The Swedish Human Proteome Resource

- Funding (non-profit) from the Wallenberg Foundation (60 FTEs)
- Hosted at the AlbaNova University Center, Stockholm, and the Uppsala University, Sweden (www.hpr.se)
- Antibodies: in-house generated mono-specific antibodies and antibodies from commercial sources (collaborative efforts)

hpr

Human Antibody Initiative (HAI)

Chair: Mathias Uhlen, Sweden

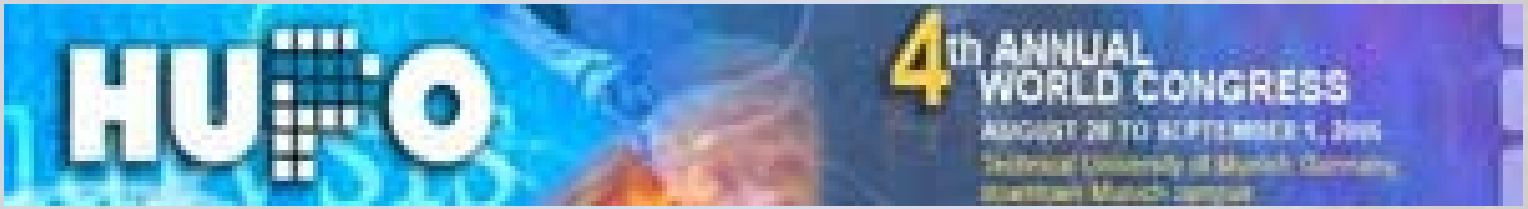
Co-chairs: to be decided

Two subprograms:

1. Protein Atlas (public database)
2. Antibody resource (public database)

Second Antibody Proteomics Workshop:

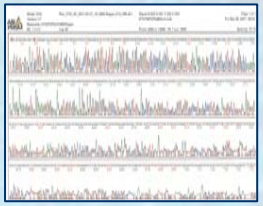
Saturday 27th of August, Munich, Germany



The Swedish HPR program

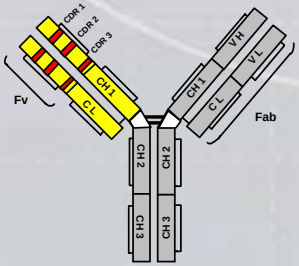
Upstream bioinformatics

Molecular biology

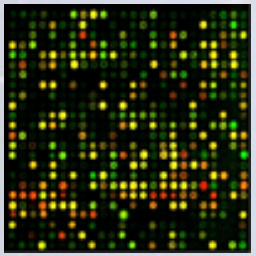


Sequence verification

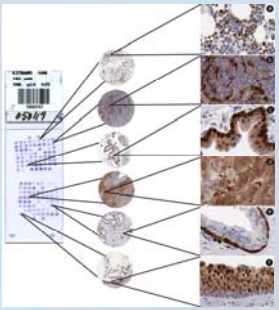
Protein Factory



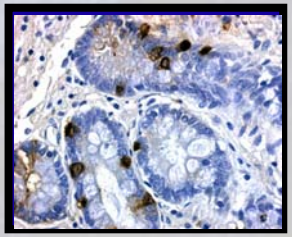
Immuno-technology



Quality assurance



Tissue Arrays



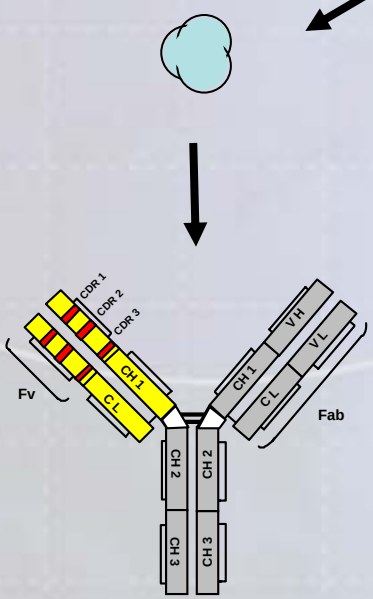
Web-based Image Database

hpr

Antigen: Protein Signature Epitope Tags (PrESTs)

....Trp Asp Trp His ... Arg Cys Ile Met Ala Leu Trp Ile His...

100 to 150 amino acids



- Avoid homology (cross-reactivity)
- Avoid transmembrane spanning regions
- Allows ”epitope” specific antibodies

- Incomplete (unknown) folding

Agaton et al (2003) Molecular and Cellular Proteomics 2, 405-416

Polyclonal or monoclonal antibodies ?

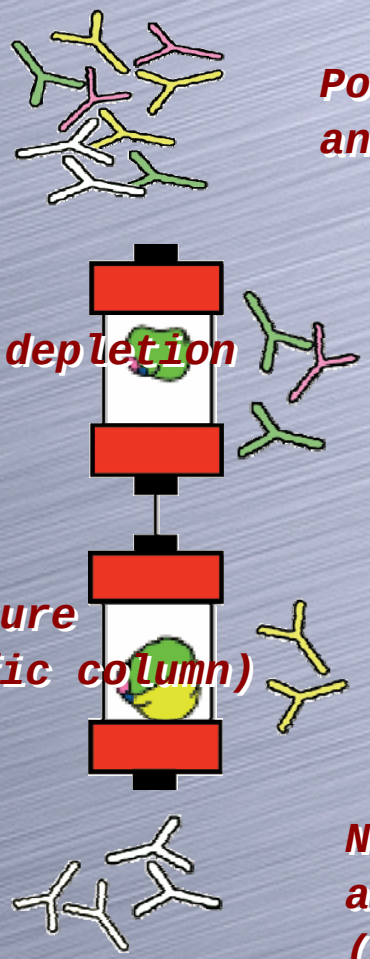
Monoclonal antibodies (or other reagents)

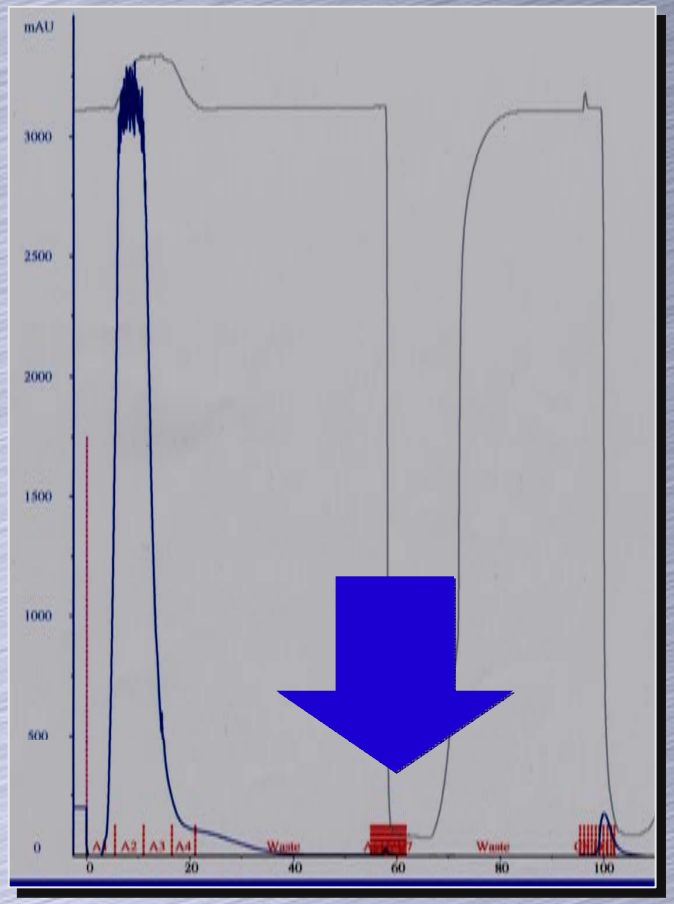
- Defined binding parameters (single -epitope
- Renewable resource

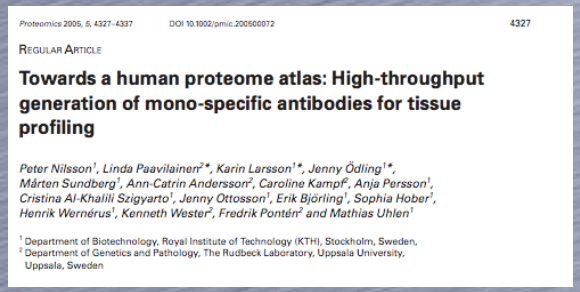
Polyclonal antibodies

- “Multi-epitope” reagents
- Inexpensive to produce

Monospecific antibodies (msAb)







Proteomics 2005, *5*, 4327–4337 DOI 10.1002/pmic.200500072 4327

REGULAR ARTICLE

Towards a human proteome atlas: High-throughput generation of mono-specific antibodies for tissue profiling

Peter Nilsson[1], Linda Paavilainen[2], Karin Larsson[1]*, Jenny Ödling[1]*, Mårten Sundberg[1], Ann-Catrin Andersson[2], Caroline Kampf[2], Anja Persson[1], Cristina Al-Khalili Szigyarto[1], Jenny Ottosson[1], Erik Björling[1], Sophia Hober[1], Henrik Wernérus[1], Kenneth Wester[2], Fredrik Pontén[2] and Mathias Uhlen[1]*

[1] Department of Biotechnology, Royal Institute of Technology (KTH), Stockholm, Sweden,
[2] Department of Genetics and Pathology, The Rudbeck Laboratory, Uppsala University, Uppsala, Sweden

Nilsson et al (2005) "Towards a Human Protein Atlas - mono-specific antibodies for tissue profiling", PROTEOMICS, in press

Assays for quality assurance

- Protein assays, such as ELISA, protein arrays, etc (with antigen as analyte)
- Adsorption assays (pre-incubate antibody with antigen)
- Western blots (size and specificity using relevant human tissue extracts)
- Immunohistochemistry (quality and bioinformatics comparison)
- RNA/protein comparisons (expression profiling of cell lines or in situ)

The ultimate quality assurance: compare the results of two or more antibodies binding to different parts (epitopes) of the same target protein

hpr

Application-specific validation

ELISA

Flow cytometry

Immunohistochemistry

Immunoprecipitation

Western blot

Affinity purification

Blocking/neutralize

Electron microscopy

Functional assays

Gel shift

Immunocytochemistry

Immunodiffusion

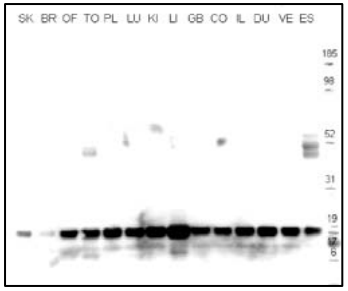
Immunofluorescence

Radioimmunoassay

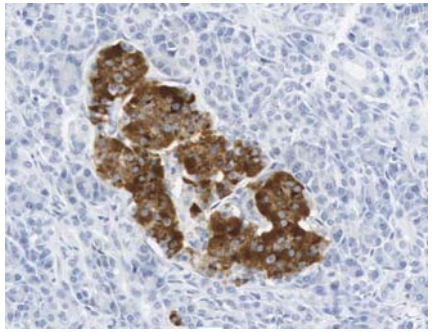
Other (?)

HPR antibody quality assurance

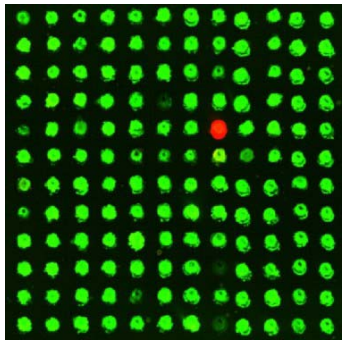
1. Western blot



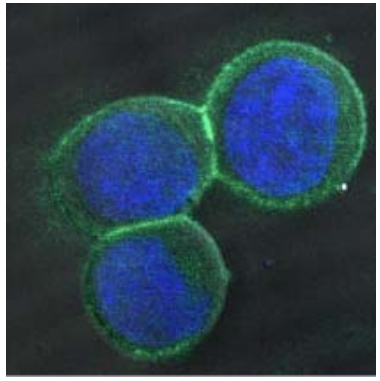
3. Immunohistochemistry (IHC)

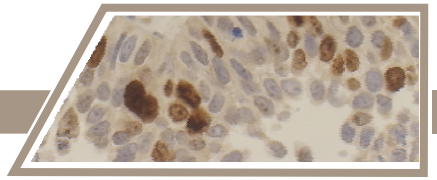


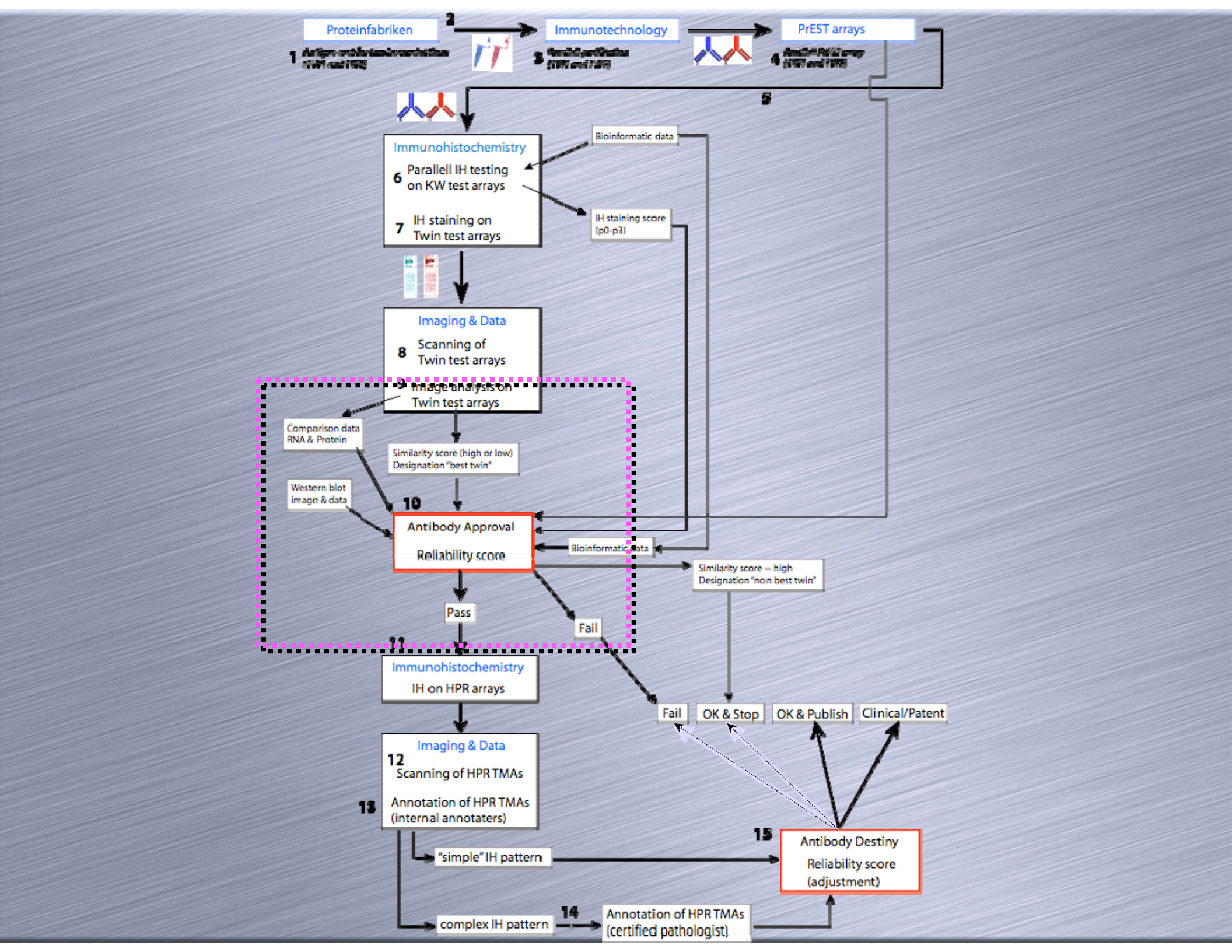
2. PrEST (protein) array



4. Confocal (fluorescent) microscopy





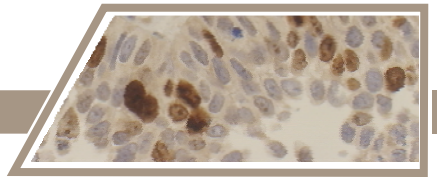

Proteinfabriken
1
2
Immunotechnology
3
PrEST arrays
4
5
Immunohistochemistry
6 Parallell IH testing on KW test arrays
7 IH staining on Twin test arrays
Bioinformatic data
IH staining score (p0-p3)
Imaging & Data
8 Scanning of Twin test arrays
9 Image analysis on Twin test arrays
Comparison data RNA & Protein
Similarity score (high or low) Designation "best twin"
Western blot image & data
10
Antibody Approval
Reliability score
Bioinformatic data
Similarity score – high Designation "non best twin"
Pass
Fail
11
Immunohistochemistry
IH on HPR arrays
Fail
OK & Stop
OK & Publish
Clinical/Patent
Imaging & Data
12 Scanning of HPR TMAs
13 Annotation of HPR TMAs (internal annotaters)
15
Antibody Destiny
Reliability score (adjustment)
"simple" IH pattern
complex IH pattern
14
Annotation of HPR TMAs (certified pathologist)

Validation score

High, Medium, Low, Not approved

High: indicates that the **quality assurance** and **literature** supports the specificity of the antibody towards the expected human target protein with high confidence

Low: shows that the antibody probably is specific to the expected target, and in this case, the **validation is less clear** and cross-reactivity cannot be excluded.



1. Tissue Western blot

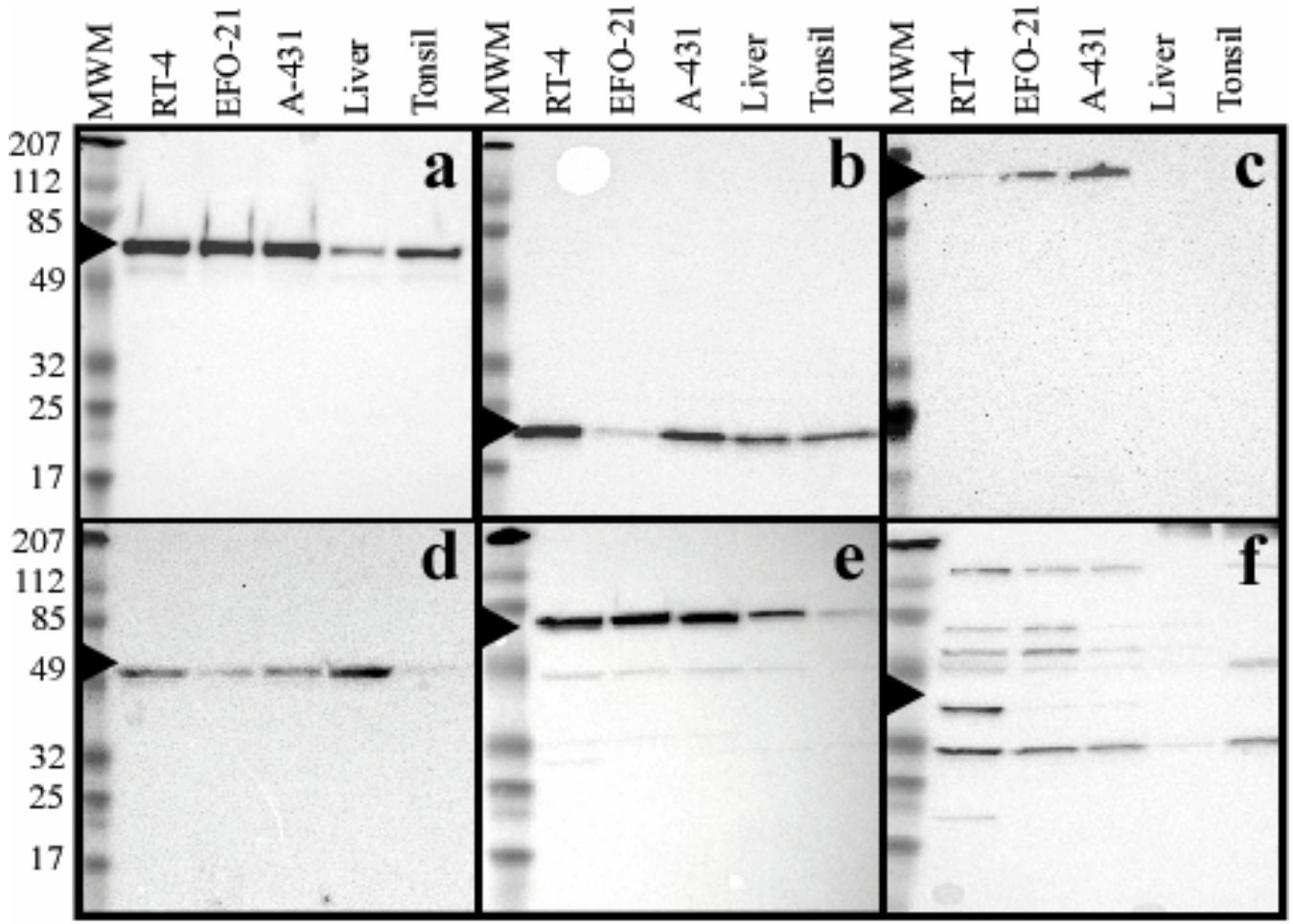
QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

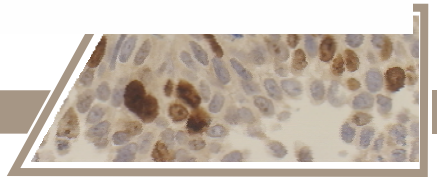
Human cell-lines:

RT-4,EFO-21,A-431

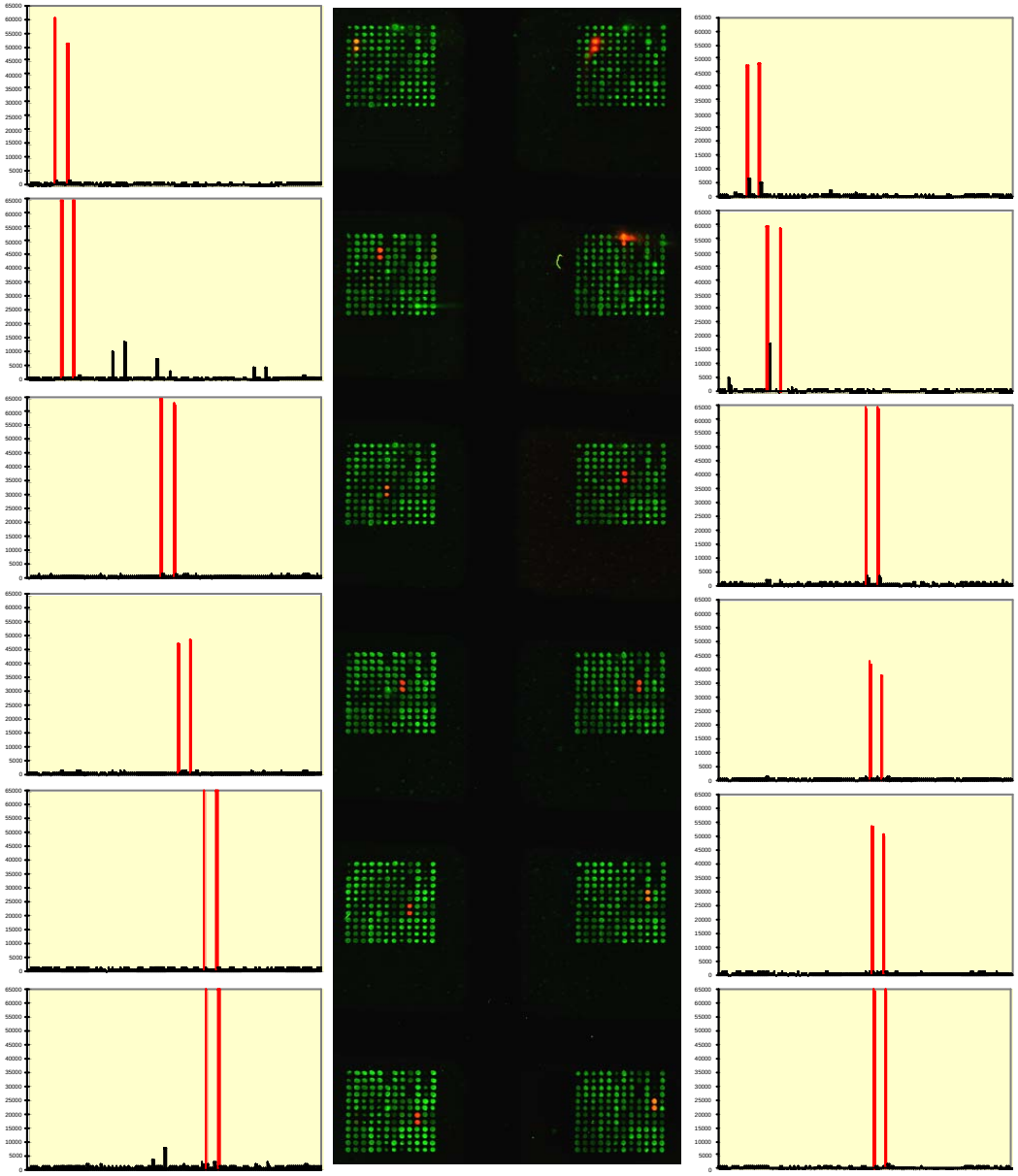
Human tissues:

liver and tonsil

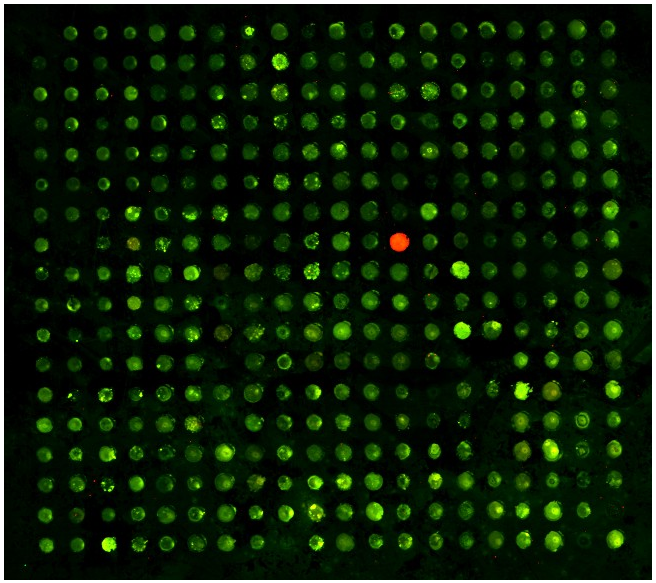




2. Protein arrays (PrESTs)

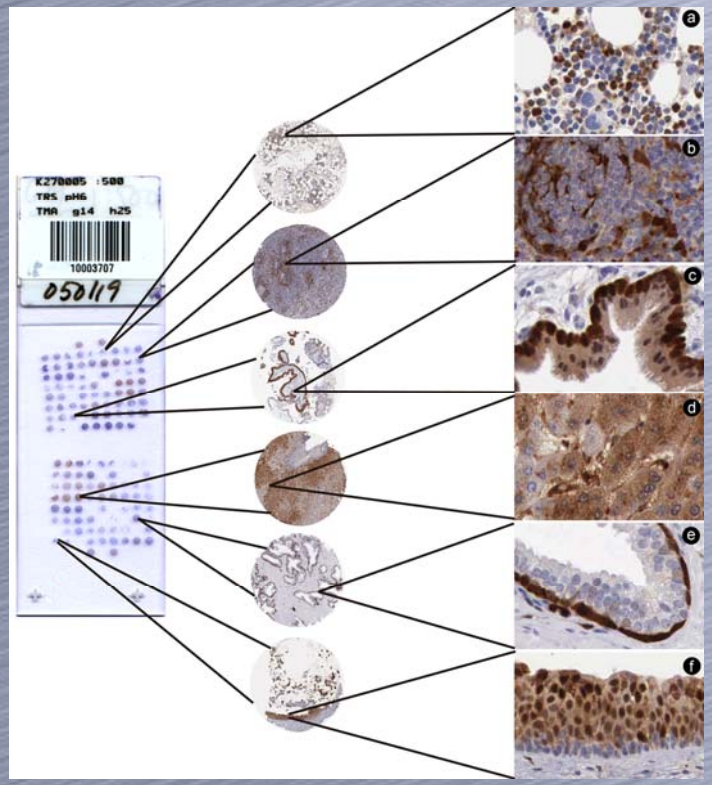


Results from one protein array slide for 12 consequtive antibodies



1400 PrESTs in triplicate on one slide

3. Immunohistochemistry (IHC)



Tissues

1. Liver
2. Tonsil
3. Colon
4. Kidney
5. Uterus
6. Placenta
7. Testis
8. CNS
9. Skin
10. Striated muscle

Cells

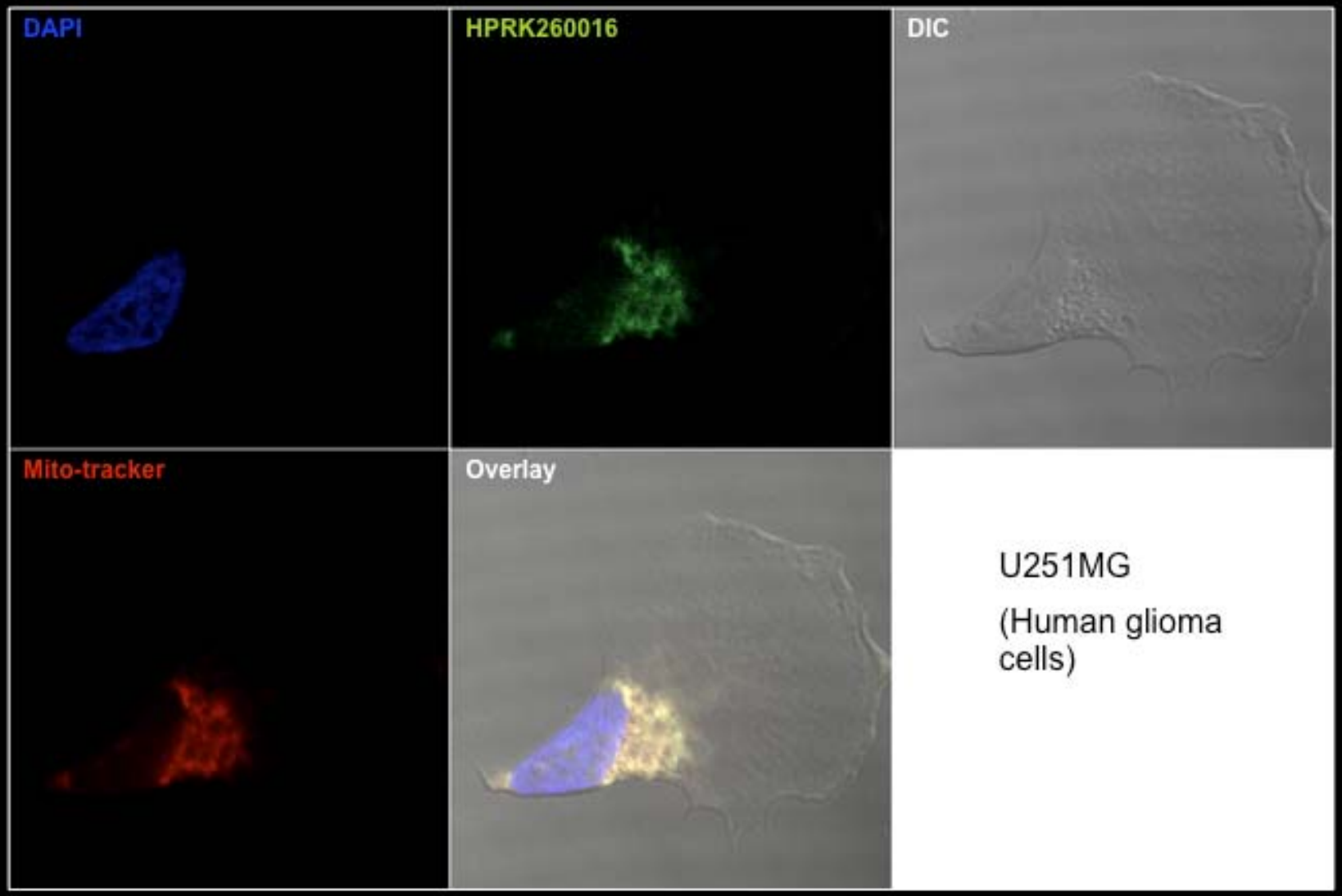
1. A-431 (skin)
2. RT-4 (urinary)
3. U251 MG (brain)
4. PC-3 (prostate)
5. SK-Mel30 (skin)
6. U-2OS (sarcoma)
7. K-562 (myeloid)
8. THP-1 (myeloid)
9. Hep G2 (liver)
10. SkBr (breast)

Duplicate samples, 0.6 mm diameter
Manual evaluation (microscope)
Scanning (time 15 min?)
TMAx analysis time ?
Output data - similar, non-similar, partly s

4. Confocal microscopy with multiple fluorescent dyes

HPRK260016: α-Stress-70 protein, mitochondrial precursor

Subcellular localisation: Mitochondria



Quality assurance

- Protein Atlas will contain antibodies with variable quality
- Objective: validate all antibodies with a standardized unit operations
- Long-term objective: two antibodies to every non-redundant protein
- Input (dialogue) from the international community necessary

PROTEOMICS

New Database to Track Protein Locations

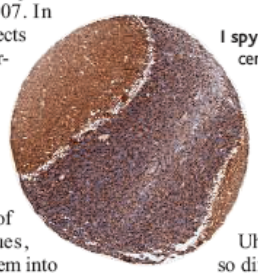
Proteomics researchers in Sweden plan to release a database next week containing hundreds of thousands of images of where different proteins are located in human cells and tissues. The database, dubbed the Protein Atlas, is intended to help biochemists identify the function of newly discovered proteins. Although the new atlas currently contains data on only some 700 proteins, the Swedish team plans to tackle some 22,000 different proteins, one for each human gene.

"That's great," says Richard Smith, a proteomics expert at the Pacific Northwest National Laboratory in Richland, Washington. "It's one of the most valuable data sets you can have," adds Michael Snyder of Yale University, who pioneered a similar large-scale effort to localize proteins in yeast. The yeast data set, for example, has proven to be an essential tool in narrowing down whether proteins operate in the nucleus, the cell membrane, or elsewhere.

As scientists began to sequence human genes in the 1990s, sorting out the cellular locations of each gene's proteins became a priority, says Mathias Uhlén, microbiologist at the Royal Institute of Technology in Stockholm, Sweden, and director of the Protein Atlas effort. "This is something that has to be done to leverage the success of the human genome project," he explains.

A pilot project 2 years ago convinced the Knut and Alice Wallenberg Foundation of Sweden to bankroll a scaled-up Protein Atlas through September 2007. In April, it became one of six projects to be coordinated by the international Human Proteome Organisation (HUPO). Uhlén says he hopes HUPO member countries will finance the completion of the Atlas, which could take another 10 years.

To track down the location of proteins inside human tissues, Uhlén's team breaks the problem into two parts—finding antibodies that target individual proteins, and then using those antibodies to hunt for proteins inside tissues. To streamline this process, Uhlén's team has created standardized arrays containing microscopic tissue samples from 48 different normal human tissues and 20 types of cancer tissue. The antibodies are tagged so they can be seen and incubated with the arrays to reveal which proteins are expressed in each of the different tissues. The tissues are then photographed at high resolution, providing for each antibody hundreds of detailed images revealing where it has bound to its target protein. For now, Uhlén says, his team of about 100 scientists is creating half a dozen antibodies a day, leading to about 30 gigabytes of data for each antibody studied (which is stored at www.proteinatlas.com).



I spy. Glycoprotein SV2A (brown) in cerebellum tissue.

Today's arsenal of drugs, Uhlén notes, targets only 500 or so different proteins. By providing clues to the function of other proteins, he says, the Atlas may accelerate their use as markers for disease or drug targets.

The Protein Atlas will have wrinkles to be ironed out. If antibodies react with more than one protein, the tissue arrays may unwittingly spotlight unintended proteins. "There are huge issues of quality assurance," Uhlén says. As a result, his team will count on outside experts to flag problems. —ROBERT F. SERVICE

CREDITS (TOP TO BOTTOM): H. K. LINDSJÖ; THE SWEDISH HUMAN PROTEOME RESOURCE

Published by AAAS

A Human Protein Atlas for Normal and Cancer Tissues Based on Antibody Proteomics*

Mathias Uhlén‡§, Erik Björling‡, Charlotta Agaton‡, Cristina Al-Khalili Szigyarto‡, Bahram Amini‡, Elisabet Andersen¶, Ann-Catrin Andersson¶, Pia Angelidou‡, Anna Asplund¶, Caroline Asplund‡, Lisa Berglund‡, Kristina Bergström‡, Harry Brumer‡, Dijana Cerjan¶, Marica Ekström‡, Adila Elobeid¶, Cecilia Eriksson‡, Linn Fagerberg‡, Ronny Falk‡, Jenny Fall‡, Mattias Forsberg‡, Marcus Gry Björklund‡, Kristoffer Gumbel¶, Asif Halimi‡, Inga Hallin¶, Carl Hamsten‡, Marianne Hansson‡, My Hedhammar‡, Görel Hercules¶, Caroline Kampf¶, Karin Larsson‡, Mats Lindskog‡, Wald Lodewyckx‡, Jan Lund‡, Joakim Lundeberg‡, Kristina Magnusson¶, Erik Malm‡, Peter Nilsson‡, Jenny Ödling‡, Per Oksvold‡, Ingmarie Olsson¶, Emma Öster‡, Jenny Ottosson‡, Linda Paavilainen¶, Anja Persson‡, Rebecca Rimini‡, Johan Rockberg‡, Marcus Runeson¶, Åsa Sivertsson‡, Anna Sköllermo‡, Johanna Steen‡, Maria Stenvall‡, Fredrik Sterky‡, Sara Strömberg¶, Mårten Sundberg‡, Hanna Tegel‡, Samuel Tourle‡, Eva Wahlund¶, Annelie Waldén‡, Jinghong Wan‡, Henrik Wernérus‡, Joakim Westberg‡, Kenneth Wester¶, Ulla Wrethagen‡, Lan Lan Xu‡, Sophia Hober‡, and Fredrik Pontén¶

Antibody-based proteomics provides a powerful approach for the functional study of the human proteome involving the systematic generation of protein-specific affinity reagents. We used this strategy to construct a comprehensive, antibody-based protein atlas for expression and localization profiles in 48 normal human tissues and 20 different cancers. Here we report a new publicly available database containing, in the first version, ~400,000 high resolution images corresponding to more than 700 antibodies toward human proteins. Each image has been annotated by a certified pathologist to provide a knowledge base for functional studies and to allow queries about protein profiles in normal and disease tissues. Our results suggest it should be possible to extend this analysis to the majority of all human proteins thus providing a valuable tool for medical and biological research. *Molecular & Cellular Proteomics 4:•••–•••, 2005.*

Antibody-based tissue profiling allows a streamlined approach for generating expression data both for normal and disease tissues (1). It is also possible to generate protein expression data on many different individual patients to evaluate heterogeneity of tissue profiles (2, 3). In addition, specific antibodies directed to a particular target protein allow numerous functional assays to be performed ranging from conventional ELISA assays to detailed localization studies using fluorescent probes and protein capture experiments ("pull-down") for purification of specific proteins and their associated complexes for structural and biochemical analyses (4, 5).

The challenge for antibody-based proteomics is to move from a conventional protein-by-protein approach into a high throughput mode to allow chromosome wide analysis (6, 7). Technical challenges involve both the antigen production and the subsequent generation and characterization of the antibodies. In addition, methods for systematic protein profiling on a whole proteome level are lacking. Agaton *et al.* (8) showed that the combination of the cloning and expression of recombinant protein fragments with immunohistochemistry analysis could be used for systematic protein expression and subcellular localization describing distribution and expression of putative gene products in normal tissues as well as in common cancers and other forms of diseased tissues. Recently Nilsson *et al.* (9) showed that this strategy could be further improved by a streamlined approach for affinity purification of the antibodies to generate monospecific antibodies (msAbs)[1]

From the ‡Department of Biotechnology, AlbaNova University Center, Royal Institute of Technology (KTH), 106 91 Stockholm, Sweden and the ¶Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, ■■■ ■■ Uppsala, Sweden

Received, August 25, 2005

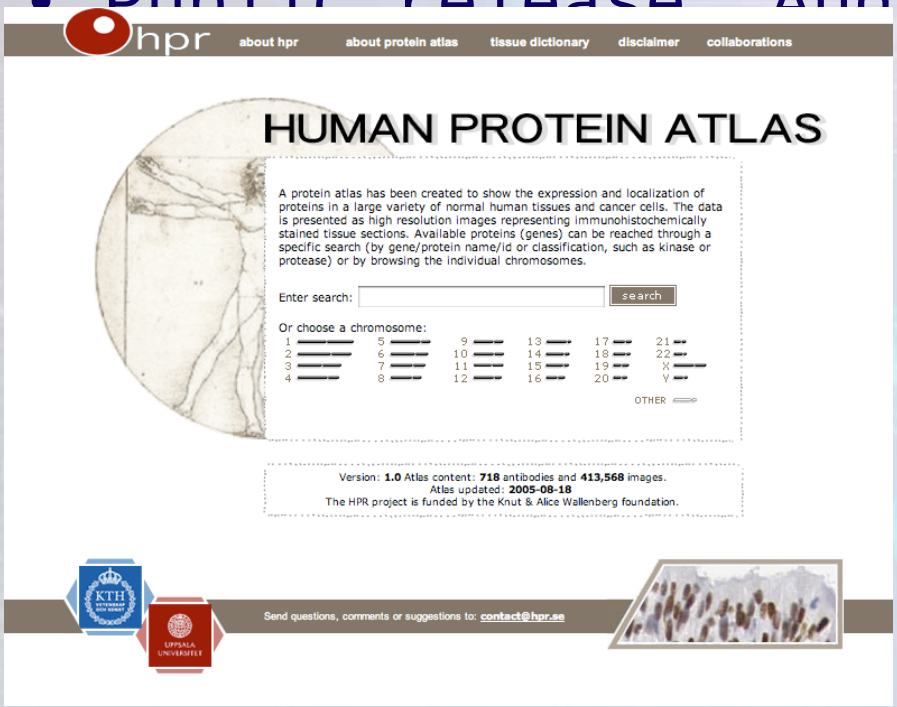
Published, MCP Papers in Press, August 27, 2005, DOI 10.1074/mcp.M500279-MCP200

[1] The abbreviations used are: msAb, monospecific antibody; TMA, tissue microarray; PrEST, protein epitope signature tag; PBST, phosphate-buffered saline supplemented with Tween; TIFF, tagged image file format; JPEG, Joint Photographic Experts Group; HPR, human proteome resource; MAT1, menage a trois-1; CRIM-1, cysteine-rich motor neuron 1; CD31, cluster of differentiation 31; PSA, prostate-specific antigen.

© 2005 by The American Society for Biochemistry and Molecular Biology, Inc.
This paper is available on line at http://www.mcponline.org

Protein Atlas

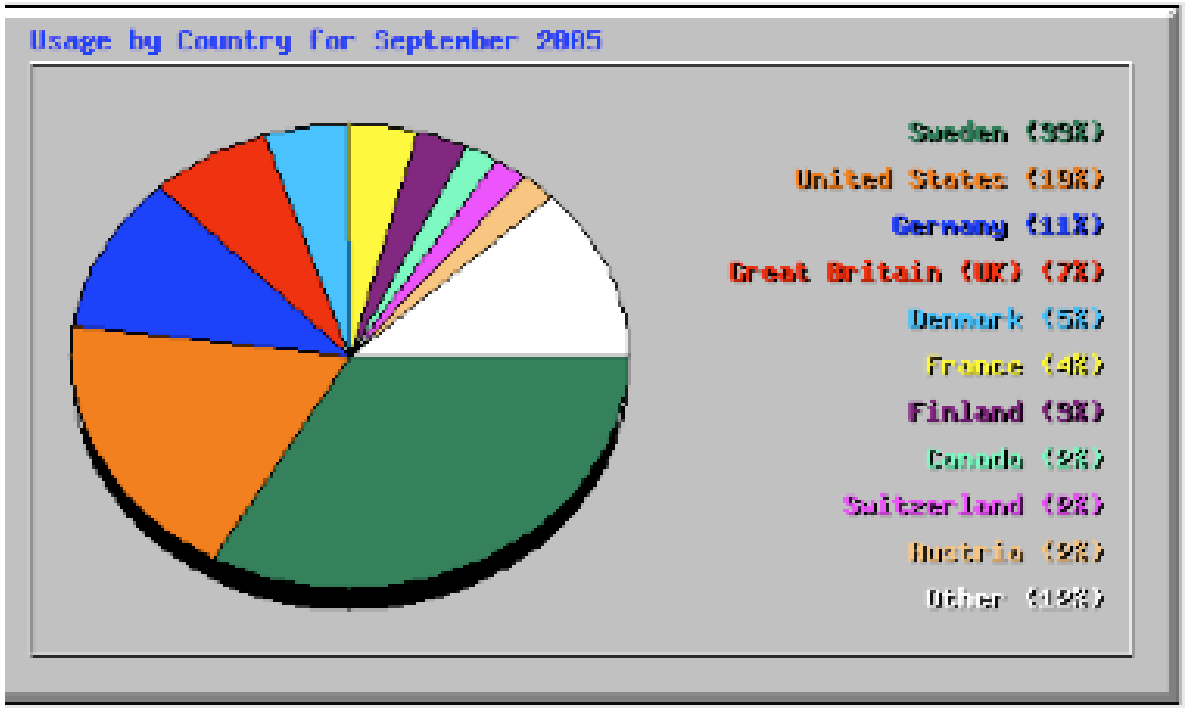
- Demo version ready, January 15
- First version released to selected groups, March 15
- Public release, August 29 (HUPO



www.proteinatlas.org

hpr

Protein Atlas - status



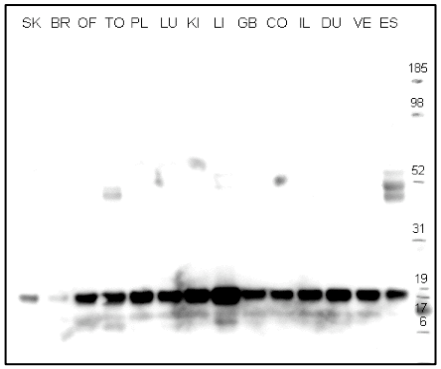
- About 30.000 hits/day
- Collaboration with Uniprot & Ensembl
- Collaboration with Swissprot

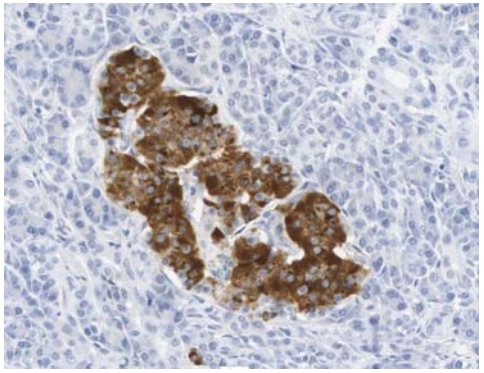
HPR status

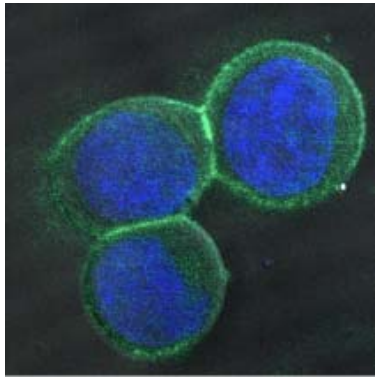
- Public release of Protein Atlas: August 29, 2005
- Focus on antibody validation
- Based on polyclonal antibodies (not renewable resource)
- 55 FTEs, but almost 100 researchers involved in the project
- Open for collaborations (to validate and use external antibodies)

hpr

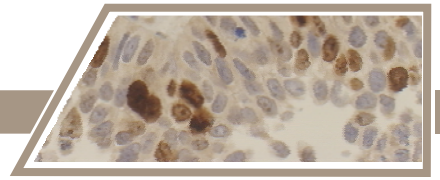
HPR validation of external antibodies







- **Target-based validation (antigen not available)**
  - **Western blots (2 cell lines, 2 tissues, plasma)**
  - **Tissue microarray (10 tissues, 10 cell lines)**
  - **Confocal microscopy (3 cell lines)**



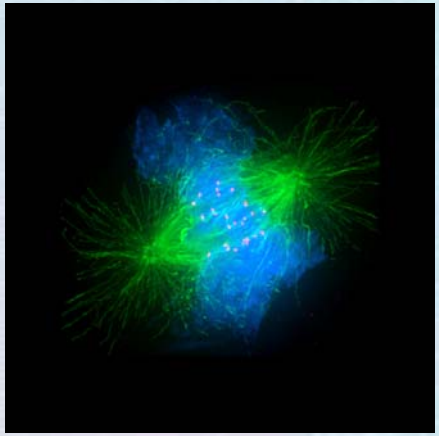
HPR validation of external antibodies

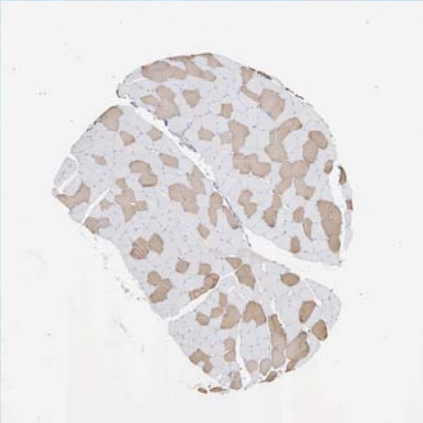
- **Possible to submit antibodies for validation (www.proteinatlas.org)**
- **Pilot project free of charge (until mid 2006)**
- **Every antibody will be validated by HPR team (WB, IHC, CM)**
- **Approved antibodies will be used for the Protein Atlas initiative**

Ensembl version:	Number of version used to identify genes and transcripts (www.ensembl.org)
Protein/gene name	Name of protein (or gene). If many alternative names, choose one.
Ensembl gene Id	Go to www.ensembl.org. Find your gene/protein. If protein is not found, write: Not found.
Ensembl transcript Id	If several splice forms exists, please indicate the relevant transcript.
Type of antigen	Specify antigen. Choose from synthetic peptide (peptide), recombinant protein fragment (PrEST), full-length recombinant protein (recombinant), purifed from natural sources (purify
Size of antigen	Give size of antigen. Choose from (i) kDa (ii) number of amino acids or (iii) full-length or (iv) not known.
Type of antibody	Specifiy type of antibody. Choose from monoclonal (mAb) polyclonal (pAb) affinity-purified polyclonal (msAb) recombinant (recAb) and other (other), please specify in Antibody co
Host species of antibody	Specify host in which the antibody is produced. Choose from (i) mouse (ii) rabbit (iii) hen (iv) recombinant or (v) other. Please specifiy in Antibody comment box.
Western blot (yes/no)	If Western blot analysis have been done succesfully, type yes.
IHC (yes/no)	If immunohistochemistry has been done succesfully, type yes. Indicate the recommended dilution factor (in Antibody comment box) for use in IHC applications, if known.
Type of validation	Indicate the type of validation. Choose from ELISA (E), protein array (A), Western blot (W), Immunohistochemistry (I), "pull-outs" (P), Flow sorting (F) and/or other (O).
Validation data (url/pdf)	Type url-link to page with validation data or name of PDF-file with data which is attached. If no validation is available, type N/A,
Link to antibody	Indicate if the antibody is available. If available give the web link (URL) to obtain antibody. If not available to the public, please indicate web link for reference/acknowledgement.
Antibody concentration	Optional. Need to be fill-in subsequently for approved antibodies
Antibody volume	Optional. Need to be fill-in subsequently for approved antibodies
Recom- storage temp	Optional. Need to be fill-in subsequently for approved antibodies
Antigen comments (optional)	Type description of antigen (N-terminal, C-terminal, peptide 23-45 etc). Indicate also if validation (such as MS) has been done.
Antibody comments (optional)	Type additional information. Lyophilized, store in glycerol etc.
References (optional)	Type appropriate reference on the use of this particular antibody.

Antibody name	Protein/Gene Name	Ensembl Gene ID	Ensembl Transcript ID (optional)	Type of antigen	Size of antigen	Type of antibody	Host species (of antibody)	Western blot (yes/no)	IHC (yes/no)	Type of validation	Validation data (url-link or pdf-file name)	Link to antibody (url)	Antib con (µg/r
PSA	PSA	ENSG00000142515	ENST00000	Purify	Full-length	mAb	mouse	no	yes	E, I	www.	www.dako.com	
EVT6	EVT6	ENSG00000139083	ENST00000	PrEST	14 kDa	msAb	rabbit	yes	yes	A, W, I	validation_EVT6.pdf	www.proteinatlas.org	
insulin	insulin	ENSG00000129965		Recomb	72 aa	pAb	hen	no	yes	E, I,	N/A	www.bioscpare.com	
Il-2	Il-2	ENSG00000109471		Synthetic	24 aa	RecAb	rec	no	no	E	validation_IL-2.pdf	www.hejsan.se	

Conclusions





- Human Protein Atlas - www.proteinatlas.org
- First version contains 400,000 images (20 Terabyte)
- Interest to validate and use external antibodies (monoclonal reagents)
- Both academic and commercial providers
- Approved antibodies will be web-linked to the provider