



UNIVERSITY
OF MANITOBA

Asia Pacific Bioinformatics Conference
Melbourne, Australia, January 17 -

BioLegato: A Programmable, Object-Oriented Graphic User Interface

Dr. Brian Fristensky
Department of Plant Science
University of Manitoba
Winnipeg, Canada



GenomeCanada



Microbial Genomics for Biofuels and
Co-Products from Biorefining Processes



GenomePrairie

Bioinformatics Team



UNIVERSITY
OF MANITOBA

Intellectual critical mass

Continuity of expertise

Synergies

Faculty

Brian Fristensky

Bioinformaticists

**Graham Alvare
Justin Zhang**

PhD Student

Abiel Roche

Research Associate

Natalie Bjorklund
Biostatistician TBA

Undergraduate Students

Dale Hamel

BioLegato

Why a Duck?

The Object-Oriented World View

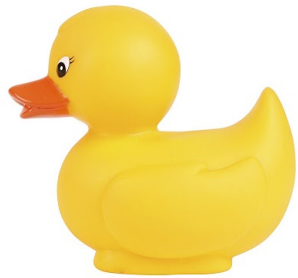


Common wisdom: "If it looks like a duck, and waddles like a duck, and quacks like a duck, then it's probably a duck."

OO: If you want a duck, then design it to look like a duck, and waddle like a duck, and quack like a duck.

BioLegato - The OO World View

1. A class is a formula for creating an object
2. Everything is an Object
3. Objects have Data and Methods
4. You can make new classes by reusing and extending an existing class



Class: DUCK
Data: Wing, Feathers, **Bill**
Methods: Waddle, **Quack, Fly, Swim**

Class: BIRD
Data: Wing, Feathers
Methods: Waddle

Reuse and extend BIRD to
to create new BIRD classes

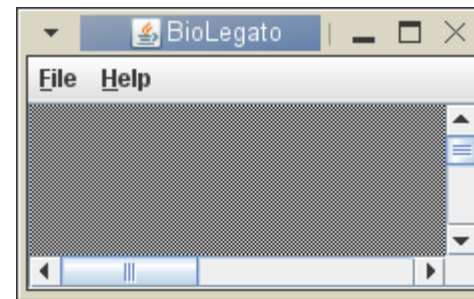


Class: PENGUIN
Data: Wing, Feathers, **Bill**
Methods: Waddle, **Honk, Swim**

The BioLegato Concept

- **Unique concept in software design.***
- **Almost no functionality is written into the program itself.**
- **Menus, and a choice of canvas to display data, are read when the program launches.**

Empty BioLegato interface: null menus, empty canvas

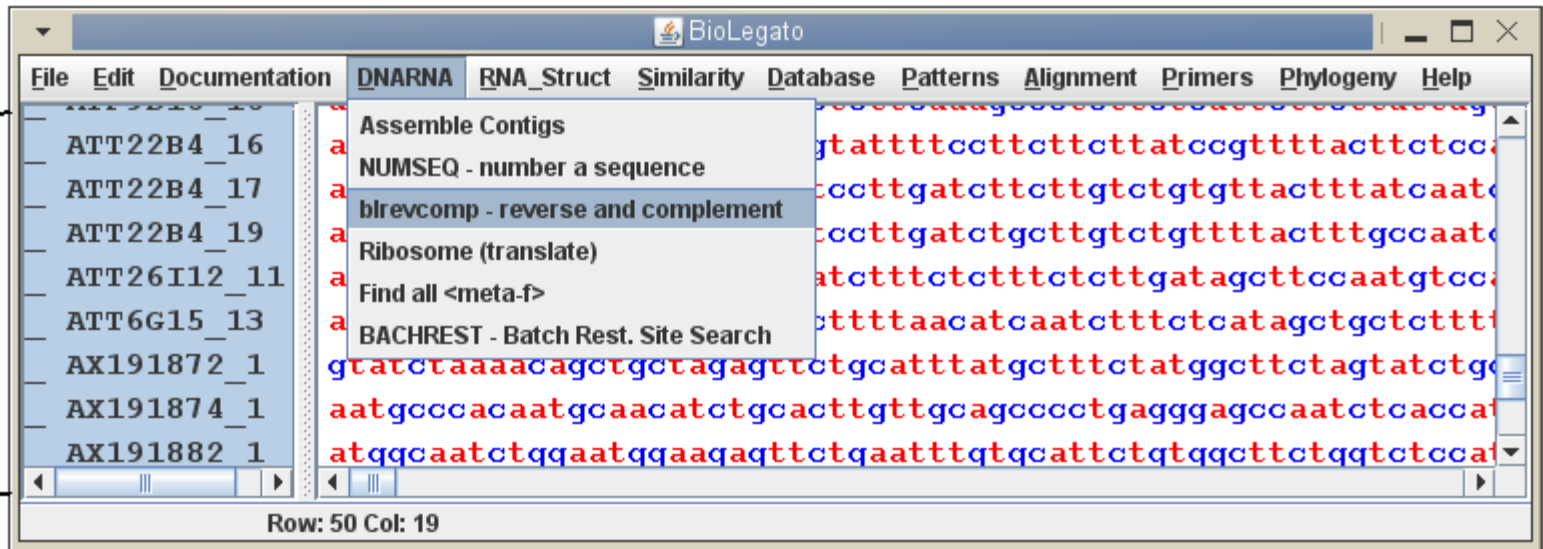


* Inspired by GDE/SeqLab of Steven Smith

BioLegato - Objects have data and methods

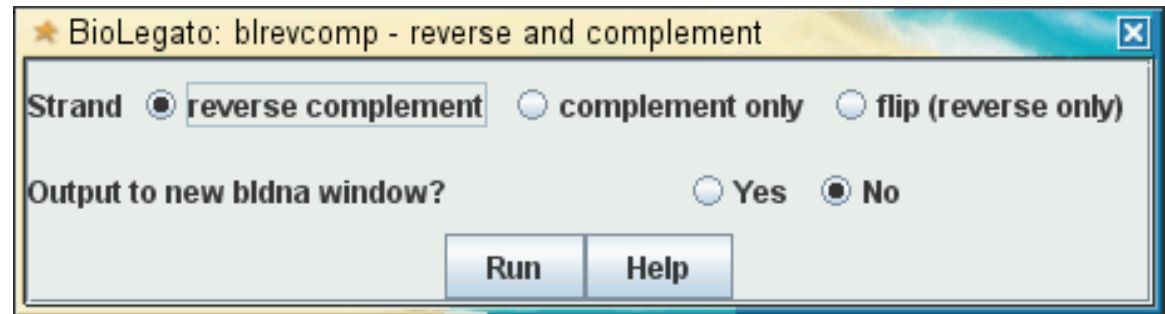
Menus
(Methods)

Canvas
(Data)



PCD - GUI scripting language

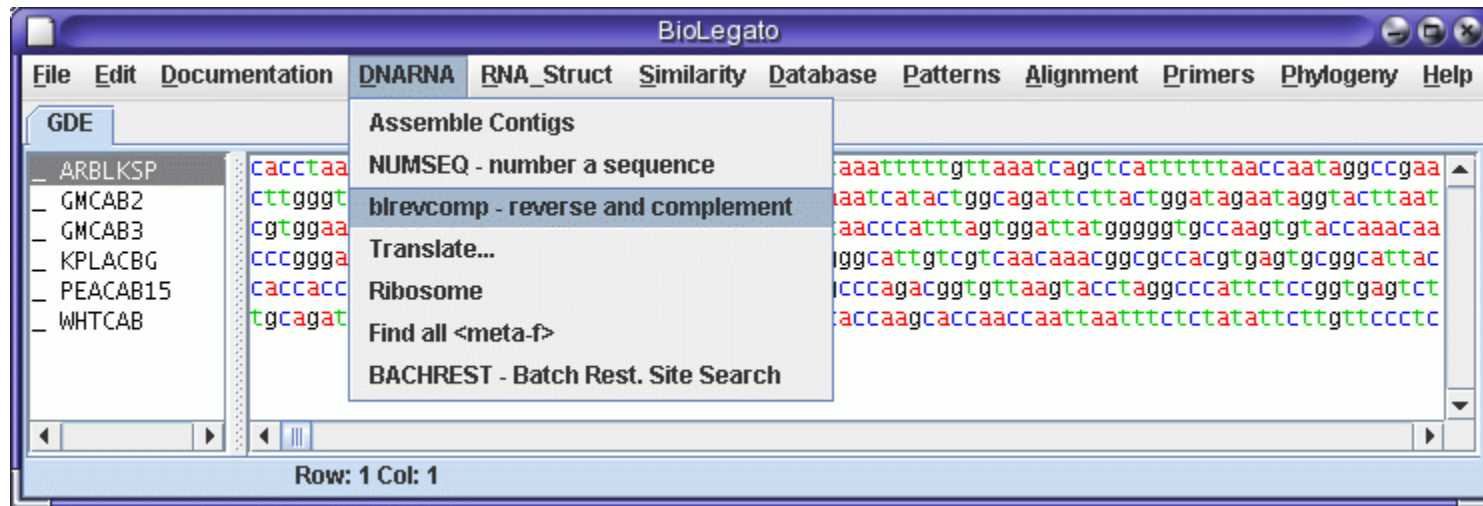
```
name          "blrevcomp - reverse and complement"
var "in1"
  type        tempfile
  direction   in
  format      flat
var "strand"
  type        chooser
  label       "Strand"
  default     0
  choices
    "reverse complement" "-r"
    "complement only"   "-c"
    "flip (reverse only)" "-f"
var "out1"
  type        tempfile
  direction   out
  format      flat
var "gdeoutput"
  type        chooser
  label       "Output to new bldna window?"
  default     1
  choices
    "Yes"      "(echo '' > %out1%; bldna %in1%.blrevcomp; $RM_CMD -f %in1%.blrevcomp) &"
    "No"       "mv %in1%.blrevcomp %out1%"
panel
  var "Run"
    type        button
    label       "Run"
    shell       "blrevcomp %STRAND% %in1% %in1%.blrevcomp; $RM_CMD %in1%; %GDEOUTPUT%"
    close      true
  var "Help"
    type        button
    label       "Help"
    shell       "$BIRCH/script/gde_help_viewer.csh null/$BIRCH/doc/BioLegato/blrevcomp.html"
    close      false
```



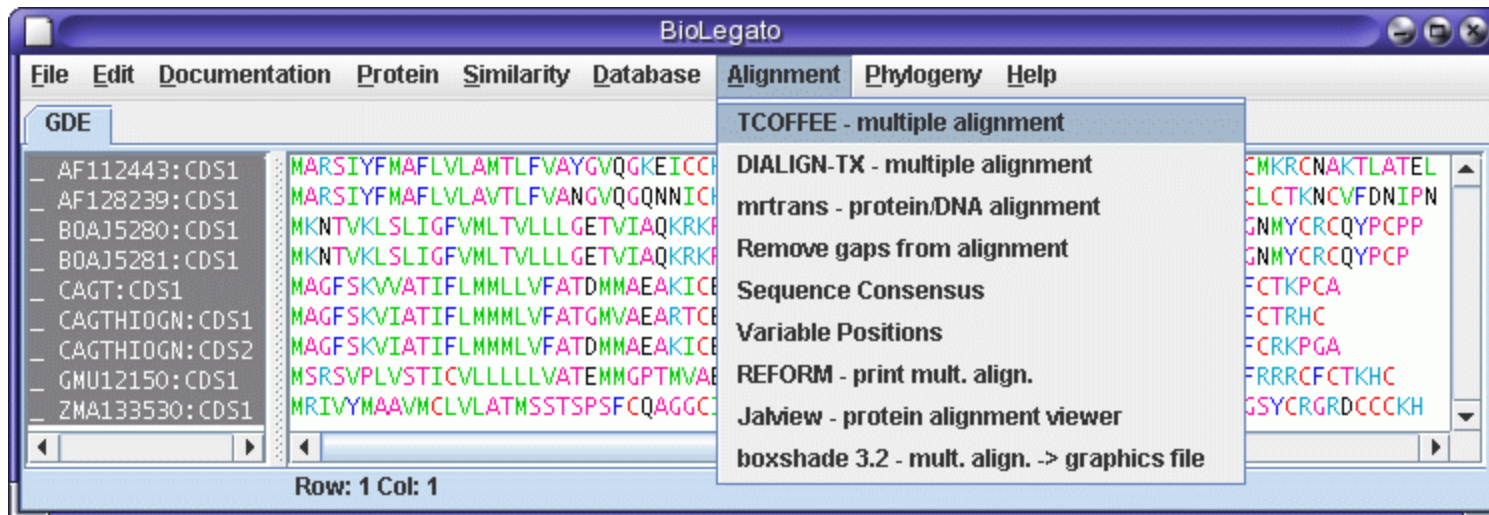
BioLegato - menus + canvas = new GUI

A GUI that will run just about any program for any kind of data can easily be made by writing new menus and choosing a canvas.

- **bldna -**
- **menus for DNA programs**
- **multiple alignment canvas**



- **blprotein -**
- **menus for protein programs**
- **multiple alignment canvas**



BioLegato - rapid deployment of new GUIs

GENOGRAPHER

birch - application launcher

DRAWTREE

TCOFFEE

ATV

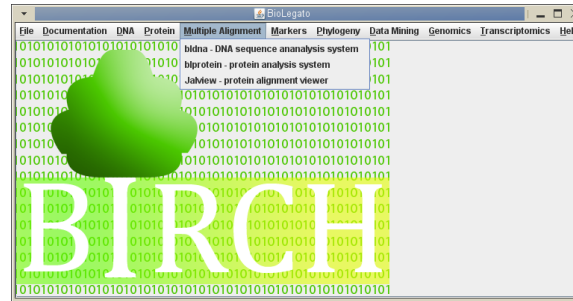
SEQUIN

BLREVCOMP

NUMSEQ

RIBOSOME

TMEV



WEIGHBOR

MRBAYES

DIALIGN

CN3D

GLSEARCH

FASTXY

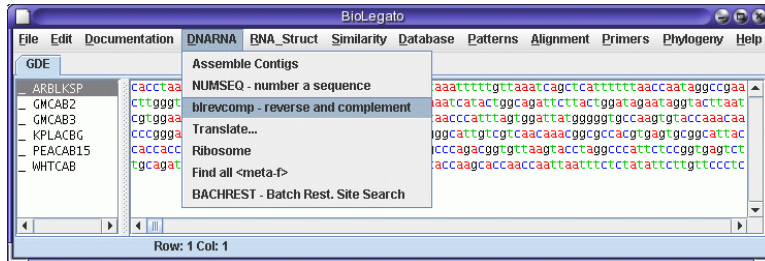
ARTEMIS

BACHREST

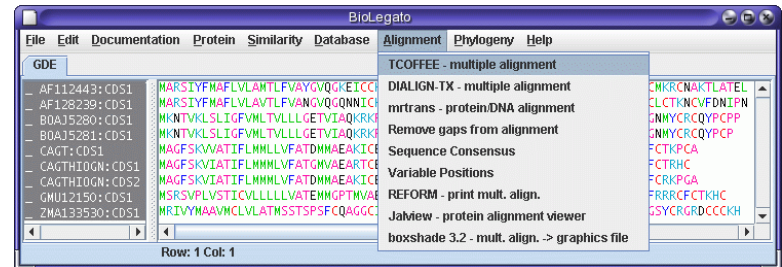
bldna - DNA sequence analysis

DIGEST

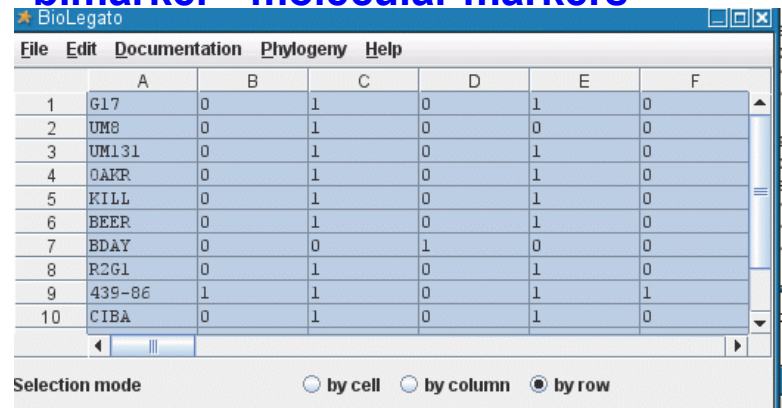
blprotein - protein sequence analysis



PRIMERS

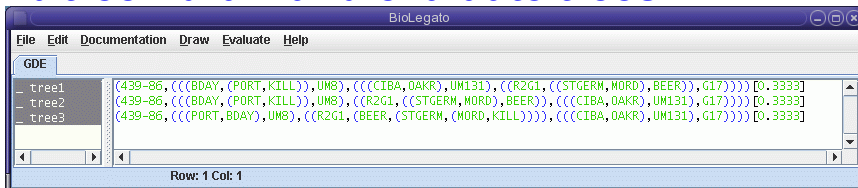


blmarker - molecular markers



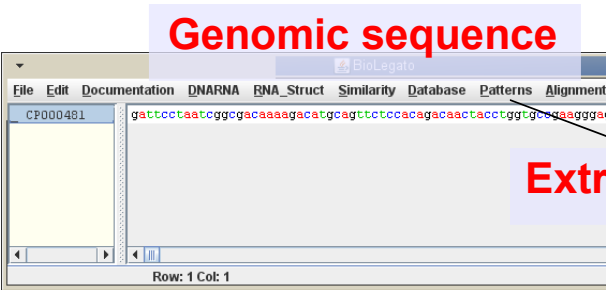
SHUFFLE

bltree - draw and evaluate trees

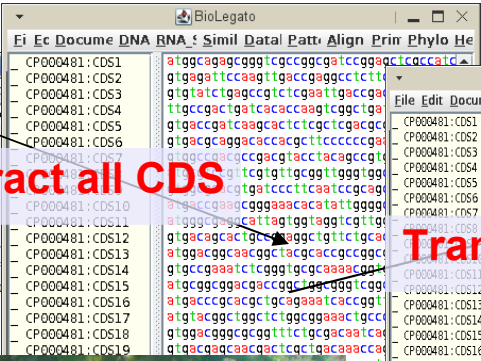


BioLegato - Point and click data pipelining

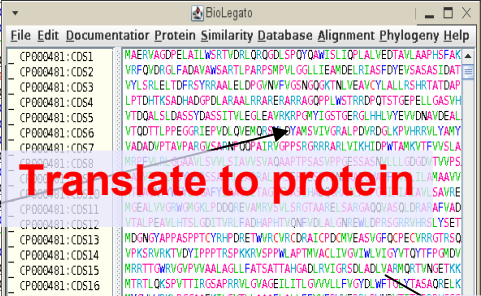
Genomic sequence



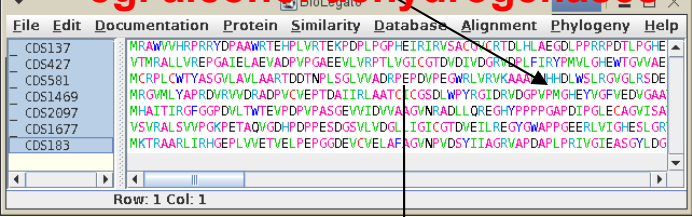
Extract all CDS



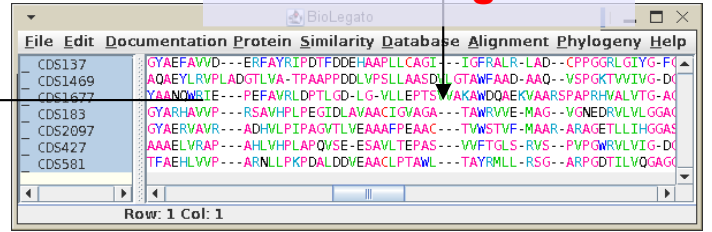
Translate to protein



**Select proteins of interest
eg. alcohol dehydrogenases**



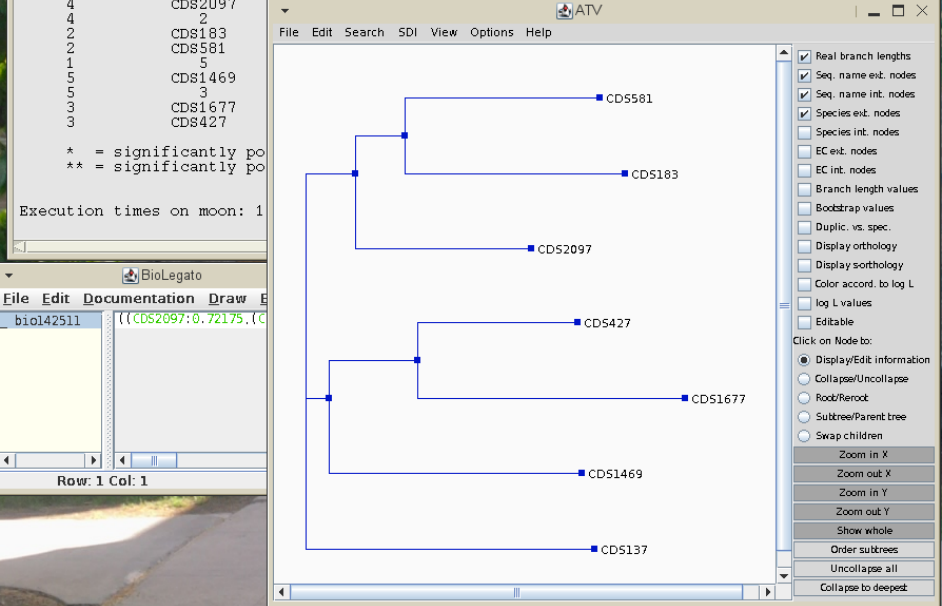
TCOFFEE alignment



**PROML
maximum likelihood phylogeny**

Ln Likelihood = -6006.54294

Between	And	Length	Approx. Confidence	Limits
1	CDS137	1.18298	(0.93840, 1.42769)	**
4	CDS2097	0.20468	(0.06749, 0.34188)	**



Bioinformatics - BioLegato Database Client

BioLegato lets us quickly build a database client tailored for the project.
Most databases just let you view and query the data.
BioLegato lets you select data, and run programs on it.

Custom Database Client for the MGCB2 project

- bldb -**
- menus for project-related programs
 - table canvas

		species_name	strain	Sy
1		cellulolyticus	11B	NUL
2		cereus	03BB102	NUL
3		subtilis	JH642	NUL
4		saccharolyticus	DSM 8903	NUL
5		NULL	T88	NUL
6		amalonaticus	Y19	NUL
7		acetobutylicum	ATCC 824	NUL
8		beijerinckii	NCIMB 8052	NUL
9		butyricum	5521	NUL
10		cellulolyticum	H10	NUL
11	45	Clostridium kluyveri	DSM 555	NUL
12	50	Clostridium perfringens	ATCC 13124	NUL
13	51	Clostridium phytofermentans	TSDg	NUL

Example: Compare pathways between two species

In the database client, select two species to compare

Choose pathway and click RUN

BioLegato calls the KEGG [3] API to create a custom KEGG map, highlighting enzymes present only in either species, or both

Click on an enzyme to display KEGG info

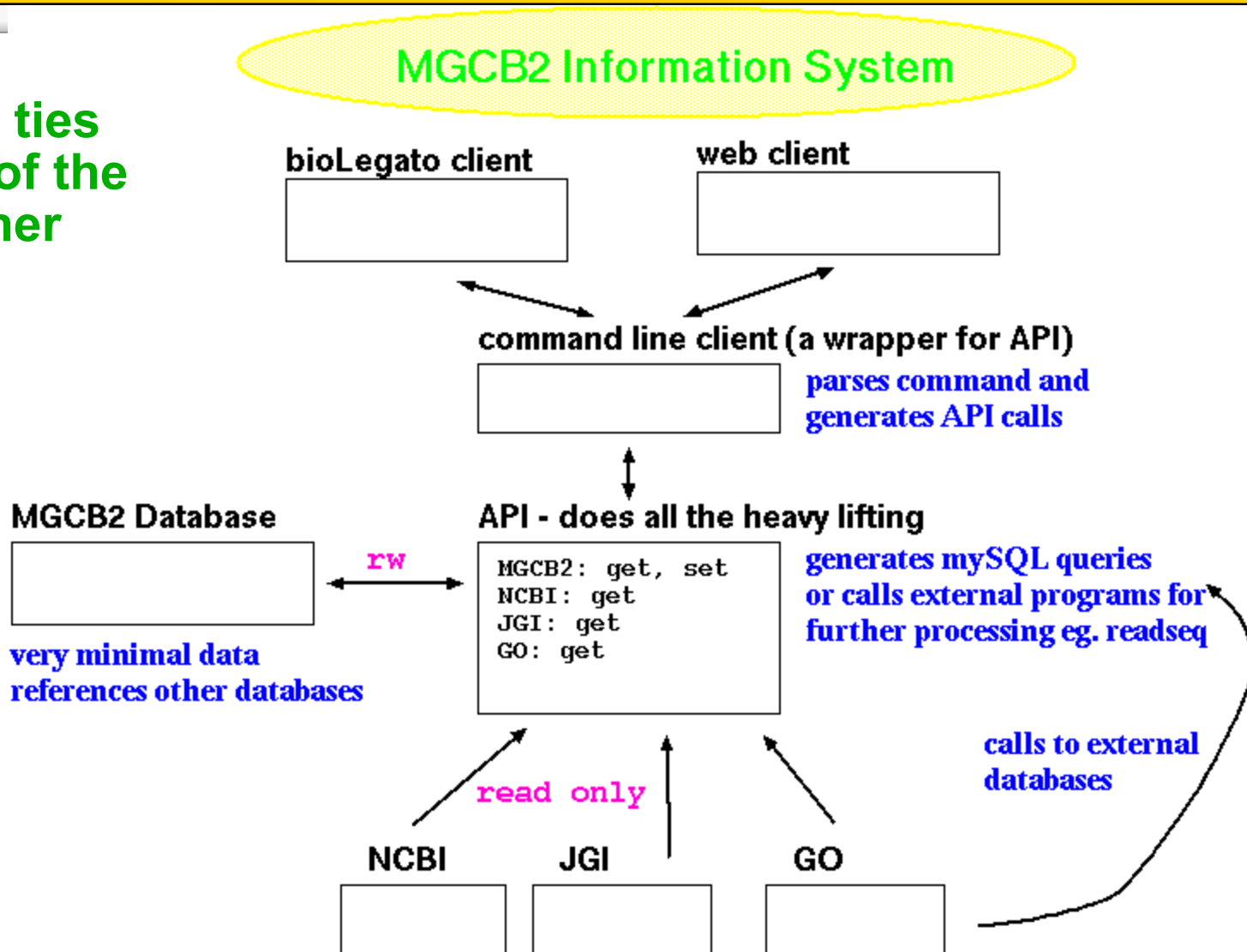
species_name	strain
cellulolyticus	11B
cereus	03BB102
subtilis	JH642
saccharolyticus	DSM 8903
NULL	T88
beijeirincii	HCIMB 8052
butyricum	5521
cellulolyticum	H10
Clostridium kluyveri	DSM 555
Clostridium perfringens	ATCC 13124
Clostridium butyfermentans	ISDa

KEGG URL: http://www.genome.jp/kegg-bin/mark_pathway_www?132588374218554map00010.args
BLUE = Clostridium kluyveri DSM 555
GREEN = Clostridium perfringens ATCC 13124
RED = both

Entry	K00163	KO
Name	aceE	
Definition	pyruvate dehydrogenase E1 component [EC:1.2.4.4]	
Pathway	ko00010 Glycolysis / Gluconeogenesis ko00020 Citrate cycle (TCA cycle) ko00290 Valine, leucine and isoleucine biosynthesis ko00620 Pyruvate metabolism ko00650 Butanoate metabolism	
Module	M00307 Pyruvate oxidation, pyruvate => acetyl	
Class	Metabolism; Carbohydrate Metabolism; Glycolysis [PATH:ko00010] Metabolism; Carbohydrate Metabolism; Citrate cycle [PATH:ko00020] Metabolism; Carbohydrate Metabolism; Pyruvate metabolism [PATH:ko00620] Metabolism; Carbohydrate Metabolism; Butanoate metabolism [PATH:ko00650] Metabolism; Amino Acid Metabolism; Valine, leucine and isoleucine biosynthesis	

Bioinformatics - Database

The database ties components of the project together



BioLegato - To Do List

- **Inheritance**
- **Relations between classes**
- **More canvases**
- **More widgets eg. menu bars, buttons, side-panels**

BioLegato/BIRCH

BioLegato is distributed with the BIRCH bioinformatics system

<http://home.cc.umanitoba.ca/~frist>

- **Solaris**
- **Linux**
- **Mac OSX**
- **Windows (2nd Quarter 2012)**