



Proteome analysis tool for Microorganisms



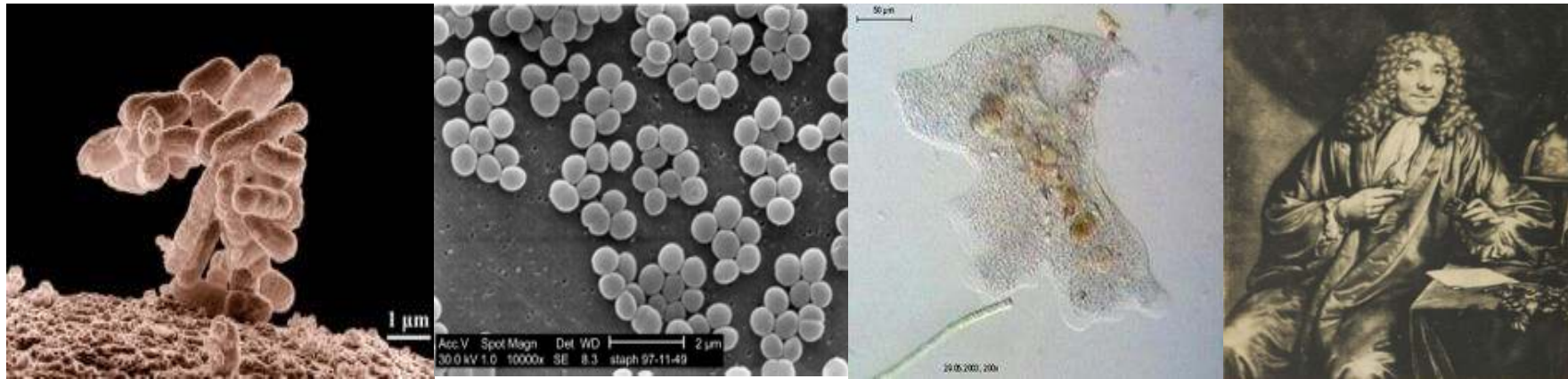
GPDB (Genome Profile DataBase)

李季青

國立清華大學 資訊工程系



Introduction: Microorganisms



- Various morphological and environmental distributing properties
- Microbial identification and classification is more difficult than higher organisms.



Conserved molecular biomarkers

- Since 1980s
- 16S rRNA gene (1500 bp)
 - ⇒ Ubiquity
 - ⇒ Conserved region
 - ⇒ Variable region
- 16S rRNA + Other biomarkers → modern taxonomy

Biochemical typing
Metabolic tests
DNA-DNA hybridization
GC contents
Chemotaxonomic markers
...

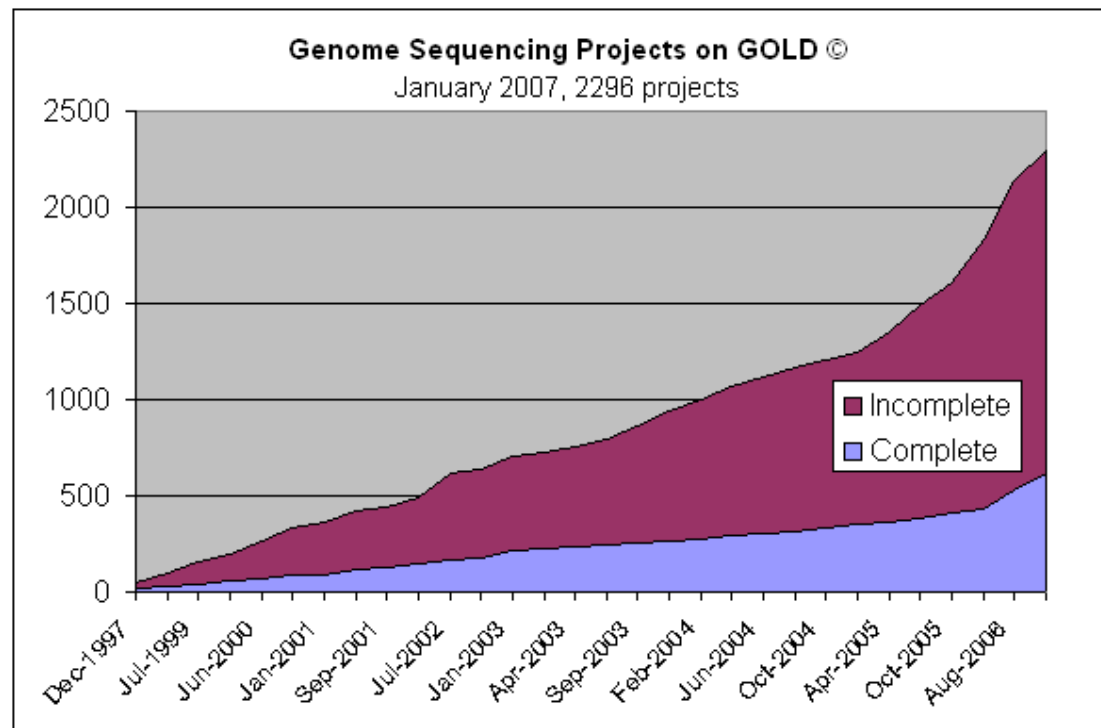


May not universal of all organisms



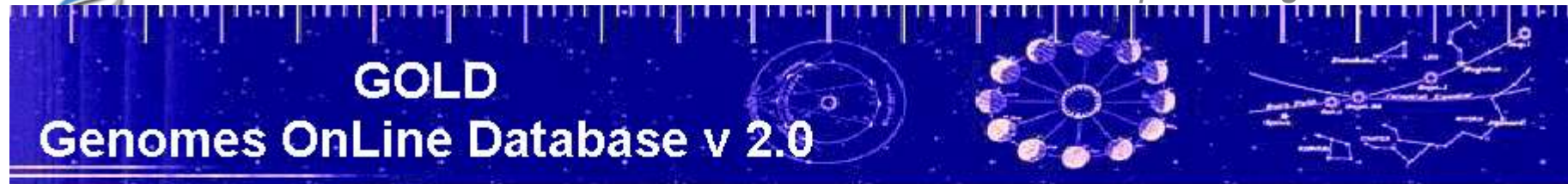
Genome sequences

How about biochemical / metabolic features, life cycle, GC-contents, type of host cell ... ?
They are determined by entire genome.





Data from: <http://www.genomesonline.org/>



Contact: Genomesonline	Last Update: July 31, 2008	Location www.genomesonline.org
842 Published Complete Genomes	Search GOLD: 3919 genome projects	130 Metagenomes
97 Archaeal Ongoing Genomes	1900 Bacterial Ongoing Genomes	950 Eukaryotic Ongoing Genomes

- * Most of complete genomes are microbes
- Valuable sources for comparative genomics studies



Genome comparison methods

Long multiple-sequence alignment
rearrangement

cause evolution events (gene order)
at such high rates.

ods

Result

$100 * (G+C) / (A+C+G+T)$
of whole genome
sequence

GC
content

shared gene
content

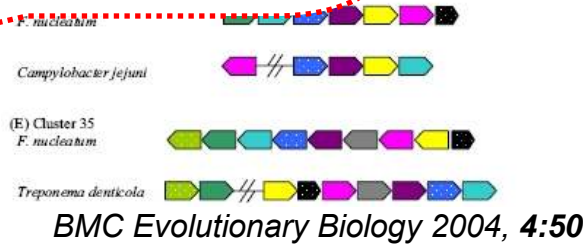
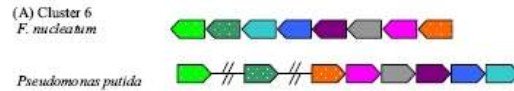
Alignment-free
statistical method

Chromosome
gene order

Whole
genome
sequence
alignment

Quick and Dirty

Computing time





How we study genomic data?

- ❑ Information derived from both **nucleotide** and **protein** sequence in a genome-wide scale.
- ❑ Provide and compare features of the fully sequenced organisms **in a graphic and easy-reading way**.
- ❑ On-line **graphic browsing interface** and **use hierarchical clustering method** to compare and view the difference between these organisms



The profile



Current Genome Profile

■ Basic information -

- ✓ Species name, taxonomy, # of chromosome/plamid , genome size, orf number...

■ Nucleotide composition -

- ✓ ATGC composition, GC/AT content, N-nucleotide frequency (n=2,3), Codon usage...

■ Amino acid composition -

- ✓ Amino acid group composition, N-peptide frequency distribution (n=1,2), Proteome length, Mw, pI distribution...



Basic information



Genome Profile DataBase

[Home](#) | [Browse](#) | [Virtual 2D](#) | [Compare](#) | [Download](#) | [Status](#) | [Help](#) | [Comment](#)

[Bacteria] - Helicobacter pylori 26695

Species Name :

Helicobacter pylori 26695 (Taxonomy id : 85962)

Genome list :

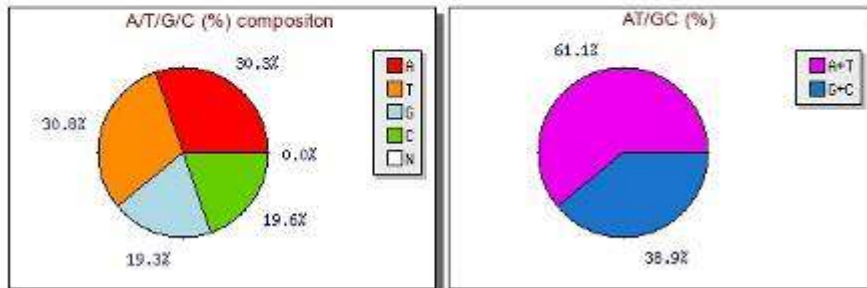
#	Accession	gi	Description	Size(bp)	Orfs
1	NC_000915	15644634	Helicobacter pylori 26695, complete genome	1667867	1576

Total Genome Size = 1.67 (Mb) , ORFs = 1576

Nucleotide composition :

A	T	G	C	N
30.3 % (505397 bp)	30.8 % (514075 bp)	19.3 % (321273 bp)	19.6 % (327080 bp)	0.0 % (42 bp)

A+T content = 61.12 % , G+C content = 38.87 %





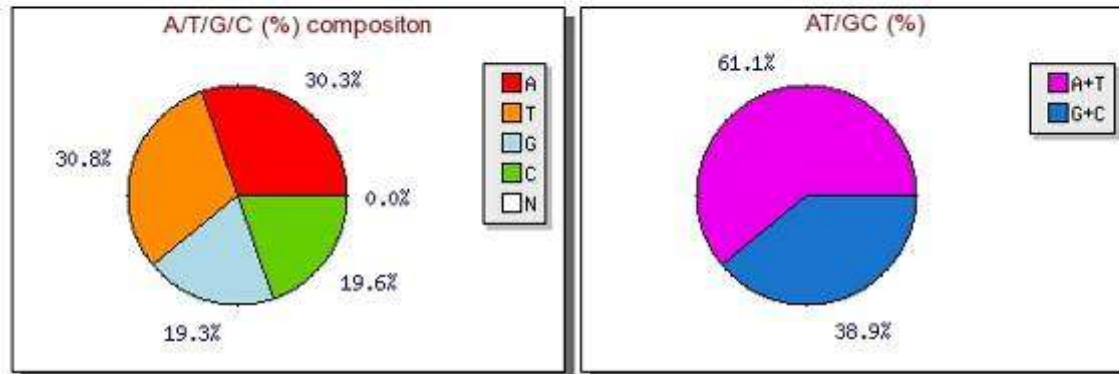
Nucleotide Composition

GC/AT content, GC/AT Skew

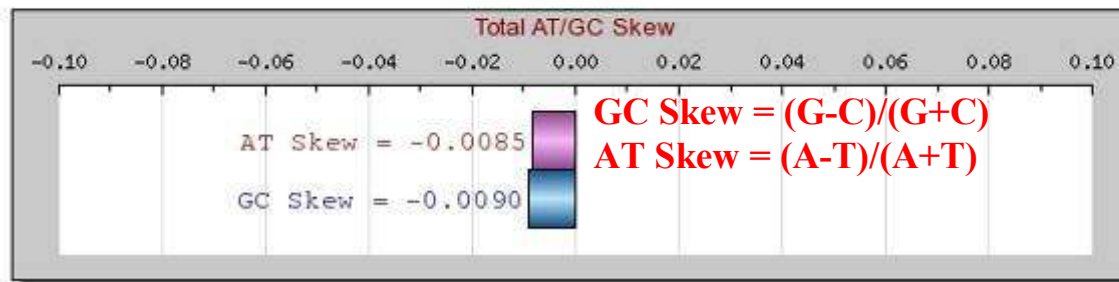
Nucleotide composition :

A	T	G	C	N
30.3 % (505397 bp)	30.8 % (514075 bp)	19.3 % (321273 bp)	19.6 % (327080 bp)	0.0 % (42 bp)

A+T content = 61.12 %, G+C content = 38.87 %



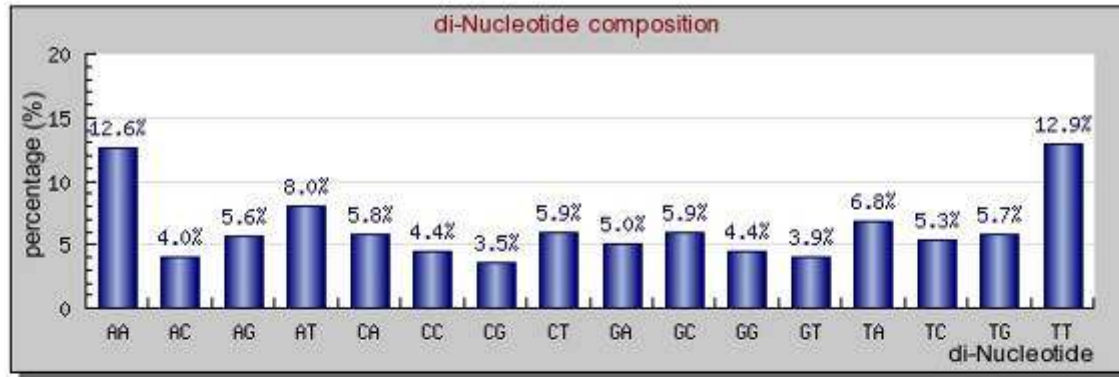
Total GC/AT Skew



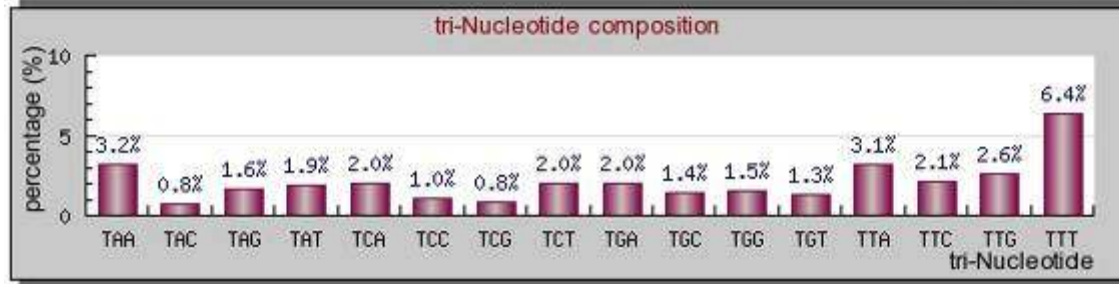
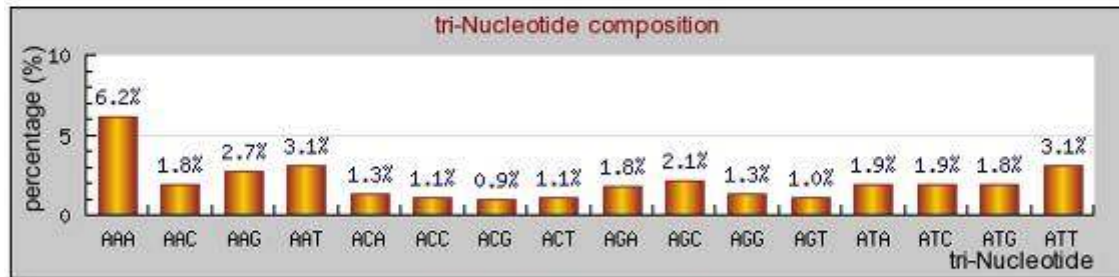


N-nucleotide Composition (N=2,3)

di-Nucleotide composition :



tri-Nucleotide composition :



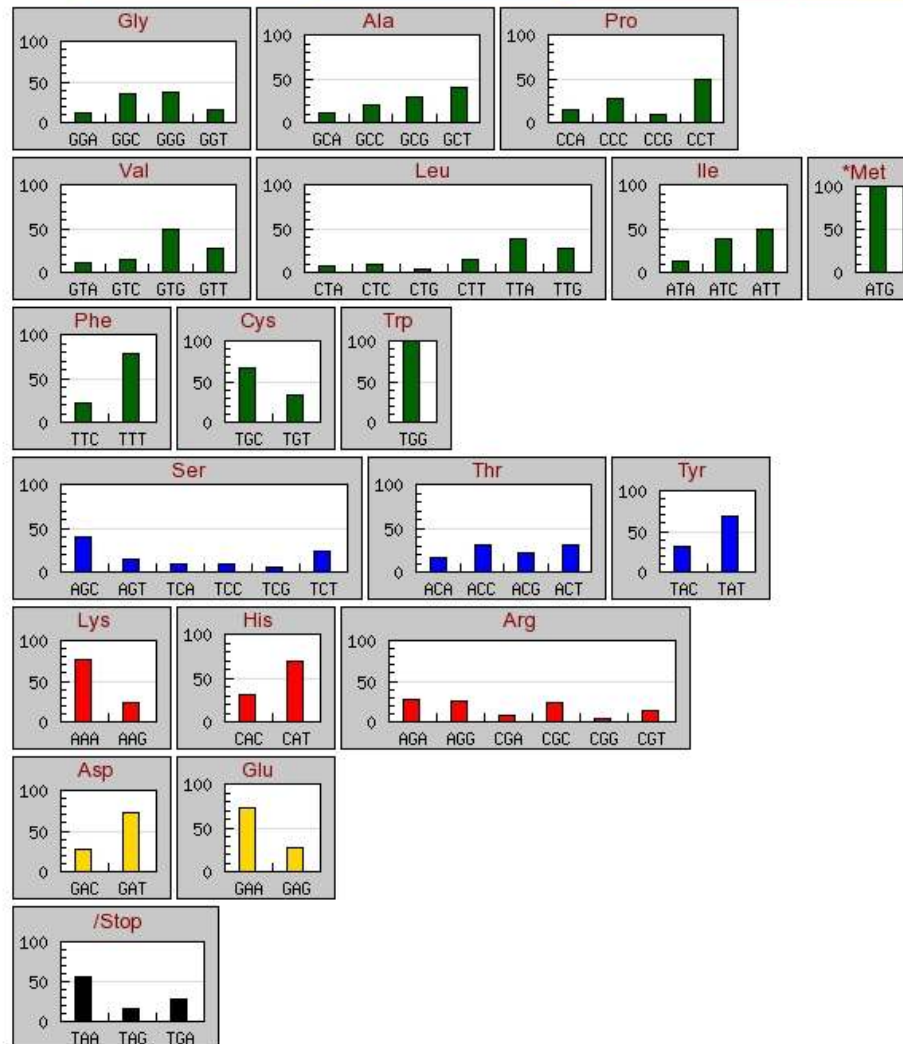


Codon Usage

Codon Usage :

◆ Total Codon Usage (%)

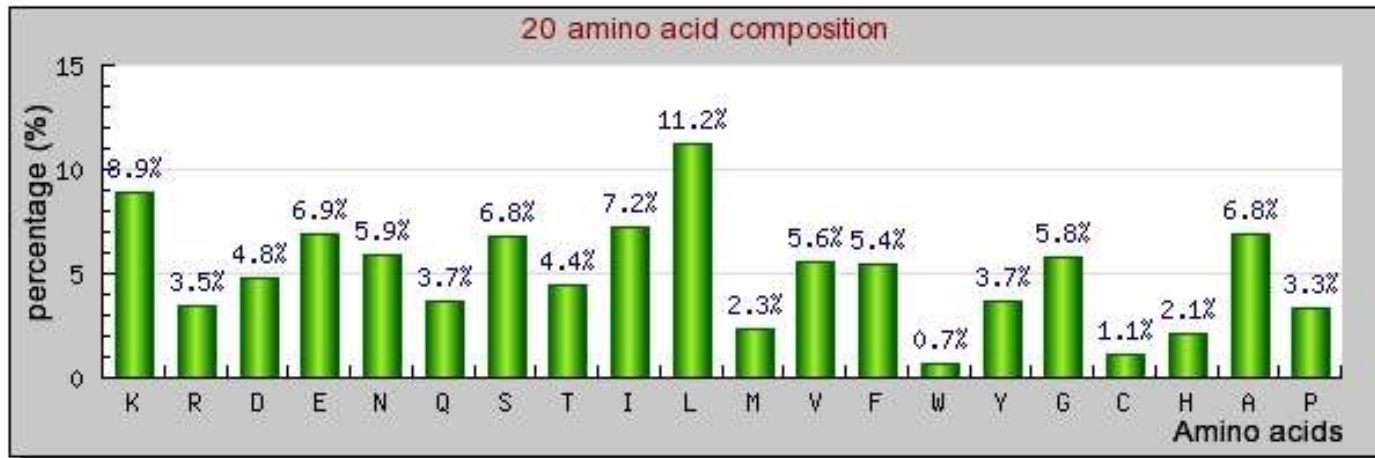
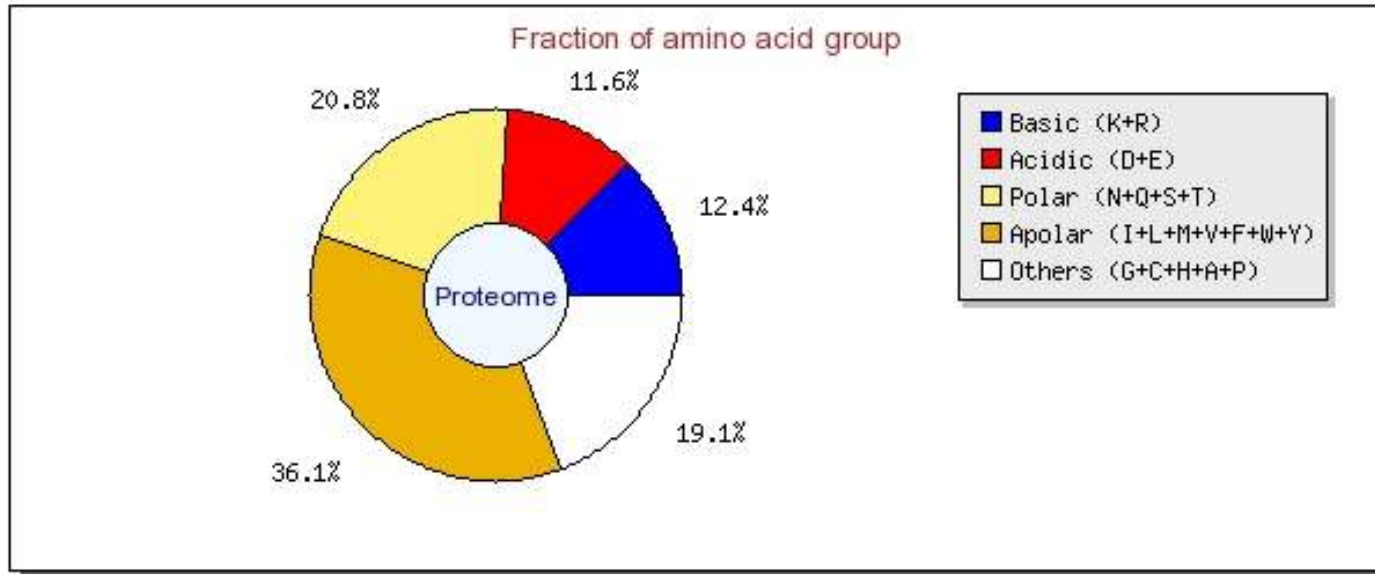
color: apolar amino-acid polar amino acid positive amino acid negative amino acid stop codon





Amino Acid Composition

Amino acid composition :

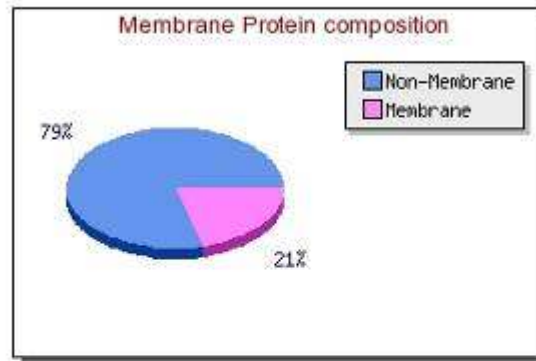




Proteome Distribution- Non-membrane/Membrane protein & Charge distribution

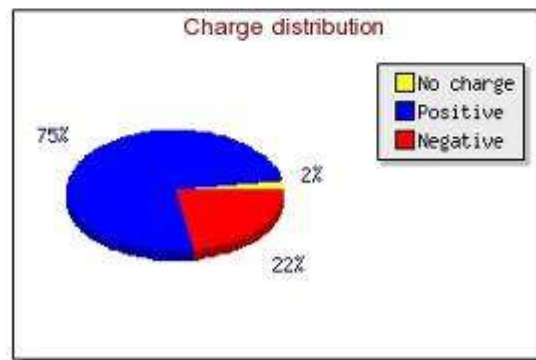
Proteome Distribution :

- o Non-Transmembrane/Transmembrane protein distribution



TMHMM prediction

- o Charge distribution



protein with

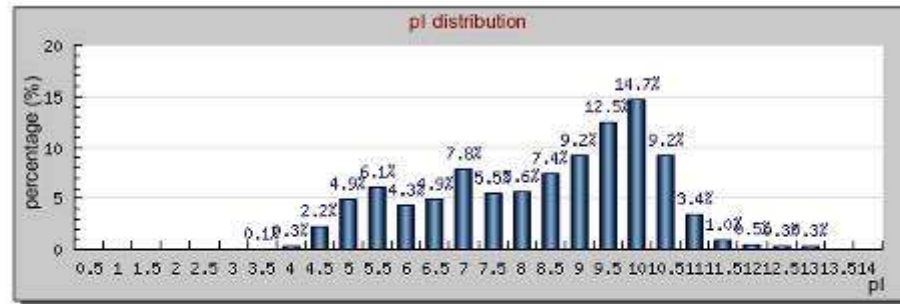
- positive charge
- negative charge
- no charge

Proteome Distribution-



pi & Mw & Length

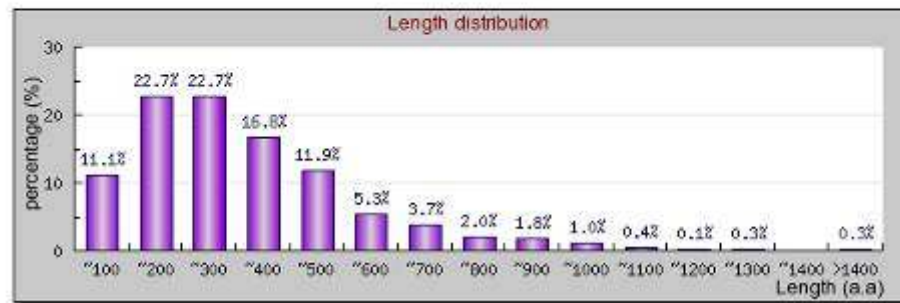
o Isoelectric point(pi) distribution

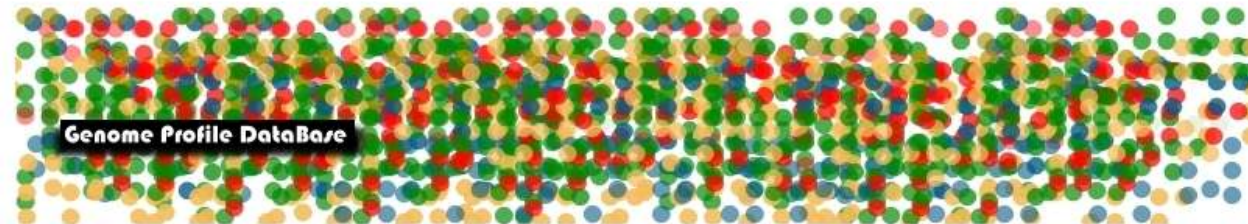


o Molecular Weight distribution



o Length distribution





What's Genome Profile DataBase (GPDB) ?

GPDB has been developed to provide and compare features of the fully sequenced organisms in a graphic and easy-reading way. We focus on these prokaryotes including bacteria and archaea. In this post-genome era, we try to grab the information derived from both nucleotide and protein sequence in a genome-wide scale. We provide some "Genome Profile", develop on-line graphic browsing interface and use hierarchical clustering method to compare and view the difference between these organisms. The original sequence data and annotations are from [NCBI GeneBank](#) and [RefSeq](#) databases. The species name system is from [Taxonomy](#) database. We wrote the `perl` program called "Genome Profile Pipeline" which can automatically mirror data from NCBI ftp site and continue to parse, caculate, and analyse data.

Current Status

Last update 2006/8/3

345 organisms (27 Archaea, 338 Bacteria)

Development Platform

Powered by Apache+PHP+MySQL under RedHat Linux 9.0.

Version 2.5 Update Log

- 2006-10-06 Add registration function.
 - 2006-08-24 Add sorting function.
 - 2006-08-20 New style.
 - 2006-08-12 New compare interface.
 - 2006-08-03 Update database to 2006/08/03.
 - 2006-07- Fix automatically update scripts
 - 2006-06-10 Java robot for collecting web resource.
 - 2006-04-30 Automatically update scripts.
 - 2006-02-06 Version 2.5 project engage.
-
- 2004-11-13 Add Virtual 2D menu to main page.
 - 2004-04-04 Add Total Codon Usage Browse & Comparison.
 - 2004-03-29 Free Registration Form added.
 - 2004-03-25 Add Total GC/AT Skew and Transmembrane protein distribution
 - 2004-03-24 Release Version 2.0
 - 2004-03-23 Virtual 2D gel completed.
 - 2004-03-19 Fix Mw/Length scale.
 - 2004-02-26 The release version of GPDB is 1.0. Add four-type comparison options.
 - 2003-12-15 Version 0.1. Platform build including basic architecture ,database design, web interface.



Browse – *Helicobacter pylori* 26695

GPDB Genome Profile DataBase

Home | **Browse** | Virtual 2D | Compare | Download | Status | Help | Comment

Browse

Select One Genome Profile to Browse:

Archaea - (Total 27 Species)

Aeropyrum pernix K1 Submit

Bacteria - (Total 338 Species)

Helicobacter pylori 26695 Submit

Fungi - (Total 5 Species)

Aspergillus fumigatus Af293 Submit

Virus - (Total 200 Species)

Acanthamoeba polyphaga mimivirus Submit

Best resolution above 1280*1024 | Copyright ©2006 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Szu-Ming Lai and Chi-ching Lee



So many genome profile, and then ?





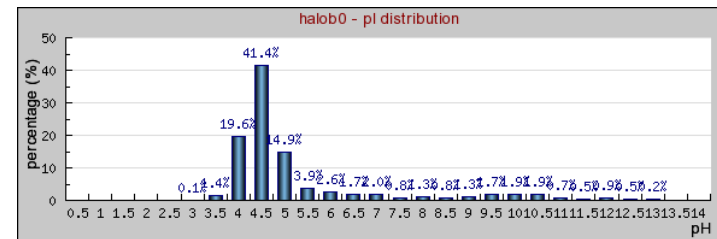
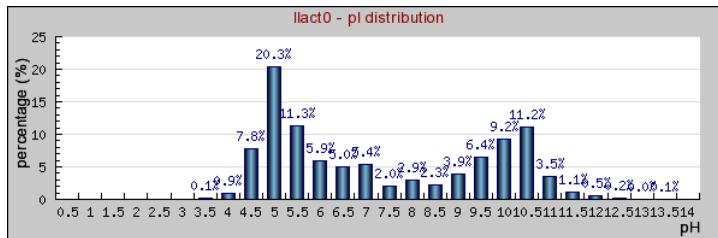
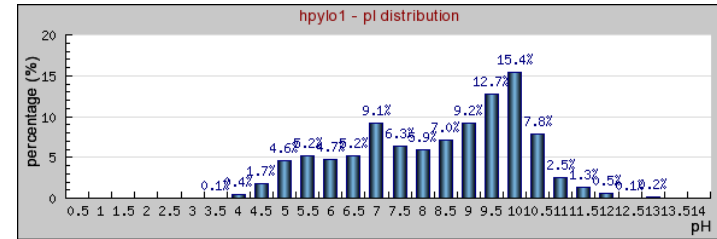
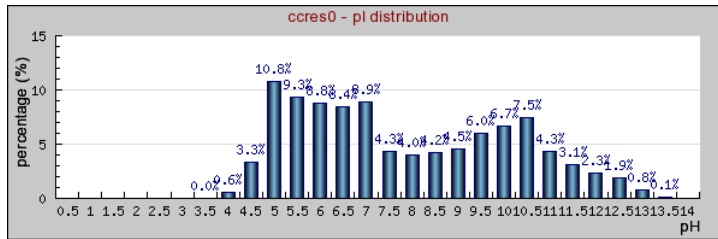
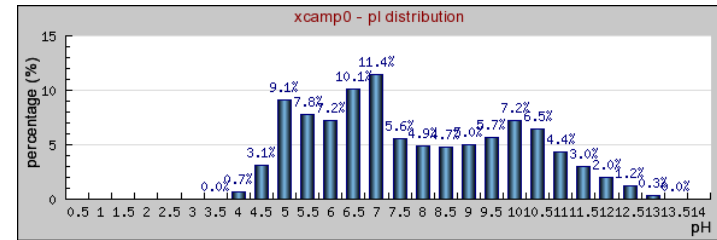
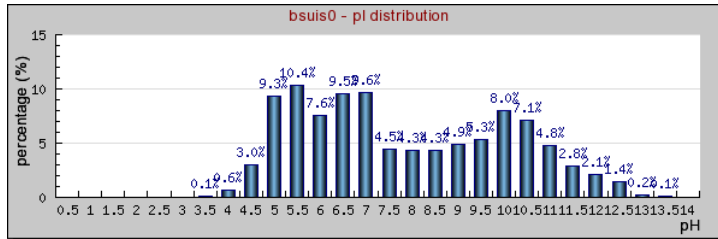
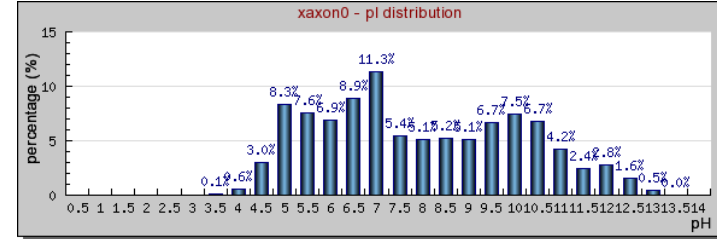
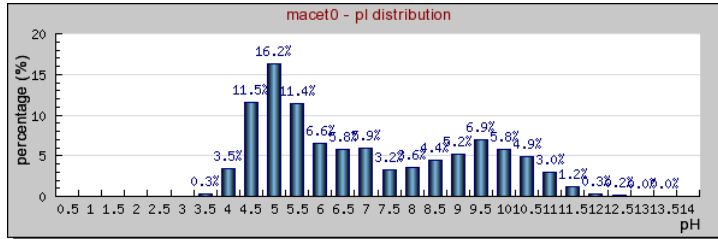
Does whole genome profile reflect phylogeny or environment or ?

- Phylogenomics- whole genome scale.
- Is it possible to explore it, using different whole genome profile.
- We don't know the answer.
- Then, we try.



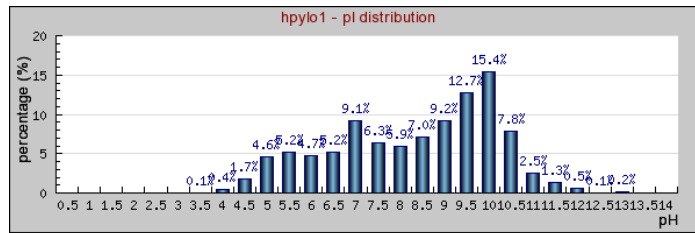
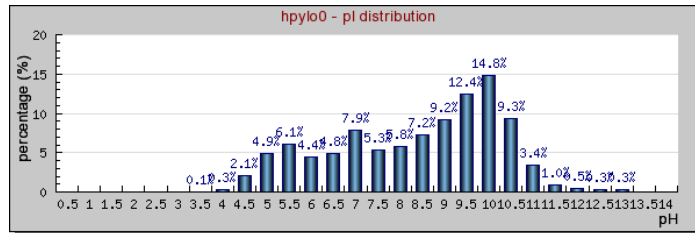


Ex: pI distribution

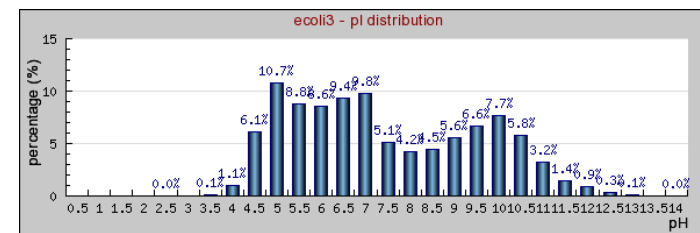
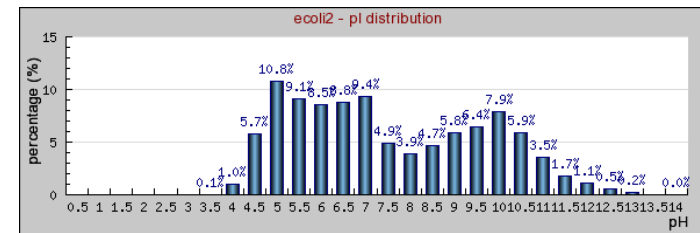
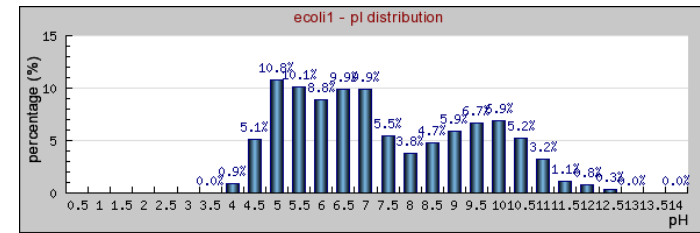
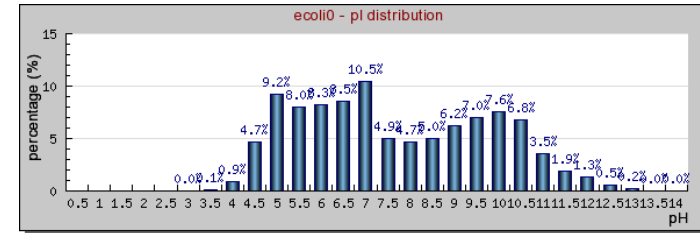




Similar ?

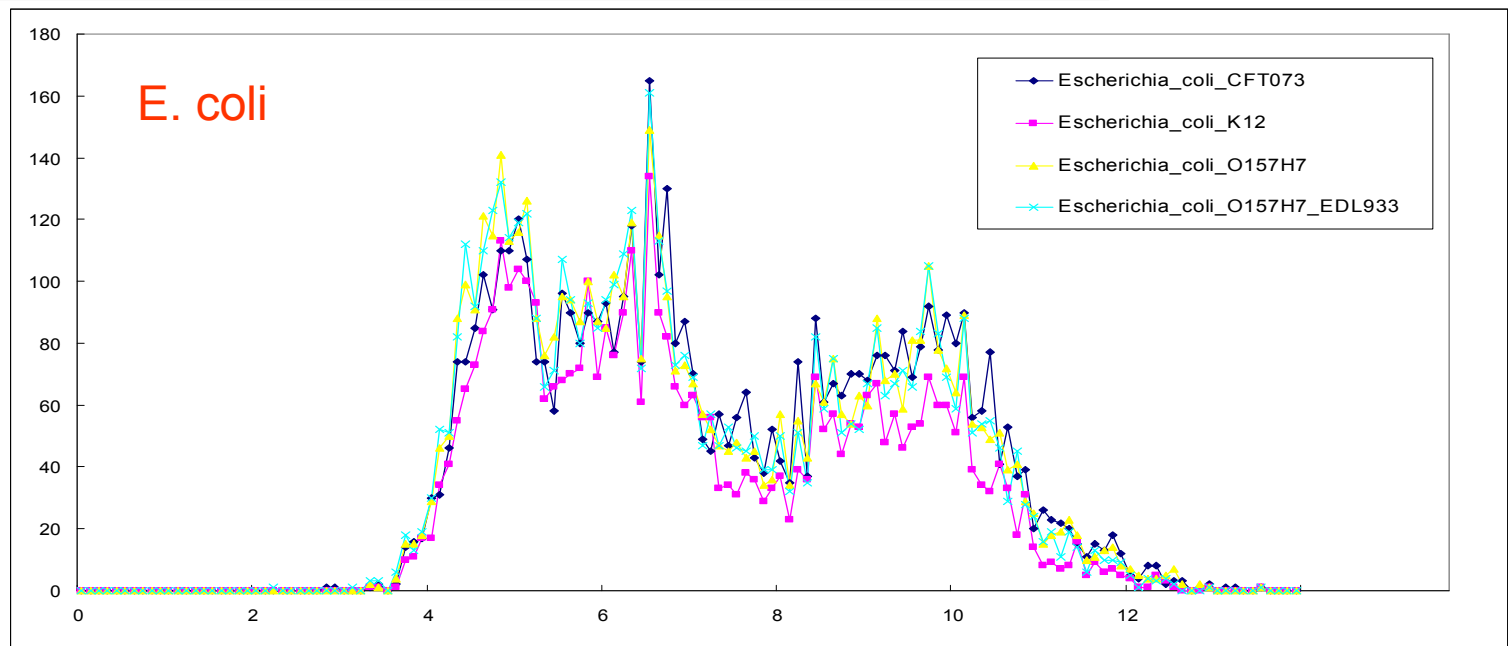
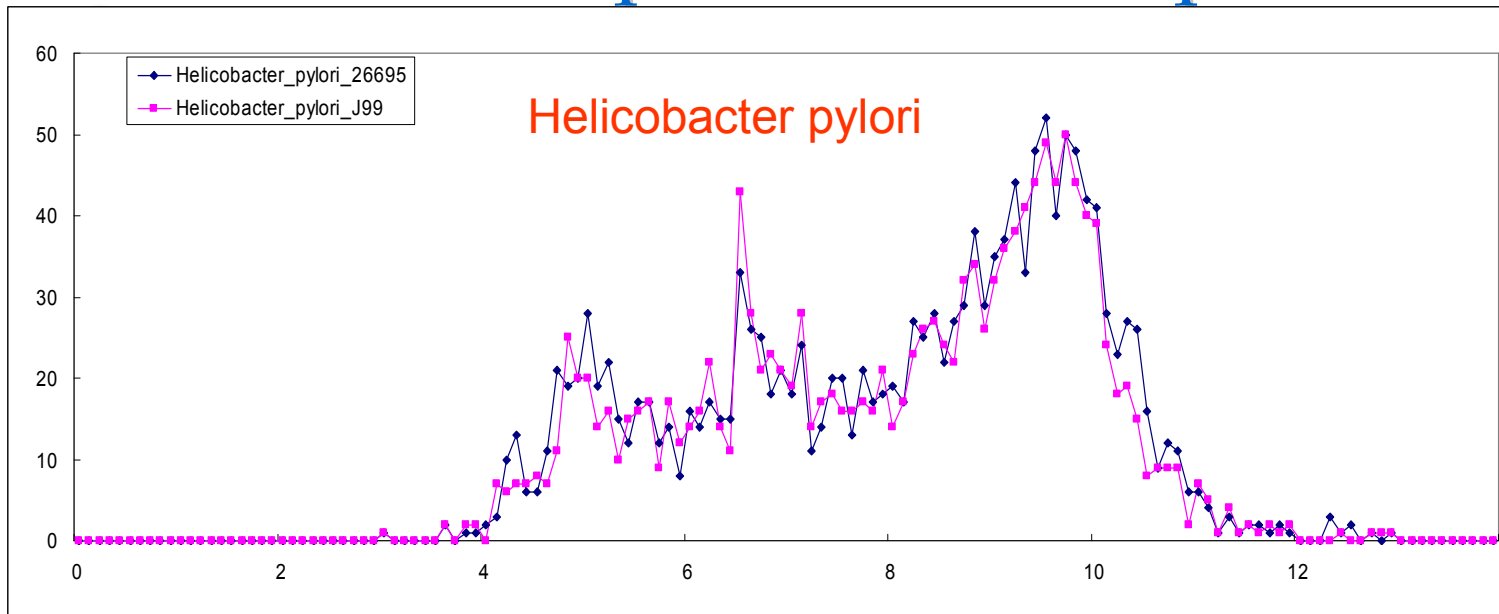


Helicobacter pylori



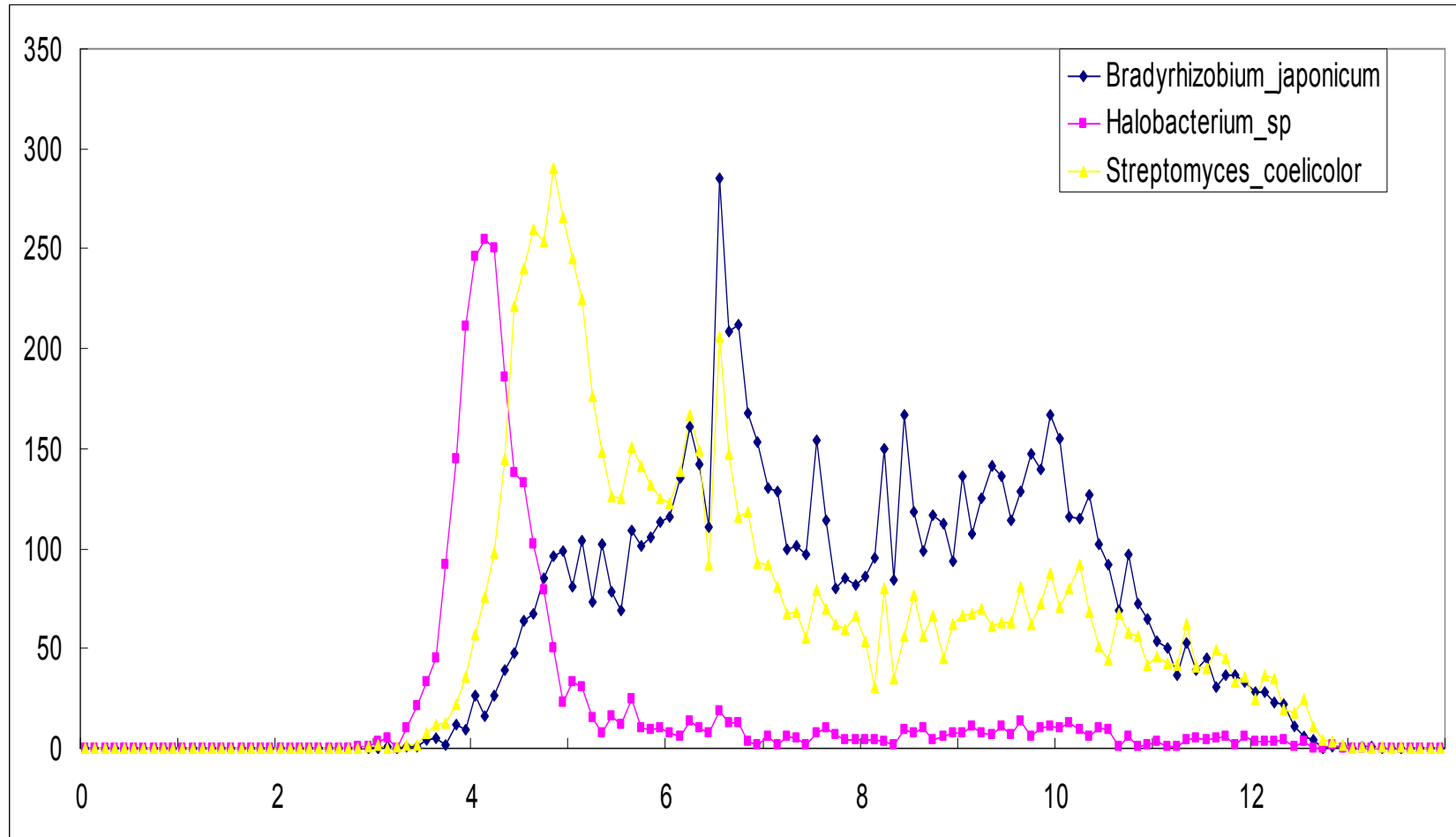
E. coli

Similar species, similar pI distribution



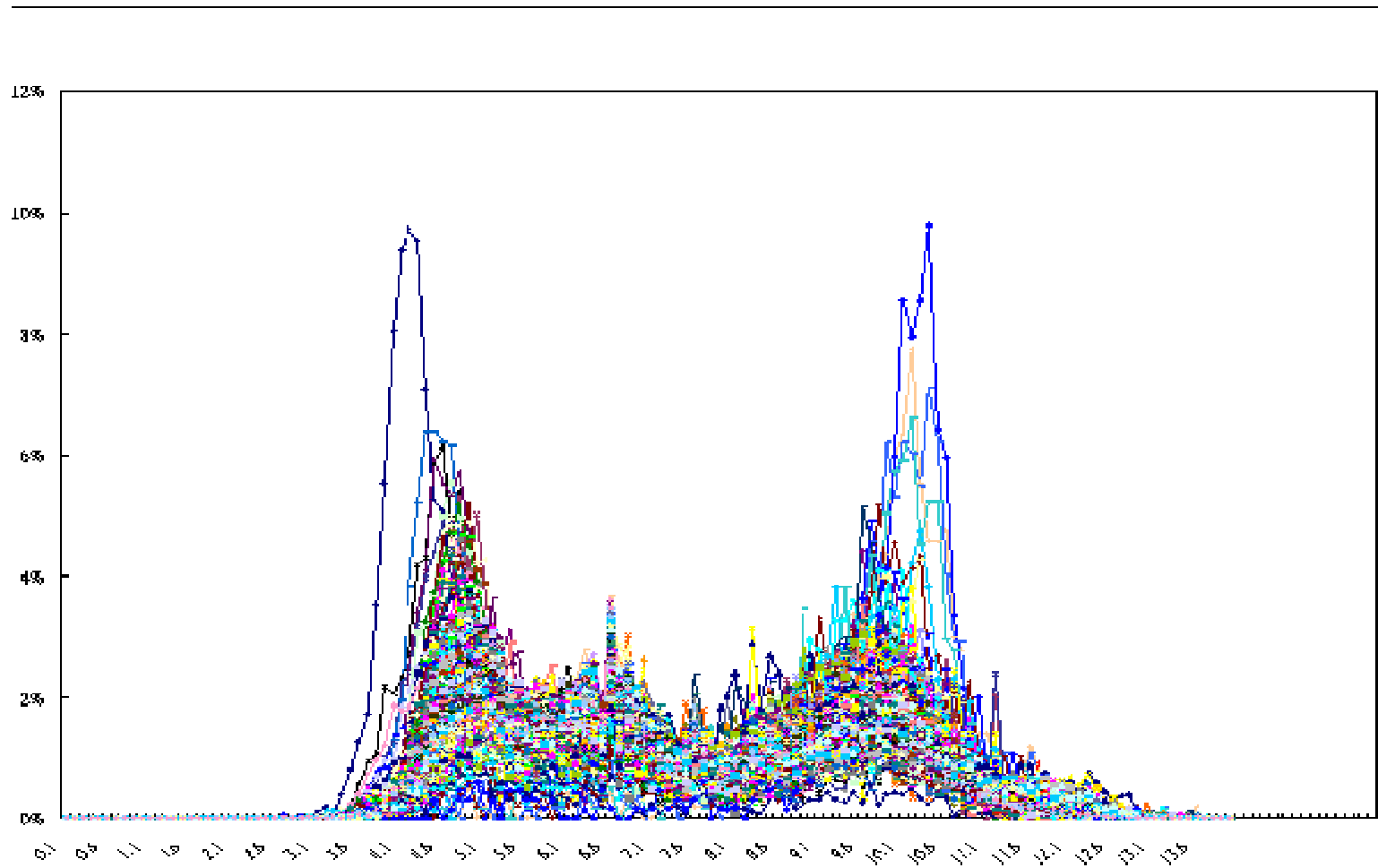


Different species, different pI distribution





pI distribution Over 100 Species





How to compare ?

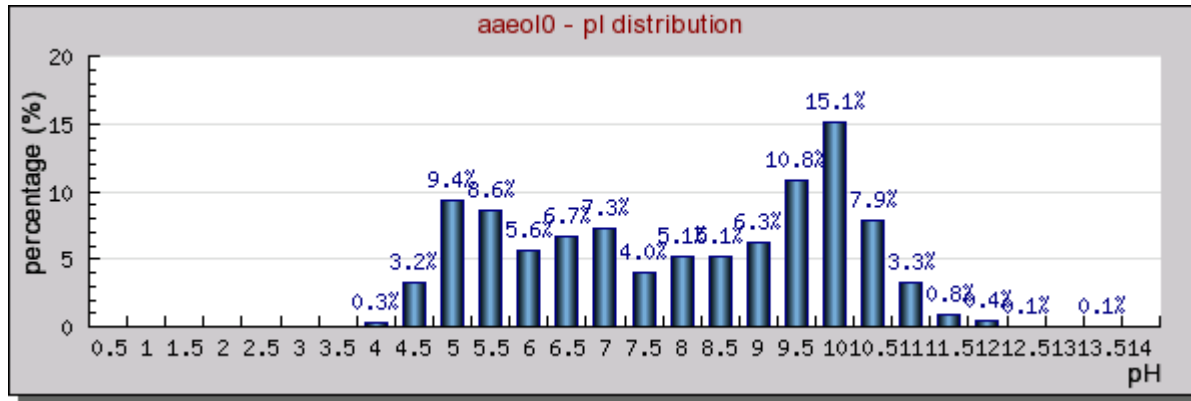
- Interactive on-line analysis to help us to explore the different combination.
- Easy-reading.





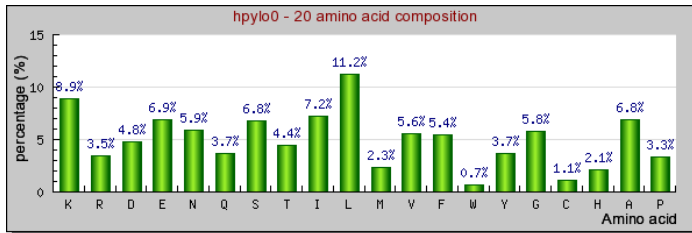
Transform

[Bacteria] - Aquifex aeolicus VF5





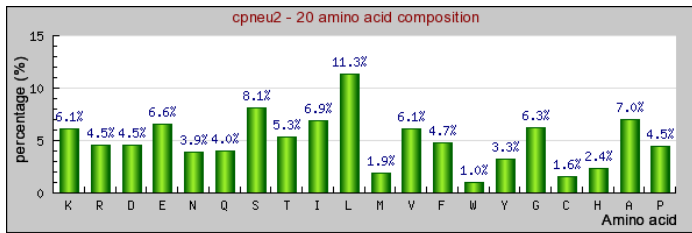
Transform



A C D W L E G I - K L M N Z P Q R S T V Y



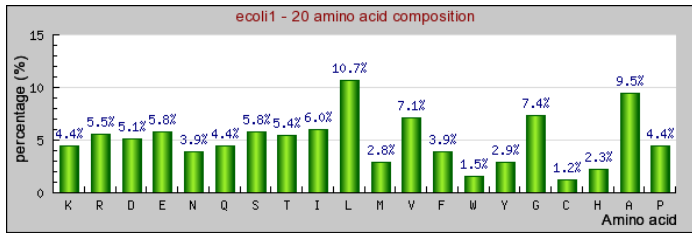
Escherichia coli K12



A C D W L E G I - K L M N Z P Q R S T V Y



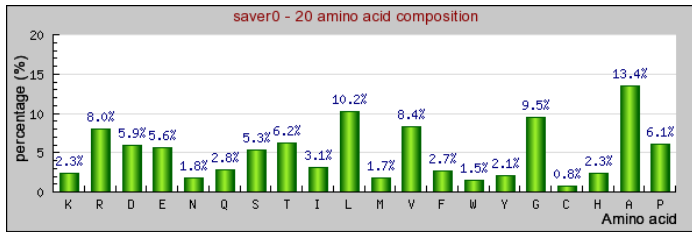
Chlamydomophila pneumoniae J138



A C D W L E G I - K L M N Z P Q R S T V Y



Helicobacter pylori 26695



A C D W L E G I - K L M N Z P Q R S T V Y



Streptomyces avermitilis MA-4680

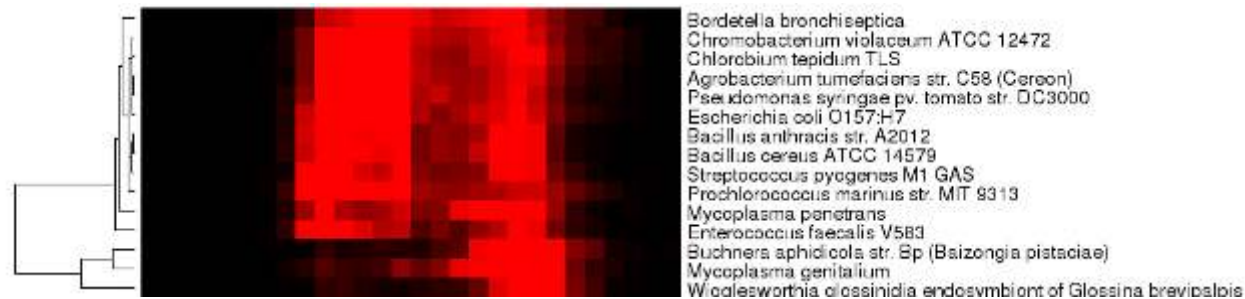


Clustering





Clustering



- There are many clustering methods.
- We use Euclidean distances for Hierarchical Clustering.
- It just a easy way to read, not the only solution!



On-line compare



Choose your interesting field

GPDB Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Compare

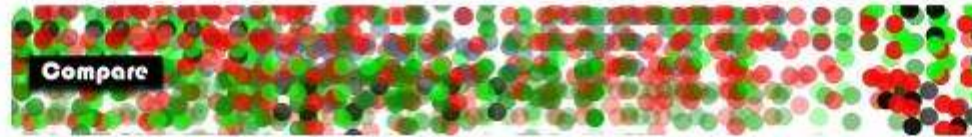
Step 1: Choose your interesting field.
Step 2: Choose compare profile and pick up organisms for compare.

tip: You can press ctrl(Windows/Linux) and command(Mac) to perform multiple choose.

Bacteria/Archaea B A <input type="checkbox"/> skip this field	Shape Coccobacillus Coccus Rod Filament Irregular coccus Sphere Rods Irregular sphere Curved Spiral <input type="checkbox"/> skip this field	Endospores No Yes <input type="checkbox"/> skip this field	Motility No Yes <input type="checkbox"/> skip this field	Salinity Non-halophilic Mesophilic Moderate halophilic Extreme halophilic <input checked="" type="checkbox"/> skip this field	Oxygen requirement Aerobic Anaerobic Facultative Microaerophilic <input checked="" type="checkbox"/> skip this field
---	--	--	--	---	--

Best resolution above 1280*1024 | Copyright ©2006 PCLy's Lab, Institute of Bioinformatics and Structural Biology, NTHU.
Maintained by Szu-Ming Lai and Chi-ching Lee

Choose Profile



Step 1: Choose your interesting field.
 Step 2: Choose compare profile and pick up organisms for compare.

Which profile do you want to compare?

- AT & GC Content
- Nucleotide composition
- Amino acid composition
- Length distribution
- Total codon usage comparison
- di-nucleotide composition
- di-peptide composition
- Molecular Weight distribution
- Isoelectric point distribution
- tri-Nucleofide composition

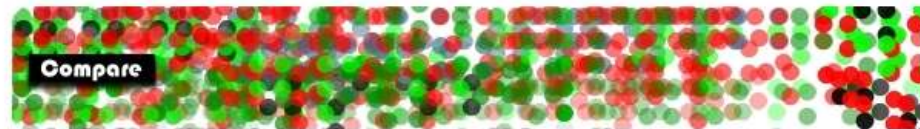
Which organisms do you want to compare?

tip: You can click title of the table to sort the organism list.

<input type="checkbox"/>	RefSeq ID	Organism	King	Shape	Endospores	Motility	Salinity	OxygenReq
<input type="checkbox"/>	NC_007963	Chromohalobacter salexigens DSM 3043	B	Rod	No	Yes	Moderate halophilic	Aerobic
<input type="checkbox"/>	NC_004557	Clostridium tetani E88	B	Rod	No	Yes	Non-halophilic	Anaerobic
<input type="checkbox"/>	NC_007298	Dechloromonas aromatica RCB	B	Rod	No	Yes		Facultative
<input type="checkbox"/>	NC_007722	Erythrobacter litoralis HTCC2594	B	Rod	No	Yes		Aerobic
<input type="checkbox"/>	NC_008228	Pseudoalteromonas atlantica T6c	B	Rod	No	Yes		Aerobic
<input type="checkbox"/>	NC_007613	Shigella boydii Sb227	B	Rod	No	Yes	Non-halophilic	Facultative
<input type="checkbox"/>	NC_007606	Shigella dysenteriae Sd197	B	Rod	No	Yes	Non-halophilic	Facultative
<input type="checkbox"/>	NC_007384	Shigella sonnei Ss046	B	Rod	No	Yes	Non-halophilic	Facultative
<input type="checkbox"/>	NC_000853	Thermotoga maritima MSB8	B	Rod	No	Yes		Anaerobic
<input type="checkbox"/>	NC_007404	Thiobacillus denitrificans ATCC 25259	B	Rod	No	Yes		Facultative



Choose Organisms



Step 1: Choose your interesting field.
 Step 2: Choose compare profile and pick up organisms for compare.

Which profile do you want to compare?

- AT & GC Content
- Nucleotide composition
- Amino acid composition
- Length distribution
- Total codon usage comparison
- di-nucleotide composition
- di-peptide composition
- Molecular Weight distribution
- tri-Nucleotide composition
- Isoelectric point distribution

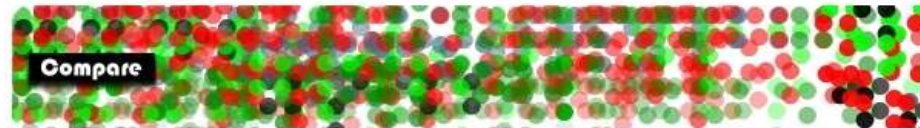
Which organisms do you want to compare?

tip: You can click title of the table to sort the organism list.

	RefSeq ID	Organism	King	Shape	Endospores	Motility	Salinity	OxygenReq
<input checked="" type="checkbox"/>	NC_007963	Chromohalobacter salexigens DSM 3043	B	Rod	No	Yes	Moderate halophilic	Aerobic
<input checked="" type="checkbox"/>	NC_004557	Clostridium tetani E88	B	Rod	No	Yes	Non-halophilic	Anaerobic
<input checked="" type="checkbox"/>	NC_007298	Dechloromonas aromatica RCB	B	Rod	No	Yes		Facultative
<input checked="" type="checkbox"/>	NC_007722	Erythrobacter litoralis HTCC2594	B	Rod	No	Yes		Aerobic
<input type="checkbox"/>	NC_008228	Pseudoalteromonas atlantica T6c	B	Rod	No	Yes		Aerobic
<input checked="" type="checkbox"/>	NC_007613	Shigella boydii Sb227	B	Rod	No	Yes	Non-halophilic	Facultative
<input checked="" type="checkbox"/>	NC_007606	Shigella dysenteriae Sd197	B	Rod	No	Yes	Non-halophilic	Facultative
<input type="checkbox"/>	NC_007384	Shigella sonnei Ss046	B	Rod	No	Yes	Non-halophilic	Facultative
<input checked="" type="checkbox"/>	NC_000853	Thermotoga maritima MSB8	B	Rod	No	Yes		Anaerobic
<input checked="" type="checkbox"/>	NC_007404	Thiobacillus denitrificans ATCC 25259	B	Rod	No	Yes		Facultative



Submit



Step 1: Choose your interesting field.
 Step 2: Choose compare profile and pick up organisms for compare.

Which profile do you want to compare?

- AT & GC Content
- Nucleotide composition
- Amino acid composition
- Length distribution
- Total codon usage comparison
- di-nucleotide composition
- di-peptide composition
- Molecular Weight distribution
- tri-Nucleotide composition
- Isoelectric point distribution

Which organisms do you want to compare?

tip: You can click title of the table to sort the organism list.

	RefSeq ID	Organism	King	Shape	Endospores	Motility	Salinity	OxygenReq
<input checked="" type="checkbox"/>	NC_007963	Chromohalobacter salexigens DSM 3043	B	Rod	No	Yes	Moderate halophilic	Aerobic
<input checked="" type="checkbox"/>	NC_004557	Clostridium tetani E88	B	Rod	No	Yes	Non-halophilic	Anaerobic
<input checked="" type="checkbox"/>	NC_007298	Dechloromonas aromatica RCB	B	Rod	No	Yes		Facultative
<input checked="" type="checkbox"/>	NC_007722	Erythrobacter litoralis HTCC2594	B	Rod	No	Yes		Aerobic
<input type="checkbox"/>	NC_008228	Pseudoalteromonas atlantica T6c	B	Rod	No	Yes		Aerobic
<input checked="" type="checkbox"/>	NC_007613	Shigella boydii Sb227	B	Rod	No	Yes	Non-halophilic	Facultative
<input checked="" type="checkbox"/>	NC_007606	Shigella dysenteriae Sd197	B	Rod	No	Yes	Non-halophilic	Facultative
<input type="checkbox"/>	NC_007384	Shigella sonnei Ss046	B	Rod	No	Yes	Non-halophilic	Facultative
<input checked="" type="checkbox"/>	NC_000853	Thermotoga maritima MSB8	B	Rod	No	Yes		Anaerobic
<input checked="" type="checkbox"/>	NC_007404	Thiobacillus denitrificans ATCC 25259	B	Rod	No	Yes		Facultative



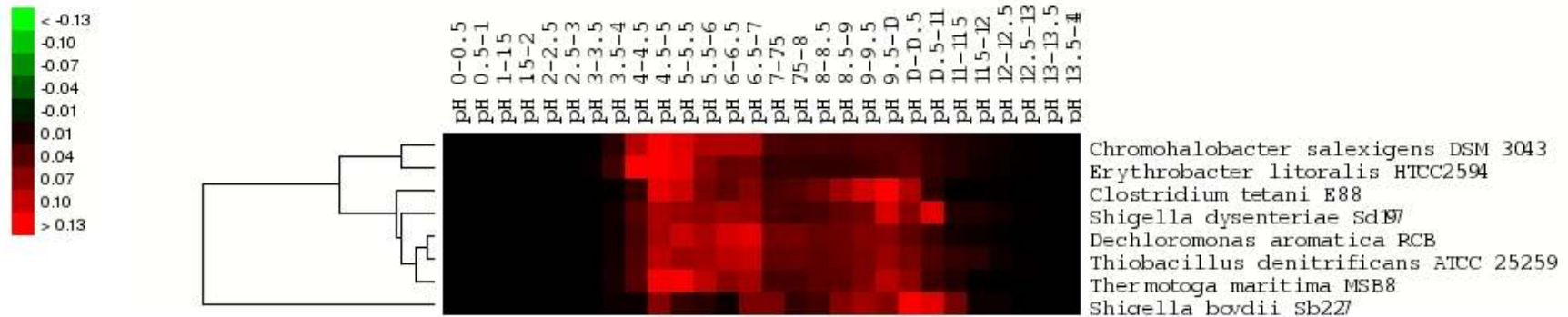
On-line Clustering



Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Genome Profile - Isoelectric point(pI) Distribution



Best resolution above 1280*1024 | Copyright ©2008 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Szu-Ming Lai and Chi-ching Lee



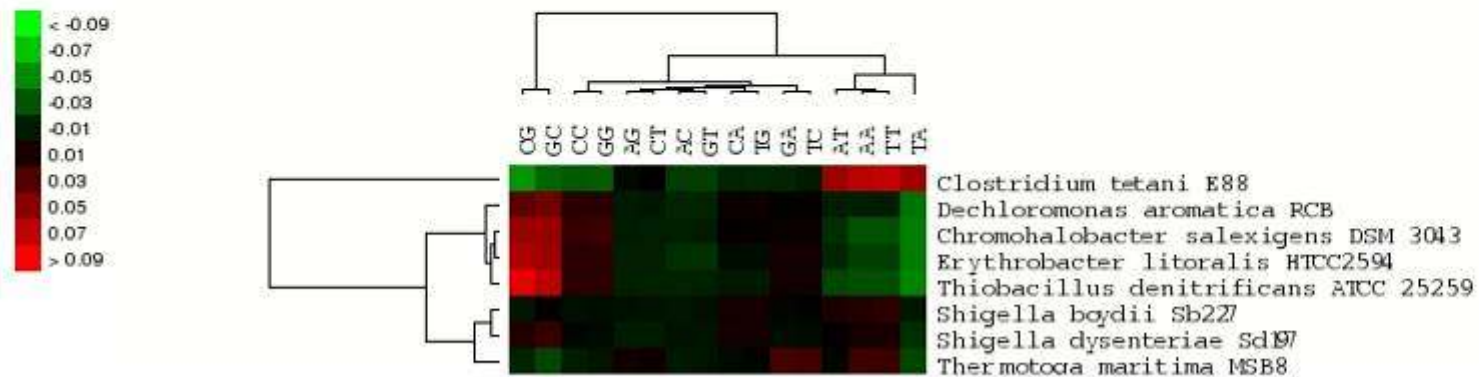
Di-nucleotide composition



Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Genome Profile - di-Nucleotide Composition



Best resolution above 1280*1024 | Copyright ©2006 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Szu-Ming Lai and Chi-ching Lee



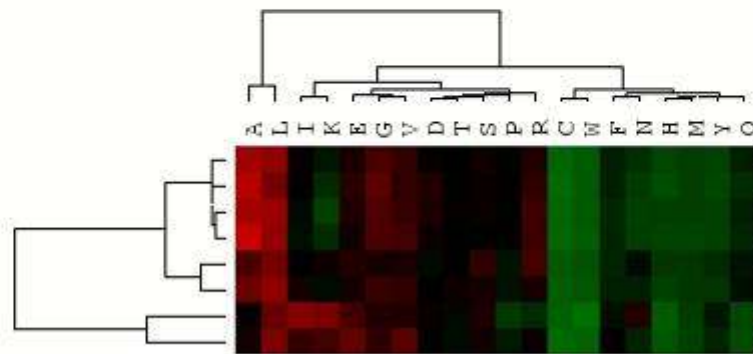
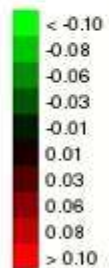
Amino acid composition



Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Genome Profile - Amino acid Composition



Dechloromonas aromatica RCB
Erythrobacter litoralis HTCC2594
Chromohalobacter salexigens DSM 3043
Thiobacillus denitrificans ATCC 25259
Shigella boydii Sb227
Shigella dysenteriae Sd197
Clostridium tetani E88
Thermotoga maritima MSB8

Best resolution above 1280*1024 | Copyright ©2006 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Su-Ming Lai and Chi-ching Lee



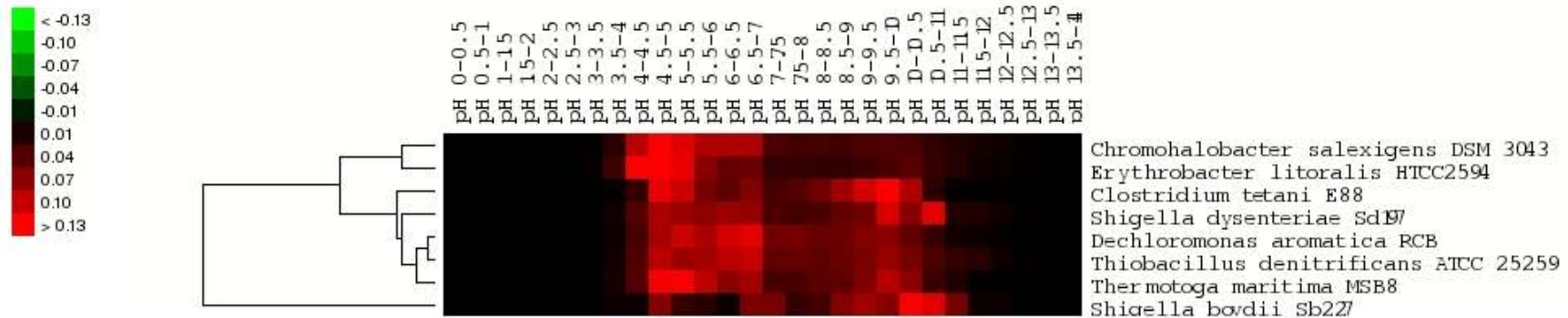
Isoelectric point distribution



Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Genome Profile - Isoelectric point(pI) Distribution



Best resolution above 1280*1024 | Copyright ©2008 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
 Maintained by Szu-Ming Lai and Chi-ching Lee



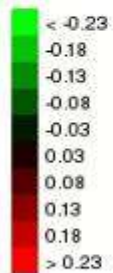
Molecular Weight distribution



Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Genome Profile - Molecular Weight Distribution



Best resolution above 1280*1024 | Copyright ©2006 PCLyu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Szu-Ming Lai and Chi-ching Lee



You can compare on profile in different combination or on the whole.

GPDB Genome Profile DataBase

Home | Browse | Virtual 2D | Compare | Download | Status | Help | Comment

Compare

Step 1: Choose your interesting field.
Step 2: Choose compare profile and pick up organisms for compare.

tip: You can press ctrl(Windows/Linux) and command(Mac) to perform multiple choose.

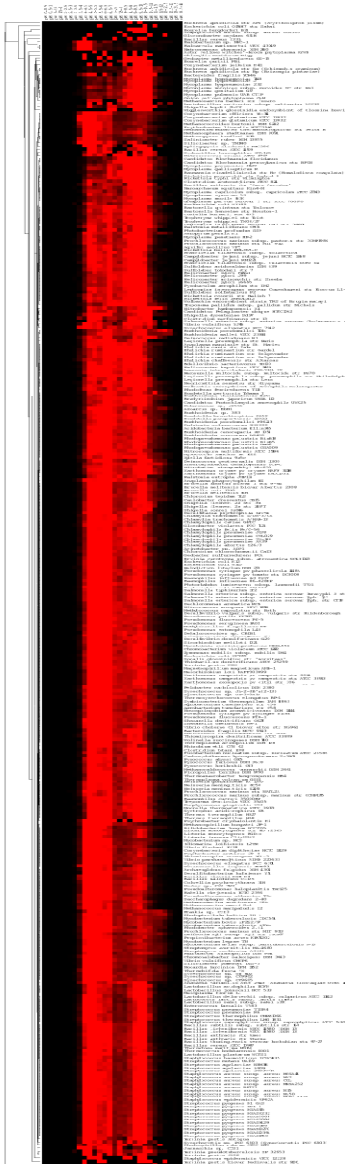
Bacteria/Archaea B A	Shape Coccobacillus Coccus Rod Filament Irregular coccus Sphere Rods Irregular sphere Curved Spiral	Endospores No Yes	Motility No Yes	Salinity Non-halophilic Mesophilic Moderate halophilic Extreme halophilic	Oxygen requirement Aerobic Anaerobic Facultative Microaerophilic
<input checked="" type="checkbox"/> skip this field	<input checked="" type="checkbox"/> skip this field	<input checked="" type="checkbox"/> skip this field	<input checked="" type="checkbox"/> skip this field	<input checked="" type="checkbox"/> skip this field	<input checked="" type="checkbox"/> skip this field

Go to step 2

Best resolution above 1280*1024 | Copyright ©2006 PCLYu's Lab, Institute of Bioinformatics and Structural Biology, NTHU
Maintained by Su-Ming Lai and Chi-ching Lee

Compare 345 organisms (27 Archaea, 338 Bacteria)

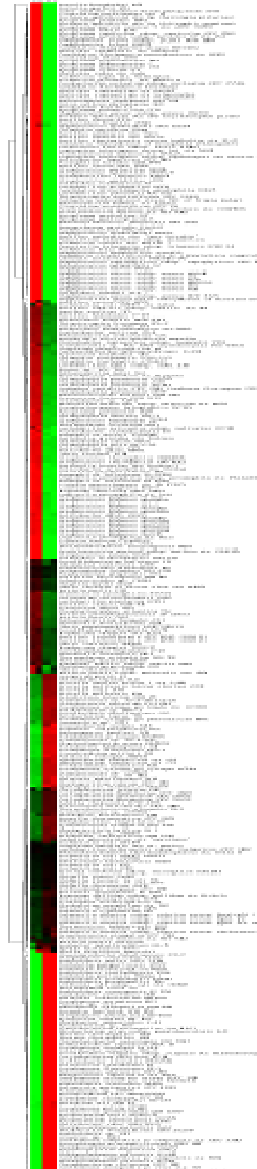
Isoelectric point(pI) Distribution



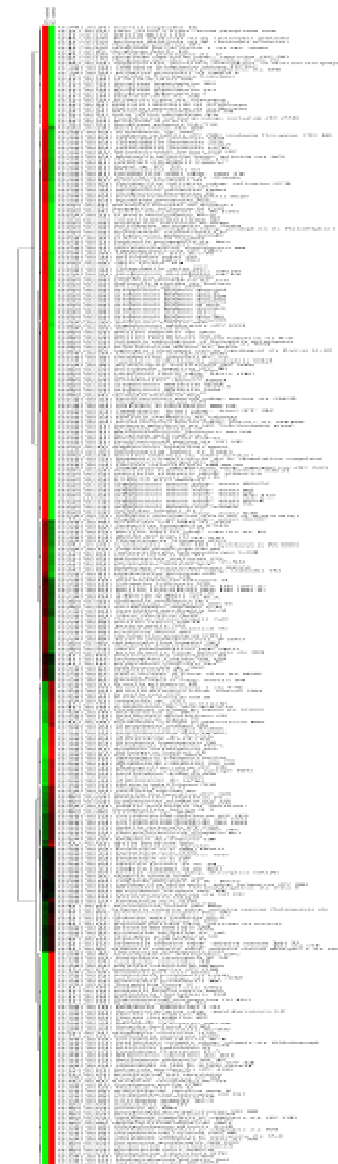
Molecular Weight Distribution



Nucleotide Composition

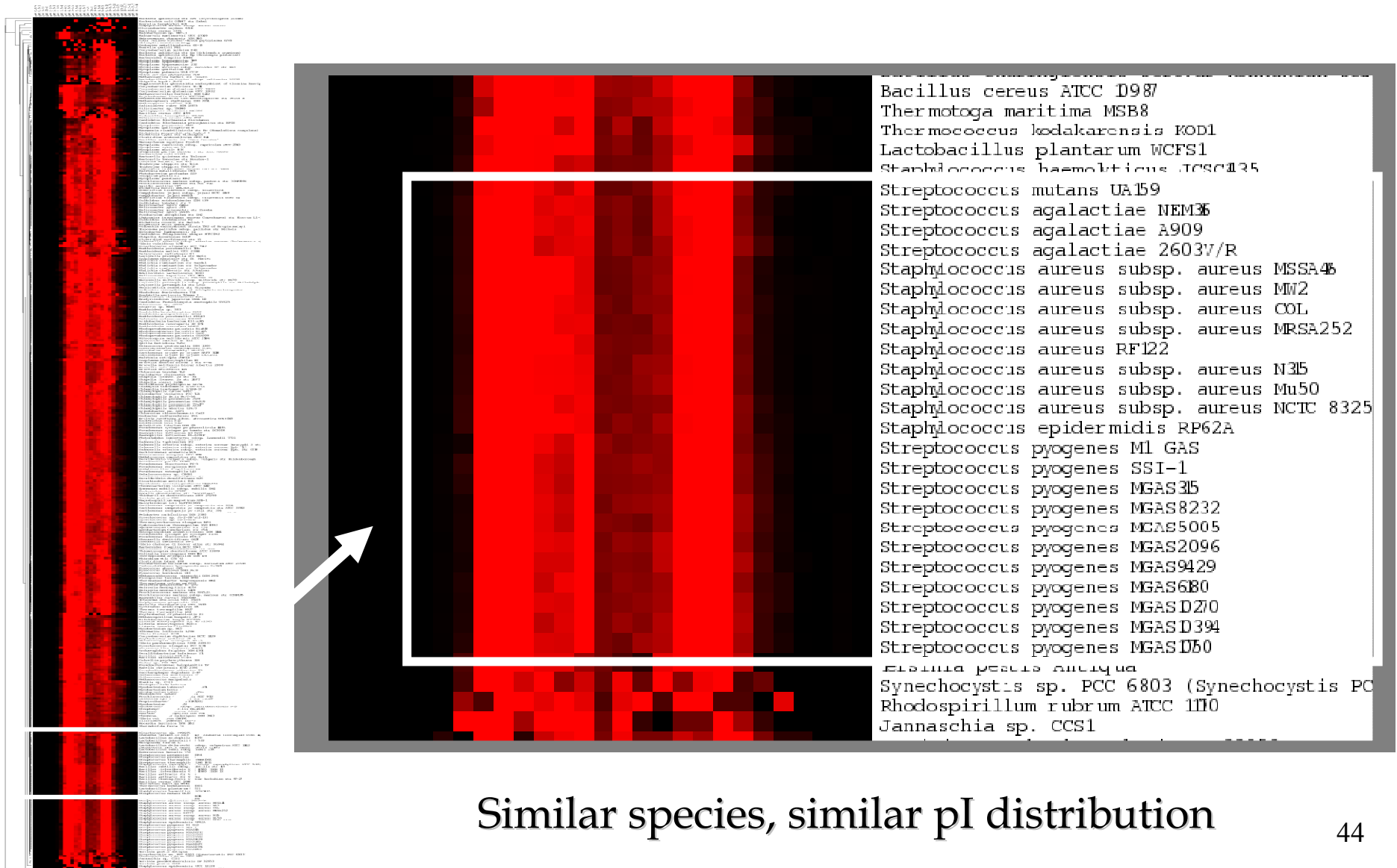


AT & GC Content



Compare 345 organisms (27 Archaea, 338 Bacteria)

Isoelectric point(pI) Distribution



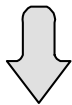
Similar species, similar pI distribution

Virtual 2D Gel Flowchart



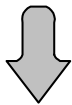
Ref Sequence
from NCBI

- Accession: NC_003902
- ftp://ftp.ncbi.nih.gov/genomes/Bacteria/Xanthomonas_campestris/NC_003902.faa



Calculate pI and MW

- EMBOSS package – pepstat
- Available at <http://www.hgmp.mrc.ac.uk/Software/EMBOSS/>

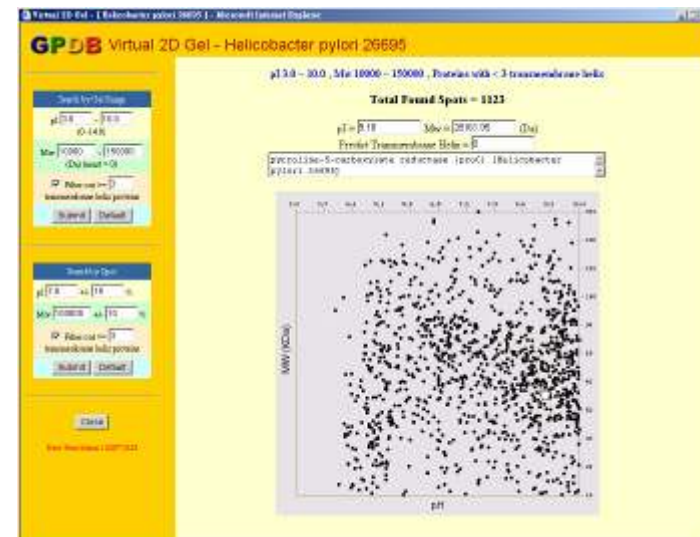


Protein records



MySQL

PHP / GD Library

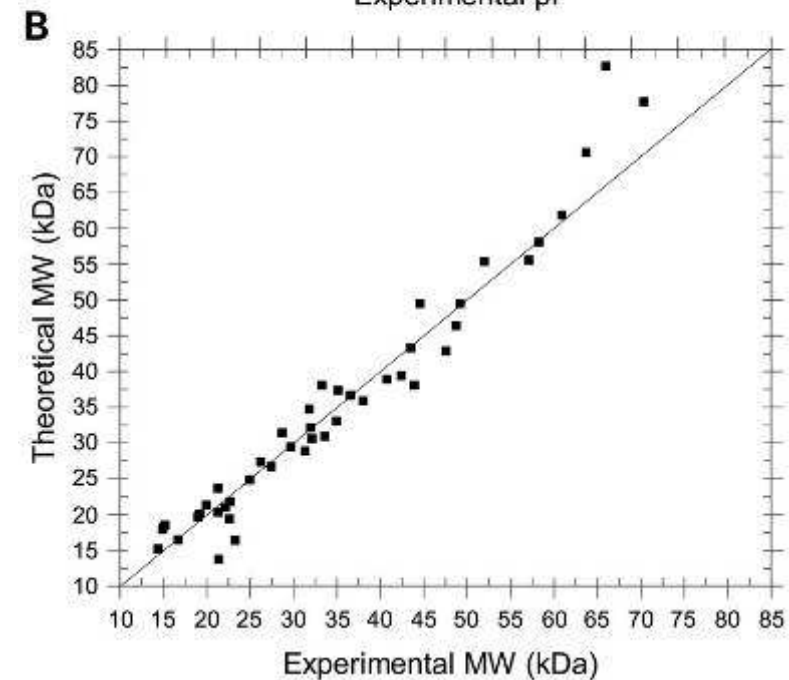
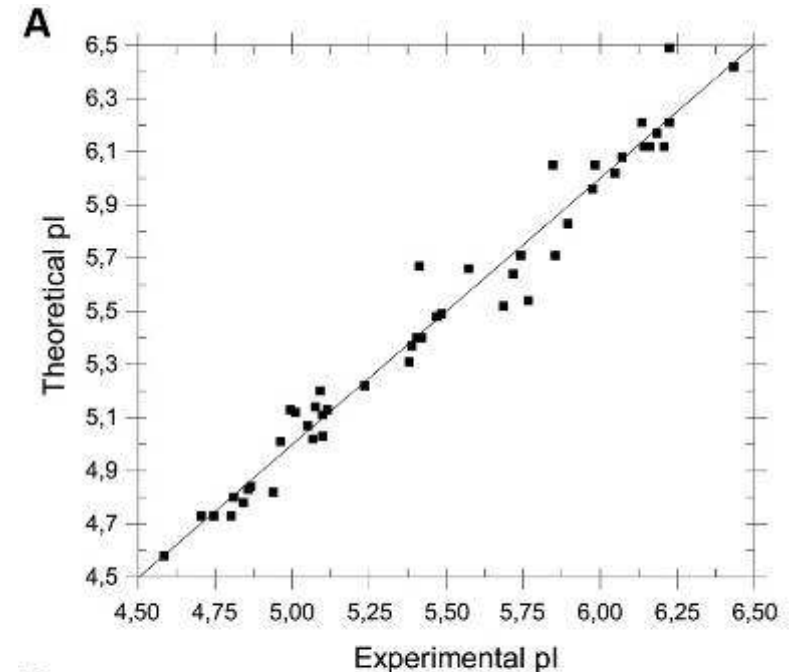




pI and MW Theoretical vs. Experimental

Comparison between calculated and experimentally obtained pI and MW values for 47 randomly selected proteins from *Pseudomonas aeruginosa*.

Nucleic Acids Research, 2003, Vol. 31, No. 13 3862-3865





Virtual 2D Gel

- Search by different pH & mw range to simulate the real 2D gel.
- Search by spot range to guess the possible spot, ex:
 - $pI = 5.7 \pm 5\%$
 - $M_w = 100K \pm 5\%$
- Filter out transmembrane proteins.
 - TMHMM program
 - E. L.L. Sonnhammer, G. von Heijne, and A. Krogh. *In J. Glasgow et al., eds., Proc. Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, 175-182. AAAI Press, 1998.
 - A. Krogh, B. Larsson, G. von Heijne, and E. L. L. Sonnhammer. *Journal of Molecular Biology*, 305(3):567-580, January 2001.
 - Moller S, Croning MD, Apweiler R. *Bioinformatics* 2002 Jan;18(1):218



Search by Gel Range without TMHMM Filter

Virtual 2D Gel - [Helicobacter pylori 26695] - Microsoft Internet Explorer

GPDB Virtual 2D Gel - Helicobacter pylori 26695

Total Found Spots = 1267

pI = 3.0 ~ 10.0 (0-14.0)
Mw 10000 ~ 150000 (Da) (must > 0)

Filter out >= 3 transmembrane helix proteins

Submit Default

pI 3-10
Mw 10000 - 15000

Search by Spot

pI 7.0 +/- 10 %
Mw 100000 +/- 10 %

Filter out >= 3 transmembrane helix proteins

Submit Default

Close

Best Resolution 1280*1024

Predict Transmembrane Helix = 0 **Transmembrane Helix**

glutamyl-tRNA reductase (hemA) [Helicobacter pylori 26695]

Annotation

Search by Gel Range with TMHMM Filter



The screenshot displays the GPDB Virtual 2D Gel interface for *Helicobacter pylori* 26695. The main window shows a search for protein E (gcpE) with 1123 found spots. The search parameters include a molecular weight (Mw) of 42486.44 Da and a membrane helix count of 0. The search results list the protein as NP_207419, protein E (gcpE) [Helicobacter pylori 26695]. The 2D gel plot shows a distribution of spots with a pH range from 5.8 to 10.0 and a molecular weight range from 10 to 150 kDa. A red circle highlights a specific spot on the gel at approximately pH 9.3 and 80 kDa. The interface also includes a sidebar with search filters for pI (3.0), Mw (100), and a checkbox for 'Filter transmembrane proteins' which is checked. The NCBI logo and search options are visible at the top of the search window.



Search by Spot Range without TMHMM Filter

Virtual 2D Gel - [Helicobacter pylori 26695] - Microsoft Internet Explorer

GPDB Virtual 2D Gel - Helicobacter pylori 26695

Total Found Spots = 11

pI = Mw = (Da)
Predict Transmembrane Helix =
cell division protein (ftsK) [Helicobacter pylori 26695]

pH	MW (KDa)
6.3	95
6.4	93
6.6	106
6.6	102
6.9	97
7.0	96.5
7.0	95
7.3	106
7.3	97
7.6	105
7.6	104

Search by Gel Range

pI ~
(0 -14.0)

Mw ~
(Da) (must > 0)

Filter out >=
transmembrane helix proteins

Search by Spot

pI +/- %

Mw +/- %

Filter out >=
transmembrane helix proteins

Best Resolution 1280*1024

Search by Spot Range with TMHMM Filter



Virtual 2D Gel - [*Helicobacter pylori* 26695] - Microsoft Internet Explorer

GPDB Virtual 2D Gel - *Helicobacter pylori* 26695

Total Found Spots = 10

pI = Mw = (Da)
Predict Transmembrane Helix =

Search by Gel Range

pI 3.0 ~ 10.0
(0 -14.0)

Mw 10000 ~ 150000
(Da) (must > 0)

Filter out >= 3
transmembrane helix proteins

Search by Spot

pI 7 +/- 10 %

Mw 100000 +/- 10 %

Filter out >= 3
transmembrane helix proteins

Best Resolution 1280*1024

pH	MW (KDa)
6.3	95.5
6.4	93.5
6.6	106.5
6.6	102.5
6.7	97.5
7.0	95.5
7.3	106.5
7.3	97.5
7.6	105.5
7.6	104.5



GPDB Current Status

- **550 organisms**
(27 Archaea, 338 Bacteria, 200 Virus, 5 Fungi)
- **880 complete sequence**
- **Total - 1,176,013 protein (ORFs)**



Conclusions

- We construct the database (GPDB), which provides many whole-genome scale features.
- GPDB can let you explore the relationship between different species using genome-wide profiles.
- It can only compare one genome profile per time now.
- We hope it can compare multiple genome profiles per time in the future.
- We hope this can help to figure out some rules.

Summary

