Validation in the eyes of a clinical scientist

perspectives from the Human Protein Atlas

HUPO 2008
Amsterdam

Professor Fredrik Ponten
Department of Genetics & Pathology
Rudbeck Laboratory
Uppsala University Hospital
Uppsala, Sweden

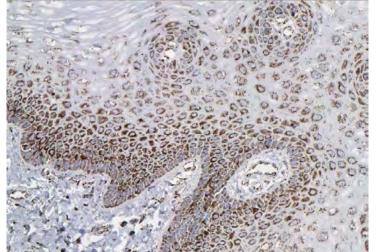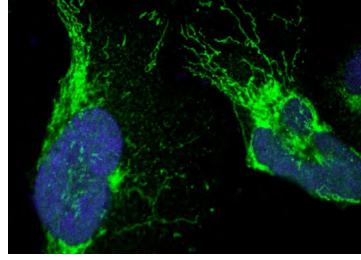# Validation

- **Antibody**

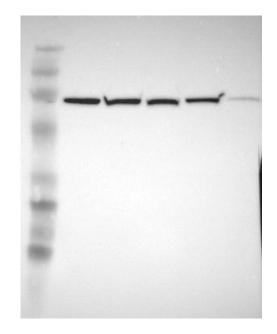 *protein arrays, WB, IF, IHC ....*

- **Expression patterns** *(protein expression vs clinical parameters)*

  <u>add</u> information to existing diagnostic and prognostic factors

  size of cohort (120, 240, >500 ......)

  quality of clinical data

  follow-up time, treatment

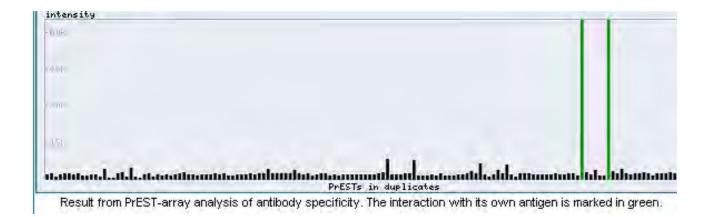 *overall survival, disease free survival, recurrence etc.*

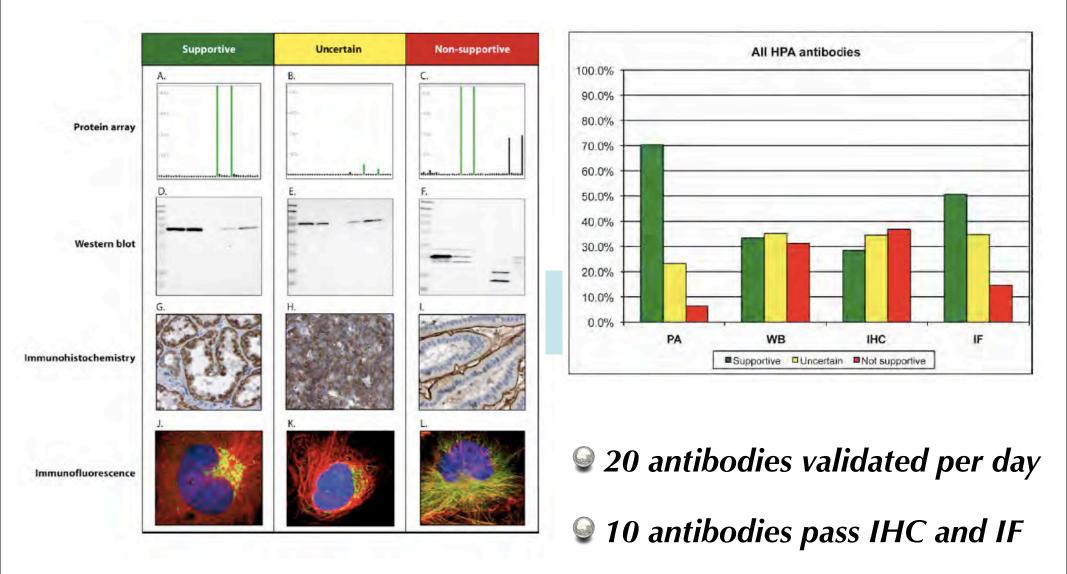# HSPA9 (HPA000898) – mitochondrial, literature supportive

Vagina

Result from PrEST-array analysis of antibody specificity. The interaction with its own antigen is marked in green.

Validation of "in-house" generated antibodies

- 20 antibodies validated per day

- 10 antibodies pass IHC and IF

The Golden Standard

Morphology to determine a firm diagnosis <u>cancer</u>, inflammatory disease, rejection etc.

Protein expression additional level of information of potential clinical relevance

screening for candidates.....

Diagnostic

Prognostic

Treatment predictive

Formalin fixed tissues
Pathology archives
Immunohistochemistry
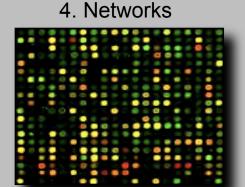
Discovery and validation of cancer biomarkers
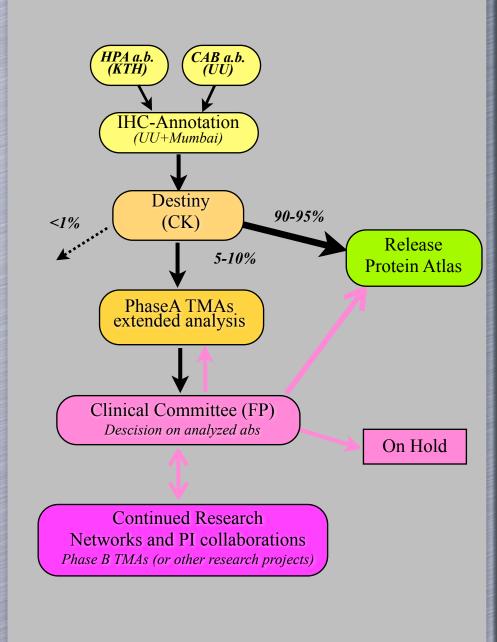
1. Annotation

2. Curation

3. Clinical committee

4. Networks

Extended TMAs - phaseA (HPR)

Extended TMAs - phaseB (external collaborators)

Functional studies, other plattforms

HPA a.b. (KTH) → *CAB a.b. (UU)* →

IHC-Annotation *(UU+Mumbai)* →

Destiny (CK)

<1%

90-95% → Release Protein Atlas

5-10% →

PhaseA TMAs extended analysis →

Clinical Committee (FP) *Descision on analyzed abs*

On Hold

Continued Research Networks and PI collaborations *Phase B TMAs (or other research projects)*

Discovery phase

Discovery phase

In silico biomarker discovery

**Variety of normal cells = negative or weak
AND
Any cancer cells = strong in > 8 tumors**

Malignant Glioma   Malignant Glioma

Liver cancer   HPA003254   Testis cancer

Endometrial cancer   Breast cancer

Ovarian cancer   HPA002110   Pancreatic cancer

Cervical cancer   Ovarian cancer
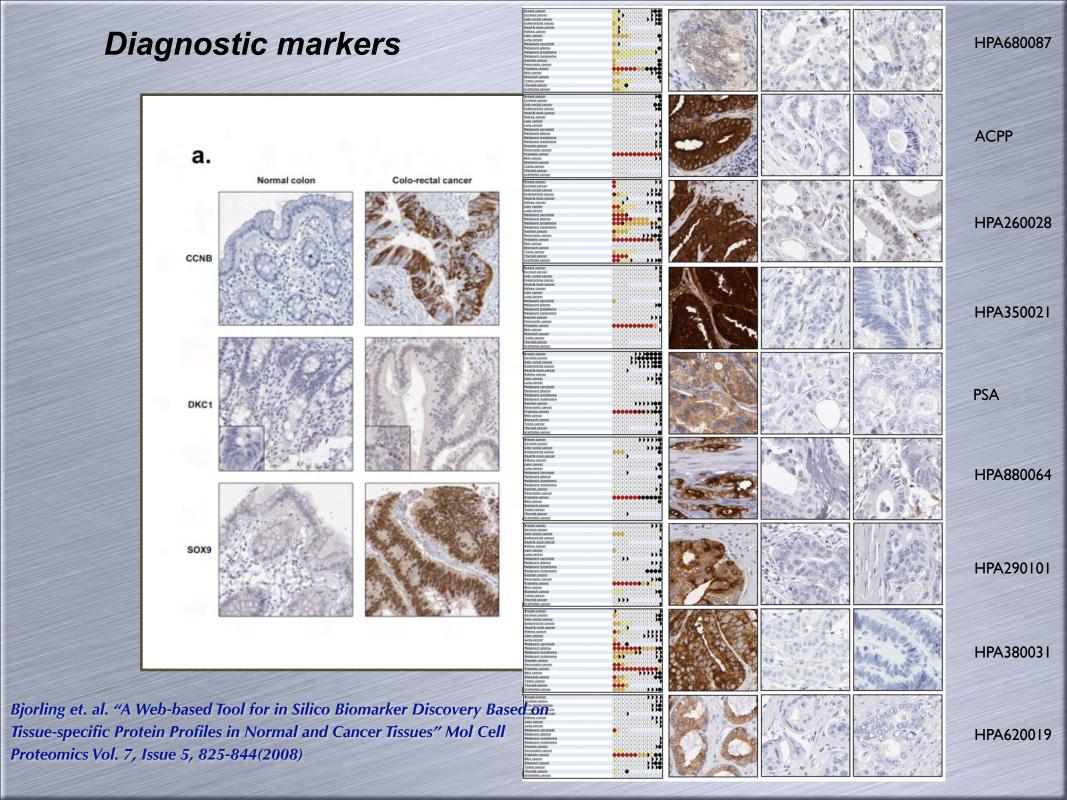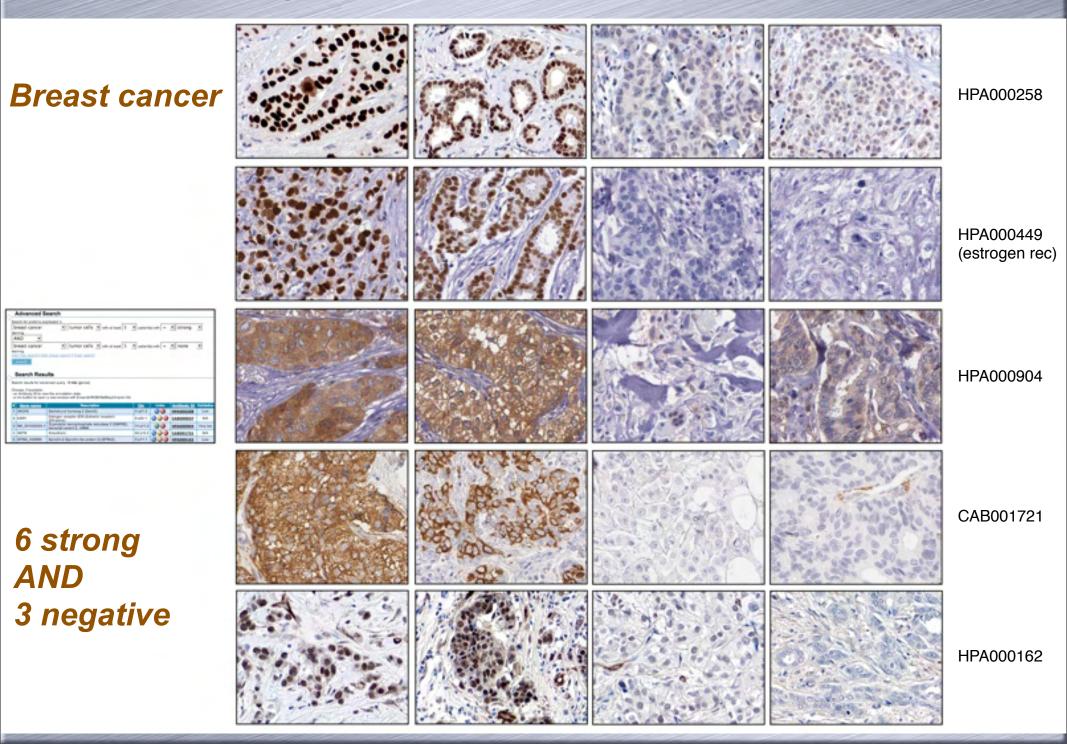
Pancreatic cancer   CAB000113   Urothelial cancer

Diagnostic markers

Bjorling et. al. "A Web-based Tool for in Silico Biomarker Discovery Based on Tissue-specific Protein Profiles in Normal and Cancer Tissues" Mol Cell Proteomics Vol. 7, Issue 5, 825-844(2008)

potential prognostic or treatment predictive markers

Breast cancer

**6 strong
AND
3 negative**

HPA000258

HPA000449
(estrogen rec)

HPA000904

CAB001721

HPA000162

Validation of tissue biomarkers in clinical TMAs

Overall Survival

Stage II
Stage I
Stage III

p= 0.0003

Months from diagnosis

- Colorectal cancer
- Prostate cancer
- Breast cancer
- Lung cancer
- Lymphoma
- Brain tumors
- Urothelial cancer
- Melanoma
- Ovarian cancer
- Renal cancer
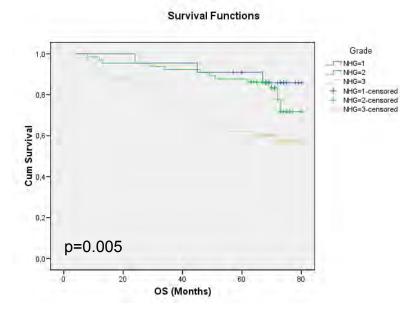
Clinical Validation

- **Collection of tissues and clinical data**

- **Generation of extended TMAs (clinical material)**

- **IHC, Annotation, Data analysis**

- **Published litterature**

- **Test set and independent validation set**

Breast cancer phase A

Histological Subtype
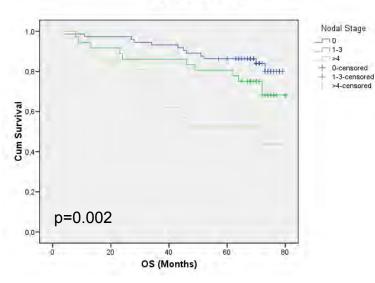
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	mixed	3	2,1	2,1	2,1
	ductal	105	72,4	72,4	74,5
	lobular	27	18,6	18,6	93,1
	tubular	7	4,8	4,8	97,9
	medullary	3	2,1	2,1	100,0
	Total	145	100,0	100,0	

p=0.005

p=0.002

Breast cancer phase A

				kontroll	kontroll				
1	18	19	32	33	45	46	59	60	71
1	18	19	32	33	45	46	59	60	71
3	16	20	31	34	44	47	58	61	70
3	16	20	31	34	44	47	58	61	70
4	15	21	30	36	43	48	57	62	69
4	15	21	30	36	43	48	57	62	69
8	14	22	28	37	42	51	56	63	68
8	14	22	28	37	42	51	56	63	68
9	13	24	27	38	41	52	55	64	67
9	13	24	27	38	41	52	55	64	67
10	11	25	26	39	40	53	54	65	66
10	11	25	26	39	40	53	54	65	66

				kontroll	kontroll				
73	85	86	99	100	116	117	128	130	150
73	85	86	99	100	116	117	128	130	150
74	83	87	98	101	114	118	127	132	149
74	83	87	98	101	114	118	127	132	149
75	82	88	97	102	111	119	126	135	148
75	82	88	97	102	111	119	126	135	148
76	81	89	96	103	110	120	125	136	147
76	81	89	96	103	110	120	125	136	147
77	80	91	95	104	109	121	124	138	146
77	80	91	95	104	109	121	124	138	146
78	79	93	94	105	107	122	123	142	145
78	79	93	94	105	107	122	123	142	145

				kontroll	kontroll				
151	164	165	178	179	MCF-7 72 h hyp	MCF-7 Estro			
151	164	165	178	179	MCF-7 48 h hyp	MCF-7 Estro+tam			
153	163	166	177	CAMA Ctr	MCF-7 24 h hyp	T47-D Ctr			
153	163	166	177	CAMA 24 h hyp	MCF-7 Ctr	T47-D 24 h hyp			
154	162	167	176	CAMA 48 h hyp	MDA-231 Estro+tam	T47-D 48 h hyp			
154	162	167	176	CAMA 72 h hyp	MDA-231 Estro	T47-D 72 h hyp			
156	161	168	175	CAMA Estro	MDA-231 72 h hyp	T47-D Estro			
156	161	168	175	CAMA Estro+tam	MDA-231 48 h hyp	T47-D Estro+tam			
157	160	169	173	MDA-468 Ctr	MDA-231 24 h hyp				
157	160	169	173	MDA-468 24 h hyp	MDA-231 Ctr				
158	159	170	172	MDA-468 48 h hyp	MDA-468 Estro+tam				
158	159	170	172	MDA-468 72 h hyp	MDA-468 Estro				

included cell lines

Some examples................

a

Tissue	Cells	
Adrenal gland	cortical cells	
	medullary cells	
Appendix	glandular cells	
	lymphoid tissue	
Bone marrow	bone marrow poetic cells	
Breast	glandular cells	
Bronchus	surface epithelial cells	
Cerebellum	cells in granular layer	
	cells in molecular layer	
	purkinje cells	
Cerebral cortex	neuronal cells	
	non-neuronal cells	
Cervix, uterine	glandular cells	
	surface epithelial cells (squamous)	
Colon	glandular cells	
Duodenum	glandular cells	
Endometrium 1	cells in endometrial stroma/ECM	
	cells in myometrium/ECM	
	glandular cells	
Endometrium 2	cells in endometrial stroma/ECM	
	cells in myometrium/ECM	
	glandular cells	
Epididymis	glandular cells	
Esophagus	surface epithelial cells	
Fallopian tube	glandular cells	
Gall bladder	glandular cells	
Heart muscle	myocytes	
Hippocampus	neuronal cells	
	non-neuronal cells	
Kidney	cells in glomeruli	
	cells in tubuli	
Lateral ventricle	neuronal cells	
	non-neuronal cells	
Liver	bile duct cells	
	hepatocytes	

Tissue	Cells	
Lung	alveolar cells	
	macrophages	
Lymph node	follicle cells (cortex)	
	non-follicle cells (paracortex)	
Nasopharynx	surface epithelial cells	
Oral mucosa	surface epithelial cells	
Ovary	follicle cells	
	ovarian stromal cells	
Pancreas	exocrine pancreas	
	islet cells	
Parathyroid gland	glandular cells	
Placenta	decidual cells	
	trophoblastic cells	
Prostate	glandular cells	
Rectum	glandular cells	
Salivary gland	glandular cells	
Seminal vesicle	glandular cells	
Skeletal muscle	myocytes	
Skin	adnexal cells	
	epidermal cells	
Small intestine	glandular cells	
Smooth muscle	smooth muscle cells	
Soft tissue 1	mesenchymal cells	
Soft tissue 2	mesenchymal cells	
Spleen	cells in red pulp	
	cells in white pulp	
Stomach 1	glandular cells	
Stomach 2	glandular cells	
Testis	cells in ductus seminiferus	
	leydig cells	
Thyroid gland	glandular cells	
Tonsil	follicle cells (cortex)	
	non-follicle cells (paracortex)	
	surface epithelial cells	
Urinary bladder	surface epithelial cells	
Vagina	surface epithelial cells	
Vulva/anal skin	surface epithelial cells	

b

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

- ● strong positivity
- ○ weak positivity
- ● moderate positivity
- ○ negative
- ● no values attributable to the given specimen

Normal

Cancer

hpr

Colon cancer - phase A TMA
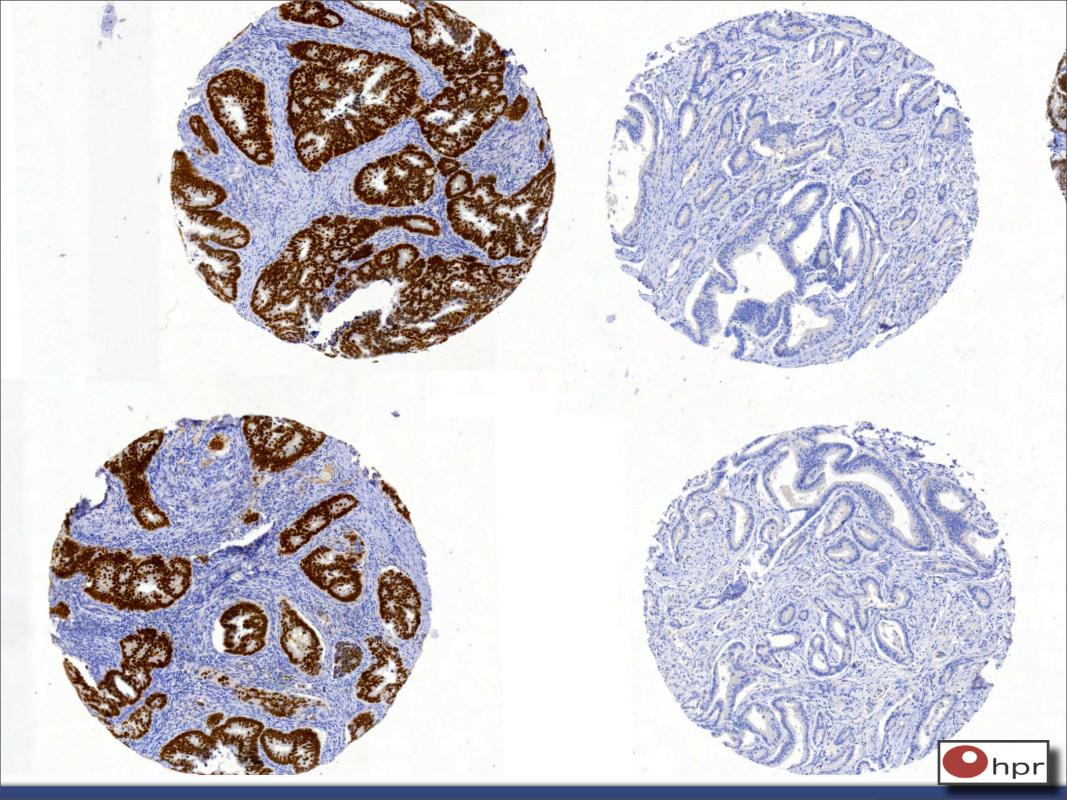
Count		DUKES			
		A	B	C	Total
SEX	female	21	24	18	63
	male	18	18	23	59
Total		39	42	41	122

Median age 75 years (32-88)
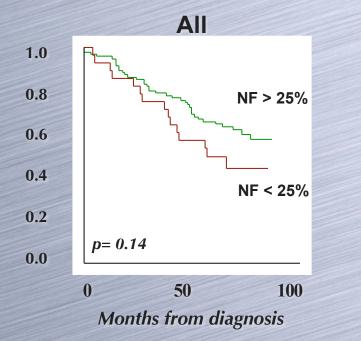
Positive nuclei in colon cancer

Sensitivity approx 80%
Specificity >95%
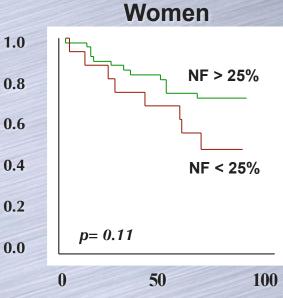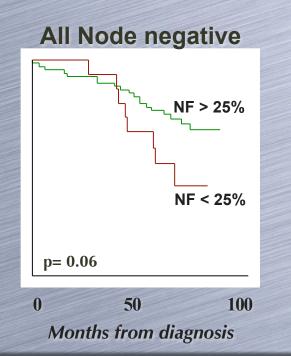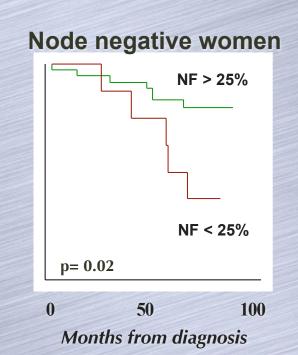
Diagnostic marker together with cytokeratin 20 (95%)
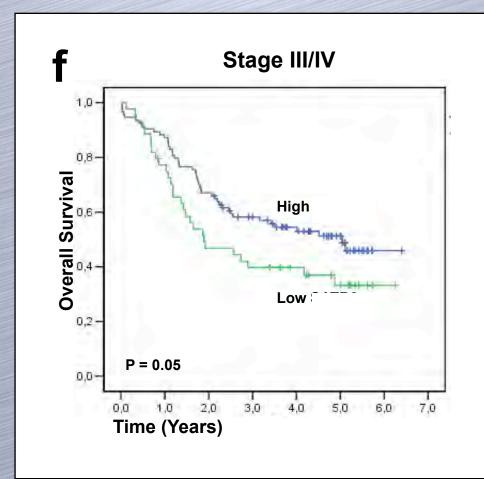

hpr

Overall Survival -phase B (validation cohort)

Overall survival - female patients
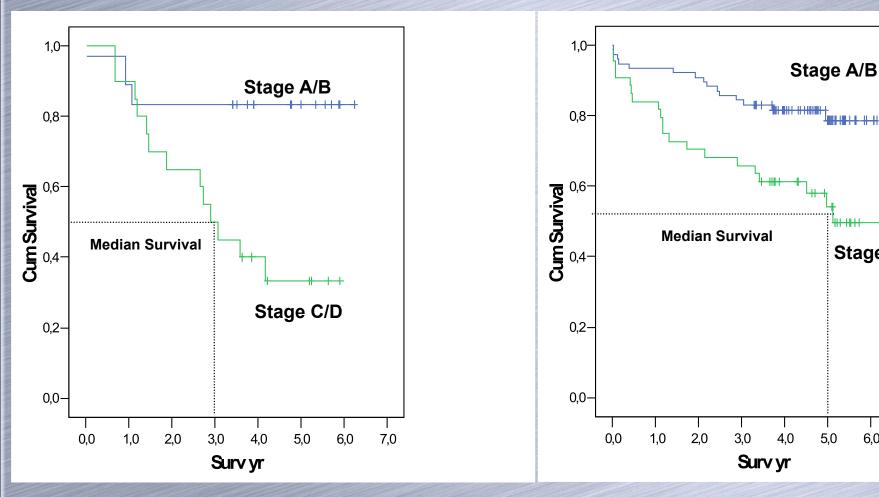
melanoma markers

- Total: 35 antibodies selected
- Immunostained melanoma TMA: 21
- Annotated: 13 (Statistics: 10)

Annotated

1. **Melan-A**
2. **Ki-67**
3. **Disc large homologue-5**
4. **Syntaxin-7**
5. **Arylsulfatase F**
6. **RING finger protein 139**
7. **TRP-1**
8. **Galectin-1**
9. **SOX-10**
10. **S100A4**
11. **MITF**
12. **N-acylglucosamine 2-epimerase**
13. **Melanocyte protein Pmel 17**

In annotation

1. HPA003985
2. HPA004823
3. HPA004838
4. HPA003648
5. HPA004622
6. HPA001369
7. HPA007410
8. HPA010026

In staining

1. HPA007234
2. HPA005740
3. HPA005459
4. HPA007398
5. HPA008232
6. HPA008755
7. HPA008572
8. HPA010639
9. CAB006265
10. HPA009076
11. HPA010655
12. HPA014375
13. CAB012962
14. CAB015178

Melanoma markers - TMAs

Screening: HPA (48 normal, 216 tumors (12 melanoma), 47 cell lines (2 melanoma lines))

Skin TMA: normal skin (location, age, gender), nevi (sub-types), melanoma (in situ, SSM, NM, LMM, ALM)

Melanoma TMA-1: 157 invasive melanomas

Melanoma TMA-2: case-control study (40 cases + 160 controls)

MelanomaTMA-1

157 patients
mean follow-up time:192 months

Table 1. Details of patients and melanomas included in the microarray

Variable	Tissue microarray	Median survival in days	
Number of patients	157	2303	
Gender			
Female	73	2478	
Male	84	2219	
Age			
<65	92	2417	
≥65	65	2303	
T-stage			
T1	64	2607	
T2	34	2393	
T3	31	1906	
T4	17	1481	
M1	5	264	
Unclassified	6	1673	
Tumor subtype			
Superficial spreading	96	2577	
Nodular melanoma	47	1930	
Lentigo maligna melanoma	4	1841	
Acral lentiginous melanoma	6	2401	
Unclassified	4	2509	
Number of patients surviving in cohort	98 (62%)		
Number of patients developing local or regional recurrence	60 (38%)		

Overall survival

T1
T2
T3
T4
T1 vs T2, p=0.56
T1 vs T3, p<0.001
T1 vs T4, p<0.001
T2 vs T3, p=0.004
T2 vs T4, p<0.001
T3 vs T4, p=0.26

Disease free survival

T1
T2
T3
T4
T1 vs T2, p=0.26
T1 vs T3, p<0.001
T1 vs T4, p<0.001
T2 vs T3, p<0.001
T2 vs T4, p<0.001
T3 vs T4, p=0.02
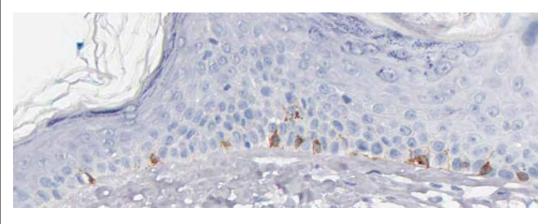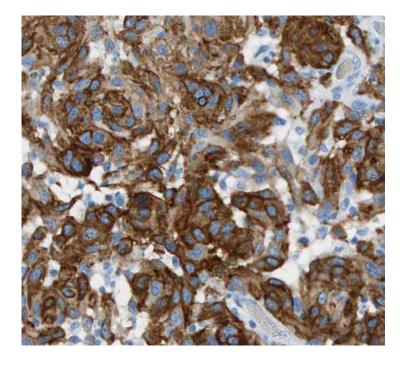
Melan-A (CAB000057)

- Melanoma antigen recognized by T-cells 1 (MART-1) (Melan-A protein) (Antigen SK29-AA) (Antigen LB39-AA).

alph. sort order	Cancer Tissues - IHC
Breast cancer	● ● ● ●
Cervical cancer	●
Colo-rectal cancer	●
Endometrial cancer	● ● ● ● ●
Head & neck cancer	● ●
Liver cancer	●
Lung cancer	●
Malignant carcinoid	
Malignant glioma	
Malignant lymphoma	
Malignant melanoma	● ● ● ● ● ● ● ● ● ● ●
Ovarian cancer	●
Pancreatic cancer	●
Prostate cancer	●
Renal cancer	●
Skin cancer	● ● ● ● ●
Stomach cancer	●
Testis cancer	●
Thyroid cancer	
Urothelial cancer	

Ki-67 *(CAB000058)*

alph. sort order	Cancer Tissues - IHC
Breast cancer	
Cervical cancer	
Colo-rectal cancer	
Endometrial cancer	
Head & neck cancer	
Liver cancer	
Lung cancer	
Malignant carcinoid	
Malignant glioma	
Malignant lymphoma	
Malignant melanoma	
Ovarian cancer	
Pancreatic cancer	
Prostate cancer	
Renal cancer	
Skin cancer	
Stomach cancer	
Testis cancer	
Thyroid cancer	
Urothelial cancer	

Ki-67

- Ki 67 and **T-stage** (R=0.29, p<0.001)
- The expression of Ki-67 was also found associated with **histology** (p<0.0001), where a larger proportion of NMM tumors showed a high Ki-67 expression in comparison with SSM tumors.
- High levels of Ki-67 expression were significantly associated with decreased **overall survival** rate (p<0.0001). In addition, a significant correlation was found between the expression of Ki-67 and disease-free survival (p<0.0002).

overall survival   **disease-free survival**

HPA001467 (Syntaxin 7)

4
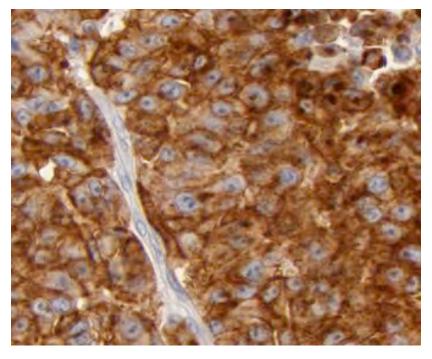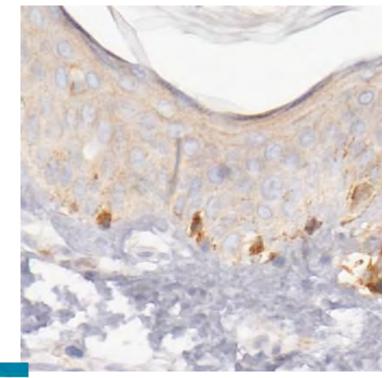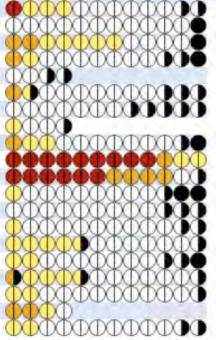
Cancer Tissues

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

IH | IF | PA | WB
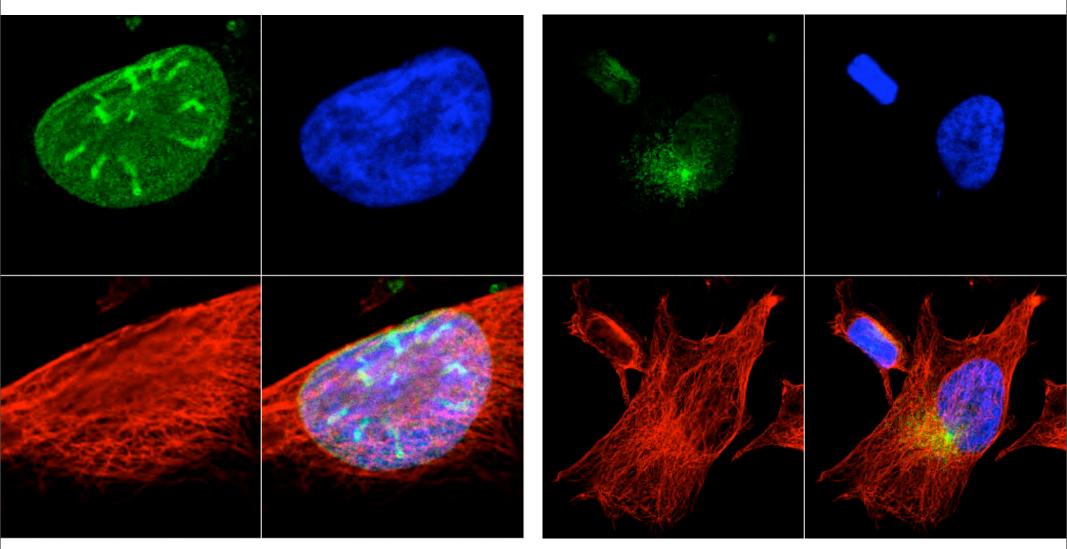
4

WM115: PFA/Triton

Ex 1

Ex 2

syntaxin-7 expression varies in relation to cell cycle phase, translocation/transport between nucleus and cytoplasm (small vesicles)

HPA001467 (Syntaxin 7)
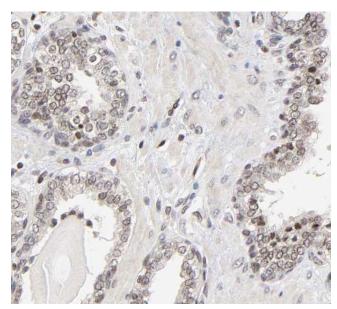
- **negative correlation** between Syntaxin-7 and **T-stage** (R=-0.16, p=0.07), with loss of high-grade expression of Syntaxin-7 in tumors with higher T-stage

- **negative correlation** between Syntaxin-7 and **Clark level**, however this was not significant (R=-0.17, p=0.05)

- trend towards an association between Syntaxin-7 and **histology** (p=0.10), with a larger proportion of SSM tumors having a high-grade expression of Syntaxin-7 compared to NMM tumors

high sensitivity
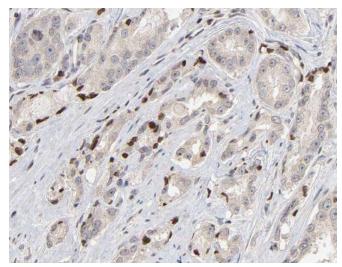
specificity - melanoma + lymphoma

melanoma 1: inverse correlation to tumor grade (p=0.07)

melanoma 2: inverse correlation to tumor grade (p=0.001)

Gene data

Normal Tissues - IHC

alph. sort order

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

Annotation Summary - IHC
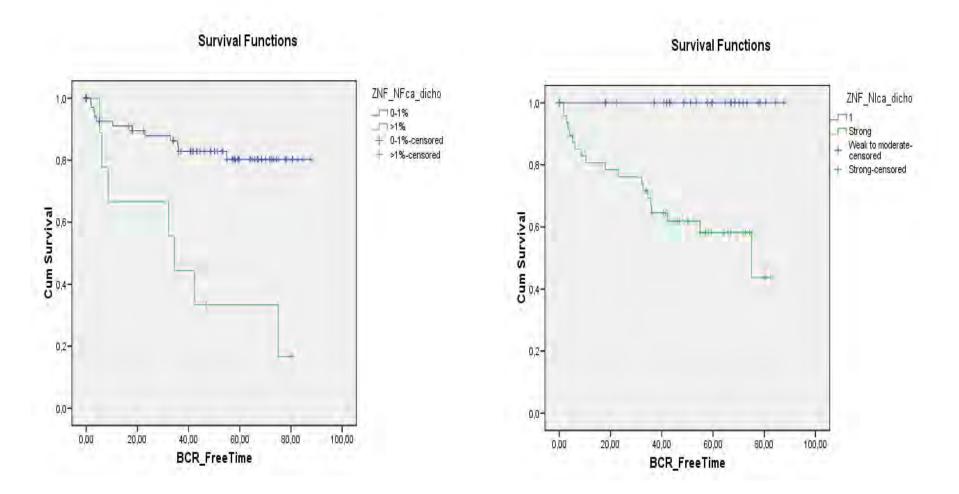
This antibody showed sometimes moderate to strong nuclear and sometimes weak to moderate cytoplasmic staining. Nuclear expression was seen in lymphoid cells, squamous epithelium, gall bladder mucosa, urothelium, trophoblast cells, ovarium, fallopian tube, breast and hematopoietic cells in bone marrow. Cytoplasmic expression was found in glandular cells of gastrointestinal tract, adrenals, lung macrophages, testis, respiratory epithelium, some neuronal cells and muscle cells. Most tumors showed negative or weak, usually cytoplasmic expression seen in some breast, testis, and colon cancers. Nuclear staining was seen in some lymphomas and prostate carcinomas.

Cancer Tissues - IHC

alph. sort order

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

A zinc finger protein (HPA 4191)

Normal prostate

Prostate cancer

A zinc finger protein
(HPA 4191)

BCR free survival

Prostate cancer

A zinc finger protein
(HPA 4191)

Prostate cancer phase A

- A low fraction of strongly stained nuclei.

- Fraction of positive cells correlated to Gleason grades and seminal vesicle invasion.

- Fraction of positive cells and intensity of staining weakly correlated to WHO grade.

- Better outcome in BCR free survival for patients with 0-1% positive nuclei, as well as negative-moderate nuclear intensity.

HPA001355 – an enzyme

- Glioma marker?!

HPA001355

HPA 0506 – unknown protein

Unknown protein HPA 0506

Glioma

Normal brain

Spleen

Kidney

Other cancer types

CD 31

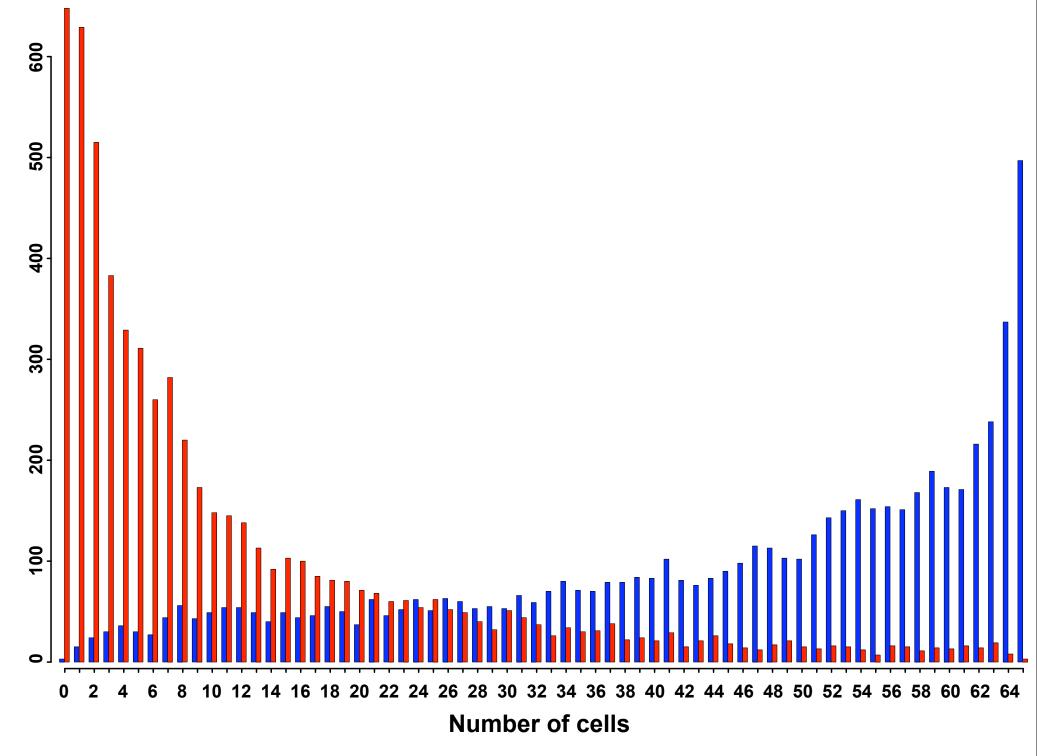# Status Clinical Validation

- Ongoing TMA production

- 12-15 new candidate markers per month

- 150 antibodies pending clinical validation

- 80 antibodies analyzed (first run)

- 25% further validation, research

- 21 proteins on "hot list

Conclusions

- **Immunohistochemistry and TMAs is a powerful tool for identification and validation of clinical biomarkers**
- **Well-defined patient cohorts are essential to draw conclusions on clinical relevance**
- **Pathology assessment for evaluation of expression patterns**
- **Functional studies needed for clinical acceptance**

Mathias Uhlen   Sophia Hober   Anna Asplund
    Erik Björling   Caroline Kampf   Kenneth Wester

Karin Jirström   Lisa Berglund   Henrik Wernerus
                      Anja Persson   Peter Nilsson
Sanjay Navani   Jenny Ottosson   Emma Lundberg
                                            Cristina Al-Khalil

Knut och Alice
Wallenbergs
Stiftelse

hpr