

Diseño de cebadores a partir del análisis bioinformático de secuencias de bacterias y hongos productores de la enzima Endoxilanas presentes en compostaje de Higuierilla.

**Alejandra Osorio González
July Marlody Tapiero Agudelo**

Trabajo de tesis para optar al título de Bacteriólogas

**Asesor
Narmer Fernando Galeano Vanegas
Investigador Facultad Ciencias de la Salud
Microbiólogo**

**Universidad Católica de Manizales
Facultad de Ciencias de la Salud
Programa de Bacteriología
2014**

Diseño de cebadores a partir del análisis bioinformático... Osorio A., Tapiero J.

Diseño de cebadores a partir del análisis bioinformático de secuencias de bacterias y hongos productores de la enzima Endoxilanas presentes en compostaje de Higuierilla.

**Alejandra Osorio González
July Marlody Tapiero Agudelo**

Trabajo de tesis para optar al título de Bacteriólogas

**Universidad Católica de Manizales
Facultad de Ciencias de la Salud
Programa de Bacteriología
2014**

NOTA DE ACEPTACIÓN

El trabajo de grado titulado “**Diseño de cebadores a partir del análisis bioinformático de secuencias de bacterias y hongos productores de la enzima Endoxilanas presentes en compostaje de Higuierilla,**” de los autores July Tapiero Agudelo y Alejandra Osorio González, cumple con los requisitos para optar a el título de bacterióloga.

Nota

Firma Directora del programa

Firma Evaluador

AGRADECIMIENTOS

A Dios por permitirnos culminar con éxito nuestra carrera.

A nuestras familias por su amor, su esfuerzo constante durante nuestra carrera, su comprensión, su ayuda, apoyo incondicional y por creer en nosotras en cada momento de esta etapa importante para nuestras vidas.

A los profesores por los conocimientos brindados durante nuestra etapa universitaria.

A los profesores Narmer Fernando Galeano, Jhon Jairo Henao y Javier Guillermo Mantilla Afanador por las asesorías brindadas, los consejos y acompañamiento en el desarrollo de esta investigación.

A la directora del programa Gloria Inés Estrada por el facilitarnos los equipos de la Universidad Católica de Manizales.

A la doctora Olga lucia Tovar Aguirre y Jennifer Gaviria quien creyó en nuestras capacidades y nos impulsó a continuar.

A nuestros compañeros y amigos por los gratos momentos compartidos y de manera muy especial a Katherine Ponce Loaiza Yised Arango Vargas y Oscar Gallego López.

TABLA DE CONTENIDO

	Pág.
Resumen	10
1. INTRODUCCIÓN	11
2. OBJETIVOS	12
2.1. Objetivo General	
2.2. Objetivos específicos	
3. MARCO TEÓRICO	13
3.1 Biodiésel	14
3.1.1 Biodiesel de primera generación	14
3.1.2 Biodiesel de segunda generación o Biocombustibles	15
3.2 Metagenómica	16
3.3 Compostaje	17
3.4 Lignocelulosa	17
3.4.1 Hemicelulosa	18
3.4.2 Lignina	19
3.4.3 Celulosa	19
3.4.4 Degradación de la lignocelulosa	19
3.4.4.1 Tratamiento Biológico	20
3.4.4.2 Xilanasas	20
3.4.4.3 Tratamiento Químico	22
3.5 Diseño de cebadores	23
3.5.1 Propiedades de oligonucleótidos que influyen en la eficiencia de la amplificación	24
3.6 MEME versión 4.4.0	27
3.7 MAST versión 4.9.0	27

3.8 CODEHOP	27
4. METODOLOGÍA PROPUESTA	27
5. RESULTADOS	29
6. DISCUSIÓN	43
7. CONCLUSIONES	46
8. PERSPECTIVAS	46
ANEXOS	47
LITERATURA CITADA	53
LITERATURA CONSULTADA	55

LISTA DE TABLAS

	Pág.
Tabla 1. Secuencia enzimática de la Endoxilanasasa	30
Tabla 2. Motivos de secuencias	30
Tabla 3. Resultados de MAST- UniProt	36
Tabla 3.1 Arquitecturas conservadas	
Tabla 4. Resultados de MAST- Uniref100	37
Tabla 4.1 Arquitecturas conservadas	
Tabla 5. Diseño de cebadores para el motivo 8	39
Tabla 6. Cebador Sentido diseñado para el motivo 8	40
Tabla 7. Alineamiento del cebador con la secuencia	40
Tabla 8. Cebador Antisentido diseñado para el motivo 8	40
Tabla 9. Alineamiento del cebador con la secuencia	41
Tabla 10. Diseño de cebadores para el motivo 9	42
Tabla 11. Cebador Sentido diseñado para el motivo 9	42
Tabla 12. Cebador Antisentido diseñado para el motivo 9	43
Tabla 13. Alineamiento del cebador con la secuencia	43

LISTA DE FIGURAS

	Pág.
Figura 1. Diagrama de flujo metodología.	29
Figura 2.	34
A. Paleta de colores.	
B. Arquitectura más conservada.	
C. Motivo 9 el más conservado al extremo 5'.	
D. Motivo 8 el más conservado al extremo 3'.	
Figura 3. Análisis de motivos y hallazgos de arquitecturas mas comunes	36
Figura 4. Arquitecturas conservadas UniRef100	38
Figura 5. Representación en logos del motivo 8 y su respectiva expresión regular	39
Figura 6. Representación en logos del motivo 9 y su respectiva expresión regular	41

ANEXOS

	Pág.
Anexo 1. Microorganismos productores de enzimas lignocelulolíticas	47
Anexo 2. Figura 3	48
Anexo 3. Figura 4	49
Anexo 4. Formato de bloques motivo 8	50
Anexo 5. Formato de bloques motivo 9	51

Diseño de cebadores a partir del análisis bioinformático de secuencias de bacterias y hongos productores de la enzima Endoxilanasas presentes en compostaje de Higuierilla.

RESUMEN

El *Ricinus communis* es una especie que pertenece a la familia Euphorbiaceae y es conocida por producir el aceite de ricino, actualmente tiene diversos usos industriales, en la producción de plásticos, cosméticos, pinturas y biocombustibles, del mismo modo conforme es explotada genera considerables residuos orgánicos que pueden ser aprovechados a nivel biotecnológico en la obtención de enzimas lignocelulolíticas. Para lograr esto, una de las alternativas es recurrir a la bioinformática y a la biología molecular. En este trabajo se realizó la búsqueda de secuencias que codifican para la enzima Endoxilanasas en bacterias, hongos y plantas empleando la base de datos nr del NCBI. Estas secuencias fueron utilizadas como base de datos de entrenamiento para el software MEME, con el fin de descubrir los motivos de la enzima, los perfiles se validaron en el MAST. Las secuencias fueron sometidas a filtros sucesivos para organizarlas y determinar de manera sencilla los motivos más conservados, los cuales se usaron para el diseño de cebadores en el programa CODEHOP.

Como resultado de la búsqueda se obtuvieron 5.239 secuencias, tras la corrida del MEME se obtuvieron 111, de estas 59 eran de bacterias, 29 de hongos y 14 de plantas, se hallaron 20 motivos, el motivo 8 y el motivo 9 se identificaron como los más conservados, seguidamente se realizó la validación con el MAST. Se obtuvieron cebadores 2 para el motivo 8 y uno para el motivo 9.

Los perfiles validados arrojaron buenos resultados, lo que indica que los motivos usados en el diseño de los cebadores son los indicados.

1 INTRODUCCIÓN

En varias regiones del país, incluyendo el Departamento de Caldas, se están llevando a cabo proyectos que posibilitan el cultivo de higuierilla, la creación de empresas como Higueroil es evidencia de ello, pues ha desarrollado cultivos experimentales en diferentes regiones del país de forma exitosa, entre ellas Casanare, Arauca, Huila, Tolima, Risaralda, Caldas, Antioquia y Córdoba, Colombiana de BIOCOMBUSTIBLES S.A es otra de las empresas que impulsa el cultivo de Higuierilla con el fin de producir aceite y biodiesel de alta calidad.

A nivel industrial el aceite de ricino es ampliamente utilizado, refinado sirve en el área farmacéutica, en bruto se utiliza como emulsificante para desinfectantes y pesticidas. Es importante mencionar que la torta de higuierilla no se utiliza como alimento animal debido a la presencia de toxinas y de sustancias alergénicas, su empleo se limita especialmente a aplicaciones de fertilizantes orgánicos. Por estas razones, el cultivo de higuierilla ha tomado gran importancia económica y ambiental, pues al igual que es explotado por su gran número de aplicaciones, también produce gran cantidad de residuos orgánicos que generalmente son tratados con ácidos, lo cual contamina aún más el entorno, aunque existen métodos enzimáticos para el tratamiento de estos residuos, actualmente no están disponibles en el país por su alto costo, motivos por los cuales se hace necesario su aprovechamiento. Existen evidencias importantes donde se menciona la producción de bioetanol a partir de material lignocelulósico, las ventajas y desventajas de los métodos de transformación de la biomasa (Sánchez et.al, 2010), y se describe la importancia de las enzimas lignocelulolíticas como método de tratamiento, sus aplicaciones y el crecimiento de la demanda en diversas industrias (Bhat MK, 2000), de manera que el interés por la implementación de metodologías enzimáticas crece día a día, brindando la posibilidad a las industrias de mejorar el manejo de los residuos lignocelulósicos y a su vez favorecer el medio ambiente.

Teniendo en cuenta una de las alternativas para la obtención de la tecnología enzimática, la biología molecular y la bioinformática se convierten en herramientas fundamentales para su implementación, al igual que la metagenómica, pues permiten estudiar de manera mucho más sencilla y eficiente gran cantidad de información genética, vital en la experimentación y análisis sobre métodos enzimáticos.

Por lo tanto estas disciplinas juegan un papel crucial en este trabajo, pues se planteó el diseño de cebadores a partir del análisis bioinformático de secuencias de bacterias, hongos y plantas productores de la enzima Endoxilanasas presentes en compostaje de Higuierilla, con el fin de sintetizar los cebadores que posteriormente permitirán la obtención del gen de la enzima a través de la PCR inversa pre-amplificada (PAI-PCR).

2 OBJETIVOS

2.1 Objetivo General

Diseñar cebadores a partir del análisis bioinformático de secuencias de bacterias y hongos productores de la enzima Endoxilanasas presentes en compostaje de Higuierilla.

2.2 Objetivos específicos

- Realizar la búsqueda de secuencias que codifican para la Endoxilanasas en bacterias y hongos en la base de datos nr del NCBI.
- Obtener perfiles proteicos de las secuencias de la Endoxilanasas empleando el software MEME.
- Determinar los motivos y arquitecturas más conservadas a partir de la salida del MEME.

- Diseñar los cebadores con los motivos más conservados en el programa CODEHOP.

3 MARCO TEÓRICO

La higuera es originaria de África tropical (Abisinia) y posiblemente de la India; extensamente introducida en las regiones cálidas y templadas cálidas de todo el mundo. Se da en un clima tropical a subtropical, con un período libre de heladas de por lo menos 5 meses y precipitaciones superiores a los 700 mm al año.

El ricino se cultiva en general como planta anual, a pesar de tratarse de una especie perenne. Requiere una buena preparación del terreno previo a su siembra, que puede ser realizada a mano o mediante sembradoras de maíz o algodón.

El fruto, es una cápsula trilocular que contiene una semilla por lóculo, exteriormente está recubierto por espinas no punzantes. La semilla es oval, de tamaño variable entre 5 y 20 mm según variedades; estas contienen aceite fijo (*oleum ricini*) en porcentaje del 35 al 55% principalmente constituido por los glicéridos de los ácidos ricinoleico, iso-ricinoleico, etc.; también ricina y ricinina, la primera es una fitotoxina sumamente venenosa por vía endovenosa y menor por vía oral, aunque también esta última vía puede ocasionar la muerte; su actividad desaparece por acción del calor moderado; el segundo es un alcaloide de fórmula $C_8H_8N_2O_2$ que es utilizado en la producción de biodiésel, según estimaciones hechas por el Ingeniero Sergio Quemé de la Dirección General de Energía, los rendimientos de aceite de higuera, podrían alcanzar hasta unos 300 galones por hectárea al año, en plantaciones manejadas, lo que proporcionaría unos 270 galones de biodiésel por hectárea al año. [1]

3.1 Biodiésel

El biodiésel es un combustible renovable que se puede obtener principalmente a partir de aceites vegetales, animales, así como de aceites reciclados.

El biodiésel tiene una ventaja ecológica, en comparación con el diésel de origen fósil, ya que reduce las emisiones de gases que provocan el efecto de invernadero. Así, por ejemplo, el uso de una tonelada de Biodiésel, evita la producción de 2.5 toneladas de dióxido de carbono (CO₂) y sobre todo elimina, si se usa el biodiésel sólo en los motores, las emisiones de azufre (SO₂) del diésel, evitando las lluvias ácidas; además, lo que es fundamental: es un combustible renovable y no finito como los hidrocarburos. [2]

3.1.1 Biodiesel de primera generación

Algunos de los insumos son de procedencia agrícola y están conformados por las partes alimenticias de las plantas, las cuales tienen un alto contenido de almidón, azúcares y aceites. Ejemplos de estas materias son el jugo de la caña de azúcar, granos de maíz, jugo de la remolacha o betabel, aceite de semilla de girasol, aceite de soya, aceite de palma, aceite de ricino, aceite de semilla de algodón, aceite de coco, aceite de maní o cacahuate, entre otros. También se emplean como insumos a las grasas animales, grasas y aceites de desecho provenientes de la cocción y elaboración de alimentos, y desperdicios sólidos orgánicos.

Los biocombustibles son producidos empleando tecnología convencional como la fermentación (para azúcares y carbohidratos), trans-esterificación (para los aceites y grasas), y la digestión anaerobia (para los desperdicios orgánicos). De estos procesos se obtiene etanol, metanol y n-butanol (a partir de azúcares), biodiesel (a partir de los aceites), y biogás (mezcla de metano y anhídrido

carbónico, también conocidos como gas natural y dióxido de carbono respectivamente, obtenida a partir de los desperdicios orgánicos).

Las ventajas de estos biocombustibles son su facilidad de procesamiento, sus bajas emisiones de gases de efecto invernadero (excepto en el caso del maíz, donde el balance de estas emisiones es casi nulo) y un balance positivo en dichas emisiones, pero tiene como desventaja el desvío de recursos alimenticios hacia la producción de energéticos. [3]

3.1.2 Biodiesel de segunda generación o Biocombustibles

Los insumos son residuos agrícolas y forestales compuestos principalmente por celulosa. Ejemplos de ellos son el bagazo de la caña de azúcar, el rastrojo de maíz (tallos, hojas y olote), paja de trigo, aserrín, hojas y ramas secas de árboles, etcétera.

Los procesos de producción tienen un nivel de complejidad más alto que los de primera generación, y como ejemplos destacan la sacarificación, fermentación y el proceso Fischer-Tropsch. Este último proceso también recibe los nombres de proceso GTL y proceso BTL, cuyas siglas en inglés provienen de "Gas-To-Liquids" y "Biomass-To-Liquids" respectivamente, los cuales consisten en la gasificación del carbón y de la materia lignocelulósica de la biomasa, para después sintetizar algún combustible líquido como el etanol.

Mediante los procesos de segunda generación se fabrica etanol, metanol, gas de síntesis (mezcla de anhídrido carbonoso, mejor conocido como monóxido de carbono, e hidrógeno), biodiesel, 2.5-dimetilfurano (DMF), entre otros.

La ventaja principal en la producción de estos biocombustibles es la inexistencia de desviaciones de alimentos provenientes de la agricultura hacia el sector energético, pero su desventaja es la poca ganancia en disminución de las emisiones de gases de efecto

invernadero durante el procesamiento de los insumos, respecto a los biocombustibles de primera generación. [3]

3.2 Metagenómica

La metagenómica es una "metodología emergente para el análisis genómico directo de una comunidad mixta de microorganismos no cultivables". Los análisis de datos existentes de metagenómica dependen en gran parte del aparato computacional diseñado originalmente para los proyectos de genómica microbiana. El desafío de armar secuencias de metagenómicas se deriva principalmente de la dificultad de estudiar especies de determinada población. Alternativamente, el microorganismo se estudia, se busca directamente en bases de datos los genes que se conocen (o proteínas) para identificar homólogos de progresión. La aproximación final puede tener una baja sensibilidad y especificidad en la identificación de secuencias homólogas, lo que puede afectar la gran variedad de análisis sucesivos. [4]

3.3 Compostaje

Proceso biológico oxidativo de residuos, provocado y controlado por microorganismos de ambiente. [5]

"La descomposición microbiológica aeróbica de residuos orgánicos de distinta procedencia en condiciones de temperatura, humedad y aireación controladas, conduce a la obtención de un producto llamado Compost o abono compuesto. (Climent, 1996)

Se distinguen las siguientes fases en el proceso de compostaje:

- **Fase inicial:** La temperatura incrementa hasta 60°C, evidenciándose un aumento de pH, hay 45 millones/g de microorganismos termofílicos, hay destrucción de microorganismos patógenos y se puede apreciar el crecimiento de hongos termofílicos. Dura entre 50 y 60 horas.

- **Segunda fase:** Aumento de temperatura la cual puede alcanzar los 70°C, hay disminución de la biomasa, destrucción de microorganismos patógenos y desaparición de hongos y bacterias termofílicas.
- **Tercera fase:** Última etapa termofílica, se presenta menor producción de calor, la temperatura disminuye, hay desarrollo de microorganismos termofílicos, los cuales atacan las sustancias lignocelulósicas.
- **Fase de maduración:** Se da a temperatura ambiente, hay desaparición de microorganismos termófilos y patógenos, aparecen los mesófilos y algunas sustancias antibióticas. [5]

3.4 Lignocelulosa

La lignocelulosa (celulosa, hemicelulosa y lignina) es el principal y más abundante componente de la biomasa producida por la fotosíntesis, anualmente se forman 200,000 millones de toneladas en el mundo. La pared celular de las plantas está formada por lignocelulosa, la composición y porcentajes de los polímeros varían entre las especies de plantas, incluso entre la edad y la etapa de crecimiento.

Existen dos tipos de sistemas enzimáticos extracelulares: los que producen hidrolasas que degradan la celulosa (celulasas) y la hemicelulosa (hemicelulasas) y los que despolimerizan la lignina por reacciones de oxidación (peroxidasas y lacasas)

Las propiedades químicas de los componentes de la lignocelulosa hacen de estos, un sustrato de enorme valor biotecnológico. Grandes cantidades de "desechos" lignocelulósicos se genera a través de las prácticas forestales y agrícolas, las industrias de papel, industrias de la madera y muchas agroindustrias, plantean un problema de contaminación del medio ambiente. Desgraciadamente, muchos de los desechos de lignocelulosa se disponen a menudo por la quema de biomasa, que no se limita a los países

en desarrollo solamente, ya que se considera un fenómeno global. Sin embargo, las enormes cantidades de biomasa vegetal residual considerados como "residuos" pueden convertirse en diversos productos de valor agregado como los biocombustibles, productos químicos, fuentes de energía barata para la fermentación, alimentos para animales y mejora de nutrientes humanos. Las enzimas lignocelulolíticas son también importantes en aplicaciones potenciales en diversas industrias, incluyendo productos químicos, combustibles, alimentos, cerveza y vino, alimentos para animales, textiles y ropa, pasta y papel, y la agricultura. [6].

3.4.1 Hemicelulosa

La hemicelulosa es un polisacárido complejo que está compuesto en su mayoría de pentosas (D-xilosa y L-arabinosa) y hexosas (D-glucosa, D-manosa y D-galactosa), son cadenas ramificadas acelitadas y conformados también por 4-O-metilglucurónico, D-galacturónico y ácido D-glucurónico, los cuales se unen mediante enlaces β -1,4 y a veces β -1,3 enlaces glucosídicos. [7]

También son considerados componente de la hemicelulosa los xilanos, los xiloglucanos, mananos, glucomananos y glucanos, los cuales se unen mediante enlaces β -1,3 o β -1,4. En cuanto a los mananos y los galactomananos, estos poseen una estructura β -1,4 unidos a residuos de manosa. Existen ciertas diferencias estructurales entre las hemicelulosas de acuerdo a la diversidad de especies y tipo de células, incluyendo los mismos individuos. [7]

La función más importante de la hemicelulosa hace referencia al mantenimiento de las fibras de lignina y celulosa unidas, dando así rigidez a la malla de celulosa-hemicelulosa-lignina, estas dos últimas están unidas por enlaces éster. [7]

3.4.2 Lignina

Polímero amorfo, tridimensional y ramificado, está compuesto por alcoholes aromáticos los cuales dan soporte estructural e impermeabilidad a la celulosa y hemicelulosa, esto la hace muy resistente a la degradación biológica y química. [7]

3.4.3 Celulosa

Polímero constituido por D-glucosa, está unido por enlaces glucosídicos β -1,4, se estructuran en cadenas lineales llamadas microfibrillas enlazadas por puentes de hidrogeno y fuerzas de van der Waals, formando una estructura cristalina resistente a la hidrólisis y regiones amorfas susceptibles a la degradación enzimática. [7]

3.4.4 Degradación de la lignocelulosa

Debido a la dificultad de disolver la lignina sin destruir algunas de sus subunidades, su estructura química exacta es difícil de determinar. En general la lignina contiene tres alcoholes aromáticos (alcohol coniférico, sinapílico y p-cumarílico).

La lignina está más vinculada a las hemicelulosas y celulosas formando un sello físico alrededor de los últimos dos componentes, una barrera impenetrable para prevenir la penetración de soluciones y enzimas. Las macromoléculas de hemicelulosa son a menudo polímeros de las hexosas (xilosa y arabinosa), de pentosas (principalmente manosa) y un número de azúcares ácidos mientras que la celulosa es un polímero homogéneo de glucosa. De los tres componentes, la lignina es la más recalcitrante a la degradación mientras que la celulosa, debido a su estructura cristalina muy ordenada, es más resistente a la hidrólisis que hemicelulosa. Métodos de hidrólisis alcalinas y ácidas se han utilizado para

degradar lignocelulosa. Pero los ácidos débiles tienden a quitar lignina, y producir resultados pobres en la hidrólisis de la celulosa. Considerando que el tratamiento con ácido se produce bajo condiciones de alta temperatura y pH relativamente extremas corrosivas se requieren la utilización de equipos costosos. Además, producen reacciones inespecíficas, produciendo subproductos distintos de glucosa. También hay preocupaciones ambientales asociados con la eliminación de ácido. Para muchos procesos las enzimas son preferidas a los procesos de ácido o alcalinos ya que son Biocatalizadores específicos, que pueden operar bajo condiciones de reacción mucho más suaves, no producen productos indeseables y son amables para el medio ambiente. [7]

3.4.4.1 Tratamiento biológico

Los tratamientos biológicos son más favorables para el medio ambiente, ya que se generan menos sustancias tóxicas y hay una demanda mínima de energía, además facilita posteriores procesos de fermentación, aunque tiene algunas limitaciones en su implementación a nivel industrial debido a la poca accesibilidad a la maquinaria enzimática. Existen algunas bacterias y hongos que tienen la capacidad de degradar la lignocelulosa gracias a la producción de enzimas lignocelulósicas, convirtiendo estos residuos en productos de valor industrial. **Ver Anexo 1.** Microorganismos productores de enzimas lignocelulolíticas

3.4.4.2 Xilanasas

Xilanasas es el nombre genérico para una familia de enzimas que catalizan la ruptura de los enlaces xilosídicos internos de la molécula de heteroxilano (xilano), un tipo de polisacárido

constituyente mayoritario de la hemicelulosa, componente importante de la biomasa de la pared celular en vegetales. Las enzimas que participan de la biodegradación del heteroxilano son las endoxilanasas, que rompen el esqueleto de xilano en pequeños oligosacáridos, y las β -xilosidasas, que a su vez los degradan a moléculas de xilosa. Su uso principal es en procesos de clarificación de vinos y jugos, y en procesos de biodecoloración de pulpa de papel. [8]

Las Xilanasas, son heteropolisacáridos y su degradación total para producir xilosa y/o arabinosa es llevada a cabo, como en la celulosa, por un grupo de enzimas que participan sinérgicamente. Las más conocidas son:

- Endo- β -D-xilanasas (E.C. 3.2.1.8): rompen al azar los enlaces glicosídicos de la cadena principal de la molécula.
- Arabinofuranosidasa (E.C. 3.2.1.55): Hidroliza las cadenas laterales de arabinosa.
- Acetil xilanesterasas (E.C. 3.1.1.72): liberan grupos acetatos.
- Glucoronidasa (E.C. 3.2.1.139): remueve las cadenas laterales de ácido glucorónico a partir de unidades de xilosa.
- β -xilosidasas (E.C. 3.2.1.37): son enzimas activas sobre oligosacáridos cortos, llevando a cabo la hidrólisis de los enlaces β -1,4-aril-xilopiranosido produciendo xilosa.

Una de las aplicaciones más importantes de las xilanasas es en la industria de la pulpa y del papel. Las xilanasas, también se usan en la industria alimentaria en la clarificación de jugos y vinos, licuefacción de mucílago de café; extracción de saborizantes y pigmentos, aceites de plantas y semillas; maceración de materia vegetal; acondicionamiento de piensos para aves y cerdos y en la industria de la panificación, entre otras. [9]

3.4.4.3 Tratamiento químico

Hidrólisis ácida: Utiliza ácidos para la transformación estructural de los polisacáridos (hemicelulosa y celulosa) en monómeros, para ello se usan ácidos como el sulfuroso, clorhídrico, sulfúrico, fosfórico, nítrico y fórmico. Pueden emplearse ácidos concentrados y ácidos diluidos.

Oxidación húmeda: Hace uso de compuestos oxidantes como el hidrógeno o el ácido peracético, los cuales son aplicados a la biomasa sumergida en el agua.

Hidrólisis con álcalis: Es realizada con hidróxido de sodio (NaOH) diluido, en el cual se sumerge la biomasa lignocelulósica, esto provoca un hinchamiento, lo cual favorece la entrada de bacterias y enzimas.

Tratamiento con solventes orgánicos: En este proceso se combina un ácido orgánico con uno inorgánico. Se usan disolventes orgánicos como metanol, etanol, acetona, etilenglicol, trietilenglicol y alcohol tetrahidrofurfurílico y ácidos inorgánicos como oxálico, acetilsalicílico y salicílico. [10]

3.5 Diseño de cebadores

El diseño cuidadoso de cebadores es uno de los aspectos más importantes de la PCR. Cebadores mal diseñados pueden amplificar otros fragmentos de ADN distintos a los buscados (amplificación inespecífica). En el diseño de los mismos algunas reglas se han demostrado como útiles, por ejemplo:

- I. Cada cebador individual debe contar con una longitud de 18-24 bases.
- II. Se debe mantener un contenido de G:C (Guanina:Citosina) entre 40 y 60%.
- III. Los dos cebadores del par deben de tener temperatura de fusión "T_m" cercanos, dentro de los 5°C.
- IV. La secuencia de los cebadores individuales debe iniciarse y terminarse con 1 o 2 bases púricas.
- V. Evitar regiones con potencialidad para formar estructuras secundarias internas.
- VI. Evitar poli X.
- VII. Secuencias adicionales pueden ser agregadas en el extremo 5' del cebador (no incluir cuando se estima la T_m del cebador).
- VIII. Se pueden agregar degeneraciones en algunas posiciones del cebador:
 - a) Se incrementa el riesgo de amplificación inespecífica.
 - b) Se disminuye la concentración en la mezcla de cada uno de los cebadores posibles
 - c) No se recomienda utilizar más de 64 cebadores diferentes en la mezcla.
 - d) Código IUPAC/IUB:

• A adenina	R	A o G
• C citosina	W	A o T
• G guanina	S	C o G
• T timidina en el ADN	Y	C o T

- Uracilo en el ARN K G o T
- N A o C o G o T H A o C o T; no G
- M A o C B C o G o T; no A
- V A o C o G; no T D A o G o T; no C [12]

3.5.1 Propiedades de oligonucleótidos que influyen en la eficiencia de la amplificación.

Composición de las bases

Contenido de G+ C debe, entre 40% y 60%, con una distribución uniforme de las cuatro bases a lo largo de la longitud del cebador (por ejemplo, no tractos polipurinas y no repetición de nucleótidos).

Longitud

La región del cebador complementario a la plantilla debe ser de 18-25 nucleótidos de longitud. Los miembros de una pareja de cebadores no deben diferir de largo por más de 3 pb.

Secuencia repetida y la auto-complementaria

No deben estar presentes secuencias repetidas invertidas o secuencias auto-complementarias de más de 3 pb de longitud. Las secuencias de este tipo tienden a formar estructuras de horquilla, que si es estable bajo condiciones de PCR, puede prevenir con la eficacia del oligonucleótido de recocido a su ADN diana.

La complementariedad entre los miembros de una pareja de cebadores

Las secuencias 3' terminales de un cebador no debe ser capaces de enlazar cualquier sitio en el otro cebador. Debido a que los cebadores están presentes en una concentración elevada en la PCR,

incluso la complementariedad débil entre ellos para la formación de híbridos y la síntesis y la consiguiente amplificación de dímeros.

Si se forman dímeros temprano en la PCR, pueden competir por la DNA polimerasa el cebador, y los nucleótidos y así puede suprimir la amplificación del ADN diana. La formación de dímeros de cebador puede ser reducido por el diseño cuidadoso del cebador, mediante el uso de arranque en caliente o de toma de contacto de PCR, y/o por el uso de la ADN polimerasa especialmente formulado (por ejemplo, Ampli Gold, Peking-Elmer) Cuando más de un par de cebadores se utiliza en un solo PCR, comprobar que ninguno de los extremos 3' se detectan complementarios a cualquier otros cebadores en la reacción.

Temperatura de Fusión o alineamiento

Los valores calculados de T_m de los miembros de una pareja de cebadores no deben diferir en > 5 °C la T_m del producto amplificado, no debe diferir de los valores de T_m de los pares de cebadores de $> 10^{\circ}\text{C}$. Este alojamiento se asegura de que el producto amplificado se desnaturaliza eficientemente durante el ciclo de la PCR.

3' Terminal

La naturaleza del extremo 3' del cebador es crucial. Si es posible, el extremo 3' base de cada cebador puede ser G o C. Sin embargo, los cebadores con un NNCG... o... en su secuencia de NNCG extremos 3' no se recomiendan debido a que el G inusualmente es alto en las bases terminales GC y promueve la formación de estructuras de horquilla y puede generar dímeros de cebadores.

Adición de sitios de restricción, promotores de bacteriófago, y la secuencia a otro extremos 5' de los cebadores

Secuencias útiles no complementarias al ADN diana se añaden comúnmente a los 5' del cebador. En general la presencia de tales secuencias no afecta significativamente el recocido del oligonucleótido a su ADN diana. Estas secuencias adicionales incluyen promotores de bacteriófago (Kain et al. 1991) y las abrazaderas GC (Sheffield et al. 1989). Los sitios de restricción son caso especial. Debido a que la eficiencia de la escisión de los sitios de restricción terminales situados en el 5' de las moléculas de ADN es pobre, el cebador debe extenderse por lo menos tres nucleótidos adicionales más allá de la secuencia de reconocimiento de la enzima de restricción.

La colocación de sitios de cebado

Dependiendo de la finalidad del experimento, la posición de sitios de cebador puede verse limitada por la ubicación de la mutación, los sitios de restricción, las secuencias codificantes, microsatélites, o elementos que actúan en cis. Al diseñar cebadores para su uso en plantillas de ADNc, lo mejor es usar cebadores de avance y retroceso que se unen a secuencias de exones diferentes. Esto permite que los productos de amplificación procedentes de ADN y cDNA genómico contaminado puedan ser fácilmente distinguibles.

Los cebadores para PCR degenerada

Cuando una secuencia corta de aminoácidos ha sido obtenida por secuenciación de una proteína purificada, pools de oligonucleótidos degenerados que contienen todas las combinaciones posibles de

codificación puede ser utilizada para amplificar el genoma correspondiente o secuencias de cDNA [11]

3.6 MEME versión 4.4.0

Es una de las herramientas más utilizadas para la búsqueda de 'señales' nuevas en conjuntos de secuencias de proteínas o ADN. El MEME funciona mediante la búsqueda de patrones de secuencias repetidas conocidas como motivos, que se producen en las secuencias de ADN o proteína proporcionados por el usuario. [12]

3.7 MAST versión 4.9.0

Herramienta de validación que permite realizar la búsqueda de motivos en otras secuencias. Emplea una base datos para dicha búsqueda y emite una salida similar a la del MEME, lo que se encontrara será una lista de las secuencias que contienen el motivo que se busca. [13]

3.8 CODEHOP

Es un programa que permite diseñar cebadores degenerados, a partir de motivos conservados de aminoácidos pertenecientes a un alineamiento múltiple de secuencias proteicas. Funciona mediante una matriz de puntuación específica de posición. [14]

4 METODOLOGÍA PROPUESTA

Lugar de estudio: Universidad Católica de Manizales

Procedimiento llevado a cabo en: Laboratorios del Grupo de Investigaciones Biológicas GIBI de la Universidad Católica de Manizales.

Construcción base de datos de entrenamiento: La búsqueda de las secuencias que codifican para la endoxilanasas en bacterias, hongos y plantas se realizó mediante una palabra clave en la base de datos de proteínas del NCBI, se filtró por los grupos taxonómicos correspondientes, las secuencias obtenidas fueron renombradas usando expresiones regulares teniendo en cuenta el siguiente orden: el identificador, nombre de la secuencia y la secuencia, también se descartaron las secuencias redundantes con una similitud del 100%, esto se realizó en el programa CD-HIT. Así mismo se eliminaron secuencias genómicas, secuencias con función putativa y con función desconocida.

Identificación por MEME: Las secuencias obtenidas fueron utilizadas como base de datos de entrenamiento en formato Fasta, el MEME se corrió utilizando el modelo de búsqueda zoops para permitir la presencia o ausencia de motivos en las secuencias, la búsqueda fue de máximo 20 motivos con un ancho mínimo de 7 y un ancho máximo de 50 aminoácidos.

Validación por MAST: Se realizó la validación de los perfiles obtenidos del MEME empleando la base de datos UniprotKDB – Swissprot y UniRef100.

Análisis de secuencias: Las secuencias enzimáticas obtenidas de la salida del MEME en formato txt fueron analizadas con el fin de hallar los motivos más conservados a ambos extremos de las secuencias, inicialmente se identificaron los motivos y luego las arquitecturas.

Diseño de cebadores: Se llevó a cabo con el programa CODEHOP, al cual se le suministro la representación en bloques de los motivos más conservados.

Validación de cebadores: Se validó la eficiencia de los cebadores diseñados en la aplicación de Java PrimerAnalisys.

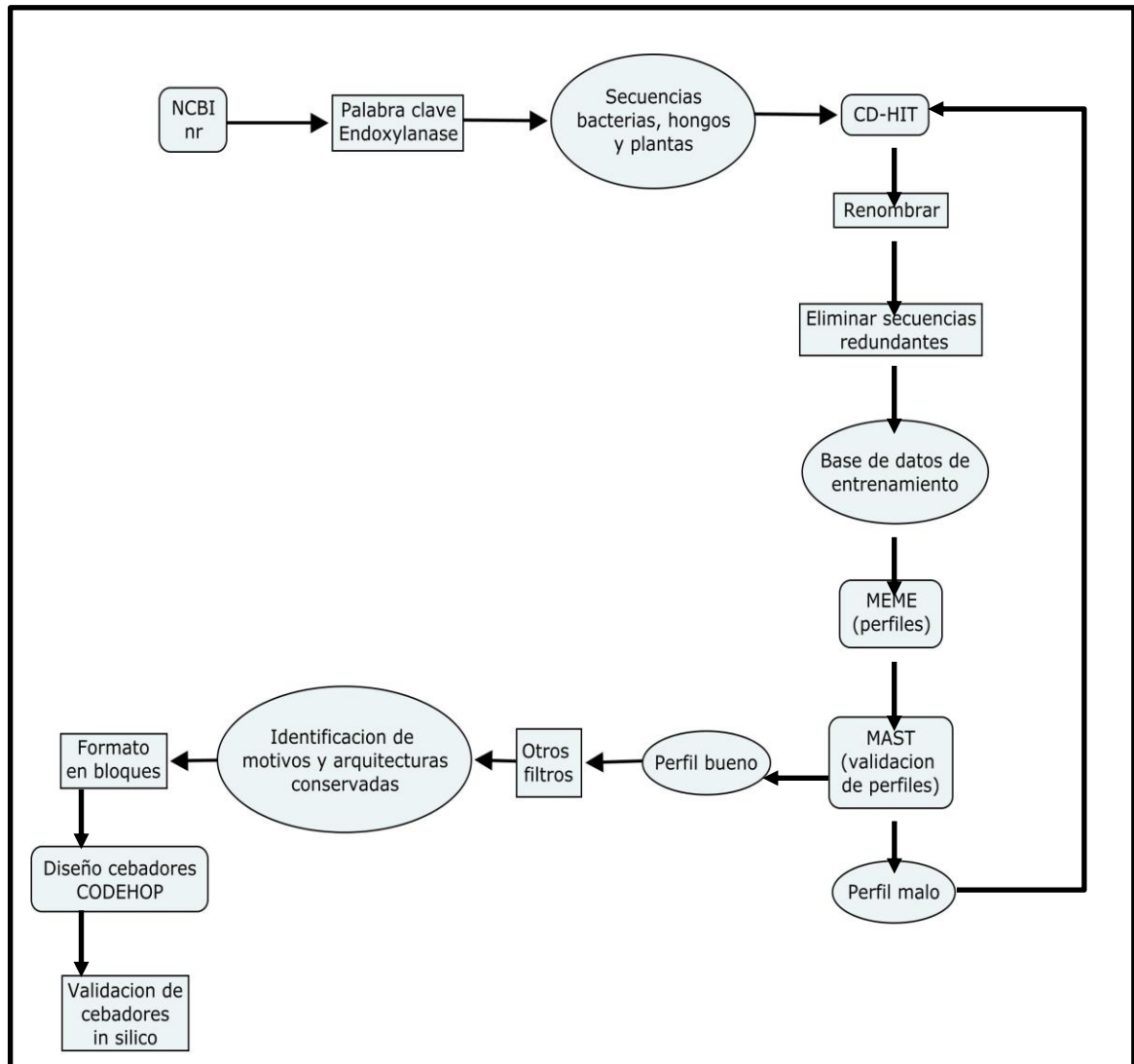


Figura 1. Diagrama de flujo metodológica.

5 RESULTADOS

Tabla 1. Secuencia enzimática de la Endoxilanas

Motivos de la secuencia enzimática de la Endoxilanas arrojados por el software MEME, editados en el programa Crimson Editor y posteriormente llevados a Excel para la organización por nombre de la enzima, distancia, probabilidad y motivos.

NOMBRE	PROBABILIDAD DE LA ARQUITECTURA	DISTANCIA 5'	MOTIVOS	PROBABILIDAD	DISTANCIA	MOTIVOS	PROBABILIDAD
Endoxylanase__thermopoly	2.25e-94	62	9	1.75e-19		14	2.72e-15
Endoxylanase__carica_pap	4.27e-120	211	18	1.91e-34	1	5	8.99e-17
Endoxylanase__thermoanae	6.91e-157	313	1	1.37e-05	53	5	2.34e-16
Endoxylanase__streptomyc	4.09e-92	12	9	2.25e-19		14	2.72e-15

Tabla 2. Motivos de secuencias

Representación con paleta de colores de los motivos en la secuencia enzimática de la Endoxilanasas.

NOMBRE DE LA SECUENCIA	MOTIVOS						
endoxylanase__thermopoly	9	14	1	5	2	8	9
endoxylanase__streptomyc	9	14	1	5	2	8	
endoxylanase__hypocrea_p	9	14	1	5	2	8	
endoxylanase__chaetomium	9	14	1	5	2	8	
endoxylanase__aspergillu	9	14	1	5	2	8	
endoxylanase__cellulosil	9	14	1	5	2	8	
endoxylanase__aspergillu	9	14	1	5	2	8	
endoxylanase_hypocrea_je	9	14	1	5	2	8	
endoxylanase__ruminococc	9	14	1	5	2	8	14 1 5
endoxylanase_glycosyl_hy	9	14	1	5	2	8	
xyn11c__chaetomium_therm	9	14	1	5	2	8	
endoxylanase_11b__chaeto	9	14	1	5	2	8	
endoxylanase_precursor__	9	14	1	5	2	8	
endoxylanase_11a__chaeto	9	14	1	5	2	8	14
endo_1,4_beta_xylanasa_1	9	14	1	5	2	8	8 9 2 8
xylanase__aspergillus_ni	9	14	1	5	2	8	
xylanase__bacillus_amylo	9	14	1	5	2	8	

beta_xylanase_precursor_	9	14	1	5	2	8				
xyn11__bacillus_lichenif	9	14	1	5	2	8				
xynbc1__botryotinia_fuck	9	14	1	5	2	8				
xylanase__synthetic_cons	9	14	1	5	2	8				
endo_1,4_beta_xylanase__	9	14	1	5	2	8				
thermostable_hybrid_xyla	9	14	1	5	2	8				
endoxylanase__neocallima	9	14	1	2	8	9	14	1	2	8
endo_1,4_beta_xylanase_1	9	1	5	2	8	14	11	15		
endoxylanase__cellulosil	9	1	5	2	8					
endoxylanase__neocallima	9	1	5	2						
endoxylanase_b__penicill	14	1	5	2	8					
endoxylanase_a__syntheti	14	1	5	2	8					
xylanase__aspergillus_ni	14	1	5	2	8					
xylanase_b__penicillium_	14	1	5	2	8					
endo_1,4_beta_xylanase_1	14	1	5	2	8					
endoxylanase__aspergillu	1	5	2	8						
endoxylanase_precursor__	2	9	14	1	5	2	8			
endoxylanase__hypocrea_p	5	3	4	6	7	10	12	11		
endoxylanase__cellulosil	5	3	4	6	7	10	12	11		
endoxylanase__paenibacil	5	3	4	6	7	10	12	11		
endoxylanase__cellvibrio	5	3	4	6	7	10	12	11		
endoxylanase__aspergillu	5	3	4	6	7	10	12	11		
endoxylanase__cellulomon	5	3	4	6	7	10	12	11	1	19
endoxylanase_precursor__	5	3	4	6	7	10	12	11		
endoxylanase_1588792_	5	3	4	6	7	10	12	11	19	19
alkaline_active_endoxyla	5	3	4	6	7	10	12	11		
alkaline_thermostable_en	5	3	4	6	7	10	12	11		

endoxylanase_c_gibberel	5	3	4	6	7	10	12	11											
endo_1,4_beta_d_xylanase	5	3	4	6	7	10	12	11											
a_chain_a_crystal_struct	5	3	4	6	7	10	12	11											
endo_1,4_beta_xylanase__	5	3	4	6	7	10	12	11											
endo_1,4_beta_xylanase__	5	3	4	6	7	10	12	11	16										
endo_1,4_beta_xylanase__	5	3	4	6	7	10	12	11	15										
xylanase__caldicellulosi	5	3	4	6	7	10	12	11	4										
a_chain_a_atomic_resolut	5	3	4	6	7	10	12	11											
xyna2__paenibacillus_sp.	5	3	4	6	7	10	12	11											
1,4_beta_d_xylan_xylanoh	5	3	4	6	7	10	12												
anther_endoxylanase__hor	5	3	4	6	7	10	13												
endoxylanase_d_gibberel	5	3	4	6	7	12	11												
xylanase__prevotella_rum	5	3	6	4	6	7	10	11											
endoxylanase__thermoanae	1	5	3	4	6	7	10	12	11	19	19	12							
endo_1,4_beta_xylanase_5	1	5	3	4	6	7	10	12	11	19	19	12							
xyna_precursor__thermoan	1	5	3	4	6	7	10	12	11	19	19	12							
xys1__streptomyces_halst	18	5	3	4	6	7	10	12	11										
endo_1,4_beta_xylanase__	18	5	3	4	6	7	10	12	11										
1,4_beta_d_xylan_xylanoh	18	5	3	4	6	7	10	12											
oryza_sativa_japonica_gr	18	5	3	4	6	7	10	13											
oryza_sativa_japonica_gr	18	5	3	4	6	7	10	13											
oryza_sativa_japonica_gr	18	5	3	4	6	7	10	13											
1,4_beta_xylanase__oryza	18	5	3	4	6	7	10	13											
1,4_beta_xylanase__oryza	18	5	3	4	6	7	10	13											
xylanase1__zea_mays_1624	18	5	3	4	6	7	10	13											
endoxylanase__carica_pap	18	5	3	4	6	7	10	13											
endoxylanase__aspergillu	12	5	3	4	6	7	10	12	11										

xyna_thene_recname:_full	3	4	6	7	10	12	11	11	19	19	12		
endoxylanase__cellulosil	3	4	16	7	10	12	11						
glucuronoxylanase_xync_b	15	7	16	17									
endoxylanase__micromonos	15	7	6	15	7								
ynff__bacillus_sp_267799	7	16	17	15									
ynff__bacillus_subtilis_	7	16	17	15									
glucuronoxylanase_126253	7	16	17	15									
glucuronoxylanase_xync__	7	7	16	17	15								
xync1_bacsu_recname:_ful	7	7	16	17	15								
putative_anther_endoxyla	6	7	10	13									
endo_1,4_beta_xylanase_7	8	19	16	5	3	4	6	7	10	12	11	5	19
c_chain_c_crystal_struct	8	9	14	1	5								
beta_1,4_xylanase__clost	8	18	5	3	4	6	7	10	12				
xylanase__caldicellulosi	4	10	5	3	4	6	7	10	12	11			

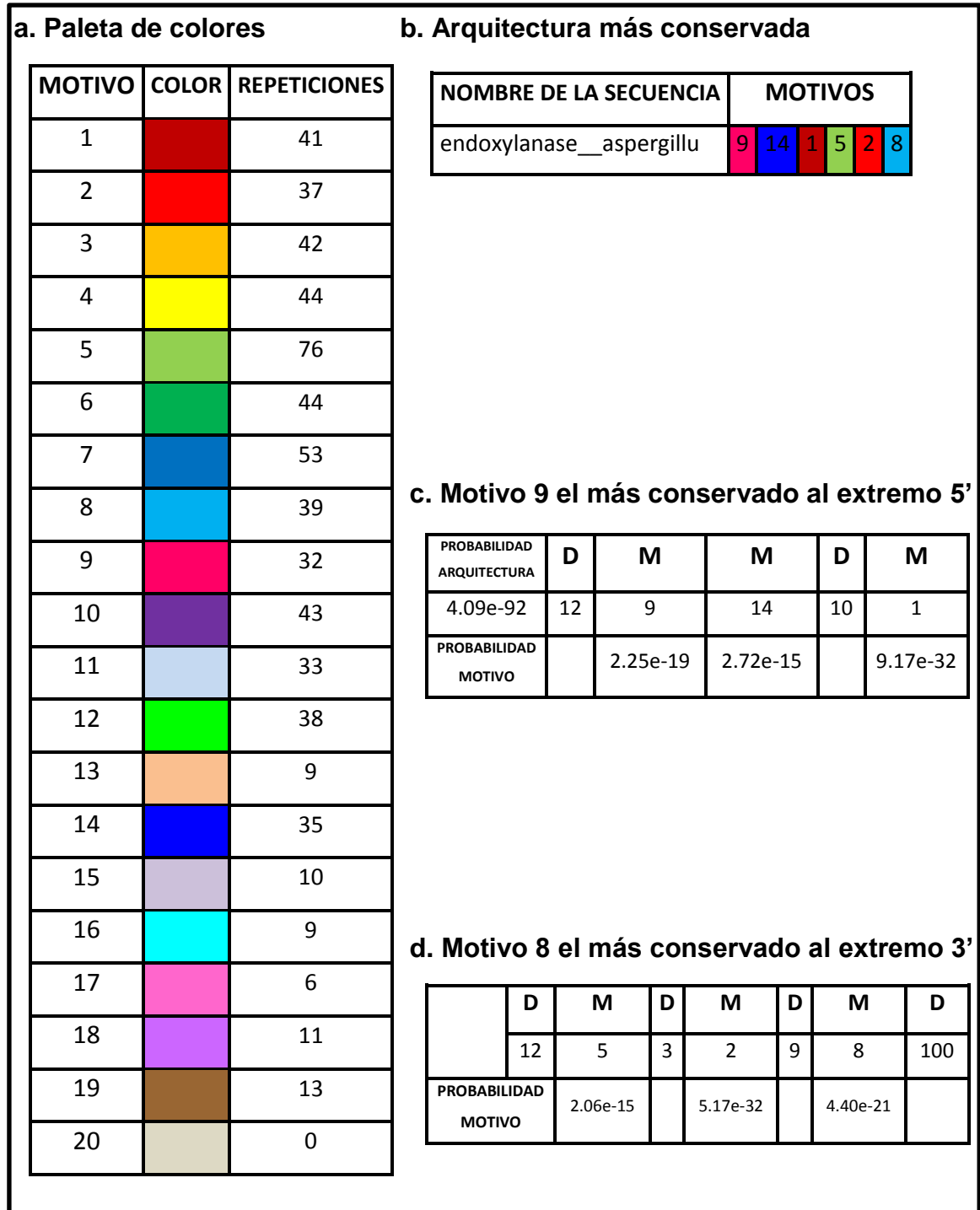
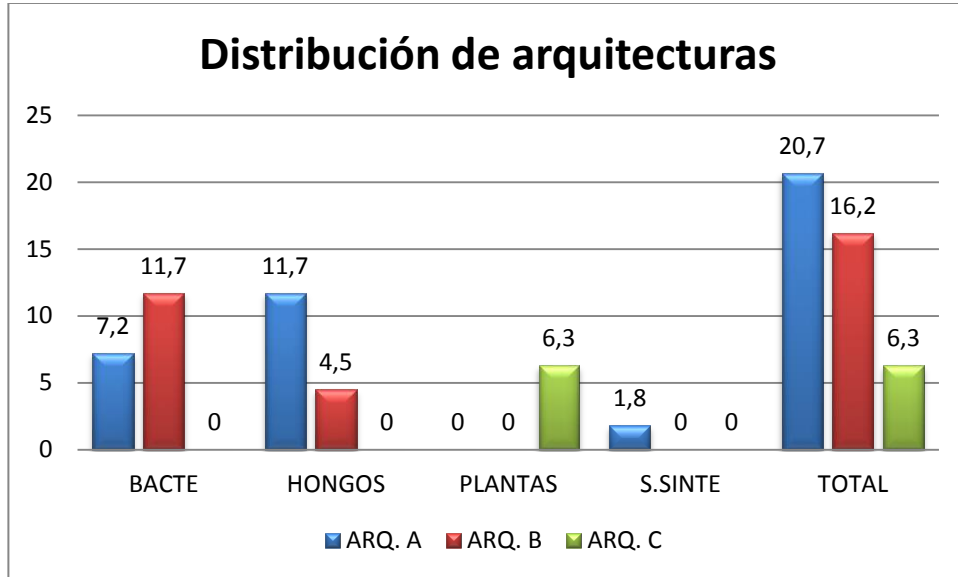


Figura 2. A. Paleta de colores: Representación por colores de cada uno de los motivos y sus repeticiones en la secuencia. **B. Arquitectura más conservada:** Esta se repite 23 de 111 en la secuencia de la enzima Endoxilanasa arrojada por el MEME versión 4.4.0. **C. Motivo 9 el más conservado al extremo 5':** El motivo

9 se repite 29 de 111 en las secuencia de la enzima Endoxilanasa arrojada por el MEME versión 4.4.0. **D. Motivo 8 el más conservado al extremo 3'**: El motivo 8 se repite 35 de 111 en las secuencia de la enzima Endoxilanasa arrojada por el MEME versión 4.4.0

A.



B.

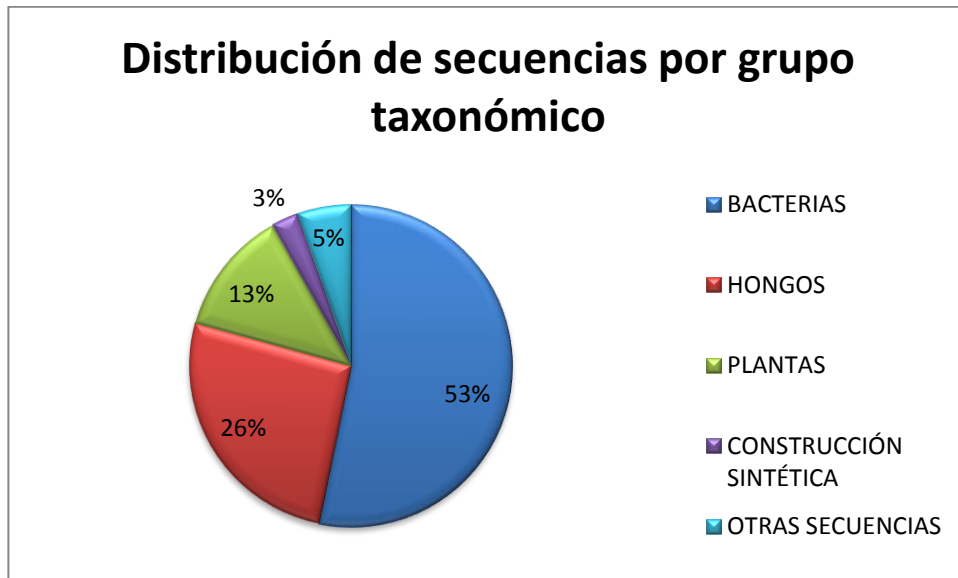


Figura 3. Análisis de motivos y hallazgo de arquitecturas más comunes. A. Arquitecturas más conservadas: Se clasificaron las arquitecturas más conservadas en A, B y C, se determinó el número de repeticiones y su porcentaje, también se determinaron las arquitecturas más conservadas en los respectivos grupos taxonómicos. **B. Distribución de secuencias por grupo taxonómico:** Se identificaron los grupos taxonómicos al que pertenecía cada secuencia y posteriormente se contaron y sacaron los respectivos porcentajes.

Tabla 3. RESULTADOS DE MAST- Uniprot

Se hallaron un total de 139 secuencias de una base de datos de 539.829 secuencias. De acuerdo al punto de corte establecido 1×10^{-5} , de las 104 secuencias 92 pertenecen a endoxilanasas, 8 pertenecen a la familia de las xilanasas, 2 tienen función similar y otras 2 son secuencias putativas.

VALIDACIÓN MAST - UniprotKDB			
Secuencias endoxilanasas	Familias Xyl	Función similar	Otras
92	8	2	2

Porcentaje de arquitecturas				
	Secuencias endoxilanasas	Familias Xyl	Función similar	Otras
ARQ. A	47,8%	12,5%	0%	0%
ARQ. B	31,5%	12,5%	50%	50%
ARQ. incompleta	20,7%	50%	50%	50%
Otras ARQ.		25%		

VALIDACIÓN MAST - UniprotKDB	
Falsos positivos	0
Falsos negativos	0

Tabla 3.1 Arquitecturas conservadas

Estas arquitecturas fueron las mismas que encontró el MEME con la base de datos de entrenamiento.




Arquitecturas mas repetidas	A		46 repeticiones (41,4%)
	B		32 repeticiones (28,8%)
Arquitectura utilizada en el diseño de cebadores			Motivos seleccionados: 8 y 9

Tabla 4. RESULTADOS DE MAST- Uniref100

Se hallaron un total de 2.645 secuencias de una base de datos de 21.026.652 secuencias. De acuerdo al punto de corte establecido 1×10^{-5} , de las 2.645 secuencias, 1.449 pertenecen a xilanasas, 434 tienen una función similar y otras 762 corresponden a otras secuencias.

VALIDACIÓN MAST – UniRef100	
Familia de xilanasas	1.449
Función similar	434
Fragmentos de secuencias	242
Secuencias Putativas	272
Proteína sin caracterizar	144
Otras secuencias	104

VALIDACIÓN MAST – Uniref100	
Falsos positivos	104
Falsos negativos	17

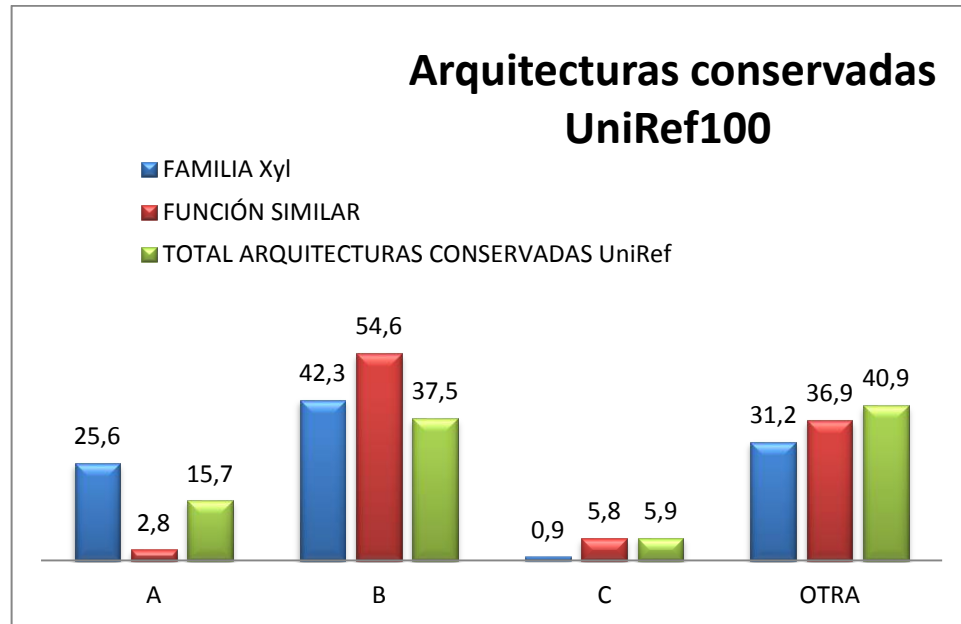


Figura 4. Arquitecturas conservadas UniRef100: Estas arquitecturas fueron las mismas que encontró el MEME con la base de datos de entrenamiento, se muestra la totalidad de ellas en porcentaje, al igual que su distribución dependiendo de la familia a la que pertenecen.

Xyl: Xilanasas

Función similar: enzimas que no pertenecen a la familia de las xilanasas pero pueden degradar enlaces B-1,4.

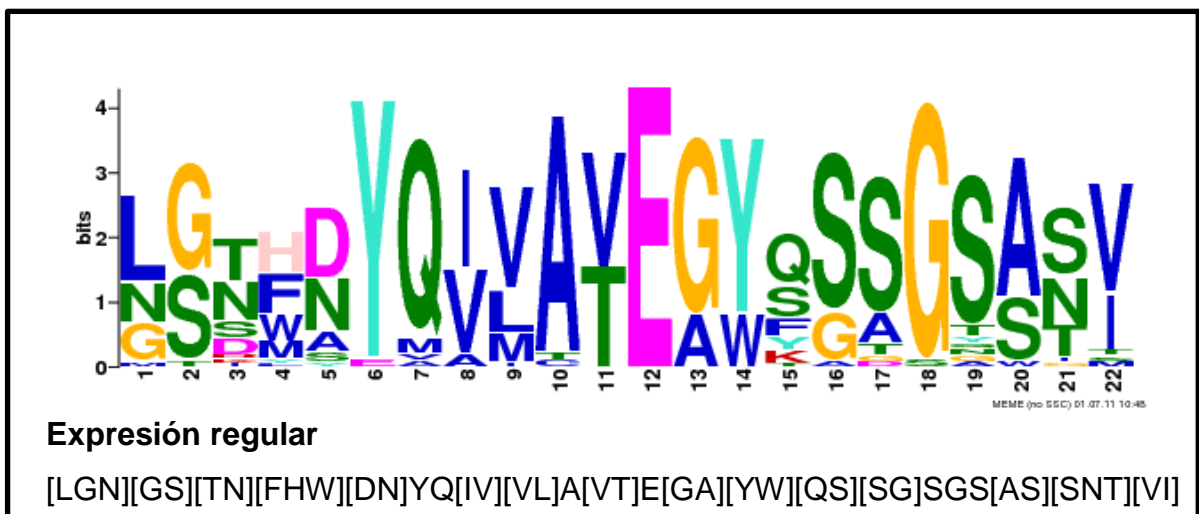


Figura 5. Representación en logos del motivo 8 y su respectiva expresión regular: La representación en logos nos muestra el nivel de conservación, de esta manera se dice que entre mayor sea el tamaño de la letra, mayor es el nivel de conservación. De esta expresión regular se obtienen los aminoácidos correspondientes para cada letra de acuerdo al código IUPAC y luego los codones correspondientes.

Tabla 5. Diseño de cebadores para el motivo 8

Secuencias utilizadas para el diseño de los cebadores del motivo 8 en el CODEHOP, se descartaron 20 secuencias para diseñar los cebadores, quedando solo 10, 3 (30%) de Bacterias y 7 (70%) de hongos.

NOMBRE DE LA SECUENCIA	MOTIVOS						
Endoxylanase__aspergillu	9	14	1	5	2	8	
Endoxylanase__thermopoly	9	14	1	5	2	8	9
Endoxylanase__chaetomium	9	14	1	5	2	8	
Endoxylanase__hypocrea_je	9	14	1	5	2	8	
Endoxylanase__hypocrea_p	9	14	1	5	2	8	
Endoxylanase__aspergillu	9	14	1	5	2	8	
Endoxylanase__streptomyc	9	14	1	5	2	8	
Endoxylanase__precursor__	9	14	1	5	2	8	
Endo_1,4_beta_xylanase__	9	14	1	5	2	8	
Endoxylanase__cellulosil	9	14	1	5	2	8	

Tabla 6. Cebador Sentido diseñado para el motivo 8

Cebador diseñado con el programa CODEHOP

Bloque Endoxilanasas Sentido Oligonucleótidos	
Secuencia de aminoácidos	L G T H D Y Q I V A T E G Y Q S S G S A T I
Región Abrazadera:	TGGACTACCAGATCGTGGCA
Región Central:	ryngarggnta -3'
Puntuación:	68
Longitud región central:	11
Degeneración:	128
Longitud región abrazadera:	20
Temperatura de Fusión o Alineamiento:	62.7 °C

Tabla 7. Alineamiento del cebador con la secuencia

Alineamiento del cebador sentido del motivo 8, con una eficiencia del 73%.

L G T H D Y Q I V A T E G Y Q S S G S A T I
5'- TGGACTACCAGATCGTGGCAryngarggnta -3'

Tabla 8. Cebador Antisentido diseñado para el motivo 8

Cebador diseñado con el programa CODEHOP

Complemento del Bloque Endoxilanasas Antisentido Oligonucleótidos	
Secuencia de aminoácidos:	L G T H D Y Q I V A T E G Y Q S S G S A T I
Región Central	ctycnattrw
Región Abrazadera	CAGAAGGCCAAGGCGATGATAC-5'
Degeneración:	64
Longitud region central:	11
Puntuación:	57
Longitud region abrazadera:	22
Temperatura de Fusión o Alineamiento:	62.2 °C

Tabla 9. Alineamiento del cebador con la secuencia

Alineamiento del cebador Antisentido del motivo 8, con una eficiencia del 77%.

L G T H D Y Q I V A T E G Y Q S S G S A T I
3'- ctyccnatrrwCAGAAGGCCAAGGCGATGATAC -5'

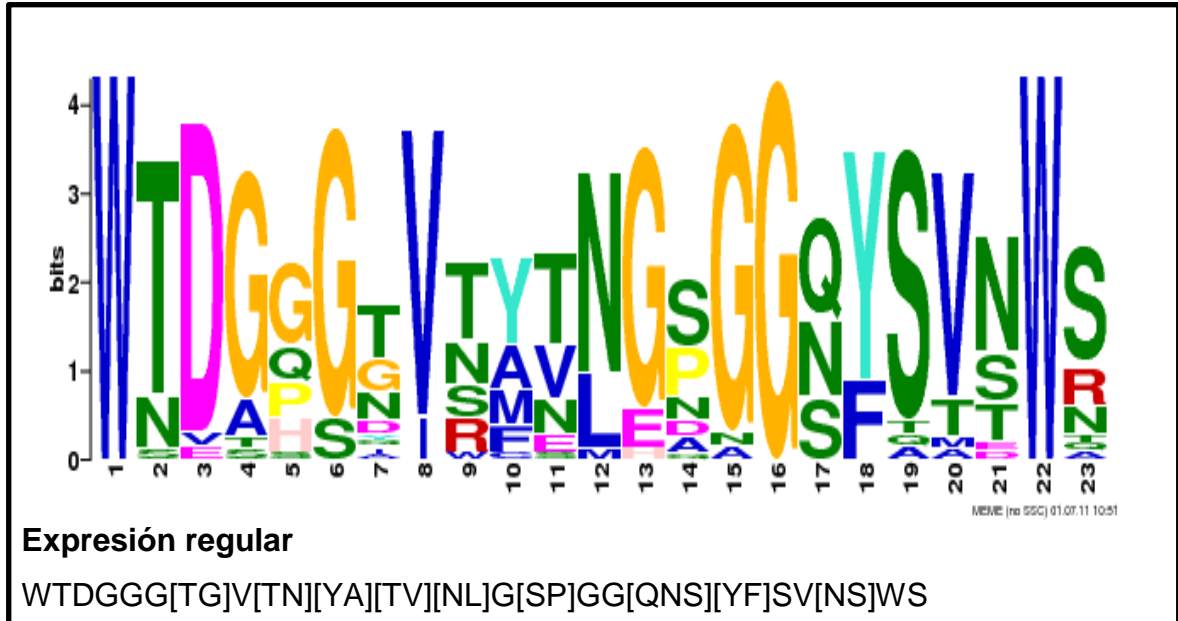


Figura 6. Representación en logos del motivo 9 y su respectiva expresión regular: La representación en logos nos muestra el nivel de conservación, de esta manera se dice que entre mayor sea el tamaño de la letra, mayor es el nivel de conservación. De esta expresión regular se obtienen los aminoácidos correspondientes para cada letra de acuerdo al código IUPAC y luego los codones correspondientes.

Tabla 10. Diseño de cebadores para el motivo 9

Secuencias utilizadas para el diseño de los cebadores del motivo 9 en el CODEHOP, se descartaron 9 secuencias para diseñar los cebadores y quedaron 14, de estas 11 (78,6%) eran de hongos, 2 (14,3%) de bacterias y una (7,1%) sintética que posteriormente se borró para mejorar la efectividad del cebador.

NOMBRE DE LA SECUENCIA	MOTIVOS						
Beta_xylanase_precursor_	9	14	1	5	2	8	
Endoxylanase_hypocrea_je	9	14	1	5	2	8	
Endoxylanase__hypocrea_p	9	14	1	5	2	8	
Endoxylanase__aspergillu	9	14	1	5	2	8	
Endoxylanase_11b__chaeto	9	14	1	5	2	8	
Endoxylanase__aspergillu	9	14	1	5	2	8	
Endoxylanase_precursor__	9	14	1	5	2	8	
Endoxylanase_11a__chaeto	9	14	1	5	2	8	14
Endoxylanase__chaetomium	9	14	1	5	2	8	
Endo_1,4_beta_xylanase__	9	14	1	5	2	8	
Endoxylanase__thermopoly	9	14	1	5	2	8	9
Endoxylanase__streptomyc	9	14	1	5	2	8	
Endoxylanase_precursor__	9	14	1	5	2	8	

Tabla 11. Cebador Sentido diseñado para el motivo 9

En este caso el programa CODEHOP no encontró ningún cebador para el motivo 9.

Bloque Endoxilanasas
No se encontraron cebadores.

Tabla 12. Cebador Antisentido diseñado para el motivo 9

Cebador diseñado con el programa CODEHOP

Complemento del Bloque Endoxilanasas Antisentido Oligonucleótidos	
Secuencia de aminoácidos:	W T D G H G N V T Y T N G S G G Q Y S V N W S
Región Central	acctknctrhs
Región Abrazadera	GGTGCCGTTGCACTCC -5'
Degeneración:	96
Longitud Región Central:	11
Puntuación:	61
Longitud Región Abrazadera:	16
Temperatura de Fusión o Alineamiento:	60.9 °C

Tabla 13. Alineamiento del cebador con la secuencia

Alineamiento del cebador Reverse del motivo 9, con una eficiencia del 82%.

W T D G H G T V T Y T N G P G G Q Y S V N W S
3'- acctknctrhsGGTGCCGTTGCACTCC -5'

6 DISCUSIÓN

Las secuencias encontradas en la base de datos de proteínas del NCBI, fueron renombradas y tratadas mediante filtros con el objetivo de descartar información no necesaria que pudiera interferir en los análisis. El método empleado para el descubrimiento de los motivos fue el algoritmo MEME, el mismo método desarrollado en el artículo de (Gorrón E. et al. 2010), el cual se basa en alineaciones múltiples locales, que permiten encontrar las regiones conservadas que servirán en el diseño de cebadores.

Tras la corrida del MEME, se analizaron los perfiles obtenidos para determinar los motivos y arquitecturas más conservadas, después de esto se evidenció que había 3 arquitecturas que se repetían significativamente a lo largo de las

secuencias, lo que indica que hay una buena homología entre ellas y que los perfiles encontrados guardan una buena similitud. A su vez esto permite considerar que uno de estos perfiles puede ser un buen candidato para la elaboración de los cebadores. De la arquitectura que más se repitió, se escogieron el motivo 8 y 9, a partir de estos se obtuvieron 2 cebadores para el motivo 8 y uno para el motivo 9.

En la validación realizada por el MAST con la base de datos Uniprot se hallaron 104 (74,8%) secuencias por debajo del punto de corte establecido 1×10^{-5} , este grupo de secuencias se clasificó de la siguiente manera: secuencias de endoxilasas, secuencias de la familia de las xilasas, secuencias con función similar y otras secuencias (putativas), el mayor porcentaje (88,4%) de secuencias correspondían a endoxilasas.

En el análisis de los perfiles hallados se observó que estos contenían los mismos motivos y arquitecturas arrojados por el MEME, además la arquitectura seleccionada para el diseño de cebadores fue la misma que se halló como la más conservada. No se identificaron enzimas correspondientes a xilasas por encima del punto de corte (falsos negativos), ni por debajo del punto de corte (falsos positivos).

Respecto a la validación realizada con el UniRef100, los resultados fueron similares pero contienen algunas variantes, probablemente debido a que esta base de datos es mucho más grande.

Se encontraron 2.645 secuencias por debajo del punto de corte establecido 1×10^{-5} , la clasificación que se hizo de las secuencias fue: Familia de enzimas xilasas, enzimas que cumplen una función similar, es decir que rompen enlaces B-1,4, las secuencias putativas, las fragmentadas, proteínas sin caracterizar y otras. Se identificaron 17 (5%) enzimas correspondientes a xilasas por encima del punto de corte (falsos negativos) y 104 (3,9%) secuencias que no correspondían a xilasas y estaban por debajo del punto de corte (falsos positivos).

Las arquitecturas y motivos encontrados son los mismos que halló el MEME, pero la arquitectura que se identificó como la más conservada no fue la misma que se empleó en el diseño de los cebadores.

Esto indica que los perfiles obtenidos son lo suficientemente buenos para el diseño de cebadores, por lo tanto se esperaría que su alineación sea específica.

La validación de los cebadores realizada arrojó para el cebador sentido del motivo 8 una eficiencia del 73% y para el antisentido una eficiencia del 77%, cumpliendo con los parámetros establecidos para el diseño de cebadores. Con relación al motivo 9, el cebador arrojó una eficiencia del 47%, no cumplió con los siguientes parámetros: la temperatura de fusión y la longitud, además podía formar 2 autodímeros, es importante tener en cuenta que el bloque de elaboración de este cebador contenía una secuencia sintética. De acuerdo a esto se realizó un segundo diseño de cebador sin la secuencia sintética, a diferencia del anterior este sí cumplía con los parámetros y tenía una eficiencia del 82%, lo que sugiere que la presencia de esta secuencia sintética influye significativamente en la eficiencia de los cebadores.

De acuerdo a lo anterior es posible considerar que los cebadores diseñados servirán posteriormente en la obtención del gen de la enzima endoxilanasas, siendo esto importante para avanzar en la búsqueda de las enzimas lignocelulolíticas en compostaje de higuera.

Con relación a la evolución de las endoxilanasas, se observó que las secuencias pertenecientes a estas, presentaban como gen predominante el *xynA*, tal como se describe en el artículo de (Briam P. et al. 1999), esto hace posible considerar que las secuencias de nucleótidos de este gen se han conservado a lo largo del tiempo.

7 CONCLUSIONES

- La elección de los perfiles fue acertada, lo que indica que el tratamiento de las secuencias mediante diversos filtros bioinformáticos funcionó al momento de descartar información no necesaria.
- Aunque los cebadores fueron diseñados a partir de perfiles, cumplieron con todos los parámetros que influyen en la amplificación, de modo que su uso es factible y es posible continuar con la parte experimental para la obtención de la enzima endoxilanasas.

9. PERSPECTIVAS

- Realizar la amplificación *in silico*, de los cebadores diseñados.
- Aplicar esta metodología para la obtención de otras enzimas lignocelulolíticas.
- Aplicar otras herramientas bioinformáticas.
- Difundir la información y el estado de los proyectos relacionados con bioinformática.

ANEXOS

Anexo 1. Microorganismos productores de enzimas lignocelulolíticas

Tipo de M.O	BACTERIAS		HONGOS		
	Aerobias mesofílicas y termofílicas	Anaerobias mesofílicas y termofílicas	Aerobios	Aerobios termofílicos	Aerobios mesofílicos
M.O	Cellulomonas sp., Cellvibriosp., Microbispora bispora y Thermomonosporasp.	Acetivibriocellulolyticus, Bacteroidessuccinogenes, Ruminococcusalbus, Ruminococcusflavifaciens y Clostridiumthermocellum	Trichoderma viride, Trichoderma reesei, Penicillium pinophilum, Sporotrichum pulverulentum, Fusarium solani, Talaromyces emersonii y Trichoderma koningi	Sporotrichum thermophile, Thermoascus aurantiacus y Humicola insolens	Neocallimastix frontalis, Piromonas communis, Sphaeromonas communis

Tomado de Sanchez, R. et al. (2010)

Anexo 2. Figura 3.

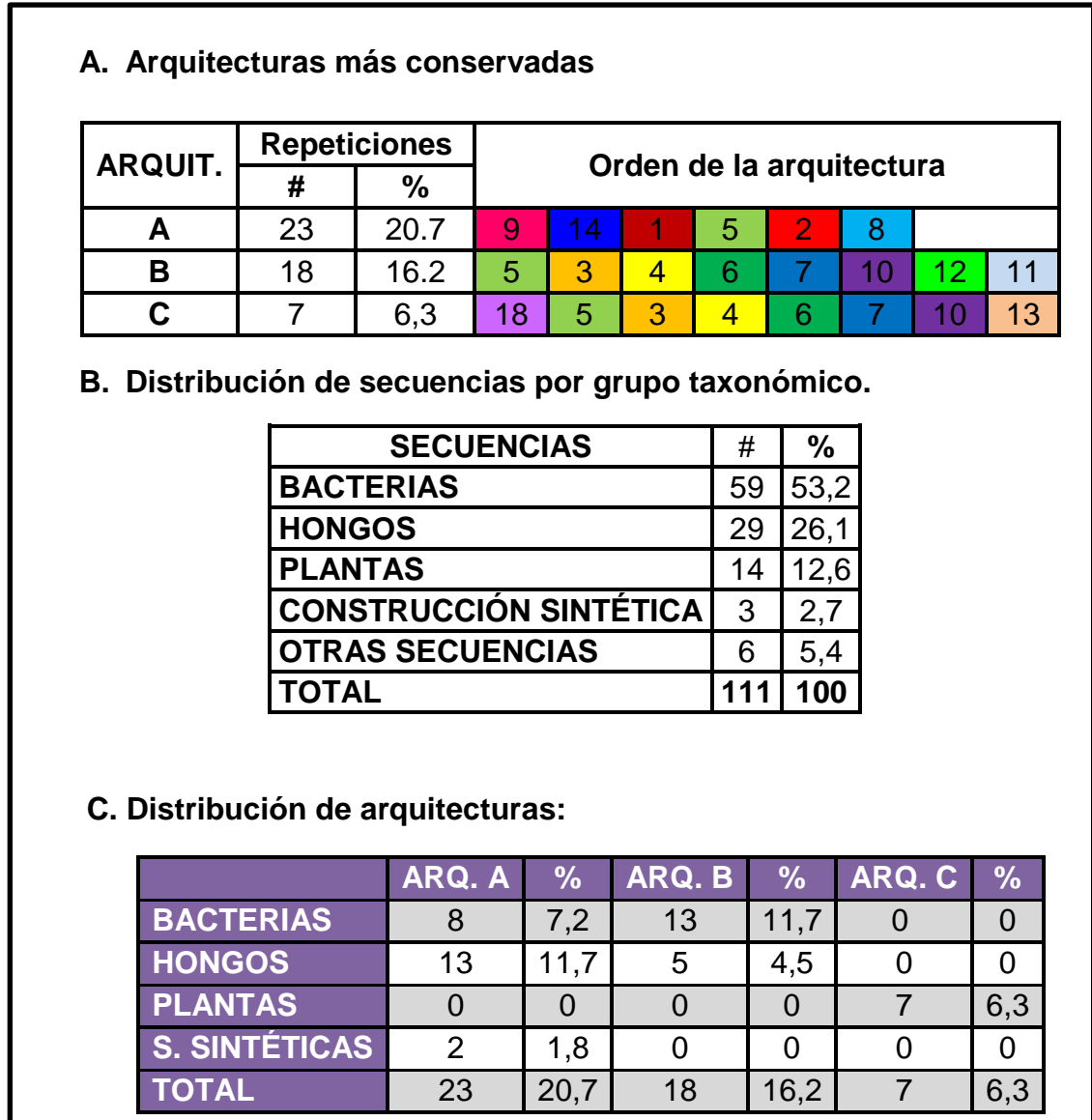


Figura 3. Análisis de motivos y hallazgo de arquitecturas más comunes. A. Arquitecturas más conservadas: Se clasificaron las arquitecturas más conservadas en A, B y C, se determinó el número de repeticiones y su porcentaje. **B. Distribución de secuencias por grupo taxonómico:** Se identificaron los grupos taxonómicos al que pertenecía cada secuencia y posteriormente se contaron y sacaron los respectivos porcentajes. **C. Distribución de**

arquitecturas: Se determinaron las arquitecturas más conservadas en los respectivos grupos taxonómicos.

Anexo 3. Figura 4

Arquitecturas más repetidas	A		415 repeticiones (15,7%)
	B		993 repeticiones (37,5%)
	C		155 repeticiones (5,9%)

/	FAMILIA Xyl	#	%	FUNCIÓN SIMILAR	#	%
ARQUITECTURAS	A	371	25,6	A	12	2,8
	B	613	42,3	B	237	54,6
	C	13	0,9	C	25	5,8
	OTRA	452	31,2	OTRA	160	36,9
	TOTAL	1449	100	TOTAL	434	100

Figura 4. Arquitecturas conservadas UniRef100: Estas arquitecturas fueron las mismas que encontró el MEME con la base de datos de entrenamiento, se muestra la totalidad de ellas, en número y porcentaje, al igual que su distribución dependiendo de la familia a la que pertenecen.

Xyl: Xilanasas

Función similar: enzimas que no pertenecen a la familia de las xilanasas pero pueden degradar enlaces B-1,4.

Anexo 4. Formato de bloques motivo 8

Este formato en bloques fue usado para el diseño de los cebadores.

ID Endoxylanase; BLOCK

AC Endoxylanase;

DE Endoxylanase

BLMOTIF 8 width=22 seqs=10

endoxylanase__aspergillus_fumigatus_125634595_ (203)

LGTHNYQIVATEGYQSSGSASI 1

endoxylanase__thermopolyspora_flexuosa_23337128_ (207)

LGSHDYQIMATEGYQSSGSSTV 1

endoxylanase__chaetomium_thermophilum_115344320_ (192)

LGTHDYQIVATEGYYSGSATV 1

endoxylanase_hypocrea_jecorina_455907_ (198)

LGTMDYQIVAVEGYFSSGSASI 1

endoxylanase__hypocrea_pseudokoningii_165906536_ (199)

LGTMDYQIVAVEGYFSSGSASI 1

endoxylanase__aspergillus_fumigatus_305377696_ (196)

LGNFDYMIVATEGYQSSGSATI 1

endoxylanase__streptomyces_thermovulgaris_104303750_ (157)

LGTFNYMILATEGYQSSGSSNI 1

endoxylanase_precursor__gibberella_zeae_56709478_ (202)

LGSHDYQILAVEGYQSSGQATM 1

endo_1,4_beta_xylanase__trichoderma_viride_6434133_ (198)

LGTLDYQIIAVEGYFSSGNANI 1

endoxylanase__cellulosilyticum_ruminicola_280977809_ (198)

MGKMYEAALTVEGYQSSGSANV 1

Anexo 5. Formato de bloques motivo 9

Este formato en bloques fue usado para el diseño de los cebadores.

ID Endoxylanase; BLOCK

AC Endoxylanase;

DE Endoxylanase

BL MOTIF 9 width=23 seqs=14

beta_xylanase_precursor__hypocrea_jecorina_780816_ (50)

WNDGHGGVTYTNGPGGQFSVNWS 1

endoxylanase_hypocrea_jecorina_455907_ (50)

WNDGHGGVTYTNGPGGQFSVNWS 1

endoxylanase__hypocrea_pseudokoningii_165906536_ (51)

WNDGHGGVTYTNGPGGQFSVNWS 1

endoxylanase__aspergillus_fumigatus_125634595_ (56)

WTDGGGDVTYTNGAGGSYSVNWR 1

endoxylanase_11b__chaetomium_thermophilum_23477176_ (56)

WTDGGGTWYTNNGGYSVNWQ 1

endoxylanase__aspergillus_fumigatus_305377696_ (49)

WTDGGGQVTYTNGGGQYQVDWN 1

endoxylanase_precursor__chaetomium_thermophilum_34499889_ (44)

WTDGQGNIRFNLESGGQYSVTWS 1

endoxylanase_11a__chaetomium_thermophilum_23477174_ (44)

WTDGQGNIRFNLESGGQYSVTWS 1

endoxylanase__chaetomium_thermophilum_115344320_ (44)

WTDGQGNIRFNLESGGQYSVTWS 1

endo_1,4_beta_xylanase__trichoderma_viride_6434133_ (50)

WNDGHSGVTYTNGAGGSFSVNWA 1

endoxylanase__thermopolyspora_flexuosa_23337128_ (62)

WTDAPGTVSMTLHSGGSYSTSWR 1

endoxylanase__streptomyces_thermovulgaris_104303750_ (12)

WTDSQGTVSMNMGSGGSYSTSWR 1

endoxylanase_precursor__gibberella_zeae_56709478_ (53)

WTDTPSAVTYTNGNGGQFSMNWN 1

LITERATURA CITADA

- [1] Quintero J.J. **Biodiésel de Higüerillo (*Ricinus communis* L.) Como Combustible en un Motor Diésel**, Facultad De Ingeniería, Universidad De San Carlos De Guatemala, Guatemala, abril de 2005. Pag. 47-48. Desde sitio Web: http://ceur.usac.edu.gt/Biocombustibles/74_biodisel_08_0467_M.pdf
- [2] Vargas G.M., Comisión nacional para el ahorro de energía (Conae). Dirección de ahorro de energía y transporte. **Biodiesel**. Diciembre de 2007. Tomado el 2 de abril de la pagina web: <http://www.conae.gob.mx/work/sites/CONAE/resources/LocalContent/466/2/biodiesel.pdf>
- [3] Álvarez M.C. **Biocombustibles: desarrollo histórico-tecnológico, mercados actuales y comercio internacional**, Revista Economía Informa, núm. 359, edición julio-agosto del 2009, pp 63-64. Desde sitio Web:http://www.economia.unam.mx/publicaciones/econinforma/pdfs/359/04_carlosalvarez.pdf
- [4] Ambrose L. Spring 2011. **Introduction, definition and process of metagenomics**. Tomado el 2 de Abril de la pagina web: <http://sgugenetics.pbworks.com/w/page/38604801/Introduction%2C-definition-and-process-of-metagenomics>
- [5] Sandoval L. Organización de Estados Americanos. **Programa Horizontal de Tecnologías Limpias y Energías Renovables. Manual de Tecnologías Limpias en PyMEs del Sector Residuos Sólidos**. Julio de 2006. Tomado el 23 de febrero de la pagina web: <http://www.redrrss.pe/material/20090128192419.pdf>
- [6] Cuervo L, Folch J, L; Quiroz R, E. **Lignocelulosa Como Fuente de Azúcares Para la Producción de Etanol**. Centro de Investigación en

Biotecnología, UAEM. Instituto de Biotecnología, UNAM. Av. Universidad 2001 Col. Chamilpa, Cuernavaca, Mor. 62209, México. Pag 2-3. Desde sitio Web: http://www.smbb.com.mx/revista/Revista_2009_3/Lignocelulosa.pdf

- [7] Howard R.L., Abotsi E., Jansen van Rensburg E.L.1 and Howard S.3; Review: **Lignocellulose biotechnology: issues of bioconversion and enzyme production**; African Journal of Biotechnology Vol. 2 (12), pp. 602-619, December 2003 Accepted 19 November 2003. Desde sitio Web: <https://tspace.library.utoronto.ca/bitstream/1807/1923/1/jb03112.pdf>
- [8] Costa M., Torres M., Magariños H. **Producción y purificación parcial de enzimas hidrolíticas de *Aspergillus ficuum* en fermentación sólida sobre residuos agroindustriales.** Revista colombiana de biotecnología Volumen12, número 2, Bogotá 2010. Disponible en página web: http://www.scielo.unal.edu.co/scielo.php?script=sci_arttext&pid=S0123-34752010000200013&lng=pt&nrm=
- [9] Ponce N.T.; Pérez A.O. **Celulasas y xilanasas en la industria.** XXX Aniversario de biotecnología y Bioingeniería. Revista Balance y perspectiva Vol. 21; pag 275-277, Septiembre-Octubre de 2002. Desde Sitio Web: http://www.docstoc.com/Docs/Document-Detail-Google.aspx?doc_id=3245669
- [10] Sánchez Riaño, A.M.; Gutiérrez Morales, A. I.; Muñoz Hernández, J. A. y Rivera Barrero, C. A. Producción de bioetanol a partir de subproductos agroindustriales lignocelulósicos. Revista *Tumbaga* (2010), 5, 61-91. Consultado en febrero 2014. Disponible en: <http://revistas.ut.edu.co/index.php/tumbaga/article/view/194/163>
- [11] Sambrook J, and Russell D, W. (2001). **Molecular Cloning a laboratory manual third edition, volume 1.**

- [12] Timothy L. Bailey, Nadya Williams, Chris Misleh and Wilfred W. Li., **MEME: discovering and analyzing DNA and protein sequence motifs**. Institute of Molecular Bioscience, The University of Queensland, Nucleic Acids Research, 2006, Vol. 34, Web Server issue.
- [13] **Centro de Bioinformática del Instituto de Biotecnología, de la Universidad Nacional de Colombia**. Consultado en febrero de 2014. <http://bioinf.ibun.unal.edu.co/documentos/multAlignmentCBIB.pdf>
- [24] Timothy M. Rose, Jorja G. Henikoff, and Steven Henikoff. Nucleic Acids Res. 2003, 1 de Julio. **CODEHOP (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) PCR primer design**. Extraído el 3 de Diciembre de 2012 desde <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC168931/>

LITERATURA CONSULTADA

- Timothy L. Bailey and Michael Gribskov, "Combining evidence using p-values: application to sequence homology searches", Bioinformatics, 14(1):48-54, 1998. **MAST versión 4.9.0** (Release date: Wed Oct 3 11:07:26 EST 2012).
- Timothy L. Bailey and Charles Elkan, "Fitting a mixture model by expectation maximization to discover motifs in biopolymers", Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology, pp. 28-36, AAAI Press, Menlo Park, California, 1994. **MEME versión 4.4.0** (Release date: Tue Apr 27 10:09:30 EST 2010).
- Martínez V., Biaani, Solís B., José, Zamarripa C., Alfredo. Calidad agroindustrial del aceite de higuera (Ricinus communis L.) para la producción de bioenergéticos. México, D.F., 29 de noviembre de 2011

<http://www.bioenergeticos.gob.mx/index.php/prensa/noticia-actual/99-calidad-agroindustrial-del-aceite-de-higuerilla-ricinus-communis-l-para-la-produccion-de-bioenergeticos.html>. Consultado el 27 de Noviembre de 2012 a las 3:30 pm.

- Duran A. Jose, Retamal P. Norma, Moratiel Y. Ruben. Cultivos energéticos alternativos. El cultivo de ricino (*Ricinus communis*) en Andalucía: Una alternativa para la producción de Biodiesel. Desde sitio Web: <http://publicaciones.pucesi.edu.ec/documentos/libros/cultivos/39-58.pdf>. Consultado el 27 de Noviembre de 2012 a las 4:30 pm.
- Mejía I. Sandra. La Higuerilla. (*Ricinus communis*) Desde sitio web: <http://www.unalmed.edu.co/~crsequed/HIGUERILLA.htm>. Consultado el 27 de Noviembre de 2012 a las 2:30 pm.
- Fernández S.J. **Estructura y función de los hidratos de carbono: azúcares, almidón, glucógeno, celulosa.** Aplicación de Tecnología de los Alimentos, Dpt. Ingeniería Química, 26 de Abril del 2005. Desde sitio Web: <http://www.ual.es/docencia/jfernand/ATA/Tema5/Tema5-HidratosCarbono.pdf>
- Carballo A.R., **Hemicelulosas de maderas;** Departamento de Química, Facultad de Forestal y Agronomía, Universidad de Pinar del Río “Hermanos Saíz Montes de Oca”, Cuba
- Clarkson, Kathleen, A.; Wang, Zhi, C. y Morgan, Andrew, J. **Nueva composición de xilanas y método para su producción.** OFICINA ESPAÑOLA DE PATENTES Y MARCAS. 16 de Noviembre de 2005.
- Bonilla R, G; Souza, V; Eguiarte L, E. **METAGENÓMICA, GENÓMICA Y ECOLOGÍA MOLECULAR: LA NUEVA ECOLOGÍA EN EL BICENTENARIO DE DARWIN.** Tip Revista Especializada en Ciencias

Químico-Biológicas, Vol. 11, Núm. 1, junio, 2008, pp. 41-51 Universidad Nacional Autónoma de México.

- Howard R.L, Abotsi E, Rensburg E.L and Howard S. (2003, Diciembre). **Lignocelluloses biotechnology: issues of bioconversion and enzyme production.** African Journal of Biotechnology Vol. 2 (12), pp. 602-619. Extraído el 22 de septiembre de 2011 desde <https://tspace.library.utoronto.ca/bitstream/1807/1923/1/jb03112.pdf>
- Benavidez A; Benjumea P; Pashova V. (2007, Noviembre). **EL BIODIESEL DE ACEITE DE HIGUERILLA COMO COMBUSTIBLE ALTERNATIVO PARA MOTORES DIESEL.** Dyna, año 74, Nro. 153, pp 141-150. Medellín. ISSN 0012-7353. Extraído el 01 de marzo de 2012 desde <http://redalyc.uaemex.mx/src/inicio/ArtPdfRed.jsp?iCve=49615302>