



## The potential clinical impact of the tissue-based map of the human proteome

Cecilia Lindskog

**To cite this article:** Cecilia Lindskog (2015) The potential clinical impact of the tissue-based map of the human proteome, Expert Review of Proteomics, 12:3, 213-215, DOI: [10.1586/14789450.2015.1040771](https://doi.org/10.1586/14789450.2015.1040771)

**To link to this article:** <https://doi.org/10.1586/14789450.2015.1040771>



Published online: 30 Apr 2015.



Submit your article to this journal [↗](#)



Article views: 2466



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 8 View citing articles [↗](#)

# The potential clinical impact of the tissue-based map of the human proteome

Expert Rev. Proteomics 12(3), 213–215 (2015)



**Cecilia Lindskog**

Department of Immunology,  
Genetics and Pathology, Science  
for Life Laboratory, Uppsala  
University, Uppsala, Sweden  
Tel.: +46 18 471 50 32  
cecilia.lindskog@igp.uu.se

Since the first draft of the human genome sequence was published, several attempts have been made to map the human proteome, the functional representation of the genome. One such initiative is the Human Protein Atlas project, which recently released a tissue-based map of the human proteome. The Human Protein Atlas is based on the combination of transcriptomics and antibody-based proteomics for mapping the human proteome down to the single cell level. The comprehensive publicly available database contains more than 13 million unique immunohistochemistry images and provides an excellent resource for exploration and investigation of future drug targets and disease biomarkers.

All living cells consist of proteins, the functional representation of the genome responsible for a multitude of essential functions needed to maintain life and make us who we are. To be able to understand the cellular interactions under normal and pathological conditions, and translate biological findings into clinical applications, in-depth knowledge of the molecular repertoire in the normal human body is crucial. Since the human genome sequence became available, several attempts have been made to create a map of the human proteome.

Two comprehensive drafts of the human proteome include the Human Proteome Map [1] and the proteomics database [2], based on mass spectrometry efforts. Another recent initiative is the tissue-based map of the human proteome, generated by the Human Protein Atlas project. Uhlen *et al.* [3] used antibody-based profiling for analysis of protein expression in a tissue context, presented in an interactive website [4] containing protein data for at least one major isoform of 85% of the translated human genome. Each of the >20,000 analyzed antibodies has been used for immunohistochemistry

on tissue microarrays containing samples from 44 different normal organs and tissues, as well as the 20 most common cancer types. The analysis is combined with RNA-seq of 32 different tissues and organs, and data from 44 cell lines at both the RNA and protein level. The public database contains more than 13 million immunohistochemically stained images available for free in a virtual microscope, which offers a unique possibility for researchers to determine the exact distribution of the protein expression *in situ*.

While mass spectrometry provides the standard for quantifying a certain set of proteins in a sample, immunohistochemistry has the advantage of adding spatial resolution and information on expression pattern in certain cells or subcellular structures. As human tissues are heterogeneous mixtures of different cell types, the opportunity to analyze the protein expression in a single cell resolution, with intact cell structure and tissue morphology, offers further understanding of the underlying biology and function of the protein.

In the Human Protein Atlas, the expression of all genes has been categorized based on expression level and tissue

EXPERT  
REVIEWS

**KEYWORDS:** immunohistochemistry • proteome • proteomics • tissue microarrays • transcriptomics

distribution, and is presented on numerous comprehensive knowledge pages summarizing the proteomes and transcriptome of each organ. A striking observation is that almost half of the proteins were shown to be expressed in all analyzed tissues, and these 'housekeeping' proteins are considered to be responsible for maintaining the basic functions for life. The housekeeping proteome includes 9000 genes encoding for example, ribosomal proteins, enzymes and mitochondrial proteins. An interesting group of proteins are proteins that were shown to be elevated in a certain tissue or group of tissues, compared with all other analyzed tissues. It can be anticipated that these tissue-elevated genes play important roles in the organ physiology. The tissue-elevated proteins are displayed in easily accessible and clickable lists, providing the basis for further studies aimed at understanding the molecular repertoire of tissues and organs in healthy and diseased states [5,6]. Exploration of the expression patterns of these proteins at the single cell level could identify targets relevant for immunohistochemical studies in clinical cohorts. These proteins may have important implications for medicine and aid in identifying and stratifying high-risk individuals, guiding treatment modalities, as well as contributing to further understanding of underlying disease mechanisms. Tissue-elevated proteins could also be searched for and analyzed in serum, for example, using affinity proteomics on suspension bead arrays, as they represent an interesting group of proteins that may be suitable for serum-based diagnostic tests. In a recent study by Haggmark *et al.* [7] focusing on plasma profiling in samples from amyotrophic lateral sclerosis, two out of the three proteins shown to be associated with the disease were tissue-elevated. Serum-based analysis can also be used to discover autoantigens for identification of potential autoimmune targets [8].

Uhlén *et al.* presents various subproteomes corresponding to particular functional groups of genes. Similar to the organ proteomes, clickable lists with genes belonging to different subproteomes are easily accessible on the website, allowing for screening and in-depth studies of proteins with a certain function. One such group of proteins is the druggable proteome. Most pharmaceutical drugs target proteins, and the US FDA has currently approved drugs targeting 620 human proteins. Interestingly, a large fraction (30%) of these proteins was shown to be expressed in all tissues and organs. A majority of the FDA-approved drug targets are secreted or membrane-bound proteins, and based on prediction methods, the Human Protein Atlas displays a separate chapter with the estimated complete set of the human secretome and membrane proteome. These approximately 3000 secreted proteins and 5500 membrane-bound proteins constitute important lists of genes for both functional and pharmaceutical studies. Another relevant group of proteins is the cancer proteome, listing 525 genes implicated in tumorigenesis, based on different genetic alterations. For many of these genes, visualization of the protein expression using immunohistochemistry adds important information as to differential expression between different forms of cancer or individual tumors within the same cancer type, as well as differences in protein expression between tumors and

the corresponding normal tissue. Other subproteomes include the regulatory proteome, focusing on 1510 known transcription factors; and the isoform proteome, discussing, for example, splice variants and post-translational modifications.

The broad spectrum of tissues analyzed in the Human Protein Atlas allows for searches of proteins expressed in certain tissues or groups of tissues [9], to generate gene lists with potential biomarker candidates that can be analyzed further on extended material. This strategy is applicable not only for cancer research but also for projects in numerous other areas, such as in the field of diabetes, in the quest for suitable candidates for imaging of beta cells [10]. The extensive amount of protein expression data in the Human Protein Atlas is available for download, and constitutes an important resource for cross-referencing of genes identified by other methods [11,12]. Moreover, the high-resolution images provided in the Human Protein Atlas may be further studied with image analysis [13], a technology likely to become increasingly important to complement the manual scoring performed by a pathologist, to produce a less subjective immunohistochemistry assessment.

The analysis of proteins in a tissue context using immunohistochemistry offers a more direct understanding of the protein function than genomes, and the method itself is easily transferable to the clinical utility. Immunohistochemistry is a widespread method used in most diagnostic laboratories, and has emerged as a validation tool in biomarker discovery. However, despite the vast number of potential prognostic and diagnostic markers being described in literature, it is surprising that few are put to practical use [14,15]. Translating a potential proteomic biomarkers to candidates used in the clinic is challenging, and on the road from discovery to approval, there are several pitfalls. This may in part be explained by inappropriate study design or lack of well-characterized, specific antibodies suitable for immunohistochemistry. Another explanation may be limited knowledge of the nature of the protein in a broader context, taking into consideration how the protein expression is distributed in a large set of different tissues and organs. The affinity-purified polyclonal antibodies used in the Human Protein Atlas have undergone several quality steps during production, and the immunohistochemical staining pattern of both tissues and cell lines has been critically evaluated in comparison with RNA expression data and previously published gene/protein characterization data from other sources. Only antibodies that passed the strict validation criteria were included in the publically available portal. For the approved antibodies, a combination of several different criteria was applied to assess the quality and give each antibody either a supportive or uncertain score. More than 5000 genes were analyzed using two or more antibodies recognizing different epitopes of the same target, adding a higher level of reliability in cases where the two antibodies showed a similar staining pattern. In the case of well characterized and differentially expressed genes, the large spectrum of tissues analyzed in the Human Protein Atlas in combination with RNA-seq data offers both positive and negative controls for immunohistochemistry quality control and

optimization. Many proteins are, however, largely uncharacterized with only limited characterization data available, and as the specificity of the staining is context dependent, all primary data is provided on the Human Protein Atlas portal, making it possible for users of the database to do an individual judgment of the antibody reliability.

The Human Protein Atlas focuses on the major isoforms of each protein, and has thus limitations in generating information on post-translational modifications, which play a key role in the regulation of biological processes and signaling pathways. For a full understanding of the functional repertoire of a certain protein or group of proteins, the information retrieved on the Human Protein Atlas should be complemented with other methods, such as proximity ligation assay [16] or mass spectrometry [17]. The tissue-based immunohistochemistry data may also be combined with imaging mass spectrometry [18], which in addition to the benefit of generating a quantitative measurement, also allows for multiplexing.

In summary, the Human Protein Atlas portal generated by Uhlen *et al.* represents an invaluable resource to gain biological insight on human proteins. The protein expression data are likely to become useful for numerous spin-off projects in basic and clinical research, and will also have a large impact on the pharmaceutical industry.

#### Acknowledgement

*The entire staff of the Human Protein Atlas project and the Science for Life Laboratory is acknowledged for valuable contributions.*

#### Financial & competing interests disclosure

*This work was supported by the Knut and Alice Wallenberg Foundation. The author has no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.*

*No writing assistance was utilized in the production of this manuscript.*

#### References

- Kim MS, Pinto SM, Getnet D, et al. A draft map of the human proteome. *Nature* 2014;509:575-81
- Wilhelm M, Schlegl J, Hahne H, et al. Mass-spectrometry-based draft of the human proteome. *Nature* 2014;509:582-7
- Uhlen M, Fagerberg L, Hallstrom BM, et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015;347:1260419
- The human protein atlas. Available from: [www.proteinatlas.org](http://www.proteinatlas.org)
- Lindskog C, Fagerberg L, Hallstrom B, et al. The lung-specific proteome defined by integration of transcriptomics and antibody-based profiling. *FASEB J* 2014;28:5184-96
- Kampf C, Mardinoglu A, Fagerberg L, et al. The human liver-specific proteome defined by transcriptomics and antibody-based profiling. *FASEB J* 2014;28:2901-14
- Haggmark A, Mikus M, Mohsenchian A, et al. Plasma profiling reveals three proteins associated to amyotrophic lateral sclerosis. *Ann Clin Transl Neurol* 2014;1:544-53
- Haggmark A, Hamsten C, Wiklundh E, et al. Proteomic profiling reveals autoimmune targets in sarcoidosis. *Am J Respir Crit Care Med* 2015;191:574-83
- Ponten F, Schwenk JM, Asplund A, Edqvist PH. The Human Protein Atlas as a proteomic resource for biomarker discovery. *J Intern Med* 2011;270:428-46
- Lindskog C, Korsgren O, Ponten F, et al. Novel pancreatic beta cell-specific proteins. antibody-based proteomics for identification of new biomarker candidates. *J Proteomics* 2012;75:2611-20
- Mardinoglu A, Agren R, Kampf C, et al. Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat Commun* 2014;5:3083
- Edlund K, Lindskog C, Saito A, et al. CD99 is a novel prognostic stromal marker in non-small cell lung cancer. *Int J Cancer* 2012;131:2264-73
- Kumar A, Rao A, Bhavani S, et al. Automated analysis of immunohistochemistry images identifies candidate location biomarkers for cancers. *Proc Natl Acad Sci USA* 2014;111:18249-54
- O'Hurley G, Sjostedt E, Rahman A, et al. Garbage in, garbage out. a critical evaluation of strategies used for validation of immunohistochemical biomarkers. *Mol Oncol* 2014;8:783-98
- Lindskog C, Edlund K, Mattsson JS, Micke P. Immunohistochemistry-based prognostic biomarkers in NSCLC. novel findings on the road to clinical use? *Expert Rev Mol Diagn* 2015;15(4):471-90
- Soderberg O, Leuchowius KJ, Gullberg M, et al. Characterizing proteins and their interactions in cells and tissues using the in situ proximity ligation assay. *Methods* 2008;45:227-32
- Witze ES, Old WM, Resing KA, Ahn NG. Mapping protein post-translational modifications with mass spectrometry. *Nat Methods* 2007;4:798-806
- Giesen C, Wang HA, Schapiro D, et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat Methods* 2014;11:417-22