



Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales  
Departamento de Química Biológica

## **Caracterización de enzimas activas sobre carbohidratos de un aislamiento de *Cellulomonas* sp. para degradación de biomasa lignocelulósica**

Tesis presentada para optar por el título de Doctor de la Universidad de Buenos Aires en el área Química Biológica

Lic. Florencia Elizabeth Piccinni

Directores de Tesis:

Dra. Eleonora Campos

Dr. Máximo L. Rivarola

Consejera de estudios:

Dra. Nancy I. López

Lugar de trabajo: Instituto de Biotecnología, CNIA-INTA

Fecha de defensa: 26 de marzo de 2019

Ciudad Autónoma de Buenos Aires

## **Agradecimientos**

A la Dra. Eleonora Campos, por todo el trabajo conjunto detrás de cada publicación y de esta Tesis y por su compromiso constante con el laboratorio y sus becarios.

Al Dr. Máximo Rivarola, por su invaluable ayuda con la "Matrix".

A la Universidad de Buenos Aires; a ANPCyT, INTA y CONICET.

A los Dres. Paola Talia, Nancy López, Diego Sauka, Pía Valacco, Fabiana Bigi y Silvina Wilkowsky por sus aportes en el área de su especialidad.

A los Dres. Oscar Taboga y Gabriela Conti, por su apoyo y sus consejos.

A la Dra. Sonia Wirth.

A todo el Laboratorio de Bioenergía. A mis compañeros Emiliano, Silvina, Ornella y Facundo, que, además de colegas, son amigos y maestros. También a Mechi y Laura N. A los virólogos Andrea, Rubén y Guido. A toda la gente de los grupos de "bacterias" (becarios, técnicos e investigadores).

Al Instituto de Biotecnología en su conjunto.

A Margarita Porta y Mónica Lahoz, que me alentaron siempre.

A María, muy especialmente.

A mis simbioses de la especie *Felis silvestris* subsp. *catus*.

A mis amigos de T.I., por acompañarme en cada etapa del doctorado. A mi amiga Marti. A Pao Manfredi.

A mi enorme y ruidosa familia (con todos mis ruidosos sobrinos), por llamarme "mi hermana, la científica" antes de que me recibiera de bióloga, o preguntarme si quiero ser doctora para curar gente. A Victoria.

A Juan, por su dulce forma de amar. A su familia, por recibirme con los brazos abiertos.

A mis padres, que me dieron todo su afecto y me transmitieron sus valores y su amor por la ciencia. Son el norte en mi vida. Les dedico esta Tesis y todo lo que soy.

<b>Índice general</b>	<b>I</b>
<b>Índice de tablas y figuras</b>	<b>IV</b>
<b>Abreviaturas y unidades de medida y concentración</b>	<b>VI</b>
<b>Resumen</b>	<b>X</b>
<b>Abstract</b>	<b>XIII</b>
<b>Publicaciones derivadas de este trabajo de Tesis</b>	<b>XVI</b>
<b>Financiamiento</b>	<b>XVII</b>

## Índice general

Introducción .....	2
1. Estructura y composición de la biomasa lignocelulósica .....	2
1.1. Pared vegetal.....	3
1.2. Componentes de la lignocelulosa.....	4
<b>1.2.a. Celulosa</b> .....	4
<b>1.2.b. Hemicelulosa</b> .....	6
<b>1.2.c. Lignina</b> .....	9
<b>1.2.d. Pectinas</b> .....	10
1.3. Bioetanol y fuentes de biomasa lignocelulósica .....	11
2. Deconstrucción enzimática de la biomasa lignocelulósica .....	12
2.1. Organismos productores de enzimas activas sobre lignocelulosa .....	12
2.2. Enzimas activas sobre lignocelulosa.....	16
<b>2.2.a. Celulasas</b> .....	16
<b>2.2.b. Hemicelulasas</b> .....	17
<b>2.2.c. Pectinasas</b> .....	18
<b>2.2.d. LPMOs</b> .....	19
<b>2.2.e. Ligninasas</b> .....	19
3. Aplicaciones industriales de las enzimas activas sobre lignocelulosa.....	19
3.1. Bioetanol de segunda generación .....	19
3.2. Otras industrias .....	22
<b>3.2.a. Fabricación de papel</b> .....	22
<b>3.2.b. Industria alimenticia humana y alimentación animal</b> .....	22
<b>3.2.c. Industria textil y detergentes para ropa</b> .....	24
4. <i>Cellulomonas</i> sp.....	24
Hipótesis y objetivos .....	26
Materiales y métodos.....	29
1. Medios de cultivo y soluciones .....	30
2. Biomosas: pretratamiento y composición.....	32
3. Cultivo del aislamiento <i>Cellulomonas</i> sp. B6 y obtención de extractos enzimáticos.....	33
4. Crecimiento en papel de filtro.....	33
5. Microscopía electrónica de barrido .....	34

6. Determinación del metabolismo de hidratos de carbono de <i>Cellulomonas</i> sp B6 .....	34
7. Determinación de la composición de ácidos grasos .....	35
8. Actividades enzimáticas del sobrenadante de cultivo .....	35
8.1. Actividades endoglucanasa (CMCasa) y xilanasas .....	35
8.2. Actividad celulasa total.....	36
8.3. Actividades $\beta$ -glucosidasa, $\beta$ -xilosidasa y celobiohidrolasa .....	36
9. Zimogramas .....	37
10. Secuenciación, ensamblado y anotación del genoma de <i>Cellulomonas</i> sp. B6 .....	38
11. Análisis filogenético en base a la secuencia de la subunidad 16S del ARN ribosomal.....	38
12. Comparación <i>in silico</i> de genomas .....	39
13. Identificación de enzimas activas sobre carbohidratos .....	39
14. Análisis del secretoma de <i>Cellulomonas</i> sp. B6.....	40
14.1 Preparación de las muestras .....	40
14.2. Detección de proteínas por espectrometría de masas (LC MS/MS) .....	40
15. Expresión recombinante de GH10xynC.....	41
15.1. Amplificación .....	41
15.2. Clonado.....	42
15.3. Expresión recombinante y purificación.....	43
15.4. SDS-PAGE.....	43
16. Análisis de la secuencia de GH10XynC .....	43
17. Análisis filogenético de GH10XynC.....	44
18. Modelado molecular de GH10XynC .....	44
19. Caracterización bioquímica de la xilanasas rGH10XynC .....	45
19.1. Actividades enzimáticas .....	45
19.2. Cinética de reacción .....	46
20. Hidrólisis de biomasa .....	46
21. Cuantificación de productos de reacción por cromatografía de alta precisión (HPLC).....	47
22. Detección de productos de reacción por cromatografía en capa delgada .....	47
Resultados .....	48
Parte 1. Actividad celulolítica y hemicelulolítica de extractos enzimáticos de <i>Cellulomonas</i> sp. B6.....	49
1.1. Selección del medio de cultivo para maximizar las actividades celulolíticas y xilanolíticas de los extractos enzimáticos de <i>Cellulomonas</i> sp. B6 .....	51
1.2. Caracterización de las actividades enzimáticas de <i>Cellulomonas</i> sp. B6 .....	52
1.3. Ensayos de actividad de los extractos enzimáticos sobre biomasa lignocelulósica.....	55
Parte 2. Caracterización genómica, bioquímica y proteómica del aislamiento <i>Cellulomonas</i> sp. B6....	57
2.1. Secuenciación del genoma y caracterización filogenética .....	58

<b>2.1.a. Genómica comparativa</b> .....	59
<b>2.1.b. Caracterización morfológica</b> .....	62
<b>2.1.c. Caracterización bioquímica</b> .....	62
2.2. Anotación del genoma de <i>Cellulomonas</i> sp. B6 .....	65
2.3. Análisis de secuencias codificantes para enzimas activas sobre polisacáridos (CAZimas) y construcción del CAZoma y secretoma teóricos .....	67
2.4. Análisis del secretoma experimental de <i>Cellulomonas</i> sp B6 .....	78
Parte 3. Expresión recombinante de una xilanasa GH10 de <i>Cellulomonas</i> sp. B6 y su aplicación a la deconstrucción de biomasa lignocelulósica .....	87
3.1. Selección de una xilanasa de interés y expresión recombinante.....	88
<b>3.1.a. Análisis de secuencia</b> .....	88
<b>3.1.b. Modelado de la estructura de GH10XynC</b> .....	90
<b>3.1.c. Expresión recombinante y purificación de rGH10XynC</b> .....	91
3.2. Caracterización de la actividad enzimática de rGH10XynC.....	92
<b>3.2.a. Identificación de los sustratos de reacción y efecto del pH y la temperatura sobre la actividad</b> .....	92
<b>3.2.b. Cinética de reacción sobre xilano</b> .....	94
<b>3.2.c. Efecto de cationes divalentes sobre la actividad</b> .....	94
<b>3.2.d. Mecanismo de actividad de la enzima</b> .....	95
3.3. Evaluación de la actividad de rGH10XynC en procesos de bioconversión de biomasa lignocelulósica .....	97
Discusión .....	100
Bibliografía.....	110
Anexo.....	123

## Índice de tablas y figuras

<b>Figura I1.</b> Estructura esquemática de la pared vegetal primaria y secundaria	4
<b>Figura I2.</b> Estructura esquemática de las microfibrillas de celulosa que componen la pared vegetal.	5
<b>Figura I3.</b> Estructura esquemática del glucuronoarabinoxilano y el glucuronoxilano.	7
<b>Figura I4.</b> Estructura esquemática del galacto(gluco)manano.	8
<b>Figura I5.</b> Estructura esquemática de los xiloglucanos.	8
<b>Figura I6.</b> Estructura esquemática de los glucanos mixtos.	9
<b>Figura I7.</b> Alcoholes fenólicos que dan origen a las diferentes subunidades de la lignina.	10
<b>Figura I8.</b> Estructura esquemática del ramnogalacturano.	11
<b>Tabla I1.</b> Composición porcentual de hemicelulosa, celulosa y lignina de biomasa lignocelulósica de distinto origen (expresada como % p/p del peso seco)	12
<b>Figura I9.</b> Estrategias de hongos y bacterias para la degradación enzimática de la pared vegetal.	14
<b>Figura I10.</b> Organismos que degradan lignocelulosa por acción de enzimas endógenas y/o exógenas, producidas por simbiosis	15
<b>Figura I11.</b> Degradación de celulosa por acción de glicosil hidrolasas.	16
<b>Figura I12.</b> Esquema de la hidrólisis enzimática de xilano.	18
<b>Figura I13.</b> Esquema general de la obtención de bioetanol y biocombustibles.	20
<b>Tabla M1.</b> Medios de cultivo y soluciones utilizados en el trabajo de Tesis.	30
<b>Tabla M2.</b> Composición de las biomásas empleadas en el trabajo de Tesis.	32
<b>Tabla M3.</b> Plásmidos utilizados en el trabajo de Tesis.	41
<b>Figura R1.</b> Crecimiento de <i>Cellulomonas</i> sp. B6 en papel de filtro en condiciones de aerobiosis y anaerobiosis.	50
<b>Figura R2.</b> Actividades enzimáticas de los extractos extracelulares de <i>Cellulomonas</i> sp. B6 en LB y MM suplementados con CMC o RAC.	51
<b>Figura R3.</b> Zimografía de sobrenadante de cultivo de <i>Cellulomonas</i> sp. B6 cultivado en MM-CMC o MM-RAC, en CMC 0,2% y xilano 0,1%.	52
<b>Figura R4.</b> Perfil de las actividades xilanasas y CMCasa del extracto extracelular de <i>Cellulomonas</i> sp. B6 en función del pH y la temperatura. Estabilidad térmica a 45 y 50°C de las actividades CMCasa y xilanasas.	53
<b>Tabla R1.</b> Actividades enzimáticas determinadas en los extractos extra e intracelular de <i>Cellulomonas</i> sp. B6 (en MM-RAC).	54
<b>Figura R5.</b> Productos de reacción sobre CMC y xilano de haya, usando sobrenadante de cultivo de <i>Cellulomonas</i> sp. B6.	55
<b>Figura R6.</b> Productos de hidrólisis de PTE usando el sobrenadante de <i>Cellulomonas</i> sp. (de MM-RAC) como extracto enzimático.	55
<b>Tabla R2.</b> Xilobiosa (X2) obtenida a partir de xilano contenido en biomasa, cuantificada por HPLC, usando los extractos extra y/o intracelular de <i>Cellulomonas</i> sp. B6.	56
<b>Tabla R3.</b> Metodología y resultados de la secuenciación por Illumina Miseq y el ensamblado del genoma de <i>Cellulomonas</i> sp. B6.	58
<b>Figura R7.</b> Filograma construido en base a la secuencia de la subunidad ribosomal 16S del aislamiento B6 y las especies del género <i>Cellulomonas</i> .	59
<b>Tabla R4.</b> Comparación a nivel genómico de <i>Cellulomonas</i> sp. B6 y de <i>C. flavigena</i> ATCC 482, <i>C. persica</i> ATCC 700642 y <i>C. fimi</i> ATCC 484.	60
<b>Figura R8.</b> Micrografías electrónicas de barrido de <i>Cellulomonas</i> sp. B6 sobre PF.	62
<b>Tabla R5.</b> Resultados de API 50 CH de <i>Cellulomonas</i> sp. B6 comparados con la información publicada para <i>C. flavigena</i> , <i>C. persica</i> y <i>C. fimi</i> .	63
<b>Tabla R6.</b> Resultados de la anotación del genoma de <i>Cellulomonas</i> sp. B6 por NCBI PGAP y RAST.	66

<b>Figura R9.</b> Número de proteínas predichas asignadas a cada subsistema de RAST.	66
<b>Figura R10.</b> CAZimas de <i>Cellulomonas</i> sp. B6, <i>C. flavigena</i> ATCC 482 y <i>C. fimi</i> ATCC 484, clasificadas según su dominio catalítico.	68
<b>Tabla R7.</b> CAZimas predichas como extracelulares por vías del péptido señal en el genoma de <i>Cellulomonas</i> sp. B6.	69
<b>Tabla R8.</b> CAZimas potencialmente secretadas por vías independientes de péptido señal.	70
<b>Tabla R9.</b> CAZimas predichas como intracelulares.	70
Tabla R10.	74
<b>Figura R11.</b> Enzimas y transportadores involucrados en la utilización de xilósidos codificados en el genoma y predichos en base a la anotación por RAST, NCBI y dbCAN.	77
<b>Figura R12.</b> Comparación de GH, AA y PL con actividad sobre celulosa y hemicelulosa identificadas en <i>Cellulomonas</i> sp. B6, <i>C. fimi</i> y <i>C. flavigena</i> .	78
<b>Tabla R11.</b> Proteínas totales y actividades CMCasa y xilanasas de los sobrenadantes de cultivo de <i>Cellulomonas</i> sp. B6 en CMC, RAC, PTE y sacarosa para ser analizados por espectrometría de masas.	79
<b>Tabla R12.</b> CAZimas y proteínas extracelulares involucradas en el transporte de azúcares identificadas por espectrometría de masas en sobrenadantes de cultivo de <i>Cellulomonas</i> sp. B6 en CMC, RAC y PTE.	80
<b>Figura R13.</b> CAZimas identificadas por espectrometría de masa en los extractos extracelulares de <i>Cellulomonas</i> sp B6, predichas como extracelulares (secretadas tanto por la vía del péptido señal como por vías no clásicas). CAZoma extracelular predicho (sólo por la vía de péptido señal). Diagrama de Venn representando el número de CAZimas predichas como extracelulares detectadas en los sobrenadantes de CMC, RAC y PTE.	82
<b>Tabla R13.</b> Abundancia relativa de las proteínas secretadas de <i>Cellulomonas</i> sp. B6.	83
<b>Figura R14</b> Modelo de degradación enzimática de celulosa, hemicelulosas, quitina y pectinas de <i>Cellulomonas</i> sp. B6.	86
<b>Figura R15.</b> Árbol filogenético del dominio GH10 de rGH10XynC.	89
<b>Figura R16.</b> Árbol filogenético del CBM2 de rGH10XynC.	89
<b>Figura R17.</b> Entorno genómico de <i>xynC</i> .	90
<b>Figura R18.</b> Modelo de la estructura tridimensional de la proteína GH10XynC correspondiente al modelo de mejor ajuste.	91
<b>Figura R19.</b> Electroforesis en gel de agarosa 0,8% de los productos de amplificación de <i>xynC</i> y SDS-PAGE 10% de las proteínas correspondientes al cultivo de <i>E. coli</i> Rossetta pET28a-XynC sin inducir e inducido, la fracción soluble, el percolado, el lavado y las fracciones de elución de la proteína rGH10XynC purificada 1-4.	92
<b>Figura R20.</b> Actividad de rGH10XynC. Efecto de: a) pH y b) temperatura; c) estabilidad térmica de rGH10XynC a 40, 45 y 50°C.	93
<b>Figura R21.</b> Ajuste de la cinética de reacción de rGH10XynC sobre xilano al modelo de Michaelis-Menten.	94
<b>Tabla R14.</b> Efecto del agregado de cationes divalentes sobre la actividad de rGH10XynC.	95
<b>Figura R22.</b> Productos de reacción de rGH10XynC sobre glucuronoxilano (de haya). Estructura del glucuronoxilano de haya (Megazymes).	95
<b>Figura R23.</b> Detección de los productos de reacción sobre a) XOS (X3-X6) y b) X2	96
<b>Figura R24.</b> Modelo del mecanismo propuesto de acción de rGH10XynC sobre xilooligosacáridos.	97
<b>Figura R25.</b> Actividad de la enzima rGH10XynC en condiciones de SSCF.	98
<b>Figura R26.</b> Cromatograma de los productos de reacción de rGH10XynC sobre PCE detectados por HPLC.	98
<b>Tabla R15.</b> Cuantificación de los productos de reacción de rGH10XynC PCE y cálculo del porcentaje de conversión de xilano a X1, X2 y X3.	99

## Abreviaturas y unidades de medida y concentración

Las siglas corresponden a las más usadas en el ámbito científico argentino, independientemente del idioma de origen. Los nombres de instituciones, programas o bases de datos se indican en el idioma original.

**% p/p:** porcentaje de peso en peso

**% p/v:** porcentaje de peso en volumen

**% v/v:** porcentaje de volumen en volumen

**<:** menor

**>:** mayor

**∑:** sumatoria

**≤:** mayor o igual

**≥:** menor o igual

**°C:** grados centígrados

**µg:** microgramos

**1G:** bioetanol de primera generación

**2G:** bioetanol de segunda generación (celulósico)

**Å:** ángstrom

**AA:** enzima con actividad auxiliar

**AAI:** *Average Aminoacid Identity* (identidad aminoacídica promedio)

**ABC:** *ATP Binding Cassette (transporter)* (transportador del sistema *cassette* de unión a ATP)

**ADN:** ácido desoxirribonucleico

**AFEX:** *Ammonia Fiber EXpansion* (expansión de fibras con amoníaco)

**ANI:** *Average Nucleotide Identity* (identidad nucleotídica promedio)

**ARN:** ácido ribonucleico

**ARNr:** ARN ribosomal

**ARNt:** ARN de transferencia

**BLAST:** *Basic Local Alignment Search Tool*

**CAZima:** *Carbohydrate-Active enZYme* (enzima activa sobre carbohidratos)

**CAZoma:** todas las CAZimas codificadas en el genoma

**CAZy:** base de datos *Carbohydrate-Active enZYmes*

**CBM:** *Carbohydrate Binding Domain* (dominio de unión a carbohidratos)

**CBP:** *Consolidated Bio-Process* (bioproceso consolidado)

**CD:** *Catalytic Domain* (dominio catalítico)

**CDD:** *Conserved Domains Database* (base de datos de dominios conservados)

**CE:** carboxiesterasa

**cm:** centímetros

**CMC:** carboximetilcelulosa

**COS:** celooligosacáridos

**dbCAN:** *DataBase for automated Carbohydrate-active enzyme Annotation* (base de datos para la anotación automática de enzimas activas sobre carbohidratos)

**DDH:** *DNA-DNA Hybridization* (hibridización de genomas *in vitro*)  
**DNS:** ácido dinitrosalicílico  
**DO:** Densidad Óptica  
**DP:** *Degree of Polymerization* (grado de polimerización)  
**SD:** *Scaffold Domain* (dominio *scaffold*, o de anclaje)  
**DTT:** ditioneotrol  
**EC:** *Enzyme Commission number* (número de comisión de actividad enzimática)  
**EDTA:** ácido etilendiamino tetra acético  
**emPAI:** *Exponentially Modified Protein Abundance Index* (índice de abundancia modificado)  
**ENA:** *European Nucleotide Archive*  
**FOS:** fructooligosacáridos  
**g:** gramos  
**G1:** glucosa  
**G2:** celobiosa  
**G3:** celotriosa  
**G4:** celotetraosa  
**GBDP:** *Genome BLAST Distance Phylogeny*  
**GDDH:** *GBDP-based DDH* (hibridización de genomas basada en GBDP)  
**GH:** glicosil hidrolasa  
**GOS:** galactooligosacáridos  
**GT:** glicosil transferasa  
**HMM:** *Hidden Markov Models* (modelos ocultos de Markov)  
**HPLC:** *High Performance Liquid Chromatography* (cromatografía líquida de alta eficiencia)  
**hs:** horas  
**HSPs:** *High Scoring Pairs* (pares de alto puntaje)  
**IPTG:** isopropil- $\beta$ -D-1-tiogalactopiranosido  
**IUBMB:** *International Union of Biochemistry and Molecular Biology*  
**IUPAC:** *International Union of Pure and Applied Chemistry*  
**kb:** kilobases  
**kDa:** kilo Dalton  
**K<sub>M</sub>:** constante de Michaelis- Menten  
**l:** litro  
**LB:** *Lysogeny Broth* (caldo de lisogenia)  
**LPMO:** *Lytic Polysaccharide Mono-Oxygenase* (monooxigenasa lítica de polisacáridos)  
**M:** concentración molar  
**Mb:** megabases  
**mg:** miligramos  
**min:** minutos  
**ml:** mililitros  
**mM:** concentración milimolar  
**MM:** medio salino mínimo

**MMD:** marlo de maíz dulce  
**MOS:** manooligosacáridos  
**mUI:** miliunidades internacionales de actividad enzimática  
**NCBI:** *National Centre for Biotechnology Information*  
**NDO:** *Non-Digestible Oligosaccharide* (oligosacárido no digerible)  
**ng:** nanogramos  
**NGS:** *Next Generation Sequencing* (secuenciación de nueva generación)  
**nl:** nanolitros  
**NREL:** *National Renewable Energy Laboratory*  
**nRIU:** *Normalized Refractive Index Units* (unidades de índice de refracción normalizadas)  
**ON:** *OverNight* (toda la noche, 16 horas aprox.)  
**PASC:** *Phosphoric Acid Swollen Cellulose* (celulosa amorfa regenerada)  
**pb:** pares de bases  
**PCE:** paja de cebada extrusada  
**PCR:** *Polymerase Chain Reaction* (reacción en cadena de la polimerasa)  
**PD:** papel de descarte  
**PDB:** *Protein Data Bank*  
**PF:** papel de filtro  
**PGAP:** *Prokaryotic Genome Annotation Pipeline*  
**pl:** punto Isoeléctrico  
**PL:** pectato liasa  
**PM:** tamaño molecular  
**PNP:** paranitrofenol  
**pNP-C:** *p*-nitrofenil β-D-celobiósido  
**pNP-G:** *p*-nitrofenil β-D-glucopiranosido  
**pNP-X:** *p*-nitrofenil β-D-xilopiranosido  
**PTE:** paja de trigo extrusada  
**RAC:** residuo agrícola de caña de azúcar  
**RAST:** *Rapid Annotation using Subsystems Technology*  
**RMSD:** *Root Mean Square Deviation* (raíz cuadrada de la desviación media de las posiciones atómicas por superposición de proteínas)  
**rpm:** revoluciones por minuto  
**SDS:** *Sodium Dodecyl Sulfate* (dodecil sulfato de sodio)  
**SDS-PAGE:** *Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis* (electroforesis en geles de poliacrilamida con SDS)  
**SEM:** *Scanning Electron Microscopy* (microscopía electrónica de barrido)  
**SOB:** *Super-Optimum Broth* (caldo súper-óptimo)  
**SOC:** *Super-Optimum broth with Catabolite Repression* (caldo súper-óptimo con represión por catabolitos)  
**SSCF:** *Simultaneous Saccharification and Co-Fermentation* (sacarificación simultánea a la co-fermentación de glucosa y xilosa)

**SSF:** *Simultaneous Saccharification and Fermentation* (sacarificación y fermentación simultáneas)

**TAE:** *buffer* Tris-acetato-EDTA

**TAT:** *Twin Arginine Translocation (pathway)* (sistema de translocación TAT de proteínas)

**TBF:** *Transformation Buffer (buffer de transformación)*

**TLC:** *Thin Layer Chromatography* (cromatografía en capa delgada)

**UFC:** unidades Formadoras de Colonias

**UI (o U):** unidades internacionales de actividad enzimática

**v:** versión

**V<sub>máx</sub>:** velocidad máxima de una reacción enzimática

**X1:** xilosa

**X2:** xilobiosa

**X3:** xilotriosa

**X4:** xilotetraosa

**X5:** xilopentaosa

**X6:** xilohexaosa

**xg:** unidades de centrifugación

**X-gal:** 5-bromo-4-cloro-3-indolil- $\beta$ -D-galactósido

**XOS:** xilooligosacáridos

**$\mu$ g:** microgramos

**$\mu$ l:** microlitros

## RESUMEN

### Caracterización de enzimas activas sobre carbohidratos de un aislamiento de *Cellulomonas* sp. para degradación de biomasa lignocelulósica

La lignocelulosa es el componente mayoritario de la biomasa vegetal. Está formada por celulosa embebida en una matriz de hemicelulosa y lignina. Los microorganismos celulolíticos aerobios secretan un conjunto de enzimas que les permiten degradar los polisacáridos estructurales (celulosa y hemicelulosas) para utilizarlos como fuente de carbono. Estas enzimas son ampliamente estudiadas y tienen gran importancia, particularmente por su potencial aplicación en el aprovechamiento de biomasa no-alimenticia, como residuos agro-forestal industriales, principalmente para obtener biocombustibles o alimento animal.

El aislamiento bacteriano *Cellulomonas* sp. B6, obtenido a partir de un consorcio celulolítico de suelo forestal, fue capaz de crecer en numerosos sustratos celulósicos, incluyendo biomasa. Los sobrenadantes de cultivo presentaron principalmente actividades endo y exoglucanasa y xilanasas y en los extractos intracelulares se detectaron actividades  $\beta$ -glucosidasa y  $\beta$ -xilosidasa. El crecimiento en biomasa lignocelulósica (residuo de cosecha de caña de azúcar, RAC) resultó en la mayor actividad enzimática en el sobrenadante (0,2 y 0,6 UI/ml para las actividades endoglucanasa y xilanasas, respectivamente). Estas actividades se mantuvieron en un rango de pH de 5 a 8 y de temperatura de 40 a 55°C, indicando la potencialidad del extracto para aplicaciones de degradación de biomasa en esas condiciones. Por zimografía, estas actividades se correlacionaron con proteínas presentes en el sobrenadante de cultivo.

El genoma de *Cellulomonas* sp. B6 fue secuenciado por Illumina MiSeq y ensamblado en 279 contigs, resultando en una cobertura de 83X. El tamaño total del genoma se estimó en 4 Mb y el contenido de GC fue de 75,1%, similar a lo reportado para cepas de referencia del género. Por análisis filogenético a partir de la secuencia de la subunidad 16S de ARN ribosomal, *Cellulomonas* sp. B6 se encuentra relacionado con las especies *Cellulomonas flavigena* y *Cellulomonas persica*, aunque presenta características genotípicas y metabólicas distintivas, lo que sugiere que podría tratarse de una nueva especie. Mediante anotación del genoma *Cellulomonas* sp. B6 por las

plataformas NCBI y RAST, seguido de análisis por la plataforma dbCAN, se identificaron 3443 secuencias codificantes para proteínas, de las cuales 205 corresponden a enzimas activas sobre carbohidratos (CAZimas). De las CAZimas predichas, 91 son glicosil hidrolasas (GH), de las cuales 32 poseen secuencia codificante para péptido señal, lo que indica que podrían ser extra-celulares.

Para identificar las proteínas responsables de la actividad enzimática previamente observada, se analizó por espectrometría de masas el secretoma completo (conjunto de proteínas extracelulares) en sobrenadantes de cultivo en celulosa (carboximetilcelulosa, CMC), residuo agrícola de caña de azúcar molido (RAC) o paja de trigo pretratada por extrusión (PTE), como únicas fuente de carbono. En el sobrenadante de cultivo en CMC se identificaron 2 exoglucanasas (GH6 and GH48) y tres xilanasas GH10, mientras que el crecimiento en biomasa (RAC o PTE) resultó en la identificación de 22 CAZimas, incluyendo endo- y exoglucanasas (2 GH6, 2 GH9 y 1 GH48), xilanasas (7 GH10 y 1 GH11), una xiloglucanasa (GH74), una arabinofuranosidasa/  $\beta$ -xilosidasa (GH43), una  $\beta$ -glucosidasa (GH3) y una  $\alpha$ -glucuronidasa (GH62), entre otras. Adicionalmente, se identificó una oxigenasa lítica de polisacáridos (LPMO) AA10, potencialmente involucrada en la deconstrucción de celulosa cristalina por un mecanismo oxidativo. La diversidad de GH10 secretadas (todas las extracelulares codificadas en el genoma) remarca la importancia de estas enzimas en la deconstrucción de biomasa.

Para comenzar a evaluar la contribución individual de las xilanasas de la familia GH10 a la actividad del extracto, una GH10 con un módulo de unión a sustrato CBM2, a la que denominamos GH10XynC, fue expresada de manera recombinante en *Escherichia coli* y purificada en forma soluble. La enzima recombinante presentó actividad endoxilanasas (EC 3.2.8.1) en el rango de 300 UI/mg, sobre xilano de abedul y sobre arabinoxilano. No presentó actividad sobre celulosa, confirmando que es una xilanasas libre de actividad celulolítica. Los productos de hidrólisis de xilano fueron xilobiosa (X2) y xilosa (X1) en menor proporción y la actividad fue óptima a 50°C, en un rango de pH de 5 a 7,5. Ensayos de hidrólisis con GH10XynC de paja de cebada y de trigo y el marlo de maíz dulce (MMD) (pre-tratados por extrusión) resultaron en la conversión a xilooligosacáridos, X2 y X1, demostrando la habilidad de la enzima de actuar sobre el xilano contenido en la biomasa.

En conjunto, estos resultados demuestran que *Cellulomonas* sp. B6 constituye una fuente de enzimas con potencial aplicación agro-industrial en el aprovechamiento de biomasa lignocelulósica. Es más, la enzima GH10XynC podría ser aplicada en la utilización de xilanos para las industrias de biocombustibles, prebióticos y alimento animal.

**Palabras clave:** CELLULOMONAS, BIOMASA LIGNOCELULOSICA, ENZIMAS, CAZIMAS, GENOMA, SECRETOMA, GLICOSIL HIDROLASAS, XILANASA, GH10.

## ABSTRACT

### **Characterization of carbohydrate active enzymes of *Cellulomonas* sp. B6 with application in deconstruction of lignocellulosic biomass.**

Lignocellulose is the major component of plant biomass. It consists of cellulose embedded in a matrix of hemicellulose and lignin. Aerobic cellulolytic microorganisms secrete a set of enzymes that allow them to degrade structural polysaccharides (cellulose and hemicelluloses) to use them as a carbon source. These enzymes are widely studied and have great importance, particularly for their potential application in the use of non-food biomass, such as agro-industrial waste, mainly to obtain biofuels or animal feed.

The bacterial isolate *Cellulomonas* sp. B6 was obtained from a forest soil cellulolytic consortium. *Cellulomonas* sp. B6 was able to grow on numerous cellulosic substrates, including biomass. The culture supernatants showed mainly endo and exo-glucanase and xylanase activities and  $\beta$ -glucosidase and  $\beta$ -xylosidase activities were detected in the intracellular extracts. Growth in lignocellulosic biomass (sugar cane harvest residue, RAC) resulted in the highest enzymatic activity in the supernatant (0.2 and 0.6 IU/ml for the endoglucanase and xylanase activities, respectively). These activities were maintained in a pH range of 5 to 8 and a temperature of 40 to 55°C, indicating the potential of the extract for applications of biomass degradation under these conditions. By zymography, these activities correlated with proteins present in the culture supernatant.

The genome of *Cellulomonas* sp. B6 was sequenced by Illumina MiSeq and assembled into 279 contigs, resulting in an 83X coverage. The total genome size was estimated at 4 Mb and the GC content was 75.1%, similar to that reported for reference strains of the genus. By phylogenetic analysis from the sequence of the 16S subunit of ribosomal RNA, it is related to *Cellulomonas flavigena* and *Cellulomonas persica* species, although *Cellulomonas* sp. B6 presents distinctive genotypic and metabolic characteristics, suggesting that it could be a new species. Annotation of the genome from *Cellulomonas* sp. B6 by the NCBI and RAST platforms, followed by analysis using the dbCAN platform, resulted in the identification of 3443 protein coding sequences, of which 205 corresponded to carbohydrates active enzymes (CAZymes).

Of the predicted CAZymes, 91 were glycosyl hydrolases (GH), of which 32 possessed, signal peptide coding sequence, which indicated that they could be extracellular.

To identify the proteins responsible for the enzymatic activity previously observed, the complete secretome (set of extracellular proteins) was analyzed by mass spectrometry in supernatants from cultures in cellulose (carboxymethylcellulose, CMC), in agricultural residue of ground sugar cane (RAC) or wheat straw pretreated by extrusion (PTE), as the sole carbon sources. In the CMC-culture supernatant, 2 exoglucanases (GH6 and GH48) and three GH10 xylanases were identified, whereas growth in biomass (RAC or PTE) resulted in the identification of 22 CAZymes, including endo- and exo-glucanases (2 GH6, 2 GH9 and 1 GH48), xylanases (7 GH10 and 1 GH11), a xyloglucanase (GH74), an arabinofuranosidase /  $\beta$ -xylosidase (GH43), a  $\beta$ -glucosidase (GH3) and an  $\alpha$ -glucuronidase (GH62), among others. Additionally, a lytic polysaccharide monooxygenase (LPMO) AA10, potentially involved in cellulose deconstruction by an oxidative mechanism, was identified. The diversity of secreted GH10 (all of the extracellular ones encoded in the genome) highlights the importance of these enzymes in the deconstruction of biomass.

To evaluate the individual contribution of the xylanases from the GH10 family to the activity of the extract, a GH10 xylanase with a CBM2substrate binding module, which we called GH10XynC, was expressed as recombinant in *Escherichia coli* and purified in soluble form. The recombinant enzyme showed endoxylanase activity (EC 3.2.8.1) in the range of 300 IU/mg, on birch xylan and on arabinoxylan. It did not present activity on cellulose, confirming that it is a xylanase free of cellulolytic activity. The hydrolysis products of xylan were xylobiose (X2) and xylose (X1) in lower proportion and the activity was optimal at 50°C, in a pH range of 5 to 7.5. Hydrolysis assays with GH10XynC of barley and wheat straw (BS, WS) and sweet corn cob (SCC) (pre-treated by extrusion) resulted in the conversion to xylooligosaccharides, X2 and X1, demonstrating the enzyme's ability to act on xylan content in biomass.

Taken together, these results show that *Cellulomonas* sp. B6 is a source of enzymes with potential agro-industrial application in the use of lignocellulosic biomass. Moreover, GH10XynC could have applications in the bioconversion of xylans for biofuels, prebiotics and feed industries.

**Keywords:** CELLULOMONAS, LIGNOCELLULOSIC BIOMASS, ENZYMES, CAZYMES, GENOME, SECRETOME, XYLANASE, GLYCOSYL HYDROLASES, GH10.

### Publicaciones derivadas de este trabajo de Tesis

- Piccinni FE, Ontañón OM, Ghio S, Sauka DH, Talia PM, Rivarola ML, Valacco MP, Campos E (2018). Secretome profile of *Cellulomonas* sp. B6 growing on lignocellulosic substrates. *J Appl Microbiol*, 126: 811—25. doi: 10.1111/jam.14176.
- Piccinni FE, Murua Y, Ghio S, Talia PM, Rivarola ML, Campos E (2016). Draft genome sequence of cellulolytic and xylanolytic *Cellulomonas* sp. strain B6 isolated from subtropical forest soil. *Genome Announc*, 4(4): e00891-16.

## **Financiamiento**

- Beca inicial de FONCyT (2014-2016) (PICT 2014-0791)
- Beca interna de finalización de doctorado de CONICET (2016-2018).
- PICT CABBIO 2016-4695 (2016-2018). IR: Dra. Eleonora Campos.
- PICT 2014-0791 (2015-2018). IR: Dra. Eleonora Campos.
- PIP 11220150100121CO (2016-2019): IR: Dra. Eleonora Campos.
- PICT 2011-2735 (2012-2016) IR: Dra. Eleonora Campos.
- INTA 1130034 (2013-2019). Coordinador: Dr. Daniel Grasso.
- BABET-REAL5 (2016-2019). Coordinador: Dr. Daniel Grasso

# Introducción

## **1. Estructura y composición de la biomasa lignocelulósica**

El término biomasa lignocelulósica se refiere a toda aquella biomasa vegetal que no es utilizada como fuente de alimento y comprende residuos de la agroindustria, cultivos energéticos, como pasturas y malezas y desechos sólidos urbanos (Den y col., 2018; Robak y Balcerak, 2018). Es un material abundante y renovable por lo que su aprovechamiento es materia de estudio a nivel global, especialmente en países con fuerte actividad agrícola y forestal (Arévalo-Gallegos y col., 2017). Una de las utilidades más investigadas en los últimos años es su aprovechamiento para la producción de etanol, dado su alto contenido de hidratos de carbono fermentables y la gran disponibilidad de biomasa (Lennartson y col., 2014). A diferencia de los biocombustibles de primera generación, derivados principalmente de almidón (como el maíz y la papa) y de la sacarosa presente en caña de azúcar, sorgo azucarero o remolacha, los biocombustibles derivados de lignocelulosa, llamados de segunda generación, provienen de fuentes no alimenticias y de bajo valor (Wyman, 2014; Lennartson y col., 2014).

Todas las características mencionadas convierten a la lignocelulosa en un recurso potencialmente aprovechable para el complemento o reemplazo de los combustibles fósiles basados en el petróleo (Arévalo-Gallegos y col., 2017).

Sin embargo, la utilización de esta biomasa para obtener bioetanol tiene aún retos técnicos para ser económicamente viable. Uno de ellos es la conversión de los polisacáridos estructurales, celulosa y hemicelulosa, en azúcares monoméricos fermentables, glucosa y xilosa, proceso que se conoce como sacarificación. Se ha establecido que una de las vías más eficiente de deconstrucción de materiales lignocelulósicos es la utilización de enzimas (glicosil hidrolasas y enzimas oxidativas) ya que permite maximizar la obtención tanto de hexosas (glucosa) como de pentosas (xilosa), para su posterior fermentación, lo cual es clave para el aprovechamiento completo de la biomasa y la sustentabilidad del proceso (Ballesteros, 2010).

En la naturaleza, la deconstrucción de lignocelulosa es llevada a cabo, en gran medida, por los hongos y bacterias saprófitas, capaces de producir enzimas activas sobre los componentes de la biomasa (Cragg y col., 2015). Estos organismos, o sus enzimas, son aprovechados en la industria de los biocombustibles para la sacarificación de la celulosa y hemicelulosa. Asimismo, muchas de las enzimas activas sobre lignocelulosa son aplicables en

otras industrias, como la alimenticia, la textil y del papel (Chakdar y col., 2016; Kumar y col., 2018; Walia y col., 2017).

La lignocelulosa es el componente estructural principal de la pared vegetal de plantas terrestres y el recurso renovable más abundante en la tierra. Su composición varía según la especie pero, a grandes rasgos, la componen los polisacáridos celulosa y hemicelulosa y la lignina, que constituyen el 20–50%, 15–35% y 10–30% del peso seco de la planta, respectivamente (Payne y col., 2015). Estos componentes interactúan a través de enlaces covalentes y no covalentes formando una estructura altamente recalcitrante a la degradación (Saini y col., 2015).

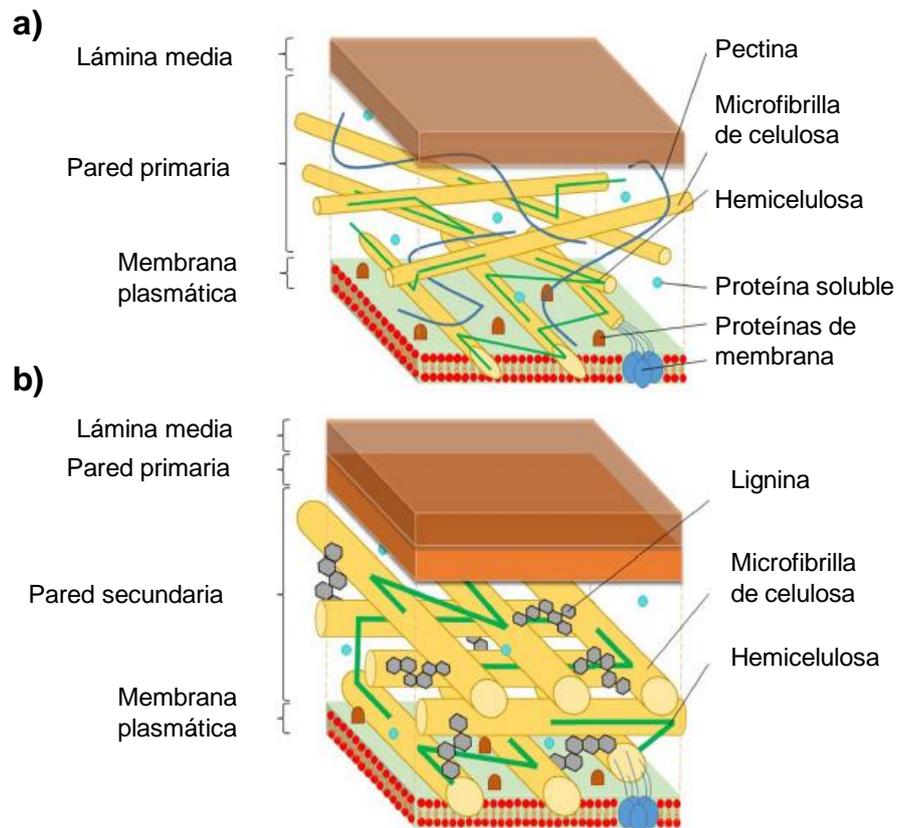
### **1.1. Pared vegetal**

La pared vegetal proporciona sostén a la planta y da forma a los tejidos y, además, constituye un mecanismo de defensa frente a ciertas condiciones ambientales, a los organismos herbívoros y a la acción de organismos patógenos (Rennie y Scheller, 2014; Scheller y Ulskov, 2010). Comienza a formarse en la división celular y se dispone en capas. La capa más alejada de la membrana plasmática es la que primero se forma y recibe el nombre de *lámina media*, que queda ubicada en la zona de la pared que está en contacto con la célula adyacente y mantiene a las células unidas. Suele tener un grosor de alrededor de 30 nm y está formada por hemicelulosa y pectina.

Las capas posteriores se depositan en el espacio comprendido entre la lámina media y la membrana plasmática. La pared celular primaria se forma luego de la lámina media en células que están en proceso de crecimiento y alcanza un grosor constante comprendido entre 0,1 y 1,0  $\mu\text{m}$  cuando el crecimiento se detiene. Está formada por celulosa, pectina y proteínas estructurales (Figura I1 a).

En algunos tipos celulares, particularmente en el esclerénquima, hay una capa adicional no deformable, la pared secundaria, normalmente más gruesa que la pared primaria y de ancho variable entorno a la membrana plasmática. Está compuesta de celulosa, hemicelulosa y lignina, y, en algunos casos, por pectina dispersa y en baja proporción (Figura I1 b) (Brett y Waldron, 1999; Zamil y Geitmann, 2017).

La hemicelulosa, la pectina y la lignina interactúan con las fibras de celulosa y aportan rigidez a la pared vegetal (Saini y col., 2015).



**Figura 11.** Estructura y composición de la pared vegetal a) primaria y b) secundaria. Fuente: Loix y col., 2017.

## 1.2. Componentes de la lignocelulosa

### 1.2.a. Celulosa

La celulosa es el material de origen biológico más abundante sobre la tierra y el componente mayoritario de la pared celular de plantas terrestres (Payne y col., 2015). Si bien se la encuentra casi exclusivamente en plantas, también es producida por tunicados (urocordados) y bacterias (Lynd y col., 2002). Está constituida por un polímero lineal anhidro de moléculas de glucopiranosas ligadas en uniones  $\beta$ -1,4. La unidad mínima de la celulosa es la celobiosa, compuesta por dos monómeros de glucosa. El grado de polimerización (DP), es decir, el número de moléculas de glucosa por cadena, varía según el origen de la celulosa en un rango de 100 a 10000 unidades. El DP es menor en la pared vegetal primaria que en la secundaria (Gorshkova y col., 2013; Kumar y Murthy, 2016; Xi y col., 2017).

Las cadenas de celulosa están estabilizadas por puentes de hidrógeno intra e intercatenarios (Lynd y col., 2002). Cada cadena tiene un extremo reductor, en que queda

expuesto el aldehído del carbono anomérico por la apertura del enlace hemiacetal, y otro no-reductor (Kumar y Murthy, 2016). Las cadenas de celulosa se disponen en forma paralela, adoptando una estructura compacta, llamada celulosa cristalina. Alrededor de 30 moléculas de celulosa se organizan formando protofibrillas, que se ensamblan en microfibrillas de entre 10 y 25 nm de diámetro. A su vez, las microfibrillas se empaquetan en macrofibrillas de aproximadamente 0,5  $\mu\text{m}$  de ancho y 4  $\mu\text{m}$  de longitud (Figura 12). Como resultado, la celulosa cristalina es un material insoluble en agua y, por lo tanto, recalcitrante a la acción enzimática (Kumar y Murthy, 2016; de Vries y Visser, 2001).

Algunas fibras de celulosa se apartan de la cristalinidad y se disponen de manera desordenada, más permeable a la entrada de agua y susceptible a la degradación enzimática (Lynd y col., 2002). Este tipo de celulosa es conocida como celulosa amorfa o paracristalina. El porcentaje de celulosa amorfa en la pared vegetal varía entre fuentes de celulosa. Una misma cadena de celulosa presenta regiones amorfas y cristalinas en forma alternada (Kumar y Murthy, 2016). Los pliegues y torsiones de microfibrillas también contribuyen a la hidratación de la celulosa y al acceso de enzimas (Lynd y col., 2002). La celulosa de la pared vegetal primaria tiene un menor grado de cristalinidad que la secundaria (Xi y col., 2017). Normalmente, la celulosa vegetal se encuentra embebida en una matriz de hemicelulosa y lignina en la gran mayoría de las plantas, salvo en las cápsulas de algodón y algunos otros casos, en los que se encuentra pura (Lynd y col., 2002).



**Figura 12.** Estructura esquemática de las microfibrillas de celulosa que componen la pared vegetal. Adaptado de la página <https://science.energy.gov/ber> del Departamento de Energía de Estados Unidos (DOE).

### 1.2.b. Hemicelulosa

La hemicelulosa es el segundo polisacárido más abundante en las plantas y, a diferencia de la celulosa, su composición varía ampliamente con el origen de la biomasa. Todos los tipos de hemicelulosas son heteropolímeros ramificados y complejos, compuestos por hexosas (D-galactosa, L-galactosa, D-manosa, L-fructosa, L-fucosa), pentosas (L-ramnosa, arabinosa, xilosa), ácido D-glucurónico y azúcares acetilados (Bosetto y col., 2016; Kucharska y col., 2018; Scheller y Ulskov, 2010).

En las maderas de angiospermas, llamadas maderas duras, el contenido de hemicelulosa es del 20% del peso seco de la planta. Las maderas de gimnospermas, llamadas blandas, están constituidas en un 5% por hemicelulosa (Singh y col., 2018).

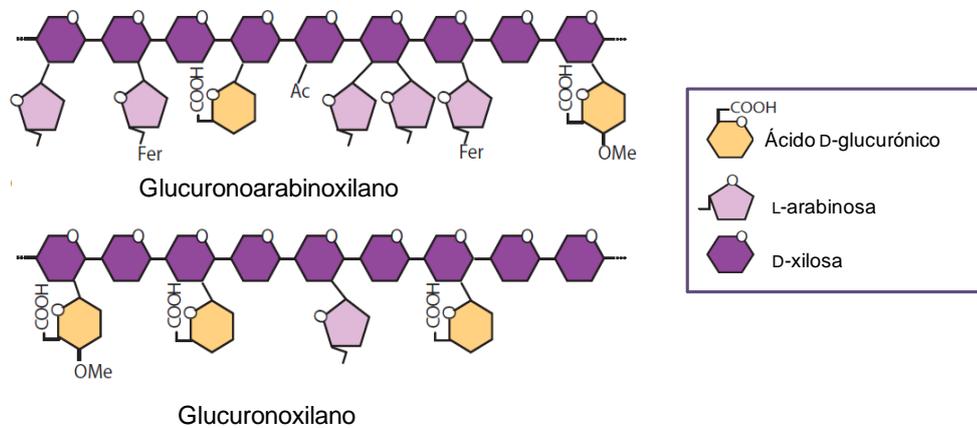
El xilano es un polímero formado por una cadena lineal de moléculas de xilopiranosas en uniones  $\beta$ -1,4, altamente ramificada, cuyos sustituyentes pueden ser acetilos, ácido glucurónico, ácido 4-O-metilglucurónico y arabinosa (Rennie y Scheller, 2014; Pauly y col., 2013). También pueden encontrarse residuos de L-ramnosa y L-fucosa (Gírio y col., 2010).

El xilano sustituido con ácido D-glucurónico, y su éter 4-O-metilico, en uniones  $\alpha$ -1-2 se conoce como glucuronoxilano (Figura I3). El glucuronoxilano es el tipo de xilano que predomina en la pared secundaria de maderas duras y en dicotiledóneas herbáceas, como *Arabidopsis thaliana* y especies del género *Populus* (Smith y col., 2017). El grado de sustitución es variable. En *Arabidopsis* sp., por ejemplo, se identificó un residuo de ácido glucurónico o metilglucurónico cada ocho moléculas de xilosa, con el doble de moléculas totales de ácido glucurónico que de metilglucurónico (Rennie y Scheller, 2014).

El xilano de monocotiledóneas commeliníneas, como las gramíneas y cereales, contiene, además, residuos de L-arabinofuranosa, por lo que recibe el nombre de glucuronoarabinoxilano (Figura I3), que constituye entre el 20-40% y el 40-50% de las paredes primaria y secundaria, respectivamente (Gírio y col., 2010; Pauly y col., 2013; Scheller y Ulskov, 2010). El xilano contenido en el endosperma de cereales tiene una abundante cantidad de L-arabinofuranosa y una baja frecuencia de residuos de ácido glucurónico y, con frecuencia, se lo llama arabinoxilano (de Vries y Visser, 2001; Scheller y Ulskov, 2010). Los restos de arabinofuranosa se unen a la xilosa en uniones  $\alpha$ -1,3 y/o  $\alpha$ -1,2 y pueden dar origen a cadenas formadas por más de un residuo de arabinofurano.

La mayoría de los xilanos están acetilados, particularmente en las dicotiledóneas, con grupos acetilo unidos al O-3 u O-2 de la xilosa. El grado de acetilación es variable. También

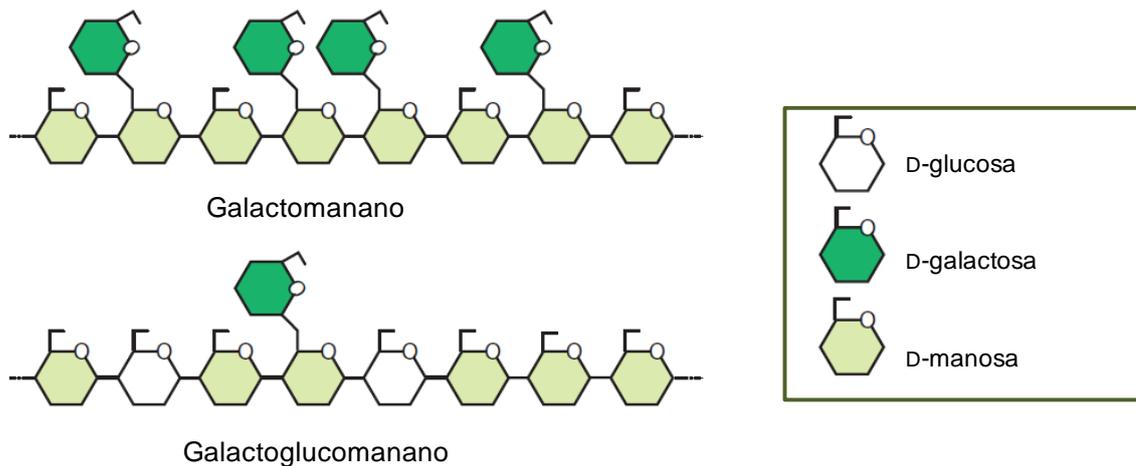
pueden encontrarse residuos aromáticos de ácido ferúlico y cumárico unidos al O-5 de arabinosas terminales (Pauly y col., 2013; Scheller y Ulvskov, 2010; Smith y col., 2017).



**Figura 13.** Estructura esquemática del glucuroarabinoxilano y el glucuronoxilano. Fer: ácido ferúlico. Ac: acetilo. Me: metilo. Fuente: Scheller y Ulvskov, 2010.

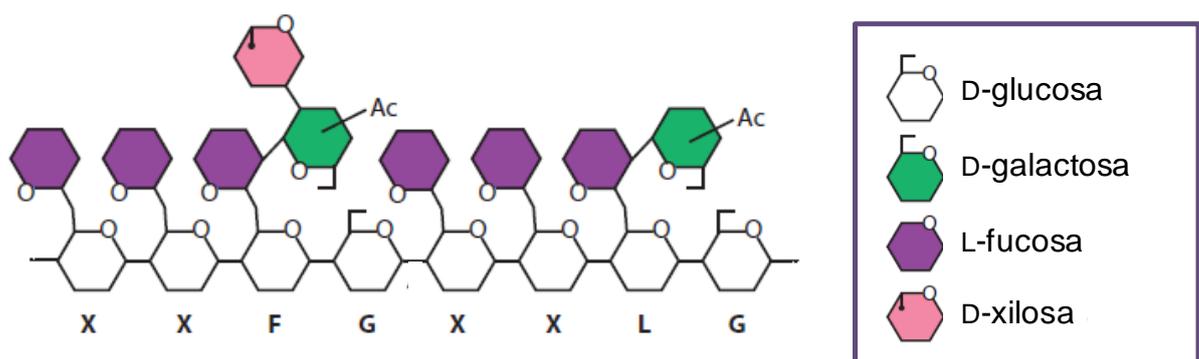
Además de los xilanos, las paredes vegetales están compuestas por otros tipos de hemicelulosa, principalmente mananos, xiloglucanos y glucanos mixtos.

Los mananos son polisacáridos que contienen manosa en uniones  $\beta$ -1,4 en su cadena principal, con sustituciones de D-galactosa en uniones  $\beta$ -1,6. En algunos casos puede haber, además, acetilaciones (Pauly y col, 2013). Cuando la manosa es el único monosacárido de la cadena principal, el polímero se denomina galactomanano (Figura 14). El galactoglucomanano, por el contrario, tiene moléculas de glucosa intercaladas con las de manosa en la cadena principal (Figura 14). Los galactomananos se encuentran típicamente en leguminosas y los glucomananos son componentes de la madera de coníferas (de Vries y Visser, 2001; Singh y col., 2018).



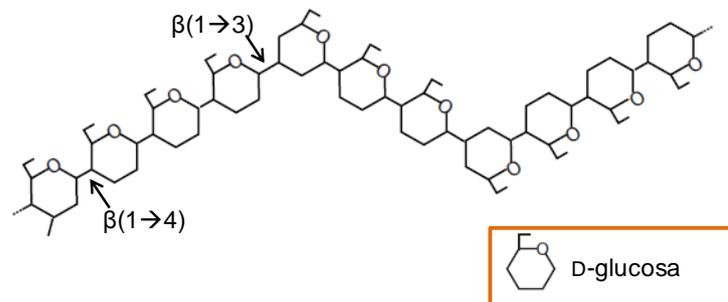
**Figura I4.** Estructura esquemática del galacto(gluco)manano. Fuente: Scheller y Ulskov 2010.

Los xiloglucanos son la hemicelulosa más abundante en la pared primaria de las dicotiledóneas, de la que constituyen hasta un 20-30% del peso seco, y parecen formar enlaces cruzados entre microfibrillas de celulosa (Pauly y col., 2013). Consisten en una cadena principal de glucosa en uniones  $\beta$ -1,4, sustituidas parcialmente por residuos de xilosa en uniones  $\alpha$ -1,6. En forma predominante, los xiloglucanos están compuestos por arreglos repetitivos de dos glucosas sustituidas seguidas de dos no sustituidas (XXGG), mayoritarios en monocotiledóneas commelínidas, y tres glucosas sustituidas consecutivas separadas por una glucosa no sustituida (XXXG). Los residuos de xilosa pueden presentar sustituciones de  $\alpha$ -L-arabinofuranosa,  $\beta$ -galactosa y  $\beta$ -galactosa ligada a  $\alpha$ -L-fucosa (Figura I5) (Scheller y Ulskov, 2010).



**Figura I5.** C Estructura esquemática de los xiloglucanos. Fuente: Scheller y Ulskov, 2010.

Los glucanos mixtos están conformados por unidades de glucosa en uniones  $\beta$ -1,3 y  $\beta$ -1,4 (Figura 16) y son característicos de las gramíneas (*Poacea*), aunque también han sido descritos en *Equisetum*, plantas hepáticas y algas (Pauly y col., 2013).



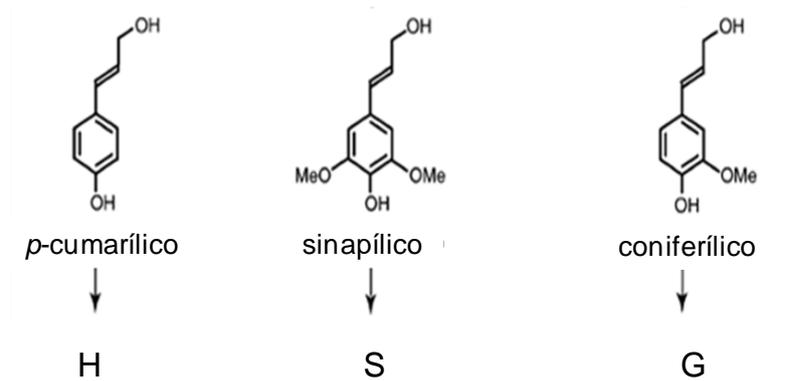
**Figura 16.** Estructura esquemática de los glucanos mixtos. El tipo de unión se indica con flechas. Fuente: Scheller y Ulaskov, 2010.

### 1.2.c. Lignina

La lignina es una de las mayores barreras para la degradación química y biológica de la biomasa. Se la encuentra principalmente en la lámina media y en la pared secundaria (Den y col., 2018; Martínez y col., 2005).

A diferencia de la celulosa y la hemicelulosa, es un polímero aromático de estructura reticulada, formado por subunidades de fenilpropanos di y monometoxilados (siringil y guaiacil propanos, S y G, respectivamente) y *p*-hidroxifenil propanos (H), derivados de los alcoholes sinapílico, coniferílico y *p*-cumarílico, respectivamente. Estas subunidades dan origen a una amplia diversidad de unidades que forman enlaces éter y C-C entre sí (Figura 17). La relación H:S:G varía según el origen de la biomasa y en las distintas capas de la pared vegetal (Martínez y col., 2005).

La lignina constituye el 16-25% y el 23-35% del peso seco de las maderas duras y blandas, respectivamente. También se encuentra en gramíneas en menor proporción (Brevu y Vasile, 2019; Den y col, 2018).



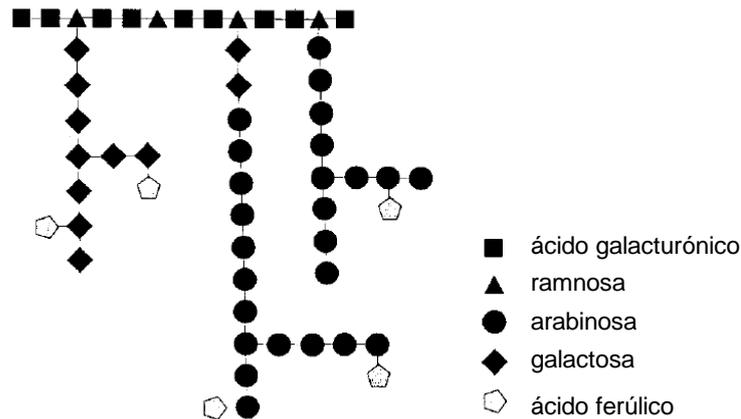
**Figura 17.** Alcoholes fenólicos que dan origen a las diferentes subunidades de la lignina. H: *p*-hidroxifenil propano, S: siringil propano, G: guaiacil propano. Adaptado de Vanholme y col., 2010.

### 1.2.d. Pectinas

La pectina constituye el 35% de la pared primaria en dicotiledóneas y en monocotiledóneas no commelínidas, el 2-10% de la pared primaria de monocotiledóneas commelínidas y hasta el 5% de la pared en maderas. Es abundante en la lámina media, inclusive, entre células con pared secundaria, como las del xilema (Mohnen, 2008).

Los polisacáridos pécticos, como el homogalacturano y el ramnogalacturano, son los más complejos de la pared vegetal (de Vries y Visser, 2001). El homogalacturano está formado por una cadena principal de ácido D-galacturónico en uniones  $\alpha$ -1,4, que puede estar acetilado en las posiciones O-2 u O-3 o metilado en O-6. En otros tipos de pectinas, la cadena principal está unida a residuos de xilosa o apiofuranosa (Harholt y col., 2010).

En el ramnogalacturano, las cadenas de ácido galacturónico están interrumpidas por residuos de ramnosa en uniones  $\alpha$ -1,2, que pueden estar ligados a cadenas laterales de arabinosa o galactosa, también ramificadas y sustituidas con residuos de ácido ferúlico, arabinosa y azúcares poco comunes, como L-fucosa (Figura 18) (de Vries y Visser, 2001).



**Figura 18.** Estructura esquemática del ramnogalacturano. Fuente: de Vries y Visser, 2001.

### 1.3. Bioetanol y fuentes de biomasa lignocelulósica

El etanol puede ser producido a partir de cualquier materia prima que contenga grandes cantidades de azúcar, como la remolacha azucarera y la caña de azúcar, o de materiales potencialmente sacarificables, como el almidón contenido en los granos de cereales (trigo, maíz, cebada). El azúcar de caña es usada principalmente por Brasil (y también por Europa), para la generación de bioetanol, mientras que Estados Unidos y países europeos de la Agencia Internacional de la Energía (IEA) utilizan mayormente granos (Fulton y col., 2004). Brasil y Estados Unidos son los mayores productores de etanol como biocombustible, ya que constituyen el 38,8% y 38,3% de la producción global, respectivamente (dato de 2007) (Kumar Saini y col., 2015).

El etanol derivado de azúcar o almidón se denomina *de primera generación* (1G). Su obtención implica el uso de cultivos alimenticios y, además, el desperdicio de otras partes de la planta, que son generalmente desechadas o utilizadas en procesos de poco valor agregado, como la quema en calderas para generación de energía eléctrica.

En consecuencia, la biomasa lignocelulósica, proveniente de partes no alimenticias de las plantas, está siendo estudiada actualmente para la producción de bioetanol de origen (hemi)celulósico, o de *segunda generación* (2G), ya que la celulosa y algunas hemicelulosas son materiales sacarificables. Las principales fuentes son los desechos agrícolas (paja de trigo, paja de cebada, cascarilla de arroz, bagazo de caña de azúcar), los residuos forestales (provenientes del raleo de especies forestales cultivadas), los residuos urbanos (papel y cartón de descarte) y cultivos destinados a la producción de energía (árboles de crecimiento

rápido, como álamos y sauces, y pastos energéticos) (Fulton y col., 2004). La composición porcentual de celulosa, hemicelulosa y lignina de algunas de estas fuentes de biomasa se describen en la Tabla I1.

**Tabla I1.** Composición porcentual de hemicelulosa, celulosa y lignina de biomasa lignocelulósica de distinto origen (expresada como % p/p del peso seco). Fuentes: Kim, 2018; Kucharska y col., 2018.

<b>Biomasa</b>	<b>Celulosa</b>	<b>Hemicelulosa</b>	<b>Lignina</b>
Bagazo de caña de azúcar	43,1	31,1	11,4
Bagazo de sorgo dulce	27,3	13,1	14,3
Paja de cebada	33,8	21,9	13,8
Paja de trigo	30,2	21,0	17,0
Madera de álamo	43,8	14,8	21,1
Madera de pino	40,0	28,4	27,7
Pasto varilla ( <i>switchgrass</i> )	39,5	20,3	17,8
Mazorca de maíz	35,0	16,8	7,0
Rastrojo de maíz	37,0	22,7	18,6
Residuos de algodón	58,5	14,4	21,5
Poda de olivo	25,0	11,1	16,2

## **2. Deconstrucción enzimática de la biomasa lignocelulósica**

### **2.1. Organismos productores de enzimas activas sobre lignocelulosa**

La biomasa lignocelulósica tiene la característica de ser recalcitrante frente a la acción enzimática. Esto se debe, principalmente, a la cristalinidad de la celulosa, la complejidad estructural de la hemicelulosa y la presencia de lignina (Alfaro y col., 2014; Cragg y col., 2015). Sin embargo, constituye la fuente principal de carbono orgánico renovable sobre la Tierra (que resulta de la fijación del carbono atmosférico a través de proceso de fotosíntesis), de manera que organismos de todos los reinos han evolucionado para aprovechar los polisacáridos que la componen como fuente de carbono y energía (Alfaro y col., 2014; Bomble y col., 2017; Cragg y col., 2015).

Los principales degradadores de biomasa lignocelulósica son hongos y bacterias saprófitos, habitantes de suelos, que son capaces de secretar enzimas activas sobre lignocelulosa que actúan en forma sinérgica y cumplen un rol fundamental en el ciclo del carbono (Bon y Ferrara, 2010; Bomble y col., 2017; Koeck y col., 2014; Wyman, 1994).

La mayoría de estas enzimas son las glicosidasas o glicosil hidrolasas que catalizan la hidrólisis de enlaces glucosídicos  $\alpha$ -1,3/1,4 y  $\beta$ -1,3/1,4. También, intervienen las carbohidrato esterasas, fosforilasas, y enzimas con actividades auxiliares (Bomble y col., 2017). Las enzimas se clasifican de acuerdo a su actividad experimentalmente comprobada mediante la asignación de un *Enzyme Commission Number* (EC) determinado por la IUBMB. Las glicosidasas se agrupan bajo el EC 3.2.1.-, siendo el último número específico para cada actividad.

Adicionalmente, las enzimas activas sobre carbohidratos, llamadas CAZimas (del inglés *Carbohydrate Active enZYmes*), fueron clasificadas en familias según la estructura de los dominios que la componen y agrupadas en la base de datos CAZy (Lombard y col., 2014), según sean glicosil hidrolasas (GH), carbohidrato esterasas (CE), glicosil transferasas (GT) polisacárido liasas (PL) o enzimas con actividades auxiliares (AA) (Levasseur y col., 2013). Las CAZimas tienen en general estructura modular y pueden estar formadas por un dominio catalítico, que puede estar ligado por secuencias enlazadoras (*linker*) a dominios de unión a sustrato (*Carbohydrate Binding Modules*, CBMs) o en algunos casos a dominios de anclaje (*scaffolding*) (DS) (Gómez del Pulgar y Saadeddin, 2014; Pàges y col., 1997). A su vez, se han descrito enzimas con más de un dominio catalítico, tanto de igual como de distinta familia.

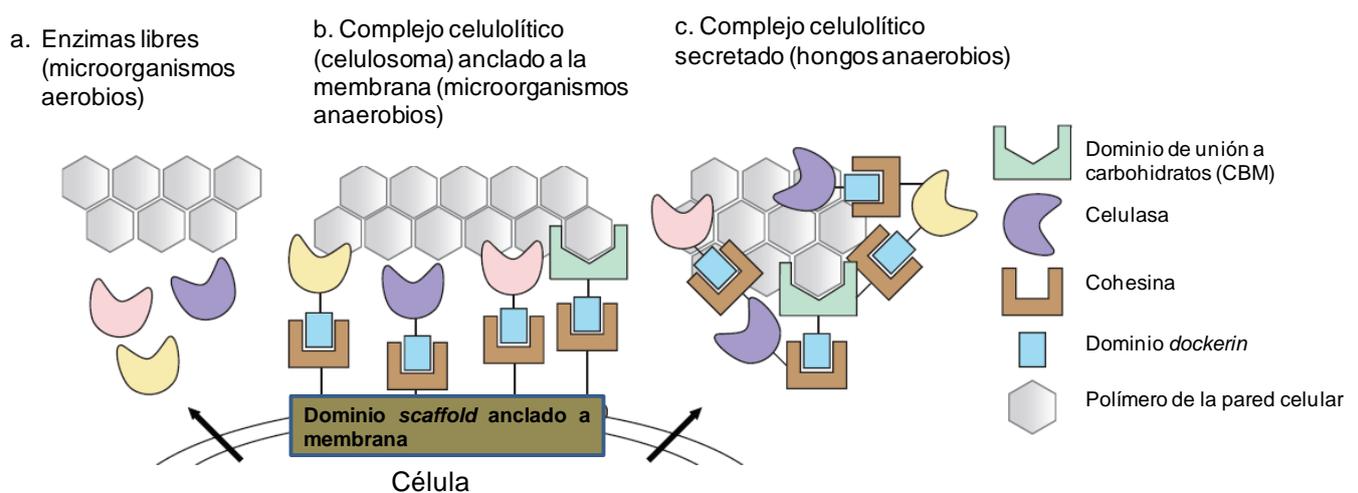
Además de la hidrólisis enzimática de los polisacáridos, la deconstrucción puede ser llevada a cabo por generación de especies oxidativas que actúan sobre la biomasa. El mecanismo oxidativo puede ser enzimático o químico (Cragg y col., 2015; Levasseur et. al, 2013). Es el caso de las monooxigenasas de polisacáridos dependientes de cobre (LPMOs), que corresponden a las familias de CAZy AA9 y AA10 (descritas en hongos y bacterias, respectivamente), y de lacasas y otras peroxidadas, producidas por algunos hongos y bacterias (Cragg y col., 2015; Levasseur y col., 2013).

La mayoría de los hongos capaces de degradar madera en descomposición son hongos que pertenecen a los filos *Ascomycota* y *Basidiomycota*. Se los clasifica en hongos de la pudrición blanca y roja según su mecanismo de acción (Lundell y col., 2010). Los hongos de la pudrición blanca degradan todos los componentes de la lignocelulosa de maderas duras y blandas, ya que secretan enzimas activas sobre (hemi)celulosa, así como lacasas y peroxidadas que deconstruyen la lignina y permiten el acceso de las (hemi)celulasas a los polisacáridos (Alfaro y col., 2014). En contraposición, los basidiomicetes son llamados hongos de la pudrición roja, dado que modifican la lignina en forma limitada, por vías

oxidativas no enzimáticas, y producen principalmente enzimas (hemi)celulolíticas (Arantes y col., 2012).

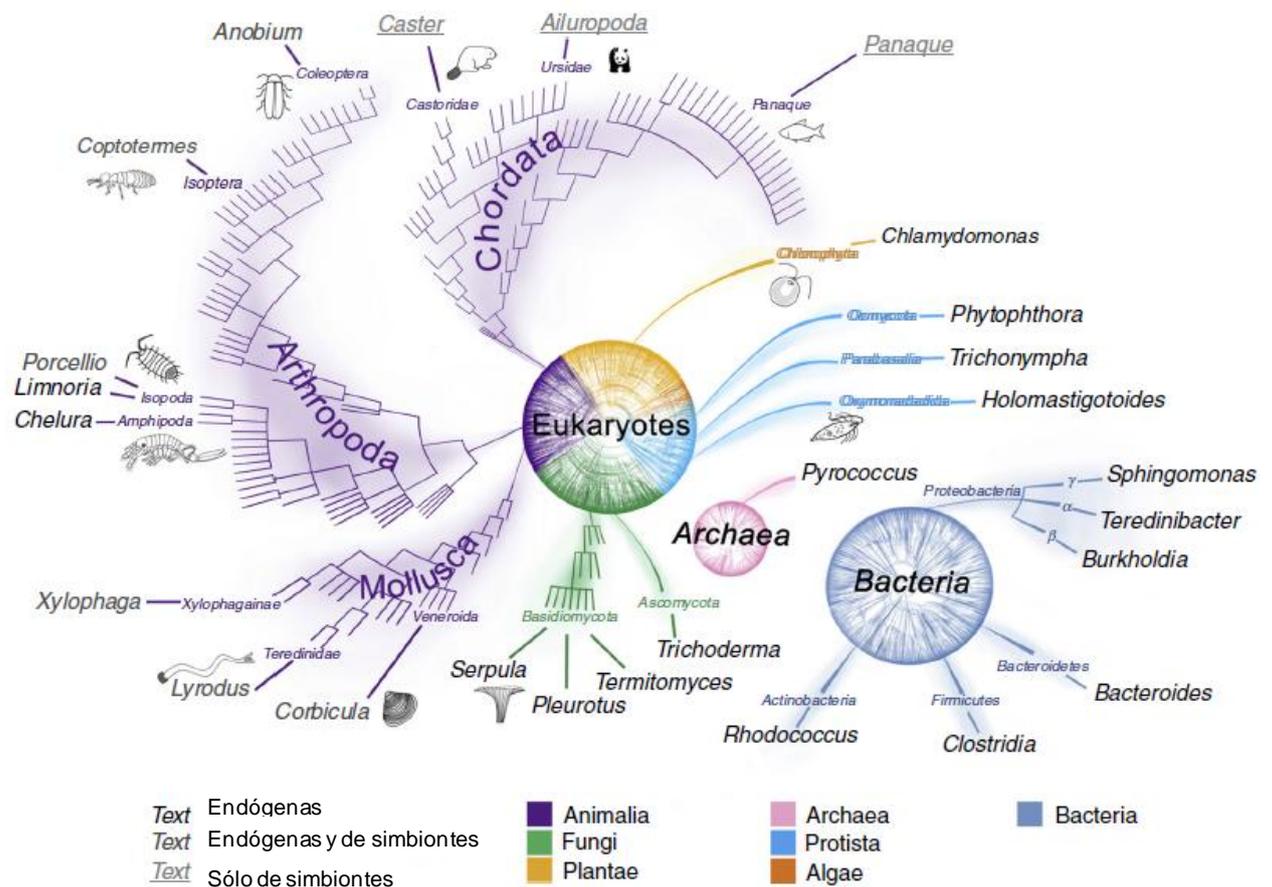
Las bacterias saprófitas pertenecen a los filos *Actinobacteria* (géneros como *Cellulomonas*, *Micromonospora*, *Streptomyces* y *Thermobifida*), *Proteobacteria* (*Gammaproteobacteria*, como las del género *Cellvibrio*) y *Bacteroidetes* (principalmente, del género *Cytophaga*) (Koeck y col., 2014). Son principalmente degradadoras de (hemi)celulosa, pero algunas actinobacterias producen lacasas y otras peroxidasas que actúan sobre lignina (*Streptomyces* sp., *Thermobifida fusca*) (Saini y col., 2015). Generalmente, se encuentran formando consorcios cuyos miembros producen enzimas con actividades complementarias.

Las estrategias bacterianas son variables en cuanto a la arquitectura de las enzimas, ya que las mismas pueden ser secretadas como unidades catalíticas independientes, o formando estructuras multi-modulares asociadas a proteínas *scaffold* (mediante dominios de anclaje, o *dockerin*, y cohesinas) llamadas “celulosomas”, o como proteínas multifuncionales que pueden hallarse libres o en celulosomas (Figura 19) (Cragg y col., 2015; Koeck y col., 2014.; Saini y col., 2015). Los celulosomas fueron descritos por primera vez en *Clostridium thermocellum* (Lamed y col., 1983) y son característicos de las bacterias anaerobias, mientras que las bacterias aerobias son secretoras de enzimas libres (Saini y col., 2015). Los hongos anaerobios pueden secretar complejos multienzimáticos, semejantes a los celulosomas (Haitjema y col., 2014).



**Figura 19.** Estrategias de hongos y bacterias para la degradación enzimática de la pared vegetal: a) secreción de enzimas libres, b) celulosomas anclados a membrana, c) complejos celulolíticos secretados. Fuente: Haitjema y col., 2014.

Si bien hongos y bacterias detritívoros producen enzimas endógenas, otros organismos se valen de las enzimas exógenas provenientes de microorganismos simbiotes, ya sea en forma exclusiva, como es el caso de los rumiantes (bacterias de los géneros *Firmicutes* y *Bacteroidetes*, principalmente) (Koeck y col., 2014; White y col., 2014) y bivalvos de la familia *Terenidae* ( $\gamma$ -*Proteobacteria*) (Cragg y col., 2015), o a fin de suplementar el repertorio endógeno, como en las termitas y otros insectos que se alimentan de maderas (Figura I10). Es interesante mencionar, además, que ha habido hallazgos novedosos de (hemi)celulasas en arqueas extremófilas (por ejemplo, del género *Pyrococcus*), en bacterias marinas (grupos *Cytophaga* y *Flavobacteria*) y en protistas de vida libre (género *Chlamydomonas*) o simbiotes de termitas (Figura I10) (Cragg et. al, 2015).



**Figura I10.** Organismos que degradan lignocelulosa por acción enzimas endógenas y/o exógenas, producidas por simbiotes. Fuente: Cragg y col., 2015.

## 2.2. Enzimas activas sobre lignocelulosa

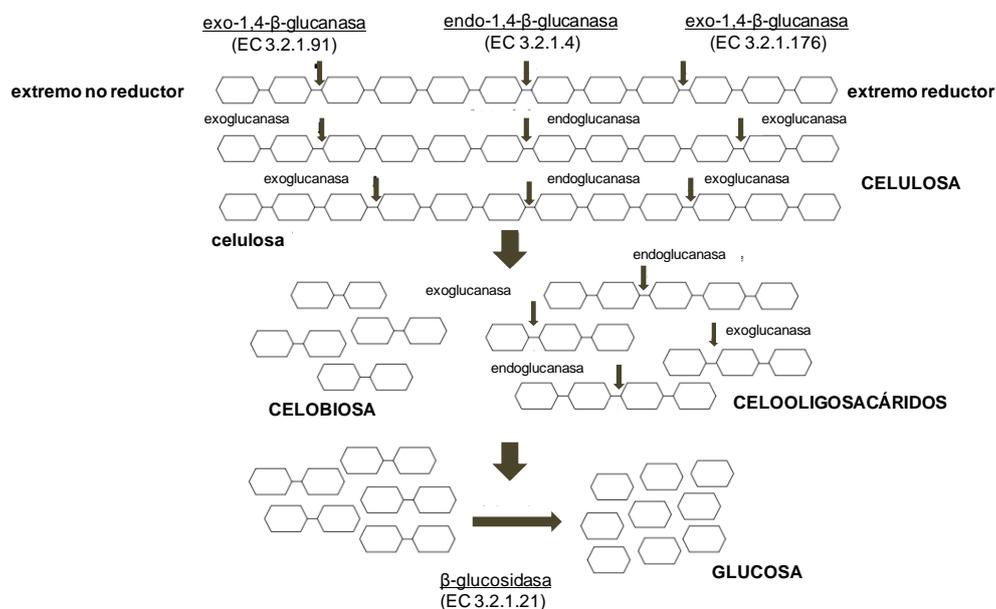
### 2.2.a. Celulasas

Las enzimas involucradas en la degradación de la celulosa se clasifican en endo- $\beta$ -1,4-glucanasas, exo-1,4- $\beta$ -glucanasas (o celobiohidrolasas) y  $\beta$ -xilosidasas.

Las endo-1,4- $\beta$ -glucanasas (EC 3.2.1.4) actúan principalmente sobre celulosa amorfa, efectuando la hidrólisis al azar de enlaces  $\beta$ -1,4 y generando, por lo tanto, extremos reductores y no reductores (Lynd y col., 2002).

Las exoglucanasas, también llamadas celobiohidrolasas, actúan liberando celobiosa, o glucosa, de los extremos reductores (EC 3.2.1.176) y no reductores (EC 3.2.1.91), generados por las endoglucanasas. Estas enzimas son fundamentales para la degradación de la celulosa cristalina y pueden tener actividad procesiva (Lynd y col., 2002; Saini y col., 2015).

La liberación de unidades de glucosa depende de la acción de las  $\beta$ -glucosidasas (EC 3.2.1.21), cuyo sustrato es la celobiosa, obtenida de la actividad concertada de endo y exoglucanasas. Es decir que la obtención de glucosa como producto de la celulosa es consecuencia del sinergismo entre endoglucanasas y exoglucanasas (sinergismo endo-exo), exoglucanasas que actúan sobre extremos reductores y no reductores (sinergismo exo-exo) y exoglucanasas y  $\beta$ -glucosidasas (Lynd y col., 2002; Saini y col., 2015) (Figura I11).



**Figura I11.** Degradación de celulosa por acción de glicosil hidrolasas. Fuente: Saini y col., 2015.

La mayoría de las glicosil hidrolasas de las familias GH5, GH6, GH7, GH8, GH9, GH12, GH44, GH15 y GH48 actúan sobre celulosa (Talamantes y col., 2016). Las exoglucanasas en general pertenecen a las familias GH6, GH7 (exclusivamente en hongos) y GH48 (en bacterias). La exoglucanasa GH7 más estudiada es Cel7a de *T. reesei*, que es la proteína mayoritaria del secretoma en condiciones de producción de celulosas (Quinlan y col., 2010). Las  $\beta$ -glucosidasas se distribuyen principalmente en las familias GH1 y GH3, pero también en GH2, GH5, GH9, GH30, GH39 y GH116 (Salgado y col., 2018). Las endoglucanasas pertenecen a las familias GH5, GH6, GH7, GH8, GH9, GH12, GH44 y GH48. También, a GH45, GH51 y GH74 (Vlasenko y col., 2010).

Los sustratos empleados en ensayos de detección de actividad sobre celulosa cristalina son, tradicionalmente, papel de filtro Whatman N°1 y Avicel (Sigma-Aldrich). La actividad endoglucanasa se evalúa usando celulosa soluble, como carboximetilcelulosa (CMC) y la actividad celulasa general, celulosa amorfa regenerada (PASC) y se mide el aumento de azúcares reductores por el método de reducción del ácido 3,5-dinitrosalicílico (DNS) (Ghose, 1987; Zhang y col., 2006).

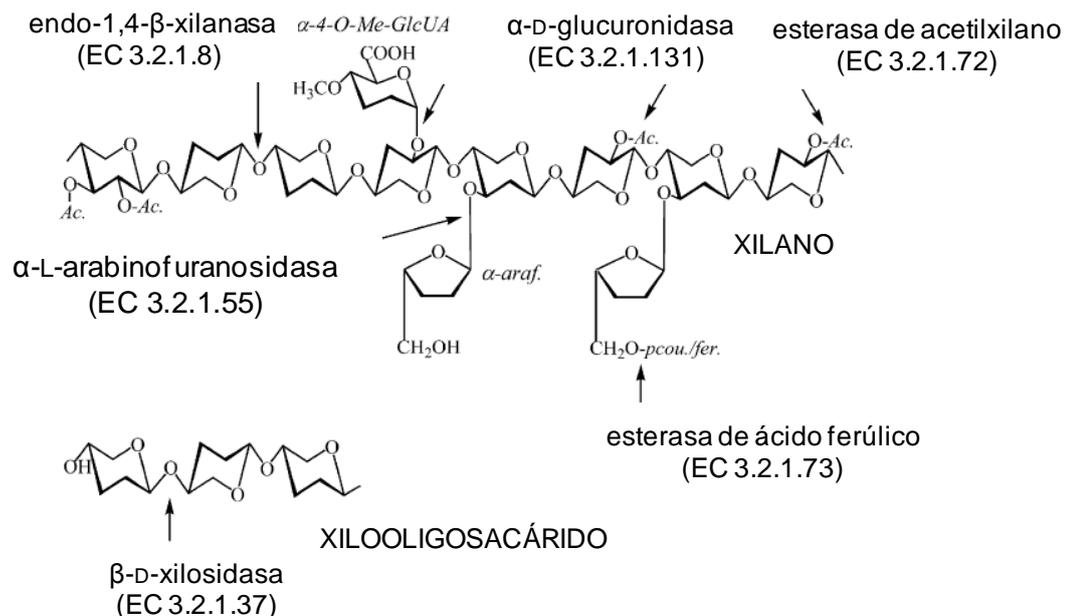
### **2.2.b. Hemicelulasas**

Las hemicelulosas son menos resistentes a la degradación física, química y/o biológica que la celulosa (Kucharska y col., 2018). La hidrólisis de las cadenas de xilano es llevada a cabo por endo- $\beta$ -1,4-xilanasas (EC 3.2.1.8), que actúan liberando xilooligosacáridos a partir de las cadenas de xilano, y por las  $\beta$ -D-xilosidasas (EC 3.2.1.37), que liberan xilosa a partir de xilobiosa. De manera similar a lo que se describe para la obtención de glucosa a partir de celulosa, dado el tamaño del sustrato, la mayoría de las xilanasas son secretadas al medio externo. Los xilooligómeros que producen son luego internalizados y degradados por  $\beta$ -xilosidasas y xilanasas intracelulares (Collins y col., 2005).

Se ha descrito actividad endoxilanasas en enzimas con módulos GH10 y GH11, principalmente, pero también GH5, GH7, GH8 y GH43. Las  $\beta$ -xilosidasas se distribuyen en las familias GH1, GH3, GH30, GH39, GH43, GH51, GH52, GH54, GH116 y GH120 (Bosetto y col., 2016; Collins y col., 2005).

Debido a que las cadenas de xilano se encuentran muy ramificadas con sustituyentes de distinta naturaleza, se requiere de la acción auxiliar de enzimas desramificantes, como las  $\alpha$ -L-arabinofuranosidasas (EC 3.2.1.55), las  $\alpha$ -D-glucuronidasas (EC 3.2.1.131), las esterasas de

ácido férulico (EC 3.1.1.73) y las esterasas de acetilxilano (EC 3.2.1.72) (Figura I12). Estas enzimas favorecen el acceso de las endoxilanasas a la cadena principal de xilano. Otros tipos de hemicelulosa involucran la actividad de enzimas como mananasas (EC 3.2.1.78) y  $\beta$ -manosidasas (EC 3.2.1.25), de  $\alpha$ -galactosidasas (EC 3.2.1.22) y/o de xiloglucanasas (EC 3.2.1.151) (Figura I12).



**Figura I12.** Esquema de la hidrólisis enzimática de xilano. Me-GlcUA: ácido metilglucurónico; araf: arabinofuranosa; *p*-cou: ácido *p*-cumarílico; fer: ácido ferúlico. Fuente: Collins y col., 2005.

La actividad hemicelulolítica general se detecta sobre la porción hemicelulósica extraída de distintas fuentes, como el glucuronoxilano de madera de abedul, o de haya, y el arabinoxilano de trigo. Otras actividades específicas se determinan con sustratos sintéticos, como el *p*-nitrofenilxilobiósido (pNPX) en el caso de la actividad  $\beta$ -xilosidasa (Ghose y Bisaria, 1987).

### 2.2.c. Pectinasas

Las enzimas que hidrolizan las cadenas de homogalacturano son las pectato y pectinaliasas (EC 4.2.2.10 y 4.2.2.11, respectivamente) y poligalacturonasas (EC 3.4.1.15 y 3.2.1.67). Múltiples enzimas actúan sobre el ramnogalacturano, como las ramnogalacturano liasas y endo/exohidrolasas (Martani y col., 2017; de Vries y Visser, 2001).

### **2.2.d. LPMOs**

Las LPMOs resultan esenciales en la degradación de polisacáridos de estructura cristalina, como la celulosa y la quitina (Morgestern y col., 2014). Actúan en forma sinérgica con las glicosil hidrolasas clásicas y podrían ser responsables del ataque inicial de los polisacáridos (Hemsworth y col., 2013). Como su mecanismo fue dilucidado recientemente, algunas LPMOs habían sido clasificadas dentro de las familias GH61 o CBM33. En el año 2013, se las re-clasificó en AA9 y AA10, respectivamente (Levasseur y col., 2013). Las AA9 son características de hongos y se ha identificado que actúan sobre celulosa, principalmente. Las AA10 son bacterianas y se ha descrito su actividad sobre celulosa y quitina (Johansen, 2016). Hoy en día, se considera que el modelo para la hidrólisis enzimática de celulosa debe incluir las glucanasas procesivas y no procesivas y las LPMOs (Morgersten y col., 2014).

### **2.2.e. Ligninasas**

Las lacasas, oxidasas que coordinan átomos de cobre (EC 1.10.32), y las hemoperoxidasas son las principales enzimas que modifican la lignina. Las hemoperoxidasas se clasifican en manganeso y lignina peroxidasas y peroxidasas versátiles (EC 1.11.1.13/14/16, respectivamente). Otra clase de ligninasas son las peroxidasas decolorantes (EC 1.11.1.19) que han sido identificadas recientemente (Martínez y col., 2005; Saini y col., 2015).

## **3. Aplicaciones industriales de las enzimas activas sobre lignocelulosa**

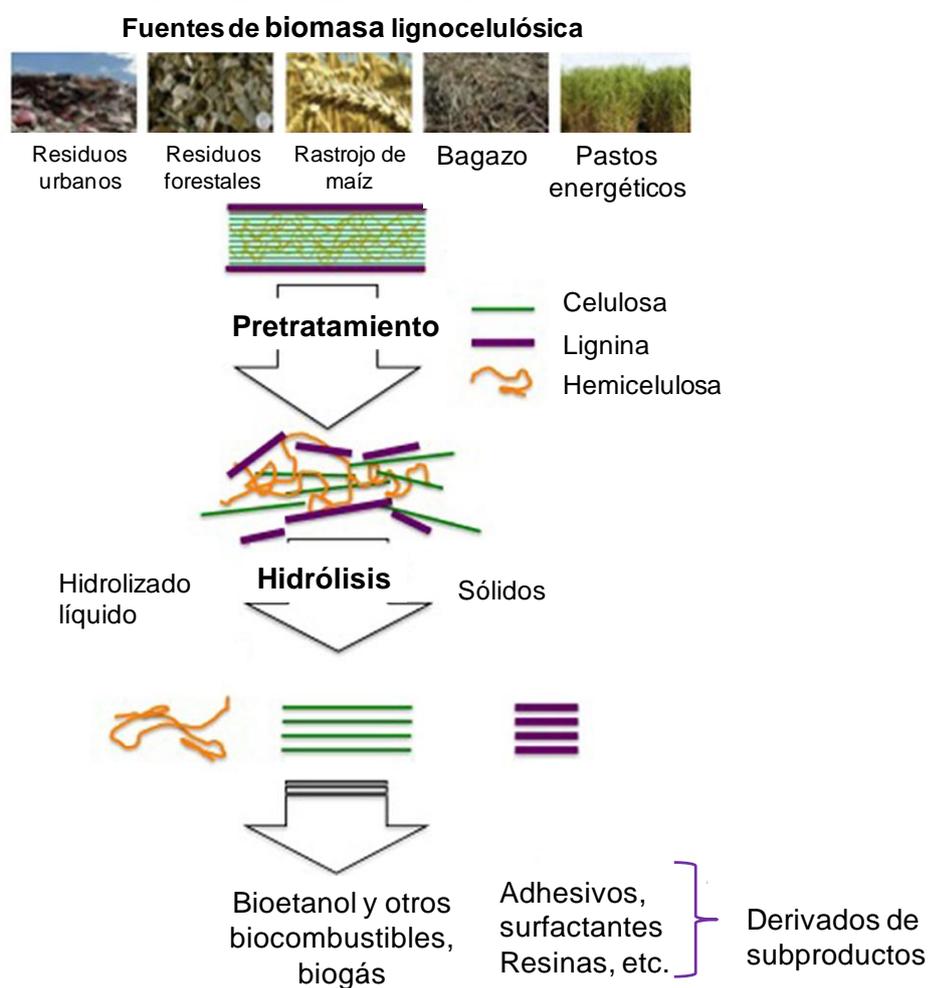
### **3.1. Bioetanol de segunda generación**

Los biocombustibles celulósicos, o de segunda generación, se basan en la sacarificación de la celulosa para la obtención de glucosa y su posterior fermentación a etanol, llevada a cabo por levaduras. La sacarificación puede efectuarse por tratamiento con ácido a altas temperaturas o por hidrólisis enzimática, siendo ésta última la metodología que ha dado mejores rendimientos (Manzanares, 2010).

Dada la resistencia de la biomasa a la degradación, es necesario realizar un pre-tratamiento disruptivo previo a la hidrólisis enzimática, a fin de separar la celulosa de la matriz de lignina y hemicelulosa. El pretratamiento puede ser físico (fragmentación

mecánica, radiación de microondas, ultrasonido, pirolisis), químico (hidrólisis ácida, pretratamiento alcalino, oxidación y ozonización, tratamiento con líquidos iónicos o solventes), físico-químico (explosión por vapor, dióxido de carbono o amonio) o biológico (descomposición por acción fúngica, bacteriana o enzimática) (Kurcharska y col., 2018).

Luego del pretratamiento y la sacarificación, los azúcares son fermentados por levaduras a etanol, que se recupera por destilación (Manzanares, 2010). A lo largo de todo el procedimiento, se generan subproductos que pueden ser aprovechados como adhesivos, resinas, materiales de construcción sintéticos, entre otras cosas (Figura I13) (Den y col., 2017).



**Figura I13.** Esquema general de la obtención de bioetanol y biocombustibles. Adaptado de: Den y col., 2017.

Nuevas estrategias consisten en integrar las etapas, lo que permite reducir el número de biorreactores y los costos de producción. Se puede realizar la *sacarificación* y

*fermentación en forma simultánea* (SSF) e, incluso, incorporar la fermentación de pentosas (derivadas de la hemicelulosa) al procedimiento, es decir, hacer la *sacarificación y co-fermentación de glucosa y xilosa en forma simultánea* (SSCF). En última instancia, lo que se busca es un *bio-proceso consolidado* (CBP), en el que el mismo microorganismo, o conjunto de microorganismos, sea capaz de sacarificar el sustrato y fermentar los azúcares (Manzanares, 2010).

Los extractos enzimáticos de celulasas que se utilizan actualmente a nivel prototípico en la industria provienen de los hongos filamentosos *Aspergillus niger* y *Trichoderma reesei*, que secretan una variedad de enzimas celulolíticas (Bhattacharya y col., 2015; Quinlan y col., 2010). En los extractos de *T. reesei*, las proteínas mayoritarias son dos celobiohidrolasas y dos endoglucanasas. Se han identificado también otras endoglucanasas (3),  $\beta$ -glucosidasas (3), xilanasas (4), arabinofuranosidasas (3), una  $\beta$ -xilosidasa, una mananasa, una  $\alpha$ -galactosidasa, una esterasa de acetilxilano y una feruloil esterasa. Además, los extractos contienen dos proteínas con dominios de unión a celulosa, una proteasa tipo tripsina y una proteína similar a las expansinas de las plantas (Quinlan y col., 2010).

Las mezclas enzimáticas producidas naturalmente por microorganismos tienen, sin embargo, un rendimiento menor que el de mezclas diseñadas específicamente. La mayoría de las mezclas naturales deben ser complementadas con enzimas adicionales. Por ejemplo, los extractos de *T. reesei* son suplementados con  $\beta$ -glucosidasas a fin de optimizar la relación entre endoglucanasas, celobiohidrolasas y  $\beta$ -glucosidasas para la degradación de cada tipo de sustrato. Otras enzimas que se agregan a las mezclas naturales son hemicelulasas (como xilanasas y  $\beta$ -xilosidasas) (Kumar y Murthy, 2017).

La sacarificación enzimática es el cuello de botella de la producción de etanol celulósico, dado el alto costo económico asociado a la producción de enzimas y la variedad de las actividades enzimáticas necesarias para la degradación completa de un sustrato complejo (Manzanares, 2010). Por lo tanto, es crucial el desarrollo de nuevos cócteles, con todas las enzimas necesarias en la proporción justa, apropiada para cada tipo de biomasa.

Además, la producción de glucosa disminuye en función del tiempo de hidrólisis, lo que se postula que puede resultar de la inactivación térmica de las enzimas, la unión irreversible al sustrato, la inhibición de las celulasas por producto y los cambios reológicos de la materia prima a lo largo del proceso (Manzanares, 2010). También se ha adjudicado la modificación del rendimiento de las enzimas a la formación de compuestos inhibitorios (como furanos,

ácidos orgánicos, fenoles) durante el pretratamiento, que interfieren en las etapas siguientes de producción de etanol celulósico, es decir, en la sacarificación y fermentación (Kim, 2018).

Por todo esto, es necesaria la prospección de enzimas novedosas, adecuadas a las condiciones del proceso (termoestables, que no presenten regulación por inhibición por producto, que se produzcan en grandes cantidades), a fin de maximizar el rendimiento y reducir los costos económicos asociados a la producción de etanol celulósico (Quinlan, Teter y Xu, 2010). Otra estrategia consiste en modificar genéticamente las enzimas ya caracterizadas para mejorar su actividad (Kim, 2018).

La sustentabilidad económica del etanol celulósico podría estar ligada, además, al diseño de biorrefinerías para la explotación conjunta tanto del biocombustible como de los subproductos, que pueden ser convertidos a productos de alto valor agregado (Arévalo-Gallegos y col., 2017).

### **3.2. Otras industrias**

Además de la industria de los biocombustibles, las enzimas activas sobre lignocelulosa tienen múltiples aplicaciones, algunas de las cuales se describen a continuación.

#### **3.2.a. Fabricación de papel**

Las lacasas y las xilanasas se pueden usar en la fabricación de papel, para hacer un blanqueamiento enzimático en reemplazo de los métodos químicos tradicionales, que involucran compuestos clorados ( $\text{ClO}_2$ ,  $\text{Cl}$ , hipoclorito) y generan derivados orgánicos de cloro, clasificados como mutagénicos o carcinogénicos potenciales. Para esta aplicación, es muy importante que los extractos enzimáticos que contengan lacasas y xilanasas y estén libres de enzimas que degraden celulosa (Kumar y col., 2016).

#### **3.2.b. Industria alimenticia humana y alimentación animal**

Las hemicelulasas pueden usarse en la producción de oligosacáridos, que han sido utilizados tradicionalmente como edulcorantes, pero que tendrían también propiedades beneficiosas para la salud. Particularmente, los *oligosacáridos que no son digeribles* por las enzimas gastrointestinales (NDOs), sirven de prebióticos para la microbiota del intestino y para probióticos, bifidobacterias y lactobacilos.

Los NDOs incluyen a los xilooligosacáridos (XOS), fructooligosacáridos (FOS), galactooligosacáridos (GOS) y manooligosacáridos (MOS), por lo tanto, para su obtención a partir de biomasa, se necesitan enzimas específicas, como xilanasas,  $\beta$ -fructofuranosidasas,  $\beta$ -galactosidasas,  $\beta$ -mananasas,  $\beta$ -glucosidasas (activas sobre galactoglucomananos) y esterases de con distinta especificidad (Rodrigues Mano y col., 2018).

Adicionalmente, las enzimas activas sobre lignocelulosa son empleadas para mejorar las características de productos de consumo humano que derivan de frutas y vegetales. Las (hemi)celulasas y pectinasas se aplican, por ejemplo, para la sacarificación de jugos y la producción de puré. También, se usan mezclas comerciales de pectinasas con bajo contenido de (hemi)celulasas para la extracción de aceites de cocina, como los aceites de oliva, canola y girasol, entre otros. En la industria vitivinícola se utilizan  $\beta$ -glucanasas, hemicelulasas y pectinasas, que mejoran ciertas características del vino, como su color y aroma (Toushik y col., 2017). Las xilanasas se usan también para mejorar las características de productos panadería como, por ejemplo, la esponjosidad del pan (Harris y Ramalingam, 2010).

Recientemente, se ha estudiado el uso de celulasas y amilasas en cócteles multi-enzimáticos diseñados para la disgregación de la matriz extracelular en biofilms de bacterias patógenas que colonizan alimentos o elementos de cocina (*Staphylococcus aureus*, *Escherichia coli*, *Salmonella enterica*, *Bacillus cereus*, *Lysteria monocytogenes*) (Galié y col., 2018).

Las enzimas fibrolíticas (celulasas,  $\beta$ -glucanasas y xilanasas, principalmente) son también de utilidad en el procesamiento de alimentos para animales monogástricos (como aves y porcinos) (Bhat, 2000; Mendoza y col., 2014), ya que colaboran en la hidrólisis de las fibras presentes en granos y vegetales, es decir, de todos los componentes que no son almidón, como celulosa, pectina y xilanos. Las fibras son consideradas factores antinutricionales. La adición de estas enzimas al alimento puede convertir los polisacáridos estructurales en un material absorbible por los animales monogástricos, lo que mejora la salud animal (Kuhad y col., 2011). Inclusive, se han observado mejoras en el crecimiento, la masa muscular y la calidad de la leche por agregado de enzimas fibrolíticas al forraje destinado a rumiantes (Mendoza y col., 2014; Sharma y col., 2016).

### 3.2.c. Industria textil y detergentes para ropa

En la industria textil, la catálisis biológica es una alternativa a los tratamientos químicos, que tienen consecuencias medioambientales negativas. Las celulasas (endo y exoglucanasas y  $\beta$ -glucosidasas) se aplican en procesos como el desgastado de telas (*biostoning*) y el acabado de fibras de lino y algodón o similares (*biopolishing*). A su vez, se utilizan pectinasas para remover pectina y ceras de las telas de algodón y en lanas (*bioscouring*).

Además, las celulasas se agregan en los detergentes para el lavado de ropa, ya que modifican las fibras de celulosa, contribuyendo a la remoción de manchas y mejorando el aspecto y la suavidad de las telas. Actualmente, se utilizan endo y exoglucanasas de hongos de los géneros *Humicola*, *Trichoderma* y *Aspergillus* y de bacterias del género *Bacillus*, entre otros. Muchas de estas enzimas son activas en condiciones levemente alcalinas y/o a baja o alta temperatura (Sharma y col., 2016).

### 4. *Cellulomonas* sp.

La mayoría de los cócteles enzimáticos comerciales para la degradación de lignocelulosa se basan en mezclas de celulasas y hemicelulasas fúngicas (Bhattacharya y col., 2015). Sin embargo, las bacterias ocupan hábitats y nichos muy diversos y específicos, por lo que son consideradas un reservorio genético y constituyen una fuente potencial de celulasas y hemicelulasas novedosas (Li y col., 2009). Estas enzimas pueden ser adicionadas a las mezclas tradicionales para mejorar el rendimiento. Se ha visto que los hongos producen principalmente celulasas, con baja secreción de hemicelulasas (<1% p/p de las proteínas secretadas). Las bacterias, si bien tienen niveles más bajos de secreción, producen una mayor variedad de hemicelulasas, que pueden ser adicionadas a los extractos fúngicos. La mezcla de celulasas de los hongos *T. reesei* y *A. niger* y hemicelulasas ( $\alpha$ -arabinofuranosidasas, xilanasas,  $\alpha$ -glucuronidasas y  $\beta$ -xilosidasas) de bacterias como *Clostridium thermocellum*, *Dictyoglomus turgidum* y *Geobacillus* sp. resultó en un mayor rendimiento de glucosa y xilosa a partir de rastrojo de maíz pretratado con amoníaco (AFEX) (Gao y col., 2010, 2011). Particularmente, los miembros del orden *Actinomycetales* han sido descritos como fuente de enzimas lignocelulolíticas. Las especies más estudiadas son *Thermobifida fusca*, *Microbispora bispora*, *Cellulomonas fimi* y, en menor medida, *Cellulomonas flavigena* (Saini y col., 2015).

Los actinomicetos miembros del género *Cellulomonas* son quemo-organotróficos, anaerobios facultativos y se caracterizan por su capacidad de degradar celulosa y hemicelulosas (Abt y col., 2010). Resultan positivos para la tinción de Gram, no forman esporas y pueden ser móviles, con presencia de un flagelo polar o subpolar, o de múltiples flagelos laterales, aunque también hay especies inmóviles (como *C. cellasea* y *C. uda*, por ejemplo). La temperatura óptima de crecimiento es de 30°C. Su genoma tiene un alto contenido de G+C, de entre 68,5 y 76% (Stackenbrandt y Schumann, 2015).

Actualmente (enero de 2019), hay 26 especies descritas en la List of Prokaryotic Names with Standing in Nomenclature (bacterio.net) (Parte, 2018) para el género *Cellulomonas*: *C. aerilata*, *C. biazotea*, *C. bogoriensis*, *C. carbonis*, *C. cartae*, *C. cellasea*, *C. cellulans*, *C. chitinilytica*, *C. composti*, *C. denverensis*, *C. fermentans*, *C. fimi*, *C. flavigena*, *C. gelida*, *C. hominis*, *C. humilata*, *C. iranensis*, *C. marina*, *C. massiliensis*, *C. oligotrophica*, *C. pakistanensis*, *C. persica*, *C. phragmiteti*, *C. soli*, *C. terrae*, *C. turbata*, *C. uda*, *C. xylanilytica*. Estas especies han sido aisladas de suelos principalmente de suelos, por ejemplo, forestales (*C. persica*, *C. iranensis*) (Elberson y col., 2000), de jardines botánicos (*C. soli*, *C. oligotrophica*) (Hatayama y col., 2013), o de minas de carbón (*C. carbonis*) (Shi y col., 2012), entre otros. También se las ha aislado de otros ambientes, como compost (*C. composti*) (Kang y col., 2007), árboles en descomposición (*C. xylanilytica*) (Rivas y col., 2004), granos de arroz (*C. pakistanensis*) (Ahmed y col., 2014), aguas marinas profundas (*C. marina*) (Zhang y col., 2013), perifiton de lagos salinos (*C. phragmiteti*) (Rusznýák y col., 2011), sedimentos en litorales (*C. bogoriensis*) (Jones y col., 2005) y aire (*C. aerilata*) (Lee y col., 2008).

# **Hipótesis y objetivos**

**Hipótesis:** Los sistemas de enzimas celulolíticas y hemicelulolíticas de un aislamiento nativo de *Cellulomonas* sp. pueden ser aplicados a la deconstrucción de la biomasa lignocelulósica para la obtención de bioetanol.

**Objetivo general:** Estudiar las bases moleculares de la actividad celulolítica y hemicelulolítica bacteriana y desarrollar complejos enzimáticos para la sacarificación de biomasa lignocelulósica.

**Objetivo específico:** Estudiar el sistema de enzimas activas sobre polisacáridos de *Cellulomonas* sp. B6 y evaluar su aplicación en la degradación de polisacáridos estructurales de biomasa para obtención de etanol 2G.

**Objetivos Particulares:**

**OP1. Evaluación de la actividad celulolítica y hemicelulolítica de extractos enzimáticos de *Cellulomonas* sp. B6.**

1.1. Selección de condiciones de cultivo para maximizar las actividades celulolíticas y xilanolíticas de los extractos enzimáticos de *Cellulomonas* sp. B6.

1.2. Caracterización de las actividades enzimáticas de *Cellulomonas* sp. B6 y ensayos de actividad de los extractos enzimáticos sobre biomasa lignocelulósica.

**OP2. Caracterización genómica, bioquímica y proteómica del aislamiento *Cellulomonas* sp. B6**

2.1. Secuenciación del genoma, caracterización filogenética y bioquímica.

2.2. Anotación del genoma y análisis de secuencias codificantes para enzimas activas sobre polisacáridos (CAZoma).

**OP3. Análisis proteómico de enzimas activas sobre carbohidratos (CAZimas) secretadas por *Cellulomonas* sp. B6.**

3.1. Identificación del conjunto de proteínas secretadas (secretoma) en diferentes condiciones de cultivo.

3.2. Comparación del secretoma experimental con el secretoma teórico.

**OP4. Expresión recombinante de una CAZima de *Cellulomonas* sp. B6 de importancia biotecnológica.**

4.1. Expresión recombinante de una enzima identificada en el secreoma y purificación de la enzimas en estado soluble y activo.

4.2. Caracterización bioquímica y evaluación de la actividad de la enzima en procesos de bioconversión de biomasa lignocelulósica.

# **Materiales y métodos**

## 1. Medios de cultivo y soluciones

Los medios de cultivo y soluciones principales que fueron utilizados en este trabajo de Tesis se describen en la tabla M1. Para medios sólidos, la concentración de ágar en medios de cultivo fue 15 g/l. Todos los medios de cultivo fueron esterilizados por autoclave (121°C, 20 min) previo a su uso y almacenados a 4°C. Los antibióticos (kanamicina 25 µg/ml, cloranfenicol 10 µg/ml y ampicilina 100 µg/ml) fueron agregados al medio de cultivo previamente autoclavado en condiciones de esterilidad. Las soluciones empleadas para la purificación de proteínas nativas por cromatografía de afinidad son los descritos en el manual QIAExpressionist<sup>®</sup>, de QIAGEN: (<https://www.qiagen.com/mx/resources/resourcedetail?id=79ca2f7d-42fe-4d62-8676-4cfa948c9435&lang=en>).

**Tabla M1.** Medios de cultivo y soluciones utilizados en el trabajo de Tesis. <sup>1</sup> Hankin y Anagnostakis, 1977; Ghio y col., 2012; <sup>2</sup> Bertani, 1951; <sup>3</sup> Hanahan, 1983; <sup>4</sup> Miller, 1959; <sup>5</sup> Green y Rogers, 2013; <sup>6</sup> Sambrook y col., 1989.

Medios de cultivo		
Medio	Reactivo	Concentración final (g/l)
Medio salino mínimo (MM) <sup>1</sup>	K <sub>2</sub> HPO <sub>4</sub>	1,67
	KH <sub>2</sub> PO <sub>4</sub>	0,87
	NaCl	0,05
	MgSO <sub>4</sub> ·7H <sub>2</sub> O	0,1
	CaCl <sub>2</sub>	0,04
	FeCl <sub>3</sub>	0,004
	NaMoO <sub>4</sub> ·2H <sub>2</sub> O	0,005
	NH <sub>4</sub> Cl	1,0
	Ácido nicotínico	0,02
	Ácido pantoténico	0,01
	Biotina	0,01
LB <sup>2</sup>	Triptona	10,0
	Extracto de levadura	5,0
	NaCl	10,0
SOB/SOC <sup>3</sup>	Triptona	20,0
	Extracto de levadura	5,0
	NaCl	0,58
	KCl	0,19
	MgCl <sub>2</sub> anhidro	0,95
	MgSO <sub>4</sub> ·7H <sub>2</sub> O	1,2
	Glucosa (para SOC)	3,6

<b>Soluciones</b>		
<b>Solución</b>	<b>Reactivo</b>	<b>Cantidad</b>
<b>Solución de ácido 3,5-dinitrosalicílico (DNS) <sup>4</sup></b>	Ácido 3,5-dinitrosalicílico (DNS)	10,6 g
	NaOH	19,8 g
	Tartrato de sodio y potasio	306 g
	Fenol	7,6 ml
	Metabisulfito de sodio	8,3 g
	H <sub>2</sub> O destilada	1,416 l
	<b>Reactivo</b>	<b>Concentración</b>
<b>TFB I pH = 5,8 <sup>5</sup></b>	Acetato de potasio	30 mM
	KCl/ RbCl	100 mM
	CaCl <sub>2</sub> .2H <sub>2</sub> O	10 mM
	MnCl <sub>2</sub> .4H <sub>2</sub> O	50 mM
<b>TFBII pH 6,0 <sup>5</sup></b>	Mops	10 mM
	KCl/RbCl	10 mM
	CaCl <sub>2</sub> .2H <sub>2</sub> O	75 mM
	Glicerol	15%
<b>Buffers (soluciones amortiguadoras)</b>		
<b>Buffer</b>	<b>Reactivo</b>	<b>Concentración</b>
<b>TAE 50X <sup>6</sup></b>	Tris-base	242 g/l
	EDTA disódico	18,51 g/l
	Ácido acético glacial	57,1 g/l
<b>LAEMMLI 10X <sup>6</sup></b>	Tris-base	30 g/l
	Glicina	144 g/l
	SDS	10 g/l
<b>Buffer de siembra para proteínas (SDS-PAGE) <sup>7</sup></b>	Tris-HCl pH 6,8	90 mM
	Glicerol	20% (p/v)
	Azul de bromofenol	0,02% (p/v)
	DTT	1 M
	SDS	2% (p/v)
<b>Solución de tinción de proteínas con Coomassie Brilliant Blue</b>	Coomassie Blue R-250 (Thermo Fisher Scientific)	2,5 g/l
	Etanol	45% (v/v)
	Ácido acético glacial	10% (v/v)
<b>Solución decolorante de geles de proteínas</b>	Etanol	20% (v/v)
	Ácido acético glacial	15% (v/v)
<b>Buffer de lisis pH 8,0 (purificación de proteínas)</b>	NaH <sub>2</sub> PO <sub>4</sub>	50 mM
	NaCl	300 mM
	Imidazol	10 mM
	Inhibidor de proteasas	1 mg/ml
<b>Buffer de lavado pH 8,0 (purificación de proteínas)</b>	NaH <sub>2</sub> PO <sub>4</sub>	50 mM
	NaCl	300 mM
	Imidazol	20 mM
<b>Buffer de elución pH 8,0 (purificación de proteínas)</b>	NaH <sub>2</sub> PO <sub>4</sub>	50 mM
	NaCl	300 mM
	Imidazol	250 mM

## 2. Biomosas: pretratamiento y composición

Las biomosas utilizadas en el presente trabajo de Tesis fueron: residuo agrícola de caña de azúcar molido (RAC), paja de trigo (PTE), paja de cebada (PCE), marlo de maíz dulce (MMD) y papel de desecho (PD).

El RAC fue provisto por el Dr. Alejandro Valeiro, del INTA EEA Famaillá (Tucumán, Argentina) y fue molido en a un tamaño de 0,3 mm en el Inst. de Suelos, CIRN, CNIA, INTA.

La paja de trigo y de cebada fueron pre-tratadas por extrusión (70°C, 6% NaOH, 2 mm) en la Unidad de Biocarburantes, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas (CIEMAT) (Madrid, España).

El MMD fue pretratado por extrusión (idénticas condiciones) en el Laboratoire de Chimie Agro-industrielle (LCA), Univ. de Toulouse, INRA-INPT (Francia).

El PD fue provisto por el laboratorio de Biorefinerías, Departamento de Biotecnología Aplicada y Ciencias de la Alimentación, de la Universidad de Tecnología y Economía de Budapest (BME)(Hungría). El PD fue procesado hasta formar una solución densa, homogénea, al 10% p/p en *buffer* acetato 0,1 M, pH=5 y autoclavado.

La composición de las biomosas (Tabla M2) fue determinada siguiendo las especificaciones estándar internacionales del National Renewable Energy Laboratory (NREL) de Estados Unidos (Sluiter y col., 2008) en los respectivos laboratorios de origen, salvo el RAC que fue analizado en el CIEMAT.

**Tabla M2.** Composición de las biomosas empleadas en el trabajo de Tesis. ND: no determinado.

Biomasa	Composición relativa (% p/p)			
	Celulosa	Hemicelulosa	Lignina total	Otros
RAC	32,6	23,9	20,0	23,6
PTE	46,8	30,0	21,1	2,2
PCE	34,5	27,8	19,7	11,6
MMD	42,1	26,3	6,3	ND
PD	58,7	13,3	11,4	12,2

### 3. Cultivo del aislamiento *Cellulomonas* sp. B6 y obtención de extractos enzimáticos

*Cellulomonas* sp. B6 fue aislado a partir de un consorcio celulolítico proveniente de suelo de bosque subtropical nativo de la provincial de Misiones, Argentina (Sabaris di Lorenzo, 2013). La información sobre la muestra fue depositada en la base de datos *Biosample* del National Center for Biotechnology Information (NCBI) (NCBI Resource Coordinators, 2017) bajo el número de acceso SAMN04287673.

Los cultivos se realizaron en LB o en MM suplementado con extracto de levadura 0,1% como fuente de nitrógeno, con el agrado de sustratos lignocelulósicos, en general al 1% p/v (10 g/l) salvo que se indique de otra manera.

Los cultivos se incubaron a 30°C en un agitador orbital (Innova 42, New Brunswick) con agitación (200 rpm). La cosecha se hizo a las 24, 48 o 72 horas por centrifugación a 5000 x g durante 20 minutos. Los sobrenadantes de cultivos en biomasa fueron filtrados con filtro de fibra de vidrio #32 (1,2 µm) (Schleicher and Schnell) para remover las trazas de biomasa. El *pellet* bacteriano se conservó a -20°C hasta su utilización. Para obtener los extractos intracelulares, se resuspendió el *pellet* en *buffer* fosfato 0,1 M a pH 7 (1/10 volumen de cultivo) y fue sometido a sonicado (6 pulsos de 10 segundos, 75 W), para la ruptura celular. La suspensión fue centrifugada a 4000 rpm y el sobrenadante fue considerado el extracto proteico intracelular. Se agregó azida de sodio (NaN<sub>3</sub>) 0,04% p/v para la preservación de los extractos, que se guardaron a 4°C.

El aislamiento fue almacenado a -80°C con glicerol 20% p/v para su conservación. La cepa fue depositada en los bancos internacionales de cepas NCIMB (número de depósito 15124) y DSMZ (Número DSM 107934).

### 4. Crecimiento en papel de filtro

Se evaluó la capacidad de crecimiento en aerobiosis y anaerobiosis de *Cellulomonas* sp. B6 en papel de filtro como única fuente de carbono por cultivo en MM líquido (en aerobiosis) o en medio de fluido ruminal modificado (Dehority, 1969) con una tira de PF de 1 cm x 6 cm durante 7 días a 30°C. La actividad se detectó como turbidez, deterioro del PF y producción de burbujas de gas.

## 5. Microscopía electrónica de barrido

Las muestras de las bacterias adheridas a papel de filtro (Materiales y Métodos, apartado 4) se visualizaron por microscopía electrónica de barrido (SEM) en un microscopio modelo FEI Quanta-250 (FEI Co., Países Bajos), en el Laboratorio de Microscopía del CICVyA, INTA. El tratamiento previo a la visualización consistió en el agregado de glutaraldehído 2,5% v/v a la muestra, seguido de la resuspensión del *pellet* y de la incubación durante toda la noche (ON, 16 horas) a 4°C. Posteriormente, se hicieron 3 lavados con *buffer* fosfato de sodio para eliminar el glutaraldehído y tiñó el papel de filtro con tetróxido de osmio (OsO<sub>4</sub>) durante 1 hora. La muestra fue lavada con agua y deshidratada con etanol a concentraciones crecientes (30 a 80% v/v) y luego cubierta con una película de oro para la visualización por SEM.

## 6. Determinación del metabolismo de hidratos de carbono de *Cellulomonas* sp B6

El metabolismo de hidratos de carbono se determinó mediante las pruebas API 50 CH (BioMérieux) según las especificaciones del fabricante. Los sustratos incluidos en estas pruebas son (en orden alfabético): ácido-2-cetoglucónico, ácido-5-cetoglucónico, adonitol, almidón, amigdalina, arbutina, β-D-metilmanopiranosido, celobiosa, D-arabinosa, D-arabitol, D-fructosa, D-fucosa, D-glucosa, D-lixosa, D-glucosa, D-lixosa, D-manosa, D-rafinosa, D-tagatosa, D-turanosa, dulcitol, D-xilosa, eritriol, esculina, galactosa, glicerol, glucógeno, inositol, inulina, lactosa, L-arabinosa, L-arabitol, L-fucosa, L-sorbosa, L-xilosa, maltosa, manitol, melibiosa, melizitosa, β-D-metilxilopiranosido, N-acetilglucosamina, gluconato de potasio, ramnosa, ribosa, sacarosa, salicina, sorbitol, trehalosa, xilitol, β-gentibiosa, β-D-metilglucopiranosido.

Las determinaciones fueron realizadas en el laboratorio de IMyZA (CNIA, INTA) y los resultados se compararon con los publicados para *Cellulomonas flavigena* ATCC 482, *Cellulomonas fimi* ATCC 484 y *Cellulomonas persica* ATCC 700642 (Ahmed y col., 2014; Elbertson y col., 2000; Stackebrandt y Schumann, 2015).

## 7. Determinación de la composición de ácidos grasos

Los ésteres metílicos de ácidos grasos fueron extraídos y analizados según los métodos descritos por Ferrari y col. (2018) e identificados usando Sherlock Microbial Identification System (MIDI) versión 4.5 (Sherlock MIDI, Microbial ID, Newark, DE, USA) (Sasser, 1990) en el Laboratorio de Bioquímica, Microbiología e Interacciones Biológicas en el Suelo, Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Buenos Aires, Argentina.

## 8. Actividades enzimáticas del sobrenadante de cultivo

### 8.1. Actividades endoglucanasa (CMCasa) y xilanasa

Para medir la actividad endoglucanasa (CMCasa) se usó CMC 1% p/v como sustrato. La actividad xilanasa se midió sobre glucuronoxilano de madera de haya (*beechwood*) o de abedul (*birchwood*) y arabinoxilano de trigo, todos al 0,5% p/v. La mezcla de reacción consistió en 0,25 ml de extracto enzimático y 0,25 ml de sustrato en el *buffer* correspondiente.

La reacción se llevó a cabo en agitación durante 30 minutos, tras lo cual se detectaron los azúcares reductores liberados por reacción con 1 ml de DNS, incubando a 100°C durante 5 minutos. Las curvas de calibración fueron construidas con glucosa, para medir actividad CMCasa, y xilosa, para la medición de la actividad xilanasa (ambas en un rango de 0,0 a 1,0 mg/ml).

Se tuvieron en cuenta los resultados correspondientes a los controles del buffer de reacción, el sustrato y la enzima, que se hicieron por duplicado. La actividad se calculó en Unidades Internacionales (UI) por ml de sobrenadante, según la ecuación 1:

$$Actividad \left( \frac{UI}{ml} \right) = \frac{Ca \left( \frac{mg}{ml} \right) + Vr (ml)}{PMa \left( \frac{mg}{\mu mol} \right) \times Ve \times minuto}$$

(Ecuación 1),

donde:

Ca: concentración de azúcar (glucosa o xilosa), Vr: volumen de reacción, PMa: peso molecular del azúcar (glucosa o xilosa), Ve: volumen de enzima purificada o de sobrenadante de cultivo.

Una UI se define como la cantidad de enzima que libera 1  $\mu\text{mol}$  de equivalentes de glucosa o xilosa por minuto en condiciones de reacción determinadas.

En el caso particular del sobrenadante de cultivo crecido en papel de desecho (PD), la reacción se llevó a cabo en un volumen final de 1,5 ml, con 0,5 ml de sustrato en buffer de reacción y 1,0 ml de extracto enzimático. Para el sobrenadante en papel de filtro, se midió la actividad xilanasa sobre arabinoxilano de trigo, glucuronoxilano y xiloglucano (0,5%). La reacción se detuvo por agregado de 3 ml de DNS y ebullición durante 5 minutos. La absorbancia se detectó luego de agregar 15 ml de agua corriente a la mezcla y se midió a 550 nm. Las UI/ml se calcularon como se describe en la ecuación 1.

## 8. 2. Actividad celulasa total

La actividad celulosa total en el sobrenadante de cultivo se midió usando celulosa amorfa regenerada (PASC) 0,5% (Zhang y col., 2006) como sustrato y se determinó en un tiempo de reacción de 24 horas. El resto del protocolo se siguió de la manera detallada en el apartado anterior (8.1) para las actividades CMCasa y xilanasa.

La PASC se preparó a partir de 200 mg de Avicel en 0,6 ml de agua destilada, mezcla a la cual se le agregó 10 ml de ácido fosfórico ( $\text{H}_2\text{PO}_4$ ) frío en forma paulatina. La mezcla fue incubada a 4°C durante una hora con agitación suave, o hasta obtener una solución transparente. Luego, se añadieron 10 ml de agua destilada fría y se centrifugó la solución a 5000 xg por 20 minutos a 4°C. Este procedimiento se repitió cuatro veces para eliminar el ácido fosfórico residual del pellet. Adicionalmente, se le agregaron 0,5 ml de carbonato de sodio ( $\text{Na}_2\text{CO}_3$ ), a fin de neutralizar el ácido fosfórico remanente. El *pellet* se resuspendió en 15 ml de agua destilada fría y se centrifugó a 5000 g durante 20 minutos a 4°C, tras lo cual se hicieron dos lavados más con agua destilada. Finalmente, se obtuvieron 15 ml de PASC 1% p/v en agua. El peso seco fue confirmado por secado en estufa hasta peso constante.

## 8. 3. Actividades $\beta$ -glucosidasa, $\beta$ -xilosidasa y celobiohidrolasa

Las actividades  $\beta$ -glucosidasa,  $\beta$ -xilosidasa y celobiohidrolasa se detectaron por un ensayo colorimétrico que consiste en la reacción con *p*-nitrofenil  $\beta$ -D-glucopiranosido

(pNP-G) 5mM, *p*-nitrofenil β-D-xilopiranosido (pNP-X en inglés) 4 mM y *p*-nitrofenil β-D-cellobiosido (pNP-C) 5 mM, respectivamente, y la posterior detección del paranitrofenol (PNP) liberado por hidrólisis de los enlaces β-1,4. El cálculo de cantidad de PNP liberado se hizo en base a una curva de calibración de la absorbancia a 410 nm en función de los μmoles totales de PNP (0,0 a 0,2 μmoles totales de PNP) (Campos y col., 2014). El tiempo de reacción fue de 30 minutos, deteniéndose la actividad por agregado de Na<sub>2</sub>CO<sub>3</sub> 2% p/v.

En todos los casos, se hicieron los controles correspondientes al buffer de reacción, el sustrato y la enzima. La actividad se expresó en μmoles de PNP liberados por minuto y ml (UI/ml).

## 9. Zimogramas

Los zimogramas se hicieron en geles de poliacrilamida 10%, con geles de separadores co-polimerizados con CMC o xilano y geles concentradores sin agregado de sustratos. Para ello, se prepararon soluciones acuosas de cada sustrato al 1%, esterilizadas por autoclavado, y se agregaron a la mezcla habitual de componentes del gel, en reemplazo del agua, para una concentración final de CMC 0,2%, o de xilano 0,1% (Ghio y col., 2016). Las muestras fueron sembradas en *buffer* de siembra con SDS pero sin DTT (Tabla M1), sin calentamiento previo. La velocidad de corrida fue de 100 V.

Una vez separadas las proteínas, se cortó la calle correspondiente al marcador de tamaño molecular y se renaturalizaron las proteínas por lavado con *buffer* TrisHCl pH 6,8/ 0,04 M (3 lavados de media hora cada uno, a temperatura ambiente, en agitación). El gel se dejó toda la noche en el mismo buffer para favorecer la renaturalización de las proteínas y, al día siguiente, se realizó un último cambio de *buffer*, seguido de incubación a 30°C durante 1 hora, para promover la actividad enzimática.

Finalmente, se lavó el gel con agua estéril y se tiñó durante 15 minutos con Rojo Congo 0,1% (p/v), tras lo cual se destiñó con durante 15 minutos con una solución de NaCl 1 M para la observación de los halos de degradación (Teather y Wood, 1982). El gel fue conservado en agua destilada hasta su documentación.

## **10. Secuenciación, ensamblado y anotación del genoma de *Cellulomonas* sp. B6**

Se construyeron dos bibliotecas genómicas a partir de ADN total cromosómico de cultivos independientes de *Cellulomonas* sp. B6 crecido en LB, que fueron secuenciadas usando la plataforma Illumina MiSeq (Servicio de Secuenciación del Instituto de Biotecnología, CICVyA, INTA). El genoma fue ensamblado *de novo* en 279 contigs con recursos computacionales de la Unidad de Bioinformática del Instituto de Biotecnología, INTA Castelar. Se utilizó la herramienta Trimmomatic v0.33 para efectuar el *trimming* de las lecturas (Bolger y col., 2014) y el ensamblado fue llevado a cabo mediante Celera Assembler v.8.2, seguido de SPAdes Genome Assembler v.3.5.0. La versión no cerrada (*draft*) del genoma, que fue depositada en el Whole Genome Submission portal del NCBI bajo el número de acceso: LNTD000000000 (RefSeq assembly accession: GCF\_001462455.1).

La anotación del genoma se hizo en forma automática mediante dos algoritmos de anotación disponibles en la web, el Prokaryotic Genome Annotation Pipeline (PGAP) del NCBI y el algoritmo de Rapid Annotation using Subsystems Technology (RAST) (<http://rast.nmpdr.org>) (Aziz y col., 2008).

## **11. Análisis filogenético en base a la secuencia de la subunidad 16S del ARN ribosomal**

En base a la secuencia completa de la subunidad ribosomal 16S obtenida por secuenciación se construyeron dos filogramas mediante los métodos *Maximum Likelihood* (Máxima Verosimilitud) y *Neighbour-Joining* (Vecino Más Cercano) (Tamura y col., 2013; Saitou y Nei, 1987) usando el programa MEGA v.7.0 (Kumar y col., 2016). Las secuencias incluidas en el análisis corresponden a las vigentes en el mes de agosto de 2017 (<http://www.bacterio.net/>) para el género *Cellulomonas* y a *Micrococcus luteus*, que se usó como *outgroup*. Las secuencias de 16S se obtuvieron de la base de datos Ribosomal Database Project (<https://rdp.cme.msu.edu/>) (Cole y col., 2014) y

fueron todas de largo  $\geq 1200$  pb. El análisis estadístico se hizo con 2000 repeticiones de *bootstrap* (Feselstein, 1985).

## 12. Comparación *in silico* de genomas

Se hizo una comparación *in silico* del genoma de *Cellulomonas* sp. B6 con los genomas de *Cellulomonas flavigena* ATCC 482 (GCA\_000092865.1), *Cellulomonas fimi* ATCC 484 (GCA\_000212695.1) y *Cellulomonas persica* ATCC 700642<sup>T</sup> (GCA\_001315245.1), cuyas secuencias fueron obtenidas de la base de datos Assembly de NCBI. La estimación de la hibridización ADN-ADN con el genoma de *Cellulomonas* sp. B6 (GCA\_001462455.1) fue realizado por la herramienta *online* Genome to Genome Distance Calculator v 2.1 (GGDC) (<http://ggdc.dsmz.de/>) (Meier-Kolthoff y col., 2013) del Instituto Leibniz DSMZ – Colección Alemana de Microorganismos y Cultivos. La identidad nucleotídica promedio (ANI) y la identidad aminoacídica promedio (AAI) se calcularon con programas disponibles en la *web* desarrollados por el Environmental Microbial Genomics Laboratory, Georgia Institute of Technology (<http://enve-omics.ce.gatech.edu/>) (Goris y col., 2007).

## 13. Identificación de enzimas activas sobre carbohidratos

Las enzimas activas sobre carbohidratos (CAZimas) codificadas por *Cellulomonas* sp. B6 se identificaron con el programa dbCAN, disponible en la *web* ([csbl.bmb.uga.edu/dbCAN/](http://csbl.bmb.uga.edu/dbCAN/)) (Yin y col., 2012), a partir de las secuencias de proteínas predichas por anotación. De esta manera, se determinaron los dominios catalíticos y de unión a sustrato correspondientes a la clasificación de la base de datos CAZy ([cazy.org](http://cazy.org)) (Lombard y col., 2014).

Posteriormente, se hizo una curación por búsqueda de proteínas homólogas mediante identidad de secuencia usando el algoritmo Basic Local Alignment Search Tool para proteínas (BLASTp) del NCBI (Atschul y col., 1990). Teniendo en cuenta los resultados de dbCAN, de la anotación por NCBI y del alineamiento de secuencias, se le asignó una actividad potencial a cada CAZima.

Las proteínas se clasificaron en intracelulares o extracelulares (secretadas) por la ausencia o presencia, respectivamente, de un péptido señal en su secuencia

aminoacídica, lo que se determinó con el programa SignalP versión 4.1 (para bacterias Gram positivas) disponible en la web (<http://www.cbs.dtu.dk/services/SignalP/>) (Petersen y col., 2011). Adicionalmente, se identificaron proteínas que podrían ser secretadas por vías independientes de péptido señal usando el servidor SecretomeP versión 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>) (Bendtsen y col., 2005). Las proteínas sin péptido señal y con un puntaje  $\geq 0,5$  en el análisis por SecretomeP (límite sugerido para bacterias Gram positivas) fueron clasificadas como potencialmente secretadas por vías alternativas al péptido señal.

## **14. Análisis del secretoma de *Cellulomonas* sp. B6**

### **14.1 Preparación de las muestras**

*Cellulomonas* sp. B6 se cultivó durante 72 horas a 30°C (200 rpm) en MM suplementado con sacarosa 0,5%, CMC 1%, RAC 1% o PTE 1% (todos p/v) y el sobrenadante de cultivo libre de células fue obtenido como se describió previamente (Materiales y Métodos, apartado 3). Los cultivos se hicieron por duplicado.

Las proteínas totales en el sobrenadante de cultivo se midieron por triplicado para cada cultivo con el *kit* comercial MicroBCA Protein Assay (Thermo Fisher Scientific). Las actividades xilanas y CMCasa también se determinaron por triplicado para cada cultivo.

Las proteínas contenidas en 10 ml de sobrenadante con fueron precipitadas con tricloroacetato 10% y resuspendidas en 1 ml de agua (18  $\Omega$ ). Estas soluciones fueron filtradas previamente a la detección de proteínas por espectrometría de masas (tamaño de poro de 0,22  $\mu\text{m}$ ).

Los sobrenadantes de cultivo crecido en sacarosa se usaron como control no inductor de la secreción de enzimas involucradas en la degradación de sustratos celulósicos.

### **14.2. Detección de proteínas por espectrometría de masas (LC MS/MS)**

La digestión y el análisis por espectrometría de masas fueron realizados por el CEQUIBIEM (Facultad de Ciencias Exactas, Universidad de Buenos Aires) mediante LC-MS/MS, es decir, por espectrometría de masas acoplada a HPLC, con un espectrómetro

modelo Q-Exactive (Thermo Fisher Scientific) y un equipo nano HPLC EASY-nLC 1000 (Thermo Fisher Scientific). Este ensayo se hizo por duplicado, con dos muestras independientes para cada medio de cultivo. Los datos fueron analizados con el software Proteome Discoverer v.2.1.1.21 (Thermo Scientific) y comparados con las secuencias de proteínas de *Cellulomonas* sp. B6 depositadas en las bases de datos de NCBI.

La abundancia relativa de cada proteína se estimó en base al índice de abundancia (emPAI) obtenido por Sequest a partir de datos de identificación de proteínas (Ishihama y col., 2005; Shinoda y col., 2010). La abundancia relativa fue expresada como mol% (emPAI%), según la ecuación 2:

$$\text{emPAI\%} = \text{emPAI} / \sum (\text{emPAI}) \times 100 \quad (\text{Ecuación 2}).$$

## 15. Expresión recombinante de GH10xynC

Los plásmidos que se usaron en el presente trabajo de Tesis se describen en la tabla M3.

**Tabla M3.** Plásmidos utilizados en el trabajo de Tesis.

Plásmido	Promotor	Resistencia a antibióticos	Procedencia
pGEM <sup>®</sup> -T easy	lacZ	Amp <sup>R</sup>	Promega
pET28a (+)	T7	Kan <sup>R</sup>	EMD Biosciences
pGEMT-GH10XynC	lacZ	AmpR	Este trabajo
pET28a-GH10XynC	T7	KanR	Este trabajo

### 15.1. Amplificación

La amplificación por reacción en cadena de la polimerasa (PCR) (Mullis, 1990) de la secuencia codificante para la enzima GH10XynC de *Cellulomonas* sp. B6 (GenBank: KSW20567.1, RefSeq: WP\_062103211.1) se realizó con *primers* diseñados mediante el programa Primer3 de la plataforma JustBio (<http://www.justbio.com/>), siendo el *primer* sentido (*forward*) gh10EMf: **5'TTGGATCCGCCGCCGCCGGCAGCACGCTCCAG3'** y el antisentido (*reverse*) gh10EMfr: **5'TTAAGCTTTTACGACGCCGAGCAGGTCAGCGTCGG3'**. Las secuencias incluyen los sitios de corte para las enzimas de restricción BamHI (gh10EMf) y de HindIII (gh10EMr)(en

negrita), que, con el agregado de un codón de terminación (subrayado) en gh10EMr, permitieron el clonado en el vector pET28a y la expresión de la proteína fusionada a una etiqueta de 6 histidinas en su extremo N-terminal. Las temperaturas de desnaturalización se calcularon con la herramienta Oligocalc, también disponible en JustBio, y fueron de 75°C y 72°C, para gh10EMf y gh10EMr, respectivamente.

Dado el alto contenido de G+C de la secuencia y las altas temperaturas de desnaturalización de los cebadores, se usó una polimerasa Taq en conjunto con la solución Q, ambas de QIAGEN, y el ciclado se realizó con una temperatura alta de hibridación. La mezcla completa de reacción para la amplificación consistió en: 10x CoralLoad *buffer* (QIAGEN) 1x, solución Q 1x, dinucleótidos trifosfatados (dNTPs) 0,4 mM, gh10EMf 0,4 µM, gh10EMr 0,4 µM, ADN genómico de *Cellulomonas* sp. B6 0,4 ng/µl, polimerasa Taq (QIAGEN) 0,05 µl/ng ADN. Luego de una desnaturalización inicial de 3 minutos a 94°C en el termociclador (BioRad), se repitió 30 un ciclo de 30 segundos a 94°C de desnaturalización, seguido de 30 segundos a 67°C para la hibridación y 2 minutos a 72°C como extensión. Una vez finalizado el ciclado, se programó un paso de extensión final de 10 minutos a 72°C.

La separación de fragmentos de ADN por electroforesis se realizó en gel de agarosa 0,8% en TAE 1X (Tabla M1), en un campo eléctrico constante (80-100 V) (Sambrook y col., 1989). Se agregó bromuro de etidio 0,5 µg/ml para la visualización de las bandas con luz UV. Las imágenes fueron documentadas mediante el equipo GelDoc (Bio-Rad). La banda correspondiente fue escindida del gel y purificada para su clonado. El tamaño molecular de las bandas de interés se estimó con el marcador 1 kb de Promega.

## 15.2. Clonado

El producto de amplificación fue clonado primero en el vector comercial pGEM<sup>®</sup>-T easy, según las especificaciones del fabricante, generando el plásmido pGEMT-GH10XynC. Se transformaron *E. coli* DH5α quimiocompetentes (Sambrook y col., 1989) para la producción de un elevado número de copias y la extracción del ADN plasmídico.

La identidad del inserto se confirmó por secuenciación en el servicio de Secuenciación de Ácidos Nucleicos por electroforesis capilar de la Unidad de Genómica

del INTA Castelar (Buenos Aires), usando los cebadores para los promotores SP6 y T7 (Axelrod y Kramer, 1985).

El inserto fue liberado utilizando las enzimas de restricción BamHI y HindIII (Promega), purificado de gel de agarosa y ligado en pET28a con la ligasa de T4 (Promega), generando el plásmido pET28a-GH10XynC. Posteriormente, se transformó *E.coli* DH5 $\alpha$  y se corroboró la inserción del inserto en el marco de lectura por secuenciación a partir del cebador para el promotor T7 (Axelrod y Kramer, 1985).

### **15.3. Expresión recombinante y purificación**

Se transformaron *E. coli* Rossetta (DE3) pLys (Novagen) quimiocompetentes con el plásmido pET28a-GH10XynC. La expresión de la proteína fue inducida cuando el cultivo alcanzó una DO = 0,4-0,6 por agregado de IPTG 0,5 mM e incubación a 20°C ON con agitación.

La purificación de la proteína recombinante se realizó por cromatografía de afinidad en columna utilizando la resina Ni-NTA (QIAGEN), de acuerdo a las especificaciones descritas en el manual QIAExpressionist<sup>®</sup> (QIAGEN), con los *buffers* de lisis, lavado y elución correspondientes (Tabla M1).

### **15.4. SDS-PAGE**

La separación de proteínas fue efectuada por electroforesis en gel de poliacrilamida, a una concentración de acrilamida –bisacrilamida (29:1) de 10 o 12%, con agregado de SDS (SDS-PAGE). Las muestras para condiciones desnaturizantes se calentaron a 100°C durante 10 minutos previo a la siembra en el gel. Las proteínas se revelaron por tinción con Coomassie *Brilliant Blue* (Tabla M1). El peso molecular de las bandas se estimó con el marcador ProteinRuler 10-180 kDa (Thermo Fisher Scientific).

## **16. Análisis de la secuencia de GH10XynC**

La secuencia de GH10XynC se analizó usando la información de la base de datos de Dominios Conservados (CDD) de NCBI y de la anotación por dbCAN y mediante la herramienta ScanProsite de Expasy (<https://prosite.expasy.org/scanprosite/>) (De Castro y col., 2006).

El tamaño molecular y el pI se predijeron usando ExPASy Compute pI/MW Tool, de ([https://web.expasy.org/compute\\_pi/](https://web.expasy.org/compute_pi/)) (Bjellqvist y col., 1993).

El entorno genómico se determinó en base a la anotación por RAST y PGAP de NCBI. La predicción de regiones promotoras se realizó con programa BPROM, de Softberry (<http://www.softberry.com/berry.phtml>) (Soloyev y Salamov, 2011).

## 17. Análisis filogenético de GH10XynC

El análisis filogenético de GH10XynC se realizó por separado para cada módulo, dado que son las regiones más conservadas en las CAZimas, a diferencia de las secuencias *linker* (Sammond y col., 2012).

Mediante BLASTp, se identificaron dominios de otras proteínas con alta identidad de secuencia (identidad  $\geq 80\%$ , cobertura  $\geq 95\%$ ) y se construyó un árbol filogenético enraizado por *Neighbour-Joining* con el programa Mega v.7.0, haciendo un análisis estadístico mediante 1000 repeticiones de *bootstrap* (Saitou y Nei, 1987; Tamura y col., 2013). Las secuencias de los dominios GH10 y CBM2 de una proteína de *Ruminococcus albus* (acceso de GenBank: ADU21425.1) fueron incluidas como *outgroup* en la construcción del árbol correspondiente.

## 18. Modelado molecular de GH10XynC

La estructura tridimensional de la proteína fue predicha a partir de su secuencia aminoacídica completa mediante la plataforma I-TASSER (homology modelling by Iterative Threading ASSEmbly Refinement) disponible en la *web* (<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>) (Yang y col., 2015). El algoritmo se describe brevemente a continuación.

El primer paso consiste en el alineamiento de regiones de la secuencia de interés con fragmentos de proteínas disponibles en la Protein Data Bank (PDB) ([rcsb.org](http://rcsb.org)) (Berman y col., 2000), cuya estructura tridimensional (o "supra-secundaria") está resuelta. El alineamiento se realiza con LOMETS, un servidor de *meta-threading*, que permite generar miles de modelos usando 9 programas distintos. Aquéllos fragmentos que se alinean con mayor score son tomados en cuenta por I-TASSER para la

reconstrucción global de la proteína. Las regiones que no resultan en ningún alineamiento significativo son modeladas por métodos *ab initio*.

Finalmente, se arma la estructura tridimensional global teniendo en cuenta las simulaciones que determinan las estructuras de menor energía libre y establecen las restricciones espaciales.

Este modelo es validado por varios parámetros: C-score (*score* de confianza del modelo), que está normalmente en un rango de [-5, 2]; TMscore (*score* de los alineamientos con los templados), que varía entre 0 y 1; y RMSD (raíz cuadrada de la desviación media de las posiciones atómicas por superposición de proteínas). Un C-score alto y un TMscore cercano a 1 indican una buena calidad de alineamiento (Roy y col., 2011).

La función biológica se predice a partir de la estructura tridimensional resultante, que se compara con modelos predichos de proteínas de las que se conoce su actividad (número EC y términos GO- *gene ontology*) y las regiones de unión a ligando.

## **19. Caracterización bioquímica de la xilanasa rGH10XynC**

### **19. 1. Actividades enzimáticas**

La concentración de proteínas de las fracciones de elución se midió con el método de Bradford (1979), usando el reactivo de Bio-Rad y adaptado a microplaca (<http://www.bio-rad.com/webroot/web/pdf/lsr/literature/4110065A.pdf>). La curva de calibración se construyó con con seroalbúmina bovina (BSA) (0 a 2 mg/ml). La concentración óptima de enzima para los ensayos de actividad se determinó por reacción sobre xilano, efectuada como se describió previamente (Materiales y Métodos, apartado 8.a), con concentraciones de enzima de 0,15; 0,3; 0,6; 0,9 y 1,2 µg/ml y se estableció en 0,6 µg/ml (concentración final en la reacción). El tiempo de reacción se determinó también por actividad sobre xilano a 0, 5, 10, 15, 20, 25 y 30 minutos de reacción y se estableció en 10 minutos.

El perfil de la actividad según el pH se ensayó en el rango de 3,0 a 10,0 (3,0 a 6,0 en *buffer* citrato; 6,5 a 8,0 en *buffer* citrato-fosfato; 8,5 a 10,0 en *buffer* glicina-NaOH) por reacción a 50°C. En el pH correspondiente a la máxima actividad, se determinó el efecto de la temperatura (40 a 65°C).

La actividades  $\beta$ -xilosidasa y endoglucanasa (medidas sobre pNP-X y CMC, respectivamente) se ensayaron a concentraciones de enzima comprendidas entre los 0,3 y los 1,2  $\mu\text{g/ml}$  y a tiempos de reacción de entre 10 y 20 minutos.

## 19.2. Cinética de reacción

La cinética de reacción de rGH10XynC purificada sobre xilano se ensayó mediante la detección de los azúcares reductores liberados en un tiempo corto de reacción (8 minutos), usando concentraciones de sustrato en un rango de 0,2 a 25  $\text{mg/ml}$ . Primero, se ensayó la reacción a tiempos cortos (0, 2, 3, 5, 8, 10, 15 y 20 minutos) y se determinó el tiempo en que la reacción transcurre en forma lineal ( $V_0$ ). Luego, se midió la actividad enzimática frente a concentraciones crecientes de xilano de abedul (0; 0,2; 0,5; 1; 2; 5; 10; 15; 20 y 25  $\text{mg/ml}$ ) al tiempo correspondiente a  $V_0$ . A partir de los resultados, se evaluó el ajuste a los modelos de Michaelis-Menten y de alosterismo con el programa GraphPad Prism version 6.04 para Windows, GraphPad Software, La Jolla California USA ([www.graphpad.com](http://www.graphpad.com)). Los parámetros cinéticos  $K_M$ , y  $V_{\text{max}}$  se calcularon con el mismo programa.

## 20. Hidrólisis de biomasa

Los ensayos de hidrólisis de biomasa se hicieron a una concentración de sólidos del 2% p/v en buffer de reacción por incubación durante 24-72 horas con agitación (200 rpm) a la temperatura y pH óptimos para los extractos enzimáticos o las proteínas purificadas empleadas. El sobrenadante de cultivo fue filtrado con filtros de nylon de 0,2  $\mu\text{m}$  antes de la detección de azúcares por cromatografía líquida de alta eficiencia (HPLC). El porcentaje de conversión a azúcares simples se calculó según la guía de procedimientos de laboratorio de NREL (Sluiter y col., 2008) (ecuación 3):

$$\% \text{ conversión} = \frac{[\text{glucosa o xilosa final (mg/ml)}] * 100}{[\text{glucano o xilano inicial (mg/ml)}] * \text{FC}} \quad \text{Ecuación 3,}$$

donde: FC: factor de corrección entre el monómero libre y cuando está formando parte del polímero (anhidro). FC para xilosa: 1,12. FC para xilobiosa =1,06.

## **21. Cuantificación de productos de reacción por cromatografía de alta precisión (HPLC)**

La detección de los productos de reacción de hidrólisis de sustratos comerciales y de biomasa fue realizada en el Laboratorio de Servicios Analíticos Especiales (LABFAUBA) con un equipo de HPLC modelo Agilent 1100 conectado a un detector de índice de refracción (RID). Se utilizó una columna Aminex 87P o Rezex RPM-monosaccharide Pb (8%) (Phenomenex), usando agua como fase móvil a un flujo de 0,4 ml/min y con un volumen de inyección de muestras de 10 µl. La columna fue mantenida a 75°C durante la corrida y la temperatura del detector fue de 35°C. La curva de calibración se construyó con los azúcares glucosa, celobiosa, xilosa, xilobiosa, xilotriosa y xilotetraosa (Sigma) en un rango de concentraciones de 0,125 a 2 mg/ml.

En todos los casos se incluyeron los controles de sustrato sin enzimas y de enzimas sin sustrato, y se restaron, de corresponder, los productos no resultantes de la hidrólisis enzimática.

## **22. Detección de productos de reacción por cromatografía en capa delgada**

Los xilooligosacáridos y celooligosacáridos producidos por acción enzimática se visualizaron mediante cromatografía en capa delgada (TLC), con una fase estacionaria de sílica gel (Merck). Como fase móvil, se usó una solución de butanol, ácido acético y agua en una relación 2:1:1 (butanol: ácido acético: agua). El revelado se hizo con una solución de ácido sulfúrico, etanol y agua en una relación 3:70:20 (ácido sulfúrico: etanol: agua) con el agregado de orcinol 1% y aplicación de calor (Ghio y col., 2016).

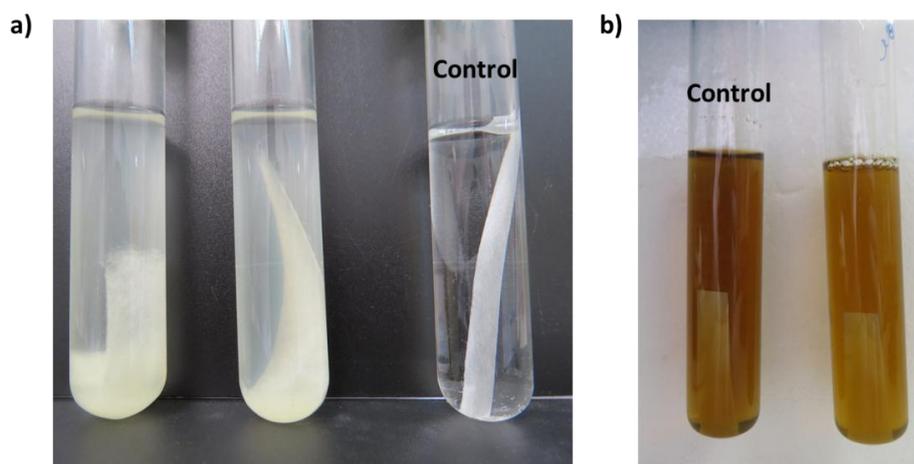
Los estándares usados fueron glucosa, celobiosa, celotriosa y celotetraosa para productos de hidrólisis de sustratos celulósicos y xilosa, xilobiosa, xilotriosa, xilotetraosa, xilopentaosa y xilohexaosa para los productos de hidrólisis de xilanos (Megazymes y Sigma). La mezcla de estándares se sembró a una concentración individual de 0,25 mg/ml.

# Resultados

**Parte 1. Actividad celulolítica y  
hemicelulolítica de extractos  
enzimáticos de *Cellulomonas* sp. B6**

En trabajos previos, se obtuvo un aislamiento bacteriano puro con actividad celulolítica y xilanolítica a partir de una muestra de suelo de bosque nativo de Misiones (Argentina). El aislamiento fue identificado como perteneciente al género *Cellulomonas* por análisis filogenético de la secuencia parcial de la subunidad ribosomal 16S (Sabaris di Lorenzo, 2013) y fue denominado *Cellulomonas* sp. B6. En este trabajo, se planteó como primer objetivo caracterizar las actividades enzimáticas de *Cellulomonas* sp B6 y optimizar las condiciones de cultivo para maximizar las actividades celulolíticas y xilanolíticas de los extractos enzimáticos.

Para confirmar la naturaleza celulolítica de *Cellulomonas* sp. B6 y su clasificación como género anaerobio facultativo, se analizó su crecimiento en medio mínimo con papel de filtro (PF) como única fuente de carbono (celulosa cristalina  $\geq 98\%$ ), en condiciones de aerobiosis y anaerobiosis durante 7 días (Figura R1).



**Figura R1.** Crecimiento de *Cellulomonas* sp. B6 en papel de filtro en condiciones de a) aerobiosis y b) anaerobiosis. En ambos casos, se muestran los signos de degradación de papel luego de 7 días de incubación y un control sin inóculo bacteriano.

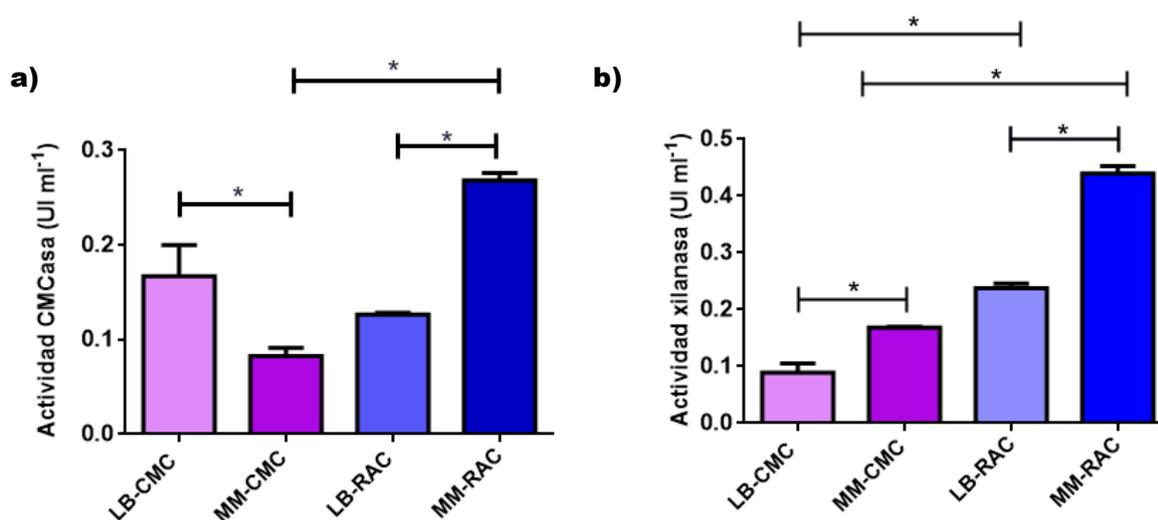
Tras siete días de incubación en aerobiosis, *Cellulomonas* sp. B6 fue capaz de degradar visiblemente el papel de filtro. En anaerobiosis, si bien la degradación fue menos visible, se observó el crecimiento del cultivo por aumento de turbidez del medio, formación de burbujas y signos de degradación del papel.

Estos resultados confirman que *Cellulomonas* sp. B6 puede degradar y utilizar celulosa como fuente de carbono, tanto en aerobiosis como en anaerobiosis.

### 1.1. Selección del medio de cultivo para maximizar las actividades celulolíticas y xilanolíticas de los extractos enzimáticos de *Cellulomonas* sp. B6

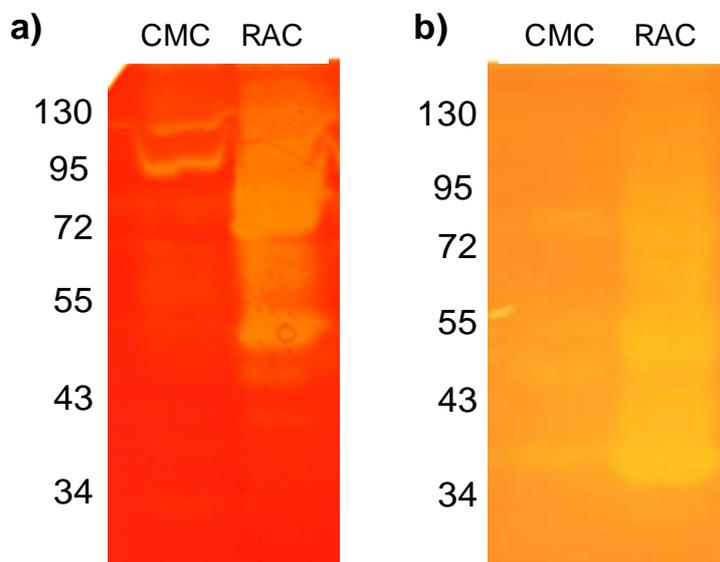
En base a resultados previos del laboratorio, se comparó la actividad enzimática extracelular de cultivos en medio mínimo (MM) o LB con residuo agrícola de caña de azúcar (RAC) o carboximetilcelulosa (CMC) como única fuente de carbono (sustrato de cultivo). Se determinó la actividad xilanasa y endoglucanasa (CMCasa) en los sobrenadantes de cultivo libre de células cosechados luego de 72 horas de incubación a 30°C (Figura R2). Las actividades endoglucanasa y xilanasa se detectaron por hidrólisis de CMC y xilano de madera de haya, respectivamente, tras incubación con el sobrenadante de cultivo (Figura R2).

Los niveles más altos de actividad xilanasa y endoglucanasa, y de proteínas extracelulares (LB-CMC: 19,5 µg/ml; MM-CMC: 3,4 µg/ml; LB-RAC: 35,6 µg/ml; MM-RAC: 51,7 µg/ml), se obtuvieron por cultivo en MM-RAC, por lo que este medio fue seleccionado para la obtención de extractos y la caracterización de las actividades enzimáticas de *Cellulomonas* sp. B6. Se muestran sólo las diferencias estadísticamente significativas de actividad entre MM o LB, para cada sustrato y entre RAC o CMC, para cada medio.



**Figura R2.** Actividades enzimáticas de los extractos extracelulares de *Cellulomonas* sp. B6 en LB y MM suplementados con CMC o RAC. a) Actividad CMCasa; b) Actividad xilanasa. Los medios de cultivo se indican en el eje X. Las diferencias significativas se determinaron por análisis de la varianza (ANOVA) unifactorial seguido del test de Tukey (\* $p < 0,01$ ).

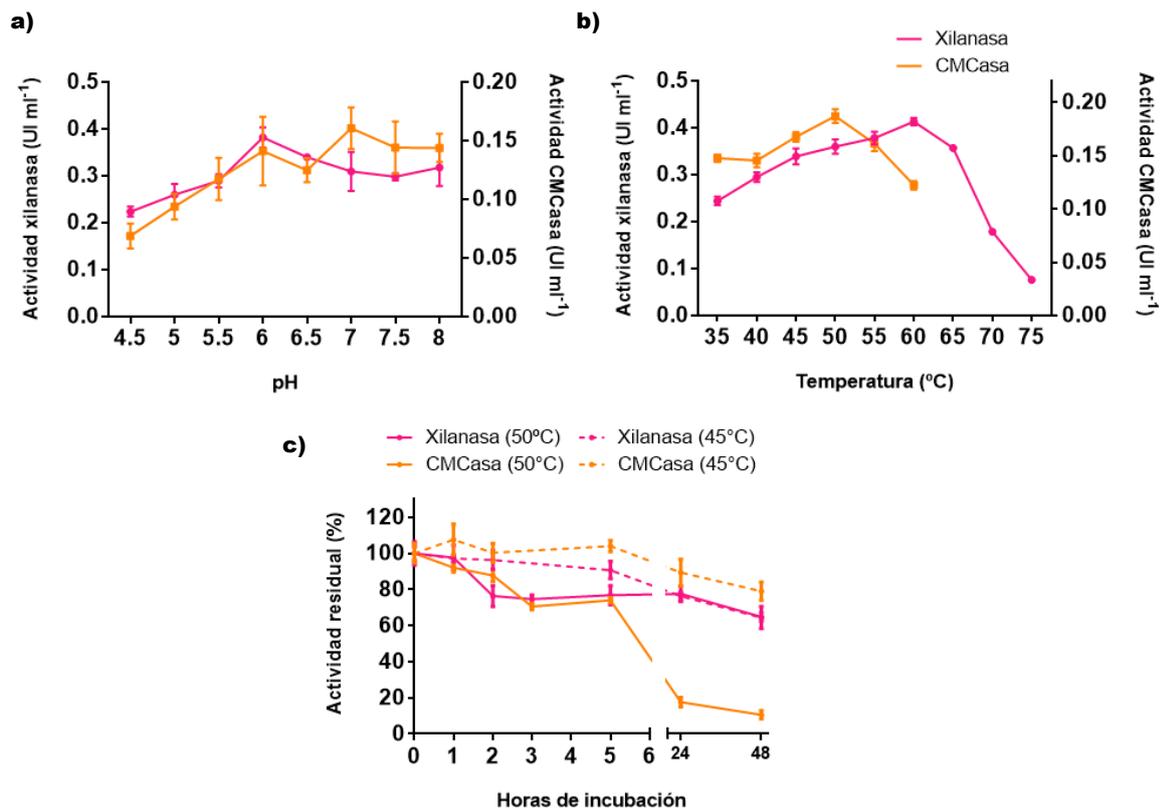
Coincidentemente, en zimogramas sobre CMC y xilano de los sobrenadantes de cultivo, la mayor intensidad de bandas se observó en el sobrenadante de cultivos en RAC (Figura R3 a y b).



**Figura R3.** Zimografía de sobrenadante de cultivo de *Cellulomonas sp.* B6 cultivado en MM-CMC o MM-RAC, en a) CMC 0,2% y b) xilano 0,1%.

## 1.2. Caracterización de las actividades enzimáticas de *Cellulomonas sp.* B6

Una vez seleccionado el medio de cultivo, se construyó el perfil de las actividades xilanasas y CMCasa del sobrenadante de cultivo en función del pH (Figura R4 a) y de la temperatura (Figura R4 b). De esta manera, se establecieron las condiciones óptimas de reacción y se evaluó la estabilidad de las actividades xilanasas y CMCasa en el extracto extracelular a la temperatura óptima y a una temperatura inferior (Figura R4 c).



**Figura R4.** Perfil de las actividades xilanasa y CMCasa del extracto extracelular de *Cellulomonas* sp. B6 en función de a) el pH y b) la temperatura. c) Estabilidad térmica a 45 y 50°C de las actividades CMCasa y xilanasa.

La actividad CMCasa fue máxima a pH 6 y la actividad xilanasa a pH 7. A pH óptimo, la actividad xilanasa fue máxima a 60°C ( $0,41 \pm 0,01$  UI/ml) y la actividad CMCasa fue máxima a 50°C ( $0,16 \pm 0,01$  UI/ml). La estabilidad térmica se estudió incubando el sobrenadante a 45°C y 50°C, pH 6 previamente a la reacción enzimática en condiciones óptimas, tomando como 100% el tiempo inicial sin pre-incubación (tiempo cero). A 50°C, la actividad xilanasa se mantuvo por encima del 60% hasta por 48 hs, mientras que la actividad CMCasa se redujo en un 80% a las 24 horas de incubación. Ambas actividades se mantuvieron por encima del 60% a las 48 hs a 45°C.

Se determinaron también las actividades celobiohidrolasa,  $\beta$ -glucosidasa y  $\beta$ -xilosidasa y celulasa total en el sobrenadante de cultivo de MM-RAC. Dado que las actividades relacionadas con la degradación final de oligosacáridos cortos se han descrito en otras bacterias como de localización intracelular, las actividades  $\beta$ -glucosidasa y  $\beta$ -xilosidasa también fueron medidas en los extractos intracelulares (Tabla R1). El ensayo se realizó a 50°C, pH 6 durante 30 minutos (con excepción de la actividad celulasa total, que se determinó a las 24 horas de reacción y a 45°C).

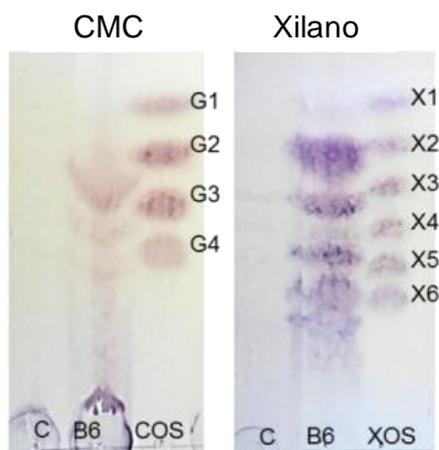
**Tabla R1.** Actividades enzimáticas determinadas en los extractos extra (E) e intracelular (I) de *Cellulomonas* sp. B6 (en MM-RAC).

Actividad (sustrato)	Extracto	mUI/ml
Xilanasa (xilano de abedul)	E	410,00 ± 10,00
Endoglucanasa (CMC)	E	160,00 ± 10,00
Celulasa total (PASC)	E	4,37 ± 1,60
β-xilosidasa (pNP-X)	E	9,14 ± 0,28
	I	277,65 ± 0,18
β-glucosidasa (pNP-G)	E	0,58 ± 0,14
	I	28,58 ± 6,05
Celobiohidrolasa (pNPC)	E	2,09 ± 0,12

Se evaluaron también otras biomásas lignocelulósicas como sustratos de cultivo y se determinaron las actividades xilanasa y CMCasa en los extractos extracelulares. En particular, la paja de trigo y de cebada, pretratadas por extrusión (PTE y PCE, respectivamente), resultaron en niveles similares de actividad (actividad CMCasa: PTE 207,0 ± 11,9 mUI/ml; PCE 85,6 ± 23,1 mUI/ml; actividad xilanasa: PTE 707,6 ± 19,4 mUI/ml; PCE 248,0 ± 42,4 mUI/ml). Paralelamente, en colaboración con el grupo de Biotecnología aplicada de la Universidad Tecnológica de Budapest (BME) (<http://biorefinery.bme.hu/en/>), se evaluó el crecimiento de *Cellulomonas* sp. B6 en papel de descarte (PD) (cartón molido y autoclavado), lo que resultó en niveles similares de actividad (actividad xilanasa: en arabinoxilano 337,0 ± 29,2 mUI/ml, en xilano de abedul 561,8 ± 4,6 mUI/ml, en xilano de haya 467,8 ± 74,3 mUI/ml; endoglucanasa 171,8 ± 1,6 mUI/ml; xiloglucanasa: 77,5 ± 1,6 mUI/ml).

Estos resultados demuestran que *Cellulomonas* sp. B6 es capaz de crecer en múltiples biomásas y secretar enzimas capaces de degradar los polisacáridos estructurales, en todos los casos.

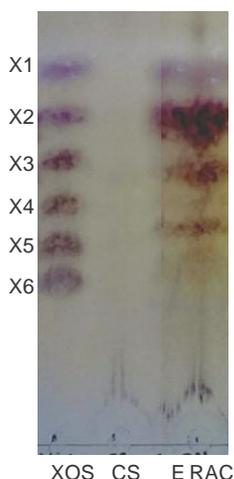
Los productos de reacción sobre CMC y xilano del extracto extracelular de cultivo en MM-RAC fueron identificados por cromatografía en capa delgada (TLC). El principal producto de degradación de CMC fue celobiosa (G2) y se observó también celotriosa (G3) y celotetraosa (G4), mientras que en la degradación de xilano se observaron xilooligosacáridos solubles de DP 2 a 6 (X2 a X6) (Figura R5). Se detectaron también trazas de xilosa, pero no de glucosa, lo cual se correlaciona con los niveles de actividad β-xilosidasa y β-glucosidasa determinados previamente.



**Figura R5.** Productos de reacción sobre CMC y xilano de haya, usando sobrenadante de cultivo de *Cellulomonas* sp. B6. Los productos fueron separados por cromatografía en capa delgada (TLC) e identificados en base a estándares. Estándares de celooligosacáridos (COS): glucosa (G1), celobiosa (G2), celotriosa (G3) y celotetraosa (G4). Estándares de xilooligosacáridos (XOS): xilosa (X1), xilobiosa (X2), xilotriosa (X3), xilotetraosa (X4), xilopentaosa (X5) y xilohexaosa (X6). En todos los casos, se muestra un control de sustrato sin agregado de enzima (C).

### 1.3. Ensayos de actividad de los extractos enzimáticos sobre biomasa lignocelulósica

Para evaluar la potencialidad de los extractos enzimáticos de *Cellulomonas* sp. B6 en la sacarificación de biomasa lignocelulósica, se evaluó la actividad del extracto obtenido en MM-RAC en la degradación de PCE o PTE. Por TLC, observamos que la xilobiosa es el producto principal, si bien se ven también xilotriosa (X3) y xilooligosacáridos (XOS) de mayor DP.



**Figura R6.** Productos de hidrólisis de PTE usando el sobrenadante de *Cellulomonas* sp. (de MM-RAC) como extracto enzimático (E RAC). CS: Control de sustrato.

Luego, se ensayó el extracto extracelular tal cual (1X) o concentrado por ultrafiltración (2X), solo o complementado con el extracto intracelular y se cuantificaron los productos de reacción por cromatografía líquida de alta eficiencia (HPLC). También, se determinaron los azúcares reductores generados a partir de los polisacáridos contenidos en la biomasa para cada ensayo por reacción con DNS. En todos los casos, el único producto cuantificable por HPLC fue X2. El mayor porcentaje de conversión de xilano a X2 se obtuvo con el extracto extracelular 2X, que fue del 12%. Se muestran los resultados para la hidrólisis de PCE (Tabla R2).

**Tabla R2.** Xilobiosa (X2) obtenida a partir de xilano contenido en biomasa, cuantificada por HPLC, usando los extractos extra (E) y/o intracelular (I) de *Cellulomonas* sp. B6. \* $p < 0,05$  (ANOVA unifactorial seguido de prueba t desapareada); los ensayos comparados se indican con el número luego del asterisco (\*).

Extracto	Azúcares reductores (mg/ml)	X2 (mg/ml)	% Conversión
E (2x)	1,66 ± 0,05	0,54 ± 0,02 * <sup>1</sup>	12,2 ± 0,4
E (1x)	1,16 ± 0,12	0,41 ± 0,02 * <sup>1*2</sup>	9,4 ± 0,5
I	1,15 ± 0,01	0,29 ± 0,04	6,6 ± 0,8
I + E (1:1)	1,39 ± 0,03	0,47 ± 0,02* <sup>2</sup>	10,7 ± 0,4

Estos resultados confirman que la principal actividad en los sobrenadantes de cultivo en biomasa de *Cellulomonas* sp. B6 es la actividad xilanasa.

**Parte 2. Caracterización genómica,  
bioquímica y proteómica del  
aislamiento *Cellulomonas* sp. B6**

## 2.1. Secuenciación del genoma y caracterización filogenética

Con el fin de explicar las actividades enzimáticas de *Cellulomonas* sp. B6 y de avanzar en la caracterización filogenética del aislamiento, se secuenció el genoma a partir de dos cultivos independientes. La secuenciación del genoma se realizó usando la plataforma Illumina MiSeq (Servicio de Secuenciación del Instituto de Bioetecnología, CICVyA, INTA). Se obtuvieron 1.532.556 lecturas apareadas, con un tamaño promedio de 500 pb, lo que corresponde a una cobertura del genoma de 83X. Las secuencias con baja calidad fueron descartadas o recortadas con el programa Trimmomatic v.3. El ensamblado se realizó con el programa Celera Assembler v.8.2 (CA) en primer lugar, que genera *contigs* de muy buena confianza, y luego con el SPAdes Genome Assembler v.3.5.0 que intenta unir *contigs* con lecturas que CA no utilizó. Se generaron 279 *contigs* ( $N_{50} = 24612$ ) con una longitud total de secuencia de 4.042.325 pb. El alto contenido de GC (75,1%), característico del género, dificultó el armado de los *contigs*.

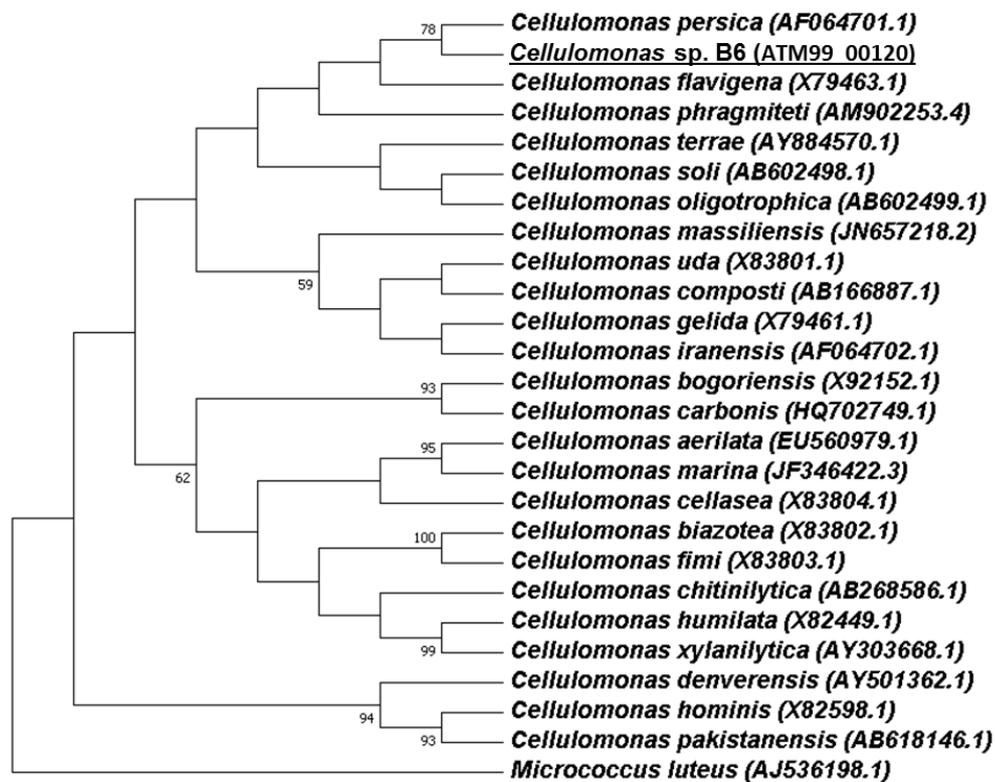
De esta manera, se obtuvo una versión preliminar (*draft*) del genoma, que fue depositada en la base de datos pública del NCBI ([ncbi.nlm.nih.gov/genbank/wgs](http://ncbi.nlm.nih.gov/genbank/wgs)) bajo el número de acceso de RefSeq Assembly: GCF\_001462455.1 (Tabla R3).

**TABLA R3.** Metodología y resultados de la secuenciación por Illumina Miseq y el ensamblado del genoma de *Cellulomonas* sp. B6.

Tecnología de secuenciación	Illumina Miseq
Método de ensamblado	Celera Assembler v. 8.2, SPAdes Genome Assembler v. 3.5.0
Número de lecturas apareadas	1.532.556
Longitud total de <i>contigs</i> (pb)	4042325
Cobertura del genoma	83X
N50	24612
L50	50
Contenido de G+C (%)	75,1
Número de acceso de NCBI	LNTD00000000 (Assembly); GCF_001462455.1 (RefSeq)

Para el análisis filogenético, se extrajo la información correspondiente a la secuencia completa de la subunidad ribosomal 16S (NCBI: ATM99\_00120) y se construyeron filogramas por los métodos *Maximum Likelihood* y *Neighbour-Joining*

con secuencias correspondientes a todas las especies del género *Cellulomonas* vigentes hasta agosto de 2017 (Figura R7).



**Figura R7.** Filograma construido en base a la secuencia de la subunidad ribosomal 16S del aislamiento B6 y las especies del género *Cellulomonas* (vigentes en agosto de 2017). El árbol presentado corresponde al método *Neighbor-Joining*. El análisis estadístico se hizo con 2000 repeticiones de *bootstrap* (sólo se muestran los valores  $\geq$  50%). El número de acceso de GenBank de cada secuencia se indica entre paréntesis.

En ambos casos, *Cellulomonas* sp. B6 se agrupó con la especie *Cellulomonas persica* ATCC 700642/ DSM-14784 en un 78% de las repeticiones de *bootstrap* (Figura R7). También podría estar emparentado con *Cellulomonas flavigena* ATCC 482/ DSM 20109, una especie más caracterizada que *C. persica*. La identidad de secuencia del ARNr 16S de B6 fue de 98,54% y 97,63% con *C. persica* y *C. flavigena*, respectivamente. En base a los resultados, no fue posible asignar a B6 a una de estas especies.

### 2.1.a. Genómica comparativa

Actualmente, con el avance de las técnicas de secuenciación de nueva generación (NGS) y la disponibilidad de información sobre genomas bacterianos completos en bases de datos públicas, se han propuesto nuevas definiciones de especie en base a la

comparación *in silico* de genomas, como alternativa a la hibridación experimental entre genomas (DDH) en la clasificación taxonómica.

Entonces, con el objetivo de asignar una especie al aislamiento *Cellulomonas* sp. B6, su genoma fue comparado *in silico* con los correspondientes a las especies tipo más emparentadas, es decir, *C. flavigena* ATCC 482 y *C. persica* ATCC 700642. Adicionalmente, se incluyó en el análisis a la especie tipo *C. fimi* ATCC 484, ya que es la más estudiada del género (Tabla R4). Las secuencias genómicas de las tres especies están disponibles en las bases de datos Assembly y Genome de NCBI.

La comparación se realizó mediante el cálculo de la identidad nucleotídica y aminoacídica promedio (ANI y AAI, respectivamente) y de la distancia entre genomas basada en BLAST, denominada Genome BLAST Distance Phylogeny (GBDP). Ambas herramientas están basadas en la búsqueda de pares de alto puntaje (*high-scoring pairs*) (HSPs) entre los genomas (que pueden estar ensamblados en *contigs*) o proteomas y son las herramientas más aceptadas por la comunidad científica para su implementación en taxonomía bacteriana y están disponibles en plataformas *online* (Mahato y col., 2017).

Una de las principales diferencias radica en que GBDP aplica una transformación de los resultados para que emulen el porcentaje de hibridación entre genomas (DDH) que se obtendría en forma experimental (Meier-Kolthoff, 2013). Por el contrario, un valor de ANI de 95% corresponde al punto de corte tradicional de DDH de 70% (Goris y col., 2007) para definir misma especie.

**Tabla R4.** Comparación a nivel genómico de *Cellulomonas* sp. B6 y de *C. flavigena* ATCC 482, *C. persica* ATCC 700642 y *C. fimi* ATCC 484. <sup>1</sup>Datos consultados en NCBI, <sup>2</sup>Resultados obtenidos por el cálculo de GBDP (programa del Leibniz Institute DSMZ).

	<i>C. flavigena</i> ATCC 482	<i>C. persica</i> ATCC 700642	<i>C. fimi</i> ATCC 484	<i>Cellulomonas</i> sp. B6
<b>GenBank</b>	GCA_000092865	GCA_001315245	GCA_000212695	GCA_001462455
<b>Nivel de ensamblado<sup>1</sup></b>	Genoma cerrado	<i>Contigs</i>	Genoma cerrado	<i>Contigs</i>
<b>Largo (pb)<sup>1</sup></b>	4.123.179	821.903	4.266.344	4.042.325
<b>% G+C<sup>1</sup></b>	70,4%	69,0%	74,7%	75,1%
<b>Distancia<sup>2</sup></b>	0,1714	0,1428	0,2023	-
<b>DDH<sup>2</sup></b>	25,4% IC: [23,0-27,8%]	29,9% IC: [27,5-32,4%]	21,70% IC: [19,4-24,1%]	-

<b>Diferencia en % G+C<sup>2</sup></b>	0,78	6,06	0,35	-
<b>ANI bilateral</b>	83,71% (DE: 4,91%)	Alineamiento insuficiente	80,95% (DE: 5,24%)	-
<b>AAI bilateral</b>	79,14% (DE: 15,09%)	Secuencias no disponibles	70,36% (DE: 15,49%)	-
<b>Identidad de ARNr 16S</b>	97,73%	98,54%	95,07%	-

Para considerar que dos cepas son de la misma especie, se propone que exista una identidad aminoacídica y nucleotídica mayores que 95%, así como un porcentaje de hibridación *in silico* de genomas (GDDH) mayor que 70%. Además, las especies de un mismo género deben formar un grupo monofilético por análisis de la secuencia de ARNr 16S (Thompson y col., 2013).

De acuerdo con estos lineamientos, *Cellulomonas* sp. B6 podría ser una especie novedosa, distinta de las especies tipo *C. flavigena* y *C. fimi*, ya que en ningún caso se obtuvo un GDDH > 70%. Los valores de ANI y AAI se encuentran muy por debajo del punto de corte (Tabla R4).

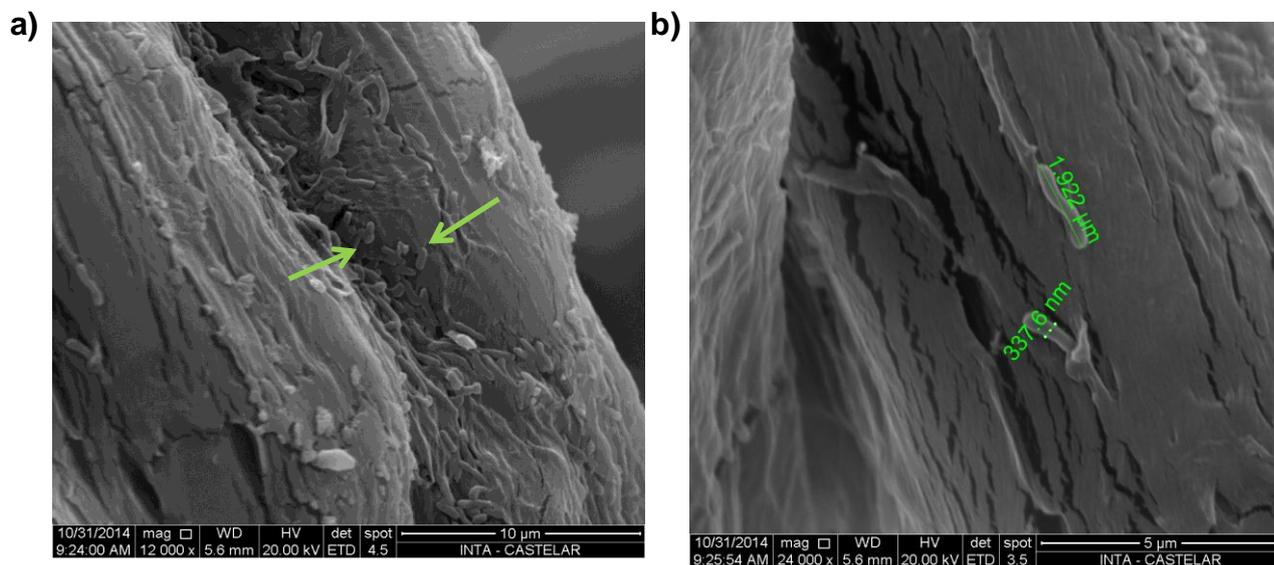
El genoma de *C. persica* no está secuenciado en su totalidad (el número de pb secuenciadas es marcadamente menor que para especies del mismo género) y las secuencias de proteínas no están disponibles, lo que impidió la comparación por ANI y AAI. Sin embargo, se tomaron en cuenta los resultados de GDDH, ya que el programa está optimizado para la comparación entre genomas incompletos (se necesita que al menos un 20% esté secuenciado). El valor entre *Cellulomonas* sp. B6 y *C. persica* de GDDH resultó muy inferior a 70% (valor de corte). Por otra parte, la diferencia de contenido G+C es mayor que 5%, valor de corte propuesto tradicionalmente para delimitación de especies (Thompson y col., 2013), aunque se debe interpretar con cuidado, dado el número de pb secuenciadas del genoma de *C. persica*.

Los genomas de *C. flavigena* y *C. fimi* presentan una diferencia de G+C < 5% con *Cellulomonas* sp. B6, pero los otros resultados validan la pertenencia a especies diferentes.

El análisis genómico fue complementado con determinaciones fenotípicas, que se describen a continuación.

### 2.1.b. Caracterización morfológica

La morfología celular de *Cellulomonas* sp. B6 fue determinada por microscopía electrónica de barrido (SEM) a partir de un cultivo crecido en anaerobiosis durante 7 días sobre papel de filtro (Figura R8).



**Figura R8.** Micrografías electrónicas de barrido de *Cellulomonas* sp. B6 sobre PF, amplificadas: a) 12000 x, las bacterias se indican con flechas, y b) 24000 x, el tamaño de la bacteria se indica en la imagen.

Las bacterias se observaron adheridas al sustrato (PF) y con forma de bacilos cortos, rectos o curvos (Figura 8 a y b) de 1,99  $\mu\text{m}$  de largo y 337,6 nm de diámetro (Figura R8 b). La morfología celular es consistente con la de todas las especies descritas del género (Stackebrandt y Schumann, 2015).

### 2.1.c. Caracterización bioquímica

El metabolismo de carbohidratos del aislamiento B6 se caracterizó mediante las pruebas API 50 CH y se comparó con la literatura de *C. flavigena* y *C. persica*, las dos especies más relacionadas filogenéticamente, y de *C. fimi* (Tabla R5).

**Tabla R5.** Resultados de API 50 CH de *Cellulomonas* sp. B6 comparados con la información publicada para *C. flavigena*, *C. persica* y *C. fimi*. (Ahmed y col., 2014; Elbertson y col., 2000; Stackebrandt and Schumann, 2015). Los resultados para *C. flavigena* fueron reportados como producción de ácidos en condiciones aeróbicas a partir de los carbohidratos correspondientes y, en el caso de *C. persica*, como crecimiento en cada sustrato, salvo que se indique lo contrario. Las diferencias entre *Cellulomonas* sp. B6 y las otras especies se indica en gris. Referencias: + positivo; - negativo; ND no determinado.

Carbohidrato	<i>Cellulomonas</i> sp. B6	<i>C. flavigena</i> ATCC 482 <sup>1</sup>	<i>C. persica</i> ATCC 700642 <sup>2</sup>	<i>C. fimi</i> ATCC 484 <sup>1</sup>
Ácido 2-cetoglucónico	-	ND	ND	ND
Ácido 5-cetoglucónico	-	-	ND	ND
Adonitol	-	ND	ND	ND
Amigdalina	+	-	ND	ND
Arbutina	+	-	ND	ND
$\beta$ -D- metilmanopiranosido	-	ND	ND	ND
Celobiosa	+	ND	ND	ND
D-arabinosa	+	ND	ND	ND
D-arabitol	-	ND	ND	ND
D-fructosa	+	ND	ND	ND
D-fucosa	-	ND	ND	ND
D-glucosa	+	ND	ND	ND
D-lixosa	-	-	ND	ND
D-manosa	+	ND	+	ND
D-rafinosa	-	-	-	-
D-tagatosa	-	-	ND	ND
D-turanosa	+	ND	ND	ND
Dulcitol	-	ND	ND	ND
D-xilosa	+	ND	ND	+
Eritritol	-	ND	ND	ND
Esculina	+	+ (hidrólisis)	ND	ND
Galactosa	+	ND	ND	ND
Glicerol	+ (débil)	ND	ND	ND
Glucógeno	+	+	ND	ND
Inositol	-	ND	ND	ND
Inulina	-	+ (débil)	ND	ND
Lactosa	+	-	-	+
L-arabinosa	+	ND	ND	ND
L-arabitol	-	ND	ND	ND
L-fucosa	-	-	ND	ND
L-sorbosa	-	ND	ND	ND

L-xilosa	-	ND	ND	ND
Maltosa	+	ND	ND	ND
Manitol	+	-	ND	-
Melibiosa	+	-	ND	ND
Melezitosa	-	-	ND	ND
$\beta$ -D-metilxilopiranosido	+	-	ND	ND
N-acetilglucosamina	-	-	ND	-
Gluconato de potasio	-	-	-	-
Ramnosa	-	-	ND	+
Ribosa	-	-	-	-
Sacarosa	+	ND	ND	ND
Salicina	+	ND	ND	ND
Sorbitol	-	-	ND	ND
Almidón	+	+	+ (hidrólisis)	ND
Trehalosa	+	ND	ND	ND
Xilitol	-	ND	ND	ND
$\beta$ -gentibiosa	+	ND	ND	ND
$\beta$ -D-metilglucopiranosido	-	-	ND	ND

*Cellulomonas* sp. B6 se distingue de *C. flavigena* por el metabolismo de amigdalina, arbutina, inulina, manitol, melibiosa,  $\beta$ -D-metilxilopiranosido y lactosa y de *C. fimi*, por el metabolismo de manitol y ramnosa. También, se diferencia de *C. persica* por su capacidad de hidrolizar lactosa.

Adicionalmente, se identificaron los ácidos grasos iso C15:0 (35,93%) y anteiso C15:0 (20,42%) como los mayoritarios en la membrana de *Cellulomonas* sp. B6, seguidos de anteiso C17:0 (4,85%) e iso C17:0 (2,92%). Si bien todos ellos han sido descritos en otros miembros del género, en *C. flavigena* predomina el ácido anteiso C15:0 y, en *C. fimi*, los más abundantes son C16:0 y anteiso C17:0. La composición de ácidos grasos *C. persica* no ha sido determinada (Stackebrandt y Schumann, 2015).

En su conjunto, todos los resultados- tanto de genómica comparativa, como de caracterización fenotípica- aportan evidencias para la posible clasificación de *Cellulomonas* sp. B6 en una especie nueva, distinta de *C. flavigena*, *C. fimi* y *C. persica*.

## 2.2. Anotación del genoma de *Cellulomonas* sp. B6

Las secuencias genómicas fueron depositadas en la base de datos de genomas del NCBI, lo que permitió la anotación automática del genoma por el Prokaryotic Genome Annotation Pipeline (PGAP). El genoma fue también anotado mediante Rapid Annotation Using Subsystems Technology (RAST).

Ambos algoritmos de anotación están basados en predicciones *ab initio* y son aptos tanto para genomas cerrados como para aquellos compuestos por múltiples *contigs* (Aziz y col., 2008; Tatusova y col., 2016). Sin embargo, presentan algunas diferencias. Por ejemplo, PGAP hace un enfoque pan-genómico, que consiste en la construcción de un conjunto de genes característicos de un clado taxonómico dado (que puede corresponder a una especie, o estar por encima de este nivel). Dentro de este conjunto, los genes se clasifican en: genes núcleo, que se encuentran en al menos el 80% de los genomas pertenecientes al clado, genes accesorios, que se identifican en 1-20% de los genomas, y genes únicos, hallados en menos del 1% de los miembros del clado. El genoma a anotar es asignado a uno de estos clados y, en un primer paso, se identifican los genes característicos del mismo. Luego, se hacen predicciones sobre otros genes que pueda haber en el genoma nuevo.

En RAST, se busca un número limitado de genes con FIGfams (construidos con modelos ocultos de Markov, HMM) establecidos, que son prácticamente universales en las bacterias, para la estimación de relaciones filogenéticas entre el genoma nuevo y los que ya están anotados y disponibles en la base de datos de RAST. Los demás FIGfams presentes en el genoma nuevo se encuentran por comparación con los genomas considerados filogenéticamente más cercanos. En el caso de *Cellulomonas* sp. B6, los genomas más relacionados ("vecinos más cercanos"), dentro de la base de datos de RAST, fueron *Cellulomonas flavigena* ATCC 482 (con un *score* de 413) y *Sanguibacter keddiei* ATCC 51767.

A pesar las diferencias en los algoritmos, los resultados de anotación obtenidos fueron muy similares.

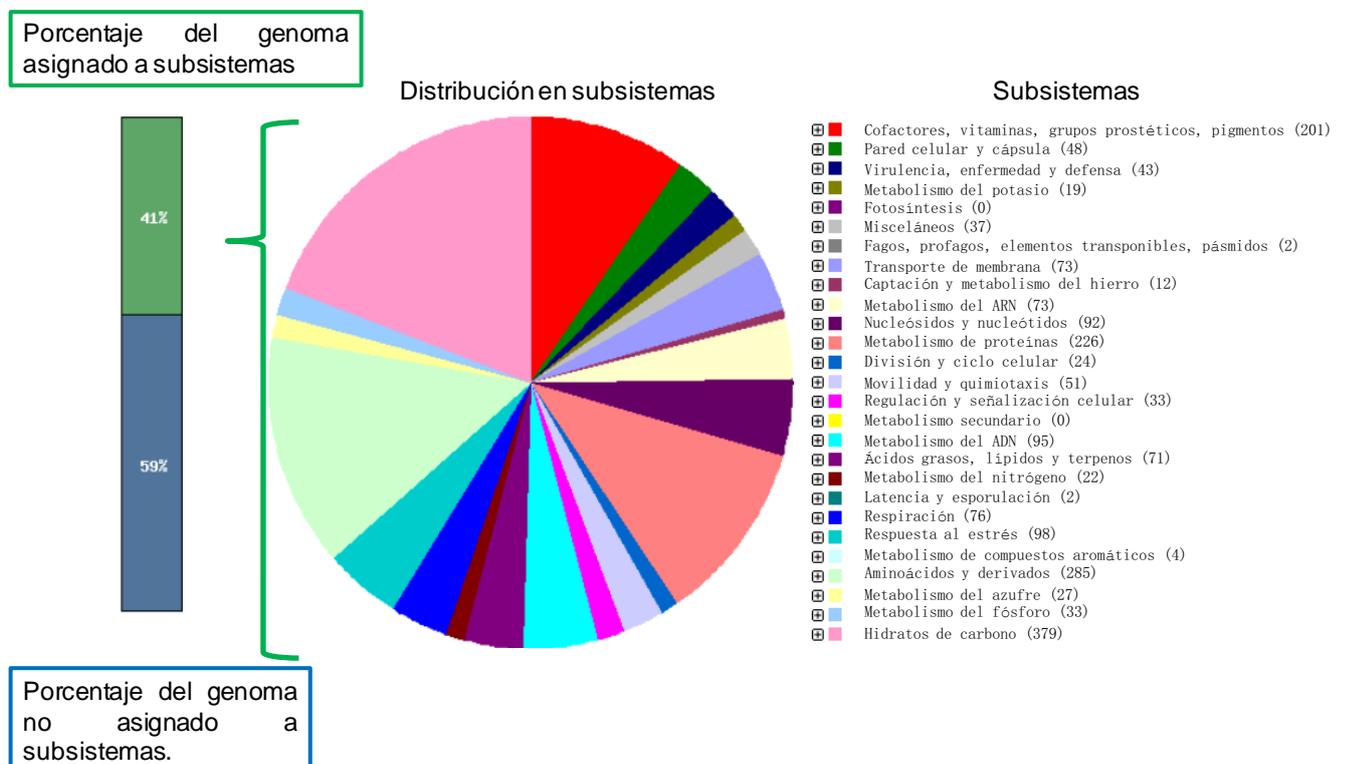
Mediante la plataforma PGAP se predijeron 3691 genes, que incluyen 3443 secuencias codificantes para proteínas, 50 pseudogenes y un ARN no codificante. También, se identificaron 50 ARNt y un set de ARNr 5S, 16S y 23S de secuencia

completa (Tabla R6). Mediante RAST se predijeron 3562 secuencias codificantes para proteínas, así como el mismo número de ARNr y de ARNr de secuencia completa que por PGAP.

**Tabla R6** Resultados de la anotación del genoma de *Cellulomonas* sp. B6 por NCBI PGAP y RAST.

	NCBI PGAP	RAST
Secuencias codificantes para proteínas	3443	3562
Pseudogenes	194	-
ARNt	50	50
ARNr (secuencia completa)	1 5S, 23S y 16S	1 5S, 23S y 16S
ARNnc (ARNasa P)	1	-
Número de subsistemas	-	374

A su vez, la plataforma RAST permite la asignación de secuencias a subsistemas, constituidos por genes de función (o "rol") relacionada y así también se pueden predecir ciertas vías metabólicas. En este caso, el 40% de las secuencias fue asignado a un subsistema (Figura R9). De éstas, 35 proteínas predichas fueron asignadas al subsistema "utilización de xilosa" y se estudiaron más detalladamente durante el desarrollo de esta Tesis.



**Figura R9.** Número de proteínas predichas asignadas a cada subsistema de RAST.

### **2.3. Análisis de secuencias codificantes para enzimas activas sobre polisacáridos (CAZimas) y construcción del CAZoma y secretoma teóricos**

Las secuencias de proteínas resultantes de la anotación por el algoritmo PGAP fueron analizadas en mayor profundidad (se eligió PGAP, ya que las bases de datos de NCBI abarcan un gran número de genomas, mayor que en el caso de RAST, y por la conexión que existe entre NCBI, ENA, Uniprot, etc.) con el objeto de analizar el potencial biotecnológico del aislamiento B6 e intentar identificar las proteínas que podrían estar involucradas en la utilización de polisacáridos estructurales (celulosa y hemicelulosa). Para ello, se analizó el genoma completo mediante la plataforma dbCAN (<http://csbl.bmb.uga.edu/dbCAN/annotate.php>) (Yin y col., 2012), que está específicamente desarrollada para este fin. Así, se identificaron todas las proteínas que contienen módulos catalíticos (CDs) pertenecientes a enzimas activas sobre carbohidratos (CAZimas) y módulos de unión a carbohidratos (CBMs) (Lombard y col., 2014). El algoritmo se basa en el reconocimiento de patrones mediante modelos ocultos de Markov (HMM) construidos para cada familia de la base de datos CAZY (actualizada regularmente), con información de la base de datos de dominios conservados de NCBI (CDD) y curados mediante análisis bibliográfico.

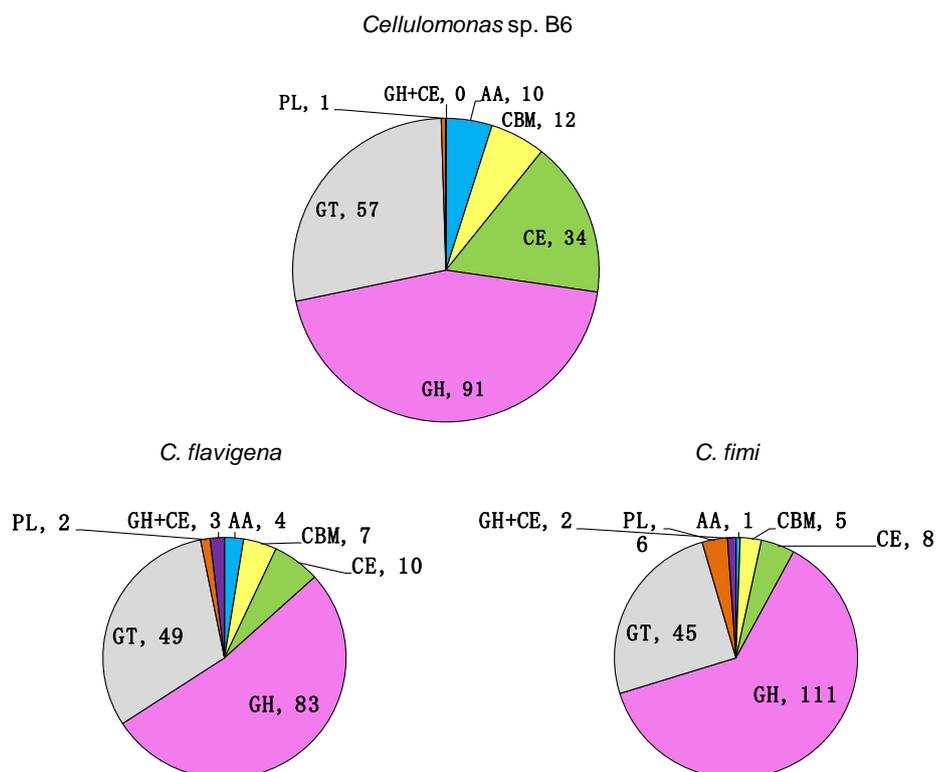
Esto permitió el armado del CAZoma teórico, es decir la identificación de todas las CAZimas codificadas en el genoma. A su vez, en base a los datos genómicos, realizamos el análisis por dbCAN de las cepas de referencia de *C. flavigena* (ATCC 482) y *C. fimi* (ATCC 484) y comparamos los CAZomas predichos para las tres especies. En el genoma de *Cellulomonas* sp. B6 se identificaron 205 CAZimas, mientras que se identificaron 158 en *C. flavigena* (ATCC 482) y 178 en *C. fimi*. En la figura R10, se muestra el número de CAZimas, clasificadas según su CD, independientemente de la presencia de CBMs asociados. También se mencionan aquéllas proteínas que presentan sólo CBMs (no asociados a ningún CD).

En las tres especies, la mayoría de las CAZimas son glicosil hidrolasas (GH), seguidas de transferasas (GT) y esterases (CE), sin embargo el número de esterases y enzimas auxiliares (AA) en *Cellulomonas* sp. B6 fue notablemente mayor que en las otras especies. Se identificaron 34 CE y 10 AA en B6, siendo 10 CE y 4 AA las descritas para *C. flavigena* y 8 CE y 1 AA para *C. fimi*. No obstante, estas especies presentan también

CAZimas con dominios GH y CE (que forman parte de la misma proteína), lo que no fue observado en B6.

Por otra parte, el número de PL fue mayor en *C. fimi* (6), mientras que se identificó una única PL en B6 y sólo dos en *C. flavigena*.

Todas las diferencias encontradas mediante este análisis aportan evidencias de que *Cellulomonas* sp. B6 podría ser una especie nueva, con características distintas de *C. flavigena* y *C. fimi*.



**Figura R10.** CAZimas de *Cellulomonas* sp. B6, *C. flavigena* ATCC 482 y *C. fimi* ATCC 484, clasificadas según su dominio catalítico. AA: enzima con actividad auxiliar, CBM: dominio de unión a carbohidratos (no asociado a ningún dominio catalítico), CE: esterasa de carbohidratos, GH: glicosil hidrolasa, GT: glicosil transferasa.

Analizamos también la posible localización, intracelular o extracelular, de las CAZimas. Para ello se utilizó el algoritmo SignalP v.4.1 para determinar la presencia de secuencia codificante para péptido señal (Petersen y col., 2011), es decir, para identificar aquellas proteínas que serían secretadas al medio extracelular por el mecanismo clásico. A su vez, se analizaron las proteínas también mediante el algoritmo SecretomeP v.2.0, que permite la predicción de proteínas extracelulares,

secretadas por mecanismos independientes de péptido señal. Se han descrito proteínas bacterianas típicamente encontradas en el medio extracelular que no portan péptido señal, pero cuyos mecanismos de secreción están todavía poco caracterizados. El algoritmo de SecretomeP fue "entrenado" con secuencias de proteínas de este tipo para la identificación de características compartidas en base a las cuales analizar nuevas secuencias. El puntaje para proteínas bacterianas secretadas por vías independientes de péptido señal resultó siempre mayor que 0,5, por lo cual se estableció este valor como umbral (Bendsten y col., 2005).

De las 205 CAZimas de *Cellulomonas* sp. B6, 48 se asignaron al extracto extracelular por identificación de un péptido señal en su secuencia, indistintamente de su puntaje en SecretomeP (Tabla R7). Cabe aclarar que, si bien la mayoría de las proteínas con péptido señal dan un puntaje que supera el umbral, puede ocurrir lo contrario (Bendsten y col., 2005).

**Tabla R7.** CAZimas predichas como extracelulares por vías del péptido señal en el genoma de *Cellulomonas* sp. B6.

Dominios CAZy	GenBank	Dominios CAZy	GenBank
AA10, CBM2	KSW19119.1	GH114	KSW20994.1
AA10, CBM2	KSW30009.1	GH13	KSW14264.1
CBM2	KSW29693.1	GH13 (2), CBM20, CBM48	KSW29886.1
CBM2, CBM63	KSW14103.1	GH15	KSW29453.1
CBM20	KSW20911.1	GH16, CBM13	KSW17651.1
CBM32	KSW30160.1	GH18, CBM2	KSW18655.1
CBM44 (2), CBM66	KSW23491.1	GH23, CBM50 (4)	KSW28633.1
CE1	KSW19118.1	GH27, CBM13	KSW30136.1
CE1	KSW19128.1	GH43, CBM6, CBM13	KSW30137.1
CE1	KSW21507.1	GH43, GH5	KSW21678.1
CE1 (2)	KSW15560.1	GH48, CBM2	KSW15916.1
CE1 (2)	KSW21470.1	GH5	KSW28923.1
CE1, CBM2	KSW29898.1	GH53, CBM61	KSW29720.1
CE10	KSW30030.1	GH6, CBM2	KSW15663.1
CE2, CBM2	KSW29963.1	GH6, CBM2	KSW30082.1
CE3, CBM13	KSW29916.1	GH6, CBM2	KSW30095.1
CE3, CBM2	KSW13732.1	GH62, GH5	KSW17752.1
GH10	KSW18330.1	GH74 (4), CBM2	KSW29726.1
GH10	KSW29966.1	GH9	KSW29417.1
GH10, CBM13	KSW21671.1	GH9, CBM2, CBM3	KSW29083.1
GH10, CBM2	KSW20567.1	GH9, CBM4 (2)	KSW29108.1

GH10, CBM9, CBM22	KSW23552.1	GH9, CBM4 (2)	KSW29539.1
GH10, CBM9, CBM22 (2)	KSW16349.1	GT26	KSW30250.1
GH11, CBM2	KSW19115.1	PL3, CBM13 (2)	KSW13730.1

En el genoma de B6, otras 26 proteínas fueron clasificadas como secretadas por vías independientes de péptido señal (Tabla R8).

**Tabla R8.** CAZimas potencialmente secretadas por vías independientes de péptido señal.

Dominios CAZy	GenBank	Dominios CAZy	GenBank
CBM2	KSW13740.1	GH5	KSW14495.1
CBM2	KSW14058.1	GT2	KSW21676.1
CBM2	KSW15603.1	GT2	KSW29278.1
CBM2	KSW29067.1	GT2	KSW29282.1
CBM32	KSW29929.1	GT2	KSW29559.1
CE1	KSW14111.1	GT2 (2)	KSW20676.1
CE1	KSW17191.1	GT51	KSW15047.1
CE1	KSW28924.1	GT51	KSW29228.1
CE7	KSW29746.1	GT76	KSW19706.1
GH3	KSW21816.1	GT83	KSW15645.1
GH3	KSW21825.1	GT83	KSW15646.1
GH43	KSW21776.1	GT83	KSW23495.1
GH43	KSW29954.1	GT83	KSW23494.1

Una CAZima cuya secuencia está incompleta (KSW14744.1) obtuvo un valor alto en SecretomeP, por lo cual se la podría considerar, en principio, como extracelular, sin clasificarla según el mecanismo de secreción. De esta manera, 75 de las 205 CAZimas fueron asignadas en forma teórica al extracto extracelular. Las CAZimas restantes se clasificaron como intracelulares por ausencia de péptido señal y puntaje bajo de SecretomeP (Tabla R9).

**Tabla R9.** CAZimas predichas como intracelulares.

Dominios CAZy	GenBank	Dominios CAZy	GenBank
AA3	KSW18930.1	GH31	KSW21932.1
AA6	KSW21003.1	GH32	KSW30258.1
AA6	KSW23546.1	GH35	KSW29984.1

AA6	KSW28595.1	GH36	KSW13763.1
AA7	KSW15162.1	GH36	KSW30131.1
AA7	KSW29399.1	GH39	KSW28373.1
AA7	KSW29433.1	GH39	KSW30127.1
AA7	KSW29579.1	GH42	KSW21933.1
CBM50	KSW13747.1	GH42	KSW28374.1
CBM50	KSW29143.1	GH42	KSW29716.1
CE1	KSW16030.1	GH42	KSW29985.1
CE1	KSW19100.1	GH43	KSW21730.1
CE1	KSW21929.1	GH43	KSW21922.1
CE1	KSW29323.1	GH43	KSW29213.1
CE1	KSW29333.1	GH51	KSW14247.1
CE1	KSW29578.1	GH51	KSW29942.1
CE10	KSW15555.1	GH67	KSW29842.1
CE10	KSW20566.1	GH77	KSW30099.1
CE10	KSW29139.1	GH92	KSW17036.1
CE10	KSW30256.1	GH94	KSW12785.1
CE14	KSW15169.1	GH95	KSW28370.1
CE14	KSW21668.1	GT2	KSW15285.1
CE14	KSW29507.1	GT2	KSW15649.1
CE14	KSW29597.1	GT2	KSW19705.1
CE3	KSW13997.1	GT2	KSW21818.1
CE3	KSW29492.1	GT2	KSW23500.1
CE4	KSW15562.1	GT2	KSW23501.1
CE7	KSW15158.1	GT2	KSW28640.1
CE7	KSW30302.1	GT2	KSW29253.1
CE9	KSW21534.1	GT2	KSW29267.1
GH1	KSW21931.1	GT2	KSW29269.1
GH1	KSW29322.1	GT2	KSW29271.1
GH1	KSW29522.1	GT2	KSW29273.1
GH10	KSW28941.1	GT2	KSW29290.1
GH109	KSW21834.1	GT2	KSW29459.1
GH109	KSW29517.1	GT2	KSW29524.1
GH109	KSW29527.1	GT2	KSW30069.1
GH120	KSW30135.1	GT2, GT27	KSW29255.1
GH127	KSW21674.1	GT2, GT4	KSW29257.1
GH127	KSW30132.1	GT20	KSW29439.1
GH13	KSW13755.1	GT28	KSW29161.1
GH13	KSW15554.1	GT35	KSW14109.1
GH13	KSW15644.1	GT39	KSW20493.1

GH13	KSW18932.1	GT4	KSW15293.1
GH13	KSW20193.1	GT4	KSW16006.1
GH13	KSW28288.1	GT4	KSW16276.1
GH13	KSW29669.1	GT4	KSW19098.1
GH13	KSW29939.1	GT4	KSW19099.1
GH13, CBM48	KSW15648.1	GT4	KSW21458.1
GH13, CBM48	KSW18933.1	GT4	KSW23499.1
GH13, CBM48 (2)	KSW13762.1	GT4	KSW23717.1
GH130	KSW19121.1	GT4	KSW28894.1
GH15	KSW18319.1	GT4	KSW28895.1
GH15	KSW21900.1	GT4	KSW28896.1
GH2	KSW13761.1	GT4	KSW29272.1
GH2	KSW13987.1	GT4	KSW29287.1
GH20	KSW20490.1	GT4	KSW29289.1
GH27 (2)	KSW21775.1	GT4	KSW29368.1
GH3	KSW19794.1	GT4	KSW30066.1
GH3	KSW21537.1	GT4, GT94	KSW29260.1
GH3	KSW21826.1	GT5	KSW21039.1
GH3	KSW28963.1	GT81	KSW30306.1
GH3	KSW29220.1	GT94	KSW23485.1
GH3	KSW29840.1	GT94	KSW28893.1
GH31	KSW19264.1	GT94	KSW28898.1

De las 57 GT totales, la mayoría (44) fue asignada al extracto intracelular, lo que es esperable, dado que las GT catalizan reacciones de síntesis como la formación de enlaces glucosídicos en los oligo y polisacáridos y la N y O-glicosilación de proteínas (Gloster, 2014).

En cuanto a las CE (34), que remueven residuos en unión éster presentes en mono, oligo y polisacáridos (Cantarel y col., 2008), 14 fueron predichas como extracelulares y podrían estar involucradas en la desacetilación de sustratos; en particular, las CE1 actúan sobre xilano acetilado.

Las enzimas de actividad auxiliar (AA) se adjudicaron en su mayoría al extracto intracelular, con la excepción 2 AA10 con dominios CBM2. Las AA intracelulares presentan dominios AA3, AA6 y AA7; que han sido identificados en enzimas con actividad oxidasa, reductasa y deshidrogenasa, algunas de las cuales estarían involucradas en la degradación y detoxificación de derivados de lignina y compuestos

aromáticos (Levasseur y col., 2013). Las AA extracelulares, en cambio, son de la familia AA10, previamente denominadas CBM33 y equivalentes a las AA9 fúngicas (ex GH61). Estas enzimas han sido relacionadas a la degradación de polisacáridos por mecanismos oxidativos y son llamadas, genéricamente, LPMOs (monooxigenasas líticas de polisacáridos) (Corrêa y col., 2016).

De las 91 GH, 85 se identificaron como secretadas, ya sea por vías clásicas (80) o independientes de péptido señal (5). Esto está relacionado con la degradación de polisacáridos. En cuanto a las intracelulares, estarían involucradas en la degradación posterior de oligosacáridos cortos en el interior celular. En forma similar a lo observado para el resto de las CAZimas de B6, la mayoría de las GH predichas como extracelulares presentaron al menos un dominio CBM.

Adicionalmente, se detectó una única PL, de la familia PL3, potencialmente secretada. Las PL están asociadas a la degradación de polisacáridos pectínicos (Saini y col., 2015).

El análisis del CAZoma teórico indica que *Cellulomonas* sp. B6 codifica para todo el repertorio de enzimas necesarias para la deconstrucción y utilización de polisacáridos estructurales presentes en la biomasa vegetal. Por lo tanto, una vez construido el CAZoma teórico completo de *Cellulomonas* sp. B6, se analizó en más detalle aquellas enzimas potencialmente involucradas en la degradación de polisacáridos estructurales (Tabla R10).

Si bien la selección de las secuencias de interés se basó, en una primera instancia, en los dominios CAZy identificados, se hizo luego una depuración manual de la anotación para asignar con mayor certeza las potenciales actividades enzimáticas. La depuración consistió en la búsqueda de homologías por alta identidad de secuencia con proteínas de la base de datos no redundante del NCBI usando el algoritmo Blastp (Atschul y col., 1990).

**Tabla R10.** Secuencias codificantes para enzimas con potencial actividad de degradación de (hemi)celulosa, quitina y pectina de *Cellulomonas* sp. B6. Los dominios CAZy fueron identificados por anotación mediante dbCAN. Los dominios múltiples se indican entre paréntesis. \*Secuencias parciales.

Actividad predicha	Dominios CAZy	GenBank	Péptido señal	Mecanismo de secreción predicho
<b><math>\alpha</math>-glucuronidasa (EC 3.2.1.139)</b>	GH67	KSW29842.1	No	-
<b><math>\alpha</math>-L-arabinofuranosidasa (EC 3.2.1.55)</b>	GH51	KSW14247.1	No	-
	GH51	KSW29942.1	No	-
	GH62, GH5	KSW17752.1	Sí	Clásico
<b>Endo-1,5-<math>\alpha</math>-L-arabinosidasa (EC 3.2.1.99)</b>	GH43	KSW21776.1	No	-
	GH43	KSW21922.1	No	-
<b><math>\beta</math>-glucosidasa (EC 3.2.1.21)</b>	GH1	KSW21931.1	No	-
	GH1	KSW29322.1	No	-
	GH1	KSW29522.1	No	-
	GH3	KSW19794.1	No	-
	GH3	KSW21537.1	No	-
	GH3	KSW21816.1	No	No clásico
	GH3	KSW21825.1	No	No clásico
	GH3	KSW21826.1	No	-
	GH3	KSW28963.1	No	-
	GH3	KSW29220.1	No	-
<b><math>\beta</math>-manosidasa (EC 3.2.1.25)</b>	GH2	KSW13761.1	No	-
	GH5	KSW28923.1*	Sí	Clásico
<b><math>\beta</math>-xilosidasa (EC 3.2.1.37)</b>	GH39	KSW30127.1	No	-
	GH39	KSW28373.1	No	-
<b><math>\beta</math>-xilosidasa (EC 3.2.1.37)/ <math>\alpha</math>-L-arabinofuranosidasa (EC 3.2.1.55)</b>	GH43	KSW29213.1	No	-
	GH43	KSW21730.1	No	-
	GH43	KSW29954.1	No	No clásico
	GH43, CBM6, CBM13	KSW30137.1	Sí	Clásico
	GH43, GH5	KSW21678.1	Sí	Clásico
<b>Quitinasa (EC 3.2.1.14)</b>	GH18, CBM2	KSW18655.1	Sí	Clásico
<b>Endo-<math>\beta</math>-1,4-glucanasa (EC 3.2.1.4)</b>	GH5	KSW14495.1	No	-
	GH6, CBM2	KSW30082.1	Sí	Clásico
	GH6, CBM2	KSW30095.1	Sí	Clásico
	GH9	KSW29417.1*	Sí	Clásico
	GH9, CBM2, CBM3	KSW29083.1	Sí	Clásico
	GH9, CBM4 (2)	KSW29108.1	Sí	Clásico
	GH9, CBM4 (2)	KSW29539.1	Sí	Clásico

<b><math>\beta</math>-1,4-xilanasas (EC 3.2.1.8)</b>	GH10, CBM22	KSW14744.1*	Secuencia incompleta	Secuencia incompleta
	GH10	KSW28941.1	No	-
	GH10	KSW18330.1	Sí	Clásico
	GH10	KSW29966.1*	Sí	Clásico
	GH10, CBM13	KSW21671.1	Sí	Clásico
	GH10, CBM2	KSW20567.1	Sí	Clásico
	GH10, CBM9, CBM22	KSW23552.1	Sí	Clásico
	GH10, CBM9, CBM22 (2)	KSW16349.1	Sí	Clásico
	GH11, CBM2	KSW19115.1	Sí	Clásico
<b>Exo-<math>\beta</math>-1,4-glucanasa (EC 3.2.1.176)</b>	GH48, CBM2	KSW15916.1	Sí	Clásico
<b>Exo-<math>\beta</math>-1,4-glucanasa (EC 3.2.1.91)</b>	GH6, CBM2	KSW15663.1	Sí	Clásico
<b>Xiloglucanasa (EC 3.2.1.151)</b>	GH74 (4), CBM2	KSW29726.1	Sí	Clásico
<b>Monooxidasa de polisacáridos (LPMO)</b>	AA10, CBM2	KSW19119.1	Sí	Clásico
	AA10, CBM2	KSW30009.1	Sí	Clásico
<b>Pectato liasa (EC 4.2.2.2)</b>	PL3, CBM13 (2)	KSW13730.1	Sí	Clásico
<b><math>\alpha</math>-galactosidasa (EC 3.2.1.22)</b>	GH27, CBM13	KSW30136.1	Sí	Clásico
	GH27	KSW21775.1	No	-

En el genoma de *Cellulomonas* sp. B6 se encuentran codificadas 9 endo- $\beta$ -1,4-xilanasas (EC 3.2.1.8), 8 de las cuales serían potencialmente secretadas (7 GH10 y una GH11). Las 8 GH10 difieren en los módulos que presentan. Dos de las enzimas están formadas sólo por el módulo catalítico GH10, mientras que el resto tiene una estructura multimodular, con variaciones en el tipo y la cantidad de módulos de unión a carbohidratos. Los CBMs identificados en las xilanasas (GH10 y GH11) de B6 han sido reportados como de unión a xilano, pero también a quitina (CBM2), celulosa (CBM2 y CBM9), glucanos mixtos (CBM22) y galactanos o arabinosa (CBM13) (cazy.org) (Fujimoto, 2013). No hay hasta el momento información sobre afinidad de los diferentes CBMs por un tipo particular de xilano, pero esta variación podría permitirle a *Cellulomonas* sp. la utilización de diversos xilanos.

Se identificaron, asimismo, GHs que podrían estar involucradas en la desramificación de xilanos, como una GH62 ( $\alpha$ -L-arabinofuranosidasa). Además, se identificaron 5 GH43, que por su estructura podrían ser  $\alpha$ -L-arabinofuranosidasas,  $\beta$ -

xilosidasas. o enzimas bifuncionales (Shallom y Shoham, 2003). De estas, dos serían secretadas en forma dependiente de péptido señal. La función de las  $\alpha$ -L-arabinofuranosidasas es la desramificación del xilano y de las  $\beta$ -xilosidasas es la degradación de xilooligosacáridos cortos a monómeros.

Además, identificamos dos  $\beta$ -xilosidasas GH39 intracelulares, que podrían estar involucradas en la utilización final de los xilooligosacáridos cortos a xilosa, para poder ser metabolizada. Cabe destacar que, en nuestros ensayos, la actividad  $\beta$ -xilosidasa fue detectada en el extracto intracelular y en el sobrenadante de cultivo libre de células, por lo que estos resultados concuerdan con lo observado experimentalmente.

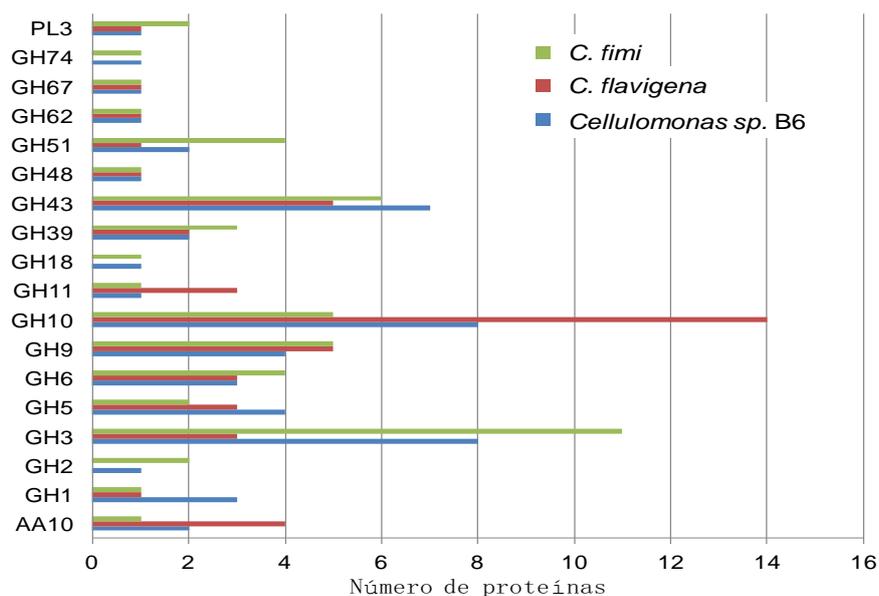
Mediante el análisis por el algoritmo dbCAN encontramos también dos proteínas extracelulares con más de un módulo catalítico: GH62-GH5 (KSW17752.1) y GH43-GH5 (KSW21678.1). Sin embargo, el análisis de módulos conservados de NCBI clasifica el módulo GH5 como un dominio de trébol-beta tipo ricina (cd00161), relacionado con la unión a carbohidratos. Hasta el momento, no hay información de proteínas homólogas caracterizadas, de manera que la función de este módulo es todavía incierta.

Cabe destacar que se identificaron además, GHs extracelulares potencialmente activas sobre otros tipos de polisacáridos, como una quitinasa (EC 3.2.1.14) GH18, una xiloglucanasa (EC 3.2.1.151) con cuatro dominios GH74 y una pectato liasa PL3 (KSW13730.1).

Estos resultados indican que *Cellulomonas* sp. B6 podría tener la capacidad de hidrolizar una variedad de polisacáridos estructurales, como glucanos (celulosa, xiloglucanos), xilanos (como arabino y glucuronoxilano) y pectina. También podría tener actividad sobre polisacáridos de reserva, como el almidón, ya que codifica para 4 potenciales amilasas extracelulares GH13 y GH15.

En base a la actividad xilanasas y el número alto de xilanasas codificadas, se analizó en mayor detalle la vía de utilización de xilósidos. El análisis combinado de la anotación por PGAP y RAST permitió la identificación de todos los componentes del sistema. De esta manera, se identificaron potenciales xilanasas, reguladores de la familia ROK de respuesta a xilosa, esterasas,  $\alpha$  y  $\beta$ - xilosidasas, xilulosa quinasas, una xilosa isomerasa, una permeasa y una proteína de unión a sustrato del transportador ABC de xilosa (KSW29515.1 y KSW28372.1, respectivamente) y una permeasa y una proteína de





**Figura R12.** Comparación de GH, AA y PL con actividad sobre celulosa y hemicelulosa identificadas en *Cellulomonas sp. B6*, *C. fimi* y *C. flavigena*. Se identificó un dominio GH10 y otro GH11 en la misma proteína de *C. flavigena*, por lo que fue tenida en cuenta en el número total de proteínas con dominios GH10 y GH11 codificadas por la especie. Lo mismo ocurrió con la proteína con un dominio GH62, que tiene un dominio GH10 asociado. La proteína que contiene un módulo GH62 y una de las GH43 presentan un dominio GH5 asociado. Estos dominios se tuvieron en cuenta para el cálculo del total de proteínas con dominios GH5.

En todos los casos, se identificó una sola proteína con un dominio GH48 y otra con un dominio GH67. En *Cellulomonas sp. B6* y *C. fimi* se identificaron proteínas con dominios GH2, GH18 y GH74, aunque la GH74 de *C. fimi* presentó un único dominio catalítico. No se encontraron GH18 ni GH74 en el CAZoma de *C. flavigena*.

Cabe destacar que, si bien en todos los casos se identificaron AA10, en *Cellulomonas sp. B6* también se encontraron proteínas auxiliares con dominios de AA3, AA6 y AA7 (Tablas R7 y R9) mientras que sólo identificaron dominios AA10 en los genomas de *C. fimi* y *C. flavigena*.

## 2.4. Análisis del secretoma experimental de *Cellulomonas sp B6*

A fin de dilucidar cuáles de las CAZimas son efectivamente secretadas y que, por lo tanto, podrían ser las responsables de las actividades observadas en los sobrenadantes de cultivo de *Cellulomonas sp. B6*, se procedió a la identificación de las mismas por

espectrometría de masas a partir de sobrenadantes completos obtenidos de cultivos crecidos en tres sustratos: CMC, RAC y PTE. CMC se eligió por fuente de carbono exclusivamente celulósica. RAC y PTE fueron elegidas como fuentes de biomasa lignocelulósicas de distinto origen y composición. A su vez, se utilizó un cultivo en sacarosa como control negativo, ya que es un disacárido no reductor que puede ser utilizado por *Cellulomonas* como fuente de carbono simple. Al ser no reductor, no interfiere con los ensayos de actividad en el sobrenadante.

A las 72 horas, se determinó el crecimiento bacteriano como UFC/ml, la cantidad de proteínas secretadas y las actividades xilanas y CMCase de los sobrenadantes de cultivo (Tabla R11).

**Tabla R11.** Proteínas totales y actividades CMCase y xilanas de los sobrenadantes de cultivo de *Cellulomonas* sp. B6 en CMC, RAC, PTE y sacarosa para ser analizados por espectrometría de masas. Se indica también el crecimiento como  $\text{Log}_{10}(\text{UFC/ml})$  al momento de la cosecha.

Sustrato de cultivo	CMC	RAC	PTE	Sacarosa
Proteínas totales (mg/ml)	0,11 ± 0,04	0,79 ± 0,10	0,57 ± 0,08	0,03 ± 0,00
Actividad CMCase (UI/ml)	0,15 ± 0,01	0,20 ± 0,01	0,19 ± 0,01	≤ 0,03
Actividad xilanas (UI/ml)	0,13 ± 0,00	0,59 ± 0,02	0,57 ± 0,03	≤ 0,03
$\text{Log}_{10}(\text{UFC/ml})$	9,77	9,82	10,27	8,74

El crecimiento fue similar en todos los sustratos, salvo en sacarosa que fue menor. Sin embargo, el cultivo en biomasa generó mayor secreción de proteínas. A su vez las actividades CMCase y xilanas fueron mayores en los extractos de cultivos crecidos sobre biomasa, respecto del sobrenadante de cultivo en CMC. El control en sacarosa resultó en una concentración de proteínas secretadas similar a la de CMC pero con resultados negativos para la actividad enzimática.

Para cada condición, las proteínas totales de los dos cultivos independientes fueron precipitadas y sometidas a espectrometría de masas, para su identificación.

Se compararon las proteínas detectadas en los sobrenadantes de cultivo con las CAZimas potencialmente secretadas codificadas en el genoma (CAZoma secretado teórico). Para el análisis, se tuvieron en cuenta todas las proteínas con dominios CAZY identificadas por, al menos, un péptido de alta confianza en una de las réplicas (Tabla

R12). Aquellas proteínas perviamente clasificadas como intracelulares (por SignalP y SecretomeP) no fueron consideradas en el análisis ya que su detección puede deberse a lisis celular causada por los procedimientos experimentales.

**Tabla R12.** CAZimas y proteínas extracelulares involucradas en el transporte de azúcares identificadas por espectrometría de masas en sobrenadantes de cultivo de *Cellulomonas* sp. B6 en CMC, RAC y PTE.

Actividad predicha	GenBank	Dominios CAZy	Sustrato de cultivo	PM[kDa]/pI
<b>Endo-1,4-<math>\beta</math>-xilanasasa (EC 3.2.1.8)</b>	KSW14744.1*	GH10, CBM22	RAC	54,8/ 4,75
	KSW29966.1*	GH10	CMC, RAC, PTE	37,8/8,51
	KSW16349.1	GH10, CBM9, CBM22 (2)	RAC, PTE	130,7/ 4,42
	KSW18330.1	GH10	RAC, PTE	65,4/5,27
	KSW19115.1	GH11, CBM2	RAC, PTE	36,5/9,36
	KSW20567.1	GH10, CBM2	RAC, PTE	50,6/8,76
	KSW23552.1	GH10, CBM9, CBM22	CMC, RAC, PTE	112,3/4,82
	KSW21671.1	GH10, CBM13	CMC, RAC	53,2/8,70
<b><math>\beta</math>-xilosidasa (EC 3.2.1.37)/ <math>\alpha</math>-L-arabinofuranosidasa (EC 3.2.1.55)</b>	KSW21678.1	GH43, GH5	RAC	76,0/9,07
<b>Endo-1,4-<math>\beta</math>-glucanasa (EC 3.2.1.4)</b>	KSW29417.1*	GH9	RAC	65,8/6,48
	KSW29083.1	GH9, CBM2, CBM3	RAC	86,4/5,68
	KSW29108.1	GH9, CBM4 (2)	RAC, PTE	114,1/4,58
	KSW29539.1	GH9, CBM4 (2)	RAC, PTE	118,8/4,58
	KSW30082.1	GH6, CBM2	PTE	47,3/8,41
<b>Exo-1,4-<math>\beta</math>-glucanasa (EC 3.2.1.91)</b>	KSW15663.1	GH6, CBM2	CMC, RAC, PTE	63,4/ 6,47
<b>Exo-1,4-<math>\beta</math>-glucanasa (EC 3.2.1.176)</b>	KSW15916.1	GH48, CBM2	CMC, RAC, PTE	90,6/ 5,33
<b><math>\beta</math>-glucosidasa (EC 3.2.1.21)</b>	KSW21825.1	GH3	RAC, PTE	83,5/4,48
<b>Xiloglucanasa (EC 3.2.1.151)</b>	KSW29726.1	GH74 (4), CBM2	RAC, PTE	95,7/5,00
<b><math>\beta</math>-1,6-glucosidasa (EC 3.2.1.20)</b>	KSW29886.1	GH13 (2), CBM20, CBM48	RAC, PTE	207,1/4,75
<b><math>\alpha</math>-galactosidasa (EC 3.2.1.22)</b>	KSW30136.1	GH27, CBM13	RAC	60,6/6,79
<b><math>\alpha</math>-L-arabinofuranosidasa (EC 3.2.1.55)</b>	KSW17752.1	GH62, GH5	RAC, PTE	53,6/8,29
<b>Quitinasa (EC 3.2.1.14)</b>	KSW18655.1	GH18, CBM2	RAC, PTE	56,3/5,77
<b>Proteína de union a carbohidratos</b>	KSW23491.1	CBM44 (2), CBM66	RAC, PTE	144,8/5,22
	KSW30160.1	CBM32	RAC, PTE	102,2/4,54
	KSW29929.1	CBM32	PTE	61,7/4,63

	KSW29693.1	CBM2	RAC, PTE	76,7/4,84
<b>LPMO</b>	KSW19119.1	AA10, CBM2	RAC	37,0/5,92
<b>Esterasa</b>	KSW21470.1	CE1 (2)	RAC, PTE	93,6/4,75
	KSW21507.1	CE1	RAC, PTE	30,2/4,4
	KSW15560.1	CE1 (2)	RAC, PTE	98,0/ 4,67
<b>Pectato liasa (EC 4.2.2.2)</b>	KSW13730.1	PL3, CBM13 (2)	RAC	30,2/4,4
<b>Proteína de unión a sustrato del transportador de azúcares ABC</b>	KSW19791.1	-	PTE	45,1/4,65
	KSW21774.1	-	RAC, PTE	48,6/4,2
	KSW21936.1	-	RAC, PTE	59,4/4,64
	KSW28372.1	-	RAC, PTE	61,0/4,61
	KSW28971.1	-	RAC, PTE	48,2/4,26
	KSW29516.1	-	RAC, PTE	60,6/4,32
	KSW29892.1	-	RAC, PTE	39,9/4,40
	KSW30268.1	-	RAC, PTE	48,6/4,46
	KSW19409.1	-	CMC	46,7/4,33

En los sobrenadantes de cultivo en sacarosa, se identificaron 31 proteínas con péptido señal, de las cuales tres presentaban dominios CAZy: una esterasa CE1, una proteína con un CBM32 (reportado para unión a quitina) y una GH114 que podría tener actividad endogalactosaminadasa (EC 3.2.1.109). Sólo una GH114 fue caracterizada hasta el momento, en la bacteria *Streptomyces griseus* (Bredamm, 1979).

Teniendo en cuenta todos los extractos extracelulares estudiados, se identificó un total de 29 de las 48 proteínas con dominios CAZy con péptido señal que habían sido predichas *in silico* (Tabla R7). Las exoglucanasas de las familias GH6 (KSW15663.1) y GH48 (KSW15916.1), así como dos xilanasas con dominios GH10 (KSW23552.1 y KSW29966.1), fueron detectadas en los sobrenadantes obtenidos por crecimiento en todos los sustratos (CMC, RAC y PTE), lo que sugiere que tienen un rol importante para la utilización de polisacáridos en *Cellulomonas* sp. B6 (Figura R13).



En PTE, se identificaron, además, una endoglucanasa GH6 y una proteína de unión a carbohidratos.

Todas las endoxilasas predichas como extracelulares codificadas en el genoma fueron efectivamente identificadas en los sobrenadantes de cultivo en biomasa, lo que explicaría la actividad xilanolítica observada en los sobrenadantes de RAC y PTE.

En estos extractos, también se encontraron 5 de las 7 endoglucanasas predichas, incluyendo todas las GH9 y una GH6 (Tabla R10). Se encontraron, también, enzimas de las familias GH62, GH18 y GH74 y 3 proteínas con dominios CE1, clasificadas como esterases putativas (KSW15560.1, KSW21470.1 y KSW21507.1). Además, se detectaron 4 proteínas con dominios CBM sin módulo catalítico asociado, ni actividad predicha (KSW23491.1, KSW29693.1, KSW30160.1 y KSW29929.1) y una posible  $\beta$ -galactosidasa de GH13 (KSW29886.1). En RAC, también se detectó una  $\alpha$ -galactosidasa potencial de la familia GH27, actividad requerida en la degradación de algunas hemicelulosas y pectinas.

Por estimación de la abundancia relativa, que fue calculada para todas las proteínas (Tabla R13, Anexo 1) la GH10 (KSW23552.1) identificada en todos los sobrenadantes de cultivo fue, asimismo, la más abundante de las CAZimas, lo que refuerza su importancia potencial en la degradación de polisacáridos.

**Tabla R13.** Abundancia relativa de las proteínas secretadas de *Cellulomonas* sp. B6. Sólo se muestran las proteínas relevantes con péptido señal o identificadas como extracelulares por SecretomeP. El índice de abundancia emPAI se calculó con Sequest y la ecuación  $\text{emPAI}/\Sigma(\text{emPAI}) \times 100$  se usó para calcular el contenido de cada proteína en mol % (emPAI%). Las diferencias observadas en los valores de emPAI% de cada muestra reflejan las diferencias en el número de proteínas totales identificadas en cada condición.

Descripción (acceso de GenBank)	PTE 1		PTE 2		RAC 1		RAC 2		CMC 1		CMC2	
	emPAI	emPAI%										
PL3, CBM13 (2) (KSW13730.1)	-	-	-	-	0,10	0,04	0,10	0,03	-	-	-	-
GH10, CBM22 (KSW14744.1)	-	-	-	-	0,12	0,06	0,12	0,04	-	-	-	-
CE1 (2) (KSW15560.1)	-	-	-	-	0,06	0,03	-	-	-	-	-	-
GH6, CBM2 (KSW15663.1)	0,43	0,22	0,43	0,22	0,19	0,09	0,43	0,13	4,38	20,64	2,16	16,21
GH48, CBM2 (KSW15916.1)	0,28	0,40	-	-	0,28	0,13	0,28	0,09	0,93	4,39	0,51	3,82

GH10, CBM9, CBM22 (2) (KSW16349.1)	0,84	0,43	0,84	0,43	0,84	0,39	0,84	0,26	-	-	-	-
GH62, GH5 (KSW17752.1)	0,11	0,05	0,11	0,05	-	-	0,49	0,15	-	-	-	-
GH10 (KSW18330.1)	0,71	0,36	0,71	0,36	1,33	0,61	1,33	0,40	-	-	-	-
GH18, CBM2 (KSW18655.1)					0,33	0,15	0,47	0,14	-	-	-	-
GH11, CBM2 (KSW19115.1)	0,47	0,24	0,47	0,24	0,78	0,36	-	-	-	-	-	-
AA10, CBM2 (KSW19119.1)	-	-	-	-	-	-	0,29	0,09	-	-	-	-
dominio de unión a sustrato de los transportadores de azúcares ABC (KSW19409.1)	-	-	-	-	-	-	-	-	7,73	36,43	5,66	42,45
dominio de unión a sustrato de los transportadores ABC (KSW19409.1)	9,00	1,27			10,45	4,83	9,00	2,74	7,73	36,43	5,66	42,45
dominio de unión a sustrato de los transportadores de azúcares ABC (KSW19791.1)	10,10	5,16	10,10	5,16	4,93	2,28	17,74	5,40	-	-	-	-
GH10, CBM2 (KSW20567.1)	0,26	0,04	-	-	0,41	0,19	0,78	0,24	-	-	-	-
CE1 (2) (KSW21470.1)	1,29	0,66	1,29	0,66	0,91	0,42			-	-	-	-
CE1 (KSW21507.1)	0,39	0,20	0,39	0,20	0,39	0,18	0,39	0,12	-	-	-	-
GH10, CBM13 (KSW21671.1)	-	-	-	-	0,09	0,04	0,19	0,06	0,09	0,42	-	-
GH43, GH5 (KSW21678.1)	-	-	-	-	0,06	0,03	0,11	0,03	-	-	-	-
CBM44 (2), CBM66 (KSW23491.1)	0,64	0,33	0,64	0,33	0,64	0,30	-	-	-	-	-	-
GH10, CBM9, CBM22 (KSW23552.1)	1,85	0,94	1,85	0,94	2,00	0,92	2,90	0,88	2,33	10,99	1,08	8,11
GH9, CBM2, CBM3 (KSW29083.1)	-	-	-	-	0,06	0,03			-	-	-	-
GH9, CBM4 (2) (KSW29108.1)	1,85	0,94	1,85	0,94	2,51	1,16	2,28	0,69	-	-	-	-
GH9 (KSW29417.1)	-	-	-	-	-	-	0,07	0,02	-	-	-	-
GH9, CBM4 (2) (KSW29539.1)	0,34	0,18	0,34	0,18	0,60	0,28	1,15	0,35	-	-	-	-
dominio de unión a sustrato de los transportadores ABC (KSW29589.1)	-	-	-	-	-	-	-	-	1,64	7,71	1,64	12,28
CBM2 (KSW29693.1)	0,07	0,04	0,07	0,04	-	-	0,15	0,05	-	-	-	-
GH74 (4), CBM2 (KSW29726.1)	0,06	0,03	0,06	0,03	0,25	0,11	0,12	0,04	-	-	-	-
dominio de unión a	4,88	2,49	4,88	2,49	-	-	3,13	0,95	-	-	-	-

sustrato de los transportadores de azúcares ABC (KSW29892.1)												
GH10 (KSW29966.1)	0,78	0,40	-	-	-	-	0,78	0,40	0,12	0,06	0,41	3,10
GH6, CBM2 (KSW30082.1)	0,16	0,08	0,16	0,08	0,08	0,04	3,13	0,95	-	-	-	-
GH27, CBM13 (KSW30136.1)	-	-	-	-	-	-	-	-	-	-	-	-
CBM32 (KSW30160.1)	0,11	0,05	0,11	0,05	0,11	0,05	0,11	0,03	-	-	-	-

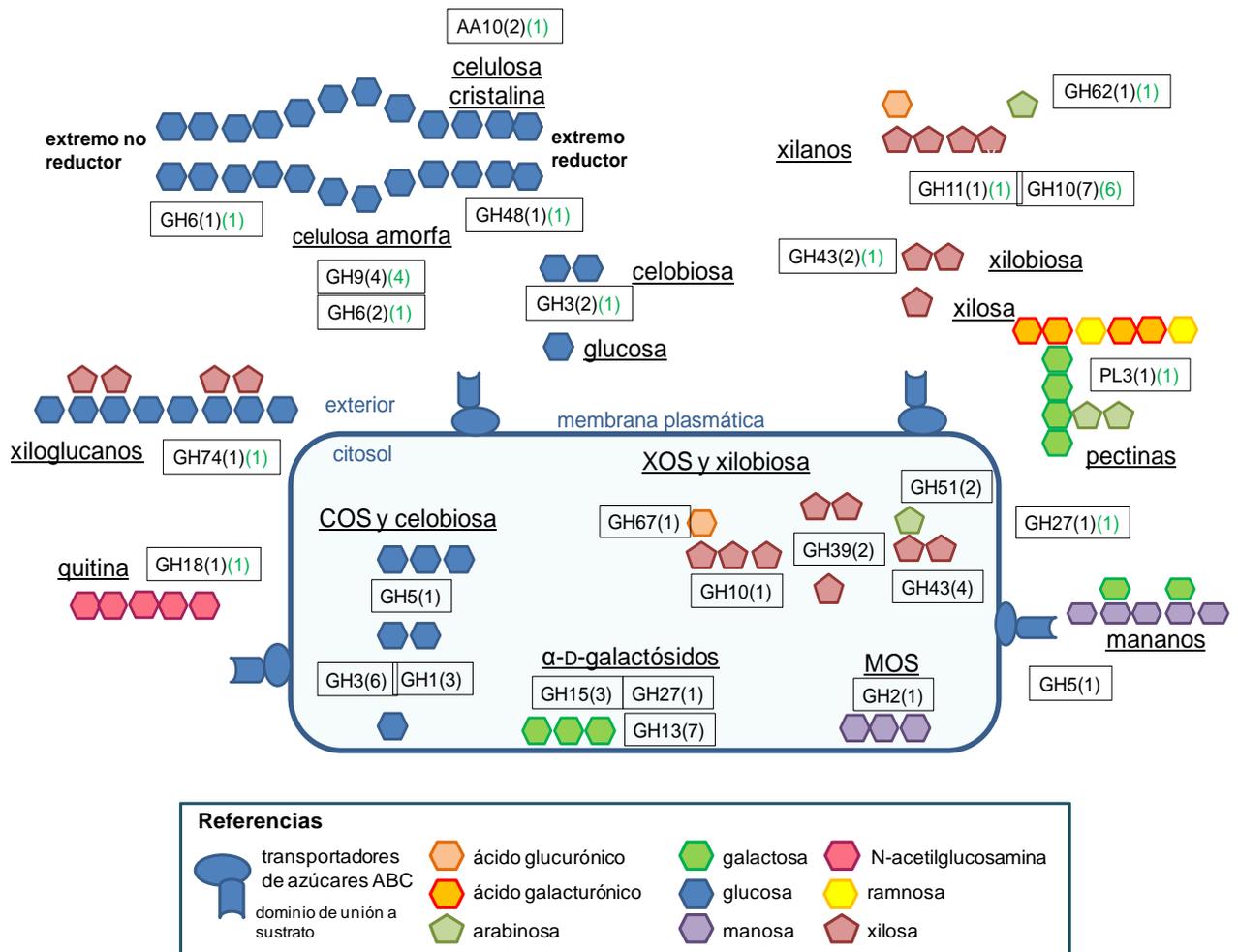
Otras proteínas que se detectaron solamente en los sobrenadantes de cultivo en RAC fueron una de las dos AA10 predichas (KSW19119.1) (Tabla R10), una  $\alpha$ -galactosidasa potencial de la familia GH27 (KSW30136.1), una GH43 (KSW21678.1) y la única PL3 codificada en el genoma (KSW13730.1). En el sobrenadante de cultivo en PTE, se identificó una  $\alpha$ -1,6-glucosidasa putativa de la familia GH13.

Cabe destacar que, en todos los sobrenadantes de cultivo, se identificaron proteínas de unión a sustrato de los transportadores de azúcares de la familia ABC con un alto índice de abundancia (Tablas R12 y R13, Anexo), que podrían estar involucradas en la captación de oligo y monosacáridos derivados de la hidrólisis (Anderson y col., 2012). El rol y la regulación de la expresión de los transportadores de carbohidratos no ha sido estudiado aún en profundidad.

Se detectaron, además, dos CAZimas posiblemente secretadas por vías independientes de péptido señal: una  $\beta$ -glucosidasa potencial de la familia GH3, presente en los sobrenadantes de cultivo de RAC y PTE, y una proteína con un dominio CBM32, identificada en el sobrenadante de cultivo en PTE.

En resumen, el cultivo en biomasa de *Cellulomonas* sp. B6 induce la secreción del repertorio completo de xilanasas y de un gran número de glucanasas. Si bien la abundancia relativa de enzimas no puede relacionarse directamente con la actividad, las proteínas secretadas identificadas se correlacionan con los resultados de actividad enzimática observados en los sobrenadantes de cultivos.

En base a la predicción de CAZimas con actividad (hemi)celulolítica y a los resultados de proteómica, se propone el siguiente modelo para *Cellulomonas* sp. B6 de utilización de polisacáridos estructurales (Figura R14).



**Figura R14.** Modelo de degradación enzimática de celulosa, hemicelulosas, quitina y pectinas de *Cellulomonas* sp. B6. El número de secuencias predichas para proteínas se indica entre paréntesis en negro. En el caso de que se hayan identificado en el secretoma, se aclara el número de proteínas detectadas al lado en letra verde.

**Parte 3. Expresión recombinante de una xilanasa GH10 de *Cellulomonas* sp. B6 y su aplicación a la deconstrucción de biomasa lignocelulósica**

### 3.1. Selección de una xilanasas de interés y expresión recombinante

La actividad xilanólítica de *Cellulomonas* sp. B6 se correlaciona con el hecho de que 8 de las 9 xilanasas identificadas en el genoma, se secretan por crecimiento en biomasa. A su vez, de las 8 xilanasas secretadas, 7 pertenecen a la familia GH10. En base a esto, se seleccionó una de las enzimas GH10 (KSW20567.1, RefSeq: WP\_062103211.1), que llamamos GH10XynC, para su expresión recombinante y caracterización bioquímica, con el objetivo de comenzar a elucidar su rol en el sistema xilanólico de *Cellulomonas* sp. B6 y evaluar su potencial biotecnológico. Esta enzima no fue estudiada previamente en ningún aislamiento de *Cellulomonas* sp.

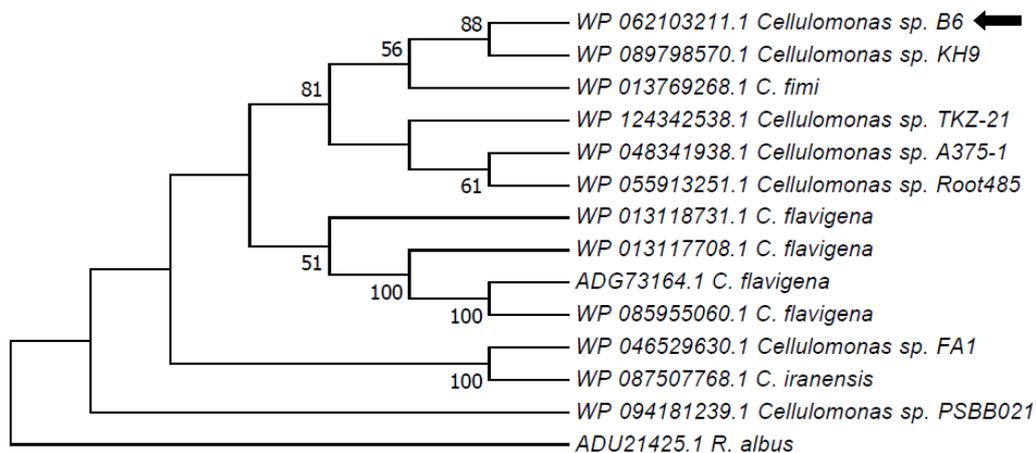
#### 3.1.a. Análisis de secuencia

A partir de los resultados de anotación, complementados con el análisis de la secuencia aminoacídica usando ScanProsite (De Castro y col., 2006), se determinó que GH10XynC tiene 471 aminoácidos y está conformada por un péptido señal (aminoácidos 1 a 33) para secreción por el sistema *Twin Arginine Translocation pathway* (TAT) (Freudl, 2018), un dominio catalítico (CD) de la familia GH10 (aminoácidos 41 a 336) y un dominio de unión a sustrato CBM2 (aminoácidos 375 a 468). El CD y CBM se encuentran unidos por una secuencia rica en treonina (T) y prolina (P) (*PT rich linker*), característica del género *Cellulomonas* (Chaudary y col., 1997). En base al alineamiento con motivos conservados de Prosite, se identificaron los residuos de ácido glutámico E163 y E271 como los aminoácidos catalíticos, dador de protones y nucleófilo, respectivamente, característicos de GH10 (Collins y col., 2005). Se predijo un tamaño molecular de tiene 47,2 kDa (438 aminoácidos) y un pI: 7,80 para la proteína madura usando Expasy Compute pI/MW Tool (Bjellqvist y col., 1993).

La secuencia completa de GH10XynC presentó mayor identidad con una xilanasas de *C. fimi* (84% con 100% cobertura) y una xilanasas de *Cellulomonas* sp. KH9 (83% con 100% cobertura), depositadas bajo los números de acceso de RefSeq: WP\_013769268.1 y WP\_089798570.1, respectivamente.

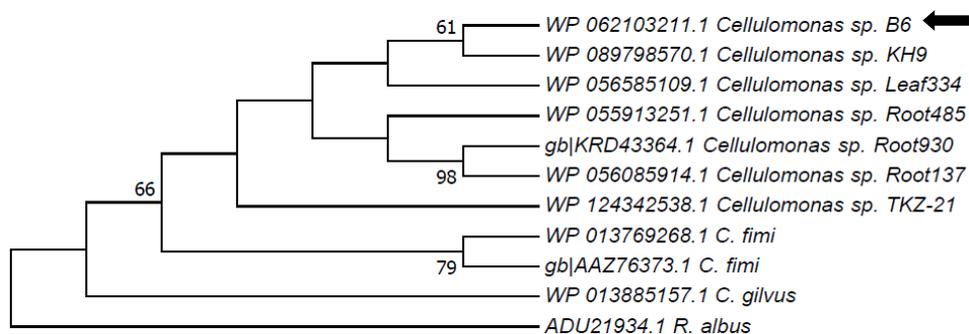
El CD presentó alta identidad de secuencia ( $\geq 80\%$ , cobertura  $\geq 95\%$ ) con dominios GH10 de otras proteínas anotadas como xilanasas de *Cellulomonas* sp. disponibles en

la base de datos de secuencias proteicas no redundantes de NCBI. Estas secuencias fueron utilizadas para el análisis filogenético de rGH10XynC por *Neighbor-Joining* (1000 repeticiones de *bootstrap*), en el que se incluyó un dominio GH10 la bacteria anaerobia celulolítica *Ruminococcus albus* DSM 20455 como *outgroup* (Figura R15).



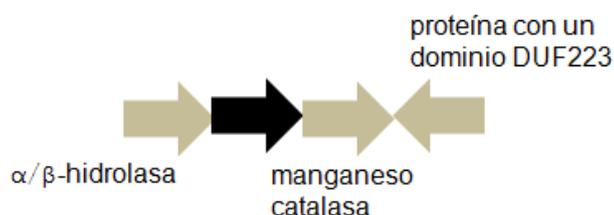
**Figura R15.** Árbol filogenético del dominio GH10 de rGH10XynC, construido por el método *Neighbor-Joining* con 1000 repeticiones de *bootstrap* (se muestran sólo los resultados  $\geq 50\%$ ), usando el programa MEGA v7.0. Se detalla el número de acceso de NCBI/GenBank para cada secuencia. La rama correspondiente a rGH10XynC se indica con una flecha. Se usó la secuencia ADU21425.1 de *R. albus* como *outgroup* para enraizar el árbol.

En base a este análisis, el dominio catalítico de rGH10XynC fue agrupado en un *cluster* con el de una  $\beta$ -1,4-xilanasas de *Cellulomonas* sp. KH9 (WP\_089798570.1) y el de una endo-1,4- $\beta$ -xilanasas de un aislamiento de *C. fimi* (WP\_013769268). De manera similar, se estudio el dominio CBM2, que formó un grupo monofilético con el CBM de la misma xilanasas de *Cellulomonas* sp. KH9 (Figura R16).



**Figura R16.** Árbol filogenético del CBM2 de rGH10XynC, construido por el método *Neighbor-Joining* con 1000 repeticiones de *bootstrap* (se muestran sólo los resultados  $\geq 49\%$ ), usando el programa MEGA 7. Se detalla el número de acceso de NCBI/GenBank para cada secuencia. La rama correspondiente a rGH10XynC se indica con una flecha. Se usó la secuencia ADU21934.1 de *R. albus* como *outgroup* para enraizar el árbol.

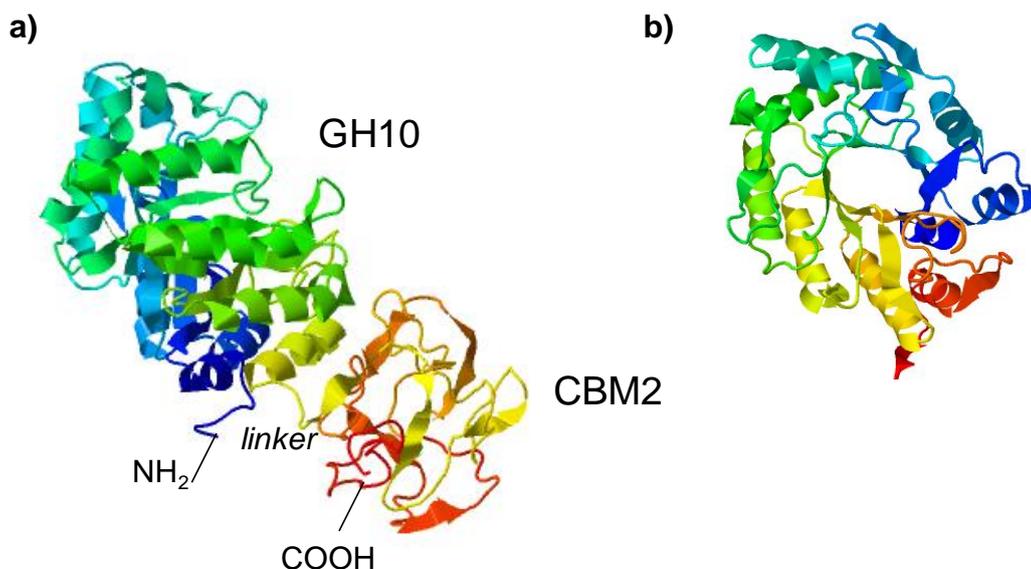
El gen que codifica para GH10XynC (*xynC*) se encuentra flanqueado por las secuencias para proteínas predichas como una  $\alpha/\beta$ -hidrolasa (RefSeq: WP\_062103210.1), una manganeso catalasa (RefSeq: WP\_062103212.1) y una proteína con un dominio no caracterizado DUF223 (GenBank: KSW20569.1) (Figura R17), según el análisis de la región por la anotación mediante RAST y NCBI. No fue posible identificar la secuencia promotora (análisis con BPROM, Softberry) (Solovyev y Salamov, 2011), probablemente porque esta secuencia está al inicio del *contig*.



**Figura R17.** Entorno genómico de *xynC*. Se indican los productos génicos predichos para cada gen identificado. El gen *xynC* se indica en negro.

### 3.1.b. Modelado de la estructura de GH10XynC

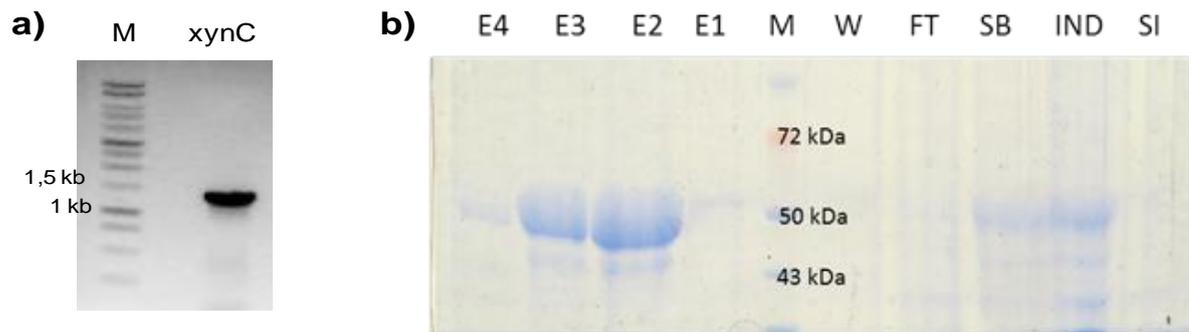
A partir de la secuencia aminoacídica de la proteína madura se generó un modelo estructural usando la plataforma I-Tasser (Roy y col., 2011). El modelo generado fue obtenido a partir de las estructuras de PDB lisyA, lisvA, 1v6wA, 2dlzA, lv0IA, 4bs0A y lv6uA (*C-score* = -0,06, *TM* =  $0,71 \pm 0,12$  y *RMSD* =  $7,1 \pm 4,2$  Å) (Figura R18 a). De esta manera, se determinó que el dominio catalítico consiste en un barril beta, en el que se alternan ocho hélices alfa y ocho hojas beta plegadas paralelas (Figura R18 b). Este dominio es típico de la familia GH10 y forma una estructura de *bowl* (Linares-Pastén, 2018). Este dominio está ligado al módulo CBM2, formado por dos hojas- $\beta$  anti-paralelas ( $\beta$ -sandwich). Adicionalmente, se generó un modelo de estructura para el CD (Figura R18 b), obtenido a partir de las estructuras de PDB 1isyA, 1nq6A, 5m0kA, 5m0k, 1isy y 4bs0A, de mejor calidad de ajuste que el modelo de la proteína completa (*C-score* = 1,54, *TM* =  $0,93 \pm 0,03$ , *RMSD* =  $3,2 \pm 2,2$  Å).



**Figura R18.** Modelo de la estructura tridimensional de la proteína GH10XynC correspondiente al modelo de mejor ajuste, generado usando la plataforma I-Tasser para: a) la proteína madura completa y b) el dominio catalítico. En la proteína completa, se indica la localización de los dominios GH10 y CBM2 y del linker que los une, así como de los extremos amino (-NH<sub>2</sub>) y carboxi (-COOH) terminales.

### 3.1.c. Expresión recombinante y purificación de rGH10XynC

La secuencia codificante para la proteína rGH10XynC madura fue amplificada por PCR a partir de ADN genómico de *Cellulomonas* sp. B6 (Figura R19 a), con los *primers* gh10EMf y gh10EMr, obteniéndose un producto de longitud acorde con la esperada (1333 pb) (Figura R19 a). Dado el alto contenido de G+C del genoma de *Cellulomonas* sp. B6, la metodología para amplificación tuvo que ser optimizada previamente, logrando establecer el protocolo adecuado. El producto de amplificación fue clonado en el vector pGEM-Teasy (Promega), y la identidad e integridad de la secuencia fueron verificadas por secuenciación del inserto completo. La secuencia codificante fue liberada del plásmido utilizando los sitios BamHI y HindIII y clonada en el vector de expresión pET28a, previamente digerido con las mismas enzimas, de manera de obtener una fusión aminoterminal de rGH10XynC a la secuencia de 6 histidinas codificada en el vector. La fusión correcta fue verificada por secuenciación de la región correspondiente del plásmido pET28a-rGH10XynC.



**Figura R19.** a) Electroforesis en gel de agarosa 0,8% de los productos de amplificación. M: Marcador de tamaño molecular 1 Kb (Promega), xynC: amplicón. b) SDS-PAGE 10% de las proteínas correspondientes a *E. coli* Rossetta pET28a-GH10XynC: cultivo sin inducir (SI) e inducido (IND), fracción soluble (SB), percolado (FT), lavado (W) y las fracciones de elución de la proteína rGH10XynC purificada 1-4 (E1, E2, E3 y E4).

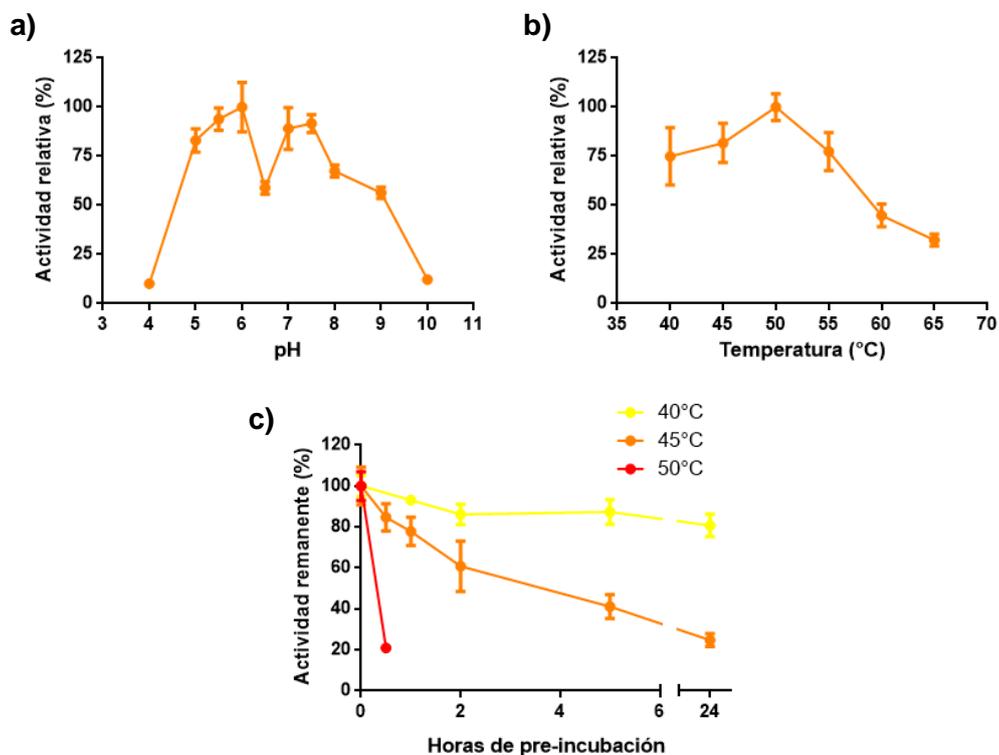
Para la expresión recombinante y purificación, el plásmido pET28a-rGH10XynC fue utilizado para transformar *E. coli* Rossetta pLys, obteniéndose 4 clones *E. coli* Rossetta pET28a-rGH10XynC. Un clon fue seleccionado para evaluar la expresión de la proteína recombinante. Se establecieron las condiciones óptimas de expresión que resultaron en la inducción en un cultivo de  $DO_{600nm}$  0,8 con 0,5 mM IPTG, a 20°C, 16 hs. En estas condiciones se obtuvo un alto porcentaje de la proteína en la fracción soluble (Figura R19 ab. La proteína fue producida y posteriormente purificada por cromatografía de afinidad usando la resina Ni-NTA (QIAGEN), con un rendimiento de 65 mg de proteína por litro de cultivo.

## 3.2. Caracterización de la actividad enzimática de rGH10XynC

### 3.2.a. Identificación de los sustratos de reacción y efecto del pH y la temperatura sobre la actividad

En una primera instancia, se evaluó la actividad de la enzima sobre xilano de madera de haya, CMC y pNP-X (para evaluar actividad  $\beta$ -xilosidasa). En concordancia con el análisis de secuencia, sólo se detectó actividad endoxilanasas y se descartaron las actividades endoglucanasas y  $\beta$ -xilosidasas. Por lo tanto se le asignó a rGH10XynC la actividad EC 3.2.1.8 (xilanasas) y es una xilanasas libre de actividad celulasa.

A fin de optimizar las condiciones de reacción, se construyeron los perfiles de actividad en función del pH (4 a 10) y de la temperatura (40 a 65°C). El efecto del pH se evaluó a 50°C y el de la temperatura, a pH 6.



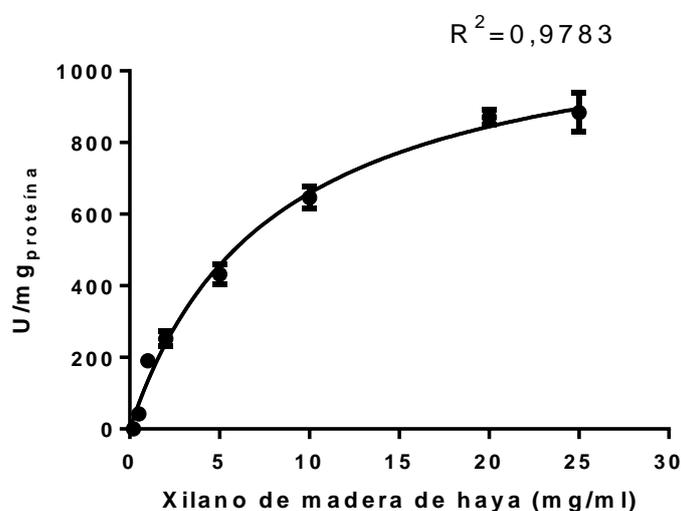
**Figura R20.** Actividad de rGH10XynC. Efecto de: a) pH y b) temperatura; c) estabilidad térmica de rGH10XynC a 40, 45 y 50°C. Condiciones de reacción: 0,5% xilano, 0,6 µg/ml rGH10XynC, 10 minutos, 400 rpm. El 100% de actividad corresponde a  $(351,5 \pm 24,2)$  UI/mg.

La actividad resultó máxima a pH 6 y 50°C y se mantuvo por encima del 50% en un rango de pH de entre 5 y 9 y entre 40 y 55°C (Figura R20 a y b). El efecto de la temperatura en la estabilidad fue estudiado por incubación de la enzima a 40, 45 y 50°C por períodos de hasta 24 horas, seguido de ensayos de actividad en condiciones óptimas (pH6, 50°C, 10 minutos de reacción). La actividad remanente se expresó en función de la actividad enzimática sin pre-incubar (Figura R20 c). rGH10XynC mostró excelente estabilidad a 40°C, manteniendo más del 80% de actividad hasta por 24 horas. A su temperatura óptima (50°C), la actividad decae al 20% en una hora de incubación. Por lo tanto, aunque las condiciones óptimas de reacción se establecieron en 50°C y pH6, para reacciones prolongadas la temperatura de elección es 40°C.

### 3.2.b. Cinética de reacción sobre xilano

Una vez establecidas las condiciones óptimas (50°C y pH 6), se evaluó la cinética enzimática sobre xilano (de madera de haya) Para ello, se incubó la enzima con concentraciones crecientes de sustrato (de 0 a 25 mg/ml) a tiempos cortos de reacción (8 minutos). Los resultados fueron analizados estadísticamente para evaluar el ajuste a los modelos de Michaelis Menten y alosterismo.

La cinética se ajustó al primer modelo, siendo  $V_{m\acute{a}x}$  (1177,00  $\pm$  58,39)  $\mu$ moles/min/mg y  $K_M$  (7,87  $\pm$  1,03) mg/ml (Figura R21). Estos resultados son coherentes con la predicción de un único sitio activo en la estructura.



**Figura R21.** Ajuste de la cinética de reacción de rGH10XynC sobre xilano al modelo de Michaelis-Menten.

### 3.2.c. Efecto de cationes divalentes sobre la actividad

Se ha reportado que la presencia de cationes divalentes, como  $Ca^{2+}$ , puede estabilizar las enzimas y mejorar su termoestabilidad (Spurway y col., 1997), o bien tener un efecto negativo sobre la actividad, como ha sido descrito para xilanasas de *C. fimi* con  $Cu^{2+}$ ,  $Ca^{2+}$  y  $Fe^{2+}$  (Khanna y Gauri, 1993). Por lo tanto, se evaluó el efecto de cationes divalentes sobre la actividad de rGH10XynC en presencia de  $Ca^{2+}$ ,  $Cu^{2+}$ ,  $Mn^{2+}$ ,  $Ni^{2+}$ ,  $Mg^{2+}$  y  $Co^{2+}$ , a una concentración de 2 mM. La actividad residual se calculó en comparación con un control sin agregado de cationes, tomado como 100% de actividad (474,3  $\pm$  12,5) UI/mg. En ningún caso la presencia de cationes resultó en un

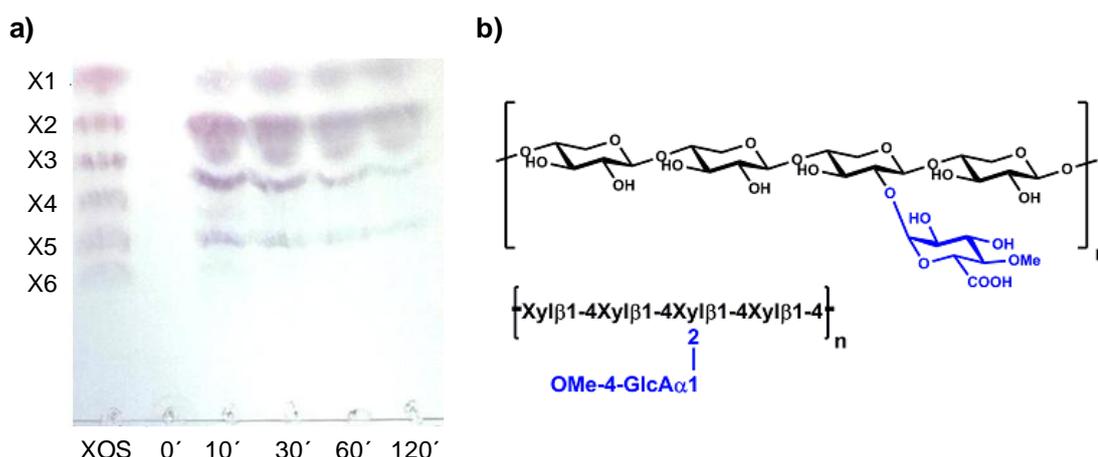
aumento de la actividad (Tabla R14). Sin embargo, la presencia de  $Mn^{2+}$  y  $Co^{2+}$  redujo la actividad xilanasa en aproximadamente 50% y 15%, respectivamente.

**Tabla R14.** Efecto del agregado de cationes divalentes sobre la actividad de rGH10XynC. Prueba t de *Student* para muestras desapareadas (GraphPad Prism v.6), diferencias significativas respecto del control sin adición de cationes: \*  $p < 0,005$ . Condiciones de reacción: 50°C, 30 min, 400 rpm, 0,6 µg/ml de enzima.

Catión (2 mM)	Actividad residual (%)
$Ca^{2+}$	107,5 ± 4,4
$Cu^{2+}$	108,2 ± 3,0
$Mn^{2+}$	52,5 ± 1,9 *
$Ni^{2+}$	98,8 ± 4,3
$Mg^{2+}$	99,6 ± 1,8
$Co^{2+}$	85,0 ± 1,7 *
-	100,0 ± 2,8

### 3.2.d. Mecanismo de actividad de la enzima

La actividad específica fue evaluada sobre arabinoxilano y xilano de haya (glucuronoxilano) y resultó de (444,3 ± 44,6) y (408,6 ± 69,8) UI/mg, respectivamente. Los principales productos de reacción se identificaron por TLC (Figura R22).

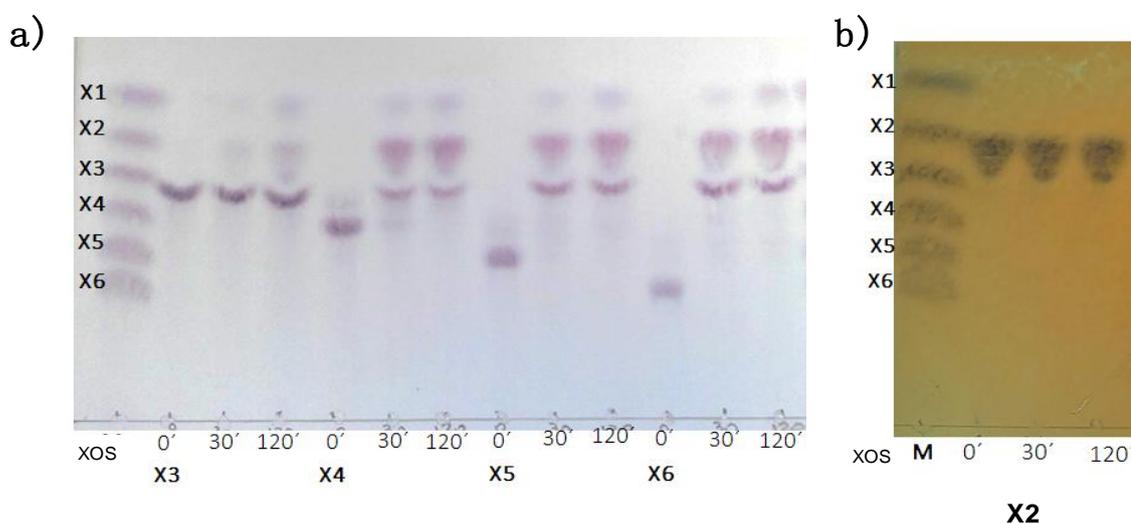


**Figura R22.** a) Productos de reacción de rGH10XynC sobre glucuronoxilano (de haya). b) Estructura del glucuronoxilano de haya (Megazymes).

El producto principal de la actividad de rGH10XynC sobre xilano fue xilobiosa. A tiempos iniciales de reacción, se observaron también xilooligosacáridos cortos X3 y X5.

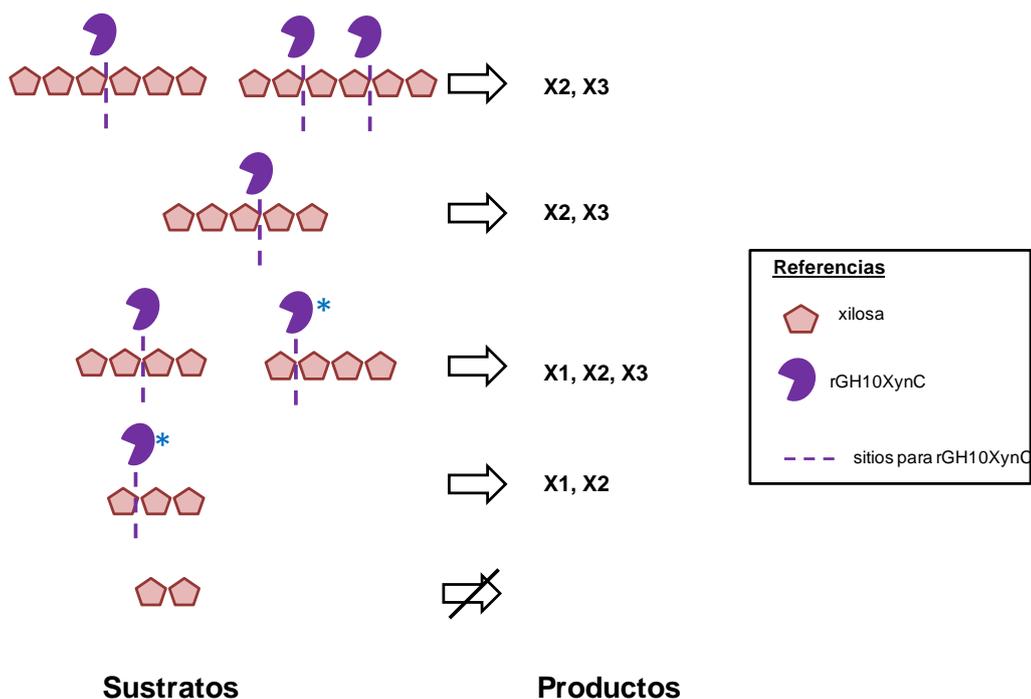
Su concentración disminuyó con el tiempo y aumentó la conversión a xilosa. Se observaron también bandas intermedias, que podrían corresponder a xilooligosacáridos sustituidos, ya que el glucuronoxilano tiene un 13% de sustituciones de ácido 4-O-metilglucurónico en uniones  $\alpha$ -1,2 con las moléculas de xilosa.

Para intentar determinar el mecanismo de acción, se evaluó la actividad de rGH10XynC XOS de grado de polimerización (DP) 2 a 6 (X2, X3, X4, X5 y X6), utilizando una concentración de XOS de 1,5 mg/ml y una concentración de enzima de 0,6  $\mu$ g/ml. Los productos de reacción de tiempo inicial (0), 30 y 120 minutos fueron identificados por TLC (Figura R23).



**Figura R23.** Detección de los productos de reacción sobre a) XOS (X3-X6) y b) X2. Condiciones de reacción: 1,5 mg/ml de XOS o xilobiosa, 5 UI<sub>enzima</sub>/mg<sub>XOS</sub>, 400 rpm a 45°C, en los tiempos indicados en la figura.

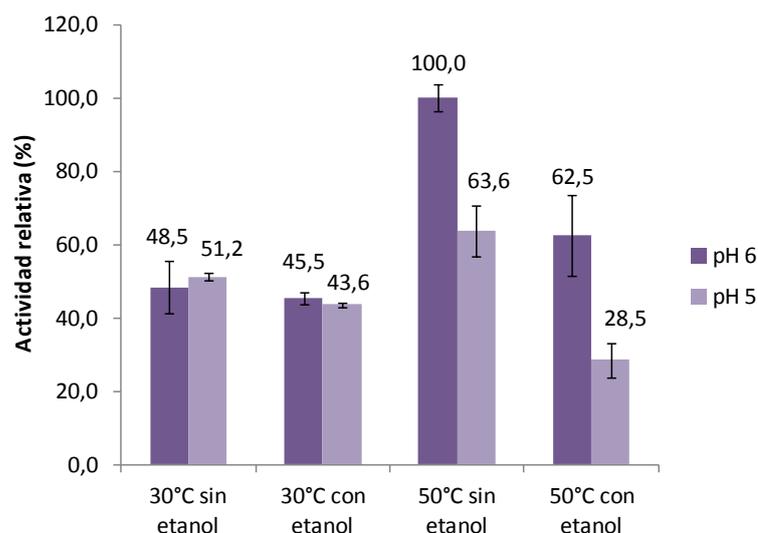
La xilanasa rGH10XynC no presentó actividad sobre X2, confirmando que no tiene actividad  $\beta$ -xilosidasa. Se observó una baja actividad sobre X3, liberando X1 y X2. El producto principal fue X2, aunque también se formaron X3 y X4 en menor proporción.



**Figura R24.** Modelo del mecanismo propuesto de acción de rGH10XynC sobre xilooligosacáridos. Los sitios no preferenciales se indican con un asterisco.

### 3.3. Evaluación de la actividad de rGH10XynC en procesos de bioconversión de biomasa lignocelulósica

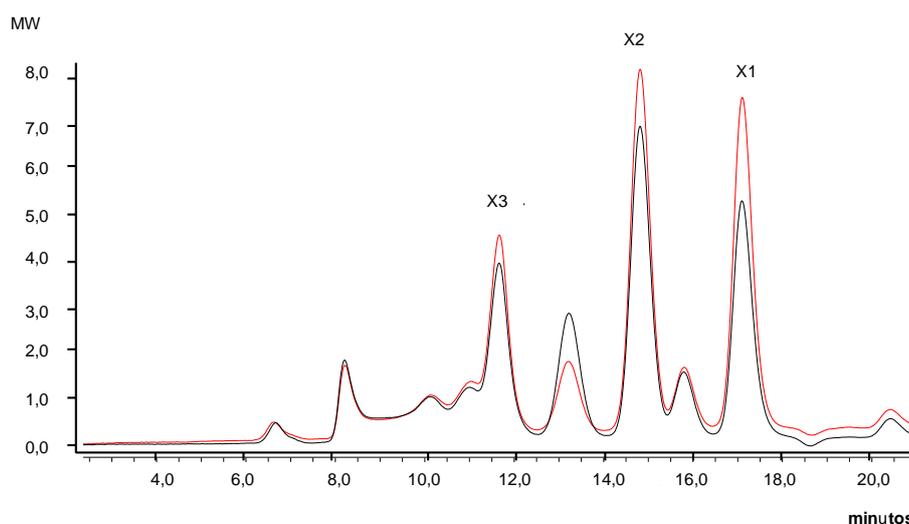
La xilanasas rGH10XynC es una enzima candidata para ser aplicada en procesos a temperaturas moderadas, con pH variable entre 5 y 9. En este sentido, en la industria del bioetanol de 2G, uno de los procesos considerados como promisorios es la sacarificación y co-fermentación simultáneas (SSCF). Los procesos SSCF ocurren a pH entre 4 y 6, alrededor de 30°C, que son las condiciones óptimas de los microorganismos fermentadores. La actividad de la enzima fue ensayada entonces a pH 5 y pH 6, a 30°C en presencia o no de etanol 5% (Figura R25).



**Figura R25.** Actividad de la enzima rGH10XynC en condiciones de SSCF. Se muestra la media de tres replicas sobre cada columna. Condiciones del ensayo: 0,5 mg/ml xilano, 0,6 µg/ml enzima, 30 minutos de incubación a la temperatura, 400 rpm.

En estas condiciones, la enzima conservó cerca de un 43% de actividad (respecto de su óptimo a 50°C), por lo que rGH10XynC podría ser una buena candidata para conformar un cóctel enzimático para SSCF.

La actividad de rGH10XynC sobre xilanos presentes en biomasa se evaluó por detección de los productos de reacción sobre PCE por HPLC (Figura R26), a una concentración de biomasa sólida (seca en estufa hasta peso constante) 2% p/v y con concentraciones de enzima de 2,5 y 5,0 UI/mg<sub>biomasa</sub>.



**Figura R26.** Cromatograma de los productos de reacción de rGH10XynC sobre PCE detectados por HPLC. Condiciones de reacción: 20 mg/ml; 2,5 o 5,0 UI/mg<sub>biomasa</sub> de rGH10XynC (en negro y rojo, respectivamente); 24 hs, 40°C, pH 6,0 y 200 rpm. La identidad de los productos se indica en la figura.

Los principales productos de reacción fueron X2, X1 y X3, siendo X2 el más abundante, lo que concuerda con los resultados obtenidos a partir de xilano (Figura R22). En base a una conversión teórica total (100%) del contenido del xilano presente en 20 mg de biomasa, se calculó la conversión a X1, X2 o X3, lográndose una conversión de aproximadamente 20% con 5 UI/mg<sub>biomasa</sub> (Tabla R15).

**Tabla R15.** Cuantificación de los productos de reacción de rGH10XynC PCE y cálculo del porcentaje de conversión de xilano a X1, X2 y X3.

IU/mg biomasa	mg/ml (conversión)		
	X1	X2	X3
2,5	0,47 ± 0,03 (12%)	0,58 ± 0,05 (16%)	0,31 ± 0,02 (6%)
5,0	0,62 ± 0,00 (14%)	0,72 ± 0,01 (19%)	0,24 ± 0,00 (7%)

Se obtuvieron resultados similares sobre PTE y marlo de maíz dulce (MMD), con porcentajes de conversión a X2 de 24 y 20%, respectivamente.

Estos resultados demuestran que la enzima rGH10XynC es capaz de hidrolizar el xilano contenido en biomasa lignocelulósica a xilooligosacáridos solubles, principalmente, a xilobiosa y xilosa.

# Discusión

En el presente trabajo de Tesis, partimos de la hipótesis de que el aislamiento *Cellulomonas* sp. B6 constituye una fuente novedosa de enzimas aplicables a procesos industriales que requieran la hidrólisis de polisacáridos estructurales, tales como la obtención de etanol celulósico. Estas enzimas podrían reemplazar o complementar a las enzimas fúngicas en varias de sus aplicaciones.

*Cellulomonas* sp. B6 había sido identificado previamente en nuestro laboratorio como el organismo responsable de la actividad endoglucanasa en un consorcio bacteriano obtenido de suelo (Sabarís di Lorenzo, 2013). El género *Cellulomonas* ha sido descrito como el único formado íntegramente por miembros capaces de degradar sustratos celulósicos en condiciones de aerobiosis y anaerobiosis (Christopherson y col., 2013; Poulsen y col., 2016).

En esta Tesis, nos propusimos estudiar el sistema de enzimas de degradación de polisacáridos de este aislamiento, con la finalidad de contribuir a entender el sistema (hemi)celulolítico del género *Cellulomonas* sp. e identificar las enzimas responsables de las actividades celulasa y xilanasas, mediante estrategias de genómica y proteómica. A su vez, nos planteamos la optimización de la secreción de enzimas hidrolíticas y la caracterización de la actividad de los extractos enzimáticos del aislamiento. Asimismo, la expresión de una xilanasas en forma heteróloga nos permitió profundizar la caracterización de su actividad enzimática. El objetivo final fue la evaluación de la capacidad de extractos enzimáticos crudos, y de enzimas de *Cellulomonas* sp. B6 purificadas, de degradar polisacáridos estructurales contenidos en la biomasa vegetal, para su potencial aplicación en bioprocesos.

Los ciclos biológicos de las bacterias son más cortos en comparación con los de los hongos, lo cual constituye una característica ventajosa para las actividades industriales basadas en lignocelulosa (Gomez Del Pulgar y Saadeddin, 2014). Por su habilidad de ocupar hábitats dentro de un amplio espectro de condiciones de pH, temperatura y salinidad, estos organismos constituyen una fuente potencial de enzimas para su aplicación industrial (Dalmaso y col., 2015).

En este trabajo demostramos que el aislamiento *Cellulomonas* sp. B6 es capaz de crecer y secretar enzimas con actividad celulasa y xilanasas por cultivo en sustratos lignocelulósicos complejos, ya sea previamente tratados por extrusión o incluso en biomasa sin pretratar, como el residuo agrícola de caña de azúcar (RAC). Es un

aislamiento mesófilo y las actividades extracelulares celulasa y xilanasas fueron máximas a temperaturas mayores (entre 40 y 55°C), con buena estabilidad térmica a 45°C. El mayor nivel de secreción de proteínas y de actividad enzimática se logró por cultivo en medio mínimo salino suplementado con biomasa lignocelulósica, siendo que las proteínas totales secretadas fueron entre 0,5-0,8 mg/ml y las actividades xilanasas y endoglucanasas en el orden de 0,5 y 0,2 UI/ml de sobrenadante, respectivamente. Estos niveles de secreción de proteínas y actividad xilanasas resultaron promisorios para la aplicación del extracto celular crudo y podrían ser optimizados por cultivo en bioreactores. Por ejemplo, una cepa considerada hipersecretora obtenida por mutagénesis química, *C. flavigena*, PR-22, produce alrededor de 0,65 mg/ml de proteínas por cultivo en *batch* (Rojas-Rejón y col., 2011).

Los principales productos de degradación de celulosa soluble y xilano usando el sobrenadante de cultivo de B6 como extracto enzimático fueron celo (G2 y G3) y xilooligosacáridos (X2 a X6), respectivamente. De acuerdo con estos resultados, las actividades  $\beta$ -glucosidasa y  $\beta$ -xilosidasa fueron detectadas mayormente en el extracto intracelular, lo que es característico de otras bacterias celulolíticas aerobias como *Actinosynnema mirum* y *Catenulispora acidiphila* (Anderson y col., 2012), que secretan enzimas capaces de degradar polisacáridos a oligosacáridos. Los oligosacáridos son internalizados vía transportadores de azúcares del sistema ABC e hidrolizados en el citosol celular (Davidson y Chen, 2004; Saini y col., 2015).

Al evaluar la actividad del extracto extracelular sobre biomasa (paja de trigo pre-tratada por extrusión, PTE), se liberaron principalmente xilo-oligosacáridos (XOS), siendo la xilobiosa el producto más abundante, y también celo-oligosacáridos y xilosa en bajas concentraciones. Esto demostró que las enzimas secretadas tienen actividad sobre los polisacáridos contenidos en la biomasa y que la principal actividad del extracto extracelular es la actividad xilanasas.

Con el objetivo de asignar las actividades observadas a enzimas particulares y de caracterizar en mayor profundidad al aislamiento B6, procedimos a la secuenciación y anotación del genoma, lo que nos permitió, adicionalmente, acercarnos a su clasificación taxonómica.

El genoma de B6 fue ensamblado en 279 contigs, un buen número para aislamientos de *Cellulomonas* sp., dado que el alto contenido de G+C representa un

obstáculo en el ensamblado de las secuencias. El genoma del aislamiento *Cellulomonas* sp. HZM, secuenciado por Illumina Miseq, se logró ensamblar en 78 *contigs* (Chua y col., 2015). Otros aislamientos del género (B12, K38, K39 y K42B) cuyo genoma fue secuenciado con la plataforma Illumina Hiseq 2000, fueron ensamblados en un número de *contigs* comprendido entre 1150 y 1800, aproximadamente (Brookshier y col., 2018).

Por análisis de la secuencia de ARN ribosomal 16S, el aislamiento B6 resultó emparentado filogenéticamente con la especie tipo *Cellulomonas flavigena* ATCC 482, ya caracterizada, y con una especie recientemente descrita, *Cellulomonas persica* ATCC 700642, sin poder asignarlo inequívocamente a alguna de las dos especies.

Sin embargo, como consecuencia de la aparición de tecnologías tipo *shotgun* para la secuenciación de genomas completos, la taxonomía microbiana se ha orientado hacia un enfoque polifásico que conjuga la información obtenida de secuencias genómicas y las caracterización fenotípica de la taxonomía clásica (Mahato y col., 2017).

En este contexto, empleamos herramientas de genómica comparativa (la identidad aminoacídica y nucleotídica promedio, o AAI y ANI, respectivamente; y el cálculo de la distancia entre genomas, o GGDH) para tratar de asignar a *Cellulomonas* sp. B6 a una especie ya descrita (Goris y col., 2007; Meier-Kolthoff y col., 2013). El valor de ANI y AAI entre los genomas de *Cellulomonas* sp. B6 y *C. flavigena* o *C. fimi* resultó, en todos los casos, menor que el valor de corte, que es de 95% (80 y 79, respectivamente, para *C. flavigena*; 81 y 70%, respectivamente, para *C. fimi*). Dado que el número de pares de bases secuenciadas del genoma de *C. persica* es bajo, no se pudo hacer el cálculo de ANI ni AAI entre *C. persica* y B6.

Sin embargo, sí pudimos estimar la GGDH de *Cellulomonas* sp. B6 y las tres cepas tipo, que fue de 25, 30 y 22% para *C. flavigena*, *C. persica* y *C. fimi*, respectivamente. Los valores de GGDH han sido ajustados para emular los de DDH determinados en forma experimental. Para distintas especies de *Cellulomonas* sp., por ejemplo *C. flavigena* y *C. fimi*, sus genomas tienen una DDH de 29%. En cambio, para cepas de una misma especie, como *C. fimi*, los valores se encuentran en el rango de 90 y 100% (Stackebrandt y Kandler, 1979). De manera que si GGDH y DDH son efectivamente comparables, los valores de GGDH obtenidos para B6 y las especies tipo resultan consistentes con los DDH determinados para especies distintas de *Cellulomonas*.

Esta conclusión fue reforzada por la caracterización fenotípica, ya que *Cellulomonas* sp. B6 difiere en cuanto al metabolismo de ciertos hidratos de carbono reportado para *C. fimi*, *C. flavigena* y *C. persica* y tiene una composición de ácidos grasos distinta de *C. flavigena* y *C. fimi* (Stackebrandt y Schumann, 2015; Elberson y col., 2000). Lamentablemente no se ha publicado información referente a la composición de ácidos grasos de *C. persica*.

Globalmente, todos estos resultados sugieren que B6 es un aislamiento novedoso respecto de las especies más estudiadas hasta el momento y que podría ser una nueva especie, si bien son necesarios estudios complementarios para poder confirmarlo.

Combinando los objetivos de adjudicar las actividades enzimáticas de B6 a enzimas particulares y de buscar de enzimas (hemi)celulolíticas potencialmente novedosas, construimos el CAZoma teórico (el conjunto de las enzimas activas sobre carbohidratos codificadas en el genoma del organismo) de *Cellulomonas* sp. B6 y lo comparamos con los reportados para *C. fimi* y *C. flavigena*.

El CAZoma de *Cellulomonas* sp. B6 consistió en 205 proteínas con dominios CAZy, de las cuales 91 fueron GHs. Estos resultados contribuyen a diferenciar a B6 de las especies tipo de *C. flavigena* y *C. fimi*, para las que se encontraron 86 y 113 GHs (base de datos CAZy, 2017). En *C. flavigena*, la ausencia de enzimas con dominios GH74, GH18 y la presencia de más de una GH11 constituyen las diferencias más interesantes. Para *C. fimi*, en contraste, si bien hay una sola GH11 y una GH74, al igual que en *Cellulomonas* sp. B6, algunas de las diferencias radican en el número de GH1 y GH3. No se identificaron dominios *scaffold*, *dockerin*, ni cohesina en el genoma de B6 que pudieran indicar la formación de celulosomas. A pesar de la capacidad de degradar celulosa en aerobiosis y anaerobiosis, lo mismo ha sido descrito para los genomas de *Cellulomonas* sp. estudiados hasta el momento (Christopherson y col., 2013), lo que indica que gran parte de las CAZimas serían secretadas en unidades libres.

Con el objetivo de estudiar qué enzimas son realmente secretadas por cultivo en sustratos celulósicos y, por lo tanto, responsables de las actividades hidrolíticas detectadas, realizamos ensayos de espectrometría de masas a partir de sobrenadantes completos de cultivos en CMC, RAC y PTE. Las actividades enzimáticas de dichos sobrenadantes coincidieron con las previamente obtenidas, es decir, la mayor actividad

y la mayor secreción de proteínas se obtuvo a partir de cultivo en biomasas, si bien los niveles de crecimiento bacteriano fueron similares en los tres sustratos.

De esta forma, identificamos 21 de las 32 GHs predichas como extracelulares. Entre ellas, una GH6, la única GH48 codificada en el genoma y 2 GH10 fueron detectadas en todas las condiciones.

Las GH48 y GH6 son exoglucanasas desde el extremo reductor y no-reductor respectivamente. Estas enzimas son las únicas exoglucanasas codificadas en el genoma. Fueron identificadas en el extracto extracelular de todas las condiciones de cultivo, tanto en biomasa como CMC, aunque no por cultivo en sacarosa, lo que sugiere su rol como las celulasas principales del género *Cellulomonas* y serían entonces las responsables de la actividad celobiohidrolasa.

Exoglucanasas GH6 y GH48 han sido descritas para en *C. fimi* (Meinke y col., 1994) y *C. flavigena* (Shen y col., 1995). Por estudios recientes de secretomas de *C. fimi*, las celobiohidrolasas CbhA (GH6) y CbhB (GH48) fueron identificadas como las CAZimas más abundantes por cultivo en CMC y en biomasa (Spertino y col., 2018). En el secretoma *C. flavigena*, los homólogos de CbhA y CbhB fueron identificados por cultivo en CMC y el homólogo de CbhA fue detectado también en xilano (Wakarchuk y col., 2016). Por lo tanto, el sistema de exoglucanasas GH48/GH6 sería fundamental para la degradación de celulosa en el género *Cellulomonas*.

Otras exoglucanasas con características similares han sido reportadas para especies de otros géneros de bacterias celulolíticas, como *Clostridium cellulolyticum* (Bélaich y col., 1997) y *Thermobifida fusca* (Gómez del Pulgar y Saadedin, 2014). La actividad de las exoglucanasas GH48, sumada a la acción de endoglucanasas GH9, se considera crucial para la degradación de celulosa cristalina (Brunecky y col., 2017). En nuestro trabajo, se detectaron todas las endoglucanasas GH9 extracelulares (4) por crecimiento en biomasa. Estos resultados sugieren que *Cellulomonas* sp. B6 requiere del sistema de exo-endoglucanasas para la deconstrucción de celulosa contenida en la biomasa.

Cabe destacar que el cultivo en un sustrato puramente celulósico, como CMC, resultó en la producción de tres xilanasas extracelulares GH10. De manera similar, muchas GH10 fueron detectadas en sobrenadantes de cultivo de *C. fimi* y *C. flavigena* crecidos en CMC (Spertino y col., 2018; Wakarchuk y col., 2016). Esto sugiere una regulación concertada de la inducción de la expresión de ciertas celulasas y xilanasas.

Además, por cultivo en biomasa, *Cellulomonas sp.* B6 secretó todo el repertorio de xilanasas extracelulares codificado en el genoma: siete GH10 y una GH11. Todavía no está clara la explicación para esta aparente redundancia de xilanasas GH10, si bien podría estar relacionada con la heterogeneidad de la composición de los xilanos o las sustituciones que presentan. Resultados similares se encontraron en *C. flavigena*, que tiene también un repertorio amplio de xilanasas. De las 16 xilanasas predichas, se identificaron 12 GH10 y las 3 GH11 en secretomas de CMC y xilano (Wakarchuk y col., 2016). Las GH10 identificadas en este trabajo tienen diferente conformación modular. Todas ellas presentan un único módulo catalítico, pero distinto número y tipo de CBM. En nuestros estudios, la GH10 más abundante fue KSW23552.1, que posee además dominios de unión a sustrato CBM9 y CBM22.

La única GH11 de *Cellulomonas sp.* B6 fue detectada exclusivamente en cultivo crecido en biomasa. En cambio, una GH11 de *C. flavigena* fue identificada por cultivo en sustratos puros (CMC y xilano) y la GH11 de *C. fimi* fue identificada en cultivo en xilano (Wakarchuk y col., 2016). En cuanto a la regulación de la expresión, se describió la regulación positiva de la transcripción de una GH11 de *C. flavigena* en presencia de biomasa lignocelulósica (Mayorga-Reyes y col., 2002).

En definitiva, en base a nuestros estudios y los antecedentes, en el género *Cellulomonas*, el crecimiento en biomasa lignocelulósica resulta en la inducción de la secreción de prácticamente todo el repertorio de xilanasas y su actividad sería necesaria para utilización de la hemicelulosa presente en la biomasa. Queda por estudiar la regulación de la expresión de las mismas y su afinidad por los sustratos, que aún no han sido dilucidados.

En hongos celulolíticos, *Trichoderma reesei* y *Aspergillus niger*, la secreción de CAZimas también se ha descrito como dependiente de la fuente de cultivo (Borin y col., 2015). Por cultivo en bagazo o tallo de caña de azúcar, se han identificado 45 y 86 CAZimas en total producidas por *T. reesei* y *A. niger*, respectivamente. En ambos casos, se detectaron  $\beta$ -glucosidasas y  $\beta$ -xilosidasas (con dominios GH1, GH3 o GH43), endoglucanasas GH5 y/o 6, celobiohidrolasas GH6 y GH7 (equivalentes a las GH48 bacterianas), xilanasas GH10, y dos quitinasas GH18 entre otras enzimas. En *T. reesei*, se identificó también una GH11 y una GH74 únicas (potencialmente activas sobre xilano y

xiloglucanos, respectivamente), entre otras CAZimas. En *A. niger*, además de una única GH74, se detectaron GH12 supuestamente activas sobre xiloglucanos.

La batería de enzimas secretadas por *Cellulomonas* sp. tiene algunos puntos en común con los secretomas de estos hongos, particularmente, con *T.reesei*, dada la producción de múltiples celulasas y xilanasas, con un número acotado de enzimas activas sobre otras hemicelulosas y pectinas. La variedad y el número xiloglucanasas y pectinasas es notablemente mayor en *A. niger*. Además, las enzimas más abundantes en el secretoma de *T. reesei* son dos celobiohidrolasas, Cel7A (GH7) y Cel6A (GH6), mientras que, en *A. niger*, las más abundantes son la  $\alpha$ -L-arabinofuranosidasa B (GH54) y la xilanasas 2 (GH11).

La GH74 y la GH18 de *Cellulomonas* sp. B6 fueron detectadas en el sobrenadante de cultivo en biomasa. En *C. fimi*, proteínas homólogas fueron identificadas también el sobrenadante de cultivo en biomasa. Si bien su actividad no ha sido descrita para *C. fimi*, sí fue caracterizada una GH18 de *C. uda* capaz de producir quito-oligosacáridos a partir de sustratos insolubles (quitina y quitosano) y se ha sugerido que los miembros de *Cellulomonas* sp., como otras actinobacterias, además de los hongos, podrían tener un rol importante en la degradación de quitina, resultante de la descomposición de artrópodos y hongos, en los ecosistemas terrestres (Reguera y Leschine, 2001 y 2003). No hay hasta el momento bibliografía sobre GH74 de *Cellulomonas* sp. que hayan sido caracterizadas, pero, dado que el extracto extracelular presenta actividad sobre xiloglucano, esta única GH74 podría ser la responsable de su degradación. Similarmente, en el sobrenadante cultivo de *T. reesei* se detectó actividad xiloglucanasa que se debería a una única GH74 (Borin y col., 2005).

Para entender la contribución de las distintas xilanasas, es necesaria la caracterización individual de las mismas. Se han expresado en forma recombinante xilanasas GH10 y GH11 de *C. flavigena* y se ha comprobado que esta última es menos activa sobre xilooligosacáridos que las GH10, lo que sugiere que difieren en la especificidad sobre xilanos y sus derivados de hidrólisis (Lisov y col., 2017).

En función de estos resultados, nos propusimos contribuir a dilucidar el rol de las xilanasas de B6. Expresamos en forma heteróloga (en *E. coli*) una de las GH10 identificada en el sobrenadante de cultivo, a la que llamamos rGH10XynC. Esta enzima

fue seleccionada porque no había sido caracterizada aún en ninguna especie de *Cellulomonas sp.* y en base a la factibilidad de su expresión recombinante.

Como la mayoría de las CAZimas secretadas por el aislamiento, rGH10XynC es una proteína multimodular, cuyo dominio catalítico está asociado a un dominio de unión a sustrato de la familia CBM2, ambos de estructura típica, revelada por modelado molecular (Collins y col., 2005, Linares-Pastén y col., 2018). El modelado molecular también permitió identificar los potenciales residuos catalíticos, característicos de las endoxilanasas de esta familia. El dominio CBM2 ha sido descrito como parte de una gran variedad de enzimas bacterianas, y en algunos casos se ha demostrado afinidad a celulosa, xilano o quitina. En este trabajo se demostró experimentalmente la actividad endoxilanasas de rGH10XynC sobre glucurono y arabinoxilano y XOS, dado que se detectó xilobiosa como producto mayoritario de hidrólisis, con liberación de oligosacáridos cortos y, en tiempos largos de reacción, de xilosa. rGH10XynC no mostró actividad sobre X2 ni pNP-X, descartando la actividad  $\beta$ -xilosidasa, y tampoco fue activa sobre CMC. En base a estos resultados, podemos confirmar que rGH10XynC es una endoxilanasas libre de actividad celulasas, asignándole el EC 3.2.1.8.

La actividad de rGH10XynC es óptima a pH 6 y una temperatura de 50°C, aunque es más estable a tiempos largos a 40°C. Estas características son similares a las reportadas para tres xilanasas GH10 de *C. flavigena* (llamadas CFXyl1, CFXyl2 y CFXyl4), no sólo en el rango de pH y temperatura en que son activas, sino también porque tienen una vida media corta a su temperatura óptima (Lisov y col., 2017). rGH10XynC presenta una baja identidad de secuencia con estas xilanasas, siendo de 35% la identidad (88% de cobertura) con CFXyl1 (WP\_013118747.1), de 69% (cobertura de 71%) con CFXyl2 (WP\_013115627.1) y de 28% (cobertura de 49%) con CFXyl3 (WP\_043598780.1).

La actividad específica de rGH10XynC sobre xilanos de distinta estructura (glucurono y arabinoxilano) en condiciones óptimas es de 400-450 UI/mg, respectivamente. Este nivel de actividad es similar al de enzimas comerciales. Por ejemplo, algunas xilanasas (EC 3.2.1.8) puras (sin agregado de enzimas con actividades complementarias), que están disponibles en el mercado (comercializadas por Megazyme <https://www.megazyme.com/>) tienen una actividad específica que varía entre los 2,5 y los 1500 U/mg, según la enzima. Estas xilanasas son principalmente GH11 de hongos (como *Trichoderma viridae*, E-XYTR1 o *Aspergillus niger*, E-XYAN4) y GH10 de

bacterias (*Aeromonas punctata*, E-XYNAP y *Thermotoga marítima*, E-XYLATM). Las xilanasas fúngicas tienen en general actividad óptima en un rango acotado de pH y las enzimas bacterianas tienen menor actividad específica que rHH10XynC, siendo que E-XYNAP de *A. punctata* (CAS: 9025-57-4) tiene una actividad sobre arabinoxilano de trigo de alrededor de 8,3 U/mg a 40°C y pH 6,5. En comparación, los valores de actividad hacen de rGH10XynC una enzima promisorio para su uso comercial, sobre todo teniendo en cuenta la actividad específica de las xilanasas bacterianas.

En resumen, *Cellulomonas* sp. B6 es capaz de secretar enzimas (hemi)celulolíticas por crecimiento en biomasa residual, que es un sustrato abundante y económico, como única fuente de carbono, incluso en residuo agrícola de caña de azúcar sin pretratamiento previo. Paralelamente, la utilidad de *Cellulomonas* sp. B6 como fuente de CAZimas quedó demostrada por la diversidad de enzimas que secreta. La actividad de la xilanasas GH10XynC, expresada en forma recombinante en *E. coli*, fue demostrada tanto sobre xilano purificado como sobre el xilano presente en biomasa lignocelulósica de distinto origen (paja de cebada, paja de trigo y marlo de maíz dulce) y se encontró en el rango de enzimas comerciales. La utilidad de GH10XynC para actuar sobre la porción hemicelulósica de la biomasa, sin afectar la celulosa es de importancia para la industria del papel (en la cual es necesario conservar la celulosa intacta). Además, por su alta actividad específica a temperaturas moderadas y tolerancia al etanol, podría formar parte de mezclas enzimáticas para la producción de etanol 2G en procesos de SSCF. Otra aplicación posible para rGH10XynC es la obtención de XOS prebióticos (DP entre 2 y 5) a partir de biomasa y el mejoramiento de la digestibilidad del alimento para animales monogástricos. Dada la facilidad de su producción y su alta actividad específica, rGH10XynC es candidata a mejoramiento mediante mutagénesis dirigida para intentar mejorar la estabilidad térmica.

Por todo esto, este trabajo, centrado en el estudio del aislamiento *Cellulomonas* sp. B6, puede considerarse una base para el desarrollo de enzimas con potenciales aplicaciones en el área agro-industrial.

# Bibliografía

- A**bt B, Foster B, Lapidus A, y col (2013). Complete genome sequence of *Cellulomonas flavigena* type strain (134). *Stand Genomic Sci*, 3: 15–25.
- Ahmed I, Kudo T, Abbas S, Ehsan M, Iino T, Fujiwara T, Ohkuma M (2013). *Cellulomonas pakistanensis* sp. nov., a moderately halotolerant *Actinobacteria*. *Int J Syst Evol Microbiol*, 64(7): 2305-11.
- Alfaro M, Oguiza JA, Ramírez L, Pisabarro AG (2014). Comparative analysis of secretomes in basidiomycete fungi. *J Proteomics*, 102: 28-43.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. *J Mol Biol*, 215(3): 403-410.
- Anderson I, Abt B, Lykidis A, Klenk HP, Kyrpides N, Ivanova N (2012). Genomics of aerobic cellulose utilization systems in actinobacteria. *PLoS One*, 7(6): e39331.
- Arantes V, Jellison J, Goodell B (2012). Peculiarities of brown-rot fungi and biochemical Fenton reaction with regard to their potential as a model for bioprocessing biomass. *Appl Microbiol Biotechnol*, 94(2): 323-38.
- Arévalo-Gallegos A, Ahmad Z, Asgher M, Parra-Saldivar R, Iqbal HMN (2017) . Lignocellulose: A sustainable material to produce value-added products with a zero waste approach-A review. *Int J Biol Macromol*, 99: 308-318.
- Axelrod VD, Kramer FR (1985). Transcription from bacteriophage T7 and SP6 RNA polymerase promoters in the presence of 3'-deoxyribonucleoside 5'-triphosphate chain terminators. *Biochemistry*, 24(21): 5716-23.
- Aziz RK, Bartels D, Best AA, y col. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9:75.
- B**allesteros M (2010). Enzymatic hydrolysis of lignocellulosic biomass. En: *Bioalcohol Production: Biochemical Conversion of Lignocellulosic Biomass*, (Ed. Waldron), Woodhead Publishing Series in Energy, pp. 159-177.
- Bélaich JP, Tardif C, Bélaich A, Gaudin C (1997). The cellulolytic system of *Clostridium cellulolyticum*. *J Biotechnol*, 57(1-3): 3-14.
- Bendtsen JD, Kiemer L, Fausbøll A, Brunak S (2005). Non-classical protein secretion in bacteria. *BMC Microbiol*, 5: 58-71.
- Berman M, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000). The Protein Data Bank. *Nucleic Acids Res*, 28: 235-242.
- Bertani G (1951). Studies on lysogenesis. The mode of phage liberation by lysogenic *Escherichia coli*. *J Bacteriol*, 62: 293–300.
- Bhat MK (2000). Cellulases and related enzymes in biotechnology. *Biotechnol Adv*, 18(5): 355-83.

Bhattacharya AS, Bhattacharya A, Pletschke BI (2015). Synergism of fungal and bacterial cellulases and hemicellulases: a novel perspective for enhanced bio-ethanol production. *Biotechnol Lett*, 37(6): 1117-1129.

Bjellqvist B, Hughes GJ, Pasquali Ch, Paquet N, Ravier F, Sanchez JC, Frutiger S, Hochstrasser DF (1993). The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis*, 14: 1023-31.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15): 2114–2120.

Bomble YJ, Lin CY, Amore A, Wei H, Holwerda EK, Ciesielski PN, Donohoe BS, Decker SR, Lynd LR, Himmel ME (2017). Lignocellulose deconstruction in the biosphere. *Curr Opin Chem Biol*, 41: 61-70.

Bon EPS, Ferrara MA (2010). Bioethanol Production via Enzymatic Hydrolysis of Cellulosic Biomass *Biotechnol*, 1: 11.

Borin GP, Sanchez CC, de Souza AP, de Santana ES, de Souza AT, Paes Leme AF, Squina FM, Buckeridge M, Goldman GH, Oliveira JV (2015). Comparative Secretome Analysis of *Trichoderma reesei* and *Aspergillus niger* during Growth on Sugarcane Biomass. *PLoS One*, 10(6): e0129275.

Bosetto A, Justo PI, Zanardi B, Venzon SS, Graciano L, dos Santos EL, Simão Rde C (2016). *Appl Biochem Biotechnol*, 178(4): 766-95.

Bradford MM (1979). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, 72: 248-254.

Brett C, Waldron K (1990). Physiology in Plant Cell Wall. En: *Topics in Plant Physiology*. Black M, Chapman J (Eds.), Unwin Hyman Ltd.

Brevu M, Vasile C (2010). Thermal degradation of lignin - A review. *Cell Chem Technol*, 44(9): 353-363.

Brookshier AM, Santo Domingo JW, Kourtev PS, Learman DR (2018). Draft Genome Sequences of Two *Bacillus* sp. Strains and Four *Cellulomonas* sp. Strains Isolated from Heavy-Metal-Contaminated Soil. *Microbiol Resour Announc*, 7(11): e01063-18.

Brunecky R, Alahuhta M, Sammond DW, Xu Q, Chen M, Wilson DB, Brady JW, Himmel ME, Bomble YJ, Lunin VV (2017). Natural diversity of glycoside hydrolase family 48 exoglucanases: insights from structure. *Biotechnol Biofuels*, 10: 274.

**C**affall KH, Mohnen D (2009). The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. *Carbohydr Res*, 344(14):1879-1900.

Campos E, Negro Alvarez MJ, Sabarís di Lorenzo G, Gonzalez S, Rorig M, Talia P, Grasso DH, Sáez F, Manzanares Secades P, Ballesteros Perdices M, Cataldi AA (2014). Purification and characterization of a GH43  $\beta$ -xylosidase from *Enterobacter* sp. identified and cloned from forest soil bacteria. *Microbiol Res*, 169(2-3): 213-220.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2008). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res*, 37: D233-8.

Chakdar H, Kumar M, Pandiyan K, Singh A, Nanjappan K, Kashyap PL, Srivastava AK (2016). Bacterial xylanases: biology to biotechnology. *3 Biotech*, 6(2): 150.

Chaudhary P, Kumar NN, Deobagkar DN (1997). The glucanases of *Cellulomonas*. *Biotechnol Adv*, 15(2):315-31.

Christopherson MR, Suen G, Bramhacharya S, Jewell KA, Aylward FO, Mead D, Brumm PJ (2013). The genome sequences of *Cellulomonas fimi* and "*Cellvibrio gilvus*" reveal the cellulolytic strategies of two facultative anaerobes, transfer of "*Cellvibrio gilvus*" to the genus *Cellulomonas*, and proposal of *Cellulomonas gilvus* sp. nov. *PLoS One*, 8(1): e53954.

Chua P, Har ZM, Austin CM, Yule CM, Dykes GA, Lee SM (2015). Genome sequencing and annotation of *Cellulomonas* sp. HZM. *Genom Data*, 5: 40–41.

Collins T, Gerday C, Feller G (2005). Xylanases, xylanase families and extremophilic xylanases. *FEMS Microbiol Rev*, 29(1): 3-23.

Corrêa TL, dos Santos LV, Pereira GA (2016). AA9 and AA10: from enigmatic to essential enzymes. *Appl Microbiol Biotechnol*, 100(1): 9-16.

Cragg SM, Beckham GT, Bruce NC, Bugg TD, Distel DL, Dupree P, Etxabe AG, Goodell BS, Jellison J9, McGeehan JE, McQueen-Mason SJ, Schnorr K, Walton PH, Watts J, Zimmer M. (2015). Lignocellulose degradation mechanisms across the Tree of Life. *Curr Opin Chem Biol*, 29: 108-19.

**D**almaso GZ, Ferreira D, Vermelho AB (2015). Marine extremophiles: a source of hydrolases for biotechnological applications. *Mar Drugs*. 13(4):1925-65

Davidson AL, Chen J (2004). ATP-binding cassette transporters in bacteria. *Annu Rev Biochem*, 73: 241-68.

De Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, Bairoch A, Hulo N (2006). ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. *Nucleic Acids Res*, 1: 34.

De Vries RP, Visser J (2001). *Aspergillus* Enzymes Involved in Degradation of Plant Cell Wall Polysaccharides. *Microbiol Mol Biol Rev*, 65(4): 497–522.

Dehority BA (1969). Pectin-fermenting bacteria isolated from the bovine rumen. *J Bacteriol*, 99(1): 189-96.

Den W, Sharma VK, Lee M, Nadadur G, Varma RS (2018). Lignocellulosic Biomass Transformations via Greener Oxidative Pretreatment Processes: Access to Energy and Value-Added Chemicals. *Front Chem*, 6: 141.

**E**lberson MA, Malekzadeh F, Yazdi MT, Kameranpour N, Noori-Dalooi MR, Matte MH, Shahamat M, Colwell RR, Sowers KR (2000). *Cellulomonas persica* sp. nov. and *Cellulomonas iranensis* sp. nov., mesophilic cellulose-degrading bacteria isolated from forest soils. *Int J Syst Evol Microbiol*, 50(3): 993-6.

**F**elsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39: 783–91.

Ferrari AE, Ravnskov S, Wall LG (2018). Crop rotation in no-till soils modifies the fatty acids signature. *Soil Use & Management*, 34: 427-436.

Freudl R (2018). Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb Cell Fact*, 17(1): 52

Fulton L, Howes T, Hardy J (2004). Biofuels for transport: an international perspective. Agencia Internacional de Energía, París, Francia. Disponible *online*: <http://www.iea.org/textbase/nppd/free/2004/biofuels2004.pdf>

**G**alié S, García-Gutiérrez C, Miguélez EM, Villar CJ, Lombó F (2018). Biofilms in the Food Industry: Health Aspects and Control Methods. *Front Microbiol*, 9: 898.

Gao D, Chundawat SP, Krishnan C, Balan V, Dale BE (2010). Mixture optimization of six core glycosyl hydrolases for maximizing saccharification of ammonia fiber expansion (AFEX) pretreated corn stover. *Bioresour Technol*, 101(8): 2770-81

Gao D, Uppugundla N, Chundawat SP, Yu X, Hermanson S, Gowda K, Brumm P, Mead D, Balan V, Dale BE (2011). Hemicellulases and auxiliary enzymes for improved conversion of lignocellulosic biomass to monosaccharides. *Biotechnol Biofuels*, 4: 5.

Ghio S, Ontañón OM, Piccini FE, Marrero Diaz de Villegas R, Talia PM, Grasso DH, Campos E (2016). *Paenibacillus* sp. A59 GH10 and GH11 extracellular endoxylanases: application in biomass bioconversion. *Bioenergy Res*, 11(1): 174-90.

Ghio S, Sabarís Di Lorenzo G, Lia V, Talia P, Cataldi A, Grasso D, Campos E (2012). Isolation of *Paenibacillus* sp. and *Variovorax* sp. strains from decaying woods and characterization of their potential for cellulose deconstruction. *Int J Biochem Mol Biol*. 3(4), 352-364.

Ghose TK (1987). Measurement of cellulase activities. *Pure Appl Chem*, 59(2): 257–68.

Ghose TK, Bisaria VS (1987). Measurement of hemicellulase activities- Part 1: xylanases. *Pure Appl Chem*, 59 (12): 1739-52.

Gírio FM, Fonseca C, Carvalheiro F, Duarte LC, Marques S, Bogel-Lukasik R (2010) Hemicelluloses for fuel ethanol: A review. *Bioresour Technol*, 101(13): 4775-800.

Gloster TM (2014). Advances in understanding glycosyltransferases from a structural perspective. *Curr Opin Struct Biol*, 28: 131-41.

Gomez Del Pulgar EM, Saadeddin A (2014). The cellulolytic system of *Thermobifida fusca*. Crit Rev Microbiol, 40: 236–247.

Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM (2007). DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol, 57: 81–91.

Gorshkova TA, Kozlova LV, Mikshina PV (2013). Spatial structure of plant cell wall polysaccharides and its functional significance. Biochemistry (Mosc), 78(7): 836-853.

Green R, Rogers EJ (2013). Transformation of chemically competent *E. coli*. Methods Enzymol, 529:329-36.

**H**aitjema CH, Solomon KV, Henske JK, Theodorou MK, O'Malley MA (2014). Anaerobic gut fungi: Advances in isolation, culture, and cellulolytic enzyme discovery for biofuel production. Biotechnol Bioeng, 111: 1471-1482.

Hanahan D (1983). Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol, 166: 557–580.

Hankin L, Anagnostakis SL (1977). Solid media containing carboxymethylcellulose to detect cellulase activity of microorganisms. J Gen Microbiol 98:109-115.

Harholt J, Suttangkakul A, Scheller HV (2010). Biosynthesis of pectin. Plant Physiol, 153(2): 384-95.

Harris AD, Ramalingam C (2010). Xylanases and its Application in Food Industry: A Review. J Exp Sci, 1(7): 1-11.

Hatayama K, Esaki K, Ide T (2013). *Cellulomonas soli* sp. nov. and *Cellulomonas oligotrophica* sp. nov., isolated from soil. Int J Syst Evol Microbiol, 63(1): 60-65.

Hemsworth GR, Davies GJ, Walton PH (2013). Recent insights into copper-containing lytic polysaccharide mono-oxygenases. Curr Opin Struct Biol, 23(5): 660-8.

**I**shihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, Mann M (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. Mol Cell Proteomics 4: 1265–72.

**J**ohansen KS (2016). Discovery and industrial applications of lytic polysaccharide mono-oxygenases. Biochem Soc Trans, 44(1): 143-9.

Jones BE, Grant WD, Duckworth AW, Schumann P, Weiss N, Stackebrandt E (2005). *Cellulomonas bogoriensis* sp. nov., an alkaliphilic cellulomonad. Int J Syst Evol Microbiol, 55(4): 1711-4.

**K**ang MS, Im WT, Jung HM, Kim MK, Goodfellow M, Kim KK, Yang HC, An DS, Lee ST (2007). *Cellulomonas composti* sp. nov., a cellulolytic bacterium isolated from cattle farm compost. Int J Syst Evol Microbiol, 57(6): 1256-60.

- Khanna S, Gauri (1993). Regulation, purification, and properties of xylanase from *Cellulomonas fimi*. *Enzyme Microb Technol*, 15: 990-995.
- Kim D (2018). Physico-Chemical Conversion of Lignocellulose: Inhibitor Effects and Detoxification Strategies: A Mini Review. *Molecules*, 23(2): e309.
- Koeck DE, Pechtl A, Zverlov VV, Schwarz WH (2014). Genomics of cellulolytic bacteria. *Curr Opin Biotechnol*, 29: 171-83.
- Kucharska K, Rybarczyk P, Hołowacz I, Łukajtis R, Glinka M, Kamiński M (2018). Pretreatment of Lignocellulosic Materials as Substrates for Fermentation Processes. *Molecules*, 23(11): E2937.
- Kuhad RC, Gupta R, Singh A (2011). Microbial cellulases and their industrial applications. *Enzyme Res*, 2011: 280696.
- Kumar D, Murthy GS (2016). Enzymatic Hydrolysis of Cellulose for Ethanol Production: Fundamentals, Optimal Enzyme Ratio, and Hydrolysis Modeling, pp 65-78. En: *New and Future Developments in Microbial Biotechnology and Bioengineering*, 1 ed. Gupta VG (Ed.), Elsevier.
- Kumar S, Stecher G, Tamura K (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol*, 33(7): 1870-4
- Kumar Saini J, Saini R, Tewari L (2015). Lignocellulosic agriculture wastes as biomass feedstocks for second-generation bioethanol production: concepts and recent developments. *3 Biotech*, 5: 337–353
- Kumar V, Dangi AK, Shukla P (2018). Engineering Thermostable Microbial Xylanases Toward its Industrial Applications. *Mol Biotechnol*, 60(3): 226-235.
- Lamed R, Setter E, Bayer E (1983). Characterization of a cellulose-binding, cellulose-containing complex in *Clostridium thermocellum*. *J Bacteriol*, 156: 828–836.
- Lee CM, Weon HY, Hong SB, Jeon YA, Schumann P, Kroppenstedt RM, Kwon SW, Stackebrandt E (2008). *Cellulomonas aerilata* sp. nov., isolated from an air sample. *Int J Syst Evol Microbiol*, 58(Pt 12): 2925-2929.
- Lennartsson PR, Erlandsson P, Taherzadeh MJ (2014). Integration of the first and second generation bioethanol processes and the importance of by-products. *Bioresour Technol*, 165: 3-8.
- Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B (2013). Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. *Biotechnol Biofuels*, 6(1): 41.
- Li LL, McCorkle SR, Monchy S, Taghavi S, van der Lelie D (2009). Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnol Biofuels*. 2:10
- Linares-Pastén JA, Aronsson A, Karlsson EN (2018). Structural Considerations on the Use of Endo-Xylanases for the Production of prebiotic Xylooligosaccharides from Biomass. *Curr Protein Pept Sci* 19(1); 48–67.

Lisov AV, Belova OV, Lisova ZA, Vinokurova NG, Nagel AS, Andreeva-Kovalevskaya ZI, Budarina ZI, Nagornykh MO, Zakharova MV, Shadrin AM, Solonin AS, Leontievsky AA. (2017). Xylanases of *Cellulomonas flavigena*: expression, biochemical characterization, and biotechnological potential. *AMB Express*, 7(1): 5.

Loix C, Huybrechts M, Vangronsveld J, Gielen M, Keunen E, Cuypers A (2017). Reciprocal Interactions between Cadmium-Induced Cell Wall Responses and Oxidative Stress in Plants. *Front Plant Sci*, 8: 1867.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*, 42: D490-495.

Lundell TK, Mäkelä MR, Hildén K (2010). Lignin-modifying enzymes in filamentous basidiomycetes – ecological, functional and phylogenetic review. *Basic Microbiol*, 50: 5–20.

Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002). Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol Rev*, 66(3): 506–577.

**M**ahato NK, Gupta V, Singh P, y col. (2017). Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. *Antonie Van Leeuwenhoek*, 110(10):1357-1371

Manzanares P (2010). Integrated hydrolysis, fermentation and co-fermentation of lignocellulosic biomass. En: *Bioalcohol Production: Biochemical Conversion of Lignocellulosic Biomass* (Ed. Waldron KW), Woodhead Publishing Series in Energy, pp. 205-23.

Martani F, Beltrametti F, Porro D, Branduardi P, Lotti M (2017). The importance of fermentative conditions for the biotechnological production of lignin modifying enzymes from white-rot fungi. *Microbiol Lett*, 364: 13.

Martínez AT, Speranza M, Ruiz-Dueñas FJ, Ferreira P, Camarero S, Guillén F, Martínez MJ, Gutiérrez A, del Río JC (2005). Biodegradation of lignocellulosics: microbial, chemical, and enzymatic aspects of the fungal attack of lignin. *Int Microbiol*, 8(3): 195-204.

Mayorga-Reyes L, Morales Y, Salgado LM, Ortega A, Ponce-Noyola T (2002). *Cellulomonas flavigena*: Characterization of an endo-1,4-xylanase tightly induced by sugarcane bagasse. *FEMS Microbiol Lett* 214: 205–9.

Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics*, 14: 60.

Meinke A, Gilkes NR, Kwan E, Kilburn DG, Warren RAJ y Miller RC (1994). Cellobiohydrolase A (CbhA) from the cellulolytic bacterium *Cellulomonas fimi* is a beta-1,4-exocellobiohydrolase analogous to *Trichoderma reesei* CBH II. *Mol Microbiol*, 12: 413-22.

Mendoza GD, Loera Corral O, Plata Pérez FX, Hernández García PA, Ramírez Mella M (2014). Considerations on the use of exogenous fibrolytic enzymes to improve forage utilization. *Scientific World Journal*, 2014: 247437.

Miller GL (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Analytical Chemistry*, 31: 426-8

Mohnen D (2008). Pectin structure and biosynthesis. *Curr Opin Plant Biol*, 11(3): 266-77.

Morgenstern I, Powlowski J, Tsang A (2014). Fungal cellulose degradation by oxidative enzymes: from dysfunctional GH61 family to powerful lytic polysaccharide monoxygenase family. *Brief Funct Genomics*, 13(6): 471-81.

Mullis KB (1990). The unusual origin of the polymerase chain reaction. *Scientific American*, 262(4): 56-61, 64-5.

**N**CBIB Resource Coordinators (2015). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*, 44(D1): D7-19.

**P**agès S, Bélaïch A, Bélaïch JP, Morag E, Lamed R, Shoham Y, Bayer EA (1997). Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. *Proteins*, 29(4): 517-27.

Parte AC (2018). LPSN - List of Prokaryotic names with Standing in Nomenclature (bacterio.net), 20 years on. *Int J Syst Evol Microbiol*, 68(6): 1825-9.

Pauly M, Gille S, Liu L, Mansoori N, de Souza A, Schultink A, Xiong G (2013). Hemicellulose biosynthesis. *Planta*, 238(4): 627-42.

Payne CM, Knott BC, Mayes HB, Hansson H, Himmel ME, Sandgren M, Ståhlberg J, Beckham GT (2015). Fungal cellulases. *Chem Rev*, 115(3): 1308-448.

Petersen TN, Brunak S, von Heijne G, Nielsen H (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*, 8: 785-6.

Poulsen HV, Willink FW, Ingvorsen K (2016). Aerobic and anaerobic cellulase production by *Cellulomonas uda*. *Archives of microbiology*, 198(8): 725-35.

**Q**uinlan RJ, Teter S, Xu F (2010). Development of cellulases to improve enzymatic hydrolysis of lignocellulosic biomass. En: *Bioalcohol Production: Biochemical Conversion of Lignocellulosic Biomass* (Ed. Waldron KW), Woodhead Publishing Series in Energy, pp. 178-204.

**R**eguera G, Leschine SB (2001). Chitin degradation by cellulolytic anaerobes and facultative aerobes from soils and sediments. *FEMS Microbiol Lett*, (2): 367-74.

Reguera G, Leschine SB (2003). Biochemical and genetic characterization of ChiA, the major enzyme component for the solubilization of chitin by *Cellulomonas uda*. *Arch Microbiol*, 180: 434-43.

Rennie EA, Scheller HV (2014). Xylan biosynthesis. *Curr Opin Biotechnol*, 26: 100-7

Rivas R, Trujillo ME, Mateos PF, Martínez-Molina E, Velázquez E (2004). *Cellulomonas xylanilytica* sp. nov., a cellulolytic and xylanolytic bacterium isolated from a decayed elm tree. *Int J Syst Evol Microbiol*, 54(2): 533-6.

Robak K, Balcerek M (2018). Review of Second Generation Bioethanol Production from Residual Biomass. *Food Technol Biotechnol*, 56(2): 174-87.

Rodrigues Mano MC, Neri-Numa IA, da Silva JB, Paulino BN, Pessoa MG, Pastore GM (2018). Oligosaccharide biotechnology: an approach of prebiotic revolution on the industry. *Appl Microbiol Biotechnol*, 102(1): 17-37.

Rojas-Rejón ÓA, Poggi-Valardo HM, Ramos-Valdivia AC, Ponce-Noyola T, Cristiani-Urbina E, Martínez A, de la Torre M. (2016) Enzymatic saccharification of sugar cane bagasse by continuous xylanase and cellulase production from *Cellulomonas flavigena* PR-22. *Biotechnol Prog*, 32(2): 321-6.

Roy A, Kucukural A, Zhang Y (2011). I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc*, 5: 725–38.

Rusznayk A, Tóth EM, Schumann P, Spröer C, Makk J, Szabó G, Vladár P, Márialigeti K, Borsodi AK (2011). *Cellulomonas phragmiteti* sp. nov., a cellulolytic bacterium isolated from reed (*Phragmites australis*) periphyton in a shallow soda pond. *Int J Syst Evol Microbiol*, 61(7): 1662-6.

**S**abarís di Lorenzo GJ (2013). Prospección, aislamiento y caracterización de bacterias celulolíticas de suelo de bosque nativo de Misiones, Argentina. Tesis de Licenciatura en Cs. Biológicas. FCEN. UBA.

Saini A, Aggarwal NK, Sharma Ayadav A (2015). *Actinomycetes: A Source of Lignocellulolytic Enzymes*. *Enzyme Res*, 2015: 279381.

Saitou N, Nei M (1987). The Neighbor-Joining Method-a New Method for Reconstructing Phylogenetic Trees. *Mol Biol Evol*, 4: 406–25.

Salgado JCS, Meleiro LP, Carli S, Ward RJ (2018). Glucose tolerant and glucose stimulated  $\beta$ -glucosidases - A review. *Bioresour Technol*, 267: 704-13.

Sambrook J, Fritsch EF, Maniatis T (1989). *Molecular cloning: a laboratory manual*, 2da ed., Cold Spring Harbor, New York.

Sammond DW, Payne CM, Brunecky R, Himmel ME, Crowley MF, Beckham GT (2012). Cellulase linkers are optimized based on domain type and function: insights from sequence analysis, biophysical measurements, and molecular simulation. *PLoS One*, 7(11): e48615.

Sasser M (1990). Identification of bacteria through fatty acid analysis. En: *Methods in Phytobacteriology*. Clement Z, Rudolph K, Sands D.C. (Eds). Akadémiai Kiadó, Budapest, pp. 199-204.

Scheller HV, Ulvskov P (2010). Hemicelluloses. *Annu Rev Plant Biol*, 61: 263–89.

Shallom D, Shoham Y (2003). Microbial hemicellulases. *Curr Opin Microbiol*, 6: 219–28.

Sharma A, Tewari R, Rana SS, Soni R, Soni SK (2016). Cellulases: Classification, Methods of Determination and Industrial Applications. *Appl Biochem Biotechnol*, 179(8): 1346-80.

Shi Z, Luo G, Wang G (2012). *Cellulomonas carbonis* sp. nov., isolated from coal mine soil. *Int J Syst Evol Microbiol*, 62(8): 2004-10.

Shinoda K, Tomita M, Ishihama Y (2010) emPAI Calc—for the estimation of protein abundance from largescale identification data by liquid chromatography-tandem mass spectrometry. *Bioinformatics* 26: 576–7.

Singh S, Singh G, Arya SK (2018). Mannans: An overview of properties and application in food products. *Int J Biol Macromol*, 119: 79-95.

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, Crocker D (2008). Determination of Structural Carbohydrates and Lignin in Biomass. Procedimiento analítico de laboratorio (LAP) de National Renewable Energy Laboratory, U.S. Department of Energy, Office of Energy Efficiency & Renewable Energy.

Smith PJ, Wang HT, York WS, Peña MJ, Urbanowicz BR (2017). Designer biomass for next-generation biorefineries: leveraging recent insights into xylan structure and biosynthesis. *Biotechnol Biofuels*, 10: 286.

Solovyev V, Salamov A (2011). Automatic Annotation of Microbial Genomes and Metagenomic Sequences. En: *Metagenomics and its Applications in Agriculture, Biomedicine and Environmental Studies* (Ed. Li RW), Nova Science Publishers, pp. 61-78.

Spertino S, Boatti L, Icardi S, Manfredi M, Cattaneo C, Marengo E, Cavaletto M (2018). *Cellulomonas fimi* secretomes : In vivo and in silico approaches for the lignocellulose bioconversion. *J Biotechnol*, 270: 21–9.

Spurway TD, Morland C, Cooper A, Sumner I, Hazlewood GP, O'Donnell AG, Pickersgill RW, Gilbert HJ (1997). Calcium protects a mesophilic xylanase from proteinase inactivation and thermal unfolding. *J Biol Chem*, 272(28): 17523-30.

Stackebrandt E, Kandler O (1979). Taxonomy of the genus *Cellulomonas*, based on phenotypic characters and deoxyribonucleic acid-deoxyribonucleic acid homology, and proposal of seven neotype strains. *Int J Syst Bacteriol*, 29(4): 273-82.

Stackebrandt E, Schumann P (2015) *Cellulomonas*. En: *Bergey's Manual of Systematics of Archaea and Bacteria* (Eds. Whitman WB, Rainey F, Kampfer P, Trujillo M, Chun J, DeVos P, Hedlund B, Dedysh S, Hoboken NJ), John Wiley & Sons Inc, pp. 1–14.

**T**alamantes D, Biabini N, Dang H, Abdoun K, Berlemont R (2016). Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol Biofuels*, 9: 133.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol*, 30(12): 2725–9.

Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res*, 44(14): 6614-24.

Teather RM, Wood PJ (1982). Use of Congo red-polysaccharide interactions in enumeration and characterization of cellulolytic bacteria from the bovine rumen. *App Environ Microbiol*, 43: 777-80.

Thompson CC, Chimetto L, Edwards RA, Swings J, Stackebrandt E, Thompson FL (2013) Microbial genomic taxonomy. *BMC Genom*, 14: 931.

Toushik SH, Lee KT, Lee JS, Kim KS (2017). Functional Applications of Lignocellulolytic Enzymes in the Fruit and Vegetable Processing Industries. *J Food Sci*, 82(3): 585-93.

**V**anholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010). Lignin Biosynthesis and Structure. *Plant Physiol*, 153(3): 895–905.

Vlasenko E, Schüle M, Cherry J, Xu F (2010). Substrate specificity of family 5, 6, 7, 9, 12, and 45 endoglucanases. *Bioresour Technol*, 101(7): 2405-11.

**W**akarchuk WW, Brochu D, Foote S, Robotham A, Saxena H, Erak T, Kelly J (2016). Proteomic analysis of the secretome of *Cellulomonas fimi* ATCC 484 and *Cellulomonas flavigena* ATCC 482. *PLoS One*, 11(3): e0151186.

Walia A, Guleria S, Mehta P, Chauhan A, Parkash J (2017). Microbial xylanases and their industrial application in pulp and paper biobleaching: a review. *3 Biotech*, 7(1): 11.

White BA, Lamed R, Bayer EA, Flint HJ (2014). Biomass utilization by gut microbiomes. *Annu Rev Microbiol*, 68: 279-96.

Wyman CE (1994). Ethanol from lignocellulosic biomass: Technology, economics, and opportunities. *Bioresour Technol*, 50: 3-15.

**X**i W, Song D, Sun J, Shen J, Li L (2017). Formation of wood secondary cell wall may involve two type cellulose synthase complexes in *Populus*. *Plant Mol Biol*, 93(4-5): 419-29.

**Y**ang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015). The I-TASSER Suite: Protein structure and function prediction. *Nature Methods*, 12: 7-8.

Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y (2012). DbCAN: A web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res*, 40: 445–51.

**Z**amil MS, Geitmann A (2017). The middle lamella-more than a glue. *Phys Biol*, 14(1): 015004.

Zhang L, Xi L, Qiu D, Song L, Dai X, Ruan J, Huang Y (2013). *Cellulomonas marina* sp. nov., isolated from deep-sea water. *Int J Syst Evol Microbiol*, 63(8): 3014-18.

Zhang YHP, Himmel ME, Mielenz JR (2006). Outlook for cellulase improvement: screening and selection strategies. *Biotechnol Adv*, 24(5): 452-81.

# **Anexo**

Proteínas detectadas por espectrometría de masas en el sobrenadante de cultivo en CMC (muestras CMC01 y CMC02), en RAC (muestras RAC01 y RAC02), en PTE (muestras PTE01 y PTE02) y sacarosa (muestras SAC01 y SAC02).

CMC01			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAI %
KSW23552.1	A0A0V8ST89	2,332	10,99
KSW15663.1	A0A0V8S5R6	4,38	20,64
KSW19409.1	A0A0V8SGG3	7,733	36,43
KSW29460.1	A0A0V8T9W0	1,754	8,26
KSW29589.1	A0A0V8TAZ1	1,637	7,71
KSW15916.1	A0A0V8S6F9	0,931	4,39
KSW29871.1	A0A0V8TBE7	0,445	2,1
KSW16982.1	A0A0V8S9L5	0,174	0,82
KSW21866.1	A0A0V8SP17	0,259	1,22
KSW16352.1	A0A0V8S7Q7	0,413	1,95
KSW30161.1	A0A0V8TCM1	0,318	1,5
KSW28998.1	A0A0V8T9B1	0,101	0,48
KSW21671.1	A0A0V8SMW4	0,089	0,42
KSW29514.1	A0A0V8TBP7	0,099	0,47
KSW29732.1	A0A0V8TAN9	0,023	0,11
KSW21563.1	A0A0V8SMI4	0,062	0,29
KSW12785.1	A0A0V8RXL6	0,062	0,29
KSW21897.1	A0A0V8SNI9	0,194	0,91
KSW21014.1	A0A0V8SL25	0,11	0,52
KSW23701.1	A0A0V8STW7	0,11	0,52
CMC02			
Acceso de GenBank	Accession	Acceso de Uniprot	emPAI %
KSW19409.1		A0A0V8SGG3	42,45
KSW15663.1		A0A0V8S5R6	16,21
KSW29589.1		A0A0V8TAZ1	12,28
KSW23552.1		A0A0V8ST89	8,11
KSW29460.1		A0A0V8T9W0	5,53
KSW15916.1		A0A0V8S6F9	3,82
KSW29966.1		A0A0V8TB53	3,1
KSW29131.1		A0A0V8T983	1,94
KSW14495.1		A0A0V8S2E5	1,59
KSW16352.1		A0A0V8S7Q7	0,91
KSW21866.1		A0A0V8SP17	0,91
KSW16982.1		A0A0V8S9L5	0,85
KSW21014.1		A0A0V8SL25	0,82
KSW28998.1		A0A0V8T9B1	0,76

KSW30161.1	A0A0V8TCM1	0,1	0,72
KSW16358.1	A0A0V8S7R0	0	0
KSW18290.1	A0A0V8SDT9	0	0
KSW19923.1	A0A0V8SHY4	0	0
RAC01			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAI %
KSW12785.1	A0A0V8RXL6	0,27	0,13
KSW13734.1	A0A0V8S074	0,15	0,07
KSW13731.1	A0A0V8S079	0,66	0,30
KSW13763.1	A0A0V8S0B6	0,4	0,19
KSW13742.1	A0A0V8S0G4	0,17	0,08
KSW13758.1	A0A0V8S0I0	0,97	0,45
KSW13738.1	A0A0V8S0K7	0,59	0,27
KSW13730.1	A0A0V8S0N0	0,1	0,04
KSW14099.1	A0A0V8S172	1,15	0,53
KSW14008.1	A0A0V8S182	0,13	0,06
KSW14007.1	A0A0V8S185	2,16	1,00
KSW14109.1	A0A0V8S193	0,05	0,02
KSW14024.1	A0A0V8S195	0,1	0,05
KSW14247.1	A0A0V8S1S2	0,39	0,18
KSW14470.1	A0A0V8S2B2	0,26	0,12
KSW14624.1	A0A0V8S2R5	0,17	0,08
KSW14704.1	A0A0V8S2Z3	1,15	0,53
KSW14798.1	A0A0V8S3A4	0,23	0,11
KSW14744.1	A0A0V8S3C7	0,12	0,06
KSW15001.1	A0A0V8S3V5	0,08	0,03
KSW15158.1	A0A0V8S487	0,78	0,36
KSW15161.1	A0A0V8S4B6	0,47	0,22
KSW15169.1	A0A0V8S4D9	0,15	0,07
KSW15162.1	A0A0V8S4M7	0,12	0,05
KSW15286.1	A0A0V8S4N4	0,31	0,14
KSW15294.1	A0A0V8S4Q8	0,27	0,13
KSW15313.1	A0A0V8S4S0	0,08	0,04
KSW15315.1	A0A0V8S4T7	1,25	0,58
KSW15288.1	A0A0V8S4U1	0,13	0,06
KSW15433.1	A0A0V8S563	0,33	0,15
KSW15434.1	A0A0V8S598	0,22	0,10
KSW15553.1	A0A0V8S5G6	0,83	0,38
KSW15558.1	A0A0V8S5H0	0,11	0,05
KSW15560.1	A0A0V8S5I5	0,06	0,03
KSW15562.1	A0A0V8S5L9	0,18	0,08
KSW15554.1	A0A0V8S5Q8	0,1	0,04
KSW15663.1	A0A0V8S5R6	0,19	0,09
KSW15610.1	A0A0V8S5S9	0,59	0,27

KSW15696.1	A0A0V8S5W2	0,42	0,19
KSW15916.1	A0A0V8S6F9	0,28	0,13
KSW15930.1	A0A0V8S6Q4	0,39	0,18
KSW16070.1	A0A0V8S6Y4	0,1	0,04
KSW16111.1	A0A0V8S708	0,44	0,20
KSW16071.1	A0A0V8S775	0,19	0,09
KSW16202.1	A0A0V8S7G9	0,12	0,05
KSW16280.1	A0A0V8S7K3	0,33	0,15
KSW16352.1	A0A0V8S7Q7	3,47	1,60
KSW16351.1	A0A0V8S7R6	0,73	0,34
KSW16349.1	A0A0V8S7T2	0,84	0,39
KSW16984.1	A0A0V8S9J4	0,93	0,43
KSW16982.1	A0A0V8S9L5	0,8	0,37
KSW17032.1	A0A0V8S9N8	0,16	0,07
KSW17192.1	A0A0V8SA53	0,59	0,27
KSW17208.1	A0A0V8SA68	1,51	0,70
KSW17209.1	A0A0V8SA74	0,93	0,43
KSW17322.1	A0A0V8SAR9	0,39	0,18
KSW17684.1	A0A0V8SBK5	1,05	0,49
KSW17696.1	A0A0V8SBK8	1,37	0,63
KSW17692.1	A0A0V8SBL0	0,33	0,15
KSW17688.1	A0A0V8SBL5	0,59	0,27
KSW17699.1	A0A0V8SBL8	1,78	0,82
KSW17957.1	A0A0V8SCF1	0,08	0,04
KSW18201.1	A0A0V8SD03	0,19	0,09
KSW18203.1	A0A0V8SD15	0,32	0,15
KSW18199.1	A0A0V8SD16	0,52	0,24
KSW18276.1	A0A0V8SDC1	0,39	0,18
KSW18330.1	A0A0V8SDQ3	1,33	0,61
KSW18495.1	A0A0V8SDU2	0,26	0,12
KSW18322.1	A0A0V8SDX3	0,39	0,18
KSW18565.1	A0A0V8SE56	0,39	0,18
KSW18506.1	A0A0V8SE91	0,12	0,05
KSW18655.1	A0A0V8SE95	0,33	0,15
KSW19086.1	A0A0V8SFI7	0,21	0,10
KSW19094.1	A0A0V8SFK9	0,36	0,17
KSW19115.1	A0A0V8SFL7	0,78	0,36
KSW19404.1	A0A0V8SGF9	0,17	0,08
KSW19409.1	A0A0V8SGG3	10,45	4,83
KSW19405.1	A0A0V8SGG7	0,09	0,04
KSW19439.1	A0A0V8SGH6	0,27	0,13
KSW19441.1	A0A0V8SGK6	0,26	0,12
KSW19469.1	A0A0V8SGN7	0,06	0,03
KSW19596.1	A0A0V8SH09	0,32	0,15
KSW19597.1	A0A0V8SH13	0,67	0,31

KSW19791.1	A0A0V8SHI8	4,93	2,28
KSW19920.1	A0A0V8SHZ5	0,3	0,14
KSW19794.1	A0A0V8SI61	0,18	0,08
KSW19866.1	A0A0V8SIB5	0,19	0,09
KSW20219.1	A0A0V8SIS6	0,26	0,12
KSW20550.1	A0A0V8SJQ2	1,28	0,59
KSW20567.1	A0A0V8SJU5	0,41	0,19
KSW20684.1	A0A0V8SK26	0,78	0,36
KSW20683.1	A0A0V8SK28	0,39	0,18
KSW20681.1	A0A0V8SK30	0,47	0,22
KSW20692.1	A0A0V8SK31	4,34	2,00
KSW20688.1	A0A0V8SK35	0,12	0,06
KSW20673.1	A0A0V8SK40	0,52	0,24
KSW20704.1	A0A0V8SK41	1,96	0,90
KSW20697.1	A0A0V8SK42	0,87	0,40
KSW20700.1	A0A0V8SK48	0,93	0,43
KSW20691.1	A0A0V8SK50	2,16	1,00
KSW20689.1	A0A0V8SK55	0,78	0,36
KSW20702.1	A0A0V8SK56	1,15	0,53
KSW20713.1	A0A0V8SK59	0,14	0,06
KSW20707.1	A0A0V8SK66	4,62	2,14
KSW20710.1	A0A0V8SK68	0,87	0,40
KSW20701.1	A0A0V8SK70	0,15	0,07
KSW20705.1	A0A0V8SK71	1,37	0,63
KSW20696.1	A0A0V8SK72	1,15	0,53
KSW20695.1	A0A0V8SK73	0,78	0,36
KSW20736.1	A0A0V8SK77	1,31	0,61
KSW20703.1	A0A0V8SK78	1,51	0,70
KSW20726.1	A0A0V8SK84	0,25	0,11
KSW20721.1	A0A0V8SK87	0,26	0,12
KSW20711.1	A0A0V8SK88	0,47	0,22
KSW20719.1	A0A0V8SK91	0,12	0,06
KSW20712.1	A0A0V8SK93	0,14	0,06
KSW20693.1	A0A0V8SKA2	0,29	0,13
KSW20735.1	A0A0V8SKA8	1,37	0,63
KSW20824.1	A0A0V8SKK2	0,05	0,02
KSW20674.1	A0A0V8SKL6	3,64	1,68
KSW20682.1	A0A0V8SKL8	0,55	0,25
KSW20698.1	A0A0V8SKN1	1,15	0,53
KSW20690.1	A0A0V8SKN7	0,54	0,25
KSW20706.1	A0A0V8SKQ6	0,36	0,17
KSW20977.1	A0A0V8SKZ6	0,17	0,08
KSW20990.1	A0A0V8SL04	0,29	0,13
KSW21020.1	A0A0V8SL08	0,4	0,18
KSW21040.1	A0A0V8SL26	0,33	0,15

KSW21039.1	A0A0V8SL49	0,5	0,23
KSW21018.1	A0A0V8SLK0	0,09	0,04
KSW21230.1	A0A0V8SLN9	0,69	0,32
KSW21273.1	A0A0V8SLQ2	0,47	0,22
KSW21229.1	A0A0V8SLR7	0,64	0,29
KSW21271.1	A0A0V8SLT1	0,67	0,31
KSW21272.1	A0A0V8SLX2	2,16	1,00
KSW21449.1	A0A0V8SM91	0,23	0,11
KSW21460.1	A0A0V8SMA2	0,11	0,05
KSW21274.1	A0A0V8SMB6	0,35	0,16
KSW21468.1	A0A0V8SMC2	0,06	0,03
KSW21470.1	A0A0V8SMC4	0,91	0,42
KSW21507.1	A0A0V8SMF4	0,39	0,18
KSW21529.1	A0A0V8SMI0	0,21	0,10
KSW21527.1	A0A0V8SMJ2	0,38	0,17
KSW21671.1	A0A0V8SMW4	0,09	0,04
KSW21678.1	A0A0V8SMY7	0,06	0,03
KSW21735.1	A0A0V8SN47	1,51	0,70
KSW21774.1	A0A0V8SN80	0,97	0,45
KSW21825.1	A0A0V8SNB8	0,18	0,09
KSW21865.1	A0A0V8SNF7	1,99	0,92
KSW21870.1	A0A0V8SNH7	0,33	0,15
KSW21840.1	A0A0V8SNH9	0,33	0,15
KSW21884.1	A0A0V8SNI3	1,15	0,53
KSW21897.1	A0A0V8SNI9	0,43	0,20
KSW21936.1	A0A0V8SNR7	0,39	0,18
KSW21893.1	A0A0V8SNT8	0,23	0,11
KSW21834.1	A0A0V8SNX9	0,41	0,19
KSW21866.1	A0A0V8SP17	7,91	3,66
KSW23491.1	A0A0V8ST30	0,64	0,30
KSW23552.1	A0A0V8ST89	2	0,92
KSW23701.1	A0A0V8STW7	0,87	0,40
KSW23777.1	A0A0V8STZ4	0,25	0,11
KSW23794.1	A0A0V8SU61	0,27	0,13
KSW28201.1	A0A0V8T686	0,47	0,22
KSW28270.1	A0A0V8T6M5	0,39	0,18
KSW28306.1	A0A0V8T6P6	0,78	0,36
KSW28272.1	A0A0V8T6T2	0,12	0,05
KSW28366.1	A0A0V8T6V5	0,11	0,05
KSW28328.1	A0A0V8T700	0,59	0,27
KSW28373.1	A0A0V8T737	0,37	0,17
KSW28305.1	A0A0V8T764	2,98	1,38
KSW28601.1	A0A0V8T771	1,91	0,88
KSW28372.1	A0A0V8T772	0,89	0,41
KSW28317.1	A0A0V8T7C1	0,33	0,15

KSW28310.1	A0A0V8T7E0	0,89	0,41
KSW28551.1	A0A0V8T7E8	0,29	0,13
KSW28654.1	A0A0V8T7H0	0,13	0,06
KSW28611.1	A0A0V8T7H6	0,32	0,15
KSW28627.1	A0A0V8T7J2	0,5	0,23
KSW28552.1	A0A0V8T7K3	0,78	0,36
KSW28560.1	A0A0V8T7L9	0,16	0,07
KSW28374.1	A0A0V8T7N1	0,07	0,03
KSW28600.1	A0A0V8T7P0	0,21	0,10
KSW28556.1	A0A0V8T7U2	0,31	0,14
KSW28591.1	A0A0V8T7Y8	1,29	0,60
KSW28775.1	A0A0V8T812	0,04	0,02
KSW29001.1	A0A0V8T881	12,9	5,96
KSW28630.1	A0A0V8T8A7	0,39	0,18
KSW28808.1	A0A0V8T8D2	1,28	0,59
KSW28914.1	A0A0V8T8E8	0,12	0,06
KSW28963.1	A0A0V8T8K6	0,32	0,15
KSW29038.1	A0A0V8T8N1	0,14	0,06
KSW29054.1	A0A0V8T8P1	0,39	0,18
KSW28971.1	A0A0V8T8Q9	0,17	0,08
KSW28554.1	A0A0V8T8R8	0,17	0,08
KSW29118.1	A0A0V8T8U7	0,3	0,14
KSW28925.1	A0A0V8T8W5	0,3	0,14
KSW29166.1	A0A0V8T8X7	0,17	0,08
KSW29139.1	A0A0V8T903	0,13	0,06
KSW29108.1	A0A0V8T913	2,51	1,16
KSW28618.1	A0A0V8T919	2,16	1,00
KSW29230.1	A0A0V8T922	2,16	1,00
KSW29083.1	A0A0V8T944	0,06	0,03
KSW29204.1	A0A0V8T993	0,14	0,06
KSW29093.1	A0A0V8T997	0,36	0,17
KSW29109.1	A0A0V8T9A8	0,47	0,22
KSW29132.1	A0A0V8T9B3	0,52	0,24
KSW29326.1	A0A0V8T9B4	0,23	0,11
KSW29030.1	A0A0V8T9E2	0,09	0,04
KSW29215.1	A0A0V8T9G7	0,36	0,17
KSW29406.1	A0A0V8T9H8	1,15	0,53
KSW29200.1	A0A0V8T9I6	2,98	1,38
KSW29110.1	A0A0V8T9J8	1,83	0,85
KSW29244.1	A0A0V8T9K2	0,59	0,27
KSW29454.1	A0A0V8T9M1	0,68	0,32
KSW29213.1	A0A0V8T9P5	0,12	0,05
KSW29229.1	A0A0V8T9R1	0,33	0,15
KSW29261.1	A0A0V8T9T5	1,42	0,66
KSW29460.1	A0A0V8T9W0	3,79	1,75

KSW29330.1	A0A0V8T9W1	0,69	0,32
KSW29518.1	A0A0V8T9Z1	0,47	0,22
KSW29378.1	A0A0V8TA08	0,52	0,24
KSW29333.1	A0A0V8TA14	1	0,46
KSW29286.1	A0A0V8TA15	0,26	0,12
KSW29302.1	A0A0V8TA31	0,23	0,11
KSW29426.1	A0A0V8TA52	0,33	0,15
KSW29539.1	A0A0V8TA57	0,6	0,28
KSW29389.1	A0A0V8TA61	1,73	0,80
KSW29367.1	A0A0V8TA69	0,29	0,13
KSW29614.1	A0A0V8TA87	0,37	0,17
KSW29437.1	A0A0V8TAA0	0,14	0,06
KSW29082.1	A0A0V8TAB3	0,15	0,07
KSW29516.1	A0A0V8TAE3	4,78	2,21
KSW29710.1	A0A0V8TAE8	0,21	0,09
KSW29726.1	A0A0V8TAG7	0,25	0,11
KSW29602.1	A0A0V8TAJ5	1,15	0,53
KSW29700.1	A0A0V8TAK9	0,5	0,23
KSW29656.1	A0A0V8TAN1	0,5	0,23
KSW29732.1	A0A0V8TAN9	0,35	0,16
KSW29527.1	A0A0V8TAP7	0,41	0,19
KSW29822.1	A0A0V8TAQ2	0,78	0,36
KSW29704.1	A0A0V8TAS0	0,07	0,03
KSW29785.1	A0A0V8TAS3	0,26	0,12
KSW29817.1	A0A0V8TAW2	0,13	0,06
KSW29589.1	A0A0V8TAZ1	3,28	1,52
KSW29696.1	A0A0V8TB09	0,09	0,04
KSW29892.1	A0A0V8TB43	1,89	0,87
KSW29966.1	A0A0V8TB53	0,12	0,06
KSW29998.1	A0A0V8TB93	0,47	0,22
KSW29842.1	A0A0V8TBB6	0,12	0,06
KSW29884.1	A0A0V8TBC4	0,87	0,40
KSW29871.1	A0A0V8TBE7	5,92	2,73
KSW29887.1	A0A0V8TBG6	0,87	0,40
KSW29903.1	A0A0V8TBH7	0,47	0,22
KSW30105.1	A0A0V8TBK0	0,57	0,26
KSW29514.1	A0A0V8TBP7	0,27	0,12
KSW29904.1	A0A0V8TBP8	0,26	0,12
KSW30107.1	A0A0V8TBQ0	0,19	0,09
KSW29937.1	A0A0V8TBS6	0,16	0,07
KSW30155.1	A0A0V8TBT6	0,13	0,06
KSW30163.1	A0A0V8TBX1	0,07	0,03
KSW29942.1	A0A0V8TBY6	0,19	0,09
KSW29985.1	A0A0V8TBZ8	0,19	0,09
KSW30136.1	A0A0V8TC25	0,08	0,04

KSW29974.1	A0A0V8TC37	0,78	0,36
KSW30184.1	A0A0V8TC60	0,04	0,02
KSW30200.1	A0A0V8TC72	0,05	0,02
KSW30037.1	A0A0V8TC78	0,62	0,28
KSW30141.1	A0A0V8TC80	0,13	0,06
KSW30054.1	A0A0V8TC91	0,16	0,07
KSW30162.1	A0A0V8TC98	0,33	0,15
KSW30135.1	A0A0V8TCB0	0,24	0,11
KSW30160.1	A0A0V8TCD5	0,11	0,05
KSW30101.1	A0A0V8TCD6	1,13	0,52
KSW30102.1	A0A0V8TCD8	0,21	0,10
KSW30081.1	A0A0V8TCF3	0,17	0,08
KSW29834.1	A0A0V8TCL7	0,33	0,15
KSW30256.1	A0A0V8TCL8	0,17	0,08
KSW30161.1	A0A0V8TCM1	1,09	0,50
KSW30268.1	A0A0V8TCM5	0,52	0,24
KSW30322.1	A0A0V8TCP4	0,62	0,29
KSW29946.1	A0A0V8TCV4	0,78	0,36
KSW30273.1	A0A0V8TD39	0,23	0,10
KSW30106.1	A0A0V8TDF3	0,47	0,22
RAC02			
GenBank Accession	Accession UNIPROT	emPAI	emPAI (%)
KSW19791.1	A0A0V8SHI8	17,74	5,4
KSW29001.1	A0A0V8T881	12,9	3,92
KSW19409.1	A0A0V8SGG3	9	2,74
KSW20704.1	A0A0V8SK41	7,73	2,35
KSW29871.1	A0A0V8TBE7	7,32	2,23
KSW20674.1	A0A0V8SKL6	6,74	2,05
KSW17208.1	A0A0V8SA68	5,31	1,62
KSW21866.1	A0A0V8SP17	5,31	1,62
KSW28305.1	A0A0V8T764	5,31	1,62
KSW29516.1	A0A0V8TAE3	4,78	1,45
KSW20692.1	A0A0V8SK31	4,34	1,32
KSW29460.1	A0A0V8T9W0	4,25	1,29
KSW16352.1	A0A0V8S7Q7	4,01	1,22
KSW20696.1	A0A0V8SK72	3,64	1,11
KSW20698.1	A0A0V8SKN1	3,64	1,11
KSW21865.1	A0A0V8SNF7	3,64	1,11
KSW28729.1	A0A0V8T7H1	3,64	1,11
KSW20707.1	A0A0V8SK66	3,22	0,98
KSW28618.1	A0A0V8T919	3,22	0,98
KSW29892.1	A0A0V8TB43	3,13	0,95
KSW30161.1	A0A0V8TCM1	2,98	0,91
KSW29110.1	A0A0V8T9J8	2,94	0,89

KSW23552.1	A0A0V8ST89	2,9	0,88
KSW28372.1	A0A0V8T772	2,86	0,87
KSW20681.1	A0A0V8SK30	2,83	0,86
KSW29589.1	A0A0V8TAZ1	2,79	0,85
KSW29822.1	A0A0V8TAQ2	2,65	0,81
KSW17699.1	A0A0V8SBL8	2,59	0,79
KSW29884.1	A0A0V8TBC4	2,51	0,76
KSW28591.1	A0A0V8T7Y8	2,31	0,7
KSW29108.1	A0A0V8T913	2,28	0,69
KSW17688.1	A0A0V8SBL5	2,16	0,66
KSW20684.1	A0A0V8SK26	2,16	0,66
KSW20691.1	A0A0V8SK50	2,16	0,66
KSW20689.1	A0A0V8SK55	2,16	0,66
KSW20705.1	A0A0V8SK71	2,16	0,66
KSW20695.1	A0A0V8SK73	2,16	0,66
KSW20735.1	A0A0V8SKA8	2,16	0,66
KSW21735.1	A0A0V8SN47	2,16	0,66
KSW29200.1	A0A0V8T9I6	2,16	0,66
KSW29974.1	A0A0V8TC37	2,16	0,66
KSW30106.1	A0A0V8TDF3	2,16	0,66
KSW29333.1	A0A0V8TA14	1,99	0,6
KSW15315.1	A0A0V8S4T7	1,96	0,59
KSW28601.1	A0A0V8T771	1,91	0,58
KSW21897.1	A0A0V8SNI9	1,89	0,58
KSW20697.1	A0A0V8SK42	1,85	0,56
KSW20736.1	A0A0V8SK77	1,85	0,56
KSW21884.1	A0A0V8SNI3	1,78	0,54
KSW13731.1	A0A0V8S079	1,74	0,53
KSW16984.1	A0A0V8S9J4	1,68	0,51
KSW20700.1	A0A0V8SK48	1,68	0,51
KSW29389.1	A0A0V8TA61	1,57	0,48
KSW13738.1	A0A0V8S0K7	1,51	0,46
KSW20703.1	A0A0V8SK78	1,51	0,46
KSW21273.1	A0A0V8SLQ2	1,51	0,46
KSW29785.1	A0A0V8TAS3	1,51	0,46
KSW20550.1	A0A0V8SJQ2	1,47	0,45
KSW14007.1	A0A0V8S185	1,37	0,42
KSW16280.1	A0A0V8S7K3	1,37	0,42
KSW28306.1	A0A0V8T6P6	1,37	0,42
KSW29230.1	A0A0V8T922	1,37	0,42
KSW30162.1	A0A0V8TC98	1,37	0,42
KSW29834.1	A0A0V8TCL7	1,37	0,42
KSW19439.1	A0A0V8SGH6	1,34	0,41
KSW23794.1	A0A0V8SU61	1,34	0,41
KSW18330.1	A0A0V8SDQ3	1,33	0,4

KSW20710.1	A0A0V8SK68	1,31	0,4
KSW21230.1	A0A0V8SLN9	1,31	0,4
KSW21893.1	A0A0V8SNT8	1,31	0,4
KSW29887.1	A0A0V8TBG6	1,31	0,4
KSW28328.1	A0A0V8T700	1,29	0,39
KSW20701.1	A0A0V8SK70	1,25	0,38
KSW14704.1	A0A0V8S2Z3	1,15	0,35
KSW16351.1	A0A0V8S7R6	1,15	0,35
KSW17755.1	A0A0V8SBS4	1,15	0,35
KSW20702.1	A0A0V8SK56	1,15	0,35
KSW20682.1	A0A0V8SKL8	1,15	0,35
KSW20706.1	A0A0V8SKQ6	1,15	0,35
KSW21020.1	A0A0V8SL08	1,15	0,35
KSW21040.1	A0A0V8SL26	1,15	0,35
KSW21470.1	A0A0V8SMC4	1,15	0,35
KSW29406.1	A0A0V8T9H8	1,15	0,35
KSW29518.1	A0A0V8T9Z1	1,15	0,35
KSW29539.1	A0A0V8TA57	1,15	0,35
KSW29602.1	A0A0V8TAJ5	1,15	0,35
KSW30101.1	A0A0V8TCD6	1,13	0,34
KSW15553.1	A0A0V8S5G6	1,07	0,33
KSW17684.1	A0A0V8SBK5	1,05	0,32
KSW20690.1	A0A0V8SKN7	1,05	0,32
KSW29261.1	A0A0V8T9T5	1,03	0,31
KSW21274.1	A0A0V8SMB6	1,02	0,31
KSW15433.1	A0A0V8S563	1	0,3
KSW28914.1	A0A0V8T8E8	1	0,3
KSW28310.1	A0A0V8T7E0	0,99	0,3
KSW13758.1	A0A0V8S0I0	0,97	0,29
KSW21774.1	A0A0V8SN80	0,97	0,29
KSW23491.1	A0A0V8ST30	0,97	0,29
KSW29454.1	A0A0V8T9M1	0,95	0,29
KSW17209.1	A0A0V8SA74	0,93	0,28
KSW20683.1	A0A0V8SK28	0,93	0,28
KSW29054.1	A0A0V8T8P1	0,93	0,28
KSW16982.1	A0A0V8S9L5	0,9	0,27
KSW29204.1	A0A0V8T993	0,9	0,27
KSW23701.1	A0A0V8STW7	0,87	0,27
KSW29378.1	A0A0V8TA08	0,87	0,27
KSW30268.1	A0A0V8TCM5	0,87	0,27
KSW30107.1	A0A0V8TBQ0	0,86	0,26
KSW20412.1	A0A0V8SJB3	0,85	0,26
KSW28971.1	A0A0V8T8Q9	0,85	0,26
KSW16349.1	A0A0V8S7T2	0,84	0,26
KSW17201.1	A0A0V8SA59	0,78	0,24

KSW17696.1	A0A0V8SBK8	0,78	0,24
KSW18495.1	A0A0V8SDU2	0,78	0,24
KSW20567.1	A0A0V8SJU5	0,78	0,24
KSW21272.1	A0A0V8SLX2	0,78	0,24
KSW28552.1	A0A0V8T7K3	0,78	0,24
KSW29903.1	A0A0V8TBH7	0,78	0,24
KSW29946.1	A0A0V8TCV4	0,78	0,24
KSW15696.1	A0A0V8S5W2	0,75	0,23
KSW28611.1	A0A0V8T7H6	0,74	0,22
KSW29330.1	A0A0V8T9W1	0,74	0,22
KSW28925.1	A0A0V8T8W5	0,69	0,21
KSW16596.1	A0A0V8S8G3	0,67	0,2
KSW19597.1	A0A0V8SH13	0,67	0,2
KSW20693.1	A0A0V8SKA2	0,67	0,2
KSW28201.1	A0A0V8T686	0,67	0,2
KSW28551.1	A0A0V8T7E8	0,67	0,2
KSW30105.1	A0A0V8TBK0	0,66	0,2
KSW18322.1	A0A0V8SDX3	0,64	0,19
KSW28808.1	A0A0V8T8D2	0,64	0,19
KSW21529.1	A0A0V8SMI0	0,62	0,19
KSW30037.1	A0A0V8TC78	0,62	0,19
KSW15610.1	A0A0V8S5S9	0,59	0,18
KSW17192.1	A0A0V8SA53	0,59	0,18
KSW17198.1	A0A0V8SA55	0,59	0,18
KSW20219.1	A0A0V8SIS6	0,59	0,18
KSW20975.1	A0A0V8SKY5	0,59	0,18
KSW28554.1	A0A0V8T8R8	0,59	0,18
KSW29166.1	A0A0V8T8X7	0,59	0,18
KSW29093.1	A0A0V8T997	0,59	0,18
KSW29215.1	A0A0V8T9G7	0,59	0,18
KSW29286.1	A0A0V8TA15	0,59	0,18
KSW29500.1	A0A0V8TAC8	0,59	0,18
KSW30256.1	A0A0V8TCL8	0,59	0,18
KSW21527.1	A0A0V8SMJ2	0,57	0,17
KSW23777.1	A0A0V8STZ4	0,55	0,17
KSW28630.1	A0A0V8T8A7	0,55	0,17
KSW28560.1	A0A0V8T7L9	0,54	0,16
KSW30135.1	A0A0V8TCB0	0,54	0,16
KSW12785.1	A0A0V8RXL6	0,53	0,16
KSW29132.1	A0A0V8T9B3	0,52	0,16
KSW29614.1	A0A0V8TA87	0,52	0,16
KSW21936.1	A0A0V8SNR7	0,51	0,15
KSW29473.1	A0A0V8TAG8	0,5	0,15
KSW29656.1	A0A0V8TAN1	0,5	0,15
KSW17752.1	A0A0V8SBR8	0,49	0,15

KSW29118.1	A0A0V8T8U7	0,48	0,15
KSW18655.1	A0A0V8SE95	0,47	0,14
KSW28963.1	A0A0V8T8K6	0,47	0,14
KSW29998.1	A0A0V8TB93	0,47	0,14
KSW30102.1	A0A0V8TCD8	0,47	0,14
KSW14798.1	A0A0V8S3A4	0,44	0,13
KSW16111.1	A0A0V8S708	0,44	0,13
KSW29690.1	A0A0V8TC65	0,44	0,13
KSW30322.1	A0A0V8TCP4	0,44	0,13
KSW15001.1	A0A0V8S3V5	0,43	0,13
KSW13991.1	A0A0V8S169	0,43	0,13
KSW15663.1	A0A0V8S5R6	0,43	0,13
KSW19920.1	A0A0V8SHZ5	0,43	0,13
KSW21834.1	A0A0V8SNX9	0,41	0,13
KSW28579.1	A0A0V8T7E1	0,41	0,13
KSW29732.1	A0A0V8TAN9	0,41	0,13
KSW29966.1	A0A0V8TB53	0,41	0,13
KSW29696.1	A0A0V8TB09	0,41	0,12
KSW13763.1	A0A0V8S0B6	0,4	0,12
KSW15046.1	A0A0V8S3X8	0,39	0,12
KSW15066.1	A0A0V8S4D5	0,39	0,12
KSW15562.1	A0A0V8S5L9	0,39	0,12
KSW15930.1	A0A0V8S6Q4	0,39	0,12
KSW17322.1	A0A0V8SAR9	0,39	0,12
KSW18565.1	A0A0V8SE56	0,39	0,12
KSW20716.1	A0A0V8SK67	0,39	0,12
KSW20726.1	A0A0V8SK84	0,39	0,12
KSW21229.1	A0A0V8SLR7	0,39	0,12
KSW21507.1	A0A0V8SMF4	0,39	0,12
KSW28270.1	A0A0V8T6M5	0,39	0,12
KSW29213.1	A0A0V8T9P5	0,39	0,12
KSW29563.1	A0A0V8TAC3	0,39	0,12
KSW29989.1	A0A0V8TC41	0,39	0,12
KSW19583.1	A0A0V8SGX2	0,37	0,11
KSW29326.1	A0A0V8T9B4	0,37	0,11
KSW19094.1	A0A0V8SFK9	0,36	0,11
KSW15158.1	A0A0V8S487	0,33	0,1
KSW16601.1	A0A0V8S8F6	0,33	0,1
KSW17692.1	A0A0V8SBL0	0,33	0,1
KSW21875.1	A0A0V8SNG4	0,33	0,1
KSW21870.1	A0A0V8SNH7	0,33	0,1
KSW21840.1	A0A0V8SNH9	0,33	0,1
KSW28317.1	A0A0V8T7C1	0,33	0,1
KSW29229.1	A0A0V8T9R1	0,33	0,1
KSW29426.1	A0A0V8TA52	0,33	0,1

KSW29937.1	A0A0V8TBS6	0,33	0,1
KSW21872.1	A0A0V8SNK9	0,32	0,1
KSW19794.1	A0A0V8SI61	0,32	0,1
KSW15286.1	A0A0V8S4N4	0,31	0,09
KSW21039.1	A0A0V8SL49	0,31	0,09
KSW28627.1	A0A0V8T7J2	0,31	0,09
KSW29700.1	A0A0V8TAK9	0,31	0,09
KSW14099.1	A0A0V8S172	0,29	0,09
KSW18502.1	A0A0V8SDW4	0,29	0,09
KSW18504.1	A0A0V8SDX0	0,29	0,09
KSW19119.1	A0A0V8SFL6	0,29	0,09
KSW20191.1	A0A0V8SIQ9	0,29	0,09
KSW21018.1	A0A0V8SLK0	0,29	0,09
KSW21271.1	A0A0V8SLT1	0,29	0,09
KSW29437.1	A0A0V8TAA0	0,29	0,09
KSW15916.1	A0A0V8S6F9	0,28	0,09
KSW15294.1	A0A0V8S4Q8	0,27	0,08
KSW28654.1	A0A0V8T7H0	0,27	0,08
KSW28620.1	A0A0V8T7Z7	0,27	0,08
KSW30141.1	A0A0V8TC80	0,27	0,08
KSW13742.1	A0A0V8S0G4	0,27	0,08
KSW28373.1	A0A0V8T737	0,27	0,08
KSW14049.1	A0A0V8S142	0,26	0,08
KSW13992.1	A0A0V8S161	0,26	0,08
KSW19441.1	A0A0V8SGK6	0,26	0,08
KSW20721.1	A0A0V8SK87	0,26	0,08
KSW29244.1	A0A0V8T9K2	0,26	0,08
KSW29527.1	A0A0V8TAP7	0,26	0,08
KSW29825.1	A0A0V8TBH8	0,26	0,08
KSW30081.1	A0A0V8TCF3	0,26	0,08
KSW18203.1	A0A0V8SD15	0,25	0,07
KSW29224.1	A0A0V8T9J2	0,25	0,07
KSW15558.1	A0A0V8S5H0	0,23	0,07
KSW15604.1	A0A0V8S5M8	0,23	0,07
KSW18199.1	A0A0V8SD16	0,23	0,07
KSW18573.1	A0A0V8SE18	0,23	0,07
KSW20673.1	A0A0V8SK40	0,23	0,07
KSW29223.1	A0A0V8T9T1	0,23	0,07
KSW29302.1	A0A0V8TA31	0,23	0,07
KSW30273.1	A0A0V8TD39	0,23	0,07
KSW15434.1	A0A0V8S598	0,22	0,07
KSW16069.1	A0A0V8S6Z1	0,22	0,07
KSW20713.1	A0A0V8SK59	0,22	0,07
KSW19086.1	A0A0V8SFI7	0,21	0,06
KSW21775.1	A0A0V8SN78	0,21	0,06

KSW29345.1	A0A0V8TA73	0,21	0,06
KSW29710.1	A0A0V8TAE8	0,21	0,06
KSW19596.1	A0A0V8SH09	0,2	0,06
KSW20712.1	A0A0V8SK93	0,2	0,06
KSW16071.1	A0A0V8S775	0,19	0,06
KSW18201.1	A0A0V8SD03	0,19	0,06
KSW19587.1	A0A0V8SGZ8	0,19	0,06
KSW21568.1	A0A0V8SMM8	0,19	0,06
KSW29829.1	A0A0V8TBN1	0,19	0,06
KSW20490.1	A0A0V8SK36	0,19	0,06
KSW21671.1	A0A0V8SMW4	0,19	0,06
KSW29942.1	A0A0V8TBY6	0,19	0,06
KSW21825.1	A0A0V8SNB8	0,18	0,06
KSW18276.1	A0A0V8SDC1	0,18	0,05
KSW18489.1	A0A0V8SDT2	0,18	0,05
KSW18985.1	A0A0V8SF91	0,18	0,05
KSW29029.1	A0A0V8T941	0,18	0,05
KSW13736.1	A0A0V8S0F9	0,17	0,05
KSW14624.1	A0A0V8S2R5	0,17	0,05
KSW21276.1	A0A0V8SLR8	0,16	0,05
KSW23601.1	A0A0V8STG8	0,16	0,05
KSW28807.1	A0A0V8T843	0,16	0,05
KSW30082.1	A0A0V8TC27	0,16	0,05
KSW14024.1	A0A0V8S195	0,15	0,05
KSW29693.1	A0A0V8TAX2	0,15	0,05
KSW13734.1	A0A0V8S074	0,15	0,04
KSW15169.1	A0A0V8S4D9	0,15	0,04
KSW21831.1	A0A0V8SNE1	0,15	0,04
KSW28331.1	A0A0V8T6X0	0,15	0,04
KSW28321.1	A0A0V8T784	0,15	0,04
KSW29082.1	A0A0V8TAB3	0,15	0,04
KSW16591.1	A0A0V8S8G7	0,14	0,04
KSW20093.1	A0A0V8SIF4	0,14	0,04
KSW29038.1	A0A0V8T8N1	0,14	0,04
KSW29192.1	A0A0V8T9G4	0,14	0,04
KSW29503.1	A0A0V8TA84	0,14	0,04
KSW14008.1	A0A0V8S182	0,13	0,04
KSW18933.1	A0A0V8SF59	0,13	0,04
KSW19472.1	A0A0V8SGN6	0,13	0,04
KSW21669.1	A0A0V8SN74	0,13	0,04
KSW28775.1	A0A0V8T812	0,13	0,04
KSW14470.1	A0A0V8S2B2	0,12	0,04
KSW14744.1	A0A0V8S3C7	0,12	0,04
KSW15898.1	A0A0V8S6M3	0,12	0,04
KSW20688.1	A0A0V8SK35	0,12	0,04

KSW23476.1	A0A0V8ST04	0,12	0,04
KSW29094.1	A0A0V8T9I4	0,12	0,04
KSW29904.1	A0A0V8TBP8	0,12	0,04
KSW15162.1	A0A0V8S4M7	0,12	0,04
KSW16202.1	A0A0V8S7G9	0,12	0,04
KSW18506.1	A0A0V8SE91	0,12	0,04
KSW28272.1	A0A0V8T6T2	0,12	0,04
KSW29726.1	A0A0V8TAG7	0,12	0,04
KSW29226.1	A0A0V8TAQ1	0,12	0,04
KSW21678.1	A0A0V8SMY7	0,11	0,03
KSW17212.1	A0A0V8SA88	0,11	0,03
KSW18781.1	A0A0V8SEP2	0,11	0,03
KSW14015.1	A0A0V8S111	0,11	0,03
KSW21460.1	A0A0V8SMA2	0,11	0,03
KSW30160.1	A0A0V8TCD5	0,11	0,03
KSW14495.1	A0A0V8S2E5	0,1	0,03
KSW28600.1	A0A0V8T7P0	0,1	0,03
KSW28603.1	A0A0V8T7R2	0,1	0,03
KSW13730.1	A0A0V8S0N0	0,1	0,03
KSW21663.1	A0A0V8SMY5	0,1	0,03
KSW29342.1	A0A0V8T9C9	0,1	0,03
KSW29809.1	A0A0V8TBG7	0,1	0,03
KSW28556.1	A0A0V8T7U2	0,09	0,03
KSW15644.1	A0A0V8S5P8	0,09	0,03
KSW29030.1	A0A0V8T9E2	0,09	0,03
KSW29709.1	A0A0V8TAZ0	0,09	0,03
KSW29818.1	A0A0V8TCJ7	0,09	0,03
KSW28631.1	A0A0V8T7M2	0,08	0,03
KSW29439.1	A0A0V8TA20	0,08	0,03
KSW17957.1	A0A0V8SCF1	0,08	0,02
KSW19795.1	A0A0V8SHJ7	0,08	0,02
KSW20732.1	A0A0V8SK80	0,08	0,02
KSW29417.1	A0A0V8T9P9	0,07	0,02
KSW28288.1	A0A0V8T6U4	0,07	0,02
KSW28374.1	A0A0V8T7N1	0,07	0,02
KSW18204.1	A0A0V8SD14	0,07	0,02
KSW20999.1	A0A0V8SL27	0,07	0,02
KSW29704.1	A0A0V8TAS0	0,07	0,02
KSW30163.1	A0A0V8TBX1	0,07	0,02
KSW15288.1	A0A0V8S4U1	0,06	0,02
KSW29139.1	A0A0V8T903	0,06	0,02
KSW19469.1	A0A0V8SGN7	0,06	0,02
KSW29842.1	A0A0V8TBB6	0,06	0,02
KSW29985.1	A0A0V8TBZ8	0,06	0,02
KSW29161.1	A0A0V8T8T7	0,05	0,01

KSW29514.1	A0A0V8TBP7	0,05	0,01
KSW29540.1	A0A0V8TA29	0,03	0,01
KSW29886.1	A0A0V8TAX3	0,03	0,01
KSW29124.1	A0A0V8T924	0,02	0,01
KSW23743.1	A0A0V8STS8	0,02	0,01
PTE01			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAi %
KSW29001.1	A0A0V8T881	267,27	37,85
KSW29460.1	A0A0V8T9W0	13,45	1,91
KSW20716.1	A0A0V8SK67	12,9	1,83
KSW23782.1	A0A0V8STY7	12,9	1,83
KSW19409.1	A0A0V8SGG3	12,11	1,72
KSW29200.1	A0A0V8T9I6	9	1,27
KSW19791.1	A0A0V8SHI8	8,01	1,13
KSW21866.1	A0A0V8SP17	7,91	1,12
KSW29516.1	A0A0V8TAE3	7,03	1
KSW20692.1	A0A0V8SK31	5,58	0,79
KSW20708.1	A0A0V8SK53	5,45	0,77
KSW17208.1	A0A0V8SA68	5,31	0,75
KSW29871.1	A0A0V8TBE7	4,75	0,67
KSW29454.1	A0A0V8T9M1	4,52	0,64
KSW20682.1	A0A0V8SKL8	4,18	0,59
KSW16352.1	A0A0V8S7Q7	4,01	0,57
KSW29892.1	A0A0V8TB43	3,92	0,56
KSW14099.1	A0A0V8S172	3,64	0,52
KSW23552.1	A0A0V8ST89	3,56	0,5
KSW20704.1	A0A0V8SK41	3,44	0,49
KSW13738.1	A0A0V8S0K7	2,98	0,42
KSW14704.1	A0A0V8S2Z3	2,98	0,42
KSW20691.1	A0A0V8SK50	2,98	0,42
KSW28305.1	A0A0V8T764	2,98	0,42
KSW29700.1	A0A0V8TAK9	2,88	0,41
KSW20681.1	A0A0V8SK30	2,83	0,4
KSW29589.1	A0A0V8TAZ1	2,79	0,4
KSW20674.1	A0A0V8SKL6	2,59	0,37
KSW21774.1	A0A0V8SN80	2,38	0,34
KSW29082.1	A0A0V8TAB3	2,38	0,34
KSW21865.1	A0A0V8SNF7	2,34	0,33
KSW29108.1	A0A0V8T9I3	2,28	0,32
KSW14007.1	A0A0V8S185	2,16	0,31
KSW17696.1	A0A0V8SBK8	2,16	0,31
KSW18491.1	A0A0V8SDU0	2,16	0,31
KSW20702.1	A0A0V8SK56	2,16	0,31
KSW20705.1	A0A0V8SK71	2,16	0,31

KSW21735.1	A0A0V8SN47	2,16	0,31
KSW28552.1	A0A0V8T7K3	2,16	0,31
KSW28601.1	A0A0V8T771	2,16	0,31
KSW28752.1	A0A0V8T8B7	2,16	0,31
KSW29230.1	A0A0V8T922	2,16	0,31
KSW30206.1	A0A0V8TBN8	2,16	0,31
KSW21470.1	A0A0V8SMC4	2,07	0,29
KSW16349.1	A0A0V8S7T2	1,95	0,28
KSW29389.1	A0A0V8TA61	1,89	0,27
KSW20673.1	A0A0V8SK40	1,85	0,26
KSW28591.1	A0A0V8T7Y8	1,75	0,25
KSW18330.1	A0A0V8SDQ3	1,71	0,24
KSW16984.1	A0A0V8S9J4	1,68	0,24
KSW29110.1	A0A0V8T9J8	1,68	0,24
KSW13743.1	A0A0V8S092	1,61	0,23
KSW20687.1	A0A0V8SK61	1,61	0,23
KSW29330.1	A0A0V8T9W1	1,59	0,23
KSW28554.1	A0A0V8T8R8	1,51	0,21
KSW21897.1	A0A0V8SNI9	1,42	0,2
KSW16601.1	A0A0V8S8F6	1,37	0,19
KSW20684.1	A0A0V8SK26	1,37	0,19
KSW20690.1	A0A0V8SKN7	1,37	0,19
KSW20707.1	A0A0V8SK66	1,37	0,19
KSW20735.1	A0A0V8SKA8	1,37	0,19
KSW28618.1	A0A0V8T919	1,37	0,19
KSW29822.1	A0A0V8TAQ2	1,37	0,19
KSW16982.1	A0A0V8S9L5	1,36	0,19
KSW15045.1	A0A0V8S422	1,31	0,19
KSW20736.1	A0A0V8SK77	1,31	0,19
KSW21527.1	A0A0V8SMJ2	1,3	0,18
KSW15930.1	A0A0V8S6Q4	1,28	0,18
KSW20550.1	A0A0V8SJQ2	1,28	0,18
KSW28808.1	A0A0V8T8D2	1,28	0,18
KSW29333.1	A0A0V8TA14	1,24	0,18
KSW28310.1	A0A0V8T7E0	1,19	0,17
KSW15606.1	A0A0V8S5M2	1,15	0,16
KSW17699.1	A0A0V8SBL8	1,15	0,16
KSW20695.1	A0A0V8SK73	1,15	0,16
KSW20696.1	A0A0V8SK72	1,15	0,16
KSW20717.1	A0A0V8SK85	1,15	0,16
KSW21884.1	A0A0V8SNI3	1,15	0,16
KSW23490.1	A0A0V8STB7	1,15	0,16
KSW28600.1	A0A0V8T7P0	1,15	0,16
KSW28729.1	A0A0V8T7H1	1,15	0,16
KSW29406.1	A0A0V8T9H8	1,15	0,16

KSW29437.1	A0A0V8TAA0	1,15	0,16
KSW29451.1	A0A0V8T9Z5	1,15	0,16
KSW29518.1	A0A0V8T9Z1	1,15	0,16
KSW29798.1	A0A0V8TBK4	1,15	0,16
KSW29903.1	A0A0V8TBH7	1,15	0,16
KSW30102.1	A0A0V8TCD8	1,15	0,16
KSW30106.1	A0A0V8TDF3	1,15	0,16
KSW30161.1	A0A0V8TCM1	1,09	0,15
KSW21230.1	A0A0V8SLN9	1,08	0,15
KSW15553.1	A0A0V8S5G6	1,07	0,15
KSW17684.1	A0A0V8SBK5	1,05	0,15
KSW15610.1	A0A0V8S5S9	1	0,14
KSW17688.1	A0A0V8SBL5	1	0,14
KSW21834.1	A0A0V8SNX9	1	0,14
KSW28273.1	A0A0V8T734	1	0,14
KSW15315.1	A0A0V8S4T7	0,97	0,14
KSW28556.1	A0A0V8T7U2	0,97	0,14
KSW21455.1	A0A0V8SMD3	0,96	0,14
KSW15046.1	A0A0V8S3X8	0,93	0,13
KSW21499.1	A0A0V8SMC9	0,93	0,13
KSW28387.1	A0A0V8T703	0,93	0,13
KSW29054.1	A0A0V8T8P1	0,93	0,13
KSW29799.1	A0A0V8TBD9	0,93	0,13
KSW29819.1	A0A0V8TB15	0,93	0,13
KSW20713.1	A0A0V8SK59	0,91	0,13
KSW13731.1	A0A0V8S079	0,91	0,13
KSW20697.1	A0A0V8SK42	0,87	0,12
KSW20710.1	A0A0V8SK68	0,87	0,12
KSW23701.1	A0A0V8STW7	0,87	0,12
KSW29132.1	A0A0V8T9B3	0,87	0,12
KSW29232.1	A0A0V8T9K7	0,87	0,12
KSW29884.1	A0A0V8TBC4	0,87	0,12
KSW29887.1	A0A0V8TBG6	0,87	0,12
KSW30268.1	A0A0V8TCM5	0,87	0,12
KSW30107.1	A0A0V8TBQ0	0,86	0,12
KSW14134.1	A0A0V8S1F3	0,85	0,12
KSW23570.1	A0A0V8STJ9	0,85	0,12
KSW29093.1	A0A0V8T997	0,85	0,12
KSW29215.1	A0A0V8T9G7	0,85	0,12
KSW30256.1	A0A0V8TCL8	0,85	0,12
KSW18268.1	A0A0V8SD74	0,83	0,12
KSW30322.1	A0A0V8TCP4	0,83	0,12
KSW15434.1	A0A0V8S598	0,82	0,12
KSW13988.1	A0A0V8S123	0,82	0,12
KSW29539.1	A0A0V8TA57	0,81	0,11

KSW14100.1	A0A0V8S1C5	0,78	0,11
KSW15433.1	A0A0V8S563	0,78	0,11
KSW16280.1	A0A0V8S7K3	0,78	0,11
KSW20689.1	A0A0V8SK55	0,78	0,11
KSW21272.1	A0A0V8SLX2	0,78	0,11
KSW21840.1	A0A0V8SNH9	0,78	0,11
KSW21879.1	A0A0V8SNI8	0,78	0,11
KSW28603.1	A0A0V8T7R2	0,78	0,11
KSW28622.1	A0A0V8T7E4	0,78	0,11
KSW28637.1	A0A0V8T865	0,78	0,11
KSW29410.1	A0A0V8TA38	0,78	0,11
KSW29695.1	A0A0V8TAU0	0,78	0,11
KSW29946.1	A0A0V8TCV4	0,78	0,11
KSW29974.1	A0A0V8TC37	0,78	0,11
KSW30135.1	A0A0V8TCB0	0,78	0,11
KSW30185.1	A0A0V8TBR6	0,78	0,11
KSW15219.1/KSW21896.1	A0A0V8SNP3	0,74	0,11
KSW28328.1	A0A0V8T700	0,74	0,1
KSW28630.1	A0A0V8T8A7	0,73	0,1
KSW30101.1	A0A0V8TCD6	0,72	0,1
KSW23491.1	A0A0V8ST30	0,72	0,1
KSW29473.1	A0A0V8TAG8	0,72	0,1
KSW20723.1	A0A0V8SK82	0,7	0,1
KSW29261.1	A0A0V8T9T5	0,7	0,1
KSW29692.1	A0A0V8TAV7	0,7	0,1
KSW30208.1	A0A0V8TCH5	0,7	0,1
KSW16067.1	A0A0V8S6Y5	0,67	0,09
KSW18504.1	A0A0V8SDX0	0,67	0,09
KSW19597.1	A0A0V8SH13	0,67	0,09
KSW20093.1	A0A0V8SIF4	0,67	0,09
KSW20394.1	A0A0V8SJT5	0,67	0,09
KSW23799.1	A0A0V8SU02	0,67	0,09
KSW29276.1	A0A0V8T9N1	0,67	0,09
KSW29367.1	A0A0V8TA69	0,67	0,09
KSW29635.1	A0A0V8TAG0	0,67	0,09
KSW21274.1	A0A0V8SMB6	0,65	0,09
KSW20712.1	A0A0V8SK93	0,64	0,09
KSW14009.1	A0A0V8S120	0,64	0,09
KSW18565.1	A0A0V8SE56	0,64	0,09
KSW21229.1	A0A0V8SLR7	0,64	0,09
KSW30128.1	A0A0V8TCA1	0,64	0,09
KSW28620.1	A0A0V8T7Z7	0,62	0,09
KSW29690.1	A0A0V8TC65	0,62	0,09
KSW30155.1	A0A0V8TBT6	0,62	0,09
KSW30332.1	A0A0V8TCQ3	0,62	0,09

KSW21529.1	A0A0V8SMI0	0,62	0,09
KSW29834.1	A0A0V8TCL7	0,62	0,09
KSW30037.1	A0A0V8TC78	0,62	0,09
KSW28372.1	A0A0V8T772	0,61	0,09
KSW30273.1	A0A0V8TD39	0,61	0,09
KSW29922.1	A0A0V8TBM6	0,6	0,09
KSW14076.1	A0A0V8S166	0,59	0,08
KSW14325.1	A0A0V8S213	0,59	0,08
KSW15225.1	A0A0V8S4I0	0,59	0,08
KSW16589.1	A0A0V8S8F3	0,59	0,08
KSW17188.1	A0A0V8SA82	0,59	0,08
KSW17192.1	A0A0V8SA53	0,59	0,08
KSW18267.1	A0A0V8SD79	0,59	0,08
KSW18518.1	A0A0V8SDY1	0,59	0,08
KSW20686.1	A0A0V8SK39	0,59	0,08
KSW20703.1	A0A0V8SK78	0,59	0,08
KSW20706.1	A0A0V8SKQ6	0,59	0,08
KSW20724.1	A0A0V8SK64	0,59	0,08
KSW21459.1	A0A0V8SMA9	0,59	0,08
KSW28646.1	A0A0V8T8C3	0,59	0,08
KSW29140.1	A0A0V8T939	0,59	0,08
KSW29244.1	A0A0V8T9K2	0,59	0,08
KSW29320.1	A0A0V8T9S7	0,59	0,08
KSW29500.1	A0A0V8TAC8	0,59	0,08
KSW29527.1	A0A0V8TAP7	0,59	0,08
KSW30049.1	A0A0V8TCB4	0,59	0,08
KSW16069.1	A0A0V8S6Z1	0,57	0,08
KSW21825.1	A0A0V8SNB8	0,57	0,08
KSW16351.1	A0A0V8S7R6	0,55	0,08
KSW23777.1	A0A0V8STZ4	0,55	0,08
KSW14629.1	A0A0V8S2W8	0,54	0,08
KSW15158.1	A0A0V8S487	0,54	0,08
KSW16073.1	A0A0V8S6Y7	0,54	0,08
KSW21020.1	A0A0V8SL08	0,54	0,08
KSW29937.1	A0A0V8TBS6	0,54	0,08
KSW30054.1	A0A0V8TC91	0,54	0,08
KSW17212.1	A0A0V8SA88	0,52	0,07
KSW18074.1	A0A0V8SCY8	0,52	0,07
KSW19621.1	A0A0V8SH48	0,52	0,07
KSW28888.1	A0A0V8T8J8	0,52	0,07
KSW29378.1	A0A0V8TA08	0,52	0,07
KSW29653.1	A0A0V8TB45	0,52	0,07
KSW13763.1	A0A0V8S0B6	0,5	0,07
KSW15154.1	A0A0V8S4M6	0,5	0,07
KSW15169.1	A0A0V8S4D9	0,5	0,07

KSW15286.1	A0A0V8S4N4	0,5	0,07
KSW21039.1	A0A0V8SL49	0,5	0,07
KSW21929.1	A0A0V8SNM2	0,5	0,07
KSW28627.1	A0A0V8T7J2	0,5	0,07
KSW30205.1	A0A0V8TCE0	0,5	0,07
KSW13742.1	A0A0V8S0G4	0,49	0,07
KSW15161.1	A0A0V8S4B6	0,47	0,07
KSW17321.1	A0A0V8SAJ4	0,47	0,07
KSW19115.1	A0A0V8SFL7	0,47	0,07
KSW20698.1	A0A0V8SKN1	0,47	0,07
KSW21273.1	A0A0V8SLQ2	0,47	0,07
KSW21462.1	A0A0V8SMB1	0,47	0,07
KSW23699.1	A0A0V8STR1	0,47	0,07
KSW28201.1	A0A0V8T686	0,47	0,07
KSW29070.1	A0A0V8T8Q7	0,47	0,07
KSW29081.1	A0A0V8T8L4	0,47	0,07
KSW29109.1	A0A0V8T9A8	0,47	0,07
KSW29148.1	A0A0V8T9C8	0,47	0,07
KSW29345.1	A0A0V8TA73	0,47	0,07
KSW29612.1	A0A0V8TAN5	0,47	0,07
KSW29704.1	A0A0V8TAS0	0,47	0,07
KSW29821.1	A0A0V8TB90	0,47	0,07
KSW30229.1	A0A0V8TCS5	0,47	0,07
KSW30260.1	A0A0V8TCB7	0,47	0,07
KSW30261.1	A0A0V8TCV7	0,47	0,07
KSW20824.1	A0A0V8SKK2	0,45	0,06
KSW28323.1	A0A0V8T6Q5	0,45	0,06
KSW29732.1	A0A0V8TAN9	0,45	0,06
KSW14798.1	A0A0V8S3A4	0,44	0,06
KSW16111.1	A0A0V8S708	0,44	0,06
KSW21669.1	A0A0V8SN74	0,44	0,06
KSW23794.1	A0A0V8SU61	0,44	0,06
KSW15001.1	A0A0V8S3V5	0,43	0,06
KSW13991.1	A0A0V8S169	0,43	0,06
KSW14796.1	A0A0V8S386	0,43	0,06
KSW16071.1	A0A0V8S775	0,43	0,06
KSW20488.1	A0A0V8SJL5	0,43	0,06
KSW28886.1	A0A0V8T914	0,43	0,06
KSW29592.1	A0A0V8TAG3	0,42	0,06
KSW16031.1	A0A0V8S6S6	0,41	0,06
KSW28914.1	A0A0V8T8E8	0,41	0,06
KSW29904.1	A0A0V8TBP8	0,41	0,06
KSW23550.1	A0A0V8STD4	0,41	0,06
KSW13752.1/ KSW14630.1	A0A0V8S2R9	0,39	0,06
KSW15900.1	A0A0V8S6E3	0,39	0,06

KSW16350.1	A0A0V8S7P9	0,39	0,06
KSW16594.1	A0A0V8S8Y1	0,39	0,06
KSW16597.1	A0A0V8S8H5	0,39	0,06
KSW17209.1	A0A0V8SA74	0,39	0,06
KSW17322.1	A0A0V8SAR9	0,39	0,06
KSW18203.1	A0A0V8SD15	0,39	0,06
KSW18322.1	A0A0V8SDX3	0,39	0,06
KSW20683.1	A0A0V8SK28	0,39	0,06
KSW20699.1	A0A0V8SK45	0,39	0,06
KSW20700.1	A0A0V8SK48	0,39	0,06
KSW21507.1	A0A0V8SMF4	0,39	0,06
KSW21821.1	A0A0V8SNG6	0,39	0,06
KSW21863.1	A0A0V8SNH2	0,39	0,06
KSW23497.1	A0A0V8ST45	0,39	0,06
KSW28303.1	A0A0V8T6S2	0,39	0,06
KSW28925.1	A0A0V8T8W5	0,39	0,06
KSW29224.1	A0A0V8T9J2	0,39	0,06
KSW29444.1	A0A0V8T9U1	0,39	0,06
KSW29713.1	A0A0V8TB48	0,39	0,06
KSW30057.1	A0A0V8TBG1	0,39	0,06
KSW30163.1	A0A0V8TBX1	0,38	0,05
KSW18199.1	A0A0V8SD16	0,37	0,05
KSW29614.1	A0A0V8TA87	0,37	0,05
KSW14624.1	A0A0V8S2R5	0,36	0,05
KSW21042.1	A0A0V8SLN7	0,36	0,05
KSW23795.1	A0A0V8STZ9	0,36	0,05
KSW28971.1	A0A0V8T8Q9	0,36	0,05
KSW29166.1	A0A0V8T8X7	0,36	0,05
KSW29734.1	A0A0V8TBF0	0,36	0,05
KSW29674.1	A0A0V8TC53	0,35	0,05
KSW13993.1	A0A0V8S107	0,33	0,05
KSW15167.1	A0A0V8S4C1	0,33	0,05
KSW15314.1	A0A0V8S504	0,33	0,05
KSW17692.1	A0A0V8SBL0	0,33	0,05
KSW21815.1	A0A0V8SNC7	0,33	0,05
KSW21870.1	A0A0V8SNH7	0,33	0,05
KSW21873.1	A0A0V8SNF9	0,33	0,05
KSW23551.1	A0A0V8STA5	0,33	0,05
KSW28306.1	A0A0V8T6P6	0,33	0,05
KSW29205.1	A0A0V8T9P1	0,33	0,05
KSW29229.1	A0A0V8T9R1	0,33	0,05
KSW29426.1	A0A0V8TA52	0,33	0,05
KSW29708.1	A0A0V8TAW6	0,33	0,05
KSW29793.1	A0A0V8TBF7	0,33	0,05
KSW30162.1	A0A0V8TC98	0,33	0,05

KSW29514.1	A0A0V8TBP7	0,33	0,05
KSW15696.1	A0A0V8S5W2	0,32	0,05
KSW16070.1	A0A0V8S6Y4	0,32	0,05
KSW19596.1	A0A0V8SH09	0,32	0,05
KSW21872.1	A0A0V8SNK9	0,32	0,05
KSW28611.1	A0A0V8T7H6	0,32	0,05
KSW29929.1	A0A0V8TB59	0,32	0,05
KSW30263.1	A0A0V8TCQ8	0,31	0,04
KSW13734.1	A0A0V8S074	0,31	0,04
KSW13758.1	A0A0V8S0I0	0,31	0,04
KSW15223.1	A0A0V8S4N0	0,31	0,04
KSW16599.1	A0A0V8S8K8	0,31	0,04
KSW19613.1	A0A0V8SH29	0,31	0,04
KSW20701.1	A0A0V8SK70	0,31	0,04
KSW20722.1	A0A0V8SKS3	0,31	0,04
KSW21556.1	A0A0V8SMI9	0,31	0,04
KSW28331.1	A0A0V8T6X0	0,31	0,04
KSW29694.1	A0A0V8TAD1	0,31	0,04
KSW15663.1	A0A0V8S5R6	0,3	0,04
KSW16357.1	A0A0V8S7T4	0,3	0,04
KSW21713.1	A0A0V8SN12	0,3	0,04
KSW13995.1	A0A0V8S186	0,29	0,04
KSW15002.1	A0A0V8S469	0,29	0,04
KSW16596.1	A0A0V8S8G3	0,29	0,04
KSW19592.1	A0A0V8SGY0	0,29	0,04
KSW20191.1	A0A0V8SIQ9	0,29	0,04
KSW21018.1	A0A0V8SLK0	0,29	0,04
KSW21779.1	A0A0V8SN77	0,29	0,04
KSW28327.1	A0A0V8T6U9	0,29	0,04
KSW28562.1	A0A0V8T7E9	0,29	0,04
KSW29705.1	A0A0V8TAK6	0,29	0,04
KSW15916.1	A0A0V8S6F9	0,28	0,04
KSW12785.1	A0A0V8RXL6	0,27	0,04
KSW14008.1	A0A0V8S182	0,27	0,04
KSW28274.1	A0A0V8T6L3	0,27	0,04
KSW29613.1	A0A0V8TAP0	0,27	0,04
KSW28373.1	A0A0V8T737	0,27	0,04
KSW14021.1	A0A0V8S144	0,26	0,04
KSW15224.1	A0A0V8S4P7	0,26	0,04
KSW15290.1	A0A0V8S4X1	0,26	0,04
KSW18495.1	A0A0V8SDU2	0,26	0,04
KSW19441.1	A0A0V8SGK6	0,26	0,04
KSW19607.1	A0A0V8SH26	0,26	0,04
KSW20219.1	A0A0V8SIS6	0,26	0,04
KSW20308.1	A0A0V8SJ26	0,26	0,04

KSW20567.1	A0A0V8SJU5	0,26	0,04
KSW20709.1	A0A0V8SKB9	0,26	0,04
KSW20721.1	A0A0V8SK87	0,26	0,04
KSW20734.1	A0A0V8SK95	0,26	0,04
KSW21716.1	A0A0V8SN10	0,26	0,04
KSW28407.1	A0A0V8T723	0,26	0,04
KSW28995.1	A0A0V8T8P5	0,26	0,04
KSW29094.1	A0A0V8T9I4	0,26	0,04
KSW29106.1	A0A0V8T969	0,26	0,04
KSW29286.1	A0A0V8TA15	0,26	0,04
KSW29356.1	A0A0V8T9V7	0,26	0,04
KSW29785.1	A0A0V8TAS3	0,26	0,04
KSW29842.1	A0A0V8TBB6	0,26	0,04
KSW30242.1	A0A0V8TCI4	0,26	0,04
KSW30270.1	A0A0V8TBV4	0,26	0,04
KSW30304.1	A0A0V8TCV3	0,26	0,04
KSW14626.1	A0A0V8S338	0,25	0,04
KSW21739.1	A0A0V8SN33	0,25	0,04
KSW15066.1	A0A0V8S4D5	0,25	0,03
KSW15162.1	A0A0V8S4M7	0,25	0,03
KSW16202.1	A0A0V8S7G9	0,25	0,03
KSW19470.1	A0A0V8SGL0	0,25	0,03
KSW19794.1	A0A0V8SI61	0,25	0,03
KSW20726.1	A0A0V8SK84	0,25	0,03
KSW19867.1	A0A0V8SHT1	0,24	0,03
KSW14024.1	A0A0V8S195	0,24	0,03
KSW13997.1	A0A0V8SOZ3	0,23	0,03
KSW14153.1	A0A0V8S1G2	0,23	0,03
KSW15304.1	A0A0V8S4V8	0,23	0,03
KSW18781.1	A0A0V8SEP2	0,23	0,03
KSW18930.1	A0A0V8SFM4	0,23	0,03
KSW19583.1	A0A0V8SGX2	0,23	0,03
KSW21449.1	A0A0V8SM91	0,23	0,03
KSW21867.1	A0A0V8SND6	0,23	0,03
KSW21893.1	A0A0V8SNT8	0,23	0,03
KSW28309.1	A0A0V8T6X9	0,23	0,03
KSW29032.1	A0A0V8T8X3	0,23	0,03
KSW29099.1	A0A0V8T957	0,23	0,03
KSW29223.1	A0A0V8T9T1	0,23	0,03
KSW29326.1	A0A0V8T9B4	0,23	0,03
KSW29578.1	A0A0V8TBV3	0,23	0,03
KSW30227.1	A0A0V8TC36	0,23	0,03
KSW21458.1	A0A0V8SMW5	0,22	0,03
KSW29535.1	A0A0V8TAD7	0,22	0,03
KSW30017.1	A0A0V8TC32	0,22	0,03

KSW30221.1	A0A0V8TCG4	0,22	0,03
KSW29118.1	A0A0V8T8U7	0,22	0,03
KSW15297.1	A0A0V8S4R3	0,21	0,03
KSW17200.1	A0A0V8SA63	0,21	0,03
KSW18655.1	A0A0V8SE95	0,21	0,03
KSW19086.1	A0A0V8SFI7	0,21	0,03
KSW19410.1	A0A0V8SH11	0,21	0,03
KSW21040.1	A0A0V8SL26	0,21	0,03
KSW28275.1	A0A0V8T6K3	0,21	0,03
KSW28998.1	A0A0V8T9B1	0,21	0,03
KSW29048.1	A0A0V8T937	0,21	0,03
KSW29295.1	A0A0V8T9N2	0,21	0,03
KSW30292.1	A0A0V8TCE7	0,21	0,03
KSW29095.1	A0A0V8T960	0,21	0,03
KSW29105.1	A0A0V8T9D3	0,21	0,03
KSW29710.1	A0A0V8TAE8	0,21	0,03
KSW30223.1	A0A0V8TCI0	0,21	0,03
KSW16426.1	A0A0V8S831	0,2	0,03
KSW29342.1	A0A0V8T9C9	0,2	0,03
KSW29651.1	A0A0V8TAI1	0,2	0,03
KSW15736.1	A0A0V8S621	0,19	0,03
KSW17199.1	A0A0V8SA64	0,19	0,03
KSW18201.1	A0A0V8SD03	0,19	0,03
KSW18709.1	A0A0V8SEJ5	0,19	0,03
KSW19587.1	A0A0V8SGZ8	0,19	0,03
KSW19866.1	A0A0V8SIB5	0,19	0,03
KSW19920.1	A0A0V8SHZ5	0,19	0,03
KSW21525.1	A0A0V8SMS1	0,19	0,03
KSW21823.1	A0A0V8SND2	0,19	0,03
KSW21928.1	A0A0V8SNS0	0,19	0,03
KSW23779.1	A0A0V8STY6	0,19	0,03
KSW28337.1	A0A0V8T7A0	0,19	0,03
KSW28906.1	A0A0V8T9W8	0,19	0,03
KSW29139.1	A0A0V8T903	0,19	0,03
KSW29388.1	A0A0V8T9Y4	0,19	0,03
KSW29829.1	A0A0V8TBN1	0,19	0,03
KSW29942.1	A0A0V8TBY6	0,19	0,03
KSW30105.1	A0A0V8TBK0	0,18	0,03
KSW13745.1	A0A0V8S0A1	0,18	0,03
KSW13753.1	A0A0V8S094	0,18	0,03
KSW15562.1	A0A0V8S5L9	0,18	0,03
KSW16068.1	A0A0V8S6X8	0,18	0,03
KSW17484.1	A0A0V8SB06	0,18	0,03
KSW17686.1	A0A0V8SBL6	0,18	0,03
KSW17706.1	A0A0V8SBW0	0,18	0,03

KSW18195.1	A0A0V8SD26	0,18	0,03
KSW18985.1	A0A0V8SF91	0,18	0,03
KSW20310.1	A0A0V8SJ27	0,18	0,03
KSW21559.1	A0A0V8SMM1	0,18	0,03
KSW21706.1	A0A0V8SNL1	0,18	0,03
KSW28575.1	A0A0V8T7X5	0,18	0,03
KSW29029.1	A0A0V8T941	0,18	0,03
KSW29509.1	A0A0V8TAR6	0,18	0,03
KSW29560.1	A0A0V8TAD3	0,18	0,03
KSW29818.1	A0A0V8TCJ7	0,18	0,03
KSW29908.1	A0A0V8TB56	0,18	0,03
KSW30127.1	A0A0V8TC84	0,18	0,03
KSW30272.1	A0A0V8TCS0	0,18	0,03
KSW28631.1	A0A0V8T7M2	0,17	0,02
KSW29165.1	A0A0V8T9K1	0,17	0,02
KSW13736.1	A0A0V8S0F9	0,17	0,02
KSW14018.1	A0A0V8S1G3	0,17	0,02
KSW18483.1	A0A0V8SDW8	0,17	0,02
KSW19094.1	A0A0V8SFK9	0,17	0,02
KSW19404.1	A0A0V8SGF9	0,17	0,02
KSW19406.1	A0A0V8SGE8	0,17	0,02
KSW19792.1	A0A0V8SHL7	0,17	0,02
KSW20977.1	A0A0V8SKZ6	0,17	0,02
KSW21446.1	A0A0V8SM94	0,17	0,02
KSW23792.1	A0A0V8STX8	0,17	0,02
KSW28656.1	A0A0V8T7W5	0,17	0,02
KSW29832.1	A0A0V8TB42	0,17	0,02
KSW29965.1	A0A0V8TBR7	0,17	0,02
KSW30078.1	A0A0V8TBE2	0,17	0,02
KSW30081.1	A0A0V8TCF3	0,17	0,02
KSW30211.1	A0A0V8TC22	0,17	0,02
KSW15435.1	A0A0V8S5B3	0,16	0,02
KSW23549.1	A0A0V8STJ6	0,16	0,02
KSW29979.1	A0A0V8TBF3	0,16	0,02
KSW15283.1	A0A0V8S4N8	0,16	0,02
KSW16064.1	A0A0V8S6W8	0,16	0,02
KSW17032.1	A0A0V8S9N8	0,16	0,02
KSW17615.1	A0A0V8SBD0	0,16	0,02
KSW21029.1	A0A0V8SLA5	0,16	0,02
KSW28555.1	A0A0V8T7M1	0,16	0,02
KSW28560.1	A0A0V8T7L9	0,16	0,02
KSW28615.1	A0A0V8T7L0	0,16	0,02
KSW28807.1	A0A0V8T843	0,16	0,02
KSW29945.1	A0A0V8TB72	0,16	0,02
KSW19440.1	A0A0V8SGJ9	0,15	0,02

KSW14011.1	A0A0V8S1A3	0,15	0,02
KSW15608.1	A0A0V8S5N8	0,15	0,02
KSW16603.1	A0A0V8S8I8	0,15	0,02
KSW20494.1	A0A0V8SJK1	0,15	0,02
KSW21263.1	A0A0V8SLT3	0,15	0,02
KSW21317.1	A0A0V8SM46	0,15	0,02
KSW21831.1	A0A0V8SNE1	0,15	0,02
KSW21868.1	A0A0V8SNG2	0,15	0,02
KSW23474.1	A0A0V8STB0	0,15	0,02
KSW28276.1	A0A0V8T6X1	0,15	0,02
KSW28764.1	A0A0V8T8C4	0,15	0,02
KSW29098.1	A0A0V8TAC7	0,15	0,02
KSW29756.1	A0A0V8TB21	0,15	0,02
KSW30200.1	A0A0V8TC72	0,15	0,02
KSW30305.1	A0A0V8TD72	0,15	0,02
KSW28374.1	A0A0V8T7N1	0,14	0,02
KSW14437.1	A0A0V8S2C5	0,14	0,02
KSW21526.1	A0A0V8SMH8	0,14	0,02
KSW23569.1	A0A0V8STE2	0,14	0,02
KSW23773.1	A0A0V8SUA5	0,14	0,02
KSW23790.1	A0A0V8SU34	0,14	0,02
KSW28551.1	A0A0V8T7E8	0,14	0,02
KSW28644.1	A0A0V8T803	0,14	0,02
KSW29000.1	A0A0V8T8V1	0,14	0,02
KSW29038.1	A0A0V8T8N1	0,14	0,02
KSW29084.1	A0A0V8T979	0,14	0,02
KSW29204.1	A0A0V8T993	0,14	0,02
KSW29435.1	A0A0V8T9X7	0,14	0,02
KSW29616.1	A0A0V8TAT3	0,14	0,02
KSW15288.1	A0A0V8S4U1	0,13	0,02
KSW15294.1	A0A0V8S4Q8	0,13	0,02
KSW18408.1	A0A0V8SDN0	0,13	0,02
KSW18484.1	A0A0V8SDY2	0,13	0,02
KSW18932.1	A0A0V8SF79	0,13	0,02
KSW19110.1	A0A0V8SFM8	0,13	0,02
KSW19439.1	A0A0V8SGH6	0,13	0,02
KSW19472.1	A0A0V8SGN6	0,13	0,02
KSW28371.1	A0A0V8T6Z5	0,13	0,02
KSW28553.1	A0A0V8T732	0,13	0,02
KSW28654.1	A0A0V8T7H0	0,13	0,02
KSW28996.1	A0A0V8T8U2	0,13	0,02
KSW29198.1	A0A0V8T8Z7	0,13	0,02
KSW29580.1	A0A0V8TAK7	0,13	0,02
KSW29872.1	A0A0V8TBJ4	0,13	0,02
KSW29938.1	A0A0V8TBP3	0,13	0,02

KSW29987.1	A0A0V8TBH1	0,13	0,02
KSW30050.1	A0A0V8TBZ9	0,13	0,02
KSW30111.1	A0A0V8TC68	0,13	0,02
KSW29228.1	A0A0V8T9I9	0,13	0,02
KSW29250.1	A0A0V8T9N6	0,13	0,02
KSW14470.1	A0A0V8S2B2	0,12	0,02
KSW15289.1	A0A0V8S4M8	0,12	0,02
KSW15898.1	A0A0V8S6M3	0,12	0,02
KSW16063.1	A0A0V8S6V9	0,12	0,02
KSW19469.1	A0A0V8SGN7	0,12	0,02
KSW20688.1	A0A0V8SK35	0,12	0,02
KSW20719.1	A0A0V8SK91	0,12	0,02
KSW21570.1	A0A0V8SN55	0,12	0,02
KSW21668.1	A0A0V8SMV7	0,12	0,02
KSW28579.1	A0A0V8T7E1	0,12	0,02
KSW28599.1	A0A0V8T7J6	0,12	0,02
KSW28660.1	A0A0V8T820	0,12	0,02
KSW29057.1	A0A0V8T996	0,12	0,02
KSW29670.1	A0A0V8TB61	0,12	0,02
KSW29742.1	A0A0V8TAI3	0,12	0,02
KSW29815.1	A0A0V8TBE9	0,12	0,02
KSW29933.1	A0A0V8TBM1	0,12	0,02
KSW29966.1	A0A0V8TB53	0,12	0,02
KSW29985.1	A0A0V8TBZ8	0,12	0,02
KSW16075.1	A0A0V8S758	0,12	0,02
KSW21250.1	A0A0V8SM71	0,12	0,02
KSW21534.1	A0A0V8SMJ3	0,12	0,02
KSW21880.1	A0A0V8SNM7	0,12	0,02
KSW28272.1	A0A0V8T6T2	0,12	0,02
KSW28963.1	A0A0V8T8K6	0,12	0,02
KSW29145.1	A0A0V8T8S0	0,12	0,02
KSW29226.1	A0A0V8TAQ1	0,12	0,02
KSW15558.1	A0A0V8S5H0	0,11	0,02
KSW18266.1	A0A0V8SDJ2	0,11	0,02
KSW18290.1	A0A0V8SDT9	0,11	0,02
KSW18573.1	A0A0V8SE18	0,11	0,02
KSW19477.1	A0A0V8SGR0	0,11	0,02
KSW28366.1	A0A0V8T6V5	0,11	0,02
KSW28643.1	A0A0V8T7K7	0,11	0,02
KSW29056.1	A0A0V8T910	0,11	0,02
KSW29252.1	A0A0V8T9D0	0,11	0,02
KSW29254.1	A0A0V8T9Y7	0,11	0,02
KSW29804.1	A0A0V8TB54	0,11	0,02
KSW29814.1	A0A0V8TBL5	0,11	0,02
KSW30156.1	A0A0V8TC56	0,11	0,02

KSW30303.1	A0A0V8TCU7	0,11	0,02
KSW14015.1	A0A0V8S111	0,11	0,01
KSW16380.1	A0A0V8S7S7	0,11	0,01
KSW17752.1	A0A0V8SBR8	0,11	0,01
KSW19915.1	A0A0V8SHX6	0,11	0,01
KSW28338.1	A0A0V8T6W5	0,11	0,01
KSW29699.1	A0A0V8TAT0	0,11	0,01
KSW30160.1	A0A0V8TCD5	0,11	0,01
KSW21036.1	A0A0V8SL24	0,1	0,01
KSW17690.1	A0A0V8SBU6	0,1	0,01
KSW18269.1	A0A0V8SD75	0,1	0,01
KSW20672.1	A0A0V8SK13	0,1	0,01
KSW21775.1	A0A0V8SN78	0,1	0,01
KSW28597.1	A0A0V8T7S1	0,1	0,01
KSW29009.1	A0A0V8T963	0,1	0,01
KSW15301.1	A0A0V8S4R4	0,1	0,01
KSW15307.1	A0A0V8S4Z8	0,1	0,01
KSW18485.1	A0A0V8SDX2	0,1	0,01
KSW18494.1	A0A0V8SDU6	0,1	0,01
KSW18514.1	A0A0V8SEI5	0,1	0,01
KSW21663.1	A0A0V8SMY5	0,1	0,01
KSW29074.1	A0A0V8T934	0,1	0,01
KSW29274.1	A0A0V8TAV0	0,1	0,01
KSW29314.1	A0A0V8T9U5	0,1	0,01
KSW29644.1	A0A0V8TAQ4	0,1	0,01
KSW14109.1	A0A0V8S193	0,09	0,01
KSW19405.1	A0A0V8SGG7	0,09	0,01
KSW21740.1	A0A0V8SN29	0,09	0,01
KSW28412.1	A0A0V8T7F8	0,09	0,01
KSW29372.1	A0A0V8T9W9	0,09	0,01
KSW14101.1	A0A0V8S1B7	0,09	0,01
KSW14469.1	A0A0V8S2F6	0,09	0,01
KSW15165.1	A0A0V8S4D7	0,09	0,01
KSW29407.1	A0A0V8T9Y6	0,09	0,01
KSW29696.1	A0A0V8TB09	0,09	0,01
KSW19473.1	A0A0V8SGN5	0,09	0,01
KSW21936.1	A0A0V8SNR7	0,09	0,01
KSW28775.1	A0A0V8T812	0,08	0,01
KSW18191.1	A0A0V8SCZ1	0,08	0,01
KSW19614.1	A0A0V8SH07	0,08	0,01
KSW23778.1	A0A0V8SU43	0,08	0,01
KSW28271.1	A0A0V8T6P4	0,08	0,01
KSW29404.1	A0A0V8T9Z8	0,08	0,01
KSW29433.1	A0A0V8T9R0	0,08	0,01
KSW29439.1	A0A0V8TA20	0,08	0,01

KSW29457.1	A0A0V8TAE5	0,08	0,01
KSW18488.1	A0A0V8SDV3	0,08	0,01
KSW19795.1	A0A0V8SHJ7	0,08	0,01
KSW19796.1	A0A0V8SHN3	0,08	0,01
KSW19923.1	A0A0V8SHY4	0,08	0,01
KSW29681.1	A0A0V8TB12	0,08	0,01
KSW20732.1	A0A0V8SK80	0,08	0,01
KSW28626.1	A0A0V8T7L4	0,08	0,01
KSW29307.1	A0A0V8T9K9	0,08	0,01
KSW29434.1	A0A0V8TBF1	0,08	0,01
KSW29600.1	A0A0V8TAS1	0,08	0,01
KSW17694.1	A0A0V8SBM0	0,07	0,01
KSW29606.1	A0A0V8TAZ8	0,07	0,01
KSW14036.1	A0A0V8S164	0,07	0,01
KSW28378.1	A0A0V8T8B4	0,07	0,01
KSW21558.1	A0A0V8SMK2	0,07	0,01
KSW28377.1	A0A0V8T6R9	0,07	0,01
KSW18204.1	A0A0V8SD14	0,07	0,01
KSW29263.1	A0A0V8T9J9	0,06	0,01
KSW15047.1	A0A0V8S451	0,06	0,01
KSW15560.1	A0A0V8S5I5	0,06	0,01
KSW15642.1	A0A0V8S5V9	0,06	0,01
KSW15461.1	A0A0V8S576	0,06	0,01
KSW21535.1	A0A0V8SMK9	0,06	0,01
KSW28912.1	A0A0V8T8Q2	0,06	0,01
KSW16378.1	A0A0V8S7Z4	0,06	0,01
KSW21468.1	A0A0V8SMC2	0,06	0,01
KSW29631.1	A0A0V8TAM2	0,06	0,01
KSW29726.1	A0A0V8TAG7	0,06	0,01
KSW13761.1	A0A0V8S0B3	0,05	0,01
KSW21037.1	A0A0V8SL63	0,05	0,01
KSW29161.1	A0A0V8T8T7	0,05	0,01
KSW21887.1	A0A0V8SNJ8	0,05	0,01
KSW15644.1	A0A0V8S5P8	0,04	0,01
KSW29853.1	A0A0V8TBE8	0,04	0,01
KSW21836.1	A0A0V8SNC5	0,04	0,01
KSW30018.1	A0A0V8TBX5	0,04	0,01
KSW23555.1	A0A0V8ST98	0,04	0,01
KSW29090.1	A0A0V8T953	0,04	0,01
KSW29886.1	A0A0V8TAX3	0,03	0
KSW15221.1	A0A0V8S4K1	0,03	0
PTE02			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAI%
KSW12785.1	A0A0V8RXL6	0,2	0,1

KSW13734.1	A0A0V8S074	0,15	0,07
KSW13731.1	A0A0V8S079	0,43	0,22
KSW13743.1	A0A0V8S092	0,21	0,11
KSW13736.1	A0A0V8S0F9	0,17	0,08
KSW13742.1	A0A0V8S0G4	0,17	0,09
KSW13758.1	A0A0V8S0I0	0,15	0,07
KSW13738.1	A0A0V8S0K7	1	0,51
KSW14009.1	A0A0V8S120	0,18	0,09
KSW13988.1	A0A0V8S123	0,09	0,05
KSW14021.1	A0A0V8S144	0,26	0,13
KSW13991.1	A0A0V8S169	0,19	0,1
KSW14099.1	A0A0V8S172	1,15	0,59
KSW14007.1	A0A0V8S185	2,16	1,1
KSW14024.1	A0A0V8S195	0,1	0,05
KSW14247.1	A0A0V8S1S2	0,09	0,04
KSW14624.1	A0A0V8S2R5	0,17	0,08
KSW14629.1	A0A0V8S2W8	0,16	0,08
KSW14704.1	A0A0V8S2Z3	3,64	1,86
KSW14666.1	A0A0V8S382	0,59	0,3
KSW14798.1	A0A0V8S3A4	0,11	0,06
KSW15001.1	A0A0V8S3V5	0,08	0,04
KSW15068.1	A0A0V8S414	0,17	0,08
KSW15283.1	A0A0V8S4N8	0,16	0,08
KSW15294.1	A0A0V8S4Q8	0,27	0,14
KSW15315.1	A0A0V8S4T7	0,15	0,07
KSW15304.1	A0A0V8S4V8	0,15	0,08
KSW15433.1	A0A0V8S563	0,19	0,1
KSW15434.1	A0A0V8S598	0,49	0,25
KSW15537.1	A0A0V8S5F2	0,09	0,05
KSW15553.1	A0A0V8S5G6	0,62	0,32
KSW15560.1	A0A0V8S5I5	0,06	0,03
KSW15606.1	A0A0V8S5M2	1,15	0,59
KSW15608.1	A0A0V8S5N8	0,15	0,07
KSW15663.1	A0A0V8S5R6	0,43	0,22
KSW15696.1	A0A0V8S5W2	0,07	0,04
KSW15930.1	A0A0V8S6Q4	0,93	0,48
KSW16069.1	A0A0V8S6Z1	0,05	0,03
KSW16111.1	A0A0V8S708	0,13	0,07
KSW16071.1	A0A0V8S775	0,19	0,1
KSW16280.1	A0A0V8S7K3	0,78	0,4
KSW16352.1	A0A0V8S7Q7	2,55	1,3
KSW16351.1	A0A0V8S7R6	0,25	0,13
KSW16349.1	A0A0V8S7T2	0,84	0,43
KSW16378.1	A0A0V8S7Z4	0,06	0,03
KSW16426.1	A0A0V8S831	0,1	0,05

KSW16596.1	A0A0V8S8G3	0,29	0,15
KSW16984.1	A0A0V8S9J4	0,93	0,48
KSW16982.1	A0A0V8S9L5	0,9	0,46
KSW17208.1	A0A0V8SA68	1,51	0,77
KSW17322.1	A0A0V8SAR9	0,39	0,2
KSW17615.1	A0A0V8SBD0	0,16	0,08
KSW17684.1	A0A0V8SBK5	0,33	0,17
KSW17696.1	A0A0V8SBK8	0,78	0,4
KSW17692.1	A0A0V8SBL0	0,33	0,17
KSW17688.1	A0A0V8SBL5	0,59	0,3
KSW17699.1	A0A0V8SBL8	0,67	0,34
KSW17752.1	A0A0V8SBR8	0,11	0,05
KSW18201.1	A0A0V8SD03	0,19	0,1
KSW18204.1	A0A0V8SD14	0,14	0,07
KSW18203.1	A0A0V8SD15	0,12	0,06
KSW18199.1	A0A0V8SD16	0,11	0,06
KSW18289.1	A0A0V8SDB6	0,23	0,12
KSW18276.1	A0A0V8SDC1	0,18	0,09
KSW18312.1	A0A0V8SDD1	0,09	0,04
KSW18330.1	A0A0V8SDQ3	0,71	0,36
KSW18491.1	A0A0V8SDU0	1,15	0,59
KSW18495.1	A0A0V8SDU2	0,12	0,06
KSW18322.1	A0A0V8SDX3	0,64	0,33
KSW18565.1	A0A0V8SE56	0,39	0,2
KSW18655.1	A0A0V8SE95	0,21	0,11
KSW19115.1	A0A0V8SFL7	0,47	0,24
KSW18930.1	A0A0V8SFM4	0,07	0,04
KSW19409.1	A0A0V8SGG3	4,82	2,46
KSW19439.1	A0A0V8SGH6	0,27	0,14
KSW19475.1	A0A0V8SGN9	0,33	0,17
KSW19583.1	A0A0V8SGX2	0,11	0,06
KSW19597.1	A0A0V8SH13	0,29	0,15
KSW19607.1	A0A0V8SH26	0,08	0,04
KSW19791.1	A0A0V8SHI8	10,1	5,16
KSW19794.1	A0A0V8SI61	0,12	0,06
KSW20093.1	A0A0V8SIF4	0,67	0,34
KSW20191.1	A0A0V8SIQ9	0,29	0,15
KSW20219.1	A0A0V8SIS6	0,26	0,13
KSW20412.1	A0A0V8SJB3	0,17	0,08
KSW20550.1	A0A0V8SJQ2	0,39	0,2
KSW20684.1	A0A0V8SK26	0,78	0,4
KSW20681.1	A0A0V8SK30	0,47	0,24
KSW20692.1	A0A0V8SK31	0,87	0,45
KSW20704.1	A0A0V8SK41	0,72	0,37
KSW20697.1	A0A0V8SK42	0,52	0,27

KSW20700.1	A0A0V8SK48	0,93	0,48
KSW20691.1	A0A0V8SK50	0,59	0,3
KSW20708.1	A0A0V8SK53	1,15	0,59
KSW20689.1	A0A0V8SK55	0,78	0,4
KSW20702.1	A0A0V8SK56	0,47	0,24
KSW20713.1	A0A0V8SK59	0,22	0,11
KSW20707.1	A0A0V8SK66	1,37	0,7
KSW20716.1	A0A0V8SK67	4,18	2,14
KSW20710.1	A0A0V8SK68	0,23	0,12
KSW20701.1	A0A0V8SK70	0,5	0,26
KSW20705.1	A0A0V8SK71	1,37	0,7
KSW20696.1	A0A0V8SK72	0,47	0,24
KSW20695.1	A0A0V8SK73	0,78	0,4
KSW20736.1	A0A0V8SK77	0,87	0,45
KSW20723.1	A0A0V8SK82	0,19	0,1
KSW20726.1	A0A0V8SK84	0,12	0,06
KSW20717.1	A0A0V8SK85	0,59	0,3
KSW20712.1	A0A0V8SK93	0,23	0,12
KSW20735.1	A0A0V8SKA8	2,16	1,1
KSW20709.1	A0A0V8SKB9	0,12	0,06
KSW20674.1	A0A0V8SKL6	0,67	0,34
KSW20682.1	A0A0V8SKL8	0,55	0,28
KSW20698.1	A0A0V8SKN1	0,47	0,24
KSW20690.1	A0A0V8SKN7	0,16	0,08
KSW20706.1	A0A0V8SKQ6	0,85	0,43
KSW21020.1	A0A0V8SL08	0,16	0,08
KSW21040.1	A0A0V8SL26	0,1	0,05
KSW21039.1	A0A0V8SL49	0,15	0,07
KSW21018.1	A0A0V8SLK0	0,09	0,05
KSW21034.1	A0A0V8SLM5	0,17	0,08
KSW21230.1	A0A0V8SLN9	0,52	0,27
KSW21273.1	A0A0V8SLQ2	0,47	0,24
KSW21229.1	A0A0V8SLR7	0,39	0,2
KSW21272.1	A0A0V8SLX2	2,16	1,1
KSW21449.1	A0A0V8SM91	0,23	0,12
KSW21467.1	A0A0V8SM98	2,16	1,1
KSW21274.1	A0A0V8SMB6	0,22	0,11
KSW21470.1	A0A0V8SMC4	1,29	0,66
KSW21455.1	A0A0V8SMD3	0,32	0,17
KSW21507.1	A0A0V8SMF4	0,39	0,2
KSW21529.1	A0A0V8SMI0	0,33	0,17
KSW21527.1	A0A0V8SMJ2	0,29	0,15
KSW21559.1	A0A0V8SMM1	0,18	0,09
KSW21716.1	A0A0V8SN10	0,12	0,06
KSW21713.1	A0A0V8SN12	0,09	0,05

KSW21735.1	A0A0V8SN47	0,59	0,3
KSW21669.1	A0A0V8SN74	0,13	0,07
KSW21774.1	A0A0V8SN80	1,25	0,64
KSW21825.1	A0A0V8SNB8	0,25	0,13
KSW21865.1	A0A0V8SNF7	0,25	0,13
KSW21884.1	A0A0V8SNI3	0,47	0,24
KSW21879.1	A0A0V8SNI8	0,28	0,14
KSW21897.1	A0A0V8SNI9	0,7	0,36
KSW21886.1	A0A0V8SNJ1	0,14	0,07
KSW21872.1	A0A0V8SNK9	0,1	0,05
KSW21929.1	A0A0V8SNM2	0,15	0,07
KSW15219.1/KSW21896.1	A0A0V8SNP3	0,17	0,09
KSW21936.1	A0A0V8SNR7	0,28	0,14
KSW21893.1	A0A0V8SNT8	0,23	0,12
KSW21834.1	A0A0V8SNX9	1	0,51
KSW21866.1	A0A0V8SP17	4,01	2,05
KSW23491.1	A0A0V8ST30	0,64	0,33
KSW23552.1	A0A0V8ST89	1,85	0,94
KSW23702.1	A0A0V8STQ3	0,09	0,05
KSW23699.1	A0A0V8STR1	0,47	0,24
KSW23701.1	A0A0V8STW7	0,52	0,27
KSW23779.1	A0A0V8STY6	0,19	0,1
KSW23782.1	A0A0V8STY7	0,93	0,48
KSW23777.1	A0A0V8STZ4	0,39	0,2
KSW23794.1	A0A0V8SU61	0,13	0,07
KSW28201.1	A0A0V8T686	0,47	0,24
KSW28178.1	A0A0V8T6B7	0,26	0,13
KSW28271.1	A0A0V8T6P4	0,08	0,04
KSW28323.1	A0A0V8T6Q5	0,1	0,05
KSW28272.1	A0A0V8T6T2	0,12	0,06
KSW28366.1	A0A0V8T6V5	0,11	0,06
KSW28328.1	A0A0V8T700	0,32	0,16
KSW28387.1	A0A0V8T703	1,68	0,86
KSW28407.1	A0A0V8T723	0,12	0,06
KSW28273.1	A0A0V8T734	0,26	0,13
KSW28373.1	A0A0V8T737	0,17	0,09
KSW28305.1	A0A0V8T764	0,59	0,3
KSW28601.1	A0A0V8T771	1,1	0,56
KSW28372.1	A0A0V8T772	1,21	0,62
KSW28310.1	A0A0V8T7E0	0,89	0,46
KSW28729.1	A0A0V8T7H1	1,15	0,59
KSW28611.1	A0A0V8T7H6	0,2	0,1
KSW28552.1	A0A0V8T7K3	0,78	0,4
KSW28626.1	A0A0V8T7L4	0,08	0,04
KSW28374.1	A0A0V8T7N1	0,14	0,07

KSW28600.1	A0A0V8T7P0	0,21	0,11
KSW28616.1	A0A0V8T7Q0	0,15	0,07
KSW28603.1	A0A0V8T7R2	0,1	0,05
KSW28556.1	A0A0V8T7U2	0,15	0,07
KSW28591.1	A0A0V8T7Y8	0,32	0,16
KSW28620.1	A0A0V8T7Z7	0,27	0,14
KSW28637.1	A0A0V8T865	0,33	0,17
KSW29001.1	A0A0V8T881	18,31	9,35
KSW28630.1	A0A0V8T8A7	0,25	0,13
KSW28378.1	A0A0V8T8B4	0,07	0,04
KSW28646.1	A0A0V8T8C3	0,1	0,05
KSW28808.1	A0A0V8T8D2	1,28	0,65
KSW28914.1	A0A0V8T8E8	0,12	0,06
KSW28963.1	A0A0V8T8K6	0,25	0,13
KSW29054.1	A0A0V8T8P1	1,68	0,86
KSW29070.1	A0A0V8T8Q7	0,21	0,11
KSW28554.1	A0A0V8T8R8	0,36	0,18
KSW29118.1	A0A0V8T8U7	0,07	0,03
KSW28925.1	A0A0V8T8W5	0,14	0,07
KSW29166.1	A0A0V8T8X7	0,17	0,08
KSW29108.1	A0A0V8T913	1,85	0,94
KSW28618.1	A0A0V8T919	0,33	0,17
KSW29230.1	A0A0V8T922	0,78	0,4
KSW29048.1	A0A0V8T937	0,47	0,24
KSW29090.1	A0A0V8T953	0,04	0,02
KSW29093.1	A0A0V8T997	0,36	0,18
KSW29132.1	A0A0V8T9B3	0,23	0,12
KSW29148.1	A0A0V8T9C8	0,47	0,24
KSW29406.1	A0A0V8T9H8	1,15	0,59
KSW29200.1	A0A0V8T9I6	5,31	2,71
KSW29224.1	A0A0V8T9J2	0,39	0,2
KSW29110.1	A0A0V8T9J8	0,83	0,42
KSW29232.1	A0A0V8T9K7	0,52	0,27
KSW29454.1	A0A0V8T9M1	1,1	0,56
KSW29185.1	A0A0V8T9Q9	0,47	0,24
KSW29229.1	A0A0V8T9R1	0,33	0,17
KSW29261.1	A0A0V8T9T5	0,43	0,22
KSW29460.1	A0A0V8T9W0	9,97	5,09
KSW29330.1	A0A0V8T9W1	0,74	0,38
KSW29287.1	A0A0V8T9Z0	0,12	0,06
KSW29378.1	A0A0V8TA08	0,23	0,12
KSW29333.1	A0A0V8TA14	0,59	0,3
KSW29286.1	A0A0V8TA15	0,26	0,13
KSW29410.1	A0A0V8TA38	0,78	0,4
KSW29539.1	A0A0V8TA57	0,34	0,18

KSW29389.1	A0A0V8TA61	1,42	0,73
KSW29367.1	A0A0V8TA69	0,29	0,15
KSW29503.1	A0A0V8TA84	0,14	0,07
KSW29614.1	A0A0V8TA87	0,23	0,12
KSW29437.1	A0A0V8TAA0	0,47	0,24
KSW29082.1	A0A0V8TAB3	0,97	0,49
KSW29560.1	A0A0V8TAD3	0,18	0,09
KSW29516.1	A0A0V8TAE3	3,16	1,61
KSW29457.1	A0A0V8TAE5	0,08	0,04
KSW29710.1	A0A0V8TAE8	0,06	0,03
KSW29726.1	A0A0V8TAG7	0,06	0,03
KSW29473.1	A0A0V8TAG8	1,25	0,64
KSW29700.1	A0A0V8TAK9	0,5	0,26
KSW29732.1	A0A0V8TAN9	0,38	0,19
KSW29527.1	A0A0V8TAP7	0,26	0,13
KSW29822.1	A0A0V8TAQ2	0,54	0,28
KSW29644.1	A0A0V8TAQ4	0,1	0,05
KSW29600.1	A0A0V8TAS1	0,08	0,04
KSW29785.1	A0A0V8TAS3	0,26	0,13
KSW29693.1	A0A0V8TAX2	0,07	0,04
KSW29589.1	A0A0V8TAZ1	2,79	1,43
KSW29696.1	A0A0V8TB09	0,19	0,1
KSW29832.1	A0A0V8TB42	0,36	0,18
KSW29892.1	A0A0V8TB43	4,88	2,49
KSW29966.1	A0A0V8TB53	0,78	0,4
KSW29929.1	A0A0V8TB59	0,1	0,05
KSW29945.1	A0A0V8TB72	0,16	0,08
KSW29998.1	A0A0V8TB93	0,47	0,24
KSW29842.1	A0A0V8TBB6	0,06	0,03
KSW29823.1	A0A0V8TBB7	0,17	0,08
KSW29884.1	A0A0V8TBC4	0,23	0,12
KSW29799.1	A0A0V8TBD9	0,93	0,48
KSW29871.1	A0A0V8TBE7	2,31	1,18
KSW29887.1	A0A0V8TBG6	0,69	0,35
KSW29903.1	A0A0V8TBH7	0,47	0,24
KSW29872.1	A0A0V8TBJ4	0,13	0,07
KSW30105.1	A0A0V8TBK0	0,06	0,03
KSW29798.1	A0A0V8TBK4	0,47	0,24
KSW29922.1	A0A0V8TBM6	0,13	0,06
KSW30206.1	A0A0V8TBN8	0,33	0,17
KSW29938.1	A0A0V8TBP3	0,44	0,22
KSW29514.1	A0A0V8TBP7	0,21	0,11
KSW30107.1	A0A0V8TBQ0	0,09	0,05
KSW30163.1	A0A0V8TBX1	0,07	0,03
KSW30219.1	A0A0V8TBY5	0,16	0,08

KSW29942.1	A0A0V8TBY6	0,09	0,05
KSW30211.1	A0A0V8TC22	0,17	0,08
KSW30082.1	A0A0V8TC27	0,16	0,08
KSW29974.1	A0A0V8TC37	0,16	0,08
KSW29690.1	A0A0V8TC65	0,27	0,14
KSW30037.1	A0A0V8TC78	0,21	0,11
KSW30127.1	A0A0V8TC84	0,09	0,04
KSW30054.1	A0A0V8TC91	0,16	0,08
KSW30162.1	A0A0V8TC98	0,33	0,17
KSW30135.1	A0A0V8TCB0	0,33	0,17
KSW30049.1	A0A0V8TCB4	0,59	0,3
KSW30260.1	A0A0V8TCB7	0,21	0,11
KSW30160.1	A0A0V8TCD5	0,11	0,05
KSW30101.1	A0A0V8TCD6	0,34	0,17
KSW30205.1	A0A0V8TCE0	0,31	0,16
KSW29834.1	A0A0V8TCL7	0,21	0,11
KSW30256.1	A0A0V8TCL8	0,17	0,08
KSW30161.1	A0A0V8TCM1	0,91	0,46
KSW30268.1	A0A0V8TCM5	0,37	0,19
KSW30322.1	A0A0V8TCP4	0,27	0,14
KSW30263.1	A0A0V8TCQ8	0,04	0,02
KSW30262.1	A0A0V8TCR1	0,39	0,2
KSW29946.1	A0A0V8TCV4	0,54	0,28
KSW30273.1	A0A0V8TD39	0,23	0,11
SAC01			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAI%
KSW29460.1	A0A0V8T9W0	12,18	20,73
KSW14037.1	A0A0V8S162	6,36	10,82
KSW23497.1	A0A0V8ST45	4,18	7,11
KSW29472.1	A0A0V8TAF4	3,64	6,2
KSW29690.1	A0A0V8TC65	2,79	4,75
KSW29589.1	A0A0V8TAZ1	2,36	4,02
KSW21841.1	A0A0V8SND4	2,16	3,68
KSW29363.1	A0A0V8T9P2	2,16	3,68
KSW18322.1	A0A0V8SDX3	1,68	2,86
KSW29001.1	A0A0V8T881	1,68	2,86
KSW16380.1	A0A0V8S7S7	1,46	2,49
KSW18201.1	A0A0V8SD03	1,42	2,42
KSW29426.1	A0A0V8TA52	1,37	2,33
KSW18187.1	A0A0V8SD02	1,15	1,96
KSW29148.1	A0A0V8T9C8	1,15	1,96
KSW29700.1	A0A0V8TAK9	0,97	1,65
KSW23782.1	A0A0V8STY7	0,93	1,58
KSW19597.1	A0A0V8SH13	0,67	1,14

KSW13738.1	A0A0V8S0K7	0,59	1
KSW29759.1	A0A0V8TB55	0,59	1
KSW15558.1	A0A0V8S5H0	0,52	0,88
KSW15606.1	A0A0V8S5M2	0,47	0,8
KSW18316.1	A0A0V8SDC8	0,47	0,8
KSW29185.1	A0A0V8T9Q9	0,47	0,8
KSW29798.1	A0A0V8TBK4	0,47	0,8
KSW20716.1	A0A0V8SK67	0,39	0,66
KSW21507.1	A0A0V8SMF4	0,39	0,66
KSW29054.1	A0A0V8T8P1	0,39	0,66
KSW15047.1	A0A0V8S451	0,35	0,6
KSW21810.1	A0A0V8SNS7	0,33	0,57
KSW23774.1	A0A0V8SU20	0,29	0,5
KSW29367.1	A0A0V8TA69	0,29	0,5
KSW29908.1	A0A0V8TB56	0,28	0,48
KSW18291.1	A0A0V8SDC6	0,27	0,47
KSW29228.1	A0A0V8T9I9	0,27	0,45
KSW18194.1	A0A0V8SDJ3	0,26	0,44
KSW20691.1	A0A0V8SK50	0,26	0,44
KSW29200.1	A0A0V8T9I6	0,26	0,44
KSW29356.1	A0A0V8T9V7	0,26	0,44
KSW29132.1	A0A0V8T9B3	0,23	0,4
KSW14010.1	A0A0V8S1E5	0,21	0,36
KSW29461.1	A0A0V8TAG9	0,2	0,34
KSW15930.1	A0A0V8S6Q4	0,18	0,3
KSW20994.1	A0A0V8SLI0	0,17	0,28
KSW28807.1	A0A0V8T843	0,16	0,26
KSW19409.1	A0A0V8SGG3	0,15	0,25
KSW21232.1	A0A0V8SLN0	0,15	0,25
KSW28764.1	A0A0V8T8C4	0,15	0,25
KSW29756.1	A0A0V8TB21	0,15	0,25
KSW29616.1	A0A0V8TAT3	0,14	0,23
KSW30064.1	A0A0V8TC48	0,12	0,2
KSW16982.1	A0A0V8S9L5	0,11	0,19
KSW30160.1	A0A0V8TCD5	0,11	0,18
KSW28336.1/KSW29753.1	A0A0V8T6X7	0,1	0,17
KSW29466.1	A0A0V8TBH9	0,1	0,17
KSW29179.1	A0A0V8T9B7	0,08	0,13
KSW20189.1	A0A0V8SIQ3	0,08	0,13
KSW29732.1	A0A0V8TAN9	0,02	0,04
SAC02			
Acceso de GenBank	Acceso de Uniprot	emPAI	emPAI%
KSW29460.1	A0A0V8T9W0	2,31	78,95
KSW21563.1	A0A0V8SMI4	0,06	2,12

KSW29804.1	A0A0V8TB54	0,11	3,76
KSW29835.1	A0A0V8TB30	0,23	7,96
KSW29385.1	A0A0V8T9L6	0,12	4,17
KSW29030.1	A0A0V8T9E2	0,09	3,04