QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Biology: the basis for smart proteomic approaches to protein analysis

Helen Kim, PhD

Department of Pharmacology & Toxicology

University of Alabama at Birmingham

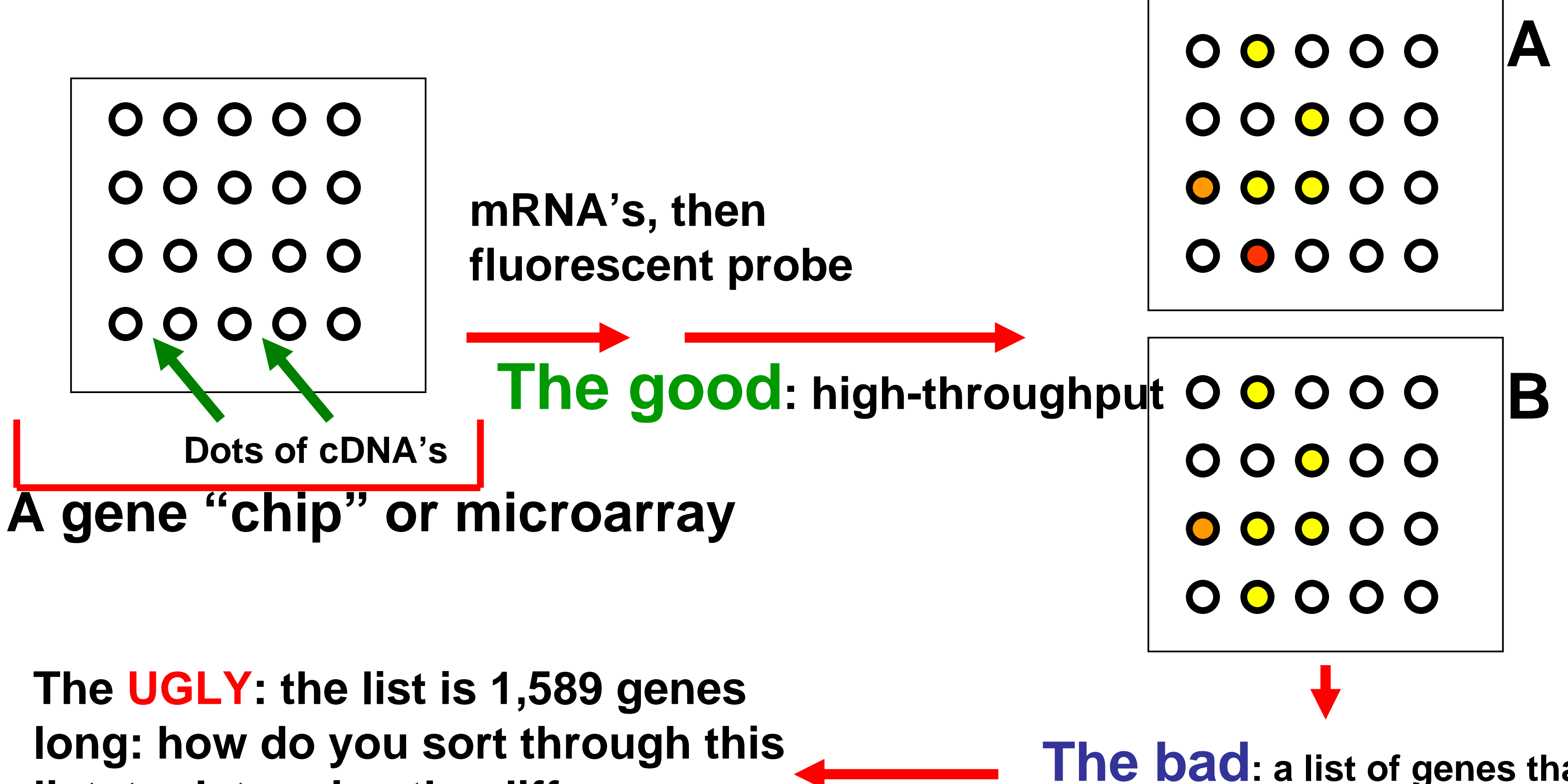
UAB Comprehensive Cancer Center Proteomic/Mass Spectrometry Shared Facility

Helenkim@uab.edu

What *is* proteomics?

- **Genomics: study of genomes of a cell or organism**
- **Proteomics:**
  - **Original definition: study of the proteins encoded by the genome of a biological sample**
  - **Current definition: study of *the whole protein complement* of a biological sample (cell, tissue, animal, biological fluid [urine, serum])**
    - **More and more, going to analysis of <u>sub-proteomes</u>**
  - **Usually involves high resolution separation of polypeptides at front-end, followed by mass spectrometry identification and analysis**

Genomics and gene microarrays: the good, the bad, and the be-careful-what-you-ask-for.



A gene "chip" or microarray

The good: high-throughput

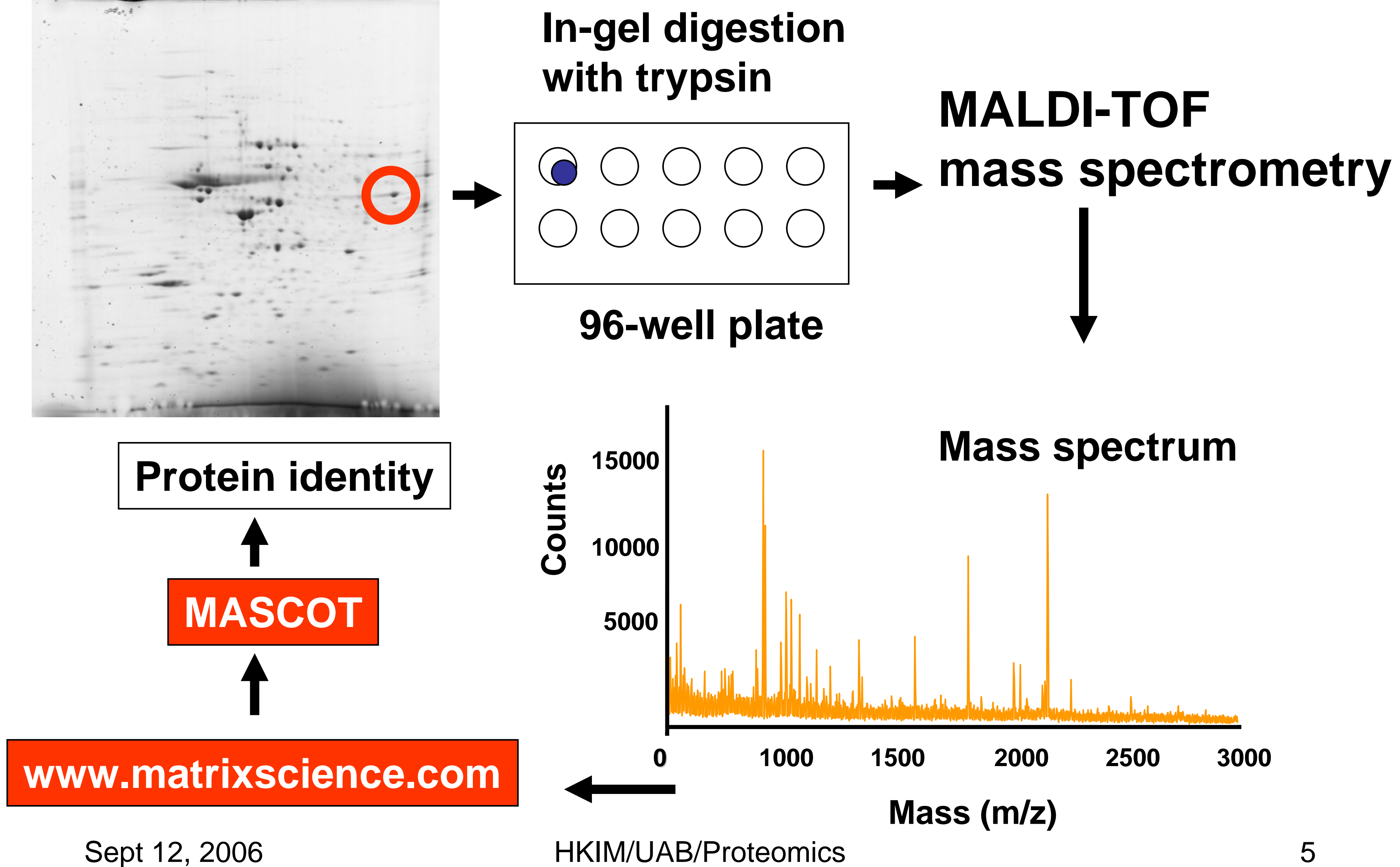
The bad: a list of genes that are "up" or "down"-regulated.

The **UGLY**: the list is 1,589 genes long: how do you sort through this list, to determine the differences relevant to the disease

If microarrays can identify hundreds, thousands of gene differences between health and disease, why do we need to bother with proteins and proteomics?

- **When you don’t have a clue;**
- **When you have only a very small clue;**
  i.e. you’ve done a microarray experiment, and you have a list of 3,284 genes that are differentially regulated in your system;
- **When you knock out a gene (and hence a protein) that you’re convinced is essential for life and health, and the animal pees as usual.**

Standard 2D gel proteomics: cut out gel spots, do MALDI, get protein ID's



Hypothesis generation or “discovery”:

Global comparison of kidney proteins from healthy mice vs from a transgenic model of kidney disease;

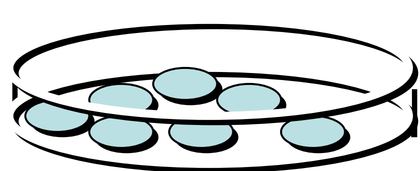
identification of protein differences can be a first step in hypothesis generation of the molecular events that lead to the disease

Hypothesis confirmation:

2D gel comparison of a membrane subfraction from kidneys from healthy mice vs from a transgenic model of kidney disease;

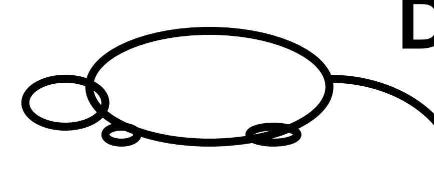
---confirm that differences in protein abundance or protein-protein interactions in the kidney membrane are correlated with disease.

Sample preparation for 2DE:



Harvest, rinse, and pellet cells;

or



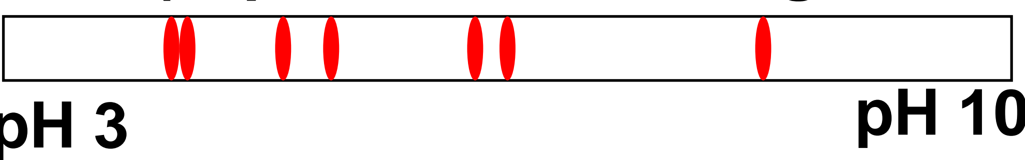
Dissect out tissue, organ, or fluids;

- **Homogenize/lyse in buffer that *unfolds the proteins w/o adding or disturbing the charges*:**
  - **High urea usually 5-8 M---unfolds the protein**
  - **1-4% detergent--keeps hydrophobic components in solution**
  - **Beta-mercaptoethanol or other reductant to reduce disulfide bonds**
  - **Inhibitors: of proteases, kinases, & phosphatases**
- **Clarify by centrifugation to get rid of insoluble/particulate matter;**
- **Protein assay to know how much and how concentrated**

2-D electrophoresis: more than the sum of its parts

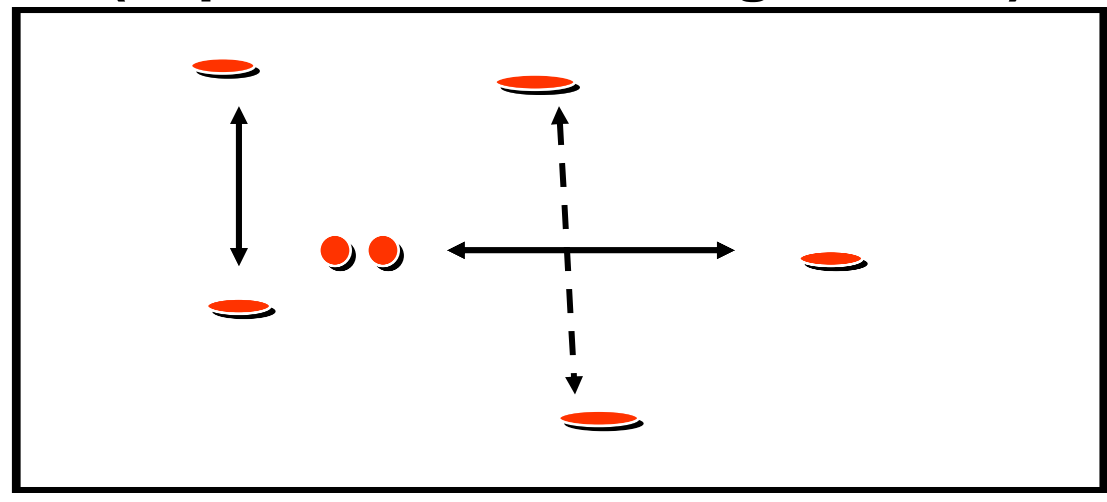
- **1st dimension: Isoelectric focusing**

  (separation according to charge)



- **2nd dimension: (SDS)-PAGE**

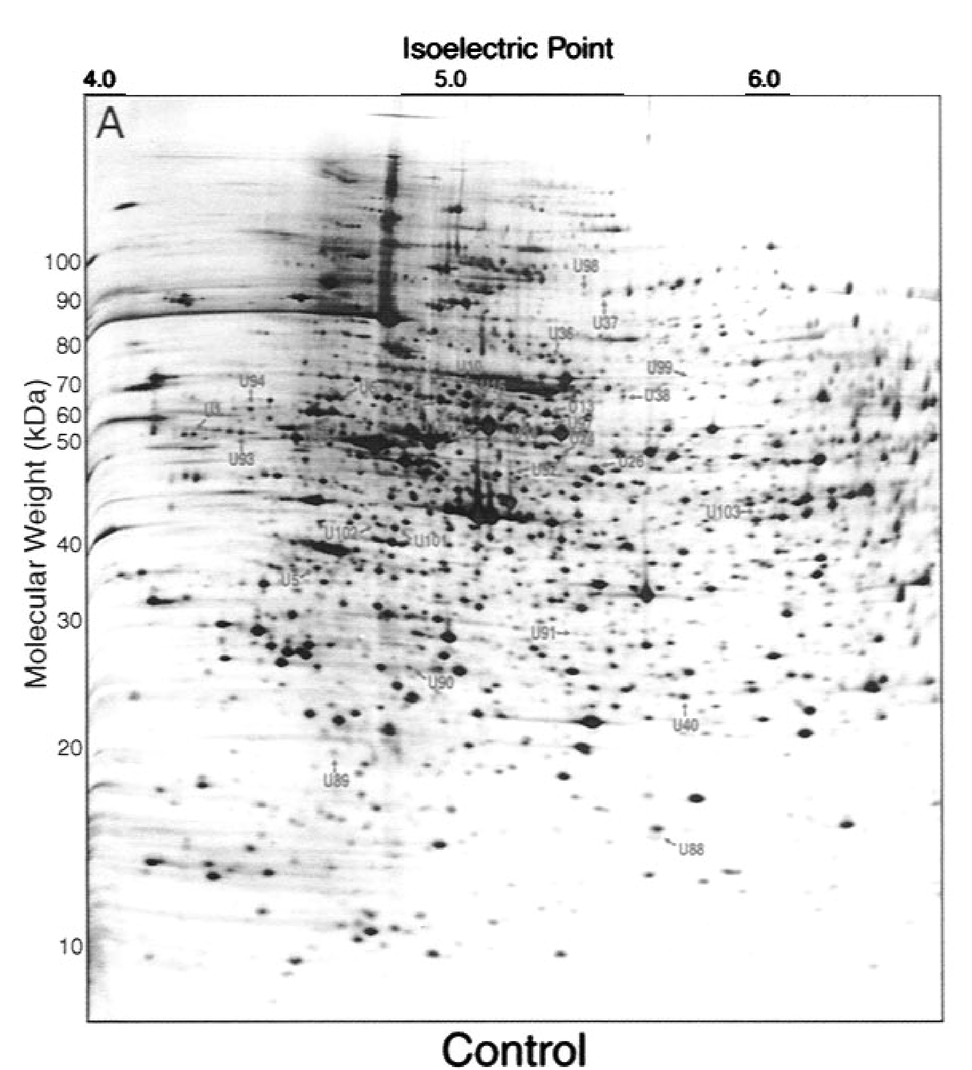
  (separation according to size)



High m.w.

Low m.w.

A real 2-D gel



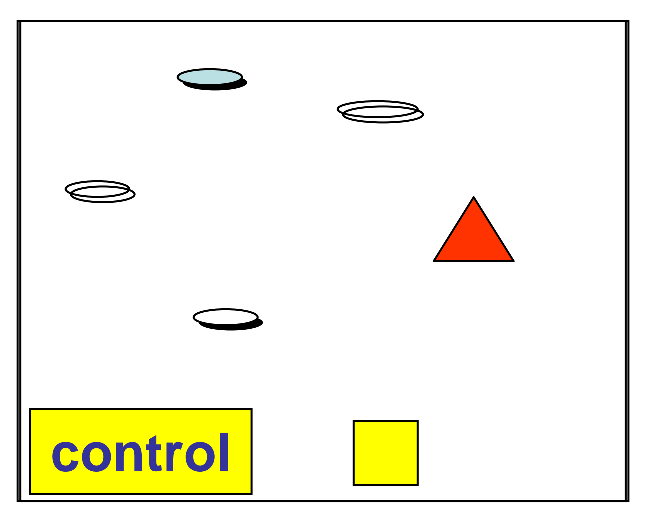
Find this and other 2-D gels at http:www.expasy.org

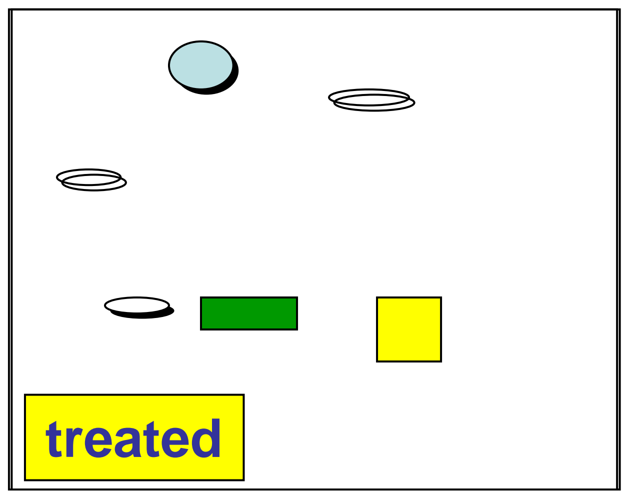
Lewis et al., [2000] Molec. Cell, 6)

(from Natalie Ahn's lab)

The pattern itself is information;
---a protein migrates to the same position under given electrophoretic conditions;
---therefore, a change in position has biological meaning.

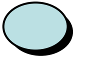
Essential part of 2-D gel proteomics: Image analysis, because your eyes are only so good.



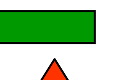


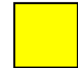
Either manually or with software: "compare" the images.

Types of information:

 ------ Upregulation of gene

------ Posttranslational modificatiion

 ------ Downregulation of gene

 ------ Aberrant processing

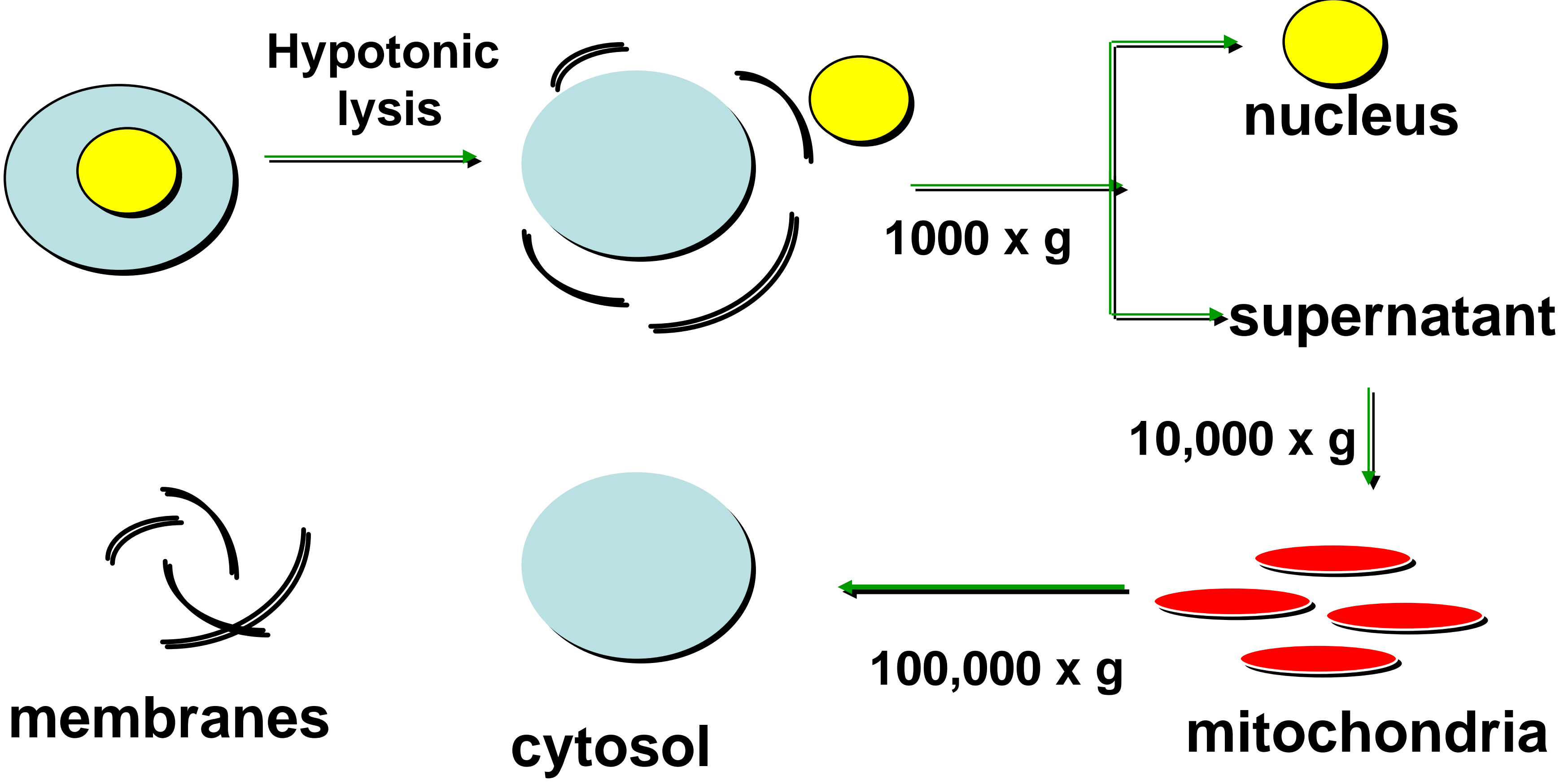
In practice, how do we deal with all those spots?

The smart answer:
as little as possible

Translation: the key to successful proteomics is working with as little complexity as possible

Various ways to reduce proteome complexity:

Subcellular fractionation by differential centrifugation

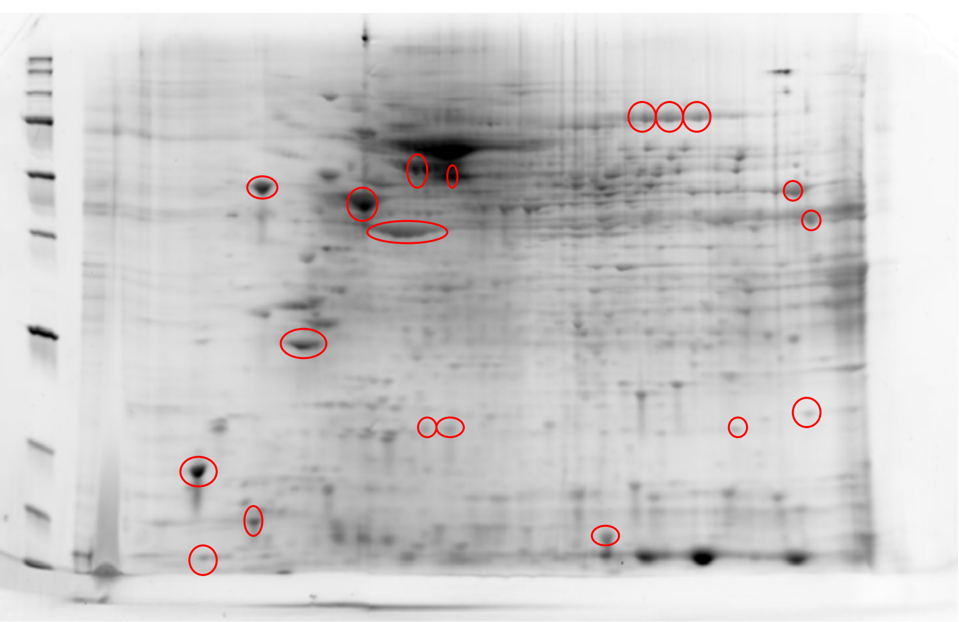


The good news: subcellular proteomes are readily "catalogued."

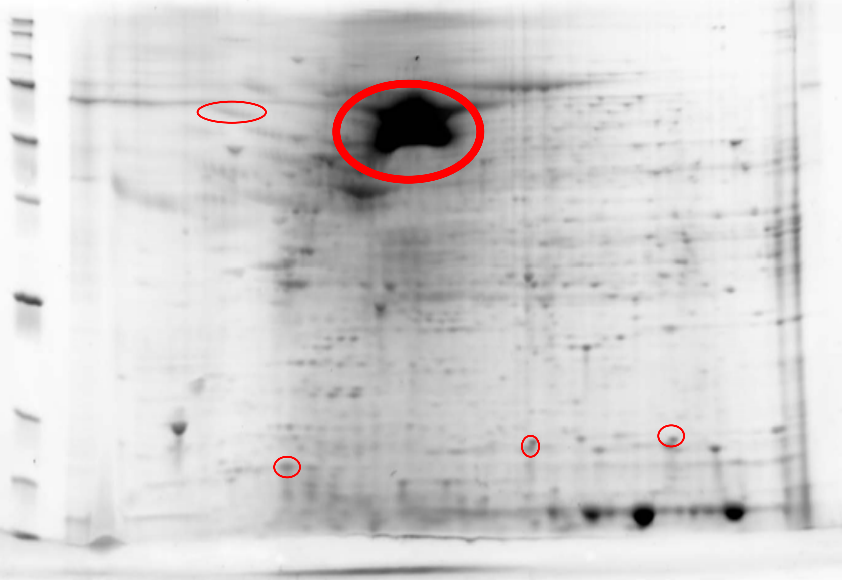
Compartment	# polypeptides in SWISSPROT as of 2000
Mitochondria (1000/cell)	269------5% of total
Lysosome (400/cell)	50------1% of total
Peroxisome	35------0.6%
ER and Golgi apparatus	157------3%
Nuclei (5% cell volume)	964-----17%
Others (cytosol, membrane)	4228----75%
	total: 5703

(Jung et al. [2000] Electrophoresis)
Note date of article: this is already old data

Preliminary 2D gel analysis of PKD-relevant protein differences



WT kidney



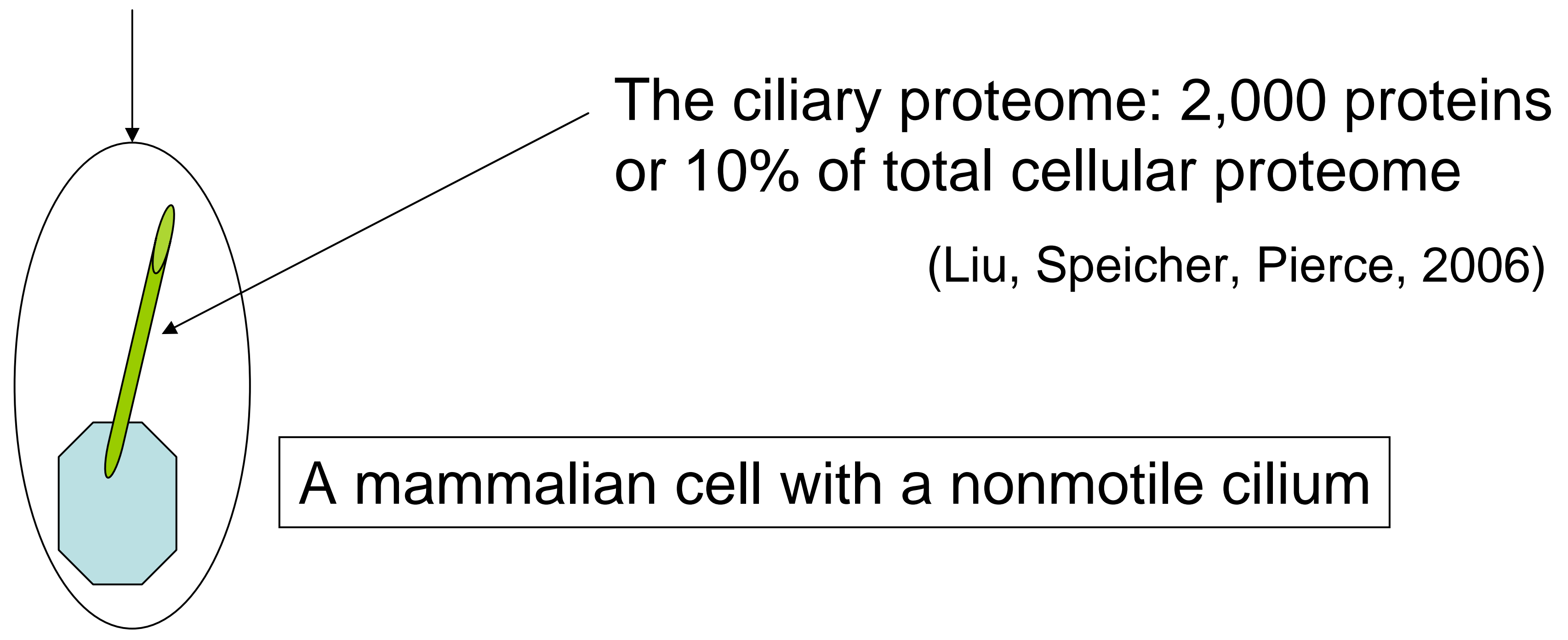
PKD kidney

proteins increased in WT

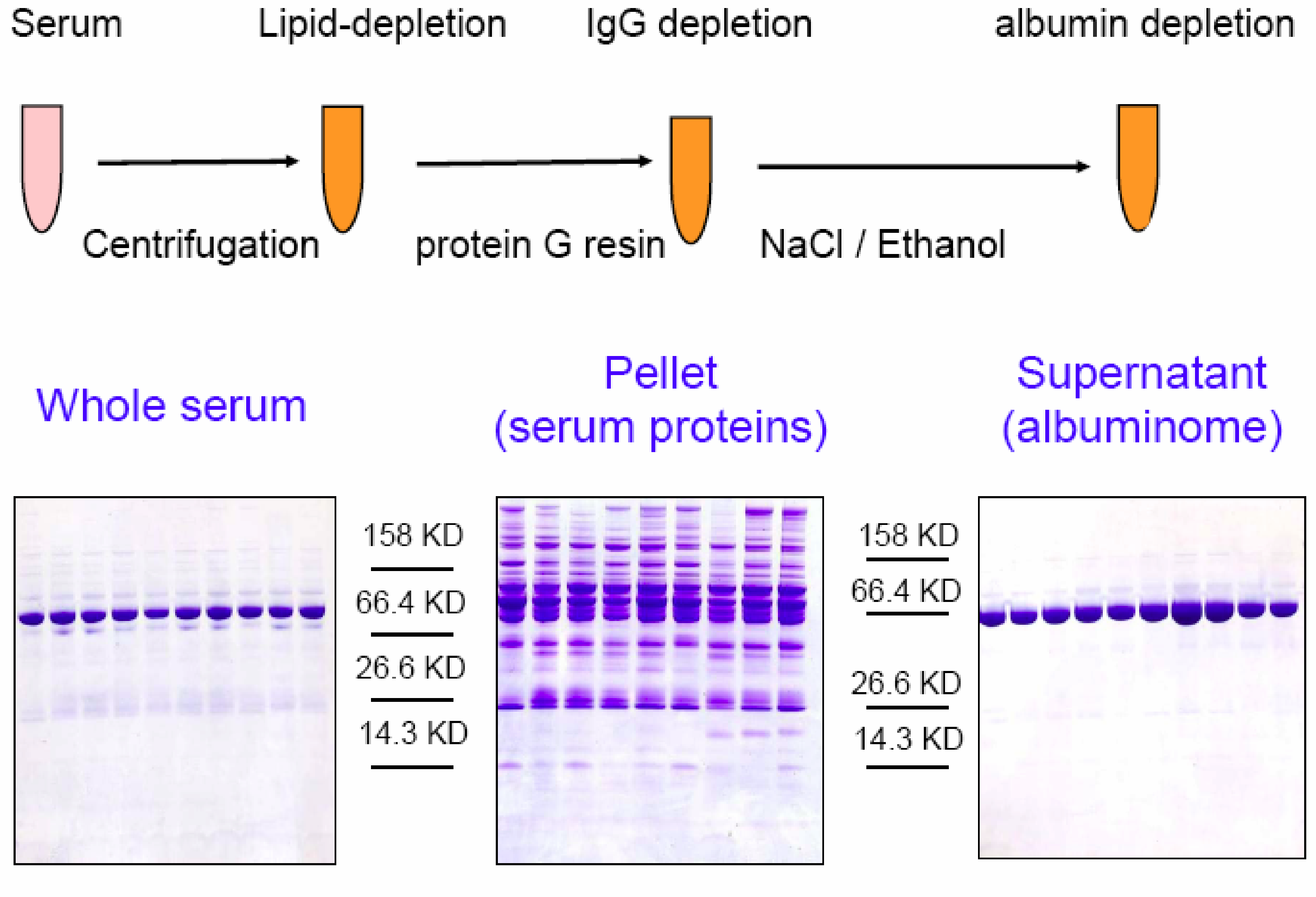
proteins increased in PKD

Smart proteomic approach to study of primary sensory cilia: isolate the cilia!!

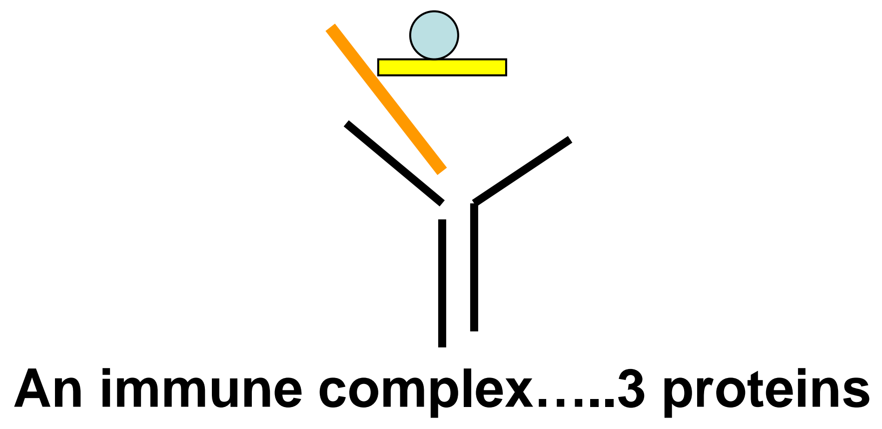
A whole cellular proteome: 20,000 proteins minimum



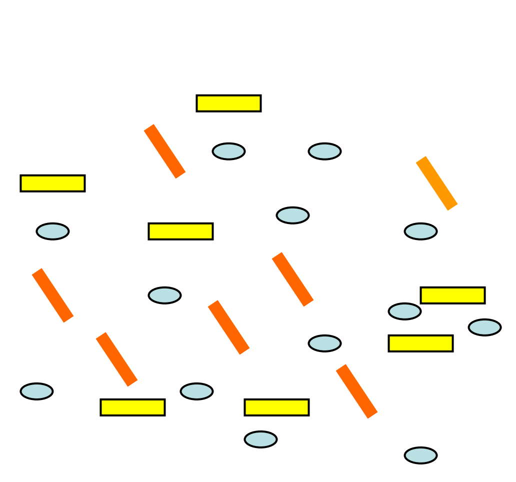
Reduce proteome complexity by depletion of the most abundant proteins



Deal with proteome complexity by increasing biological specificity, and therefore information



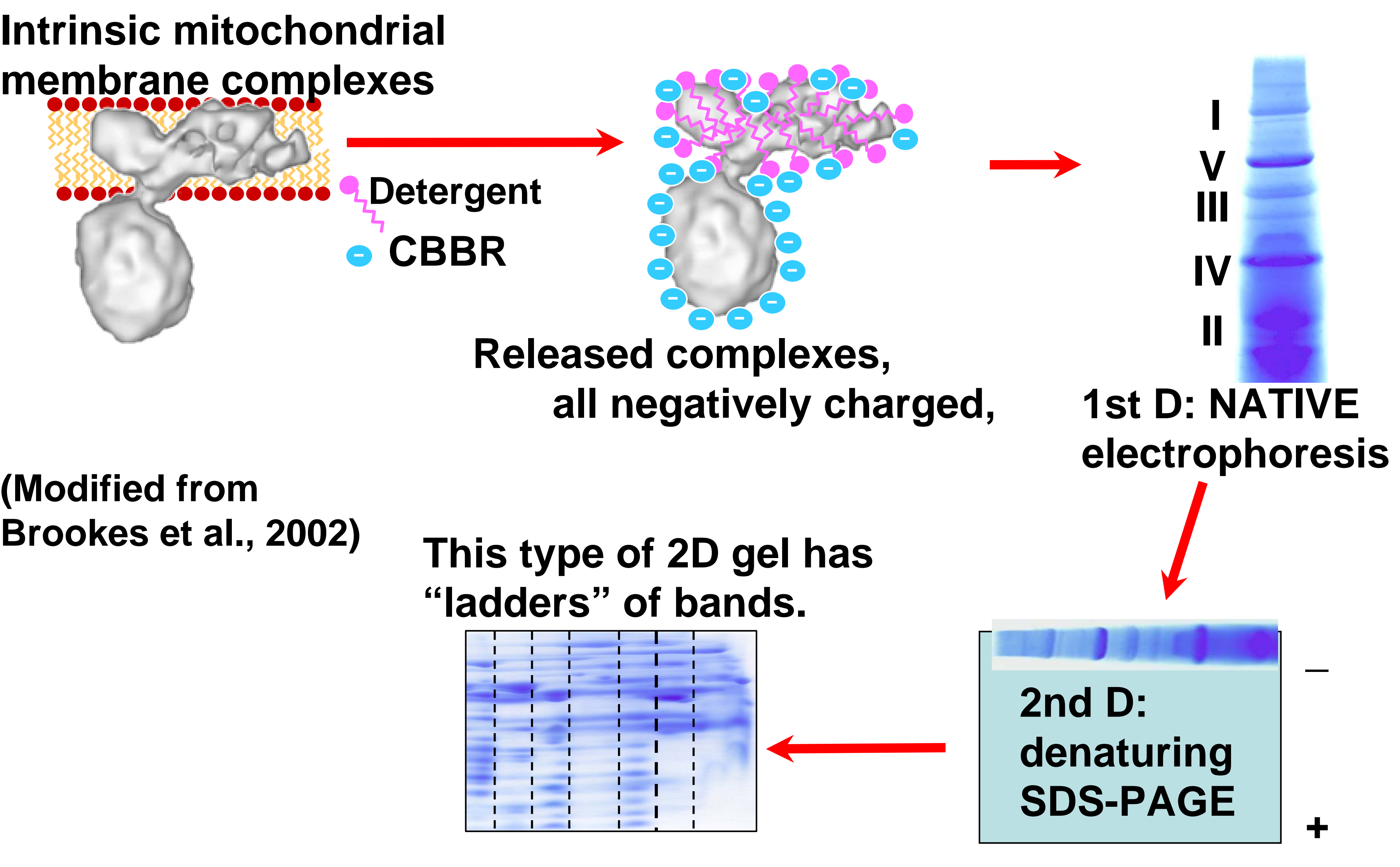
An immune complex…..3 proteins



A cell lysate…….6,973 proteins

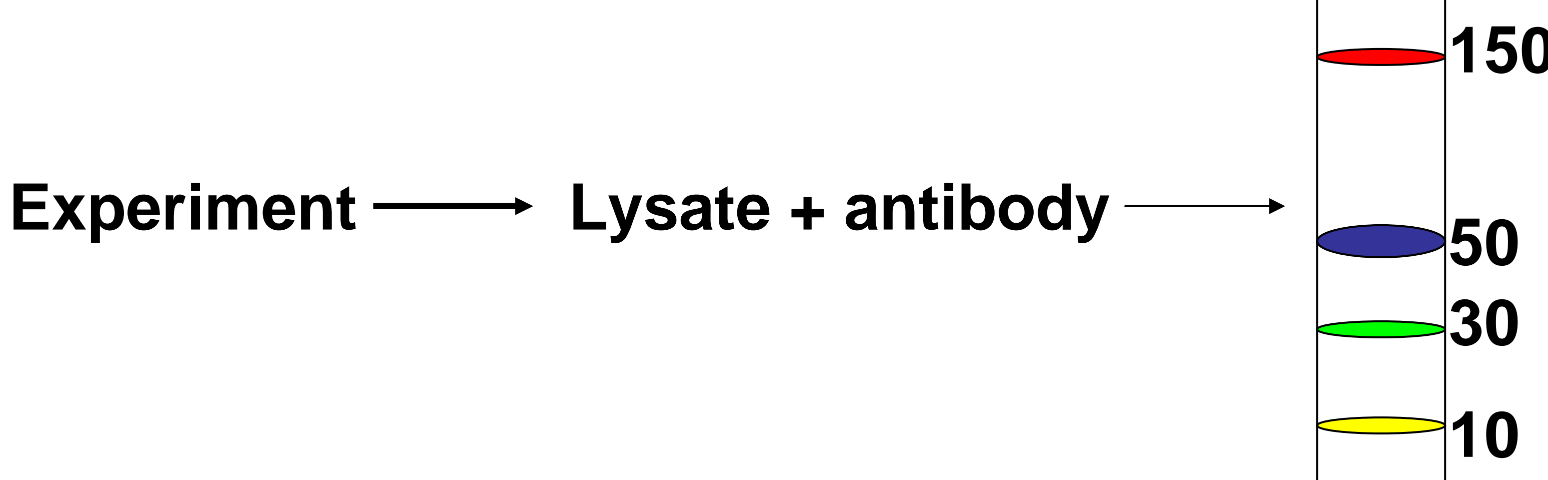
Which would you rather try and analyze on a 2D gel?

2D-Blue-Native gels: for protein-protein interactions in membrane environments



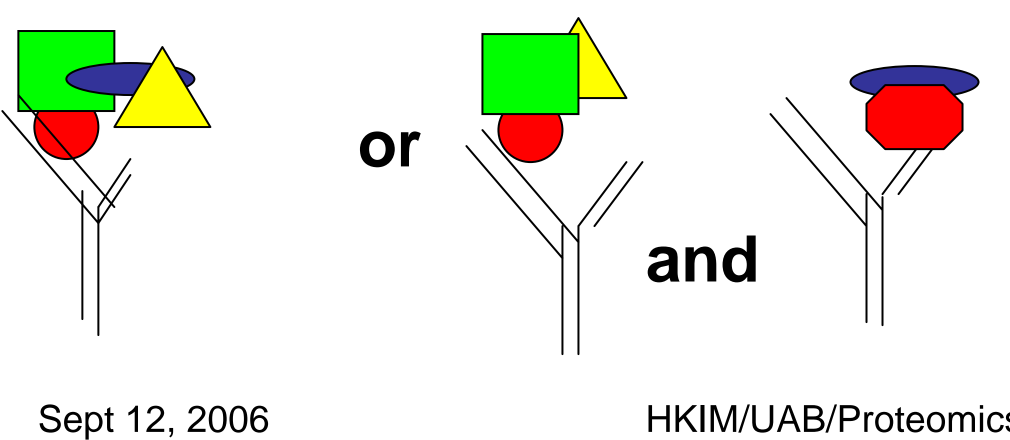
Aside from membrane complexes, BN gels can also address:

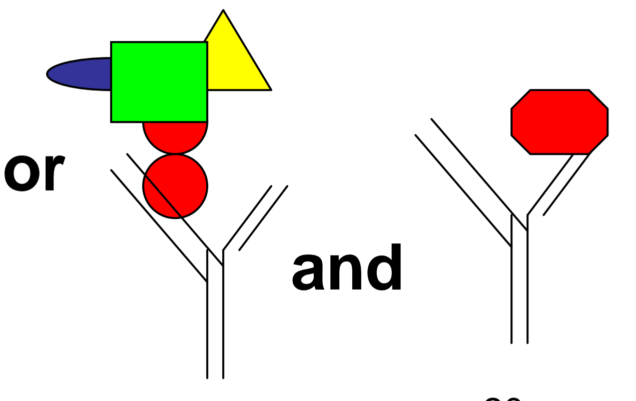
Which proteins are interacting with which?

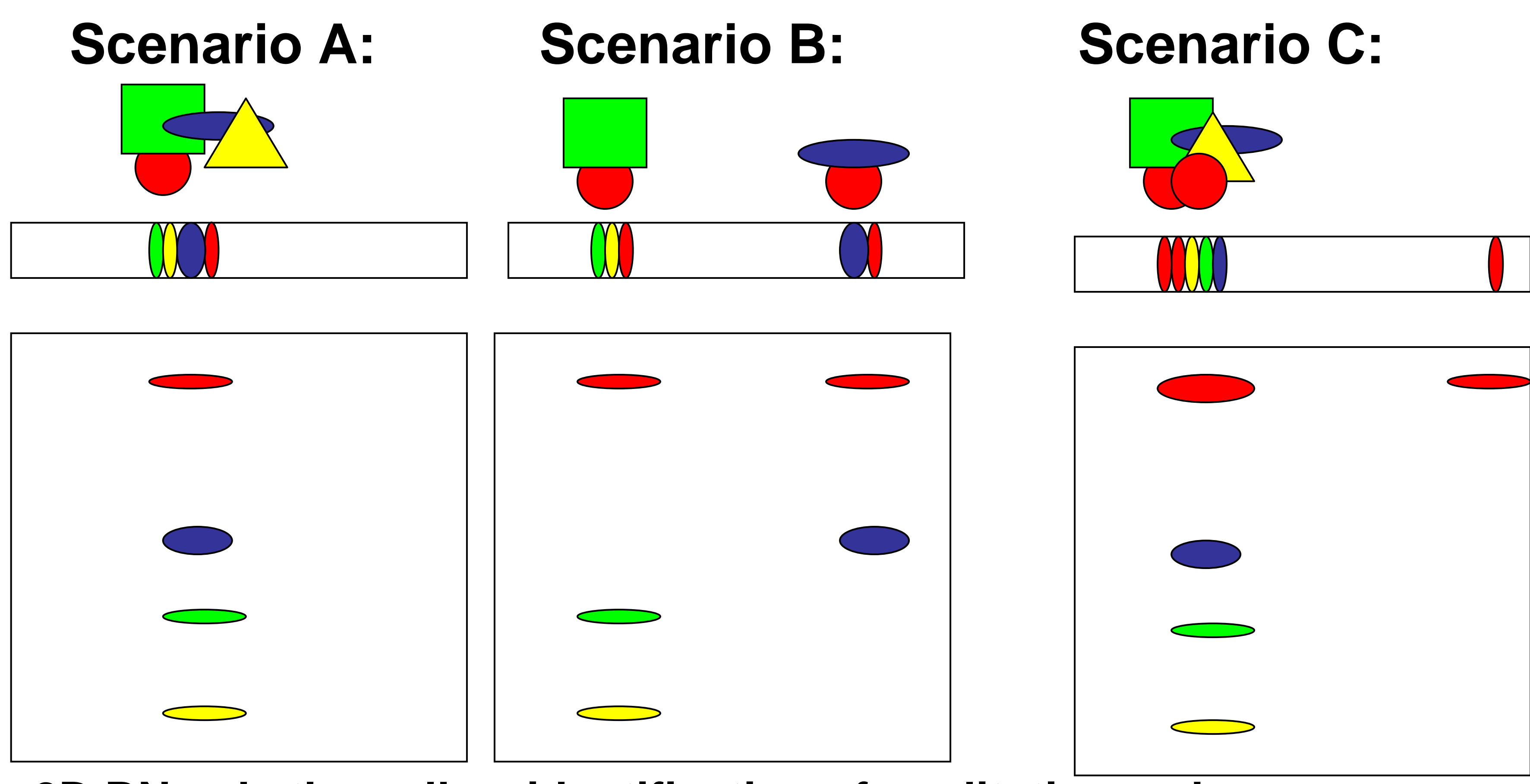


The immunoprecipitate

But who's binding to whom:

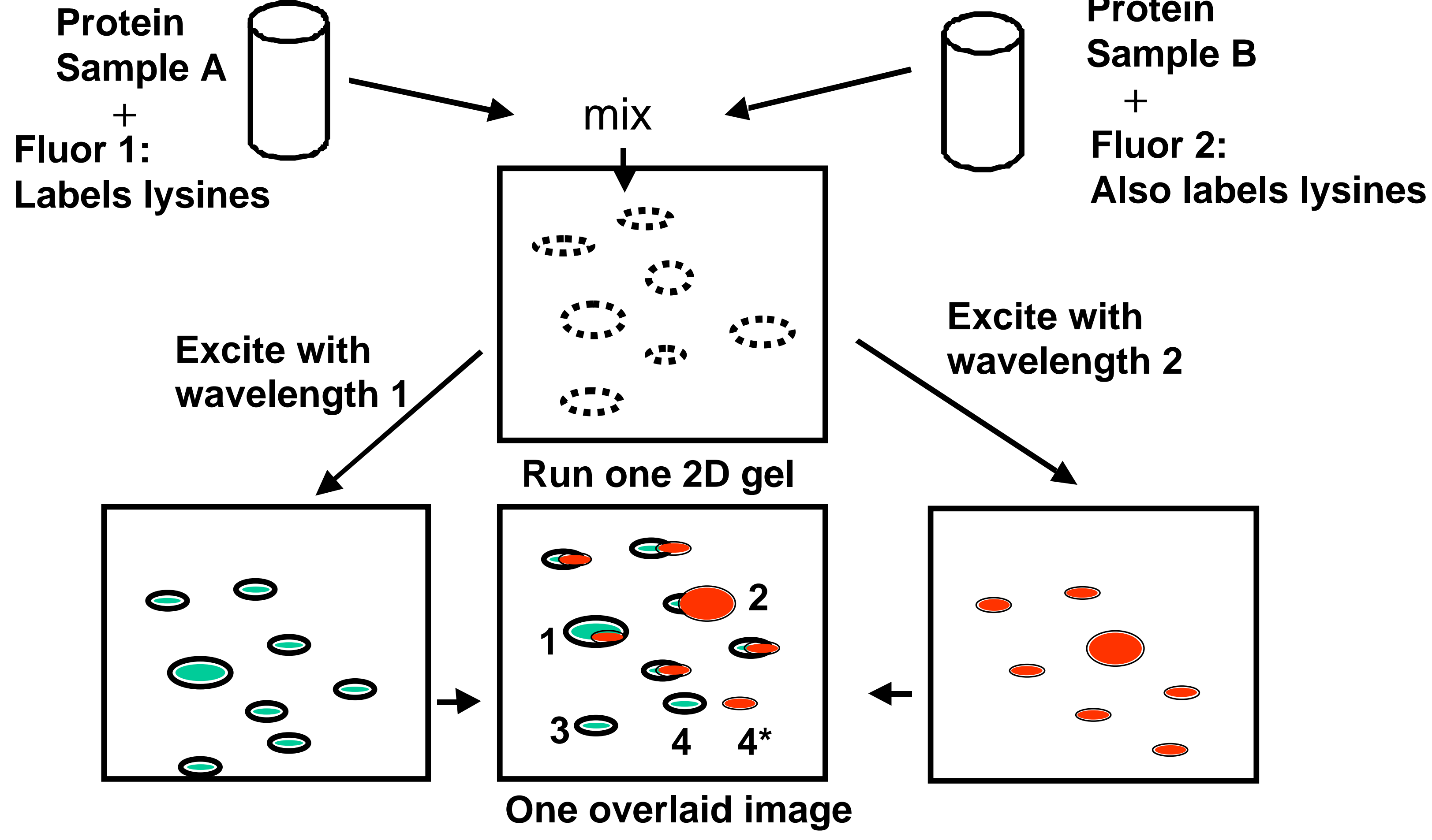






2D BN gels thus allow identification of qualitative and quantitative aspects of protein interactions that might be difficult to assess by any other method

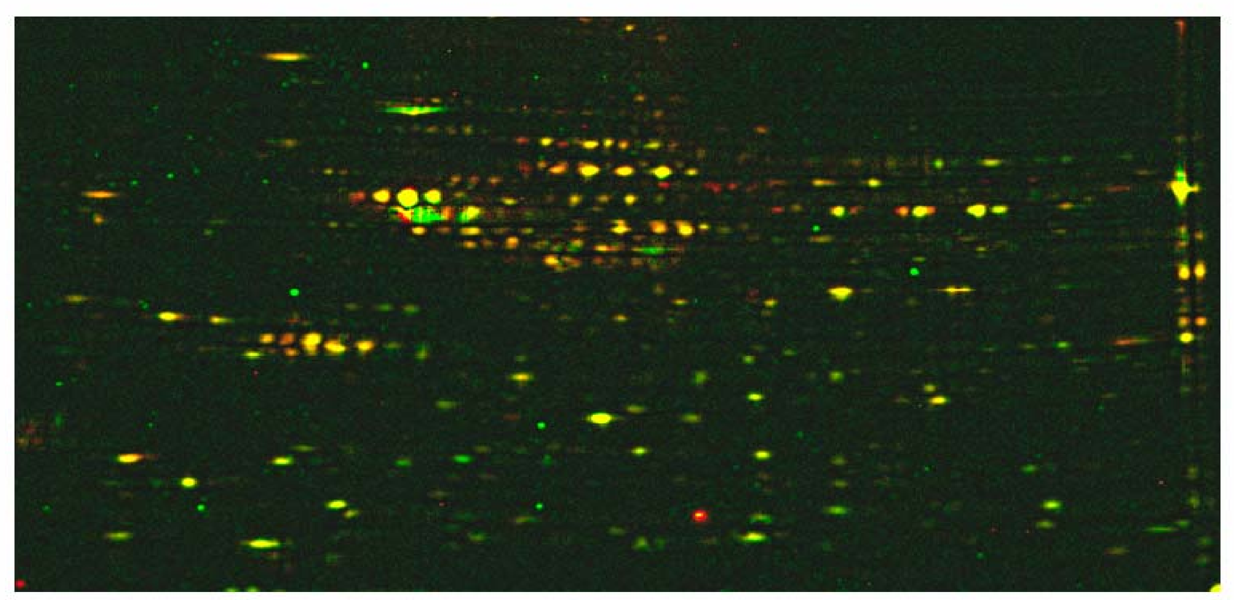
Difference gel electrophoresis (DIGE); where a protein migrates with itself---- can really be valuable.



Single biggest bottleneck, even with DIGE: even commercially made gels (1st dimension, 2nd dimension) are not perfect. 2D displays of proteins across multiple gels stil landmark spots matched across gels.

Dealing with proteome complexity

Differential protein labeling with Cy3 and Cy5
Superimposed images from the same gel
of normal and cancer cell lines from the breast



Visually: 100% green spots are specific to normal cells, 100% red spots are specific to cancer cells, degrees of yellow-orange indicate differential expression

Image analysis software uses pixel intensities, does not need colors.

Make use of databases and the internet:

I. Check existing databases and web-links:

www.expasy.org

many are annotated

helpful links:

proteomics tools

II. Keep up with the literature/ competition:

Electrophoresis

Proteomics

Molecular & Cellular Proteomics

J. Proteome Research

III. Use genomics information when available:

The polypeptide sequence (from the cDNA) can predict electrophoretic parameters-- m.w. & pI;

helpful in setting up 2D gel conditions

Future directions in intact protein analysis approaches

I. DIGE and Cy-dye labelling will enhance 2D gel analysis of complex proteomes;
II. Subcellular fractionation will be "re-invented" and applied to reduce proteome complexity;
III. While automated 2D LC/LC-MS/MS can be used for "discovery" and "profiling", every spot resolved from another on a 2D gel is a discovery;
IV. 2D gel positional information, without protein identity, is information itself.
V. 2D gels will become more valuable also for validation of results obtained by profiling approaches;
VI. Once the nature of a protein difference is defined, a gel approach, either IEF or SDS-PAGE, could be used in a high-throughput manner for diagnostic or screening purposes.

Final thoughts

- What proteomics technology gives back is like any other analytical approach: it’s as good as what you put in;
- Be mindful of distinguishing between low abundance proteins vs low level contaminants;
- Keep in mind “conventional” approaches like Western blotting to validate proteomic results;
- Purify, purify, purify before running any proteomic experiment.
- Formulating a hypothesis forces you to incorporate biological information that can enhance proteomic analysis.