



Christine Hoogland Swiss Institute of Bioinformatics

Proteo E Infor atics group



- Ron D. Appel
- Pierre-Alain Binz
- Gérard Bouchet
- Sébastien Catherinet
- Robin Gras
- Céline Hernandez
- David Hernandez
- Patricia Hernandez
- Christine Hoogland

- Frédérique Lisacek
- Khaled Mostaguir
- Markus Müller
- Patricia Palagi
- Stéphane Pelhatre
- Marc Tuloup
- Daniel Walther
- Nadine Zangger



Proteomics pathway





LIMS

- Close collaboration with BPRG
- Applied Biosystems
 - SQL*LIMS, v4.0 (sample plans, projects, samples)
 - RIS (Proteomics module: Melanie image logging for 1-DE, 2-DE, spot logging, MS peaklists uploads, identification launches, validation of protein identification, ELISA, immunoblots)
 - SQL GT (plate management)



Logical Principle

- Projects/submissions
 contain
- Samples
- to which one can assess
- Tasks/Methods
- that produce
- Results

🗜 S.	aureus	Attributes	: Ge	el Image	MS Data 8	Protein	s						
•- 🎳 •	N315 N315/Total extract	4	[]> î	i B	V	ersion:	2/2	Statu	IS: ON_GOING_VALIDATION USER	: PENLIMS	Date	: <mark>11/26/</mark> :	2003
🌳 🎳 N315/Total membrane extracts		Rank	Accession No.		Mass	lass Score		Protein Name			Masses	M V	/alidate
🗢 🔟 CENTRIFUGE METHOD		1		Q895Z	1 46.456	i 18			(Q895Z1) Cell division	protein ftsA.	4	/ 20	
	🌳 🎳 N315/Membrane Soluble	2		Q8P5	6 25.896	6 10	10 (Q8P5I6) Acetoacetyl-coA reductase.			3	/ 20		
	💁 🕼 CENTRIFUGE METHOD	3		Q92FA	3 63.221	6	(Q92	292FA3) Putative peptidoglycan bound protein (LPXTG motif).			2	/ 20	1
	🕑 攝 N315/Soluble-2D02-0757	4		Q883Y	7 43.193	8 6		(Q883Y7) 3-oxoacyl-(acyl-carrier-protein) synthase I.			3	/ 20	V
	💁 🌄 N315/Soluble 1D Gel Aliquot	5		Q87V2	6 46.976	6 6		(Q87V26) Gluconate transporter family protein.			1	/ 20	
		6		Q8UIG	10.969	9 5		(Q8UIQ7) Hypothetical protein Atu0236.			2	20	- 님
	. CENTRIFUGE METHOD	CENTRIFLIGE METHOD 7 Q10884 53.457 4 (Q10884) Probable formate hydrogeniyas			e CY251.05. 44 0) LIESOU	1	720	-⊢					
				Q5011	9 0.551	3			(@50	119) Ob50y.		.) 20	
	0 八 N215(Mambrana Chlarapharma	Peptide D	ata										
		Chec I	Mass I	Matched N	lass Ca	Delta	St	End	Sequence	Modifica	Miss	Score	Validate
				1.079,47	2.155,11	1,81	2	22	TLRIAYVTSGMGSVGTAICQK	<	1	8	
el View	er de la companya de			1.079,63	2.157,09	0,15	103	121	DDWNAVIASNLHSLFNITK	<	0	0	
Image				1.531,49	3.062,37	-1,40	47	74	EQRELGEDEVASEGNAADWDSTMAA		1	1	
$I \rightarrow$													
		MS Data-											
- 11		Peak Id	on Mas	s Intensi	ty Area	Charg	e						
4.3		1	876,	90 8	70 980,	1 0				1 1			Labels
		2	1.079_{c}	47 3.0	26 844,	5 0		× a.	000				
	THE PARTY OF	3	1.079,	63 3.2	38 696,	4 0		su					
- 1		4	1.379,	40 1.6	10 704,	3 0		⊕ 2. ₩	000				
		5	1.379,0	60 1.3 col c	13 181,	1 U	-111	- 1.	000				
		0	1.395,1	00 5 40 4.6	21 240,		-111						
		0	1.001,4	49 1.0 71 2.2	201,100 201,1		-8		1000 1250 1500 1750 2000 2250	2500 2750 3 Aass	000 3250	3500 37	50 4000
		WSW5 Da	la			1							
		lon Mas	S	ntensity	Area	Charge						Γ	Labels
	-	04	00	1.320	003,4	4 I 2 1	864	> 25	500				
	100 MM	86	,00 NG	1 612	904,2	2 1	001	± 20	100				
	-	101	06	508	1047	7 1		⊂ 15 ©	500				
		104	.03	567	132.0	3 1	-111	E 10	200				
0.		110	,05	206	962,9	3 1							
-		120	,09	225	913,4	4 1			0 100 200 300 400 500 600 700	800 900 10	00 1100 1	200 130(0 1400 1500
1		129	,11	474	032,9	3 1	-		N	lass			

🔄 Data Browser

🦲 Gel Viewer -Gel Image -

forcer on _ D ×



Proteomics pathway



The MELANIE 2D gel analysis software



Benefits:

Accurately and rapidly identifies and quantifies your spots, matches your gel images, and finds significant protein expression changes.

Offers numerous filtering, querying, reporting, statistical and graphing options so that you can easily view, compare, analyse and present your results.









Integration into the laboratory workflow

- Integration with Amersham Biosciences' ImageScanner[™] via LabScan[™] 5.0, also developed by the SIB team
- Direct image acquisition from Twain compatible scanners
- Support of multiple file formats (.tif, .gel, .img, .png)
- Unrivalled annotation capabilities with the possibility to link gel objects to external query engines or data sources of any format (text, html, spreadsheet, multimedia, 2DE database entry, etc.), located locally or on the Internet
- Fully automated integration with spot-picking robots
- Clipboard support to copy images, graphics and tables to other programs



- Unique identifier for each spot, gel, and report to ensure data integrity and consistency
- Backup and restore functions
- Sophisticated Undo/Redo function
- History function tracks all operations for control and quality assurance





→~29

HPLC-ESI/QTOF



0.05 - [400-1200] m/z

 \rightarrow 16000 mesures



Proteomics pathway



MS Interpretation











- Java interface (standalone)
 - Creation and management of projects
 - Batch process capabilities
 - Various possibilities of result visualization
 - Spectra manipulation features

MS-MS Interpretation









ExPASy Proteomics tools

- Take into account Swiss-Prot annotations concerning PTMs and splice variants.
- Closely integrated and hyperlinked with Swiss-Prot and TrEMBL entries on ExPASy, and among each other.
- http://www.expasy.org/tools/



ExPASy Proteomics tools

- Protein identification
 - AACompIdent, AACompSim, MultiIdent, Peptident/Aldente, TagIdent
- Protein characterization
 - FindMod, GlycoMod, FindPept, Myristoylator, Sulfinator
- Primary structure analysis
 - Protparam, Compute pI/Mw, ProtScale
- Tertiary structure
 - SWISS-MODEL



Proteomics pathway







SWISS-2DPAGE

Release 17.0 (Mar 2004) : 1265 entries (3971 spots) in 36 reference maps:

Human: 18 maps (Liver, Plasma, Red blood cells, ...), *Mus musculus*: 6 maps (Liver, Islets, ...), *Arabidopsis thaliana, Dictyostellium discoideum, Saccharomyces cerevisiae, Staphylococcus aureus Escherichia coli:* 8 maps (for several pl ranges)

Structure of SWISS-2DPAGE



Suis Nice2DPage View of SWISS-2DPAGE:

P02679

NiceProt View of SWISS-PROT: P02679

Quick BlastP search

[General] [Name and origin] [References] [Comments] [Cross-references] [Keywords] [Features] [Sequence] [Tools]

				General information	about the entry		
				Entry name		FIBG_HUMAN	
				Primary accession nu	umber	P02679	
Concerl information about	t the entry			Secondary accession	number(s)	None	
General Information about	t the entry			Entered in SWISS-PR	lOT in	Release 01, July 1986	
View entry in original SWIS:	S-2DPAGE format	PROC. NATL. ACAD. SCI. U	J.S.A. 74:5421-5425(1977).	Sequence was last mo	odified in	Release 05, August 1987	
Entry name	FIRG HIMAN	2D PAGE maps for identified	l proteins	Annotations were las	st modified in	Release 40, October 2000	
		Comparing the discounting Lot D for		Protoin nome	не рготети	FIRDINGGEN GAMMAA A CHAIN Brownood	
Primary accession number	P02679	Compute the theoretical pl/MW		Suponum(s)		None	
Entered in	D A A A A	Cerebrospinal_Fluid		Gene name(s)		FGG	
SWISS-2DPAGE in	Kelease 00, August			From		Homo sapiens (Human)	
T	D. 1 . 00. 17	Shaff -	MAPLOCATIONS: • SPOT <u>2D-000BY</u>	p.		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; I	dutheria; Primates
Last modified in	Release 08, Novem	and the second s	 SPOT <u>2D-000BYY</u>: pI=5.53, Mw=51657 	1 axonomy		Catarrhini; Hominidae; Homo.	
Name and origin of the pro	otein	"HALE ICH	 SPOT <u>2D-000BZ6</u>: pI=5.59, Mw=51526 	NCBI TaxID		9606;	
Description	EIBRINGGEN GA	The second s		References			
Description	FIDIGINO OLIN OA	······································	 MAPPING: MATCHING WITH THE PL. 	A.S. SEQUENCE FROM N	I.A.		
From	HOMO SAPIENS			MEDLINE=85252774 Rixon M.W., Chung I	[NCBI, ExPASy, Isre D.W., Davie E.W.;	<u>ael, Japan];</u> PubMed= <u>2990550;</u>	
Town	EUKARYOTA; M	HepG2_Secreted_Proteins		"Nucleotide sequenc	e of the gene for the	gamma chain of human fibrinogen.";	
Taxonomy	MAMMALIA; EU		MAPLOCATIONS: • SPOT 2D-00		Conservation (1918)		
Poferenzas		Non- Comment	SPOT 2D 00001T al=5 27 Mar=46	- 59		A REAL PROPERTY AND DESCRIPTION OF THE OWNER OWN	
References		1272 127 128 128 128 128 128 128 128 128 128 128	CDOT 2D 0000111 I 5 45 M				
[1]		Hand St. Contraction	• SPOT <u>ZD-00091V</u> : pt=5.45, MW=46		**********		
MAPPING ON GEL.			MAPPING: MATCHING WITH TH			and the second se	
MEDLINE; 96007934. [NO	CBI, ExPASy, Israel, J						
SANCHEZ JC., APPEL F	LD., GOLAZ O.G., F			e -	11.	the state of the s	
DE			MAPLUCATIONS: • SPOT <u>2D-000</u>				
FI FOTROPHORESIS 16:1	131-1151/1995)		 SPOT <u>2D-0004 V 0</u>: pI=5.03, Mw=51 			A PATENCE AND A PATENCE	
	.151-1151(1775).		 SPOT <u>2D-0004VY</u>: pI=5.13, Mw=51 	8 -			
[2]			 SPOT <u>2D-0004VZ</u>: pI=5.16, Mw=50 	110	5 2 5 2 2 4		
MAPPING ON GEL.		Plasma	 SPOT <u>2D-0004WN</u>: pI=5.44, Mw=4 	11	• • • • • •		
MEDLINE; 78094420. [NC	<u>CBI, ExPASy, Israel, J</u>		• SPOT 2D-0004WQ: pI=5.34, Mw=4 9	1456			
ANDERSON N.L. ANDE	RSON N.G.:	A CONTRACTOR OF	• SPOT 2D-06 pI=5 56 Mw=4		31 - J. B.		
		2 mil.	• SPOT 2D-000 VT pI=5 24 Mw=49		·	The second se	
		ar a water	STOT 2D 0004WILL pI-5.65 Mar-4	π -	Sales Same	and the state of t	
		1.1 · · · ·	CDOT 2D 0004W0 PI-5.05, WW-4		1. 1. 1.		
			• SPOT <u>ZD-0004AS</u> : p1=3.44, MW=46				
			 SPOT <u>2D-0004XV</u>: pI=0.06, Mw=48 	8 -	• • •		
			 SPOT <u>2D-0004ZZ</u>: pI=5.56, Mw=44 		•		
			 SPOT <u>2D-000500</u>: pI=5.62, Mw=447 				
			MAPPING: MATCHING WITH TH				
		Copyright		12			
		This SWICE ODDA CE	a constitute the Carries Institute of Disinformation	<u> </u>	• 13 - •	· · · · · · · · · · · · · · · · · · ·	
		THIS SWISS-ZDPAGE entry is	copyright the Swiss institute of Biomformatics	1			
	I	Unon-protit institutions as long as	are content is in no way modified and this state				
				4.0 4.5	5.0 5.5	6.0 6.5 7.07.58.0	

pH. This image is copyright by SIB. Usage by and for commercial entities requires a license agreement.



Make 2D-DB II

- Relational format structure (PostgreSQL)
- Conversion of already existing federated 2-DE databases or creation of new ones
- Automatic updates related to external data



Make 2D-DB II

- Facilitate the maintenance through Web interfaces (queries, updates)
- Flexible structure (facilitate extention to new information types)
- Interconnect remote databases in order to build a virtual global one
- Generation of flat file and XML formats



2D-PAGE 2D-PAGE Test Database Test DB This is a fictive database to test the Make2D-DB II too

This is a fictive database to test the Make2D-DB II tool

Search by:	Search proteins by pl/Mw range					
- accession number	pl Range:					
- description, ID or gene	min. => 0 max. => 15					
- author	Mw Range (K.Daltons):					
- spot ID / serial number	min.=> 0 max.=> 250					
> pI/Mw range Limit to Map: All Maps						
combined fields	Please give a pl range for your spots. For example, you may type 7.25 or <i>just</i> 7. Give also a range for Mw in Kilo Daltons . You may type, for example, 90.550 or 110. The search will include all maps, except when you limit your search to a particular map.					
proteins in a map	Execute query Reset					
navigate the maps						
	2D-PAGE Test Database					
Gateways to other related servers						

SWISS-2DPAGE - The two-dimensional polyacrylamide gel electrophoresis database of the Swiss Institute of Bioinformatics in Geneva

- WORLD-2DPAGE Index to other Federated 2-D PAGE databases
- ExPASy The proteomics web server of the Swiss Institute of Bioinformatics, in Geneva

This database was constructed using the Make2D-DB II package (ver. 0.39) from the WORLD-2DPAGE of the ExPASy web server



EBP federated 2D-PAGE Viewer: <u>Swiss-2DPAGE ECOLI</u> ECOLI-DIGE4.5-6.5 (All spots)

This is an EBP internal server to query 'EBP federated 2D-PAGE' in relational (PostgreSQL) format









Integration of various





Thank you for your atttention

Proteo E Infor atics group SIB

