



GELBANK

A database of annotated 2-dimensional gel electrophoresis patterns of biological systems with completed genomes

By Alexander Nezhinsky & Thomas Steenbergen

Introduction

- Gelbank is database with 2-dimensional gel electrophoresis (2DE) of proteomes from organisms with known genome information.
- The website contains 233 identifications for 81 gel patterns for Homo Sapiens e.a.

What are proteomes?
What is a 2-dimensional gel electrophoresis?

Proteomes

The term "proteome" has been used to refer to the collection of proteins found in a particular cell type under a particular set of environmental conditions . (Wikipedia.org)

The proteome of an organism is the collection of potential open reading frames (ORFs)

Some Theory, ORFs

By examining the DNA sequence alone we can determine the sequence of amino acids that will appear in the final protein.

In translation codons of three nucleotides determine which amino acid will be added next in the growing protein chain. But you will need to decide on which nucleotide to start translation, and when to stop, this is called an *open reading frame*.

Example ORF

```
atgcccaagctgaatagcgtagaggggtttcatcattgaggacgatgataa
```

```
1 atg ccc aag ctg aat agc gta gag ggg ttt tca tca ttt gag gac gat gta taa
  M P K L N S V E G F S S F E D D V *
2 tgc cca agc tga ata gcg tag agg ggt ttt cat cat ttg agg acg atg tat
  C P S * I A * R G F H H L R T M Y
3 gcc caa gct gaa tag cgt aga ggg gtt ttc atc att tga gga cga tgt ata
  A Q A E * R R G V F I I * G R C I
```

ATG is a start condon,
TAA, TGA & TAG are stop condons

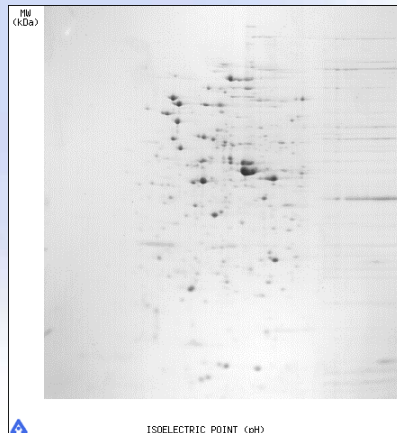
Proteome Analysis

Proteome Analysis requires the isolation of the complete proteome, separation of complex protein mixtures into discrete protein components measurement of relative abundance and identification of each protein Component.

The most widely used separation method is:

2-dimensional gel electrophoresis

What is 2D Gel Electrophoresis?



Back to 2DE

2DE is a separation method for proteome-analysis. Proteins are separated according to their isoelectric piont (pI) and their molecular weight (MW).

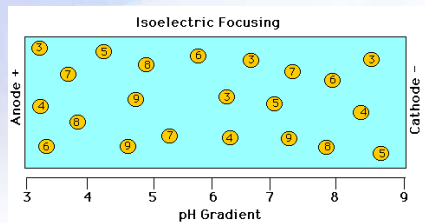
It's a combination of 2 techniques:

- Isoelectric focusing
- SDS Polyacrylamide Gel Electrophoresis

Isoelectric focusing

Because of the amino acids in proteins, they have amphoteric properties and will be positively charged at pH values below their pI and negatively charged above.

This means that proteins will migrate toward their pI. Most proteins have a pI in the range of 5 to 8.5.

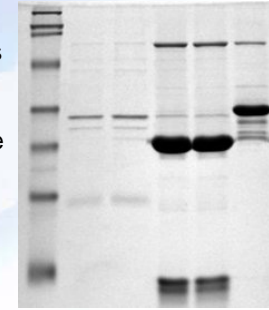


SDS Polyacrylamide Gel Electrophoresis

The proteins are inserted into a (flat) gel and an electrical current is applied.

Due to their linear form and negative charge, the proteins move through the gel.

Large proteins will move slowly through the gel, constrained by bulk, while small proteins will slip from pore to pore easily and move quickly.



Thus proteins may be separated roughly according to size (molecular weight).

Proteome Analysis

Then we can identify our proteins with peptide mass fingerprinting by comparing their peptide masses of our proteins with the predicted peptide masses.

Gelbank can now answer 2 questions:

Given a genome and potential ORF information, what is the most likely identity of a protein at a given location on a gel?

What are the predicted proteins resolved in a given pI and MW range?

Gelbank features

1. Proteomes: collection of interfaces for proteome queries.
2. 2D Gels: interfaces for 2D gel queries.
3. Tools: tools for proteome analysis by 2DE.
4. Upload: registered users can manipulate their 2DE patterns.
5. Bio-bag: a "shopping" basket for easy storage of items from the website.
6. FTP site: for access of data in various formats

Proteomes

- Search by description

Protein Search (description) 203 records (1-10)

Query: tyrosine kinase src

Species: Homo sapiens, Helicobacter pylori, Helicobacter pylori, Helicobacter pylori, Homo sapiens

GI	Description	Species	MW (kDa)	pI
4303823	protein-tyrosine kinase fyn isoform a; proto-oncogene tyrosine-protein kinase fyn; src/yes-related novel gene; src-like kinase; c-syn proto-oncogene; tyrosine kinase p59fyn(T); OKT3-induced calcium influx regulator	Homo sapiens (genome)	60.762	6.229
4788078	c-src tyrosine kinase	Homo sapiens (genome)	58.704	4.431
	proto-oncogene tyrosine-protein kinase SRC; proto-oncogene SRC; Retv proto-oncogene; c-syn;avian sarcoma (Schmid-Sharp A-2) GAT AAGAAAGAA TAAAGAT; Retv;src;src-kinase; SRC-1	Homo sapiens	59.935	7.109

Proteomes

- Search by sequence fragment

Protein Search (sequence fragment) 111 records (1-10)

Peptide sequence: TALC

Species: Agrobacterium sp. (ADP1), Aeropyrum pernix, Agrobacterium tumefaciens (C58), Agrobacterium tumefaciens (C58 (U. Washington))

GI	Description	Species	MW (kDa)	pI
4503681				
4503683				
4503965				
4503987				
4504611				
4505245				
4506131				
4506597				
4506871				
4507197				

Proteomes

- Search by mW and pI

Protein Search (MW and pI) 20 records (1-20)

Molecular weight (Daltons): min 100000, max 300000, display 1kDa

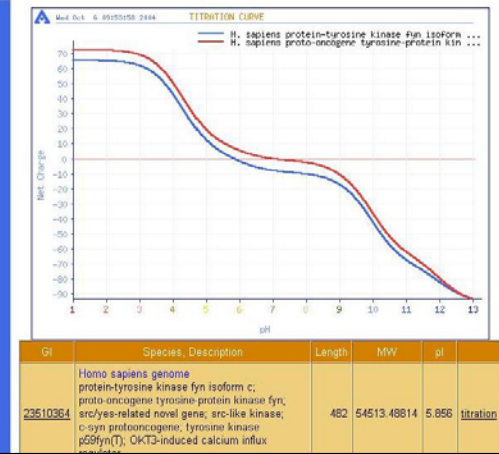
isoelectric point: min 3, max 12

Species: Agrobacterium sp. (ADP1), Aeropyrum pernix, Agrobacterium tumefaciens (C58), Agrobacterium tumefaciens (C58 (U. Washington)), Apis mellifera (DH4)

GI	Description	Species	MW (kDa)	pI
20083217	isoacyl-tRNA synthetase	Agrobacterium sp. (ADP1) (chromosome)	106	5.433
20083452	conserved hypothetical protein; putative VGR-related protein	Agrobacterium sp. (ADP1) (chromosome)	123	5.252
20083579	DNA-directed RNA polymerase beta chain (Transcriptase beta chain) (RNA polymerase beta subunit)	Agrobacterium sp. (ADP1) (chromosome)	155	5.236
20083580	DNA-directed RNA polymerase beta' chain (Transcriptase beta' chain) (RNA polymerase beta' subunit)	Agrobacterium sp. (ADP1) (chromosome)	157	7.593

Use of tools

- ClustalW
- Titration



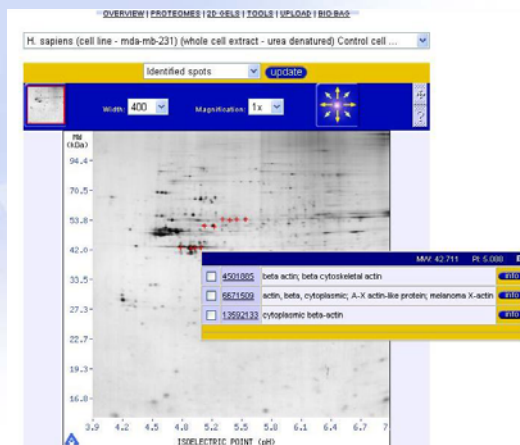
2DE gels

2DE gels-Simple search

GEL	Description
	<p><i>Homo sapiens</i> cell line - md7 whole cell extract - urea denatured Control cells (ganarstein treatment) Staining: silver IEF: carrier ampholyte (4-7) DALY: 10-176 ANL Protein Mapping Group <gababnica@anl.gov> (arsteomevab.anl.gov) 03/07/03</p>
	<p><i>Homo sapiens</i> cell line - hah293 ptyr column eluate Control Staining: silver IEF: carrier ampholyte (50/50 5-7/3-10) DALY: 10-176 ANL Protein Mapping Group <gababnica@anl.gov> (arsteomevab.anl.gov) 03/28/02</p>
	<p><i>Homo sapiens</i> cell line - hah293 ptyr column eluate Thapsigargin treated Staining: silver IEF: carrier ampholyte (50/50 5-7/3-10) DALY: 10-176</p>

2DE gels-Simple search

- Profiled Gel
- ClustaW
- Titration
- Size till 1500
- animation



2DE gels – Complex search

- Species
- Tissue
- Sample type
- Staining
- First Dimension
- PH range
- Second Dimension

Tools



Bio-Bag

- Storage of object for later use
- Deleted after 30 min of inactivity
- Tools: multiple sequence alignment, titration curves ORFS, animation, etc.
- Contents can be deleted

Tools

- Selecting and saving patterns
- ClustalW (alignment)
- Titration curve
- Images animation
- Selecting sub-region
- Saving animations to BioBag

Upload

- PNG (8-bit), 1,4Mb max
- Profiling (markers, assigning MW pl values)
- ORACLE9i
- Assigning ORF's to a protein spot
- Adding new properties

FTP-site

- Entire content website
- Tab-delimited file
- Flat text format
- ORACLE9i database dumps
- 2DE patterns
- Weekly updated

References

- http://www.aber.ac.uk/parasitology/Proteome/Tut_2D.html#Section%201
- http://bioweb.uwlax.edu/GenWeb/Molecular/Seq_Anal/Translation/translation.html
- <http://www.kennislink.nl/web/show?id=89064&showframe=content&vensterid=70344&prev=89059>
- <http://http://gelbank.anl.gov/>
- www.ipn.uni-kiel.de/eibe/UNIT14NL.PDF
- <http://en.wikipedia.org>