



Exploring microbial diversity to further the bioethanol supply chain in Brazil

Dr. rer. nat. Diego Mauricio Riaño Pachón
Evolutionary and Computational Biology Group
Centro Nacional de Pesquisa em Energia e Materiais
Laboratório Nacional de Ciência e Tecnologia do
Bioetanol

<http://bce.bioetanol.cnpem.br/>

Brazilian Bioethanol Science and Technology Lab



FIFA WORLD CUP
Brasil



Rio 2016



Campinas, SP

Brazilian Centre of Research in Energy and Materials (CNPEM)

CTBE is part of Ministry of Science and Technology with other three National Labs



Brazilian Centre of Research in Energy and Materials (CNPEM)

CTBE is part of Ministry of Science and Technology with other three National Labs



Sirius: 3G under construction (2015)



Brazilian Synchrotron Light Lab

Set of scientific instrumentation for the application of UV and X-ray on studies of Advanced Materials, Nanosciences, Biological Materials, Oil Industry, Metallurgy



Brazilian Nanotechnology National Laboratory

Create new advanced functional materials required for sustainability



Brazilian Biosciences National Lab

Biotechnology and drug development



Brazilian Bioethanol Science
and Technology Laboratory

www.bioethanol.org.br



Ministry of Science
and Technology





Brazilian Bioethanol Science
and Technology Laboratory

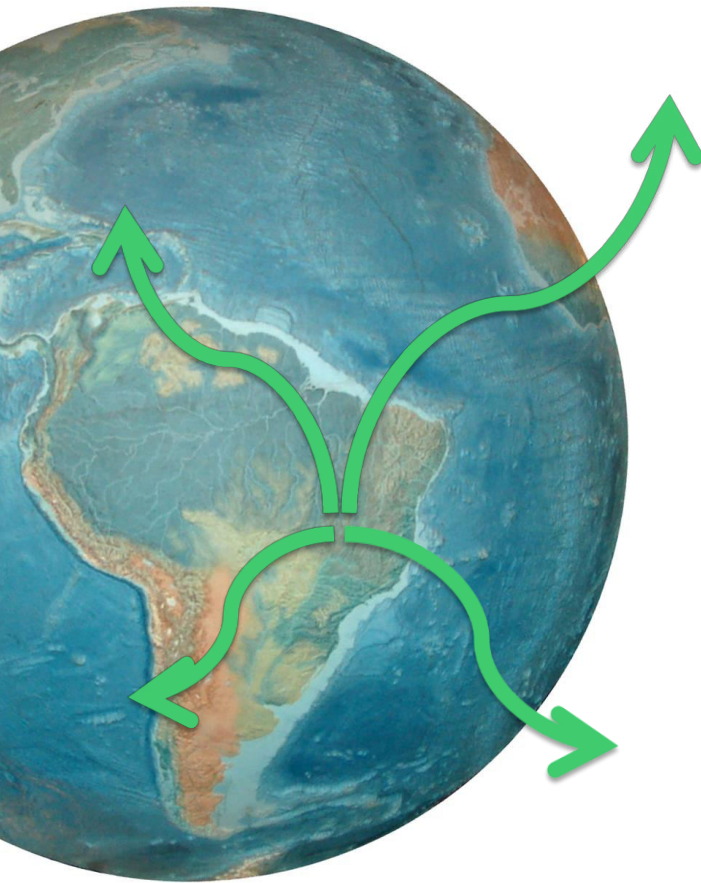
*Basic science and innovation:
Fundamental initiatives to keep Brazilian
leadership in sugarcane/ethanol
production cycle.*



Ministry of Science
and Technology




Replacing 10% of the world demand for gasoline by Brazilian ethanol in 2025



Ethanol Project (NIPE/Unicamp-CGEE/MCT)

A production of 250 billion liters of ethanol could generate in Brazil:

- Over **9 million new jobs** (direct, indirect and induced).
- **A raise of 13% in the GDP.**
- **1000 new distilleries**



Strategy: To create a National Laboratory that can produce scientific knowledge on the bioethanol production cycle, and able to face technological bottlenecks.



Brazilian Bioethanol Science
and Technology Laboratory

A National Laboratory on Bioethanol

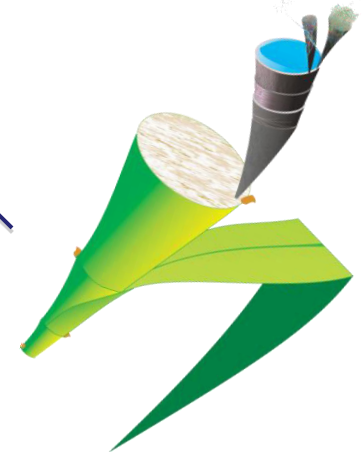
Dedicated to Research,
Development and
Innovation

**Infrastructure
available to research
institutions**

In touch with industrial
problems and solutions

Numbers:

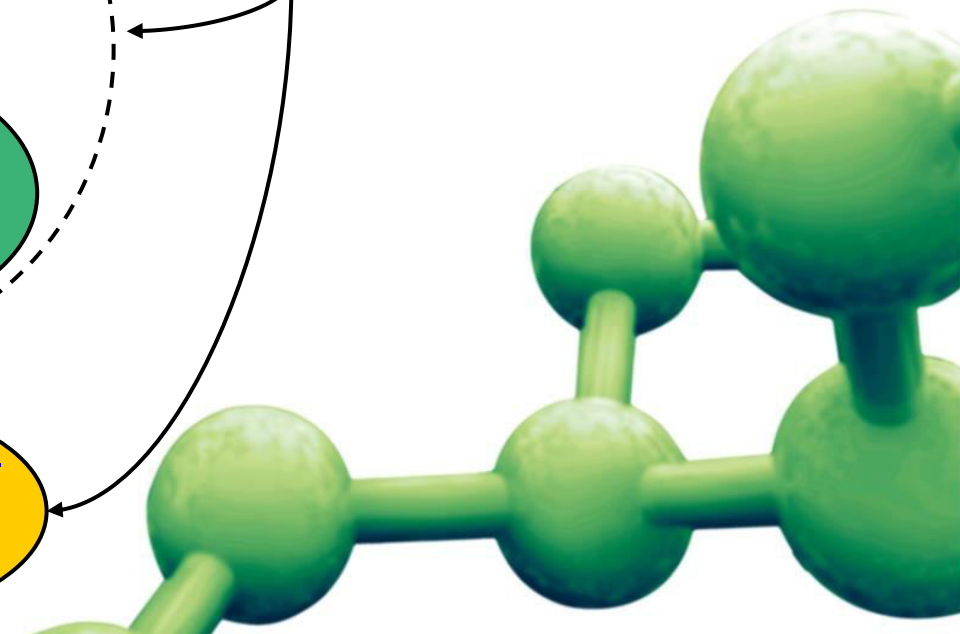
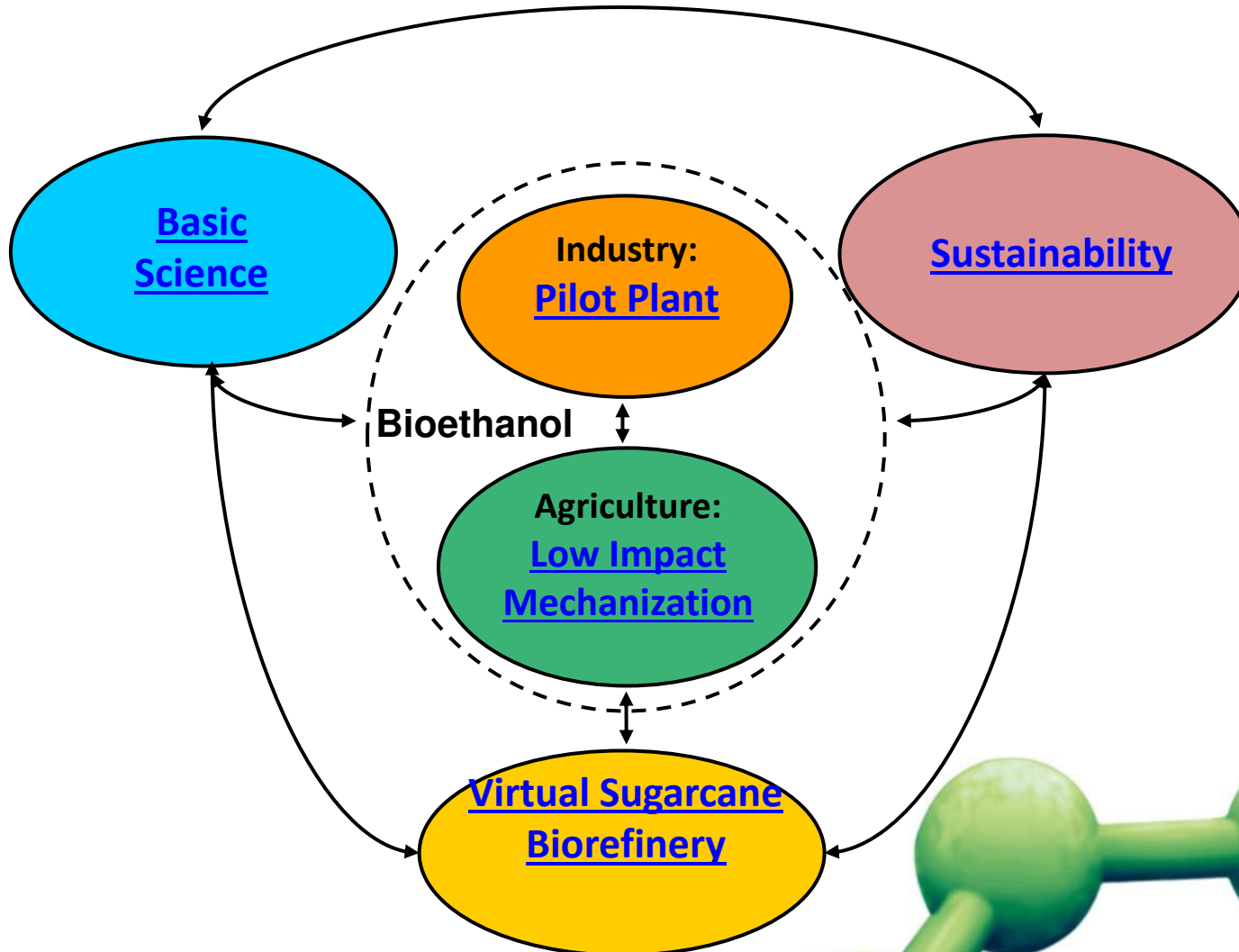
- . Federal Funding: **~R\$ 90 million**
- . Buildings: **8.722,28 m²**
- . Research team by 2014 (biologists, physicists, chemists and engineers): **170 employees**



Ministry of Science
and Technology



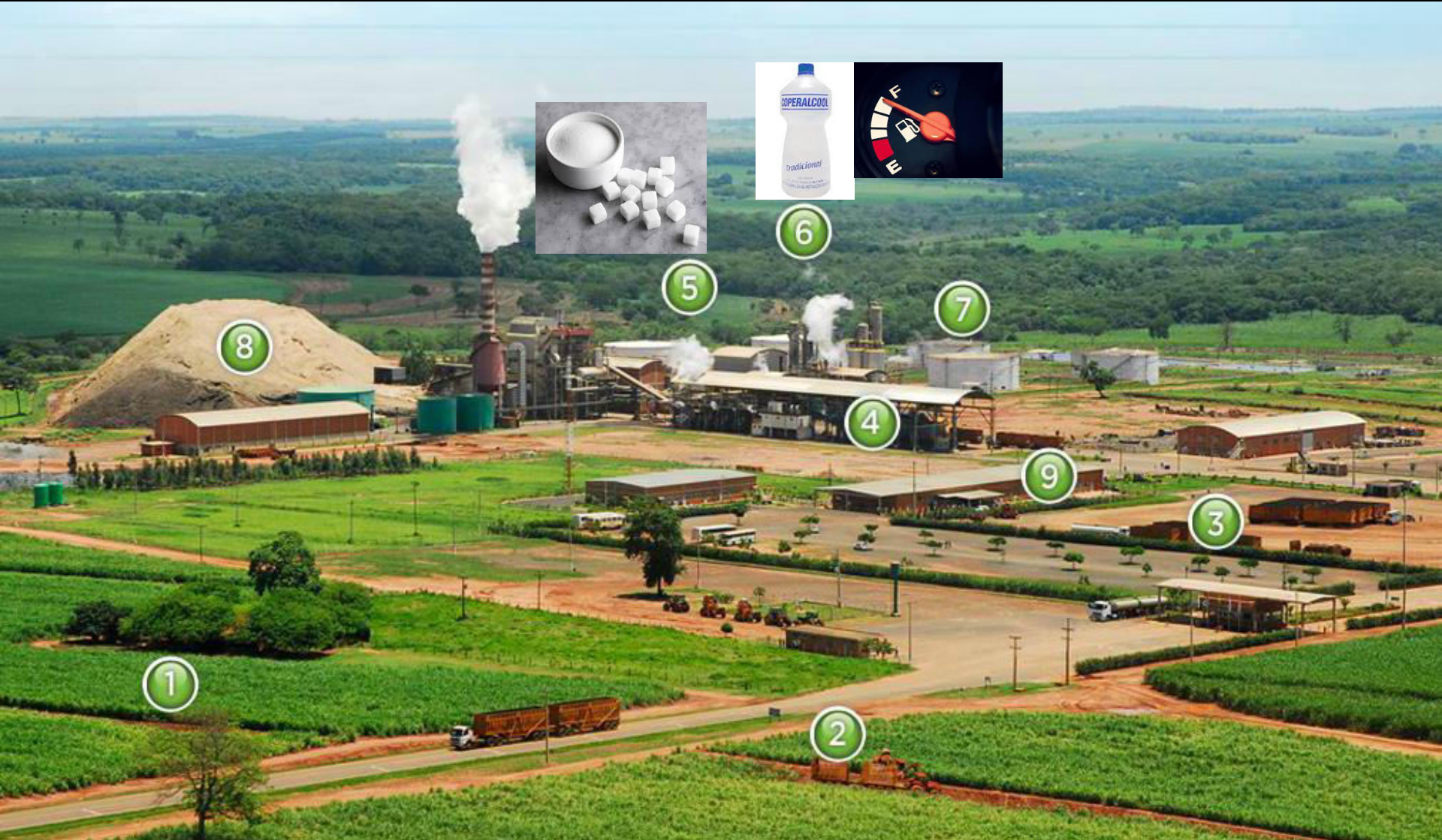
Research Programs



Exploring microbial diversity to further bioethanol production in Brazil

Current **collaboration** with Prof. Dr.
Gustavo H. Goldman and Dr. Juliana
Velasco

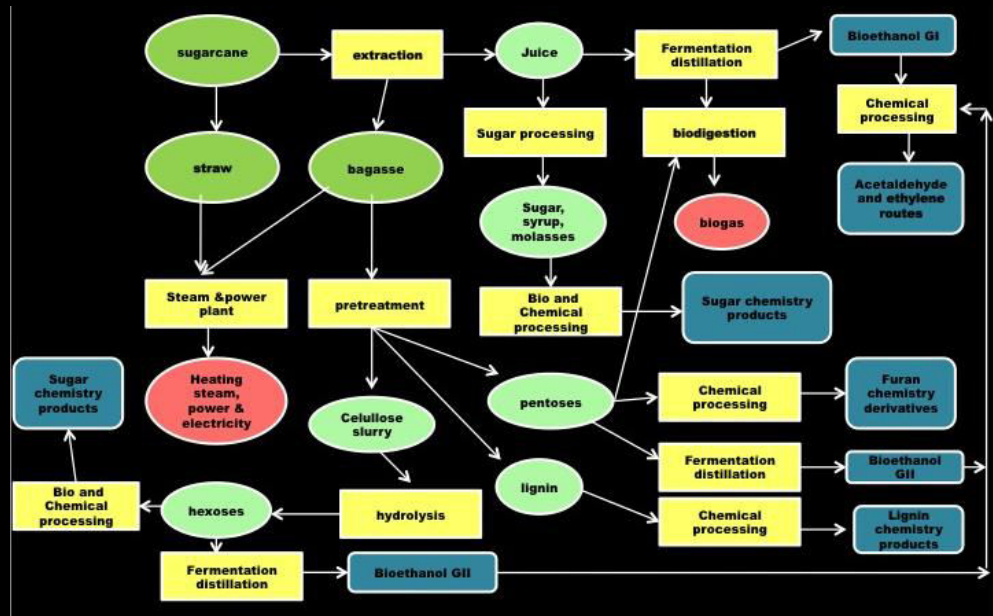
Sugarcane processing Usinas/Mills/Ingenios



Other products from the bioethanol/sugar supply chain



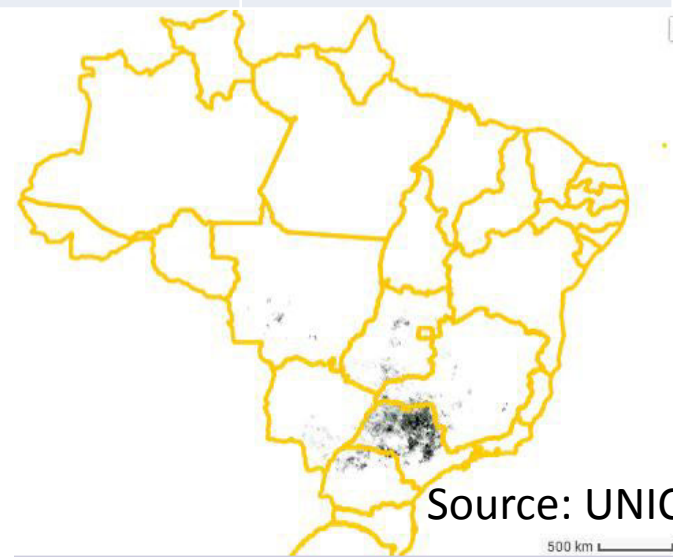
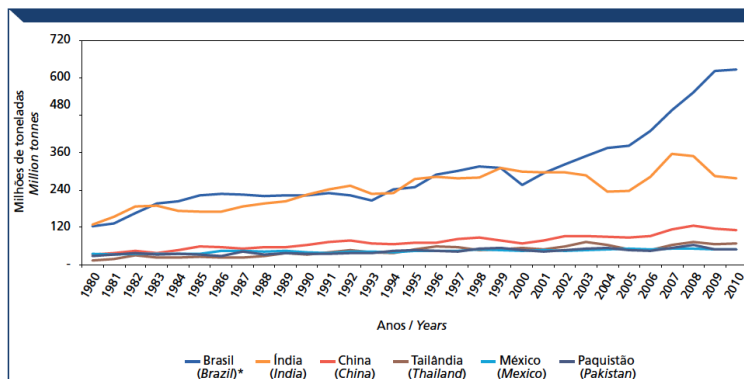
Biorefinery



Some numbers about sugarcane in Brazil (South-Central region)

Product		Harvest		
		2011/2012	2012/2013	2013/2014
Sugarcane ¹		493.159	532.758	594.101
Sugar ¹		31.304	34.068	34.265
Total Ethanol ²		20.542	21.285	25.373
Mix (%)	Sugar	48,44%	49,54%	45,36%
	Ethanol	51,56%	50,46%	54,64%

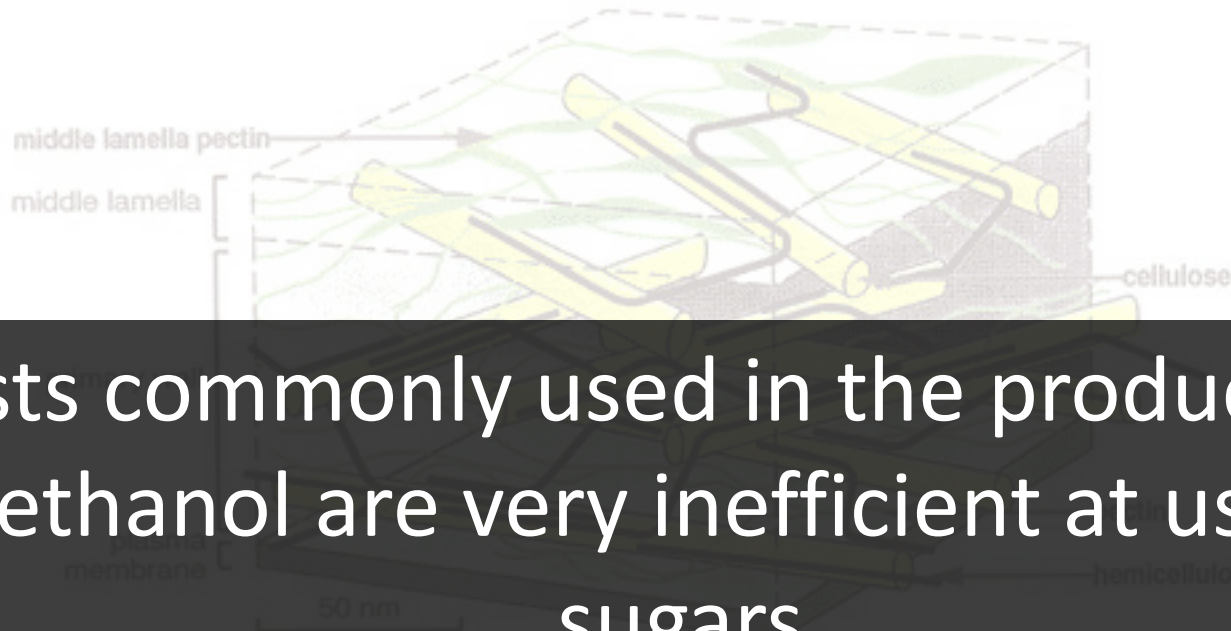
Brazil World Lider in Sugarcane production



Source: UNICA

1. Thousand tons
2. Million liters

Lignocellulosic ethanol (2G)



Yeasts commonly used in the production of bioethanol are very inefficient at using 5C sugars

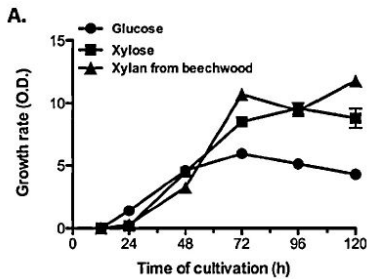
Table 1 Occurrence of hemicelluloses in primary and secondary walls of plants

Polysaccharide	Dicot walls		Grass walls		Conifer walls	
	Primary	Secondary	Primary	Secondary	Primary	Secondary
Xyloglucan	20–25	Minor	2–5	Minor	10	– ^b
Glucuronoxylan	–	20–30	–	–	–	–
5C Glucuronoarabinoxylan	5	–	20–40	40–50	2	5–15
(Gluco)mannan	3–5	2–5	2	0–5	–	–
Galactoglucomannan	–	0–3	–	–	+ ^b	10–30
β -(\rightarrow 3,1 \rightarrow 4)-glucan	Absent	Absent	2–15	Minor	Absent	Absent

Where can we find
organisms capable of using
5C sugars?

Let's look at the
gastrointestinal track of
insects living on/in
sugarcane soils/roots

Organisms Selection



Xylanase activity

Borges *et al.*, 2014. Process Biochemistry 49:77-83

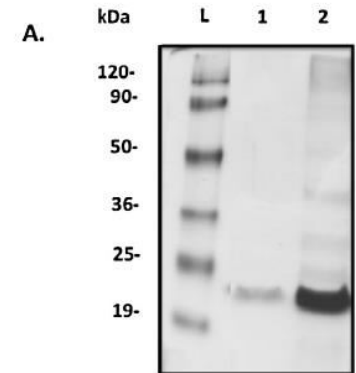


Table 2
Properties of G11 eukaryotic endo-1,4-xylanases.

Xylanase	Species	Mass (kDa)	V_{max} (μ moles product/min/mg protein)
PbXynA	<i>Pseudozyma brasiliensis</i> sp. nov.	24	2792.0
-	<i>Aspergillus ficuum</i> AF-98	35.0	18.38
XynG2	<i>Aspergillus oryzae</i> RIB40	21	123.0
EX2	<i>Trichoderma reesei</i> K1201	20	526.0
XylG	<i>Pachydictyonella</i> sp.	45	17.7

Cloning/PCR of the gene: Unsuccessful

Genome sequencing

Characteristic	<i>P. brasiliensis</i>	<i>P. antarctica</i>	<i>P. aphidis</i>	<i>P. flocculosa</i>	<i>P. hubeiensis</i>	<i>U. maydis</i>	<i>M. globosa</i>
Assembly statistics							
Total scaffold length (Mbp)	17.33	18.07	17.92	23.31	18.44	19.74	8.96
N50 scaffolds (Mbp)	2.21	2.38	2.38	1.98	1.40	0.51	1.41
N90 scaffolds (Mbp)	1.72	2.07	2.07	1.48	1.09	0.48	1.01
GC content (%)							
Assembled	33.85	42.03	37.51	36.12	39.33	34.13	37.00
Protein coding	58.10	60.28	60.50	64.30	56.49	53.95	52.06
Number of protein coding genes	5768	6640	6011	6877	7472	6522	4286

?

<http://bce.bioetanol.cnpem.br/cgi-bin/gb2/gbrowse/pseudozyma/>



Genome sequence available in Genbank WGS: **AWXO01000000**

Short reads available at SRA: **ERX311387**

Identification of isolates

Be careful with BLASTphemy!

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

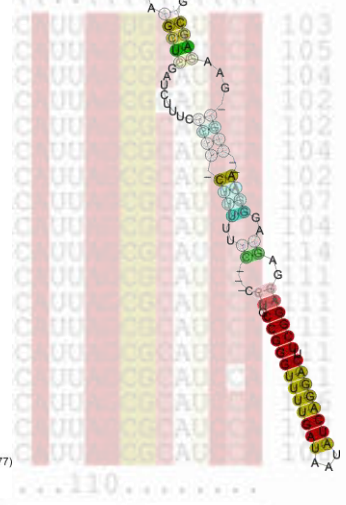
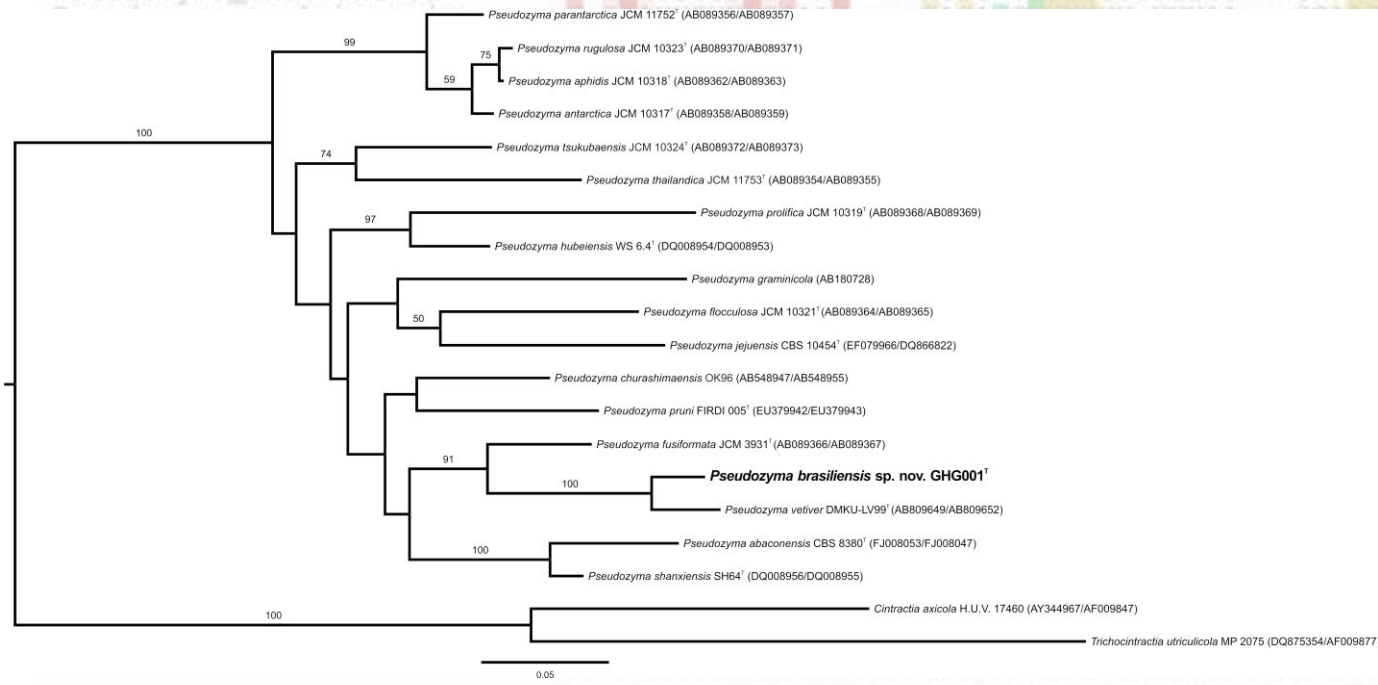
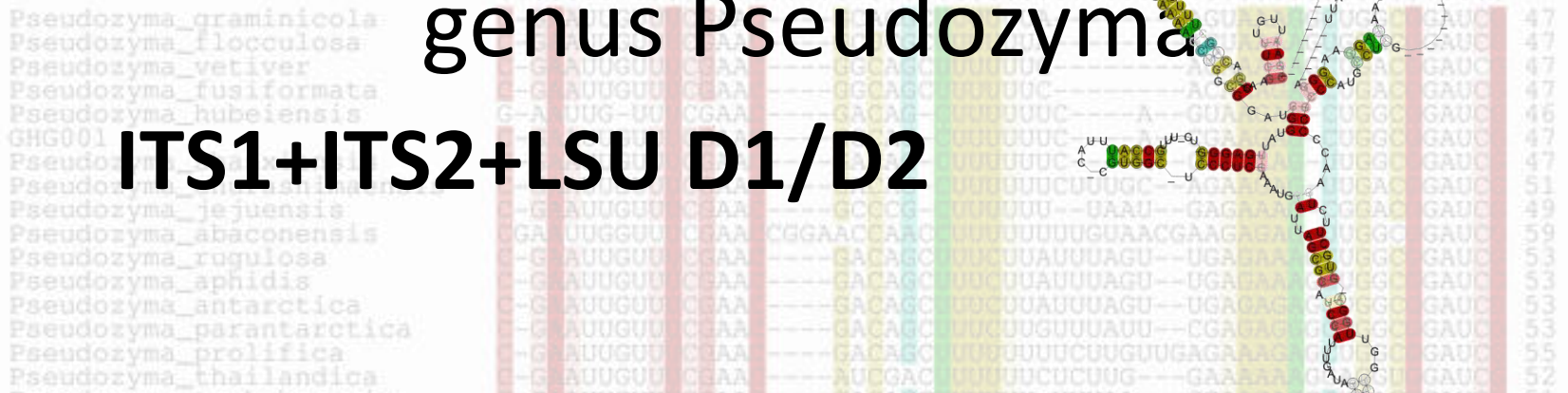
Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
AM901705.1	Uncultured basidiomycete ITS region including 18S rRNA gene, ITS1,	520	520	92%	5e-144	89%	
GU056026.1	Uncultured Ustilaginaceae isolate 24.27c 18S ribosomal RNA gene, pa	516	516	93%	6e-143	89%	
HQ832814.1	Pseudozyma hubeiensis isolate LH146 18S ribosomal RNA gene, parti	514	514	93%	2e-142	89%	
FJ515182.1	Pseudozyma shanxiensis strain SN37 18S ribosomal RNA gene, partial	508	508	93%	1e-140	88%	
AY740055.1	Sporisorium nealii internal transcribed spacer 1, partial sequence; 5.8	503	503	93%	5e-139	88%	
AY740043.1	Sporisorium andropogonis internal transcribed spacer 1, partial seque	499	499	93%	6e-138	88%	
AY740164.1	Macalpinomyces tristachyae internal transcribed spacer 1, partial se	497	497	93%	2e-137	88%	
AY740168.1	Ustilago cynodontis internal transcribed spacer 1, partial sequence;	494	494	93%	3e-136	88%	
HM143013.1	Ustilago cynodontis isolate UE internal transcribed spacer 1, partial s	492	492	93%	1e-135	88%	
FJ919774.1	Pseudozyma fusiformata isolate AP6 internal transcribed spacer 1, pa	490	490	89%	4e-135	89%	
AY740172.1	Ustilago striiformis internal transcribed spacer 1, partial sequence; 5	488	488	93%	1e-134	88%	
AF038825.1	Ustilago cynodontis internal transcribed spacer 1, 5.8S ribosomal RN	483	483	93%	5e-133	87%	
EU480258.1	Uncultured soil fungus clone CS2M5c37P internal transcribed spacer	481	481	93%	2e-132	87%	
AY740165.1	Ustilago alcornii internal transcribed spacer 1, partial sequence; 5.8S	481	481	93%	2e-132	87%	
GU139172.1	Sporisorium spinulosum voucher HMAS 193085 18S ribosomal RNA gen	479	479	93%	8e-132	87%	
EU516997.1	Uncultured Ustilago clone IIN4-14 18S ribosomal RNA gene, partial se	479	479	93%	8e-132	87%	
AB180728.1	Pseudozyma graminicola genes for ITS1, 5.8S rRNA, ITS2	479	479	93%	8e-132	87%	
AY740054.1	Sporisorium modestum internal transcribed spacer 1, partial sequ	479	479	93%	8e-132	87%	
AY740152.1	Macalpinomyces simplex internal transcribed spacer 1, partial sequ	473	473	93%	4e-130	87%	
AY188382.1	Ustilaginomycete sp. DD-0128.9 internal transcribed spacer 1, parti	472	472	93%	1e-129	87%	
AY344970.1	Sporisorium aegypticum internal transcribed spacer 1, partial sequ	468	468	93%	2e-128	87%	
AY740073.1	Ustilago trichophora internal transcribed spacer 1, partial sequ	468	468	93%	2e-128	87%	

WRONG!

Identification of a new species in the genus *Pseudozyma*

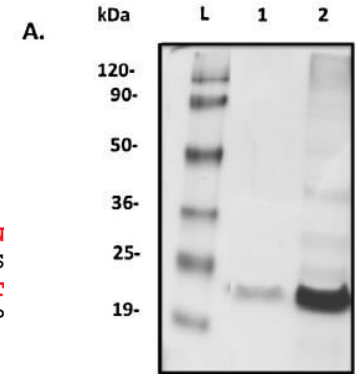
ITS1+ITS2+LSU D1/D2



Identifying the gene responsible for the function – exploiting the genome information

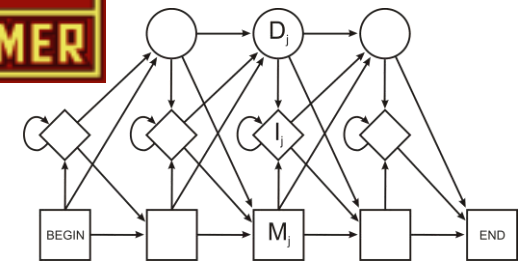
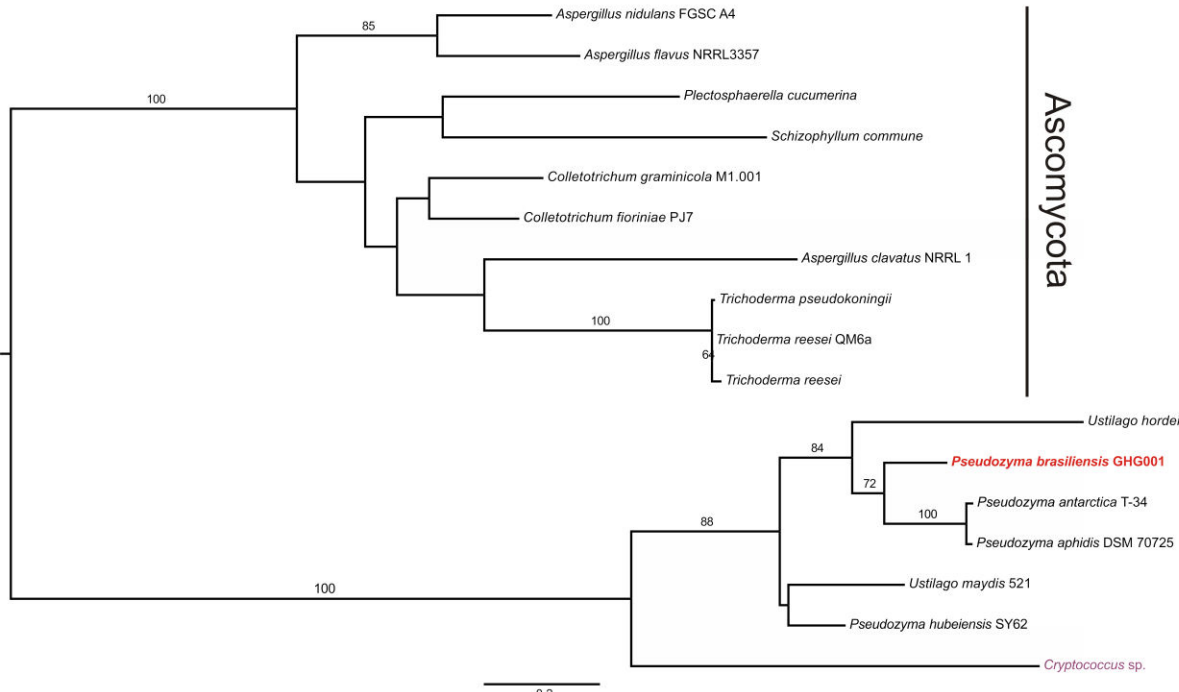
Mass Spectrometry: Coverage 24%

1 MKFSTILAFV GAAGAVAASP IAAPSGESAG LEKR**Q**SIN**Y**V **Q**NYNGN
 51 **K**YNEGAGTYS GNWNNP**S**DFV IGLGWSTGTS NRVIN**F**NGNY QSN**Q**GS
 101 YGWLNNPLTE **Y**VVENYSYD PCSV**S**NTQVV GSVYSDGASY KICK**H**T
 151 **P**SI**Q**GT**K**T**F**G **Q**Y**F**SV**R**SSKR **S**SG**S**VT**L**AN**H** **F**N**A**N**K**K**Y**G**F**A NGATNP
 201 QVFATEAFSG QGSVSTTISG



Endo-1,4- β -xylanase

RAxML WAG+I+ Γ 500 bootstrap replicates

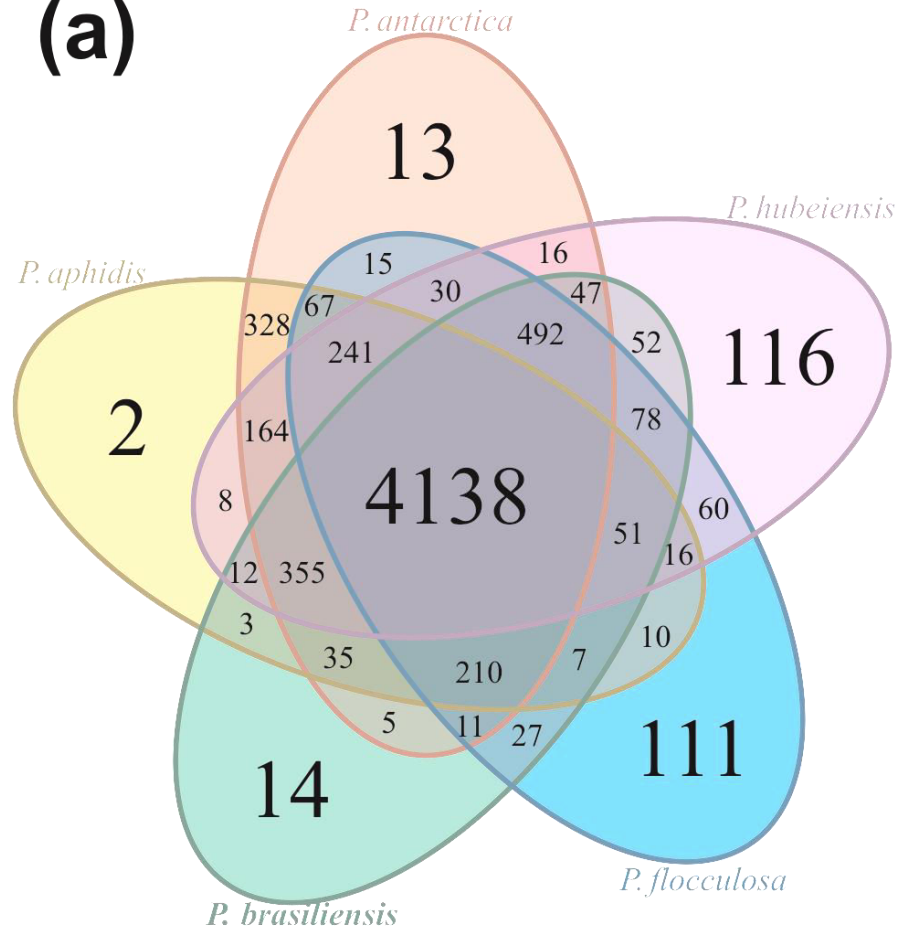


Basidiomycota

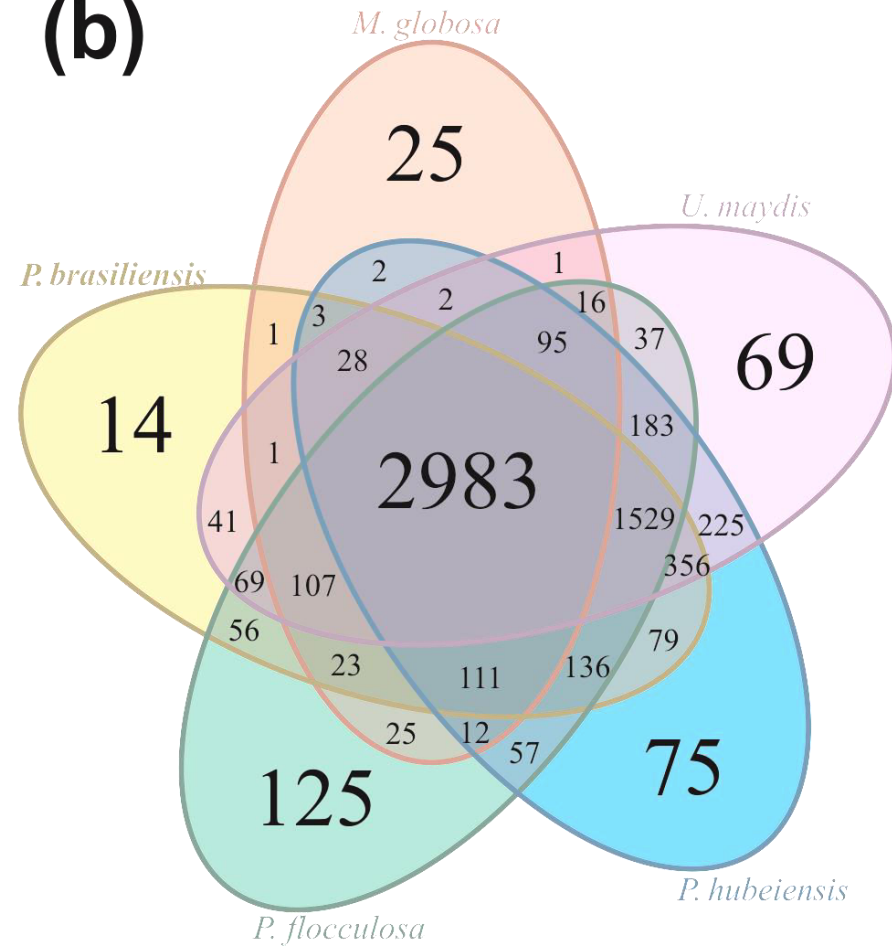


Gene families in Pseudozyma

(a)



(b)



NEXT:

Xylose transportes?

Evolution of xylose uptake in Basidiomycetes

Some other projects

- Comparative genomics of cellulose producing bacteria
- Genome sequencing, assembly and annotation of sugarcane cultivars
 - Deciphering gene regulatory networks in C4 photosynthesis

Want more details?
Interested in joining
the group?

<http://bce.bioetanol.cnpem.br/>

<http://www.bioetanol.cnpem.br/>

diego.riano@bioetanol.org.br

Call for data analysis papers

Community standards for data access, interoperability and metadata only make sense if data are creatively reused to further research. We are therefore inviting the submission of Analysis papers that reformat and integrate existing data sets to generate substantial novel insights into gene expression in cell differentiation transitions and different cell fates.

THANKS

<http://bce.bioetanol.cnpem.br/>

diego.riano@bioetanol.org.br