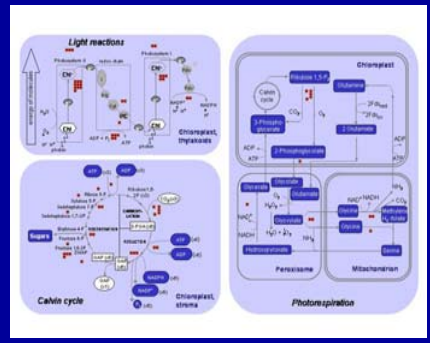
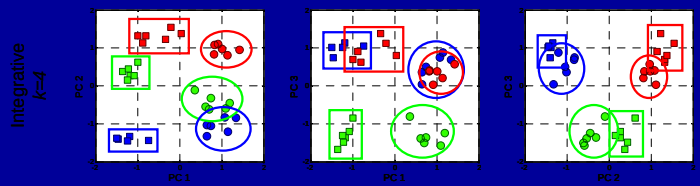
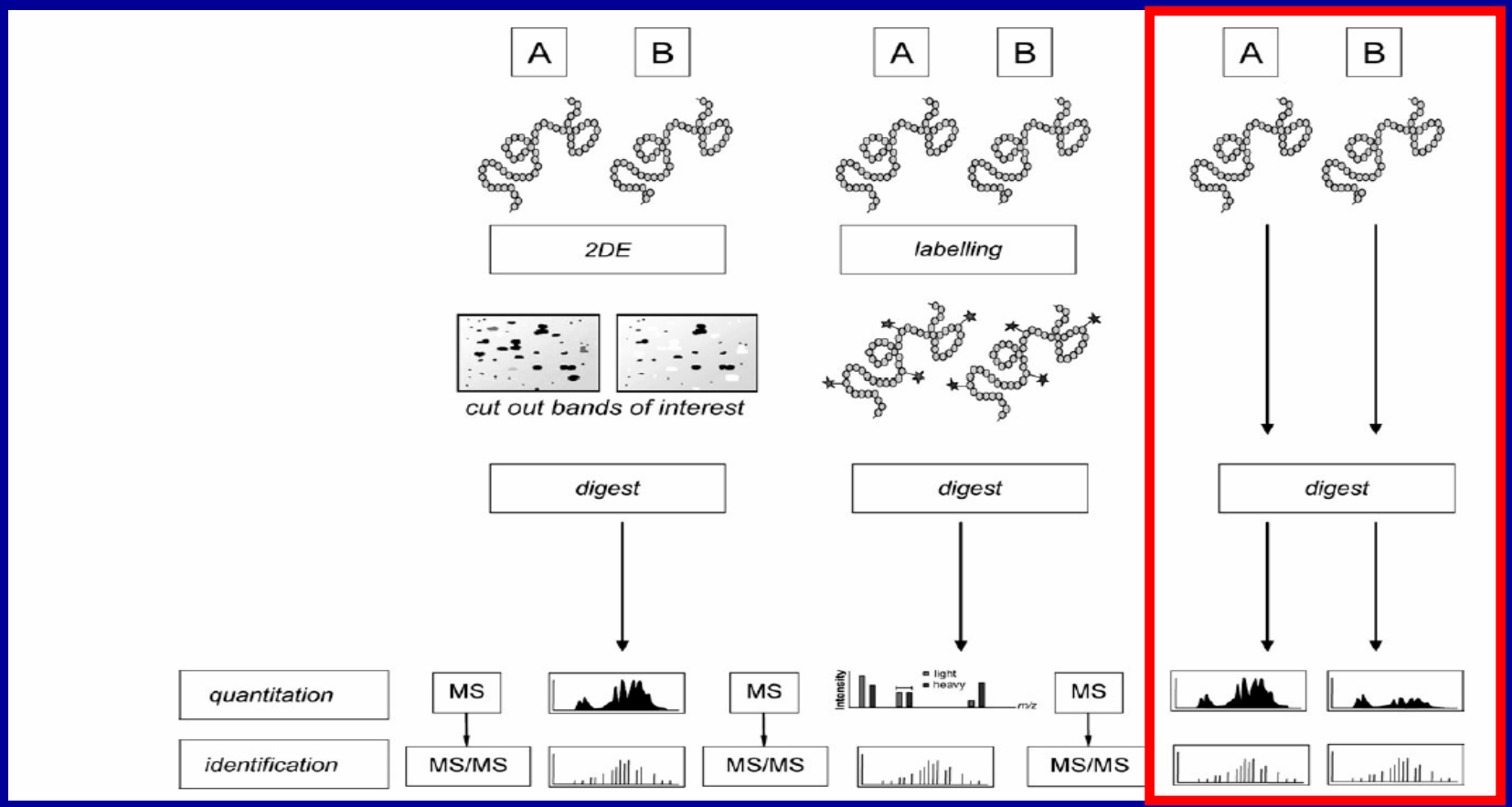
Quantitative Shotgun Proteomics

Wolfram Weckwerth
MPI Molekulare Pflanzenphysiologie



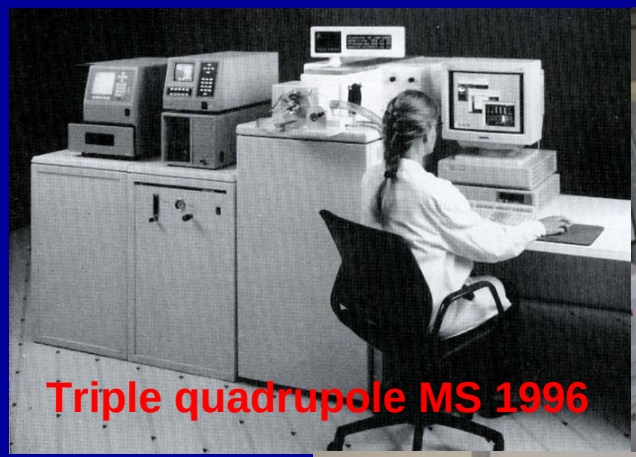


Label-free quantitative proteomics – “nongel” approach



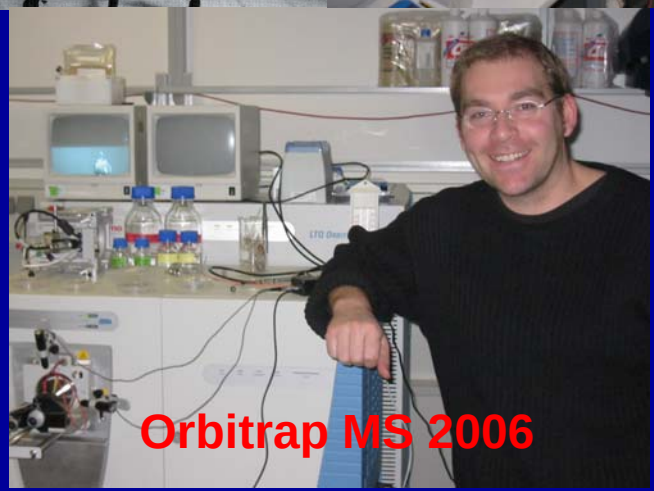
Glinski & Weckwerth (2006) Mass Spectrometry Reviews

Developments of Mass analyzers



Triple quadrupole MS 1996

Iontrap MS 2004



Orbitrap MS 2006

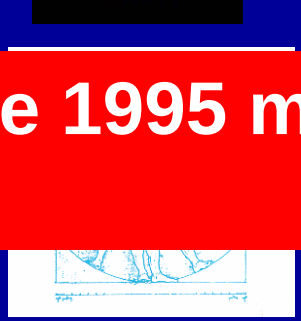
Model Systems

Arabidopsis

Tomate/ Kartoffel

Synechocystis

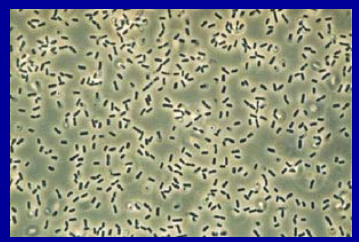
since 1995 more than 300 organisms sequenced ....trend increasing



Homo sapiens

Familiaris

Medicago



Bacillus

Chlamydomonas

Maus

„Shotgun proteomics/phosphoproteomics“

- „non-gel“ based protein analysis

Complex Protein Sample

↓ **Proteolytic digest**

HOMOGENEOUS Peptides

↓

Chromatography → **ESI** → **MS MS/MS**

LC/MS

↓

Search against Genomic Database (MASCOT, Sequest)

↓

Peptide Identification and Quantification

←

QUANTITATIVE Protein Data

Weckwerth et al. 2004 Proteomics
Wienkoop et al. 2004 RCM
Wienkoop et al. 2004 Phytochemistry
Morgenthal et al. 2005 Metabolomics
Wienkoop and Weckwerth 2006 JXB
Wienkoop et al. 2007 JSS

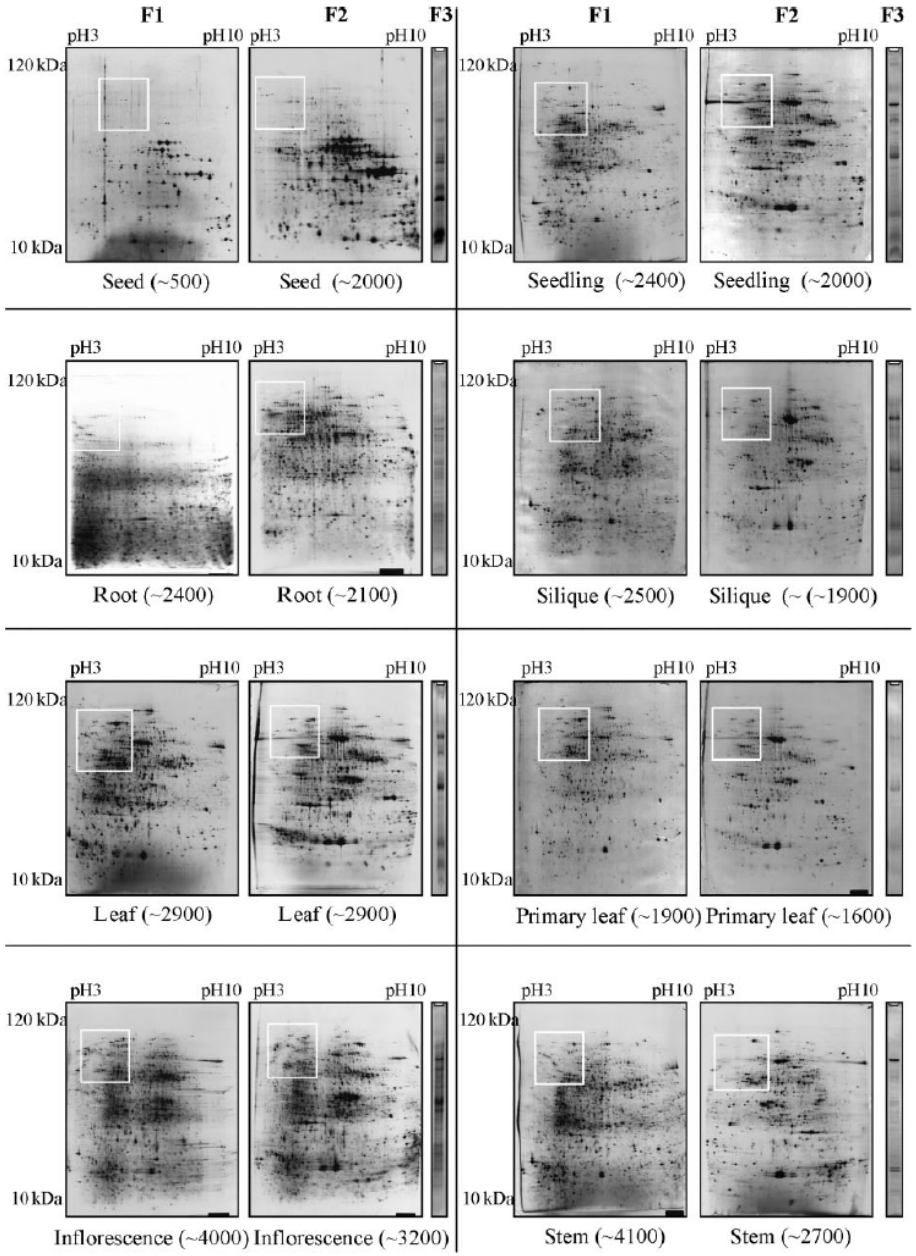
Phosphoproteomics:
Glinski et al. 2003 RCM
Glinski an Weckwerth 2005 Molec & Cell Prot
Wolschin et al. 2005 Proteomics
Wolschin et al. 2005 RCM
Wolschin and Weckwerth 2005 Plant Methods
Wolschin and Weckwerth 2006 RCM

Qualitative Proteinanalyse

2DE versus Shotgun

- **Complementary techniques** (protein ID, soluble fraction, membrane proteins)
- Automation/throughput
- Protein species versus unique proteins

2DE



Giavalisco et al. 2005

A. thaliana

6000 spots ausgeschnitten
MALDI-TOF 2900 identifiziert
660 „unique proteins“

⇒Studie hat ca. 2-3 Jahre gedauert!
⇒(Optimierung der Extraktion + 2DE,
⇒spots ausschneiden, analysieren etc.)

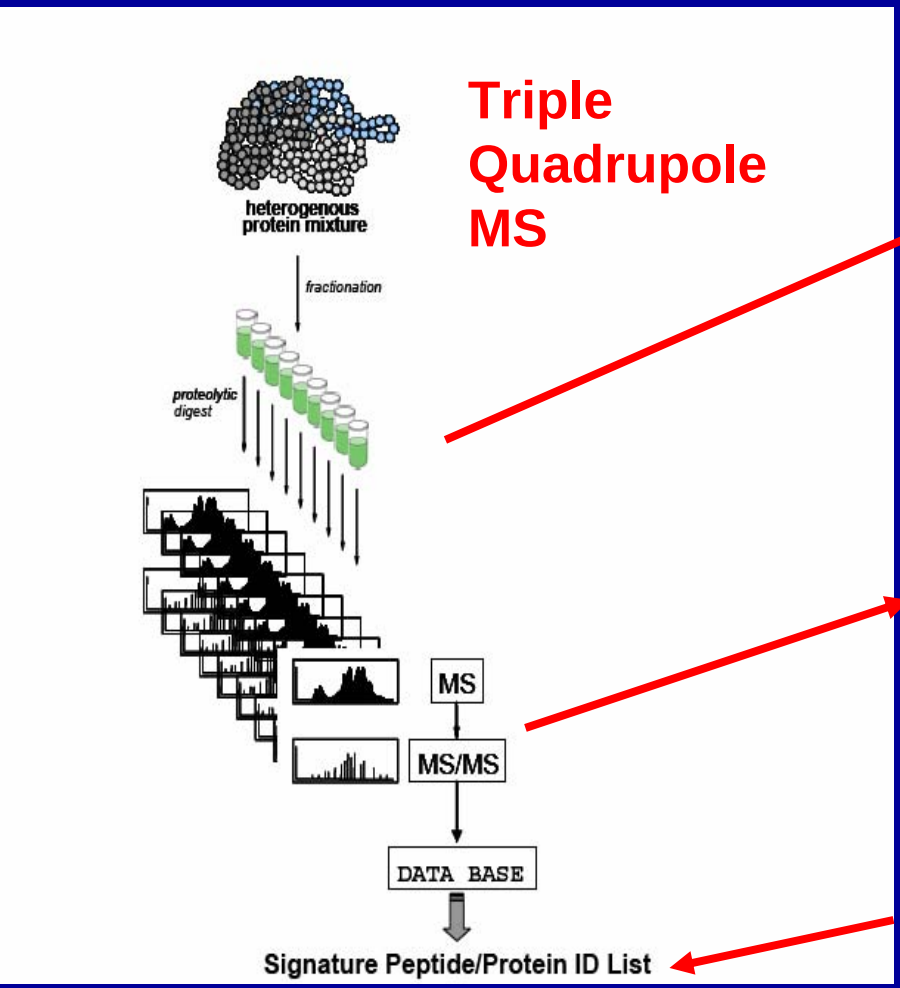
⇒spot => „Protein-Spezies“
(Proteinmodifikation: Phosphorylaierung, Acetylierung, Glycosilierung, etc.)

Jungblut & Thiede (1997)
Mass Spectrometry Reviews

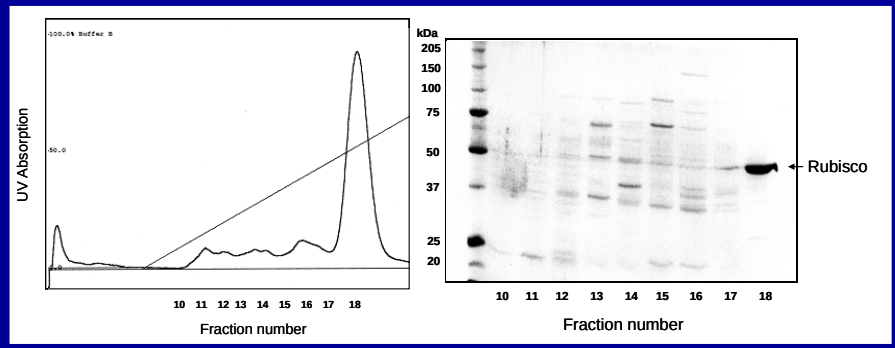
Shotgun proteomics

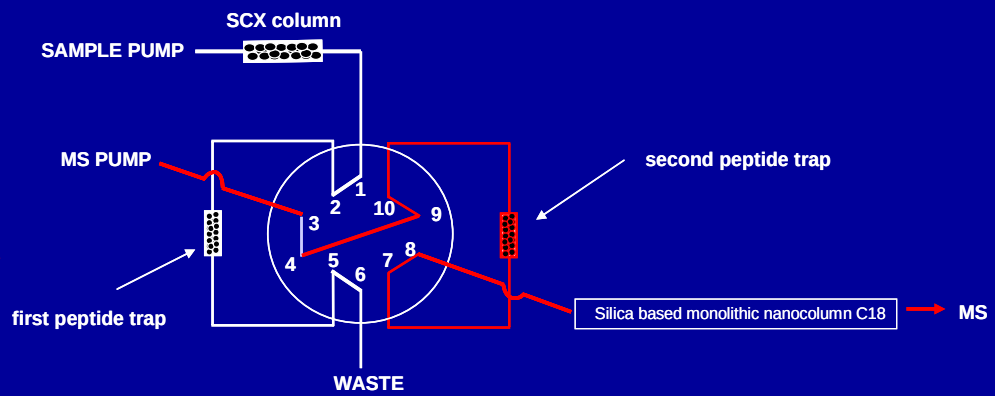
- Automation and throughput:
  - 1D LC/MS => up to 200 unique proteins (120 min per analysis, protein modifications/splicing/protein species hard to identify, identification and quantification parallel)
  - 2D LC/MS => up to 500 unique proteins (6 cycles á 90 min => 9 hours/sample, identification and quantification parallel)
  - Rapid Screening (of all mass spectra combined with biostatistics, identification process only for significant changes)

Shotgun protein Masterlists
=> signature peptides = target proteins



Triple Quadrupole MS





⇒Digestion of protein fractions
⇒2dimensional nano LC/LC/MS/MS analysis
⇒Masterlist of proteins for LC/MS detection methods
⇒ ~1032 unique proteins Arabidopsis thaliana
⇒ ~ 1 week/ sample (*Arabidopsis thaliana* + *Medicago truncatula*)

Wienkoop et al. RCM 2004

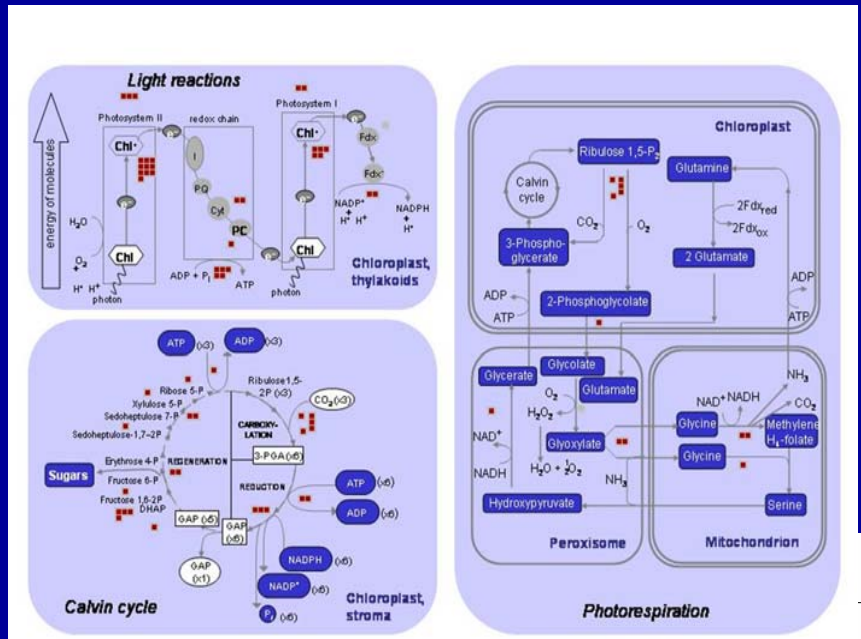
CID-Mass Spectral Libraries and Metainformation in a **Reference Database** for Proteomics Experiments

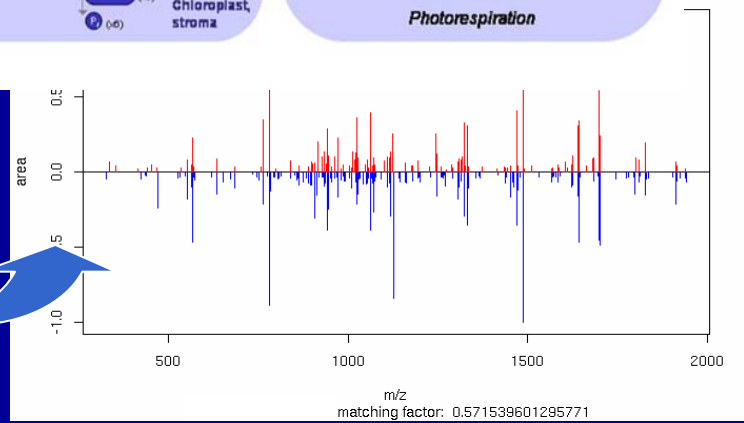
Protein/Peptide list
Reference compounds
Unknowns

Collision-induced dissociation (CID)
MS^n

Mass spectral libraries with native molecular ion and fragmentation pathways

Measurement of unknown samples and matching with Mass spectral library (similar to GC/MS)





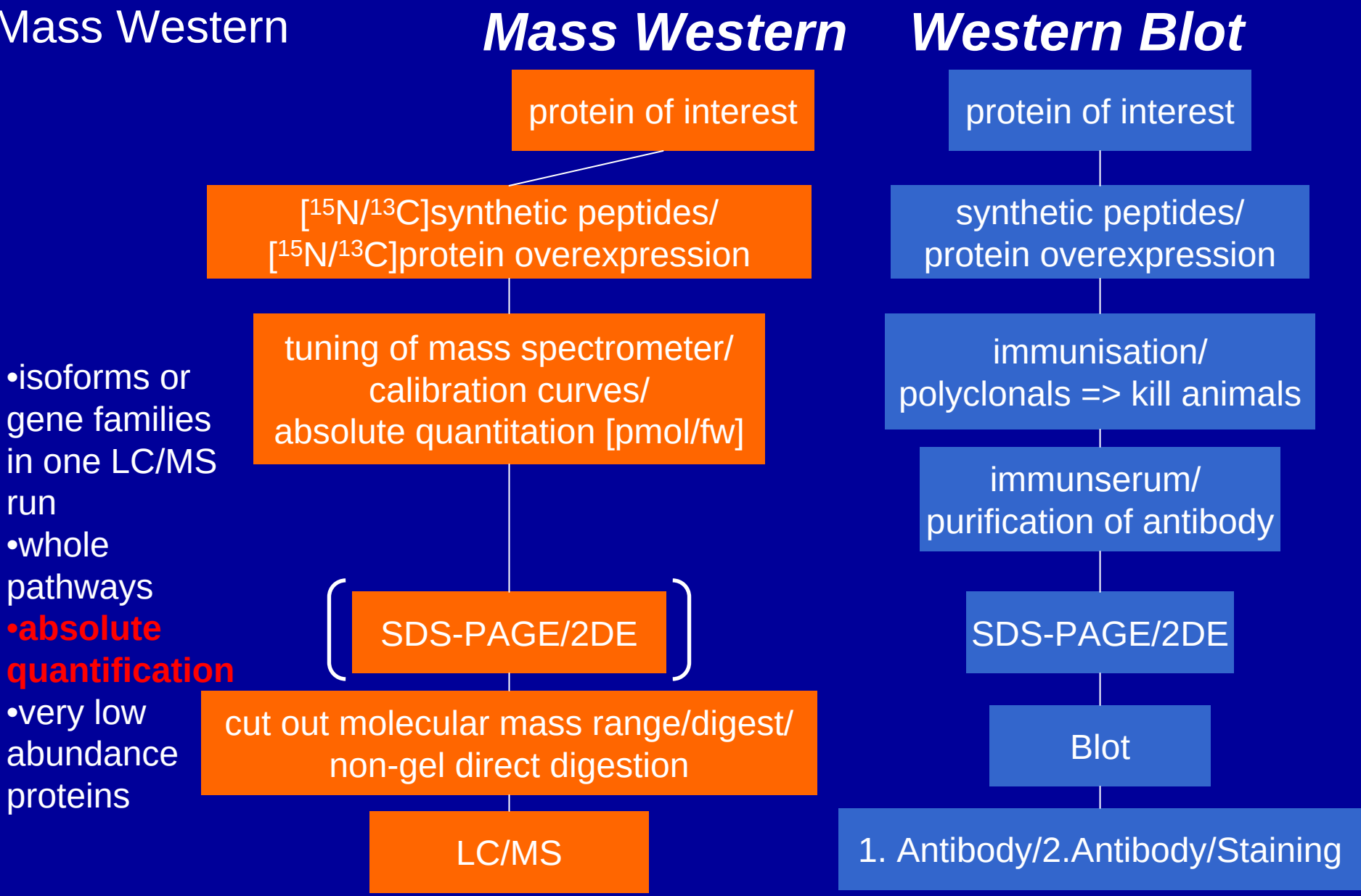
LC/MS Mass spectral library ProMEX

Hummel et al. submitted

2748 peptides *Arabidopsis thaliana*
842 peptides *Medicago truncatula*
486 peptides *Chlamydomonas reinhardtii*
~300 phosphoproteins *Arabidopsis thaliana* and *Chlamydomonas reinhardtii*

Quantitative Proteinanalyse

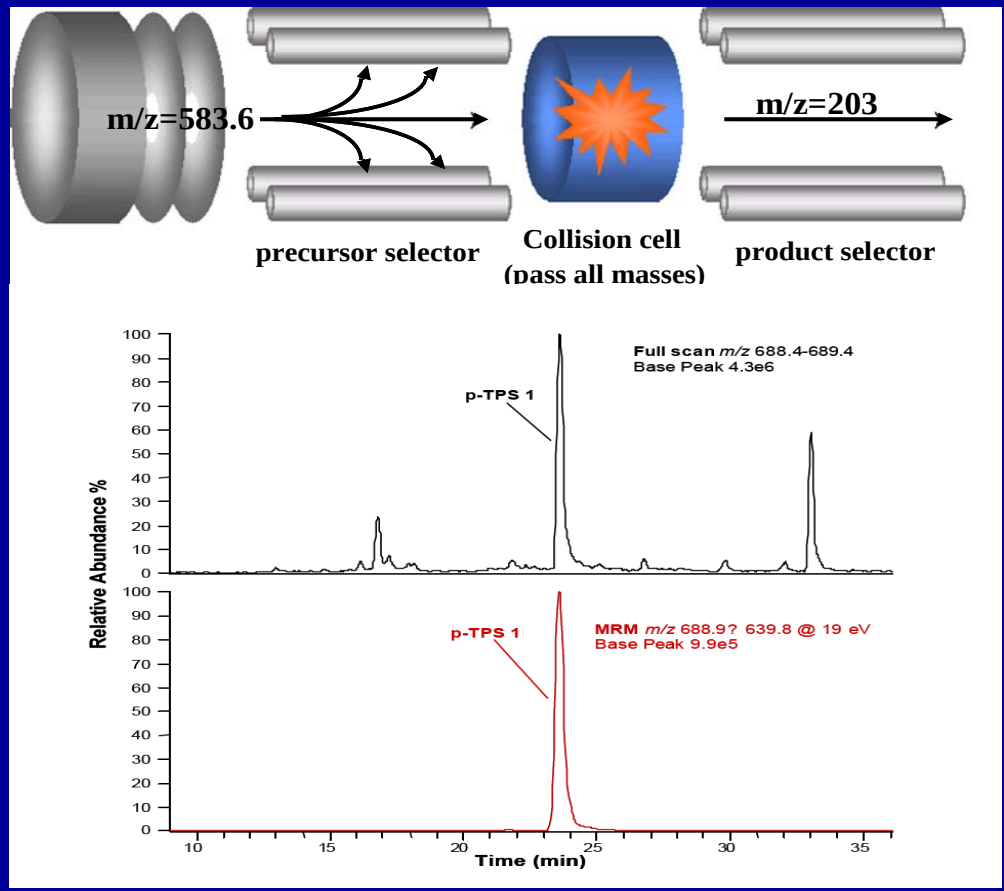
Signature peptides - Mass Western



- isoforms or gene families in one LC/MS run
- whole pathways
- **absolute quantification**
- very low abundance proteins

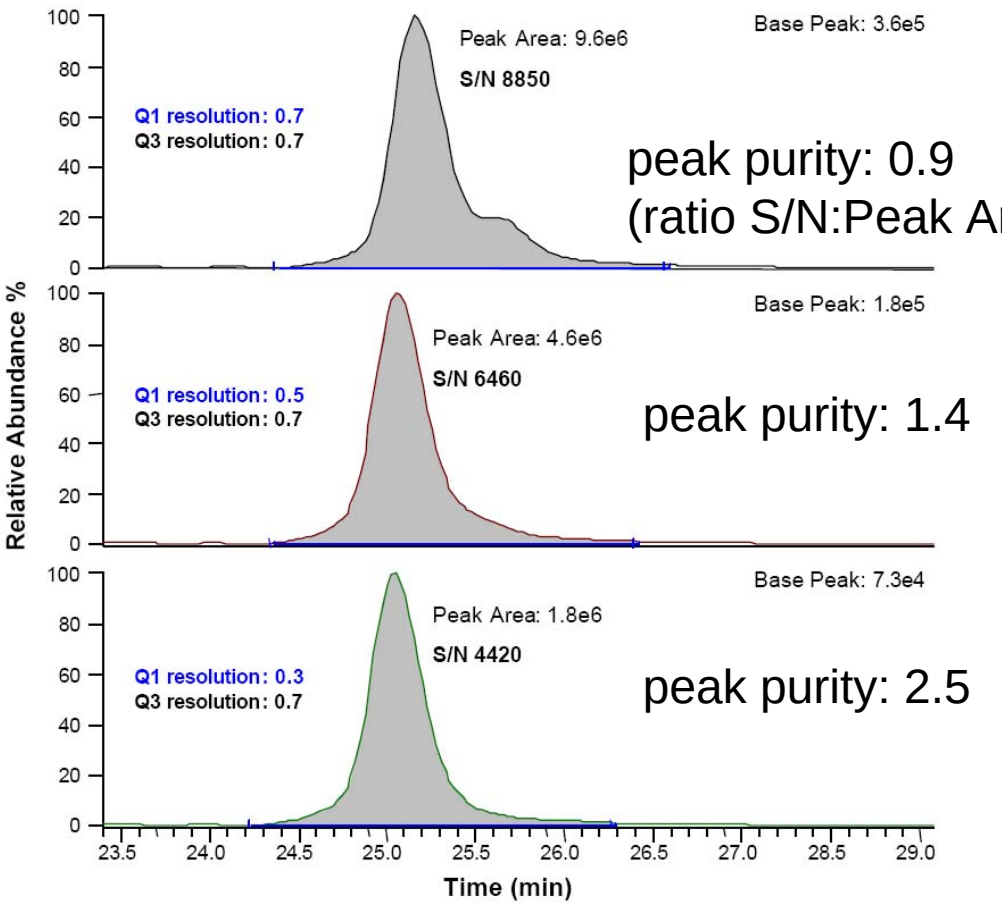
Wienkoop & Weckwerth JXB 2006

„Peptide Single Reaction Monitoring“ in a triple quadrupole MS



Glinski & Weckwerth (2005) Molecular & Cellular Proteomics

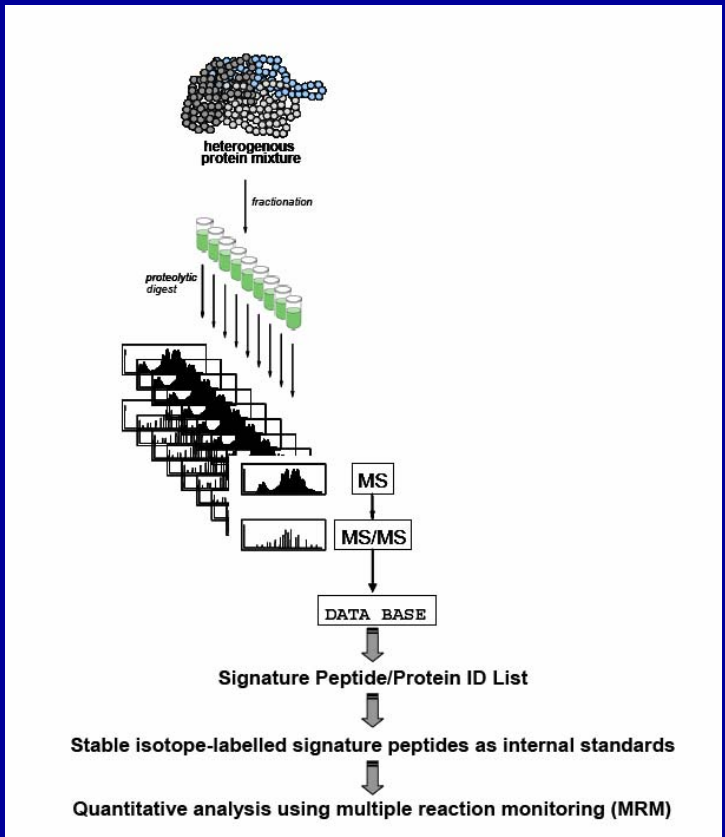
High resolution triple quadrupole MS enhances peak purity in SRM mode

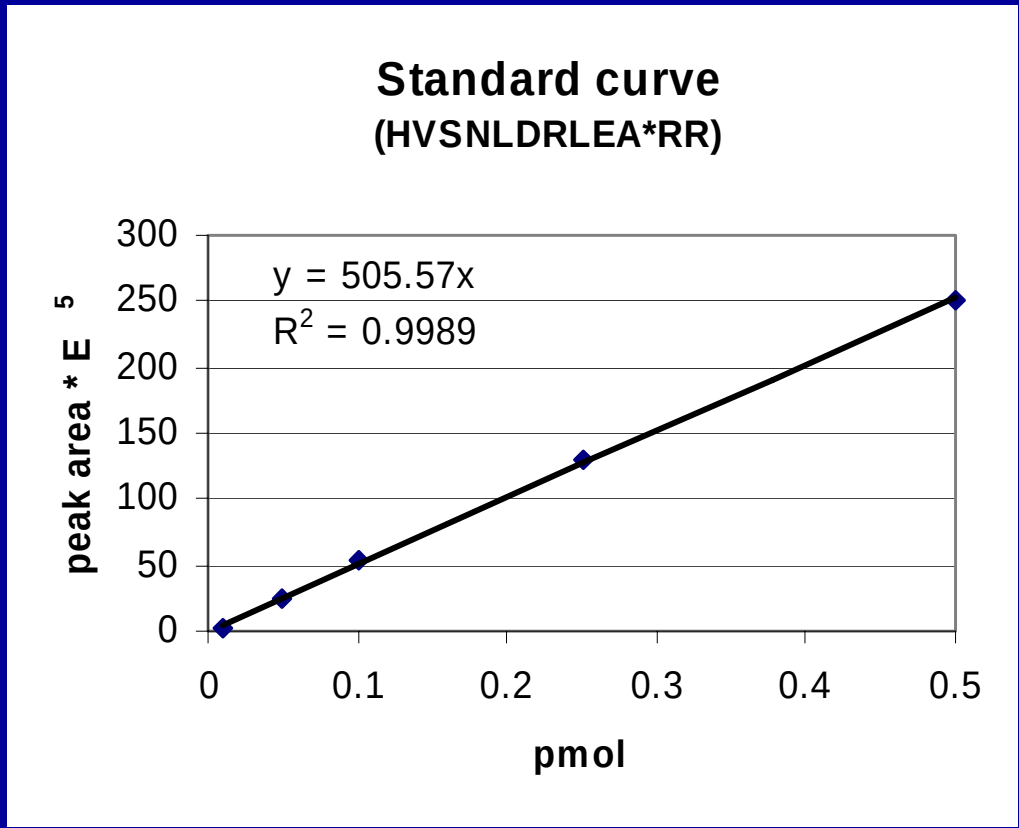


Glinski & Weckwerth (2005) Molecular & Cellular Proteomics

Relative and absolute quantitative shotgun proteomics: targeting low-abundance proteins in *Arabidopsis thaliana*

Wienkoop & Weckwerth JXB 2006





Sucrose Synthase in Komplex-Probe
=> 2fmol/ mg Frischgewicht

Quantitative Phosphoproteomics

Workflow

Starting Material:
Arabidopsis thaliana

Grinding

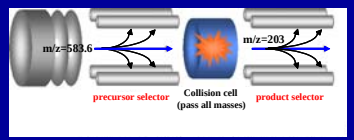
Extraction

Sample Preparation and Incubation

In vitro Phosphorylation

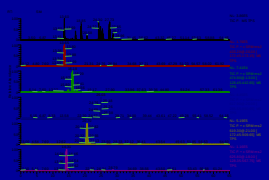
in vitro
phosphorylation

ESI LC-MS/MS

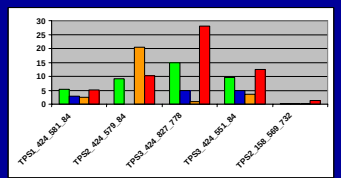


Fully automated MRM Analysis

MRM Transitions



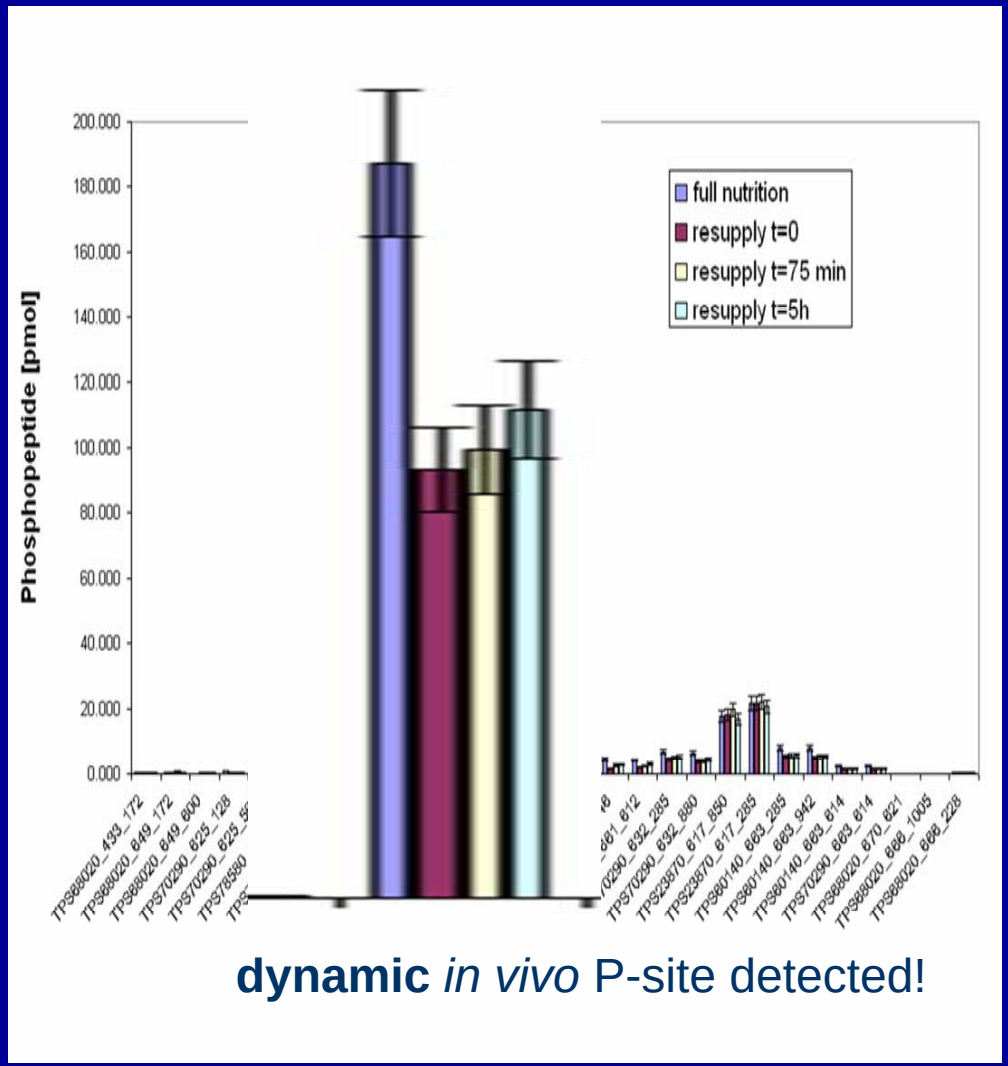
Data Evaluation



Identification and Quantitation of Phosphorylated Peptides

Multiplex kinase target analysis

Trehalose-6-phosphate Synthase gene family: Calcium - dependent and – independent multisite phosphorylation



dynamic *in vivo* P-site detected!

Glinski & Weckwerth Mol. Cell. Prot. 2005

Strategy for the identification and relative quantification of **site-specific** protein phosphorylation using LC/MS^2/MS^3 in an Iontrap MS

Sucrose-Phosphate Synthase in Arabidopsis thaliana

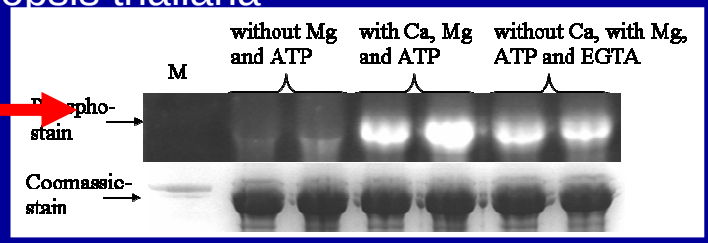
Phosphorylation of the protein of interest under different conditions

Step 1: 'SDS-PAGE and in-gel detection of phosphorylation using a fluorescent dye'

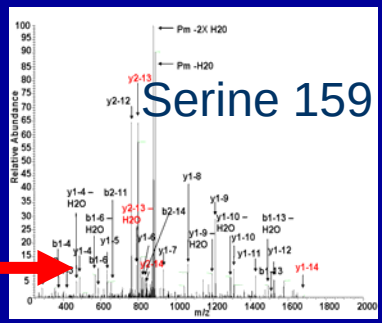
Step 2: 'Determination of phosphorylation sites using neutral loss driven MS^3'

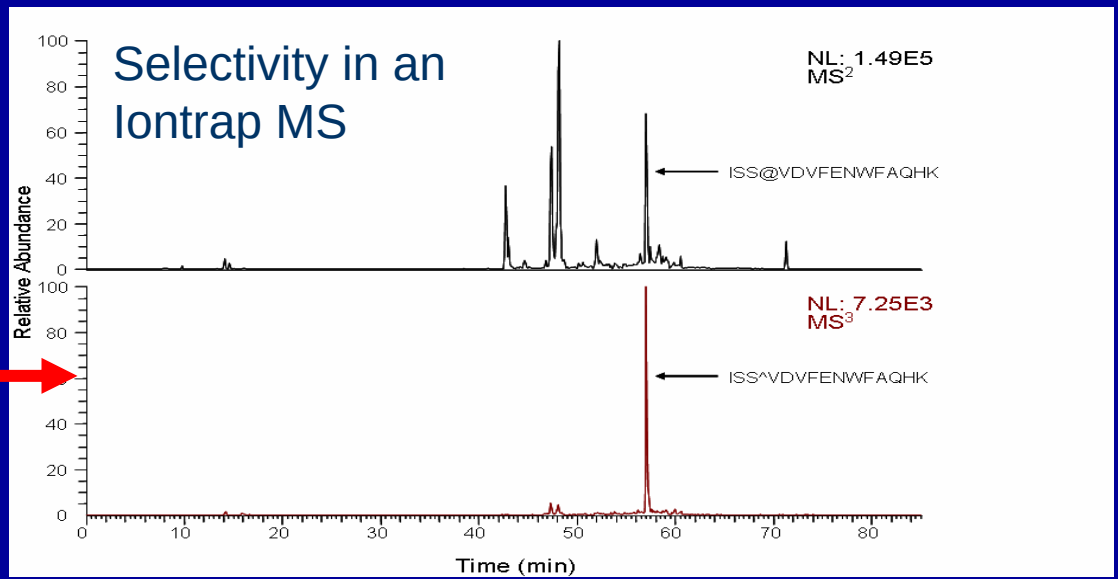
Step 3: 'Library of phosphorylation sites and suitable internal standards'

Step 4: 'Site-specific relative quantification of phosphorylation using MS^2 and MS^3 data'



yeast

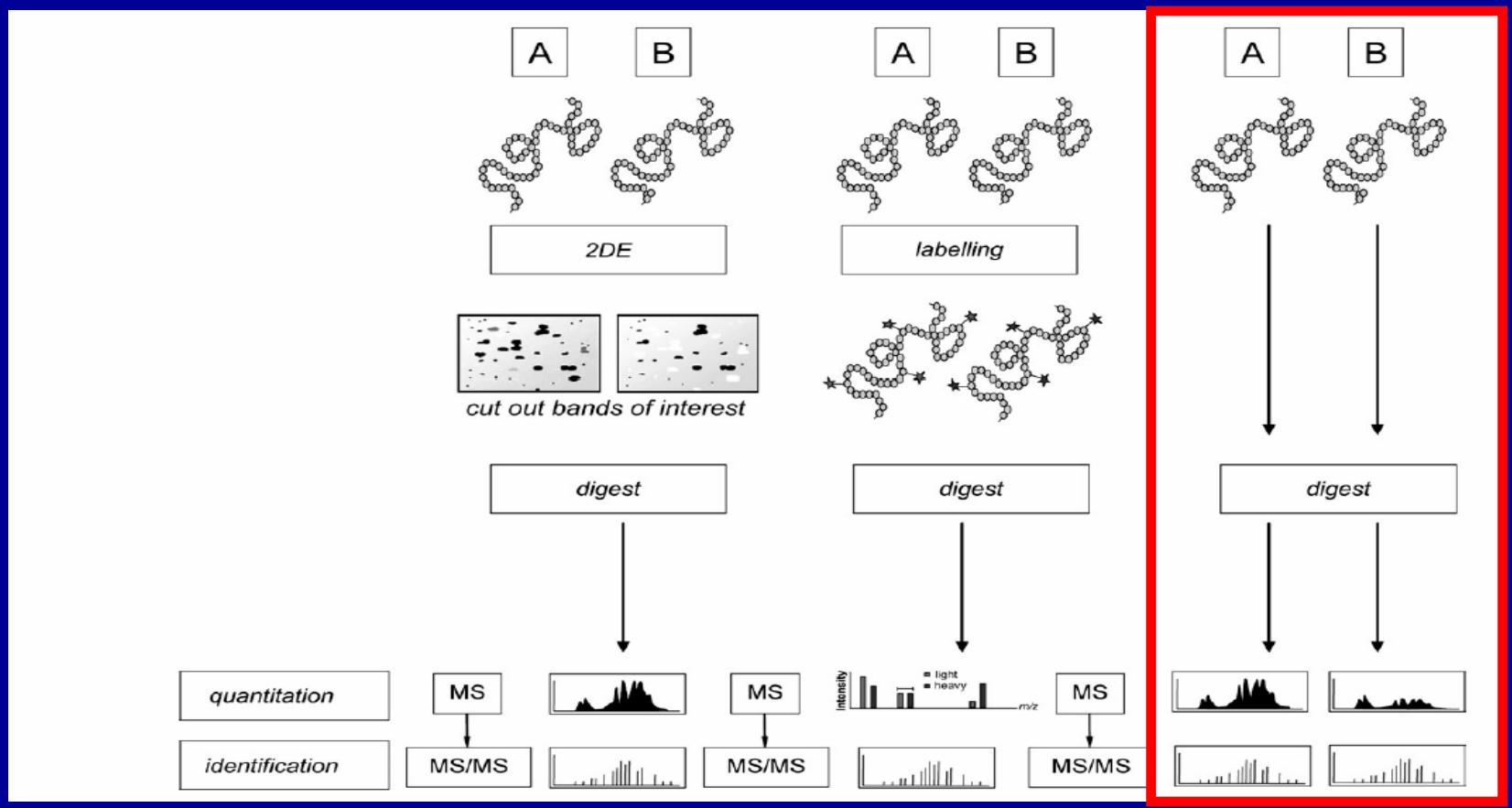




Wolschin et al. 2005 RCM

Wolfgang-Paul-Promotionspreis-Vortrag von Florian Wolschin/Phosphoproteomics am Dienstag, 9.15 Uhr

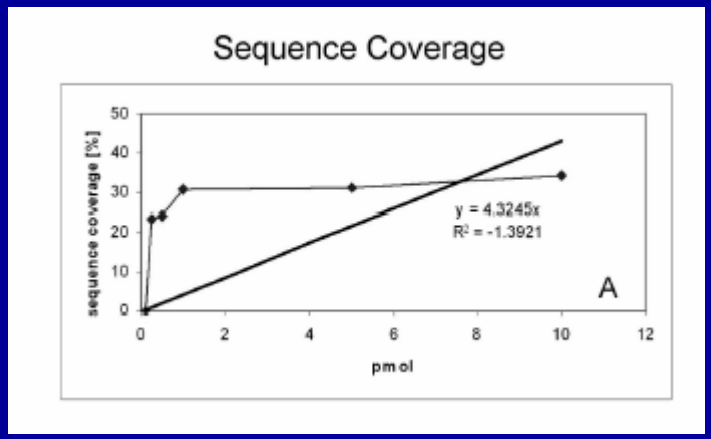
Label-free quantitative proteomics – **non-targeted approach**

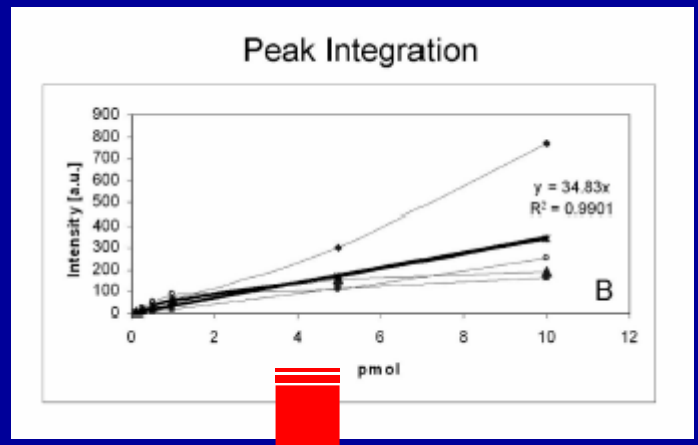


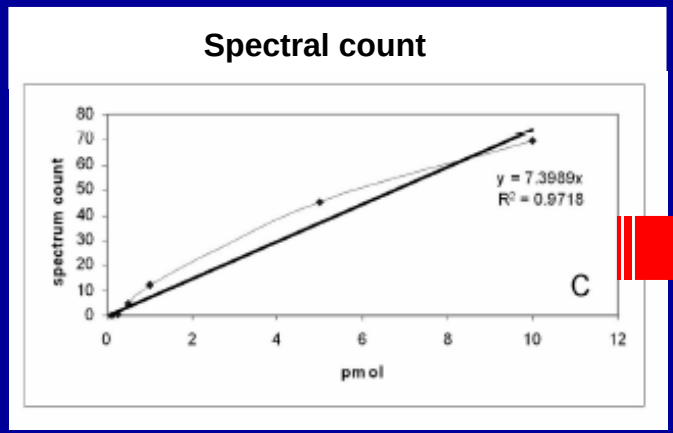
Weckwerth et al. 2004 Proteomics
Morgenthal et al. 2005 Metabolomics
Wienkoop & Weckwerth 2006 JXB
Wienkoop et al. 2006 Journal of Separation Science

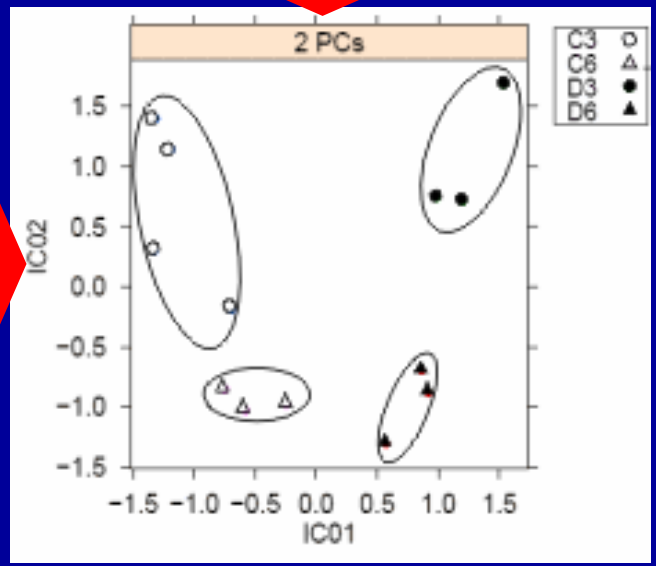
Stable isotope-free quantitative shotgun proteomics combined with sample pattern recognition for rapid diagnostics

- Systematic comparison of protein sequence coverage, peak integration and spectral count





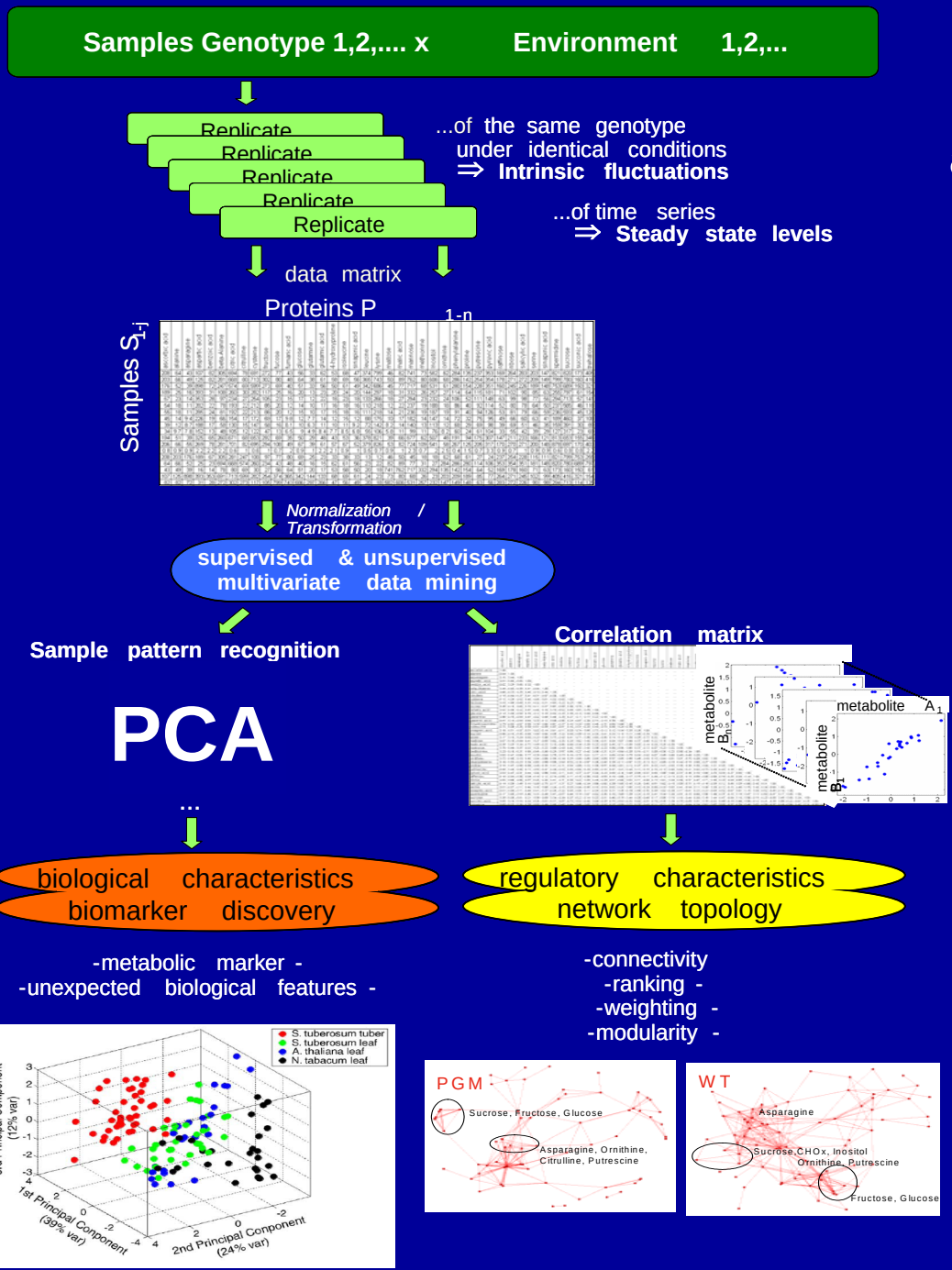




Wienkoop et al. 2006
Journal of Separation Science

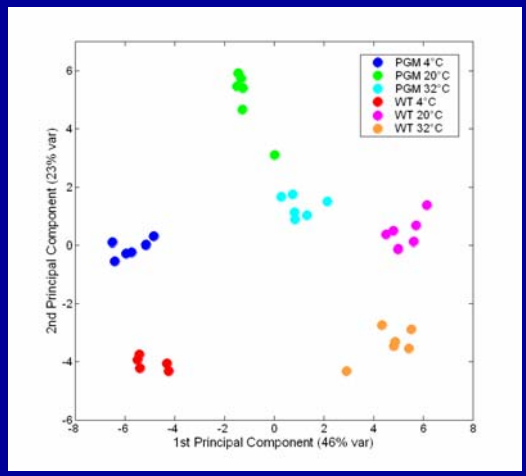
Exkurs: Pattern recognition and biological interpretation

Weckwerth 2003 Annual Review of Plant Biology
Weckwerth et al. 2004 Proteomics
Weckwerth & Morgenthal Drug Discovery Today 2005



Exkurs: Hauptkomponentenanalyse (Principal component analysis PCA)

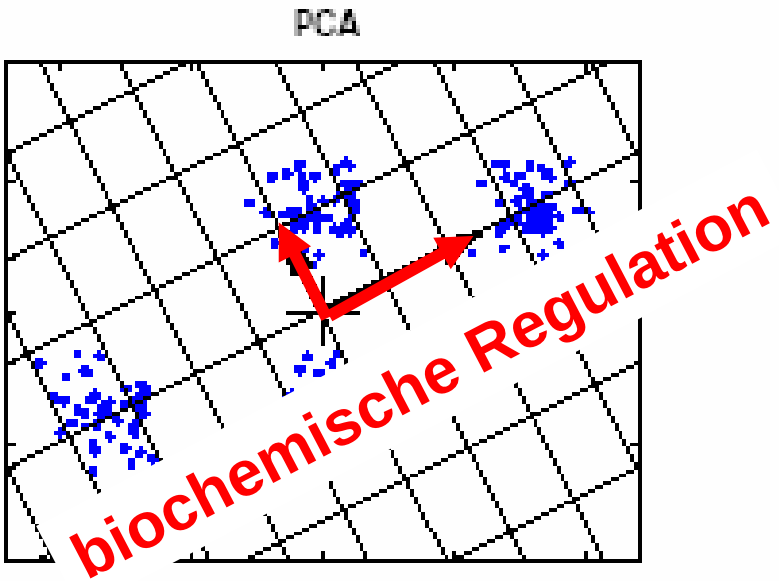
- Die Dimensionalität der Daten wird reduziert, um eine Visualisierung zu ermöglichen (üblicherweise 2-3 neue Komponenten).
- Linearkombinationen aus den Proteinen/Metaboliten (Variablen) mit neuen Faktoren in Abhängigkeit der Varianz (größte Varianz -> hohe Faktoren)
- Diese Linearkombinationen bilden ein neues Koordinatensystem, in dem bis zu 95% der Gesamtvarianz abgebildet ist. Proben, die sich durch hohe Kovarianz ausweisen, werden zusammen gruppiert.
- **Die Korrelation einzelner Proteine wird als Information erhalten – im Gegensatz zur univariaten Mittelwertsanalyse.**

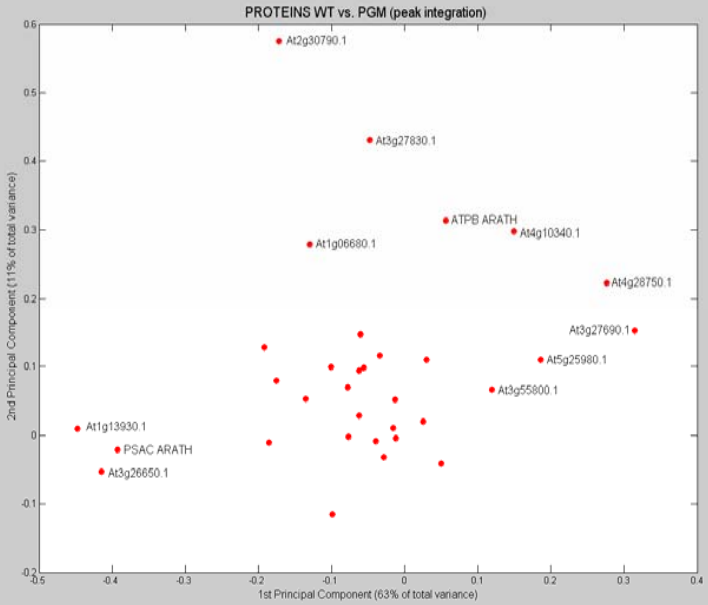


=> Mustererkennung

Analyse der Kovarianz-/Korrelationsstruktur der experimentellen Daten mithilfe der Hauptkomponentenanalyse

=> Mustererkennung in molekularen Daten



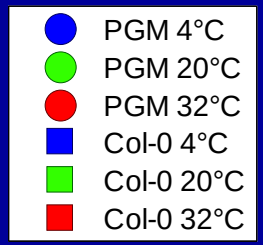


⇒Scores (Mustererkennung) => Loadings (Proteine)

Weckwerth & Morgenthal 2005 Drug Discovery Today

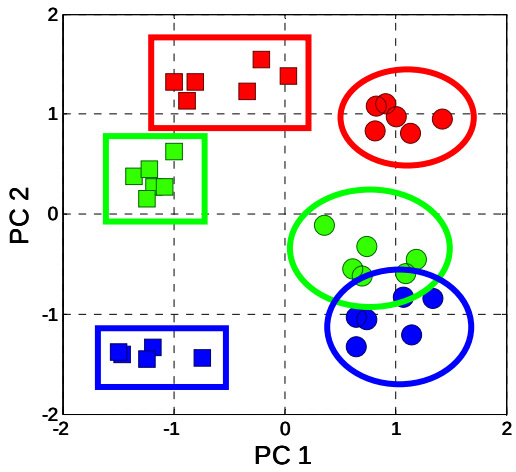
Proteomics-Metabolomics-Datenintegration

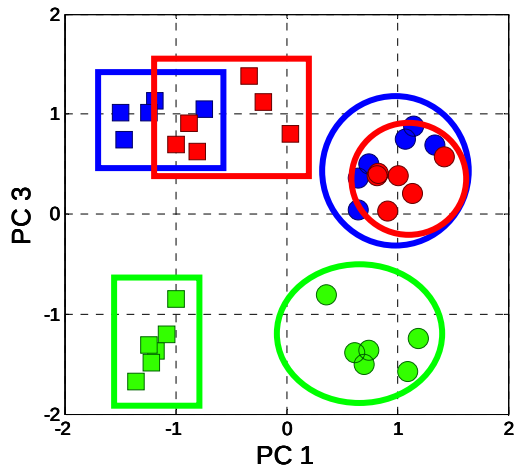
Temperatur-Genotyp-Relationship
Matrix: 36 Proben x 300 Metabolite x 180 Proteine (identified and quantified)

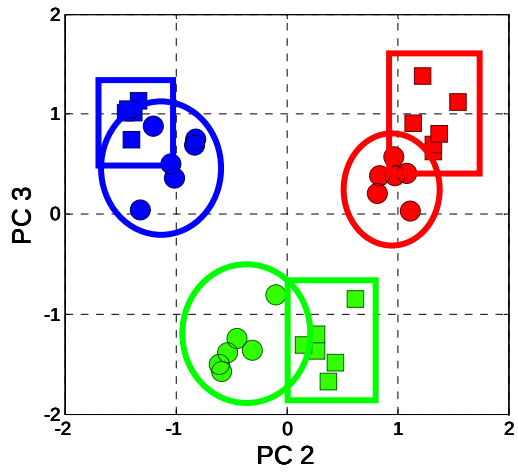


PC1 PGM vs WT
PC2 temperature gradient
PC3 general temperature stress

Integrative *k=4*







=> spezifische Protein-Metabolit-Korrelationen erklären die Komponenten!

Loadings

Metabolites

Proteins