

Chinook: A collaborative system for bioinformatics analysis.

VanBUG October 2004

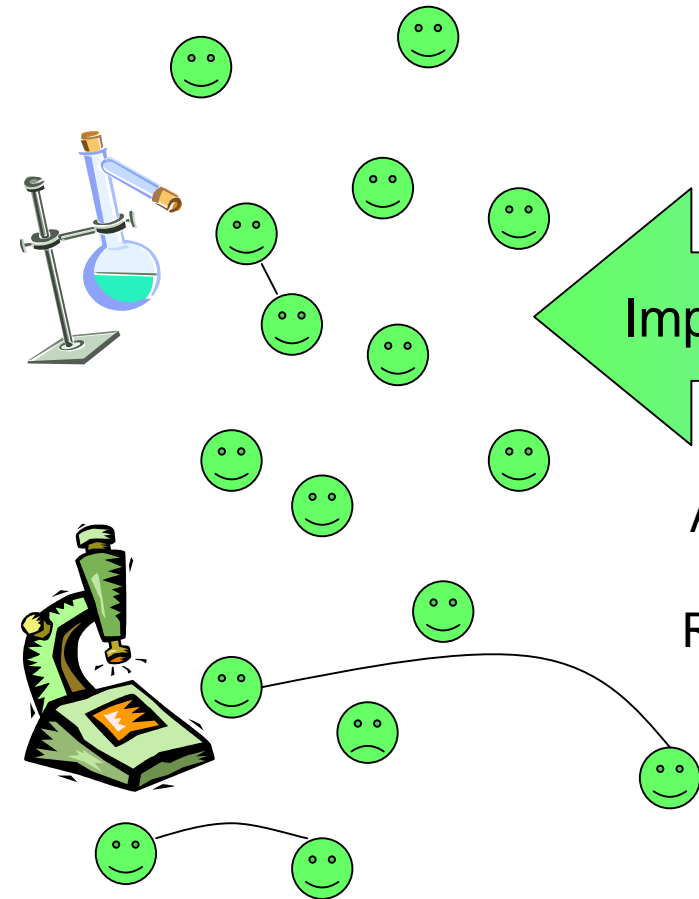
Stephen Montgomery

Things get more complicated...

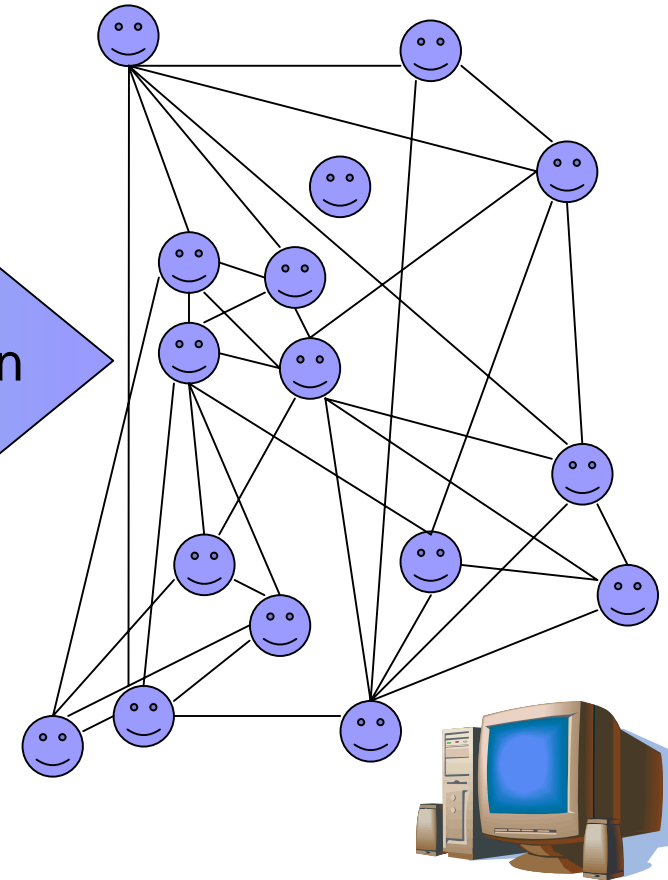
- Each individual has:
 - Access to different resources
 - Computational / Monetary / Personnel
 - Finite time available
 - A different social network
 - Professional obligations
- Each group
 - Organizational boundaries
 - Toolkit (suites and scripts)
 - Method of providing tools (OS, Internet, Interfaces)

An improved interaction map of **biologists** and **bioinformaticians**

Biologists



Bioinformaticians

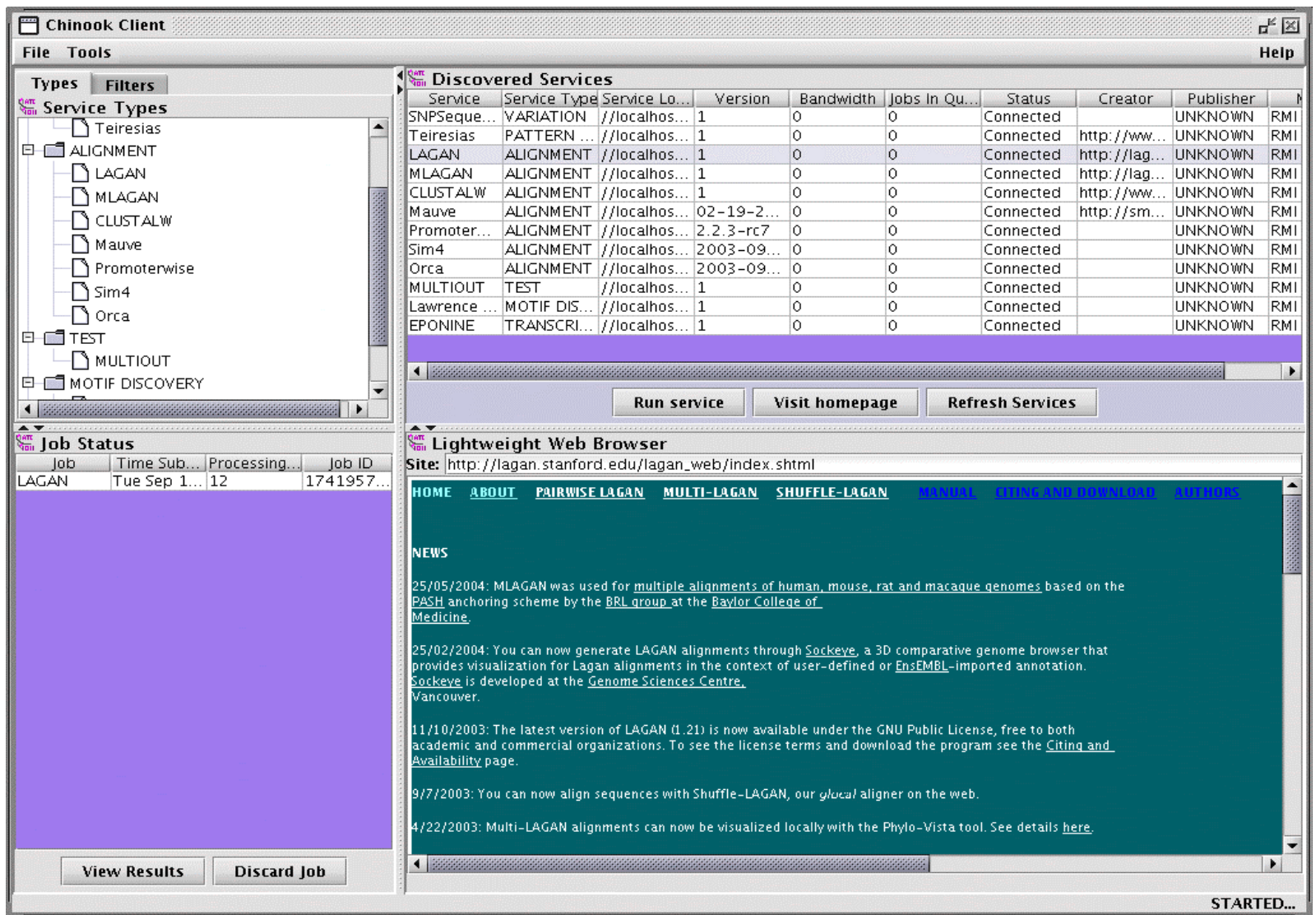


Access to communities
Access to resources
Retain sub-organization

A community-based approach to bioinformatics analysis

- Use the principles of peer-to-peer technology
 - Allow biologists to easily discover and run bioinformatics tools
 - Create a dynamic, reliable network for analysis
 - Reduce overlapping integration efforts
 - Improve communication within/outside organizations
 - Address problems relevant to bioinformatics
 - Attribution
 - Data management
 - Resource distribution





discover and run jobs here or through **bioperl**

Configure service

Data entry

DATA NAME: dna_sequence Edit data

Parameter entry

Name	Value	Type	Description
translate	<input type="checkbox"/>	BOOLEAN	Use transl...
rc	<input type="checkbox"/>	BOOLEAN	Reverse c...
fastreject	<input type="checkbox"/>	BOOLEAN	Abandon ...
chaos		STRING	The conte...
order		STRING	The conte...
recurfl		STRING	Used in re...

OK Cancel



Enter data

Name: A End: 2000
 Start: 1000 Species: Rattus_norvegicus
 Database: rattus_norvegicus_core_22_3b
 Chromosome: 1
 Strand: 1 DatabaseType: ENSEMBL

EMPTY

Add Edit Remove

OK Cancel

CURRENT: 2 MIN: 2 MAX: 2



Enter data

Select retrieval mode
 DNA_LOCATION

Name: A
 Chromosome: 2
 Start: 1000000
 End: 1001000
 Strand: 1
 DatabaseType: ENSEMBL
 Species: Rattus_norvegicus
 Database: rattus_norvegicus_core_22_3b

DATA VALID

OK Cancel

Result information

Job Status

JOB COMPLETED STATUS: success
 Download result files from table

Preview results

Stdout Stderr

Result files

File ID	Description	Downloaded	Local File...
6322.178...	Standard ...	false	
374.1798...	Standard ...	false	

Download

OK Cancel

Preview

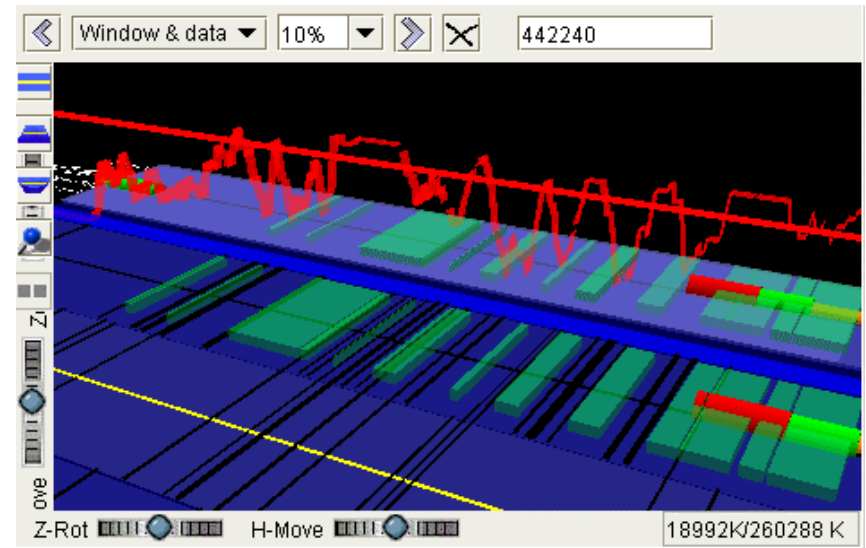
```
Using /tmp/chinook.tmp.-1758111813 /tmp/chinook.tmp.-314156771
(12, 0, 25, 0) x
Using /tmp/chinook.tmp.-1758111813 /tmp/chinook.tmp.-314156771
(13, 1, 30, 0) x
Using /tmp/chinook.tmp.-1758111813 /tmp/chinook.tmp.-314156771
(8, 1, 30, 0) x
Using /tmp/chinook.tmp.-1758111813 /tmp/chinook.tmp.-314156771
(7, 1, 30, 0) x
0 cumulative anchors
read 0 anchs
Total size = 3 * 10^6
```

Algorithms integrated into Chinook

ClustalW	Genscan
Conreal	Sim4
DIALIGN	MSCAN
LAGAN	ANN-Spec
Mauve	Recursive Gibbs Motif Sampler
ORCA	MEME
Shuffle-LAGAN	Motifsampler
T-Coffee	RSAT oligo analysis
Promoterwise	STUBB
Primer3	Teiresias
Eponine	wConsensus
ELPH	

Projects involving Chinook

- **OrthoSEQ** plans to provide analysis through the Chinook/BioPerl Perl API.
- **Sockeye** uses Chinook to deliver state-of-the-art alignment, PCR prediction, and regulatory analysis
- **Pegasys** plans to provide pipeline management to subset of services advertised by Chinook.
- **Bio-Linux** planning to integrate a subset of their algorithms.





Use cases of Chinook

- Grid/Cluster computing.
- Internally connect teams / individuals.
- Collaborate with remote individuals.
- Provide an API layer to your algorithms.
- Insert bioinformatics analysis into applications.
- Show off your tools.

Future Plans

Political Map of the World



Acknowledgements

GENOME SCIENCES CENTRE

- Steve Jones
- Tony Fu
- Jun Guan
- Keven Lin

- Asim Siddiqui
- Genereg team @ GSC

- Mark Mayo
- Bernard Li

CENTRE FOR MOLECULAR MEDICINE AND THERAPEUTICS

- Wyeth Wasserman
- Jonathan Lim

UBC BIOINFORMATICS CENTRE

- Francis Ouellette
- Graham McVicker
- Sohrab Shah

Funding: MSFHR, Genome Canada

VISIT <http://smweb.bcgsc.bc.ca> OR <http://www.bcgsc.bc.ca/chinook>