

**Protein Profiling und Metabolite Profiling  
in *Arabidopsis thaliana*  
im systembiologischen Kontext**

vorgelegt von  
Dipl.-Ing.  
Katja Morgenthal

von der Fakultät III Prozesswissenschaften der Technischen Universität Berlin  
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften  
– Dr.-Ing. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. Roland Lauster  
Berichter: Prof. Dr. Roland Tressl  
Dr. Wolfram Weckwerth

Tag der wissenschaftlichen Aussprache: 21. Dezember 2006

Berlin 2007  
D 83

## Danksagung

Die vorliegende Dissertation wurde im Zeitraum von März 2003 bis September 2006 am Max-Planck-Institut für Molekulare Pflanzenphysiologie in Potsdam angefertigt. Für sein mir entgegen gebrachtes Vertrauen und die Möglichkeit, diese Arbeit dort durchzuführen, danke ich herzlich Herrn Prof. Dr. Mark Stitt.

Herrn Dr. Wolfram Weckwerth vom Institut für Molekulare Pflanzenphysiologie in Potsdam möchte ich für die Bereitstellung des Themas und die gewissenhafte Betreuung dieser Arbeit danken - es war nicht leicht, die kleinen Moleküle zu zügeln.

Mein besonderer Dank gilt Herrn Prof. Dr. Roland Tressl vom Institut für Biotechnologie der Technischen Universität, Fachgebiet Molekularanalytik, für die nun schon viele Jahre andauernde großzügige Unterstützung und sein fortwährendes Interesse an meiner Arbeit.

Herzlich danken möchte ich ferner Herrn Dr. Leif-Alexander Garbe. Mit großem Interesse und vielen wertvollen Anregungen hat er den Fortgang dieser Arbeit bei jeder Gelegenheit unterstützt.

Schliesslich danke ich allen Mitgliedern der Arbeitsgruppe, welche durch ihre rege und stete Diskussionsbereitschaft wesentlich zum Gelingen dieser Arbeit beigetragen haben.

# Inhalt

<b>1</b>	<b>Einleitung und Problemstellung.....</b>	<b>9</b>
1.1	<i>Arabidopsis thaliana</i> – Der Modellorganismus der Pflanzen- biochemie.....	12
1.1.1	Die Anatomie der Pflanze.....	13
1.2	Die Analyse des Proteoms – Proteomics.....	16
1.2.1	Gelbasierte quantitative Proteomanalyse .....	22
1.2.2	Gelfreie quantitative Proteomanalyse .....	23
1.3	Die Analyse des Metaboloms – Metabolomics.....	25
1.4	Experimentelles Design und Chemometrik .....	33
1.4.1	Dimensionsreduzierende Verfahren.....	34
1.4.2	Klasseneinteilung .....	39
<b>2</b>	<b>Metabolische Netzwerke .....</b>	<b>42</b>
2.1	Die Organisation metabolischer Netzwerke .....	42
2.2	Metabolite Profiling pflanzlichen Untersuchungsmaterials – Methodenevaluierung.....	43
2.3	Metabolitvariabilität und Korrelationen.....	47
2.4	Differentielle und konservierte Korrelationen .....	48
2.5	Der Ursprung beobachteter Korrelationen – ein numerisches Beispiel .....	51
2.6	Umkehrungen von Korrelationen – ein weiteres numerisches Beispiel .....	54
2.7	Einfluss extrinsischer Parameter auf die Metabolitvariationen .....	55
2.8	Diskussion.....	56
<b>3</b>	<b>Untersuchungen der zeitabhängigen systemischen Regulation des Metabolismus von <i>Arabidopsis thaliana</i>.....</b>	<b>62</b>
3.1	Der circadiane Rhythmus der Pflanze .....	62
3.2	Protein Profiling des circadianen Rhythmus von <i>A. thaliana</i> .....	64
3.2.1	Chemometrische Analyse der Proteindaten.....	67
3.3	Metabolite Profiling des circadianen Rhythmus von <i>Arabidopsis thaliana</i> .....	70
3.3.1	Chemometrische Analyse der Metabolitdaten .....	70
3.4	Der circadiane Rhythmus von <i>A. thaliana</i> im Kontext metabolischer Netzwerke .....	76
3.5	Diskussion.....	78

<b>4</b>	<b>Systemische temperaturbedingte abiotische Stressresponse in <i>Arabidopsis thaliana</i></b> .....	<b>96</b>
4.1	Die Stressbiologie der Pflanze.....	96
4.2	Protein Profiling temperaturgestresster <i>A. thaliana</i> Col-0 und PGM Pflanzen .....	98
4.2.1	Phenolische Extraktion löslicher Proteine mit Saccharose .....	99
4.2.2	Chemometrische Analyse der Proteindaten.....	100
4.3	Metabolit Profiling temperaturgestresster <i>A. thaliana</i> Col-0 und PGM Pflanzen .....	106
4.3.1	Chemometrischen Analyse der Metabolitdaten .....	107
4.4	Diskussion.....	113
<b>5</b>	<b>Integrative Datenanalyse zur Mustererkennung und Identifikation von potenziellen Biomarkern in <i>Arabidopsis thaliana</i></b> .....	<b>130</b>
5.1	Strategien zur Identifikation von Biomarkern .....	130
5.2	Mustererkennung am Beispiel der zeitabhängigen systemischen Regulation des Metabolismus von <i>Arabidopsis thaliana</i> .....	131
5.3	Potenzielle Biomarker der zeitabhängigen systemischen Regulation des Metabolismus von <i>Arabidopsis thaliana</i> .....	135
5.4	Mustererkennung am Beispiel der systemischen temperaturbedingten abiotische Stressresponse in <i>Arabidopsis thaliana</i> .....	137
5.5	Potenzielle Biomarker der systemischen temperaturbedingten abiotische Stressresponse in <i>Arabidopsis thaliana</i> .....	139
5.6	Diskussion.....	142
<b>6</b>	<b>Gerichtete quantitative Proteomics</b> .....	<b>144</b>
6.1	Strategie der gerichteten Proteinanalyse .....	144
6.2	Diskussion.....	147
<b>7</b>	<b>Zusammenfassung und Ausblick</b> .....	<b>150</b>
<b>8</b>	<b>Material und Methoden</b> .....	<b>154</b>
8.1	Materialien und Equipment.....	154
8.1.1	Chemikalien und Referenzverbindungen .....	154
8.1.2	Analytisches Equipment .....	154
8.1.3	Pflanzliches Untersuchungsmaterial und entwicklungs-synchronisierte Anzucht. ....	155
8.1.4	Zellaufschluß .....	156
8.2	Software .....	156
8.3	Integrative Extraktion von Metaboliten und Proteinen .....	156
8.3.1	Metabolitextraktion mit CH <sub>3</sub> OH/CHCl <sub>3</sub> /H <sub>2</sub> O.....	156
8.3.2	Phenolische Proteinextraktion mit Urea .....	157

8.3.3	Phenolische Proteinextraktion mit Saccharose.....	157
8.4	Metabolite Profiling.....	157
8.4.1	Derivatisierung.....	157
8.4.2	GC-TOF-MS basierte Metabolitanalyse.....	157
8.4.2.1	GC-TOF-MS Setup.....	157
8.4.2.2	Qualitätskontrolle.....	158
8.4.2.3	Spektrbibliothek.....	158
8.4.2.4	Semi-automatische Peakidentifikation und Quantifizierung.....	158
8.5	Protein Profiling.....	159
8.5.1	Enzymatischer Proteinverdau.....	159
8.5.2	nanoLC-MS-basierte Shotgun-Proteinanalyse.....	159
8.5.2.1	LC-MS Setup.....	159
8.5.2.2	Proteinidentifikation.....	160
8.5.2.3	Quantifizierung über Peakintegration.....	160
8.5.2.4	Quantifizierung über Spektren-Count.....	160
8.5.3	LC-MS basiertes Multiple Reaction Monitoring.....	161
8.5.3.1	LC-MS Setup.....	161
8.6	Chemometrische Analyse.....	162
8.6.1	Konstruktion metabolischer Netzwerke.....	162
8.6.2	Dimensionsreduktion.....	162
8.6.3	Clustering.....	163
8.6.4	Varianzanalyse.....	163
<b>9</b>	<b>Referenzen.....</b>	<b>164</b>
<b>10</b>	<b>Anhang.....</b>	<b>185</b>

## Abkürzungen

ACL	ATL-Citrat Lyase
ACN	Acetonitril
AGI	Arabidopsis Genome Initiative
AIC	Analytical Ion Chromatogram
ANOVA	Varianzanalyse (Analysis of Variance)
APCI	Athmospheric Pressure Chemical-Ionization
APPI	Athmospheric Pressure Photo-Ionization
APX	Ascorbat Peroxidase
CHO	Carbohydrat
CI	chemische Ionisierung
Col-0	Columbia-0
COR	cold responsive genes
CV	Variationskoeffizient (coefficient of variation)
2D	zweidimensional
DMFA	Dimethylformamide
Da	Dalton
DESI	Desorptions-Electrospray-Ionization
DIGE	2D Difference Gel Electrophoresis
DTT	Dithiothreitol
EI	Elektronenstoss-Ionisierung
ESI	Electrospray Ionisation
eV	Elektronenvolt
FBPase	Fructose 1.6-bisphosphatase
FT	Fourier-Transformation
FW	Frischgewicht
GAP-DH	Glyceraldehyd 3-phosphat Dehydrogenase
GC	Gaschromatographie
GR	Glutathion Reduktase
GR-RBP	Glycin-reiches RNA-bindendes Protein
GST	Glutathion S-Transferase
h	Plancksches Wirkungsquantum $h=6,6260755 \pm 0,0000040 \cdot 10^{-34}$ J-sek
HPLC	Hochleistungsflüssigchromatographie
HSP	Hitzeschockprotein
IAA	Iodoacetamid
ICA bzw. IC	Unabhängige Komponentenanalyse (Independent Components Analysis) bzw. Unabhängige Komponente
ICAT™	Isotope-Coded Affinity Tag
ICP	Inductively Coupled Plasma
ICPL™	Isotope-Coded Protein Labeling
ICR	Ionen Cyclotron Resonanz
ID	Innendurchmesser
IEF	isoelektrische Fokussierung
J	Joule
kV	Kilovolt
LIT	lineare Ionenfalle (linear iontrap)
m/z	Masse zu Ladungs-Verhältnis
MALDI	Matrix-assisted Laser Desorption Ionization

min	Minuten
MRM	Multiple Reaction Monitoring
MS	Massenspektrometrie
MSTFA	<i>N</i> -Methyl- <i>N</i> -trimethylsilyltrifluoracetamid
p	Irrtumswahrscheinlichkeit (zur Verwerfung der Nullhypothese)
PAGE	Polyacrylamidgelelektrophorese
PAP	plastidiäres Lipid-assoziiertes Protein
PCA bzw. PC	Hauptkomponentenanalyse (Principal Components Analysis) bzw. Hauptkomponente
PGM	Phosphoglucomutase
Q	Quadrupol
R	Regressionskoeffizient
$r_{xy}$	Pearson'scher Produktmoment-Korrelationskoeffizient
RI	Retentionsindex
ROS	Reaktive Sauerstoffspezies (reactive oxygen species)
RT	Raumtemperatur oder Retentionszeit
RubisCO	Ribulose 1,5-bisphosphat Carboxylase/Oxygenase
SDS	sodium dodecylsulfat
sek	Sekunde
SOD	Superoxid Dismutase
TIC	Total Ion Chromatogram
TOF	Flugzeit (Time-of-flight)
UPLC™	Ultrahochdruckflüssigchromatographie
UV	ultraviolett
Xcorr	Cross Correlation Coefficient
XIC	Extracted Ion Chromatogram

### Einbuchstabencode der wichtigsten proteinogenen Aminosäuren

A	Alanin	G	Glycin
V	Valin	S	Serin
L	Leucin	T	Threonin
I	Isoleucin	C	Cystein
F	Phenylalanin	Y	Tyrosin
W	Tryptophan	N	Asparagin
M	Methionin	Q	Glutamin
P	Prolin	K	Lysin
D	Asparaginsäure	R	Arginin
E	Glutaminsäure	H	Histidin



## 1 Einleitung und Problemstellung

Die natürliche Variabilität zwischen individuellen Organismen war lange Zeit Gegenstand der genetischen Forschung. Mit der Entschlüsselung einiger tausend Nucleotide, welche erstmals 1977 in den von Sanger *et al.* durchgeführten Untersuchungen des Bakteriophagen phiX174 gelang [1], begann eine Ära der systematischen Genomanalyse verschiedenster Organismen. Diese vielleicht wichtigste Methode der Molekularbiologie, die Sequenzanalyse, wurde 1980 mit dem Nobelpreis ausgezeichnet. Im Jahre 1982 konnten Sanger *et al.* die vollständige DNA-Sequenz des Bakteriophagen Lambda, welcher über 48.502 Nucleotidpaare verfügt, entschlüsseln [2]. Die erste komplette Nucleotidsequenz eines Organismus wurde 1995 veröffentlicht; das 1.830.137 Nucleotidpaare umfassende Genom des Gram-negativen Bakteriums *Haemophilus influenza* war entschlüsselt [3]. Seither sind die Genomsequenzen von 413 Organismen vollständig bekannt (<http://www.genomesonline.org/>, Stand 30. August 2006), darunter u.a. das Genom des *Escherichia coli* Stammes K12, welches für ca. 4.300 Gene kodiert [4], das Genom der Bäckerhefe *Saccharomyces cerevisiae* mit etwa 6.000 Genen [5,6], das von *Arabidopsis thaliana*, welches für ca. 25.000 Gene kodiert [7] sowie das Genom des Menschen, von dem zunächst angenommen wurde, dass dieses 35.000 Gene umfasst. Heute jedoch geht man davon aus, dass das menschliche Genom für ca. 25.000 Gene kodiert [8,9]. Damit entspricht die Anzahl der den Menschen charakterisierenden Gene der in *Arabidopsis* und es wird unmittelbar deutlich, dass die Anzahl der Gene keine Aussagen über den Grad der Komplexität eines Organismus zulässt.

In der Zelle wird die in der DNA niedergeschriebene Information mit Hilfe von RNA-Polymerasen zunächst auf die mRNA übertragen, welche als molekulare Zwischenstufe das Template für die Proteinbiosynthese an den Ribosomen darstellt. Mit der Entdeckung des alternativen Splicings der RNA der Eukaryoten Ende der siebziger

Jahre [10], welches die wahlweise Entfernung einzelner Introns oder ganzer Intron-Exon-Intron-Komplexe bezeichnet, so dass aus einem Gen verschiedene mRNA's und somit Proteine entstehen können, ist auch die von Beadle *et* Tatum nach Experimenten mit dem Pilz *Neurospora crassa* 1941 formulierte Theorie der ein-Gen-ein-Enzym-Hypothese [11], welche wenig später auf die ein-Gen-ein-Protein-Hypothese erweitert wurde, nicht mehr haltbar. Man nimmt heute an, dass etwa 50-60% der im humanen Genom codierten Gene im Rahmen der Transkription durch unterschiedliche Splicevorgänge prozessiert werden [12]. In Pflanzen kommt diese Form der RNA-Prozessierung weniger häufig vor. Man hat bis dato für etwa 7% der in *Arabidopsis* vorkommenden Gene Splicevarianten nachgewiesen [13]. In jedem Falle übersteigt jedoch die Anzahl der verschiedenen Proteinspezies die der proteincodierenden Gene um ein Vielfaches. Die Entdeckung der reversen Transkriptase von Baltimore und Temin im Jahre 1970 [14,15] hat weiter dazu geführt, dass auch das zentrale Dogma der Molekularbiologie, demzufolge die genetische Information von der DNA über die mRNA auf die Proteine übertragen wird und nicht umgekehrt, seinen dogmatischen Charakter verloren hat. Somit wurde deutlich, dass die Diversität der auf der Erde vorkommenden Organismen nicht allein durch die in der DNA festgeschriebenen Informationen zu erklären ist.

Durch die zahlreichen abgeschlossenen Genomsequenzierungsprojekte wuchs das Wissen über DNA- und Proteinsequenzen in den letzten Jahren deutlich schneller, als das über die jeweiligen biochemischen Funktionen. Ein prominentes Beispiel dieser Diskrepanz stellt die Ackerschmalwand *A. thaliana* dar. Seit der vollständigen Sequenzierung des Genoms dieser zur Familie der *Brassicaceae* gehörenden Pflanze im Jahre 2000 sind bis heute nur etwa 30% ihrer Gene funktionell charakterisiert [7]. In der sog. Post-Genom-Ära geht es nun um die Aufklärung von Gen-Funktions-Zusammenhängen. Vollständige Genomsequenzen bieten dabei die Möglichkeit, die in der DNA verschlüsselten Informationen über die Struktur der mRNA und der Proteine zu lesen. Da jedoch in jeder einzelnen Zelle eines Organismus immer nur ein Bruchteil der Gene in Abhängigkeit vom momentanen Wachstums-, Differenzierungs- oder physiologischen Zustand aktiv ist, erfordert die Funktionsaufklärung zunächst die Ermittlung, welche Gene unter welchen Bedingungen exprimiert werden. Mit Hilfe der in den letzten Jahren durchgeführten und zumeist auf Microarray Technologien beruhenden Genexpressionsanalysen konnten in verschiedenen Organismen zahlreiche, unter bestimmten Bedingungen differentiell exprimierte Gene identifiziert werden [16-20]. Der Vergleich dieser Analysen mit den mittels der Proteomics-Technologie, deren Ziel es ist, die Vielfalt der in einem Organismus vorkommenden Proteine zu erfassen und deren Funktionen und komplexes Zusammenspiel zu ermitteln, gewonnenen Daten erbrachte jedoch, dass der mRNA-Level nicht notwendigerweise mit der Abundanz oder der

Aktivität eines Proteins korreliert [21,22]. Dieses liegt zum einen daran, dass aus den oben erwähnten alternativen Splicevorgängen zahlreiche Proteinvarianten eines einzigen Genabschnittes resultieren können, zum anderen wird die Komplexität der Stoffwechselebene der Proteine weiter erhöht durch posttranslationale Modifikationen wie Phosphorylierungen, Glycosylierungen, Sulfatierungen, Methylierungen, Ubiquitinylie-runen, Acetylierungen oder Myristoylierungen [23-26].

Neben den Proteinen, deren Bedeutung für die zelluläre Organisation, Regulation und Signaltransduktion unumstritten ist, spielen ebenso die Metabolite eine zentrale Rolle im komplexen Zusammenspiel zwischen den Genen, Transkripten und Proteinen, welches letztendlich den Phänotyp eines Organismus bestimmt. Die Metabolite stellen die Endprodukte der Genexpression dar und nehmen unmittelbar Einfluss auf die Mechanismen zur Regulation der Expression von Genen und Proteinen, so dass die Organisation des Stoffwechsels über ein komplexes Netzwerk aus sich überlappenden Aktivitäten von Metaboliten und Enzymen, deren Struktur in den Genen codiert ist und deren Substrat wiederum die Metabolite darstellen, erfolgt. Das globale Verständnis dieses komplexen Zusammenspiels, welches letztlich die Transition vom Genotyp zum Phänotyp bedingt, stellt die wohl grösste Herausforderung der modernen Biowissenschaften dar. Die Vorhersage des Phänotyps eines Organismus wird dabei jedoch häufig durch ein hohes Mass an phänotypischer Plastizität, welche das Vermögen eines Organismus beschreibt, auf bestimmte Umweltbedingungen mit einem veränderten Phänotyp zu reagieren, erschwert [27].

Mit der Erkenntnis, dass biologische Abläufe nur über eine umfassende Analyse des gesamten komplexen biologischen Systems erklärt werden können, entwickelte sich das noch junge Wissenschaftsgebiet der Systembiologie, dessen Ziel es ist, ein integriertes Bild über alle regulatorischen Prozesse eines Organismus zu bekommen. Die Analyse des Proteoms sowie die des Metaboloms stellen zwei Schlüsseltechnologien der Systembiologie dar. Als unverzichtbares Verfahren im Hinblick auf die qualitative und quantitative Analyse dieser beiden Stoffwechselebenen erwies sich die Massenspektrometrie. Im Verlaufe der letzten Jahre haben sich verschiedene Massenspektrometrie-basierte Strategien zur Protein- und Metabolitanalyse biologischen Materials herausgebildet. In Analogie zu der von Fiehn geprägten Nomenklatur [28] enthält die Tabelle 1.1 eine Übersicht über die verschiedenen Strategien zur Analyse des Metaboloms, welche bedingt auch bei die Analyse des Proteoms Anwendung finden.

In den folgenden Absätzen werden einige bedeutende Aspekte der Protein- und Metabolitanalyse pflanzlichen Untersuchungsmaterials diskutiert. Der Fokus dieser Ausführungen liegt dabei auf der Analyse des Modellorganismus *A. thaliana*, welche einen Schwerpunkt der vorliegenden Arbeit darstellt.

Tabelle 1.1: Strategien zur Analyse des Metaboloms

Strategie	Beschreibung
Metabolomics	Quantitative oder qualitative Analyse aller in einer biologischen Probe vorkommenden Metabolite
Metabonomics	Quantitative Analyse der Response auf genetische Modifikationen oder pathophysiologische Stimuli
Metabolit Target Analyse	gezielte Quantifizierung ausgewählter Metabolite
Metabolite Profiling	Quantitative oder qualitative Analyse einer Gruppe von strukturell ähnlichen oder in einem ausgewählten biochemischen Pathway vorkommenden Metabolite
Metabolic Fingerprinting	Klassifizierung von Proben durch schnelle, globale Analyse biologischen Materials ohne umfassende Identifizierung einzelner Metabolite
Metabolic Footprinting	Klassifizierung von Metaboliten, welche aus dem Intrazellularräum von Mikroorganismen ins Medium sekretiert werden

## 1.1 *Arabidopsis thaliana* – Der Modellorganismus der Pflanzenbiochemie

*Arabidopsis thaliana* ist ein zur Familie der *Brassicaceae* gehörendes, einjähriges blühendes Wildkraut (Abbildung 1.1.1), dessen erste Beschreibungen auf das 16. Jahrhundert zurückgehen. Nachdem Friedrich Laibach bereits 1907 die korrekte Anzahl der in *Arabidopsis* vorkommenden Chromosomen veröffentlichte [29], fasste er 1943 das Potenzial von *A. thaliana* als Modellorganismus für die genetische Forschung zusammen. Laibach erkannte die Vorzüge der hohen Fruchtbarkeit, des kurzen Generationscyclus, der leichten Kultivierbarkeit und der geringen Raumbeanspruchung für genetische und entwicklungsphysiologische Untersuchungen und bezeichnete *A. thaliana* als ein Untersuchungsobjekt, welches "in mancher Beziehung dem Paradeobjekt der Genetiker, der *Drosophila*, nahe kommt" [30].



**Abbildung 1.1.1: Verschiedene Entwicklungsstadien von *A. thaliana* Col-0.** Gezeigt sind das etwa 6 Wochen nach der Aussaat erreichte Rosettenstadium (**Abb. 1.1.1-A**), das etwa 8 Wochen nach der Aussaat erreichte frühe Blühstadium (**Abb. 1.1.1-B**) sowie das etwa 9 bis 10 Wochen nach der Aussaat erreichte späte Blühstadium (**Abb. 1.1.1-C**).

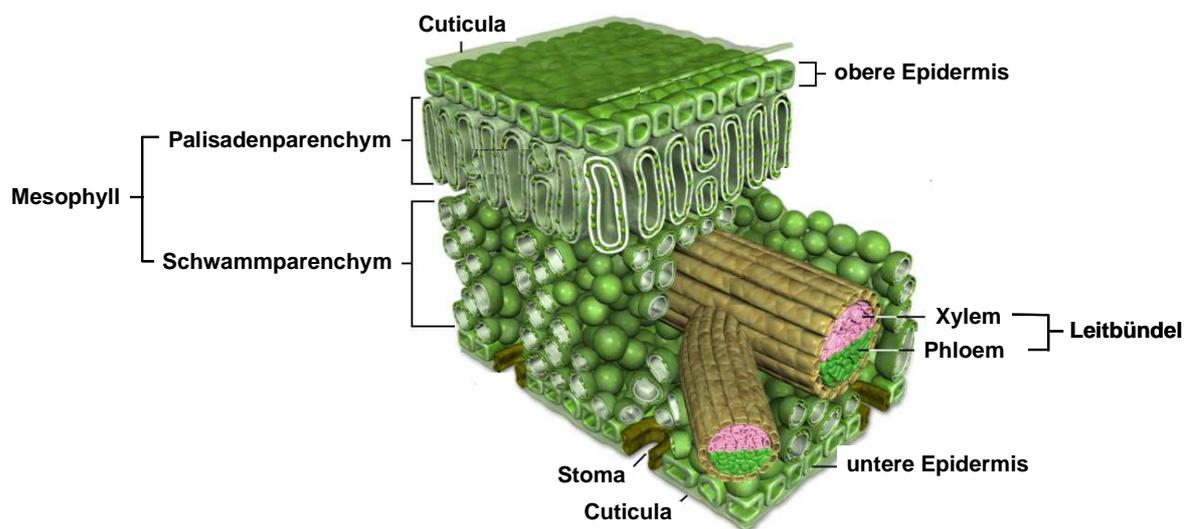
Mit der Veröffentlichung der vollständigen Nucleotidsequenz des Genoms im Jahre 2000 [7], welche das Ergebnis einer beispiellosen Kooperation von Pflanzenforschern aus der ganzen Welt darstellte, erlebte *A. thaliana* ein Comeback. Das kleine, 5 Chromosomen umfassende Genom ist heute ausführlich kartiert und die zur Verfügung stehenden effizienten Transformationstechniken mit *Agrobacterium tumefaciens* ermöglichen die gezielte Mutagenese [31,32].

Wenngleich *A. thaliana* keine grosse wirtschaftliche Bedeutung zukommt können aufgrund der engen Verwandtschaft mit Kulturarten wie Kohl (*Brassica oleracea*) oder Raps (*Brassica napus*) die für *Arabidopsis* erarbeiteten Informationen über funktionale Zusammenhänge zwischen Genotyp, Umwelt und Phänotyp auf Nutzpflanzen übertragen werden.

### 1.1.1 Die Anatomie der Pflanze

Im Gegensatz zu einzelligen Organismen weisen Pflanzen drei verschiedene Organe, namentlich Blatt, Stängel und Wurzel auf. Jedes dieser Organe setzt sich weiter aus verschiedenen Gewebeverbänden zusammen, innerhalb derer die einzelnen Zellen weiter kompartimentiert sind. Die wichtigsten zur Photosynthese befähigten Organe stellen die Blätter dar. Am Aufbau eines Laubblattes sind unabhängig vom Blatttyp die charakteristischen Elemente Epidermis, Mesophyll und Leitbündel beteiligt (Abbildung 1.1.1.1). Die meist einschichtige Epidermis bildet dabei den Abschluss des Blattes nach aussen und besitzt eine wasserundurchlässige Wachsschicht, die Cuticula, welche die Pflanze vor Austrocknung durch unkontrollierte Transpiration schützt. Die Epidermiszellen besitzen in der Regel keine Chloroplasten. Eine Ausnahme stellen die meist in die untere Epidermis eingebetteten Schliesszellen, die sog. Stomata, dar, welche der Regulation des Gasaustausches dienen und immer auch Chloroplasten enthalten. Das zwischen der oberen und der unteren Epidermis gelegene Blattgewebe bilden chloroplastenhaltige Mesophyllzellen, in denen das durch die Stomata aufgenommene CO<sub>2</sub> mit Hilfe der Lichtenergie im Zuge der Photosynthese und des anschliessenden Calvin Cyclus assimiliert wird. Das Mesophyll, auch Blattparenchym genannt, setzt sich aus dem unter der oberen Epidermis gelegenen Palisadenparenchym, welches aus mehreren Lagen langgestreckter chloroplastenreicher Zellen besteht, und dem darunter gelegenen Schwammparenchym aus unregelmässig geformten Zellen mit grossen Interzellularräumen zusammen. Das gesamte Blatt ist dicht von Leitbündeln durchzogen, welche von der Sprossachse abzweigen und blind im Mesophyll enden. Die Hauptaufgabe der Leitbündel, welche aus den zwei Gewebeverbänden Xylem und Phloem zusammengesetzt sind, besteht im An- und Abtransport von Wasser und Nährstoffen. Auch die in höheren Pflanzen vorrangig im photosynthetisch hochaktiven

Palisadenparenchym synthetisierten Assimilate werden in geeigneter chemischer Form, z.B. als Saccharose, vom Ort ihrer Entstehung, den sog. Source-Bereichen, zu den Orten ihres Verbrauches, den sog. Sink-Bereichen des Vegetationskörpers wie etwa den Wurzeln transportiert [33-35].



**Abbildung 1.1.1.1: Schematische Darstellung eines Blattquerschnittes.**

(aus <http://micro.magnet.fsu.edu/cells/leaftissue/leaftissuelarge.html> )

### **Die Zelle als morphologisches System**

Durch fortwährende Wachstums- und Differenzierungsvorgänge ist die pflanzliche Zelle zeitlebens durch morphologische Veränderungen gekennzeichnet. Die pflanzliche Zelle stellt ein komplexes, aus mehreren Kompartimenten mit unterschiedlicher metabolischer Funktion zusammengesetztes Gebilde dar. Im Gegensatz zur tierischen Zelle, welche mit ihrer Umgebung weitgehend isoosmotisch ist, bedarf die pflanzliche Zelle, welche in ihrer Größe der tierischen Zelle weit überlegen ist, zu ihrer mechanischen Stabilisierung einer reissfesten Zellwand. Diese ist vorrangig aus Polysacchariden und Proteinen aufgebaut und stellenweise porenartig durch Plasmodesmen unterbrochen. Durch die Plasmodesmen sind nahezu alle Zellen eines Gewebeverbandes miteinander verbunden. Grundsätzlich können in den Verbänden pflanzlicher Gewebe zwei Räume unterschieden werden. Während die Gesamtheit aller Protoplasten einschliesslich der sie verbindenden Plasmodesmen als Symplast bezeichnet wird, fasst man die Gesamtheit der extraplasmatischen Räume unter dem Begriff Apoplast zusammen. Der den Symplast

umgebende Apoplast enthält die Zellwände und stellt einen Diffusionsraum für Wasser und kleine bis mittlere Moleküle dar.

Im Verlaufe der Evolution entstanden Pflanzenzellen aus endosymbiontischen Systemen und stellen somit genetische Chimäre dar, innerhalb derer die drei Genome der Ausgangszellen, das Kerngenom, das Mitochondriengenom und das Plastidengenom, weitgehend erhalten geblieben sind. Innerhalb des zellinneren Cytoplasmas ist die pflanzliche Zelle durch eine Reihe von membranumschlossenen Zellorganellen gekennzeichnet, denen bestimmte metabolische Funktionen zukommen.

Das grösste Organell der ausgewachsenen Zelle stellt eine zentrale Vakuole dar, welche bis zu 90% des gesamten Zellvolumens einnehmen kann. Die Vakuole dient der Pflanze als vorübergehender Speicher von Reservestoffen wie Zuckern, Aminosäuren oder Lipiden. Eine weitere bedeutende Funktion der Vakuole besteht in der Aufnahme von Stoffen, die in der Zelle als Abfallprodukte anfallen und zum Teil toxisch sind wie etwa Alkaloide oder Phenole. Neben der Funktion als intrazelluläres Wasser- und Stoffreservoir ist die Vakuole massgeblich an der Strukturgebung der gesamten Pflanze beteiligt. Mit der Aufrechterhaltung des Zellinnendruckes, welcher in mechanischer Wechselwirkung mit dem Gegendruck der Zellwand die Turgeszenz der Pflanzen bestimmt, erfüllt die Vakuole essentielle Funktionen für das Streckungswachstum. Die reine Volumenzunahme durch die Vergrösserung der Vakuole stellt für die Pflanze eine kostengünstige und effiziente Form des Wachstums dar.

Das zweitgrösste Organell stellen die Chloroplasten dar, welche in ausgewachsenen Zellen etwa 20% des Zellvolumens einnehmen und denen als Träger der photosynthetischen Komponenten neben der Synthese von Stärke eine besondere Bedeutung zukommt.

Daneben verfügt die pflanzliche Zelle über Mitochondrien, deren Hautaufgabe in der Bereitstellung von Reduktionsäquivalenten und ATP im Rahmen des Citratcyclus und der Zellatmung zu sehen ist, und über Peroxisomen, in denen der oxidative Abbau der Fettsäuren unter Freisetzung von Acetyl-CoA stattfindet. Darüber hinaus findet man in pflanzlichen Zellen weitere Plastiden wie z.B. Leukoplasten, welche eine Vielzahl von ausdifferenzierten Plastiden mit unterschiedlicher Funktion umfassen, Amyloplasten, welchen die Funktion als Stärkespeicher in nicht-grünen Geweben wie Wurzeln, Knollen oder Samen zukommt, oder Chromoplasten, welche aufgrund ihres hohen Gehaltes an Carotinoiden rot, orange oder gelb gefärbt sind und dessen Funktion bislang unklar ist [33-35].

### ***Die Zelle als metabolisches System***

Die die Zelle aufbauenden Moleküle und Molekülaggregate haben eine Lebensdauer, die meist sehr viel kürzer ist, als die Lebensdauer der Zelle selbst. Diese Tatsache bedingt

einen beständigen Auf- und Abbau verschiedenster Inhaltsstoffe, welche die Zelle zu einem stofflich hochgradig dynamischen Gebilde machen. Im Hinblick auf das anabolische und katabolische Stoffwechselgeschehen hat auch die Kompartimentierung der pflanzlichen Zelle weit reichende Konsequenzen. So können beispielsweise Aminosäuren in relativ grossen Mengen in der Vakuole gespeichert werden, wo sie metabolisch weitgehend inaktiv sind. Im Cytoplasma hingegen stellen sie hochreaktive Ausgangsverbindungen für die Proteinbiosynthese dar. Neben den morphologischen Kompartimenten existieren auch metabolische Kompartimente, welche nicht immer membranumschlossen sein müssen. Auch innerhalb der meisten Organellen wie Chloroplasten oder Mitochondrien lassen sich metabolische Kompartimente definieren, welche Zellbereiche darstellen, innerhalb derer für bestimmte Moleküle homogene Reaktionsbedingungen herrschen. Diese Kompartimentierung dient vorrangig der Ordnung und Kanalisation des Stoffwechselgeschehens und hat zur Folge, dass ein und die selbe Substanz in der Zelle in mehreren Pools, wie etwa an der Oberfläche von Membranen oder im Inneren von Organellen, in verschiedenen Konzentrationen vorkommen kann. Somit geht mit der quantitativen Bestimmung der Metabolit- und Proteinkonzentration eines biologischen Extraktes, welcher durch Homogenisierung ganzer Zellen, Zellverbände oder Organe hergestellt wird, immer auch ein Informationsverlust einher. Auch Experimente zur Aufklärung von Stoffwechselwegen mit radioaktiv markierten Vorstufen können im Hinblick auf die Aussagefähigkeit der erzielten Ergebnisse problematisch sein, da die markierten Vorstufen von aussen appliziert werden müssen und der Grad der dadurch hervorgerufenen Verdünnung in den morphologischen oder metabolischen Kompartimenten schwer abzuschätzen ist [33-35]. Da diese Limitierungen jedoch mit den heutigen präparativen Techniken nur schwer umgangen werden können, sind an die Interpretation der zumeist durch die Analyse von Extrakten ganzer Organe oder Zellverbände gewonnenen Ergebnisse besondere Anforderungen gestellt.

## 1.2 Die Analyse des Proteoms – Proteomics

Im Gegensatz zum Genom, welches ein statisches Gebilde darstellt und durch die Abfolge, Art und Anzahl seiner Nucleotide genau definiert ist, ist das Proteom durch ein hohes Mass an Dynamik gekennzeichnet. Da in einem Organismus immer nur ein Teil der Gene aktiv ist, können die aktuell vorhandene Menge der Proteine, die Anzahl der unterschiedlichen Proteinspezies sowie der Grad der posttranslationalen Modifikationen als Sensor für verschiedene physiologische Zustände oder Umweltbedingungen aufgefasst werden.

Das Interesse für die Analyse der komplexen Funktionszusammenhänge von Proteinen weckten O'Farrell und Klose 1975 mit ihren Veröffentlichungen der ersten zweidimensionalen Elektropherogramme [36,37]. Die Charakterisierung und systematische Analyse der aufgetrennten Proteine erwies sich jedoch mit den damals zur Verfügung stehenden analytischen Methoden als nahezu unmöglich. Erst in den letzten Jahren führten die enormen Entwicklungen auf dem Gebiet der instrumentellen Analytik, insbesondere der Massenspektrometrie, zu signifikanten Fortschritten auf dem Gebiet der Proteomanalyse [38,39].

An die sinnvolle Analyse des Proteoms, welches das Proteinäquivalent eines Genoms umfasst und eine quantitative Darstellung des Proteinexpressionsmusters einer Zelle, eines Organismus oder einer komplexen Zellflüssigkeit unter definierten Bedingungen ermöglichen soll, sind verschiedene Anforderungen gestellt.

Im ersten Schritt jeder Analyse müssen die Proteine aus Zellen oder Zellverbänden in Lösung gebracht und von den unlöslichen Bestandteilen der Zelle abgetrennt werden. Da sich die physikalischen und chemischen Eigenschaften von Proteinen jedoch deutlich unterscheiden können, muss zunächst eine geeignete Methode gefunden werden, um die interessierenden Analyten aus ihrer natürlichen Umgebung quantitativ zu extrahieren. Saravanan *et al.* (2004) stellten eindrucksvoll den Einfluss verschiedener Methoden zur Proteinextraktion aus pflanzlichem Gewebe auf das resultierende Proteinexpressionsmuster dar [40].

Für die Zielsetzung der Bestimmung der quantitativen Verhältnisse von Proteinen zueinander sind die klassischen Verfahren der Proteinreinigung häufig ungeeignet. So bedingen mehrstufige Isolationsstrategien immer auch eine Abreicherung bestimmter Proteine. Auch die Fraktionierung von Proteinen stellt aufgrund der Tatsache, dass einzelne Proteine oft in mehrere Fraktionen anzutreffen sind, eine für die quantitative Proteomanalyse wenig geeignete Methode dar. Daneben erschweren artifizielle Proteinmodifikationen z.B. die Oxidation von Methionin [41], die  $\beta$ -Eliminierung von Cysteinseitenketten bei hohen pH-Werten, was in einer falsch-positiven Identifikation von Dehydroalanin resultieren kann [42] oder die unkontrollierte Proteolyse [43] die Identifizierung und Quantifizierung. Die angewandte Methode der Proteinextraktion sollte daher nur wenige Arbeitsschritte erfordern und zugleich eine möglichst quantitative Extraktion der interessierenden Proteine gewährleisten. Zudem müssen die verwendeten Probenpuffer mit den sich anschließenden Verfahren der Proteinauftrennung und -identifikation kompatibel sein.

Ist eine robuste und reproduzierbare Extraktionsmethode gefunden, erfordert eine erfolgreiche Proteomanalyse aufgrund der angesprochenen Komplexität der

Stoffwechselebene der Proteine zunächst eine Auftrennung der Analyten. Diese kann gelbasiert oder gelfrei erfolgen. Die höchste Auflösung lässt sich heute mit der zwei-dimensionalen Polyacrylamidgelelektrophorese (2D-PAGE) erzielen. In der ersten Dimension erfolgt dabei eine isoelektrische Fokussierung (IEF) der Proteine an immobilisierten pH-Gradienten, die Trennung in der zweiten Dimension erfolgt elektrophoretisch in orthogonaler Richtung nach der Grösse. Mit dieser Technik sind heute bis zu 10.000 Proteine auftrennbar [44]. Der hohen Auflösung stehen jedoch einige unbestreitbare Limitierungen entgegen. Neben der fehlenden Automatisierbarkeit, welche eine begrenzte Reproduzierbarkeit bedingt, ist der Transfer von der ersten IEF-Dimension auf die zweite Dimension häufig unvollständig, so dass Proteinverluste unvermeidbar sind [45]. Zur Sichtbarmachung der elektrophoretisch aufgetrennten Proteine stehen verschiedene Färbetechniken zur Verfügung, welche sich hinsichtlich der Sensitivität stark unterscheiden. Auch die Möglichkeiten der Identifizierung von Proteinen aus der Gelmatrix sind begrenzt, so dass die aufgetrennten Proteine zur weiteren Analyse aus der nicht-inerten Gelmatrix gebracht werden müssen. Die begrenzte Kompatibilität einzelner Färbemethoden mit weiteren, für die Proteinidentifizierung erforderlichen Verfahren stellt eine weitere Limitierung der gelbasierten Proteinanalytik dar. So sind beispielsweise Silber-gefärbte Gelspots wegen der teilweise irreversiblen Fixierung der Proteine zur quantitativen massenspektrometrischen Analyse nicht geeignet. Zudem werden gleiche Mengen verschiedener Proteine häufig mit unterschiedlicher Intensität gefärbt, so dass keine Proportionalität zwischen der Proteinkonzentration und der Färbeintensität gegeben ist [46-48]. Auch im Hinblick auf die Analyse posttranslationaler Modifikationen erwies sich die klassische SDS-PAGE als kritisch. So wiesen beispielsweise Sun *et al.* nach, dass während der elektrophoretischen Auftrennung häufig artifizielle Proteinoxidationen auftreten [49].

Neben der 2D-PAGE steht für die Auftrennung von intakten Proteinen mit der Hochleistungsflüssigchromatographie (HPLC) eine gelfreie Methode mit gutem Auflösungsvermögen zur Verfügung. Als Standardverfahren kommen dabei vorrangig die Gelfiltration, die Affinitätschromatographie, die Ionenaustauschchromatographie oder die reversed phase Chromatographie zum Einsatz. Der Nachweis der aufgetrennten Proteine kann mit Hilfe eines UV-Detektors oder massenspektrometrisch erfolgen, wobei die Massenspektrometrie für die Bestimmung des Molekulargewichts und der Sekundärstruktur von Proteinen heute die Methode der Wahl darstellt. Die Voraussetzung dafür schaffte die Entwicklung schonender Ionisierungsverfahren wie die Matrix-assisted Laser Desorption Ionisation (MALDI) [50] oder die Elektrospray Ionisation (ESI) [51] in den 80er Jahren. Die Möglichkeit der Ionisierung von Proteinen und Polypeptiden katalysierte eine enorme Weiterentwicklung bestehender Gerätelösungen, so dass heute komplexe Instrumente wie z.B. Hybrid Quadrupol Time-of-Flight (Q-TOF)- oder Tandem Time-of-

Flight (TOF-TOF)-Massenanalysatoren für die Proteinanalyse kommerziell verfügbar sind. In der Kombination von MALDI mit TOF- oder TOF-TOF-, sowie ESI mit Ionenfallen-, Q- oder Q-TOF-Massenanalysatoren können heute Proteine aus hochkomplexen Proteinlösungen zuverlässig identifiziert werden [39].

Einen kurzen Überblick über die in der Protein-, aber auch in der Metabolitanalytik häufig Anwendung findenden Massenanalysatoren und Gerätelösungen sowie neueste instrumentelle Entwicklungen und Applikationen geben die folgenden Absätze.

### ***Triple-Quadrupol Massenspektrometer***

Mit Hilfe von aus drei Quadrupolen Q1 bis Q3 aufgebauten Triple-Quadrupol Massenspektrometern ist es möglich, gezielte Sequenzinformationen von Peptiden zu erhalten. Vor der Detektion der Analytionen oder daraus resultierender Fragmentionen müssen diese nacheinander einen Hilfsquadrupol oder Octapolfilter Q0, welcher den Eintritt der Ionen in den analytischen Ionenkanal ermöglicht, einen ersten Messquadrupol Q1, einen als Kollisionszelle fungierenden Quadrupol Q2 sowie einen zweiten Messquadrupol Q3 passieren. Die beiden Messquadrupole Q1 und Q3 können optional als selektive Massenfilter oder als über einen definierten  $m/z$ -Bereich scannende Massenanalysatoren betrieben werden. Als Stossgas für die Fragmentierung im Q2 kann Stickstoff, Helium oder Argon dienen. Je nach Einsatz der Messquadrupole lassen sich mit Triple-Quadrupol Massenspektrometern verschiedene  $MS^2$  Experimente durchführen (Abbildung 1.2.1) [52]. Werden im Q1 definierte Vorläuferionen isoliert, deren in Q2 erzeugte Fragmentionen im Q3 gescannt werden, spricht man von einer Produkt-Ionen-Analyse. Bei der Vorläufer-Ionen-Analyse werden die Quadrupole in umgekehrter Richtung betrieben. Während Q1 den Transfer aller Vorläuferionen über einen definierten  $m/z$ -Bereich in die Kollisionszelle Q2 erlaubt, wo diese sequenziell fragmentiert werden, können nur ausgewählte Fragmentionen Q3 passieren und zum Detektor gelangen. Eine weitere Möglichkeit der Nutzung von Triple-Quadrupol Massenspektrometern besteht in der Durchführung von Neutralverlust-Analysen. Hierbei werden, ähnlich der Vorläufer-Ionen-Analyse, Ionen über einen definierten Massenbereich in die Kollisionszelle transferiert und dort fragmentiert. Der zweite Messquadrupol scannt um die Masse des Neutralverlusts zurückversetzt synchron zum ersten Messquadrupol. Dabei gelangen ausschliesslich Fragmentionen zum Detektor, welche die geforderte Massendifferenz zu ihren Vorläuferionen aufweisen. Die Neutralverlust-Analyse bietet sich für die gezielte Analyse posttranslatinaler Proteinmodifikationen an [53,54]. Beim Multiple Reaction Monitoring (MRM) hingegen fungieren beide Messquadrupole als Massenfilter, wobei definierte Vorläuferionen im Q1 isoliert werden und ausschliesslich deren spezifische Fragmentionen Q3 passieren können [55]. Der MRM-Modus ist hervorragend für die

hochselektive Detektion spezifischer Analyten mit bekannter Fragmentierung aus komplexen Proben geeignet und lässt sich ebenso wie die Neutralverlust-Analyse zur Analyse posttranslationaler Modifikationen nutzen [56-58].

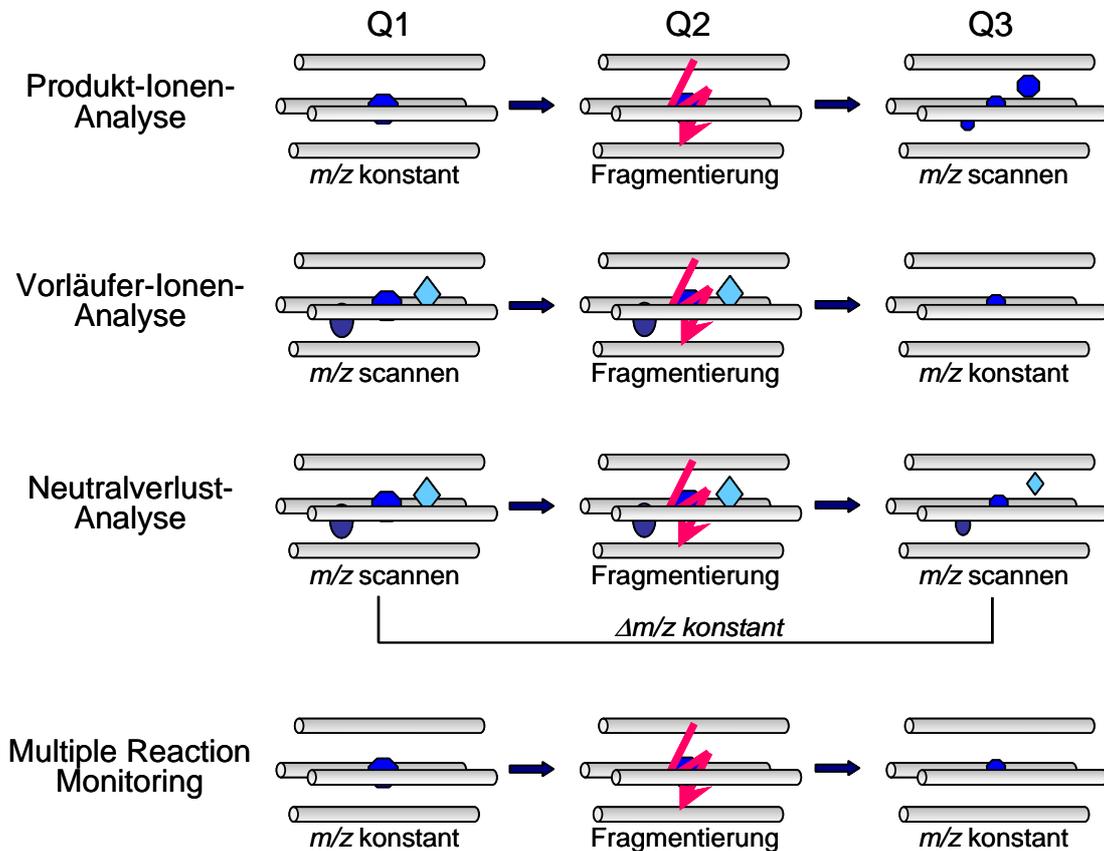


Abbildung 1.2.1: Prinzipien der mittels Triple-Quadrupol Massenspektrometer realisierbaren MS<sup>2</sup>-Experimente

### Ionenfallen Massenanalysatoren

Mittels 3D-Quadrupol Ionenfallen-Massenanalysatoren werden Ionen über einen definierten Zeitbereich spezifisch in einem elektrischen Feld eingefangen, welches durch rotationssymmetrische Elektroden erzeugt wird. Durch die kontinuierliche Erhöhung des zwischen den Elektroden angelegten Wechselspannungsanteils können die eingefangenen Ionen massenselektiv aus der Ionenfalle ejiziert werden und/oder Vorläuferionen spezifisch isoliert werden. Durch die Kollision der isolierten Ionen mit Helium können diese zur Fragmentierung gebracht und die entstehenden Fragmente selektiv aus der Falle ejiziert und detektiert werden. Alternativ können einzelne Fragmente weiter in der Falle isoliert und weiteren Dissoziationen unterzogen werden, womit MS<sup>n</sup>-Fragmentierungen möglich sind. Neuere Ionenfallen Massenanalysatoren sind mit

linearen Ionenfallen (LIT) ausgestattet, welche aufgrund ihrer Geometrie über ein grösseres Volumen verfügen. Im Gegensatz zu 3D-Quadrupol Ionenfallen sind lineare Ionenfallen mit zwei Detektoren ausgestattet, was die Sensitivität weiter deutlich verbessert [52,59].

Neuartige Gerätelösungen wie das Q-Q-LIT verbinden Quadrupol-Massenanalysatoren mit LIT, wodurch die Vorzüge der Triple-Quadrupol Instrumente (Abbildung 1.2.1) durch die implementierte LIT um die Möglichkeit zur Durchführung von  $MS^n$ -Experimenten erweitert werden und ein quasi-kontinuierlicher Betrieb möglich wird [60-63].

### ***TOF- und Hybrid-TOF- Instrumente***

Moderne TOF-Massenanalysatoren, welchen die Bestimmung des Masse-zu-Ladung-Verhältnisses durch Messung der Flugzeit von Ionen nach deren Beschleunigung im elektrischen Feld durch eine definierte Flugstrecke im Hochvakuum zugrunde liegt, zeichnen sich durch eine hohe Auflösung von bis zu 20.000 sowie eine sehr hohe Massengenauigkeit aus [64]. In Hybridgeräten werden Ionen durch einen Quadrupol (Q-TOF) oder zwei (Q-Q-TOF) zum TOF beschleunigt, wo die eigentliche Massenanalyse stattfindet. Im Q-Q-TOF werden typischerweise mehrfach geladene Vorläuferionen im ersten Quadrupol selektiert, im zweiten Quadrupol unterlaufen diese einer Kollisions-induzierten Fragmentierung und die resultierenden Produktionen werden im angeschlossenen TOF-Massenanalysator getrennt [65]. Die auch im  $MS^2$ -Modus zu erreichende hohe Massengenauigkeit und –auflösung resultiert in einer zuverlässigen Aufklärung der Primär- und Sekundärstruktur von Proteinen und Peptiden und erlaubt darüber hinaus die *de novo*-Sequenzierung.

Eine spezielle Anwendung von TOF-Massenanalysatoren stellt die Kopplung mit der sehr sensitiven und im Vergleich zur ESI gegenüber Kontaminationen wie Salzen oder geringen Konzentrationen an Detergenzien deutlich toleranteren MALDI dar. Neueste Gerätelösungen ermöglichen eine Kopplung von MALDI mit Q-Q-TOF oder TOF-TOF-Massenanalysatoren. Der Vorteil dieser Kopplungen besteht in der Fragmentierung der zumeist einfach geladenen Vorläuferionen, wodurch leicht interpretierbare  $MS^2$ -Spektren mit hoher Massengenauigkeit erhalten werden können [66-68].

### ***Ionen Cyclotron Resonanz und Orbitrap Massenanlysatoren***

Eine durch exzellente Massenauflösung sowie –genauigkeit gekennzeichnete Gerätelösung stellen die unter Verwendung der Fourier-Transformation betriebenen Ionen Cyclotron Resonanz (FT-ICR) Massenspektrometer dar. Im Ultrahochvakuum werden dabei Ionen in einem starken magnetischen Feld (typischerweise 9 oder 12 Tesla) eingefangen und zwischen einem Kondensatorplattenpaar auf Kreisbahnen gehalten. Die

Kreisfrequenz des von den geladenen Teilchen erzeugten oszillierenden Bildstroms ist massenabhängig und kann mit Hilfe der Fourier-Transformation ermittelt werden. Den mit FT-ICR Massenanalytoren zu erzielenden Massengenauigkeiten von unter 1ppm und der extrem hohen Massenauflösung von bis zu 1.000.000 stehen als Nachteil die relativ geringe Scanrate (ein Scan dauert mehrere Sekunden) und damit verbunden eine vergleichsweise geringe Sensitivität gegenüber. LIT-ICR Hybridgeräte ermöglichen die parallele Aufnahme qualitativ hochwertiger MS- und MS<sup>2</sup>-Spektren [24,69-72].

Bezüglich der Massenauflösung und –genauigkeit mit der FT-ICR-MS vergleichbare Charakteristika weisen die erstmals 2005 vorgestellten Orbitrap Massenanalytoren auf [73]. Ionen werden hier in einem elektrostatischen Feld eingefangen, wo sie um eine Zentralelektrode rotierend auf axial oszillierenden Bahnen gehalten werden. Ionen mit gleichem  $m/z$ -Verhältnis induzieren einen zwischen den äusseren Hälften der Orbitrap detektierbaren Sinusstrom mit spezifischer Frequenz. Befinden sich gleichzeitig Ionen verschiedener Masse und/oder Ladung in der Orbitrap, erzeugen diese ein aus der Frequenzüberlagerung resultierendes komplexes Signal. Auch hier können die Frequenzen und somit die Massen der einzelnen Ionen mit Hilfe der Fourier-Transformation bestimmt werden. Mit den als Hybridgeräte in Verbindung mit einer linearen Ionenfalle verfügbaren Orbitrap Massenanalytoren steht eine verhältnismässig kostengünstige magnet- und cryogenfreie FT-MS Gerätelösung zur Verfügung, welche mit einer zu erreichenden Massengenauigkeit von 2ppm an die Spezifikationen der FT-ICR Massenanalytoren heranreicht [74,75].

Während sich die klassische Proteomforschung lange auf den Nachweis und die Katalogisierung von Proteinen beschränkte, erlauben diese technischen Entwicklungen zunehmend auch die quantitative Proteinanalyse. Den Massenspektrometrie-basierten Methoden zur Proteinquantifizierung stehen verschiedene gelbasierte Methoden gegenüber. Beide Verfahren haben ihre individuellen Vor- und Nachteile und stellen somit komplementäre Plattformen der quantitativen Proteomics dar. Die folgenden Absätze geben einen kurzen Einblick in verschiedene Methoden der quantitativen Proteinanalytik.

### 1.2.1 Gelbasierte quantitative Proteomanalyse

Für die Quantifizierung von mittels 2D-PAGE aufgetrennten Proteinen stehen verschiedene Methoden zur Verfügung. Ein interessantes Verfahren stellt die im Jahre 2002 entwickelte DIGE (2D Difference Gel Electrophoresis)-Technologie dar [76]. Bei dieser Methode werden zwei unterschiedliche Proteinextrakte vor der elektrophoretischen Auftrennung durch die kovalente Bindung verschiedener fluoreszierender Cyaninfarbstoffe markiert. Dabei kann die Markierung minimal an den Lysinresten der Proteine erfolgen,

wobei etwa 5% eines Proteins markiert werden, oder aber die Markierung erfolgt quantitativ an den Cysteinresten [77]. Da die Fluoreszenzfarbstoffe unterschiedliche Wellenlängen emittieren, ist die relative Quantifizierung von Proteinen unterschiedlicher Proben aus einem Gel möglich. Alternative Techniken zur relativen Quantifizierung von Proteinen aus Polyacrylamidgelen beruhen auf der Markierung von Cysteinresten mit stabilen Isotopen wie  $^1\text{H}_3/{}^2\text{H}_3$ -Acrylamid [78,79] oder  $^1\text{H}_4/{}^2\text{H}_4$ -2-Vinylpyridin [79]. Zur Identifizierung und Quantifizierung müssen die aufgetrennten Proteine jedoch aus der Gelmatrix gebracht und einer massenspektrometrischen Analyse zugeführt werden.

### 1.2.2 Gelfreie quantitative Proteomanalyse

Neben der Möglichkeit der Auftrennung komplexer Proteinextrakte an Polyacrylamidgelen existieren verschiedene, direkt an die MS gekoppelte chromatographische Methoden. Im Gegensatz zu gelbasierten Techniken erfolgt die Aufspreizung der komplexen Proteinextrakte hier vornehmlich auf der Ebene von Peptiden, welche aus dem enzymatischen Proteinverdau resultieren. Für die relative oder auch absolute Quantifizierung von Proteinen, welche unterschiedliche physiologische Zustände oder genetische Modifikationen anzeigen, kommen auch hierbei häufig isotoopenmarkierte Verbindungen zum Einsatz. Eine Übersicht über nachfolgend diskutierten Strategien der Proteinquantifizierung gibt die Abbildung 1.2.2.1.

Das wohl bekannteste Verfahren zur parallelen relativen Quantifizierung von Proteinen nach Isotoopenmarkierung stellt die von Gygi *et al.* (1999) vorgestellte ICAT<sup>TM</sup> (Isotope-Coded Affinity Tag)-Technologie dar [80]. Diese beruht auf der chemischen Derivatisierung von Cysteinresten mit dem sog. ICAT<sup>®</sup>-Reagenz (Fa. Applied Biosystems, Darmstadt, Deutschland). Markiert werden hierbei ganze Proteine, deren korrespondierende biotinylierte Peptide nach einem enzymatischen Verdau affinitätschromatographisch über Avidin aufgereinigt werden. Die Komplexität einer Peptidprobe kann auf diese Weise auf etwa ein Zehntel reduziert werden. Dieses ist jedoch auch mit einem erheblichen Verlust an Sequenzinformationen verbunden, da ausschliesslich cysteinhaltige Peptide biotinyliert sind. Die relative Quantifizierung der Peptide erfolgt über die Massendifferenz von 9 Da, welche sich aus der Differenz zwischen dem  $^{13}\text{C}_9$ -markierten und dem unmarkierten Linker zwischen der Cysteinreaktiven Gruppe und dem Biotinrest des ICAT<sup>®</sup>-Reagenz ergibt. Um eine quantitative Derivatisierung hochkomplexer Proteinproben zu gewährleisten, muss das ICAT<sup>®</sup>-Reagenz in grossem Überschuss zugesetzt werden, was mit erheblichen Kosten verbunden ist. Zudem muss überschüssiges ICAT<sup>®</sup>-Reagenz vor der Anreicherung der biotinylierten Peptide in einem zusätzlichen Schritt chromatographisch entfernt werden.

Als weiterer Nachteil dieser Methode erwies sich die Tatsache, dass auch nicht derivatisierte Peptide häufig unspezifisch an Avidin binden [81].

Eine ähnliche Methode zur relativen Quantifizierung von Proteinen bietet das 2004 unter dem Begriff ICPL<sup>TM</sup> (Isotope-Coded Protein Labeling) von Schmidt *et al.* vorgestellte Verfahren zur Markierung freier Aminogruppen intakter Proteine mit dem Aminspezifischen Reagenz *N*-Nicotinyl-succinimid (Fa. Serva, Heidelberg, Deutschland) [82].

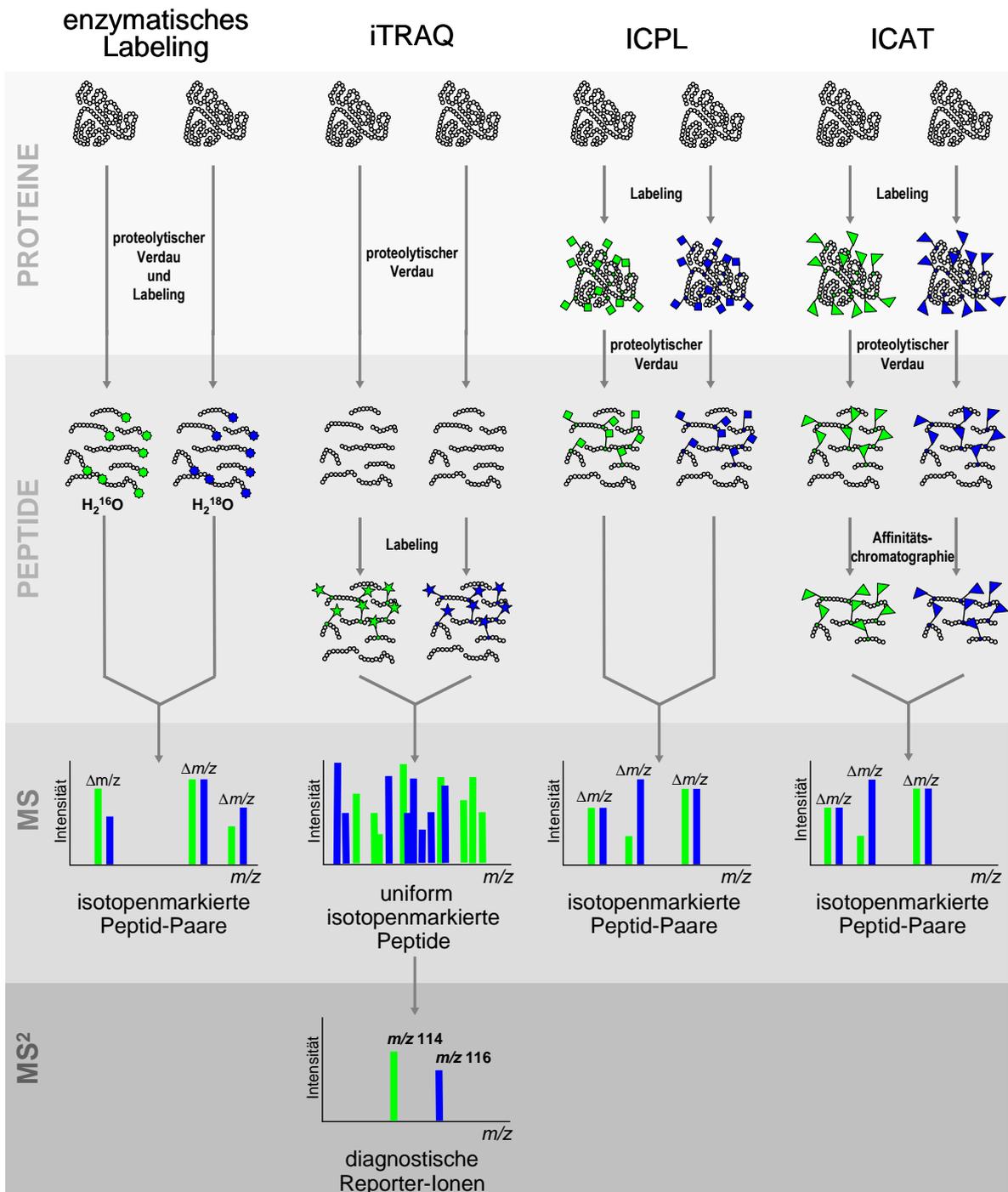


Abbildung 1.2.2.1: Strategien der Massenspektrometrie-basierte Proteinquantifizierung.

Neben ICAT<sup>TM</sup> und ICPL<sup>TM</sup> existieren weitere Methoden zur Markierung von Proteinen, wie das metabolische *in vivo* Labeling, wobei Organismen auf <sup>15</sup>N- oder <sup>13</sup>C-angereicherten Medien angezogen werden, das enzymatische <sup>18</sup>O-Labeling, wobei Proteinextrakte in H<sub>2</sub><sup>16</sup>O und H<sub>2</sub><sup>18</sup>O enzymatisch verdaut werden [81,83,84], sowie diverse Methoden der chemischen Derivatisierung von Peptiden mit isotopenmarkierten Reagenzien, wie z.B. die Alkylierung reduzierter Cysteinreste mit markiertem und unmarkiertem Dithiothreitol oder die Methylierung der freien Carboxylgruppen saurer Aminosäuren mit C<sup>1</sup>H<sub>3</sub>OH sowie C<sup>2</sup>H<sub>3</sub>OH. Eine sehr junge Methode zur indirekten Proteinquantifizierung stellt die von Ross *et al.* (2004) vorgestellte Methode zur Markierung von proteolytischen Peptiden mit maximal vier aminspezifischen isobaren Reagenzien dar [85]. Diese als iTRAQ<sup>TM</sup> bekannte Technologie beruht auf der Derivatisierung freier Aminogruppen mit iTRAQ<sup>TM</sup> Reagenzien (Fa. Applied Biosystems, Darmstadt, Deutschland), welche identische MS-Spektren liefern, jedoch nach der MS<sup>2</sup>-Fragmentierung individuelle, sog. diagnostischen Reporter-Ionen aufweisen. Mit dieser Technologie können heute vier unterschiedliche Proteinextrakte individuell markiert und parallel *via* Massenspektrometrie sowohl relativ als auch absolut quantifiziert werden.

Den Vorteilen der allen vorgestellten Methoden gemeinsamen, auch als Multiplexing bezeichneten Parallelisierung der vergleichenden Proteinanalyse nach der Markierung mit stabilen Isotopenverbindungen, steht jedoch der Nachteil der limitierten Anzahl der zu vereinigenden Proben gegenüber. Für die vergleichende Analyse von mehr als vier Zuständen steht bis dato keine Technologie kommerziell zur Verfügung.

### 1.3 Die Analyse des Metaboloms – Metabolomics

In begrifflicher Analogie zum Proteom wird die quantitative Gesamtheit der in einer Zelle oder einem Organismus vorkommenden kleinen Moleküle, welche an Reaktionen zur Aufrechterhaltung lebenswichtiger und für das Wachstum essentieller Stoffwechselfvorgänge beteiligt sind, als Metabolom bezeichnet [28,86,87]. Die genaue Anzahl der das Metabolom verschiedener Spezies umfassenden Intermediate ist ungewiss. Schätzungen gehen von etwa 600 Metaboliten in der Hefe *S. cerevisiae* [88] und bis zu 200.000 Metaboliten im Pflanzenreich [89] aus. Eine besondere Herausforderung bei der Analyse des Metaboloms stellt die Tatsache dar, dass Metabolite enzymatisch und chemisch ineinander umwandelbar sind. Weder auf der Ebene der Transkripte, noch auf der Ebene der Proteine finden Reaktionen statt, welche unmittelbar mit einem Massentransfer verbunden sind.

Die methodischen Herangehensweisen an die Analyse des Proteoms und die des Metaboloms sind im Wesentlichen identisch. Auch bei der Metabolomanalyse stellt die angewandte Extraktionsmethode zur quantitativen Herauslösung der kleinen Moleküle aus Zellen, Zellorganellen oder Gewebeverbänden einen das Ergebnis unter Umständen drastisch beeinflussenden, essentiellen Schritt dar [90-92], dem sich die Detektion und die Identifikation der einzelnen Analyten anschliesst. In Abhängigkeit von der verfolgten Strategie zur Analyse des Metaboloms ergeben sich im Umgang mit dem Untersuchungsmaterial Unterschiede. Während Strategien wie das Metabolic Footprinting (siehe Tabelle 1.1), welche die Analyse des Exometaboloms zum Gegenstand haben, relativ einfach zu realisieren sind, erfordern Analysen des Endometaboloms bezüglich der Probenpräparation einen deutlich höheren Aufwand und grössere Sorgfalt. So müssen hierbei Enzymaktivitäten unmittelbar nach dem Ernten des Untersuchungsmaterials gequencht werden, bevor intrazelluläre Metabolite geeignet extrahiert werden können. Pflanzliches Material wird dazu zumeist mittels flüssigen Stickstoffs schockgefroren.

In der Literatur sind zahlreiche Protokolle für die Extraktion intrazellulärer Metabolite aus diversen Organismen beschrieben, welche sich hinsichtlich der Eignung für die Extraktion bestimmter Stoffklassen zum Teil erheblich unterscheiden. So unterliegen beispielsweise die in Lösung ohnehin nicht sonderlich stabilen Triosephosphate in alkalischen Lösungen einer raschen Degradation. Die Extraktion von Metaboliten aus gefrorenem Gewebe erfolgt typischerweise in wässrigen Lösungen wie z.B. die Extraktion säurestabiler Metabolite mit Perchlorsäure oder Trichloressigsäure, die Extraktion mit hochmolaren Kaliumhydroxid-Lösungen, oder auch mit kochenden oder tiefkalten Wasser-Alkohol-Gemischen [90,92,93]. Für die Wahl der geeigneten Extraktionsmethode ist neben der Stabilität der Zielanalyten die Kompatibilität mit den für die Identifikation und Quantifizierung zur Anwendung kommenden Techniken entscheidend. Ein für die simultane Extraktion von Proteinen und Metaboliten aus gefrorenem Pflanzenmaterial geeignetes und zugleich MS-kompatibles Protokoll stellten Weckwerth *et al.* (2004) vor [94].

Die ausserordentlich grosse Diversität der Metabolite, welche zudem innerhalb der Zellen, Gewebeverbände oder einzelner Organe in ihren Konzentrationen um bis zu neun Grössenordnungen (pM bis mM) variieren können [95], macht eine Reduktion der Komplexität von Metabolitextrakten aus biologischen Proben durch geeignete, zumeist chromatographische Trennmethode und/oder die Analyse von Verdünnungsserien unabdingbar. Wenngleich die Methoden zur Auftrennung der in den Extrakten gelösten Verbindungen variieren können, stellt die Massenspektrometrie eine Schlüsseltechnologie für die Identifizierung und Quantifizierung der kleinen Moleküle dar [38,95]. Die wichtigsten Massenspektrometrie-basierten Methoden zur Analyse des Metaboloms werden in den nächsten Absätzen vorgestellt.

### ***Gaschromatographie-basierte qualitative und quantitative Analyse des Metaboloms***

Für die Identifizierung und Quantifizierung von Metaboliten stellt die Kopplung der Gaschromatographie mit der Massenspektrometrie traditionell die Methode der Wahl dar. Bereits im Jahre 1954 nutzten James *et* Martin die Gaschromatographie als Technik zur Analyse und Identifikation flüchtiger Verbindungen [96]. Nur vier Jahre später gelang Holmes *et* Morrell die erste GC/MS-Kopplung [97]. Ungeachtet der neuesten Entwicklungen analytischer Plattformen und Gerätelösungen (siehe Kapitel 1.2) gehören GC/MS gekoppelte Systeme nach wie vor zu den am häufigsten eingesetzten Methoden in der Metabolitanalytik. Dieses liegt vor allem in der hohen Trennleistung und Sensitivität sowie in den kurzen Analysenzeiten begründet. Die notwendige Voraussetzung für eine gaschromatographische Analyse stellt die Verdampfbarkeit und eine ausreichende thermische Stabilität der Zielanalyten dar. Nicht-flüchtige Verbindungen müssen, um diesen Anforderungen zu genügen, zunächst chemisch geeignet derivatisiert werden. Generell umfasst die chemische Derivatisierung vor der Substitution aktiver Protonen durch weniger polare Trialkyl-, meist Trimethylsilylgruppen, die Oximbildung mit O-Alkylhydroxylaminen. In Abhängigkeit von der Struktur der Ausgangsverbindungen entstehen dabei jedoch häufig mehrere Derivate, was die Komplexität über das Mass der in der Ausgangsprobe tatsächlich vorkommenden Metabolite erhöht. Da die Silylierung insbesondere in Anwesenheit von Wasser eine reversible Reaktion darstellt, müssen die zu analysierenden Extrakte vor der Derivatisierung bis zur völligen Trockne eingeeengt werden. Dieses bedingt jedoch den Verlust leicht flüchtiger Komponenten. Für die Analyse ausgewählter Zielverbindungen existieren daneben weitere, spezifische Methoden für die Derivatisierung funktioneller Gruppen wie etwa die Alkylveresterung von Carboxylgruppen.

Neben der Flüchtigkeit und thermischen Stabilität erfordert die GC/MS niedrige bis mittlere Molekulargewichte der Analyten. Verbindungen mit einem Molekulargewicht von über 500 Da sind im Allgemeinen einer gaschromatographischen Analyse nicht zugänglich. Diesen Limitierungen der GC/MS stehen jedoch die aus den mit den heute existierenden Gerätelösungen realisierbaren extrem hohen Scanraten resultierenden kurzen Analysenzeiten gegenüber.

### ***GC/MS Systeme für die schnelle Analyse komplexer biologischer Proben***

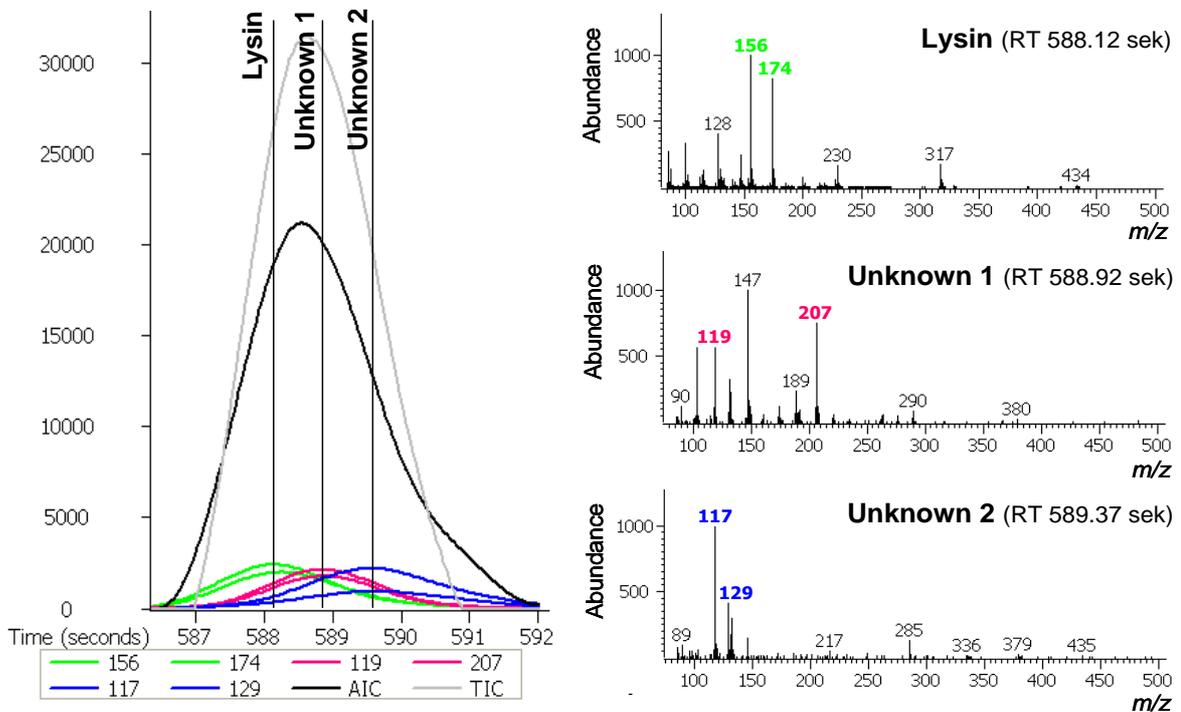
Eine insbesondere für schnelle gaschromatographische Analysen komplexer Gemische entwickelte Gerätelösung stellt das von der Firma LECO® (LECO Corporation, St. Joseph, MI, USA) entwickelte Pegasus® II GC/TOF System dar, welches die analytischen Techniken GC und MS unter Verwendung eines Reflektor-TOF-Massenanalysators verbindet. Die Besonderheit dieses Instruments besteht in einem sog. High Speed Data

Akquisition System. Konventionell werden bei gaschromatographischen Analysen Signale mit Peakbreiten von einigen Sekunden erhalten, welche mit Quadrupol- oder Sektorfeld-Massenanalytoren problemlos analysiert werden können. Um eine adäquate chromatographische Auflösung zu erlangen, sind hierbei jedoch relativ lange Analysezeiten erforderlich. Eine vergleichbare chromatographische Auflösung bei kurzen Analysezeiten erfordert schmalere Peaks. Die Identifizierung und Quantifizierung stellt sich jedoch aufgrund der Tatsache, dass die Scanraten konventioneller Massenspektrometer hierbei nur eine geringe Anzahl von Datenpunkten (Spektren) entlang eines Peaks offerieren, als schwierig dar. Mit dem High Speed, High Dynamic Range Data Akquisition System des Pegasus®II werden von der Ionenquelle Spektren, die sog. Transients, mit einer Frequenz von 5.000 sek<sup>-1</sup> über einen Massenbereich von  $m/z$  5 bis 1000 erzeugt. Ein eigentliches Massenspektrum ergibt sich aus der Summe von 10 Transients, so dass eine maximale Spektrenakquisitionsrate von 500 Massenspektren- $\text{sek}^{-1}$  erreicht wird. Wird das System mit einer geringeren Spektrenakquisitionsrate betrieben, erhöht sich die kumulative Anzahl der Transients zu Gunsten eines verbesserten Signal-zu-Rausch-Verhältnisses. Auch der dynamische Bereich lässt sich auf diese Weise für Ionen, welche mit einer Rate von durchschnittlich <1 pro Transient erzeugt werden, deutlich verbessern. Daneben sind mit der auf der Annahme, dass das Fragmentspektrum einer reinen Verbindung innerhalb eines chromatographischen Peaks unveränderlich ist, beruhenden, implementierten Peak-Deconvolutions-Softwarelösung auch co-eluierende Verbindungen detektierbar und quantifizierbar, sofern diese unterschiedliche Fragmentspektren aufweisen (Abbildung 1.3.1).

So können heute mit einem Pegasus®II TOF Massenspektrometer aus einer einzigen biologischen Probe innerhalb einer 20-minütigen Chromatographie über 1000 Peaks detektiert und über die Massenspuren der für die jeweiligen Ausgangsverbindungen spezifischen Fragmentionen quantifiziert werden [98].

Für die Analyse kleiner Molekülen von Vorteil und zugleich limitierend stellt sich die kommerzielle Verfügbarkeit zahlreicher chemischer Substanzen mit hohem Reinheitsgrad dar, welche zur Erstellung von Referenz-Spektrenbibliotheken herangezogen werden können. So können heute etwa 200 der das Metabolom von *S. cerevisiae* umfassenden Metabolite käuflich erworben werden [95]. Da jedoch mit zunehmender Komplexität des Metaboloms der prozentuale Anteil der verfügbaren Referenzsubstanzen abnimmt und deren Synthese sehr anspruchsvoll und zeitaufwendig ist, ist eine absolute Quantifizierung in umfangreichen Metabolomanalysen nicht realisierbar und auf die Metabolit Target Analyse ausgewählter Verbindungen beschränkt.

Auch in Verbindung mit der Gaschromatographie stehen heute moderne Gerätelösungen wie die Kopplungen mit Triple-Quadrupol, Ionenfallen oder TOF-Massenanalysatoren zur Verfügung, mit denen eine Vielzahl an MS<sup>2</sup>- oder MS<sup>n</sup>-Experimenten wie beispielsweise



**Abbildung 1.3.1: Peak-Deconvolution der LECO®Pegasus®ChromaTOF™-Software am Beispiel einer GC/TOF/MS-Analyse polarer Metabolite aus *A. thaliana* Blattmaterial.** Aufgrund der unterschiedlichen Fragmentspektren sind die drei co-eluierenden Verbindungen Lysin, Unknown 1 und Unknown 2 identifizierbar und anhand der Massenspuren spezifischer Fragment-Ionen quantifizierbar. AIC: Analytical Ion Chromatogram, TIC: Total Ion Chromatogram.

MRM-Experimente realisierbar sind [99-101]. Die gerätetechnische Grundlage für die Durchführbarkeit von Tandem-MS-Experimenten in Verbindung mit der Gaschromatographie stellen unterschiedliche, implementierte Ionisierungstechniken dar. Die Ionisierung der Zielanalyten erfolgt im Hochvakuum, wobei sich der schonenden chemischen Ionisierung (CI) mit Reaktandgasen eine Elektronenstoss-Ionisierung (EI) anschliesst. Durch starken Elektronenbeschuss (150 eV) werden die Reaktandgasmoleküle hierbei zunächst zur Bildung positiv geladener Primärionen angeregt. Die eigentliche Ionisation der Analyten erfolgt durch stabile Plasmaionen, welche durch eine Serie von Stößen der Primärionen mit weiteren Reaktandgasmolekülen entstehen. Mit abnehmender Protonenaffinität des Reaktandgases erfolgt eine zunehmend schonendere positive Ionisierung der Zielanalyten durch Ladungsaustausch mit dem Reaktandgas, so dass das Ausmass der

Fragmentierung durch die Wahl des Reaktandgases gesteuert werden kann und Fragmentspektren erhalten werden, welche positiv geladene Quasi-Molekülonen  $[M]^{+}$  aufweisen. Durch Zusammenstöße von Primärelektronen mit Gasmolekülen entstehen im Reaktandgasplasma auch niederenergetische Sekundärelektronen, welche insbesondere von Verbindungen mit hoher Elektronegativität wie beispielsweise chlorhaltigen organischen Substanzen mit hoher Effizienz eingefangen werden können, was in der Entstehung negativ geladener Quasi-Molekülonen  $[M]^{-}$  resultiert. Diese können dann gegenüber chemisch positiv ionisierten Analyten mit deutlich höherer Empfindlichkeit aus komplexen Proben nachgewiesen werden. Nach vorangegangener Derivatisierung wie beispielsweise die Umsetzung von Alkoholen und Aminen mit Trifluoracetylchlorid oder Pentafluorbenzylbromid lässt sich dieser Vorteil der negativen CI auch für weniger elektronegative Analyten nutzbar machen [102-105]. Mit der Möglichkeit der Bestimmung des Molekulargewichtes mittels CI und der Aufnahme von Fragmentspektren der chemisch ionisierten Analyten durch Elektronenstoss eignen sich Tandem-GC-MS Systeme hervorragend für die Strukturaufklärung unbekannter Verbindungen.

Ein wesentliches Kriterium für die Wahl des für die Analyse bestimmter Zielanalyten geeigneten Massenspektrometers stellt das Auflösungsvermögen dar. Dieses ist durch die Breite eines in  $m/z$  homogenen Ionensignals auf der Massenskala definiert. Für die hochauflösende massenspektrometrische Analyse, welche beispielsweise vom Gesetzgeber für die Rückstandsanalytik von Pestiziden gefordert ist, stellen Kopplungen von Gaschromatographen mit FT-ICP-Massenspektrometern oder doppelt-fokussierenden Sektorfeld-Massenspektrometern probate gerätetechnische Lösungen dar, mit denen Auflösungen von bis zu 100.000 erzielt werden können. Für Sektorfeld-Massenspektrometer gelten zwei Peaks als getrennt, sofern das Tal zwischen diesen 10% des Peaks mit der geringeren Signalintensität ausmacht, wobei die Peakbreite mit der Masse wächst, so dass die eingestellte Auflösung im gesamten Massenbereich gilt. Im Gegensatz dazu ist die Peakbreite von Quadrupol- oder Ionenfallen-Massenspektrometern auf der Massenskala konstant und meist auf 50% Tal eingestellt. Das Kernprinzip hochauflösender Sektorfeld-Massenspektrometer besteht in der doppelten Fokussierung des Ionenstrahls durch ein Magnetfeld und ein elektrisches Feld, welche nacheinander angeordnet sind.

Eine interessante und zu den konventionellen GC/MS-Systemen, welche die Extraktion der Zielanalyten und deren Derivatisierung erfordern, komplementäre Technik stellt die Entwicklung neuer Einlasssysteme dar, mit deren Hilfe die Kopplung von Headspace-Extraktionssystemen mit der GC möglich ist. Diese Systeme ermöglichen die Analyse

leicht flüchtiger Komponenten aus komplexen Matrices ohne eine vorherige aufwendige Probenvorbereitung [104,106-108].

Aufgrund der bereits angesprochenen eingeschränkten Anwendbarkeit der GC/MS für Verbindungen mit Molekulargewichten bis zu 500 Da finden auch komplementäre Flüssigchromatographie-basierte Methoden eine breite Anwendung in der Metabolomanalyse.

### ***Flüssigchromatographie-basierte qualitative und quantitative Analyse des Metaboloms***

Neben der Möglichkeit der Auftrennung aus biologischem Material extrahierter Metabolite in der Gasphase können diese auch in gelöster, flüssiger Form, ähnlich den Peptiden bei der Proteomanalyse, chromatographisch aufgetrennt und anschliessend massenspektrometrisch analysiert werden [109-111]. Die Voraussetzung für eine LC/MS-basierte Analyse von Metaboliten stellt ihre positive oder negative Ionisierbarkeit dar. Während die Ionisation von Peptiden aufgrund ihrer funktionellen Gruppen mit der ESI zumeist problemlos realisierbar ist, stellte sich die Ionisierung stark hydrophober Metabolite lange als problematisch dar. Durch die Entwicklungen alternativer Ionisierungstechniken unter Atmosphärendruck wie etwa der chemischen Ionisation (Atmospheric Pressure Chemical Ionization, APCI) oder der Photo-Ionisation (Atmospheric Pressure Photo-Ionization, APPI) sind jedoch auch diese Verbindungen heute einer LC/MS-Analyse problemlos zugänglich [112-116]. Beide Ionisationstechniken sind schonend. Die Wahl der geeigneten Ionisationsquelle hängt in erster Linie von den chemischen Eigenschaften der Zielanalyten ab. Im Gegensatz zur ESI, bei der die Ionisierung in der flüssigen Phase stattfindet, werden bei der APCI mittels Stickstoffstrom zerstäubte solvatisierte Verbindungen durch eine auf ca. 400°C temperierte Keramik geführt, wobei das Lösungsmittel und die Zielanalyten vollständig verdampfen. Die vaporisierten Zielanalyten werden unter Hochspannung von ca. 5 kV über eine Corona-Entladungsnadel in ein Plasma überführt, wobei die Molekülonen des Eluenten in Reaktandgasionen überführt werden, welche ihre Ladungen schliesslich auf die Analytmoleküle übertragen. Die Ionenbildung hängt in starkem Masse von den Eigenschaften der Analyten und den verwendeten Lösungsmitteln ab [117,118]. Während basische Analyten ausschliesslich positiv geladene Ionen und Kationenaddukte bilden, spalten saure Analyten Protonen ab und können Anionen anlagern. Ähnlich der negativen CI begünstigen elektronegative Substituenten die Bildung negativer Ionen [119,120]. Im Gegensatz zur ESI, wo die Anzahl der auf die Analyten übertragenen Ladungen vom Molekulargewicht und den funktionellen Gruppen der Analyten abhängt, treten bei der APCI ausschliesslich einfach geladene Ionen auf [121-123]. Anwendung findet die APCI in der Analyse gering bis

mässig polarer Verbindungen bis zu einem Molekulargewicht von ca. 2000 Da wie z.B. Fettsäuren und Lipide [116,124,125], Propanoide [126], Amphetamine und Alkaloide [127-129], Carotinoide [130-132] Aldehyde und Ketone [133,134], oder auch in der Aminosäure- und Peptidanalytik [135].

Unpolare Substanzen sind jedoch auch mit der APCI häufig nicht oder nur unzureichend ionisierbar. Für diese Verbindungen steht mit der APPI eine alternative Ionisierungstechnik zur Verfügung [112,113,136,137]. Hierbei ersetzt eine Photonenemittierende Kryptonlampe die Corona-Entladungsnadel. Die Ionisierung der Zielanalyten kann durch die direkte Interaktion der Moleküle mit den emittierten Photonen erfolgen, wobei unter Elektronenabgabe positiv geladene Molekülionen entstehen, sofern das Ionisierungspotenzial der Zielanalyten geringer ist als die Photonenenergie  $h \cdot \nu$  [138]. Erfolgt die Verdampfung der Zielanalyten in protischen Lösungsmitteln werden die Zielanalyten durch Protonenübertragung der durch Photoionisation entstandenen Radikalkationen des Lösungsmittels ionisiert [139]. Ist die Protonenaffinität der Analytmoleküle höher als die der deprotonierten Radikalkationen, kann die Ionisierung auch durch Ladungsübertragung auf die Zielanalyten erfolgen [140]. Durch Ladungsaustausch, Elektroneneinfang oder auch Substitutionsreaktionen können mit der APPI saure Analyten mit einer gegenüber der APCI deutlich höheren Effizienz auch unter Bildung negativ geladenen Ionen fragmentiert werden [141,142]. Im Vergleich zur APCI erfordert die Desolvatation bei der APPI geringere Temperaturen, so dass auch thermisch instabilere Analyten ionisiert werden können [143]. Als weitere Vorteile der APPI sind die gegenüber ESI oder auch APCI deutlich geringeren Matrixeffekte, eine verbesserte Sensitivität und die Erniedrigung des Detektionslimits zu nennen [112,114,144]. Für die APPI sind zahlreiche Applikationen, darunter u.a. die Analyse polyaromatischer Kohlenwasserstoffe

[145,146], Steroide [147,148], Mycotoxine [149,150], Aldehyde und Ketone [151], Antibiotika [152], Fettsäuren und Acyllipiden [115] oder auch die Analyse von Pestiziden [144] sowie hydrophober Peptide [153] beschrieben.

In Abhängigkeit von der Geräteausstattung finden auch bei der Identifikation und Quantifikation von Metaboliten verschiedenste, darunter die in Kapitel 1.2 angeführten MS-, MS<sup>2</sup> oder MS<sup>n</sup>-Experimente Anwendung. Mit einigen Massenspektrometern ist zudem ein optionaler kontinuierlicher Wechsel von positivem zu negativem Ionenmodus innerhalb von Millisekunden realisierbar, wodurch die Anzahl der nachweisbaren Analyten deutlich erhöht werden kann [154].

Eine neben der HPLC für proteolytische Peptide und Metabolite gleichermassen geeignete, neuartige Gerätelösung für die Auftrennung komplexer Proben stellt die Ultrahochdruckflüssigchromatographie (Ultrahigh Pressure Liquid Chromatography, UPLC<sup>TM</sup>, Fa. Waters Corporation, Milford, MA, USA). Hierbei kommen Trennsäulen zum

Einsatz, welche mit Sorbentien extrem geringer Partikelgrösse ( $<2\mu\text{m}$ ) gefüllt sind und bei Betriebsdrücken von bis zu 15.000 psi durch herausragende Trennleistungen gekennzeichnet sind [155].

LC/MS-basierte Analysen finden vorrangig bei Analyse der grossen Anzahl pflanzlicher Sekundärmetabolite wie Polyphenole oder Steroide Anwendung [156-159].

Eine viel versprechende Alternative zur Analyse von Inhaltsstoffen biologischen Gewebes ohne eine aufwendige Probenvorbereitung stellt die 2005 vorgestellte Desorptions-Elektrospray-Ionisation (DESI) in Verbindung mit der MS dar [160]. Bei diesem Verfahren wird ein feines Spray geladener Lösungsmitteltröpfchen unter Atmosphärendruck mittels eines Hilfsgases direkt auf eine biologische Gewebeprobe gerichtet, wobei Ionen von der Gewebeoberfläche desorbiert und der massenspektrometrischen Analyse direkt zugeführt werden. Wenngleich die räumliche Auflösung dieser Anwendung heute bei 0.5 bis 1 mm liegt, verspricht dieses Ionisationsverfahren in naher Zukunft eine breite Anwendung auf dem Gebiet des Tissue Imaging zur systematischen Charakterisierung von Gewebeoberflächen oder histologischen Gewebeschnitten. Erste Anwendungen der DESI-MS auf dem Gebiet der Protein- und Metabolitanalytik sind beschrieben [160-162].

#### **1.4 Experimentelles Design und Chemometrik**

Abhängig von der Art und Herkunft der zu analysierenden Proben ergeben sich mitunter drastische Unterschiede. Während sich im Rahmen von Untersuchungen des Stoffwechsels von Mikroorganismen aufgrund der Möglichkeit der synchronen Anzucht unter streng kontrollierten Wachstumsbedingungen Anhand der Analyseergebnisse einer minimierten Anzahl von Proben bereits signifikante Aussagen treffen lassen, erfordert die Analyse des Stoffwechsels höherer Organismen, welche eine deutlich grössere biologische Varianz aufweisen, die Untersuchung weitaus umfangreicherer Stichproben zur Erhebung valider Daten. Mit den heutigen gerätetechnischen Entwicklungen kann dieses im Rahmen von immer häufiger Anwendung findenden Hochdurchsatzanalysemethoden problemlos realisiert werden. Die Analyse grosser Probensets birgt jedoch auch Hindernisse. Neben der Tatsache, dass viele Proben häufig in mehreren Batches analysiert werden müssen, was zu artifiziellen 'Batch-zu-Batch'-Unterschieden führen kann, erfordern allein die innerhalb kurzer Zeit anfallenden grossen Datenmengen eine zumeist anspruchsvolle chemometrische Analytik [95,163-165]. Zudem stellt die Anwendung der sog. 'Omics'-Technologien häufig den ersten Schritt des Prozesses der Erkenntnisgewinnung dar, so dass noch keine sachbezogene Hypothese vorliegt. Zur Lösung von Problemen dieser Art werden typischerweise zwei oder mehrere biologischer Systeme durch möglichst viele Variablen charakterisiert und die erhaltenen

Resultate miteinander verglichen. Unterschiede sind dabei jedoch nicht immer offensichtlich. Häufig liegen diese in den hochdimensionalen Datensätzen verborgen und treten erst nach einer zuweilen recht anspruchsvollen chemometrischen Analyse zu Tage. Chemometrische Analysen dienen zwei wesentlichen Zielen: experimentelle Daten sollen unter Erhalt relevanter Informationen visualisiert und/oder klassifiziert werden. Für die Visualisierung, welche der Erkennung von Strukturen oder spezifischen Mustern anhand der Verteilung der analysierten Stichproben verschiedener Grundgesamtheiten dient, müssen die zumeist hochdimensionalen Datensätze zunächst reduziert werden.

#### 1.4.1 Dimensionsreduzierende Verfahren

Bei der Dimensionsreduktion sucht man nach einer zwei- oder dreidimensionalen Darstellung, welche die wesentliche Verteilung oder auch Struktur der  $n$ -dimensionalen Daten möglichst gut wieder gibt. Ein häufig Anwendung findendes Verfahren zur Dimensionsreduktion stellt die Hauptkomponentenanalyse (Principal Components Analysis, PCA) dar [166].

##### **Hauptkomponentenanalyse (PCA)**

Ziel der PCA ist es, die Originalvariablen eines Datensatzes unter Erhalt der relevanten Informationen durch eine geringere Anzahl von Variablen zu ersetzen. Enthält ein Datensatz stark korrelierte Variablen, so haben diese im Wesentlichen den selben Informationsgehalt und es geht nur wenig Information verloren, wenn eine von zwei stark korrelierten Variablen aus einem Originaldatensatz entfernt wird.

Durch eine lineare Transformation des Originaldatensatzes wird bei der PCA versucht, die Anzahl korrelierter Variablen auf eine geringe Anzahl unkorrelierter Variablen zu reduzieren, wobei korrelierte Variablen in neuen, unkorrelierten Variablen, den sog. Hauptkomponenten zusammengefasst werden, welche redundanzfrei den grössten Teil der Gesamtvarianz erklären. Man versucht also, aus einem Originaldatensatz mit vielen Variablen einige wenige latente Faktoren zu extrahieren, welche für die Eigenschaften der Variablen bestimmend sind. Der Originaldatensatz wird dazu unter Minimierung der Korrelationen zwischen den Variablen mathematisch in einen Vektorraum mit neuer Basis überführt, wobei die berechneten Hauptkomponenten Linearkombinationen der ursprünglichen Variablen darstellen. Dieses erfolgt analog zum dreidimensionalen Raum, in dem sich jeder beliebige Punkt  $b$  durch die Linearkombination der drei orthogonalen,

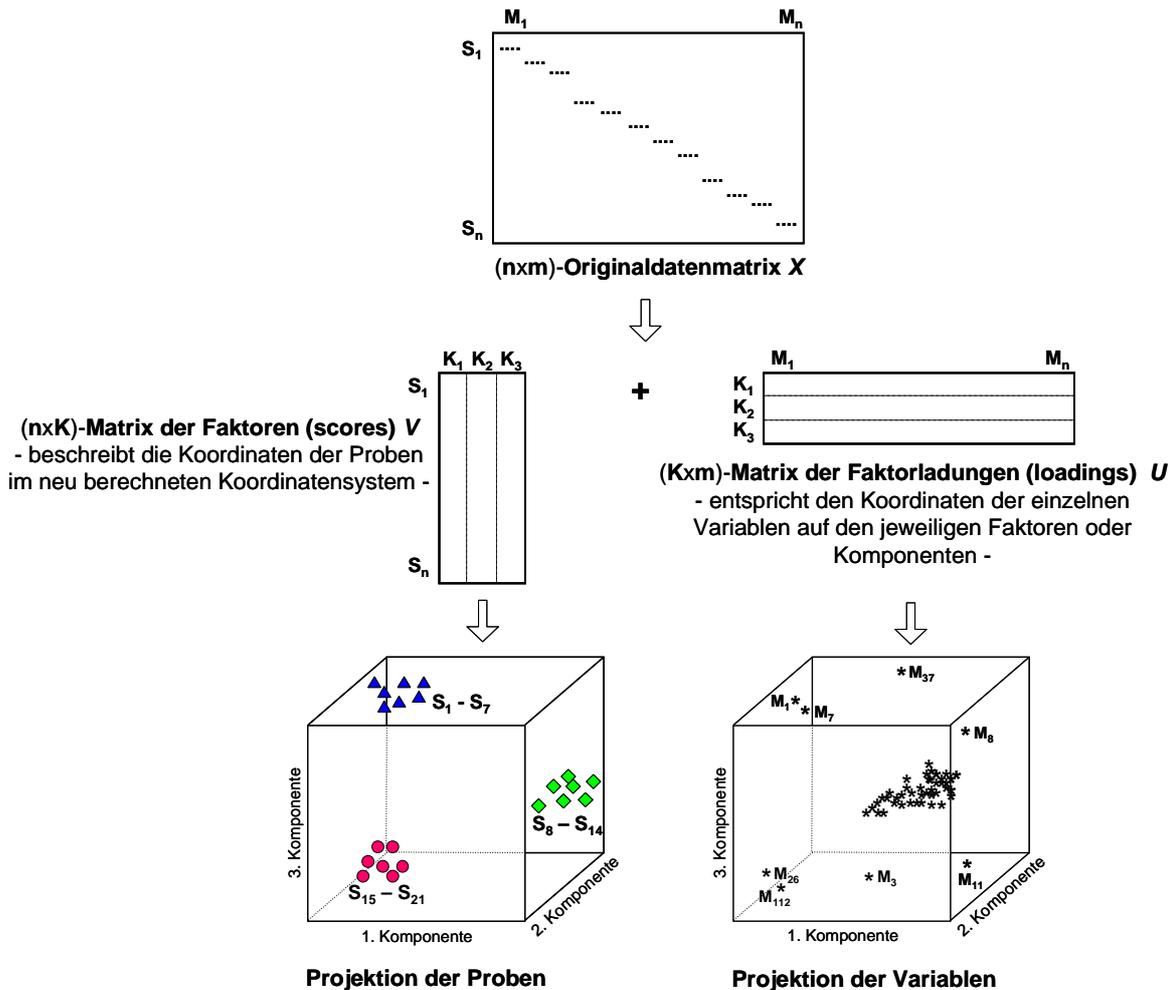
und damit gänzlich unkorrelierten Basisvektoren  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  und  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  mit einem Vektor  $\begin{pmatrix} bx \\ by \\ bz \end{pmatrix}$ ,

welcher die Koordinaten des Punktes  $b$  auf den Basisvektoren enthält, beschreiben lässt. Die Basisvektoren der PCA werden jedoch nicht willkürlich gewählt, sondern ausgehend von einer n-dimensionalen Punktwolke so festgelegt, dass die erste Achse des neu berechneten Koordinatensystems der Richtung der grössten Varianz entspricht. Alle weiteren Achsen werden orthogonal in Richtung abnehmender Varianz angeordnet. Für n-dimensionale Datensätze ergeben sich dadurch grundsätzlich n orthogonal zueinander angeordnete Achsen. Geometrisch betrachtet entspricht dieses der Suche nach einem neuen Koordinatensystem, welches durch Rotation des ursprünglichen Koordinatensystems gewonnen wird. Die neuen Koordinatenachsen entsprechen dabei den Richtungen der Hauptkomponenten. Aus der Summe der entlang jeder Achse repräsentierten Varianz ergibt sich die Gesamtvarianz. Ordnet man die Achsen nach absteigender Varianz, lassen sich die durch eben diese Achsen repräsentierten Faktoren oder Hauptkomponenten bestimmen, welche zusammen einen Grossteil der Gesamtvarianz (im Allgemeinen 95%) erklären. Ist auf diese Weise eine neue Basis gefunden, können die transformierten Originaldaten anhand weniger Hauptkomponenten (Dimensionen) adäquat abgebildet werden. Als Ergebnis der PCA erhält man zwei Matrizen, die Matrix der Faktoren (scores), welche die Koordinaten der Proben im neu berechneten Raumgitter der Hauptkomponenten beschreibt, und die Matrix der Faktorladungen (loadings), welche die Richtung der jeweiligen Faktoren im originalen Vektorraum als Vektor-Koordinaten angeben (Abbildung 1.4.1.1).

Mathematisch passiert bei der PCA folgendes:

Als Ausgangspunkt für die Berechnung dient die n-dimensionale Datenmatrix  $\mathbf{X}$  der Gestalt  $\mathbf{n}$  (Variablen, Merkmale)  $\times$   $\mathbf{m}$  (Beobachtungen, Proben). Durch spaltenweise Subtraktion des Durchschnittsvektors, dessen Elemente den Mittelwerten der Zeilen von  $\mathbf{X}$  entsprechen, erhält man zunächst einen zentrierten Datensatz  $\mathbf{A}$ . Anschliessend wird die Kovarianzmatrix  $\mathbf{K}$  von  $\mathbf{A}$  gemäss  $\mathbf{K} = \mathbf{AA}^T$  berechnet. Nach Eigenwertzerlegung der erhaltenen Kovarianzmatrix  $\mathbf{K}$ , welche die Dimension  $\mathbf{m} \times \mathbf{m}$  aufweist, erhält man zwei Matrizen, eine Diagonalmatrix  $\mathbf{Y}$  mit den  $\mathbf{m}$  Eigenwerten, anhand derer sich ableiten lässt, wie viele Vektoren oder Dimensionen zur Erklärung der geforderten Varianz nötig sind, sowie eine auch als Matrix der Faktoren bezeichnete Eigenvektormatrix  $\mathbf{V}$ , welche die Koordinaten der Proben auf den neu berechneten Basisvektoren bzw. Hauptkomponenten enthält. Beide Matrizen haben die Dimension  $\mathbf{m} \times \mathbf{m}$ . Ist die Matrix der Faktoren bekannt, ist die Bestimmung der Koordinaten der Variablen auf den Faktoren durch Multiplikation

von  $V$  mit der zentrierten Datenmatrix  $A$  realisierbar, wobei ebenso viele Eigenvektoren oder Hauptkomponenten berechnet werden, wie Eigenwerte nötig sind, um die geforderte Varianz zu erklären. Die spaltenweise Anordnung der normierten Eigenvektoren ergibt die gesuchte Koordinatenmatrix der Faktorladungen  $U$ , welche den Einfluss der einzelnen Variablen/Metabolite auf die durch die jeweiligen Basisvektoren/Hauptkomponenten repräsentierte Varianz anzeigt [166].



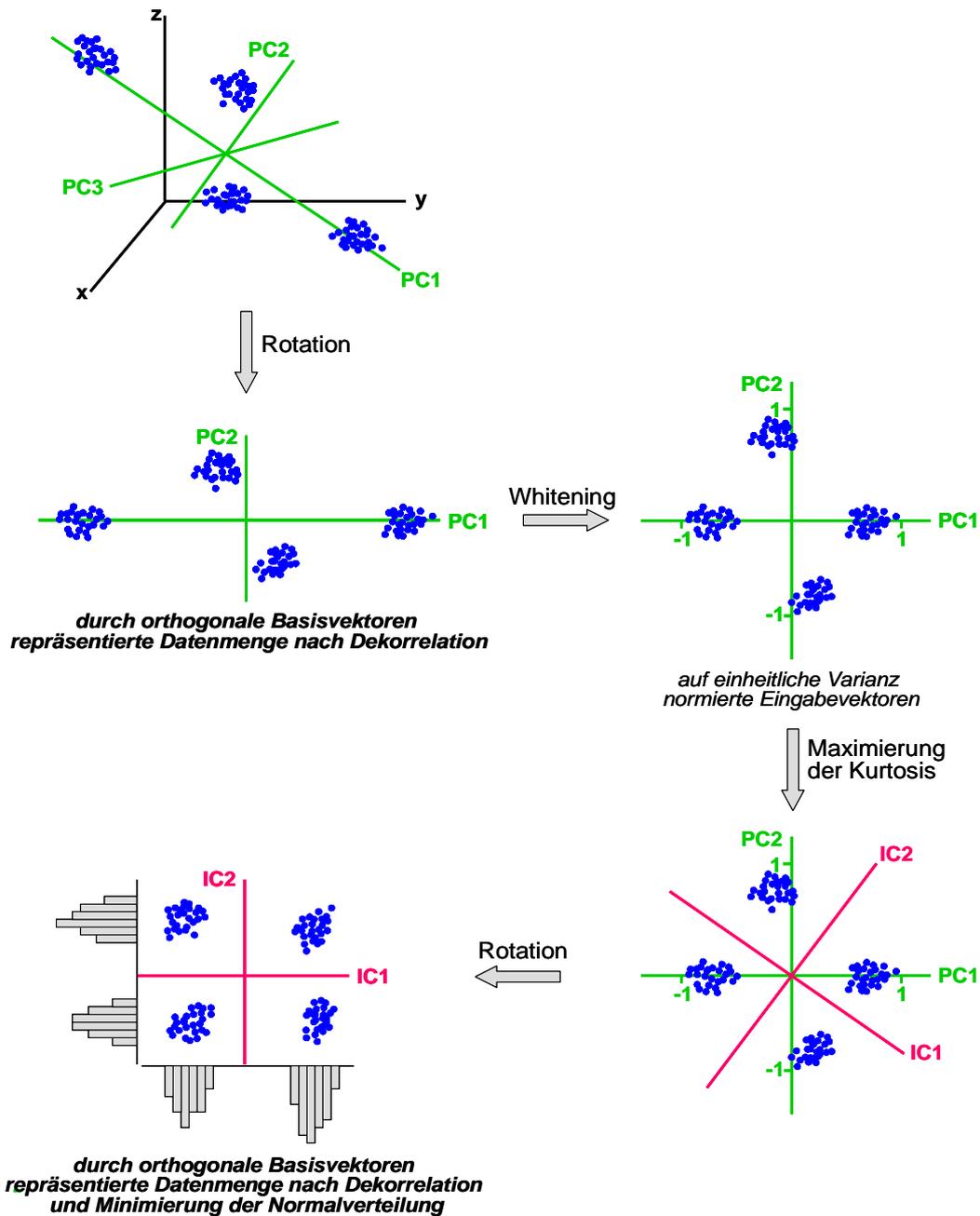
**Abbildung 1.4.1.1: Schema der Dimensionsreduktion durch die Hauptkomponentenanalyse PCA.** Ausgehend von einer  $n$ -dimensionalen Originaldatenmatrix  $X$  erfolgt die Dimensionsreduktion bei der Hauptkomponentenanalyse durch die Überführung der Originalvariablen in einen neu berechneten Vektorraum. Basierend auf der Annahme, dass die Richtung der grössten Varianz die meisten Informationen beinhaltet, werden die Basisvektoren in Richtung absteigender Varianz orthogonal angeordnet. Nach einer Eigenwertzerlegung der die Lage aller Datenpunkte im neuen Vektorraum beinhaltenden Transformationsmatrix lässt sich diese in eine Matrix der Faktoren (scores)  $V$  überführen, welche die Lage der Datenpunkte im neu berechneten Vektorraum wieder gibt, und der Anteil der einzelnen Basisvektoren an der Gesamtvarianz berechnen. Da häufig ein Grossteil der Gesamtvarianz durch nur wenige Faktoren/Dimensionen erklärt wird (hier  $K_1$  bis  $K_3$ ), können die Datenpunkte visualisiert werden. Den Beitrag der einzelnen Variablen auf die von den jeweiligen Basisvektoren repräsentierte Varianz ergibt sich aus der erhaltenen Matrix der Faktorenladungen (loadings)  $U$ , welche die Koordinaten der Variablen auf den Faktoren beinhaltet, so dass auch diese abgebildet werden können. *Abkürzungen:* M, Metabolit; S, Sample/Probe; K, (Haupt)Komponente.

### **Unabhängige Komponentenanalyse (ICA)**

Neben der PCA steht mit der Unabhängigen Komponentenanalyse (Independent Components Analysis, ICA) eine weiterführende dimensionsreduzierende Technik zur Verfügung, mit dessen Hilfe hochdimensionale Datensätze linear so transformiert werden, dass diese nicht nur wie im Falle der PCA dekorreliert, sondern auch statistisch unabhängige Gruppen voneinander separiert werden [167,168]. Die statistische Unabhängigkeit kann anhand Verteilung beurteilt werden, wobei gilt: je nicht-gaussischer die Verteilung, desto unabhängiger die Gruppen. Ein wesentlicher Unterschied zwischen der PCA und der ICA besteht darin, dass auch nicht-orthogonale Strukturen in n-dimensionalen Originaldatensätzen erkannt, und nach der Transformation in einem neu berechneten orthogonalen Vektorraum abgebildet werden können. Ermöglicht wird dieses durch das auch als Whitening bezeichnete stochastische Strecken oder Verzerren der zentrierten Originaldaten. Dazu wird der zu transformierende Datensatz zunächst einer PCA unterzogen und durch Rotation des ursprünglichen Koordinatensystems in einen neuen Vektorraum überführt. Sind die neu berechneten Faktoren in Richtung abnehmender Varianz festgelegt, erfolgt eine lineare Transformation der bereits zentrierten Eingabevektoren in orthogonale Vektoren, so dass deren Komponenten eine auf den Wert 1 normierte Varianz aufweisen (Abbildung 1.4.1.2). Basierend auf der so erhaltenen Verteilung werden die Datenpunkte durch Rotation nochmals in einen wieder neu berechneten Vektorraum überführt, wobei dessen Basisvektoren, welche den Unabhängigen Komponenten entsprechen, so angeordnet werden, dass die Normalverteilung in Richtung der Vektoren minimiert wird [168]. Ein klassisches Mass für die Beurteilung der Ähnlichkeit einer Verteilung zur Normalverteilung stellt die Kurtosis dar.

Diese berechnet sich nach  $kurtosis(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^4}{(n-1)\delta^4} - 3$ , wobei  $X = (X_1 \dots X_n)$  die

Variablen bzw. Komponenten,  $\bar{X}$  deren Mittelwert sowie  $\delta$  der Standardabweichung entspricht. Die Kurtosis für Normalverteilungen beträgt 3. Super-gaussische Verteilungen nehmen positive, sub-gaussische Verteilungen negative Werte an. Auch bimodale Verteilungen oder Clusterstrukturen haben eine negative Kurtosis. Daher werden die neu berechneten unabhängigen Komponenten bei der ICA ausgehend von der am stärksten negativen Kurtosis, welche der ersten unabhängigen Komponente entspricht, aufsteigend sortiert. Sind die unabhängigen Komponenten gefunden, kann nach gleichem Vorgehen wie bei der PCA die Matrix der Faktorladungen berechnet werden, so dass sowohl die Verteilung der Proben als auch der Beitrag der einzelnen Variablen zwei- oder dreidimensional abgebildet werden können (Abbildung 1.4.1.2).



**Abbildung 1.4.1.2: Prinzip der Unabhängigen Komponentenanalyse ICA.** Durch vorangegangene Hauptkomponentenanalyse werden Datenpunkte zunächst durch Rotation in einen neuen Vektorraum überführt. Durch die anschließende Normierung der Eingabevektoren auf einheitliche Varianz (Whitening) werden die Datenpunkte verzerrt. Basierend auf der so erhaltenen Verteilung erfolgt eine zweite Transformation durch Überführung in einen orthogonalen Vektorraum, dessen Basisvektoren so angeordnet werden, dass die nicht-Gaussheit maximiert wird, wodurch statistisch unabhängige Richtungen bestimmt werden können.

### 1.4.2 Klasseneinteilung

Neben der Mustererkennung ist es häufig sinnvoll, eine bestehende Datenmenge anhand charakteristischer Merkmale in voneinander verschiedene, homogene Gruppen oder Klassen einzuteilen. Eine häufig verwendete Methode zur Klasseneinteilung stellt die Clusteranalyse dar. Dabei werden ausgehend von einer gegebenen Menge von Objekten homogene Gruppen ermittelt, wobei der Abstand der Objekte innerhalb einer Klasse oder eines Clusters zueinander geringer ist, als der Abstand zu Objekten anderer Cluster. Cluster können auch als Gruppen von Objekten definiert werden, die untereinander oder in Bezug auf einen berechneten Schwerpunkt einen minimalen Abstand aufweisen.

Grundsätzlich lassen sich zwei wesentliche Konstruktionsverfahren von Clustern unterscheiden. Zum einen gibt es die sog. k-means Verfahren, welche eine Einteilung der Objekte in eine zuvor definierte Anzahl von Clustern liefern, wobei entsprechend eines globalen Masses eine optimale Homogenität unter den einzelnen Clustern erreicht wird, zum anderen existieren Verfahren zur Bildung hierarchischer Systeme von Gruppen von Objekten, welche von der feinsten Einteilung bis zur größten Gruppierung reichen. Diese Einteilung kann mit dem hierarchischen Clustering erzielt werden. In jedem Fall ist jedoch vor der Analyse festzulegen, bezüglich welcher Eigenschaften die Objekte miteinander verglichen werden sollen bzw. anhand welchen Masses die Ähnlichkeit bzw. Unähnlichkeit zwischen den Objekten numerisch ausgedrückt werden soll.

#### ***Partitionierende oder k-means Clusteringverfahren***

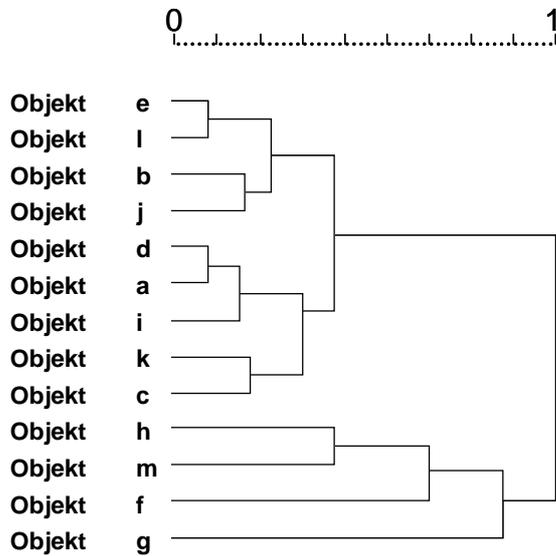
Im Gegensatz zu agglomerativ hierarchischen Clusteringverfahren, bei dem alle Elemente gemäss des zuvor definierten Abstandsmasses zu einem einzigen Cluster zusammengefügt werden, wird beim k-means Clustering eine gewünschte Anzahl von  $n$  Clustern vorgegeben. Zunächst erfolgt die Auswahl der  $n$  Clusterzentren zufällig und jedes der zu klassifizierenden Objekte wird den ihm am nächsten liegenden Clusterzentrum zugeordnet. Im Anschluss daran wird für jeden Cluster ein neues Clusterzentrum berechnet und die Zuordnung der Objekte überprüft. Ändert sich die Zuordnung, wird so lange mit der Berechnung eines neuen Clusterzentrums fortgefahren, bis ein quasi-stationärer Zustand bezüglich der Zuordnung der Objekte erreicht wird. Allerdings muss dieser Algorithmus nicht notwendigerweise konvergieren. Wird ein Cluster in einem Schritt der Zuordnung der Objekte nach Berechnung eines neuen Clusterzentrums nicht mehr gefüllt, kann für diesen nun leeren Cluster im nächsten Schritt kein neues Zentrum berechnet werden. Als einziger Ausweg einer solchen Situation bleibt

häufig nur der Neustart der Clusteranalyse in der Hoffnung, dass die erste zufällige Anordnung der Clusterzentren in einer 'sinnvollen' Anordnung der zu klassifizierenden Objekte resultiert. Als weiterer Nachteil der partitionierenden Clusteringverfahren gilt die Tatsache, dass die Anzahl der Cluster vor der Analyse definiert werden muss, was häufig aufgrund fehlender Erwartungswerte in einer artifiziellen Einteilungen der Objekte resultiert. Für ausführliche Beschreibungen der verschiedenen Verfahren zur Klasseneinteilung sei auf die zahlreich erscheinenden Publikationen und Lehrbücher zum Thema, wie etwa [169-171], verwiesen.

### ***Agglomerativ-hierarchische Clusteringverfahren***

Bei den agglomerativen oder sich anhäufenden Verfahren des hierarchischen Clusterings werden einzelne Objekte schrittweise zu einem einzigen Cluster zusammengefasst. Dabei wird zunächst jedes Objekt als ein eigener Cluster mit einem Element aufgefasst. Anschliessend werden diejenigen Objekte zusammengeführt, welche die grösste Ähnlichkeit zueinander aufweisen. Nach diesem Prinzip verfährt man so lange, bis alle Objekte zu einem einzigen Cluster angeordnet sind. Als Ergebnis erhält man ein hierarchisches System, welches als Dendrogramm darstellbar ist (Abbildung 1.4.2.1). Vor der Analyse ist festzulegen, aufgrund welcher Merkmale die Objekte miteinander verglichen werden sollen. Häufig dient dabei der Pearsonsche Produktmoment-Korrelations-Koeffizient  $r_{xy}$ , welcher sich sehr einfach und schnell berechnen lässt (siehe Kapitel 2), als Mass zur Anordnung der Objekte. Die sehr übersichtliche Anordnung der einzelnen Objekte in einem Dendrogramm, dessen Inklusionsrelationen üblicherweise auf einer Anordnung einer Gruppe der  $i$ -ten Subgruppe oberhalb der Gruppe einer  $(i-1)$ -ten Subgruppe basiert, ermöglicht das Finden einer Lösung des Clusterproblems. Einem Dendrogramm sind neben den mengentheoretischen Inklusionsbeziehungen immer auch die Abstandsrelationen zwischen den Objekten, und somit auch die Reihenfolge der Vereinigungen zu entnehmen. Zudem gibt ein Dendrogramm Aufschluss über die Kompaktheit von Clustern. Je früher zwei Objekte zusammengefügt werden, desto kürzer ist die Distanz der Knoten, welche als vereinigte Klammern in X-Richtung dargestellt werden, woraus folgt: je kürzer die Klammer, desto kompakter der Cluster. Für die üblichen Verfahren des hierarchischen Clusterings gilt eine Monotonieeigenschaft. Dieses hat zur Folge, dass die Abstände der einzelnen Objekte im Verlauf der Aggregation steigen. Wenngleich bei der Agglomeration immer die Objekte zusammen geführt werden, welche bezüglich des zuvor festgelegten Ähnlichkeitsmasses den geringsten Abstand zueinander aufweisen, ist dieser Abstand mindestens so gross wie die Abstände aller

Paare, welche zuvor vereinigt wurden. Mit zunehmendem Abstand der Vereinigungspaare wächst folglich die Heterogenität der zusammengeführten Objekte.



**Abbildung 1.4.2.1: Beispiel einer agglomerativ-hierarchischen Clusteranalyse.** Zur Clusterbildung wurden die 13 Objekte a bis m verwendet. Das hierarchisch-agglomerative Verfahren der Vereinigung dieser 13 Objekte resultierte in einer Anordnung, welcher die Ähnlichkeit oder Unähnlichkeit anhand der in Richtung der X-Achse angeordneten Klammerlängen, hier anhand der Skala von 0 bis 1 am oberen Rand der Grafik, zu entnehmen ist, wobei die Heterogenität zunimmt, je weiter die Klammer nach Rechts ( $\rightarrow$ 1) reicht.

## 2 Metabolische Netzwerke

### 2.1 Die Organisation metabolischer Netzwerke

Systeme, welche über einen Mechanismus zu ihrer Organisation verfügen, dessen zugrunde liegende Struktur sich mathematisch modellieren lässt, werden allgemein als Netzwerke bezeichnet. Auch im Metabolismus von Organismen jeglicher Art existieren dynamische Abhängigkeiten und Verbindungen von Genen, Proteinen sowie Metaboliten, welche sich mathematisch z.B. durch Kinetiken beschreiben lassen. Die Entwicklungen verschiedenster analytischer Plattformen in den letzten Jahren machen es heute möglich, unterschiedliche Stoffwechselebenen in grossem Umfang zu analysieren und die so gewonnenen Informationen zu biomolekularen Netzwerken zusammenzufassen. So können z.B. die aus ChIP-Chip Experimenten gewonnenen Daten zur Konstruktion von Genregulationsnetzwerken herangezogen werden [172,173], Protein-Protein-Interaktionsnetzwerke können basierend auf der Yeast two-hybrid Technologie [174,175] oder Signalling Netzwerke nach der Analyse posttranslationaler Proteinmodifikationen konstruiert werden [176].

Technische Entwicklungen auf dem Gebiet der Massenspektrometrie haben in den letzten Jahren dazu geführt, dass immer kleinere Probenmengen mit immer höherer Genauigkeit analysiert werden können. Insbesondere die Kopplung der Gaschromatographie an schnell scannende, hochauflösende TOF-Massenanalysatoren besticht durch ihre Sensitivität und Akkuratess und die im Vergleich zur Flüssigchromatographie kurzen Analysezeiten. Daher ist das Metabolite Profiling mittels GC-TOF-MS hervorragend zur Konstruktion metabolischer Netzwerke, welche die Metabolitdynamik bzw. zelluläre Fluktuationen durch biochemische Pathways hindurch beschreiben, geeignet.

## 2.2 Metabolite Profiling pflanzlichen Untersuchungsmaterials – Methodenevaluierung

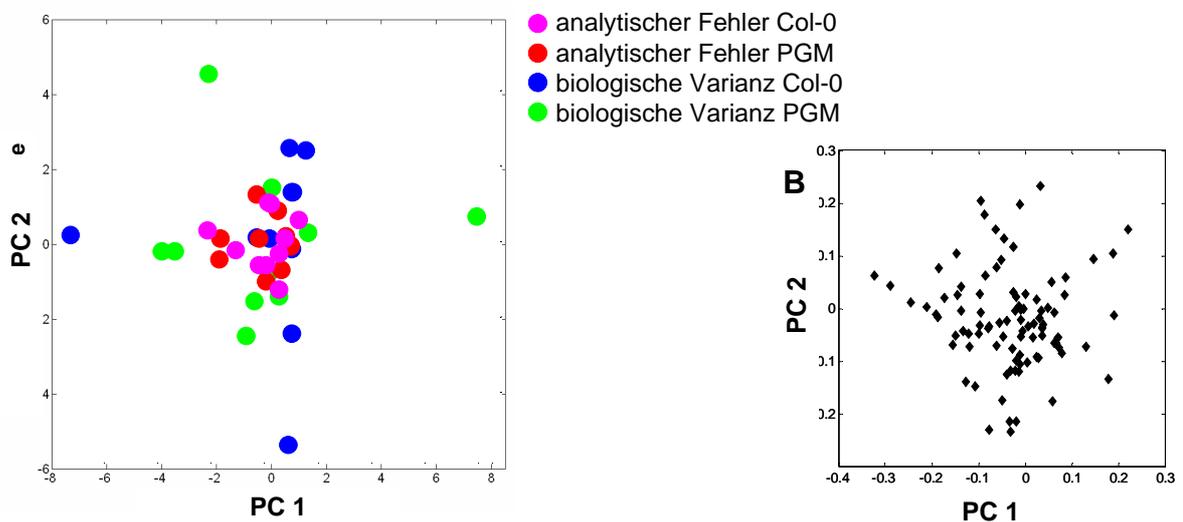
Profilanalysen biologischer Systeme setzen die akkurate Bestimmung möglichst vieler Stoffwechselprodukte in einem angemessenen Zeitraum voraus. Um aus den gewonnenen Daten biologisch relevante Informationen extrahieren zu können, sind ferner Replikatanalysen für die Durchführung statistischer Testverfahren unerlässlich.

Die hier vorgestellten Analysen zur Identifizierung und Quantifizierung polarer Inhaltsstoffe pflanzlichen Gewebes erfolgten gaschromatographisch an einem HP 6890, gekoppelt an ein Pegasus® II Time-of-Flight Massenspektrometer. Dieses System wurde mit einer Akquisitionsrate von 20 Spektren/sek bei nominaler Massenauflösung in einem Massenbereich von 85 – 600 m/z betrieben. Um die eindeutige Identifikation verschiedener Inhaltsstoffe aus den komplexen Matrices der polaren Extrakte sicherzustellen, wurde zunächst eine Spektrenbibliothek erstellt. Dazu wurde das Retentionsverhalten von ca. 400 Standardverbindungen basierend auf dem Retentionsindex-System von Kovats standardisiert. Als Retentionszeit-Standards dienten sowohl homologe n-Alkane als auch n-Fettsäuremethylester.

Um die Präzision der hier angewandten Metabolitanalyse zu überprüfen, wurden technische und biologische Replikate von *Arabidopsis thaliana* Pflanzen des Genotyps Columbia-0 (Col-0) sowie der stärkelosen, chloroplastidiären Phosphoglucomutase knock-out Mutante (PGM) gaschromatographisch analysiert. Dazu wurde nach Erreichen des Entwicklungsstadiums 5.1 [177] (siehe auch Anhang A-1) jeweils 20 Pflanzen eines jeden Genotyps geerntet. Alle Pflanzen wurden simultan unter Kurztagbedingungen angezogen. Die Ernte erfolgte 4h nach Beginn der 8-stündigen Photoperiode. Aus einem Pool von jeweils 10 Pflanzen wurden 10 Aliquots extrahiert und analysiert. Daneben wurden jeweils 10 Pflanzen Col-0 und PGM individuell analysiert. Alle Proben wurden dem Metabolite Profiling entsprechend 8.4 unterzogen. Dabei konnten 80 Verbindungen in allen Chromatogrammen anhand ihrer Fragmentspektren sowie ihres Retentionsverhaltens eindeutig und zumeist dem Primärstoffwechsel zugeordnet werden. Für die akkurate relative Quantifizierung wurde die Eignung der drei isotoopenmarkierten Verbindungen D-Sorbitol-<sup>13</sup>C<sub>6</sub>, D,L-Leucin-2,3,3-D<sub>3</sub> und L-Asparaginsäure-2,3,3-D<sub>3</sub> als Interne Standards getestet. Als Mass für die Eignung diente der aus der Analyse der jeweils 10 technischen Replikate resultierende Variationskoeffizient eines jeden Metaboliten nach der Normalisierung der Flächenwerte auf die Signalintensitäten der isotoopenmarkierten internen Standards. Dabei zeigte sich, dass für die Quantifizierung von Kohlenhydraten D-Sorbitol-<sup>13</sup>C<sub>6</sub>, für die Quantifizierung von Aminosäuren D,L-Leucin-2,3,3-D<sub>3</sub> und für die Quantifizierung von Intermediaten des Citrat Cyclus L-Asparaginsäure-2,3,3-D<sub>3</sub> als Interne

Standards am besten geeignet waren. Die resultierenden Variationskoeffizienten einzelner Metabolite lagen zumeist unter 25% (siehe Tabelle 2.2.1). Diese Werte umfassen die Summe aller möglichen Fehler, welche während der Extraktion, der Derivatisierung, der GC-TOF-MS Analyse sowie der Datenauswertung auftraten und repräsentieren somit den analytischen Fehler. Daneben wurde die biologische Variabilität durch die individuelle Analyse von biologischen Replikaten beider Genotypen ermittelt. Dabei zeigte sich, wie in Tabelle 2.2.1 aufgeführt, dass diese den analytischen Fehler um ein Vielfaches überstieg.

Nach einer Dimensionsreduktion des vorliegenden, hochdimensionalen Datensatzes durch die Hauptkomponentenanalyse (Principal Components Analysis, PCA), mit Hilfe derer alle vorhandenen Datenpunkte in ein neues, orthogonales Koordinatensystem sortiert wurden, wobei die Orientierung der Richtung der grössten Varianz angepasst wurde, konnten die biologische Varianz und der analytischer Fehler graphisch dargestellt werden. Abbildung 2.2.1-A zeigt, dass die auf die Varianz zurückzuführende Streuung in den biologischen Replikaten um ein Vielfaches grösser war, während die technischen Replikate eine wesentlich geringere Streuung und somit Varianz aufwiesen. Dass diese Beobachtung nicht nur auf einige wenige Metabolite zurückzuführen ist, zeigt Abbildung 2.2.1-B. Hier wird deutlich, dass nach vorangegangener Dimensionsreduktion sehr viele



**Abbildung 2.2.1 : Gegenüberstellung von analytischem Fehler und biologischer Varianz** (aus [178]). Nach einer Dimensionsreduktion durch die Hauptkomponentenanalyse  $\log_{10}$ -transformierter Metabolitdaten (normalisiert auf Frischgewicht und Interne Standards) von *A. thaliana* Wildtyp Col-0 und PGM konnte die hochdimensionale Datenmatrix visualisiert werden. **Abb. 2.2.1-A:** Streuung der technischen und biologischen Replikate resultierend aus der Analyse von jeweils 10 unabhängigen Metabolitprofilen. Jeder Punkt repräsentiert 80 quantitativ bestimmte Metabolite aus einer Probe. **Abb. 2.2.1-B:** Score-plot der zugehörigen Faktorladungen, welcher den Einfluss der einzelnen Metabolite auf die Anordnung der Proben (Abb. 2.2.1-A) entlang der jeweiligen Hauptkomponente widerspiegelt .

Faktoren bzw. Metabolite die biologische Varianz in Richtung der ersten beiden Hauptkomponenten bestimmten. Zudem erbrachte diese Analyse, dass der Einfluss von Matrixeffekten auf die Akkuratess der Metabolitquantifizierung eine zu vernachlässigende Grösse darstellte.

**Tabelle 2.2.1: Gegenüberstellung von analytischem Fehler und biologischer Varianz.** Resultierend aus der Analyse von jeweils 10 technischen Replikaten der *A. thaliana* Genotypen Col-0 sowie der stärkelosen PGM Mutante wurde der analytische Fehler ermittelt. Die biologischen Varianzen resultieren aus der Analyse von jeweils 10 individuellen Pflanzen, welche simultan unter identischen Bedingungen kultiviert wurden.

Metabolit	analytischer Fehler CV [%]		biologische Varianz CV [%]	
	COL-0	PGM	COL-0	PGM
O-Acetylserin	26	21	35	29
4-Aminobutyrat	22	18	26	26
Alanin	13	13	46	48
β-Alanin	7	7	21	42
Allantoin	23	27	41	53
Arabinose	10	10	31	17
Arginin	13	12	32	47
Ascorbat	45	38	54	47
Asparagin	12	12	47	51
Aspartat	10	10	27	37
Benzoessäure	10	10	57	58
Citramalat	14	12	54	56
Citrat	19	19	38	46
Citrullin	12	12	30	27
Cystein	21	19	46	39
Dehydroascorbat	24	24	41	48
Erythritol/Threitol	9	12	50	47
Erythronat/Threonat	41	39	68	64
Erythronat-/Threonat-, 1,4-lacton	45	52	65	65
Ethanolamin	29	29	31	46
3-Ethyl-3-hydroxy-3-methylvalerat	21	23	31	29
Fructose	12	9	47	31
Fructose 6-phosphat	19	17	42	35
Fucose	7	7	22	17
Fumarat	6	6	52	111
Galactinol	12	15	71	62
Galactonat/Gluconat	10	10	24	18
Glucose	6	6	30	29
Glucose 1-phosphat	12	12	72	74
Glucose 6-phosphat	16	16	69	89
Glutamat	24	22	41	43
Glutamin	32	34	34	36
Glycerat	7	7	32	28
Glycerol	38	36	85	73
Glycerolphosphat	16	16	44	43
Glycin	9	9	68	28
Homoserin	16	14	44	41
4-Hydroxybenzoat	23	26	43	42

**Tabelle 2.2.1 : Gegenüberstellung von analytischem Fehler und biologischer Varianz.**

(Fortsetzung) Resultierend aus der Analyse von jeweils 10 technischen Replikaten der *A. thaliana* Genotypen Col-0 sowie der stärkelosen PGM Mutante wurde der analytische Fehler ermittelt. Die biologischen Varianzen resultieren aus der Analyse von jeweils 10 individuellen Pflanzen, welche simultan unter identischen Bedingungen kultiviert wurden.

Metabolit	analytischer Fehler CV [%]		biologische Varianz CV [%]	
	COL-0	PGM	COL-0	PGM
6-Hydroxynicotinat	23	24	59	65
4-Hydroxyprolin	19	21	61	67
Hydroxylamin	20	17	53	42
2-Ketoglutarat	23	27	31	35
Indol 3-acetonitril	14	14	38	32
<i>myo</i> - Inositol	6	6	37	43
Isoleucin	8	8	42	21
Leucin	14	14	21	24
Lysin	8	8	36	41
Lyxose	12	13	43	42
Malat	11	11	23	25
Maltose	27	27	35	42
Mannose	21	18	43	32
Methionin	11	10	21	22
3-Methyl-3-hydroxyglutarat	19	19	30	23
Ornithin	11	11	29	107
Oxalat	21	19	28	31
Phenylalanin	10	10	41	34
Phosphat	23	23	55	27
Prolin	11	11	70	42
Propylamin 2,3-diol	22	14	29	33
Psicose	14	14	41	23
Putrescin	22	22	57	65
Pyroglutamat	12	12	36	36
Pyruvat	15	15	40	23
Raffinose	14	14	98	135
Ribonat	15	16	33	34
Ribose	12	11	45	34
Ribulose/Xylulose	14	17	27	29
Saccharose	11	11	48	23
Salicylat	12	14	31	36
Serin	8	8	43	18
Shikimat	14	19	29	32
<i>cis</i> - Sinapat	11	11	28	26
<i>trans</i> - Sinapat	9	9	50	77
Sorbitol	11	15	29	32
Spermidin	9	9	43	34
Spermin	9	11	38	31
Suberylglycin	43	31	45	43
Succinat	16	16	29	33
Threonin	8	8	33	32
Trehalose	7	7	47	17
Tyrosin	6	6	32	28
Uracil	14	15	63	54
Urat	21	19	50	56
Urea	25	24	68	73
Valin	9	9	23	25
Xylose	16	13	49	54

### 2.3 Metabolitvariabilität und Korrelationen

Wie im Kapitel 2.2 gezeigt, weisen Metabolite eine bemerkenswerte Variabilität in ihrer Abundanz auf. Dies gilt auch, wenn Pflanzen des gleichen Genotyps unter identischen Bedingungen kultiviert werden. Diese Variabilität ermöglicht die Konstruktion komplexer Netzwerke, welche wiederum auf Korrelationen zwischen einzelnen Metabolitpaaren zurückgeführt werden können. Da die Verknüpfung von beobachteten Korrelationsmustern oder auch der Topologie von Korrelationsnetzwerken mit dem heutigen Wissen über biochemische Stoffwechselwege bisher erst im Ansatz gelang, wurde eine vergleichende, speziesübergreifende Analyse dieser Netzwerke durchgeführt. Dazu wurde Blattgewebe von *A. thaliana*, *Nicotiana tabacum* und *Solanum tuberosum*, sowie Knollengewebe von *S. tuberosum* dem Metabolite Profiling nach 8.4 unterzogen. Die verschiedenen Spezies wurden jeweils unter identischen Bedingungen kultiviert (siehe 8.1.3) und die spezifischen Organe jeder Pflanze individuell analysiert. In den resultierenden Datensätze wurde der lineare Zusammenhang zwischen zwei Metaboliten A und B durch Pearson'schen Produktmoment-Korrelationskoeffizienten gemäss

$$(1) \quad r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}}$$

berechnet. Die Bewertung der statistischen Signifikanz der Korrelationen erfolgte nach der Student t-Transformation gemäss

$$(2) \quad t = r_{xy} \frac{\sqrt{N-2}}{\sqrt{1-r_{xy}^2}},$$

wobei N der Anzahl der Messpunkte eines jeden Metabolitpaares entspricht. Die Signifikanzschwelle zur Verwerfung der Nullhypothese unkorrelierter Metabolitpaare wurde auf eine Irrtumswahrscheinlichkeit von  $p \leq 0.001$  festgesetzt. Aus der im vorangegangenen Kapitel beschriebenen Analyse biologischer und technischer Replikate konnte ein zusätzliches Kriterium zur Bewertung der Signifikanz paarweiser Metabolitkorrelationen abgeleitet werden. Der aus den Analysen der untersuchten Spezies und Gewebe resultierende Variationskoeffizient eines jeden Metaboliten musste dabei den im vorangegangenen Kapitel festgeschriebenen analytischen Fehler übersteigen.

Die Unterschiede zwischen Korrelationen in den verschiedenen Organen der einzelnen Spezies wurden gemäss der z-Transformation nach

$$(3) \quad z = \frac{1}{2} \frac{\log\left(\frac{1+r_1}{1-r_1}\right) - \log\left(\frac{1+r_2}{1-r_2}\right)}{\sqrt{\frac{1}{N_1-3} + \frac{1}{N_2-3}}}$$

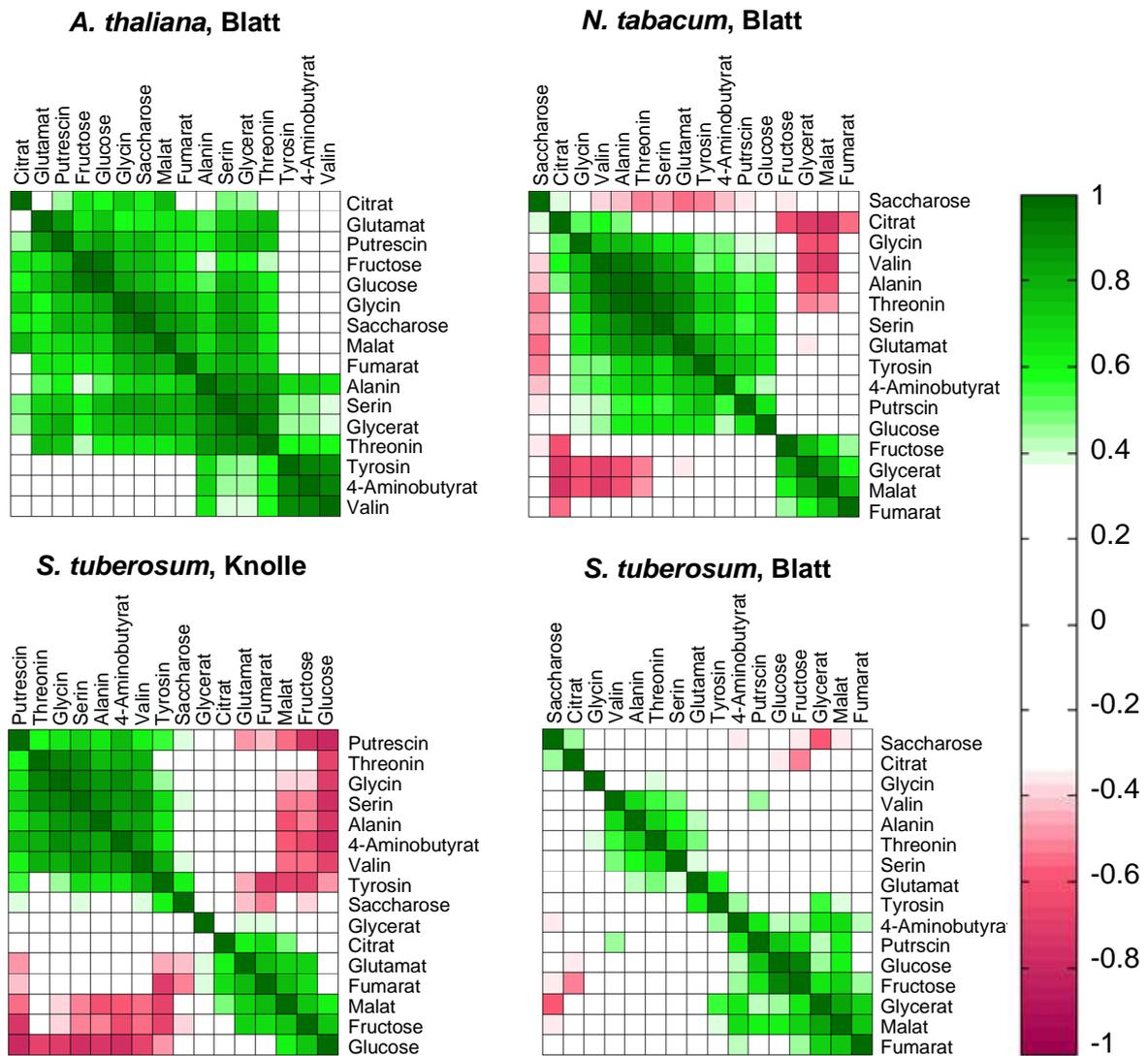
getestet, wobei  $r_1$  und  $r_2$  die Pearson'schen Produktmoment-Korrelationskoeffizienten in den unterschiedlichen Organen darstellen und  $N_1$  und  $N_2$  der Anzahl der gemessenen Metabolitpaare und somit der Anzahl der analysierten biologischen Replikate in diesen Organen entspricht.

Abbildung 2.3.1 visualisiert die erhaltenen, auf biologische Variabilität beruhenden Korrelationsprofile ausgewählter Metabolite der einzelnen Gewebe. Die farbcodierte Darstellung macht deutlich, dass ausgeprägte Unterschiede in den exprimierten Netzwerken der einzelnen Spezies und/oder Organen existierten.

## 2.4 Differentielle und konservierte Korrelationen

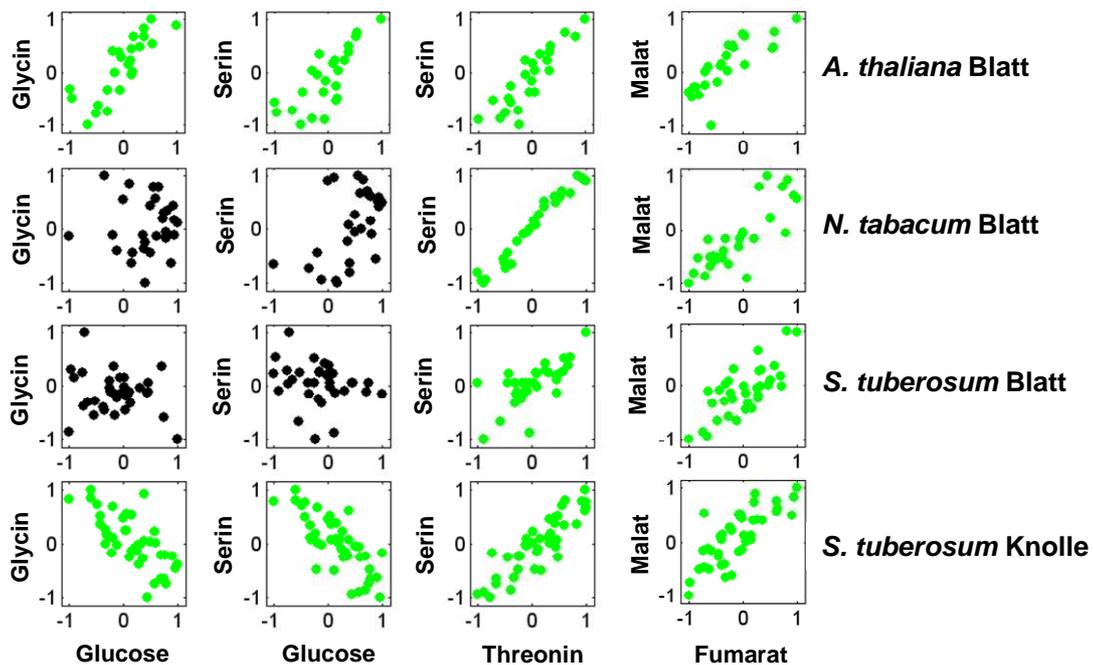
Wenngleich Korrelationen zwischen Metabolitpaaren bis dato durch zwischen verschiedenen Genotypen oder experimentellen Bedingungen vergleichende Analysen im Ansatz gedeutet werden können, ist eine weit reichende Interpretation dieses Phänomens bislang nicht gelungen. Daher wurde eine detaillierte Analyse von Korrelationen zwischen Metabolitpaaren verschiedener Gewebetypen unterschiedlicher Spezies vorgenommen. Dazu wurden, wie im vorangegangenen Kapitel beschrieben, vier prototypische Beispiele getestet: neben Blattmaterial von *A. thaliana*, *N. tabacum* und *S. tuberosum* welches als Source-Gewebe durch CO<sub>2</sub>-Fixierung und Saccharosesynthese dominiert ist, wurden *S. tuberosum* Knollen stellvertretend für nicht-grünes Sink-Gewebe, welches auf den Import von Photoassimilaten aus den voll entwickelten Source-Blättern abhängig ist, analysiert.

In Übereinstimmung zu früheren Studien [94,98] ergab diese Analyse, dass die meisten Metabolite ihr korrelatives Verhalten gegenüber anderen Metaboliten spezies- und gewebeübergreifend nur leicht ändern. Daneben konnten signifikante Korrelationen zwischen Metabolitpaaren identifiziert werden, welche in allen vier untersuchten Datensätzen konserviert waren. Dem gegenüber zeigten andere Metabolite deutlich veränderte Korrelationsmuster. Solche differentiellen Korrelationen, wie sie z.B. zwischen Glycin und Glucose oder Serin und Glucose beobachtet wurden, können intuitiv auf



**Abbildung 2.3.1:** Farbcodierte Darstellung der Korrelationsprofile in Blattgewebe von *A. thaliana*, *N. tabacum* und *S. tuberosum* sowie *S. tuberosum* Knollengewebe (aus [179]). Dargestellt sind die Korrelationsprofile ausgewählter paarweiser Metabolitkorrelationen, welche nach dem Pearson'schen Produktmoment-Korrelationskoeffizienten berechnet und auf ihre Signifikanz getestet wurden. Mit ( $p \leq 0.001$ ) signifikante positive Korrelationen sind dabei rot, negative Korrelationen grün dargestellt. Deutlich sichtbar sind die zwischen den einzelnen Spezies und/oder Organen existierenden ausgeprägten Unterschiede des aktiv exprimierten metabolischen Netzwerks.

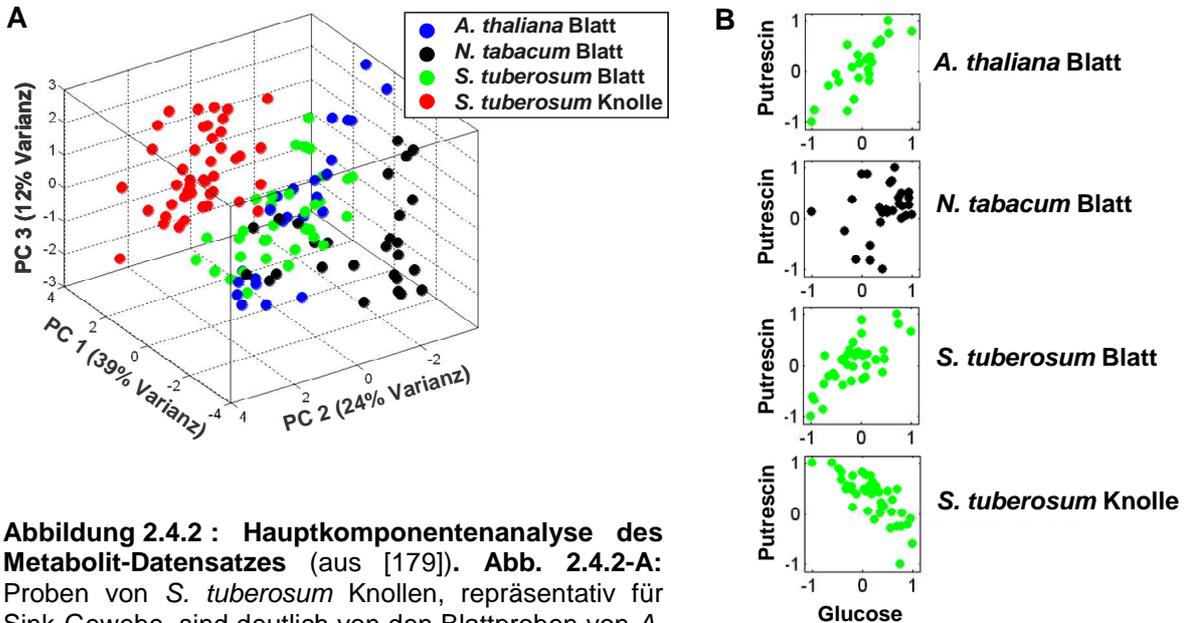
unterschiedliche physiologische Zusammenhänge des C/N-Metabolismus, welcher in Pflanzen stark verknüpft und in Sink-Geweben im Vergleich zu Source-Geweben revers reguliert ist [180,181], verstanden werden. Diese in Abbildung 2.4.1 zusammengefassten Ergebnisse lassen erkennen, dass sowohl konservierte als auch differentielle Korrelationen zwischen einzelnen Metabolitpaaren in den analysierten Geweben auftraten. Daneben wurden deutliche Unterschiede zwischen den Strukturen der Netzwerke der Source (Blatt)- und Sink (Knollen)-Gewebe offenbar.



**Abbildung 2.4.1 : Beispiele paarweiser Metabolitkorrelationen** (aus [179]). Analysiert wurde Blattmaterial von *A. thaliana* (24 Proben), *N. tabacum* (29 Proben), *S. tuberosum* (34 Proben) sowie *S. tuberosum* Knollenmaterial (43 Proben). Aus den resultierenden Datensätzen wurde der lineare Zusammenhang zweier Metabolite nach dem Pearson'schen Produktmoment-Korrelationskoeffizienten errechnet. Signifikante Korrelationen ( $p \leq 0.001$ ) sind grün gefärbt, Streudiagramme von Metabolitpaaren, die dieses Kriterium nicht erfüllten, sind schwarz dargestellt. Während in den ersten beiden Spalten Beispiele differentieller Korrelationen gezeigt sind, werden konservierte Korrelationen durch die dritte und vierte Spalte repräsentiert.

Nach vorangegangener Dimensionsreduktion durch eine Hauptkomponentenanalyse konnten die vorliegenden Daten visualisiert werden. Abbildung 2.4.2-A zeigt die Verschiedenartigkeit der einzelnen Gewebe. Die das Sink-Gewebe repräsentierenden Proben der Kartoffelknollen weisen eine deutliche Distanz zu den einen überlappenden Cluster bildenden Blattproben auf, welche stellvertretend für Source-Gewebe stehen. Anhand der Faktorladungen der ersten Hauptkomponente, welche der Wichtung der Variablen auf die Varianz in Richtung dieser Achse entsprechen, konnte Glucose als der Metabolit mit dem grössten Einfluss auf die Separation der Sink- von den Source-Geweben identifiziert werden. Die Korrelation dieses Metaboliten mit Putrescin in den unterschiedlichen Gewebetypen liess ein weiteres Phänomen erkennbar werden. Während in *S. tuberosum* Blattgewebe eine positive Korrelation beobachtet wurde, zeigte sich in den Sink-Gewebe repräsentierenden Proben aus den Kartoffelknollen eine negative Korrelation (Abbildung 2.4.2-B). Somit konnte ein weiteres Beispiel differentieller Korrelationen identifiziert werden, wobei sich der lineare Zusammenhang zwischen den Metaboliten bildlich durch unterschiedliche Steigungen der resultierenden Regressions-

geraden im zweidimensionalen kartesischen Koordinatensystem darstellte. Mögliche Ursachen dieser experimentell beobachteten Umkehrkorrelationen werden in den folgenden Kapiteln aufgezeigt.



**Abbildung 2.4.2 : Hauptkomponentenanalyse des Metabolit-Datensatzes** (aus [179]). **Abb. 2.4.2-A:** Proben von *S. tuberosum* Knollen, repräsentativ für Sink-Gewebe, sind deutlich von den Blattproben von *A. thaliana*, *N. tabacum* und *S. tuberosum* getrennt. Die Analyse der Faktorladungen ergab, dass Glucose den grössten Beitrag zur Varianz entlang der ersten Hauptkomponente leistete und somit entscheidend zur Trennung der Klassen beitrug. **Abb. 2.4.2-B:** Korrelationen zwischen Glucose und Putrescine in einzelnen Geweben verschiedener Spezies (mit  $p < 0.001$  signifikante Korrelationen sind grün dargestellt). Während beide Metabolite in Source-Geweben positiv oder nicht korreliert sind, kehrt sich das Korrelationsverhalten im Falle des Sink-Gewebes um, hier liegt eine negative Korrelation vor. Somit konnte ein weiteres Beispiel differenzieller Korrelationen identifiziert werden.

## 2.5 Der Ursprung beobachteter Korrelationen – ein numerisches Beispiel

An einem vereinfachten Beispiel des Calvin Zyclus mit anschliessender cytosolischer Saccharosesynthese lassen sich mögliche Ursachen auftretender Korrelationen ableiten. Tatsächlich ist auch unter konstanten experimentellen Bedingungen das in Abbildung 2.5.1-A dargestellte Reaktionssystem durch zeitlich variierende Faktoren beeinflusst, welche zu einer gewissen Variabilität einzelner Metabolitlevel führen. Um diesen Effekt in einem numerischen Beispiel zu simulieren wurde angenommen, dass die Lichtintensität, und damit verbunden die effektive ATP-Produktion, die fluktuierenden Komponenten darstellen. Diese simulierten Fluktuationen pertubieren durch den gesamten Pathway und



$$(2) \quad v_{\text{Licht}} = V_{\text{Licht}} [\text{ADP}] \cdot [\text{P}_{\text{ch}}]$$

angenommen, wobei  $V_{\text{Licht}}$  stellvertretend für die effektive Lichtintensität steht. Die Synthese von Saccharose 6-phosphat wurde nach

$$(3) \quad v_3 = k_3 [\text{F6P}_{\text{cyt}}] \cdot [\text{UDP-Glc}] \cdot f_1([\text{P}_{\text{cyt}}]) \cdot f_2([\text{G6P}_{\text{cyt}}])$$

mit  $f_1 = (1 + [\text{P}_{\text{cyt}}] / K_p)^{-1}$  und  $f_2 = (1 + [\text{G6P}_{\text{cyt}}] / K_g)^{-1}$  berechnet. Den plastidiären Reaktionen wurde eine weitere Reaktion (4) zugefügt, um den ATP-Verbrauch der Konversionsreaktion von 3-Phosphoglycerat zu 1,3-Bisphosphoglycerat zu berücksichtigen:



Das System umfasste weiterhin 3 Konservierungen: neben  $\text{ADP} + \text{ATP} = A_{\text{total}} = 2$  wurde das Gesamtphosphat sowohl im Chloroplasten als auch im Cytosol ( $\text{P}_{\text{ch,total}} = 25$  und  $\text{P}_{\text{cyt,total}} = 5$ ) als konstant angenommen. Der Faktor 5 der Stöchiometrie trägt den Volumenverhältnissen innerhalb der Zelle Rechnung; 20% des Zellvolumens wurden den Chloroplasten zugerechnet. Da das hier beschriebene Modell vorrangig illustrativen Zwecken dienen sollte, flossen alle Geschwindigkeitskonstanten sowie Parameter der in Abbildung 2.5.1-A gezeigten Reaktionen arbiträr mit  $k_{+1} = 1.0$ ,  $k_1 = k_{+1}/q_1$  mit  $q_1 = 7.1$ ,  $k_2 = 1.0$ ,  $k_3 = 1.0$ ,  $K_p = 1.0$ ,  $K_g = 1.0$ ,  $k_{+4} = 10.0$ ,  $k_4 = k_{+4}/q_4$  mit  $q_4 = 2.3$ ,  $k_5 = 1.0$ ,  $k_6 = 1.0$ ,  $k_7 = 1.0$ ,  $k_8 = 0.1$ ,  $k_9 = 1.0$ ,  $k_{10} = 1.0$ ,  $k_{11} = 1.0$ ,  $k_{12} = 3.0$ ,  $k_{13} = 3.0$ ,  $k_{\pm 14} = k_{\pm 4}$ ,  $k_{15} = 0.1$ ,  $k_{16} = 1.0$  und  $k_{17} = 1.0$  in die Kalkulationen ein. Das gesamte, hier beschriebene System konnte durch insgesamt 15 unabhängige Differentialgleichungen beschrieben werden.

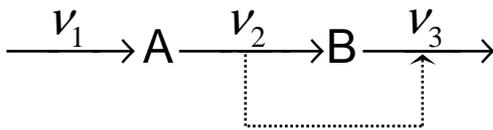
Dieses vereinfachte numerische Beispiel sollte dazu dienen, einige bedeutende Aspekte, die Interpretation der beobachteten Korrelationen betreffend, zu demonstrieren. Wie in Abbildung 2.5.1 gezeigt, besteht zwischen auftretenden Korrelationen und dem zugrunde liegenden biochemischen Pathway keine direkte Verbindung. Während einige Korrelationen in der Tat intuitiv verstanden werden können, wie z.B. der starke lineare Zusammenhang zwischen G6P und F6P (Abbildung 2.5.1-B), sind die meisten Korrelationen, insbesondere zwischen im Pathway weit voneinander entfernt liegenden Metaboliten, nicht auf der Basis einfacher heuristischer Zusammenhänge zu erklären.

Detaillierte numerische Modelle des Calvin Cyclus können den vorangegangenen Arbeiten von Petterson *et Ryde-Petterson* [183], Poolmann *et al.* [184] oder Noe *et Girsch* [185] entnommen werden.

## 2.6 Umkehrungen von Korrelationen – ein weiteres numerisches Beispiel

Wie in Kapitel 2.4 beschrieben, wurden neben den in allen untersuchten Datensätzen konservierten Korrelationen auch differentielle Korrelationen zwischen Metabolitpaaren identifiziert, welche sich je nach Herkunft des analysierten Gewebes unterschieden. Von besonderem Interesse waren dabei differentielle Korrelationen mit veränderlichem Richtungskoeffizienten. Solch eine Umkehr von positiver zu negativer Korrelation konnte unter anderem in *S. tuberosum* Sink- und Source-Geweben für Glucose und Putrescin beobachtet werden (Abbildung 2.4.1-B).

Um zu verstehen, wie oder warum solche Umkehrungen innerhalb zellulärer oder metabolischer Netzwerke auftreten, wurde ein weiteres numerisches Beispiel zur Hilfe genommen. Dazu genügte ein minimalistisches Reaktionssystem, wie in Abbildung 2.6.1 gezeigt, welches aus nur zwei Metaboliten A und B zusammengesetzt ist, wobei angenommen wurde, dass die Degradation des Metaboliten B durch einen feed-forward Mechanismus in Abhängigkeit von seiner Konzentration inhibiert wurde.



**Abbildung 2.6.1: Modell des minimalistischen, der numerischen Simulation zugrunde liegenden metabolischen Reaktionssystems.** Für die Degradation des Metaboliten B wurde angenommen, dass diese konzentrationsabhängig nach einem feed-forward Mechanismus reguliert wird.

Für das sich aus folgenden Ratengleichungen (1) und Parametern (2)

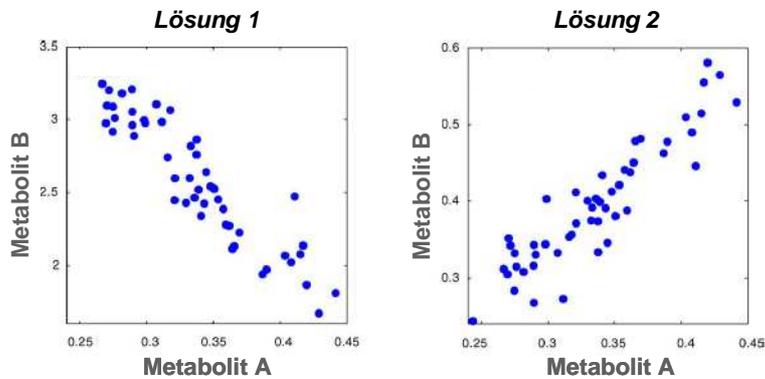
$$(1) \quad \frac{d}{dt} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} +1 & -1 & 0 \\ 0 & +1 & -1 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}, \text{ mit } v_1 = \text{konst.}, v_2 = k_2 \cdot A, v_3 = \frac{k_3 B}{1 + \left(\frac{B}{K_1}\right)^n}$$

$$(2) \quad k_2=1, k_3=1, K_1=1 \text{ und } n=2$$

zusammen setzende Differentialgleichungssystem ergaben sich zwei unabhängige steady-state Lösungen:  $A^s = c/k_2$  und  $B_{1,2}^s = K_1 \left( 1 \pm \sqrt{1 - 4\alpha^2} / 2\alpha \right)$  mit  $\alpha = c/K_1 k_3$ .

Biologische Replikate wurden für jede dieser beiden Lösungen durch zufällige Fluktuationen des Influxparameters  $c$  gemäss einer Gaussischen-Normalverteilung um

den Mittelwert 0.35 simuliert. Die aus beiden Lösungen, welche die steady-state Bedingungen erfüllen, resultierenden Korrelationen sind in Abbildung 2.6.2 gezeigt.



**Abbildung 2.6.2 : Beispiel differentieller Metabolit-Korrelationen, numerisch durch Simulation eines minimalen Reaktionssystems erzeugt.** Die resultierenden Korrelationsmuster der beiden steady-state Lösungen gleichen den experimentell beobachteten Umkehrkorrelationen.

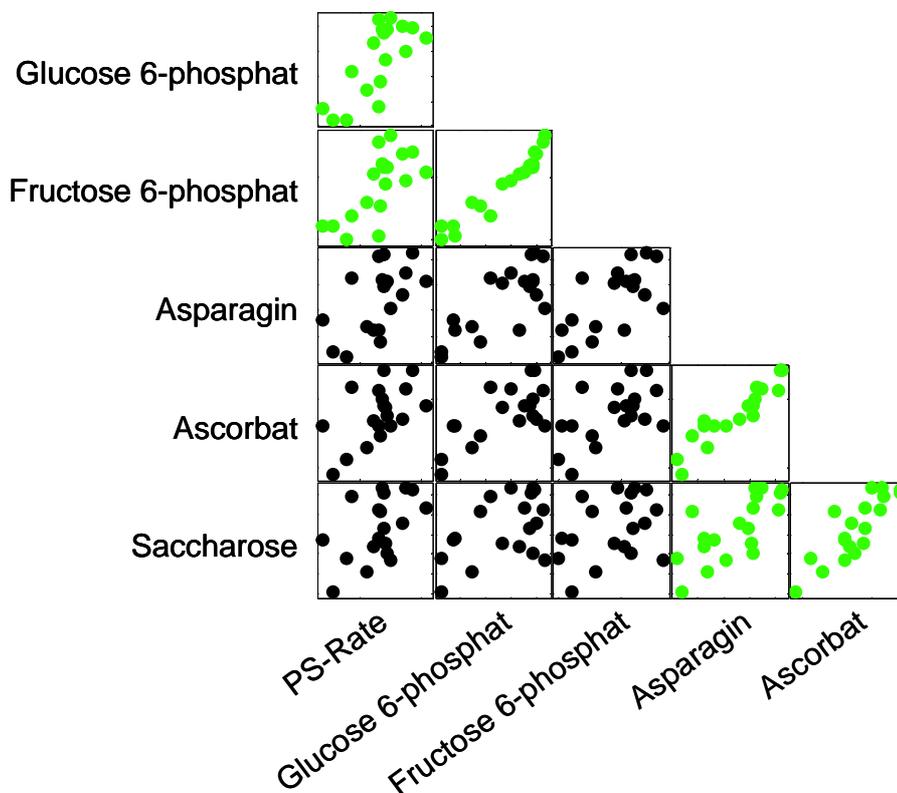
Somit konnte gezeigt werden, dass ein einfaches System genügt, um Umkehrkorrelationen numerisch zu erzeugen, welche abhängig vom physiologischen Status des untersuchten biologischen Gewebes auch experimentell beobachtet wurden. In der vorgestellten Simulation resultierten die multistationären Lösungen des sich aus diversen Parametern zusammensetzenden Differentialgleichungssystems zur Beschreibung des minimalen Reaktionssystems in einem Wechsel der paarweisen linearen Abhängigkeit beteiligter Metabolite von positiv zu negativ. Numerisch induziert wurden die Korrelationen, wie auch im vorangegangenen Beispiel, durch minimale Fluktuationen eines extrinsischen Parameters.

## 2.7 Einfluss extrinsischer Parameter auf die Metabolitvariationen

In den vorangegangenen numerischen Beispielen induzierten minimale Fluktuationen extrinsischer Parameter unterschiedliche Ausprägungen der zugrunde liegenden metabolischen Netzwerke. Um die Existenz dieses postulierten Phänomens experimentell zu bestätigen, wurden von insgesamt 20 voll expandierten *A. thaliana* Col-0 Source-Blättern die Photosyntheseraten in der Mitte der 12-stündigen Photoperiode fluorometrisch mit dem IMAGING-PAM Fluorometer (Fa. Walz, Effeltrich, Germany) gemäss den Angaben des Herstellers ermittelt. Im Anschluss wurde jedes Rosettenblatt mittels GC-TOF-MS individuell analysiert. Die ermittelten Konzentrationsprofile von insgesamt 82 identifizierten Metaboliten wurden auf ihre Beziehungen zur Photosyntheserate, deren Werte eine gewisse Variabilität aufwiesen, getestet. Die Photosyntheserate spiegelt unmittelbar die auf jedes Blatt individuell einwirkende

Lichtintensität wieder. Als Mass für den linearen Zusammenhang zwischen den einzelnen Variablen diene wieder der Pearson'sche Produktmoment-Korrelationskoeffizient.

Tatsächlich konnte der stärkste lineare Zusammenhang zwischen der Photosyntheserate und Fructose 6-phosphat sowie Glucose 6-phosphat nachgewiesen werden, während die meisten der insgesamt 82 in 20 Replikaten quantifizierte Metabolite keine mit  $p \leq 0.05$  signifikante Co-Response zeigten (Abbildung 2.7.1). Somit konnte der Einfluss extrinsischer Parameter auf die Variabilität bestimmter Metabolite, welche sich gemäss der vorgestellten numerischen Simulationen in ganz bestimmten Korrelationsmustern manifestiert, erstmals experimentell nachgewiesen werden.



**Abbildung 2.7.1 Beziehungen zwischen der Photosyntheserate und der Variabilität einzelner Metabolitkonzentrationen in *A. thaliana*.** Dargestellt ist die Scatterplotmatrix der 5 Metabolite, welche die gemäss des Pearson'schen Produktmoment-Korrelationskoeffizienten stärkste Co-Response zur Photosyntheserate (PS-Rate) aufwiesen. Signifikante Korrelationen ( $p \leq 0.05$ ) sind grün dargestellt. Neben Glucose 6-phosphat und Fructose 6-phosphat konnte kein weiterer Metabolit identifiziert werden, welcher mit der Photosyntheserate korrelierte.

## 2.8 Diskussion

Die Bestimmung und Interpretation der *in vivo* auftretenden Metabolitdynamik stellt heute eine der grössten Herausforderungen der modernen Biochemie dar. Da die Menge an

verfügbaren biologischen Daten durch den Einsatz neuartiger high-throughput Methoden stetig wächst, wird auch die Interpretation dieser Datenmengen immer anspruchsvoller. Eine gängige Methode der Verarbeitung der mit den so genannten 'Omics'-Technologien gewonnenen Daten stellen Netzwerkanalysen dar.

Kose *et al.* schlugen 2001 erstmals eine netzwerkbasierte Methode vor, mit deren Hilfe multiparallel gewonnene Metabolitdaten pflanzlichen Gewebes visualisiert werden können [186]. Diese Methode basiert auf der Konnektivität von Metaboliten, wobei ausgehend von linearen Korrelationen zwischen Metabolitpaaren Datensätze algorithmisch nach Knotenpunkten durchsucht werden. Von diesen sog. Cliques, welche Cluster von Metaboliten mit höchster Verknüpfung darstellen, wurde angenommen, dass diese potenzielle Verzweigungsstellen oder Verbindungsglieder diskreter Stoffwechselwege innerhalb des aktiv exprimierten Netzwerkes darstellen. Die Interpretation dieser Cliques stellte sich jedoch als äusserst schwierig dar.

Weckwerth *et al.* schlugen daher später vor, die Metabolitdynamik basierend auf vergleichenden Korrelationsanalysen zwischen Metabolitleveln von Kontrollpflanzen und behandelten oder genetisch modifizierten Pflanzen zu untersuchen [98,165].

Im vorliegenden Kapitel wurde dieser Ansatz erweitert und eine vergleichende Analyse von Korrelationsnetzwerken verschiedener Pflanzenspezies und Gewebetypen vorgestellt. Da die inhärente Variabilität unabhängiger biologischer Proben die Grundlage für die Konstruktion solcher Netzwerke darstellt, wurde zunächst die Eignung der vorliegenden, mittels GC-TOF-MS gewonnenen Metabolitdaten für diese Art der Netzwerkanalyse getestet.

### **Korrelationen zwischen Metabolitpaaren haben biologische Ursachen**

Durch die Analyse technischer Replikate gelang es, die in biologischen Replikaten mit der hier angewandten Methode des Metabolite Profiling messbaren Variationen einzelner Metabolitkonzentrationen zweifelsfrei auf die biologische Varianz zurückzuführen (Abbildung 2.2.1). Basierend auf dieser biologischen Varianz wurden Metabolit-Datensätze von *A. thaliana*, *N. tabacum* und *S. tuberosum* sowie Knollengewebe von *S. tuberosum*, welche aus der Analyse biologischer Replikate jeder Spezies, welche unter konstanten Bedingungen kultiviert wurden, resultierten, nach paarweisen Korrelationen durchsucht.

Die komparative Analyse dieser Korrelationsnetzwerke erbrachte in Übereinstimmung zu früheren Studien, dass die nachweisbare Konnektivität zwischen Metaboliten auf eine geringe, jedoch signifikante Anzahl beschränkt ist und je nach Herkunft des untersuchten Gewebes variiert [98,187,188]. Durch die detaillierte Korrelationsanalyse konnten Wechselbeziehungen zwischen ausgewählten Metabolitpaaren nachgewiesen werden,

welche in den Geweben aller vier untersuchten Spezies konserviert waren. Diese Beobachtung schliesst weiter aus, dass Korrelationen die Folge zufälliger Messunsicherheiten sind und somit artifizielle Phänomene darstellen. Daneben konnten auch systematische Messunsicherheiten als Ursache beobachteter Korrelationen ausgeschlossen werden, da die Mehrheit der Metabolitpaare unkorreliert erschien und darüber hinaus differentielle Wechselbeziehungen beobachtet wurden. Insbesondere das beobachtete Phänomen der Umkehrung linearer Abhängigkeiten von positiv zu negativ zeigt, dass intrinsische Effekte der Messmethodik und Korrelationen zwischen Metabolitpaaren keine kausalen Ereignisse darstellen (Abbildung 2.4.1 und 2.4.2).

Gleichzeitig machten die Korrelationsprofile deutlich, dass die intuitive Annahme, dass Korrelationen vornehmlich zwischen Metaboliten auftreten, welche im biochemischen Reaktionsnetzwerk benachbart sind, nicht zutrifft (Abbildung 2.5.1). Auch in früheren Untersuchungen unterschiedlicher Organismen mit komplementären Analysetechniken wurden signifikante Korrelationen zwischen Metaboliten beobachtet, welche im biochemischen Reaktionsnetzwerk sehr weit voneinander entfernt liegen, während zwischen benachbarten Metaboliten häufig kein linearer Zusammenhang nachweisbar war [94,98,165,178]. Dieses erschwert bis heute die Interpretation von Korrelationsnetzwerken erheblich und erfordert zum Verständnis der beobachteten Korrelationsmuster die Verwendung stochastischer Modelle.

***Korrelationen zwischen Metabolitpaaren sind für definierte physiologische Zustände charakteristisch***

Durch die numerische Simulation eines detaillierten Modells zur Beschreibung glycolytischer Oszillationen in Hefe konnten Steuer *et al.* (2003) zeigen, dass minimale Fluktuationen in biochemischen Netzwerken Korrelationsmuster erzeugen, welche intuitiv nicht interpretierbar sind, aber dennoch als Konsequenz des zugrunde liegenden, komplexen enzymatischen Reaktionsnetzwerkes ausgeprägt werden [189]. Die entstehenden Korrelationsmuster wurden auf eine Kombination stöchiometrischer und kinetischer Effekte der Gestalt  $\mathbf{J}\mathbf{\Gamma} + \mathbf{\Gamma}\mathbf{J}^T = -2\mathbf{D}$  zurückgeführt, wobei  $\mathbf{\Gamma}$  der Kovarianz-Matrix und somit den Korrelationen,  $\mathbf{J}$  der Jacobi-Matrix sowie  $\mathbf{D}$  der Fluktuationsmatrix entspricht. Aus diesem Zusammenhang zwischen experimentell beobachtbaren Korrelationen und dem zugrunde liegenden Reaktionsnetzwerk folgt unmittelbar, dass sich die Behandlung einer Pflanze oder eine genetische Modifikation, welche ein verändertes enzymatisches Reaktionsnetzwerk bedingt, in einer veränderten Jacobi-Matrix und schliesslich auch einer veränderten Kovarianz- und Korrelationsmatrix manifestiert. Somit können beobachtete Korrelationen in den vorliegenden Metabolit-Datensätzen als direkte Ausprägung des genetischen Hintergrundes interpretiert werden und stellen 'Fingerprints' genau definierter physiologischer Zustände dar. Gleichzeitig wird

jedoch deutlich, dass Korrelationen per se keine Aussage über die Verknüpfung einzelner Stoffwechselprodukte in metabolischen Pathways zulassen. Vielmehr sind Kovarianz und Korrelationen ein Ergebnis des komplexen Zusammenspiels aller Reaktionen und regulatorischen Interaktionen innerhalb eines biologischen Systems.

***Korrelationen zwischen Metabolitpaaren werden durch minimale Fluktuationen extrinsischer Parameter verursacht***

Camacho *et al.* (2005) führten basierend auf numerischen Simulationen sehr starke Korrelationen auf das chemisches Gleichgewicht zwischen den beteiligten Metaboliten zurück und postulierten, dass Metabolite mit ähnlichen physikalischen Eigenschaften auch stark korrelieren. Als häufigste Ursache beobachteter Korrelationen wurden eine ausserordentlich hohe Varianz eines spezifischen Enzyms innerhalb des zugrunde liegenden Systems oder aber die Response zweier Metabolite auf nur einen dominierenden Parameter vorgeschlagen, wobei die Möglichkeit der Identifizierung des entsprechenden Enzyms anhand der Metabolitprofile als unmöglich bewertet wurde [190].

In der vorliegenden Arbeit wurden an numerischen Beispielen weitere mögliche Ursachen der Entstehung linearer Abhängigkeiten aufgezeigt. Dabei konnte die Variabilität einzelner Metabolitkonzentrationen auf minimale Fluktuationen externer Parameter zurückgeführt werden, wie sie in der Natur unvermeidlich auftreten. Dieses können z.B. Variationen in der Lichtintensität, der Temperatur oder der Nährstoffverfügbarkeit sein. Daneben können geringe Unterschiede in der Konzentration oder Aktivität einzelner Enzyme oder auch metabolische Oszillationen zu einer gewissen Variabilität einzelner Metabolitlevel führen. Numerisch konnte weiter gezeigt werden, dass Variationen in den Konzentrationen einzelner Metabolite durch das komplexe biochemische Reaktionsnetzwerk hindurch pertubieren und spezifische Muster paarweiser Metabolitkorrelationen erzeugen (Abbildung 2.5.1).

Die experimentelle Bestätigung, dass geringste Fluktuationen externer Parameter die Konzentrationen bestimmter Metabolite beeinflussen, gelang durch die Korrelation der Photosyntheseraten einzelner Rosettenblätter von *A. thaliana* mit den zugehörigen Metabolitprofilen (Abbildung 2.7.1). Da die Photosyntheserate die Lichtintensität unmittelbar reflektiert, spiegelt die Korrelation dieser Grösse mit Glucose 6-phosphat und Fructose 6-phosphat die Kausalität zwischen der Variation dieser Metabolitkonzentrationen und der Fluktuation eines extrinsischen Parameters wieder, welche den numerischen Simulationen zugrunde lag. In allen untersuchten Datensätzen wurde eine starke Korrelation zwischen Glucose 6-phosphat und Fructose 6-phosphat beobachtet. Nachdem die Varianz dieser Metabolitkonzentrationen auf Fluktuationen der Lichtinten-

sität zurückgeführt wurde, lässt sich der lineare Zusammenhang durch die rasche Interkonversion zwischen diesen Intermediaten zur Aufrechterhaltung des chemischen Gleichgewichtes erklären. Auch die beobachtete konservierte Korrelation zwischen Malat und Fumarat könnte in der raschen Umsetzung innerhalb des TCA-Cyclus begründet liegen.

### ***Differentielle Korrelationen deuten auf regulatorische Knotenpunkte innerhalb des enzymatischen Netzwerkes hin***

Durch die vergleichende, spezieübergreifende Analyse konnte aufgezeigt werden, dass regulatorische Unterschiede innerhalb metabolischer Routen zu einer Inversion der linearen Beziehung zwischen Metaboliten führen können. Numerisch wurde dieser Effekt auf Multistationarität, wie z.B. die Existenz mehrerer steady-state Lösungen eines Differentialgleichungssystems, zurückgeführt. Mit Hilfe solcher Multistationaritäten werden in Simulationen zur Beschreibung zellulärer Regulationsmechanismen häufig beobachtete 'Switching'-Prozesse zwischen unterschiedlichen metabolischen Zuständen erklärt [191]. Auch in Modellen der Regulation von Oszillationen im Zellcyclus während der Mitose oder der Ausdifferenzierung von Zellen in mehrzelligen Organismen wurden biologische Schaltungen mit Hilfe multipler steady-state Zustände beschrieben [192,193]. Poolman *et al.* fanden z.B. in Modellen des photosynthetischen Calvin Cyclus Hinweise auf solch bi-stabiles Verhalten [184], und konnten dieses wenig später auch experimentell bestätigen [194]. Der Übergang zwischen zwei steady-state Zuständen wurde in diesen Studien an die Verfügbarkeit anorganischen Phosphates geknüpft, welche den Kohlenstoffbedarf der Zelle anzeigt. Interpretiert wurde dieses Phänomen als Ausprägung unterschiedlicher metabolischer Zustände in jungen Blättern, welche Kohlenstoffsenken, und älteren Blättern, welche Kohlenstoffquellen darstellen. Auch der in dieser Arbeit angestellte Vergleich der Metabolitprofile von Blatt- und Knollenmaterial reflektiert die Unterschiede zwischen den physiologischen Zuständen in Sink- und Source-Gewebe. In diesem Sinne können die experimentell gefundenen Beispiele von Umkehrungen der linearen Beziehung zwischen Metabolitpaaren, wie z.B. von Glucose und Putrescin in *S. tuberosum* Blatt- und Knollengewebe (Abbildungen 2.4.1 und 2.4.2-B), auf ausgeprägte Unterschiede in der Regulation der beiden Gewebetypen zugrunde liegenden metabolischen Netzwerke zurückgeführt werden. Tatsächlich ist die reverse Regulation des C-/N Metabolismus in Sink- und Source-Geweben literaturbekannt [180,181].

Wenngleich Korrelationen per se lediglich statistische Zusammenhänge aufzeigen und nicht als Beweis für Kausalität betrachtet werden dürfen, konnte gezeigt werden, dass die vergleichende Analyse von Korrelationsnetzwerken das Auffinden potenzieller regulatorischer Knotenpunkte ermöglicht. Dieses erlaubt Rückschlüsse auf die Metabolitdynamik

anhand von Metabolitprofilen, welche den steady-state Zustand eines Organismus reflektieren. Somit supplementiert diese Methode der Datenanalyse die häufig angewandten Verfahren zum Testen auf Mittelwertsunterschiede oder zur Dimensionsreduktion, mit denen immer auch der Verlust von Informationen einhergeht. Der Zusammenhang zwischen der Topologie eines Korrelationsnetzwerkes und dem enzymatischen Netzwerk bietet die Möglichkeit, die Metabolitdynamik zu erfassen und regulatorische Veränderungen in metabolischen Netzwerken zu identifizieren (siehe auch Abbildung 6.1).

Darüber hinaus kann die hier vorgestellte Methode problemlos zur Auswertung von Transkript- oder Proteinexpressions-Datensätzen genutzt werden. Insbesondere die integrative Analyse komplementärer Stoffwechselebenen stellt eine viel versprechende Möglichkeit dar, Co-Regulationen von Metaboliten, Transkripten und Proteinen innerhalb biochemischer Netzwerke aufzudecken und Einblicke in die hinter den linearen Strukturen metabolischer Pathways verborgene Organisation des Metabolismus zu erhalten.

### **3 Untersuchungen der zeitabhängigen systemischen Regulation des Metabolismus von *Arabidopsis thaliana***

#### **3.1 Der circadiane Rhythmus der Pflanze**

Je nach Periodenlänge biologischer Rhythmen unterscheidet man infradiane Rhythmen, deren Frequenz unter der einen Tages liegt, ultradiane Rhythmen, deren Frequenz über der einer Tageslänge liegt sowie circadiane Rhythmen, welche eine Periodenlänge von ungefähr 24 Stunden aufweisen. Der circadiane Rhythmus grüner Pflanzen ist durch eine Vielzahl metabolischer Faktoren und physiologischer Prozesse beeinflusst. So wird beispielsweise ein grosser Teil der während der kurzen Photoperiode unter Mitwirkung von Lichtenergie und Chlorophyll assimilierten Kohlenstoffe normalerweise in Form von transitorischer Stärke in den Chloroplasten grüner Pflanzen kurzfristig gespeichert und dient in der anschliessenden Dunkelperiode als Energielieferant zur Aufrechterhaltung lebenswichtiger Zellfunktionen.

Im vorangegangenen Kapitel 2 zur Untersuchung metabolischer Netzwerke diente die numerische Simulation des photosynthetischen Calvin Cyclus dazu, Beziehungen zwischen Metabolitkonzentrationen zu einem definierten Zeitpunkt herzuleiten. In einem hier vorgestellten Experiment wurden die mit der circadianen Rhythmik verbundenen Auswirkungen auf das metabolische Netzwerk der stärkelosen PGM Mutante untersucht. Dazu wurden die Rosettenblätter von jeweils 10 individuellen, simultan unter Kurztagbedingungen kultivierten *A. thaliana* Pflanzen des Wildtyps Col-0 sowie der plastidiären knock-out Mutante PGM nach Erreichen des Entwicklungsstadiums 5.1 [177] zu 6 verschiedenen Zeitpunkten eines Tages geerntet. Das Blattmaterial wurde zur vergleichenden Analyse komplementärer biochemischer Stoffwechselebenen heran-

gezogen. Da die Quantifizierung sowohl der Proteine als auch der Metabolite mit den hier angewandten massenspektrometrischen Methoden relativ erfolgte, wurden die circadianen Veränderungen im Stoffwechsel zweier *Arabidopsis thaliana* Genotypen vergleichend analysiert. Neben dem Wildtyp Col-0 wurde die chloroplastidiäre PGM Mutante als Modellorganismus gewählt. Diese recht gut charakterisierte Mutante ist durch eine loss-of-function Mutation der chloroplastidiären Phosphoglucomutase (PGM) gekennzeichnet. Dieses Enzym katalysiert die reversible Interkonversion zwischen Glucose 1-phosphat und Glucose 6-phosphat und stellt somit eines der vier Enzyme dar, welche die Stärkesynthese ausgehend von Fructose 6-phosphat in der Pflanze katalysieren. Im Gegensatz zu allen anderen diploiden Pflanzenspezies, welche über zwei Isoformen der PGM verfügen, sind in *A. thaliana* neben der chloroplastidiären zwei weitere Isoformen bekannt, welche wahrscheinlich in den cytoplasmatischen Metabolismus involviert sind [195]. Unter Lang- und/oder Kurztagbedingungen angezogen ist die PGM durch ein im Vergleich zum Wildtyp Col-0 vermindertes Wachstum gekennzeichnet (Abbildung 3.1.1).



**Abbildung 3.1.1: Gegenüberstellung der Phänotypen von *A. thaliana* Col-0 und PGM.** Gezeigt sind die etwa 6 Wochen nach der Aussaat erreichten Entwicklungsstadien der simultan unter Kurztagbedingungen (8h Licht/16h Dunkel) kultivierten Genotypen Col-0 (**Abb. 3.1.1-A**) und PGM (**Abb. 3.1.1-B**).

### ***Integrative Protein- und Metabolitanalyse***

Sollen intrazelluläre Proteine und Metabolite simultan aus einer biologischen Probe extrahiert werden, müssen verschiedene Aspekte der Präparation berücksichtigt werden. Zunächst müssen möglichst alle enzymatischen Aktivitäten umgehend gequenchet werden, um eine weitere Metabolisierung sowie Proteindegradationen zu vermeiden. Daneben sollten möglichst viele Stoff- und Proteinklassen extrahiert werden, um den metabolischen Zustand so umfassend wie möglich charakterisieren zu können. Die gesamte Prozedur der Extraktion sollte so kurz wie möglich gehalten werden, um artifizielle Modifikationen von Proteinen und Metaboliten, wie z.B. Oxidationen, zu vermeiden [90,165,196]. Insbesondere das Profiling des pflanzlichen Metabolismus erfordert aufgrund der ungeheuren Anzahl inhärenter Komponenten eine besondere Sorgfalt bei der Präparation

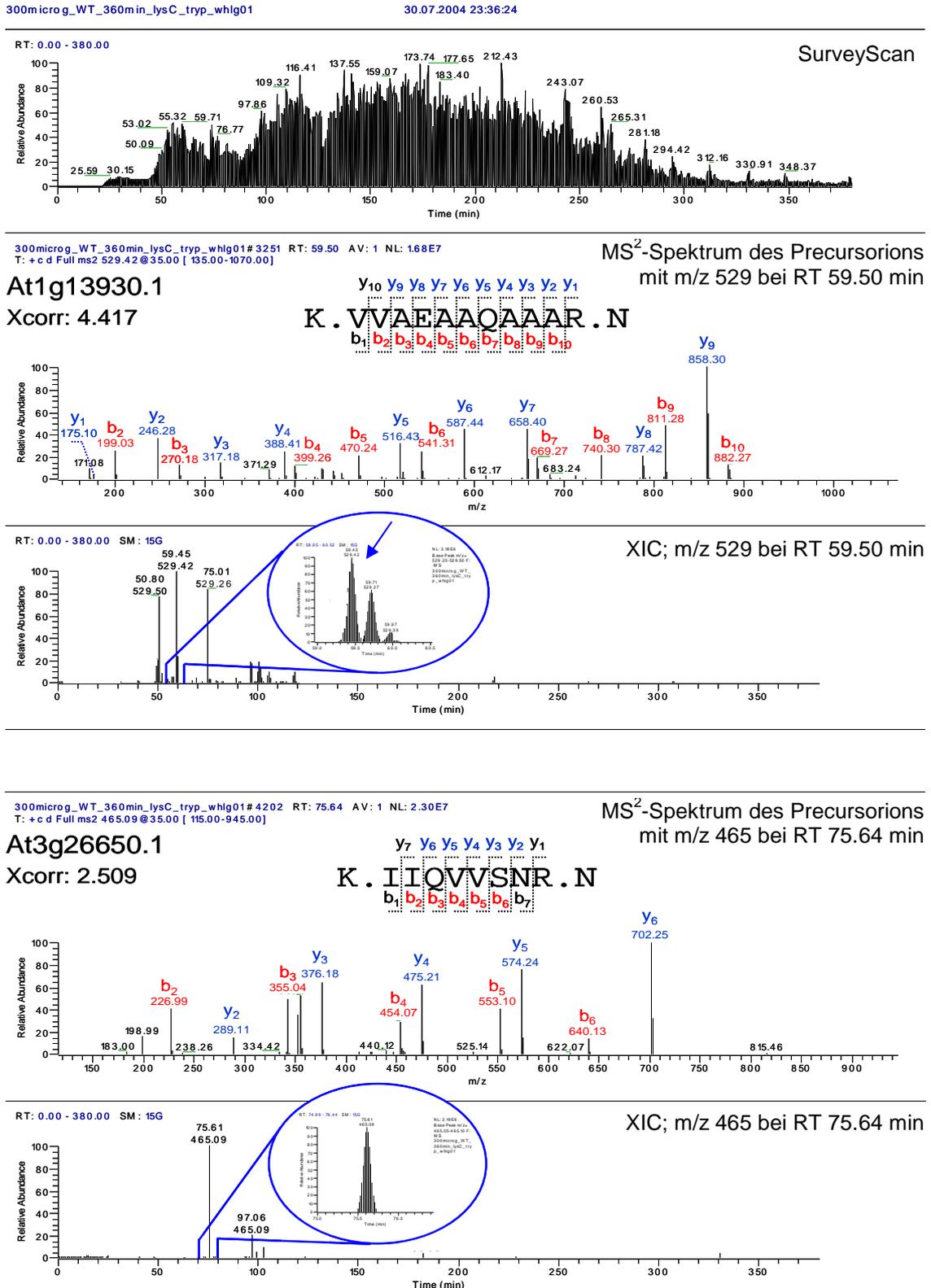
[98,196]. Die hier vorgestellte Prozedur der integrativen Extraktion von Metaboliten und Proteinen ist an die von Weckwerth *et al.* (2003) beschriebene Methode angelehnt [94], in welcher Urea als chaotropes Reagenz zur Denaturierung und Lösung von Proteinen verwendet wird. Arbeiten zur Untersuchung posttranslationaler Modifikationen an Hämoglobin haben jedoch gezeigt, dass unter Verwendung von Urea häufig signifikante Anteile carbamylierter Cysteinreste detektiert wurden, welche nicht-natürlichen Ursprungs waren. Zurückgeführt wurde dieser Effekt auf eine Reaktion der Sulfhydrylgruppe des Cysteins mit Isocyanat, welches hitzekatalysiert als Degradationsprodukt aus Urea entsteht. Isocyanat reagiert ebenso mit primären Aminogruppen von Proteinen oder Aminosäuren unter Bildung von S-Carboxymethylverbindungen, welche die N-terminale Sequenzierung blockieren und die eindeutige Sequenzierung bzw. Identifikation und vor allem die Quantifizierung mittels der Shotgun-Proteinanalyse erschweren [197]. Daher wurde dem verwendeten Proteinextraktionspuffer Methylamin, welches als Scavenger von Isocyanat fungiert, gemäss 8.3.2 zugesetzt.

Jede Probe wurde individuell dem Metabolite Profiling gemäss 8.4 unterzogen. Dabei konnten 80 strukturbekannte Metabolite in allen Messungen identifiziert und relativ quantifiziert werden. Proteine wurden aus denselben Proben phenolisch mit Ureapuffer extrahiert. Die relativ geringen Proteinausbeuten von ca. 1 mg/g FW erforderten vor der Shotgun-Proteinanalyse die Vereinigung von jeweils 3 Proteinpellets. Alle Messungen wurden in Triplika durchgeführt wobei 40 Proteine über die Peakintegration der korrespondierenden Peptide quantifiziert werden konnten.

Die erhaltenen Daten beider biologischer Stoffwechselebenen wurden im einzelnen sowie integrativ analysiert. In den folgenden Kapiteln erfolgt eine detaillierte Gegenüberstellung der durch verschiedene angewandte Verfahren des data minings aus den erhaltenen Daten extrahierten Informationen über die mit der circadianen Rhythmik verbundenen Änderungen im pflanzlichen Stoffwechsel.

### **3.2 Protein Profiling des circadianen Rhythmus von *A. thaliana***

Zur Analyse des Proteoms wurden die durch tryptischen Verdau erhaltenen Peptide zunächst einer datenabhängigen Shotgun-Analyse unterzogen. Die Auftrennung der hochkomplexen Peptidproben erfolgte dabei mittels nanoHPLC in einer 4-stündigen Chromatographie an einer 50cm langen monolithischen C18 reversed phase Kapillarsäule. Die im MS-SurveyScan detektierten drei abundantesten Peptide wurden anschliessend softwaregesteuert in einem MS<sup>2</sup>-Experiment fragmentiert. Über einen Datenbankabgleich der erhaltenen Peptidspektren gelang die Identifizierung der korrespondierenden Proteine.



**Abbildung 3.2.1 : Darstellung der Strategie der Proteinquantifizierung über Peakintegration an zwei Beispielen.** Nach der Identifizierung der Proteine At1g13930.1 und At3g26650.1 anhand der MS<sup>2</sup>-Spektra der tryptischen Peptide durch einen Datenbankabgleich wurden die Ionenpeaks der Vorläuferpeptide aus den MS-SurveyScans extrahiert. Die anschließende Peakintegration der jeweiligen Ionenpeaks erlaubte die Quantifizierung der korrespondierenden Proteine. Detektierte y- und b-Ionen sind farbig gekennzeichnet. *Abkürzungen:* RT, Retentionszeit; XIC, Extracted Ion Chromatogram; Xcorr, Cross Correlation coefficient, Einbuchstabencode der Aminosäuresequenz gemäss Abkürzungsverzeichnis

Nach Extraktion der Ionenspuren der Vorläuferpeptide bekannter Masse und Retentionszeit aus den MS-Läufen konnten über Peakintegration quantitative Daten erhoben werden. Abbildung 3.2.1 verdeutlicht diese Strategie an zwei Beispielen. Dabei wurde das Protein At1g13930.1, ein bisher in *Arabidopsis* funktionell uncharakterisiertes Protein, anhand eines 9 Aminosäuren langen Peptides identifiziert. Durch die Serie der detektierten  $\gamma$ - und  $\beta$ -Ionen eines weiteren, 5 Aminosäuren langen Peptides gelang nach Extraktion der zugehörigen Vorläuferionenspur die relative Quantifizierung der chloroplastidiären Glyceraldehyd 3-phosphat Dehydrogenase (GAP-DH) At3g26659.1 (siehe auch Kapitel 8.5.2.3.1, Material und Methoden).

Mit dieser Methode konnten insgesamt 40 Proteine aus hochkomplexen Peptidproben relativ quantifiziert werden. Etwa die Hälfte dieser Proteine ist laut Annotation in den Chloroplasten lokalisiert. Unter diesen 22 plastidiären Proteinen konnten neben 9 Proteinen des photosynthetischen Apparates 5 Enzyme identifiziert werden, welche am Calvin Zyklus beteiligt sind (Tabelle 3.2.1). Einige dieser Proteine konnten jedoch in keiner von der PGM Mutante stammenden Probe detektiert werden, diese sind mit \* gekennzeichnet. Eine weitere Schwierigkeit bei der Analyse der vorliegenden Daten ergab sich aus der geringen Anzahl gemessener biologischer Replikate. Daher wurden zur Auswertung lediglich die Median-Werte der quantifizierten Proteine ohne besondere Berücksichtigung von Variationskoeffizienten oder Standardfehlern herangezogen.

**Tabelle 3.2.1 : AGI-Code und funktionelle Beschreibung der via Peakintegration in Blattgewebe von *A. thaliana* Col-0 und PGM quantifizierten Proteine.**  
Ausführliche Annotationen sind dem Anhang A-2 zu entnehmen.

AGI-Code	Funktionelle Beschreibung
At1g06680.1	photosystem II oxygen-evolving complex 23, chloroplast
At1g13440.1	glyceraldehyde 3-phosphate dehydrogenase, cytosolic
At1g13930.1	expressed protein, similar to drought-induced sunflower protein
At1g19100.1	ATP-binding region
At1g51965.1	pentatricopeptide (PPR) repeat-containing protein
At1g67090.1	ribulose biphosphate carboxylase small chain 1A, chloroplast
At2g21170.1	triosephosphate isomerase, chloroplast
At2g28190.1*	superoxide dismutase [Cu-Zn], chloroplast
At2g30790.1	photosystem II oxygen-evolving complex 23, chloroplast
At2g35370.1	glycine cleavage system H protein 1, mitochondrial
At2g37220.1	29 kDa ribonucleoprotein, chloroplast
At2g39730.1*	ribulose biphosphate carboxylase/oxygenase activase, chloroplast
At3g01500.1	carbonic anhydrase 1, chloroplast
At3g14210.1*	myrosinase-associated protein
At3g15360.1	thioredoxin M-type 4, chloroplast
At3g16890.1	ATP synthase beta chain

Die mit \* gekennzeichneten Proteine konnten in keiner der PGM Proben detektiert werden.

**Tabelle 3.2.1: AGI-Code und funktionelle Beschreibung der via Peakintegration in Blattgewebe von *A. thaliana* Col-0 und PGM quantifizierten Proteine.**  
(Fortsetzung) Ausführliche Annotationen sind dem Anhang A-2 zu entnehmen.

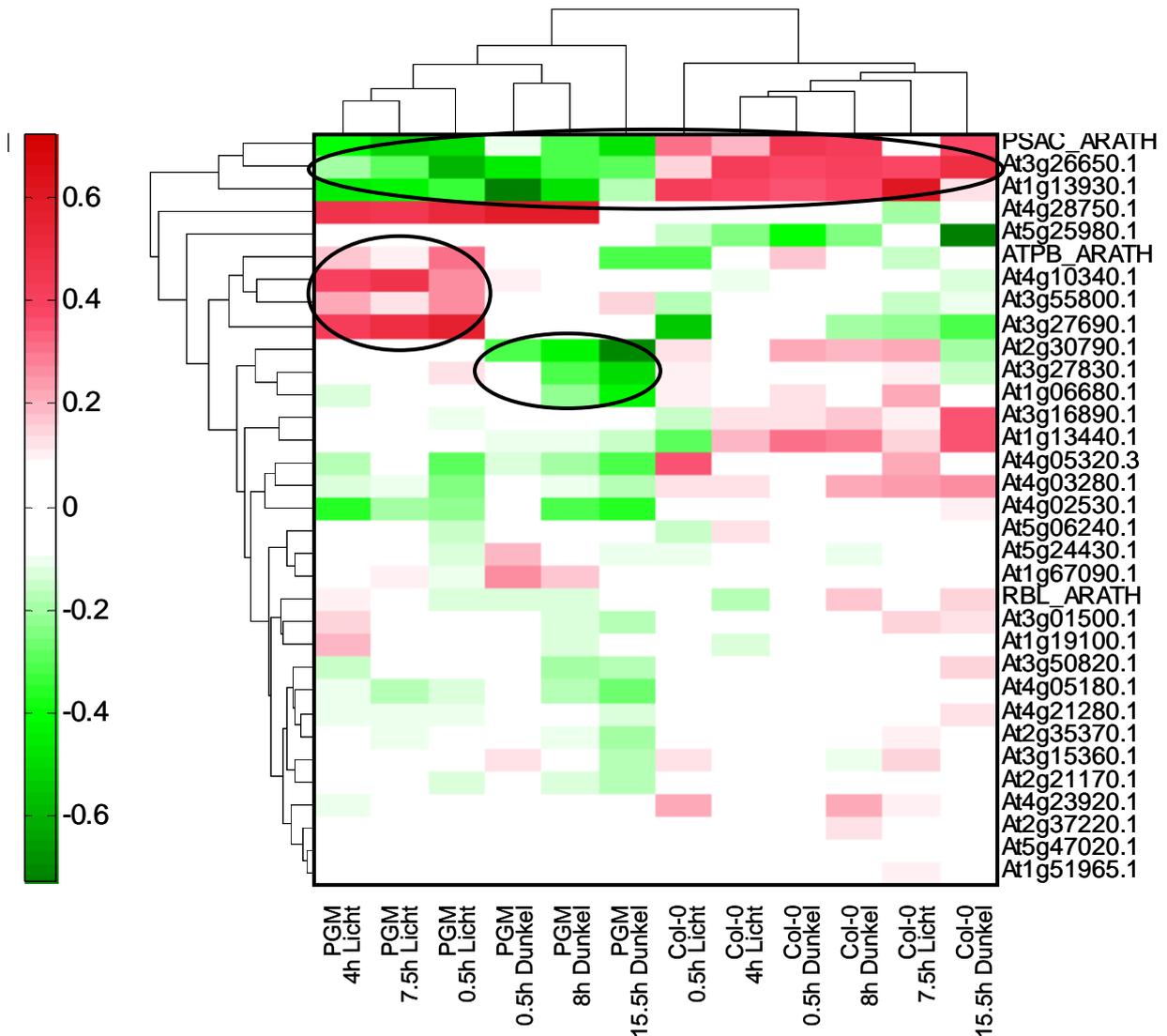
AGI-Code	Funktionelle Beschreibung
At3g26650.1	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast
At3g27690.1	chlorophyll A-B binding protein, chloroplast
At3g27830.1	50S ribosomal protein L12-1, chloroplast
At3g47070.1*	expressed protein
At3g50820.1	oxygen-evolving enhancer protein, chloroplast
At3g55800.1	glycine cleavage system H protein 1, mitochondrial
At4g02530.1	thylakoid lumen protein, chloroplast
At4g03280.1	cytochrome B6-F complex iron-sulfur subunit, chloroplast
At4g05180.1	oxygen-evolving enhancer protein 3, chloroplast
At4g05320.3	polyubiquitin (UBQ10) (SEN3) senescence-associated protein
At4g10340.1	chlorophyll A-B binding protein CP26, chloroplast
At4g21280.1	oxygen-evolving enhancer protein 3, chloroplast
At4g23920.1	UDP-glucose 4-epimerase, putative
At4g28660.1*	photosystem II reaction centre W protein, chloroplast
At4g28750.1	photosystem I reaction center subunit IV, chloroplast
At5g06240.1	expressed protein
At5g18740.1*	expressed protein, predicted
At5g24430.1	calcium-dependent protein kinase, putative
At5g25980.1	glycosyl hydrolase family 1 protein
At5g47020.1	glycine-rich protein
At5g66570.1*	oxygen-evolving enhancer protein 1-1, chloroplast
ATPB_ARATH	ATP synthase beta chain
PSAC_ARATH	Photosystem I iron-sulfur center, chloroplast
RBL_ARATH	Ribulose biphosphate carboxylase large chain precursor, chloroplast

Die mit \* gekennzeichneten Proteine konnten in keiner der PGM Proben detektiert werden.

### 3.2.1 Chemometrische Analyse der Proteindaten

Um Unterschiede zwischen beiden Genotypen und zugleich die mit der diurnalen Rhythmik verbundenen Veränderungen im pflanzlichen Stoffwechsel aufzudecken, wurde zunächst ein Bi-Clustering der geeignet normalisierten Protein-Datenmatrix mit dem Ziel, die Daten durch Charakteristika von sog. Clustern zu ersetzen, vorgenommen. Dabei erfolgte ein hierarchisches Clustering sowohl der experimentellen Bedingungen, als auch der durch relative Quantifizierung bestimmten Proteinexpressionslevel. Als Abstands- oder Ähnlichkeitsmass diente der Pearson'sche Produktmoment-Korrelationskoeffizient.

Durch Clustering der experimentellen Bedingungen wurde deutlich, dass zwischen beiden analysierten Genotypen deutliche Unterschiede existierten und beide Genotypen über einen individuell ausgeprägten Proteinhaushalt verfügten (Abbildung 3.2.1.1).

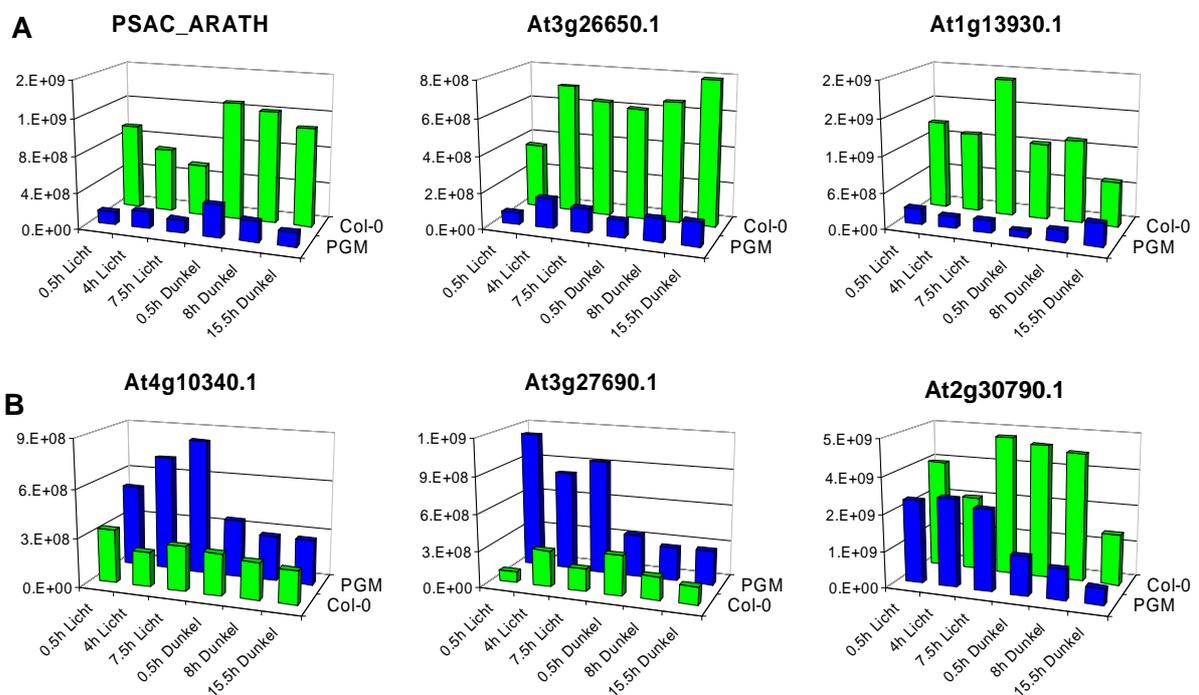


**Abbildung 3.2.1.1 : Bi-Clustering und farbcodierte Darstellung relativer Änderungen der Proteinkonzentrationen in *A. thaliana* Col-0 und PGM Pflanzen.** Dargestellt sind die aus drei biologischen Replikaten resultierenden Median-Werte jeder Stichprobe. Das Bi-Clustering erfolgte mit dem Pearson'schen Produktmoment-Korrelationskoeffizienten als Abstandsmass, die Farbcodierung erfolgte nach  $\log_{10}$ -Transformation der auf den Wildtyp Col-0 normierten Expressionen, wobei erhöhte Expressionslevel rot, verminderte Expressionslevel grün dargestellt sind. Diese Darstellung offenbart, dass anhand einiger in der PGM Mutante konstitutiv reprimierter Proteine (PSAC\_ARATH, Photosystem I; At3g26650.1, GAP-DH A, chloroplastidiär; At1g13930.1, unbekanntes Protein mit Homologie zu einem bei Trockenheit in Sonnenblumen exprimierten Protein) eine Unterscheidung zwischen den analysierten Genotypen Col-0 und PGM möglich ist. Daneben konnten anhand der farbcodierten Darstellung Kandidaten identifiziert werden, welche gemäss ihrer Proteinprofile in Abhängigkeit von der Photoperiode reguliert wurden (z.B. At4g10340.1 und At3g27690.1, Chlorophyll-bindende Proteine; At2g30790.1, Photosystem II).

So konnten neben Proteinen des photosynthetischen Apparates (PSAC\_ARATH, Photosystem I) auch am Calvin Zyklus beteiligte Enzyme (At3g26650.1, chloroplastidiäre Glycerinaldehyd 3-phosphat Dehydrogenase A, GAP-DH A) identifiziert werden, welche in der PGM Mutante deutlich schwächer exprimiert wurden. Daneben fiel das Protein

At1g13930.1, dessen funktionelle Charakterisierung bislang nicht erfolgte, durch eine vergleichsweise hohe Abundanz im Wildtyp auf. Dieses Protein weist eine schwache Homologie zu einem durch Trockenstress in der Sonnenblume *Helianthus annuus* induzierten Genprodukt unbekannter Funktion auf (Abbildung 3.2.1.2-A).

Des Weiteren fiel auf, dass die cytosolische GAP-DH A (At1g13440.1) eine grosse Distanz zur plastidiären Isoform (At3g26650.1), deren Expression in der Mutante auf etwa 1/6 vermindert war, aufwies. Hier war die relative Konzentration in der Mutante im Mittel auf die Hälfte der des Wildtyps vermindert, wobei wiederum keine Response auf die Photoperiode beobachtet wurde (Daten nicht gezeigt).



**Abbildung 3.2.1.2 : Expressionsprofile ausgewählter Proteine in *A. thaliana*.** Während zur Unterscheidung der Genotypen Col-0 und PGM neben Vertretern des photosynthetischen Apparates (PSAC\_ARATH, Photosystem I) und des Calvin Cyclus (At3g26650.2, chloroplastidiäre GAP-DH A) auch funktionell bisher nicht eindeutig charakterisierte Proteine (At1g13930.1, exprimiertes Protein mit Homologie zu einem unter Trockenstress in Sonnenblumen induziertem Protein) beitragen (**Abb. 3.2.1.2-A**), konnte die diurnale Rhythmik durch unterschiedliche Expressionslevel vorrangig von Proteinen des photosynthetischen Apparates (At4g10340.1 und At3g27690.1, Chlorophyll-bindende Proteine; At2g30790.1, Photosystem II) abgebildet werden (**Abb. 3.2.1.2-B**).

Aus dem Dendrogramm der experimentellen Bedingungen wird weiter unmittelbar ersichtlich, dass die Proteinexpressionen der PGM Mutante die diurnale Rhythmik sehr schön widerspiegeln. Hier wurden die die Photoperiode reflektierenden Proben zu einem

separaten Cluster zusammengefasst. Als Charakteristika dieser Klasseneinteilung fielen dabei vorrangig Proteine des photosynthetischen Apparates (At4g10340.1, At3g27690.1, At2g30790.1) durch verstärkte Expressionen im Licht auf (Abbildung 3.2.1.2-B). Der Wildtyp zeigte hingegen ein homogeneres Proteinexpressionsmuster, was in einer weitaus geringeren Distanz zwischen den Stichproben aller 6 Erntezeitpunkte resultierte. Hier ergab sich für keines der identifizierten Proteine ein deutlicher Hinweis auf eine lichtabhängige Regulation der Expression, was jedoch in der geringen Anzahl quantifizierter Proteine begründet sein kann.

### **3.3 Metabolite Profiling des circadianen Rhythmus von *Arabidopsis thaliana***

Neben der im vorangegangenen Kapitel beschriebenen Analyse der Proteinexpressionen wurde auch die komplementäre Stoffwechselebene der Metabolite mittels GC-TOF-MS untersucht.

Die hohen Konzentrationen der löslichen Carbohydrate Glucose, Fructose und Saccharose erforderte aufgrund des limitierten dynamischen Bereiches des TOF-Detektors für die Metabolitanalyse der während der Photoperiode geernteten Pflanzen des Genotyps PGM zusätzlich die Analyse von Verdünnungsserien. Die erhaltene Metabolit-Datenmatrix wies eine im Vergleich zu den Proteindaten weitaus höhere Komplexität auf. Es wurden insgesamt 80 strukturbekannte Metabolite in den zwei Genotypen Col-0 und PGM zu 6 verschiedenen Zeitpunkten innerhalb eines Tages quantifiziert. Daraus ergäben sich für die Auswertung dieses Datensatzes mindestens  $80 \times 2 \times (6-1) = 800$  sinnvolle Vergleiche. Daher musste zunächst eine geeignete Methode gefunden werden, mit deren Hilfe die Datenmenge unter Erhalt der biologisch relevanten Informationen reduziert werden konnte.

#### **3.3.1 Chemometrische Analyse der Metabolitdaten**

Da hier im Gegensatz zu der im vorangegangenen Kapitel beschriebenen Proteinanalyse zu jedem Zeitpunkt 10 biologische Replikate eines jeden Genotyps analysiert werden konnten, war es möglich, eine Varianzanalyse durchzuführen. Diese Methode der Datenanalyse beantwortet die Frage, ob Mittelwertsunterschiede zwischen Gruppen von Stichproben nachgewiesen werden können, oder ob die Nullhypothese, dass alle Beobachtungen der gleichen Verteilung entsprechen, mit den Daten verträglich ist. Als Testverfahren kam der Kruskal-Wallis-Test, welcher ein Verfahren der einfaktoriellen Varianzanalyse (one-way Analysis of Variance, ANOVA) auf Gleichheit der Lage-

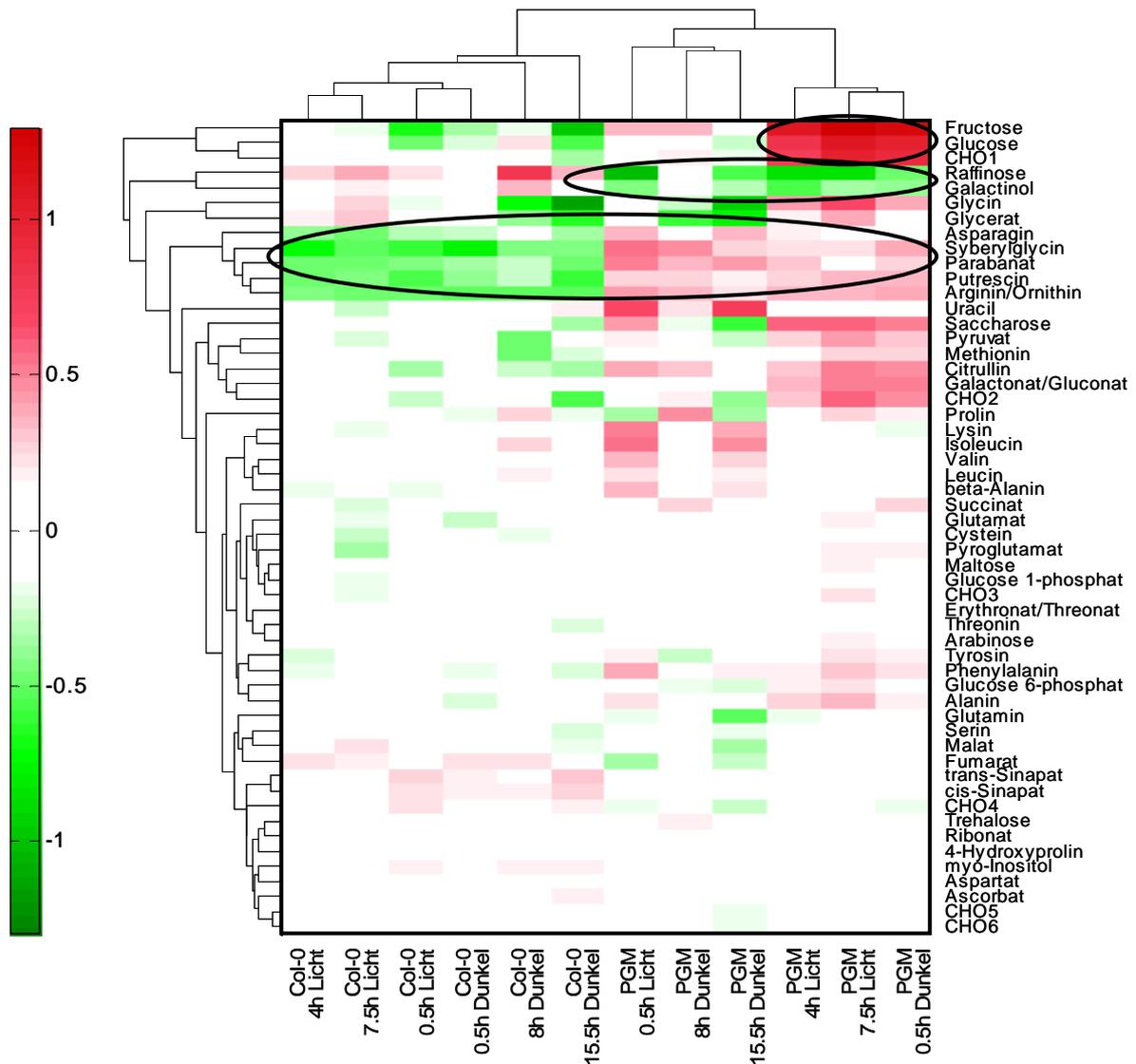
parameter darstellt, zum Einsatz. Dieser verteilungsfreie, nicht-parametrische Test erfordert im Gegensatz zu parametrischen Tests keine Normalverteilung der zu untersuchenden Daten und ist deutlich robuster gegen Ausreisser, welche einen erheblichen Einfluss auf den Mittelwert der betreffenden Stichprobe haben.

Um die innerhalb eines Tagesganges auftretenden metabolischen Veränderungen der Pflanze aufzudecken, wurde zunächst getestet, für welche Metabolite quantitative Unterschiede nachweisbar waren. Mit dieser Analyse konnte der vorliegende Datensatz auf 54 Metabolite reduziert werden, deren Konzentrationen sich in Abhängigkeit vom Erntezeitpunkt mit einer Irrtumswahrscheinlichkeit von  $p \leq 0.05$  signifikant unterschieden. Zur Visualisierung wurde der reduzierte Datensatz einem Bi-Clustering unterzogen, wobei die aus den 10 biologischen Replikaten einer jeden Gruppe resultierenden Mittelwerte der einzelnen Metabolite dargestellt wurden. Die resultierende Abbildung 3.3.1.1 macht deutlich, dass auch auf Metabolitebene die PGM Mutante gegenüber dem Wildtyp ein nahezu systemisch verändertes Profil in den untersuchten Blattproben aufwies.

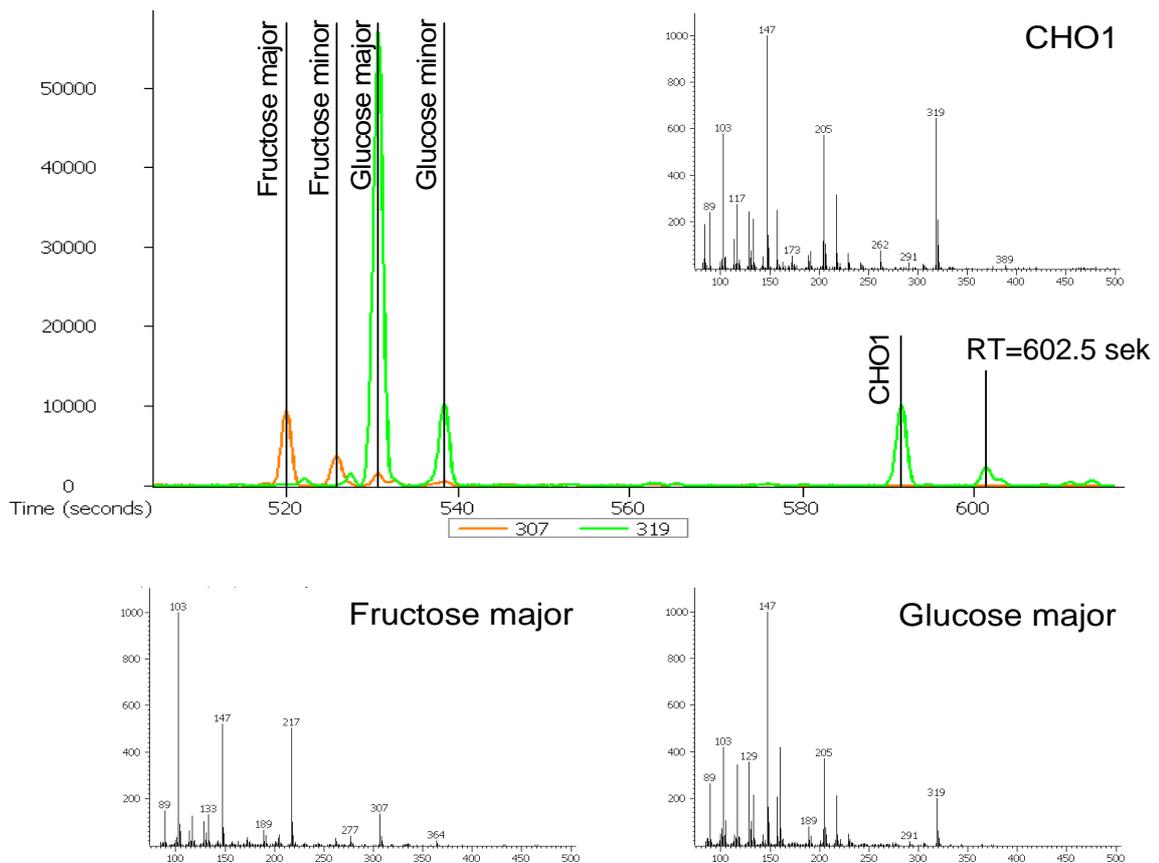
Aus der Hierarchie der experimentellen Bedingungen ist zudem der Einfluss der Photoperiode insbesondere auf den Stoffwechsel der PGM Mutante erkennbar, welcher von der starken Akkumulation der löslichen Carbohydrate Glucose und Fructose dominiert wird. Ein ähnlicher Effekt wurde für eine bisher nicht eindeutig identifizierte Verbindung CHO1 beobachtet, deren Massenspektrum eine grosse Ähnlichkeit mit den für Aldohexosen typischen Fragmenten aufwies. Auch liess das Retentionsverhalten dieses Analyten auf ein Carbohydrat schliessen (Abbildung 3.3.1.2)

Eine ebenfalls starke Response auf die Photoperiode zeigten Intermediate des Photorespiratorischen Cyclus. Diese blieben jedoch von den bis zu 40-fach erhöhten Konzentrationen der Carbohydrate Glucose, Fructose und Saccharose weitgehend unbeeinflusst. Serin, Glycin und Glycerat wiesen in beiden Genotypen ähnliche Abundanzen auf. Des Weiteren liess die farbcodierte Darstellung eine deutliche Herabregulation von Intermediaten des Raffinosepathways in der PGM Mutante erkennen. Im Gegensatz zu den akkumulierten Hexosen ergab sich für die Trisaccharide Galactinol und Raffinose zudem kein Hinweis auf eine lichtabhängige Regulation dieses Stoffwechselweges (Abbildung 3.3.1.3).

Die Varianzanalyse erbrachte weiter, dass auch die Sekundärmetabolite Benzoat und Salicylat sowie das am Auxin-Metabolismus beteiligte Indol 3-acetonitril, ähnlich dem bekannten Metaboliten Citramalat, dessen Funktion in Pflanzen bisher gänzlich unbekannt ist, unabhängig von der Photoperiode erhöhte Konzentrationen in der PGM Mutante aufwiesen (Abbildung 3.3.1.4).

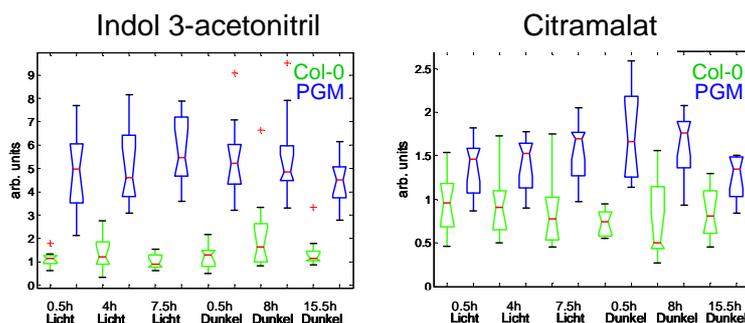
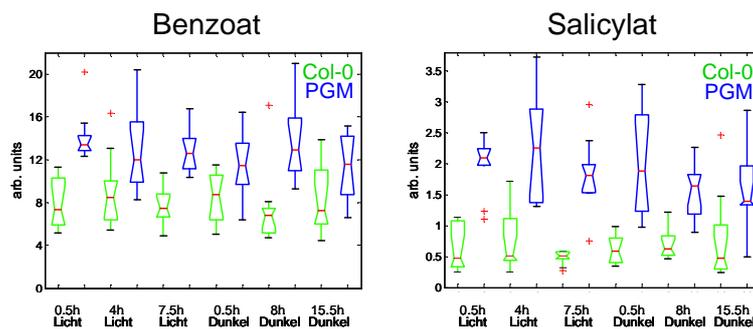
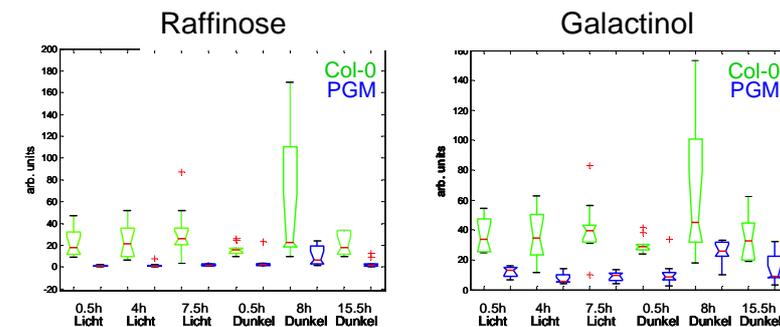
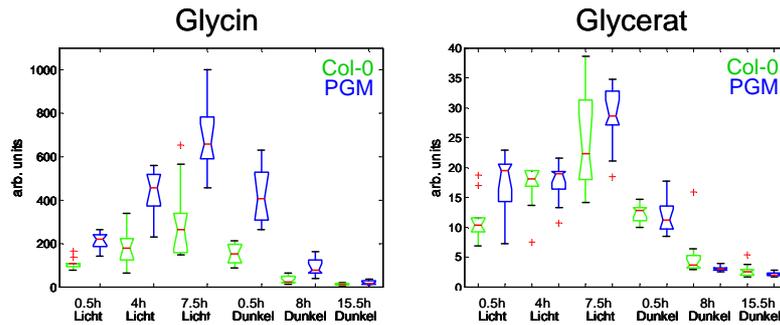
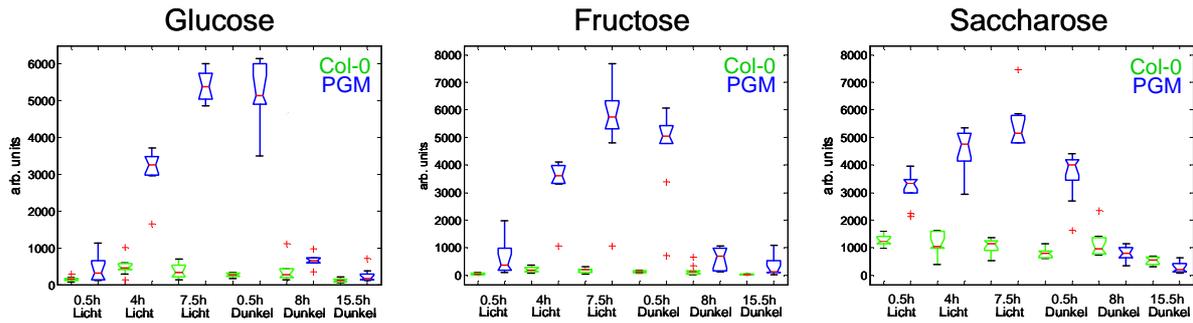


**Abbildung 3.3.1.1: Unterschiede der circadianen Metabolitprofile von *A. thaliana* Col-0, und PGM.** Dargestellt sind die aus dem Bi-Clustering der experimentellen Bedingungen sowie ausgewählter, relativ quantifizierter Metabolite resultierenden zeitabhängigen metabolischen Veränderungen in Blattproben von *A. thaliana* Col-0 Wildtyp sowie PGM Mutante. Das Clustering erfolgte mit dem Pearson'schen Produktmoment-Korrelationskoeffizienten als Abstandsmass, die Farbcodierung erfolgte nach  $\log_{10}$ -Transformation der auf den Median normierten Metabolitdaten, wobei erhöhte Konzentrationen rot, verminderte Konzentrationen grün dargestellt sind. Anhand dieser Darstellung konnten metabolische Marker identifiziert werden, welche den physiologischen Zustand oder auch den genetischen Hintergrund des analysierten Materials anzeigen. Glucose und Fructose formen zusammen mit einem bislang nicht eindeutig identifizierten Carbohydrat CHO1 einen unabhängigen Cluster, welcher die drastische Akkumulation löslicher Carbohydrate während der Lichtperiode in der PGM Mutante reflektiert. Des Weiteren offenbart die farbcodierte Darstellung, dass Intermediate des Raffinosepathways in der PGM in konstitutiv geringeren Konzentrationen vorlagen, während Blattkonzentrationen von Polyaminen sowie Intermediaten des Harnstoffcyclus in der Mutante erhöht waren.



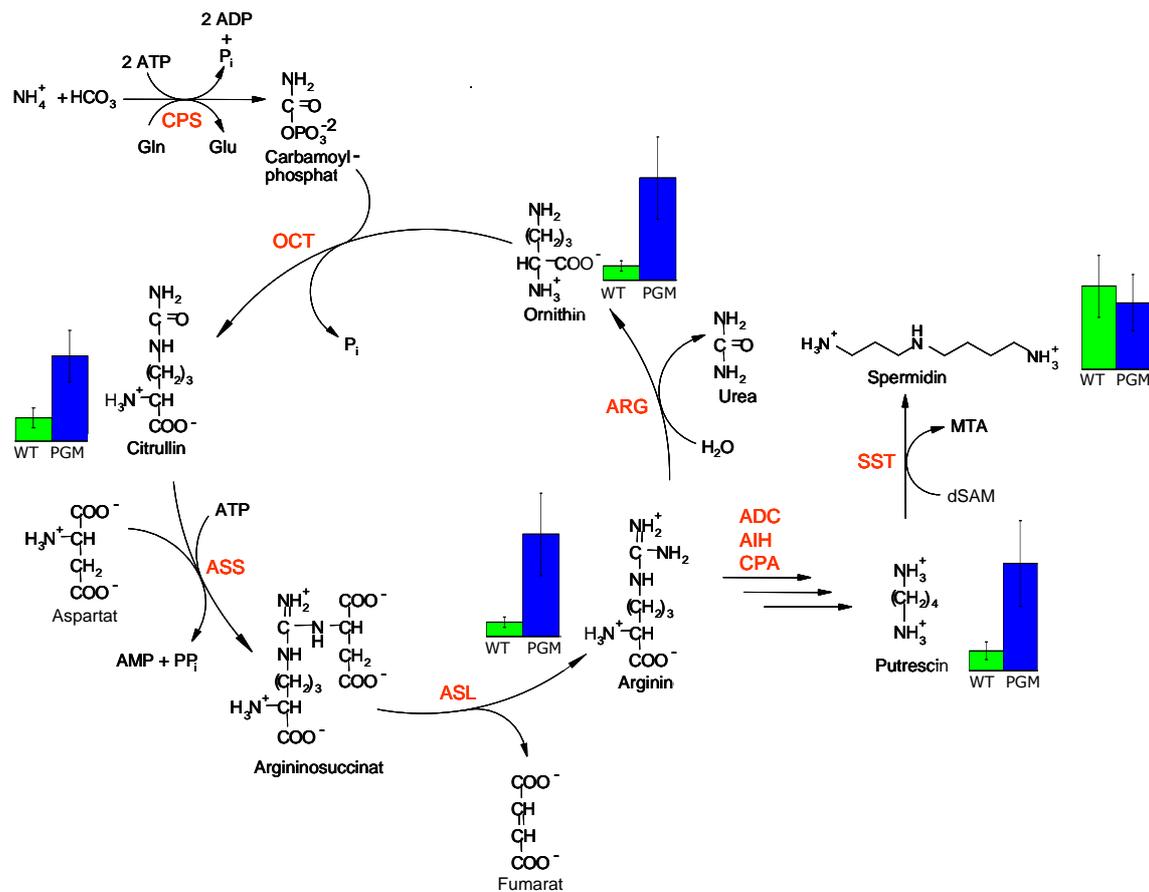
**Abbildung 3.3.1.2: Fragmentspektrum und Retentionsverhalten der Verbindung CHO1.** Retentionsverhalten und Massenspektrum der Verbindung CHO1 deuten auf eine Carbohydrat-Verbindung hin. Die Präsenz der Fragmente mit  $m/z$  205 und 319 lassen weiter auf eine Aldohexose schließen während die für Keto-hexosen charakteristischen Fragmente  $m/z$  217 und 307 nicht detektiert wurden. Die Bezeichnungen major und minor kennzeichnen die aus der Oximierung resultierenden Konfigurationsisomere der offenkettig fixierten Konformationen der Carbohydrate. Bei dem bei 602.5 sek eluierten Peak könnte es sich um ein eben solches Isomer der postulierten Aldohexose CHO1 handeln.

Insgesamt erbrachte die Analyse, dass die Mehrheit der identifizierten Metabolite in der PGM Mutante deutlich erhöhte Konzentrationen aufwies. So erbrachte die Analyse auch für Polyamine und Intermediate des Harnstoffzyklus über den gesamten diurnalen Verlauf im Mittel 5- bis 7-fach erhöhte Konzentrationen in der PGM. Für den peripheren Metaboliten Spermidin konnte dieser Effekt jedoch nicht beobachtet werden. Hier konnte zwischen dem Wildtyp und der Mutante kein signifikanter Unterschied in der Konzentration nachgewiesen werden (Abbildung 3.3.1.5).



**Abbildung 3.3.1.3: Einfluss der Photoperiode auf den Primärstoffwechsel von *A. thaliana*.** Dargestellt sind die aus 10 biologischen Replikaten resultierenden relativen Änderungen ausgewählter Metabolite. Während die Carbohydrate Glucose, Fructose und Saccharose in den PGM Pflanzen, hervorgerufen durch das Unvermögen der Synthese von Stärke, während der Lichtperiode stark akkumulierten, weisen die Intermediate des Photorespiratorischen Cyclus in den untersuchten Genotypen ähnliche Konzentrationen auf. Die Konzentrationsprofile von Raffinose und Galactinol hingegen deuten auf eine Herabregulierung des Raffinosepathways in der PGM Mutante hin.

**Abbildung 3.3.1.4: Einfluss der Photoperiode auf den Sekundärstoffwechsel von *A. thaliana*.** Dargestellt sind die Konzentrationsprofile der Sekundärmetabolite Benzoat und Salicylat, welche ähnlich Indol 3-acetonitril, einem Intermediat des pflanzlichen Hormonmetabolismus, unabhängig von der Photoperiode deutlich erhöhte Konzentrationen in der PGM Mutante erkennen lassen. Auch der bisher funktionell nicht charakterisierte Metabolit Citramalat wies in der Mutante gegenüber dem Wildtyp systemisch erhöhte Konzentrationen auf.



**Abbildung 3.3.1.5: Harnstoffzyklus in *A. thaliana*** (aus [178]). Die ermittelten Konzentrationen einzelner Intermediate, dargestellt als Medianwerte aller Zeitpunkte, zeigen eine Hochregulation dieses Stoffwechselweges in der PGM Mutante an. Abkürzungen: CPS, Carbamylphosphat Synthetase, kleine Untereinheit (AT3g27740), Carbamylsynthetase (At1g29900); OCT, Ornithin Carbamyltransferase (At1g75330); ASS, Argininosuccinat Synthase, (At4g24830); ASL, Argininosuccinat Lyase (At5g10920); ARG, Arginase, (At4g08900, At4g08870); ADC, Arginin Decarboxylase (At4g34710, At2g16500); AIH, Agmatin Iminohydrolase (At5g08170); CPA, *N*-Carbamylputrescine Amidohydrolase (At2g27450); SST, Spermidin Synthase (At5g53120, At1g23820, At1g70310); MTA, 5-Methylthioadenosine; dSAM, S-Adenosylmethionin, decarboxyliert.

Im Rahmen dieser Analyse konnten somit Stoffwechselwege und Intermediate des Primär-, sowie des Sekundärstoffwechsels identifiziert werden, welche in der PGM Mutante einer systemisch veränderten Regulation unterliegen oder als Folge einer solch veränderten Regulation signifikante Unterschiede in ihren Konzentrationen aufwiesen. Daneben gelang es, Metabolite mit Spezifität zu bestimmten physiologischen Stati der einzelnen Genotypen zu determinieren.

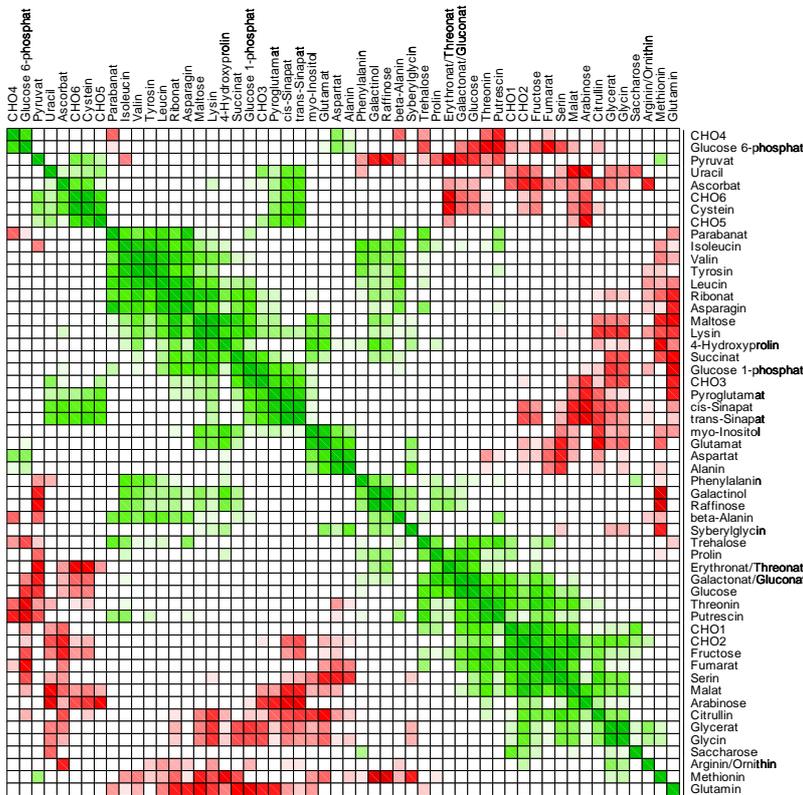
### 3.4 Der circadiane Rhythmus von *A. thaliana* im Kontext metabolischer Netzwerke

Zur Darstellung der innerhalb des circadianen Rhythmus auftretenden linearen Beziehungen einzelner Stoffwechselzwischenprodukte wurde eine Netzwerkanalyse basierend auf den Metabolit-Datensätzen der untersuchten Genotypen durchgeführt. Die linearen Beziehungen zwischen Metaboliten, welche zwischen den Erntezeitpunkten gemäss der ANOVA signifikante Unterschiede in ihren Konzentrationen aufwiesen, wurden in Abbildung 3.4.1 visualisiert. Da es sich bei der hier vorliegenden Analyse um eine Zeitreihen-Analyse handelt, können die beobachteten Korrelationen als Indizien für eine Co-Regulation der entsprechenden Stoffwechselprodukte bewertet werden. Unterschiede in den Netzwerken deuten folglich auf Unterschiede im regulatorischen Enzymnetzwerk beider Genotypen hin.

Tatsächlich offenbart die farbcodierte Darstellung des Korrelationsnetzwerkes der plastidiären PGM Mutante einen Cluster negativ korrelierter Aminosäuren und Carbohydrate, welcher im korrespondierenden Wildtyp nicht beobachtet wurde.

Die eingehende Analyse der Korrelationsnetzwerke erbrachte weiter, dass die in der PGM Mutante während der Photoperiode stark akkumulierten Carbohydrate gut korrelierten. Abbildung 3.4.2 verdeutlicht diese Beobachtung. Während Saccharose mit Fructose oder Glucose im Wildtyp nur schwach bis mässig korrelieren, konnte ein starker Zusammenhang zwischen diesen Variablen in der PGM Mutante nachgewiesen werden. So korrelierten Fructose und Saccharose stark in den bei Licht ( $r_{\text{Fructose-Saccharose}} = 0.915$ , mit  $p=2.26 \times 10^{-11}$ ) und bei Dunkelheit ( $r_{\text{Fructose-Saccharose}} = 0.98$ , mit  $p=1.84 \times 10^{-16}$ ) geernteten Proben der Mutante (zum Vergleich: Col-0 Licht  $r_{\text{Fructose-Saccharose}} = 0.43$ ; Col-0 Dunkel  $r_{\text{Fructose-Saccharose}} = 0.48$ ). Auch Glucose und Fructose korrelierten in der PGM Mutante deutlich stärker als im korrespondierenden Wildtyp Col-0 (PGM Licht  $r_{\text{Glucose-Fructose}} = 0.97$  mit  $p=8.48 \times 10^{-19}$ ; Col-0 Licht  $r_{\text{Glucose-Fructose}} = 0.88$  mit  $p=9.25 \times 10^{-11}$ ; PGM Dunkel  $r_{\text{Glucose-Fructose}} = 0.91$  mit  $p=1.41 \times 10^{-9}$ ; Col-0 Dunkel  $r_{\text{Glucose-Fructose}} = 0.43$  mit  $p=0.018$ ). Bei der Bewertung der linearen Zusammenhänge zwischen diesen Metaboliten fiel weiter auf, dass die Korrelationen zwischen diesen Metaboliten in der PGM Mutante unterschiedliche Muster aufwiesen. Ähnlich dem in Kapitel 2.6 diskutierten Phänomen der Umkehrung signifikanter Korrelation wurde aus der Scatterplot-Darstellung deutlich, dass der Anstieg der resultierenden Regressionsgeraden je nach Zeitpunkt der Probennahme unterschiedlich war.

Col-0



PGM

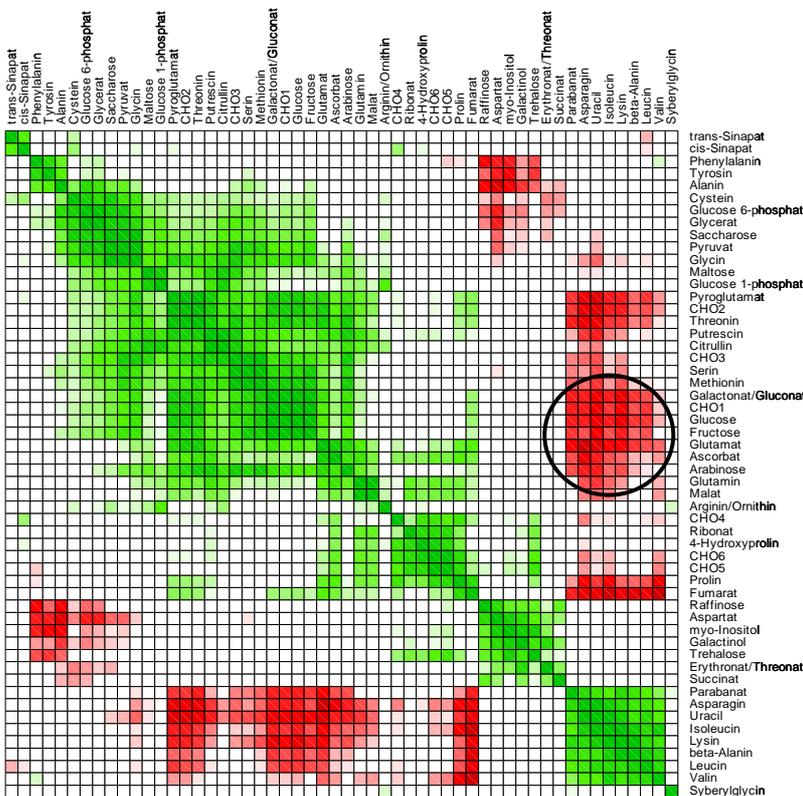
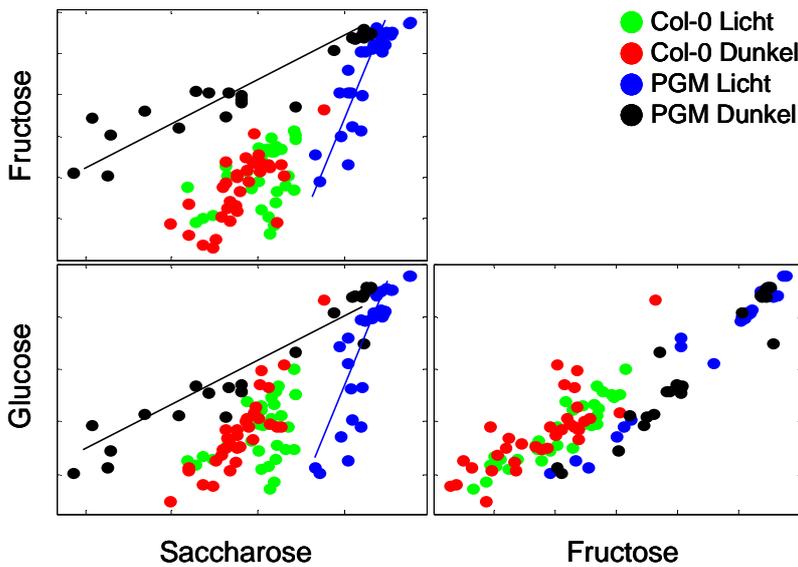


Abbildung 3.4.1: Farbcodierte Darstellung des Korrelationsnetzwerkes der diurnalen Veränderungen im Blattgewebe von *A. thaliana*. Dargestellt sind Korrelationen zwischen Metaboliten, welche nach vorangegangener Varianzanalyse mit  $p \leq 0.05$  signifikante Mittelwertsunterschiede in ihren Konzentrationen zu 6 verschiedenen Zeitpunkten innerhalb eines Tages aufwiesen (siehe dazu Kapitel 3.3). Paarweise Metabolitkorrelationen wurden nach dem Pearson'schen Produktmoment-Korrelationskoeffizienten berechnet und auf ihre Signifikanz getestet ( $p \leq 0.001$ ). Zwischen den die Tagesrhythmik reflektierenden Netzwerken beider Genotypen Col-0 und PGM sind deutliche Unterschiede erkennbar. Unter anderem weist das Profil der PGM Mutante im Gegensatz zu dem des korrespondierenden Wildtyps Col-0 einen invers korrelierten Cluster von Aminosäuren und Carbohydraten auf.



**Abbildung 3.4.2 :** Vergleich der Korrelationen zwischen Glucose, Fructose und Saccharose unter besonderer Berücksichtigung der Photoperiode. Während in Col-0 schwache bis mässige Korrelationen zwischen diesen Carbohydraten beobachtet wurden, korrelierten diese in der PGM Mutante stark. Die bei Licht und bei Dunkelheit geernteten Proben der PGM weisen zudem deutliche Unterschiede im Verlauf der Wertepaare auf.

Somit konnte die in Kapitel 2 beschriebene Beobachtung der variablen Ausprägung von Korrelationen in Abhängigkeit vom zugrunde liegenden enzymatischen oder Reaktionsnetzwerk um ein zusätzliches Beispiel experimentell nachweisbarer, differentieller Korrelationen erweitert werden.

### 3.5 Diskussion

Sollen unterschiedliche Stoffwechselebenen simultan analysiert werden, so erfordert dies eine sorgfältige Präparation der Analyten aus der komplexen zellulären Matrix des Untersuchungsmaterials. Im vorangegangenen Abschnitt wurden die Ergebnisse der Analyse der innerhalb des diurnalen Rhythmus auftretenden Veränderungen des Proteoms sowie des Metaboloms zweier *Arabidopsis* Genotypen vorgestellt, wobei die Extraktion der Komponenten beider Stoffwechselebenen aus identischem Blattmaterial erfolgte. Die Analyse kleiner Moleküle erfolgte gaschromatographisch, die Proteinanalytik flüssigchromatographisch. Im Gegensatz zur Gaschromatographie, bei der die Ionisation der Analyten vor deren massenspektrometrischen Analyse mittels Elektronenstoss erfolgt, treten bei der Elektrospray-Ionisation häufig Ionensuppressionen auf, welche die Quantifizierung einzelner Analyten zum Teil drastisch behindern. Diese Limitierung der LC-MS kann nur durch eine zureichende chromatographische Auftrennung umgangen werden. In der vorgestellten Methode erfolgte dies durch eine 4-stündige Chromatographie an einer 50 cm langen monolithischen reversed phase C-18 Kapillarsäule mit einem Durchmesser von 100 µm. Im Gegensatz zu gepackten Säulen,

welche mit porösen Sorbentien gefüllt sind, ist das Säulenbett monolithischer Kapillarsäulen von starrer Struktur und weist keine Partikelzwischenräume auf. Die hochpermeablen, durchgängigen Poren setzen der mobilen Phase einen deutlich geringeren Widerstand entgegen und erlauben somit wesentlich höhere Flussraten, wodurch die Analysezeiten wesentlich verkürzt werden können [198]. Somit kam eine für die Auftrennung der komplexen Peptidproben bestens geeignete Trennsäule zum Einsatz. Weitere Möglichkeiten zur Verbesserung der Auftrennung hochkomplexer Peptidproben stellen vorangehende Fraktionierungen, z.B. durch Grössenausschlusschromatographie, dar. Dieses erfordert jedoch eine deutlich grössere Ausgangsmenge an Gesamtprotein und verlängert zudem die Gesamtanalysezeiten erheblich, wodurch auch die Gefahr artifizierender Proteinmodifikationen und -degradationen erheblich steigt. Daher wurde auf zusätzliche Fraktionierungen der Proteinextrakte verzichtet und die Eignung einer relativ kurzen Methode im Hinblick auf die quantitative Protein-Profilanalyse, neben der Metabolit-Profilanalyse, zur Charakterisierung verschiedener *Arabidopsis* Genotypen getestet.

### **Die zentralen Energiewandlungsprozesse der Pflanze unterliegen definierten Regelmechanismen**

Durch ihre vornehmlich sessile Lebensweisen sind Pflanzen zum Teil beträchtlichen Schwankungen der ihre natürliche Umgebung charakterisierenden Faktoren, wie z.B. Nährstoffangebot oder die Verfügbarkeit von Wasser, ausgesetzt. Der wohl stärksten Fluktuation unterliegt, bedingt durch den periodischen Wechsel von Tag und Nacht, die Verfügbarkeit von Lichtenergie. In den Chloroplasten wird die Energie des Lichtes zunächst in ATP umgewandelt und in der anschliessenden Reaktionskette des Calvin Cyclus durch die Assimilation von CO<sub>2</sub> chemisch in Form von Carbohydraten gebunden. Ein Teil der Photoassimilate wird überwiegend in Form von Dihydroxyacetonphosphat im Austausch gegen anorganisches Phosphat aus den Chloroplasten exportiert. Die Fixierung des atmosphärischen Kohlenstoffes wird durch die Ribulose 1,5-bisphosphat Carboxylase/Oxygenase (RubisCO) mit einer vergleichsweise geringen Aktivität (drei Umsätze pro Sekunde [199]) katalysiert. Die RubisCO liegt in hohen Konzentrationen gelöst im Stroma der Chloroplasten vor, nur ein geringer Teil ist an die Thyllakoidmembranen gebunden, macht etwa 50-70% der löslichen Proteine im Blatt aus (siehe auch Abbildung 4.2.1.1) und ist mit einer Grösse von ca. 500 kDa zugleich eines der grössten in Pflanzen vorkommenden Proteine.

Neben der Möglichkeit der CO<sub>2</sub>-Fixierung im Calvin-Cyclus katalysiert die RubisCO auch die konkurrierende Reaktion der O<sub>2</sub>-Fixierung. Ob Kohlendioxid oder Sauerstoff fixiert wird, hängt hauptsächlich vom Partialdruck beider Gase ab. Im Gegensatz zur Carboxylase-

funktion der RubisCO, welche die Bildung von Triosephosphaten katalysiert, führt die Oxygenasefunktion zur Bildung von Phosphoglycerat sowie Phosphoglycolat. Das entstandene Phosphoglycolat kann in den Chloroplasten nicht metabolisiert werden und wird daher in dephosphorylierter Form zunächst in die Peroxisomen transportiert. Nach Oxidation und Transaminierung entsteht Glycin, welches in die Mitochondrien weitertransportiert und dort unter Desaminierung und Decarboxylierung zu Serin umgesetzt wird. Auf diesem als Photorespiration bezeichnetem Weg werden ca. 50% des assimilierten Kohlenstoffes energetisch nutzlos umgesetzt. Wenngleich der physiologische Nutzen dieser Reaktionskette bis dato unklar ist, gilt die Photorespiration als Hauptquelle für die Bildung der Aminosäuren Glycin und Serin. Somit stellen die über den Calvin Cyclus in Form von Carbohydraten fixierten Kohlenstoffverbindungen das Substrat der Respiration dar.

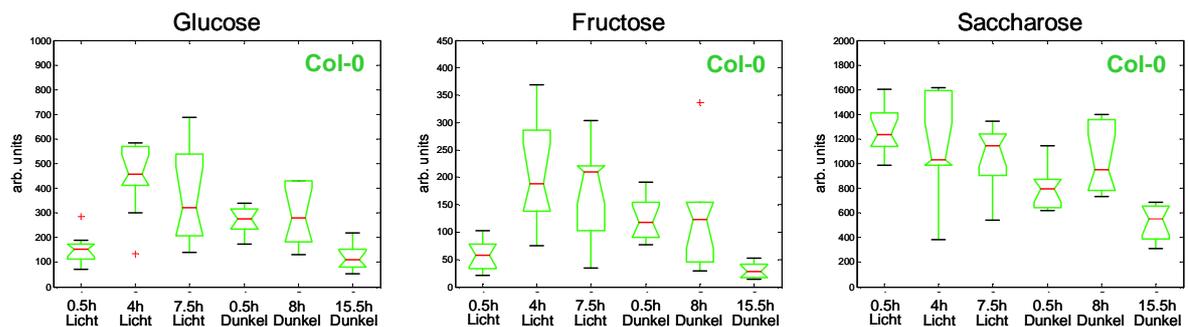
Neben der Photorespiration unterliegen Carbohydrate weiteren oxidativen Abbauprozessen zur Gewinnung von Energie- und Reduktionsäquivalenten. Der Abbau der Carbohydrate erfolgt in mehreren Schritten. Während der Glycolyse wird Glucose im Cytoplasma zu Pyruvat abgebaut. Nach dem Transport in die Mitochondrien wird Pyruvat oxidativ zu Acetyl-CoA decarboxyliert und diese Verbindung im Rahmen des Citratcyclus weiteren oxidativen Decarboxylierungsreaktionen unterworfen. Insbesondere der oxidative Pentosephosphatweg, ein Nebenweg der Glycolyse, liefert den Hauptteil der für die reduktive Biosynthese von Fettsäuren und Aminosäuren erforderlichen Reduktionsäquivalente in Form von NADPH und stellt die für die Synthese der Nucleotide und Nucleinsäuren wichtige Verbindung Ribose 5-phosphat bereit. Der Hauptgewinn an chemischer Energie wird an der inneren Mitochondrienmembran durch Endoxidation in der Atmungskette erzielt.

Daneben beeinflussen Carbohydrate auch die Fixierung des Stickstoffs und damit die Aminosäure- und Proteinbiosynthese. Nach ihrer Umwandlung in Carbonsäuren dienen die Assimilate der Photosynthese als Kohlenstoffgerüst für die Assimilation von Stickstoff, welcher in Form von  $\text{NO}_3^-$  oder  $\text{NH}_3$  über die Wurzeln der Pflanze aufgenommen wird. Dieser Prozess erfordert einen beträchtlichen Teil der durch die Respiration freigesetzten Energie. Alternativ kann  $\text{NH}_3$  auch im Zuge der Photorespiration aus der Aminosäure Glycin freigesetzt und anschliessend re-assimiliert werden. Dieser Prozess übersteigt mitunter die *de novo*-Assimilation von  $\text{NH}_3$  um das zehnfache [200].

### **Umweltbedingte Schwankungen der Lichtenergie werden durch die Dissimilation von Carbohydraten kompensiert**

Der Katabolismus der Carbohydrate stellt während der Nacht die einzige Quelle für die Aufrechterhaltung der für das Wachstum nötigen Energie- und Reduktionsäquivalente dar. Daher ist der pflanzliche Stoffwechsel normalerweise so organisiert, dass ein grosser Teil der Photoassimilate in Form von transitorischer Stärke im Blatt verbleibt und in der anschliessenden Dunkelperiode re-mobilisiert wird, um die Respiration sowie die Saccharosesynthese und den –export aufrecht zu erhalten. Somit kommt der Stärke eine zentrale Bedeutung für den pflanzlichen Stoffwechsel und das Wachstum zu [201,202].

In Übereinstimmung mit diesem Sachverhalt offenbaren die im Rahmen dieser Arbeit aufgenommenen Metabolitprofile im Wildtyp auch nach 8h Dunkelheit der Lichtperiode vergleichbare Konzentrationen der Carbohydrate Glucose, Fructose und Saccharose (Abbildung 3.5.1, Auszug aus Abbildung 3.3.1.3). Erst zum Ende der Nacht ist eine deutliche Abnahme der Konzentrationen zu beobachten, wobei die reduzierenden Zucker



**Abbildung 3.5.1 : Einfluss der Photoperiode auf den Carbohydratstatus von *A. thaliana* Col-0.** Dargestellt sind die diurnalen Konzentrationsprofile der löslichen Carbohydrate Glucose, Fructose und Saccharose (Auszug aus Abb. 3.3.1.3). Eine deutliche Konzentrationsabnahme dieser Verbindungen ist aufgrund der Mobilisierung transitorischer Stärke erst zum Ende der 16-stündigen Dunkelperiode erkennbar.

auch nach 30 min Licht noch ähnlich verminderte Konzentrationen aufweisen. Dieses Ergebnis wurde in einer unabhängigen Studie von Gibon *et al.* (2004) zur Untersuchung des Einflusses der Tageslänge auf den Carbon- und Stärkemetabolismus in *A. thaliana* durch enzymatische Assays bestätigt [203]. Gleichzeitig beobachteten die Autoren, dass die transitorische Stärke im Blatt zum Ende der mit 16h relativ langen Nacht vollständig re-mobilisiert wurde.

Die Intermediate des photorespiratorischen Cyclus zeigten erwartungsgemäss hohe Konzentrationen während der Lichtperiode (Abbildung 3.3.1.3).

### **Die optimale Modulation des Stoffwechsels der Pflanze erfolgt durch die photosynthetische Lichtwirkung**

Der Einfluss des Lichtes auf die Expression zahlreicher Gene in *Arabidopsis* ist mehrfach untersucht worden. Schaffer *et al.* (2001) untersuchten mit Hilfe der mRNA Microarray Technologie die Expression von ca. 11.000 Genen, was etwa 50% aller in *Arabidopsis* vorkommenden Gene entspricht, und fanden, dass 11% dieser Gene diurnale Veränderungen in ihrer Expression aufwiesen [204]. Dieses Ergebnis überraschte. Insbesondere im Hinblick auf die oben diskutierte Tatsache, dass das Licht die einzige Energiequelle der Pflanzen darstellt, erscheint die Anzahl der circadian oszillierenden Gene unerwartet niedrig. Auch Bläsing *et al.* (2005) analysierten die Genexpression in *Arabidopsis*, verwendeten jedoch den das vollständige *Arabidopsis* Genom umfassenden 22K ATH1 Microarray (Fa. Affymetrix GeneChips, Santa Clara, CA, USA). Diese Analysen erbrachten, dass 30-50% aller in *Arabidopsis* vorkommenden Gene diurnalen Veränderungen ihrer Transkriptlevel im Blatt von *Arabidopsis* unterliegen [205]. Beiden Experimenten lag eine 12h Licht / 12h Dunkelperiode zugrunde, jedoch analysierten Schaffer *et al.* die Expression zu 4 unterschiedlichen Zeitpunkten, während Bläsing *et al.* die Genexpression zu 6 verschiedenen Zeitpunkten in drei biologischen Replikaten analysierten. Die letztgenannten Autoren identifizierten Sets von oszillierenden Genen und ordneten diese funktionellen Kategorien zu, wobei die stärksten diurnalen Veränderungen auf Transkriptebeine für Gene des Saccharose- und Stärkestoffwechsels, der Assimilation von Nährstoffen und der Redox-Regulation beobachtet wurden. Für einen Grossteil dieser Gene wurde das Maximum der Transkriptmenge im Blatt zum Ende der Nacht beobachtet.

Die Analyse der im Rahmen dieser Arbeit erstellten Proteinprofile erbrachte, dass die Proteinmengen insbesondere im Wildtyp nur schwache diurnale Oszillationen aufwiesen (Abbildungen 3.2.1.1 und 3.2.1.2). Dieses kann zum einen in der relativ geringen Anzahl der hier quantitativ erfassten Proteine begründet sein, zum anderen erlauben die Daten aufgrund der zugrunde liegenden Analyseverfahren keine Aussagen über die Aktivität der einzelnen Proteine.

Darüber hinaus muss eine diurnale Veränderung in der Transkription nicht notwendigerweise einen Einfluss auf biologische oder biochemische Prozesse bedingen. Von diurnalen Veränderungen der Transkriptmengen bestimmter Gene lassen sich ausschliesslich Informationen über die Regulation der Genexpression ableiten. Wenngleich der Einfluss der Transkription auf die Proteinmenge einzelner Gene beschrieben ist, welche z.B. mit der Regulation der inneren Uhr [206-208], der Response auf die Photoperiode [209] oder auch der Assimilation von Nährstoffen [210,211] assoziiert sind, zeigte sich für ebenso viele Gene unterschiedlicher metabolischer

Funktionen keine Co-Expression von Transkript und Protein. So beobachteten beispielsweise Smith *et al.* (2004) oder Lu *et al.* (2005) auf Transkriptebene starke diurnale Veränderungen von Enzymen, welche im Stärkeabbau involviert sind, während die Menge der korrespondierenden Proteine nahezu unverändert blieb [212,213]. Auch Ideker *et al.* (2001) beobachteten in ihren Untersuchungen zur Regulation des Galactosestoffwechsels in Hefe, dass die Korrelationen zwischen der mRNA- und der Proteinexpression nur minimal ist [22].

cDNA Microarray Analysen erbrachten weiter, dass die Transkriptmengen der RubisCO sowie der RubisCO Aktivase infolge der circadianen Regulation einem raschen Turnover unterliegen [214]. Betrachtet man jedoch die Halbwertszeit des Proteinturnover der RubisCO, welche fünf bis sechs Tage beträgt [215], wird schnell deutlich, dass eine Veränderung der Proteinmenge für die Pflanze in vielen Fällen kontraproduktiv ist und neben einer Verschwendung von Energie, insbesondere bei langsamer Synthese und Turnover, die Anpassung eher an bereits vergangene, als an momentane physiologische Situationen bedingt. Daher erscheint es plausibel, dass die Regulation des komplexen metabolischen Netzwerkes zu einem erheblichen Teil über posttranslationale Modifikationen erfolgt.

### ***Die genetischen Eigenschaften von Organismen spiegeln sich in spezifischen Proteinexpressionsmustern und Metabolitprofilen wider***

Ein anderer Aspekt der Bedeutung der Proteinmenge ergibt sich aus der vergleichenden Analyse von Wildtyp und Mutante. Ein Eingriff in das Genom eines Organismus und damit verbunden eine Veränderung des zugrunde liegenden genetischen Netzwerkes lässt auch eine globale Veränderung des Proteinnetzwerkes erwarten. Viele Untersuchungen haben gezeigt, dass die Regulation zahlreicher Gene, welche u.a. für Komponenten des photosynthetischen Apparates, der Pathogenabwehr, des abiotischen Stresses sowie des Stärke-, Lipid- oder auch Nitratmetabolismus kodieren, über Carbohydrate erfolgt [216-220]. Daneben beeinflusst die Verfügbarkeit von photosynthetischem Kohlenstoff die Proteinsynthese und –remobilisierung [217,218,221,222].

In Übereinstimmung mit diesen Studien erbrachte ein Vergleich der metabolischen Profile, dass sich die plastidiäre Mutation in der stärkelosen, zuckerakkumulierenden PGM Mutante in globalen Veränderungen auf der Ebene der Proteine sowie der Metabolite auswirkte (Abbildungen 3.2.1.1 und 3.3.1.1). Als problematisch im Hinblick auf die Aussagefähigkeit der Ergebnisse erwies sich jedoch die Tatsache, dass die Proteinprofile der untersuchten Genotypen zu den verschiedenen Zeitpunkten innerhalb des diurnalen Rhythmus in nur drei Replikate mit einem Variationskoeffizienten von ca. 50% ermittelt werden konnten. Dennoch konnten mit Hilfe des hierarchischen Clusterings auch Markerproteine identifiziert werden, welche unabhängig vom Erntezeitpunkt in beiden

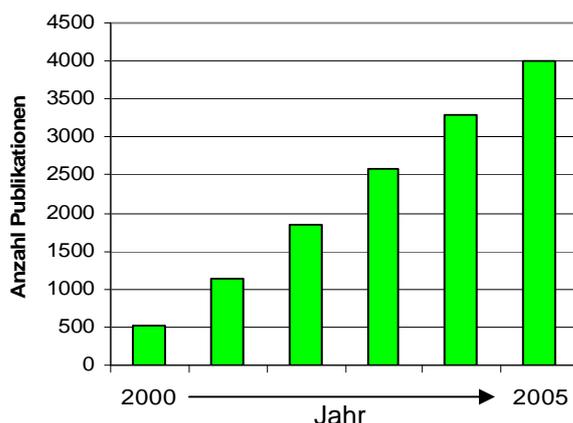
Genotypen unterschiedlich stark exprimiert wurden oder aber in Abhängigkeit von der Photoperiode diurnale Unterschiede in ihrer Expression aufwiesen.

### **Die optimale Modulation der photosynthetischen Energiegewinnung erfolgt durch Carbohydrate**

Aufgrund ihres Unvermögens der Konversion von Glucose 1-phosphat zu Glucose 6-phosphat mangelt es der PGM Mutante an der Fähigkeit, Fluktuationen in der Carbohydratbilanz durch die Re-Mobilisierung von transitorischer Stärke auszugleichen [195]. Als Konsequenz dieses Unvermögens unterliegen die diurnalen Veränderungen der intrazellulären Konzentrationen löslicher Zucker drastischen Schwankungen. Während der Lichtperiode akkumulieren lösliche Carbohydrate im Blatt der PGM bis auf das 40-fache der Konzentrationen im korrespondierenden Wildtyp. Interessanterweise konnten für die Intermediate des TCA-Cyclus sowie für Aminosäuren keine Veränderung der Oszillation nachgewiesen werden (Abbildung 3.3.1.1). Dieses unerwartete Ergebnis indiziert eine verminderte Metabolisierung der photosynthetischen Assimilate in der PGM.

Lösliche Carbohydrate, vor allem Glucose, Fructose und Saccharose spielen eine zentrale Rolle in der Regulation des pflanzlichen Metabolismus und fungieren als Energieträger und Transportmetabolite, sind massgeblich am Aufbau der pflanzlichen Zellwand beteiligt und haben vor allem aufgrund ihrer Funktion als Signalmoleküle einen entscheidenden Einfluss auf die Regulation der Genexpression. Sowohl die Synthese und Speicherung als auch der Abbau der Carbohydrate stellen hochdynamische Prozesse dar, welche eng mit dem physiologischen Status der Pflanze, der Verfügbarkeit von Nährstoffen und den jeweils herrschenden Umweltbedingungen verknüpft sind.

Wenngleich die jährlich erscheinende Anzahl von Publikationen zur Untersuchung des Einflusses von Carbohydraten auf den pflanzlichen Metabolismus enorm ist (Abbildung 3.5.2), ist bis heute unklar, wie Pflanzen diese Verbindungen wahrnehmen und auf welchem Wege die Signaltransduktion erfolgt.



**Abbildung 3.5.2: Übersicht über die zum Einfluss von Carbohydraten auf den pflanzlichen Metabolismus erschienenen Publikationen der Jahre 2000 bis 2005.** Gezeigt ist das Ergebnis der Suchanfrage nach Artikeln mit den Schlüsselwörtern *PLANT METABOLISM* und *CARBOHYDRATES* der PubMed MEDLINE®-Datenbank der National Library of Medicine (NLM)

Viele der in Col-0 und PGM detektierten differentiell exprimierten Proteine sind funktionell mit dem photosynthetischen Apparat assoziiert (Tabelle 3.2.1). Der Einfluss von Carbohydraten auf die Expression der in der Photosynthese involvierten Gene ist mehrfach beschrieben worden. Steht photosynthetischer Kohlenstoff in hohen Konzentrationen zur Verfügung, wird die Expression von Genen, deren Produkte in verschiedenste Prozesse der Verwertung und Speicherung des Kohlenstoffs involviert sind, induziert, wobei photosynthetische Gene gleichzeitig reprimiert werden. Insbesondere die Feedback-Inhibierung der Photosynthese durch die Akkumulierung von Saccharose im Blatt der Pflanze ist literaturbekannt [219,223-225]. Ein Mangel an Carbohydraten führt dagegen zur Aktivierung der Expression photosynthetischer Gene [216,226].

Eine besondere Charakteristik der PGM Mutante stellt die Tatsache dar, dass diese bei der Anzucht unter kontinuierlichem Licht keinen Phänotyp aufweist [195]. Dieses deutet darauf hin, dass der Carbohydratmangel während der Dunkelperiode hauptsächlich für das verminderte Wachstum der PGM Mutante unter Kurztagbedingungen ist. Auch Bläsing *et al.* (2005) identifizierten in ihrer auf Affymetrix Arrays basierenden Analyse der Genexpression in *Arabidopsis* Sets von Genen, welche der diurnalen Regulation unterliegen, und führten dies vorrangig auf die intrazelluläre Konzentration von Carbohydraten zurück [205].

#### **Die Modulation des bifunktionellen Enzyms RubisCO kann über die Konzentration von Carbohydraten erfolgen**

Bezüglich der CO<sub>2</sub>-Aufnahme erbrachten vorangegangene Untersuchungen der stärkelosen *Arabidopsis thaliana* Mutante PGM einen nur geringfügigen Einfluss hoher Carbohydratkonzentrationen. Caspar *et al.* (1986) beschreiben eine nur leicht verminderten CO<sub>2</sub>-Aufnahme in der stärkelosen *Arabidopsis* Mutante während der Lichtperiode [195]. Auch Gibon *et al.* (2004) konnten keine drastische Verminderung der CO<sub>2</sub>-Aufnahme im Verlauf der Lichtperiode nachweisen und stellten eine im Vergleich zum Wildtyp nur leicht verminderte Photosyntheseleistung in der PGM fest [203]. Die Autoren konnten zudem keinen Einfluss der während der Lichtperiode akkumulierten Carbohydrate Saccharose, Glucose und Fructose auf die CO<sub>2</sub>-Aufnahme nachweisen. Die im Rahmen dieser Arbeit erstellten Proteinprofile wiesen jedoch für mehrere Proteine des photosynthetischen Apparates deutlich verminderte Expressionen in der PGM Mutante auf (Abbildung 3.2.1.2-A).

Die photosynthetische CO<sub>2</sub>-Fixierung wird durch das bifunktionelle Enzym RubisCO initiiert. Die RubisCO wird über zwei durch Splicevariationen in *Arabidopsis* vorkommende

Isoformen der RubisCO Aktivase reguliert wird. Durch die Dissoziation von Ribulose 1,5-bisphosphat (RuBP) wirken die RubisCO Aktivasen der Inhibierung der RubisCO-Aktivität entgegen. RuBP ist zum einen Substrat der RubisCO, kann jedoch gleichzeitig an die inaktivierte Form der RubisCO binden und somit deren Aktivierung durch die Carbamylierung eines Lysinrestes blockieren.

Mittels der hier vorgestellten Methode der Proteinquantifizierung über die Peakintegration korrespondierender Peptide konnte die RubisCO Aktivase (At2g39730.1), welche der grösseren der zwei Isoformen entspricht, in der PGM Mutante zu keinem Erntezeitpunkt detektiert werden (Tabelle 3.2.1).

Zhang *et al.* konnten nachweisen, dass diese 46 kDa grosse Isoform der RubisCO Aktivase massgeblich an der Regulation der RubisCO beteiligt ist [227]. Transgene *Arabidopsis* Pflanzen, welche diese Isoform nicht exprimierten, hatten die Kapazität für die Herabregulation der Carboxylaseaktivität der RubisCO bei Dunkelheit verloren. Wird die Carboxylaseaktivität der RubisCO nicht herabreguliert, kann der Fluss der Carboxylierung weiterhin durch den RubisCO-Inhibitor RuBP reguliert werden. Als regulierender Parameter der Aktivität der RubisCO Aktivase gilt das ADP/ATP-Verhältnis im Stroma des Chloroplasten, wobei ADP als Inhibitor der RubisCO Aktivase-Aktivität wirkt [228-230]. Unter steady-state Bedingungen variiert das ADP/ATP-Verhältnis jedoch kaum mit einer veränderlichen Lichtintensität [231,232], so dass die Regulation der RubisCO Aktivase nicht allein durch die Photoperiode erfolgen kann.

Brouquisse *et al.* (1991) beobachteten bei extremem Glucosemangel eine signifikant erhöhte ADP-Regeneration in Mais [233]. Da die während der Lichtperiode im Blatt der PGM Mutante akkumulierte Carbohydrate in der darauf folgenden Dunkelperiode rasch verbraucht werden, ist die zweite Hälfte der Dunkelperiode durch einen dramatischen Carbohydratmangel gekennzeichnet. Dieser Zustand könnte ein für die Aktivierung der RubisCO Aktivase ungünstiges ADP/ATP-Verhältnis bis in die beginnende Lichtperiode hinein bedingen. Inaktive oder auch nicht-assemblierte Proteine unterliegen in der pflanzlichen Zelle einer raschen Degradation, vorausgesetzt, die Raten der Proteinsynthese und des -turnover sind entsprechend hoch, so dass als mögliche Folge die Konzentration der RubisCO Aktivase in der PGM in einer im Vergleich zum Wildtyp deutlich verminderten Konzentration vorlag und somit nicht mehr detektiert werden konnte. Folglich würde jedoch die Carboxylaseaktivität der RubisCO nun nicht mehr wie im Wildtyp herabreguliert, wodurch die Rate der CO<sub>2</sub>-Fixierung ungeachtet der verminderten Expression von Proteinen des photosynthetischen Apparates nur geringfügig erniedrigt wird.

Da die RubisCO auch die mit der CO<sub>2</sub>-Fixierung konkurrierende Reaktion der Sauerstoff-fixierung katalysiert, postulierten Caspar *et al.*, dass das verminderte Wachstum der PGM

unter Kurztagbedingungen in einem respiratorischen Verlust von CO<sub>2</sub> begründet liegt [195]. Die Metabolitprofile des diurnalen Rhythmus von Col-0 und PGM wiesen jedoch ähnliche Konzentrationen der photorespiratorischen Intermediate Glycin, Glycerat und Serin auf (Abbildung 3.3.1.3). Dieses liegt in einer weiteren möglichen Modulation der RubisCO in der PGM während der Lichtperiode, welche durch hohe Konzentrationen an Carbohydraten gekennzeichnet ist, begründet.

Untersuchungen an salzgestressten Pflanzen erbrachten, dass erhöhte Level an Saccharose und verwandten Zuckeralkoholen, wie Inositol oder Mannitol, das Potenzial besitzen, die Oxygenaseaktivität der RubisCO zu inhibieren, während die Carboxygenaseaktivität von diesen gelösten Substanzen weitgehend unbeeinflusst bleibt [234].

Legt man diese Annahme zugrunde, ist der Stoffwechsel der PGM während der Lichtperiode durch eine Inhibierung der Oxygenaseaktivität der RubisCO gekennzeichnet, so dass die Photorespiration reprimiert wird.

### ***Hohe Konzentrationen von Carbohydraten haben einen möglichen inhibitorischen Einfluss auf die photosynthetische Elektronentransportkette und das Reduktionspotenzial der Chloroplasten***

Die Analyse der Proteindaten erbrachte weiter, dass das Vorläufermolekül des Apoproteins PSAC, welches als Eisen-Schwefel-Zentrum eine Untereinheit des Photosystems I darstellt, eine deutlich verminderte Proteinexpression in der zuckerakkumulierenden Mutante aufwies (Abbildung 3.2.1.2). Auch für den Cytochrom b6-f-Komplexes (At4g03280.1) erbrachte die Peakintegration eine um 50% verminderte Expression in der Mutante (Abbildung 3.2.1.1). Diese Ergebnisse indizieren einen verminderten photosynthetischen Elektronentransport und werden durch die Ergebnisse von Ott *et al.* (1999) der Untersuchung des photosynthetischen Elektronentransportkette der roten Lichtnelke *Silene dioica* gestützt, welche erbrachten, dass hohe Saccharosekonzentrationen den Elektronenfluss durch das Photosystem I inhibieren [235]. Auch Eichelmann *et al.* (1994) postulierten nach Untersuchungen einer stärkeleeren chloroplastidiären Mutante von *Nicotiana sylvestris* einen verminderten Elektronenfluss durch das Photosystem I [236].

Beide Enzyme, sowohl die RubisCO als auch die RubisCO Aktivase, weisen in ihrer Aminosäuresequenz viele Cysteinreste auf. Die bereits erwähnten Untersuchungen von Zhang *et al.* zur Regulation der RubisCO erbrachten weiter, dass auch eine Cystein → Alanin Substitution in der C-terminalen Extension der grösseren Isoform der RubisCO Aktivase in dem Verlust der Kapazität der Herabregulation der Carboxylaseaktivität der RubisCO bei Dunkelheit resultiert [227]. Somit muss die Regulation der RubisCO auch über den Redox-Status der RubisCO Aktivase erfolgen. In Folge eines verminderten

Reduktionspotenzials im Stroma der Chloroplasten, welches aus einem verminderten photosynthetischen Elektronentransport resultieren kann, könnte die RubisCO Aktivase während der Lichtperiode in der PGM nur bedingt durch die Reduktion des Cysteins aktiviert werden, was ebenfalls eine Modulation der RubisCO und somit der photosynthetischen Energiegewinnung zur Folge hätte. Zur Überprüfung dieser These sind jedoch die Bestimmungen des stomatalen Redoxpotenzials sowie der photosynthetischen Elektronentransportrate nötig. Darüber hinaus gäben vergleichende Messungen der Aktivität der RubisCO Aktivase im Wildtyp sowie der PGM weiteren Aufschluss über die Regulation der Photosynthese und den angeschlossenen Calvin Zyclus in der Mutante.

### **Die chloroplastidiäre Mutation der Phosphoglucomutase bedingt eine verminderte Proteinexpression der in den Calvin Zyclus involvierten plastidiären GAP-DH**

Das photosynthetische Elektronentransportsystem und der Calvin Zyclus werden parallel reguliert, um die ATP- und NADPH-Synthese mit der Rate der Metabolisierung dieser Komponenten zu koordinieren und eine Überreduktion zu verhindern [237]. Sheen konnte zeigen, dass Saccharose sowie Glucose die Transkription der Promotoren von 7 photosynthetischen Genen in Mais inhibieren [238].

Bei der Analyse der Proteinexpressionen fiel die am photosynthetischen Calvin Zyclus beteiligte chloroplastidiäre GAP-DH A (At3g26650.1) durch eine deutlich verminderte Expression in der Mutante auf (Abbildung 3.2.1.2-A). Das Genom von *A. thaliana* kerncodiert für 7 Isoformen der phosphorylierenden GAP-DH. Drei Gene kodieren für die zwei NADH- bzw. NADPH-abhängigen Untereinheiten GAPA und GAPB, welche zusammen mit der Phosphoribulokinase (PRK) und dem chloroplastidiären CP12 Peptid einen supramolekularen Komplex mit postulierter regulatorischer Funktion der CO<sub>2</sub>-Assimilation formen [239]. Beide Isoformen liegen in hohen Konzentrationen im Stroma der Chloroplasten vor. Daneben kodieren zwei Gene für die cytosolische, NAD-abhängige GAP-DH, welche in den glycolytischen Pathway sowie die Gluconeogenese involviert ist. Untersuchungen der Regulation der Gen-Expression lieferten bislang konträre Ergebnisse. Marri *et al.* (2005) beschrieben eine drastische Abnahme der Konzentration des *GapA* Transkriptes in Blättern adulter *Arabidopsis* Pflanzen bei kontinuierlicher Dunkelheit sowie nach Inkubation einzelner Blätter mit Saccharose [240]. Die Untersuchungen von Shih *et al.* Goodman (1988) und Yang *et al.* (1993) ergaben hingegen, dass nicht die Expression des *GapA* Gens, sondern vielmehr die des *GapC* Gens, welches für die cytosolische GAP-DH codiert, in *Arabidopsis* durch Saccharose induziert wird [241,242].

Die Proteinprofile beider Enzyme wiesen in der PGM Mutante verminderte Expressionen auf, wenngleich dieser Effekt für die chloroplastidiäre GAP-DH, deren Konzentration in der PGM Mutante etwa 1/7 der des Wildtyps betrug, deutlicher ausgeprägt war. Die Konzentration der cytosolischen GAP-DH war in der PGM Mutante etwa auf die Hälfte der

Konzentration im korrespondierenden Wildtyp vermindert. Ein Effekt der Photoperiode auf die Proteinexpressionen konnte für beide Enzyme nicht beobachtet werden. Diese Ergebnisse deuten darauf hin, dass hohe Carbohydratkonzentrationen und/oder ein verminderter Elektronentransport durch das PS I eine mögliche inhibitorische Wirkung auf die Proteinexpression der in den Calvin-Cyclus involvierten Enzyme haben. Die verminderte Proteinexpression der plastidiären GAP-DH in der PGM könnte allerdings auch auf dem Fehlen einzelner Assemblierungspartner des über Protein-Protein-Interaktionen organisierten supramolekularen GAP-DH/CP12/PRK Komplexes in der PGM und der damit verbundenen Initiation der Degradation beruhen. Des Weiteren kann ein pleiotroper Effekt der Mutation der chloroplastidiären GAP-DH als Ursache für die verminderte Proteinexpression der plastidiären GAP-DH in der PGM nicht ausgeschlossen werden.

### ***Verschiedene metabolische Marker zeigen einen erhöhten Stress-Status der PGM an***

Das hierarchische Clustering der relativen Änderungen der ermittelten Proteinkonzentrationen offenbarte, dass neben der plastidiären GAP-DH auch das in *Arabidopsis* bis dato uncharakterisierte Protein At1g13930.1 mit Homologie zu einem durch Trockenstress induzierten Protein der Sonnenblume in der PGM deutlich verminderte Konzentrationen aufwies (Abbildung 3.2.1.2-A).

Die Proteinexpression des für At1g13930.1 kodierenden Genabschnittes beschrieben Leonhardt *et al.* (2004) als in den Schliesszellen von *Arabidopsis* durch Abscisinsäure induzierbar [243]. Die Sesquiterpenverbindung Abscisinsäure erfüllt in der Pflanze die Funktion eines Stresshormons, welches als Reaktion auf verschiedene abiotische Stressfaktoren vermehrt gebildet wird und unter anderem das Wachstum hemmt. Während der Trockenheit triggert dieses in der Wurzel synthetisierte Phytohormon nach dem Transport ins Blatt unter anderem das Schliessen der Stomata, um die Pflanze vor Wasserverlust zu schützen und ist somit an der Osmoregulation beteiligt.

Wenngleich die Interpretation der verminderten Expression von At1g13930.1 in der PGM Mutante durch die Tatsache erschwert wird, dass mittels der Metabolitanalyse das in nur sehr geringen Konzentrationen im Blatt der Pflanze vorkommende Phytohormon Abscisinsäure nicht analysiert werden konnte, ergibt sich hieraus ein Hinweis auf einen veränderten Wasserhaushalt der PGM.

Neben Abscisinsäure ist auch der Sekundärmetabolit Salicylsäure in der Regulation der Stomataöffnung involviert. Als Response auf erhöhte Konzentrationen wurde mehrfach das Schliessen der Stomata beobachtet [244-246]. Die Metabolitanalysen über den

diurnalen Rhythmus erbrachten konstitutiv erhöhte Salicylsäurekonzentrationen in der PGM Mutante (Abbildung 3.3.1.5).

Im Zusammenhang mit der Akkumulation von Salicylsäure wurde mehrfach eine Verminderung des Wachstums beschrieben [247-249]. Mateo *et al.* (2006) stellten zudem einen Zusammenhang zwischen der Photosynthese und dem Salicylsäuregehalt in *Arabidopsis* her [250]. In Salicylsäure akkumulierenden Mutanten beobachteten die Autoren einen reduzierten photosynthetischen Elektronentransport am Photosystem II und führten das verminderte Wachstum der Mutanten auf einen verminderten Gasaustausch zurück.

Diese Ergebnisse indiziert, dass das komplexe Regelsystem der Öffnung der Stomata in der PGM über einen modulierten Mechanismus kontrolliert wird.

Limitierter Gasaustausch und eine beeinträchtigte Photosyntheseleistung stellen Faktoren dar, welche den oxidativen Stress fördern [251,252]. Mateo *et al.* (2006) beobachteten in den Salicylsäure akkumulierenden Mutanten weiter, dass diese durch einen Anstieg der  $H_2O_2$ -Konzentrationen gekennzeichnet sind [250]. Auch bei Carbohydratmangel, welcher sich während der Dunkelperiode in der PGM Mutante einstellt, entsteht vermehrt  $H_2O_2$  als Artefakt der peroxisomalen  $\beta$ -Oxidation der Fettsäuren [253,254]. Zugleich sind Fettsäuremobilisation und  $\beta$ -Oxidation unter hohen Carbohydratkonzentrationen, welche während der Lichtperiode in der PGM vorliegen, reprimiert [233] und es werden vermehrt Reduktions-äquivalente durch die Glycolyse und den oxidativen Pentosephosphat-Cyclus produziert [255,256], welche als Scavenger reaktiver Sauerstoffspezies (ROS) wirksam sind. Daneben ist Glucose ein Precursor der Synthese der als Antioxidantien wirksamen Carotinoide sowie der Ascorbinsäure [257,258]. Auch die Synthese der Aminosäuren Cystein, Glutamat und Glycin wird durch erhöhte Glucosekonzentrationen induziert, um die reduktive Detoxifikation von ROS über Glutathion zu begünstigen [259]. Gleichzeitig unterliegen reduzierende Zucker einer raschen Autoxidation durch Triplett-Sauerstoff. Als Folge dieser Autoxidationen treten vermehrt partiell reduzierte Sauerstoffspezies, wie z.B. Superoxid-Anionen ( $O_2^{\bullet-}$ ) oder Hydroxylradikale ( $HO^{\bullet}$ ) auf, welche die Zelle durch ihre hohe Reaktivität in vielerlei Hinsicht schädigen können. Somit nehmen Carbohydrate eine zentrale Rolle in der Balance der ROS ein.

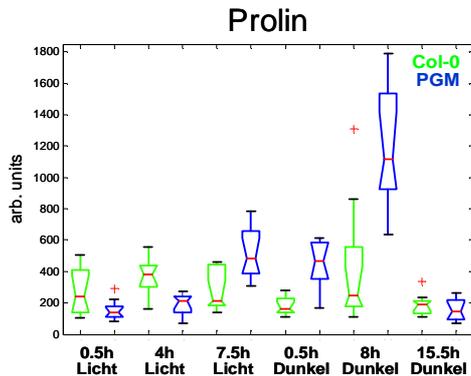
Die Metabolitprofile der Ascorbinsäure liessen keine erhöhten Konzentrationen dieses Akzeptors der Autoxidation in der Mutante erkennen, was jedoch in der raschen Oxidierbarkeit begründet sein kann. Wenngleich für die Aminosäuren Cystein, Glutamat und Glycin ebenfalls keine signifikant erhöhten Konzentrationen in der PGM nachgewiesen werden konnten, ist eine Erhöhung der Syntheserate dieser Bausteine nicht ausgeschlossen, da das Tripeptid Glutathion mittels den angewandten

Analysemethoden nicht detektiert werden konnte. Auch die limitierte Anzahl quantifizierter Proteine lässt leider keine näheren Aussagen über den oxidativen Stressstatus der Pflanze zu.

Als metabolische Stressmarker gelten weiter Polyamine. Für zahlreiche Intermediate des Harnstoffzyklus erbrachte die Metabolitanalyse signifikant erhöhte Konzentrationen, wobei der Einfluss der Photoperiode vernachlässigbar war (Abbildung 3.3.1.4). Dieses deutet auf eine ganztägige Stresssituation in der PGM hin und kann als Indiz für die vermehrte Entstehung von ROS durch die Autoxidation der akkumulierten Zucker Glucose und Fructose während der Lichtperiode, sowie auf die durch Carbohydramangel während der Dunkelperiode induzierte peroxisomalen  $\beta$ -Oxidation bzw. des dadurch vermehrt auftretenden  $H_2O_2$  in der PGM Mutante gewertet werden.

Auch die Akkumulationen der Chlorophyll-bindenden Proteine At4g10340.1 sowie At3g27690.1 (Abbildung 3.2.1.2-B) sind möglicherweise durch Stresssituationen in der PGM Mutante bedingt. Beide Proteine stellen Komponenten des mengenmässig am häufigsten vorkommenden Lichtsammelkomplexes LHCIIb dar, welcher sich aus etwa einem Drittel aller Proteine der Thylakoidmembranen sowie 40% des Gesamtchlorophylls zusammensetzt. Wenngleich anhand der oben diskutierten Profile von einer verminderten Photosyntheseleistung in der PGM auszugehen ist, schliesst dieses nicht aus, dass die lichtabhängige Biosynthese des Chlorophylls in der PGM Mutante hochreguliert ist und damit verbunden der Anteil Chlorophyll-assoziiierter Proteine zunimmt. Auch Chlorophyll-Fluoreszenz-Messungen von *Nicotiana sylvestris* Mutanten mit einer loss-of-function Mutation der plastidiären Phosphoglucomutase erbrachten einen gegenüber dem Wildtyp erhöhten Chlorophyllgehalt in den voll expandierten Source-Blättern bei gleichzeitig verminderter Elektronentransferrate in der Mutante [236]. Die Intermediate der Chlorophyllbiosynthese sind mehrfach als wichtige Signalmoleküle der Signalübermittlung zwischen Kern und Chloroplasten beschrieben worden. Unter Stressbedingungen wird über diese Verbindungen die Transkription diverser kerncodierter Gene initiiert und somit eine rasche Response auf veränderte Umweltbedingungen sichergestellt [260,261].

Möglich ist jedoch auch, dass der Carbohydramangel in der zweiten Hälfte der Dunkelperiode die Metabolisierung der photosynthetischen Assimilate in der darauf folgenden Lichtperiode über bisher unbekannte Mechanismen reprimiert. Dass der sich während der Dunkelperiode einstellende Carbohydramangel in der PGM eine für das Wachstum der Pflanze deutlich ungünstigere Situation darstellt, als die hohen Carbohydramkonzentrationen während der Lichtperiode, zeigen auch die Metabolitprofile der Aminosäure Prolin. Der Abbildung 3.5.3 entsprechend wiesen die Blätter der PGM bei Dunkelheit deutlich erhöhte Konzentrationen dieser bei Stress gebildeten bzw. eine Stresssituation der Pflanze anzeigenden Aminosäure auf.



**Abbildung 3.5.3: Diurnale Veränderungen von Prolin in *A. thaliana* Col-0 und PGM Blattgewebe.** Die deutliche Akkumulation dieser Aminosäure indiziert eine erhöhte Stress-situation in der PGM während der von Carbohydramangel dominierten Dunkelperiode.

### **Regulatorische Veränderungen des Stoffwechsels bedingen differentielle Korrelationsnetzwerke**

Einen weiteren, wichtigen Hinweis auf eine Veränderung in der Regulation des Stoffwechsels der PGM Mutante erbrachte die Korrelationsanalyse. Die farbcodierte Darstellung des Korrelationsnetzwerkes der Mutante offenbarte ein Cluster negativ korrelierter Aminosäuren und Carbohydrate (Abbildung 3.4.1), welches eine inverse Regulation des Stickstoff- und Kohlenstoffhaushaltes in der Mutante indiziert. Die Regulation dieser gekoppelten Systeme erfolgt über das Verhältnis von Stickstoff- und Carbohydratverbindungen zueinander. Stehen ausreichend Carbohydratverbindungen zur Verfügung, wird die Assimilation und Allokation von Stickstoffverbindungen initiiert. Die Produkte der Stickstoffassimilation hingegen hemmen die Prozesse der Stickstoffaufnahme, -reduktion und -assimilation [262]. Unter stickstofflimitierenden Bedingungen wird bei hohen Carbohydratkonzentrationen zusätzlich die Expression der RubisCO unterdrückt [263]. Als metabolische Marker für den Stickstoffstatus der Pflanze gelten die Aminosäuren Glutamin und Glutaminsäure. Sowohl die Proteinprofile der grossen (RBP\_ARATH), als auch der kleinen Untereinheit der RubisCO (At1g67090.1) sowie die Metabolitprofile der Endprodukte der Nitratassimilation (Abbildungen 3.2.1.1 und 3.3.1.1) offenbarten ähnliche Konzentrationen in beiden analysierten Genotypen. Dieses Ergebnis indiziert einen vergleichbaren Stickstoff-Status in beiden Genotypen und ist mit den Beobachtungen von Bläsing *et al.* konsistent. Die Autoren wiesen ähnliche Nitratkonzentrationen in den Blättern von *Arabidopsis* Col-0 und PGM nach [205]. Gleichzeitig wird deutlich, dass die Balance zwischen stickstoffhaltigen und kohlenstoffhaltigen Verbindungen in der PGM Mutante gestört ist. Eine denkbare Folge dieses Ungleichgewichts ist die Repression der Expression oder Translation bestimmter Gene, was sich insgesamt negativ auf das Wachstum der PGM auswirkt.

Auch die eingehende Analyse der Korrelationsprofile der akkumulierten Carbohydrate indiziert regulatorische Veränderungen im Stoffwechsel der PGM Mutante. Während Glucose und Fructose im Wildtyp über den diurnalen Rhythmus nur einen schwachen linearen Zusammenhang mit Saccharose zeigten, weisen diese Carbohydrate abhängig von Erntezeitpunkt starke, differentielle Korrelationen in der PGM auf (Abbildungen 3.4.1 und 3.4.2). Wie im Kapitel 2 diskutiert, können sich regulatorische Unterschiede innerhalb metabolischer Routen in veränderten, paarweisen Korrelationsmustern manifestieren.

Caspar *et al.* wiesen in der chloroplastidiären PGM Mutante eine erhöhte Aktivität der sauren Invertase nach [195]. Saure Invertasen nehmen im pflanzlichen Stoffwechsel ebenso wie Carbohydrate eine zentrale regulatorische Funktion ein und beeinflussen u.a. die Assimilatverteilung. Wenngleich die genaue Ursache der unterschiedlichen Steigungen der Regressionsgeraden aufgrund der Vielzahl regulatorischer Prozesse, welche im pflanzlichen Stoffwechsel über die Konzentrationen von Carbohydraten reguliert werden, nicht benannt werden kann, ist ein Zusammenhang zwischen der erhöhten Aktivität der sauren Invertase in PGM und dem beobachteten Korrelationsprofil durchaus denkbar.

### **Zusammenfassung**

Zusammenfassend zeigen die vorangegangenen diskutierten Ergebnisse, dass die im Rahmen dieses Kapitels vorgestellten Strategien zur Analyse des pflanzlichen Stoffwechsels essentielle Methoden zur Charakterisierung verschiedener *Arabidopsis* Spezies darstellen. Darüber hinaus konnte demonstriert werden, dass die angewandten nicht-zielgerichteten, multiparallelen Analysemethoden geeignete Methoden für die Stoffwechselanalyse von *Arabidopsis thaliana* darstellen und die Identifizierung neuer, potenzieller Biomarker oder Zielverbindungen für weitergehende, zielgerichtete Untersuchungen zur Regulation und Organisation des Stoffwechsels der chloroplastidiären *A. thaliana* PGM Mutante ermöglichen.

Wenngleich die Anzahl der über die Peakintegration relativ quantifizierten Proteine eher gering war, reflektieren die vergleichenden Analysen der beiden Genotypen sowohl auf Metabolit-, als auch auf Proteinebene literaturbekannte Veränderungen im Stoffwechsel der Mutante. Daneben ergaben sich aus den erhaltenen Profilen Hinweise auf bisher unbekannt, mögliche regulatorische Mechanismen.

Die im Rahmen dieser Arbeit erstellten Proteinprofile indizieren in Anbetracht der Tatsache, dass die vorangegangenen Untersuchungen von Caspar *et al.* [195] sowie Gibon *et al.* [203,210] zur CO<sub>2</sub>-Aufnahme der PGM keine Hinweise auf eine Reprimierung der Carboxylaseaktivität der RubisCO durch hohe Carbohydratkonzentrationen erbrachten, einen erhöhten Aktivierungsgrad dieses Enzyms in der zuckerakkumulieren-

den Mutante. Diese Annahme wird durch die Proteinexpressionsprofile der RubisCO Aktivase At2g39730.1, welche in der PGM zu keinem Erntezeitpunkt detektiert werden konnte, gestützt.

Anhand der erstellten Proteinprofile des Eisen-Schwefel-Zentrums der Photosystem I-Untereinheit PSAC\_ARATH sowie des Cytochrom b6-f-Komplexes (At4g03280.1) ergaben sich für die PGM zudem Hinweise auf einen verminderten photosynthetischen Elektronentransport sowie ein vermindertes Reduktionspotenzial im Stroma der Chloroplasten.

Als mögliche Response auf die Limitierung der photosynthetisierten Reduktionsäquivalente zeigte die chloroplastidiäre GAP-DH eine deutlich verminderte Proteinexpression, was auf eine Herabregulierung des reduktiven Calvin Zyclus in der PGM hindeutet.

Die erhöhten Salicylsäuregehalte in der PGM bedingen möglicherweise eine Veränderung in der Regulation der Stomatabewegung in der Mutante. In diesem Zusammenhang stellt auch das funktionell bisher nicht charakterisierte Protein At1g13930.1 einen interessanten potenziellen Biomarker dar.

Polyamine und Intermediate des Harnstoffzyclus deuten auf eine durchgehende Stresssituation der PGM Mutante hin, welche durch die Carbohydratakkumulation während der Licht-, sowie den -mangel während der Dunkelperiode bedingt ist. Dabei kommt als mögliche Ursache des verminderten Wachstums der PGM unter Kurz- und/oder Langtagbedingungen eine vermehrte Entstehung von ROS in den Peroxisomen durch die  $\beta$ -Oxidation von Fettsäuren während der durch Carbohydratmangel gekennzeichneten Dunkelperiode in Betracht. Dieses ist jedoch nur durch die Inkubation isotope markierter Verbindungen zu klären.

Des Weiteren konnte an zwei Beispielen gezeigt werden, dass die vergleichende Analyse von Korrelationsnetzwerken eine sinnvolle Ergänzung zu den konventionellen datenanalytischen Verfahren, wie z.B. dem auf die Erkennung von Mittelwertsunterschieden basierenden T-Test, darstellt und Hinweise auf regulatorische Veränderungen in der komplexen Organisation des Stoffwechsels geben kann.

Dennoch ist die Aussagefähigkeit der mittels der Shotgun-Proteinanalytik gewonnenen Daten im Hinblick auf regulatorische Mechanismen aufgrund der fehlenden Information über die Proteinaktivität begrenzt, so dass die Messung von Enzymaktivitäten nach wie vor eine wichtige Methode zur Untersuchung des Stoffwechsels darstellt.

Beide Methoden, sowohl die Analyse kleiner Moleküle mittels GC-TOF-MS, als auch die LC-MS basierte Methode zur indirekten, relativen Quantifizierung von Proteinen aus komplexen Proben, können problemlos zur Analyse von Untersuchungsmaterial verschiedenen Ursprungs verwendet werden. Ein besonderer Vorteil liegt dabei in der kurzen Analysezeit insbesondere der Metabolitanalytik (ca. 25 min), welche zudem nur geringe Probenmengen (1-10 mg Blattmaterial) erfordert.

## **4 Systemische temperaturbedingte abiotische Stressresponse in *Arabidopsis thaliana***

### **4.1 Die Stressbiologie der Pflanze**

Pflanzen sind in ihrer natürlichen Umgebung vielfältigen Umwelteinflüssen wie Wasserangebot, Temperatur oder UV-Strahlung ausgesetzt, welche tages- und jahreszeitlichen Schwankungen unterliegen. Bei annähernd optimaler Anpassung an die Umgebung, wie sie im Verlaufe der Evolution erreicht wurde, tritt normalerweise kein Stress auf. Plötzliche Veränderungen der den Standort charakterisierenden Umweltfaktoren führen jedoch häufig dazu, dass der zumeist eng begrenzte Bereich der optimalen Anpassung überschritten wird, was in einer Beeinträchtigung des Stoffwechsels oder der Entwicklung der Pflanze resultiert. Als Folge dieser physiologischen Belastung stellt sich Stress ein. Während abiotischer Stress durch physikalische oder chemische Ökofaktoren wie z.B. Kälte, Hitze, Trockenheit oder auch Nährstoffmangel hervorgerufen wird, fasst man die im Zusammenhang mit dem Angriff von Pathogenen, Parasiten oder anderen konkurrierenden Organismen hervorgerufenen Stresssituationen unter dem Begriff biotischer Stress zusammen [33-35].

Die meisten Pflanzen verfügen über eine Vielzahl mehr oder weniger spezifischer Abwehrmechanismen, welche ihnen das Überleben auch unter ungünstigen Umweltbedingungen ermöglicht. Wirkt Stress auf eine Pflanze ein, so reagiert diese im Hinblick auf den Stressfaktor sowohl mit spezifischen, als auch mit unspezifischen Reaktionen. Jede Art von Stress beeinflusst die Pflanze nicht nur durch den von ihr hervorgerufenen Primärstress, sondern auch durch Folgereaktionen. So tritt beispielsweise Wasserstress als Primärstress bei Trockenheit, als Sekundärstress bei hohen und niedrigen Temperaturen sowie Salzstress, oder auch als Tertiärstress nach starker ionisierender

Strahlung auf [264,265]. Dieses komplexe Zusammenwirken von spezifischen und unspezifischen Stressantworten lässt häufig keinen Schluss auf die eigentliche 'Schwachstelle' in der pflanzlichen Stressabwehr zu, so dass in den meisten Fällen nur der Grad der durch Stress verursachten Schädigung quantifizierbar ist.

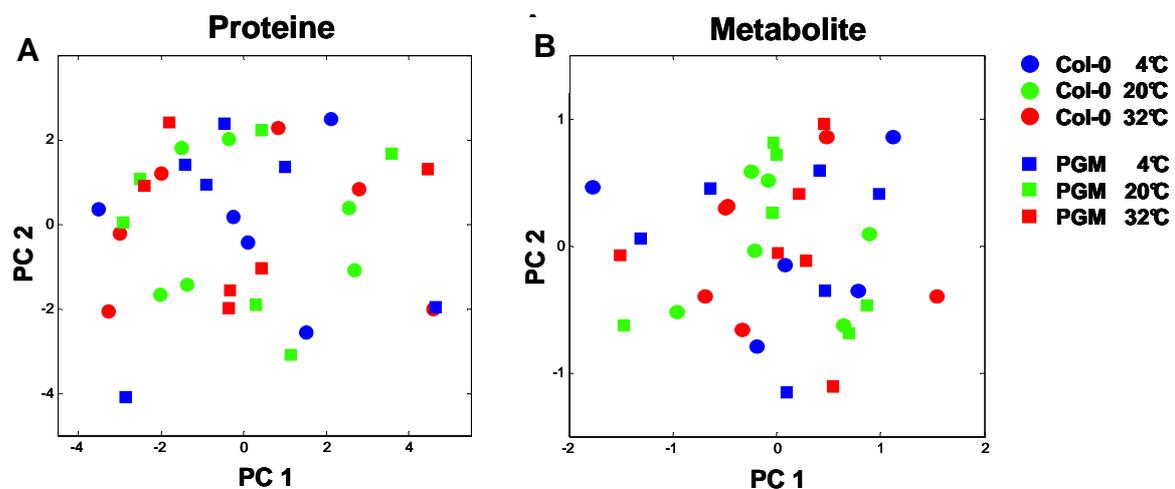
Zur Erweiterung der im Kapitel 2.8 analysierten zeitabhängigen Regulation des Metabolismus von *A. thaliana* wurden die Auswirkungen von temperaturbedingtem abiotischen Stress auf den Stoffwechsel der zuckerakkumulierenden, stärkefreien PGM Mutante im Vergleich zum korrespondierenden Wildtyp Col-0 analysiert. Hierzu wurden beide Genotypen nach Erreichen des Entwicklungsstadiums 5.1 [177] in 6 biologischen Replikaten für jeweils 3 Tage den Temperaturen 4°C und 32°C ausgesetzt und zur integrativen Analyse herangezogen. Um die Adaption an niedrige und hohe Temperaturen zu analysieren und allgemeine, durch Temperaturstress bedingte Marker in Abhängigkeit vom Genotyp zu identifizieren, wurden Kontrollpflanzen bei 20°C kultiviert.

Polare Metabolite wurden mit CH<sub>3</sub>OH/CHCl<sub>3</sub>/H<sub>2</sub>O entsprechend 8.3.1 extrahiert. Korrespondierende Proteine wurden aus denselben Proben nach einem neu entwickelten Protokoll phenolisch mit Saccharose gemäss 8.3.3 extrahiert. Aufgrund erhöhter Ausbeuten ermöglichte dieses Protokoll die individuelle Analyse jeder einzelnen Pflanze. Die Quantifizierung identifizierter Proteine erfolgte über den Spektren-Count, was zudem zu einer deutlichen Erhöhung der Anzahl quantifizierter Proteine führte. Somit konnten neben 80 mit GC-TOF-MS quantifizierten Metaboliten zumeist bekannter biochemischer Funktion 179 Proteine mittels eindimensionaler LC-MS Shotgun-Proteinanalyse relativ quantifiziert werden.

Bedingt durch die vornehmlich sessile Lebensweise der Pflanzen hat sich im Verlaufe der Evolution im Pflanzenreich ein beachtliches Arsenal hochwirksamer Mechanismen zur Erlangung einer gewissen Stressresistenz herausgebildet. In vielen Fällen wird diese Resistenz erst durch den Einfluss eines oder mehrerer Stressfaktoren induziert, so dass eine erhöhte Stresstoleranz eine Stressakklimatisierung voraussetzt. Sofern das Einwirken eines Stressors die Pflanze nicht irreversibel schädigt, kann dieses eine morphologische, physiologische und/oder biochemische Anpassung der Pflanze auslösen, was in einer gewissen Abhärtung gegenüber suboptimalen Bedingungen resultiert. Die mit der Akklimatisierung an eine bestimmte Stresssituation verbundenen regulatorischen Veränderungen können wenige Stunden bis Tage andauern [33-35].

Um den Grad der Anpassung der im Rahmen dieser Arbeit analysierten Pflanzen an die jeweiligen Temperaturbedingungen abzuschätzen, wurde zunächst die inhärente biologische Varianz gestresster und ungestresster Pflanzen verglichen. Dazu wurden die erhaltenen Protein- und Metabolit-Datenmatrices beider Genotypen mit Hilfe der PCA in

ihrer Dimension zunächst reduziert, so dass diese anschliessend visualisiert werden konnten. Wie Abbildung 4.1.1 zeigt, wiesen die temperaturgestressten Pflanzen beider Genotypen sowohl auf der Ebene der Proteine, als auch auf der Ebene der Metabolite biologische Variationen auf, welche denen der korrespondierenden ungestressten Kontrollpflanzen entsprachen. Dieses Ergebnis indiziert, dass während der 72-stündigen Exposition bei 4°C und/oder 32°C die mit dem Erwerb einer erhöhten Stresstoleranz verbundenen physiologischen und biochemischen Veränderungen weitgehend abgeschlossen und die Pflanzen somit an die jeweiligen Temperaturen angepasst waren.



**Abbildung 4.1.1 :** Gegenüberstellung der biologischen Varianz der Protein- (A) und Metabolitkonzentrationen (B) der gestressten und ungestressten *A. thaliana* Genotypen Col-0 und PGM. Dargestellt sind die aus der Analyse von jeweils 6 biologischen Replikaten resultierenden biologischen Varianzen von insgesamt 179 relativ quantifizierten Proteinen (Abb. 4.1.1-A) sowie 80 relativ quantifizierten Metaboliten (Abb. 4.1.1-B) nach Dimensionsreduktion durch die PCA, basierend auf  $\log_{10}$ -transformierten Daten, wobei jeder Punkt eine Probe repräsentiert. Ein Vergleich der Streuung der Datenpunkte macht deutlich, dass die für 72 Stunden bei 4°C und 32°C exponierten Pflanzen beider Genotypen bezüglich der erhobenen Protein- und Metabolitdaten eine ähnliche biologische Varianz aufwiesen wie die ungestressten Kontrollpflanzen, was eine Adaption an diese Temperaturen indiziert.

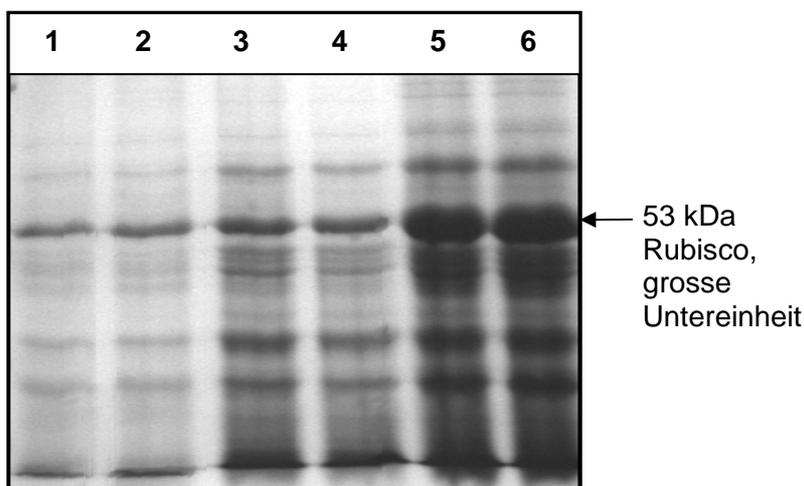
## 4.2 Protein Profiling temperaturgestresster *A. thaliana* Col-0 und PGM Pflanzen

Im vorangegangenen Kapitel 3.2 wurden verbunden mit der Proteinanalytik verschiedene Schwierigkeiten offenbar. Die relativ geringen Ausbeuten erforderten nach der Anwendung des dort verwendeten Extraktionsprotokolls die Vereinigung von Proteinpellets, was zu einer deutlich verminderten Anzahl analysierter Replikate führte und die Auswertung erheblich erschwerte. Um die Ausbeuten der extrahierten löslichen Proteine zu verbessern und somit die Dimensionalität der integrativ auswertbaren Datenmatrix zu

erhöhen, wurden verschiedene Extraktionsmethoden getestet. Daneben galt es, Urea, welches häufig als chaotropes Reagenz zur Lösung und Denaturierung von Proteinen verwendet wird, aber zugleich einer temperaturkatalysierten Degradation unter Bildung von Isocyanat unterliegt, wenn möglich geeignet zu substituieren.

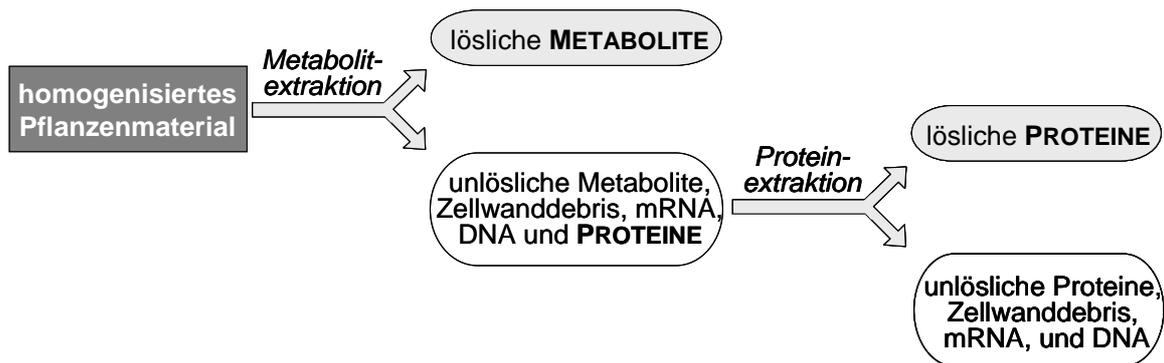
#### 4.2.1 Phenolische Extraktion löslicher Proteine mit Saccharose

Die phenolische Extraktion mit Urea [94] wurde zunächst dahingehend modifiziert, dass der 30 minütige Schritt der Resuspendierung präzipitierter Proteine eingespart und die Zugabe von 500 µL TE-Puffer (10 mM TRIS, 1 mM EDTA-Na<sub>2</sub>)-gesättigtem Phenol direkt ohne vorherige Inkubation erfolgte. Daneben wurde eine neue Extraktionsmethode getestet, in welcher Urea durch Saccharose in hohen Konzentrationen (40% w/v) substituiert wurde (Kapitel 8.3.3) durch konnten die Proteinausbeuten im Vergleich zur Extraktion mit Urea um ca. 150% auf etwa 2.5 µg/mg FW gesteigert werden. Abbildung 4.2.1.1 veranschaulicht dieses Ergebnis; hier wurden, ausgehend von gleichen Frischgewichten, Gesamtproteinfraktionen resultierend aus den getesteten Extraktionsmethoden in Doppelbestimmung auf einem 10%-SDS-Gel aufgetrennt. Deutlich zu erkennen ist die signifikante Erhöhung der Ausbeute an Gesamtprotein bei Extraktion mit Saccharose (Spur 5 und 6), wobei die relativen Verhältnisse einzelner Proteine zueinander invariant blieben. Aus den Proteinausbeuten der verkürzten phenolischen Extraktion mit Urea (Spur 3 und 4) ergab sich keine in solchem Masse signifikante Verbesserung. Im Folgenden wurden Proteine daher ausschliesslich phenolisch mit Saccharose extrahiert.



**Abbildung 4.2.1.1: Gegenüberstellung löslicher Gesamtproteinfraktionen aus *A. thaliana* Col-0.** Dargestellt ist die Auftrennung von Gesamtproteinfraktionen über SDS-PAGE (10%). Die Proteinextraktion erfolgte gemäss drei unterschiedlichen Protokollen: 1+2: phenolische Extraktion mit Urea entsprechend Kapitel 8.3.2, 3+4: phenolische Extraktion mit Urea nach verkürztem Protokoll (siehe Text), 5+6: phenolische Extraktion mit Saccharose gemäss 8.3.3.

Ein weiterer Vorteil dieser Extraktionsmethode lag neben der deutlich verbesserten Proteinausbeute in der Praktikabilität und Schnelligkeit dieses Protokolls. Durch die hohe Saccharosekonzentration wurde eine Phasenumkehr induziert, infolge derer die phenolische die obere Phase darstellte (Abbildung 4.2.1.2). Dies erleichterte die Separation der gelösten Proteine von unlöslichen Bestandteilen. Im Vergleich zur phenolischen Extraktion mit Urea konnte die gesamte Prozedur der Extraktion zudem um ca. 60 min verkürzt werden.



**Abbildung 4.2.1.2: Flowchart der integrativen Extraktion von Metaboliten und Proteinen aus einer biologischen Probe.** Ausgehend von homogenisiertem Pflanzenmaterial wurden zunächst polare Metabolite extrahiert, anschließend konnten lösliche Proteine aus dem verbliebenen heterogenen Pellet in die phenolische (obere) Phase isoliert werden.

#### 4.2.2 Chemometrische Analyse der Proteindaten

Um die im Rahmen der Shotgun-Analyse ermittelten relativen Änderungen einzelner Proteinkonzentrationen den experimentellen Bedingungen zuzuordnen, wurde die erhaltene Datenmatrix, wie schon im Kapitel 3, zunächst einer Varianzanalyse unterzogen. Mit diesem Verfahren wurde geprüft, ob und für welche Proteine die Varianz zwischen den untersuchten Stichproben grösser war als die Varianz innerhalb der Stichproben. Der dazu angewandte Kruskal-Wallis-Test ergab, dass mit einer Irrtumswahrscheinlichkeit von  $p \leq 0.05$  Unterschiede zwischen den 6 Stichproben (Col-0: 4°C, 20°C, 32°C; PGM: 4°C, 20°C, 32°C) der untersuchten Grundgesamtheiten in der Expression von 30 der insgesamt 179 detektierten Proteine nachweisbar waren. Gemäss Ihrer funktionellen Charakterisierung sind viele dieser Proteine in der Stressresponse involviert oder besitzen zumindest eine starke Homologie zu solch bekannten Proteinen. Diese sind der Tabelle A-2.1 im Anhang zu entnehmen.

Einige, darunter weitere in der Stressresponse involvierte Proteine wurden unter bestimmten experimentellen Bedingungen nicht detektiert. Diese, innerhalb einer Stichprobe numerische Werte von Null und somit eine identische Ausprägung aufweisenden Variablen stellen Grenzfälle dar, welche aufgrund ihrer Verteilung einer Varianzanalyse nicht zugänglich sind. Unter den in Tabelle 4.2.2.1 aufgeführten Variablen fielen dennoch weitere Proteine auf, deren Expressionsprofile im Einklang mit dem experimentellen Setup standen. So konnten das 'cold-responsive' Protein At2g42530.1, das 'stress-induced/stress-responsive' Protein At5g15960.1 sowie das 'low-temperature-responsive' Protein At5g52310 ausschliesslich in den kältegestressten Pflanzen beider Genotypen nachgewiesen werden. Des Weiteren wiesen ribosomale 60S Untereinheiten bildende Proteine eine differentielle Expressionen auf, was auf eine durch Temperaturstress hervorgerufene veränderte Regulation der Translation hindeutet.

Zur Visualisierung und Klassifizierung der durch Kälte- und/oder Hitzestress hervorgerufenen Veränderungen wurden die Proteine mit einer nachweisbar hohen Varianz zwischen den analysierten Stichproben zusammen mit den partikulär nicht detektierten Proteinen einem Bi-Clustering unterzogen (Abbildung 4.2.2.1).

Die resultierende Hierarchie lässt multifaktorielle Unterschiede zwischen den experimentellen Bedingungen erkennen. Wenngleich drei verschiedene Cluster erkannt werden können, welche darauf hindeuten, dass sich die Expressionsprofile der bei 4°C und 20°C geernteten Pflanzen beider Genotypen ähneln und deutlich von denen der hitzestressten Pflanzen verschieden sind, machen die grossen Distanzen zwischen den Stichproben ihre Verschiedenartigkeit deutlich. Gleichzeitig können aus der Hierarchie der Variablen Klassen von Proteinen identifiziert werden, deren Expression als mögliche Response auf Temperaturstress veränderlich waren. So wurde beispielsweise ein Cluster von Proteinen offenbar, deren Expressionsprofile auf eine Hochregulation bei 32°C hindeuteten. Daneben wies die Peptidyl-prolyl *cis-trans* Isomerase At2g21130.1 ebenso wie die weder bei 4°C im Wildtyp, noch bei 20°C in der Mutante detektierte Isoform At3g38740.1 deutlich erhöhte Konzentrationen bei 32°C auf. Ähnlich erhöhte Expressionen bei 32°C zeigten auch die zwei Isoformen eines 'major latex-related' Proteins (At3g26450.1 und At4g23670) sowie das bei 4°C im Wildtyp undetektierte, als Antioxidant annotierte Protein At1g65980.1 (Abbildung 4.2.2.1).

Eine verminderte Expression nach Hitzestress wiesen zwei kerncodierte Ribonucleoproteine mit chloroplastidiärer Signalsequenz (At2g37220.1, At3g53460.1), das in der Mutante bei 32°C nicht nachweisbare Hitzeschock-Protein At3g09440.1 sowie das im Wildtyp nach Hitzestress nicht detektierte ribosomale 60S-Protein At2g39460.1 auf.

Tabelle 4.2.2.1: Durch Temperaturstress veränderte Proteinexpressionen von *A. thaliana*

**Col-0 und PGM Pflanzen.** Angegeben sind die resultierenden Mittelwerte (MW) aus den in 6 biologischen Replikaten über den Spektren-Count quantifizierter Proteine einer jeden Stichprobe mit Standardfehler (SE). Die unter bestimmten experimentellen Bedingungen nicht detektierten Proteine sind mit n.d. gekennzeichnet. Die mit  $p \leq 0.05$  von den bei 20°C kultivierten Kontrollpflanzen signifikant unterschiedlichen Konzentrationen einzelner Proteine nach entsprechendem Temperaturstress sind farblich hervorgehoben.

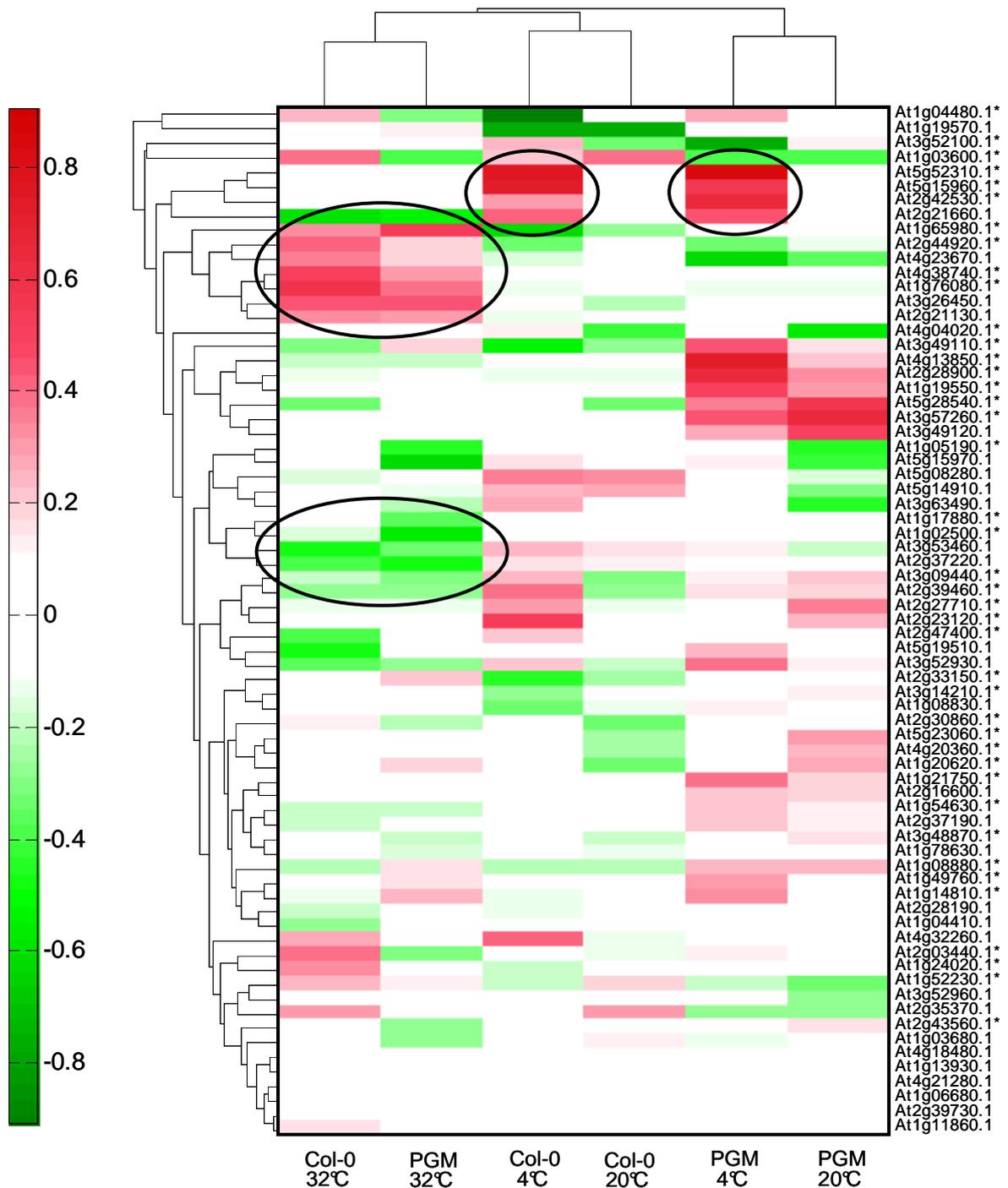
	Col-0 4°C MW ± SE	Col-0 20°C MW ± SE	Col-0 32°C MW ± SE	PGM 4°C MW ± SE	PGM 20°C MW ± SE	PGM 32°C MW ± SE
<b>Response auf Kälte- und Hitzestress</b>						
At2g30860.1	2.0 ± 0.4	n.d.	2.8 ± 0.4	2.3 ± 0.3	2.5 ± 0.3	1.3 ± 0.1
At1g20620.1	2.0 ± 0.3	n.d.	2.0 ± 0.2	2.3 ± 0.2	4.2 ± 0.3	3.3 ± 0.3
At4g20360.1	1.6 ± 0.2	n.d.	2.0 ± 0.4	1.8 ± 0.2	3.0 ± 0.3	1.7 ± 0.2
At5g23060.1	1.6 ± 0.2	n.d.	1.5 ± 0.2	2.0 ± 0.2	3.3 ± 0.3	1.8 ± 0.3
At1g14810.1	n.d.	1.3 ± 0.1	n.d.	2.8 ± 0.2	1.3 ± 0.1	2.3 ± 0.3
At4g04020.1	4.8 ± 0.8	1.3 ± 0.1	4.5 ± 0.6	4.0 ± 0.4	n.d.	3.2 ± 0.4
At1g52230.1	1.4 ± 0.2	3.2 ± 0.6	3.7 ± 0.6	1.3 ± 0.1	n.d.	2.8 ± 0.4
<b>Response auf Kältestress</b>						
At5g52310.1	5.8 ± 0.5	n.d.	n.d.	7.0 ± 0.2	n.d.	n.d.
At5g15960.1	5.4 ± 0.7	n.d.	n.d.	3.7 ± 0.4	n.d.	n.d.
At2g42530.1	2.0 ± 0.3	n.d.	n.d.	4.5 ± 0.4	n.d.	n.d.
At2g44920.1	n.d.	2.7 ± 0.4	5.3 ± 0.8	n.d.	1.5 ± 0.2	3.2 ± 0.8
At5g28540.1	2.0 ± 0.3	n.d.	n.d.	4.7 ± 0.5	7.5 ± 0.5	2.2 ± 0.4
At1g04480.1	n.d.	8.8 ± 1.5	14.3 ± 2.5	15.5 ± 2.3	7.8 ± 0.9	4.2 ± 0.8
At1g24020.1	n.d.	1.8 ± 0.2	3.5 ± 0.4	1.7 ± 0.2	1.5 ± 0.2	1.3 ± 0.1
At2g33150.1	n.d.	1.7 ± 0.2	2.7 ± 0.2	3.0 ± 0.4	3.0 ± 0.3	4.7 ± 0.5
At3g14210.1	n.d.	1.7 ± 0.2	2.0 ± 0.2	2.3 ± 0.3	2.5 ± 0.3	1.8 ± 0.2
At3g49110.1	n.d.	1.8 ± 0.2	1.7 ± 0.2	9.8 ± 1.2	5.0 ± 0.6	5.2 ± 0.4
At1g05190.1	3.2 ± 0.4	3.2 ± 0.4	2.7 ± 0.4	3.0 ± 0.2	n.d.	n.d.
At3g52100.1	10.0 ± 1.6	2.7 ± 0.6	6.7 ± 1.0	n.d.	8.0 ± 1.3	4.8 ± 0.9
<b>Response auf Hitzestress</b>						
At2g03440.1	1.8 ± 0.3	1.5 ± 0.2	5.0 ± 0.5	2.7 ± 0.4	2.2 ± 0.3	n.d.
At1g54630.1	1.8 ± 0.3	1.3 ± 0.1	n.d.	2.5 ± 0.3	2.2 ± 0.3	n.d.
At2g47400.1	4.0 ± 0.4	2.7 ± 0.3	n.d.	2.5 ± 0.4	2.3 ± 0.3	2.3 ± 0.3
At1g02500.1	3.6 ± 0.3	3.8 ± 0.4	2.7 ± 0.3	4.0 ± 0.5	4.3 ± 0.3	n.d.
At1g17880.1	2.2 ± 0.2	2.5 ± 0.3	1.8 ± 0.2	2.7 ± 0.2	2.3 ± 0.2	n.d.
At2g43560.1	2.0 ± 0.3	1.7 ± 0.2	2.0 ± 0.4	1.5 ± 0.2	2.7 ± 0.3	n.d.
At1g21750.1	n.d.	n.d.	n.d.	2.5 ± 0.3	1.5 ± 0.2	n.d.
At3g57260.1	n.d.	n.d.	n.d.	2.8 ± 0.4	4.5 ± 0.4	n.d.
<b>Differentielle Response</b>						
At3g09440.1	3.6 ± 0.4	n.d.	1.3 ± 0.1	2.7 ± 0.3	3.3 ± 0.3	n.d.
At3g48870.1	1.6 ± 0.2	n.d.	1.5 ± 0.2	1.8 ± 0.2	2.2 ± 0.3	n.d.
At2g23120.1	4.2 ± 0.4	n.d.	n.d.	1.5 ± 0.2	2.2 ± 0.3	n.d.
At2g27710.1	2.6 ± 0.4	n.d.	n.d.	1.7 ± 0.2	3.0 ± 0.3	n.d.
At2g39460.1	4.4 ± 0.4	n.d.	n.d.	2.7 ± 0.3	2.8 ± 0.3	n.d.
At4g13850.1	1.6 ± 0.2	1.5 ± 0.2	n.d.	8.5 ± 1.2	2.5 ± 0.3	n.d.
At1g76080.1	n.d.	1.7 ± 0.2	5.3 ± 0.5	n.d.	n.d.	3.2 ± 0.3
At4g38740.1	n.d.	1.3 ± 0.1	3.8 ± 0.3	n.d.	n.d.	2.3 ± 0.2
At1g65980.1	n.d.	2.2 ± 0.4	8.8 ± 1.4	4.8 ± 1.0	3.3 ± 0.7	14.0 ± 1.1
At1g08880.1	n.d.	n.d.	n.d.	3.0 ± 0.3	3.0 ± 0.4	2.3 ± 0.3
At1g19550.1	n.d.	n.d.	n.d.	3.8 ± 0.4	2.3 ± 0.3	1.3 ± 0.1
At1g49760.1	n.d.	n.d.	n.d.	2.3 ± 0.2	1.3 ± 0.1	1.7 ± 0.2
At2g28900.1	n.d.	n.d.	n.d.	5.8 ± 0.5	2.8 ± 0.4	1.7 ± 0.2
At1g03600.1	4.0 ± 0.8	6.2 ± 0.9	6.0 ± 0.9	n.d.	n.d.	n.d.

glutathione S-transferase  
catalase 3 (SEN2)  
elongation factor Tu / EF-Tu  
expressed protein  
semialdehyde dehydrogenase family protein  
plastid-lipid associated protein PAP, putative  
photosystem I reaction center subunit VI, chloroplast

low-temperature-responsive protein 78 / desiccation-responsive protein 29A  
stress-responsive protein / stress-induced protein (KIN1)  
cold-responsive protein / cold-regulated protein  
thylakoid lumenal 15 kDa protein, chloroplast  
luminal binding protein 1  
60S ribosomal protein L23  
Bet v I allergen family protein similar to major pollen allergen  
acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiolase, putative  
myrosinase-associated protein  
peroxidase 33 / neutral peroxidase C precursor  
ribosomal protein L6 family protein  
PHD finger family protein contains Pfam profile

nodulin-related similar to Early nodulin 12B precursor  
acyl carrier protein 3, chloroplast  
CP12 domain-containing protein  
S-adenosylmethionine synthetase 1  
nascent polypeptide-associated complex domain-containing protein / BTF3b-like transcription factor  
immunophilin / FKBP-type peptidyl-prolyl cis-trans isomerase family protein  
protein disulfide isomerase  
glycosyl hydrolase family 17 protein

heat shock cognate 70 kDa protein 3  
ATP-dependent Clp protease ATP-binding subunit  
expressed protein  
60S acidic ribosomal protein P2  
60S ribosomal protein L23A  
glycine-rich RNA-binding protein  
thioredoxin family protein  
peptidyl-prolyl cis-trans isomerase  
peroxiredoxin type 2 (alkyl hydroperoxide reductase and thiol-specific antioxidant)  
histone H2A  
dehydroascorbate reductase, putative  
polyadenylate-binding protein  
mitochondrial import inner membrane translocase subunit  
photosystem II family protein



**Abbildung 4.2.2.1: Bi-Clustering und farbcodierte Darstellung der 30 nach Temperaturstress in *A. thaliana* Col-0 und/oder PGM Pflanzen mit  $p \leq 0.05$  relevante relative Änderungen in ihren Konzentrationen aufweisenden Proteine sowie partikulär undetektierete Proteine.** Das Bi-Clustering erfolgte nach  $\log_{10}$ -Transformation der Median-normalisierten Proteinkonzentrationen mit dem Pearson'schen Produktmoment-Korrelationskoeffizienten als Abstandsmass. Erhöhte Expressionen innerhalb der standardisierten Datenmatrix sind rot, verminderte Expressionslevel sind grün dargestellt. Die mit \* gekennzeichneten Proteine wurden unter bestimmten experimentellen Bedingungen detektiert (vgl. dazu Tabelle 4.2.2.1). Diese farbcodierte Darstellung lässt verschiedene Proteine erkennen, deren Expression Kältestress-induziert in beiden Genotypen hochreguliert war, darunter u.a. das kältere-regulierte Protein At5g15970.1. Daneben werden Cluster von Proteinen mit differentieller Expression nach Hitzestress deutlich.

Des Weiteren waren 4 Proteine in beiden Genotypen bei Kältestress hochreguliert. Neben den Kälte-induzierten Proteinen At2g42530.1 und At5g52310.1 sowie dem 'stress-responsive/stress-induced' Protein At5g15960.1, welche ausschliesslich in den an 4°C adaptierten Pflanzen beider Genotypen nachweisbar waren, fiel in der farbcodierten Darstellung auch das Glycin-reiche RNA-bindende Protein At2g21660.1 durch erhöhte Expression bei 4°C auf (Abbildung 4.2.2.1).

Um eine weitergehende Zuordnung der Proteine mit nachweisbar veränderten Expressionsleveln zu den jeweiligen Stressfaktoren vornehmen zu können wurde getestet, welche Unterschiede innerhalb der Genotypen zwischen den Temperaturbehandlungen nachweisbar waren. Dazu wurden die Expressionen der jeweils bei 4°C und 32°C geernteten Pflanzen mit denen der bei 20°C kultivierten Pflanzen verglichen, wobei die Nullhypothese gleicher Expressionen bei  $p \leq 0.05$  verworfen wurde. Mit Hilfe der auf dem Kruskal-Wallis-Test basierenden multiplen Varianzanalyse konnte weiter nachgewiesen werden, dass die Unterschiede in den Expressionen von 24 Proteinen auf Stressreaktionen zurückzuführen sind. Unter diesen sind laut Annotation 9 Proteine in der Stressresponse involviert. Deren relativen Änderungen sind in Tabelle 4.2.2.2 zusammengestellt. Zugleich wurde deutlich, dass beide untersuchten Genotypen sehr unterschiedlich auf Temperaturstress reagierten. Somit konnten sowohl genotypspezifische Marker als auch für die jeweiligen Stresssituationen spezifische Markerproteine identifiziert werden. Die Expression des Glycin-reichen RNA-bindenden Proteins At2g21660.1 wies in beiden Genotypen eine starke Response sowohl auf Kälte-, als auch auf Hitzestress auf, wobei die Expression mit steigender Temperatur abnahm.

Kältestress induzierte in beiden untersuchten Genotypen individuelle Stressreaktionen. Während die chloroplastidiäre ATP-Synthase At4g32260.1 neben der cytosolischen Fructose 1,6-bisphosphat Aldolase At3g52930.1 in gestressten Wildtyp Pflanzen hochreguliert war, konnten das 'cold-responsive/cold-regulated' Protein At5g15970.1 sowie ein weiteres Glycin-reiches RNA-bindendes Protein At1g13850.1 als Markerproteine für kältegestresste PGM Pflanzen identifiziert werden. Beide Enzyme waren hier ebenfalls hochreguliert.

Als Reaktionen auf Hitzestress war in beiden Genotypen eine verstärkte Expression eines 'major-late' ähnlichen Proteins (At3g26450.1) sowie einer Peptidyl-prolyl *cis-trans* Isomerase zu verzeichnen. Das Ribonucleoprotein At2g21130.1 mit chloroplastidiärer Signalsequenz hingegen erschien in beiden Genotypen durch Hitzestress herabreguliert. Eine Wildtyp-spezifische Response auf Hitzestress liess sich von den Expressionsprofilen der Dehydroascorbat Reduktase At1g19570.1 und einer weiteren Isoform eines 29kDa Ribonucleoproteins mit chloroplastidiärer Zielsequenz (At3g53460.1) ableiten. In den

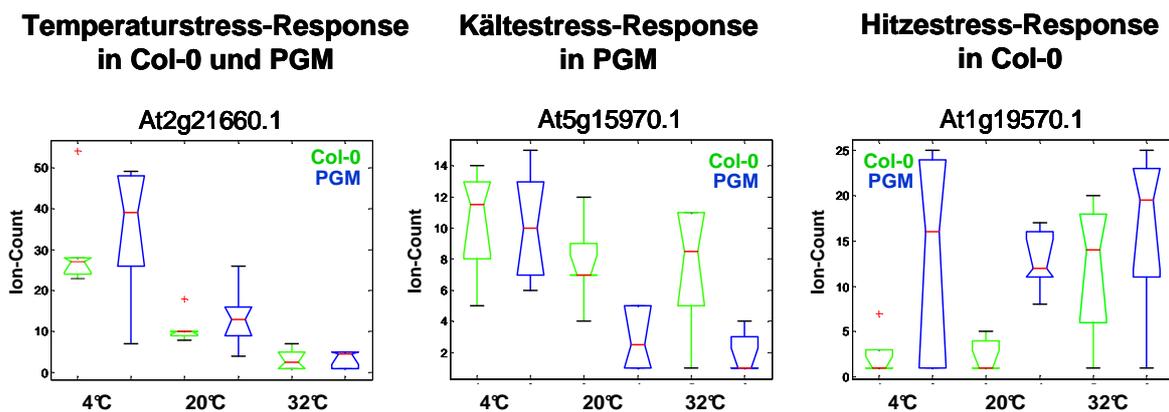
Tabelle 4.2.2.2:

**Einfluss von temperaturbedingtem abiotischen Stress auf die Protein-expressionen in *A. thaliana* Col-0 und PGM.** Angegeben sind die mit  $p \leq 0.05$  signifikanten relativen Änderungen einzelner Proteinkonzentrationen zu den Referenzwerten der bei 20°C kultivierten Kontrollpflanzen beider Genotypen, dargestellt als Verhältnisse (X-fold) mit Irrtumswahrscheinlichkeit (p-Wert).

	Col-0 4°C X-fold (p-Wert)	Col-0 32°C X-fold (p-Wert)	PGM 4°C X-fold (p-Wert)	PGM 32°C X-fold (p-Wert)	
<b>Response auf Kälte- und Hitzestress</b>					
At2g21660.1	<b>2.9</b> ( 0.006 )	<b>-3.4</b> ( 0.003 )	<b>2.6</b> ( 0.045 )	<b>-3.9</b> ( 0.019 )	glycine-rich RNA-binding protein
<b>Response auf Kältestress</b>					
At4g32260.1	<b>3.6</b> ( 0.007 )				ATP synthase family, chloroplast fructose-bisphosphate aldolase, cytosolic stress-responsive protein (KIN2) / stress-induced protein (KIN2) (cold-induced) expressed protein weakly similar to drought-induced sunflower protein SDi-6 superoxide dismutase
At3g52930.1	<b>2.4</b> ( 0.019 )				
At5g15970.1			<b>3.6</b> ( 0.004 )		
At1g13930.1			<b>1.3</b> ( 0.019 )		
At1g08830.1			<b>1.3</b> ( 0.036 )		
<b>Response auf Hitzestress</b>					
At3g26450.1		<b>4.4</b> ( 0.004 )		<b>2.9</b> ( 0.006 )	major latex protein-related protein nodulin-related similar to Early nodulin 12B precursor peptidyl-prolyl cis-trans isomerase 29 kDa ribonucleoprotein, chloroplast precursor dehydroascorbate reductase oxygen-evolving enhancer protein 3, chloroplast 29 kDa ribonucleoprotein, chloroplast precursor hydroxymethylbilane synthase / porphobilinogen deaminase, chloroplast malate dehydrogenase, cytosolic magnesium-chelatase subunit chlI, chloroplast peroxiredoxin type 2 (alkyl hydroperoxide reductase and thiol-specific antioxidant) peroxidase peptidyl-prolyl cis-trans isomerase, cytosolic
At2g03440.1		<b>3.3</b> ( 0.045 )		<b>n.d.</b>	
At2g21130.1		<b>2.2</b> ( 0.022 )		<b>2.6</b> ( 0.027 )	
At2g37220.1		<b>-3.2</b> ( 0.004 )		<b>-2.5</b> ( 0.004 )	
At1g19570.1		<b>5.6</b> ( 0.020 )			
At4g21280.1		<b>1.3</b> ( 0.016 )			
At3g53460.1		<b>-4.2</b> ( 0.007 )			
At5g08280.1		<b>-3.1</b> ( 0.016 )			
At1g04410.1		<b>-1.9</b> ( 0.005 )			
At4g18480.1		<b>-1.3</b> ( 0.019 )			
At3g52960.1				<b>2.1</b> ( 0.016 )	
At3g49120.1				<b>-3.6</b> ( 0.021 )	
At2g16600.1				<b>-1.8</b> ( 0.044 )	
<b>Differentielle Response</b>					
At4g23670.1	<b>-1.8</b> ( 0.006 )	<b>1.8</b> ( 0.023 )		<b>3.5</b> ( 0.004 )	major latex protein-related protein ribosomal protein L13 family protein photosystem II oxygen-evolving complex 23 peroxiredoxin type 2 (alkyl hydroperoxide reductase and thiol-specific antioxidant) glycine-rich RNA-binding protein
At1g78630.1	<b>1.8</b> ( 0.017 )			<b>-1.8</b> ( 0.043 )	
At1g06680.1	<b>-1.2</b> ( 0.011 )			<b>1.1</b> ( 0.037 )	
At1g65980.1	<b>n.d.</b>			<b>4.2</b> ( 0.031 )	
At4g13850.1		<b>n.d.</b>	<b>3.4</b> ( 0.033 )		

hitze gestressten PGM Pflanzen fielen die Isoformen einer Alkyl-hydroperoxid Reduktase (At1g65980.1 und At3g52960.1) durch eine erhöhte Proteinexpression auf. Im Gegensatz dazu war die Peroxidase At3g49120.1 bei 32°C deutlich herabreguliert (Tabelle 4.2.2.2).

Eine Auswahl der mit Hilfe der Varianzanalyse identifizierten Proteinmarker, deren Expressionen in den untersuchten Genotypen Col-0 und PGM bedingt durch Temperaturstress spezifisch reguliert wurden, ist in Abbildung 4.2.2.2 zusammengestellt.



**Abbildung 4.2.2.2.: Auswahl spezifischer Markerproteine für Temperaturstress in *A. thaliana* Col-0 und PGM Pflanzen.** Dargestellt sind die auf der Analyse von 6 biologischen Replikate beruhenden Expressionsunterschiede in beiden Genotypen. Während die Expression des Glycinreichen RNA-bindenden Proteins At2g21660.1 linear mit steigender Temperatur abnahm, konnte das als 'cold-responsive/cold-regulated' annotierte Protein At5g15970.1 aufgrund der verstärkten Expression bei 4°C als Marker-Protein kältegestresster PGM Pflanzen identifiziert werden. Eine Wildtyp-spezifische Reaktion auf Hitzestress wies die bei 32°C deutlich hochregulierte Dehydroascorbat Reduktase At1g19570.1 auf.

### 4.3 Metabolit Profiling temperaturgestresster *A. thaliana* Col-0 und PGM Pflanzen

Um neben den im vorangegangenen Kapitel aufgeführten potenziellen Proteinmarkern für temperaturbedingten Stress auch Metabolite identifizieren zu können, welche ebenso als Stressmarker dieser komplementären Stoffwechselebene dienen können, wurden ausgehend von gleichem pflanzlichen Untersuchungsmaterial die kleinen Moleküle mittels GC-TOF-MS analysiert. Daraus resultierte eine Datenmatrix, welche die relativen Konzentrationen von 80 polaren Metaboliten umfasste.

### 4.3.1 Chemometrischen Analyse der Metabolitdaten

Zur Analyse der Metabolitdaten wurden diese, wie auch die Proteindaten, dem Kruskal-Wallis-Tests unterzogen, welcher ergab, dass 61 Variablen bzw. Metabolite signifikante Konzentrationsunterschiede ( $p \leq 0.05$ ) zwischen gestressten und ungestressten Pflanzen beider Genotypen aufwiesen. Damit beliefen sich die nachweisbaren Auswirkungen von Temperaturstress auf 75% der im Rahmen dieser Analyse erfassten Repräsentanten des pflanzlichen Stoffwechsels. Dieses Ergebnis zeigt, dass, wie auch von Cook *et al.* [266]. oder Kaplan *et al.* [267] beschrieben, die Auswirkungen von Temperaturstress nahezu systemische Veränderungen im Metabolom der Pflanze hervorrufen. Die betreffenden Metabolite sind in Tabelle 4.3.1.1 aufgeschlüsselt.

**Tabelle 4.3.1.1: Einfluss von temperaturbedingtem abiotischen Stress auf das Metabolom von *A. thaliana* Col-0 und PGM.** Angegeben sind relative Änderungen zu den Referenzwerten der bei 20°C kultivierten Kontrollpflanzen beider Genotypen (X-fold) mit Irrtumswahrscheinlichkeit (p-Wert).

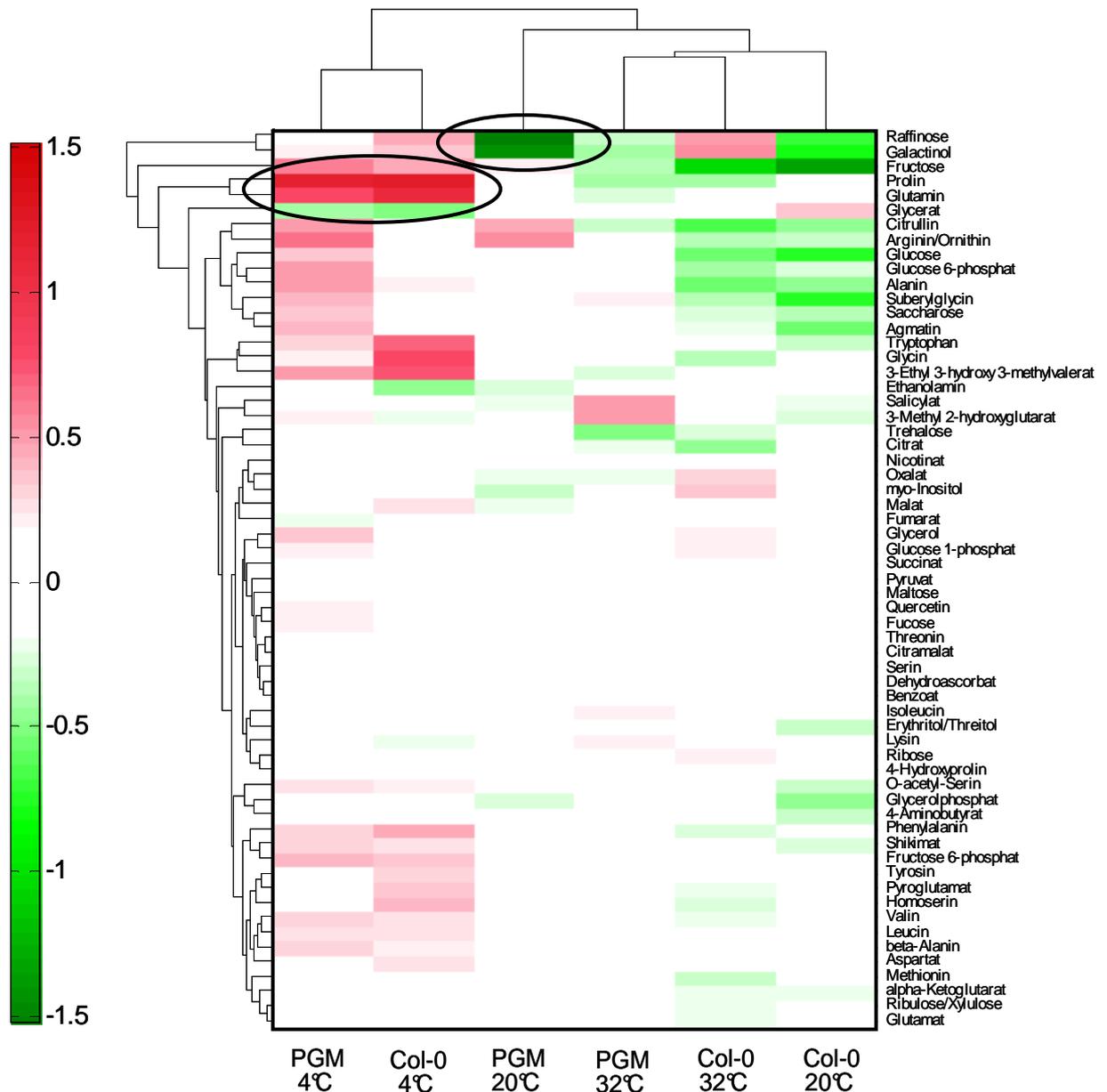
	Col-0 4°C X-fold ( p-Wert)	Col-0 32°C X-fold ( p-Wert)	PGM 4°C X-fold ( p-Wert)	PGM 32°C X-fold ( p-Wert)
<b>Response auf Kälte- und Hitzestress</b>				
Raffinose	13.6 ( 0.006 )	16.7 ( 0.004 )	52.5 ( 0.004 )	17.2 ( 0.004 )
Galactinol	13.1 ( 0.006 )	21.0 ( 0.004 )	41.6 ( 0.004 )	10.4 ( 0.004 )
4-Aminobutyrat	3.1 ( 0.006 )	2.1 ( 0.004 )	1.5 ( 0.004 )	2.0 ( 0.004 )
Glycerolphosphat	4.0 ( 0.006 )	3.3 ( 0.004 )	1.9 ( 0.006 )	1.9 ( 0.010 )
O-Acetylserin	3.7 ( 0.006 )	1.8 ( 0.004 )	2.2 ( 0.004 )	1.4 ( 0.004 )
Glycerat	-6.8 ( 0.006 )	-1.7 ( 0.004 )	-3.4 ( 0.004 )	-1.9 ( 0.004 )
Prolin	17.6 ( 0.006 )	-2.4 ( 0.010 )	14.3 ( 0.004 )	-2.8 ( 0.004 )
Phenylalanin	3.9 ( 0.006 )	-1.4 ( 0.010 )	1.6 ( 0.004 )	-1.6 ( 0.004 )
beta-Alanin	1.7 ( 0.028 )	-1.3 ( 0.010 )	1.9 ( 0.004 )	-1.3 ( 0.004 )
Aspartat	1.7 ( 0.006 )	-1.5 ( 0.016 )	1.7 ( 0.004 )	-1.2 ( 0.025 )
Lysin	-1.3 ( 0.028 )	1.4 ( 0.016 )	-1.6 ( 0.004 )	1.2 ( 0.037 )
Fructose	64.9 ( 0.006 )	1.9 ( 0.006 )	2.4 ( 0.004 )	-3.8 ( 0.004 )
Glucose	8.7 ( 0.006 )	1.6 ( 0.006 )	2.0 ( 0.004 )	-1.6 ( 0.004 )
Fructose 6-phosphat	3.4 ( 0.006 )	1.3 ( 0.037 )	2.2 ( 0.004 )	-1.7 ( 0.006 )
Shikimat	3.4 ( 0.006 )	1.3 ( 0.004 )	1.7 ( 0.004 )	-1.7 ( 0.004 )
Ribose	-1.5 ( 0.045 )	1.5 ( 0.004 )		
<b>Response auf Kältestress</b>				
Tryptophan	10.6 ( 0.006 )		2.4 ( 0.010 )	
Alanin	4.9 ( 0.011 )		3.1 ( 0.004 )	
Methionin	2.0 ( 0.006 )		1.9 ( 0.004 )	
Salicylat	1.9 ( 0.011 )		2.2 ( 0.004 )	
Quercetin	1.4 ( 0.045 )		1.7 ( 0.010 )	
Tyrosin	3.1 ( 0.006 )			
alpha-Ketoglutarat	2.4 ( 0.006 )			
Ethanolamin	-2.9 ( 0.006 )			
Dehydroascorbat			1.6 ( 0.004 )	
Benzoat			1.3 ( 0.025 )	
Citramalat			1.3 ( 0.010 )	

**Tabelle 4.3.1.1: Einfluss von temperaturbedingtem abiotischen Stress auf das Metabolom von *A. thaliana* Col-0 und PGM.** Angegeben sind relative Änderungen zu den Referenzwerten der bei 20°C kultivierten Kontrollpflanzen beider Genotypen (X-fold) mit Irrtumswahrscheinlichkeit (p-Wert).

	Col-0 4°C X-fold ( p-Wert)	Col-0 32°C X-fold ( p-Wert)	PGM 4°C X-fold ( p-Wert)	PGM 32°C X-fold ( p-Wert)
<b>Response auf Hitzestress</b>				
Oxalat		1.8 ( 0.037 )		
4-Hydroxyprolin		1.8 ( 0.004 )		
Trehalose				-3.9 ( 0.016 )
Maltose				-1.3 ( 0.025 )
<b>Differentielle Response</b>				
Suberylglycin	4.7 ( 0.006 )	2.5 ( 0.004 )	2.1 ( 0.004 )	
Agmatin	2.7 ( 0.011 )	2.2 ( 0.010 )	2.1 ( 0.004 )	
Saccharose	1.9 ( 0.006 )	1.2 ( 0.016 )	1.7 ( 0.004 )	
Glycin	7.5 ( 0.006 )	-1.9 ( 0.004 )	1.5 ( 0.004 )	
Valin	2.4 ( 0.006 )	-1.2 ( 0.006 )	2.3 ( 0.004 )	
Leucin	1.9 ( 0.011 )	-1.3 ( 0.004 )	2.0 ( 0.004 )	
Fumarat	1.3 ( 0.011 )	1.7 ( 0.004 )		2.0 ( 0.004 )
Citrullin	4.2 ( 0.006 )	-1.7 ( 0.037 )		-6.1 ( 0.004 )
Homoserin	2.6 ( 0.006 )	-1.8 ( 0.004 )		-1.3 ( 0.016 )
Glutamat	1.6 ( 0.006 )	-1.2 ( 0.037 )		-1.2 ( 0.016 )
Glutamin	14.2 ( 0.006 )		5.4 ( 0.004 )	-2.0 ( 0.037 )
3-Ethyl 3-hydroxy 3-methylvalerat	6.1 ( 0.006 )		2.9 ( 0.004 )	-1.9 ( 0.004 )
Pyroglutamat	3.4 ( 0.006 )		1.3 ( 0.025 )	-1.3 ( 0.006 )
Glucose 6-phosphat	2.8 ( 0.006 )		2.4 ( 0.004 )	-2.1 ( 0.010 )
Malat	1.8 ( 0.006 )		1.7 ( 0.004 )	1.4 ( 0.016 )
Nicotinat	-1.3 ( 0.006 )		1.2 ( 0.025 )	-1.2 ( 0.025 )
<b>Differentielle Response</b>				
Erythritol/Threitol		2.3 ( 0.004 )	1.4 ( 0.010 )	1.8 ( 0.004 )
3-Methyl 2-hydroxyglutarat		1.9 ( 0.004 )	1.7 ( 0.025 )	3.4 ( 0.004 )
myo-Inositol		1.9 ( 0.004 )	1.6 ( 0.037 )	1.7 ( 0.025 )
Isoleucin		1.2 ( 0.006 )	1.5 ( 0.010 )	2.2 ( 0.004 )
Citrat		-2.8 ( 0.004 )	-1.3 ( 0.037 )	-2.1 ( 0.004 )
Pyruvat		-1.4 ( 0.025 )	1.5 ( 0.010 )	-1.5 ( 0.025 )
Glycerol		1.7 ( 0.004 )	2.9 ( 0.004 )	
Glucose 1-phosphat		1.6 ( 0.004 )	1.6 ( 0.004 )	
Fucose		1.3 ( 0.004 )	1.4 ( 0.004 )	
Succinat		-1.6 ( 0.004 )	-1.7 ( 0.004 )	
Serin		-1.2 ( 0.037 )	1.3 ( 0.025 )	
Arginin/Omithin	2.2 ( 0.006 )			-4.0 ( 0.004 )
Ribulose/Xylulose	1.8 ( 0.006 )			-1.2 ( 0.025 )
Threonin	1.2 ( 0.045 )			-1.1 ( 0.006 )

Nach einer auf den reduzierten Datensatz angewandten Clusteranalyse wurden die multifaktoriellen, durch Temperaturstress hervorgerufenen Unterschiede in den Metabolprofilen visualisiert (Abbildung 4.3.1.1). Die Klasseneinteilung der experimentellen Bedingungen macht dabei deutlich, dass das Stoffwechselgeschehen insbesondere durch Erniedrigung der Temperatur erheblich beeinflusst wurde.

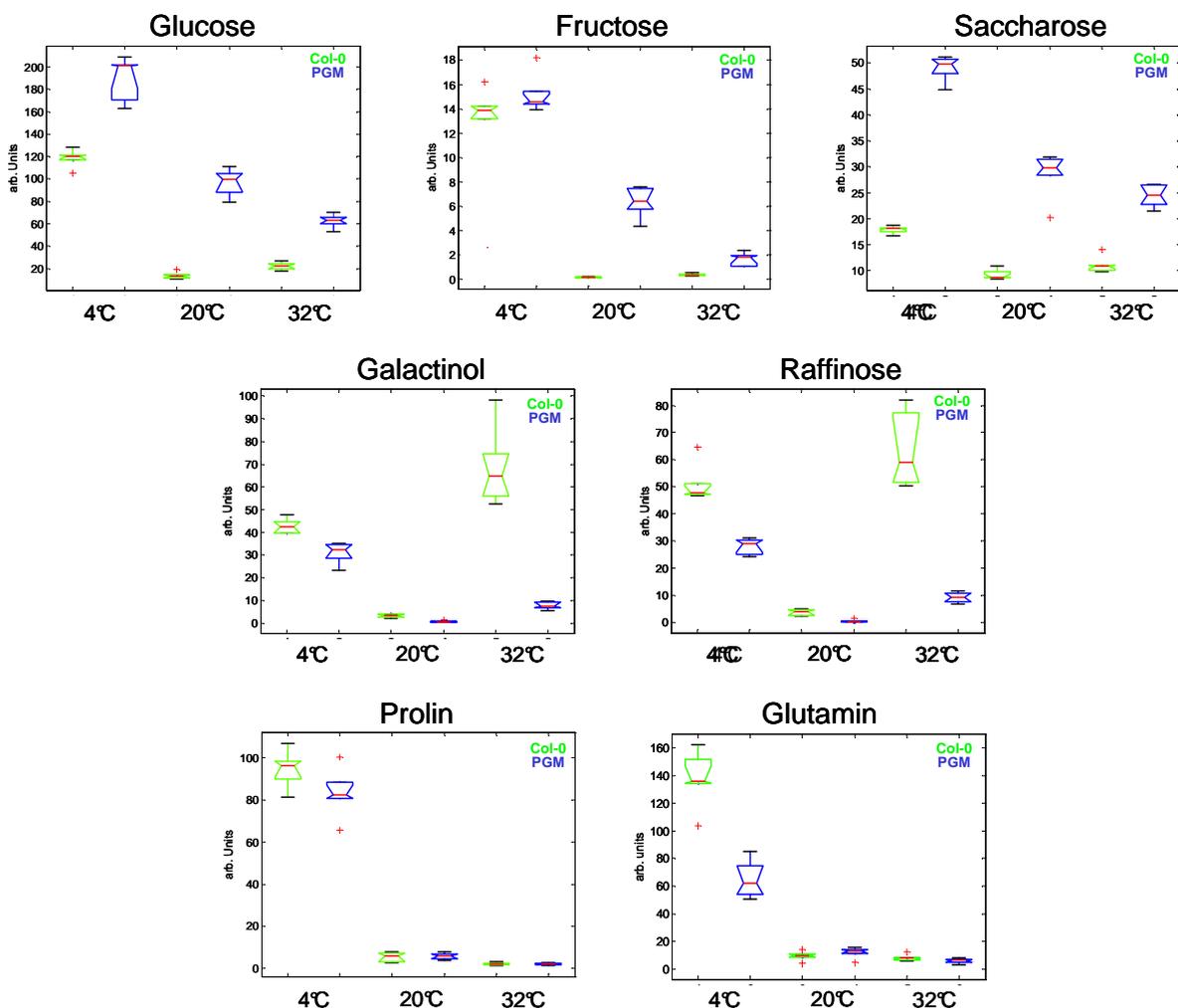
In beiden Genotypen resultierte Kältestress in einer deutlichen Akkumulation zahlreicher Verbindungen, vorrangig solcher mit kryoprotektiven Eigenschaften. Erste Beobachtungen des Phänomens der Speicherung löslicher Zucker nach Temperaturniedrigung gehen bereits auf das letzte Jahrhundert zurück [268]. Von besonderem Interesse hier waren



**Abbildung 4.3.1.1: Bi-Clustering und farbcodierte Darstellung relativer Änderungen ermittelter Metabolitkonzentrationen in *A. thaliana* Col-0 und PGM Pflanzen nach abiotischem Temperaturstress.** Das Bi-Clustering erfolgte mit dem Pearson'schen Produktmoment-Korrelationskoeffizienten als Abstandsmass, die Farbcodierung erfolgte nach  $\log_{10}$ -Transformation der auf die Kontrolle (Col-0, 20°C) normierten relativen Metabolitkonzentrationen, wobei erhöhte Konzentrationen rot, verminderte Konzentrationen grün dargestellt sind. Obgleich diese Darstellung nahezu systemische Veränderungen der Metabolitprofile nach Temperaturstress erkennen lässt, können zwei interessante Cluster erkannt werden. Während Galactinol und Raffinose in beiden Genotypen stark korrelieren, fallen die Aminosäuren Prolin und Glutamin durch deutlich erhöhte Konzentrationen in den kältegestressten Pflanzen auf.

daher die temperaturbedingten Veränderungen der Metabolitgehalte in der stärkelosen, zuckerakkumulierenden PGM Mutante. Tatsächlich ergab sich verbunden mit der Akklimatisierung an 4°C ein weiterer Anstieg der ohnehin im Vergleich zum Wildtyp stark erhöhten Konzentrationen von Glucose, Fructose und Saccharose, wenngleich die kälteinduzierten Stressreaktionen bezüglich dieser Metabolite im Wildtyp stärker ausgeprägt waren (Abbildung 4.3.1.2). Des Weiteren zeichnete sich für die PGM Mutante eine Abnahme dieser Metabolitgehalte mit steigender Temperatur ab.

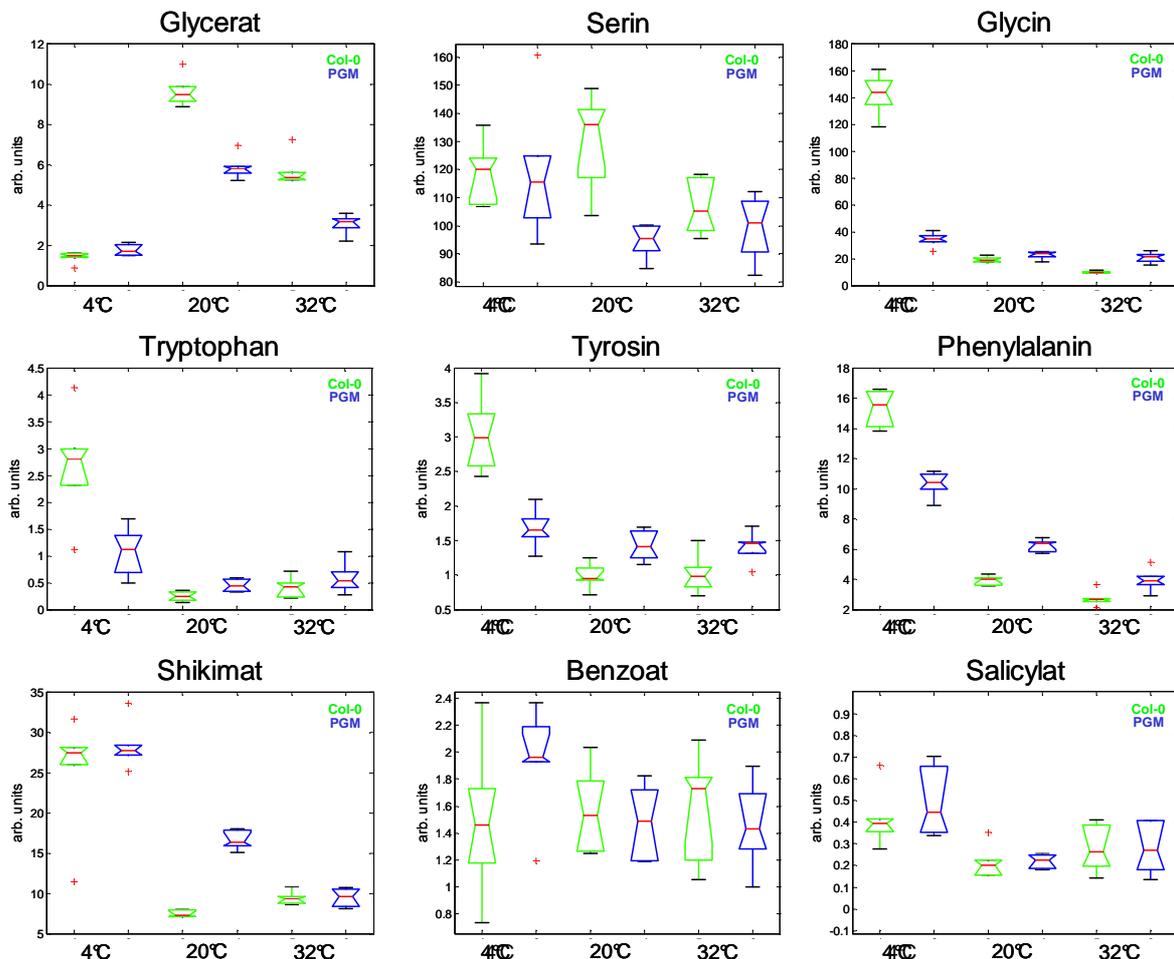
Darüber hinaus offenbarte die farbcodierte Darstellung einen deutlichen Anstieg der Konzentrationen der Galactoside Raffinose und Galactinol in den Blattgeweben kälte-



**Abbildung 4.3.1.2: Konzentrationsprofile ausgewählter Metabolite in *A. thaliana* Col-0 und PGM Pflanzen nach Temperaturstress.** Kältestress induzierte sowohl im Wildtyp, als auch in der stärkelosen PGM Mutante eine deutliche Akkumulation der löslichen Carbohydrate Glucose, Fructose und Saccharose. Die Konzentrationsprofile der Galactoside Raffinose und Galactinol lassen eine Stressreaktion sowohl auf Kälte-, als auch auf Hitzestress in beiden Genotypen erkennen. Zudem resultierte Kältestress in einer deutlichen Akkumulation von Prolin und Glutamin in den Blattgeweben der analysierten *A. thaliana* Genotypen Col-0 und PGM.

gestresster Pflanzen. Interessanterweise wiesen auch die Profile der hitzestressierten Pflanzen, und hierbei insbesondere die des Wildtyps, eine deutliche Akkumulation dieser Metabolite auf (Abbildungen 4.3.1.1. und 4.3.1.2). Im Hinblick auf die durch Kältestress induzierten Veränderungen zeigten auch die Aminosäuren Prolin und Glutamin eine starke Response. Die Konzentrationen beider Aminosäuren stiegen in den Blattgeweben der bei 4°C kultivierten Pflanzen beider Genotypen signifikant an (Abbildung 4.3.1.2).

Bezüglich der an der Photorespiration beteiligten Intermediate ergab sich eine differentielle Stressresponse. Während sich die Stressreaktionen in beiden Genotypen in einer deutlichen Herabregulation von Glycerat manifestierten, blieben die Konzentrationen von Serin nahezu unverändert. Glycin hingegen wurde unter Kältestress im Wildtyp stark akkumuliert (Abbildung 4.3.1.3).

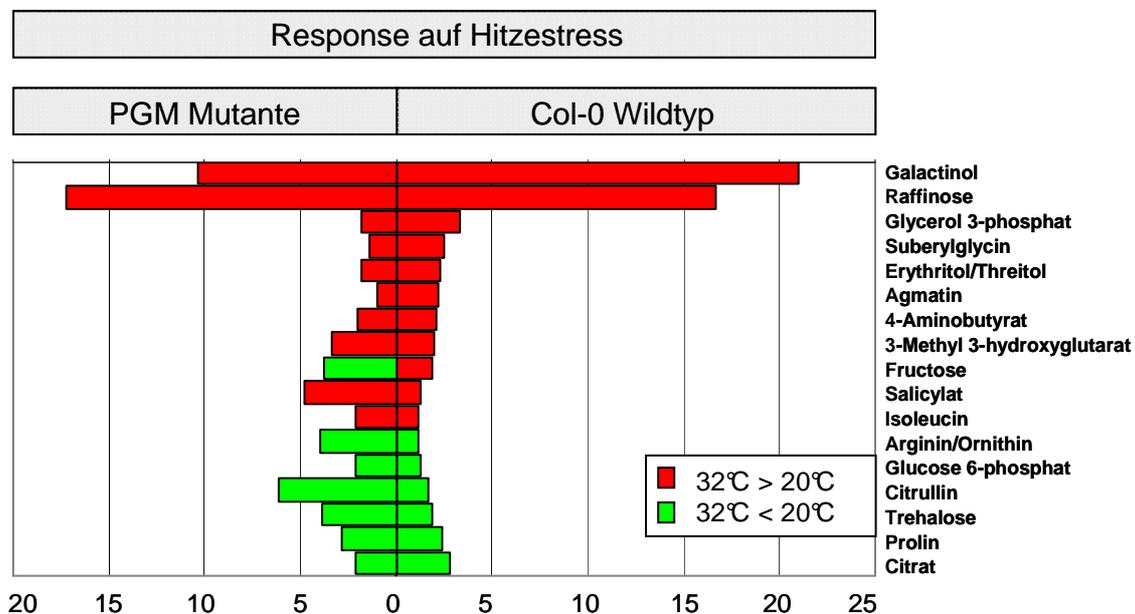


**Abbildung 4.3.1.3: Einfluss der Temperatur auf ausgewählte Stoffwechselwege in *A. thaliana*.** Während die Response der Intermediate des Photorespiratorischen Cyclus keinen eindeutigen Trend aufwiesen, induzierte Kältestress in Col-0 und PGM eine Hochregulierung der Biosynthese aromatischer Aminosäuren. Diese wirkte sich jedoch nicht auf die Sekundärmetabolite Benzoat und Salicylat aus.

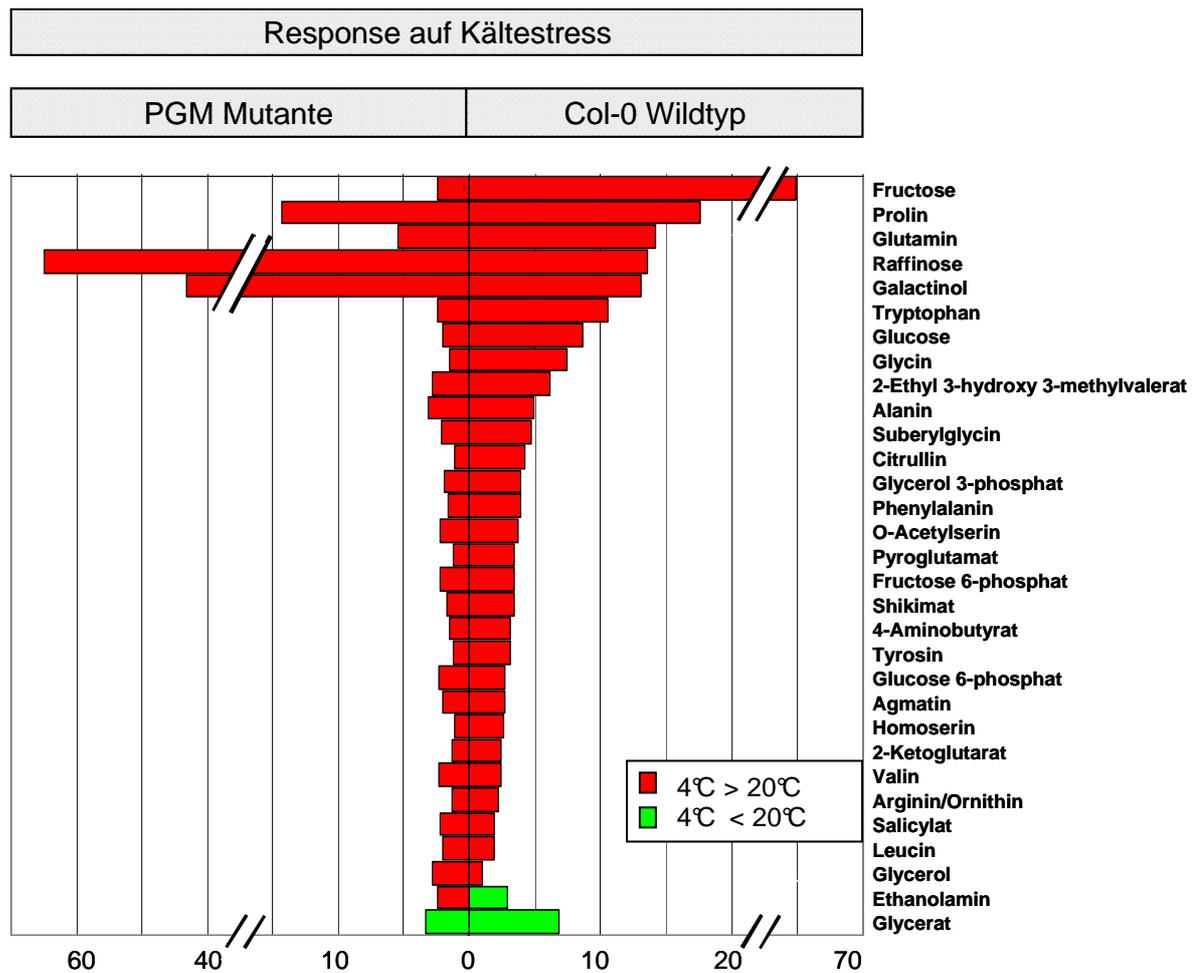
Daneben liessen die Konzentrationsprofile von Tryptophan, Tyrosin und Phenylalanin auf eine Kältestress-induzierte Hochregulation des Syntheseweges aromatischer Aminosäuren schliessen. Auch die für Shikimat in den Blattgeweben kältegestresster Pflanzen beobachtete Akkumulation indiziert eine verstärkte Synthese dieser Stoffklasse. Eine Auswirkung auf die Konzentrationen der Sekundärmetabolite Benzoat und Salicylat konnte jedoch nicht festgestellt werden (Abbildung 4.3.1.3). Auch für den Precursor der Indol 3-essigsäure, Indol 3-acetonitril, erbrachte die Analyse ähnliche Konzentrationen in den gestressten und ungestressten Pflanzen (Daten nicht gezeigt).

Zusammenfassend offenbarte die vergleichende Analyse der Temperaturstressresponse, dass durch Temperaturerniedrigung der pflanzliche Metabolismus stärker beeinflusst wurde, als durch Temperaturerhöhung. Daneben konnten einige Metabolite mit spezifischer Response auf die einzelnen Stressbedingungen sowie Genotypspezifitäten identifiziert werden. Einen Interessanten Aspekt stellte die Beobachtung einer gemeinsamen Response auf Kälte- und Hitzestress dar, welche für einen grossen Teil der Metabolite nachweisbar war.

Die Abbildungen 4.3.1.4 und 4.3.1.5 fassen die deutlichsten metabolischen Veränderungen im Zusammenhang mit abiotischem Temperaturstress zusammen.



**Abbildung 4.3.1.4: Deutlichste, durch Hitzestress induzierte metabolische Stressantworten in *A. thaliana* Col-0 und PGM.** Angegeben sind relative Änderungen zu den bei 20°C kultivierten Kontrollpflanzen beider Genotypen, wobei relative Konzentrationserhöhungen rot, relative Erniedrigungen der Konzentrationen ausgewählter Metabolite grün dargestellt sind.



**Abbildung 4.3.1.5: Deutlichste, durch Kältestress induzierte metabolische Stressantworten in *A. thaliana* Col-0 und PGM.** Angegeben sind relative Änderungen zu den bei 20°C kultivierten Kontrollpflanzen beider Genotypen, wobei relative Konzentrationserhöhungen rot, relative Erniedrigungen der Konzentrationen ausgewählter Metabolite grün dargestellt sind.

#### 4.4 Diskussion

Durch die jüngeren Entwicklungen auf dem Gebiet der Gensequenzierung und die Verfügbarkeit von Techniken zur Analyse der globalen Genexpression wurde die Response von Pflanzen auf verschiedenste Stressbedingungen auf Genexpressions-ebene ausführlich analysiert. Diese Analysen erbrachten, dass, wenngleich biotischer und abiotischer Stress die Pflanze durch individuelle Mechanismen schädigen, die Stressreaktionen zumeist sehr ähnlich sind [269-272]. So resultieren sowohl biotischer, als auch abiotischer Stress u.a. in der Bildung freier Radikale, welche die Zelle in mannigfaltiger Weise schädigen können [273,274]. Wenngleich der Einfluss diverser abiotischer Stressoren auf die Transkription insbesondere in *Arabidopsis* mehrfach

analysiert wurde [270,272,275-277], ist bis dato über die Auswirkungen von abiotischem Stress auf den Protein- oder Metabolithaushalt der Pflanze nur wenig bekannt. Die zumeist auf Microarrays beruhenden Analysen der Stressresponse erbrachten übereinstimmend, dass mehrere hundert Gene in die zelluläre Stressantwort involviert sind. Daher gilt es heute als evident, dass die Akklimatisierung an extreme Temperaturen aus einem komplexen Zusammenspiel zahlreicher physiologischer und biochemischer Veränderungen des Primär- und Sekundärmetabolismus resultiert [278,279].

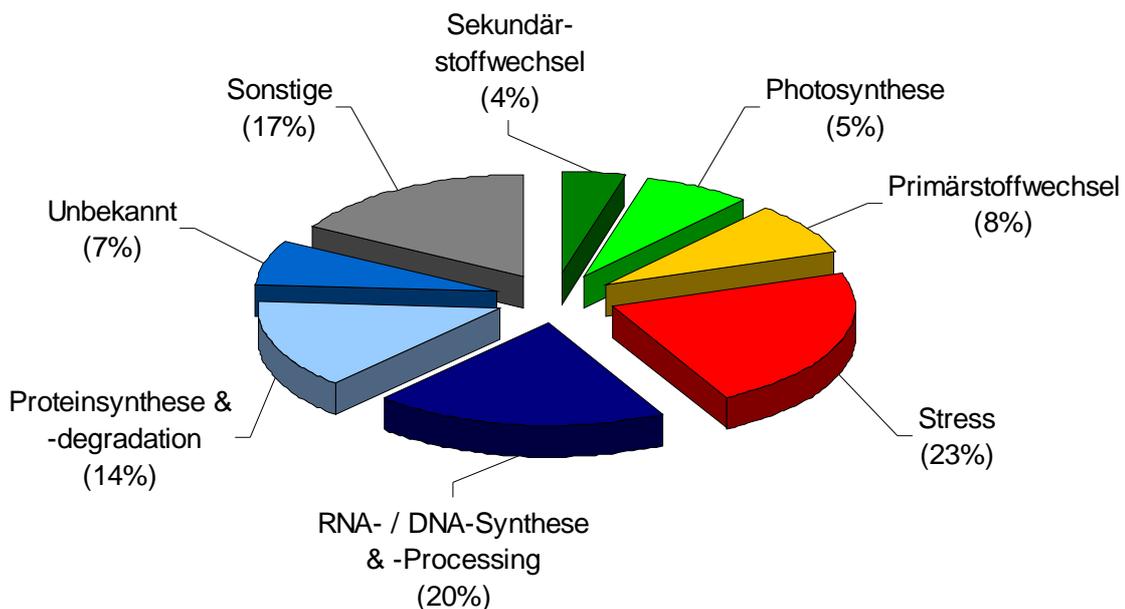
Im Rahmen dieser Arbeit sollten mit der Analyse der Temperaturstressresponse auf Metabolit- und Proteinebene zwei Fragestellungen untersucht werden. Die Analyse der diurnalen Veränderungen erbrachte, wie im vorangegangenen Kapitel 3 zur Untersuchung der mit der circadianen Rhythmik verbundenen Änderungen im pflanzlichen Stoffwechsel diskutiert, dass die Expression einzelner Proteine im Verlaufe des diurnalen Rhythmus nur schwach oszilliert. Daher sollte untersucht werden, welchen Einfluss die Temperatur auf die Proteinexpression ausübt.

Zudem sind die kryoprotektiven Eigenschaften von Carbohydraten literaturbekannt [268,270,276,277]. Durch die vergleichende Analyse der *Arabidopsis thaliana* Genotypen Col-0 und PGM sollte weiter untersucht werden, ob die zuckerakkumulierende PGM Mutante durch eine konstitutiv höhere Toleranz gegenüber Temperaturstress gekennzeichnet ist.

Bei der Analyse der zeitabhängigen Regulation des Stoffwechsels der untersuchten *Arabidopsis* Spezies wurden Proteine relativ über die Integration der Signale tryptischer Peptide quantifiziert. Diese Methode erwies sich jedoch zum einen als sehr zeitaufwendig, da die Ionenspuren der Vorläuferpeptide manuell aus den MS-SurveyScans extrahiert und anschliessend integriert werden mussten, zum anderen stellte sich die Anzahl der nach diesem Vorgehen quantifizierbaren Proteine als relativ gering dar, was die Interpretation des Einflusses des Lichts auf den Stoffwechsel von Col-0 und PGM erheblich erschwerte. Mit einer neu entwickelten Methode zur Extraktion löslicher Proteine sowie der relativen Quantifizierung über der Spektren-Count konnten im Rahmen dieses Versuches die Dimensionalität der resultierenden Protein-Datenmatrix erheblich erhöht werden. So gelang mit diesem Vorgehen die relative Quantifizierung von insgesamt 179 Proteinen. Der Spektren-Count, welcher der kumulativen Anzahl der nach der MS<sup>2</sup>-Fragmentierung eines Vorläuferpeptides detektierten Fragmentspektren eines jeden Proteins entspricht, ist der Abundanz eines Proteins in sehr guter Näherung proportional, wie Liu *et al.* (2004) zeigen konnten [280]. Gleichzeitig konnten mittels GC-TOF-MS 80 strukturbekannt polare Metabolite in allen Proben relativ quantifiziert werden.

### **Temperaturstress induziert globale Veränderungen im Proteom und Metabolom der Pflanze**

Die vergleichende Analyse der gewonnenen Daten erbrachte, dass die deutlichsten Veränderungen nach 3-tägiger Akklimation an 4°C bzw. 32°C im Metabolom der beiden analysierten Genotypen, Col-0 sowie PGM, nachweisbar waren; für etwa 75% der analysierten Metabolite erbrachte die Varianzanalyse signifikante Unterschiede zwischen den Blattkonzentrationen gestresster und ungestresster Pflanzen (Tabelle 4.3.1.1 und Abbildung 4.3.1.1). Auf Proteinebene belief sich die Temperaturstressresponse auf ca. 25% der detektierten Proteine (Tabellen 4.2.2.1 und 4.2.2.2 sowie Abbildung 4.2.2.1). Wenngleich die Anzahl der nach Temperaturstress differentiell exprimierten Proteine unerwartet gering ausfiel, konnte für viele der gemäss Annotation in der Response auf Temperaturstress involvierten Proteine eine differentielle Expression nach 72-stündiger Akklimation an 4°C bzw. 32°C (Abbildung 4.4.1) nachgewiesen werden, was die Eignung der angewandten Methode der instrumentellen Analytik, sowie der datenanalytischen Verfahren bestätigt.

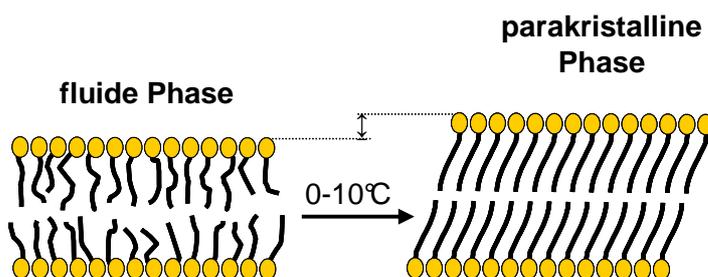


**Abbildung 4.4.1: Funktionelle Klassifizierung der durch Temperaturstress induzierten Proteinexpression.** Dargestellt sind Proteine mit differentiellen Expressionslevel als Response auf Temperaturstress nach ihrer Klassifikation gemäss MAPMAN-Bin-code (<http://gabi.rzpd.de/projects/MapMan/>).

**Der Grad der Kältetoleranz wird durch die Funktionalität und Fluidität der Biomembranen bestimmt**

Die Empfindlichkeit von Pflanzen gegenüber Kälte liegt vor allem in der funktionellen Störung bestimmter Membranfunktionen begründet. Bei niedrigen Temperaturen, meist im Bereich von 0°C bis 10°C, unterliegen Membranlipide einem Phasenübergang von der fluiden oder auch flüssigkristallinen Phase, welche durch ein hohes Mass an Beweglichkeit der einzelnen Lipidmoleküle gekennzeichnet ist, in die Gel- oder auch parakristalline Phase, in welcher die Lipide nahezu unbeweglich sind (Abbildung 4.4.2). Die Temperatur dieses sog. Hauptphasenübergangs hängt von der Lipidzusammensetzung ab und ist für jede Membran charakteristisch. Die abrupte Änderung in der Permeabilität von Membranen sowie der Aktivität von Membranenzymen bei Unterschreiten der Phasenübergangstemperatur führt häufig zu einem Ausströmen von Ionen und Metaboliten, so dass Membranläsionen zum Verlust der Selektivität von Permeationsvorgängen und zur Aufhebung der Kompartimentierung der Zelle führen [281-285]. Gleichzeitig nimmt bei sinkender Temperatur die Viskosität des Wassers zu und der hydroaktive Stomatenschluss wird verhindert oder findet nur noch stark verzögert statt [286-289], was letztendlich trotz optimaler Wasserversorgung der Pflanze dazu führt, dass sich Trockenstress als Sekundärstress bei Kältestress einstellt.

Die Erhöhung der Synthese sowohl von gesättigten, als auch von ungesättigten Fettsäuren, welche den Membranen auch bei niedrigen Temperaturen eine gewisse Fluidität verleihen, ist als Reaktion von Pflanzen auf Kältestress gut dokumentiert. Die Synthese der Lipide erfolgt in Pflanzen über zwei unterschiedliche Biosynthesewege. Neben dem prokaryotischen Pfad der Synthese in den Plastiden können diese auch im Endoplasmatischen Retikulum über einen eukaryotischen Syntheseweg gebildet werden [290]. Während Palmitin- und Stearinsäure beispielsweise in den chloroplastidiären Kompartimenten synthetisiert werden, wobei die Kettenverlängerung an kovalent gebundenen Acyl-Carrier Proteinen erfolgt, wird Ölsäure durch die plastidiäre Acyl-Lipid Desaturase gebildet und kann nach dem Export zum endoplasmatischen Retikulum durch die mikrosomale Acyl-Lipid Desaturase weiter desaturiert werden [291-293].



**Abbildung 4.4.2: Hauptphasenübergang der Membranlipide.** Während die Lipidmoleküle in der fluiden Phase beweglich sind, zeichnet sich die sich bei niedrigen Temperaturen einstellende parakristalline Phase durch ein hohes Mass an Ordnung aus. Die Lipidmoleküle sind hier nahezu unbeweglich.

Im Zusammenhang mit dem Lipidstoffwechsel erbrachte die Proteinanalyse, dass das in der Synthese und Elongation von Fettsäuren involvierte Acyl-Carrier-Protein At1g54630.1 in den Blättern kältegestresster Pflanzen, jedoch nicht in denen hitzegepresster Pflanzen detektiert wurde (Tabelle 4.2.2.1). Wenngleich der Grad der Sättigung der Fettsäuren nach Kältestress mit den angewandten analytischen Methoden nicht bestimmt werden konnte und die Quantifizierung über den Spektren-Count keine Aussagen über die Aktivität dieses Enzyms zulassen, kann dieses Ergebnis dennoch als Hinweis auf die verstärkte Synthese von gesättigten und ungesättigten Fettsäuren, welche durch die Erniedrigung der Phasenübergangstemperatur wesentlich zur Aufrechterhaltung der Membranfunktionen beitragen, in den kältegestressten Pflanzen gedeutet werden.

***Citrat ist ein potenzieller metabolischer Marker der Biosynthese von Fettsäuren im cytosolischen Kompartiment der Pflanze***

Die wichtigste Ausgangsverbindung für die Synthese von Fettsäuren stellt das Acetyl-CoA dar. Da dieser Acylgruppendonator nicht membrangängig ist und die Synthese und Prozessierung der Fettsäuren in unterschiedlichen Kompartimenten der pflanzlichen Zelle erfolgt, muss auch das Acetyl-CoA kompartimentspezifisch generiert werden. In den Chloroplasten kann die Biogenese des Acetyl-CoA's ausgehend von Acetat oder von Pyruvat erfolgen. Cytosolisch wird derzeit die von der ATP-Citrat Lyase (ACL) katalysierte Biogenese von Acetyl-CoA diskutiert [294,295]. Jüngste Untersuchungen an transgenen *Arabidopsis* Pflanzen deuten weiter darauf hin, dass die ATP-abhängige Spaltung von Citrat die einzige Acetyl-CoA erzeugende Reaktion im Cytosol darstellt [295].

Von dieser Annahme ausgehend indizieren auch die Metabolitprofile eine verstärkte Fettsäuresynthese in den kältegestressten Pflanzen. Während die Konzentrationen von Malat im Vergleich zu den Kontrollpflanzen nach Akklimatisierung an 4°C in beiden Genotypen signifikant erhöht waren, konnte eine Erhöhung der Citratkonzentrationen nicht nachgewiesen werden. Möglicherweise wurde unter Kältestress ein erheblicher Teil des mitochondrial synthetisierten Citrats durch einen von Fatland *et al.* (2002) postulierten Citrat-Shunt [295] ins Cytosol geschleust, dort ACL-katalysiert nicht-hydrolytisch gespalten und somit das zur Elongation und Synthese von mehrfach ungesättigten Fettsäuren essentielle Acetyl-CoA bereitgestellt. Gleichwohl kann eine Inhibierung der Citrat Synthetase oder die oxidative Decarboxylierung von Pyruvat als Ursache für die im Vergleich zu Malat geringen Citratkonzentration ebenso wie eine Hochregulation der anaplerotischen Reaktionen nicht ausgeschlossen werden. Hierzu wären Untersuchungen des metabolischen Flusses nach Inkubation mit beispielsweise <sup>13</sup>C-markierter Glucose oder ähnlichen Carbohydratverbindungen nötig.

**Die Akkumulation kryoprotektiver Substanzen wird nicht über die Konzentrationen löslicher Carbohydrate reguliert**

Neben der verstärkten Synthese ungesättigter Fettsäuren gilt die Akkumulation löslicher Carbohydratverbindungen als eine der deutlichsten Reaktionen auf Kältestress. Durch die bei plötzlicher starker Abkühlung unter den Gefrierpunkt häufig auftretende Bildung von Eiskristallen können Cytoplasmamembranen irreversible Schäden erleiden, welche letztlich zu einem Absterben der Zelle führen. Bei langsamer Abkühlung erfolgt die Eisbildung vorrangig in den Interzellularräumen, was die Membranen schont, allerdings aufgrund des niedrigen Wasserdampfdruck des Eises in einer enormen Dehydratisierung resultiert. Dieser Wasserverlust bedingt einen deutlichen Anstieg der Konzentrationen zellulärer Lösungen. Die dabei auftretenden hohen Konzentrationen insbesondere von Salzen können in einer Inaktivierung von Enzymen resultieren, wodurch sich das Membranpotenzial verschiebt und die Hydratisierung der Zellmembranen dramatisch verschlechtert [296-298]. Zur Vermeidung von Gefrierschäden akkumulieren Pflanzen verschiedene Gefrierschutzstoffe, bei denen es sich vorrangig um Carbohydratverbindungen handelt. Diese erfüllen zwei wichtige Funktionen: zum einen resultiert die Akkumulation dieser kryoprotektiven Substanzen in einem Anstieg des osmotischen Potenzials und damit verbunden in einer Erhöhung der Wasserrückhaltekapazität, andererseits kommt es zu einer Verdünnung der Ionenkonzentrationen an den Oberflächen der Membranen, wodurch die Hauptphasenübergangstemperatur erniedrigt wird [299]. Ein weiterer stabilisierender Effekt resultiert aus der Erhaltung der Struktur und Funktion labiler Proteine. Aufgrund ihrer reichlich vorhandenen Hydroxylgruppen tragen Carbohydrate wesentlich dazu bei, eine geordnete Hydrathülle an der Oberfläche von Proteinen aufrecht zu erhalten und somit konformationsschädigende Wechselwirkungen von Proteinen mit ihrer Umgebung zu minimieren. Somit kommt diesen kryoprotektiven Verbindungen eine Bedeutung zu, die denen der Chaperonine gleicht [300,301].

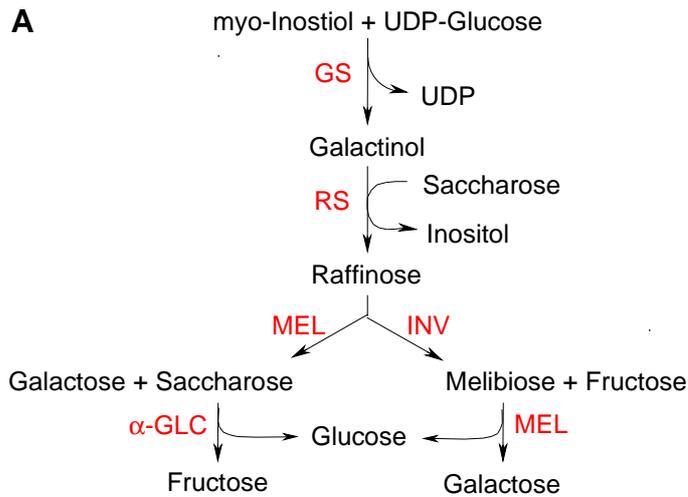
In Übereinstimmung zu diesen Mechanismen zur Erhöhung der Stresstoleranz erbrachte die Analyse der Metabolitdaten signifikant erhöhte Konzentrationen an Glucose, Fructose und Saccharose in den kältegestressten Blättern beider untersuchten Genotypen. Auch die Blattkonzentrationen der zuckerakkumulierenden PGM Mutante stiegen nach Kältestress im Vergleich zu den Kontrollpflanzen weiter signifikant an (Abbildung 4.3.1.2). Dieses Ergebnis zeigt, dass die Signaltransduktion bezüglich der Regulation der Akkumulation kryoprotektiver Substanzen bei Kältestress nicht allein durch die Konzentration der Carbohydrate erfolgen kann. Als mögliche Ursache für die Akkumulation löslicher Carbohydrate, welche wiederum die Inhibierung der Expression der in die Photosynthese und den Calvin Cyclus involvierten Gene bedingt, kommt eine Verminderung des Phloemtransportes bei niedrigen Temperaturen in Betracht [302,303].

**Die chloroplastidiäre Mutation der Phosphoglucomutase bedingt eine differentielle Regulation des Raffinosestoffwechsels**

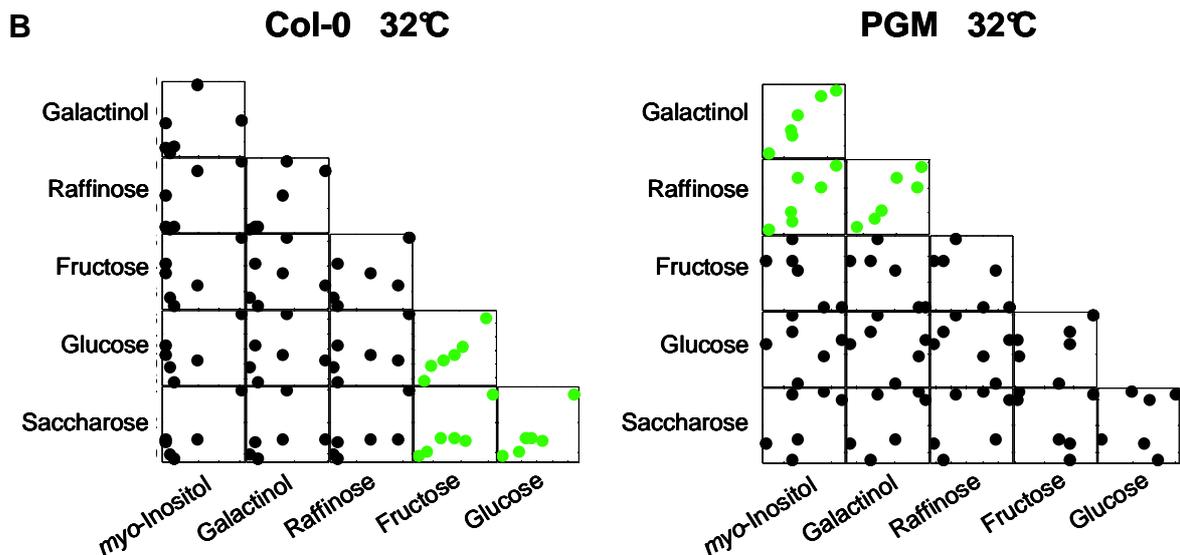
Neben löslichen Carbohydraten akkumulieren Pflanzen bei niedrigen Temperaturen auch Galactoside wie Raffinose und Galactinol. Die Bedeutung dieser zur Familie der Raffinose-Oligosaccharide gehörenden Verbindungen bezüglich der Vermittlung von Kälteresistenz ist bisher jedoch unklar. Taji *et al.* (2002) beobachteten die Akkumulation von Raffinose und Galactinol in *Arabidopsis* nach Trocken-, Salz- sowie Kältestress, wobei die Autoren weder Raffinose noch Galactinol in ungestressten Pflanzen detektieren konnten [304]. Auch auf Transkriptebene beobachteten Liu *et al.* (1998) einen Anstieg der Galactinol Synthase und stellten eine erhöhte Aktivität dieses Enzyms nach Kältestress fest [305]. Zuther *et al.* (2004) konnten jedoch in einer Galactinol Synthase überexprimierenden sowie einer Raffinose Synthase knock-out Mutante von *Arabidopsis* gegenüber dem Wildtyp keine Veränderung der Kältetoleranz nachweisen [306].

Die Analyse der im Rahmen dieser Arbeit erstellten Metabolitprofile temperaturgestresster *Arabidopsis* Genotypen erbrachte, dass neben der drastischen Akkumulation der löslichen Carbohydrate die Veränderungen in den Konzentrationen von Raffinose und Galactinol sowohl nach Hitze-, als auch nach Kältestress die deutlichsten metabolischen Veränderungen darstellten (Abbildung 4.3.1.2). Jedoch wiesen die Profile dieser Verbindungen deutliche Unterschiede zwischen dem Wildtyp Col-0 und der stärkelosen PGM Mutante auf. Die Konzentrationen von Glucose, Fructose und Saccharose nahmen im Wildtyp sowohl nach Hitze-, als auch nach Kältestress zu. Dieses Ergebnis korreliert mit den von Kaplan *et al.* (2004) angestellten Untersuchungen zur Analyse der Temperaturstress-Metaboloms in *Arabidopsis*. Die Autoren identifizierten ebenfalls Saccharose, Glucose und Fructose als metabolische Marker für Hitze- und Kältestress in Col-0 [267].

In der stärkelosen PGM Mutante hingegen nahmen die Konzentrationen dieser Carbohydrate mit steigender Temperatur ab. Für die Galactoside Raffinose und Galactinol erbrachte die Analyse sowohl nach Kälte-, als auch nach Hitzestress in beiden Genotypen erhöhte Blattkonzentrationen (Abbildung 4.3.1.2). Inwieweit die Abnahme der Carbohydratkonzentrationen nach Hitzestress in der stärkelosen Mutante durch eine veränderte Regulation des Raffinosepathways bedingt ist, kann hier nicht beantwortet werden. Die Korrelationsprofile deuten jedoch darauf hin, dass das der PGM Mutante zugrunde liegende enzymatische Netzwerk zumindest nach Hitzestress Veränderungen gegenüber dem des Wildtyps aufweist. Wie Abbildung 4.4.3 zeigt, korrelieren die Intermediate des Raffinosestoffwechsels in der Mutante, nicht jedoch im Wildtyp. Dieses deutet, wie im Kapitel 2 – Metabolische Netzwerke – diskutiert, auf eine Veränderung des enzymatischen Haushaltes in der PGM Mutante nach Hitzestress hin.



**Abbildung 4.4.3: Raffinosepathway und Co-Response beteiligter Intermediate in *Arabidopsis* nach Hitze-stress.** Dargestellt sind die Raffinose-synthese und -degradation in *Arabidopsis* (aus [266]) (**Abb. 4.4.3-A**), sowie die Scatterplotmatrix, welche die Co-Response der Intermediate des Raffinosepathways abbildet (**Abb. 4.4.3-B**). Metabolitpaare, welche gemäss dem Pearson'schen Produktmoment-Korrelationskoeffizienten nach Hitze-stress (32°C) signifikant korrelierten ( $p \leq 0.05$ ), sind grün dargestellt. Der Vergleich der Co-Response offenbart ausgeprägte Unterschiede zwischen Col-0 und PGM. Abkürzungen: GS, Galactinol Synthase; RS, Raffinose Synthase; MEL, Melibiase; INV, Invertase;  $\alpha$ -GLC,  $\alpha$ -Glucosidase.



***Kältestress induziert die Akkumulation der kompatiblen Substanz Prolin als mögliche Folge einer erhöhten Biosyntheserate über den Glutamatweg***

Als Reaktion auf Trockenstress, welcher sich als Sekundärstress bei niedrigen Temperaturen einstellt, ist die Synthese von Metaboliten mit osmotischer Funktion, welche das osmotische Potenzial innerhalb der Zelle senken, in Pflanzen literaturbekannt. Zu diesen sog. kompatiblen Substanzen zählen neben Zuckeralkoholen tertiäre Sulfonverbindungen wie z.B. Dimethylsulfonpropionat, quarternäre Ammoniumverbindungen wie Glycin-Betain auch einfache Aminosäuren [274,307,308].

Im Zusammenhang mit der Response auf Temperaturstress erwies sich neben den löslichen Carbohydraten und Galactosiden auch die Aminosäure Prolin aufgrund deutlich erhöhter Konzentrationen nach Adaption an 4°C als metabolischer Marker für Kältestress (Abbildung 4.3.1.2).

Diverse Studien an transgenen Pflanzen erbrachten, dass hohe Konzentrationen an freiem Prolin in einer erhöhten Toleranz gegenüber vielen abiotischen Stressfaktoren wie Trocken-, Salz- und/oder Kältestress resultiert [304,309-312]. Obwohl die osmoprotektive Wirkung von Prolin unumstritten ist, ist die genaue Funktion dieser Verbindung bisher unklar. Neben einer proteinstabilisierenden Wirkung wurde auch die Interkalierung von Prolin in die Phospholipidmembranen sowie eine Wirkung als Radikalfänger diskutiert [313-315].

In höheren Pflanzen sind zwei mögliche Biosynthesewege zur Bildung von Prolin beschrieben. Als Ausgangsprodukte gelten Glutamat oder Ornithin, wobei die vom Glutamat ausgehende Synthese, welche auch dem Biosyntheseweg in *E. coli* entspricht, überwiegt [316]. An der Synthese des Prolins über den Glutamatweg sind zwei Enzyme beteiligt; die  $\Delta^1$ -Pyrrolin 5-carboxylat Synthetase (P5CS) und  $\Delta^1$ -Pyrrolin 5-carboxylat Reduktase (P5CR).

An transgenen Tabakpflanzen, welche P5CS überexprimierten, beobachteten Kishor *et al.* (1995) eine erhöhte Prolinbiosynthese, konnten jedoch keine Aussagen über die limitierende Komponente der Prolinsynthese machen [309]. Brugière *et al.* (1999) wiesen mit Hilfe von  $^{15}\text{N}$ -Labeling-Experimenten ebenfalls an transgenen Tabakpflanzen nach, dass die Prolinbiosynthese im Wesentlichen durch die Aktivität der Glutamin Synthetase kontrolliert wird. Die Autoren stellten weiter fest, dass nicht Glutamat das wichtigste Substrat für die Synthese von Prolin darstellt, sondern vielmehr freies Glutamin [317].

Mit diesen Ergebnissen korrelieren auch die mittels GC-TOF-MS erstellten Metabolitprofile der untersuchten *Arabidopsis* Genotypen. Sowohl Col-0 als auch PGM wiesen nach Kältestress neben deutlich erhöhten Prolingehalten ebenso erhöhte Konzentrationen an Glutamin auf (Abbildung 4.3.1.2).

#### ***COR-Proteine stellen wichtige Markerproteine für abiotischen Temperaturstress dar***

Eine weitere, mit der Verschiebung von warmen zu kalten Temperaturen verbundene regulatorischen Veränderung des pflanzlichen Stoffwechsels stellt die Synthese neuer Proteine dar. Eine der ersten Arbeiten zur Untersuchung der Response auf Kältestress stellten Guy *et al.* im Jahre 1985 vor [318]. Die seither vorrangig auf Genexpressionsanalysen beruhenden Untersuchungen zahlreicher Arbeitsgruppen resultierten in der Identifizierung mehrerer hundert Gene, welche durch Kältestress induziert werden, darunter *LTI* (low temperature induced)-, *KIN* (cold-inducible)-, *RD*

(responsive to desiccation)-, *ERD* (early dehydration-inducible)- oder *LEA* (late embryogenesis abundant)-Gene, welche häufig als *COR* (cold responsive)-Gene zusammengefasst werden [271,272,275,298,319]. Eine besondere Charakteristik dieser kalteinduzierten Proteine stellen ihre kurze Sequenz (z.B. *COR15am* codiert für ein 9.5 kDa Produkt), wonach sie mitunter auch als Polypeptide bezeichnet werden, sowie ihre extreme Hydrophilie dar [298].

Mittels der in diesem Kapitel vorgestellten Methode zur relativen Quantifizierung von Proteinen über den Spektren-Count konnten diverse Vertreter der Familie dieser *COR*-Proteine detektiert werden. Zudem erbrachte die auf der Varianzanalyse beruhende Auswertung der erhobenen Daten, dass diese Proteine sich in Abhängigkeit von den jeweiligen Stressbedingungen in ihrer Abundanz signifikant unterschieden. So konnten beispielsweise die als 'low-temperature responsive', 'stress responsive' oder 'cold-induced' annotierten Proteine At5g52310.1, At5g15960.1 sowie At2g42530.1 nur in den kältegestressten Blattgeweben der untersuchten Genotypen detektiert werden (Tabelle 4.2.2.1). Auch das Stress-induzierte Protein At5g15970.1 wurde in kältegestressten PGM Pflanzen deutlich stärker exprimiert als in den Kontrollpflanzen (Abbildung 4.2.2.2).

Diese Ergebnisse korrelieren mit unabhängigen, gelbasierten Proteinanalysen der Kältestress-Response. In ihren auf der DIGE-Technologie basierenden Untersuchungen bestätigten Amme *et al.* (2006) die Zunahme des als 'low-temperature responsive' annotierten Proteins 78 (At5g52310.1) in *A. thaliana* bei niedrigen Temperaturen [320]. Dieses zeigt, dass die in diesem Kapitel vorgestellte massenspektrometrische Methode zur relativen Quantifizierung von Proteinen eine wichtige komplementäre Alternative zu den gelbasierten Techniken zur Bestimmung von Fluktuationen auf Proteinebene darstellt. Wenngleich die Klasse der *COR*-Proteine oder -polypeptide der hier angewandten Extraktionsmethode aufgrund ihrer Primärstruktur besonders gut zugänglich waren, reflektieren deren Profile sehr schön die experimentellen Bedingungen, was die Eignung des Spektren-Counts als indirektes Mass für die Abundanz von Peptiden bzw. korrespondierenden Proteinen anzeigt.

#### ***RNA-bindenden Proteinen kommt eine entscheidende Rolle in der Vermittlung einer erhöhten Kältetoleranz zu***

Mit Hilfe der Varianzanalyse gelang die Identifizierung von temperaturspezifisch exprimierten Proteinen und Enzymen, deren Funktion im Hinblick auf eine spezifische Stressresponse bisher wenig beschrieben ist. Als Markerprotein sowohl für Hitze-, als auch für Kältestress fiel das Glycin-reiche RNA-bindende Protein (GR-RBP) At2g21660.1 auf. Die Expression dieses Proteins nahm in beiden Genotypen mit steigender Temperatur nahezu linear ab (Abbildung 4.2.2.2).

Die Identifizierung eines Gens, welches für ein GR-RBP codiert, gelang erstmals 1988 aus Mais [321]. Seither wurden zahlreiche GR-RBP-kodierende Gene in verschiedenen Pflanzenspezies identifiziert [322-326]. Allen bisher bekannten GR-RBP's ist ein RNA-Erkennungsmotiv am N-Terminus, sowie eine Glycin-reiche Region am C-terminalen Ende gemeinsam. Jüngere Beobachtungen deuten darauf hin, dass diese Proteine auch in der Stressresponse involviert sind. Wenngleich einige Untersuchungen zur Expression auf spezifische Funktionen der GR-RBP's schliessen lassen [327,328], ist die genaue Bedeutung dieser Klasse von Proteinen bis heute unverstanden [329]. Allgemein wird angenommen, dass den GR-RBP's eine wichtige Rolle bei der Prozessierung der RNA sowie post-transkriptionellen Modifikationen zukommt.

An transgenen *Arabidopsis*-Linien, welche das GR-RBP AtRZ-1a, ein Homolog des im Rahmen dieser Arbeit identifizierten GR-RBP's At2g21660.1, überexprimierten, beobachteten Kim *et al.* (2005) eine erhöhte Kältetoleranz [330]. Nach weiterführenden Untersuchungen zur Aufklärung des Wirkungsmechanismus dieses Genproduktes postulierten die Autoren, dass die AtRZ-1a-induzierte Hochregulation der Expressionen verschiedenster COR- und weiterer in Kältestress involvierter Proteine posttranslational erfolgen muss [331]. Unter den in den transgenen Linien co-exprimierten Proteinen identifizierten die Autoren u.a. die in der cytosolischen Acetyl-CoA Synthese involvierte ATP-Citrat Lyase.

Die vergleichende Analyse der Kälte-, sowie Hitze-induzierten Proteinprofile offenbarte für beide der im Rahmen dieser Arbeit analysierten *Arabidopsis* Genotypen einen ähnlichen Trend der Proteinexpression von At2g21660.1 sowie des chloroplastidiären Acyl-Carrier Proteins At1g54630.1 (Tabellen 4.2.2.1 und 4.2.2.2). Die Konzentrationen beider Proteine nahmen mit steigender Temperatur ab, so dass ein möglicher Einfluss von GR-RBP's auf eine posttranskriptionelle Modulation der Lipidbiosynthese in Betracht kommt.

Als Markerproteine fielen weiter zwei kerncodierte Ribonucleoproteine mit chloroplastidiärer Zielsequenz (At2g37220.1 sowie At3g53460.1) auf. Die ermittelten Konzentrationen dieser Proteine offenbarten eine im Vergleich zu den Kontrollpflanzen deutlich verminderte Expression nach Hitzestress (Tabelle 4.2.2.2). Über die Funktionen von At2g37220.1 sowie At3g53460.1 in Verbindung mit Temperaturstress herrscht bis dato ebenso Unklarheit, wie über die Mechanismen der Vermittlung der Stresstoleranz durch die GR-RBP's. Auch diese Proteine sind durch eine RNA-bindende Domäne gekennzeichnet, welche ursprünglich aus Tabak isoliert wurde [332]. Somit könnte diesen Proteinen eine Funktion ähnlich denen der Chaperonine zugrunde liegen.

***Hitzeschockproteine verleihen der Pflanze eine gewisse Elastizität gegenüber hohen Temperaturen und sind für das Überleben bei starker Wärmebelastung essentiell***

Bei hohen Temperaturen kommt es in Pflanzen, aber auch in Mikroorganismen oder Tieren, zur Aktivierung einer evolutionär konservierten Gruppe von Genen sowie zur transienten Expression der korrespondierenden Proteine. In Pflanzen sind bis dato etwa 30 dieser induzierbaren sog. Hitzeschockproteine (HSP's) bekannt. HSP's verhindern die Denaturierung, falsche Faltung sowie die Aggregation denaturierter Proteine, womit ihnen die Funktion von Chaperoninen zukommt [333].

Während unter den detektierten Proteinen zahlreiche kälteinduzierte oder auch Kälteresistenz vermittelnden Proteine vertreten waren, erwies sich die Anzahl detektierter Hitzeschockproteinen als äusserst gering. Von drei detektierten Vertretern dieser Proteinklasse zeigte allein das 'heat shock cognate 70 kDa Protein 3' (At3g09440.1) eine temperaturabhängige Expression. Im Wildtyp Col-0 konnte dieses zur Familie der HSP70 gehörende Proteine sowohl im hitze-, als auch im kältegestressten Blattgewebe detektiert werden, während in den Kontrollpflanzen dieses Genotyps keine Expression nachgewiesen werden konnte (Tabelle 4.2.2.1). Interessanterweise konnte dieses Protein nicht in den hitzegepressten Blattgeweben der PGM Mutante detektiert werden.

In höheren Pflanzen setzt die Synthese der HSP's im Allgemeinen bereits wenige Minuten nach der Erhöhung der Temperatur in den Bereich von etwa 38-40°C in allen Geweben ein [296]. Eine Induktion der vorrangig als Chaperonine wirksamen HSP's auch durch andere als Hitzestressfaktoren ist literaturbekannt [296,334,335]. Wenngleich in dem hier vorgestellten Versuch die Temperaturbehandlung mit 32°C unterhalb des kritischen Bereiches lag, liegt diesem Ergebnis möglicherweise eine veränderte Regulation der Expression dieser Proteinklasse in der PGM Mutante zugrunde. Möglich ist jedoch auch, dass die gemäss ihrer Funktion an Komponenten denaturierten Proteins gebundenen Hitzeschockproteine mit dem verwendeten Extraktionsprotokoll aufgrund des Fehlens von Urea nicht kompatibel sind.

***Die Bestimmung des sekundären oder tertiären Stress-Status anhand von Proteinexpressionsprofilen wird durch ein breites Arsenal von in Arabidopsis vorkommenden (Iso)Enzymen erschwert***

In Verbindung mit Temperaturstress stellt sich als Sekundär- oder Tertiärstress immer auch oxidativer Stress ein. An der Detoxifizierung reaktiver Sauerstoffspezies sind folgende Enzyme wesentlich beteiligt: Superoxid Dismutasen (SOD), welche die Dismutation von Superoxidradikalen zur Hydrogenperoxiden und molekularem Sauerstoff katalysieren, Ascorbat Peroxidasen (APX) und Catalasen (CAT), welche ebenfalls an der Detoxifizierung der Hydrogenperoxide beteiligt sind, sowie Dehydroascorbat Reduktasen

und Glutathion Reduktasen (GR), welche die Regenerierung von Ascorbat, welches APX-katalysiert oxidiert wird, katalysieren [336]. Neben den ausschliesslich in Mikrobodies lokalisierten Catalasen kommen in Pflanzen weitere, in den Chloroplasten, den Mitochondrien sowie im Cytosol lokalisierte detoxifizierende Enzyme vor, von denen zahlreiche Isoformen existieren. Die Aktivierung spezifischer detoxifizierender Enzyme als Reaktion auf Temperaturstress ist speziesübergreifend nicht einheitlich. Während in Tabak bei niedrigen Temperaturen die chloroplastidiäre und cytosolische Fe-SOD induziert wird [337], erfolgt die Aktivierung von SOD und APX in Reis nach dem Transfer von kalten zu ambienten Temperaturen [338]. In *Arabidopsis* induziert Kältestress vor allem die Aktivierung von APX und GR [339]. Daneben resultiert Hitzestress in einer auffälligen Akkumulation der cytosolischen Fe-SOD. Auch cytosolische APX werden in *Arabidopsis* bei hohen Temperaturen verstärkt exprimiert, wobei die Induktion der Genexpression über Hitze Schock Elemente erfolgt [340,341].

Ein weiteres, in der Detoxifizierung von ROS involviertes Protein stellt die Glutathion S-Transferase (GST) dar [259]. Die Isoform At2g30860.1 dieses Enzyms konnte im Rahmen der hier vorgestellten Analyse der Temperaturstressresponse der *A. thaliana* Genotypen Col-0 und PGM ausschliesslich in den temperaturgestressten Blattgeweben des Wildtyps Col-0 detektiert werden, während die Profile der GST in der PGM Mutante auf invariante Konzentrationen schliessen lassen. Ähnliche Ergebnisse erbrachte die Analyse für die Catalase 3 (At1g20620.1). Auch dieses Enzym konnte in den ungestressten Wildtyp Pflanzen nicht detektiert werden, während in der PGM Mutante keine Induktion der Expression durch hohe oder niedrige Temperaturen erfolgte (Tabelle 4.2.2.1). Die ermittelten Konzentrationen der Isoform At4g04020.1 eines plastidiären Lipid-assoziierten Proteins (PAP) deuten jedoch darauf hin, dass dessen Expression in beiden Genotypen sowohl durch die Temperaturniedrigung, als auch durch die Temperaturerhöhung induziert wurde (Tabelle 4.2.2.1). Auch für PAP ist eine Akkumulation unter biotischen sowie abiotischen Stresssituationen, insbesondere bei oxidativem Stress, in den Blättern der *Solanaceae* nachgewiesen [342].

Die Varianzanalyse erbrachte weiter, dass zwei putative Dehydroascorbat Reduktase-Isoformen differentiell exprimiert wurden. Während At1g19550.1 ausschliesslich in der PGM detektiert wurde, offenbarten die Expressionsprofile der Isoform At1g19570.1 deutlich erhöhte Konzentrationen in den hitzestressten Blattgeweben des Wildtyps (Tabellen 4.2.2.1 und 4.2.2.2). Wenngleich relativ viele der differentiell exprimierten Proteine gemäss ihrer Annotation in der Entgiftung von ROS involviert sind, lassen die Ergebnisse der hier vorgestellten Analysen keine Aussage über der oxidativen Stress-Status in den untersuchten Genotypen zu. Dieses Ergebnis ist besonders vor dem Hintergrund der drastischen Akkumulation löslicher Carbohydrate, welche einer leichten Autoxidierbarkeit unter Bildung zahlreicher Produkte, welche ihrerseits autoxidatorische

Wirkungen aufweisen, in der PGM sowie nach Kältestress in beiden untersuchten Genotypen zunächst erstaunlich. Dennoch sind für dieses Ergebnis mehrere Gründe vorstellbar.

Der während der Photosynthese entstehende Disauerstoff stellt neben der Grundlage des Energiestoffwechsels auto- und heterotropher Organismen, welche O<sub>2</sub> als Elektronenakzeptor der Atmungskette nutzen, ein wirksames Zellgift dar. Während der Photosynthese herrschen in den Chloroplasten der Pflanzen hohe O<sub>2</sub>-Konzentrationen. Gleichzeitig wird durch die photosynthetische Elektronentransportkette ein hohes Mass an energiereichen Reduktionsäquivalenten bereitgestellt. Wird im Photosystem II mehr Lichtenergie absorbiert, als im anschließenden Calvin Cyclus nutzbar gemacht werden kann, erfolgt eine Spinumkehr des Triplett-Sauerstoffs in den angeregten Singulett-Sauerstoff. Um diese extrem reaktive Sauerstoffspezies und die aus dessen Folgereaktionen resultierenden ROS zu entgiften, verfügen Pflanzen über zahlreiche Isoformen detoxifizierender Enzyme. Die TAIR (The Arabidopsis Information Resource)-Datenbank enthält beispielsweise 8 SOD's, 10 Isoformen der APX sowie 12 GR-Isoformen (<http://www.arabidopsis.org>). Die im Rahmen dieser Arbeit detektierten Proteine stellen somit nur einen Bruchteil der in *Arabidopsis* vorkommenden Enzyme, welche in der ROS-Detoxifizierung involviert sind, dar. Zudem liegen keine Informationen über deren Enzymaktivitäten vor. Auch zeigten verschiedene Bemühungen, transgene Pflanzen mit erhöhter Stresstoleranz zu kreieren, dass dieses Strategien erfordert, welche das komplexe ROS-detoxifizierende System berühren. So konnte für GR-überexprimierende transgene Tomatenpflanzen trotz erhöhter Aktivität des GR-Proteins keine erhöhte Toleranz gegenüber niedrigen Temperaturen nachgewiesen werden [343].

#### ***Temperaturstress induziert die Akkumulation zahlreicher primärmetabolischer Signalmoleküle sowie Precursor-Metabolite des Sekundärstoffwechsels***

Neben vielen Proteinen, welchen eine regulatorische Funktion in der Stressresponse zukommt, agieren zahlreiche Metabolite als Initiatoren von Abwehrmechanismen zur Verringerung oder Vermeidung von Schädigungen oder auch zu Erhöhung der Stresstoleranz. Wenngleich diese Abwehrmechanismen vorrangig auf der Ebene des Sekundärstoffwechsels wie z.B. im Rahmen des Phenylpropanoid-, des Isoprenoid- oder auch des Alkaloidstoffwechsels erfolgen, wird die Induktion dieser Mechanismen häufig durch einen veränderten Status spezifischer Intermediate des Primärstoffwechsels angezeigt. Zu diesen Precursoren zählen beispielsweise verzweigte Aminosäuren, welche die Vorstufen cyanogener Glycoside darstellen [344].

Die in der vorliegenden Arbeit vorgestellten Metabolitprofile offenbaren nach Kältestress insbesondere im Wildtyp deutlich erhöhte Konzentrationen der aromatischen Amino-

säuren Tryptophan, Phenylalanin und Tyrosin (Abbildungen 4.3.1.3 und 4.3.1.5), welche die Ausgangsverbindungen diverser sekundärmetabolischer Routen darstellen. Neben ihrer antioxidativen Wirkung bilden phenolische Intermediate Vorstufen u.a. für die Biogenese von Phytoalexinen, welche postinfektionell zur Abwehr von Schadorganismen (meist Pilze) gebildet werden, Ligninen, denen eine Funktion bei der Pathogenabwehr zugesagt wird, Flavanen und Flavonoiden, welche die Pflanze vor Schädigungen durch ionisierende Strahlung bewahren sowie Cumarinen und Ligninen, welche sich von der Zimtsäure ableiten [345]. Neben den aromatischen Stoffwechselintermediaten ergaben sich auch für Vorstufen der cyanogenen Glycoside, Leucin, Valin und Isoleucin [344] signifikant erhöhte Konzentrationen in den Blättern gestresster Pflanzen (Abbildung 4.3.1.5). Zusammenfassend indizieren diese Ergebnisse eine erhöhte Produktion von Sekundärmetaboliten als Response oder auch Abwehrmechanismus zur Vermeidung von Schädigungen infolge von Kältestress.

Im Zusammenhang mit der allgemeinen Response auf Stress wird in der Literatur häufig die Bedeutung von Salicylsäure diskutiert [249,346-348]. In den Untersuchungen zur Temperaturstressresponse von *Arabidopsis* stellten Kaplan *et al.* (2004) bezüglich dieses Metaboliten eine transiente Erhöhung bei Hitzestress (innerhalb von ca. 10 min) sowie bei Kältestress (innerhalb von 12h) fest und identifizierten Salicylsäure als einen Schlüsselmetaboliten der frühen Hitzeschockresponse [267]. Die massgebliche Rolle der Salicylsäure in der systemisch erworbenen Resistenz gegen Pathogene wie Bakterien Pilze oder Viren gilt als evident [245,349-351]. Auch im Zusammenhang mit der Toleranz gegenüber Temperaturstress kommt der Salicylsäure eine grosse Bedeutung zu [352]. In Übereinstimmung mit Studien, in denen eine erhöhte Kältetoleranz nach externer Salicylsäureapplikation beobachtet wurde [352-354], offenbarten auch die im Rahmen dieser Arbeit erstellten Metabolitprofile erhöhte Konzentrationen dieser Verbindung in den kältegestressten Blättern von Col-0 und PGM (Abbildung 4.3.1.3). Zudem indiziert der erhöhte Salicylsäuregehalt die verstärkte Synthese weiterer Phytohormone wie Jasmonsäure und deren Derivate, Abscisinsäure oder Ethylen. In der Literatur finden sich zahlreiche Hinweise auf eine gegenseitige Beeinflussung der Signaltransduktionswege dieser Verbindungen [355-357].

### **Zusammenfassung**

Zusammenfassend erbrachte die hier vorgestellte Analyse der Temperaturstressresponse von *A. thaliana*, dass, ähnlich den Ergebnissen der im Kapitel 3 vorgestellten Analyse der zeitabhängigen Regulation des pflanzlichen Stoffwechsels, die deutlichsten Veränderungen auf der Ebene der Metabolite nachweisbar waren. Die Analysen beider Stoffwechsel-

ebenen, Proteine und Metabolite, offenbarte neben literaturbekannten Markern, was die Güte der Analysen bestätigt, auch unerwartete, bisher unbekannte Temperaturstress-induzierte Veränderungen in den *A. thaliana* Genotypen Col-0 und PGM.

Die Variationskoeffizienten der jeweils in 6 biologischen Replikate gemessenen Metabolite zeigten, dass die analysierten Pflanzen nach 72-stündiger Exposition bei 4°C und 32°C an die jeweiligen Stressbedingungen sowohl auf Metabolit-, als auch auf Proteinebene weitgehend angepasst waren.

Für etwa 75% der erfassten kleinen Moleküle ergaben sich nach Temperaturstress signifikante Unterschiede in ihrer Konzentration. Als Ergebnis dieser zu einem erheblichen Teil unspezifischen Response auf den Sekundärreiz der Temperaturveränderung in einen physiologisch ungünstigen Bereich konnten verschiedene Metabolite als stressspezifische Markersubstanzen identifiziert werden.

So erbrachte die Analyse der vorliegenden Daten, dass nach Kältestress Konzentrationen typische Substanzen mit osmoprotektiver Wirkung, wie z.B. Glucose, Fructose und Saccharose, drastisch akkumuliert wurden. Daneben wurde eine Konzentrationserhöhung weiterer kompatibler Substanzen beobachtet. Sowohl Prolin als auch Glutamin akkumulierten bei Kältestress deutlich. Dieses Ergebnis deutet zugleich darauf hin, dass die Prolinbiosynthese, wie von Brugière postuliert [317], über die Glutamin Synthetase reguliert wird.

In Übereinstimmung mit den Arbeiten von Kaplan *et al.* [267] wurden die Galactoside Raffinose und Galactinol als metabolische Marker sowohl für Kälte-, als auch für Hitzestress in beiden analysierten *Arabidopsis* Genotypen, Col-0 und PGM, identifiziert. Darüber hinaus deuten die Korrelationsprofile der Intermediate des Raffinosepathways auf eine veränderte Regulation dieses Stoffwechselweges in der zuckerakkumulierenden Mutante hin.

Entgegen dem auf Transkriptebene häufig beobachteten Effekt der verstärkten Transkription einer Vielzahl von Genen als Reaktion der Pflanze auf einen bestimmten Stressfaktor, belief sich die Anzahl der Proteine, welche als Response auf den abiotischen Stressfaktor Temperatur eine nachweisbar veränderte Expression aufwiesen, mit 25% als vergleichsweise gering. Im Hinblick auf die Möglichkeit der Regulation der Proteinfunktion über diverse posttranslationale Modifikationen und unter Berücksichtigung des energetischen Aufwandes, den die Proteinsynthese erfordert, erscheint dieses Ergebnis jedoch plausibel.

Die mit der Reaktion auf Temperaturstress einhergehende morphologische Veränderung der Lipidmembranen wurde durch die Expression des Acyl-Carrier Proteins At1g54630.1 sowie durch erhöhte Konzentrationen an Citrat, die wohl einzige Ausgangsverbindung für die Synthese von cytosolischem Acetyl-CoA, in den kältegestressten Pflanzen angezeigt.

Als Proteinmarker für Kältestress wurden vorrangig Vertreter der COR-Proteine identifiziert, deren Expressionsprofile sehr gut mit dem experimentellen Setup korrelierten.

Durch verminderte Expression in den hitzestressten Pflanzen fielen weiter Proteine mit RNA-bindenden Domänen auf. Neben zwei kerncodierten Proteinen mit chloroplastidiärer Zielsequenz (At2g37220.1 sowie At3g53460.1) konnte das Glycin-reiche RNA-bindende Protein At2g21660.1, dessen Expression in beiden Genotypen mit steigender Temperatur abnahm, als Markerprotein identifiziert werden. Dieses indiziert eine Beteiligung dieser funktionell bisher nicht charakterisierten Proteine an der Regulation der Transkription oder auch posttranskriptionellen Modifikation von Genprodukten, welche kälteinduziert exprimiert werden.

Des Weiteren deuten die erhöhten Konzentrationen verzweigter und aromatischer Aminosäuren auf eine verstärkte Synthese von Sekundärmetaboliten hin, welchen eine bedeutende Rolle im Zusammenhang mit der Erhöhung der Stresstoleranz zukommt. Ebenso erwiesen sich die erhöhten Salicylsäuregehalte nach Kältestress als Indikatoren für die Synthese von Phytohormonen, welche systemische Stressantworten induzieren.

## **5 Integrative Datenanalyse zur Mustererkennung und Identifikation von potenziellen Biomarkern in *Arabidopsis thaliana***

### **5.1 Strategien zur Identifikation von Biomarkern**

Unter einem Biomarker wird ein objektiv messbarer Parameter verstanden, der einen physiologischen oder pathologischen Zustand oder auch die Antwort auf einen pharmakologischen Eingriff anzeigt [358]. Die Strategien zur Entdeckung von Biomarkern gehen von der Annahme aus, dass durch die Abweichung eines extrinsischen oder auch intrinsischen Parameters vom Normal- oder Grundzustand eine physiologische Reaktion hervorgerufen wird, die eine spezifische Änderung im Genexpressions-, Protein- oder Metabolitmuster erzeugt. Da an einen Biomarker die Anforderung der allgemeinen Gültigkeit gestellt ist, müssen die Strategien zur Identifizierung solcher Marker ein umfassendes Bild des physiologischen Status der zu untersuchenden Objekte abgeben und gleichzeitig einen hohen Probendurchsatz erlauben, um eine hinreichende Spezifität des identifizierten Markers nachzuweisen. Diesen Anforderungen werden molekulare Profiling-Technologien wie das Metabolite- oder Protein Profiling gerecht.

Im ersten Schritt der Entdeckung eines Biomarkers gilt es, aus den mit multiparallelen Technologien gewonnenen hochdimensionalen Datensätzen relevante Informationen herauszufiltern. Da ein einzelnes Objekt in diesen hochdimensionalen Datensätzen durch eine Vielzahl von Merkmalen repräsentiert wird, muss die Dimensionalität, welche durch die Anzahl der Variablen gegeben ist, geeignet reduziert werden. Dazu bieten sich dimensionsreduzierende datenanalytische Techniken wie die Haupt- oder die Unabhängige Komponentenanalyse an, mit deren Hilfe der Originaldatensatz unter grösstmöglichem Erhalt der inhärenten Varianz in einen wenig dimensionalen Unterraum zerlegt werden

kann. Eine weitere Möglichkeit zur Identifizierung von Biomarkern stellt die Reduktion der Dimensionen durch Eliminierung von irrelevanten Merkmalen dar. Hierbei wird jede Dimension separat bewertet; irrelevante Attribute werden aus dem Datensatz entfernt.

Der nach Anwendung einer oder der Kombination beider Verfahren erhaltene reduzierte Datensatz beinhaltet dann Variablen, welche als Indikatoren für die Abweichung vom Normalzustand oder einen Belastungszustand gelten können. Die Eignung dieser Variablen als Biomarker, anhand derer sich ein qualitativer und quantitativer Bezug zwischen der Expositionzeit und der physiologischen Antwort eines Organismus auf die vom Normalzustand abweichenden Veränderungen herstellen lassen muss, muss jedoch im Einzelfall validiert werden.

In den vorangegangenen Kapiteln wurde deutlich, dass die Auswertung der mit den sog. 'Omics'-Technologien gewonnene Daten zuweilen recht aufwendig ist und fundierte Kenntnisse über verschiedene Verfahren der multivariaten Datenauswertung erfordert. Die Verfügbarkeit von Informationen über komplementäre Stoffwechselebenen besitzt das Potenzial, Stoffwechselvorgänge innerhalb eines Organismus als komplexe Vorgänge erklären zu können und die gegenseitige Beeinflussung von transkribierten Genen, translatierten Proteinen und beteiligten Metaboliten zu erfassen. Hieran sind jedoch auch neue Anforderungen an die Datenverarbeitung und –auswertung geknüpft.

Das folgende Kapitel befasst sich mit der Auswertung und Visualisierung multivariater Daten. Ziel war dabei das Auffinden einer datenanalytischen Methode, welche es ermöglicht, biologisch relevante Informationen aus hochdimensionalen Datensätzen schnell zu extrahieren, diese graphisch adäquat abzubilden und dem jeweils zugrunde liegenden experimentellen Design entsprechende, potenzielle Biomarker zu identifizieren. Als Testsets dienten die Datensätze der vorangegangenen Kapitel. Dabei wurden die mit unterschiedlichen Technologien gewonnen Protein- und Metabolitdaten integrativ analysiert, um eine interpretierbare und im Einklang mit den zugrunde liegenden experimentellen Bedingungen stehende Verbindung zwischen diesen beiden Stoffwechselebenen herzustellen.

## **5.2 Mustererkennung am Beispiel der zeitabhängigen systemischen Regulation des Metabolismus von *Arabidopsis thaliana***

Die Analysen der im vorangegangenen Kapitel vorgestellten Ergebnisse zur Untersuchung der zeitabhängigen systemischen Regulation des Metabolismus von *A. thaliana* haben gezeigt, dass der diurnale Rhythmus beider untersuchten Genotypen Col-0 und PGM durch die Dynamik vieler der erfassten Metabolite reflektiert wurde. Auch konnten in Kapitel 3 Proteine identifiziert werden, deren Expressionen in Abhängigkeit von der

Photoperiode reguliert wurden. In einer weiterführenden Analyse wurde daher getestet, wie die Dynamik beider Stoffwechselebenen in Verbindung mit der entsprechenden Klassifizierung der einzelnen Proben abgebildet werden kann. Da sowohl die Protein- als auch die Metabolitdaten hochdimensionale Räume darstellen, mussten diese zunächst komprimiert werden. Eine im Feld der Metabolomics-Technologien häufig angewandte Methode zur Dimensionsreduktion und Mustererkennung ist die Hauptkomponentenanalyse (Principal Components Analysis, PCA). Mit deren Hilfe ist es möglich, hochdimensionale Merkmalsräume, wie sie durch die grosse Anzahl der gemessenen Metabolite als Variablen entstehen, in einen niedrigdimensionalen Unterraum zu transformieren, in dem der Hauptteil der Datenvarianz liegt. Dabei wird die Datenmatrix in sog. Hauptkomponenten zerlegt, welche aus der Korrelations- oder Kovarianz-matrix der ursprünglichen Daten extrahiert werden und jeweils Linearkombinationen der ursprünglichen Matrixelemente darstellen. Die erste Hauptkomponente beinhaltet dabei die grösste Varianz innerhalb des Datensatzes. Alle zwingend orthogonal zueinander liegenden Hauptkomponenten oder auch Faktoren werden nacheinander in Richtung abnehmender Varianz bestimmt. Bei ausreichend repräsentativen Daten und linearen Zusammenhängen können die Hauptkomponenten häufig interpretiert werden.

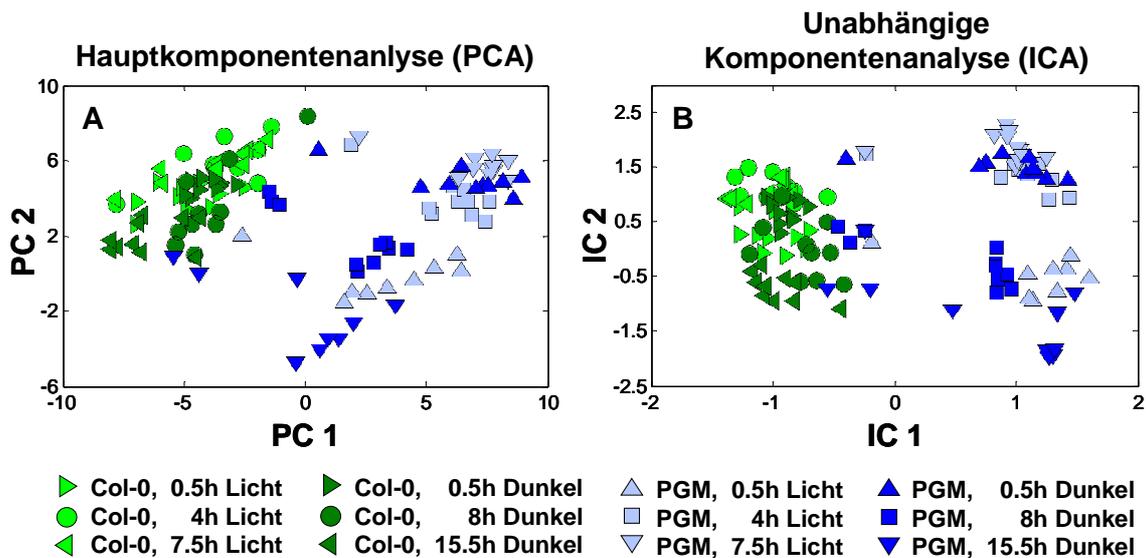
Sollen Daten, welche durch unterschiedliche Technologien gewonnen wurden, integrativ analysiert werden, müssen diese vor deren eigentlichen Verarbeitung normalisiert werden. Um eine vergleichbare Skalierung der mittels GC-TOF-MS und LC-MS gewonnenen quantitativen Daten, welche keine absoluten Konzentration sondern willkürliche Einheiten darstellen, zu gewährleisten, wurden diese zunächst auf den Median aller Messungen normiert. Dieses erlaubt einen direkten Vergleich der Fluktuationen von Metabolit- oder Proteinleveln, wobei starke relative Veränderungen weiterhin durch eine grosse Varianz repräsentiert werden. Durch anschliessende  $\log_{10}$ -Transformation wurden die aus der Median-Normierung resultierenden schiefwinklig verteilten Variablen einer Gaussischen Verteilung angenähert.

Die zweidimensionale Projektion der Metabolitdaten des im Kapitel 3 diskutierten Experimentes zur Bestimmung der zeitabhängigen systemischen Regulation lieferte nach vorangegangener Hauptkomponentenanalyse die in Abbildung 5.2.1-A dargestellte Klassifizierung der zu unterschiedlichen Zeitpunkten innerhalb eines Tages geernteten Proben der analysierten Genotypen. Anhand der Lage der Datenpunkte wird deutlich, dass die Varianz zwischen den Datensätzen der analysierten Genotypen Col-0 und PGM maximal ist. Darüber hinaus lässt Abbildung 5.2.1-A erkennen, dass die zweite Hauptkomponente den Erntezeitpunkt der Proben reflektiert. Dennoch erbrachte die PCA nur eine unzureichende Separation. Während die Unterschiede zwischen den Genotypen durch die erste Hauptkomponente noch adäquat abgebildet wurden, gelang die

Beschreibung der mit der circadianen Rhythmik verbundenen Veränderungen im pflanzlichen Stoffwechsel durch die zweite Hauptkomponente nur bedingt. Dieses Ergebnis offenbart zugleich einen Nachteil der PCA: schiefwinklige Verteilungen können aufgrund der Tatsache, dass die Lage der Datenpunkte im neu berechneten Vektorraum allein durch Rotation verändert wird, nur schlecht approximiert werden.

Um die Mustererkennung oder Merkmalstrennung zu verbessern, wurde eine Kombination aus PCA und einer weiteren Methode der Dimensionsreduktion, der Unabhängigen Komponentenanalyse (Independent Components Analysis, ICA) vorgeschlagen. Mit Hilfe der ICA wurden die Variablen des Originaldatensatzes, ähnlich der PCA, nach einer Hauptachsentransformation in einen Vektorraum mit neuer Basis überführt, der durch statistisch unabhängige Komponenten oder Gruppen in den Daten aufgespannt wird [168]. Auch hierbei stellen die neu berechneten Komponenten Linearkombinationen der untransformierten Variablen dar, jedoch können hierbei im Gegensatz zu PCA auch nicht-orthogonale Strukturen in untransformierten Originaldatensätzen erfasst werden, wodurch die Separation von Proben mitunter deutlich verbessert werden kann. Gewöhnlich wird die ICA auf Datensätze angewendet, deren Anzahl der Replikate oder Proben die Anzahl der Variablen deutlich übersteigt. Im vorliegenden Datensatz jedoch war die Anzahl der Variablen (80 quantifizierte Metabolite) relativ hoch gegenüber der Anzahl gemessener Proben. Häufig liefert die Anwendung der ICA auf solch hochdimensionale Datensätze, welche zumeist eine beträchtliche Streuung der Datenpunkte aufweisen, keine adäquate Entmischung der statistisch unabhängigen Gruppen. Daher wurde die Dimensionalität des vorliegenden Datensatz zunächst durch Anwendung der PCA reduziert, wobei die Varianz, welche ein Mass für den Informationsgehalt der Daten darstellt, erhalten blieb. Nach Anwendung der ICA auf diesen reduzierten Datensatz wurden die neu berechneten unabhängigen Komponenten anhand der Kurtosis, welche ein Mass für die nicht-Gaussische Verteilung darstellt, bewertet.

Die zweidimensionale Projektion der unabhängigen Komponenten mit der höchsten Kurtosis offenbart, dass die Proben der beiden Genotypen Col-0 und PGM zwei statistisch unabhängige Datensätze darstellten (IC1), während die zweite Dimension bzw. die zweite unabhängige Komponente (IC2) eine Abhängigkeit vom Erntezeitpunkt erkennen lässt (Abbildung 5.2.1-B). Dabei bilden die Proben der PGM Mutante zwei unabhängige Gruppen, während die Proben des Wildtyps überlappen. Zugleich wird ein zeitlicher Versatz deutlich. Die jeweils in der Mitte und am Ende der Licht- bzw. Dunkelperiode geernteten Proben der PGM Mutante weisen eine starke Ähnlichkeit mit den 30 min nach dem Wechsel der jeweiligen Periode geernteten Proben auf.



**Abbildung 5.2.1 : Visualisierung circadianer Unterschiede in den Metabolitprofilen von *A. thaliana* Col-0 und PGM Blattmaterial.** Zu 6 verschiedenen Zeitpunkten des Tages wurden jeweils 10 Pflanzen eines jeden Genotyps geerntet und individuell analysiert. Dieses umfasste die Identifikation und Quantifizierung von 80 strukturbekanntem Metaboliten. Nach einer Dimensionsreduktion des Datensatzes durch die Hauptkomponentenanalyse (**Abb. 2.4.1.1-A**) sowie die Unabhängige Komponentenanalyse basierend auf 4 Hauptkomponenten (**Abb. 2.4.1.1-B**) repräsentiert jeder Punkt eine Probe bzw. Pflanze.

Die Analyse der Proteindaten erbrachte nach gleichem Vorgehen die in Abbildung 5.2.2-A dargestellte Klassifizierung. Wieder erwies sich die ICA als das datenanalytische Verfahren, welches nach der Dimensionsreduktion eine Visualisierung des vorliegenden multivariaten Datensatzes ermöglichte, welche die experimentellen Bedingungen am deutlichsten reflektiert. Die zweidimensionale Projektion des von den ersten beiden Unabhängigen Komponenten aufgespannten Datenraumes offenbart, dass die Entwirrung der vier experimentellen Bedingungen Col-0, PGM, Licht und Dunkel im Ansatz gelang. Eine deutliche Verbesserung der Klassifizierung erbrachte die Transformation des integrativen Protein-Metabolit-Datensatzes durch die ICA. Nach entsprechender Hauptachsentransformation wurde von den neu berechneten ersten beiden unabhängigen Komponenten ein Datenraum aufgespannt, in dem die Proben sinnvoll angeordnet wurden. Die zweidimensionale Anordnung offenbart vier voneinander unabhängige Gruppen, deren Klassifizierung das experimentelle Design weitestgehend reflektiert. Die erste unabhängige Komponente (IC1) wurde dabei von genotyp-spezifischen Merkmalen dominiert, während die zweite unabhängige Komponente (IC2) in beiden Genotypen eine Abhängigkeit von der Photoperiode erkennen lässt (Abbildung 5.2.2-C).

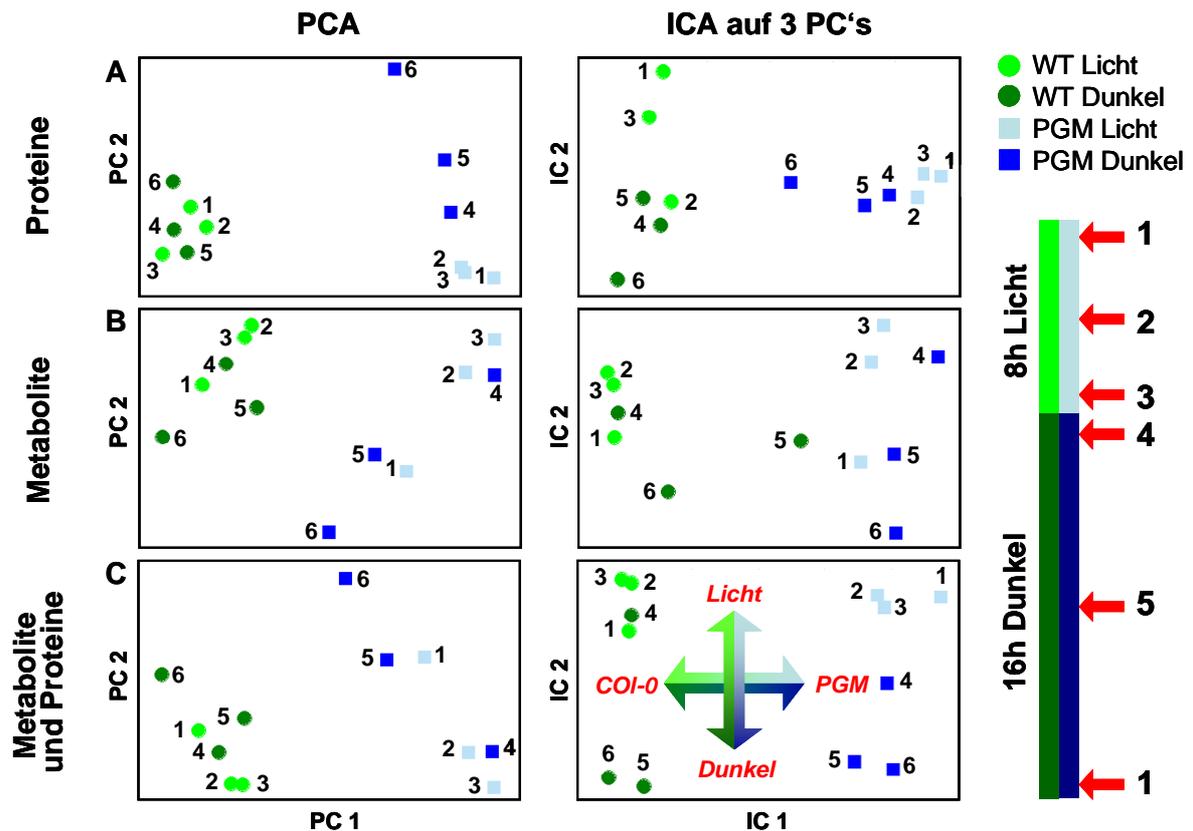


Abbildung 5.2.2 : Gegenüberstellung von PCA und ICA der Proteindaten (Abb. 5.2.2-A), der Metabolitdaten (Abb. 5.2-B) und des integrativen Protein-Metabolit-Datensatzes (Abb. 5.2.2-C). *Arabidopsis thaliana* Pflanzen der Genotypen Col-0 Wildtyp sowie PGM Mutante wurden zu 6 verschiedenen Zeitpunkten eines Tages geerntet und analysiert: 1, 0,5h Licht; 2, 4h Licht; 3, 7,5h Licht; 4, 0,5h Dunkel; 5, 8h Dunkel; 6, 15,5h Dunkel. Die deutlichste Diskriminierung zwischen den 4 Bedingungen (Col-0, PGM, Licht, Dunkel) gelang durch die Anwendung der ICA auf den durch vorangegangene PCA prä-reduzierten integrativen Protein-Metabolit-Datensatz.

### 5.3 Potenzielle Biomarker der zeitabhängigen systemischen Regulation des Metabolismus von *Arabidopsis thaliana*

Der im vorangegangenen Kapitel 5.2 vorgestellte Vergleich dimensionsreduzierender Verfahren offenbarte nach anschließender Visualisierung eine deutliche Abhängigkeit der Klassifizierung der Datenpunkte vom zugrunde liegenden Datensatz. Die das experimentelle Setup am deutlichsten reflektierende Klassifizierung erbrachte die Anwendung der ICA auf den integrativen Datensatz.

Die Faktorladungen, welche die Relevanz einzelner Variablen entlang der jeweiligen unabhängigen Komponenten beschreiben, sind in Tabelle 5.3.1 aufgeführt und verdeutlichen, dass insbesondere die Metabolite eine hohe Korrelation mit den

Komponenten aufweisen. Die Faktorladungen offenbaren weiter, dass die Wichtung der einzelnen Variablen an die dimensionsreduzierende Methode geknüpft sind.

Während die aufgrund des genetischen Hintergrundes in der PGM Mutante akkumulierten löslichen Carbohydrate Glucose, Fructose und Saccharose stark auf die ersten beiden Hauptkomponenten (PC1 und PC2) laden, kommt diesen Metabolite ein deutlich geringerer Einfluss auf die Klassifizierung der Proben nach der Unabhängigen Komponentenanalyse zu. Mit der im Kapitel 3.3 vorgestellten hierarchischen Clusteranalyse konnten diese Metabolite jedoch eindeutig als Marker zur Unterscheidung des Wildtyps Col-0 von der stärkelosen PGM Mutante identifiziert werden.

Die Clusteranalyse erbrachte weiter, dass die Intermediate des photorespiratorischen Cyclus in beiden Genotypen einem circadianen Verlauf folgten (Abbildungen 3.3.1.3). Dieses wird durch die Faktorladungen beider dimensionsreduzierender Methoden widergespiegelt. Sowohl Glycin als auch Glycerat laden am stärksten auf die zweite Hauptkomponente (PC2) sowie auf die zweite Unabhängige Komponente (IC2), entlang derer eine Separation der Proben in Abhängigkeit von der Photoperiode erfolgte.

Anhand der Faktorladungen der ersten Unabhängigen Komponente wird weiter ein starker Zusammenhang zwischen den Metaboliten Asparagin, Arginin/Ornithin, Putrescin sowie Glycin und der Separation der Genotypen deutlich. Daneben trugen auch die drei Proteine (PSAC\_ARATH, At3g26650.1, sowie At3g27690.1) zur Gruppierung entlang dieser Komponente bei. Auch diese Metabolite und Proteine konnten mit den vorangegangenen datenanalytischen Methoden identifiziert werden (vgl. dazu Kapitel 3.2 und 3.3). Anhand der Faktorladungen konnte eine zusätzliche Verbindung identifiziert werden, welche deutlich zur Diskriminierung zwischen beiden Genotypen beitrug. Dabei handelte es sich um Suberylglycin, ein Amid der Korksäure, dessen Abundanz in der PGM Mutante im Vergleich zum korrespondierenden Wildtyp etwa 5-fach erhöht war.

Den Faktorladungen zufolge ist die Klassifizierung entlang der zweiten Haupt- (PC2) sowie Unabhängigen Komponente (IC2) vorrangig durch die Metabolitdynamik von Intermediaten des photorespiratorischen Cyclus dominiert. Auch dieses Ergebnis steht in Einklang mit den zuvor durchgeführten Auswertungen. Im Kapitel 3.2 wurden zudem drei Proteine identifiziert, deren Expressionsprofile die circadiane Rhythmik in der PGM Mutante reflektierten (At4g10340.1 und At3g27690.1, Chlorophyll-bindende Proteine; At2g30790.1, Photosystem II; Abbildung 3.2.1.2). Tatsächlich wiesen diese Kandidaten unter den Proteinen die weiteste Distanz vom Ursprung in Richtung der zweiten unabhängigen Komponente auf. Bei genauerer Betrachtung der Faktorladungen wird jedoch deutlich, dass diese Proteine im Vergleich zu den Metaboliten Glycerat und Glycin in einem weitaus geringerer Masse auf die zweite Komponente laden.

**Tabelle 5.3.1: Gegenüberstellung der Faktorladungen der PCA und ICA des integrativen Protein-Metabolit-Datensatzes**

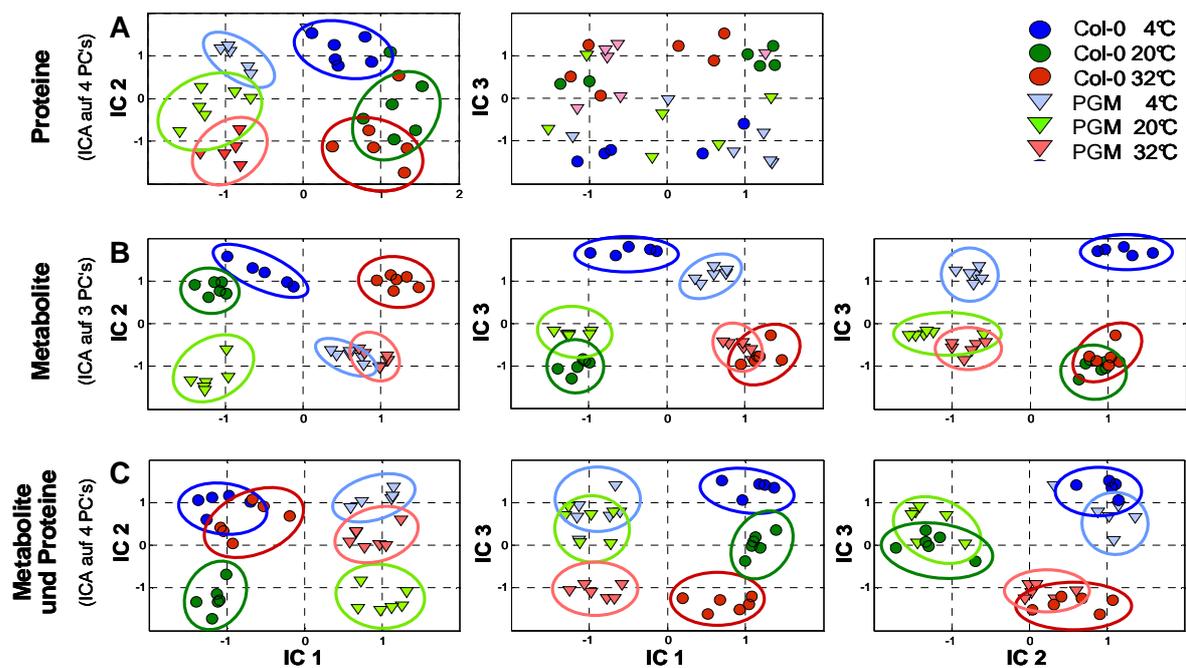
Hauptkomponentenanalyse		Unabhängige Komponentenanalyse	
Faktorladungen PC 1 Separation Col-0 - PGM	Faktorladungen PC 2 Separation Licht - Dunkel	Faktorladungen IC 1 Separation Col-0 - PGM	Faktorladungen IC 2 Separation Licht - Dunkel
-0.32 Fructose	0.34 Glycin	0.04 Suberylglycin	-0.17 Glycin
-0.22 Glucose	0.27 Glycerat	0.04 PSAC_ARATH	-0.16 Glycerat
-0.20 Glycin	0.21 Glucose	-0.04 At3g26650.1	0.08 Prolin
-0.19 Arginin/Ornithin	-0.20 Asparagin	-0.04 Arginin/Ornithin	0.07 Urea
0.19 At1g13930.1	0.19 Suberylglycin	0.04 Asparagin	0.07 Succinat
-0.18 Asparagin	0.18 Saccharose	0.03 Putrescin	-0.07 At5g25980.1
-0.18 CHO1	0.17 CHO1	0.03 At3g27690.1	-0.07 At3g27690.1
-0.17 Putrescin	0.17 At2g30790.1	0.03 Glycin	-0.06 At2g30790.1
0.16 At3g26650.1	-0.16 Lysin	-0.03 At1g13930.1	-0.06 Saccharose
0.15 PSAC_ARATH	0.15 CHO2	0.02 Fructose	-0.06 Pyruvat
-0.14 Citrullin	0.14 Fructose	0.02 Lysin	0.05 Glucose
-0.13 Indol 3-acetonitril	-0.13 Isoleucin	-0.02 Fumarat	-0.05 At1g13930.1
-0.12 CHO2	0.12 At3g26650.1	0.02 Isoleucin	-0.05 At3g27830.1
-0.12 At4g28750.1	0.12 At3g27830.1	0.02 Salicylat	-0.05 Methionin
-0.12 At3g27690.1	0.11 At1g06680.1	0.02 Indol 3-acetonitril	0.05 PSAC_ARATH
-0.12 Saccharose	0.11 Fumarat	0.02 Citrullin	0.04 Glutamat
-0.11 Salicylat	0.11 Glutamin	-0.02 At4g03280.1	-0.04 At4g10340.1
-0.09 Urea	-0.11 Arginin/Ornithin	-0.02 Prolin	0.04 Indol 3-acetonitril
-0.08 Phenylalanin	0.10 PSAC_ARATH	-0.02 At4g05320.3	-0.04 At1g06680.1
-0.08 Pyruvat	0.10 At4g05320.3	0.02 $\beta$ -Alanin	0.04 At3g16890.1
-0.07 At4g10340.1	0.09 Malat	-0.02 At4g02530.1	0.03 RBL_ARATH
-0.07 Methionin	0.09 At1g13930.1	0.02 Phenylalanin	0.03 <i>myo</i> -Inositol
-0.07 At5g25980.1	-0.09 $\beta$ -Alanin	-0.02 At2g30790.1	0.03 Asparagin
0.07 At4g03280.1	0.09 Serin	0.02 Valin	-0.03 Glucose 6-phosphat
0.07 At4g05320.3	-0.09 Indol 3-acetonitril	-0.02 Glutamin	0.03 Aspartat
0.06 At4g02530.1	0.09 Prolin	-0.02 At1g13440.1	-0.03 ATPB_ARATH
-0.06 Alanin	-0.08 Urea	0.01 At4g28750.1	-0.03 Alanin
-0.06 Pyroglutamat	0.08 Pyruvat	-0.01 Glucose	0.03 Suberylglycin
-0.05 $\beta$ -Alanin	-0.08 Valin	-0.01 Malat	0.02 At4g03280.1
0.05 At1g13440.1	-0.08 Putrescin	0.02 At3g26650.1	0.02 Ascorbat
-0.05 Glycerat	0.07 At4g02530.1	0.02 Pyroglutamat	0.02 At1g67090.1
0.05 <i>myo</i> -Inositol	-0.07 Leucin	-0.02 Serin	-0.02 Citrullin
0.05 Sinapat	0.07 At5g25980.1	-0.01 Sinapat	-0.02 Spermidin

#### 5.4 Mustererkennung am Beispiel der systemischen temperaturbedingten abiotische Stressresponse in *Arabidopsis thaliana*

Analog zum vorangegangenen Kapitel wurde die Qualität der Mustererkennung und Biomarkeridentifizierung durch die Anwendung dimensionsreduzierender, datenanalytischer Verfahren am Beispiel des Temperaturstress-Datensatzes erneut getestet. Dieser wies, bedingt durch die Anwendung des weiterentwickelten Proteinextraktionsprotokolls und einer neuartigen Methode zur Quantifizierung von Proteinen aus hochkomplexen Proben, eine deutlich höhere Dimensionalität auf. Auch hier lieferte die Unabhängige Komponentenanalyse sinnvolle Klassifizierungen der einzelnen Proben (Abbildung 5.4.1).

Die Neuordnung der Proben anhand der Proteindaten lässt eine Separation der Genotypen (Col-0 und PGM) entlang der ersten Unabhängigen Komponente, sowie die Separation der experimentellen Bedingungen (4°C, 20°C und 32°C) entlang der zweiten Unabhängigen Komponente erkennen (Abbildung 5.4.1-A).

Die ICA der Metabolitdaten resultierte hingegen in unterschiedlichen Gruppierungen der Proben. So offenbarte die zweidimensionale Darstellung des von der ersten und dritten Unabhängigen Komponente aufgespannten Raumes Ähnlichkeiten zwischen den jeweils bei 20°C und 32°C kultivierten Pflanzen beider Genotypen. Betrachtet man die Anordnung der Proben zwischen der zweiten und dritten unabhängigen Komponente, so überlappen die bei 20°C und 32°C geernteten Pflanzen des Wildtyps (Abbildung 5.4.1-B). Dieses erschwert die Interpretation der Unabhängigen Komponenten.



**Abbildung 5.4.1: PCA und ICA der Proteindaten (A), der Metabolitdaten (B) und des integrativen Protein-Metabolit-Datensatzes (C).** Abhängig vom zugrunde liegenden Originaldatensatz lieferte die Dimensionsreduktion durch eine Unabhängige Komponentenanalyse verschiedene Klassifikationen von temperaturgestressten *A. thaliana* Col-0 und PGM Proben. Während die Klassifizierung anhand der Proteindaten den genetischen Hintergrund der analysierten Genotypen entlang der IC1, sowie den Einfluss der Temperatur entlang der IC2 erkennen lässt (**Abb. 2.5.2.1-A**), verdeutlicht die Klassifizierung beruhend auf dem Metabolit-Datensatz, dass verschiedene Stresssituationen ähnliche Antworten im Metabolom der Pflanze hervorrufen können (**Abb. 2.5.2.1-B**). Die Klassifizierung basierend auf dem integrativen Protein-Metabolit-Datensatz liess sowohl individuelle, als auch den beiden Genotypen gemeinsame Stressantworten erkennen (**Abb. 2.5.2.1-C**).

Wurde jedoch der integrative Protein-Metabolit-Datensatz dem dimensionsreduzierenden Verfahren zugrunde gelegt, konnten die experimentellen Bedingungen durch nur drei Unabhängige Komponenten weitestgehend abgebildet werden, welche auch interpretiert werden konnten. So offenbarte die zweidimensionale Darstellung die Möglichkeit der Unterscheidung der Genotypen entlang der ersten, die zuvor im Kapitel 4 diskutierten inhärenten Stressreaktionen, welche gleichermaßen durch niedrige sowie durch hohe Temperaturen induziert wurden, entlang der zweiten, sowie den graduellen Einfluss der Temperatur entlang der dritten Unabhängigen Komponente (Abbildung 5.4.1-C). Somit konnten sowohl die individuellen Stressantworten der untersuchten Genotypen als auch generell durch Temperaturstress hervorgerufenen Veränderungen im pflanzlichen Stoffwechsel aufgelöst werden.

Die Anordnung der Proben in dem von der ersten und dritten Unabhängigen Komponente aufgespannten Datenraum offenbart die aufgrund der akkumulierten löslichen Carbohydrate Glucose, Fructose und Saccharose erwartete überlappende Anordnung der die bei 20°C und 4°C kultivierten Pflanzen repräsentierenden Datenpunkte. Die Analyse der erhaltenen Protein- und Metabolitdaten mit Hilfe der Varianzanalyse erbrachte, wie im Kapitel 4 diskutiert, dass einige Metabolite und Proteine in beiden Genotypen gleichermaßen reguliert wurden, wie z.B. Prolin und Glutamin (Abbildung 4.3.1.2), oder auch das Glycin-reiche RNA-bindende Protein At2g21660.1 (Abbildung 4.2.2.2), deren Konzentrationen mit steigender Temperatur abnahmen. Diese genotypübergreifende, universelle Response auf Temperaturstress wird durch die überlappende Anordnung der die einzelnen Pflanzen repräsentierenden Datenpunkte zwischen der zweiten und dritten Unabhängigen Komponente abgebildet.

Diese Beispiele zeigen, dass die Applikation der Unabhängigen Komponentenanalyse auf den integrativen Protein-Metabolit-Datensatz eine gegenüber der Hauptkomponentenanalyse deutlich verbesserte Methode zur Visualisierung hochdimensionaler Daten darstellt, wobei die inhärenten grundlegenden Strukturen des Originaldatensatzes durch nur wenige Komponenten adäquat abgebildet werden können.

## **5.5 Potenzielle Biomarker der systemischen temperaturbedingten abiotische Stressresponse in *Arabidopsis thaliana***

Aus dem Ranking der Faktorladungen der ICA, welche dem Einfluss eines jeden Metaboliten oder Proteins auf die beobachtete Klassifizierung der Proben in Richtung der jeweiligen Komponente proportional sind, liessen sich diverse potenzielle Biomarker für die jeweiligen Stresssituationen ableiten. Dabei zeigte sich, dass viele der im Kapitel 4 mit Hilfe der Varianzanalyse identifizierten Marker auch erheblich zur Gruppierung der

Datenpunkte entlang der entsprechenden Unabhängigen Komponenten beitragen. Dieses Ergebnis mag zunächst trivial erscheinen, zeigt jedoch aufgrund der vielfältigen Möglichkeiten der Transformation der durch unterschiedliche Technologien gewonnenen Daten die Eignung der  $\log_{10}$ -Transformation zur Identifizierung potenzieller Biomarker durch die dimensionsreduzierenden Techniken PCA und ICA (einige Beispiele des Einflusses der Datentransformation auf Gruppierung der Variablen und die Wichtung der Variablen sind im Anhang A-3 aufgeführt).

So weisen die hohen Faktorladungen der zweiten Unabhängigen Komponente auf einen starken Zusammenhang zwischen den Konzentrationen von Raffinose sowie Galactinol und einer allgemeinen Stressresponse in beiden Genotypen hin (Tabelle 5.5.1). Dieses Ergebnis korreliert mit den Ergebnissen der Varianzanalyse, welche erbrachte, dass die Galactoside Raffinose und Galactinol sowohl in den kälte-, als auch in den hitzestressierten Blattgeweben der beiden analysierten *A. thaliana* Genotypen Col-0 und PGM akkumulierten. Wenngleich sich der Einfluss der verbliebenen Variablen auf die Separation entlang der zweiten Unabhängigen Komponente als wesentlich geringer darstellte (zum Vergleich: Faktorloadings IC2 Galactinol -0.13, At2g44650.1 -0.05), wurde dem chloroplastidiären Protein At2g44650.1 mit Chaperonin-Funktion (Cpn10) die höchste Faktorladung auf der Ebene der Proteine zugeordnet. Wenngleich über die genaue Funktion dieses Proteins in Pflanzen bisher nur wenig bekannt ist, deuten jüngste Arbeiten darauf hin, dass dieses Protein in die Temperaturstressresponse involviert ist. So stellten Stocchi *et al.* (2006) durch Untersuchungen der Kältstressresponse in *E. coli* einen Zusammenhang zwischen einem dem At2g44650.1 homologen Protein und der Adaption an niedrige Temperaturen her [359]. Der Kruskal-Wallis-Test offenbarte jedoch keine mit  $p \leq 0.05$  signifikanten Unterschiede in der Expression dieses Proteins zwischen den gestressten und ungestressten Pflanzen der analysierten Genotypen (vergleiche dazu Abbildung 5.5.1). Dieses Ergebnis zeigt, dass das Ranking der Faktorladungen bei unterschreiten eines gewissen Absolutwertes, welcher jedoch im Einzelnen mit Hilfe zusätzlicher statistischer Verfahren wie z.B. T-Test oder ANOVA ermittelt werden muss, eher willkürlich erfolgt. Die Faktorladungen an sich beinhalten keine Informationen über die Signifikanz des Beitrages der einzelnen Variablen auf die beobachtete Separation der Datenpunkte.

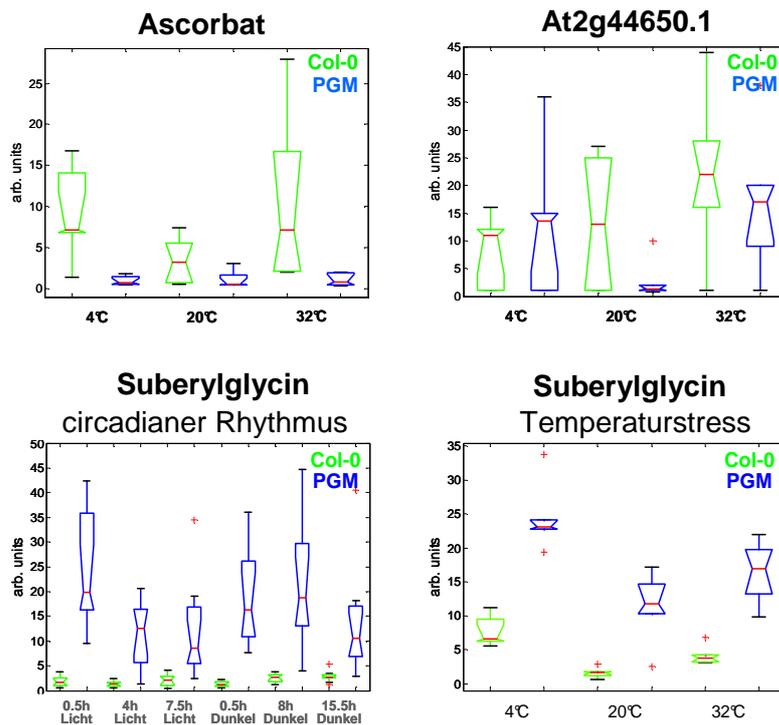
Die Faktorladungen der dritten Unabhängigen Komponente, entlang derer der graduelle Einfluss der Temperatur auf den pflanzlichen Metabolismus deutlich wird, offenbaren Stressantworten beider Stoffwechselebenen. Neben Prolin, welches als kompatible Substanz gilt, und Glutamin offenbaren die Faktorladungen eine spezifische Response des Glycin-reichen RNA-bindenden Proteins At2g21660.1, der Hydroperoxid Reduktase At1g65980.1 sowie der zuvor als nach Hitzestress spezifisch exprimiert identifizierten

kerncodierten Ribonucleoproteine mit chloroplastidiärer Signalsequenz (At2g37220.1 und At3g53460.1).

**Tabelle 5.5.1: Faktorladungen der ICA des integrativen Protein-Metabolit-Datensatzes**

Faktorladungen IC 1		Faktorladungen IC 2		Faktorladungen IC 3	
Separation Col-0 - PGM		allgemeine Stressantwort		spezifische Stressantwort	
0.07	Asparagin	-0.13	Raffinose	0.06	Prolin
-0.06	Ascorbat	-0.12	Galactinol	0.06	At2g21660.1
0.05	Suberylglycin	-0.05	At2g44650.1	-0.05	At1g65980.1
0.05	At1g19570.1	-0.04	At1g65980.1	0.04	Glutamin
0.05	Fructose	0.04	Glycerat	0.04	At3g53460.1
0.04	Agmatin	-0.04	Prolin	0.04	At2g37220.1
0.04	At3g49110.1	-0.04	Glycerolphosphat	-0.04	At3g26450.1
0.04	Glucose	-0.04	Fructose	-0.04	Raffinose
0.04	At1g65980.1	-0.04	Tryptophan	0.03	Citrullin
-0.04	At2g35370.1	-0.04	Suberylglycin	-0.03	Galactinol
0.04	3-Methyl-3-hydroxyglutarat	-0.03	Fructose	-0.03	At1g19570.1
0.04	Arginin/Ornithin	-0.03	At5g52310.1	0.03	Glycin
-0.04	At3g63490.1	-0.03	Glutamin	-0.03	At2g21130.1
-0.03	Raffinose	0.03	At1g74470.1	0.03	Fructose
-0.03	At5g14910.1	0.03	At3g01480.1	-0.03	At1g76080.1
-0.03	At5g15970.1	-0.03	Ascorbat	0.03	2-Ethyl 3-hydroxy 3-methyl-valerat
0.03	Alanin	-0.03	At4g04020.1	-0.03	Glutarat
-0.03	At5g08280.1	-0.03	O-Acetylserin	-0.03	At2g44650.1
0.03	At1g08830.1	-0.03	Agmatin	-0.03	At3g01480.1
-0.03	At1g03680.1	-0.03	4-Aminobutyrat	0.03	At3g52930.1
-0.03	At1g03600.1	-0.03	At3g26450.1	0.03	At3g52930.1
0.03	Saccharose	-0.03	At2g21130.1	-0.02	At4g23670.1
-0.03	Gakactinol	-0.03	Glucose	0.02	At5g15970.1
0.03	Citrullin	0.03	At4g28750.1	0.02	At5g14910.1
0.03	At5g28540.1	-0.02	3-Methyl-3-hydroxyglutarat	0.02	At5g19510.1

Zur Separation entlang der ersten Unabhängigen Komponente, welche die genetischen Hintergründe der analysierten *A. thaliana* Genotypen reflektiert, trugen in besonderem Masse Asparagin und Ascorbat bei. Auch Thimm *et al.* (2004) stellten in ihrer vergleichenden Studien der Genexpression von *A. thaliana* Col-0 und PGM fest, dass die stärkelose Mutante durch eine Hochregulation der in die Argininbiosynthese involvierten Gene gekennzeichnet ist [360]. Somit stellt dieser Metabolit einen potenziellen Marker für die Mutation des chloroplastidiär codierten Phosphoglucomutasegens dar. Daneben indiziert Ascorbat einen erhöhten oxidativen Stressstatus in der Mutante. Aufgrund der leichten Oxidierbarkeit dieser Verbindung zeigen die aus den jeweils 6 biologischen Replikaten resultierenden Daten jedoch eine erhebliche Streuung (Abbildung 5.5.1). Zudem kann nicht ausgeschlossen werden, dass die in der Mutante erniedrigten Konzentrationen auf eine matrixvermittelte Oxidation während der Extraktion zurückzuführen ist. Dieses könnte beispielsweise durch eine Standardaddition während der Extraktion geprüft werden. Dennoch deuten auch die Faktorladungen der in der Detoxifizierung reaktiver Sauerstoffspezies beteiligten Enzyme At1g19570.1, ein Isoenzym der Dehydroascorbat Reduktase, sowie die Peroxidase At3g49110.1 auf einen erhöhten Stress-Status in der PGM Mutante hin (vergleiche dazu auch Abbildung 4.2.2.2).



**Abbildung 5.5.1: Konzentrationsprofile ausgewählter, anhand der Faktorladungen der ICA des integrativen Protein-Metabolit-Datensatzes identifizierter potenzieller Biomarker von *A. thaliana*.** Die hohe Faktorladung des Ascorbats der IC1 kann als Hinweis auf eine erhöhten oxidativen Stressstatus im Wildtyp gedeutet werden. Die Konzentrationsprofile der jeweils 6 biologischen Replikate zeigen jedoch, dass diese aufgrund der leichten Oxidierbarkeit der gaschromatographischen Analyse nur schwer zugängliche Verbindung eine beträchtliche Streuung der Messwerte aufwies. Daher können Matrixeffekte als Ursache für die erhöhten Ascorbatkonzentrationen im Wildtyp nicht ausgeschlossen werden.

Das chloroplastidiäre Protein At2g44650.1 mit Chaperonin-Funktion (Cpn10) zeigte eine relativ hohe Faktorladung der IC2. Jedoch zeigt diese Abbildung, dass zwischen den ermittelten, relativen Konzentrationen der jeweiligen Stichproben keine signifikanten Unterschiede existieren. Sowohl bei der Analyse der zeitabhängigen systemischen Regulation des Stoffwechsels, als auch bei der Analyse der Temperaturstressresponse fiel Suberylglycin durch hohe Faktorladungen der die Genotypen separierenden Komponente auf. Die Konzentrationen dieses Korksäureamids war in der PGM Mutante reproduzierbar etwa 5-fach erhöht.

Erneut fiel im Zusammenhang mit der Separation zwischen den analysierten Genotypen Col-0 und PGM das Korksäureamid Suberylglycin durch signifikant erhöhte Konzentrationen in der PGM auf (Abbildung 5.5.1). Offenbar stellt dieser Metabolit, über dessen Biogenese in Pflanzen bisher in der Literatur keine Hinweise zu finden sind, ein wichtiges Intermediat des pflanzlichen Stoffwechsels dar, welches womöglich in der Signaltransduktion der Carbohydrate involviert ist.

## 5.6 Diskussion

Die vergleichende Analyse verschiedener Strategien zur Mustererkennung erbrachte, dass diese durch die Integration von Datensätzen, welche komplementäre Stoffwechselebene repräsentieren, deutlich verbessert werden kann. An zwei unabhängigen Datensätzen konnte gezeigt werden, dass sich die Co-Varianz der Metabolite sowie der Proteine im

Einzelnen in einer individuellen Gruppierung von Proben mit unterschiedlichen genetischen Hintergründen und/oder physiologischen Stati manifestiert, welche nur bedingt interpretiert werden können. Dieses macht zugleich die Identifizierung von Biomarkern unmöglich.

Der integrative Protein-Metabolit-Datensatz konnte hingegen mit den dimensionsreduzierenden Techniken PCA und ICA so prozessiert werden, dass sowohl die genetischen Hintergründe als auch die experimentellen Bedingungen zweidimensional projiziert werden konnten. Dieses Ergebnis zeigt das Potenzial der integrativen Analyse umfangreicher Datensätze verschiedener Stoffwechselebenen, welche mit den heute zur Verfügung stehenden instrumentellen Analysemethoden schnell erhoben werden können. Die integrative Analyse multiparallel gewonnener Daten kann somit Hinweise auf Genfunktionen liefern und ermöglicht die Identifikation von Targets für eine gezielte genetische Manipulation. Dabei gilt es jedoch zu berücksichtigen, dass sowohl die PCA als auch die ICA durch individuelle Vor- und Nachteile gekennzeichnet sind, so dass ein Vergleich der Resultate beider Techniken unabdingbar erscheint. Insbesondere hinsichtlich der Identifikation von Biomarkern stehen die dimensionsreduzierenden Techniken alternativen Verfahren wie etwa der Varianzanalyse im Einzelfall jedoch nach, da Variablen, welche geringe jedoch signifikante Unterschiede in ihrer Abundanz aufweisen, nach deren Transformation in den neu berechneten Unterraum häufig von den Variablen, welche grosse Unterschiede zwischen zwei oder mehreren Stichproben aufweisen, dominiert werden, so dass sich anhand der Faktorladungen ausschliesslich die Variablen identifizieren lassen, welche als Response auf einen Reiz die grössten Unterschiede in ihrer Abundanz aufweisen. Dieses stellt eine offensichtliche Schwachstelle dieser Verfahren dar, da insbesondere den Sekundärmetaboliten und Phytohormonen trotz ihrer geringen Abundanz eine entscheidende Bedeutung in der Regulation des pflanzlichen Stoffwechsels zukommt und schon geringe Konzentrationsänderungen einen weit reichenden Einfluss auf das Stoffwechselgeschehen haben können.

Eine eher auf das Auffinden von Unterschieden als von Ähnlichkeiten beruhende Methode stellt die multivariate Varianzanalyse dar. Im Gegensatz zu den dimensionsreduzierenden Techniken, deren Ziel die Reduktion einer Vielzahl von Variablen auf einige wenige latente Komponenten darstellt, erlaubt dieses Verfahren die Identifizierung aller sich zwischen zwei oder mehreren Stichproben signifikant ändernden Variablen, jedoch können die hochdimensionalen Daten nicht visualisiert werden. Da beide datenanalytischen Methoden durch den Verlust von Informationen gekennzeichnet sind, stellt die Kombination dieser Verfahren für die Identifikation von Biomarkern das Mittel der Wahl dar.

## 6 Gerichtete quantitative Proteomics

In den vorangegangenen Kapiteln wurden zwei ungerichtete, Shotgun-basierte Methoden zur Erhebung quantitativer Proteindaten vorgestellt. Neben der Quantifizierung von Proteinen über die Peakintegration der Vorläuferionenspuren tryptischer Peptide (Kapitel 3.2) wurden Proteine auch über die kumulative Anzahl detektierter Peptide (Kapitel 4.2) quantifiziert. Beide Methoden stellen ungerichtete Verfahren dar.

Mit der im Kapitel 1.2 vorgestellten Performance von Triple-Quadrupol Massenspektrometern erscheinen auch gerichtete Proteinanalysen realisierbar. Diese Möglichkeit der Proteinanalyse wurde an einem in den folgenden Absätzen vorgestellten Experiment getestet.

### 6.1 Strategie der gerichteten Proteinanalyse

Durch die Möglichkeit des Einsatzes der zwei Messquadrupole eines Triple-Quadrupol Massenspektrometers als Massenfilter lassen sich auch aus hochkomplexen Proben spezifische Zielanalyten mit bekanntem Fragmentierungsverhalten sehr selektiv analysieren. Diese auch als Multiple Reaction Monitoring bekannte Methode zur Durchführung von MS<sup>2</sup>-Experimenten wurde in Anlehnung an Arbeiten von Glinski *et al.* zur Analyse posttranslationaler Modifikationen [57,361] zur Quantifizierung der cytosolischen Fructose 1,6-bisphosphatase At1g43670.1 aus hochkomplexen Peptidproben, welche aus dem tryptischen Verdau von löslichen Gesamtproteinfraktionen aus *A. thaliana* Col-0 Blattmaterial resultierten, getestet. Dazu wurden zunächst die gerätetechnischen Parameter Tube Lens, Sheath Gas Pressure, Auxillary Gas Pressure, Spray Voltage sowie Transfer Capillary Temperature für die Detektion von synthetisch hergestellten Peptiden mit den in Tabelle 6.1.1 angegebenen Sequenzen optimiert (siehe auch Kapitel 8.5.3). Im

Anschluss wurde die optimale Kollisionsenergie für die Fragmentierung dieser Peptide ermittelt und die resultierenden Fragmentionen erfasst (Tabelle 6.1.1).

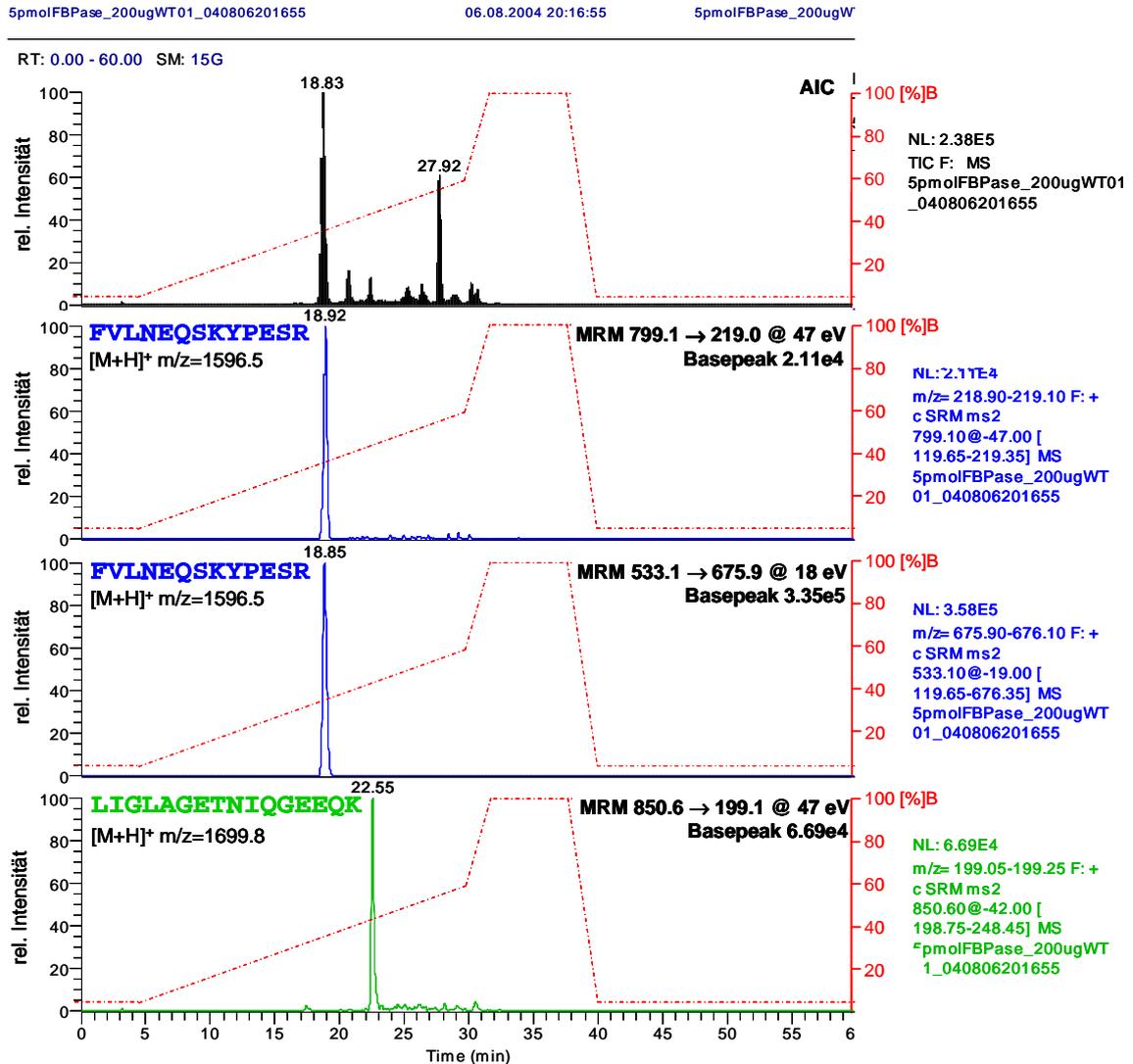
**Tabelle 6.1.1: Peptidzielsequenzen und deren massenspektrometrischen Analyse mittels Multiple Reaction Monitoring zugrunde liegende Parameter.** MS<sup>2</sup>-Experimente, welche nach Standardaddition zu 200 µg Gesamtprotein aus Blattmaterial von *A. thaliana* gute Linearitäten aufwiesen, sind blau gekennzeichnet.

Peptid Sequenz	detektierte Vorläuferionen (m/z)	Kollisionsenergie (eV)	detektierte Produktionen (m/z)
<b>FVLNEQSKYPESR</b> <i>m/z</i> [M+H] <sup>+</sup> = 1596.5	<b>799.1</b> [M+2H] <sup>2+</sup>	<b>47</b> 57	<b>219.0</b> 120.0
	<b>533.1</b> [M+3H] <sup>3+</sup>	<b>19</b> 43	<b>675.9</b> 120.0
<b>LIGLAGETNIQGEEQK</b> <i>m/z</i> [M+H] <sup>+</sup> = 1699.8	<b>850.6</b> [M+2H] <sup>2+</sup>	<b>48</b> 43	<b>199.1</b> 86.1
	567.4 [M+3H] <sup>3+</sup>		

Um eine eindeutige Identifizierung zu gewährleisten, wurde das Retentionsverhalten der synthetischen Peptide an einer C18 reversed phase Säule ermittelt (Abbildung 6.1.1). Zur Erstellung von Kalibriergeraden erfolgte zunächst eine Standardaddition von gemäss Kapitel 8.3.3 phenolisch mit Saccharose extrahierten und gemäss 8.5.1 proteolytisch verdauten löslichen Gesamtproteinfraktionen aus *Arabidopsis thaliana* Col-0 Blattmaterial nach Erreichen des Entwicklungsstadiums 5.1 [177]. Dazu wurden 200 µg extrahiertes Gesamtprotein nach dem proteolytischen Verdau mit 5, 25, 50, 100 und 150 pmol des jeweiligen synthetischen Peptides versetzt und die erhaltenen Proben massenspektrometrisch über die in Tabelle 6.1.1 angegebenen MRM-Experimente analysiert.

Gute Linearitäten wiesen die Resultate der MRM-Experimente des zweifach geladenen Vorläuferpeptides FVLNEQSKYPESR bei einer Kollisionsenergie von 47 kV, was in dem Fragment mit *m/z* 219 resultierte, die Fragmentierung des dreifach geladenen Vorläuferpeptides FVLNEQSKYPESR bei 19 eV, was in dem Fragment mit *m/z* 675.9 resultierte, sowie die Fragmentierung des in einem Fragment mit *m/z* 199.1 resultierenden zweifach geladenen Vorläuferpeptides LIGLAGETNIQGEEQK bei einer Kollisionsenergie von 46 kV auf (Abbildung 6.1.2). Die Analyse erfolgte in Triplika.

Im Anschluss daran wurden aus Blattmaterial von *A. thaliana* Col-0 nach Erreichen des Entwicklungsstadiums 5.1 [177] extrahierte Gesamtproteinfraktionen nach gleichem Vorgehen analysiert. Anhand der erhaltenen Signalintensitäten der bei ca. 18.8 min und



**Abbildung 6.1.1: Chromatogramme der MRM-Experimente nach Standardaddition synthetischer Peptide der cytosolischen FBPase aus *A. thaliana*.** Dargestellt sind die aus der Addition von jeweils 5 pmol des jeweiligen synthetischen Peptides zu 200 µg proteolytisch verdautem löslichen Gesamtprotein aus *A. thaliana* Col-0 Blattmaterial resultierenden Signalintensitäten der in Tabelle 6.1.1 angegebenen MRM-Experimente.

22.5 min eluierten Verbindungen gelang es, tryptische Peptide aus hochkomplexen Gesamtproteinfraktionen nach dem proteolytischen Verdau zweifelsfrei als Fragmente der FBPase zu identifizieren (Abbildung 6.1.3). Ein Abgleich der aus den Analysen der extrahierten Blattproteinfraktionen resultierenden Signalintensitäten mit den mittels Standardaddition der synthetischen Peptide erstellten Kalibriergeraden erbrachte, dass die in der Pflanze zum Erntezeitpunkt tatsächlich vorliegende Konzentration der FBPase deutlich unter 25 pmol /g Gesamtprotein lag.

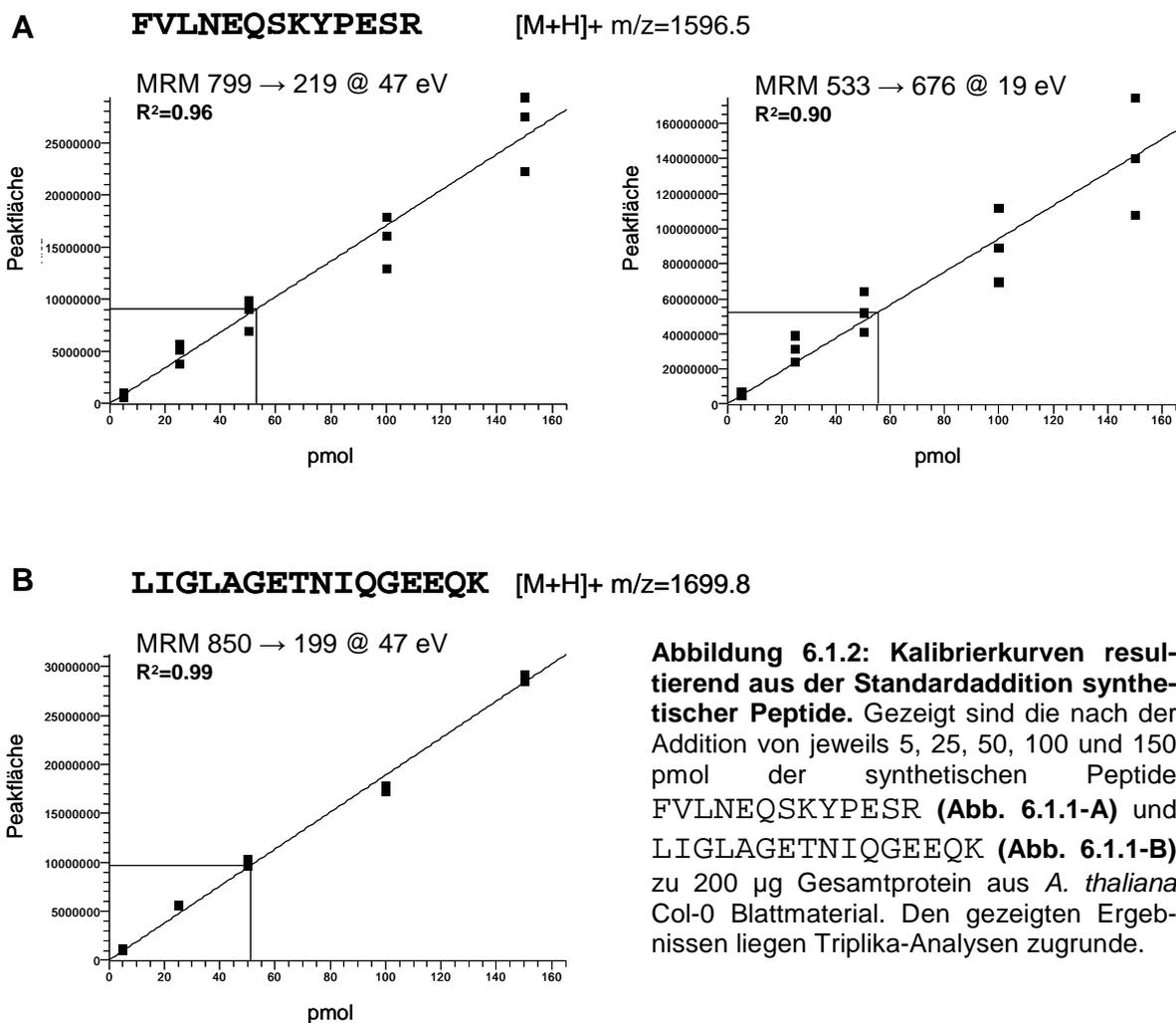
Die ermittelten Konzentrationen für die jeweiligen Peptide betragen im Einzelnen für die Sequenz **FVLNEQSKYPE**SR :

MRM 799.1 [M+2H]<sup>2+</sup> → 219.0 @ 47 eV: 0.22 pmol/200µg Gesamtprotein ± CV 37.4 %

MRM 533.1 [M+3H]<sup>3+</sup> → 675.9 @ 19 eV: 0.27 pmol/200µg Gesamtprotein ± CV 9.4 %

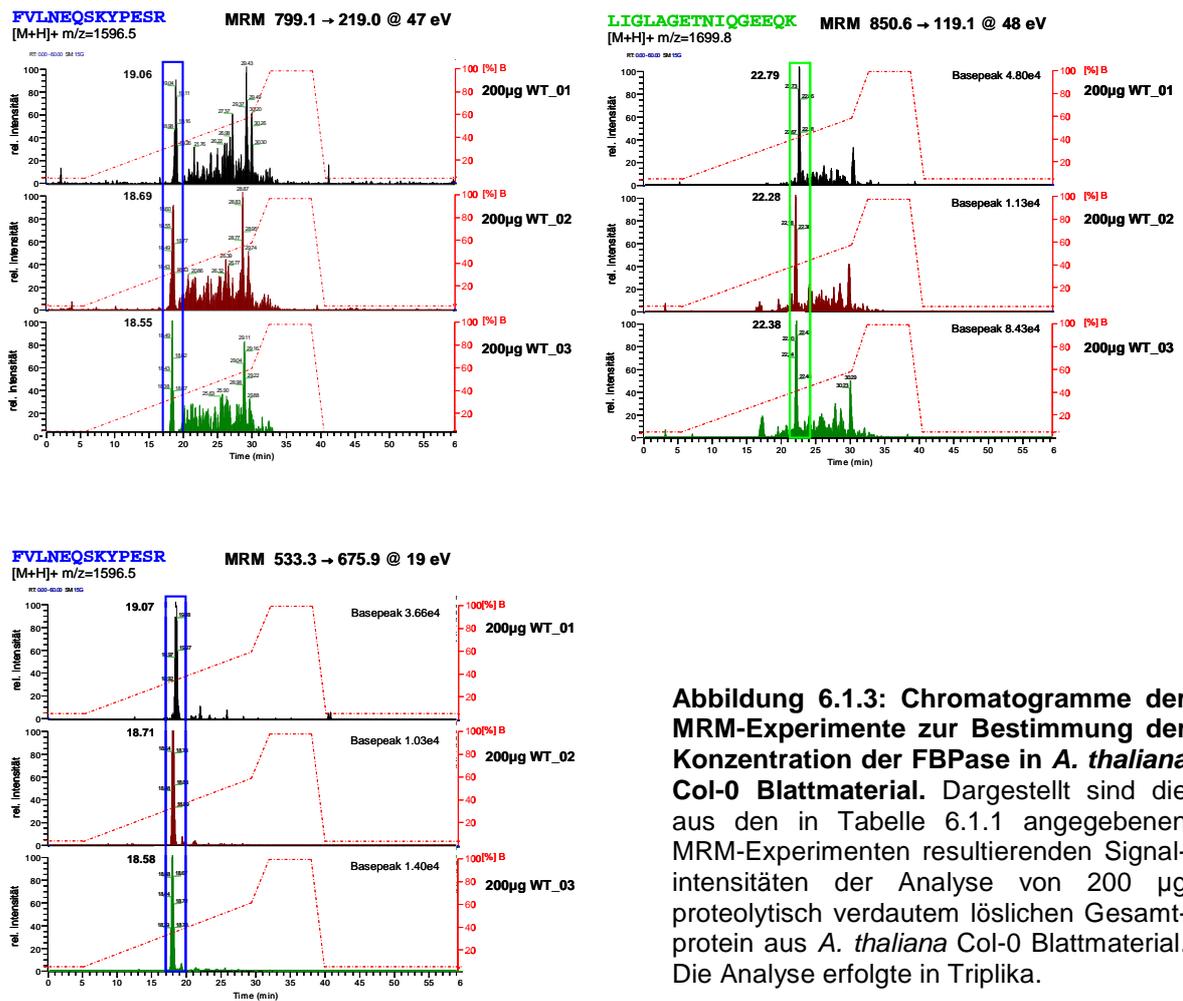
für die Sequenz **LIGLAGETNIQ**GEEQK :

MRM 533.1 [M+2H]<sup>2+</sup> → 199.1 @ 48 eV: 0.79 pmol/200µg Gesamtprotein ± CV 10.9 %



## 6.2 Diskussion

Die durch den Abgleich der aus der Analyse von aus Blattproben isolierten Proteinfractionen resultierenden Signalintensitäten mit den zuvor erstellten Kalibriergeraden erbrachten eine Konzentration der cytosolischen FBPase von unter 5pmol/mg extrahiertem Gesamtprotein.



**Abbildung 6.1.3: Chromatogramme der MRM-Experimente zur Bestimmung der Konzentration der FBPase in *A. thaliana* Col-0 Blattmaterial.** Dargestellt sind die aus den in Tabelle 6.1.1 angegebenen MRM-Experimenten resultierenden Signalintensitäten der Analyse von 200 µg proteolytisch verdautem löslichen Gesamtprotein aus *A. thaliana* Col-0 Blattmaterial. Die Analyse erfolgte in Triplika.

In Anbetracht der Tatsache, dass die Kalibrierkurven zur Konzentrationsbestimmung der cytosolischen FBPase aus *A. thaliana* Blattmaterial durch eine Standardaddition synthetischer Peptide mit Sequenzhomologie, jedoch ohne Isotopenmarkierung erfolgte, muss die tatsächlich vorliegende Konzentration dieses Enzyms weitaus geringer sein. Eine Literaturrecherche zur Ermittlung eines Vergleichswertes der in *A. thaliana* üblicherweise vorliegenden Konzentrationen dieses Enzyms erbrachte leider keinen Referenzwert.

Wenngleich die exakte Konzentrationsbestimmung der in *Arabidopsis thaliana* Col-0 Blattmaterial nach Erreichen des Entwicklungsstadiums 5.1 [177] vorliegenden Konzentrationen der FBPase aufgrund der Standardaddition verschiedener Konzentrationen unmarkierter synthetischer Peptide nicht möglich ist, zeigt das Ergebnis dieses Versuches das Potenzial dieser Vorgehensweise für die gerichtete quantitative Analyse ausgewählter Proteine. Durch den Einsatz  $^{15}\text{N}$ - oder  $^{13}\text{C}$ -isotopenmarkierter Peptide kann diese Limitierung jedoch so umgangen werden, dass nach der Aufnahme von Kalibrierkurven basierend auf der Signalintensität der isotopenmarkierten

Referenzpeptide durch Spiken in die Matrix die tatsächlich in der zu analysierenden biologischen Probe vorliegenden Proteinkonzentrationen anhand der Signalintensitäten der unmarkierten Peptide abgeschätzt werden kann. Durch die hohe Selektivität, welche aus dem Betrieb der beiden Messquadrupole eines Triple-Quadrupol Massenspektrometers als selektive Massenfilter resultiert, ist dieses auch aus hochkomplexen biologischen Proben ohne aufwendige Fraktionierungen möglich. Zusätzlich kann eine entscheidende Verbesserung der Sensitivität durch den Einsatz der nanoHPLC erreicht werden. Zusammenfassend bietet dieses, wenn auch relativ kostengünstige, Vorgehen die Möglichkeit, die Konzentrationen ausgewählter Proteine aus komplexen biologischen Proben ohne eine aufwendige Probenvorbereitung oder Fraktionierung nicht nur relativ, sondern auch absolut zu bestimmen und stellt somit einen zukunftsorientierten Ansatz der modernen Bioanalytik dar.

## 7 Zusammenfassung und Ausblick

Wesentliche Fortschritte in der instrumentellen Analysetechnik erlauben heute die simultane Messung einer Vielzahl von Metaboliten oder Peptiden aus komplexen biologischen Proben. Der geeignete Umgang mit den anfallenden grossen Datenmengen und deren Interpretation stellen gegenwärtig die wohl grössten Herausforderungen der Systembiologie, deren Ziel es ist, über einen interdisziplinären Ansatz mit den Methoden und Konzepten der Molekularbiologie, der Systemwissenschaften und der Informatik ein integriertes Bild über die regulatorischen Prozesse aller Stoffwechselebenen zu bekommen, dar. Die methodischen Ansätze der Systembiologie unterteilen sich in den Top-down-Ansatz, mit dessen Hilfe komplexe, phänotypische Beobachtungen in immer detailliertere Teilprozesse zerlegt werden, und den Bottom-up-Ansatz, der auf dem Zusammenführen biologischer Basisprozesse zu komplexen Einheiten beruht.

Die unverzichtbare Voraussetzung für systembiologische Ansätze zur Stoffwechselanalytik stellt die zuverlässige Identifizierung und Quantifizierung möglichst vieler Intermediate und Produkte dar. Dieser auch als Schwerpunkt der vorliegenden Arbeit bearbeitete Sachverhalt erfordert neben der fundierten Kenntnis über chemische Reaktionen ein hohes technisches Verständnis der Funktionsweise der zahlreichen, heute verfügbaren Gerätelösungen. Durch den Einsatz moderner Massenspektrometer wie das Leco Pegasus II TOF System (LECO Corporation, St. Joseph, MI, USA), das Linear Ion Trap (LTQ) Massenspektrometer (ThermoElectron Corporation, San Jose, CA, USA) und das Triple Stage Quadrupole (TSQ) Discovery MAX Massenspektrometer (Fa. ThermoElectron Corporation, San Jose, CA, USA) konnten im Rahmen dieser Arbeit 80 strukturbekannte Metabolite sowie 188 Proteine aus komplexen biologischen Proben von *A. thaliana* eindeutig identifiziert und deren Konzentrationen relativ bestimmt werden. Mit Hilfe datenanalytischer Verfahren gelang es, mit physiologischen Bedingungen assoziierte

Veränderungen des Stoffwechselgeschehens der Pflanze aufzudecken und eine Vielzahl potenzieller Biomarker und Zielanalyten für weiterführende Untersuchungen zur Organisation des pflanzlichen Stoffwechsels zu identifizieren.

An unabhängigen Beispielen wurde das Potenzial der Korrelationsanalyse sowie das der integrativen Analyse von Metabolit- und Proteindaten im Hinblick auf den Top-down-Ansatz aufgezeigt. Eine zusammenfassende Übersicht über die dabei verfolgten Strategien und die daraus resultierenden Konsequenzen bezüglich der Interpretation ist in Abbildung 6.1 wiedergegeben.

Die Voraussetzung für diese Analysen stellte die Etablierung robuster, sensitiver massenspektrometrischer Methoden zur relativen Quantifizierung von Metaboliten mittels GC-TOF-MS, sowie zur LC-MS basierten relativen indirekten Quantifizierung von Proteinen über korrespondierende Peptide aus komplexen Proben dar. Die entwickelten Methoden sind problemlos für die Analyse biologischen Materials jeglicher Herkunft anwendbar.

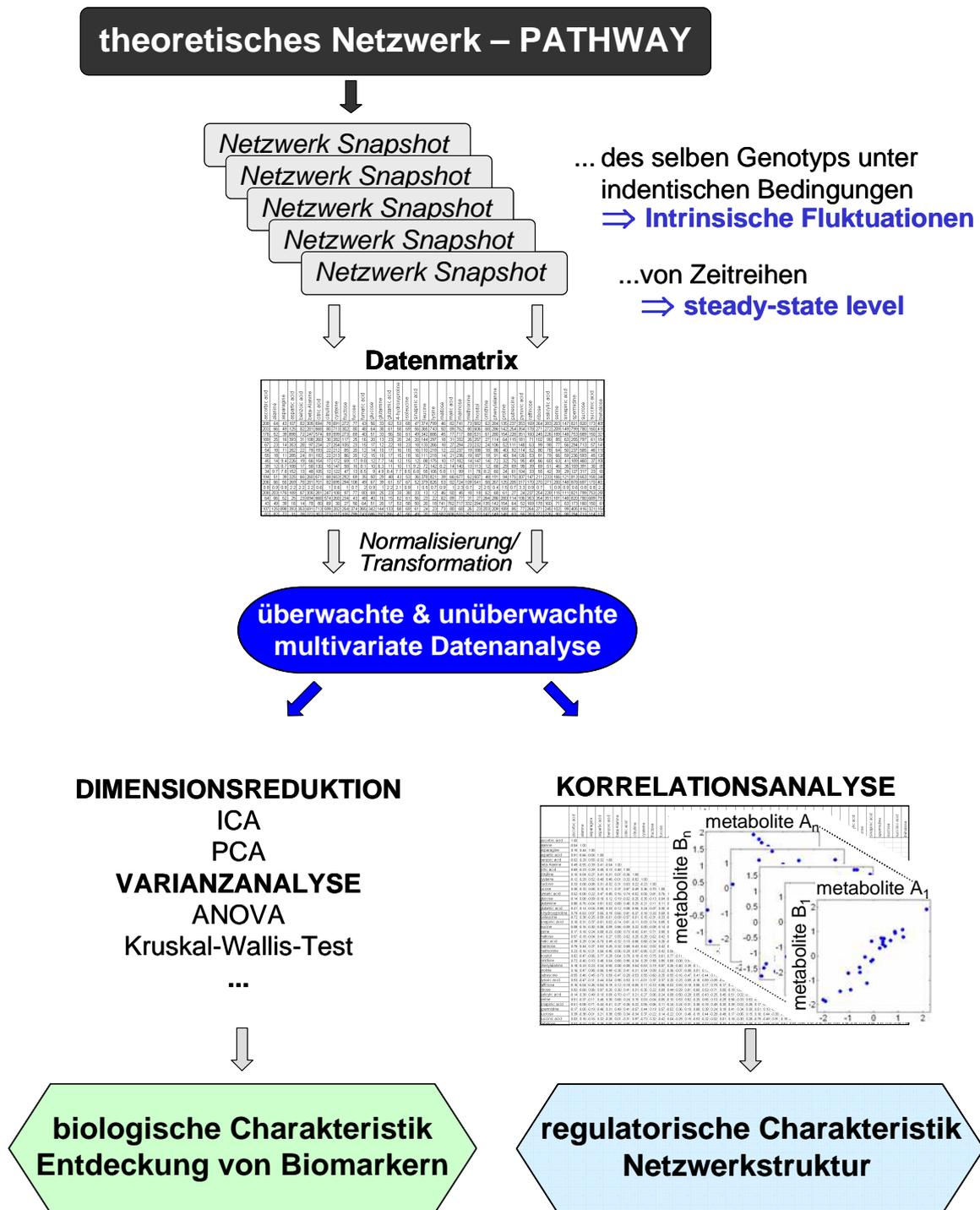
Die auf dem Metabolite Profiling beruhenden vergleichenden Analysen des Stoffwechsels verschiedener Organismen erfassen zumeist steady-state Bedingungen und reflektieren nicht den metabolischen Fluss durch das metabolische Netzwerk. Um Veränderungen des zugrunde liegenden enzymatischen Netzwerkes aufzudecken, welche die Regulation des Stoffwechsels beeinflussen, sind jedoch Informationen über die metabolischen Flüsse essentiell. Die Analyse metabolischer Flüsse erfordert streng kontrollierbare biologische Systeme, in denen die zu untersuchenden Organismen jedoch ihrer natürlichen Umgebung entzogen sind wie z.B. Mikroorganismen- oder Zell- und Gewebekulturen. Diese Systeme sind jedoch immer artifiziell, so dass die *in vivo* Dynamik damit nicht bestimmbar ist. Daher besteht eine grosse Herausforderung darin, Veränderungen in der Regulation enzymatischer Reaktionen auch *in vivo* zu ermitteln.

Frühere Arbeiten von Arkin *et al.* zeigten, dass zum Verständnis der Organisation metabolischer Netzwerke anhand von steady-state Messungen stochastische Modelle unerlässlich sind [362,363]. Von diesen Studien und vorangegangenen Arbeiten von Weckwerth *et al.* und Steuer *et al.*, in denen der Zusammenhang zwischen der Korrelation von Metabolitpaaren mit dem zugrunde liegenden Netzwerk hergeleitet wurde [98,165,189,364,365], inspiriert, wurden im Rahmen dieser Arbeit verschiedene Korrelationsnetzwerke speziesübergreifend analysiert. Dazu wurden die linearen Zusammenhänge der mittels GC-TOF-MS gewonnen experimentellen Daten mit numerisch erzeugten Korrelationsprofilen verglichen. Dabei gelang der experimentelle Nachweis, dass minimale Fluktuationen extrinsischer Parameter, im speziellen Falle der Lichtenergie, die Konzentrationen einiger weniger Metabolite direkt beeinflussen und diese Fluktuationen spezifische Korrelationen erzeugen. Daneben konnten durch die

vergleichende, speziesübergreifende Datenanalyse von Metabolitprofilen regulatorische Veränderungen und potenzielle Knotenpunkte der enzymatischen Netzwerke der drei Spezies *A. thaliana*, *N. tabacum* und *S. tuberosum* identifiziert werden.

In jüngster Zeit kommt dem Gebiet der 'Functional Genomis' eine immer grössere Bedeutung zu. Die drei Schlüsseltechnologien dieses Forschungsgebietes, Transcriptomics, Proteomics und Metabolomics, sind durch individuelle Charakteristika gekennzeichneten, welche jede einzelne dieser 'Omics'-Technologien zur Beantwortung einer speziellen Fragestellung mehr oder weniger geeignet erscheinen lassen. Dieses liegt vor allem in der geringen Korrelation zwischen diesen unterschiedlichen Stoffwechselebenen begründet. Mit zunehmender Distanz der analysierten Ebene vom Phänotyp verringert sich die Chance, relevante Targets für die gezielte Veränderung von metabolischen Schwachstellen zu identifizieren. Somit kommt der Metabolitanalyse im Hinblick auf Gentechnik eine entscheidende Bedeutung zu. Gleichzeitig sind an die Interpretation von Metabolitdaten besondere Herausforderungen gestellt. Da zwischen dem Metabolom und dem Genom nur eine indirekte Beziehung besteht, ist eine Vorhersage, welches Gen manipuliert werden muss, um den gewünschten knock-out-Effekt oder die Überexpression eines bestimmten Genproduktes zu erzielen, äusserst schwierig.

Neben der Identifizierung von Targets für die gezielte Manipulation von Organismen stellt auch die Entdeckung von Biomarkern zur Charakterisierung eines physiologischen oder pathologischen Zustandes sowie der pharmakologischen Antwort eines Organismus auf einen therapeutischen Eingriff eine bedeutende Applikation der 'Omics'-Technologien dar. Die Strategien zur Identifizierung von Bio- oder diagnostischen Markern beinhalten die umfassende, vergleichende Analyse von biologischem Material behandelter oder modifizierter Organismen und unbehandelter Kontrollproben in einer adäquaten Anzahl biologischer Replikate. Nach der Aufbereitung der gewonnenen Daten mittels verschiedener Methoden des data minings können potenzielle Biomarker identifiziert werden, welche eine Charakterisierung der analysierten Populationen erlauben (Abbildung 6.1). Diese kann umso präziser erfolgen, je mehr Variablen erfasst werden. Somit stellt die Entwicklung analytischer Methoden zur Analyse von Stoffwechselintermediaten nach wie vor einen Wissenschaftszweig mit entscheidender Bedeutung dar.



**Abbildung 6.1: Multivariate Analyse von 'Omics'-Daten im systembiologischen Kontext** (aus [366]). Ausgehend von Metabolit- und/oder Proteinprofilen, welche den Steady-state-Level des analysierten Untersuchungsmaterials beschreiben, können in Abhängigkeit vom experimentellen Design nach der Anwendung dimensionsreduzierender Techniken biologische Charakterisierungen erfolgen und/oder potenzielle Biomarker identifiziert werden oder aber regulatorische Charakteristiken erstellt und die Netzwerktopologie analysiert werden. Die Voraussetzung dieser Analysen stellt in jedem Fall die Analyse einer adäquaten Anzahl biologischer Replikate dar.

## 8 Material und Methoden

### 8.1 Materialien und Equipment

#### 8.1.1 Chemikalien und Referenzverbindungen

Standardchemikalien wurden, soweit nicht gesondert ausgewiesen, in analytischer Reinheit von den Firmen Fluka (Buchs, Schweiz), Sigma (Taufkirchen, Deutschland), Merck (Darmstadt, Deutschland), Riedel-de-Haën (Seelze, Deutschland), Roche (Penzberg, Deutschland) und Roth (Karlsruhe, Deutschland) bezogen.

Synthetische Peptide wurden in einer Reinheit von mindestens 80% von der Firma Peptides & Elephants GMBH (Nuthetal, Deutschland) bezogen.

Wasser wurde mittels hausinterner Milli-Q-Anlage (Fa. Millipore, Billerica, MA, USA) entionisiert und ausschliesslich in dieser Form verwendet.

Puffer und Lösungen wurden stets am Tag ihrer Verwendung frisch zubereitet.

#### 8.1.2 Analytisches Equipment

HP 5980 Gaschromatograph gekoppelt an ein Pegasus II TOF Massenspektrometer	(Fa. Agilent, Waldbronn, Deutschland) (LECO Corporation, St. Joseph, MI, USA)
Agilent nanoHPLC System 1100	(Fa. Agilent, Böblingen, Deutschland)
Linear Ion Trap (LTQ) Massenspektrometer	(ThermoElectron Corporation, San Jose, CA, USA)
Triple Stage Quadrupole (TSQ) Discovery MAX Massenspektrometer	(ThermoElectron Corporation, San Jose, CA, USA)
IMAGING-PAM Fluorometer	(Fa. Walz, Effeltrich, Germany)

### 8.1.3 Pflanzliches Untersuchungsmaterial und entwicklungs-synchronisierte Anzucht

***Arabidopsis thaliana* Columbia-0** (Col-0, Wildtyp) und **PGM** (loss-of-function Mutation der chloroplastidiären Phosphoglucomutase) Samen wurden auf Standarderde GS 90 (2:1 [v/v] Vermiculite), angestaut mit 0.15% [v/v] Previcur N, ausgesät und unter Langtagbedingungen (16h Licht: 145  $\mu$ E, 20°C, 75% relative Luftfeuchte; 8h Dunkelheit: 0  $\mu$ E, 6°C, 75% relative Luftfeuchte) stratifiziert. Die nach 7 Tagen erhaltenen Keimlinge wurden für weitere 14 Tage unter Kurztagbedingungen (8h Licht: 145  $\mu$ E mit je 0.5h 50% Licht am Morgen sowie am Abend, 20°C, 60% relative Luftfeuchte; 16h Dunkelheit: 0  $\mu$ E, 16°C, 75% relative Luftfeuchte) angezogen. So erhaltene Jungpflanzen wurden bis zur Ernte für ca. 6 Wochen unter genau kontrollierten Wachstumsbedingungen (siehe Tabelle 8.1.4.1) in Phytotronen angezogen. Die Bewässerung erfolgte mit Leitungswasser, wobei die Bewässerungsmenge dem Wachstumsstadium der Pflanzen angepasst wurde (Jungpflanzen ca. 20 mL/Tag ansteigend auf ca. 250 mL/Tag bei ausgewachsenen Pflanzen).

**Tabelle 8.1.4.1: Wachstumsbedingungen für *A. thaliana* Jungpflanzen** (beginnend am 21 Tag nach Aussaat bis zur Erntereife, falls nicht gesondert angegeben)

Bedingung	Lichtperiode	Temp. Tag	Temp. Nacht	rel. Luftfeuchtigkeit
Kurztag	8h, 160 $\mu$ E	20 °C	20°C	70 %
Langtag	12h, 160 $\mu$ E	20 °C	20°C	70 %

***Nicotiana tabacum* Samsun NN (SNN)** Samen wurden auf Pikiererde, angestaut mit 0.15% [v/v] Previcur N, ausgesät und 14 Tage unter Langtagbedingungen (16h Licht: 120  $\mu$ E, 22°C, 70% relative Luftfeuchte; 8h Dunkelheit: 0  $\mu$ E, 18°C, 70% relative Luftfeuchte) stratifiziert. Aufgelaufene Keimlinge wurden für weitere 4 Wochen unter diesen Bedingungen gehalten und die erhaltenen Jungpflanzen in ein Erde-Sand-Gemisch (Mischungsverhältnis 2:1) pikiert und bis zur Ernte für weitere 6 Wochen unter Langtagbedingungen (16h Licht: 145  $\mu$ E, 25°C, 60-70% relative Luftfeuchte; 8h Dunkelheit: 0  $\mu$ E, 20°C, 60-70% relative Luftfeuchte) im Gewächshaus gehalten. Die Bewässerung erfolgte computergesteuert mit integrierter Düngung (Hakaphos spezial: 16% N, 8% P, 22% K, 3% Mg; 1 g/L; Fa. COMPO GmbH&Co. KG, Münster, Deutschland), dabei wurde die Bewässerungsmenge dem Wachstumsstadium der Pflanzen angepasst (Jungpflanzen ca. 50 mL/Tag ansteigend auf ca. 500 mL/Tag bei ausgewachsenen Pflanzen).

***Solanum tuberosum* cv. Desiree** Gewebekulturstecklinge wurden in Pikiererde, angestaut mit 0.15% [v/v] Previcur N, gesetzt und 3 Wochen unter Langtagbedingungen (16h Licht: 120  $\mu$ E, 22°C, 70% relative Luftfeuchte; 8h Dunkelheit: 0  $\mu$ E, 18°C, 70% relative Luftfeuchte) gehalten. Die erhaltenen Jungpflanzen wurden anschliessend in ein Erde-Sand-Gemisch (Mischungsverhältnis 2:1) getopft und bis zur Ernte für weitere 6 Wochen unter Langtagbedingungen (16h Licht: 145  $\mu$ E, 20°C, 50-60% relative Luftfeuchte; 8h Dunkelheit: 0  $\mu$ E, 18°C, 50-60% relative Luftfeuchte) im Gewächshaus gehalten. Die Bewässerung erfolgte computergesteuert mit integrierter Düngung (Hakaphos rot: 8% N, 12% P, 24% K, 4% Mg; 0.7 g/L; Fa. COMPO GmbH&Co. KG, Münster, Deutschland), dabei wurde die Bewässerungsmenge dem Wachstumsstadium der Pflanzen angepasst (Jungpflanzen ca. 20 mL/Tag ansteigend auf ca. 250 mL/Tag bei ausgewachsenen Pflanzen).

**Blattmaterial** erntereifer Pflanzen wurde unmittelbar nach der Ernte zu definierten Zeitpunkten in flüssigem Stickstoff schockgefroren und bis zur weiteren Analyse bei  $-80^{\circ}\text{C}$  gelagert.

**Knollenmaterial** wurden mit  $\text{H}_2\text{O}$  gewaschen und mittels eines Korkbohrers (5 mm ID) Proben quer zur Sprossachse ausgestanzt. Aus diesen Stanzstücken wurden mit einem Skalpell bis zu drei 1.5 mm dicke Scheiben 3 mm unterhalb der Schale präpariert und diese bis zur weiteren Analyse bei  $-80^{\circ}\text{C}$  gelagert.

#### 8.1.4 Zellaufschluß

Geerntetes Blattmaterial wurde mit einem Mörser unter ständiger Kühlung durch flüssigen Stickstoff sehr fein gemörsert, 10-15 mg des erhaltenen Pulvers unter weiterer Kühlung abgewogen und das Material nochmals mit einer Retschmühle für 30 sek bei einer Frequenz von  $20\text{ s}^{-1}$  homogenisiert.

Die Homogenisierung der aus Kartoffelknollen präparierten Proben erfolgte durch die Retschmühle (zwei Zyklen, 30 sek,  $20\text{ sek}^{-1}$ ).

## 8.2 Software

LECO <sup>®</sup> ChromaTOF <sup>™</sup> Version 1.61	(LECO Corporation, St. Joseph, MI, USA)
Chemstation <sup>®</sup> Version 2.2	(Fa. Agilent, Böblingen, Deutschland)
Xcalibur <sup>™</sup> Version 1.4	(ThermoElectron Corporation, San Jose, CA, USA)
Bioworks <sup>™</sup> Version 3.1	(ThermoElectron Corporation, San Jose, CA, USA)
Contrast <sup>™</sup> Version 1.9	(ThermoElectron Corporation, San Jose, CA, USA)
LCQUAN <sup>™</sup> Version 2.0	(ThermoElectron Corporation, San Jose, CA, USA)
MATLAB <sup>®</sup> Version 7.0	(The Mathworks Inc., Natick, MA, USA)

## 8.3 Integrative Extraktion von Metaboliten und Proteinen

Die Extraktion von Metaboliten und/oder Proteinen erfolgte grundsätzlich nach dem integrativen Ansatz, welcher die simultane Extraktion beider Stoffklassen aus einer Probe erlaubt. Dabei wurden verschiedene Methoden der Metabolit- und Proteinextraktion in Kombination angewandt.

### 8.3.1 Metabolitextraktion mit $\text{CH}_3\text{OH}/\text{CHCl}_3/\text{H}_2\text{O}$

Homogenisiertes Pflanzenmaterial (10-15 mg) wurden mit 1 mL einer eiskalten ( $-20^{\circ}\text{C}$ ) Extraktionslösung ( $\text{CH}_3\text{OH}/\text{CHCl}_3/\text{H}_2\text{O}$  [v/v/v = 2.5:1:0.5]) nach Zugabe von 10  $\mu\text{L}$  einer internen Standardlösung bestehend aus je 200  $\mu\text{g}$  D-Sorbitol- $^{13}\text{C}_6$ , D,L-Leucin-2,3,3- $\text{D}_3$ , L-Asparaginsäure-2,3,3- $\text{D}_3$ /mL  $\text{H}_2\text{O}$  (Fa. ISOTECH, Miamisburg, OH, USA) für 8 min unter Schütteln bei  $4^{\circ}\text{C}$  extrahiert. Nach anschließender Zentrifugation für 6 min bei 15.000 g und  $4^{\circ}\text{C}$  erfolgte eine Phasentrennung des erhaltenen Überstandes durch Zugabe von 500  $\mu\text{L}$   $\text{H}_2\text{O}$ , vortexen für 1-2 min und Zentrifugation für 2 min bei 15.000 g und  $4^{\circ}\text{C}$ . Für

anschließende Metabolitanalysen wurde die polare Phase komplett bis zur vollständigen Trockene im Vakuum eingeengt.

Die im Pellet verbliebenen löslichen Proteine wurden gemäss 8.3.2 bzw. 8.3.3 extrahiert.

### 8.3.2 Phenolische Proteinextraktion mit Urea

Die präzipitierten löslichen Proteine wurden zunächst in 1 mL eines frisch zubereiteten Ureapuffers (8 M Urea, 50 mM Ammoniumbicarbonat, 200 mM Methylamin, 1% [v/v]  $\beta$ -Mercaptoethanol, pH 7.5) durch sanftes Schütteln bei 4°C für 30 min resuspendiert, anschliessend bei 5.000 g und 4°C für 10 min zentrifugiert und der Überstand nach Zugabe von 500  $\mu$ L TE-Puffer (10 mM TRIS, 1 mM EDTA- $\text{Na}_2$ )-gesättigtem Phenol (Fa. Roth, Karlsruhe, Deutschland) für 30 min bei RT sanft geschüttelt. Nach durch Zugabe von 200  $\mu$ L eines 50  $\mu$ M Tris-Puffers (pH 7.5) und Zentrifugation bei 5.000 g für 5 min erfolgter Phasentrennung wurden die in die organische Phase extrahierten Proteine über Nacht in 5 Volumen eiskaltem Aceton gefällt (-20°C) und die Präzipitate zweimal mit eiskaltem  $\text{CH}_3\text{OH}$  (-20°C) gewaschen.

### 8.3.3 Phenolische Proteinextraktion mit Saccharose

Die im Pellet verbliebenen löslichen Proteine wurden durch Zugabe von 200  $\mu$ L Saccharosepuffer (40% Saccharose [w:v], 50 mM HEPES-KOH, pH 7.5) sowie 600  $\mu$ L TE-Puffer (10 mM TRIS, 1 mM EDTA- $\text{Na}_2$ )-gesättigtem Phenol (Fa. Roth, Karlsruhe, Deutschland) unter 30-minütigem Schütteln bei 4°C extrahiert. Nach Zentrifugation für 5 min bei 5.000 g und 4°C wurden die nun in der organischen Phase befindlichen Proteine über Nacht in 5 Volumen eiskaltem Aceton (-20°C) präzipitiert und das nach anschliessender Zentrifugation bei 5.000 g und 4°C für 10 min erhaltene Proteinpellet zweimal mit eiskaltem  $\text{CH}_3\text{OH}$  (-20°C) gewaschen (siehe auch Abbildung 4.2.1.2).

## 8.4 Metabolite Profiling

### 8.4.1 Derivatisierung

Eingeengte Metabolitfraktionen wurden unter starkem Schütteln zunächst mit 15 mL einer frisch angesetzten Methoxyaminhydrochloridlösung (40 mg/mL Pyridin) für 90 min bei 30°C, anschliessend mit weiteren 135  $\mu$ L MSTFA (Fa. Machery-Nagel, Düren, Deutschland) für 30 min bei 37°C inkubiert. Vor der Injektion wurden die derivatisierten Proben bei 14.000 g für 5 min zentrifugiert und anschliessend mit 1  $\mu$ L Retentionsindex-Markerlösungen (n-Alkane C12-C26, je 2  $\mu$ L/mL in Pyridin sowie n-Fettsäuremethylester-Standardlösung C8-C30, je 200  $\mu$ g/mL  $\text{CHCl}_3$ ) versetzt.

### 8.4.2 GC-TOF-MS basierte Metabolitanalyse

#### 8.4.2.1 GC-TOF-MS Setup

Alle Metabolitanalysen erfolgten mit einem HP 6890 Gaschromatographen mit konischen, deaktivierten Standard-Split/Splitless-Linern, gekoppelt an ein Pegasus II TOF Massenspektrometer. Nach splittloser Injektion von 1  $\mu$ L derivatisierter Probe bei 230°C erfolgte

die Trennung an einer MDN-35 (35%-Phenyl-methylpolysiloxan) Kapillarsäule, 30 m x 0,32 mm ID, 0,25 µm Film (Fa. SUPELCO, Bellefonte, PA, USA) bei konstantem Fluss von 2 mL/min Helium mit temperaturprogrammierter Verdampfung (80°C, 2 min isotherm; 80-360°C, 15°C/min; 360°C, 5 min isotherm). Vor jeder Injektion wurde der Liner durch Injektion von 1 µL MSTFA gespült; nach ca. 80 Injektionen erfolgte ein manueller Linerwechsel.

Die Datenakquisition erfolgte softwaregesteuert mit Hilfe der ChromaTOF™ Software bei einer Scanrate von 20 s<sup>-1</sup> in einem Massenbereich von m/z 85-600 nach Elektronenstossionisation bei 70 eV. Als Referenzgas für das (Auto)Tuning diente Trisperflourbutylamin (FC43) (Fa. PerkinElmer Inc., Wellesley, MA, USA).

Die Analysen des während der Lichtperiode geernteten Blattmaterials der stärkeleeren *A. thaliana* Mutante PGM erforderte für die Quantifizierung löslicher Carbohydrate aufgrund der drastischen Akkumulation dieser Verbindungen zusätzlich die Analyse von Verdünnungsserien. Dazu wurden derivatisierte Proben mit Derivatisierungsreagenzien im Verhältnis 1:5, ggf. 1:10 verdünnt.

#### 8.4.2.2 Qualitätskontrolle

Zu Beginn und am Ende jedes Probensets wurde ein Quality Control-Standardmix (QC-Mix) bestehend aus 28 Standardsubstanzen (Tabelle 8.4.2.2.1) analytischer Reinheit nach entsprechender Derivatisierung in einer Konzentration von je 100 pmol injiziert. Zusätzlich wurden Verdünnungen (10%, 30%, 50% [v/v]) des QC-Mixes injiziert und die Güte der Analyse anhand der resultierenden Linearität und der Intensitäten der für die jeweiligen Standardsubstanzen erhaltenen Signale bewertet, wobei folgende Kriterien zugrunde gelegt wurden: Abweichung der Peakfläche ≤ 5%, R<sup>2</sup> ≥ 0.95.

**Tabelle 8.4.2.2.1: Zusammensetzung des QC-Mixes**

Alanin	4-Hydroxybenzoesäure	Lanosterol	Ribose
Cholesterol	3-Hydroxypyridine	Maltose	Sorbitol
Citramalat	Isoleucin	Maltotriose	Sorbose
Citrat	Kaffeat	Palatinose	Stigmasterol
Fucose	α-Ketoglutarat	Panthothenat	Threitol
Glucose	Lactitol	Putrescin	Urea
Glycin	Lactose	Ribitol	Valin

#### 8.4.2.3 Spektrenbibliothek

Zur Erstellung einer Bibliothek wurden die Spektren von ca. 400 zumeist im pflanzlichen Stoffwechsel vorkommenden chemischen Verbindungen durch Injektion von je 100 pmol nach der Derivatisierung aufgenommen und deren Retentionsindex basierend auf dem Retentionsverhalten der zugesetzten homologen n-Alkane und/oder n-Fettsäuremethyl-ester errechnet und eine benutzerdefinierte Spektrenbibliothek erstellt.

#### 8.4.2.4 Semi-automatische Peakidentifikation und Quantifizierung

Alle zu einer Probencharge gehörenden Chromatogramme wurden zunächst mit Hilfe der LECO®ChromaTOF™ Software mit der benutzerdefinierte Spektrenbibliothek abgeglichen.

Abhängig von der Anzahl der gefundenen Peaks in jedem Chromatogramm und von der Güte der dazugehörigen Spektren wurde ein Chromatogramm ausgewählt, welches im weiteren Verlauf als Referenzchromatogramm dienen sollte. Alle Annotationen wurden im Referenzchromatogramm manuell unter besonderer Berücksichtigung der Retentionszeit bzw. des Retentionsindex evaluiert und gleichzeitig geeignete Fragmentmassen, spezifisch für den jeweiligen Analyten, definiert. Durch den softwaregesteuerten Abgleich aller Chromatogramme mit dem Referenzchromatogramm unter Einhaltung spezifischer Schwellenwerte (Spektrenübereinstimmung mind. 80%, Retentionszeitabweichung <2 sek) erfolgte die Peakannotation und Quantifizierung basierend auf den definierten Fragmentmassen. Die Identifikationen sowie die Integrationsgrenzen eines jeden Peaks wurden in allen Chromatogrammen manuell überprüft.

## 8.5 Protein Profiling

### 8.5.1 Enzymatischer Proteinverdau

Die gemäss Kapitel 8.3 erhaltenen Proteinpellets wurde in 20 µL Solubilisierungspuffer (8 M Urea, 50 mM TRIS-HCl, 200 mM Methylamin, 0.1% [v/v] β-Mercaptoethanol, pH 7.5) resuspendiert und der Proteingehalt nach Bradford [367] bestimmt.

Nach Zugabe von Dithiothreitol (DTT) bis zu einer Endkonzentration von 5mM wurden die Proteine unter sanftem Schütteln für 30 min bei 37°C zunächst reduziert und anschliessend durch weitere Zugabe einer frisch zubereiteten Iodoacetamid (IAA)-Lösung (Endkonzentration 100 mM) und Inkubation für 60 min im Dunkeln bei RT alkyliert. Die verbliebene IAA-Aktivität wurde durch Zugabe von DTT in einer Endkonzentration von 10 mM gequencht.

Unter sanftem Schütteln für 5h bei 37°C wurden die Proteine zunächst mit Endoproteinase Lys-C, sequencing grade (Fa. Roche Applied Sciences, Penzberg, Deutschland) (Enzym:Substrat = 1:1000) verdaut. Die gesamte Probe wurde anschliessend mit Trypsinpuffer (50 mM TRIS, 10% ACN, 10 mM CaCl<sub>2</sub>; pH 7.5) vierfach verdünnt und nach Zugabe von Poroszyme™ Bulk Immobilized Trypsin (Fa. Applied Biosystems, Darmstadt, Deutschland) in einem exakten Verhältnis Enzym:Substrat = 1:20 erneut für 12h bei 37°C sanft schüttelnd inkubiert. Erhaltene Peptide wurde mit SPEC® C18 96-Well Plates (Fa. Varian, Darmstadt, Deutschland) den Herstellerangaben entsprechend entsalzt und nach Lyophilisation bis zur massenspektrometrischen Analyse bei -20°C gelagert.

### 8.5.2 nanoLC-MS-basierte Shotgun-Proteinanalyse

#### 8.5.2.1 LC-MS Setup

Jeweils 50 µg des nach 8.5.1 proteolytisch verdauten Gesamtproteins wurden vor der massenspektrometrischen Analyse in 10-20 µL einer 5% [v/v] Ameisensäurelösung gelöst. Bei einer Flußrate von 10 µL/min wurde diese mittels Agilent nanoHPLC System 1100 auf einer Peptide-Trap Vorsäule (Fa. Michrom, Auburn, CA, USA) zunächst aufkonzentriert. Nach 10-minütigem Waschen mit HPLC-Laufmittel **A** (5% [v/v] ACN, 0.1% [v/v] Ameisensäure) bei einer Flußrate von 10 µL/min wurden die Peptide anschliessend auf eine monolithische C18 reversed phase Kapillarsäule (0.5 m, 100µm ID) geladen. Die Elution erfolgte mittels nanoHPLC durch eine 4-stündige lineare Gradientenelution mit HPLC-Laufmittel **B** (90% [v/v] ACN, 0.1% [v/v] Ameisensäure) bei einer Flußrate von 400 nL/min. Eluierte Peptide wurden mit einem Finnigan Linear Ion Trap Massenspektrometer (LTQ) datenabhängig analysiert. Dabei wurden die drei abundantesten Peptidionen eines

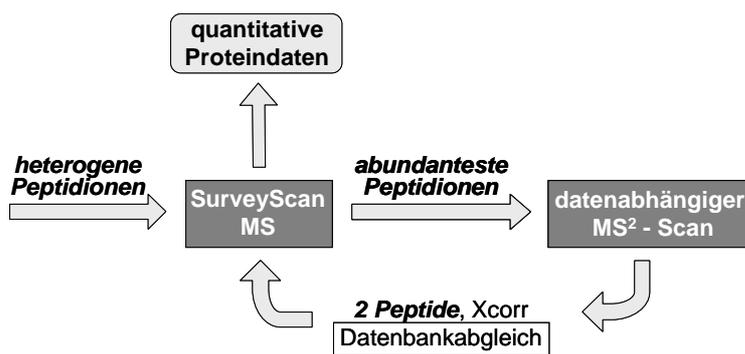
jeden MS-SurveyScans dynamisch einer kollisionsinduzierten Dissoziation bei einer standardisierten Kollisionenergie von 35% mit anschließendem MS<sup>2</sup>-Scan unterworfen. Die Kapillartemperatur wurde auf 150°C, die Electrospray Voltage auf 1.9 kV festgesetzt.

### 8.5.2.2 Proteinidentifikation

Aus den erhaltenen Raw Files wurden mit Bioworks<sup>TM</sup> softwaregesteuert DTA Files erstellt und mit für *Arabidopsis* spezifischen Datenbanken (<http://www.arabidopsis.org/> oder <ftp://ftpmips.gsf.de/cress/>) abgeglichen [94]. Ein Protein galt als eindeutig identifiziert, wenn mind. 2 verschiedene Peptide detektiert wurden, welche den folgenden Kriterien genügten: Xcorr -1 2.0, -2 1.5, -3 3.3 [368].

### 8.5.2.3 Quantifizierung über Peakintegration

Die Ionenspuren der Vorläuferionen der nach MS<sup>2</sup>-Fragmentierung identifizierten Peptide wurden aus den MS-SurveyScans extrahiert und Peakflächen integriert (siehe Abbildung 8.5.2.3.1).



**Abbildung 8.5.2.3.1: Strategie der Proteinquantifizierung über Peakintegration.** Nach der Identifizierung korrespondierender Proteine anhand der MS<sup>2</sup>-Spektren tryptischer Peptide durch einen Datenbankabgleich wurden Ionenspuren der Vorläuferpeptide aus den MS-Survey-Scans extrahiert. Durch Peakintegration der jeweiligen Ionenspuren konnten quantitative Proteindaten erhalten werden.

Der Vergleich der in den einzelnen Proben detektierten Peptide erfolgte softwaregesteuert mit Contrast<sup>TM</sup>.

### 8.5.2.4 Quantifizierung über Spektren-Count

Zur Quantifizierung von Proteinen über den Spektren-Count wurde die kumulative Anzahl der nach der MS<sup>2</sup>-Fragmentierung eines Vorläuferpeptides detektierten Fragmentenspektren erfasst. Gemäss Arbeiten von Liu *et al.* (2004) ist diese der Abundanz eines Proteins in guter Näherung proportional [280].

Der Vergleich der in den einzelnen Proben detektierten Peptide erfolgte softwaregesteuert mit Contrast<sup>TM</sup>.

### 8.5.3 LC-MS basiertes Multiple Reaction Monitoring

#### 8.5.3.1 LC-MS Setup

Jeweils 200 µg des nach 8.5.1 proteolytisch verdauten Gesamtproteins wurden vor der massenspektrometrischen Analyse in 10-20 µL einer 50% ACN [v/v], 5% [v/v] Ameisensäurelösung gelöst. Die erhaltenen Proben wurden anschliessend chromatographisch an einer Luna C18 reversed phase Säule (100mm, 3µm ID) (Fa. Phenomenex, Aschaffenburg, Deutschland) durch eine lineare Gradientenelution aufgetrennt. Nach einer 5-minütigen Haltezeit von 95% [v/v] HPLC-Laufmittel **A** (0.1% [v/v] Ameisensäure) wurde dazu innerhalb von 25 min bei einer Flussrate von 40 µL/min linear auf 60% HPLC-Laufmittel **B** (20% [v/v] ACN, 0.1% [v/v] Ameisensäure) umgestellt. Zur Spülung und Re-Äquilibration der Trennsäule wurde der Anteil von HPLC-Laufmittel **B** innerhalb von 2 min auf 100% erhöht, dieses für 10 min gehalten, innerhalb von weiteren 3 min auf die Anfangsbedingung (95% [v/v] HPLC-Laufmittel **A**) zurückgestellt und die Säule für 20 min äquilibriert, so dass die duty cycle time insgesamt 60 min betrug.

#### Tuning

Eluierte Peptide wurden mit einem Triple Stage Quadrupole (TSQ) Discovery MAX Massenspektrometer, ausgerüstet mit einer Ion Max ESI-Quelle unter Verwendung einer 32-Gauge Metallnadel, analysiert. Die gerätetechnischen Parameter wurden auf die Detektion der synthetischen Peptide der Sequenzen FVLNEQSKYPEER und LIGLAGETNIQGEEQK, welche von der Aminosäuresequenz der cytosolischen Fructose 1,6-bisphosphatase (At1g46370.1) abgeleitet wurden, optimiert. Dazu wurden Peptidstandardlösungen der Konzentration 10pmol/µL in 50% [v/v] ACN, 0.1% [v/v] Ameisensäure bei einer Flussrate von 40 µL/min verwendet. Nach massenspektrometrischer Charakterisierung der synthetischen Peptide wurden das Fragmentierungsverhalten analysiert und die notwendige Kollisionsenergie optimiert. Als Kollisionsgas diente Argon. Die massenspektrometrische Analyse der Peptide erfolgte im MRM Modus bei folgenden Parametern: Tube Lens 170 (arbiträre Einheit), Sheath Gas Pressure 28 (arbiträre Einheit), Spray Voltage 3.8 kV. Die Temperatur der beheizten Transferkappilare betrug 270°C. Die kollisionsinduzierte Fragmentierung der zwei- bzw. dreifach geladenen Vorläuferpeptide erbrachte die in Tabelle 8.5.4.1.1 aufgeführten vier abundantesten Fragmente bei den angegebenen Kollisionsenergien.

**Tabelle 8.5.3.1.1: Detektierte Vorläuferionen und Fragmentionen der Peptidzielsequenzen mit erforderlicher Kollisionsenergie.** Die zugrunde liegenden MS<sup>2</sup>-Experimente erfolgten an einem Triple-Quadrupol Massenspektrometer im MRM-Modus bei den oben angegebenen gerätetechnischen Parametern.

Peptid Sequenz	detektierte Vorläuferionen (m/z)	Kollisionsenergie (eV)	detektierte Produktionen (m/z)
FVLNEQSKYPEER m/z [M+H] <sup>+</sup> = 1596.5	799.1 [M+2H] <sup>2+</sup>	47	219.0
		57	120.0
		42	261.9
		42	247.0
	533.1 [M+3H] <sup>3+</sup>	19	675.9
		43	120.0
		22	219.0
		31	262.1

**Tabelle 8.5.3.1.1: Detektierte Vorläuferionen und Fragmentionen der Peptidzielsequenzen mit (Fortsetzung) erforderlicher Kollisionsenergie.** Die zugrunde liegenden MS<sup>2</sup>-Experimente erfolgten an einem Triple-Quadrupol Massenspektrometer im MRM-Modus bei den oben angegebenen gerätetechnischen Parametern.

Peptid Sequenz	detektierte Vorläuferionen (m/z)	Kollisionsenergie (eV)	detektierte Produktionen (m/z)	
<b>LIGLAGETNIQGEEQK</b> <i>m/z</i> [M+H] <sup>+</sup> = 1699.8	850.6 [M+2H] <sup>2+</sup>	10	850.5	
		48	199.1	
		43	284.1	
		52	86.0	
	567.4 [M+3H] <sup>3+</sup>	43	86.1	
		19	590.2	
		28	199.1	
		56	84.0	

Die Scanweite für alle durchgeführten MRM-Experimente wurde auf 0.5 Da festgesetzt. Die Auflösungen des ersten, als Massenfilter betriebenen Messquadrupols Q1 betrug 0.5, die des zweiten, ebenfalls als Massenfilter betriebenen Messquadrupols Q3 wurde auf 0.7 festgesetzt.

### **Proteinquantifizierung**

Zur Analyse der in *A. thaliana* Blattmaterial vorliegenden Proteinkonzentrationen der Fructose 1,6-bisphosphatase (At1g46370.1) wurde aufgeschlossenes Blattmaterial gemäss 8.3.3 phenolisch mit Saccharose extrahiert und im Anschluss gemäss 8.5.1 proteolytisch verdaut. Zur Validierung der Methode wurden zunächst Kalibriergeraden resultierend aus der Analyse von durch die Addition von jeweils 5, 25, 50, 100 und 150 pmol der synthetischen Standardpeptide zu 200 µg verdauten Gesamtproteinfraktionen gemäss den oben aufgeführten MRM-Parametern erstellt. Lyophilisierte Peptidpellets wurden dazu in 5% [v/v] Ameisensäure gelöst. Die resultierenden Ergebnisse sind Kapitel 6 zu entnehmen.

Die Auswertung der analysierten Proben erfolgte Softwareunterstützt mit LCQUAN<sup>TM</sup>.

## **8.6 Chemometrische Analyse**

### **8.6.1 Konstruktion metabolischer Netzwerke**

Zur Konstruktion metabolischer Netzwerke wurde der paarweise lineare Zusammenhang zwischen Metabolitkonzentrationen anhand des Pearson'schen Produktmoment-Korrelationskoeffizienten berechnet. Abhängig von der Anzahl analysierter Proben wurde die Signifikanz von Korrelationen nach der Student t-Transformation abgeschätzt. Als Schwellenwert zur Verwerfung der Nullhypothese unkorrelierter Metabolitpaare wurde eine Irrtumswahrscheinlichkeit von  $p \leq 0.001$  zugrunde gelegt. Die farbcodierte Darstellung der resultierenden Korrelationsnetzwerke sowie die Darstellung ausgewählter Metabolitkorrelationen erfolgten in MATLAB<sup>®</sup>.

### **8.6.2 Dimensionsreduktion**

#### **Hauptkomponentenanalyse PCA**

Die Dimensionsreduktion der aus den Metabolit- und/oder Proteinanalysen resultierenden hochdimensionalen Datensätze zur zwei- oder dreidimensionalen Darstellung der Datenmengen mittels Hauptkomponentenanalyse erfolgte nach vorangegangener  $\log_{10}$ -Transformation in MATLAB<sup>®</sup> basierend auf der Kovarianzmatrix gemäss des implementierten PCA-Scriptes PRINCOMP.

### **Unabhängige Komponentenanalyse ICA**

Die Dimensionsreduktion der aus den Metabolit- und/oder Proteinanalysen resultierenden hochdimensionalen Datensätze zur zwei- oder dreidimensionalen Darstellung der Datenmengen mittels Unabhängiger Komponentenanalyse erfolgte nach vorangegangener  $\log_{10}$ -Transformation in MATLAB<sup>®</sup> unter Verwendung des von Scholz *et al.* modifizierten FastICA-Algorithmus nach Hyvärinen *et Oja* [168,370]

### **8.6.3 Clustering**

Bi-Clustering und farbcodierte Darstellung relativer Änderungen der Protein- und/oder Metabolitkonzentrationen sowie der experimentellen Bedingungen erfolgte nach Median-Normierung und  $\log_{10}$ -Transformation der erhaltenen Daten mit Hilfe des in MATLAB<sup>®</sup> implementierten CLUSTERDATA-Algorithmus. Als Abstands- oder Ähnlichkeitsmass diente der Pearson'sche Produktmoment-Korrelationskoeffizient.

### **8.6.4 Varianzanalyse**

Die Varianzanalyse (one-way Analysis of Variance, ANOVA) zum Testen auf Mittelwertsunterschiede zwischen Stichproben verschiedener Grundgesamtheiten erfolgte unter Verwendung des in MATLAB<sup>®</sup> implementierten nicht-parametrischen Kruskal-Wallis-Tests. Die zur Verwerfung der Nullhypothese, dass die  $n$  untersuchten Stichproben sich nicht unterscheiden, angenommenen Irrtumswahrscheinlichkeiten  $p$  sind den einzelnen Kapiteln zu entnehmen.

## 9 Referenzen

1. Sanger F, Air GM, Barrell BG, Brown NL, Coulson AR, Fiddes C, Hutchison III A, Slocombe PM, Smith M: **Nucleotide sequencing of bacteriophage phi X174 DNA.** *Nature* 1977, **265**:687-695.
2. Sanger F, Coulson AR, Hong GF, Petersen GB: **Nucleotide sequence of bacteriophage lambda DNA.** *J Mol Biol* 1982, **162**:729-773.
3. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM, et al.: **Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.** *Science* 1995, **269**:496-512.
4. Blattner FR, Plunkett G, 3rd, Bloch CA, Perna NT, Burland V, Riley M, Collado-Vides J, Glasner JD, Rode CK, Mayhew GF, et al.: **The complete genome sequence of Escherichia coli K-12.** *Science* 1997, **277**:1453-1474.
5. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, et al.: **Life with 6000 genes.** *Science* 1996, **274**:546, 563-547.
6. Mewes HW, Albermann K, Bahr M, Frishman D, Gleissner A, Hani J, Heumann K, Kleine K, Maierl A, Oliver SG, et al.: **Overview of the yeast genome.** *Nature* 1997, **387**:7-65.
7. The Arabidopsis Genome Initiative: **Analysis of the genome sequence of the flowering plant Arabidopsis thaliana.** *Nature* 2000, **408**:796-815.
8. McPherson JD, Marra M, Hillier L, Waterston RH, Chinwalla A, Wallis J, Sekhon M, Wylie K, Mardis ER, Wilson RK, et al.: **A physical map of the human genome.** *Nature* 2001, **409**:934-941.
9. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, et al.: **The sequence of the human genome.** *Science* 2001, **291**:1304-1351.
10. Berk A, Sharp P: **Structure of the Adenovirus 2 Early mRNA's.** *Cell* 1978, **14**:695-711.
11. Beadle GW, Tatum EL: **Genetic Control of Biochemical Reactions in Neurospora.** *Proc Natl Acad Sci U S A* 1941, **27**:499-506.
12. Itoh H, Washio T, Tomita M: **Computational comparative analyses of alternative splicing regulation using full-length cDNA of various eukaryotes.** *RNA* 2004, **10**:1005-1018.
13. Alexandrov NN, Troukhan ME, Brover VV, Tatarinova T, Flavell RB, Feldmann KA: **Features of Arabidopsis genes and genome discovered using full-length cDNAs.** *Plant Mol Biol* 2006, **60**:69-85.

14. Baltimore D: **RNA-dependent DNA polymerase in virions of RNA tumor viruses.** *Nature* 1970, **226**:1209-1211.
15. Temin HM, Mizutani S: **RNA-directed DNA polymerase in virions of Rous sarcoma virus.** *Nature* 1970, **226**:1211-1213.
16. Branco-Price C, Kawaguchi R, Ferreira RB, Baily-Serres J: **GEgenome-wide analysis of transcript abundance and translation in Arabidopsis seedlings subjected to oxygen deprivation.** *Annals of Botany* 2005, **96**:647-660.
17. Zhang CK, Lang P, Dane F, Ebel RC, Singh NK, Locy RD, Dozier WA: **Cold acclimation induced genes of trifoliolate orange (Poncirus trifoliata).** *Plant Cell Reports* 2005, **23**:764-769.
18. Way H, Chapman S, McIntyre L, Casu R, Xue GP, Manners J, Shorter R: **Identification of differentially expressed genes in wheat undergoing gradual water deficit stress using a subtractive hybridisation approach.** *Plant Science* 2005, **168**:661-670.
19. Wu J, Zhang NS, Hayes A, Panoutsopoulou K, Oliver SG: **Global analysis of nutrient control of gene expression in Saccharomyces cerevisiae during growth and starvation.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:3148-3153.
20. Zamboni N, Fischer E, Muffler A, Wyss M, Hohmann HP, Sauer U: **Transient expression and flux changes during a shift from high to low riboflavin production in continuous cultures of Bacillus subtilis.** *Biotechnology and Bioengineering* 2005, **89**:219-232.
21. Gygi SP, Rochon Y, Franza BR, Aebersold R: **Correlation between protein and mRNA abundance in yeast.** *Molecular and Cellular Biology* 1999, **19**:1720-1730.
22. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L: **Integrated genomic and proteomic analyses of a systematically perturbed metabolic network.** *Science* 2001, **292**:929-934.
23. MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, Gould KL, Wolters D, Washburn M, Weiss A, et al.: **Shotgun identification of protein modifications from protein complexes and lens tissue.** *Proceedings of the National Academy of Sciences of the United States of America* 2002, **99**:7900-7905.
24. Meng F, Forbes AJ, Miller LM, Kelleher NL: **Detection and localization of protein modifications by high resolution tandem mass spectrometry.** *Mass Spectrometry Reviews* 2005, **24**:126-134.
25. Schachter H, Chen SH, Zhang WL, Spence AM, Zhu SX, Callahan JW, Mahuran DJ, Fan XL, Bagshaw RD, She YM, et al.: **Functional post-translational proteomics approach to study the role of N-glycans in the development of Caenorhabditis elegans.** In *Glycogenomics: The Impact of Genomics and Informatics on Glycobiology*. Edited by: PORTLAND PRESS LTD; 2002:1-21. Biochemical Society Symposium
26. Cantin GT, Yates JR, 3rd: **Strategies for shotgun identification of post-translational modifications by mass spectrometry.** *Journal of Chromatography A* 2004, **1053**:7-14.
27. Bradshaw AD: **Unravelling phenotyping plasticity -- Why should we bother?** *New Phytology* 2006, **170**:644-648.
28. Fiehn O: **Metabolomics--the link between genotypes and phenotypes.** *Plant Mol Biol* 2002, **48**:155-171.
29. Laibach F: **Zur Frage nach der Individualität der Chromosomen im Pflanzenreich.** *Botanisches Centralblatt Beihefte 1* 1907, **22**:191-219.
30. Laibach F: **Arabidopsis thaliana (L) Heynh. als Objekt für genetische und entwicklungsphysiologische Untersuchungen.** *Botanisches Archiv* 1943, **44**:439-455.
31. Katavic V, Haughn GW, Reed D, Martin M, Kunst L: **In planta transformation of Arabidopsis thaliana.** *Molecular Genetics and Genomics* 1994, **245**:363-370.

32. Barghchi M: **High-frequency and efficient Agrobacterium-mediated transformation of Arabidopsis thaliana ecotypes "C24" and "Landsberg erecta" using Agrobacterium tumefaciens.** *Methods in Molecular Biology* 1995, **44**:135-147.
33. Schopfer P, Brennicke P: *Pflanzenphysiologie*. Heidelberg: Spektrum Akademischer Verlag; 2006.
34. Plant physiology online on World Wide Web URL: <http://3e.plantphys.net/>
35. The Arabidopsis Book on World Wide Web URL: <http://www.aspb.org/publications/arabidopsis/>
36. O'Farrell PH: **High resolution two-dimensional electrophoresis of proteins.** *J Biol Chem* 1975, **250**:4007-4021.
37. Klose J: **Protein mapping by combined isoelectric focussing and electrophoresis of mouse tissue: A novel approach to testing for induced point mutations in mammals.** *Humangenetik* 1975, **26**:231-243.
38. Glinski M, Weckwerth W: **The Role of Mass Spectrometry in Plant Biology.** *Mass Spectrometry Reviews* 2005, in press.
39. Domon B, Aebersold R: **Mass spectrometry and protein analysis.** *Science* 2006, **312**:212-217.
40. Saravanan RS, Rose JKC: **A critical evaluation of sample extraction techniques for enhanced proteomic analysis of recalcitrant plant tissues.** *Proteomics* 2004, **4**:2522-2532.
41. Hollemeyer K, Heinzle E, Tholey A: **Identification of oxidized methionine residues in peptides containing two methionine residues by derivatization and matrix-assisted laser desorption mass spectrometry.** *Proteomics* 2002, **2**:1524-1531.
42. Herbert B, Hopwood F, Oxley D, McCarthy J, Laver M, Grinyer J, Goodall A, Williams K, Castagna A, Righetti PG: **Beta-elimination: an unexpected artefact in proteome analysis.** *Proteomics* 2003, **3**:826-831.
43. Jones EW: **Vacuolar proteases and proteolytic artifacts in Saccharomyces cerevisiae.** *Methods in Enzymology* 2002, **351**:127-150.
44. Klose J, Kobalz U: **Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome.** *Electrophoresis* 1995, **16**:1034-1059.
45. Lottspeich F, Zorbach H: *Bioanalytik*. Heidelberg: Spektrum Akademischer Verlag; 1998.
46. Lauber WM, Carroll JA, Dufield DR, Kiesel JR, Radabaugh MR, Malone JP: **Mass spectrometry compatibility of two-dimensional gel protein stains.** *Electrophoresis* 2001, **22**:906-918.
47. Scheler C, Lamer S, Pan ZM, Li XP, Salnikow J, Jungblut P: **Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser desorption/ionization mass spectrometry (MALDI-MS).** *Electrophoresis* 1998, **19**:918-927.
48. Gygi SP, Corthals GL, Zhang Y, Rochon Y, Aebersold R: **Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology.** *Proceedings of the National Academy of Sciences of the United States of America* 2000, **97**:9390-9395.
49. Sun G, Anderson VE: **Prevention of artifactual protein oxidation generated during sodium dodecyl sulfate-gel electrophoresis.** *Electrophoresis* 2004, **25**:959-965.
50. Karas M, Hillenkamp F: **Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons.** *Anal Chem* 1988, **60**:2299-2301.
51. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM: **Electrospray ionization for mass spectrometry of large biomolecules.** *Science* 1989, **246**:64-71.
52. Tozuka Z, Kaneko H, Shiraga T, Mitani Y, Beppu M, Terashita S, Kawamura A, Kagayama A: **Strategy for structural elucidation of drugs and drug metabolites using (MS)<sup>n</sup> fragmentation in an electrospray ion trap.** *Journal of Mass Spectrometry* 2003, **38**:793-808.

53. Swiderek KM, Davis MT, Lee TD: **The identification of peptide modifications derived from gel-separated proteins using electrospray triple quadrupole and ion trap analyses.** *Electrophoresis* 1998, **19**:989-997.
54. Hoffman MD, Sniatynski MJ, Rogalski JC, Le Blanc JC, Kast J: **Multiple neutral loss monitoring (MNM): a multiple method for post-translational modification screening.** *Journal of the American Society for Mass Spectrometry* 2006, **17**:307-317.
55. Mulholland M, Whelan TJ, Rose H, Keegan J: **Direct identification and quantitation of prednisone in the presence of overlapping hydrocortisone by liquid chromatography with electrospray and atmospheric-pressure chemical-ionisation mass spectrometry.** *Journal of Chromatography A* 2000, **870**:135-141.
56. Yang L, Amad M, Winnik WM, Schoen AE, Schweingruber H, Mylchreest I, Rudewicz PJ: **Investigation of an enhanced resolution triple quadrupole mass spectrometer for high-throughput liquid chromatography/tandem mass spectrometry assays.** *Rapid Communications in Mass Spectrometry* 2002, **16**:2060-2066.
57. Glinski M, Romeis T, Witte CP, Wienkoop S, Weckwerth W: **Stable isotope labeling of phosphopeptides for multiparallel kinase target analysis and identification of phosphorylation sites.** *Rapid Commun Mass Spectrom* 2003, **17**:1579-1584.
58. Cox DM, Zhong F, Du M, Duchoslav E, Sakuma T, McDermott JC: **Multiple reaction monitoring as a method for identifying posttranslational modifications.** *Journal of Biomolecular Techniques* 2005, **16**:83-90.
59. Hager JW: **A new linear ion trap mass spectrometer.** *Rapid Communications in Mass Spectrometry* 2002, **16**:512-526.
60. Hager JW, Yves Le Blanc JC: **Product ion scanning using a Q-q-Q linear ion trap (Q TRAP) mass spectrometer.** *Rapid Communications in Mass Spectrometry* 2003, **17**:1056-1064.
61. Xia YQ, Miller JD, Bakhtiar R, Franklin RB, Liu DQ: **Use of a quadrupole linear ion trap mass spectrometer in metabolite identification and bioanalysis.** *Rapid Commun Mass Spectrom* 2003, **17**:1137-1145.
62. Smyth WF, Joyce C, Ramachandran VN, O' Kane E, Coulter D: **Characterisation of selected hypnotic drugs and their metabolites using electrospray ionisation with ion trap mass spectrometry and with quadrupole time-of-flight mass spectrometry and their determination by liquid chromatography-electrospray ionisation-ion trap mass spectrometry.** *Analytica Chimica Acta* 2004, **506**:203-214.
63. Hopfgartner G, Varesio E, Tschappat V, Grivet C, Bourgoigne E, Leuthold LA: **Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules.** *Journal of Mass Spectrometry* 2004, **39**:845-855.
64. Balogh MP: **Debating resolution and mass accuracy.** *Spectroscopy* 2004, **19**:34-40.
65. Ens W, Standing KG: **Hybrid quadrupole/time-of-flight mass spectrometers for analysis of biomolecules.** *Biological Mass Spectrometry* 2005, **402**:49-78.
66. Bykova NV, Rampitsch C, Krokhnin O, Standing KG, Ens W: **Determination and characterization of site specific N-glycosylation using MALDI-Qq-TOF tandem mass spectrometry: case study with a plant protease.** *Analytical Chemistry* 2006, **78**:1093-1103.
67. Liu ZY, Schey KL: **Optimization of a MALDI TOF-TOF mass spectrometer for intact protein analysis.** *Journal of the American Society for Mass Spectrometry* 2005, **16**:482-490.
68. Bennett KL, Stensballe A, Podtelejnikov AV, Moniatte M, Jensen ON: **Phosphopeptide detection and sequencing by matrix-assisted laser desorption/ionization quadrupole time-of-flight tandem mass spectrometry.** *Journal of Mass Spectrometry* 2002, **37**:179-190.

69. Sanders M, Shipkova PA, Zhang H, Warrack BM: **Utility of the hybrid LTQ-FTMS for drug metabolism applications.** *Current Drug Metabolism* 2006, **7**:547-555.
70. Grosshans PB, Chen R, Limbach PA, Marshall AG: **Linear Excitation and Detection in Fourier-Transform Ion-Cyclotron Resonance Mass-Spectrometry.** *International Journal of Mass Spectrometry and Ion Processes* 1994, **139**:169-189.
71. Zabrouskov V, Giacomelli L, van Wijk KJ, McLafferty FW: **New approach for plant proteomics - Characterization of chloroplast proteins of Arabidopsis thaliana by top-down mass spectrometry.** *Molecular & Cellular Proteomics* 2003, **2**:1253-1260.
72. Aharoni A, Ric de Vos CH, Verhoeven HA, Maliepaard CA, Kruppa G, Bino R, Goodenowe DB: **Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry.** *Omics* 2002, **6**:217-234.
73. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Cooks RG: **The Orbitrap: a new mass spectrometer.** *Journal of Mass Spectrometry* 2005, **40**:420-440.
74. Olsen JV, de Godoy LM, Li G, Macek B, Mortensen P, Pesch R, Makarov A, Lange O, Horning S, Mann M: **Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap.** *Molecular and Cellular Proteomics* 2005, **4**:2010-2021.
75. Makarov A, Denisov E, Lange O, Horning S: **Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer.** *Journal of the American Society for Mass Spectrometry* 2006, **17**:977-982.
76. Patton WF: **Detection technologies in proteome analysis.** *Journal of Chromatography B* 2002, **771**:3-31.
77. Shaw J, Rowlinson R, Nickson J, Stone T, Sweet A, Williams K, Tonge R: **Evaluation of saturation labelling two-dimensional difference gel electrophoresis fluorescent dyes.** *Proteomics* 2003, **3**:1181-1195.
78. Gehanne S, Ceconi D, Carboni L, Righetti PG, Domenici E, Hamdan M: **Quantitative analysis of two-dimensional gel-separated proteins using isotopically marked alkylating agents and matrix-assisted laser desorption/ionization mass spectrometry.** *Rapid Commun Mass Spectrom* 2002, **16**:1692-1698.
79. Righetti PG, Castagna A, Antonucci F, Piubelli C, Ceconi D, Campostrini N, Antonioli P, Astner H, Hamdan M: **Critical survey of quantitative proteomics in two-dimensional electrophoretic approaches.** *J Chromatogr A* 2004, **1051**:3-17.
80. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R: **Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.** *Nature Biotechnology* 1999, **17**:994-999.
81. Sakai J, Kojima S, Yanagi K, Kanaoka M: **<sup>18</sup>O-labeling quantitative proteomics using an ion trap mass spectrometer.** *Proteomics* 2005:16-23.
82. Schmidt A, Kellermann J, Lottspeich F: **A novel strategy for quantitative proteomics using isotope-coded protein labels.** *Proteomics* 2004.
83. Stewart II, Thomson T, Figeys D: **<sup>18</sup>O labeling: a tool for proteomics.** *Rapid Communications in Mass Spectrometry* 2001, **15**:2456-2465.
84. Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C: **Proteolytic <sup>18</sup>O labeling for comparative proteomics: model studies with two serotypes of adenovirus.** *Analytical Chemistry* 2001, **73**:2836-2842.
85. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, et al.: **Multiplexed protein quantitation in saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.** *Mol Cell Proteomics* 2004, **3**:1154-1169.
86. Oliver SG, Winson MK, Kell DB, Baganz F: **Systematic functional analysis of the yeast genome.** *Trends in Biotechnology* 1998, **16**:373-378.
87. Goodacre R, Vaidyanathan S, Dunn WB, Harrigan GG, Kell DB: **Metabolomics by numbers: acquiring and understanding global metabolite data.** *Trends Biotechnol* 2004, **22**:245-252.

88. Foerster J, Famili I, Fu P, Passon BO, Nielsen J: **Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network.** *Genome research* 2003, **13**:244-253.
89. Fiehn O: **Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks.** *Comparative and Functional Genomics* 2001, **2**:155-168.
90. Maharjan RP, Ferenci T: **Global metabolite analysis: the influence of extraction methodology on metabolome profiles of *Escherichia coli*.** *Analytical Biochemistry* 2003, **313**:145-154.
91. Starmans DAJ, Nijhuis HH: **Extraction of secondary metabolites from plant material: A review.** *Trends in Food Science & Technology* 1996, **7**:191-197.
92. Gullberg J, Jonsson P, Nordstrom A, Sjostrom M, Moritz T: **Design of experiments: an efficient strategy to identify factors influencing extraction and derivatization of *Arabidopsis thaliana* samples in metabolomic studies with gas chromatography/mass spectrometry.** *Analytical Biochemistry* 2004, **331**:283-295.
93. Hausler RE, Fischer KL, Flugge UI: **Determination of low-abundant metabolites in plant extracts by NAD(P)H fluorescence with a microtiter plate reader.** *Anal Biochem* 2000, **281**:1-8.
94. Weckwerth W, Wenzel K, Fiehn O: **Process for the integrated extraction, identification and quantification of metabolites, proteins and RNA to reveal their co-regulation in biochemical networks.** *Proteomics* 2004, **4**:78-83.
95. Dunn WB, Bailey NJ, Johnson HE: **Measuring the metabolome: current analytical technologies.** *Analyst* 2005, **130**:606-625.
96. James AT, Martin AJP: **Gas-Liquid Chromatography - a Technique for the Analysis and Identification of Volatile Materials.** *British Medical Bulletin* 1954, **10**:170-176.
97. Holmes JC, Morrell FA: **Oscillographic Mass Spectrometric Monitoring of Gas Chromatography.** *Applied Spectroscopy* 1957, **11**:86-87.
98. Weckwerth W, Loureiro ME, Wenzel K, Fiehn O: **Differential metabolic networks unravel the effects of silent plant phenotypes.** *Proc Natl Acad Sci U S A* 2004, **101**:7809-7814.
99. Pollmann S, Muller A, Pietrowski M, Weiler EW: **Occurrence and formation of indole-3-acetamide in *Arabidopsis thaliana*.** *Planta* 2002, **216**:155-161.
100. Ferreira V, Jarauta I, Lopez R, Cacho J: **Quantitative determination of sotolon, maltol and free furaneol in wine by solid-phase extraction and gas chromatography-ion trap mass spectrometry.** *Journal of Chromatography A* 2003, **1010**:95-103.
101. Martinez Vidal JL, Arrebola Liebanas FJ, Gonzalez Rodriguez MJ, Garrido Frenich A, Fernandez Moreno JL: **Validation of gas chromatography/triple quadrupole mass spectrometry based method for the quantification of pesticides in food commodities.** *Rapid Communication in Mass Spectrometry* 2006, **20**:365-375.
102. Nakamura S, Takino M, Daishima S: **Trace level determination of phenols as pentafluorobenzyl derivatives by gas chromatography-negative-ion chemical ionization mass spectrometry.** *Analyst* 2001, **126**:835-839.
103. Zaikin VG, Halket JM: **Review: derivatization in mass spectrometry-6. Formation of mixed derivatives of polyfunctional compounds.** *Eur J Mass Spectrom (Chichester, Eng)* 2005, **11**:611-636.
104. Li Z, Jacobus LK, Wuelfing WP, Golden M, Martin GP, Reed RA: **Detection and quantification of low-molecular-weight aldehydes in pharmaceutical excipients by headspace gas chromatography.** *J Chromatogr A* 2006, **1104**:1-10.
105. Gopaul SV, Farrell K, Abbott FS: **A chromatography/negative ion chemical ionization mass spectrometry and liquid chromatography/electrospray ionization tandem mass spectrometry quantitative profiling of N-**

- acetylcysteine conjugates of valproic acid in urine: application in drug metabolism studies in human. *Journal of Mass Spectrometry* 2000, **35**:698-704.
106. Tholl D, Boland W, Hansel A, Loreto F, Rose US, Schnitzler JP: **Practical approaches to plant volatile analysis.** *Plant J* 2006, **45**:540-560.
  107. Demyttenaere JC, Morina RM, De Kimpe N, Sandra P: **Use of headspace solid-phase microextraction and headspace sorptive extraction for the detection of the volatile metabolites produced by toxigenic Fusarium species.** *J Chromatogr A* 2004, **1027**:147-154.
  108. Pellati F, Benvenuti S, Yoshizaki F, Bertelli D, Rossi MC: **Headspace solid-phase microextraction-gas chromatography--mass spectrometry analysis of the volatile compounds of Evodia species fruits.** *J Chromatogr A* 2005, **1087**:265-273.
  109. Yu NH, Ho ENM, Leung DKK, Wan TSM: **Screening of anabolic steroids in horse urine by liquid chromatography-tandem mass spectrometry.** *Journal of Pharmaceutical and Biomedical Analysis* 2005, **37**:1031-1038.
  110. Cuyckens F, Claeys M: **Mass spectrometry in the structural analysis of flavonoids.** *J Mass Spectrom* 2004, **39**:1-15.
  111. Zywicki B, Catchpole G, Draper J, Fiehn O: **Comparison of rapid liquid chromatography-electrospray ionization-tandem mass spectrometry methods for determination of glycoalkaloids in transgenic field-grown potatoes.** *Analytical Biochemistry* 2005, **336**:178-186.
  112. Cai SS, Syage JA: **Comparison of atmospheric pressure photoionization, atmospheric pressure chemical ionization, and electrospray ionization mass spectrometry for analysis of lipids.** *Anal Chem* 2006, **78**:1191-1199.
  113. Raffaelli A, Saba A: **Atmospheric pressure photoionization mass spectrometry.** *Mass Spectrometry Reviews* 2003, **22**:318-331.
  114. Hanold KA, Fischer SM, Cormia PH, Miller CE, Syage JA: **Atmospheric pressure photoionization. 1. General properties for LC/MS.** *Anal Chem* 2004, **76**:2842-2851.
  115. Cai SS, Syage JA: **Atmospheric pressure photoionization mass spectrometry for analysis of fatty acid and acylglycerol lipids.** *J Chromatogr A* 2006, **1110**:15-26.
  116. Byrdwell WC: **Atmospheric pressure chemical ionization mass spectrometry for analysis of lipids.** *Lipids* 2001, **36**:327-346.
  117. Rauha JP, Vuorela H, Kostianen R: **Effect of eluent on the ionization efficiency of flavonoids by ion spray, atmospheric pressure chemical ionization, and atmospheric pressure photoionization mass spectrometry.** *J Mass Spectrom* 2001, **36**:1269-1280.
  118. Higashi T, Shimada K: **Derivatization of neutral steroids to enhance their detection characteristics in liquid chromatography-mass spectrometry.** *Analytical and Bioanalytical Chemistry* 2004, **378**:875-882.
  119. Hayen H, Jachmann N, Vogel M, Karst U: **LC-electron capture-APCI(-)-MS determination of nitrobenzoxadiazole derivatives.** *Analyst* 2003, **128**:1365-1372.
  120. Higashi T, Takido N, Yamauchi A, Shimada K: **Electron-capturing derivatization of neutral steroids for increasing sensitivity in liquid chromatography/negative atmospheric pressure chemical ionization-mass spectrometry.** *Anal Sci* 2002, **18**:1301-1307.
  121. Dobberstein P, Muenster H: **Application of a New Atmospheric-Pressure Ionization Source for Double Focusing Sector Instruments.** *Journal of Chromatography A* 1995, **712**:3-15.
  122. Fenn JB: **Ion Formation from Charged Droplets - Roles of Geometry, Energy, and Time.** *Journal of the American Society for Mass Spectrometry* 1993, **4**:524-535.
  123. McLuckey SA, Stephenson JL: **Ion ion chemistry of high-mass multiply charged ions.** *Mass Spectrometry Reviews* 1998, **17**:369-407.

124. List GR, Byrdwell WC, Steidley KR, Adlof RO, Neff WE: **Triacylglycerol structure and composition of hydrogenated soybean oil margarine and shortening basestocks.** *J Agric Food Chem* 2005, **53**:4692-4695.
125. Nichols DS, Davies NW: **Improved detection of polyunsaturated fatty acids as phenacyl esters using liquid chromatography-ion trap mass spectrometry.** *J Microbiol Methods* 2002, **50**:103-113.
126. Bily AC, Burt AJ, Ramputh AI, Livesey J, Regnault-Roger C, Philogene BR, Arnason JT: **HPLC-PAD-APCI/MS assay of phenylpropanoids in cereals.** *Phytochem Anal* 2004, **15**:9-15.
127. Stanaszek R, Piekoszewski W: **Simultaneous determination of eight underivatized amphetamines in hair by high-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry (HPLC-APCI-MS).** *J Anal Toxicol* 2004, **28**:77-85.
128. Barker SA, Littlefield-Chabaud MA, David C: **Distribution of the hallucinogens N,N-dimethyltryptamine and 5-methoxy-N,N-dimethyltryptamine in rat brain following intraperitoneal injection: application of a new solid-phase extraction LC-APCI-MS-MS-isotope dilution method.** *J Chromatogr B Biomed Sci Appl* 2001, **751**:37-47.
129. Wada K, Mori T, Kawahara N: **Application of liquid chromatography-atmospheric pressure chemical ionization mass spectrometry to the differentiation of stereoisomeric C19-norditerpenoid alkaloids.** *Chem Pharm Bull (Tokyo)* 2000, **48**:660-668.
130. Frenich AG, Torres ME, Vega AB, Vidal JL, Bolanos PP: **Determination of ascorbic acid and carotenoids in food commodities by liquid chromatography with mass spectrometry detection.** *J Agric Food Chem* 2005, **53**:7371-7376.
131. Schweiggert U, Kammerer DR, Carle R, Schieber A: **Characterization of carotenoids and carotenoid esters in red pepper pods (*Capsicum annuum* L.) by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry.** *Rapid Commun Mass Spectrom* 2005, **19**:2617-2628.
132. Breithaupt DE, Bamedi A: **Carotenoids and carotenoid esters in potatoes (*Solanum tuberosum* L.): new insights into an ancient vegetable.** *J Agric Food Chem* 2002, **50**:7175-7181.
133. Karst U: **Liquid chromatography-electron capture-atmospheric pressure chemical ionization-mass spectrometry.** *Analytical and Bioanalytical Chemistry* 2005, **382**:1744-1746.
134. Hayen H, Jachmann N, Vogel M, Karst U: **LC-Electron capture APCI-MS for the determination of nitroaromatic compounds.** *Analyst* 2002, **127**:1027-1030.
135. Desai MJ, Armstrong DW: **Analysis of native amino acid and peptide enantiomers by high-performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry.** *Journal of Mass Spectrometry* 2004, **39**:177-187.
136. Cai YX, Kingery D, McConnell O, Bach AC: **Advantages of atmospheric pressure photoionization mass spectrometry in support of drug discovery.** *Rapid Communications in Mass Spectrometry* 2005, **19**:1717-1724.
137. Yoshioka N, Akiyama Y, Teranishi K: **Rapid simultaneous determination of o-phenylphenol, diphenyl, thiabendazole, imazalil and its major metabolite in citrus fruits by liquid chromatography-mass spectrometry using atmospheric pressure photoionization.** *Journal of Chromatography A* 2004, **1022**:145-150.
138. Syage JA, Evans MD, Hanold KA: **Photoionization mass spectrometry.** *American Laboratory* 2000, **32**:24-+.
139. Robb DB, Covey TR, Bruins AP: **Atmospheric pressure photoionisation: An ionization method for liquid chromatography-mass spectrometry.** *Analytical Chemistry* 2000, **72**:3653-3659.

140. Kauppila TJ, Kuuranne T, Meurer EC, Eberlin MN, Kotiaho T, Kostianen R: **Atmospheric pressure photoionization mass spectrometry. Ionization mechanism and the effect of solvent on the ionization of naphthalenes.** *Anal Chem* 2002, **74**:5470-5479.
141. Kauppila TJ, Kotiaho T, Kostianen R, Bruins AP: **Negative ion-atmospheric pressure photoionization-mass spectrometry.** *J Am Soc Mass Spectrom* 2004, **15**:203-211.
142. Bos SJ, van Leeuwen SM, Karst U: **From fundamentals to applications: recent developments in atmospheric pressure photoionization mass spectrometry.** *Analytical and Bioanalytical Chemistry* 2006, **384**:85-99.
143. Greig MJ, Bolanos B, Quenzer T, Bylund JM: **Fourier transform ion cyclotron resonance mass spectrometry using atmospheric pressure photoionization for high-resolution analyses of corticosteroids.** *Rapid Commun Mass Spectrom* 2003, **17**:2763-2768.
144. Takino M, Yamaguchi K, Nakahara T: **Determination of carbamate pesticide residues in vegetables and fruits by liquid chromatography-atmospheric pressure photoionization-mass spectrometry and atmospheric pressure chemical ionization-mass spectrometry.** *J Agric Food Chem* 2004, **52**:727-735.
145. Straube E, Dekant W, Volkel W: **Comparison of different LC-MS/MS methods for the determination of nitro aromatic compounds and corresponding reduced metabolites.** *Naunyn-Schmiedeberg's Archives of Pharmacology* 2004, **369**:R140-R140.
146. Moriwaki H, Ishitake M, Yoshikawa S, Miyakoda H, Alary JF: **Determination of polycyclic aromatic hydrocarbons in sediment by liquid chromatography-atmospheric pressure photoionization-mass spectrometry.** *Anal Sci* 2004, **20**:375-377.
147. Trosken ER, Straube E, Lutz WK, Volkel W, Patten C: **Quantitation of lanosterol and its major metabolite FF-MAS in an inhibition assay of CYP51 by azoles with atmospheric pressure photoionization based LC-MS/MS.** *J Am Soc Mass Spectrom* 2004, **15**:1216-1221.
148. Lembcke J, Ceglarek U, Fiedler GM, Baumann S, Leichtle A, Thiery J: **Rapid quantification of free and esterified phytosterols in human serum using APPI-LC-MS/MS.** *J Lipid Res* 2005, **46**:21-26.
149. Sforza S, Dall'asta C, Marchelli R: **Recent advances in mycotoxin determination in food and feed by hyphenated chromatographic techniques/mass spectrometry.** *Mass Spectrom Rev* 2006, **25**:54-76.
150. Cavaliere C, Foglia P, Pastorini E, Samperi R, Lagana A: **Liquid chromatography/tandem mass spectrometric confirmatory method for determining aflatoxin M1 in cow milk - Comparison between electrospray and atmospheric pressure photoionization sources.** *Journal of Chromatography A* 2006, **1101**:69-78.
151. van Leeuwen SM, Hendriksen L, Karst U: **Determination of aldehydes and ketones using derivatization with 2,4-dinitrophenylhydrazine and liquid chromatography-atmospheric pressure photoionization-mass spectrometry.** *J Chromatogr A* 2004, **1058**:107-112.
152. Takino M, Daishima S, Nakahara T: **Determination of chloramphenicol residues in fish meats by liquid chromatography-atmospheric pressure photoionization mass spectrometry.** *J Chromatogr A* 2003, **1011**:67-75.
153. Bolanos B, Greig M, Ventura M, Farrell W, Aurigemma CM, Li H, Quenzer TL, Tivel K, Bylund JMR, Tran P, et al.: **SFC/MS in drug discovery at Pfizer, La Jolla.** *International Journal of Mass Spectrometry* 2004, **238**:85-97.
154. Chen Z, Chen B, Yao SZ: **High-performance liquid chromatography/electrospray ionization-mass spectrometry for simultaneous determination of taurine and**

- 10 water-soluble vitamins in multivitamin tablets.** *Analytica Chimica Acta* 2006, **569**:169-175.
155. Plumb R, Castro-Perez J, Granger J, Beattie I, Joncour K, Wright A: **Ultra-performance liquid chromatography coupled to quadrupole-orthogonal time-of-flight mass spectrometry.** *Rapid Commun Mass Spectrom* 2004, **18**:2331-2337.
156. Zhang X, Cambrai A, Miesch M, Roussi S, Raul F, Aoude-Werner D, Marchioni E: **Separation of Delta5- und Delta7-phytosterols by adsorption chromatography and semipreparative reversed phase high-performance liquid chromatography for quantitative analysis of phytosterols in food.** *Journal of Agricultural and Food Chemistry* 2006, **54**:1196-1202.
157. Louter AJ: **Determination of plant sterol oxidation products in plant sterol enriched spreads, fat blends, and plant sterol concentrates.** *Journal of AOAC International* 2004, **87**:485-492.
158. Kahle K, Kraus M, Richling E: **Polyphenol profiles of apple juices.** *Molecular Nutrition and Food Research* 2005, **49**:797-806.
159. Vrhovsek U, Rigo A, Tonon D, Mattivi F: **Quantitation of polyphenols in different apple varieties.** *Journal of Agricultural and Food Chemistry* 2004, **52**:6532-6538.
160. Takats Z, Wiseman JM, Cooks RG: **Ambient mass spectrometry using desorption electrospray ionization (DESI): instrumentation, mechanisms and applications in forensics, chemistry, and biology.** *Journal of Mass Spectrometry* 2005, **40**:1261-1275.
161. Cooks RG, Ouyang Z, Takats Z, Wiseman JM: **Detection Technologies. Ambient mass spectrometry.** *Science* 2006, **311**:1566-1570.
162. Myung S, Wiseman JM, Valentine SJ, Takats Z, Cooks RG, Clemmer DE: **Coupling desorption electrospray ionization with ion mobility/mass spectrometry for analysis of protein structure: evidence for desorption of folded and denaturated States.** *Journal of Physical Chemistry B* 2006, **110**:5045-5051.
163. Bijlsma S, Boboldijk I, Verheij ER, Ramaker R, Kochhar S, Macdonald IA, van Ommen B, Smilde AK: **Large-Scale Human Metabolomics Studies: A strategy for Data (Pre-) Processing and Validation.** *Analytical Chemistry* 2006, **78**:567-574.
164. Bhalla R, Narasimhan K, Swarup S: **Metabolomics and its role in understanding cellular responses in plants.** *Plant Cell Rep* 2005, **24**:562-571.
165. Weckwerth W: **Metabolomics in systems biology.** *Annu Rev Plant Biol* 2003, **54**:669-689.
166. Diamantaras K, Kung S: *Principal Components Neural Networks.* New York: Wiley; 1996.
167. Hyvärinen A, Karhunen J, Oja E: *Independent Component Analysis.* New York: Wiley; 2001.
168. Scholz M, Gatzek S, Sterling A, Fiehn O, Selbig J: **Metabolite fingerprinting: detecting biological features by independent component analysis.** *Bioinformatics* 2004, **20**:2447-2454.
169. Bortz J: *Statistik für Sozialwissenschaftler.* Berlin: Springer Verlag; 1999.
170. Bacher J: *Clusteranalyse.* München: Oldenbourg Verlag; 1994.
171. Kaufman L, Rousseeuw PJ: *Finding groups in data.* New York: Wiley; 1990.
172. Galbraith SJ, Tran LM, Liao JC: **Transcriptome network component analysis with limited microarray data.** *Bioinformatics* 2006:in press.
173. Scott MS, Perkins T, Bunnell S, Pepin F, Thomas DY, Hallett M: **Identifying regulatory subnetworks for a set of genes.** *Molecular and Cellular Proteomics* 2005, **4**:683-692.
174. Uhrig JF: **Protein interaction networks in plants.** *Planta* 2006:in press.

175. Lubovac Z, Gamalielsson J, Olsson B: **Combining functional and topological properties to identify core modules in protein interaction networks.** *Proteins* 2006:in press.
176. Mumby M, Brekken D: **Phosphoproteomics: new insights into cellular signalling.** *Genome Biology* 2005, **9**:2301-2308.
177. Boyes DC, Zayed AM, Ascenzi R, McCaskill AJ, Hoffman NE, Davis KR, Gorlach J: **Growth stage-based phenotypic analysis of arabidopsis: A model for high throughput functional genomics in plants.** *Plant Cell* 2001, **13**:1499-1510.
178. Morgenthal K, Wienkoop S, Scholz M, Selbig J, Weckwerth W: **Correlative GC/TOF/MS based metabolite profiling and LC/MS based protein profiling reveal time-related systemic regulation of metabolite-protein networks and improve pattern recognition for multiple biomarker selection.** *Metabolomics* 2005.
179. Morgenthal K, Weckwerth W, Steuer R: **Metabolomic networks in plants: Transitions from pattern recognition to biological interpretation.** *Biosystems* 2005.
180. Harn C, Khayat E, Daie J: **Expression Dynamics of Genes Encoding Key Carbon Metabolism Enzymes During Sink to Source Transition of Developing Leaves.** *Plant and Cell Physiology* 1993, **34**:1045-1053.
181. Jeong ML, Jiang H, Chen H-S, Tsai C-J, Harding SA: **Metabolic Profiling of the Sink-to-Source Transition in Developing Leaves of Quaking Aspen.** *Plant Physiol.* 2004, **136**:3364-3375.
182. Garcia-Ojalvo J, Sancho JM: *Noise in Spatially Extended Systems.* New York: Springer-Verlag; 1999.
183. Pettersson G, Ryde-Pettersson U: **A mathematical model of the Calvin photosynthesis cycle.** *Eur J Biochem* 1988, **175**:661-672.
184. Poolman MG, Fell DA, Thomas S: **Modelling photosynthesis and its control.** *J Exp Bot* 2000, **51 Spec No**:319-328.
185. Noe SM, Giersch C: **A simple dynamic model of photosynthesis in oak leaves: coupling leaf conductance and photosynthetic carbon fixation by variable intracellular CO<sub>2</sub> pool.** *Funct Plant Biol* 2004, **31**:1195-1204.
186. Kose F, Weckwerth W, Linke T, Fiehn O: **Visualizing plant metabolomic correlation networks using clique-metabolite matrices.** *Bioinformatics* 2001, **17**:1198-1208.
187. Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR: **Metabolic Profiling Allows Comprehensive Phenotyping of Genetically or Environmentally Modified Plant Systems.** *Plant Cell* 2001, **13**:11-29.
188. Martins AM, Camacho D, Shuman J, Sha W, Mendes P, Shulaev V: **A systems biology study of two distinct growth phases of *Saccharomyces cerevisiae* cultures.** *Current Genomics* 2004, **5**:649-663.
189. Steuer R, Kurths J, Fiehn O, Weckwerth W: **Observing and interpreting correlations in metabolomic networks.** *Bioinformatics* 2003, **19**:1019-1026.
190. Camacho D, de la Fuente A, Mendes P: **The origin of correlations in metabolomic data.** *Metabolomics* 2005, **1**:53-63.
191. Heinrich R, Schuster S: *The Regulation of Cellular Systems.* New York: Chapman & Hall; 1996.
192. Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A: **Multistability in the lactose utilization network of *Escherichia coli*.** *Nature* 2004, **427**:737-740.
193. Kauffman SA: **Metabolic stability and epigenesis in randomly constructed genetic nets.** *J Theor Biol* 1969, **22**:437-467.
194. Poolman MG, Olcer H, Lloyd JC, Raines CA, Fell DA: **Computer modelling and experimental evidence for two steady states in the photosynthetic Calvin cycle.** *Eur J Biochem* 2001, **268**:2810-2816.

195. Caspar T, Huber SC, Somerville C: **Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis-Thaliana (L) Deficient in Chloroplast Phosphoglucomutase Activity.** *Plant Physiology* 1985, **79**:11-17.
196. Kopka J, Fernie A, Weckwerth W, Gibon Y, Stitt M: **Metabolite profiling in plant biology: platforms and destinations.** *Genome Biol* 2004, **5**:109.
197. Lippincott J, Apostol I: **Carbamylation of cysteine: a potential artifact in peptide mapping of hemoglobins in the presence of urea.** *Anal Biochem* 1999, **267**:57-64.
198. Tanaka N, Kobayashi H: **Monolithic columns for liquid chromatography.** *Anal Bioanal Chem* 2003, **376**:298-301.
199. Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson J: *Molekulare Organisation von Zellen.* Wienheim: VCH Verlagsgesellschaft; 1994.
200. Lea P: *Primary nitrogen metabolism.* Edited by B. HJ. New York: Academic Press; 1997.
201. Geiger DR, Servaites JC: **Diurnal regulation of photosynthetic carbon metabolism in C3 plants.** *Annu Rev Plant Biol* 1994, **45**:235-256.
202. Geiger DR, Servaites JC, Fuchs MA: **Role of starch in carbon translocation and partitioning at the plant level.** *J Plant Physiol* 2000, **27**:571-582.
203. Gibon Y, Blasing OE, Palacios-Rojas N, Pankovic D, Hendriks JH, Fisahn J, Hohne M, Gunther M, Stitt M: **Adjustment of diurnal starch turnover to short days: depletion of sugar during the night leads to a temporary inhibition of carbohydrate utilization, accumulation of sugars and post-translational activation of ADP-glucose pyrophosphorylase in the following light period.** *Plant J* 2004, **39**:847-862.
204. Schaffer R, Landgraf J, Accerbi M, Simon V, Larson M, Wisman E: **Microarray Analysis of Diurnal and Circadian-Regulated Genes in Arabidopsis.** *Plant Cell* 2001, **13**:113-123.
205. Blasing OE, Gibon Y, Gunther M, Hohne M, Morcuende R, Osuna D, Thimm O, Usadel B, Scheible WR, Stitt M: **Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis.** *Plant Cell* 2005, **17**:3257-3281.
206. Kim JY, Song HR, Taylor BL, Carre IA: **Light-regulated translation mediates gated induction of the Arabidopsis clock protein LHY.** *EMBO J* 2003, **22**:925-944.
207. Mas P, Kim WY, Somers DE, Kay SA: **Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana.** *Nature* 2003, **426**:567-570.
208. Wang ZY, Tobin EM: **Constitutively expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own degradation.** *Cell* 1998, **93**:1207-1217.
209. Heintzen C, Nater M, Apel K, Staiger D: **AtGRP7, a nuclear RNA-binding Protein as a component of a circadian-regulated negative feedback loop in Arabidopsis thaliana.** *Proc Natl Acad Sci U S A* 1997, **94**:8515-8520.
210. Gibon Y, Blasing OE, Hannemann J, Carillo P, Hohne M, Hendriks JH, Palacios N, Cross J, Selbig J, Stitt M: **A Robot-Based Platform to Measure Multiple Enzyme Activities in Arabidopsis Using a Set of Cycling Assays: Comparison of Changes of Enzyme Activities and Transcript Levels during Diurnal Cycles and in Prolonged Darkness.** *Plant Cell* 2004.
211. Scheible WR, Gonzalez-Fontes A, Morcuende R, Lauerer M, Geiger M, Glaab J, Schulze ED, Stitt M: **Tobacco mutants with a decreased number of functional nia-genes compensate by modifying the diurnal regulation transcription, post-translational modification and turnover of nitrate reductase.** *Planta* 1997, **203**:305-319.
212. Smith SM, Fulton DC, Chia T, Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM: **Diurnal Changes in the Transcriptome Encoding**

- Enzymes of Starch Metabolism Provide Evidence for Both Transcriptional and Posttranscriptional Regulation of Starch Metabolism in Arabidopsis Leaves.** *Plant Physiol.* 2004, **136**:2687-2699.
213. Lu Y, Gehan JP, Sharkey TD: **Daylength and circadian effects on starch degradation and maltose metabolism.** *Plant Physiol* 2005, **138**:2280-2291.
214. Gutierrez RA, Ewing RM, Cherry JM, Green PJ: **Identification of unstable transcripts in Arabidopsis by cDNA microarray analysis: Rapid decay is associated with a group of touch- and specific clock-controlled genes.** *PNAS* 2002, **99**:11513-11518.
215. Simpson E: **Measurement of protein degradation in leaves of Zea mays using [<sup>3</sup>H]acetic anhydrid and tritiated water.** *Plant Physiol* 1981, **67**:1214-1219.
216. Stitt M: **Rising CO<sub>2</sub> levels and their potential significance for carbon flow in photosynthetic cells.** *Plant Cell Environment* 1991, **14**:741-762.
217. Pego JV, Kortstee AJ, Huijser C, Smeekens SCM: **Photosynthesis, sugars and the regulation of gene expression.** *J. Exp. Bot.* 2000, **51**:407-416.
218. Jang JC, Sheen J: **Sugar sensing in higher plants.** *Plant Cell* 1994, **6**:1665-1679.
219. Krapp A, Hofmann B, Schafer C, Stitt M: **Regulation of the Expression of Rbcs and Other Photosynthetic Genes by Carbohydrates - a Mechanism for the Sink Regulation of Photosynthesis.** *Plant Journal* 1993, **3**:817-828.
220. Fonseca F, Bowsher C, Stulen I: **Impact of elevated atmospheric carbon dioxide on nitrate reductase transcription and activity in leaves and roots of *Plantago major*.** *Physiologia Plantarum* 1997, **100**:940-948.
221. Cheng CL, Acedo GN, Cristinsin M, Conkling MA: **Sucrose mimics the light induction of Arabidopsis nitrate reductase gene transcription.** *Proc Natl Acad Sci U S A* 1992, **89**:1861-1864.
222. Koch KE: **Carbohydrate-Modulated Gene Expression in Plants.** *Annu Rev Plant Physiol Plant Mol Biol* 1996, **47**:509-540.
223. Dijkwel PP, Kock P, Bezemer R, Weisbeek PJ, Smeekens S: **Sucrose Represses the Developmentally Controlled Transient Activation of the Plastocyanin Gene in Arabidopsis thaliana Seedlings.** *Plant Physiol* 1996, **110**:455-463.
224. Goldschmidt EE, Huber SC: **Regulation of Photosynthesis by End-Product Accumulation in Leaves of Plants Storing Starch, Sucrose, and Hexose Sugars.** *Plant Physiology* 1992, **99**:1443-1448.
225. Vanoosten JJ, Wilkins D, Besford RT: **Regulation of the Expression of Photosynthetic Nuclear Genes by CO<sub>2</sub> Is Mimicked by Regulation by Carbohydrates - a Mechanism for the Acclimation of Photosynthesis to High CO<sub>2</sub>.** *Plant Cell and Environment* 1994, **17**:913-923.
226. Krapp A, Hofmann B, Schäfer C, Stitt M: **Regulation of the expression of rbcS and other photosynthetic genes by carbohydrates: a mechanism for the 'sink regulation' of photosynthesis?** *Plant J* 1993, **3**:817-828.
227. Zhang N, Kallis RP, Ewy RG, Portis AR, Jr.: **From the Cover: Light modulation of Rubisco in Arabidopsis requires a capacity for redox regulation of the larger Rubisco activase isoform.** *PNAS* 2002, **99**:3330-3334.
228. Sage RF, Sharkey TD, Seemann JR: **Regulation of Ribulose-1,5-Bisphosphate Carboxylase Activity in Response to Light Intensity and CO<sub>2</sub> in the C(3) Annuals *Chenopodium album* L. and *Phaseolus vulgaris* L.** *Plant Physiol* 1990, **94**:1735-1742.
229. Portis AR, Jr.: **Rubisco activase - Rubisco's catalytic chaperone.** *Photosynth Res* 2003, **75**:11-27.
230. Zhang N, Portis AR, Jr.: **Mechanism of light regulation of Rubisco: a specific role for the larger Rubisco activase isoform involving reductive activation by thioredoxin-f.** *Proc Natl Acad Sci U S A* 1999, **96**:9438-9443.

231. Brooks A, Portis AR: **Protein-Bound Ribulose Bisphosphate Correlates with Deactivation of Ribulose Bisphosphate Carboxylase in Leaves.** *Plant Physiol* 1988, **87**:244-249.
232. Perchorowicz JT, Raynes DA, Jensen RG: **Light limitation of photosynthesis and activation of ribulose bisphosphate carboxylase in wheat seedlings.** *Proc Natl Acad Sci U S A* 1981, **78**:2985-2989.
233. Brouquisse R, James F, Raymond P, Pradet A: **Study of Glucose Starvation in Excised Maize Root Tips.** *Plant Physiol* 1991, **96**:619-626.
234. Sivakumar P, Sharmila P, Jain V, Pardha Saradhi P: **Sugars have potential to curtail oxygenase activity of Rubisco.** *Biochem Biophys Res Commun* 2002, **298**:247-250.
235. Ott T, Clarke J, Birks K, Johnson G: **Regulation of the photosynthetic electron transport chain.** *Planta* 1999, **209**:250-258.
236. Eichelmann H, Laisk A: **CO<sub>2</sub> Uptake and Electron Transport Rates in Wild-Type and a Starchless Mutant of *Nicotiana sylvestris* (The Role and Regulation of Starch Synthesis at Saturating CO<sub>2</sub> Concentrations).** *Plant Physiol* 1994, **106**:679-687.
237. Foyer C, Furbank R, Harbinson J, Horton P: **The Mechanisms Contributing to Photosynthetic Control of Electron-Transport by Carbon Assimilation in Leaves.** *Photosynthesis Research* 1990, **25**:83-100.
238. Sheen J: **Metabolic repression of transcription in higher plants.** *Plant Cell* 1990, **2**:1027-1038.
239. Graciet E, Lebreton S, Gontero B: **Emergence of new regulatory mechanisms in the Benson-Calvin pathway via protein-protein interactions: a glyceraldehyde-3-phosphate dehydrogenase/CP12/phosphoribulokinase complex.** *J. Exp. Bot.* 2004, **55**:1245-1254.
240. Marri L, Sparla F, Pupillo P, Trost P: **Co-ordinated gene expression of photosynthetic glyceraldehyde-3-phosphate dehydrogenase, phosphoribulokinase, and CP12 in *Arabidopsis thaliana*.** *J Exp Bot* 2005, **56**:73-80.
241. Shih M-C, Goodman, H.M: **Differential light-regulated expression of nuclear genes encoding chlrotoplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase in *Nicotiana tabacum*.** *EMBO J* 1988, **7**:893-898.
242. Yang Y, Kwon H-B, Peng H-P, Shih M-C: **Stress response and metabolic regulation of glyceraldehyde-3-phosphate dehydrogenase genes in *Arabidopsis*.** *Plant Physiol.* 1993, **101**:209-216.
243. Leonhardt N, Kwak JM, Robert N, Waner D, Leonhardt G, Schroeder JI: **Microarray Expression Analyses of *Arabidopsis* Guard Cells and Isolation of a Recessive Abscisic Acid Hypersensitive Protein Phosphatase 2C Mutant.** *Plant Cell* 2004, **16**:596-615.
244. Chen Z, Silva H, Klessig DF: **Active oxygen species in the induction of plant systemic acquired resistance by salicylic acid.** *Science* 1993, **262**:1883-1886.
245. Shirasu K, Nakajima H, Rajasekhar VK, Dixon RA, Lamb C: **Salicylic acid potentiates an agonist-dependent gain control that amplifies pathogen signals in the activation of defense mechanisms.** *Plant Cell* 1997, **9**:261-270.
246. Mateo A, Muhlenbock P, Rusterucci C, Chang CC, Miszalski Z, Karpinska B, Parker JE, Mullineaux PM, Karpinski S: **LESION SIMULATING DISEASE 1 is required for acclimation to conditions that promote excess excitation energy.** *Plant Physiol* 2004, **136**:2818-2830.
247. Bowling SA, Guo A, Cao H, Gordon AS, Klessig DF, Dong X: **A mutation in *Arabidopsis* that leads to constitutive expression of systemic acquired resistance.** *Plant Cell* 1994, **6**:1845-1857.

248. Mauch F, Mauch-Mani B, Gaille C, Kull B, Haas D, Reimann C: **Manipulation of salicylate content in Arabidopsis thaliana by the expression of an engineered bacterial salicylate synthase.** *Plant J* 2001, **25**:67-77.
249. Scott IM, Clarke SM, Wood JE, Mur LA: **Salicylate accumulation inhibits growth at chilling temperature in Arabidopsis.** *Plant Physiol* 2004, **135**:1040-1049.
250. Mateo A, Funck D, Muhlenbock P, Kular B, Mullineaux PM, Karpinski S: **Controlled levels of salicylic acid are required for optimal photosynthesis and redox homeostasis.** *J Exp Bot* 2006.
251. van Rensen JJS, Xu CH, Govindjee: **Role of bicarbonate in photosystem II, the water-plastoquinone oxido-reductase of plant photosynthesis.** *Physiologia Plantarum* 1999, **105**:585-592.
252. Fryer MJ, Ball L, Oxborough K, Karpinski S, Mullineaux PM, Baker NR: **Control of Ascorbate Peroxidase 2 expression by hydrogen peroxide and leaf water status during excess light stress reveals a functional organisation of Arabidopsis leaves.** *Plant J* 2003, **33**:691-705.
253. Contento AL, Kim SJ, Bassham DC: **Transcriptome profiling of the response of Arabidopsis suspension culture cells to Suc starvation.** *Plant Physiology* 2004, **135**:2330-2347.
254. Dieuaide M, Brouquisse R, Pradet A, Raymond P: **Increased Fatty Acid beta-Oxidation after Glucose Starvation in Maize Root Tips.** *Plant Physiol* 1992, **99**:595-600.
255. Averill-Bates DA, Przybytkowski E: **The role of glucose in cellular defences against cytotoxicity of hydrogen peroxide in Chinese hamster ovary cells.** *Arch Biochem Biophys* 1994, **312**:52-58.
256. Zhang YM, Wong TY, Chen LY, Lin CS, Liu JK: **Induction of a futile Embden-Meyerhof-Parnas pathway in Deinococcus radiodurans by Mn: possible role of the pentose phosphate pathway in cell survival.** *Appl Environ Microbiol* 2000, **66**:105-112.
257. Foyer C: *Ascorbic acid.* Edited by Hess JL. Boca Raton: CRC Press; 1993.
258. Pallett KE, Young AJ: *Carotenoids.* Edited by Hess JL. Boca Raton: CRC Press; 1993.
259. Noctor G, Foyer CH: **ASCORBATE AND GLUTATHIONE: Keeping Active Oxygen Under Control.** *Annu Rev Plant Physiol Plant Mol Biol* 1998, **49**:249-279.
260. Strand A: **Plastid-to-nucleus signalling.** *Curr Opin Plant Biol* 2004, **7**:621-625.
261. Johannningmeier U: **Possible control of transcript levels by chlorophyll precursors in Chlamydomonas.** *Eur J Biochem* 1988, **177**:417-424.
262. Coruzzi G, Bush DR: **Nitrogen and carbon nutrient and metabolite signaling in plants.** *Plant Physiology* 2001, **125**:61-64.
263. Paul MJ, Driscoll SP: **Sugar repression of photosynthesis: The role of carbohydrates in signalling nitrogen deficiency through source:sink imbalance.** *Plant Cell and Environment* 1997, **20**:110-116.
264. Levitt J: *Responses of plants to environmental stress.* New York: Academic Press; 1972.
265. Shinozaki K, Yamaguchi-Shinozaki K: **Gene expression and signal transduction in water-stress response.** *Plant Physiol* 1997, **115**.
266. Cook D, Fowler S, Fiehn O, Thomashow MF: **A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis.** *Proceedings of the National Academy of Sciences of the United States of America* 2004, **101**:15243-15248.
267. Kaplan F, Kopka J, Haskell DW, Zhao W, Schiller KC, Gatzke N, Sung DY, Guy CL: **Exploring the temperature-stress metabolome of Arabidopsis.** *Plant Physiol* 2004, **136**:4159-4168.
268. Müller-Thurgau H: *Über das Gefrieren und Erfrieren der Pflanzen*, vol XV. Berlin: Landwirtschaftliche Jahrbücher; 1886.

269. Rabbani MA, Maruyama K, Abe H, Khan MA, Katsura K, Ito Y, Yoshiwara K, Seki M, Shinozaki K, Yamaguchi-Shinozaki K: **Monitoring expression profiles of rice genes under cold, drought, and high-salinity stresses and abscisic acid application using cDNA microarray and RNA gel-blot analysis.** *Plant Physiol* 2003, **2003**:1755-1767.
270. Rizhsky L, Liang H, Shuman J, Shulaev V, Davletova S, Mittler R: **When defense pathways collide. The response of *Arabidopsis* to a combination of drought and heat stress.** *Plant Physiol* 2004, **134**:1683-1696.
271. Hughes MA, Dunn MA: **The molecular biology of plant acclimation to low temperature.** *J Exp Bot* 1996, **47**:291-305.
272. Kreps JA, Wu Y, Chang H-S, Zhu T, Wang X, Harper JF: **Transcriptome Changes for *Arabidopsis* in Response to Salt, Osmotic, and Cold Stress.** *Plant Physiol* 2002, **130**:2129-2141.
273. Mittler R: **Oxidative stress, antioxidants and stress tolerance.** *Trends Plant Sci* 2002, **9**:405-410.
274. Smirnov N: **Plant resistance to environmental stress.** *Curr Opin Biotechnol* 1998, **9**:214-219.
275. Seki M, Narusaka M, Abe H, Kasuga M, Yamaguchi-Shinozaki K, Carninci P, Hayashizaki Y, Shinozaki K: **Monitoring the expression pattern of 1300 *Arabidopsis* genes under drought and cold stresses by using a full-length cDNA microarray.** *Plant Cell* 2002, **13**:61-72.
276. Guy CL, Huber JLA, Huber SC: **Sucrose phosphate synthase and sucrose accumulation at low temperature.** *Plant Physiol* 1992, **100**:502-508.
277. Fan TW-M, Colmer TD, Lane AN, Higashi RM: **Determination of metabolites by <sup>1</sup>H NMR and GC: analysis for organic osmolytes in crude tissue extracts.** *Anal Biochem* 1993, **214**:260-271.
278. Gilmour SJ, Sebolt AM, Salazar MP, Everard JD, Thomashow MF: **Overexpression of the *Arabidopsis* CBF3 transcriptional activator mimics multiple biochemical changes associated with cold acclimation.** *Plant Physiol* 2000, **124**:1854-1865.
279. Shinozaki K, Yamaguchi-Shinozaki K, Seki M: **Regulatory network of gene expression in the drought and cold responses.** *Curr Opin Plant Biol* 2003, **6**:410-417.
280. Liu H, Sadygov RG, Yates RI: **A model for random sampling and estimation of relative protein abundance in shotgun proteomics.** *Anal Chem* 2004, **76**:4193-4201.
281. Misquitta Y, Caffrey M: **Rational design of lipid molecular structure: a case study involving the C19:1c10 monoacylglycerol.** *Biophys J* 2001, **81**:1047-1058.
282. Binder WH, Barragan V, Menger FM: **Domains and rafts in lipid membranes.** *Angew Chem Int Ed Engl* 2003, **42**:5802-5827.
283. Drobnis EZ, Crowe LM, Berger T, Anchordoguy TJ, Overstreet JW, Crowe JH: **Cold shock damage is due to lipid phase transitions in cell membranes: a demonstration using sperm as a model.** *J Exp Zool* 1993, **265**:432-437.
284. Sinensky M: **Homeoviscous adaptation--a homeostatic process that regulates the viscosity of membrane lipids in *Escherichia coli*.** *Proc Natl Acad Sci U S A* 1974, **71**:522-525.
285. Gordon-Kamm WJ, Steponkus PL: **Lamellar-to-hexagonal<sub>II</sub> phase transitions in the plasma membrane of isolated protoplasts after freeze-induced dehydration.** *Proc Natl Acad Sci U S A* 1984, **81**:6373-6377.
286. Pardossi A, Vernieri P, Tognoni F: **Involvement of Abscisic Acid in Regulating Water Status in *Phaseolus vulgaris* L. during Chilling.** *Plant Physiol* 1992, **100**:1243-1250.

287. Lee TM, Lur HS, Chu C: **Role of Abscisic-Acid in Chilling Tolerance of Rice (*Oryza-Sativa L*) Seedlings .1. Endogenous Abscisic-Acid Levels.** *Plant Cell and Environment* 1993, **16**:481-490.
288. Fennell A, Markhart AH: **Rapid acclimation of root hydraulic conductivity to low temperature.** *Journal of Experimental Botany* 1998, **49**:879-884.
289. Matzner S, Comstock J: **The temperature dependence of shoot hydraulic resistance: implications for stomatal behaviour and hydraulic limitation.** *Plant Cell and Environment* 2001, **24**:1299-1307.
290. Ohlrogge J, Browse J: **Lipid biosynthesis.** *Plant Cell* 1995, **7**:957-970.
291. Harwood JL: **Recent advances in the biosynthesis of plant fatty acids.** *Biochim Biophys Acta* 1996, **1301**:7-56.
292. Browse J, Somerville C: **Glycerolipid Synthesis - Biochemistry and Regulation.** *Annual Review of Plant Physiology and Plant Molecular Biology* 1991, **42**:467-506.
293. Somerville C, Browse J: **Plant Lipids - Metabolism, Mutants, and Membranes.** *Science* 1991, **252**:80-87.
294. Fatland BL, Ke J, Anderson MD, Mentzen WI, Cui LW, Allred CC, Johnston JL, Nikolau BJ, Wurtele ES: **Molecular characterization of a heteromeric ATP-citrat lyase that generates cytosolic acetyl-coenzym A in Arabidopsis.** *Plant Physiol* 2002, **130**:740-756.
295. Fatland BL, Nikolau BJ, Wurtele ES: **Reverse genetic characterization of cytosolic Acety-CoA generation by ATP-citrat lyase in Arabidopsis.** *Plant Cell* 2005, **17**:182-203.
296. Iba K: **Acclimative response to temperature stress in higher plants: approaches of gene engineering for temperature tolerance.** *Annu Rev Plant Biol* 2002, **53**:225-245.
297. Thomashow MF: **Role of cold-responsive genes in plant freezing tolerance.** *Plant Physiol* 1998, **118**:1-8.
298. Thomashow MF: **PLANT COLD ACCLIMATION: Freezing Tolerance Genes and Regulatory Mechanisms.** *Annu Rev Plant Physiol Plant Mol Biol* 1999, **50**:571-599.
299. Qiunn PJ: **A lipid phase separation model of low-temperature damage to biological membranes.** *Cryobiol* 1985, **22**:28-46.
300. Singer MA, Linquist S: **Multiple effects of trehalose on protein folding *in vitro* and *in vivo*.** *Mol Cell* 1998, **639-648**.
301. Diamant S, Eliahu N, Rosenthal D, Goloubinoff P: **Chemical chaperones regulate molecular chaperones *in vitro* and in cells under combined salt and heat stresses.** *J Biol Chem* 2001, **276**:39586-39591.
302. Strand A, Hurry V, Gustafsson P, Gardestrom P: **Development of Arabidopsis thaliana leaves at low temperature releases the suppression of photosynthesis and photosynthetic gene expression despite the accumulation of soluble carbohydrates.** *Plant J* 1997, **12**:605-614.
303. Krapp A, Stitt M: **An evaluation of direct and indirect mechanisms for the "sink-regulation" of photosynthesis in spinach: changes in gas exchange, carbohydrates, metabolites, enzyme activities and steady-state transcript levels after cold-gridling source leaves.** *Planta* 1995, **1995**:313-323.
304. Taji T, Ohsumi C, Iuchi S, Seki M, Kasuga M, Kobayashi M, Yamaguchi-Shinozaki K, Shinozaki K: **Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in Arabidopsis thaliana.** *Plant J* 2002, **29**:417-426.
305. Liu JJJ, Krenz DC, Galvez AF, de Lumen BO: **Galactinol synthase (GS): increased enzyme activity and levels of mRNA due to cold and desiccation.** *Plant Science* 1998, **134**:11-20.

306. Zuther E, Buchel K, Hundertmark M, Stitt M, Hinch DK, Heyer AG: **The role of raffinose in the cold acclimation response of *Arabidopsis thaliana*.** *FEBS Lett* 2004, **576**:169-173.
307. Hare PD, Cress WA, Van Staden J: **Dissecting the roles of osmolyte accumulation during stress.** *Plant Cell and Environment* 1998, **21**:535-553.
308. Rontein D, Basset G, Hanson AD: **Metabolic engineering of osmoprotectant accumulation in plants.** *Metab Eng* 2002, **4**:49-56.
309. Kishor P, Hong Z, Miao GH, Hu C, Verma D: **Overexpression of [delta]-Pyrroline-5-Carboxylate Synthetase Increases Proline Production and Confers Osmotolerance in Transgenic Plants.** *Plant Physiol* 1995, **108**:1387-1394.
310. Nanjo T, Kobayashi M, Yoshiba Y, Sanada Y, Wada K, Tsukaya H, Kakubari Y, Yamaguchi-Shinozaki K, Shinozaki K: **Biological functions of proline in morphogenesis and osmotolerance revealed in antisense transgenic *Arabidopsis thaliana*.** *Plant J* 1999, **18**:185-193.
311. Parvanova D, Ivanov S, Konstantinova T, Karanov E, Atanassov A, Tsvetkov T, Alexieva V, Djilianov D: **Transgenic tobacco plants accumulating osmolytes show reduced oxidative damage under freezing stress.** *Plant Physiol Biochem* 2004, **42**:57-63.
312. Kolodiaznaia Ia S, Titov SE, Kochetov AV, Komarova ML, Romanova AV, Koval VS, Shumnyi VK: **[Evaluation of salt tolerance in *Nicotiana tabacum* plants bearing an antisense suppressor of the proline dehydrogenase gene].** *Genetika* 2006, **42**:278-281.
313. van Rensburg L, Krüger GHJ, Krüger H: **Proline accumulation as drought-tolerance selection criterion: its relationship to membrane integrity and chloroplast ultrastructure in *Nicotiana tabacum*.** *J Plant Physiol* 1993, **141**:188-194.
314. Fricke W, Pahlich E: **The Effect of Water-Stress on the Vacuole-Extravacuole Compartmentation of Proline in Potato Cell-Suspension Cultures.** *Physiologia Plantarum* 1990, **78**:374-378.
315. Hare PD, Cress WA: **Metabolic implications of stress-induced proline accumulation in plants.** *Plant Growth Regulation* 1997, **21**:79-102.
316. Delauney AJ, Verma DPS: **Proline Biosynthesis and Osmoregulation in Plants.** *Plant Journal* 1993, **4**:215-223.
317. Brugiere N, Dubois F, Limami AM, Lelandais M, Roux Y, Sangwan RS, Hirel B: **Glutamine synthetase in the phloem plays a major role in controlling proline production.** *Plant Cell* 1999, **11**:1995-2011.
318. Guy CL, Niemi KJ, Brambl R: **Altered gene expression during cold acclimation in spinach.** *Proceedings of the National Academy of Sciences of the United States of America* 1985, **82**:3673-3677.
319. Thomashow MF: **Molecular genetics of cold acclimation in higher plants.** *Adv Genet* 1990, **28**:99-131.
320. Amme S, Matros A, Schlesier B, Mock H-P: **Proteome analysis of cold stress response in *Arabidopsis thaliana* using DIGE-technology.** *J Exp Bot* 2006, **57**:1537-1546.
321. Gomez J, Sanchez-Martinez D, Stiefel V, Rigau J, Puigdomenech P, Pages M: **A gene induced by the plant hormone abscisic acid in response to water stress encodes a glycine-rich protein.** *Nature* 1988, **334**:262-264.
322. Hanano S, Sugita M, Sugiura M: **Isolation of a novel RNA-binding protein and its association with a large ribonucleoprotein particle present in the nucleoplasm of tobacco cells.** *Plant Mol Biol* 1996, **31**:57-68.
323. Hirose T, Sugita M, Sugiura M: **cDNA structure, expression and nucleic acid-binding properties of three RNA-binding proteins in tobacco: occurrence of tissue-specific alternative splicing.** *Nucleic Acids Res* 1993, **21**:3981-3987.

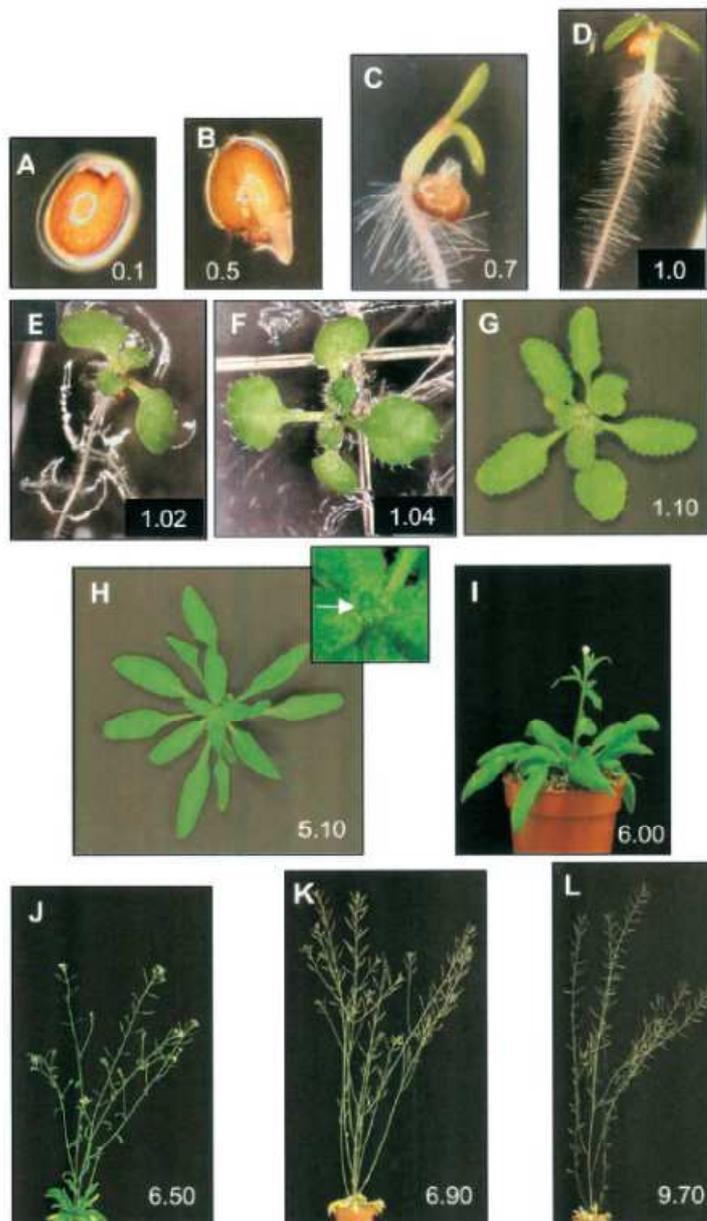
324. Ludevid MD, Freire MA, Gomez J, Burd CG, Albericio F, Giralt E, Dreyfuss G, Pages M: **RNA binding characteristics of a 16 kDa glycine-rich protein from maize.** *Plant J* 1992, **2**:999-1003.
325. Bergeron D, Beauseigle D, Bellemare G: **Sequence and expression of a gene encoding a protein with RNA-binding and glycine-rich domains in Brassica napus.** *Biochim Biophys Acta* 1993, **1216**:123-125.
326. Cretin C, Puigdomenech P: **Glycine-rich RNA-binding proteins from Sorghum vulgare.** *Plant Mol Biol* 1990, **15**:783-785.
327. Kwak KJ, Kim YO, Kang H: **Characterization of transgenic Arabidopsis plants overexpressing GR-RBP4 under high salinity, dehydration, or cold stress.** *J Exp Bot* 2005, **56**:3007-3016.
328. Karlson D, Nakaminami K, Toyomasu T, Imai R: **A cold-regulated nucleic acid-binding protein of winter wheat shares a domain with bacterial cold shock proteins.** *J Biol Chem* 2002, **277**:35248-35256.
329. Sachetto-Martins G, Franco LO, de Oliveira DE: **Plant glycine-rich proteins: a family or just proteins with a common motif?** *Biochim Biophys Acta* 2000, **1492**:1-14.
330. Kim YO, Kim JS, Kang H: **Cold-inducible zinc finger-containing glycine-rich RNA-binding protein contributes to the enhancement of freezing tolerance in Arabidopsis thaliana.** *Plant J* 2005, **42**:890-900.
331. Kim YO, Kang H: **The Role of a Zinc Finger-Containing Glycine-Rich RNA-Binding Protein During Cold Adaptation Process in Arabidopsis thaliana.** *Plant Cell Physiol* 2006.
332. Ohta M, Sugita M, Sugiura M: **3 types of nuclear genes encoding chloroplast RNA-binding proteins (Cp29, Cp31 and Cp33) are present in Arabidopsis thaliana - Presence of Cp31 in chloroplasts and its homolog in nuclei / cytoplasms.** *Plant Mol Biol* 1995, **27**:529-539.
333. Morimoto RI, Tissieres A, Georgopoulos C: *The biology of heat shock proteins and molecular chaperones.* New York: Cold Spring Harbor Lab. Press; 1994.
334. Wang W, Vinocur B, Shoseyov O, Altman A: **Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response.** *Trends Plant Sci* 2004, **9**:244-252.
335. Vierling E: **The Roles of Heat-Shock Proteins in Plants.** *Annual Review of Plant Physiology and Plant Molecular Biology* 1991, **42**:579-620.
336. Tun K, Dhar PK, Palumbo MC, Giuliani A: **Metabolic pathways variability and sequence/networks comparisons.** *BMC Bioinformatics* 2006, **7**:24.
337. Tsang EW, Bowler C, Herouart D, Van Camp W, Villarroel R, Genetello C, Van Montagu M, Inze D: **Differential regulation of superoxide dismutases in plants exposed to environmental stress.** *Plant Cell* 1991, **3**:783-792.
338. Saruyama H, Tanida M: **Effect of chilling on activated oxygen-scavenging enzymes in low-temperature sensitive and tolerant cultivars of rice (Oryza sativa L.).** *Plant Sci* 1995, **109**:105-113.
339. O'Kane D, Gill V, Boyd P, Bourdon R: **Chilling, oxidative stress and antioxidant responses in Arabidopsis thaliana callus.** *Planta* 1996, **198**:371-377.
340. Mittler R, Zilinskas BA: **Molecular cloning and characterization of a gene encoding pea cytosolic ascorbate peroxidase.** *J Biol Chem* 1992, **267**:21802-21807.
341. Storozhenko S, de Pauw P, van Montagu M, Inze D, Kushnir S: **The heat shock element is a functional component of the Arabidopsis APX1 gene promoter.** *Plant Physiol* 1998, **118**:1005-1014.
342. Langenkamper G, Manac'h N, Broin M, Cuine S, Becuwe N, Kuntz M, Rey P: **Accumulation of plastid lipid-associated proteins (fibrillin/CDSP34) upon oxidative stress, ageing and biotic stress in Solanaceae and in response to drought in other species.** *J. Exp. Bot.* 2001, **52**:1545-1554.

343. Bruggemann W, Beyel V, Brodka M, Poth H, Weil M, Stockhaus J: **Antioxidants and antioxidative enzymes in wild-type and transgenic *Lycopersicon* genotypes of different chilling tolerance.** *Plant Science* 1999, **140**:145-154.
344. Vetter J: **Plant cyanogenic glycosides.** *Toxicon* 2000, **38**:11-36.
345. Dixon AA: **Natural products in plant disease resistance.** *Nature* 2001, **411**:843-847.
346. Klee H: **Hormones are in the air.** *Proc Natl Acad Sci U S A* 2003, **100**:10144-10145.
347. Mettraux JP, Signer H, Ryals J, Ward E, Wyss-Benz M, Gaudin J, Raschdorf K, Schmid E, Blum W, Inverardi B: **Increase in salicylic acid at the onset of systemic resistance in cucumber.** *Science* 1990, **1990**:1004-1006.
348. Malamy J, Carr JP, Klessig DF, Raskin I: **Salicylic acid: a like endogenous signal in the resistance response of tobacco to viral infection.** *Science* 1990, **1990**:1002-1004.
349. Mettraux JP, Signer H, Ryals J, Ward E, Wyss-Benz M, Gaudin J, Raschdorf K, Schmidt E, Blum W, Inverardi B: **Increase in salicylic acid at the onset of systemic acquired resistance in cucumber.** *Science* 1990, **250**:1004-1006.
350. Schmelz EA, Engelberth J, Alborn HT, O'Donnell P, Sammons M, Toshima H, Tumlinson JHI: **Simultaneous analysis of phytohormones, phytotoxins, and volatile organic compounds in plants.** *Proc Natl Acad Sci U S A* 2003, **100**:10552-10557.
351. Sassenrath GF, Ort DR: **The relationship between inhibition of photosynthesis at low temperature and inhibition of photosynthesis after rewarming in chilling-sensitive tomato.** *Plant Physiol Biochem* 1990, **28**:457-465.
352. Clarke SM, Mur LA, Wood JE, Scott IM: **Salicylic acid dependent signalling promotes basal thermotolerance in *Arabidopsis thaliana*.** *Plant J* 2004, **38**:432-447.
353. Dat JF, Foyer CH, Scott IM: **Changes in salicylic acid and antioxidants during induced thermotolerance in mustard seedlings.** *1998* 1998, **118**:1455-1461.
354. Senaratna T, Touchell D, Bunn E, Dixon K: **Acetyl salicylic acid (aspirin) and salicylic acid induce multiple stress tolerance in bean and tomato plants.** *Plant Growth Regulation* 2000, **2000**:30.
355. Feys BJ, Parker JE: **Interplay of signalling pathways in plant disease resistance.** *Trends Genet* 2000.
356. Ton J, Van Pelt JA, Van Loon LC, Pieterse CM: **Differential effectiveness of salicylate-dependent and jasmonate/ethylene-dependent induced resistance in *Arabidopsis*.** *Mol Plant Microbe Interact* 2002, **15**:27-34.
357. Nurnberger T, Scheel D: **Signal transmission in the plant immune response.** *Trends Plant Sci* 2001, **6**:372-379.
358. Zolg JW, Langen H: **How industrie is approaching the search for new diagnostic markers and biomarkers.** *Mol Cell Proteomics* 2004, **3**:345-354.
359. Strocchi M, Ferrer M, Timmis KN, Golyshin PN: **Low temperature-induced systems failure in *Escherichia coli*: Insights from rescue by cold-adapted chaperones.** *Proteomics* 2006, **6**:193-206.
360. Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M: **MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes.** *Plant J* 2004, **37**:914-939.
361. Glinski M, Weckwerth W: **Differential Multisite Phosphorylation of the Trehalose-6-phosphate Synthase Gene Family in *Arabidopsis thaliana*.** *Molecular and Cellular Proteomics* 2005, **4**:1614-1625.
362. Arkin A, Ross J, McAdams HH: **Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells.** *Genetics* 1998:1633-1648.
363. Arkin A, Shen PD, Ross J: **A test case of correlation metric construction of a reaction pathway from measurements.** *Science* 1997, **277**:1275-1279.

364. Steuer R, Kurths J, Fiehn O, Weckwerth W: **Interpreting correlations in metabolomic networks.** *Biochem Soc Trans* 2003, **31**:1476-1478.
365. Weckwerth W, Fiehn O: **Can we discover novel pathways using metabolomic analysis?** *Curr Opin Biotechnol* 2002, **13**:156-160.
366. Weckwerth W, Morgenthal K: **Metabolomics: from pattern recognition to biological interpretation.** *Drug Discov Today* 2005, **10**:1551-1558.
367. Bradford MM: **Rapid and Sensitive Method for Quantitation of Microgram Quantities of Protein Utilizing Principle of Protein-Dye Binding.** *Analytical Biochemistry* 1976, **72**:248-254.
368. Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP: **Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome.** *J Proteome Res* 2003, **2**:43-50.
369. Tabb DL, McDonald WH, Yates JR: **DTASelect and contrast: Tools for assembling and comparing protein identifications from shotgun proteomics.** *Journal of Proteome Research* 2002, **1**:21-26.
370. Hyvärinen A, Oja E: **Independent component analysis: Algorithms and applications.** *Neural Networks* 2000, **4-5**:411-430.

## 10 Anhang

<b>Anhang A-1:</b>	<b><i>Arabidopsis thaliana</i> Entwicklungsstadien.....</b>	<b>179</b>
<b>Anhang A-2:</b>	<b>AGI-Code und funktionelle Beschreibung der im Rahmen der Analyse der abiotischen Temperaturstress-Response in <i>A. thaliana</i> Col-0 und PGM detektierten Proteine .....</b>	<b>180</b>
<b>Anhang A-3:</b>	<b>Einfluss der Datentransformation auf die ICA am Beispiel des integrativen Protein-Metabolit-Datensatzes der temperaturbedingten abiotischen Stressresponse von <i>A. thaliana</i> Col-0 und PGM .....</b>	<b>190</b>

Anhang A-1: Entwicklungsstadien von *A. thaliana* Col-0 (nach [177]).

**Abbildung A-1.1: Phänotypische Entwicklungsstadien von *A. thaliana* Col-0.** Dargestellt sind die verschiedenen Stadien der Samenkeimung (**Abb. A-1.1 A bis C**), der Blattentwicklung (**Abb. A-1.1 D bis G**), sowie einzelne Stadien der Blütenentwicklung und Samenreife (**Abb. A-1.1 H bis L**) für die Anzucht auf Erde.

Für weiterführende Charakteristika der einzelnen Stadien siehe Tabelle A-1.1.

**Tabelle A-1.1: *Arabidopsis* Entwicklungsstadien und Charakteristika für die Anzucht auf Erde**

Stadium - Charakteristik	Stadium - Charakteristik	Stadium - Charakteristik
<i>Samenkeimung</i>	<i>Blattentwicklung</i>	<i>Blütenbildung</i>
<b>0.1</b> Samenquellung	<b>1.0</b> vollständig geöffnete Keimblätter	<b>5.10</b> beginnende Knospung
<b>0.5</b> Entstehung der Keimwurzel	<b>1.02</b> zwei Rosettenblätter (Länge >1 mm)	<b>6.00</b> erste Blüte geöffnet
<b>0.7</b> Bildung von Keimblatt und Wurzel	<b>1.04</b> vier Rosettenblätter (Länge >1 mm)	<b>6.50</b> mittleres Blühstadium
	<b>1.10</b> zehn Rosettenblätter (Länge >1 mm)	<b>6.90</b> spätes Blühstadium
		<b>9.70</b> Senescenz

## Anhang A-2/

**Tabelle A-2.1 : AGI-Code und funktionelle Beschreibung der im Rahmen der Analyse der abiotischen Temperaturstress-Response in *A. thaliana* detektierten Proteine.** Die gemäss Kruskal-Wallis-Test nach Temperaturstress (4°C und/oder 32°C) mit  $p \leq 0.05$  im Vergleich zu den bei 20°C kultivierten Kontrollpflanzen der Genotypen Col-0 und PGM signifikante Unterschiede in ihrer Expression aufweisenden Proteine sind Fett geschrieben. Proteine, welche unter bestimmten experimentellen Bedingungen nicht detektiert werden konnten, sind grau hinterlegt. Blau hinterlegte Zeilen kennzeichnen Proteine mit nachweisbar signifikanten Unterschieden (Kruskal-Wallis-Test; Irrtumswahrscheinlichkeit  $p \leq 0.05$ ) zwischen den 6 untersuchten Gruppen (Col-0 4°C, 20°C, 32°C und PGM 4°C, 20°C, 32°C).

AGI-Code	Funktionelle Beschreibung
At1g02500.1	S-adenosylmethionine synthetase 1 (SAM1) identical to S-adenosylmethionine synthetase 1 (Methionine adenosyltransferase 1, AdoMet synthetase 1) [Arabidopsis thaliana] SWISS-PROT:P23686
At1g03130.1	photosystem I reaction center subunit II, chloroplast, putative / photosystem I 20 kDa subunit, putative / PSI-D, putative (PSAD2) similar to SP P12353 Photosystem I reaction center subunit II, chloroplast precursor (Photosystem I 20 kDa subunit) (PSI-D) [Spinacia oleracea]; contains Pfam profile PF02531: PsaD
At1g03600.1	photosystem II family protein similar to SP:P74367 [Synechocystis sp.]; similar to ESTs emb Z27038, gb AA451546, emb Z29876, gb T45359 and gb R90316
At1g03680.1	thioredoxin M-type 1, chloroplast (TRX-M1) nearly identical to SP O48737 Thioredoxin M-type 1, chloroplast precursor (TRX-M1) [Arabidopsis thaliana]; similar to ESTs gb T13714, gb H76398, gb N37762, gb AA042639, gb T21104, emb Z30901
At1g04410.1	<b>malate dehydrogenase, cytosolic, putative strong similarity to malate dehydrogenase from Mesembryanthemum crystallinum [SP O24047], Medicago sativa [SP O48905], Prunus persica [GI:15982948]; contains InterPro entry IPR001236: Lactate/malate dehydrogenase</b>
At1g04480.1	60S ribosomal protein L23 (RPL23A) identical to GB:AAB80655
At1g05190.1	ribosomal protein L6 family protein Similar to Mycobacterium RlpF (gb Z84395). ESTs gb T75785,gb R30580,gb T04698 come from this gene
At1g06430.1	FtsH protease, putative similar to zinc dependent protease GI:7650138 from [Arabidopsis thaliana]
At1g06680.1	<b>photosystem II oxygen-evolving complex 23 (OEC23) JBC 14:211-238 (2002); identical to 23 kDa polypeptide of oxygen-evolving complex (OEC) GB:CAA66785 GI:1769905 [Arabidopsis thaliana]</b>
At1g07890.1	L-ascorbate peroxidase 1, cytosolic (APX1) identical to SP Q05431 L-ascorbate peroxidase, cytosolic (EC 1.11.1.11) (AP) [Arabidopsis thaliana], L-ascorbate peroxidase [Arabidopsis thaliana] gi 16173 emb CAA42168; strong similarity to cytosolic ascorbate peroxidase [Spinacia oleracea] gi 1384110 dbj BAA12890
At1g07920.1	elongation factor 1-alpha / EF-1-alpha identical to GB:CAA34456 from [Arabidopsis thaliana] (Plant Mol. Biol. 14 (1), 107-110 (1990))
At1g08830.1	<b>superoxide dismutase [Cu-Zn] (SODCC) / copper/zinc superoxide dismutase (CSD1) identical to SWISS-PROT: P24704</b>
At1g08880.1	histone H2A, putative Strong similarity to histone H2A Cicer arietinum SP O65759, Picea abies SP P35063; contains Pfam profile PF00125 Core histone H2A/H2B/H3/H4; ESTs gb ATTS3874,gb T46627,gb T14194 come from this gene
At1g09340.1	expressed protein
At1g11430.1	plastid developmental protein DAG, putative similar to DAG protein, chloroplast precursor [Garden snapdragon] SWISS-PROT:Q38732
At1g11860.1	aminomethyltransferase, putative similar to aminomethyltransferase, mitochondrial precursor SP:O49849 from [Flaveria anomala]

At1g12900.1	glyceraldehyde 3-phosphate dehydrogenase, chloroplast, putative / NADP-dependent glyceraldehydephosphate dehydrogenase, putative similar to SP P25856 Glyceraldehyde 3-phosphate dehydrogenase A, chloroplast precursor (EC 1.2.1.13) (NADP-dependent glyceraldehydephosphate dehydrogenase subunit A) [Arabidopsis thaliana]; contains Pfam profiles PF02800: Glyceraldehyde 3-phosphate dehydrogenase C-terminal domain, PF00044: Glyceraldehyde 3-phosphate dehydrogenase NAD binding domain
At1g13440.1	glyceraldehyde 3-phosphate dehydrogenase, cytosolic, putative / NAD-dependent glyceraldehyde-3-phosphate dehydrogenase, putative very strong similarity to SP P25858 Glyceraldehyde 3-phosphate dehydrogenase, cytosolic (EC 1.2.1.12) [Arabidopsis thaliana]; contains Pfam profiles PF02800: Glyceraldehyde 3-phosphate dehydrogenase C-terminal domain, PF00044: Glyceraldehyde 3-phosphate dehydrogenase NAD binding domain
<b>At1g13930.1</b>	<b>expressed protein weakly similar to drought-induced protein SDi-6 (PIR:S71562) common sunflower (fragment)</b>
At1g14810.1	semialdehyde dehydrogenase family protein similar to SP:O31219 Aspartate-semialdehyde dehydrogenase (EC 1.2.1.11) (ASA dehydrogenase) (ASADH) [Legionella pneumophila]; contains Pfam profiles PF02774: Semialdehyde dehydrogenase dimerisation domain, PF01118: Semialdehyde dehydrogenase NAD binding domain
At1g17880.1	nascent polypeptide-associated complex (NAC) domain-containing protein / BTF3b-like transcription factor, putative similar to SP P20290 Transcription factor BTF3 (RNA polymerase B transcription factor 3) [Homo sapiens]; contains Pfam profile PF01849: NAC domain; identical to cDNA BTF3b-like factor GI:5912423
At1g19550.1	dehydroascorbate reductase, putative similar to dehydroascorbate reductase [Arabidopsis thaliana] gi 10952514 gb AAG24946
<b>At1g19570.1</b>	<b>dehydroascorbate reductase, putative similar to GB:BAA90672 from (Oryza sativa)</b>
At1g20260.1	vacuolar ATP synthase subunit B, putative / V-ATPase B subunit, putative / vacuolar proton pump B subunit, putative / V-ATPase 57 kDa subunit, putative strong similarity to SP P11574 Vacuolar ATP synthase subunit B (EC 3.6.3.14) (V-ATPase B subunit) (Vacuolar proton pump B subunit) (V-ATPase 57 kDa subunit) [Arabidopsis thaliana]; contains Pfam profiles PF00006: ATP synthase alpha/beta family nucleotide-binding domain, PF02874: ATP synthase alpha/beta family beta-barrel domain
At1g20340.1	plastocyanin similar to plastocyanin GI:1865683 from [Arabidopsis thaliana]
At1g20620.1	catalase 3 (SEN2) almost identical to catalase 3 SP:Q42547, GI:3123188 from [Arabidopsis thaliana]; identical to catalase 3 (SEN2) mRNA, partial cds GI:3158369
At1g21750.1	protein disulfide isomerase, putative similar to SP P29828 Protein disulfide isomerase precursor (PDI) (EC 5.3.4.1) [Medicago sativa]; isoform contains non-consensus GA donor splice site at intron 9
At1g23410.1	ubiquitin extension protein, putative / 40S ribosomal protein S27A (RPS27aA) strong similarity to ubiquitin extension protein (UBQ5) GB:AAA32906 GI:166934 from (Arabidopsis thaliana)
At1g23740.1	oxidoreductase, zinc-binding dehydrogenase family protein contains Pfam profile PF00107: oxidoreductase, zinc-binding dehydrogenase family
At1g24020.1	Bet v I allergen family protein similar to major pollen allergen Bet v 1 GB:CAA96544 GI:1321726 from [Betula pendula]; contains Pfam profile PF00407: Pathogenesis-related protein Bet v I family
At1g26630.1	eukaryotic translation initiation factor 5A, putative / eIF-5A, putative strong similarity to SP Q9AXQ6 Eukaryotic translation initiation factor 5A-1 (eIF-5A 1) [Lycopersicon esculentum]
At1g27400.1	60S ribosomal protein L17 (RPL17A) similar to GB:P51413 from [Arabidopsis thaliana]; similar to ESTs gb L33542 and gb AA660016
At1g31812.1	acyl-CoA binding protein / ACBP identical to acyl-CoA-binding protein (ACBP) [Arabidopsis thaliana] SWISS-PROT:P57752

At1g32470.1	glycine cleavage system H protein, mitochondrial, putative similar to SP P25855 Glycine cleavage system H protein 1, mitochondrial precursor [Arabidopsis thaliana]; contains Pfam profile PF01597: Glycine cleavage H-protein
At1g32990.1	ribosomal protein L11 family protein similar to chloroplast ribosomal protein L11 GI:21312 from [Spinacia oleracea]
At1g42970.1	glyceraldehyde-3-phosphate dehydrogenase B, chloroplast (GAPB) / NADP-dependent glyceraldehydephosphate dehydrogenase subunit B identical to SP P25857 Glyceraldehyde 3-phosphate dehydrogenase B, chloroplast precursor (EC 1.2.1.13) (NADP-dependent glyceraldehydephosphate dehydrogenase subunit B) [Arabidopsis thaliana]
At1g47128.1	cysteine proteinase (RD21A) / thiol protease identical to SP P43297 Cysteine proteinase RD21A precursor (EC 3.4.22.-) [Arabidopsis thaliana], thiol protease RD21A SP:P43297 from [Arabidopsis thaliana]
At1g48030.1	dihydrolipoamide dehydrogenase 1, mitochondrial / lipoamide dehydrogenase 1 (MTLPD1) identical to GB:AAF34795 [gi:12704696] from [Arabidopsis thaliana]
At1g48350.1	ribosomal protein L18 family protein similar to ribosomal protein L18 GI:3980238 from [Thermotoga maritima]
At1g49760.1	polyadenylate-binding protein, putative / PABP, putative similar to poly(A)-binding protein GB:AAF66825 GI:7673359 from [Nicotiana tabacum]
At1g52230.1	photosystem I reaction center subunit VI, chloroplast, putative / PSI-H, putative (PSAH2) identical to SP Q9SUI6; similar to PSI-H precursor [Nicotiana sylvestris] GI:407355; contains Pfam profile PF03244: Photosystem I reaction centre subunit VI
At1g53240.1	malate dehydrogenase [NAD], mitochondrial identical to mitochondrial NAD-dependent malate dehydrogenase GI:3929649 SP Q9ZP06 from [Arabidopsis thaliana]; contains InterPro entry IPR001236: Lactate/malate dehydrogenase
At1g54630.1	acyl carrier protein 3, chloroplast (ACP-3) nearly identical to SP P25702 Acyl carrier protein 3, chloroplast precursor (ACP) [Arabidopsis thaliana]
At1g55480.1	expressed protein
At1g55490.1	RuBisCO subunit binding-protein beta subunit, chloroplast / 60 kDa chaperonin beta subunit / CPN-60 beta identical to SWISS-PROT:P21240- RuBisCO subunit binding-protein beta subunit, chloroplast precursor (60 kDa chaperonin beta subunit, CPN-60 beta) [Arabidopsis thaliana]
At1g56190.1	phosphoglycerate kinase, putative similar to SP P41758 Phosphoglycerate kinase, chloroplast precursor (EC 2.7.2.3) [Chlamydomonas reinhardtii]; contains Pfam profile PF00162: phosphoglycerate kinase
At1g64200.1	vacuolar ATP synthase subunit E, putative / V-ATPase E subunit, putative / vacuolar proton pump E subunit, putative similar to SP Q39258 Vacuolar ATP synthase subunit E (EC 3.6.3.14) (V-ATPase E subunit) (Vacuolar proton pump E subunit) [Arabidopsis thaliana]; contains Pfam profile PF01991: ATP synthase (E/31 kDa) subunit
<b>At1g65980.1</b>	<b>peroxiredoxin type 2, putative strong similarity to type 2 peroxiredoxin [Brassica rapa subsp. pekinensis] GI:4928472; contains Pfam profile: PF00578 AhpC/TSA (alkyl hydroperoxide reductase and thiol-specific antioxidant) family</b>
At1g67090.1	ribulose biphosphate carboxylase small chain 1A / RuBisCO small subunit 1A (RBCS-1A) (ATS1A) identical to SP P10795 Ribulose biphosphate carboxylase small chain 1A, chloroplast precursor (EC 4.1.1.39) (RuBisCO small subunit 1A) [Arabidopsis thaliana]
At1g73230.1	nascent polypeptide-associated complex (NAC) domain-containing protein similar to SP P20290 Transcription factor BTF3 (RNA polymerase B transcription factor 3) [Homo sapiens]; contains Pfam profile PF01849: NAC domain
At1g74470.1	geranylgeranyl reductase identical to geranylgeranyl reductase GB:Y14044 [Arabidopsis thaliana] (involvement: chlorophyll, the tocopherol and the phylloquinone pathways Eur J Biochem 1998 Jan 15;251(1-2):413-7)
At1g75350.1	ribosomal protein L31 family protein similar to SP:O46917 from [Guillardia theta]
At1g76080.1	thioredoxin family protein low similarity to thioredoxin (TRX) [Fasciola hepatica] GI:6687568; contains Pfam profile PF00085: Thioredoxin

At1g78630.1	<b>ribosomal protein L13 family protein similar to ribosomal protein L13 GI:170132 from [Spinacia oleracea]</b>
At1g79040.1	photosystem II 10 kDa polypeptide identical to photosystem II 10 kDa polypeptide, chloroplast [precursor] SP:P27202 from [Arabidopsis thaliana]; contains Pfam profile: PF04725 photosystem II 10 kDa polypeptide PsbR
At2g03440.1	<b>nodulin-related similar to Early nodulin 12B precursor (N-12B) (Swiss-Prot:Q40339) [Medicago sativa]</b>
At2g16600.1	<b>peptidyl-prolyl cis-trans isomerase, cytosolic / cyclophilin / rotamase (ROC3) identical to cytosolic cyclophilin [Arabidopsis thaliana] GI:1305455</b>
At2g19760.1	profilin 1 (PRO1) (PFN1) (PRF1) / allergen Ara t 8 identical to profilin 1 (Allergen Ara t 8) SP:Q42449 GI:1353770 from [Arabidopsis thaliana]
At2g20260.1	photosystem I reaction center subunit IV, chloroplast, putative / PSI-E, putative (PSAE2) identical to SP Q9S714; similar to SP P12354 Photosystem I reaction center subunit IV, chloroplast precursor (PSI-E) [Spinacia oleracea]; contains Pfam profile PF02427: Photosystem I reaction centre subunit IV / PsaE
At2g21130.1	<b>peptidyl-prolyl cis-trans isomerase / cyclophilin (CYP2) / rotamase identical to cyclophilin [Arabidopsis thaliana] gi 2443757 gb AAB71402; similar to peptidyl-prolyl cis-trans isomerase, PPIase (cyclophilin, cyclosporin A-binding protein) [Arabidopsis thaliana] SWISS-PROT:P34790</b>
At2g21170.1	triosephosphate isomerase, chloroplast, putative similar to Triosephosphate isomerase, chloroplast precursor: SP P48496 from Spinacia oleracea, SP P46225 from Secale cereale
At2g21640.1	expressed protein
At2g21660.1	<b>glycine-rich RNA-binding protein (GRP7) SP Q03250 Glycine-rich RNA-binding protein 7 [Arabidopsis thaliana]</b>
At2g21870.1	expressed protein
At2g23120.1	expressed protein
At2g24020.1	expressed protein contains Pfam domain PF02575: Uncharacterized BCR, YbaB family COG0718
At2g27710.1	60S acidic ribosomal protein P2 (RPP2B)
At2g28000.1	RuBisCO subunit binding-protein alpha subunit, chloroplast / 60 kDa chaperonin alpha subunit / CPN-60 alpha identical to SWISS-PROT:P21238- RuBisCO subunit binding-protein alpha subunit, chloroplast precursor (60 kDa chaperonin alpha subunit, CPN-60 alpha) [Arabidopsis thaliana]
At2g28190.1	<b>superoxide dismutase [Cu-Zn], chloroplast (SODCP) / copper/zinc superoxide dismutase (CSD2) identical to GP:3273753:AF061519</b>
At2g28900.1	mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein contains Pfam PF02466: Mitochondrial import inner membrane translocase subunit Tim17
At2g30860.1	glutathione S-transferase, putative identical to GB:Y12295
At2g33150.1	acetyl-CoA C-acyltransferase, putative / 3-ketoacyl-CoA thiolase, putative similar to 3-ketoacyl-CoA thiolase (E.C. 2.3.1.16) from [Arabidopsis thaliana] GI:2981616, [Cucumis sativus] GI:393707, [Cucurbita cv. Kurokawa Amakuri] GI:1694621; contains InterPro accession IPR002155: Thiolase
At2g33800.1	ribosomal protein S5 family protein contains Pfam profiles PF03719: Ribosomal protein S5, C-terminal domain, PF00333: Ribosomal protein S5, N-terminal domain
At2g35370.1	<b>glycine cleavage system H protein 1, mitochondrial (GDCSH) (GCDH) identical to SP P25855 Glycine cleavage system H protein 1, mitochondrial precursor [Arabidopsis thaliana]</b>
At2g36460.1	fructose-bisphosphate aldolase, putative similar to PIR S65073 fructose-bisphosphate aldolase (EC 4.1.2.13) isoenzyme C-1, cytosolic [Oryza sativa]; contains Pfam profile PF00274 Fructose-bisphosphate aldolase class-I
At2g36530.1	enolase identical to SWISS-PROT:P25696 enolase (EC 4.2.1.11) (2-phosphoglycerate dehydratase)(2-phospho-D- glycerate hydro-lyase) [Arabidopsis thaliana]
At2g37190.1	ribosomal protein L12 (RPL12A)

At2g37220.1	<b>29 kDa ribonucleoprotein, chloroplast, putative / RNA-binding protein cp29, putative similar to SP Q43349 29 kDa ribonucleoprotein, chloroplast precursor (RNA-binding protein cp29) [Arabidopsis thaliana]</b>
At2g37660.1	expressed protein
At2g39460.1	60S ribosomal protein L23A (RPL23aA) identical to GB:AF034694
At2g39730.1	ribulose biphosphate carboxylase/oxygenase activase / RuBisCO activase identical to SWISS-PROT:P10896 ribulose biphosphate carboxylase/oxygenase activase, chloroplast precursor (RuBisCO activase, RA)[Arabidopsis thaliana]
At2g42530.1	cold-responsive protein / cold-regulated protein (cor15b) nearly identical to cold-regulated gene cor15b [Arabidopsis thaliana] GI:456016; contains Pfam profile PF02987: Late embryogenesis abundant protein
At2g43560.1	immunophilin / FKBP-type peptidyl-prolyl cis-trans isomerase family protein identical to probable FKBP-type peptidyl-prolyl cis-trans isomerase 2, chloroplast precursor (Ppiase) (Rotamase) (SP:022870)[Arabidopsis thaliana], contains Pfam PF00254: peptidyl-prolyl cis-trans isomerase, FKBP-type
At2g44650.1	chloroplast chaperonin 10 (cpn10) identical to chloroplast chaperonin 10 GI:14041813 from [Arabidopsis thaliana]
At2g44920.1	thylakoid luminal 15 kDa protein, chloroplast identical to SP O22160 Thylakoid luminal 15 kDa protein, chloroplast precursor (p15) [Arabidopsis thaliana]; contains 8 pentapeptide repeats
At2g47400.1	CP12 domain-containing protein contains Pfam profile: PF02672 CP12 domain
At3g01280.1	porin, putative similar to SP P42055 34 kDa outer mitochondrial membrane protein porin (Voltage-dependent anion-selective channel protein) (VDAC) [Solanum tuberosum]; contains Pfam profile PF01459: Eukaryotic porin
At3g01390.1	vacuolar ATP synthase subunit G 1 (VATG1) / V-ATPase G subunit 1 (VAG1) / vacuolar proton pump G subunit 1 (VMA10) identical to SWISS-PROT:O82628 vacuolar ATP synthase subunit G 1 (V-ATPase G subunit 1, Vacuolar proton pump G subunit 1) [Arabidopsis thaliana]
At3g01480.1	peptidyl-prolyl cis-trans isomerase, putative / cyclophilin, putative / rotamase, putative similar to peptidyl-prolyl cis-trans isomerase, chloroplast precursor (40 kDa thylakoid lumen PPIase, 40 kDa thylakoid lumen rotamase) [Spinacia oleracea] SWISS-PROT:O49939
At3g01500.1	carbonic anhydrase 1, chloroplast / carbonate dehydratase 1 (CA1) nearly identical to SP P27140 Carbonic anhydrase, chloroplast precursor (EC 4.2.1.1) (Carbonate dehydratase) [Arabidopsis thaliana]
At3g04790.1	ribose 5-phosphate isomerase-related similar to ribose-5-phosphate isomerase GI:18654317 from [Spinacia oleracea]
At3g06610.1	DNA-binding enhancer protein-related similar to huntingtin interacting protein HYPK (GI:3329429) [Homo sapiens]; identical to Egd2p (GI:172043) [Saccharomyces cerevisiae] similar to EGD2 protein (GAL4 DNA-binding enhancer protein 2) (Swiss-Prot:P38879) [Saccharomyces cerevisiae]
At3g09440.1	heat shock cognate 70 kDa protein 3 (HSC70-3) (HSP70-3) identical to SP O65719 Heat shock cognate 70 kDa protein 3 (Hsc70.3) [Arabidopsis thaliana]
At3g12780.1	phosphoglycerate kinase, putative similar to SP P41758 Phosphoglycerate kinase, chloroplast precursor (EC 2.7.2.3) [Chlamydomonas reinhardtii]; contains Pfam profile PF00162: phosphoglycerate kinase
At3g14210.1	myrosinase-associated protein, putative similar to GB:CAA71238 from [Brassica napus]; contains Pfam profile:PF00657 Lipase/Acylhydrolase with GDSL-like motif
At3g20390.1	endoribonuclease L-PSP family protein contains Pfam domain PF01042: Endoribonuclease L-PSP
At3g25920.1	50S ribosomal protein L15, chloroplast (CL15) identical to GB:P25873 from [Arabidopsis thaliana]
At3g26060.1	peroxiredoxin Q, putative similar to peroxiredoxin Q [Sedum lineare] GI:6899842; contains Pfam profile: PF00578 AhpC/TSA (alkyl hydroperoxide reductase and thiol-specific antioxidant) family

<b>At3g26450.1</b>	<b>major latex protein-related / MLP-related low similarity to major latex protein [Papaver somniferum][GI:294060] ; contains Pfam profile PF00407: Pathogenesis-related protein Bet v I family</b>
At3g26650.1	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast (GAPA) / NADP-dependent glyceraldehydephosphate dehydrogenase subunit A identical to SP P25856 Glyceraldehyde 3-phosphate dehydrogenase A, chloroplast precursor (EC 1.2.1.13) (NADP-dependent glyceraldehydephosphate dehydrogenase subunit A) [Arabidopsis thaliana]
At3g27830.1	50S ribosomal protein L12-1, chloroplast (CL12-A) identical to ribosomal protein L12 GB:X68046 [Arabidopsis thaliana] (J. Biol. Chem. 269 (10), 7330-7336 (1994))
At3g46780.1	expressed protein
At3g47070.1	expressed protein
At3g48870.1	ATP-dependent Clp protease ATP-binding subunit (ClpC) identical to AtClpC GI:5360574 from [Arabidopsis thaliana]; contains Pfam profiles PF02861: Clp amino terminal domain and PF02151: UvrB/uvrC motif
At3g49110.1	peroxidase 33 (PER33) (P33) (PRXCA) / neutral peroxidase C (PERC) identical to SP P24101 Peroxidase 33 precursor (EC 1.11.1.7) (Atperox P33) (ATPCa) (Neutral peroxidase C) (PERC) [Arabidopsis thaliana]
<b>At3g49120.1</b>	<b>peroxidase, putative identical to peroxidase [Arabidopsis thaliana] gi 405611 emb CAA50677</b>
At3g50820.1	oxygen-evolving enhancer protein, chloroplast, putative / 33 kDa subunit of oxygen evolving system of photosystem II, putative (PSBO2) identical to SP:Q9S841 Oxygen-evolving enhancer protein 1-2, chloroplast precursor (OEE1) [Arabidopsis thaliana]; strong similarity to SP P23321 Oxygen-evolving enhancer protein 1-1, chloroplast precursor (OEE1) (33 kDa subunit of oxygen evolving system of photosystem II) (OEC 33 kDa subunit) (33 kDa thylakoid membrane protein) [Arabidopsis thaliana]
At3g52100.1	PHD finger family protein contains Pfam profile PF00628: PHD-finger
At3g52150.1	RNA recognition motif (RRM)-containing protein similar to chloroplast RNA-binding protein cp33 [Arabidopsis thaliana] GI:681912; contains InterPro entry IPR000504: RNA-binding region RNP-1 (RNA recognition motif) (RRM) domain
<b>At3g52930.1</b>	<b>fructose-bisphosphate aldolase, putative similar to SP O65735 ALF_CICAR Fructose-bisphosphate aldolase, cytoplasmic isozyme [Cicer arietinum], cytosolic aldolase [Fragaria x ananassa] GI:10645188; contains Pfam profile PF00274 Fructose-bisphosphate aldolase class-I</b>
<b>At3g52960.1</b>	<b>peroxiredoxin type 2, putative similar to type 2 peroxiredoxin [Brassica rapa subsp. pekinensis] GI:4928472; contains Pfam profile: PF00578 AhpC/TSA (alkyl hydroperoxide reductase and thiol-specific antioxidant) family</b>
<b>At3g53460.1</b>	<b>29 kDa ribonucleoprotein, chloroplast / RNA-binding protein cp 29 nearly identical to SP Q43349 29 kDa ribonucleoprotein, chloroplast precursor (RNA-binding protein cp29) [Arabidopsis thaliana]</b>
At3g55440.1	triosephosphate isomerase, cytosolic, putative strong similarity to triosephosphate isomerase, cytosolic from Petunia hybrida [SP P48495], from Coptis japonica [SP P21820]
At3g57260.1	glycosyl hydrolase family 17 protein similar to glucan endo-1,3-beta-glucosidase, acidic isoform precursor SP:P33157 from [Arabidopsis thaliana]
At3g60750.1	transketolase, putative strong similarity to transketolase 1 [Capsicum annuum] GI:3559814; contains Pfam profiles PF02779: Transketolase, pyridine binding domain, PF02780: Transketolase, C-terminal domain, PF00456: Transketolase, thiamine diphosphate binding domain
At3g62030.1	peptidyl-prolyl cis-trans isomerase, chloroplast / cyclophilin / rotamase / cyclosporin A-binding protein (ROC4) identical to peptidyl-prolyl cis-trans isomerase, chloroplast precursor, PPIase (cyclophilin, cyclosporin A-binding protein) [Arabidopsis thaliana] SWISS-PROT:P34791; identical to cDNA nuclear-encoded chloroplast stromal cyclophilin (ROC4) GI:405130

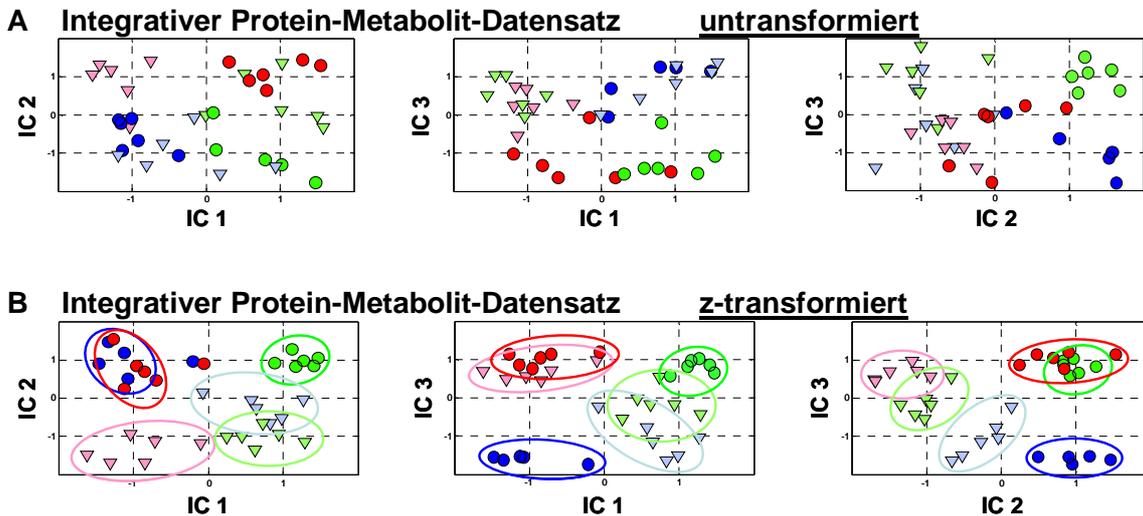
At3g63190.1	ribosome recycling factor, chloroplast, putative / ribosome releasing factor, chloroplast, putative similar to SP P82231 Ribosome recycling factor, chloroplast precursor (Ribosome releasing factor, chloroplast) (RRF) (CpFrr) (RRFHCP) [Spinacia oleracea]; contains Pfam profile PF01765: ribosome recycling factor
At3g63490.1	ribosomal protein L1 family protein ribosomal protein L1, S.oleracea, EMBL:SORPL1
At4g01310.1	ribosomal protein L5 family protein contains Pfam profiles PF00673: ribosomal L5P family C-terminus, PF00281: ribosomal protein L5
At4g02530.1	chloroplast thylakoid lumen protein SP:022773 ;TL16_ARATH
At4g03280.1	cytochrome B6-F complex iron-sulfur subunit, chloroplast / Rieske iron-sulfur protein / plastoquinol-plastocyanin reductase (petC) identical to gi:9843639; identical to cDNA rieske iron-sulfur protein precursor (petC) GI:5725449
At4g03520.1	thioredoxin M-type 2, chloroplast (TRX-M2) nearly identical to SP Q9SEU8 Thioredoxin M-type 2, chloroplast precursor (TRX-M2) [Arabidopsis thaliana]
At4g04020.1	plastid-lipid associated protein PAP, putative / fibrillin, putative strong similarity to plastid-lipid associated proteins PAP1 GI:14248554, PAP2 GI:14248556 from [Brassica rapa], fibrillin [Brassica napus] GI:4139097; contains Pfam profile PF04755: PAP_fibrillin
At4g05180.1	oxygen-evolving enhancer protein 3, chloroplast, putative (PSBQ2) identical to SP Q41932 Oxygen-evolving enhancer protein 3-2, chloroplast precursor (OEE3) (16 kDa subunit of oxygen evolving system of photosystem II) (OEC 16 kDa subunit) [Arabidopsis thaliana]; similar to SP P12301 Oxygen-evolving enhancer protein 3, chloroplast precursor (OEE3) (16 kDa subunit of oxygen evolving system of photosystem II) (OEC 16 kDa subunit) [Spinacia oleracea]; contains Pfam profile PF05757: Oxygen evolving enhancer protein 3 (PsbQ)
At4g08390.1	L-ascorbate peroxidase, stromal (sAPX) identical to stromal ascorbate peroxidase [Arabidopsis thaliana] gi 1419388 emb CAA67425
At4g13850.1	<b>glycine-rich RNA-binding protein (GRP2) glycine-rich RNA binding protein 2 AtGRP2 [Arabidopsis thaliana] GI:2826811</b>
At4g18480.1	<b>magnesium-chelatase subunit chlI, chloroplast / Mg-protoporphyrin IX chelatase (CHLI) (CS) (CH42) identical to SP P161127 Magnesium-chelatase subunit chlI, chloroplast precursor (Mg-protoporphyrin IX chelatase) (Protein CS/CH-42) [Arabidopsis thaliana]</b>
At4g20260.1	DREPP plasma membrane polypeptide family protein contains Pfam profile: PF05558 DREPP plasma membrane polypeptide
At4g20360.1	elongation factor Tu / EF-Tu (TUFA) identical to SWISS-PROT:P17745 elongation factor Tu, chloroplast precursor (EF-Tu) [Arabidopsis thaliana]
At4g21280.1	<b>oxygen-evolving enhancer protein 3, chloroplast, putative (PSBQ1) (PSBQ) identical to SP Q9XFT3 Oxygen-evolving enhancer protein 3-1, chloroplast precursor (OEE3) (16 kDa subunit of oxygen evolving system of photosystem II) (OEC 16 kDa subunit) [Arabidopsis thaliana]; similar to SP P12301 Oxygen-evolving enhancer protein 3, chloroplast precursor (OEE3) (16 kDa subunit of oxygen evolving system of photosystem II) (OEC 16 kDa subunit) [Spinacia oleracea]; contains Pfam profile PF05757: Oxygen evolving enhancer protein 3 (PsbQ)</b>
At4g23670.1	<b>major latex protein-related / MLP-related low similarity to major latex protein [Papaver somniferum][GI:294060] contains Pfam profile PF00407: Pathogenesis-related protein Bet v I family</b>
At4g24280.1	heat shock protein 70, putative / HSP70, putative strong similarity to heat shock protein 70 [Arabidopsis thaliana] GI:6746592; similar to heat shock 70 protein - Spinacia oleracea,PID:g2654208
At4g24770.1	31 kDa ribonucleoprotein, chloroplast, putative / RNA-binding protein RNP-T, putative / RNA-binding protein 1/2/3, putative / RNA-binding protein cp31, putative similar to SP Q04836 31 kDa ribonucleoprotein, chloroplast precursor (RNA-binding protein RNP-T) (RNA-binding protein 1/2/3) (AtRBP33) (RNA-binding protein cp31) [Arabidopsis thaliana]; contains InterPro entry IPR000504: RNA-binding region RNP-1 (RNA recognition motif) (RRM)

At4g25050.1	acyl carrier family protein / ACP family protein similar to Acyl carrier protein, chloroplast precursor from [Spinacia oleracea] SP P23235, [Casuarina glauca] SP P93092; contains InterPro accession IPR003881: Isochorismatase
At4g28750.1	photosystem I reaction center subunit IV, chloroplast, putative / PSI-E, putative (PSAE1) identical to SP Q9S831; similar to SP P12354 Photosystem I reaction center subunit IV, chloroplast precursor (PSI-E) [Spinacia oleracea]; contains Pfam profile PF02427: Photosystem I reaction centre subunit IV / PsaE
At4g29060.1	elongation factor Ts family protein similar to SP P35019 Elongation factor Ts (EF-Ts) [Galdieria sulphuraria]; contains Pfam profiles PF00627: UBA/TS-N domain, PF00889: Elongation factor TS, PF00575: S1 RNA binding domain
At4g29350.1	profilin 2 (PRO2) (PFN2) (PRF2) identical to profilin 2 SP:Q42418 GI:1353772 from [Arabidopsis thaliana]; identical to cDNA profilin (PRF2) GI:9965570
<b>At4g32260.1</b>	<b>ATP synthase family contains Pfam profile: PF00430 ATP synthase B/B' CF(0); identical to cDNA chloroplast ATP synthase beta chain precursor (atpG) GI:5730140</b>
At4g34620.1	ribosomal protein S16 family protein ribosomal protein S16, Neurospora crassa, PIR2:A29927
At4g34870.1	peptidyl-prolyl cis-trans isomerase / cyclophilin (CYP1) / rotamase identical to cyclophilin (CYP1) gi 992643 gb AAA75512; similar to peptidyl-prolyl cis-trans isomerase, PPIase (cyclophilin, cyclosporin A-binding protein) [Catharanthus roseus] SWISS-PROT:Q39613
At4g35450.1	ankyrin repeat family protein / AFT protein (AFT) contains ankyrin repeats, Pfam:PF00023; identical to cDNA AFT protein (AFT) GI:3478699
At4g37930.1	glycine hydroxymethyltransferase / serine hydroxymethyltransferase / serine/threonine aldolase (SHM1) identical to serine hydroxymethyl transferase [Arabidopsis thaliana] GI:6899945
At4g38680.1	cold-shock DNA-binding family protein contains Pfam domains PF00313: 'Cold-shock' DNA-binding domain and PF00098: Zinc knuckle
At4g38740.1	peptidyl-prolyl cis-trans isomerase / cyclophilin / rotamase / cyclosporin A-binding protein (ROC1) identical to SP P34790 Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (PPIase) (Rotamase) (Cyclophilin) (Cyclosporin A-binding protein) [Arabidopsis thaliana]
At4g38970.1	fructose-bisphosphate aldolase, putative strong similarity to plastidic fructose-bisphosphate aldolase (EC 4.1.2.13) from Nicotiana paniculata (NPALDP1) [GI:4827251], Oryza sativa, PIR2:T02057 [SP Q40677]
At4g39260.1	glycine-rich RNA-binding protein 8 (GRP8) (CCR1) SP Q03251 Glycine-rich RNA-binding protein 8 (CCR1 protein) (GRP8) [Arabidopsis thaliana] isoform contains a non-consensus CG acceptor splice site at intron 2
At5g02240.1	expressed protein
At5g02500.1	heat shock cognate 70 kDa protein 1 (HSC70-1) (HSP70-1) identical to SP P22953 Heat shock cognate 70 kDa protein 1 (Hsc70.1) [Arabidopsis thaliana]
<b>At5g08280.1</b>	<b>hydroxymethylbilane synthase / porphobilinogen deaminase, chloroplast / pre-uroporphyrinogen synthase identical to SP Q43316</b>
At5g08670.1	ATP synthase beta chain 1, mitochondrial identical to SP P83483 ATP synthase beta chain 1, mitochondrial precursor (EC 3.6.3.14) [Arabidopsis thaliana]; strong similarity to SP P17614 ATP synthase beta chain, mitochondrial precursor (EC 3.6.3.14) [Nicotiana plumbaginifolia]; contains Pfam profiles PF00006: ATP synthase alpha/beta family nucleotide-binding domain, PF00306: ATP synthase ab C terminal, PF02874: ATP synthase alpha/beta family beta-barrel domain; supporting cDNA gi 26452102 dbj AK118538.1
At5g09650.1	inorganic pyrophosphatase family protein similar to SP Q15181 Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate [Homo sapiens]; contains Pfam profile PF00719: inorganic pyrophosphatase
At5g12140.1	cysteine protease inhibitor, putative / cystatin, putative similar to SP P31726 Cystatin I precursor (CORN kernel cysteine proteinase inhibitor) [Zea mays]; contains Pfam profile PF00031: Cystatin domain

At5g14740.1	carbonic anhydrase 2 / carbonate dehydratase 2 (CA2) (CA18) nearly identical to SP P42737 Carbonic anhydrase 2 (EC 4.2.1.1) (Carbonate dehydratase 2) [Arabidopsis thaliana]
At5g14910.1	heavy-metal-associated domain-containing protein Pfam profile PF00403: Heavy-metal-associated domain
At5g15090.1	porin, putative / voltage-dependent anion-selective channel protein, putative similar to SP P42055 34 kDa outer mitochondrial membrane protein porin (Voltage-dependent anion-selective channel protein) (VDAC) [Solanum tuberosum]; contains Pfam profile PF01459: Eukaryotic porin; identical to cDNA voltage-dependent anion-selective channel protein GI:4006940
At5g15960.1	stress-responsive protein (KIN1) / stress-induced protein (KIN1) identical to SP P18612 Stress-induced KIN1 protein [Arabidopsis thaliana]
At5g15970.1	<b>stress-responsive protein (KIN2) / stress-induced protein (KIN2) / cold-responsive protein (COR6.6) / cold-regulated protein (COR6.6) identical to SP P31169 Stress-induced KIN2 protein (Cold-induced COR6.6 protein) [Arabidopsis thaliana]</b>
At5g16710.1	dehydroascorbate reductase, putative Strong similarity to dehydroascorbate reductase [Spinacia oleracea] gi:10952512 gb:AAG24945
At5g17170.1	rubredoxin family protein contains Pfam profile PF00301: Rubredoxin
At5g19510.1	elongation factor 1B alpha-subunit 2 (eEF1Balpha2) identical to elongation factor 1B alpha-subunit [Arabidopsis thaliana] GI:6686821
At5g20720.1	20 kDa chaperonin, chloroplast (CPN21) (CHCPN10) (CPN20) identical to chloroplast 20 kDa chaperonin, chloroplast precursor (Protein Cpn21), chloroplast protein Cpn10, chloroplast chaperonin 10 (Ch-CPN10), SP O65282 from [Arabidopsis thaliana]; identical to cDNA chaperonin 20 GI:14587372
At5g23060.1	expressed protein
At5g28540.1	luminal binding protein 1 (BiP-1) (BP1) SWISS-PROT:Q9LKR3 PMID:8888624
At5g35630.1	glutamine synthetase (GS2) identical to glutamine synthetase, chloroplast precursor (glutamate-- ammonia ligase, GS2) [Arabidopsis thaliana] SWISS-PROT:Q43127
At5g38410.1	ribulose biphosphate carboxylase small chain 3B / RuBisCO small subunit 3B (RBCS-3B) (ATS3B) identical to SP P10798 Ribulose biphosphate carboxylase small chain 3B, chloroplast precursor (EC 4.1.1.39) (RuBisCO small subunit 3B) [Arabidopsis thaliana]
At5g38430.1	ribulose biphosphate carboxylase small chain 1B / RuBisCO small subunit 1B (RBCS-1B) (ATS1B) identical to SP P10796 Ribulose biphosphate carboxylase small chain 1B, chloroplast precursor (EC 4.1.1.39) (RuBisCO small subunit 1B) [Arabidopsis thaliana]
At5g50920.1	ATP-dependent Clp protease ATP-binding subunit / ClpC almost identical to ClpC GI:2921158 from [Arabidopsis thaliana]; contains Pfam profile PF02861: Clp amino terminal domain; contains Pfam profile PF00004: ATPase, AAA family; contains Pfam profile PF02151: UvrB/uvrC motif
At5g52310.1	low-temperature-responsive protein 78 (LTI78) / desiccation-responsive protein 29A (RD29A)
At5g54600.1	50S ribosomal protein L24, chloroplast (CL24) identical to SP P92959 50S ribosomal protein L24, chloroplast precursor [Arabidopsis thaliana]
At5g54770.1	thiazole biosynthetic enzyme, chloroplast (ARA6) (THI1) (THI4) identical to SP Q38814 Thiazole biosynthetic enzyme, chloroplast precursor (ARA6) [Arabidopsis thaliana]
At5g61410.1	ribulose-phosphate 3-epimerase, chloroplast, putative / pentose-5-phosphate 3-epimerase, putative strong similarity to SP Q43157 Ribulose-phosphate 3-epimerase, chloroplast precursor (EC 5.1.3.1) (Pentose-5-phosphate 3-epimerase) (PPE) (RPE) (R5P3E) [Spinacia oleracea]; contains Pfam profile PF00834: Ribulose-phosphate 3 epimerase family
At5g63400.1	adenylate kinase identical to adenylate kinase (ATP-AMP transphosphorylase) [Arabidopsis thaliana] SWISS-PROT:O82514

At5g66570.1	oxygen-evolving enhancer protein 1-1, chloroplast / 33 kDa subunit of oxygen evolving system of photosystem II (PSBO1) (PSBO) identical to SP:P23321 Oxygen-evolving enhancer protein 1-1, chloroplast precursor (OEE1) (33 kDa subunit of oxygen evolving system of photosystem II) (OEC 33 kDa subunit) (33 kDa thylakoid membrane protein) [Arabidopsis thaliana]
ATPA_ARATH	ATP synthase alpha chain (EC 3.6.1.34) (P56757)
ATPB_ARATH	ATP synthase beta chain (P19366)
ATPE_ARATH	ATP synthase epsilon chain (EC 3.6.1.34) (P09468)
CYF_ARATH	Apocytochrome F precursor (P56771)
PSBC_ARATH	Photosystem II 44 kDa reaction center protein (P6 protein) (CP43) (P56778)
RBL_ARATH	Ribulose biphosphate carboxylase large chain precursor (EC 4.1.1.39) (RuBisCO large subunit) (O03042)
At1g02500.1	S-adenosylmethionine synthetase 1 (SAM1) identical to S-adenosylmethionine synthetase 1 (Methionine adenosyltransferase 1, AdoMet synthetase 1) [Arabidopsis thaliana] SWISS-PROT:P23686
At1g03130.1	photosystem I reaction center subunit II, chloroplast, putative / photosystem I 20 kDa subunit, putative / PSI-D, putative (PSAD2) similar to SPIP12353 Photosystem I reaction center subunit II, chloroplast precursor (Photosystem I 20 kDa subunit) (PSI-D) [Spinacia oleracea]; contains Pfam profile PF02531: PsaD

**Anhang A-3: Einfluss der Datentransformation auf die ICA am Beispiel des integrativen Protein-Metabolit-Datensatzes der temperaturbedingten abiotischen Stressresponse von *A. thaliana* Col-0 und PGM**



**Abbildung A-3.1: Einfluss der Datentransformation auf die ICA.** Abhängig von der Transformation des Originaldatensatzes resultiert aus der ICA eine unterschiedliche Gruppierung der Proben im neu berechneten Datenraum. **Abb. A-3.1 A:** Werden die Originaldaten nicht transformiert, ergibt sich keine sinnvolle Klassifizierung der Proben. **Abb. A-3.1 B:** Die Anwendung der ICA auf den Metabolit-weise z-transformierten Datensatz (Mittelwert=0, Standardabweichung=1) resultiert in einer schwer interpretierbaren Anordnung der Proben. Auch die Faktorladungen der ersten Unabhängigen Komponenten (**Tabelle A-3.1**) reflektieren weder den genetischen Hintergrund der PGM Mutante, noch die deutlichsten durch Temperaturstress hervorgerufenen Veränderungen im Protein- und Metabolit-haushalt der analysierten Genotypen.

● Col-0 4°C  
 ● Col-0 20°C  
 ● Col-0 32°C  
 ▼ PGM 4°C  
 ▼ PGM 20°C  
 ▼ PGM 32°C

**Tabelle A-3.1: Faktorladungen der ICA des integrativen Protein-Metabolit-Datensatzes nach z-Transformation**

Faktorladungen IC 1	Faktorladungen IC 2	Faktorladungen IC 3
-0.06 4-Aminobutyrat	0.05 Nicotinat	-0.02 Phenylalanin
-0.05 Fumarat	0.03 Sinapat	-0.02 Glutamin
-0.04 Isoleucin	0.03 Serin	-0.02 At2g21660.1
0.04 Glycerat	0.03 Erythronat/Threonat	-0.02 Pyroglutamat
-0.04 Erythritol/Threitol	-0.03 At1g19570.1	-0.02 Glycin
-0.04 At3g26450.1	-0.03 3-Methyl-3-hydroxy-glutarat	-0.02 Shikimat
-0.03 3-Methyl-3-hydroxy-glutarat	0.03 Ferulat	0.02 Glycerat
-0.03 Tyrosin	-0.03 Lysin	-0.02 Tyrosin
0.03 At2g37220.1	-0.03 At4g24280.1	-0.02 Fructose 6-phosphat
-0.03 At5g20720.1	-0.03 At4g39260.1	-0.02 Fructose
-0.03 At2g37660.1	0.03 At5g15970.1	-0.02 Prolin
-0.03 At4g24280.1	0.03 myo-Inositol	-0.02 Tryptophan
-0.03 Raffinose	0.03 Dehydroascorbat	-0.02 At5g15960.1
0.03 Trehalose	-0.03 At4g34870.1	-0.02 At2g39460.1
0.03 At2g16600.1	-0.03 At1g20620.1	-0.02 2-Ethyl-3-hydroxy-3-methyl-valerat