The potential of fungal genetics and transformation for production of tools for cellulosic ethanol production

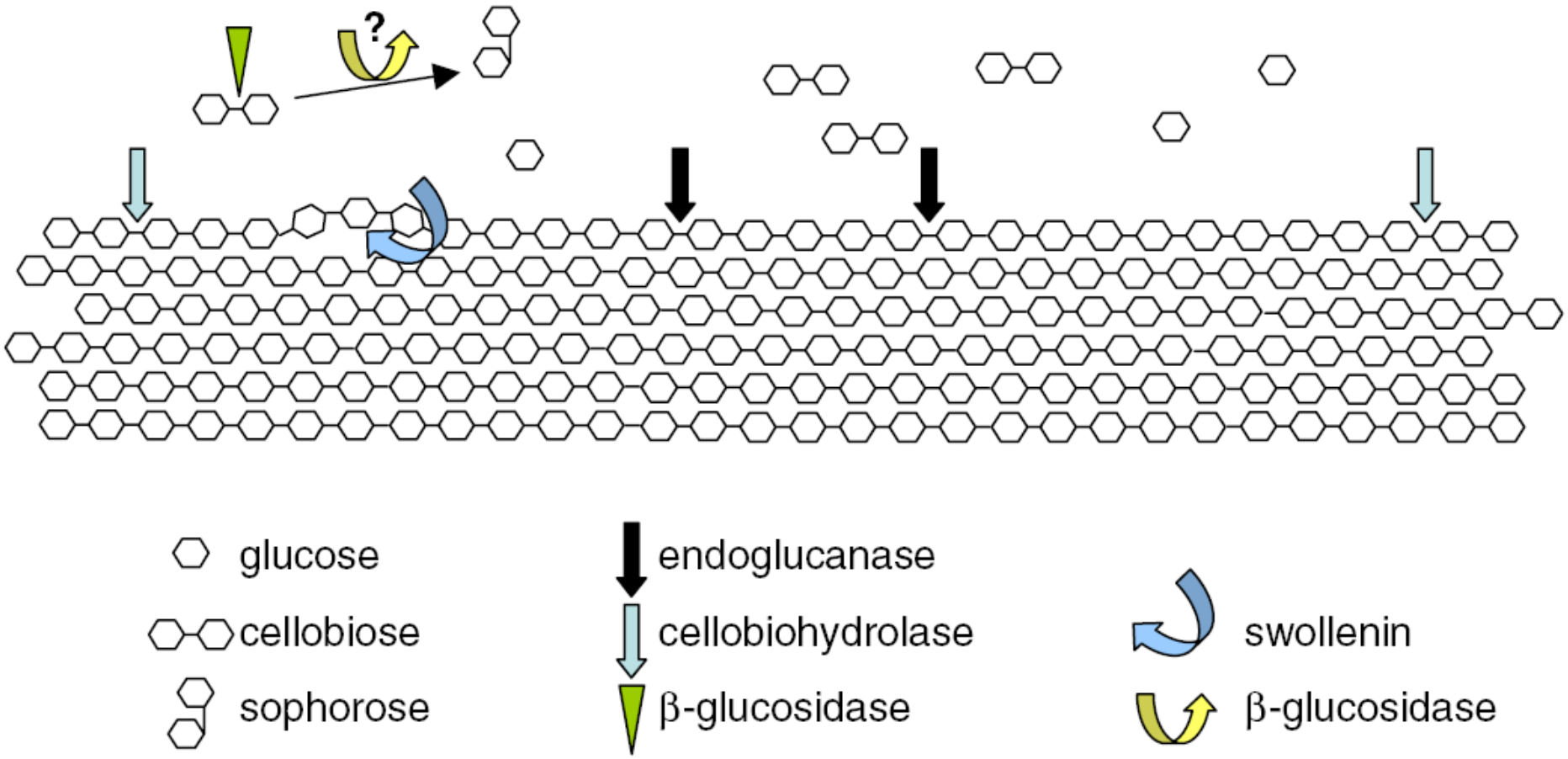
Gustavo H. Goldman
FCFRP-Universidade de São Paulo

- It is estimated that capital costs associated with lignocellulosic ethanol are about US$ 4 per gallon and that these need to be reduced by more than half to be economically sustainable

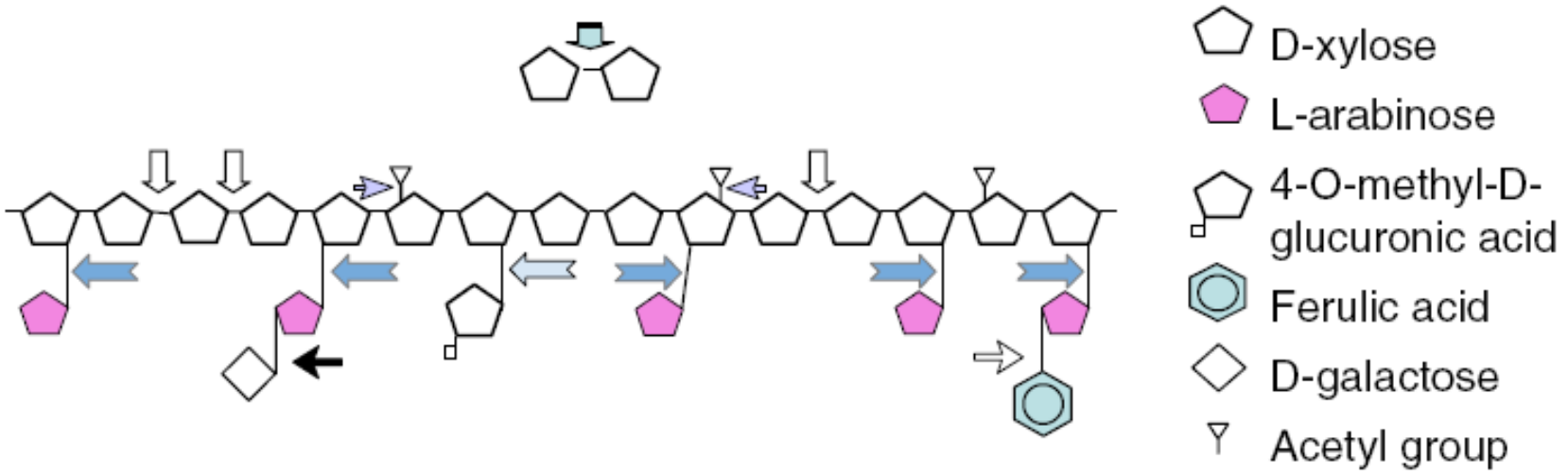
- Complete substrate utilization is one of the prerequisites to render lignocellulosic ethanol processes economically competitive

- This means that **ALL** types of sugars in cellulose and hemicellulose must be converted to ethanol, and that microorganisms must be obtained that efficiently perform this conversion under industrial conditions

Degradation of cellulose



Degradation of hemicellulose



SYMBOL	ENZYME	EC-NUMBER	LINKAGE HYDROLYSED
⇩	Endo-1,4-β-xylanase	3.2.1.8	β-1,4
⬓	β-D-xylosidase	3.2.1.37	β-1,4
➡	α-L-arabinofuranosidase Arabinoxylan arabino-furanohydrolases	3.2.1.55	α-L-1,2; α-L-1,3 and α-L-1,5
⇨	α-D-glucuronidase	3.2.1	α-1,2
⇨	feruloyl esterase	3.2.1	ester bond
➡	α-D-galactosidase	3.2.1.22	α-1,6
⇨	asetylxylan esterase	3.1.1.6	ester bond

- Biomass is composed of cellulose (40-50 %), hemicellulose (25-35 %) and lignin (15-20 %)

- Glucose constitutes about 60 % of the total sugars available in cellulosic biomaterial

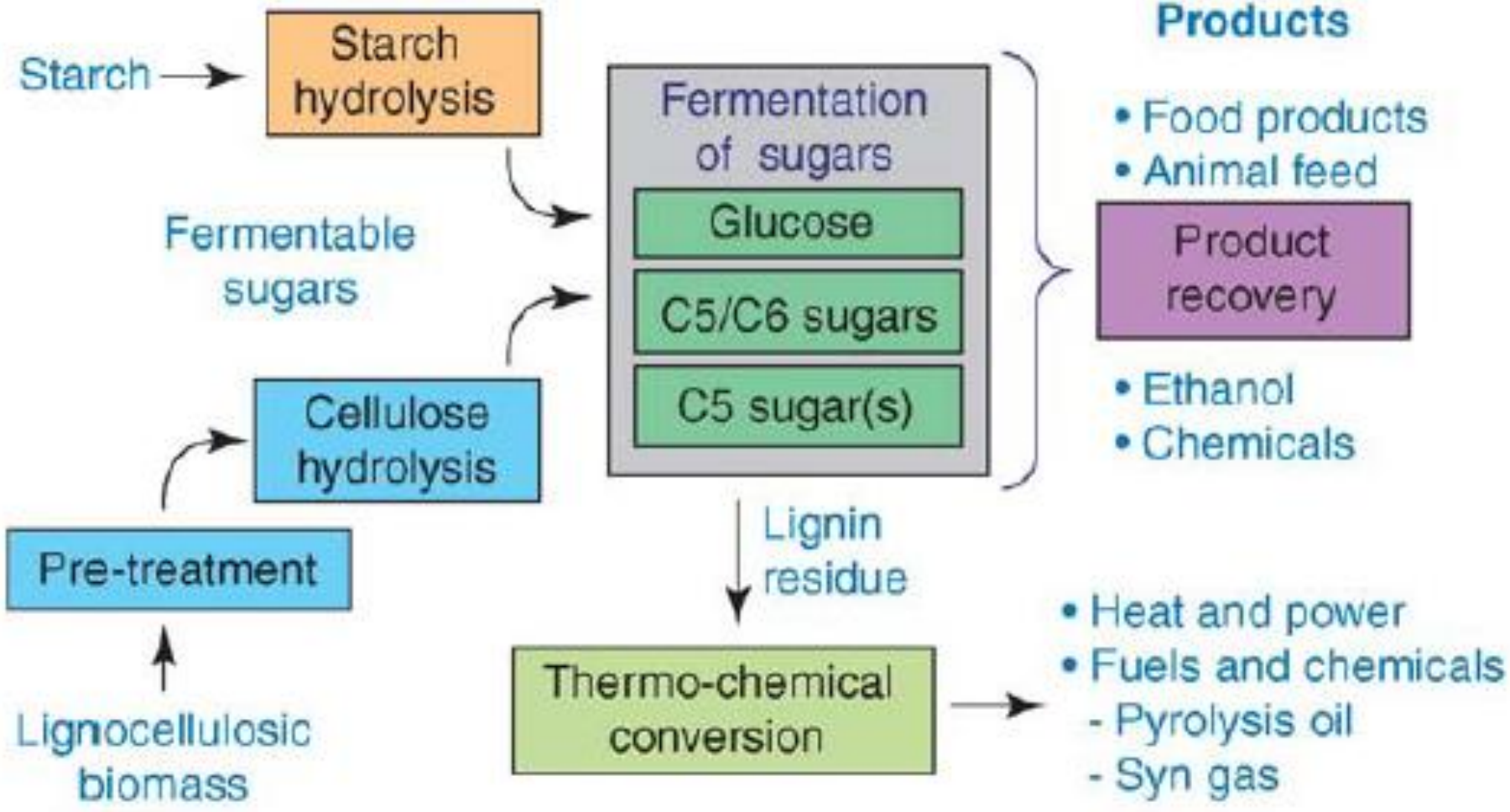
- Fermentation of the available sugars in cellulosic biomass presents a unique challenge because of the presence of other sugars such as xylose and arabinose (C5 sugars)

- The degree of branching and identity of the minor sugars in hemicellulose tends to vary depending upon the type of plant

- The conversion of biomass to useable energy is not economical, however, unless hemicellulose is used in addition to cellulose

- Lignin can be covalently linked to hemicellulose via ferulic acid ester linkages

Schematic of biomass and starch processing that could occur in a biorefinery



- Pretreatment chemistries vary from very acidic to quite alkaline

- The acidic treatments (usually sulfuric acids) will hydrolyze the hemicellulose fraction while leaving the cellulose and lignin intact in the residual solids

- The alkaline approaches tend to have more of an effect on the lignin component and leave both cellulose and hemicellulose intact

- Acidic treatments may result in high concentrations of furfurals in the liquid phase, whereas the alkaline methods may result in ferulate and acetate in the hydrolysate.

Pros and cons of various natural microorganisms with regard to industrial ethanol production

Organism	Natural sugar utilization pathways					Major products		Tolerance			O_2 needed	pH range
	Glu	Man	Gal	Xyl	Ara	EtOH	Others	Alcohols	Acids	Hydrolysate		
Anaerobic bacteria	+	+	+	+	+	+	+	−	−	−	−	Neutral
E. coli	+	+	+	+	+	−	+	−	−	−	−	Neutral
Z. mobilis	+	−	−	−	−	+	−	+	−	−	−	Neutral
S. cerevisiae	+	+	+	−	−	+	−	++	++	++	−	Acidic
P. stipitis	+	+	+	+	+	+	−	−	−	−	+	Acidic
Filamentous fungi	+	+	+	+	+	+	−	++	++	++	−	Acidic

Outline

- Engineering *S. cerevisiae* for xylose fermentation

- *Trichoderma reesei* and *Aspergillus niger* as model cellulolytic microorganisms

- Regulation of cellulases/hemicellulases expression at transcriptional level

- Ideal microorganism envisioned for use in biomass conversion should:

(i) produces ethanol with a high yield and specific productivity;
(ii) has a broad substrate range and high ethanol tolerance;
(iii) and be tolerant to the inhibitors present in lignocellulosic hydrolysates

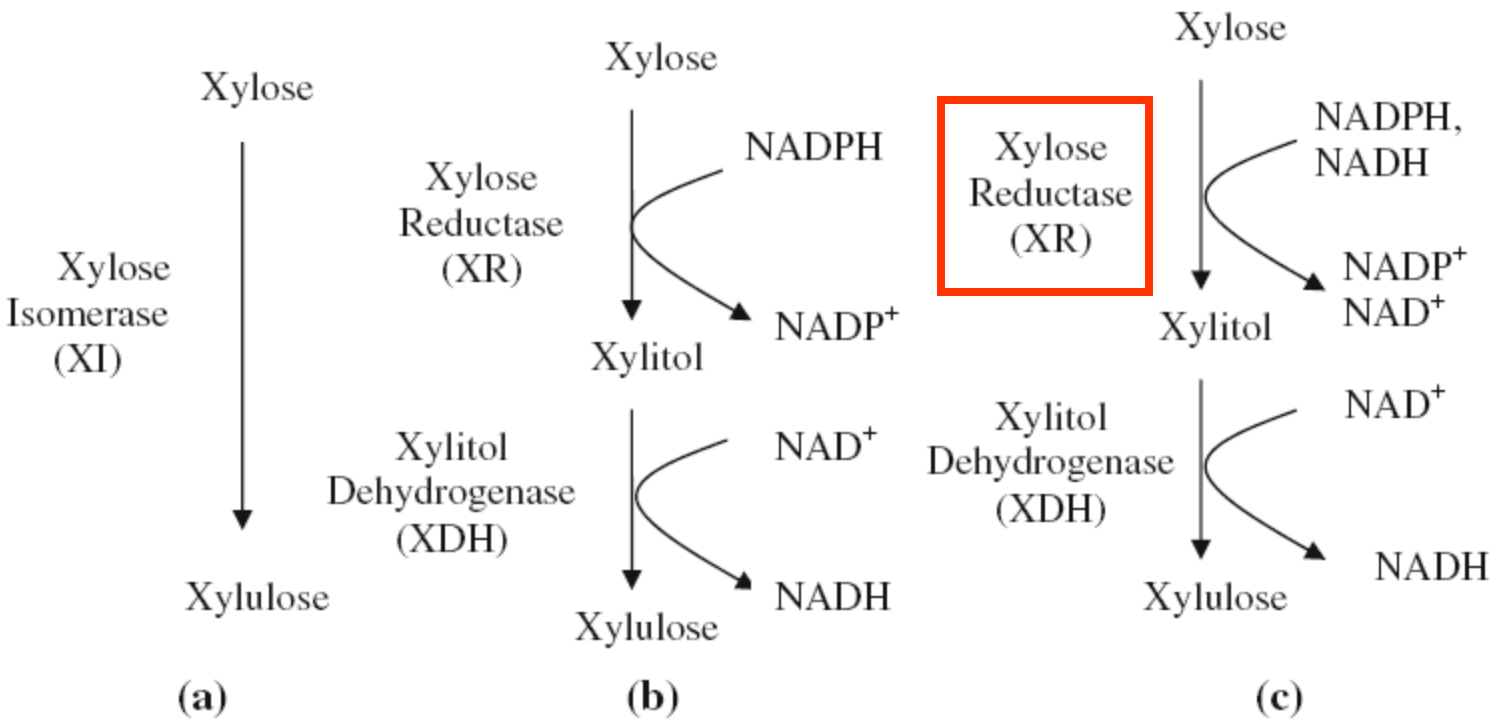
- *S. cerevisiae* meets many of the criteria above **BUT** it is unable to effectively use xylose as a sole carbon source

- Organisms to ferment the C5 sugars in cellulosic biomass can be divided into two subgroups, namely naturally-occurring and genetically engineered microorganisms.

- The naturally-occurring microorganisms include *Pichia stipitis*, *Candida shehatae*, and *Pachystolen tannophilus*

- Genetically-engineered organisms with C5 fermenting capabilities include *Saccharomyces cerevisiae, E. coli*, and *Zymomonas mobilis.*

Xylose fermentation pathway



- The difference in cofactor preference of XR and XDH leads to the formation of xylitol under anaerobic conditions. Xylitol is therefore a byproduct in ethanol fermentation and its production reduces the final ethanol yield.

- *Pichia stipitis* is a haploid, homothalic, hemiascomycetous yeast that has the **highest native capacity for xylose fermentation of any known microbe**

- Fed batch cultures of *Pichia stipitis* produce almost **50 g/l of ethanol from xylose with yields of 0.35 to 0.44 g/g xylose** and they can ferment hydrolysates at 80 % of the maximum theoretical capacity

- *S. cerevisiae* regulates fermentation by sensing the presence of glucose, whereas *P. stipitis* induces fermentative activity in response to oxygen limitation.

- In yeasts, such as *S. cerevisiae*, ethanol is produced when sugar concentrations are relatively low, even under aerobic conditions. (Crabtree effect)

- Unlike *S. cerevisiae*, *P. stipitis* is a respiratory yeast, which does not produce ethanol under aerobic conditions, even in the presence of excess sugars

- The requirement for electron acceptors translates to very low, carefully controlled, levels of oxygen required for maximum ethanol production from arabinose and xylose by these yeasts

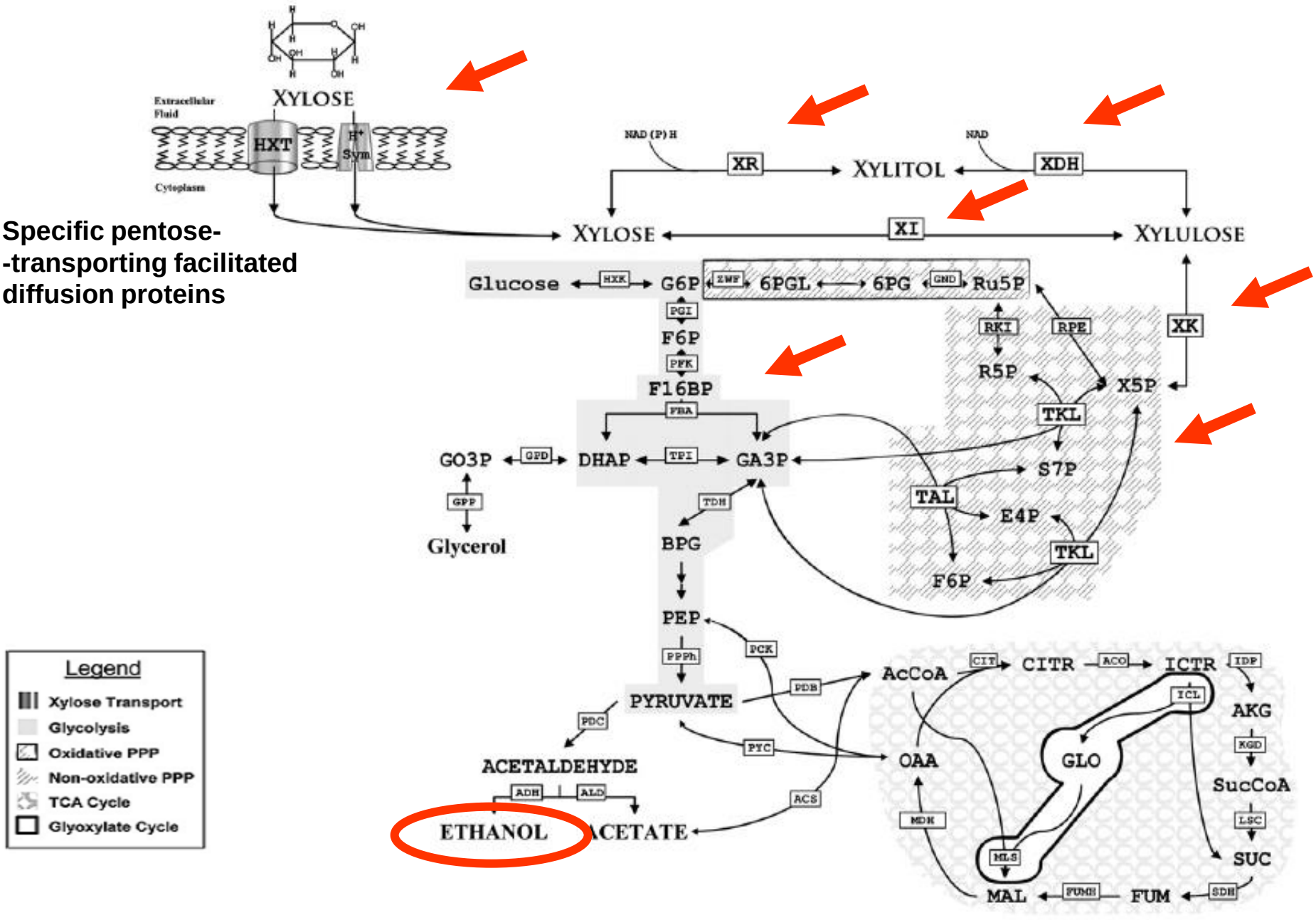
- While *S. cerevisiae* naturally harbors genes for xylose utilization, these are expressed at such low levels that they do not support growth on xylose

- However, *S. cerevisiae* grows on D-xylulose and ferments it to ethanol, albeit with ten times lower rate than for glucose

- The inability of *S. cerevisiae* to grown on xylose has been attributed to **inefficient xylose uptake, a redox imbalance generated in the first two steps of xylose metabolism, insufficient Xylulose Kinase (XK) activity, and an inefficient PPP**.

Summary of the genetically engineered *S. cerevisiae* xylose-fermenting strain

Specific pentose--transporting facilitated diffusion proteins



Characteristics of industrial xylose-fermenting *S. cerevisiae* strains

Strain	Genotype	Strain Background
TMB 3400	XR/XDH/XK integrated and random mutagenesis	USM21 (van der Westhuizen and Pretorius, 1992)
TMB 3006	**XR/XDH/XK integrated in**	**TMB 3000 (Linden et al., 1992)**
Adapted TMB 3006	Improved hydrolysate tolerance by evolutionary engineering	TMB 3000 (Linden et al., 1992)
F 12	XR/XDH/XK integrated	*S. cerevisiae* F
BH42	Bred strain	-
A4	XR/XDH/XK integrated in *S. cerevisiae* A^+	*S. cerevisiae* A^-
424A LNH-ST	XR/XDH/XK multiply integrated	*S. cerevisiae* ATCC 4124
I400 (pLNH32)	XR/XDH/XK on plasmid	*S. diastaticus/* *S. uvarum*

Hahn-Hägerdal *et al.* Appl Microbiol Biotechnol (2007) 74:937–953

Fermentation performance of industrial xylose-fermenting *S. cerevisiae* strains in lignocellulose hydrolysates

Strain	Hydrolysate	Detoxification	Setup	Ethanol productivity	Ethanol yield on total sugar	Reference
TMB 3400	Spruce	Nondetoxified	Fed-batch	0.25	0.43	(Hahn-Hägerdal and Pamment 2004)
TMB 3006	Spruce	Nondetoxified	Fed-batch	0.66	0.37	(Hahn-Hägerdal and Pamment 2004)
424A LNH-ST	Corn stover	Overliming	Batch	–	0.41	(Sedlak and Ho 2004)
424A LNH-ST	Corn stover	Not known	Batch	–	0.44	(Sedlak and Ho 2004)
Adapted TMB 3006	Northern spruce 70%	Nondetoxified	continuous, D 0.1	–	0.41 (on glucose)	(Hahn-Hägerdal et al. 2005)
MT8-1/Xyl/ BGL	wood chip hydrolysate	Overliming	Batch	0.42	0.41	(Katahira et al. 2006)
F 12	Still bottoms fermentation residue (vinasse)	Not known	Batch	0.005–0.24	0.27[a]	(Olsson et al. 2006)
TMB 3400	Corn stover, steam pretreatment	Nondetoxified	Batch SSF	–	0.32	(Öhgren et al. 2006)
TMB 3400	Corn stover, steam pretreatment	Nondetoxified	Fed-batch SSF	–	0.30	(Öhgren et al. 2006)

- Fed batch cultures of *Pichia stipitis*: 50 g/l of ethanol from xylose with yields of 0.35 to 0.44 g/g xylose capacity

Hahn-Hägerdal *et al.* Appl Microbiol Biotechnol (2007) 74:937–953

-Many microorganisms have the capacity to degrade plant biomass. Filamentous fungi like ***Hypocrea jecorina*** **(*Trichoderma reesei*)** and ***Aspergillus niger*** have been shown to produce a wide spectrum of polysaccharide-hydrolytic enzymes

- Of these fungi, *A. niger* is industrially used as a producer of many **pectinases and hemicellulose degrading enzymes, like xylanases and arabinases**

General features of fungal genomes compared with that of *T. reesei* and *A. niger*

Organism	Size (Mb)	No. genes	% coding	% GC
T. reesei	**33.9**	**9,129**	**40.40**	**52.0**
A. niger	**33.9**	**14,165**	**~50.00**	**50.4**
F. graminearum	36.1	11,640	56.24	48.3
N. crassa	38.7	10,620	38.50	49.6
M. grisea	39.4	12,841	50.40	52.0
A. nidulans	30.1	10,701	58.80	50.3
S. cerevisiae	12	5,885	72.55	38.3
P. chrysosporium	34.5	10,048	42.22	56.8

T. reesei

- Strikingly, fewer glycoside hydrolases, polysaccharide lyases, cellulases, hemicellulases and pectinases are present than in the other fungal genomes examined

- A considerable proportion (41 %) of cellulases and hemicellulases are located in clusters, and approximately two-thirds and some of these genes show evidence for co-expression

- Analyses of the protein secretion revealed a higher level of diversity among the membrane-trafficking components compared with the budding yeast *S. cerevisiae*

Sizes of CAZyme families, by class, in the 13 fungal genomes analyzed

Lineages		Species	GH	Avg. GH	GT	Avg. GT	CBM	Avg. CBM	CE	Avg. CE	PL	Avg. PL
Ascomycetes	Eurotio.	*A.nid.*	247	265	91	103	36	40	29	28	19	18
		A.fum.	263		103		55		29		13	
		A.ory.	285		114		30		26		21	
	Sordario.	*M.gris.*	231	211	94	96	58	49	47	32	4	8
		N.cra.	171		76		39		21		3	
		T.ree.	**200**		**103**		**36**		**16**		**3**	
		F.gra.	243		110		61		42		20	
	Saccharo.	*C.alb.*	58	47	69	70	4	9	3	3	0	0
		S.cer.	45		67		12		3		0	
		C.gla.	38		73		12		3		0	
	Archiasco.	*S.pom.*	46	46	61	61	5	5	5	5	0	0
Basidio.		*C.neo.*	75		68		10		9		3	

Enzymes: GH, glycoside hydrolase; GT, glycosyltransferase; CBM, carbohydrate-binding module; CE, carbohydrate esterase; PL, polysaccharide lyase.

Cellulolytic enzymes encoded in the *T. reesei* genome

Cellulase type[a]	CBH1 (Cel7A)	CBH2 (Cel6)	EG1 (Cel7B)	EG2 (Cel5)	EG3 (Cel12)	EG4 (Cel61)	EG5 (Cel45)	
Species[b]								Sum
A.nid.	2	2	1	2	1	9	1	18
A.fum.	2	1	2	3	3	7	1	19
A.ory.	2	1	1	2	2	8	0	16
M.gris.	3	2	2	2	3	17	1	30
N.cra.	2	2	3	1	0	14	1	23
T.ree.	**1**	**1**	**1**	**2**	**1**	**3**	**1**	**10**
F.gra.	1	0	1	2	2	13	1	20
C.alb.	0	0	0	0	0	0	0	0
S.cer.	0	0	0	0	0	0	0	0
C.gla.	0	0	0	0	0	0	0	0
S.pom.	0	0	0	0	0	0	0	0
C.neo.	0	0	0	0	0	1	0	1
P.chr.	7	1	2	2	1	14	0	27

Enzymes: CBH1, exocellobiohydrolase I, GH7; CBH2, exocellobiohydrolase II, GH6; EG1, endoglucanase I, GH7; EG2, endoglucanase II, GH5_5; EG3, endoglucanase III, GH12_1; EG4, glycoside hydrolase family, Cel61, GH61; EG5, endoglucanase V, Cel45.

Hemicellulose-degrading enzymes encoded in *T. reesei* genome, arranged by GH (Glycoside Hydrolase) family

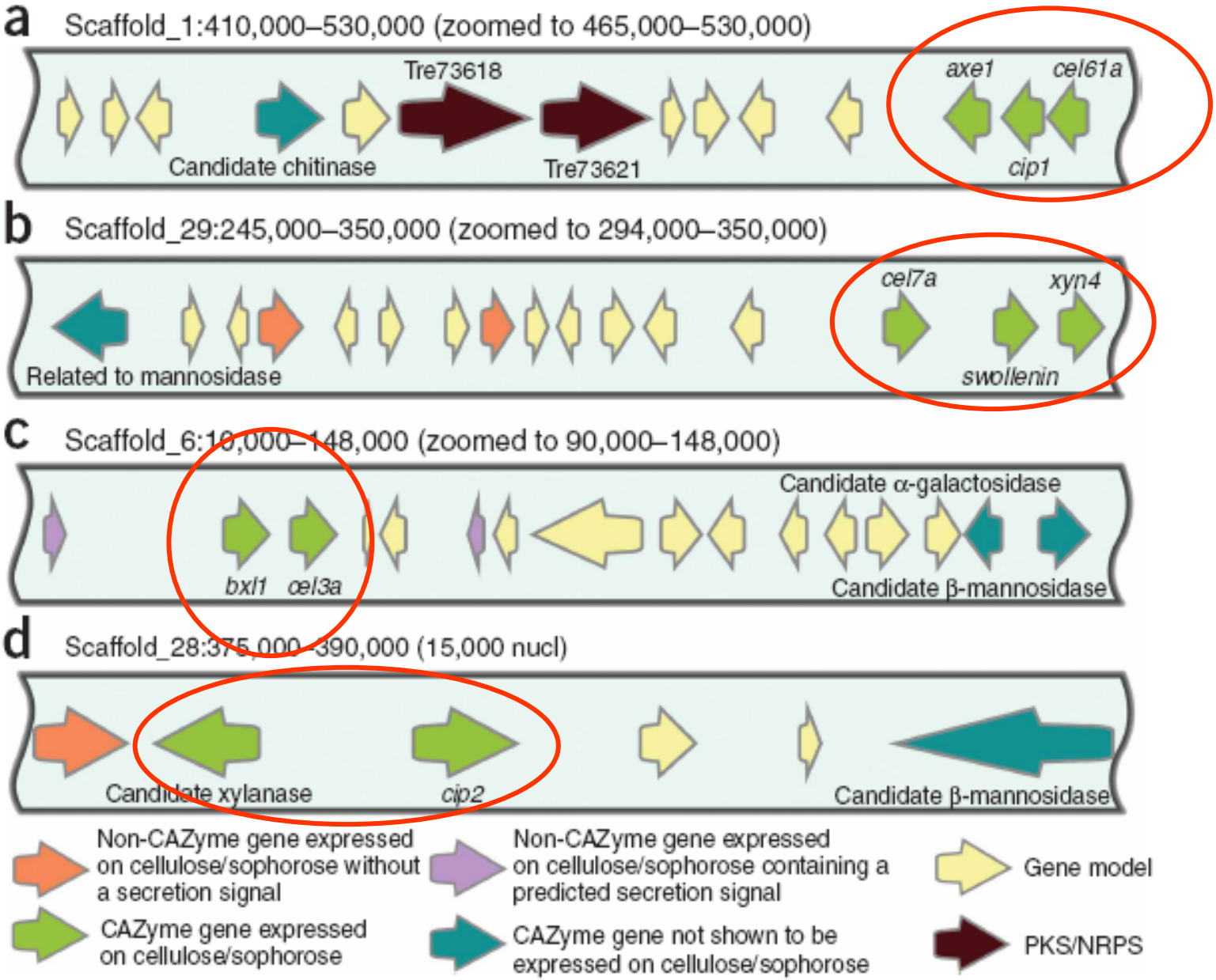
Family[a] Species[b]	GH43	GH10	GH11	GH51	GH74	GH62	GH53	GH54	GH67	GH29	GH26	GH95	Total
A.nid.	15	3	2	2	2	2	1	1	1	0	3	3	35
A.fum.	18	4	3	2	2	2	1	1	1	0	0	2	36
A.ory.	20	4	4	3	0	2	1	1	1	0	1	3	40
M.gris.	19	5	5	3	1	3	1	1	1	4	0	1	44
N.cra.	7	4	2	1	1	0	1	1	1	0	1	0	19
T.ree.	**2**	**1**	**4**	**0**	**1**	**1**	**0**	**2**	**1**	**0**	**0**	**4**	**16**
F.gra.	16	5	3	2	1	1	1	1	1	1	0	2	34
C.alb.	0	0	0	0	0	0	0	0	0	0	0	0	0
S.cer.	0	0	0	0	0	0	0	0	0	0	0	0	0
C.gla.	0	0	0	0	0	0	0	0	0	0	0	0	0
S.pom.	0	0	0	0	0	0	0	0	0	0	0	0	0
C.neo.	0	0	0	1	0	0	0	0	0	0	0	0	1
P.chr.	4	6	1	2	4	0	1	0	0	0	0	1	19

- 130 of the 316 (41 %) CAZyme genes are found in 25 discrete regions ranging from 14 kb to 275 kb in length (roughly 2.4 Mb, or 7 % of the genome)

-These regions contain a density of CAZenzyme genes averaging fivefold greater than the expected density for randomly distributed genes

- Gene movement is the major factor in the organization of the clusters, whereas gene duplications have a minor role

Regions of increased CAZyme density

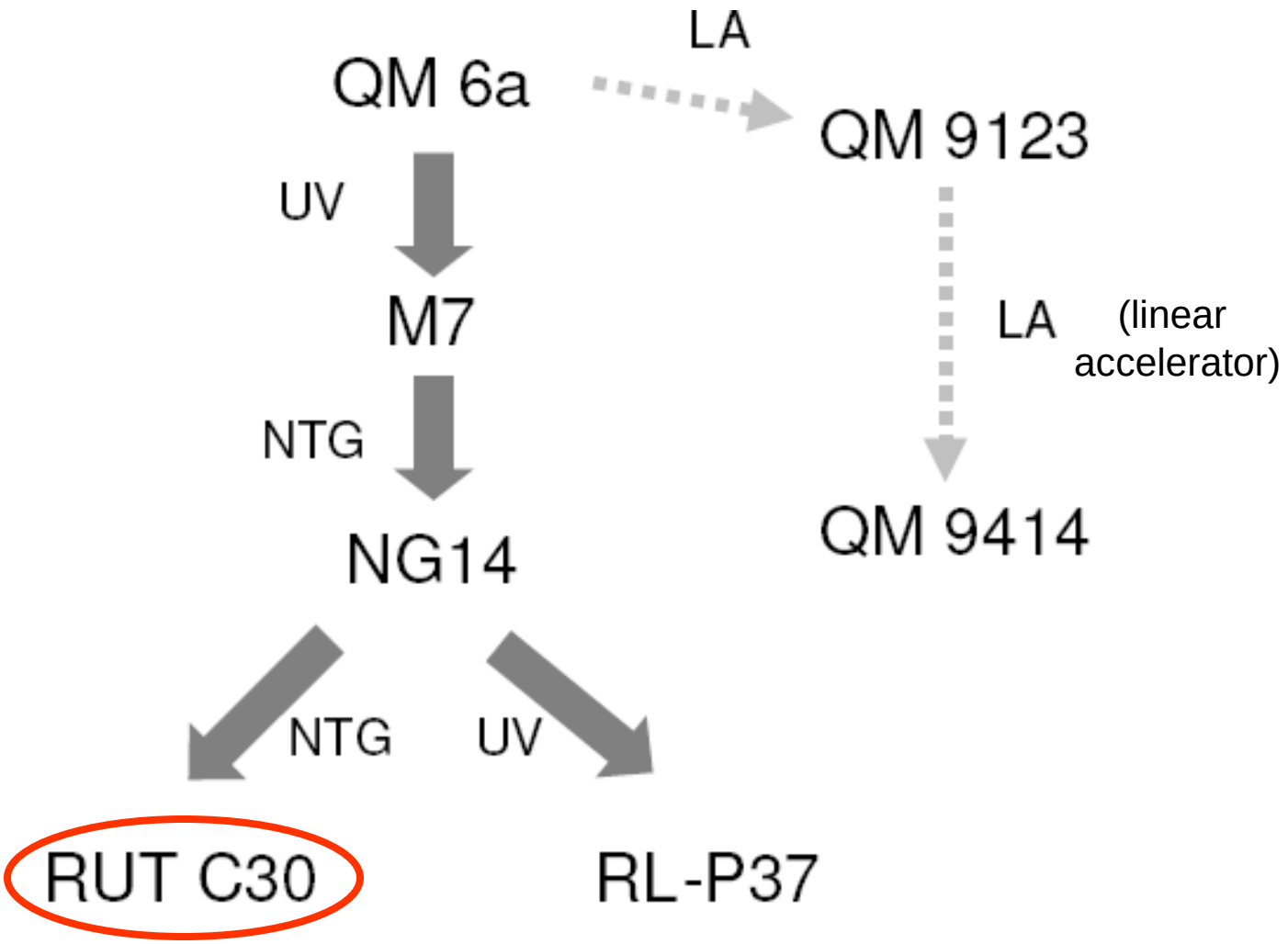


- Although not all of the clustered GH genes were coexpressed, four examples were found in which adjacent or nearly adjacent genes were coexpressed.

- Notably, in these regions there is no syntenic signal with any of the other fungal genomes, suggesting that these genes are reordered in *T. reesei* and that this organization is evolutionarily advantageous for the fungus.

- It is reasonable to conclude that the clustering of the CAZyme genes is favored by selection for the enhanced degradative efficiency and coordinated regulation that a colocalization strategy may offer.

Pedigree of strain *T. reesei* RUT C30 and its relationship to the wild-type isolate QM6a*



*All of the strains that are currently used on a commercial scale have been ultimately derived from one single isolate which was collected on the Solomon islands during World War II

- **RUT C30**:

(i) a truncation in the *cre1* gene, which renders this strain carbon catabolite derepressed

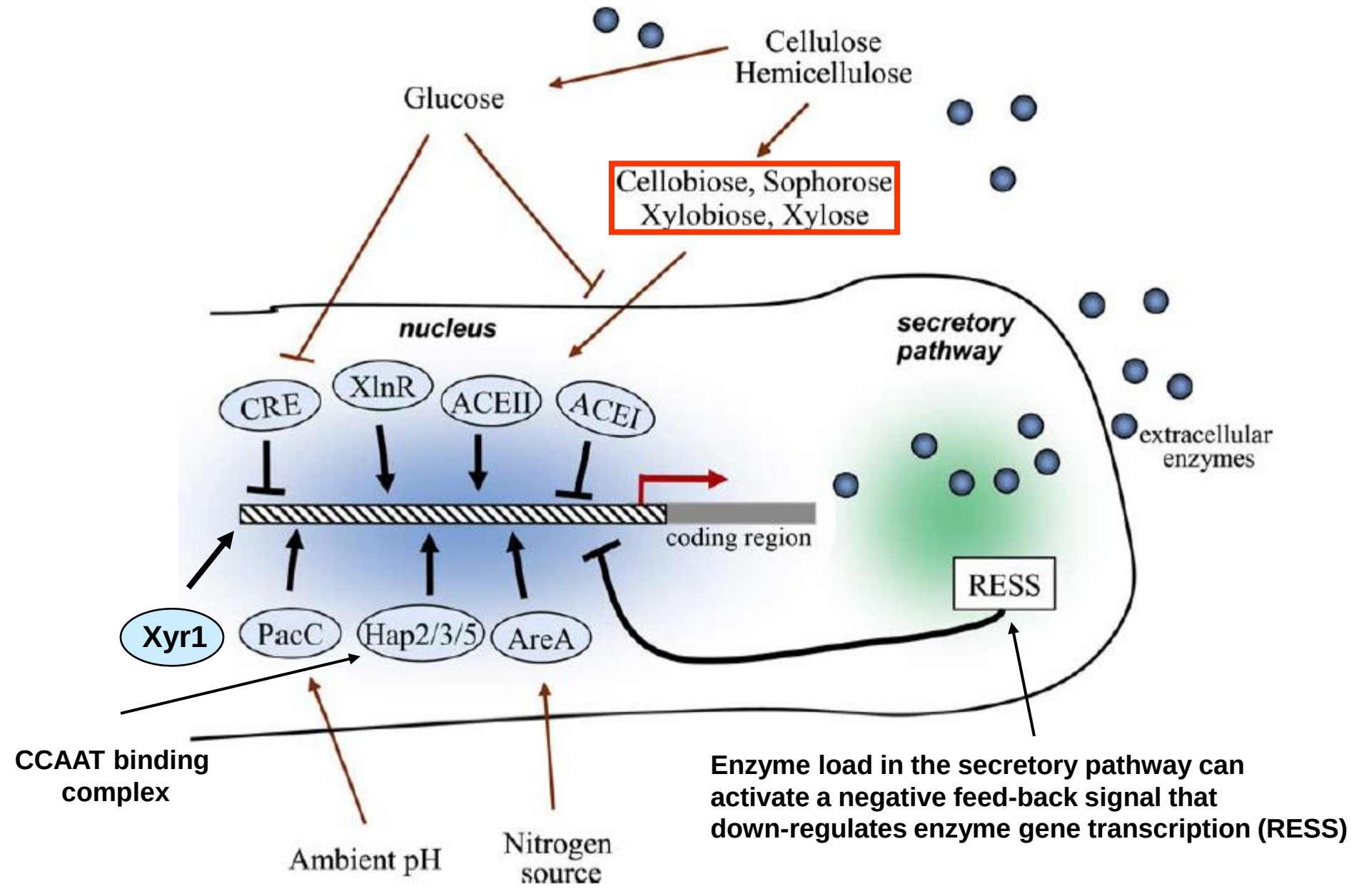
(ii) Frameshift mutation in the glycoprotein processing β-glucosidase II encoding gene;

(iii) Two largest chromosomes are somewhat smaller, whereas the other five chromosomes are somewhat larger, resulting in a total increase in genome size from 32.5 to 34.7 Mbps.

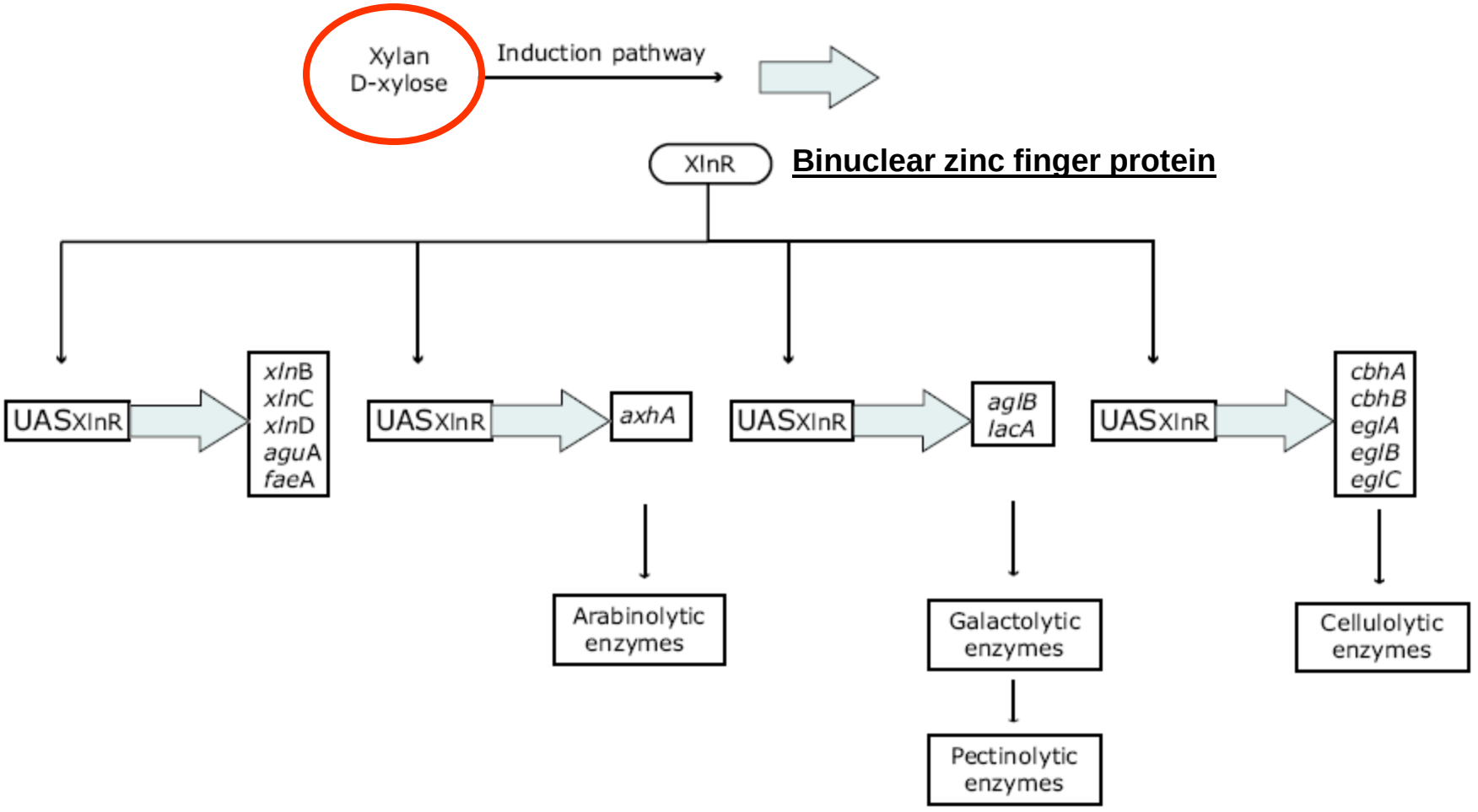
(iv) Lacks a large (85 kb) segment of genes present on scaffold 15 of the genomic sequence of the wild type strain *H. jecorina* QM6a.

- How is the fungus able to sense the insoluble polymeric substrate? What are the signalling molecules? How is the system regulated ?

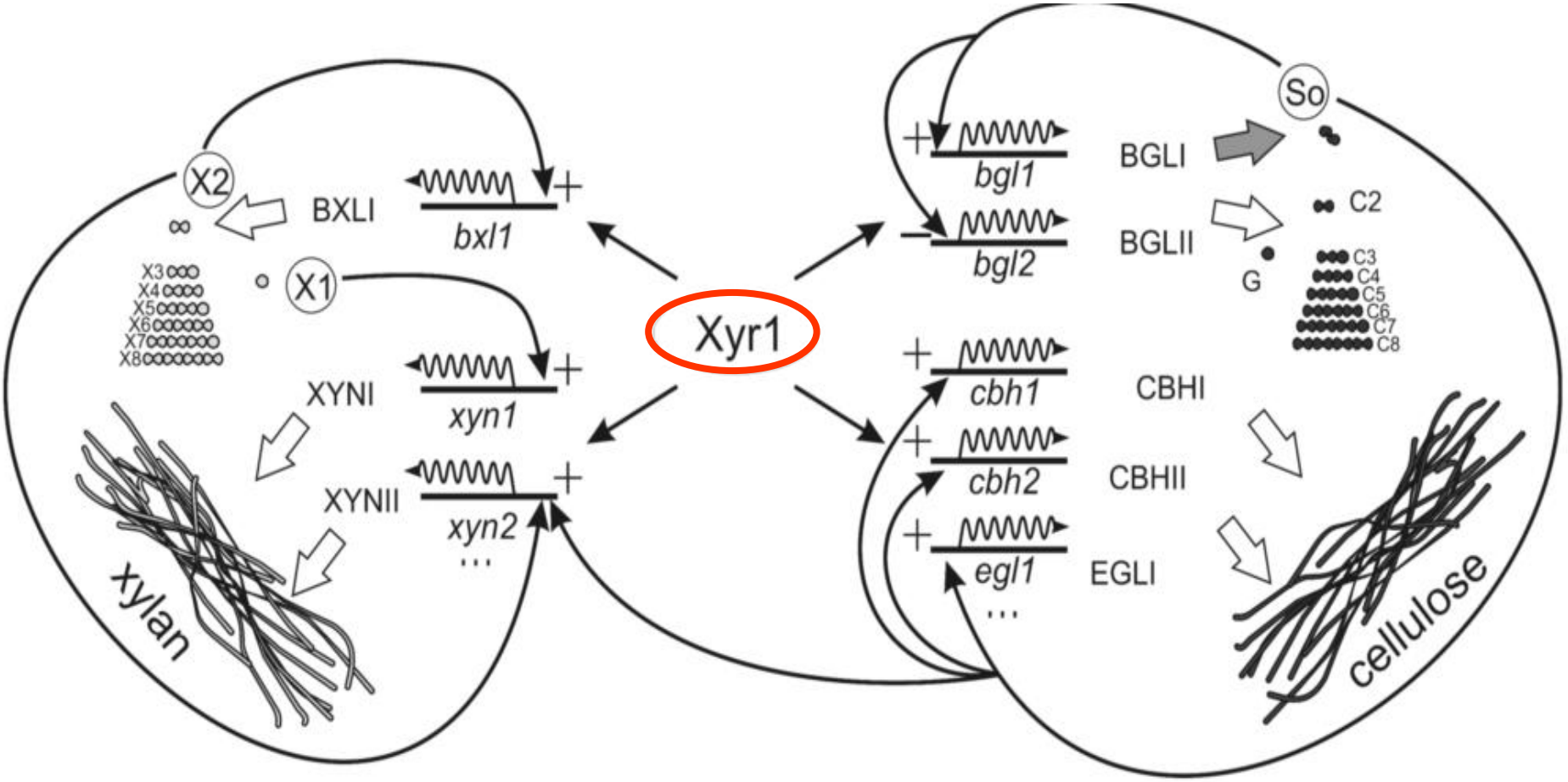
Representation of the different fungal trans-acting factors and regulatory responses affecting cellulase and xylanase expression



Model for the regulation of genes encoding enzymes involved in the degradation of plant cell wall polysaccharides in *A. niger*

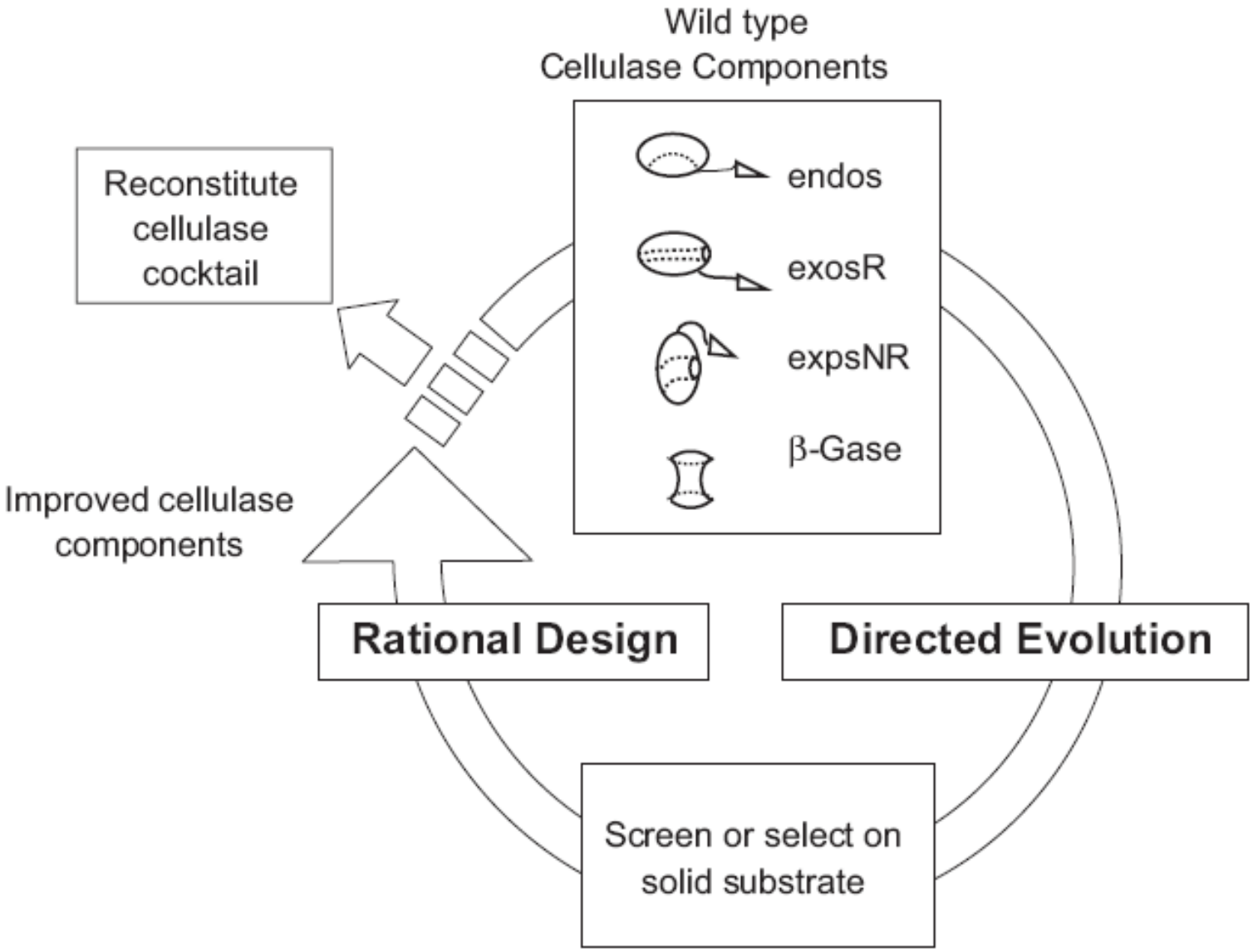


The central role of **Xyr1** in the hydrolytic enzyme system of *T. reesei*



xylanases (as XYNI, XYNII), cellulases (as CBHI, CBHII, EGLI), D-xylose (X1), β-xylosidase I (BXLI), sophorose (SO), β-glucosidase I (BGLI).

Scheme of cellulase engineering for non-complexed cellulases



Topics to be considered

- To improve the molecular and genetic toolbox for model cellulolytic microorganisms

- To undertand microbial metabolism and physiology of these microorganisms

- To use metabolic control analysis and either site-directed mutagenesis or directed-evolution to “pull” the metabolic flow

- To exploit microbial diversity