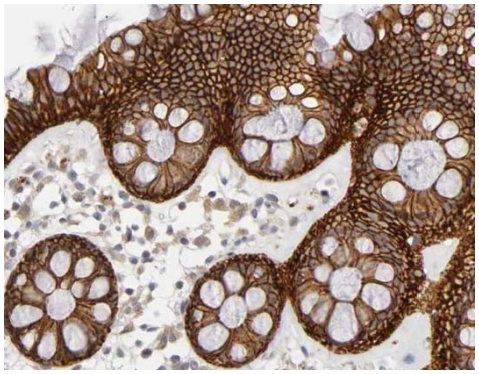
A human protein atlas for normal and cancer tissues

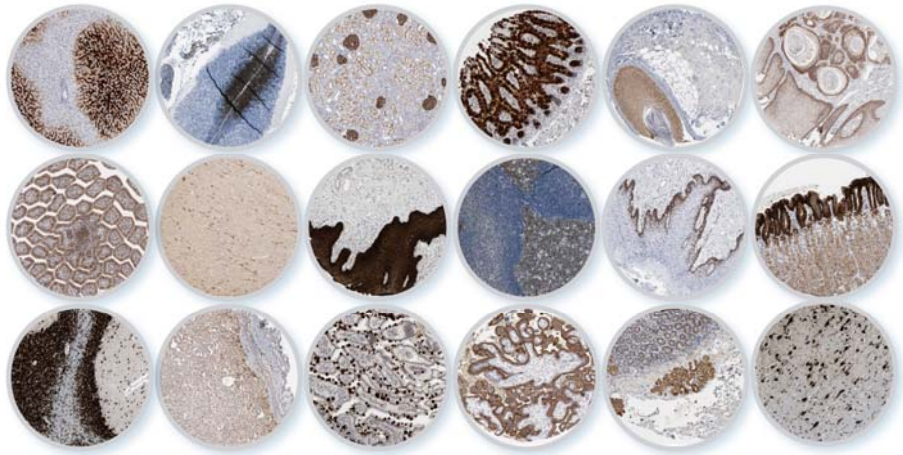
Peter Nilsson
Dept. of Proteomics
School of Biotechnology
KTH – Royal Institute of Technology
AlbaNova University Center
Stockholm, Sweden
nipe@kth.se

June 13, 2008



Antibody based proteomics

The systematic generation and use of antibodies to functionally explore the proteome



Human Protein Atlas (www.proteinatlas.org)
– a public database with expression profiles in normal and cancer tissues

The Swedish Human Proteome Resource (HPR)

- Program director: Mathias Uhlén
- Funding from Knut and Alice Wallenberg Foundation started 2003
- Seven sites: KTH, Uppsala, KI, Malmö, Seoul, Beijing and Mumbai
- 80 employees at KTH and Uppsala University (altogether 200 researchers)
- A publicly available Human Protein Atlas
- Current throughput: ~10 new validated antibodies per day
- Aim to analyze 10,000 human proteins by the end of 2009
- First draft of the human proteome by 2014

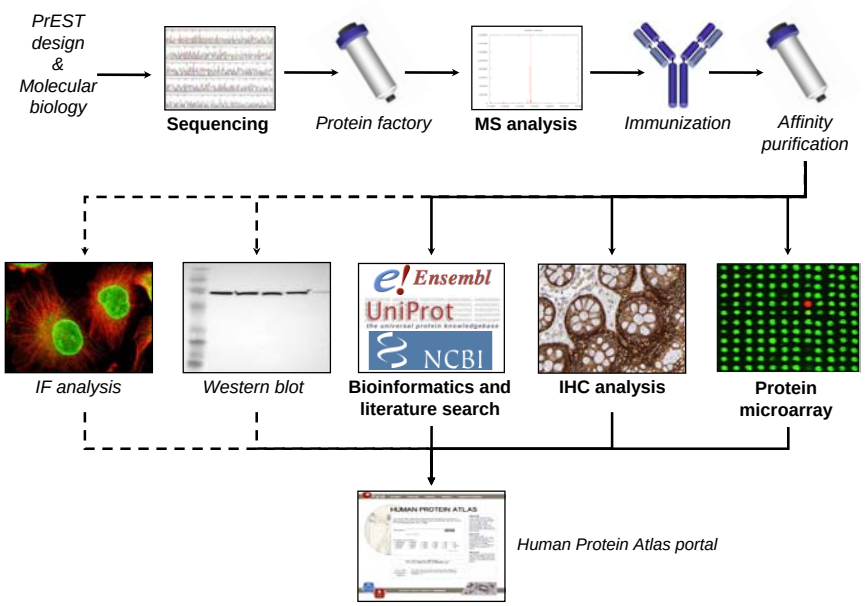
How large is the human proteome?

Non-redundant protein	20, 488*	A representative protein from every gene locus
Protein variants	>200,000	Different protein fragments (splice variants or proteolytic fragments)
Protein species (isoforms)	>100,000	Protein that differ in post-translational modifications (PTM)
Combinatorial variants	>10 million	Different proteins created by somatic DNA rearrangements
Protein allels	>75,000	Proteins that differ by genetic variation (coding SNPs)

- *18,609 annotated by SwissProt*
- *10,721 evidence at protein level*

** Clamp et al (2007) Distinguishing protein-coding and noncoding genes in the human genome PNAS 104:49, p19428-433.*

A multi-disciplinary program



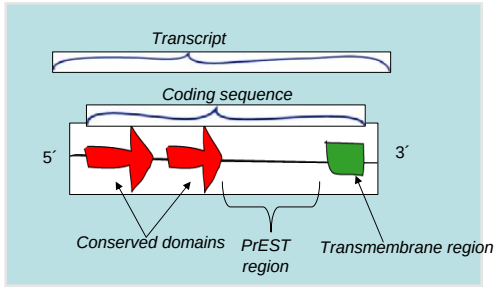
PrEST design

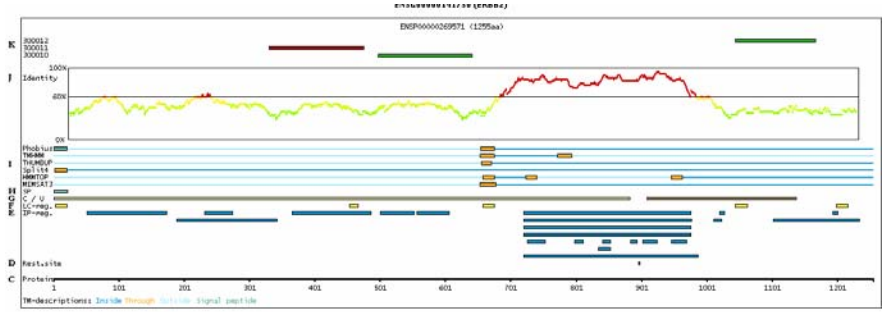
PrEST = Protein Epitope Signature Tag. Selected fragment (25-150 amino acids) of a protein. Represents the target protein.

As sequence-dissimilar as possible to other proteins. Avoid homology

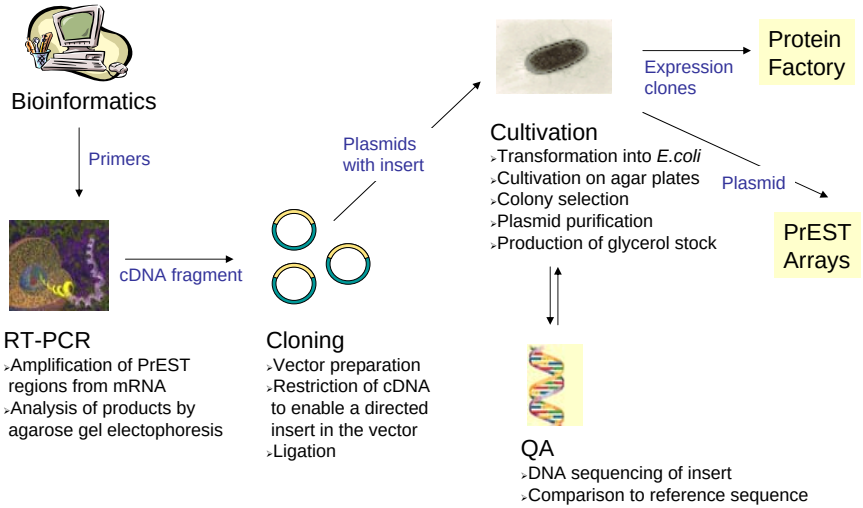
Avoid transmembarne regions

If possible design four PrESTs per protein





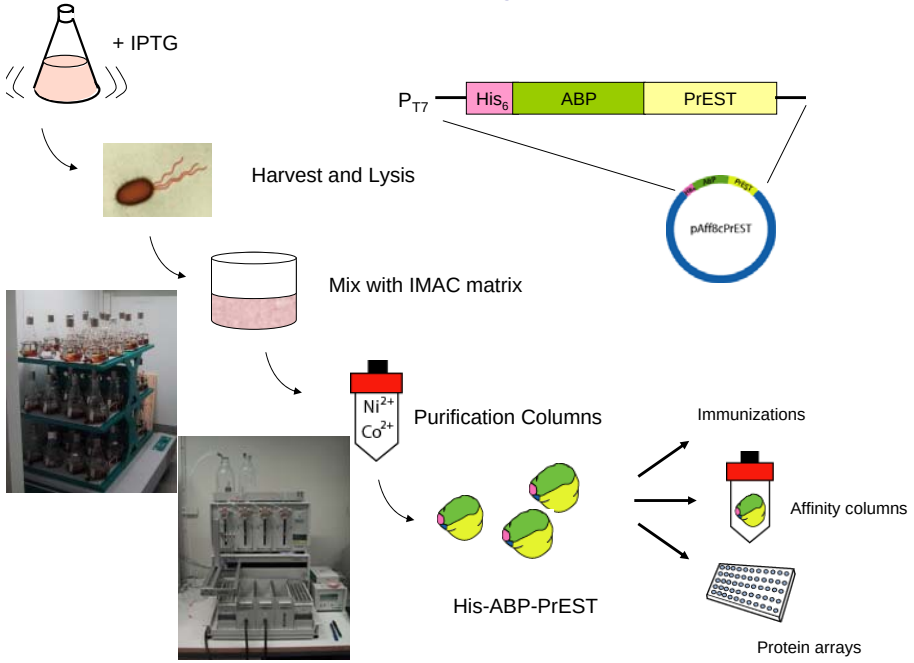
Molecular Biology



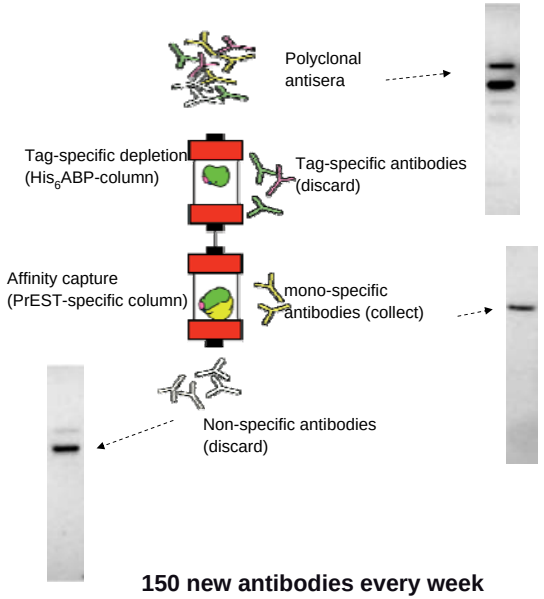
250 clones/week

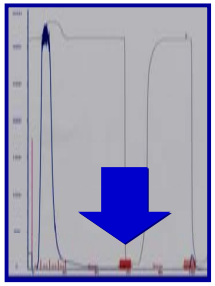
Protein factory

200 purified proteins/week



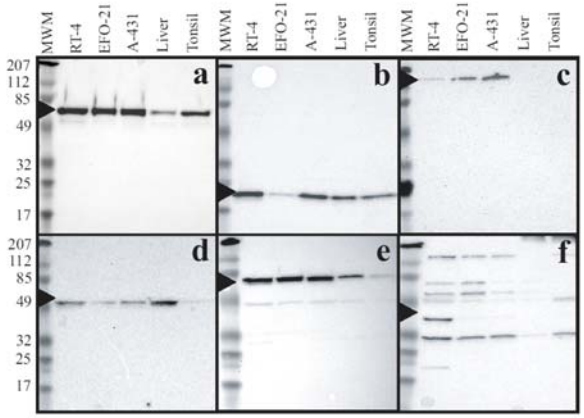
Mono-specific antibodies (msAb)





150 new antibodies every week

Western blots



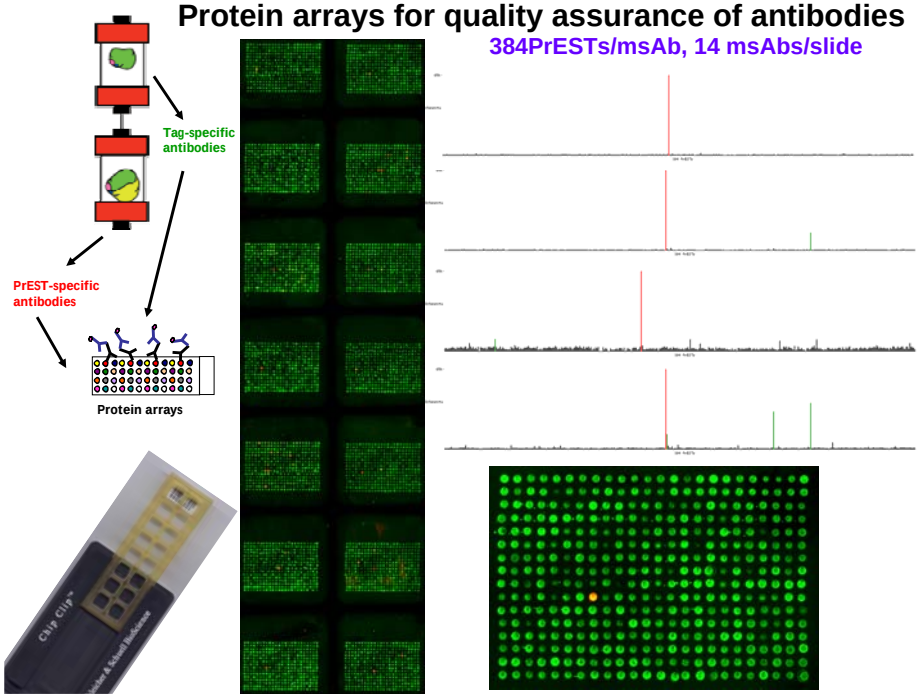
"Tissue Westerns" gives:

- ***Estimation of cross-reactivity***
- ***Semi-quantitative expression profile***
- ***Size variants (splice variants etc)***

The arrow shows the expected size (based on gene sequence prediction)

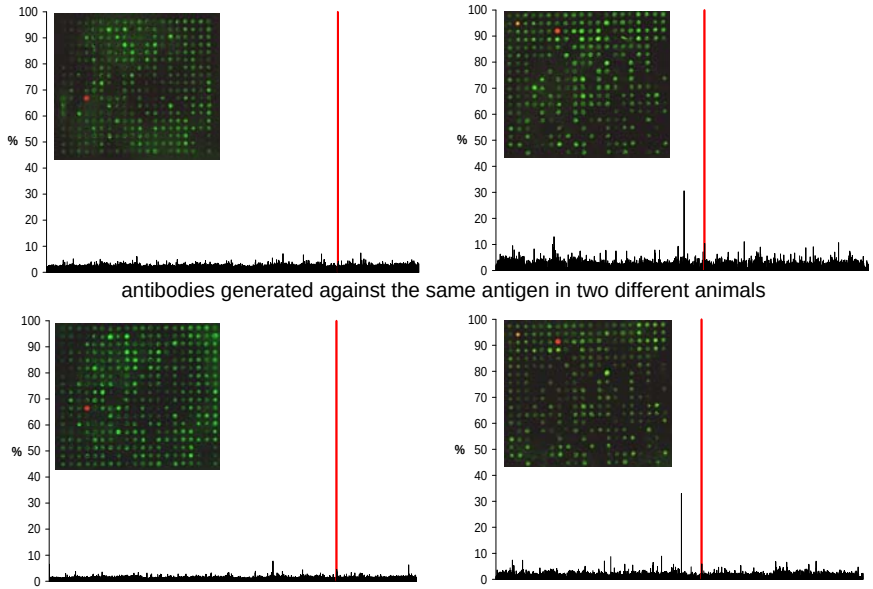
Protein arrays for quality assurance of antibodies

384PrESTs/msAb, 14 msAbs/slide



Protein microarrays for validation of HPR antibodies

1440 proteins spotted in triplicate



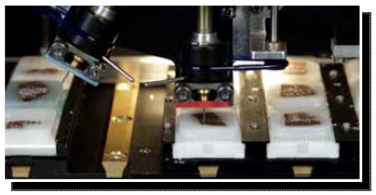
antibodies generated against the same antigen in two different animals

Tissue Microarrays

Collection of tissues

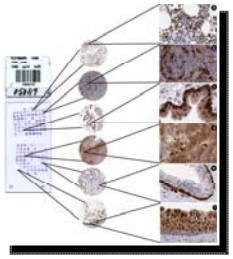
Annotation

TMA design



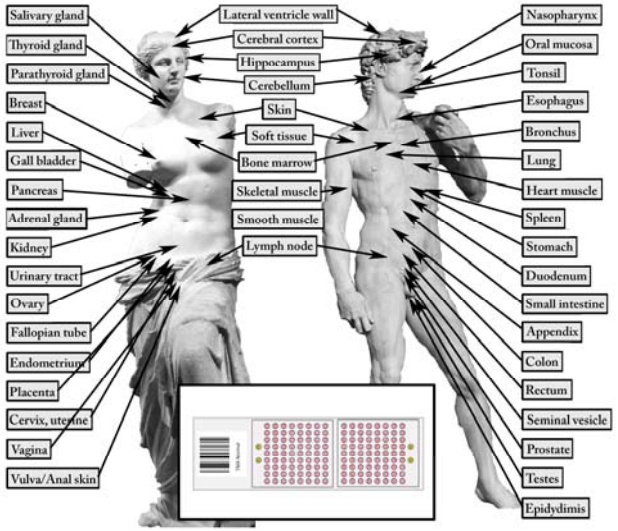
TMA construction

TMA sectioning



Analysis

Normal tissue profiling



Surgical specimens
Normal Morphology

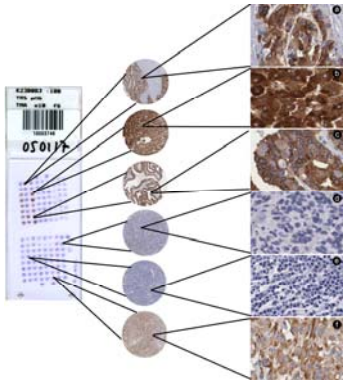
48 different tissue types
(triplicate samples)

Kampf et. al. *Clin. Proteomics 1 285-299*
Antibody-based tissue profiling as a tool in clinical proteomics (2004)

Cancer profiling

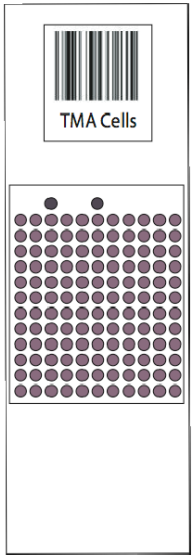
Breast cancer	12
Carcinoid	4
Cervical cancer	12
Colo-rectal cancer	12
Endometrial cancer	12
Glioma	12
Head-Neck cancer	4
Kidney cancer	12
Liver cancer	12
Lung cancer	12
Lymphoma	12
Melanoma	12
Ovarian cancer	12
Pancreatic cancer	12
Prostate cancer	12
Skin cancer	12
Stomach cancer	12
Testis cancer	12
Thyroid cancer	4
Urothelial cancer	12

20 cancer types



432 samples from 216 patients

Cell tissue microarray



HEMATOPOIETIC TUMORS

MYELOID
K-562 (Myeloid leukemia)
HL-60 (Promyelocytic leukemia)
NB-4 (Promyelocytic leukemia)
U937 (Myelomonocytic lymphoma)
THP-1 (Monocytic leukemia)
HEL (Erythroleukemia)
HMC-1 (Mastcell leukemia)

LYMPHOID
KM3 (B-cell leukemia)
MOLT-4 (Lymphoblastic leukemia)
Daudi (Burkitt lymphoma)
U-698 (B-cell lymphoma)
RPMI 8226 (Myeloma)
LP-1 (Myeloma)
Karpas 707 (Myeloma)
HDML-2 (Hodgkin lymphoma)
U-266 1970 (Myeloma)
U-266 1984 (Myeloma)

CLINICAL CELL SAMPLES
CLL (x4)
ALL (x2)
CML (x2)
AML (x2)
PBMC (x2)

SOLID TUMORS

BRAIN
U-87 MG (Glioblastoma)
U-138 MG (Glioblastoma)
U251 MG (Glioblastoma)
SH-SY5Y (Neuroblastoma)
D341 Med (Medulloblastoma)

LUNG
A-549 (Non small cell lung carcinoma)
SCLC-21H (Small cell lung carcinoma)

ABDOMINAL
Hep-G2 (Hepatocellular carcinoma)
CAPAN-2 (Pancreatic carcinoma)
CACO-2 (Colon carcinoma)

BREAST, FEMALE REPRODUCTIVE SYSTEM
SK-Br-3 (Breast carcinoma)
MCF-7 (Breast carcinoma)
AN3 CA (Endometrial carcinoma)
SiHa (Cervical carcinoma)
HeLa (Cervical carcinoma)
EFO-21 (Ovarian carcinoma)
BEWO (Choriocarcinoma)

URINARY, MALE REPRODUCTIVE SYSTEM
PC-3 (Prostate carcinoma)
NTERA-2 (Testicular carcinoma)
HEK 293 (Embryonal kidney cell line)
RT-4 (Urinary bladder carcinoma)

SKIN
A431 (Epidermoid carcinoma)
HaCaT (Keratinocyte cell line)
SK-MEL-30 (Melanoma)
WM 115 (Melanoma)

SARCOMA
U-2 OS (Osteosarcoma)
RH-30 (Rhabdomyosarcoma)
U-2197 (Histiocytic sarcoma)

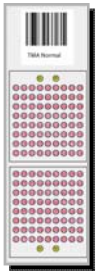
MISCELLANEOUS
HTH-83 (Thyroid carcinoma)
TIME-cell (Endothelial cell line)

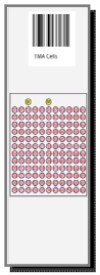
hpr

48 Normal tissues | 20 Cancer tissues | 12 Cells and 46 Cell lines

Surgical specimens

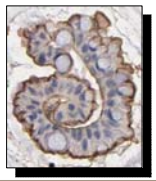
Normal morphology



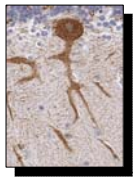


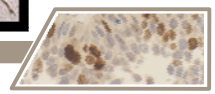
Clinical blood cells

432 samples in duplicate from 216 patients



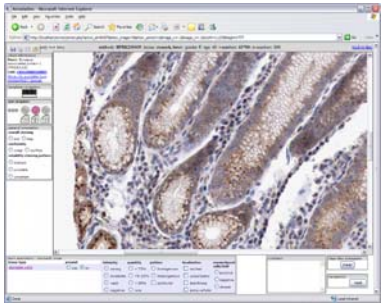
708 TMA cores/images for each antibody > 200 Gb/day





Annotation – a major challenge

How to annotate all the generated immunohistochemical images?



All images are looked at by an experienced pathologist (576 images/antibody).

Annotation of pattern, distribution, intensity, localisation

Development of web-based annotation software

The Mumbai HPR team

Dr. Sanjay Navani

- ***10 pathologists***
- ***All images available through internet***
- ***Web-based annotation system***

hpr

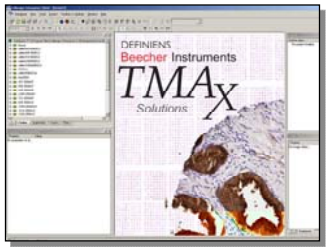
Curation

Tissues

Manual annotation

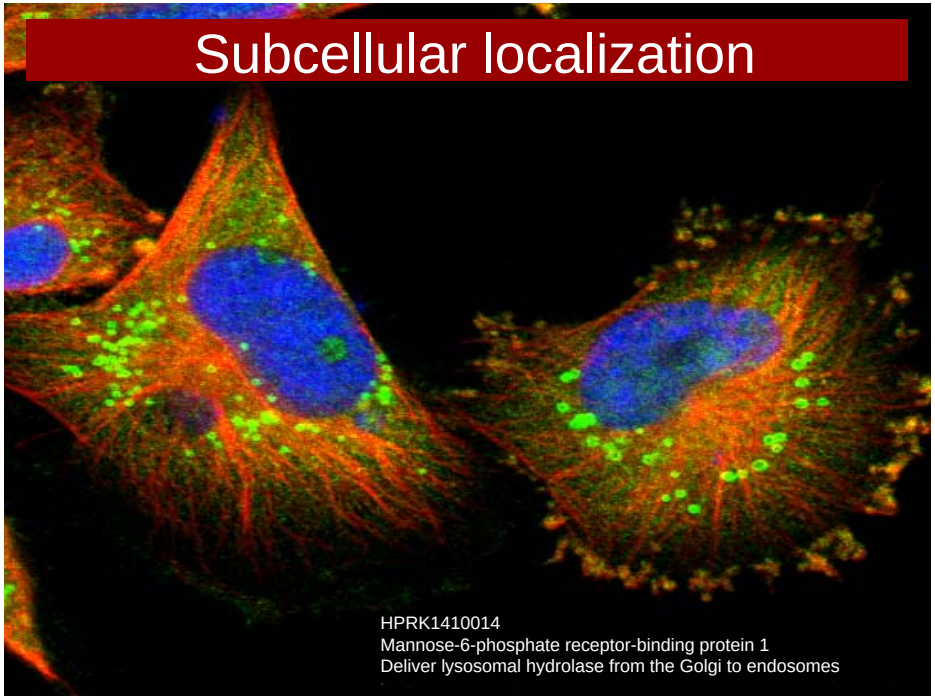
Cells

Automated annotation



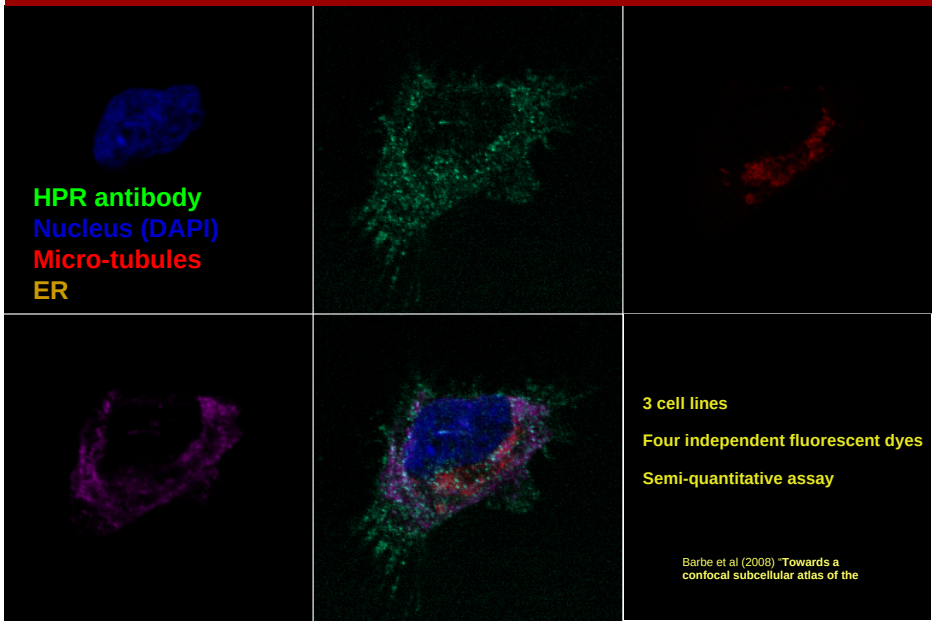
- ***>10,000 images annotated per day***
- ***Six software engineers to support the work flow***

Subcellular localization

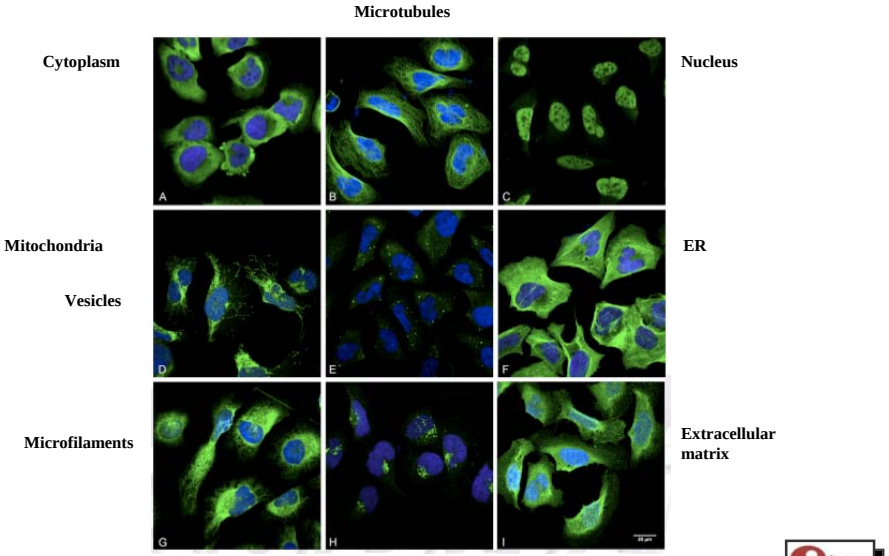


HPRK1410014
Mannose-6-phosphate receptor-binding protein 1
Deliver lysosomal hydrolase from the Golgi to endosomes

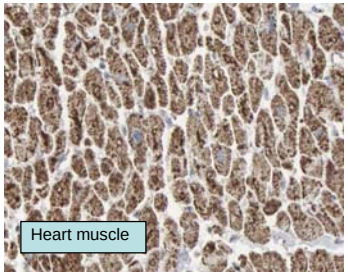
Confocal microscopy with multiple dyes

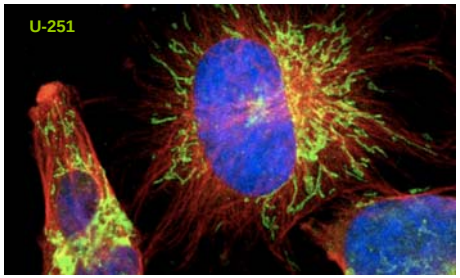


Subcellular structures

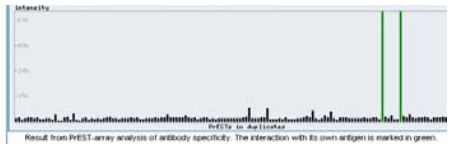


Stress 70 protein - "mortalin" (HPA000898)

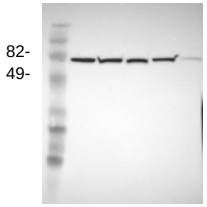




Mitochondria (supported by literature)

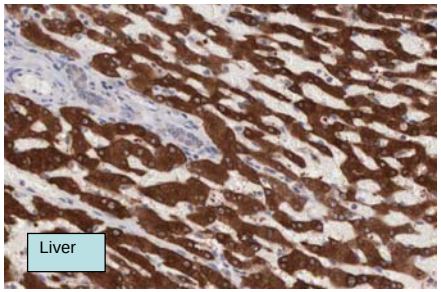


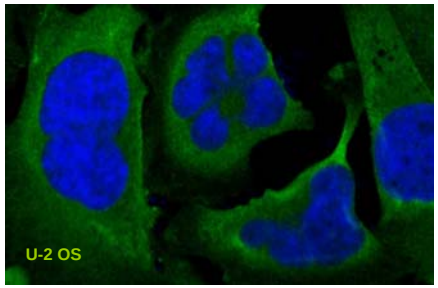
Protein array - specific



Mw according to gene prediction: 74K

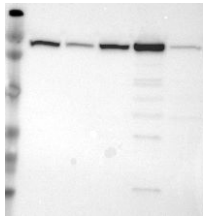
MTHFD1 (HPA000704)





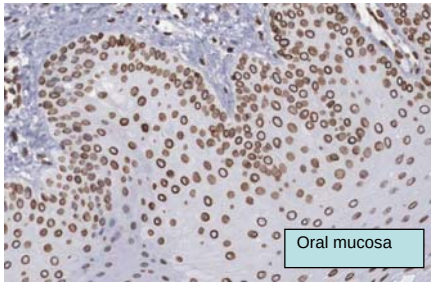
Cytoplasmic (supported by literature)

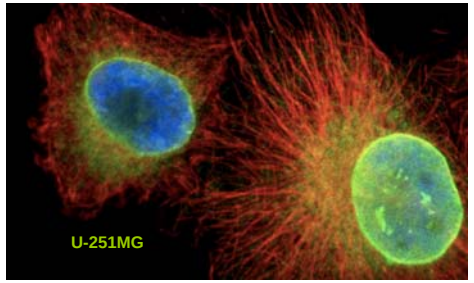
Protein array - specific



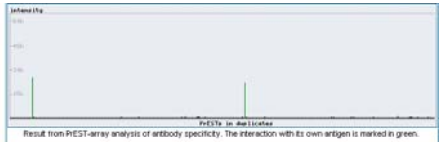
Two splice forms (25K and 29K)

Emerin (HPA000609)

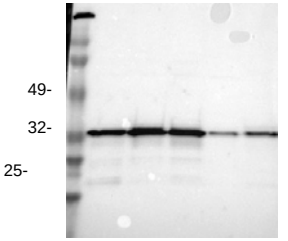




Nuclear membrane (supported by literature)

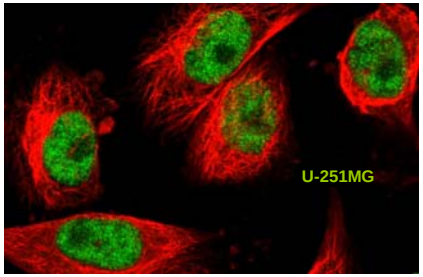


Protein array - specific

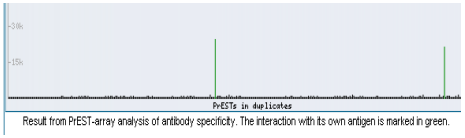


Mw according to gene prediction: 25K and 29K

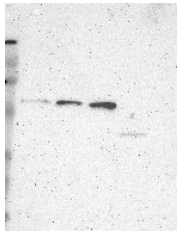
G patch domain and KOW motifs protein (HPA000287)



Nuclear (no literature available)



Protein array - specific

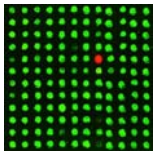


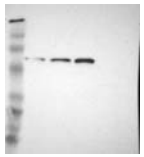
Two splice forms (both 52K)

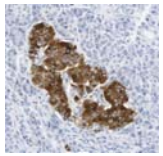
hpr

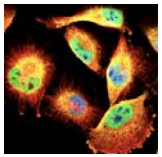
Antibody quality assessment

Protein array | Western blot | IHC | Immunofluorescence | Literature







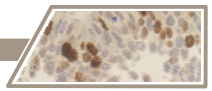


UniProt
the universal protein resource

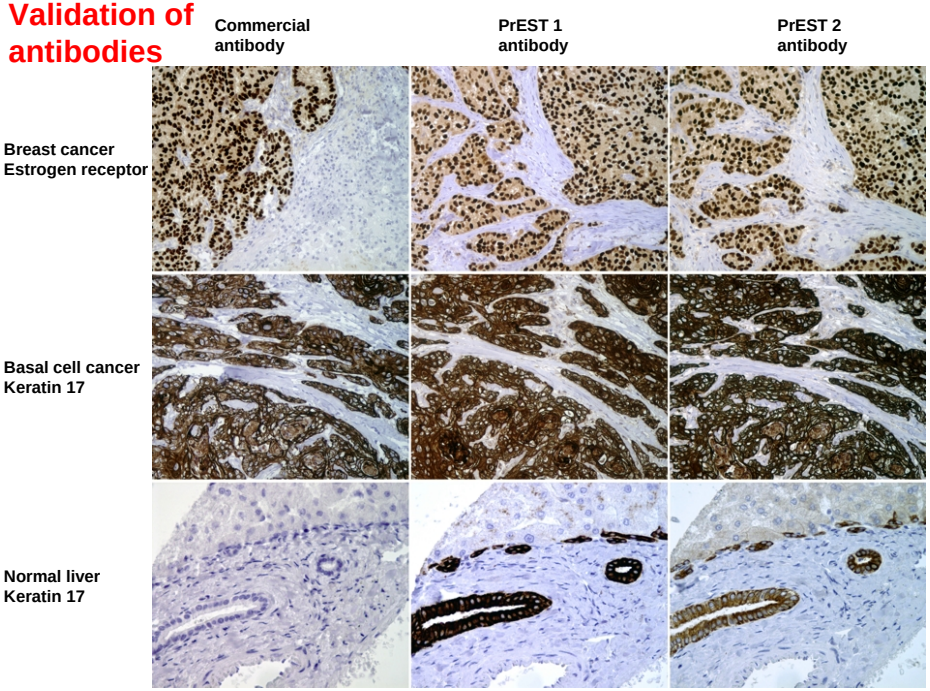
20 years swissprot

Validation score:

High	Two independent antibodies with same staining patterns
Medium	Consistent with bioinfomatics and litterature (if available)
Low	At least some evidence that supports the staining patterns
Very low	No evidence that the antibody is correct
Failed	The antibody is (most likely) not correct



Validation of antibodies



HPR status (June 13, 2008)

≈ 18.300 genes initiated (~60% of all human genes)

≈ 36.800 PrESTs designed

≈ 24.400 PrEST clones delivered (sequence verified)

≈ 18.500 (13.100 unique) antigens sent for immunizations

≈ 10.600 array verified antibodies

≈ 9.700 antibodies analyzed on western blot

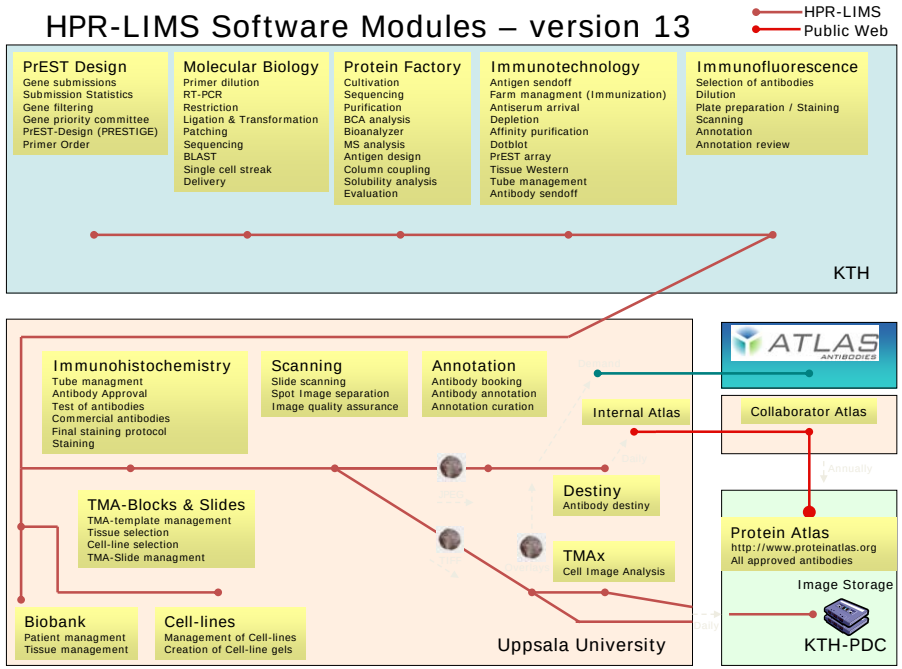
≈ 7.000 annotated and curated antibodies (4.600 HPR ab)

Internal atlas content: 7.065 antibodies (5604 genes)

4,069,440 images

~10 new validated antibodies every day

HPR-LIMS Software Modules – version 13



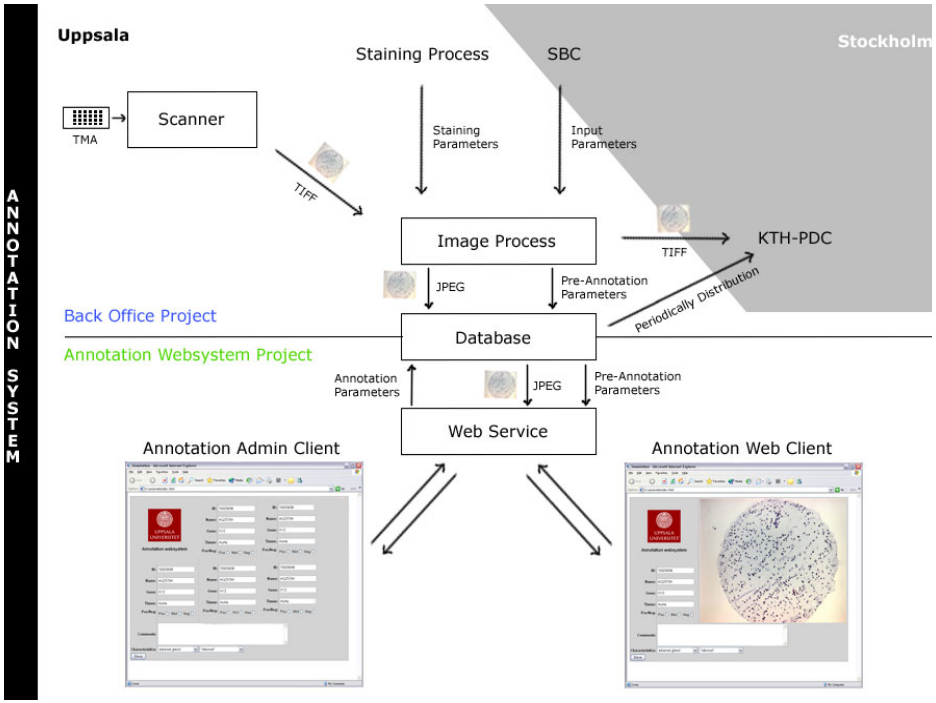
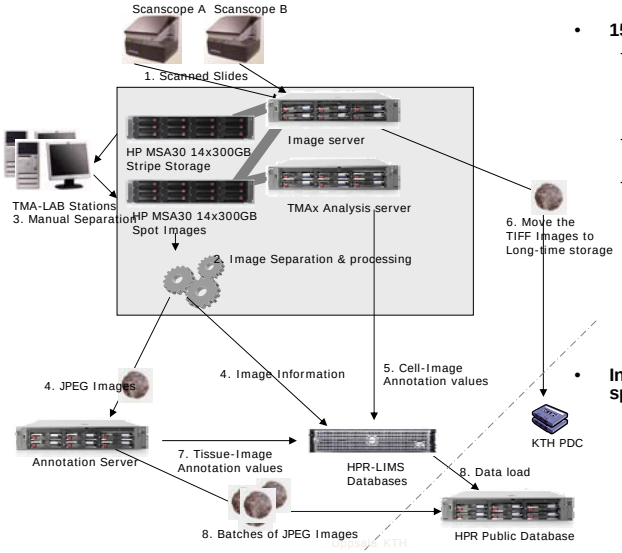
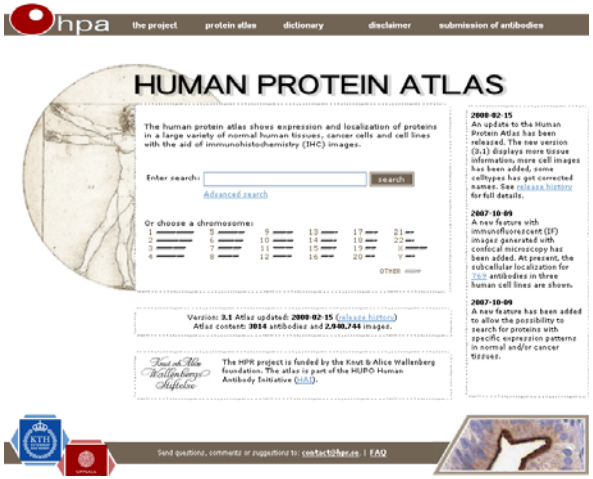


Image Systems *v5*

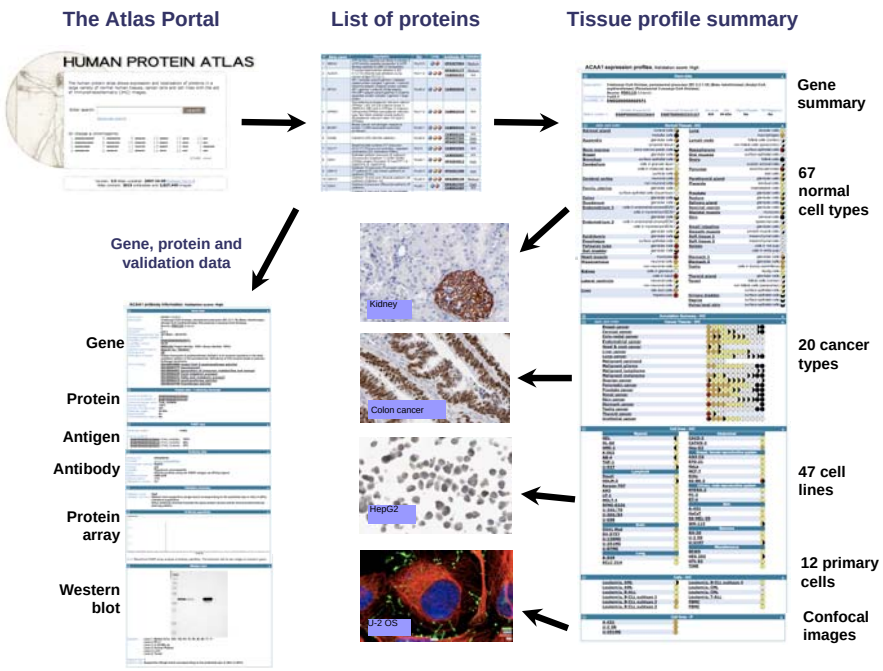


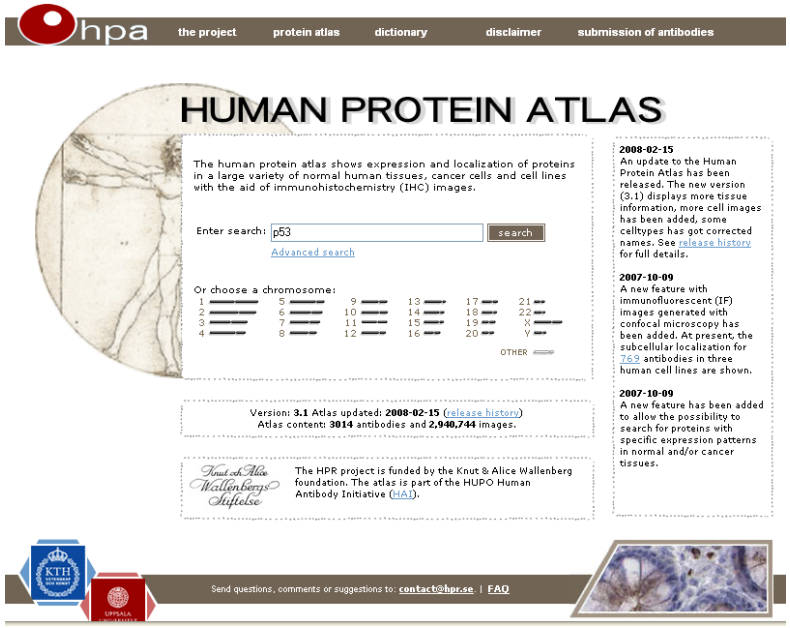
- 15 Servers handling all the data
  - KTH-PDC
    - 2 load balancers
    - 3 web/database
    - 1 image server
  - KTH-Albanova
    - 1 image backup server
  - UU-Rudbeck
    - HPR-LIMS
    - Annotation/HPR-external
    - Image server
    - TMAx-server
    - HPR-DEV1
    - HPR-DEV2
    - HPR-Demo
    - Oldannotations
- In total 21 Terabytes (!!!) of harddisc space on 70 harddrives

Erik Björling 2006-08-24



1st release Aug 29 2005	*HUPO Munich*	*(718 antibodies)*
2nd release Oct 30 2006	*HUPO Long Beach*	*(1514 antibodies)*
3rd release Oct 08 2007	*HUPO Seoul*	*(3015 antibodies)*
4th release Aug 18 2008	*HUPO Amsterdam*	*(~6000 antibodies)*
5th release Sep 28 2009	*HUPO Toronto*	*(~9000 antibodies??)*
6th release Sep 21 2010	*HUPO Sydney*	*(~12000 antibodies??)*
7th release Aug 27 2011	*HUPO Geneva*	*(~15000 antibodies??)*





Normal tissues - p53

TP53 in normal tissues. Reliability: N/A

Annotation Summary

Normal cells were unstained, except some scattered moderately stained nuclei in proliferating epithelia as seen in gastrointestinal canal and in basal layer of squamous epithelium. In malignant tumors a distinct nuclear staining was seen in many cases, but not in all tumor types. In none of the tumor groups were all individual cases positive. If positive, either all tumor nuclei were strongly stained or some of the nuclei were moderately or weakly stained. Highly malignant tumor types were more often stained than low malignant types. For example, prostate carcinomas were all negative. Some highly malignant cases were however also negative, as lymphomas and renal cell carcinomas. Tumor stroma was generally negative.

Adrenal gland	cortical cells	Lung	alveolar cells
	medullar cells		macrophages
Appendix	glandular cells	Lymph node	follicle cells (cortex)
	lymphoid tissue		non-follicle cells (paracortex)
Bone marrow	bone marrow poetic cells	Nasopharynx	surface epithelial cells
Breast	glandular cells	Oral mucosa	surface epithelial cells
Bronchus	surface epithelial cells	Ovary	follicle cells
Cerebellum	cells in granular layer		ovarian stromal cells
	cells in molecular layer	Pancreas	exocrine pancreas
	purkinje cells		islet cells
Cerebral cortex	neuronal cells	Parathyroid gland	glandular cells
	non-neuronal cells	Placenta	decidual cells
Cervix, uterine	glandular cells		trophoblastic cells
	surface epithelial cells (squamous)	Prostate	glandular cells
Colon	glandular cells	Rectum	glandular cells
Duodenum	glandular cells	Salivary gland	glandular cells
Endometrium 1	cells in endometrial stroma/ECM	Seminal vescicle	glandular cells
	cells in myometrium/ECM	Skeletal muscle	myocytes
	glandular cells	Skin	adnexal cells
Endometrium 2	cells in endometrial stroma/ECM		epidermal cells
	cells in myometrium/ECM	Small intestine	glandular cells
	glandular cells	Smooth muscle	smooth muscle cells
Epididymis	glandular cells	Soft tissue 1	mesenchymal cells
Esophagus	surface epithelial cells	Soft tissue 2	mesenchymal cells
Fallopian tube	glandular cells	Spleen	cells in red pulp
Gall bladder	glandular cells		cells in white pulp
Heart muscle	myocytes	Stomach 1	glandular cells
Hippocampus	neuronal cells	Stomach 2	glandular cells
	non-neuronal cells	Testis	cells in ductus seminiferus
Kidney	cells in glomeruli		leydig cells
	cells in tubuli	Thyroid gland	glandular cells
Lateral ventricle	neuronal cells	Tonsil	follicle cells (cortex)
	non-neuronal cells		non-follicle cells (paracortex)
Liver	bile duct cells		surface epithelial cells
	hepatocytes	Urinary bladder	surface epithelial cells
		Vagina	surface epithelial cells
		Vulva/anal skin	surface epithelial cells

Strong staining
Medium staining
Weak staining
Negative staining

Not representative

Cancer view

TP53 in cancer tissues. Reliability: N/A

Annotation Summary

Normal cells were unstained, except some scattered moderately stained nuclei in proliferating epithelia as seen in gastrointestinal canal and in basal layer of squamous epithelium. In malignant tumors a distinct nuclear staining was seen in many cases, but not in all tumor types. In none of the tumor groups were all individual cases positive. If positive, either all tumor nuclei were strongly stained or some of the nuclei were moderately or weakly stained. Highly malignant tumor types were more often stained than low malignant types. For example, prostate carcinomas were all negative. Some highly malignant cases were however also negative, as lymphomas and renal cell carcinomas. Tumor stroma was generally negative.

Breast cancer	
Cervical cancer	
Colo-rectal cancer	
Endometrial cancer	
Head & neck cancer	
Kidney cancer	
Liver cancer	
Lung cancer	
Malignant carcinoid	
Malignant glioma	
Malignant lymphoma	
Malignant melanoma	
Ovarian cancer	
Pancreatic cancer	
Prostate cancer	
Skin cancer	
Stomach cancer	
Testis cancer	
Thyroid cancer	
Urothelial cancer	

Strong staining
Medium staining
Weak staining
Negative staining

Not representative

Colo-rectal cancer [TP53]

Annotation Data

Overall staining: 12/12

Tumor cells

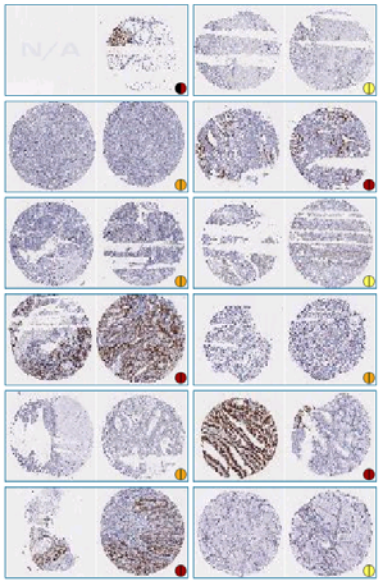
Intensity		Quantity	
Strong:	6/12	>75%:	5/12
Moderate:	3/12	75%-25%:	1/12
Weak:	3/12	<25%:	6/12
Negative:	0/12	Rare:	0/12

Localization
Nuclear: 12/12
Cytoplasmic/Membrane: 0/12

Tissue data

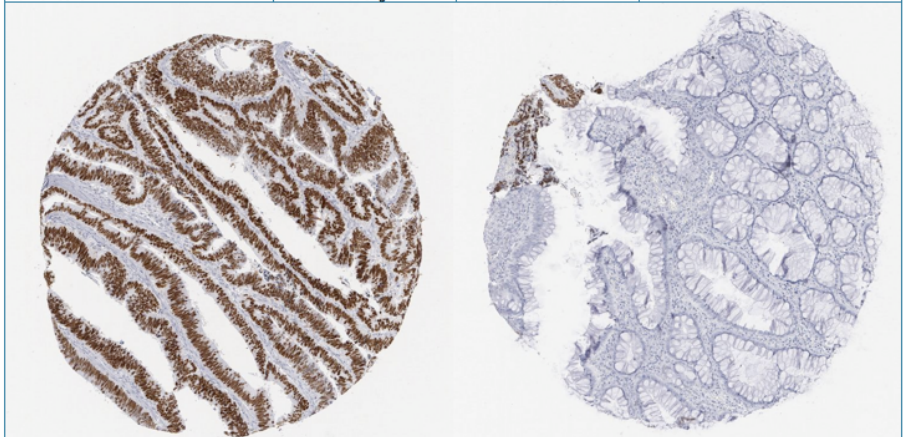
Male:	6/12	Age >60:	9/12
Female:	6/12	Age >25, <60:	2/12
		Age <25:	1/12

> Clear <



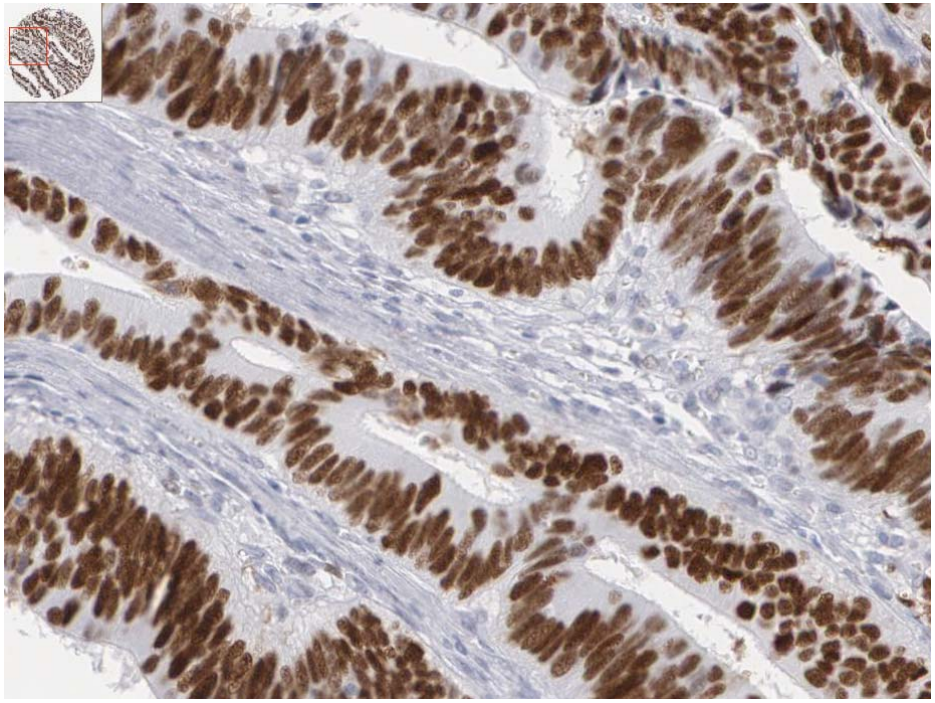
Colo-rectal cancer [TP53]

Cell Type	Intensity	Quantity	Localization
Tumor cells	strong	>75%	nuclear

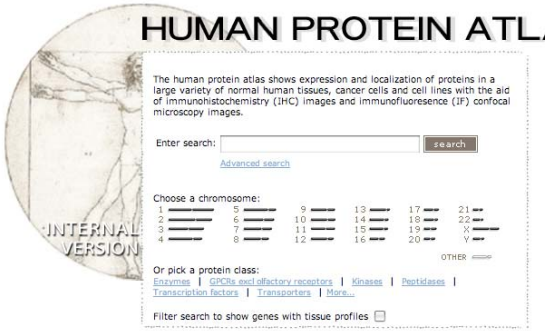


Male, age 72 | Male, age 72

Brown color indicates presence of protein, blue color shows cell nuclei. Image Usage Policy



Version 4.0 - will be launch August 18, 2008



- *6,000 antibodies corresponding to 5,200 genes*
- *>5 million images (majority manually annotated)*
- *Double the content from previous version (25% of all genes)*

Protein classes

Protein Classes

Main classes

Protein class	No of genes	Code	No of genes with tissue profiles	Source
Kinases	503	Ki	277	UniProt - Swiss-Prot Protein Knowledgebase
CD markers	365	Cd	222	UniProt - Swiss-Prot Protein Knowledgebase
Blood group antigen proteins	29	Ba	14	UniProt - Swiss-Prot Protein Knowledgebase
Transporters	430	Tr	156	Transport Classification Database (TCDB)
Ribosomal proteins	160	Ri	30	UniProt - Swiss-Prot Protein Knowledgebase
Peptidases	500	Pe	189	Merops / UniProt - Swiss-Prot Protein Knowledgebase
G-protein coupled receptors	762	Gr	127	UniProt - Swiss-Prot Protein Knowledgebase
Transcription factors	1443	Tf	424	DBD: Transcription factor prediction database
Enzymes	2493	Ez	980	ENZYME nomenclature database

Project related classes

Protein class	No of genes	Code	No of genes with tissue profiles	Source
Candidate cancer biomarkers	1064	Cb	674	Plasma Proteome Institute
Plasma proteins	682	Pp	384	Plasma Proteome Institute
Candidate cardiovascular disease genes	164	Cc	110	Plasma Proteome Institute

- *Allow searches on protein classes*
- *Also project related protein classes*

All genes included

Search Results

Search results for protein class: **503 hits**

kinases

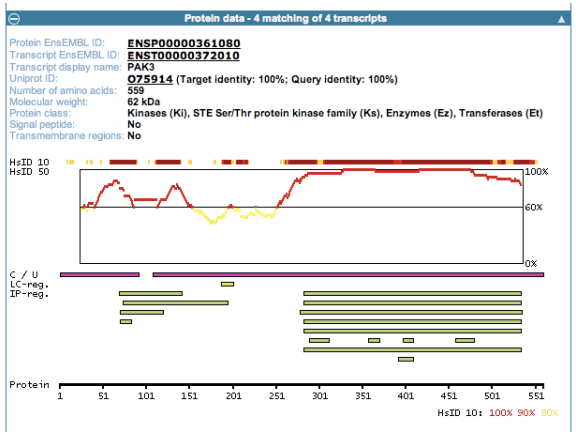
Choose, if available:
- an Antibody ID to view the annotation data
- a link button to open a new window with Ensembl/NCBI/Uniprot info

#	Gene name	Description	Chr	Links	Class	Antibody ID	Validation
301	NUAK1	NUAK family SNF1-like kinase 1 (EC 2.7.11.1) (AMPK-related protein kinase 5).	12:q23.3		Kc Ez Ki Et		
302	NUAK2	NUAK family SNF1-like kinase 2 (EC 2.7.11.1) (SNF1/AMP kinase-related kinase) (SNARK).	1:q32.1		Kc Ez Ki Et	HPA008958	Medium
303	OBSCN	Obscurin (Obscurin-myosin light chain kinase) (Obscurin-MLCK) (Obscurin-RhoGEF).	1:q42.13		Kc Ez Ki Et		
304	OXSR1	Serine/threonine-protein kinase OSR1 (EC 2.7.11.1) (Oxidative stress- responsive 1 protein).	3:p22.3		Ez Ki Kc Et	HPA008237 CAB017181	Medium N/A
305	PAK1	Serine/threonine-protein kinase PAK 1 (EC 2.7.11.1) (p21-activated kinase 1) (PAK-1) (p65-PAK) (Alpha-PAK).	11:q13.5		Ez Ki Kc Et	HPA003565 CAB005312	Medium N/A
306	PAK2	Serine/threonine-protein kinase PAK 2 (EC 2.7.11.1) (p21-activated kinase 2) (PAK-2) (Gamma-PAK) (PAK65) (S6/H4 kinase).	3:q29		Ez Ki Kc Et	CAB007794	N/A
307	PAK3	Serine/threonine-protein kinase PAK 3 (EC 2.7.11.1) (p21-activated kinase 3) (PAK-3) (Beta-PAK) (Oligophrenin-3).	X:q23		Ez Ki Kc Et		

Gene protein information

Tissue profile

Protein data visualization



Protein summary

Homology (10 window size)

Homology (50 window size)

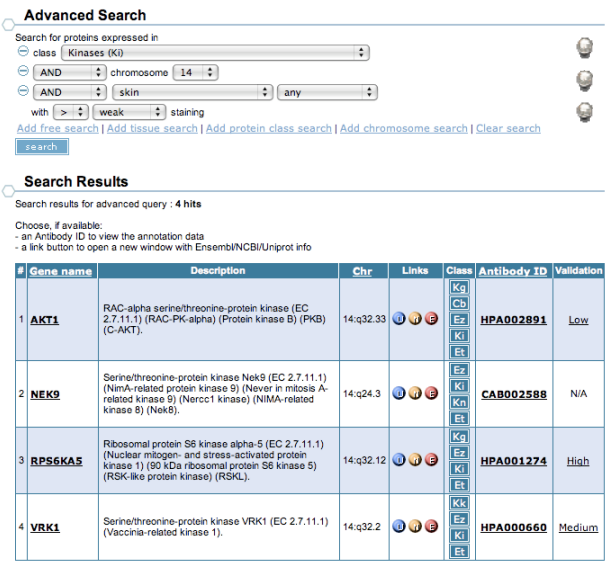
Splice variants
Low complexy regions

InterPro regions

Protein (amino acids)

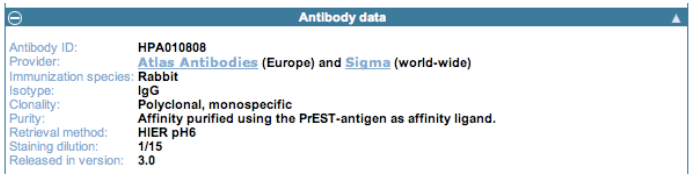
- ***Shown for all (20,488) genes (PRESTIGE software)***

Advanced queries



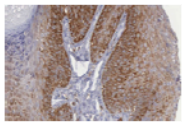
- ***Kinases***
- ***Chromosome 14***
- ***Expressed in skin***

HPA antibodies are available to the public



Prestige Antibodies™

- Available through Atlas Antibodies in Europe
- Available through Sigma-Aldrich worldwide



Public interest/User statistics

Top 25 countries during 2007

Countries	Pages	Hits	Bandwidth
United States	1425838	5862080	1630.19 GB
Germany	280384	711065	15.73 GB
Sweden	135332	1120823	18.35 GB
Great Britain	57920	556328	9.07 GB
Canada	36949	278617	2.92 GB
Japan	33884	291003	13.40 GB
Netherlands	28846	311513	4.53 GB
Denmark	25086	139058	2.10 GB
South Korea	19996	292802	1.83 GB
France	19694	184763	2.85 GB
China	19057	201179	2.59 GB
Spain	12847	140282	1.91 GB
Australia	12489	107181	1.60 GB
Unknown	12274	140637	1.09 GB
Italy	11861	103517	1.52 GB
Finland	11204	132129	2.11 GB
Norway	10295	89272	1.98 GB
Switzerland	10168	92347	1.78 GB
Belgium	8471	89621	1.49 GB
Brazil	8369	51459	745.57 MB
India	8111	93958	1.16 GB
Austria	7429	80987	1.37 GB
Taiwan	6480	79302	967.24 MB
Ireland	6258	56764	890.11 MB
Ukraine	4104	6502	178.07 MB

2007:
Approximately 13,000 visits/month
Visitors from 174 ”countries”
> 2,2 x 10^6 visited pages
>1,7 Tbytes of data downloaded

26-65	66-105	106-145	146-174
Singapore	Vietnam	Guam (USA)	Cameroon
New Zealand	Venezuela	Guatemala	Botswana
Portugal	Morocco	Dominican Republic	Saint Lucia
Turkey	Cambodia	Faroe Islands	Andorra
Israel	Latvia	Bahamas	Dominica
Romania	Cyprus	Mauritius	Cayman Islands
Poland	Cuba	Ghana	Reunion (French)
Greece	Costa Rica	Malawi	Virgin Islands (British)
Russian Federation	Sri Lanka	Nicaragua	Paraguay
Czech Republic	Nigeria	Georgia	Liechtenstein
Mexico	Syria	Panama	Saint Kitts & Nevis Anguilla
Estonia	Senegal	Albania	Micronesia
Hong Kong	Macedonia	Honduras	Polynesia (French)
Argentina	Bosnia-Herzegovina	Zambia	Madagascar
Bulgaria	Bangladesh	Moldova	Azerbaidjan
South Africa	European country	Grenada	Bermuda
Hungary	Algeria	Antigua and Barbuda	Uzbekistan
Malaysia	Trinidad and Tobago	Tanzania	Cape Verde
Philippines	Palestinian Territories	El Salvador	Zimbabwe
Thailand	Qatar	Fiji	Rwanda
Iran	Jamaica	Virgin Islands (USA)	Burkina Faso
Iceland	Bahrain	Ethiopia	Papua New Guinea
Pakistan	Sudan	Macau	Kyrgyzstan
Egypt	Yemen	Belize	Armenia
Slovak Republic	Lebanon	Aruba	Northern Mariana Islands
Slovenia	Malta	Uganda	Guyana
Chile	Maldives	Bolivia	Somalia
Indonesia	Ecuador	Kazakhstan	Tadjikistan
United Arab Emirates	Netherlands Antilles	Gambia	
Saudi Arabia	Kenya	Monaco	
Luxembourg	Belarus	Laos	
Croatia	Uruguay	Bhutan	
Lithuania	Libya	Mongolia	
Kuwait	Namibia	Saint Vincent & Grenadines	
Jordan	Brunei Darussalam	Ivory Coast (Cote D'Ivoire)	
Puerto Rico	Mozambique	Benin	
Colombia	Tunisia	San Marino	
Nepal	Barbados	Myanmar	
Peru	Iraq	New Caledonia (French)	
Satellite access host	Oman	Suriname	

In silico biomarker discovery

Advanced Search

Search for proteins expressed in

breast cancer | tumor cells | with at least 8 | patient(s)with | = | strong | staining

Add free search | Add tissue search | Clear search

search

Search Results

Search results for advanced query : **247 hits** (genes)

Advanced Search

Search for proteins expressed in

breast cancer | tumor cells | with at least 8 | patient(s)with | = | strong | staining

AND | breast | glandular cells | with | < | moderate | staining

Add free search | Add tissue search | Clear search

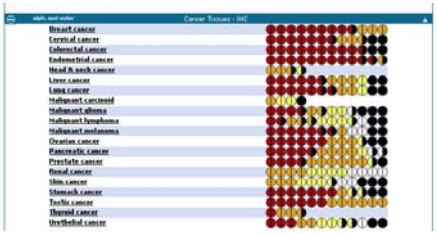
search

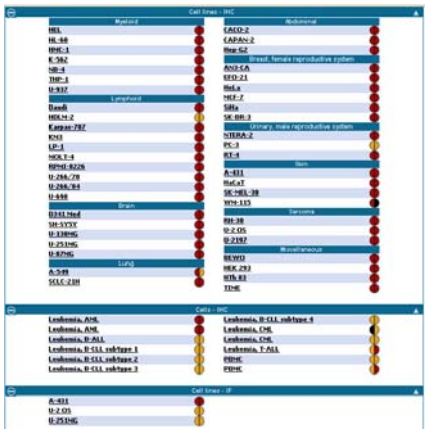
Search Results

Search results for advanced query : **5 hits** (genes)

Choose, if available:
- an Antibody ID to view the annotation data
- a link button to open a new window with Ensembl/NCBI/Uniprot info

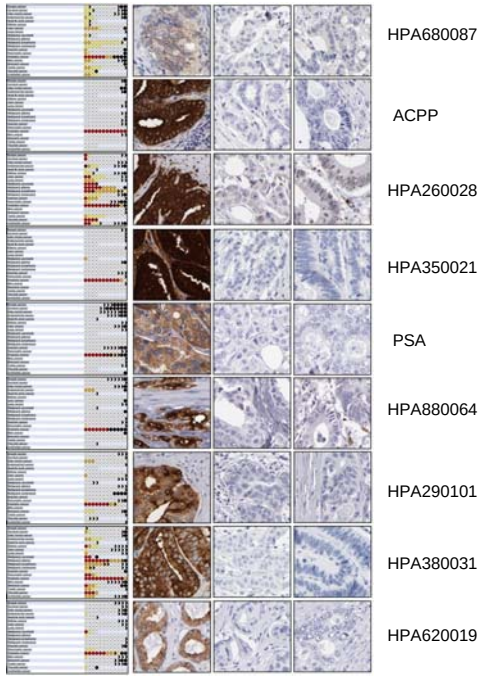
#	Gene name	Description	Chr	Links	Antibody ID	Validation
1	CLPP	Putative ATP-dependent Clp protease proteolytic subunit, mitochondrial precursor (EC 3.4.21.92) (Endopeptidase Clp).	19:p13.3	U N E	HPA010649	Medium
2	MPO	Myeloperoxidase precursor (EC 1.11.1.7) (MPO) [Contains: 89 kDa myeloperoxidase; 84 kDa myeloperoxidase; Myeloperoxidase light chain; Myeloperoxidase heavy chain].	17:q23.2	U E	CAB000059	N/A
3	MRPL40	39S ribosomal protein L40, mitochondrial precursor (L40mt) (MRP-40) (Nuclear localization signal-containing protein deleted in velocardiofacial syndrome) (Up-regulated in metastasis).	22:q11.21	U N E	HPA006181	High
4	MYOM2	Myomesin-2 (M-protein) (165 kDa titin-associated protein) (165 kDa connectin-associated protein).	8:p23.3	U N E	HPA001765	Low
5	USP51	Ubiquitin carboxyl-terminal hydrolase 51 (EC 3.1.2.15) (Ubiquitin thioesterase 51) (Ubiquitin-specific-processing protease 51) (Deubiquitinating enzyme 51).	X:p11.21	U N E	HPA001942	Low



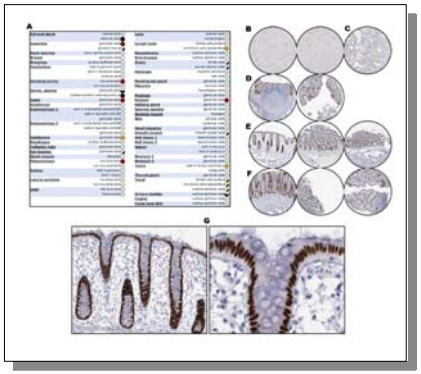


In silico biomarker discovery

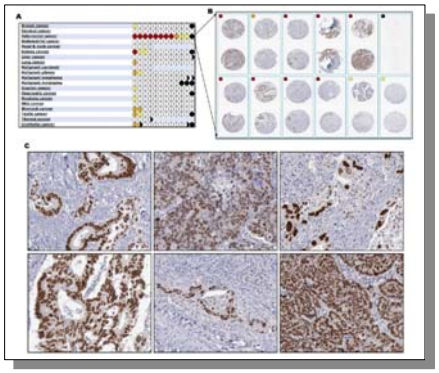
Prostate cancer



Biomarker discovery – prognostics



Normal



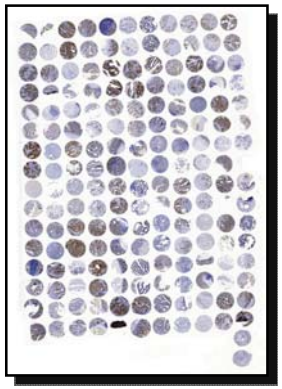
Colon cancer

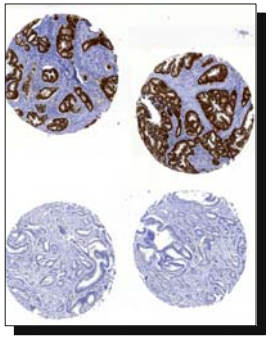
Validation - Colon cancer

Median age 75 years (32-88)

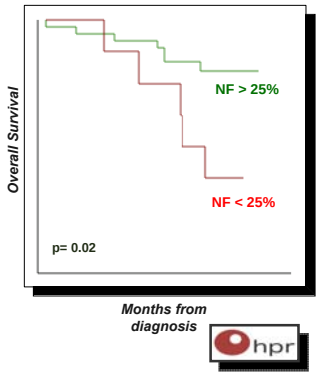
Count

		DUKES			
		A	B	C	Total
SEX	female	21	24	18	63
	male	18	18	23	59
Total		39	42	41	122



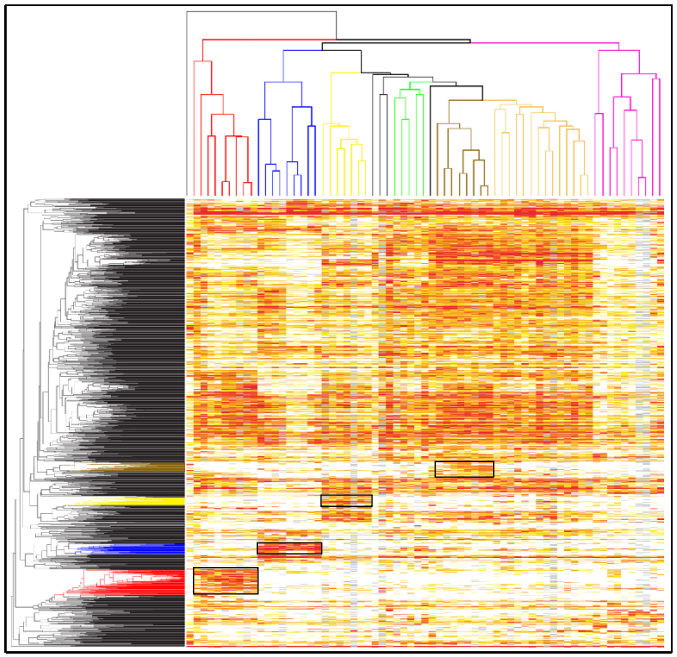


Survival



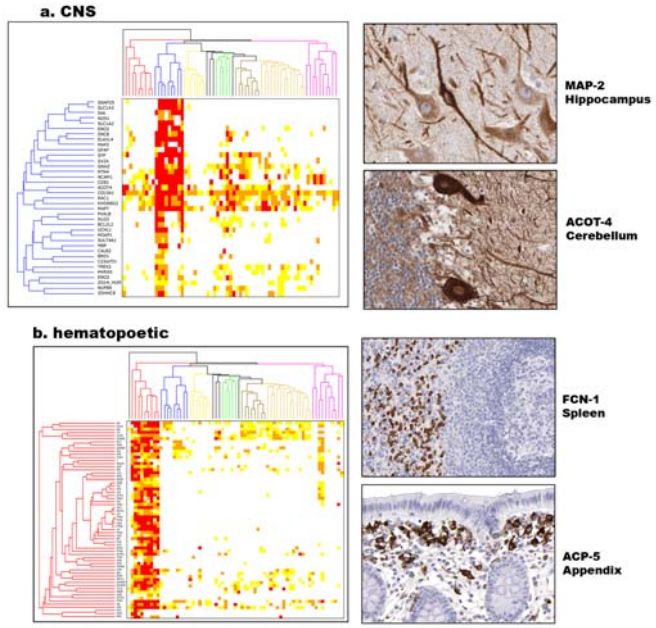
visualisation of all protein expressions and localisations

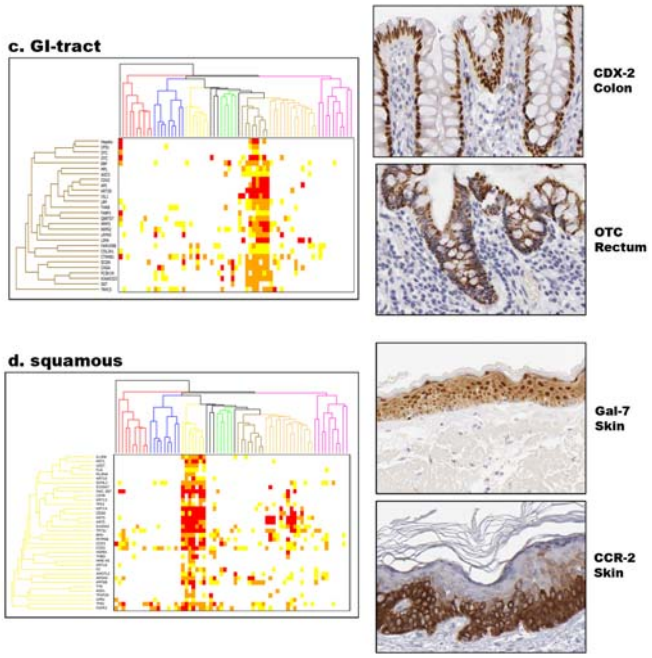
> 500.000 images
> 1025 antibodies



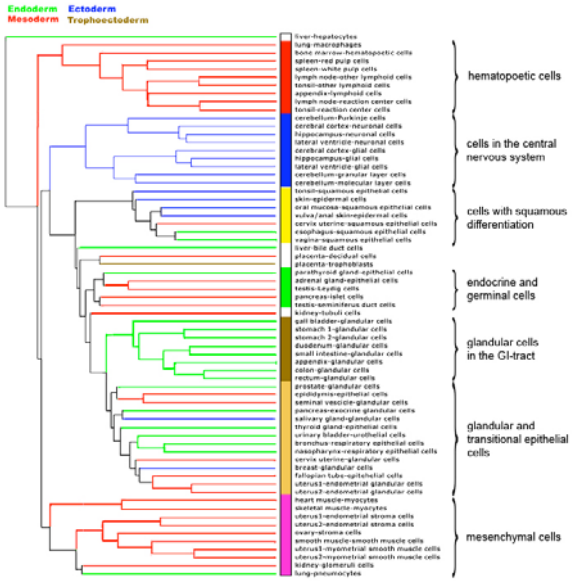
1. Lymphoid
2. Brain
3. Squamous
4. GI-tract

Protein composition typical of CNS and hematopoietic cells

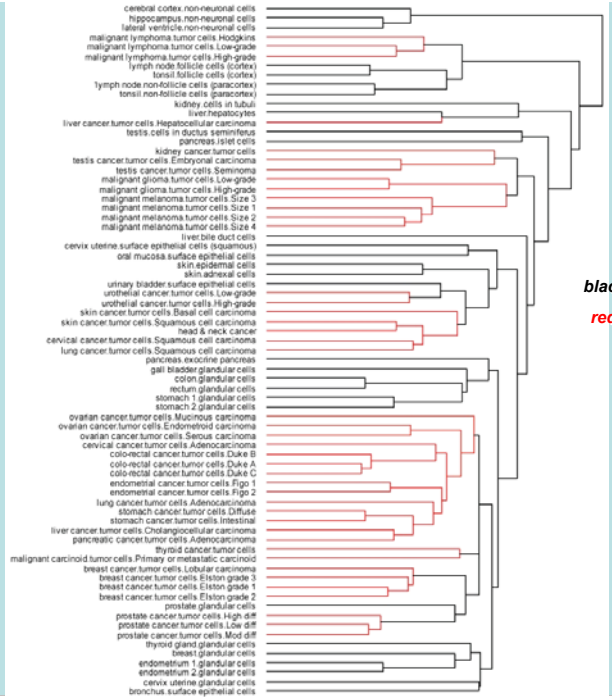




Clustering of cell types and origin from embryonal germ cell layers



Cell types

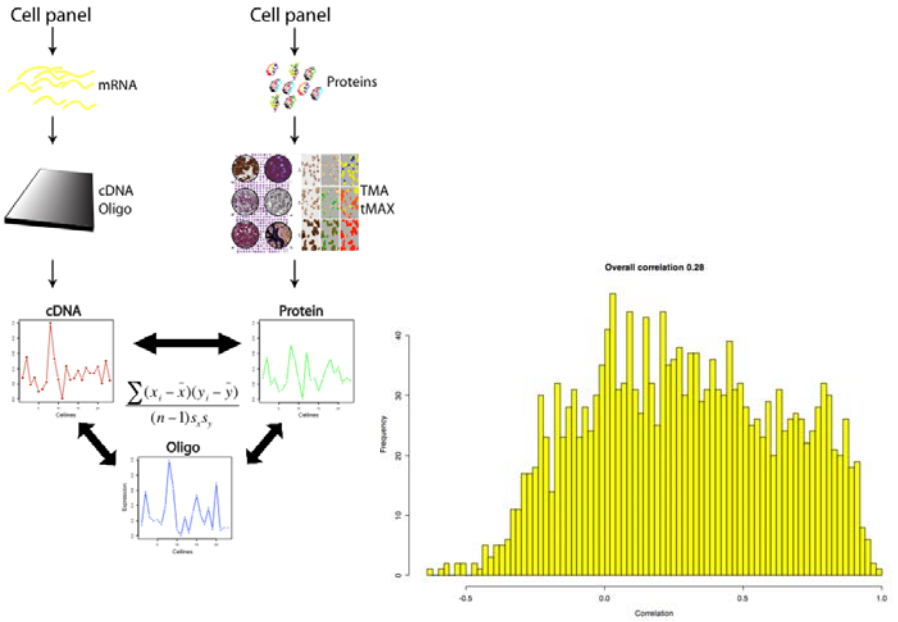


Dendrogram
Cells from normal and cancer tissues

black = normal cells

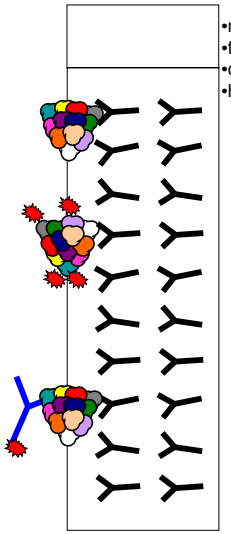
red = cancer cells

Comparative RNA and protein expression analysis



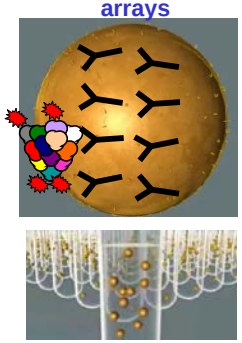
Protein microarrays for serum anaylysis

antibody microarrays

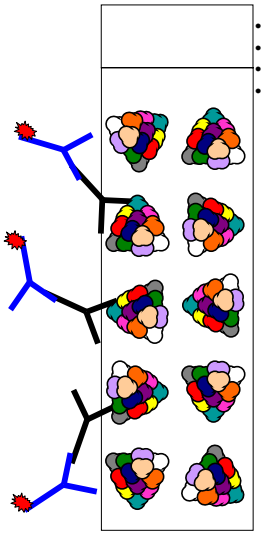


- many analytes
- few samples
- complex detection
- high sensitivity

suspension bead arrays

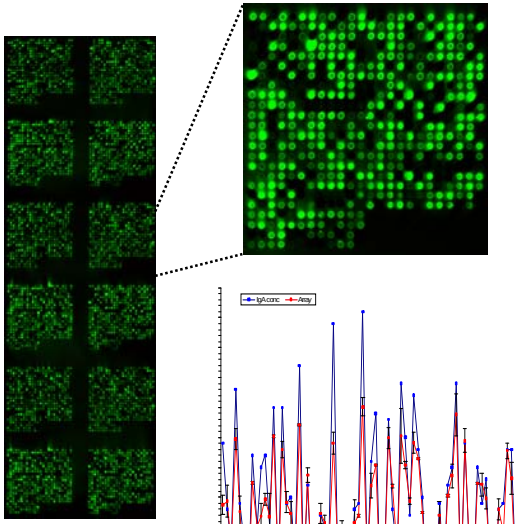


reverse phase serum microarrays



- many samples
- few analytes
- simple detection
- medium sensitivity

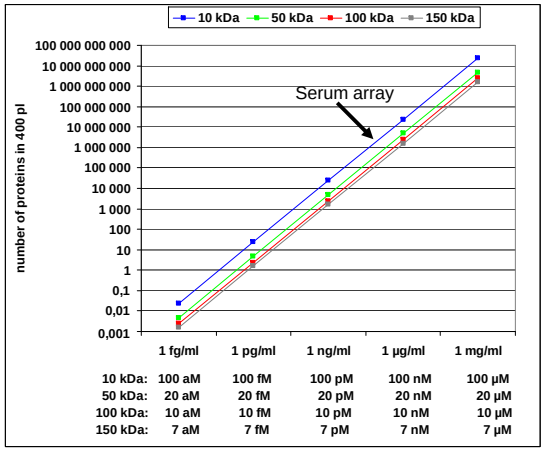
Analysis of IgA levels in 2000 patients

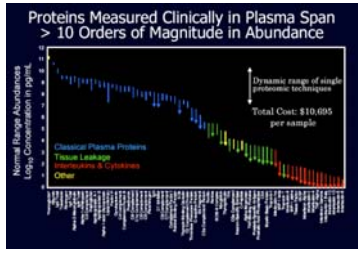


- 2000 serum samples
- Spotted in triplicate
- 2 nl of serum per spot

Janzi et al, 2005
Mol Cell Proteomics

Inherent challenging sensitivity issue for reverse phase arrays



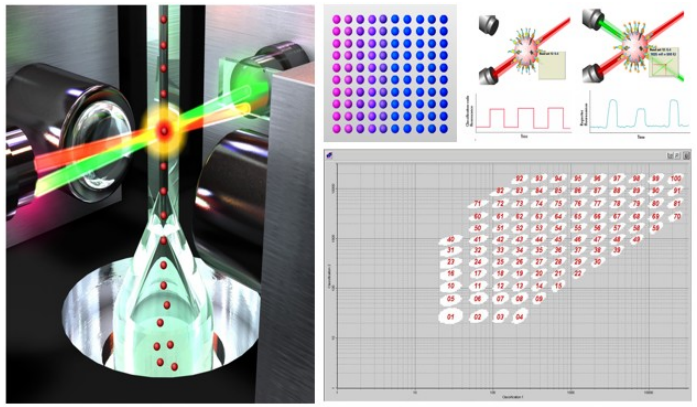


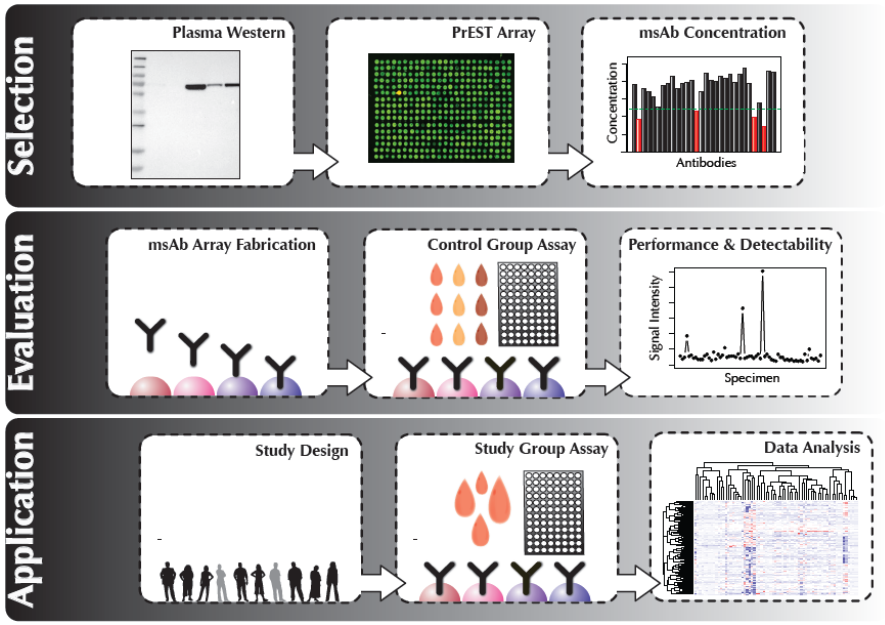
0.4 nl of plasma/serum

50 kDa protein
1 pg/ml (20 fM)

8 protein molecules

Suspension Bead Array Technology

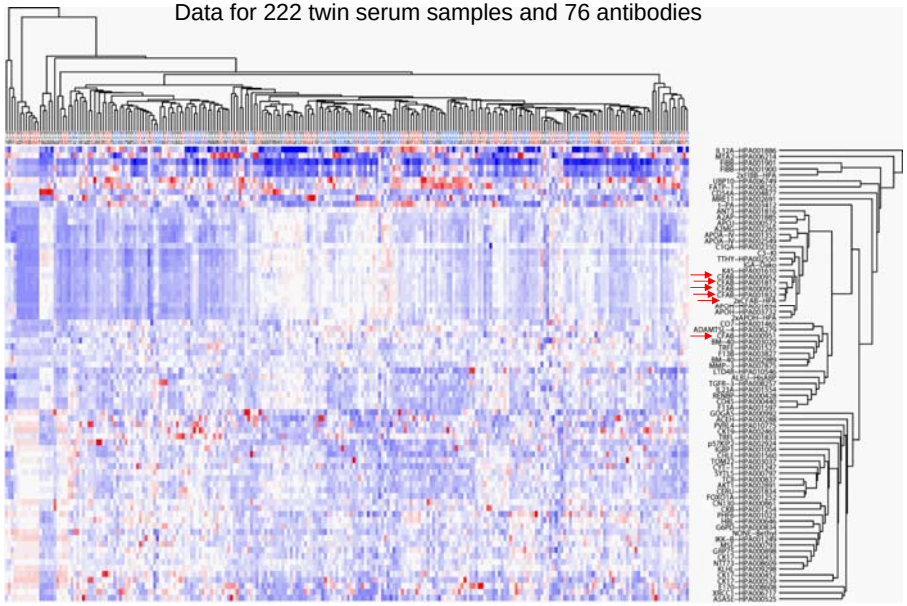




Jochen Schwenk

Bead-based suspension arrays (Luminex) – ratios vs reference pool

Data for 222 twin serum samples and 76 antibodies



Acknowledgements

Fredrik Pontén Caroline Kampf Erik Björling Sophia Hober **Mathias Uhlén**

The entire HPR crew

Stockholm
Henrik Wernerus
Lisa Berglund
Anja Persson
Jenny Ottosson
Peter Nilsson
Cristina Al-Khalili Szigyarto
Emma Lundberg

Uppsala
Kenneth Wester
Anna Asplund