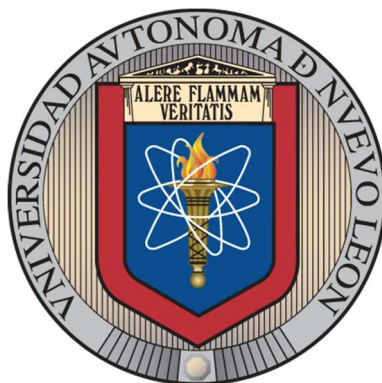


UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE CIENCIAS BIOLÓGICAS



EFECTO DE LA REGIÓN TERMINADORA DEL CASETE DE EXPRESIÓN EN LOS NIVELES DE EXPRESIÓN DEL GEN HETERÓLOGO *FTEII* REGULADO POR EL PROMOTOR DEL GEN *GAPDH* EN CEPAS RECOMBINANTES DE *Komagataella phaffii* (*Pichia pastoris*)

Por

MC. NANCY PENTÓN PIÑA

Como requisito parcial para obtener el Grado de
DOCTOR EN CIENCIAS CON ORIENTACIÓN
EN BIOTECNOLOGÍA

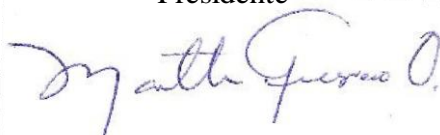
2026

EFFECTO DE LA REGIÓN TERMINADORA DEL CASETE DE EXPRESIÓN EN
LOS NIVELES DE EXPRESIÓN DEL GEN HETERÓLOGO *FTEII* REGULADO POR
EL PROMOTOR DEL GEN *GAPDH* EN CEPAS RECOMBINANTES DE
Komagataella phaffii (*Pichia pastoris*)


Comité de Tesis



DR. JOSÉ M. VIADER SALVADÓ
Presidente



DRA. MARTHA GUERRERO OLAZARÁN
Secretario



DR. BENITO PEREYRA ALFÉREZ
Vocal 1



DR. JUAN ANTONIO GALLEGOS LÓPEZ
Vocal 2



DR. HAMLET AVILÉS ARNAUT
Vocal 3



DRA. KATIUSKA AREVALO NIÑO
Subdirectora de Posgrado



SUBDIRECCIÓN
DE POSGRADO

EFFECTO DE LA REGIÓN TERMINADORA DEL CASETE DE EXPRESIÓN EN
LOS NIVELES DE EXPRESIÓN DEL GEN HETERÓLOGO *FTEII* REGULADO POR
EL PROMOTOR DEL GEN *GAPDH* EN CEPAS RECOMBINANTES DE
Komagataella phaffii (*Pichia pastoris*)

Dirección de Tesis



DR. JOSÉ MARÍA VIADER SALVADÓ
Director



DRA. MARTHA GUERRERO OLAZARÁN
Co-Director

AVISOS 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 videos, 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

El Consejo Nacional de Ciencia y Tecnología (CONACYT) financió los estudios de doctorado de la estudiante Nancy Pentón Piña, con número de matrícula de la UANL 1940832, mediante la beca identificada con número CVU 893044. Este proyecto fue financiado por el Fondo Sectorial de Investigación para la Educación de la Secretaría de Educación Pública (SEP) y el Consejo Nacional de Ciencia y Tecnología (CONACYT) con la clave CB-2016-01-286093.

Parte del presente trabajo se presentó en dos congresos nacionales y en un congreso internacional:

- N. Pentón-Piña, A.L. Herrera-Estala, M. Guerrero-Olazarán, J.M. Viader-Salvadó. Identificación de la secuencia del terminador transcripcional del gen *GAP* de *Pichia pastoris* (*Komagataella* spp.). XVIII Congreso Nacional de Biotecnología y Bioingeniería, Sociedad Mexicana de Biotecnología y Bioingeniería A.C., León, Guanajuato (México), 23-28 de junio de 2019.
- N. Pentón-Piña, M. Guerrero-Olazarán, J.M. Viader-Salvadó. Funcionalidad de la región terminadora del gen *GAP* en un casete de expresión con promotor *GAP* en cepas recombinantes de *Pichia pastoris* (*Komagataella* spp.). XIX Congreso Nacional de Biotecnología y Bioingeniería, Sociedad Mexicana de Biotecnología y Bioingeniería A.C. 27 de septiembre – 1 de octubre de 2021.
- M. Guerrero-Olazarán, N. Pentón-Piña, Y. Robainas-del-Pino, J.A. Fuentes-Garibay, J.M. Viader-Salvadó. Effect of *AOXI* and *GAP* transcriptional terminators on transcript levels of both the heterologous and the *GAPDH* genes and the extracellular $Y_{p/x}$ in *GAP* promoter-based *Komagataella phaffii* strains. Pichia 2024. San Diego, CA, USA. 24-27 de marzo de 2024.

Además, el presente trabajo se publicó el 26 de septiembre de 2024 en la revista *PeerJ*:

- J.M. Viader-Salvadó, N. Pentón-Piña, Y. Robainas-del-Pino, J.A. Fuentes-Garibay, M. Guerrero-Olazarán. 2024. Effect of *AOXI* and *GAP* transcriptional terminators on transcript levels of both the heterologous and the *GAPDH* genes and the extracellular $Y_{p/x}$ in *GAP* promoter-based *Komagataella phaffii* strains. PeerJ 12:e18181 <https://doi.org/10.7717/peerj.18181>.

AGRADECIMIENTOS

Al Fondo Sectorial de Investigación para la Educación de la Secretaría de Educación Pública (SEP) y al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el financiamiento del presente proyecto con la clave CB-2016-01-286093.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por otorgarme la beca que me permitió realizar este proyecto. Beca otorgada No. CVU 893044.

A la Universidad Autónoma de Nuevo León, México y al Instituto de Biotecnología por permitirme ser parte de su estudiantado con número de matrícula 1940832.

Al Dr. José María Viader Salvadó y a la Dra. Martha Guerrero Olazarán, por permitirme el honor de ser uno de sus estudiantes de doctorado, por sus orientaciones como directores de tesis, por sus enseñanzas, dedicación, paciencia y su gran esfuerzo para poder lograr la culminación del presente proyecto de tesis.

A mi familia y amistades que siempre me han apoyado y han estado al pendiente.

A mi mamá que siempre está presente cuando la he necesitado.

A la Dra. Socorro, a la Dra. Miriam por sus enseñanzas y constante atención a nuestra familia.

A los profesores y personal administrativo del IB que tanto nos ayudaron desde nuestra llegada a México y por ser parte importante en mi desarrollo profesional.

A mis compañeros de generación del Doctorado: Janeth, Daniel, Mirian y Fabiola.

A Yane por ser mi compañera de laboratorio, de generación del Doctorado y por su apoyo incondicional siempre que necesité de su ayuda profesional y personal.

A mis compañeros del laboratorio Ana Herrera, Dr. José Antonio, David, Karla, Daniela, Noé, Mariana, Albert, Fabián, Rebeca, Daniel, Diana, Uxue y Giselle.

A los profesores y estudiantes del laboratorio L9 y L4, donde tantas veces fuimos para usar sus equipos de laboratorio y guardar la conserva de nuestras clonas.

A todos lo que de algún modo contribuyeron al desarrollo del presente trabajo.

Gracias.

DEDICATORIAS

A mis niños

Quiero que se sientan orgullosos de su mami

Mis niños, no se pongan techo en su desarrollo profesional, todo es posible

Agarren alas y sean lo que quieran ser en la vida

A mi mamá

Mami no lo hubiera logrado sin ti

A mi esposo

Por haberme acompañado durante más de la mitad de mi vida

ÍNDICE

INDICE DE TABLAS	XI
INDICE DE FIGURAS.....	XIII
LISTA DE SIMBOLOS Y ABREVIATURAS.....	XV
RESUMEN	XVIII
ABSTRACT.....	XIX
1 INTRODUCCIÓN.....	1
2 ANTECEDENTES.....	3
2.1 Importancia de las proteínas recombinantes en la industria y exigencias del mercado	3
2.2 <i>K. phaffii</i> como hospedero para la producción de proteínas heterólogas	4
2.2.1 Importancia.....	4
2.2.2 Secreción de proteínas heterólogas	5
2.2.3 Cepas de <i>K. phaffii</i> utilizadas en la expresión de proteínas heterólogas.....	5
2.2.4 Vectores de expresión	7
2.3 Transformación.....	9
2.4 Integración de vectores en el genoma de <i>K. phaffii</i>.....	9
2.5 Estrategias para una selección eficiente de transformantes.....	10
2.6 Integración de vectores en el genoma de <i>K. phaffii</i>.....	10
2.7 Los promotores transcripcionales	12
2.8 Los terminadores transcripcionales	14
2.8.1 Generación de los extremos 3' del ARNm por cortes enzimáticos y poliadenilación	14
2.8.2 Funcionamiento de los terminadores transcripcionales poliadenilados	17
2.8.3 Señales de poliadenilación en la región 3'UTR	18
2.9 Modificaciones Postraduccionales.....	19
2.10 Cultivo celular. Fuente de carbono, velocidad específica de crecimiento y cinética de producción.....	20
2.11 Antecedentes directos	22
3 JUSTIFICACIÓN	24
4 HIPÓTESIS.....	25
5 OBJETIVO DEL TRABAJO.....	26
5.1 Objetivo general.....	26

5.2	Objetivos específicos.....	26
6	MATERIALES Y MÉTODOS.....	27
6.1	Lugar de trabajo.....	27
6.2	Estrategia general.....	27
6.3	Cepas, plásmidos, composición del medio, productos químicos y enzimas.	29
6.4	Secuencia del terminador transcripcional <i>GAPDH</i> (T_{GAP}).....	30
6.5	Modelo de señales de poliadenilación en T_{GAP}	32
6.6	Preparación de las células calcio competentes de la cepa TOP10 de <i>E. coli</i> .	32
6.7	Transformación de las células competentes de la cepa TOP10 de <i>E. coli</i>	32
6.8	Preparación del BCT de las células TOP10 de <i>E. coli</i>	33
6.9	Purificación de ADN plasmídico.....	33
6.10	Caracterización del plásmido pUCIDT GAP por digestión enzimática.....	34
6.11	Aislamiento y purificación del vector pGAHFTEII y el inserto T_{GAP} a escala preparativa.....	35
6.12	Construcción del vector pP GAP -FTEII- T_{GAP}	36
6.13	Caracterización del vector pP GAP -FTEII- T_{GAP}	37
6.14	Linealización y diafiltración del vector pP GAP -FTEII- T_{GAP}	38
6.15	Transformación de las células electrocompetentes de la cepa KM71 de <i>K. phaffii</i>	39
6.16	Caracterización de las cepas P GAP - T_{AOX1} y P GAP - T_{GAP} de <i>K. phaffii</i>	41
6.17	Funcionalidad de los elementos reguladores P GAP - T_{AOX1} y P GAP - T_{GAP}	42
6.18	Cinética de crecimiento.....	45
6.19	Determinación de dosis génica.....	45
6.20	Determinación de concentración de proteínas totales (método de Bradford)	48
6.21	Análisis de proteínas por SDS-PAGE.....	49
6.22	Determinación de actividad enzimática.....	50
6.23	Efecto del terminador transcripcional en el gen heterólogo sobre los niveles de transcripción del gen heterólogo <i>FTEII</i> y el gen endógeno <i>GAPDH</i> , y en el Yp/x extracelular.....	52
6.24	Análisis estadístico.....	54
7	RESULTADOS.....	55
7.1	Diseño y síntesis de la secuencia nucleotídica del terminador del gen <i>GAPDH</i> denominada T_{GAP}	55

7.2	Modelo de señales de poliadenilación en el terminador <i>GAPDH</i>	59
7.3	Construcción del vector de expresión pP_{GAP}-FTEII-T_{GAP}	60
7.3.1	Clonaje y digestión <i>in silico</i>	60
7.3.2	Ligación vector-inserto	61
7.4	Construcción de las cepas de <i>K. phaffii</i> KM71/P_{GAP}-FTEII-T_{GAP} y KM71/P_{GAP}-FTEII-T_{AOX1}	65
7.4.1	Linealización con Sall	65
7.4.2	Transformación de la cepa KM71 de <i>K. phaffii</i> con los vectores pGAHFTEII y pP _{GAP} -FTEII-T _{GAP}	66
7.5	Cultivo celular	70
7.6	Selección de clonas posibles unicopias	72
7.7	Determinación de la dosis génica	74
7.8	Funcionalidad de los elementos reguladores P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}	75
7.8.1	Detección de transcritos del gen heterólogo <i>FTEII</i> por RT-PCR	75
7.8.2	Detección de producto proteico del gen heterólogo <i>FTEII</i>	76
7.9	Efecto del terminador transcripcional en el gen heterólogo sobre los niveles de transcripción del gen heterólogo <i>FTEII</i> y el gen endógeno <i>GAPDH</i>, y en el Yp/x extracelular	78
7.9.1	Cinética de crecimiento	78
7.9.2	Niveles de transcripción de <i>FTEII</i> y <i>GAPDH</i> y Y p/x extracelular a lo largo del tiempo de cultivo	79
7.9.3	Comparación de los niveles de transcripción de <i>FTEII</i> y <i>GAPDH</i> de la misma cepa cultivada en la misma fuente de carbono	81
7.9.4	Efecto de la fuente de carbono en los niveles de transcripción de <i>FTEII</i> y <i>GAPDH</i>	82
7.9.5	Comparación de los niveles de transcripción entre cepas cultivadas en la misma fuente de carbono	82
7.9.6	Comparación de la Yp/x extracelular entre cepas y el efecto de la fuente de carbono	83
7.9.7	Comparación de los cocientes de niveles de transcripción de <i>FTEII</i> y cocientes de Yp/x extracelulares	83
8	DISCUSIÓN	86
9	CONCLUSIONES	102
10	PERSPECTIVAS	106
11	BIBLIOGRAFÍA	107

INDICE DE TABLAS

TABLA 1. CEPAS DE <i>K. PHAFFII</i> UTILIZADAS EN LA EXPRESIÓN DE PROTEÍNAS HETERÓLOGAS	6
TABLA 2. VECTORES DE EXPRESIÓN UTILIZADOS EN LA EXPRESIÓN DE PROTEÍNAS HETERÓLOGAS EN <i>K. PHAFFII</i>	8
TABLA 3. PROMOTORES MÁS FRECUENTES UTILIZADOS PARA LA PRODUCCIÓN DE PROTEÍNAS RECOMBINANTES EN LEVADURAS	12
TABLA 4. VOLÚMENES DE CADA COMPONENTE DE LA MEZCLA DE REACCIÓN PARA LA CARACTERIZACIÓN POR DIGESTIÓN ENZIMÁTICA DE pUCIDT-T _{GAP}	34
TABLA 5. TAMAÑO ESPERADO DE LOS FRAGMENTOS DE INTERÉS PARA LA CARACTERIZACIÓN DE pUCIDT-AMP POR DIGESTIÓN ENZIMÁTICA CON NOTI-HF Y/O BSU36I	35
TABLA 6. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN PARA LA DIGESTIÓN ENZIMÁTICA CON NOTI-HF Y/O BSU36I	35
TABLA 7. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN PARA LA LIGACIÓN DEL VECTOR DIGERIDO Y EL INSERTO	36
TABLA 8. MEZCLA DE REACCIÓN PARA CARACTERIZAR EL VECTOR pP _{GAP} -FTEII-T _{GAP} POR PCR.....	37
TABLA 9. CONDICIONES DEL PROGRAMA DE PCR PARA CARACTERIZAR EL VECTOR pP _{GAP} -FTEII-T _{GAP}	37
TABLA 10. MEZCLA DE REACCIÓN PARA LINEALIZAR LOS VECTORES pP _{GAP} -FTEII-T _{GAP} O pGAHFTEII CON SALI	38
TABLA 11. CONDICIONES DE OPERACIÓN DEL ELECTROPORADOR PARA LA TRANSFORMACIÓN DE CÉLULAS ELECTROCOMPETENTES DE LA CEPA KM71 DE <i>K. PHAFFII</i>	40
TABLA 12. CONDICIONES DEL PROGRAMA DE PCR PARA CARACTERIZAR LA CEPA P _{GAP} -T _{AOX1} Y P _{GAP} -T _{GAP} CON DOBLE PAREJA DE OLIGONUCLEÓTIDOS	41
TABLA 13. TAMAÑO ESPERADO DE LOS FRAGMENTOS AMPLIFICADOS EN EL PCR DE INTEGRACIÓN CON DOBLE PAREJA DE OLIGONUCLEÓTIDOS.....	42
TABLA 14. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN PARA EL TRATAMIENTO DEL RNA TOTAL CON DNAsa RQ1	43
TABLA 15. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN LA RETRO TRANSCRIPCIÓN DEL ARNm	43
TABLA 16. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN PARA EL PCR DE EVALUACIÓN DE FUNCIONALIDAD DE LOS ELEMENTOS REGULADORES P _{GAP} -T _{AOX1} Y P _{GAP} -T _{GAP}	44
TABLA 17. CONDICIONES DEL PROGRAMA DEL RT-PCR.....	44
TABLA 18. COMPOSICIÓN DE LA MEZCLA DE REACCIÓN PARA EL qPCR DE DETERMINACIÓN DE DOSIS GÉNICA	47
TABLA 19. CURVA DE CALIBRACIÓN DEL MÉTODO DE BRADFORD CON BSA	48
TABLA 20. COMPOSICIÓN DE REACTIVOS QUE CONFORMAN EL GEL SEPARADOR PREPARADO AL 12 % DE ACRILAMIDA/BISACRILAMIDA PARA LA ELECTROFORESIS DE PROTEÍNA POR SDS-PAGE	49

TABLA 21. COMPOSICIÓN DE REACTIVOS QUE CONFORMAN EL GEL CONCENTRADOR DE LA ELECTROFORESIS DE PROTEÍNA POR SDS-PAGE	49
TABLA 22. VOLÚMENES PARA LA PREPARACIÓN DE LOS ESTÁNDARES PARA LA CURVA DE CALIBRACIÓN.....	50
TABLA 23. MEZCLA DE REACCIÓN PARA LA DETERMINACIÓN DE ACTIVIDAD DE FITASA ..	51
TABLA 24. TAMAÑO ESPERADO DE LOS FRAGMENTOS GENERADOS PRODUCTO DE LA DIGESTIÓN DE PUCIDT _{GAP} CON LAS ENZIMAS NOTI-HF Y/O BSU36I.	58
TABLA 25. DOSIS GÉNICA DE LAS CLONAS MULTICOPIAS Y UNICOPIAS DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	75
TABLA 26. DETECCIÓN DE PROTEÍNAS TOTALES, RENDIMIENTO PRODUCTO POR UNIDAD DE BIOMASA Y ACTIVIDAD ENZIMÁTICA DE FTEII EN LOS SOBRENADANTES DE CULTIVO DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	78

INDICE DE FIGURAS

FIGURA 1. LA SECUENCIA AAUAAA ES NECESARIA PARA EL CORTE ENZIMÁTICO Y LA POLIADENILACIÓN.....	14
FIGURA 2. COMPLEJO DE PROCESAMIENTO DEL EXTREMO 3' DEL ARN.	16
FIGURA 3. FACTORES DE INICIACIÓN DE LA TRADUCCIÓN.	17
FIGURA 4. FASES DE CRECIMIENTO MICROBIANO.	20
FIGURA 5. REPRESENTACIÓN DE LA VELOCIDAD ESPECÍFICA DE CRECIMIENTO.	21
FIGURA 6. ESTRATEGIA GENERAL DEL PROYECTO.....	29
FIGURA 7. MAPEO DEL TRANSCRIPTOMA DE TRES CEPAS KM71.	55
FIGURA 8. ELECTROFORESIS EN GEL DE AGAROSA AL 1.5 % DEL FRAGMENTO DE ADNc AMPLIFICADO, PRODUCTO DE LA PCR DEL 3'RACE.....	56
FIGURA 9. ELECTROFEROGRAMA Y SECUENCIA PRODUCTO DE LA SECUENCIACIÓN DEL FRAGMENTO AMPLIFICADO POR PCR RESULTANTE DEL 3'RACE.	56
FIGURA 10. REPRESENTACIÓN ESQUEMÁTICA DEL PLÁSMIDO pUCIDT _{GAP}	57
FIGURA 11. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % PARA CARACTERIZAR EL PLÁSMIDO pUCIDT _{GAP} PROVENIENTE DE LA TRANSFORMACIÓN EN LA CEPA TOP10 DE E. COLI.....	58
FIGURA 12. REPRESENTACIÓN ESQUEMÁTICA DE LA SECUENCIA T _{GAP}	59
FIGURA 13. REPRESENTACIÓN ESQUEMÁTICA DEL VECTOR DE EXPRESIÓN pGAHFTEII. .	60
FIGURA 14. REPRESENTACIÓN ESQUEMÁTICA DEL VECTOR DE EXPRESIÓN pP _{GAP} -FTEII-T _{GAP}	60
FIGURA 15. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % A ESCALA PREPARATIVA DE LOS FRAGMENTOS DE INTERÉS OBTENIDOS DE LA DOBLE DIGESTIÓN ENZIMÁTICA CON NOTIHF Y Bsu36I (NEB, UK) DE LOS PLÁSMIDOS PURIFICADOS.	61
FIGURA 16. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % DE LOS FRAGMENTOS PURIFICADOS (VECTOR E INSERTO) OBTENIDOS A ESCALA PREPARATIVA.	62
FIGURA 17. SIEMBRA EN PLACAS DE LA CEPA TOP10 DE E. COLI TRANSFORMADAS CON EL VECTOR pP _{GAP} -FTEII-T _{GAP}	62
FIGURA 18. REPRESENTACIÓN ESQUEMÁTICA DEL CASETE DE EXPRESIÓN DE LOS VECTORES pGAHFTEII (P _{GAP} -T _{AOX1}) Y pP _{GAP} -FTEII-T _{GAP} (P _{GAP} -T _{GAP}).	63
FIGURA 19. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % PARA LA CARACTERIZACIÓN DE pP _{GAP} -FTEII-T _{GAP}	63
FIGURA 20. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % DEL PLÁSMIDO pP _{GAP} -FTEII-T _{GAP}	64
FIGURA 21. RESULTADO DE LA COMPARACIÓN POR BLASTN DE LAS SECUENCIAS NUCLEOTÍDICAS DEL TERMINADOR GAPDH (TDH3) DE LA CEPA CBS 7435 DE K. PHAFFII Y LA SECUENCIA OBTENIDA POR SECUENCIACIÓN DE pP _{GAP} -FTEII-T _{GAP}	64
FIGURA 22. ELECTROFORESIS DE AGAROSA AL 0.8 % DE LOS PLÁSMIDOS PURIFICADOS: A. pGAHFTEII Y B. pP _{GAP} -FTEII-T _{GAP}	65
FIGURA 23. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % DE LA LINEALIZACIÓN CON SALI.....	66

FIGURA 24. SIEMBRA EN PLACAS DE LA CEPA KM71 DE <i>K. PHAFFII</i> TRANSFORMADA CON EL PLÁSMIDO P _{GAP} H _{FTEII}	66
FIGURA 25. SIEMBRA EN PLACAS CON LA CEPA KM71 DE <i>K. PHAFFII</i> TRANSFORMADA CON EL PLÁSMIDO P _{GAP} -F _{FTEII} -T _{GAP}	67
FIGURA 26. A. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % DE LA CARACTERIZACIÓN DE LA CEPA P _{GAP} -T _{AOXI} POR PCR EMPLEANDO EL PAR DE OLIGONUCLEÓTIDOS FTE1/3TH, DIRIGIDOS A LA REGIÓN DEL TERMINADOR DEL GEN HETERÓLOGO.	67
FIGURA 27. A. ELECTROFORESIS EN GEL DE AGAROSA AL 0.8 % DE LA CARACTERIZACIÓN DE LA CEPA P _{GAP} -T _{GAP} POR PCR EMPLEANDO CON EL PAR DE OLIGONUCLEÓTIDOS FTE1/3TH, DIRIGIDOS A LA REGIÓN DEL TERMINADOR DEL GEN HETERÓLOGO.	68
FIGURA 28. A. ELECTROFORESIS EN GEL DE AGAROSA PARA LA VERIFICACIÓN POR PCR DE LA INTEGRACIÓN DEL CASETE DE EXPRESIÓN EN EL GENOMA DE <i>K. PHAFFII</i>	69
FIGURA 29. GRÁFICA DEL LOGARITMO NATURAL DE LA DENSIDAD CELULAR (DC) EN LA FASE EXPONENCIAL DEL CULTIVO CELULAR DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	70
FIGURA 30. A Y B. GRÁFICAS DE DISPERSIÓN DEL RENDIMIENTO DE PROTEÍNA TOTALES POR UNIDAD DE BIOMASA VS DC.	71
FIGURA 31. GRÁFICA DE CAJAS Y BIGOTES PARA EL ANÁLISIS DE LA VARIABILIDAD DEL RENDIMIENTO PROTEÍNA/BIOMASA ENTRE 18 CLONAS DE LA CEPA P _{GAP} -T _{AOXI} Y 25 CLONAS DE LA CEPA P _{GAP} -T _{GAP}	73
FIGURA 32. RECTAS DE EFICIENCIAS DEL qPCR. A. PARA EL GEN HETERÓLOGO FTEII Y B. PARA EL GEN GAPDH ENDÓGENO.	74
FIGURA 33. ELECTROFORESIS EN GEL DE AGAROSA PARA VERIFICAR LA CALIDAD DEL ARN EXTRAÍDO DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	75
FIGURA 34. A. ELECTROFORESIS EN GEL DE AGAROSA AL 2 % PARA DETERMINACIÓN DE FUNCIONALIDAD.	76
FIGURA 35. ELECTROFORESIS EN GEL DE POLIACRILAMIDA AL 12% DEL SOBRENADANTE LIBRE DE CÉLULAS DEL CULTIVO DE LAS CEPAS BASADAS EN P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	77
FIGURA 36. PERFIL CROMATOGRÁFICO DE LA CROMATOGRAFÍA DE EXCLUSIÓN MOLECULAR EN COLUMNA PD-10 SEPHADEX G25 DEL SOBRENADANTE DE CULTIVO DE LAS CEPAS.	77
FIGURA 37. A. CINÉTICA DE CRECIMIENTO Y B. VELOCIDAD ESPECÍFICA DE CRECIMIENTO (μ) EN DIFERENTES TIEMPOS DE CULTIVO PARA P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP} EN GLUCOSA (BMGLC) Y GLICEROL (BMGLY).	79
FIGURA 38. NIVELES DE TRANSCRITOS DE FTEII Y GAPDH DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	80
FIGURA 39. COCIENTES DE NIVELES DE TRANSCRITOS DE FTEII Y COCIENTES DE Y _{P/X} EXTRACELULAR DE LA CEPA P _{GAP} -T _{AOXI} VS. P _{GAP} -T _{GAP} CULTIVADAS EN GLUCOSA (BMGLC) Y GLICEROL (BMGLY).	84
FIGURA 40. COCIENTES DE NIVELES DE TRANSCRITOS DE FTEII Y COCIENTES DE Y _{P/X} EXTRACELULAR RESPECTO A LAS FUENTES DE CARBONO GLUCOSA VS GLICEROL DE LAS CEPAS P _{GAP} -T _{AOXI} Y P _{GAP} -T _{GAP}	85

LISTA DE SIMBOLOS Y ABREVIATURAS

Abreviatura	Significado
aa	Aminoácido
ADN	Ácido desoxirribonucleico
ADNc	Ácido desoxirribonucleico complementario
ADNg	Ácido desoxirribonucleico genómico
Amp	Ampicilina
<i>AOXI</i>	Gen codificante para la alcohol oxidasa 1
BCT	Banco de células de trabajo
BMG	Medio mínimo amortiguado con glicerol
BSA	Albúmina de suero bovino
CaCl ₂	Cloruro de calcio
Ct	Ciclo umbral de fluorescencia
DC	Densidad celular
DO _{600nm}	Densidad óptica a 600 nm
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	Ácido etilendiaminotetra-acético
EtOH	Alcohol etílico
GRAS	“Generally recognized as safe”: Generalmente reconocido como seguro
HCl	Ácido clorhídrico
HEPES	(ácido 4- (2-hidroxietil) -1-piperazinetanosulfónico) es un agente tampón de ácido sulfónico de ion híbrido
<i>HIS4</i>	Gen de la histidinol deshidrogenasa defectuoso o mutado
kb	Kilobase = Mil pares de bases (10 ³ pb)
kDa	Kilodaltones (10 ³ Da)
KM71/P _{GAP} -FTEII-T _{AOXI}	Cepa recombinante de <i>K. phaffii</i> productora y secretora de una fitasa termoestable regulada por el promotor constitutivo <i>GAP</i> y terminador <i>AOXI</i>
KM71/P _{GAP} -FTEII-T _{GAP}	Cepa recombinante de <i>K. phaffii</i> productora y secretora de una fitasa termoestable regulada por el promotor y terminador <i>GAP</i>
LB	Medio de cultivo Luria-Bertani
M	Marcador de peso molecular o concentración expresada en molar
mA	Miliamperios
MeOH	Metanol

MLFB (en inglés)	Lote alimentado con metanol limitante
mM	Concentración milimolar (10^{-3} M)
Mut	Fenotipo de utilización de metanol
Mut ⁺	Fenotipo Mut positivo (posee los dos genes AOX funcionales)
Mut ^s	Mutante que carece del gen <i>AOX1</i> funcional
NCBI (por sus siglas en inglés)	Centro Nacional de Información Biotecnológica
ng	Nanogramos (10^{-9} g)
NL	Nuevo León
nm	Nanómetro (10^{-9} m)
nt	Nucleotidos
pb	Pares de base
PCR	Reacción en cadena de la polimerasa
pH	Potencial de iones hidrógeno
PIB	Producto interno bruto
PSA	Persulfato de amonio.
PSA	Persulfato de amonio
Qp	Productividad volumétrica
RBD	Medio de regeneración con dextrosa
RBD His ⁻	Medio de regeneración con dextrosa sin histidina
RE	Retículo endoplásmico
RP-HPLC o HPLC-RP	Cromatografía líquida de alta resolución en fase reversa.
rpm	Revoluciones por minuto
SDS	Dodecilsulfato de sodio
SDS-PAGE	Electroforesis en gel de poliacrilamida en condiciones desnaturalizantes
SEVAG	Cloroformo-Alcohol isoamílico 24:1
SMC	Sitio múltiple de clonación
SOC o SOB	“Super Optimal Broth”: Medio de crecimiento bacteriano rico en nutrientes utilizado para el cultivo microbiológico, generalmente de <i>Escherichia coli</i>
Sol. I	Tris-HCl, EDTA y glucosa
Sol. II	Solución de lisis celular
Sol. III	Solución neutralizante
TA	Temperatura ambiente (22-24°C)
TBE	Solución amortiguadora (Tris/ácido bórico/EDTA)

TDT	Ditiotreitol
TE	Solución amortiguadora (Tris/EDTA)
TEMED	Tetrametiletilendiamina
TFA	Ácido trifluoroacético
TSNT	Solución amortiguadora de lisis
UANL	Universidad Autónoma de Nuevo León
UI	Unidades internacionales
UV	Luz ultravioleta
<i>vs</i>	"versus"
<i>et al. (et alii)</i>	y otros
YNB	Base nitrogenada de levaduras
$Y_{p/x}$	Rendimiento de la proteína extracelular/biomasa
YPD	Extracto de levadura-peptona-dextrosa

RESUMEN

El promotor constitutivo del gen de la gliceraldehído-3-fosfato deshidrogenasa (*GAPDH*) se considera una opción adecuada para la producción sin metanol de proteínas heterólogas en *Komagataella phaffii* (syn. *Pichia pastoris*). El interés en la combinación promotor-terminador ha crecido al reconocerse el papel regulador de los terminadores transcripcionales. Aún no se ha caracterizado completamente, en cepas con promotor *GAPDH* (P_{GAP}), el efecto del terminador de alcohol oxidasa 1 (T_{AOX1}) o *GAPDH* (T_{GAP}) sobre la transcripción y el rendimiento producto/biomasa ($Y_{p/x}$). En este estudio, se diseñaron dos cepas de *K. phaffii*, cada una con una copia del cassette heterólogo con las combinaciones P_{GAP} - T_{AOX1} y P_{GAP} - T_{GAP} . Se compararon niveles de transcripción del gen reportero y del *GAPDH* endógeno y los valores de $Y_{p/x}$ extracelular, estudiados en medios de cultivos suplementados con glucosa o glicerol y en tres tiempos de crecimiento celular. Se observó que la regulación transcripcional y los valores de $Y_{p/x}$ dependieron de la combinación P_{GAP} -terminador, de la fuente de carbono y de la etapa de crecimiento. Los niveles de transcripción más altos se obtuvieron en glucosa a alta velocidad de crecimiento (0.253 h^{-1}). Los valores de $Y_{p/x}$ extracelular aumentaron en el tiempo, alcanzando los valores máximos con glucosa y la combinación P_{GAP} - T_{AOX1} . El T_{AOX1} o T_{GAP} en la estructura del gen heterólogo activaron distintos elementos reguladores génicos en cada cepa, modulando de forma diferencial la expresión del gen heterólogo respecto al gen *GAPDH* endógeno. El T_{AOX1} indujo una regulación competitiva de la actividad transcripcional, condujo a una mayor estabilidad del ARNm y desencadenó distintos mecanismos de regulación metabólica en comparación con T_{GAP} . Además, mejoró la traducción y/o la secreción de la proteína heteróloga a una alta velocidad específica de crecimiento (0.253 h^{-1}), mientras que T_{GAP} fue más efectivo para mejorar la actividad postranscripcional a una velocidad específica de crecimiento baja (0.030 h^{-1}), independientemente de la fuente de carbono. El mayor $Y_{p/x}$ extracelular se obtuvo con la cepa KM71/ P_{GAP} -FTEII- T_{AOX1} cuando el cultivo se realizó a una velocidad específica de crecimiento baja (0.030 h^{-1}) utilizando glucosa como fuente de carbono. Así, la optimización de elementos reguladores y de estrategias de crecimiento son parámetros claves para mejorar la producción de biomoléculas de interés en *K. phaffii*.

ABSTRACT

The constitutive promoter of the glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene is a suitable option for methanol-free production of heterologous proteins in *Komagataella phaffii* (syn. *Pichia pastoris*). Interest in promoter-terminator combination has increased after recognizing the regulatory role of transcriptional terminators to modulate the expression of heterologous genes. In strain based on the *GAPDH* promoter (P_{GAP}), the effect of the alcohol oxidase 1 gene terminator (T_{AOX1}) or the *GAPDH* (T_{GAP}) on transcriptional activity and extracellular product/biomass yield ($Y_{p/x}$) has not yet been fully characterized. In the present study, two strains of *K. phaffii* were constructed, each with a single copy of the heterologous gene structure formed by the combination P_{GAP} - T_{AOX1} and P_{GAP} - T_{GAP} . A comparative analysis was carried out in terms of the transcription levels of the reporter gene and the endogenous *GAPDH* gene, and the extracellular $Y_{p/x}$ values; studied in culture media supplemented with glucose or glycerol and in three different stages of cell growth. The results showed that both the regulation of the transcription of heterologous gene and the endogenous *GAPDH* gene, as well as the values of $Y_{p/x}$ were influenced by the P_{GAP} -transcriptional terminator combination, by the carbon source and by the stage of cell growth. The highest transcription levels of heterologous gene and endogenous *GAPDH* were observed in glucose cultures at a high specific growth rate (0.253 h^{-1}). The extracellular values of $Y_{p/x}$ showed an increasing trend as the culture time progressed, with the highest values observed in glucose cultures and with the P_{GAP} - T_{AOX1} combination. The T_{AOX1} or T_{GAP} within the heterologous gene structure activated distinct gene regulatory elements in each strain, leading to a differential modulation of gene regulation for the heterologous respect to *GAPDH* genes. T_{AOX1} induced a competitive transcriptional regulation, improved mRNA stability and improved translation and/or secretion activity of heterologous proteins at a high specific growth rate (0.253 h^{-1}), while T_{GAP} was more effective in enhancing post-transcriptional activity at a low specific growth rate (0.030 h^{-1}), regardless of carbon source. The highest extracellular $Y_{p/x}$ was obtained with the KM71/ P_{GAP} -FTEII- T_{AOX1} strain when the culture was performed at a low specific growth rate (0.030 h^{-1}) using glucose as a carbon source. Therefore, the optimization of regulatory elements and cell growth strategies are key parameters in improving the production of biomolecules of interest in *K. phaffii*.

1 INTRODUCCIÓN

El sistema de expresión de *Pichia pastoris*, reclasificada desde 2009 como *Komagataella phaffii* (*K. phaffii*) (Kurtzman 2009) ha emergido con gran interés en los últimos tiempos. Las ventajas que ofrece esta levadura se deben a su eficiente sistema de secreción de proteínas y su habilidad para realizar modificaciones postraduccionales en medio mínimo definidos, por lo que es actualmente uno de los hospederos más utilizados para la producción de proteínas recombinantes (Vieira-Gomes et al. 2018; Cereghino y Cregg 2000). La diversificación de su aplicación es, entre otros aspectos, por la existencia de vectores con promotores fuertes. El promotor constitutivo del gen de la gliceraldehído-3-fosfato deshidrogenasa (*GAPDH*) ha surgido como una opción adecuada para la producción de proteínas en sistemas de expresión de *K. phaffii* libres de metanol (Looser et al. 2015; García-Ortega et al. 2019). *GAPDH* es una enzima clave tanto en la glucólisis como en la gluconeogénesis, y por lo tanto en el metabolismo de la glucosa y del glicerol. Es por ello que tanto la glucosa como el glicerol constituyen fuentes de carbono que activan constitutivamente y a altos niveles el gen *GAPDH*.

A pesar del uso generalizado del sistemas basados en el promotor constitutivo P_{GAP} para la producción de proteínas en *K. phaffii*, aún no se ha caracterizado integralmente el efecto de los factores ambientales y moleculares sobre los resultados de la producción de proteínas. Entre estos factores el efecto de la fuente de carbono, la velocidad específica de crecimiento y la combinación de elementos reguladores (promotor-terminador) desempeñan un papel crucial en la modulación de la expresión de genes y en el rendimiento proteína/biomasa ($Y_{p/x}$). Sin embargo su efecto continua siendo motivo de debate. Algunos autores sugieren que la glucosa es la fuente de carbono más favorable, dado que la actividad transcripcional de P_{GAP} es mayor en células cultivadas en glucosa que en glicerol (Waterham et al. 1997; Cereghino y Cregg 2000). Otros autores proponen el glicerol como el sustrato más adecuado para la fase de cultivo en lote (“*batch*”), y la glucosa para la fase de cultivo alimentado (“*feed-batch*”) (García-Ortega et al. 2013), o recomiendan una velocidad específica de crecimiento (μ) elevada para obtener altos niveles de producción de proteína (García-Ortega et al. 2013; Looser et al. 2015). Sin

embargo, en nuestro laboratorio se ha observado que a μ baja se favorece la secreción de la proteína heteróloga y el rendimiento extracelular $Y_{p/x}$ (Herrera-Estala et al. 2022).

Recientemente los terminadores transcripcionales también han surgido como un factor molecular importante en la regulación de la expresión génica heteróloga, ya que además de su función principal en la terminación transcripcional, la región 3' no traducible (3'UTR) dentro de la secuencia del terminador influye en la estabilidad del ARNm, así como en las eficiencias transcripcional y traduccional (Mayr 2019; Kuersten y Goodwin 2003). Por lo tanto, la combinación de terminadores con un promotor adecuado ha demostrado ser una estrategia eficaz para sintonizar la expresión génica (Currant et al. 2013; Vogl et al. 2016; Ramakrishnan et al. 2020; Ito et al. 2020; Robainas-del-Pino et al. 2023). Los conocimientos sobre la funcionalidad de los terminadores de *K. phaffii* son todavía limitados (Ito et al. 2020). Se desconoce el efecto del terminador transcripcional del gen de la alcohol oxidasa 1 (T_{AOX1}) o del terminador transcripcional *GAP* (T_{GAP}) del gen *GAPDH* en la estructura génica heteróloga sobre la actividad transcripcional en una cepa basada en P_{GAP} , y no se ha caracterizado completamente el impacto sobre el rendimiento de producto extracelular/biomasa ($Y_{p/x}$). Esto se debe principalmente a que los vectores basados en P_{GAP} disponibles en el mercado sólo albergan el T_{AOX1} .

En este trabajo, identificamos una secuencia de T_{GAP} y construimos dos cepas de *K. phaffii*, cada una de las cuales contiene una sola copia de una combinación diferente de los elementos reguladores de ADN (P_{GAP} - T_{AOX1} o P_{GAP} - T_{GAP}) dentro de la estructura génica heteróloga para producir extracelularmente la fitasa FTEII. Además, evaluamos el impacto de las combinaciones de elementos reguladores, junto con la fuente de carbono (glucosa o glicerol) y la fase de crecimiento celular, en los niveles de transcripción del gen heterólogos *FTEII* y del gen *GAPDH* endógenos y de los valores extracelulares de $Y_{p/x}$ en las cepas recombinantes de *K. phaffii* denominadas KM71/ P_{GAP} -FTEII- T_{AOX1} y KM71/ P_{GAP} -FTEII- T_{GAP} . Los resultados indican que la regulación de la transcripción del gen heterólogo y el gen endógeno *GAPDH*, los valores $Y_{p/x}$ extracelular y la traducción y/o secreción de la proteína heteróloga estuvieron influenciados por la combinación P_{GAP} -terminador transcripcional, actuando como moduladores la fuente de carbono y la fase de crecimiento celular.

2 ANTECEDENTES

2.1 Importancia de las proteínas recombinantes en la industria y exigencias del mercado

La producción de proteínas recombinantes surgió a principios de la década de 1980 con el desarrollo de herramientas de ingeniería genética, que representaban una alternativa convincente a la extracción de proteínas a partir de fuentes naturales. Con el paso de los años, se hicieron posibles altos niveles de proteínas heterólogas en una variedad de hospederos que van desde la bacteria *Escherichia coli* (*E. coli*) hasta las células de mamíferos. La importancia de las proteínas recombinantes está representada por el tamaño de su mercado. Se prevé que el tamaño del mercado mundial de expresión de proteínas alcance los 4820 millones de dólares para 2029 desde los 3410 millones de dólares de 2024, a una tasa de crecimiento anual del 7.1% entre 2024 y 2029 (Markets 2024). Su aplicación se extiende a las industrias de alimentos, detergentes, papel, productos químicos y de cosméticos, siendo las empresas farmacéuticas y biotecnológicas el motor impulsor de este mercado denominado “mercado de expresión de proteínas”.

Este mercado exige de altos niveles de producción de las proteínas recombinantes. Para tal fin son varios los hospederos que se utilizan para la producción de proteínas recombinantes. La bacteria *E. coli* es ampliamente conocida como hospedero preferencial debido a la gran disponibilidad de herramientas genéticas, crecimiento rápido y técnicas simples de cultivo. Sin embargo, dependiendo de las propiedades de la proteína deseada, la utilización de *E. coli* podría ser laboriosa o incluso imposible, particularmente cuando se requiere de modificaciones postraduccionales. Por lo tanto, como alternativa se han utilizado hospederos eucariotas tales como levaduras, hongos filamentosos, células de insectos y líneas celulares de plantas y mamíferos. Entre ellos, las levaduras combinan la simplicidad de un organismo unicelular, tiene menores demandas nutricionales en comparación con las líneas celulares de insectos y mamíferos, y posee la capacidad de realizar la mayoría de las modificaciones postraduccionales requeridas para una proteína recombinante biológicamente activa. Por lo tanto, las levaduras se han utilizado para la producción de diversas proteínas farmacéuticas (Nielsen 2013). Se han comercializado varias proteínas recombinantes que usan a *Saccharomyces cerevisiae* como hospedero

(Jozala y Pessoa 2016). Sin embargo, en los últimos años, muchas proteínas comercialmente disponibles utilizan otras levaduras como hospederos. Entre ellas, *K. phaffii* ha llamado la atención por la clonación y expresión de más de 1000 proteínas heterólogas, como fragmentos de anticuerpos, precursor de la insulina, antígeno de superficie de la Hepatitis B, factor de necrosis tumoral, el antígeno de *cetomegalovirus pp* 55, enzimas líticas ampliamente utilizadas en la industria papelera, textil, agroalimentaria, química y farmacéutica (Vieira Gomes et al. 2018; Serrano-Rivero et al. 2016).

2.2 *K. phaffii* como hospedero para la producción de proteínas heterólogas

2.2.1 Importancia

K. phaffii se ha convertido en un sistema de gran éxito para la expresión de genes heterólogos. Varios factores han contribuido a su rápida aceptación, ya que cuenta con un promotor derivado del gen alcohol oxidasa I (*AOX1*) que es especialmente adecuado para la expresión controlada de genes heterólogos; es posible emplear las técnicas necesarias para la manipulación genética molecular de *S. cerevisiae*, uno de los sistemas experimentales mejor caracterizados de la biología moderna; crece en presencia de oxígeno, un factor fisiológico clave que facilita su cultivo a altas densidades celulares en relación con las levaduras fermentativas; y una decisión en 1993 de Phillips Petroleum Company de liberar el sistema de expresión de *K. phaffii* a los laboratorios de investigación académica, cuya consecuencia ha sido una explosión en la base de conocimiento del sistema. Más de 1000 proteínas diferentes se han producido con éxito en *K. phaffii*.

Como levadura, *K. phaffii* es un microorganismo unicelular fácil de manipular y cultivar. Sin embargo, también es un eucariota capaz de realizar muchas de las modificaciones postraduccionales de las células eucariotas superiores, como plegamiento, formación de enlaces disulfuro y glicosilación (Cereghino y Cregg 2000; Li et al. 2007). Así, muchas proteínas que terminan como cuerpos de inclusión inactivos en bacterias se producen como moléculas biológicamente activas en *K. phaffii*. El sistema de *K. phaffii* también se considera generalmente como más rápido, más fácil y menos costoso de usar (Macauley-Patrick et al. 2005) que los sistemas de expresión derivados de eucariotas superiores (Higgins y Cregg 1998), como los sistemas de células de cultivo de tejidos de insectos y

mamíferos, y por lo general posibilita mayores niveles de expresión (Romanos et al. 1992; Li et al. 2007).

2.2.2 Secreción de proteínas heterólogas

Con *K. phaffii*, las proteínas heterólogas pueden expresarse intracelularmente o secretarse en el medio. *K. phaffii* secreta niveles bajos de proteínas endógenas y debido a que su medio de cultivo no contiene proteínas añadidas, una proteína heteróloga secretada comprende la gran mayoría del total de proteínas en el medio. Por lo tanto, la secreción sirve como un primer paso importante en purificación, separando la proteína heteróloga del grueso de las proteínas celulares. La secreción requiere la presencia de una secuencia señal en la proteína heteróloga para dirigirla hacia la vía secretora. Diferentes secuencias señales de secreción se han utilizado con éxito, aunque el éxito ha sido variable. La secuencia prepro del factor alfa de *S. cerevisiae* se ha utilizado como señal de secreción con mucho éxito (Higgins y Cregg 1998).

2.2.3 Cepas de *K. phaffii* utilizadas en la expresión de proteínas heterólogas

Todas las cepas de expresión de *K. phaffii* son derivados de NRRL-Y 11430 (Laboratorios de Investigación Regional del Norte, Peoria, IL) (Tabla 1). La mayoría tiene una mutación en el gen de la histidinol deshidrogenasa (*HIS4*) para permitir la selección de vectores de expresión que contienen *HIS4* tras la transformación. También están disponibles otras combinaciones de genes biosintéticos/huéspedes mutantes auxotróficos, pero se usa con menos frecuencia. Todas estas cepas crecen en medios complejos, y requieren suplementación con histidina (u otro nutriente apropiado) para crecimiento en medios mínimos.

Hay tres tipos de cepas huésped disponibles resultante de deleciones en uno o ambos genes *AOX* que varían con respecto a su capacidad para utilizar metanol. Las cepas con genes *AOX* eliminados a veces son mejores productores de una proteína heteróloga que las cepas de tipo salvaje. Estas cepas también requieren mucho menos metanol para inducir la expresión, lo que puede ser útil en fermentación a gran escala donde una gran cantidad de metanol a veces se considera un riesgo de incendio importante (Higgins y Cregg 1998).

Sin embargo, uno de los sistemas de expresión más utilizado emplea la cepa GS115 (*his4*), que es de tipo salvaje con respecto a los genes *AOX1* y *AOX2* y crece en metanol, de forma similar a la de tipo salvaje (fenotipo de utilización de metanol más [*Mut*⁺]). KM71 (*his4 arg4 aox16::ARG4*) es una cepa en la que el gen cromosómico *AOX1* es mayoritariamente eliminado y reemplazado con el gen *ARG4* de *S. cerevisiae*. Como resultado, esta cepa debe depender del gen *AOX2* que es mucho más débil que *AOX1* y crece en metanol a un ritmo lento [fenotipo de utilización de la vía de metanol lenta (*Mut*^s)]. Con muchos vectores de expresión de *K. phaffii*, es posible insertar un casete de expresión y eliminar simultáneamente el gen *AOX1* de una cepa *Mut*⁺. La cepa MC100-3 (*his4 arg4 aox16::SARG4 aox26::Phis4*) tiene ambos genes *AOX* eliminados y es totalmente incapaz de crecer en metanol [fenotipo de utilización de la vía del metanol menos (*Mut*⁻)] (Serrano-Rivero et al., 2016; Higgins y Cregg 1998). Algunas proteínas heterólogas secretadas son inestables en el medio de cultivo de *K. phaffii* en el que son rápidamente degradados por proteasas. Principales proteasas vacuolares parecen ser un factor significativo en la degradación, particularmente en cultivos de fermentadores, debido al entorno de alta densidad celular en combinación con la lisis de un pequeño porcentaje de células. El uso de cepas huésped que son defectuosas en estas proteasas han demostrado ayudar a reducir la degradación en varios casos. SMD1163 (*his4 pep4 prb1*), SMD1165 (*his4 prb1*), y SMD1168 (*his4 pep4*) son cepas deficientes en proteasa que pueden proporcionar un entorno más adecuado para la expresión de ciertas proteínas heterólogas.

Tabla 1. Cepas de *K. phaffii* utilizadas en la expresión de proteínas heterólogas

Cepa	Genotipo	Fenotipo	Características
Y-11430	Silvestre	Silvestre	Consumo de metanol alto
X33	Silvestre	Silvestre	Consumo de metanol alto
GS115	<i>his4</i>	<i>Mut</i> ⁺ , <i>su</i> ⁻	Consumo de metanol alto, auxótrofas a histidina
KM71	<i>his4, arg4, aox1::SARG4</i>	<i>Mut</i> ^s , <i>his</i> ⁻	Consumo de metanol bajo, auxótrofas a histidina
MC100-3	<i>his4 arg4 aox1Δ :: SARG4 aox2Δ :: Phis4</i>	<i>Mut</i> ⁻ , <i>su</i> ⁻	Incapaz de consumir el metanol, auxótrofas a histidina

SMD1168	<i>pep4, his4</i>	Mut ⁺ , His ⁻ deficiente en proteasas	Consumo de metano alto, auxótrofa histidina, no produce proteinasa A
SMD1165	<i>PRB1, HIS4</i>	Mut ⁺ , His ⁻ deficiente en proteasas	Consumo de metano alto, auxótrofa histidina, no produce proteinasa B
SMD1163	<i>pep4, prb1, his4</i>	Mut ⁺ , His ⁻ deficiente en proteasas	Consumo de metano alto, auxótrofa histidina, no produce proteinasa A, proteinasa B

2.2.4 Vectores de expresión

Los vectores diseñados para la expresión de proteínas heterólogas en *K. phaffii* tienen varias características comunes (Tabla 2). El casete de expresión de genes heterólogos está compuesto por secuencias de ADN que contienen el promotor *AOX1*, seguido de uno o más sitios de restricción únicos para la inserción del gen heterólogo, seguido de la secuencia de terminación transcripcional del gen *AOX1* de *K. phaffii* que dirige el procesamiento 3' y la poliadenilación eficiente de los ARNm. Muchos de estos vectores también incluyen el gen *HIS4* de *K. phaffii* como un marcador seleccionable para la transformación en huéspedes mutantes *his4*, así como por las secuencias necesarias para la replicación y el mantenimiento de plásmidos en bacterias (es decir, origen de replicación de ColE1 y gen de resistencia a ampicilina). Algunos vectores también contienen secuencias flanqueantes 3' de *AOX1* que se derivan de una región del genoma de *K. phaffii* y se encuentra inmediatamente después de la región 3' del gen *AOX1* y se puede utilizar para dirigir fragmentos que contienen un casete de expresión génica para la integración por reemplazo de genes en el locus *AOX1*.

Características adicionales que están presentes en ciertos vectores de expresión de *K. phaffii* sirven como herramientas para funciones especializadas. Para la secreción de proteínas heterólogas, se han construido vectores que contienen una secuencia de ADN que sigue inmediatamente al promotor *AOX1* que codifica una señal de secreción. La más frecuente utilizado de estos es la secuencia de prepro del factor alfa de *S. cerevisiae*. Sin embargo, los vectores que contienen la secuencia señal derivada del gen de la fosfatasa ácida de *K. phaffii* (PHO1) también está disponible.

Se han desarrollado vectores con marcadores dominantes de resistencia a fármacos que permiten el enriquecimiento de cepas que reciben múltiples copias de casetes de expresión durante la transformación. Un conjunto de vectores (pPIC3K y pPIC9K) contiene el gen bacteriano de resistencia a la kanamicina y confiere resistencia a altas niveles de G418 en cepas que contienen múltiples copias de estos vectores. Otro conjunto de vectores (la serie pPICZ) contiene el gen *Sh ble* de *Streptoalloteichus hindustanus*. Este gen es pequeño (375 pb) y confiere resistencia al fármaco zeocina en *Escherichia coli* (Higgins y Cregg 1998).

Tabla 2. Vectores de expresión utilizados en la expresión de proteínas heterólogas en *K. phaffii*

Vectores de expresión	Marcador de selección	Características
pPIC3K	<i>HIS4</i> y <i>kanr</i>	Sitio múltiple de clonación; selección de cepas multicopias con G418.
pPICZ	<i>Bler</i>	Sitio múltiple de clonación; selección de cepas multicopias con zeocina; fusión de la proteína heteróloga con etiqueta de seis histidinas.
pGAPZ	<i>Bler</i>	P _{GAP} , sitio múltiple de clonación, selección con zeocina, fusión de la proteína heteróloga con etiqueta de seis histidinas.
pPIC9	<i>HIS4</i>	P _{AOX1} acoplado a factor de secreción α -MF, sitios XhoI, EcoRI, NotI, SnaBI y AvrII para inserción.
pPICZ α	<i>Bler</i>	P _{AOX1} acoplado a factor de secreción α -MF, sitio múltiple de restricción, generación de multicopias, selección con zeocina, fusión de la proteína heteróloga a una cola de seis histidinas.
pGAPZ α	<i>Bler</i>	P _{GAP} acoplado a factor de secreción α -MF, sitio múltiple de restricción, selección con zeocina, fusión de la proteína heteróloga a una cola de seis aa histidinas.

Otra característica presente en ciertos vectores (por ejemplo, pAO815 y pPICZ serie de vectores) es que están diseñados para facilitar la construcción de vectores de expresión con múltiples copias del casete de expresión. Múltiples copias de un casete de expresión se introducen en estos vectores mediante la inserción de un casete de expresión unido por un sitio BamHI y BglIII, en el sitio BamHI de un vector que ya contiene una sola copia de casete de expresión. La unión entre los dos casetes ya no puede ser escindida por ninguna de los dos enzimas que permiten la inserción de otro casete unido a BamHI-BglIII en el

mismo vector para generar un vector con tres copias de casete. El proceso de adición se repite hasta que haya entre 6 y 8 copias de un casete en un solo vector que luego se transforma en la cepa huésped de *K. phaffii*. Finalmente, los vectores que contienen un promotor constitutivo de *K. phaffii* derivado del gen de la gliceraldehído-3-fosfato deshidrogenasa (*GAPDH*) de *K. phaffii* se encuentran también disponibles. El P_{GAP} es una alternativa conveniente al promotor *AOX1* para la expresión de genes cuyos productos no son tóxicos para *K. phaffii*. Además, su uso no implica el uso de metanol, que puede ser problemático en algunos casos (Higgins y Cregg 1998).

2.3 Transformación

Los métodos empleados para la transformación mediada por ADN de *K. phaffii* son similares a los de *S. cerevisiae* y conducen a resultados similares. *K. phaffii* puede ser transformado por fusión de esferoplastos, o por métodos de células enteras como electroporación o coprecipitación de ADN con cloruro de litio, cloruro de calcio o polietilenglicol. La electroporación es un método simple y rápido para transformar *K. phaffii*, es el método recomendado y comúnmente usado (Cregg et al. 2018).

2.4 Integración de vectores en el genoma de *K. phaffii*

Como en *S. cerevisiae*, los ADN de vectores lineales pueden generar transformantes estables de *K. phaffii* vía recombinación homóloga entre secuencias compartidas por el vector y genoma huésped (Cregg et al. 1987). Tales transformantes muestran fuerte estabilidad en ausencia de presión selectiva, incluso cuando se presenta en múltiples copias. Todos los vectores de expresión de *K. phaffii* portan al menos un segmento de ADN (el fragmento del promotor *AOX1* o *GAP*) con sitios de restricción únicos que pueden usarse para dirigir el vector para que se integre en el genoma del huésped por un solo evento de inserción de tipo cruzado. Los vectores que contienen el gen *HIS4* de *K. phaffii* también se pueden dirigir para que se integren en locus *his4* genómico de *K. phaffii* (Invitrogen 2010).

2.5 Estrategias para una selección eficiente de transformantes

Una variedad de marcadores seleccionables ha sido caracterizada e incorporada a los vectores de expresión de *K. phaffii*. Estos marcadores se pueden dividir en dos subgrupos: (a) marcadores biosintéticos, incluidos *HIS4*, *ARG4*, *ADE1*, *URA3*; y (b) marcadores de resistencia a fármacos, incluido el gen *kanR* y el gen *Sh ble* que confieren resistencia a G418 y zeocina, respectivamente. Combinaciones de estos diferentes marcadores se emplean para la selección eficiente de transformantes verdaderos.

La primera estrategia adoptada para generar cepas recombinantes fue la transposición de una sola copia del gen foráneo en el sitio *AOXI*, ya que este tipo de transformante es el más estable, y en algunos de los primeros estudios produjo niveles razonables de producto. Se han acumulado numerosos ejemplos donde se han utilizado transformantes multicopia para aumentar los rendimientos $Y_{p/x}$. Un aspecto importante de la selección eficiente de transformantes de *K. phaffii* es la selección de transformantes multicopia y, posteriormente, la determinación del número exacto de copias.

Para determinar el número absoluto de copias del vector, el ADN genómico total de cepas transformadas se aísla primero, y luego se utiliza el análisis de transferencia de Southern para determinar la estructura cromosómica del ADN del vector integrado (es decir, el sitio de integración, el número copia de genes *AOXI* o *HIS4*, y si se ha producido el reemplazo del gen *AOXI*). También se han desarrollado métodos por qPCR para la determinar el número de copias del casete de expresión integrados en el genoma de la levadura (Abad et al. 2010). En los casos en que se disponga de un buen ensayo para la detección de proteína expresada, la selección puede llevarse a cabo directamente mediante los niveles de producción sin ningún conocimiento previo del número de copias (Laroche et al. 1994).

2.6 Integración de vectores en el genoma de *K. phaffii*

Los elementos reguladores trabajan en conjunto para garantizar que el gen se exprese en las condiciones adecuadas dependiendo del estado metabólico y ambiental de las células. Los elementos reguladores *-cis* son secuencias específicas de ADN que se encuentran cerca del gen de interés y controlan directamente su transcripción, es decir, son elementos

que actúan como “interruptores locales” para controlar directamente la actividad del gen. Dentro de esta clasificación se encuentran:

- El promotor que contiene sitios específicos donde se une la maquinaria de transcripción, como la RNA polimerasa II, para iniciar la transcripción del gen.
- Los *enhancers* son secuencias activadoras que incrementan la eficiencia de la transcripción al reclutar factores de transcripción específicos.
- Los *silencers* son secuencias que pueden inhibir la transcripción al interactuar con proteínas represoras.
- La caja TATA es una región clave dentro del promotor que guía la formación del complejo de iniciación de la transcripción.
- Las UAS (secuencias de activación *Upstream*) en levaduras, son regiones río arriba del gen que actúan como activadores transcripcionales al atraer factores de transcripción.

Los elementos reguladores *-trans* son factores moleculares (proteínas o ARN) que interactúan con los elementos *-cis* para regular la expresión del gen, es decir, son los “reguladores móviles” que interactúan con los elementos *-cis* para activar o reprimir la transcripción, dependiendo de las señales celulares y ambientales. Dentro de esta clasificación se encuentran:

- Los factores de transcripción activadores. En el gen *GAPDH*, factores específicos como los activadores de genes implicados en el metabolismo de carbohidratos (p. ej., HAP, GCR1/GCR2 en levadura) pueden unirse a las UAS para potenciar la transcripción. Estos activadores están regulados por la disponibilidad de glucosa y otras fuentes de carbono, lo que garantiza que el gen *GAPDH* se exprese cuando es necesario para la glucólisis.
- Los factores de transcripción represores en condiciones de escasez de carbono o estrés pueden unirse a elementos *-cis* específicos para reducir la transcripción.
- Los cofactores y remodeladores de cromatina son proteínas que alteran la estructura de la cromatina cerca del gen *GAPDH*, facilitando o dificultando el acceso de la RNA polimerasa II y otros componentes.

Debido a que el gen *GAPDH* desempeña un papel clave en la glucólisis, su expresión está fuertemente regulada por las condiciones metabólicas. Por ejemplo:

- En presencia de glucosa, los factores activadores como GCR1 y GCR2 incrementan su transcripción.
- En ausencia de glucosa, el gen puede ser reprimido o su transcripción modulada según las necesidades celulares.

Los elementos reguladores aseguran que la expresión del gen sea eficiente y adaptativa, ajustándose a las necesidades energéticas de la célula (Struhl 1989).

2.7 Los promotores transcripcionales

Los promotores son un elemento crítico en el diseño de vectores de expresión y por tanto son probablemente la parte genética más caracterizada y diseñada en muchos sistemas de levaduras. Los promotores constitutivos o inducibles bien caracterizados con una fuerte actividad transcripcional se usan para lograr la sobreproducción de la proteína recombinante. Los promotores más utilizados en la producción de proteínas recombinantes en levaduras se muestran en la Tabla 3.

Tabla 3. Promotores más frecuentes utilizados para la producción de proteínas recombinantes en levaduras

Hospedero	Promotores constitutivos	Promotores inducibles
<i>S. cerevisiae</i>	ADH1, GAPDH, PGK1, TPI, ENO, PYK1, TEF	GAL1-10, CUP1, ADH2
<i>K. lactis</i>	PGK	LAC4, ADH4
<i>Y. lipolytica</i>	TEF, RPS7, XPR2/hp4d	POX2, POT1, ICL1
<i>K. phaffii</i>	GAP, TEF, PGK, YPT1	AOX1, FLD1, PEX8

En *K. phaffii* usualmente se utilizan los vectores de expresión con el promotor del gen *AOX1* (P_{AOX1}) inducible con metanol. Hasta la fecha se han producido más de 300 proteínas recombinantes usando este promotor. Pese a que se tienen altos rendimientos empleando el P_{AOX1} , el uso de metanol resulta ser peligroso y su manejo en grandes volúmenes es costoso. Además, el metabolismo del metanol genera una gran cantidad de

calor y una alta demanda de oxígeno puro durante el cultivo que puede llegar a limitar el crecimiento celular, incrementa el costo productivo y limita la capacidad de fermentación a gran escala.

El metanol es un sustrato barato y fácilmente disponible en el momento en que se desarrolló el sistema *K. phaffii*. Sin embargo, la amplia gama de aplicaciones en el campo de la producción de proteínas que han sido desarrolladas a lo largo de los años ha puesto de manifiesto la necesidad de un sustrato controlable, menos volátil y menos inflamable para la inducción. Una estrategia para eludir la necesidad de metanol es la ingeniería del P_{AOX1} para aislar variantes que permitan niveles transcripcionales sustanciales en glucosa o desrepresión de glicerol, como el promotor MOX de *Hansenula* polimorfa (Sudbery et al. 1988). Además, también existe la necesidad de promotores con fuerza variable. En este contexto, Hartner et al. en 2008 (Hartner et al. 2008) construyeron una mutante del P_{AOX1} de diferente fuerza transcripcional. Además de *AOX1*, *K. phaffii* tiene una segunda alcohol oxidasa que se expresa a bajos niveles, pero no ha dado buenos resultados en la producción de altos niveles de la proteína recombinante, por lo que se ha incrementado la búsqueda de promotores alternativos al P_{AOX1} (Çalik et al. 2015), siendo el promotor del gen *GAPDH* (P_{GAP}) uno de los promotores alternativos más empleados (Vogl y Glieder 2013).

El P_{GAP} promueve la transcripción de la enzima gliceraldehído-3-fosfato deshidrogenasa (*GAPDH*) que es una enzima clave en las rutas de la glicólisis y la gluconeogénesis. Con P_{GAP} se reduce el tiempo de cultivo, lo que simplifica el proceso para la producción de proteínas. En función de la fuente de carbono empleada, se pueden tener niveles de producción de proteínas similares e incluso mayores respecto al sistema controlado por el P_{AOX1} (Vogl y Glieder 2013). Por ejemplo, se ha producido hasta 400 mg/L de quitinasa recombinante activa en fermentador de 1.5 L (Goodrick et al. 2001). Además, se ha reportado que la pro-acualisina con 38 kDa se secretaba a niveles de 1 g/L en experimentos con matraces (Olezdzka et al. 2003). En comparación con P_{AOX1} , P_{GAP} permite niveles de expresión más altos en algunos casos. Por ejemplo, la expresión del péptido de mamífero transportadores hPEPT1 y rPEPT2 usando el P_{GAP} rindió cinco veces más niveles de expresión que usando el clásico P_{AOX1} (Döring et al. 1998).

2.8 Los terminadores transcripcionales

2.8.1 Generación de los extremos 3' del ARNm por cortes enzimáticos y poliadenilación

Los extremos 3' de los ARNm se generan mediante una escisión que es seguida de la adición de poli (A). Experimentos previos han demostrado que la adición de 3'-desoxiadenosina, también conocida como cordicepina, no detiene la transcripción del ARN nuclear, pero sí la aparición de ARNm en el citoplasma. Esto muestra que la poliadenilación es necesaria para la maduración del ARNm.

La generación del extremo 3' se ilustra en la Figura 1. La secuencia AAUAAA es necesaria para el corte enzimático y la poliadenilación. La ARN polimerasa transcribe pasado el sitio correspondiente al extremo 3', y las secuencias en el ARN se reconocen como dianas para un corte endonucleolítico seguido de poliadenilación. Un solo complejo de procesamiento realiza tanto el corte como la poliadenilación. La secuencia poli (A) estabiliza el ARNm contra la degradación del extremo 3'. Su extremo 5' ya está estabilizado por CAP. La polimerasa de ARN continúa la transcripción después de la escisión, pero el extremo 5' generado por la escisión es desprotegido y degradado (Lewin 2012).

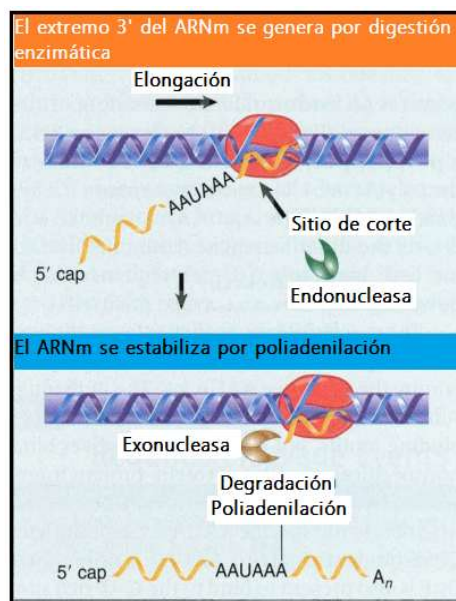


Figura 1. La secuencia AAUAAA es necesaria para el corte enzimático y la poliadenilación.

El evento de escisión proporciona un desencadenante indirecto para la terminación de la transcripción por la ARN polimerasa II. Una exonucleasa se une al extremo 5' del ARN que continúa siendo transcrito después de la escisión, y degrada el ARN más rápido de lo que se sintetiza por la ARN polimerasa. Luego interactúa con proteínas auxiliares que se unen al dominio carboxi-terminal de la polimerasa, y esta interacción desencadena la liberación de la ARN polimerasa del ADN, causando la terminación de la transcripción. El modelo general es similar al de la proteína rho en la terminación de la transcripción por la ARN polimerasa bacteriana. Esto explica por qué los sitios de terminación para la ARN polimerasa II no están bien definidos, pero puede ocurrir en diferentes lugares dentro de la región ubicada río abajo del sitio de corte y poliadenilación. Una característica común de los ARNm en eucariotes superiores (pero no en levadura) es la presencia de la secuencia altamente conservada AAUAAA en la región de 11 a 30 nucleótidos río arriba del sitio de la adición de poli (A). La delección o mutación del hexámero AAUAAA evita la generación del extremo 3' poliadenilado. La señal es necesaria tanto para la escisión como para la poliadenilación.

Estudios *in vitro* demostraron la secuencia de reacciones (Figura 2) que dan lugar a la generación de la adecuada estructura del extremo 3'. Se requiere de una endonucleasa que se compone de CFI y CFII para escindir el ARN, de una poli (A) polimerasa (PAP) para sintetizar la cola de poli (A) y un componente de especificidad (CPSF) que reconoce la secuencia AAUAAA y dirige las demás actividades. Un factor estimulante, CstP, se une a una secuencia rica en G-U que se encuentra río abajo del propio sitio de escisión. El factor de especificidad contiene cuatro subunidades, que juntas se unen específicamente al ARN que contiene la secuencia AAUAAA. Las subunidades son proteínas que tienen motivos de unión a ARN comunes, pero que por sí mismas se unen inespecíficamente al ARN. Son necesarias las interacciones proteína-proteína entre las subunidades para generar el sitio de unión a AAUAAA específico. CPSF se une fuertemente a AAUAAA solo cuando CstF también está presente para unirse al sitio rico en G-U. El factor de especificidad es necesario para las reacciones de escisión y de poliadenilación.

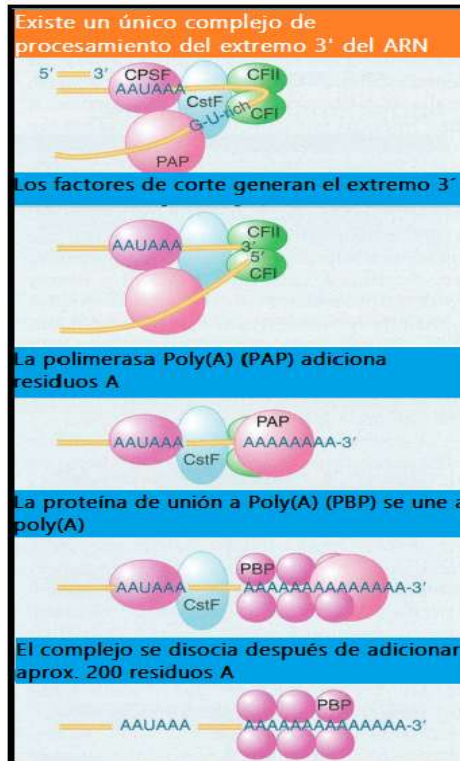


Figura 2. Complejo de procesamiento del extremo 3' del ARN. CPSF (componente de especificidad) y CstF (factor estimulante) constan cada uno de varias subunidades; los otros componentes son monoméricos. La masa total es > 900 kDa.

Se forma un complejo entre la endonucleasa y la poli (A) polimerasa, y este complejo generalmente realiza la escisión seguida de la poliadenilación de una manera estrechamente acoplada. Los dos componentes CFI y CFII (factores de escisión I y II), junto con el factor de especificidad son necesarios y suficientes para la escisión endonucleolítica. La poli (A) polimerasa tiene una actividad catalítica con efecto inespecífico; pero cuando se combina con los otros componentes, la reacción sintética se vuelve específica para el ARN que contiene la secuencia AAUAAA. La reacción de poliadenilación pasa por dos etapas. Primero, se agrega secuencia corta de oligo (A) (-10 residuos) hasta el final de 3'. Esta reacción es absolutamente dependiente de la secuencia AAUAAA, y la poli (A) polimerasa lo realiza bajo la guía del factor de especificidad. En la segunda fase, la cola del oligo (A) se extiende hasta una longitud total de 200 residuos. Esta reacción requiere del factor estimulante que reconoce la cola con el oligo (A) y dirige la poli (A) polimerasa específicamente para extender el extremo 3' de la secuencia poli (A).

La poli (A) polimerasa por sí misma agrega residuos A individualmente a la posición 3'. Su modo de acción es distributivo; se disocia después de que se haya agregado cada nucleótido. Sin embargo, en presencia de CPSF y PABP (unión poli (A) proteína), funciona procesivamente para extender una cadena de poli (A) individual. El PABP es una proteína de 33 kDa que se une estequiométricamente al tramo de poli (A). La longitud de poli (A) es controlado por el PABp, que de alguna manera limita la acción de la poli (A) polimerasa a -200 adiciones de residuos A. El límite puede representar la acumulación de una masa crítica de PABP en la cadena de poli (A). PABP se une al factor de iniciación de la traducción eIF4G, generando así un circuito cerrado en el que un complejo de proteínas contiene los extremos 5' y 3' del ARNm.

2.8.2 Funcionamiento de los terminadores transcripcionales poliadenilados

La presencia de poli (A) en la cola de 3' de un ARNm estimula la unión de un complejo de iniciación de la traducción eIF4F en el extremo de 5' (heterotrímero conformado por eIF4G (proteína andamio o scaffold protein), eIF4E (se une al extremo 5' cap-metilado y eIF4A (helicasa que desune la doble cadena del extremo 5')). Para este efecto se requiere de la proteína de unión a poli (A) (Pab Ip en levadura). Pab Ip se une a la proteína eIF4G. Esto implica que el ARNm tendrá una organización circular tanto tiempo como eIFG esté unida a los extremos 5' y 3' del mRNA.

La unión de la PABP al factor de iniciación eIF4G genera un bucle cerrado, en el que los extremos 5' y 3' del ARNm se acercan por estar retenidos en el mismo complejo proteico (Figura 3).

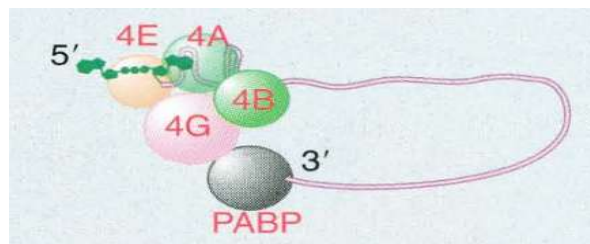


Figura 3. Factores de iniciación de la traducción. eIF4F (heterotrímero conformado por eIF4E, eIF4A y eIF4G).

La importancia de la formación de este circuito cerrado no está clara, aunque podría tener varios efectos, como:

- Estimular el inicio de la traducción. Cuando el proceso de traducción termina en el extremo 3', las subunidades ribosomales liberadas ya están en la proximidad del extremo 5'.
- Estabilizar el ARNm contra la degradación.
- Permitir la unión de factores al extremo 3' para regular el inicio de la traducción.

La eliminación de poli (A) inhibe el inicio de la traducción *in vitro* y el agotamiento de PABP tiene el mismo efecto *in vivo* en levaduras.

Estos efectos podrían depender de la unión de PABP al complejo de iniciación en el extremo 5' del ARNm. Hay muchos ejemplos en el desarrollo embrionario temprano en los que la poli (A) de un ARNm particular se correlaciona con su traducción. En algunos casos, los ARNm se almacenan en forma no poliadenilada y se agrega poli (A) cuando se requiere su traducción; en otros casos, el ARNm poli (A)⁺ se de-adenilan y su traducción se reduce.

2.8.3 Señales de poliadenilación en la región 3'UTR

El estudio de unos pocos genes ha permitido identificar tres elementos que constituyen una señal de poliadenilación de levadura: el elemento de eficiencia, el elemento de posicionamiento y el sitio real para la escisión y la poliadenilación.

La formación del extremo 3' del ARNm en células eucariotas implica la escisión endonucleolítica en un sitio específico del ARNm precursor acoplado en la mayoría de los casos a la polimerización de una cola de poli (A) sobre el fragmento de escisión río arriba (Zhao et al. 1999). Este proceso es un paso esencial en la síntesis de ARNm eucariótico porque la cola poli (A) funciona en la renovación del ARNm (Beelman y Parker 1995) y en la traducción (Sachs et al. 1997).

El minucioso análisis realizado por Guo y Sherman (Guo y Sherman 1996) en el gen *CYCI* de *S. cerevisiae* ha permitido la identificación de tres elementos trabajando en conjunto, que no sólo son necesarios sino también suficientes para constituir una señal de poliadenilación de levadura. El elemento de eficiencia (EE), cuya supresión disminuye la eficiencia del procesamiento. Es un elemento rico en TA. El elemento TAGTATGTA fue el primer hexanucleótido propuesto como una señal de poliadenilación, revelando que

TAYRTA es esencial para esta función y que la secuencia TATATA tiene la mejor capacidad para la formación de extremos 3'. Además, los dos residuos T en la primera y quinta posición son los nucleótidos más esenciales en esta secuencia (Irniger y Braus 1994).

El segundo elemento de la señal de poliadenilación es un elemento de posicionamiento rico en A, cuyo reconocimiento parece estar estrechamente relacionado con la escisión en el sitio PyAn rio abajo más cercano, generalmente ubicado entre 13 y 27 nucleótidos rio abajo. Las secuencias identificadas hasta ahora para esta función han sido AAAAAAAAA, TTAAGAAC, AAAA, AATAATGA y AATAAA, siendo esta última la señal más fuerte y posible consenso (Van Helden et al. 2000).

2.9 Modificaciones Postraduccionales

K. phaffii tiene el potencial de realizar muchas de las modificaciones postraduccionales típicamente asociadas con los eucariotas superiores. Estos incluyen procesamiento de secuencias señal (tanto pre- como tipo prepro), plegamiento, formación de puentes disulfuro y O- y N-glicosilaciones.

La glicosilación de proteínas eucariotas heterólogas secretadas por *K. phaffii* y otros hongos pueden ser problemáticos. En los mamíferos, los O-oligosacáridos se componen de una variedad de azúcares, incluyendo N-acetilgalactosamina, galactosa y ácido siálico. Por el contrario, los eucariotas inferiores, incluido *K. phaffii*, agregan O-oligosacáridos compuestos únicamente de residuos de manosa (Cregg et al. 2000).

Existen enzimas disponibles comercialmente que reconocen la estructura del núcleo con alto contenido de manosa y catalizan la hidrólisis del enlace $\text{GlcNAc}\beta\text{-1,4GlcNAc}$ en el núcleo, resultando en la liberación del oligosacárido, generalmente resultando la disminución suficiente en la masa aparente que se puede detectar utilizando SDS-PAGE. La Endo- β -N-acetilglucosaminidasa H (Endo H) es una de esas enzimas que se ha utilizado ampliamente para este objetivo. El otro tipo de enzima utilizada para la liberación del oligosacárido ligado a N no es específico para el alto contenido de manosa o estructuras de tipo complejo, y cataliza la hidrólisis de los enlaces $\text{GlcNAc}\beta\text{-Asn}$ (Bretthauer y Castellino 1999).

2.10 Cultivo celular. Fuente de carbono, velocidad específica de crecimiento y cinética de producción

K. phaffii es capaz de crecer en muchas fuentes de carbono y energía (Inan y Meagher 2001), de las cuales el glicerol, la glucosa y el metanol son los más utilizados. En comparación con los promotores inducibles, existe mayor disponibilidad de fuentes de carbono para la expresión génica heteróloga en *K. phaffii* con sistema de expresión constitutivo. La glucosa, el glicerol, el ácido oleico y el metanol son las fuentes de carbono para el sistema de expresión P_{GAP} . Aunque sigue siendo una disputa sobre cuál es la mejor para la expresión, la glucosa y el glicerol son los preferibles para este sistema de expresión (Zhang et al. 2007; Garcia-Ortega et al. 2013, Looser et al. 2015; Zepeda et al. 2018; Liu et al. 2016; Dou et al. 2021).

Durante el cultivo de microorganismos ocurren varias fases de crecimiento caracterizadas por las variaciones en la velocidad específica de crecimiento. Como definición clásica se asume la siguiente consecución de fases (Figura 4): La fase de retraso donde la velocidad específica de crecimiento es nula, en la fase de aceleración aumenta la velocidad específica de crecimiento, en la fase exponencial la velocidad específica de crecimiento es constante y máxima, en la fase de retardo disminuye la velocidad específica de crecimiento, luego le sucede la fase estacionaria con velocidad específica de crecimiento nula y la fase de declive con velocidad específica de crecimiento negativa (Monod 1949).

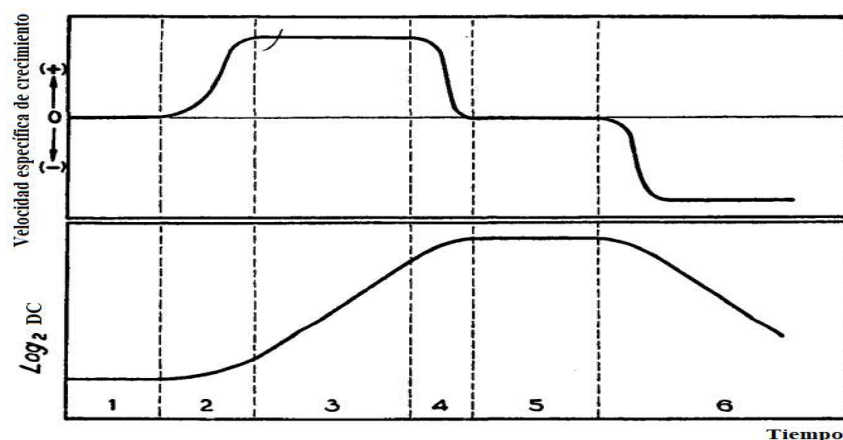


Figura 4. Fases de crecimiento microbiano.

Curva inferior: logaritmo de la densidad celular. Curva superior: variaciones en la velocidad específica de crecimiento. Las líneas punteadas verticales marcan los límites de las fases (Monod 1949).

Para la determinación de la velocidad específica de crecimiento máxima (μ) se calcula la pendiente de la curva de crecimiento expresada como logaritmo natural de la densidad celular vs tiempo en su fase exponencial (Figura 5).

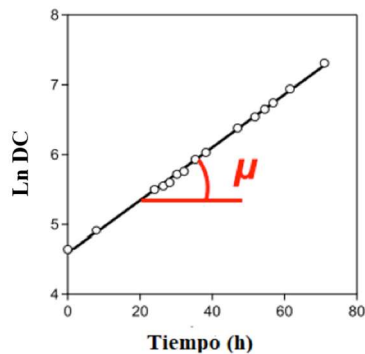


Figura 5. Representación de la velocidad específica de crecimiento (μ) como la pendiente del \ln (DC) vs el tiempo (Looser et al. 2015).

La determinación cuidadosa de μ es esencial para el diseño de las estrategias de cultivo en las que el crecimiento se mantiene en un cierto valor óptimo mediante la adición controlada del sustrato de la fuente carbono en el modo de lote alimentado (Looser et al. 2015).

Referente a este tema, varios han sido los estudios realizados con el sistema de expresión de *K. phaffii* para evaluar la influencia de las fuentes de carbono más comúnmente usadas (glucosa y glicerol) y la velocidad específica de crecimiento (μ) en la producción de proteína recombinante bajo el promotor P_{GAP} . Los resultados reportados no son definitivos porque aún son motivo de debate. Algunos autores han revelado que la fortaleza de P_{GAP} varía en función de la fuente de carbono empleada durante el crecimiento celular. En presencia de glucosa la expresión de β -lactamasa resultó significativamente mayor comparada con el usual P_{AOX1} en cultivos con metanol (Waterham et al. 1997). En la producción de una proteína factor estimulante de colonias de macrófagos (hGM-CSF) observaron que con la variación de parámetros operacionales (pH, temperatura) o incluso a diferentes velocidades específicas de crecimiento (μ) el rendimiento $Y_{p/x}$ no tuvo variaciones significativas con valores constantes alrededor de 2.5 mg/g y solo la fuente de carbono fue el parámetro que influyó significativamente en $Y_{p/x}$, siendo mejor en presencia de glucosa que en glicerol (Yogender et al. 2006). Sin embargo, en la evaluación de diferentes fuentes de carbono para la producción de angiostatina resultó mejor el

glicerol que la glucosa o el metanol (Zhang et al. 2007). Otros autores recomiendan el empleo de ambas fuentes de carbono en las distintas etapas del proceso de cultivo empleando el glicerol para la primera etapa de cultivo por lote y el empleo de glucosa para la etapa de cultivo alimentado, aunque esta estrategia solo incrementó en 1.1 el $Y_{p/x}$ respecto a la combinación glicerol-glicerol, la decisión se basó en el empleo de la glucosa, debido a que proporciona un menor generación de calor y demanda de oxígeno por unidad de biomasa en comparación con el glicerol (García-Ortega et al. 2013).

La posible diferencia en el incremento de los $Y_{p/x}$ en presencia de glucosa respecto a glicerol pudiera deberse entre otros aspectos a que el glicerol no es una fuente de carbono fermentable y *K. phaffii* no se considera una levadura fermentativa, sin embargo hay acumulación de etanol durante el cultivo cuando se usa glicerol a velocidades altas de alimentación (Zepeda et al. 2018). Se ha observado que la actividad de P_{GAP} en células creciendo con glicerol y metanol es aproximadamente de dos a un tercio los niveles observados en glucosa, respectivamente (Waterham et al. 1997, Cereghino y Cregg 2000, Zepeda et al. 2018).

Se ha reportado además altos niveles de transcritos del gen heterólogo en glucosa y los mayores niveles de $Y_{p/x}$ extracelular en glicerol indicando que los niveles de transcritos por ellos mismos no son suficientes para predecir los niveles de producción de proteína (Liu et al. 2016; Dou et al. 2021). Este fenómeno pudiera explicarse por el hecho de que las condiciones de cultivo que producen los menores niveles de transcritos también producen los mayores $Y_{p/x}$, por ejemplo en la producción de FTEII bajo la regulación de P_{GAP} (Herrera-Estala et al. 2022). Esto se debe a una menor represión catabólica que permite la sobre regulación de los genes *KAR2* y *PSAI-1*, que son genes relacionados con la vía de secreción más que con el incremento de la transcripción de los genes heterólogos.

2.11 Antecedentes directos

En nuestro laboratorio se han realizado estudios de expresión de la proteína heteróloga FTEII en *K. phaffii* con la combinación de elementos reguladores P_{AOXI} - T_{AOXI} (Guerrero-Olazarán et al. 2010) y más recientemente, por las ventajas que ofrece el empleo del P_{GAP} , se realizaron análisis de expresión de *FTEII* con la combinación

$P_{GAP-T_{AOX1}}$ (Herrera-Estala et al. 2022). El estudio se realizó con una cepa multicopia del gen heterólogo para la producción de FTEII en condiciones de cultivo optimizadas, empleando glicerol como fuente de carbono. Los resultados indicaron una relación inversamente proporcional entre la Q_p y la velocidad específica de crecimiento. Los cultivos realizados con μ bajas (0.054 h^{-1}) mostraron los mayores valores de Q_p y $Y_{p/x}$ extracelular. Estos resultados difieren de los previamente reportados por otros autores (Looser et al. 2017, García-Ortega et al. 2013) para cepas productoras de proteínas recombinantes empleando el P_{GAP} . A pesar de los esfuerzos en la optimización de los parámetros del proceso de producción, el rendimiento $Y_{p/x}$ resultó 5.3 veces menor que los reportados por Guerrero-Olazarán et al. en 2010.

Estos resultados son indicio de que la expresión del gen heterólogo *FTEII* en cepas de *K. phaffii* con P_{GAP} pudiera estar influenciada por condiciones ambientales, por factores moleculares o por la combinación de ambos, indicando la necesidad de estudiar el $Y_{p/x}$ extracelular, y los niveles de expresión de *FTEII* en *K. phaffii* que contengan el gen heterólogo con la combinación de los elementos reguladores $P_{GAP-T_{AOX1}}$ o $P_{GAP-T_{GAP}}$ en cepas con igual número de copias y transformadas en el mismo sitio de inserción del genoma de la levadura.

3 JUSTIFICACIÓN

El desarrollo de la ingeniería genética ha favorecido el avance de la industria biotecnológica mediante la obtención de proteínas recombinantes con aplicación en los sectores farmacéutica, agropecuaria, acuícola. Entre los sistemas de expresión más utilizados destaca la levadura *K. phaffii*, debido al promotor fuerte y regulable por metanol de la enzima alcohol oxidasa (P_{AOX1}), y a su capacidad para alcanzar altas densidades celulares en biorreactores con medios económicos.

Sin embargo, el uso del metanol presenta importantes limitaciones: es altamente inflamable, lo que dificulta su almacenamiento en grandes cantidades; es un derivado del petróleo por lo que se prohíbe su empleo en la obtención de alimentos o suplementos nutricionales. Para evitar o al menos reducir su empleo se han desarrollado promotores alternativos. Uno de los más utilizados es el de la enzima gliceraldehido-3-fosfato deshidrogenasa (P_{GAP}); que es un promotor constitutivo y las estrategias de producción resultan más sencillas que con el P_{AOX1} .

A pesar del gran número de trabajos científicos que reportan la expresión de genes heterólogos y la producción de proteínas recombinantes en el sistema de expresión de *K. phaffii* basados en P_{GAP} aún hay aspectos ambientales y moleculares que han sido poco estudiados, como el empleo de diferentes terminadores transcripcionales para potenciar la expresión génica. Además factores como la influencia de la fuente de carbono o la velocidad específica de crecimiento continúan siendo temas controversiales en cuanto a su impacto en la productividad volumétrica (Q_p), especialmente en el caso de proteínas heterólogas secretadas al medio de cultivo.

El presente trabajo contribuye al aporte de conocimiento científico relacionado con el impacto de las combinaciones de los elementos reguladores $P_{GAP}-T_{AOX1}$ y $P_{GAP}-T_{GAP}$, de la fuente de carbono (glucosa o glicerol), de la fase de crecimiento celular, y de la interacción de todos estos factores en los niveles de transcripción de un gen heterólogos reportero y del gen *GAPDH* endógeno, así como en los valores de $Y_{p/x}$ extracelular en la cepa recombinante KM71 de *K. phaffii*. Estos conocimientos resultan de utilidad en la optimización de procesos de producción de biomoléculas de interés.

4 HIPÓTESIS

Es posible alterar los niveles de expresión del gen heterólogo regulado por el P_{GAP} teniendo como terminador transcripcional la región terminadora del gen *GAPDH* en lugar de la región terminadora del gen *AOX1* de *K. phaffii*.

5 OBJETIVO DEL TRABAJO

5.1 Objetivo general

Evaluar y comparar el efecto de los terminadores transcripcionales *AOXI* y *GAP* sobre los niveles de transcritos del gen heterólogo *FTEII* y del gen endógeno *GAPDH*, así como sobre el rendimiento extracelular $Y_{p/x}$, en cepas recombinantes de *K. phaffii* que emplean el promotor *GAP* para la expresión del gen heterólogo.

5.2 Objetivos específicos

Para dar cumplimiento al Objetivo general se propusieron los siguientes objetivos específicos:

1. Construir un vector de expresión de *K. phaffii* conteniendo un casete de expresión conformado por el promotor constitutivo P_{GAP} , una región codificante de la proteína recombinante *FTEII* asociada en su extremo N-terminal con el péptido prepro del factor α de *S. cerevisiae*, y la región terminadora del gen *GAPDH*.
2. Construir y seleccionar cepas recombinantes de *K. phaffii* KM71 portadoras del casete de expresión conformado por el promotor constitutivo P_{GAP} , una región codificante de la proteína recombinante *FTEII* asociada en su extremo N-terminal con el péptido prepro del factor α de *S. cerevisiae* y la región terminadora del gen *GAPDH* de *K. phaffii*, y cepas de *K. phaffii* con un casete de expresión similar al anterior, pero con el terminador de *AOXI* en lugar del terminador del gen *GAPDH*.
3. Comparar los niveles de expresión del gen heterólogo y del gen *GAPDH* nativo; entre la cepa que contiene el promotor y el terminador *GAP*, con la cepa que contiene el gen heterólogo regulado por P_{GAP} y la región terminadora del gen *AOXI*, y correlacionar los niveles de expresión del gen *FTEII* con el rendimiento extracelular $Y_{p/x}$.

6 MATERIALES Y MÉTODOS

6.1 Lugar de trabajo

El proyecto se desarrolló en el Laboratorio de Biotecnología Molecular (L5) perteneciente al Instituto de Biotecnología (IB) de la Facultad de Ciencias Biológicas (FCB) de la Universidad Autónoma de Nuevo León (UANL).

6.2 Estrategia general

El estudio del efecto del terminador del gen *GAPDH* en la expresión de un gen heterólogo con el promotor del propio gen P_{GAP} , se inició con la construcción del vector de expresión pP_{GAP} -FTEII- T_{GAP} (vector con promotor y terminador del gen *GAPDH* de *K. phaffii*). Para ello se realizó el diseño de una secuencia nucleotídica denominada T_{GAP} , que se conformó con la región terminadora (que incluyó la región 3'UTR) del gen *GAPDH*, flanqueada en sus extremos con los sitios de restricción de las enzimas *NotI* y *Bsu36I*. El servicio de síntesis de T_{GAP} se contrató a Integrated DNA Technologies Inc. (Coralville, IA, EUA), quienes proporcionaron la secuencia sintetizada insertada en el vector pUCIDT-AMP al que se denominó pUCIDT $_{GAP}$. Este vector se clonó en la cepa TOP10 de *E. coli* para su conservación y propagación. Los vectores pUCIDT $_{GAP}$ y pGAHFTEII se digirieron con las enzimas de restricción *NotI*-*HF* y *Bsu36I*. El vector pGAHFTEII (del cepario del laboratorio L5 del Instituto de Biotecnología de la UANL) se conforma por el P_{GAP} , la secuencia que codifica para la región prepro del factor α de *S. cerevisiae*, el gen heterólogo *FTEII* (secuencias de interés), y la región terminadora del gen *AOX1* (T_{AOX1}). De la digestión enzimática se obtuvieron los fragmentos de interés necesarios, que se ligaron con la enzima ADN T4 ligasa. Así se obtuvo el vector de expresión pP_{GAP} -FTEII- T_{GAP} conformado por: el P_{GAP} , la secuencia que codifica para región prepro del factor α de *S. cerevisiae*, el gen heterólogo, y la región terminadora del gen *GAPDH*. Su correcta construcción se caracterizó por secuenciación y por la presencia de secuencias consenso que constituyen señales para la poliadenilación.

Luego se construyeron las cepas KM71/ P_{GAP} -FTEII- T_{GAP} y KM71/ P_{GAP} -FTEII- T_{AOX1} de *K. phaffii*. El proceso de transformación se realizó con células electrocompetentes de la

cepa nativa KM71 de *K. phaffii* y cada uno de los vectores de expresión (pP_{GAP}-FTEII-T_{GAP} o pGAHFTEII) previamente linealizados con la enzima de restricción *Sall*. Las clonas transformadas se seleccionaron por prototrofia a histidina por su capacidad de crecimiento en medio selectivo His⁻. Se comprobó la integración del casete de expresión al genoma de las células transformadas por PCR de integración con doble pareja de oligonucleótidos, y su funcionalidad por RT-PCR para detectar la presencia de transcritos expresados por el gen heterólogo *FTEII*. Las cepas transformadas se denominaron cepas P_{GAP}-T_{GAP} y P_{GAP}-T_{AOX1}, según corresponde.

Finalmente, se evaluó el efecto de la región terminadora del gen *GAPDH* en los niveles de expresión del gen heterólogo *FTEII* en comparación con la región terminadora del gen *AOX1*. Este análisis se realizó entre las cepas P_{GAP}-T_{GAP} y P_{GAP}-T_{AOX1} con igualdad de dosis génica del gen heterólogo *FTEII*. La selección por dosis génica se realizó por qPCR, utilizando sondas de hidrólisis específicas a una secuencia del gen heterólogo *FTEII* y al CDS del gen *GAPDH* nativo. En cada cepa se seleccionó una clona unicopia y se comparó el nivel de expresión del gen heterólogo *FTEII* con el del gen *GAPDH* nativo. Además, se analizó el nivel de expresión del gen heterólogo y su correlación con el Y_{p/x} extracelular entre las cepas P_{GAP}-T_{GAP} y P_{GAP}-T_{AOX1}, cultivadas tanto en presencia de glucosa como de glicerol como fuente de carbono. La estrategia general del proyecto se muestra en la Figura 6.

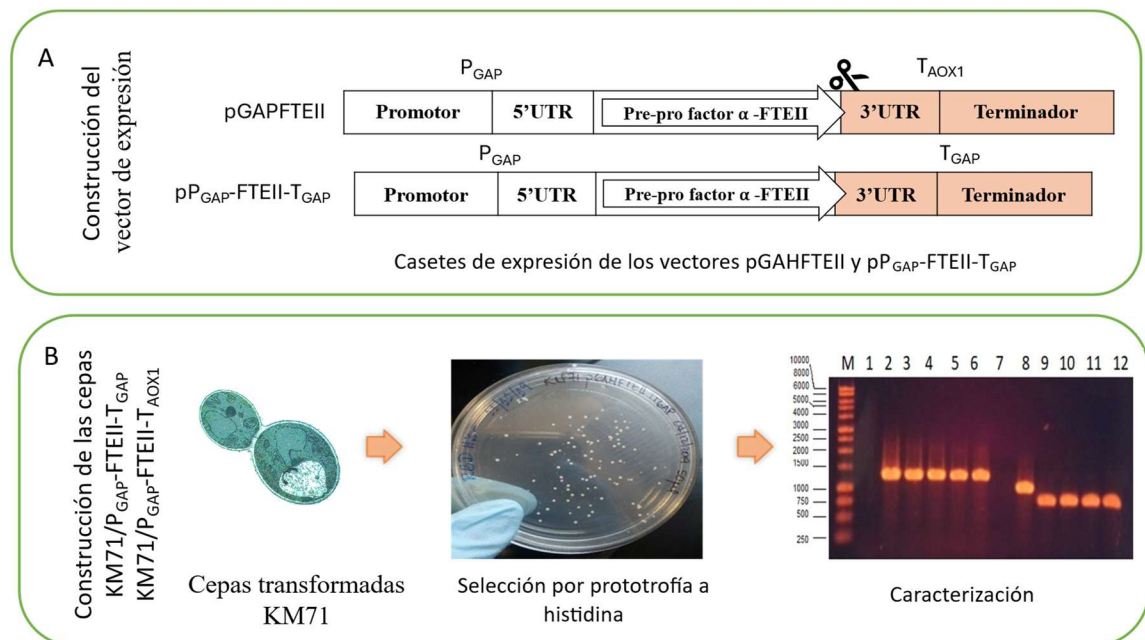




Figura 6. Estrategia general del proyecto.

A. Construcción del vector de expresión pP_{GAP}-FTEII-T_{GAP} con la subclonación de T_{GAP} en pGAHFTEII. **B.** Construcción de las cepas recombinantes de *K. phaffii* KM71/P_{GAP}-FTEII-T_{GAP} y KM71/P_{GAP}-FTEII-T_{AOX1}, selección de las clones transformados y verificación de la integración del casete de expresión en el genoma de *K. phaffii*. **C.** Evaluación comparativa de la expresión del gen heterólogo FTEII en clones unicopias KM71/P_{GAP}-FTEII-T_{GAP} y KM71/P_{GAP}-FTEII-T_{AOX1}.

6.3 Cepas, plásmidos, composición del medio, productos químicos y enzimas

La cepa de *K. phaffii* KM71 Muts (*his4*) se obtuvieron de Thermo Fisher Scientific (Waltham, MA, EUA). El vector de clonación pUCIDT-AMP fue suministrado por Integrated DNA Technologies Inc. (Coralville, IA, EUA). La cepa KM71GAHFTEII de *K. phaffii* y el plásmido pP_{GAP}-FTEII-T_{AOX1} (inicialmente llamado pGAHFTEII) se obtuvo previamente en nuestro laboratorio (Herrera-Estala et al. 2022). Este plásmido deriva del vector pPIC9 (Thermo Fisher Scientific), y contiene un casete de expresión que incluye la secuencia P_{GAP}, seguido de la secuencia que codifica para el péptido prepro del factor α de *S. cerevisiae*, y una secuencia nucleotídica que codifica para la fitasa FTEII optimizada con codones preferenciales de *K. phaffii* (Viader-Salvadó et al. 2010), además de la región terminadora del gen *AOX1* (nucleótidos 240823 a 241156 de la cepa CBS 7435 de *K. phaffii* cromosoma 4 acceso del GenBank no. FR839631.1) y una copia funcional del gen histidinol deshidrogenasa (*HIS4*) para restaurar la auxotrofia de la célula hospedera. Las endonucleasas de restricción *Bsu36I* y *NotI-HF*, la ADN polimerasa Q5 Hot Start de alta fidelidad, la Endo Hf glicosidasa se adquirieron en New England Biolabs (Beverly, MA, EUA). La enzima de restricción *Sall* se adquirió de Clontech (Palo Alto, CA, USA). Las enzimas reversotranscriptasa M-MLV, la DNasa libre de RNasa RQ1, la GoTaq ADN polimerasa, y los oligo(dT)₁₅ fueron proporcionados por Promega (Madison, WI, USA). La reversotranscriptasa SCRIPT, el oligo(dT)₂₀ y el SybrMaster mix fueron adquiridos a Jena Bioscience GmbH (Jena, Germany). Otros oligonucleótidos y sondas de qPCR PrimeTime provinieron de Integrated DNA Technologies, Inc.

(Coralville, IA, USA). La solución RNAlater fue de Ambion (Grand Island, NY, USA). Las placas de agar Luria-Bertani (LB), caldo LB, el extracto de levadura-peptonadextrosa (YPD), el medio mínimo amortiguado con glicerol (BMG) se prepararon de acuerdo con el manual del kit de expresión de *K. phaffii* (Invitrogen, San Diego, CA, EUA). Las modificaciones al medio BMG standard, etiquetados como BMGlc y BMGly contuvieron 30 mM de glucosa (0.54 % [p/v] o 30 mM de glicerol (0.28 % [p/v]), en lugar del 1 % (p/v) de glicerol y también se suplementó con 0.1 % (p/v) de CaCl₂. La cepa TOP10 de *Escherichia coli*, usada como hospedero para la subclonación, se obtuvo de Promega (Madison, WI, EUA). El marcador de tamaño molecular Benchtop 1 Kb DNA Ladder, el kit para aislamiento de DNA plasmídico Wizard® Plus SV Minipreps DNA Purification system fueron suministrados por Promega (Madison, WI, EUA). La Acrilamida/bisacrilamida se suministró por la casa comercial Bio-Rad (Bio-Rad Laboratories SA, CA, EE. UU). Todos los reactivos químicos fueron de grado analítico y comprados en Sigma Aldrich Co. (St. Louis, MO) o de Productos Químicos Monterrey (Monterrey, N.L., México).

6.4 Secuencia del terminador transcripcional *GAPDH* (T_{GAP})

Para determinar la secuencia de T_{GAP}, inicialmente se utilizó la información de tres transcriptomas de cepas KM71 de *K. phaffii* crecidas en glicerol o metanol previamente obtenidos en el laboratorio, disponibles en NCBI Sequence Read Archive (SRA) con el número de acceso de BioProject PRJNA930494. Los transcriptomas correspondieron a las cepas KM71 con vector pPIC9, con vector pPIC9 conteniendo la secuencia del gen heterólogo para la expresión de FTEII (pPIC9FTEII) y con vector pPIC9 conteniendo la secuencia del gen heterólogo de una enzima que degrada la lactosa (pPIC9Lac412). Las lecturas de los archivos FASTQ se mapearon contra la secuencia inter-CDS del gen *GAPDH* y del gen contiguo localizado río abajo de *GAPDH* (denominado *NAB6*) del genoma de la cepa CBS7435 de *K. phaffii*. Para ello se utilizó el programa HISAT2 v2.1 (Kim, Langmead and Salzberg, 2015) (Galaxy version 2.1.0) con el objetivo de detectar la región intergénica sin alineamiento de lecturas y por lo tanto sin transcritos, que se correspondería con el fin de la secuencia 3'UTR del gen *GAPDH*. Finalmente se

determinó la secuencia 3'UTR del gen *GAPDH* con la técnica 3'RACE que amplifica la región 3' de un ADNc.

Primero se aisló el ARN total de la cepa KM71GAHFTEII de *K. phaffii* usando el sistema de aislamiento de ARN total SV, y el ARN purificado se trató con ADNasa RQ1 libre de ARNasa (Promega, Madison, WI, EUA). Luego se sintetizó el ADNc mediante transcripción inversa utilizando 14,7 μ M del oligonucleótido T17AP (5'-GACTCGAGTCGACATCGAT17-3') y 2 μ l de transcriptasa inversa M-MLV (Promega, Madison, WI, EUA) en un volumen de reacción final de 30 μ L. El ADNc se amplificó mediante PCR utilizando el oligonucleótido RACEAP (5'-GACTCGAGTCGACATCG-3') y el oligonucleótido 5qGAP (5'-GCAGACTACCTACGAGGAGATCAA-3'), específico al CDS del gen *GAPDH*. La PCR se realizó en un termociclador PCR Multigene Mini (Labnet International Inc., Edison, NJ, EUA). Se utilizó un volumen de reacción de 25 μ L que contenía 5 μ M de cada cebador, 0.24 mM de dNTP, 1 U de la DNA Polimerasa Q5® High-Fidelity (New England Bio-labs, Beverly, MA, EUA) y 1 μ L de ADNc. Se programaron 30 ciclos de amplificación en las siguientes condiciones: 98 °C durante 10 s, 54 °C durante 30 s y 72 °C durante 30 s, con un primer paso de desnaturalización a 98 °C durante 30 s y un paso de extensión final a 72 °C durante 2 min.

El ADNc amplificado se purificó usando Wizard® SV Gel y PCR Clean-Up System (Promega, Madison, WI, EUA) y se secuenció usando el oligonucleótido específico al CDS del gen *GAPDH* denominado 5qGAP. Se utilizó un secuenciador ABI Prism 310 (Apply Biosystems, Foster City, CA, EUA) en la Unidad de Biología Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México. La secuencia obtenida se comparó con la secuencia reportada en la base de datos GenBank del gen *GAPDH* (*TDH3*) de la cepa CBS 7435 (acceso de Gene Bank no. FR839631.1) utilizando la herramienta Blastn del Centro Nacional de Información Biotecnológica (Bethesda, MD, EUA), y así identificar la región 3'UTR. La secuencia 3'UTR de *GAPDH* y 50 nucleótidos río abajo de esta región se consideró la secuencia del terminador transcripcional T_{GAP} conformada por 101 pb.

Se sintetizó un fragmento de ADN conteniendo el sitio de restricción de *NotI*, seguido de 11 nucleótidos espaciadores, la secuencia T_{GAP}, y 507 nucleótidos comprendidos desde el

final de la región 3' de T_{AOX1} hasta el sitio de restricción *Bsu36I* del vector pPIC9, se insertó en el plásmido pUCIDT-AMP y se secuenció por Integrated DNA Technologies, Inc. para generar el plásmido pUCIDT_{GAP}.

6.5 Modelo de señales de poliadenilación en T_{GAP}

Se utilizaron 51 hexanucleótidos reportados en la literatura (Van Helden et al. 2000) como posibles señales de poliadenilación para levadura y se alinearon con la secuencia de 101 pb seleccionada como terminador en T_{GAP}. El alineamiento se realizó con el módulo CAP del programa BioEdit, en donde se fijó el número mínimo de solapamiento de bases en 4 y un mínimo de 85 % de similitud. Se propuso un modelo de señales poliadenilación en el terminador del gen *GAPDH* de *K. phaffii* basado en la estructura propuesta por Graber et al. para levaduras (Graber et al. 1999).

6.6 Preparación de las células calcio competentes de la cepa TOP10 de *E. coli*

Las conservas en glicerol de las células de la cepa TOP10 de *E. coli* almacenadas a -20°C se reactivaron a razón de 100 µL de células en 10 mL del medio de cultivo LB. Las células se incubaron a 37 °C y 250 rpm durante toda la noche. Seguidamente, se midió la densidad celular y se volvió a inocular en 10 mL de medio de cultivo LB a una DO_{600nm} inicial entre 0.01-0.05. El cultivo se incubó a 37 °C y 250 rpm hasta alcanzar una DO_{600nm} final de 0.4-0.6. Las células se colectaron a razón de 2 mL por vial y se centrifugaron a 2000 x g por 15 min a 4 °C. Se desechó el sobrenadante y el paquete celular se puso en hielo durante 30 min. Luego se resuspendió suavemente en 200 µL (relación 1:10 respecto al volumen inicial) de 0.1 mol/L de CaCl₂ (esta solución se colocó en hielo antes de usarse). Las células resuspendidas se incubaron por 10 min en hielo. Se centrifugaron los viales a 2000 x g por 10 min a 4 °C. Se descartó el sobrenadante y las células se resuspendieron con 100 µL de CaCl₂ 0.1 mol/L (relación 1:20 respecto al volumen inicial). Las células se almacenaron a 4 °C durante toda la noche antes de ser usadas en la transformación celular.

6.7 Transformación de las células competentes de la cepa TOP10 de *E. coli*

Las células de *E. coli* TOP10 calcio-competentes se transformaron con el plásmido pUCIDT-T_{GAP} que contiene la secuencia del terminador T_{GAP}. Se aplicaron 4 µL del

plásmido control pGAHFTEII y 2 μL del plásmido pUCIDT-AMP cada uno en 100 μL de células competentes. El material se incubó a 4 °C durante 30 min. Posteriormente, se incubó a 42 °C durante 45 s e inmediatamente después se incubó a 4 °C durante 2 min. Luego se añadieron 900 μL de medio SOC (que estaba a TA) y las células transformadas se incubaron a 37 °C y 250 rpm durante 1 h. Luego, se realizó la siembra en placas a razón de 100 μL del cultivo de células transformadas por placa. Las placas se incubaron en una incubadora estática a 37 °C hasta que se visualizaran las colonias transformadas (24-36 h).

6.8 Preparación del BCT de las células TOP10 de *E. coli*

Se adicionaron 50 μL de cultivo en 4 mL de medio de cultivo LB suplementado con ampicilina 100 $\mu\text{g}/\text{mL}$. El cultivo se incubó a 37 °C y 250 rpm durante 12-14 h. Seguidamente, se midió la densidad celular y se volvió a inocular en 5 mL de medio de cultivo LB suplementado con ampicilina a una $\text{DO}_{600\text{nm}}$ inicial entre 0.01-0.05. El cultivo se incubó a 37 °C y 250 rpm hasta alcanzar una $\text{DO}_{600\text{nm}}$ final de 0.4-0.6 (4-6 h). Luego se centrifugó a 3350 x g por 15 min a 4 °C. Se desechó el sobrenadante y el paquete celular se resuspendió, suavemente con la ayuda de una micropipeta, en medio LB suplementado con glicerol al 50 % (v/v) y se almacenó a -70 °C hasta su uso.

6.9 Purificación de ADN plasmídico

La purificación y extracción del ADN plasmídico se realizó con el sistema de purificación DNA Wizard (Wizard Plus SV Minipreps, Promega Num. Cat. A1330), utilizando el procedimiento recomendado por el fabricante. Se empleó el método de lisis alcalina y purificación por membrana. Se partió de 2 mL del cultivo en medio LB con ampicilina. Las células se centrifugaron a 8000 rpm por 5 min, el sobrenadante se desechó y el paquete celular se resuspendió con la solución de resuspensión (Tris-HCl 50 mM pH 7.5, EDTA 10 mM, RNAasa A 100 $\mu\text{g}/\text{mL}$). Se añadieron 250 μL de solución de lisis celular (NaOH 200 mM, SDS 1 %), la mezcla se agitó por inversión. Se adicionaron 10 μL de solución alcalina de proteasas, se agitó por inversión y se incubó a temperatura ambiente por 5 min. Luego se añadieron 350 μL de solución neutralizante (hidrocloruro de guanidina 4.09 M, acetato de potasio 759 mM, ácido acético glacial 2.12 M, pH 4.2). La mezcla se agitó y centrifugó a 13000 rpm durante 10 min a temperatura ambiente. El sobrenadante se

decantó en una de las columnas del sistema Wizard Plus SV DNA y se centrifugó a 13000 rpm por 1 min a temperatura ambiente. Se añadieron 750 μL de solución de lavado (acetato de potasio 60 mM, Tris-HCl 8.3 mM pH 7.5, EDTA 0.04 mM pH 8.0, etanol 60%), se centrifugó a 13000 rpm por 1 min, se adicionaron 250 μL de la solución de lavado y se centrifugó a 13000 rpm por 2 min a temperatura ambiente. La columna se transfirió a un tubo nuevo y estéril y se centrifugó a 13000 rpm por 1 min, se añadieron 100 μL de agua libre de nucleasas y se centrifugó a 13000 rpm por 1 min a temperatura ambiente. El plásmido purificado se analizó en un gel de agarosa al 0.8 %.

6.10 Caracterización del plásmido pUCIDT_{GAP} por digestión enzimática

El plásmido pUCIDT_{GAP} aislado de la cepa transformada TOP10 de *E. coli* se caracterizó por digestión enzimática con las enzimas de restricción *Bsu36I* y/o *NotI-HF*. Primero se preparó una mezcla de reacción para cada enzima en dos viales de 200 μL para la muestra y para el control (Tabla 4).

Tabla 4. Volúmenes de cada componente de la mezcla de reacción para la caracterización por digestión enzimática de pUCIDT-T_{GAP}

Reactivo	Muestra	Conc. inicial	Volumen (μL)	Conc. final
Agua	-	-	c.s.	-
CutSmart	-	10 X	2	1 X
ADN	pUCIDT-T _{GAP}	61 ng/ μL	8	24.5 ng/ μL
	pGAHFTEII*	37 ng/ μL	13.5	

*: Control positivo del ensayo. c.s.: cantidades suficientes de agua libre de nucleasas para completar el volumen total de reacción de 19.9 μL .

A cada vial con 19.9 μL de la mezcla de reacción se le adicionó 0.1 μL de las enzimas *Bsu36I* o *NotI-HF*. Los viales se incubaron a 37 °C por 1 h. Luego se inactivó la acción de la enzima *NotI-HF* a 65 °C por 20 min y la *Bsu36I* a 80 °C por 20 min. Se visualizó la eficiencia de corte en una electroforesis de agarosa al 0.8 %. En la caracterización por doble digestión enzimática se realizó el mismo procedimiento, pero adicionando 0.1 μL de la segunda enzima según corresponda. Se estimó el tamaño del fragmento liberado por

electroforesis en gel de agarosa al 0.8 % y se comparó con los tamaños esperados (Tabla 5).

Tabla 5. Tamaño esperado de los fragmentos de interés para la caracterización de pUCIDT-AMP por digestión enzimática con *NotI-HF* y/o *Bsu36I*

Enzima	Banda esperada	Tamaño esperado (pb)	
		pUCIDT-T _{GAP}	pGAHFTEII*
<i>NotI-HF</i> o <i>Bsu 36I</i>	Plásmido linealizado	3379	8580
<i>NotI-HF</i> + <i>Bsu 36I</i>	Fragmento del vector	2759	7727
	Insertar fragmento	620	853

*: Control positivo del ensayo.

6.11 Aislamiento y purificación del vector pGAHFTEII y el inserto T_{GAP} a escala preparativa

Se purificaron los plásmidos pGAHFTEII y pUCIDT_{GAP} con el sistema Wizard-Plus-SV-minipreps-DNA-purification-system (Promega, Madison, WI, EUA). Se cuantificó la concentración y calidad de ambos plásmidos. Se realizó la mezcla de reacción (Tabla 6) y los plásmidos purificados se trataron con las enzimas de restricción *NotI-HF* y *Bsu36I*.

Tabla 6. Composición de la mezcla de reacción para la digestión enzimática con *NotI-HF* y/o *Bsu36I*

Reactivos	Conc. final	Volumen de reacción (98 µL)	
		T _{GAP} (inserto) Volumen (µL)	pGAHFTEII Volumen (µL)
Agua	---	55.2	28
Buffer 10 X	1 X	10	10
ADN	0.02 µg/µL	32.8	60

Se dividieron los 98 µL a razón de 45 µL en dos viales de 0.2 mL, uno para la enzima *NotI-HF* y el otro para la enzima *Bsu36I*. A cada vial se le adicionó 1.5 µL de una de las enzimas, se incubó durante 1 h a 37 °C. Luego se inactivó la enzima de restricción (*Bsu36I* a 80 °C por 20 min y *NotI-HF* a 65 °C por 20 min). Se visualizó la eficiencia del corte en una electroforesis de agarosa al 0.8 % a escala analítica. Se realizó el mismo procedimiento para la digestión con la segunda enzima, según corresponda.

Cada uno de los plásmidos doblemente digeridos (con volumen aproximado de 50 µL) se aplicaron a la electroforesis de agarosa 0.8 % a escala preparativa. Se establecieron las

condiciones de la corrida en 100 V, 85 mA y 75 min. Los fragmentos se cortaron cuidadosamente en el transiluminador empleando una longitud de onda de 365 nm. A cada uno de los fragmentos de interés cortados se les determinó el peso y posteriormente se purificaron por separado con el Kit Wizard® SV gel and PCR Clean-up System (Promega, Madison, WI, EUA). El tamaño de los fragmentos cortados se analizó en una electroforesis al 0.8 % a escala analítica.

6.12 Construcción del vector pP_{GAP}-FTEII-T_{GAP}

Primero se realizó la ligación de los fragmentos aislados y purificados, y para ello inicialmente se estimó la concentración de los fragmentos. Se realizaron los cálculos para tener una relación vector-inserto 3:1 según NEB calculator (<https://nebiocalculator.neb.com/#!/ligation>) lo que correspondió a 50 ng de vector y 20 ng de inserto. Tomando en cuenta estas cantidades se realizó la mezcla de reacción de la muestra y los controles (Tabla 7).

Tabla 7. Composición de la mezcla de reacción para la ligación del vector digerido y el inserto

Reactivos	Control I (μL)	Control II (μL)	Muestra (μL)
Agua	7	2	1
Buffer de ligación	2	2	2
Vector (5 ng/μL)	10	10	10
Inserto (2 ng/μL)	-	6	6
Incubación a 45 °C por 5 min	Sí	No	Sí
T4 ligasa	1	No	1

Control I: Vector sin inserto con ligasa. **Control II:** Vector con inserto sin ligasa. **Muestra:** Vector con inserto y con ligasa.

La reacción de ligación se incubó toda la noche a 4 °C. Posteriormente se verificó la eficiencia de la ligación por electroforesis de agarosa al 0.8 % y el vector obtenido denominado pP_{GAP}-FTEII-T_{GAP} se preservó en células de la cepa TOP10 de *E. coli*, siguiendo los procedimientos previamente descritos en los acápites 6.6 y 6.7.

6.13 Caracterización del vector pP_{GAP}-FTEII-T_{GAP}

Las cepas transformadas crecidas en placas con medio LB y ampicilina, se inocularon en medio LB líquido con ampicilina (triptona 1%, extracto de levadura 0.5 %, NaCl 1 %, ampicilina 100 µg/mL), y se incubaron de 14 a 18 h a 37 °C y 250 rpm. La purificación del ADN plasmídico se realizó por el método de lisis alcalina según se describió en la sección 6.8. El plásmido purificado se analizó por electroforesis en gel de agarosa al 0.8 %.

La caracterización del vector pP_{GAP}-FTEII-T_{GAP} se realizó por doble digestión enzimática según el procedimiento descrito en acápite 6.9, por PCR y por secuenciación. Para realizar la PCR primero se extrajo y purificó el vector por el método de lisis alcalina. Se preparó la mezcla de reacción según se muestra en la Tabla 8 y se añadió 1 µL de plásmido (volumen final de 25 µL).

Tabla 8. Mezcla de reacción para caracterizar el vector pP_{GAP}-FTEII-T_{GAP} por PCR

Reactivo	Conc. Final (µM)	Volumen en la reacción (µL)
Agua	-	13.3
Buffer Green	1 x	5
Primer 5' FTE1	0.5	2.5
Primer 3' 3TH	0.5	2.5
dNTPs	200	0.5
Go Taq	0.04 U/µL (1 U)	0.2

Se utilizó el programa para el PCR que se muestra en la Tabla 9.

Tabla 9. Condiciones del programa de PCR para caracterizar el vector pP_{GAP}-FTEII-T_{GAP}

Paso	T (°C)	T (min)	No. ciclos
Desnaturalización inicial	95	1	1
Desnaturalización	95	1	
Hibridación	56	0.5	30
Extensión	72	1.5	
Extensión final	72	5	1
Enfriamiento	4	∞	1

Para la caracterización por secuenciación se utilizaron los oligonucleótidos FTE1 y 3TH y un secuenciador ABI Prism 310 (Apply Biosystems, Foster City, CA, EUA) de la Unidad de Biología Molecular, del Instituto de Fisiología Celular, en la Universidad Nacional Autónoma de México.

6.14 Linealización y diafiltración del vector pP_{GAP}-FTEII-T_{GAP}

El vector pP_{GAP}-FTEII-T_{GAP} se purificó con el sistema Pure YieldTM plasmid Midiprep System (Promega, Madison, WI, EUA). Se determinó la concentración por espectrofotometría a DO_{260nm} y el plásmido contenido en aproximadamente 250 µL (104 ng/µL) se concentró en un sistema de concentración al vacío “*speed back*” a un volumen final aproximado de 10 µL, en un tiempo de 1:20 h. Se determinó la concentración final por espectrofotometría a 260 nm (3.99 µg/µL para pP_{GAP}-FTEII-T_{GAP} y 2.7 µg/µL para pGAHFTEII), la calidad del ADN purificado por la relación de absorbancia DO_{260nm}/DO_{280nm} y DO_{260nm}/DO_{230nm}, y la integridad del material genético por electroforesis en gel de agarosa al 0.8 %. Los vectores pP_{GAP}-FTEII-T_{GAP} y pGAHFTEII purificados y concentrado se digirió con *Sall*. La mezcla de reacción se preparó según se muestra en la siguiente Tabla 10.

Tabla 10. Mezcla de reacción para linealizar los vectores pP_{GAP}-FTEII-T_{GAP} o pGAHFTEII con *Sall*

Reactivo	Muestra	Conc. final	Volumen de reacción (20 µL)
ADNp	pP _{GAP} -FTEII-T _{GAP}	1.3 µg/µL	6.5
	pGAHFTEII	(26 µg)	10
Buffer	-	1X	2
<i>Sall</i>	-	50 U/µL	1
Agua	-	-	c.s

c.s. Cantidades suficientes de agua libre de nucleasas para completar el volumen total de reacción.

La mezcla se incubó a 37 °C por 16 h. La inactivación se realizó a 65 °C por 20 min. La eficiencia del corte se checó por electroforesis en gel de agarosa al 0.8 %.

El plásmido linealizado se diafiltró en un cartucho Centricón con 3K de tamaño de poro (Merck, NY, EUA). El lavado se llevó a cabo en 3 ciclos de diafiltración con 10 diavolumenes de agua. La centrifugación se realizó a 14000 x g y 4 °C. Terminado el tercer ciclo se colectaron 100 µL de la muestra por inversión del vial y centrifugación a 1000 x g durante 2 min a 4 °C. El volumen colectado se concentró en un sistema de concentración al vacío “*speed back*” a un volumen final aproximado de 10 µL.

Este procedimiento se utilizó también para linealizar y diafiltrar el vector pGAHFTEII.

6.15 Transformación de las células electrocompetentes de la cepa KM71 de *K. phaffii*

A partir de una conserva de células KM71 (Invitrogen Num. Cat. K1730-01), se inocularon 50 µL de células en 10 mL de caldo YPD (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %), y se incubó toda la noche a 30 °C y agitación de 250 rpm. Las células reactivadas, se inocularon en 50 mL de caldo YPD fresco contenidos en un matraz de 250 mL (Pyrex No. 4980 Stopper No. 10) a una DO_{600nm} inicial de 0.1. Las células se cultivaron a 30 °C y en agitación a 250 rpm hasta alcanzar una DO_{600nm} entre 0.6-1.0 (tiempo aproximado de 4-6 h). El cultivo se centrifugó durante 5 min a 2000 x g y 4 °C. Se decantó el sobrenadante y el paquete celular se resuspendió en 10 mL de YPD suplementado con buffer HEPES (Research Organics Num. Cat. 9592H-100GR) (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %, Hepes 20 mM). El paquete celular resuspendido con YPD-HEPES se vertió en un matraz de 250 mL (Pyrex No. 4080 Stopper No. 9). Luego se añadieron cuidadosamente 250 µL de DTT a 1 mol/L (Promega Num. Cat. V3151) y se incubó a 30 °C y 250 rpm por 15 min. Posteriormente, se agregaron 40 mL de agua fría estéril al matraz con la mezcla y se centrifugó el volumen total hasta obtener el paquete celular. El paquete celular se resuspendió en 25 mL de agua fría estéril y se centrifugó nuevamente. Ambos lavados con el agua estéril se realizaron a 4 °C y 250 rpm por 10 min. El paquete celular obtenido se resuspendió en 2 mL de sorbitol 1 mol/L frío y estéril (Sigma Num. Cat. S1876), y se centrifugó a 2000 x g por 10 min a 4 °C. Se decantó el sobrenadante y el paquete celular se resuspendió en 100 µL de sorbitol 1 mol/L frío y estéril.

Se realizaron alícuotas de 40 μL de células competentes y se colocaron en tubos de microcentrífuga de 2.0 mL (MBP Num. Cat. 3434) estériles. Las alícuotas empleadas en la transformación se almacenaron por 12 h a $-20\text{ }^{\circ}\text{C}$. El resto de las alícuotas se preservaron envueltas en varias capas de papel secante a $-70\text{ }^{\circ}\text{C}$.

Previo a la transformación de *K. phaffii*, las celdas de electroporación (BioRad Num. Cat. 1652086) se colocaron en hielo por 2 min. Para la transformación, a una de las alícuotas con las células competentes se le añadió 4 μL del plásmido pP_{GAP}-FTEII-T_{GAP} o de pGAHFTEII previamente linealizados con *Sall* y diafiltrados. La mezcla de células y ADN se colocó en las celdas (en condiciones ambientales de un flujo laminar) y éstas a su vez se colocaron cuidadosamente en el tren de carga del electroporador (BioRad Gene Pulser II) habiendo ajustado previamente las condiciones eléctricas (Tabla 11).

Tabla 11. Condiciones de operación del electroporador para la transformación de células electrocompetentes de la cepa KM71 de *K. phaffii*

Celda (mm)	Volumen de células (μL)	Voltaje (V)	Capacitancia (μF)	Resistencia (Ω)	Fuerza del campo (kV/cm)	Tiempo de pulso (mseg)
2	40	1500	25	400	7500	5

Una vez que se llevó a cabo la descarga eléctrica, se retiró cuidadosamente la celda, bajo flujo laminar se le añadió 1 mL de solución de sorbitol a 1 mol/L (previamente esterilizada y enfriada a $4\text{ }^{\circ}\text{C}$), y se homogenizó por inversión de 5 a 10 veces o hasta que la mezcla quedara homogénea. Posteriormente, se tomaron 50, 100, 150, 200 y 500 μL de la mezcla y se sembraron en placas Petri conteniendo medio sólido RBD His⁻ (sorbitol 1 M, dextrosa 2 %, YNB 1.34 %, biotina 4×10^{-5} %, aminoácidos 0.005 %, agar 2 %). Adicionalmente, 10 μL de la mezcla se sembraron en placas Petri conteniendo medio sólido YPD (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %, agar 2 %). Como control positivo de la viabilidad celular se sembraron 10 μL de las células electrocompetentes sin transformar en placas Petri con medio sólido YPD y como control negativo de la transformación se sembraron 50 μL de las células electrocompetentes sin transformar en placas Petri con medio sólido RBD His⁻.

Para todas las variantes se utilizó la técnica de extensión en placa con un asa de Digrafsky (coloquialmente conocida como asa de extensión). Las placas Petri se incubaron a 30 °C por 4-5 días, hasta observar la aparición de colonias blancas y redondas.

Del proceso de transformación y selección se generaron las cepas KM71/P_{GAP}-FTEII-T_{AOXI} (KM71 transformada con el vector pGAHFTEII) y KM71/P_{GAP}-FTEII-T_{GAP} (KM71 transformada con el vector pP_{GAP}-FTEII-T_{GAP}) a las que se les llamó también cepas con la combinación de elementos reguladores P_{GAP}-T_{AOXI} y P_{GAP}-T_{GAP}, respectivamente.

6.16 Caracterización de las cepas P_{GAP}-T_{AOXI} y P_{GAP}-T_{GAP} de *K. phaffii*

La caracterización de las cepas se realizó por PCR de integración. Para realizar la PCR primero se extrajo y purificó el ADN genómico según un procedimiento modificado de la técnica TSNT (Viader-Salvadó et al. 2006). Se preparó la mezcla de reacción de forma independiente por pareja de oligonucleótidos, para ambos se utilizó la misma composición en la mezcla de reactivos según se muestra en la Tabla 8. Los parámetros del programa de PCR se describen en la Tabla 12.

La caracterización se realizó con una pareja de oligonucleótidos GAPF y FTE2 dirigidos a la región del promotor del gen heterólogo y la segunda pareja de oligonucleótidos FTE1 y 3TH, dirigidos a la región del terminador de dicho gen.

Tabla 12. Condiciones del programa de PCR para caracterizar la cepa P_{GAP}-T_{AOXI} y P_{GAP}-T_{GAP} con doble pareja de oligonucleótidos

Paso	T (°C)	T (min)	No. ciclos
Desnaturalización inicial	94	2	1
Desnaturalización	94	1	
Hibridación	53	1	30
Extensión	72	1.5	
Extensión final	72	10	1
Enfriamiento	4	∞	1

El tamaño esperado del fragmento amplificado por pareja de oligonucleótidos se muestra en la Tabla 13.

Tabla 13. Tamaño esperado de los fragmentos amplificados en el PCR de integración con doble pareja de oligonucleotidos

Cepa KM71	Tamaño esperado del amplicón (pb)	
	Oligos: GAPF y FTE2	Oligos: FTE1 y 3TH
P _{GAP} -T _{AOX1}	1311	1047
P _{GAP} -T _{GAP}	1311	814

6.17 Funcionalidad de los elementos reguladores P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}

Se seleccionó una clona de cada una de las cepas KM71 de *K. phaffii*. Las células reactivadas se inocularon en medio YPD con una DO_{600nm} de 0.6. Se incubaron a 30 °C con agitación a 250 rpm hasta alcanzar una DO_{600nm} final entre 6.0 y 8.0. Luego se tomó la cantidad de células necesarias para una DO_{600nm} inicial de 1.4, se centrifugaron a 2000 x g e inocularon en 100 mL de medio BMG-CaCl₂. El cultivo se incubó durante 24 h a 30 °C y con agitación a 250 rpm, con adición de glicerol 1 % a las 14 h de cultivo. A las 10 h de cultivo, previo a la adición del glicerol se tomó muestra para evaluar la funcionalidad de P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP} en la expresión del gen heterólogo *FTEII*.

Se determinó el volumen de cultivo para un total de 50 mg de peso húmedo de biomasa, previamente centrifugada a 3000 x g durante 10 minutos a 4 °C. El paquete celular se resuspendió en el volumen equivalente a 5 veces el volumen del paquete celular con el reactivo RNeasy lysis solution® (Ambion) y se incubó toda la noche a 4 °C. Posteriormente se centrifugaron las células a 3000 x g durante 10 minutos a 4 °C y se decantó el sobrenadante. Finalmente, las muestras fueron etiquetadas y almacenadas a -70 °C para extraer RNA total.

La funcionalidad de los elementos reguladores se comprobó analizando la presencia de transcritos del gen heterólogo *FTEII* por RT-PCR. Por lo que se extrajo y purificó el RNA total por el procedimiento propuesto en el manual del kit SV total RNA Isolation System (Promega, Madison, WI, EUA).

Para llevar a cabo el RT-PCR, el RNA total se trató con DNasa RQ1 libre de RNAsa (Promega, Madison, WI, EUA) (Tabla 14).

Tabla 14. Composición de la mezcla de reacción para el tratamiento del RNA total con DNasa RQ1

Reactivo	Volumen (30 μ L)
Solución de RNA	24 (8 μ L x 3)*
Buffer de DNasa RQ1 Promega (10X)	3
DNasa RQ1 Promega (1 U/ μ L)	3

*: Se utilizó 8 μ L de solución de RNA equivalente a 1 μ g de RNA con 1 μ L de RNasa cada 10 μ L de reacción (según recomendación del fabricante).

La mezcla de reacción se incubó durante 1 h a 37 °C y la reacción se detuvo con la adición de solución “Stop” (Promega, Madison, WI, EUA) a 65 °C por 15 min. La integridad del RNA se evaluó por electroforesis en gel de agarosa al 0.8 %.

Se realizó la retro transcripción con la enzima M-MLV. La mezcla de reacción se preparó según se muestra en la Tabla 15, siguiendo el procedimiento que se describe a continuación: Primero se mezcló el ARN previamente tratado con DNasa RQ1 con el Oligonucleótido (dT)₁₅, la mezcla se calentó a 70 °C por 5 min e inmediatamente se colocó en hielo; luego se agregó el resto de los reactivos de la reacción de transcripción inversa y la mezcla se incubó a 42 °C por 60 min. La actividad enzimática se detuvo por calentamiento de la mezcla a 65 °C durante 15 min. Finalizado el procedimiento las muestras se almacenaron a -20 °C.

Tabla 15. Composición de la mezcla de reacción la retro transcripción del ARNm

Reactivo	Concentración inicial	Volumen (μ L)	Concentración final
Oligonucleótido (dT) ₁₅	0.5 μ g/ μ L	1.0	0.025 μ g/ μ L
Solución de RNA	61.2 ng/ μ L	16.0	32.72 ng/ μ L
Agua tratada con DEPC	0.1%	2.5	0.015%
Buffer de la transcriptasa reversa M-MLV	5X	6.0	1X
dNTP's	10 mM	1.5	0.5 mM
TDT	100 mM	1.0	3 mM
Enzima transcriptasa inversa M-MLV*	200 U/ μ L	2.0	13.3 U/ μ L

*: Para obtener el control negativo del ensayo no se adicionó la enzima transcriptasa inversa M-MLV.

De la retro transcripción se obtuvo el ADNc y para el PCR se realizaron dos mezclas de reacción: una con oligonucleótidos dirigidos a *FTEII* denominados FTE1 (5'-CGCTGACAAGAACGGTTACA-3') y FTE2 (5'-CTTGTGGGTGGAAACCAATC-3') y la otra mezcla de reacción con oligonucleótidos dirigidos al gen endógeno de β -actina, oligonucleótido 5ACT (5'-GCCCCAGAAGAGCACCCAGT-3') y oligonucleótido 3ACT (5'-ACCGGAAGCGTACAGGGACA-3'). La mezcla de reacción se realizó según se muestra en la Tabla 16 a la que se añadió 1 μ L de ADNc.

Tabla 16. Composición de la mezcla de reacción para el PCR de evaluación de funcionalidad de los elementos reguladores P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}

Reactivo	Concentración Inicial	Volumen (μ L)	Concentración final
Agua	-	13.3	-
Green Buffer	5X	5	1X
Oligonucleótido 3'	5 μ M	2.5	0.5 μ M
Oligonucleótido 5'	5 μ M	2.5	0.5 μ M
dNTPs	10 mM	0.5	10 mM
Enzima Go Taq	5 U/ μ L	0.2	0.04 U/ μ L

Se utilizó el programa para el PCR que se muestra en la Tabla 17.

Tabla 17. Condiciones del programa del RT-PCR

Paso	T ($^{\circ}$ C)	Tiempo (min)	No. ciclos
Desnaturalización inicial	95	2	1
Desnaturalización	95	1	
Hibridación para oligonucleótido de <i>FTEII</i>	50	1	
Hibridación para oligonucleótido de β -actina	56	1	30
Extensión	72	2	
Extensión final	72	5	1
Enfriamiento	4	∞	1

El tamaño de los fragmentos amplificados se comparó con los esperados: con los oligonucleótidos dirigidos al gen heterólogo *FTEII* se esperó obtener un tamaño de 536 pb y con los oligonucleótidos a β -actina de 150 pb.

6.18 Cinética de crecimiento

La cinética de crecimiento se realizó en matraz de 250 mL con 25 mL de medio de cultivo con agitación constante, se cultivaron 18 clonas de la cepa P_{GAP}-T_{AOX1} y 25 clonas de la cepa P_{GAP}-T_{GAP}. Se utilizó medio BMG-CaCl₂ con las cinco cepas con mayor rendimiento. Se determinó la densidad celular cada 2 h del cultivo durante 24 h.

Primero se inocularon 100 μ L de cada clona de *K. phaffii* en 10 mL de medio YPD contenidos en tubos de 50 mL, se incubaron a 30 °C con agitación a 250 rpm hasta una densidad óptica final entre 6-8. Se calculó la cantidad suficiente para comenzar el cultivo en medio BMG-CaCl₂ con DO_{600nm} inicial de 1.4. Se tomó el volumen necesario del cultivo en YPD y se centrifugó a 1500 x g por 5 min, se desechó el sobrenadante de cultivo y se resuspendió en medio BMG-CaCl₂ para inocular en 25 mL de este medio. Se incubó a 30 °C durante 24 h a 250 rpm. Se tomaron muestras cada dos horas y se determinó densidad celular.

6.19 Determinación de dosis génica

Las clonas de las cepas P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP} con mayor rendimiento proteínas/biomasa se seleccionaron como posibles multicopias y las clonas de menor rendimiento proteínas/biomasa se seleccionaron como posibles unicopias. Las clonas seleccionadas se cultivaron a nivel de matraz. Primero se inocularon 100 μ L de cada clona en 10 mL de medio YPD y se incubaron a 30 °C y 250 rpm hasta que alcanzaron una DO_{600nm} 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, que se almacenaron a -20 °C para ser empleados en la preparación de ADN genómico (ADNg).

La extracción de ADNg se realizó por el método de extracción con solventes orgánicos (TSNT) (Viader-Salvadó et al. 2006). Se purificó el ADNg obtenido, primero se realizó un tratamiento con la RNase One (Promega) para eliminar el RNA presente en las muestras, añadiendo 1 μ L de la enzima, y se incubó a 37 °C por 1 hora. Posteriormente, se realizó una diafiltración de las muestras en los dispositivos de filtración Amicon Ultra 3-kDa de 0.5 mL (Millipore, MA, EUA). Se adicionaron 50 μ L de cada muestra en los cartuchos acoplados dentro de tubos de 2.0 mL y se añadieron 400 μ L de agua estéril libre de nucleasas a cada uno. Se centrifugaron a 12000 rpm por 10 min a temperatura ambiente y siguiendo el mismo procedimiento, se hicieron 3 lavados adicionales. Finalmente, para recuperar el DNAg 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 ADNg colectado se calentaron a 65 °C por 20 minutos. Se tomaron 3 μ L de cada muestra, se analizaron en un gel de agarosa al 0.8 % y se cuantificaron por espectrofotometría.

La determinación de dosis génica se realizó por qPCR con el sistema de cuantificación Mx3005P (Agilent Technologies, Santa Clara, CA, USA). Se utilizaron sondas de hidrólisis dirigidas a las regiones CDS del gen heterólogo *FTEII* y al gen *GAPDH* endógeno, este último se utilizó como muestra estándar. Se estableció una metodología para el desarrollo del qPCR que se muestra en el Anexo 1.

Las rectas para determinar la eficiencia de amplificación de los genes *FTEII* y *GAPDH* se realizaron con 4 puntos con ADNg 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 μ L, con la composición de la mezcla que se muestra en Tabla 18. Se realizó un ensayo negativo de la reacción de qPCR con el objetivo de verificar la ausencia de ADNg contaminante.

Tabla 18. Composición de la mezcla de reacción para el qPCR de determinación de dosis génica

Reactivo	Volumen (20 μ L)
Agua	2
Master Mix	10
ADNg	4
Prime Time <i>FTEII</i> o <i>GAPDH</i>	4

Con las curvas de eficiencia de amplificación determinadas para *FTEII* y *GAPDH*, se analizó el ADNg de las clonas seleccionadas. El análisis se realizó con 1 ng de ADNg en el ensayo de qPCR y se estimó la cantidad de ADNg del gen *FTEII* y *GAPDH* en el intervalo de linealidad de acuerdo con el valor de Ct. El valor de *FTEII* se normalizó respecto al obtenido para *GAPDH*. El valor normalizado de *FTEII* cercano a 1 se consideró 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 ADNg (eje de ordenadas) y el valor de Ct (eje de las abscisas) de las muestras standards. Se utilizó la ecuación (1):

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

Donde E es la eficiencia y m es la pendiente de la recta del gráfico de logaritmo de la cantidad de ADNg contra valor de Ct.

La cantidad del ADNg de los genes *FTEII* y *GAPDH* en las clonas seleccionadas como posibles multicopias o posibles unicopias (muestras problemas) se calculó a partir de la ecuación (2):

$$ADNg_{GAPDH \text{ o } FTEII} = 10^{(Ct-b)/m} \quad (2)$$

Donde $ADNg_{GAPDH \text{ o } FTEII}$ es la cantidad de ADN del gen endógeno *GAPDH* o del gen heterólogo *FTEII* de las muestras problema. Ct es el valor promedio del ciclo humbral de la muestra problema, m y b son la pendiente y el intercepto de la recta del logaritmo de la cantidad de ADNg vs Ct de las muestras standard.

El valor de dosis génica se determinó por la ecuación (3):

$$Dosis\ g\acute{e}nica = \frac{ADNg_{FTEII}}{ADNg_{GAPDH}} \quad (3)$$

Donde $ADNg_{GAPDH}$ o $FTEII$ es la cantidad de ADN del gen end3geno *GAPDH* o del gen heter3logo *FTEII* de las muestras problema calculado por la ecuaci3n (2).

6.20 Determinaci3n de concentraci3n de prote3nas totales (m3todo de Bradford)

Para la determinaci3n de la concentraci3n de prote3nas totales extracelulares se utilizaron los sobrenadantes de cultivos en BMG de los cultivos de *K. phaffii* P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}}. La concentraci3n de prote3nas totales se determin3 miendiendo la absorbancia a 595 nm en un NanoPhotometer® Pearl (NanoPhotometer Pearl, Implen GmbH, Munich, Alemania) y correlacion3ndola con los valores de concentraci3n de BSA de la curva est3ndar (Tabla 19).

Tabla 19. Curva de calibraci3n del m3todo de Bradford con BSA

Curva est3ndar* (μg/mL)	BSA (50 μg/mL) (μL)	Agua destilada (μL)
Blanco para la medici3n de la absorbancia	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

*: En cada uno de los puntos de la curva adem3s del volumen de BSA y de agua a3adido, se a3icion3 800 μL de reactivo de Bradford para un volumen total de 1000 μL

Cada 200 μL de diluciones del standard o de muestra se le a3icionaron 800 μL del reactivo Bradford (las preparaciones se realizaron por triplicado). El reactivo de Bradford se prepar3 con 20 mg de colorante azul de Coomassie G-250 mezclado con agitaci3n vigorosa con 10 mL de etanol al 95 %. A la mezcla se le a3icionaron 20 mL de 3cido fosf3rico al 85 % y se agit3 vigorosamente por 20 o 30 min para finalmente completar con agua destilada a un volumen final de 200 mL. El reactivo de Bradford se filtr3 con filtro

de papel Whatman no 1, se almacenó a temperatura ambiente con protección de la luz solar y se usó en un tiempo igual o menor de 15 días.

6.21 Análisis de proteínas por SDS-PAGE

La electroforesis de proteína se realizó en condiciones desnaturalizante por la presencia de β -mercaptoetanol en la solución tampón para la dilución de las muestras, denominado “buffer sample”. El gel separador se preparó al 12 % de acrilamida/bisacrilamida (Tabla 20).

Tabla 20. Composición de reactivos que conforman el gel separador preparado al 12 % de acrilamida/bisacrilamida para la electroforesis de proteína por SDS-PAGE

Reactivo	Volumen (μ L)
Agua destilada	3100
Acrilamida/bisacrilamida	3200
Tris-HCl 2 M, pH 8.8	1500
Hoja de datos de seguridad (FDS) 10%	80
PSA 10%	50
TEMED	10

El gel concentrador se preparó con la mezcla de reactivos que se muestran en la siguiente Tabla (Tabla 21).

Tabla 21. Composición de reactivos que conforman el gel concentrador de la electroforesis de proteína por SDS-PAGE

Reactivo	Volumen (μ L)
Agua destilada	3050
Acrilamida/bisacrilamida	600
Tris-HCl 2 M, pH 6.8	250
Hoja de datos de seguridad (FDS) 10%	40
PSA 10 %	10
TEMED	5

Para el gel tapón se utilizó 1 mL del gel separador, 4 μ L de PSA 10 % y 8 μ L de TEMED.

La corrida electroforética se realizó en un tiempo aproximado de 1 h a 25 mA y de 6 a 8 °C, en la solución de corrida diluida a partir de una solución tampón 5 veces concentrada compuesta por Tris-HCl 0.125 M, glicina 0.96 M, SDS 0.5 % a pH 8.8.

Terminada la corrida el gel se sumergió en solución de tinción preparada con azul brillante de Coomassie R250 al 0.20 %, metanol al 50% y ácido acético al 10 %. Una vez teñido, el gel se lavó con agua destilada y se destiñó la zona del gel no ocupada por proteínas con solución decolorante preparada con metanol al 5% y ácido acético al 7 %.

6.22 Determinación de actividad enzimática

A las muestras del sobrenadante de cultivo se les realizó un cambio de tampón por cromatografía de exclusión molecular empleando una columna PD-10 Sephadex G-25. La columna se equilibró con solución tampón Tris-HCl 100 mM pH 8.5, 50 mM de NaCl y 2 % de glicerol. Una vez colectada la muestra de interés se le adicionó CaCl_2 para una concentración final de 1 mM. La determinación de la actividad de fitasa se llevó a cabo mediante el método descrito por Guerrero-Olazarán et al. (Guerrero-Olazarán et al. 2010).

Se empleó como sustrato fitato de sodio (sal dodecasódica del ácido hexafosfórico de inositol) (Sigma Aldrich Corporation, PM: 770.04). Se preparó buffer de actividad (100 mM de Tris-HCl pH 7.5 y 1 mM de CaCl_2). Antes de comenzar, el buffer de actividad se precalentó a 37 °C por 30 min. Se realizaron tres réplicas para cada punto de la recta de calibración. Se utilizó una solución stock de fosfato de potasio monobásico 1.64 mM con buffer de actividad y se prepararon los estándares de fosfato según se muestra en la tabla Tabla 22.

Tabla 22. Volúmenes para la preparación de los estándares para la curva de calibración

Estándar de fosfato (mM)	Volumen requerido del estándar (μ L)	Volumen de buffer de actividad (μ L)
0.0	0	100
0.57	35	65
1.02	62	38
1.64	100	0

Los estándares de fosfato se incubaron a 37 °C durante 30 minutos en un Thermomixer R (Eppendorf). Se agregaron 600 µL de TCA al 15 %, y posteriormente 800 µL de colorante recién preparado (ácido ascórbico 10%, molibdato de amonio 2.5 %, ácido sulfúrico 6 N). Se incubaron los tubos de los estándares a 37 °C por 45 min, y se midió la absorbancia a 820 nm en un espectrofotómetro Beckman Coulter DU650. Se determinó la recta de calibración, para la interpolación de los valores de las muestras. En el ensayo se consideraron dos blancos: el blanco de sustrato (sólo sustrato sin fitasa) y el blanco de muestra (sin el sustrato).

Para la reacción enzimática se realizaron mezclas de reacción en tubos de 2.0 mL, por triplicado según los volúmenes que se describen en la Tabla 23.

Tabla 23. Mezcla de reacción para la determinación de actividad de fitasa

Blanco de sustrato	Volumen (µL)	Blanco de muestra	Blanco de muestra	Muestra	Volumen (µL)
Buffer de actividad	40	Muestra	40	Muestra	40
Sustrato	60	Buffer de actividad	60	Sustrato	60

Las mezclas y los blancos correspondientes se incubaron a 37 °C durante 30 minutos en un Thermomixer R (Eppendorf). La reacción se detuvo agregando 600 µL de TCA al 15 %, posteriormente se agregaron 800 µL de colorante recién preparado. Se incubaron a 37 °C por 45 min, se midió la absorbancia a 820 nm en un espectrofotómetro Beckman Coulter DU650.

Cálculo de la actividad volumétrica enzimática empleando las siguientes ecuaciones:

$$APL = APM - ABF - ABM \quad (4)$$

Cálculo de la absorbancia generada por el fosfato liberado en la mezcla de reacción. 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 (ecuación 5) 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.

$$x = \frac{(y-b)}{m} \quad (5)$$

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 y y m es la pendiente de la recta de calibración.

Finalmente se calculó la actividad volumétrica de la Fitasa por la ecuación (6):

$$AV_f (U/mL) = \frac{x}{t} * FD * 2.5 \quad (6)$$

Donde AV_f es la actividad volumétrica de fitasa en la muestra, x es la concentración (mM) de fosfato libre en la mezcla de reacción calculada por la ecuación 5, 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 μ L de muestra en un volumen final de reacción de 100 μ L. El valor de APM debe estar por debajo de la absorbancia del estándar de mayor valor de concentración y que x/t este comprendido entre 0.0050 y 0.083.

Para el cálculo del rendimiento actividad fitasa/biomasa ($Y_{p/x}$), se dividió la actividad volumétrica de fitasa final en U/L entre la densidad celular en g/L de células secas. El $Y_{p/x}$ se midió en unidades de fitasa producidas por gramo de biomasa (U/g).

6.23 Efecto del terminador transcripcional en el gen heterólogo sobre los niveles de transcripción del gen heterólogo *FTEII* y el gen endógeno *GAPDH*, y en el $Y_{p/x}$ extracelular

Las dos cepas de copia única se cultivaron en medios BMGlc y BMGly (DO_{600nm} inicial de 0.2) durante 24 h a 30 °C y 250 rpm. Las muestras se recolectaron en cinco momentos de cultivo (3, 6, 12, 18 y 24 h) para evaluar el crecimiento por DO_{600nm} . Las concentraciones de biomasa en g células secas/L se estimaron teniendo en cuenta que 1 DO_{600nm} son aproximadamente 0.3 g/L de células secas. Se estimaron las velocidades específicas de crecimiento (μ) en la fase de crecimiento exponencial (3 a 12 h) a partir de

la pendiente de la curva del logaritmo natural de la concentración de biomasa vs. tiempo de cultivo. Los datos de concentración de biomasa vs. tiempo de cultivo también se ajustaron a la solución de integración de la ecuación logística de Verhulst-Pearl (López et al. 2010) minimizando la suma residual de cuadrados entre la ecuación integrada y las concentraciones de biomasa, utilizando la herramienta Solver de Microsoft Excel (Microsoft Co., Redmond, Washington, EE.UU.). Las concentraciones máximas de biomasa (X_{max}) se obtuvieron directamente de las ecuaciones ajustadas. Se estimaron valores puntuales de μ a las 6, 12 y 18 h de cultivo a partir de la relación entre la primera derivada de las curvas ajustadas de células de crecimiento y la concentración de biomasa en cada momento, tal como se describió anteriormente (Cos et al. 2005) y considerando los volúmenes de muestreo como despreciables.

Las actividades de las combinaciones de elementos reguladores ($P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$) sobre la expresión de genes heterólogo *FTEII* y *GAPDH* endógeno en cada cepa de unicopia se determinaron mediante la PCR cuantitativa con transcripción inversa (RT-qPCR) utilizando el mismo termociclador que se empleó para la determinación del número de copias de genes. Se utilizó el gen *YPT1* como normalizador. Las células de 6, 12 y 18 horas de cultivo se recolectaron por centrifugación, y se conservaron en una solución de RNlater hasta realizar el aislamiento de ARN. La síntesis de ADNc se realizó a partir de 1 μ g de ARN utilizando el kit de síntesis de ADNc SCRIPT y oligo(dT)₂₀, siguiendo las instrucciones del fabricante. Los ADNc *FTEII* y *GAPDH* de cada muestra se amplificaron por duplicado utilizando el sistema PrimeTime que contiene primers y sondas de hidrólisis dirigidas a las regiones CDS del gen heterólogo *FTEII* o al gen *GAPDH* endógeno. El ADNc de *YPT1* se amplificó utilizando la mezcla de qPCR SybrMaster y los cebadores 5qYPT1 y 3qYPT1 como se describió anteriormente (Caballero-Pérez et al. 2021). Los niveles de transcripción de *FTEII* y *GAPDH* se calcularon como “fold change” del nivel de transcripción de *YPT1* de la misma muestra utilizando una curvas de calibración previamente determinadas con ADN genómico para cada amplificación de ADNc de cada gen.

Además, los sobrenadantes en cada momento de muestreo se concentraron y diafiltraron por ultrafiltración como se describió anteriormente (Robainas-del-Pino et al. 2023). El

$Y_{p/x}$ extracelular se estimó como la relación entre la actividad volumétrica de la fitasa extracelular y el aumento de la concentración de biomasa desde el inicio del cultivo.

Se calcularon cocientes de los niveles de transcripción de *FTEII* y cocientes de $Y_{p/x}$ extracelulares para comparar los valores obtenidos de los cultivos de la cepa P_{GAP}-T_{AOXI} con los de la cepa P_{GAP}-T_{GAP} utilizando la misma fuente de carbono (cocientes inter cepas), y utilizando la misma cepa y diferentes fuentes de carbono como la glucosa y el glicerol (cocientes fuente de carbono) para dilucidar el impacto del terminador transcripcional dentro de la estructura génica heteróloga influenciado por la fuente de carbono en los eventos postranscripcionales, como la traducción y/o la secreción de proteínas heterólogas.

6.24 Análisis estadístico

El análisis estadístico se realizó con el programa “Statgraphics Centurion XVI.II” y análisis de datos del programa Microsoft Excel (Microsoft Co., Redmond, Washington, EE.UU.). Se realizó análisis de correlación lineal y evaluación de parámetros de estadística descriptiva. Los valores de densidad celular se evaluaron mediante una prueba de bondad de ajuste para verificar su conformidad con una distribución normal. Adicionalmente, se calcularon los intervalos de confianza al 95 % para la media poblacional y la desviación estándar verdadera con la finalidad de determinar si las clonas seleccionadas como muestras pueden considerarse representativas de la población. Los valores de μ , los niveles de transcripción, los valores de $Y_{p/x}$ extracelular, las proporciones de niveles de transcripción *FTEII* y las relaciones $Y_{p/x}$ extracelular entre cepas unicopias se compararon estadísticamente mediante una prueba t de Student con un nivel de significancia de 0.05.

7 RESULTADOS

7.1 Diseño y síntesis de la secuencia nucleotídica del terminador del gen *GAPDH* denominada T_{GAP}

El análisis de mapeo de transcritos en tres cepas KM71 contra el genoma de la cepa CBS7435 de *K. phaffii*, mostró la presencia de lecturas de ARN a lo largo de toda la región inter-CDS de los genes *GAPDH* y *NAB6*, lo que imposibilitó definir la región del terminador del gen *GAPDH* (Figura 7).

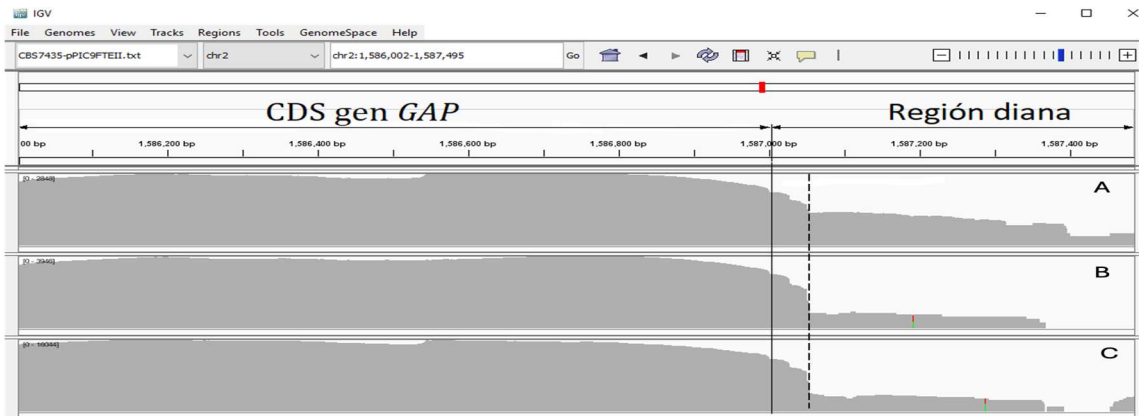


Figura 7. Mapeo del transcriptoma de tres cepas KM71.

Cepas transformadas con tres vectores diferentes contra el genoma de la cepa CBS7435. Se representan las regiones: CDS del gen *GAPDH* desde la posición 1586002 pb a 1587004 pb, la región diana desde la posición 1587005 pb a 1587495 pb y la región UTR del gen *GAPDH* desde la posición 1587005 pb a 15870058 pb [representada con (- -)]. A. transcriptoma de la cepa KM71 con vector pPIC9FTEII. B. transcriptoma de la cepa KM71 con vector pPIC9 sin gen heterólogo. C. transcriptoma de la cepa KM71 con vector pPIC9Lac412.

En este contexto, el ensayo 3'RACE fue concluyente para definir la región 3'UTR del gen *GAPDH* en cepas KM71. En la primera etapa de este ensayo, se obtuvo ARN total purificado y tratado con DNasa RQ1 con niveles de pureza que cumplieron con los parámetros de calidad establecidos, con una relación $DO_{260nm}/280nm$ de 2.037 (criterio de aceptación de aproximadamente 2.0) y una relación $DO_{260nm}/230nm$ de 1.628 (rango permitido de 1.8 a 2.2). La retrotranscripción del ARN utilizando el oligonucleótido T17AP permitió la obtención del ADNc de los genes con poli (A). La PCR realizada con los oligonucleótidos dirigidos al CDS del gen *GAPDH* y al adaptador incorporado por PCR en la cola poli (A), produjo un fragmento amplificado de aproximadamente 350 pb (Figura 8).

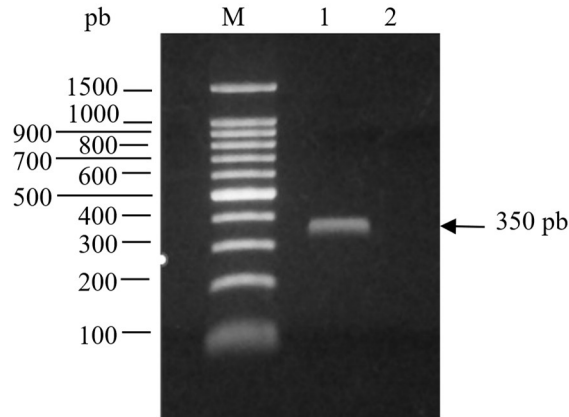


Figura 8. Electroforesis en gel de agarosa al 1.5 % del fragmento de ADNc amplificado, producto de la PCR del 3'RACE.

Fragmento enviado a secuenciar para identificar la secuencia de la región 3'UTR del gen *GAPDH*. Carril **M**. Marcador de peso molecular 100 pb DNA Ladder (Promega). Carril **1**. Fragmento amplificado del ADNc purificado (aprox. 350 pb). Carril **2**. Control negativo del PCR. La flecha indica el fragmento de ADNc amplificado por PCR con oligonucleótidos específicos a la región terminadora del gen *GAPDH*.

El análisis de este fragmento mediante secuenciación empleando un oligonucleótido específico dirigido al CDS del gen *GAPDH*, reveló el punto exacto de la secuencia nucleotídica donde se realizó el corte enzimático y se agregó la cola poli (A) (Figura 9). Esto permitió identificar con precisión el sitio de corte y poliadenilación del terminador *GAPDH*. Dicho sitio, con secuencia "TA", se localizó a 51 nt río abajo del CDS del gen *GAPDH*. La secuencia de estos 51 nt resultó 100 % idéntica a la secuencia 3'UTR del gen *GAPDH* (*TDH3*) de la cepa CBS 7435 de *K. phaffii*.

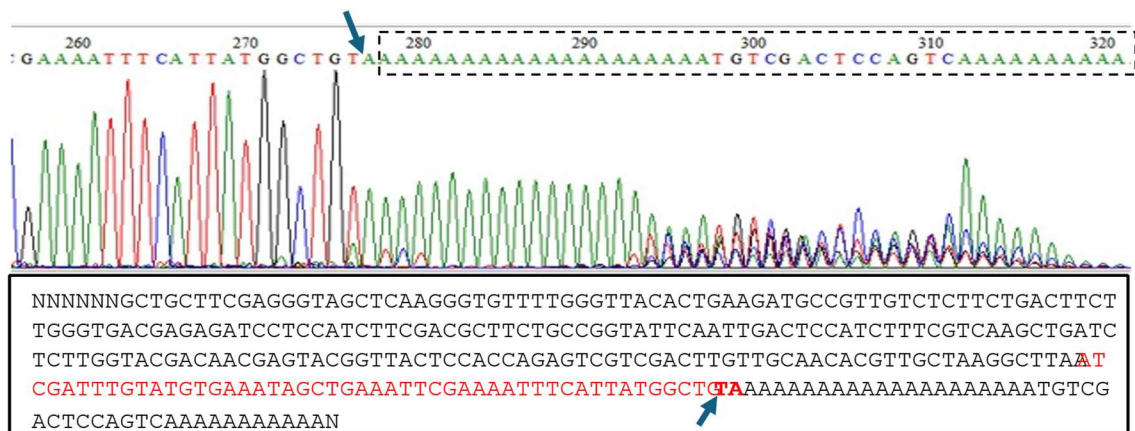


Figura 9. Electroferograma y secuencia producto de la secuenciación del fragmento amplificado por PCR resultante del 3'RACE.

Se indica la secuencia "TA" como sitio de corte y poliadenilación del gen *GAPDH* (señalada con flecha), se enmarca con líneas discontinuas la cola de poli (A) y sitio de unión del oligonucleótido T17AP. Se presenta la secuencia nucleotídica de la región 3'UTR del gen *GAPDH* (señalada en rojo).

Para el diseño de la región T_{GAP}, a la secuencia nucleotídica de la región 3'UTR se le agregaron 50 nt, utilizando la misma secuencia que la anotada para la cepa CBS 7435 río abajo del sitio de poliadenilación, por lo que se consideró como terminador del gen *GAPDH* la región de 101 nt río debajo de su CDS.

En el diseño de la secuencia sintética se incluyeron varias secuencias: la de los sitios de corte de las enzimas *NotI* y *Bsu36I* (15 nt), la comprendida entre los sitios de corte en el vector pGAHFTEII (510 nt), y la del terminador T_{GAP} (101 nt), por lo que, la secuencia a sintetizar resultó con 620 nt y quedó diseñada con la secuencia que se muestra a la derecha en la Figura 10.

La secuencia diseñada fue sintetizada y clonada en el plásmido pUCIDT-AMP por la compañía IDT, EUA resultando el plásmido pUCIDT_{GAP} de 3379 pb (Figura 10).

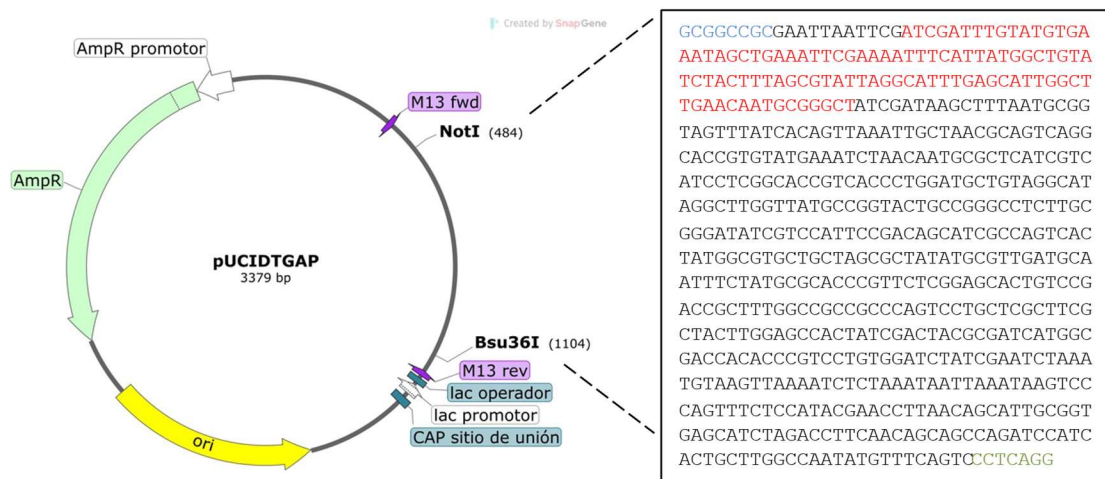


Figura 10. Representación esquemática del plásmido pUCIDT_{GAP}.

En pUCIDT_{GAP} se suministró la secuencia sintética T_{GAP} por la compañía IDT Inc. (Coralville, IA, EUA). A la derecha se muestra la secuencia sintética que contiene el terminador T_{GAP} y se representan con colores cada una de las secuencias que la componen: en azul el sitio de corte para la enzima de restricción *NotI*, en negro parte de la secuencia del vector pGAHFTEII (vector de expresión empleado en el subclonaje de T_{GAP}), en color rojo la secuencia T_{GAP} (terminador del gen *GAPDH*) y en color verde el sitio de corte para la enzima *Bsu36I*.

Los resultados del análisis *in silico* de la digestión enzimática de pUCIDT_{GAP} muestran los tamaños esperados de los fragmentos digeridos (Tabla 24).

Tabla 24. Tamaño esperado de los fragmentos generados producto de la digestión de pUCIDT_{GAP} con las enzimas *NotI-HF* y/o *Bsu36I*.

Ezima	Fragmento	Tamaño esperado (pb)	
		pGAHFTEII*	pUCIDT _{GAP}
<i>NotI-HF</i> o <i>Bsu36I</i>	Plásmido linealizado	8586	3379
	Vector	7733	2759
<i>NotI-HF</i> + <i>Bsu36I</i>	Inserto	853**	620**

*: Control positivo del ensayo. **: La diferencia de tamaño entre los insertos de ambos vectores es de 233 pb, diferencia de tamaño apreciable y útil para caracterización de los vectores por digestión enzimática.

Estos resultados esperados se obtuvieron luego de la caracterización enzimática de T_{GAP} (Figura 11).

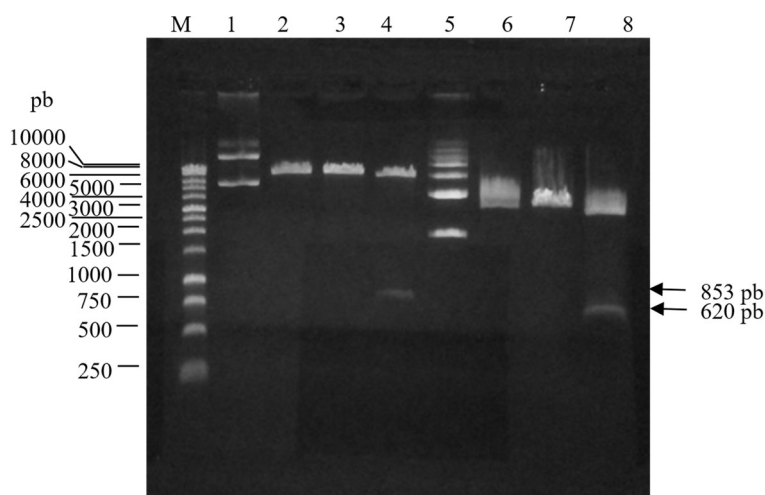


Figura 11. Electroforesis en gel de agarosa al 0.8 % para caracterizar el plásmido pUCIDT_{GAP} proveniente de la transformación en la cepa TOP10 de *E. coli*.

Carril **M**. Marcador de tamaño molecular de 1 kb (Promega, EUA). Carril **1**. Plásmido pGAHFTEII purificado (Control positivo). Carril **2**. pGAHFTEII digerido con *NotI-HF* (NEB, UK). Carril **3**. pGAHFTEII digerido con *Bsu36I* (NEB, UK). Carril **4**. pGAHFTEII con doble digestión enzimática. Carril **5**. pUCIDT_{GAP} extraído y purificado de la cepa TOP10 de *E. coli*. Carril **6**. pUCIDT_{GAP} digerido con *NotI-HF* (NEB, UK). Carril **7**. pUCIDT_{GAP} digerido con *Bsu36I* (NEB, UK). Carril **8**. pUCIDT_{GAP} con doble digestión enzimática.

En la Figura 11 carril 6 y 7 se observa que el plásmido pUCIDT_{GAP} linealizado tiene un tamaño aproximado entre 3000 pb y 4000 pb, lo que concuerda con el tamaño esperado de 3379 pb. Luego de la doble digestión enzimática se desprendió un fragmento de 620 pb, de esta forma se corroboró la presencia de T_{GAP} en pUCIDT_{GAP}.

7.2 Modelo de señales de poliadenilación en el terminador *GAPDH*

En el análisis de alineamiento de 51 hexanucleótidos reportados en la literatura como secuencias conservadas presentes en los terminadores de genes de levaduras (Van Helden et al. 2000) con los 101 nt de T_{GAP} se identificaron los cuatro elementos de procesamiento que en conjunto conforman las señales de poliadenilación del terminador del gen *GAPDH* (Figura 12).

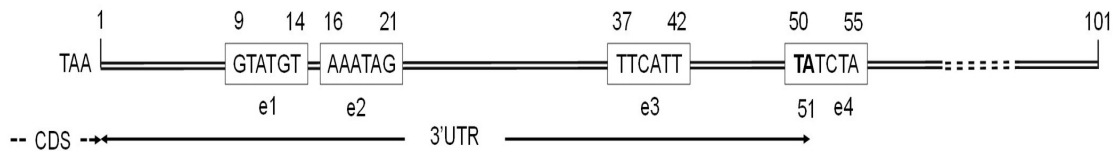


Figura 12. Representación esquemática de la secuencia T_{GAP}.

La región 3'UTR se indica con una flecha de dos puntas. Los números corresponden a las posiciones de los nucleótidos del sitio de inicio 3'UTR. Los supuestos elementos de procesamiento 3' e1, e2, e3 y e4, también conocidos como elementos de eficiencia, posicionamiento, próximo río arriba del sitio de corte, y próximo río abajo del sitio de corte, respectivamente, se muestran en rectángulos. También se muestra el codón de stop (TAA) de la secuencia codificante (CDS) de gen *GAPDH* y la secuencia TA (en negrita) para el sitio de corte y poliadenilación del ARN.

Según reportes de Van Helden et al. y Graber et al. (Van Helden et al. 2000; Graber et al. 1999) el elemento de eficiencia (e1) dentro de la secuencia 3'UTR es el más alejado río arriba del sitio de corte y poliadenilación. Se ha reportado que los residuos T ubicados en posiciones intercaladas son los nucleótidos más esenciales en esta secuencia. El elemento de posicionamiento (e2) es una zona rica en A, que se localiza usualmente entre 13 y 27 nt río arriba de sitio de corte y poliadenilación. La distancia entre e1 y e2 es variable. El sitio de corte y poliadenilación es una secuencia del tipo PyAn (TA o CA) y por último una zona rica en U que flanquea el sitio de corte y poliadenilación. En T_{GAP} (Figura 12) se observa claramente cada uno de estos cuatro elementos. La secuencia consenso GTATGT como elemento de eficiencia (e1). La secuencia AAATAG como elemento de posicionamiento (e2) ubicada aproximadamente a 27 nucleótidos río arriba del sitio de corte y poliadenilación. La secuencia TA como el sitio de corte y poliadenilación, y la zona rica en U que flanquea el sitio de corte y poliadenilación (e3 y e4) con las secuencias TTCATT y TATCTA.

7.3 Construcción del vector de expresión pP_{GAP}-FTEII-T_{GAP}

7.3.1 Clonaje y digestión *in silico*

Del análisis *in silico* del clonaje de T_{GAP} en pGAHFTEII (Figura 13) resultó el vector pP_{GAP}-FTEII-T_{GAP} (Figura 14) con 8353 pb (7733 pb del vector y 620 pb del inserto T_{GAP}).

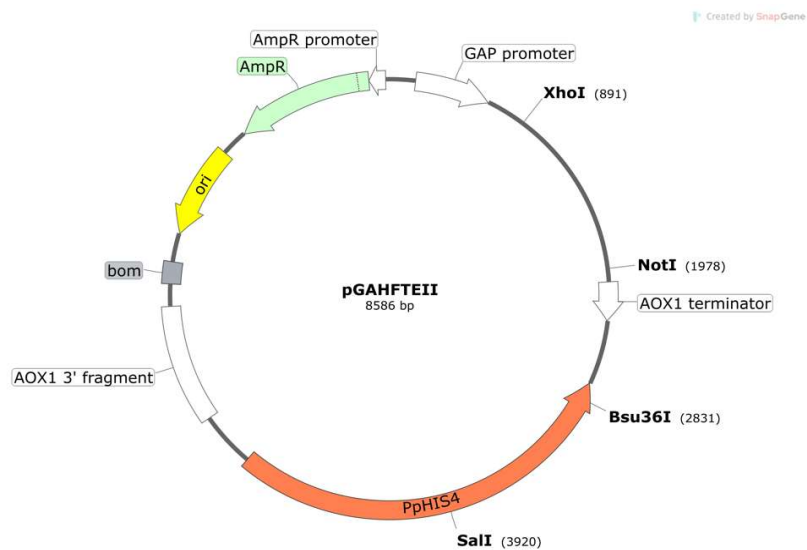


Figura 13. Representación esquemática del vector de expresión pGAHFTEII.

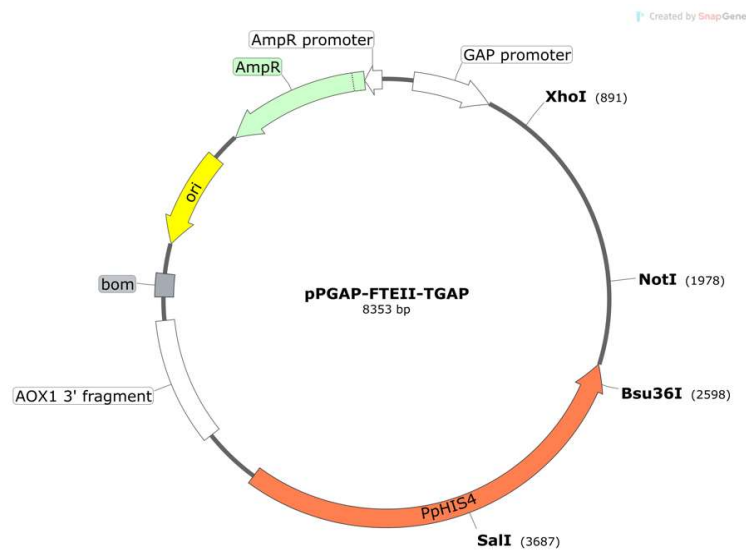


Figura 14. Representación esquemática del vector de expresión pP_{GAP}-FTEII-T_{GAP}.

Del análisis de la digestión enzimática *in silico* con *Bsu36I* y *NotI-HF* (Tabla 24) se apreció que existe una diferencia de 233 pb entre los fragmentos liberados (inserto) en el vector pGAHFTEII respecto al pP_{GAP}-FTEII-T_{GAP}, lo cual es favorable para poder analizar la correcta construcción de pP_{GAP}-FTEII-T_{GAP} mediante caracterización enzimática en un gel de agarosa.

7.3.2 Ligación vector-inserto

A partir de una electroforesis de agarosa a escala preparativa se aislaron, purificaron y prepararon para la ligación los fragmentos: vector pGAHFTEII e inserto T_{GAP} (Figura 15 A y B, respectivamente).

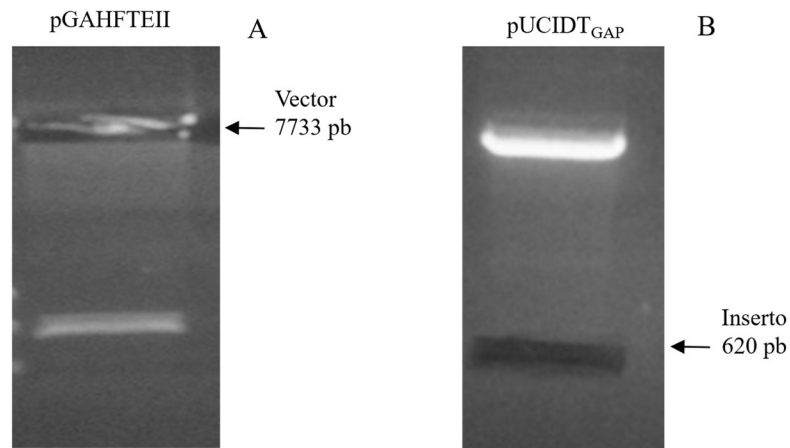


Figura 15. Electroforesis en gel de agarosa al 0.8 % a escala preparativa de los fragmentos de interés obtenidos de la doble digestión enzimática con *NotI-HF* y *Bsu36I* (NEB, UK) de los plásmidos purificados. **A.** pGAHFTEII. **B.** pUCIDT_{GAP}. Las flechas indican los fragmentos de interés cortados para la construcción de pP_{GAP}-FTEII-T_{GAP}.

Luego de purificados los fragmentos con el Kit Wizard® SV gel and PCR Clean-up System (Promega, EUA) se analizaron en un gel de agarosa al 0.8 % a escala analítica. Se observó que las bandas de los fragmentos purificados se encontraban en el tamaño esperado, 7733 pb para el vector y 620 pb para el inserto (Figura 16).

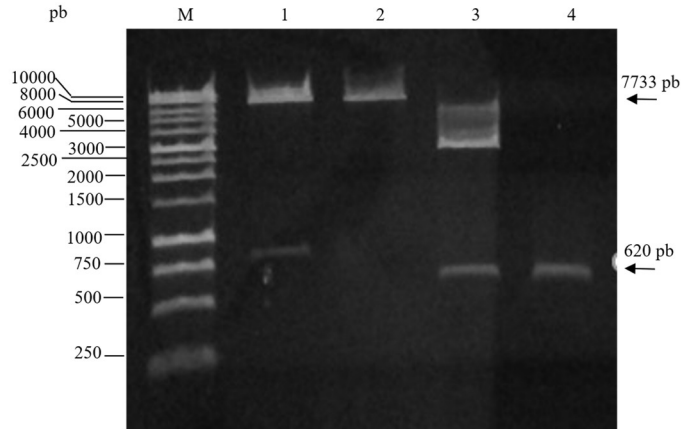


Figura 16. Electroforesis en gel de agarosa al 0.8 % de los fragmentos purificados (vector e inserto) obtenidos a escala preparativa.

Carril **M**. Marcador de tamaño molecular 1Kb DNA Ladder (Promega, EUA). Carril **1**. Plásmido pGAHFTEII purificado y digerido con las enzimas de restricción NotI -HF y Bsu36I (NEB, UK). Carril **2**. vector purificado proveniente de pGAHFTEII. Carril **3**. Plásmido pUCIDT_{GAP} purificado y digerido con las enzimas de restricción NotI-HF y Bsu36I (NEB, UK). Carril **4**. Inserto purificado obtenido de pUCIDT_{GAP}.

Para la ligación y obtención de pP_{GAP}-FTEII-T_{GAP} se consideró la concentración estimada de 5 ng/μL para el vector y de 2 ng/uL para el inserto. Se realizaron los cálculos para 20 μL de reacción y una relación 3:1 vector-inserto.

Una vez realizada la ligación, se transformaron las células competentes de la cepa TOP10 de *E. coli* (Figura 17) con el vector resultante pP_{GAP}-FTEII-T_{GAP} para su preservación y propagación.

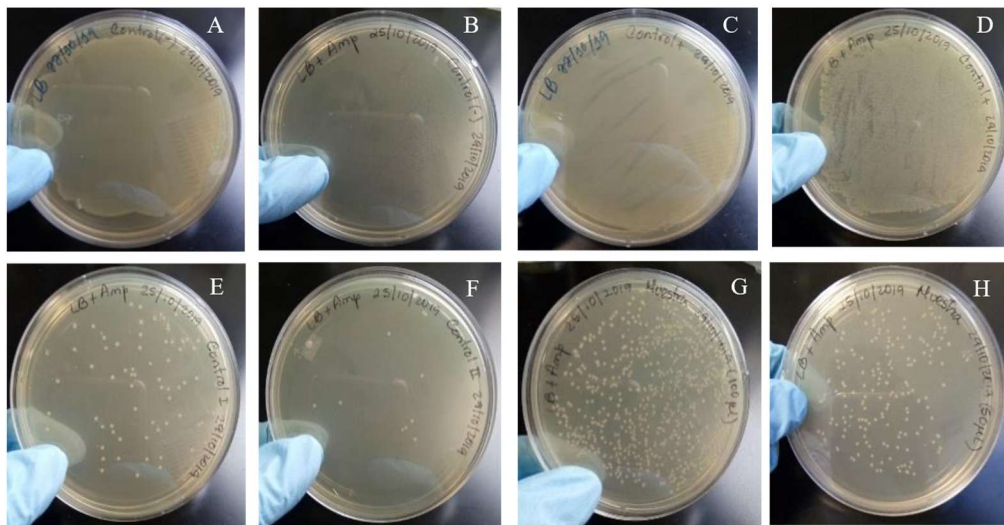


Figura 17. Siembra en placas de la cepa TOP10 de *E. coli* transformadas con el vector pP_{GAP}-FTEII-T_{GAP}. **A y B**. Control negativo en medio LB y en medio LB con ampicilina, respectivamente. **C y D**. Control positivo en medio LB y en medio LB con ampicilina, respectivamente. **E**. Control I (reacción de ligación

con vector, sin inserto y con ligasa) en medio LB con ampicilina. **F.** Control II (con vector con inserto sin ligasa) en medio LB con ampicilina; **G y H.** Muestra (con vector, inserto y ligasa) en medio LB con ampicilina aplicada a 100 μ L y 50 μ L, respectivamente.

El vector pP_{GAP}-FTEII-T_{GAP} resultó con un casete de expresión conformado por el promotor constitutivo del gen *GAPDH*, una región codificante de la proteína recombinante FTEII asociada en su extremo N-terminal con el péptido prepro del factor α de *S. cerevisiae* y el terminador del gen *GAPDH* (Figura 18). A esta construcción se le denominó P_{GAP}-T_{GAP}.

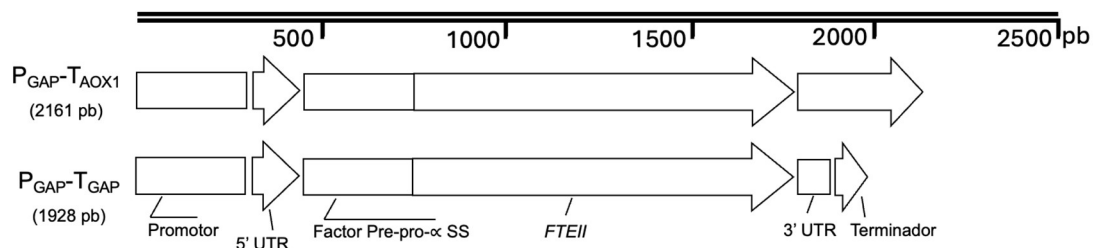


Figura 18. Representación esquemática del casete de expresión de los vectores pGAP-FTEII (P_{GAP}-T_{AOX1}) y pP_{GAP}-FTEII-T_{GAP} (P_{GAP}-T_{GAP}).

Los resultados de la caracterización por PCR y corte enzimático de pP_{GAP}-FTEII-T_{GAP} se muestran en las Figura 19 A y B.

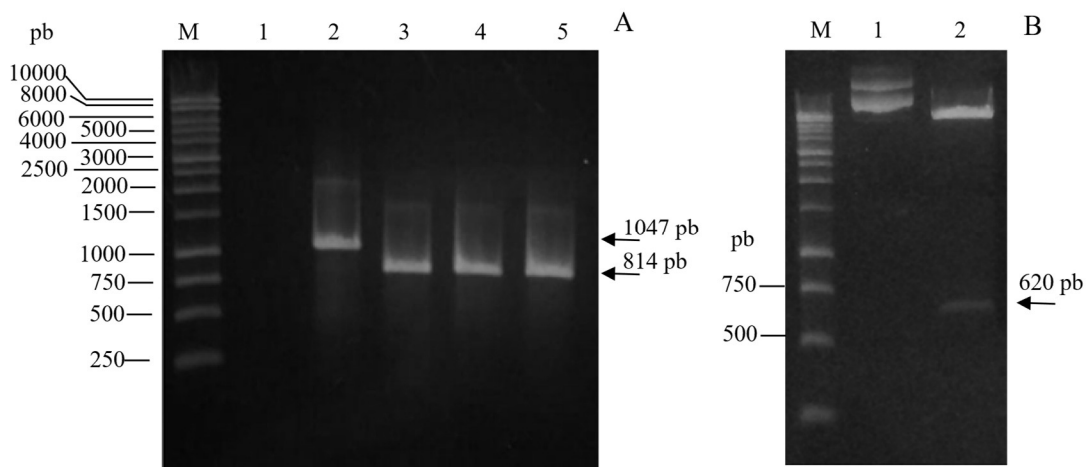


Figura 19. Electroforesis en gel de agarosa al 0.8 % para la caracterización de pP_{GAP}-FTEII-T_{GAP}. **A.** PCR con pareja de oligos FTE1/3TH dirigidos a la región del terminador. Carril **M.** Marcador de tamaño molecular 1 Kb DNA Ladder (Promega, EUA). Carril **1.** Control negativo de la PCR, Carril **2.** Control positivo de la PCR, vector pGAHFTEII. Carriles del **3 - 5.** Vector purificado pP_{GAP}-FTEII-T_{GAP} proveniente de tres colonias seleccionadas. **B.** digestión enzimática con *NotI-HF* y *Bsu36I*. Carril **M.** Marcador de tamaño molecular 1 Kb DNA Ladder (Promega, EUA). Carril **1.** Plásmido pP_{GAP}-FTEII-T_{GAP} purificado. Carril **2.** Plásmido pP_{GAP}-FTEII-T_{GAP} digerido con *NotI-HF* y *Bsu36I*.

De ambos análisis se obtuvieron los resultados esperados: un fragmento en la región del terminador amplificado por PCR, con tamaño de 814 pb y luego de la doble digestión enzimática se desprendió un fragmento con el tamaño esperada de 620 pb, lo cual evidenció la correcta construcción del vector pP_{GAP}-FTEII-T_{GAP}.

Se seleccionó una de las clonas evaluadas y luego de purificar el vector pP_{GAP}-FTEII-T_{GAP} (Figura 20) se secuenció con los oligonucleótidos FTE1 y 3TH.

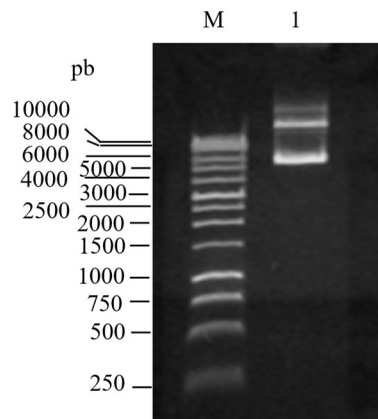


Figura 20. Electroforesis en gel de agarosa al 0.8 % del plásmido pP_{GAP}-FTEII-T_{GAP}. Carril M. Marcador de peso molecular 1 Kb DNA Ladder (Promega, EUA). Carril 1. Plásmido extraído de la cepa TOP10 de *E. coli* y purificado por columna con el sistema Wizard-plus-sv-minipreps-dna-purification (Promega, EUA).

El resultado de la secuenciación corroboró que después del subclonaje de T_{GAP} en el vector pP_{GAP}-FTEII-T_{GAP} su secuencia nucleotídica, de 101 pb, se mantuvo 100 % idéntica a la secuencia del terminador del gen *GAPDH* (*TDH3*) de la cepa CBS 7435 de *K. phaffii* (Figura 21).

Sequence ID: Query_528005 Length: 101 Number of Matches: 1

Range 1: 1 to 101 [Graphics](#) ▼ Next Match ▲

Score	Expect	Identities	Gaps	Strand
187 bits(101)	4e-52	101/101(100%)	0/101(0%)	Plus/Plus
Query 897	ATCGATTTGTATGTGAAATAGCTGAAATTCGAAAATTTTCATTATGGCTGTATCTACTTTA	956		
Sbjct 1	ATCGATTTGTATGTGAAATAGCTGAAATTCGAAAATTTTCATTATGGCTGTATCTACTTTA	60		
Query 957	GCGTATTAGGCATTTGAGCATTGGCTTGAACAATGCGGGCT	997		
Sbjct 61	GCGTATTAGGCATTTGAGCATTGGCTTGAACAATGCGGGCT	101		

Figura 21. Resultado de la comparación por Blastn de las secuencias nucleotídicas del terminador *GAPDH* (*TDH3*) de la cepa CBS 7435 de *K. phaffii* y la secuencia obtenida por secuenciación de pP_{GAP}-FTEII-T_{GAP}.

7.4 Construcción de las cepas de *K. phaffii* KM71/P_{GAP}-FTEII-T_{GAP} y KM71/P_{GAP}-FTEII-T_{AOX1}

7.4.1 Linealización con *Sall*

Los plásmidos pGAHFTEII y pP_{GAP}-FTEII-T_{GAP}, después de ser extraídos y purificados, resultaron con concentraciones de 125 ng/μL y 104 ng/μL, respectivamente; ambos en un volumen de 530 μL y con buena calidad (DO_{260nm/280nm} de 1.786 y 1.756, respectivamente). Ambas muestras se concentraron en un CentriVap (LABCONCO, EUA) hasta un volumen final aproximado de 10 μL, obteniéndose concentraciones de 2690 ng/μL para pGAHFTEII y 3990 ng/μL para pP_{GAP}-FTEII-T_{GAP}, y una relación DO_{260nm/280nm} de 1.815 y 1.814, respectivamente (Figura 22).

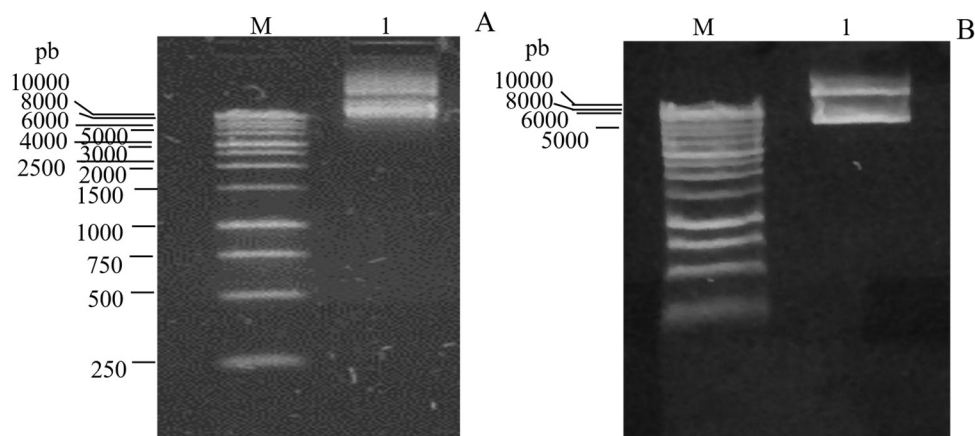


Figura 22. Electroforesis de agarosa al 0.8 % de los plásmidos purificados: **A.** pGAHFTEII y **B.** pP_{GAP}-FTEII-T_{GAP}. Carriles **M.** Marcador de tamaño molecular 1 Kb DNA Ladder (Promega). Carriles **1.** ADN plasmídico purificado.

Se linealizaron ambos plásmidos con *Sall* (Figura 23). Después de concentrar cada plásmido linealizado se diafiltraron contra agua y se obtuvieron con una concentración de 288 ng/μL y 245 ng/μL en 10 μL y con una relación DO_{260nm/280nm} de 1.789 y 1.885 para pGAHFTEII y pP_{GAP}-FTEII-T_{GAP}, respectivamente.

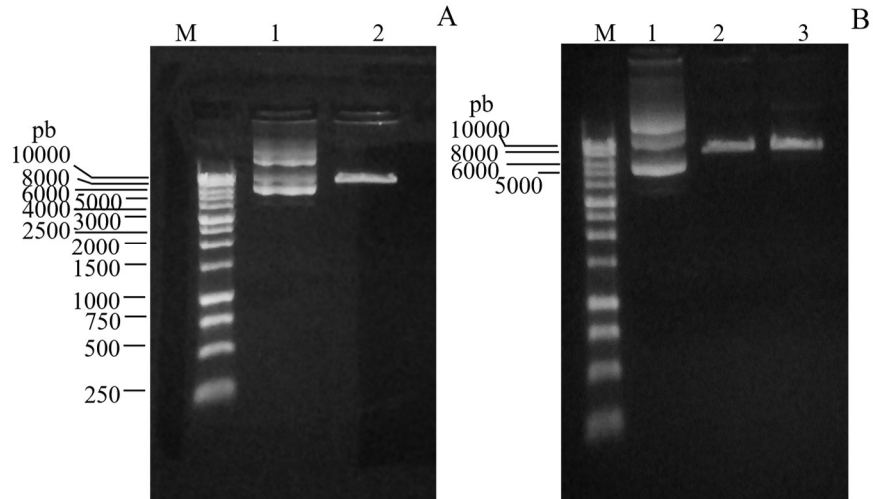


Figura 23. Electroforesis en gel de agarosa al 0.8 % de la linealización con *Sall*.

A. pGAHFTEII y **B.** pP_{GAP}-FTEII-T_{GAP}. Carriles **M.** Patrón de tamaño molecular de 1kb DNA Ladder (Promega, EUA). Carriles **1.** ADN plasmídico. Carriles **2.** ADN plasmídico digerido con la enzima *Sall*. Carril **3.** ADN plasmídico digerido con la enzima *Sall*, concentrado y diafiltrado (solo en el caso de la figura B).

7.4.2 Transformación de la cepa KM71 de *K. phaffii* con los vectores pGAHFTEII y pP_{GAP}-FTEII-T_{GAP}

Se transformaron las células electrocompetentes KM71 de *K. phaffii* con los vectores linealizados pGAHFTEII y pP_{GAP}-FTEII-T_{GAP}. Las cepas transformadas se seleccionaron por prototofía a histidina al crecer en placas con medio RBD His⁻ (Figuras 24 y 25).

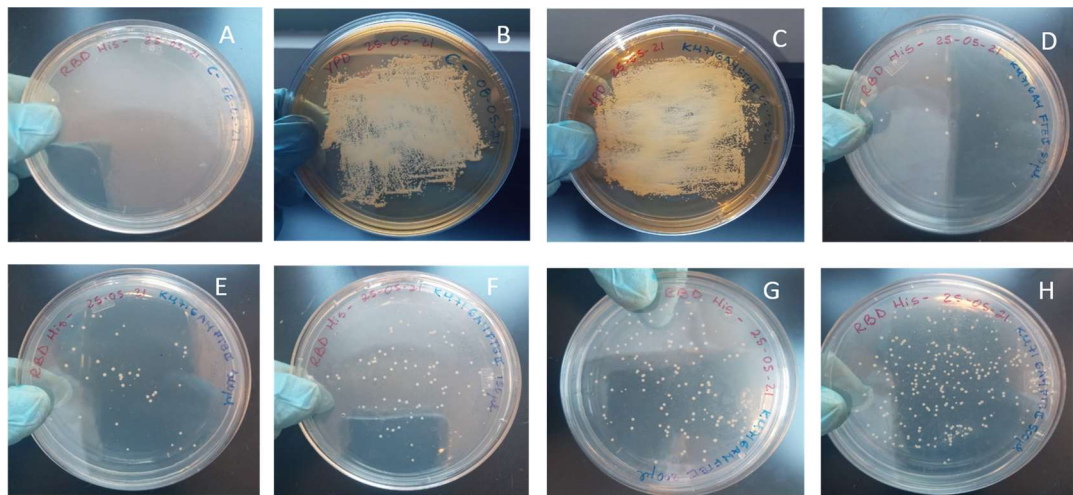


Figura 24. Siembra en placas de la cepa KM71 de *K. phaffii* transformada con el plásmido pGAHFTEII.

A. placa con medio RBD His⁻ con células sin transformar. **B.** Placa con medio YPD agar con células sin transformar. **C.** Placa con medio YPD agar con células transformadas. **D - H.** Placas con medio RBD His⁻ con células transformadas (siembra de 50, 100, 150, 200 y 500 µL, respectivamente).

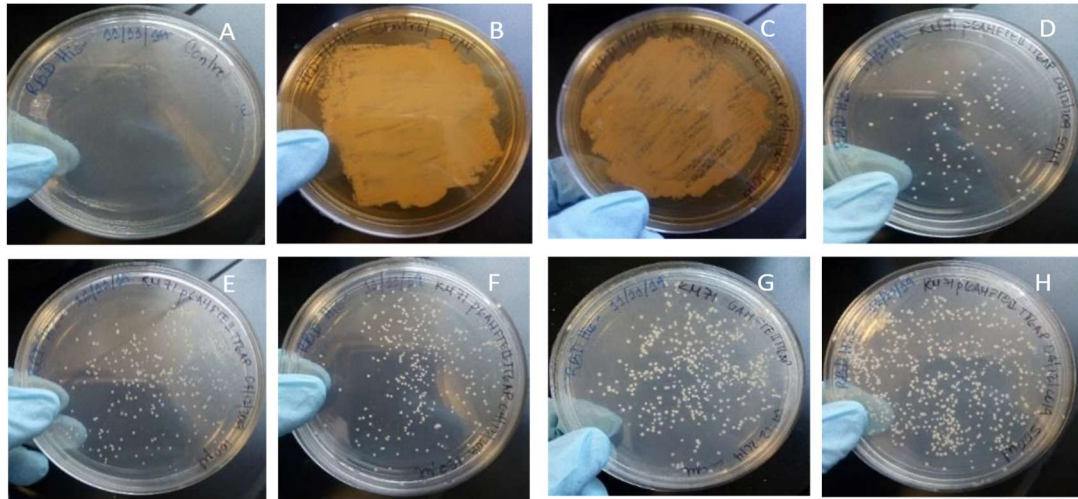


Figura 25. Siembra en placas con la cepa KM71 de *K. phaffii* transformada con el plásmido pP_{GAP}-FTEII-T_{GAP}. **A.** placa con medio RBD His⁻ con células sin transformar. **B.** YPD agar con células sin transformar. **C.** YPD agar con células transformadas. **D - H.** Placa con medio RBD His⁻ con células transformadas (siembra de 50, 100, 150, 200 y 500 µL, respectivamente).

Se obtuvieron más de 1000 clonas producto de la transformación que crecieron en medio RBD His⁻. Las cepas transformadas KM71/P_{GAP}-FTEII-T_{AOX1} y KM71/P_{GAP}-FTEII-T_{GAP} seleccionadas por prototofía a histidina se caracterizaron por PCR con el par de oligonucleótidos FTE1/3TH dirigidos al terminador del gen heterólogo (Figuras 26 y 27) integrado al genoma de *K. phaffii*. Estas cepas las denominamos P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}, respectivamente.

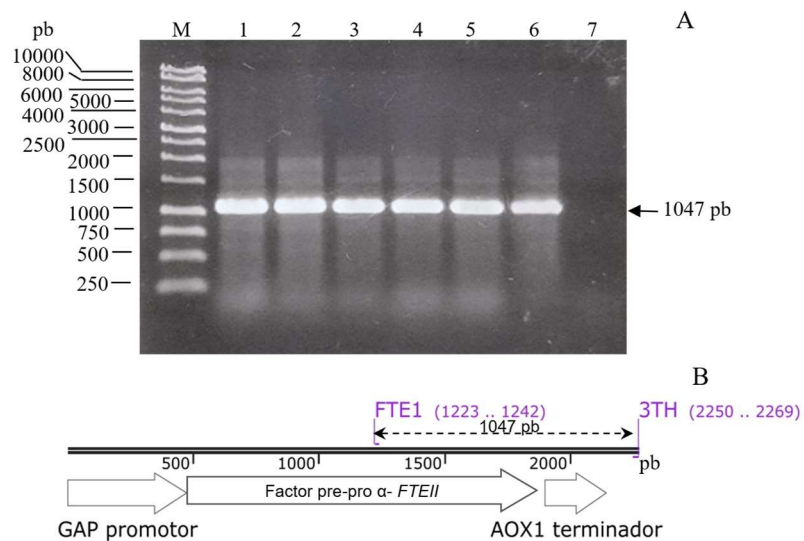


Figura 26. **A.** Electroforesis en gel de agarosa al 0.8 % de la caracterización de la cepa P_{GAP}-T_{AOX1} por PCR empleando el par de oligonucleótidos FTE1/3TH, dirigidos a la región del terminador del gen heterólogo.

Carril **M**. Patrón de tamaño molecular de 1kb DNA Ladder (Promega, EUA). Carriles **1 – 5**. Producto amplificado del ADN_g de cinco clonas transformantes P_{GAP}-T_{AOX1}. Carril **6**. Producto amplificado del plásmido pGAHFTEII como control positivo del ensayo. Carril **7**. Control negativo de PCR. **B**. Representación esquemática de los sitios de hibridación de los oligonucleótidos FTE1/3TH en el casete de expresión integrado en el genoma de la cepa P_{GAP}-T_{AOX1}.

En cinco de las clonas transformadas P_{GAP}-T_{AOX1} y en la clona control positivo proveniente de la cepa KM71GAHFTEII con casete de expresión con P_{GAP}-T_{AOX1} se amplificó una banda de tamaño esperado de 1047 pb. Lo que indica la presencia de T_{AOX1} en el gen heterólogo incorporado al genoma de las cepas transformadas y presente en la cepa control positivo.

En cuatro de las clonas transformadas P_{GAP}-T_{GAP} se amplificó una banda de tamaño esperado para el terminador T_{GAP}, de 814 pb (Figura 27). Este terminador es 233 pb menor que el T_{AOX1} y es la diferencia observada en los tamaños de los fragmentos amplificados para la región del terminador entre las cepas transformadas y la cepa control positivo (con 1047 pb).

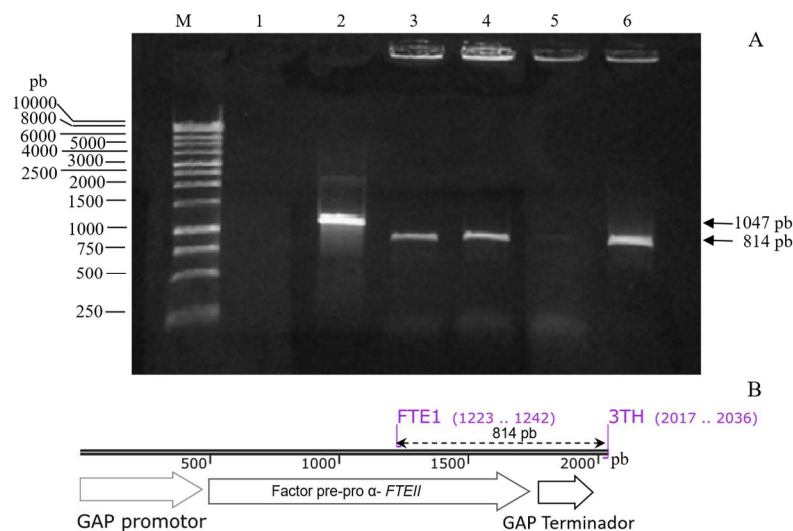


Figura 27. **A**. Electroforesis en gel de agarosa al 0.8 % de la caracterización de la cepa P_{GAP}-T_{GAP} por PCR empleando con el par de oligonucleótidos FTE1/3TH, dirigidos a la región del terminador del gen heterólogo.

Carril **M**. Patrón de tamaño molecular de 1kb DNA Ladder (Promega, EUA). Carril **1**. Control negativo de PCR. Carril **2**. Producto amplificado del plásmido pGAHFTEII como control positivo del ensayo. Carriles del **3 – 6**. Producto amplificado del ADN_g de cuatro clonas transformantes de la cepa P_{GAP}-T_{GAP}. **B**. Representación esquemática de los sitios de hibridación de los oligonucleótidos FTE1/3TH en el casete de expresión integrado en el genoma de la cepa P_{GAP}-T_{GAP}.

Estos resultados confirmaron la presencia del casete de expresión con el terminador T_{AOX1} o T_{GAP} insertado en el ADN_g de las cepas transformantes P_{GAP}-T_{AOX1} o P_{GAP}-T_{GAP}, según corresponde.

El ADN_g extraído de ambas cepas también se analizó por PCR de integración con doble pareja de oligonucleótidos (Figura 28). Se utilizaron los oligonucleótidos GAPF/FTE2 dirigidos al promotor y los oligonucleótidos FTE1/3TH dirigidos al terminador del gen heterólogo *FTEII*. La región del promotor del casete de expresión en las cepas P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP} es idéntica, es por ello que en ambas cepas se amplificó un fragmento de tamaño esperado de 1311 pb con los oligonucleótidos dirigidos a la región del promotor; mientras que con los oligonucleótidos dirigidos a la región del terminador el tamaño del fragmento amplificado resultó de 1047 pb para T_{AOX1} y de 814 pb para T_{GAP}, con una diferencia de 233 pb que es la diferencia de longitud entre ambos terminadores.

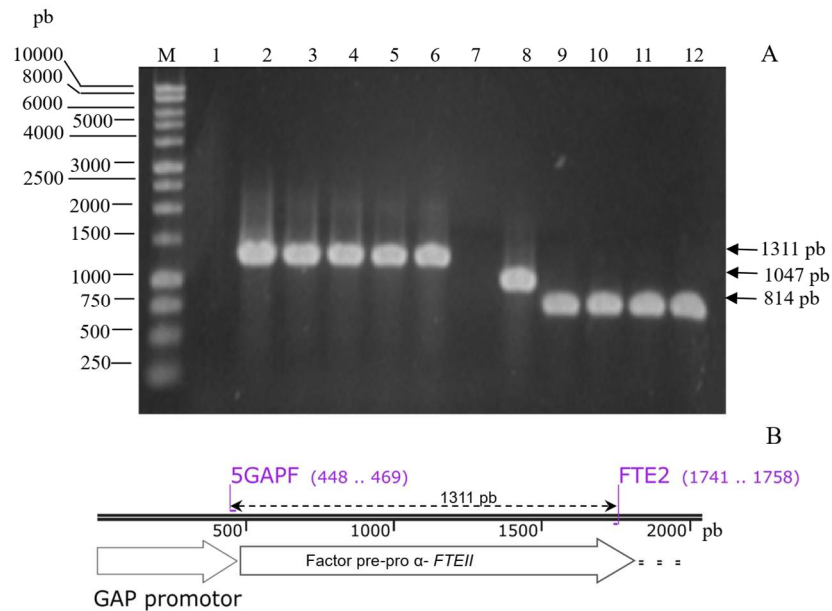


Figura 28. **A.** Electroforesis en gel de agarosa para la verificación por PCR de la integración del casete de expresión en el genoma de *K. phaffii*.

Carril **M.** marcador de tamaño molecular de 1kb DNA Ladder; carriles **1 y 7**, controles negativos de PCR. Carril **2**, producto de PCR de ADN genómico de la cepa basada en P_{GAP}-T_{AOX1} con el par de oligonucleótidos GAPF/FTE2. Carriles **3 a 6**, producto de PCR de ADN genómico de cuatro clones de la cepa basada en P_{GAP}-T_{GAP} con el par de oligonucleótidos GAPF/FTE2. Carril **8**, producto de PCR de ADN genómico de la cepa basada en P_{GAP}-T_{AOX1} con el par de oligonucleótidos FTE1/3TH. Carriles **9 a 12**, producto de PCR de ADN genómico de cuatro clones de la cepa basada en P_{GAP}-T_{GAP} con el par de oligonucleótidos FTE1/3TH. **B.** Representación esquemática de los sitios de hibridación del par de oligonucleótidos GAPF/FTE2 en el casete de expresión integrado en el genoma de las cepas P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}.

Estos resultados confirmaron la correcta construcción e inserción de los casetes de expresión P_{GAP}-T_{AOX1} o P_{GAP}-T_{GAP} en las cepas transformadas KM71 de *K. phaffii*.

7.5 Cultivo celular

Se seleccionaron de forma aleatoria e independiente 18 clonas de la cepa $P_{GAP-TAOX1}$ y 25 clonas de la cepa $P_{GAP-TGAP}$. Todas las clonas en ambas cepas se cultivaron en igualdad de condiciones: en medio BMG suplementado con 0.1 % de $CaCl_2$ a temperatura de 30 °C y agitación a 250 rpm; a partir de una DO_{600nm} inicial aproximada de 1.4, con adición de glicerol al 1 % a las 14 h de cultivo y durante 24 h. En estas condiciones de cultivo se observó similitud en el crecimiento celular para ambas cepas. La fase de crecimiento exponencial (donde la velocidad específica de crecimiento es constante y máxima) ocurrió entre las 2 y 10 h de cultivo a una velocidad específica de crecimiento aproximado de 0.19 h^{-1} (Figura 29). A las 24 h de cultivo el valor de la concentración de biomasa de la cepa $P_{GAP-TAOX1}$ estuvo entre 7.22 y 9.30 g/L y en la cepa $P_{GAP-TGAP}$ entre 7.39 y 9.23 g/L (Figura 30 A y B, respectivamente; valores resaltados en rojo).

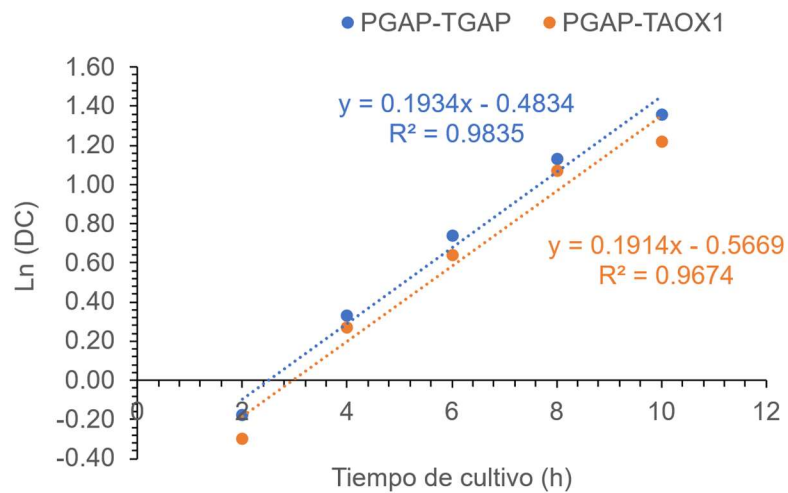


Figura 29. Gráfica del logaritmo natural de la densidad celular (DC) en la fase exponencial del cultivo celular de las cepas $P_{GAP-TAOX1}$ y $P_{GAP-TGAP}$. La pendiente de las rectas representa la velocidad específica de crecimiento máxima (μ) para ambas cepas.

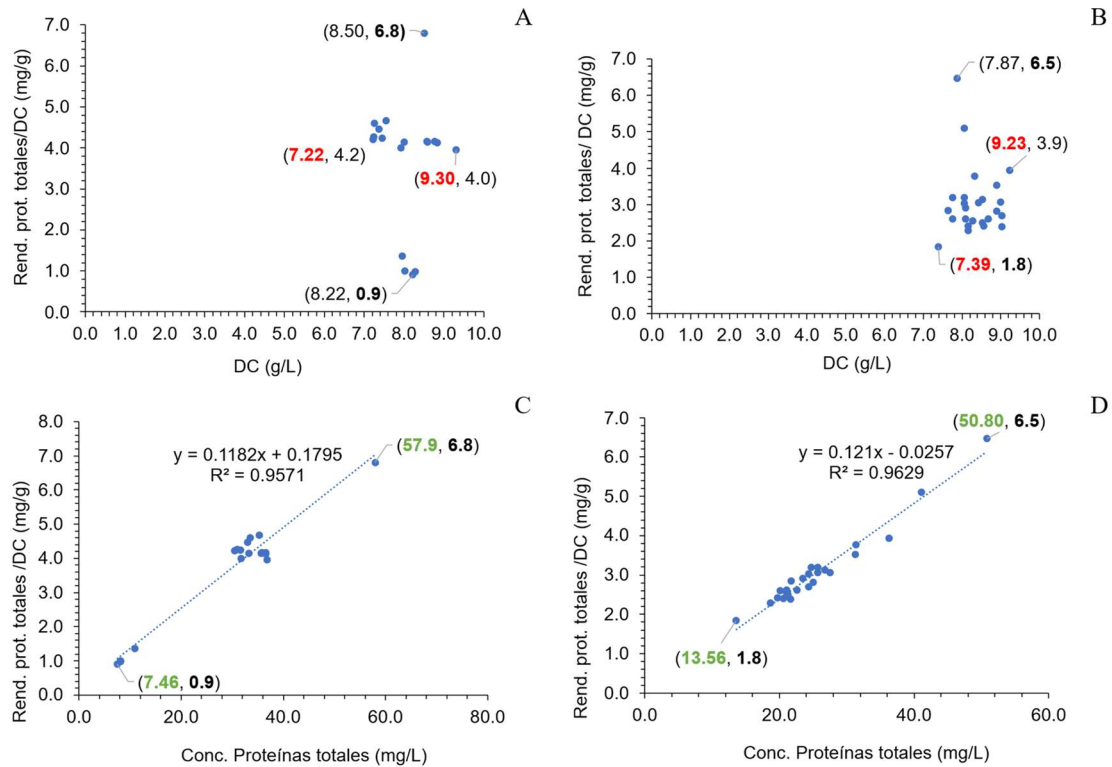


Figura 30. **A y B.** Gráficas de dispersión del rendimiento de proteína totales por unidad de biomasa vs DC. **A.** cepa $P_{GAP-T_{AOX1}}$ y **B.** cepa $P_{GAP-T_{GAP}}$. Se representa el par ordenado (x,y) de los valores extremos de la DC (x) resaltados en rojo y del rendimiento proteínas totales por unidad de biomasa (y) resaltado en negro. **C y D.** Gráficas de dispersión del rendimiento de proteínas totales por unidad de biomasa vs concentración de proteínas totales extracelulares de: **C.** cepa $P_{GAP-T_{AOX1}}$ y **D.** cepa $P_{GAP-T_{GAP}}$. Se representa el par ordenado (x,y) de los valores extremos de la concentración de proteínas totales extracelulares (x) resaltados en verde y del rendimiento proteínas totales por unidad de biomasa (y) resaltado en negro. Se representa además la línea de tendencia, la ecuación y el valor de R^2 del análisis de correlación lineal.

Según el análisis estadístico (Anexo 4 y 5), se encontró que la media poblacional de la densidad celular para la cepa $P_{GAP-T_{AOX1}}$ se pudo estimar entre 7.75 g/L y 8.37 g/L, con 95 % de confianza y un error relativo de 3.9 %; mientras que la desviación estándar estimada se situó en el rango de 0.469 g/L y 0.937 g/L. El valor bajo del error relativo de la media (< 5 %) y la demostración de que los valores de densidad celular provienen de una distribución normal sugirieron que la muestra de 18 clonas de la cepa $P_{GAP-T_{AOX1}}$ es adecuada para representar, en términos de la media de la densidad celular, a la población de más de 1000 transformantes obtenidos y seleccionados en medio RBD His^r.

Resultados similares se observaron en el análisis de las 25 clonas seleccionadas de la cepa $P_{GAP-T_{GAP}}$. El análisis estadístico (Anexos 6 y 7) evidenció que es posible estimar con 95 % de confianza, que la media poblacional de la densidad celular se encuentra dentro

del intervalo de confianza comprendido entre 8.14 g/L y 8.54 g/L, en tanto que la desviación estándar estimada está dentro del intervalo de confianza entre 0.382 g/L y 0.681 g/L. El intervalo de confianza para la media se estimó con un error relativo de 2.4 %. Este valor bajo del error relativo de la media ($< 5 \%$) y la demostración de que los valores de densidad celular provienen de una distribución normal sugirieron que la muestra de 25 clonas de la cepa $P_{GAP-T_{GAP}}$ es adecuada para representar, en términos de la media de la densidad celular, a la población de más de 1000 transformantes obtenidos y seleccionados en medio RBD His⁻ para esta cepa.

En cuanto a la concentración de proteína extracelular, los valores oscilaron entre 7.46 y 57.90 mg/L para $P_{GAP-T_{AOX1}}$ (Anexo 2) y entre 13.56 y 50.80 mg/L para $P_{GAP-T_{GAP}}$ (Anexo 3) (Figura 30 C y D, valores resaltados en verde). Los valores de rendimiento proteína/biomasa estuvieron en un rango de 0.9 a 6.8 mg/g y de 1.8 a 6.5 mg/g, respectivamente (Figura 30, valores resaltados en negro).

Del análisis de los resultados es destacable observar que a pesar de que la dispersión de los valores de densidad celular fue baja con valores de coeficiente de variación de 7.8 % (Anexo 4) para $P_{GAP-T_{AOX1}}$ y de 5.9 % para $P_{GAP-T_{GAP}}$ (Anexo 6). Los valores de concentración de proteína extracelular y de rendimiento proteína biomasa mostraron una marcada variabilidad. El coeficiente de variación de la concentración de proteína y del rendimiento proteína biomasa para la cepa $P_{GAP-T_{AOX1}}$ fue de 30.3 % y de 31.2 %, respectivamente y para la cepa $P_{GAP-T_{GAP}}$ fue de 43.9 % y de 42.7 %, respectivamente.

Se identificó que esta variabilidad del rendimiento proteína/biomasa entre las clonas de una misma cepa no está correlacionada con la variabilidad en los valores de densidad celular (Figura 30 A y B) pero si está directamente correlacionada (con R^2 de 96 %) con la variabilidad en la concentración de proteína extracelular (Figura 30 C y D).

7.6 Selección de clonas posibles unicopias

El primer criterio de selección de clonas posibles unicopias fue el rendimiento proteína extracelular/biomasa. Se observó gran variabilidad de este parámetro entre clonas para ambas cepas según se describió en el acápite anterior y se observa en la Figura 31.

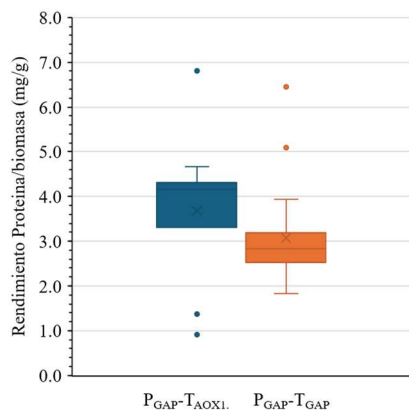


Figura 31. Gráfica de cajas y bigotes para el análisis de la variabilidad del rendimiento proteína/biomasa entre 18 clones de la cepa $P_{GAP}-T_{AOX1}$ y 25 clones de la cepa $P_{GAP}-T_{GAP}$.

Se observaron clones con valores atípicos de rendimiento proteína/biomasa. La clona de mayor rendimiento con valor de 6.8 mg/g en la cepa $P_{GAP}-T_{AOX1}$ es una clona previamente definida como multicopia (Herrera-Estala et al. 2020) de la cepa KM71GAHFTEII. Entonces por inferencia se definieron las clones con valores de rendimientos altos y atípicos de 6.5 y 5.1 mg/g de la cepa $P_{GAP}-T_{GAP}$ como posibles clones multicopias. Las clones con menor rendimiento proteína/biomasa y valores atípicos alrededor de 1 mg/g en la cepa $P_{GAP}-T_{AOX1}$ son clones revertantes. Estas clones, a pesar de crecer en medio selectivo RBD His⁻, cuando se evaluaron por PCR de integración se observó una banda muy tenue o ausencia de producto amplificado en el PCR de integración. Probablemente esto se deba a que no incorporaron el casete de expresión en su genoma o lo perdieron por generación, a medida que se realizan los cultivos celulares.

Es conocido que entre el 90 y el 99 % de las clones transformadas resultan ser unicopias (Higging et al. 1998), por lo que se decidió escoger dentro de la región intercuartil una clona de la cepa $P_{GAP}-T_{AOX1}$ (identificada como CN11) y dos clones de la cepa $P_{GAP}-T_{GAP}$ como posibles unicopias (identificadas como CN4 y CN2). Adicionalmente, se escogió la clona multicopia $P_{GAP}-T_{AOX1}$ (C⁺) y una de las clones posible multicopias de $P_{GAP}-T_{GAP}$ (CN8). Todas estas clones se analizaron por dosis génica.

7.7 Determinación de la dosis génica

Las clonas seleccionadas como posibles unicopias y las clonas de mayor rendimiento en cada cepa se evaluaron por dosis génica. Durante el desarrollo del qPCR, en la obtención de la recta de eficiencia los valores del ciclo umbral (Ct) de la amplificación del gen heterólogo *FTEII* y del gen *GAPDH* endógeno, mostraron linealidad en el intervalo de 0.2 a 25 ng de ADNg. Ambas rectas se obtuvieron con valor de regresión lineal del 99 % indicando que con el valor de Ct es posible predecir con precisión la cantidad de ADNg. Además, las pendientes de las rectas se obtuvieron con valores de - 3.4 y - 3.5 (Figura 32), y los valores de eficiencia del qPCR fueron de 94.9 % y 92.0 % para los genes *FTEII* y *GAPDH*, respectivamente por lo que se cumplió con los parámetros de calidad de ejecución del qPCR (R^2 cercano a 100 %, pendiente de -3.3 ± 10 % y eficiencia de 100 ± 10 %).

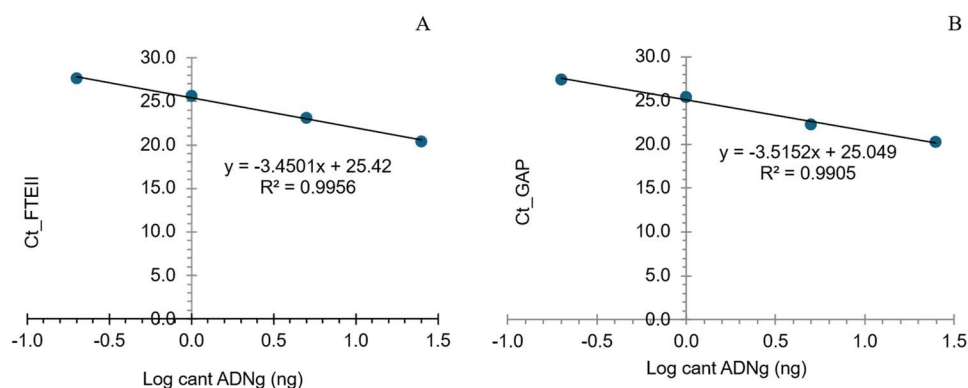


Figura 32. Rectas de eficiencias del qPCR. **A.** para el gen heterólogo *FTEII* y **B.** para el gen *GAPDH* endógeno.

En la evaluación de las clonas seleccionadas se confirmó que las clonas de mayores rendimiento proteína/biomasa en ambas cepas son multicopias. La clona de la cepa $P_{GAP-T_{AOX1}}$ resultó con tres copias del gen heterólogo *FTEII* y la cepa $P_{GAP-T_{GAP}}$ con dos copias de dicho gen. Las tres clonas seleccionadas como posibles unicopias (una clona de la cepa $P_{GAP-T_{AOX1}}$ y dos clonas de la cepa $P_{GAP-T_{GAP}}$) resultaron tener una única copia del gen heterólogo *FTEII* (Tabla 25). La clona unicopia de $P_{GAP-T_{AOX1}}$ (denotada como CN11) y una de las dos clonas unicopias de $P_{GAP-T_{GAP}}$ (denotada como CN2) se escogieron para los análisis comparativos de expresión de genes.

Tabla 25. Dosis génica de las clonas multicopias y unicopias de las cepas P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}}

Cepa	Clona	Gen <i>FTEII</i>		Gen <i>GAPDH</i>		Dosis génica
		Ct	ADNg (ng)	Ct	ADNg (ng)	
P _{GAP-T_{AOX1}}	C+	23.11	4.67	24.21	1.73	3
	CN11*	24.52	1.82	24.05	1.93	1
P _{GAP-T_{GAP}}	CN8**	23.78	3.00	24.39	1.54	2
	CN4*	24.30	2.12	23.50	2.77	1
	CN2*	24.93	1.39	24.22	1.73	1

Ct. Ciclo umbral. C+. Clona multicopia de la cepa KM71GAHFTEII (Herrera-Estala et al. 2020). *. Clonas unicopias. **. Clona multicopia.

7.8 Funcionalidad de los elementos reguladores P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}}

7.8.1 Detección de transcritos del gen heterólogo *FTEII* por RT-PCR

El ARN extraído de las muestras tomadas a las 10 h del cultivo de P_{GAP-T_{GAP}} y P_{GAP-T_{AOX1}} se obtuvo con una concentración de 124 ng/μL y 113 ng/μL, respectivamente y una relación de DO_{260nm/280nm} de 2.160 y de 2.014, respectivamente. Ambos con buena calidad del ARN por tener un valor aproximado de 2.0 en la relación de DO_{260nm/280nm}. Los resultados de la electroforesis de agarosa corroboran la calidad del ARN obtenido (Figura 33) al definirse las dos bandas de ARN ribosomal 28S y 18S y no observarse degradación de ARN que pudiera haberse generado durante el proceso de extracción.

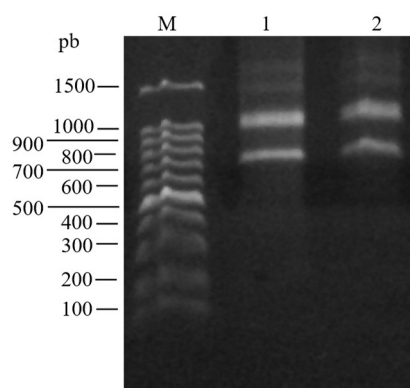


Figura 33. Electroforesis en gel de agarosa para verificar la calidad del ARN extraído de las cepas P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}}.

Carril **M**. Patrón de tamaño molecular de 100 pb ADN Ladder (Promega, EUA). Carril **1**. ARN extraído de la muestra tomada a las 10 h de cultivo de la cepa P_{GAP-T_{GAP}}. Carril **2**. ARN extraído de la muestra tomada a las 10 h de cultivo de la cepa P_{GAP-T_{AOX1}}.

El análisis de RT-PCR para las cepas basadas en P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}} mostró la banda esperada de 536 pb correspondiente a un fragmento del ADNc obtenido por

retrotranscripción del gen heterólogo de *FTEII* (Figura 34), lo que demuestra la funcionalidad de la combinación de elementos reguladores promotor-terminador en ambas cepas.

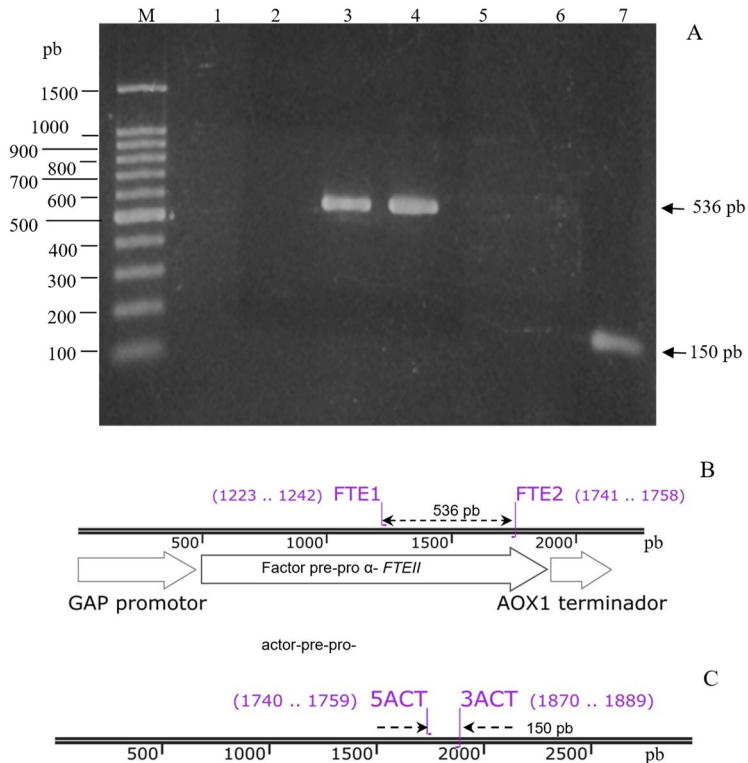


Figura 34. **A.** Electroforesis en gel de agarosa al 2 % para determinación de funcionalidad.

Carril **M.** Marcador de tamaño molecular 100 pb ADN Ladder (Promega, EUA). Carriles **1 al 4.** Ensayos con los oligonucleótidos FTE1 y FTE2. Carril **2.** Control negativo del paso de retrotranscripción. Carril **3.** Control negativo de la PCR. Carril **4.** Producto RT-PCR de la cepa basada en $P_{GAP-T_{AOX1}}$. Carril **5.** producto RT-PCR de la cepa basada en $P_{GAP-T_{GAP}}$. Carriles **5 a 7.** Ensayos con los oligonucleótidos dirigidos al gen endógeno de actina (5ACT y 3ACT). Carril **5.** Control negativo del paso de retrotranscripción. Carril **6.** control negativo PCR. Carril **7.** Producto amplificado por RT-PCR (control positivo). **B.** Representación esquemática de los sitios de hibridación del par de oligonucleótidos FTE1/FTE2 en el casete de expresión integrado en el genoma de las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$. **C.** Representación esquemática de los sitios de hibridación del par de oligonucleótidos 5ACT/3ACT dirigidos al gen endógeno de actina en *K. phaffii*. La secuencia de actina se obtuvo de la base de datos del GenBank “*Pichia pastoris* actin, complete cds accession AF216956.1”.

7.8.2 Detección de producto proteico del gen heterólogo *FTEII*

La SDS-PAGE del sobrenadante libre de células de los cultivos de las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$ reveló un barrido por encima de 39 kDa, que es típico de la proteína recombinante fitasa FTEII (Figura 35). Tras la N-desglicosilación por Endo Hf, este barrido se desplazó a una banda de 39 kDa (Figura 35, carriles 2 y 4) que corresponde a la masa molecular predicha de la fitasa FTEII a partir de su secuencia de aminoácidos.

Este resultado corrobora la secreción de fitasa FTEII como una proteína altamente N-glicosilada, como se observó previamente (Viader-Salvadó et al., 2010).

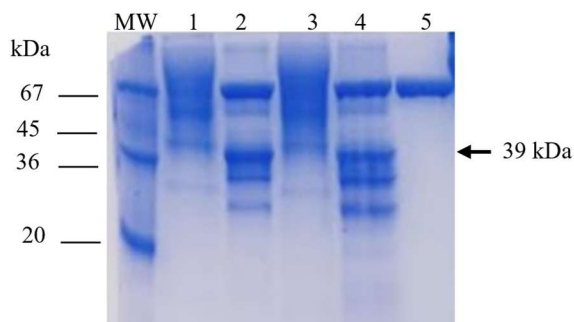


Figura 35. Electroforesis en gel de poliacrilamida al 12% del sobrenadante libre de células del cultivo de las cepas basadas en $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$. Carril MW. Marcador de peso molecular. Carriles 1 y 2. Proteínas sin y con tratamiento con Endo Hf glicosidasa de la cepa basada en $P_{GAP-T_{AOX1}}$. Carriles 3 y 4. Proteínas sin y con tratamiento con Endo Hf glicosidasa de las cepas $P_{GAP-T_{GAP}}$. Carril 5. Proteína Endo Hf glicosidasa. La flecha indica a la proteína recombinante fitasa FTEII N-desglicosilada.

Previo a la determinación de actividad enzimática de fitasa, el sobrenadante de los cultivos de las cepas basadas en $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$ se desalaron por cambio de tampón, en cromatografía de gel filtración (Figura 36).

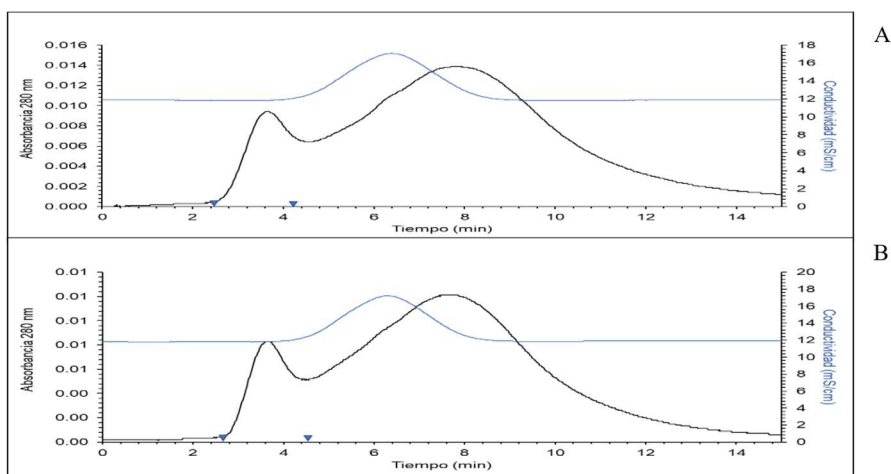


Figura 36. Perfil cromatográfico de la cromatografía de exclusión molecular en columna PD-10 Sephadex G25 del sobrenadante de cultivo de las cepas. A. $P_{GAP-T_{GAP}}$ y B. $P_{GAP-T_{AOX1}}$. Las marcas en el eje horizontal indican el tiempo de colección del pico de interés.

Los sobrenadantes de cultivo desalados de las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$ exhibieron actividades de fitasa de 0.34 y 0.25 U/mL y rendimiento producto/biomasa de 87.0 y 63.2 U/g, respectivamente (Tabla 26).

Tabla 26. Detección de proteínas totales, rendimiento producto por unidad de biomasa y actividad enzimática de FTEII en los sobrenadantes de cultivo de las cepas P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}.

Cepas	Proteínas (mg/L)	Y _{p/x} * (U/g)	Actividad volumétrica de Fitasa (U/mL)
P _{GAP} - T _{AOX1}	29.5 ± 0.4	87.0 ± 0.1	0.34 ± 0.013
P _{GAP} -T _{GAP}	23.0 ± 0.2	63.2 ± 0.1	0.25 ± 0.006

*: Rendimiento producto-biomasa.

Los resultados combinados de la detección de transcritos de *FTEII* y la proteína FTEII a través de RT-PCR, SDS-PAGE y la actividad de la enzima fitasa corroboraron la funcionalidad adecuada de las combinaciones de los elementos reguladores (es decir, P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}).

7.9 Efecto del terminador transcripcional en el gen heterólogo sobre los niveles de transcripción del gen heterólogo *FTEII* y el gen endógeno *GAPDH*, y en el Y_{p/x} extracelular

7.9.1 Cinética de crecimiento

La cinética de crecimiento de las dos cepas (P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}), utilizando glucosa o glicerol como fuentes de carbono (Figura 37A), mostró el característico crecimiento poblacional sigmoide (logístico). Las concentraciones de biomasa mostraron un crecimiento exponencial de 3 a 12 h de cultivo, con valores de μ que oscilaron entre 0.213 ± 0.023 y $0.249 \pm 0.013 \text{ h}^{-1}$. Se obtuvo una concentración de biomasa máxima (X_{max}) más alta en los cultivos con glucosa en comparación con los que usaron glicerol (1.6 y 1.8 veces mayor para las cepas basadas en P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}, respectivamente). Esto resultaba esperado, ya que todos los cultivos fueron equimolares en la fuente de carbono, y una molécula de glucosa aporta seis carbonos, mientras que una molécula de glicerol aporta sólo tres. Además, la glucosa es una fuente de carbono que se incorpora directamente en la glucólisis para generar energía y biomasa, mientras que el glicerol es una fuente de carbono que entra en la gluconeogénesis, y en *K. phaffii* se ha reportado un flujo de glicerol a 1,3-bisfosfoglicerato del 64 % en promedio (Tomàs-Gamisans et al., 2019), lo que sugiere que no todo el glicerol se aprovecha para generar biomasa y energía.

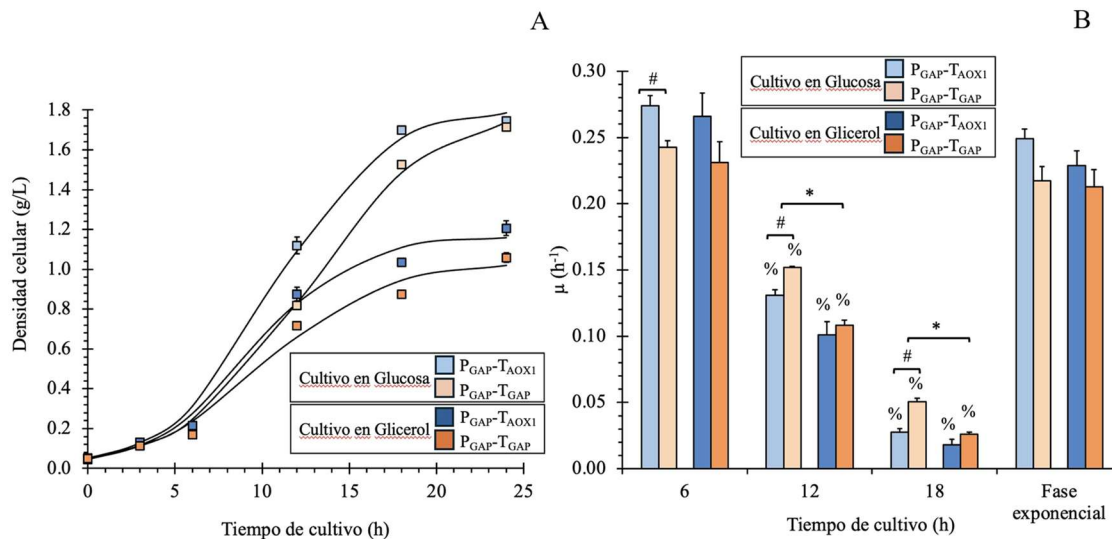


Figura 37. **A.** Cinética de crecimiento y **B.** Velocidad específica de crecimiento (μ) en diferentes tiempos de cultivo para P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP} en glucosa (BMGlc) y glicerol (BMGly). Los datos se presentan como la media \pm error estándar de tres cinéticas realizadas de forma independiente. La diferencia significativa ($p < 0.05$) se representa con (%) entre dos tiempos consecutivos de cultivo, (*) para las células de la misma cepa crecidas en glucosa o glicerol y (#) para las dos cepas crecidas en la misma fuente de carbono.

Los valores puntuales de μ mostraron disminuciones durante el tiempo de cultivo de 2.1 y 4.4 veces menores en promedio de 6 a 12 y de 12 a 18 h de cultivo, respectivamente (Figura 37B). Por lo tanto, los valores puntuales de μ se pueden clasificar como μ alto, medio y bajo. Aunque no se observaron diferencias significativas en los valores de μ durante la fase de crecimiento exponencial, ya sea al comparar la misma cepa cultivada en diferentes fuentes de carbono o al evaluar las dos cepas cultivadas bajo la misma fuente de carbono, se observaron diferencias significativas en los valores de μ puntuales entre las dos cepas en los cultivos de glucosa y entre los cultivos de glucosa y glicerol para la cepa basada en P_{GAP}-T_{GAP} a las 12 y 18 h de cultivo (Figura 37B).

7.9.2 Niveles de transcripción de *FTEII* y *GAPDH* y Y p/x extracelular a lo largo del tiempo de cultivo

Las Figuras 38A y 38B muestran los niveles de transcripción de *FTEII* y *GAPDH* de las dos cepas en los tres tiempos de muestreo para células cultivadas con glucosa (Figura 38A) o glicerol (Figura 38B).

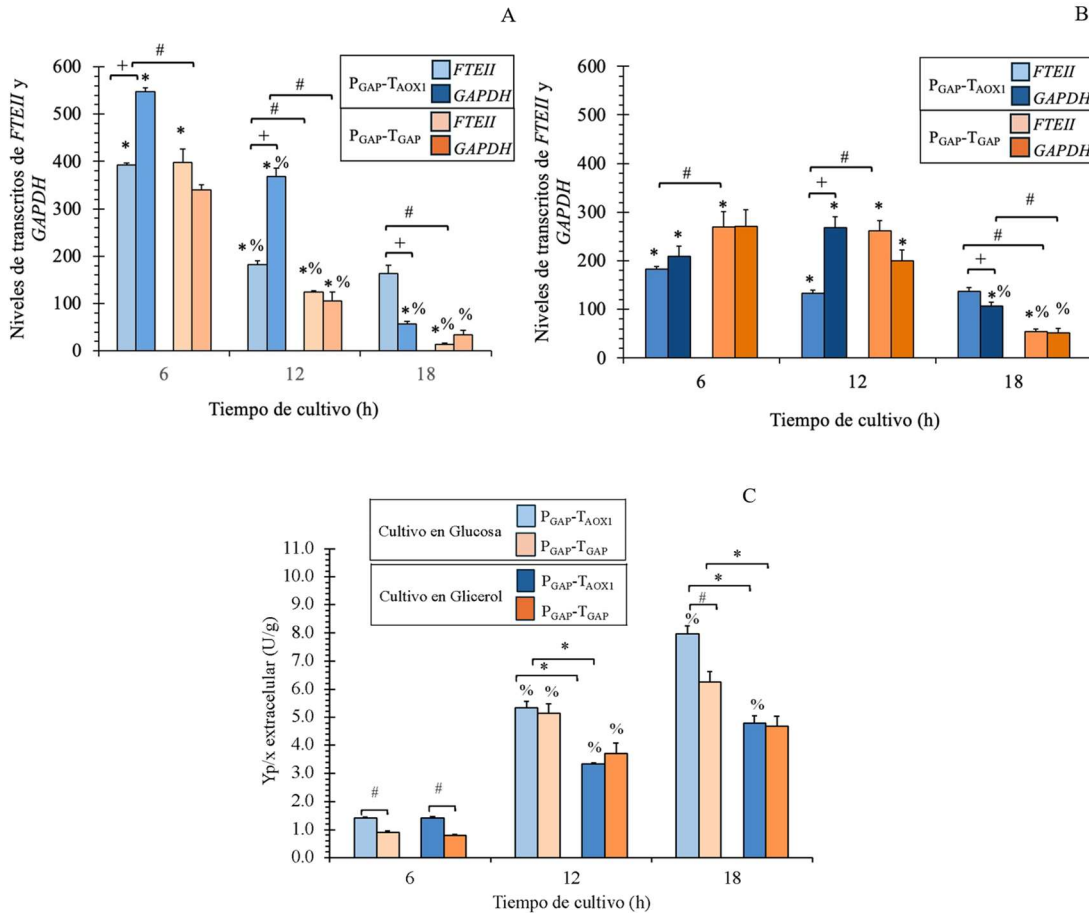


Figura 38. Niveles de transcritos de *FTEII* y *GAPDH* de las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$. **A.** glucosa y **B.** glicerol. **C.** Rendimiento $Y_{p/x}$ extracelular de ambas cepas en glucosa (BMGlc) y glicerol (BMGly). Los datos se presentan como la media \pm error estándar de tres cinéticas realizadas de forma independiente. Las diferencias significativas ($p < 0.05$) se representa con (%) entre dos tiempos consecutivos de cultivo, (+) niveles de transcritos de *FTEII* y *GAPDH* para la misma cepa en la misma fuente de carbono, (*) para las células de la misma cepa crecidas en glucosa o glicerol, y (#) para las dos cepas crecidas en la misma fuente de carbono.

En la cepa $P_{GAP-T_{AOX1}}$, los niveles de transcripción de *FTEII* alcanzaron su punto máximo a las 6 h de cultivo, coincidiendo con la μ más alta (0.253 h^{-1}), independientemente de si las células se cultivaron en glucosa o glicerol (Figura 37B). De 12 a 18 h de cultivo los niveles disminuyeron y se mantuvieron estables. Por su parte, los niveles de transcripción de *GAPDH* mostraron una tendencia similar a los niveles de transcripción de *FTEII* hasta las 12 h en los cultivos de glucosa (Figura 38A), aunque no se observaron diferencias significativas en los niveles de transcripción de *GAPDH* de 6 a 12 h en los cultivos de glicerol (Figura 38B). A las 18 h de cultivo, cuando la glucosa o el glicerol estaban casi agotados, el gen *GAPDH* mostró una disminución significativa, mostrando niveles de transcripción más bajos que los del gen *FTEII*.

Aunque la tendencia de la actividad transcripcional de los genes *FTEII* y *GAPDH* en las células $P_{GAP-T_{GAP}}$ cultivadas con glucosa fueron similares a las de las células $P_{GAP-T_{AOX1}}$ hasta 12 h de cultivo, los niveles de transcripción permanecieron sin cambios en las células $P_{GAP-T_{GAP}}$ cultivadas con glicerol de 6 a 12 h de cultivo. Además, en la cepa $P_{GAP-T_{GAP}}$, ambos genes mostraron una disminución apreciable en los niveles de transcripción a las 18 h de cultivo con cualquiera de las fuentes de carbono analizadas.

A lo largo de la progresión del cultivo (es decir, disminuyendo la μ), los valores de $Y_{p/x}$ extracelular aumentaron consistentemente para la cepa $P_{GAP-T_{AOX1}}$ en las dos fuentes de carbono evaluadas (Figura 38C). De manera similar, los valores de $Y_{p/x}$ extracelular para los cultivos celulares $P_{GAP-T_{GAP}}$ aumentaron de 6 a 12 h en las dos fuentes de carbono; sin embargo, no se observaron diferencias significativas en el $Y_{p/x}$ extracelular de 12 a 18 h tanto en los cultivos de glucosa como en los de glicerol (Figura 38C). En general, los niveles de transcripción de *FTEII* y el $Y_{p/x}$ extracelular se relacionaron de manera inversamente proporcional.

7.9.3 Comparación de los niveles de transcripción de *FTEII* y *GAPDH* de la misma cepa cultivada en la misma fuente de carbono

Aunque tanto el gen *FTEII* como el *GAPDH* estuvieron regulados por P_{GAP} en las dos cepas, los niveles de transcripción del gen endógeno *GAPDH* para la cepa $P_{GAP-T_{AOX1}}$ fueron mayores que los del gen *FTEII* a las 6 h y 12 h de cultivo en glucosa (1.7 veces mayor en promedio) (Figura 38A) y a las 12 h de cultivo en glicerol (2.0 veces mayor) (Figura 38B). Sin embargo, *FTEII* exhibió en promedio niveles de transcripción 2.1 veces más altos, en comparación con *GAPDH* a las 18 h de cultivo en las dos fuentes de carbono (Figura 38A y 38B). No se observaron diferencias significativas entre los niveles de transcripción de *FTEII* y *GAPDH* para la cepa $P_{GAP-T_{AOX1}}$ a las 6 h de cultivo en glicerol (Figura 38B). No se observaron diferencias significativas en los niveles de transcripción en la cepa $P_{GAP-T_{GAP}}$ entre los genes heterólogos *FTEII* y *GAPDH* endógenos en los tres tiempos de muestreo de cultivo en ninguna de las fuentes de carbono.

7.9.4 Efecto de la fuente de carbono en los niveles de transcripción de *FTEII* y *GAPDH*

Los niveles de transcripción de *FTEII* y *GAPDH* para la cepa P_{GAP-T_{AOX1}} fueron más altos en las células cultivadas en glucosa (Figura 38A) en comparación con las células cultivadas en glicerol (Figura 38B) a las 6 h (2.2 y 2.6 veces, respectivamente) y a las 12 h de cultivo (1.4 veces para los dos genes). Sin embargo, no se observaron diferencias significativas en los niveles de transcripción de *FTEII* a las 18 h de cultivo entre las células cultivadas en glucosa o glicerol, mientras que los niveles de transcripción de *GAPDH* fueron 1.9 veces mayores en las células cultivadas con glicerol en comparación con las células cultivadas con glucosa.

Aunque los niveles de transcripción de *FTEII* y *GAPDH* para las células P_{GAP-T_{AOX1}} fueron más altos en los cultivos de glucosa (Figura 38A) que en los de glicerol (Figura 38B) durante la fase de crecimiento exponencial (μ alta y media), la mayor parte del tiempo los niveles de transcripción de *FTEII* y *GAPDH* de las células P_{GAP-T_{GAP}} fueron más altos en los cultivos con glicerol (Figura 38A), en lugar de en los cultivos con glucosa (Figura 38B). Específicamente, las células P_{GAP-T_{GAP}} exhibieron niveles más altos de transcripción de *FTEII* en glicerol que en los cultivos con glucosa a las 12 y 18 h de cultivo (2.1 y 4.1 veces, respectivamente), mientras que a las 6 h de cultivo, estas células mostraron niveles más altos de transcripción de *FTEII* en glucosa (1.5 veces) que en cultivos con glicerol. Los niveles de transcripción de *GAPDH* fueron 1.9 veces más altos en las células cultivadas con glicerol que en las cultivadas con glucosa a las 12 h de cultivo, aunque no se observaron diferencias significativas a las 6 h y 18 h de cultivo.

7.9.5 Comparación de los niveles de transcripción entre cepas cultivadas en la misma fuente de carbono

Los niveles de transcripción de *FTEII* entre las dos cepas (P_{GAP-T_{AOX1}} y P_{GAP-T_{GAP}}) a las 6 h de cultivo en glucosa no mostraron diferencias significativas (Figura 38A), aunque a las 12 h y a las 18 h de cultivo, los niveles de transcritos de *FTEII* fueron más altos en células de P_{GAP-T_{AOX1}} en comparación con células de P_{GAP-T_{GAP}}, con valores 1.5 veces y 12.3 veces mayores, respectivamente. Sin embargo, en los cultivos con glicerol a las 6 h y a las 12 h de cultivo se observaron mayores niveles de transcritos de *FTEII* en la cepa

$P_{GAP-T_{GAP}}$ con valores de 1.5 veces y 2.0 veces mayores, respectivamente, comparado con la cepa $P_{GAP-T_{AOX1}}$ también crecida en glicerol.

Los niveles de transcritos de *GAPDH* en las células $P_{GAP-T_{AOX1}}$ a las 6, 12 h de cultivo en glucosa y las 18 h en glicerol fueron 1.6, 3.5 y 2.0 veces más altos, respectivamente, comparado con las células $P_{GAP-T_{GAP}}$ (Figura 38A y 38B). A las 6 y 12 h de cultivo en glicerol no se observaron diferencias significativas entre cepas. Tampoco se observaron diferencias significativas en los niveles de transcripción de *GAPDH* entre las dos cepas cultivadas en glucosa a las 18 h de cultivo.

7.9.6 Comparación de la $Y_{p/x}$ extracelular entre cepas y el efecto de la fuente de carbono

Los cultivos de células $P_{GAP-T_{AOX1}}$ en ambas fuentes de carbono mostraron un valor promedio de $Y_{p/x}$ extracelular que fue 1.3 veces mayor que el de los cultivos de células $P_{GAP-T_{GAP}}$ (Figura 38C). Además, no se observaron diferencias significativas en el $Y_{p/x}$ extracelular entre los cultivos de glucosa y glicerol a las 6 h de cultivo para ninguna de las dos cepas. En promedio, la $Y_{p/x}$ extracelular fue 1.6 veces mayor en glucosa en comparación con los cultivos de glicerol a las 12 y 18 h de cultivo para la cepa $P_{GAP-T_{AOX1}}$ y 1.4 veces mayor para la cepa $P_{GAP-T_{GAP}}$.

7.9.7 Comparación de los cocientes de niveles de transcripción de *FTEII* y cocientes de $Y_{p/x}$ extracelulares

Los cocientes de los niveles de transcripción de *FTEII* y los cocientes de $Y_{p/x}$ extracelulares entre cepas fueron diferentes, y también diferentes en cuanto al tiempo de cultivo (Figura 39). Los cocientes de $Y_{p/x}$ extracelulares entre cepas fueron mayores que los cocientes entre los niveles de transcripción de *FTEII* entre cepas a las 6 h de cultivo en las fuentes de carbono evaluadas. Por el contrario, a la μ media (0.123 h^{-1}), el cociente de $Y_{p/x}$ extracelular entre cepas fue mayor que el cociente de los niveles de transcripción de *FTEII* solo en los cultivos de glicerol, mientras que a las 18 h de cultivo, el cociente de los niveles de transcripción de *FTEII* entre cepas fue mayor que el cociente de $Y_{p/x}$ extracelular entre cepas en ambas fuentes de carbono. Estos resultados sugieren que la actividad postranscripcional se vio fuertemente afectada por la presencia de T_{AOX1} dentro

de la estructura del gen heterólogo y dependió de la fuente de carbono y la etapa de crecimiento celular.

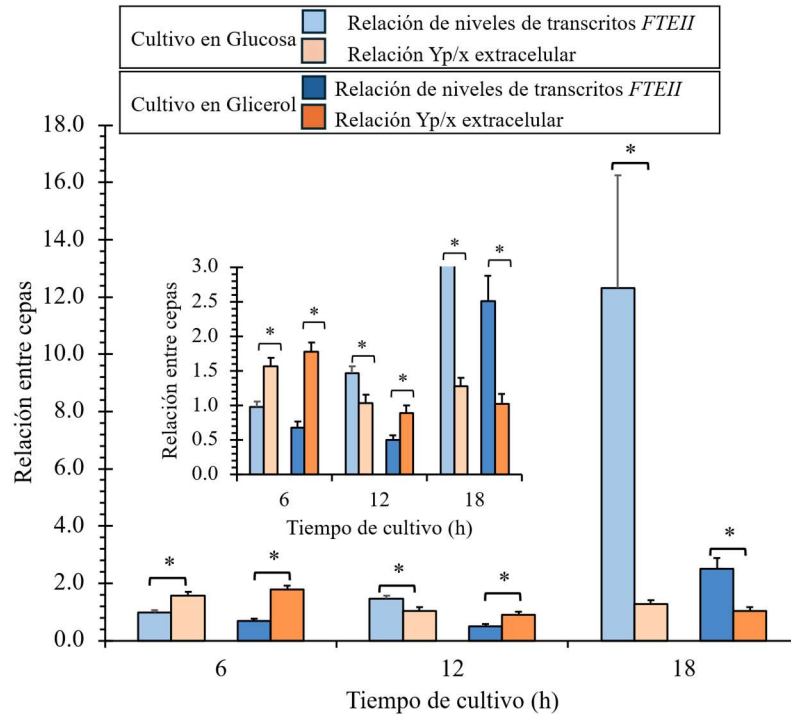


Figura 39. Cocientes de niveles de transcritos de *FTEII* y cocientes de $Y_{p/x}$ extracelular de la cepa $P_{GAP-T_{AOX1}}$ vs. $P_{GAP-T_{GAP}}$ cultivadas en glucosa (BMGlc) y glicerol (BMGly).

Los datos se presentan como la media \pm error estándar de tres cultivos independiente. La diferencia significativa ($p < 0.05$) se representa con (*) para los cocientes de niveles de transcritos de *FTEII* y cocientes de $Y_{p/x}$ entre cepas crecidas en la misma fuente de carbono.

Los cocientes de los niveles de transcripción de *FTEII* y los cocientes de $Y_{p/x}$ extracelulares de las células cultivadas en glucosa entre los valores de las células cultivadas en glicerol fueron diferentes, y también diferentes en términos del tiempo de cultivo (Figura 40). A las 6 h de cultivo, el cociente de niveles de transcripción de *FTEII* de la cepa $P_{GAP-T_{AOX1}}$ respecto a fuente de carbono fue mayor que el correspondiente cociente $Y_{p/x}$ extracelular, mientras que no se observaron diferencias significativas entre los dos cocientes para los cultivos de células $P_{GAP-T_{GAP}}$. Además, no se observaron diferencias significativas entre los dos cocientes a las 12 h de cultivo para la cepa $P_{GAP-T_{AOX1}}$. Por el contrario, los cocientes de $Y_{p/x}$ extracelulares respecto a la fuente de carbono fueron mayores a las 12 h de cultivo para la cepa en $P_{GAP-T_{GAP}}$ y a las 18 h de cultivo para las dos cepas que los correspondientes cocientes de los niveles de transcripción de *FTEII*.

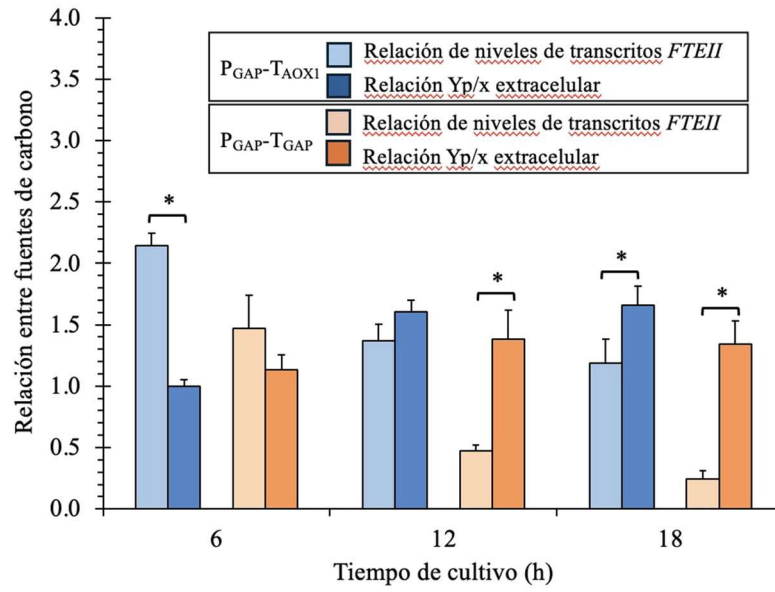


Figura 40. Cocientes de niveles de transcritos de *FTEII* y cocientes de $Y_{p/x}$ extracelular respecto a las fuentes de carbono glucosa vs glicerol de las cepas P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP}. Los datos se presentan como la media \pm error estándar de tres cultivos independiente. La diferencia significativa ($p < 0.05$) se representa con (*) para los cocientes de niveles de transcritos de *FTEII* y cocientes de $Y_{p/x}$ entre la misma cepa crecida en diferente fuente de carbono.

8 DISCUSIÓN

En los últimos años se han realizado varios reportes referentes a la funcionalidad de los terminadores transcripcionales en la regulación de la expresión de genes heterólogos. Curran et al. (Curran et al. 2013) reportaron por primera vez el uso de los terminadores transcripcionales para incrementar la expresión de genes en *S. cerevisiae*. Estudiaron un total de 30 terminadores con aplicación en ingeniería metabólica. Determinaron que los cambios en la vida media del ARNm son la principal causa de los niveles de expresión de la transcripción. Demostraron que la diferencia en el nivel de transcripción puede ser más de 6.5 veces incluso para promotores fuertes. Que la influencia de la selección del terminador aumentó cuando se combinó con un promotor de baja expresión, con una diferencia máxima en la expresión de proteínas de 11 veces entre un terminador que mejora la expresión y el terminador del plásmido parental y una diferencia de más de 35 veces cuando se comparó con la línea base (sin terminador). Evidenciaron la utilidad de la selección de terminadores para la ingeniería metabólica. Destacaron la utilidad de los elementos de terminación para controlar las vías metabólicas de la misma forma que los promotores se utilizan tradicionalmente en los sistemas de expresión en levaduras. Concluyeron que los terminadores serían una parte importante de la expresión génica heteróloga y en la ingeniería para levadura en el futuro.

Vogl et al. (Vogl et al. 2016) en un estudio de la ruta metabólica de utilización del metanol (MUT) en *K. phaffii* para la identificación de promotores útiles en la expresión génica, utilizó promotores fuertemente regulados junto con 20 nuevos terminadores como una herramienta útil para la expresión de una vía biosintética heteróloga. En *K. phaffii*, el terminador del gen *AOX1* se ha utilizado exclusivamente para terminar la transcripción de genes heterólogos. Sin embargo, en su estudio evaluaron quince terminadores de genes de la vía MUT fuertemente expresados, y terminadores de los genes constitutivos (*GAPDH*, *GCW14*, *TEF1*). Por las propiedades conservadas de los terminadores de levadura (ricos en AT), también evaluaron cinco terminadores heterólogos de *S. cerevisiae* para evitar similitudes de secuencia con el genoma del hospedero. Clonaron los terminadores manteniendo la secuencia natural hasta el codón de parada, aguas abajo de un gen reportero bajo el control de P_{AOX1} . Los terminadores de *K. phaffii* mostraron capacidades

de terminación de la transcripción comparables a las del terminador *AOXI* (57-89 % de fluorescencia de la proteína reportera). Las construcciones con terminadores de *S. cerevisiae* dieron una fluorescencia de la proteína reportera similar a la de los terminadores de *K. phaffii* (62-78 %). Reportaron que la secuencia del terminador *AOXI* en los vectores de expresión contienen algunas mutaciones (lo denominaron terminador *AOXI**) en comparación con la secuenciación del genoma. Atribuyeron las diferencias a la posibilidad de haber surgido inicialmente de una secuenciación o amplificación por PCR errónea de esta región altamente rica en AT. No se observaron diferencias en las mediciones de fluorescencia entre el terminador *AOXI* con y sin mutaciones. Sugirieron que las mutaciones no tuvieron efecto sobre la fuerza de terminación. También evaluaron el efecto de un sitio de restricción *NotI* que se usa normalmente para insertar genes en vectores que contienen el terminador *AOXI** y reportaron que asombrosamente, el terminador con sitio *NotI* mostró un 37 % más de fluorescencia del gen reportero, sugiriendo un efecto beneficioso que pudo ser causada por una estructura secundaria, que se forma por la secuencia palindrómica larga y rica en GC de *NotI*.

Ramakrishnan et al. (Ramakrishnan et al. 2020) reportaron el control transcripcional de la expresión de genes heterólogos en *K. phaffii* mediante la manipulación de los terminadores utilizando P_{AOXI} y P_{GAP} . Mencionó que los reportes relacionados con el incremento de la expresión de genes se han enfocado en la búsqueda de promotores y en el número de copias del gen; y que existen pocos estudios recientes sobre el efecto de los terminadores relacionados a este tema. Mencionaron que los terminadores son componentes esenciales en los casetes de expresión, que influyen en el procesamiento del extremo 3' del ARNm, la estabilidad del ARNm y la eficiencia de traducción, y que además pueden modular la producción de proteínas. Enfatizó que en *K. phaffii* se ha prestado poca atención a la selección del terminador para aumentar la expresión de genes heterólogos. Para explorar la aplicación potencial de las regiones de terminación transcripcional para aumentar la producción secretora de un gen reportero, evaluaron la capacidad de tres clases diferentes de regiones terminadoras: (1) regiones terminadoras de genes de la vía de oxidación del metanol de *K. phaffii*; (2) regiones terminadoras de genes bien expresados de *K. phaffii*; y (3) terminadores de otros genes de levadura como *S. cerevisiae*. Observaron que los terminador de la dihidroxiacetona sintasa, un gen de alta

expresión en la vía de utilización del metanol, mostró niveles de expresión de la proteína heteróloga similar al terminador *AOXI* bajo el control del P_{AOXI} y tres veces mayor en la expresión constitutiva de la proteína heteróloga bajo el control del P_{GAP} . La abundancia de la transcripción del gen heterólogo se encontró que se correlacionó con la expresión de proteínas. Además, la determinación de la vida media del ARNm mostró una correlación directa con una mayor tasa de transcripción.

Todos estos estudios han evidenciado la importancia y la actual carencia de conocimientos en la temática relacionada con la influencia de los terminadores transcripcionales en los niveles de transcritos, la expresión de un gen heterólogo, y la producción de proteína heteróloga intra o extracelular en el sistema de expresión de *K. phaffii*. En relación a esto nuestro trabajo se enfocó en estudiar el efecto del terminador *GAPDH*, comparado con el usual terminador del gen *AOXI* en la expresión del gen heterólogo con promotor P_{GAP} con el supuesto de que pudiera existir una alteración en la producción de la proteína FTEII al medio extracelular. Los resultados generaron puntos de discusión importante relacionados con la correlación entre los elementos reguladores transcripcionales ($P_{GAP-T_{AOXI}}$ y $P_{GAP-T_{GAP}}$), los niveles de transcritos, la producción extracelular de FTEII y la etapa del cultivo celular en glucosa o glicerol.

En primera instancia para definir la secuencia del terminador del gen *GAPDH* se ejecutó un análisis de mapeo de RNA para detectar una zona inter-CDS entre el gen *GAPDH* y su gen contiguo río abajo denominado *NAB6*, con la finalidad de encontrar la zona no transcribible (sin presencia de RNA) del gen *GAPDH*. El análisis se llevó a cabo con transcriptomas de tres cepas diferentes. Sin embargo, en todos los casos se observó la presencia de RNA en toda la zona inter-CDS por lo que no fue posible definir la región no transcribible. Esto probablemente fue debido a la presencia en esta zona de transcritos crípticos inestables (CUTs, por sus siglas en inglés) y/o transcritos no caracterizados estables (SUTs, por sus siglas en inglés) (Marquardt et al. 2011). Los CUTs son un subconjunto de ARN no codificantes producidos a partir de regiones intergénicas e intragénicas y se han descrito como el tipo principal de transcritos de ARN polimerasa II en *S. cerevisiae* (Wyers et al. 2005). Los SUT son similares a los CUT, aunque son menos susceptibles a la degradación que los CUT. Estos últimos son propios de genes

involucrados en las rutas metabólicas del catabolismo de la glucosa (Neil et al. 2009), como el gen *GAPDH* que codifica para la gliceraldehido-3-fosfato deshidrogenasa.

Con la técnica 3'RACE se definió la secuencia de 3'UTR del gen *GAPDH*. Esto permitió caracterizar al terminador T_{GAP} como un terminador transcripcional poliadenilado, con un sitio de corte y poliadenilación con secuencia "TA", secuencia que coincide con el motivo de secuencia $Py(A)_n$ reportada en la literatura para levaduras (Guo et al. 1996). La posición en la que se encontró el sitio "TA" definió la longitud de la región 3'UTR con 51 nt. A los 51 nt de la región 3'UTR se le adicionaron 50 nt río abajo del sitio de poliadenilación para una longitud total de T_{GAP} de 101 pb. El criterio de adición de 50 nt se basó en el diseño de terminadores propuesto por Curran et al. (Curran et al. 2013) para *S. cerevisiae*, que reporta la adición de 50 pb después del final del 3'UTR.

El criterio de longitud de los terminadores y sus regiones 3'UTR a ser subclonados en un casete de expresión ha sido variable. Tuller et al. (Tuller et al. 2009) midieron la longitud de cada elemento de los genes en la levadura *S. cerevisiae* y reportaron un valor promedio de 136 pb para la región 3'UTR con una desviación estándar, en el mismo orden de magnitud, de 138 pb y para el terminador un valor promedio de 275 pb con 765 pb de desviación estándar. En este sentido, los diseños de terminadores de *K. phaffii* se han basado en estudios en *S. cerevisiae*. Por ejemplo, Vogl et al. (Vogl et al. 2016) utilizó una longitud de 250 pb para la región 3'UTR. Ramakrishnan et al. (Ramakrishnan et al. 2020) utilizó 500 pb de longitud total del terminador. Ito et al. (Ito et al. 2020) determinó por 3'RACE la longitud de la región 3'UTR de 10 terminadores de *K. phaffii* con valores entre 57 y 173 nt. Como dato de interés durante el desarrollo del presente estudio se realizó un análisis de distribución de frecuencia de las longitudes de la región 3'UTR de los 3244 genes reportados en la base de datos GenBank (Anexo 8) para *K. phaffii* y se observó que las mayores frecuencias de longitud de la región 3'UTR en *K. phaffii* se encuentran entre 25 y 72 nt.

En 2015 Curran et al. (Curran et al. 2015) reportaron el empleo de terminadores sintéticos cortos de 35 a 70 pb, con utilidad en la expresión de genes heterólogos en *S. cerevisiae*. Para el diseño de los terminadores utilizaron elementos consenso que forman parte de la arquitectura de los terminadores transcripcionales. Estos elementos son: de eficiencia, de

posicionamiento y próximo que flanquea el sitio de corte y poliadenilación. Para las secuencias consenso, se basaron en reportes de la literatura como el de Guo y Sherman (Guo y Sherman 1996), Van Helden et al. (Van Helden et al. 2000) y Graber et al. (Graber et al. 1999). Con ello lograron terminadores funcionales con longitud promedio de 49 pb, longitud relativamente corta cuando se compara con los terminadores nativos. Según Tuller et al. (Tuller et al. 2009), reportaron que los terminadores sintéticos obtuvieron 3.7 veces más producción de proteína fluorescente y 4.4 veces más transcritos que con el terminador CYC1 y obtuvieron características de funcionalidad similares o incrementadas respecto a los terminadores nativos en el contexto de las rutas metabólicas. Los terminadores sintéticos también se evaluaron y se comprobó su funcionalidad en la levadura *Yarrowia lipolytica*. Con una producción de más del 60 % de una proteína fluorescente cuando se compara con el terminador CYC1 en *Yarrowia lipolytica*; demostraron que los terminadores sintéticos cortos, diseñados con los elementos reguladores, son transferibles entre diversas especies de levaduras y altamente funcionales. De esta forma, demostraron que los tres elementos de eficiencia, posicionamiento y próximo, que flanquean, el sitio de corte y poliadenilación constituyen secuencias conservadas y son señales necesarias y suficientes en los terminadores poliadenilados. El terminador del gen *GAPDH* definido en el presente proyecto, al compararse con los reportados en la literatura resultó ser un terminador corto. Con tan solo 101 pb, contiene en su región 3'UTR (51 pb) los elementos de eficiencia (e1) con la secuencia consenso GTATGT, el elemento de posicionamiento (e2) con la secuencia consenso AAATAG, el elemento próximo río arriba del sitio de corte (e3) con secuencia TTCATT y el sitio de corte definido con la secuencia TA. Justo seguido al sitio de corte y por lo tanto justo después de la región 3'UTR se encuentra el elemento próximo río abajo del sitio de corte (e4) con secuencia TATCTA. Su longitud resultó ser suficiente para adoptar una conformación en forma de horquilla y ser funcional para la terminación y como elemento regulador de la transcripción.

En el diseño de la secuencia nucleotídica a sintetizar, además de los 101 pb del terminador *GAPDH* (T_{GAP}), fue necesario incorporar los sitios de restricción para las enzimas *NotI* y *Bsu36I*. Esto permitió verificar la presencia de T_{GAP} en el plásmido pUCIDT_{GAP}, subclonar el inserto T_{GAP} en pGAHFTEII y caracterizar el vector pP_{GAP}-FTEII- T_{GAP} . La secuencia

diseñada además de los 15 nt de los sitios de restricción, incorporó 510 nt de la estructura del vector pGAHFTEII, y los 101 nt del terminador T_{GAP}, para un total de 620 pb. Con el subclonaje de T_{GAP} se sustituyó la secuencia de 334 pb del terminador *AOXI* en el casete de expresión de pGAHFTEII. De esta forma el plásmido pP_{GAP}-FTEII-T_{GAP} resultó con 8353 pb; 233 pb menos que el vector de origen pGAHFTEII, con 8586 pb. Así quedó construido el casete de expresión en pP_{GAP}-FTEII-T_{GAP} con la combinación de los elementos reguladores promotor y terminador del gen *GAPDH*.

Estudios previos de expresión del gen heterólogos *FTEII* en *K. phaffii* realizados en nuestro laboratorio han utilizado la combinación de promotor y terminador *AOXI* y recientemente Herrera-Estala et al. (Herrera-Estala et al. 2020) en la búsqueda de utilizar promotores alternativos a los inducibles con metanol, reportaron el estudio de expresión de *FTEII* empleando un casete de expresión con P_{GAP} y terminador *AOXI*. Transformaron la cepa KM71 de *K. phaffii* con el vector de expresión pGAHFTEII, y obtuvieron alrededor de 200 His⁺ transformantes; a varios de los cuales se les analizó el ADN genómico por PCR de integración con oligonucleótidos dirigidos a región del P_{GAP} y al gen *FTEII*, observándose un producto amplificado con 1311 pb. Este resultado también se observó en nuestro estudio cuando se analizaron por PCR de integración con doble pareja de oligonucleótidos las construcciones KM71/P_{GAP}-FTEII-T_{AOXI} (cepa KM71 transformada con el vector de expresión pGAHFTEII) y KM71/P_{GAP}-FTEII-T_{GAP} (cepa KM71 transformada con pGAHFTEII modificado por subclonaje de T_{GAP} en sustitución de T_{AOXI}). Esto indicó que la región del P_{GAP} permaneció invariable luego de todo el proceso de subclonaje, inserción del vector y construcción de ambas cepas.

Las condiciones de cultivo en matraz propuestas por Herrera-Estala et al. (Herrera-Estala et al. 2020) son similares a las utilizadas en nuestro estudio: medio BMG, suplementado con 0.1 % de CaCl₂ y DO_{600nm} inicial de 1.4, incubación a 30 °C y 250 rpm, pero con adición de glicerol a las 14 h de cultivo en lugar de a las 20 h porque a las 14 h de cultivo se observó reducción del crecimiento celular por agotamiento de la fuente de carbono. Otros autores (Ramakrishnan et al., 2020) han utilizado medio BMGY (con 2 % de glicerol) en el cultivo de cepas de *K. phaffii* con P_{GAP}. La incorporación del extracto de levadura en el medio de cultivo propicia mayor cantidad de nutrientes para el crecimiento

celular, pero en nuestro caso no resultó factible su uso porque interfiere en la cuantificación de proteínas totales por el método de Bradford y en la electroforesis de proteínas. No obstante, con las condiciones de cultivo establecidas, el valor de la velocidad específica de crecimiento en la fase exponencial de las dos construcciones realizadas fue de 0.19 h^{-1} . Este valor también se ha reportado por varios autores cuando se utiliza casete de expresión con promotor del gen *GAPDH*, independientemente del tipo de cepa de *K. phaffii*, del medio de cultivo y de las fuentes de carbono empleada: glucosa o glicerol (García-Ortega et al. 2013; Looser et al. 2015). De esta forma se lograron cultivos celulares con poca variabilidad de densidad celular entre réplicas, entre clonas y entre cepas con valor cercano a 8.0 g/L . Estos resultados de conjunto con la demostración de que las muestras analizadas, 18 clonas de la cepa KM71/ $P_{\text{GAP-FTEII-T}_{\text{AOX1}}}$ y 25 clonas de KM71/ $P_{\text{GAP-FTEII-T}_{\text{GAP}}}$, resultaron ser representativas (con 95 % de confianza) de los más de 1000 transformantes obtenidos, nos permitió realizar una adecuada selección de la clona unicopia utilizando el criterio de rendimiento proteína extracelular/biomasa.

Como resultado del análisis de los valores de rendimiento proteína extracelular/biomasa se observó gran variabilidad de este parámetro entre las clonas de una misma cepa. El coeficiente de variación del rendimiento proteína extracelular/biomasa para la cepa KM71 con casete de expresión basado en $P_{\text{GAP-T}_{\text{AOX1}}}$ fue de 31.2 % y para $P_{\text{GAP-T}_{\text{GAP}}}$ fue de 42.7 %. Del análisis de correlación se evidenció que la variabilidad del rendimiento no estaba asociado a la variabilidad en la densidad celular entre las clonas (con coeficiente de variación de 5.9 % para $P_{\text{GAP-T}_{\text{AOX1}}}$ y de 7.8 % para $P_{\text{GAP-T}_{\text{GAP}}}$), sino que se asoció a la variabilidad de la secreción de proteína al medio extracelular, con un coeficiente de correlación del 96 %. El hecho de que el incremento de la secreción de proteínas heterólogas al medio extracelular de algunas clonas es significativamente diferente al resto de las clonas de una misma cepa en igualdad de condiciones de clonaje y ambientales tiene su explicación en el concepto de variación clonal en *K. phaffii*.

Este concepto sugiere que la variación en la producción de la proteína heteróloga entre clonas de una misma cepa se relaciona a la dosis génica del gen heterólogo insertado durante el proceso de transformación al genoma del hospedero (Aw et al. 2017). Existen varios reportes que intentan explicar este fenómeno de variación clonal descrito como el

fenómeno de mayor influencia en la expresión de un gen heterólogo (Viader-Salvadó et al. 2006). Aunque también existen otros factores que influyen como las deleciones o polimorfismos de un solo nucleótido, reordenamientos genómicos, y el lugar del sitio del locus de integración (Vogl et al. 2018). Es por ello que para realizar estudios comparativos entre cepas, como en nuestro caso, se debe llevar a cabo un proceso de selección de la cepa por dosis génica y realizar las comparaciones entre cepas con igualdad de dosis génica. Usualmente las comparaciones se realizan entre cepas unicopias.

Tanto en la construcción de la cepa P_{GAP}-T_{AOX1} como de la P_{GAP}-T_{GAP} se obtuvieron más de 1000 transformantes. Por lo que para seleccionar una clona unicopia inicialmente se estableció como criterio de selección el rendimiento proteína extracelular/biomasa. Bajo este criterio también se seleccionó la clona de mayor rendimiento proteína extracelular/biomasa como posible multicopia y posible sobre productora. Este criterio se ha usado acertadamente en estudios realizados en el laboratorio (Caballero-Pérez et al. 2021; Herrera-Estala et al. 2020). Para la determinación de la concentración de proteínas totales presentes en el medio de cultivo, se utilizó la determinación por el método de Bradford. Método adecuado para la selección de las clonas, considerando que en *K. phaffii* la mayor cantidad de proteínas excretadas pertenecen a la proteína heteróloga. Utilizando esta misma metodología de análisis en la selección de posibles clonas unicopias, inicialmente seleccionamos las de menor rendimiento proteína extracelular/biomasa. Pero las clonas de menor rendimiento resultaron ser clonas revertantes. Estas clonas perdieron o no hubo una adecuada inserción del casete de expresión en su genoma. Es recomendable, para la selección de clonas posibles unicopia realizar un diagrama de cajas y bigotes con los valores de rendimiento proteína/biomasa y excluir las clonas de menor rendimiento con valores atípico, y escoger una de las clonas que se encuentren dentro del rango del diagrama, pertenecientes a los cuartiles 2 y 3. Las clonas seleccionadas bajo este criterio resultaron ser clonas unicopias, demostrado por dosis génica. Este resultado concuerda con lo reportado por Higging et al. (Higging et al. 1998): “la inserción de múltiples genes representa del 1 al 10 % de los transformantes”, indicando que la mayoría de las clonas transformadas son unicopias y representan del 90 al 99 % de los transformantes. Una vez que se seleccionaron las clonas unicopias de las dos cepas de *K. phaffii*, se evaluó el efecto del terminador en la expresión de *FTEII* mediante: la evaluación de la funcionalidad de

los elementos reguladores por la presencia y el análisis de los niveles de transcritos del gen heterólogo y del $Y_{p/x}$ extracelular en diferentes condiciones de cultivo y a diferentes velocidades específicas de crecimiento.

Con la funcionalidad de los elementos reguladores se constató que el casete de expresión con la combinación P_{GAP} - T_{GAP} es capaz de generar transcritos del gen heterólogo. A este resultado se llegó a través de la amplificación de un fragmento del ADNc obtenido por retrotranscripción, amplificación que se generó por la hibridación de los oligonucleótidos específicos al gen heterólogo *FTEII*. Los controles negativos evidenciaron la ausencia de inespecificidades y la amplificación del gen endógeno de β -actina evidenció la correcta ejecución de la consecución de pasos del RT-PCR (cultivo celular, extracción y purificación de ARN, tratamiento con DNasa, retrotranscripción y PCR).

Con el análisis de los niveles de transcritos del gen heterólogo y del $Y_{p/x}$ extracelular en diferentes condiciones de cultivo y a diferentes velocidades específicas de crecimiento se estudió el efecto de la combinación de los elementos reguladores provenientes de un mismo gen y la influencia de la fuente de carbono glucosa o el glicerol en la expresión génica. Este tema ha resultado controversial, estudios en *S. cerevisiae* compararon terminadores múltiples con promotores nativos y no nativos y no se observó un efecto claro sobre la expresión génica. Vogl et al. (Vogl et al. 2016) obtuvieron los mayores niveles de expresión con la combinación promotor terminador *AOXI*. Sin embargo, estudios recientes publicados durante el desarrollo de nuestro trabajo de tesis han demostrado que el acoplamiento del P_{GAP} con el terminador *GAPDH* no mejoró la expresión de la proteína reportera (Ramakrishnan et al., 2020). Probablemente el efecto de la alta expresión en la combinación de elementos reguladores para un mismo gen nativo, como en el caso del promotor y terminador *AOXI*, se deba a la fortaleza del promotor más que al efecto que pueda ejercer el terminador ya que se ha demostrado que cuando el promotor es fuerte e inducible, como el promotor *AOXI* no se observa marcada diferencia entre los distintos terminadores (Ramakrishnan et al. 2020). Los resultados en nuestro estudio concuerdan con estos resultados y coinciden con los reportados recientemente por Ito et al. 2020 (Ito et al. 2020). En su estudio evaluó el efecto de 72 diferentes terminadores en un casete de expresión con P_{GAP} , se observó que la producción

de la proteína heteróloga fue 1.4 veces superior con terminador *AOX1* que con terminador *GAPDH*.

Aunque normalmente se utiliza una proteína verde fluorescente intracelular como proteína reportera para medir la fuerza del promotor (Hartner et al. 2008; Qin et al. 2011; Prielhofer et al. 2013), este enfoque mide la transcripción junto con el proceso de traducción, y no evalúa el proceso de secreción de proteínas, a pesar de que el sistema de expresión en *K. phaffii* se emplea frecuentemente para la producción extracelular de proteínas recombinantes. Por lo tanto, en el presente trabajo decidimos evaluar los niveles de transcritos del gen heterólogo mediante RT-qPCR y calcular el valor extracelular de $Y_{p/x}$ utilizando un gen reportero heterólogo que codifica para una proteína secretada al medio de cultivo. Esta estrategia es similar a la utilizada previamente para estudiar el promotor y el terminador transcripcional del gen *1033* de *K. phaffii*, que presenta una actividad transcripcional constitutiva, débil y no dependiente de metanol (Robainas-del-Pino et al. 2023). Además, este enfoque resulta útil no solo para evaluar los efectos del promotor y el terminador transcripcional sobre la transcripción del gen heterólogo y la estabilidad del transcrito, sino también para analizar la influencia de los elementos regulatorios sobre la traducción y la secreción de proteínas.

Los perfiles de actividad transcripcional de *FTEII* y *GAPDH* mostraron patrones diferentes en respuesta a las fuentes de carbono, los valores de μ y la combinación P_{GAP} -terminador transcripcional presente en la estructura del gen heterólogo. Específicamente, las células P_{GAP} - T_{AOX1} mostraron patrones transcripcionales distintos para los genes *FTEII* y *GAPDH* al ser cultivadas en glucosa o glicerol. Estos resultados indican una probable regulación competitiva de la actividad transcripcional entre ambos genes debido a la presencia de T_{AOX1} en la estructura génica de *FTEII*. Esta regulación competitiva podría dar lugar a diferencias en los mecanismos de regulación *cis* y/o *trans*, y a niveles de transcritos más altos del gen endógeno *GAPDH* en comparación con el gen heterólogo *FTEII* a las 6 y 12 h (valores de μ altos y medios, respectivamente) en cultivos tanto con glucosa como con glicerol. Una competencia regulatoria similar entre un gen heterólogo y la expresión del gen *GAPDH* utilizando distintas combinaciones

P_{GAP} -terminador transcripcional ya ha sido reportada previamente (Dou et al. 2021; Robainas-del-Pino et al. 2023).

Los niveles de transcritos más altos de ambos genes observados en células cultivadas con glucosa en comparación con las cultivadas con glicerol durante la fase de crecimiento exponencial (6 y 12 h de cultivo) en la cepa $P_{GAP-T_{AOX1}}$ coinciden con reportes previos donde se ha observado que la actividad transcripcional de P_{GAP} en células cultivadas con glucosa es mayor que en aquellas cultivadas con glicerol (Waterham et al. 1997).

En la cepa $P_{GAP-T_{AOX1}}$, cuando la glucosa o el glicerol comenzaban a agotarse y la μ disminuía, los genes *FTEII* y *GAPDH* mostraron una regulación negativa significativa. Este patrón transcripcional en ambos genes probablemente fue provocado por el agotamiento de la fuente de carbono. Sin embargo, los niveles de transcritos más altos de *FTEII* en comparación con *GAPDH* en células cultivadas con glucosa y con glicerol a las 18 h de cultivo sugieren que la regulación negativa inducida por el agotamiento de la fuente de carbono fue más intensa en presencia de T_{GAP} que en presencia de T_{AOX1} . Además, los niveles de transcritos de *GAPDH* más altos a las 18 h de cultivo en glicerol en comparación con los cultivos en glucosa indican que la regulación negativa inducida por el agotamiento de la fuente de carbono en presencia de T_{GAP} fue aún más intensa en los cultivos con glucosa que en los de glicerol. Estos hallazgos indican la existencia de un mecanismo diferente de regulación metabólica negativa y/o una mayor estabilidad del ARNm debido a la presencia de T_{AOX1} (Ramakrishnan et al. 2020; Ito et al. 2020; Herrera-Estala et al. 2022). Por lo tanto, la regulación metabólica negativa de ambos genes dependió de la combinación promotor P_{GAP} -terminador transcripcional dentro de la estructura génica y de la fuente de carbono.

Cuando los genes heterólogo *FTEII* y endógeno *GAPDH* compartieron la misma combinación promotor-terminador en sus estructuras génicas (es decir, en la cepa $P_{GAP-T_{GAP}}$), se observó una actividad transcripcional comparable de ambos genes en cultivos con glucosa y glicerol. Estos resultados respaldan la idea de que la presencia de T_{GAP} en las estructuras génicas de *FTEII* y *GAPDH* desencadenó una competencia equilibrada en su regulación transcripcional a través de elementos regulatorios *cis* y *trans*. Los niveles de transcritos de *FTEII* y *GAPDH* mostraron una respuesta dinámica frente a

la μ y a las fuentes de carbono. Los niveles más altos de transcritos se alcanzaron en cultivos con glucosa a una μ alta (0.253 h^{-1}). Sin embargo, los niveles de transcritos fueron más altos en cultivos con glicerol en comparación con los de glucosa cuando la μ disminuyó (a las 12 y 18 h de cultivo). Estos hallazgos podrían estar relacionados con la presencia de T_{GAP} en ambos genes, lo que provoca una regulación génica dependiente no solo de la fuente de carbono sino también del valor de μ . Al igual que las células basadas en $P_{GAP-T_{AOX1}}$, las células $P_{GAP-T_{GAP}}$ también mostraron una regulación negativa de ambos genes a medida que la μ disminuía y las fuentes de carbono se agotaban. La regulación negativa metabólica fue más pronunciada en las células cultivadas con glucosa, lo que resultó en niveles de transcritos más bajos para ambos genes en comparación con las células cultivadas con glicerol a las 18 h de cultivo.

Esta investigación reveló patrones distintos en la actividad transcripcional de *FTEII* bajo el control de P_{GAP} entre las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$ a lo largo del tiempo de cultivo en glucosa y glicerol. Estos perfiles de actividad transcripcional fueron influenciados por la combinación del promotor P_{GAP} con el terminador transcripcional, las fuentes de carbono y las etapas del crecimiento celular (velocidad específica de crecimiento). Los niveles de transcritos de *FTEII* más bajos observados en las células $P_{GAP-T_{GAP}}$ a las 12 y 18 h de cultivo, en comparación con las células $P_{GAP-T_{AOX1}}$ cultivadas en glucosa, probablemente se debieron a una regulación negativa más fuerte inducida por la disminución de los niveles de glucosa en presencia de T_{GAP} , o a una mayor estabilidad del ARNm de *FTEII* conferida por T_{AOX1} . En consecuencia, los niveles de transcritos de *FTEII* se mantuvieron sin cambios entre las 12 y 18 h de cultivo en la cepa $P_{GAP-T_{AOX1}}$. Por el contrario, los niveles más altos de transcritos de *FTEII* observados en las células $P_{GAP-T_{GAP}}$ a las 6 y 12 h de cultivo en glicerol, en comparación con las células $P_{GAP-T_{AOX1}}$, sugieren una mayor actividad transcripcional de P_{GAP} cuando T_{GAP} estaba presente en la estructura génica de *FTEII*, en comparación con T_{AOX1} cuando se utilizó glicerol como fuente de carbono. Sin embargo, la cepa $P_{GAP-T_{AOX1}}$ mostró los niveles más altos de transcritos de *FTEII* a las 18 h en glicerol, lo que sugiere una regulación negativa más intensa debida al agotamiento del glicerol en presencia de T_{GAP} como terminador transcripcional en la estructura génica de *FTEII* en la cepa $P_{GAP-T_{GAP}}$, o una mayor estabilidad de los transcritos de *FTEII* conferida por T_{AOX1} .

La actividad transcripcional del gen *GAPDH* en las células $P_{GAP-T_{GAP}}$ cultivadas con glucosa tuvo un patrón similar al de la cepa $P_{GAP-T_{AOX1}}$. Los niveles de transcritos del gen *GAPDH* más bajos en las células $P_{GAP-T_{GAP}}$ durante las etapas de crecimiento con μ alta y media (0.253 y 0.123 h^{-1} , respectivamente), en comparación con las células $P_{GAP-T_{AOX1}}$, respaldan la idea de que la presencia de T_{AOX1} en la estructura génica de *FTEII* indujo una regulación distinta del gen *GAPDH*, lo que resultó en una mayor actividad transcripcional en las células $P_{GAP-T_{AOX1}}$ cultivadas con glucosa. Este efecto no se observó en los cultivos con glicerol, lo que sugiere que la presencia de T_{AOX1} en la estructura génica de *FTEII* no aumentó la actividad transcripcional del gen *GAPDH* a μ alta (0.253 h^{-1}) cuando se utilizó glicerol como fuente de carbono. Además, los niveles de transcritos de *GAPDH* fueron comparables entre ambas cepas a las 6 y 12 h de cultivo en glicerol, pero más altos para las células $P_{GAP-T_{AOX1}}$ a las 18 h en glicerol en comparación con las células $P_{GAP-T_{GAP}}$. Los niveles de transcritos de *GAPDH* disminuyeron en ambas cepas a medida que se agotaba la fuente de carbono entre las 12 y 18 h de cultivo. La regulación negativa metabólica fue más pronunciada en las células $P_{GAP-T_{AOX1}}$ cultivadas con glucosa, pero en la cepa $P_{GAP-T_{GAP}}$ más intensa cuando fue cultivada con glicerol. En conjunto, estos resultados indican que la combinación específica P_{GAP} -terminador dentro de la estructura del gen heterólogo, la fuente de carbono y la etapa del crecimiento celular influyen en la regulación transcripcional del gen *GAPDH*.

En general, los valores de $Y_{p/x}$ extracelular fueron más altos en la cepa $P_{GAP-T_{AOX1}}$ en comparación con la cepa $P_{GAP-T_{GAP}}$. Estos valores mostraron una tendencia creciente en la cepa $P_{GAP-T_{AOX1}}$ a medida que progresaba el cultivo, y hasta las 12 h de cultivo en la cepa $P_{GAP-T_{GAP}}$, con valores más altos en glucosa que en glicerol. Los resultados revelaron una relación directamente proporcional entre los niveles de transcritos de *FTEII* regulados por P_{GAP} y los valores de μ en la mayoría de los casos, mientras que se observó una relación inversamente proporcional entre los valores de $Y_{p/x}$ extracelular y los valores de μ . En consecuencia, se evidenció una relación inversamente proporcional entre los niveles de transcritos de *FTEII* y los valores de $Y_{p/x}$ extracelular. Estos hallazgos concuerdan con reportes previos sobre una mayor actividad transcripcional de P_{GAP} a μ alta (Looser et al., 2015), y con nuestros resultados anteriores (Herrera-Estala et al. 2022), donde mostramos que los valores de $Y_{p/x}$ extracelular para la misma proteína heteróloga (la fitasa *FTEII*)

bajo el control de P_{GAP} y T_{AOX1} fueron más altos bajo condiciones de μ baja y niveles bajos de transcritos del gen heterólogo. Estos resultados confirman aún más nuestra hipótesis previa de que el plegamiento adecuado de la proteína heteróloga y/o el procesamiento postraduccionnal dentro de la vía secretora representa el paso limitante para alcanzar altos rendimientos de proteínas extracelulares (Herrera-Estala et al. 2022).

La comparación de los cocientes de niveles de transcritos de *FTEII* entre las cepas $P_{GAP-T_{AOX1}}$ y $P_{GAP-T_{GAP}}$ cultivadas con la misma fuente de carbono, así como los cocientes correspondientes de $Y_{p/x}$ extracelular entre cepas, permitió conocer el efecto del terminador transcripcional dentro de la estructura del gen heterólogo sobre la actividad celular post-transcripcional de la proteína heteróloga (la traducción y/o la actividad de la vía de secreción de la proteína). Cocientes de $Y_{p/x}$ extracelular entre cepas más altos que los cocientes correspondientes de niveles de transcritos de *FTEII* indican que T_{AOX1} mejoró la actividad post-transcripcional en comparación con T_{GAP} ; en cambio, si el cociente es menor, T_{GAP} mejoró la actividad post-transcripcional en comparación con T_{AOX1} . La ausencia de diferencias entre ambos cocientes indica un efecto comparable del terminador sobre la actividad post-transcripcional.

Nuestros resultados revelan diferencias entre los cocientes entre cepas, lo que indica que el terminador transcripcional dentro de la estructura génica de *FTEII* desempeñó un papel fundamental al influir en la traducción y/o la secreción de la proteína heteróloga, como se ha descrito anteriormente para otros terminadores transcripcionales (Mayr 2019; Kuersten y Goodwin 2003). Además, los hallazgos indican que la presencia de T_{AOX1} dentro de la estructura del gen heterólogo mejoró significativamente la traducción y/o la secreción de la proteína heteróloga en cultivos con glucosa a μ alta (0.253 h^{-1}), así como en cultivos con glicerol a μ alta y media (0.253 y 0.123 h^{-1} , respectivamente), en comparación con T_{GAP} . En contraste, T_{GAP} fue más efectivo que T_{AOX1} al mejorar la actividad post-transcripcional en cultivos con glucosa a μ media (0.123 h^{-1}) y en cultivos con cualquiera de las fuentes de carbono ensayadas a μ baja (0.030 h^{-1}).

De forma similar a la comparación entre cepas, la comparación entre los cocientes de niveles de transcritos de *FTEII* en células cultivadas con glucosa y con glicerol, y los cocientes correspondientes de $Y_{p/x}$ extracelular según la fuente de carbono, ofrece

información sobre el efecto de la fuente de carbono en la actividad celular post-transcripcional de la proteína heteróloga (traducción y/o actividad de la vía de secreción de la proteína heteróloga). Se observaron diferencias entre los cocientes según la fuente de carbono, lo que indica que la actividad post-transcripcional fue influenciada por la fuente de carbono analizada. Los resultados indican que el glicerol mejoró la actividad post-transcripcional de la proteína heteróloga en comparación con la glucosa en cultivos a μ alta (0.253 h^{-1}) cuando la combinación $P_{GAP}\text{-}T_{AOX1}$ estaba presente en la estructura del gen heterólogo, tal como observamos anteriormente con otro par promotor-terminador usando el mismo gen reportero (Robainas-del-Pino et al., 2023). En contraste, la glucosa mejoró la actividad post-transcripcional de la proteína heteróloga en comparación con el glicerol a μ baja (0.030 h^{-1}) cuando estaba presente la combinación $P_{GAP}\text{-}T_{AOX1}$, y a μ media y baja (0.123 h^{-1} y 0.030 h^{-1} , respectivamente) cuando la combinación $P_{GAP}\text{-}T_{GAP}$ estaba presente en la estructura del gen heterólogo.

La caracterización bioquímica de la proteína heteróloga, fitasa FTEII, secretada al medio extracelular se reportó previamente por Viader et al. (Viader et al. 2010). Es característico para la FTEII observar un barrido en el gel SDS-PAGE en el rango de 45 a 66 kDa y después del tratamiento de desglicosilación con Endo Hf en el gel de electroforesis observar una banda con una masa molecular aparente de 39 kDa. Este resultado indicó claramente que la fitasa recombinante FTEII expresada en *K.phaffii* está altamente N-glicosilada. En la electroforesis de proteína realizada en el presente estudio se corrieron los sobrenadantes de cultivo tanto de la cepa $P_{GAP}\text{-}T_{AOX1}$ como $P_{GAP}\text{-}T_{GAP}$ sin tratamiento con Endo Hf y se observó un barrido en la zona alrededor de los 45 a 67 kDa. Una vez que las proteínas en el sobrenadante se trataron con Endo Hf apareció una banda por encima de los 36 kDa con peso molecular aparente de 39 kDa, constatando que en ambas cepas con la construcción $P_{GAP}\text{-}T_{AOX1}$ y con la construcción $P_{GAP}\text{-}T_{GAP}$ se secreta la proteína recombinante FTEII de forma glicosilada al medio extracelular. Además, se detectó actividad enzimática de fitasa en los sobrenadantes de cultivo. La fitasa es una enzima que hidroliza el fitato y se libera fosfato inorgánico libre. La detección de actividad enzimática en el sobrenadante de cultivo corroboró la expresión de fitasa FTEII biológicamente activa en el medio de cultivo. Los resultados evidenciaron presencia de actividad enzimática en el sobrenadante de cultivo de las cepas unicopias $P_{GAP}\text{-}T_{GAP}$ y

$P_{GAP-T_{AOX1}}$ con actividad volumétrica de fitasa con valor promedio de 0.25 U/mL y 0.34 U/mL, respectivamente.

9 CONCLUSIONES

1. Se construyó el vector de expresión pP_{GAP}-FTEII-T_{GAP} conformado por los elementos reguladores del gen *GAPDH* de *K. phaffii*.
 - La definición del terminador del gen *GAPDH* (T_{GAP}) como un terminador corto de tan solo 101 nt, que incluye la región 3'UTR de 51 nt que contiene las secuencias consenso necesarias de un terminador poliadenilado demostró ser suficiente para funcionar como un terminador transcripcional activo. Lo que resulta de interés para el diseño de vectores optimizados en el sistema de expresión de *K. phaffii*.
 - El subclonaje en el vector pGAHFTEII de una secuencia de 620 nt que contiene al terminador T_{GAP}, flanqueada por los sitios de corte *NotI* y *Bsu36I*, permitió reemplazar el terminador del gen *AOX1* (T_{AOX1}) por el terminador del gen *GAPDH* (T_{GAP}), posibilitando la construcción de un nuevo vector de expresión denominado pP_{GAP}-FTEII-T_{GAP}, regulado por la combinación P_{GAP}-T_{GAP}. Este vector resultó ser una herramienta clave para la evaluación comparativa del impacto de la combinación de los elementos reguladores P_{GAP}-T_{GAP} frente a P_{GAP}-T_{AOX1} en la expresión de un gen heterólogo en *K. phaffii*.
 - Los vectores de expresión pP_{GAP}-FTEII-T_{GAP} (8353 pb) y pGAHFTEII (8586 pb) contienen un casete de expresión con la región codificante de la proteína recombinante FTEII fucionada en su extremo N-terminal con el péptido prepro del factor α de *S. cerevisiae*. Esta construcción permitió la secreción de la proteína recombinante al medio extracelular y con ello la evaluación del impacto de la combinación de los elementos reguladores P_{GAP}-T_{GAP} y P_{GAP}-T_{AOX1} en el rendimiento extracelular FTEII/biomasa (Y_{p/x}).
2. Se construyeron las cepas KM71/P_{GAP}-FTEII-T_{AOX1} con el vector de expresión pGAHFTEII y KM71/P_{GAP}-FTEII-T_{GAP} con el vector de expresión pP_{GAP}-FTEII-T_{GAP}.
 - La transformación de la cepa KM71 de *K. phaffii* con los vectores linealizados pGAHFTEII o pP_{GAP}-FTEII-T_{GAP} permitió en cada caso la obtención de más de 1000 clonas transformantes seleccionadas por prototrofia a histidina. El cultivo celular de clonas representativas de la cepa KM71/P_{GAP}-FTEII-T_{AOX1} y de la cepa

KM71/P_{GAP}-FTEII-T_{GAP} mostraron un crecimiento exponencial definido para cepas bajo la regulación del promotor P_{GAP} con una velocidad específica de crecimiento de 0.19 h⁻¹ y poca variabilidad en la densidad celular durante 24 h de cultivo (coeficientes de variación inferiores al 8 %) lo que constata similitud fenotípica entre ambas construcciones.

- En ambas cepas se identificó una alta variabilidad en el rendimiento proteína/biomasa, con coeficientes de variación superiores al 30 %. Esta variabilidad no presentó correlación con la densidad celular pero mostró una correlación lineal directa con la concentración de proteína extracelular. Este resultado evidenció que el rendimiento de proteína está más influenciado por la variación de dosis génica del gen heterólogo *FTEII* insertado en el genoma de la levadura y por la eficiencia de secreción en cada clona que por la biomasa generada.
 - El análisis de dosis génica por qPCR confirmó en ambas cepas que las clonas de mayores rendimientos de proteínas extracelular/biomasa son clonas multicopias (con dos o tres copias del gen *FTEII*), mientras que las clonas de menor rendimiento proteína extracelular/biomasa, poseían una sola copia del gen reportero. Este análisis posibilitó la selección de clonas unicopias para ambas cepas, eliminó el efecto de la variación de dosis génica y estableció una base sólida para comparar el impacto de los terminadores T_{AOX1} y T_{GAP} en la expresión de genes y en la producción de proteínas recombinantes al medio extracelular en *K. phaffii*.
3. Se evaluó el efecto de la región terminadora del gen *GAPDH* sobre los niveles de transcritos de un gen heterólogo regulado por el promotor P_{GAP} y del gen endógeno *GAPDH*, y sobre el rendimiento producto/biomasa (Y_{p/x}) extracelular, en comparación con la región terminadora del gen *AOX1*, en cepas recombinantes de *K. phaffii*.
- La presencia de transcritos de *FTEII*, y la presencia de Fitasa en el sobrenadante de cultivo demostraron la funcionalidad de los elementos reguladores P_{GAP}-T_{AOX1} y P_{GAP}-T_{GAP} integrados en el genoma de las cepas de levadura KM71/P_{GAP}-FTEII-T_{AOX1} y KM71/P_{GAP}-FTEII-T_{GAP}, respectivamente.

- En ambas cepas (KM71/P_{GAP}-FTEII-T_{AOX1} y KM71/P_{GAP}-FTEII-T_{GAP}) se demostró que la velocidad específica de crecimiento en la fase exponencial, fue de aproximadamente 0.200 h⁻¹ lo que concuerda con el valor típico en la etapa de mayor crecimiento celular reportado para cepas de *K. phaffii* reguladas bajo el promotor P_{GAP}. Valor que se obtuvo independientemente de la fuente de carbono empleada (glucosa o glicerol).
- La regulación transcripcional tanto del gen heterólogo *FTEII* como del gen endógeno *GAPDH*, los valores de Y_{p/x} extracelular, y la traducción y/o secreción de la proteína heteróloga estuvieron influenciados por la combinación P_{GAP}-terminador transcripcional, actuando la fuente de carbono y la etapa de crecimiento celular como factores moduladores.
- Los mayores niveles de transcritos del gen heterólogo y del gen *GAPDH* endógeno se observaron en los cultivos en glucosa a alta velocidad específica de crecimiento (0.253 h⁻¹) y el mayor rendimiento extracelular Y_{p/x} con el terminador T_{AOX1}, en cultivos en glucosa y baja μ (0.030 h⁻¹).
- La presencia de T_{AOX1} o T_{GAP} dentro de la estructura génica heteróloga activó distintos elementos reguladores génicos en cada cepa, lo que condujo a una modulación diferencial de la regulación génica para los genes heterólogo y endógeno *GAPDH*, a pesar de que ambos genes estaban bajo el control del mismo promotor (P_{GAP}).
- El T_{AOX1} indujo una regulación competitiva de la actividad transcripcional entre el gen heterólogo y el gen *GAPDH* endógeno, lo que resultó en una mayor actividad transcripcional del gen *GAPDH*.
- El T_{AOX1} condujo a una mayor estabilidad del ARNm y desencadenó distintos mecanismos de regulación a la baja metabólica debido al agotamiento de la fuente de carbono en comparación con T_{GAP}.
- El T_{AOX1} mejoró la traducción y/o la actividad de secreción de proteínas heterólogas a una alta velocidad específica de crecimiento (0.253 h⁻¹), mientras que T_{GAP} fue más efectivo para mejorar la actividad postranscripcional a una velocidad específica de crecimiento baja (0.030 h⁻¹), independientemente de la fuente de carbono.

- Para aumentar el $Y_{p/x}$ extracelular de un sistema de *K. phaffii* basado en el promotor P_{GAP} , T_{AOX1} resultó ser un terminador transcripcional más adecuado que T_{GAP} , y el cultivo debe llevarse a cabo a una velocidad específica de crecimiento baja utilizando glucosa como fuente de carbono.

10 PERSPECTIVAS

El presente trabajo propone como perspectiva para el incremento del rendimiento $Y_{p/x}$ de FTEII emplear la cepa de *K. phaffii* basada en la combinación de elementos reguladores $P_{GAP}-T_{AOX1}$, con cultivos en glucosa como fuente de carbono y con regulación de la velocidad específica de crecimiento a niveles bajos en el orden de 0.030 h^{-1} . Se recomienda emplear una cepa multicopia. Los resultados del presente trabajo ofrecen además nuevas perspectivas al conocimiento general sobre el efecto de la combinación de elementos reguladores promotor-terminador en la expresión y secreción de proteínas heterólogas en *K. phaffii* y brinda evidencias de que estos pueden ser modulados por la fuente de carbono y la velocidad específica de crecimiento. Estos hallazgos contribuyen al conocimiento y al desarrollo de estrategias de investigación para lograr ajustes finos en el incremento de la eficiencia en bioprocesos.

11 BIBLIOGRAFÍA

Aw R, Barton GR, Leak DJ. 2017. Insights into the prevalence and underlying causes of clonal variation through transcriptomic analysis in *Pichia pastoris*. *Applied microbiology and biotechnology* 101:5045-5058. DOI 10.1007/s00253-017-8317-2.

Beelman C, Parker R. 1995. Degradation of mRNA in eukaryotes. *Cell* 81:179-183. DOI: 10.1016/0092-8674(95)90326-7.

Bollok M, Resina D, Valero F, Ferrer P. 2009. Recent patents on *Pichia pastoris* expression system: expanding the toolbox for recombinant protein production. *Recent patent on biotechnology* 3:192-201. DOI: 10.2174/187220809789389126.

Bretthauer RK, Castellino FJ. 1999. Review. Glycosylation of *Pichia pastoris*-derived proteins. *Biotechnol. Appl. Biochem.* 30:193–200. DOI 10.1111/j.1470-8744.1999.tb00770.x.

Caballero-Perez A, Viader-Salvadó JM, Herrera-Estala AL, Fuentes-Garibay JA, Guerrero-Olazarán M. 2021. Buried Kex2 Sites in Glargine Precursor Aggregates Prevent Its Intracellular Processing in *Pichia pastoris* Muts Strains and the Effect of Methanol-Feeding Strategy and Induction Temperature on Glargine Precursor Production Parameters. *Applied Biochemistry and Biotechnology* 193: 2806-2829. DOI: 10.1007/s12010-021-03567-z.

Çalik P, Atab Ö, Güneş H, Massahib A, Boya E, Keskin A, Öztü S, Zerze GH, Özdamarc 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. DOI 10.1016/j.bej.2014.12.003.

Castillo-Galván M. 2014. Análisis de la respuesta metabólica involucrada en la optimización de condiciones de cultivo para incrementar la producción de proteínas heterólogas en la levadura metilotrófica *Pichia pastoris*. UANL.

Cereghino JL, Cregg JM. 2000. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS microbiology reviews 24:45-66. DOI: 10.1111/j.1574-6976.2000.tb00532.x.

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 116(4):321-335. DOI 10.1016/j.jbiotec.2004.12.010.

Cregg JM, Cereghino JL, Shi J, Higgins DR. 2000. Recombinant Protein Expression in *Pichia pastoris*. Molecular Biotechnology 16:23-52. DOI: 10.1385/MB:16:1:23.

Cregg JM, Tolstorukov I, Kusari A, Sunga AJ, Madden K, Chappell T. 2018. Expression of recombinant genes in the yeast *Pichia pastoris*. Current Protocols Essential Laboratory Techniques, e25.

Cregg JM, Tschopp JF, Stillman C, Siegel R, Akong M, Craig WS, Buckholz RG, Madden KR, Kellaris PA, Davis GR, Smiley BL, Cruze J, Torregrossa R, Velicelebi G, Thill GP. 1987. High-level expression and efficient assembly of hepatitis B surface antigen in the methylotrophic yeast *Pichia pastoris*. Bio/Technology 5:479-485.

Cregg JM, Vedvick TS, Raschke WC. 1993. Recent advances in the expression of foreign genes in *Pichia pastoris*. Biotechnology 11:905-910. DOI: 10.1038/nbt0893-905.

Curran KA, Karim AS, Gupta A, Alper HS. 2013. Use of expression-enhancing terminators in *S. cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. Metabolic Engineering 19:88-97. DOI: 10.1016/j.ymben.2013.07.001.

Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. 2015. Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. ACS Synthetic Biology 4(7):824-832. DOI: 10.1021/sb5003357.

Döring F, Theis S, Daniel H. 1998. Use of the glyceraldehydes-3-phosphate dehydrogenase promoter for production of functional mammalian membrane transport

proteins in the yeast *Pichia pastoris*. *Biochemical and Biophysical Research Communications* 250:531-535. DOI: 10.1006/bbrc.1998.9342.

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):1-12. DOI 10.1186/s12934-021-01648-6.

Fernández KB. 2015. Análisis transcripcional de la vía de secreción y metabolismo del metanol durante la producción de una proteína heteróloga en *Pichia pastoris*. UANL.

Garcia-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. DOI: 10.1016/j.nbt.2019.06.002.

Garcia-Ortega X, Ferrer P, Montesinos 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. DOI: 10.1016/j.bej.2013.07.013.

Goodrick JC, Xu M, Finnegan R, Schilling BM, Schiavi S, Hoppe H, Wan NC. 2001. High-level expression and stabilization of recombinant human chitinase produced in a continuous constitutive *Pichia pastoris* expression system. *Biotechnology and Bioengineering* 74:492-497. DOI: 10.1002/bit.1140.

Graber JH, Cantor CR, Mohr SC, Smith TF. 1999. Genomic detection of new yeast pre-mRNA 3'-end-processing signals. *Nucleic Acids Research* 27:888-894. DOI: 10.1093/nar/27.3.888.

Guerrero-Olazarán M, Cabrera-Barrera EL, Galán-Wong LJ, Viader-Salvadó JM. 2004. Biotecnología de Proteínas Recombinantes para la aplicación en acuicultura. In: *Memorias del VII Simposio Internacional de Nutrición Acuícola*, México, Noviembre 16-18.

Guerrero-Olazarán M, Rodríguez-Blanco L, Garreon-Treviño JG, Gallegos-López JA, Viader-Salvadó JM. 2010. Expression of a *Bacillus* Phytase C Gene in *Pichia pastoris*

and Properties of the Recombinant Enzyme. *Apply Environmental Microbiology* 76: 5601-5608. DOI: 10.1128/AEM.00762-10.

Guo Z, Sherman F. 1996. 3'-end-forming signals of yeast mRNA. *Molecular and cellular biology* 15:5983-5990. DOI 10.1016/s0968-0004(96)10057-8.

Hartner FS, Ruth C, Langenegger D, Johnson SN, Hyka P, Lin-Cereghino GP, Lin-Cereghino J, Kovar K, Cregg JM, Glieder A. 2008. Promoter library designed for fine-tuned gene expression in *Pichia pastoris*. *Nucleic acids research* 36:e76-e76. DOI: 10.1093/nar/gkn369.

Helden JV, Olmo MLD, Pérez-Ortín JE. 2000. Statistical analysis of yeast genomic downstream sequences reveals putative polyadenylation signals. *Nucleic Acids Research* 28:1000-1010. DOI: 10.1093/nar/28.4.1000.

Herrera-Estala AL, Fuentes-Garibay JA, Guerreo-Olazarán M, Viader-Salvadó M. 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 *PSAI-1* gene expression yielding enhanced extracellular productivity. *Journal of Biotechnology* 352: 59-67. DOI: 10.1016/j.jbiotec.2022.05.010.

Higgins DR, Cregg JM. 1998. Methods in molecular biology: *Pichia* protocols. Production of recombinant proteins in fermenter cultures of the yeast *Pichia pastoris*. *Current Opinion in Biotechnology* 13:329-332. DOI: 10.1016/S0958-1669(02)00330-0.

Inan M, Meagher MM. 2001. The effect of ethanol and acetate on protein expression in *Pichia pastoris*. *Journal Bioscience Bioengineering* 92:337-41. DOI: 10.1016/S1389-1723(01)80236-X.

Invitrogen. 2010. pGAP Z A, B, and C pGAP Z α A, B, and C *Pichia* expression vectors for constitutive expression and purification of recombinant proteins, *Invitrogen manual* 25:25-0174.

Irniger S, Braus GH. 1994. Saturation mutagenesis of a polyadenylation signal reveals a hexanucleotide element essential for mRNA 3' end formation in *S. cerevisiae*. *Proc. Natl Acad. Sci. USA* 91:257-261. DOI:10.1073/pnas.91.1.257.

- Ito Y, Terai G, Ishigami M, Hashiba N, Nakamura Y, Bamba T, 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. DOI: 10.1093/nar/gkaa1066.
- Jozala AF, Pessoa A. 2016. Biopharmaceuticals from microorganisms: from production to purification. *Brazilian Journal of Microbiology* 47:51-63. <https://www.scielo.br/j/bjm/a/q5VcnRQ9b888WW6XL7X7tTc/?format=html&lang=en>.
- Kuersten S, Goodwin EB. 2003. The power of the 3'UTR: translational control and development. *Nature Reviews Genetics* 4(8):626-637. DOI 10.1038/nrg1125.
- 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. DOI: 10.1007/s10295-009-0638-4.
- Laroche Y, Storme V, De Meutter J, Messens J, Lauwereys M. 1994. High-level secretion and very efficient isotopic labeling of tick anticoagulant peptide (TAP) expressed in the methylotrophic yeast, *Pichia pastoris*. *Bio/Technology* 12:1119-1124. DOI: 10.1038/nbt1194-1119.
- Lewin B. 2012. mRNA stability and Localization. In: *Lewin's Genes XI*. Krebs JE, Goldstein ES, Kilpatrick ST (eds). Unite State of America, pp. 622-646.
- Li P, Anumanthan A, Gao XG, Ilangovan K, Suzara VV, Duzgune N, Renugopalakrishnan V. 2007. Expression of recombinant proteins in *Pichia pastoris*. *Applied Microbiology and Biotechnology* 142:105-124. DOI: 10.1007/s12010-007-0003-x.
- Liu Y, Beyer A, Aebersold R. 2016. On the dependency of cellular protein levels on mRNA abundance. *Cell* 165: 535–550. DOI: 10.1016/j.cell.2016.03.014.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25:402-408. DOI: 10.1006/meth.2001.1262.

Looser V, Lüthy D, Straumann M, Hecht K, Melzoch K, Kovar1 K. 2017. Effects of glycerol supply and specific growth rate on methanol-free production of CALB by *P. pastoris*: functional characterization of a novel promoter. *Appl Microbiol Biotechnol* 101:3163–3176. DOI 10.1007/s00253-017-8123-x.

Looser V, Bruhlmann F, Bumbaka C, Stengera M, Costab A, Camattaric D, Fotiadisb KK. 2015. Cultivation strategies to enhance productivity of *Pichia pastoris*: A review. *Biotechnology Advances* 33:1177-1193. DOI: 10.1016/j.biotechadv.2015.05.008.

López M, Loera O, Guerrero-Olazarán M, Viader-Salvadó JM, Gallegos-López JA, Fernández FJ, Favela-Torres E, Viniegra-González G. 2010. Cell growth and *Trametes versicolor* laccase production in transformed *Pichia pastoris* cultured by solid-state or submerged fermentations. *Journal of Chemical Technology and Biotechnology* 85(4):435-440. DOI 10.1002/jctb.2306.

Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM. 2005. Heterologous protein production using the *Pichia pastoris* expression system. *Yeast* 22:249-270. DOI: 10.1002/yea.1208.

Markets M. 2024. Protein Expression Market. Available at: <http://www.marketsandmarkets.com/PressReleases/protein-expression.asp>. Acceso Noviembre 2024.

Marquardt S, Hazelbaker DZ, Buratowski S. 2011. Distinct RNA degradation pathways and 3' extensions of yeast non-coding RNA species. *Transcription* 2(3):145-154. DOI 10.4161/trns.2.3.16298.

Mattanovich D, Jungo C, Wenger J, Dabros M, Maurer M, 2014. Yeast suspension culture. In: Meyer H, Schmidhalter DR, eds. *Industrial scale suspension culture of living cells*. Hoboken, NJ: Wiley-Blackwel, 94–129. DOI: 10.1002/9783527683321.ch02.

Mayr C. 2019. What are 3'UTRs doing?. *Cold Spring Harbor Perspectives in Biology* 11(10): a034728. DOI 10.1101/cshperspect.a034728.

Monod J. 1949. The growth of bacterial cultures. *Annual Review of Microbiology* 3: 371-394.

- Neil H, Malabat C, d'Aubenton-Carafa Y, Xu Z, Steinmetz LM, Jacquier A. 2009. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* 457(7232):1038-1042. DOI 10.1038/nature07747.
- Nielsen J. 2013. Production of biopharmaceutical proteins by yeast, *Bioengineered* 4:207-211. DOI: doi.org/10.4161/bioe.22856.
- Olezdzka G, Dazbrowski S, Kur J. 2003. High levels expression, secretion and purification of the thermostable aqualysin I from *Thermus aquaticus* YT-1 in *Pichia pastoris*. *Protein Expression and Purification* 29:223-229. DOI: 10.1016/S1046-5928(03)00060-3.
- Prielhofer R, Maurer M, Klein J, Wenger J, Kiziak C, Gasser B, Mattanovich D. 2013. Induction without methanol: novel regulated promoters enable high-level expression in *Pichia pastoris*. *Microbial Cell Factories* 12(5). DOI: 10.1186/1475-2859-12-5.
- Qin X, Qian J, Yao G, Zhuang Y, Zhang S, Chu J. 2011. GAP promoter library for fine-tuning of gene expression in *Pichia pastoris*. *Applied and Environmental Microbiology* 77(11):3600-3608. DOI: 10.1128/AEM.02843-10.
- Ramakrishnan K, Prattipati 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. DOI: 10.1007/s00253-020-10785-8.
- Robainas-Del-Pino Y, Viader-Salvadó JM, Herrera-Estala AL, Guerrero-Olazarán M. 2023. Functional characterization of the *Komagataella phaffii* 1033 gene promoter and transcriptional terminator. *World Journal of Microbiology and Biotechnology* 39(9):246. DOI: 10.1007/s11274-023-03682-5.
- Romanos M. 1995. Advance in the use of *Pichia pastoris* for high level gene expression. *Current Opinion in Biotechnology* 6:527-533. DOI: 10.1016/0958-1669(95)80087-5.
- Romanos MA, Scorer CA, Clare JJ. 1992. Foreign gene expression in yeast: a review. *Yeast* 8:423-488.
- Sachs A, Sarnow P, Hentze MW. 1997. Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* 89:831-838.

Serrano-Rivero Y, Fando-Calzada K, Rafael MD. 2016. *Pichia pastoris*: una plataforma para la producción de proteínas heterólogas. CENIC Ciencias Biológicas 47:67–77.

Struhl K. 1989. Molecular mechanisms of transcriptional regulation in yeast. Annual Review of Biochemistry. 58:1051-77.

Sudbery PE, Gleeson MA, Veale RA, Ledebor AM, Zoetmulder MC. 1988. *Hansenula polymorpha* as a novel yeast system for the expression of heterologous genes. Biochemical Society Transactions 16:1081-1083. DOI: doi.org/10.1042/bst0161081a.

Tuller T, Ruppin E, Kupiec M. 2009. Properties of untranslated regions of the *S. cerevisiae* genome. BMC Genomics 10:1-11. DOI: doi.org/10.1186/1471-2164-10-391.

Van Helden J, del Olmo M, Pérez-Ortí JE. 2000. Statistical analysis of yeast genomic downstream sequences reveals putative polyadenylation signals. Nucleic Acids Research 28:1000-1010. DOI: 10.1093/nar/28.4.1000.

Viader-Salvadó JM, Cab-Barrera EL, Galán-Wong LJ, Guerrero-Olazarán M. 2006. Genotyping of recombinant *Pichia pastoris* strains 11:348-359. DOI: 10.2478/s11658-006-0029-z.

Viader-Salvadó JM, Castillo-Galván M, Fuentes-Garibay JA, Iracheta-Cárdenas MM, Guerrero-Olazarán M. 2013. Optimization of five environmental factors to increase beta-propeller phytase production in *Pichia pastoris* and impact on the physiological response of the host. Biotechnology Progress 29:1377-1385. DOI:10.1002/btpr.1822.

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(19):6423-6430. DOI 10.1128/AEM.00253-10.

Viader-Salvadó JM, Pentón-Piña N, Robainas-del-Pino Y, Fuentes-Garibay JA, Guerrero-Olazarán M. 2024. Effect of AOX1 and GAP transcriptional terminators on transcript

levels of both the heterologous and the GAPDH genes and the extracellular $Y_{p/x}$ in GAP promoter-based *Komagataella phaffii* strains. PeerJ 12:e18181 DOI 10.7717/peerj.18181.

Viera-Gomes AM, Souza-Carmo T, Silva-Carvalho L, Mendonça-Bahia F, Parachin NS. 2018. Comparison of Yeasts as Hosts for Recombinant Protein Production. Microorganisms 6:38. DOI: 10.3390/microorganisms6020038.

Vogl T, Gebbie L, Palfreyman RW, Speight R. 2018. Effect of Plasmid Design and Type of Integration Event on Recombinant Protein Expression in *Pichia pastoris*. Applied and Environmental Microbiology 84:2712-17. DOI: doi.org/10.1128/AEM.02712-17.

Vogl T, Glieder A. 2013. Regulation of *Pichia pastoris* promoters and its consequences for protein production. New Biotechnology. 30:385-404. DOI: 10.1016/j.nbt.2012.11.010.

Vogl T, Sturmberger L, Kickenweiz T, Wasmayer R, Schmid Ch, Hatzl AM, Gerstmann MA, Pitzer J, Wagner M, Thallinger GG, Geier M, Glieder A. 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. DOI: 10.1021/acssynbio.5b00199.

Waterham H R, Digan ME, Koutz P J, 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. DOI: 10.1016/S0378-1119(96)00675-0.

Wyers F, Rougemaille M, Badis G, Rousselle JC, Dufour ME, Boulay J, Régnault B, Devaux F, Namane A, Séraphin B, Libri D, Jacquier A. 2005. Cryptic Pol II transcripts are degraded by a nuclear quality control pathway involving a new Poly(A) Polymerase. Cell 121(5):725-737. DOI 10.1016/j.cell.2005.04.030.

Yogender P, Amardeep K, Mukherjee KJ. 2006. Process optimization of constitutive human granulocyte– macrophage colony-stimulating factor (hGM-CSF) expression in *Pichia pastoris* fed-batch culture. Biotechnological products and process engineering 69:650-657. DOI: 10.1007/s00253-005-0018-6.

Zepeda AB, Pessoa Jr A, Fariás JG. 2018. Carbon metabolism influenced for promoters and temperature used in the heterologous protein production using *Pichia pastoris* yeast. *Brazilian Journal of Microbiology* 49S:119-127. DOI: doi.org/10.1016/j.bjm.2018.03.010.

Zhang A L, Zhang T Y, Luo J X, Chen S C, Guan W J, Fu C Y, Peng S Q, Li H L. 2007. Constitutive expression of human angiostatin in *Pichia pastoris* by high-density cell culture. *Journal of Industrial Microbiology and Biotechnology* 34:117-122. DOI: 10.1007/s10295-006-0175-3.

Zhao J, Hyman L, Moore C. 1999. Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis. *Microbiology* 63:405-445. DOI: 10.1128/mubr.63.2.405-445.1999.

ANEXOS

Anexo 1. Metodología de desarrollo del qPCR para determinación de dosis génica.

Planificación Miras problemas						
En este esquema en el coctel son 4 uL de ADN y 2 uL de agua para 20 uL totales						
Cocxel para dosis génica						
			Tubos		Réplicas	Total Tubos
DNÁreferencia	DNÁref	puntos FTEII	1	X	2	2
	DNÁref	puntos GAP	1	X	2	2
		Negativo FTEII	0.5	X	1	0.5
		Negativo GAP	0.5	X	1	0.5
DNÁproblema	DNÁprob	FTEII	7	X	2	14
	DNÁprob	GAP	7	X	2	14
	Cruzando dedos	Sin testigo positivo				
		Tubo extra	4	X	1	4
			TOTAL			37

ETAPA 1		Comunes a todo	Master Mix	Vol por reacción		Vol Total
				uL	uL	
			agua	10	2	370
						74
				TOTAL		444
	Regalo	2	4.50	4.50	4.50	4.50
						TOTAL
						18.00

ETAPA 2		4 tubos				TOTAL
		Clona 1 TTGAP(1 ng)	Clona 2 TTGAP(1 ng)	Clona 3 TTGAP(1 ng)	Tubo C+ TTAOX1 (1 ng)	
	Comunes a todo	540	540	540	540	216.0
	DNÁ mues tra/ref	18.0	18.0	18.0	18.0	
	Água	-	-	-	-	
	TOTAL	72.0	72.0	72.0	72.0	288.0
	Regalo	1	4.25	4.25	4.25	4.25
						17.00

ETAPA 3		Cada tubo dividir en 2 tubos			
		Dividir en viales Prime Time FTEII o GAP	340	340	340
		8.5	8.5	8.5	8.5
	TOTAL	42.5	42.5	42.5	42.5

ETAPA 4	Réplicas en viales de qPCR	Cada vial dividir en 2 tubos			
		Réplica 1	20.0	20.0	20.0
	Réplica 2	20.0	20.0	20.0	20.0
	TOTAL	40.0	40.0	40.0	40.0

Anexo 2. Datos primarios de densidad celular (DC) y rendimiento proteína extracelular-biomasa a las 24 h de cultivo de las 18 clonas de la cepa KM71/P_{GAP}-FTEII-T_{AOX1}.

No	Colonia	DC (g/L)	Conc. Prot. Total (µg/ml)	Rend. Prot. Total/DC (mg/g)
1	F3	8.22	7.46	0.9
2	F15	8.27	8.11	1.0
3	F1	8.01	8.08	1.0
4	F13	7.95	10.85	1.4
5	F24	9.30	36.75	4.0
6	F10	7.92	31.65	4.0
7	F2	8.83	36.45	4.1
8	F11*	8.00	33.15	4.1
9	F4	8.59	35.65	4.2
10	F6	8.77	36.50	4.2
11	F8	8.60	35.75	4.2
12	F16	7.22	30.45	4.2
13	F30	7.45	31.60	4.2
14	F14	7.24	30.90	4.3
15	F29	7.37	32.90	4.5
16	F9	7.25	33.40	4.6
17	F22	7.54	35.25	4.7
18	C+	8.50	57.90	6.8

C+: Clona multicopia de la cepa KM71GAHFTEII (Herrera-Estala A. et al. 2020). *. Clona F11 (clona unicopia denotada también como CN11 en Tabla 24. Los valores resaltados con diferente color o en letra bold enmarcan el rango de valores [mínimo-máximo] de cada variable representados en la Figura 30.

Anexo 3. Datos primarios de densidad celular y rendimiento proteína extracelular-biomasa a las 24 h de cultivo de la cepa KM71/P_{GAP}-FTEII-T_{GAP}.

No.	Colonia	DC (g/L)	Conc. Prot. Total (µg/ml)	Rend. Prot. Total/DC (mg/g)
1	5	7.39	13.56	1.8
2	27	8.15	18.62	2.3
3	2*	9.03	21.60	2.4
4	11	8.57	20.56	2.4
5	10	8.16	19.69	2.4
6	19	8.53	21.30	2.5
7	36	8.28	21.10	2.6
8	17	7.75	20.13	2.6
9	14	8.67	22.60	2.6
10	6	8.08	21.10	2.6
11	28	9.04	24.30	2.7
12	23	8.90	25.00	2.8
13	34	7.64	21.70	2.8
14	4*	8.09	23.50	2.9
15	32	8.06	24.35	3.0
16	26	8.42	25.70	3.1
17	31	8.99	27.55	3.1
18	33	8.53	26.70	3.1
19	3	8.07	25.70	3.2
20	29	7.76	24.75	3.2
21	21	8.89	31.30	3.5
22	12	8.33	31.40	3.8
23	1	9.23	36.30	3.9
24	9	8.07	41.10	5.1
25	8**	7.87	50.80	6.5

*. Clonas 2 y 4 [clonas unicopias denotadas también como CN2 y CN4 en Tabla 24 y en el reporte de la revista PeerJ (Viader-Salvadó et al, 2024)]. **. Clona 8 [clona multicopia denotadas también como CN8 en Tabla 24]. Los valores resaltados con diferente color o en letra bold enmarcan el rango de valores [mínimo-máximo] de cada variable representados en la Figura 30.

Anexo 4. Análisis estadístico de la densidad celular de 18 clonas de la cepa KM71/P_{GAP}-FTEII-T_{AOX1} para determinar si la muestra es representativa de la población.

Análisis de Una Variable - DC KM71/P_{GAP}-FTEII-T_{AOX1}

Datos/Variable: DC_KM71/P_{GAP}-FTEII -T_{AOX1} (g/L)

18 valores con rango desde 7.22 a 9.30

El StatAdvisor

Este procedimiento está diseñado para resumir una sola muestra de datos. Calculará varios estadísticos y gráficas. También incluidos en el procedimiento están los intervalos de confianza y las pruebas de hipótesis. Use los botones de Opciones Tabulares y de Opciones Gráficas en la barra de herramientas para análisis, para acceder a estos diferentes procedimientos.

Resumen Estadístico para DC_KM71/P_{GAP}-FTEII-T_{AOX1}

Recuento	18
Promedio	8.05722
Desviación Estándar	0.625947
Coefficiente de Variación	7.76877%
Mínimo	7.22
Máximo	9.3
Rango	2.08
Sesgo Estandarizado	0.376426
Curtosis Estandarizada	-0.763301

El StatAdvisor

Esta tabla muestra los estadísticos de resumen para DC_KM71/P_{GAP}-FTEII-T_{AOX1}. Incluye medidas de tendencia central, medidas de variabilidad y medidas de forma. De particular interés aquí son el sesgo estandarizado y la curtosis estandarizada, las cuales pueden utilizarse para determinar si la muestra proviene de una distribución normal. Valores de estos estadísticos fuera del rango de -2 a +2 indican desviaciones significativas de la normalidad, lo que tendería a invalidar cualquier prueba estadística con referencia a la desviación estándar. En este caso, el valor del sesgo estandarizado se encuentra dentro del rango esperado para datos provenientes una distribución normal. El valor de curtosis estandarizada se encuentra dentro del rango esperado para datos provenientes de una distribución normal.

Intervalos de Confianza para DC_KM71/P_{GAP}-FTEII-T_{AOX1}

Intervalos de confianza del 95.0% para la media: 8.05722 +/- 0.311277 [7.74595, 8.3685]

Intervalos de confianza del 95.0% para la desviación estándar: [0.469703, 0.938384]

El StatAdvisor

Este panel muestra los intervalos de confianza del 95.0% para la media y la desviación estándar de DC_KM71/P_{GAP}-FTEII-T_{AOX1}. La interpretación clásica de estos intervalos es que, en muestreos repetidos, estos intervalos contendrán la media verdadera ó la desviación estándar verdadera de la población de la que fueron extraídas las muestras, el

95.0% de las veces. En términos prácticos, puede establecerse con 95.0% de confianza, que la media verdadera de DC_KM71/P_{GAP}-FTEII-T_{AOXI} se encuentra en algún lugar entre 7.74595 y 8.3685, en tanto que la desviación estándar verdadera está en algún lugar entre 0.469703 y 0.938384.

Ambos intervalos asumen que la población de la cual proviene la muestra puede representarse por la distribución normal. Mientras que el intervalo de confianza para la media es bastante robusto y no muy sensible a violaciones de este supuesto, los intervalos de confianza para la desviación estándar son muy sensibles. Si los datos no provienen de una distribución normal, el intervalo para la desviación estándar puede ser incorrecto. Para verificar si los datos provienen de una distribución normal, seleccione Resumen Estadístico de la lista de Opciones Tabulares, ó escoja Gráfica de Probabilidad Normal de la lista de Opciones Gráficas.

Anexo 5. Análisis estadístico de la densidad celular de $KM71/P_{GAP-FTEII-T_{AOX1}}$ para determinar si los datos ajustan a una distribución normal.

Ajuste de Datos No Censurados - DC $KM71/P_{GAP-FTEII-T_{AOX1}}$

Datos/Variable: DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$ (g/L)

18 valores con rango desde 7.22 a 9.30

Distribuciones Ajustadas

<i>Normal</i>
medios = 8.05722
desviación estándar = 0.625947

El StatAdvisor

Este análisis muestra los resultados de ajustar una distribución normal a los datos de DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$. Los parámetros estimados para la distribución ajustada se muestran arriba. Se puede evaluar si la distribución normal ajusta los datos adecuadamente, seleccionando Pruebas de Bondad de Ajuste de la lista de Opciones Tabulares. También puede evaluarse visualmente que tan bien la distribución normal se ajusta, seleccionando Histogramas de Frecuencia de la lista de Opciones Gráficas. Otras opciones dentro el procedimiento permiten calcular y desplegar áreas de colas y valores críticos para la distribución. Para seleccionar una distribución diferente, presione el botón secundario del ratón y seleccione Opciones de Análisis.

Pruebas de Normalidad para DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$

<i>Prueba</i>	<i>Estadístico</i>	<i>Valor P</i>
Estadístico W de Shapiro-Wilk	0.945782	0.365622

El StatAdvisor

Esta ventana muestra los resultados de diversas pruebas realizadas para determinar si DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$ puede modelarse adecuadamente con una distribución normal. La prueba de Shapiro-Wilk está basada en la comparación de los cuartiles de la distribución normal ajustada a los datos.

Debido a que el valor-P más pequeño de las pruebas realizadas es mayor ó igual a 0.05, no se puede rechazar la idea de que DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$ proviene de una distribución normal con 95% de confianza.

Pruebas de Bondad-de-Ajuste para DC_ $KM71/P_{GAP-FTEII-T_{AOX1}}$

Prueba de Kolmogorov-Smirnov

	<i>Normal</i>
DMAS	0.129018
DMENOS	0.0936671
DN	0.129018
Valor P	0.925425

El StatAdvisor

Esta ventana muestra los resultados de diversas pruebas realizadas para determinar si $DC_KM71/P_{GAP-FTEII-T_{AOX1}}$ puede modelarse adecuadamente con una distribución normal.

Debido a que el valor-P más pequeño de las pruebas realizadas es mayor ó igual a 0.05, no se puede rechazar la idea de que $DC_KM71/P_{GAP-FTEII-T_{AOX1}}$ proviene de una distribución normal con 95% de confianza.

Anexo 6. Análisis estadístico de la densidad celular de $KM71/P_{GAP}$ -FTEII- T_{GAP} para determinar si la muestra es representativa de la población.

Análisis de Una Variable - DC $KM71/P_{GAP}$ -FTEII- T_{GAP}

Datos/Variable: DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} (g/L)

25 valores con rango desde 7.39 a 9.23

El StatAdvisor

Este procedimiento está diseñado para resumir una sola muestra de datos. Calculará varios estadísticos y gráficas. También incluidos en el procedimiento están los intervalos de confianza y las pruebas de hipótesis. Use los botones de Opciones Tabulares y de Opciones Gráficas en la barra de herramientas para análisis, para acceder a estos diferentes procedimientos.

Resumen Estadístico para DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP}

Recuento	25
Promedio	8.34
Desviación Estándar	0.489966
Coefficiente de Variación	5.87489%
Mínimo	7.39
Máximo	9.23
Rango	1.84
Sesgo Estandarizado	0.191232
Curtosis Estandarizada	-0.788524

El StatAdvisor

Esta tabla muestra los estadísticos de resumen para DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} . Incluye medidas de tendencia central, medidas de variabilidad y medidas de forma. De particular interés aquí son el sesgo estandarizado y la curtosis estandarizada, las cuales pueden utilizarse para determinar si la muestra proviene de una distribución normal. Valores de estos estadísticos fuera del rango de -2 a +2 indican desviaciones significativas de la normalidad, lo que tendería a invalidar cualquier prueba estadística con referencia a la desviación estándar. En este caso, el valor del sesgo estandarizado se encuentra dentro del rango esperado para datos provenientes una distribución normal. El valor de curtosis estandarizada se encuentra dentro del rango esperado para datos provenientes de una distribución normal.

Intervalos de Confianza para DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP}

Intervalos de confianza del 95.0% para la media: 8.34 +/- 0.202248 [8.13775, 8.54225]

Intervalos de confianza del 95.0% para la desviación estándar: [0.382579, 0.681617]

El StatAdvisor

Este panel muestra los intervalos de confianza del 95.0% para la media y la desviación estándar de DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} . La interpretación clásica de estos intervalos es que, en muestreos repetidos, estos intervalos contendrán la media verdadera ó la

desviación estándar verdadera de la población de la que fueron extraídas las muestras, el 95.0% de las veces. En términos prácticos, puede establecerse con 95.0% de confianza, que la media verdadera de $DC_KM71/P_{GAP}-FTEII-T_{GAP}$ se encuentra en algún lugar entre 8.13775 y 8.54225, en tanto que la desviación estándar verdadera está en algún lugar entre 0.382579 y 0.681617.

Ambos intervalos asumen que la población de la cual proviene la muestra puede representarse por la distribución normal. Mientras que el intervalo de confianza para la media es bastante robusto y no muy sensible a violaciones de este supuesto, los intervalos de confianza para la desviación estándar son muy sensibles. Si los datos no provienen de una distribución normal, el intervalo para la desviación estándar puede ser incorrecto. Para verificar si los datos provienen de una distribución normal, seleccione Resumen Estadístico de la lista de Opciones Tabulares, ó escoja Gráfica de Probabilidad Normal de la lista de Opciones Gráficas.

Anexo 7. Análisis estadístico de la densidad celular de $KM71/P_{GAP}$ -FTEII- T_{GAP} para determinar si los datos ajustan a una distribución normal.

Ajuste de Datos No Censurados - DC $KM71/P_{GAP}$ -FTEII- T_{GAP}

Datos/Variable: DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} (g/L)

25 valores con rango desde 7.39 a 9.23

Distribuciones Ajustadas

<i>Normal</i>
Medios = 8.34
desviación estándar = 0.489966

El StatAdvisor

Este análisis muestra los resultados de ajustar una distribución normal a los datos de DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} . Los parámetros estimados para la distribución ajustada se muestran arriba. Se puede evaluar si la distribución normal ajusta los datos adecuadamente, seleccionando Pruebas de Bondad de Ajuste de la lista de Opciones Tabulares. También puede evaluarse visualmente que tan bien la distribución normal se ajusta, seleccionando Histogramas de Frecuencia de la lista de Opciones Gráficas. Otras opciones dentro el procedimiento permiten calcular y desplegar áreas de colas y valores críticos para la distribución. Para seleccionar una distribución diferente, presione el botón secundario del ratón y seleccione Opciones de Análisis.

Pruebas de Normalidad para DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP}

<i>Prueba</i>	<i>Estadístico</i>	<i>Valor P</i>
Estadístico W de Shapiro-Wilk	0.967876	0.600525

El StatAdvisor

Esta ventana muestra los resultados de diversas pruebas realizadas para determinar si DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} puede modelarse adecuadamente con una distribución normal. La prueba de Shapiro-Wilk está basada en la comparación de los cuartiles de la distribución normal ajustada a los datos.

Debido a que el valor-P más pequeño de las pruebas realizadas es mayor ó igual a 0.05, no se puede rechazar la idea de que DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP} proviene de una distribución normal con 95% de confianza.

Pruebas de Bondad-de-Ajuste para DC_ $KM71/P_{GAP}$ -FTEII- T_{GAP}

Prueba de Kolmogorov-Smirnov

	<i>Normal</i>
DMAS	0.123331
DMENOS	0.109181
DN	0.123331
Valor P	0.84149

El StatAdvisor

Esta ventana muestra los resultados de diversas pruebas realizadas para determinar si $DC_KM71/P_{GAP}-FTEII-T_{GAP}$ puede modelarse adecuadamente con una distribución normal.

Debido a que el valor-P más pequeño de las pruebas realizadas es mayor ó igual a 0.05, no se puede rechazar la idea de que $DC_KM71/P_{GAP}-FTEII-T_{GAP}$ proviene de una distribución normal con 95% de confianza.

Anexo 8. Análisis de distribución de frecuencia de las regiones 3'UTR anotadas para 3244 regiones 3' UTR anotadas en la base de datos del GenBank para *K. phaffii*.

