FEBS Letters 430 (1998) 64–72 FEBS 20292

Minireview

Human and mouse proteomic databases: novel resources in the protein universe

Julio E. Celis*, Morten Østergaard, Niels A. Jensen, Irina Gromova, Hanne Holm Rasmussen, Pavel Gromov

Department of Medical Biochemistry and Danish Centres for Human Genome Research and Molecular Gerontology, Ole Worms Allé Building 170, The University of Aarhus, DK-8000 Aarhus C, Denmark

Received 24 April 1998

Abstract Proteomics[1] is an emerging area of research of the post-genomic era that deals with the global analysis of gene expression using a plethora of techniques to resolve (high resolution two-dimensional polyacrylamide gel electrophoresis, 2D PAGE), identify (peptide sequencing by Edman degradation, mass spectrometry, Western immunoblotting, etc.), quantitate and characterize proteins, as well as to store (comprehensive 2D PAGE databases), communicate and interlink protein and DNA sequence and mapping information from genome projects. Here we review the current status as well as applications of human and mouse proteomic 2D PAGE databases that are being systematically constructed for the global analysis of gene expression in both health and disease (http://biobase.dk/cgi-bin/celis). Furthermore, we discuss the problems one faces when using powerful proteomic technology to study heterogeneous tissue and tumor biopsies, and emphasize the importance of building comprehensive databases that contain a critical mass of information for both known and novel proteins in normal and disease conditions.

Key words: Human genome; Mouse genome; Proteomics; Proteome profiling; Two-dimensional polyacrylamide gel electrophoresis database; Disease; Cancer; Tissue heterogeneity; Transgene

1. Introduction

Proteomics is an emerging area of research of the post-genomic era that deals with the global analysis of gene expression using a combination of techniques to resolve (high resolution two-dimensional polyacrylamide gel electrophoresis, 2D PAGE), identify (peptide sequencing by Edman degradation, mass spectrometry, Western immunoblotting, etc.), quantitate (scanners, phosphorimager, etc.) and characterize proteins, as well as to store (comprehensive 2D PAGE databases), communicate and interlink protein and forthcoming DNA sequence and mapping information from genome projects (http://biobase.dk/cgi-bin/celis) [1,2]. Each one of these technologies can be applied independently, although their impact can be maximized when used in concert to the study of complex biological problems.

For the last 23 years, high resolution 2D PAGE has been the technique of choice to analyze the protein composition of a given cell type and for monitoring changes in gene activity through the quantitative and qualitative analysis of the thousands of proteins that orchestrate various cellular functions ([3–5] and references therein). Proteins are usually the functional molecules and, therefore, the most likely components to reflect qualitative (expression of new proteins, post-translational modifications) and quantitative (up and down regulation, coordinated expression) differences in gene expression. Genes may be present, they may be mutated, but they are not necessarily transcribed. Some genes are transcribed but not translated, and the number of mRNA copies per cell does not necessarily reflects the number of functional protein molecules [6]. For these reasons focusing on the proteins has certain advantages as compared to the mRNA based technology.

The 2D gel technology, which was originally described by O'Farrell [7,8] and Klose [9], separates proteins in terms of both their isoelectric point (p*I*) and molecular weight. Usually, one chooses a condition of interest and lets the cell reveal the global protein expression response, as all detected proteins can be analyzed both qualitatively and quantitatively (relative abundance, post-translational modifications, co-regulated proteins, etc.) in relation to each other. Currently, the carrier ampholyte based 2D gel technology [7,8,10] can routinely resolve nearly 4000 [^{35}S]methionine labeled polypeptides from unfractionated cell extracts ([10], see also procedures and videos at http://biobase.dk/cgi-bin/celis). The lowest level of detection for [^{35}S]methionine labeled proteins correspond to polypeptides that are present in about 40 000 molecules per cell, although immunoblotting in combination with enhanced chemiluminescence (ECL) detection can reveal polypeptides that are present in as few as 1000 molecules per cell [11]. Low abundance proteins can be visualized either by increasing the sensitivity of the detection procedures, and/or by the analysis of subcellular fractions or partially purified protein samples. Recently, thanks to the advent of immobilized pH gradients (IPGs) [12–14], which are an integral part of the polyacrylamide matrix, one can generate more reproducible focusing patterns to obtain enhanced resolution, as very narrow pH gradients of about 0.05 pH/cm can be established. In the near future, 2D gels will continue to provide the foundation for a high throughput proteomic approach, although other alternatives are being considered [2].

*Corresponding author. Fax: (45) 8613 1160.
E-mail: jec@biokemi.au.dk

[1] The term proteome was coined by Wilkins et al., Biotech. Genet. Eng. Rev. 13 (1996) 19–50.

Abbreviations: 2D, two-dimensional; PAGE, polyacrylamide gel electrophoresis

PII: S0014-5793(98)00527-4

Through the years, major advances in hardware and software development for scanning and analyzing 2D gels ([15,16] and references therein), as well as the development of sensitive tools for rapid protein identification (immunoblotting, Edman degradation peptide sequencing, mass spectrometry) [17–24], have stimulated the creation of comprehensive proteomic 2D PAGE databases that aim at interfacing protein information with forthcoming DNA mapping and sequence data from genome projects, and that offer a global approach to the study of gene expression both in health and disease (http://biobase.dk/cgi-bin/celis) ([3,4,25–27] and references therein). Besides annotating genomes, these databases are expected to address problems that cannot be approached by DNA analysis, namely, relative abundance of the protein product, post-translational modifications, subcellular localization, turnover, interaction with other proteins as well as functional aspects. Today, using the integrated approach offered by proteomic 2D PAGE databases it is possible to reveal phenotype-associated proteins, microsequence them, search for homology with previously identified proteins, clone the cDNAs, assign partial protein sequences to genes for which the full DNA sequence and the chromosome location are known, and study the regulatory properties and function of the thousands of protein components of organelles, pathways, and cytoskeletal systems in health and disease [28]. Proteomics is only one of the many technologies that are currently available for the study of global cell regulation, and we envision that in the near future complex biological problems will be addressed using a plethora of technology from the realm of functional genomics [29,30] that include genomics, high throughput proteomics, cDNA arrays, phage antibody libraries as well as transgenic animals (Fig. 1).

The purpose of this short article is to review the current status as well as applications of human and mouse proteomic 2D PAGE databases which are being systematically constructed for the study of global gene expression both in health and disease (http://biobase.dk/cgi-bin/celis). Furthermore, we discuss the problems one faces when using powerful proteomic technology to study heterogeneous tissues and tumors, and emphasize the importance of building comprehensive databases that contain a critical mass of information for both known and novel proteins in normal and disease conditions.

2. Aarhus proteomic databases on the Internet: http://biobase.dk/cgi-bin/celis

2.1. The keratinocyte 2D PAGE database

The keratinocyte 2D PAGE database [31], which is being systematically constructed using carrier ampholytes, is the largest of its kind and currently lists 3625 cellular (2313 isoelectric focusing, IEF; 954 non-equilibrium pH gradient electrophoresis, NEPHGE) and externalized polypeptides (358, IEF) of which 1285 have been identified using a combination of techniques including immunoblotting [32], Edman degradation of internal peptides [33,34], and mass spectrometry [35]. It should be noted that immunoblotting was the first tool available for the rapid identification of proteins resolved by 2D gels [36], and in a way marked the beginning of what we now call the proteomic era.

Fig. 2 shows the synthetic master 2D PAGE image (IEF) of non-cultured human keratinocyte proteins as depicted in the World Wide Web (http://biobase.dk/cgi-bin/celis). Proteins flagged with a red cross correspond to known polypeptides. Today, mass spectrometry represents the technique of choice for rapid protein identification, as it is fast and requires only picomolar amounts of proteins [21,23,24]. In addition, peptide sequencing is now possible using mass spectrometry, a fact that will greatly speed up the identification of hitherto unknown proteins on 2D gels as well as the cloning of genes coding for novel proteins [19,20,22,23].

Information obtained on any given polypeptide, known or unknown, can be easily retrieved from the database by clicking on the corresponding spot (in this case the spots indicated in Fig. 2; proliferating cell nuclear antigen, PCNA, and the unknown protein IEF SSP 8509). A file listing all of the information entered for the particular protein, mostly obtained from experiments performed in our laboratory, is displayed in the format shown in Fig. 3 (PCNA) and Fig. 4 (IEF SSP 8509) (only a fraction of the files are shown). PCNA/cyclin [37,38] was perhaps the first interesting protein characterized

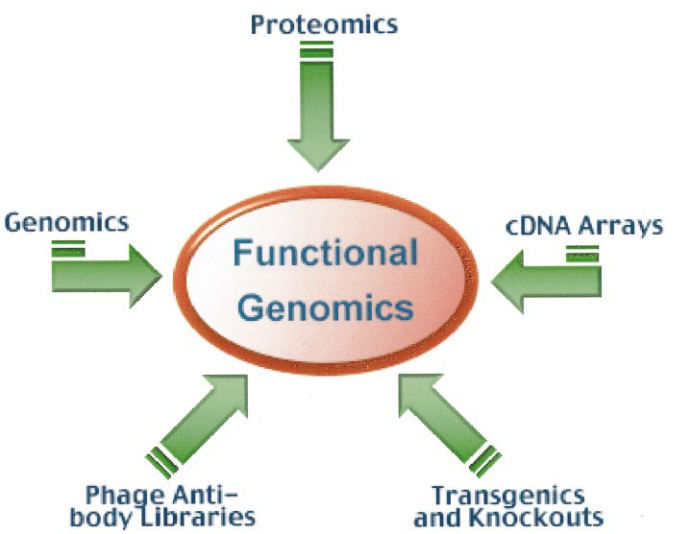


Fig. 1. Technology in functional genomics.

using the 2D PAGE technology. Files for known proteinscontain links to a subset of Medline (http://www.ncbi.nlm.nih.gov/PubMed/), SwissProt (http://expasy.hcuge.ch/sprot/sprot-top.html) and PDB (http://www.pdb.bnl.gov). Other links include OMIM (http://www.ncbi.nlm.nih.gov/Omim/), GeneCards (http://bioinformatics.weizmann.ac.il/cards), UniGene (http://www.ncbi.nlm.nih.gov/UniGene/index.html) and other Web sites such as CySPID (cytoskeletal protein database; http://paella.med.yale.edu/~panzer/cytoskdb/index.html), metabolic pathways (compiled by KEGG; http://www.genome.ad.jp/kegg/), the cytokine explorer (http://kbot.mig.missouri.edu:443/cytokines/explorer.html), histology images (http://biosun.biobase.dk/~pdi/jecelis/micrographs.html), etc. In the future, as new databases and related Web sites become available, it will be possible to navigate throughout various databanks containing complementary information on nucleic acid and protein sequences, genome mapping, diseases, protein structure, post-translational modifications, antibodies and cellular localization of the antigen, signalling pathways, histology, etc.

Functions to query the databases include search by name, protein number or keywords, molecular weight and p*I*, as well as organelle or cellular component (Fig. 2, right-hand panel). As an example, Fig. 5 shows a list of all proteins recorded in the keratinocyte database that can be retrieved when searching for proteins having M_rs between 50 and 55 kDa and p*I*s between 5.0 and 5.5. The list can be extended to other databases available at the site, provided one indicates so when performing the search. By using other functions listed in Fig. 2 it is possible to flip (important as gels are presented and analyzed in different orientations by various laboratories) and enlarge the master synthetic image, as well as to retrieve a list of known proteins and information categories available in the database. Categories or entries are created so as to gather information on physical, chemical, biochemical, physiological, genetic, architectural as well as biological properties of proteins. In general, entries reflect the type of specific biological problem that is being studied using the database approach.

At present, about 80 information categories are available in the World Wide Web version of the IEF keratinocyte database, including protein name, cellular localization, proteins differentially regulated in differentiated keratinocytes, proteins

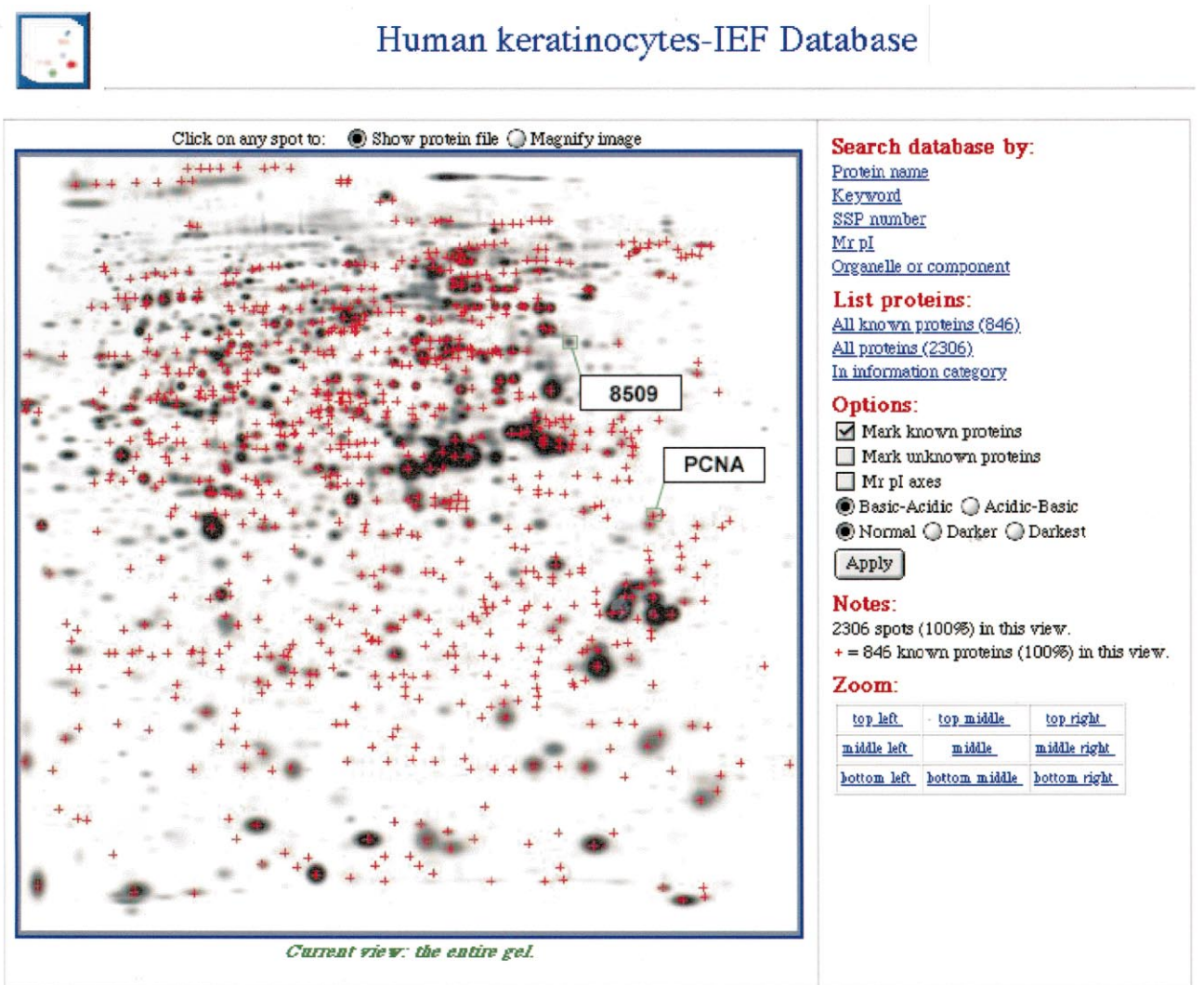


Fig. 2. Master synthetic image of human keratinocyte proteins separated by IEF 2D PAGE as depicted in the World Wide Web. Proteins flagged with a red cross correspond to known proteins. By clicking on any spot it is possible to obtain a file containing information about the protein as well as links to other sites in the World Wide Web.

expressed in other cell types (HeLa cells, fibroblasts, urothelial cells), heat shock proteins, proteins affected by interleukins (1α, 1β, 2, 3, 6, 7 and 8), proteins affected by interferons (α, β and γ), proteins differentially regulated in keratinocytes treated with various chemicals (calcium, okadaic acid, phorbol myristate acetate, retinoic acid, staurosporin), transformation sensitive proteins, levels in fetal human tissues, etc. Some of these entries are also available in other databases accessible at the Web site.

To facilitate the comparison between keratinocyte proteins gels run with carrier ampholytes and IPGs we have provided in the gel gallery section of our Web site IPG gel images that can be connected directly to the master keratinocyte image, by clicking 100 known polypeptides selected as reference. In addition, we provide protocols and videos that can be used to reproduce the data displayed. Finally, the Web site also offers a gallery of 2D gel immunoblots that underline the value of the 2D gel technology to determine antibody specificity as well as to reveal post-translational modifications. 2D blots of keratinocyte proteins reacted with antibodies raised against low abundance proteins which exhibit extensive modifications (SHC, Fig. 6A; cortactin, Fig. 6B), and developed using the ECL procedure are shown in Fig. 6. Clearly, this technology reveals the extent of the modification(s) and can be effectively complemented by mass spectrometry and labeling with specific radioactive precursors [39] to determine the nature of the modification(s).

2.2. *Applications of the keratinocyte database to the study of bladder squamous cell carcinomas*

So far, the keratinocyte databank has been instrumental for the study of fresh bladder squamous cell carcinomas (SCC), as the latter cells closely resemble keratinocytes both in morphology and protein expression profiles [40–42]. Thus, by using the information already stored in the master database, in particular information concerning differentiation markers, tissue distribution and externalized proteins, it has been possible to identify protein markers that define the degree of differentiation of these lesions [43] as well as a single biomarker, the calcium binding and chemoattractant protein psoriasin [42–45], which is expressed specifically by some differentiated tumor cells and is externalized to the urine of SCC bearing

Spot information: SSP 9226

Proliferating cell nuclear antigen (PCNA). PCNA/cyclin.; Mr = 37.4 kDa ; pI = 4.4

For more information, see also MEDLINE or the gopher servers GenBank, Swiss-Prot, PIR, and PDB
Note: Netscape may not return from the Gopher Menu page using 'Go Back'.
Use the Window menu History command to return from the Gopher search.

Category	Description	Annotation Entry	Link
genecards	Link to GeneCards.		GeneCards
humanchrom	Human Chromosome.		Human/Mouse Homology
omim	Link to OMIM.		OMIM
prosite	Link to Prosite.		PROSITE
swissprot	Link to Swiss-Prot.		SWISS-PROT
unigene			UniGene
protname	Protein Name.	Proliferating cell nuclear antigen (PCNA). PCNA/cyclin.	
autoantigen		Autoantigen.	
bfgf	Keratinocyte Proteins Affected by bFGF.	Not affected. J.E.Celis, unpublished observation.	Cytokine Explorer
cellcycle		Increase synthesis during S-phase.	Brite
cellsignetdatab		Listed in CSNDB.	Cell Signalling Network Database (CSNDB)
cellulloc	Cellular Localization.	Nucleus.	
cellulrole	Cellular Role According to EGAD.	DNA replication.	
credinv	Credit to Investigator that Aided the Identification.	R.Bravo and J.E.Celis, Aarhus.	
difregdiffker	Proteins Differentially Regulated in In Vitro Differentiated Keratinoytes.	Downregulated. Ratio differentiated/proliferating = 0.7. E.Olsen et al., Electrophoresis 16, 2241, 1995.	
distriexker	Distribution in Triton Extracted Keratinocytes.	Triton cytoskeleton and supernatant. J.E.Celis, unpublished observation.	

Fig. 3. File for the proliferating cell nuclear antigen (PCNA). Only some of the entries available in the Internet version of the database are shown.

patients [42,43]. Psoriasin, alone or in combination with other biomarkers, may be useful for the follow-up of SCC bearing patients, in particular males, as we have previously shown that the presence of stratified squamous epithelia in the female trigone may lead to false positives [42].

Since SCCs have a very bad prognosis, one long-term aim of our studies is to identify premalignant lesions – and eventually protein markers that are externalized to the urine – that may pinpoint individuals at risk very early during the development of the disease. This goal is being pursued systematically by applying proteomic technology in combination with immunohistochemistry to the analysis of tumors and urothelial tissue biopsies obtained from patients that have undergone removal of the bladder due to invasive disease (cystectomy). Considering that bladder cancer is a field disease [46], that is, a large part of the bladder urothelium is at risk of developing disease, we surmised that some areas in the involved urothelium of SCC patients may exhibit protein expression profiles and immunofluorescence staining patterns that closely resemble those observed in the invasive tumors. These lesions, if detected early enough in the disease process, may be crucial in determining treatment modality.

To illustrate the approach, Fig. 7 shows immunofluorescence staining of cryostat sections of an SCC tumor (Fig. 7A,C,E,G) and of the involved urothelium from the same patient (Fig. 7B,D,F,H) reacted with antibodies against keratins 19, 14, 13 and the basal antigen BG3C8 [47], respectively. To decide which antibody to include in the immunohistochemistry battery is a long process, which requires first the analysis of the protein profiles of many tumors and biopsy material exhibiting various degrees of atypia. Once major protein differences are found, the next step consists in identifying the polypeptide markers by using proteomic technology such as microsequencing or mass spectrometry. If the proteins are known, we usually obtain the antibodies from commercial sources. If on the other hand they are novel, they are eluted from the 2D gels and used to immunize rabbits and mice for antibody production. In future, phage antibody libraries may prove to be instrumental for the rapid production of a large number of specific antibodies for diagnostic pathology.

As depicted in Fig. 7, there is only one small region in the abnormal urothelium of this particular section (indicated with an arrow) that resembles the staining patterns of the tumor, and that may be considered a putative premalignant lesion.

Spot information: SSP 8509

Name unknown; Mr = 67.5 kDa ; pI = 4.7

Category	Description	Annotation Entry	Link	
difregdiffker	Proteins Differentially Regulated in In Vitro Differentiated Keratinoytes.	Downregulated. E.Olsen et al., unpublished observation.		
distriexker	Distribution in Triton Extracted Keratinocytes.	Triton supernatant. J.E.Celis, unpublished observation.		
expinhelacells	Expressed in Hela Cells.	Expressed. J.E.Celis, unpublihed observation.		
il1alpha	Keratinocyte Proteins Affected by Interleukin 1 Alpha.	Not affected. J.E.Celis, unpublished observation.	Cytokine Explorer	
interfalpha	Proteins Affected by Interferon Alpha.	Not affected. J.E.Celis, unpublished observation.	Signaling pathway	
interfgam	Keratinocyte Proteins Affected by Interferon Gamma.	Not affected. J.E.Celis, unpublished observation.	Cytokine Explorer	Signaling pathway
levelsk14	Protein Levels in SV40 Transformed K14 Keratinocytes.	Ratio transformed/normal proliferating = 0.76. J.E.Celis and E.Olsen, Electrophoresis 15, 309, 1994.		
levfehutis	Radioactive Levels in Fetal Human Tissues.	Adrenal glands = +; brain = +; cerebellum = +;diaphragm = +; ear = +; eye = +; gall bladder = +; hypophysis = +; kidney = +; large intestine = +; liver = +; lung = +; meninges = +; mesonephric tissue = +; pancreas = +; skeletal muscle = +; skin = +; small intestine = +; spleen = +; stomach = +; submandibular gland = +; thymus = +; tongue = +; umbilical cord = +; ureter = +. J.E.Celis and A. Celis, unpublished observations.	Normal Histology	
okadaicacid	Keratinocyte Proteins Affected by Okadaic Acid.	Not affected. O.K.Vintermyr and J.E.Celis, unpublished observation.		
pma	Keratinocyte Proteins Affected by PMA.	Not affected. J.E.Celis, unpublished observation.		

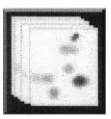
Fig. 4. File for the unknown protein IEF SSP 8509. Only some of the entries available in the Internet are shown.

Other areas of the involved urothelium show immunofluorescence patterns of antigen staining that are clearly different from those observed in the invasive tumor (indicated with small arrows). Clearly, the heterogeneity of the tissue biopsy is such that it would be unwise to use 2D PAGE alone to analyze it in the absence of histological data. The latter may also be true for approaches involving sensitive mRNA based technology.

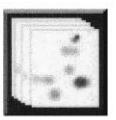
It should be stressed that our main project in bladder cancer centers around transitional cell carcinomas (TCCs), which comprise 95% of all bladder lesions in the western hemisphere. Our goal is to use proteomic technology to define 'normal' and to identify markers for the objective classification of these lesions, as well as to reveal specific proteins in the urine (tumor markers, degradation products of matrix proteins) that may identify patients at risk of developing invasive disease [48]. To support these studies we are building large 2D PAGE databases of tumors and urine (http://biobase.dk/cgi-bin/celis) in an effort to identify signalling pathways and components, that are affected at a given stage of tumor development. In the long run, these studies are expected to provide a global picture of the gene expression changes accompanying the malignant process, and may provide with novel targets for drug discovery.

2.3. The mouse kidney 2D PAGE database

The newborn mouse kidney (non-cultured) 2D PAGE database represents the more recent addition to the set of databanks that we have made available through the Internet (http://biobase.dk/cgi-bin/celis). To date, the database lists 1779 proteins (IEF), of which 223 have been identified using 2D PAGE immunoblotting, Edman degradation of internal peptide sequences, and mass spectrometry. Information categories available so far in this database include cellular localization, knockouts (see Fig. 8), as well as expression in other mouse tissues (brain, bladder, cerebellum, ear, heart, large intestine, liver, lung, optical nerve, skeletal muscle, skin, small intestine, spleen, stomach, tail, thymus and tongue). To facilitate the comparison of protein expression profiles between various tissues we have included at the Web site a gallery of



Search Results



Search criteria:

Proteins found by this search:

Click any protein to view its name on the gel.

- **(M 1) Human keratinocytes-IEF Database** (proteins found = 14)

Protein name	SSP	Mr	pI
Ornithine decarboxylase variant.	5430	53.3	5.32
Unknown	6302	50.1	5.24
Unknown	6307	50.7	5.33
Keratin 7 variant.	6321	51.5	5.38
Unknown	6401	54.6	5.21
Unknown	6405	52.5	5.25
Untransformed steroid receptor complex (p56 component).	6413	53.9	5.44
Untransformed steroid receptor complex (p56 component).	6414	53.4	5.35
Unknown	6427	54.7	5.04
Unknown	7313	51.4	5.19
Unknown	7314	51.0	5.03
Unknown	7315	51.3	5.11
Regulatory subunit of cAMP-dependent protein kinase (type 1).	7411	53.3	5.16
Regulatory subunit of cAMP-dependent protein kinase (type 1).	7412	53.5	5.07

This page produced using custom web software based on PDQUEST.

Gels analyzed using PDQUEST from Bio-Rad (formerly PDI).

Fig. 5. Proteins in the keratinocyte database with M_rs between 50 and 55 kDa and p*I*s between 5.0 and 5.5.

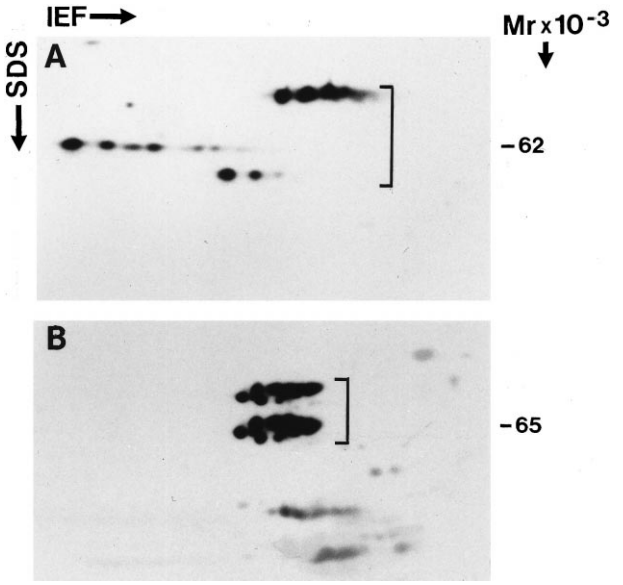


Fig. 6. Proteins detected in 2D blots of whole keratinocyte proteins reacted with antibodies against (A) SHC and (B) cortactin, and detected using the ECL procedure.

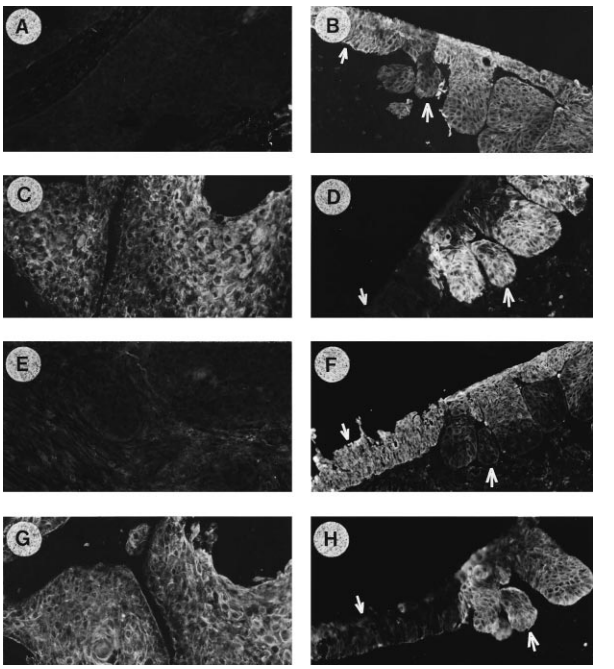


Fig. 7. Immunofluorescence pictures of formaldehyde fixed cryostat sections of SCC 798-1 (A, C, E, and G) and of the involved urothelium from the same patient (B, D, F and H), reacted with antibodies against keratin 19 (A, B), keratin 14 (C, D), keratin 13 (E, F) and the basal antigen BG3C8 (G, H). From Celis et al. [52].

reference 2D gel autoradiograms. In general, the protein expression profiles of tissues are qualitative rather similar (compare Fig. 9A, kidney and Fig. 9B, lung), and as in the case of human cells and tissues one would expect that a large fraction of all expressed proteins are shared by all cell types [31].

Clearly, the database is still in its infancy but in due course it is expected to become a major resource for the analysis of transgenics [49–51], as one may envisage compensatory changes in protein expression taking place as a result of the loss or overexpression of a given gene product.

3. Perspective

A great deal of work is being devoted to the development of high throughput technology to resolve, visualize, quantitate and characterize proteins in an effort to facilitate the study of diseases and to expedite the process of drug discovery [2]. Today, the rapid development in high throughput proteomics is being catalyzed in part by the pharmaceutical industry, who sees a role for this technology in tuning up the various steps involved in drug discovery which include target identification and validation, drug discovery (efficacy, selectivity and mode of action), in vivo properties as well as clinical trials.

With the 2D PAGE technology available today it is rather straightforward to reveal major protein expression changes associated with a given disease, or with the response to a given effector (growth and differentiation factors, drugs,

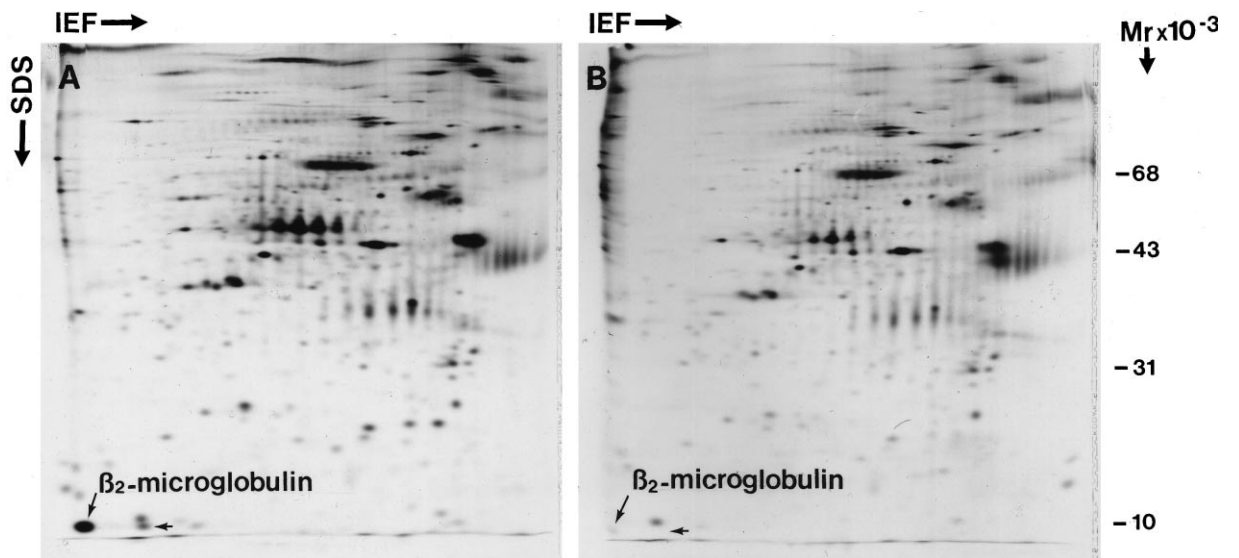


Fig. 8. IEF 2D gels of [^{35}S]methionine labeled externalized proteins from primary adult kidney cells (overnight culture) from (A) normal and (B) β_2 microglobulin knockout mice. The arrowhead indicates the position of an acidic variant of β_2 microglobulin. From Celis et al. [10].

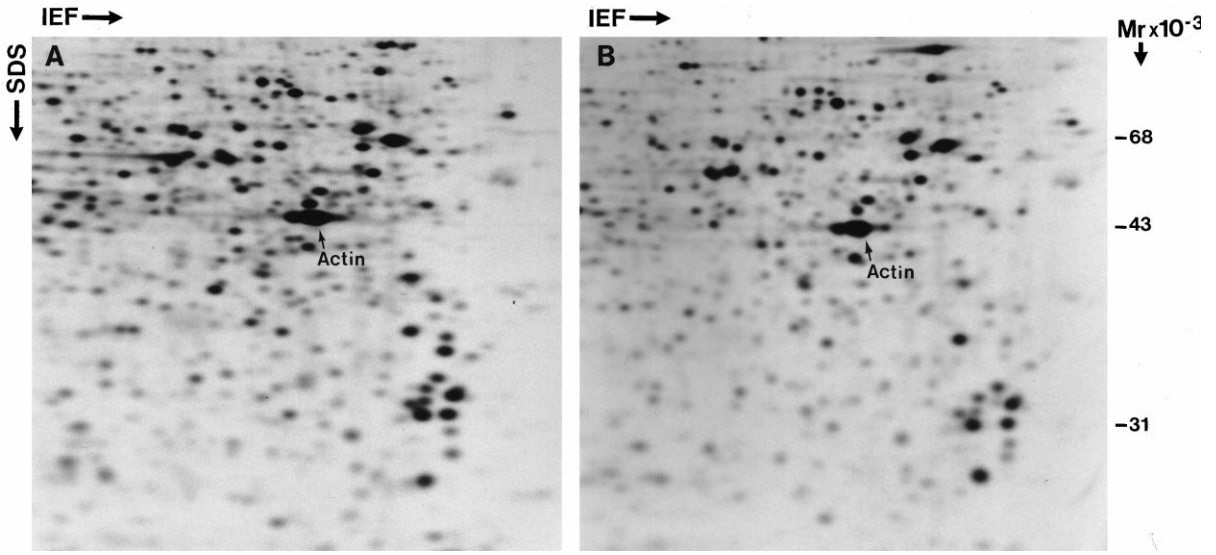


Fig. 9. [^{35}S]Methionine labeled proteins from newborn mouse kidney (A) and (B) lung. Only a fraction of the IEF gels are shown.

etc.), but the question still remains which of these changes are primarily associated with the disease or the effector. Thus, having at hand additional information to narrow the possibilities seems crucial at this time, and it is in this context that comprehensive proteomic 2D PAGE databases containing a critical mass of information may prove to be invaluable in rapidly sorting out the putative targets of interest. Making a 2D PAGE database of a cell line is not a problem as the cultures are rather homogeneous, but building databanks of normal tissues and disease biopsy material is a different proposition as the heterogeneity of the samples requires first hand knowledge of the biology of the tissue in question, as well as probes (antibodies) to identify the various cell types that contribute to the overall protein expression pattern. We believe that in the near future phage antibody libraries ([53], and references therein) may become instrumental for the rapid generation of specific antibodies against proteins of interest revealed by the 2D PAGE technology, and in combination, they may offer a unique platform for building up the field of molecular pathology.

Clearly, there are many limitations associated with building up large 2D PAGE databases that go beyond the heterogeneity of the samples. These include the detection of very low abundance polypeptides, the resolution of very basic and high molecular weight proteins, as well as the lack of satisfactory quantitation procedures for the analysis of all the proteins resolved in the gels. Some of these shortcomings are being addressed by developing larger and better gel separation systems, more sensitive dyes for protein detection, by the analysis of fractionated protein samples and subcellular fractions, as well as by the development of better software for spot detection and quantitation. In addition, there are efforts being devoted to standardizing the gel technology in general, so that information gathered in any given laboratory can be readily used by others in the scientific community.

Acknowledgements: We would like to thank Ariana Celis, Bodil Basse, Jette B. Lauridsen, Inger Andersen, Gitte Ratz, Bente Hein and Pamela Celis for expert technical assistance and Torben F. Ørntorft and Hans Wolf for stimulating discussions. We are indebted to J. Vandekerckhove's group for support throughout the years. The work was supported by grants from the Danish Biotechnology Programme, the Danish Cancer Society and the Danish Centre for Molecular Gerontology.

References

[1] Wilkins, K. (1997) BioRadiations 99, 4–6.
[2] Celis, J.E. (1998) Drug Discovery Today 3, 193–195.
[3] Celis, J.E. (Ed.) (1996) Electrophoresis 17, 1653–1798.
[4] Humphery-Smith, I. (Ed.) (1997) Electrophoresis 18, 1205–1498.
[5] Huber, L.A. (Ed.) (1997) Electrophoresis 18, 2505–2700.
[6] Anderson, L. and Seilhamer, J. (1997) Electrophoresis 18, 533–537.
[7] O'Farrell, P. (1975) J. Biol. Chem. 250, 4007–4021.
[8] O'Farrell, P.Z., Goodman, H.M. and O'Farrell, P. (1977) Cell 12, 1133–1141.
[9] Klose, J. (1975) Humangenetik 26, 231–243.
[10] Celis, J.E., Ratz, G., Basse, B., Lauridsen, J.B., Celis, A., Jensen, N.A. and Gromov, P. (1997) in: Cell Biology: A Laboratory Handbook (Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K. and Small, J.V., Eds.), Vol. IV, pp. 375–386, Academic Press, San Diego, CA.
[11] Celis, J.E., Gromov, P., Østergaard, M., Gromova, I., Madsen, P., Varmark, H., Jensen, N.A. and Rasmussen, H.H. (1998) Prog. Biophys. Mol. Biol. (in press).
[12] Bjellqvist, B., Ek, K., Righetti, P.G., Gianazza, E., Görg, A., Westermeier, R. and Postel, W. (1982) J. Biochem. Biophys. Methods 6, 317–339.
[13] Görg, A., Pistel, W. and Günther, S. (1988) Electrophoresis 9, 531–546.
[14] Righetti, P.G. (1990) Immobolized pH Gradients: Theory and Methodology, Elsevier, Amsterdam.
[15] Miller, M.J. (1989) Adv. Electrophoresis 3, 182–217.
[16] Dunn, M.J. (1992) in: Computers in Biochemistry: A Practical Approach (Bryce, C.F.A., Ed.), pp. 215–245, IRL Press, Oxford.
[17] Vandekerckhove, J. and Rasmussen, H.H. (1997) in: Cell Biology: A Laboratory Handbook (Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K. and Small, J.V., Eds.), Vol. IV, pp. 505–513, Academic Press, San Diego, CA.
[18] Nika, H. and Aebersold, R. (1997) in: Cell Biology: A Laboratory Handbook (Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K. and Small, J.V., Eds.), Vol. IV, pp. 514–525, Academic Press, San Diego, CA.
[19] Wilm, M., Shevchenko, Y., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, L. and Mann, M. (1996) Nature 379, 466–469.
[20] Yates, J.R. (1996) Methods Enzymol. 271, 351–377.
[21] Patterson, S.D., Thomas, D. and Bradshaw, R.A. (1996) Electrophoresis 17, 877–891.

[22] Mann, M. (1996) Trends Biochem. Sci. 21, 494–495.
[23] Pappin, D.J. (1997) Methods Mol. Biol. 64, 165–173.
[24] Roepstorff, P. (1997) Curr. Opin. Biotechnol. 8, 6–13.
[25] Celis, J.E., Rasmussen, H.H., Leffers, H., Madsen, P., Homoré, B., Gesser, B., Dejgaard, K. and Vandekerckhove, J. (1991) FASEB J. 5, 2200–2208.
[26] Celis, J.E., Rasmussen, H.H., Olsen, E., Madsen, P., Leffers, H., Honoré, B., Dejgaard, K., Gromov, P., Hoffmann, H.J., Nielsen, M., Vassilev, A., Vintermyr, O., Hao, J., Celis, A., Basse, B., Lauridsen, J.B., Ratz, G.P., Andersen, A.H., Walbum, E., Kjærgaard, I., Puype, M., Van Damme, J. and Vandekerckhove, J. (1993) Electrophoresis 14, 1091–1198.
[27] Appel, R.D., Vargas, J.R., Palagi, P.M., Walther, D. and Hochstrasser, D.F. (1997) Electrophoresis 18, 2735–2748.
[28] Celis, J.E., Gromov, P., Østergaard, M., Madsen, P., Honoré, B., Dejgaard, K., Olsen, E., Vorum, H., Kristensen, D.B., Gromova, I., Haunsø, A., Van Damme, J., Puype, M., Vandekerckhove, J. and Rasmussen, H.H. (1996) FEBS Lett. 398, 129–134.
[29] Miklos, G.L. and Rubin, G.M. (1996) Cell 86, 521–529.
[30] Oliver, S.G. (1996) Nature 379, 597–600.
[31] Celis, J.E., Rasmussen, H.H., Gromov, P., Olsen, E., Madsen, P., Honoré, B., Dejgaard, K., Vorum, H., Kristensen, D.B., Østergaard, M., Jensen, N.A., Celis, A., Basse, B., Lauridsen, J., Ratx, G.P., Walbum, E., Andersen, I., Puype, M., Van Damme, J. and Vandekerckhove, J. (1995) Electrophoresis 16, 2177–2240.
[32] Celis, J.E., Lauridsen, J.B. and Basse, B. (1997) in: Cell Biology: A Laboratory Handbook (Celis, J.E., Carter, N., Hunter, T., Shotton, D., Simons, K. and Small, J.V., Eds.), Vol. I, pp. 429–437, Academic Press, San Diego, CA.
[33] Bauw, G., Van Damme, J., Puype, M., Vandekerckhove, J., Gesser, B., Ratz, G.P., Lauridsen, J.B. and Celis, J.E. (1989) Proc. Natl. Acad. Sci. USA 86, 7701–7705.
[34] Rasmussen, H.H., Van Damme, J., Bauw, G., Puype, M., Celis, J.E. and Vandekerckhove, J. (1991) in: Methods in Protein Sequence Analysis, (Jörnvall, H., Höög, J.-O. and Gustavsson, A.-M., Eds.), pp. 103–104, Birkhäuser, Basel.
[35] Rasmussen, H.H., Mörtz, E., Mann, M., Roepstorff, P. and Celis, J.E. (1994) Electrophoresis 15, 406–416.
[36] Bravo, R. and Celis, J.E. (1982) Clin. Chem. 28, 766–781.
[37] Bravo, R., Fey, S.J., Bellatin, J., Larsen, P.M., Arevalo, J. and Celis, J.E. (1981) Exp. Cell. Res. 136, 311–319.
[38] Celis, J.E. and Celis, A. (1985) Proc. Natl. Acad. Sci. USA 82, 3262–3266.
[39] Gromov, P. and Celis, J.E. (1997) in: Cell Biology: A Laboratory Handbook (Celis, J.E., Ed.), Vol. IV, pp. 409–417, Academic Press, San Diego, CA.
[40] Koss, L.G. (1975) Atlas of Tumor Pathology, Fasc. 11, Armed Forces Institute of Pathology, Washington, DC.
[41] Olsen, S. (1984) Tumors of the Kidney and the Urinary Tract. Colour Atlas and Textbook, Munksgaard, Copenhagen.
[42] Celis, J.E., Rasmussen, H.H., Vorum, H., Madsen, P., Honore, B., Wolf, H. and Ørntoft, T.F. (1996) J. Urol. 155, 2105–2112.
[43] Østergaard, M., Rasmussen, H.H., Nielsen, H.V., Vorum, H., Ørntoft, T.F., Wolf, H. and Celis, J.E. (1997) Cancer Res. 57, 4111–4117.
[44] Madsen, P., Rasmussen, H.H., Leffers, H., Honoré, B., Dejgaard, K., Olsen, E., Kiil, J., Walbum, E., Andersen, A.H., Basse, B., Lauridsen, J.B., Ratz, G.P., Celis, A., Vandekerckhove, J. and Celis, J.E. (1991) J. Invest. Dermatol. 97, 701–712.
[45] Jinquan, T., Vorum, H., Larsen, C.G., Madsen, P., Rasmussen, H.H., Gesser, B., Etzerodt, M., Honoré, B., Celis, J.E. and Thestrup-Pedersen, K. (1996) J. Invest. Dermatol. 107, 5–10.
[46] Slaughter, D.P., Southwick, H.W. and Smejkal, W. (1953) Cancer 6, 963–968.
[47] Pallesen, G., Nielsen, S. and Celis, J.E. (1987) Histopathology 11, 591–601.
[48] Celis, J.E., Østergaard, M., Basse, B., Celis, A., Lauridsen, J.B., Ratz, G.P., Andersen, I., Hein, B., Wolf, H., Ørntoft, T.F. and Rasmussen, H.H. (1996) Cancer Res. 56, 4782–4790.
[49] Jensen, N.A., Pedersen, K.M., Celis, J.E. and West, M.J. (1998) J. Clin. Invest. 101, 1292–1299.
[50] Jensen, N.A., Pedersen, K.M., Celis, J.E. and West, M.J. (1998) Oncogene 16 (16), 2123–2129.
[51] Jensen, N.A. and Celis, J.E. (1998) Electrophoresis (in press).
[52] Celis, J.E., Østergaard, M., Celis, P., Rasmussen, H.H. and Wolf, H. (1998) BioRadiations (in press).
[53] McCafferty, J., Hoogenboom, H.R. and Chiswell, D.J. (Eds.) (1996) Antibody Engineering, Oxford University Press, New York.