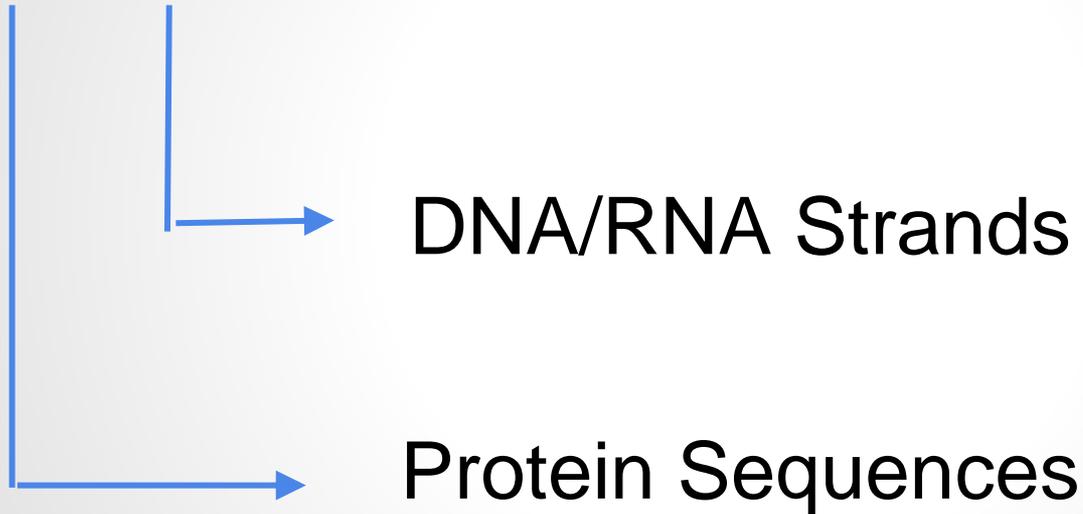


Bioinformatics Applications & Analysis using PoPLAR Gateway

Nyalia Lui, Yuan Liu,
and Bhanu Rekepalli PhD

Bioinformatics Focus

Genomic Sequencing



Problems



Data Growth

Abundance of Data

Unorganized Data

Knowledge Discovery

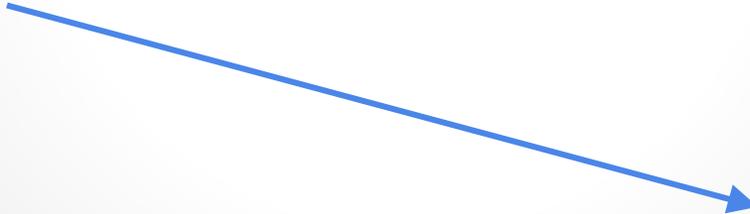
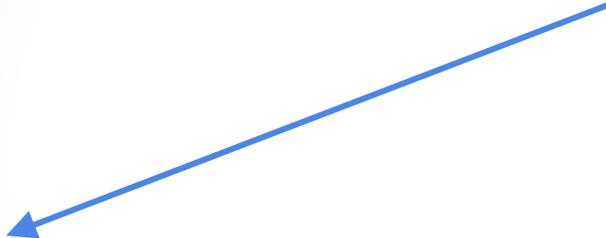
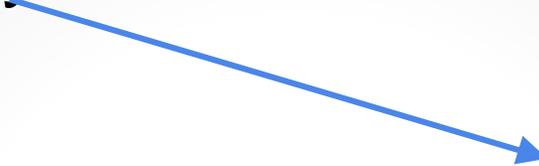
Explored Aspects of Bioinformatics

Genome Assembly

Sequence Analysis

Structures

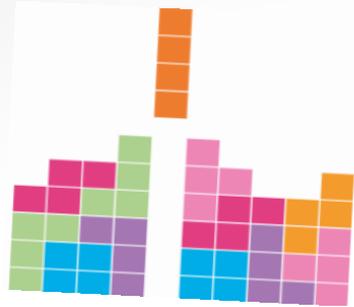
Drug Discovery



Applications

BLAST

(Basic Local Alignment
Search Tool)



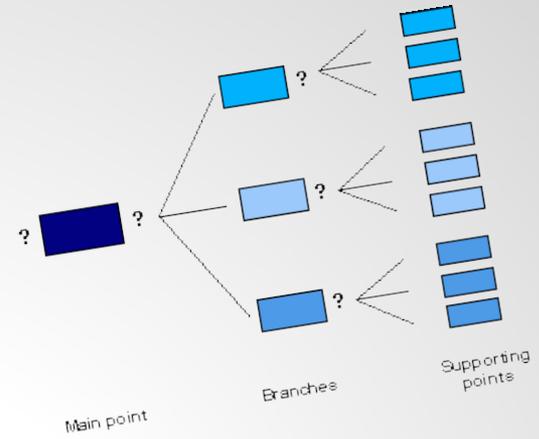
MUMMER

HMMER



PhyML

- RAxML



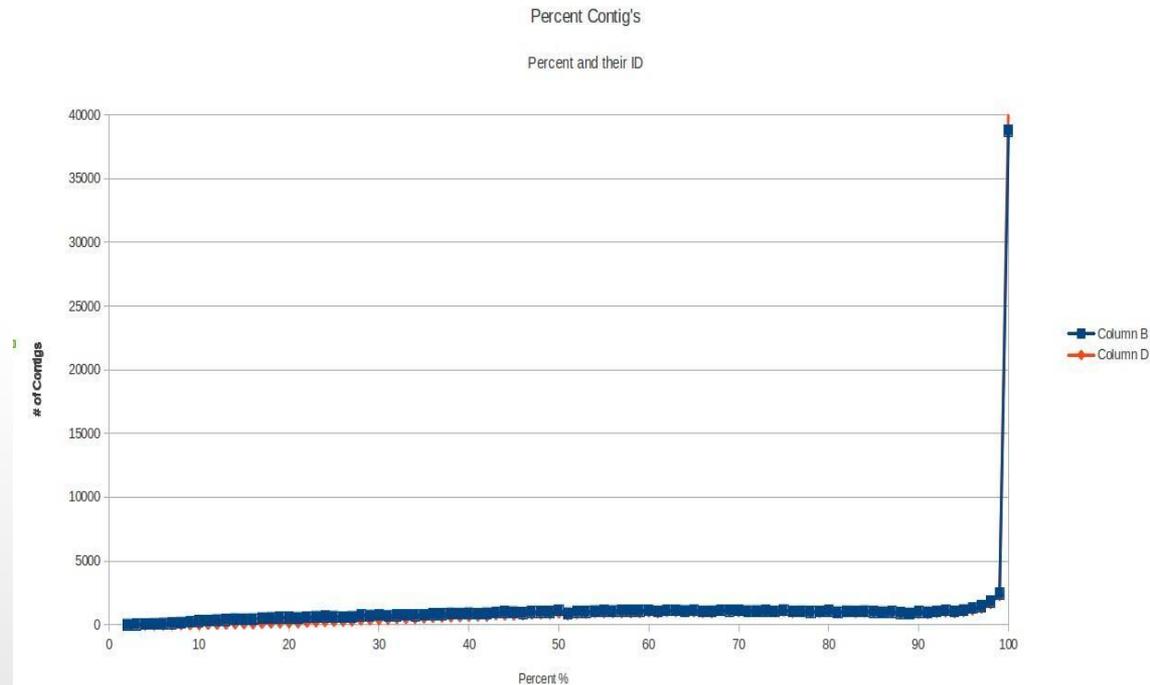
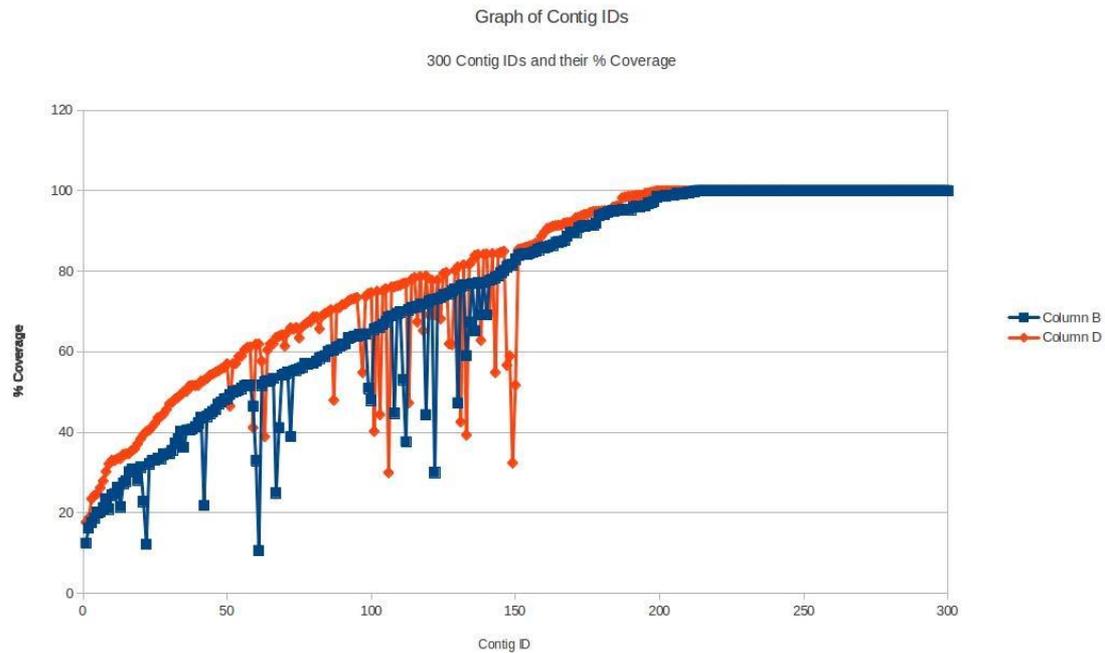
MUSCLE

ClustalW
AABCGLK
OKBCGKL
• |BCG|



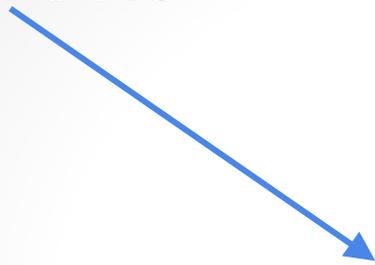
BLAST vs MUMMER

- Contigs vs Nodes
- What are we comparing?
- What is coverage?
- BLAST vs MUMMER

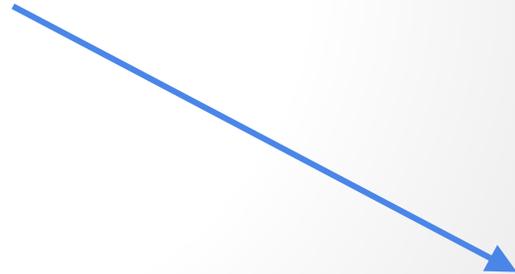


Our Focus

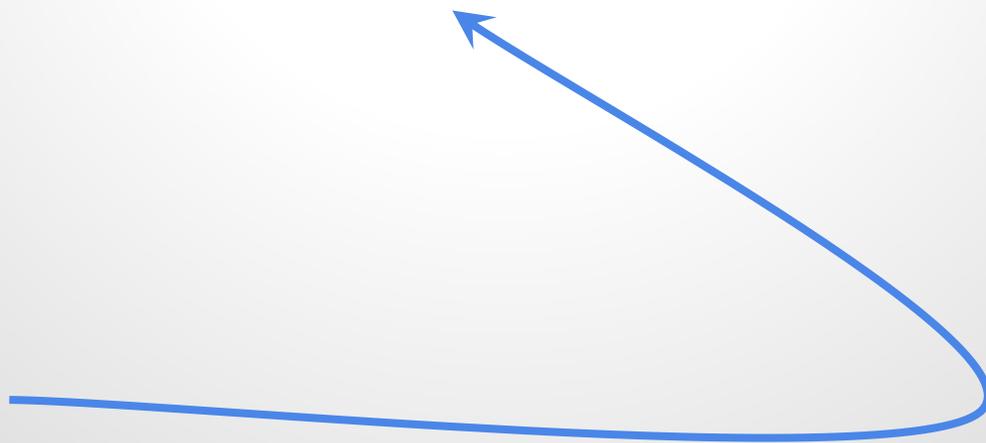
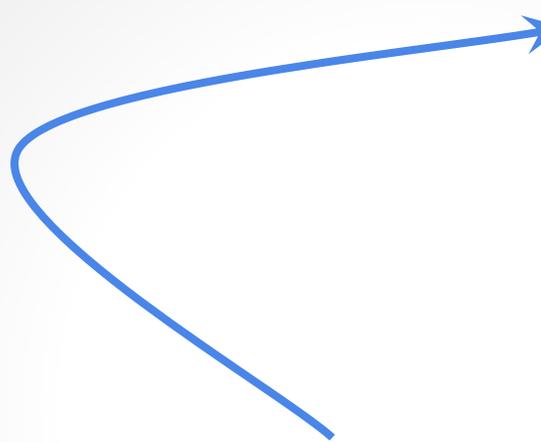
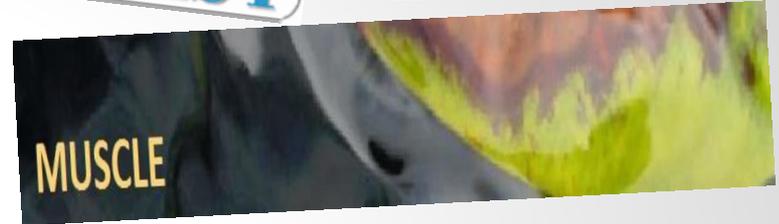
Analytical Tools



Develop Resources



PoPLAR Science Gateway





PoPLAR

Portal for Petascale Lifescience Applications & Research

Home

Toolkit

My Workbench

My Profile

Help

How to Cite Us

Logout

Folders

- Tester**
 - Data (4)
 - Tasks (1)

Tasks

[Refresh Tasks](#)

Current CPU Hr Usage: 0 [Explain this?](#)

There is currently 1 data item in this tab.

Show records on each page

[<<](#) Page 1 of 1 [>>](#)

<input type="checkbox"/>	Select all	Label	Tool	Input	Parameters	Output	Date Created	Action
<input type="checkbox"/>	Clone	Test BLAST	HSP BLAST on Kraken	View (1)	View (5)	View (2)	7/29/13, 16:03	View Output

Move selected to [GO](#)

[Create New Task](#)

[Delete Selected](#)

[Report an Issue](#)



NICS - University of Tennessee - Oak Ridge National Laboratory
PO Box 2008, BLDG 5100 - Oak Ridge, TN 37831-6173

Valid: [xhtml](#) | [css](#)





PoPLAR

Portal for Petascale Lifescience Applications & Research

Home

Toolkit

My Workbench

My Profile

Help

How to Cite Us

Logout

Folders



Tester

Data (3)

Tasks (1)

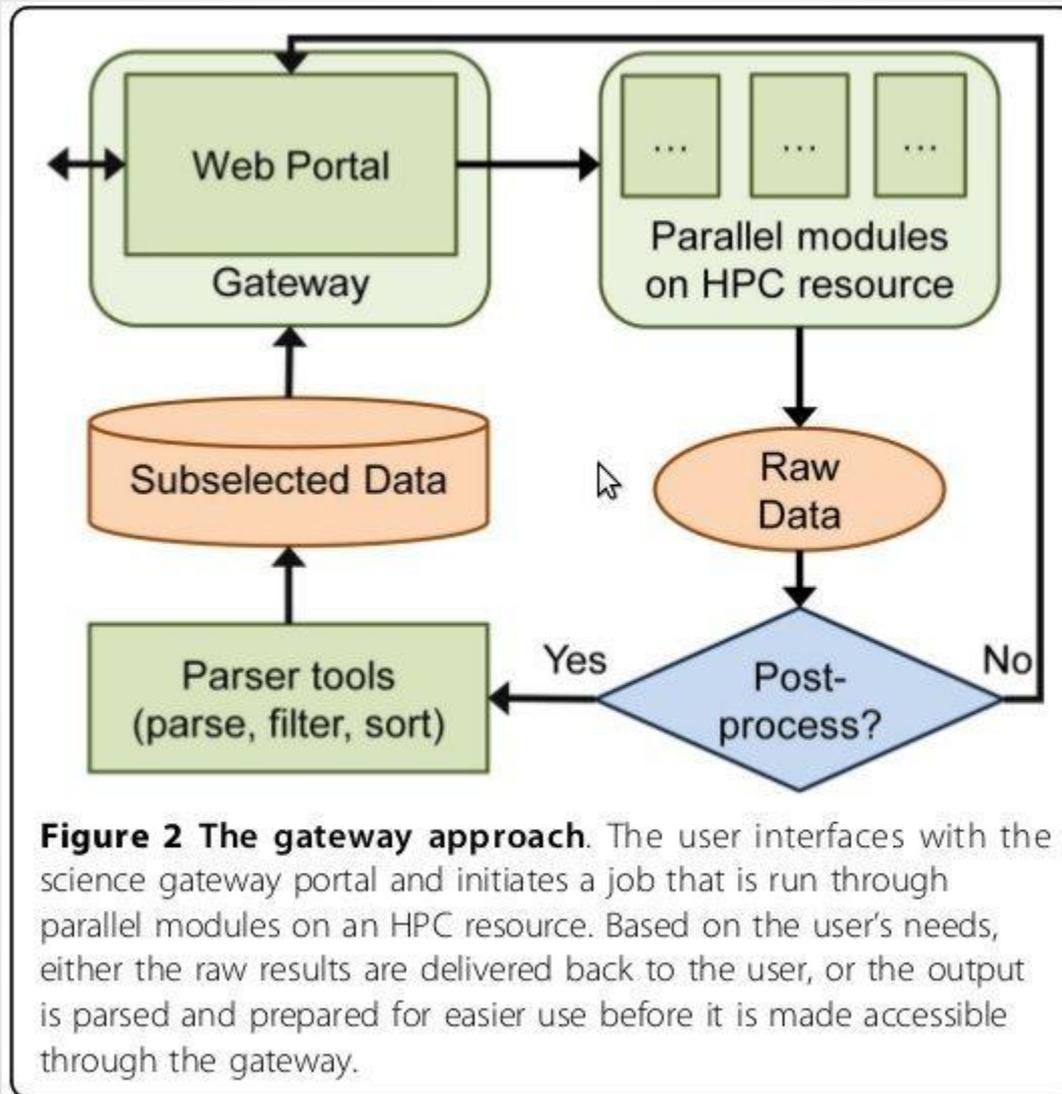
View Data Details

Label:	Test BLAST_output
Owner:	nlul1
Group:	nlul1
Bytes:	1048861
Date Created:	7/30/13, 09:10
Data Format:	Blast XML Output

Show/Hide Data Contents

NODE_7_length_72_cov_2.875000	gi 315587265 gb CP002343.1	87.95	83	10	0	1	83	1261023	1260941	6e-14
NODE_7_length_72_cov_2.875000	gi 315587265 gb CP002343.1	86.90	84	11	0	1	84	2326016	2325933	4e-12
NODE_7_length_72_cov_2.875000	gi 170774137 gb CP001013.1	100.00	25	0	0	63	87	901037	901061	0.003
NODE_7_length_72_cov_2.875000	gi 269303491 gb CP001821.1	100.00	22	0	0	64	85	3142236	3142257	0.21
NODE_7_length_72_cov_2.875000	gi 158107272 gb CP000820.1	82.43	74	13	0	1	74	558528	558601	0.21
NODE_7_length_72_cov_2.875000	gi 158107272 gb CP000820.1	82.43	74	13	0	1	74	3658835	3658762	0.21
NODE_7_length_72_cov_2.875000	gi 158107272 gb CP000820.1	82.43	74	13	0	1	74	5513285	5513212	0.21
NODE_7_length_72_cov_2.875000	gi 336284912 dbj AB611743.1	100.00	21	0	0	31	51	112	92	0.85
NODE_7_length_72_cov_2.875000	gi 284012950 gb CP001860.1	100.00	21	0	0	60	80	2764371	2764351	0.85
NODE_7_length_72_cov_2.875000	gi 226237899 dbj AP011115.1	86.67	45	6	0	31	75	7906142	7906186	0.85
NODE_7_length_72_cov_2.875000	gi 378401447 emb HE663493.1	100.00	20	0	0	38	57	1018485	1018504	3.4
NODE_7_length_72_cov_2.875000	gi 365267655 gb HQ852724.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267657 gb HQ852725.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267653 gb HQ852723.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267651 gb HQ852722.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267649 gb HQ852721.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267647 gb HQ852720.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267645 gb HQ852719.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267615 gb HQ852704.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267641 gb HQ852717.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267623 gb HQ852708.1	100.00	20	0	0	35	54	99	80	3.4
NODE_7_length_72_cov_2.875000	gi 365267637 gb HQ852715.1	100.00	20	0	0	35	54	99	80	3.4

The life of a running task



Research Use



Provide easier access to command line tools through PoPLAR Gateway.

Educate Biologists on accessing HPC resources for large scale data analysis.



Future Work



Adapt more bioinformatics tools into parallel applications.

Incorporate the other aspects of bioinformatics

Thank You

- Questions?

Acknowledgements

