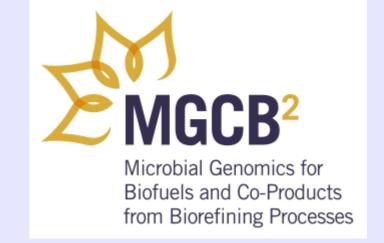
Microbial Genomics - Annual General Meeting Winnipeg, Manitoba, November 7 - 8, 2011



BioLegato: A Programmable, Object-Oriented Graphic User Interface



GenomePrairie

S

Graham Alvare¹, Abiel Roche², Brian Fristensky^{1*} Departments of Plant Science¹ and Computer Science², University of Manitoba, Winnipeg, Canada R2T 2N2

*Corresponding author frist@cc.umanitoba.ca

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.

The Object-Oriented 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: BIRD

ABSTRACT

Object-oriented (OO) concepts are built into most modern programming languages and databases. BioLegato brings OO concepts to the laboratory biologist. BioLegato is best thought of as a GUI that launches other programs. As in most GUIs, each bioLegato window consists of a canvas for displaying and manipulating data, and a set of pull-down menus. In the OO paradigm, the canvas corresponds to the data, and the menus correspond to the methods associated with that data. Rather than being hard-coded, menus and canvas are read at runtime. To specify the contents of menus, we have created a small language, PCD. Canvases are implemented as Java plugins. Thus, a new GUI can be created by choosing a canvas appropriate to the data and writing menus in PCD. By analogy, bioLegato can be thought of as a generic Object class, which can be reused and extended.

Programs in the BIRCH bioinformatics system can be run through five bioLegato interfaces: bldna, for DNA sequences; blprotein for proteins; blmarker for molecular markers; bltree for phylogenetic trees, and birch, a launcher for programs in BIRCH. Virtually all methods in bioLegato call external programs, meaning that bioLegato can run either locally-installed programs or remote web services. Wherever possible, the output appears in a new bioLegato window, allowing *ad hoc* pipelining. For example, when bldna translates a DNA sequence, the protein output pops up in a blprotein window. The programmability of bioLegato speeds development of new GUIs for almost any sort of data, giving biologists objects that behave like the real-world entities they represent.

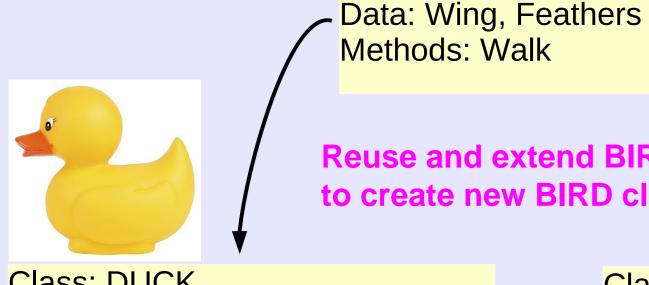
Point and click data pipelining: Run programs directly from the database

Most database clients only search and display data. The bioLegato database client lets you select data, and launch external programs, including other bioLegato programs.

	•			🛃 BioLegato	.	- - ×
	<u>F</u> ile <u>E</u> d	it <u>T</u> asks <u>A</u>	dd <u>H</u> elp			
		strain_id	genus	species_name	strain	Synor
	1	36	Acidothermus	cellulolyticus	118	NULL 🔺
	2	37	Bacillus	cereus	03BB102	NULL 😑
	3	98	Bacillus	subtilis	JH642	NULL
	4	55	Caldicellulosiruptor	saccharolyticus	DSM 8903	NULL
	5	75	Chloroflexus	NULL	T88	NULL
P	lec	t₁a sn	ecies in the	database	Y19	NULL
Ĭ	7	28	Clostridium	acetobutylicum	ATCC 824	NULL
	8	33	Clostridium	beijerinckii	NCIMB 8052	NULL
	9	34	Clostridium	butyricum	5521	NULL
	10	35	Clostridium	cellulolyticum	H10	NULL _
		4		· · ·		

Biole Genome is retrieved to bioLegato <u>File Edit D</u>ocumentation <u>D</u>NARNA <u>R</u>NA_Struct <u>S</u>imilarity





Reuse and extend BIRD to to create new BIRD classes

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

Class: PENGUIN Data: Wing, Feathers, Bill Methods: Walk, Honk, Swim

BioLegato Objects have Data and Methods

In BioLegato, the menus correspond to methods, and the canvas corresponds to the data. Both menus and canvas are read when the program is launched. Canvases are Java plugins, while menus are written in PCD, a new, compact language that describes the contents of menus, such as buttons, sliders and file choosers. Thus, new bioLegato classes can be created by reusing existing menus and canvases, or creating new menus and canvases.

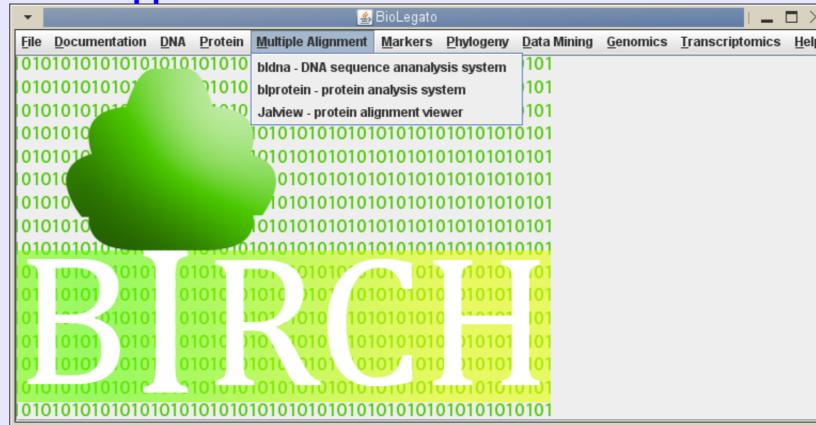
(Methods) Tile Edit Documentation DNARNA RNA_Struct Similarity Database Patterns	<u>A</u> lignment <u>P</u> rimers <u>P</u> hylogeny <u>H</u> elp
ATT22B4_17 ATT22B4_17 Image: Contrast of the second contrest of the second contrest of the second contrest of the second c	gcccctgagggagccaatctcacca

PCD - A new language for creating bioLegato "methods"

BIRCH Web site: http://home.cc.umanitoba.ca/~psgendb

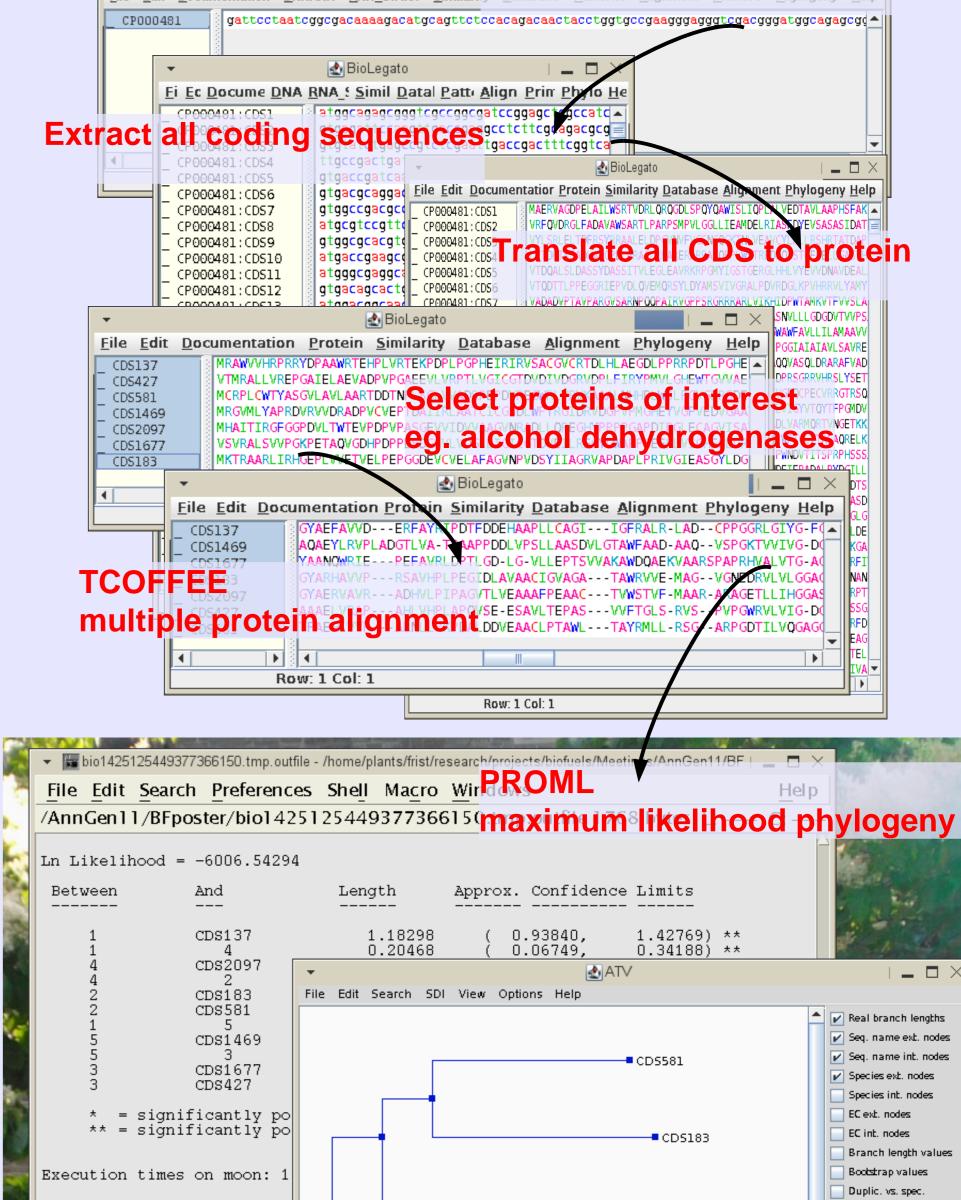
Biolegato reuses and extends canvases and menus to create new biological objects

birch - application launcher



bldna - DNA sequence analysis

					-			-					
						BioLega	ito						08
<u>F</u> ile	Edit	Docum	nentation	DNARNA	RNA_Struct	<u>S</u> imilarity	<u>D</u> atab	ase	Patterns	<u>A</u> lignment	Primers	Phylogeny	Help
GD)E			Assembl	e Contigs								
_ A	RBLKS	•	cacctaa	NUMSEQ	- number a se	equence			-	_		caataggeeg	
	MCAB2		cttgggt	NILCYCOT	np - reverse ar	nd complem	IGIIL			-		ataggtactta	
	MCAB3 PLACB	G	cccggga cccggga	Tranclate	e							ytgtaccaaad agtgcggcatt	
	EACAB:		caccacc	Ribosom	e			iccca	lg <mark>ac</mark> ggtgt1	taagtaccta	gg <mark>ccca</mark> tte	.tccggtgagt	tet
– W	HTCAB		tgcagat	Find all <	meta-f>			acca	agcaccaad	caattaatt	tetetatat	tettgtteed	tc
			8	DACUDE		t Cito Coor	ab						



All it takes to add a new program as a "method"...

🖈 BioLegato: blrevcomp - rev	erse and	compleme	nt	Þ
Strand	ent 🔾 c	omplement	only) flip (reverse only)
Output to new bldna window?		0	Yes	No
	Run	Help		

... is to write a menu in PCD. Thus, any program installed on your computer, or any program on the Web, can become a bioLegato method.

name "blrevcomp var "in1"	o - reverse and complement"
type tempfile	
direction in	Most of the DCD and a provifies "widerate" like
format flat	Most of the PCD code specifies "widgets", like
var "strand"	buttons, sliders, and file choosers, that set the
type chooser label "Strand"	
default 0	parameters for any program that bioLegato runs
choices	
"reverse complet	ment" "-r"
"complement onl	
"flip (reverse only	^{y)" "-f"} This line tells bioLegato,
var "out1" type tempfile	
direction out	which program to run, and
format flat	which parameters to pass
var "gdeoutput"	
type chooser	to the program, when you
	new bldna window?"
default 1 choices	click off the Run button.
	" > %out1%; bldna %in1%.blrevcomp; \$RM_CMD -f %in1%.blrevcomp)&"
	in1%.blrevcomp %out1%"
panel	
var "Run"	
type button label "Run"	
	omp %STRAND% %in1% %in1%.blrevcomp; \$RM CMD %in1%; %GDEOUTPUT%"
close true	
var "Help"	
type button	
label "Help"	
shell "\$BIRC close false	CH/script/gde_help_viewer.csh null/\$BIRCH/doc/bioLegato/blrevcomp.html"

BACHREST - Batch Rest. Site Search	-	
Row: 1 Col: 1		

blprotein - protein sequence analysis

1	BioLe	egato	999
<u>File Edit D</u> ocumentation <u>P</u> r	rotein <u>S</u> imilarity <u>D</u> atabase	Alignment Phylogeny Help	
GDE		TCOFFEE - multiple alignment	
_ AF112443:CDS1	FMAFLVLAMTLFVAYGVQGKEICC	DIALIGN-TX - multiple alignment	CMKRCNAKTLATEL 🔺
	FMAFLVLAVTLFVANGVQGQNNIC	mrtrans - protein/DNA alignment	
	LSLIGFVMLTVLLLGETVIAQKRKA LSLIGFVMLTVLLLGETVIAQKRKA	Remove gaps from alignment	GNMYCRCQYPCPP GNMYCRCQYPCP
	VVATIFLMMLLVFATDMMAEAKIC	Sequence Consensus	СТКРСА
- CHAILTOAN CODT	VIATIFLMMMLVFATGMVAEARTC	Variable Dositions	FCTRHC
- CHAILITOAN CAPE N	VIATIFLMMMLVFATDMMAEAKICE LVSTICVLLLLLVATEMMGPTMVAE		FCRKPGA FRRRCFCTKHC
	AAVMCLVLATMSSTSPSFCQAGGC		GSYCRGRDCCCKH 🖵
		boxshade 3.2 - mult. align> graphics file	► T
Row: 1 C	Col: 1]

bltree - display and analyze phylogenetic trees

	BioLegato	- • ×
<u>File</u> <u>E</u> dit	Documentation Draw Evaluate Help	
GDE		
_ tree1 _ tree2 _ tree3	(439-86,(((BDAY,(PORT,KILL)),UM8),(((CIBA,OAKR),UM131),((R2G1,((STGERM,MORD),BEER)),G17))))[0.3333] (439-86,(((BDAY,(PORT,KILL)),UM8),((R2G1,((STGERM,MORD),BEER)),(((CIBA,OAKR),UM131),G17))))[0.3333] (439-86,(((PORT,BDAY),UM8),((R2G1,(BEER,(STGERM,(MORD,KILL)))),(((CIBA,OAKR),UM131),G17))))[0.3333]	
4		•
	Row: 1 Col: 1	

blmarker - analyze molecular marker data

	A	В	С	D	E	F
1	G17	0	1	0	1	0
2	UM8	0	1	0	0	0
3	UM131	0	1	0	1	0
4	OAKR	0	1	0	1	0
5	KILL	0	1	0	1	0
6	BEER	0	1	0	1	0
7	BDAY	0	0	1	0	0
8	R2G1	0	1	0	1	0
9	439-86	1	1	0	1	1
10	CIBA	0	1	0	1	0
	•				1	

	CD52097	Display orthology
- Piel agata		Display sorthology
✓ SioLegato		Color accord, to log L
<u>File Edit Documentation Draw E</u>		🔤 🔄 log L values
bio142511 ((CDS2097:0.72175.(C	CD5427	Editable
		Click on Node to:
		Display/Edit information
		Collapse/Uncollapse
	CD51677	Root/Reroot
		Subtree/Parent tree
		 Swap children
		Zoom in X
Row: 1 Col: 1	CD51469	Zoom out X
		Zoom in Y
		Zoom out Y
and the second se		Show whole
	CD5137	Order subtrees
	-	Uncollapse all
the second secon		Collapse to deepest

SUMMARY

•One GUI unifies most data and programs

•BioLegato can utilize any data, any programs, from anywhere

•OO approach models the GUI to biological concepts, which are intuitive to the scientist

REFERENCES

.1. Fristensky B (2007) BIRCH: A user-oriented, locally-customizable, bioinformatics system. BMC **Bioinformatics**, 8:54 2. Smith, S.W., Overbeek, R., Woese, C.R., Gilbert, W., and Gillevet, P.M. (1994) The genetic data environment: an expandable GUI for multiple sequence analysis. Computer Appl. in the Biosciences 10:671-675.

