. Insights on heterologous expression of fungal cellulases in *Pichia pastoris*. *Biochemistry and Molecular Biology* 3: 15-35.
- Karbalaei M, Rezaee SA, Farsiani H. 2020. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *Journal of Cellular Physiology* 1:1-15.
- Kim D, Paggi JM, Park Ch, Bennett Ch, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* 37: 907–915.
- Küberl A, Schneider J, Thallinger GG, Anderl I, Wibberg D, Hajek T, Jaenicke S. 2011. High-quality genome sequence of *Pichia pastoris* CBS7435. *Journal of Biotechnology* 154: 312–320.
- Kuersten S, Goodwin EB. 2003. The power of the 3' UTR: translational control and development. *Nature Reviews Genetics* 4: 626–37.
- Kurtzman CP. 2009. Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. *Journal of Industrial Microbiology and Biotechnology* 36: 1435–1438.
- Lee JY, Chen H, Liu A, Alba BM., Lim ACh. 2017. Auto-induction of *Pichia pastoris* AOX1 promoter for membrane protein expression. *Protein Expression and Purification* 137: 7–12.
- Li R, Yu Ch, Li Y, Lam TW, Yiu SM, Kristiansen K, Wang J. 2009. SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics* 25: 1966–1967.
- Liang S, Zou Ch, Lin Y, Zhang X, Ye Y. 2013. Identification and characterization of *PGCW14*: a novel, strong constitutive promoter of *Pichia pastoris*. *Biotechnology Letters* 35: 1865–1871.
- Liu Y, Beyer A, Aebersold R. 2016. On the dependency of cellular protein levels on mRNA abundance.” *Cell* 165: 535–550.
- Looser V, Bruhlmann B, Bumbak F, Stenger C, Costa M, Camattari A, Fotiadis D, Kovar

- K. 2014. Cultivation strategies to enhance productivity of *Pichia pastoris*: a review. *Biotechnology Advances* 33: 1177–1193.
- Lu Z, Lin Z. 2021. The origin and evolution of a distinct mechanism of transcription initiation in yeasts. *Genome Research* 31: 51–63.
- Lyer V, Struhl K. 1995. Poly(DA:DT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO Journal* 14: 2570–2579.
- Matsuyama T. 2019. Recent developments in terminator technology in *Saccharomyces cerevisiae*. *Journal of Bioscience and Bioengineering* 128: 655–661.
- Matthews CB., Wright Ch, Kuo A, Colant N, Westoby M, Love JCh. 2017. Reexamining opportunities for therapeutic protein production in eukaryotic microorganisms. *Biotechnology and Bioengineering* 114: 2432–2444.
- Mavrich TN., Ioshikhes IP, Venters BJ, Jiang C, Tomsho LP, Qi J, Schuster SC, Albert I, Pugh BF. 2008. A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Research* 18: 1073–1083.
- Mayr Ch. 2019. What are 3' Utrs doing? *Cold Spring Harbor Perspectives in Biology* 11: 1-15.
- Menendez J, Valdes I, Cabrera N. 2003. The *ICLI* gene of *Pichia pastoris*, transcriptional regulation and use of its promoter. *Yeast* 20: 1097–1108.
- Mittal N, Pieters J, Nimwegen EV, Gruber AJ, Keller W, Müller P, Jayachandran R. 2014. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nature Communications* 5: 1-10.
- Mombeni M, Arjmand S, Omid S, Siadat R, Alizadeh H. 2020. Enzyme and microbial technology *PMOX*: a new powerful promoter for recombinant protein production in yeast *Pichia pastoris*. *Enzyme and Microbial Technology* 139: 1-8.
- Moreira de Almeida JR, Pepe de Moraes LM, Gonçalves Torres FA. 2005. Molecular characterization of the 3-phosphoglycerate kinase gene (*PGKI*) from the methylotrophic yeast *Pichia pastoris*. *Yeast* 22: 725–737.
- Nordén Kristina, Agemark M, Danielson JA, Alexandersson E, Kjellbom P, Johanson U. 2011. Increasing gene dosage greatly enhances recombinant expression of aquaporins in *Pichia pastoris*. *BMC Biotechnology* 11: 1-12.
- Pérez A, López C. 2013. Y predicción bioinformática de factores de transcripción involucrados en la regulación de miARNs en plantas identification of cis-regulatory elements and bioinformatic prediction of transcriptional factors involved in regulation of miRNAs in plants. *Acta Biológica Colombiana* 18: 107–120.
- Pesole G, Liuni S. 1999. Internet resources for the functional analysis of 5' and 3' untranslated regions of eukaryotic mRNAs. *Trends in Genetics* 15: 378.
- Pesole G, Liuni S, Grillo G, Licciulli F, Mignone F, Gissi C, Saccone C. 2002. UTRdb and UTRsite: specialized databases of sequences and functional elements of 5' and 3' untranslated regions of eukaryotic mRNAs. Update 2002. *Nucleic Acids Research*

30: 335–340.

- Pfaffl MW, Georgieva TM, Georgiev IP, Ontsouka E, Hageleit M, Blum JW. 2002. Real-time RT-PCR quantification of insulin-like growth factor (IGF)-1, IGF-1 receptor, IGF-2, IGF-2 receptor, Insulin receptor, growth hormone receptor, IGF-binding proteins 1, 2 and 3 in the bovine species. *Domestic Animal Endocrinology* 22: 91–102.
- Pfaffl MW. 2004. Quantification strategies in real-time PCR. In *A-Z of Quantitative PCR*, edited by S.A. Bustin, 87–112. La Jolla, CA.
- Pfaffl MW, Hageleit M. 2001. Validities of mRNA quantification using recombinant RNA and recombinant DNA external calibration curves in real-time RT-PCR. *Biotechnology Letters* 23: 275–282.
- Porro D, Gasser B, Fossati T, Maurer M, Branduardi P, Sauer M, Mattanovich D. 2011. Production of recombinant proteins and metabolites in yeasts. *Applied Microbiology and Biotechnology* 89: 939–948.
- Porro D, Mattanovich D. 2004. Recombinant protein production in yeasts. *Methods in Molecular Biology* (Clifton, N.J.) 267: 241–258.
- Potvin G, Ahmad A, Zhang Z. 2012. Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: a review. *Biochemical Engineering Journal* 64: 91–105.
- Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak Ch, Gasser B, Mattanovich D. 2013. Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microbial Cell Factories* 12: 1-10.
- Priest HD, Filichkin SA, Mockler TC. 2009. Cis-regulatory elements in plant cell signaling. *Current Opinion in Plant Biology* 12: 643–649.
- Ramakrishnan K, Pratiapati M, Samuel P, Sankaranarayanan M. 2020. Transcriptional control of gene expression in *Pichia pastoris* by manipulation of terminators.” *Applied Microbiology and Biotechnology* 104: 7841–7851.
- Raveh-Sadka Tali, Levo M, Shabi U, Shany B, Keren L, Lotan-Pompan M, Zeevi D, Sharon E, Weinberger A, Segal E. 2012. Manipulating nucleosome disfavoring sequences allows fine-tune regulation of gene expression in yeast. *Nature Genetics* 44: 743–750.
- Rebnegger C, Graf AB, Valli M, Steiger MG, Gasser B, Maurer M, Mattanovich D. 2014. In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. *Biotechnology Journal* 9: 511–525.
- Rhee HS, Pugh BF. 2011. Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution. *Cell* 147: 1408–1419.
- Roider HG, Kanhere A, Manke T, Vingron M. 2007. Predicting transcription factor affinities to DNA from a biophysical model. *Bioinformatics* 23: 134–141.
- Rollano Peñazola O, Mollinedo Portugal P. 2017. Bioinformatic analysis of Rna-Seq with

- a perspective for Bolivia. *Revista Boliviana de Química* 34: 50–55.
- Sakai Y, Yurimoto H, Matsuo H, Kato N. 1998. Regulation of peroxisomal proteins and organelle proliferation by multiple carbon sources in the methylotrophic yeast, *Candida boidinii*. *Yeast* 14: 1175–1187.
- Schmittgen, T. 2006. Quantitative gene expression by real-time PCR: a complete protocol. In *Real-Time PCR* (1st Edition). Taylor & Francis, pp. 155-166.
- Schüller HJ. 2003. Transcriptional control of nonfermentative metabolism in the yeast *Saccharomyces cerevisiae*. *Current Genetics* 43: 139–160.
- Schutter KD, Lin YCh, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, Rouzé P, De Peer YV, Callewaert N. 2009. Genome sequence of the recombinant protein production host *Pichia pastoris*. *Nature Biotechnology* 27: 561–566.
- Sears IB, O'Connor J, Rossanese OW, Glick BS. 1998. A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*. *Yeast* 14: 783–790.
- Segal, Widom J. 2009. Poly(DA:DT) Tracts: major determinants of nucleosome organization. *Current Opinion in Structural Biology* 19: 65–71.
- Shen S, Sulter G, Jeffries TW, Cregg JM. 1998. A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Gene* 216: 93–102.
- Stadlmayr G, Mecklenbräuker A, Rothmüller A, Maurer M, Sauer M, Mattanovich D, Gasser B. 2010. Identification and characterisation of novel *Pichia pastoris* promoters for heterologous protein production. *Journal of Biotechnology* 150: 519–529.
- Struhl K. 1985. Elements for constitutive transcription in yeast I- II1. *Biochemistry* 82: 8419–8423.
- Struhl K. 1986. Constitutive and inducible *Saccharomyces cerevisiae* promoters: evidence for two distinct molecular mechanisms. *Molecular and Cellular Biology* 6: 3847–3853.
- Theron ChW, Berrios J, Delvigne F, Fickers P. 2018. Integrating metabolic modeling and population heterogeneity analysis into optimizing recombinant protein production by *Komagataella (Pichia) pastoris*. *Applied Microbiology and Biotechnology* 102: 63–80.
- Thomas-Chollier M, Hufton A, Heinig M, O'Keefe S, El Masri N, Roider HG, Manke T, Vingron M. 2011. Transcription factor binding predictions using TRAP for the analysis of ChIP-Seq data and regulatory SNPs. *Nature Protocols* 6: 1860–1869.
- Thorvaldsdóttir H, Robinson JT, Mesirov JP. 2013. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* 14: 178–192.
- Turcotte B, Liang XB, Robert F, Soontorngun N. 2010. Transcriptional regulation of nonfermentable carbon utilization in budding yeast. *FEMS Yeast Research* 10: 2–13.

- Vanz AL, Lünsdorf H, Adnan A, Nimtz M, Gurramkonda Ch, Khanna N, Rinas U. 2012. Physiological response of *Pichia pastoris* GS115 to methanol-induced high level production of the Hepatitis B surface antigen: catabolic adaptation, stress responses, and autophagic processes. *Microbial Cell Factories* 11: 1–11.
- Viader-Salvadó, JM, Gallegos-López JA, Carreón-Treviño JG, Castillo-Galván M, Rojo-Domínguez A, Guerrero-Olazarán M. 2010. Design of thermostable beta-propeller phytases with activity over a broad range of pHs and their overproduction by *Pichia pastoris*. *Applied and Environmental Microbiology* 76: 6423–6430.
- Vogl T, Glieder A. 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnology* 30: 385–404.
- Vogl T, Sturmberger L, Kickenweiz T, Wasmayer R, Schmid Ch, Hatzl AM, Gerstmann MA. 2016. A toolbox of diverse promoters related to methanol utilization: functionally verified parts for heterologous pathway expression in *Pichia pastoris*. *ACS Synthetic Biology* 5: 172–186.
- Waterham HR, Digan ME, Koutz PJ, Lair SV, Cregg JM. 1997. Isolation of the *Pichia pastoris* Glyceraldehyde-3-Phosphate Dehydrogenase gene and regulation and use of its promoter. *Gene* 186: 37–44.
- Westers L, Westers H, Quax WJ. 2004. *Bacillus subtilis* as cell factory for pharmaceutical proteins: a biotechnological approach to optimize the host organism. *Biochimica et Biophysica Acta - Molecular Cell Research* 1694: 299–310.
- Wingender E, Dietze P, Karas H, Knüppel R. 1996. TRANSFAC: a database on transcription factors and their DNA binding sites. *Nucleic Acids Research* 24: 238–241.
- Wiseman DN, Otchere A, Patel JH, Uddin R, Pollock NL, Routledge SJ, Rothnie AJ, Slacka C, Poynera DR, Billa RM, Goddarda AD. 2020. Expression and purification of recombinant G Protein-coupled receptors: a review. *Protein Expression and Purification* 167: 105524.
- Yamanishi M, Ito Y, Kintaka R, Imamura Ch, Katahira S, Ikeuchi A, Moriya H, and Matsuyama T. 2013. A genome-wide activity assessment of terminator regions in *Saccharomyces cerevisiae* provides a ‘Terminatome’ toolbox. *ACS Synthetic Biology* 2: 337–347.
- Yang Ch, Bolotin E, Jiang T, Sladek FM, Martinez E. 2007. Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene* 389: 52–65.
- Yang Z, Zhang Z. 2018. Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris*: a review. *Biotechnology Advances* 36: 182–195.
- Yin J, Li G, Ren X, Herrler G. 2007. Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes. *Journal of Biotechnology* 127: 335–347.
- Yurimoto H, Hasegawa T, Sakai Y, Kato N. 2001. Characterization and high-level

production of D-amino acid oxidase in *Candida boidinii*. *Biosci Biotechnol Biochem* 65: 627–633.

Zhu T, Sun H, Li P, Xue Y, Li Y, Ma Y. 2014. Constitutive expression of alkaline  $\beta$ -mannanase in recombinant *Pichia pastoris*. *Process Biochemistry* 49: 2025–2029.