e-Science as a tool for manipulating large data sets to understand the responses of plants to climate change and their use for bioenergy

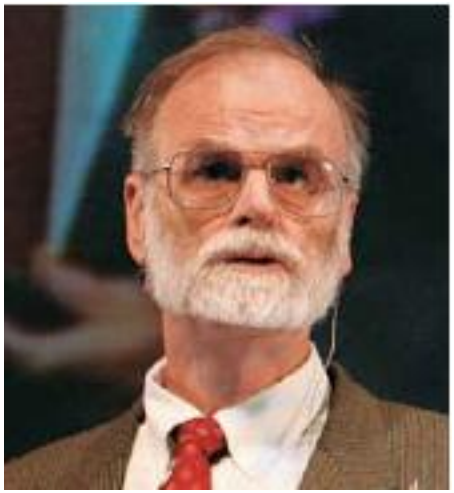
Marcos Buckeridge

LAFIECO, IB USP/CTBE Campinas

What is eScience ?

Jim Gray on eScience: A Transformed Scientific Method

Based on the transcript of a talk given by Jim Gray to the NRC-CSTB[1] in Mountain View, CA, on January 11, 2007[2]



"eScience is where "IT meets scientists." ……

When the data finally shows up in your computer, what do you do with all this information that is now in your digital shoebox?"

Science Paradigms

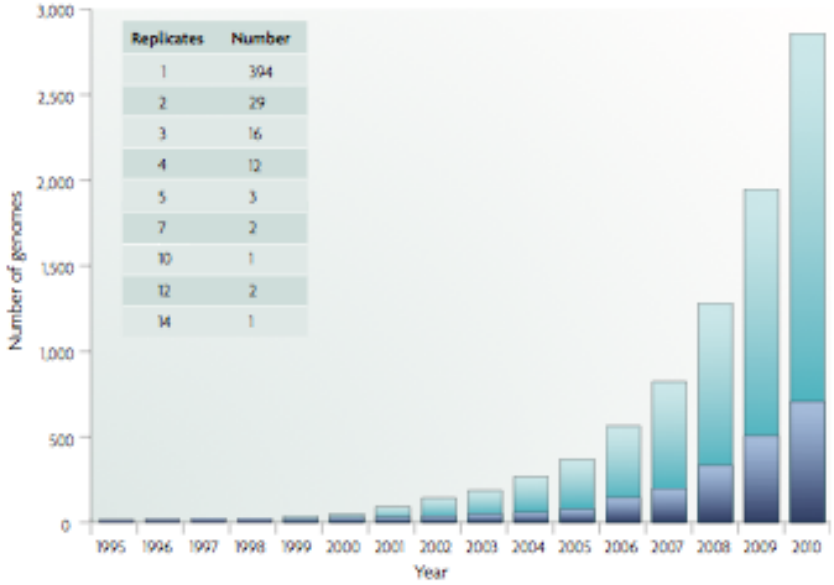
- Thousand years ago: science was **empirical**, describing natural phenomena
- Last few hundred years: **theoretical** branch, using models, generalizations
- Last few dacades: a **computational** branch, stimulating complex phenomenaa
- Today: **data exploration** (**eScience**), unify theory, experiment and simulation

  - *Data captured by instruments or generated by a simulator*
  - *Processed by software*
  - *Informational knowledge stored in computer*
  - *Scientist analyzes database/files using data management and statistics*

Jim Gray on eScience: A Transformed Scientific Method

Based on the transcript of a talk given by Jim Gray to the NRC-CSTB[1] in Mountain View, CA, on January 11, 2007[2]

e-Science: *we are flooded with information*



Replicates	Number
1	394
2	29
3	16
4	12
5	3
7	2
10	1
12	2
14	1

Figure 1 | **Publicly available bacterial genomes.** Light-blue bars show the accumulated number of sequences; dark-blue bars show the number of sequences published in the indicated year. The table shows the frequency with which replicate genomes of the same species have been sequenced (data from KEGG (Kyoto Encyclopedia of Genes and Genomes) bacterial genomes; see Further information).

Craddock et al. 2008. Nature Reviews

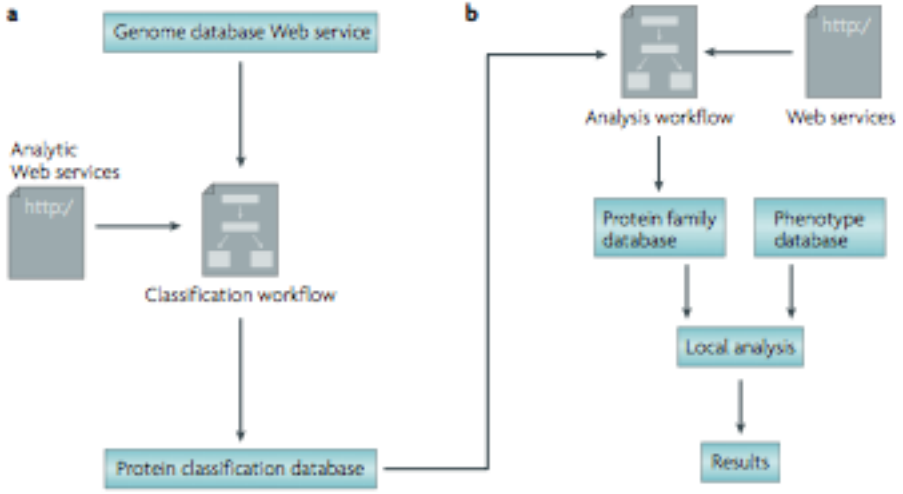


Figure 2 | **Analysis of the *Bacillus subtilis* secretome as an *in silico* experiment. a** | Classification of *B. subtilis* proteins. Data from a genome database that is exposed as a Web service are read into a workflow that incorporates various distributed Web services that implement analytical tools. The output is stored in a local database. **b** | Analysis of protein families. Data from the protein classification database are analysed using a different set of analytical Web services and the results are combined with phenotypic information for analysis and visualization on a local machine.

myGrid

omii-uk
www.omii.ac.uk

www.mygrid.org.uk

Taverna

utopia

Design, editing and execution of workflows

Tools for analysing protein sequence and structure

myexperiment

BioCatalogue beta
"The Life Science Web Service Registry"

Sharing of workflows and related data

SysMO SEK
Systems Biology of Microorganisms

Finding, sharing and exchanging data, models and processes in Systems Biology

RightField

Sharing the meaning of your data

Cataloging and annotation of services

MethodBox

Browse datasets and share knowledge.

What are we doing at LAFIECO that requires eScience ?

- Trying to understand how plants respond to the Global Climate Change
- Trying to produce more renewable fuels in a sustainable way

Laboratório de Fisiologia Ecológica de Plantas (LAFIECO)

Founded: Setember/2007
Team Leader: Marcos S. Buckeridge
Depto. Botânica, IB/USP.

Funding:

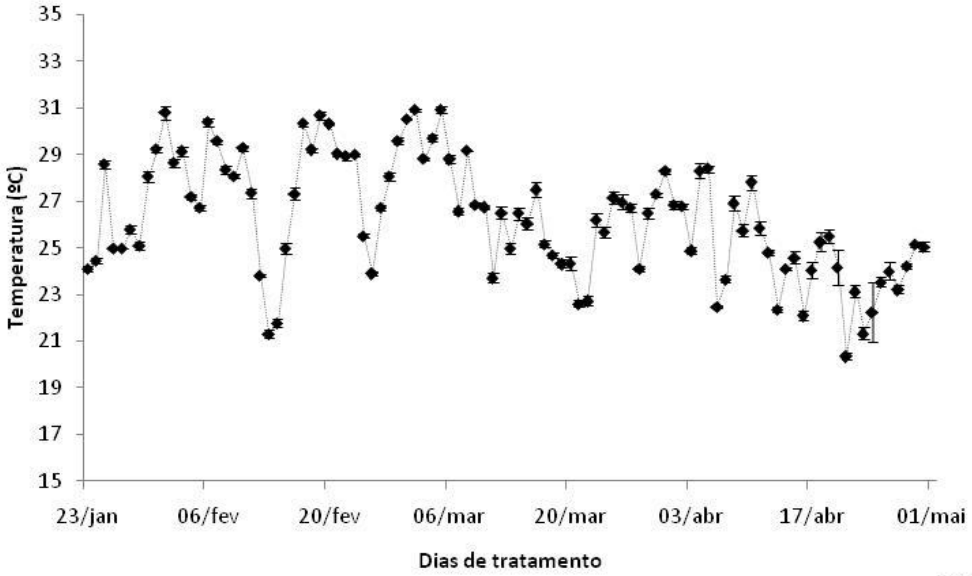
- **Infrastructure -** Eletronorte/Pará, Ministério de Ciência e Tecnologia (MCT)

- **Equipment –** FAPESP, CNPq

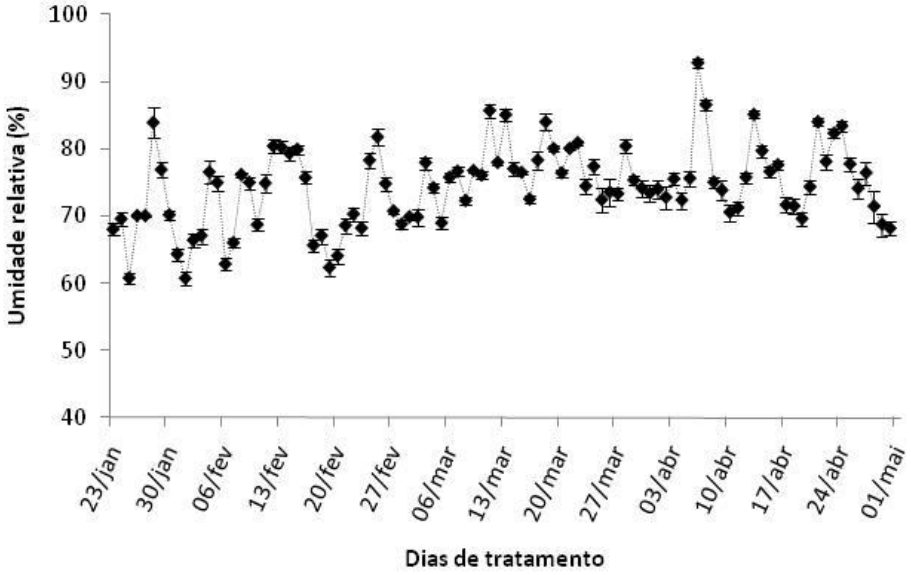
RICS (REMOTE INTEGRATED CONTROL SYSTEM)
MICROCLIMATIC DATA

Lafieco
Departamento de Botânica
IB-USP
Laboratório de Fisiologia Ecológica de Plantas

DADOS MICROCLIMÁTICOS



Experiment with *S. reticulata*
January to May 2009



Laboratório de Fisiologia Ecológica de Plantas

MISSION

Study plant species in the context of the Global Climatic Changes

Parameters

CO_2 H_20 T°C

(eco) Physiology | **Biochemistry** | **Gene expression**

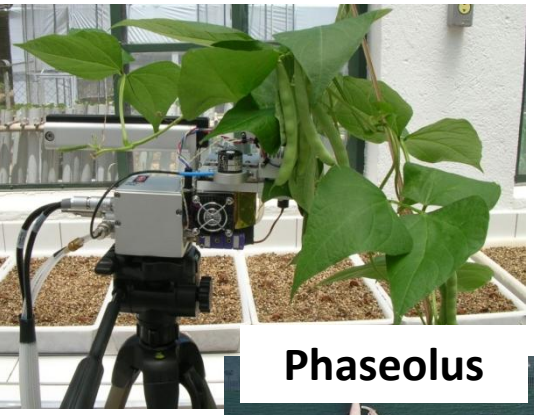
Native Species

Crop Species

Contribute for forecasting models of behavior of plant species withing their biomes

Understand and model plant behaviour so that we can contribute to predictive models of crop behavior under the Climate changes.

Some species used in research projects developed by the **LAFIECO** team



Phaseolus

sugarcane

Matapasto

soybean

Senna alata

Assai palm

Marcos Buckeridge – LAFIECO-USP

Matapasto (*Senna reticulata)*, an Amazonian legume tree, growing under elevated CO_2

Experimental procedure

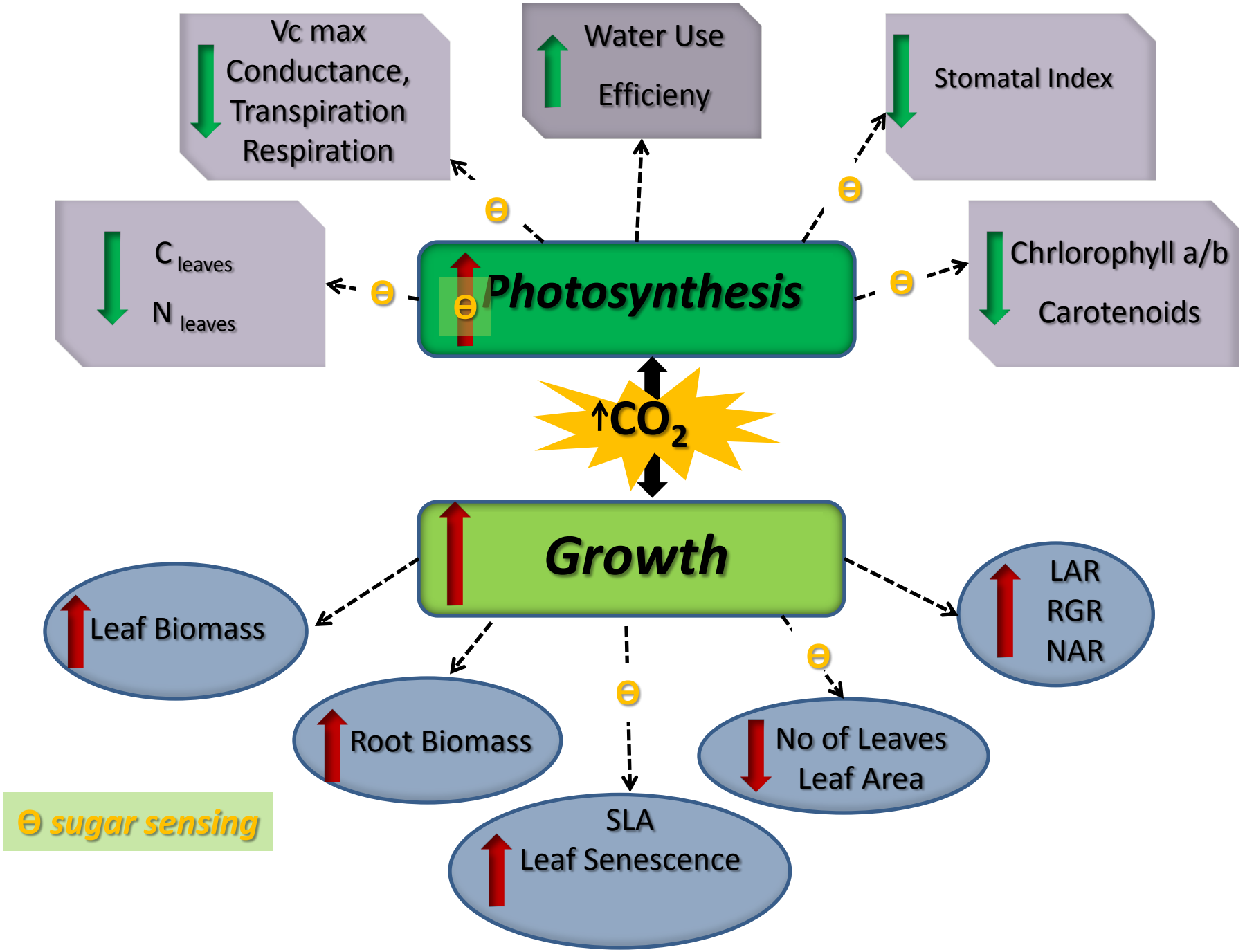
Seeds collected in Belém/PA

Mechanical scarification

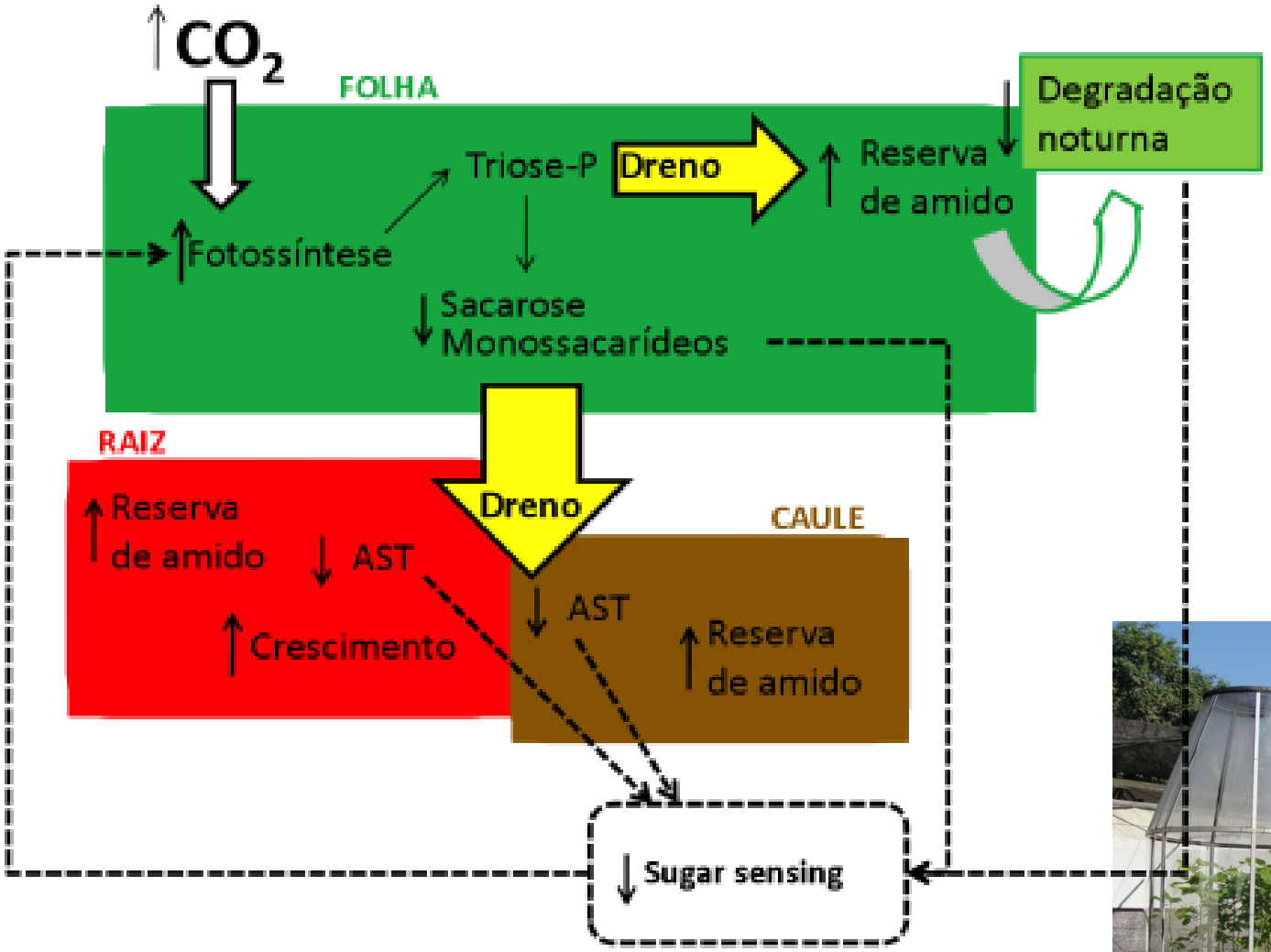
Germination at 28°C

After 10 days, plants are transferred to 10L pots and will then continue to grow in elevated CO2

Harvesting times

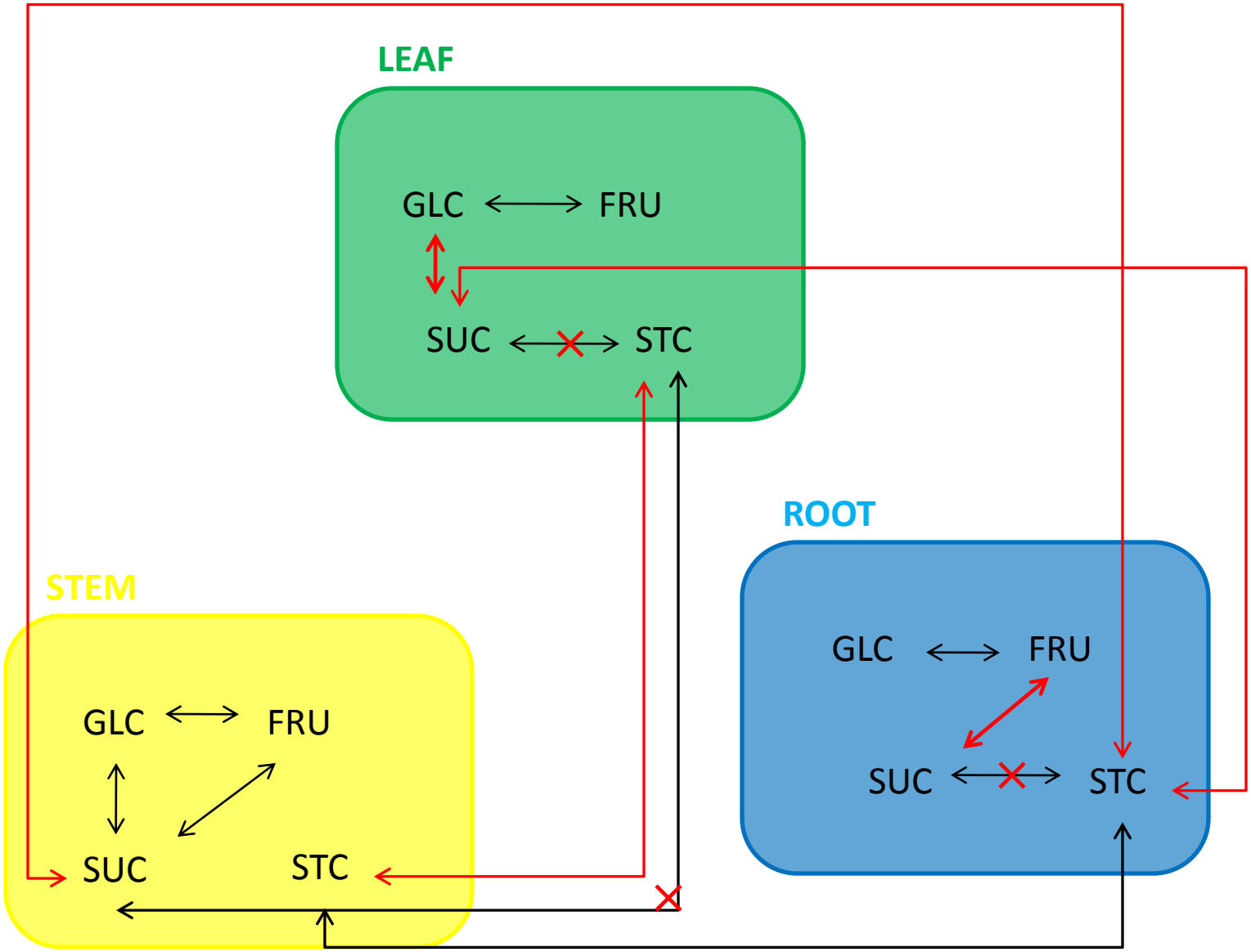

Vc max
Conductance,
Transpiration
Respiration
Water Use
Efficieny
Stomatal Index
C leaves
N leaves
Photosynthesis
Chrlorophyll a/b
Carotenoids
↑CO2
Growth
Leaf Biomass
LAR
RGR
NAR
Root Biomass
No of Leaves
Leaf Area
SLA
Leaf Senescence
θ sugar sensing

Metabolismo de carboidratos na leguminosa amazônica *Senna reticulata* em alto CO_2



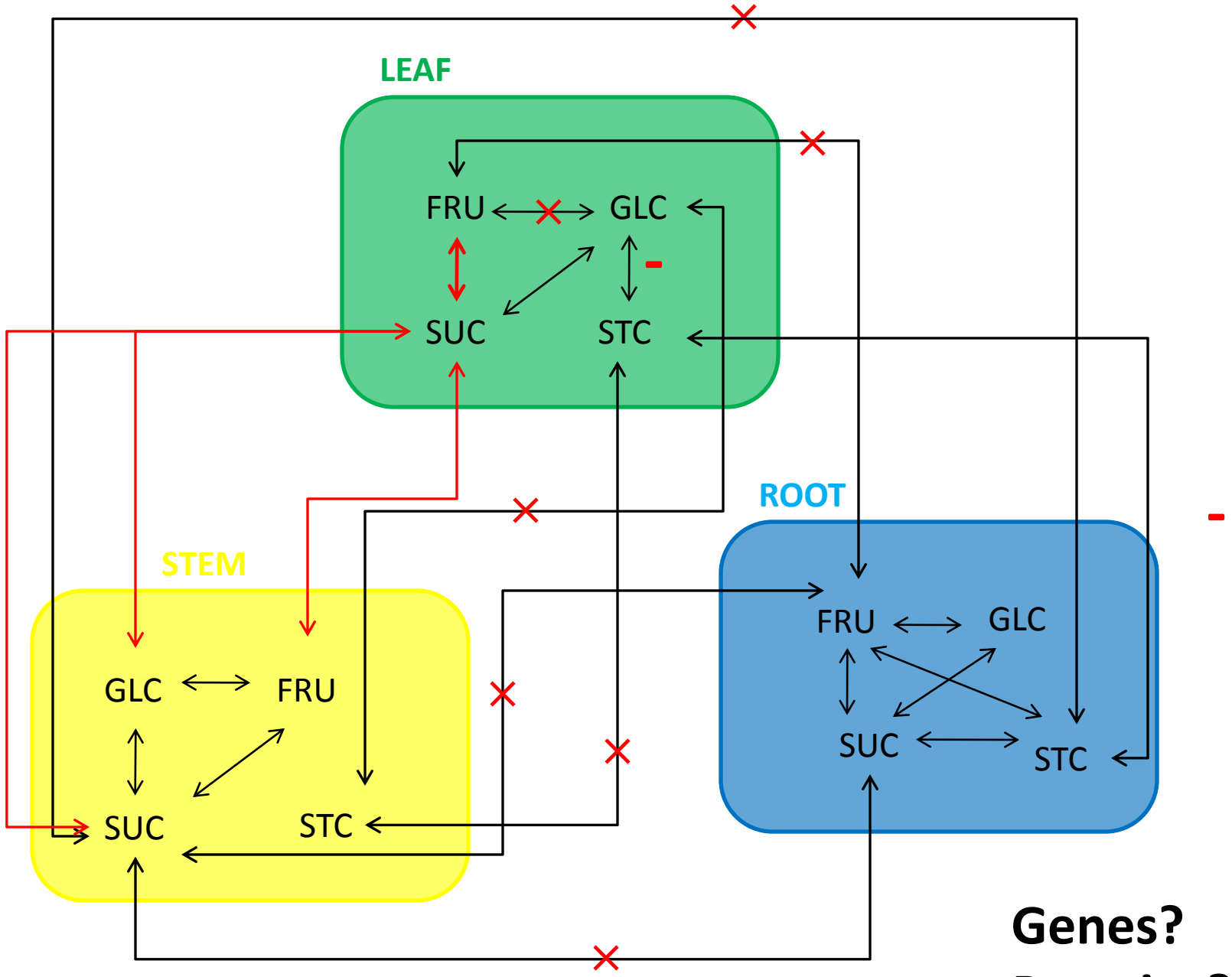
Senna reticulata crescendo em Câmaras de Topo Aberto no **LAFIECO**

Aluna: Bruna C. Arenque, Orientador: Marcos Buckeridge
Trabalho: **Metabolismo de carboidratos da espécie amazônica *Senna reticulata* sob cultivo em alto CO_2 *Senna reticulata***



Genes?
Proteins?

During Growth



During 24 hours

ECOLOGICAL SUCCESSION

Sesbania virgata

19 Kg per Ton
(70 Kg of CO2 per ton)

Schyzolobium parahyba

137 Kg per Ton
(487 Kg of CO2 per ton)

Piptadenia gonoacantha

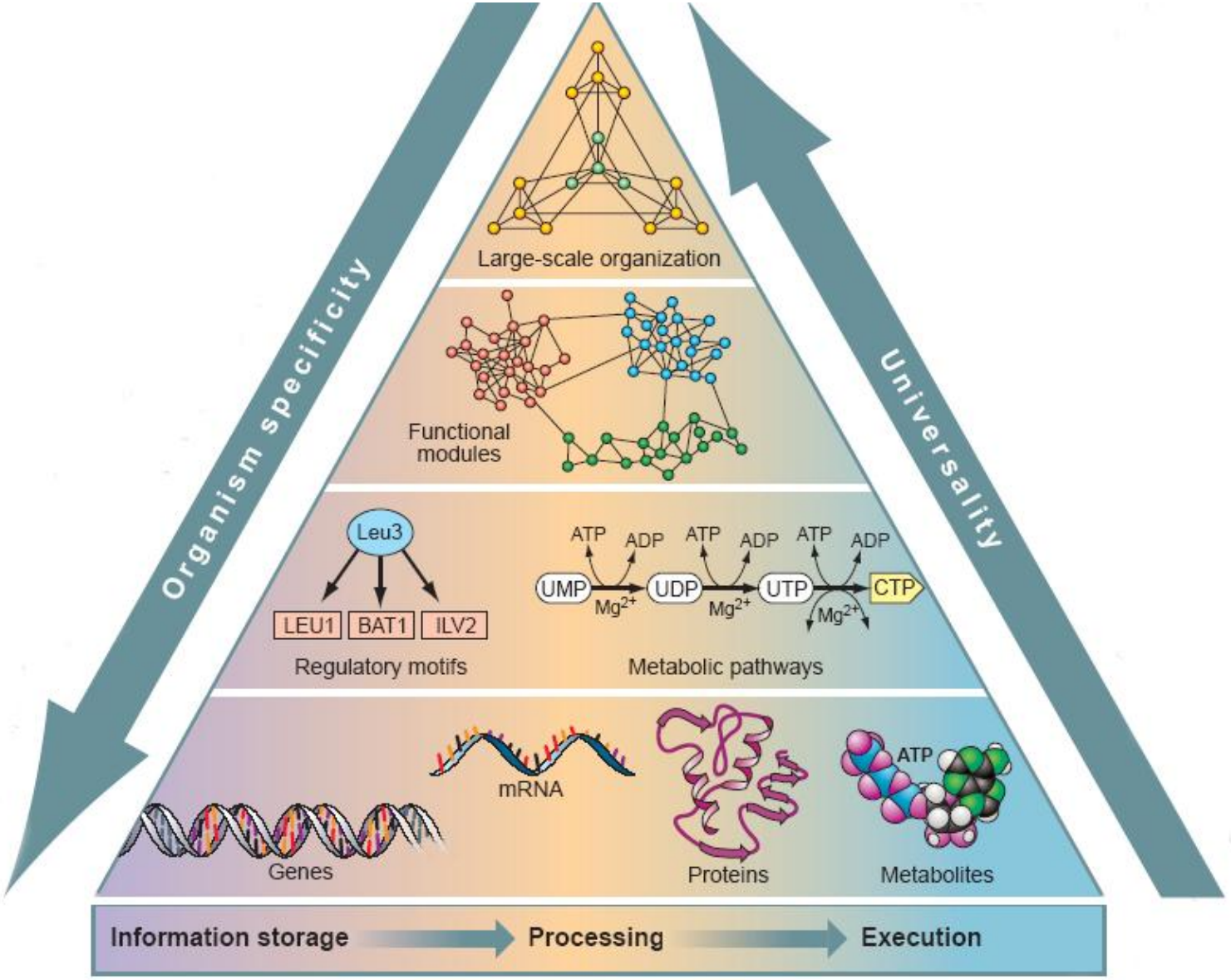
23 Kg per Ton
(84 Kg of CO2 per ton)

Dalbergia nigra

14 Kg per Ton
(51 Kg of CO2 per ton)

Hymenaea courbaril

In prep



Universality Pyramid (Oltvai, Z.N. & Barabási, A.L. 2002, Science 298: 763)

Can we use sensors to automatically mearure phenology, growth and metabolism?

- *We could test such sensors in the laboratory and when they work we could start using it in the forest*

- *If this works, (eco)physiological data could be included in **workflows** and processed to find conections with metabolic and genetic layers of the system*

- *We are discussing the development of a **sensor of sugar** that could send a signal in highthroughput so that we could analyse metabolic data under the light of the environmental data.*

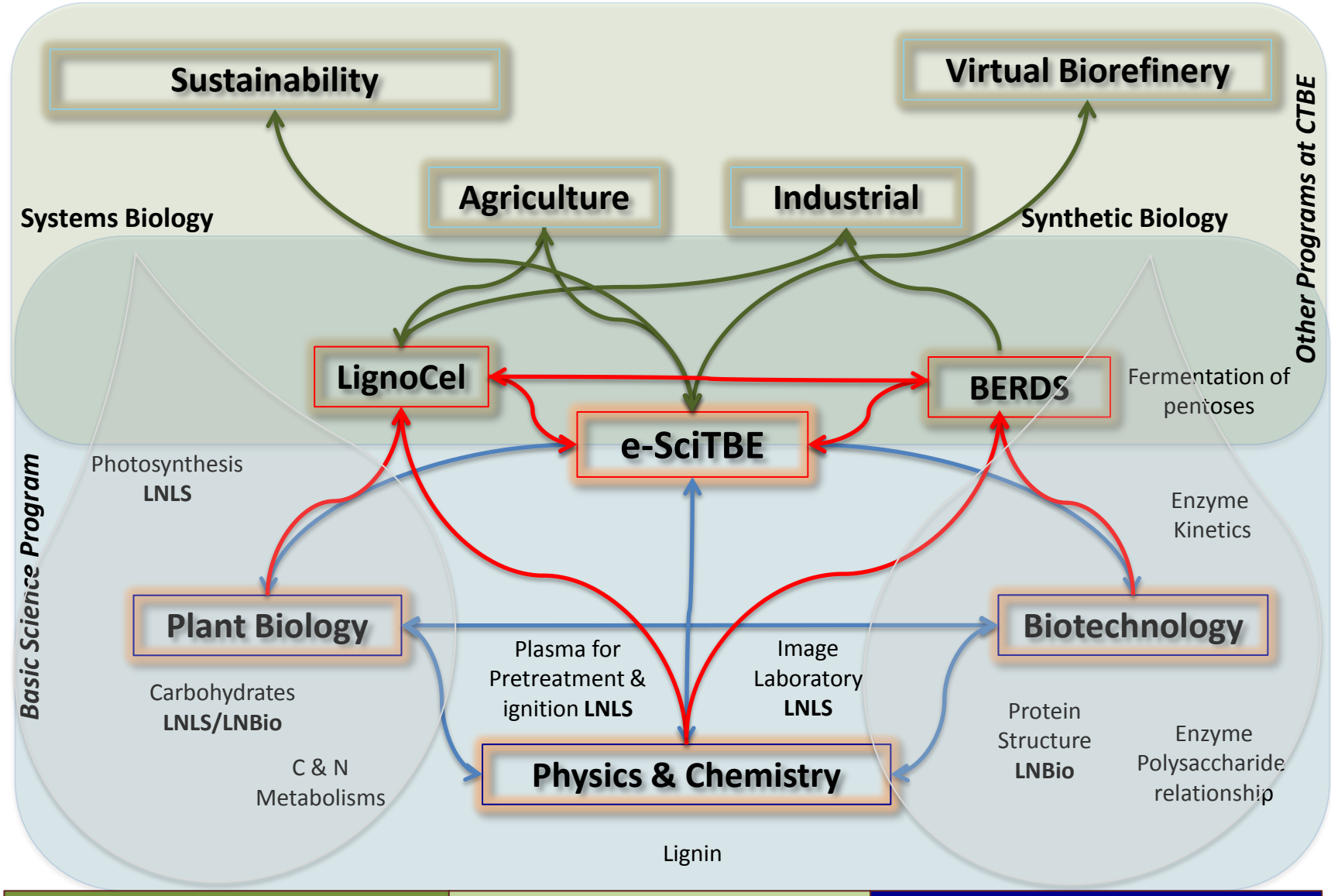
THANK YOU

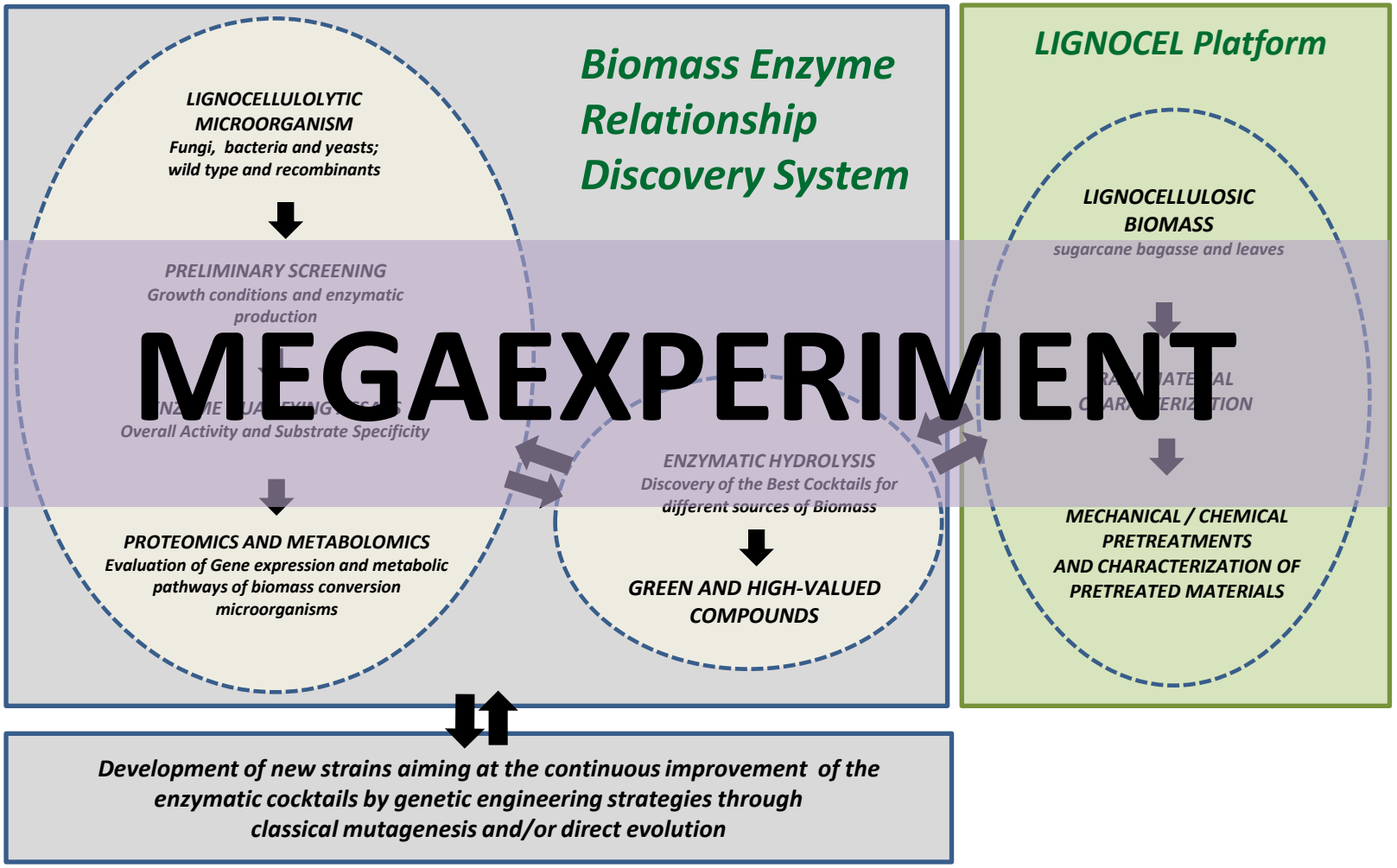
msbuck@usp.br

The real-world research problems that scientists address rarely arise within orderly disciplinary categories, and neither do their solutions (Carole L. Palmer, 2001)

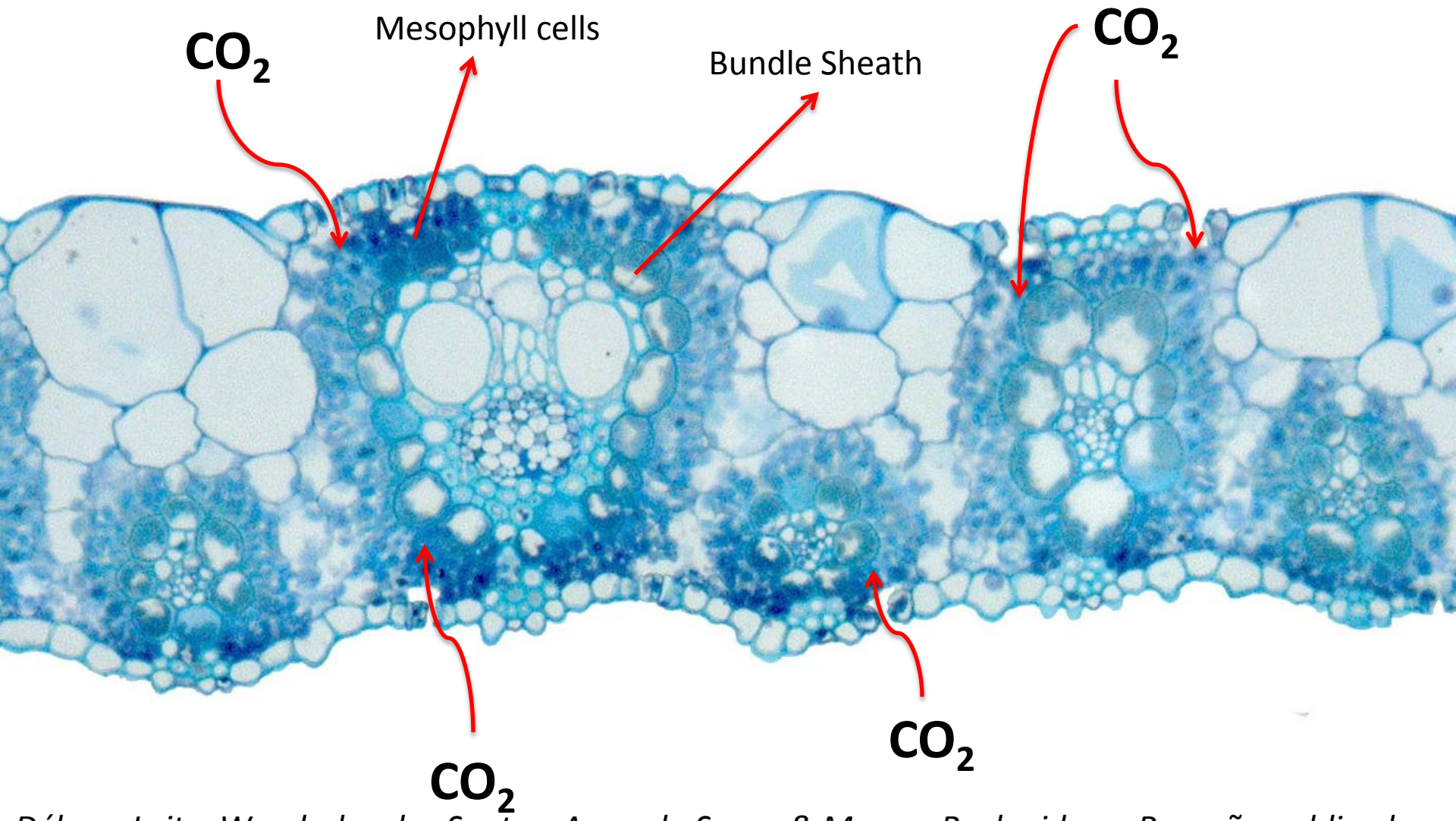
Brazilian Bioethanol Science and Technology Center

CTBE - *Campinas*

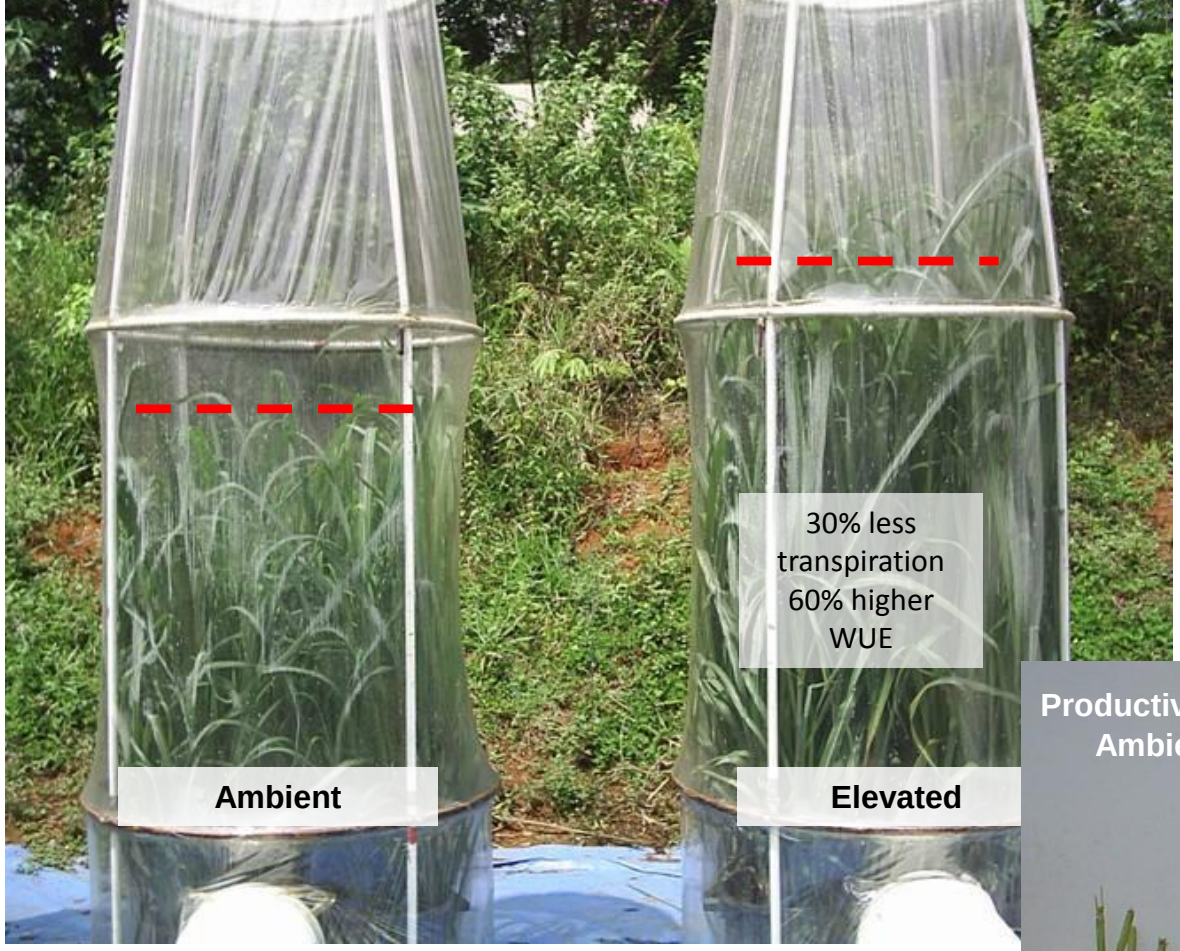

Sustainability
Virtual Biorefinery
Agriculture
Industrial
Systems Biology
Synthetic Biology
Other Programs at CTBE
LignoCel
BERDS
Fermentation of pentoses
e-SciTBE
Basic Science Program
Photosynthesis
LNLS
Enzyme Kinetics
Plant Biology
Biotechnology
Plasma for Pretreatment & ignition LNLS
Image Laboratory LNLS
Carbohydrates
LNLS/LNBio
Protein Structure
LNBio
Enzyme Polysaccharide relationship
C & N Metabolisms
Physics & Chemistry
Lignin
Plant
Cell Wall
Microorganisms


Biomass Enzyme Relationship Discovery System
LIGNOCELLULOLYTIC MICROORGANISM
Fungi, bacteria and yeasts; wild type and recombinants
PRELIMINARY SCREENING
Growth conditions and enzymatic production
Overall Activity and Substrate Specificity
PROTEOMICS AND METABOLOMICS
Evaluation of Gene expression and metabolic pathways of biomass conversion microorganisms
ENZYMATIC HYDROLYSIS
Discovery of the Best Cocktails for different sources of Biomass
GREEN AND HIGH-VALUED COMPOUNDS
LIGNOCEL Platform
LIGNOCELLULOSIC BIOMASS
sugarcane bagasse and leaves
MECHANICAL / CHEMICAL PRETREATMENTS AND CHARACTERIZATION OF PRETREATED MATERIALS
MEGAEXPERIMENT
Development of new strains aiming at the continuous improvement of the enzymatic cocktails by genetic engineering strategies through classical mutagenesis and/or direct evolution

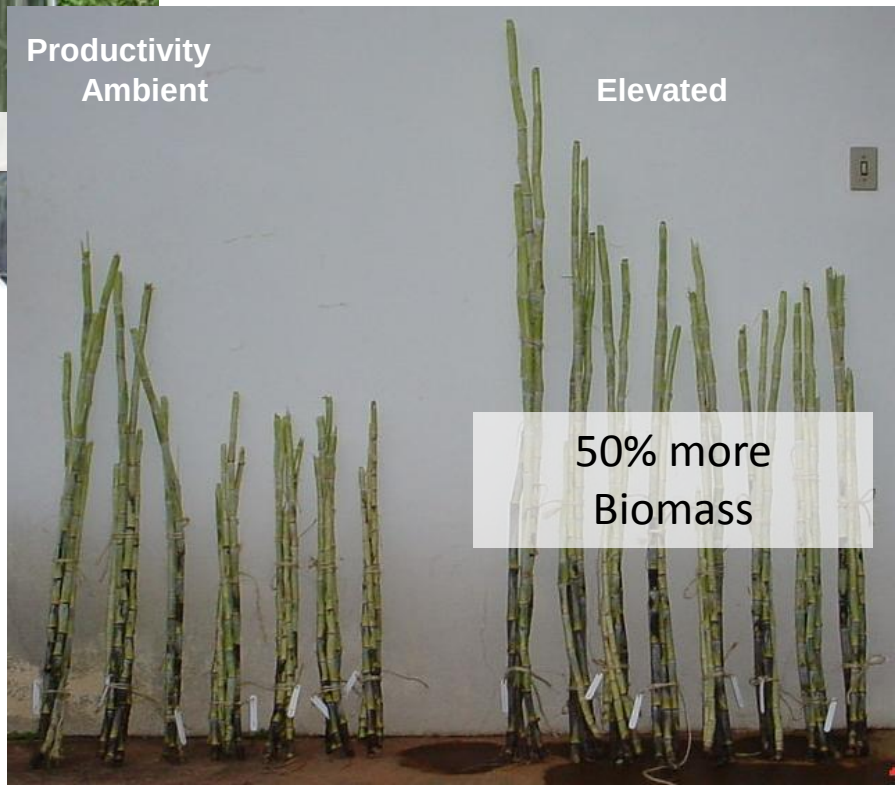
Sugarcane leaves performe C_4 photosynthesis



Débora Leite, Wanderley dos Santos, Amanda Souza & Marcos Buckeridge – Res. não publicados



De Souza et al. 2008
Plant Cell & Environment,
Volume 31, pg 1116



2005

1 year in elevated CO_2