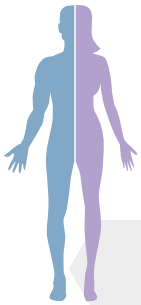
The Power of Proteins

Unraveling Human Complexity Probing the Proteome



When the first near-complete human genome sequence was published in 2004, many researchers were surprised to find only around 20,000 coding genes. This is about the same number as the 1,000-cell nematode, *Caenorhabditis elegans*. It was clear that there is not a direct correlation between the complexity of an organism and its gene count.

We are undoubtedly more than just our DNA sequence. But how much more? Although each cell in our body contains our genetic material, we are built out of proteins. They give us form and function—the house that is built from a DNA blueprint. Thus, understanding the proteome—more specifically, the variety of different proteomes within a single organism—is fundamental to fully elucidating biological function and dysfunction.

Proteomes vary at the organ and tissue level as well as the cellular level. So the proteome for the liver will be markedly different to that of the brain. At the same time, the proteome of a neuron will differ from the proteome of a glial cell. A diseased cell's proteome will not only be different from a normal cell, but the differences could provide clues to what is causing the disease phenotype. When it comes to cancer, a tumor will have its own unique proteome and may even contain multiple proteomes present in distinct clonal outgrowths. Characterizing all of these proteomes can help us to understand why a cell is cancerous, how aggressive a particular cancer may be, and even what drugs might work best against a particular tumor. Identifying and being able to track biomarkers of disease is therefore an important step toward generating targeted and effective therapies. It may also enable treatments to be tailored to specific people (and their own unique proteomes) and specific cancer types and sub-types within those individuals.

Studying the proteome is no simple task. Results must be comparable across the world as well as over time. The techniques used to analyze, characterize, and measure individual proteins must be robust and reliable. Many groups are working toward building a comprehensive proteome database that sets a baseline for normal tissue as well as characterizing the proteomes of tissues in a diseased state. One such project is the Human Protein Atlas (www.proteinatlas.org), a collaboration between the Royal Institute of Technology (KTH), Uppsala University, and the Science for Life Laboratory in Sweden.

In the poster on the reverse of this page, we have summarized and provided examples of some of the rich data housed in the Human Protein Atlas database. As data is added to such repositories, and more powerful computers are used to annotate and mine this data, scientists will gain a clearer sense of what the proteome can tell us and how this information can be used to improve human health and disease treatment.

Sean Sanders, Ph.D.
Editor, Custom Publishing, *Science*

Writers: Mathias Uhlén, Ph.D.; Caroline Kampf, Ph.D.; Fredrik Pontén, M.D., Ph.D. ⬢ Illustrator/Designer: Luca Marziani ⬢ Editor: Sean Sanders, Ph.D.

Sponsored by

Knut och Alice Wallenbergs Stiftelse

THE HUMAN PROTEIN ATLAS

Science AAAS

Produced by the *Science*/AAAS Custom Publishing Office

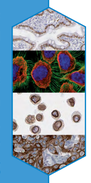
Published on 7 November 2014

THE HUMAN PROTEOME

The Power of Proteins. *The human genome consists of approximately 20,000 protein-coding genes. If DNA can be equated with the blueprint for a home, then proteins can be thought of as the bricks and mortar, plumbing, and paint—essentially everything that makes up the house. This poster summarizes the multiple ongoing antibody- and transcriptomics-based proteome projects and where in the human body this research is focused. For more detailed information, visit: 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 divided into four parts: normal tissue, subcellular, cell lines, and cancer. Transcriptomics data provides gene expression information across different tissues and organs, while antibody-based protein profiles show cell-level localization for the corresponding protein. The Human Protein Atlas (version 13.0) contains protein data for 85% of the translated human genome and includes 13 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 such 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 HOUSEKEEPING PROTEOME

Transcriptomics analysis suggests that some 9,000 genes code for so-called housekeeping proteins, located in every cell, that 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

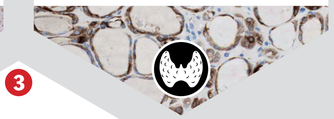
All processes within a living cell are highly regulated, including cell proliferation, differentiation, and death. Regulatory mechanisms include the control of gene expression as well as posttranslational modifications that can regulate protein activity, stability, localization, or degradation. Transcription 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 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.

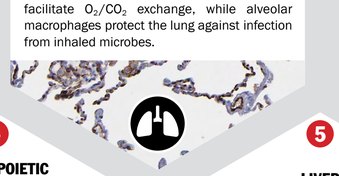
2 ENDOCRINE ORGANS

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.



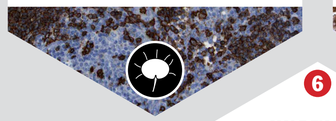
3 LUNG

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.



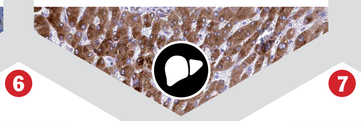
4 HEMATOPOIETIC TISSUES

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 occur in secondary lymphoid organs including the spleen, lymph nodes, and lymphoid tissue located in various organs such as the appendix.



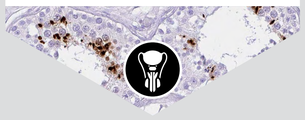
5 LIVER

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.



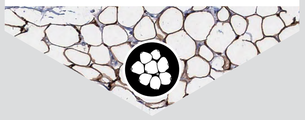
6 MALE TISSUES

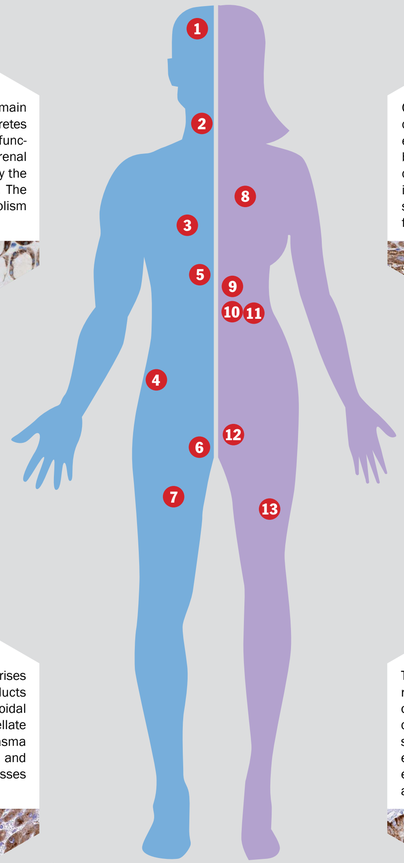
The testes and prostate gland represent the male reproductive tissues. The testes produce sperm and hormones important for the development of male sex characteristics. The prostate secretes semen components that support sperm viability. Many of the testes-specific proteins are involved in the reproductive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



7 ADIPOSE TISSUE

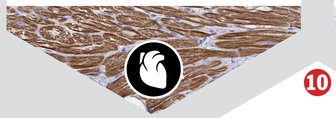
Adipose, or fat, tissue is highly specialized to store energy, secrete hormones, and protect the body. It surrounds important organs (visceral fat) and also forms an insulating layer beneath the skin (subcutaneous fat). Proteins involved in lipid metabolic processes, secretion, and transport are found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





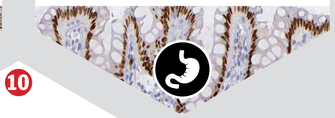
8 CARDIAC AND SKELETAL MUSCLE

Cardiac muscle, comprising 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.



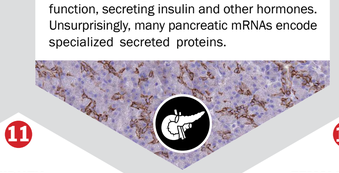
9 GASTROINTESTINAL TISSUES

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.



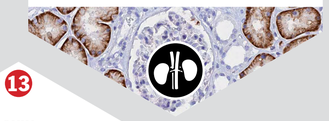
10 PANCREAS

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.



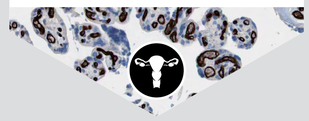
11 KIDNEY

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.



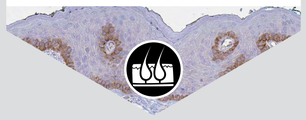
12 FEMALE TISSUES

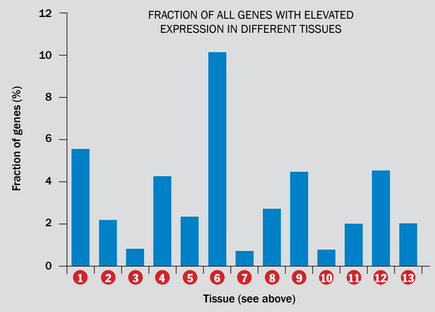
The female reproductive tissues include the ovaries, which produce female reproductive hormones and germ cells; the endometrium, site of fertilized egg implantation; and the placenta. The endometrial mucosa undergoes hormone-driven regeneration during the menstrual cycle, providing the necessary environment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



13 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 the tumorigenesis process. 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.

Sponsored by

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 USE OF ANTIBODIES TO STUDY THE HUMAN PROTEOME

antibodypedia

www.antibodypedia.com

THE POWER OF ANTIBODIES

Antibodies, also known as immunoglobulins, are Y-shaped proteins, which are used by the immune system to identify and destroy foreign objects such as bacteria and viruses. The antibody recognizes a unique part of the foreign target, the antigen. The unique properties of antibodies are used in a wide range of therapeutic and research applications. This poster describes some of the most common techniques.

WESTERN BLOT

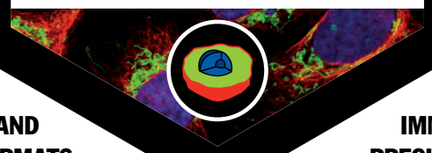
Western blot is an analytical technique used to detect specific proteins in a sample. Proteins are separated on a gel and the result visualized on a membrane using labeled antibodies. It is a common method and almost all available commercial antibodies are validated using this method.

IMMUNOHISTOCHEMISTRY

Immunohistochemistry is a microscopy based technique for visualizing cellular macromolecules, such as proteins, in complex tissues. By using specific antibodies to generate a colored precipitate in the tissue, a visual output of the existence and localization of the target molecule is generated.

IMMUNOCYTOCHEMISTRY

Immunocytochemistry (ICC) is a technique for the visualization of proteins and peptides in cells. In ICC the extracellular matrix around the cells is removed and, by using an antibody linked to a reporter (e.g., a fluorophore), the sub-cellular localization may be seen through a microscope.



FLOW CYTOMETRY

Flow cytometry is a laser-based, biophysical technology used to count, measure size, and detect properties of particles in suspension. A sample of suspended particles is separated through a narrow nozzle, and a laser enables detection of properties of individual particles in the sample.

ANTIBODYPEDIA

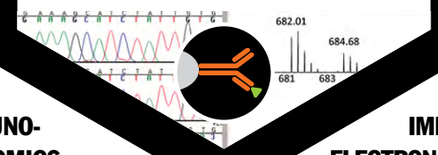
Antibodypedia is a web-based knowledge resource with annotated and scored antibodies from commercial and academic providers. All information is free and accessible in the database. With the knowledge in Antibodypedia you have the power to select the right antibody for the right application.

ELISA AND ARRAY FORMATS

Immunosorbent methods use antibodies and reporters to detect a substance. A common technique is the “enzyme-linked immunosorbent assay” (ELISA) that uses an enzymatic reaction as reporter. The immunoassay format may be miniaturized on microarrays to allow multiplexing for multi-parameter analysis.

IMMUNO-PRECIPITATION

Immunoprecipitation uses antibodies to isolate and concentrate a protein out of a solution containing thousands of proteins. A solid support is used to allow precipitation of the antibody-protein complex. An advantage is that the natural functionality of the native protein is preserved.

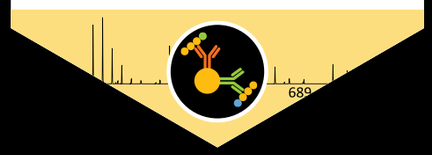


PROXIMITY LIGATION ASSAY

A proximity ligation assay uses a pair of oligonucleotide labeled antibodies binding to different epitopes on a protein, or to epitopes in close proximity on two proteins in a complex. Used for detection, visualization and quantification of single proteins or protein-protein interactions.

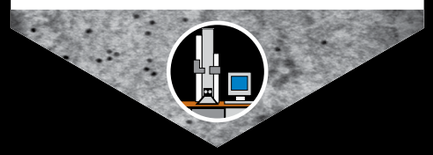
IMMUNO-PROTEOMICS

Immunoproteomics combines the use of antibodies and mass spectrometry to study large sets of proteins. Immuno-affinity enrichment may be used to reduce the large dynamic range in biological samples before MS-analysis. Immunoproteomics is a useful tool within quantitative proteomics.



IMMUNO-ELECTRON MICROSCOPY

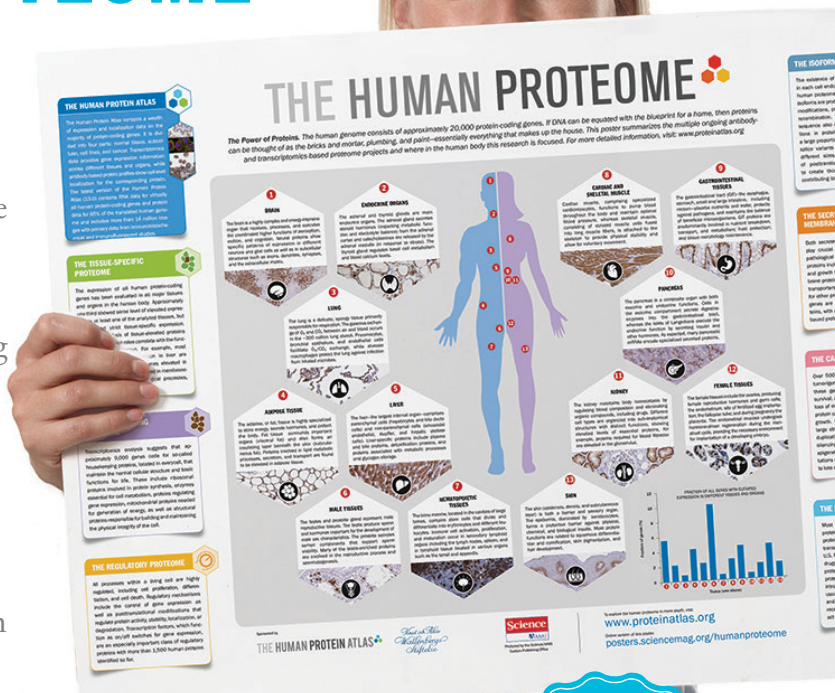
Immunoelectron microscopy combines the ability of an antibody to specifically bind a protein with the high spatial resolution of an electron microscope. Detection of the antibody’s sub-cellular localization in the sample is made by conjugating the antibody with colloidal gold particles.



HUMAAAN PROTEOME

Interested in the human proteome? Then open this poster to discover how leading scientists have mapped the complete human proteome using Triple A Polyclonals. Today we provide over 17,000 antibodies, which target more than 75% of the human protein coding genes to researchers worldwide. Triple A Polyclonals offer advanced levels of specificity, reliability and versatility.

Want to use them in your research? Learn more today at atlasantibodies.com



ATLAS ANTIBODIES
Totally human