From metagenomic gene discovery to enzymatic breakdown of plant cell wall biopolymers

Dominic Wong

Bioproducts Chemistry & Engineering
Western Regional Research Center
Agricultural Research Service
U.S. Department of Agriculture
Albany, California

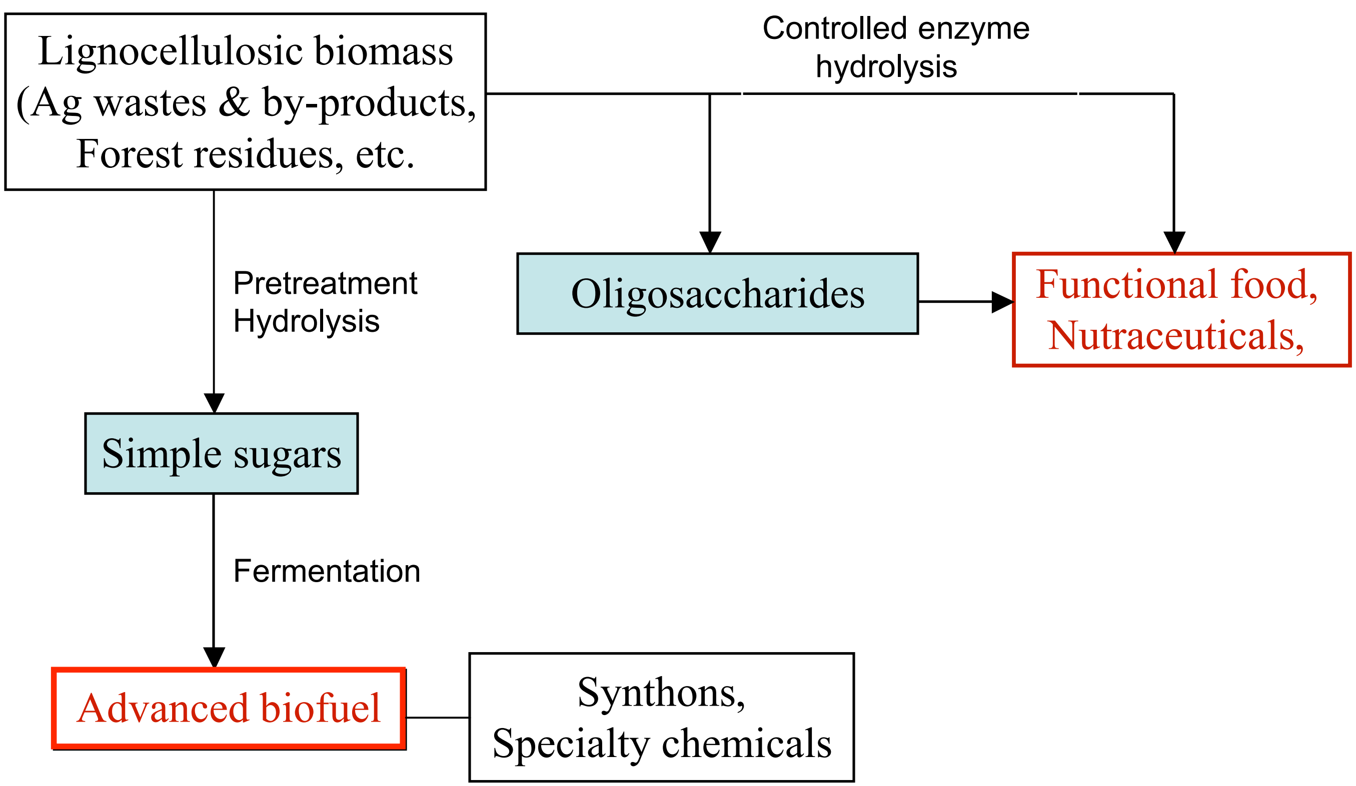
Research Objectives:

- *Research key enzymes and their collaborative mode of action on breaking down plant cell wall polymers.*

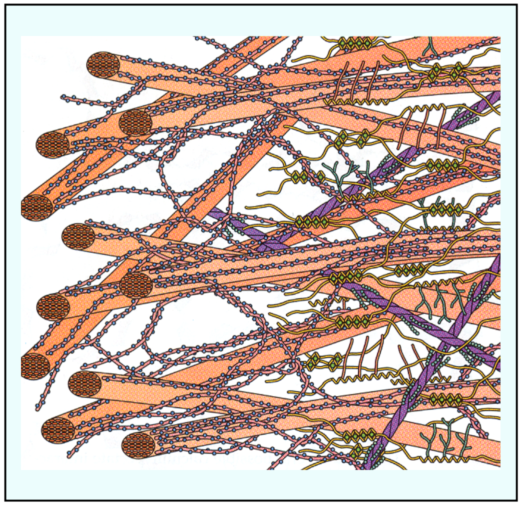
- *Research the efficacy of using hemicellulases and other non-cellulase enzymes prior to, during, or after pretreatment.*

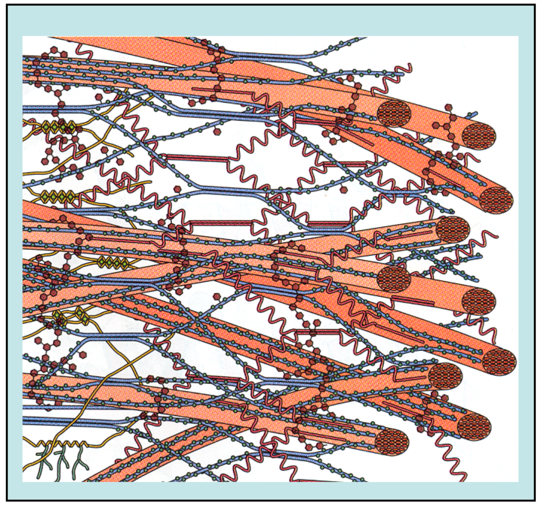
<u>For conversion of agricultural wastes to biofuel and high-value products</u>

Biomass conversion to high value products



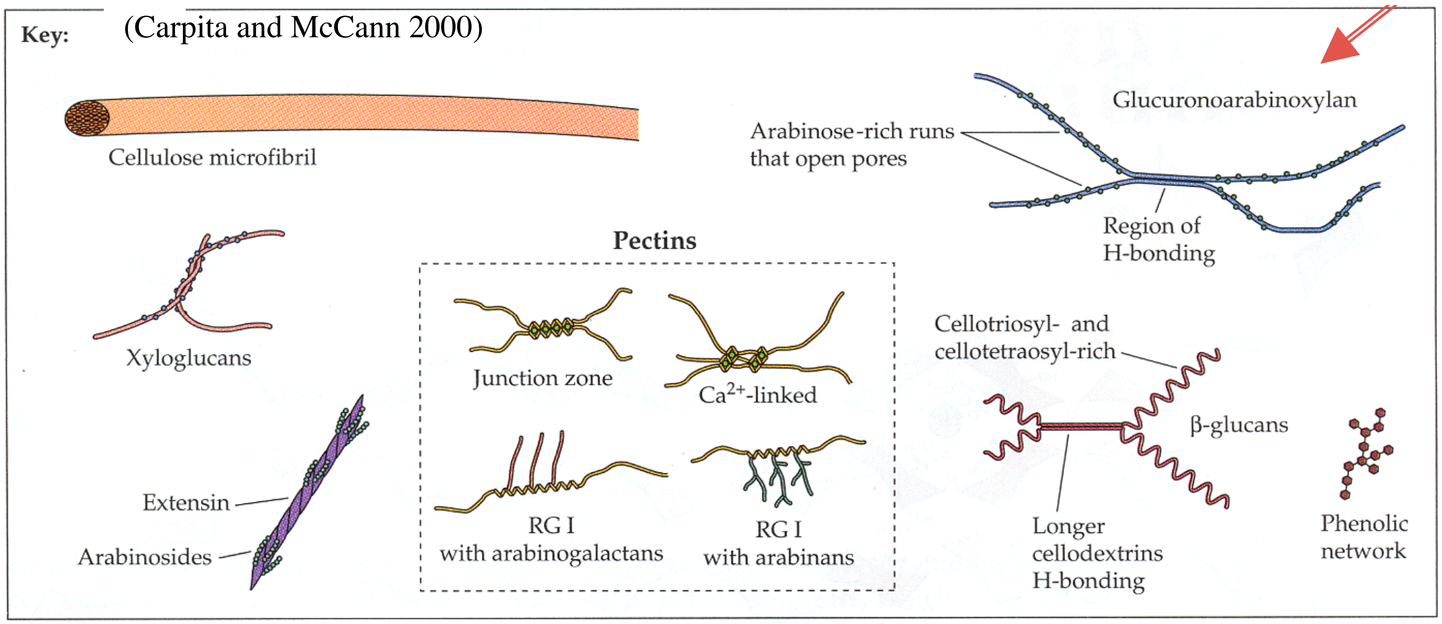
Cell wall polymers





Biomass source	Cellulose	Hemi-	Lignin
	(% dry weight)		
Agricultural residues	38	32	17
Corn fiber	15	35	8
Corn cob	45	35	15
Corn stover	40	25	17
Rice straw	35	25	12
Wheat straw	30	50	20
Sugarcane bagasse	40	24	25
Herbaceous energy crops	45	30	15
Switch grass	45	30	12
Coastal Bermuda grasss	25	35	6
Sorted municipal solid waste	45	9	10
Underutilized & short rotation hardwoods	50	23	22

(Saha 2003; Wyman 1994)



(Carpita and McCann 2000)

Type I -
All dicots & one-half of monocots (roses, sunflowers, maples, oaks beans, spinach, sugar beet)

Type II -
The "commelinoid" of monocots (palms, gingers, cypresses, and grasses- including grain crops)

Biomass Degradation *- Known enzymes found in nature*

Cellulose: (3)

endoglucanase, exoglucanase, cellobiosidase

Hemicellulose: (15):

Glucuronoarabinoxylan: (6) *endoxylanase, β-xylosidase, α-L-arabinoxylanosidase, α-glucuronisdase, acetylxylan esterase, feruloyl esterase*

Xyloglucan: (6) *Endoglucanase, β-glucosidase, xyloglucosidase, α-xylosidase, β-galactosidase, fucosidase*

β-Glucan: (3) *β-glucan endohydrolases, β-glucan exohydrolases, β-glucosidase*

Pectins: (17)

Hydrolases: (5) *endopolygalacturonase, exopolygalacturonase, rhamnogalacturonan hydrolase, rhamnogalacturonan rhamnohydrolase, xylogalacturonanhydrolase*

Lyases: (4) *endopectate lyase, exopectate lyase, pectin lyase (endo), rhamnogalacturonan lyase*

Esterases: (3) *pectin methyl esterase, pectin acetyl esterase, rhamnogalacturonan acetyl esterase*

Auxiliary enzymes: (5) *galactanase, arabinanase, β-galactosidase, α-L-furanosidase, feruloyl esterases*

Lignin: (3) *lignin peroxidase, manganese-dependent peroxidase, laccase*

Bioconversion of hemicellulose

Hemicellulose

enzymatic hydrolysis

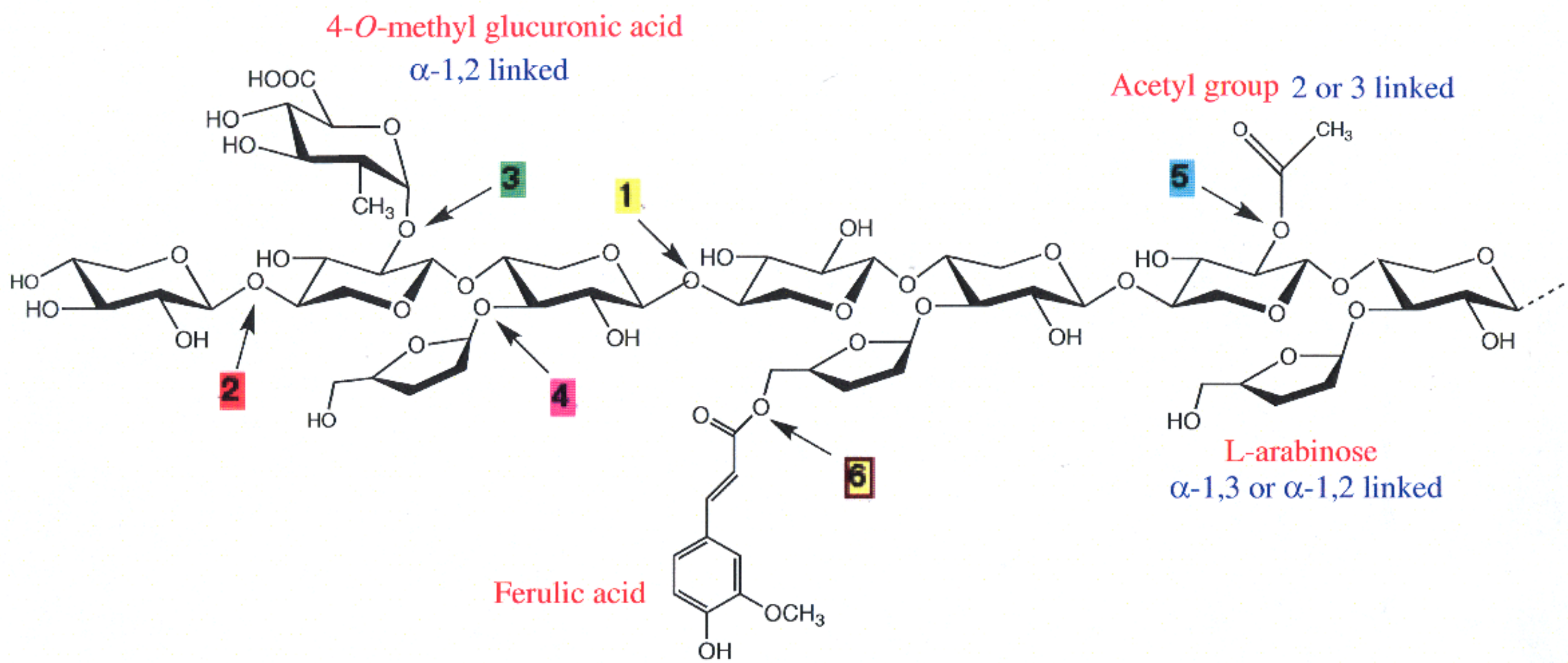
- Building blocks
  Monosaccharides
- Functional oligosaccharides
  Prebiotics
  Antioxidants

enzymatic/chemical modification

- Biodegradable composites
- Functional food additives
  Thickeners
  Gelling agents
  Coating agents

USDA WRRC

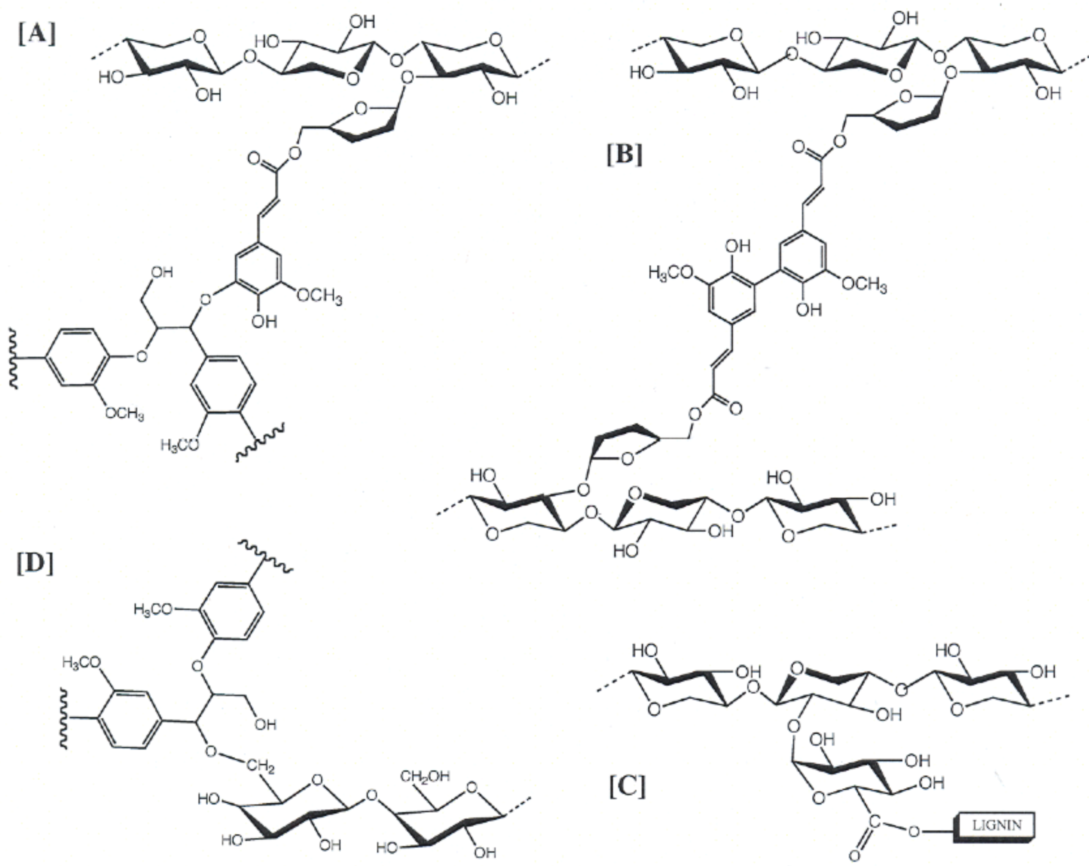
Glucuronoarabinoxylan



1 endo-β-1,4-xylanase (EC3.2.1.8)
2 β-xylosidase (EC 3.2.1.37)
3 α-glucuronidase (EC 3.2.1.139)
4 α-L-arabinofuranosidase (EC 3.2.1.55)
5 acetylxylan esterase (EC 3.1.1.72)
6 feruloyl esterase (EC 3.1.1.79)

USDA WRRC

Chemical structures of crosslinks



(A) xylan-ferulic acid-lignin; (B) arabinoxylan-diferulic acid-arabinoxylan
(C) xylan-glucuronic acid-lignin; (D) pectin-galactan-lignin

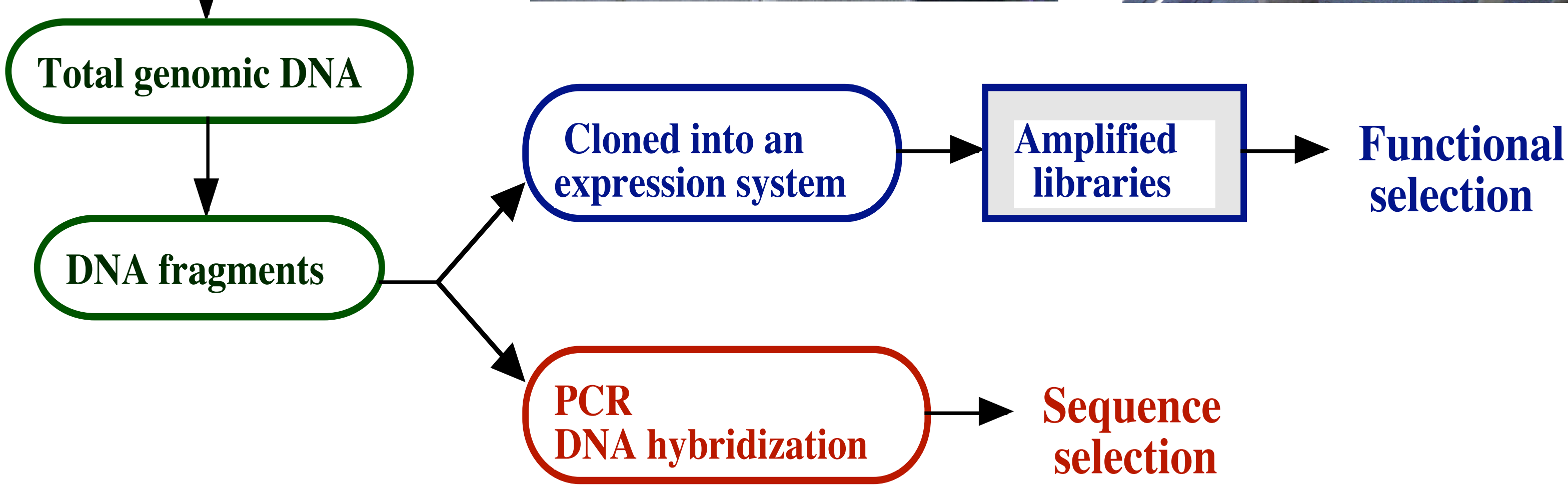
Covalent crosslinks of cell wall polymers

Polymer	Crosslink	Polymer	Bond types	References
arabinoxylan	diferulic acid	arabinoxylan	ester	Markwalder and Neukom 1975
arabinoxylan	diferulic acid	pectic galactan	ester	Saulnier and Thibault 1999
arabinoxylan	xylose C2/C3-OH	lignin	ether	Watanabe et al. 1989; Kishimoto et al. 2002
arabinoxylan	ferulic acid	lignin	ester/ether	Scalbert et al. 1985; Iiyama et al. 1990
arabinoxylan	triferulic/tetraferulic acid	arabinoxylan	ester	Fry et al. 2000; Bunzel et al. 2006
glucuronoxylan	4-*O*-methyl-D-glucuronic acid	lignin	ester	Watanabe and Koshijima 1988
homogalacturonan	galacturonic acid	polysaccharide	ester	Brown and fry 1993
rhamnogalacturonan I	galactan/arabinan	xyloglucan	ether?	Keegstra et al. 1973; Brett et al. 2005
β-1,4-galactan	galactose C6-OH	lignin	ether	Watanabe et al. 1989
pectic arabinan	diferulic acid	pectic galactan	ester	Colquhoun et al. 1994; Ralet et al. 2005
pectic polysaccharide	arabinan	protein	glycosidic	Keegstra et al. 1973
cellulose	xyloglucan	cellulose		Pauly et al. 1999
acetylglucomannan	glucose C6-OH	lignin	ether	Watanabe et al. 1989
galactomannan	galactose C3-OH	lignin	ether	Ericksson et al. 1980
protein	tyrosine/Cysteine	lignin/Protein	ether	Iiyama et al. 1994; Oudgenoeg et al. 2002
polysaccharide	D-*lyxo*-5-hexosulopyranuronic acid	polysaccharide	glycosidic	Painter 1983

Note: In some cases, the "linkage group" (the chemical fragment bearing the elements of the polymeric structure joint by the covalent linkage) has been actually isolated and structurally identified.

Novel Genes from Metagenomes

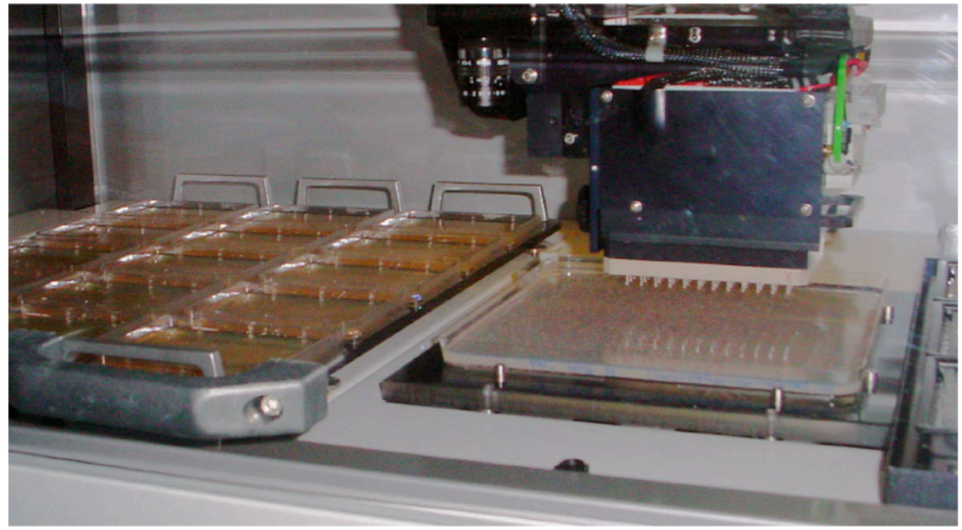
“Metagenome” = the collective genomes of all microorganisms in a given habitat (Handerisman et al. 1998; Schmidt et al. 1991)



D. Wog (2010) *In*: Metagenomics: Theory, Methods, and Applications

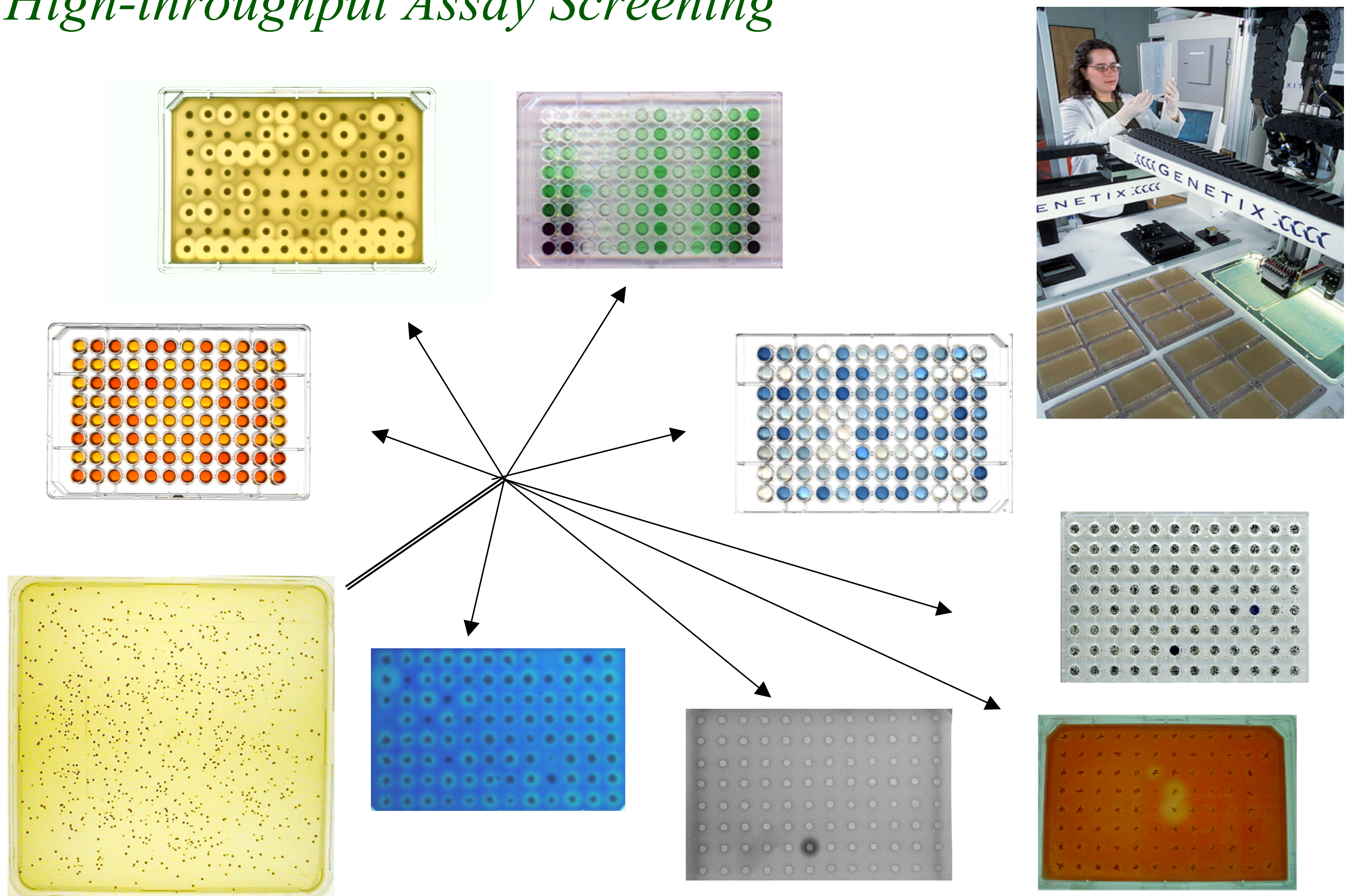
Q-Bot

Q-Pix



USDA WRRC

High-throughput Assay Screening



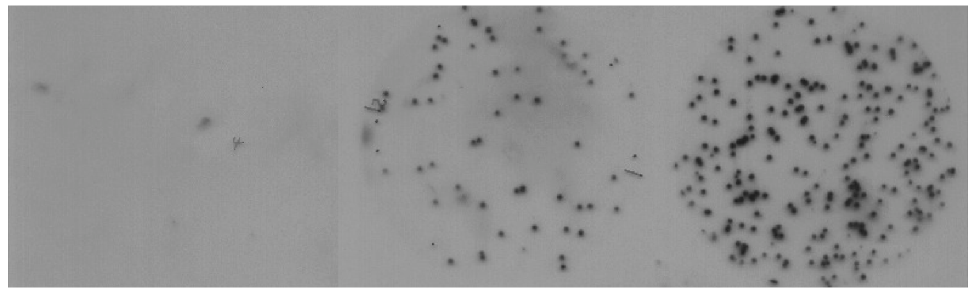
Q-tray

Wong et al. US Patent 8,361,764 B1

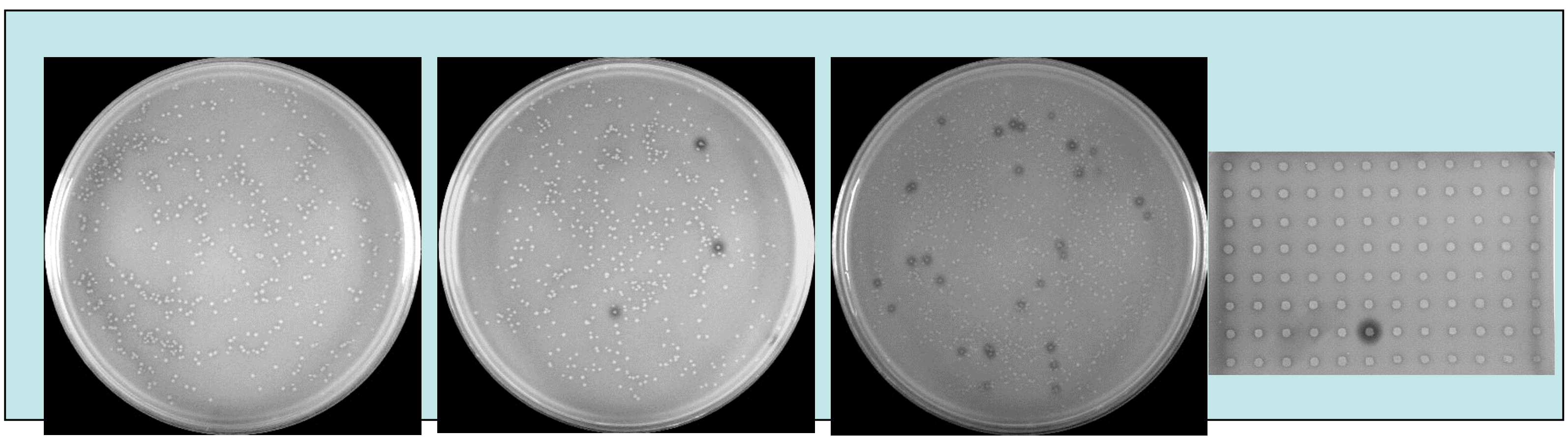
High-throughput Gene Screening

Screening for xylanase genes from metagenomic libraries

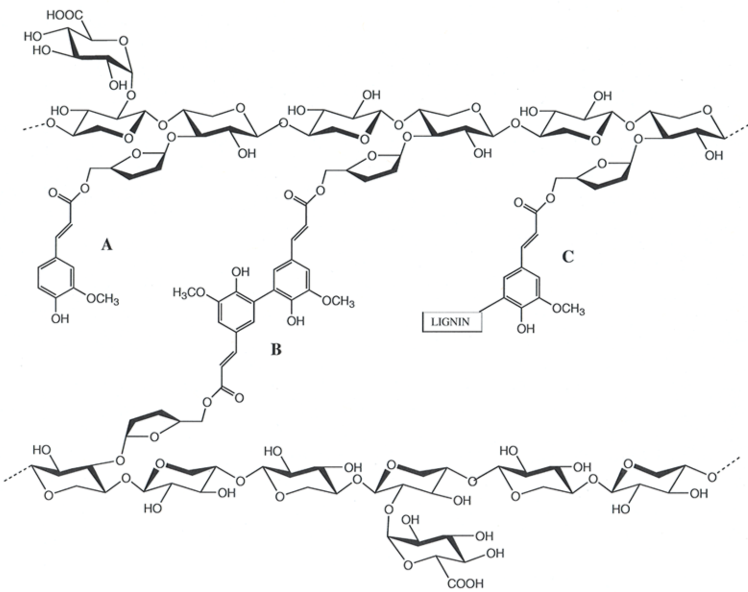
Screening for amylase genes from cDNA (λZapII) libraries

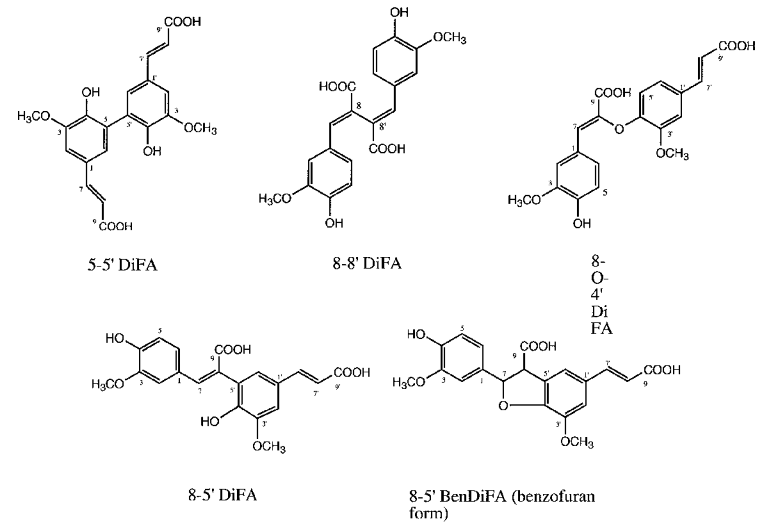


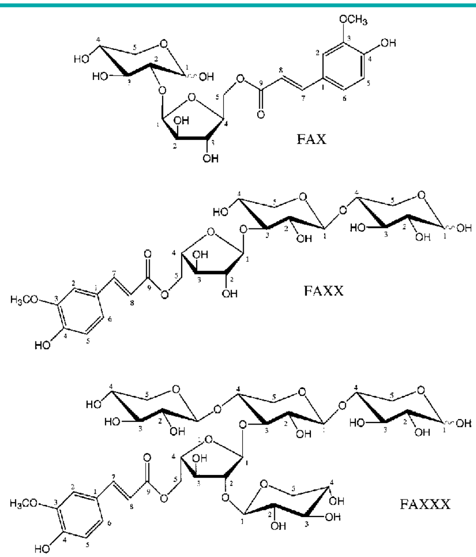
Screening for feruloyl esterase genes from metagenomic libraries

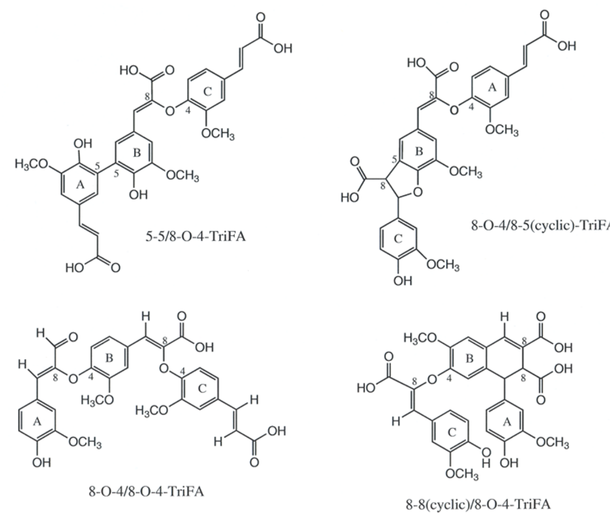


Feruloylation of arabinoxylans



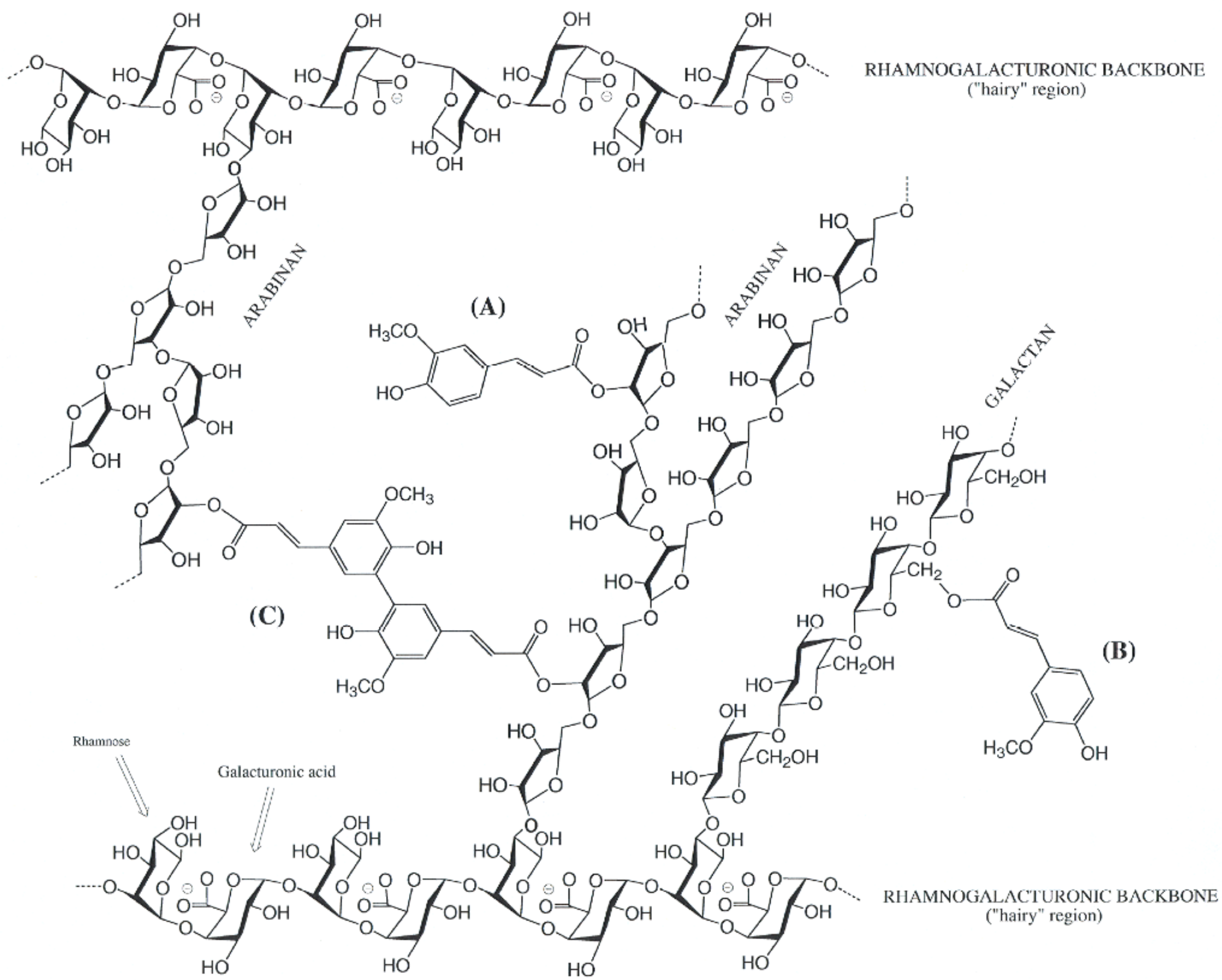






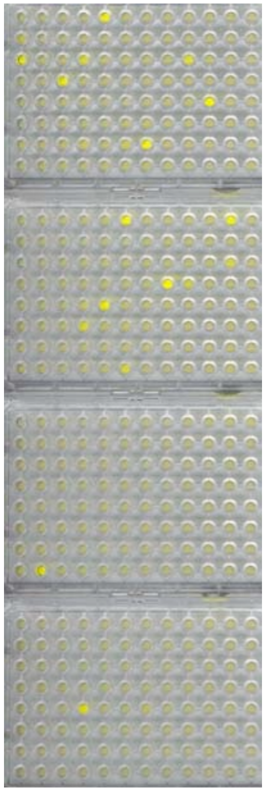
D. Wong (2006) (2009) Applied Biochem. Biotechnol

Feruloylation of rhamnified pectins



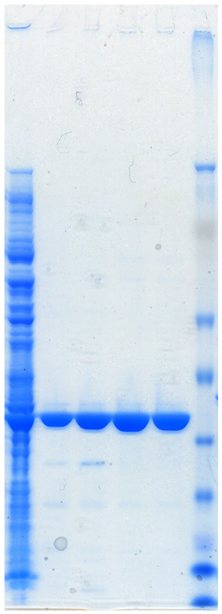
D. Wong (2008) Protein Journal

FAE gene isolation and expression

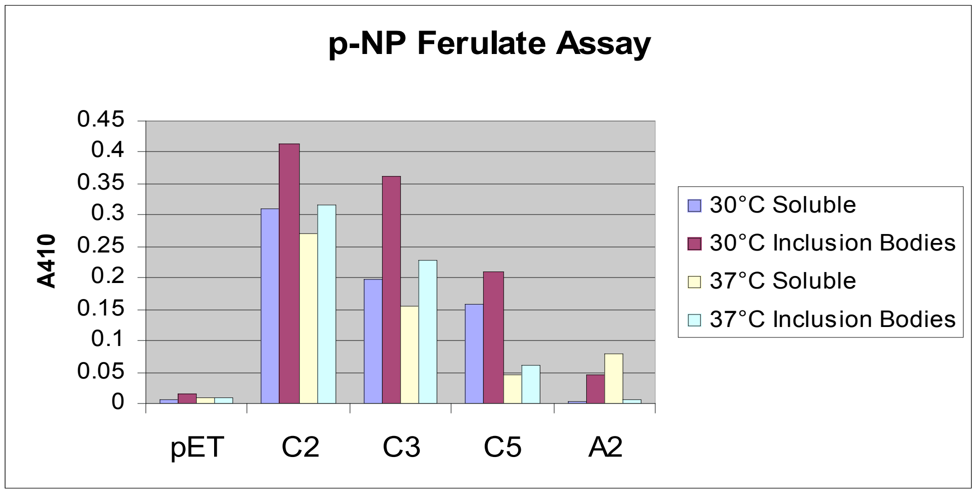


High-throughput screening

Rescreen positives on ethyl ferulate



pET vector/ *E. coli* BL21



Wong et al. (2013) US Patent 8,361,764 B1

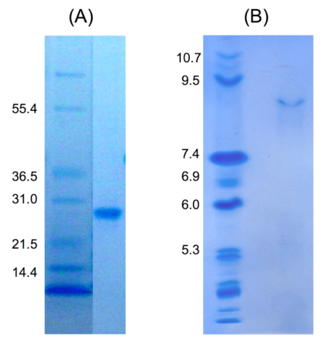
FAE genes and proteins characterization						
Clone Name	Gene Location (Insert Size)	Reference Sequence	% Identity to Ref Seq	Number of Residues (Includes tag)	SEQ ID NO	Recombinant Size (kDa)
Rum-C1	1811-2497 (687 bp)	ZP_01061020.1 Hydrolase of ‾ Family	45	229	1	25.2
Rum-C2	587-1396 (810 bp)	ZP_01061020.1 Hydrolase of ‾ Family	45	269	2	29.7
Rum-C3	663-1343 (681 bp)	ZP_01061020.1 Hydrolase of ‾ Family	46	381	3	41.9
Rum6-C4	112-831 (720 bp)	AAF70241 Feruloyl Esterase A	41	246	4	27.1
Rum-C5	60-1001 (942 bp)	ZP_01122099.1 probable lipase/esterase	32	416	5	45
Rum-C6	2871-3827 (957 bp)	YP_169760.1 hypothetical protein lipase/esterase	34	422	6	45.6
Rum-C7	447-1277 (831 bp)	ZP_01777732 putative esterase	48	277	7	32.3
Rum-C8	1849-3207 (1359 bp)	ZP_01061020.1 Hydrolase of ‾ Family	47	382	8	41.9
Rum-C9	66-1403 (1338 bp)	ZP_01061020.1 Hydrolase of ‾ Family	47	383	9	41.9
Mul-C10	181-1281 (1101 bp)	ZP_01421329 Beta-lactamase	44	367	10	39.7
Mul-C11	116-1057 (942 bp)	YP_583166 alpha/beta hydrolase	54	322	11	34.7
Mul-C12	1495-2412 (918 bp)	YP_578455 metallophosphoesterase	43	319	12	34.8

Genbank 15, WRRC 12

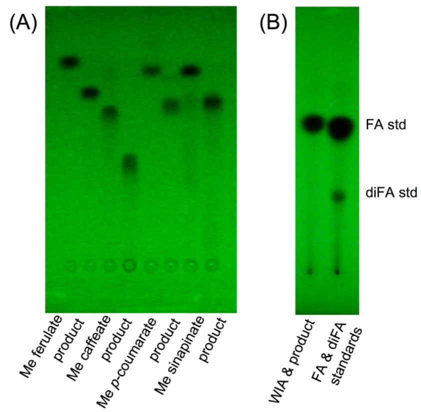
Wong et al. (2013) US Patent 8,361,764 B1

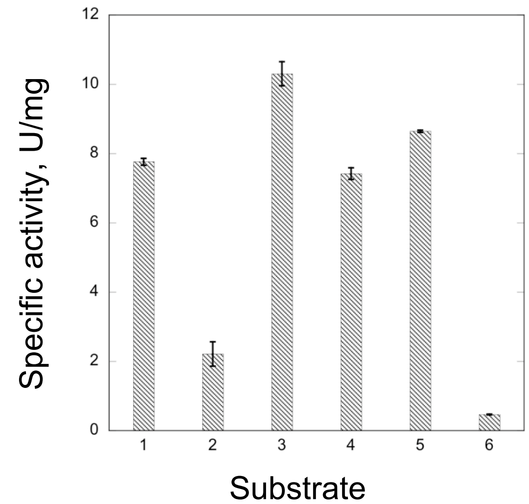
Novel Feruloyl Esterases - Phylogenetic Tree

Fae clone 2 enzyme characterization



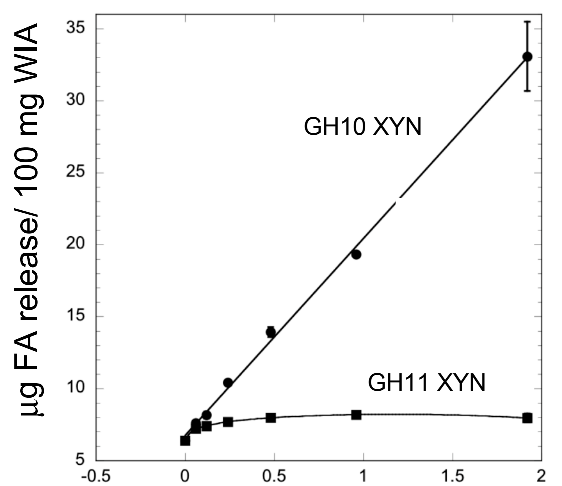
SDS-PAGE IEF



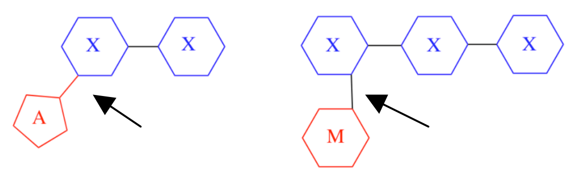


1. Corn fiber
2. Corn bran
3. Wheat bran
4. Wheat arabinoxylan
5. Rice bran
6. switchgrass

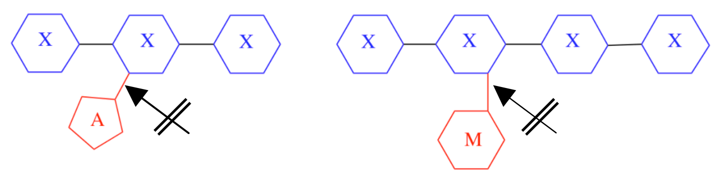
Synergistic effect of endoxylanase on Fae release of FA



nmole xylanase added with esterase



GH10 hydrolysis products with end-linked side groups



GH11 products : internal side groups

The Fae 2 enzyme can hydrolyze only or at a much higher rate araf-ferulic acid side groups linked to the non-reducing end.

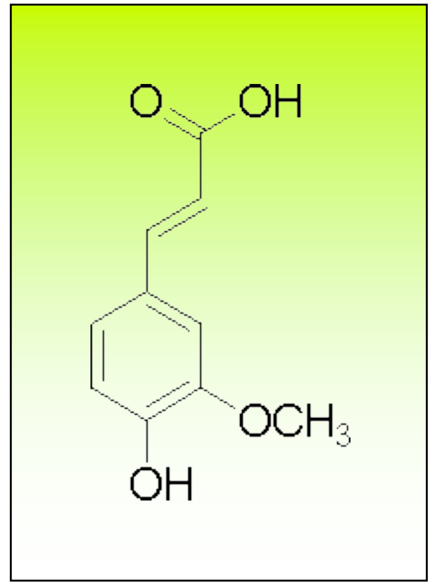
Wong et al. (2013) J. Ind. Microbiol. Biotechnol.

Ferulic acid and derivatives

Properties:
- improve anti-oxidative activity
- increase lipid solubility
- enhance gelling effect

Applications:
- protect against photo-oxidative damage
- enhance wound management aid
- synthesis of vanillic acid
- produce thermal stable biogel

Release of ferulic acid:
- prevent filter blockage in brewing
- increase loaf volume in baking
- increase biomass degradation



Barley: 0.14% dry w/w

Wheat bran: 0.66%

Sugar beet pulp: 0.8%

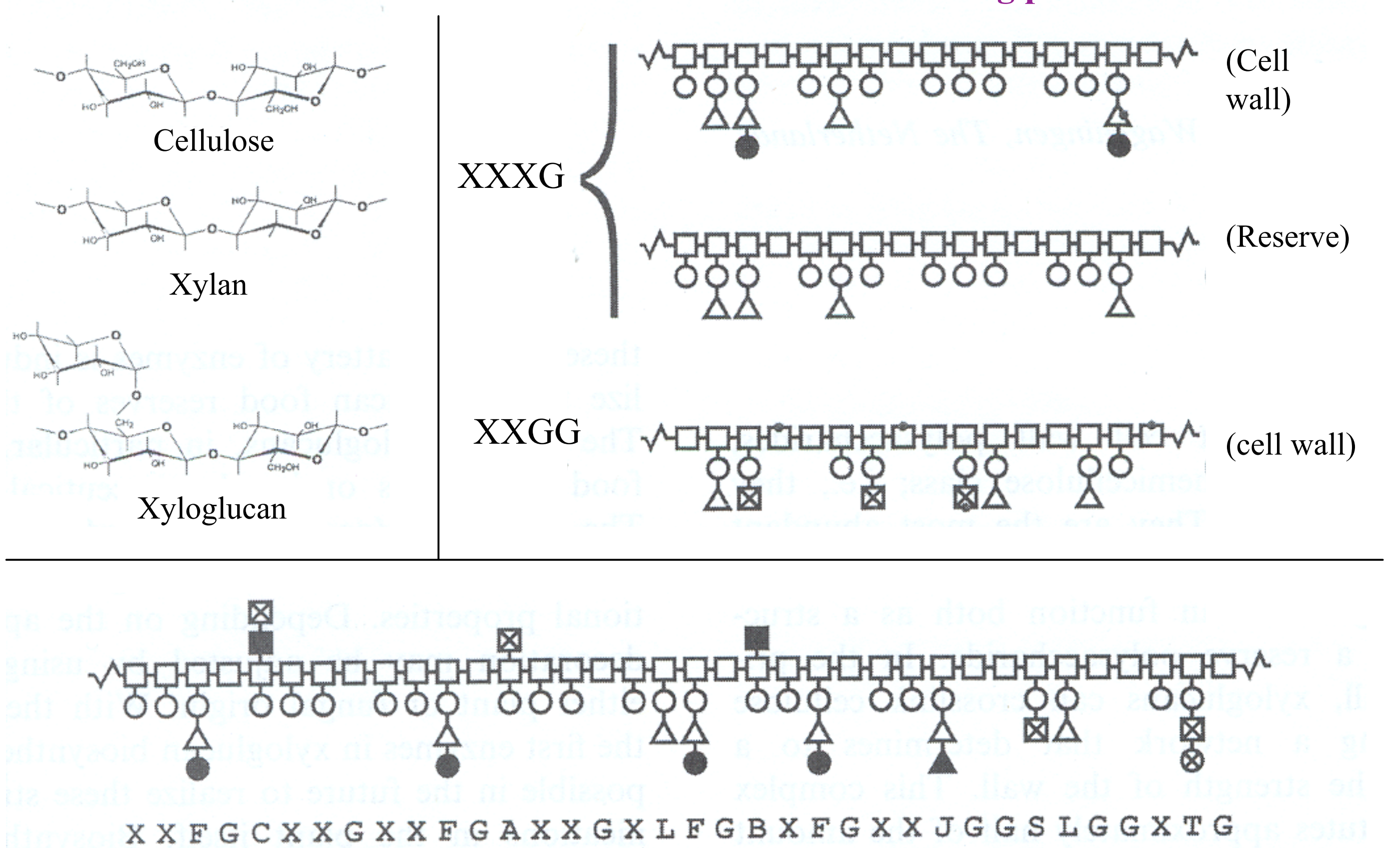
Corn bran: 3.1%

Corn fiber: 3-6%

Switchgrass: 0.3-0.5%

Xyloglucan

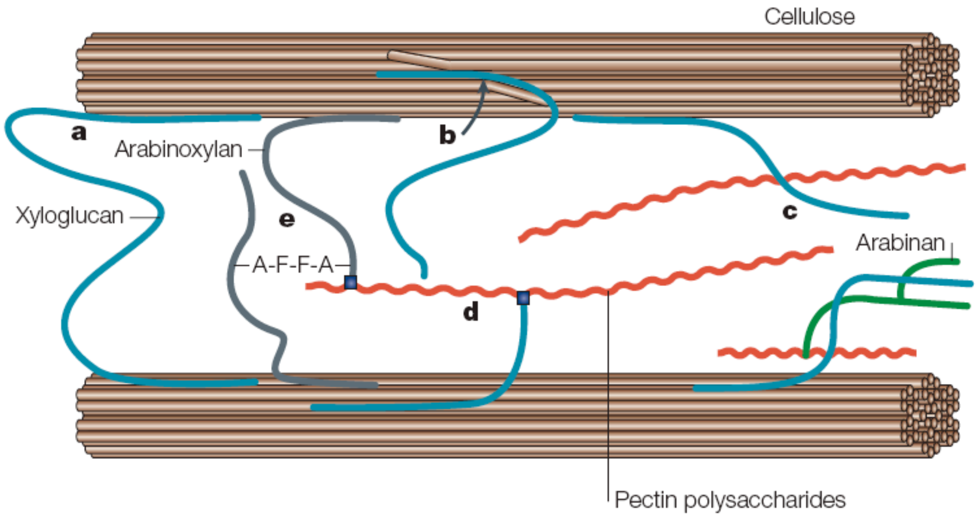
General branching pattern



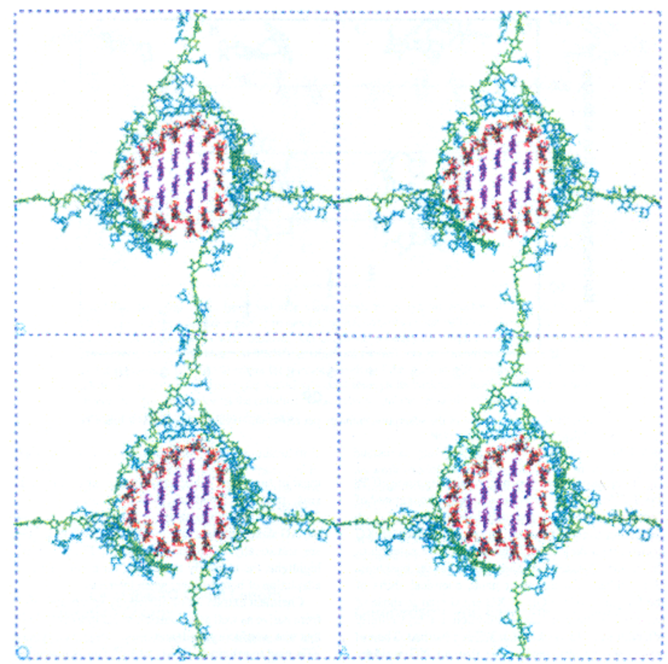
Hypothetical xyloglucan

Vincken (2003) In: *Handbook of Food Enzymology*

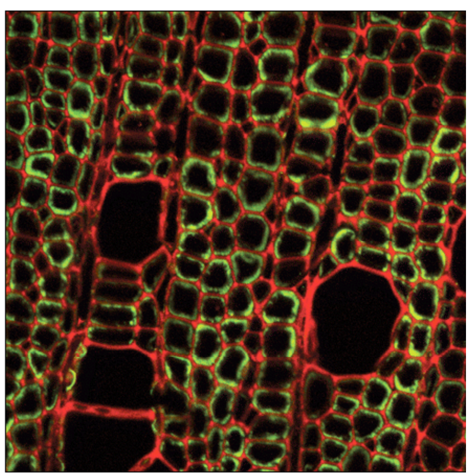
Xyloglucan in cell wall



Cosgrove, D. Nature Rev. Mol. Cell Biol. 2005



Hanus, J & Mazeau, K. Biopolymers 2006

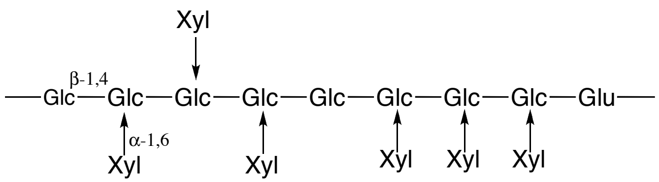


Mellerowicz et al. Ann. Bot. 2008

XEG5A: xyloglucan-specific endo-β-1,4-glucanase

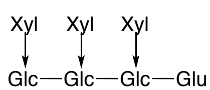
Substrate	**Relative activity**
Tamarind xylogluca n	5.6
HDP xyloglucan oligosaccharide	2.6
CMC	1.0
Cellulose, xylan, arabinoxylan, pectin	n.d.

Tamarind xyloglucan



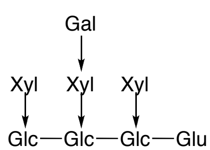
Blocks

(1)

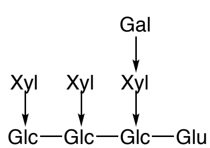


XXXG (heptasaccharide), XG7, DP7

(2)

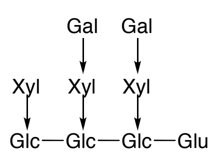


XLXG Ooctasaccharide), XG8, DP8



XXLG Ooctasaccharide), XG8, DP8

(3)

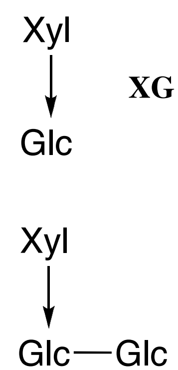


XXLG (nonasaccharide). XG9. DP9

Tamarind xyloglucan *or*
HDP xyloglucan oligosaccharide

$\downarrow$ XEG5A

XXXG, XLXG/XXLG, XLLG

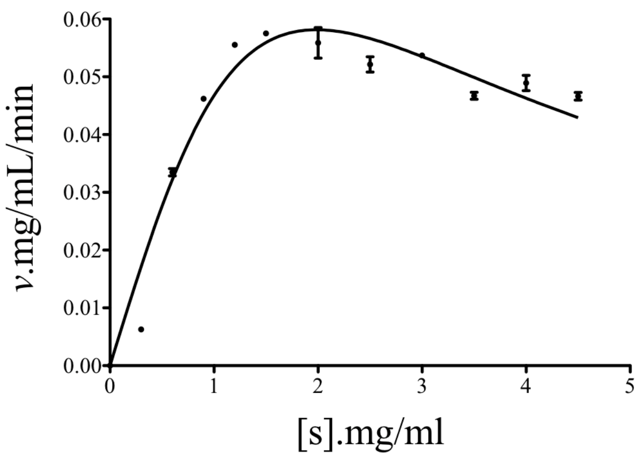


XG (isoprimeverose); XG2, DP2

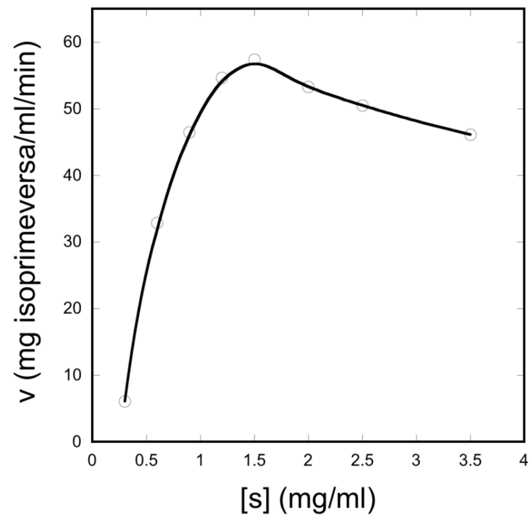
XGG (xylosyl-cellobiose); XG3, DP3

Wong et al. (2010) *Applied Microbiol. Biotechnol.*

XEG5A: xyloglucan-specific endo-β-1,4-glucanase

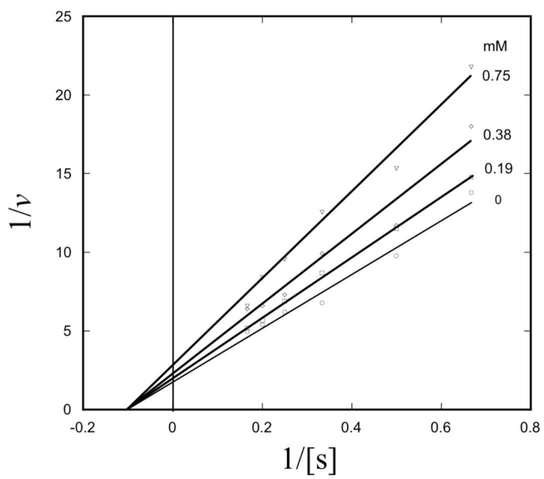


Xyloglucan oligosaccharide HDP-XGP

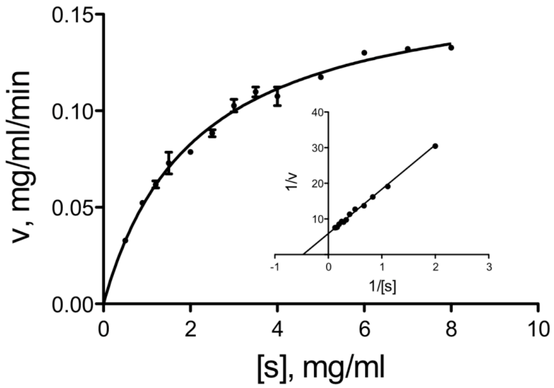


Xyloglucan oligosaccharide DP14

Xyloglucan oligosaccharid DP7 (XXXG):
K_i = 1.46 ± 0.13 mM



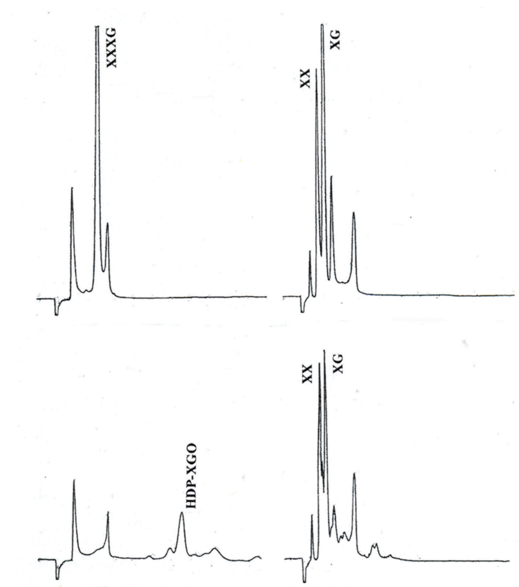
XEG5B: xyloglucan-specific exo-β-1,4-glucanase [EC 3.2.1.155]



K_m = 2.12 ± 0.13
K_{cat} = 47.6 min^{-1}
(Tamarind xyloglucan, 37^{o}C, pH 6.0)

Active only on xyloglucan & HDP-XGO
Activity not detected -CMC, Avicel, acid cellulose, lichenan, laminarin, xylan, arabinoxylan

HPLC analysis of end-products



Tamarind xyloglucan *or*
HDP xyloglucan oligosaccharide

↓ XEG5B

XXXG, XLXG/XXLG, XLLG ?

↓

Xyl
↓
Glc

isoprimeverose; XG2, DP2

Xyl
↓
Glc—Glc

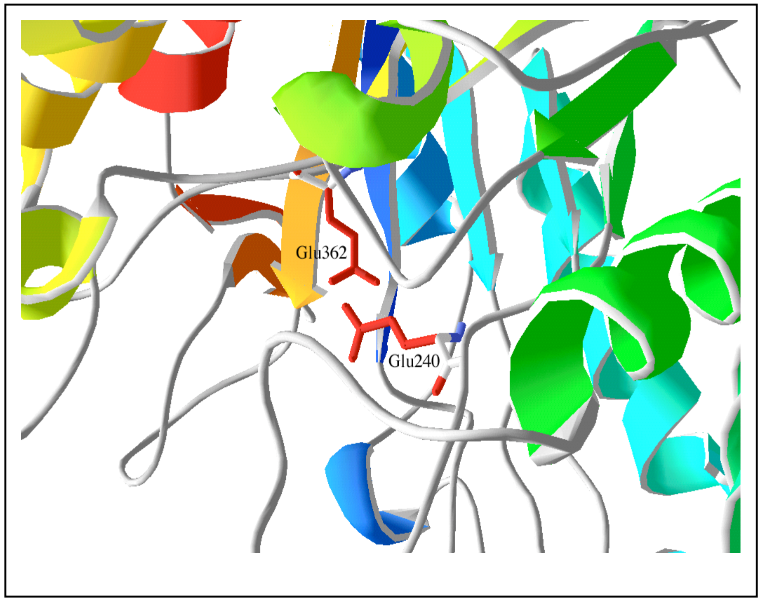
xylosyl-cellobiose; XG3, DP3

Wong et al. (2010) *Protein & Peptide Lett.*

XEG 5A & 5B Structures

Gene: 1293 bp
Protein: 431 aa (20 aa signal peptide, 45 kD)
Structure: $(\alpha/\beta)_8$ fold, GH family 5
Catalytic: Glu240, Glu362, retaining

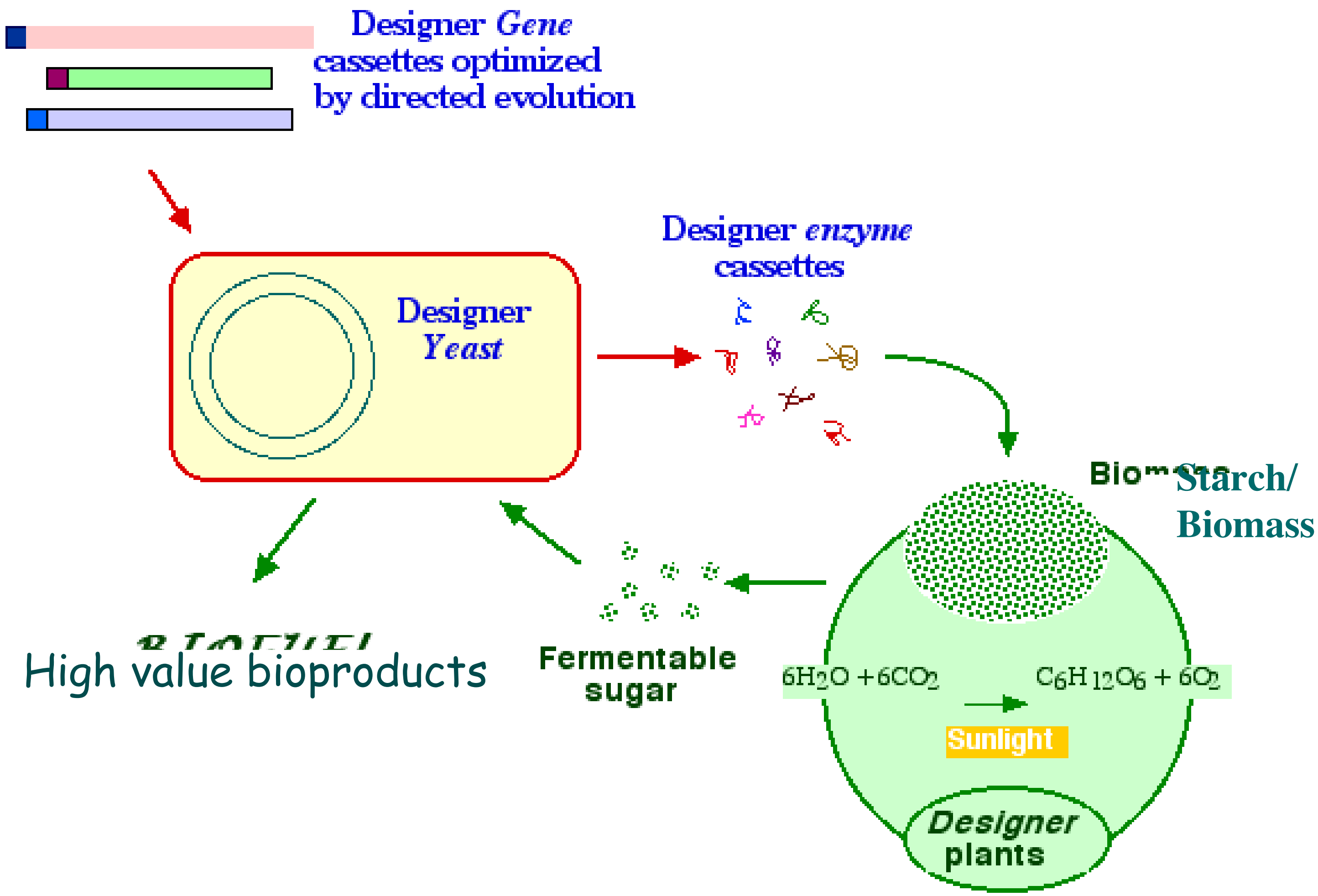
Gene: 1548 bp
Protein: 516 aa (18 aa signal peptide, 45.5 kD)
Structure: $(\alpha/\beta)_8$ fold GH family 5
Catalytic: Glu431, Glu293, retaining



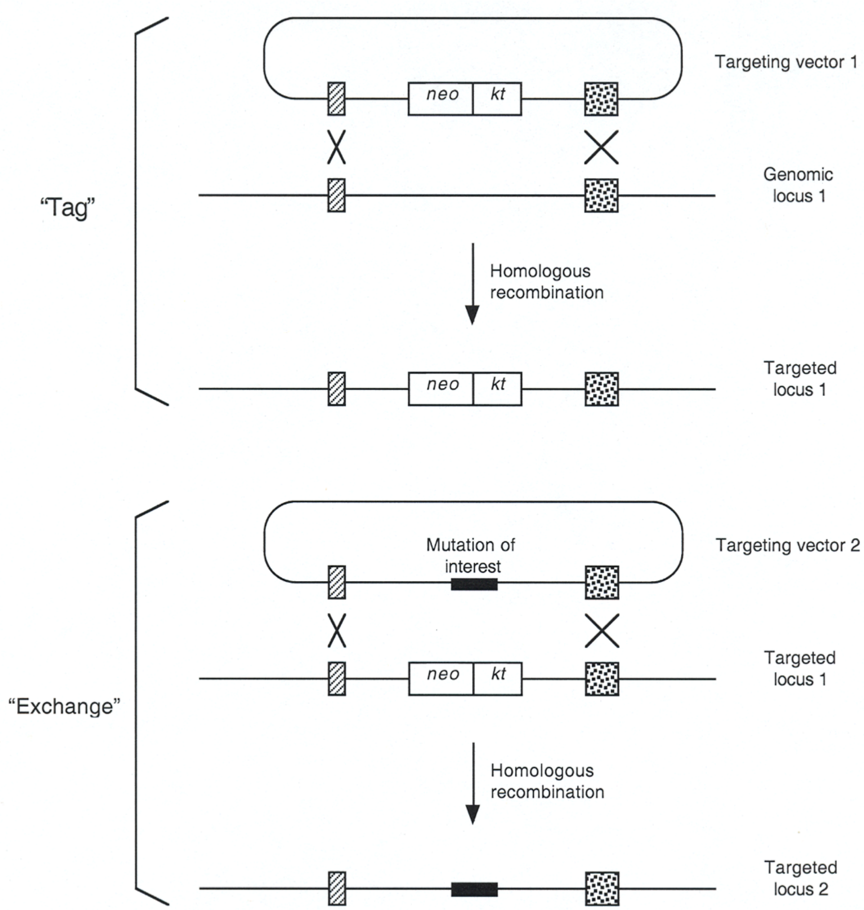
XEG5A - endo

XEG5B - exo

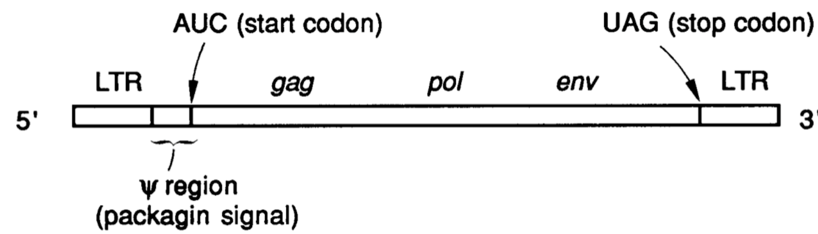
One-pot bioconversion



Homologous recombination



Double-hit replacement



Retrovirus RNA genome

LTR = long terminal repeats
gag = viral core protein
pol = reverse transcriptase
env = envelop
psi region = packaging signal

D. Wong *(2006) The ABCs of Gene Cloning*

Ty elements in *Saccharomyces cerevisiae*

Table 1. Distribution of Ty Elements by Chromosome

Chr. no.	Number of insertions[a] Ty1	Ty2	Ty3	Ty4	Ty5	Total Ty insertions on chromosome
I	6 (1)	1	1	0	0	8
II	13 (2)	2 (1)	1	2	0	18
III	11	2 (1)	0	1	2 (1)	16
IV	20 (5)	3 (3)	6	2	0	31
V	18 (2)	5	4	3	2	32
VI	8	1 (1)	0	1	0	10
VII	25 (3)	2 (2)	6 (1)	4	1	38
VIII	14 (1)	1	3	3 (1)	1	22
IX	5	1	2 (1)	1	0	9
X	14 (2)	4	1	3 (1)	0	22
XI	11	1	1	0	1	14
XII	18 (4)	3 (2)	4	1	0	26
XIII	15 (4)	2	0	3	0	20
XIV	5 (2)	1 (1)	4	3	0	13
XV	18 (2)	3 (2)	4	2	0	27
XVI	16 (4)	2	4	3 (1)	0	25
Genome total	217 (32)	34 (13)	41 (2)	32 (3)	7 (1)	331

[a]Insertions include full-length elements, solo LTRs, and LTR fragments. Numbers of full-length elements are shown in parentheses.

From *Genome Res*. 8:464-478 (1998)

Single clones:

endoglucanase, exoglucanase, b-glucosidase, feruloyl esterase, endoxylaase, α-amylase, glucoamylase,

Double clones (co-integration):

(1) endoxylanase + feruloyl esterase

(2) α-amylase/glucoamylase

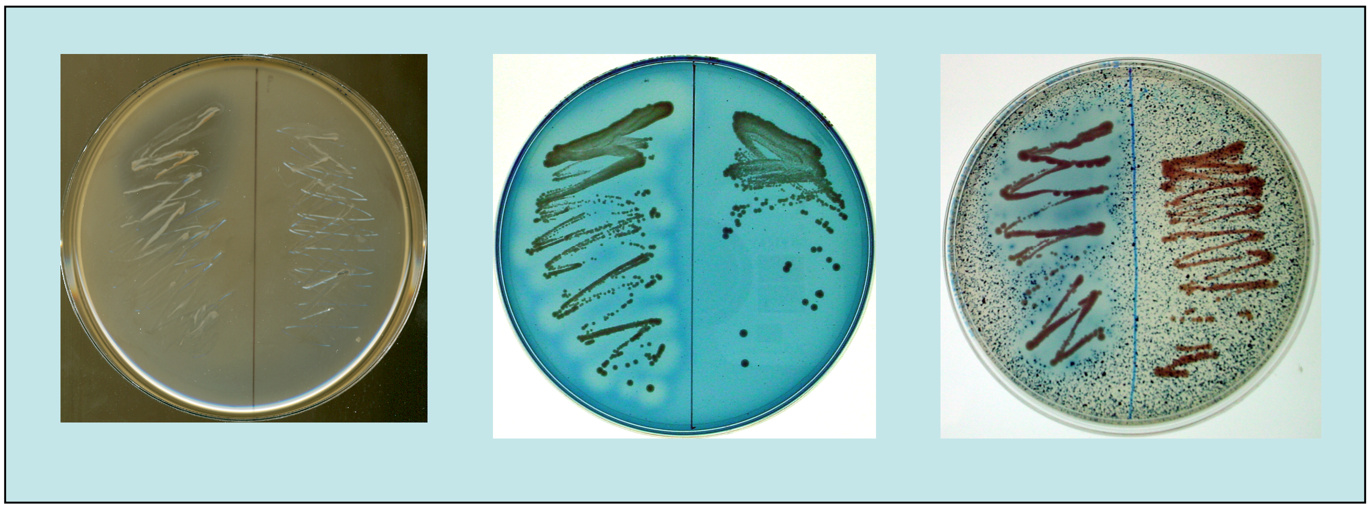
Triple clones:

(1) endoglucanase + exoglucanase + β-glucosidase

(2) endoglucanase + endoxylanase + feruloyl esterase

Integration of genes in yeast

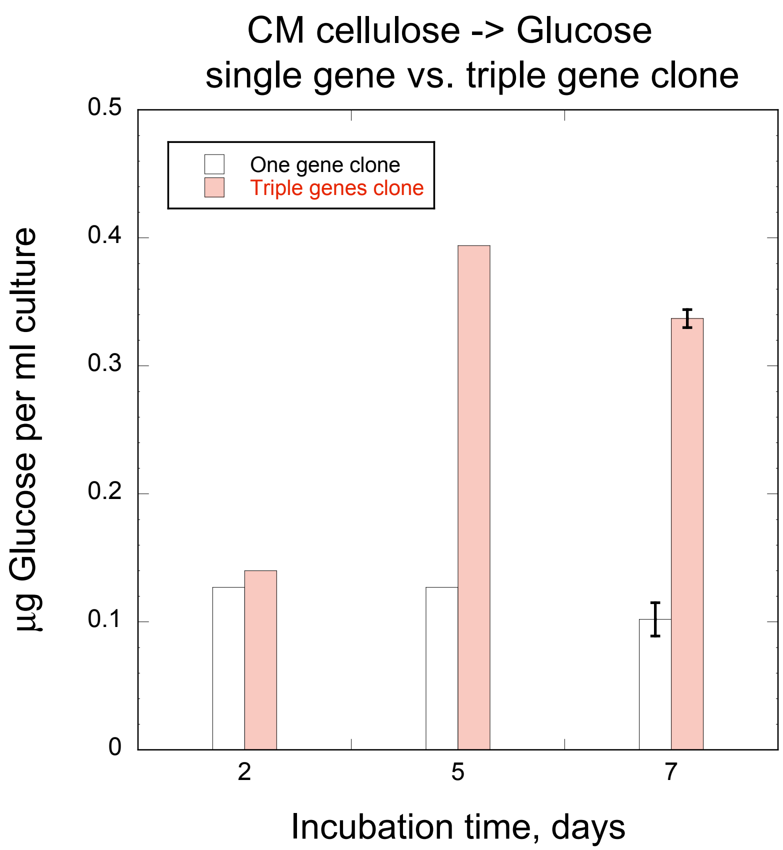
- Integration of multiple genes
- at precise locations in the yeast chromosome
- resulting in stable clones
- rapid testing of enzyme combinations in a cell-based system



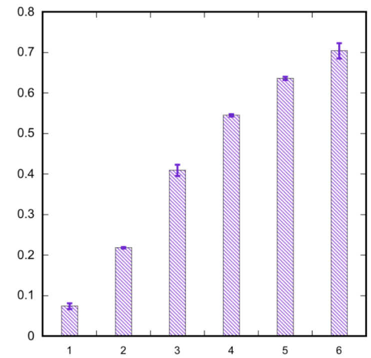
Feruloyl esterase *endo-Xylanase* *endo-Cellulase*

Triple clones -

endoglucanase + exoglucanase + β-glucosidase

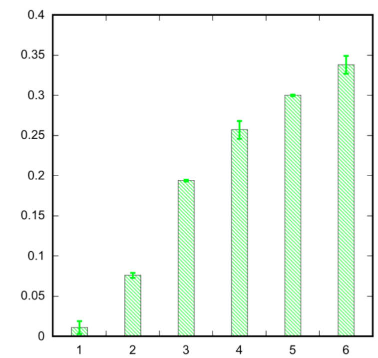


endoglucanase + endoxylanase + feruloyl esterase

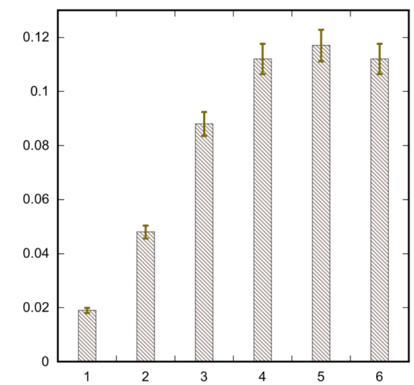


CM celllulose activity

Arabinoxylan activity



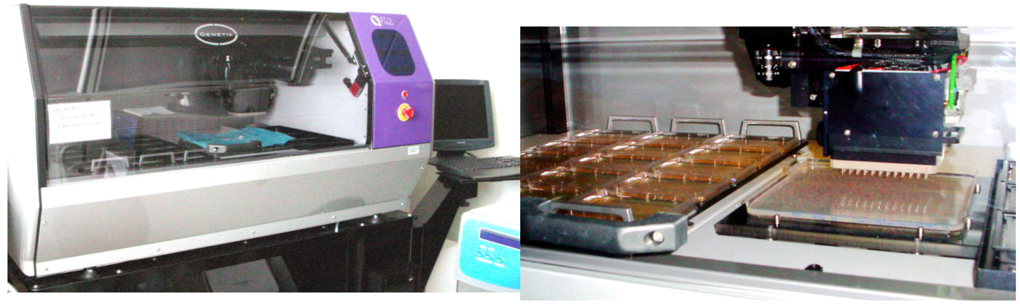
NP ferulate activity



High-Throughput Functional Screening for active clones in Libraries

Metagenomic libraries

- Gene discovery
- Gene variants



Q-Pix

Liquid handling

Capabilities

Incubator-shaker
for microplates

Microplate
reader

- Assay development
- Assay under bioprocess condition

USDA WRRC

WRRC Genes and Enzymes

Enzymes	*Genes*	*Expressed*	*Purified*	*Characterized*
ferulic acid esterase*	12	10	10	10
α-glucuronidase*	(1)	1	1	1
glucuronoyl esterase*	(1)	1	1	1
arabinanase*	9	4	3	2
xyloglucanase*	6	2	1	1
mannanase-glucanase (bifunctional)*	1	1	1	1
rhamnogalacturonanase	1	1	0	0
galactanase	4			
arabinofuranosidase	2			
endoxylanase	2			
Galactomannanase	1			
acetylxylan esterase	2			

Yeast vectors

episomal, integrative

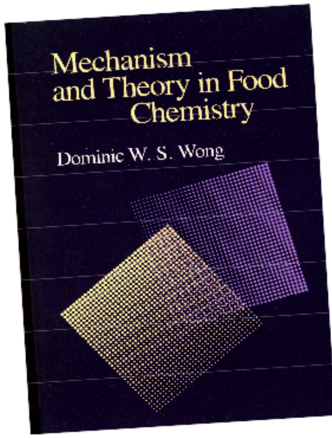
Yeast clones

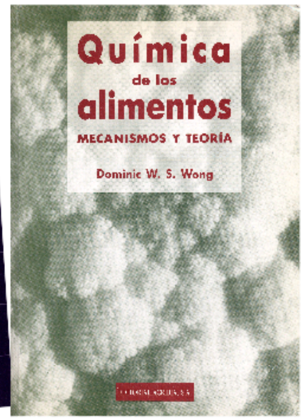
single gene, double genes, triple genes

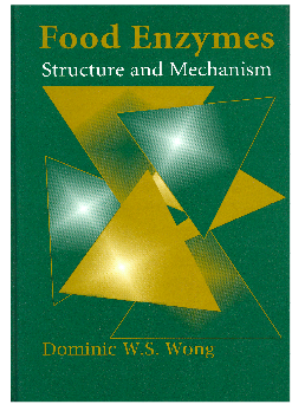
ADDITIONAL ENZYME/PROTEIN RESEARCH AREAS:

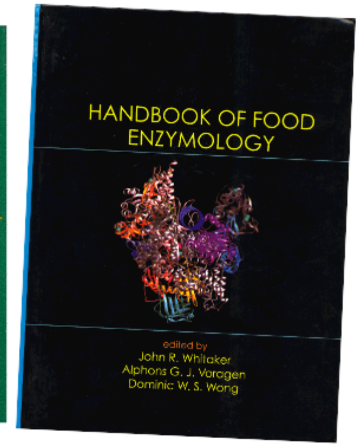
- Directed evolution of enzymes
- Engineering *yeast* for consolidated bioprocessing
- Enzymatic fragmentation of crystalline and complex biopolymers
- Enzymatic crosslinking of food proteins

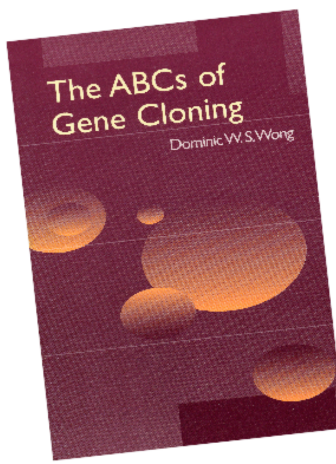
Backgrounds

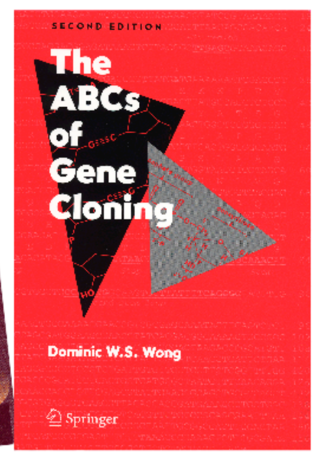




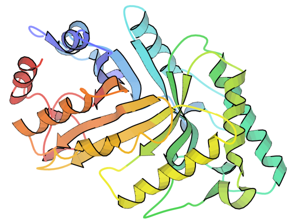




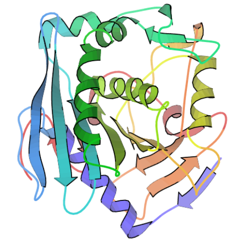




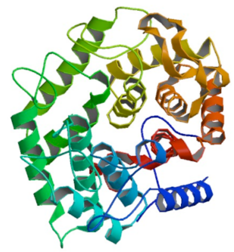
Enzymes.......



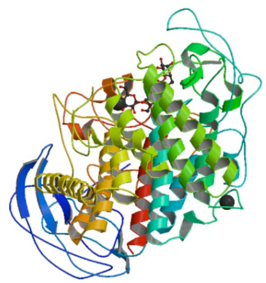
α–amylase



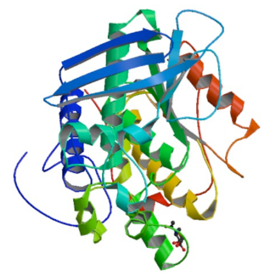
Feruloyl esterase



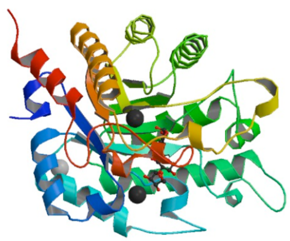
Endoglucanase



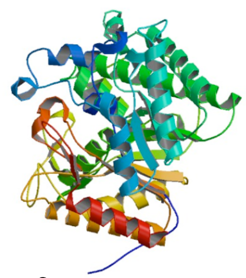
Cellobiohydrolase



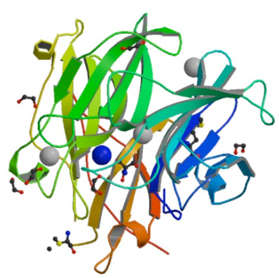
Glucuronoyl esterase



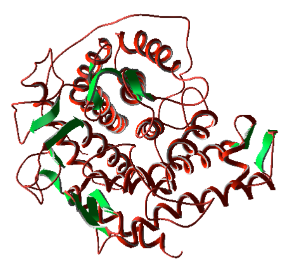
Endoxylanase



β-Glucosidase



β-Xylosidase



Glucoamylase
(catalytic domain)

- Sarah Batt
- Victor chan
- Kurt Wagschal
- Rena Kibblewhite
- Charles Lee
- Amanda McCormack
- Meiling Shang
- David Wan
- Dominic Wong

- Dr. Peter Biely, Slovak Academy of Science
- Dr. Mario Murakami, LNBio, Brazil
- Dr. Mary Jo Zidwick, Cargill Biotech Center
- Professor Roy Doi, UC Davis
- Dr. Hans Liao, OPX Biotechnologies

USDA WRRC

Thank You.......