Mar 23,2010
UAB BMG744

The best proteomics: combination of good gels, addressing quality control, and mining the data

Helen Kim, Ph.D.

Dept of Pharmacology & Toxicology
University of Alabama at Birmingham
helenkim@uab.edu
205-934-3880

OUTLINE

I. **Proteomics analysis of actions of a dietary supplement, grape seed extract (GSE), in the brain;**

II. **DIGE analysis of protein differences in rat tissues at different developmental stages to determine basis for cancer risk**

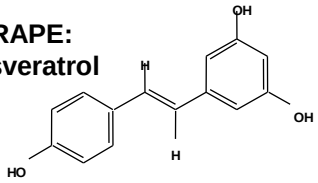
Our principal goal: to understand the molecular basis of human chronic conditions/diseases, to develop prevention or therapies.

Strategy: a *proteomics* approach

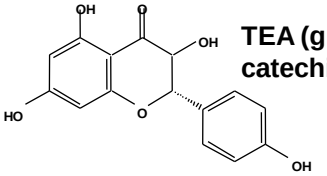
Hypothesis: Actions of "beneficial" agents such as dietary anti-oxidants in normal and disease tissue will reveal subproteomes of proteins "at risk" for disease-relevant changes.

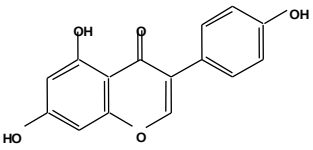
POLYPHENOLS: similar structures among themselves, and with 17β-estradiol

GRAPE: resveratrol

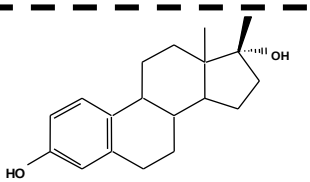


TEA (grapes): catechin





SOY: genistein



17β-estradiol

RATIONALE for predicting grape seed extract (GSE) polyphenols will have neuroprotective actions:

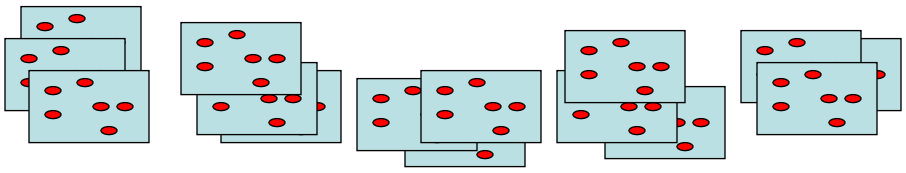
- Joseph et al., 1999;
  - blueberry extract supplement protected against age-related cognitive impairment
- Pan et al., 2000;
  - soy isoflavones protected against ovariectomy-induced cognitive dysfunction
- Peng et al., 2005;
  - our own studies showed GSE enhanced cognition in estrogen-depleted (ovx'd) rats;

The experiment, and the dataset:

AIN-76A diet; normal adult rats (n = 5)

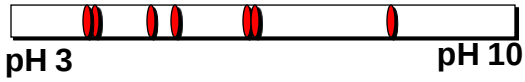
AIN-76A + 5% GSE; normal adult rats (n = 5)

whole brain homogenates, triplicate 2D gels

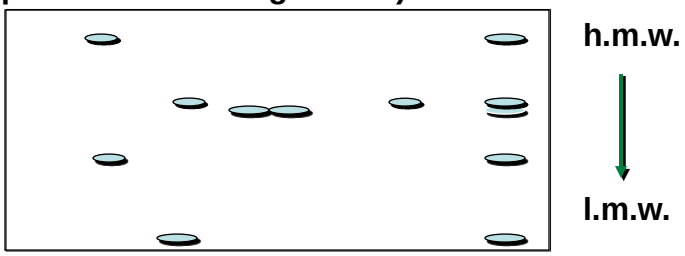


Basics of 2-D electrophoresis

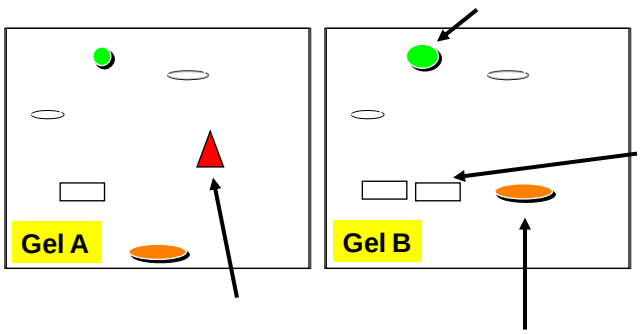
- **1st dimension: isoelectric focusing on flat plastic strips** ***containing immobilized pH gradients (IPG)*** **(separation according to charge)**



- **2nd dimension: (SDS)-polyacrylamide gel electrophoresis (separation according to size)**

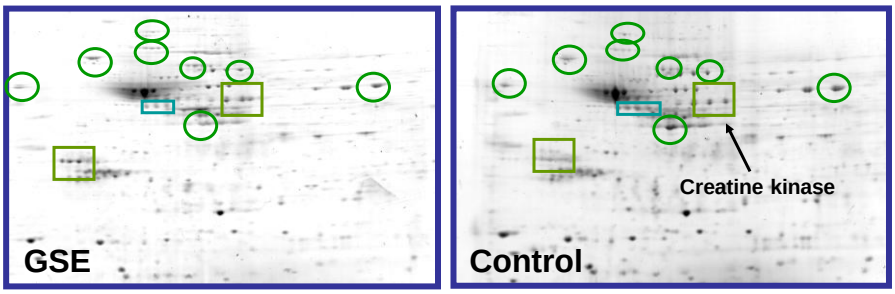


Critical part of 2-D gel proteomics: Image analysis



The arrows indicate the types of information *suggested* by image analysis data.

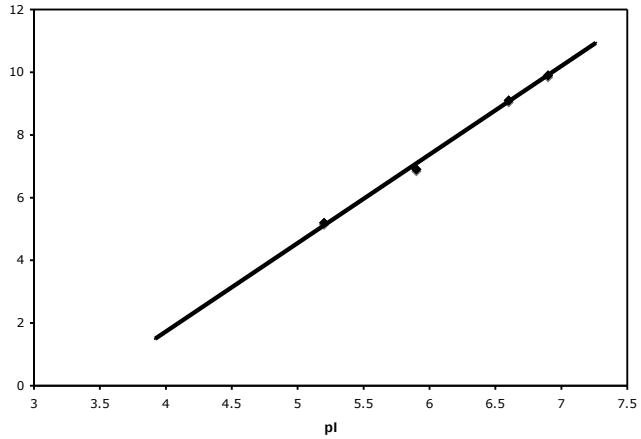
Image analysis indicated spots that differed between the two sets of gels;



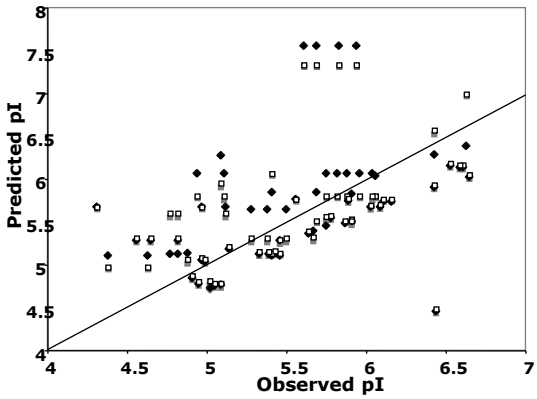
○ Different in intensity

□ Different in position (variability)

How do we know the IPG strips work?

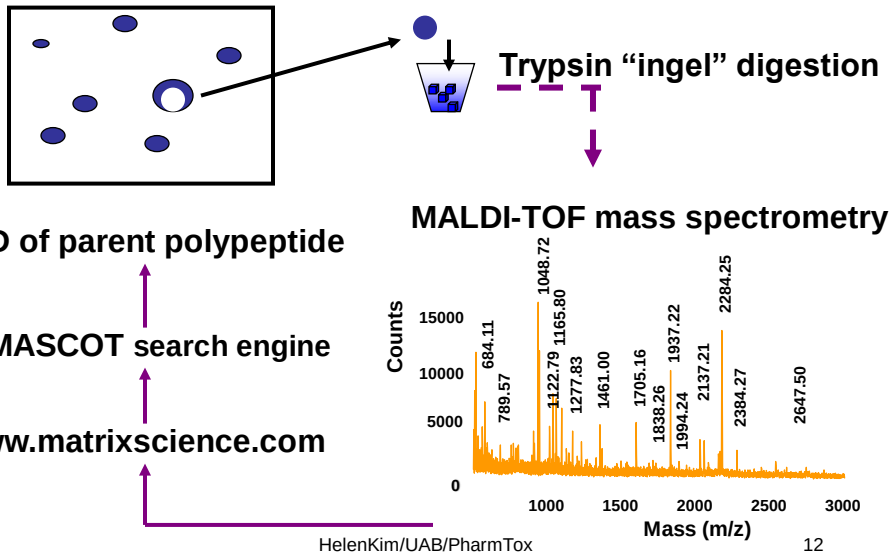


Most proteins focus near but not at their predicted isoelectric point. Why?

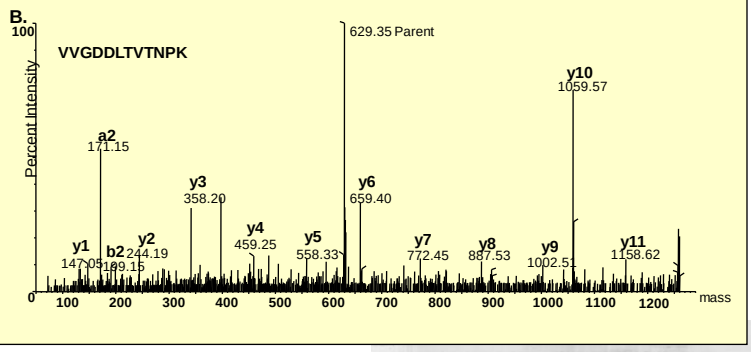


What would it mean if ALL the proteins focussed higher or lower than predicted pIs.

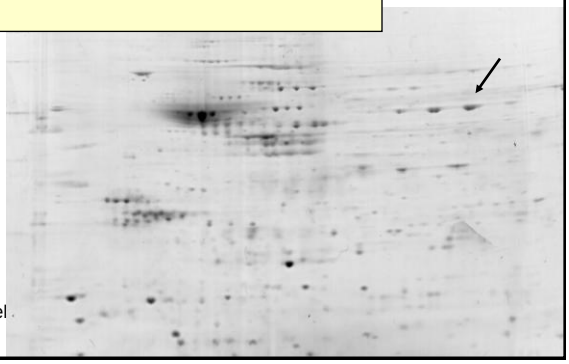
Standard followup procedure once "significant" spots are determined by image and statistical analysis: Peptide mass fingerprint analysis by MALDI-TOF MS



(Kim et al., in Tollefsboll, Methods in Aging Research, in press)



α-enolase was confirmed by LC-MS/MS



Hel

Database of protein differences in GSE vs CT brains

Protein Name	#matched pep	Accession #	MOWSE	Obs m.w.	Pred m.w.	Obs pI	Pred pI	Nature of change in GSE brains
Mitochondrial matrix proteinprecursorP60	10	P19227	*1.26E+04	64900	60956	5.6	5.9	+1.5
Creatine Kinase BB chain	12	P07335	*1.66E+05	45600	42712	5.45	5.3	+1.52 Translocation to Acidic pH
Actin	8	P10365	*2.18E+05	42000	41636	5.3	5.4	Less complex
GFAP	20	P47819	*9.67E+09	49000	49943	5.4	5.3	- 1.6
14-3-3 epsilon	10	P42655	*1.41E+09	31900	29174	4.49	4.6	- 2.1
Alpha Enolase	9	P04764	*6.64E+05	46000	46985	6.0	6.2	Less complex
Gamma Enolase	10	P07323	95	47000	47111	5.12	5.03	Less complex
RIKEN cDNA (NM 025994)	9	NP080270	169 95	26000 26000	25084 25084	5.0 5.1	5.0	-1.56
HSC-70	12	gi4103877	110	70321	42455	5.9	6.64	+1.63
HSC-71	16	gi123644	105	70386	71195	5.43	5.49	+1.91
Neurofilament L Triplet protein	14	gi13929098	120	61025	61298	4.61	4.63	+1.63
Neurofilament M triplet protein	19	gi8393823	153	95086	95591	4.75	4.76	+1.73
Vimentin	10	gi202368	93	53600	53641	5.09	5.06	-1.52

(Deshane et al., 2004. J. Agric. Food Chem.)

Initial conclusion: GSE is neuroprotective, since its effects on proteins are counter to the directions of change for the same proteins in disease.

Mitochondrial matrix protein precursor HSP-60 ↑
14-3-3 epsilon protein ↓
Heat shock cognate 70 ↑
Heat shock cognate 71 ↑
Neurofilament triplet protein L ↑
Neurofilament triplet protein M ↑
GFAP ↓
Vimentin ↓
Creatine kinase Brain isoform ↑
Gamma enolase
Alpha enolase
Actin
RIKEN cDNA NM 025944

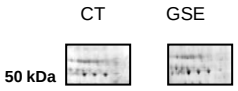
GREEN: folding/ stress response
PURPLE : energy
BLUE: new
ORANGE: cytoskeleton

(Schonberger et al., 2001)
(Tilleman et al., 2002a)
(Tilleman et al., 2002b)

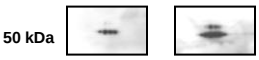
(Deshane et al., 2004. J. Agric. Food Chem.)

Western blot analysis of 2D gel image and statistical analysis

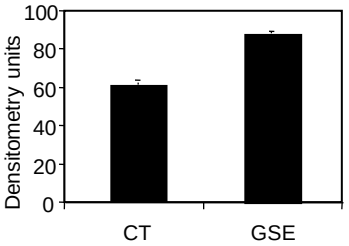
A. Stained gel for HSP-60



B. Western Blots



C. Quantitative Densitometry

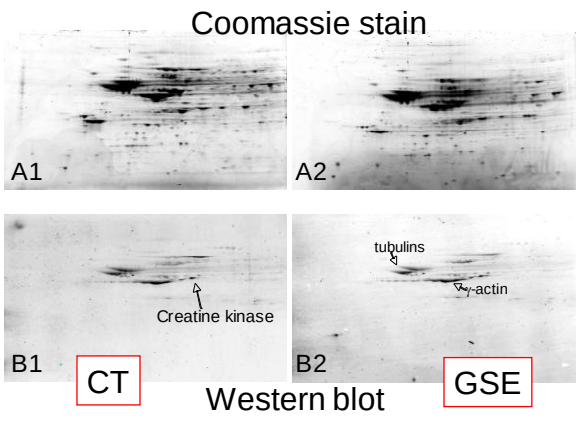


Validation of protein identifications and quantitations

Protein	LC-MS/MS	Western blot
CK-BB	+	+
Hsp60	+	+
GFAP	--	
Actin	--	+
NFL-M	+	--
α-enolase	+	--
γ-enolase	+	--
Hsc70	--	--
Hsc71	--	--
14-3-3e	--	+
NFL-L	--	+

(Kim et al., 2005, in Luo and Packer, Oxidative Stress & Neurodegeneration)

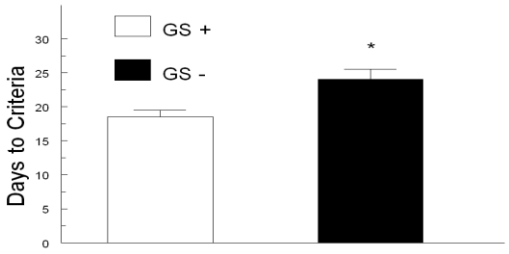
2D gel proteomics of actions of GSE on protein oxidations



Use chemistry to study subproteomes, such as oxidations

QuickTime™ and a
decompressor
are needed to see this picture.

The strongest validation of proteomic data comes from completely different experimental procedures that usually test for function.



In this case, GSE fed to young female rats enhanced learning and memory over those fed control diets. This data was consistent with GSE having neuroprotective actions.

(Peng et al., 2005)

DIGE (difference gel electrophoresis): powerful, but also requires addressing of quality control issues to ensure data is robust.

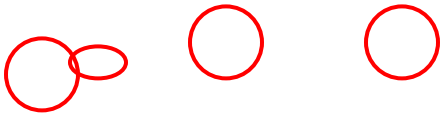
One issue addressed here: does the same sample bind the two main cy-dyes, cy3 and cy5, equivalently?

How do you deal with running multiple samples on 2D gels as objectively as possible

QuickTime™ and a
decompressor
are needed to see this picture.

Grid assigns random pairs of samples per gel;
In this experiment, the different days were swapped to make sure that there wasn't preferential dye binding of one day over the other.

Real data from a DIGE experiment



QuickTime™ and a
decompressor
are needed to see this picture.

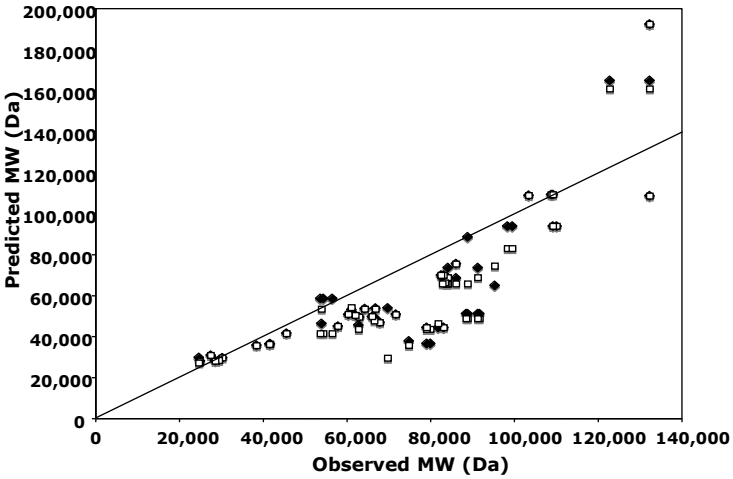
(from Kim et al., submitted)

Image analysis enables “picking” of gel spots for MS analysis

QuickTime™ and a
decompressor
are needed to see this picture.

In this study, half the spots on the Sypro stained gel didn’t contain enough protein for analysis; solution(s)?

Analysis of observed mw over predicted mw: many polypeptides run at mw higher than predicted.



What would it mean if every single spot ran at a higher mw than predicted?

Some protein markers for rat development

QuickTime™ and a
decompressor
are needed to see this picture.

DeCyder visualization tools confirm quantitative data.

QuickTime™ and a
decompressor
are needed to see this picture.

Coefficient of variance: statistical tool that assesses the extent of variation over the range of samples:

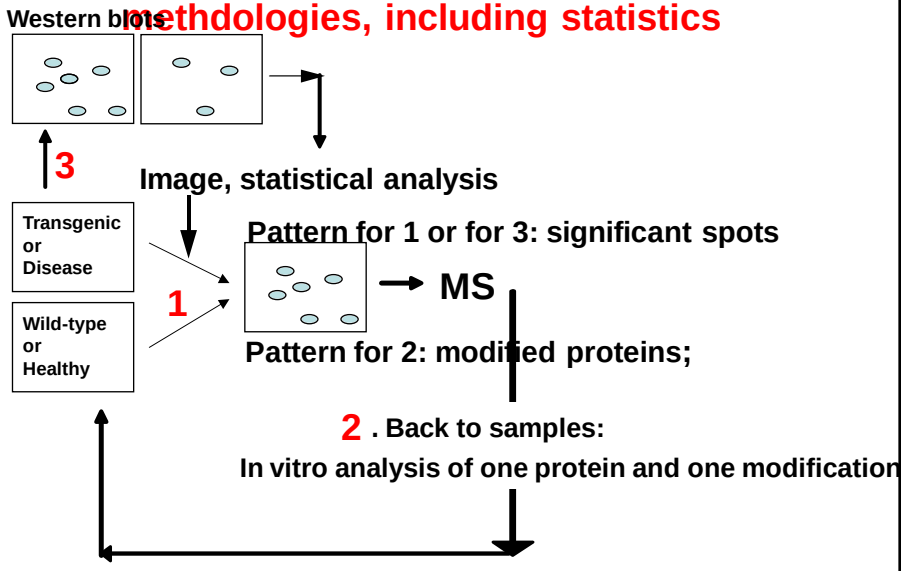
QuickTime™ and a
decompressor
are needed to see this picture.

Conclusion: there was no difference in CV between the two groups, thus the differences were due to the true means. **What does this graph tell you about how to keep your CV low?**

Take home lessons, part I

- **Proteomics suggests GSE has pleiotropic effects in the brain:**
  - **gene expression/protein turnover;**
    - **(Deshane et al., J. Agric. Food Chem.,2004)**
  - **protein oxidations;**
  - **Informatics helps you relate your data to the rest of the world; in the case of actions of GSE, we know that many of the proteins affected in normal brain are also differentially expressed in AD brain.**
    - ***How do we PROVE that any one effect of GSE actually prevents cognitive impairment?***

Optimal data analysis: Iterative process may start with proteomics but can/should include other methdologies, including statistics



(Kim et al., 2005, in Luo and Packer, Oxidative stress & neurodegeneration)

Thoughts for future experiments

- **DIGE: powerful, but room for lots of quality control;**
- **In the case of the data presented, these were differences in protein abundance between two developmental ages in female rat tissues;**
  - **Next: which are important in the actions of a carcinogen which is given at the later age?**
  - **Next: which of THESE are the ones affected by compounds that have anti-cancer activity.**
- **How do we zero in on epithelial cell specific differences, and minimize blood protein (albumin) differences?**