



Diversity search for biomass-degrading enzymes

April 3, 2008

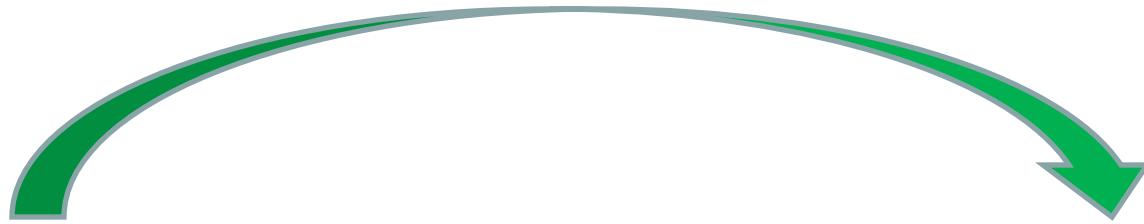
Clean Technology Forum

Adrian Tsang

<https://fungalgenomics.concordia.ca>

Liquid fuels and chemicals

Conversion



Biomass

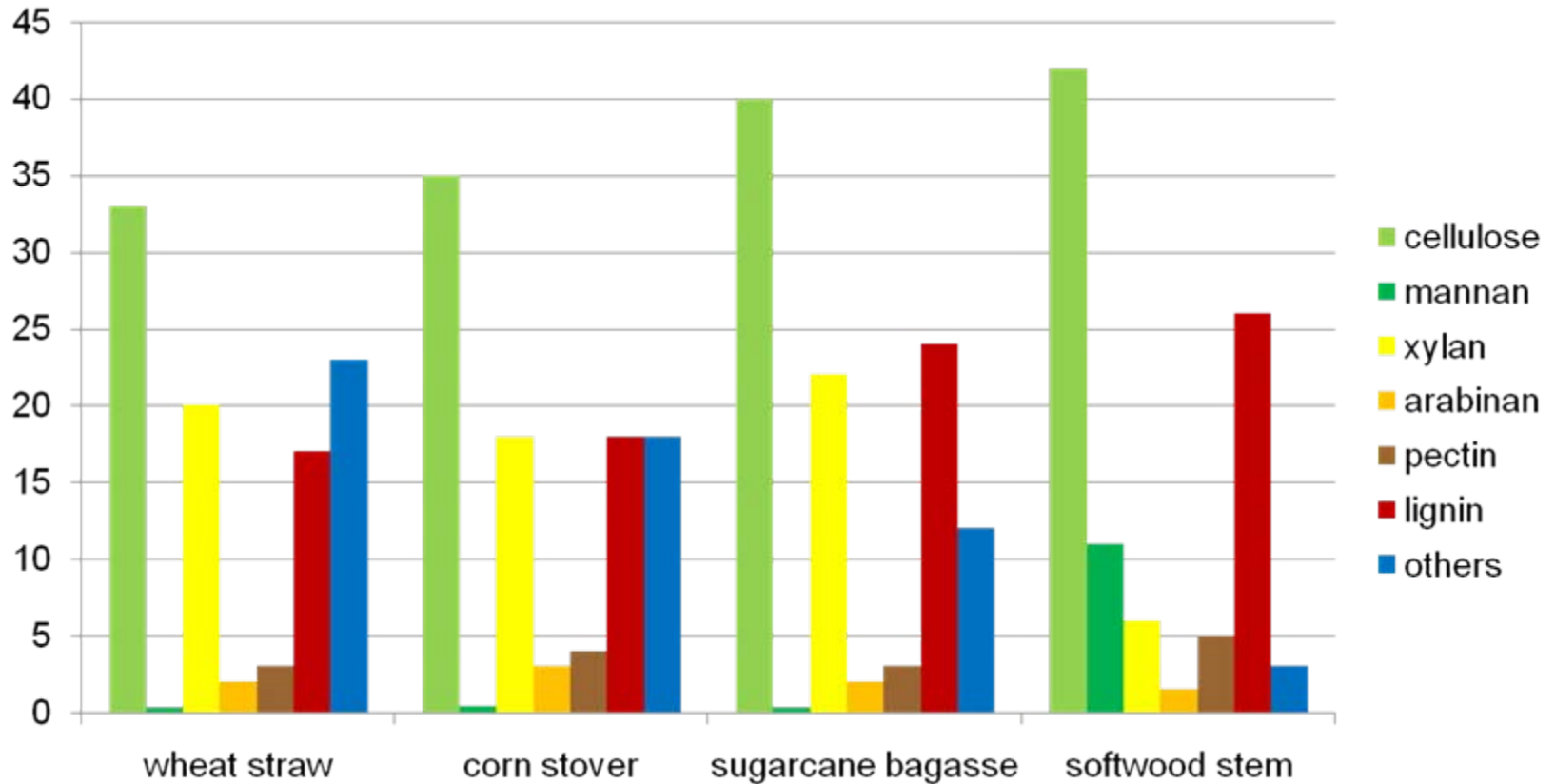


Bioproducts



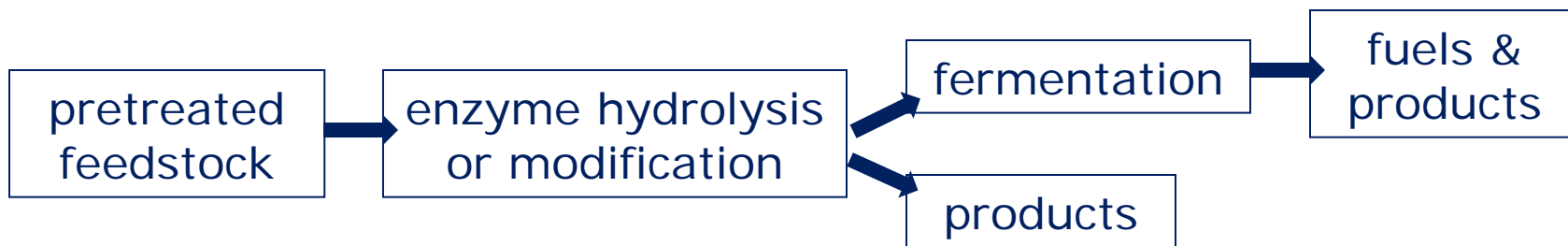
- high volume, low value
- high value, low volume

Lignocellulose composition



Source: US DOE Energy Efficiency and Renewable Energy:
http://www1.eere.energy.gov/biomass/feedstock_databases.html

Sustainable conversion

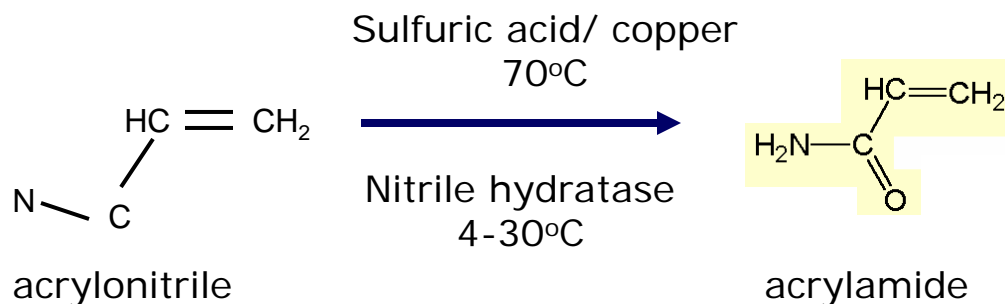
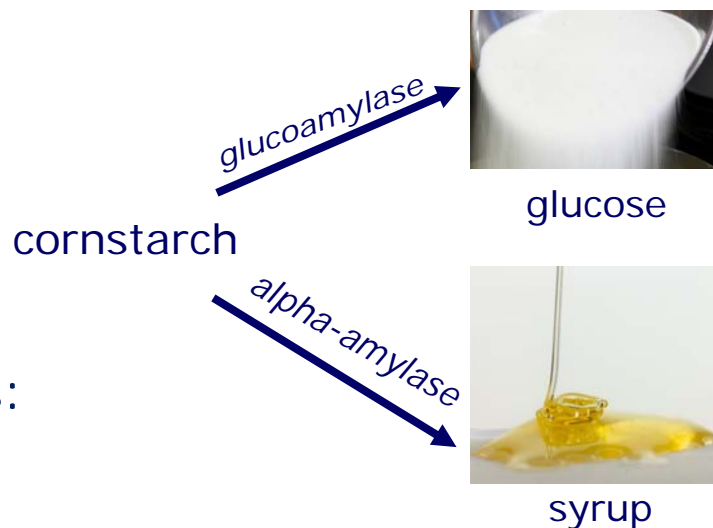


Enzyme advantages:

- specific
- low energy input
- biodegradable

Commercially available enzymes:

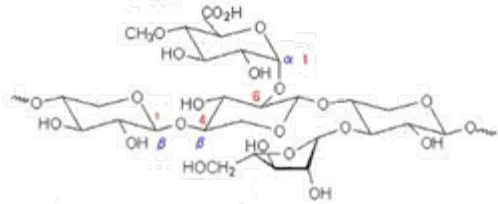
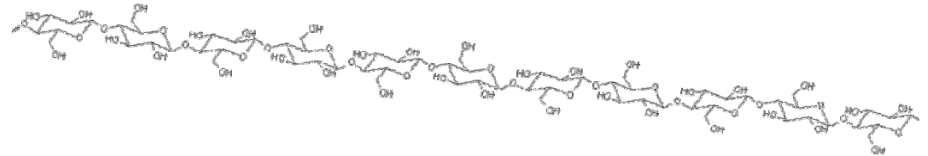
- limited functionality
- long reaction time
- cost



Lignocellulolytic enzymes

Cellulose

- Endoglucanase, exoglucanase, beta-glucosidase
- Sowellinin, expansin

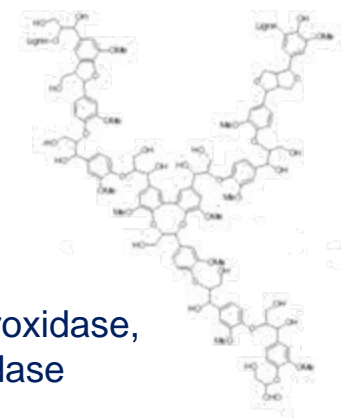
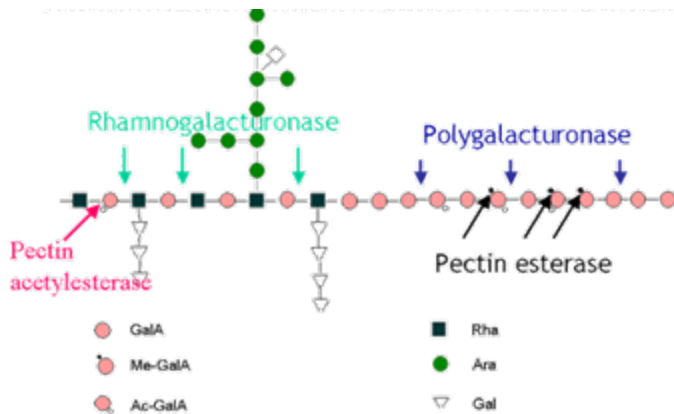


Hemicellulose

Xylan – endoxylanase, xylosidase, acetylxylanesterase, alpha-glucuronidase, feruloly esterase

Mannan – mannanase, mannosidase

Arabinan – arabinase, arabinofuranosidase



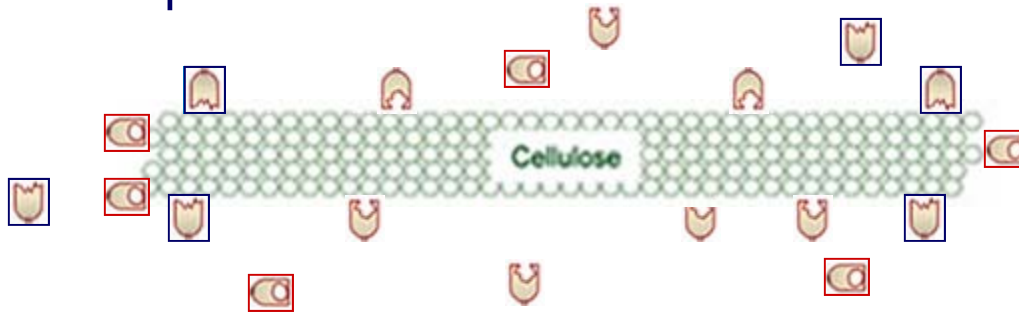
Lignin

- Laccase, lignin peroxidase, manganese peroxidase

Enzyme source

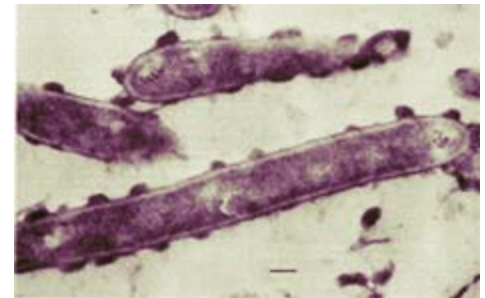
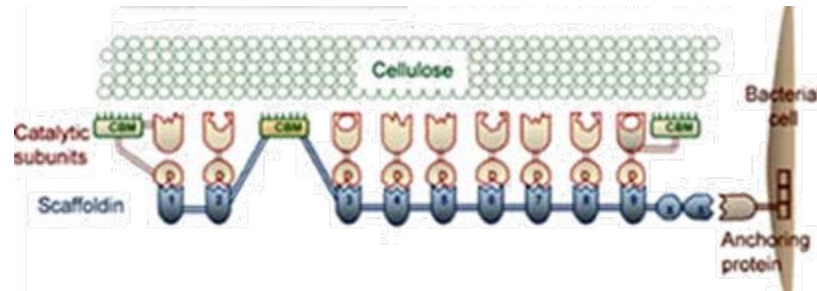
Fungi:

- major decomposers of the biosphere
- secreted proteins

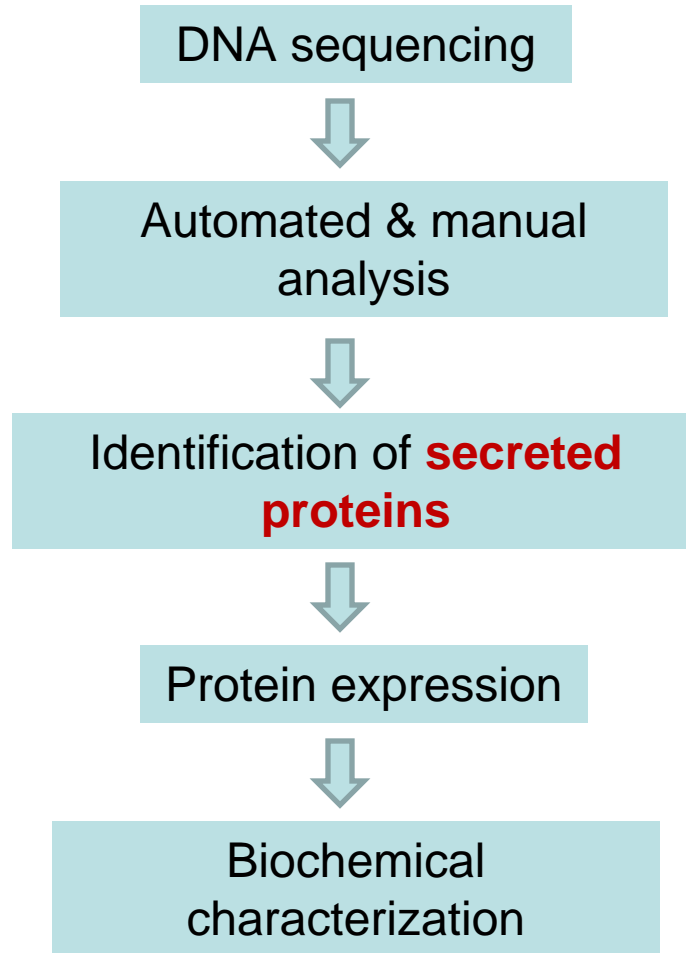


Bacteria:

- decomposers in aquatic environment
- cellulosomes



Genes to enzymes



Functional prediction by sequence comparison

Low-hanging fruits



Lucas Cranach's
Adam and Eve

Putative conserved domains have been detected



```

Query 146 R-----PQRDGPALRATAMIGFGQWLLDNGYTSTATDIVWPLVRNDLSYVAQYWNQTGYD 200
          R-----PQ-DGPALRA-A+I-+----L+D-----AT+-VWP+++DL-Y-A+-WN-TG+D
Sbjct 145 RTLTERPQHDGPALRAIALINYANHLVDTDNSTWATERVWPVIKKDLDTAKNWNNTGFD 204

Query 201 LWEEVNGSSFFFTIAVQHRALVEGSAFATAVGSSCSWCDSQAPEILCYLQSFWTGS--FIL 258
          LWEEV-SSFFT-AVQ--RAL-EG---A-----+-----Q---+-C+LQS+W-----FI
Sbjct 205 LWEEVATSSFFFTSAVQLRALKEGVTLADRLQQNSAPYVEQVERVHCFLQSYWNPEQGFIT 264
  
```

Score	E	Sequences producing significant alignments:	(Bits)	Value
gb AAF75523.1 		glucoamylase [<i>Lentinula edodes</i>]	533	2e-149
dbj BAE47183.1 		glucoamylase [<i>Fomitopsis palustris</i>]	530	2e-148
ref XP_001590395.1 		hypothetical protein SS1G_08135 [<i>Scleroti...</i>]	530	2e-148
emb CAP70294.1 		unnamed protein product [<i>Podospora anserina</i>] ...	525	5e-147

Pioneer genes

Function Unknown (FUN) genes

[GENE ID: 2789037 FG07260.1](#) | hypothetical protein [Gibberella zeae PH-1] Score = 277 bits (709), Expect = 1e-72, Method: Compositional matrix adjust. Identities = 209/547 (38%), Positives = 296/547 (54%), Gaps = 66/547 (12%)

```
Query 1 MDAQYPPFASRDDIWRVFEELKELHATQFEQAERIARLERRRDEDAKLRSVWGPLSPFPSS 60
      MD++++-SR+D++++-E+K++++-Q---AER+-RLE+++++DA-L+SVW---SPFP
Sbjct 1 MDPFRF-VVSREDLYTLQMEVKQVQYAQSNHAERLLRLEKKQADDAALKSVWN--SPFPGV 57

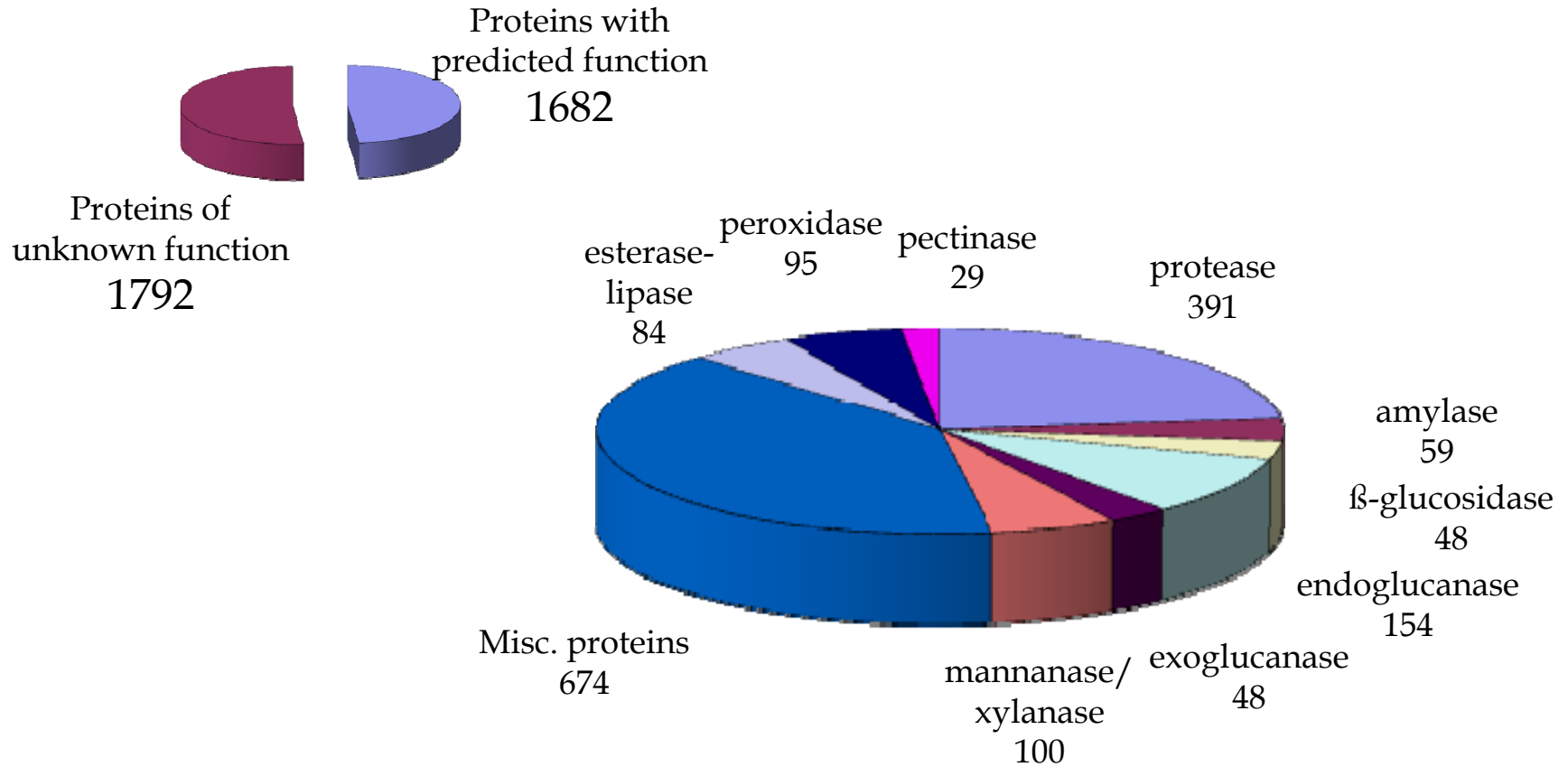
Query 61 VGGAIPTPEVPVFHTPTDAFKGFD-QGHHHGAVSTMGLDNEEEPRRGT-SRANSVRFDESAI 118
      +-G---T-PV-----D-F---D-QG---+G-----EE-RRG---SRANSVRFDESA+
Sbjct 58 LSGTPQTGPVSI PHNDMFDDLDEQGEELLGSLHLGPAEEEPVRRGAASRANSVRFDESAL 117
```



No putative conserved domains have been detected

Score	E Sequences producing significant alignments:	(Bits)	Value
ref XP_965701.2 	hypothetical protein NCU02570 [Neurospora cr...	249	2e-64
ref XP_361068.2 	hypothetical protein MGG_03611 [Magnaporthe ...	238	7e-61
ref XP_389498.1 	hypothetical protein FG09322.1 [Gibberella z...	237	1e-60

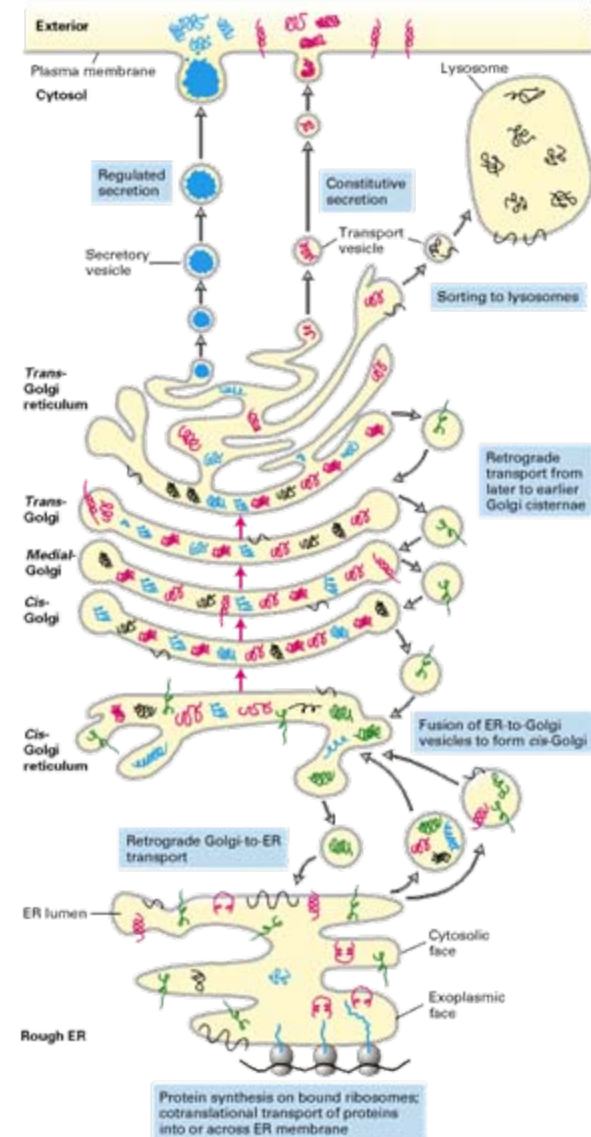
Secretomes of the 15 target fungi



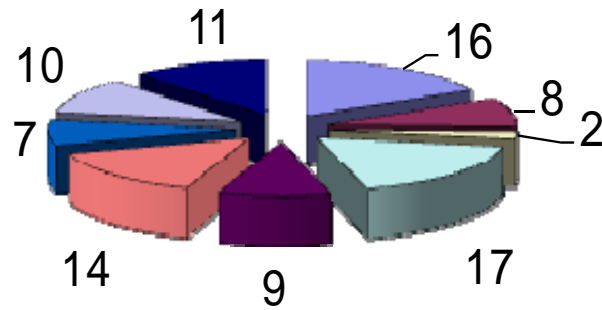
Classical secretory pathway

Proteins with classical signal peptide:

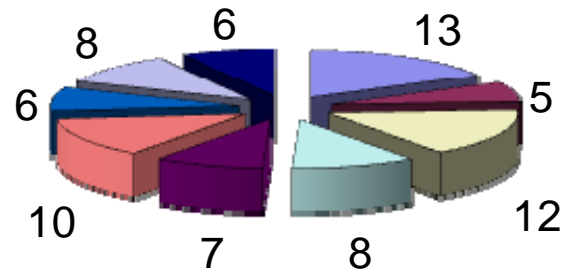
- resident proteins in the lumen and membranes of the endoplasmic reticulum, Golgi complex, lysosomes, vacuoles, and other vesicles
- plasma membrane
- extracellular (secreted) proteins: soluble and cell wall



Mass spectrometric analysis of *A. niger* secretome



xylan



pectin

chitinase/cell wall glucanase

hypothetical

oxidoreductase

xylanase/arabinase

cellulase/cellobiase

lipase/esterase

pectinase

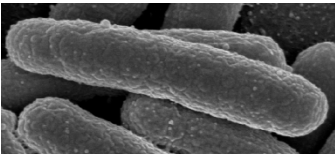
protease/peptidase

misc. hydrolase

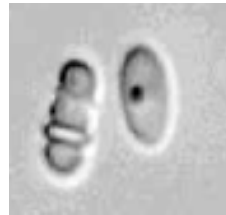
Protein production hosts

Laboratory workhorses:

E. coli

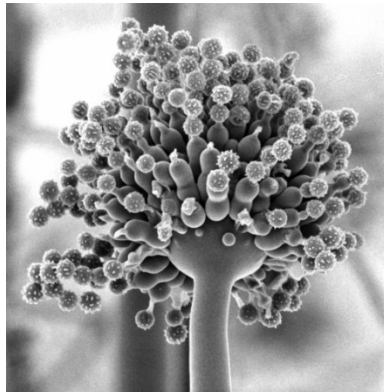


Pichia pastoris



Industrial workhorses:

Aspergillus niger

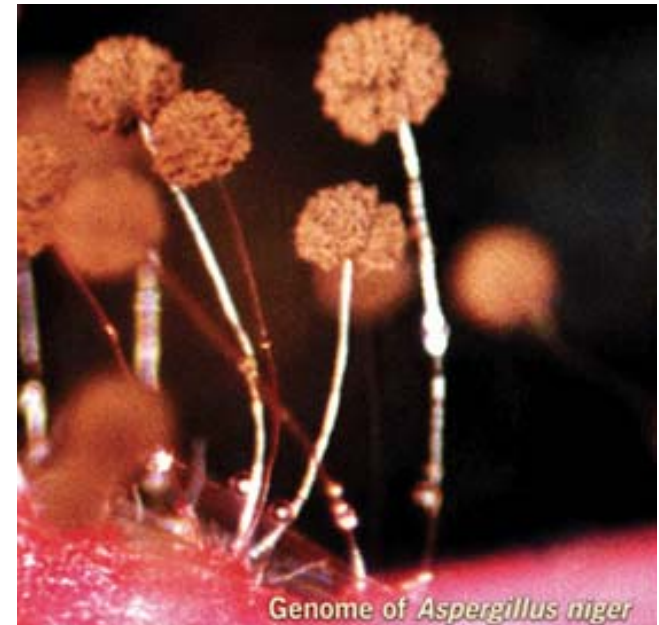


Trichoderma reesei



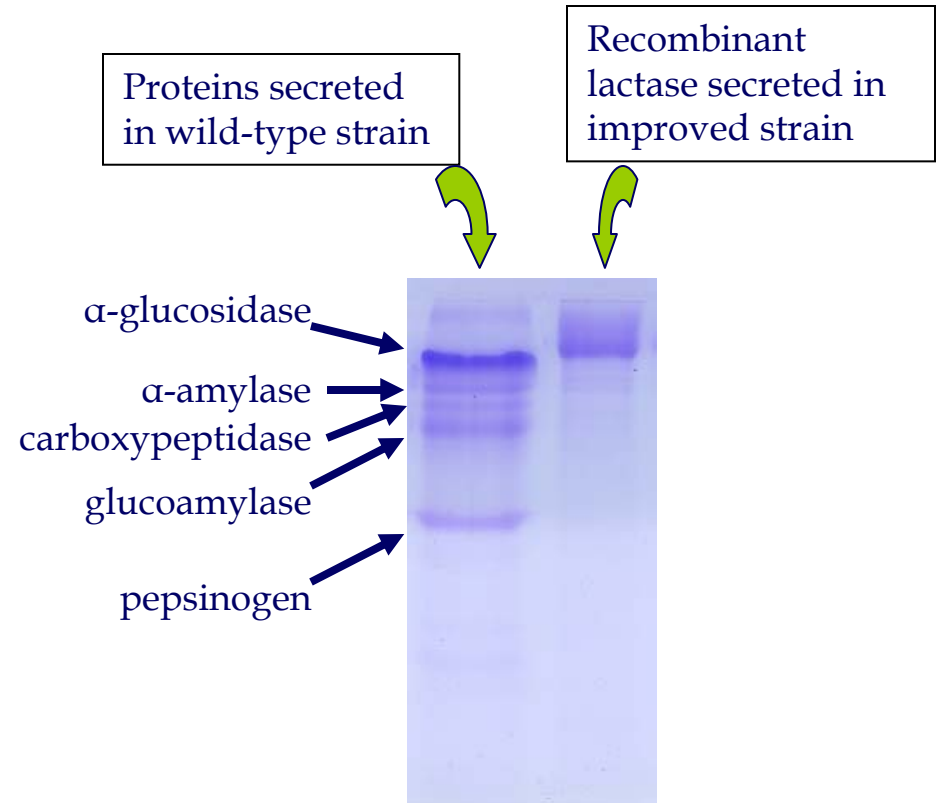
Aspergillus niger as expression host

- Secretes large quantities of recombinant proteins, upwards of 30 grams of protein/litre
- Genome sequence of two strains publicly available
- Used as production host by major enzyme companies
- Most commonly used host for the production of recombinant industrial enzymes



Development of “clean” host strains

- Disrupt genes encoding major secreted proteins and contaminating activities
- Benefits:
 - Reduce background levels in screening
 - Enhance secretion of recombinant proteins

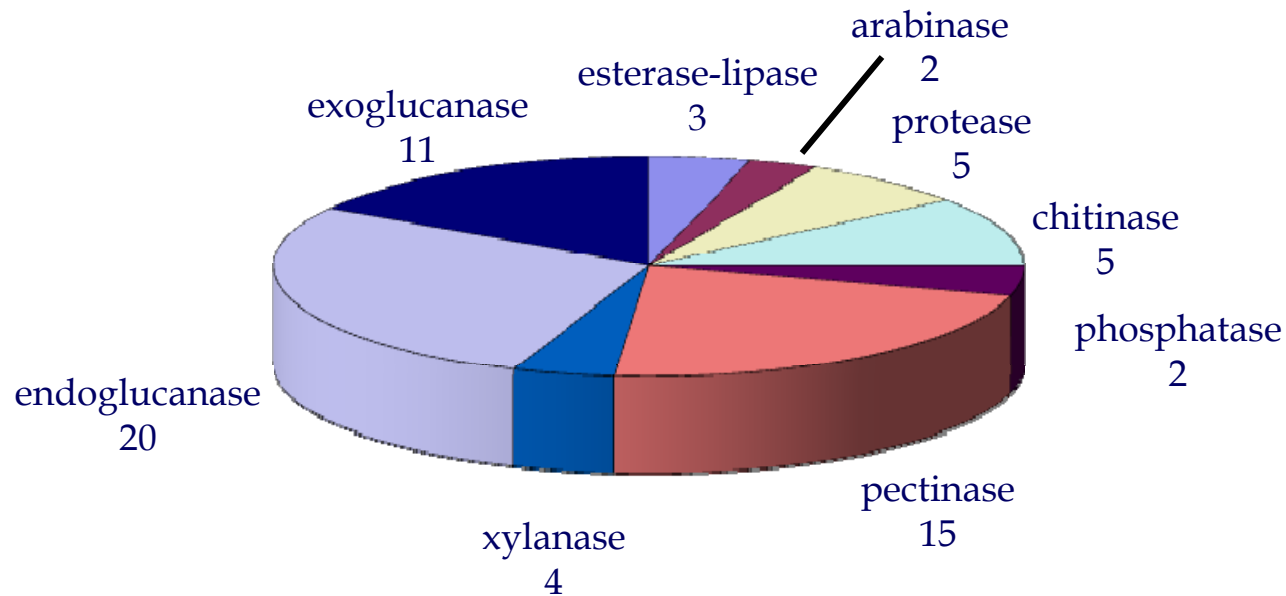


Summary of expression - 2007

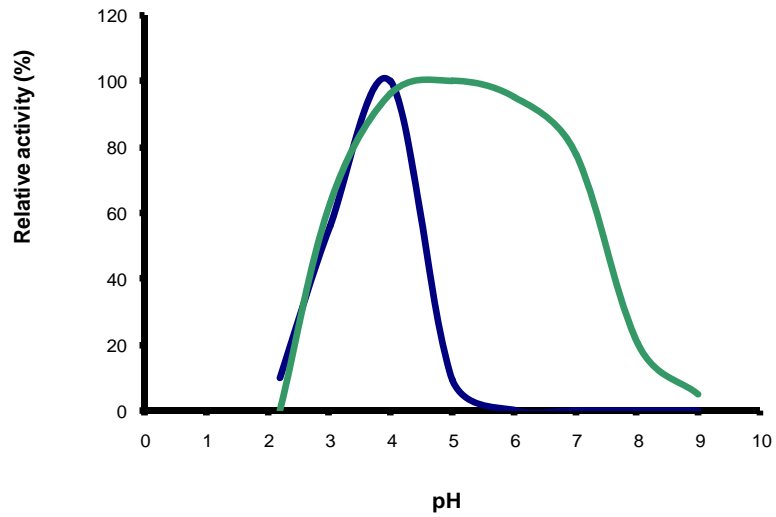
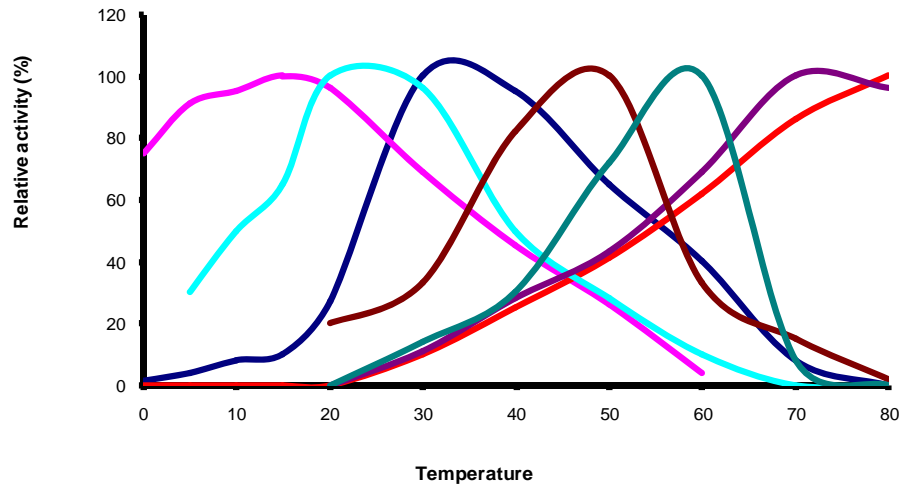
	# of genes		Success rate (%)
	transformed	expressed	
Genes predicted to encode extracellular proteins			
With putative function	708	308	44
With unknown function	1260	332	26
Total	1968	640	32

Function of pioneer genes

Recombinant hypothetical secreted proteins are screened for activity on 36 natural and synthetic substrates.

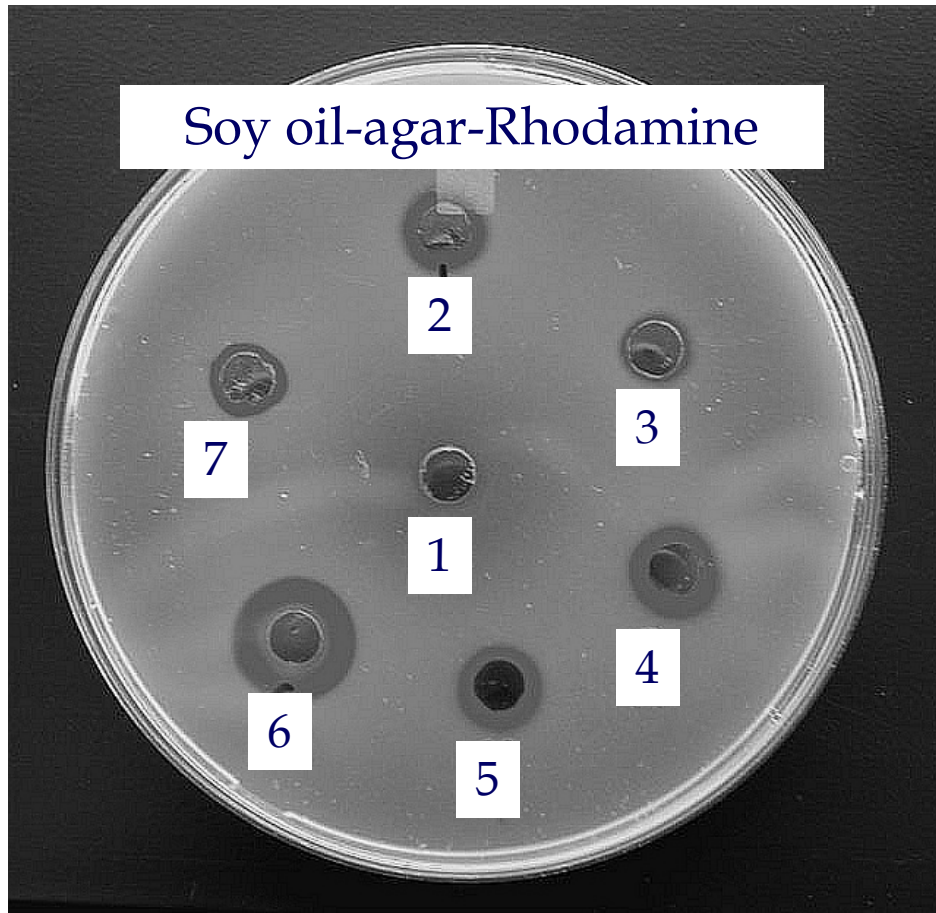


Enzymes with diverse properties



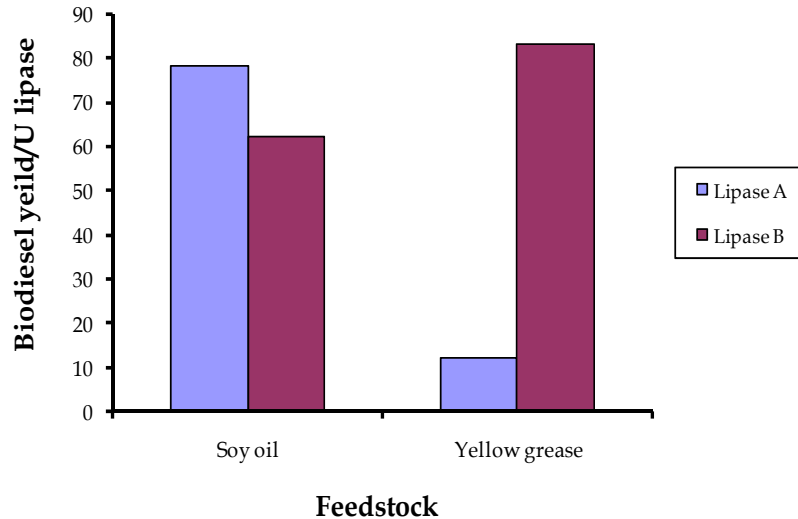
Natural vs synthetic substrates

Activity on synthetic substrate does not always reflect activity on natural substrates.



1. No enzyme
2. 2500 mU *C. rugosa* lipase
3. 250 mU *C. rugosa* lipase
4. 250 mU lipase A
5. 125 mU lipase A
6. 17 mU lipase B
7. 3 mU lipase B

Low-hanging fruits can be delicious



Conversion of fatty biomass into diesel – highly similar enzymes display different conversion properties

```

putative_lipase_A      SNVICTADACPSVEEASTKMLLEFDLTNNFGGTAGFLAADNTNKRLVVAFRGSSTIKNWI 117
putative_lipase_B      TTLTCDVGNCPLEAAGATTIDEFDDTSSYGDPFGFIAVDPTNELIVLSFRGSSDLSNWI 119
Thermomyces_lipase     TNITCTGNACPEVEKADATFLYSFEDSG-VGDVTGFLALDNTNKLIVLSFRGSRSIENWI 112
                        :::** . ** ** *.: : .*: :. * .:** * ** : :::* ** * :***
    
```

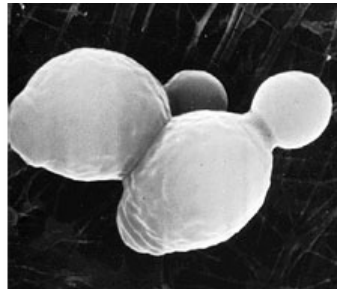
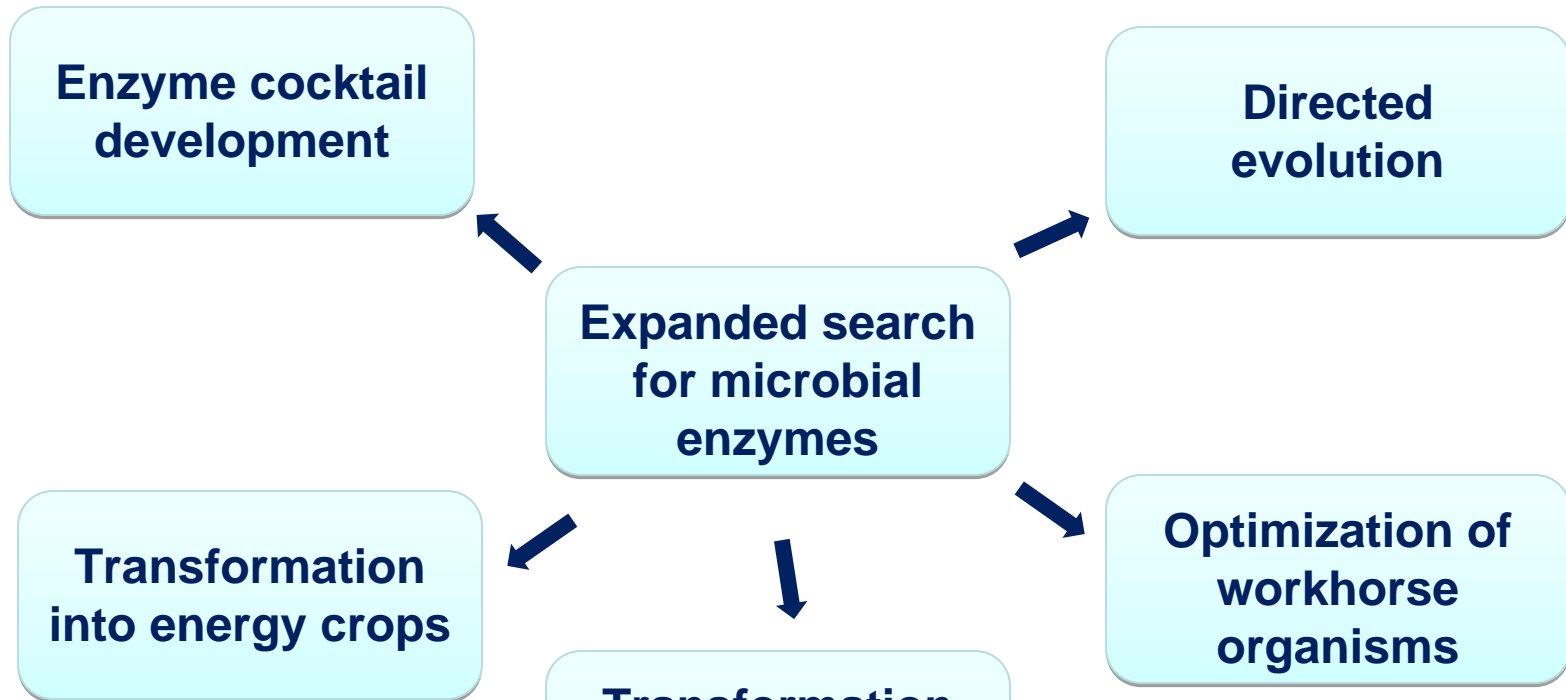
```

putative_lipase_A      ADLDFILQDNDLCTGCKVHTGFWKAWAAAADNLTSKIKSAMSTYSGYTLVFTGHS LGGA 177
putative_lipase_B      ADLDFGLTSVSSICDGCCEMHKGFYEAWVIADTITSKVEAAVSSYPDYTLVFTGHSYGAA 179
Thermomyces_lipase     GNLNFDLKEINDICSGCRGHGFTSSWRVADTLRQKVEDAVREHPDYRVVFTGHS LGGA 172
                        .:** * . .: ** * ** .: . ** .: .: * : .: * : ***** **
    
```

```

putative_lipase_A      LATLGATVLRNDGYSVELYTYGCPRVGNLYALAEHITSQSGGANFRVTHLNDIVPRLPPMD 237
putative_lipase_B      LAAVAATVLRNAGYTLDLNFGQPRIGNLALADYITGQNMGSNYRVVTHDDIVPKLPPEL 239
Thermomyces_lipase     LATVAGADLRNGYDIDVFSYGAPRVGNRAFAEFLTVQGTGGLYRITHNTDIVPRLPPRE 232
                        **::.: ** . ** :::: .: ** ** *:*: * * * : ** * : ***** **
    
```

Current and future activities



Conclusions

- Genomics can fast-track the transition to the bio-based economy by contributing to three areas:
 - Feedstock improvement
 - Sustainable microbial conversion
 - Development of value-added bioproducts
- The industrial workhorse *Aspergillus niger* be adapted for global expression of secreted proteins
- Most extracellular proteins of fungi encode degradative enzymes
- Biochemical properties on synthetic substrates are not necessarily good predictors of activities on natural substrates