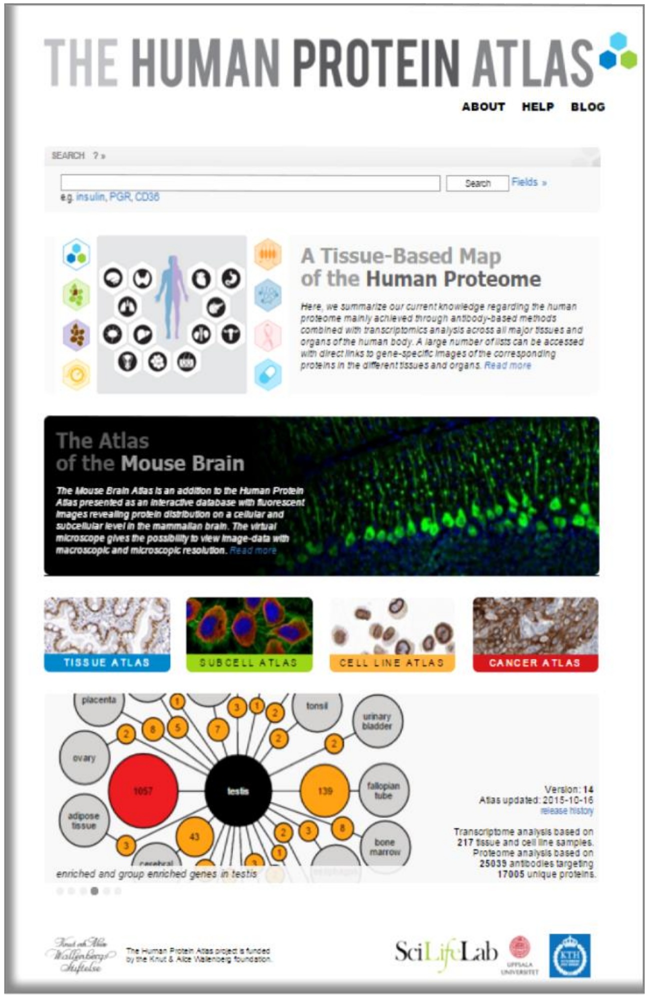
THE HUMAN PROTEIN ATLAS

Per-Henrik Edqvist, Ph.D., Associate professor

www.proteinatlas.org



A Tissue-Based Map of the Human Proteome

Here, we summarize our current knowledge regarding the human proteome mainly achieved through antibody-based methods combined with transcriptomics analysis across all major tissues and organs of the human body. A large number of lists can be accessed with direct links to gene-specific images of the corresponding proteins in the different tissues and organs. Read more

The Atlas of the Mouse Brain

The Mouse Brain Atlas is an addition to the Human Protein Atlas presented as an interactive database with fluorescent images revealing protein distribution on a cellular and subcellular level in the mammalian brain. The virtual microscope gives the possibility to view image-data with macroscopic and microscopic resolution. Read more

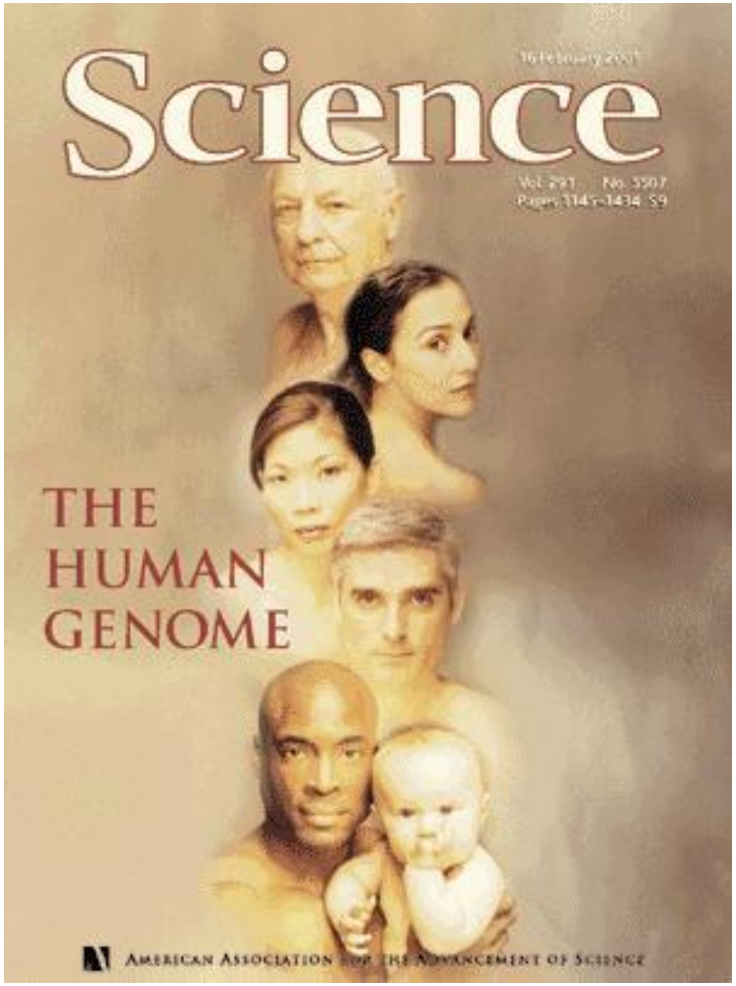
enriched and group enriched genes in testis

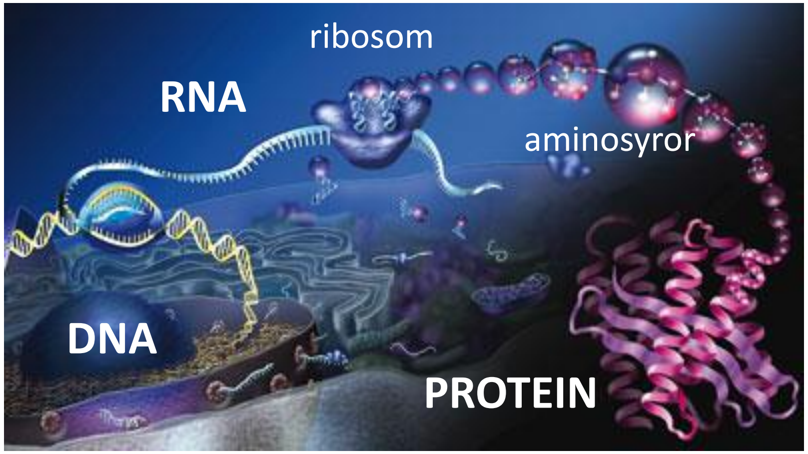
Version: **14**
Atlas updated: 2015-10-16
release history

Transcriptome analysis based on **217** tissue and cell line samples.
Proteome analysis based on **25039** antibodies targeting **17005** unique proteins.

The Human Protein Atlas project is funded by the Knut & Alice Wallenberg foundation.

26 juni 2000





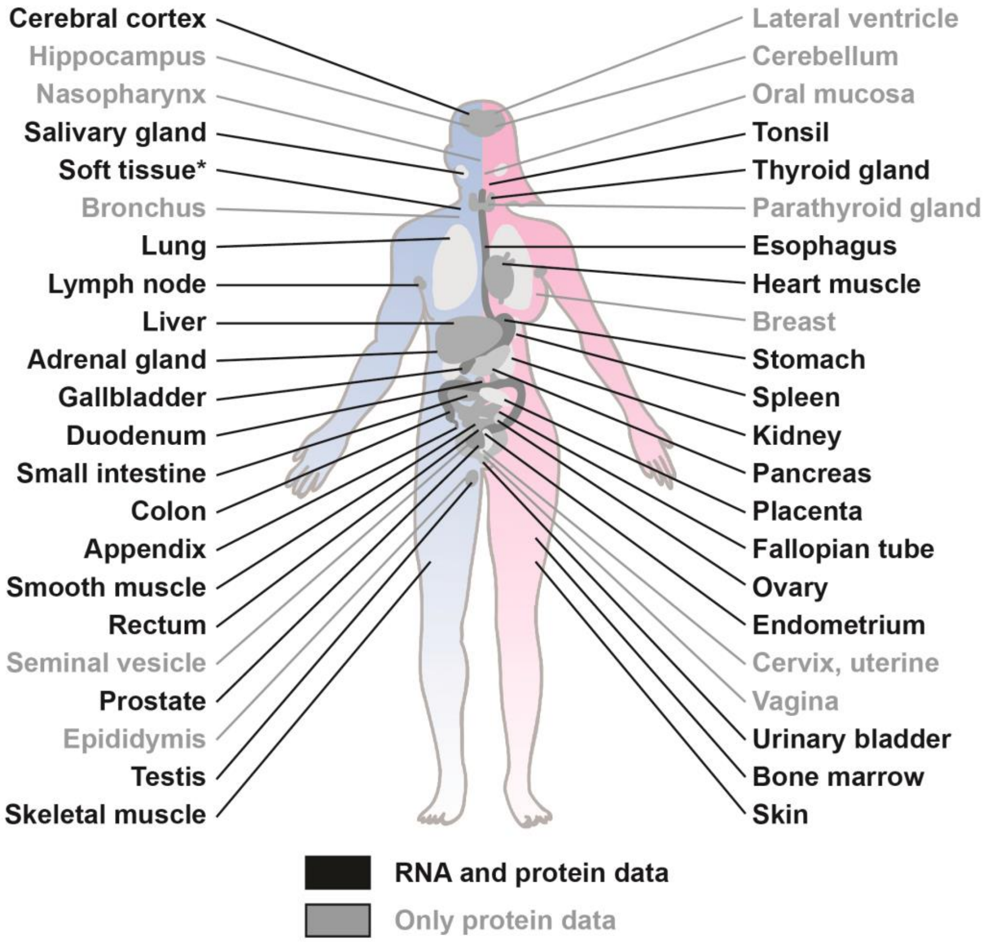
”The central dogma”

23 januari 2015

Science

AAAS

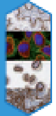
RNA and protein expression data



THE HUMAN PROTEOME

The Power of Proteins. *If DNA can be equated with the blueprint for a home, then proteins provide the bricks and mortar, plumbing, paint—essentially everything that makes up the house. The human genome consists of approximately 20,000 protein-coding genes. This poster summarizes the multiple on-going antibody- and transcriptomics-based proteome projects and where in the human body this research is focused. For more detailed information, view: www.proteinatlas.org*

THE HUMAN PROTEIN ATLAS



The Human Protein Atlas contains a wealth of expression and localization data on the majority of protein-coding genes. It is subdivided into four parts: normal tissue, cancer, subcellular, and cell lines. Transcriptomics data provides gene expression information across the tissues and organs, while antibody-based protein profiles show cell-level localization for the corresponding protein. The latest version of the Human Protein Atlas (13.0) contains RNA data for 91% and protein data for 85% of the translated human genome and includes more than 14 million images with primary data from immunohistochemical and immunofluorescent studies.

THE TISSUE-SPECIFIC PROTEOME

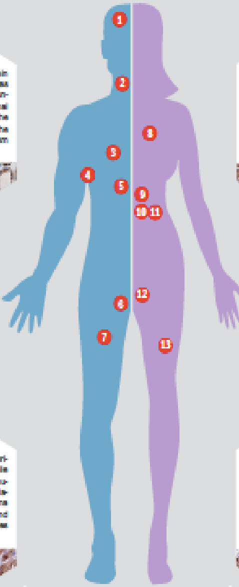
The expression of all human protein-encoding genes has been measured in samples representing all major tissues and organs in the human body. Approximately one-third showed some level of elevated expression in at least one of the analyzed tissues, but few showed strict tissue-specific expression. Functional analysis has shown that the role of proteins with a tissue-elevated expression correlates with the tissue/organ function. For example, the liver produces large numbers of secreted proteins, while the kidney expresses many membrane-bound transport proteins, and the brain harbors a preponderance of neurological proteins.

THE HOUSE-KEEPING PROTEOME

Transcriptomics analysis suggests that some 9,000 genes code for so-called house-keeping proteins that are needed in all cells to maintain the normal cellular structure and basic functions for life. These include ribosomal proteins involved in protein synthesis, enzymes essential for cell metabolism and gene expression, and mitochondrial proteins needed for energy generation, as well as structural proteins responsible for building and maintaining the physical integrity of the cell.

THE REGULATORY PROTEOME

All processes within a living cell are highly regulated, including cell proliferation, differentiation, and death. Regulatory mechanisms include the control of gene expression as well as posttranslational modifications that can regulate protein activity, stability, localization, or degradation. Transcription factors, of which 1,500 human proteins have been identified, are an especially important class of regulatory proteins as they function as the on/off switch for gene expression.



1 THE BRAIN

The brain is a highly complex and energy-intensive organ that receives, processes, and executes the coordinated higher functions of perception, motion, and cognition. Neural proteins show specific patterns of expression in different cells as well as in subcellular structures such as axons, dendrites, synapses, and the extracellular matrix.

2 THE ENDOCRINE ORGANS

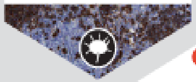
The adrenal and thyroid glands are the main endocrine organs. The adrenal gland secretes steroid hormones (impacting metabolic function and electrolyte balance) from the adrenal cortex and catecholamines are released by the adrenal medulla (in response to stress). The thyroid gland regulates basal cell metabolism and blood calcium levels.

3 THE LUNG

The lung is a delicate, spongy tissue primarily responsible for respiration, the gaseous exchange of O_2 and CO_2 between air and blood that occurs at the ~300 million alveoli. Pneumocytes, bronchial epithelium, and endothelial cells facilitate O_2-CO_2 exchange, while alveolar macrophages protect the lung against infection from inhaled microbes.

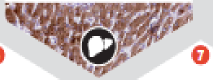
4 THE HEMATOPOIETIC TISSUES

Bone marrow, located in the cavities of large bones, contains stem cells that divide and differentiate into mature erythrocytes and leukocytes. Immune cell activation, proliferation, and maturation occurs in secondary lymphoid organs including the spleen, lymph nodes, and lymphoid tissue located in various organs such as the appendix.



5 THE LIVER

The liver—the largest internal organ—comprises parenchymal cells (hepatocytes and bile ducts cells) and non-parenchymal cells (sinusoidal endothelial, Kupffer, and hepatic stellate cells). Liver-specific proteins include plasma and bile proteins, detoxification proteins, and proteins associated with metabolic processes and glycogen storage.



6 THE MALE TISSUES

The testes and prostate gland represent the male reproductive tissues. The testes produce sperm and hormones important for the development of male sex characteristics. The prostate secretes semen components that support sperm viability. Many of the testes-specific proteins are involved in the reproductive process and spermatogenesis.

7 THE ADIPOSE TISSUE

Adipose, or fat, tissue is highly specialized to store energy, secrete hormones, and protect the body. It surrounds important organs (visceral fat) and also forms an insulating layer beneath the skin (subcutaneous fat). Proteins involved in lipid metabolic processes, secretion, and transport are found to be elevated in adipose tissue.

8 THE CARDIAC AND SKELETAL MUSCLE

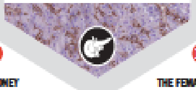
Cardiac muscle, comprised of specialized cardiomyocytes, functions to pump blood to all extremities of the body and maintain optimal blood pressure, whereas skeletal muscle, consisting of striated muscle cells fused into long muscle fibers and attached to the skeleton, provides physical stability and allows for voluntary movement.

9 THE GASTROINTESTINAL TISSUES

The gastrointestinal tract (GIT)—the esophagus, stomach, small and large intestines, and rectum—absorbs nutrients and water, protects against pathogens, and maintains the balance of beneficial microorganisms. GIT proteins are predominantly involved in nutrient breakdown, transport, and metabolism; host protection; and tissue morphology maintenance.

10 THE PANCREAS

The pancreas carries out both exocrine and endocrine functions. Glandular cells in the exocrine compartment secrete digestive enzymes into the gastrointestinal tract, while the islets of Langerhans execute the endocrine function, secreting insulin and other hormones. Unsurprisingly, many pancreatic mRNAs encode specialized secreted proteins.



11 THE KIDNEY

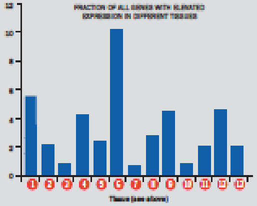
The kidney maintains body homeostasis by regulating blood composition and eliminating organic compounds, including drugs. Different cell types are organized into sub-anatomical structures with distinct functions, showing elevated levels of essential proteins—for example, proteins required for blood filtration are elevated in the glomerulus.

12 THE FEMALE TISSUES

The female reproductive tissues include the ovary, which produces female reproductive hormones and germ cells; the endometrium, site of fertilized embryo implantation; and the placenta. The endometrial mucosa undergoes hormone-driven regeneration during the menstrual cycle, providing the necessary environment for implantation of a developing embryo.

13 THE SKIN

The skin (epidermis, dermis, and subcutaneous layer) is both a protective barrier and sensory organ. The epidermis (mostly keratinocytes) forms a protective barrier against physical, chemical, and biological insults. Most protein functions are related to squamous cell differentiation and cornification, skin pigmentation, and hair development.



THE SECRETOME AND MEMBRANE PROTEOME

Both secreted and membrane-bound proteins play crucial roles in many physiological and pathological processes. Important secreted proteins include cytokines, coagulation factors, and growth factors, among others, while membrane proteins include ion channels or molecular transporters, enzymes, receptors and anchors for other proteins. Approximately 3,000 human genes are predicted to encode secreted proteins, with another 5,500 encoding membrane-bound proteins.

THE ISOFORM PROTEOME

The existence of a variety of protein isoforms in each cell endows the structural space of the human proteome with breadth and complexity. Isoforms are produced through posttranslational modifications, proteolytic cleavage, or somatic recombination. Variations in the amino acid sequence also result from local genetic variations in protein-coding regions. Additionally, a large proportion of protein-coding genes have splice variants that yield protein products of different sizes. The almost limitless variety of posttranslational modifications combine to create thousands of additional variants, contributing to a rich and diverse proteome.

THE CANCER PROTEOME

Over 500 genes have been implicated in malignant transformation. Normal expression of these genes is essential for orderly growth, survival, and function. However, overexpression, loss of expression, or expression of a defective protein can contribute to dysfunction and tumor growth. Dysregulated expression results from large structural rearrangements, chromosomal duplication, specific gene amplifications, or silencing of transcription through mutations or epigenetic mechanisms. Furthermore, point mutations or small insertions or deletions can lead to loss or gain of function in the affected protein.

THE DRUGGABLE PROTEOME

Most pharmaceutical drugs act by targeting proteins and modulating their activity. Target proteins belong to four main families: enzymes, transporters, ion channels, and receptors. The U.S. Food and Drug Administration has approved drugs targeting approximately 600 human proteins, with most acting on signal transduction proteins that convert extracellular signals into intracellular responses. Antibody-based drugs usually cannot penetrate the plasma membrane and therefore target cell surface proteins such as receptors, while small molecule drugs are able to act on both intracellular and extracellular targets.

THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

Science
AAAS

Produced by the Science/AAAS Custom Publishing Office

To explore the human proteome in more depth, visit:
www.proteinatlas.org

Online version of this poster:
posters.sciencemag.org/humanproteome

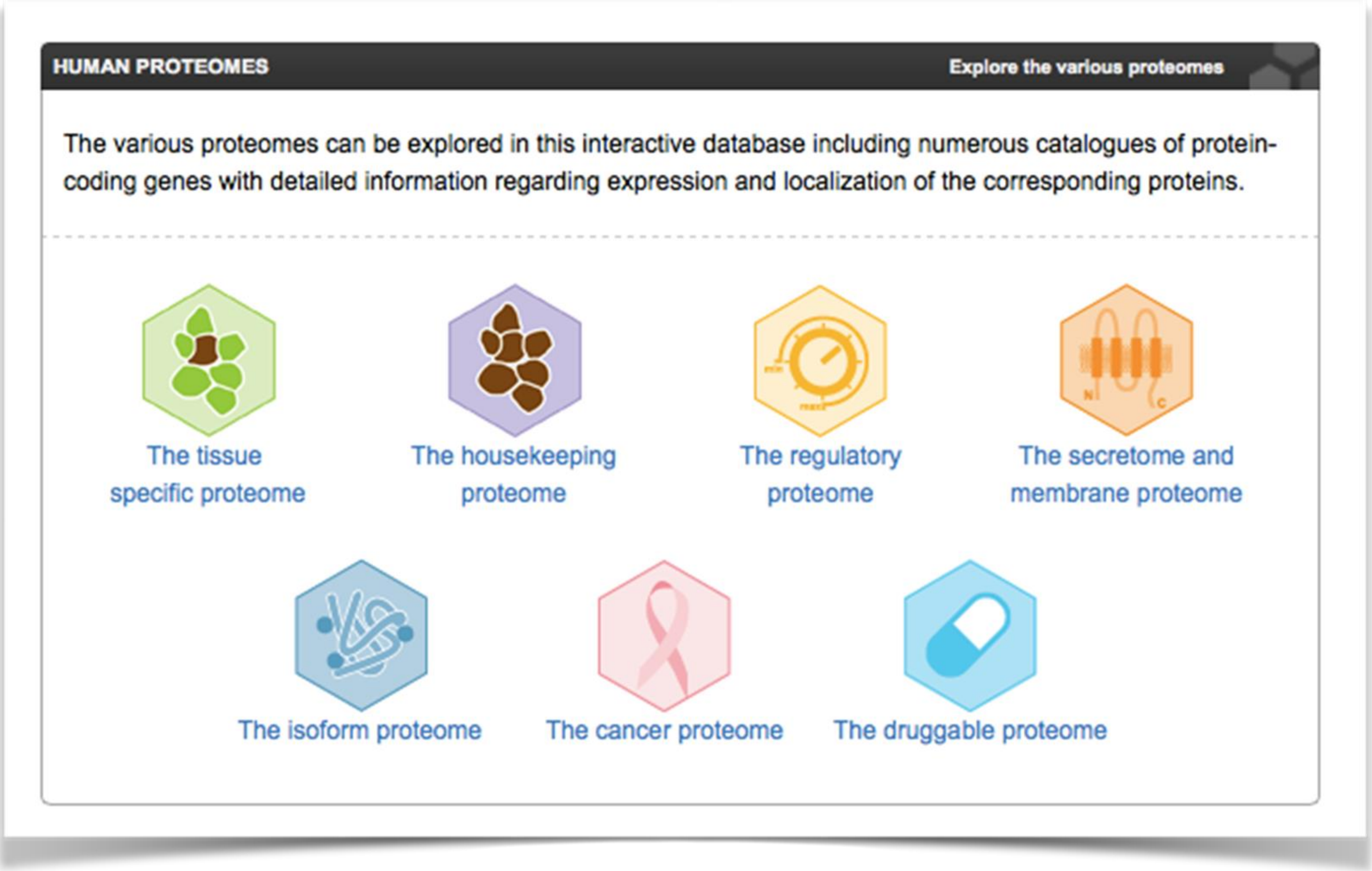
THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

KTH VETENSKAP OCH KONST

UPPSALA UNIVERSITET

Human proteome or Human proteomes?



HUMAN PROTEOMES

Explore the various proteomes

The various proteomes can be explored in this interactive database including numerous catalogues of protein-coding genes with detailed information regarding expression and localization of the corresponding proteins.

The Human Protein Atlas project

History

Started in 2003 as a collaboration between KTH and Uppsala University, and made possible by a generous grant from the Knut and Alice Wallenberg foundation.

Vision

Generate validated protein probes to all the non-redundant proteins encoded by the human genome and use these to functionally explore the human proteome.

Goals

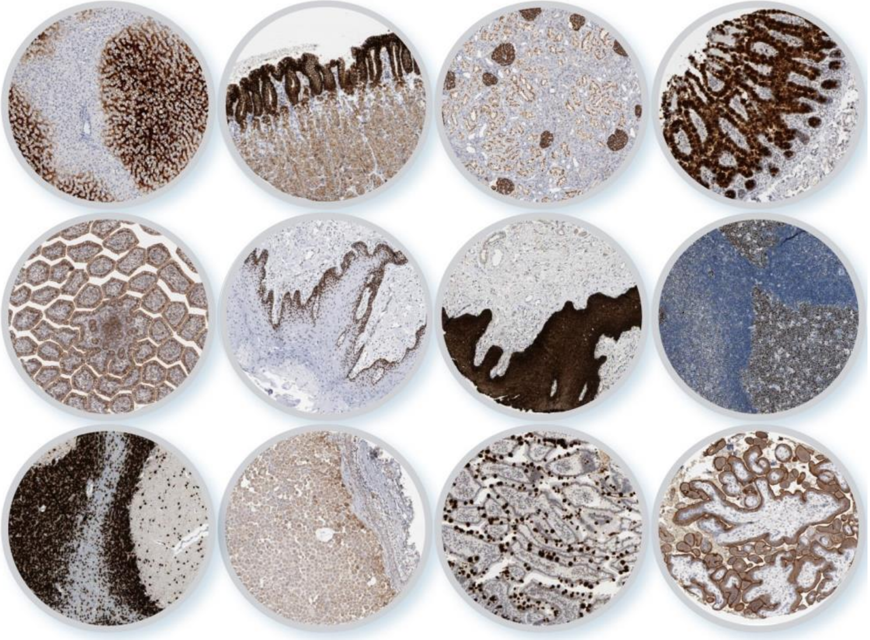
Generate and validate antibodies towards all human proteins (a major isoform for each protein encoding gene).

Create a map of protein expression patterns in cells and in human tissues.

Utilize reagents and protein expression data to identify and explore biomarkers of clinical relevance.

Output at peak production

8-10 new human proteins/day
10,000 images/day
500 Gigabyte of data/day

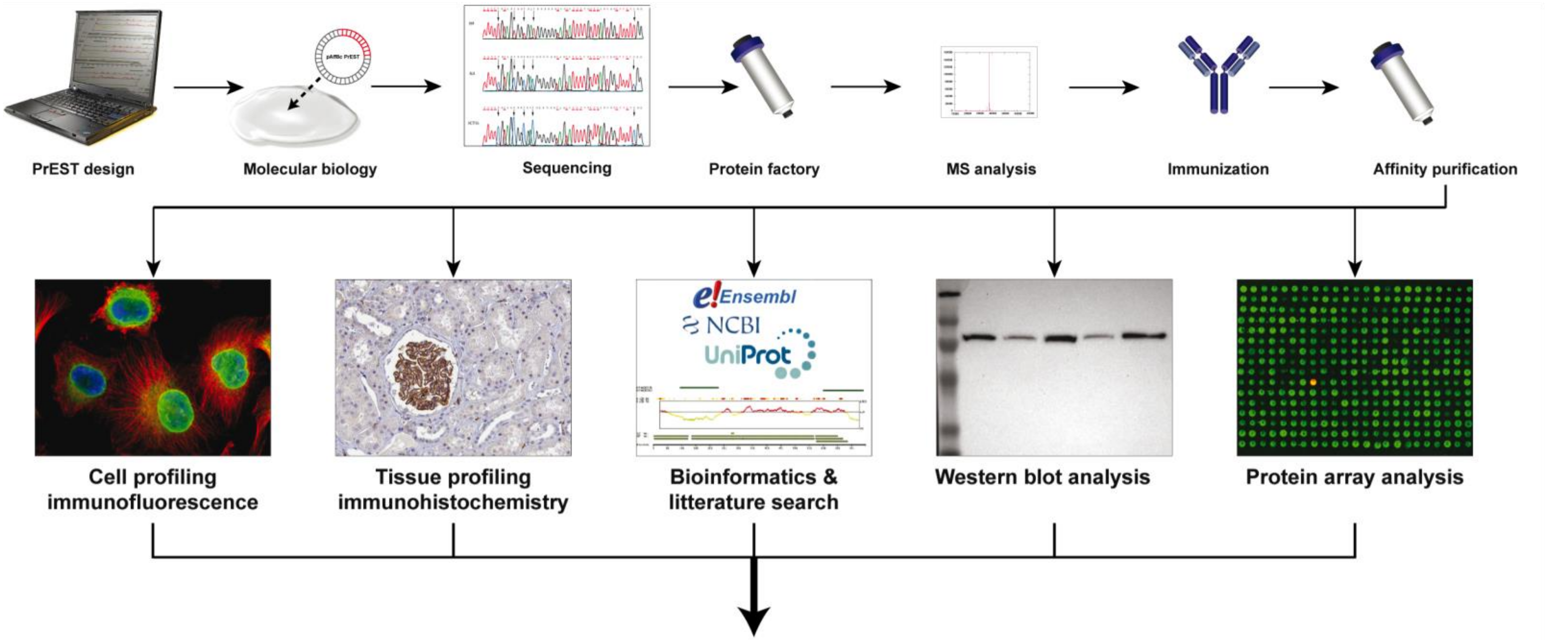


High throughput technologies

The Human Protein Atlas would not be possible without:

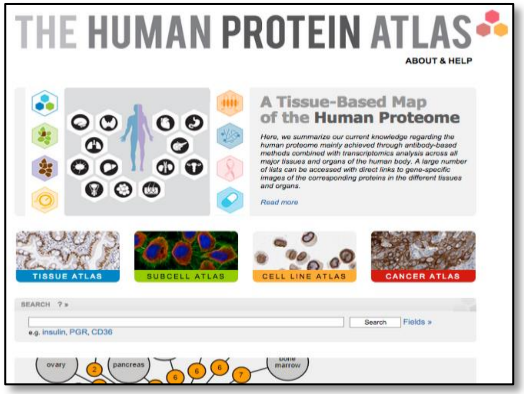
- Industry-scale antibody production and validation
- Tissue microarray technology
- Laboratory information management system (LIMS)

The Workflow within HPA



PrEST design → Molecular biology → Sequencing → Protein factory → MS analysis → Immunization → Affinity purification

- Cell profiling immunofluorescence
- Tissue profiling immunohistochemistry
- Bioinformatics & litterature search
- Western blot analysis
- Protein array analysis



THE HUMAN PROTEIN ATLAS

ABOUT & HELP

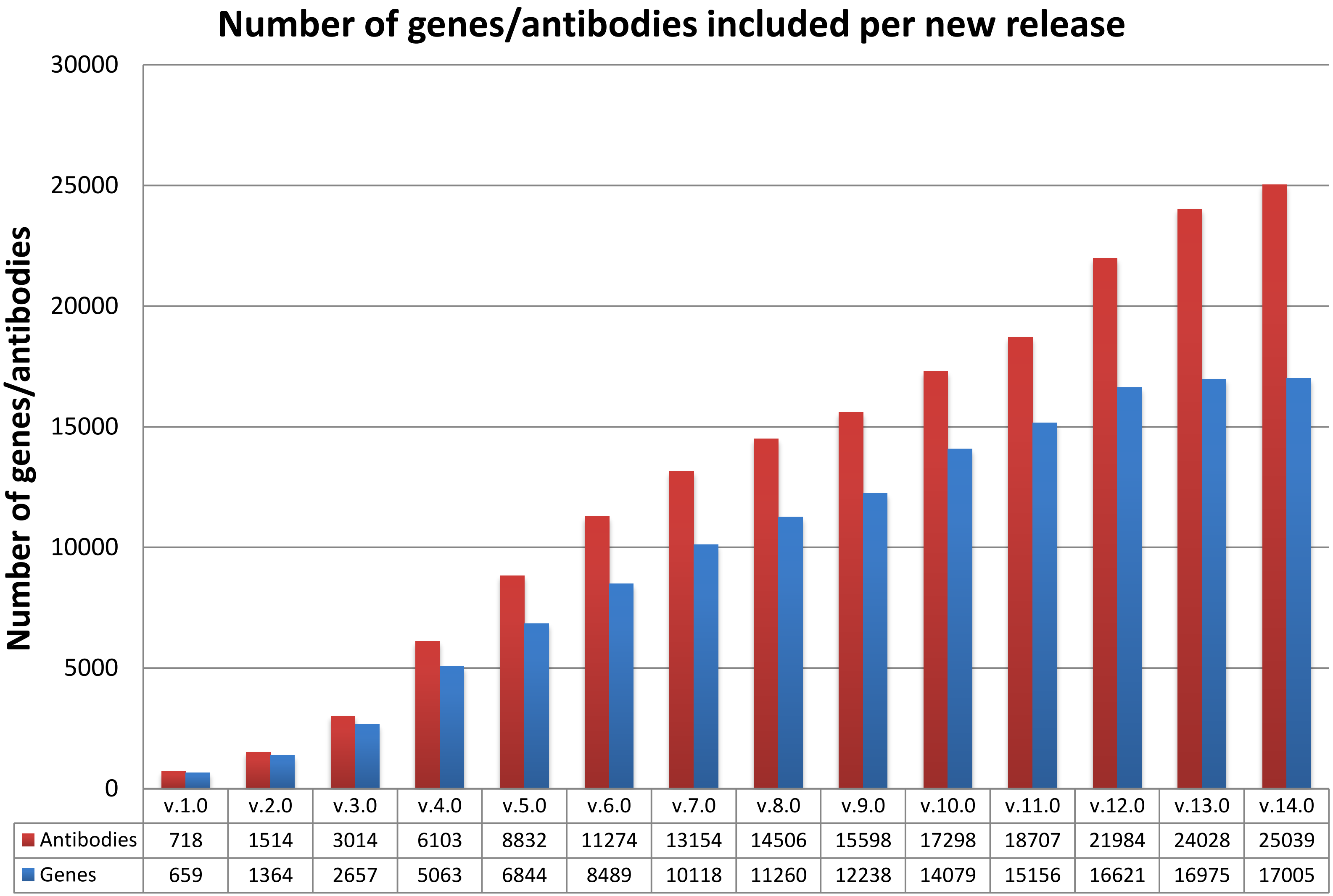
A Tissue-Based Map of the Human Proteome

Here, we summarize our current knowledge regarding the human proteome mainly achieved through antibody-based methods combined with transcriptomics analysis across all major tissues and organs of the human body. A large number of lists can be accessed with direct links to gene-specific images of the corresponding proteins in the different tissues and organs.

Read more

Human proteome (20356 protein coding genes)	
PrEST Design	98%
Cloning	97%
Antigens	96%
Antibodies	95%
Antibody-based profiling	92%

Release history: number of genes/antibodies



Number of genes/antibodies included per new release

	v.1.0	v.2.0	v.3.0	v.4.0	v.5.0	v.6.0	v.7.0	v.8.0	v.9.0	v.10.0	v.11.0	v.12.0	v.13.0	v.14.0
🟥 Antibodies	718	1514	3014	6103	8832	11274	13154	14506	15598	17298	18707	21984	24028	25039
🟦 Genes	659	1364	2657	5063	6844	8489	10118	11260	12238	14079	15156	16621	16975	17005

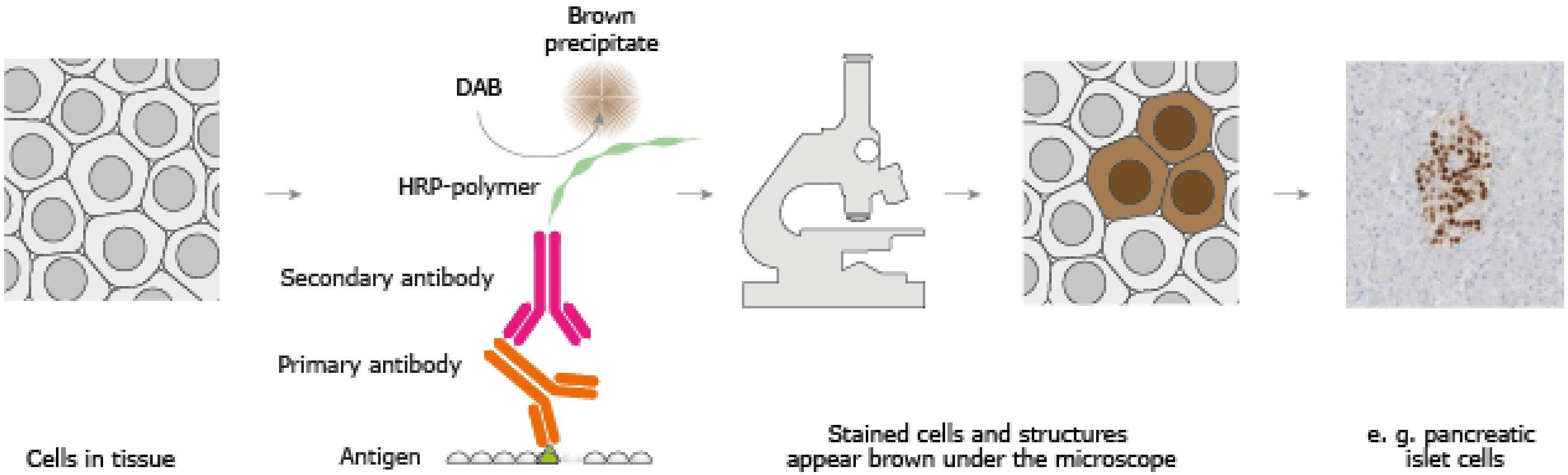
THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

KTH VETENSKAP OCH KONST

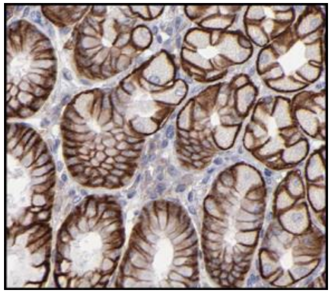
UPPSALA UNIVERSITET

Immunohistochemistry

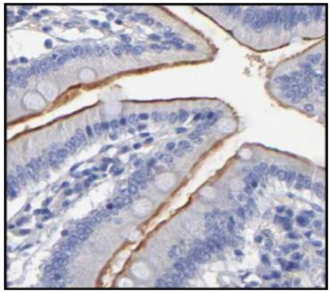


Immunohistochemistry

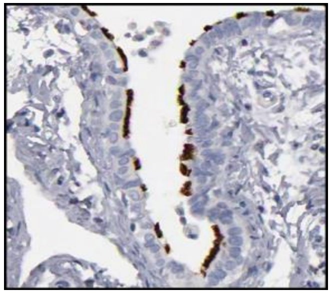
Plasma membrane



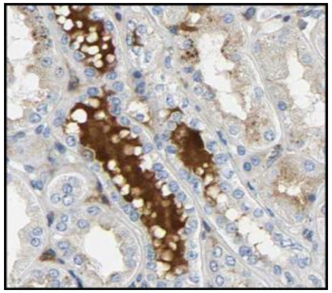
Microvilli



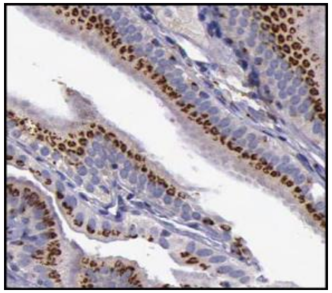
Cilia



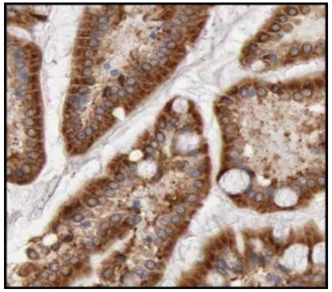
Extra cellular



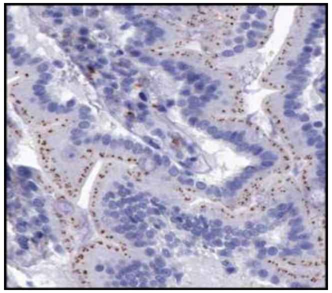
Golgi



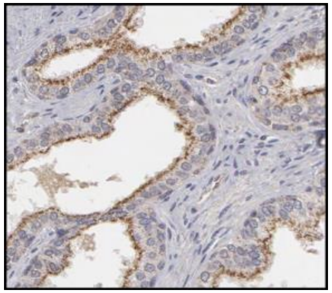
Endoplasmic reticulum



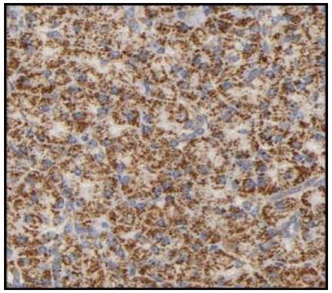
Lysosomes



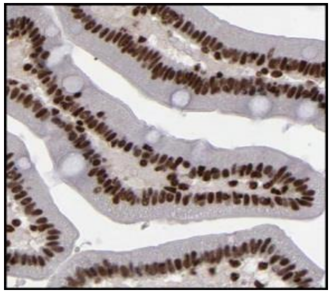
Peroxisomes



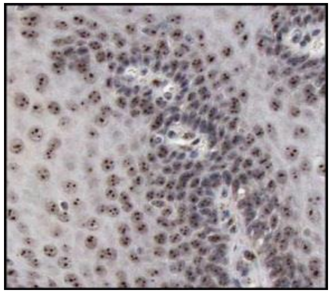
Mitochondria



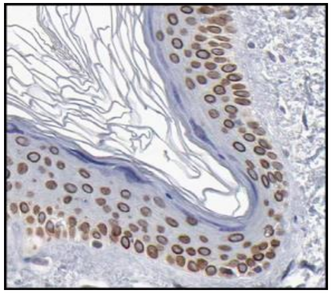
Nuclei



Nucleoli



Nuclear membranes



THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

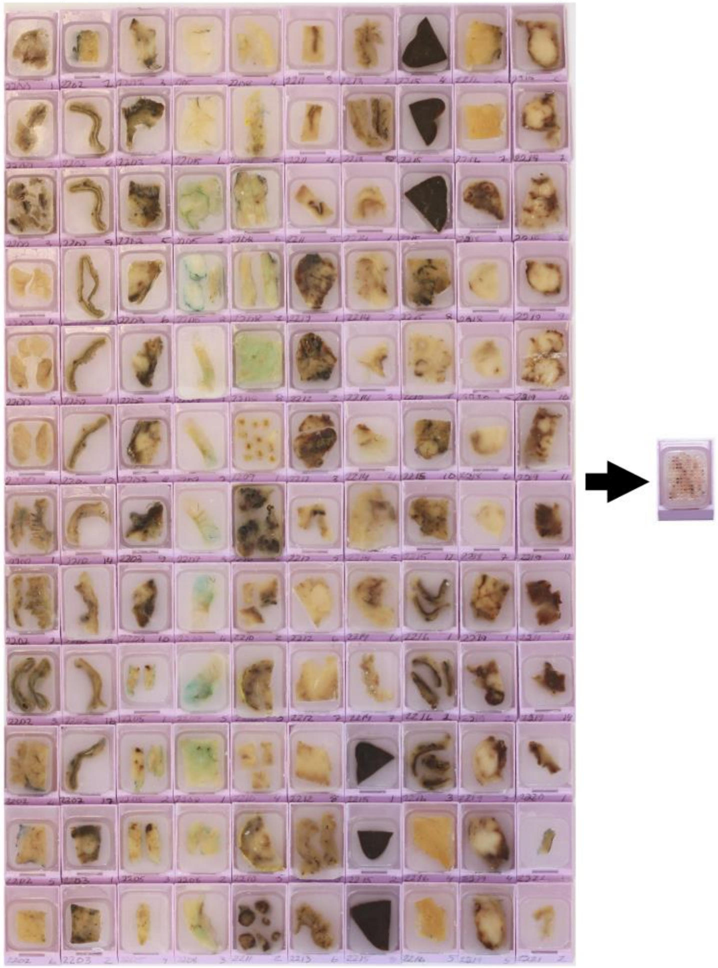
KTH VETENSKAP OCH KONST

UPPSALA UNIVERSITET

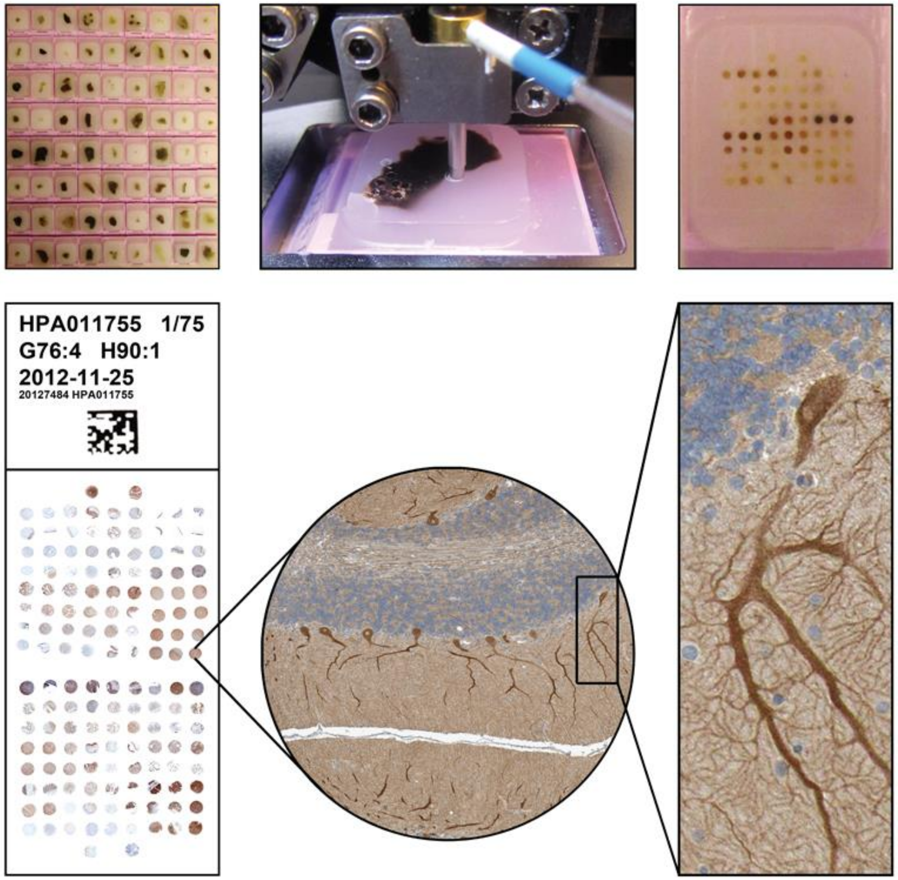
Tissue microarray production

THE PRINCIPLE OF TISSUE MICROARRAYS

Samples from 120 donor blocks can be collected into one single reciever block.



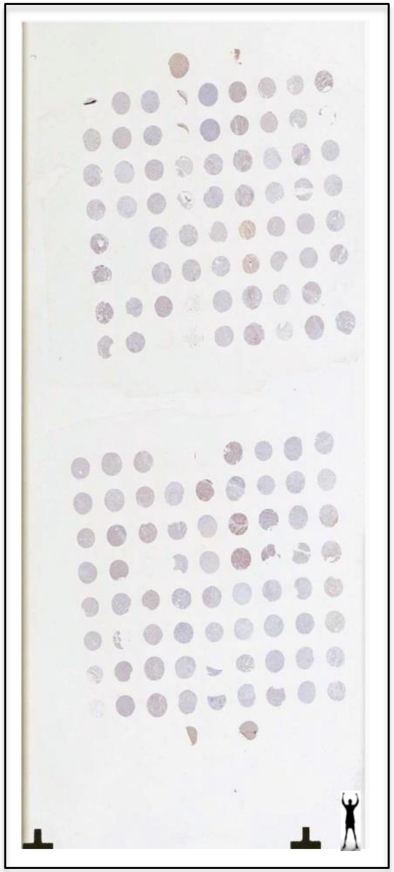
Tissue microarrays



Automated immunohistochemistry

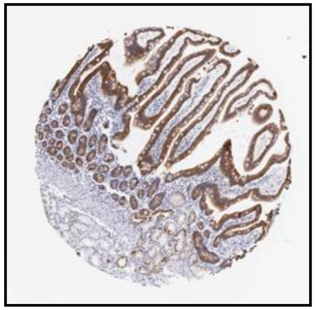
Digital scanning of immunostained TMAs

20x scanning: 50.000 pixels/inch (one HD-TV per mm)



Annotation of images

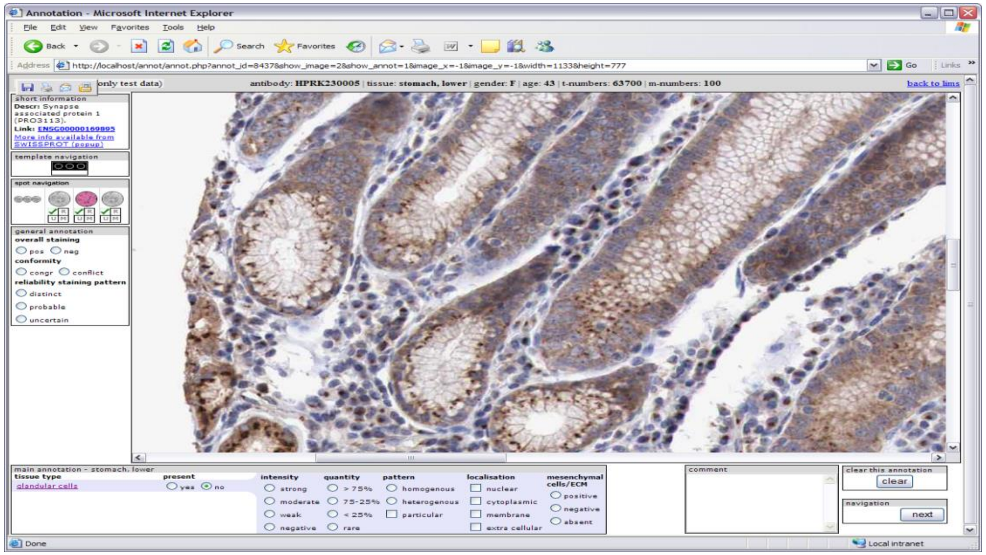
Tissues



Scanning

Cells

Manual annotation



Annotation - Microsoft Internet Explorer

Address http://localhost/annot/annot.php?annot_id=8437&show_image=2&show_annot=1&image_x=-1&image_y=-1&width=1133&height=777

only test data)

antibody: **HPRK230005** | tissue: **stomach, lower** | gender: **F** | age: **43** | t-numbers: **63700** | m-numbers: **100**

back to lims

short information

Descr: Synapse associated protein 1 (PRO3113).

Link: ENSG00000169895

More info available from SWISSPROT (popup)

template navigation

spot navigation

general annotation

overall staining

○ pos ○ neg

conformity

○ congr ○ conflict

reliability staining pattern

○ distinct

○ probable

○ uncertain

main annotation - stomach, lower

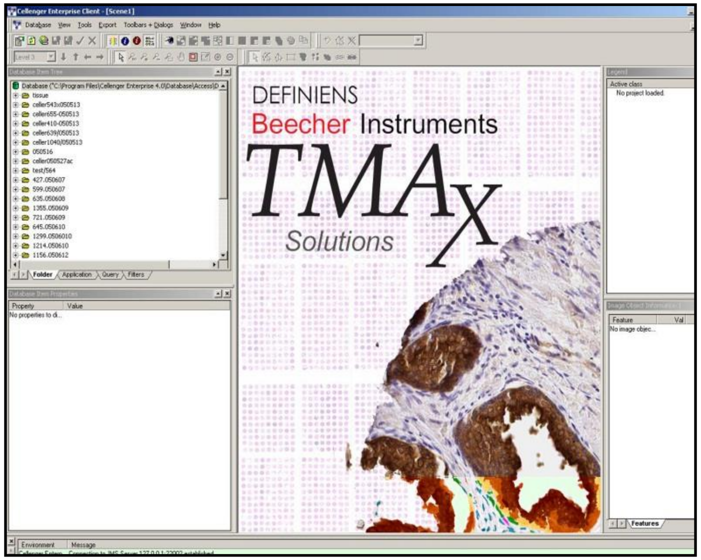
tissue type	present	intensity	quantity	pattern	localisation	mesenchymal cells/ECM
glandular cells	○ yes ● no	○ strong ○ moderate ○ weak ○ negative	○ > 75% ○ 75-25% ○ < 25% ○ rare	○ homogenous ○ heterogenous ☐ particular	☐ nuclear ☐ cytoplasmic ☐ membrane ☐ extra cellular	○ positive ○ negative ○ absent

comment

clear this annotation

navigation

Automated annotation



Cellenger Enterprise Client - [Scene1]

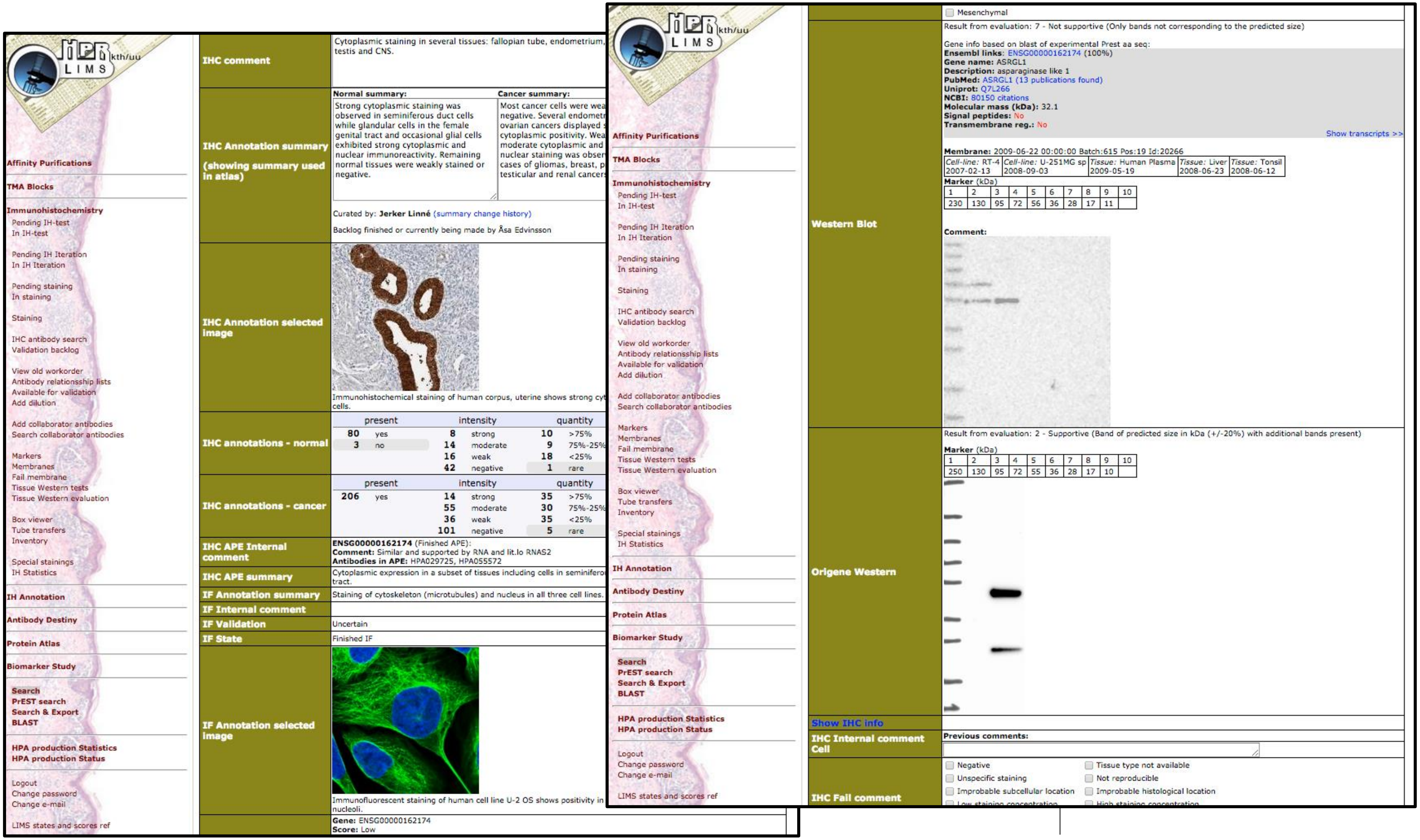
Database ("C:\Program Files\Cellenger Enterprise 4.0\Database\Access\D
- tissue
- celler543x050513
- celler655-050513
- celler410-050513
- celler639/050513
- celler1040/050513
- 050516
- celler050527ac
- test/564
- 427.050607
- 599.050607
- 635.050608
- 1355.050609
- 721.050609
- 645.050610
- 1299.0506010
- 1214.050610
- 1156.050612

Property	Value
No properties to di...	

No project loaded.

Feature	Val
No image objec...	

”Destiny”



IHC comment

Cytoplasmic staining in several tissues: fallopian tube, endometrium, testis and CNS.

IHC Annotation summary (showing summary used in atlas)

Normal summary:

Strong cytoplasmic staining was observed in seminiferous duct cells while glandular cells in the female genital tract and occasional glial cells exhibited strong cytoplasmic and nuclear immunoreactivity. Remaining normal tissues were weakly stained or negative.

Cancer summary:

Most cancer cells were wea negative. Several endometr ovarian cancers displayed s cytoplasmic positivity. Wea moderate cytoplasmic and nuclear staining was obser cases of gliomas, breast, p testicular and renal cancer

Curated by: **Jerker Linné** (summary change history)

Backlog finished or currently being made by Åsa Edvinsson

IHC Annotation selected image

Immunohistochemical staining of human corpus, uterine shows strong cyt cells.

IHC annotations - normal

present		intensity		quantity	
80	yes	8	strong	10	>75%
3	no	14	moderate	9	75%-25%
		16	weak	18	<25%
		42	negative	1	rare

IHC annotations - cancer

present		intensity		quantity	
206	yes	14	strong	35	>75%
		55	moderate	30	75%-25%
		36	weak	35	<25%
		101	negative	5	rare

IHC APE Internal comment

ENSG00000162174 (Finished APE):
Comment: Similar and supported by RNA and lit.lo RNAS2
Antibodies in APE: HPA029725, HPA055572

IHC APE summary

Cytoplasmic expression in a subset of tissues including cells in seminifero tract.

IF Annotation summary

Staining of cytoskeleton (microtubules) and nucleus in all three cell lines.

IF Internal comment

IF Validation

Uncertain

IF State

Finished IF

IF Annotation selected image

Immunofluorescent staining of human cell line U-2 OS shows positivity in nucleoli.

Gene: ENSG00000162174
Score: Low

☐ Mesenchymal

Western Blot

Result from evaluation: 7 - Not supportive (Only bands not corresponding to the predicted size)

Gene info based on blast of experimental Prest aa seq:
Ensembl links: ENSG00000162174 (100%)
Gene name: ASRGL1
Description: asparaginase like 1
PubMed: ASRGL1 (13 publications found)
Uniprot: Q7L266
NCBI: 80150 citations
Molecular mass (kDa): 32.1
Signal peptides: No
Transmembrane reg.: No

Show transcripts >>

Membrane: 2009-06-22 00:00:00 Batch:615 Pos:19 Id:20266

Cell-line: RT-4 2007-02-13	*Cell-line:* U-251MG sp 2008-09-03	*Tissue:* Human Plasma 2009-05-19	*Tissue:* Liver 2008-06-23	*Tissue:* Tonsil 2008-06-12

Marker (kDa)

1	2	3	4	5	6	7	8	9	10
230	130	95	72	56	36	28	17	11	

Comment:

Origene Western

Result from evaluation: 2 - Supportive (Band of predicted size in kDa (+/-20%) with additional bands present)

Marker (kDa)

1	2	3	4	5	6	7	8	9	10
250	130	95	72	55	36	28	17	10	

Show IHC info

IHC Internal comment Cell

Previous comments:

IHC Fail comment

☐ Negative
☐ Unspecific staining
☐ Improbable subcellular location
☐ Low staining concentration
☐ Tissue type not available
☐ Not reproducible
☐ Improbable histological location
☐ High staining concentration

Curation and Annotated protein expression

APE state: **Finished APE**
View in internal atlas | View in RNAseq Genes

HPA029725			
Nasopharynx	respiratory epithelial cells	🟧🟧	🟦
Bronchus	respiratory epithelial cells	🟧🟧	🟦
Lung	macrophages	⬜🟨	⬜
	pneumocytes	⬜⬜	⬜
Oral mucosa	squamous epithelial cells	⬜⬜	⬜
Salivary gland	glandular cells	⬜🟧	🟦
Esophagus	squamous epithelial cells	⬜🟨	⬜
Stomach 1	glandular cells	⬜🟧	🟦
Stomach 2	glandular cells	⬜🟧	🟦
Duodenum	glandular cells	🟨🟨	🟦
Small intestine	glandular cells	⬜🟧	🟦
Colon	endothelial cells	⬜⬛	⬜
	glandular cells	🟨🟨	🟦
	peripheral nerve/ganglion	⬜⬛	⬜
Rectum	glandular cells	⬜🟨	⬜
Appendix	glandular cells	🟨🟧	🟦
	lymphoid tissue	⬜⬜	⬜
Liver	bile duct cells	⬜🟨	⬜
	hepatocytes	⬜🟨	⬜
Gallbladder	glandular cells	🟧🟧	🟦
Pancreas	exocrine glandular cells	🟨🟨	🟦
	islets of Langerhans	⬜⬜	⬜
Thyroid gland	glandular cells	⬜🟧	🟦
Parathyroid gland	glandular cells	⬜⬜	⬜
Adrenal gland	glandular cells	⬜🟨	⬜
Kidney	cells in glomeruli	⬜⬜	⬜
	cells in tubules	🟨🟧	🟦
Urinary bladder	urothelial cells	🟨🟨	🟦
Vagina	squamous epithelial cells	⬜🟨	⬜
Skin 2	epidermal cells	⬜⬜	⬜
Skin 1	fibroblasts	⬜⬜	⬜
	keratinocytes	⬜⬜	⬜
	Langerhans	⬜⬜	⬜
	melanocytes	⬜⬜	⬜
Ovary	follicle cells	⬜⬜	⬜
	ovarian stroma cells	⬜🟨	⬜
Fallopian tube	glandular cells	🟥🟧	🟦
Uterus 2	cells in endometrial stroma	⬜⬜	⬜
	glandular cells	🟥🟥	🟦
Uterus 1	cells in endometrial stroma	⬜⬜	⬜
	glandular cells	🟥🟥	🟦
Cervix, uterine	glandular cells	🟥🟧	🟦
	squamous epithelial cells	⬜⬜	⬜

HPA055572			
Placenta	decidual cells	⬜⬛	⬜
	trophoblastic cells	🟨🟧	🟦
Breast	adipocytes	⬜⬜	⬜
	glandular cells	⬜🟨	🟦
	myoepithelial cells	⬜🟨	🟦
Testis	cells in seminiferous ducts	🟥🟥	🟦
	Leydig cells	🟨🟧	🟦
Epididymis	glandular cells	🟥🟧	🟦
Seminal vesicle	glandular cells	🟨🟨	🟦
Prostate	glandular cells	🟨🟨	🟦
Cerebral cortex	endothelial cells	⬜🟧	🟦
	glial cells	🟨🟧	🟦
	neuronal cells	🟨🟧	🟦
	neuropil	🟨🟨	🟦
Cerebellum	cells in granular layer	🟨🟨	🟦
	cells in molecular layer	⬜🟨	⬜
	Purkinje cells	🟧⬜	⬜
Hippocampus	glial cells	🟧🟧	🟦
	neuronal cells	🟧🟨	🟦
Lateral ventricle	glial cells	🟥🟧	🟦
	neuronal cells	🟨🟨	🟦
Spleen	cells in red pulp	⬜⬜	⬜
	cells in white pulp	⬜⬜	⬜
Lymph node	germinal center cells	⬜⬜	⬜
	non-germinal center cells	⬜⬜	⬜
Tonsil	germinal center cells	⬜⬜	⬜
	non-germinal center cells	⬜⬜	⬜
	squamous epithelial cells	⬜🟨	⬜
Bone marrow	hematopoietic cells	⬜⬜	⬜
Heart muscle	myocytes	⬜🟨	⬜
Smooth muscle	smooth muscle cells	🟨🟨	🟦
Skeletal muscle	myocytes	⬜🟧	🟦
Soft tissue 1	adipocytes	⬜⬜	⬜
	chondrocytes	⬛⬜	⬜
	fibroblasts	⬜🟨	⬜
	peripheral nerve	⬛⬜	⬜
Soft tissue 2	adipocytes	⬜⬜	⬜
	chondrocytes	⬜⬛	⬜
	fibroblasts	⬜⬜	⬜
	peripheral nerve	⬛⬜	⬜

UniProt

Uniprot ID: Q7L266
Protein name: ASGL1_HUMAN
Full name: Isoaspartyl peptidase/L-asparaginase
Protein existence: evidence at protein level
Function: Has both L-asparaginase and beta-aspartyl peptidase activity. May be involved in the production of L-aspartate, which can act as an excitatory neurotransmitter in some brain regions. Is highly active with L-Asp beta-methyl ester. Besides, has catalytic activity toward beta-aspartyl dipeptides and their methyl esters, including beta-L-Asp-L-Phe, beta-L-Asp-L-Phe methyl ester (aspartame), beta-L-Asp-L-Ala, beta-L-Asp-L-Leu and beta-L-Asp-L-Lys. Does not have aspartylglucosaminidase activity and is inactive toward GlcNAc-L-Asn. Likewise, has no activity toward glutamine.
Subcellular location: Cytoplasm. *NOTE*: Midpiece of sperm tail.
Tissue specificity: Expressed in brain, kidney, testis and tissues of the gastrointestinal tract. Present in sperm (at protein level). Over-expressed in uterine, mammary, prostatic and ovarian carcinoma.

RNAseq data

Category: Tissue enhanced
Highest: testis **91.4** fpkm
Lowest: skin 1 **0.6** fpkm

Edited by Linda Oskarsson at 2014-04-24 Annotated protein expression summary:

Cytoplasmic expression in a subset of tissues including cells in seminiferous ducts and glandular cells of female genital tract.

Reliability: High

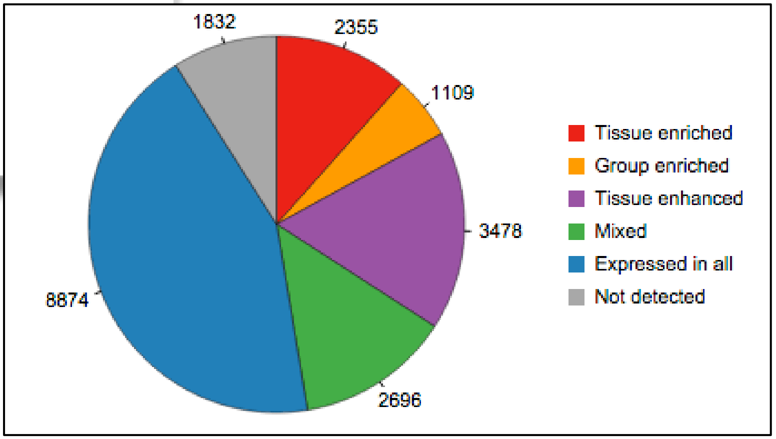
Internal comment:

Similar and supported by RNA and lit.lo
RNAS2

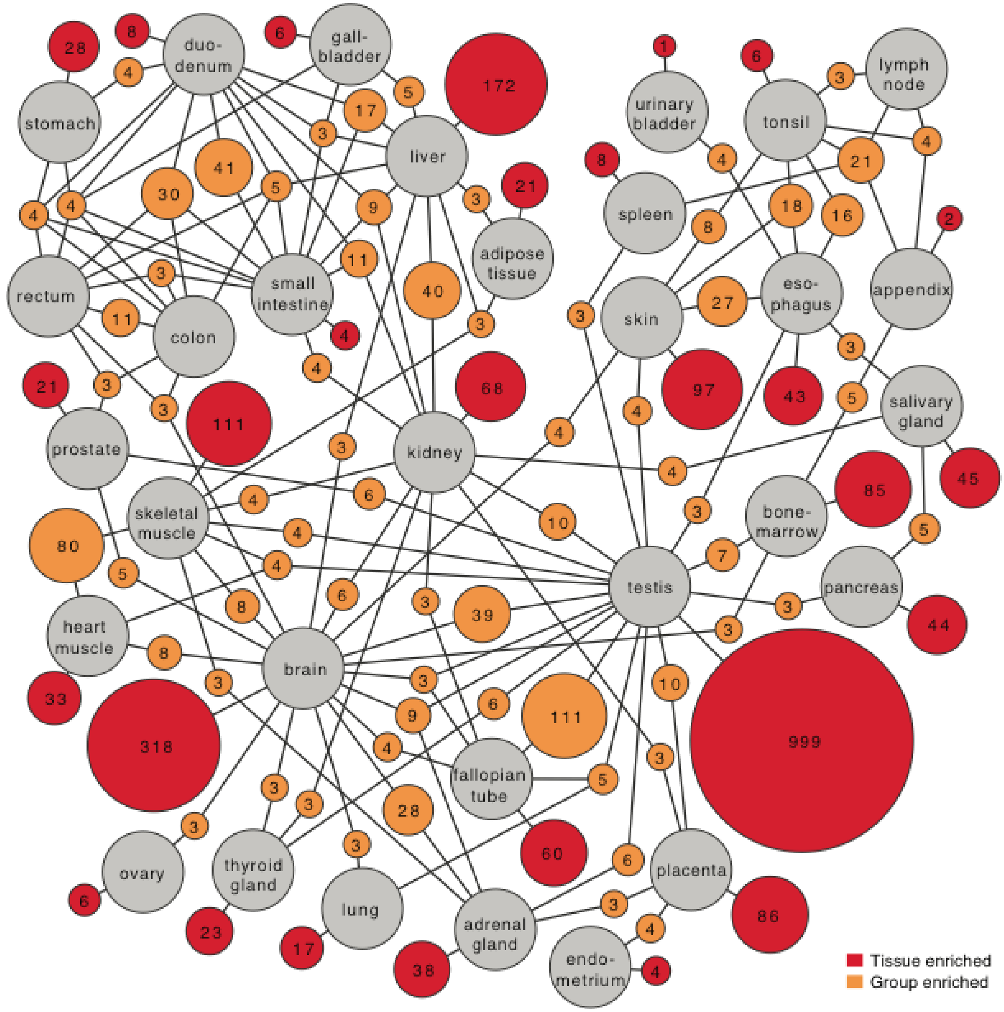
Status: Finished

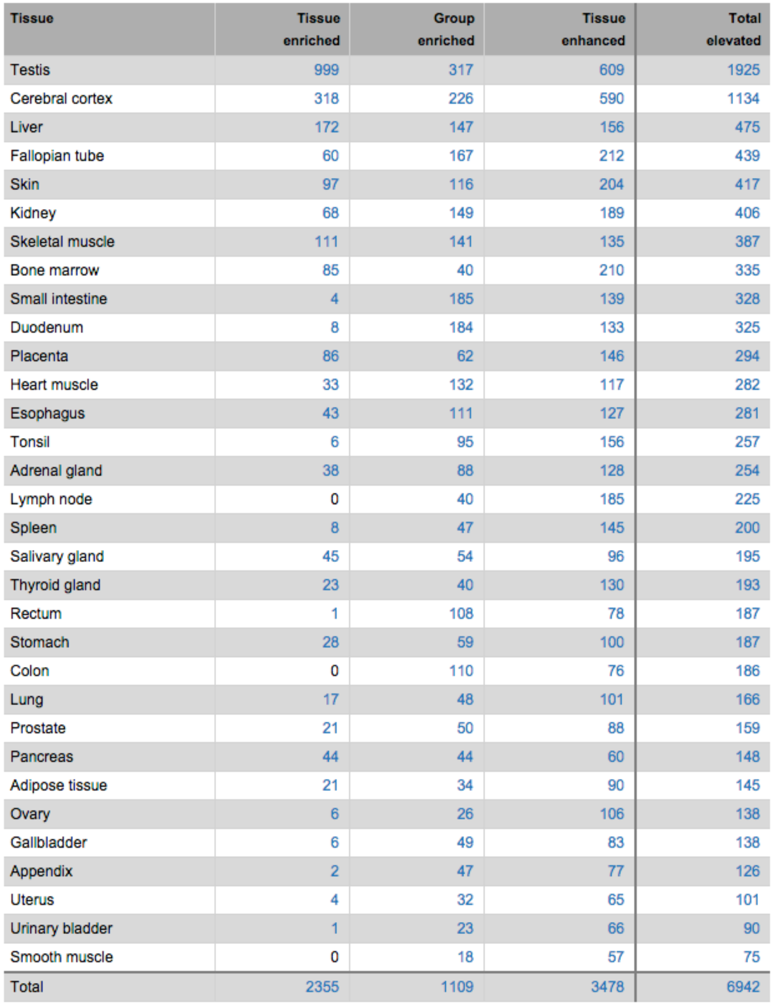
RNA expression profiling of human tissues

Category	Description	Number of genes	Fraction (%)
Tissue enriched	At least five-fold higher mRNA levels in a particular tissue as compared to all other tissues	2,355	12
Group enriched	At least five-fold higher mRNA levels in a group of 2-7 tissues	1,109	5
Tissue enhanced	At least five-fold higher mRNA levels in a particular tissue as compared to average levels in all tissues	3,478	17
Expressed in all	Detected in all tissues (FPKM > 1)	8,874	44
Mixed	Detected in 2-31 tissues but not elevated in any tissue	2,696	13
Not detected	FPKM less than 1 in all tissues	1,832	9
Total	Total number of genes analyzed with RNA-seq	20,344	
Total elevated	Total number of tissue enriched, group enriched and tissue enhanced genes	6,942	



RNA expression profiling of human tissues





Tissue	Tissue enriched	Group enriched	Tissue enhanced	Total elevated
Testis	999	317	609	1925
Cerebral cortex	318	226	590	1134
Liver	172	147	156	475
Fallopian tube	60	167	212	439
Skin	97	116	204	417
Kidney	68	149	189	406
Skeletal muscle	111	141	135	387
Bone marrow	85	40	210	335
Small intestine	4	185	139	328
Duodenum	8	184	133	325
Placenta	86	62	146	294
Heart muscle	33	132	117	282
Esophagus	43	111	127	281
Tonsil	6	95	156	257
Adrenal gland	38	88	128	254
Lymph node	0	40	185	225
Spleen	8	47	145	200
Salivary gland	45	54	96	195
Thyroid gland	23	40	130	193
Rectum	1	108	78	187
Stomach	28	59	100	187
Colon	0	110	76	186
Lung	17	48	101	166
Prostate	21	50	88	159
Pancreas	44	44	60	148
Adipose tissue	21	34	90	145
Ovary	6	26	106	138
Gallbladder	6	49	83	138
Appendix	2	47	77	126
Uterus	4	32	65	101
Urinary bladder	1	23	66	90
Smooth muscle	0	18	57	75
Total	2355	1109	3478	6942

THE HUMAN PROTEIN ATLAS

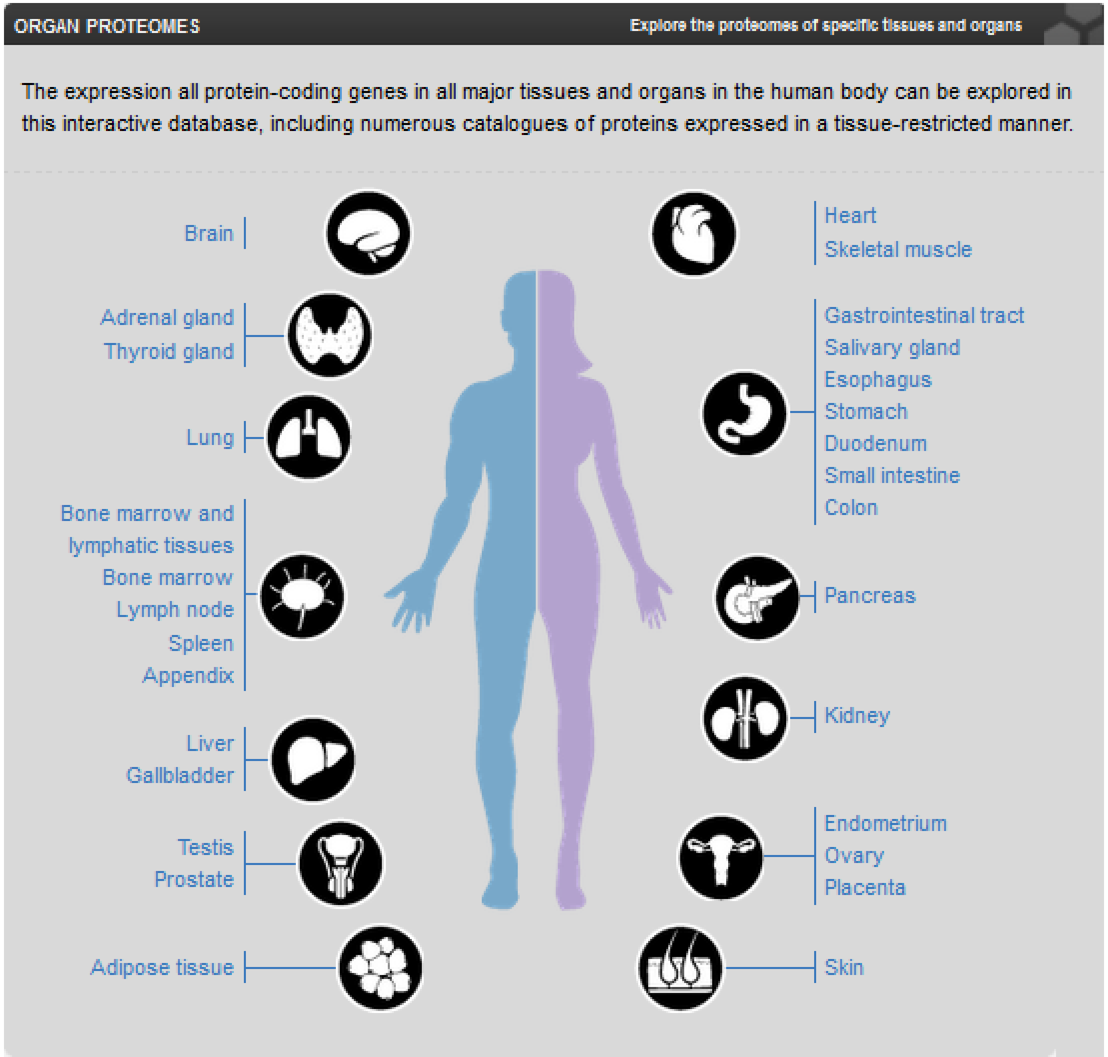
Knut och Alice Wallenbergs Stiftelse

KTH VETENSKAP OCH KONST

UPPSALA UNIVERSITET

THE TISSUE-SPECIFIC PROTEOME

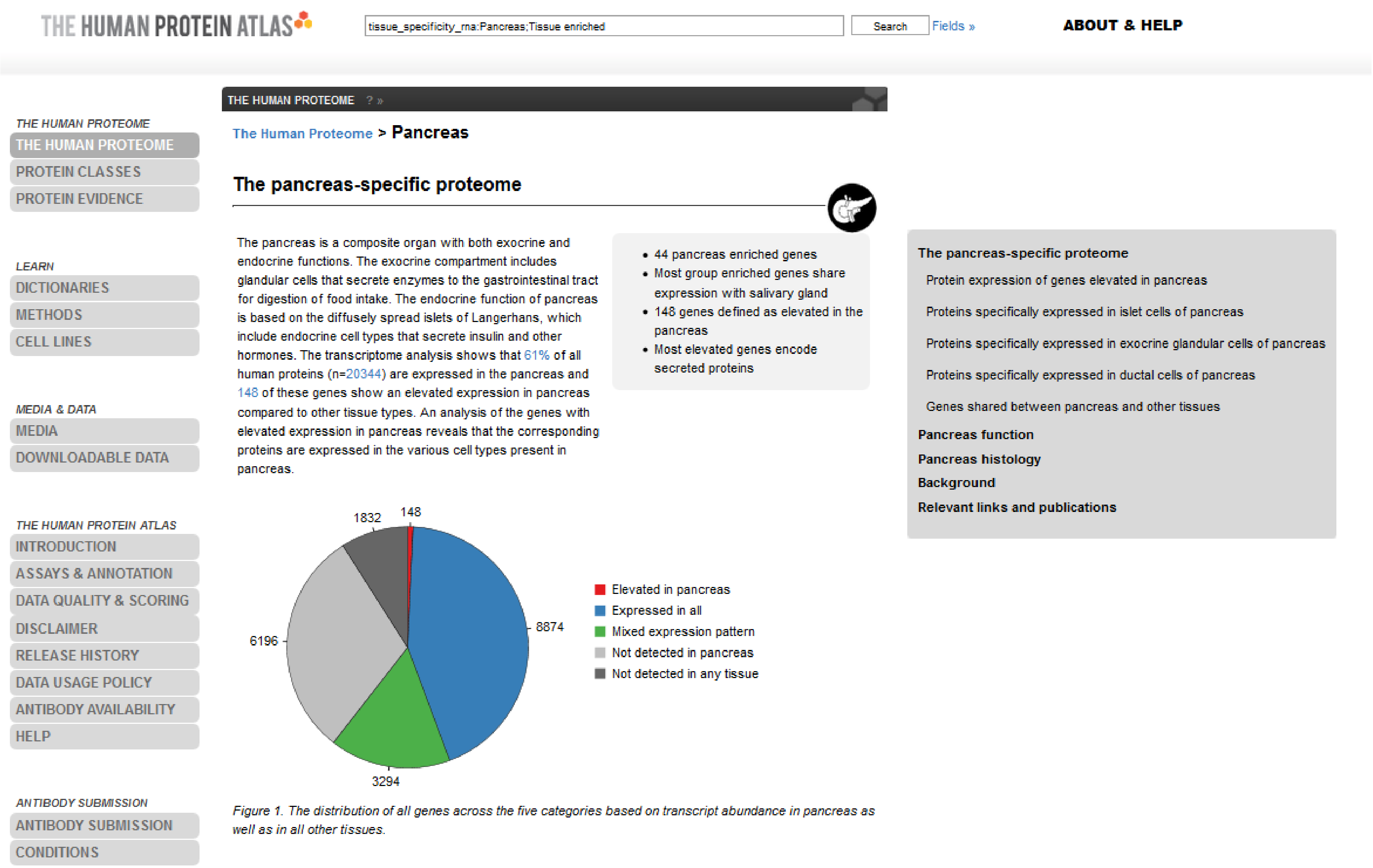
The organ proteomes



ORGAN PROTEOMES

Explore the proteomes of specific tissues and organs

The expression all protein-coding genes in all major tissues and organs in the human body can be explored in this interactive database, including numerous catalogues of proteins expressed in a tissue-restricted manner.



The Human Proteome > **Pancreas**

The pancreas-specific proteome

The pancreas is a composite organ with both exocrine and endocrine functions. The exocrine compartment includes glandular cells that secrete enzymes to the gastrointestinal tract for digestion of food intake. The endocrine function of pancreas is based on the diffusely spread islets of Langerhans, which include endocrine cell types that secrete insulin and other hormones. The transcriptome analysis shows that 61% of all human proteins (n=20344) are expressed in the pancreas and 148 of these genes show an elevated expression in pancreas compared to other tissue types. An analysis of the genes with elevated expression in pancreas reveals that the corresponding proteins are expressed in the various cell types present in pancreas.

- 44 pancreas enriched genes
- Most group enriched genes share expression with salivary gland
- 148 genes defined as elevated in the pancreas
- Most elevated genes encode secreted proteins

Figure 1. The distribution of all genes across the five categories based on transcript abundance in pancreas as well as in all other tissues.

Table 1. The genes with elevated expression in pancreas

Category	Number of genes	Description
Tissue enriched	44	At least five-fold higher mRNA levels in a particular tissue as compared to all other tissues
Group enriched	44	At least five-fold higher mRNA levels in a group of 2-7 tissues
Tissue enhanced	60	At least five-fold higher mRNA levels in a particular tissue as compared to average levels in all tissues
Total	148	Total number of elevated genes in pancreas

Table 2. The 12 genes with the highest level of enriched expression in pancreas. "Predicted localization" shows the classification of each gene into three main classes: Secreted, Membrane, and Intracellular, where the latter consists of genes without any predicted membrane and secreted features. "mRNA (tissue)" shows the transcript level as FPKM values, TS-score (Tissue Specificity score) corresponds to the score calculated as the fold change to the second highest tissue.

Gene	Description	Predicted localization	mRNA (tissue)	TS-score
CTRB2	chymotrypsinogen B2	Secreted	52212.1	863
PNLIPRP1	pancreatic lipase-related protein 1	Secreted	4011.7	837
CTRB1	chymotrypsinogen B1	Secreted	58491.4	788
CELA2B	chymotrypsin-like elastase family, member 2B	Secreted	7104.7	709
CUZD1	CUB and zona pellucida-like domains 1	Membrane,Secreted	981.3	690
SYCN	syncollin	Secreted	5584.1	545
CELA3A	chymotrypsin-like elastase family, member 3A	Secreted	51227.1	519
CELA2A	chymotrypsin-like elastase family, member 2A	Secreted	25078.3	518
CELA3B	chymotrypsin-like elastase family, member 3B	Secreted	17378.1	496
CTRC	chymotrypsin C (caldecrin)	Secreted	13991.8	401
CEL	carboxyl ester lipase	Secreted	17834.4	387
CTRL	chymotrypsin-like	Secreted	2349.5	344

Some of the proteins predicted to be membrane-spanning are intracellular, e.g., in the Golgi or mitochondrial membranes, and some of the proteins predicted to be secreted can potentially be retained in a compartment belonging to the secretory pathway, such as the ER, or remain attached to the outer face of the cell membrane by a GPI anchor.

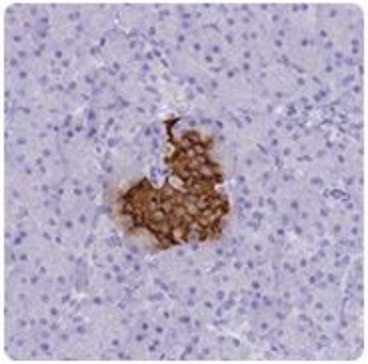
Protein expression of genes elevated in pancreas

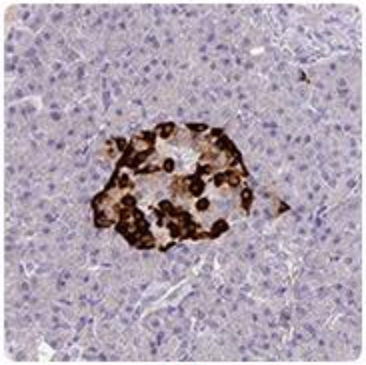
Gene Ontology-based analysis of all the 148 genes elevated in pancreas indicates a clear overrepresentation of proteins associated with metabolic processes, proteolysis and digestion. A majority of the 148 genes encode secreted proteins.

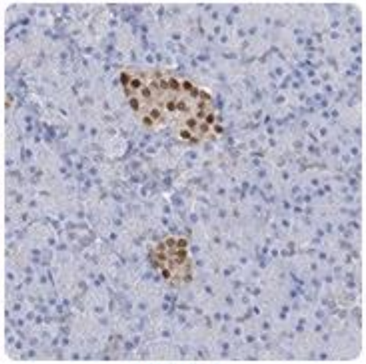
In-depth analysis of the elevated genes in pancreas using antibody-based protein profiling allowed us to create a map of where these proteins are expressed within the pancreas.

Proteins specifically expressed in islet cells of pancreas

The islet cells constitute 2% of the pancreas, and are responsible for maintaining a steady blood glucose level by secreting hormones regulating uptake and release of glucose. Examples of proteins expressed in islet cells include INS, which is secreted following elevated blood glucose levels, and GCG, which acts in the opposite direct raising blood glucose levels. Another important protein is the transcription factor NKX6-1, essential for development of beta cells.

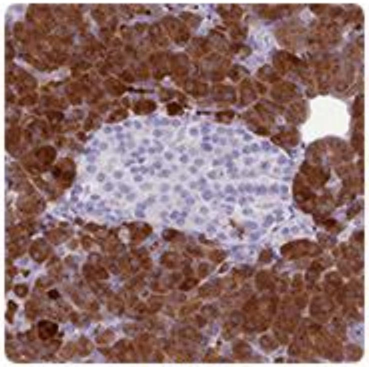

INS

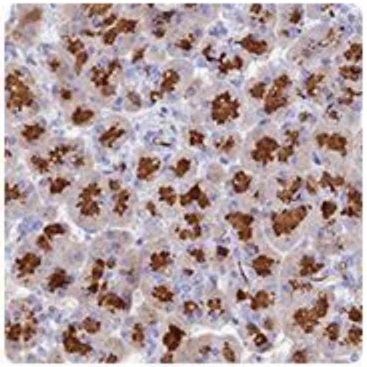

GCG


NKX6-1

Proteins specifically expressed in exocrine glandular cells of pancreas

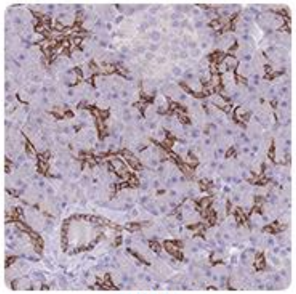
The exocrine part of the pancreas is mainly composed of exocrine glandular cells and ductal cells. The main function of the exocrine glandular cells is to produce digestive enzymes and Cl^- rich fluid for transportation of enzymes. Examples of proteins expressed in exocrine glandular cells include AMY2A and CELA3A, both involved in enzymatic digestion of proteins and lipids.


AMY2A


CELA3A

Proteins specifically expressed in ductal cells of pancreas

The enzymes secreted by exocrine glandular cells are transported through the pancreatic ductal system into the main pancreatic duct that ends in the duodenum. The ductal epithelium secretes HCO_3^- rich fluid for regulation of the pH. One example of a protein expressed in ductal cells is SLC4A4, which acts as a Na^+/HCO_3^- cotransporter.

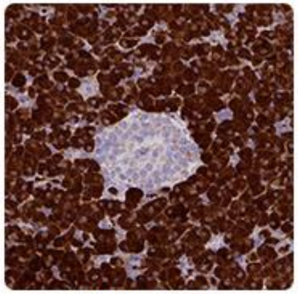


SLC4A4

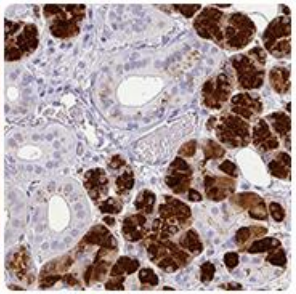
Genes shared between pancreas and other tissues

There are 44 group enriched genes expressed in the pancreas. Group enriched genes are defined as genes showing a 5-fold higher average level of mRNA expression in a group of 2-7 tissues, including pancreas, compared to all other tissues.

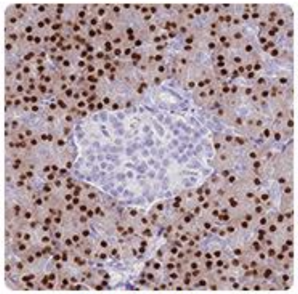
Pancreas shares nine genes with salivary gland, an organ with exocrine function highly correlated with the exocrine pancreas. One example of a group enriched gene shared between pancreas and salivary gland is the digestion enzyme AMY1C. Another example is BHLHA15, a transcription factor regulating acinar cell function.



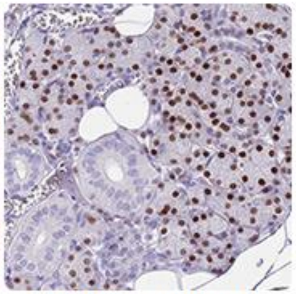
AMY1C - pancreas



AMY1C - salivary gland



BHLHA15 - pancreas



BHLHA15 - salivary gland

Pancreas function

The pancreas is a mixed exocrine/endocrine gland with dual functions essential for maintaining physiological levels of blood glucose and for digestion of food intake. The underlying cell types that execute these diverse functions are exocrine cells, responsible for storing enzymes, and cells in islets of Langerhans, synthesizing different hormones.

Pancreas histology

The exocrine component is composed of lobular units of acini, that discharge their secretions into progressively larger ducts that finally merge into the main pancreatic duct, which ends in duodenum. The pyramidal-shaped acinar cells are filled by eosinophilic zymogen granules.

The islets of Langerhans, which constitute 1-2% of the cell mass in the adult pancreas, represent the endocrine component of the pancreas. The islets are round, compact structures that are highly vascularised with sparse connective tissue. The main cell types in the islets are beta cells - responsible for insulin production, alpha cells - responsible for glucagon secretion, delta cells - responsible for somatostatin secretion and PP cells - the pancreatic polypeptide secreting cells.

The histology of human pancreas including detailed images and information about the different cell types can be viewed in the Protein Atlas Histology Dictionary.

Background

Here, the protein-coding genes expressed in the pancreas are described and characterized, together with examples of immunohistochemically stained tissue sections that visualize protein expression patterns of proteins that correspond to genes with elevated expression in the pancreas.

Transcript profiling and RNA-data analyses based on normal human tissues have been described previously (Fagerberg et al., 2013). Analyses of mRNA expression including over 99% of all human protein-coding genes was performed using deep RNA sequencing of 122 individual samples corresponding to 32 different human normal tissue types. RNA sequencing results of 2 fresh frozen tissues representing normal pancreas was compared to 120 other tissue samples corresponding to 31 tissue types, in order to determine genes with elevated expression in pancreas. A tissue-specific score, defined as the ratio between mRNA levels in pancreas compared to the mRNA levels in all other tissues, was used to divide the genes into different categories of expression. These categories include: genes with elevated expression in pancreas, genes expressed in all tissues, genes with a mixed expression pattern, genes not expressed in pancreas, and genes not expressed in any tissue. Genes with elevated expression in pancreas were further sub-categorized as i) genes with enriched expression in pancreas, ii) genes with group enriched expression including pancreas and iii) genes with enhanced expression in pancreas.

Human tissue samples used for protein and mRNA expression analyses were collected and handled in accordance with Swedish laws and regulation and obtained from the Department of Pathology, Uppsala University Hospital, Uppsala, Sweden as part of the sample collection governed by the Uppsala Biobank. All human tissue samples used in the present study were anonymized in accordance with approval and advisory report from the Uppsala Ethical Review Board.

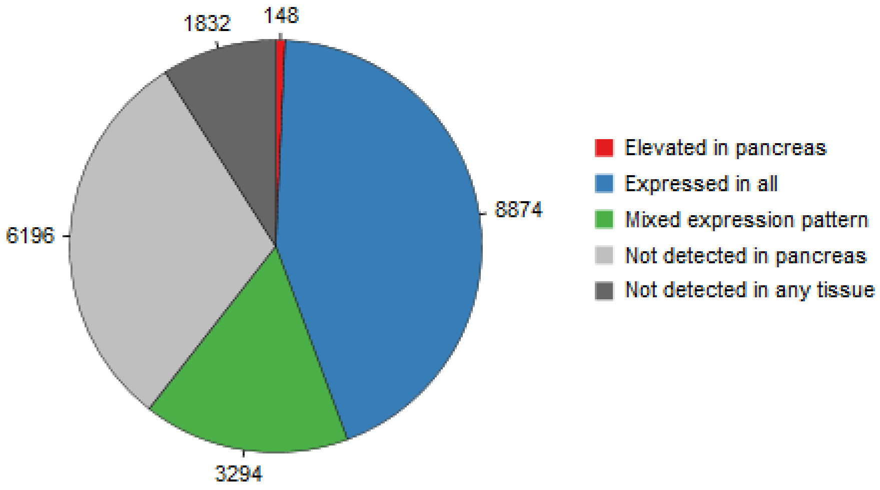
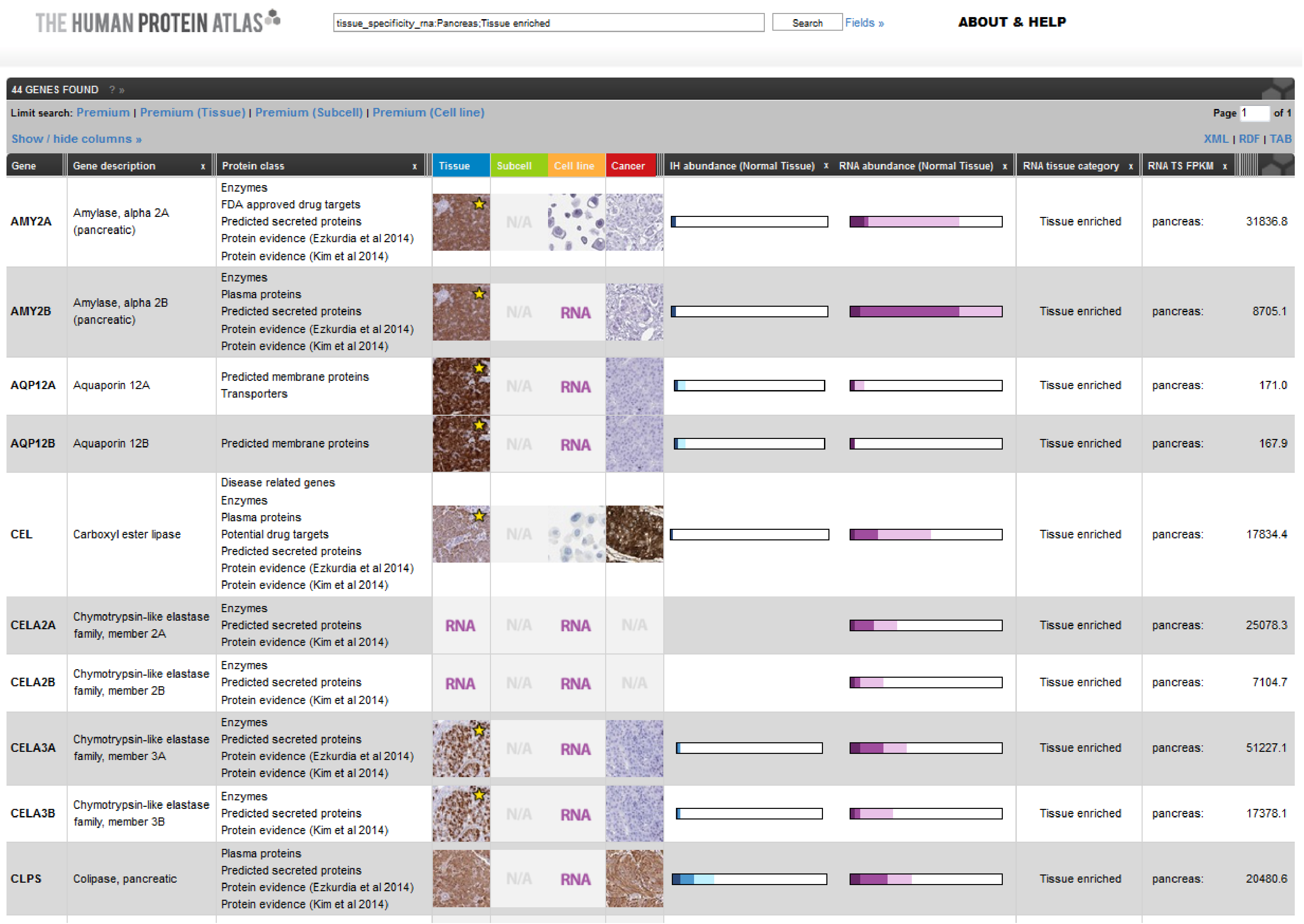


Figure 1. The distribution of all genes across the five categories based on transcript abundance in pancreas as well as in all other tissues.

148 genes show some level of elevated expression in the pancreas compared to other tissues. The three categories of genes with elevated expression in pancreas compared to other organs are shown in Table 1.

Table 1. The genes with elevated expression in pancreas

Category	Number of genes	Description
Tissue enriched	44	At least five-fold higher mRNA levels in a particular tissue as compared to all other tissues
Group enriched	44	At least five-fold higher mRNA levels in a group of 2-7 tissues
Tissue enhanced	60	At least five-fold higher mRNA levels in a particular tissue as compared to average levels in all tissues
Total	148	Total number of elevated genes in pancreas



44 GENES FOUND ? »

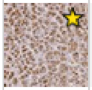
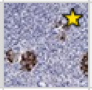
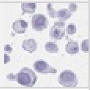
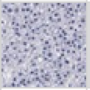
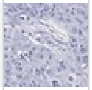
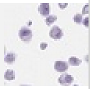
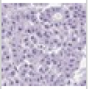
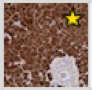
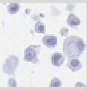
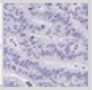
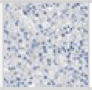
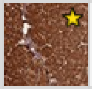
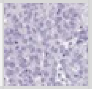
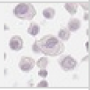
Limit search: Premium | Premium (Tissue) | Premium (Subcell) | Premium (Cell line)

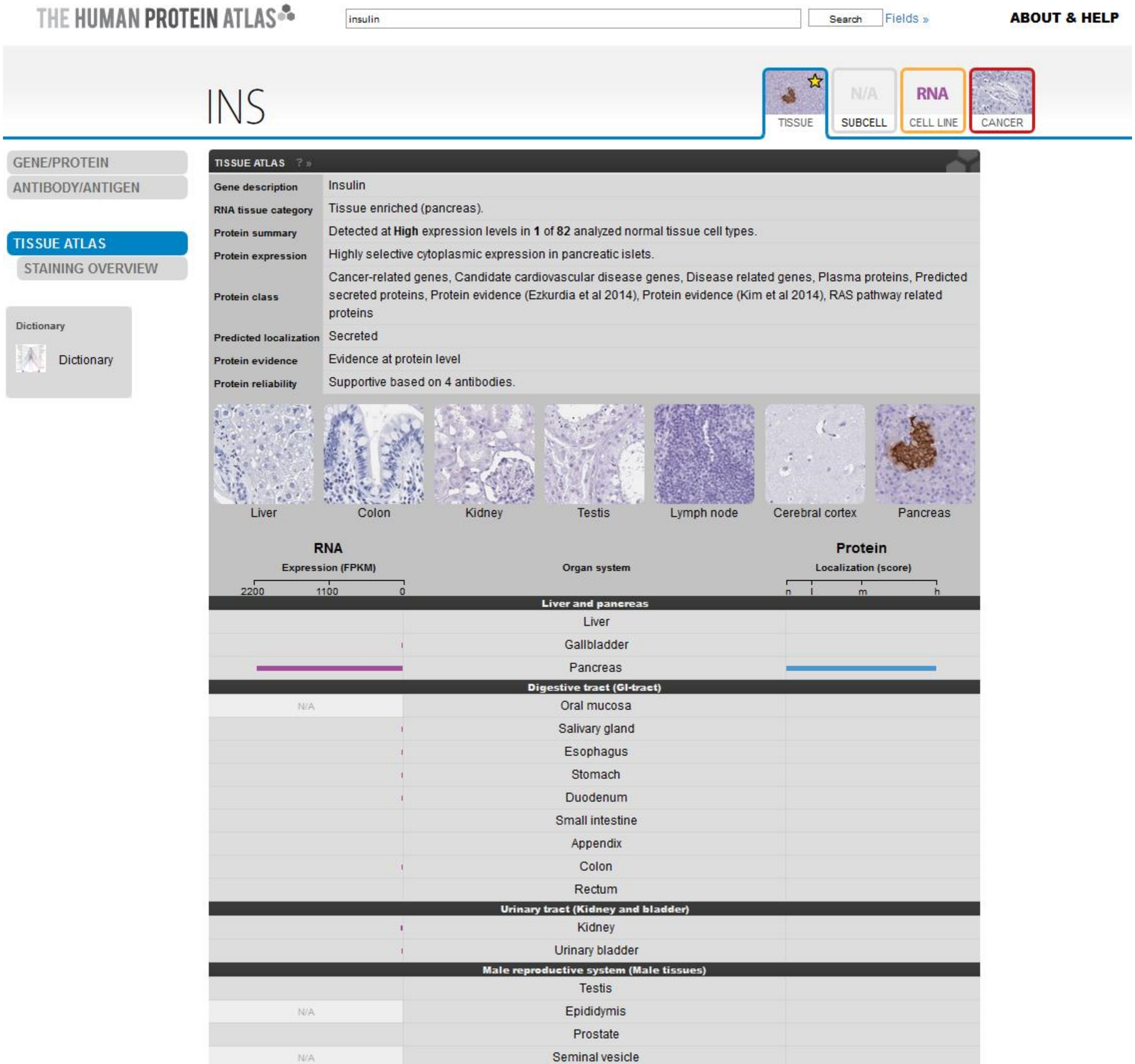
Page 1 of 1

Show / hide columns »

XML | RDF | TAB

Gene	Gene description	Protein class	Tissue	Subcell	Cell line	Cancer	IH abundance (Normal Tissue)	RNA abundance (Normal Tissue)	RNA tissue category	RNA TS FPKM
AMY2A	Amylase, alpha 2A (pancreatic)	Enzymes FDA approved drug targets Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A					Tissue enriched	pancreas: 31836.8
AMY2B	Amylase, alpha 2B (pancreatic)	Enzymes Plasma proteins Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A	RNA				Tissue enriched	pancreas: 8705.1
AQP12A	Aquaporin 12A	Predicted membrane proteins Transporters		N/A	RNA				Tissue enriched	pancreas: 171.0
AQP12B	Aquaporin 12B	Predicted membrane proteins		N/A	RNA				Tissue enriched	pancreas: 167.9
CEL	Carboxyl ester lipase	Disease related genes Enzymes Plasma proteins Potential drug targets Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A					Tissue enriched	pancreas: 17834.4
CELA2A	Chymotrypsin-like elastase family, member 2A	Enzymes Predicted secreted proteins Protein evidence (Kim et al 2014)	RNA	N/A	RNA	N/A			Tissue enriched	pancreas: 25078.3
CELA2B	Chymotrypsin-like elastase family, member 2B	Enzymes Predicted secreted proteins Protein evidence (Kim et al 2014)	RNA	N/A	RNA	N/A			Tissue enriched	pancreas: 7104.7
CELA3A	Chymotrypsin-like elastase family, member 3A	Enzymes Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A	RNA				Tissue enriched	pancreas: 51227.1
CELA3B	Chymotrypsin-like elastase family, member 3B	Enzymes Predicted secreted proteins Protein evidence (Kim et al 2014)		N/A	RNA				Tissue enriched	pancreas: 17378.1
CLPS	Colipase, pancreatic	Plasma proteins Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A	RNA				Tissue enriched	pancreas: 20480.6

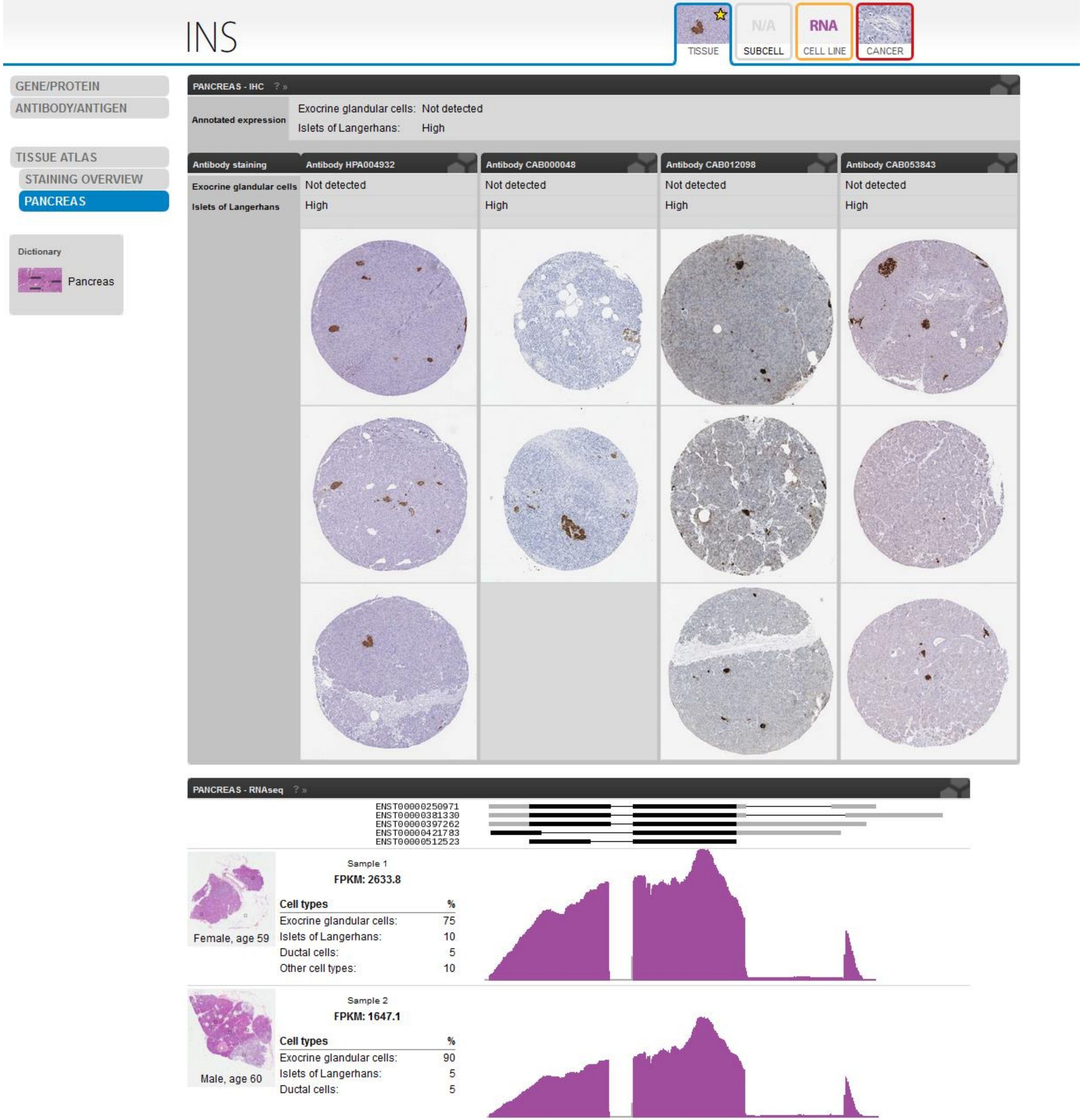
GP2	Glycoprotein 2 (zymogen granule membrane)	Predicted membrane proteins Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A	RNA		Tissue enriched	pancreas:	8760.6
GPHA2	Glycoprotein hormone alpha 2	Predicted secreted proteins	RNA	N/A	RNA	N/A	Tissue enriched	pancreas:	164.4
GRPR	Gastrin-releasing peptide receptor	G-protein coupled receptors Predicted membrane proteins Protein evidence (Ezkurdia et al 2014)	RNA	N/A	RNA	N/A	Tissue enriched	pancreas:	8.9
IAPP	Islet amyloid polypeptide	Predicted secreted proteins		N/A			Tissue enriched	pancreas:	77.1
INS	Insulin	Cancer-related genes Candidate cardiovascular disease genes Disease related genes Plasma proteins Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) RAS pathway related proteins		N/A	RNA		Tissue enriched	pancreas:	2140.4
KIRREL2	Kin of IRRE like 2 (Drosophila)	Predicted membrane proteins Predicted secreted proteins Protein evidence (Kim et al 2014)	RNA	N/A	RNA	N/A	Tissue enriched	pancreas:	23.0
PDIA2	Protein disulfide isomerase family A, member 2	Enzymes Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A			Tissue enriched	pancreas:	915.7
PLA2G1B	Phospholipase A2, group IB (pancreas)	Candidate cardiovascular disease genes Enzymes FDA approved drug targets Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) RAS pathway related proteins		N/A			Tissue enriched	pancreas:	13646.6
PM20D1	Peptidase M20 domain containing 1	Enzymes Predicted secreted proteins		N/A	RNA		Tissue enriched	pancreas:	42.9
PNLIP	Pancreatic lipase	Enzymes FDA approved drug targets Plasma proteins Predicted secreted proteins Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014)		N/A	RNA		Tissue enriched	pancreas:	33795.0
PNLIPRP1	Pancreatic lipase-related protein 1	Predicted secreted proteins Protein evidence (Kim et al 2014)	RNA	N/A	RNA	N/A	Tissue enriched	pancreas:	4011.7
PPY	Pancreatic polypeptide	Cancer-related genes Predicted secreted proteins		N/A			Tissue enriched	pancreas:	82.5



INS

Gene description	Insulin
RNA tissue category	Tissue enriched (pancreas).
Protein summary	Detected at **High** expression levels in **1** of **82** analyzed normal tissue cell types.
Protein expression	Highly selective cytoplasmic expression in pancreatic islets.
Protein class	Cancer-related genes, Candidate cardiovascular disease genes, Disease related genes, Plasma proteins, Predicted secreted proteins, Protein evidence (Ezkurdia et al 2014), Protein evidence (Kim et al 2014), RAS pathway related proteins
Predicted localization	Secreted
Protein evidence	Evidence at protein level
Protein reliability	Supportive based on 4 antibodies.

RNA	Organ system	Protein
	Liver and pancreas	
	Liver	
	Gallbladder	
	Pancreas	
	Digestive tract (GI-tract)	
N/A	Oral mucosa	
	Salivary gland	
	Esophagus	
	Stomach	
	Duodenum	
	Small intestine	
	Appendix	
	Colon	
	Rectum	
	Urinary tract (Kidney and bladder)	
	Kidney	
	Urinary bladder	
	Male reproductive system (Male tissues)	
	Testis	
N/A	Epididymis	
	Prostate	
N/A	Seminal vesicle	



INS

PANCREAS - IHC

Annotated expression	
Exocrine glandular cells:	Not detected
Islets of Langerhans:	High

Antibody staining	Antibody HPA004932	Antibody CAB000048	Antibody CAB012098	Antibody CAB053843
Exocrine glandular cells	Not detected	Not detected	Not detected	Not detected
Islets of Langerhans	High	High	High	High

PANCREAS - RNAseq

Sample 1

FPKM: 2633.8

Female, age 59

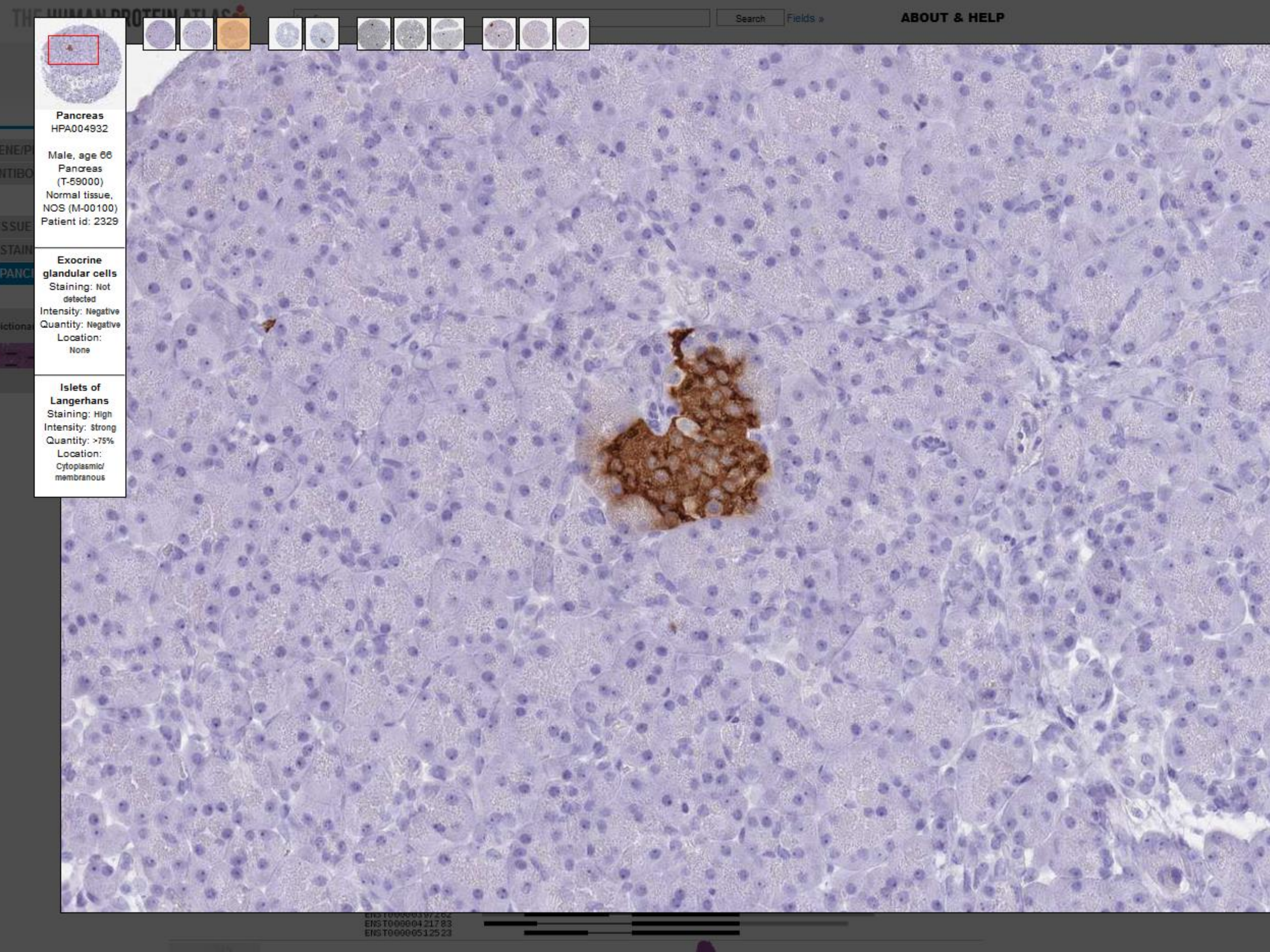
Cell types	%
Exocrine glandular cells:	75
Islets of Langerhans:	10
Ductal cells:	5
Other cell types:	10

Sample 2

FPKM: 1647.1

Male, age 60

Cell types	%
Exocrine glandular cells:	90
Islets of Langerhans:	5
Ductal cells:	5

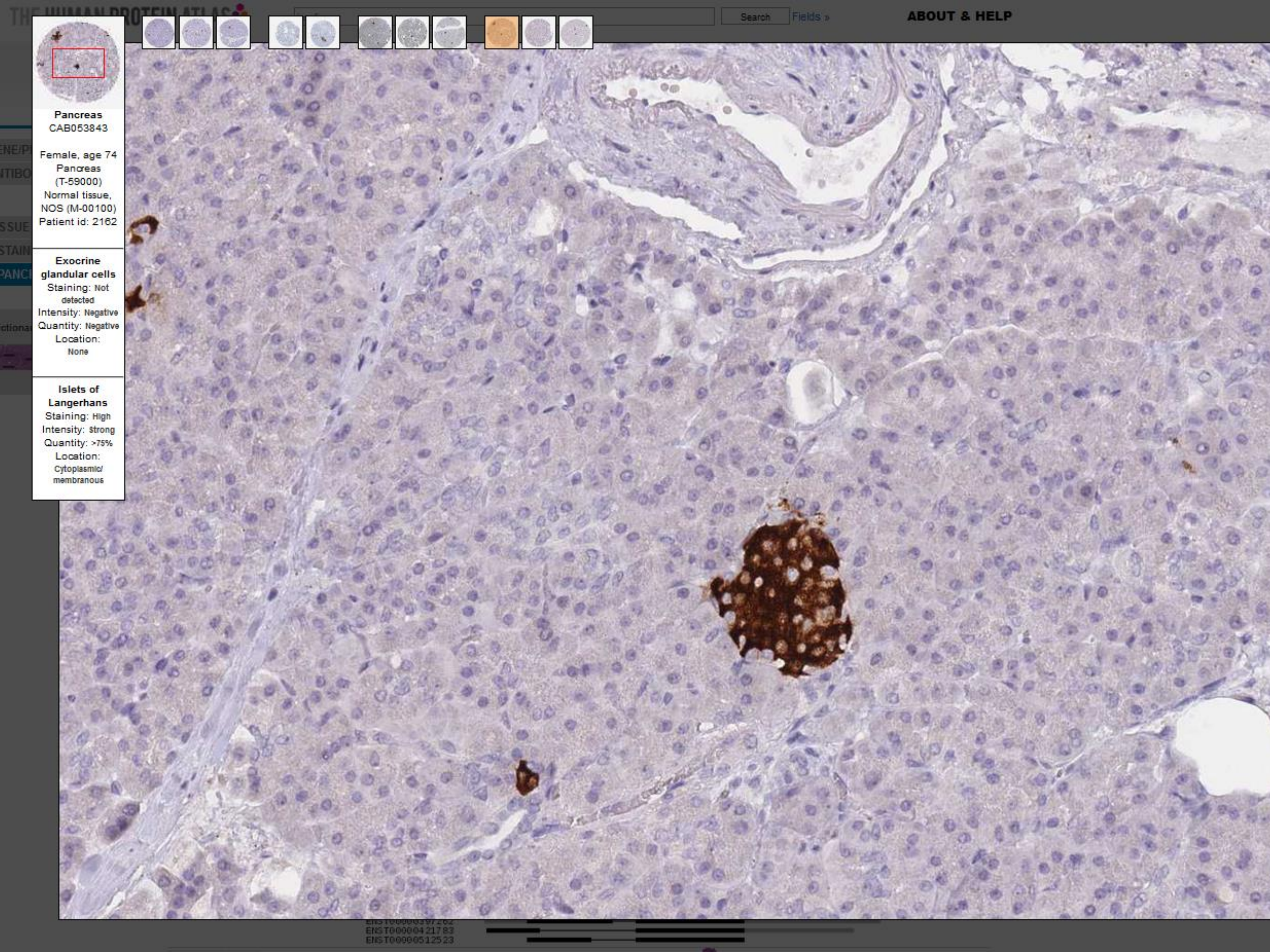


Pancreas
HPA004932

Male, age 66
Pancreas
(T-59000)
Normal tissue,
NOS (M-00100)
Patient id: 2329

Exocrine glandular cells
Staining: Not detected
Intensity: Negative
Quantity: Negative
Location:
None

Islets of Langerhans
Staining: High
Intensity: Strong
Quantity: >75%
Location:
Cytoplasmic/membranous

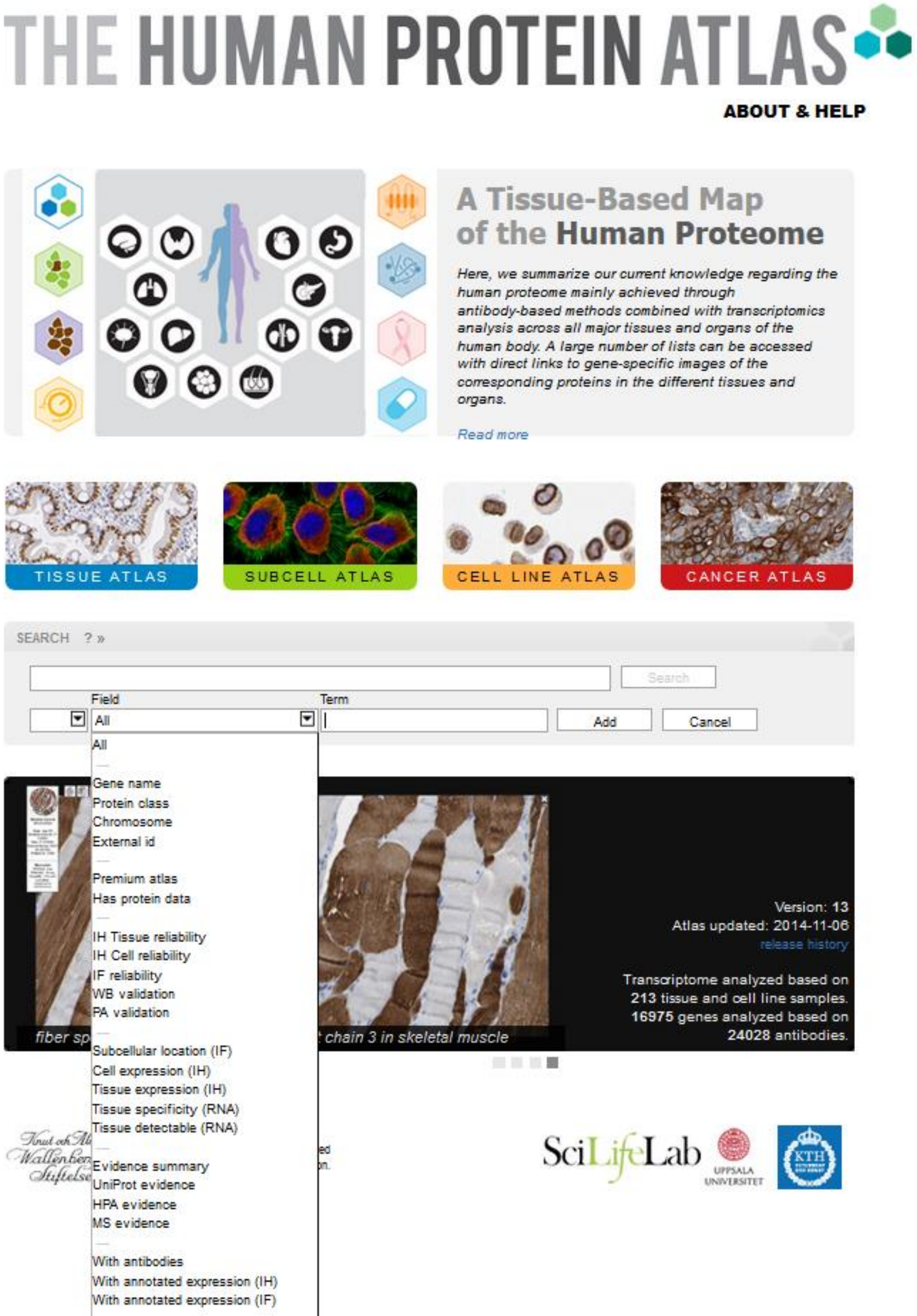


Pancreas
CAB053843

Female, age 74
Pancreas
(T-59000)
Normal tissue,
NOS (M-00100)
Patient id: 2162

Exocrine glandular cells
Staining: Not detected
Intensity: Negative
Quantity: Negative
Location:
None

Islets of Langerhans
Staining: High
Intensity: Strong
Quantity: >75%
Location:
Cytoplasmic/membranous



THE HUMAN PROTEIN ATLAS

ABOUT & HELP

A Tissue-Based Map of the Human Proteome

Here, we summarize our current knowledge regarding the human proteome mainly achieved through antibody-based methods combined with transcriptomics analysis across all major tissues and organs of the human body. A large number of lists can be accessed with direct links to gene-specific images of the corresponding proteins in the different tissues and organs.

Read more

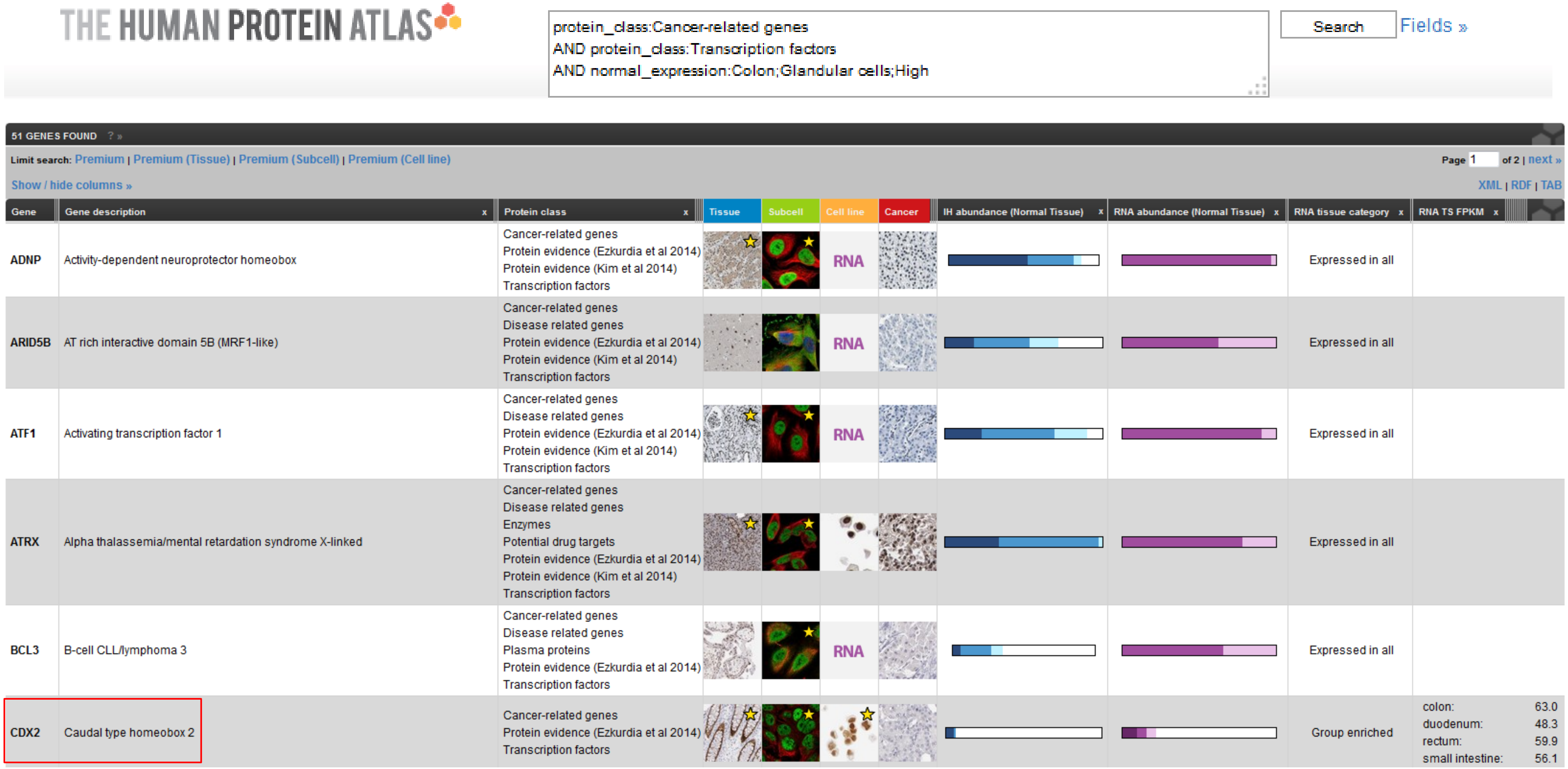
Field: All

- All
- Gene name
- Protein class
- Chromosome
- External id
- Premium atlas
- Has protein data
- IH Tissue reliability
- IH Cell reliability
- IF reliability
- WB validation
- PA validation
- Subcellular location (IF)
- Cell expression (IH)
- Tissue expression (IH)
- Tissue specificity (RNA)
- Tissue detectable (RNA)
- Evidence summary
- UniProt evidence
- HPA evidence
- MS evidence
- With antibodies
- With annotated expression (IH)
- With annotated expression (IF)

fiber sp … *chain 3 in skeletal muscle*

Version: 13
Atlas updated: 2014-11-06
release history

Transcriptome analyzed based on **213** tissue and cell line samples. **16975** genes analyzed based on **24028** antibodies.



protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon;Glandular cells;High

51 GENES FOUND

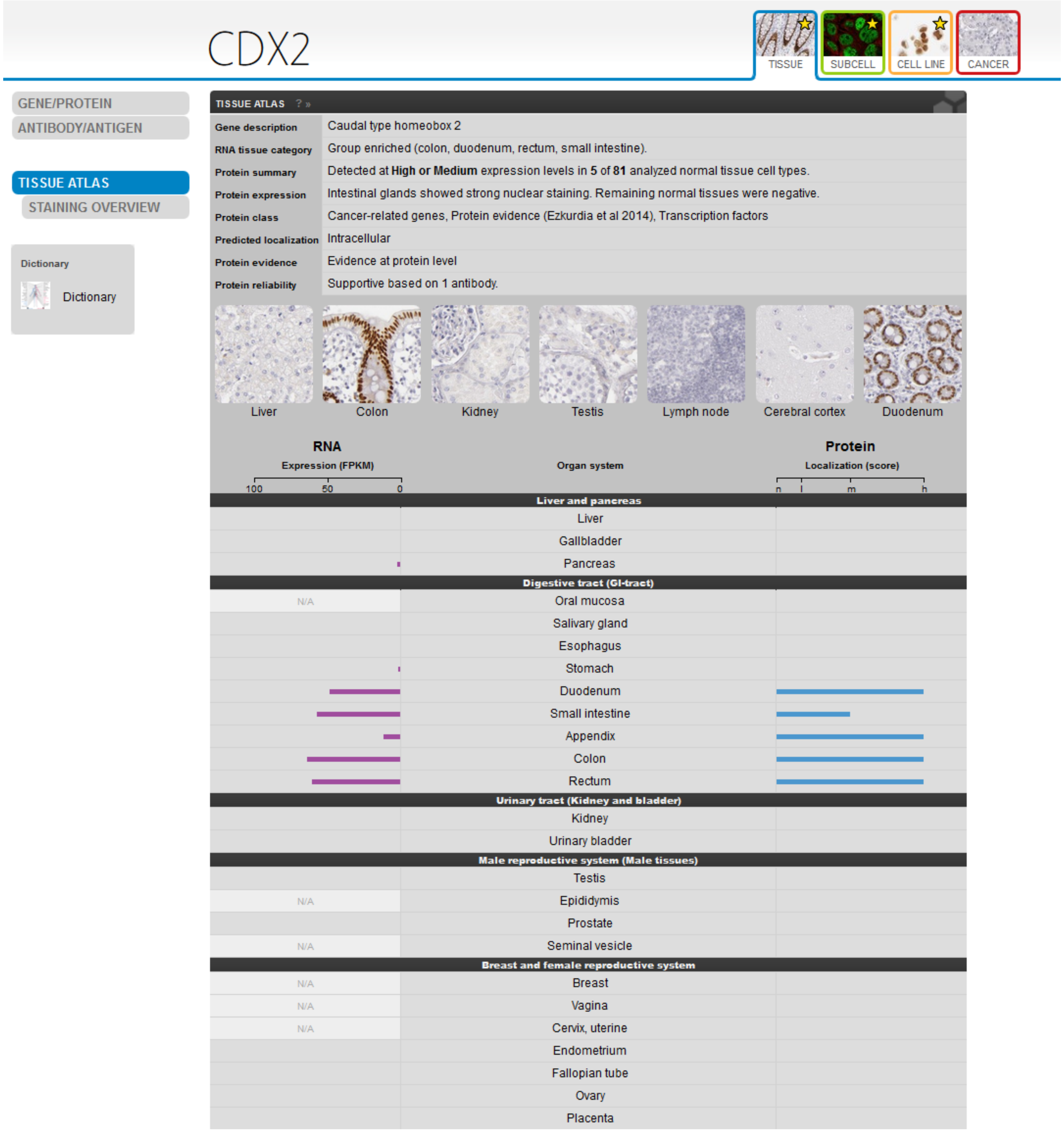
Limit search: Premium | Premium (Tissue) | Premium (Subcell) | Premium (Cell line)

Page 1 of 2 | next »

Show / hide columns »

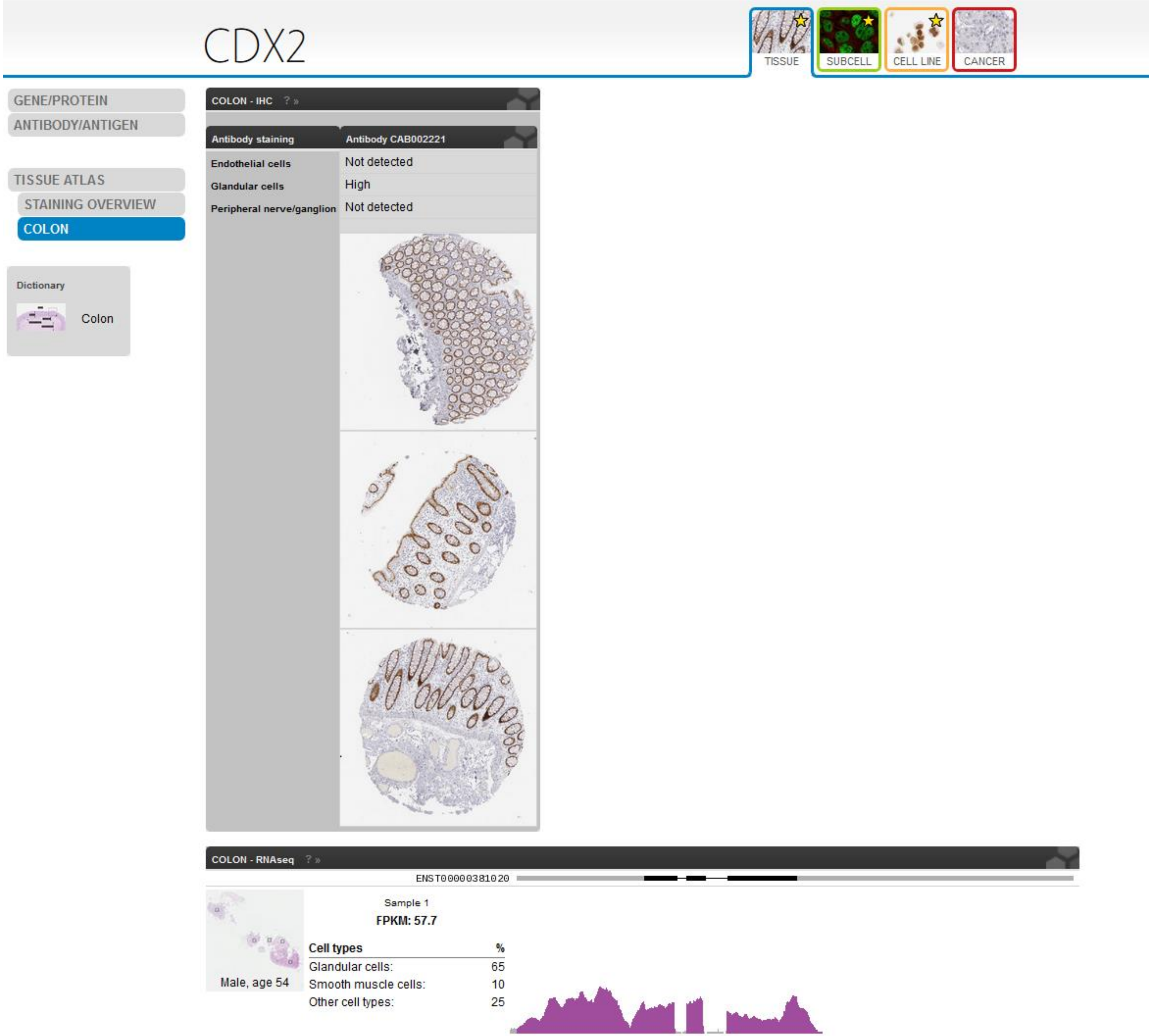
XML | RDF | TAB

Gene	Gene description	Protein class	Tissue	Subcell	Cell line	Cancer	IH abundance (Normal Tissue)	RNA abundance (Normal Tissue)	RNA tissue category	RNA TS FPKM
ADNP	Activity-dependent neuroprotector homeobox	Cancer-related genes Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) Transcription factors			RNA				Expressed in all	
ARID5B	AT rich interactive domain 5B (MRF1-like)	Cancer-related genes Disease related genes Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) Transcription factors			RNA				Expressed in all	
ATF1	Activating transcription factor 1	Cancer-related genes Disease related genes Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) Transcription factors			RNA				Expressed in all	
ATRX	Alpha thalassemia/mental retardation syndrome X-linked	Cancer-related genes Disease related genes Enzymes Potential drug targets Protein evidence (Ezkurdia et al 2014) Protein evidence (Kim et al 2014) Transcription factors							Expressed in all	
BCL3	B-cell CLL/lymphoma 3	Cancer-related genes Disease related genes Plasma proteins Protein evidence (Ezkurdia et al 2014) Transcription factors			RNA				Expressed in all	
CDX2	Caudal type homeobox 2	Cancer-related genes Protein evidence (Ezkurdia et al 2014) Transcription factors							Group enriched	colon: 63.0 duodenum: 48.3 rectum: 59.9 small intestine: 56.1



Gene description	Caudal type homeobox 2
RNA tissue category	Group enriched (colon, duodenum, rectum, small intestine).
Protein summary	Detected at **High or Medium** expression levels in **5** of **81** analyzed normal tissue cell types.
Protein expression	Intestinal glands showed strong nuclear staining. Remaining normal tissues were negative.
Protein class	Cancer-related genes, Protein evidence (Ezkurdia et al 2014), Transcription factors
Predicted localization	Intracellular
Protein evidence	Evidence at protein level
Protein reliability	Supportive based on 1 antibody.

RNA	Organ system	Protein
	Liver and pancreas	
	Liver	
	Gallbladder	
	Pancreas	
	Digestive tract (GI-tract)	
N/A	Oral mucosa	
	Salivary gland	
	Esophagus	
	Stomach	
	Duodenum	
	Small intestine	
	Appendix	
	Colon	
	Rectum	
	Urinary tract (Kidney and bladder)	
	Kidney	
	Urinary bladder	
	Male reproductive system (Male tissues)	
	Testis	
N/A	Epididymis	
	Prostate	
N/A	Seminal vesicle	
	Breast and female reproductive system	
N/A	Breast	
N/A	Vagina	
N/A	Cervix, uterine	
	Endometrium	
	Fallopian tube	
	Ovary	
	Placenta	



CDX2

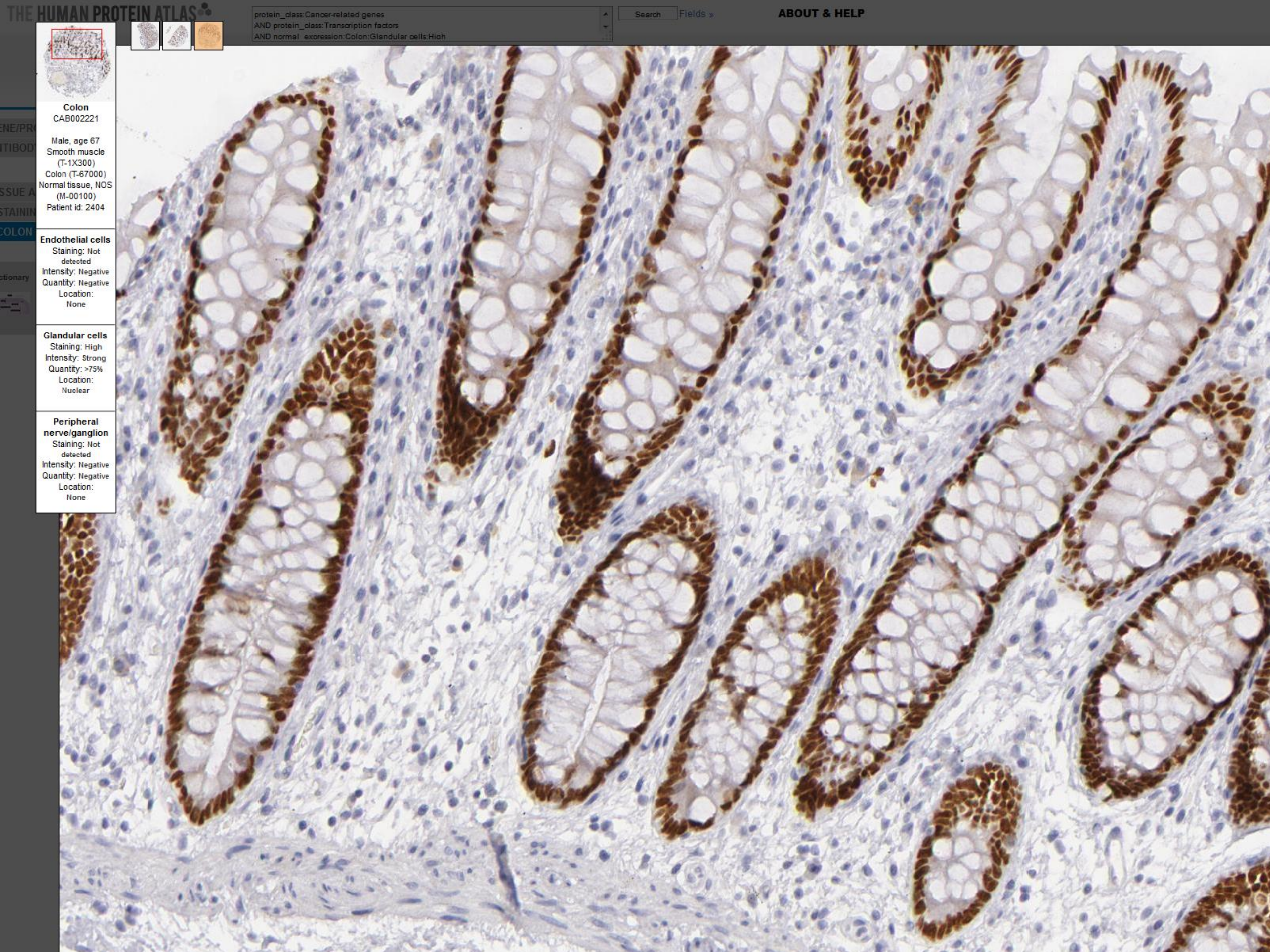
Antibody staining	Antibody CAB002221
Endothelial cells	Not detected
Glandular cells	High
Peripheral nerve/ganglion	Not detected

Male, age 54

Sample 1

FPKM: 57.7

Cell types	%
Glandular cells:	65
Smooth muscle cells:	10
Other cell types:	25



protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon:Glandular cells:High

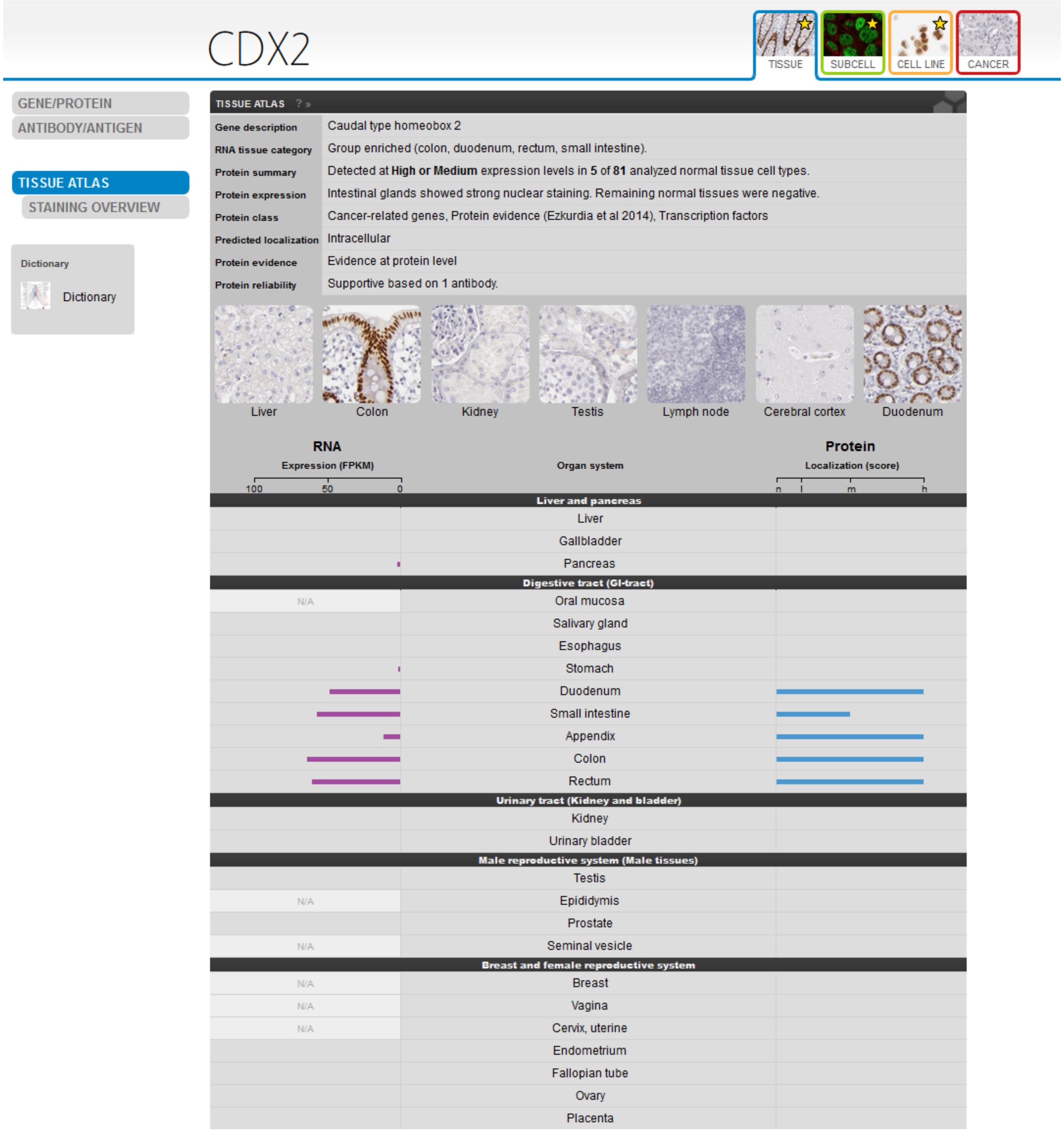
Colon
CAB002221

Male, age 67
Smooth muscle
(T-1X300)
Colon (T-67000)
Normal tissue, NOS
(M-00100)
Patient id: 2404

Endothelial cells
Staining: Not detected
Intensity: Negative
Quantity: Negative
Location: None

Glandular cells
Staining: High
Intensity: Strong
Quantity: >75%
Location: Nuclear

Peripheral nerve/ganglion
Staining: Not detected
Intensity: Negative
Quantity: Negative
Location: None



CDX2

Gene description	Caudal type homeobox 2
RNA tissue category	Group enriched (colon, duodenum, rectum, small intestine).
Protein summary	Detected at **High or Medium** expression levels in **5** of **81** analyzed normal tissue cell types.
Protein expression	Intestinal glands showed strong nuclear staining. Remaining normal tissues were negative.
Protein class	Cancer-related genes, Protein evidence (Ezkurdia et al 2014), Transcription factors
Predicted localization	Intracellular
Protein evidence	Evidence at protein level
Protein reliability	Supportive based on 1 antibody.

RNA Expression (FPKM)	Organ system	Protein Localization (score)
	Liver and pancreas	
	Liver	
	Gallbladder	
	Pancreas	
	Digestive tract (GI-tract)	
N/A	Oral mucosa	
	Salivary gland	
	Esophagus	
	Stomach	
	Duodenum	
	Small intestine	
	Appendix	
	Colon	
	Rectum	
	Urinary tract (Kidney and bladder)	
	Kidney	
	Urinary bladder	
	Male reproductive system (Male tissues)	
	Testis	
N/A	Epididymis	
	Prostate	
N/A	Seminal vesicle	
	Breast and female reproductive system	
N/A	Breast	
N/A	Vagina	
N/A	Cervix, uterine	
	Endometrium	
	Fallopian tube	
	Ovary	
	Placenta	

```
protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon;Glandular cells;High
```

CDX2

Summary	Localized to the nucleus but excluded from the nucleoli.
Main location	Nucleus but not nucleoli
Reliability	Supportive
Protein evidence	Evidence at protein level
Assay summary	Analysis based on two antibodies, HPA045669 and HPA049580, using immunoflourescence in human and mouse cells

Dictionary

Nucleus but not nucleoli

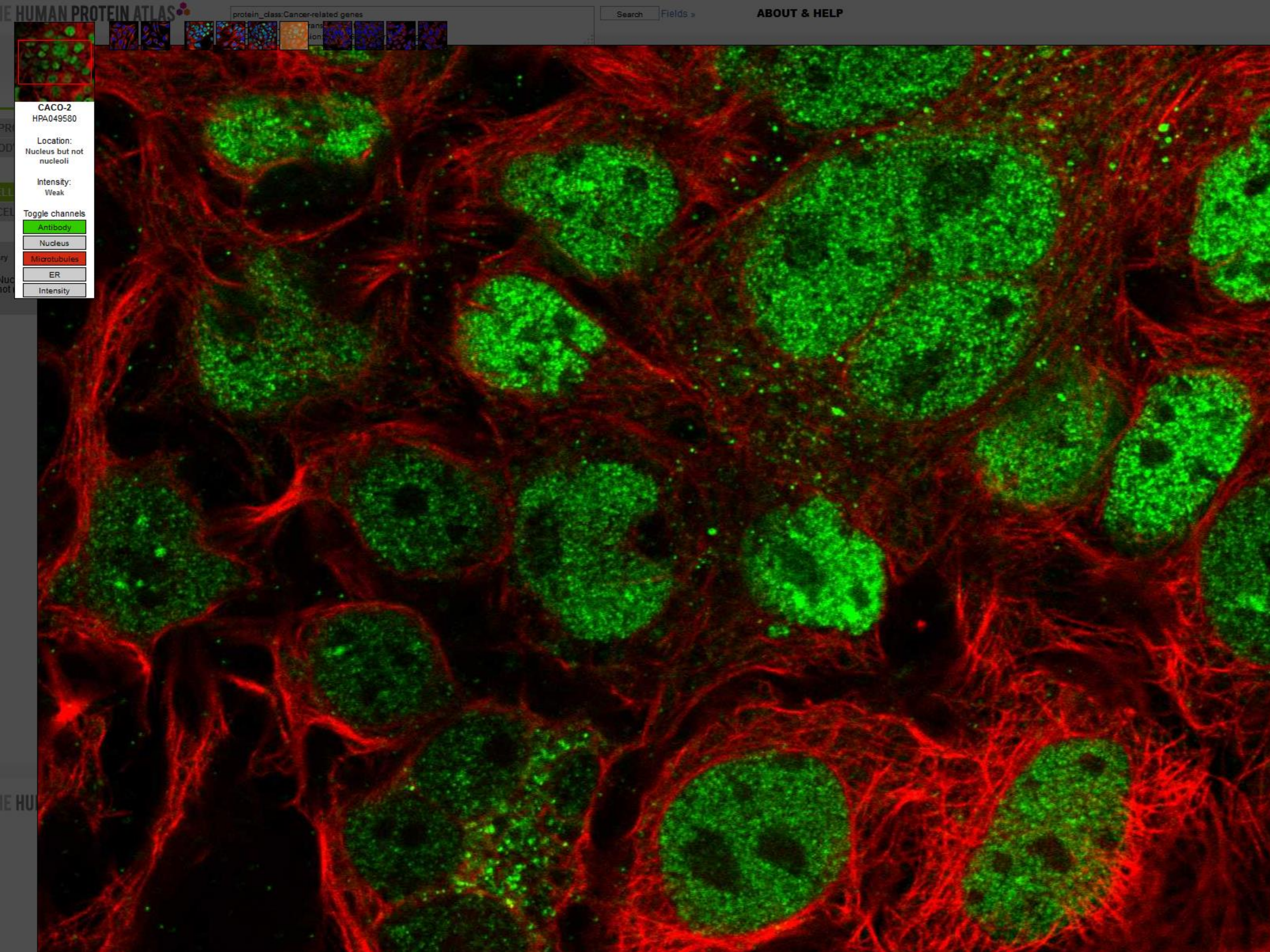
HPA049580: CACO-2

HPA049580: CACO-2

HPA049580: HEK 293

Toggle channels

HPA045669	Cell line: **U-2 OS** Location: **Not available** Literature validation: **Uncertain**	Cell line: **CACO-2** Location: **Nucleus but not nucleoli** Literature validation: **Supportive**	Cell line: **HEK 293** Location: **Not available** Literature validation: **Uncertain**
HPA049580	HPA049580 stained U-2 OS cells unspecifically	Cell line: **CACO-2** Location: **Nucleus but not nucleoli** Literature validation: **Supportive**	Cell line: **HEK 293** Location: **Not available** Literature validation: **Uncertain**

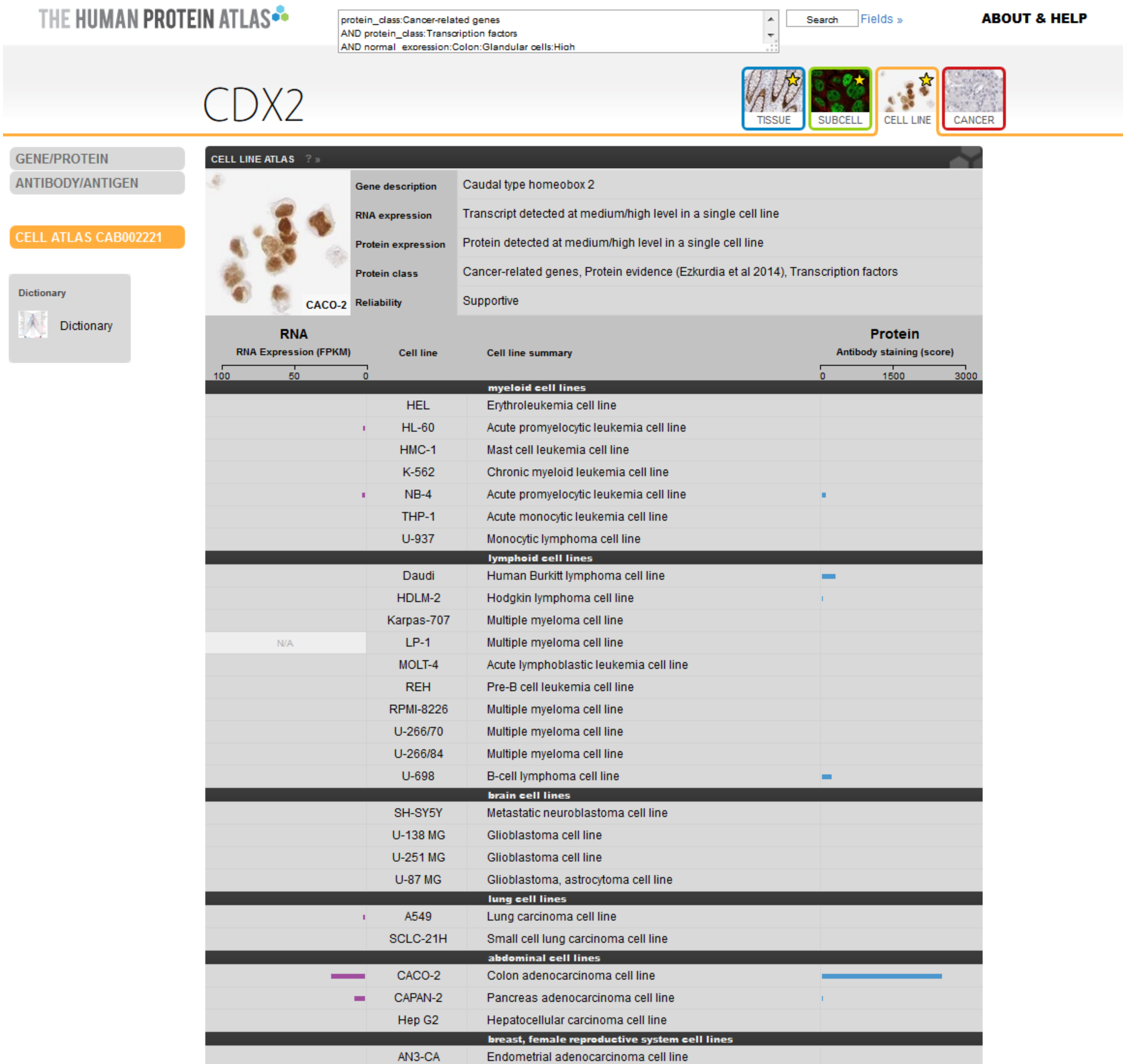


CACO-2
HPA049580

Location:
Nucleus but not nucleoli

Intensity:
Weak

protein_class:Cancer-related genes

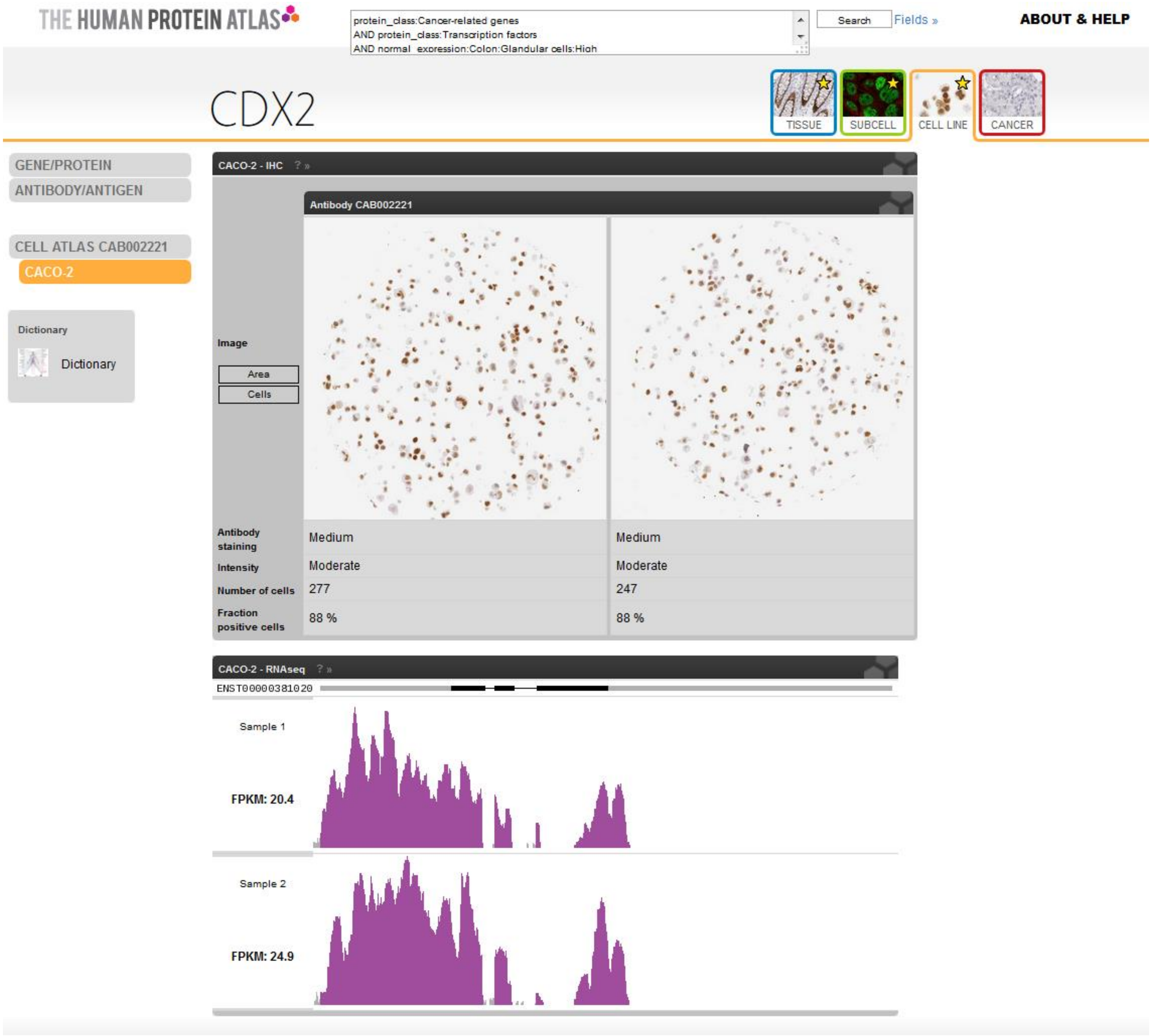


protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon:Glandular cells:High

CDX2

Gene description	Caudal type homeobox 2
RNA expression	Transcript detected at medium/high level in a single cell line
Protein expression	Protein detected at medium/high level in a single cell line
Protein class	Cancer-related genes, Protein evidence (Ezkurdia et al 2014), Transcription factors
Reliability	Supportive

RNA RNA Expression (FPKM)	Cell line	Cell line summary	**Protein** Antibody staining (score)
		myeloid cell lines	
	HEL	Erythroleukemia cell line	
	HL-60	Acute promyelocytic leukemia cell line	
	HMC-1	Mast cell leukemia cell line	
	K-562	Chronic myeloid leukemia cell line	
	NB-4	Acute promyelocytic leukemia cell line	
	THP-1	Acute monocytic leukemia cell line	
	U-937	Monocytic lymphoma cell line	
		lymphoid cell lines	
	Daudi	Human Burkitt lymphoma cell line	
	HDLM-2	Hodgkin lymphoma cell line	
	Karpas-707	Multiple myeloma cell line	
N/A	LP-1	Multiple myeloma cell line	
	MOLT-4	Acute lymphoblastic leukemia cell line	
	REH	Pre-B cell leukemia cell line	
	RPMI-8226	Multiple myeloma cell line	
	U-266/70	Multiple myeloma cell line	
	U-266/84	Multiple myeloma cell line	
	U-698	B-cell lymphoma cell line	
		brain cell lines	
	SH-SY5Y	Metastatic neuroblastoma cell line	
	U-138 MG	Glioblastoma cell line	
	U-251 MG	Glioblastoma cell line	
	U-87 MG	Glioblastoma, astrocytoma cell line	
		lung cell lines	
	A549	Lung carcinoma cell line	
	SCLC-21H	Small cell lung carcinoma cell line	
		abdominal cell lines	
	CACO-2	Colon adenocarcinoma cell line	
	CAPAN-2	Pancreas adenocarcinoma cell line	
	Hep G2	Hepatocellular carcinoma cell line	
		breast, female reproductive system cell lines	
	AN3-CA	Endometrial adenocarcinoma cell line	



protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon:Glandular cells:High

CDX2

CELL ATLAS CAB002221

CACO-2

CACO-2 - IHC

Antibody CAB002221

Antibody staining	Medium	Medium
Intensity	Moderate	Moderate
Number of cells	277	247
Fraction positive cells	88 %	88 %

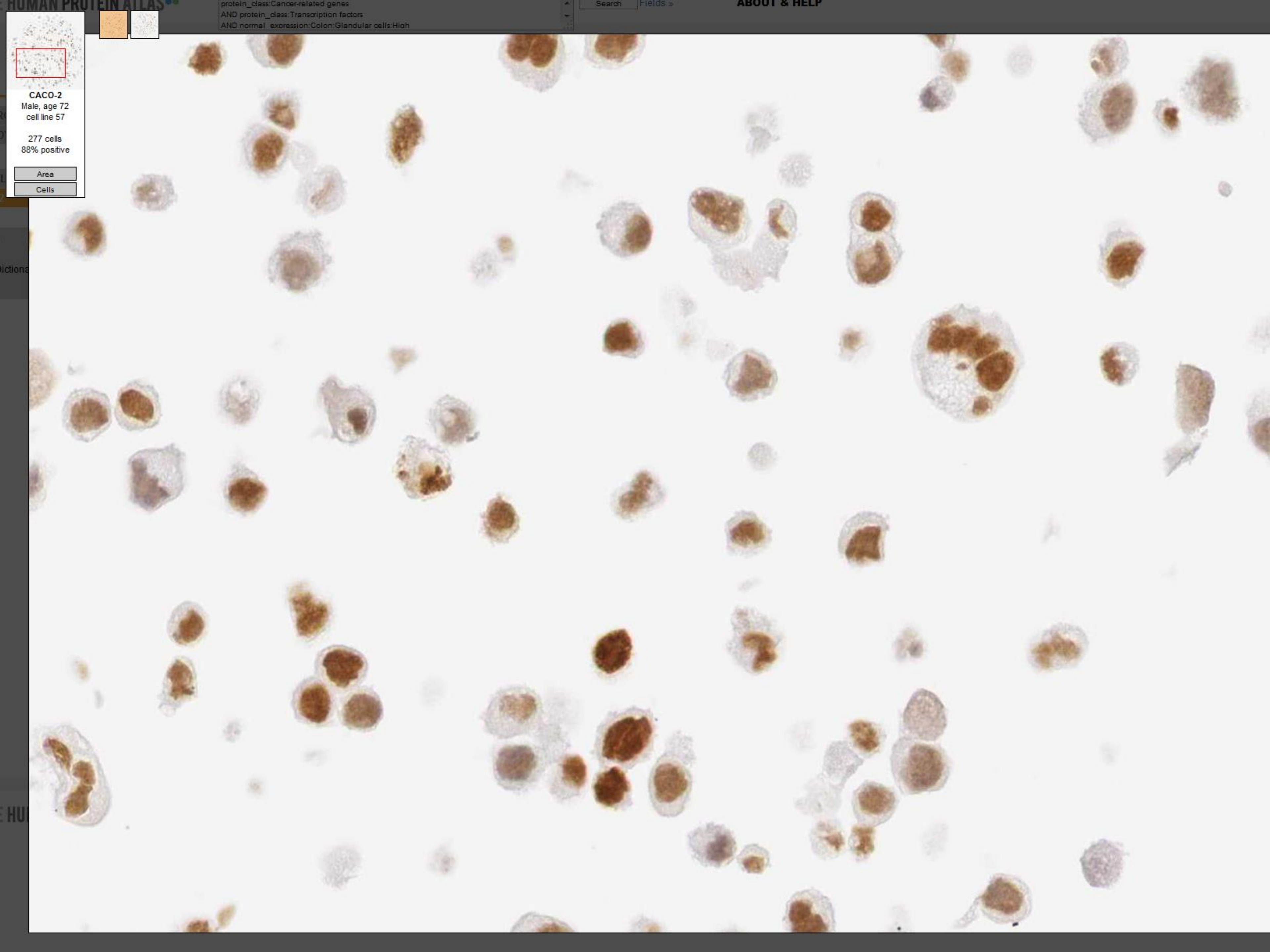
CACO-2 - RNAseq

Sample 1

FPKM: 20.4

Sample 2

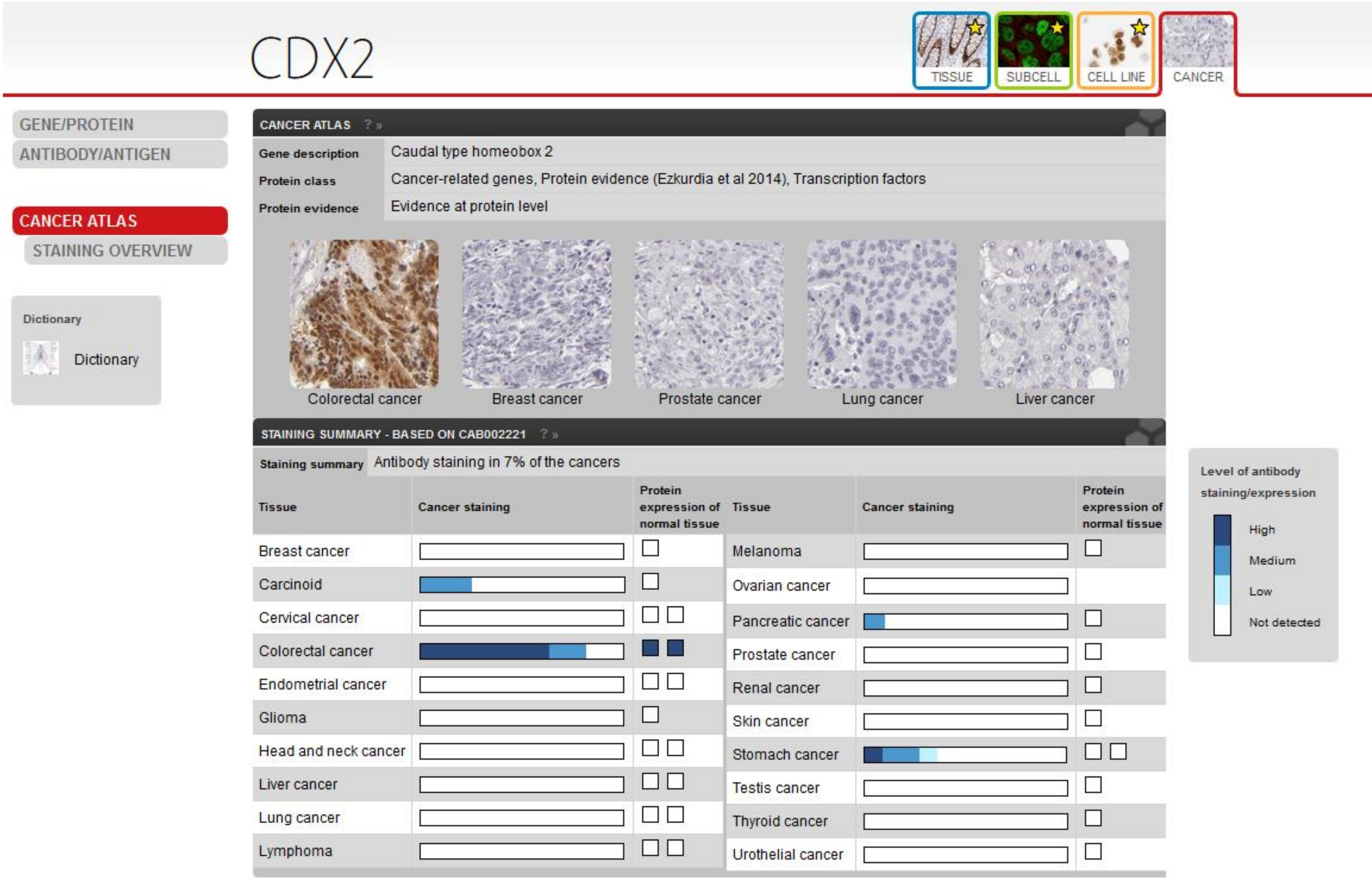
FPKM: 24.9



protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon:Glandular cells:High

CACO-2
Male, age 72
cell line 57

277 cells
88% positive



CDX2

CANCER ATLAS

Gene description	Caudal type homeobox 2
Protein class	Cancer-related genes, Protein evidence (Ezkurdia et al 2014), Transcription factors
Protein evidence	Evidence at protein level

Colorectal cancer Breast cancer Prostate cancer Lung cancer Liver cancer

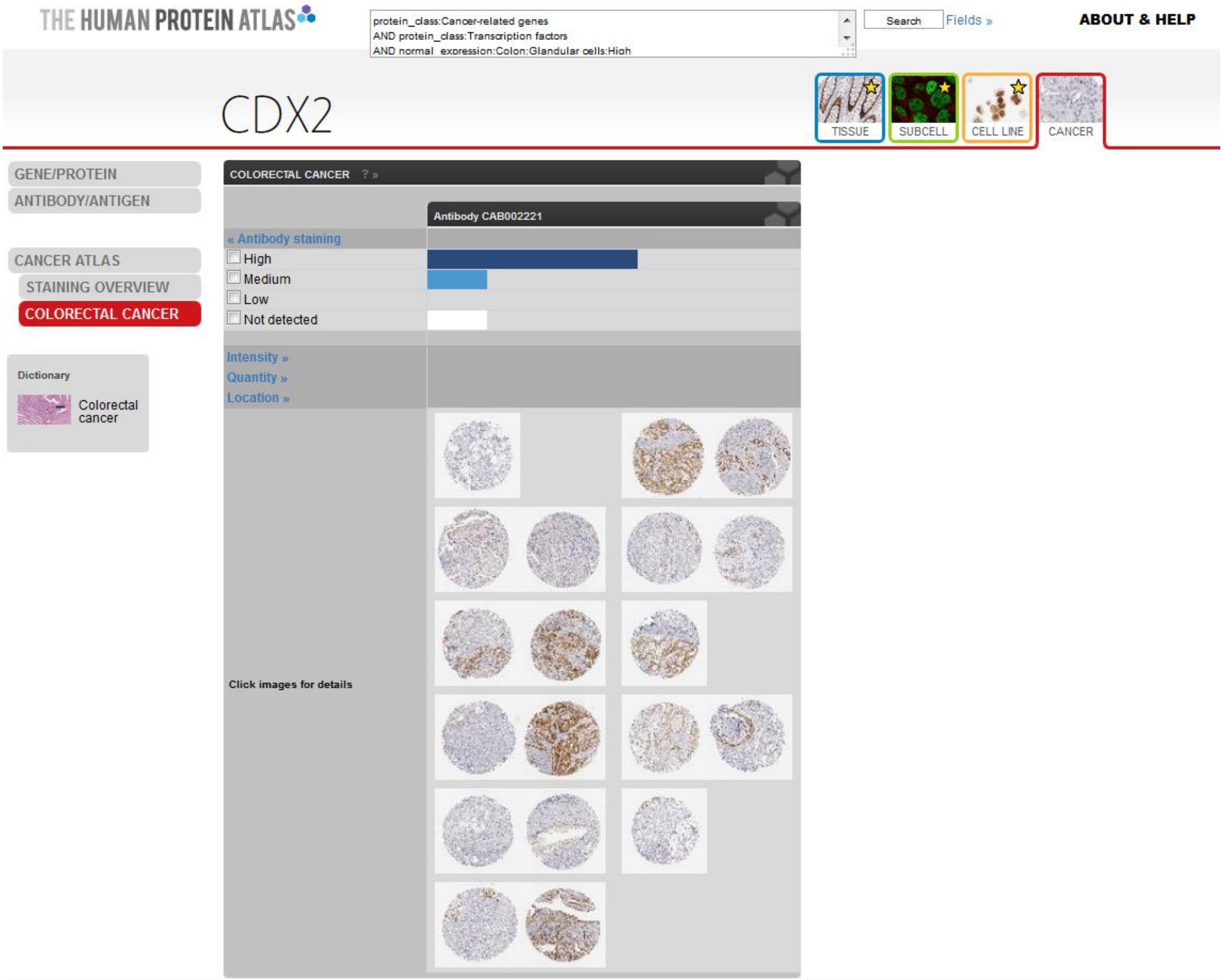
STAINING SUMMARY - BASED ON CAB002221

Staining summary Antibody staining in 7% of the cancers

Tissue	Cancer staining	Protein expression of normal tissue	Tissue	Cancer staining	Protein expression of normal tissue
Breast cancer		☐	Melanoma		☐
Carcinoid		☐	Ovarian cancer		
Cervical cancer		☐ ☐	Pancreatic cancer		☐
Colorectal cancer		■ ■	Prostate cancer		☐
Endometrial cancer		☐ ☐	Renal cancer		☐
Glioma		☐	Skin cancer		☐
Head and neck cancer		☐ ☐	Stomach cancer		☐ ☐
Liver cancer		☐ ☐	Testis cancer		☐
Lung cancer		☐ ☐	Thyroid cancer		☐
Lymphoma		☐ ☐	Urothelial cancer		☐

Level of antibody staining/expression

- High
- Medium
- Low
- Not detected



THE HUMAN PROTEIN ATLAS

protein_class:Cancer-related genes
AND protein_class:Transcription factors
AND normal_expression:Colon:Glandular cells:High

CDX2

COLORECTAL CANCER ? »

Antibody CAB002221

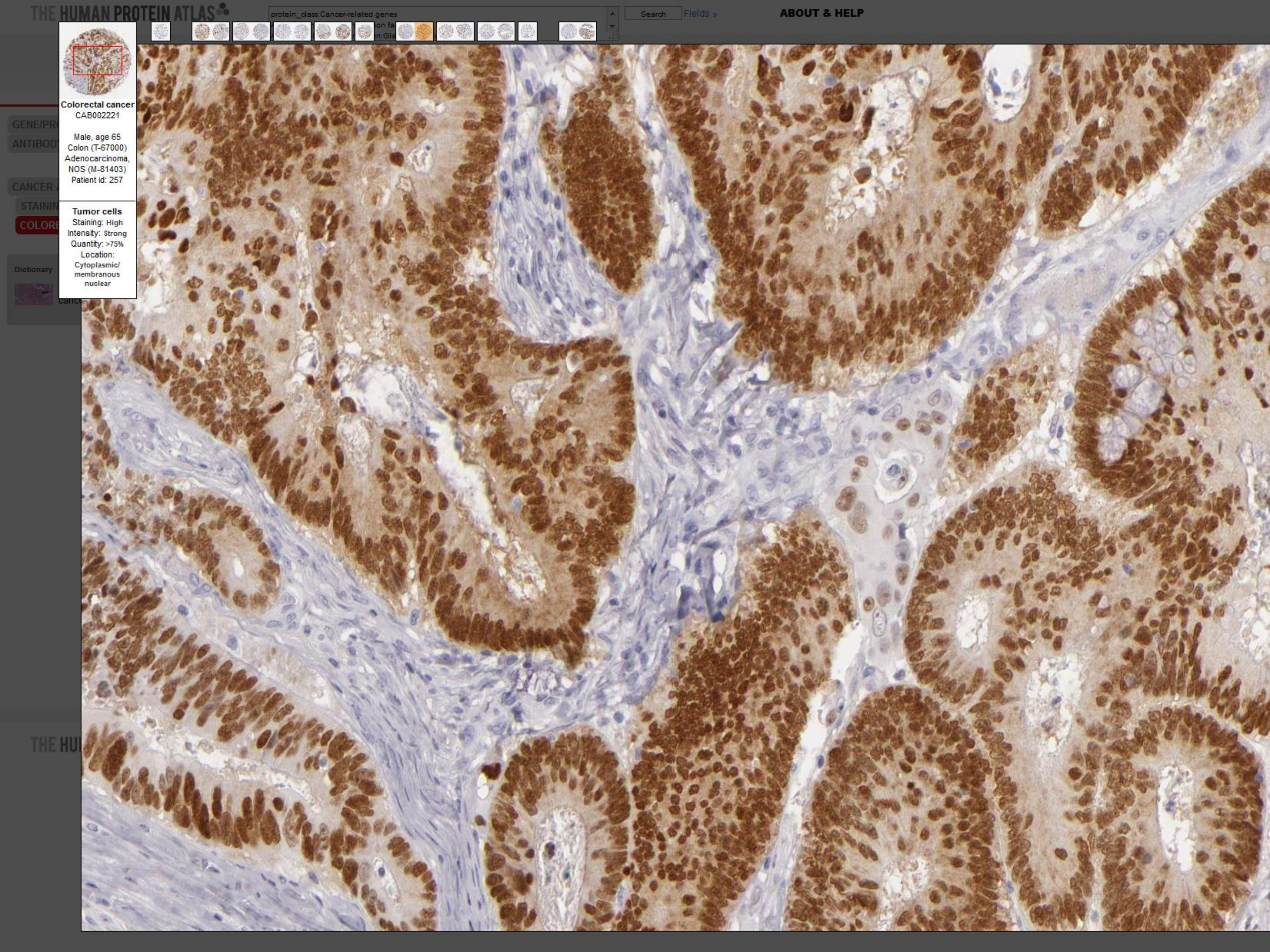
« Antibody staining

- ☐ High
- ☐ Medium
- ☐ Low
- ☐ Not detected

Click images for details

Dictionary

Colorectal cancer

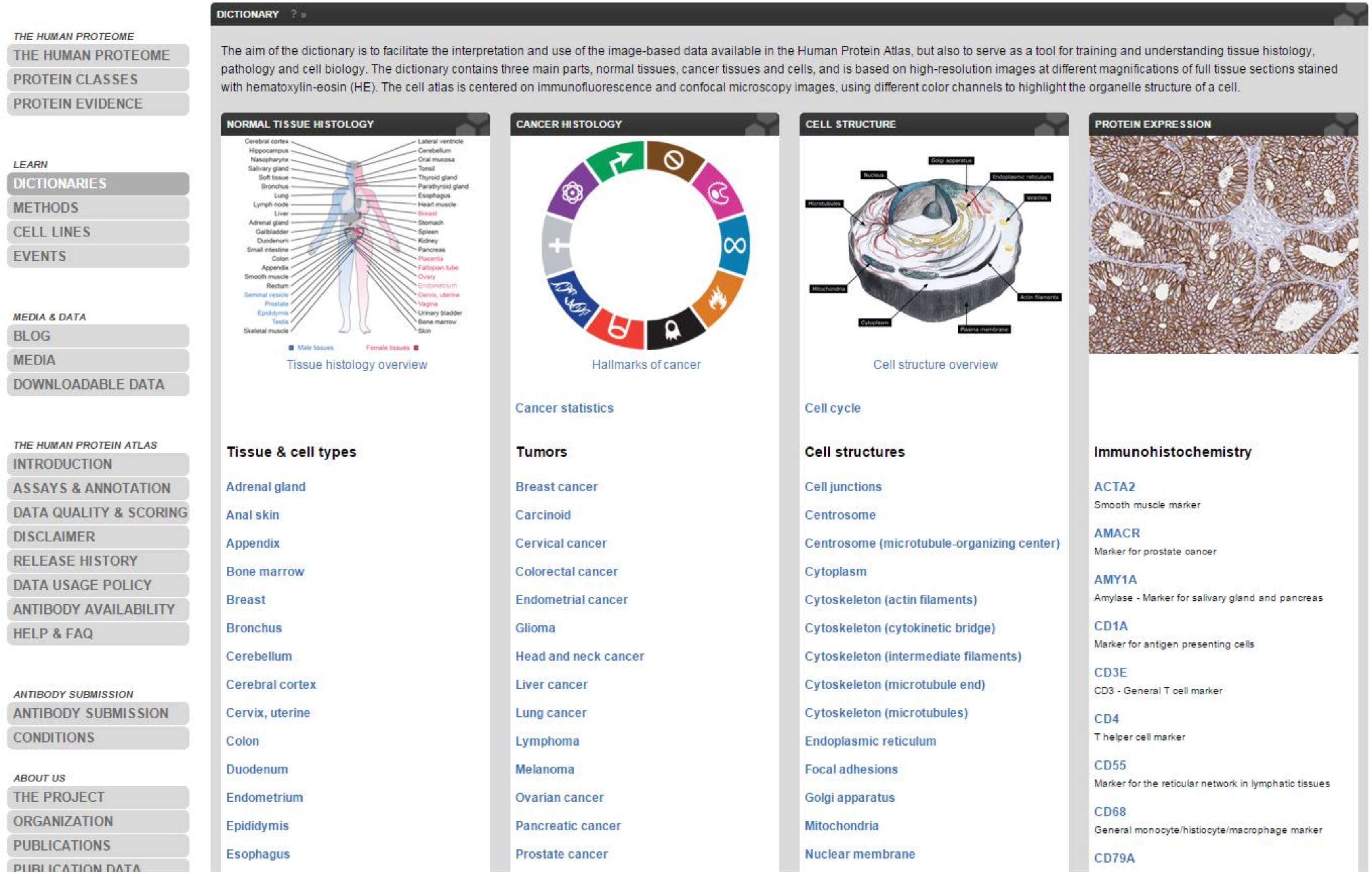


Colorectal cancer
CAB002221

Male, age 65
Colon (T-67000)
Adenocarcinoma, NOS (M-81403)
Patient id: 257

Tumor cells
Staining: High
Intensity: Strong
Quantity: >75%
Location: Cytoplasmic/ membranous nuclear

Dictionary



The aim of the dictionary is to facilitate the interpretation and use of the image-based data available in the Human Protein Atlas, but also to serve as a tool for training and understanding tissue histology, pathology and cell biology. The dictionary contains three main parts, normal tissues, cancer tissues and cells, and is based on high-resolution images at different magnifications of full tissue sections stained with hematoxylin-eosin (HE). The cell atlas is centered on immunofluorescence and confocal microscopy images, using different color channels to highlight the organelle structure of a cell.

Tissue histology overview

Tissue & cell types

- Adrenal gland
- Anal skin
- Appendix
- Bone marrow
- Breast
- Bronchus
- Cerebellum
- Cerebral cortex
- Cervix, uterine
- Colon
- Duodenum
- Endometrium
- Epididymis
- Esophagus

Hallmarks of cancer

Cancer statistics

Tumors

- Breast cancer
- Carcinoid
- Cervical cancer
- Colorectal cancer
- Endometrial cancer
- Glioma
- Head and neck cancer
- Liver cancer
- Lung cancer
- Lymphoma
- Melanoma
- Ovarian cancer
- Pancreatic cancer
- Prostate cancer

Cell structure overview

Cell cycle

Cell structures

- Cell junctions
- Centrosome
- Centrosome (microtubule-organizing center)
- Cytoplasm
- Cytoskeleton (actin filaments)
- Cytoskeleton (cytokinetic bridge)
- Cytoskeleton (intermediate filaments)
- Cytoskeleton (microtubule end)
- Cytoskeleton (microtubules)
- Endoplasmic reticulum
- Focal adhesions
- Golgi apparatus
- Mitochondria
- Nuclear membrane

Immunohistochemistry

ACTA2	Smooth muscle marker
AMACR	Marker for prostate cancer
AMY1A	Amylase - Marker for salivary gland and pancreas
CD1A	Marker for antigen presenting cells
CD3E	CD3 - General T cell marker
CD4	T helper cell marker
CD55	Marker for the reticular network in lymphatic tissues
CD68	General monocyte/histiocyte/macrophage marker
CD79A	

THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

KTH VETENSKAP OCH KONST

UPPSALA UNIVERSITET

THE HUMAN PROTEOME
THE HUMAN PROTEOME
PROTEIN CLASSES
PROTEIN EVIDENCE

LEARN
DICTIONARIES
METHODS
CELL LINES

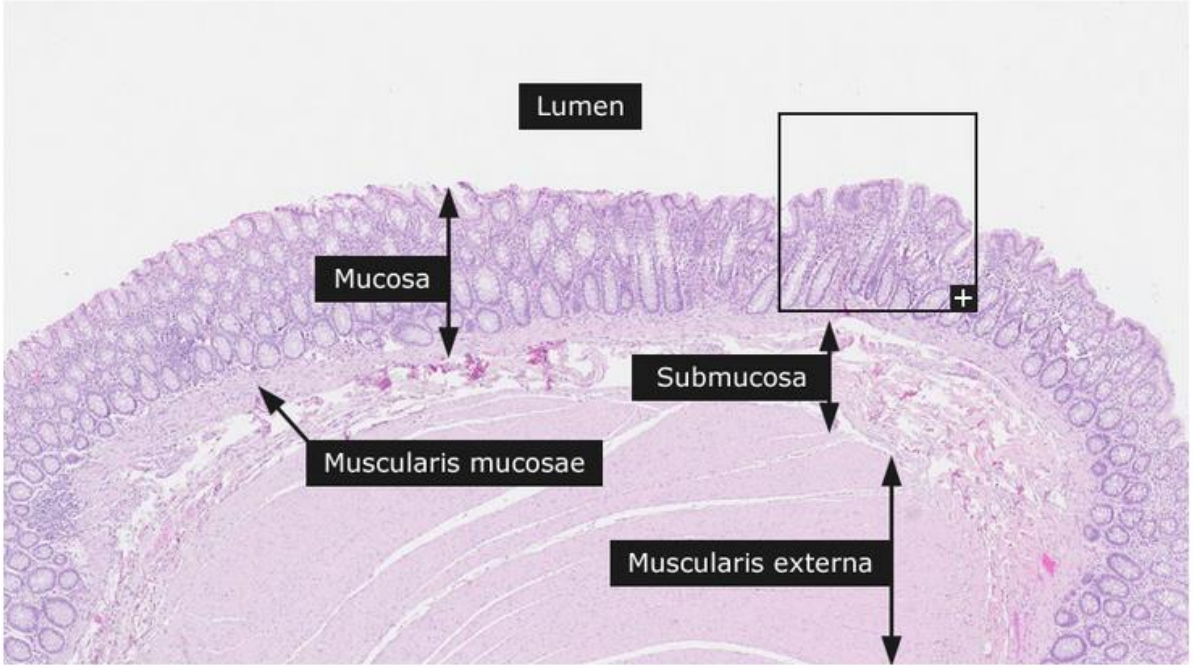
MEDIA & DATA
MEDIA
DOWNLOADABLE DATA

THE HUMAN PROTEIN ATLAS
INTRODUCTION
ASSAYS & ANNOTATION
DATA QUALITY & SCORING
DISCLAIMER
RELEASE HISTORY
DATA USAGE POLICY
ANTIBODY AVAILABILITY
HELP

ANTIBODY SUBMISSION
ANTIBODY SUBMISSION
CONDITIONS

ABOUT US
THE PROJECT
ORGANIZATION
PUBLICATIONS
LINKS
CONTACT

Dictionary > Normal tissue > **Colon**



Colon

The colon is divided into four parts, the ascending, transverse, descending and sigmoid colon and is on average 1,5 meters long. Its main function is reassertion of fluid, electrolytes, and vitamins.

Since the large intestine has no villi or plica circularis the ***mucosa*** is smooth. Simple tubular ***intestinal glands*** (crypts of Lieberkuhn) extend through the entire thickness of the mucosa. The surface ***columnar epithelium*** and the cells lining the crypts are ***enterocytes***, with an oval basal nucleus and apical ***brush border***, the microscopic representation of microvilli. There are also numerous mucous secreting ***goblet cells*** recognized by their content of a large mucous globule. The ***lamina propria*** with connective tissue and inflammatory cells surround the crypts. A thin smooth muscular layer, the ***lamina muscularis mucosae*** marks the border between the mucosa and submucosa.

The ***submucosa*** consists of loose connective tissue with vessels and nerves. Some solitary ***lymph follicles*** are also seen.

The muscular layer (***muscularis externa***) consists of an inner circular smooth muscle layer, the outer longitudinal muscle layer is not continuous as in the rest of the gastrointestinal tract. It is divided into three thickened muscular bands, called teniae coli.

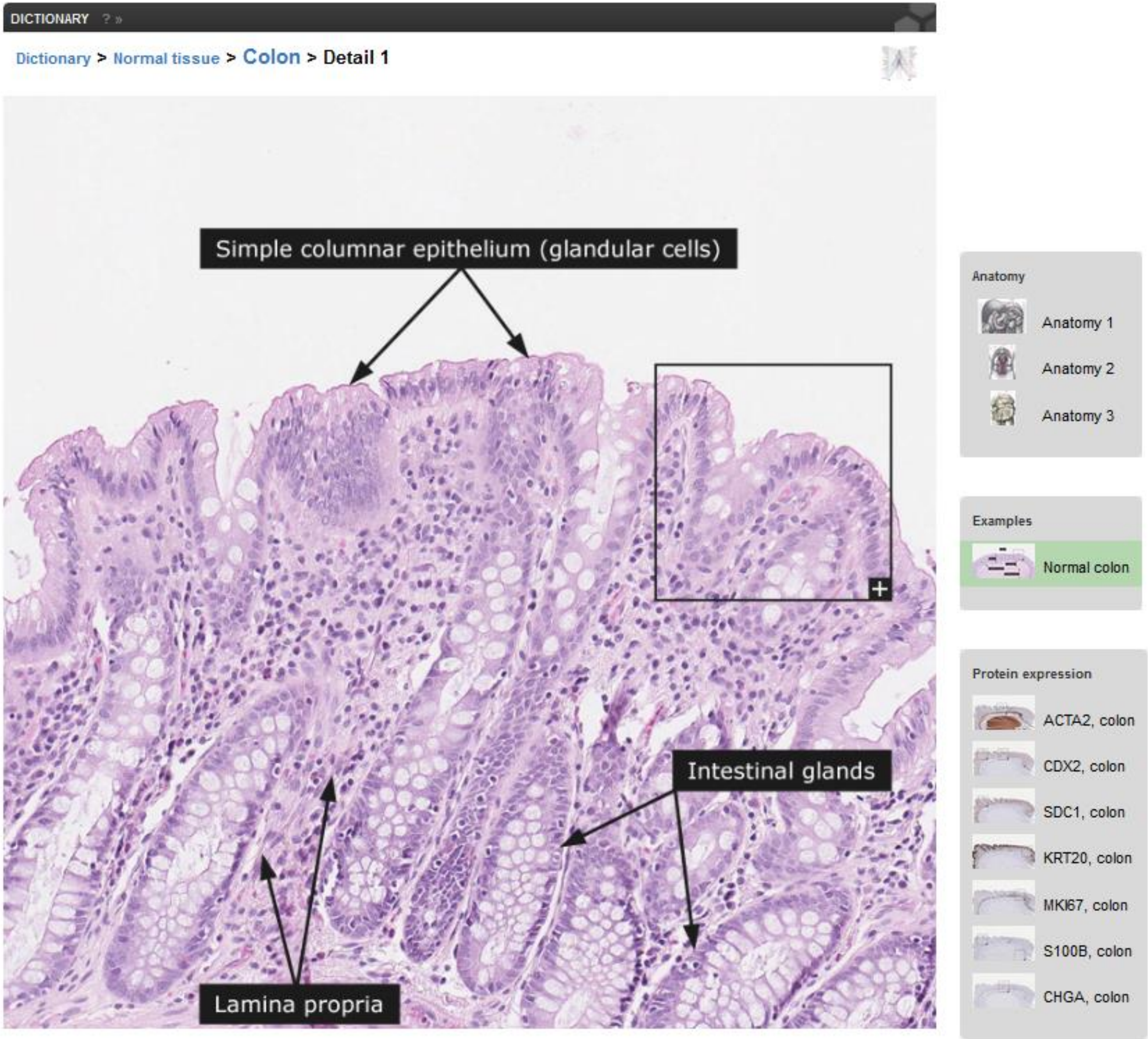
General histology of gastro-intestinal tract (GI-tract)

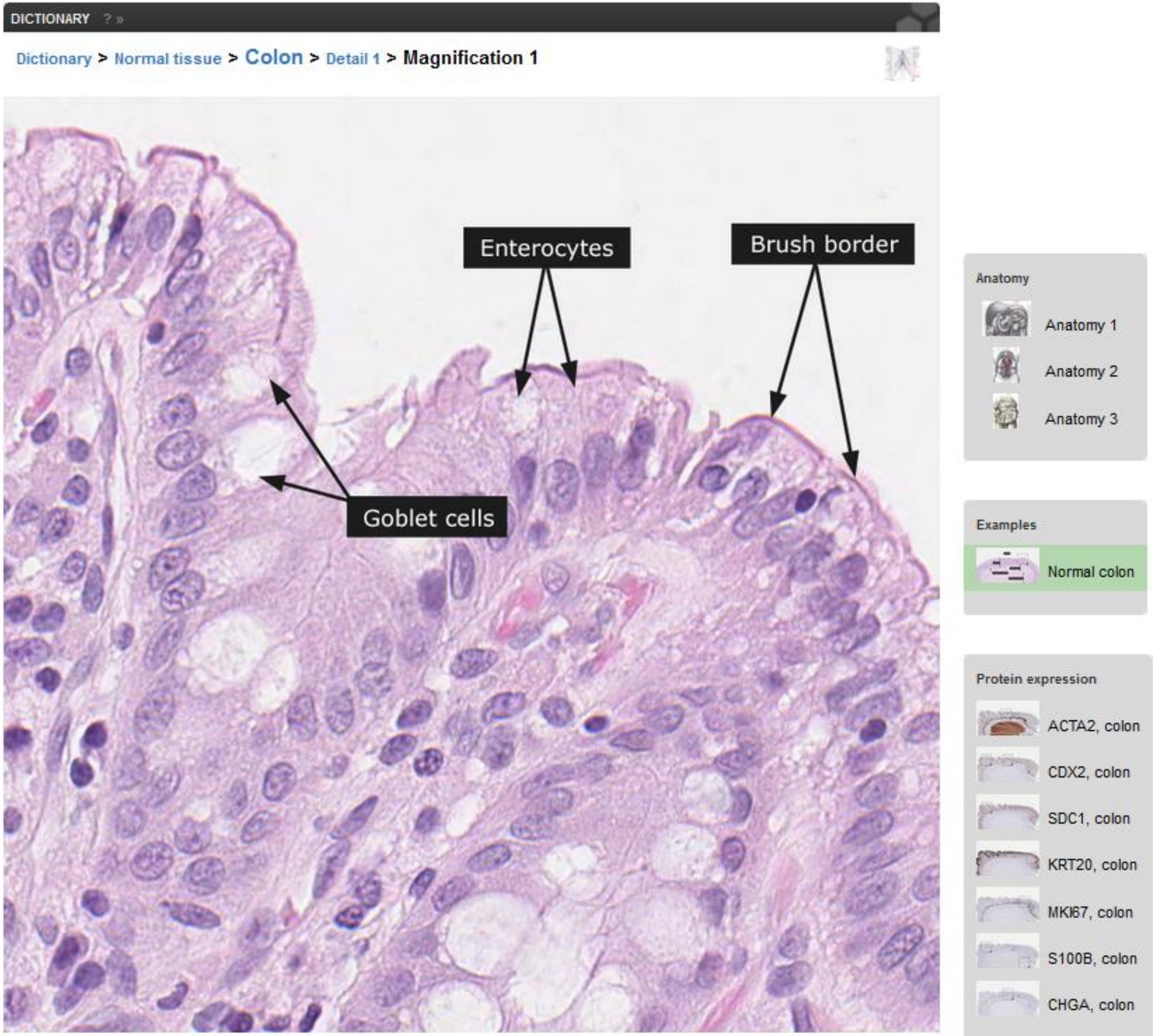
The gastrointestinal canal consists of the ***esophagus***, ***stomach***, ***duodenum***, ***jejunum***, ***ileum***, ***colon***, ***rectum*** and ***anal canal***. It is best viewed as a long tube passing from the oral to the anal opening. It supplies the body with water, electrolytes and nutrients from the food we eat. Our main sources of food are carbohydrates, proteins and fats, which in general cannot be absorbed in the form they are ingested. First they have to be broken down into small enough compounds. The process of digestion and absorption is carried out in a stepwise fashion as the food passes down the different parts of the gastrointestinal tract.

Anatomy
Anatomy 1
Anatomy 2
Anatomy 3

Examples
Normal colon

Protein expression
ACTA2, colon
CDX2, colon
SDC1, colon
KRT20, colon
MKI67, colon
S100B, colon
CHGA, colon


DICTIONARY ? »
Dictionary > Normal tissue > Colon > Detail 1
Simple columnar epithelium (glandular cells)
Intestinal glands
Lamina propria
Anatomy
Anatomy 1
Anatomy 2
Anatomy 3
Examples
Normal colon
Protein expression
ACTA2, colon
CDX2, colon
SDC1, colon
KRT20, colon
MKI67, colon
S100B, colon
CHGA, colon



Dictionary > Normal tissue > Colon > Detail 1 > Magnification 1

Anatomy

- Anatomy 1
- Anatomy 2
- Anatomy 3

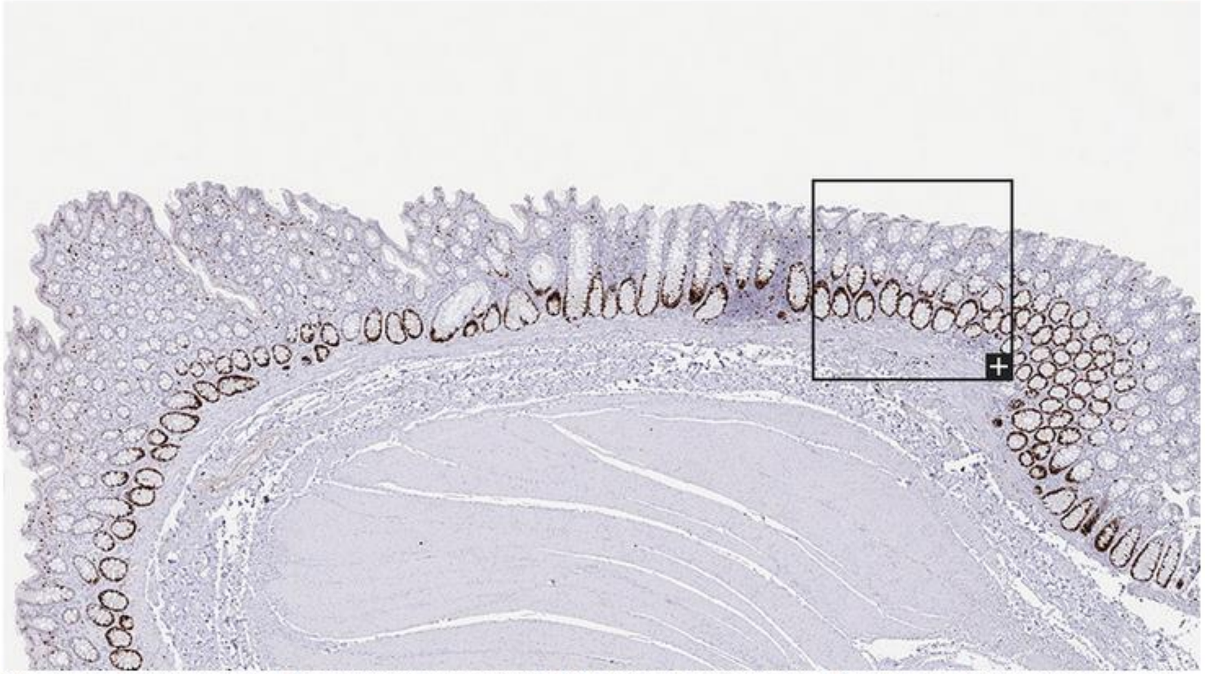
Examples

- Normal colon

Protein expression

- ACTA2, colon
- CDX2, colon
- SDC1, colon
- KRT20, colon
- MKI67, colon
- S100B, colon
- CHGA, colon

Dictionary > Protein expression > **MKI67**



Normal colon showing expression of the proliferation marker Ki-67 (MKI67) in the nuclei of proliferating glandular cells. Ki-67 expression is mainly found in deeper parts of colonic crypts. Ki-67 expression is additionally seen in a few inflammatory cells present in lamina propria.

Ki-67 (MKI67) - general marker for proliferation

The MKI67 gene encodes a widely known yet functionally poorly characterized protein called Ki-67 that specifically labels the nuclei and chromosomes of cells actively undergoing proliferation, but is not detected in cells that are in resting G0-phase. Thus, Ki-67 is a very general marker for actively proliferating cells and this characteristic has been heavily exploited in both research and clinical work. It is commonly used to assess the proliferative activity in tumors, which may be a measure of aggressiveness in certain types of cancer.

Show MKI67 in the protein atlas.

Normal tissue - Protein expression

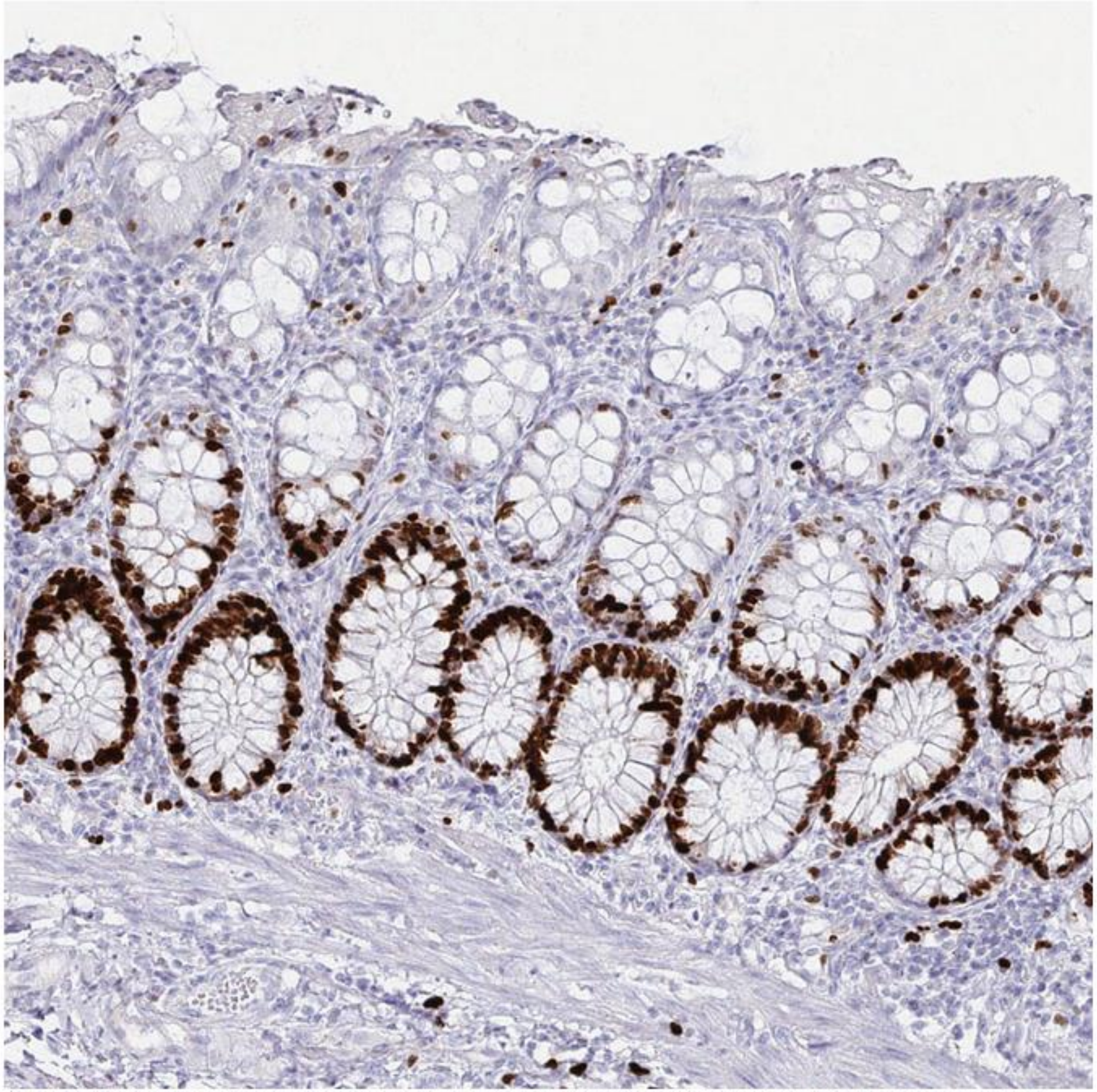
- Colon
- Lymph node
- Skin
- Testis
- Tonsil
- Urinary bladder

Cancer tissue - Protein expression

- Breast cancer 1, ductal carcinoma
- Breast cancer 2, ductal carcinoma
- Carcinoid 1
- Carcinoid 2
- Colorectal cancer 1, adenocarcinoma
- Colorectal cancer 2, adenocarcinoma
- Colorectal cancer 3, adenocarcinoma
- Lung cancer 1, adenocarcinoma
- Lung cancer 2, squamous cell carcinoma
- Lung cancer 3, small cell lung carcinoma
- Lung cancer 4, adenocarcinoma
- Lymphoma 1, B-CLL

DICTIONARY ? »

Dictionary > Protein expression > MKI67 > Detail 1



High magnification of normal colon mucosa showing abundant expression of the proliferation marker Ki-67 in the deeper regions of colonic crypts, where proliferating cells are present. Scarce inflammatory cells in the upper regions of the colon mucosa are also seen expressing Ki-67.

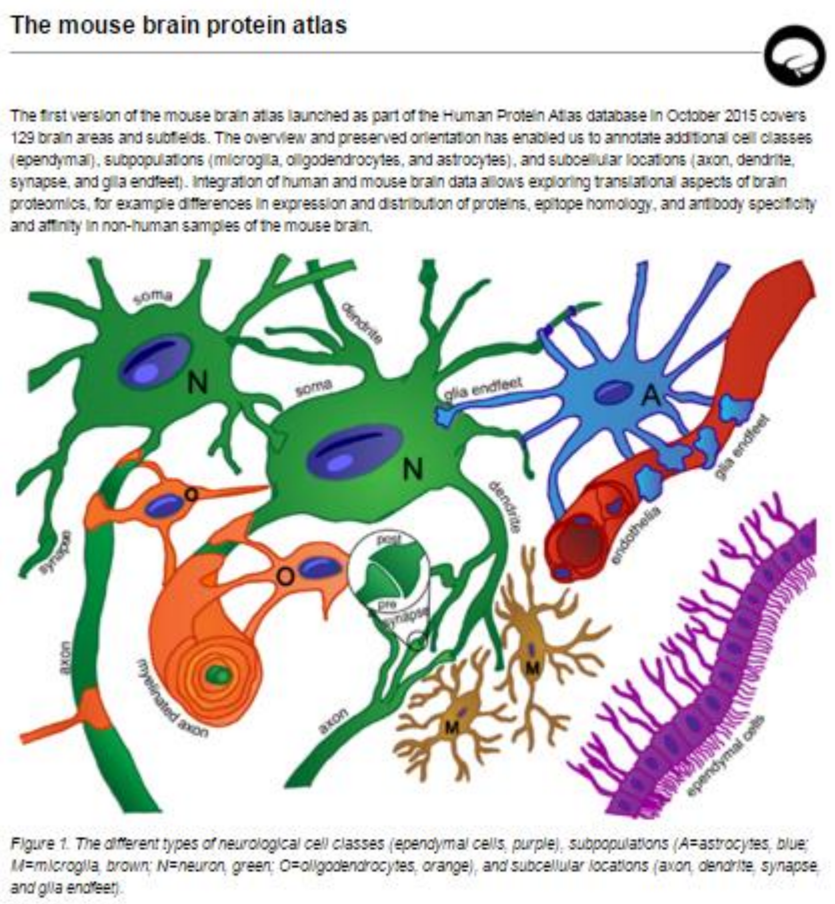
Normal tissue - Protein expression

- Colon
- Lymph node
- Skin
- Testis
- Tonsil
- Urinary bladder

Cancer tissue - Protein expression

- Breast cancer 1, ductal carcinoma
- Breast cancer 2, ductal carcinoma
- Carcinoid 1
- Carcinoid 2
- Colorectal cancer 1, adenocarcinoma
- Colorectal cancer 2, adenocarcinoma
- Colorectal cancer 3, adenocarcinoma
- Lung cancer 1, adenocarcinoma
- Lung cancer 2, squamous cell carcinoma
- Lung cancer 3, small cell lung carcinoma

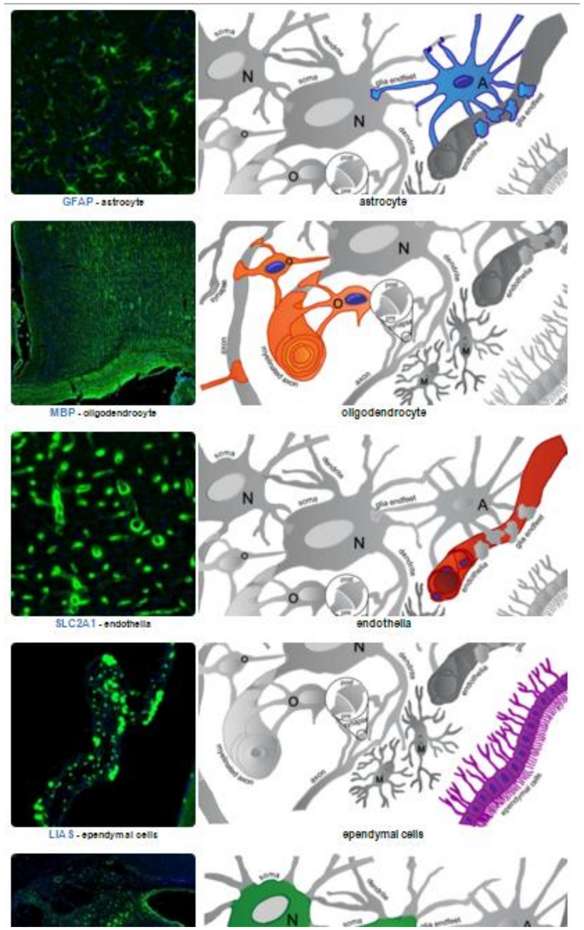
The Mouse Brain Atlas (new v.14)

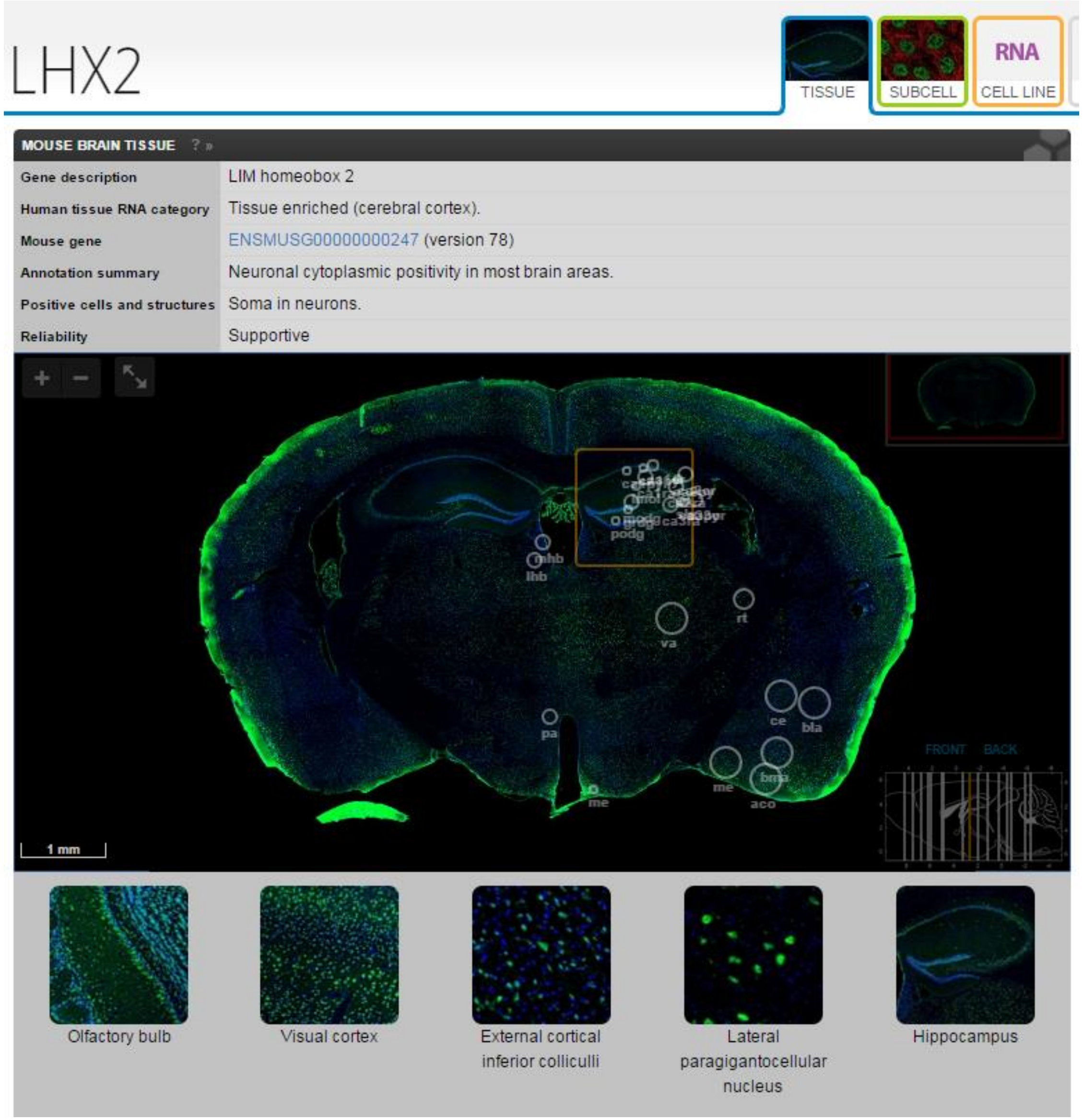


The mouse brain protein atlas

The first version of the mouse brain atlas launched as part of the Human Protein Atlas database in October 2015 covers 129 brain areas and subfields. The overview and preserved orientation has enabled us to annotate additional cell classes (ependymal), subpopulations (microglia, oligodendrocytes, and astrocytes), and subcellular locations (axon, dendrite, synapse, and glia endfeet). Integration of human and mouse brain data allows exploring translational aspects of brain proteomics, for example differences in expression and distribution of proteins, epitope homology, and antibody specificity and affinity in non-human samples of the mouse brain.

Figure 1. The different types of neurological cell classes (ependymal cells, purple), subpopulations (A=astrocytes, blue; M=microglia, brown; N=neuron, green; O=oligodendrocytes, orange), and subcellular locations (axon, dendrite, synapse, and glia endfeet).





LHX2

MOUSE BRAIN TISSUE

Gene description	LIM homeobox 2
Human tissue RNA category	Tissue enriched (cerebral cortex).
Mouse gene	ENSMUSG00000000247 (version 78)
Annotation summary	Neuronal cytoplasmic positivity in most brain areas.
Positive cells and structures	Soma in neurons.
Reliability	Supportive

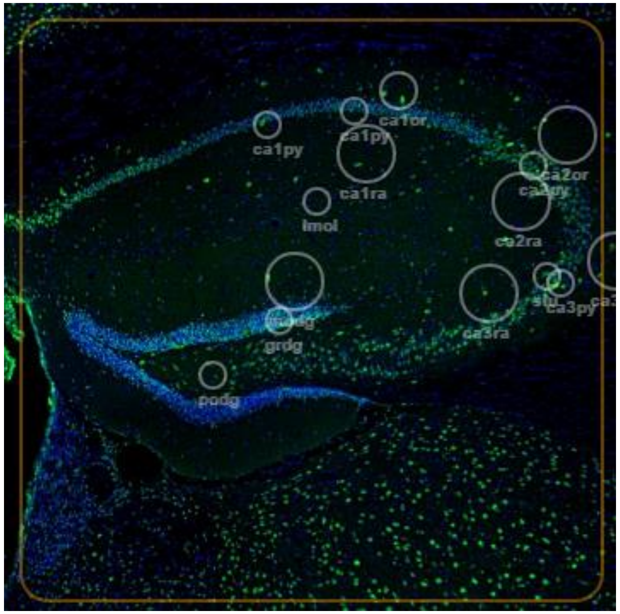
Olfactory bulb

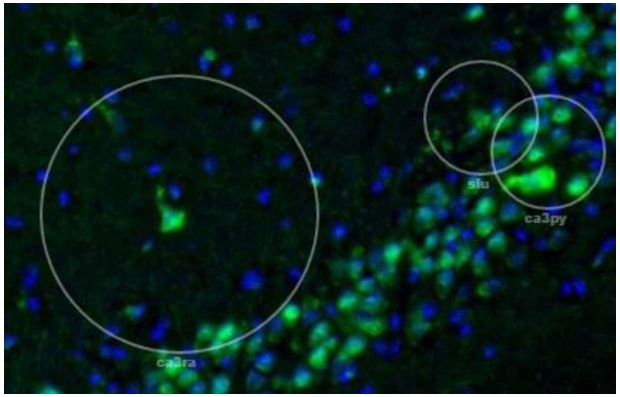
Visual cortex

External cortical inferior colliculi

Lateral paragigantocellular nucleus

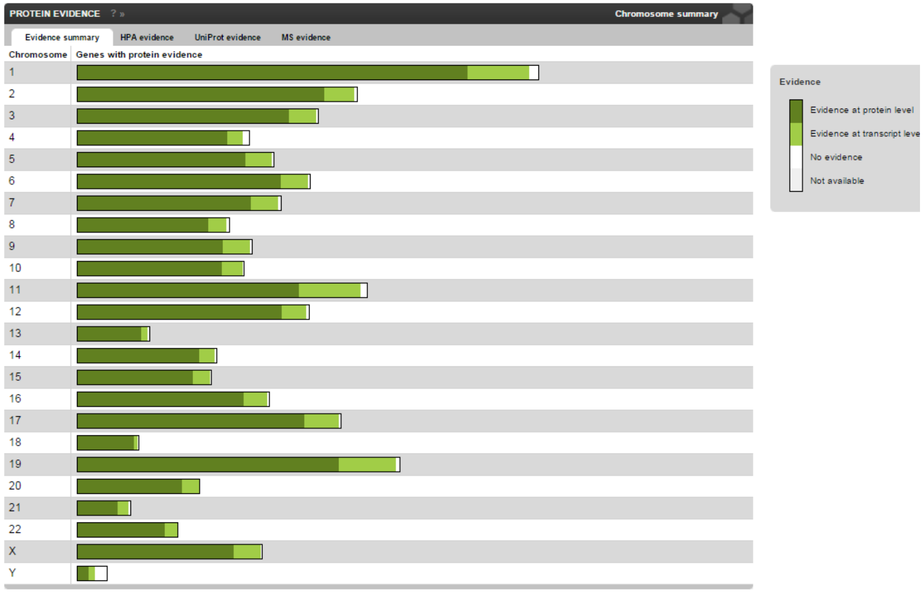
Hippocampus

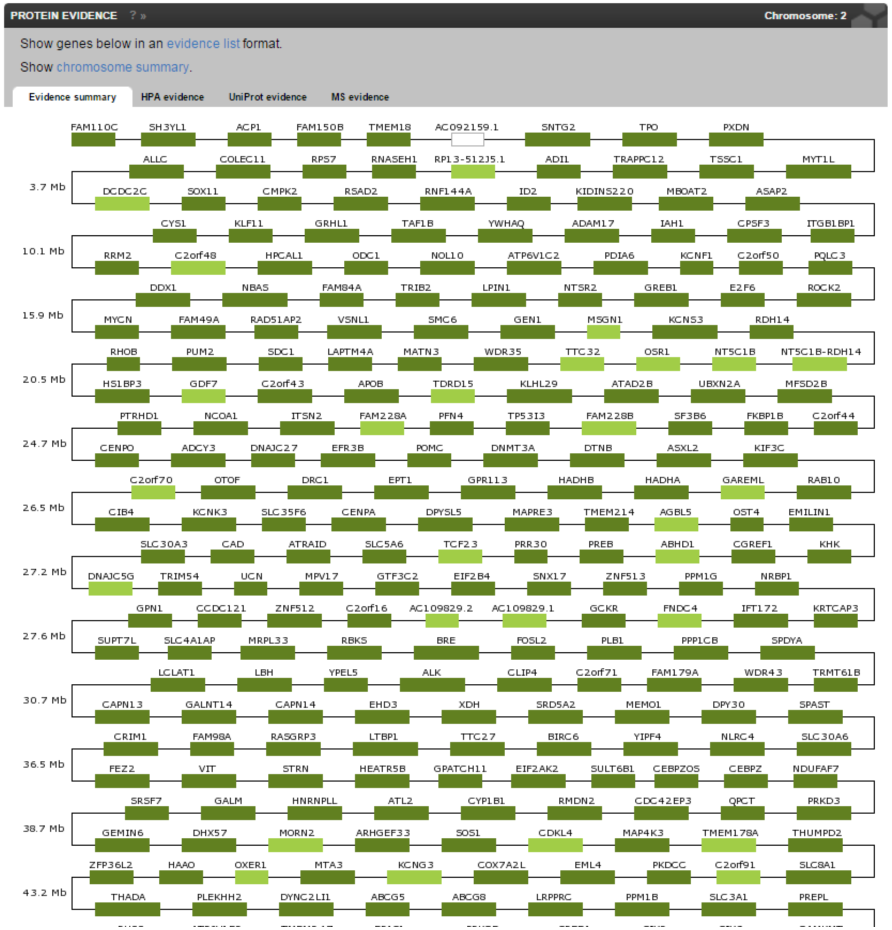





Protein
Relative fluorescence intensity
Mouse brain region
0
255
olfactory bulb, forebrain
Anterior olfactory nucleus (aon)
Granule cell layer (gro)
Internal plexiform layer (ipl)
Mitral cell layer (mi)
Glomerular layer (gl)
Rostral migratory stream (rms)
External plexiform layer (epl)
External plexiform layer of the accessory OB (epla)
N/A
Granule cell layer of the accessory OB (gra)
Glomerular layer of the accessory OB (gla)
basal forebrain, forebrain
Dorsal tenia tecta (dtt)
Caudate putamen (cpu)
Accumbens nucleus, core (acbc)
Accumbens nucleus, shell (acbsh)
Island of Calleja (icj)
Ventral pallidum (vp)
Medial septum (ms)
Nucleus of the vertical limb of the diagonal band (vdb)
Lateral septum (ls)
Nucleus of the horizontal limb of the diagonal band (hdb)
Globus pallidus (gp)
cerebral cortex, forebrain
Frontal association cortex (fra)
Motor cortex (m)
Cingulate cortex (cg)
Piriform cortex, L1 (pirl1)
Piriform cortex, L2 (pirl2)

Evidence map of human proteins





Show genes below in an evidence list format.

Show chromosome summary.

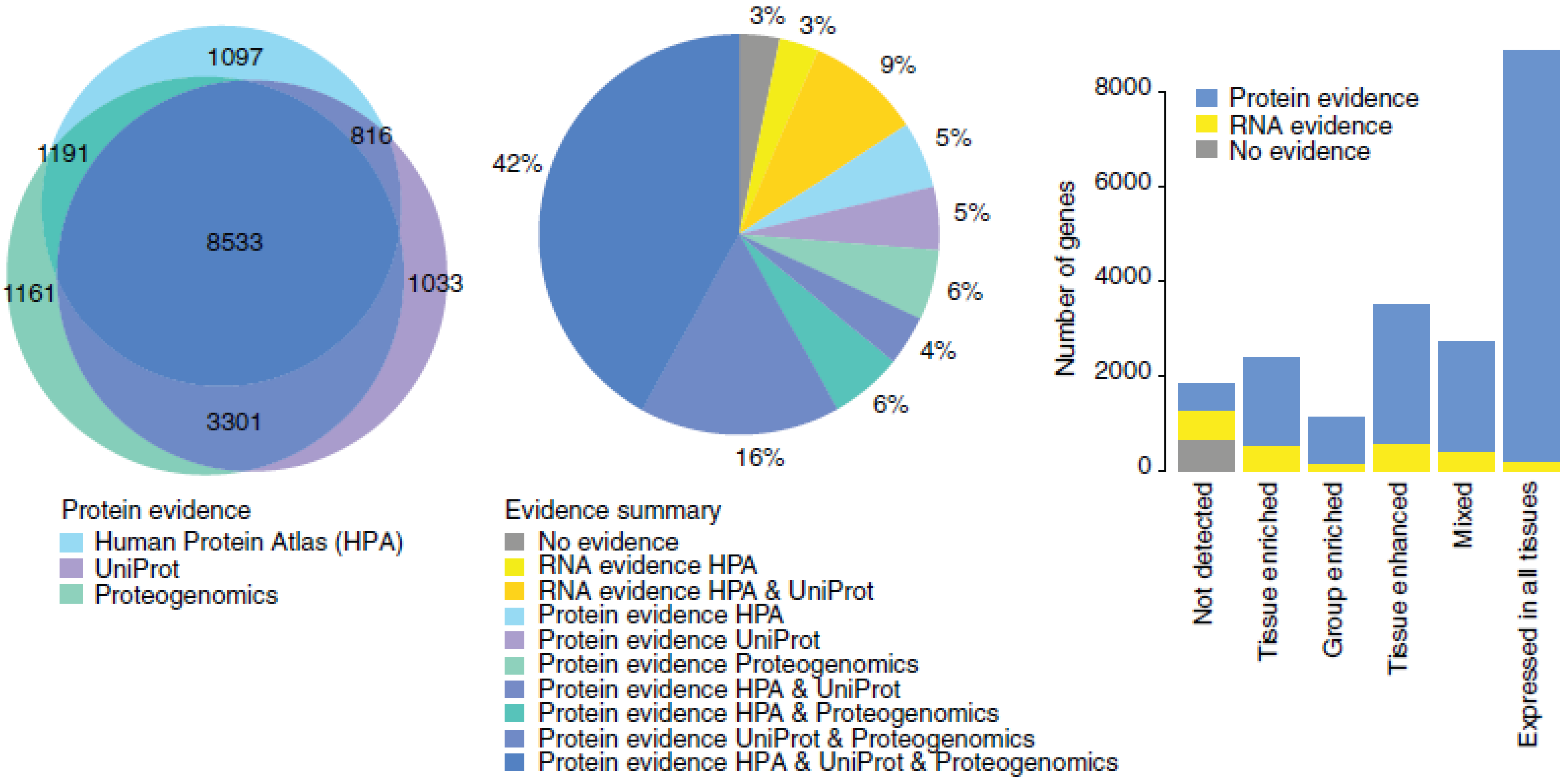
THE HUMAN PROTEIN ATLAS

Knut och Alice Wallenbergs Stiftelse

KTH VETENSKAP OCH KONST

UPPSALA UNIVERSITET

Protein evidence



- ***17,132 protein-coding genes with proteins identified from at least one effort***
- ***677 genes (3.3%) lacking experimental evidence of existence***
- ***18% of genes with elevated expression lack evidence on the protein level***

100 antibodies sent every day

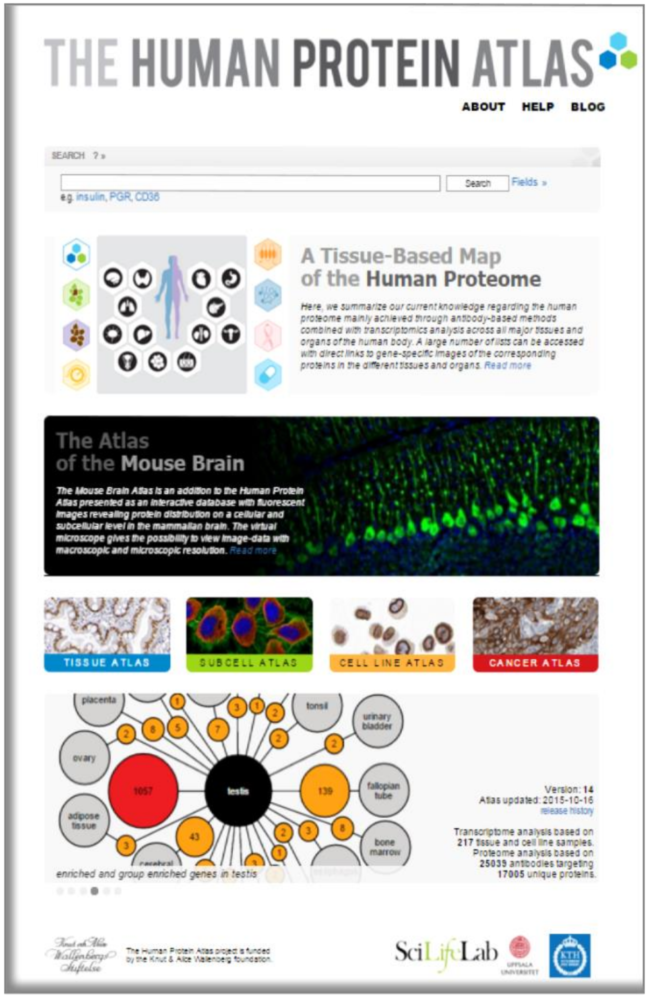
AAA
POLYCLONALS
Designed by researchers, for researchers.

Approximately two publications from external users using the HPA antibodies every day

ATLAS ANTIBODIES

THE HUMAN PROTEIN ATLAS

www.proteinatlas.org



THE HUMAN PROTEIN ATLAS

e.g. insulin, PGR, CD36

A Tissue-Based Map of the Human Proteome

Here, we summarize our current knowledge regarding the human proteome mainly achieved through antibody-based methods combined with transcriptomics analysis across all major tissues and organs of the human body. A large number of lists can be accessed with direct links to gene-specific images of the corresponding proteins in the different tissues and organs. Read more

The Atlas of the Mouse Brain

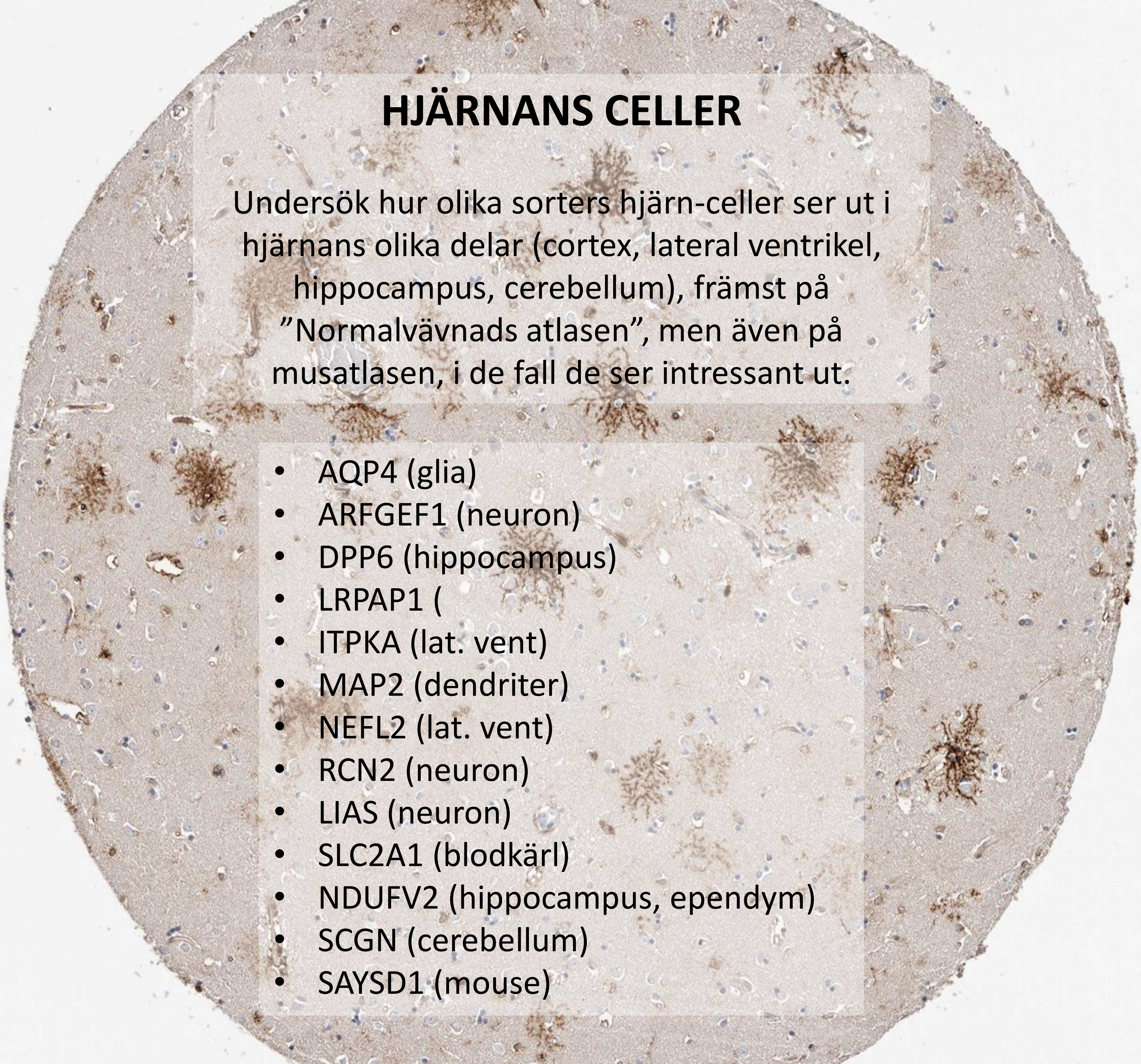
The Mouse Brain Atlas is an addition to the Human Protein Atlas presented as an interactive database with fluorescent images revealing protein distribution on a cellular and subcellular level in the mammalian brain. The virtual microscope gives the possibility to view image-data with macroscopic and microscopic resolution. Read more

enriched and group enriched genes in testis

Version: **14**
Atlas updated: 2015-10-16
release history

Transcriptome analysis based on **217** tissue and cell line samples.
Proteome analysis based on **25039** antibodies targeting **17005** unique proteins.

The Human Protein Atlas project is funded by the Knut & Alice Wallenberg foundation.



HJÄRNANS CELLER

Undersök hur olika sorters hjärn-celler ser ut i hjärnans olika delar (cortex, lateral ventrikel, hippocampus, cerebellum), främst på ”Normalvävnads atlasen”, men även på musatlasen, i de fall de ser intressant ut.

- AQP4 (glia)
- ARFGEF1 (neuron)
- DPP6 (hippocampus)
- LRPAP1 (
- ITPKA (lat. vent)
- MAP2 (dendriter)
- NEFL2 (lat. vent)
- RCN2 (neuron)
- LIAS (neuron)
- SLC2A1 (blodkärl)
- NDUFV2 (hippocampus, ependym)
- SCGN (cerebellum)
- SAYSD1 (mouse)