Danseed Symposium
Kobæk Strand, Skelskør
14 March, 2017

What are proteomics? And what can they tell us about seed maturation and germination?

Ian Max Møller
Department of Molecular Biology and Genetics
Aarhus University
Flakkebjerg, Denmark

Summary

- Proteomics still require access to relatively sophisticated (expensive) equipment – HPLC and mass spectrometers – and specialist operators
- For the best results, it requires that the DNA of that species has been fully sequenced
- Using proteomics, it is possible to separate, quantify and identify hundreds, or even thousands, of proteins in a sample
- The actual bioinformatic analyses afterwards can take much more time than the experimental analyses
- The information can be useful in plant breeding

Overview

- **Proteomics – General**
  - What are proteomics?
  - Separation
  - Quantification
  - Identification
  - Gel-based vs gel-free proteomics

- **An example: The proteomics of desiccation tolerance in maize embryos**

  (Huang et al. 2012, J. Proteom.)

Acknowledgements

- **Chinese Academy of Sciences**
  - For giving me a Visiting Professorship (6 months 2010-2014)
- **Institute of Botany, CAS, Beijing**
  - Professor Song-Quan Song
  - Dr. Wei-Qing Wang
  - All the other members of Professor Song's group
- **Department of Biochemistry and Molecular Biology at SDU, Odense**
  - Ole Nørregaard Jensen
  - Adelina Rogowska-Wrzesinska

What are proteomics?

- **The Proteome**
  - the complete profile of proteins expressed in a given cell, tissue, species or biological system at a given time
- **Proteomics**
  - The systematic analysis of the proteome
  - The techniques used to analyze the proteome

- Genomics era and now the post-genomics era
- Many ‘omics: Transcriptomics, proteomics, metabolomics
- Proteomics only possible since year 2000
  - Better mass spectrometers
  - Better databases (many fully sequenced species including many of agricultural interest)
  - Faster computers
  - Better search programmes

Why do proteomics?

- Genes are merely the blueprint for the proteins, which do all the work
- Transcription profiles do not reflect protein profiles
  - The central dogma

    ”DNA makes RNA makes proteins”

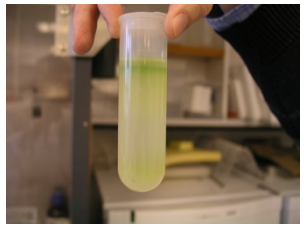
    is only qualitatively correct
- Find proteins (and therefore their genes) involved in specific processes
  - For seeds - desiccation tolerance, vigor, ageing, priming, etc.
- Find new functions

Some numbers

- Plant genomes have at least 27 000 protein-encoding genes
- Many proteins are post-translationally modified, e.g.
  - For example - phosphorylation
- > 300 different post-translational modifications
- Several modifications can occur on the same protein
- Each protein can exist in many (perhaps hundreds or thousands) different forms with slightly different properties!

Conclusions

- The study of post-translational protein modification will be (is) the next ”omics” revolution
- Fractionation of the proteins is usually a crucial first step in proteomics



Generalised work-flow

Complex protein mixtures

↓

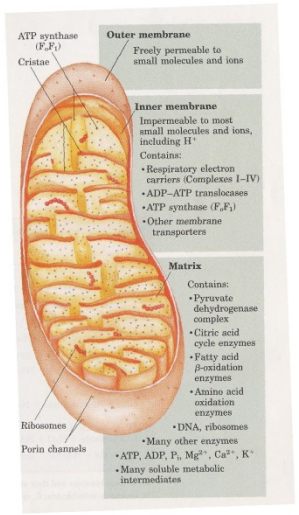
Separation

↓

Quantification

↓

Identification

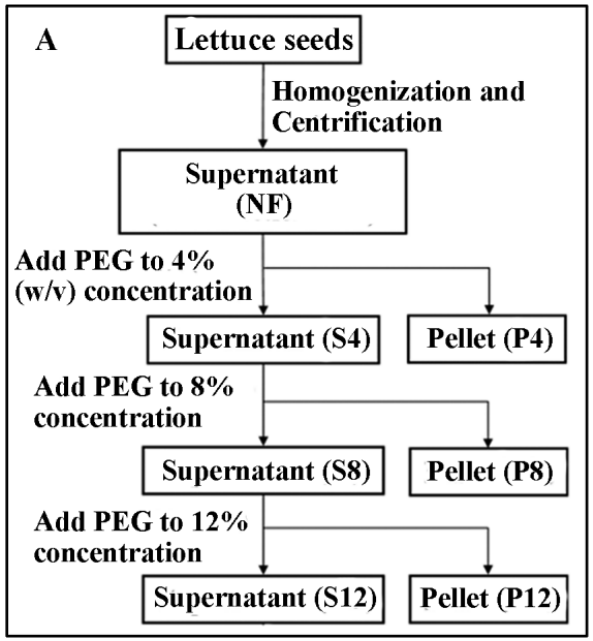


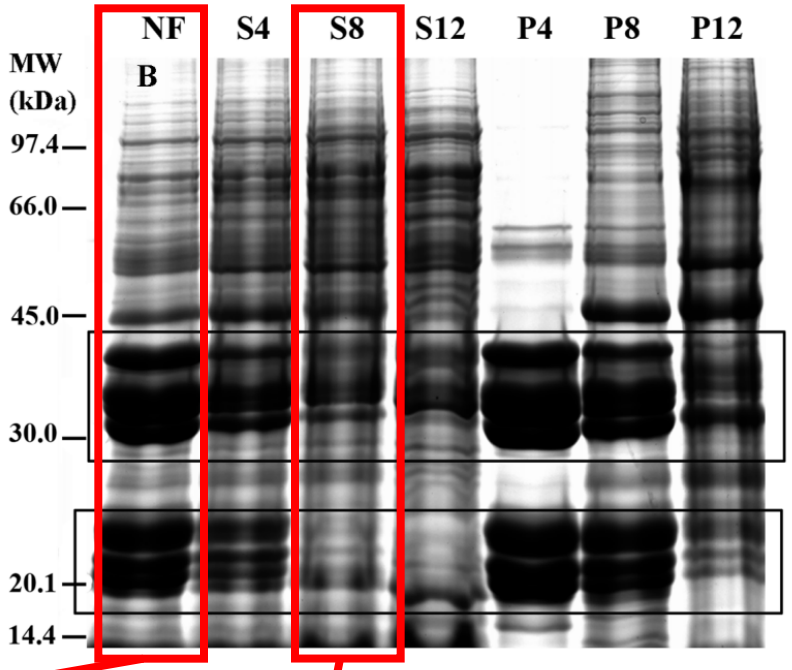
Separation techniques

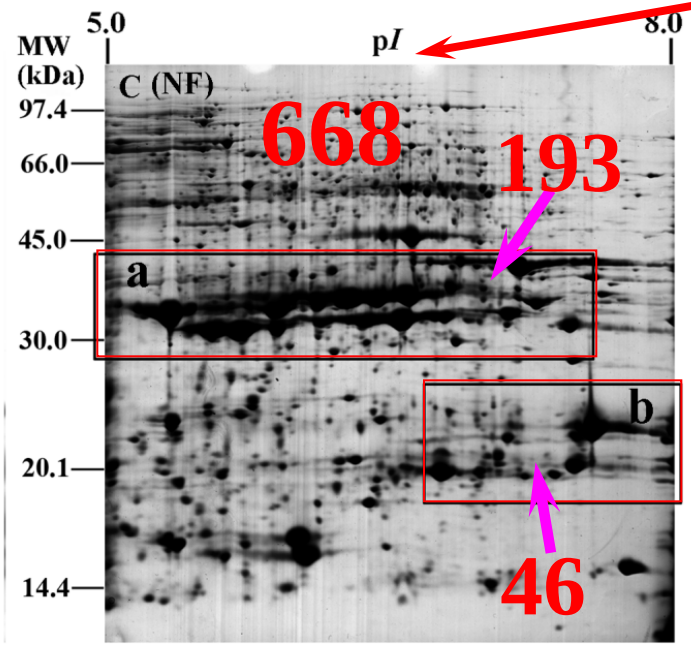
- **Protein fractionation**
  - PEG precipitation of superabundant proteins, e.g. storage proteins
  - Subcellular fractionation

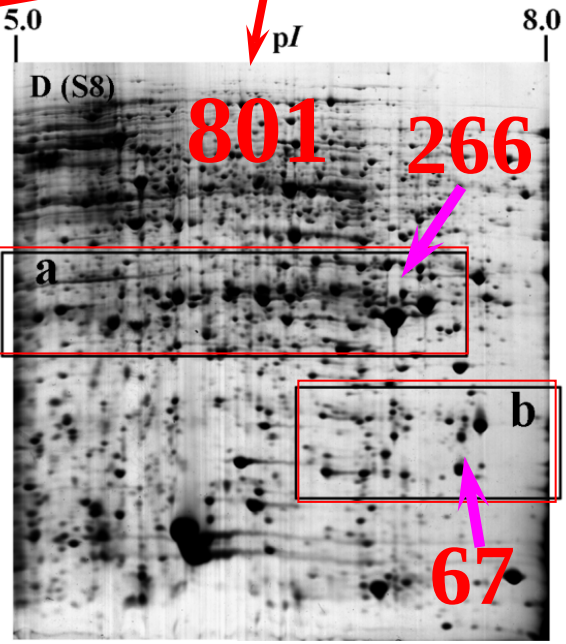
- **Electrophoresis – one-, two- and three-dimensional (1D, 2D, 3D)**
  - 2D and 3D - Image analysis program
  - 2D and 3D - Spot volume - number of pixels in a spot is a measure of the amount of protein
  - A 2D gel can separate 1000-1500 protein spots

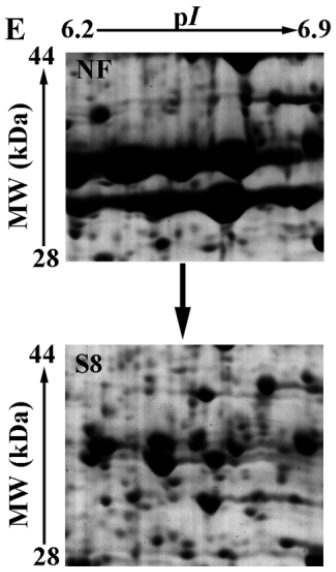
PEG precipitation to remove highly abundant storage proteins











Wang et al. 2015 Plant Physiology

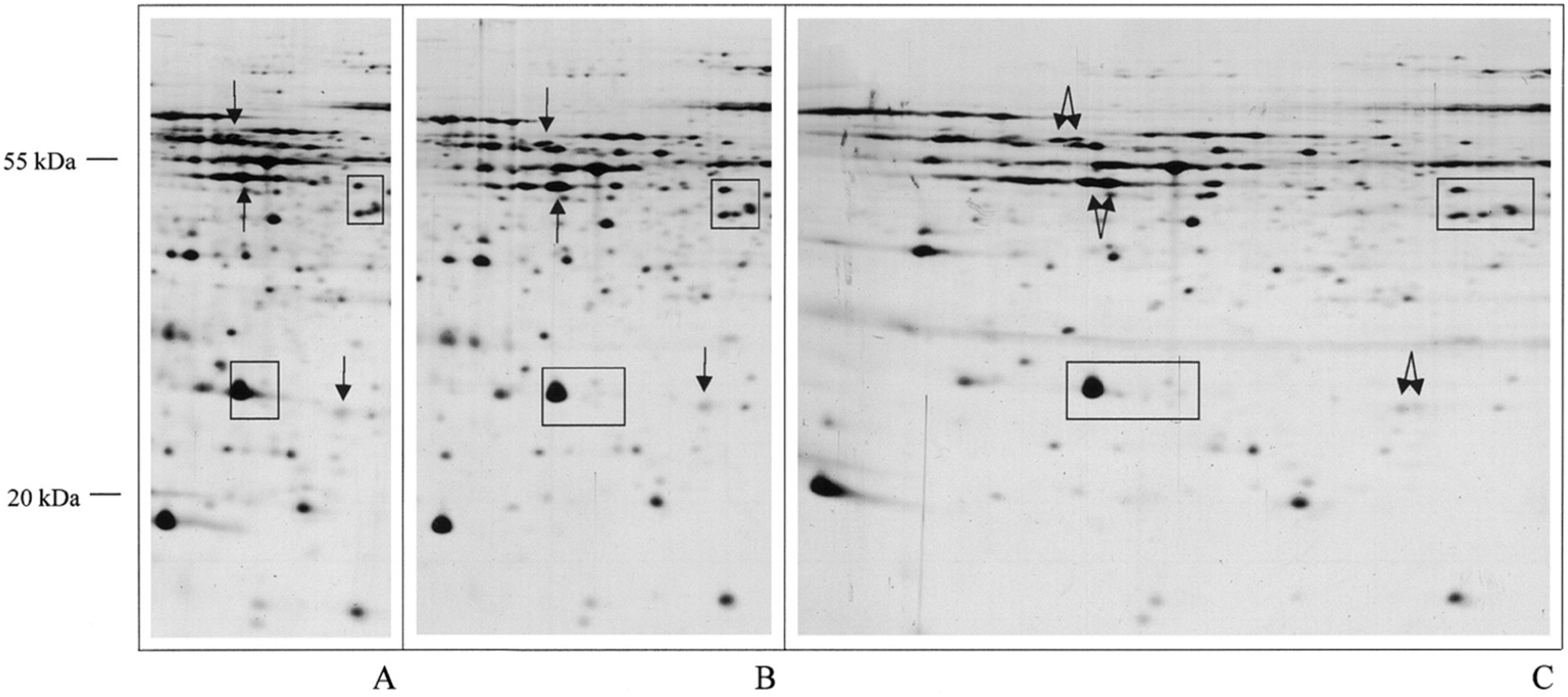
2D-gels followed by mass spectrometry

- 2D gels give much useful information about each spot/protein
  - Relative amount
  - Size
  - Isoelectric point
- Remember: only the most abundant ca 1000 proteins (e.g. pI 5-8) are visible
- Requires 200-500 μg protein
- Spots changing in amount can be identified using specialized image analysis programs

- Faster gel-free techniques are now available, but 2D gels will continue to be used

The effect of pH range in the 1st dimension on resolution

A, B and C all show the range pI 5.1-5.9



pH 3-10 pH 4-7 pH 5-6

The pH-range of the complete gel

Advantage – better resolution; disadvantage – loss of overview

Kruft et al. 2001

Protein identification in gel spots by mass spectrometry (MS)

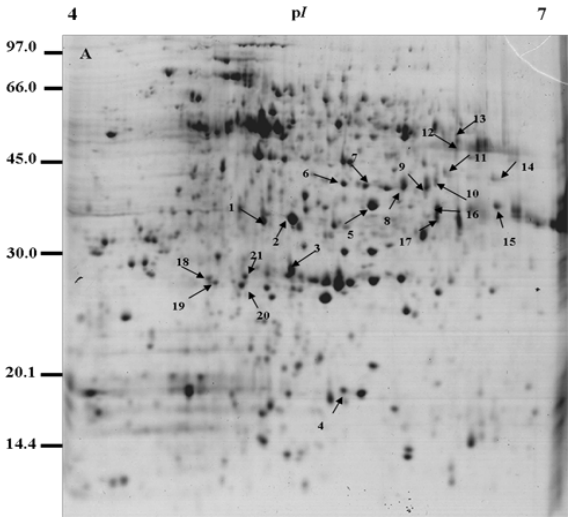
- Trypsin digestion of proteins in excised gel pieces (protein spots)
- Electrospray mass spectrometry
  - 1D-liquid chromatography (LC)-MS/MS
  - 2D-LC-MS/MS
- This gives the mass of the individual peptides (MS) as well as their fragmentation pattern (MS/MS)
- Database search (e.g. Mascot) in DNA/protein database of that species (or closely related species)

Gel-free proteomics
2D-Liquid chromatography-mass spectrometry

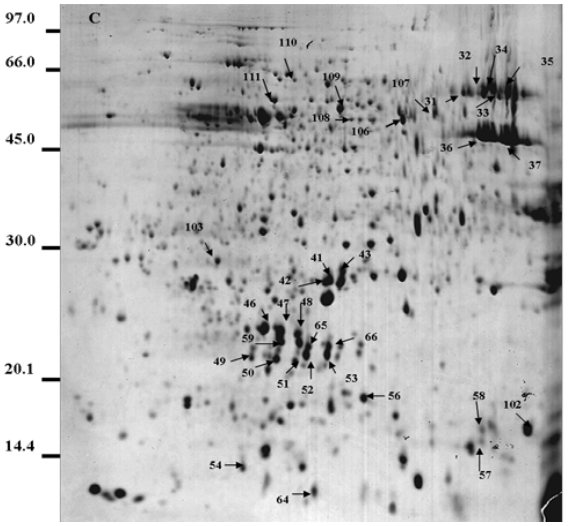
- The method:
  - A complex protein mixture is treated with trypsin
  - The peptides are separated by 2D liquid chromatography
  - The peptides from the column enter the mass spectrometer on line
  - The peptides are characterized by MS/MS
  - The proteins to which the peptides belong are identified by database search
- Faster, but does not give the same amount of information
- More proteins with extreme properties identified
- Very sensitive – requires only 5 μg protein
- Quantitative MS – labelling is necessary
  - MS analysis is not in itself quantitative - different peptides have different abilities to become vaporized and therefore give different signal size

2D-gels of maize embryo proteins

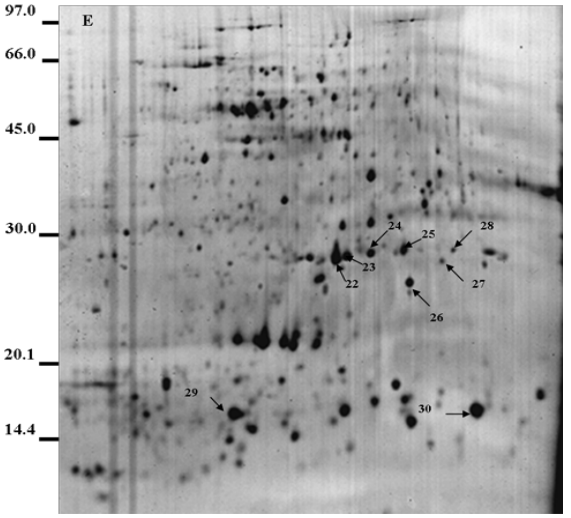
28N



52N

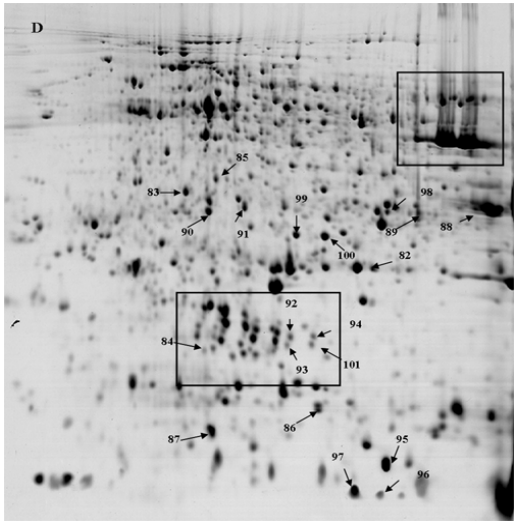


72HN



1st dimension – isoelectric focusing (pI)

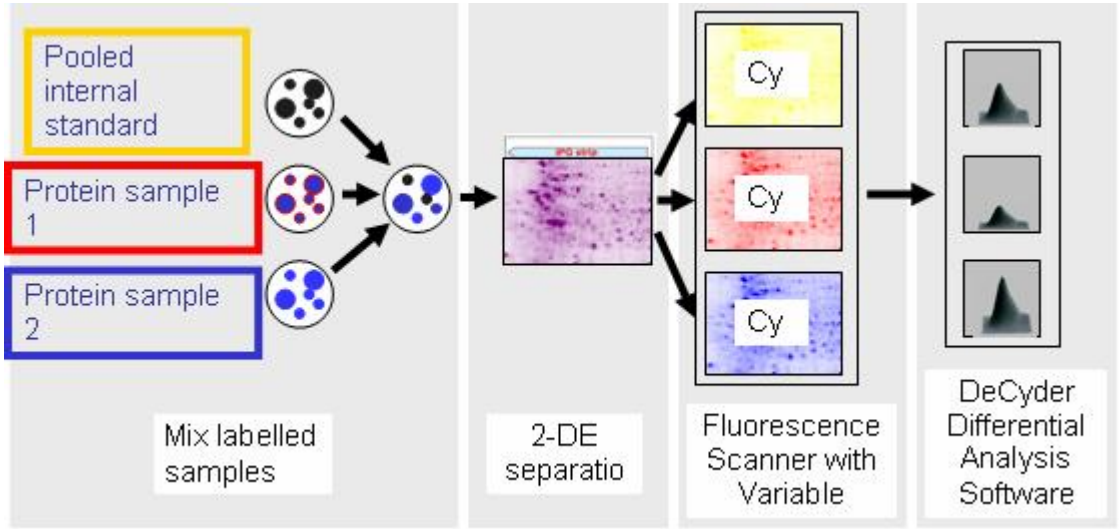
52D

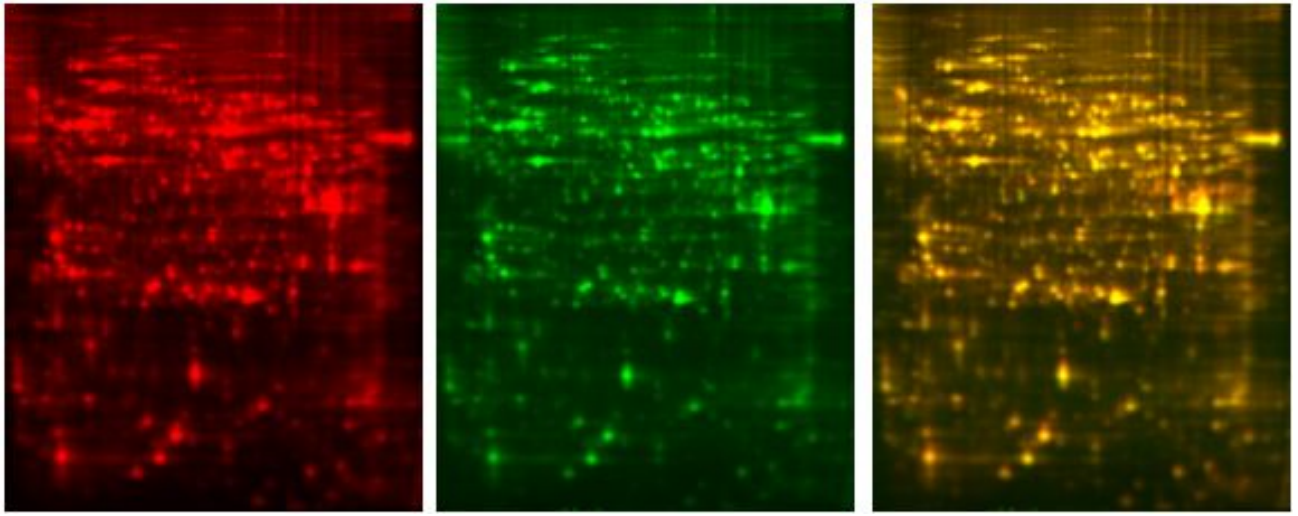


2nd dimension – SDS-PAGE – protein Size (kDa)

Huang et al. 2012 J Proteom

DIGE – Differential in gel electrophoresis



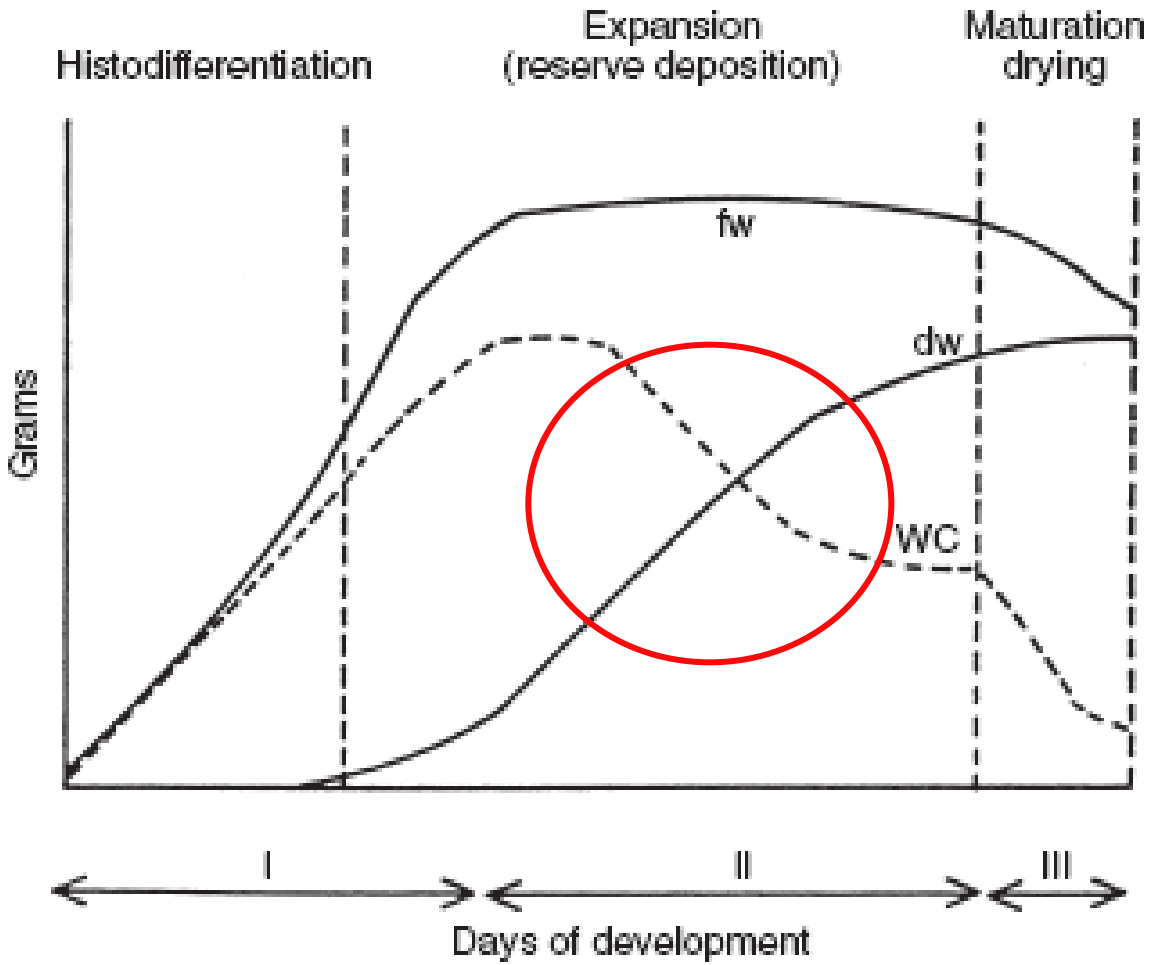


An example

Proteomics of desiccation tolerance in maize

(Huang, Møller and Song 2012 Journal of Proteomics)

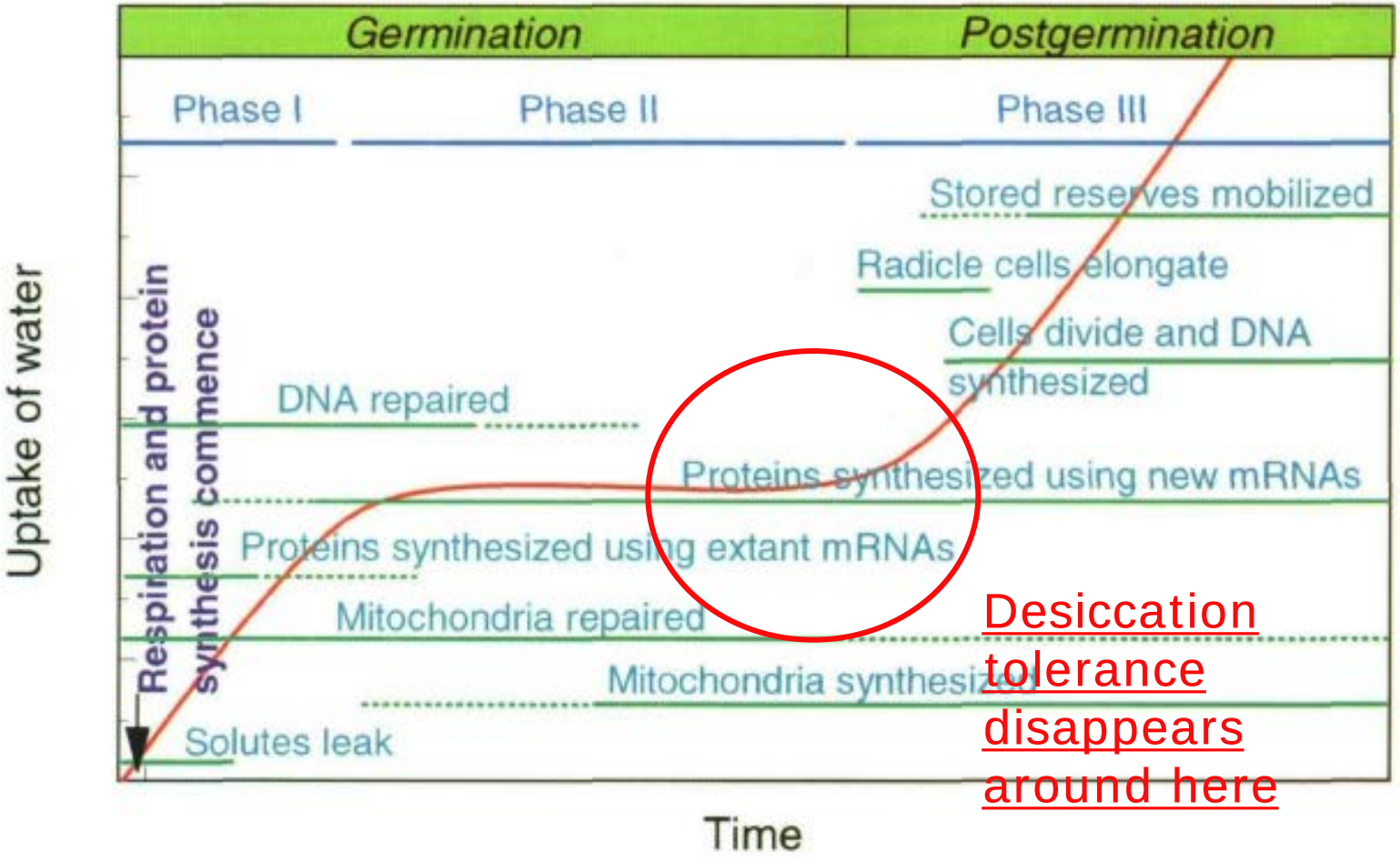
Development and maturation of orthodox seeds



Desiccation tolerance develops around here

Kermode and Finch-Savage 2002

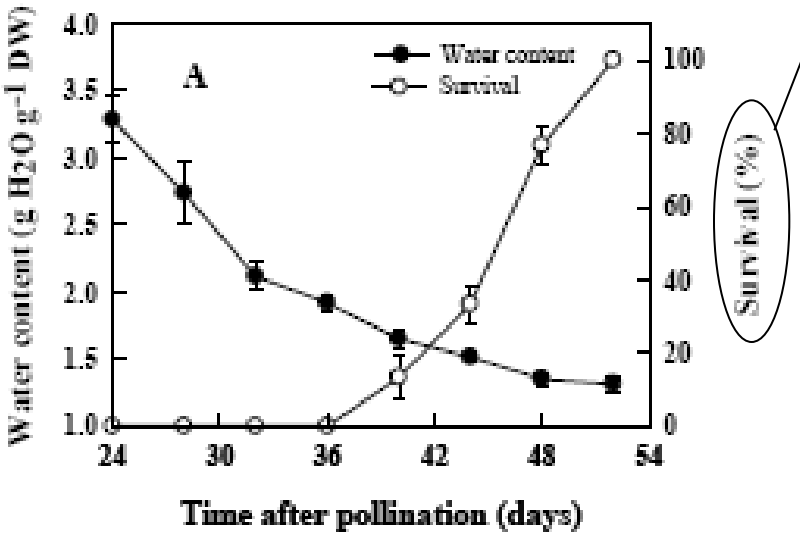
Seed germination

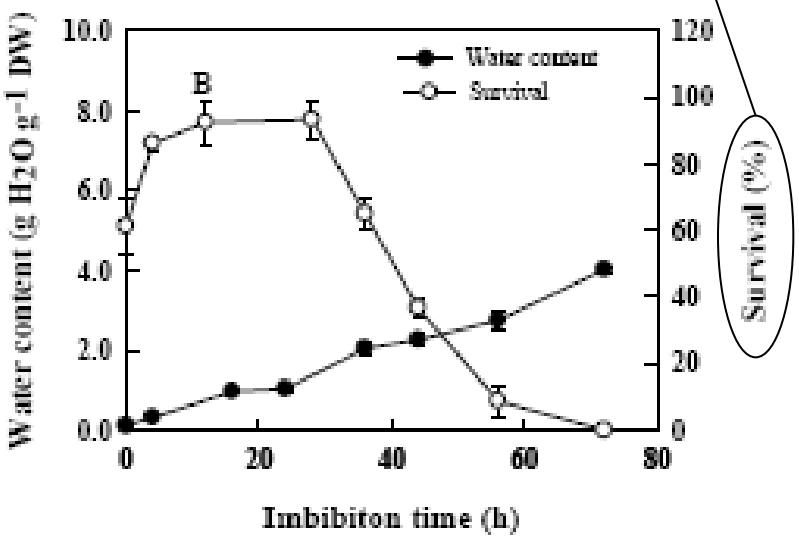


Bewley 1997 Plant Cell

Basic properties of developing and germinating maize embryos

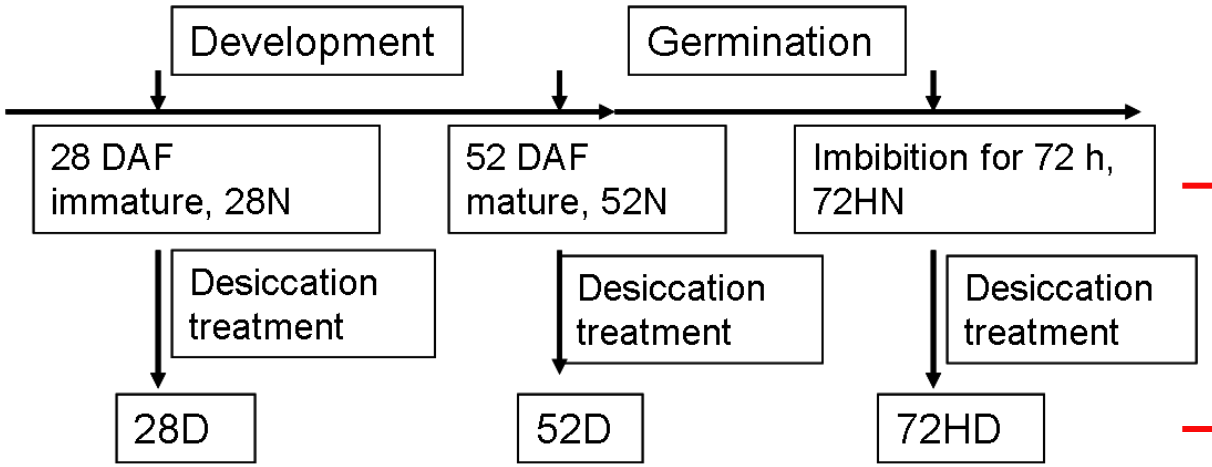
After desiccation





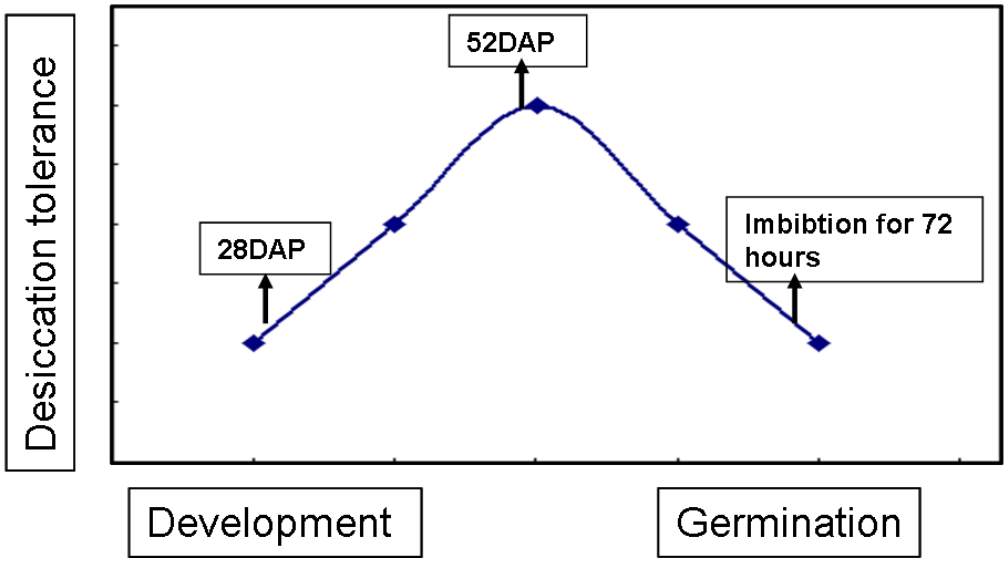
Huang et al. 2012 J Proteom

Experimental protocol for the study of desiccation tolerance in maize embryos



Protein extraction

Protein extraction



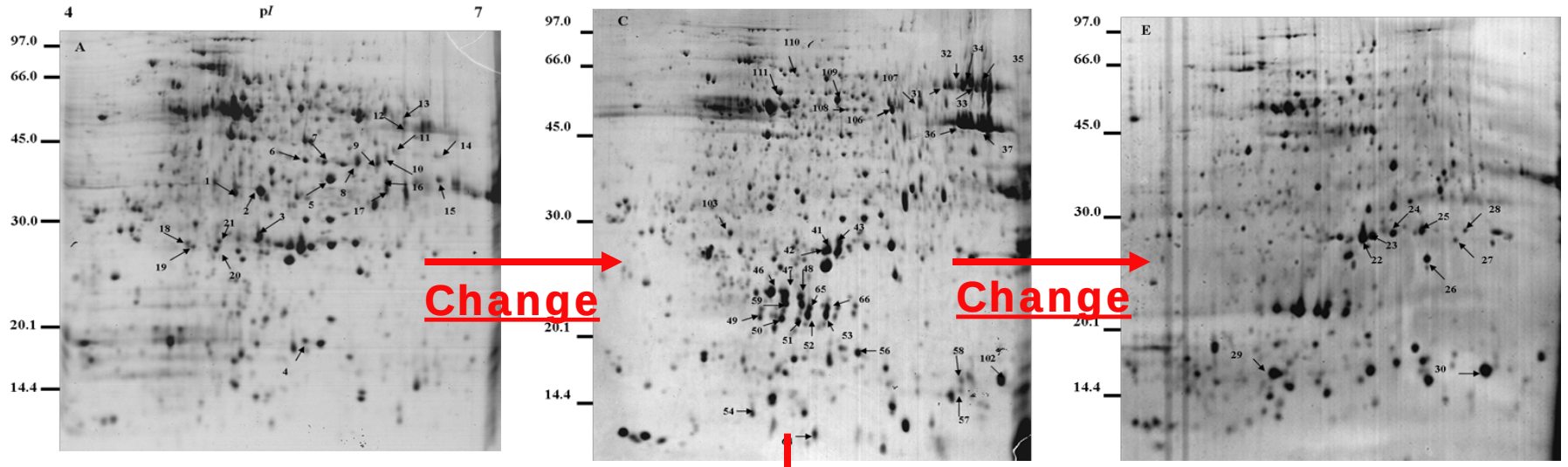
Huang et al. 2012 J Proteom

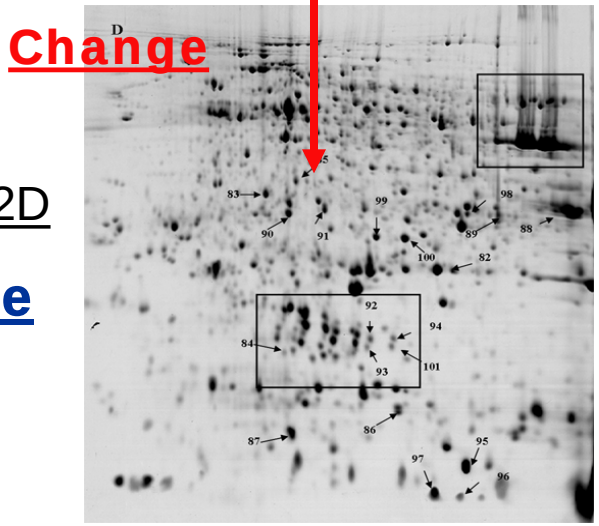
2D-gels of maize embryo proteins – Finding changes

28N

52N

72HN



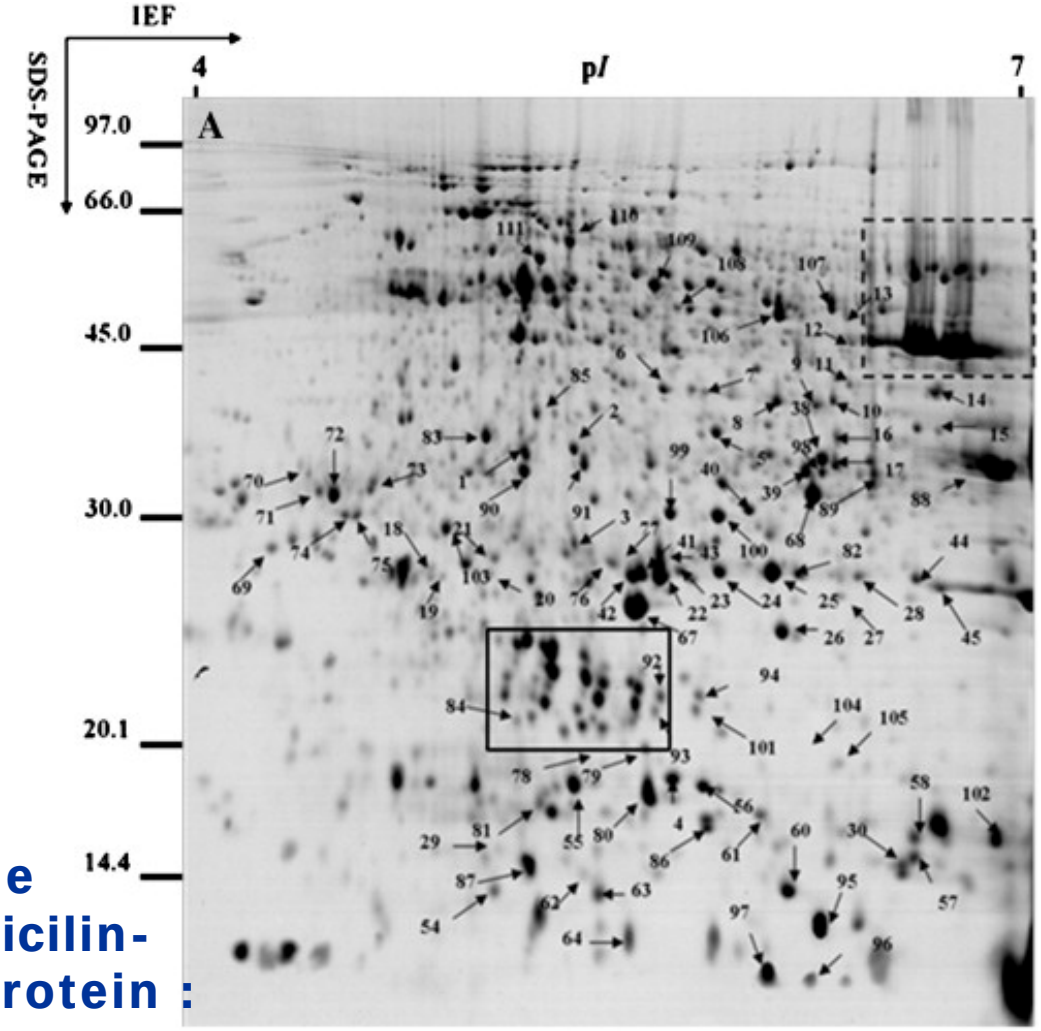


52D

Significant change
Increase 1.5x
Decrease 1.5x

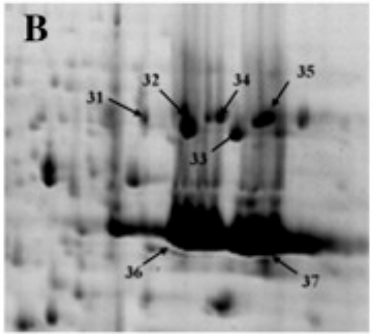
Huang et al. 2012 J Proteom

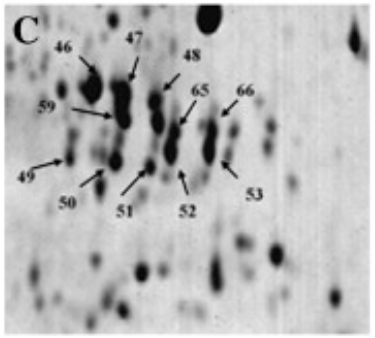
2-DE analysis of maize embryo proteins (52D)



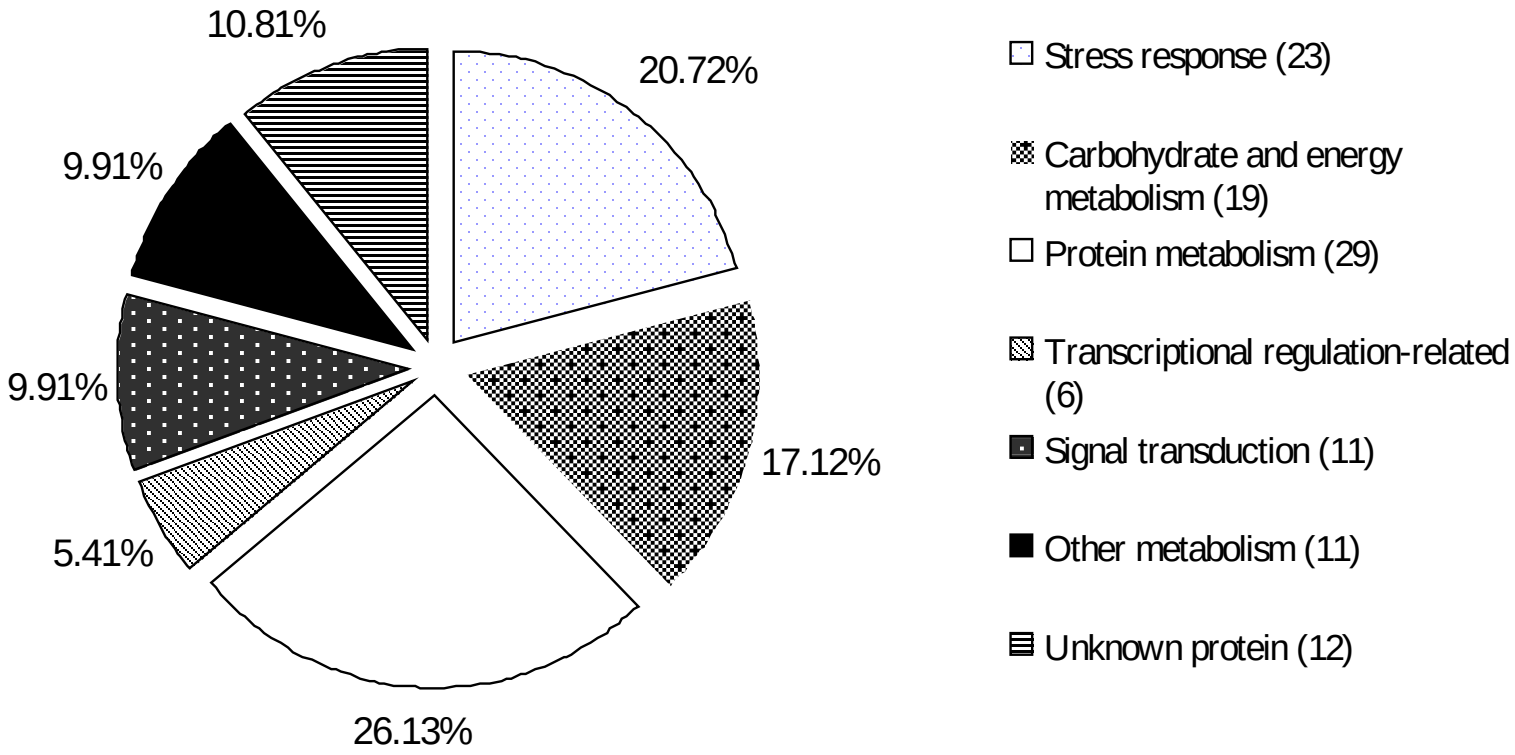
Many spots contain the same gene product (vicilin-like embryo storage protein :

Spots 31-35 – full size
Spots 46, 59, 68 – truncated





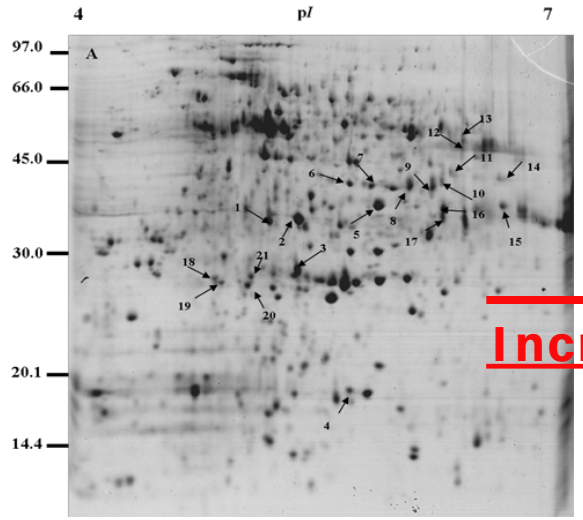
Functional classification of all 111 identified changing protein spots



Huang et al. 2012 J Proteom

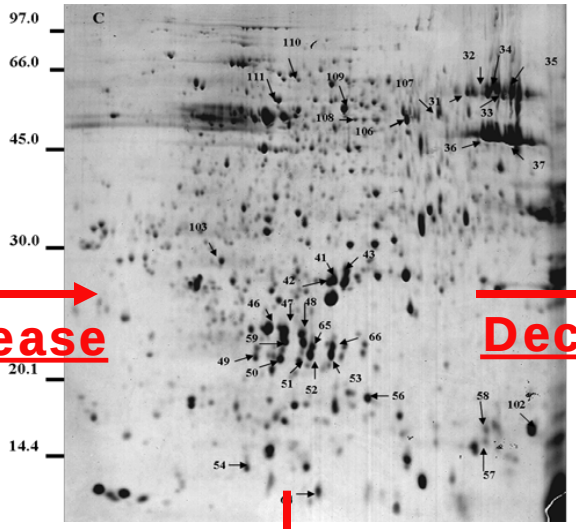
Proteins of particular interest

28N



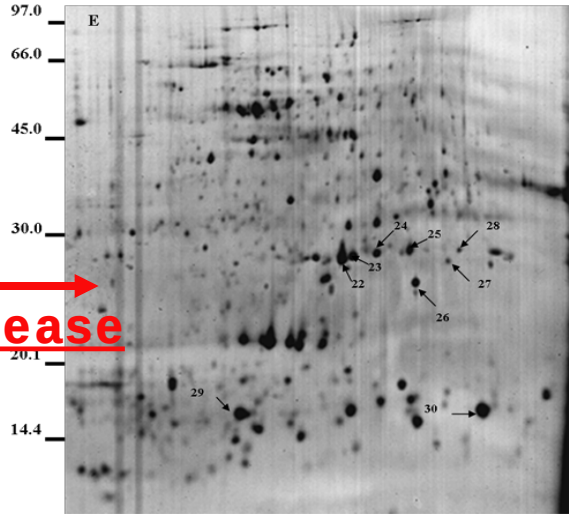
Increase

52N



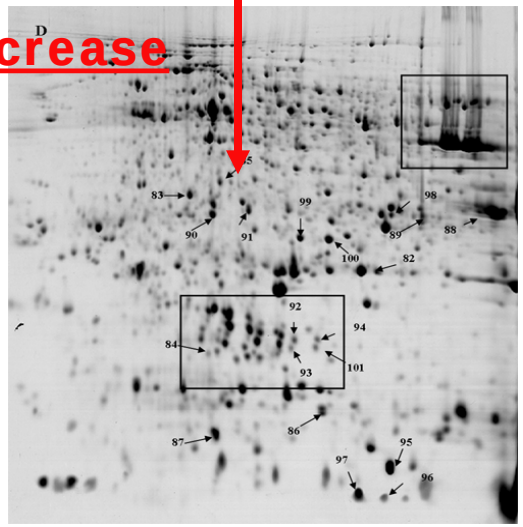
Decrease

72HN



Increase

52D



Significant change
Increase 1.5x
Decrease 1.5x

Or the inverse pattern (for proteins which make the embryos less desiccation tolerant)

Huang et al. 2012 J Proteom

Conclusions

- We identified nine proteins potentially involved in helping maize embryos become desiccation tolerant
- We identified two proteins, which may make the embryos less desiccation tolerant
- Most of these proteins (seven) are stress-related
- The next step would be to investigate the expression of these target proteins in various tissues during seed maturation and germination

Huang et al. 2012 J Proteom

Proteins possibly involved in conferring desiccation tolerance in maize embryos

Increased 28N→52N, decreased 52N→72HN, increased 52N→52D

- 17.4kDa Class I heat shock protein 3 (spot 55),
- late embryogenesis abundant protein EMB564 (spot 57)
- OmpA/MotB family outer membrane protein, (spot 58)
- globulin 2 (spot 66)
- TPA: putative cystatin (spot 82)
- NBS-LRR resistance-like protein RGC456 (spot 86)
- stress responsive protein (spot 88)
- major allergen Bet v 1.01C (spot 96)
- proteasome subunit alpha type1 (spot 97)

Decreased 28N→52N, increased 52N→72HN, decreased 52N→52D

- Rhd6-like 2 (spot 29)
- low-molecular-weight heat shock protein precursor (spot 78)

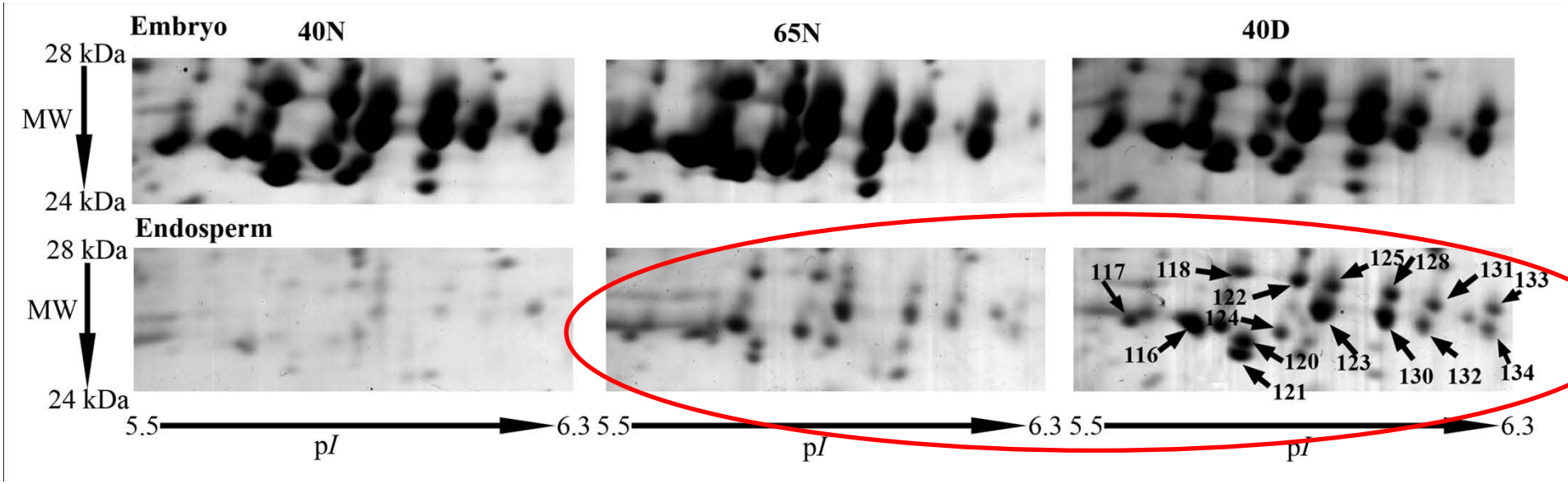
Huang et al. 2012 J Proteom

Summary

- Proteomics still require access to relatively sophisticated (expensive) equipment – HPLC and mass spectrometers – and specialist operators
- For the best results, it requires that the DNA of that species has been fully sequenced
- Using proteomics, it is possible to separate, quantify and identify hundreds, or even thousands, of proteins in a sample
- The actual bioinformatic analyses afterwards can take much more time than the experimental analyses
- The information can be useful in plant breeding

Thank you for your attention!

Accumulation of low-molecular weight storage proteins



Accumulation of low-molecular weight storage proteins in the endosperm may assist in germination

Wang et al. 2014 J Proteome Res

The sequence coverage of an endosperm storage protein - globulin-1 S allele precursor

Wang et al. 2014
J Proteome Res

Sequenced peptides from all the spots containing the protein (62% overall coverage)

Example of the full-sized protein (spot 55 – 56 kDa) with sequenced peptides matching both the N- and C- terminal of the sequence (38% overall coverage)

Example of the shortened protein (spot 132 – 27 kDa) with sequenced peptides matching only to the N-terminal of the sequence (31% overall coverage)

Example of the shortened protein (spot 153 – 15 kDa)) with sequenced peptides matching only to the C-terminal of the sequence (12% overall coverage)