

MAX-PLANCK-INSTITUT FÜR MOLEKULARE GENETIK



**Hochdurchsatz Generierung und Analyse von
Arabidopsis thaliana-Proteinen**

Dissertation

zur Erlangung der Doktorwürde

des Fachbereichs Biologie, Chemie, Pharmazie

der Freien Universität Berlin

vorgelegt von Tanja Feilner

geb. am 07.07.1972
in Bayreuth

Berlin, September 2004

Erstgutachter:

Prof. Dr. Hans Lehrach

Zweitgutachter:

Prof. Dr. Thomas Schmülling

Tag der Disputation: 08.02.2005

1	Einleitung	1
1.1	Arabidopsis thaliana und die verwendeten Gewebe	1
1.1.1	Arabidopsis thaliana (Ackerschmalwand)	1
1.1.2	Die verwendeten Gewebe	2
1.2	Proteomics	3
1.3	Gewinnung von Proteinen im Hochdurchsatz	5
1.3.1	Strategien zur Hochdurchsatz-Klonierung	6
1.3.2	Proteinexpression im Hochdurchsatz	8
1.3.3	Proteinaufreinigung im Hochdurchsatz	9
1.4	Protein-Arrays und deren Anwendung	10
1.4.1	Grundlagen zur Array-Technologie	10
1.4.2	Oberflächen	12
1.4.3	Detektion	13
1.4.4	Analytische und funktionelle Protein-Microarrays	14
1.5	Phosphorylierung als eine posttranslationale Proteinmodifikation	14
1.5.1	Mitogen-aktivierte Proteinkinasen (MAPKs)	15
1.6	Zielsetzung	18
2	Material und Methoden	20
2.1	Material	20
2.1.1	Chemikalien	20
2.1.2	Proteine, Antikörper und Enzyme	21
2.1.3	Synthetische Oligonukleotide	22
2.1.4	Andere Materialien	22
2.1.5	Labor-Ausrüstung	23
2.1.6	Kits	24
2.1.7	Medien, Puffer und Lösungen	25
2.1.7.1	Medien	25
2.1.7.2	Kompetente Zellen	26
2.1.7.3	Polymerase-Kettenreaktion (PCR) und Agarosegele	26
2.1.7.4	Prozessierung der Koloniefilter	27
2.1.7.5	Antikörper-Bindungstests mit Filtern	27
2.1.7.6	SDS-Gelelektrophorese und Westernblot	28
2.1.7.7	Proteinreinigung unter denaturierenden Bedingungen	28
2.1.7.8	Proteinreinigung unter nativen Bedingungen	29
2.1.7.9	Phosphorylierungs-Assay	29
2.1.8	Software	29
2.1.9	Stämme	30
2.1.10	Plasmid-Konstrukte	31
2.1.10.1	pQE30NASTattB	32
2.1.10.2	pQE30NASTDV	33
2.1.10.3	pENTR1a	34
2.2	Methoden	35
2.2.1	Herstellung von zwei Arabidopsis cDNA-Bibliotheken	35
2.2.1.1	cDNA-Synthese und Größenfraktionierung	35
2.2.1.2	Fällung der cDNA	37
2.2.1.3	Vektorpräparationen	37

2.2.1.4	Ligation der cDNAs mit den Vektoren	38
2.2.1.5	Herstellung elektrokompetenter E. coli SCS1/pSE111-Zellen	39
2.2.1.6	Ermittlung der Transformationskompetenz der kompetenten Zellen	39
2.2.1.7	Transformation von E. coli mittels Elektroporation	40
2.2.1.8	Anordnung und Replikation einer cDNA-Bibliothek	40
2.2.2	Herstellung eines Proteinexpressionssubsets der cDNA-Expressionsbibliothek	42
2.2.2.1	Herstellen von Hochdichte-Proteinfiltern der cDNA-Expressionsbibliothek	42
2.2.2.2	Proteinexpression und Prozessierung der Filter	43
2.2.2.3	Immunofärbung und Auswertung der Proteinfilter	44
2.2.2.4	Neuordnung einer cDNA-Bibliothek zur Erstellung einer Unterbibliothek mit potentiellen Proteinexpressionsklonen	45
2.2.3	Herstellung eines Uniklonsets einer Arabidopsis cDNA-Expressionsbibliothek	45
2.2.3.1	Sequenzierung	45
2.2.3.2	Qualitätskontrolle der Sequenzen	45
2.2.3.3	BLAST-Analyse der Sequenzen	47
2.2.3.4	Cluster-Analyse der Sequenzen	48
2.2.3.5	Selektion der Klone für das Uniklonset	49
2.2.3.6	Anordnung des Uniklonsets	50
2.2.3.7	Erweiterung des Uniklonsets durch vollständige cDNA-Expressionsklone	50
2.2.4	Weitere nukleinsäuretechnische Methoden	52
2.2.4.1	Plasmidpräparation	52
2.2.4.2	Sequenzierung und Sequenzanalyse	53
2.2.4.3	PCR zur Amplifikation von DNA-Fragmenten	53
2.2.4.4	Gelelektrophoretische Auftrennung von DNA	55
2.2.4.5	Isolierung von DNA-Fragmenten aus Agarosegelen	55
2.2.5	Weitere proteinchemische Methoden	56
2.2.5.1	SDS-Polyacrylamidgel-Elektrophorese (SDS-PAGE)	56
2.2.5.2	Westernblot	58
2.2.5.3	Proteinexpression in E. coli in 96 well Format (1 ml Kultur)	58
2.2.5.4	Denaturierende Aufreinigung im 96 well Format von Proteinen mit RGS His ₆ -Tag	59
2.2.5.5	Bradford-Test	59
2.2.6	Spotten von aufgereinigten Proteinen auf beschichtete Glas-Microarrays (FAST TM -Slides)	60
2.2.6.1	Herstellung eines Test-Microarrays	60
2.2.6.2	Spotten der Proteine des Erweiterten Uniklonset	61
2.2.6.3	Antikörper-screening der gespotteten Protein-Microarrays	64
2.2.6.4	Kinase-Assays auf den Protein-Microarrays des Erweiterten Uniklonsets mit radioaktivem γ -ATP	65
2.2.6.5	Detektion und Quantifizierung der potentiellen Phosphorylierungs-Targets	67
3	Ergebnisse	71
3.1	Herstellung der cDNA-Bibliotheken	71
3.1.1	Verwendete Strategien	71
3.1.2	cDNA-Synthese	72
3.1.3	Vektorpräparation	73

3.1.4	Ligation, Transformation beider cDNA-Bibliotheken und Anordnung der Infloreszenz cDNA-Bibliothek	74
3.2	Charakterisierung der cDNA-Bibliotheken	75
3.2.1	Bestimmung der durchschnittlichen Insertlängen	75
3.2.1.1	Pistill cDNA-Bibliothek	75
3.2.1.2	Infloreszenz cDNA-Bibliothek	77
3.2.2	Sequenzanalysen beider cDNA-Bibliotheken	78
3.2.2.1	Testsequenzierung der Pistill cDNA-Bibliothek	78
3.2.2.2	Testsequenzierung der Infloreszenz cDNA-Bibliothek	78
3.3	Erstellung einer Proteinexpressions-Unterbibliothek der ATM1 cDNA-Bibliothek	80
3.3.1	Identifizierung potentieller Expressionsklone der ATM1 Bibliothek mit Hilfe von Hochdichte-Proteinfiltern	80
3.4	Erstellung des Uniklonsets der ATM1 cDNA-Bibliothek	82
3.4.1	Sequenzierung	82
3.4.2	Sequenzanalyse	84
3.4.2.1	BLAST-Analyse	84
3.4.2.2	Cluster-Analyse	85
3.4.3	Selektion der Klone für das Uniklonset	86
3.4.4	Erweiterung des Uniklonsets durch cDNA Expressionsklone, die in voller Länge vorlagen	88
3.4.5	Reinigung von 96 rekombinanten Proteinen des Uniklonsets und Erstellung eines Testmicroarrays	89
3.4.6	Bradford-Test	93
3.5	Spotten und Phosphorylierung des Erweiterten Uniklonsets	93
3.5.1	Detektion potentieller Phosphorylierungs-Targets	94
3.5.1.1	Vorversuche	94
3.5.1.2	Etablierung der Quantifizierung potentieller Targets im dichten Spotpattern	94
3.5.2	Vergleich potentieller Phosphorylierungs-Targets der verschiedenen Kinasen	109
4	Diskussion	113
4.1	Gewinnung rekombinanter Arabidopsis Proteine	113
4.1.1	Herstellung und Charakterisierung von zwei cDNA-Bibliotheken	113
4.1.2	Identifizierung potentieller Expressionsklone und Herstellung der Proteinexpressions-Unterbibliothek	117
4.1.3	Charakterisierung der Proteinexpressions-Unterbibliothek	118
4.1.4	Erstellung des Uniklonsets	120
4.1.5	Herstellung von Protein-Microarrays des Erweiterten Uniklonsets	122
4.2	Phosphorylierungsstudien als Funktionelle Analyse	124
4.2.1	Vergleich von Phosphorylierungsstudien	124
4.2.2	Phosphorylierungsstudien mit den Proteinen des Erweiterten Uniklonsets	127
4.2.2.1	Quantifizierung mittels Proteinkinase A	127
4.2.2.2	Phosphorylierungsstudien unter Verwendung der Arabidopsis Kinasen MPK3 und MPK6	129
5	Literaturverzeichnis	136
6	Abkürzungsverzeichnis	146

7	Zusammenfassung149
8	Summary151
9	Anhang153
10	Danksagung170

1 Einleitung

1.1 *Arabidopsis thaliana* und die verwendeten Gewebe

1.1.1 *Arabidopsis thaliana* (Ackerschmalwand)

Arabidopsis thaliana zählt zu den Wildkräutern und ist eine einjährige Pflanze. Sie gehört zu der Familie der Brassicaceae (Kreuzblütengewächse), ist ein Kohlgewächs und ein naher Verwandter von Senf, Raps und Blumenkohl. *Arabidopsis* ist heimisch in West-Eurasien, wurde durch den Menschen jedoch auch in Ostasien (Japan) und Amerika, sowie in geringerem Maße auch in Australien, Neuseeland und Chile verbreitet.

Arabidopsis thaliana hat sich aus vielerlei Gründen zu einem der wichtigsten Modellorganismen für klassische und molekulargenetische Studien entwickelt.

Diese Pflanze weist mit 10.000 – 30.000 Samen pro Pflanze eine hohe Reproduktionsrate auf, ist relativ klein (Höhe: 30 – 40 cm) und benötigt relativ wenig Platz (300 Pflanzen/m²). Zudem lässt sie sich mit einer Generationszeit von etwa sechs bis acht Wochen schnell kultivieren.

Darüber hinaus wurde im Dezember 2000 die vollständige Genomsequenz von *Arabidopsis thaliana* in „Nature“ veröffentlicht (*Arabidopsis*-Genome-Initiative, 2000). Das Genom ist gegenüber anderen Pflanzen mit ungefähr 130 Mbp relativ klein und besteht aus schätzungsweise 28.000 Genen, welche sich auf 5 Chromosomen verteilen. Die Genomgröße bei Pflanzen weist eine hohe Varianz auf: Raps hat eine Genomgröße von 1,2 Gbp, Mais von 3 Gbp, Gerste von 5 Gbp und Weizen sogar eine Genomgröße von 16 Gbp. Obwohl die Genomgröße von *Arabidopsis* vergleichbar ist mit der von *Drosophila melanogaster* und *Caenorhabditis elegans*, ist bei dieser Pflanze eine größere Anzahl von Genen vorzufinden. Das Genom von *C. elegans* mit ungefähr 100 Mbp kodiert für etwa 19.000 Gene und das von *Drosophila* mit ca. 120 Mbp für schätzungsweise 13.600 Gene (Adams *et al.*, 2000; *C.elegans*-Sequencing-Consortium, 1998). Bei den beiden weiteren bekannten Genomen von Reis konnte eine Genomgröße von ungefähr 430 Mbp festgestellt werden, die für ca. 50.000 Gene kodieren (Goff *et al.*, 2002; Yu *et al.*, 2002).

1.1.2 Die verwendeten Gewebe

Für die vorliegende Arbeit wurden zwei Gewebe, Pistill (Blütenstempel) und Infloreszenz (Blütenstand), ausgewählt, von denen erwartet wurde, dass sie eine hohe Transkriptions- und Translationsaktivität aufweisen.

Pistill ist das weibliche Sexualorgan der Pflanze. Hier erfolgt die Befruchtung durch den wind- oder insektenvermittelten Transport von Pollen, den männlichen Gameten.

Die Pistill besteht aus Narbe, Griffel und Fruchtknoten:

- Die Narbe (Stigma) dient zum Auffangen und Keimen des Pollens.
- Der Griffel (Stylus) ist der Bereich, in dem die Pollenschläuche mit männlichen Keimzellen dem Fruchtknoten entgegen wachsen.
- Der Fruchtknoten (Ovarium) enthält die Samenanlagen.

Die **Infloreszenz** wurde gewählt, da die Blütenbildung den Übergang von der vegetativen zur reproduktiven Entwicklungsphase darstellt. Aus diesem Grunde sind in diesem, sowohl Proteine der vegetativen Phase als auch der reproduktiven Phase zu erwarten.

Arabidopsis thaliana folgt in seiner Entwicklung den gleichen Mustern wie alle Angiospermen (Bedecktsamer). Die Entwicklung besteht aus drei sich überlappenden Entwicklungsphasen, die alle im Sprossapikalmeristem zentriert sind:

- vegetatives Meristem, welches Blätter produziert,
- Infloreszenz, welches das Blühmeristem produziert,
- Blühmeristem, welches sich bei einer normalen *Arabidopsis*-Blüte aus vier Kelchblättern (Sepalen), vier weißen Blütenblättern (Petalen), sechs Staubblättern (Stamina) und aus einem Stempel (Pistillum), bestehend aus zwei Fruchtblättern (Karpelle), zusammensetzt

Das Sprossmeristem wird dazu veranlasst, an Stelle der Laubblätter Kelchblätter, Blütenblätter, Staubblätter und Fruchtblätter auszubilden. Dieser Übergang kann erst in einem bestimmten Lebensalter der Pflanze erfolgen, welches innerhalb bestimmter Grenzen genetisch fixiert ist: die Pflanze muss das Stadium der Blühreife erreicht haben. Wann das der Fall ist, ist von Art zu Art sehr unterschiedlich. Ist die Pflanze blühreif, so kann sie zur Blütenbildung gebracht werden. Dabei können zwei Etappen voneinander unterschieden werden, die Blühinduktion und die Differenzierung der Blüten und Blütenstände (Heß, 1991).

Vegetative Meristeme können sich, wenn die Pflanze eine Blühinduktion erfährt, direkt in Blühmeristeme umwandeln. In vielen Fällen, wie z.B. bei *Arabidopsis*, werden vegetative

Meristeme nicht direkt in Blühmeristeme umgewandelt; stattdessen wird das vegetative Meristem erst in ein Infloreszenzmeristem überführt.

1.2 Proteomics

Proteomics ist eine wesentliche Forschungsrichtung der modernen Biowissenschaften. Neben **Proteomics** (Studie der Proteine) kann man noch zwischen **Genomics** (Studie der Geninformation) und **Transkriptomics** (Studie der mRNA-Expression) unterscheiden:

Unter **Genomics** versteht man die Ermittlung (DNA-Sequenzierung) und die Analyse der Informationen, die in der genomischen DNA eines Organismus (im Genom) vorhanden ist. Das gesamte Genom beinhaltet alle Möglichkeiten und Grenzen, die ein Organismus hat, um sich zu entwickeln, zu differenzieren, zu interagieren und auf extrazelluläre und intrazelluläre Änderungen zu reagieren. Diese zellulären Reaktionen werden u.a. mit Techniken aus Transkriptomics und Proteomics intensiv untersucht, und die resultierenden Daten im Kontext der biologischen Fragestellung diskutiert.

Transkriptomics untersucht die Gesamtheit der nach der Transkription von Genen entstehenden mRNA und damit die spezifische Nutzung der durch das Genom gegebenen Möglichkeiten, bezogen auf ein Teilkompartiment (Organ, Gewebe und Zelle) und einen bestimmten Zustand (situationsabhängig). Diese Untersuchungen können u. a. mit Hilfe von DNA-Microarrays erfolgen. Die Anwendung dieser Methode erlaubt Aussagen über die Genexpression sowie über die Variation innerhalb von Gensequenzen im Hochdurchsatzformat (DeRisi *et al.*, 1997; Fields *et al.*, 1999). Jedoch ist es mittels DNA-Microarrays nur begrenzt möglich, posttranskriptionelle (z.B. splicing) bzw. nicht möglich, posttranslatorische Modifikationen nachzuweisen.

Dieses wird insbesondere mit Hilfe von verschiedenen **Proteomics**-Techniken ermöglicht. Dabei bauen Proteomics und Transkriptomics auf die im Genomics- Bereich gewonnenen Informationen auf. Unter Proteomics versteht man das Studium der Gesamtheit der Proteine, die in einer Organelle, einer Zelle, einem Gewebe oder einem Organismus unter definierten Bedingungen vorhanden sind.

Dabei wird ein umfassendes Verständnis von zellulären Vorgängen in Zukunft nur durch die Verknüpfung der gewonnen Daten aller drei Methoden möglich sein.

Mit unterschiedlichsten Proteomics-Techniken können bestimmte zeit- und ortsbedingte Proteingehalte ermittelt werden. Änderungen des Proteingehaltes, Proteinmodifikationen (wie Glykosylierungen, Acetylierungen und Phosphorylierungen), die Bildung von Multiproteinkomplexen sowie Interaktionen der Proteine untereinander und mit anderen Molekülen, lassen Rückschlüsse auf die Funktion dieser Proteine zu.

Als eine klassische Proteomics-Methode kann die 2-dimensionale Gel- Elektrophorese (2DE) angesehen werden (Klose, 1975; Klose *et al.*, 2002; O'Farrell, 1975). Hier erfolgt die Trennung der Proteine in zwei unabhängigen Polyacrylamid-Gelelektrophoresesystemen. Im ersten Trennsystem werden die Proteine durch isoelektrische Fokussierung mittels pH-Gradienten, im zweiten Trennsystem durch SDS-PAGE getrennt. Limitierende Faktoren dieser Methode wurden ausführlich von Kersten *et al.* 2004a beschrieben. Dazu gehört, dass einerseits der Arbeitsaufwand sehr hoch ist und die Standardisierung der Extraktpräparation noch optimierungsbedürftig ist. Andererseits hängt die Genauigkeit der Methode sehr stark von der Qualität der 2DE-Gele und den Laufbedingungen ab, was zur Folge hat, dass diese Methode schwer reproduzierbar ist. Ein weiterer Nachteil ist, dass nur ein Teil der in der Probe enthaltenen Proteine nachgewiesen werden kann. Sehr große und sehr kleine, ebenso wie sehr basische und sehr saure Proteine können nicht erfasst werden. Die Proteine können anschließend mittels verschiedener Färbetechniken wie z.B. Coomassie-Färbung (Diezel *et al.*, 1972), Silber-Färbung (Heukeshoven und Dernick, 1985) oder Fluoreszenz-Färbung (Patton, 2000) visualisiert werden.

Allerdings kann man mit Hilfe der 2DE-Methode allein keine Aussagen über die Identität bzw. Aminosäuresequenz der Proteine machen. Zur Identifizierung der aufgetrennten Proteine wird heutzutage die 2DE-Methode mit einer anschließenden Massenspektroskopie (MS) kombiniert (Andersen *et al.*, 1996; Hillenkamp und Karas, 1990).

Zwar können die aufgetrennten Proteine mittels 2DE in hoher Qualität identifiziert und katalogisiert werden, jedoch liegen die Proteine anschließend in denaturierter Form vor und können somit nicht für zukünftige funktionelle Studien eingesetzt werden.

Für funktionelle Studien ist die Expression der zu untersuchenden Proteine in rekombinanter Form häufig unerlässlich (Kersten *et al.*, 2002). Um Protein-Protein-Interaktionen bzw. um Interaktionspartner zu einem bekannten Protein zu finden, können sowohl die *Two-Hybrid* Systeme (Chien *et al.*, 1991; Dove und Hochschild, 2004; Fields und Sternglanz, 1994; Joung *et al.*, 2000) als auch Protein-Microarrays (MacBeath und Schreiber, 2000; Martinsky, 2004) eingesetzt werden.

Das *Yeast-Two-Hybrid-System* (Y2H) beruht auf der Rekonstitution eines Transkriptionsfaktors in Hefe durch die Wechselwirkung zweier Proteine, von denen das eine Protein (*bait*) an die DNA-Bindungsdomäne (DBD) und das andere Protein (*prey*) an die Transkriptionsaktivierungsdomäne (TAD) des Transkriptionsfaktors fusioniert ist. Diese Interaktion führt innerhalb eines geeigneten Hefe-Stammes zur Transkription und Translation von Reportergenen, welche die Interaktion von zwei Proteinen über detektierbare Signale anzeigen. Häufig erfolgt die Detektion durch die bei der Interaktion ermöglichte Expression eines lacZ Gens. Die Hefe exprimiert β -Galactosidase, wodurch es in einem X-Gal-haltigen Medium zur Spaltung des X-Gals und somit zu einer Blaufärbung kommt. Als Problem bei dieser Methode kann die Detektionsempfindlichkeit betrachtet werden. Die Detektionssignale werden über zwei Stufen verstärkt. Eine Interaktion zwischen zwei Proteinen bewirkt zunächst die Transkription vieler fusionierter Detektionsproteine (z.B. β -Galactosidase), die ihrerseits viele Substratmoleküle umsetzt. Die so erreichte Empfindlichkeit begünstigt jedoch auch das Auftreten von falschpositiven Klonen. Hochdurchsatzmethoden zur Anwendung des Yeast-Two-Hybrid wurden unter anderen von Pandey und Mann 2000 beschrieben.

Als eine weitere vielversprechende Proteomics-Methode kristallisiert sich in den letzten Jahren die Protein-Microarraytechnologie heraus (MacBeath und Schreiber, 2000; Templin *et al.*, 2003; Zhu und Snyder, 2003a). Ein wesentlicher Vorteil dieser Methode ist die gleichzeitige Untersuchung einer hohen Anzahl zu untersuchender Proteine auf kleinsten Raum, wie z. B. enzymatische Umsetzung und Wechselwirkungsanalysen mit unterschiedlichsten potentiellen Interaktionspartnern.

Hierbei ist die Etablierung von Versuchsapparaturen unerlässlich, die es zulassen, den Versuchsablauf zu miniaturisieren und simultan zu gestalten, um den Versuchsaufwand, die Kosten sowie den Verbrauch des Probenmaterials zu reduzieren. Bei dieser Methode werden die (meist rekombinanten) Proteine auf der Oberfläche von Objektträgern angeordnet und immobilisiert. Durch die gerichtete Anordnung können die korrespondierenden DNA-Sequenzen direkt mit den gespotteten Proteinen verknüpft werden (Cahill, 2001).

1.3 Gewinnung von Proteinen im Hochdurchsatz

Die Gewinnung von rekombinanten Proteinen im Hochdurchsatz beinhaltet zunächst die Synthese der cDNA-Moleküle aus mRNA, welche anschließend mittels Hochdurchsatz-Klonierung in geeignete Vektorsysteme überführt werden.

1.3.1 Strategien zur Hochdurchsatz-Klonierung

ORF-Ansatz

Hierbei handelt es sich um eine parallel gerichtete Klonierung von *open reading frames* (ORFs) in möglichst voller Länge nach Amplifikation ausgehend von cDNA oder genomischer DNA (bei Genen ohne Introns) mittels genspezifischer Primer.

Diese gerichtete Klonierung kann mittels Restriktionsenzym-Spaltung und anschließender Ligation oder mittels Rekombination, z.B. mit der GATEWAY™-Technologie, erfolgen. Letztere Methode ist besonders gut für den Hochdurchsatz geeignet (Heyman *et al.*, 1999; Walhout *et al.*, 2000).

Die GATEWAY-Technologie basiert im wesentlichen auf zwei Rekombinationen (BP- und LR-Reaktion, Abb. 1.1) und beruhend auf der Verwendung von Lambda-Rekombinase und Rekombinationsstellen. Die kodierende Sequenz der zu untersuchenden Gene kann als PCR-Produkt oder von einer cDNA-Bibliothek durch sequenzspezifische Rekombination (BP-Reaktion) in den sog. Entryvektor übertragen werden. Mit Hilfe des Entryvektors ist es dann möglich, das Gen in verschiedene Destinationsvektoren und Expressionssysteme zu überführen.

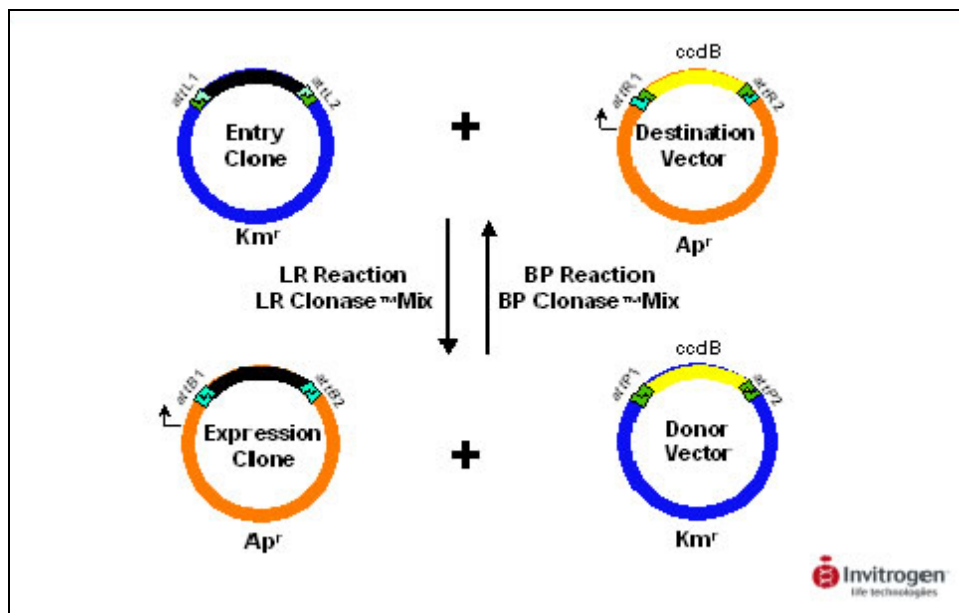


Abb. 1.1: Schematische Darstellung der Rekombinationstechnologie (GATEWAY-Technologie).

Im wesentlichen verlaufen die BP- und die LR- Reaktionen folgendermaßen:

Bei der **BP-Reaktion** wird das Gen mit zwei attB-Rekombinationsstellen flankiert und mit den beiden attP-Rekombinationsstellen des Donorvektors rekombiniert. Damit das Gen in der richtigen Orientierung in den Donorvektor rekombiniert wird, gibt es zwischen den

attB1- und attB2-Rekombinationsstellen, wie auch zwischen den attP1- und attP2-Rekombinationsstellen geringe Unterschiede in der DNA-Sequenz, die dazu führen, dass attB1 nur mit attP1 und attB2 nur mit attP2 rekombinieren können. Die bei dieser Rekombination entstandenen Eingangsklone (Entryklone) enthalten das Gen, welches jetzt von den attL1- und attL2-Rekombinationsstellen flankiert ist.

Die **LR-Reaktion** beinhaltet den Transfer des Gens, flankiert mit attL-Rekombinationsstellen, in einen Expressionsvektor und funktioniert nach dem gleichen Prinzip. Die zwei attL-Rekombinationsstellen werden nun mit den beiden attR-Rekombinationsstellen des Destinationsvektors rekombiniert. Die richtige Orientierung wird auch hier wieder durch kleine DNA-Sequenz Unterschiede in den entsprechenden Rekombinationsstellen gewährleistet.

Sowohl die Entryklone als auch die Expressionsklone können aufgrund variierender Resistenzen selektiert werden. Jedoch kommt bei beiden Reaktionen (BP und LR) noch eine weitere Selektionsmöglichkeit hinzu. Sowohl beim Donorvektor (BP-Reaktion) als auch beim Destinationsvektor (LR-Reaktion) befindet sich zwischen den Rekombinationsstellen das *ccdB*-Gen. Das *ccdB*-Gen ist ein „Suizid-Gen“, dessen Genprodukt mit der *E. coli*-DNA-Gyrase interferiert (Bernard und Couturier, 1992). Die Selektion basiert auf dem Austausch des *ccdB*-Gens gegen das einzubringende Gen. Nur die Plasmide ohne *ccdB*-Gen, aber mit der entsprechenden Resistenz können Kolonien bilden.

Bibliotheken-Ansatz

Eine zweite Strategie zur Hochdurchsatz-Klonierung beruht darauf, dass man die Gesamt-RNA eines definierten Gewebes, Zelltypus oder Organismus', nach Entfernung der genomischen DNA, gewinnt. Danach wird die mRNA isoliert und davon ausgehend die cDNA synthetisiert. Die generierten cDNA-Moleküle werden anschließend in einen geeigneten Expressionsvektor kloniert.

Diese Vektoren ermöglichen die Expression von Proteinen, welche an einem Tag fusioniert sind. Mit Hilfe dieses Tags können die potentiellen Expressionsklone selektiert werden und darauf basierend eine cDNA Expressionsbank hergestellt werden (Büssow *et al.*, 1998; Büssow *et al.*, 2000; Clark *et al.*, 1999; Holz *et al.*, 2001).

Mit dem ORF-Ansatz können die Gene im richtigen Leserahmen (*reading frame*) und in voller Länge kloniert werden. Jedoch ist dieser Ansatz von der Verfügbarkeit der Sequenzinformation sowie der Qualität der Genannotation abhängig. Beim Bibliotheken-

Ansatz sind diese Informationen nicht notwendig, dafür wird jedoch nur ein Teil der Gene im richtigen Leserahmen kloniert. Einige selten vorkommende RNAs werden nicht unbedingt erfasst und nicht alle cDNAs können in voller Länge kloniert werden. Weiterhin können Klone auch einen Teil der 5'-untranslatierten Region aufweisen, die eine artifizielle zusätzliche Proteinsequenz bilden und damit weitere Untersuchungen beeinflussen kann. Des Weiteren können interne Stopcodonen der 5'-UTR die Proteinexpression verhindern. Im Rahmen dieser Arbeit wurde zur Hochdurchsatzherstellung von Expressionsklonen der Bibliotheken-Ansatz verwendet.

1.3.2 Proteinexpression im Hochdurchsatz

Nach der Wahl des Klonierungssystems steht man vor der Frage, welches Expressionssystem das geeignetste ist. Die wohl bekanntesten Expressionssysteme sind das bakterielle System *Escherichia coli* und die eukaryontischen Hefesysteme *Saccharomyces cerevisiae* und *Pichia pastoris*.

Dabei werden hier andere Expressionssysteme als *E. coli* zwar kurz mit Vor- und Nachteilen erläutert, jedoch kamen diese für diese Arbeit nicht in Frage, da die Expressions- und Aufreinigungssysteme für *E. coli* in der Arbeitsgruppe bereits etabliert waren.

E. coli Zellen sind leicht genetisch zu manipulieren und zeichnen sich durch eine hohe Wachstumsrate, niedrige Nährstoffansprüche und eine hohe Kapazität der Proteinsynthese aus. Darüber hinaus sind sie einfach zu kultivieren und kostengünstig. Es konnte gezeigt werden, dass auch menschliche Proteine in löslicher Form in diesem System aufgereinigt werden können. So lagen bei Braun *et al.* 2002 46% der 326 getesteten humanen Proteine nach der Expression in löslicher Form vor. Jedoch gibt es auch hier Nachteile: Da *E. coli* als prokaryontische Organismen über keine membrangetrennten Zellorganellen verfügen, werden die Proteine direkt im reduzierenden Zytosol synthetisiert, und eukaryontische Proteine werden dann häufig unvollständig gefaltet. Darüber hinaus können die überproduzierten und falsch gefalteten Proteine unlösliche und funktionslose Einschlusskörperchen (*inclusion bodies*) bilden. Des Weiteren sind die bakteriellen Expressionssysteme nicht in der Lage, solche für die eukaryontischen Proteine wichtigen posttranslationalen Modifikationen, wie z.B. die Ausbildung von Disulfidbrücken, Glykosylierung oder Phosphorylierung, durchzuführen. Bis heute wurden etwa 200 posttranslationale Modifikationen beschrieben, wobei diese teilweise großen Einfluss auf die zellulären Prozesse und die Signalkaskaden haben (Meri und Baumann, 2001; Parekh und Rohlf, 1997). Demgegenüber stehen die eukaryontischen Expressionssysteme (wie Säugertier-, Insekten- und Hefezellen), bei denen die Zellen einen vom Zytoplasma

abgegrenzten Zellkern sowie Zellorganellen mit eigener Membranen besitzen. Die Stoffwechselreaktionen sind nun nicht mehr nur auf das Zytoplasma begrenzt, sondern auf die unterschiedlichen Zellorganellen aufgeteilt. Dadurch sind komplexere Stoffwechsel- und Syntheseschritte, z.B. die Glykosylierung, möglich. Darüber hinaus kann in diesen Systemen die Sekretion der Proteine erfolgen, welche für eine korrekte Faltung der eukaryontischen Proteine häufig unerlässlich ist.

Mit dem Säugetierzellen-Expressionssystem kann man somit einen hohen Anteil an funktionellen Proteinen erhalten. Dieses System produziert die gleichen post-translationalen Modifikationen und erkennt die gleichen Signale für Synthese, Weiterverarbeitung und Sekretion wie die Zellen, aus denen die Sequenzen ursprünglich stammen. Jedoch ist dieses System durch eine niedrige Expressionsrate gekennzeichnet. Darüber hinaus ist es sehr komplex, schwer zu konstruieren, instabil und teuer.

Dagegen erlaubt das Insektenzellen-Expressionssystem die Expression von einer vergleichsweise höhere Ausbeute an Proteinen. Aber auch dieses System hat den Nachteil, dass es unter aufwendigen Wachstumsbedingungen zu kultivieren ist.

Im Gegensatz dazu stellen Hefen ein einfaches eukaryontisches Expressionssystem dar, welches sich durch ein schnelles Wachstum und eine hohe Ausbeute auszeichnet. In diesem System können einige eukaryontische posttranslationale Modifikationen stattfinden. Zudem konnte gezeigt werden, dass ein Großteil der zu exprimierenden Proteine in diesem System löslich vorliegt (Harashima, 1994).

Neben der Klonierung im Hochdurchsatz wurden Techniken entwickelt, die eine parallele Expression von vielen Proteinen ermöglichen (Büssow *et al.*, 1998; Lüking *et al.*, 1999). Für diesen Zweck sind *E. coli* und Hefe als Expressionssysteme besonders gut geeignet. Um Proteine sowohl in *E. coli* als auch in *P. pastoris* exprimieren zu können, haben Lüking *et al.* 2003a ein duales Vektorsystem entwickelt, welches die Expression rekombinater Proteine in beiden Systemen ermöglicht.

1.3.3 Proteinaufreinigung im Hochdurchsatz

Für die Aufreinigung von Proteinen im Hochdurchsatz ist die Anwendung von Affinitätstags weitverbreitet. Die für die Tags kodierenden Bereiche werden zusammen mit den Bereichen für das zu exprimierende Protein transkribiert und translatiert. Aufgrund der hochaffinen Bindung der Tags an bestimmte Trägermaterialien ist die Aufreinigung von Proteinen im Hochdurchsatz möglich (Nilsson *et al.*, 1997). Häufig verwendete Tags sind z.B. die Glutathion S-Transferase (GST)-, Histidin (His)-, FLAG-Tag, Biotin-Tag und Protein A (Phizicky *et al.*; 2003; Schatz *et al.*; 1993; Scheich *et al.*, 2003). Ein Nachteil der

Verwendung von Tags ist, dass es durch den Tag zu einer Veränderung der Struktur und der Funktion der Proteine kommen kann.

Braun *et al.* 2002 haben die Effizienz der Hochdurchsatzaufreinigung von humanen Proteinen aus Bakterienzellen mit Hilfe von vier unterschiedlichen Tags miteinander verglichen. Sie konnten zeigen, dass GST-Fusionsproteine eine höhere Expressionsrate und Aufreinigungseffizienz unter denaturierenden Bedingungen aufwiesen, als die His-Fusionsproteine. Scheich *et al.* 2003 verglichen mit Hilfe eines N-terminalen Tandemtags (His₆- und GST-Tag) die Effizienz der Hochdurchsatz-Proteinaufreinigung unabhängig von der Expressionsrate. Sie konnten zeigen, dass die Proteinaufreinigung über den His₆-Tag unter Verwendung der Ni-NTA Aagrose effizienter war als über den GST-Tag.

1.4 Protein-Arrays und deren Anwendung

1.4.1 Grundlagen zur Array-Technologie

Das Prinzip eines Arrays basiert auf der systematischen Anordnung von Akzeptormolekülen auf einer Fläche. Damit erlaubt ein Array die parallele Analyse einer hohen Anzahl von Molekülen und die Zuordnung der Ergebnisse zu den immobilisierten Molekülen.

Methodisch basieren Protein-Arrays auf den technischen Entwicklungen, die zuvor für DNA-Arrays erfolgten. Bei den DNA-Microarrays (Hochdichte DNA-Filter) werden cDNAs in Form rekombinanter *E. coli* Kolonien auf eine Membran gespottet; hierbei handelt es sich im weitesten Sinne um eine Art cDNA-Dot Blot. Das Ergebnis waren die sogenannten ‚gridded libraries‘, bei denen jeder Klon eine definierte Position auf der Filtermembran hat (Büssow *et al.*, 1998). Eine weitere Verwendungsmöglichkeit für diese gespotteten Membranen war die Hybridisierung mit einer DNA-Sonde.

Büssow *et al.* 1998 zeigten eine weitere Anwendungsmöglichkeit dieses Arrayformates, die Verwendung als Protein-Microarrays (Hochdichte Protein-Filter). Bei beiden werden cDNA enthaltende Bakterien in hoher Dichte auf PVDF-Membranen gebracht. Allerdings werden nach dem Spotten die Filter unterschiedlich behandelt, so dass man einerseits cDNA-Filter und andererseits Protein-Filter erhalten kann. Auf diese Weise kann man die Filter für ein paralleles Screening mit DNA-Sonden bzw. Antikörpern benutzen.

Aufgrund der hohen Kosten bezüglich der Herstellung bzw. des Kaufes von solchen Membranen und des hohen Verbrauches an Probenmaterialien auf den Membranen, mit einer Größe von 22 cm x 22 cm, wurde die Weiterentwicklung zu Microarrays unumgänglich.

Microarrays bieten dabei generell eine Reihe von Vorteilen gegenüber Macroarrays oder anderen Proteomics-Technologien. Sie zeichnen sich durch eine hohe Dichte von immobilisierten Molekülen auf kleinstem Raum, sowie durch eine hohe Reproduzierbarkeit aus. Weitere Vorteile von Microarrays sind:

- der Einsatz sehr kleiner Volumina, teils limitierender Substanzen (wie z.B. Serum, Enzyme), was eine hohe Anzahl von unabhängigen Wiederholungen ermöglicht
- der Einsatz kleinster Mengen der gespotteten Moleküle, was ebenfalls eine hohe Anzahl von unabhängigen Wiederholungen ermöglicht
- die systematische Anordnung vieler Moleküle, was eine direkte Zuordnung der Signale ermöglicht
- stärkere Signale durch konzentrierte und fokussierte Auftragung (Spotten) der Moleküle
- mögliche Quantitative Analysen im Gegensatz zu Macroarrays, bei denen eine nicht quantifizierbare *in situ* Proteinexpression stattfindet
- statistische Datenerhebung durch mögliche Spotreplikationen auf einem Microarray und einer hohen Anzahl von verwendeten Microarrays

Ein Durchbruch in diese Richtung gelang damit, dass statt der flexiblen Nylonmembran festes Trägermaterial wie beschichtetes Glas oder aminiertes Polypropylen verwendet wurde (Skena *et al.*, 1995). Zudem erleichtert das feste Trägermaterial die Konstruktion kleinster Hybridisierungskammern und die Automatisierung des gesamten Herstellungs- und Detektionsprozesses. Außerdem kann es bei flexiblen Oberflächen dazu kommen, dass das gespottete Material sich so weit ausdehnt, dass es einerseits zu einer Kontamination mit den benachbarten Spots kommen kann und andererseits keine exakte Aussage über Signal-zu-Hintergrundverhältnis gemacht werden kann (Haab *et al.*, 2001; Lüking *et al.*, 1999; Templin *et al.*, 2003). Zusätzlich kam es mit der Entwicklung neuer Oberflächen zu einer verbesserten Immobilisierung der Moleküle an den Oberflächen.

Aufgrund der Erfolge der analytischen Studien mit Hilfe der DNA-Microarraytechnologie lag der Versuch der Etablierung der Protein-Microarraytechnologie für funktionelle Studien nahe. Bei der Etablierung dieser Technologie müssen jedoch viele Aspekte berücksichtigt werden.

DNA ist leicht zu extrahieren, relativ unempfindlich und somit ein vergleichsweise leicht zu handhabendes Probenmaterial, während das Exprimieren und das Isolieren von Proteinen, sowie das Aufrechterhalten der Funktionalität von Proteinen problematisch sind. Des

weiteren unterscheiden sich DNA und Proteine in ihrem Aufbau und ihrer Komplexität. Die DNA ist relativ einfach und einheitlich aufgebaut und besteht aus vier variablen Bausteinen, den Nukleotiden. Nukleinsäuren sind hydrophil und wegen der Phosphatgruppen im Zuckerrückgrat negativ geladen. Proteine dagegen bestehen aus 20 variablen Bausteinen, den Aminosäuren. Aufgrund dieser unterschiedlich aufgebauten Säuren und der variierenden Zusammensetzung der Aminosäuren bei den einzelnen Proteinen, können hier diverseste Moleküle gebildet werden, die sich in Struktur und Ladung stark voneinander unterscheiden können. Proteine können somit u.a. hydrophil oder hydrophob sowie basisch (positiv geladen) oder sauer (negativ geladen) sein.

Die Berücksichtigung dieser Aspekte macht offensichtlich, dass für die Immobilisierung von Nukleinsäuren eine einheitliche Arrayoberfläche ausreichend ist, um optimale Bindungsbedingungen mit höchster Bindungskapazität zu erreichen. Bei den Proteinen dagegen wird deutlich, dass die Wahl des geeigneten Oberflächenmaterials wesentlich anspruchsvoller ist. Die Oberflächen lassen sich in zwei Kategorien unterteilen: planare, zweidimensionale und in räumliche, dreidimensionale Oberflächen.

1.4.2 Oberflächen

Zur ersten Kategorie gehören die planaren, nicht-gelbeschichteten Microarrays, deren Glasoberfläche mit unterschiedlichen Chemikalien modifiziert wurde. Dazu gehören u.a. Aldehyd- (MacBeath und Schreiber, 2000), Epoxy- (Angenendt *et al.*, 2003b), nickelbeschichtete (MacBeath und Schreiber, 2000) und poly-L-Lysin-Oberflächen (Ge, 2000; Snappyan *et al.*, 2003).

Die dreidimensionalen Microarrays dagegen weisen eine gelbeschichtete Oberfläche, wie Polyacrylamid (Arenkov *et al.*, 2000) oder Agarose (Afanassiev *et al.*, 2000) auf. Hierzu gehören auch die, in dieser Arbeit verwendeten FASTTM-Slides von Schleicher & Schuell; diese haben eine Nitrozellulose ähnliche Oberfläche.

Dreidimensionale Microarrays haben im Vergleich zu den nicht-gelbeschichteten Microarrays eine höhere Kapazität, um die Proteine zu immobilisieren. Die Microarrays ermöglichen eine einheitliche Verteilung der Feuchtigkeit, was die Proteindenaturierung verringert. Dies wiederum kann die Funktionalität und die Aktivität der gespotteten Proteine positiv beeinflussen (Angenendt *et al.*, 2002).

1.4.3 Detektion

Neben der direkt markierten Detektion gibt es noch die markierungsfreie Detektion. Da in der vorliegenden Arbeit ausschließlich mit der markierten Detektion gearbeitet wurde, wird im Folgenden nur auf diese detaillierter eingegangen. Zusammenfassungen der markierungsfreien Detektion sind in Glökler und Angenendt 2003 und Zhu *et al.* 2003b zu finden.

Bei der markierten Detektion können die Interaktionspartner entweder direkt markiert oder mit spezifisch markierten Antikörpern detektiert werden. Für die direkte Markierung werden häufig fluoreszierende Farbstoffe (z.B. Cy3 und Cy5) eingesetzt. Diese sind leicht in der Handhabung, extrem sensitiv und ermöglichen eine Auswertung in hoher Spotdichte (Haab *et al.*, 2001; Templin *et al.*, 2002). Eine weitere fluoreszenzbasierte Detektion ist die *rolling circle amplification* (RCA) (Kingsmore und Patel, 2003; Schweitzer *et al.*, 2002). Hier wird der Detektionsantikörper mit einem DNA-Primer markiert. Nachdem der Antikörper an den immobilisierten Interaktionspartner gebunden hat, findet in Anwesenheit von fluoreszierenden Nukleotiden, DNA-Polymerase und einer zirkulären DNA eine sich im Kreis drehende Replikation statt. Die Fluoreszenz, die anschließend gemessen werden kann, ist direkt proportional zu der Menge der spezifischen Proteine in der Originalprobe (Schweitzer *et al.*, 2002). Diese Methode ist äußerst sensitiv. Aber auch andere sensitive Detektionsmethoden sind zu erwähnen, wie z.B. Radioaktivität (MacBeath und Schreiber, 2000; Zhu *et al.*, 2000).

Die Detektion unter Verwendung spezifisch markierter Antikörper wurde als erstes eingesetzt, um Proteine auf Filterarrays (Büssow *et al.*, 1998; Joos *et al.*, 2000) sowie Glassarrays (Guschin *et al.*, 1997) zu detektieren.

Im wesentlichen gibt es hier drei Strategien für die Detektion:

Beim direkten Immunoassay werden die Proteine als Akzeptormoleküle direkt auf der Oberfläche immobilisiert und mit Hilfe der Interaktion mit einem markierten Detektionsantikörper sichtbar gemacht.

In Gegensatz zum ersten Assay werden beim Antigen capture Immunoassay die Antikörper als Akzeptormoleküle auf der Oberfläche immobilisiert und die Proteine, die eine Interaktion mit den Antikörpern eingehen, sind direkt markiert.

Beim Sandwichimmunoassay werden ebenfalls die Antikörper als Akzeptormoleküle auf die Oberfläche gebracht, anschließend werden die Proteine gebunden und diese Bindung kann dann mit einem zweiten markierten Detektionsantikörper gezeigt werden.

1.4.4 Analytische und funktionelle Protein-Microarrays

Hinsichtlich der Anwendung von Protein-Microarrays unterscheidet man zwischen analytischen Protein-Microarrays und funktionellen Protein-Microarrays.

Bei den analytischen Protein-Microarrays werden entweder die Antigene oder die Antikörper in hoher Dichte auf der Oberfläche immobilisiert und z.B. zur Ermittlung der Existenz und der Konzentration von Antikörpern bzw. Antigenen in komplexen Lösungen, wie Serum, eingesetzt. Basierend auf Serum-*screening* Versuchen können so z.B. diagnostische Proteinmarker identifiziert werden (Cahill, 2001; Glökler und Angenendt, 2003; Petricoin *et al.*, 2002a; Petricoin *et al.*, 2002b; Walter *et al.*, 2002). Des Weiteren können die Antigen-Microarrays eingesetzt werden, um die Spezifität bzw. Kreuzreaktivität von Antikörpern zu bestimmen (Kersten *et al.*, 2003; MacBeath und Schreiber, 2000; Michaud *et al.*, 2003).

Für funktionelle Protein-Microarrays wird eine große Anzahl an aufgereinigten, rekombinanten und vorzugsweise nativen Proteinen von einer Zelle, einem Gewebe oder einem Organismus auf der Oberfläche immobilisiert. Diese Arrays können verwendet werden, um eine Wechselwirkung der immobilisierten Proteine mit anderen Proteinen (MacBeath und Schreiber, 2000; Zhu *et al.*, 2001), DNA (Ge, 2000; Kersten *et al.*, 2004b; Snapyan *et al.*, 2003), kleinen Molekülen (Burton und McGehee, 2004; MacBeath und Schreiber, 2000; Zhu *et al.*, 2001) oder die Modifikation der Proteine durch Enzyme (Arenkov *et al.*, 2000; MacBeath und Schreiber, 2000; Zhu *et al.*, 2000, Kramer *et al.*, 2004) zu identifizieren.

1.5 Phosphorylierung als eine posttranslationale Proteinmodifikation

Die Phosphorylierung von Proteinen ist eine wichtige posttranslationale Proteinmodifikation. Dabei erfolgt die Übertragung eines negativ geladenen γ -Phosphatrestes eines Phosphatgruppendonors (meist Adenosin- und in seltenen Fällen Guanosintriphosphat) auf eine der Akzeptoraminosäuren. Durch Phosphorylierungen kann es zu Konformationsänderungen der Polypeptidketten und somit zu einer Veränderung der Proteineigenschaften kommen. Der Vorgang ist reversibel.

Phosphorylierungen sind wesentlich an der Kontrolle von intrazellulären Prozessen in Eukaryonten, wie Stoffwechsel, Transkription, Translation, Zellteilung und Zellwachstum sowie Membrantransport und Signalübertragung beteiligt. Sehr häufig reguliert die Phosphorylierung die Aktivität von Enzymen. Darüber hinaus kann die Interaktionsfähigkeit

mit anderen Proteinen, DNA sowie die Lokalisation von Proteinen in der Zelle beeinflusst werden.

Die Stärke und Dauer der Phosphorylierung der Proteine wird einerseits durch die Proteinkinasen, welche die Phosphatreste übertragen und andererseits durch die Proteinphosphatasen, welche die Phosphorsäureester-Bindungen hydrolysieren, reguliert.

Es gibt zwei Gruppen von Kinasen, die über ihre Lokalisation definiert werden: Rezeptorkinasen, welche sich in der Membran befinden und zytosolische Proteinkinasen, welche sich frei im Zytosol bewegen.

Die Unterteilung der Proteinkinaseklassen kann nach ihren Akzeptoraminosäuren in Serin/Threonin-, Tyrosin-, Histidin/Arginin/Lysin-, Cystein- und Aspartat/Glutamat-Kinasen erfolgen. Während in den Prokaryonten die Phosphorylierung an Histidinresten überwiegt, werden in den Eukaryonten fast ausschließlich Serin/Threonin- und Tyrosin-Phosphorylierungen beobachtet (Kerk et al, 2002; Luan, 1998). In Pflanzen sind bisher hauptsächlich Serin/Threonin spezifische Proteinkinasen identifiziert worden (Hardie, 1999; Mundy und Schneitz, 2002).

Die herausragende Bedeutung der Kinasen wird durch den hohen Anteil von Proteinkinasegenen im eukaryontischen Genom verdeutlicht. In *S. cerevisiae*, *C. elegans* und in *D. melongaster* gehören Proteinkinasen zu den Proteinklassen mit den häufigsten Vertretern. Es wird angenommen, dass durchschnittlich 2% aller eukaryotischen Gene dieser drei Organismen für Proteinkinasen codieren (Lander et al., 2001; Rubin et al., 2000). Champion et al. 2004 gehen davon aus, dass etwa 4% des *Arabidopsis* Genoms Serin/Threonin-Proteinkinasen repräsentiert.

Da im Rahmen der vorliegenden Arbeit ausschließlich die Phosphorylierungsspezifität von MAP-Kinasen untersucht wurde, wird im weiteren Verlauf der Arbeit nur auf diese Familie der Kinasen detaillierter eingegangen. Umfassende Zusammenfassungen aller Familien der pflanzlichen Proteinkinasen sind in Champion et al., 2004 und Hardie, 1999 zu finden.

1.5.1 Mitogen-aktivierte Proteinkinasen (MAPKs)

Bei den MAP-Kinasen erfolgt die Übertragung des Phosphatrestes über eine Kaskade. Beim Übergang von einem Element zum nächsten, kann es zu einer Verstärkung des Signals kommen. Wirken einige Signale auf mehrere Zielproteine, kann es zu einer Verzweigung der Übertragungswege kommen, wodurch sich die Vielfalt der Reaktionsmöglichkeiten, ausgehend von dem ursprünglichen Reiz, erhöhen kann.

MAP-Kaskaden sind sowohl bei der Signalübertragung verschiedener biotischer und abiotischer Stressfaktoren, wie Temperatur- oder Trockenheitsstress, Pathogeninfektion und

Verwundung als auch bei der Weitergabe von Pflanzenhormonen, wie Ethylen, Abszinsäure und Auxin beteiligt (Jonak *et al.*, 2002; Lu *et al.*, 2002; Tena *et al.*, 2001). MAP-Kinasen versetzen die Zelle in die Lage, auf Änderungen in der extrazellulären Umgebung zu reagieren (Abb. 1.2). Dabei führt die Bindung der Signalmoleküle an membranständige Rezeptoren zu deren Konformationsänderung, wodurch das Signal über die intrazelluläre Domäne in das Zellinnere gelangt. Die anschließende Signalübertragung zum Zellkern erfolgt dann in der Regel über eine dreistufige Phosphorylierungskaskade (Asai *et al.*, 2002; Jonak *et al.*, 2002; Tena *et al.*, 2001). Dabei phosphoryliert eine Serin/Threonin-Proteinkinase (Mitogen-aktivierte Proteinkinase Kinase Kinase, MAPKKK) eine nachfolgende dualspezifische Kinase (MAPKK), die wiederum eine MAP-Kinase (MAPK) aktiviert. Es kommt nur zu einer Aktivierung der MAPK, wenn sowohl ein Threonin- als auch das Tyrosinrest in der „Aktivierungsschleife“ durch die MAPKK phosphoryliert wird (Keyse, 2000; Nühse *et al.*, 2000). Diese Reste in der „Aktivierungsschleife“ (Hardie, 1999; Tena *et al.*, 2001; Zhang und Klessig, 2001) sind nur durch eine Aminosäure (Thr-X-Tyr) voneinander getrennt.

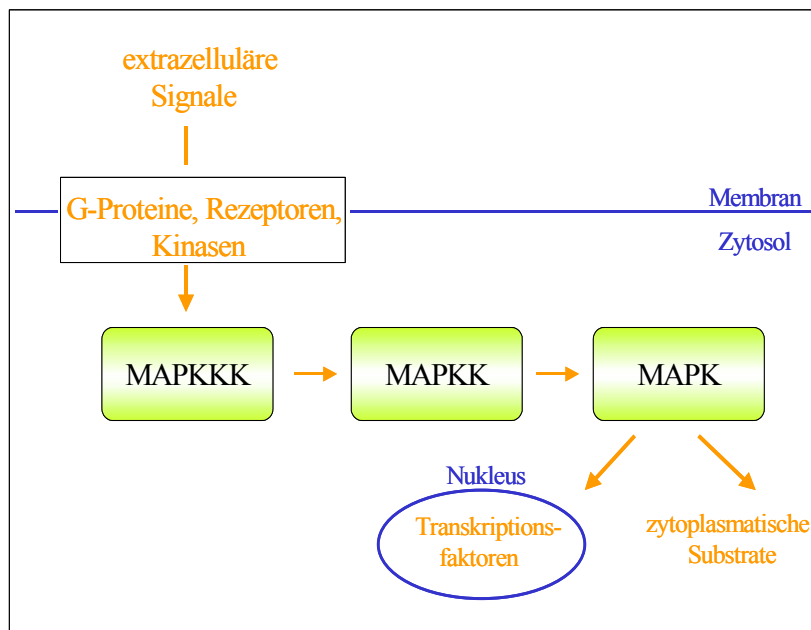


Abb. 1.2: Schematische Darstellung des MAP-Kinasen Signalweges.

Die MAP-Kinase steht am Ende verschiedener Signaltransduktionsketten, die Wachstums- und Differenzierungssignale durch Phosphorylierung von der Plasmamembran zum Kern weiterleiten. Über die Substrate der MAP-Kinasen ist erst wenig bekannt. Substrate, welche die MAP-Kinasen *in vitro* phosphorylieren, könnten Aufschluss über eine mögliche

Funktion dieser Kinasen im Zellzyklus geben. Aktivierte MAP-Kinasen können beispielsweise in den Zellkern gelangen und dort Transkriptionsfaktoren phosphorylieren und damit deren DNA Bindung regulieren. Somit stellt die MAP-Kinase eine Verbindung zwischen Zytosol und Nukleus dar, indem es durch extrazelluläre Stimulation zu einer Änderung in der Genexpression kommen kann.

MAPKgroup 2002 beschreiben bei *Arabidopsis* die Existenz von bisher 20 MAPKs, (MPKs) 10 MAPKKs und 60 MAPKKKs.

Die MPKs werden in vier Gruppen unterteilt (A-D), wobei aufgrund von Sequenzvergleichen innerhalb der Konsensussequenz die MPKs aus zwei Übergruppen bestehen. Die eine Übergruppe (Gruppen A - C) weist an der Phosphorylierungsstelle das Aminosäuremotiv TEY (Threonin-Glutamat-Tyrosin) und die andere Übergruppe (Gruppe D) das Motiv TDY (Threonin-Asparagat-Tyrosin) auf.

Bei den pflanzlichen MAPKKs besteht die Phosphorylierungsstelle aus dem Aminosäuremotiv T/SxxxxTS und unterscheidet sich von den tierischen Aminosäuremotiv: T/SxxxTS (MAPKgroup, 2002).

Die Regulation der MAPKKK ist dagegen deutlich komplexer, da es hier eine große Anzahl von regulatorischen Motiven gibt, z.B. Bindestellen für GTP-bindende Proteine oder Serin/Threonin- und Tyrosin Phosphorylierungsstellen.

In der vorliegenden Arbeit wurden die *Arabidopsis thaliana* MPK3 und MPK6 verwendet. Beide MAP-Kinasen gehören zu der Gruppe A, d.h. bei beiden besteht die Phosphorylierungsstelle aus dem Aminosäuremotive TEY. Die Kinasen dieser Gruppe sind überwiegend an Signalübertragungen beteiligt, die durch Änderungen der Umweltbedingungen und des Hormonhaushaltes ausgelöst werden.

1.6 Zielsetzung

Mit dem zunehmenden Wissen über die DNA-Information (Genomics) von immer mehr Lebewesen, steigt auch das Interesse an der Untersuchung der Genfunktionen durch Anwendung funktioneller Proteomstudien. In der vorliegenden Arbeit wurde dazu eine relativ neue Proteomics-Technologie, die Protein-Microarraytechnologie gewählt. Diese Technologie erlaubt es, funktionelle Studien miniaturisiert und parallelisiert durchzuführen. Im Rahmen dieser Arbeit sollte die Protein-Microarraytechnologie angewandt werden, um erste funktionelle Analysen von *Arabidopsis thaliana*-Proteinen im hohen Durchsatz durchzuführen. Das setzt die Expression rekombinanter *Arabidopsis thaliana*-Proteine voraus, so dass sich für diese Arbeit zwei Teilziele ergaben:

1. Expression von *Arabidopsis thaliana*-Proteinen nach cDNA-Klonierung
2. nachfolgende Analyse dieser Proteine unter Verwendung von Protein-Microarrays

Zu Gewinnung rekombinanter Proteine von *Arabidopsis thaliana* im Hochdurchsatz wurden zwei Gewebe als interessante Untersuchungsobjekte ausgesucht, Pistill und Infloreszenz. Aus beiden sollten, ausgehend von den entsprechenden mRNAs, die cDNAs synthetisiert, im Hochdurchsatz kloniert und damit zwei unterschiedliche cDNA-Bibliotheken konstruiert werden. Die Konstruktion der Pistill cDNA-Bibliothek sollte in einem GATEWAY-Entryvektor erfolgen. Mit diesem System besteht die Möglichkeit, Gene mittels gerichteter Rekombination in unterschiedliche Expressions- und Analysesysteme zu transferieren. Dagegen sollte die Infloreszenz cDNA-Bibliothek direkt in einen *E. coli* Protein-Expressionsvektor kloniert werden. Dieser Vektor ermöglicht die Hochdurchsatzexpression von Fusionsproteinen mit einem N-terminalen RGS His₆-Tag. Die potentiellen Expressionsklone dieser neu gewonnenen cDNA-Bibliothek sollten, unter Verwendung von Hochdichte-Proteinfiltern, sogenannten Protein-Macroarrays, ermittelt und in einer Proteinexpressions-Unterbibliothek neu angeordnet werden. Basierend auf den Sequenzinformationen der Proteinexpressions-Unterbibliothek sollte ein Uniklonset erstellt werden, in dem jeder Klon möglichst nur einmal vorläge. Die Proteine dieses Uniklonsets sollten exprimiert und aufgereinigt werden.

Im zweiten Teil der Arbeit sollten alle *Arabidopsis*-Proteine des neu zusammengestellten Uniklonsets in einer ersten funktionellen Studie auf Protein-Microarrays verwendet werden, um potentielle Phosphorylierungs-Targets von zwei interessanten *Arabidopsis*-MAP-Kinasen, MPK 3 und 6, zu identifizieren. Dazu sollten die Proteine auf Microarrays angeordnet werden. Ein Nachweis der exprimierten Proteine auf den Protein-Microarrays

sollte mit Hilfe eines anti-RGS His₆-Antikörpers durchgeführt werden. Auch die eigentlichen *in vitro* Phosphorylierungsstudien sollten auf diesen Microarrays stattfinden. Zu diesem Zwecke sollte zunächst ein Quantifizierungssystem zur statistischen Analyse der Microarrays nach Phosphorylierung etabliert werden. Dieses Quantifizierungssystem sollte der Identifizierung potentieller Targets der beiden MAP-Kinasen dienen. Die identifizierten potentiellen Phosphorylierungs-Targets sollten in einem unabhängigen *in vitro* Testsystem, mittels Phosphorylierungen auf PVDF-Membranen verifiziert werden.

Mit dem etablierten Phosphorylierungs-Assay sollte eine erste funktionelle Studie mit *Arabidopsis thaliana*-Proteinen unter Verwendung von Protein-Microarrays, mit dem in dieser Arbeit erstellten Uniklonsets, erfolgen.

2 Material und Methoden

2.1 Material

2.1.1 Chemikalien

[γ - ³³ P]ATP, 250 μ Ci/ml	Amersham Pharmacia Biotech Europe GmbH, Freiburg
Agarose	Invitrogen GmbH, Karlsruhe
Ammoniumpersulfat	Bio-Rad Laboratories GmbH, München
Ampicillin	Sigma-Aldrich Chemie, München
Attophos (2'-[2-benzothiazoyl]-6'-hydroxybenzothiazol-Phosphat [BBTP])	JBL Scientific, San Luis Obispo, USA
β -Mercaptoethanol	Sigma-Aldrich Chemie, München
Bromphenol Blau	Sigma-Aldrich Chemie, München
Bradford-Reagenz	Bio-Rad, Laboratories GmbH, München
Coomassie Brilliant Blue G/R 250	Pierce Biotechnology, Rockford, USA
D-+-Glukose Monohydrat	Merck, Darmstadt
dATP, dCTP, dGTP, dTTP	Boehringer Mannheim GmbH, Mannheim
Dithiothreitol (DTT)	Serva, Heidelberg
DNA-Längenstandards:	
- <i>Gene ruler</i> DNA 1 Kbp ladder, mix, ready to use	MBI Fermentas, St.Leon-Rot
- <i>Mass ruler</i> DNA, mix, ready to use	MBI Fermentas, St.Leon-Rot
Dimethyl Sulfoxide (DMSO)	Sigma-Aldrich Chemie, München
Ethylendiamintetraessigsäure (EDTA)	Merck, Darmstadt
Ethidiumbromid [10 mg/ml]	Sigma-Aldrich Chemie, München
Glycerin	Merck, Darmstadt
Glycin	Merck, Darmstadt
Guanidin Hydrochlorid	Sigma-Aldrich Chemie, München
Harnstoff	Merck, Darmstadt
Hefeextrakt	Difco Laboratories, Detroit, USA
Imidazol	Sigma-Aldrich Chemie, München
Isopropyl- β -D-Thiogalactopyranosid (IPTG)	MBI Fermentas, St.Leon-Rot
Kanamycin	Sigma-Aldrich Chemie, München
Magermilchpulver	Bio Rad Laboratories GmbH, München
Myelin-assoziiertes basisches Protein (MBP)	Sigma-Aldrich Chemie, München
Molekulargewichtsmarker:	
- <i>Rainbow coloured protein marker</i> RPN 756	Amersham Pharmacia Biotech Europe GmbH, Freiburg
Natriumdodecylsulfat (SDS)	Sigma-Aldrich Chemie, München
NBT/BCIP für Westenblots (Nitroblau-Tetrazoliumchlorid/5-Bromo-4-Chloro-3-Indolyl-Phosphat-Toluidinsalz)	Sigma-Aldrich Chemie, München
Ni-NTA Agaroselösung	Qiagen GmbH, Hilden
Polyethylenglycol (PEG)	Sigma-Aldrich Chemie, München

Protease Inhibitor Cocktail	Roche Diagnostics GmbH, Mannheim
Rotiphorese Gel 30 (29% Acrylamid, 0,8% Bisacrylamid)	Carl Roth GmbH, Karlsruhe
Tetramethylethyldiamin (TEMED)	Invitrogen GmbH, Karlsruhe
Triton X-100	Sigma-Aldrich Chemie, München
Trypton, Bacto	Difco Laboratories, Detroit, USA
Tween 20	Sigma-Aldrich Chemie, München

Anorganische Salze, Säuren, Basen und Alkohole wurden in *pro Analysis* Qualität von Merck, Darmstadt bezogen.

2.1.2 Proteine, Antikörper und Enzyme

Alkalische Phosphatase	Sigma-Aldrich Chemie, München
anti-Maus IgG-Cy3-Konjugat	Jackson Immuno Research Laboratories; Vertrieb durch Dianova, Hamburg
Maus IgG	Sigma-Aldrich Chemie, München
Kaninchen IgG, normal	Sigma-Aldrich Chemie, München
Ratten IgG	Sigma-Aldrich Chemie, München
anti-RGS His ₆ -Antikörper, Maus, monoklonal	Qiagen GmbH, Hilden
anti-Maus IgG-AP-Konjugat	Sigma-Aldrich Chemie, München
anti-PhyB-Antikörper, Maus, Pea 25	FU Berlin, Prof. Dr. John Hughes
Lysozym	Boehringer Mannheim GmbH, Mannheim
GATEWAY TM -LR Clonase Enzym Mix	Invitrogen GmbH, Karlsruhe
Restriktionsenzyme	New England Biolabs GmbH, Schwalbach/ Taunus
Taq-DNA-Polymerase	über die Serviceeinheit des MPI-MG von verschiedenen Herstellern bezogen und zur Verfügung gestellt
T4 DNA-Ligase	New England Biolabs GmbH, Schwalbach/ Taunus
Rinder-Serum-Albumin BSA (Standard für Protein-Microarrays)	Sigma-Aldrich Chemie, München
MAP-Kinasen	IPB Halle, Prof. Dr. Dierk Scheel
Proteinkinase A, Maus	New England Biolabs GmbH, Schwalbach/ Taunus

2.1.3 Synthetische Oligonukleotide

Die verwendeten Oligonukleotide wurden von der Firma Metabion (Martinsried) mit einer eingestellten Konzentration von 100 pmol/ μ l bezogen.

Oligonukleotid	Sequenz (5' → 3')	Anlagerungs- temperatur
PQE65	TGA GCG GAT AAC AAT TTC ACA CAG	60°C
PQE276	GGC AAC CGA GCG TTC TGA AC	60°C
PENTR1a-5'	CGG ATG GCC TTT TTG CGT TTC TAG	63°C
PENTR1a-3'	TGC AAT GTA ACA TCA GAG ATT TTG AGA CAC	63°C

2.1.4 Andere Materialien

Agarplatten, Bio Assay Dish, 22 cm x 22 cm	Nunc GmbH & Co. KG, Wiesbaden
Adhesive Sealing Sheets	Abgene Germany, Hamburg
Whatman-Papiere, 3MM	Whatman GmbH, Göttingen
Deckgläschen, 24 x 60 mm	Carl Roth GmbH, Karlsruhe
Elektroporationsküvetten 0,2 cm	Bio-Rad, Labratories GmbH, München
Eppendorf Gefäße: 0,5, 1,5 und 2,0 ml	Eppendorf AG, Hamburg
Faltenfilter Nr. 595	Schleicher & Schuell, Dassel
FAST TM -Slides	Schleicher & Schuell, Dassel
Filterplatten, 96 well MADV N 65	Millipore GmbH, Eschborn
<i>Imaging plate</i> : BAS-IP SR 0813	Fujifilm, Japan; Vertrieb durch Raytest GmbH, Straubenhardt
Kassette (Hypercassette)	Amersham Pharmacia Biotech Europe GmbH, Freiburg
Kimwipes Handtücher	Kimberly-Clark Corporation, Irving, TX, USA
Kryoröhrchen	Greiner bio-one GmbH, Frickenhausen
96er PCR-Platten	Abgene Germany, Hamburg
96er Mikrotiterplatten	Greiner bio-one GmbH, Frickenhausen
96er Mikrotiterplatten, <i>deep-well</i> -Platten	Qiagen GmbH, Hilden
384er Mikrotiterplatten	Genetix, New Milton, Hampshire, GB
Mikrotiterplatten-Abklebefolie	Bio-Stat Diagnostics, Stockport, GB
Nitrozellulose-Dialyse-Plättchen	Millipore GmbH, Eschborn
Parafilm	Pechiney Plastic Packaging; Neenah, WI, USA
1 ml Polypropylensäulen	Qiagen GmbH, Hilden
PVDF-Filter 22 cm x 22 cm, Immobilon P	Millipore GmbH, Eschborn
Petrischalen, <i>Bio Assay Dish</i>	Nunc GmbH & Co. KG, Wiesbaden
Plastik-Inkubationsboxen	Stewart Company, GB
Reaktionsgefäße (<i>falcon tubes</i>): 15 und 50 ml	Greiner bio-one GmbH, Frickenhausen
Replikatoren, 96 <i>pin</i>	Nunc GmbH & Co. KG, Wiesbaden
Replikatoren, 384 <i>pin</i>	Genetix, New Milton, Hampshire, GB
Saranfolie	Saran, Dow Chemical, Michigan, USA

Schlauchfolie	Carl Roth GmbH, Karlsruhe
Sterilfilter 0,2 µm	Sarstedt, Nürnberg
2.1.5 Labor-Ausrüstung	
Barcode:	
- LYNX 6000	Lynk Data Systems Ltd, Newcastle, GB
Gel-Elektrophorese-Apparaturen:	
- Transformator (Energie Versorgung)	Bio-Rad Laboratories GmbH, München
- Agarose-Gel-Elektrophorese-Apparatur	MBI Fermentas, St.Leon-Rot
- Gel-Dokumentation für Ethidiumbromid gefärbte Agarose-Gele	Herolab GmbH, Wiesloch
- SDS-Polyacrylamidgel-Elektrophorese Apparatur, Hoefer SE 200	Amersham Pharmacia Biotech Europe GmbH, Freiburg
- Blot-Apparatur, TransBlot SD transfer cell	Bio-Rad, Laboratories GmbH, München
Einbändelmaschine f. Mikrotiterplatten:	
- Tippy Pack (<i>binding machine</i>), Spot	Manfred Pütz GmbH, Kerpen
Einschweißmaschine f. Mikrotiterplatten:	
- Lady Pack (<i>shrink wrap</i>)	Pactur S.r.L., Bologna, I
Elektroporator (Gene Pulser II)	Bio-Rad, Laboratories GmbH, München
Kühl-/ Gefrierschränke:	
- Forma (-80°C)	ThermoQuest Analytische Systeme GmbH, Egelsbach
- Freezer (-4°C bis -20°C)	Robert Bosch, Gerlingen-Schillerhöhe
- Kühlschrank (+2°C bis +8°C)	Robert Bosch, Gerlingen-Schillerhöhe
Inkubatoren:	
- Inkubator	Heraeus Instruments GmbH, Hanau
- Inkubator mit Schüttler	New Brunswick Scientific GmbH, Frickenhausen
- Schüttelinkubator Tritramax	Heidolph Instruments Laboratory Products GmbH, Schwabach
Magnetrührer	Heidolph Instruments Laboratory Products GmbH, Schwabach
Mikrotiterplatten-Photometer	Spektramax 250, Molecular Devices, Sunnyvale, CA, USA
PCR-Gerät:	MJ Research, Inc., Watertown, USA
- PTC100	
pH-Meter:	
- MP 220	Mettler-Toledo GmbH, Giessen
Pipetten:	
- verstellbar:	Abimed Analyse-Technik GmbH, Langenfeld
2 µl, 20 µl, 200 µl und 1000 µl	
- 8-Kanal, verstellbar, ohne Ladeadapter:	
10 µl, 100 µl und 300 µl	Eppendorf AG, Hamburg
- 8-Kanal, verstellbar, mit Ladeadapter:	
50 – 1200 µl	Eppendorf AG, Hamburg
Reagenzien-Reservoir	Corning, NY, USA
Roboter:	
- QFill II-Maschine und Zubehör	Genetix, New Milton, Hampshire, GB

- für das „Spotten“ von Klonen auf Filter („Medusa“)	Entwicklung in der Abteilung Lehrach MPI-MG Berlin
-für das „Anordnen“ und das „Picken“ von putativen Proteinexpressionsklonen („Mercury“)	Entwicklung in der Abteilung Lehrach MPI-MG Berlin
-zur Herstellung von Protein-Microarrays, „QArray“	Genetix, New Milton, Hampshire, GB
Scanner und Kamera:	
- Für Filter: LAS-1000 Ch	Fujifilm, Japan; Vertrieb durch Raytest GmbH, Straubenhardt
- für Proteingele: HP ScanJet 5370C	Hewlett-Packard Corp., Palo Alto, CA, USA
- für Protein-Microarray; nach Antikörperinkubation: Array Scanner 428	Affymetrix; Santa Clara, CA, USA
- für Protein-Microarray; nach Phosphorylierung: BioImagers FLA 8000	Fujifilm, Japan; Vertrieb durch Raytest GmbH, Straubenhardt
Schüttler:	
- für Filter (Rocky)	Fröbel Labortechnik, Wasserburg
- Horizontalschüttler	IKA Labortechnik GmbH + CoKG, Staufen
- Rollenmischer RM5	Greiner bio-one GmbH, Frickenhausen
Spektrophotometer:	
- Ultraspec 3100 pro	Amersham Pharmacia Biotech Europe GmbH, Freiburg
Sterile Arbeitsflächen (<i>clean benches</i>)	Heraeus Instruments GmbH, Hanau
Thermomixer Compact	Eppendorf AG, Hamburg
Ultraschall-Homogenisierer:	
- Sonifier W 250	Ultraschall und Labortechnik, G. Heinemann, Schwäbisch Gmund
Vortexer:	
- MS1 Minishaker	IKA-Werke, Staufen
Vakuumpumpe	Millipore GmbH, Schwalbach
Waagen	
- Präzisionswaage: PB 1501	Mettler-Toledo GmbH, Giessen
- Portable Waage: PC 2000	Mettler-Toledo GmbH, Giessen
- Wasserbad	Köttermann GmbH + CokG, Uetze/Härisgen
Zentrifugen:	
- Avanti-J20	Beckmann Coulter, Fullerton, CA, USA
- Avanti-J25	Beckmann Coulter, Fullerton, CA, USA
- Sigma 4K15	Sigma-Aldrich Chemie Laborzentrifugen GmbH, Osterode
- Mikrotiterplatten-Zentrifuge 5810R	Eppendorf AG, Hamburg
- Tischzentrifuge 5415D	Eppendorf AG, Hamburg
2.1.6 Kits	
cDNA Superscript Plasmid Systems	Invitrogen GmbH, Karlsruhe
Plasmid-Päparations-Kits	Qiagen GmbH, Hilden
Qiaex II Gelextraktions-Kit	Qiagen GmbH, Hilden
QIAquick PCR Purification-Kit	Qiagen GmbH, Hilden

2.1.7 Medien, Puffer und Lösungen

2.1.7.1 Medien

2LB Broth

40 g vorgemischter 2 LB Broth
(10 g Trypton, 5 g Hefeextrakt,
10 g NaCl)
- in 1 l ddH₂O Wasser gelöst und durch
Autoklavierung bei 120°C für 20 min
sterilisiert.

2YT Broth

31 g vorgemischter 2YT Broth
(16 g Trypton, 10 g Hefeextrakt,
5 g NaCl)
- in 1 l ddH₂O Wasser gelöst und durch
Autoklavierung bei 120°C für 20 min
sterilisiert.

2YT Broth Agar

46 g vorgemischter 2YT Broth Agar
(16 g Trypton, 10 g Hefeextrakt,
5 g NaCl, 15 g Agar)
- in 1 l ddH₂O gelöst und durch
Autoklavierung bei 120°C für 20 min
sterilisiert.

SB-Medium

Komponente A (4 x SB-Konzentrat):
48 g/l Bacto-Trypton
96 g/l Hefeextrakt
0,8% (v/v) Glycerin
- durch Autoklavierung bei 120°C für 20
min sterilisiert.
Komponente B
(20 x Kaliumphosphatlösung):
17 mM KH₂ PO₄
72 mM K₂HPO₄
- durch Filtration mittels Filter mit
0,2 µm Porengröße sterilisiert.
- Komponente A wird 1:4 und
Komponente B wird 1: 20 in ddH₂O
verdünnt.

Gefriermix (10 x HMFM)

Komponente A:
20 mM Na₃-Citrat
10 mM MgSO₄·7H₂O
44% (v/v) Glycerin
- durch Autoklavierung bei 120°C für 20
min sterilisiert.

	<u>Komponente B:</u> 36 mM K ₂ HPO ₄ 13 mM KH ₂ PO ₄ - durch Filtration mittels Filter mit 0,2 µm Porengröße sterilisiert. - Komponenten A und B wurden nach dem Sterilisieren im Verhältnis 4:1 vereinigt.
SOB	20 g/l Trypton 5 g/l Hefeextrakt 10 mM NaCl 10 mM KCl - durch Autoklavierung bei 120°C für 20 min sterilisiert.
SOC	20 g/l Trypton 5 g/l Hefeextrakt 10 mM NaCl 10 mM KCl 10 mM MgCl ₂ 10 mM MgSO ₄ 20 mM Glukose - durch Autoklavierung ohne Glukose bei 120°C für 20 min sterilisiert - sterilfiltrierte 40%ige (w/v) Glukose nach dem Autoklavieren hinzugeben.

2.1.7.2 Kompetente Zellen

TSS	85% (v/v) LB Broth 10% (v/w) PEG8000 5% (w/v) DMSO 50 mM MgCl ₂ - durch Filtration mittels Filter mit 0,2 µm Porengröße sterilisiert.
------------	---

2.1.7.3 Polymerase-Kettenreaktion (PCR) und Agarosegele

10 x PCR-Puffer	1,5 mM MgCl ₂ 50 mM KCl 15 mM Tris-HCl 35 mM Tris-Base 0,1% (v/v) Tween 20 0,015 mM Cresolrot - durch Autoklavierung ohne Cresolrot bei 120°C für 20 min sterilisiert. - sterilfiltrierte Cresolrot nach dem Autoklavieren hinzugeben
------------------------	--

2.1.7.4 Prozessierung der Koloniefilter

Denaturierungslösung	0,5 M NaOH 1,5 M NaCl
Neutralisierungslösung	1 M Tris-HCl, pH 7,4 1,5 M NaCl
20 x SSC	3 M NaCl 0,1 M Na ₃ -Citrat pH 7,0

2.1.7.5 Antikörper-Bindungstests mit Filtern

PBS	12 mM NaH ₂ PO ₄ 1,76 mM KH ₂ PO ₄ 137 mM NaCl 27 mM KCl pH 7,5
PBST	12 mM NaH ₂ PO ₄ 1,76 mM KH ₂ PO ₄ 137 mM NaCl 27 mM KCl pH 7,5 0,1% (v/v) Tween
TBS	10 mM Tris-HCl, pH 7,5 150 mM NaCl
TBST	10 mM Tris-HCl, pH 7,5 150 mM NaCl 0,1% (v/v) Tween
TBSTT	20 mM Tris-HCl, pH 7,5 0,5 M NaCl 0,1% (v/v) Tween 20 0,5% (v/v) Triton X-100
Alkalische Phosphatasepuffer	1 mM MgCl ₂ 100 mM Tris-HCl, pH 9,5
Attophos-Stammlösung	2,4 M Diethanolamin 5 mM Attophos 0,23 mM MgCl ₂ - pH-Wert mit HCl auf 9,2 einstellen - durch Filtration mittels Filter mit 0,2 µm Porengröße sterilisiert.

2.1.7.6 SDS-Gelelektrophorese und Westernblot

Coomassie-Blau Färbelösung	1,25 g Coomassie Brillant Blue G 250 wurden in 225 ml Ethanol gelöst. 225 ml ddH ₂ O und 50 ml Essigsäure wurden hinzugefügt. Die Mischung wurde für 2 h gerührt und durch einen gefalteten Filter gegeben.
5 x Elektrophoresepuffer	0,125 M Tris-Base 1,25 M Glycin 0,5% (w/v) SDS
Entfärber	20% (v/v) Ethanol 20% (v/v) Essigsäure
Laufpuffer	25 mM Tris 192 mM Glycin 0,1% (w/v) SDS pH 8,3
SDS-Ladepuffer (4 x)	0,2 M Tris-HCl, pH 6,8 8% (w/v) SDS 40% (w/v) Glycerin 0,004% Bromphenol Blau - 0,1 M DTT oder 20% (v/v) β-Mercaptoethanol wurden vor Proteindenaturierung separat dazugeben. - Aliquots wurden bei – 20°C eingefroren.
Blotpuffer	20 mM Tris, pH 8,0 150 mM Glycin 20% (v/v) Methanol
Blocklösung für PVDF-Membran	3% (w/v) Magermilchpulver in TBS 0,1% (v/v) Tween 20

2.1.7.7 Proteinreinigung unter denaturierenden Bedingungen

Lysepuffer	6 M Guanidinium Hydrochlorid 0,1 M NaH ₂ PO ₄ 0,01 M Tris - pH-Wert mit NaOH auf 8,0 einstellen
Waschpuffer	8 M Harnstoff 0,1 M NaH ₂ PO ₄ 0,01 M Tris - pH-Wert mit HCl auf 6,3 einstellen

Elutionspuffer 8 M Harnstoff
0,1 M NaH₂PO₄
0,01 M Tris
- pH-Wert mit HCl auf 4,5 einstellen

2.1.7.8 Proteinreinigung unter nativen Bedingungen

Lysepuffer 300 mM NaCl
50 mM NaH₂PO₄
10 mM Imidazol
- pH-Wert mit NaOH auf 8,0 einstellen

Waschpuffer 300 mM NaCl
50 mM NaH₂PO₄
20 mM Imidazol
- pH-Wert mit NaOH auf 8,0 einstellen

Elutionspuffer 300 mM NaCl
50 mM NaH₂PO₄
250 mM Imidazol
- pH-Wert mit NaOH auf 8,0 einstellen

2.1.7.9 Phosphorylierungs-Assay

Blocklösung für Protein-Microarrays 2% (w/v) BSA in TBST

MAP-Kinase-Puffer 25 mM Tris pH 7,5
1mM EGTA
1mM DTT
5mM MgCl₂
1 mM MnCl₂,
20µM ATP
- durch Filtration mittels Filter mit
0,2 µm Porengröße sterilisiert.

Proteinkinase A-Puffer 50 mM Tris-HCl
10 mM MgCl₂
pH 7,5

2.1.8 Software

Programme zur Microarray-Auswertung:
Aida Array Metrix und Aida Array Compare
VectorNTI 6

Raytest GmbH, Straubenhardt

Informax Inc., Bethesda, USA

Programm zur Filterauswertung:
Visual Grid 2.01
pregap 1.2

Im Haus entwickelt, MPI-MG

www.mrc-lmb.cam.ac.uk/pubseq/staden_home.html

Phred	www.mrc-lmb.cam.ac.uk/pubseq/staden_home.html
Transeq	http://www.hgmp.mrc.ac.uk/Software/EMBOSS/
CrossMatch	http://genome.uc.edu/genome/HelpPages/phred-phrap-polyphred/swat-crossmatch.html
RepeatMasker	http://repeatmasker.genome.washington.edu
d2_cluster	http://www.sanbi.ac.za/bench.html
BLAST	http://www.mips.gsf.de/proj/thal/ftp://ftp.tigr.org/pub/data/a_thaliana/ath1/SEQUENCES

2.1.9 Stämme

E. coli SCS1/pSE111

hsdR17($r_k^- m_k^+$) *recA1* *endA1* *gyrA96* *thi-1* *relA1* *supE44* (Stratagene) enthält Plasmid pSE111 mit *lacI^Q* Repressorgen und *argU* Gen für die seltene ArginintRNA (Büssow *et al.*, 1998)

E. coli DH5 α

Invitrogen GmbH, Karlsruhe
 $F^- \Phi 80lacZ\Delta M15 \Delta(lacZYA-argF)U169$
recA1 *endA1* *hsdR17*(r_k^-, m_k^+) *phoA* *supE44* *thi-1* *gyrA96* *relA1* λ^-)

Das Hilfsplasmid pSE111, welches stabil in die SCS1-Zellen transformiert wurde, wurde von Dr. Eberhard Scherzinger am MPI-MG entwickelt (Abb. 2.1).

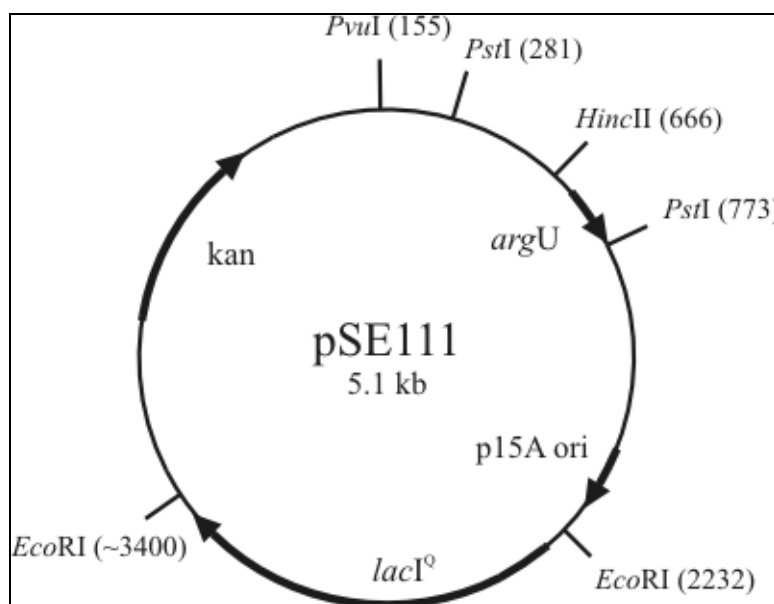


Abb. 2.1: Graphische Darstellung des Hilfs-Plasmids pSE111 mit seinen Besonderheiten.

Das Hilfsplasmid enthält das *argU*-Gen für eine seltene tRNA und das *lacI^Q*-Gen. Durch das *lacI^Q*-Gen wird der *lac*-Repressor konstitutiv exprimiert und somit die vorzeitige Induktion des rekombinanten Proteins verhindert (*leaky expression*). Des weiteren besitzt das Plasmid pSE111 eine Kanamycin-Resistenz (Neomycin-Phospho-Transferase-Gen).

2.1.10 Plasmid-Konstrukte

- | | |
|------------------|---|
| puc 18 | Invitrogen GmbH, Karlsruhe |
| 1. pQE30NASTattB | -pQE30-Derivat von Qiagen GmbH, Hilden
-modifiziert am MPI-MG, Berlin
-Überführung in einen Gateway-kompatiblen Vektor durch Invitrogen GmbH, Karlsruhe
-GenBank- <i>Accession</i> -Nummer: AY386205 |
| 2. pQE30NASTDV | -pQE30-Derivat von Qiagen GmbH, Hilden
-modifiziert am MPI-MG, Berlin
-Überführung in einen Gateway-kompatiblen Vektor durch Invitrogen GmbH, Karlsruhe |
| 3. pENTR1a | Invitrogen GmbH, Karlsruhe |

2.1.10.1 pQE30NASTattB

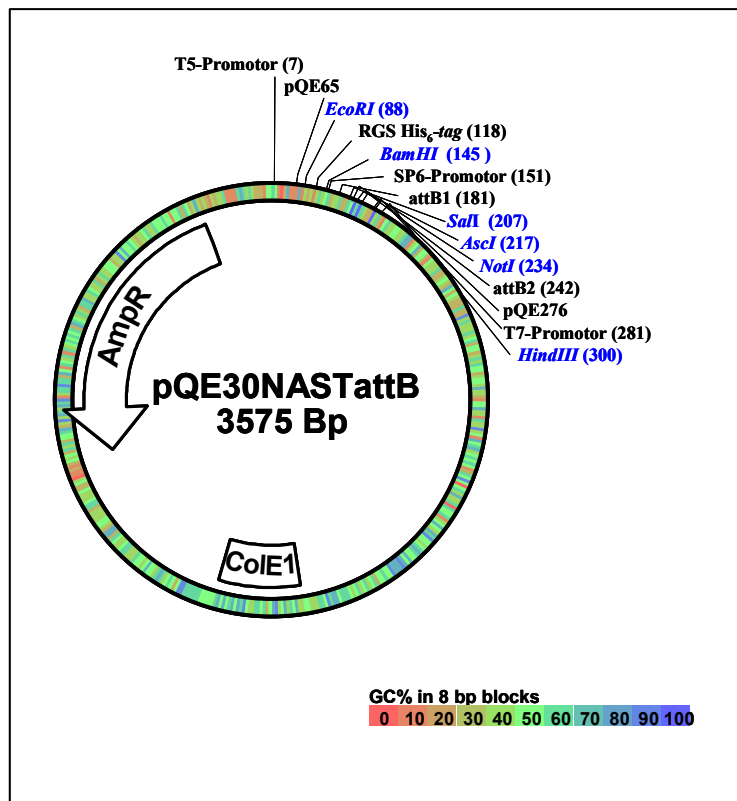


Abb. 2.2: Vektorkarte des *E. coli* Expressionsvektors pQE30NASTattB. Der 3.575 Bp große Vektor enthält eine Ampicillin-Resistenz (AmpR), ein *Origin of Replication* (ColE1), zwei attB-Rekombinationsstellen, ein RGS His₆-Tag, die Bindestellen für die Vektorprimer (pQE65 und pQE276), einen T5- und einen T7-Promotor und eine *multiple cloning site* mit den Schnittstellen *SalI* und *NotI*.

Der *E. coli* Expressionsvektor pQE30NASTattB (Abb.2.2, GenBank-Accession-Nummer: AY386205) enthält einen T5-Promotor (vom T5-Phagen) und zwei *lac*-Operatoren, die der IPTG-induzierten Expression von rekombinanten Proteinen dienen. Aufgrund der Gegenwart der zwei *lac*-Operatoren wird die Effektivität der Bindung des *lacI*-Repressors erhöht, so dass in Abwesenheit des Induktors IPTG die Repression des T5-Promotors sichergestellt wird.

Zwischen dem T5-Promotor und der *multiple cloning site* liegt außerdem noch die Sequenz für ein RGS His₆-Epitop. Dadurch erhalten rekombinant exprimierte Proteine N-terminal einen solchen Tag. Der RGS His₆-Tag ermöglicht sowohl die Reinigung der rekombinanten Proteine über Metall-Ionen-Affinitäts-Chromatographie, als auch die Detektion mittels Antikörper. Es stehen außerdem ein SP6- und ein T7-Promotor zur Verfügung. Über *SalI* und *NotI* in den Vektor klonierte Inserts können dadurch *in vitro* in *sense*-Orientierung mit Hilfe der SP6-RNA-Polymerase bzw. in *anti sense*-Orientierung mit Hilfe der T7-Polymerase transkribiert werden. Zusätzlich besitzt das Plasmid pQE30NASTattB eine

Ampicillin-Resistenz (β -Lactamase-Gen). Weiterhin enthält der Vektor attB-Rekombinationsstellen, diese flankieren das klonierte Gen und ermöglichen einen Transfer dieses Gens in andere Expressionssysteme.

2.1.10.2 pQE30NASTDV

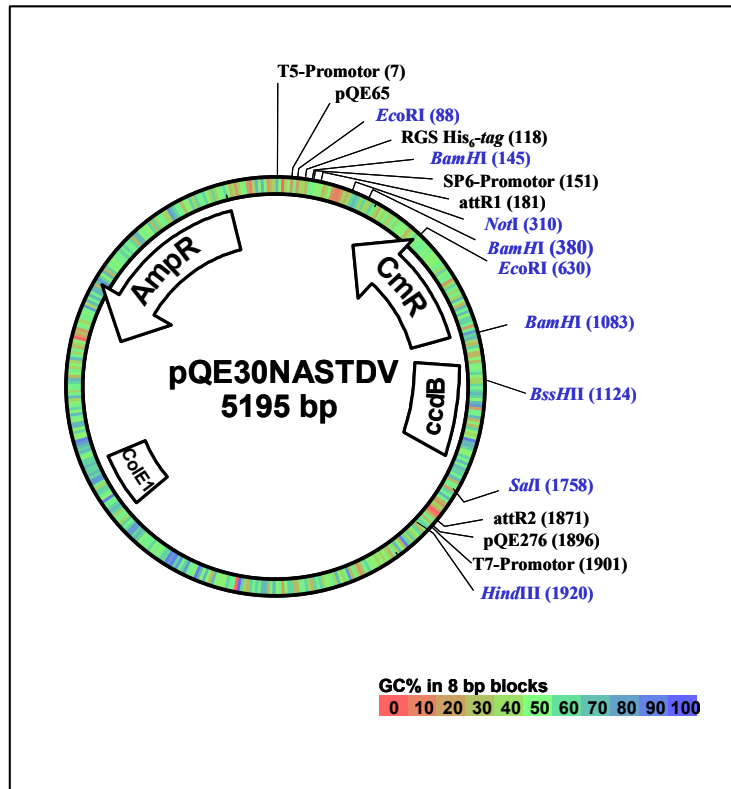


Abb. 2.3: Vektorkarte des *E. coli* Destinationsvektors pQE30NASTDV. Der 5.195 Bp große Vektor enthält eine Ampicillin-Resistenz (AmpR), eine Chloramphenicol-Resistenz (CmR), einen *Origin of Replication* (ColE1), zwei attR-Rekombinationsstellen, ein RGS His₆-Tag, die Bindestellen für die Vektorprimer (pQE65 und pQE276), das *ccdB*-Gen und eine *multiple cloning site* mit den Schnittstellen *SaII* und *NotI*.

Der *E. coli* Destinationsvektor pQE30NASTDV (Abb. 2.3) ist ähnlich aufgebaut wie der Vektor pQE30NAStattB. Auch der Destinationsvektor enthält einen T5-Promotor (vom T5-Phagen) und zwei lac-Operatoren. Des weiteren enthält der pQE30NASTDV ein RGS His₆-Epitop, ein SP6-, ein T7-Promotor und eine *multiple cloning site* mit den Schnittstellen *SaII* und *NotI*. Zudem enthält der Vektor die attR-Rekombinationsstellen. Diese flankieren ebenso das Gen und ermöglichen einen Transfer des Gens in andere Expressionssysteme. Zwischen den beiden attR-Rekombinationsstellen und der *multiple cloning site* liegt das *ccdB*-Gen, welches bei der Rekombination gegen das Insert ausgetauscht wird. Bei einem nichterfolgten Austausch kommt es aufgrund der Suizid-Funktion des *ccdB*-Gens zu einer Negativselektion, d.h. alle Klone, die nach der Rekombination noch ein *ccdB*-Gen enthalten,

sterben. Zusätzlich zu der Ampicillin-Resistenz enthält dieser Vektor eine Chloramphenicol-Resistenz.

2.1.10.3 pENTR1a

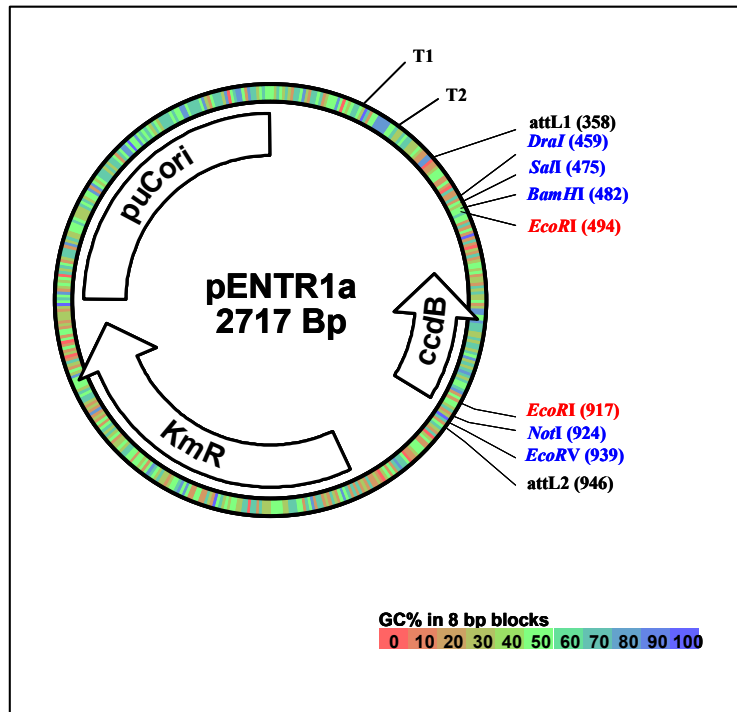


Abb. 2.4: Vektorkarte des *E. coli* Entryvektors pENTR1a. Der 2.717 Bp große Vektor enthält eine Kanamycinresistenz (KmR), ein *Origin of Replication* (puC ori), zwei attL-Rekombinationsstellen, zwei rrn B-Operons: T1 und T2, ein *ccdB*-Gen und eine *multiple cloning site* mit den Schnittstellen *SalI* und *NotI*.

Der *E. coli* Entryvektor pENTR1a (Abb. 2.4) enthält zwei rrn B-Operons (*transcription termination sequences*) T1 und T2; diese verhindern eine eventuelle Expression aufgrund vektorkodierender Promotoren. Des Weiteren stehen die attL1- und attL2-Rekombinationsstellen zur Verfügung. Mit Hilfe dieser beider Rekombinationsstellen ist unter Verwendung des Destinationsvektors die Rekombination des Gens durchführbar und somit die Erstellung eines Expressionsklons.

Zwischen den beiden attL-Rekombinationsstellen und der *multiple cloning site* liegt hier, wie auch bei dem pQE30NASTDV, das *ccdB*-Gen, welches bei einem nichterfolgten Austausch eine Negativselektion ermöglicht. Als einen weiteren Selektionsfaktor besitzt der Entryklon eine Kanamycin-Resistenz (Neomycin-Phospho-Transferase-Gen). Hinzu kommt ein puCori für die Replikation in *E. coli*.

2.2 Methoden

2.2.1 Herstellung von zwei *Arabidopsis* cDNA-Bibliotheken

Die Abkürzung cDNA steht für *complementary* DNA, welche entsteht, wenn man mit Hilfe der Reversen Transkriptase (RT) *messenger* RNA (mRNA) kopiert. Im Wesentlichen werden cDNA-Moleküle als Ausgangsmaterial für die PCR und für die Herstellung von cDNA-Bibliotheken verwendet.

In der vorliegenden Arbeit wurden ausgehend von zwei *Arabidopsis*-Geweben zwei cDNA-Bibliotheken hergestellt: Pistill und Infloreszenz. In beiden Fällen wurden die *Arabidopsis*-Pflanzen unter Standardbedingungen (14 h/10 h [Tag/Nacht] Rhythmus bei 22°C/18°C) im Gewächshaus angezogen.

Zur Erstellung der Pistill cDNA-Bibliothek ist die Wahl auf zwei verschiedene *Arabidopsis*-Mutanten gefallen, bei denen das pi (pata) Gen bzw. die beiden Gene ap2-5 und sap ausgeschaltet wurden. Diese Mutanten wurden ausgewählt, da bei beiden eine verstärkte Bildung des Fruchtknotens vorliegt, welcher ein Forschungsobjekt unserer Kooperationspartner Dr. Richard Immink und Prof. Dr. Gerco Angenent (Plant Research International, Wageningen, Niederlande) ist. Pi ist ein B-Typ MADS-Box-Protein und führt zu Mutationen in den Wirteln 3 und 4 der Blüte; dabei werden Staubbeutel/-blätter in Pistill umgewandelt. Auch in den ap2-5/sap-Mutanten werden vermehrt Pistills in der Blüte gebildet (Byzova *et al.*, 1999).

Zur Erstellung der Infloreszenz cDNA-Bibliothek wurden die apikalen Bereiche jungen Infloreszenzmeristems (ca. 1 Woche nach der Schosserbildung) des Wildtyps *Arabidopsis thaliana* Columbia Ökotyps Col-1 geerntet.

Die mRNA-Moleküle beider Gewebe wurden von den oben erwähnten niederländischen Kooperationspartnern isoliert. Zu diesem Zwecke erfolgte zunächst die Gesamt-RNA Gewinnung nach (Verwoerd *et al.*, 1989). Aus der Gesamt-RNA wurde mRNA unter Verwendung des mRNA Aufreinigungs-Kits von Amersham (27-9258-02) isoliert.

2.2.1.1 cDNA-Synthese und Größenfraktionierung

Die cDNA-Synthese wurde in beiden Fällen mit Hilfe des cDNA Superscript Plasmid Systems für cDNA-Synthese der Firma Life Technologies GmbH nach den Angaben des Herstellers durchgeführt. Der Schlüssel zu dieser Methode ist das Enzym Reverse Transkriptase (RT), welches ein DNA-Polynukleotid komplementär zu einem bereits vorhandenen RNA-Molekül synthetisiert. Die Superscript II RT stammt aus dem *Moloney*

murine leukemia virus (MMLV-RT) und ist eine RNA-abhängige DNA-Polymerase. Aufgrund einer zusätzlichen Modifikation besitzt die Superskript II RT, im Gegensatz zu anderen RTs, keine RNase H-Aktivität. Dadurch kann sowohl die Ausbeute der cDNA-Synthese als auch die Länge der Transkripte erhöht werden.

Die Synthese des cDNA-Erststrangs begann mit der Anlagerung der Oligo(dT)-*NotI*-Primer am Poly-A⁺-Schwanz und setzte sich in Richtung 5'-Ende der mRNA fort. Diese Primer dienten einerseits als Ausgangspunkt für die Superskript II RT, andererseits befand sich am 5'-Ende der Primer eine Restriktions-Schnittstelle (in der vorliegenden Arbeit: *NotI*), die für die anschließende Klonierung der cDNA elementar war. Nach Beendigung der Erststrangsynthese wurde zunächst eine Analyse der Erststrangsynthese durchgeführt. Dazu wurden 1/6 der Erststrang-Reaktion einer gelelektrophoretischen Auftrennung unterzogen (Kap. 2.2.4.4). Anschließend erfolgte die Zweitstrangsynthese, hier wurden die RNA-Bestandteile des RNA-DNA-Hybrides durch DNA ersetzt. Diese Zweitstrangsynthese erfolgte nach der Methode von Gubler und Hoffman (1983) und wurde durch die Enzyme *E. coli* RNase H, *E. coli* DNA-Polymerase I und *E. coli* DNA-Ligase ermöglicht. Durch *E. coli* RNase H-Zugabe kam es zu Brüchen und Lücken in den RNA-Molekülen des Hybrides. Kurze RNA-Stücke, die nicht von dem Enzym verdaut wurden, verblieben an dem cDNA-Erststrang und dienten als Primer für die *E. coli* DNA-Polymerase I. Die Polymerase füllte die entstandenen Lücken mit den komplementären Nukleotiden auf. Am Ende der Zweitstrangsynthese wurden die benachbarten Fragmente, die bis jetzt noch durch fehlende Bindungen zwischen den benachbarten Nukleotiden getrennt sind, mittels der *E. coli* DNA-Ligase verbunden.

Nach Beendigung der Zweitstrangsynthese wurden sogenannte glatte Enden (*blunt ends*) generiert. Hierbei wurden die überhängenden Enden (*sticky ends*) der DNA-Moleküle durch die Zugabe der T4 Polymerase mit den entsprechenden Nukleotiden aufgefüllt. An die glatten Enden wurden dann *SalI*-Adaptoren mittels der T4 DNA-Ligase ligiert. Nach diesem Schritt enthielt nun jedes der generierten cDNA-Moleküle an beiden Enden einen *SalI*-Adaptor. Um eine gerichtete Klonierung zu ermöglichen, folgte nach der Adaptorligation ein *NotI*-Verdau am 5'-Ende des oligo(dT)-*NotI*-Primers. Am Ende standen doppelsträngige cDNA-Moleküle mit 5'-*SalI*- und 3'-*NotI*-Überhängen zur Verfügung, die in die entsprechenden Vektoren ligiert werden konnten.

Vor der Vektorligation wurden die cDNA-Moleküle einer Größenfraktionierung mittels Sepharose-Säulen-Chromatographie unterzogen; einerseits, um nicht-ligierte *SalI*-

Adaptoren, welche in einem hohen Überschuss vorlagen und Primer-Adapter-Fragmente zu eliminieren und andererseits, um cDNA-Moleküle kleiner als 500 Bp auszuschließen.

Dazu wurde von der Sepharose-Säule zunächst die überschüssige Flüssigkeit (20% Ethanol) entfernt, indem die Säule 4 mal mit 0,8 ml TEN-Puffer gewaschen wurde. Anschließend wurde die in 100 µl TEN-Puffer gelöste synthetisierte cDNA auf die Säule gegeben und das Eluat in einem Reaktionsgefäß (Nr. 1) aufgefangen. Danach wurde mit 100 µl TEN-Puffer eluiert und das Eluat im sterilen Reaktionsgefäß (Nr. 2) aufgefangen. Darauf folgend wurden wieder 100 µl TEN-Puffer auf die Säule gegeben, wobei dann je ein Tropfen (~ 35 µl) in jeweils einem sterilem Reaktionsgefäß aufgefangen wurde. Es wurden immer wieder 100 µl TEN-Puffer auf die Säule gegeben, bis die Reaktionsgefäße (Nr. 3 – 20) mit jeweils einem Tropfen befüllt waren. Die Fraktionsvolumina aller Reaktionsgefäße (Nr. 1 – 20) wurden ermittelt. Diese dienten zur Ermittlung des Gesamtvolumens, welches zur Ermittlung des *cutoff* Wertes (550 µl) benötigt wurden. Nach der Bestimmung des *cutoff* Wertes wurden vier bis fünf Fraktionen (entsprechend den Angaben des Herstellers) unterhalb des *cutoff* Wertes vereinigt (siehe Tab. 3.1).

2.2.1.2 Fällung der cDNA

Zur Konzentrierung bzw. Entsalzung der vereinigten DNAs wurde eine Fällung mit 1/10 Vol. (v/v) Natriumacetat (3 M; pH 4,8) und 1,3 µl Glykogen (20 mg/ml) durchgeführt. Die Präzipitation erfolgte durch die Zugabe von 2,5 Vol. eiskaltem 99%igen Ethanol (v/v). Nach vorsichtigem Durchmischen wurde die DNA über Nacht (ÜN) bei – 20°C gefällt. Nach der Zentrifugation (20 min, 14.000 x g, Raumtemperatur) wurde das DNA-Pellet getrocknet und in 13 µl DEPC (Diethylpyrocarbonat) behandeltem Wasser resuspendiert. Für Testligationen wurden 3 µl eingesetzt.

2.2.1.3 Vektorpräparationen

Die Vektoren wurden mittels Restriktionsendonukleasen für die Ligation mit den cDNA-Molekülen vorbereitet. Restriktionsendonukleasen Typ II erkennen für sie spezifische kurze DNA-Sequenzmotive von 4 bis ungefähr 10 Bp (meist 6). Der DNA-Strang wird innerhalb des Erkennungsmotivs oder in einem Abstand von einigen Bp hydrolytisch gespalten, wobei die Schnittstelle je nach Spezifität der Endonuklease glatt (*blunt*) oder versetzt (*sticky*) sein kann.

Zur Herstellung der Infloreszenz cDNA-Bibliothek wurden 30 µg des pQE30NASTattB-Vektors (GenBank-*Accession*-Nummer: AY386205) verwendet, welcher zusätzlich ein 1.665 Bp großes, über *SalI/NotI* kloniertes Insert, besass. Dieses Insert ist ein aus

Arabidopsis thaliana stammendes Gen, welches für das Phytochrom B (*phyB*) kodiert. Der Vektor wurde mit 300 U *SalI*, 150 U *NotI*, 75 U *BglII*, 1 x *SalI*-Puffer in Gegenwart von BSA verdaut. Mit den Enzymen *SalI* und *NotI* wurde geschnitten, um die Ligation (Kap. 2.2.1.4) zu ermöglichen. *BglII* dagegen wurde verwendet, um innerhalb des Insert (*phyB*) zu schneiden. Die Restriktion mittels *BglII* minimierte eine evtl. Religation bzw. den evtl. Anteil an ungeschnittenem Vektor. Alle Enzymmengen wurde in drei Teile geteilt (3 x 100 U *SalI*, 3 x 50 U *NotI*, 3 x 25 U *BglII*), welche nach verschiedenen Inkubationszeiten bei 37°C dazugegeben wurden. Ausgegangen wurde von 1/3 *SalI* für 3 h, gefolgt von 1/3 *SalI* ÜN. Am nächsten Morgen folgte die Zugabe des letzten Drittels *SalI* für 3 h. Anschließend folgte die Inkubation mit jeweils 1/3 *NotI* und 1/3 *BglII* ebenfalls in den Schritten 2 x 3 h und 1 x ÜN.

Zur Herstellung der Pistill cDNA-Bibliothek betrug die Ausgangsmenge des Vektors pENTR1a (<http://www.invitrogen.com/>) ebenfalls 30 µg. In diesem Fall wurde mit den Enzymen *SalI* (100 U) und *NotI* (60 U) innerhalb der *multiple cloning site* des Vektors geschnitten. Bei diesem Ansatz wurden die Enzymmengen in zwei Teile geteilt. Zuerst wurde auch hier wieder mit *SalI* restringiert. Dazu wurden 80 U hinzugegeben und für 3 h bei 37° inkubiert, anschließend wurde die restliche Menge *SalI* (20 U) gleichzeitig mit dem ersten Anteil von *NotI* (40 U) zugegeben und ÜN inkubiert. Am nächsten Morgen folgte der letzte dreistündige Inkubationsschritt mit *NotI* (20 U). Anschließend wurden die jeweilig geschnittenen Vektoren mittels Gelextraktion (Kap. 2.2.4.5) isoliert.

2.2.1.4 Ligation der cDNAs mit den Vektoren

Die cDNA-Moleküle wurden mit Hilfe des Enzyms T4 DNA-Ligase über die Schnittstellen *SalI* und *NotI* in den Vektor ligiert. Für die Ligationsansätze (siehe unten) wurden der geschnittene und aufgereinigte Vektor und die entsprechenden cDNA-Fragmente vereinigt. Zur Ermittlung der Religationsfrequenz des Vektors wurden für diese Religationsansätze anstelle der cDNAs die entsprechende Menge ddH₂O hinzugegeben. Alle Ansätze wurden für 45 s bei 70°C erhitzt. Anschließend wurden die Ansätze auf Eis gestellt und dort erfolgte die Zugabe der T4 DNA-Ligase und des dazugehörigen Puffers. Die Ligation wurde ÜN bei 16°C durchgeführt. Am nächsten Tag wurde die cDNA in 3 Vol. eiskaltem 99%igem Ethanol und 1/10 Vol. 7,5 M Natriumacetat, pH 4,8 für 2 h bei - 20°C gefällt und anschließend für 30 min zentrifugiert (20.800 x g, 4°C). Das Pellet wurde mit 70%igem Ethanol gewaschen, für 10 min (20.800 x g, 4°C) zentrifugiert, dann getrocknet und in 10 µl ddH₂O resuspendiert. Vor der Elektroporation (Kap. 2.2.1.7) wurden die Ligationsansätze

halbiert (5 μ l) und auf Nitrozellulose-Plättchen gegen ddH₂O dialysiert. Nach einer 30 minütigen Dialyse wurden die Aliquots in neue sterile Reaktionsgefäße überführt.

Ligationsansätze (10 μ l):

- Infloreszenz cDNA:
 - 50 ng Vektor-DNA (pQE30NASTattB)
 - 3 μ l der vereinigten Infloreszenz cDNA-Fractionen
 - 1 x T4 DNA-Ligasepuffer
 - 200 U T4 DNA-Ligase

- Pistill-cDNA:
 - 50 ng Vektor-DNA (pENTR1a)
 - 1 μ l der vereinigten Pistill cDNA-Fractionen
 - 1 x T4 DNA-Ligasepuffer
 - 200 U T4 DNA-Ligase

In beiden Fällen wurden 3 Ligationen durchgeführt.

2.2.1.5 Herstellung elektrokompenter *E. coli* SCS1/pSE111-Zellen

50 ml LB-Medium (mit 30 μ g/ml Kanamycin) wurden mit einer Einzelkolonie *E. coli* SCS1/pSE111 inokuliert und ÜN bei 37°C geschüttelt. 1% (v/v) der Vorkultur wurden in 2 x 2,4 l SOB-Medium überführt und bis zu einer OD₆₀₀ von 0,4 – 0,5 bei 37°C inkubiert. Die Bakterien, die sich jetzt in der logarithmischen Wachstumsphase befanden, wurden 30 min auf Eis unter regelmäßigem Schütteln gekühlt und anschließend 10 min bei 1.850 x g und 4°C zentrifugiert. Das Zellpellet wurde vorsichtig in je 5 - 10 ml gekühltem 10%igen (v/v) autoklavierten Glycerin resuspendiert und auf 400 ml mit gekühltem 10%igem (v/v) autoklavierten Glycerin aufgefüllt. Nach erneuter Zentrifugation (10 min, 1.850 x g, 4°C) wurden die Pellets wie oben resuspendiert, je zwei Pellets wurden vereinigt und mit je 100 ml gekühltem 10%igem (v/v) autoklavierten Glycerin aufgefüllt und erneut zentrifugiert (10 min, 1.850 x g, 4°C). Anschließend wurden die Pellets wie oben resuspendiert, vereinigt und in 50 ml Reaktionsgefäße überführt und mit gekühltem 10%igem (v/v) autoklavierten Glycerin auf 50 ml aufgefüllt und zentrifugiert (10 min, 1.350 x g, 4°C). Der Überstand wurde abgossen, die Pellets in ca. 2 - 5 ml des verbleibenden Glycerins resuspendiert und in je ~ 70 μ l in 0,5 ml Reaktionsgefäße aliquotiert. Die Aliquots wurden in flüssigem Stickstoff schockgefroren und bei - 80°C bis zur weiteren Verwendung gelagert.

2.2.1.6 Ermittlung der Transformationskompetenz der kompetenten Zellen

Die Kompetenz der hergestellten kompetenten Zellen wurde mittels Transformation eines Kontrollplasmids (pUC 18; Konzentration: 1 pg/Transformationsansatz) berechnet.

2.2.1.7 Transformation von *E. coli* mittels Elektroporation

Für die Transformation wurden 30 µl der elektrokompenten Zellen (Kap. 2.2.1.5) auf Eis aufgetaut und mit 5 µl Ligationsgemisch in einem Reaktionsgefäß auf Eis vermischt. Im Falle der Infloreszenz wurden *E. coli* SCS1/pSE111-Zellen verwendet, während für Pistill *E. coli* DH5α-Zellen eingesetzt wurden. Anschließend wurden die entsprechenden Ansätze in gekühlte Elektroporationsküvetten überführt und, in beiden Fällen einem elektrischen Puls von 1,67 kV, 25 µF und 200 Ω ausgesetzt. Zur Regeneration der Zellen wurden diese mit 1 ml SOC-Medium versetzt und bei 37°C für 45 min schüttelnd (~ 650 upm) inkubiert. Die Transformationsansätze wurden auf Anzuchtplatten (22 cm x 22 cm; für Infloreszenz (*E. coli* SCS1/pSE111-Zellen): 2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin; für Pistill (*E. coli* DH5α-Zellen): 2YT, 50 µg/ml Kanamycin) aufgebracht und ÜN bei 37°C inkubiert. Die Pistill-Klone wurden anschließend vereinigt. Ein Hälfte der vereinigten Klone wurde mit Glycerin versetzt (Endkonzentration: 20% (v/v)), bei – 80°C gelagert und die andere Hälfte wurde zur Plasmid-Maxipräparation (Kap. 2.2.4.1) verwendet. Die Infloreszenz-Klone wurden gepickt, dazu sollten die Klone in einer Dichte von 1.000 - 2.000 Kolonien/Agarselektionsplatte vorliegen.

2.2.1.8 Anordnung und Replikation einer cDNA-Bibliothek

Die nach der Transformation der Infloreszenz cDNA-Bibliothek erhaltenen Klone wurden mit Hilfe des Pick-Roboters “Mercury” in beschriftete und mediumbefüllte 384er Mikrotiterselektionsplatten transferiert und parallel kultiviert.

Befüllt wurden die Mikrotiterselektionsplatten mit Gefriermedium (2YT, 1 x HMF, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) mit Hilfe des Qfill-Gerätes.

Die Beschriftung erfolgte mit dem Barcoder LINX 6000 und setzte sich wie folgt zusammen:

- Bibliotheks-Nummer (erteilt vom RZPD, Deutsches Ressourcenzentrum für Genomforschung GmbH)
- Name der Bibliothek (im Falle der Infloreszenz-Bibliothek: ATM1)
- aktuelles Datum
- Barcode
- Kopieanzahl (Q00 steht für die Originalplatten, Q01 für die Masterplatten, Q02 für die Kopieplatten 1 usw.)
- Plattennummer

Der Roboterkopf des Pick-Roboters "Mercury" ist aus einem herausnehmbaren 96er *pin-gadget* (*pin*-Ø: 0,5 mm, *pin*-Abstand: 9 mm) und einer CCD (*charge coupled devies*)-Kamera zusammengesetzt. Der Roboterkopf kann durch Linearachsen in drei Richtungen (x, y und z) bewegt werden. Die Ortsauflösung des Achsensystems beträgt 5 µm und erreicht eine Schnelligkeit von 2 m/s. Das *gadget* besteht aus einem Kunststoffblock, in dem die *pins* aus Edelstahl positioniert sind und jeder *pin* separat auszufahren ist. Es können 5.000 Klone/h von den Anzuchtplatten in die Mikrotiterselektionsplatten gepickt werden. Mit Hilfe der CCD-Kamera können die Klone mit einem Durchmesser von ungefähr 1 mm ermittelt werden. Zur Optimierung der Auflösung und der Pickeffizienz unterteilt die Kamera die Anzuchtplatten (22 cm x 22 cm) in 54 Felder. Jedes dieser Felder wird automatisch kalibriert, so dass z.B. Hintergrundvariationen für jedes Feld separat einbezogen werden. Vor dem eigentlichen Picken wurden der Innenraum des Roboters und das *gadget* 20 min mit UV-Licht bestrahlt. Zudem wurde die Bürste des Ethanolbades mit 3%igem H₂O₂ (v/v) sterilisiert. Anschließend wurde das *gadget* in den Roboterkopf eingesetzt, die Bürste mit 2 l ddH₂O gespült, mit 80%igem Ethanol befüllt und in der entsprechenden Halterung befestigt. Dieses Ethanolbad dient zur Sterilisierung des *gadgets* zwischen den Inokulationsschritten.

Nachdem die Kulturen gepickt, in den 384er Mikrotiterselektionsplatten angeimpft und ÜN bei 37°C inkubiert wurden, wurden die Originalplatten repliziert.

Von jeder Platte wurden eine Masterplatte und zwei bis drei Kopien angelegt. Die Replikation erfolgte per Hand mittels 384er *pin* Einweg-Replikatoren. Nach Beendigung der ÜN-Inkubation bei 37°C wurden die Platten bei - 80°C gelagert.

2.2.2 Herstellung eines Proteinexpressionssubsets der cDNA-Expressionsbibliothek

2.2.2.1 Herstellen von Hochdichte-Proteinfiltern der cDNA-Expressionsbibliothek

Alle Klone der geordneten Infloreszenz cDNA-Bibliothek wurden mit Hilfe des Spot-Roboters “Medusa” auf PVDF-Membranen übertragen. Hierzu wurden die Klone in mediumbefüllte 384er Mikrotiterselektionsplatten (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) frisch repliziert und ÜN bei 37°C inkubiert.

Die Übertragung der Kulturen aus der Mikrotiterselektionsplatte auf die Membran erfolgte mit einem 384er *pin-gadget* (*pin*-Ø: 0,45 mm, *pin*-Abstand: 4,5 mm). Die *pins* dieses *gadgets* sind nicht separat beweglich und werden gleichzeitig in den Mikrotiterplatten inokuliert. Die Klone jeder Plattenkoordinate wurden auf jeweils zwei verschiedene Positionen der PVDF-Membran gespottet. Auf diese Weise ist später jeder Klon durch zwei Spots repräsentiert. Der Abstand zwischen den einzelnen Spots auf der Membran beträgt 0,9 mm.

Auch hier wurden der Roboter, das *gadget* und die Bürste des Ethanolbades, wie in Kap. 2.2.1.8 beschrieben, vor dem Spotten sterilisiert.

Zur Vorbereitung der Filter wurden pro Membran zwei Whatman-Papiere zugeschnitten (22 cm x 22 cm) und in 1 x PBS geschwenkt. Die mit 99%igem Ethanol und ddH₂O aktivierte Membran wurde ebenfalls in 1 x PBS überführt. Für das eigentliche Spot-Verfahren wurden je zwei getränkte Whatman-Papiere und eine Membran auf eine Plexiglasplatte luftblasenfrei gelegt, wobei die Whatman-Papiere zur Vermeidung der Austrocknung der Membran dienen. Die Plexiglasplatten wurden nebeneinander in den Innenraum des Roboters gelegt. Die 384er Mikrotiterplatten der Infloreszenz cDNA-Bibliothek wurden in eine spezielle Halterung geschoben. Beim Start der Spot-Routine nimmt der Roboter die Platten einzeln aus der Halterung und überträgt die Klone einer jeden Platte mit Hilfe des *gadgets* auf die Membranen. Das Spotten der Klone von einzelnen Plattenkoordinaten auf die Membranen erfolgte in einem definierten Muster (5 x 5-Pattern, Abb. 2.5). Zur Erleichterung der Identifizierung und der späteren Auswertung wurde in das Zentrum eines jeden Patterns ein Tintenpunkt (*guide dot*) gespottet. Dazu wurde am Ende des Spottens eine 384er Mikrotiterplatte mit Tinte gefüllt, in die Halterung geschoben und die *guide dots* auf die entsprechende Position überführt. Die Filter wurden in sechs Felder unterteilt und in jedes dieser Felder konnten jeweils die Klone von zwölf Mikrotiterplatten transferiert werden. Somit konnten pro Filter (22 cm x 22 cm) 55.296 Spots in Duplikaten von 27.648 individuelle Klone gespottet werden.

Feld 2					Feld 6					Feld 3				
8	32	68	8	14	12	36	72	12	18	9	33	69	9	15
2	56	20	62	68	6	60	24	66	72	3	57	21	63	69
50	32	●	38	56	54	36	●	42	60	51	33	●	39	57
38	44	62	14	26	42	48	66	18	30	39	45	63	15	27
2	26	50	20	44	6	30	54	24	48	3	27	51	21	45
10	34	70	10	16	7	31	67	7	13	11	35	71	11	17
4	58	22	64	70	1	55	19	61	67	5	59	23	65	71
52	34	●	40	58	49	31	●	37	55	53	35	●	41	59
40	46	64	16	28	37	43	61	13	25	41	47	65	17	29
4	28	52	22	46	1	25	49	19	43	5	29	53	23	47
Feld 4					Feld 1					Feld 5				

Abb. 2.5: Anordnung der Mikrotiterselektionsplatten auf einem 22 cm x 22 cm großen Filter. Die Filter sind in sechs Felder unterteilt, wobei in jedes Feld die entsprechend nummerierten Platten als Duplikate in einem 5 x 5-Pattern gespottet wurden, d.h. pro Feld 12 Platten. Insgesamt konnten auf jeden Filter 72 Platten angeordnet werden. In der Mitte eines jeden Feldes befindet sich der *guide dot*.

2.2.2.2 Proteinexpression und Prozessierung der Filter

Nach dem Spotten wurden die Filter mit der bespotteten Seite nach oben auf Anzuchtplatten (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) luftblasenfrei aufgelegt und ÜN bei 37°C inkubiert. Zur Induktion der Proteinexpression wurden die Filter von den Anzuchtplatten luftblasenfrei auf Induktionsplatten (2YT, 1 mM IPTG, 100 µg/ml Ampicillin, 25 µg/ml Kanamycin) überführt und für 4 h bei 37°C inkubiert.

Zum Prozessieren der Filter unter denaturierenden Bedingungen wurden die Filter mit dem anhaftenden Zellmaterial von den Induktionsplatten abgenommen und nacheinander auf unterschiedlich getränkte Whatman-Papiere überführt. Die Whatman-Papiere waren jeweils mit den folgenden Lösungen getränkt:

10 min	Denaturierungslösung
2 x 5 min	Neutralisierungslösung
15 min	2 x SSC

Danach wurden die Filter bei Raumtemperatur getrocknet und gelagert oder direkt mit dem Detektionsantikörper hybridisiert.

2.2.2.3 Immunofärbung und Auswertung der Proteinfiler

Von den frisch induzierten und prozessierten Proteinfiltern wurden die Zellbestandteile der Bakterien-Kolonien mit in TBSTT getränkten Papiertüchern vorsichtig abgewischt. Anschließend wurden die Filter wie folgt gewaschen:

2 x 20 min TBSTT

1 x 10 min TBS

Die Proteinfiler wurden dann in Blocklösung für 2 h bei Raumtemperatur geblockt. Danach wurden die Filter mit dem ersten Antikörper (anti-RGS His₆, 1:2.000 verdünnt in Blocklösung) ÜN inkubiert und anschließend gewaschen:

3 x 20 min TBST

Als Zweitantikörper wurde ein enzymgekoppelter Antikörper (anti-Maus-AP-Konjugat; 1:5.000 verdünnt in Blocklösung) eingesetzt. Der Filter wurde für 1 h inkubiert und dann wiederum gewaschen:

2 x 20 min TBST

1 x 20 min TBS

Den Waschschritten folgte eine zehnminütige Inkubation mit Alkalischen-Phosphatasepuffer (im Dunkeln). Zum Nachweis für die Alkalische-Phosphatase wurden die Filter mit 0,125 mM Attophos für 5 min inkubiert. Attophos ist ein fluoreszierendes Substrat der alkalischen Phosphatase und dient zur Detektion der RGS His₆ getagten Proteine. Anschließend wurden die Filter unter UV-Licht (bei 477 nm) mit der CCD Fluoreszenz-Video-Dokumentation photographiert.

Das Fluoreszenzbild der Proteinfiler wurde mit Hilfe des Programms „Visual Grid 2.01“ ausgewertet. Das Programm ermöglicht zum einen, dass die Positionen der verschiedenen Klone zugeordnet und ausgewertet werden können und zum anderen die Erstellung einer Liste von potentiellen Proteinexpressionsklonen. Für die Zuordnung wurde auf die Filter (bestehend aus sechs Feldern, Abb. 2.5) ein *grid* gelegt, ausgehend von den vier Eck-*guide dots*. Dieses *grid* besteht aus 48 x 48 Unterfeldern. Jedes Unterfeld besteht aus 25 Spots, welche aus dem 5 x 5-Pattern resultieren. In der Mitte eines jeden Unterfeldes befindet sich der *guide dot*. Somit enthält jedes Unterfeld zwölf verschiedene Klone, die in Duplikaten gespottet wurden (Abb. 2.6). Ein Klon wurde als anti-RGS His₆ positiv und damit als potentieller Proteinexpressionsklon gewertet, wenn die beiden Spots eines Klons innerhalb eines Unterfeldes als leuchtende Punkte visuell zu erkennen waren. Anschließend erstellte das Programm mit den genannten Daten eine Liste mit den Plattenkoordinaten, in denen sich potentielle Proteinexpressionsklone befanden.

2	6	12	2	3
1	10	4	11	12
9	6	•	7	10
7	8	11	3	5
1	5	9	4	8

Abb. 2.6: Anordnung der Klone der Mikrotiterselektionsplatten innerhalb eines Unterfeldes im *grid*. Die Filter (bestehend aus sechs Feldern) wurden in 48 x 48 Unterfelder unterteilt, wobei in jedem Unterfeld die jeweiligen Klone der entsprechenden Platten als Duplikate in einem 5 x 5-Pattern um einen *guide dot* angeordnet sind. Insgesamt wurden pro Unterfeld 12 verschiedene Klone angeordnet. In der Mitte eines jeden Feldes befindet sich der *guide dot*.

2.2.2.4 Neuordnung einer cDNA-Bibliothek zur Erstellung einer Unterbibliothek mit potentiellen Proteinexpressionsklonen

Mit Hilfe des „Mercury“-Roboters wurden die potentiellen Proteinexpressionsklone, entsprechend der im vorherigen Kapitel erstellten Liste, in mediumbefüllte 384er Mikrotiterselektionsplatten neu angeordnet. Dazu wurde die cDNA-Bibliothek frisch repliziert. Nach der Neuordnung wurde die Unterbibliothek ÜN bei 37°C inkubiert, repliziert, erneut kultiviert, eingefroren und bei - 80°C gelagert (Kap. 2.2.1.8).

Alle Replikationsschritte erfolgten in 384er Mikrotiterselektionsplatten, welche mit 2YT, 1 x HMFM, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin befüllt waren.

2.2.3 Herstellung eines Uniklonsets einer *Arabidopsis* cDNA-Expressionsbibliothek

2.2.3.1 Sequenzierung

Alle Sequenzierungen wurden von der AGOWA GmbH (Berlin) unter Verwendung der Gefrierkulturen der entsprechenden Klone durchgeführt. Dabei wurde ausschließlich vom 5'-Ende mit Hilfe des Vektorprimers pQE65 ansequenziert (Kap. 2.2.4.2).

2.2.3.2 Qualitätskontrolle der Sequenzen

Die Sequenzen wurden mit Hilfe des Staden-Software-Pakets, Version 2001 (Staden *et al.*, 2000) bearbeitet. Dieses Paket beinhaltet verschiedene Programme zur Qualitätskontrolle und zum Selektieren, Aufbereiten, Assemblieren und Editieren von DNA-Sequenzen. Eine ausführliche Zusammenfassung des Staden-Software-Pakets kann www.mrc-lmb.cam.ac.uk/pubseq/staden_home.html entnommen werden.

Als erstes erfolgte mittels des *Basecaller*-Programms „Phred“ das *base calling* der Rohsequenzen, d.h. die Fluoreszenzsignale und –intensitäten wurden analysiert und in DNA-Sequenzen umgerechnet (Ewing und Green, 1998). „Phred“ liest die Rohdaten (Sequenz Chromatogramme), analysiert die Basenabfolge der DNA und führt eine Qualitätskontrolle anhand der Summe der Flächen, der im Chromatogramm befindlichen Peaks durch und teilt den Sequenzen einen Qualitätswert zu.

Mit Hilfe des Programmes „Pregap 4 1.2“ wurden Vektor- (pQE30NASTattB) und Adaptersequenzanteile aus den DNA-Sequenzen entfernt (Bonfield und Staden, 1996).

Alle Sequenzen, die nach der Qualitätskontrolle durch „Phred“ bzw. dem Entfernen aller Vektor- und Adapteranteile kürzer als 100 Nukleotide waren, wurden verworfen.

Bei der Erstellung der Bibliothek kann es zu Kontaminationen mit der DNA des Wirtsstammes (*E. coli*) kommen. Aus diesem Grund wurden die Sequenzen über „BLASTN 2.2.6“ gegen die genomische *E. coli* Nukleotid-Sequenzdatenbank (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA>) abgeglichen und Sequenzen mit guter Übereinstimmung zu *E. coli* Sequenzen, d.h. mit einem E-Wert (Irrtumswahrscheinlichkeit) größer als e^{-20} , eliminiert.

Ausgehend von den Sequenzen mit ausreichender Qualität wurden BLAST- und Cluster-Analysen durchgeführt. Die Ergebnisse dieser Analysen wurden in einer Tabelle zusammengefasst, welche sich aufgrund ihrer Größe auf folgender HTML-Seite befindet:

- HTML-Tabelle 1:

http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionCloneSet.shtml

Nach der BLAST- und Cluster-Analyse und der Erstellung der HTML-Tabelle 1 folgte die Erstellung des Uniklonsets. Die Ergebnisse wurden in einer weiteren Tabelle zusammengefasst, welche ebenfalls aufgrund ihrer Größe auf folgender HTML-Seite zu finden sind:

- HTML-Tabelle 2:

http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionUniCloneSet.shtml

Zur Veranschaulichung des Aufbaus dieser Tabellen wurden jeweils die ersten Zeilen einer jeden Tabelle im Anhang dargestellt (siehe Anhang 5.3. und 5.4).

Da ein Ziel dieser Arbeit die Erstellung eines Uniklonsets war, beziehen sich alle Tabellenverweise der folgenden Kapitel (2.2.3.3, 2.2.3.4 und 2.2.3.5) auf die HTML-Tabelle 2.

2.2.3.3 BLAST-Analyse der Sequenzen

Mit den nach der Qualitätskontrolle verbliebenen Sequenzen wurde mit dem Programm „BLASTX 2.2.6“ (Altschul *et al.*, 1997) eine BLAST-Analyse gegen alle sich in der MIPS-*Arabidopsis*-Datenbank (<http://www.mips.gsf.de/proj/thal/>) befindlichen Proteinsequenzen durchgeführt (Version: arabi_all_proteins_v11103; <ftp://ftpmips.gsf.de/cress/arabiprot>).

Mit Hilfe eines Perl-Skripts wurden die Parameter aus den BLAST-Ergebnissen ausgelesen. Dabei wurden sowohl die Sequenzen mit einem BLAST-Hit als auch die Sequenzen ohne einem BLAST-Hit in die oben erwähnte HTML-Tabelle 2 übertragen. Allen Sequenzen mit BLAST-Hit wurde der AGI *gene code* des besten Hits zugeordnet. Die Tabelle setzt sich aus folgenden Parametern zusammen:

- Sequenz-Name (Spalte B)
- GenBank-*Accession* Nummer (Spalte C)
- BLAST-Hit (AGI (*Arabidopsis* Genom Initiative) *gene code*) auf MIPS-Ebene (Spalte H)
- putative Genfunktion auf MIPS-Ebene (Spalte I)
- E-Wert (Irrtumswahrscheinlichkeit) auf MIPS-Ebene (Spalte J)
- Startpunkt des Query (Abfrage-Sequenz) [Nukleotiden] (Spalte K)
- Startpunkt des Subjekt (Treffer-Sequenz) [Aminosäuren] (Spalte L)
- Subjektlänge der vollständigen cDNA-Sequenz [Aminosäuren], (Spalte M)

Neben den oben genannten Parametern sind der HTML-Tabelle 2 noch weitere Parameter hinzugefügt:

- Unigen-Name (Spalte A)
- Aminosäureposition, bei der das erste Stopcodon identifiziert werden konnte (Spalte G)

Zur Ermittlung des letzten Parameters wurden die Nukleotidsequenzen, bei denen der Vektor- und Adapteranteil entfernt wurde, im Leserahmen 1 in Aminosäuresequenzen translatiert. Dies wurde unter der Verwendung des Programms „Transeq“ (translates nucleic acid sequences to the corresponding peptide sequences) durchgeführt, welches Bestandteil des EMBOS-Paketes ist (Rice *et al.*, 2000). Die translatierten Stopcodone werden mit Hilfe

des Programms mit einem "*" gekennzeichnet. Die Identifizierung der Position des ersten Stopcodons erfolgte mit Hilfe eines Perl-Skriptes.

2.2.3.4 Cluster-Analyse der Sequenzen

Der Cluster-Analyse der DNA-Sequenzen vorhergehend erfolgte eine Überprüfung der Sequenzen auf repetitive Bereiche und Elemente geringer Komplexität durch das Programm „RepeatMasker“ (<http://repeatmasker.genome.washington.edu>) unter der Verwendung der Referenzdatenbank Athrep.ref (RepBase7.7.fasta: athrep.ref, Version 6.4.0). Diese enthält alle bekannten hoch-repetitiven Sequenzen von *Arabidopsis thaliana*. „RepeatMasker“ vergleicht mit Hilfe des Programmes „CrossMatch“ (Ewing *et al.*, 1998) (phred1), <http://genome.uc.edu/genome/HelpPages/phred-phrap-polyphred/swat-crossmatch.html>) die Eingangssequenzen mit den Sequenzen der Referenzdatenbank. Die so ermittelten hoch-repetitiven Sequenzabschnitte wurden im weiteren mit „RepeatMasker“ maskiert, indem die Nukleotide dieser Abschnitte durch „N“ ersetzt wurden.

Anschließend erfolgte die Cluster-Analyse der Sequenzen mit dem Programm „d2_cluster“ (Burke *et al.*, 1999). Ziel einer Cluster-Analyse ist, die zu klassifizierenden Objekte in Klassen zusammenzufassen. Dabei sollen die Objekte innerhalb einer Klasse möglichst ähnlich sein, während die Klassen untereinander möglichst unähnlich sein sollten.

Alle sich im Set befindlichen Sequenzen wurden zusammen mit allen kodierenden Sequenzen der MIPS-*Arabidopsis*-Datenbank (Version: arabi_cds_v110903) geclustert (<ftp://ftpmips.gsf.de/cress/arabidna>). Es wurden die voreingestellten Parameter für die Cluster-Analyse gewählt. Dabei werden zwei Sequenzen dem gleichen Cluster zugeordnet, wenn diese beiden Sequenzen um mehr als 100 Nukleotide überlappen und in diesem Überlappungsbereich die Ähnlichkeit größer als 90 % ist.

Die folgende Informationen wurden mittels Perl-Skripte aus der Cluster-Analyse ausgelesen und ebenfalls in die oben genannte HTML-Tabelle 2 übertragen:

- Cluster-Nummer (Spalte D)
- Vergleich Cluster-AGI *gene code* und BLAST-AGI *gene code* (Spalte E)
- Cluster AGI *gene codes*, die im jeweiligen Cluster enthalten ist/sind (Spalte F)

Unabhängig von den oben genannten Parametern sind in der oben genannten HTML-Tabelle 2 zusätzlich noch enthalten:

- BLAST-Hit auf TIGR-Ebene (Spalte N)
- putative Genfunktion auf TIGR-Ebene (Spalte O)
- E-Wert auf TIGR-Ebene (Spalte P)

Diese letzten drei Spalten (N – P) der HTML-Tabelle wurden zu Aktualisierung der Daten kurz vor Beendigung der vorliegenden Arbeit (August 2004) eingefügt. Hier wurde von den Sequenzen eine erneute BLAST-Analyse gegen alle sich in der TIGR-*Arabidopsis*-Datenbank befindlichen Proteinsequenzen (Version: 5.0) durchgeführt (ftp://ftp.tigr.org/pub/data/a_thaliana/ath1/SEQUENCES).

Die Selektionen der Klone für das Uniklonset erfolgten alle basierend auf den Parametern der BLAST-Analyse gegen MIPS.

2.2.3.5 Selektion der Klone für das Uniklonset

Das erste Selektionskriterium war die Ermittlung der Position des ersten Stopcodons innerhalb der Sequenzen und die Eliminierung von Sequenzen, die ein Stopcodon bei ≤ 70 Aminosäuren besaßen.

Aufgrund der Cluster- und BLAST-Analysen ergaben sich zwei große Sequenzgruppen:

1. Sequenzen, die bei der Cluster- und BLAST-Analyse gleiche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit aufwiesen (Spalte F mit Eintrag und Spalte E mit Plus gekennzeichnet).
2. Sequenzen, die nicht mit einer annotierten *Arabidopsis*-Sequenz zusammen in ein Cluster fielen (Spalte F ohne Eintrag) bzw. Sequenzen, die bei der Cluster- bzw. BLAST-Analyse unterschiedliche *Arabidopsis*-Sequenzen als Clusterpartner bzw. besten BLAST-Hit hatten (Spalte E mit Minus gekennzeichnet).

Diese beide großen Sequenzgruppen wurden bezüglich der Selektion unterschiedlich behandelt. Sequenzen für das Uniklonset wurden aus Sequenzgruppe 1 nach folgenden Kriterien ausgewählt:

- Nur Klone, die sich im Leserahmen 1 befinden, wurden weiter selektiert.
- Von diesen Klonen wurden nur die Klone einbezogen, die einen E-Wert (Irrtumswahrscheinlichkeit) besser (d.h. kleiner) als e^{-20} aufwiesen.

- Klone die nur einmal vorkamen (*Singeltons*) wurden direkt in das Uniklonset transferiert. Klone, die in Clustern vorlagen, unterlagen weiteren Selektionskriterien. Die Kriterien für die Klone innerhalb eines Clusters waren, beginnend mit dem wichtigsten, die folgenden:
 - Startpunkt des Subjekt (Treffer-Sequenz) [Aminosäuren] (Spalte K)
 - Startpunkt des Query (Abfrage-Sequenz) [Nukleotiden] (Spalte L)
 - E-Wert (Irrtumswahrscheinlichkeit) (Spalte J)

Die Werte der zuletzt genannten drei Parameter sollten so klein wie möglich sein, immer an der hierarchischen Anordnung der Kriterien orientiert. Aufgrund dieser Selektion sollten je Cluster ein Klon in das Uniklonset transferiert werden.

Sequenzen, die mit mehreren annotierten *Arabidopsis*-Sequenzen in ein Cluster fielen, wurden nach ihren BLAST-Hits gruppiert und für jeden unterschiedlichen BLAST-Hit wurde ein Klon entsprechend der oben beschriebenen Selektionskriterien in das Uniklonset übernommen.

Bei den Sequenzen der Sequenzgruppen 2 wurden, nach der oben beschriebenen Stopcodonselektion, die verbliebenen *Singeltons*, direkt in das Uniklonset transferiert. Für Klone, die in dieser Gruppe Cluster bildeten, wurde als weiteres Kriterium die Clustergröße einbezogen. Von Clustern, die aus weniger als sechs Sequenzen bestanden, wurden alle Klone in das Uniklonset übertragen. Bei Clustern, die sich aus sechs oder mehreren Sequenzen zusammensetzten, wurden nur die Klone verwendet, die kein Stopcodon hatten.

2.2.3.6 Anordnung des Uniklonsets

Die selektierten Klone wurden basierend auf der in Kap. 2.2.3.5 erstellten Liste mit Hilfe des „Mercury“-Pick-Roboters in beschriftete und mediumbefüllte 384er Mikrotiterplatten (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin), wie in Kap. 2.2.2.4 erläutert, neu angeordnet.

2.2.3.7 Erweiterung des Uniklonsets durch vollständige cDNA-Expressionsklone

Das Uniklonset wurde durch 192 vollständige cDNA-Expressionsklone ergänzt. 96 dieser Klone wurden durch restriktionsabhängige Klonierung (*AscI/NotI*) gewonnen (Kersten *et al.*, 2003). Bei den restlichen 96 Klonen handelte es sich um Transkriptionsfaktor-Klone, welche freundlicherweise von Dr. Ralf Stracke und Prof. Dr. Bernd Weisshaar (Max Planck Institut für Züchtungsforschung, Köln) zur Verfügung gestellt wurden. Diese lagen als GATEWAY-

Entryklone vor und wurden im Rahmen dieser Arbeit im 96 *well* Format mittels der LR-Reaktion in den *E. coli* Destinationsvektor (pQE30NASTDV) überführt.

2.2.3.7.1 LR-Reaktion

Das im Entryklon befindliche Gen ist flankiert mit attL-Stellen und kann in einen Expressionsvektor transferiert werden. Dabei werden die attL-Stellen mit den attR-Stellen des Destinationsvektors (pQE30NASTDV) mit Hilfe der GATEWAYTM-Technologie rekombiniert (siehe Abb. 1.1).

Zunächst wurde von dem Destinationsvektor eine Vektorpräparation (Kap. 2.2.4.1) mit einer anschließenden Vektorlinearisierung durchgeführt. Diese Linearisierung hatte den Effekt, die Aktivität der Suizid-Funktion des *ccdB*-Gens zu vermindern. Dazu wurden 6 µg des Vektors mit dem Enzym *Bss*HIII (4 U) bei 50°C für 3 h geschnitten. Anschließend wurde der linearisierte Vektor mittels Gelextraktion isoliert (Kap. 2.2.4.5). Ausgehend von dem linearisierten und isolierten Vektor erfolgte die LR-Reaktion im 96 *well* Format (Tab. 2.1).

Tab. 2.1: Reaktionsansätze der LR-Reaktion je Entryklon.

		Negativkontrolle
Entryklone	4 µl	-
pQE30NASTDV (c = 30 ng/µl)	4 µl	4 µl
LR-Puffer	2 µl	2 µl
1 x TE	-	4 µl
LR-Klonase Enzym-Mix	1 µl	1 µl
Proteinkinase K	1 µl	1 µl

In dieser Tabelle sind die Reaktionsansätze der LR-Reaktion dargestellt. Die Ansätze wurden im 96 *well* Format durchgeführt.

Die entsprechenden Komponenten der Reaktionsansätze wurden bis auf die Proteinkinase K vermischt und bei 26°C für 1 h inkubiert. Anschließend erfolgte die Zugabe der Proteinkinase K und eine zehnmütige Inkubation bei 37°. Dann wurden die Ansätze auf Eis überführt und standen für die Transformation zur Verfügung.

2.2.3.7.2 Herstellung hitzeschockkompetenter *E. coli* SCS1/pSE111-Zellen

10 ml LB-Medium (mit 30 µg/ml Kanamycin) wurden mit einer Einzelkolonie *E. coli* SCS1/pSE111 inokuliert und ÜN bei 37°C geschüttelt. 1 ml der Vorkultur wurde in 100 ml LB-Medium überführt und bei 37°C bis zu einer OD₆₀₀ von 0,5 – 0,7 inkubiert. Danach folgte eine fünfminütige Zentrifugation (1.850 x g) bei 4°C. Das Zellpellet wurde in 2 ml TSS auf Eis resuspendiert und anschließend mit 200 µl gekühltem 87%igem (v/v) autoklavierten Glycerin versetzt, aliquotiert (je ~ 70 µl in 0,5 ml Reaktionsgefäße) und bei -80°C bis zur weiteren Verwendung gelagert. Die Transformationskompetenz wurde mittels pUC 18 ermittelt (Kap. 2.2.1.6).

2.2.3.7.3 Transformation von *E. coli* mittels Hitzeschock

Für die Transformation wurden pro *well* 30 µl der hitzeschockkompetenten Zellen (Kap. 2.2.3.7.2) in einer 96er Mikrotiterplatte vorgelegt und mit 1 µl des LR-Reaktionsgemisches (Kap. 2.2.3.7.1) auf Eis vermischt und für 30 min inkubiert. Danach erfolgte der Hitzeschock bei 42°C für 30 s. Zur Regeneration der Zellen wurden diese mit 1 ml 2YT-Medium (*deep-well*-Platten) versetzt und bei 37°C für 1 h unter Schütteln (~ 650 upm) inkubiert.

Anschließend wurden die Transformationsansätze einzeln auf 6 *well* Zellkulturplatten (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) aufgebracht und ÜN bei 37°C inkubiert. Je *well* wurden zwei Klone gepickt und in zwei befüllte 96er Mikrotiterselektionsplatten (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) transferiert und ÜN bei 37°C im Schüttler (~ 650 upm) inkubiert. Die Klone wurden mit Glycerin (Endkonzentration: 20% (v/w)) versetzt und bei -80°C gelagert.

2.2.4 Weitere nukleinsäuretechnische Methoden

2.2.4.1 Plasmidpräparation

Die Isolierung qualitativ hochwertiger DNA wurde unter Verwendung der Qiagen Mini-, Midi- und Maxi-Kits, gemäß den Angaben des Herstellers, durchgeführt.

Dazu wurden transformierte *E. coli* ÜN Kulturen (Kulturvolumen pro Säule: ca. 3 ml für Minipräparation, ca. 50 ml für Midipräparation und ca. 200 ml für Maxipräparation) für etwa 20 min bei 6.000 x g zentrifugiert und das Bakterienpellet anschließend in einem Resuspensionspuffer, der 100 µg/ml RNase A enthielt, aufgenommen. Nach der Zell-Lyse wurde die Suspension neutralisiert und zentrifugiert (30 min, 20.000 x g, 4°C). Der

Überstand, der die Plasmid-DNA enthielt, wurde dann auf eine aktivierte Säule gegeben. Bei dem Säulenmaterial handelt es sich um ein Anionen-Austauscher Harz, an das DNA bei geringer Salzkonzentration und niedrigem pH bindet. Durch eine Erhöhung des pH Wertes und der Salzkonzentration kann die Plasmid-DNA eluiert werden. Die gebundene DNA wurde in einem Waschschrift von Zellmetaboliten gereinigt, anschließend mit einem Elutionspuffer eluiert, mit 0,7 Vol. Isopropanol gefällt und 20 - 30 min bei 4°C pelletiert (20.000 x g). Danach wurde die DNA mit einem geeigneten Vol. 70% (v/v) EtOH gewaschen und nach Zentrifugation getrocknet, in Elutionspuffer resuspendiert und bis zur Verwendung bei -20°C gelagert.

Für die Vektorpräparation der Vektoren pQE30NASTattb und pENTR1a wurde eine ÜN Kultur von 800 ml (für pQE30NASTattb: 2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin, für pENTR1a: 50 µg/ml Kanamycin) eingesetzt. Die Aufreinigung erfolgte nach dem oben erläuterten Qiagen-Protokoll. Für 800 ml ÜN-Kultur wurden vier Maxi-Säulen verwendet, von denen die DNA mit je 200 µl Elutionspuffer eluiert wurde. Um sicherzustellen, dass die Eluate homogen gelöst waren, folgte eine dreistündige Inkubation bei 60°C. Die erhaltenen Eluate dienten als Ausgangsmaterial für die Vektorpräparation (Kap. 2.2.1.3).

2.2.4.2 Sequenzierung und Sequenzanalyse

Die Sequenzierungen der potentiellen Proteinexpressionsklone der Infloreszenz cDNA-Bibliothek wurden von der AGOWA GmbH (Berlin) durchgeführt. Alle anderen Sequenzierungen erfolgten bei DLMBC (Berlin). Es wurde ausschließlich vom 5'-Ende ansequenziert. Die verwendeten Oligonukleotide waren pQE65 (Infloreszenz-Klone) und pENTR1a-5' (Pistill-Klone) von der Firma Metabion, mit einer eingestellten Konzentration von 100 pmol/µl.

2.2.4.3 PCR zur Amplifikation von DNA-Fragmenten

Mit der Methode der PCR (Polymerase-Kettenreaktion) werden DNA-Matrizen exponentiell amplifiziert. Dazu wird die doppelsträngige DNA zuerst hitzedenaturiert, um dann einzelsträngig in einer Anlagerungsreaktion mit sequenzspezifischen Primern partielle doppelsträngige DNA-Matrizen für eine thermostabile Polymerase zu erzeugen (Annealing). Diese füllt in einer spezifischen Kettenreaktion (in Richtung 5' nach 3') die Einzelstränge zu Doppelsträngen auf (Verlängerung, Elongation). Diese findet bei 72°C (Temperaturoptimum der Taq-DNA-Polymerase) statt.

Nachdem der erste Zyklus beendet ist, wird die Temperatur wieder auf 94°C erhöht. Die DNA wird erneut denaturiert und der Prozess beginnt von neuem. In 20 - 35 solcher Zyklen kann so auch von kleinsten Mengen Template-DNA eine große Anzahl von Kopien hergestellt werden. Die Anzahl der Zyklen wurde beschränkt, da bei zu einer hohen Anzahl an Zyklen die Lesegenauigkeit der Taq-DNA-Polymerase unzureichend wird.

Die Reaktionsbedingungen hängen von verschiedenen Parametern ab. Sie richten sich u. a. nach der Länge der zu amplifizierenden Sequenz und nach der spezifischen Anlagerungstemperatur der eingesetzten Primerpaare. Die Anlagerungstemperatur beider Primer sollten ähnlich sein, und beide Primer sollten in gleicher Konzentration vorliegen. Die Reaktionszyklen wurden in automatischen Thermocyclern durchgeführt. Meist wurde nach Abschluss der Zyklen ein verlängerter Elongationsschritt von 3 - 7 min (72°C) zum Vervollständigen unvollständiger Amplifikate angefügt. Ein Reaktionsansatz (50 µl) zur Amplifikation von DNA-Fragmenten enthielt in der Regel folgende Komponenten:

- 1 x PCR-Puffer
- 0,2 mM dNTPs (dATP, dCTP, dGTP, dTTP)
- 0,2 µM Vorwärtsprimer
- 0,2 µM Rückwärtsprimer
- 2,5 U Taq-DNA-Polymerase
- ddH₂O
- DNA-Template

Für die PCR wurde folgendes Temperatur-Programm verwendet:

Denaturierung: 94°C für 4 min
Denaturierung: 94°C für 30 s —————
Anlagerung: x°C für 1 min 24 x
Strangsynthese: 72°C für 1 min 30 s ———
Strangsynthese: 72°C für 7 min

Die Anlagerungstemperaturen (x°C) für die verwendeten Primer sind dem Kap. 2.1.3 zu entnehmen.

Nach Beendigung der Amplifikation wurde ein Aliquot im Agarosegel überprüft.

2.2.4.4 Gelelektrophoretische Auftrennung von DNA

DNA-Fragmente werden zur Analyse ihrer Molekülgröße üblicherweise in Agarosegelen (0,6% - 1,5% (w/v)) elektrophoretisch aufgetrennt. Auf Grund der negativen Ladung von Nukleinsäuren lassen sich diese in einem elektrischen Feld bewegen. Basierend auf dieser Bewegung erfolgt die Auftrennung der Nukleinsäuren. Durch die Vernetzung der Agarose werden größere Moleküle stärker zurückgehalten als die kleineren und bewegen sich deshalb langsamer in Richtung Kathode. Die Laufstrecke der Nukleinsäuren in einer bestimmten Zeit ist dabei umgekehrt proportional zum Logarithmus ihrer Größe.

0,8% bzw. 1,0% (w/v) Agarose wurde durch kurzzeitiges Kochen in 0,5 x TBE-Puffer gelöst und nach Abkühlen (~ 60°C) mit Ethidiumbromid (c = 0,5 µg/ml) versetzt und in die entsprechenden horizontalen Gelträger gegossen. In den so hergestellten Gelen wurden die Proben in 0,5 x TBE-Puffer elektrophoretisch aufgetrennt. Die Spannung betrug 80 V, die Laufzeit variierte zwischen 20 - 30 min. Die DNA-Fragmente wurden durch Interkalation mit dem Fluoreszenzfarbstoff Ethidiumbromid unter UV-Licht bei einer Wellenlänge von 254 nm sichtbar gemacht und photographiert. Der parallel zu den Proben aufgetragene DNA-Längenstandard ermöglicht die Größenzuordnung der DNA-Fragmente [Bp].

2.2.4.5 Isolierung von DNA-Fragmenten aus Agarosegelen

Zur Isolierung von DNA aus Agarosegelen wurden diese mit Hilfe von UV-Licht (366 nm) visualisiert. Die zu isolierenden Banden wurden mit einem sterilen Skalpell aus dem Gel ausgeschnitten und die DNA unter Verwendung des QiaexII Gelextraktions-Systems daraus isoliert. Dieses System macht sich die Eigenschaft von Agarose zunutze, sich in Gegenwart von hohen NaJ-Konzentrationen aufzulösen. Die so freigesetzte DNA bindet an die zugesetzten Glaspartikel und kann so von der Agarose abgetrennt und aufgearbeitet werden. Gemäß den Angaben des Herstellers wurden die aus dem Gel ausgeschnittenen Banden in ein Reaktionsgefäß überführt, gewogen und pro 100 mg Gel wurden 300 µl des NaJ-haltigen Puffers QX1, sowie pro 5 µg erwarteter DNA 10 µl Glasmilch zugesetzt. Es wurde 10 min bei 50°C und häufigem Mischen inkubiert, anschließend die Glasmilch mit gebundener DNA 10 s bei 16.000 x g zentrifugiert. Das Pellet wurde 2 x in 500 µl eines hochsalzhaltigen Waschpuffers (QX2) gewaschen und zentrifugiert (16.000 x g). Im Anschluss wurde das Pellet 2 x mit einem weiteren Waschpuffer (PE) gewaschen und anschließend für 15 min bei Raumtemperatur getrocknet. Zur Elution der DNA wurde das trockene Glasmilchpellet in 30 µl TE-Puffer aufgenommen und 10 min bei 50°C inkubiert. Die Glasmilch wurde dann 2 min bei 16.000 x g abzentrifugiert und der Überstand in ein

neues Reaktionsgefäß überführt. Zum vollständigen Entfernen von Glasmilchresten wurde dieser Vorgang wiederholt.

2.2.5 Weitere proteinchemische Methoden

2.2.5.1 SDS-Polyacrylamidgel-Elektrophorese (SDS-PAGE)

Bei dieser Methode erfolgt die Auftrennung der Proteine unter Verwendung von Polyacrylamidgelen. Dabei werden die Proteingemische, in Abhängigkeit von der zu erwartenden Größe der Proteine, auf Gele mit variierenden Polyacrylamidkonzentrationen (8 - 15% (w/v)) aufgetragen. Durch Inkubation der Proteine mit SDS (Sodiumdodecylsulfat) und mit Disulfidbrücken-reduzierenden Mitteln (z.B. β -Mercaptoethanol (20% (v/v)) oder DTT (0,1 M)) wurde, während eines fünf- bis zehnminütigen Aufkochens der Proben, eine vollständige Denaturierung und Solubilisierung der Proteine erreicht. Das SDS bindet an die hydrophoben Regionen der Proteine und SDS-Protein-Komplexe werden gebildet. Aufgrund der negativen Ladung des SDS wandern die SDS-Protein-Komplexe daher im elektrischen Feld in Richtung Anode, und die Auftrennung der Proteine erfolgt entsprechend ihrer unterschiedlichen Molekulargewichte.

Als SDS-Gelelektrophoresesystem wurde das diskontinuierliche Lämmli-System mit Tris-HCl-Puffern verwendet. In diesem System wird das Trenngel von einem Sammelgel überschichtet. Das Sammelgel dient zur Fokussierung der Proteine. Im Trenngel findet die eigentliche Trennung der Proteine statt. Dabei hängt der Auftrennungsbereich von der Polyacrylamidkonzentration ab und kann der Tab. 2.2 entnommen werden.

Tab. 2.2: Auftrennungsbereiche der Trenngelschicht in Abhängigkeit von der Polyacrylamidkonzentration.

Trenngel	8%	10%	12%	15%
Auftrennungsbereich (kDa)	15 - 300	5 - 200	5 - 150	2 - 100
Optimale Auftrennung (kDa)	25 - 150	15 - 100	10 - 70	5 - 50

In dieser Tabelle Auftrennungsbereiche der Trenngelschicht in Abhängigkeit von der Polyacrylamidkonzentration dargestellt

Im Rahmen dieser Arbeit wurden ausschließlich 15% Gele verwendet, die sich aus folgenden Komponenten zusammensetzten:

	Trenngel	Sammelgel
SDS	0,1% (w/v)	0,1% (w/v)
Acrylamid/Bis-Lösung	15 % (v/v)	3,8% (v/v)
Tris-HCl	0,38 M; pH 8,8	0,125 M; pH 6,8
Amoniumpersulfat	0,1% (w/v)	0,1% (w/v)
TEMED	0,03% (v/v)	0,1% (v/v)

Zur Herstellung der Gele wurden die entsprechenden Komponenten vereinigt und gut gemischt. Die Polymerisation der Gele wurde durch die Zugabe der jeweiligen Menge TEMED gestartet. Zuerst wurde das Trenngel zwischen zwei gesäuberte Glasplatten, welche durch *spacer* voneinander getrennt wurden, gefüllt und mit Isopropanol überschichtet. Nach abgeschlossener Polymerisation wurde das Isopropanol abgegossen und das Sammelgel über das Trenngel geschichtet. Der „Gelkamm“ wurde luftblasenfrei eingesetzt. Vor dem Auftragen der Proteinproben wurden die auspolymerisierten Gele an die Elektrophoresekammer angeschlossen und mit Elektrophoresepuffer überschichtet. Pro Tasche wurden 10 µl Probe aufgetragen. Zur Bestimmung der Proteingrößen [kDa] wurde auf den SDS-Polyacrylamidgele ein vorgefärbter Molekulargewichtsmarker (*rainbow coloured protein marker*) mitgeführt.

Für die SDS-Polyacrylamidgele von Gesamt-Zell-Extrakten wurden nach der Induktion je Kultur 100 µl Probe entnommen und zentrifugiert (20 min, 4.000 x g , 4°C). Die Pellets wurden in 20 µl PBS resuspendiert und mit 5 µl 5 x SDS-Ladepuffer versetzt. Anschließend wurden die Proben 5 - 10 min bei 100°C denaturiert und auf die Gele aufgetragen.

Um die Elektrophorese zu starten, wurde zunächst eine Spannung von 80 V angelegt. Nachdem die Lauffront die Grenze zwischen Sammel- und Trenngel passiert hatte, wurde die Spannung auf 150 V erhöht. Wenn die Lauffront die untere Trenngelgrenze erreicht hatte, wurde der Lauf beendet.

Die SDS-Polyacrylamidgele wurden standardmäßig mit Coomassie Brilliant Blau R250 gefärbt. Nach der Elektrophorese wurde das Gel ca. 30 min in der Färbelösung geschwenkt. Zur Differenzierung wurde das Gel anschließend für mehrere Stunden in Entfärbelösung schwenkend inkubiert, wobei der Entfärbler mehrmals gewechselt wurde.

2.2.5.2 Westernblot

2.2.5.2.1 *Semi-dry Blot von Proteinen*

Bei diesem Verfahren wurden die mittels eines SDS-Polyacrylamidgels elektrophoretisch aufgetrennten Proteine (Kap. 2.2.5.1) auf eine PVDF-Membran transferiert. Dazu wurde das Gel von den Glasplatten gelöst, das Sammelgel abgetrennt und verworfen und das Trenngel in den Blotpuffer überführt. Die Membran und sechs Whatmann-Papiere wurden entsprechend der Gelgröße zugeschnitten. Die mit 99%igem Ethanol und ddH₂O aktivierte Membran sowie die Whatmann-Papiere wurden ebenfalls in Blotpuffer geschwenkt. Für den *semi-dry* Blot wurden erst drei getränkte Whatman-Papiere, die Membran, das Gel und anschließend wiederum drei getränkte Whatman-Papiere übereinander luftblasenfrei auf die Anodenplatte gelegt. Nachdem die Kathodenplatte aufgebracht wurde, erfolgte der Transfer für 1 h bei 0,8 mA/cm². Der mitgeführte vorgefärbte Molekulargewichtsmarker diente neben der Bestimmung der Proteingrößen auch zur Kontrolle des Transfers. Anschließend konnte der Blot getrocknet oder direkt für weitere Analysen verwendet werden.

2.2.5.2.2 *Immunofärbung des Blots*

Vor der Hybridisierung mit Antikörpern wurde die Membran ÜN in Blocklösung geschwenkt, um die verbliebenen Bindungsstellen abzusättigen.

Am darauf folgenden Tag wurde die Membran für 2 h mit dem Erstantikörper (anti-RGS His₆, 1:2.000 verdünnt in Blocklösung) bei Raumtemperatur inkubiert. Die Membran wurde 3 x 10 min in TBST gewaschen und dann für 1 h mit dem Zweitantikörper (anti-Maus-AP-Konjugat, 1:5.000 verdünnt in Blocklösung) bei Raumtemperatur inkubiert. Es folgte ein dreimaliges Waschen der Filter in TBST.

Die Immunfärbung des Blotes mit dem anti-RGS His₆-Antikörper erlaubt den Nachweis rekombinanter His-getagter Proteine. Detektiert wurde der gebundene Zweitantikörper mittels Färbung der Membran mit NBT/BCIP (Substrat der alkalischen Phosphatase). Dazu wurde eine Tablette in 10 ml ddH₂O gelöst und lichtgeschützt auf der ausgebreiteten Membran verteilt. Die Färbung wurde durch Waschen der Membran mit ddH₂O gestoppt.

2.2.5.3 Proteinexpression in *E. coli* in 96 well Format (1 ml Kultur)

Für die Proteinexpression wurden die Klone in 100 µl Medium (2YT, 2% (w/v) Glukose, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin) in 96 *deep-well*-Platten angeimpft. Das Animpfen der Kulturen erfolgte aus 384er oder 96er Mikrotiterselektionsplatten mit

Bakterien, die bei - 80°C gelagert wurden, mit Hilfe eines wiederverwendbaren 96er Stahl-Replikators.

Die *deep-well*-Platten wurden ÜN bei 37°C im Schüttler (~ 650 upm) inkubiert. Am nächsten Morgen wurden die Kulturen mit vorgewärmtem Medium (SB, 100 µg/ml Ampicillin, 15 µg/ml Kanamycin, 1x KBP-Puffer und 20 µg/ml Thiamin) 1/10 in den Platten verdünnt und für weitere 2 h inkubiert. Zur Einleitung der Induktion wurde IPTG in einer Endkonzentration von 1 mM hinzugegeben. Nach einer weiteren Inkubation bei 37°C für 4 - 5 h wurden die Platten bei 4.000 x g für 10 – 20 min zentrifugiert, das überständige Medium wurde entfernt und die Zellpellets bei - 80°C eingefroren.

2.2.5.4 Denaturierende Aufreinigung im 96 *well* Format von Proteinen mit RGS His₆-Tag

Die bei - 80°C eingefrorenen Zellpellets (Kap. 2.2.5.3) wurden aufgetaut, und mit je 150 µl Lysepuffer versetzt, resuspendiert und für 30 min bei Raumtemperatur schüttelnd (~ 650 upm) inkubiert. Die gelösten Pellets wurden dann bei 1.900 x g bei 4°C für 30 min zentrifugiert. Der Überstand wurde vorsichtig abgenommen, in eine 96 *well* Filterplatte überführt und mit Hilfe einer Vakuumpumpe in eine zweite Filterplatte filtriert.

Je Lysat wurden 30 µl einer 50%igen (w/v) Ni-NTA Agaroselösung (in Lysepuffer verdünnt) hinzugegeben und für 1 h unter Schütteln inkubiert. Durch eine anschließende Vakuumfiltration wurde die Flüssigkeit entfernt. Danach wurden drei Waschschriffe für je 5 min mit jeweils 100 µl Waschpuffer durchgeführt. Zwischen den einzelnen Waschschriffen wurde die Flüssigkeit ebenfalls wieder vakuumfiltriert. Für die Elution der Proteine von der Ni-NTA-Agarose wurden je *well* 80 µl Elutionspuffer zugegeben und für 20 min bei ~ 650 upm schüttelnd inkubiert. Dann wurden die Eluate in eine neue, sterile Mikrotiterplatte filtriert. Die Proteine wurden bei - 20°C gelagert.

2.2.5.5 Bradford-Test

Mit Hilfe dieses Tests (Bradford, 1976) ist es möglich, die Proteinkonzentration von Lösungen zu bestimmen. Dabei wird der Effekt ausgenutzt, dass sich bei der Bindung von Coomassie-Brilliant-Blau an Proteine das Absorptionsmaximum des Farbstoffes von 465 nm nach 595 nm verschiebt. Die Zunahme der Absorption bei 595 nm ist proportional zur Protein-Konzentration der Lösung.

Zunächst wurde eine Eichreihe mit BSA hergestellt (0,01 mg/ml, 0,02 mg/ml, 0,04 mg/ml, 0,06 mg/ml, 0,08 mg/ml, 0,1 mg/ml und 0,2 mg/ml). Zur Erstellung der Eichreihe wurde das BSA mit dem bei der denaturierenden Aufreinigung verwendeten Elutionspuffer versetzt.

Zu 100 μ l Bradford-Reagenz wurden 10 μ l Proteinlösung der Probe bzw. der Eichreihe in eine 96er Assay-Platte gegeben.

Nach 10 min wurde mit einem Mikrotiterplatten-Photometer die Absorption bei 595 nm gemessen. Ausgehend von der mittels BSA erstellten Eichgerade wurden die Proteinkonzentrationen der Proben bestimmt.

2.2.6 Spotten von aufgereinigten Proteinen auf beschichtete Glas-Microarrays (FASTTM-Slides)

Zur Herstellung von Protein-Microarrays wurden die Proteine denaturierend aufgereinigt und mit einem „QArray“-Microarrayer in Duplikaten auf die Nitrozellulose ähnliche Oberfläche der FASTTM-Slides transferiert. Die Übertragung der Proteine erfolgte mit einem 16er *pin-gadget* (*pin*-Ø: 150 μ m, *pin*-Abstand: 4,5 mm). Die *pins* dieses *gadgets* wurden gleichzeitig in die Mikrotiterplatten eingetaucht.

20 μ l der aufgereinigten Proteine (Kap. 2.2.5.4) und der Kontrollen wurden in eine 384er Mikrotiterplatte überführt. Die Luftfeuchtigkeit im Inneren des Roboters wurde auf 70% eingestellt. Der Roboter taucht mit seinen *pins* in die mit Proteinlösung gefüllten *wells* der 384er Platte und transferiert die an den *pins* verbliebenen Proteinlösungen anschließend nach einem definierten Pattern auf die Microarraysubstrate. Zwischen dem Aufbringen von verschiedenen Proteinen wurden die *pins* in der Waschstation des Roboters mit destilliertem Wasser und 80% Ethanol intensiv gewaschen und anschließend mit Druckluft getrocknet.

2.2.6.1 Herstellung eines Test-Microarrays

Bei dem Test-Microarray mit 96 aufgereinigten Proteinen des Uniklonsets der ATM1 cDNA-Bibliothek (Kap. 2.2.3) wurden die Proteine in einem 4 x 4-Pattern in horizontalen Duplikaten auf die Microarrays gebracht. Alle Proteine wurden für diesen Prototypen in ein Feld gespottet (Abb. 2.7). Der Abstand zwischen den Spots beträgt bei einem 4 x 4-Pattern 1.125 μ m.

96	96	95	95	94	94	93	93	92	92	91	91	leer	leer	leer	Leer
90	90	89	89	88	88	87	87	86	86	85	85	leer	leer	leer	Leer
84	84	83	83	82	82	81	81	80	80	79	79	leer	leer	leer	Leer
78	78	77	77	76	76	75	75	74	74	73	73	leer	leer	leer	Leer
72	72	71	71	70	70	69	69	68	68	67	67	leer	leer	leer	Leer
66	66	65	65	64	64	63	63	62	62	61	61	leer	leer	leer	Leer
60	60	59	59	58	58	57	57	56	56	55	55	leer	leer	leer	Leer
54	54	53	53	52	52	51	51	50	50	49	49	leer	leer	leer	Leer
48	48	47	47	46	46	45	45	44	44	43	43	leer	leer	leer	leer
42	42	41	41	40	40	39	39	38	38	37	37	leer	leer	leer	leer
36	36	35	35	34	34	33	33	32	32	31	31	leer	leer	leer	leer
30	30	29	29	28	28	27	27	26	26	25	25	leer	leer	leer	leer
24	24	23	23	22	22	21	21	20	20	19	19	a	a	b	b
18	18	17	17	16	16	15	15	14	14	13	13	c	c	d	d
12	12	11	11	10	10	9	9	8	8	7	7	e	e	f	f
6	6	5	5	4	4	3	3	2	2	1	1	g	g	h	h

Abb. 2.7: Schema des 4 x 4-Pattern zum Spotten von 96 zufällig gewählten Proteinen des Uniklonsets für den Test-Microarray. Die 96 Proteine wurden in einem Feld in horizontalen Duplikaten gespottet und sind mit Proteinnummern (Zahlen in den Kästchen) gekennzeichnet. Die Kontrollen waren: a) denaturierender Elutionspuffer; b) PBS; c) Kaninichen-anti-Maus-IgG-Cy3-Konjugat, 1:25 in PBS verdünnt; d) BSA, c = 20 pmol/μl in PBS; e) Kaninichen IgG, 1:10 in PBS verdünnt; f) Ratten IgG, 1:10 in PBS verdünnt; g) Maus anti-RGS His₆-Antikörper, 1:10 in PBS verdünnt und h) ein Maus IgG; 1:10 in PBS verdünnt.

2.2.6.2 Spotten der Proteine des Erweiterten Uniklonset

Die aufgereinigten Proteine des gesamten Erweiterten Uniklonsets der ATM1 cDNA-Bibliothek wurden auf FASTTM-Slides transferiert.

Das Uniklonset bestand aus fünf 384er Mikrotiterplatten (Plattennummer: 311-315), wovon die letzte nur halb befüllt war. Die Plattenpositionen einer 384er Mikrotiterplatte wurden entsprechend dem 16 *pin-gadget* 24 Koordinaten zugeordnet. In Abb. 2.8 ist die Zuordnung für Platte 311 als Beispiel dargestellt. Diese Platte begann mit der Koordinate 1 und endete mit der Koordinate 24. Platte 312 wurden die Koordinaten 25 bis 48 zugeordnet, Platte 313 die Koordinaten 49 bis 72 und Platte 314 die Koordinaten 73 bis 96. Die letzte Platte 315, welche nur halb befüllt war, begann mit der Koordinate 97 und endete mit der Koordinate 108.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
A																									
B																									
C																									
D				1				5				9				13				17				21	
E																									
F																									
G																									
H				2				6				10				14				18				22	
I																									
J																									
K																									
L				3				7				11				15				19				23	
M																									
N																									
O																									
P				4				8				12				16				20				24	

Abb. 2.8: Zuordnung der 24 Koordinaten zu einer 384er Mikrotiterplatte (Plattenummer: 311) des Unikonsets der ATM1 cDNA-Expressionsbibliothek, basierend auf einem 16 *pin-gadget*. Die Koordinaten dieser Platte begannen mit 1 und endeten mit 24.

Die Kontrollen wurden in eine weitere 384er Mikrotiterplatte (Plattenummer: 316) pipettiert (Abb. 2.9). Dabei wurden die *guide dots* (Koordinate 121: [Myelin-assoziiertes basisches Protein (MBP)] = 2.000 ng/μl, versetzt mit anti-RGS His₆ Antikörper, in einer Endkonzentration von 1:10) bzw. die Kontrollen der Koordinate 122 - 127 in jedes Unterfeld gespottet, während die verschiedenen Kontrollen der Koordinate 128 pro Feld dreimal vorkamen. Damit die *guide dots* sowohl für die Phosphorylierungsstudien als auch für das Antikörper-*screening* verwendet werden konnten, wurden in die entsprechenden *wells*, sowohl MBP (artifizielles Substrat der Kinase), als auch anti-RGS His₆-Antikörper (Detektion der His₆-getagten Proteine) pipettiert.

	1	2	3	4	5	6	7	8	9	23	24
A											
B	MBP	C =	2000	ng/ μ l	MBP	C =	125	ng/ μ l			
C	Maus	anti-	RGS	His ₆							
D				121				125			
E											
F	MBP	C =	1000	ng/ μ l	denat.	Puffer					
G											
H				122				126			
I											
J	MBP	C =	500	ng/ μ l	nicht	befüllt					
K											
L				123				127			
M					a ₁	B ₁	C ₁	d ₁			
N	MBP:	c =	250	ng/ μ l	e ₁	A ₂	B ₂	c ₂			
O					d ₂	E ₂	A ₃	b ₃			
P				124	c ₃	D ₃	E ₃	128			

Abb. 2.9: Schematische Darstellung des Pipettierschemas für die Kontrollplatte (Plattennummer: 316). Die Kontrollen waren: Koordinate 121: [MBP] = 2.000 ng/ μ l, versetzt mit anti-RGS His₆ Antikörper, in einer Endkonzentration von 1:10; 122: [MBP] = 1.000 ng/ μ l; 123: [MBP] = 500 ng/ μ l; 124: [MBP] = 250 ng/ μ l; 125: [MBP] = 125 ng/ μ l; 126: denaturierender Elutionspuffer; 127: nicht befüllt. In Koordinate 128 wurden verschiedene Kontrollen in Triplets vorgelegt: a₁ - a₃) destilliertes Wasser; b₁ - b₃) PBS; c₁ - c₃) BSA, c = 20 pmol/ μ l; d₁ - d₃) Maus anti-RGS His₆ Antikörper, 1:10 in PBS verdünnt; e₁ - e₃) Kaninchen anti-Maus IgG3-Cy3, 1:25 in PBS verdünnt.

Die Proteine des Erweiterten Uniklonsets der ATM1 cDNA-Bibliothek wurden in einem 11 x 11-Pattern, und mit einem Spotabstand von 410 μ m gespottet. Es wurden auf jeden Microarray zwei identische Felder gespottet, wobei sich in jedem Feld ein Spot pro Protein befand. Zur Erleichterung der Identifizierung und der späteren Auswertung wurden, wie bei den Proteinfiltren (Kap. 2.2.2.3), *guide dots* gespottet. Allerdings wurden diese im Falle der Microarrays nicht in das Zentrum gespottet, sondern an die Seiten eines jeden Unterfeldes (Abb. 2.10). Nach dem Spotten konnten die Microarrays bei 4°C gelagert oder sofort für weitere Versuche verwendet werden.

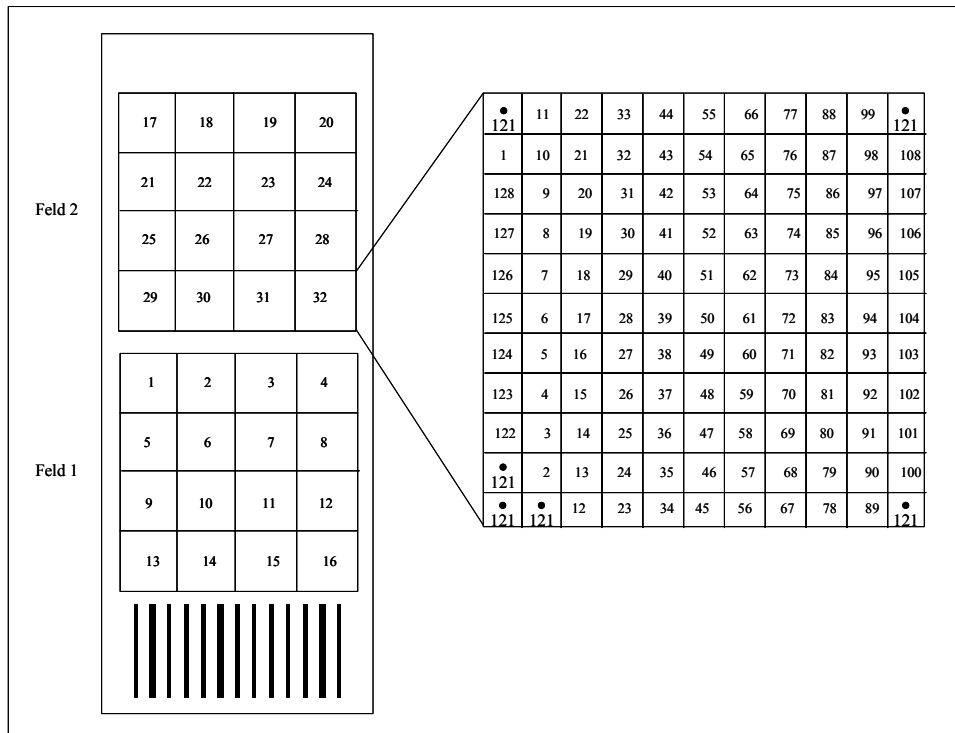


Abb. 2.10: 11 x 11 Pattern zum Spotten der Proteine des Unikonsets der ATM1 cDNA-Bibliothek. Pro Microarray wurden zwei identische Felder gespottet. Jedes Feld bestand aus 16 Unterfeldern. Pro Unterfeld wurde von den Koordinaten 1 – 108 je ein Protein gespottet, d.h. in jedem Feld befand sich jeweils ein Protein des Unikonsets. Die Koordinate 121 wurde je Unterfeld 6 x gespottet und repräsentiert die *guide dots* [MBP] = 2.000 ng/μl, versetzt mit anti-RGS His₆ Antikörper, in einer Endkonzentration von 1:10). Des weiteren wurden die Koordinaten 122 – 127 je Unterfeld einmal gespottet. Die Koordinaten 122 – 125 setzen sich aus einer Verdünnungsreihe des MBPs zusammen (122: [MBP] = 1.000 ng/μl; 123: [MBP] = 500 ng/μl; 124: [MBP] = 250 ng/μl; 125: [MBP] = 125 ng/μl). Die Koordinate 126 enthält den denaturierende Elutionspuffer und die Koordinate 127 wurde nicht befüllt. Die Koordinate 128 setzt sich aus verschiedenen Kontrollen (vgl.: Abb. 2.10) zusammen, von denen jede pro Feld dreimal gespottet wurde.

2.2.6.3 Antikörper-*screening* der gespotteten Protein-Microarrays

Zunächst wurden die Microarrays für 45 min mit TBST gewaschen. Danach folgte ein einstündiges Blockens mit 2% (w/v) BSA in TBST. Um ein Austrocknen der Microarrays bei der Antikörper-Inkubation zu vermeiden, wurden die Microarrays auf TBST-getränkte Whatman-Papiere gelegt. Von dem Erstantikörper (anti-RGS His₆, 1:2.000 in 2% (w/v) BSA in TBST verdünnt) wurden 200 μl auf die Microarrayoberfläche pipettiert, mit einem Deckgläschen (24 x 60 mm) abgedeckt und für 1 h inkubiert. Anschließend wurden die Microarrays zweimal 15 min in TBST gewaschen und dann für 1 h mit einem fluoreszenzmarkierten Zweitantikörper (anti-Maus-Cy3-Konjugat, 1:800 in 2% (w/v) BSA in TBST verdünnt) ebenfalls mit 200 μl je Microarray inkubiert. Mit Hilfe der Fluoreszenzmarkierung kann der gebundene Zweitantikörper auf dem Microarray nachgewiesen werden. Wegen der Lichtempfindlichkeit des Farbstoffs wurden ab diesen letzten Schritt und alle folgenden Schritte lichtgeschützt durchgeführt. Nach drei weiteren

Wachschritten mit PBST (je 15 min) wurden die Microarrays kurz in destilliertem Wasser geschwenkt, anschließend getrocknet und mit einem Laserscanner bei 532 nm (Cy3) gescannt. Alle Inkubations- und Waschschrte wurden bei Raumtemperatur durchgeföhrt.

Die Auswertung der Spots der eingescannten Protein-Microarrays erfolgte mit dem Programmen „Aida Array Metrix“ (Version 3.44) und der Vergleich der Quantifizierung „Aida Array Compare“ (Version 3.53.47) (Aida: Advanced Image Data Anlyzer).

Um sicherzugehen, dass der Zweitantikörper spezifisch an den Erstantikörper bindet und nicht evtl. unspezifisch an die gespotteten Proteine, wurde bei den durchgeföhrt Microarray-Versuchen als Kontrolle jeweils ein Microarray nur mit dem Zweitantikörper inkubiert.

Das Prinzip der Detektion von Fluoreszenz markierten Antikörpern auf Microarrays basiert auf der Anregung mit einer für den Farbstoff passenden Anregungswellenlänge (Excitation) des Farbstoff markierten Antikörpers mit Hilfe eines Laserstrahls. Der Farbstoff emittiert daraufhin Licht mit einer längeren Wellenlänge (Emissions-Wellenlänge). Das entstehende Licht wird dabei Punkt für Punkt über den gesamten Microarray gemessen und durch die Ansteuersoftware und Elektronik des Microarrayscanners zu einem Bild in 16 Bit *tif Format umgesetzt.

2.2.6.4 Kinase-Assays auf den Protein-Microarrays des Erweiterten Uniklosets mit radioaktivem γ -ATP

Zunächst wurden die Microarrays für 45 min mit TBST gewaschen. Danach folgte ein einstündiges Blocken mit 2% BSA (w/v) in TBST. Um ein Austrocknen der Microarrays bei der Inkubation mit den entsprechenden Kinasen zu vermeiden, wurden die Microarrays auf TBST-getränkte Whatman-Papiere gelegt. Die Reaktionsansätze (250 μ l) der entsprechenden Kinaselösungen Tab. 2.3 und Tab. 2.4 wurden auf die Microarrayoberfläche pipettiert, mit einem Deckgläschen abgedeckt und für 30 min inkubiert.

Tab. 2.3: Reaktionsansätze der entsprechenden MAP-Kinasen, mit denen die Microarrays inkubiert wurden.

	MPK3	MPK6	Negativkontrolle
Kinaselösung	25 µl (c = 100 ng/µl)	25 µl (c = 200 ng/µl)	-
5 x Kinasepuffer	50 µl	50 µl	50 µl
ddH ₂ O	172,5 µl	172,5 µl	197,5 µl
γ-ATP (³³ P) (250 µCi/ml)	2,5 µl	2,5 µl	2,5 µl

In dieser Tabelle sind die Reaktionsansätze der entsprechenden MPKs, mit denen die Microarrays inkubiert wurden, dargestellt. MPK = Mitogen-aktivierte Proteinkinase.

Tab. 2.4: Reaktionsansatz der PKA, mit dem die Microarrays inkubiert wurden.

	PKA
Kinaselösung	5 µl (c = 12,5 µl/ml)
10 x Kinasepuffer	25 µl
ddH ₂ O	217,5 µl
γ-ATP (³³ P) (250 µCi/ml)	2,5 µl

In dieser Tabelle ist der Reaktionsansatz der PKA, mit dem die Microarrays inkubiert wurden, dargestellt. PKA = Proteinkinase A.

Anschließend wurden die Microarrays wie folgt gewaschen:

2 x 15 min	2,0 x PBST
2 x 15 min	1,0 x PBST
1 x 30 min	1,0 x PBST
1 x 30 min	0,5 x PBST
1 x 30 min	0,1 x PBST
1 x 30 min	0,1 x PBS

Den Waschschritten folgend wurden die Microarrays getrocknet, mit Saranfolie abgedeckt und in eine Spezialkassette für *Imaging plates* gelegt. Radioaktiv phosphorylierte Proteine wurden auf den Microarrays mittels Phosphorimager detektiert.

2.2.6.5 Detektion und Quantifizierung der potentiellen Phosphorylierungs-Targets

Im Rahmen dieser Arbeit erfolgte die Analyse der Microarrays nach der Kinasierung unter Verwendung eines BioImagers. Die Positionen und der Abstand der phosphorylierten Proteinspots auf den Microarrays wurden mit Hilfe einer *super resolution (SR) Imaging plate* in Verhältnis 1:1 detektiert. Dazu wurde die generell zuvor gelöschte *Imaging plate* in den MPK3-Versuchen 6 bzw. 12 h, im MPK6-Versuch 48h und im PKA-Versuch 12 h auf die Microarrays gelegt. Anschließend wurde die *Imaging plate* bei 640 nm mit einer Auflösung von 10 μm auf dem Bioimager eingescannt.

Die gewonnenen *Images* wurde anschließend mit Hilfe der Programme „Aida Array Metrix“ und „Aida Array Compare“ ausgewertet.

„Aida Array Metrix“ ermöglicht die Quantifizierung der Spots eines Microarrays. „Aida Array Compare“ kann anschließend verwendet werden, um die ermittelten quantitativen Daten mehrerer Microarrays zu vergleichen. Programm-Parameter, auf die im Folgenden nicht detaillierter eingegangen wird, wurden auf die im Programm vorgesehenen Standardwerte belassen.

In „Aida Array Metrix“ wurde das eingescannte Image der *Imaging plate* geöffnet. Anschließend wurden die dazugehörige Konfigurationsdatei (*GAL file*) importiert. Der *GAL file (Gene Pix Array List)* ist ein von der Firma Axon entwickeltes Dateiformat und ermöglicht die Zuordnung der Spotpositionen zu den Plattenkoordinaten. Der *GAL file* wird von der Kontrollsoftware des Microarrayers erstellt. Nach Import des *GAL files* in „Aida Array Metrix“ werden die Positionsangaben der Spots schematisch in einem *grid* dargestellt. Dieses *grid* wird anhand der *guide dots dots* auf dem Image ausgerichtet. Im *GAL file* Format ist jedoch keine Definition der Replikatpositionen enthalten. Diese relative Anordnung der Wiederholungen eines Proteins auf den Microarrays wurde anhand des Spot-Patterns in die Definition des *grids* in „Aida Array Metrix“ entsprechend dem Manual festgelegt.

Für alle Spots wurde ein einheitlicher Durchmesser von 320 μm zur Quantifizierung festgelegt. Anschließend wurde der *mode of non spot* (mit einen um 4 μm vergrößerten Spotdurchmesser) zur Ermittlung des lokalen Hintergrundsignals innerhalb eines jeden Unterfelds gewählt. Diese Definitionen wurde in einem *grid*-Template (*.atf Format) abgespeichert. Dieses Template wurde für alle *Image*-Quantifizierungen verwendet. Die endgültige Anordnung der zu qualifizierenden Spots (*region of interest*) wurde mit Hilfe eines implementierten Algorithmus optimiert.

Mit den so definierten Werten wurden alle Spots des Microarrays quantifiziert und anschließend die Signalintensitäten der beiden Spots eines Proteins miteinander verglichen. Dazu wurden Signalintensitäten aller Duplikate in einem Scatterplot in „Aida Array Matrix“ gegeneinander aufgetragen. Mit Hilfe dieser Duplikatkorrelations-Plots wurden die Datenpunkte, die mehr als 25% von ihrem Mittelwert abwichen, manuell markiert (flaggen) und somit von der weiteren Auswertung ausgeschlossen. Die Qualität der Spots kann durch verschiedene Faktoren beeinflusst werden, wie z.B. Artefakte aufgrund radioaktiver Präzipitate oder Überstrahlungseffekte. Des Weiteren wurden auch die zur Ausrichtung des *grids* verwendeten *guide dots* von der Auswertung ausgeschlossen.

Zusätzlich wurden alle Datenpunkte, die von „Aida Array Matrix“ automatisch, aufgrund mangelnder Homogenität, als nicht auszuwertende Spots markiert wurden, manuell kontrolliert und gegebenenfalls der Ausschluss der Spots von der Analyse rückgängig gemacht. Die gesamte Auswertung (*grid*, Duplikatkorrelation, Hintergrundkorrektur etc.) wurde in dem von Raytest entwickelten *.adf Format gespeichert, wodurch das eingescannte Ausgangsbild unverändert bleibt.

Die *.adf Dateien von mehreren Microarrays unterschiedlicher Versuchsreihen wurden in „Aida Array Compare“ geladen und die potentiellen Phosphorylierungs-Targets des Erweiterten Unikonsets der entsprechenden Kinasen identifiziert. Die Kriterien für die Selektion der Proteine waren folgende:

Von zwei Microarrays eines Experimentes wurden nur die Proteine in die Endauswertung einbezogen, bei denen die hintergrundkorrigierten, gemittelten Intensitäten mit einer Reproduzierbarkeit von einem Faktor 2 bestimmt wurden. Des Weiteren gingen nur die Proteine in die Endauswertung ein, deren hintergrundkorrigierten Intensitäten über der 10fachen Standardabweichung des mittleren Hintergrundsignals aller Unterfelder des Microarrays ($Hg - 10 \times \delta$) lagen. Nur Proteine, die von beiden Microarrays eines Experimentes in die Endauswertung gingen, wurden als potentielle Phosphorylierungs-Targets identifiziert. Dazu wurde die Schnittmenge dieser Proteine beider Microarrays gebildet. Im Falle der Durchführung von mehreren Versuchen wurden nur jene Proteine als potentielle Phosphorylierungs-Targets bezeichnet, die in allen Experimenten mit einer Kinase als Targets bestimmt wurden (erneute Schnittmengenbildung).

Um die ermittelten potentiellen Phosphorylierungs-Targets mit einer unabhängigen Methode zu bestätigen, wurden diese erneut exprimiert und aufgereinigt. Anschließend wurden die aufgereinigten Proteine unter Verwendung 15%iger SDS-Polyacrylamidgelelelektrophoretisch aufgetrennt und mittels Blot auf PVDF-Membranen transferiert

(Kap. 2.2.5.1 und 2.2.5.2). Anschließend erfolgte die enzymatische Phosphorylierung der Proteine auf den Blots (Tab. 2.5 und Tab. 2.6).

Die Phosphorylierung und alle nachfolgenden Schritte erfolgten analog dem Phosphorylierungsprotokoll der Protein-Microarrays (siehe Beginn dieses Kapitels) in einem angepassten vergrößertem Volumen.

Tab. 2.5: Reaktionsansatz der MPK3 mit dem die Blots inkubiert wurden.

	MPK3
Kinaselösung	400 µl (c = 100 ng/µl)
5 x Kinasepuffer	800 µl
ddH ₂ O	2760 µl
γ-ATP (³³ P) (250 µCi/ml)	40 µl

In dieser Tabelle ist der Reaktionsansatz der MPK3, mit denen die Blots inkubiert wurden, dargestellt. MPK3 = Mitogen-aktivierte Proteinkinase 3.

Tab. 2.6: Reaktionsansätze der PKA, mit dem die Blots inkubiert wurden.

	PKA
Kinaselösung	80 µl
10 x Kinasepuffer	400 µl
ddH ₂ O	3480 µl
γ-ATP (³³ P) (250 µCi/ml)	40 µl

In dieser Tabelle ist der Reaktionsansatz der PKA, mit dem die Blots inkubiert wurden, dargestellt. PKA = Proteinkinase A.

Zusätzlich wurden die Sequenzen der potentiellen Phosphorylierungs-Targets der entsprechenden Kinasen hinsichtlich bekannter Konsensussequenzen und ihren Serin- und Threonengehalt untersucht.

Ein Problem hierbei war, dass die Klone ausgehend von ihrem 5'-Ende sequenziert wurden und somit nicht die vollständige Sequenzinformation der Klone vorlagen. Da die cDNAs unter Verwendung eines Oligo(dT)-Primers synthetisiert wurden, und so von einem vollständigem 3'-Ende ausgegangen werden konnte, wurden die vollständigen Sequenzinformationen der MIPS-Datenbank hinzugezogen. Von diesen vollständigen Datenbank-Sequenzen wurde das gegebenenfalls fehlende 5'-Ende abgezogen. So wurden

die tatsächlich ermittelten Informationen der 5`-Enden und die durch die Homologiesuche hinzugezogenen Informationen des 3`-Endes für die Motivsuche bzw. die Bestimmung des Serin- und Threoningehaltes genutzt (HTML-Tabelle 2: http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionUniCloneSet.shtml, Start Subject > 1).

3 Ergebnisse

3.1 Herstellung der cDNA-Bibliotheken

3.1.1 Verwendete Strategien

Es wurden zwei *Arabidopsis thaliana* cDNA-Bibliotheken, aus den Geweben Pistill und Infloreszenz, synthetisiert. Die cDNA-Synthesen erfolgten für beide cDNA-Bibliotheken mittels Oligo(dT)-*NotI*-Primer (3'-Ende) und *SalI*-Adapter (5'-Ende). Anschließend wurden die cDNA-Bibliotheken in unterschiedliche *E. coli* Vektorsysteme mittels restriktionsabhängiger Klonierung (*SalI/NotI*) transformiert (vgl. 2.2.1).

Die cDNA-Bibliothek von Pistill wurde in einem Entryvektor (pENTR1a) des GATEWAYTM-Klonierungssystems (Invitrogen) konstruiert. Ausgehend vom Entryvektor können die klonierten cDNAs mittels sequenzspezifischer DNA-Rekombination, die durch spezifische Rekombinasen katalysiert wird, in verschiedene GATEWAYTM-Expressionsvektoren überführt werden.

Nach der Klonierung der Pistill-cDNAs in den Entryvektor, der Transformation in *E. coli* DH5 α -Zellen und anschließendem Ausplattieren der Zellen auf den Anzuchtplatten, wurden alle Klone vereinigt und Plasmid-Präparationen durchgeführt (vgl. 2.2.4.1). Die vereinigten Plasmid-DNAs wurden eingefroren und stehen somit als Ressource für weitere Studien, wie u.a. für die unseres Kooperationspartners (Dr. Richard Immink und Prof. Dr. Gerco Angenent, Plant Research International, Niederlande) zur Verfügung.

Die zweite cDNA-Bibliothek, aus dem Infloreszenzgewebe, wurde im Rahmen dieser Arbeit als eine geordnete cDNA-Expressionsbibliothek direkt in einem *E. coli* Expressionsvektor hergestellt. Der verwendete Vektor pQE30NASTattB ermöglicht die Expression der Proteine, fusioniert an einen N-terminalen RGS His₆-Tag (vgl. 2.2.1). Die Transformation der Konstrukte erfolgte hier in *E. coli* SCS1/pSE111-Zellen. Nach der Wachstumsphase auf den Anzuchtplatten wurden die Klone gepickt und in medienbefüllten 384er Mikrotiterselektionsplatten parallel kultiviert. Von der so erhaltenen geordneten cDNA-Bibliothek erfolgte anschließend das *screening* auf potentielle Expressionsklone unter Verwendung von Hochdicht-Proteinfiltren. Die potentiellen Expressionsklone wurden anschließend in einer cDNA-Expressionsbibliothek neu angeordnet (vgl. 2.2.2 und Kap. 3.3). Von dieser Proteinexpressions-Unterbibliothek wurde dann auf der Grundlage der 5'-Sequenzen ein Uniklonset zusammengestellt (vgl. 2.2.3 und Kap.3.4).

Die im Hochdurchsatzformat durchgeführten Schritte, wie das Picken der Klone, das Spotten der gepickten Klone auf Membranen sowie das Transferieren der potentiellen Expressionsklone von einer geordneten cDNA-Bibliothek in eine cDNA-Expressionsbibliothek und das Transferieren der selektierten Klone von der cDNA-Expressionsbibliothek in ein Unikonset wurde mit Hilfe von für diese Zwecke entwickelten Robotern (Büssow *et al.*, 1998; Lehrach *et al.*, 1997) durchgeführt.

3.1.2 cDNA-Synthese

Die cDNAs der beiden Gewebe Pistill und Infloreszenz wurden unter Verwendung des cDNA Superscript Plasmid Systems synthetisiert (vgl. 2.2.1.1 und 2.2.1.2). Für beide wurde als Ausgangsmaterial Poly-A⁺RNA (mRNA) verwendet.

Nach der Erststrangsynthese unter Verwendung eines Oligo(dT)-*NotI*-Primers wurden die beiden Produkte mit Hilfe von 1,0%igen (w/v) Agarosegelen gelelektrophoretisch aufgetrennt (vgl. 2.2.4.4). Die synthetisierten cDNAs waren als ein „Schmier“ im Agarosegel mit einer kontinuierlichen Größenverteilung im Bereich von ungefähr 350 Bp – 4.000 Bp erkennbar.

Nach der Zweitstrangsynthese erfolgte eine Größenfraktionierung mittels Sepharose-Säulen-Chromatographie. Auf diese Weise wurden, neben überschüssigen Adaptoren, Fragmente, die kleiner als 500 Bp waren, eliminiert. Zu diesem Zweck wurden die Volumina der einzelnen Fraktionen bestimmt (Tab. 3.1). Entsprechend den Angaben des Herstellers lassen sich DNA-Fragmente größer als 500 Bp gewinnen, wenn ein *cutoff*-Wert von 550 µl Gesamtvolumen angesetzt wird und die vier bis fünf Fraktionen (je Fraktion ungefähr 35 µl) unterhalb des *cutoff*-wertes vereinigt und verwendet werden.

Bei den Pistill-cDNAs umfasste diese Vereinigung die Fraktionen 10 - 14 (149 µl) und bei der Infloreszenz-cDNAs die Fraktionen 9 - 12 (145 µl). Anschließend wurden die vereinigten Fraktionen gefällt und standen damit der Ligation mit dem für sie vorgesehenen Vektor zur Verfügung.

Tab. 3.1: Größenfraktionierung der Pistill- und der Infloreszenz- cDNAs mittels Säulenchromatographie. Die Bestimmung der Gesamtvolumina diente der Ermittlung des *cutoff* Wertes (550 μ l). 4 bzw. 5 Fraktionen unter dem *cutoff* Wert wurden verwendet (entsprechend den Angaben des Herstellers). Die grün bzw. rot markierten Bereiche stellen die Fraktionen dar, die für die Herstellung der entsprechenden cDNA-Bibliotheken eingesetzt wurden.

Pistill			Infloreszenz		
Nr.	Fraktionsvolumen (μ l)	Gesamtvolumen (μ l)	Nr.	Fraktionsvolumen (μ l)	Gesamtvolumen (μ l)
1	82	82	1	96	96
2	108	190	2	100	196
3	36	226	3	35	231
4	32	258	4	28	259
5	19	277	5	36	295
6	29	306	6	37	332
7	38	344	7	29	361
8	26	370	8	35	396
9	30	400	9	36	432
10	20	420	10	36	468
11	35	455	11	36	504
12	21	476	12	37	541
13	35	511	13	36	577
14	38	549	14	38	615
15	35	584	15	37	652
16	35	619	16	32	684
17	39	658	17	37	721
18	39	697	18	32	753
19	28	725	19	37	790
20	26	751	20	32	822

In der Tabelle ist die Größenfraktionierung der Pistill- und der Infloreszenz- cDNAs mittels Säulenchromatographie dargestellt.

3.1.3 Vektorpräparation

Die Vektoren pENTR1a (Pistill; <http://www.invitrogen.com/>) bzw. pQE30NASTattB (Infloreszenz; GenBank-Accession-Nummer: AY386205) wurden zunächst mit den Enzymen *SalI* und *NotI* geschnitten, wobei der Vektor pQE30NASTattB ein über *SalI* und *NotI* kloniertes Insert (*phyB*) besass (vgl. 2.2.1.3). Das heißt, für diesen Vektor waren nach einer vollständigen Restriktion zwei Fragmente in dem 0,8%igen (w/v) Agraosegel zu erkennen (pQE30NASTattB: 3.573 Bp und *phyB*: 1.665 Bp). Der vollständig geschnittene Vektor ließ sich damit im Agarosegel eindeutig als Bande bei 3.573 Bp von dem einfach geschnittenen Vektor bei 5.238 Bp unterscheiden. Zudem konnten, im Falle einer unvollständigen Restriktion, bzw. bei einer Religation des Vektors (Ligationsansatz kontaminiert mit *phyB* Insert) die *PhyB*-Expressionsklone unter Verwendung des

entsprechenden anti-PhyB Pea 25 Antikörpers identifiziert werden (Cordonnier *et al.*, 1986). Der Antikörper wurde von Prof. Dr. John Hughes (Freie Universität Berlin) zur Verfügung gestellt.

Der Vektor pENTR1a besass dagegen kein Insert. Jedoch sollte hier die Bildung von Religanden aufgrund der Aktivität des sich in der *multiple cloning site* des Vektors befindlichen *ccdB*-Gens minimiert werden. Das *ccdB*-Gen ist ein „Suizid-Gen“, dessen Genprodukt mit der *E. coli*-DNA-Gyrase interferiert (Bernard und Couturier, 1992). Die Selektion basiert auf dem Austausch des einzubringenden Gens mit dem *ccdB*-Gen. Nur Zellen, die Plasmide ohne *ccdB*-Gen, aber mit der entsprechenden Resistenz tragen, können Kolonien bilden, für die anderen Zellen wirkt das *ccdB*-Genprodukt letal.

Die geschnittenen Vektoren (pENTR1a bei 2.717 Bp und pQE30NASTattB bei 3.575 Bp) wurden aus dem Agarosegel eluiert und für die Ligationen mit den jeweiligen cDNA-Fractionen der cDNA der entsprechenden Gewebe verwendet (Kap. 3.1.2).

3.1.4 Ligation, Transformation beider cDNA-Bibliotheken und Anordnung der Infloreszenz cDNA-Bibliothek

Die cDNA-Fractionen der unterschiedlichen Gewebe wurden mit dem jeweiligen präparierten Vektor GATEWAY-Entryvektor pENRT1a (Pistill) bzw. Proteinexpressionsvektor pQE30NASTattB (Infloreszenz) ligiert (vgl. 2.2.1.4). Darauffolgend wurden die Ligationsgemische der jeweiligen cDNA-Fractionen mittels Elektroporation in die entsprechenden kompetenten Zellen transformiert (vgl. 2.2.1.7). Die verwendeten Zellen waren DH5 α für Pistill bzw. SCS1/pSE111 für Infloreszenz, wobei die Transformationseffizienz der DH5 α -Zellen bei $\geq 1 \times 10^8$ koloniebildende Einheiten (KBE)/ μg pUC 18 und der SCS1/pSE111-Zellen bei $\geq 3,7 \times 10^7$ KBE/ μg pUC 18 lag. Als Kontrolle wurden Religationsansätze der geschnittenen Vektoren ohne cDNA mitgeführt. Die ermittelten Religanden lagen bei dem pENTR1a-Vektor bei 5,0% und bei dem pQE30NASTattB-Vektor bei 6,5%.

Die Transformationsansätze wurden nach der Regenerationszeit auf Anzuchtplatten ausplattiert und inkubiert. Für die Pistill cDNA-Bibliothek wurden auf diese Weise 660.800 Klone und für die Infloreszenz cDNA-Bibliothek 40.000 Klone gewonnen.

Um von der Infloreszenz eine geordnete Proteinexpressions-Bibliothek zu erstellen, wurden alle 40.000 Klone der cDNA-Bibliothek nach der Wachstumsphase auf den Anzuchtplatten mit Hilfe des Pick-Roboters „Mercury“ gepickt und in medienbefüllten 384er Mikrotiterselektionsplatten parallel kultiviert. Auf diese Weise wurde eine geordnete

cDNA-Bibliothek hergestellt (Plattenummer 1-103, 106-108), die dann als Grundlage für die Erstellung einer geordneten Expressionsbibliothek (Kap. 3.3), sowie eines Uniklonsets (Kap. 3.4) diene.

3.2 Charakterisierung der cDNA-Bibliotheken

3.2.1 Bestimmung der durchschnittlichen Insertlängen

Zur Beurteilung der Qualität beider cDNA-Bibliotheken wurden jeweils 96 Klone gepickt und mittels der entsprechenden Vektorprimer Kolonie-PCRs durchgeführt (vgl. 2.2.4.3). Die Vektorprimer binden dabei flankierend zu den Klonierungsstellen. Für den Vektor pENTR1a ergibt sich ein PCR-Amplifikat ohne Insert von 850 Bp, für den Vektor pQE30NASTattB ergibt sich ein PCR-Amplifikat ohne Insert von 318 Bp, für potentielle Religanden von pQE30NASTattB-phyB ein PCR-Amplifikat von 2.000 Bp.

3.2.1.1 Pistill cDNA-Bibliothek

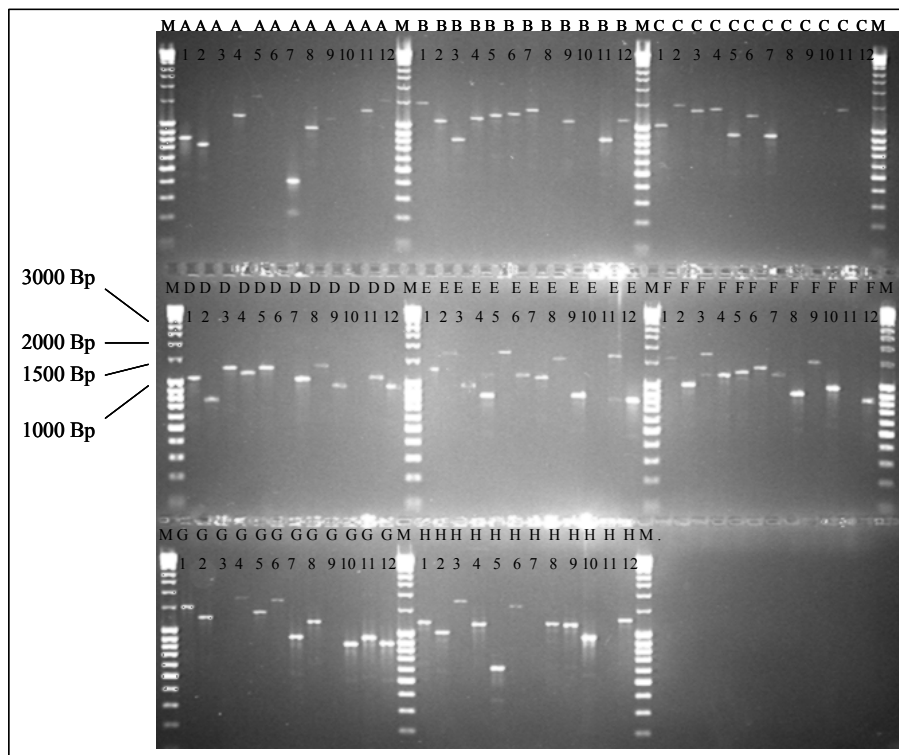


Abb. 3.1: Insertlängenvergleich von 95 zufällig gepickten Klonen aus der Pistill cDNA-Bibliothek. 10 µl PCR-Produkt aus Kolonie-PCRs wurden auf einem 1,0%igen (w/v) Agarosegel elektrophoretisch aufgetrennt. Die Plattenpositionen der 96 *well* PCR-Platten sind oberhalb der Laufspur zu entnehmen. M: Marker, Bp= Basenpaare; A1= Positivkontrolle (Ausgangsvektor, pENTR1a).

Von den 95 getesteten Klonen der Pistill cDNA-Bibliothek lieferten 77 ein PCR-Produkt, welches im Agarosegel sichtbar war (Abb. 3.1). Die Insertlängen dieser Klone wurden bestimmt, indem von der jeweiligen Länge des PCR-Produktes der Vektoranteil abgezogen

wurde. Bei 5 der 95 Klone handelt es sich um potentielle Religanden, da die PCR-Produkte eine Länge von 850 Bp aufwiesen, welche der theoretischen Länge der Religanden entsprach. Darüber hinaus lieferten 18 Klone keine PCR-Produkte. Um die Insertlänge von diesen 18 Klonen zu ermitteln, wurden von diesen Klonen Plasmidpräparationen durchgeführt (vgl. 2.2.4.1). Nach *SalI*- und *NotI*-Verdau und elektrophoretischer Trennung konnte die Insertlänge von 15 dieser Klone ermittelt werden. 3 der 18 isolierten und verdauten Klone lieferten kein Insert. Bei einem der Klone (H7) könnte es sich mit einer Länge von 850 Bp wiederum um einen Religanden handeln (Tab. 3.2).

Tab. 3.2: Ermittlung der Insertlängen der Pistill-Klone anhand des Verdauens der Plasmide. Die 18 Pistill-Klone die bei der PCR kein Ergebnis zeigten, wurden mittels Plasmidpräparation, *SalI*-/*NotI*-Verdau und anschließender Gelelektrophorese kontrolliert. Für 15 der 18 Klone konnten Inserts ermittelt werden.

Koordinaten	[Bp]
A3	300
A6	1.400
A10	1.700
B8	1.300
B10	kein Insert
C8	1.300
C9	1.300
C10	1,000
C12	1.700
D6	1.500
D10	1.500
E2	1.200
E10	kein Insert
F11	1.300
G3	1.600
G9	kein Insert
H7	850
H11	1.650

In der Tabelle ist die Ermittlung der Insertlängen der 18 Pistill-Klone anhand des Verdauens der Plasmide dargestellt.

In die Bestimmung der durchschnittlichen Insertlänge wurden alle Klone außer potentielle Religanden (6,3%) und die Klone ohne Insert (drei Klone) einbezogen. Davon ausgehend wurde für die Pistill cDNA-Bibliothek eine durchschnittlich Insertlänge von 1,09 Kbp bestimmt.

3.2.1.2 Infloreszenz cDNA-Bibliothek

Im Falle der Infloreszenz cDNA-Bibliothek lieferten 94 von 96 Klonen ein PCR-Produkt (Abb. 3.2). 7 Klone wurden aufgrund der PCR-Produktlänge (2.000 Bp) als potentielle Religanden identifiziert. Von diesen Klonen erfolgte eine 2. Kolonie-PCR mit genspezifischen Primern für *phyB* (Insert im Ausgangsvektor). Dabei zeigte sich, dass von den 7 potentiellen Religanden 2 ein PCR-Produkt aufwiesen und somit als Religanden bestätigt wurden.

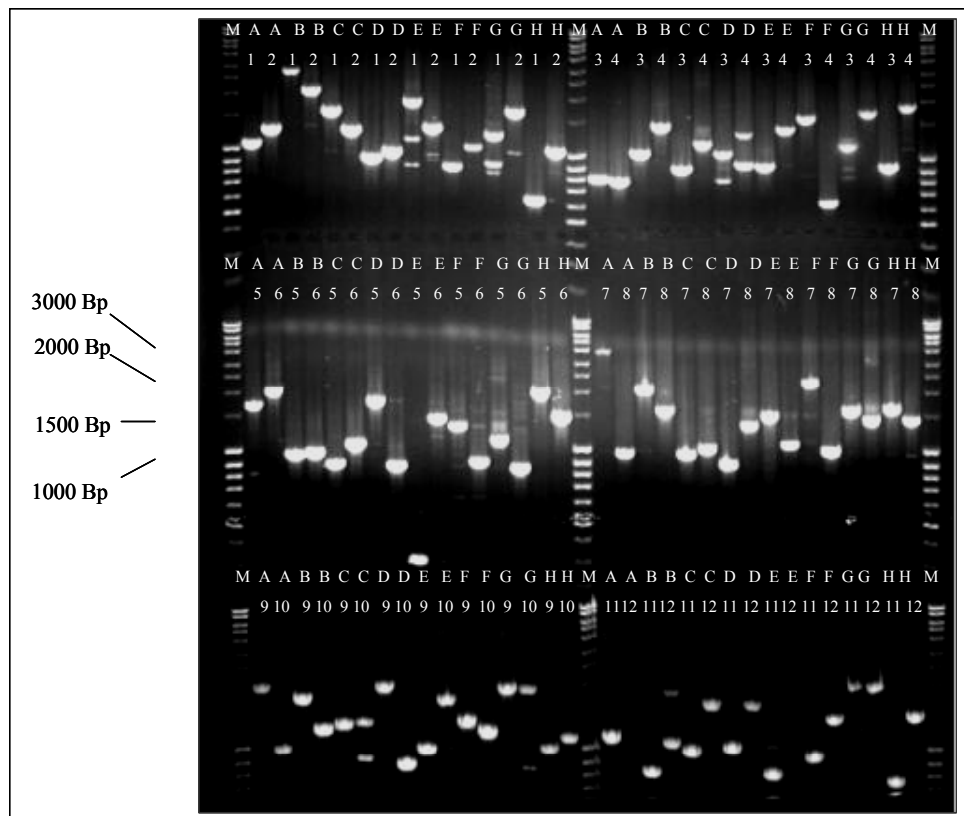


Abb. 3.2: Insertlängenvergleich von 96 zufällig gepickten Klonen aus der ATM1 cDNA-Bibliothek. 10 µl PCR-Produkt aus Kolonie-PCRs wurden auf einem 1,0%igen (w/v) Agarosegel elektrophoretisch aufgetrennt. Die Plattenpositionen der 96 well PCR-Platten sind oberhalb der Laufspur zu entnehmen. M: Marker, Bp= Basenpaare.

In die Bestimmung der durchschnittlichen Insertlänge wurden die 94 Klone abzüglich Religanden (2,1%) einbezogen. Damit wurde für die Infloreszenz cDNA-Bibliothek eine durchschnittliche Insertlänge von 1,06 Kbp ermittelt.

3.2.2 Sequenzanalysen beider cDNA-Bibliotheken

3.2.2.1 Testsequenzierung der Pistill cDNA-Bibliothek

Im Falle der Pistill cDNA-Bibliothek wurden 48 Klone gepickt und vom 5'-Ende mit Hilfe des Vektorprimers pENTR1a-5' ansequenziert (DLMBC, Berlin) (vgl. 2.2.4.2).

Mit den Sequenzen wurde eine BLAST-Analyse gegen die kodierenden Sequenzen der MIPS (*munich information center for protein sequences*) *Arabidopsis*-Datenbank durchgeführt. Das BLAST-Programm (*Basic Local Alignment Search Tool*, (Altschul *et al.*, 1990; Altschul *et al.*, 1997)), ist ein Ähnlichkeits-Suchprogramm, das mit einer spezifischen Suchsequenz die zur Verfügung stehenden Nukleotid- oder Proteinsequenz-Datenbanken nach Sequenzähnlichkeiten durchsucht. Alle 48 Pistill-Klone konnten einem bekannten AGI (*Arabidopsis* Genome Initiative) *gene code* zugeordnet werden (siehe Anhang 8.1). Es zeigte sich, dass ein Pistill-Klon, bezogen auf den AGI *gene code* zweimal vorkam (40 ribosomales Protein S16; At5g18380). Alle anderen Klone kamen einmalig vor. 11 der 47 Klone beinhalten ein vollständiges Gen, d.h. der Start des Subjekts ist gleich 1 (23,4%). Bei dem hier bezeichnet Subjekt (Treffer-Sequenz) handelt es sich um eine aus der Proteinexpressions-Unterbibliothek stammende 5'-Ende ansequenzierte Sequenz, von welcher anhand der BLAST-Analyse eine Suche auf Sequenzähnlichkeiten durchgeführt wurde. Die bei dieser Suche gefundenen ähnlichen Datenbanksequenzen werden als Query (Abfrage-Sequenz) bezeichnet. Wenn der Start des Subjekts gleich eins ist, ist der Beginn des Subjekts im Query enthalten. Das kann als ein Indiz dafür angesehen werden, dass die Reverse Transkriptase das Gen in voller Länge transkribiert hat. Darüber hinaus lagen 15 der 48 Klone im richtigen Leserahmen vor, d.h. in dem Leserahmen, in dem sich das RGS His₆-Tag befindet (31,3%) (siehe Anhang 8.1).

3.2.2.2 Testsequenzierung der Infloreszenz cDNA-Bibliothek

Im Falle der Infloreszenz cDNA-Bibliothek erfolgte eine Sequenzierung im größeren Maßstab als für die Pistill cDNA-Bibliothek. 576 (6 x 96) Infloreszenz-Klone wurden gepickt und vom 5'-Ende mittels des Oligonukleotids pQE65 ansequenziert (DLMBC, Berlin) (vgl. 2.2.4.2).

Von diesen Klonen lieferten 434 Klone eine Sequenz. 13 der 434 Sequenzinformationen wurden als nicht ausreichend befunden, da die Sequenzen zu mehr als 20% aus undefinierten Basen bestanden. Der Abgleich der restlichen 421 Sequenzen mit der MIPS-*Arabidopsis*-Datenbank ergab, dass für 3 der 421 Sequenzen keine BLAST-Hits mit

bekanntes AGI *gene codes* gefunden werden konnten. Den restlichen 418 Sequenzen konnten AGI *gene codes* zugeordnet werden (siehe Anhang 8.2).

Anhand der Tabelle (siehe Anhang 8.2) wurde festgestellt, dass es sich bei 7 von 418 Klonen (1,7%) um Religanden handelt, d.h. diese Klone haben ein *phyB*-Insert, welches im Ausgangsvektor enthalten war.

Aufgrund der Anzahl der Klone, die sequenziert wurden, war es möglich, eine erste Aussage über die Redundanz der Bibliothek zu machen. Dazu wurde das Set, bestehend aus 411 Klonen (ohne *phyB* Religanden), unter Einbeziehung des AGI *gene codes* sortiert. 297 Klone (72,3%) kamen einmal vor, und die restlichen 115 Klone lagen mehrmals vor und bildeten 49 Cluster (Tab. 3.3). Es konnten 192 der 411 Klone in voller Länge ermittelt werden (46,7%). Des Weiteren lagen 128 der 411 Klone im richtigen Leserahmen vor (31,1%).

Das größte Cluster bestand aus 9 Sequenzen, welche als hypothetische Proteine identifiziert wurden (Tab. 3.3). Beim nächst größeren Cluster handelte es sich um das Gen für *phyB*. Es stellte den Religandenanteil, bestehend aus sieben Sequenzen, dar. Außerdem gab es zwei Cluster, die sich aus je vier Sequenzen und neun Cluster, die sich aus je drei Sequenzen zusammensetzten. Cluster die sich aus je zwei Sequenzen zusammenfassen ließen, kamen 36 mal vor (Tab. 3.3).

Tab. 3.3: Ergebnisse der Cluster-Analyse der redundanten Sequenzen der Testsequenzierung der Infloreszenz cDNA-Bibliothek.

Clustergröße	Clusterfrequenz	Sequenzanzahl je Clustergruppe
2	36	72
3	9	18
4	2	8
7	1	7
9	1	9
Gesamt	49	114

In der Tabelle sind die Ergebnisse der Cluster-Analyse der redundanten Sequenzen der Testsequenzierung der Infloreszenz cDNA-Bibliothek dargestellt.

3.3 Erstellung einer Proteinexpressions-Unterbibliothek der ATM1 cDNA-Bibliothek

3.3.1 Identifizierung potentieller Expressionsklone der ATM1 Bibliothek mit Hilfe von Hochdichte-Proteinfiltern

Als potentielle Expressionsklone wurden hier Klone bezeichnet, die mit einem anti-RGS His₆-Antikörper detektiert werden konnten. Klone, deren Inserts im Leserahmen zum RGS His₆-Tag lagen, lieferten ein getagtes Protein und damit ein Signal mit dem tagspezifischen Antikörper. Klone, deren Inserts nicht im Leserahmen zum RGS His₆-Tag lagen, besitzen meist zufallsbedingte Stopcodone im Anfangsbereich des Inserts (Lüking *et al.*, 2003b). Diese Klone bilden kurze Polypeptide, welche sich häufig nicht in stabile Strukturen falten können, deshalb in den meisten Fällen von der Wirtszelle abgebaut werden und somit kein Signal liefern. Obwohl sich mit diesem Immunoassay falschpositive Klone nicht generell ausschließen lassen, kommt es zu einer deutlichen Anreicherung von Klonen, deren Insert sich im gleichen Leserahmen mit dem RGS His₆-Tag befindet.

Für die Identifizierung der potentiellen Expressionsklone wurden die 40.000 Infloreszenz-Klone der geordneten cDNA-Bibliothek ATM1 mit Hilfe des Spot-Roboters „Medusa“ auf PVDF-Membranen gespottet (vgl. 2.2.2.1). Die Klone wurden in Duplikaten in einer Dichte von 27.648 unterschiedlichen Klonen pro Filter (20 cm x 20 cm) aufgetragen. Für die 40.000 Klone wurden 2 Membranen benötigt, wobei eine Membran nur halb bespottet wurde. Nach der Proteinexpression und der Prozessierung der Filter wurden die proteinexprimierenden Klone unter Verwendung eines Detektionsantikörpers (anti-RGS His₆) in Verbindung mit einem enzymgekoppelten Antikörper (anti-Maus-AP-Konjugat) nachgewiesen. Dazu wurden die Filter unter UV-Licht mit der Fluoreszenz-Video-Dokumentation photographiert. Die Identifizierung der anti-RGS His₆ positiven Klone auf dem *Image* erfolgte anschließend unter Verwendung des Programms „Visual Grid 2.01“ (vgl. 2.2.2). Anti-RGS His₆ positive Klone sind als weiße Spots auf dem Filter zu erkennen und konnten visuell identifiziert werden (Abb. 3.3). Die Verwendung von Duplikaten jedes Klons gewährleistet, dass eventuell auftretende falschpositive Ergebnisse minimiert werden. Es wurden nur die Klone als potentiellen Expressionsklone bewertet, bei denen beide Spots eines Duplikats detektiert werden konnten. So konnten 7.000 potentielle Proteinexpressionsklone identifiziert werden. Parallel dazu wurden die Hochdichte-Proteinfilter zur Identifizierung von Religanden mit einem Antikörper gegen PhyB gescreent

(Insert des Ausgangsvektors). Das *screening* ergab, dass 2.000 der 40.000 Klone (5%) das Protein PhyB exprimiert und damit als Religanden klassifiziert werden konnten.

Die ermittelten Plattenkoordinaten der 5.000 potentiellen Proteinexpressionsklone (ohne Religanden, 12,5%) wurden in einer Liste zusammengefasst. Ausgehend von dieser Liste konnte die Neuordnung der cDNA-Bibliothek zur Erstellung einer Proteinexpressions-Unterbibliothek erfolgen (Plattenummer: 201-219).

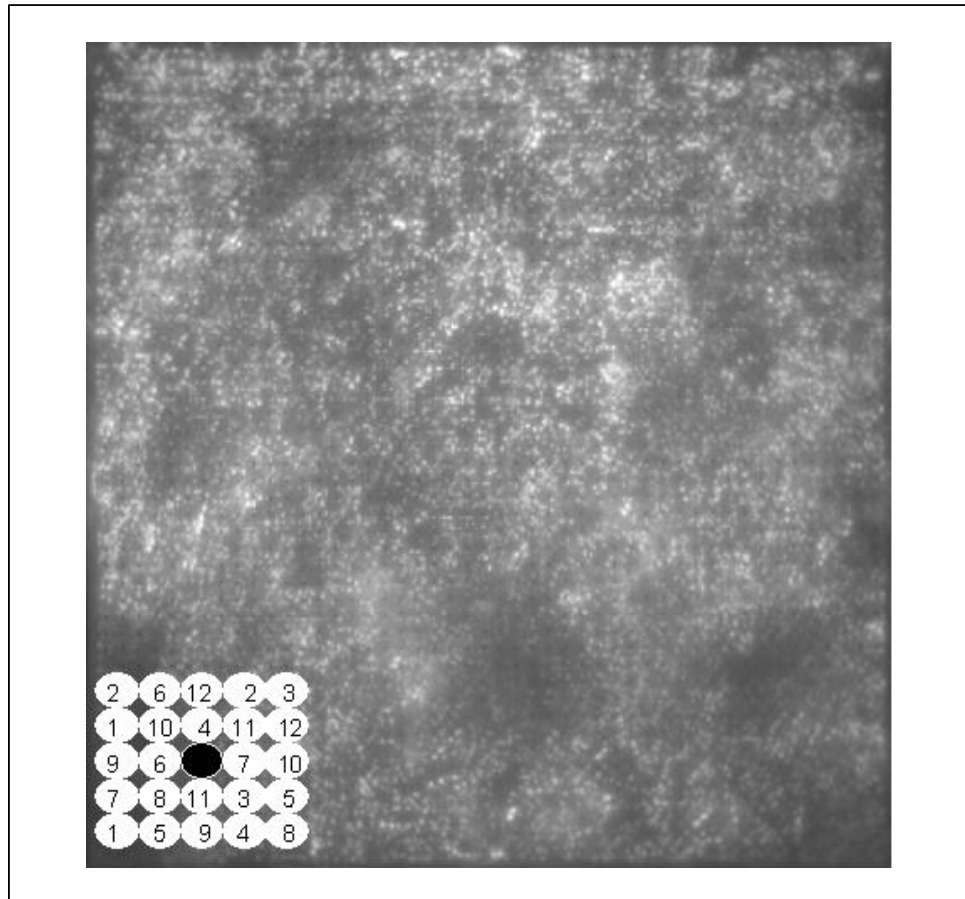


Abb. 3.3: Detektion putativer Expressionsklone. 27.648 Klone der Infloreszenz cDNA-Bibliothek (ATM1) wurden in Duplikaten auf eine PVDF-Membran gespottet. Proteine wurden *in situ* exprimiert und anschließend gegen einen anti-RGS His₆-Antikörper hybridisiert. Klone wurden als positiv bewertet, wenn die beiden Spots eines Duplikats als weiße Spots zu erkennen waren. Unten links im Bild ist schematisch das 5 x 5-Spotpattern dargestellt, wobei in der Mitte der *guide dot* zu sehen ist.

3.4 Erstellung des Uniklonsets der ATM1 cDNA-Bibliothek

3.4.1 Sequenzierung

Nach der Neuordnung der Unterbibliothek der potentiellen Proteinexpressionsklone der Infloreszenz cDNA-Bibliothek (Kap. 3.3) wurde als nächstes ein Uniklonset, ausgehend von diesen neu angeordneten Klonen, erstellt (vgl. 2.2.3). Diese weitere Neuordnung diente dazu, ein Set zu erstellen, in dem jeder Klon möglichst nur einmal vorkommt. Dazu wurden zunächst alle Klone der Proteinexpressions-Unterbibliothek sequenziert.

Es wurden 4.999 Infloreszenz-Klone der Unterbibliothek, unter Verwendung des Oligonukleotides pQE65, vom 5'-Ende ansequenziert (AGOWA GmbH, Berlin) (vgl. 2.2.3.1). Basierend auf der Sequenzinformation wurde anschließend das Uniklonset erstellt. Im Weiteren wird die Selektion der Klone für das Uniklonset erläutert. Zur Verdeutlichung sind die folgenden Selektionsschritte in Abb. 3.4 in einem Flussschema dargestellt. Bei den mittels Sequenzierung erhaltenen Sequenzen handelt es sich um ESTs (*expressed sequence tagged site*), die im nachfolgenden Text als Sequenzen bezeichnet werden.

Nach der Sequenzierung wurden die Sequenzen hinsichtlich ihrer Qualität kontrolliert (vgl. 2.2.3.2). Dazu wurden zunächst die Vektor- und Adaptersequenzen entfernt. Danach wurden 494 Sequenzen, welche weniger als 100 Nukleotide besaßen, aussortiert. Des Weiteren wurden 34 Sequenzen aufgrund der hohen Ähnlichkeit mit *E. coli* (<ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA>) eliminiert. Zusätzlich wurden 73 Sequenzen eliminiert, die nach Maskierung der Sequenzen mit Hilfe des „RepeatMakers“ aus mehr als 40% „Ns“ bestanden. Die Maskierung diente zur Ermittlung und Entfernung von hochrepetitiven Sequenzabschnitten. Insgesamt wurden damit von den 4.999 Sequenzen 601 Sequenzen mit nicht ausreichender Qualität ausgeschlossen. Die restlichen 4.398 Sequenzen gingen in die Erstellung des Uniklonsets ein.

Durch die dargestellte Überarbeitung der Sequenzen wurden diese auf die Ähnlichkeitssuche in den entsprechenden Datenbanken vorbereitet. Zudem konnte anhand der Entfernung des 5'-Vektoranteils eine Aussage darüber getroffen werden, ob sich die Sequenz im richtigen Leserahmen, d.h. dem Leserahmen, in dem der RGS His₆-Tag vorliegt, befand.

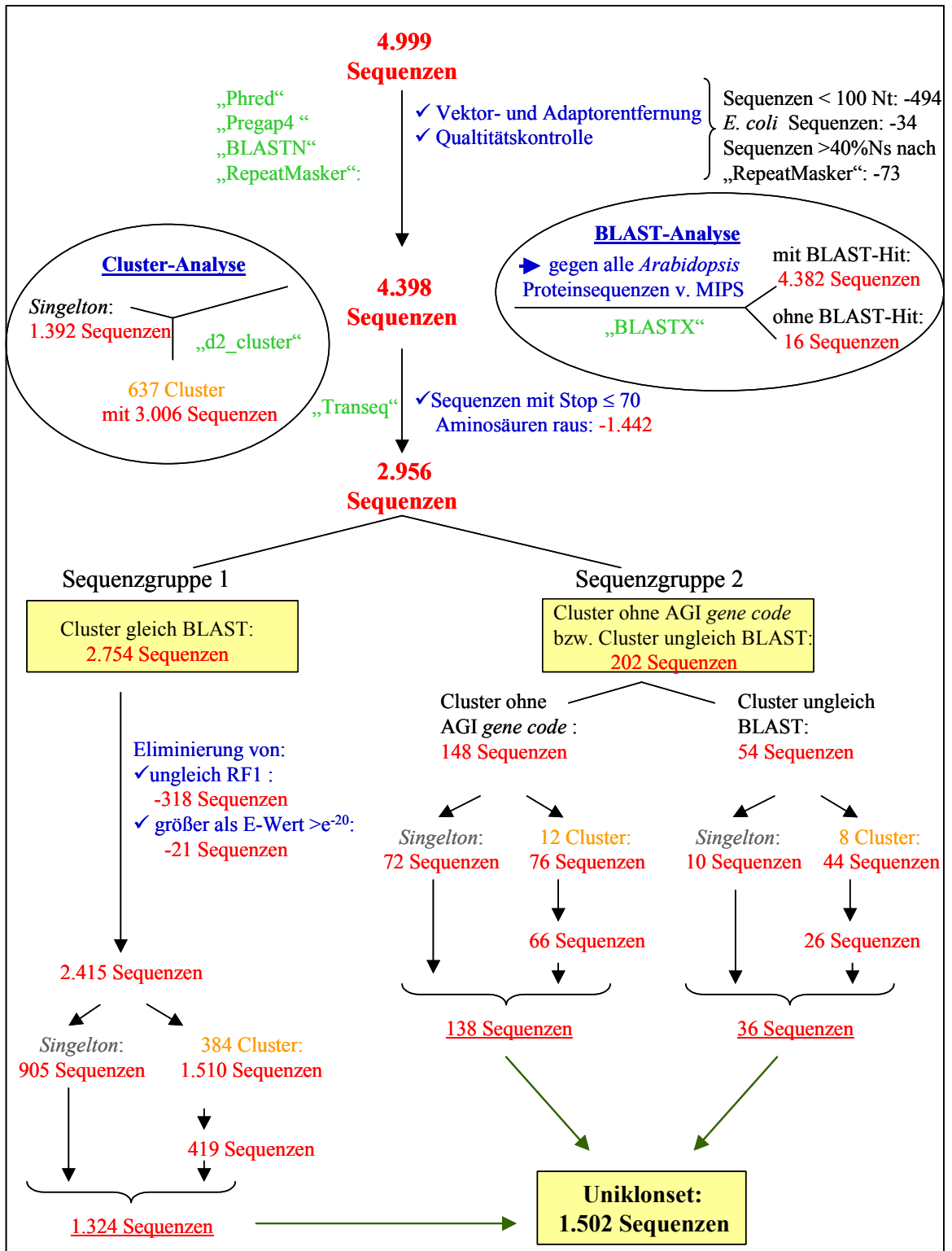


Abb. 3.4: Flussschema zur Erstellung des Uniklonsets der ATM1 cDNA-Bibliothek. Ausgehend von 4.999 potentiellen Expressionsklonen wurde ein Uniklonset, bestehend aus 1.502 Klonen, erstellt. Die blauen Markierungen bezeichnen die durchgeführten Selektionskriterien bzw. Analysen. Die roten bzw. die orangen Zahlen stellen die jeweilige Anzahl der Sequenzen bzw. Cluster dar. Hellgrün markiert sind die verwendeten Programme. Die gelben Kästen markieren sowohl die Anzahl der Sequenzen, von denen ausgegangen wurde, als auch die Anzahl der Sequenzen, die in das Uniklonset eingingen.

3.4.2 Sequenzanalyse

Mit den 4.398 Sequenzen, die den Qualitätskriterien entsprachen, wurde sowohl eine BLAST-Analyse (vgl. 2.2.3.3) als auch eine Cluster-Analyse (vgl. 2.2.3.4) durchgeführt.

Die Ergebnisse der folgenden Abschnitte sind in zwei Tabellen zusammengefasst, welche sich aufgrund ihrer Größen auf folgenden HTML-Seiten befinden:

- HTML-Tabelle 1:
http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionCloneSet.shtml

- HTML-Tabelle 2:
http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionUniCloneSet.shtml

Ausgehend von den Daten der HTML-Tabelle 1 wurde das Uniklonset erstellt. Die Daten dieses Sets wurde in der HTML-Tabelle 2 zusammengefasst. Der Aufbau beider Tabellen ist im Anhang (8.3 und 8.4) schematisch dargestellt.

Alle Tabellenverweise in den folgenden Kapiteln (Kap. 3.4.2 und 3.4.3) beziehen sich auf die HTML-Tabelle 1.

3.4.2.1 BLAST-Analyse

Die BLAST-Analyse wurde gegen die MIPS-*Arabidopsis*-Protein-Datenbank durchgeführt. Für 4.382 Sequenzen wurden Übereinstimmungen mit den kodierenden MIPS-*Arabidopsis*-Sequenzen gefunden; diesen konnten AGI *gene codes* zugeordnet werden. Dagegen konnte für 16 Sequenzen kein BLAST-Hit ermittelt werden (HTML-Tabelle 1, Spalte H). Basierend auf der BLAST-Analyse konnte eine erste Aussage über den Anteil der vollständigen cDNA-Klone und der Klone im richtigen Leserahmen getroffen werden. 1.732 der 4.398 (39,4%) sequenzierten potentiellen Expressionsklone lagen in voller Länge vor. 2.715 Klone der 4.398 (61,8%) Ausgangsklone befanden sich im richtigen Leserahmen.

Um eine erste Information über die potentielle Länge der exprimierten Proteine zu erhalten, wurden die in Kap. 3.4.1 erhaltenen 4.398 Nukleotidsequenzen in Aminosäuresequenzen translatiert und anschließend die Position der Stopcodone ermittelt. Aus der HTML-Tabelle 1 ist auch die Position des ersten Stopcodons (Spalte G) zu entnehmen. Es zeigte sich, dass für 1.159 (26,4%) Sequenzen kein Stopcodon zu identifizieren war. Dagegen wiesen 3.239 Sequenzen ein oder mehrere Stopcodone auf.

3.4.2.2 Cluster-Analyse

Nach der BLAST-Analyse der Sequenzen wurden die gewonnenen Ergebnisse in der HTML-Tabelle 1 zusammengefasst. Die Sequenzen wurden maskiert und auf DNA-Ebene gemeinsam mit allen codierenden Sequenzen des *Arabidopsis*-Genoms geclustert. Die Cluster-Analyse zeigte, dass 1.392 Klone je einmal vorkamen (31,7%). Die restlichen 3.006 Sequenzen waren redundant und konnten in 637 Cluster (Gruppen mit hoher Sequenzähnlichkeit) zusammengefasst werden (Tab. 3.4). Dabei wurden den unterschiedlichen Cluster entsprechend ihrer Zusammengehörigkeit Nummern zugeordnet, diese sind der HTML-Tabelle 1 (Spalte D) zu entnehmen.

Tab. 3.4: Ergebnisse der Cluster-Analyse der redundanten Sequenzen der potentiellen Proteinexpressionsklone der ATM1 cDNA-Bibliothek.

Clustergröße	Clusterfrequenz	Sequenzanzahl je Clustergruppe
2	327	654
3	124	372
4	57	228
5	32	160
6	15	90
7	18	126
8	13	104
9	9	81
10	10	100
11	8	88
12	2	24
13	3	39
14	3	42
18	2	36
19	1	19
20	1	20
21	1	21
22	1	22
23	1	23
26	1	26
27	1	27
28	1	28
31	1	31
32	1	32
33	1	33
90	1	90
163	1	163
327	1	327
Gesamt	637	3.006

In der Tabelle sind die Ergebnisse der Cluster-Analyse der redundanten Sequenzen der potentiellen Proteinexpressionsklone der ATM1 cDNA-Bibliothek dargestellt.

3.4.3 Selektion der Klone für das Uniklonset

Das erste Selektionskriterium war die Position des ersten Stopcodons innerhalb einer Sequenz. Es wurden 1.442 Sequenzen, die ein Stopcodon bei ≤ 70 Aminosäuren besaßen, aussortiert. Innerhalb der verbliebenen 2.956 Sequenzen ergaben sich aufgrund der Cluster- und der BLAST-Analyse zwei große Sequenzgruppen:

1. 2.754 Sequenzen, die bei der Cluster- und BLAST-Analyse gleiche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit aufwiesen (HTML-Tabelle 1, Spalte E mit Plus und Spalte F mit Eintrag).
2. Diese Sequenzgruppe bestand aus 202 Sequenzen: 148 Sequenzen, denen bei der Cluster-Analyse keine AGI *gene codes* zugeordnet werden konnten (HTML-Tabelle 1, Spalte F ohne Eintrag) bzw. 54 Sequenzen, die bei der Cluster- bzw. BLAST-Analyse unterschiedliche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit hatten (HTML-Tabelle 1, Spalte E mit Minus).

Von den 2.754 Sequenzen der Sequenzgruppe 1, deren Cluster-AGI *gene code* dem BLAST-AGI *gene code* entsprach (HTML-Tabelle 1, Spalte E mit Plus und Spalte F mit Eintrag), wurde als erstes der Leserahmen ermittelt, in dem sich die Sequenzen in Relation zum RGS His₆-Tag befanden. Dabei entspricht der Leserahmen 1 dem Leserahmen des RGS His₆-Tags. Damit wurden 318 *non-sense* Proteine, die im falschen Leserahmen abgelesen wurden, ausgeschlossen. Um die Wahrscheinlichkeit der Aussage der BLAST-Ergebnisse zu erhöhen, wurden zudem noch 21 Sequenzen mit einem E-Wert (Irrtumswahrscheinlichkeit) schlechter (also größer) als e^{-20} eliminiert (HTML-Tabelle 1, Spalte J). Von den nun verbliebenen 2.415 Sequenzen waren 905 Sequenzen Klone, die nur einmal vorkamen (*Singletons*); diese wurden direkt in das Uniklonset transferiert. Die restlichen 1.510 Sequenzen bildeten 384 unterschiedliche Cluster. Für diese Sequenzen mussten weitere Selektionskriterien festgelegt werden.

Diese Kriterien lauteten wie folgt, beginnend mit dem Wichtigsten:

- der Startpunkt des Subjekts (HTML-Tabelle 1, Spalte K), der Startpunkt des Querys (HTML-Tabelle 1, Spalte L) und der E-Wert (HTML-Tabelle 1, Spalte J)

Die Werte dieser Parameter sollten so klein wie möglich sein.

Aus den 384 Clustern wurden insgesamt 419 Klone für das Uniklonset ausgewählt. Von den 384 Clustern sind 363 Cluster durch jeweils einen Klon vertreten, während 21 Cluster durch mehr als einen Klon (zwischen zwei und acht Klonen) im Uniklonset repräsentiert sind.

Diese mehrfache Repräsentation ergab sich daraus, dass die Sequenzen dieser 21 Cluster jeweils mit mehreren *Arabidopsis*-Genen in ein Cluster fielen. Dies führte dazu, dass es für die Sequenzen eines Clusters unterschiedliche Treffer in der BLAST-Analyse gab. Entsprechend der unterschiedlichen BLAST-Hits wurden aus einem Cluster mehrere Klone selektiert (ein Klon pro BLAST-Hit).

Die Sequenzen der Sequenzgruppe 2 setzten sich aus 148 Sequenzen ohne Cluster-AGI *gene code* (HTML-Tabelle 1, Spalte F ohne Eintrag) bzw. 54 Sequenzen, deren Cluster-AGI *gene code* nicht dem BLAST-AGI *gene code* entsprach (HTML-Tabelle 1, Spalte E mit Minus), zusammen.

Da es bei allen Sequenzen der Sequenzgruppe 2 fraglich ist, ob man sich auf die Parameter des BLAST-Ergebnisses und somit auch auf die Leserahmendaussage beziehen kann, wurden hier nur die Aussagen der Stopcodone und der Clustergröße betrachtet. Von den 148 Sequenzen ohne Cluster-AGI *gene code* bestanden 72 Sequenzen aus Clustern mit je einer Sequenz und wurden direkt, ohne Stopcodon-Selektion, in das Uniklonset transferiert. Die restlichen 76 Sequenzen gehörten zu 12 Clustern mit mehr als einer Sequenz. Auf die Klonauswahl aus diesen 12 Clustern wird später näher eingegangen.

Von den 54 Sequenzen, deren Cluster-AGI *gene code* nicht dem BLAST-AGI *gene code* entsprach, gehörten 10 Sequenzen zu Clustern bestehend aus einer Sequenz, die direkt, ohne Stopcodon-Selektion, in das Uniklonset aufgenommen wurden. Die restlichen 44 Sequenzen bildeten 8 Cluster mit mehr als einer Sequenz.

Für die Auswahl der Sequenzen für das Uniklonset aus den Clustern (> 1 Sequenz) der Sequenzgruppe 2 wurde als Kriterium die Clustergröße wie folgt angewandt: Von Clustern, die sich aus weniger als sechs Sequenzen zusammensetzten, wurden alle Klone in das Uniklonset transferiert. Bei den Clustern, die sich aus sechs oder mehr Sequenzen zusammensetzten, wurden nur die Sequenzen in das Uniklonset mit einbezogen, die kein Stopcodon hatten.

Die 12 Cluster der 76 Sequenzen ohne Cluster-AGI *gene code* setzten sich aus 10 Clustern mit weniger als 6 Sequenzen und 2 Clustern mit mehr als 6 Sequenzen zusammen. Von den ersten Clustern (< 6 Sequenzen) gingen alle 23 Klone und von den zweiten Clustern (> 6 Sequenzen) gingen von 53 Sequenzen 43 in das Uniklonset ein.

Die 8 Cluster der 44 Sequenzen, bei denen der Cluster-AGI *gene code* nicht dem BLAST-AGI *gene code* entsprach, bestanden aus 7 Clustern mit weniger als 6 Sequenzen und 1 Cluster mit mehr als 6 Sequenzen. Von diesen gingen jeweils 22 bzw. 4 (von 22) Klone in das Uniklonset ein.

Zusammenfassend kann man sagen, dass ausgehend von 4.999 Sequenzen 4.398 Sequenzen als Sequenzen mit ausreichender Qualität befunden wurden. Diese Sequenzen wurden BLAST- und Cluster-Analysen unterzogen. Basierend auf den Ergebnissen dieser Analysen wurde ein Uniklonset erstellt, welches aus 1.502 Sequenzen besteht (Plattennummer: 311-314).

3.4.4 Erweiterung des Uniklonsets durch cDNA Expressionsklone, die in voller Länge vorlagen

Das Uniklonset wurde um weitere 192 vollständige cDNA-Klone erweitert. 96 dieser Klone wurden durch restriktionsabhängige Klonierung (*AscI/NotI*) gewonnen (Kersten *et al.*, 2003). Bei den restlichen 96 Klone handelte es sich um die Transkriptionsfaktoren, die als Entryklone vorlagen (zur Verfügung gestellt von Dr. Ralf Stracke und Prof. Dr. Bernd Weisshaar, Max Planck Institut für Züchtungsforschung, Köln). Diese wurden im Rahmen dieser Arbeit mittels LR-Reaktion im 96 *well* Format in den *E. coli* Destinationsvektor pQE30NASTDV überführt (vgl 2.2.3.7; Plattennummer: 315).

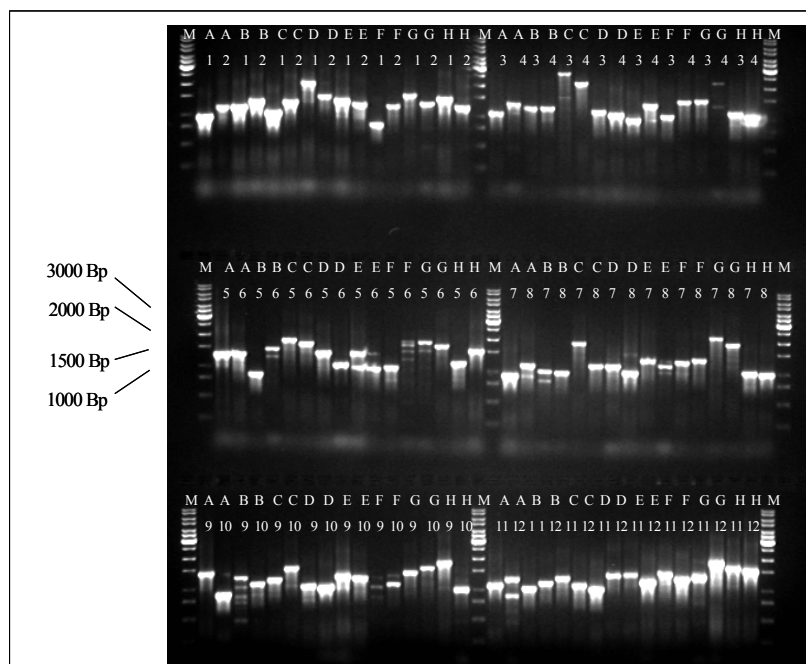


Abb. 3.5: Ermittlung der Insertlängen der 96 vollständige cDNA-Transkriptionsfaktoren. 10 μ l PCR-Produkt aus Kolonie-PCRs wurden auf einem 1,0%igen (w/v) Agarosegel elektrophoretisch aufgetrennt. Die Plattenpositionen der 96 *well* PCR-Platten sind oberhalb der Laufspur zu entnehmen. M = Marker, Bp = Basenpaare.

Zur Charakterisierung der gewonnenen Klone wurde jeweils ein rekombinanter Klon aus ursprünglich einer LR-Reaktion in einem 96 *well* PCR-Ansatz unter Verwendung der pQE-Vektorprimer eingesetzt. Das Ergebnis des PCR-*screenings* lieferte für alle 96 Klone PCR-Produkte in der zu erwartenden Insertlänge (Abb. 3.5). Das bedeutet, dass der parallele Transfer von 96 cDNAs aus dem Entryvektor in einen *E. coli* Expressionsvektor im 96 *well* Format erfolgreich war.

Somit lagen zusätzlich zu den 1.502 Uniklonset-Klonen 192 vollständige cDNA-Klone vor. Diese nunmehr insgesamt 1.694 Klone bildeten zusammen das Erweiterte Uniklonset, das im Weiteren verwendet wurde.

Die Daten des Uniklonsets sind, wie bereits oben erwähnt, in der HTML-Tabelle 2 zusammengefaßt und im Anhang exemplarisch dargestellt (siehe Anhang 8.4).

3.4.5 Reinigung von 96 rekombinanten Proteinen des Uniklonsets und Erstellung eines Testmicroarrays

Um eine erste Analyse des neu erstellten Uniklonsets durchzuführen, wurden 96 Klone des Uniklonsets der Mikrotiterplatte 311 Quadrant A1 exprimiert und im 96 *well* Format aufgereinigt (vgl. 2.2.5.4). Dazu wurde zunächst ein Teil der Zellpellets nach der Expression mittels SDS-Polyacrylamid-Elektrophorese untersucht (vgl. 2.2.5.1).

Von den 96 getesteten Klonen des Uniklonsets war bei 83% das spezifische überexprimierte Protein in den mit Coomassie-Blau gefärbten Gelen erkennbar (Abb. 3.6). Die Größe der Proteine variierte dabei zwischen 14 und 50 kDa.

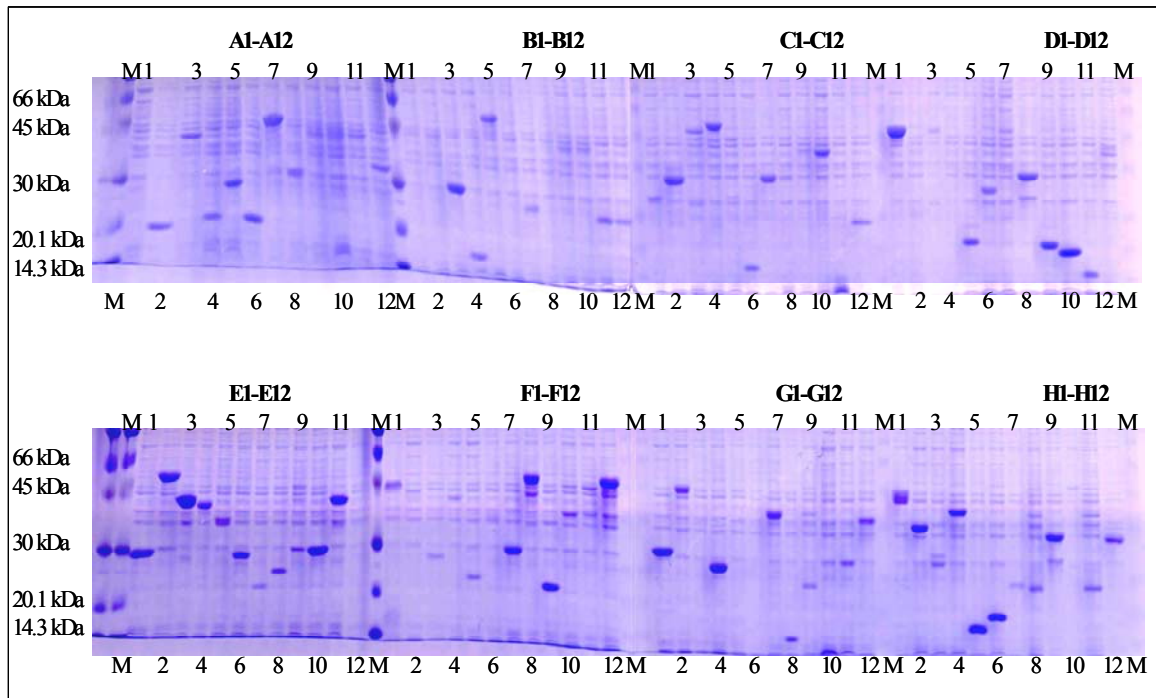


Abb. 3.6: 15% SDS-Polyacrylamidgel mit den Gesamtlysaten von 96 Klonen des Uniklonsets der ATM1 Bibliothek. Die Klone stammen aus der Platte 311, Quadrant A1. Die Plattenpositionen sind oberhalb der Laufspur zu entnehmen. M = Marker, die Markergrößen sind in kDa angegeben. Das Gel wurde mit Coomassie-Brilliant-Blau gefärbt.

Anschließend wurden die Zellpellets im 96 *well* Format mittels NiNTA-Agaroselösung über Filterplatten unter denaturierenden Bedingungen aufgereinigt. Ein Teil der Eluate, welche die gereinigten Proteine enthalten, wurde auf SDS-Polyacrylamidgele aufgetragen. Ein Gel wurde direkt nach der Elektrophorese mit Coomassie-Brilliant-Blau Färbelösung angefärbt (Abb. 3.6). Aus diesen Eluaten erfolgte die Aufreinigung unter denaturierenden Bedingungen. Die aufgereinigten Proteine sind in Abb. 3.6 dargestellt. Von den 96 aufgetragenen Proteinen sind 73 eindeutig detektierbar, während 4 weitere Proteine nur schwach zu erkennen sind. Somit waren ungefähr 80% der aufgereinigten Proteine im SDS-Polyacrylamidgel visuell zu identifizieren. Ein weiteres Gel wurde durch Westernblot auf eine PVDF-Membran transferiert und mit einem anti-RGS His₆-Antikörpers hybridisiert (vgl 2.2.5.1 und 2.2.5.2). Auf dem Westernblot konnten die gleichen Banden, wie auf dem gefärbten Coomassie-Brilliant-Blau visuell identifiziert werden (Abb. 3.7).

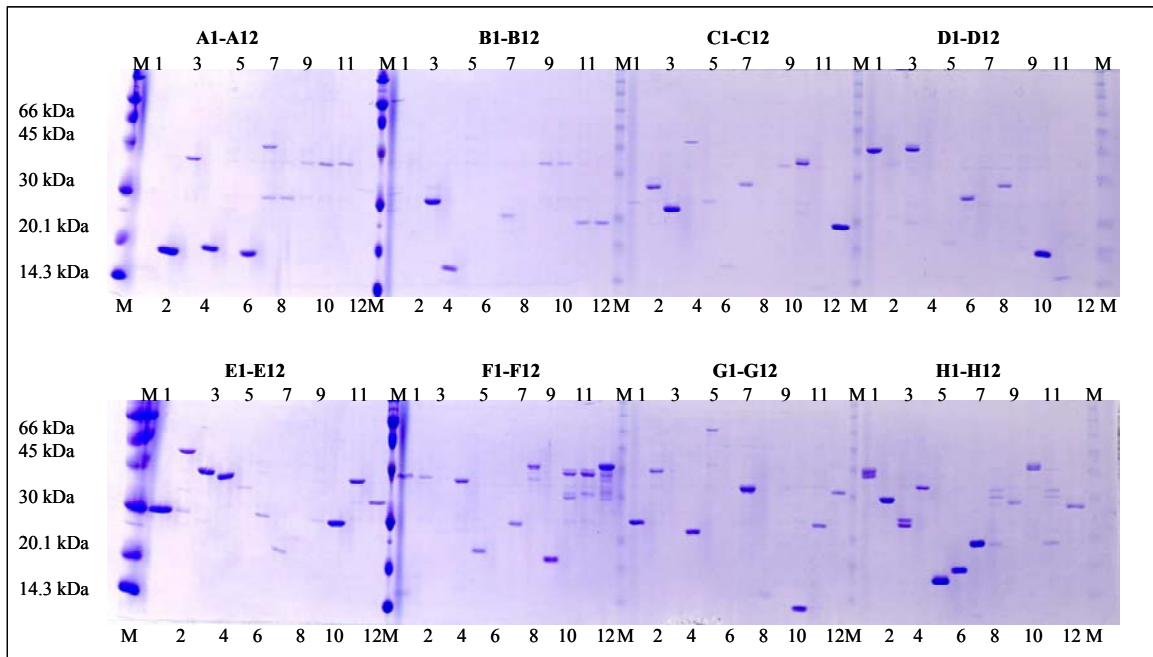


Abb. 3.7: 15% SDS-Polyacrylamidgel der 96 aufgereinigten Proteine des Uniklonsets der ATM1 Bibliothek. Die Klone stammen aus der Platte 311, Quadrant A1. Die Plattenpositionen sind oberhalb der Laufspur zu entnehmen. M = Marker, die Markergrößen sind in kDa angegeben. Das Gel wurde mit Coomassie-Brilliant-Blau gefärbt.

Im Folgenden wurden die 96 zufällig gewählten, aufgereinigten und in Abb. 3.6 und 3.7 dargestellten Proteine des Uniklonsets der ATM1 cDNA-Bibliothek auf einen Testmicroarray in einem 4 x 4-Pattern (Spotabstand: 1.125 μm) gespottet (Abb.3.8). Der Microarray wurde zunächst mit einem anti-RGS His₆-Antikörper und anschließend mit einem anti-Maus-Cy3 Antikörper inkubiert (vgl. 2.2.6.1 und 2.2.6.3).

Mit dem Testmicroarray sollte zunächst überprüft werden, ob die im Hochdurchsatz exprimierten und aufgereinigten Proteine des Uniklonsets auf den Microarrays detektierbar waren. Als Negativkontrollen wurden denaturierender Elutionspuffer, PBS, BSA (Protein ohne ein RGS His₆-Tag) und Kaninchen IgG eingesetzt. Diese Negativkontrollen lieferten keine Signale. Als eine weitere Negativkontrolle wurde Ratten IgG mitgeführt. Hier konnte ein Signal detektiert werden, was auf die Kreuzreaktivität des Zweitantikörpers, anti-Maus-Cy3 mit Ratten IgG zurückzuführen ist. Die Positivkontrollen waren ein Cy3 markierter Antikörper, der Maus anti-RGS His₆-Antikörper und Maus IgG. Für alle drei Positivkontrollen konnten Signale ermittelt werden.

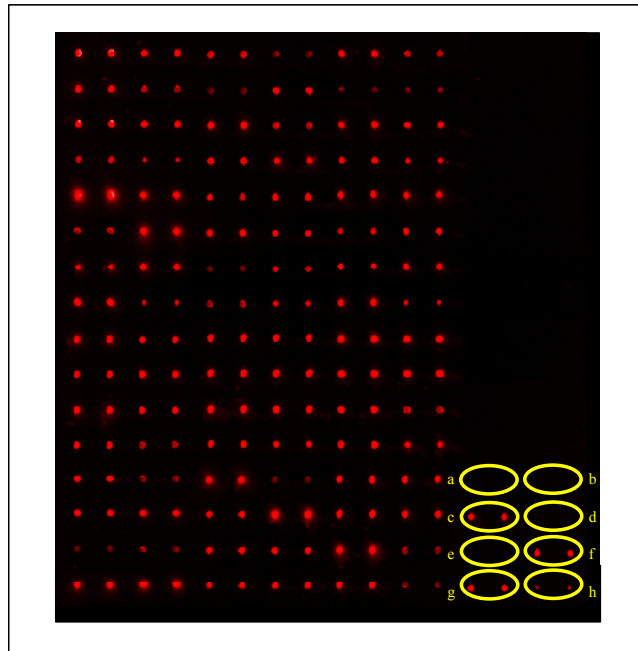


Abb.3.8: Testmicroarray der 96 Proteine des Uniklonsets der ATM1 cDNA-Bibliothek. Die Proteine wurden in horizontalen Duplikaten in einem 4 x 4-Pattern gespottet. In der rechten unteren Ecke sind die Kontrollen zu finden (a - h). Die Kontrollen waren: a) Elutionspuffer; b) PBS; c) Kaninchen-anti-Maus-IgG-Cy3-Konjugat, 1:25 in PBS verdünnt; d) BSA, $c = 20 \text{ pmol}/\mu\text{l}$ in PBS; e) Kaninchen IgG, 1:10 in PBS verdünnt; f) Ratten IgG, 1:10 in PBS verdünnt; g) Maus anti-RGS His₆-Antikörper, 1:10 in PBS verdünnt und h) Maus IgG, 1:10 in PBS verdünnt.

Für die Duplikate aller 96 Proteine waren Signale erkennbar, d.h. die Ausbeute der verwendeten Expressions- und Aufreinigungsschritte waren ausreichend, um alle hier zu untersuchenden rekombinanten Proteine zu detektieren (Abb.3.8). Im Vergleich zum Coomassie-Brilliant-Blau gefärbten Gel (80%) konnten auf dem Microarray mittels Antikörperdetektion alle rekombinanten Proteine nachgewiesen werden.

Nach der erfolgreichen Generierung des Testmicroarrays wurden nun alle 1.694 Klone des Erweiterten Uniklonsets im Hochdurchsatz unter denaturierenden Bedingungen aufgereinigt. Anschließend wurden die aufgereinigten Proteine auf FASTTM-Slides gespottet und für Phosphorylierungsversuche eingesetzt.

3.4.6 Bradford-Test

Um einen Eindruck zu bekommen, wie effizient die Aufreinigung des Erweiterten Uniklonsets war, und in welchen Konzentrationen die aufgereinigten Proteine vorlagen, wurde von drei 96er Platten (311-313, jeweils die Quadranten A2; vgl. 2.2.5.5) ein Bradford-Test durchgeführt (siehe Anhang 8.5).

Die mittels Bradford-Test ermittelte durchschnittliche Proteinkonzentration lag bei 105,96 µg/ml.

3.5 Spotten und Phosphorylierung des Erweiterten Uniklonsets

Nachdem die aufgereinigten Proteine des Erweiterten Uniklonsets auf die Microarrays angeordnet wurden (vgl. 2.2.6.2), erfolgte der Nachweis der rekombinanten Proteine auf einem Microarray unter Verwendung eines anti-RGS His₆-Antikörpers, gefolgt von einem Fluoreszenz markierten Zweitantikörper (anti-Maus-IgG-Cy3-Konjugat; vgl. 2.2.6.3), (Abb. 3.9).

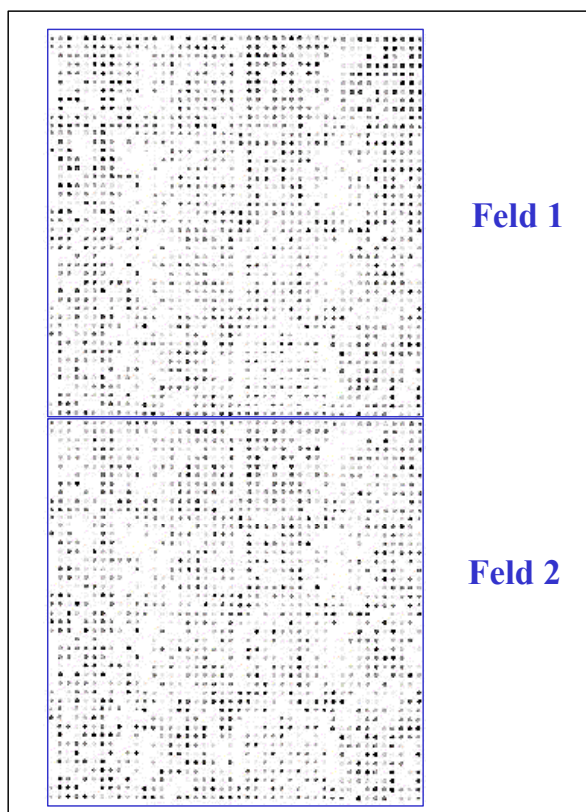


Abb. 3.9: Images des eingescannten Microarrays des Erweiterten Uniklonsets, nach Inkubation mit dem anti-RGS His₆-Antikörper. Die Proteine sind jeweils in zwei identischen Feldern gespottet worden. Der Nachweis rekombinanter Proteine erfolgte mit Hilfe eines anti-RGS His₆-Antikörpers gefolgt von einem Cy3-konjugierten Zweitantikörper. Fluoreszenzsignale wurden bei 532 nm detektiert.

Als Kontrolle wurde noch ein weiterer gespotteter Microarray nur mit diesem Zweitantikörper inkubiert. In diesem Fall wurden außer bei den gespotteten Positivkontrollen (Maus anti-RGS His₆-Antikörper und Kaninchen anti-Maus IgG3-Cy3) keine Signale detektiert. Damit wurde gezeigt, dass keine unspezifische Bindung des Zweitantikörpers an die rekombinaten Proteine erfolgt.

Im Fluoreszenz-*Image*, das nach der Inkubation mit dem anti-RGS His₆-Antikörper gewonnen wurde (Abb. 3.9), wird deutlich, dass 95% der gespotteten Proteine ein spezifisches Signal mit diesem Antikörper liefert. Dieses Ergebnis ist neben dem Bradford-Test ein Beleg dafür, dass die Hochdurchsatz-Proteinaufreinigung effizient genug war, um die gewonnen Proteine auf Microarrays nachzuweisen.

3.5.1 Detektion potentieller Phosphorylierungs-Targets

3.5.1.1 Vorversuche

Die Etablierung eines Microarray-basierten Phosphorylierungs-Assays zur qualitativen Ermittlung potentieller Targetproteine im kleinen Maßstab erfolgte in unserer Arbeitsgruppe unter Verwendung von rekombinaten Gerste-Proteinen (Kramer *et al.*, 2004). Diese *proof-of-principle* Studie wurde mit 768 gespotteten Proteinen durchgeführt, wobei potentielle Targets der Casein Kinase2 α identifiziert wurden.

3.5.1.2 Etablierung der Quantifizierung potentieller Targets im dichten Spotpattern

In der vorliegenden Arbeit wurde der oben genannte Phosphorylierungs-Assay dahingehend weiterentwickelt, dass eine wesentlich höhere Anzahl von Targets parallel analysiert und darüber hinaus quantifiziert werden kann. Als Targets wurden die 1.694 rekombinant exprimierten Proteine der Klone des Erweiterten Unikonsets eingesetzt. Nach Etablierung des Assays und des Quantifizierungssystems wurden Phosphorylierungsstudien unter Verwendung der noch wenig untersuchten *Arabidopsis* MAP-Kinasen 3 und 6 durchgeführt. Zur Etablierung der Quantifizierung wurde die Proteinkinase A (PKA) hinzugezogen. Diese Kinase gehört zu den am besten untersuchten Proteinkinasen. Ihr phylogenetisches Vorkommen reicht von den Hefen bis zu den Säugern, scheint jedoch bei den Prokaryonten und den höheren Pflanzen zu fehlen (Kallus, 2001). Aufgrund ihrer hohen Aktivität und ihrer kommerziellen Verfügbarkeit mit sehr konstanter spezifischer Aktivität ist diese Kinase, auch wenn sie in Pflanzen nicht vorkommt, sehr gut zur Etablierung des im Folgenden beschriebenen Quantifizierungssystems geeignet.

3.5.1.2.1 Ermittlung der Spot-Dichte

Zunächst wurde anhand von Phosphorylierungsstudien mit der Proteinkinase A die maximale Spotdichte ermittelt, bei der bei einem Phosphorylierungs-Assay einzelne Signale noch unterscheidbar sind.

Das Erweiterte Uniklonset wurde zu diesen Zwecke in einem 14 x 14-Pattern (Spotabstand: 321 μm) und in einem 15 x 15-Pattern (Spotabstand: 300 μm) gespottet. Bei beiden Spotdichten war eine Detektion der Proteine unter Verwendung des anti-RGS His₆-Antikörpers möglich. Nach Phosphorylierung der Microarrays mit Proteinkinase A wurden die radioaktiven Signale unter Verwendung einer *Imaging Plate* detektiert (vgl. 2.2.6.5). Es zeigte sich, dass bei den in dieser Dichte gespotteten Proteinen, die Proteine mit einer hohen Signalintensität zu starken Überstrahlungseffekten führten. Dies führte zu einer fehlerhaften Auswertung der Signalintensitäten benachbarter Proteine und zu fehlerhaften Hintergrundkorrekturen. Daher wurde für alle weiteren Versuche, welche in die Phosphorylierungsstudien eingingen, ein Spot-Pattern von 11 x 11 (Spotabstand: 410 μm) festgelegt. Dieses Pattern wurde generell in zwei identischen Feldern gespottet, wobei jedes Protein einmal pro Feld gespottet wurde und damit als Duplikat vorlag.

3.5.1.2.2 Quantifizierung mittels Proteinkinase A

Diese Versuche zur Etablierung der Methode erfolgten ausgehend von je zwei Microarrays aus unabhängigen Versuchsreihen. Zu jedem Phosphorylierungs-Versuch (siehe auch Kap. 0) wurde generell ein Kontroll-Microarray mitgeführt. Bei den Kontrollen enthielt der Phosphorylierungsansatz γ -ATP (³³P), aber keine Kinase. Die Analyse der *Images* dieser Kontroll-Microarrays lieferte keine Signale.

Abb. 3.10 zeigt die phosphorylierten Microarrays nach der Umsetzung mit der Proteinkinase A. Die Detektion der radioaktiven Signale erfolgte mit einer *Imaging Plate*, die wie im Kap. 2.2.6.5 beschrieben, ausgelesen wurden. Im rechten Bild (B) wurde zusätzlich ein *grid* aufgelegt und ausgerichtet (vgl. 2.2.6.5). Die Ausrichtung des *grids* erfolgte unter Verwendung der *grid dots* ([Myelin-assoziiertes basisches Protein (MBP)] = 2000 ng/ μl , versetzt mit Maus anti-RGS His₆-Antikörper, mit einer Endkonzentrationen von 1:10). Die Anordnung der *guide dots* innerhalb eines der 32 Unterfelder ist in Abb. 3.10 dargestellt.

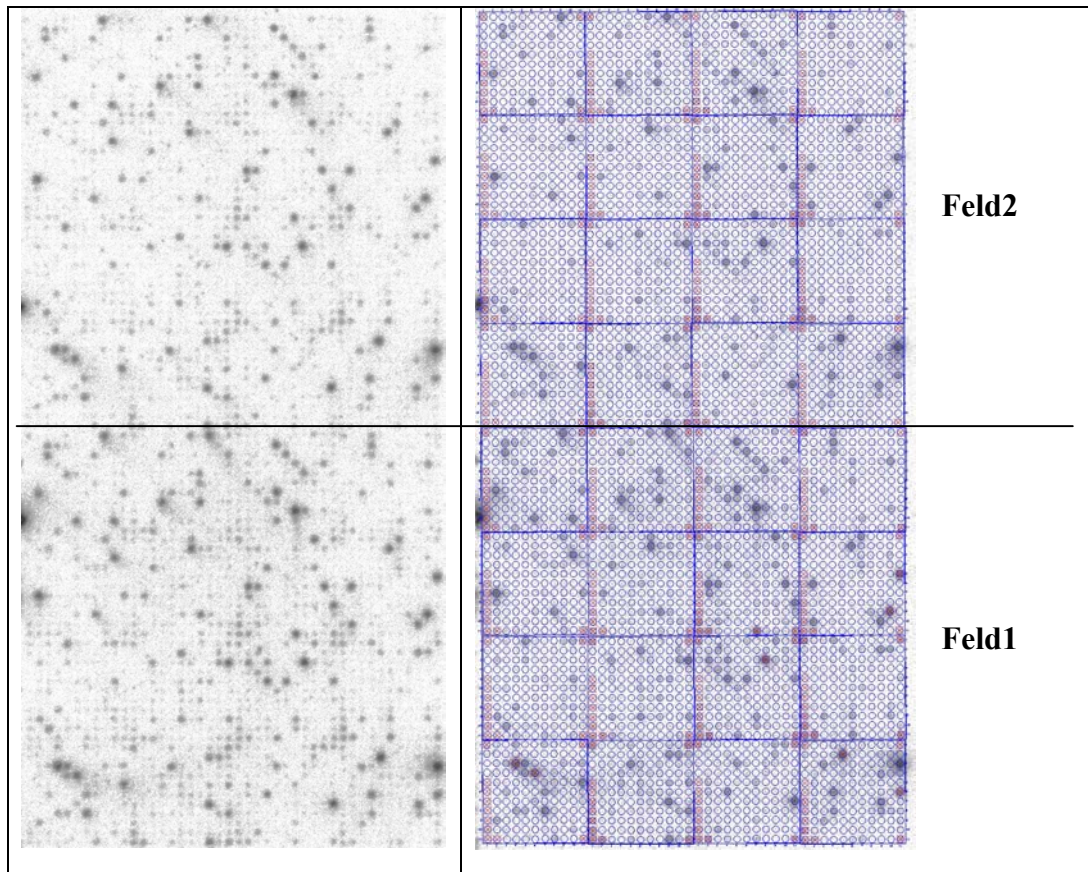


Abb. 3.10: Image eines eingescannten Microarrays mit den gespotteten Proteinen des Erweiterten Unikonsets nach der Phosphorylierung durch die PKA. Die Detektion erfolgte mittels *Imaging Plate*, die nach der Exposition (12 h) ausgelesen wurde (vgl. 2.2.6.5). Darstellung mit einem dynamischer Bereich von 0 – 65.536. Die Proteine wurden in zwei identische Felder auf den Microarrays angeordnet. Linkes Bild: zeigt einen mit PKA phosphorylierten Microarray. Rechtes Bild: entspricht dem linken Bild mit aufgelegten und ausgerichteten *grid* (vgl 2.2.6.5 und Abb. 2.2.6).

Anschließend wurde für jeden Microarray eine Feld zu Feld-Korrelation mit den Rohdaten der Duplikate durchgeführt. Eine dieser Rohdaten-Korrelationen ist in beispielhaft in Abb. 3.11 dargestellt.

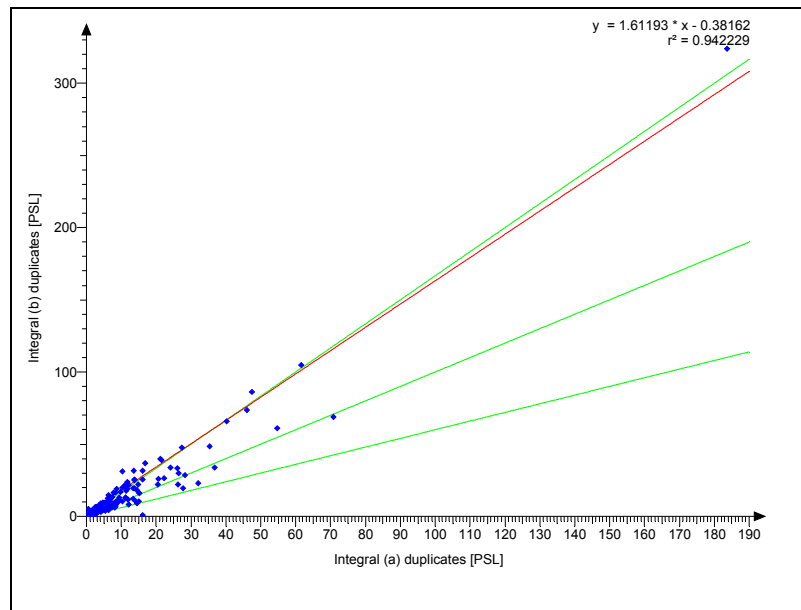


Abb. 3.11: Duplikatkorrelation der Signalintensitäten, der mit der PKA phosphorylierten und gespotteten Proteine auf dem Microarray. Jeder Datenpunkt in dieser Feld zu Feld-Korrelation (aus „Aida Array Matrix“) präsentiert den Signalintensität eines Proteins nach Phosphorylierung. Die Position eines Datenpunktes ergibt sich aus der Intensität der beiden Duplikate eines Proteins auf dem Microarray. Die äußeren beiden grünen Linien begrenzen den Bereich innerhalb dessen die beiden Duplikate weniger als 25% von ihrem Mittelwert abweichen. Alle Datenpunkte außerhalb dieses Bereichs wurden von der weiteren Auswertung ausgeschlossen (2.2.6.5). Die mittlere grüne Gerade zeigt schematisch die Idealgerade an, bei der alle Duplikate 100%ige reproduzierbar wären. Die rote Gerade ist die Regressionsgerade der Datenpunkte, deren Parameter rechts oben im Bild angegeben sind. PSL = photostimulierte Lumineszenz.

Die Datenpunkte, die mehr als 25% von ihrem Mittelwert abwichen, wurden manuell markiert und von der weiteren Auswertung ausgeschlossen. Des weiteren wurden Datenpunkte, die von „Aida Array Matrix“ automatisch als nicht auswertbar markiert wurden, manuell kontrolliert und gegebenenfalls ihr Ausschluss von der Auswertung rückgängig gemacht. Qualitätsgerechte Daten (nichtmarkierte Daten), die in die weitere Auswertung eingingen, wurden erneut korreliert. Die dabei zu verzeichnende Erhöhung des Korrelationskoeffizienten im Vergleich zur Rohdaten-Korrelation zeigt, dass durch die zuvor durchgeführte Selektion der Daten die Auswertung optimiert wurde.

Nach der Quantifizierung der einzelnen Microarrays, wurden die Mittelwerte der hintergrundkorrigierten Signalintensitäten der entsprechenden Proteine gewonnen. Diese wurden jeweils für die Microarray-Replikate eines Versuches miteinander korreliert (Microarray zu Microarray-Korrelation), um die Inter-Microarray-Variabilität der Signale eines Proteins innerhalb eines Versuches zu ermitteln (vgl. 2.2.6.5).

Zur Selektion potentieller Phosphorylierungs-Targets wurde ein *cutoff* Wert definiert. Zunächst gingen in diese Selektion nur die Proteine ein, deren gemittelte Werte bei einem

Microarray zu Microarray Vergleich mit einer Reproduzierbarkeit von 2 bestimmt werden konnten. Von den Proteinen wurden nur diejenigen Proteine als Kandidatenproteine ausgewählt, deren mittlere hintergrundkorrigierten Signalintensität auf einem Microarray über der 10fachen Standardabweichung des mittleren Hintergrundsignals aller Unterfelder des entsprechenden Microarrays ($Hg - 10 \times \delta$) lagen. Im Anschluss wurde die Schnittmenge der selektierten Proteine von mehreren Microarrays eines Experimentes gebildet.

Im Falle der Proteinkinase A heißt das, dass von 1.761 bzw. 1.752 auszuwertende Spots von zwei Microarrays 45 bzw. 35 Proteine als potentielle Phosphorylierungs-Targets ermittelt wurden. Basierend auf diesen Ergebnissen wurde die Schnittmenge dieser Targets ermittelt. In Abb. 3.12 ist die Reproduzierbarkeit der Targets in zwei unabhängigen mit Proteinkinase A inkubierten Microarrays dargestellt.

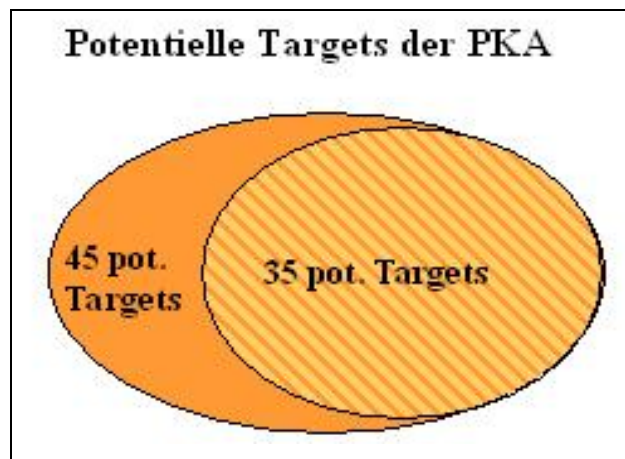


Abb. 3.12: Schematische Darstellung zur Ermittlung der potentiellen Targets der PKA. Die Ermittlung erfolgte ausgehend von zwei Microarrays. Pot. = potentiell.

Wie aus Abb. 3.12 ersichtlich wird, konnten ausgehend von den gespotteten Proteinen des Erweiterten Unikonsets 35 potentielle Phosphorylierungs-Targets (PKA-Set) detektiert werden. Diese sind in Tab. 3.5 zusammengefasst.

Tab. 3.5: Potentielle Phosphorylierungs-Targets der PKA.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
311 D05	At5g47390	Myb-related transcription activator-like	myb family transcription factor
311 D06	At1g68520	putative B-box zinc finger protein	zinc finger (B-box type) family protein
311 D15	At2g05990	enoyl-ACP reductase (ENR1)	enoyl-(acyl-carrier protein) reductase (NADH), chloroplast, putative / NADH-dependent enoyl-ACP reductase, putative
311 D21	At4g31580	splicing factor 9G8-like SR protein / SRZ-22	splicing factor RSZp22 (RSZP22) / 9G8-like SR protein (SRZ22)
311 M20	At2g30520	unknown protein	signal transducer of phototropic response (RPT2)
311 O09	At5g67350	unknown protein	expressed protein
311 P17	At3g11630	putative 2-cys peroxiredoxin	2-cys peroxiredoxin, chloroplast (BAS1)
312 A23	At4g14870	putative protein	expressed protein
312 I20	At1g22160	unknown protein	senescence-associated protein-related
312 K05	At4g16430	putative transcription factor BHLH3	basic helix-loop-helix (bHLH) family protein
312 K23	At1g56200	unknown protein	expressed protein
312 L08	At1g78020	unknown protein	senescence-associated protein-related
312 P13	At5g17710	chloroplast GrpE protein	co-chaperone grpE family protein
313 B21	At1g22610	Highly similar to phospho ribosylanthranilate transferase-	C2 domain-containing protein
313 D16	At4g15540	hypothetical protein	nodulin-related
313 D22	At5g63810	beta-galactosidase (emb/CAB64746.1)	beta-galactosidase, putative / lactase, putative
313 G17	At5g03040	unknown protein	calmodulin-binding family protein
313 G18	At5g17870	plastid-specific ribosomal protein	plastid-specific ribosomal protein-related
		6 precursor (Psrp-6) - like	
313 J03	At3g21055	photosystem II 5 kD protein precursor	photosystem II 5 kD protein, putative
313 N06	At5g54630	unknown protein	zinc finger protein-related
313 O21	At1g16870	Unknown protein (F611.13)	mitochondrial 28S ribosomal protein S29-related
313 P04	At4g33890	unknown protein	expressed protein
314 A16	At3g45010	carboxypeptidase precursor-like protein	serine carboxypeptidase III, putative
314 I06	At4g29080	phytochrome-associated protein PAP2	auxin-responsive AUX/IAA family protein
314 I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
314 J03	At1g07350	transformer-SR ribonucleoprotein, putative	transformer serine/arginine-rich ribonucleoprotein, putative
314 M21	At3g43510	putative protein	copa-like retrotransposon family
314 P13	At5g08670	H ⁺ -transporting ATP synthase beta chain (mitochondrial) -like protein	ATP synthase beta chain 1, mitochondrial
315 A10	At1g35680	clp. ribosomal large SU protein L21	50S ribosomal protein L21, chloroplast / CL21 (RPL21)
315 E03	At3g49120	Peroxidase	peroxidase, putative
315 E12	At4g38970	Putative fructose-bisphosphate aldolase	fructose-bisphosphate aldolase, putative
315 L05	At4g17880	AtbHLH 003	basic helix-loop-helix (bHLH) family protein
315 L06	At5g46760	AtbHLH 005	basic helix-loop-helix (bHLH) family protein
315 M08	At5g12030	17.6kDa HSP protein	17.7 kDa class II heat shock protein 17.6A (HSP17.7-CII)
315 M12	At5g11520	Aspartate aminotransferase (Asp3)	aspartate aminotransferase, chloroplast / transaminase A (ASP3) (YLS4)

In der Tabelle sind die potentiellen Phosphorylierungs-Targets des PKA-Sets dargestellt. PKA = Proteinkinase A.

Zur unabhängigen experimentellen Bestätigung dieser 35 potentiellen Phosphorylierungs-Targets des PKA-Sets und somit auch zur Bestätigung der ausgewählten Kriterien zur Quantifizierung, wurden die Proteine des PKA-Sets erneut exprimiert, aufgereinigt und auf ein 15%iges Polyacrylamidgel aufgetragen. Anschließend wurden die Proteine auf eine PVDF-Membran geblottet und dort ebenfalls unter Verwendung der Proteinkinase A phosphoryliert (Abb. 3.13).

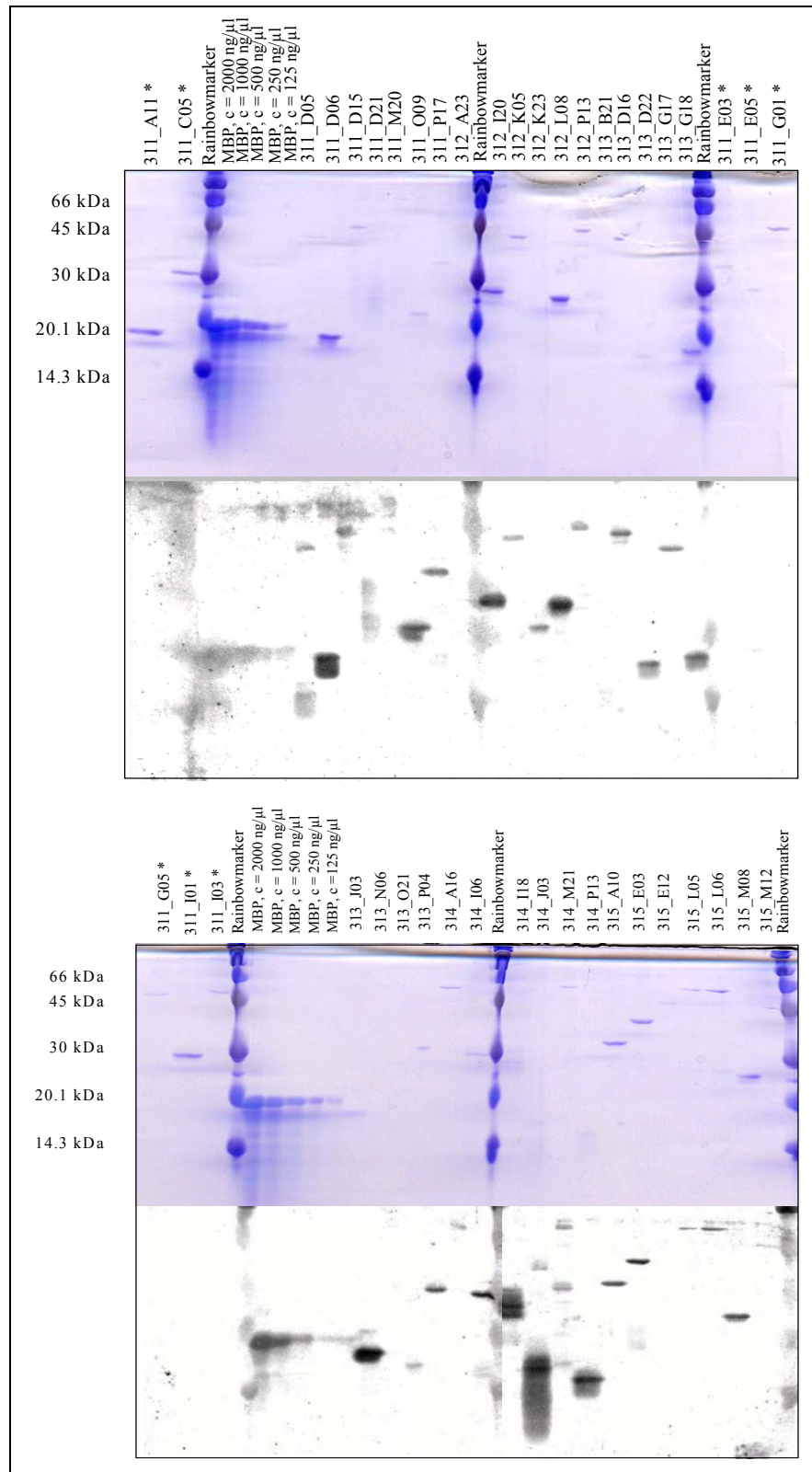


Abb. 3.13: 15% SDS-Polyacrylamidgele und phosphorylierte Blots der identifizierten potentiellen Phosphorylierungs-Targets des PKA-Sets. In den jeweiligen oberen Gelen wurden die aufgereinigten Proteine mit Coomassie-Brilliant-Blau gefärbt. Parallel wurden die Proteine mittels Blot auf PVDF-Membran transferiert und dort unter Verwendung der PKA phosphoryliert (jeweiligen unteren Teile der Abb.). Die Plattenpositionen des Unikonsets sind oberhalb der Laufspur zu entnehmen. * = Negativkontrollen.

In Abb. 3.13 ist zu erkennen, dass von den 35 potentiellen Phosphorylierungs-Targets des PKA-Sets in den Coomassie-Brilliant-Blau gefärbten Gelen 26 Proteine und in den phosphorylierten Blots 28 Proteine visuell zu identifizieren waren. In den jeweiligen äußeren Banden wurden die Negativkontrollen aufgetragen (in Abb. 3.13 mit * markiert). Diese konnten im Coomassie-Brilliant-Blau gefärbten Gel visualisiert, aber nach der Phosphorylierung des Blots nicht detektiert werden. Bei den Negativkontrollen handelt es sich um Proteine, die bei der Ermittlung der potentiellen Phosphorylierungs-Targets aufgrund der niedrigen Signalintensitäten beider Duplikate auf den Microarrays nicht als Targets identifiziert wurden.

Des weiteren wurden das PKA-Set und das Uniklonset auf die folgenden vier bekannten Konsensussequenzen von Targetproteinen der Proteinkinase A untersucht (Kallus, 2001):

- RXS*/T*
- RR/KX/S*/T*
- RXXS*
- KRXXS*

(R = Arginin, K = Lysin, S = Serin, T = Threonin, X = beliebige Aminosäure, * = Phosphorylierungsstelle)

In Tab. 3.6 werden die Ergebnisse dieser Motivsuche zusammengefasst. Die Tabelle zeigt, dass der Anteil dieser Motive im PKA-Set im Vergleich zum Uniklonset zunahm.

Tab. 3.6: Vergleichende Motivsuche des PKA-Sets und des Uniklonsets.

	PKA-Set	Uniklonset
RXS*/T*	89%	73%
RR/KX/S*/T*	80	22%
RXXS*	89%	59%
KRXXS*	20%	9%

In der Tabelle sind die Ergebnisse der Motivsuche unter Verwendung des PKA-Sets dargestellt. PKA = Proteinkinase A.

Da die Proteinkinase A zu den Serin/Threonin-Proteinkinasen gehört, wurden zusätzlich die Serin- und Threoningehalte, sowohl im PKA-Set als auch im Uniklonset, bestimmt (Tab. 3.7). Anhand der Tabelle ist eine Zunahme des Seringehaltes im PKA-Set im Vergleich zum Uniklonset zu sehen. Bezogen auf den Threoningehalt kam es zu keiner Änderung.

Tab. 3.7: Vergleich der Serin- und der Threoningehalte des PKA- und des Uniklonsets.

	PKA-Set	Uniklonset
Seringehalt	10%	8%
Threoningehalt	5%	5%

In der Tabelle sind die Ergebnisse des Vergleichs der Serin- und Threoningehalte des PKA-Sets und des Uniklonsets dargestellt. PKA = Proteinkinase A.

3.5.1.2.3 Identifizierung potentieller Phosphorylierungs-Targets der MAP-Kinasen MPK3 und MPK6

Nachdem ein Quantifizierungssystem etabliert worden war, wurden die Phosphorylierungs-Targets der *Arabidopsis* MAP-Kinase 3 (MPK3) und der MAP-Kinase 6 (MPK6) ermittelt. Beide Kinasen wurden in aufgereinigter und aktivierter Form von Dr. Justin Lee und Prof. Dr. Dierk Scheel, Institut für Pflanzenbiochemie (IPB), Halle, zur Verfügung gestellt.

Jeder Kinase-Versuch wurde auf jeweils zwei Microarrays in zwei unabhängigen Versuchsreihen durchgeführt. Anhand von Vorversuchen wurde eine geeignete Konzentration der einzusetzenden MPKs ermittelt. Dabei zeigte sich, dass die MPK3 (Konzentration: 100 ng/μl) in geringeren Konzentrationen eingesetzt werden konnte, als die MPK6 (Konzentration: 200 ng/μl).

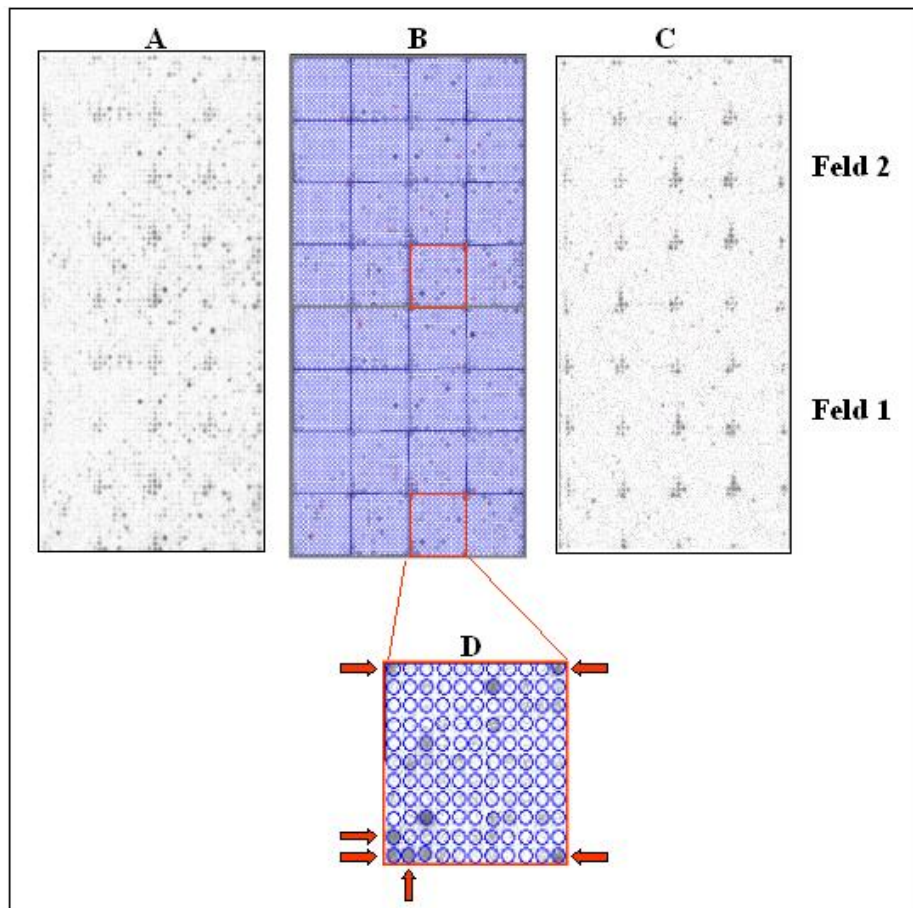


Abb. 3.14: Image der eingescannten Microarrays mit den gespotteten Proteinen des Erweiterten Unikonsets nach der Phosphorylierung durch die MPK3 bzw. MPK6. Die Proteine wurden in zwei identischen Feldern auf den Microarrays angeordnet. Die Detektion erfolgte mit einer Imaging Plate, die nach der Exposition (MPK3: 6 h und MPK6: 48 h) ausgelesen wurde (vgl. 2.2.6.5). Darstellung mit einem dynamischer Bereich von 0 - 65.536. A: zeigt einen mit MPK3 phosphorylierten Microarray. B: entspricht A mit dem aufgelegten und ausgerichteten grid (vgl. 2.2.6.5). C: präsentiert den mit MPK6 phosphorylierten Microarray. D: stellt eins der 16 Unterfelder der beiden identischen Felder dar. Die roten Pfeile markieren die guide dots ([Myelin-assoziiertes basisches Protein (MBP)] = 2.000 ng/ μ l, versetzt mit anti-RGS His₆-Antikörper, mit einer Endkonzentrationen von 1:10).

Abb. 3.14 zeigt beispielhaft die auf den Microarrays durchgeführte Phosphorylierung von Proteinen durch MPK3 (Bild A und Bild B) bzw. MPK6 (Bild C). Die Detektion wurde wie im Kap. 2.2.6.5 beschrieben durchgeführt. Auch in diesem Experiment erfolgte zunächst die Feld zu Feld-Korrelationen, beispielhaft in Abb. 3.15 dargestellt, das Markieren der nicht auszuwertenden Spots und die anschließend Microarray zu Microarray-Korrelation (Kap. 3.5.1.2.2).

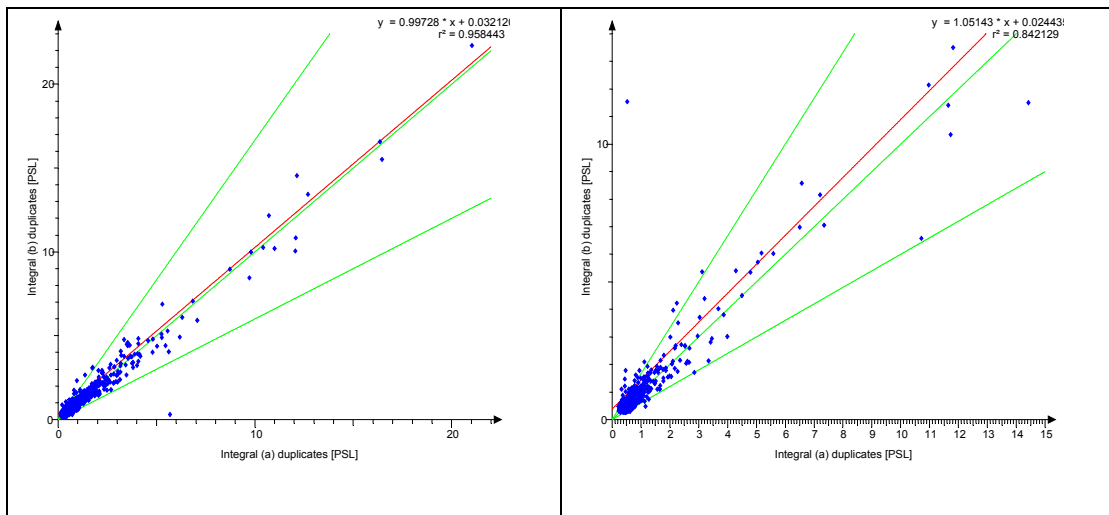


Abb. 3.15: Duplikatkorrelation der Signalintensitäten der mit MPK3 (links) und MPK6 (rechts) phosphorylierten und gespotteten Proteine auf den Microarrays. Die Rohdaten wurden mit dem Programm „Aida Array Matrix“ korreliert. Hierbei werden die beiden Spots, die ein Protein repräsentieren, miteinander verglichen. Detaillierte Beschreibung zur Abb. siehe Abb. 3.11. PSL = photostimulierte Lumineszenz.

Die Feld zu Feld-Korrelation in Abb. 3.15 zeigt, dass die Signalintensitäten der Duplikate auf allen Microarrays der MPK3 in einer hohen Reproduzierbarkeit vorlagen und somit alle Microarrays in die weitere Auswertung mit einbezogen wurden. Bei der MPK6 konnte anhand der Feld zu Feld-Korrelation festgestellt werden, dass die Daten zweier Microarrays aus dem einem Experiment in nicht ausreichender Reproduzierbarkeit vorlagen. Leider lag die MPK6 in nicht ausreichender Aktivität vor, um diesen Versuch im Rahmen der vorliegenden Arbeit noch einmal zu wiederholen (persönliche Mitteilung Dr. Justin Lee und Prof. Dr. Dierk Schell, IBP, Halle). Somit konnten die Schnittmengen der MPK3, sowohl von den einzelnen Experimenten als auch von den Experimenten untereinander gebildet werden. Die Ermittlung der MPK6 Schnittmengen dagegen ging von einem Experiment aus.

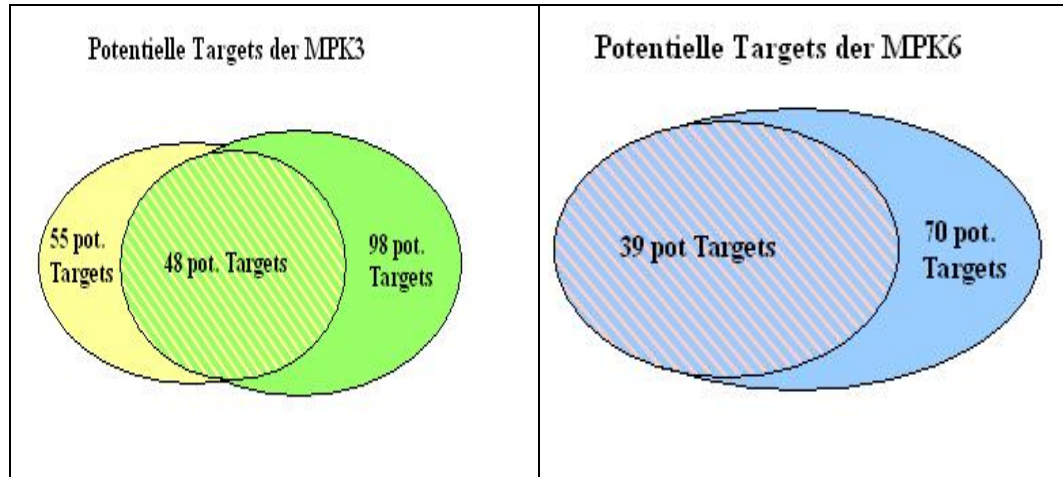


Abb. 3.16: Schematische Darstellung zur Ermittlung der potentiellen Targets der MPK3 (links) und der MPK6 (rechts). Im Falle der MPK3 wurde die Schnittmenge ausgehend von zwei Experimenten mit insgesamt vier Microarrays gebildet. Bei der MPK6 wurde die Schnittmenge ausgehend von zwei Microarrays eines Experimentes gewonnen. Pot. = potentiell.

Wie aus Abb. 3.16 ersichtlich wird, konnten ausgehend von den gespotteten Proteinen des Erweiterten Unikonsets 48 potentielle Phosphorylierungs-Targets der MPK3 (MPK3-Set, linkes Bild in Abb. 3.16) und 39 potentielle Phosphorylierungs-Targets der MPK6 (MPK6-Set, rechtes Bild in Abb. 3.16) identifiziert werden. Die potentiellen Phosphorylierungs-Targets des MPK3-Sets bzw. des MPK6-Sets sind in Tab. 3.8 bzw. 3.9 zusammengefasst.

Tab. 3.8: Zusammenfassung der potentiellen Targets der MPK3.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
311 B05	At5g23200	unknown protein	expressed protein
311 B06	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
311 F11	At5g58620	putative protein	zinc finger (CCCH-type) family protein
311 F16	At5g67360	cucumisin-like serine protease (gb AAC18851.1)	cucumisin-like serine protease (ARA12)
311 H10	At2g40510	40S ribosomal protein S26	40S ribosomal protein S26 (RPS26A)
312 A24	At5g02560	unknown protein	histone H2A, putative
312 B03	At1g26740	unknown protein	expressed protein
312 B13	At1g77450	GRAB1-like protein	no apical meristem (NAM) family protein
312 D20	At1g02840	ribonucleoprotein SF-2 like protein	pre-mRNA splicing factor SF2 (SF2) / SR1 protein
312 F05	At3g11510	putative 40S ribosomal protein s14	40S ribosomal protein S14 (RPS14B)
312 G12	At5g19290	phospholipase - like protein	esterase/lipase/thioesterase family protein
312 G21	At2g19730	putative ribosomal protein L28	60S ribosomal protein L28 (RPL28A)
312 J15	At4g03260	putative protein phosphatase regulatory subunit	leucine-rich repeat family protein
312 P16	At4g39880	unknown protein	ribosomal protein L23 family protein
313 B11	At4g15000	ribosomal protein	60S ribosomal protein L27 (RPL27C)
313 E14	At5g48990	unknown protein (At5g48990)	kelch repeat-containing F-box family protein
313 F05	At5g62070	unknown protein	calmodulin-binding family protein
313 F12	At3g48930	cytosolic ribosomal protein S11	40S ribosomal protein S11 (RPS11A)
313 G02	At1g64370	unknown protein	expressed protein
313 G18	At5g17870	plastid-specific ribosomal protein 6 precursor (Psrp-6) - like	plastid-specific ribosomal protein-related
313 H08	At2g39460	60S ribosomal protein L23A	60S ribosomal protein L23A (RPL23aA)
313 J07	At1g52740	putative histone H2A	histone H2A, putative
313 K04	At3g52580	putative ribosomal protein S14	40S ribosomal protein S14 (RPS14C)
313 K23	At5g48760	60S ribosomal protein L13a	60S ribosomal protein L13a (RPL13aD)
313 L13	At3g06730	thioredoxin, putative	thioredoxin family protein
313 N14	At4g39200	ribosomal protein S25	40S ribosomal protein S25 (RPS25E)
313 O05	At1g65480	flowering time locus T (FT)	flowering locus T protein (FT)
314 E05	At5g65360	histone H3 (sp P05203)	histone H3
314 F02	At5g65360	histone H3 (sp P05203)	histone H3
314 G05	At5g45775	ribosomal protein L11-like	60S ribosomal protein L11 (RPL11D)
314 G08	At5g10360	40S ribosomal protein S6	40S ribosomal protein S6 (RPS6B)
314 G10	At5g44100	casein kinase I	casein kinase, putative
314 G18	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 H17	At3g07350	unknown protein	expressed protein
314 I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
314 J03	At1g07350	transformer-SR ribonucleoprotein, putative	transformer serine/ arginine-rich ribonucleoprotein, putative
314 L17	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 L19	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 N09	At5g65360	histone H3 (sp P05203)	histone H3
314 O11	At1g02070	unknown protein	expressed protein
314 P19	At4g17390	60S ribosomal protein L15 homolog	60S ribosomal protein L15 (RPL15B)
315 A04	At3g58700	Ribosomal protein L16	60S ribosomal protein L11 (RPL11B)
315 A05	At5g14320	30S ribosomal protein S13	30S ribosomal protein S13, chloroplast (CS13)
315 C07	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
315 E08	At1g03680	Thioredoxin m1 (m-type precursor)	thioredoxin M-type 1, chloroplast (TRX-M1)
315 F05	At2g02820	AtMYB 88	myb family transcription factor (MYB88)
315 G09	At5g66940	DNA binding protein-like	Dof-type zinc finger domain-containing protein
315 I01	At1g16700	NADH:ubiquinone oxidoreductase (complex I)	NADH-ubiquinone oxidoreductase 23 kDa subunit, mitochondrial, putative

In der Tabelle sind die potentielle Phosphorylierungs-Targets des MPK3-Sets dargestellt. MPK3 = Mitogen-aktivierten Proteinkinase 3.

Tab. 3.9: Zusammenfassung der potentiellen Targets der MPK6.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
311_B06	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
311_F11	At5g58620	putative protein	zinc finger (CCCH-type) family protein
311_F16	At5g67360	cucumis-like serine protease	cucumis-like serine protease (ARA12)
311_H07	At5g47570	unknown protein	expressed protein
311_H10	At2g40510	40S ribosomal protein S26	40S ribosomal protein S26 (RPS26A)
311_P20	At3g07110	putative 60S ribosomal protein L13A	60S ribosomal protein L13A (RPL13aA)
312_B13	At1g77450	GRAB1-like protein	no apical meristem (NAM) family protein
312_D20	At1g02840	ribonucleoprotein SF-2 like protein	pre-mRNA splicing factor SF2 (SF2) / SR1 protein
312_F05	At3g11510	putative 40S ribosomal protein s14	40S ribosomal protein S14 (RPS14B)
312_G01	At1g56220	unknown protein	dormancy/auxin associated family protein
312_G12	At5g19290	phospholipase - like protein	esterase/lipase/thioesterase family protein
312_I20	At1g22160	unknown protein	senescence-associated protein-related
312_P16	At4g39880	unknown protein	ribosomal protein L23 family protein
313_B05	At2g46020	putative SNF2 subfamily transcriptional activator	transcription regulatory protein SNF2, putative
313_E14	At5g48990	unknown protein (At5g48990)	kelch repeat-containing F-box family protein
313_G13	At4g11280	ACC synthase (AtACS-6)	1-aminocyclopropane-1-carboxylate synthase 6 / ACC synthase 6 (ACS6)
313_H08	At2g39460	60S ribosomal protein L23A	60S ribosomal protein L23A (RPL23aA)
313_K04	At3g52580	putative ribosomal protein S14	40S ribosomal protein S14 (RPS14C)
313_K23	At5g48760	60S ribosomal protein L13a	60S ribosomal protein L13A (RPL13aD)
313_L02	At3g22845	unknown protein	emp24/gp25L/p24 protein-related
313_L13	At3g06730	thioredoxin, putative	thioredoxin family protein
313_N06	At5g54630	unknown protein	zinc finger protein-related
313_O05	At1g65480	flowering time locus T (FT)	flowering locus T protein (FT)
313_P22	At5g06140	sorting nexin-like protein	phox (PX) domain-containing protein
314_F02	At5g65360	histone H3 (sp P05203)	histone H3
314_G05	At5g45775	ribosomal protein L11-like	60S ribosomal protein L11 (RPL11D)
314_G10	At5g44100	casein kinase I	casein kinase, putative
314_H17	At3g07350	unknown protein	expressed protein
314_I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
314_L23	At5g21100	ascorbate oxidase-like protein	L-ascorbate oxidase, putative
314_N09	At5g65360	histone H3 (sp P05203)	histone H3
314_N21	At3g60390	homeobox-leucine zipper protein HAT3	homeobox-leucine zipper protein 3 (HAT3) / HD-ZIP protein 3
314_P19	At4g17390	60S ribosomal protein L15 homolog	60S ribosomal protein L15 (RPL15B)
315_A04	At3g58700	Ribosomal protein L16	60S ribosomal protein L11 (RPL11B)
315_A05	At5g14320	30S ribosomal protein S13	30S ribosomal protein S13, chloroplast (CS13)
315_A10	At1g35680	clp. ribosomal large SU protein L21	50S ribosomal protein L21, chloroplast / CL21 (RPL21)
315_E08	At1g03680	Thioredoxin ml (m-type precursor)	thioredoxin M-type 1, chloroplast (TRX-M1)
315_I01	At1g16700	NADH:ubiquinone oxidoreductase (complex I)	NADH-ubiquinone oxidoreductase 23 kDa subunit, mitochondrial, putative
315_P07	At1g06170	putative bHLH transcription factor (bHLH089)	basic helix-loop-helix (bHLH) family protein

In der Tabelle sind die potentielle Phosphorylierungs-Targets des MPK6-Sets dargestellt. MPK 6 = Mitogen-aktivierten Proteinkinase 6.

Um die identifizierten potentiellen Phosphorylierungs-Targets zu bestätigen, wurden die entsprechenden Klone des MPK3-Sets selektiert, kultiviert und die Proteine nach der Aufreinigung (vgl. 2.2.5.4) auf ein 15%iges SDS-Polyacrylamidgel aufgetragen (vgl. 2.2.5.1). Anschließend wurden die aufgetrennten Proteine mittels Blot auf eine PVDF-Membran transferiert (vgl. 2.2.5.2). Danach wurden die transferierten Protein unter Verwendung der MPK3 phosphoryliert (Abb. 3.17).

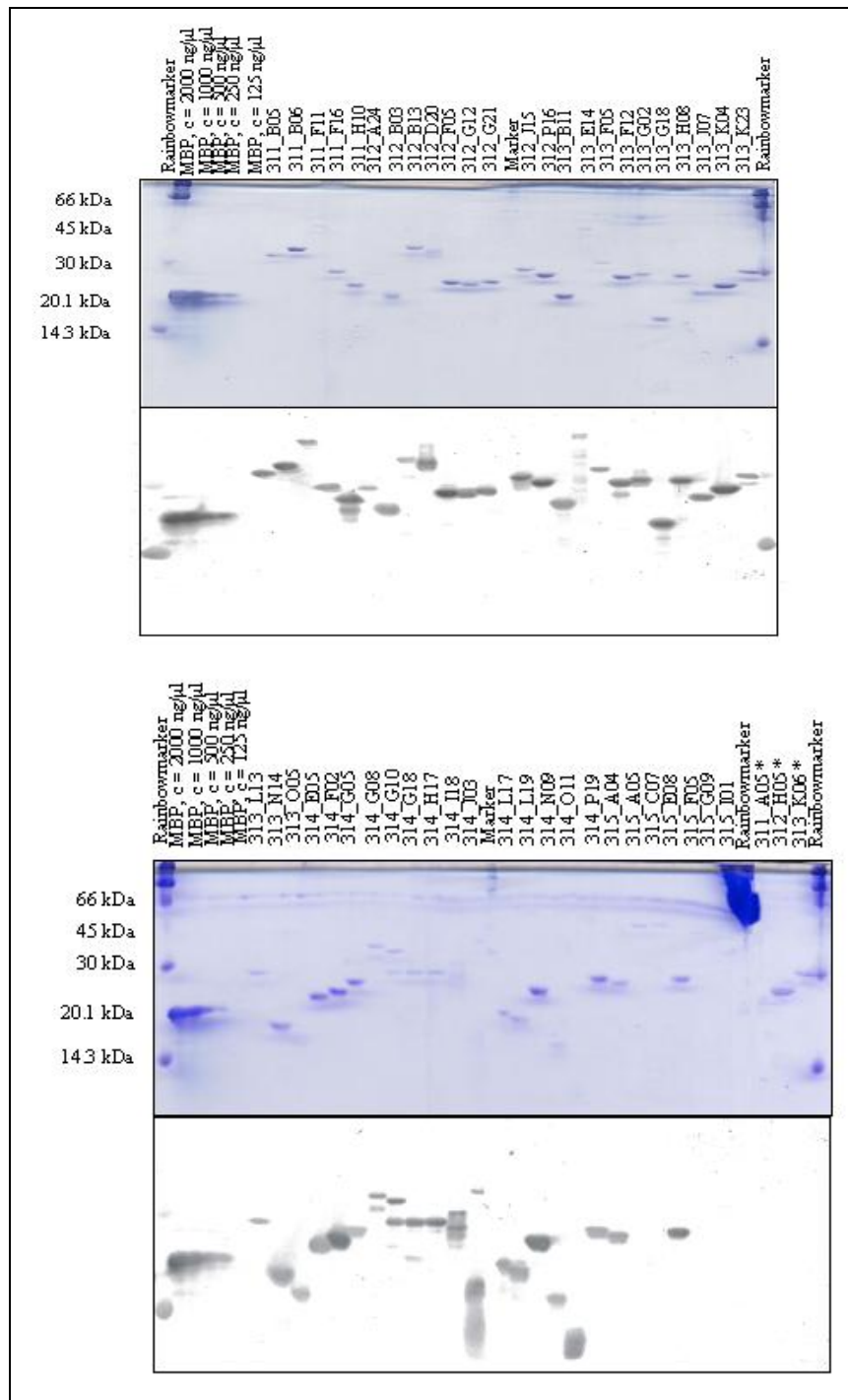


Abb. 3.17: 15% SDS-Polyacrylamidgele und phosphorylierte Blots der identifizierten potentiellen Phosphorylierungs-Targets des MPK3-Sets. In den jeweiligen oberen Gelen wurden die aufgereinigten Proteine mit Coomassie-Brilliant-Blau gefärbt. Parallel wurden die Proteine mittels Blot auf PVDF-Membran transferiert und dort unter Verwendung der MPK3 phosphoryliert (jeweiligen unteren Teile der Abb.). Die Plattenpositionen des Uniklonsets sind oberhalb der Laufspur zu entnehmen. * = Negativkontrollen.

Abb. 3.17 zeigt, dass von 48 potentiellen Phosphorylierungs-Targets des MPK3-Sets 45 sowohl im Coomassie-Brilliant-Blue gefärbten Gel als auch in den phosphorylierten Blots visuell bestätigt werden konnten. Auf beiden Gelen wurde neben dem

Molekulargewichtsmarker (links) das MBP (Myelin-assoziiertes basisches Protein) als Positivkontrolle in verschiedenen Verdünnungsstufen eingesetzt (2.000 ng/ μ l, 1.000 ng/ μ l, 500 ng/ μ l, 250 ng/ μ l, 125 ng/ μ l), wobei alle bis auf die letzte Verdünnungsstufe detektiert wurden. Im unteren Gel wurden rechts die Negativkontrollen aufgetragen (in Abb. 3.17 mit * markiert). Diese konnten im Coomassie-Brilliant-Blau-Gel visualisiert, aber nach der Phosphorylierung des Blots nicht detektiert werden. Bei den Negativkontrollen handelt es sich ebenfalls wieder um Proteine, die bei der Ermittlung der potentiellen Phosphorylierungs-Targets aufgrund der niedrigen Signalintensitäten nicht als Targets identifiziert wurden.

3.5.2 Vergleich potentieller Phosphorylierungs-Targets der verschiedenen Kinasen

Da die MPK3 und die MPK6 zu einer Familie der Proteinkinasen gehören, wurden zur Ermittlung gemeinsamer potentieller Phosphorylierungs-Targets dieser beiden Kinasen die Schnittmenge der Targets beider Kinasen gebildet (Abb. 3.18).

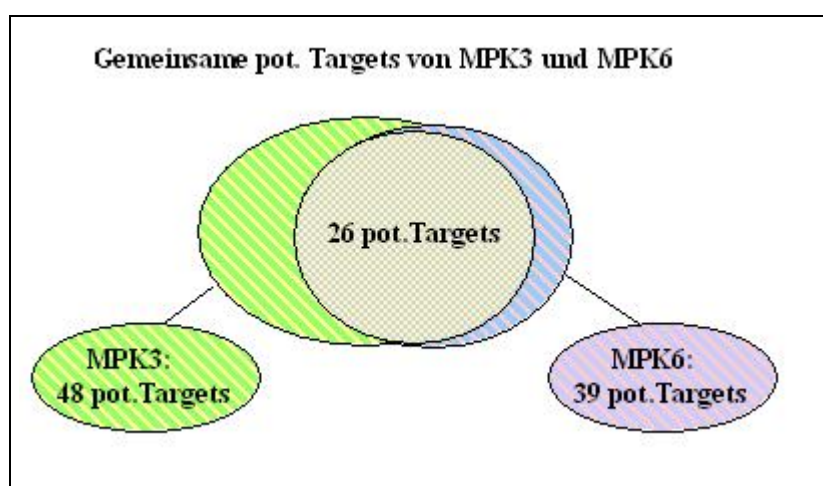


Abb. 3.18: Schematische Darstellung zur Ermittlung der gemeinsamen potentiellen Targets der MPK3 und der MPK6. Pot. = potentiell.

Der Abb. 3.18 kann entnommen werden, dass 26 gemeinsame potentielle Phosphorylierungs-Targets der MPK3 und der MPK6 ermittelt wurden, welche in Tab. 3.10 zusammengefasst werden. Zusätzlich werden in zwei weiteren Tabellen (Tab.3.11 und Tab. 3.12) die potentiellen Phosphorylierungs-Targets dargestellt, die spezifisch sind für MPK3 bzw. MPK6.

Tab. 3.10: Zusammenfassung der gemeinsamen potentiellen Targets der MPK3 und MPK6.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
311 B06	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
311 F11	At5g58620	putative protein	zinc finger (CCCH-type) family protein
311 F16	At5g67360	cucumisin-like serine protease	cucumisin-like serine protease (ARA12)
311 H10	At2g40510	40S ribosomal protein S26	40S ribosomal protein S26 (RPS26A)
312 B13	At1g77450	GRAB1-like protein	no apical meristem (NAM) family protein
312 D20	At1g02840	ribonucleoprotein SF-2 like protein	pre-mRNA splicing factor SF2 (SF2) / SR1 protein
312 F05	At3g11510	putative 40S ribosomal protein s14	40S ribosomal protein S14 (RPS14B)
312 G12	At5g19290	phospholipase - like protein	esterase/lipase/thioesterase family protein
312 P16	At4g39880	unknown protein	ribosomal protein L23 family protein
313 E14	At5g48990	unknown protein (At5g48990)	kelch repeat-containing F-box family protein
313 H08	At2g39460	60S ribosomal protein L23A	60S ribosomal protein L23A (RPL23aA)
313 K04	At3g52580	putative ribosomal protein S14	40S ribosomal protein S14 (RPS14C)
313 K23	At5g48760	60S ribosomal protein L13a	60S ribosomal protein L13A (RPL13aD)
313 L13	At3g06730	thioredoxin, putative	thioredoxin family protein
313 O05	At1g65480	flowering time locus T (FT)	flowering locus T protein (FT)
314 F02	At5g65360	histone H3 (sp P05203)	histone H3
314 G05	At5g45775	ribosomal protein L11-like	60S ribosomal protein L11 (RPL11D)
314 G10	At5g44100	casein kinase I	casein kinase, putative
314 H17	At3g07350	unknown protein	expressed protein
314 I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
314 N09	At5g65360	histone H3 (sp P05203)	histone H3
314 P19	At4g17390	60S ribosomal protein L15 homolog	60S ribosomal protein L15 (RPL15B)
315 A04	At3g58700	Ribosomal protein L16	60S ribosomal protein L11 (RPL11B)
315 A05	At5g14320	30S ribosomal protein S13	30S ribosomal protein S13, chloroplast (CS13)
315 E08	At1g03680	Thioredoxin m1 (m-type precursor)	thioredoxin M-type 1, chloroplast (TRX-M1)
315 I01	At1g16700	NADH:ubiquinone oxidoreductase (complex I	NADH-ubiquinone oxidoreductase 23 kDa subunit, mitochondrial, putative

In der Tabelle sind die gemeinsamen potentiellen Phosphorylierungs-Targets des MPK3 und MPK6-Sets dargestellt. MPK = Mitogen-aktivierten Proteinkinase.

Tab. 3.11: Potentielle Targets, die spezifisch für MPK3 sind.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
312 A24	At5g02560	unknown protein	histone H2A, putative
312 B03	At1g26740	unknown protein	expressed protein
312 G21	At2g19730	putative ribosomal protein L28	60S ribosomal protein L28 (RPL28A)
312 J15	At4g03260	putative protein phosphatase regulatory subunit	leucine-rich repeat family protein
313 B11	At4g15000	ribosomal protein	60S ribosomal protein L27 (RPL27C)
313 F05	At5g62070	unknown protein	calmodulin-binding family protein
313 F12	At3g48930	cytosolic ribosomal protein S11	40S ribosomal protein S11 (RPS11A)
313 G18	At5g17870	plastid-specific ribosomal protein 6 precursor (Psrp-6) - like	plastid-specific ribosomal protein-related
313 J07	At1g52740	putative histone H2A	histone H2A, putative
313 N14	At4g39200	ribosomal protein S25	40S ribosomal protein S25 (RPS25E)
314 E05	At5g65360	histone H3 (sp P05203)	histone H3
314 G08	At5g10360	40S ribosomal protein S6	40S ribosomal protein S6 (RPS6B)
314 G18	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 J03	At1g07350	transformer-SR ribonucleoprotein, putative	transformer serine/arginine-rich ribonucleoprotein, putative
314 L17	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 L19	At3g04400	60S ribosomal protein L17	60S ribosomal protein L23 (RPL23C)
314 O11	At1g02070	unknown protein	expressed protein
315 C07	At2g18020	60S ribosomal protein L2	60S ribosomal protein L8 (RPL8A)
315 F05	At2g02820	AtMYB 88	myb family transcription factor (MYB88)
315 G09	At5g66940	DNA binding protein-like	Dof-type zinc finger domain-containing protein

In der Tabelle sind die spezifischen potentiellen Phosphorylierungs-Targets des MPK3-Sets dargestellt. MPK3 = Mitogen-aktivierten Proteinkinase 3.

Tab. 3.12: Potentielle Targets, die spezifisch für MPK6 sind.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
311_H07	At5g47570	unknown protein	expressed protein
311_P20	At3g07110	putative 60S ribosomal protein L13A	60S ribosomal protein L13A (RPL13aA)
312_G01	At1g56220	unknown protein	dormancy/auxin associated family protein
312_I20	At1g22160	unknown protein	senescence-associated protein-related
313_B05	At2g46020	putative SNF2 subfamily transcriptional activator	transcription regulatory protein SNF2, putative
313_G13	At4g11280	ACC synthase (AtACS-6)	l-aminocyclopropane-1-carboxylate synthase 6 / ACC synthase 6 (ACS6)
313_L02	At3g22845	unknown protein	emp24/gp25L/p24 protein-related
313_N06	At5g54630	unknown protein	zinc finger protein-related
313_P22	At5g06140	sorting nexin-like protein	phox (PX) domain-containing protein
314_L23	At5g21100	ascorbate oxidase-like protein	L-ascorbate oxidase, putative
314_N21	At3g60390	homeobox-leucine zipper protein HAT3	homeobox-leucine zipper protein 3 (HAT3) / HD-ZIP protein 3
315_A10	At1g35680	clp. ribosomal large SU protein L21	50S ribosomal protein L21, chloroplast / CL21 (RPL21)
315_P07	At1g06170	putative bHLH transcription factor (089)	basic helix-loop-helix (bHLH) family protein

In der Tabelle sind die spezifischen potentiellen Phosphorylierungs-Targets des MPK6-Sets dargestellt. MPK6 = Mitogen-aktivierten Proteinkinase 6.

Als eine weitere Kontrolle hinsichtlich der Spezifität der Targetermittlung wurde zusätzlich zu der Ermittlung der gemeinsamen potentiellen Phosphorylierungs-Targets, der zu einer Familie gehörenden MPK3 und der MPK6, die jeweilige Schnittmenge mit den Targets der Proteinkinase A gebildet.

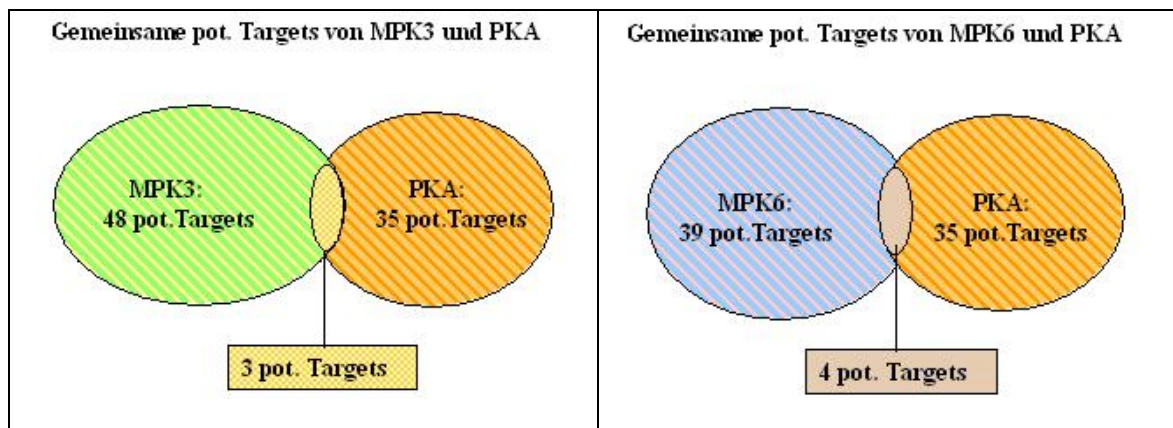


Abb. 3.19: Schematische Darstellung der gemeinsamen Targets der PKA mit MPK3 (linke Seite) bzw. MPK6 (rechte Seite). pot. = potentiell.

Abb. 3.19 ist zu entnehmen, dass die Proteinkinase A nur wenige gemeinsame Targets mit denen der MPK3 (3 gemeinsame pot. Targets) bzw. MPK6 (4 gemeinsame pot. Targets) hat. Dies zeigt, dass die Ermittlung der potentiellen Phosphorylierungs-Targets der verschiedenen Kinasen spezifisch ist. Die ermittelten gemeinsamen Phosphorylierungs-Targets sind in Tab. 3.13 (PKA und MPK3) und Tab. 3.14 (PKA und MPK6) zusammengefasst.

Tab. 3.13: Zusammenfassung der gemeinsamen potentiellen Targets der PKA und der MPK3.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
313_G18	At5g17870	plastid-specific ribosomal protein 6 precursor (Psrp-6) - like	plastid-specific ribosomal protein-related
314_I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
314_J03	At1g07350	transformer-SR ribonucleoprotein, putative	transformer serine/arginine-rich ribonucleoprotein, putative

In der Tabelle sind die gemeinsamen potentiellen Phosphorylierungs-Targets des PKA- und des MPK3-Sets dargestellt. MPK3 = Mitogen-aktivierten Proteinkinase 3; PKA = Proteinkinase A.

Tab. 3.14: Zusammenfassung der gemeinsamen potentiellen Targets der PKA und der MPK6.

unigene name	BLAST-Hit	putative Genfunktion (MIPS)	Genfunktion (TIGR 5)
312_I20	At1g22160	unknown protein	senescence-associated protein-related
313_N06	At5g54630	unknown protein	zinc finger protein-related
314_I18	At1g23860	9G8-like splicing factor / SRZ-21	splicing factor RSZp21 (RSZP21) / 9G8-like SR protein (SRZ21)
315_A10	At1g35680	clp. ribosomal large SU protein L21	50S ribosomal protein L21, chloroplast / CL21 (RPL21)

In der Tabelle sind die gemeinsamen potentiellen Phosphorylierungs-Targets des PKA- und des MPK6-Sets dargestellt. MPK6 = Mitogen-aktivierten Proteinkinase 6; PKA = Proteinkinase A.

Den Tab. 3.13 und 3.14 ist zu entnehmen, dass der Klon 311_I18 sowohl in dem PKA-MPK3-Set als auch in dem PKA-MPK6-Set vorkommt.

4 Diskussion

Ziel der Arbeit war erstens die Herstellung von rekombinanten *Arabidopsis thaliana*-Proteinen und zweitens die Etablierung von funktionellen Studien im Hochdurchsatz. Aus diesem Grunde wurden diese Themen in zwei Abschnitten zusammengefasst und diskutiert. Zur Gewinnung rekombinanter Proteine im ersten Teil der Arbeit wurden cDNA-Bibliotheken aus der Pistill und der Infloreszenz von *Arabidopsis thaliana* erstellt und charakterisiert. Ausgehend von der Infloreszenz cDNA-Bibliothek wurde ein Uniklonset von Expressionsklonen erstellt. Zu diesem Zweck wurden potentielle Expressionsklone dieser Bibliothek unter Verwendung von Hochdichte-Proteinfiltern selektiert und in einer Proteinexpressions-Unterbibliothek neu angeordnet. Basierend auf BLAST- und Cluster-Analysen dieser Klone, erfolgte anschließend die Erstellung eines Uniklonsets aus 1.502 Klonen, das um 192 vollständige cDNA-Klone erweitert wurde. Alle 1.694 Klone dieses Erweiterten Uniklonsets wurden im Hochdurchsatz exprimiert und die entsprechenden rekombinanten Proteine aufgereinigt.

Im zweiten Abschnitt wurden mit den aufgereinigten Proteinen *Arabidopsis*-Protein-Microarrays erstellt, welche anschließend für Phosphorylierungsstudien *in vitro* verwendet wurden. Diese Studien wurden unter Verwendung verschiedener Kinasen durchgeführt: einerseits mit der gut beschriebenen Maus Proteinkinase A (PKA) und andererseits mit zwei *Arabidopsis* MAP-Kinasen (MPK3 und 6), die eine wichtige Rolle in der Signaltransduktion spielen. Ziel war es, nach Weiterentwicklung des Testsystems, unter Verwendung der PKA, potentielle Phosphorylierungs-Targets der MAP-Kinasen zu identifizieren.

4.1 Gewinnung rekombinanter *Arabidopsis* Proteine

4.1.1 Herstellung und Charakterisierung von zwei cDNA-Bibliotheken

Die cDNA-Bibliotheken repräsentieren im Idealfall die Gesamtheit der exprimierenden Gene eines Gewebes, eines Zelltyps oder eines Organismus' zu einem bestimmten Zeitpunkt. Basierend auf der Analyse dieser Bibliotheken, kann man wichtige Einblicke in den Ablauf biologischer Prozesse bekommen. So können z.B. neue Gene und deren endgültige Spleißprodukte identifiziert werden, aber auch die Expressionsintensität der Gene des entsprechenden Stadiums des jeweiligen Gewebes, Zelltyps bzw. Organismus' analysiert werden (Clark *et al.*, 1999).

cDNA-Bibliotheken können auch als Quelle zur Expression rekombinanter Proteine des entsprechenden Gewebes und Stadiums genutzt werden. Dazu werden die cDNAs entweder

direkt in einen geeigneten Expressionsvektor (Büssow *et al.*, 1998; Gutjahr, 2003; Kramer, 2003) oder zunächst in einen Entryvektor des GATEWAY-Systems kloniert. Ausgehend von diesem Entryvektor kann dann die cDNA über Rekombination je nach Bedarf in unterschiedliche Expressionsvektoren transferiert werden (Heyman *et al.*, 1999; Walhout *et al.*, 2000). Im Rahmen dieser Arbeit wurden beide Wege zur Herstellung von cDNA-Bibliotheken verwendet. Die Infloreszenz cDNA-Bibliothek wurde direkt im *E. coli* Expressionsvektor pQE30NASTattB hergestellt. Dieses bakterielle Expressionssystem wurde wegen seiner guten Eignung für den Hochdurchsatz ausgewählt (vgl. 1.3.2). Eine weitere cDNA-Bibliothek aus dem Pistillgewebe von *Arabidopsis* wurde in den Entryvektor pENTR1a des GATEWAY-Systems kloniert und steht für zukünftige funktionelle Studien in verschiedensten Expressionssystemen zur Verfügung.

Die Charakterisierung von jeweils 96 Klonen beider cDNA-Bibliotheken bezüglich der durchschnittlichen Insertlänge lieferte Werte von 1,01 Kbp (Pistill) bzw. 1,06 Kbp (Infloreszenz). Diese Werte liegen damit relativ nahe an der durchschnittlichen Genlänge von 1,3 Kbp des *Arabidopsis*-Genoms (*Arabidopsis*-Genome-Initiative, 2000), was für eine gute Qualität der hergestellten cDNA-Bibliotheken spricht. Ähnliche durchschnittliche Genlängen wurden auch in den Genomen von *Oryza sativa* L. *ssp. indica* (1,1 Kbp; (Yu *et al.*, 2002)) und *Oryza sativa ssp. japonica* (1,3 Kbp; (Goff *et al.*, 2002)) gefunden. Auch die beiden von (Kramer, 2003) erstellten Gerste cDNA-Bibliotheken ergaben eine durchschnittliche Insertlänge von ungefähr 1 Kbp (Perikarp: 996 Bp und Embryosack: 932 Bp).

Der Anteil an Religanden konnte für 96 Klone der Infloreszenz cDNA-Bibliothek mit Hilfe einer PCR mittels genspezifischer Primer für *phyB* (Insert im Ausgangsvektor) ermittelt werden und lag bei 2,1%. Diese Religationen können sowohl auf eine unvollständige Restriktion des Vektors pQE30NASTattB als auch auf eine Kontamination des geschnittenen Vektors mit dem ausgeschnittenen Insert (*phyB*) zurückzuführen sein.

Dagegen wurde der Anteil der Religanden der 95 Klone der Pistill cDNA-Bibliothek nur ausgehend von der theoretischen PCR-Produktlänge der Religanden ermittelt und lag bei 6,3%. Dabei ist es nicht auszuschließen, dass sich unter den 6,3 % einige Klone befinden, die keine Religanden sind, aber zufällig ein PCR-Produkt in derselben Größe lieferten. Aufgrund der unterschiedlichen Ermittlung der Religationenanteile sind die beiden Werte nicht unmittelbar miteinander vergleichbar.

Im Falle des Vektors pENTR1a wäre eine Negativselektion der Religanden aufgrund der Suizid-Funktion des *ccdB*-Genes zu erwarten gewesen; der hier erhaltene hohe Anteil von Religanden (6,3%) könnte jedoch auf unvollständige Aktivität des *ccdB*-Genes zurückzuführen sein.

Aufgrund der Testsequenzierung von 48 Pistill-Klonen wurde ersichtlich, dass basierend auf einer BLAST-Analyse allen sequenzierten Klonen einem AGI *gene code* zugeordnet werden konnten. Dabei zeigte sich die hohe Diversität der Klone. Nur ein Klon kam zweimal vor, die restlichen Klone (95,8%) waren je einmal vertreten (*Singeltons*). Der Anteil an Klonen in voller Länge lag bei 23,4%. Darüber hinaus zeigte sich, dass sich 31,3% der Klone im richtigen Leserahmen (d.h. der Leserahmen des RGS His₆-Tags) befanden.

In die Testsequenzierung der Infloreszenz cDNA-Bibliothek wurde eine größere Anzahl von Klonen (576) einbezogen, da diese cDNA-Bibliothek der Ursprung für ein Proteinexpressions-Uniklonset war. 418 Sequenzen mit ausreichender Qualität konnte ein AGI *gene code* zugeordnet werden. Bei 7 Sequenzen handelte es sich um *phyB*, welche den Religandenanteil darstellen und nicht in die weitere Auswertung eingingen. Die Analyse der Diversität der Klone mittels Cluster-Analyse, auf der Grundlage der AGI-*gene codes* ergab, dass 72,3% der Klone einmal vorkamen (*Singeltons*), während die restlichen 27,7% sich in 49 Cluster zusammenfassen ließen. 192 der 411 Klone (46,7%) lagen in voller Länge und 128 Klone (31,1%) im richtigen Leserahmen vor.

Ein Vergleich der Diversität der Infloreszenz cDNA-Bibliothek (72,3% *Singeltons*) mit der Pistill cDNA-Bibliothek (95,8% *Singeltons*) ist nicht sinnvoll, da die Anzahl der analysierten Klone der Pistill cDNA-Bibliothek zu gering dazu war. Dies ist auch eine mögliche Ursache für die Differenz der Klone, die in voller Länge vorliegen (Pistill: 23,4%; Infloreszenz: 46,7%). Weitere Gründe für einen geringeren Anteil an kompletten cDNA-Klonen können auch auf die mRNA Qualität (mögliche Degradation) und die cDNA-Synthese zurückzuführen sein. Bei der Synthese der Infloreszenz cDNA könnte die Reverse Transkriptase effektiver gewesen sein, welches die Häufigkeit erhöhte, ein Gen in voller Länge zu transkribieren.

Bezogen auf den Anteil der Klone, die sich im richtigen Leserahmen befanden, ist für beide in dieser Arbeit erstellten Bibliotheken (Pistill: 31,3% und Infloreszenz: 32,3%) kein direkter Vergleich mit anderen cDNA-Bibliotheken möglich, da für keine der bisher beschriebenen cDNA-Bibliotheken vor Neuordnung der Expressionsklone eine aussagekräftige Leserahmennaussage auf Sequenzbasis beschrieben wurde. Kramer 2003 testeten die Expression von 96 Klonen aus einer Gerste-Embryosack cDNA-Bibliothek und

zeigten, dass 18% dieser Klone eine im SDS-Polyacrylamidgel sichtbare Überexpressionsbande lieferten. Das lässt vermuten, dass in dieser Bibliothek mindestens 18% der Klone im richtigen Leserahmen vorliegen. Jedoch kann davon ausgegangen werden, dass der Anteil höher ist, da bei dieser Methode die schwach translatierten Proteine nicht erfasst werden. Im Gegensatz dazu kann es bei der Ermittlung des richtigen Leserahmens auf Sequenzbasis dazu kommen, dass Klone miterfasst werden, die zwar im richtigen Leserahmen sind aber aufgrund von Stopcodonen in der 5'-untranslatierte Region nicht exprimiert werden.

Die Pistill cDNA-Bibliothek setzte sich mit 660.800 Klonen aus deutlich mehr Klonen zusammen, als die Infloreszenz cDNA-Bibliothek, die aus 40.000 Klonen besteht. Als eine Hauptursache hierfür kann die Verwendung unterschiedlicher kompetenter Zellen angesehen werden. Die Transformationseffizienz der käuflich erworbenen DH5 α -Zellen war um den Faktor 10 höher als die der selbsthergestellten kompetenten SCS1/pSE111-Zellen. Weitere Faktoren können sowohl in den unterschiedlichen Vektoren, z.B. die variierende Vektorgröße (pENTR1a: 2.717 Bp und pQE30NASTattB: 3.575 Bp) als auch in der Ligationseffizienz zu finden sein. Ein weiterer, jedoch nicht so gewichtiger Faktor, um die unterschiedlichen Bibliotheksgrößen zu erklären, ist eine mögliche negative Selektion einiger Proteinexpressionsklone im pQE30NASTattB Vektor durch toxische oder das Wachstum beeinträchtigende Proteine, bei einer nicht vollständig unterdrückten Expression (*leaky expression*).

Da jedoch nur die Infloreszenz cDNA-Bibliothek in einen Expressionsvektor (pQE30NASTattB) kloniert wurde und der zeitliche und finanzielle Aufwand zu hoch gewesen wäre, alle 660.800 Pistill-Klone in einen Expressionsvektor zu überführen, wurden im Folgenden ausschließlich die Klone der Infloreszenz cDNA-Bibliothek verwendet.

Die Pistill cDNA-Bibliothek wurden in einen Entryvektor (pENTR1a) kloniert und vereinigt. Diese Klone stehen somit z. B. als Ressource für ein späteres Überführen in unterschiedliche Expressionsvektoren zur Verfügung.

Die positiven Ergebnisse, die aus der Testsequenzierung der Infloreszenz Klone gewonnen wurden, belegen die gute Qualität dieser cDNA-Bibliothek. Ausgehend von dieser cDNA-Bibliothek wurden im weiteren Verlauf potentielle Expressionsklone zur Erstellung einer Proteinexpressions-Unterbibliothek detektiert.

4.1.2 Identifizierung potentieller Expressionsklone und Herstellung der Proteinexpressions-Unterbibliothek

Die Klonierung der Infloreszenz cDNA-Bibliothek erfolgte in den *E. coli* Expressionsvektor pQE30NASTattB. *E. coli* wurde als Expressionswirt verwendet, da er eine stabile Expression ermöglicht, auf genetischer und biochemischer Ebene gut charakterisiert ist und eine hohe Reproduktivität aufweist. Aufgrund dieser Eigenschaften ist *E. coli* sehr gut für den hohen Durchsatz geeignet. Außerdem ist die im Haus entwickelte Pick- und Spotttroutine mit diesem Expressionssystem bereits etabliert.

Der verwendete Expressionsvektor ermöglicht die rekombinante Proteinexpression über Induktion des T5-Promoters mit IPTG. Darüber hinaus enthält der Vektor ein N-terminales RGS His₆-Tag, womit die Proteine nach der Expression fusioniert sind. Damit können die Fusionsproteine über *Immobilised Metal Affinity Chromatography* (IMAC), sowohl unter nativen, als auch unter denaturierenden Bedingungen aufgereinigt werden. Aufgrund der geringen Größe des Tags geht man davon aus, dass die Struktur und Funktion des fusionierten Proteins nicht beeinträchtigt wird und das Tag demzufolge nach der Aufreinigung nicht entfernt werden muss. Des Weiteren kann das Tag auch noch zur Detektion der rekombinanten Proteine mit Hilfe eines anti-RGS His₆-Antikörpers verwendet werden.

Die Notwendigkeit der Detektion der potentiellen Expressionsklone aus einer cDNA-Bibliothek ist darauf zurückzuführen, dass beim Klonieren der cDNA kein Einfluss darauf genommen werden kann, in welchem der drei Leserahmen die cDNA transferiert wird. Somit ist zu erwarten, dass etwa nur ein Drittel der klonierten Gene im richtigen Leserahmen vorliegen und daher als potentielle Expressionsklone identifiziert werden können. Dies konnte anhand der Testsequenzierung für die im Rahmen dieser Arbeit erstellten cDNA-Bibliotheken bestätigt werden (Pistill: 31,3% und Infloreszenz: 32,3% lagen im richtigen Leserahmen vor).

Die Klone der Infloreszenz cDNA-Bibliothek wurden mittels Robotertechnologie auf PVDF-Membranen gespottet. Anschließend wurde *in situ* die Expression der rekombinanten Fusionsproteine induziert, welche anschließend mit Hilfe des spezifischen anti-RGS His₆-Antikörpers identifiziert werden konnten.

Basierend auf dem Antikörper-*screening* der Hochdichte-Proteinfiler konnten 7.000 Klone als potentielle Expressionsklone identifiziert werden. Das entspricht einem Anteil von 17,5% der 40.000 Klone aus der cDNA-Bibliothek. Auch in den bereits oben erwähnten

Bibliotheken von (Büssow *et al.*, 1998; Gutjahr, 2003; Kramer, 2003) lag der Anteil an potentiellen Expressionsklonen bei ungefähr 20%. Mögliche Ursache dafür, dass die Anzahl an potentiellen Expressionsklonen niedriger ist als die zu erwartende Anzahl (ein Drittel) kann u.a. sein, dass gering exprimierende Proteine gar nicht oder in zu geringer Menge synthetisiert werden, um sie als Expressionsklone zu identifizieren. Berücksichtigt werden muss auch, dass ein Teil der exprimierten Proteine toxisch auf *E. coli* wirken kann bzw. die Proteine durch die Zellen degradiert werden können. Des Weiteren kann bei Klonen, die in voller Länge vorliegen, die 5'-untranslatierte Region zusätzliche Stopcodone enthalten, so dass der Klon das enthaltene Gen nicht exprimiert.

Neben dem *screening* der Hochdichte-Proteinfilter mit dem anti-RGS His₆-Detektionsantikörper, wurde ein weiterer Filtersatz mit einem anti-PhyB-Antikörper inkubiert. Somit konnten die Religanden (Ausgangsvektor mit Insert *phyB*) identifiziert und bei der Erstellung der entsprechenden Unterbibliotheken ausgeschlossen werden. 2.000 der 7.000 potentiellen Expressionsklone wurden als Religanden identifiziert. Dieser hohe Anteil an Religanden kann auf einen nicht vollständig geschnittenen Ausgangsvektor (*SalI* und *NotI*) zurückzuführen sein.

Aufgrund der systematischen Anordnung auf den klonalen Proteinfiltern war es möglich, die 5.000 potentiellen Expressionsklone (abzüglich der Religanden) zu identifizieren und diese in einer Proteinexpressions-Unterbibliothek neu anzuordnen. Um eine definierte Aussage über die Diversität der Bibliothek machen zu können und um die Erstellung eines Uniklonsets zu ermöglichen, wurde die gesamte Proteinexpressions-Unterbibliothek ausgehend vom 5'-Ende ansequenziert und analysiert.

4.1.3 Charakterisierung der Proteinexpressions-Unterbibliothek

Ausgehend von der Sequenzinformation wurde zunächst eine Qualitätskontrolle durchgeführt. 4.398 Sequenzen wurden hinsichtlich ihrer Qualität als ausreichend befunden. Von diesen Sequenzen wurde eine BLAST-Analyse gegen die MIPS-*Arabidopsis*-Protein-Datenbank durchgeführt. 99,6% der Sequenzen konnte ein AGI *gene code* zugeordnet werden. Bei den restlichen Sequenzen, denen kein AGI *gene code* zugeordnet werden konnte, könnte es sich um bisher noch nicht beschriebene Gene handeln.

Basierend auf der BLAST-Analyse konnte gezeigt werden, dass sich 61,8% der sequenzierten Ausgangsklone in Bezug auf den RGS His₆-Tag im richtigen Leserahmen befanden und dass 39,4% der Proteinexpressions-Unterbibliotheks-Klone in voller Länge vorlagen. Dabei zeigte sich, dass sich im Vergleich zur Testsequenzierung (vor dem Rearray) der Anteil vollständiger cDNA-Klone etwas verringert hat, von 46,7%

(vor dem Rearray) auf 39,4% (nach dem Rearray). Eine mögliche Ursache hierfür kann sein, dass bei einigen vollständigen cDNA-Klonen die 5'-untranslatierte Region transkribiert wird, welche Stopcodonen aufweisen kann. Diese Klone liefern kein Fusionsprotein mit dem RGS His₆-Tag und sind damit nicht detektierbar und wurden deshalb nicht in die Proteinexpressions-Unterbibliothek aufgenommen. Der Anteil der Klone im richtigen Leserahmen hat sich dagegen fast verdoppelt, von 32,3% (vor dem Rearray) auf 61,8% (nach dem Rearray). Diese Anreicherung zeigt, dass sowohl die Identifizierung der potentiellen Expressionsklone, als auch deren erneute Anordnung in eine Proteinexpressions-Unterbibliothek außerordentlich wichtig ist. Vergleichbare Studien, in denen auch eine große Anzahl von Expressionsklonen auf Sequenzbasis analysiert wurde, finden sich bei Kramer 2003 und Lüking *et al.* 2003b.

Kramer 2003 sequenzierten 4.100 putative Expressionsklone der Gerste-Embryosack cDNA-Bibliothek. Von 2.416 Sequenzen mit ausreichender Qualität lagen 43% im richtigen Leserahmen vor.

In der Studie von Lüking *et al.* 2003b wurden 2.413 Proteine der humanen fötalen Hirngewebe Bibliothek (Büssow *et al.*, 1998) auf einen Microarray gespottet. Dafür wurden die verwendeten Klone sequenziert und 2.303 der 2.413 Klone konnten einer Sequenz zugeordnet werden. Basierend auf einer BLAST-Analyse (gegen die nichtredundanten Proteindatenbank NRPROT) konnte für 1.603 Klone der Leserahmen ermittelt werden. Von diesen 1.603 Klonen waren 55% im richtigen Leserahmen. Somit liegt der in der vorliegenden Arbeit erzielte Anteil von 61,8% Klonen mit richtigem Leserahmen deutlich über den Werten aus den Studien von Kramer 2003 und Lüking *et al.* 2003b, womit die gute Qualität des in dieser Arbeit durchgeführten Rearrays unterstrichen wird.

In der Studie von Lüking *et al.* 2003b lagen von den 882 Klonen, die sich im richtigen Leserahmen befanden, 382 (43%) in voller Länge vor. Im Vergleich dazu befanden sich in der vorliegenden Arbeit 926 Klone (34,1%) in voller Länge, wenn man von den 2.715 Sequenzen ausgeht, die im richtigen Leserahmen vorlagen. Somit war im Vergleich zu Lüking *et al.* 2003b zwar der Anteil der Klone, die sich im richtigen Leserahmen befanden höher; allerdings war der Anteil der Klone, die in voller Länge vorlagen, geringer. Eine mögliche Begründung für den geringeren Anteil an Klonen, die in voller Länge vorliegen, könnte sein, dass die Reverse Transkriptase nicht so effektiv war bzw. bei der Größenfraktionierung der cDNA auch große cDNAs verloren gingen.

Eine mögliche Ursache dafür, dass sich nach der Neuordnung der Expressionsklone nicht alle Klone im richtigen Leserahmen befinden, liegt darin, dass einige der kleinen Proteine,

die nicht im richtigen Leserahmen vorliegen, exprimiert werden. Wenn die Länge dieser Proteine ausreicht, um nicht von *E. coli* abgebaut zu werden, können diese mit dem anti-RGS His₆-Antikörper detektiert werden und gelangen somit in das potentielle Expressionssubset (Büssow *et al.*, 1998; Lüking *et al.*, 2003b).

Zur weiteren Analyse wurden die 4.398 *Arabidopsis*-Sequenzen mit allen kodierenden Sequenzen von *Arabidopsis* zusammen geclustert. Diese Analyse zeigte, dass 31,7% der Expressionsklone einmal vorkamen, während die restlichen Klone mehrfach repräsentiert waren und sich in 637 Cluster zusammenfassen lassen.

Basierend auf diesen Analysen wurde aus der Proteinexpressions-Unterbibliothek das Uniklonset erstellt.

4.1.4 Erstellung des Uniklonsets

Als erstes Selektionskriterium für die Erstellung des Uniklonsets wurde die Position des ersten Stopcodons einbezogen. Dazu wurden die DNA-Sequenzen translatiert.

Basierend auf der Position des ersten Stopcodons wurden alle Sequenzen eliminiert, die ein Stopcodon bei ≤ 70 Aminosäuren aufwiesen (1.442 Klone, entspricht 32,8% der Expressionsklone). Somit wurde gewährleistet, dass nur solche Klone einbezogen wurden, die Proteine mit einer Größe von mindestens 7,7 kDa exprimieren. Diese Größe wurde gewählt, da Lüking *et al.* 2003b zeigten, dass bei der Identifizierung von Expressionsklonen auf klonalen Proteinfiltren auch Peptide (welche im Leserahmen des RGS His₆-Tags vorlagen) mit einer Größe von 5 – 8 kDa zu detektieren waren.

Die verbliebenen 2.956 Sequenzen wurden in die folgenden zwei Sequenzgruppen unterteilt:

1. Sequenzen, die bei der Cluster- und BLAST-Analyse gleiche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit aufwiesen (2.754 Sequenzen)
2. Sequenzen, die basierend auf der Cluster-Analyse keinem AGI *gene code* zugeordnet werden konnten bzw. Sequenzen, die bei der Cluster- bzw. BLAST-Analyse unterschiedliche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit hatten (202 Sequenzen).

Diese Sequenzgruppen sind die Grundlage zur Erstellung des Uniklonsets. Die Selektionskriterien, mit denen die Sequenzen für das Uniklonset aus diesen beiden Sequenzgruppen ausgewählt wurden, sind im folgenden kurz dargestellt, um die Daten im Anschluss daran besser diskutieren zu können.

Die Selektionskriterien waren für die Sequenzen der beiden Sequenzgruppen unterschiedlich. Da bei den Sequenzen der ersten Sequenzgruppe BLAST- und Cluster-AGI *gene code* identisch waren, erfolgte hier die Selektion basierend auf der BLAST- und der

Cluster- BLAST-Analyse. Dagegen war in der zweiten Sequenzgruppe (BLAST- und Cluster-AGI *gene code* nicht identisch) nicht sicher, dass die Auskunft der BLAST-Analyse korrekt war und somit wurden die Parameter dieser Analyse nicht mit einbezogen. Hier gingen als Selektionskriterien das Stopcodon und die Clustergröße ein.

Bei den Sequenzen der ersten Gruppe wurden alle Klone, die nicht im Leserahmen des RGS His₆-Tags (11,6%) vorlagen, sowie alle Klone, welche einen E-Wert (Irrtumswahrscheinlichkeit) über e^{-20} aufwiesen und somit als nicht sicher genug galten, nicht berücksichtigt.

Ausgehend von der Cluster-Analyse wurden alle Klone die je einmal vorkamen (37,5%; *Singeltons*), direkt, ohne weitere Selektion, in das Uniklonset transferiert. Die restlichen Sequenzen ließen sich in 384 Cluster zusammenfassen. Da in dem zu erstellenden Uniklonset jedes Gen möglichst nur einmal vertreten sein sollte, waren hier weitere Selektionskriterien erforderlich. Diese waren der Startpunkt des Subjekts (Treffer-Sequenz), der Startpunkt des Querys (Abfrage-Sequenz) und der E-Wert aus der BLAST-Analyse. Die Kriterien waren in der eben beschriebenen Reihenfolge festgelegt und sollten alle so klein wie möglich sein. Im Falle des Startpunktes des Subjekts heißt das, je kleiner der Wert, desto länger ist der Klon bzw. im Idealfall, wenn der Wert gleich eins ist, liegt das klonierte Gen in voller Länge vor. Beim Startpunkt des Querys bedeutet der kleinste Wert, dass der Klon mit der kürzesten 5'-untranslatierten Region genommen wurde. Sollte aufgrund der Aussage bezüglich des Startpunktes des Subjekts und dem Startpunkt des Querys keine Sequenzauswahl innerhalb eines Clusters möglich sein, wurde als weiteres Kriterium der E-Wert nochmals einbezogen. Es wurde der Klon mit dem kleinsten E-Wert aus dem Cluster genommen. Damit wurde der Grad der Wahrscheinlichkeit, dass es sich bei dem ausgewählten Klon um das bei der BLAST-Analyse gefundene Gen handelt, nochmals erhöht (vgl. 3.4.2.2).

Ausgehend von den 384 Clustern wurden jedoch mehr als 384 Klone in das Uniklonset transferiert. Insgesamt wurden 419 Klone für das Uniklonset ausgewählt. Das ist darauf zurückzuführen, dass im Falle von 21 Clustern den Sequenzen innerhalb eines Clusters mehr als ein BLAST-AGI *gene code* zugeordnet werden konnte und dass aus diesen Clustern mehrere Klone (ein Klon pro BLAST-Hit) nach den oben genannten Kriterien selektiert wurden. Mit dieser Vorgehensweise sollten die wenigen Fälle berücksichtigt werden, bei denen unterschiedliche Gene mittels Cluster-Analyse aufgrund hoher Sequenzähnlichkeiten in eine Cluster eingeordnet wurden. Diese verschiedenen Gene sollten im Uniklonset erfasst werden. Jedoch wird hierbei auch in Kauf genommen, dass einige wenige Gene mehrfach in

dem Uniklonset repräsentiert sind. Diese Mehrfachrepräsentierung kommt dadurch zustande, dass im Laufe der Evolution einige der Gene auf den fünf Chromsomen mehrfach verliegen und somit unterschiedlichen *Accession*-Nummern zugeordnet werden können (Mayer, 2001).

Von den Sequenzen der ersten Sequenzgruppe gingen somit 48,1% der Klone in das Uniklonset ein.

Wie bereits beschrieben, ging die Selektion der Sequenzen der zweiten Sequenzgruppe von weniger Kriterien aus. Die Kriterien waren hier: die Position des Stopcodons und die Clustergröße. Die zweite Sequenzgruppe ließ sich in zwei weitere Untergruppen unterteilen. In der einen Untergruppe konnten den Sequenzen mittels Cluster-Analyse kein AGI *gene code* zugeordnet werden. Für die Sequenzen der anderen Untergruppe konnte zwar auf Grundlage der Cluster-Analyse ein AGI *gene code* ermittelt werden, dieser war jedoch nicht identisch mit dem der BLAST-Analyse.

Von beiden Untergruppen wurden auch wieder alle *Singeltons* direkt in das Uniklonset transferiert (82 der 202 Klone). Die restlichen Klone lagen in Clustern vor. Von den Clustern, die sich aus weniger als sechs Sequenzen zusammensetzten, gingen alle Klone in das Uniklonset ein. Dagegen gingen bei Clustern mit mehr als sechs Klonen nur die Klone ein, die kein Stopcodon besaßen. Zusätzlich zu den 82 *Singeltons* gingen somit noch 92 weitere Klone in das Uniklonset ein. Das bedeutet, dass von Clustern der Sequenzgruppe 2 teilweise mehr als ein Klon in das Uniklonset eingehen konnte. Diese Vorgehensweise sollte ermöglichen, dass sich unter den ausgewählten Klone auch Klone im richtigen Leserahmen befanden, da hier keine Selektion bezüglich der Leserahmen möglich war.

Ausgehend von 2.956 potentiellen Expressionsklonen der Proteinexpressions-Unterbibliothek, deren Sequenzen qualitätsgerecht waren und kein Stopcodon bei ≤ 70 Aminosäuren besaßen, gingen 1.502 Klone (50,7%) in das Uniklonset ein. Zusätzlich zu den 1.502 Uniklonset-Klonen kamen noch 192 Klone ((Kersten *et al.*, 2003) und Klone von Transkriptionsfaktoren) hinzu, die in voller Länge vorlagen. Diese nunmehr insgesamt 1.694 Klone bildeten zusammen das Erweiterte Uniklonset.

4.1.5 Herstellung von Protein-Microarrays des Erweiterten Uniklonsets

Alle Klone des Erweiterten Uniklonsets wurden im 96 *well* Format exprimiert und die Proteine unter denaturierenden Bedingungen aufgereinigt. Ausgehend von drei zufällig ausgewählten 96 *well* Mikrotiterplatten wurde mit Hilfe des Bradford-Tests eine Proteinausbeute von durchschnittlich 8,5 μg bestimmt. Diese lag jedoch unter der erzielten durchschnittlichen Ausbeute von (Kersten *et al.*, 2003) mit 24,5 μg bei 96 *Arabidopsis*-

Klonen. Diese Diskrepanz könnte auf das vergleichsweise kleinere Elutionsvolumen zurückzuführen sein, welches in der vorliegenden Arbeit während der manuellen Reinigung eingesetzt wurde (80 µl Elutionspuffer). Die Proteine von Kersten *et al.* 2003 dagegen wurden unter Verwendung des BioRoboter 8000 (Qiagen) in 350 µl eluiert. Trotz der etwas niedrigeren Gesamtausbeute erfolgte in der vorliegenden Arbeit eine bewusste Entscheidung zugunsten der manuellen Aufreinigung, da hiermit höhere Proteinkonzentrationen erzielt werden können (Uniklonset: 106 µg/ml, Kersten *et al.* 2003: 70 µg/ml). Somit können Microarrays mit einem höheren Proteingehalt generiert werden.

Die 1.694 Proteine des Erweiterten Uniklonsets wurden in einem 11 x 11-Pattern gespottet. Nach dem Spotten der Proteine konnten durchschnittlich 95% der His-getagten Proteine auf den Microarrays mit Hilfe eines anti-RGS His₆-Antikörpers nachgewiesen werden. Dieses Ergebnis belegt, neben den ermittelten Proteinkonzentrationen unter Verwendung des Bradford-Test (Kap. 4.2), dass die Hochdurchsatz-Proteinaufreinigung effizient genug war, um die gewonnen Proteine auf Microarrays nachzuweisen.

4.2 Phosphorylierungsstudien als Funktionelle Analyse

Ausgehend von den aufgereinigten Proteinen des Erweiterten Uniklonsets erfolgten funktionelle Analysen der Proteine unter Verwendung der Microarray-Technologie. Dazu wurden die Proteine des Erweiterten Uniklonsets mittels eines Microarrayers auf Nitrozellulose ähnliche FASTTM-Slides gespottet. Anschließend sollte die Phosphorylierung der gespotteten Proteine mit ausgewählten Kinasen untersucht werden.

4.2.1 Vergleich von Phosphorylierungsstudien

Beim Vergleich der bisher entwickelten Protein-Microarray-basierten Phosphorylierungs-Assays mit dem im Rahmen dieser Arbeit entwickelten Assay, konnten einige Unterschiede festgestellt werden. Diese Unterschiede beziehen sich auf die Oberflächen der Microarrays, die Anzahl der gespotteten Proteine und die Detektion der phosphorylierten Targets.

(Zhu *et al.*, 2000) verwendeten für die Phosphorylierungs-Assays PDMS-beschichtete *nanowell*-Microarrays. Bei diesem Assay wurden 119 Hefe Proteinkinasen und 17 verschiedene Substrate eingesetzt. Die in den *nanowells* kovalent gebundenen Substrate wurden unter Verwendung der Hefe Kinasen und γ -ATP (³³P) phosphoryliert. Sie konnten, unter Verwendung eines Phosphorimager zeigen, dass 32 der 119 Kinasen ein bis zwei der 17 Substrate phosphorylieren (26,9%).

MacBeath und Schreiber 2000 führten ihre Phosphorylierungsstudien auf BSA-N-hydroxysuccinimid (BSA-NHS) beschichteten Microarrays durch. Dabei wurden drei unterschiedliche Kinase-Substrat-Interaktionspaare untersucht, welche jeweils in vier identischen Spots pro Microarray vorlagen. Anschließend wurde jeder dieser Microarrays mit einer der drei Kinasen und γ -ATP (³³P) phosphoryliert. Dabei konnten sie zeigen, dass jeweils nur die spezifischen Substrate der entsprechenden Kinasen phosphoryliert wurden. Zur Detektion der phosphorylierten Substrate verwendeten MacBeath und Schreiber 2000 eine Photoemulsion, in welche die Microarrays gegeben wurden. Anschließend wurden die Microarrays wie ein Foto entwickelt. Diese Microarrays wurden dann mit einem Lichtmikroskop in Teilbereichen aufgenommen, und diese später zusammengefügt. MacBeath und Schreiber 2000 diskutieren diese recht aufwendige und nicht für eine Untersuchung im Hochdurchsatz geeignete Detektion, damit dass zu dem damaligen Zeitpunkt eine ausreichende Auflösung weder mittels Röntgenfilm noch mittels Phosphorimager erzielt werden konnte.

Bei den Phosphorylierungsstudien von Kramer *et al.* 2004 wurden 768 rekombinante Gerste-Proteine auf FASTTM-Slides gespottet und anschließend mit Gerste Casine Kinase2 α

phosphoryliert. Dabei wurden 384 Proteine jeweils vierfach in einem 10 x 10-Pattern auf einen Microarray transferiert. Bei der Detektion der beiden Microarrays mittels Röntgenfilm konnten 21 potentielle Targets identifiziert werden (2,7%).

Im Rahmen der vorliegenden Arbeit wurden 1.694 *Arabidopsis*-Proteine FAST™-Slides gespottet. Die Proteine wurden in einem 11 x 11-Pattern gespottet. Dabei wurden auf jeden Microarray zwei identische Felder gespottet, wobei sich in jedem Feld ein Spot pro Protein befand. Die phosphorylierten Proteine wurden anschließend unter Verwendung eines Phosphorimagers (Bioimager FLA8000) detektiert. Diese Detektion wurde gewählt, da die Detektion mittels Phosphorimager 10 – 100 x empfindlicher als der Röntgenfilm ist, und daher eine wesentlich kürzere Expositionszeit benötigt wird bzw. auch schwächere Signale detektiert werden können. Zudem ist bei dieser Methode das System weniger schnell gesättigt, so dass sowohl sehr starke als auch sehr schwache Signale mit einer Exposition erfasst werden können. Außerdem geht der lineare Signalbereich über 5 Größenordnungen (100.000 : 1). Somit ist der Phosphorimager für Quantifizierungen besser geeignet als die Verwendung von Röntgenfilmen. Hiermit konnte eine der Aufgaben dieser Arbeit dahingehend gelöst werden, die Signalerfassung der phosphorylierten *Arabidopsis*-Proteine mit einem Phosphorimager durchzuführen. Diese Signalerfassung stellt die Grundlage des ebenfalls in dieser Arbeit entwickelten Quantifizierungssystems dar.

Die bisher beschriebene qualitative und leichter visuell zu detektierende Variante, die Duplikate eines Proteins nebeneinander zu spotten (Kramer *et al.*, 2004; MacBeath und Schreiber, 2000), konnte in der vorliegenden Arbeit aufgrund der Optimierung der Reproduzierbarkeit und Quantifizierung, nicht beibehalten werden. In der vorliegenden Arbeit wurden die *Arabidopsis*-Proteine zunächst in einem dichten Pattern (14 x 14 und 15 x 15) auch jeweils vierfach in ein Feld gespottet, dabei zeigte sich bei intensiven radioaktiven Signalen ein Überstrahlungseffekt. Dieser hatte zur Folge, dass Spots mit intensiven Signalen nicht quantifiziert werden konnten. Um diesen Effekt zu vermeiden, wurden die *Arabidopsis*-Proteine in den weiteren Versuchen zwar in Duplikaten, jedoch räumlich voneinander getrennt auf die Microarrays gespottet. Des Weiteren wurde als Spotdichte ein 11 x 11-Pattern festgelegt. Mit diesem verwendeten Spotpattern ist eine parallele Analyse von 1.728 Proteinen (3.456 Spots, ohne *guide dots* und Kontrollen) pro Microarray möglich. Im Gegensatz zu Kramer *et al.* 2004, hier wurden 384 Proteine auf einen Microarray gespottet und phosphoryliert, ist somit eine effektivere Identifikation von potentiellen Phosphorylierungs-Targets möglich.

Alle oben beschriebenen Phosphorylierungs-Detektionen erfolgten unter Verwendung von radioaktiv markiertem ATP. Da diese Methode immer noch die einfachste und sensitivste Detektions-Methode ist, jedoch gesundheitliche Risiken mit sich bringt, werden Detektions-Alternativen gesucht. So verwendeten Houseman *et al.* 2002 zur Detektion von phosphorylierten Substraten auf den Microarrays einen anti-Phosphotyrosin-Antikörper, welcher anschließend mit einem zweiten Fluoreszenz markierten Antikörper inkubiert und detektiert wurde. Auch Lesacherre *et al.* 2002 setzen Fluoreszenz markierte anti-Phosphoserin- und anti-Phosphotyrosin-Antikörper zur Detektion von phosphorylierten Peptiden ein.

Vorteile der Detektion phosphorylierter Substrate mittels Fluoreszenz sind die Einschränkung der gesundheitlichen Risiken und die Möglichkeit, herkömmliche Microarrayscanner zur Auswertung von Phosphorylierungsstudien einzusetzen. Jedoch ist diese Detektions-Methode noch in der Entwicklung und noch nicht mit radioaktiv markierter Detektion zu vergleichen. Des Weiteren muss bedacht werden, dass bei der Detektion mittels Antikörper nur die Targets ermittelt werden können, die an einem bestimmten Epitop phosphoryliert werden. Somit ist diese Antikörpermethode zwar zur Identifizierung potentieller Phosphorylierungs-Targets mit bekannten Epitopen geeignet, aber nicht als *screening tool* zur Erfassung möglichst vieler Phosphorylierungs-Targets.

Um die Spezifität von Kinasen zu ermitteln, können auch Peptid-Microarrays verwendet werden (Falsey *et al.*, 2001; Houseman *et al.*, 2002; Lesacherre *et al.*, 2002; Lizcano *et al.*, 2002).

Lizcano *et al.* 2002 untersuchten die Substratspezifität der humane Serin/Threonin Proteinkinase Nek6 unter Verwendung von Peptid-Microarrays. Bekannt ist, dass diese Kinase *in vivo* die hydrophoben Motive der 3 N-terminalen Leuzin-Reste der S6K (p70 *ribosomal S6 kinase*) und SGK (*serum- and glucocorticoid-induced protein kinase*) phosphoryliert. Auf den Peptid-Arrays mit 710 Peptiden, mit bekannten Phosphorylierungsstellen humaner Proteine, konnten die Phosphorylierung von 7 Peptiden durch die Nek6 gezeigt werden.

Houseman *et al.* 2002 charakterisierten die Spezifität der c-Scr (*nonreceptor tyrosin kinase*), indem sie drei unterschiedliche Peptide auf Microarrays transferierten, wobei ein Peptid ein bekanntes Substrat der c-Scr darstellt. Unter Verwendung eines spezifischen Phosphotyrosin-Antikörpers konnte die spezifische Phosphorylierung des Substrat-Peptids gezeigt werden.

Ein Vorteil der von uns verwendeten Protein-Arrays im Vergleich zu der Peptid-Arrays ist, dass keine Kenntnisse über die Sequenzinformation der Targetsequenzen vorhanden sein müssen. Darüber hinaus sind die Arrays mit rekombinaten Proteinen nicht so teuer wie die Arrays mit Peptiden. Des weiteren sind Protein-Arrays universeller einsetzbar, da diese auch für andere funktionelle Studien, wie zur Untersuchung der Wechselwirkung der Proteine mit anderen Molekülen wie Proteine, DNA und kleine Moleküle, verwendet werden können.

4.2.2 Phosphorylierungsstudien mit den Proteinen des Erweiterten Uniklonsets

Für die in dieser Arbeit durchgeführten Phosphorylierungsstudien wurden die Proteinkinasen PKA, MPK3 und MPK6 gewählt.

Für die Phosphorylierungsstudien wurde zunächst der von Kramer *et al.* 2004 etablierte Microarray-basierte Phosphorylierungs-Assay weiterentwickelt, wobei unter Verwendung der Proteinkinase A ein Quantifizierungssystem etabliert wurde. Anschließend erfolgte die Identifizierung potentieller Phosphorylierungs-Targets der *Arabidopsis* Mitogen-aktivierten Proteinkinasen (MPK3 und MPK6) ausgehend von allen gespotteten Proteinen des Erweiterten Uniklonsets. Zu jedem Phosphorylierungs-Versuch wurde generell ein Kontroll-Microarray mitgeführt. Dieser Phosphorylierungsansatz enthielt γ -ATP (^{33}P), aber keine Kinase. Da in keinem dieser Kontroll-Microarrays Signale detektiert wurden, kann davon ausgegangen werden, dass unter den Versuchsbedingungen keine Bindung des γ -ATP (^{33}P) an die gespotteten Proteine bzw. keine Autophosphorylierung erfolgte.

Für alle drei Kinasen wurde je ein Set zusammengestellt (PKA-Set, MPK3-Set und MPK6-Set), welches aus den identifizierten potentiellen Phosphorylierungs-Targets der entsprechenden Kinasen bestand.

4.2.2.1 Quantifizierung mittels Proteinkinase A

In den Phosphorylierungsstudien unter Verwendung der Proteinkinase A, welche zur Etablierung des Quantifizierungssystems eingesetzt wurden, konnten 35 potentielle Phosphorylierungs-Targets ermittelt werden. Zur Bestätigung wurden diese erneut kultiviert, aufgereinigt, auf eine PVDF-Membran transferiert und dort erneut phosphoryliert. Von 35 potentiellen Phosphorylierungs-Targets konnten 28 bestätigt werden. Dabei können falschpositive Proteine ausgeschlossen werden, da alle im Coomassie-Brilliant-Blau gefärbten Gel identifizierten Protein phosphoryliert wurden, d.h. jedes im Coomassie-Brilliant-Blau gefärbten Gel visuell identifizierte Proteine wurde phosphoryliert. Für die Proteine, die weder im Coomassie-Brilliant-Blau gefärbten Gel noch im phosphorylierten Blot detektiert werden konnte, kann keine Aussage gemacht werden. Mögliche Ursachen

dafür, dass 7 Targets nicht visuell bestätigt werden konnten, kann einerseits sein, dass die radioaktiven Signale der phosphorylierten Proteine auf dem Blot zu gering waren, um detektiert zu werden. Andererseits kann es sein, dass die Proteine für diese erneute Aufreinigung entweder nicht ausreichend exprimiert wurden, die Aufreinigung nicht effizient war oder dass die aufgetragene Proteinmenge nicht ausreichend war. Ein weiterer Hinweis für die Richtigkeit der identifizierten Targets ist, dass alle im Coomassie-Brilliant-Blau gefärbten Gel sichtbaren Banden der Negativkontrollen keine Signale im Radiogramm zeigten. Dabei ist zu beachten, dass die Negativkontrollen aufgrund der für den Phosphorylierungs-Assay gewählten Quantifizierungsmethoden als negativ gewertet wurden.

Zusätzlich wurde die Qualität der identifizierten Targets durch den Vergleich der sich in dem PKA- und in dem Uniklonset befindlichen Sequenzen kontrolliert. Dabei wurden sowohl eine Motivsuche von vier bekannten Konsensussequenzen (Kallus, 2001) durchgeführt als auch eine Kontrolle der Serin- und Threoninegehalte. Bei der Motivsuche konnte für alle vier Konsensussequenzen eine Anreicherung der entsprechenden Motivgehalte im PKA-Set ermittelt werden. Diese Anreicherung zeigt, dass die Identifizierung der potentiellen Targets generell möglich ist. Jedoch zeigte sich auch, dass jede der vier Konsensussequenzen häufiger als 35 x (Anzahl der Targets im PKA-Set) im Uniklonset vorkommen, d.h. nicht jedes Protein, welches eines dieser Motive aufwies, wurde phosphoryliert. Das kann darauf zurückzuführen sein, dass lediglich das Vorliegen des Motivs nicht ausreichend für die jeweilige Phosphorylierung ist. Es ist aber auch nicht auszuschließen, dass das gewählte quantitative Kriterium zur Charakterisierung eines Proteins als Target (nur die Proteine, deren mittlere Signalintensität über der 10fachen Standardabweichung des mittleren Hintergrundsignals aller Unterfelder lagen, wurden als potentielle Phosphorylierung-Targets identifiziert) evtl. zu streng ist. Da es jedoch Ziel dieser Arbeit war, ein *screening tool* zu etablieren, welches zu einer quantifizierten Eingrenzung der Anzahl von potentiellen Phosphorylierungs-Targets dienen sollte, die anschließend unabhängig verifiziert werden sollten, wurde dieses strenge Kriterium beibehalten. Bei dieser Vorgehensweise kann man annehmen, dass auch der Anteil an falschpositiven Targets eingeschränkt wird. In weiteren vergleichenden Studien sollte man versuchen, für dieses Kriterium den exakten Grenzwert zu ermitteln, wodurch es zu einer Maximierung der Selektion aller sich im Set befindlichen potentiellen Targets kommen könnte.

Der Vergleich des Serin- und Threoningehaltes im PKA- und Uniklonset zeigte eine Anreicherung des Seringehaltes, was als Hinweis darauf angesehen werden könnte, dass die Proteinkinase A zwar zu Serin/Threonin-Proteinkinase gehört, jedoch die Serin-Reste vielleicht bevorzugt phosphoryliert.

4.2.2.2 Phosphorylierungsstudien unter Verwendung der *Arabidopsis* Kinasen MPK3 und MPK6

Wie bereits in der Einleitung erläutert, gehören die MAP-Kinasen zu den Serin/Threonin-Proteinkinasen. Beide MAP-Kinasen (MPK) MPK3 und MPK6 werden in eine der vier Gruppen (Gruppe A), in welche die MPKs unterteilt werden, eingeordnet (vgl. 1.5.1; (MAPKgroup, 2002). Die MPKs sind Bestandteile von Signaltransduktionskaskaden, den MAPK-Kaskaden. Diese Kaskaden bestehen aus MAPKK-Kinasen (MAPKKK), MAPK-Kinasen (MAPKK) und MAP-Kinasen (MPK) und sind an der Weiterleitung der durch Rezeptoren oder Sensoren erkannten externen Stimuli von der Zelloberfläche zum Zellkern beteiligt. Im Zellkern kann es dann zur Phosphorylierung von Transkriptionsfaktoren kommen. Die Weiterleitung des Stimulus erfolgte über sogenannte *second messenger*, durch die es u. a. zu Änderungen in der Ionenpermeabilität der Plasmamembran (Nünberger und Scheel, 2001) oder zu Akkumulationen von reaktiven Sauerstoffspezies (ROS, z.B. Superoxidanionen, Hydroxylradikale oder Wasserstoffperoxid) kommen kann (Neill *et al.*, 2002). Die ROS können direkt toxisch auf das Phytopathogen wirken oder in die Verstärkung der pflanzlichen Zellwand involviert sein (Lamb und Dixon, 1997).

Durch diese Weiterleitung kommt es zur Aktivierung von Abwehrmechanismen (Zwerger und Hirt, 2001). Solche Kaskaden ermöglichen es den Pflanzen, auf verschiedene biotische und abiotische Stressfaktoren zu reagieren. Eine Beteiligung von MAP-Kinasen an der pflanzlichen Signaltransduktion konnte für eine große Anzahl von Stimuli beschrieben werden. Beispielsweise für mechanischen und osmotischen Stress (Bogre *et al.*, 1996; Ichimura *et al.*, 2000; Mizoguchi *et al.*, 1996; Tena und Renaudin, 1998), Trockenheit und Kälte (Ichimura *et al.*, 2000; Jonak *et al.*, 1996; Mizoguchi *et al.*, 1996), sowie Elicitor- als auch Pathogenkontakt (Asai *et al.*, 2002; Nühse *et al.*, 2000; Romeis *et al.*, 1999; Zhang *et al.*, 1998). Des Weiteren sind sie auch bei der Weitergabe der Pflanzenhormone, wie Ethylen, Abszissinsäure und Auxin beteiligt (Jonak *et al.*, 2002; Lu *et al.*, 2002; Tena *et al.*, 2001).

Bei den in der vorliegenden Arbeit durchgeführten Phosphorylierungsstudien konnten für das MPK3-Set 48 potentielle Phosphorylierungs-Targets und für das MPK6-Set 39 potentielle Phosphorylierungs-Targets identifiziert werden. Die geringere Anzahl an

identifizierten Targets bei dem MPK6-Set kann einerseits darauf zurückzuführen sein, dass die Kinase weniger Targets phosphoryliert oder für die MPK6 weniger potentielle Phosphorylierungs-Targets im Erweiterten Unikonset vorlagen, als für die MPK3. Andererseits bleibt zu bedenken, dass die rekombinate MPK6 mit einer geringeren Aktivität vorlag (persönliche Mitteilung unserer Kooperationspartner Dr. Justin Lee und Prof. Dr. Dierk Scheel), daher zwar eine höhere Konzentration eingesetzt wurde, diese aber vielleicht nicht ausreichend war, um eine zur MPK3 vergleichbare Aktivität zu erreichen. Es kann aber auch sein, dass die in der vorliegenden Arbeit verwendeten Assay-Bedingungen (wie z. B. Kofaktoren) für die MPK6 weniger optimal waren, als für die MPK3, da die Bedingungen ausgehend vom Myelin-assoziiertes basisches Protein (MBP), einem artifiziellen Substrat der Kinasen, optimiert wurden.

Zur Bestätigung der identifizierten Targets in einem unabhängigen *in vitro* Testsystem wurden die 48 potentiellen Phosphorylierungs-Targets des MPK3-Sets erneut aufgereinigt, auf PVDF-Membranen transferiert und unter Verwendung der MPK3 phosphoryliert. Dies war leider nur für die MPK3 möglich, da zu diesem Zeitpunkt keine ausreichend aktive MPK6 mehr zugänglich war. Von 48 potentiellen Phosphorylierungs-Targets konnten 45 bestätigt werden. Mögliche Ursachen für die drei nicht identifizierten Targets sind wie bei dem Phosphorylierungs-Blot der Proteinkinase A zu diskutieren.

Bei den Literaturrecherchen zu den identifizierten potentiellen Phosphorylierungs-Targets zeigte sich, dass für die *Arabidopsis* MPK noch nicht bekannt ist, welche Targetproteine direkt phosphoryliert werden. Bisher gibt es nur Untersuchungen der durch Stress ausgelösten MAP-Kinasen-Signaltransduktionskaskade hinsichtlich bestimmter Gentranskripte, aus denen man erste Hinweis auf mögliche Targets gewinnen kann (z.B. Kimura *et al.*, 2003; Kreps *et al.*, 2002; Seki *et al.*, 2002).

Beispielsweise konnten sowohl die Aktivierung von MPK3 durch Trockenheit, Kälte, Salzstress und mechanische Beanspruchung (Mizoguchi *et al.*, 1996) als auch eine posttranslationale Aktivierung von MPK3 durch H₂O₂ beobachtet werden.

Des weiteren konnte für die Transkripte der MEKK1, MPK3 und MPK1 eine Induktion durch osmotischen Stress und mechanische Beanspruchung gezeigt werden. Zusätzlich sind die Transkripte der MEKK1 und MPK3 durch Kälte induzierbar [Mizoguchi *et al.*, 1996]. Darüber hinaus konnten die Transkripte der MPK3 und MPK4 durch Bestrahlung mit UV-Licht induziert werden (Ulm *et al.*, 2002). Teige *et al.* 2004 konnten sowohl eine Aktivierung der MKK2 durch Kälte, Salzstress oder durch die stressinduzierte MEKK1 als auch die davon ausgehende Induktion der MAPK4 und MAPK6 zeigen.

Zudem konnte eine Aktivierung der MAP-Kinasen nicht nur bei abiotischen sondern auch bei den biotischen Stressfaktoren beobachtet werden. Asai *et al.* 2002 beschrieben bei *Arabidopsis* die vollständige MAP-Kinasen-Signaltransduktionskaskade, welche durch den bakteriellen Elicitor Flg22 ausgelöst wird. Diese MAPK-Kaskade besteht aus den Komponenten MEKK1, MKK4/5 und MPK3/6, welche die transkriptionelle Aktivierung der WRKY-Transkriptionsfaktoren WRY22/29 induzieren. Die Übertragung des durch den Elicitor Flg22 ausgelösten Signals erfolgt über FLS2 (*flagellin sensitive 2*), eine LRR (*leucin-rich repeat*) Rezeptor ähnliche Kinase (Gómez-Gómez und Boller, 2000).

Neben dem eben beschriebenen bakteriellen Elicitor bei *Arabidopsis* wurden für die MPK6 Orthologen in *Medicago sativa*, SIMK (*stress induced MAPK*), und in Tabak, SIPK (*salicylate induced MAPK*), weitere Elicitoren beschrieben, welche zur Aktivierung dieser Kinasen führen. Die SIMK-Aktivierung erfolgt durch unterschiedliche pilzliche Elicitoren (Cardinale *et al.*, 2000). Für SIPK konnte eine Aktivierung durch allgemeine Elicitoren (Droillard *et al.*, 2000; Lee *et al.*, 2001) und durch TMV (Tabak-Mosaik-Virus)-Infektion (Zhang *et al.*, 1998) beschrieben werden. Auch bei Petersilie konnte eine Aktivierung der PcMPK6 durch die Elicitoren PEP-13 und HrpZ beobachtet werden (Kroj *et al.*, 2003).

Wan *et al.* 2004 vermutet eine Überlappung von der von Asai *et al.* 2002 beschriebenen Signaltransduktionskaskaden mit der Signaltransduktion von Chitin Elicitoren, welche bei der Übertragung des Chitin-Stimulus von den Zellwänden pathogener Pilze und Insekten ausgelöst wird.

Einige in dieser Arbeit identifizierten potentiellen Phosphorylierungs-Targets stehen im Zusammenhang mit bereits beschriebenen Komponenten der MAP-Kaskaden. Aus diesem Grund ist ihre Phosphorylierung durch die MPKs nicht unwahrscheinlich.

Das als potentielles Phosphorylierungs-Target der MPK3 identifizierte Protein: *leucine-rich repeat family protein contains leucine rich repeat (LRR) domains* (At4g03260; exprimiert von Klon: 312_J15 des Uniklonsets) könnte, aufgrund seiner LRR-Domäne, an einer von der Elicitoren ausgelösten-Signaltransduktionskaskade beteiligt sein (Asai *et al.*, 2002; Gómez-Gómez und Boller, 2000; Wan *et al.*, 2004). Als ein weiteres interessantes Target konnte das durch Klon 313_G13 exprimierte Protein (At4g11280) bei Phosphorylierung mit der MPK6 detektiert werden. Hierbei handelt es sich um die ACC Synthase 6 (*1-aminocyclopropane-1-carboxylate synthase 6*). Ouaked *et al.* 2003 konnten zeigen, dass eine MAPK-Kaskade an der Ethylen-Signaltransduktion beteiligt ist, in dem bestimmte MPKs in *Arabidopsis* und *Medicago* durch die Vorläufersubstanz des Ethylens ACC aktiviert werden.

Dieses sind die aktivierten MPKs SIMK (*stress induced MAPK*) und MMK (*Medicago MAPK2*) in *Medicago* und in *Arabidopsis* die MPK6 sowie eine weitere MAP-Kinase.

Des weiteren wurden die Histonen H2A (im MPK3-Set) und H3 (in beiden Sets) als potentielle Phosphorylierungs-Targets identifiziert. Für Histone ist bekannt, dass diese als Reaktion auf extrazelluläre Signale phosphoryliert werden. Yamagata *et al.* 2001 zeigten eine Phosphorylierung von Histonen durch die *MAPK kinase-like protein kinase* in Zellkultur von Sojabohnen. Ob die Histone tatsächlich in der Signaltransduktion der MPKs involviert sind, bleibt noch zu zeigen.

Auch ein Protein des MPK3-Sets (At5g62070; *calmodulin-binding family protein*; exprimiert durch Klon 313_F05) kann ein Bestandteil der MAPK-Kaskaden sein. Perruc *et al.* 2004 zeigten, dass bei *Arabidopsis* ein Calmodulin bindendes Protein (At2g41010) als Negativregulator bei der Stresstoleranz gegenüber Salz- und osmotischem Stress beteiligt ist. Ein eventueller Zusammenhang zwischen einer möglichen Phosphorylierung von Calmodulin-bindenden Proteinen durch MPKs und deren Rolle in der Regulation der Stresstoleranz muss noch bestätigt werden. Calmodulin könnte über die Bindung an das Calmodulin-bindende Protein oder aber auch über die Aktivierung der Ca²⁺-calmodulinabhängigen Proteinkinase (CaM-Kinase) einen Einfluss auf die Stresstoleranz haben. CaM-Kinasen können z.B. Serin- und Threoninreste in ihren Zielproteinen phosphorylieren und damit eine Enzymaktivierung bewirken. Daran könnten ebenfalls auch die MPKs beteiligt sein.

Die Lokalisierung der aktivierten MPKs im Zellkern läßt Transkriptionsfaktoren als Targets vermuten. Diese Vermutung wird auch durch die in dieser Arbeit als Phosphorylierungs-Targets gefundenen Transkriptionsfaktoren bestätigt. So konnten folgende Transkriptionsfaktoren identifiziert werden: Der Klon des MPK3-Sets 315_G09 (At5g66940; *Dof-type zinc finger domain-containing protein*), sowie die Klone des MPK6-Sets 315_P07 (At1g06170; *basic helix-loop-helix (bHLH) family protein*), 314_N21 (At3g60390; *homeobox-leucine zipper protein HAT3*), 313_N06 (At5g54630; *zinc finger protein-related*) und 311_F11 (At5g58620; *zinc finger (CCCH-type)*), wobei letzterer zusätzlich auch im MPK3-Set vorkommt und mit seinem Zinkfingermotiv eine DNA-bindende Domäne aufweist.

Wie sich anhand der Histone und der Transkriptionsfaktoren zeigte, wurden auch gemeinsame potentielle Phosphorylierungs-Targets für beide MPKs identifiziert. Insgesamt waren das 26 potentielle Phosphorylierungs-Targets. Dieses wurde jedoch auch vermutet, da beide MPKs zu einer Untergruppe einer Proteinklasse gehören.

Als weiteres gemeinsames potentiell Phosphorylierungs-Target wurde z. B. die Casein Kinase (CK; 314_G10; At5g44100) ermittelt. Eine direkte Beteiligung der Casein Kinase in der Signaltransduktion der MAPK-Kaskaden ist für Pflanzen noch nicht beschrieben. In HeLa-Zellen wurde jedoch eine direkte Wechselwirkung der p38 MAP-Kinase mit der CK2 und deren stressinduzierten Aktivierung durch diese MAP-Kinase gezeigt (Sayed *et al.*, 2000).

Ein weiterer Klon, der in beiden Sets zu finden war, ist der Klon: 313_E14 (At5g48990; *kelch repeat-containing F-box family protein*). F-Box Proteine sind häufig an der Degradation von cytosolischen Proteinen beteiligt (Andrade *et al.*, 2001). F-Box Proteine erkennen und binden die abzubauenen Proteine, welche häufig phosphoryliert sind (Willems *et al.*, 1999). Das am F-Box-Protein gebundene Protein wird anschließend mehrfach ubiquitiniert und zum Proteasomen transportiert. Diese Polyubiquitinierung dient als Degradationssignal für das Proteasomen. Ob die Phosphorylierung der F-Box-Proteine *in vivo* eine Rolle spielt, bleibt noch zu klären.

Als eine weitere Kontrolle hinsichtlich der Spezifität der identifizierten Targets der jeweiligen Kinasen wurden neben den gemeinsamen Phosphorylierungs-Targets der beiden MPKs (26 insgesamt) auch die gemeinsamen Targets der PKA mit denen der MPK3 bzw. der MPK6 bestimmt. Für die MPK3 wurden drei und für die MPK6 vier gemeinsame Targets mit der PKA ermittelt. Wie zu erwarten war, lieferte der Vergleich der unterschiedlichen Serin/Threonin-Proteinkinasen einen niedrigeren Anteil an gemeinsamen Targets, als der Vergleich der beiden MAP-Kinasen, die zu einer Kinasegruppe gehören.

Da alle drei Kinasen zu den Serin/Threonin-Proteinkinasen gehören, liegt die Vermutung nahe, dass bei den ermittelten gemeinsamen potentiellen Phosphorylierungs-Targets beider Gruppen ausschließlich Serin- und Threonin-Reste phosphoryliert wurden, was aber in der vorliegenden Arbeit nicht weiter untersucht wurde. Die genaue Ermittlung der phosphorylierten Aminosäure ist prinzipiell entweder mit Massenspektroskopie (Glinski *et al.*, 2003) oder mit phospho-spezifischen Antikörpern (Houseman *et al.*, 2002; Lesaichere *et al.*, 2002) möglich.

Die Ergebnisse der Phosphorylierungsstudien zeigen, dass sowohl der etablierte Kinase-Assay als auch das dazu gehörende Quantifizierungssystem reproduzierbar und spezifisch sind.

Des Weiteren sind alle durchgeführten Kontrollen ein Hinweis darauf, dass alle Schritte, die zur Entstehung des Uniklonsets (Detektion der potentiellen Expressionsklone mittels Hochdichte Proteinfiltern, die Sequenzanalyse und davon ausgehend die Erstellung des

Uniklonsets) erfolgreich waren. Außerdem belegen sowohl die Kontrollen als auch die Ergebnisse dieser Phosphorylierungsstudien, dass die im Rahmen dieser Arbeit verwendete Aufreinigungstechnik für den Hochdurchsatz geeignet war.

An dieser Stelle soll nochmals betont werden, dass es sich bei dem hier weiterentwickelten Mikroarray-basierten Phosphorylierungssystem nur um ein *screening tool* handelt, welches zu einer *in vitro* Identifizierung der potentiellen Kinasetargets dient. Die mit diesem Testsystem identifizierten potentiellen Phosphorylierung-Targets müssen jedoch in anschließenden Folgeversuchen *in vivo* weiter verifiziert werden. Das ist aus folgenden Gründen erforderlich: In der vorliegenden Arbeit wurden die Targets unter denaturierenden Bedingungen aufgereinigt. Es kann also nicht ausgeschlossen werden, dass unter nativen Bedingungen die Phosphorylierungsstellen für die Kinasen nicht zugänglich sind. Des Weiteren besteht auch die Möglichkeit, dass *in vivo* die hier identifizierten Targets nicht mit den Kinasen interagieren können, da sie in unterschiedlichen Zellkompartimenten lokalisiert sein könnten. Dennoch zeigt die Identifikation von bekannten Phosphorylierung-Targets, dass diese Methode als eine Möglichkeit zur Ermittlung von unbekanntem potentiellen Phosphorylierungs-Targets angesehen werden kann.

In vivo Methoden, die angewandt werden können, um die identifizierten potentiellen Phosphorylierungs-Targets zu verifizieren, sind z.B. Methoden zur Untersuchung von Protein-Protein Wechselwirkungen, z.B. in Hefe (Yeast-Two-Hybrid) oder in Pflanzen direkt. Hierbei kommt es zur Verwendung von mikrospektroskopischen Ansätzen wie Fluoreszenz-Resonanz-Energie-Transfer (FRET) und Fluoreszenz-Korrelations-Spektroskopie (FCS) (Hink *et al.*, 2002). Mittels FRET-Methoden kann eine abstandsabhängige Wechselwirkung zwischen zwei angeregten Farbstoffmolekülen detektiert werden, bei der die Anregungsenergie von einem Donormolekül auf eine Akzeptormolekül übertragen wird, ohne daß Photonen abgestrahlt werden. Bei FCS wird meist nur eine Molekül farblich markiert. Bei dieser Methode kann man die Bewegung einzelner Moleküle beobachten, so bewegen sich z. B. gebundene Moleküle vergleichsweise langsamer.

Eine weitere *in vivo* Methode zur Bestätigung der potentiellen Phosphorylierungs-Targets stellt die Klonierung der potentiellen Targets in Epitop-getagte Pflanzenexpressionsvektoren dar. Dazu könnte das GATEWAY-System benutzt werden, um eine effiziente und schnelle Subklonierung zu ermöglichen. Anschließend werden pflanzliche Protoplasten mit diesen Plasmiden transformiert, und die Phosphorylierung dieser Targets kann durch Zugabe von radioaktiv markiertem Phosphat zum Medium untersucht werden. Bedingungen, die zu einer

Aktivierung der MAP-Kinase führen (Elicitor, abiotische Stressbehandlung oder Kotransfektion in eine aktive *upstream* MAP Kinase Kinase) können genutzt werden, um die Beteiligung von MAP-Kinase bei der Phosphorylierung der potentiellen Targets zu untersuchen.

Bei der Validierung von Phosphorylierungs-Targets mittels Protein-Protein-Interaktionsstudien bleibt aber zu beachten, dass die Wechselwirkung einer Kinase mit einem Protein nicht zur Phosphorylierung des Proteins führen muß.

Die in der Arbeit hergestellten Klone stellen die bisher umfangreichste Sammlung von rekombinaten *Arabidopsis thaliana* Expressionsklonen dar und stehen für weitere Studien zur Verfügung.

Diese Arbeit, ist die erste systematische Studie zur Identifizierung von Phosphorylierungs-Targets der MAP-Kinasen im Hochdurchsatz. Zur Identifizierung wurde ein Quantifizierungssystem entwickelt, bei dem nur phosphorylierten Signale, die sich deutlich vom Hintergrund unterschieden, verwendet wurden, um die dazugehörigen Proteine als Phosphorylierungs-Targets zu identifizieren. Eine Lockerung des in der vorliegenden Arbeit sehr streng verwendeten *cutoff* Wertes für die Targetselektion, könnte die Anzahl der identifizierten Phosphorylierungs-Targets in Zukunft evtl. erhöhen.

Desweiteren konnten in der vorliegenden Arbeit bisher noch nicht beschriebene Targets der MAP-Kinasen identifiziert werden. Diese *in vitro* identifizierten Phosphorylierungs-Targets liefern eine interessante Grundlage für weitere *in vivo* Untersuchungen. Der Phosphorylierungs-Assay ist für verschiedene Kinasen anwendbar und kann in Zukunft zur Identifizierung von Phosphorylierungs-Targets weiterer *Arabidopsis thaliana* Kinasen verwendet werden.

4 Diskussion

Ziel der Arbeit war erstens die Herstellung von rekombinanten *Arabidopsis thaliana*-Proteinen und zweitens die Etablierung von funktionellen Studien im Hochdurchsatz. Aus diesem Grunde wurden diese Themen in zwei Abschnitten zusammengefasst und diskutiert. Zur Gewinnung rekombinanter Proteine im ersten Teil der Arbeit wurden cDNA-Bibliotheken aus der Pistill und der Infloreszenz von *Arabidopsis thaliana* erstellt und charakterisiert. Ausgehend von der Infloreszenz cDNA-Bibliothek wurde ein Uniklonset von Expressionsklonen erstellt. Zu diesem Zweck wurden potentielle Expressionsklone dieser Bibliothek unter Verwendung von Hochdichte-Proteinfiltern selektiert und in einer Proteinexpressions-Unterbibliothek neu angeordnet. Basierend auf BLAST- und Cluster-Analysen dieser Klone, erfolgte anschließend die Erstellung eines Uniklonsets aus 1.502 Klonen, das um 192 vollständige cDNA-Klone erweitert wurde. Alle 1.694 Klone dieses Erweiterten Uniklonsets wurden im Hochdurchsatz exprimiert und die entsprechenden rekombinanten Proteine aufgereinigt.

Im zweiten Abschnitt wurden mit den aufgereinigten Proteinen *Arabidopsis*-Protein-Microarrays erstellt, welche anschließend für Phosphorylierungsstudien *in vitro* verwendet wurden. Diese Studien wurden unter Verwendung verschiedener Kinasen durchgeführt: einerseits mit der gut beschriebenen Maus Proteinkinase A (PKA) und andererseits mit zwei *Arabidopsis* MAP-Kinasen (MPK3 und 6), die eine wichtige Rolle in der Signaltransduktion spielen. Ziel war es, nach Weiterentwicklung des Testsystems, unter Verwendung der PKA, potentielle Phosphorylierungs-Targets der MAP-Kinasen zu identifizieren.

4.1 Gewinnung rekombinanter *Arabidopsis* Proteine

4.1.1 Herstellung und Charakterisierung von zwei cDNA-Bibliotheken

Die cDNA-Bibliotheken repräsentieren im Idealfall die Gesamtheit der exprimierenden Gene eines Gewebes, eines Zelltyps oder eines Organismus' zu einem bestimmten Zeitpunkt. Basierend auf der Analyse dieser Bibliotheken, kann man wichtige Einblicke in den Ablauf biologischer Prozesse bekommen. So können z.B. neue Gene und deren endgültige Spleißprodukte identifiziert werden, aber auch die Expressionsintensität der Gene des entsprechenden Stadiums des jeweiligen Gewebes, Zelltyps bzw. Organismus' analysiert werden (Clark *et al.*, 1999).

cDNA-Bibliotheken können auch als Quelle zur Expression rekombinanter Proteine des entsprechenden Gewebes und Stadiums genutzt werden. Dazu werden die cDNAs entweder

direkt in einen geeigneten Expressionsvektor (Büssow *et al.*, 1998; Gutjahr, 2003; Kramer, 2003) oder zunächst in einen Entryvektor des GATEWAY-Systems kloniert. Ausgehend von diesem Entryvektor kann dann die cDNA über Rekombination je nach Bedarf in unterschiedliche Expressionsvektoren transferiert werden (Heyman *et al.*, 1999; Walhout *et al.*, 2000). Im Rahmen dieser Arbeit wurden beide Wege zur Herstellung von cDNA-Bibliotheken verwendet. Die Infloreszenz cDNA-Bibliothek wurde direkt im *E. coli* Expressionsvektor pQE30NASTattB hergestellt. Dieses bakterielle Expressionssystem wurde wegen seiner guten Eignung für den Hochdurchsatz ausgewählt (vgl. 1.3.2). Eine weitere cDNA-Bibliothek aus dem Pistillgewebe von *Arabidopsis* wurde in den Entryvektor pENTR1a des GATEWAY-Systems kloniert und steht für zukünftige funktionelle Studien in verschiedensten Expressionssystemen zur Verfügung.

Die Charakterisierung von jeweils 96 Klonen beider cDNA-Bibliotheken bezüglich der durchschnittlichen Insertlänge lieferte Werte von 1,01 Kbp (Pistill) bzw. 1,06 Kbp (Infloreszenz). Diese Werte liegen damit relativ nahe an der durchschnittlichen Genlänge von 1,3 Kbp des *Arabidopsis*-Genoms (*Arabidopsis*-Genome-Initiative, 2000), was für eine gute Qualität der hergestellten cDNA-Bibliotheken spricht. Ähnliche durchschnittliche Genlängen wurden auch in den Genomen von *Oryza sativa* L. *ssp. indica* (1,1 Kbp; (Yu *et al.*, 2002)) und *Oryza sativa ssp. japonica* (1,3 Kbp; (Goff *et al.*, 2002)) gefunden. Auch die beiden von (Kramer, 2003) erstellten Gerste cDNA-Bibliotheken ergaben eine durchschnittliche Insertlänge von ungefähr 1 Kbp (Perikarp: 996 Bp und Embryosack: 932 Bp).

Der Anteil an Religanden konnte für 96 Klone der Infloreszenz cDNA-Bibliothek mit Hilfe einer PCR mittels genspezifischer Primer für *phyB* (Insert im Ausgangsvektor) ermittelt werden und lag bei 2,1%. Diese Religationen können sowohl auf eine unvollständige Restriktion des Vektors pQE30NASTattB als auch auf eine Kontamination des geschnittenen Vektors mit dem ausgeschnittenen Insert (*phyB*) zurückzuführen sein.

Dagegen wurde der Anteil der Religanden der 95 Klone der Pistill cDNA-Bibliothek nur ausgehend von der theoretischen PCR-Produktlänge der Religanden ermittelt und lag bei 6,3%. Dabei ist es nicht auszuschließen, dass sich unter den 6,3 % einige Klone befinden, die keine Religanden sind, aber zufällig ein PCR-Produkt in derselben Größe lieferten. Aufgrund der unterschiedlichen Ermittlung der Religationenanteile sind die beiden Werte nicht unmittelbar miteinander vergleichbar.

Im Falle des Vektors pENTR1a wäre eine Negativselektion der Religanden aufgrund der Suizid-Funktion des *ccdB*-Genes zu erwarten gewesen; der hier erhaltene hohe Anteil von Religanden (6,3%) könnte jedoch auf unvollständige Aktivität des *ccdB*-Genes zurückzuführen sein.

Aufgrund der Testsequenzierung von 48 Pistill-Klonen wurde ersichtlich, dass basierend auf einer BLAST-Analyse allen sequenzierten Klonen einem AGI *gene code* zugeordnet werden konnten. Dabei zeigte sich die hohe Diversität der Klone. Nur ein Klon kam zweimal vor, die restlichen Klone (95,8%) waren je einmal vertreten (*Singeltons*). Der Anteil an Klonen in voller Länge lag bei 23,4%. Darüber hinaus zeigte sich, dass sich 31,3% der Klone im richtigen Leserahmen (d.h. der Leserahmen des RGS His₆-Tags) befanden.

In die Testsequenzierung der Infloreszenz cDNA-Bibliothek wurde eine größere Anzahl von Klonen (576) einbezogen, da diese cDNA-Bibliothek der Ursprung für ein Proteinexpressions-Uniklonset war. 418 Sequenzen mit ausreichender Qualität konnte ein AGI *gene code* zugeordnet werden. Bei 7 Sequenzen handelte es sich um *phyB*, welche den Religandenanteil darstellen und nicht in die weitere Auswertung eingingen. Die Analyse der Diversität der Klone mittels Cluster-Analyse, auf der Grundlage der AGI-*gene codes* ergab, dass 72,3% der Klone einmal vorkamen (*Singeltons*), während die restlichen 27,7% sich in 49 Cluster zusammenfassen ließen. 192 der 411 Klone (46,7%) lagen in voller Länge und 128 Klone (31,1%) im richtigen Leserahmen vor.

Ein Vergleich der Diversität der Infloreszenz cDNA-Bibliothek (72,3% *Singeltons*) mit der Pistill cDNA-Bibliothek (95,8% *Singeltons*) ist nicht sinnvoll, da die Anzahl der analysierten Klone der Pistill cDNA-Bibliothek zu gering dazu war. Dies ist auch eine mögliche Ursache für die Differenz der Klone, die in voller Länge vorliegen (Pistill: 23,4%; Infloreszenz: 46,7%). Weitere Gründe für einen geringeren Anteil an kompletten cDNA-Klonen können auch auf die mRNA Qualität (mögliche Degradation) und die cDNA-Synthese zurückzuführen sein. Bei der Synthese der Infloreszenz cDNA könnte die Reverse Transkriptase effektiver gewesen sein, welches die Häufigkeit erhöhte, ein Gen in voller Länge zu transkribieren.

Bezogen auf den Anteil der Klone, die sich im richtigen Leserahmen befanden, ist für beide in dieser Arbeit erstellten Bibliotheken (Pistill: 31,3% und Infloreszenz: 32,3%) kein direkter Vergleich mit anderen cDNA-Bibliotheken möglich, da für keine der bisher beschriebenen cDNA-Bibliotheken vor Neuordnung der Expressionsklone eine aussagekräftige Leserahmennaussage auf Sequenzbasis beschrieben wurde. Kramer 2003 testeten die Expression von 96 Klonen aus einer Gerste-Embryosack cDNA-Bibliothek und

zeigten, dass 18% dieser Klone eine im SDS-Polyacrylamidgel sichtbare Überexpressionsbande lieferten. Das lässt vermuten, dass in dieser Bibliothek mindestens 18% der Klone im richtigen Leserahmen vorliegen. Jedoch kann davon ausgegangen werden, dass der Anteil höher ist, da bei dieser Methode die schwach translatierten Proteine nicht erfasst werden. Im Gegensatz dazu kann es bei der Ermittlung des richtigen Leserahmens auf Sequenzbasis dazu kommen, dass Klone miterfasst werden, die zwar im richtigen Leserahmen sind aber aufgrund von Stopcodonen in der 5'-untranslatierte Region nicht exprimiert werden.

Die Pistill cDNA-Bibliothek setzte sich mit 660.800 Klonen aus deutlich mehr Klonen zusammen, als die Infloreszenz cDNA-Bibliothek, die aus 40.000 Klonen besteht. Als eine Hauptursache hierfür kann die Verwendung unterschiedlicher kompetenter Zellen angesehen werden. Die Transformationseffizienz der käuflich erworbenen DH5 α -Zellen war um den Faktor 10 höher als die der selbsthergestellten kompetenten SCS1/pSE111-Zellen. Weitere Faktoren können sowohl in den unterschiedlichen Vektoren, z.B. die variierende Vektorgröße (pENTR1a: 2.717 Bp und pQE30NASTattB: 3.575 Bp) als auch in der Ligationseffizienz zu finden sein. Ein weiterer, jedoch nicht so gewichtiger Faktor, um die unterschiedlichen Bibliotheksgrößen zu erklären, ist eine mögliche negative Selektion einiger Proteinexpressionsklone im pQE30NASTattB Vektor durch toxische oder das Wachstum beeinträchtigende Proteine, bei einer nicht vollständig unterdrückten Expression (*leaky expression*).

Da jedoch nur die Infloreszenz cDNA-Bibliothek in einen Expressionsvektor (pQE30NASTattB) kloniert wurde und der zeitliche und finanzielle Aufwand zu hoch gewesen wäre, alle 660.800 Pistill-Klone in einen Expressionsvektor zu überführen, wurden im Folgenden ausschließlich die Klone der Infloreszenz cDNA-Bibliothek verwendet.

Die Pistill cDNA-Bibliothek wurden in einen Entryvektor (pENTR1a) kloniert und vereinigt. Diese Klone stehen somit z. B. als Ressource für ein späteres Überführen in unterschiedliche Expressionsvektoren zur Verfügung.

Die positiven Ergebnisse, die aus der Testsequenzierung der Infloreszenz Klone gewonnen wurden, belegen die gute Qualität dieser cDNA-Bibliothek. Ausgehend von dieser cDNA-Bibliothek wurden im weiteren Verlauf potentielle Expressionsklone zur Erstellung einer Proteinexpressions-Unterbibliothek detektiert.

4.1.2 Identifizierung potentieller Expressionsklone und Herstellung der Proteinexpressions-Unterbibliothek

Die Klonierung der Infloreszenz cDNA-Bibliothek erfolgte in den *E. coli* Expressionsvektor pQE30NASTattB. *E. coli* wurde als Expressionswirt verwendet, da er eine stabile Expression ermöglicht, auf genetischer und biochemischer Ebene gut charakterisiert ist und eine hohe Reproduktivität aufweist. Aufgrund dieser Eigenschaften ist *E. coli* sehr gut für den hohen Durchsatz geeignet. Außerdem ist die im Haus entwickelte Pick- und Spotttroutine mit diesem Expressionssystem bereits etabliert.

Der verwendete Expressionsvektor ermöglicht die rekombinante Proteinexpression über Induktion des T5-Promoters mit IPTG. Darüber hinaus enthält der Vektor ein N-terminales RGS His₆-Tag, womit die Proteine nach der Expression fusioniert sind. Damit können die Fusionsproteine über *Immobilised Metal Affinity Chromatography* (IMAC), sowohl unter nativen, als auch unter denaturierenden Bedingungen aufgereinigt werden. Aufgrund der geringen Größe des Tags geht man davon aus, dass die Struktur und Funktion des fusionierten Proteins nicht beeinträchtigt wird und das Tag demzufolge nach der Aufreinigung nicht entfernt werden muss. Des Weiteren kann das Tag auch noch zur Detektion der rekombinanten Proteine mit Hilfe eines anti-RGS His₆-Antikörpers verwendet werden.

Die Notwendigkeit der Detektion der potentiellen Expressionsklone aus einer cDNA-Bibliothek ist darauf zurückzuführen, dass beim Klonieren der cDNA kein Einfluss darauf genommen werden kann, in welchem der drei Leserahmen die cDNA transferiert wird. Somit ist zu erwarten, dass etwa nur ein Drittel der klonierten Gene im richtigen Leserahmen vorliegen und daher als potentielle Expressionsklone identifiziert werden können. Dies konnte anhand der Testsequenzierung für die im Rahmen dieser Arbeit erstellten cDNA-Bibliotheken bestätigt werden (Pistill: 31,3% und Infloreszenz: 32,3% lagen im richtigen Leserahmen vor).

Die Klone der Infloreszenz cDNA-Bibliothek wurden mittels Robotertechnologie auf PVDF-Membranen gespottet. Anschließend wurde *in situ* die Expression der rekombinanten Fusionsproteine induziert, welche anschließend mit Hilfe des spezifischen anti-RGS His₆-Antikörpers identifiziert werden konnten.

Basierend auf dem Antikörper-*screening* der Hochdichte-Proteinfilter konnten 7.000 Klone als potentielle Expressionsklone identifiziert werden. Das entspricht einem Anteil von 17,5% der 40.000 Klone aus der cDNA-Bibliothek. Auch in den bereits oben erwähnten

Bibliotheken von (Büssow *et al.*, 1998; Gutjahr, 2003; Kramer, 2003) lag der Anteil an potentiellen Expressionsklonen bei ungefähr 20%. Mögliche Ursache dafür, dass die Anzahl an potentiellen Expressionsklonen niedriger ist als die zu erwartende Anzahl (ein Drittel) kann u.a. sein, dass gering exprimierende Proteine gar nicht oder in zu geringer Menge synthetisiert werden, um sie als Expressionsklone zu identifizieren. Berücksichtigt werden muss auch, dass ein Teil der exprimierten Proteine toxisch auf *E. coli* wirken kann bzw. die Proteine durch die Zellen degradiert werden können. Des Weiteren kann bei Klonen, die in voller Länge vorliegen, die 5'-untranslatierte Region zusätzliche Stopcodone enthalten, so dass der Klon das enthaltene Gen nicht exprimiert.

Neben dem *screening* der Hochdichte-Proteinfilter mit dem anti-RGS His₆-Detektionsantikörper, wurde ein weiterer Filtersatz mit einem anti-PhyB-Antikörper inkubiert. Somit konnten die Religanden (Ausgangsvektor mit Insert *phyB*) identifiziert und bei der Erstellung der entsprechenden Unterbibliotheken ausgeschlossen werden. 2.000 der 7.000 potentiellen Expressionsklone wurden als Religanden identifiziert. Dieser hohe Anteil an Religanden kann auf einen nicht vollständig geschnittenen Ausgangsvektor (*SalI* und *NotI*) zurückzuführen sein.

Aufgrund der systematischen Anordnung auf den klonalen Proteinfiltern war es möglich, die 5.000 potentiellen Expressionsklone (abzüglich der Religanden) zu identifizieren und diese in einer Proteinexpressions-Unterbibliothek neu anzuordnen. Um eine definierte Aussage über die Diversität der Bibliothek machen zu können und um die Erstellung eines Uniklonsets zu ermöglichen, wurde die gesamte Proteinexpressions-Unterbibliothek ausgehend vom 5'-Ende ansequenziert und analysiert.

4.1.3 Charakterisierung der Proteinexpressions-Unterbibliothek

Ausgehend von der Sequenzinformation wurde zunächst eine Qualitätskontrolle durchgeführt. 4.398 Sequenzen wurden hinsichtlich ihrer Qualität als ausreichend befunden. Von diesen Sequenzen wurde eine BLAST-Analyse gegen die MIPS-*Arabidopsis*-Protein-Datenbank durchgeführt. 99,6% der Sequenzen konnte ein AGI *gene code* zugeordnet werden. Bei den restlichen Sequenzen, denen kein AGI *gene code* zugeordnet werden konnte, könnte es sich um bisher noch nicht beschriebene Gene handeln.

Basierend auf der BLAST-Analyse konnte gezeigt werden, dass sich 61,8% der sequenzierten Ausgangsklone in Bezug auf den RGS His₆-Tag im richtigen Leserahmen befanden und dass 39,4% der Proteinexpressions-Unterbibliotheks-Klone in voller Länge vorlagen. Dabei zeigte sich, dass sich im Vergleich zur Testsequenzierung (vor dem Rearray) der Anteil vollständiger cDNA-Klone etwas verringert hat, von 46,7%

(vor dem Rearray) auf 39,4% (nach dem Rearray). Eine mögliche Ursache hierfür kann sein, dass bei einigen vollständigen cDNA-Klonen die 5'-untranslatierte Region transkribiert wird, welche Stopcodonen aufweisen kann. Diese Klone liefern kein Fusionsprotein mit dem RGS His₆-Tag und sind damit nicht detektierbar und wurden deshalb nicht in die Proteinexpressions-Unterbibliothek aufgenommen. Der Anteil der Klone im richtigen Leserahmen hat sich dagegen fast verdoppelt, von 32,3% (vor dem Rearray) auf 61,8% (nach dem Rearray). Diese Anreicherung zeigt, dass sowohl die Identifizierung der potentiellen Expressionsklone, als auch deren erneute Anordnung in eine Proteinexpressions-Unterbibliothek außerordentlich wichtig ist. Vergleichbare Studien, in denen auch eine große Anzahl von Expressionsklonen auf Sequenzbasis analysiert wurde, finden sich bei Kramer 2003 und Lüking *et al.* 2003b.

Kramer 2003 sequenzierten 4.100 putative Expressionsklone der Gerste-Embryosack cDNA-Bibliothek. Von 2.416 Sequenzen mit ausreichender Qualität lagen 43% im richtigen Leserahmen vor.

In der Studie von Lüking *et al.* 2003b wurden 2.413 Proteine der humanen fötalen Hirngewebe Bibliothek (Büssow *et al.*, 1998) auf einen Microarray gespottet. Dafür wurden die verwendeten Klone sequenziert und 2.303 der 2.413 Klone konnten einer Sequenz zugeordnet werden. Basierend auf einer BLAST-Analyse (gegen die nichtredundanten Proteindatenbank NRPROT) konnte für 1.603 Klone der Leserahmen ermittelt werden. Von diesen 1.603 Klonen waren 55% im richtigen Leserahmen. Somit liegt der in der vorliegenden Arbeit erzielte Anteil von 61,8% Klonen mit richtigem Leserahmen deutlich über den Werten aus den Studien von Kramer 2003 und Lüking *et al.* 2003b, womit die gute Qualität des in dieser Arbeit durchgeführten Rearrays unterstrichen wird.

In der Studie von Lüking *et al.* 2003b lagen von den 882 Klonen, die sich im richtigen Leserahmen befanden, 382 (43%) in voller Länge vor. Im Vergleich dazu befanden sich in der vorliegenden Arbeit 926 Klone (34,1%) in voller Länge, wenn man von den 2.715 Sequenzen ausgeht, die im richtigen Leserahmen vorlagen. Somit war im Vergleich zu Lüking *et al.* 2003b zwar der Anteil der Klone, die sich im richtigen Leserahmen befanden höher; allerdings war der Anteil der Klone, die in voller Länge vorlagen, geringer. Eine mögliche Begründung für den geringeren Anteil an Klonen, die in voller Länge vorliegen, könnte sein, dass die Reverse Transkriptase nicht so effektiv war bzw. bei der Größenfraktionierung der cDNA auch große cDNAs verloren gingen.

Eine mögliche Ursache dafür, dass sich nach der Neuordnung der Expressionsklone nicht alle Klone im richtigen Leserahmen befinden, liegt darin, dass einige der kleinen Proteine,

die nicht im richtigen Leserahmen vorliegen, exprimiert werden. Wenn die Länge dieser Proteine ausreicht, um nicht von *E. coli* abgebaut zu werden, können diese mit dem anti-RGS His₆-Antikörper detektiert werden und gelangen somit in das potentielle Expressionssubset (Büssow *et al.*, 1998; Lüking *et al.*, 2003b).

Zur weiteren Analyse wurden die 4.398 *Arabidopsis*-Sequenzen mit allen kodierenden Sequenzen von *Arabidopsis* zusammen geclustert. Diese Analyse zeigte, dass 31,7% der Expressionsklone einmal vorkamen, während die restlichen Klone mehrfach repräsentiert waren und sich in 637 Cluster zusammenfassen lassen.

Basierend auf diesen Analysen wurde aus der Proteinexpressions-Unterbibliothek das Uniklonset erstellt.

4.1.4 Erstellung des Uniklonsets

Als erstes Selektionskriterium für die Erstellung des Uniklonsets wurde die Position des ersten Stopcodons einbezogen. Dazu wurden die DNA-Sequenzen translatiert.

Basierend auf der Position des ersten Stopcodons wurden alle Sequenzen eliminiert, die ein Stopcodon bei ≤ 70 Aminosäuren aufwiesen (1.442 Klone, entspricht 32,8% der Expressionsklone). Somit wurde gewährleistet, dass nur solche Klone einbezogen wurden, die Proteine mit einer Größe von mindestens 7,7 kDa exprimieren. Diese Größe wurde gewählt, da Lüking *et al.* 2003b zeigten, dass bei der Identifizierung von Expressionsklonen auf klonalen Proteinfiltren auch Peptide (welche im Leserahmen des RGS His₆-Tags vorlagen) mit einer Größe von 5 – 8 kDa zu detektieren waren.

Die verbliebenen 2.956 Sequenzen wurden in die folgenden zwei Sequenzgruppen unterteilt:

1. Sequenzen, die bei der Cluster- und BLAST-Analyse gleiche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit aufwiesen (2.754 Sequenzen)
2. Sequenzen, die basierend auf der Cluster-Analyse keinem AGI *gene code* zugeordnet werden konnten bzw. Sequenzen, die bei der Cluster- bzw. BLAST-Analyse unterschiedliche AGI *gene codes* als Clusterpartner bzw. besten BLAST-Hit hatten (202 Sequenzen).

Diese Sequenzgruppen sind die Grundlage zur Erstellung des Uniklonsets. Die Selektionskriterien, mit denen die Sequenzen für das Uniklonset aus diesen beiden Sequenzgruppen ausgewählt wurden, sind im folgenden kurz dargestellt, um die Daten im Anschluss daran besser diskutieren zu können.

Die Selektionskriterien waren für die Sequenzen der beiden Sequenzgruppen unterschiedlich. Da bei den Sequenzen der ersten Sequenzgruppe BLAST- und Cluster-AGI *gene code* identisch waren, erfolgte hier die Selektion basierend auf der BLAST- und der

Cluster- BLAST-Analyse. Dagegen war in der zweiten Sequenzgruppe (BLAST- und Cluster-AGI *gene code* nicht identisch) nicht sicher, dass die Auskunft der BLAST-Analyse korrekt war und somit wurden die Parameter dieser Analyse nicht mit einbezogen. Hier gingen als Selektionskriterien das Stopcodon und die Clustergröße ein.

Bei den Sequenzen der ersten Gruppe wurden alle Klone, die nicht im Leserahmen des RGS His₆-Tags (11,6%) vorlagen, sowie alle Klone, welche einen E-Wert (Irrtumswahrscheinlichkeit) über e^{-20} aufwiesen und somit als nicht sicher genug galten, nicht berücksichtigt.

Ausgehend von der Cluster-Analyse wurden alle Klone die je einmal vorkamen (37,5%; *Singeltons*), direkt, ohne weitere Selektion, in das Uniklonset transferiert. Die restlichen Sequenzen ließen sich in 384 Cluster zusammenfassen. Da in dem zu erstellenden Uniklonset jedes Gen möglichst nur einmal vertreten sein sollte, waren hier weitere Selektionskriterien erforderlich. Diese waren der Startpunkt des Subjekts (Treffer-Sequenz), der Startpunkt des Querys (Abfrage-Sequenz) und der E-Wert aus der BLAST-Analyse. Die Kriterien waren in der eben beschriebenen Reihenfolge festgelegt und sollten alle so klein wie möglich sein. Im Falle des Startpunktes des Subjekts heißt das, je kleiner der Wert, desto länger ist der Klon bzw. im Idealfall, wenn der Wert gleich eins ist, liegt das klonierte Gen in voller Länge vor. Beim Startpunkt des Querys bedeutet der kleinste Wert, dass der Klon mit der kürzesten 5'-untranslatierten Region genommen wurde. Sollte aufgrund der Aussage bezüglich des Startpunktes des Subjekts und dem Startpunkt des Querys keine Sequenzauswahl innerhalb eines Clusters möglich sein, wurde als weiteres Kriterium der E-Wert nochmals einbezogen. Es wurde der Klon mit dem kleinsten E-Wert aus dem Cluster genommen. Damit wurde der Grad der Wahrscheinlichkeit, dass es sich bei dem ausgewählten Klon um das bei der BLAST-Analyse gefundene Gen handelt, nochmals erhöht (vgl. 3.4.2.2).

Ausgehend von den 384 Clustern wurden jedoch mehr als 384 Klone in das Uniklonset transferiert. Insgesamt wurden 419 Klone für das Uniklonset ausgewählt. Das ist darauf zurückzuführen, dass im Falle von 21 Clustern den Sequenzen innerhalb eines Clusters mehr als ein BLAST-AGI *gene code* zugeordnet werden konnte und dass aus diesen Clustern mehrere Klone (ein Klon pro BLAST-Hit) nach den oben genannten Kriterien selektiert wurden. Mit dieser Vorgehensweise sollten die wenigen Fälle berücksichtigt werden, bei denen unterschiedliche Gene mittels Cluster-Analyse aufgrund hoher Sequenzähnlichkeiten in eine Cluster eingeordnet wurden. Diese verschiedenen Gene sollten im Uniklonset erfasst werden. Jedoch wird hierbei auch in Kauf genommen, dass einige wenige Gene mehrfach in

dem Uniklonset repräsentiert sind. Diese Mehrfachrepräsentierung kommt dadurch zustande, dass im Laufe der Evolution einige der Gene auf den fünf Chromsomen mehrfach verliegen und somit unterschiedlichen *Accession*-Nummern zugeordnet werden können (Mayer, 2001).

Von den Sequenzen der ersten Sequenzgruppe gingen somit 48,1% der Klone in das Uniklonset ein.

Wie bereits beschrieben, ging die Selektion der Sequenzen der zweiten Sequenzgruppe von weniger Kriterien aus. Die Kriterien waren hier: die Position des Stopcodons und die Clustergröße. Die zweite Sequenzgruppe ließ sich in zwei weitere Untergruppen unterteilen. In der einen Untergruppe konnten den Sequenzen mittels Cluster-Analyse kein AGI *gene code* zugeordnet werden. Für die Sequenzen der anderen Untergruppe konnte zwar auf Grundlage der Cluster-Analyse ein AGI *gene code* ermittelt werden, dieser war jedoch nicht identisch mit dem der BLAST-Analyse.

Von beiden Untergruppen wurden auch wieder alle *Singeltons* direkt in das Uniklonset transferiert (82 der 202 Klone). Die restlichen Klone lagen in Clustern vor. Von den Clustern, die sich aus weniger als sechs Sequenzen zusammensetzten, gingen alle Klone in das Uniklonset ein. Dagegen gingen bei Clustern mit mehr als sechs Klonen nur die Klone ein, die kein Stopcodon besaßen. Zusätzlich zu den 82 *Singeltons* gingen somit noch 92 weitere Klone in das Uniklonset ein. Das bedeutet, dass von Clustern der Sequenzgruppe 2 teilweise mehr als ein Klon in das Uniklonset eingehen konnte. Diese Vorgehensweise sollte ermöglichen, dass sich unter den ausgewählten Klone auch Klone im richtigen Leserahmen befanden, da hier keine Selektion bezüglich der Leserahmen möglich war.

Ausgehend von 2.956 potentiellen Expressionsklonen der Proteinexpressions-Unterbibliothek, deren Sequenzen qualitätsgerecht waren und kein Stopcodon bei ≤ 70 Aminosäuren besaßen, gingen 1.502 Klone (50,7%) in das Uniklonset ein. Zusätzlich zu den 1.502 Uniklonset-Klonen kamen noch 192 Klone ((Kersten *et al.*, 2003) und Klone von Transkriptionsfaktoren) hinzu, die in voller Länge vorlagen. Diese nunmehr insgesamt 1.694 Klone bildeten zusammen das Erweiterte Uniklonset.

4.1.5 Herstellung von Protein-Microarrays des Erweiterten Uniklonsets

Alle Klone des Erweiterten Uniklonsets wurden im 96 *well* Format exprimiert und die Proteine unter denaturierenden Bedingungen aufgereinigt. Ausgehend von drei zufällig ausgewählten 96 *well* Mikrotiterplatten wurde mit Hilfe des Bradford-Tests eine Proteinausbeute von durchschnittlich 8,5 μg bestimmt. Diese lag jedoch unter der erzielten durchschnittlichen Ausbeute von (Kersten *et al.*, 2003) mit 24,5 μg bei 96 *Arabidopsis*-

Klonen. Diese Diskrepanz könnte auf das vergleichsweise kleinere Elutionsvolumen zurückzuführen sein, welches in der vorliegenden Arbeit während der manuellen Reinigung eingesetzt wurde (80 µl Elutionspuffer). Die Proteine von Kersten *et al.* 2003 dagegen wurden unter Verwendung des BioRoboter 8000 (Qiagen) in 350 µl eluiert. Trotz der etwas niedrigeren Gesamtausbeute erfolgte in der vorliegenden Arbeit eine bewusste Entscheidung zugunsten der manuellen Aufreinigung, da hiermit höhere Proteinkonzentrationen erzielt werden können (Uniklonset: 106 µg/ml, Kersten *et al.* 2003: 70 µg/ml). Somit können Microarrays mit einem höheren Proteingehalt generiert werden.

Die 1.694 Proteine des Erweiterten Uniklonsets wurden in einem 11 x 11-Pattern gespottet. Nach dem Spotten der Proteine konnten durchschnittlich 95% der His-getagten Proteine auf den Microarrays mit Hilfe eines anti-RGS His₆-Antikörpers nachgewiesen werden. Dieses Ergebnis belegt, neben den ermittelten Proteinkonzentrationen unter Verwendung des Bradford-Test (Kap. 4.2), dass die Hochdurchsatz-Proteinaufreinigung effizient genug war, um die gewonnen Proteine auf Microarrays nachzuweisen.

4.2 Phosphorylierungsstudien als Funktionelle Analyse

Ausgehend von den aufgereinigten Proteinen des Erweiterten Uniklonsets erfolgten funktionelle Analysen der Proteine unter Verwendung der Microarray-Technologie. Dazu wurden die Proteine des Erweiterten Uniklonsets mittels eines Microarrayers auf Nitrozellulose ähnliche FASTTM-Slides gespottet. Anschließend sollte die Phosphorylierung der gespotteten Proteine mit ausgewählten Kinasen untersucht werden.

4.2.1 Vergleich von Phosphorylierungsstudien

Beim Vergleich der bisher entwickelten Protein-Microarray-basierten Phosphorylierungs-Assays mit dem im Rahmen dieser Arbeit entwickelten Assay, konnten einige Unterschiede festgestellt werden. Diese Unterschiede beziehen sich auf die Oberflächen der Microarrays, die Anzahl der gespotteten Proteine und die Detektion der phosphorylierten Targets.

(Zhu *et al.*, 2000) verwendeten für die Phosphorylierungs-Assays PDMS-beschichtete *nanowell*-Microarrays. Bei diesem Assay wurden 119 Hefe Proteinkinasen und 17 verschiedene Substrate eingesetzt. Die in den *nanowells* kovalent gebundenen Substrate wurden unter Verwendung der Hefe Kinasen und γ -ATP (³³P) phosphoryliert. Sie konnten, unter Verwendung eines Phosphorimager zeigen, dass 32 der 119 Kinasen ein bis zwei der 17 Substrate phosphorylieren (26,9%).

MacBeath und Schreiber 2000 führten ihre Phosphorylierungsstudien auf BSA-N-hydroxysuccinimid (BSA-NHS) beschichteten Microarrays durch. Dabei wurden drei unterschiedliche Kinase-Substrat-Interaktionspaare untersucht, welche jeweils in vier identischen Spots pro Microarray vorlagen. Anschließend wurde jeder dieser Microarrays mit einer der drei Kinasen und γ -ATP (³³P) phosphoryliert. Dabei konnten sie zeigen, dass jeweils nur die spezifischen Substrate der entsprechenden Kinasen phosphoryliert wurden. Zur Detektion der phosphorylierten Substrate verwendeten MacBeath und Schreiber 2000 eine Photoemulsion, in welche die Microarrays gegeben wurden. Anschließend wurden die Microarrays wie ein Foto entwickelt. Diese Microarrays wurden dann mit einem Lichtmikroskop in Teilbereichen aufgenommen, und diese später zusammengefügt. MacBeath und Schreiber 2000 diskutieren diese recht aufwendige und nicht für eine Untersuchung im Hochdurchsatz geeignete Detektion, damit dass zu dem damaligen Zeitpunkt eine ausreichende Auflösung weder mittels Röntgenfilm noch mittels Phosphorimager erzielt werden konnte.

Bei den Phosphorylierungsstudien von Kramer *et al.* 2004 wurden 768 rekombinante Gerste-Proteine auf FASTTM-Slides gespottet und anschließend mit Gerste Casine Kinase2 α

phosphoryliert. Dabei wurden 384 Proteine jeweils vierfach in einem 10 x 10-Pattern auf einen Microarray transferiert. Bei der Detektion der beiden Microarrays mittels Röntgenfilm konnten 21 potentielle Targets identifiziert werden (2,7%).

Im Rahmen der vorliegenden Arbeit wurden 1.694 *Arabidopsis*-Proteine FAST™-Slides gespottet. Die Proteine wurden in einem 11 x 11-Pattern gespottet. Dabei wurden auf jeden Microarray zwei identische Felder gespottet, wobei sich in jedem Feld ein Spot pro Protein befand. Die phosphorylierten Proteine wurden anschließend unter Verwendung eines Phosphorimagers (Bioimager FLA8000) detektiert. Diese Detektion wurde gewählt, da die Detektion mittels Phosphorimager 10 – 100 x empfindlicher als der Röntgenfilm ist, und daher eine wesentlich kürzere Expositionszeit benötigt wird bzw. auch schwächere Signale detektiert werden können. Zudem ist bei dieser Methode das System weniger schnell gesättigt, so dass sowohl sehr starke als auch sehr schwache Signale mit einer Exposition erfasst werden können. Außerdem geht der lineare Signalbereich über 5 Größenordnungen (100.000 : 1). Somit ist der Phosphorimager für Quantifizierungen besser geeignet als die Verwendung von Röntgenfilmen. Hiermit konnte eine der Aufgaben dieser Arbeit dahingehend gelöst werden, die Signalerfassung der phosphorylierten *Arabidopsis*-Proteine mit einem Phosphorimager durchzuführen. Diese Signalerfassung stellt die Grundlage des ebenfalls in dieser Arbeit entwickelten Quantifizierungssystems dar.

Die bisher beschriebene qualitative und leichter visuell zu detektierende Variante, die Duplikate eines Proteins nebeneinander zu spotten (Kramer *et al.*, 2004; MacBeath und Schreiber, 2000), konnte in der vorliegenden Arbeit aufgrund der Optimierung der Reproduzierbarkeit und Quantifizierung, nicht beibehalten werden. In der vorliegenden Arbeit wurden die *Arabidopsis*-Proteine zunächst in einem dichten Pattern (14 x 14 und 15 x 15) auch jeweils vierfach in ein Feld gespottet, dabei zeigte sich bei intensiven radioaktiven Signalen ein Überstrahlungseffekt. Dieser hatte zur Folge, dass Spots mit intensiven Signalen nicht quantifiziert werden konnten. Um diesen Effekt zu vermeiden, wurden die *Arabidopsis*-Proteine in den weiteren Versuchen zwar in Duplikaten, jedoch räumlich voneinander getrennt auf die Microarrays gespottet. Des Weiteren wurde als Spotdichte ein 11 x 11-Pattern festgelegt. Mit diesem verwendeten Spotpattern ist eine parallele Analyse von 1.728 Proteinen (3.456 Spots, ohne *guide dots* und Kontrollen) pro Microarray möglich. Im Gegensatz zu Kramer *et al.* 2004, hier wurden 384 Proteine auf einen Microarray gespottet und phosphoryliert, ist somit eine effektivere Identifikation von potentiellen Phosphorylierungs-Targets möglich.

Alle oben beschriebenen Phosphorylierungs-Detektionen erfolgten unter Verwendung von radioaktiv markiertem ATP. Da diese Methode immer noch die einfachste und sensitivste Detektions-Methode ist, jedoch gesundheitliche Risiken mit sich bringt, werden Detektions-Alternativen gesucht. So verwendeten Houseman *et al.* 2002 zur Detektion von phosphorylierten Substraten auf den Microarrays einen anti-Phosphotyrosin-Antikörper, welcher anschließend mit einem zweiten Fluoreszenz markierten Antikörper inkubiert und detektiert wurde. Auch Lesacherre *et al.* 2002 setzen Fluoreszenz markierte anti-Phosphoserin- und anti-Phosphotyrosin-Antikörper zur Detektion von phosphorylierten Peptiden ein.

Vorteile der Detektion phosphorylierter Substrate mittels Fluoreszenz sind die Einschränkung der gesundheitlichen Risiken und die Möglichkeit, herkömmliche Microarrayscanner zur Auswertung von Phosphorylierungsstudien einzusetzen. Jedoch ist diese Detektions-Methode noch in der Entwicklung und noch nicht mit radioaktiv markierter Detektion zu vergleichen. Des Weiteren muss bedacht werden, dass bei der Detektion mittels Antikörper nur die Targets ermittelt werden können, die an einem bestimmten Epitop phosphoryliert werden. Somit ist diese Antikörpermethode zwar zur Identifizierung potentieller Phosphorylierungs-Targets mit bekannten Epitopen geeignet, aber nicht als *screening tool* zur Erfassung möglichst vieler Phosphorylierungs-Targets.

Um die Spezifität von Kinasen zu ermitteln, können auch Peptid-Microarrays verwendet werden (Falsey *et al.*, 2001; Houseman *et al.*, 2002; Lesacherre *et al.*, 2002; Lizcano *et al.*, 2002).

Lizcano *et al.* 2002 untersuchten die Substratspezifität der humane Serin/Threonin Proteinkinase Nek6 unter Verwendung von Peptid-Microarrays. Bekannt ist, dass diese Kinase *in vivo* die hydrophoben Motive der 3 N-terminalen Leuzin-Reste der S6K (p70 *ribosomal S6 kinase*) und SGK (*serum- and glucocorticoid-induced protein kinase*) phosphoryliert. Auf den Peptid-Arrays mit 710 Peptiden, mit bekannten Phosphorylierungsstellen humaner Proteine, konnten die Phosphorylierung von 7 Peptiden durch die Nek6 gezeigt werden.

Houseman *et al.* 2002 charakterisierten die Spezifität der c-Scr (*nonreceptor tyrosin kinase*), indem sie drei unterschiedliche Peptide auf Microarrays transferierten, wobei ein Peptid ein bekanntes Substrat der c-Scr darstellt. Unter Verwendung eines spezifischen Phosphotyrosin-Antikörpers konnte die spezifische Phosphorylierung des Substrat-Peptids gezeigt werden.

Ein Vorteil der von uns verwendeten Protein-Arrays im Vergleich zu der Peptid-Arrays ist, dass keine Kenntnisse über die Sequenzinformation der Targetsequenzen vorhanden sein müssen. Darüber hinaus sind die Arrays mit rekombinaten Proteinen nicht so teuer wie die Arrays mit Peptiden. Des weiteren sind Protein-Arrays universeller einsetzbar, da diese auch für andere funktionelle Studien, wie zur Untersuchung der Wechselwirkung der Proteine mit anderen Molekülen wie Proteine, DNA und kleine Moleküle, verwendet werden können.

4.2.2 Phosphorylierungsstudien mit den Proteinen des Erweiterten Uniklonsets

Für die in dieser Arbeit durchgeführten Phosphorylierungsstudien wurden die Proteinkinasen PKA, MPK3 und MPK6 gewählt.

Für die Phosphorylierungsstudien wurde zunächst der von Kramer *et al.* 2004 etablierte Microarray-basierte Phosphorylierungs-Assay weiterentwickelt, wobei unter Verwendung der Proteinkinase A ein Quantifizierungssystem etabliert wurde. Anschließend erfolgte die Identifizierung potentieller Phosphorylierungs-Targets der *Arabidopsis* Mitogen-aktivierten Proteinkinasen (MPK3 und MPK6) ausgehend von allen gespotteten Proteinen des Erweiterten Uniklonsets. Zu jedem Phosphorylierungs-Versuch wurde generell ein Kontroll-Microarray mitgeführt. Dieser Phosphorylierungsansatz enthielt γ -ATP (^{33}P), aber keine Kinase. Da in keinem dieser Kontroll-Microarrays Signale detektiert wurden, kann davon ausgegangen werden, dass unter den Versuchsbedingungen keine Bindung des γ -ATP (^{33}P) an die gespotteten Proteine bzw. keine Autophosphorylierung erfolgte.

Für alle drei Kinasen wurde je ein Set zusammengestellt (PKA-Set, MPK3-Set und MPK6-Set), welches aus den identifizierten potentiellen Phosphorylierungs-Targets der entsprechenden Kinasen bestand.

4.2.2.1 Quantifizierung mittels Proteinkinase A

In den Phosphorylierungsstudien unter Verwendung der Proteinkinase A, welche zur Etablierung des Quantifizierungssystems eingesetzt wurden, konnten 35 potentielle Phosphorylierungs-Targets ermittelt werden. Zur Bestätigung wurden diese erneut kultiviert, aufgereinigt, auf eine PVDF-Membran transferiert und dort erneut phosphoryliert. Von 35 potentiellen Phosphorylierungs-Targets konnten 28 bestätigt werden. Dabei können falschpositive Proteine ausgeschlossen werden, da alle im Coomassie-Brilliant-Blau gefärbten Gel identifizierten Protein phosphoryliert wurden, d.h. jedes im Coomassie-Brilliant-Blau gefärbten Gel visuell identifizierte Proteine wurde phosphoryliert. Für die Proteine, die weder im Coomassie-Brilliant-Blau gefärbten Gel noch im phosphorylierten Blot detektiert werden konnte, kann keine Aussage gemacht werden. Mögliche Ursachen

dafür, dass 7 Targets nicht visuell bestätigt werden konnten, kann einerseits sein, dass die radioaktiven Signale der phosphorylierten Proteine auf dem Blot zu gering waren, um detektiert zu werden. Andererseits kann es sein, dass die Proteine für diese erneute Aufreinigung entweder nicht ausreichend exprimiert wurden, die Aufreinigung nicht effizient war oder dass die aufgetragene Proteinmenge nicht ausreichend war. Ein weiterer Hinweis für die Richtigkeit der identifizierten Targets ist, dass alle im Coomassie-Brilliant-Blau gefärbten Gel sichtbaren Banden der Negativkontrollen keine Signale im Radiogramm zeigten. Dabei ist zu beachten, dass die Negativkontrollen aufgrund der für den Phosphorylierungs-Assay gewählten Quantifizierungsmethoden als negativ gewertet wurden.

Zusätzlich wurde die Qualität der identifizierten Targets durch den Vergleich der sich in dem PKA- und in dem Uniklonset befindlichen Sequenzen kontrolliert. Dabei wurden sowohl eine Motivsuche von vier bekannten Konsensussequenzen (Kallus, 2001) durchgeführt als auch eine Kontrolle der Serin- und Threoninegehalte. Bei der Motivsuche konnte für alle vier Konsensussequenzen eine Anreicherung der entsprechenden Motivgehalte im PKA-Set ermittelt werden. Diese Anreicherung zeigt, dass die Identifizierung der potentiellen Targets generell möglich ist. Jedoch zeigte sich auch, dass jede der vier Konsensussequenzen häufiger als 35 x (Anzahl der Targets im PKA-Set) im Uniklonset vorkommen, d.h. nicht jedes Protein, welches eines dieser Motive aufwies, wurde phosphoryliert. Das kann darauf zurückzuführen sein, dass lediglich das Vorliegen des Motivs nicht ausreichend für die jeweilige Phosphorylierung ist. Es ist aber auch nicht auszuschließen, dass das gewählte quantitative Kriterium zur Charakterisierung eines Proteins als Target (nur die Proteine, deren mittlere Signalintensität über der 10fachen Standardabweichung des mittleren Hintergrundsignals aller Unterfelder lagen, wurden als potentielle Phosphorylierung-Targets identifiziert) evtl. zu streng ist. Da es jedoch Ziel dieser Arbeit war, ein *screening tool* zu etablieren, welches zu einer quantifizierten Eingrenzung der Anzahl von potentiellen Phosphorylierungs-Targets dienen sollte, die anschließend unabhängig verifiziert werden sollten, wurde dieses strenge Kriterium beibehalten. Bei dieser Vorgehensweise kann man annehmen, dass auch der Anteil an falschpositiven Targets eingeschränkt wird. In weiteren vergleichenden Studien sollte man versuchen, für dieses Kriterium den exakten Grenzwert zu ermitteln, wodurch es zu einer Maximierung der Selektion aller sich im Set befindlichen potentiellen Targets kommen könnte.

Der Vergleich des Serin- und Threoningehaltes im PKA- und Uniklonset zeigte eine Anreicherung des Seringehaltes, was als Hinweis darauf angesehen werden könnte, dass die Proteinkinase A zwar zu Serin/Threonin-Proteinkinase gehört, jedoch die Serin-Reste vielleicht bevorzugt phosphoryliert.

4.2.2.2 Phosphorylierungsstudien unter Verwendung der *Arabidopsis* Kinasen MPK3 und MPK6

Wie bereits in der Einleitung erläutert, gehören die MAP-Kinasen zu den Serin/Threonin-Proteinkinasen. Beide MAP-Kinasen (MPK) MPK3 und MPK6 werden in eine der vier Gruppen (Gruppe A), in welche die MPKs unterteilt werden, eingeordnet (vgl. 1.5.1; (MAPKgroup, 2002). Die MPKs sind Bestandteile von Signaltransduktionskaskaden, den MAPK-Kaskaden. Diese Kaskaden bestehen aus MAPKK-Kinasen (MAPKKK), MAPK-Kinasen (MAPKK) und MAP-Kinasen (MPK) und sind an der Weiterleitung der durch Rezeptoren oder Sensoren erkannten externen Stimuli von der Zelloberfläche zum Zellkern beteiligt. Im Zellkern kann es dann zur Phosphorylierung von Transkriptionsfaktoren kommen. Die Weiterleitung des Stimulus erfolgte über sogenannte *second messenger*, durch die es u. a. zu Änderungen in der Ionenpermeabilität der Plasmamembran (Nünberger und Scheel, 2001) oder zu Akkumulationen von reaktiven Sauerstoffspezies (ROS, z.B. Superoxidanionen, Hydroxylradikale oder Wasserstoffperoxid) kommen kann (Neill *et al.*, 2002). Die ROS können direkt toxisch auf das Phytopathogen wirken oder in die Verstärkung der pflanzlichen Zellwand involviert sein (Lamb und Dixon, 1997).

Durch diese Weiterleitung kommt es zur Aktivierung von Abwehrmechanismen (Zwerger und Hirt, 2001). Solche Kaskaden ermöglichen es den Pflanzen, auf verschiedene biotische und abiotische Stressfaktoren zu reagieren. Eine Beteiligung von MAP-Kinasen an der pflanzlichen Signaltransduktion konnte für eine große Anzahl von Stimuli beschrieben werden. Beispielsweise für mechanischen und osmotischen Stress (Bogre *et al.*, 1996; Ichimura *et al.*, 2000; Mizoguchi *et al.*, 1996; Tena und Renaudin, 1998), Trockenheit und Kälte (Ichimura *et al.*, 2000; Jonak *et al.*, 1996; Mizoguchi *et al.*, 1996), sowie Elicitor- als auch Pathogenkontakt (Asai *et al.*, 2002; Nühse *et al.*, 2000; Romeis *et al.*, 1999; Zhang *et al.*, 1998). Des Weiteren sind sie auch bei der Weitergabe der Pflanzenhormone, wie Ethylen, Abszissinsäure und Auxin beteiligt (Jonak *et al.*, 2002; Lu *et al.*, 2002; Tena *et al.*, 2001).

Bei den in der vorliegenden Arbeit durchgeführten Phosphorylierungsstudien konnten für das MPK3-Set 48 potentielle Phosphorylierungs-Targets und für das MPK6-Set 39 potentielle Phosphorylierungs-Targets identifiziert werden. Die geringere Anzahl an

identifizierten Targets bei dem MPK6-Set kann einerseits darauf zurückzuführen sein, dass die Kinase weniger Targets phosphoryliert oder für die MPK6 weniger potentielle Phosphorylierungs-Targets im Erweiterten Unikonset vorlagen, als für die MPK3. Andererseits bleibt zu bedenken, dass die rekombinate MPK6 mit einer geringeren Aktivität vorlag (persönliche Mitteilung unserer Kooperationspartner Dr. Justin Lee und Prof. Dr. Dierk Scheel), daher zwar eine höhere Konzentration eingesetzt wurde, diese aber vielleicht nicht ausreichend war, um eine zur MPK3 vergleichbare Aktivität zu erreichen. Es kann aber auch sein, dass die in der vorliegenden Arbeit verwendeten Assay-Bedingungen (wie z. B. Kofaktoren) für die MPK6 weniger optimal waren, als für die MPK3, da die Bedingungen ausgehend vom Myelin-assoziiertes basisches Protein (MBP), einem artifiziellen Substrat der Kinasen, optimiert wurden.

Zur Bestätigung der identifizierten Targets in einem unabhängigen *in vitro* Testsystem wurden die 48 potentiellen Phosphorylierungs-Targets des MPK3-Sets erneut aufgereinigt, auf PVDF-Membranen transferiert und unter Verwendung der MPK3 phosphoryliert. Dies war leider nur für die MPK3 möglich, da zu diesem Zeitpunkt keine ausreichend aktive MPK6 mehr zugänglich war. Von 48 potentiellen Phosphorylierungs-Targets konnten 45 bestätigt werden. Mögliche Ursachen für die drei nicht identifizierten Targets sind wie bei dem Phosphorylierungs-Blot der Proteinkinase A zu diskutieren.

Bei den Literaturrecherchen zu den identifizierten potentiellen Phosphorylierungs-Targets zeigte sich, dass für die *Arabidopsis* MPK noch nicht bekannt ist, welche Targetproteine direkt phosphoryliert werden. Bisher gibt es nur Untersuchungen der durch Stress ausgelösten MAP-Kinasen-Signaltransduktionskaskade hinsichtlich bestimmter Gentranskripte, aus denen man erste Hinweis auf mögliche Targets gewinnen kann (z.B. Kimura *et al.*, 2003; Kreps *et al.*, 2002; Seki *et al.*, 2002).

Beispielsweise konnten sowohl die Aktivierung von MPK3 durch Trockenheit, Kälte, Salzstress und mechanische Beanspruchung (Mizoguchi *et al.*, 1996) als auch eine posttranslationale Aktivierung von MPK3 durch H₂O₂ beobachtet werden.

Des Weiteren konnte für die Transkripte der MEKK1, MPK3 und MPK1 eine Induktion durch osmotischen Stress und mechanische Beanspruchung gezeigt werden. Zusätzlich sind die Transkripte der MEKK1 und MPK3 durch Kälte induzierbar [Mizoguchi *et al.*, 1996]. Darüber hinaus konnten die Transkripte der MPK3 und MPK4 durch Bestrahlung mit UV-Licht induziert werden (Ulm *et al.*, 2002). Teige *et al.* 2004 konnten sowohl eine Aktivierung der MKK2 durch Kälte, Salzstress oder durch die stressinduzierte MEKK1 als auch die davon ausgehende Induktion der MAPK4 und MAPK6 zeigen.

Zudem konnte eine Aktivierung der MAP-Kinasen nicht nur bei abiotischen sondern auch bei den biotischen Stressfaktoren beobachtet werden. Asai *et al.* 2002 beschrieben bei *Arabidopsis* die vollständige MAP-Kinasen-Signaltransduktionskaskade, welche durch den bakteriellen Elicitor Flg22 ausgelöst wird. Diese MAPK-Kaskade besteht aus den Komponenten MEKK1, MKK4/5 und MPK3/6, welche die transkriptionelle Aktivierung der WRKY-Transkriptionsfaktoren WRY22/29 induzieren. Die Übertragung des durch den Elicitor Flg22 ausgelösten Signals erfolgt über FLS2 (*flagellin sensitive 2*), eine LRR (*leucin-rich repeat*) Rezeptor ähnliche Kinase (Gómez-Gómez und Boller, 2000).

Neben dem eben beschriebenen bakteriellen Elicitor bei *Arabidopsis* wurden für die MPK6 Orthologen in *Medicago sativa*, SIMK (*stress induced MAPK*), und in Tabak, SIPK (*salicylate induced MAPK*), weitere Elicitoren beschrieben, welche zur Aktivierung dieser Kinasen führen. Die SIMK-Aktivierung erfolgt durch unterschiedliche pilzliche Elicitoren (Cardinale *et al.*, 2000). Für SIPK konnte eine Aktivierung durch allgemeine Elicitoren (Droillard *et al.*, 2000; Lee *et al.*, 2001) und durch TMV (Tabak-Mosaik-Virus)-Infektion (Zhang *et al.*, 1998) beschrieben werden. Auch bei Petersilie konnte eine Aktivierung der PcMPK6 durch die Elicitoren PEP-13 und HrpZ beobachtet werden (Kroj *et al.*, 2003).

Wan *et al.* 2004 vermutet eine Überlappung von der von Asai *et al.* 2002 beschriebenen Signaltransduktionskaskaden mit der Signaltransduktion von Chitin Elicitoren, welche bei der Übertragung des Chitin-Stimulus von den Zellwänden pathogener Pilze und Insekten ausgelöst wird.

Einige in dieser Arbeit identifizierten potentiellen Phosphorylierungs-Targets stehen im Zusammenhang mit bereits beschriebenen Komponenten der MAP-Kaskaden. Aus diesem Grund ist ihre Phosphorylierung durch die MPKs nicht unwahrscheinlich.

Das als potentielles Phosphorylierungs-Target der MPK3 identifizierte Protein: *leucine-rich repeat family protein contains leucine rich repeat (LRR) domains* (At4g03260; exprimiert von Klon: 312_J15 des Uniklonsets) könnte, aufgrund seiner LRR-Domäne, an einer von der Elicitoren ausgelösten-Signaltransduktionskaskade beteiligt sein (Asai *et al.*, 2002; Gómez-Gómez und Boller, 2000; Wan *et al.*, 2004). Als ein weiteres interessantes Target konnte das durch Klon 313_G13 exprimierte Protein (At4g11280) bei Phosphorylierung mit der MPK6 detektiert werden. Hierbei handelt es sich um die ACC Synthase 6 (*1-aminocyclopropane-1-carboxylate synthase 6*). Ouaked *et al.* 2003 konnten zeigen, dass eine MAPK-Kaskade an der Ethylen-Signaltransduktion beteiligt ist, in dem bestimmte MPKs in *Arabidopsis* und *Medicago* durch die Vorläufersubstanz des Ethylens ACC aktiviert werden.

Dieses sind die aktivierten MPKs SIMK (*stress induced MAPK*) und MMK (*Medicago MAPK2*) in *Medicago* und in *Arabidopsis* die MPK6 sowie eine weitere MAP-Kinase.

Des weiteren wurden die Histonen H2A (im MPK3-Set) und H3 (in beiden Sets) als potentielle Phosphorylierungs-Targets identifiziert. Für Histone ist bekannt, dass diese als Reaktion auf extrazelluläre Signale phosphoryliert werden. Yamagata *et al.* 2001 zeigten eine Phosphorylierung von Histonen durch die *MAPK kinase-like protein kinase* in Zellkultur von Sojabohnen. Ob die Histone tatsächlich in der Signaltransduktion der MPKs involviert sind, bleibt noch zu zeigen.

Auch ein Protein des MPK3-Sets (At5g62070; *calmodulin-binding family protein*; exprimiert durch Klon 313_F05) kann ein Bestandteil der MAPK-Kaskaden sein. Perruc *et al.* 2004 zeigten, dass bei *Arabidopsis* ein Calmodulin bindendes Protein (At2g41010) als Negativregulator bei der Stresstoleranz gegenüber Salz- und osmotischem Stress beteiligt ist. Ein eventueller Zusammenhang zwischen einer möglichen Phosphorylierung von Calmodulin-bindenden Proteinen durch MPKs und deren Rolle in der Regulation der Stresstoleranz muss noch bestätigt werden. Calmodulin könnte über die Bindung an das Calmodulin-bindende Protein oder aber auch über die Aktivierung der Ca²⁺-calmodulinabhängigen Proteinkinase (CaM-Kinase) einen Einfluss auf die Stresstoleranz haben. CaM-Kinasen können z.B. Serin- und Threoninreste in ihren Zielproteinen phosphorylieren und damit eine Enzymaktivierung bewirken. Daran könnten ebenfalls auch die MPKs beteiligt sein.

Die Lokalisierung der aktivierten MPKs im Zellkern läßt Transkriptionsfaktoren als Targets vermuten. Diese Vermutung wird auch durch die in dieser Arbeit als Phosphorylierungs-Targets gefundenen Transkriptionsfaktoren bestätigt. So konnten folgende Transkriptionsfaktoren identifiziert werden: Der Klon des MPK3-Sets 315_G09 (At5g66940; *Dof-type zinc finger domain-containing protein*), sowie die Klone des MPK6-Sets 315_P07 (At1g06170; *basic helix-loop-helix (bHLH) family protein*), 314_N21 (At3g60390; *homeobox-leucine zipper protein HAT3*), 313_N06 (At5g54630; *zinc finger protein-related*) und 311_F11 (At5g58620; *zinc finger (CCCH-type)*), wobei letzterer zusätzlich auch im MPK3-Set vorkommt und mit seinem Zinkfingermotiv eine DNA-bindende Domäne aufweist.

Wie sich anhand der Histone und der Transkriptionsfaktoren zeigte, wurden auch gemeinsame potentielle Phosphorylierungs-Targets für beide MPKs identifiziert. Insgesamt waren das 26 potentielle Phosphorylierungs-Targets. Dieses wurde jedoch auch vermutet, da beide MPKs zu einer Untergruppe einer Proteinklasse gehören.

Als weiteres gemeinsames potentiell Phosphorylierungs-Target wurde z. B. die Casein Kinase (CK; 314_G10; At5g44100) ermittelt. Eine direkte Beteiligung der Casein Kinase in der Signaltransduktion der MAPK-Kaskaden ist für Pflanzen noch nicht beschrieben. In HeLa-Zellen wurde jedoch eine direkte Wechselwirkung der p38 MAP-Kinase mit der CK2 und deren stressinduzierten Aktivierung durch diese MAP-Kinase gezeigt (Sayed *et al.*, 2000).

Ein weiterer Klon, der in beiden Sets zu finden war, ist der Klon: 313_E14 (At5g48990; *kelch repeat-containing F-box family protein*). F-Box Proteine sind häufig an der Degradation von cytosolischen Proteinen beteiligt (Andrade *et al.*, 2001). F-Box Proteine erkennen und binden die abzubauenen Proteine, welche häufig phosphoryliert sind (Willems *et al.*, 1999). Das am F-Box-Protein gebundene Protein wird anschließend mehrfach ubiquitiniert und zum Proteasomen transportiert. Diese Polyubiquitinierung dient als Degradationssignal für das Proteasomen. Ob die Phosphorylierung der F-Box-Proteine *in vivo* eine Rolle spielt, bleibt noch zu klären.

Als eine weitere Kontrolle hinsichtlich der Spezifität der identifizierten Targets der jeweiligen Kinasen wurden neben den gemeinsamen Phosphorylierungs-Targets der beiden MPKs (26 insgesamt) auch die gemeinsamen Targets der PKA mit denen der MPK3 bzw. der MPK6 bestimmt. Für die MPK3 wurden drei und für die MPK6 vier gemeinsame Targets mit der PKA ermittelt. Wie zu erwarten war, lieferte der Vergleich der unterschiedlichen Serin/Threonin-Proteinkinasen einen niedrigeren Anteil an gemeinsamen Targets, als der Vergleich der beiden MAP-Kinasen, die zu einer Kinasegruppe gehören.

Da alle drei Kinasen zu den Serin/Threonin-Proteinkinasen gehören, liegt die Vermutung nahe, dass bei den ermittelten gemeinsamen potentiellen Phosphorylierungs-Targets beider Gruppen ausschließlich Serin- und Threonin-Reste phosphoryliert wurden, was aber in der vorliegenden Arbeit nicht weiter untersucht wurde. Die genaue Ermittlung der phosphorylierten Aminosäure ist prinzipiell entweder mit Massenspektroskopie (Glinski *et al.*, 2003) oder mit phospho-spezifischen Antikörpern (Houseman *et al.*, 2002; Lesaiquerre *et al.*, 2002) möglich.

Die Ergebnisse der Phosphorylierungsstudien zeigen, dass sowohl der etablierte Kinase-Assay als auch das dazu gehörende Quantifizierungssystem reproduzierbar und spezifisch sind.

Des Weiteren sind alle durchgeführten Kontrollen ein Hinweis darauf, dass alle Schritte, die zur Entstehung des Uniklonsets (Detektion der potentiellen Expressionsklone mittels Hochdichte Proteinfiltern, die Sequenzanalyse und davon ausgehend die Erstellung des

Uniklonsets) erfolgreich waren. Außerdem belegen sowohl die Kontrollen als auch die Ergebnisse dieser Phosphorylierungsstudien, dass die im Rahmen dieser Arbeit verwendete Aufreinigungstechnik für den Hochdurchsatz geeignet war.

An dieser Stelle soll nochmals betont werden, dass es sich bei dem hier weiterentwickelten Mikroarray-basierten Phosphorylierungssystem nur um ein *screening tool* handelt, welches zu einer *in vitro* Identifizierung der potentiellen Kinasetargets dient. Die mit diesem Testsystem identifizierten potentiellen Phosphorylierung-Targets müssen jedoch in anschließenden Folgeversuchen *in vivo* weiter verifiziert werden. Das ist aus folgenden Gründen erforderlich: In der vorliegenden Arbeit wurden die Targets unter denaturierenden Bedingungen aufgereinigt. Es kann also nicht ausgeschlossen werden, dass unter nativen Bedingungen die Phosphorylierungsstellen für die Kinasen nicht zugänglich sind. Des Weiteren besteht auch die Möglichkeit, dass *in vivo* die hier identifizierten Targets nicht mit den Kinasen interagieren können, da sie in unterschiedlichen Zellkompartimenten lokalisiert sein könnten. Dennoch zeigt die Identifikation von bekannten Phosphorylierung-Targets, dass diese Methode als eine Möglichkeit zur Ermittlung von unbekanntem potentiellen Phosphorylierungs-Targets angesehen werden kann.

In vivo Methoden, die angewandt werden können, um die identifizierten potentiellen Phosphorylierungs-Targets zu verifizieren, sind z.B. Methoden zur Untersuchung von Protein-Protein Wechselwirkungen, z.B. in Hefe (Yeast-Two-Hybrid) oder in Pflanzen direkt. Hierbei kommt es zur Verwendung von mikrospektroskopischen Ansätzen wie Fluoreszenz-Resonanz-Energie-Transfer (FRET) und Fluoreszenz-Korrelations-Spektroskopie (FCS) (Hink *et al.*, 2002). Mittels FRET-Methoden kann eine abstandsabhängige Wechselwirkung zwischen zwei angeregten Farbstoffmolekülen detektiert werden, bei der die Anregungsenergie von einem Donormolekül auf eine Akzeptormolekül übertragen wird, ohne daß Photonen abgestrahlt werden. Bei FCS wird meist nur eine Molekül farblich markiert. Bei dieser Methode kann man die Bewegung einzelner Moleküle beobachten, so bewegen sich z. B. gebundene Moleküle vergleichsweise langsamer.

Eine weitere *in vivo* Methode zur Bestätigung der potentiellen Phosphorylierungs-Targets stellt die Klonierung der potentiellen Targets in Epitop-getagte Pflanzenexpressionsvektoren dar. Dazu könnte das GATEWAY-System benutzt werden, um eine effiziente und schnelle Subklonierung zu ermöglichen. Anschließend werden pflanzliche Protoplasten mit diesen Plasmiden transformiert, und die Phosphorylierung dieser Targets kann durch Zugabe von radioaktiv markiertem Phosphat zum Medium untersucht werden. Bedingungen, die zu einer

Aktivierung der MAP-Kinase führen (Elicitor, abiotische Stressbehandlung oder Kotransfektion in eine aktive *upstream* MAP Kinase Kinase) können genutzt werden, um die Beteiligung von MAP-Kinase bei der Phosphorylierung der potentiellen Targets zu untersuchen.

Bei der Validierung von Phosphorylierungs-Targets mittels Protein-Protein-Interaktionsstudien bleibt aber zu beachten, dass die Wechselwirkung einer Kinase mit einem Protein nicht zur Phosphorylierung des Proteins führen muß.

Die in der Arbeit hergestellten Klone stellen die bisher umfangreichste Sammlung von rekombinaten *Arabidopsis thaliana* Expressionsklonen dar und stehen für weitere Studien zur Verfügung.

Diese Arbeit, ist die erste systematische Studie zur Identifizierung von Phosphorylierungs-Targets der MAP-Kinasen im Hochdurchsatz. Zur Identifizierung wurde ein Quantifizierungssystem entwickelt, bei dem nur phosphorylierten Signale, die sich deutlich vom Hintergrund unterschieden, verwendet wurden, um die dazugehörigen Proteine als Phosphorylierungs-Targets zu identifizieren. Eine Lockerung des in der vorliegenden Arbeit sehr streng verwendeten *cutoff* Wertes für die Targetselektion, könnte die Anzahl der identifizierten Phosphorylierungs-Targets in Zukunft evtl. erhöhen.

Desweiteren konnten in der vorliegenden Arbeit bisher noch nicht beschriebene Targets der MAP-Kinasen identifiziert werden. Diese *in vitro* identifizierten Phosphorylierungs-Targets liefern eine interessante Grundlage für weitere *in vivo* Untersuchungen. Der Phosphorylierungs-Assay ist für verschiedene Kinasen anwendbar und kann in Zukunft zur Identifizierung von Phosphorylierungs-Targets weiterer *Arabidopsis thaliana* Kinasen verwendet werden.

5 Literatur

- Adams, M.D., Celniker, S.E., Holt, R.A., et al. (2000) The genome sequence of *Drosophila melanogaster*. *Science*, **287**, 2185-95.
- Afanassiev, V., Hanemann, V. und Wolfl, S. (2000) Preparation of DNA and protein micro arrays on glass slides coated with an agarose film. *Nucleic Acids Res*, **28**, E66.
- Altschul, S.F., Gish, W., Miller, W., et al. (1990) Basic local alignment search tool. *J Mol Biol*, **215**, 403-10.
- Altschul, S.F., Madden, T.L., Schaffer, A.A., et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-402.
- Andersen, J.S., Svensson, B. und Roepstorff, P. (1996) Electrospray ionization and matrix assisted laser desorption/ionization mass spectrometry: powerful analytical tools in recombinant protein chemistry. *Nat Biotechnol*, **14**, 449-57.
- Andrade, M.A., Gonzalez-Guzman, M., Serrano, R., et al. (2001) A combination of the F-box motif and kelch repeats defines a large Arabidopsis family of F-box proteins. *Plant Mol Biol*, **46**, 603-14.
- Angenendt, P., Glokler, J., Murphy, D., et al. (2002) Toward optimized antibody microarrays: a comparison of current microarray support materials. *Analytical Biochemistry*, **309**, 253-60.
- Angenendt, P., Glokler, J., Sobek, J., et al. (2003b) Next generation of protein microarray support materials: evaluation for protein and antibody microarray applications. *J Chromatogr A*, **1009**, 97-104.
- Arabidopsis-Genome-Initiative. (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796-815.
- Arenkov, P., Kukhtin, A., Gemmell, A., et al. (2000) Protein microchips: use for immunoassay and enzymatic reactions. *Analytical Biochemistry*, **278**, 123-31.
- Asai, T., Tena, G., Plotnikova, J., et al. (2002) MAP kinase signalling cascade in *Arabidopsis* innate immunity. *Nature*, **415**, 977-83.
- Bernard, P. und Couturier, M. (1992) Cell killing by the F plasmid CcdB protein involves poisoning of DNA-topoisomerase II complexes. *J Mol Biol*, **226**, 735-45.
- Bogre, L., Ligterink, W., Heberle-Bors, E., et al. (1996) Mechanosensors in plants. *Nature*, **383**, 489-90.

- Bonfield, J.K. und Staden, R. (1996) Experiment files and their application during large-scale sequencing projects. *DNA Seq*, **6**, 109-17.
- Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem*, **72**, 248-54.
- Braun, P., Hu, Y., Shen, B., et al. (2002) Proteome-scale purification of human proteins from bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, **99**, 2654-9.
- Burke, J., Davison, D. und Hide, W. (1999) d2_cluster: a validated method for clustering EST and full-length cDNA sequences. *Genome Res*, **9**, 1135-42.
- Burton, G.R. und McGehee, R.E., Jr. (2004) Identification of candidate genes involved in the regulation of adipocyte differentiation using microarray-based gene expression profiling. *Nutrition*, **20**, 109-14.
- Büssow, K., Cahill, D., Nietfeld, W., et al. (1998) A method for global protein expression and antibody screening on high - density filters of an arrayed cDNA library. *Nucleic Acids Research*, **26**, 5007-5008.
- Büssow, K., Nordhoff, E., Lübbert, C., et al. (2000) A human cDNA library for high-throughput protein expression screening. *Genomics*, **65**, 1-8.
- Byzova, M.V., Franken, J., Aarts, M.G., et al. (1999) Arabidopsis STERILE APETALA, a multifunctional gene regulating inflorescence, flower, and ovule development. *Genes Dev*, **13**, 1002-14.
- C.elegans-Sequencing-Consortium. (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium. *Science*, **282**, 2012-8.
- Cahill, D.J. (2001) Protein and antibody arrays and their medical applications. *Journal of Immunological Methods*, **250**, 81-91.
- Cardinale, F., Jonak, C., Ligterink, W., et al. (2000) Differential activation of four specific MAPK pathways by distinct elicitors. *J Biol Chem*, **275**, 36734-40.
- Champion, A., Kreis, M., Mockaitis, K., et al. (2004) Arabidopsis kinome: after the casting. *Funct Integr Genomics*, **4**, 163-87. Epub 2004 Jan 22.
- Chien, C.T., Bartel, P.L., Sternglanz, R., et al. (1991) The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A*, **88**, 9578-82.

- Clark, M.D., Panopoulou, G.D., Cahill, D.J., et al. (1999) Construction and analysis of arrayed cDNA libraries [Review]. *Methods in Enzymology*, **303**, 205-233.
- Cordonnier, M.M., Greppin, H. und Pratt, L.H. (1986) Identification of a highly conserved domain on phytochrome from angiosperms to algae. *Plant Physiology*, **80**, 982-987.
- DeRisi, J.L., Iyer, V.R. und Brown, P.O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science*, **278**, 680-686.
- Diezel, W., Kopperschlager, G. und Hofmann, E. (1972) An improved procedure for protein staining in polyacrylamide gels with a new type of Coomassie Brilliant Blue. *Anal Biochem*, **48**, 617-20.
- Dove, S.L. und Hochschild, A. (2004) A bacterial two-hybrid system based on transcription activation. *Methods Mol Biol*, **261**, 231-46.
- Droillard, M.J., Thibivilliers, S., Cazale, A.C., et al. (2000) Protein kinases induced by osmotic stresses and elicitor molecules in tobacco cell suspensions: two crossroad MAP kinases and one osmoregulation-specific protein kinase. *FEBS Lett*, **474**, 217-22.
- Ewing, B. und Green, P. (1998) Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res*, **8**, 186-94.
- Ewing, B., Hillier, L., Wendl, M.C., et al. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res*, **8**, 175-85.
- Falsey, J.R., Renil, M., Park, S., et al. (2001) Peptide and small molecule microarray for high throughput cell adhesion and functional assays. *Bioconjug Chem*, **12**, 346-53.
- Fields, S., Kohara, Y. und Lockhart, D.J. (1999) Functional genomics. *Proc Natl Acad Sci U S A*, **96**, 8825-6.
- Fields, S. und Sternglanz, R. (1994) The two-hybrid system: an assay for protein-protein interactions. *Trends Genet*, **10**, 286-92.
- Ge, H. (2000) UPA, a universal protein array system for quantitative detection of protein-protein, protein-DNA, protein-RNA and protein-ligand interactions. *Nucleic Acids Research*, **28**, e3.
- Glinski, M., Romeis, T., Witte, C.P., et al. (2003) Stable isotope labeling of phosphopeptides for multiparallel kinase target analysis and identification of phosphorylation sites. *Rapid Commun Mass Spectrom*, **17**, 1579-84.
- Glöckler, J. und Angenendt, P. (2003) Protein and antibody microarray technology. *J Chromatogr B Analyt Technol Biomed Life Sci*, **797**, 229-40.

- Goff, S.A., Ricke, D., Lan, T.H., et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). *Science*, **296**, 92-100.
- Gómez-Gómez, L. und Boller, T. (2000) FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Mol Cell*, **5**, 1003-11.
- Guschin, D., Yershov, G., Zaslavsky, A., et al. (1997) Manual manufacturing of oligonucleotide, DNA, and protein microchips. *Anal Biochem*, **250**, 203-11.
- Gutjahr, C. (2003) Generierung und Applikation von Hochdichte-Proteinarrays für die Identifizierung von Autoantigenen bei systemischen Erkrankungen. *Dissertation, Freien Universität Berlin*.
- Haab, B.B., Dunham, M.J. und Brown, P.O. (2001) Protein microarrays for highly parallel detection and quantitation of specific proteins and antibodies in complex solutions. *Genome Biology*, **2**, RESEARCH0004.
- Harashima, S. (1994) Heterologous protein production by yeast host-vector systems. *Bioprocess Technol*, **19**, 137-58.
- Hardie, D.G. (1999) PLANT PROTEIN SERINE/THREONINE KINASES: Classification and Functions. *Annu Rev Plant Physiol Plant Mol Biol*, **50**, 97-131.
- Heß, D. (1991) Pflanzenphysiologie. *UTB für Wissenschaft*, **9. Auflage**.
- Heukeshoven, J. und Dernick, R. (1985) Simplified method for silverstaining of proteins in polyacrylamide gels and the mechanism of silver staining. *Electrophoresis*, **6**, 103-12.
- Heyman, J.A., Cornthwaite, J., Foncerrada, L., et al. (1999) Genome-scale cloning and expression of individual open reading frames using topoisomerase I-mediated ligation. *Genome Research*, **9**, 383-392.
- Heyman, J.A., Cornthwaite, J., Foncerrada, L., et al. (1999) Genome-scale cloning and expression of individual open reading frames using topoisomerase I-mediated ligation. *Genome Research*, **9**, 383-392.
- Holz, C., Lueking, A., Bovekamp, L., Gutjahr, C., Bolotina, N., Lehrach, H., Cahill, D. J., et al. (2001) A human cDNA expression library in yeast enriched for open reading frames. *Genome Research*, **11**; 1730-5.
- Hillenkamp, F. und Karas, M. (1990) Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods Enzymol*, **193**, 280-95.
- Hink, M.A., Bisselin, T. und Visser, A.J. (2002) Imaging protein-protein interactions in living cells. *Plant Mol Biol*, **50**, 871-83.

- Houseman, B.T., Huh, J.H., Kron, S.J., et al. (2002) Peptide chips for the quantitative evaluation of protein kinase activity. *Nat Biotechnol*, **20**, 270-4.
- Ichimura, K., Mizoguchi, T., Yoshida, R., et al. (2000) Various abiotic stresses rapidly activate Arabidopsis MAP kinases ATMPK4 and ATMPK6. *Plant J*, **24**, 655-65.
- Jonak, C., Kiegerl, S., Ligterink, W., et al. (1996) Stress signaling in plants: a mitogen-activated protein kinase pathway is activated by cold and drought. *Proc Natl Acad Sci U S A*, **93**, 11274-9.
- Jonak, C., Okresz, L., Bogre, L., et al. (2002) Complexity, cross talk and integration of plant MAP kinase signalling. *Curr Opin Plant Biol*, **5**, 415-24.
- Joos, T.O., Schrenk, M., Hopfl, P., et al. (2000) A microarray enzyme-linked immunosorbent assay for autoimmune diagnostics. *Electrophoresis*, **21**, 2641-50.
- Joung, J.K., Ramm, E.I. und Pabo, C.O. (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A*, **97**, 7382-7.
- Kallus, B. (2001) cAMP-abhängige Phosphorylierung der Cytochrom c Oxidase in Rattenhepatozyten. *Dissertation, Heinrich-Heine-Universität Düsseldorf*.
- Kersten, B., Burkle, L., Kuhn, E.J., et al. (2002) Large-scale plant proteomics. *Plant Mol Biol*, **48**, 133-41.
- Kersten, B., Feilner, T., Angenendt, P., et al. (2004a) Proteomic Approaches in Plant Biology. *Current Proteomics*, **1**, 131-144.
- Kersten, B., Feilner, T., Kramer, A., et al. (2003) Generation of Arabidopsis protein chips for antibody and serum screening. *Plant Molecular Biology*, **52**, 999-1010.
- Kersten, B., Possling, A., Blaesing, F., et al. (2004b) Protein microarray technology and ultraviolet crosslinking combined with mass spectrometry for the analysis of protein-DNA interactions. *Anal Biochem*, **331**, 303-13.
- Keyse, S.M. (2000) Protein phosphatases and the regulation of mitogen-activated protein kinase signalling. *Curr Opin Cell Biol*, **12**, 186-92.
- Kimura, M., Yamamoto, Y.Y., Seki, M., et al. (2003) Identification of Arabidopsis genes regulated by high light-stress using cDNA microarray. *Photochem Photobiol*, **77**, 226-33.
- Kingsmore, S.F. und Patel, D.D. (2003) Multiplexed protein profiling on antibody-based microarrays by rolling circle amplification. *Curr Opin Biotechnol*, **14**, 74-81.

- Klose, J. (1975) Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik*, **26**, 231-43.
- Klose, J., Nock, C., Herrmann, M., et al. (2002) Genetic analysis of the mouse brain proteome. *Nat Genet*, **30**, 385-93.
- Kramer, A. (2003) Herstellung von Gerste-Protein-Chips mit Proteinen aus cDNA-Expressions-Bibliotheken und Testung ihrer Anwendbarkeit für Phosphorylierungsstudien *in vitro*. *Diplomarbeit, Freie Universität Berlin*.
- Kramer, A., Feilner, T., Possling, A., et al. (2004) Identification of barley CK2alpha targets by using the protein microarray technology. *Phytochemistry*, **65**, 1777-84.
- Kreps, J.A., Wu, Y., Chang, H.S., et al. (2002) Transcriptome changes for Arabidopsis in response to salt, osmotic, and cold stress. *Plant Physiol*, **130**, 2129-41.
- Kroj, T., Rudd, J.J., Nurnberger, T., et al. (2003) Mitogen-activated protein kinases play an essential role in oxidative burst-independent expression of pathogenesis-related genes in parsley. *J Biol Chem*, **278**, 2256-64. Epub 2002 Nov 7.
- Lamb, C. und Dixon, R.A. (1997) The Oxidative Burst in Plant Disease Resistance. *Annu Rev Plant Physiol Plant Mol Biol*, **48**, 251-275.
- Lander, E.S., Linton, L.M., Birren, B., et al. (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.
- Lee, J., Klessig, D.F. und Nurnberger, T. (2001) A harpin binding site in tobacco plasma membranes mediates activation of the pathogenesis-related gene HIN1 independent of extracellular calcium but dependent on mitogen-activated protein kinase activity. *Plant Cell*, **13**, 1079-93.
- Lehrach, H., Bancroft, D. und and Maier, E. (1997) Robotics, computing, and biology. *Interdisciplinary Science Review*, **22**, 37-44.
- Lesaicherre, M.L., Uttamchandani, M., Chen, G.Y., et al. (2002) Antibody-based fluorescence detection of kinase activity on a peptide array. *Bioorg Med Chem Lett*, **12**, 2085-8.
- Lizcano, J.M., Deak, M., Morrice, N., et al. (2002) Molecular basis for the substrate specificity of NIMA-related kinase-6 (NEK6). Evidence that NEK6 does not phosphorylate the hydrophobic motif of ribosomal S6 protein kinase and serum- and glucocorticoid-induced protein kinase *in vivo*. *J Biol Chem*, **277**, 27839-49.

- Lu, C., Han, M.H., Guevara-Garcia, A., et al. (2002) Mitogen-activated protein kinase signaling in postgermination arrest of development by abscisic acid. *Proc Natl Acad Sci U S A*, **99**, 15812-7. Epub 2002 Nov 14.
- Lüking, A., Horn, M., Eickhoff, H., et al. (1999) Protein microarrays for gene expression and antibody screening. *Analytical Biochemistry*, **270**, 103-11.
- Lüking, A., Horn, S., Lehrach, H., et al. (2003a) A dual-expression vector allowing expression in *E. coli* and *P. pastoris*, including new modifications. *Methods Mol Biol*, **205**, 31-42.
- Lüking, A., Possling, A., Huber, O., et al. (2003b) A Nonredundant Human Protein Chip for Antibody Screening and Serum Profiling. *Mol Cell Proteomics*, **2**, 1342-1349.
- MacBeath, G. und Schreiber, S.L. (2000) Printing proteins as microarrays for high-throughput function determination. *Science*, **289**, 1760-1763.
- MAPKgroup. (2002) Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends Plant Sci*, **7**, 301-8.
- Martinsky, T. (2004) Protein Microarray Manufacturing. *PharmaGenomics*, **3**, 42-48.
- Mayer, K. (2001) DAS ARABIDOPSIS GENOM— Eine Blaupause für die Genome höherer Pflanzen. *GenomXPress*, **1**, 3-5.
- Meri, S. und Baumann, M. (2001) Proteomics: posttranslational modifications, immune responses and current analytical tools. *Biomol Eng*, **18**, 213-20.
- Michaud, G.A., Salcius, M., Zhou, F., et al. (2003) Analyzing antibody specificity with whole proteome microarrays. *Nat Biotechnol*, **21**, 1509-12. Epub 2003 Nov 9.
- Mizoguchi, T., Irie, K., Hirayama, T., et al. (1996) A gene encoding a mitogen-activated protein kinase kinase kinase is induced simultaneously with genes for a mitogen-activated protein kinase and an S6 ribosomal protein kinase by touch, cold, and water stress in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A*, **93**, 765-9.
- Mundy, J. und Schneitz, K. (2002) Protein phosphorylation in and around signal transduction. *Trends Plant Sci*, **7**, 54-5.
- Neill, S.J., Desikan, R., Clarke, A., et al. (2002) Nitric oxide is a novel component of abscisic acid signaling in stomatal guard cells. *Plant Physiol*, **128**, 13-6.
- Nilsson, J., Stahl, S., Lundeberg, J., et al. (1997) Affinity fusion strategies for detection, purification, and immobilization of recombinant proteins. *Protein Expr Purif*, **11**, 1-16.
- Nühse, T.S., Peck, S.C., Hirt, H., et al. (2000) Microbial elicitors induce activation and dual phosphorylation of the *Arabidopsis thaliana* MAPK 6. *J Biol Chem*, **275**, 7521-6.

- Nürnbergger, T. und Scheel, D. (2001) Signal transmission in the plant immune response. *Trends Plant Sci*, **6**, 372-9.
- O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, **250**, 4007-21.
- Ouaked, F., Rozhon, W., Lecourieux, D., et al. (2003) A MAPK pathway mediates ethylene signaling in plants. *Embo J*, **22**, 1282-8.
- Pandey, A. und Mann, M. (2000) Proteomics to study genes and genomes. *Nature*, **405**, 837-46.
- Parekh, R.B. und Rohlff, C. (1997) Post-translational modification of proteins and the discovery of new medicine. *Curr Opin Biotechnol*, **8**, 718-23.
- Patton, W.F. (2000) A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics. *Electrophoresis*, **21**, 1123-44.
- Perruc, E., Charpentreau, M., Ramirez, B.C., et al. (2004) A novel calmodulin-binding protein functions as a negative regulator of osmotic stress tolerance in *Arabidopsis thaliana* seedlings. *Plant J*, **38**, 410-20.
- Petricoin, E.F., 3rd, Hackett, J.L., Lesko, L.J., et al. (2002a) Medical applications of microarray technologies: a regulatory science perspective. *Nature Genetics*, **32 Suppl**, 474-9.
- Petricoin, E.F., Zoon, K.C., Kohn, E.C., et al. (2002b) Clinical proteomics: translating benchside promise into bedside reality. *Nature Reviews. Drug Discovery*, **1**, 683-95.
- Phizicky, E., Bastiaens, P.I., Zhu, H., et al. (2003) Protein analysis on a proteomic scale. *Nature*, **422**, 208-15.
- Rice, P., Longden, I. und Bleasby, A. (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet*, **16**, 276-7.
- Romeis, T., Piedras, P., Zhang, S., et al. (1999) Rapid Avr9- and Cf-9 -dependent activation of MAP kinases in tobacco cell cultures and leaves: convergence of resistance gene, elicitor, wound, and salicylate responses. *Plant Cell*, **11**, 273-87.
- Rubin, G.M., Yandell, M.D., Wortman, J.R., et al. (2000) Comparative genomics of the eukaryotes. *Science*, **287**, 2204-15.
- Sayed, M., Kim, S.O., Salh, B.S., et al. (2000) Stress-induced activation of protein kinase CK2 by direct interaction with p38 mitogen-activated protein kinase. *J Biol Chem*, **275**, 16569-73.

- Schatz, P.J. (1993) Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: a 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Biotechnology (N Y)*, **11**, 1138-43.
- Scheich, C., Sievert, V. und Bussow, K. (2003) An automated method for high-throughput protein purification applied to a comparison of His-tag and GST-tag affinity chromatography. *BMC Biotechnol*, **3**, 12.
- Schena, M., Shalon, D., Davis, R.W., et al. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467-70.
- Schweitzer, B., Roberts, S., Grimwade, B., et al. (2002) Multiplexed protein profiling on microarrays by rolling-circle amplification. *Nature Biotechnology*, **20**, 359-65.
- Seki, M., Ishida, J., Narusaka, M., et al. (2002) Monitoring the expression pattern of around 7,000 Arabidopsis genes under ABA treatments using a full-length cDNA microarray. *Funct Integr Genomics*, **2**, 282-91. Epub 2002 Aug 21.
- Snappyan, M., Lecocq, M., Guevel, L., et al. (2003) Dissecting DNA-protein and protein-protein interactions involved in bacterial transcriptional regulation by a sensitive protein array method combining a near-infrared fluorescence detection. *Proteomics*, **3**, 647-57.
- Staden, R., Beal, K.F. und Bonfield, J.K. (2000) The Staden package, 1998. *Methods Mol Biol*, **132**, 115-30.
- Teige, M., Scheickl, E., Eulgem, T., et al. (2004) The MKK2 pathway mediates cold and salt stress signaling in Arabidopsis. *Mol Cell*, **15**, 141-52.
- Templin, M.F., Stoll, D., Schrenk, M., et al. (2002) Protein microarray technology. *Drug Discov Today*, **7**, 815-22.
- Templin, M.F., Stoll, D., Schwenk, J.M., et al. (2003) Protein microarrays: promising tools for proteomic research. *Proteomics*, **3**, 2155-66.
- Tena, G., Asai, T., Chiu, W.L., et al. (2001) Plant mitogen-activated protein kinase signaling cascades. *Curr Opin Plant Biol*, **4**, 392-400.
- Tena, G. und Renaudin, J.P. (1998) Cytosolic acidification but not auxin at physiological concentration is an activator of MAP kinases in tobacco cells. *Plant J*, **16**, 173-82.
- Ulm, R., Ichimura, K., Mizoguchi, T., et al. (2002) Distinct regulation of salinity and genotoxic stress responses by Arabidopsis MAP kinase phosphatase 1. *Embo J*, **21**, 6483-93.
- Verwoerd, T.C., Dekker, B.M. und Hoekema, A. (1989) A small-scale procedure for the rapid isolation of plant RNAs. *Nucleic Acids Res*, **17**, 2362.

- Walhout, A.J., Temple, G.F., Brasch, M.A., et al. (2000) GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. *Methods in Enzymology*, **328**, 575-92.
- Walter, G., Bussow, K., Lueking, A., et al. (2002) High-throughput protein arrays: prospects for molecular diagnostics. *Trends in Molecular Medicine*, **8**, 250-3.
- Wan, J., Zhang, S. und Stacey, G. (2004) Activation of a mitogen-activated protein kinase pathway in Arabidopsis by chitin. *Molecular Plant Pathology*, **5**, 125.
- Willems, A.R., Goh, T., Taylor, L., et al. (1999) SCF ubiquitin protein ligases and phosphorylation-dependent proteolysis. *Philos Trans R Soc Lond B Biol Sci*, **354**, 1533-50.
- Yamagata, H., Saka, K., Tanaka, T., et al. (2001) Light activates a 46-kDa MAP kinase-like protein kinase in soybean cell culture. *FEBS Lett*, **494**, 24-9.
- Yu, J., Hu, S., Wang, J., et al. (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science*, **296**, 79-92.
- Zhang, S., Du, H. und Klessig, D.F. (1998) Activation of the tobacco SIP kinase by both a cell wall-derived carbohydrate elicitor and purified proteinaceous elicitors from *Phytophthora* spp. *Plant Cell*, **10**, 435-50.
- Zhang, S. und Klessig, D.F. (2001) MAPK cascades in plant defense signaling. *Trends Plant Sci*, **6**, 520-7.
- Zhu, H., Bilgin, M., Bangham, R., et al. (2001) Global analysis of protein activities using proteome chips. *Science*, **293**, 2101-5.
- Zhu, H., Bilgin, M. und Snyder, M. (2003b) Proteomics. *Annu Rev Biochem*, **72**, 783-812.
- Zhu, H., Klemic, J.F., Chang, S., et al. (2000) Analysis of yeast protein kinases using protein chips. *Nature Genetics*, **26**, 283-289.
- Zhu, H. und Snyder, M. (2003a) Protein chip technology. *Current Opinion in Chemical Biology*, **7**, 55-63.
- Zwerger, K. und Hirt, H. (2001) Recent advances in plant MAP kinase signalling. *Biol Chem*, **382**, 1123-31.

6 Abkürzungsverzeichnis

Abb.	Abbildung
ACC	1-aminocyclopropane-1-carboxylate
Aida	<i>Advanced Image Data Anlayzer</i>
AGI	<i>Arabidopsis</i> Genom Initiative
amol, fmol, nmol	attomol, femtomol, nanomol
AP	Alkalische Phosphatase
ATP	Adenosintriphosphat
ATM1	<i>Arabidopsis thaliana</i> meristem 1
BLAST	<i>Basic Lokal Alignment Search Tool</i> , Ähnlichkeits-Suchprogramm
Bp, Kbp, Mbp	Basenpaare, Kilobasenpaare, Megabasenpaare
BSA	Rinder-Serum-Albumin
bzw.	beziehungsweise
C	Konzentration
°C	Grad Celsius
ca.	circa
CaM-Kinase	calmodulinabhängige Proteinkinase
<i>ccd</i>	<i>control of cell death</i>
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
Ci	Curie
CK2	Casein-Kinase 2
ddH ₂ O	doppelt destilliertes Wasser
DEPC	Diethylpyrocarbonat
d. h.	das heißt
DNA, cDNA	Desoxyribonukleinsäure, <i>complementary</i> - Desoxyribonukleinsäure
dNTP	Desoxyribonukleosidtriphosphat (dATP, dCTP, dGTP, dTTP)
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DMSO	Dimethyl Sulfoxid
DTT	Dithiothreitol
<i>E. coli</i>	<i>Escherichia coli</i>
<i>et al.</i>	<i>et alias</i> (und andere)
EDTA	Ethylendiamintetraacetat
EGTA	Ethylenglycerol-bis-(β-Aminoethylether)-Tetracetat
EtOH	Ethanol
EST	Expressed Sequeunce Tag
evtl.	eventuell
g	Gramm
x g	x-fache der Erdbeschleunigung
h	Stunden
Hg	<i>Hintergrund</i>
HMFM	Hogness Modified Freezing Medium (<i>Gefriermix</i>)
i. V.	im Verhältnis
IPB	Institut für Pflanzenbiochemie, Halle
IgG	Immunglobulin G
IPTG	Isopropyl-β-D-thiogalacto-pyranosid
kDa	Kilodalton
L	Liter
LRR	<i>leucin-rich repeat</i>

Abkürzungsverzeichnis

m	Meter
m-	milli- (1×10^{-3})
μ -	mikro- (1×10^{-6})
M	Molar (mol/Liter)
MAPKK	Mitogen-aktivierte Proteinkinase-Kinase
MAPKKK	Mitogen-aktivierte Proteinkinase-Kinase-Kinase
MBP	Myelin-assoziiertes basisches Protein
min	Minuten
MPK	Mitogen-aktivierte Proteinkinase
MEK	<i>Mitogen-activated or extracellular signal regulated protein kinase kinase</i>
MPI-MG	Max-Planck-Institut für molekulare Genetik
n-	nano- (1×10^{-9})
N-terminal	aminoterminal
NBT/BCIP	Nitro-Blau-Tetrazolium-chlorid/5-bromo-4-chloro-3-Indolylphosphat-Toluidin-Salz
NiNTA	Nickel-Nitrilotriessigsäure
Nr.	Nummer
nt	Nukleotid
ODx	Optische Dichte bei einer Wellenlänge von x nm
ORF	<i>open reading frame</i>
PAA	Polyacrylamid
PAGE	Polyacrylamid-Gel-Elektrophorese
PBS	Phosphat gepufferte Salzlösung
PCR	Polymerasekettenreaktion
PEG	Polyethylenglycol
<i>phyB</i>	Phytochrom B
PKA	Proteinkinase A
PBST	Phosphat gepufferte Salzlösung mit Tween 20
pH	pH-Wert
PMSF	Phenylmethylsulfonylfluorid
PVDF	Polyvinylidenfluorid
<i>Query</i>	Trefferdie bei der Sequenzähnlichkeitssuche mittels BLAST-Analyse gefundene und ähnlichste Datenbanksequenz
RNA, mRNA	Ribonukleinsäure, <i>messenger</i> -Ribonukleinsäure
RNase	Ribonuklease
ROS	reaktive Sauerstoffspezies
RT	Reverse Transkriptase
s	Sekunde
SDS	Natrium-Dodecylsulfat
SIMK	<i>stress induced</i> MAPK
SIPK	<i>salicylate induced</i> MAPK
SSC	Standard-Salzlösung-Citrat
<i>Subject</i>	Abfrage-Sequenz; mit der die BLAST-Analyse durchgeführt wird
Tab.	Tabelle
TBS, TBST, TBST-T	Tris gepufferte Salzlösung, Tris gepufferte Salzlösung mit Tween 20, Tris gepufferte <i>Salzlösung</i> mit Tween 20 und Triton X 100
TBE	Tris-Borate-Na ₂ EDTA
TE	Tris-Na ₂ EDTA
TEMED	Tetramethylethylendiamin
TEN	Tris gepufferte Salzlösung-EDTA-Natrium Chlorid

Abkürzungsverzeichnis

TMV	Tabak Mosaikvirus
TRIS	Tris-(hydroxyl)-Aminomethan
ÜN	über Nacht
U	Unit (Einheit der Enzymaktivität)
u. a.	unter anderen
Upm	Umdrehungen pro Minute
UTR	untranslatierte Region
UV-Licht	Ultraviolettes Licht
V	Volt
vgl.	vergleiche
Vol.	Volumen
v/v	Volumen pro Volumen
w/v	Gewicht pro Volumen
z. B.	zum Beispiel
2DE	2-dimensionale Elektrophorese

7 Zusammenfassung

Ziel dieser Arbeit war es, sowohl eine Ressource an *Arabidopsis thaliana*-Proteinen zu schaffen, als auch Protein-Microarrays herzustellen und diese für funktionelle Studien einzusetzen. Als Plattform für diese funktionellen Analysen wurde die Protein-Microarray Methode gewählt. Diese Methode gehört neben den Macroarrays zu den Arraytechnologien. In beiden Fällen werden Proteine in systematischer Anordnung auf eine Oberfläche gebunden. Diese Technologien, als Weiterentwicklung der DNA Arraytechnologien, machen nun miniaturisierte und parallelisierte funktionelle Studien von Proteinen möglich.

Als Ressource zur Gewinnung von *Arabidopsis thaliana*-Proteinen wurden in der vorliegenden Arbeit zwei cDNA-Bibliotheken, ausgehend von mRNA aus den Geweben Pistill und Infloreszenz, hergestellt.

Die Pistill cDNA-Bibliothek besteht aus 660.800 Klonen in einem GATEWAY kompatiblen Klonierungsvektor und bildet damit die Grundlage für weitere genomweite Studien in verschiedenen Expressionssystemen.

Die Infloreszenz cDNA-Bibliothek (ATM1) besteht aus 40.000 Klonen in einem Protein Expressionsvektor. Diese Bibliothek wurde, unter Verwendung von Hochdichte Proteinfiltren (Macroarrays), auf proteinexprimierende Klone untersucht. Ausgehend davon wurde eine Proteinexpressions-Unterbibliothek hergestellt. Dieses besteht aus 5.000 Proteinexpressionsklonen. Diese Klone wurden zur weiteren Analyse vom 5'-Ende ansequenziert. Aus diesem Expressionssubset wurde nach umfangreicher bioinformatischer Analyse ein nahezu nicht redundantes, sogenanntes Uniklonset bestehend aus 1.502 Klonen erstellt. Mit diesem nicht redundanten Set sind Analysen mit dem gleichen Informationsgehalt, wie mit der redundanten Proteinexpressions-Unterbibliothek möglich. Dieses Uniklonset (1.502 Klone) wurde für weitere Analysen um 192 vollständige cDNA-Klone erweitert. Alle 1.694 Proteine, die von den Klonen dieses Erweiterten Uniklonsets exprimiert werden, wurden in hohem Durchsatz aufgereinigt.

Als eine erste funktionelle Studie dieser ca. 1.700 *Arabidopsis*-Proteine des Erweiterten Uniklonsets wurde die Phosphorylierung durch zwei *Arabidopsis* MAP-Kinasen unter Verwendung von Protein-Microarrays getestet. Nach Etablierung der ersten pflanzlichen Protein-Microarrays in unserer Arbeitsgruppe (Kersten, Feilner *et*

al., 2003) wurde eine Protein-Microarray basierten Proteomics Methode zur Untersuchung von Gerste Proteinen entwickelt (Kramer, Feilner *et al.*, 2004). Aufbauend auf dieser Methode wurde in der vorliegenden Arbeit ein Kinase Assay für *Arabidopsis*-Proteine etabliert. Dazu wurden alle ca. 1.700 gereinigten *Arabidopsis*-Proteine auf Nitrozellulose ähnliche beschichtete Microarrays gespottet. Diese wurden zuerst dazu verwendet, um ein Quantifizierungssystem unter Verwendung der Proteinkinase A zu etablieren, welches eine Selektion potentieller Phosphorylierungs-Targets ermöglicht. Mit diesem Quantifizierungssystem wurden dann Microarrays ausgewertet, die mit je einer der beiden *Arabidopsis* MAP-Kinasen 3 (MPK3) und 6 (MPK6) inkubiert worden waren. In diesen Studien wurden aus den ca. 1.700 *Arabidopsis*-Proteinen ca. 50 Targets der MPK3 und ca. 40 Targets der MPK6 identifiziert. Diese Ergebnisse bilden einerseits die Grundlage für zukünftige gezielte *in vivo* Untersuchung der neu identifizierten Targets. Andererseits kann dieser etablierte Phospholierungs-Assay auch für weitere Studien des Erweiterten Unikonsets mit anderen Kinasen, als auch generell für Proteine anderer Herkunft genutzt werden.

8 Summary

Aim of this work was to develop a resource of *Arabidopsis thaliana* proteins as well as creating protein-microarrays for use of functional studies. Protein-microarrays were chosen as a base for these functional analyses. Besides macroarrays this method belongs to the array technologies. In both cases proteins are arrayed onto surfaced in systematic arrays. These technologies, as further enhancement of DNA array technologies, are enabling miniaturised and parallelised functional studies of proteins. To receive the resource of *Arabidopsis thaliana*-proteins for this work, we generated two cDNA libraries, starting from mRNA of pistil and inflorescence.

The pistil cDNA-library consists of 660,800 clones in a GATEWAY compatible cloning vector and generates the base for further genome wide studies in different expression and analysis systems.

The inflorescence cDNA-library (ATM1) consists of 40,000 clones in a protein expression vector. This library was examined for protein expressing clones, using high-density protein-macroarrays. Based on this library we developed a protein expression sub-library. This sub-library consists of 5,000 protein expression clones. These clones were sequenced from the 5'-end. After extensive bioinformatics analysis we created from this expression subset a non redundant so-called unicloneset consisting of 1,502 clones. With this non-redundant set analyses are possible with the same information, as analyses with the redundant protein expression sub-library. This unicloneset (1,502 clones) was extended by 192 full length cDNA-clones for further analysis. From all 1,694 clones of the "extended unicloneset" proteins were expressed and purified in high-throughput.

The first functional study of the 1,700 *Arabidopsis*-proteins of the extended uniclone set, was the phosphorylation by two *Arabidopsis* MAP-kinases using protein-microarrays. After establishing the first plant protein-microarray (Kersten, Feilner *et al.*, 2003), we developed a proteomic method based on protein-microarrays (Kramer, Feilner *et al.*, 2004). Based on this method, we established a kinase assay for *Arabidopsis* proteins. Therefore all 1,700 purified *Arabidopsis*-proteins were spotted onto nitrocellulose coated microarrays surfaces. These were used to establish a quantification system using protein kinase A (PKA), which enables a selection of potential phosphorylation targets. Microarrays which were incubated with either

Arabidopsis MAP-Kinase 3 (MPK3) or 6 (MPK6) were analysed using this quantification system. In this study starting from ca. 1,700 *Arabidopsis*-proteins ca. 50 targets of MPK3 and ca. 40 targets of MPK6 were identified.

These results are a base for *in vivo* research of new identified targets and the established kinase assay can be used in particular for research of the extended unicloneset with other kinases as in general for studding further proteins, independent of origin.

9 Anhang

Anhang 8.1:

Auf den folgenden zwei Seiten (Seite 150-151) sind die Ergebnisse der BLAST-Analyse der Testsequenzierung der 48 Pistill cNDA-Klone dargestellt (vgl. 3.2.2.1).

sequence name	cluster size	BLAST-Hit	putative gene/function	e-value	start query [nt]	start subject [aa]	length subject [aa]	frame
P-c4		At1g05680	putative indole-3-acetate beta-glucosyltransferase	1,70E-65	164	144	453	2
P-h11		At1g06680	23 kDa polypeptide of oxygen-evolving complex (OEC)	1,70E-90	165	16	263	3
P-h12		At1g09640	unknown protein	2,70E-37	200	340	414	2
P-g7		At1g10360	putative glutathione S-transferase TSI-1	2,30E-31	178	95	227	1
P-g1		At1g12900	glyceraldehyde 3-phosphate dehydrogenase A, chloroplast precursor,	1,50E-82	199	75	399	1
P-f4		At1g21461	unknown protein	1,90E-68	197	44	247	2
P-c6		At1g25310	hypothetical protein	2,80E-42	188	63	151	2
P-a2		At1g42970	putative glyceraldehyde-3-phosphate dehydrogenase	1,50E-19	187	405	447	1
P-h4		At1g55460	putative protein	9,40E-67	199	221	411	1
P-b5	1	At1g60000	unknown protein	1,20E-66	171	25	258	3
P-b6		At1g60000	unknown protein	4,90E-86	189	25	258	3
P-c3		At1g61680	hypothetical protein	4,70E-81	183	166	474	3
P-b1		At1g72640	hypothetical protein	7,60E-22	222	1	153	3
P-c2		At1g73450		4,30E-67	196	820	1155	1
P-h1		At2g01930	unknown protein	1,60E-78	163	81	283	1
P-f3		At2g18170	MAP kinase (ATMPK7)	1,00E-91	156	9	368	3
P-h9		At2g21330	fructose biphosphate aldolase like protein	9,40E-83	180	171	399	3
P-d7		At2g27030	calmodulin	2,00E-29	192	6	149	3
P-e1		At2g28790	thaumatin like protein	2,30E-70	168	7	249	3
P-d11		At2g36800	putative glucosyl transferase	1,20E-65	496	105	495	-3
P-g11		At2g42530	cold-regulated protein cor1.5b precursor	3,90E-68	242	1	141	1
P-d2		At3g06510	beta glucosidase like protein	4,40E-22	196	572	622	1
P-d3		At3g10920	putative [Mn] superoxide dismutase	2,70E-85	189	1	231	3
P-e6		At3g10985	unknown protein	2,40E-52	273	1	110	3
P-b7		At3g13200	unknown protein	2,60E-72	242	1	230	2
P-f5		At3g25070	unknown protein	8,70E-97	224	1	211	2
P-b2		At3g26070	unknown protein	1,80E-65	136	36	242	1
P-d1		At3g44190	unknown protein	1,20E-74	208	217	367	1
P-h8		At3g47470	light-harvesting chlorophyll a/b-binding protein	7,40E-90	160	29	251	1
P-g2		At3g47810	putative protein	1,80E-49	332	1	190	2
P-b4		At3g55430	beta-1,3-glucanase - like protein	2,20E-85	197	249	449	2
P-d12		At3g62800	unknown protein	7,30E-67	194	228	355	2
P-a8		At4g01710	unknown protein	4,80E-65	195	2	132	3

P-d5	At4g03430	putative pre-mRNA splicing factor	1,80E-68	183	810	1029	3
P-g5	At4g08810	putative protein	1,90E-52	189	269	552	3
P-h5	At4g10280	unknown protein	3,40E-07	244	40	140	1
P-f9	At4g24770	31 kDa RNA binding protein (rbp31)	6,00E-65	196	6	329	1
P-e5	At4g35980	putative protein	2,20E-74	179	222	512	2
P-f6	At5g01530	chlorophyll a/b-binding protein CP29	3,40E-60	221	1	290	2
P-h2	At5g02960	unknown protein	3,70E-71	255	1	142	3
P-d4	At5g05230	unknown protein	4,10E-40	197	163	248	2
P-e1	At5g18380	40S RIBOSOMAL PROTEIN S16	5,80E-67	233	1	146	2
P-f2	At5g18380	40S RIBOSOMAL PROTEIN S16	1,10E-71	247	1	146	1
P-g8	At5g40340	unknown protein	2,70E-64	197	421	1008	2
P-e7	At5g47070	protein serine/threonine kinase-like	1,90E-61	197	218	410	2
P-f8	At5g47570	unknown protein	8,70E-64	202	1	125	1
P-a11	At5g61410	ribulose-5-phosphate-3-epimerase	7,70E-72	183	109	281	3
P-e5	At5g66730	zinc finger protein	7,90E-86	192	99	500	3

Anhang 8.2:

Auf den folgenden zwölf Seiten (Seite 152 - 163) sind die Ergebnisse der BLAST-Analyse der Testsequenzierung der 576 Infloreszenz cNDA-Klone dargestellt (vgl. 3.2.2.2).

sequence name	cluster size	BLAST-Hit	putative gene/function	e-value	start query [nt]	start subject [aa]	length subject [aa]	frame
atm6_E1	1	At1g01620	hypothetical protein	3,00E-82	41	1	286	2
atm4_D5	1	At1g02780	ribosomal protein L19, putative	1,00E-77	6	1	214	3
atm3_H9	1	At1g03430	putative AHP2	3,00E-19	260	1	192	2
atm4_F8	1	At1g03680	putative thioredoxin-m	3,00E-79	81	1	179	3
atm5_B4	3	At1g04270	putative 40S ribosomal protein S15	1,00E-61	69	1	152	3
atm5_C3	1	At1g04270	putative 40S ribosomal protein S15	4,00E-68	69	1	152	3
atm5_C4	1	At1g04270	putative 40S ribosomal protein S15	4,00E-74	69	1	152	3
atm5_E12	1	At1g04770	hypothetical protein	3,00E-68	48	1	302	3
atm1_E1	1	At1g05850	class I chitinase, putative	1,00E-87	115	1	321	1
atm2_G12	1	At1g06680	23 kDa polypeptide of oxygen-evolving complex (OEC)	6,00E-55	34	1	263	1
atm1_G7	2	At1g06760	histone H1, putative	1,00E-31	65	1	274	2
atm1_G8	1	At1g06760	histone H1, putative	2,00E-34	65	1	274	2
atm5_B9	2	At1g07890	L-ascorbate peroxidase	7,00E-81	78	1	250	3
atm5_C2	1	At1g07890	L-ascorbate peroxidase	1,00E-76	28	1	250	1
atm1_H4	1	At1g08110	glyoxalase I, putative	1,00E-87	45	1	185	3
atm4_C12	3	At1g09310	unknown protein	3,00E-64	10	1	179	1
atm5_A5	1	At1g09310	unknown protein	6,00E-66	28	1	179	1
atm5_H12	1	At1g09310	unknown protein	1,00E-59	38	1	179	2
atm1_D11	1	At1g12080	unknown protein	4,00E-04	61	1	138	1
atm3_E12	2	At1g13930	putative hydroxyproline-rich glycoprotein	1,00E-58	36	1	155	3
atm5_F9	1	At1g13930	putative hydroxyproline-rich glycoprotein	8,00E-58	39	1	155	3
atm1_F1	1	At1g14320	hypothetical protein	e-103	9	1	220	3
atm1_G10	3	At1g15820	hypothetical protein	5,00E-78	213	1	258	3
atm3_D9	1	At1g15820	hypothetical protein	3,00E-58	23	1	258	2
atm5_D9	1	At1g15820	hypothetical protein	2,00E-75	23	1	258	2
atm6_A11	1	At1g16460	putative thiosulfate sulfurtransferase	1,00E-56	163	1	318	1
atm1_A9	2	At1g20630	hypothetical protein	e-113	73	1	492	1
atm3_A11	1	At1g20630	hypothetical protein	1,00E-69	63	1	492	3
atm5_F4	1	At1g24020	pollen allergen-like protein	1,00E-73	45	1	155	3
atm4_F11	1	At1g27950	hypothetical protein	7,00E-74	13	1	193	1
atm3_H5	1	At1g29910	chlorophyll a/b-binding protein	6,00E-73	22	1	267	1
atm4_F7	1	At1g31330	photosystem I subunit III precursor, putative	5,00E-65	61	1	221	1
atm4_C6	1	At1g36240	60S ribosomal protein L30, putative	1,00E-56	41	1	112	2

atm3_B4	1	At1g52230	photosystem I subunit VI precursor	1,00E-59	10	1	145	1
atm5_E5	1	At1g52300	60S ribosomal protein L37, putative	6,00E-35	36	1	95	3
atm2_H10	1	At1g52600		1,00E-45	99	1	180	3
atm3_C12	1	At1g55360	unknown protein	2,00E-54	220	1	422	1
atm2_D3	1	At1g55520	transcription initiation factor TFIIID-2	1,00E-53	82	1	200	1
atm6_A6	2	At1g56580	unknown protein	3,00E-60	41	1	166	2
atm6_A7		At1g56580	unknown protein	3,00E-63	41	1	166	2
atm1_B4	1	At1g61170		2,00E-50	54	1	251	3
atm1_C5	2	At1g62370	hypothetical protein	1,00E-79	127	1	204	1
atm1_C6		At1g62370	hypothetical protein	1,00E-62	127	1	204	1
atm1_C7	1	At1g62370	hypothetical protein	1,00E-41	127	1	204	1
atm2_F11	2	At1g63290	putative D-ribose-5-phosphate	2,00E-66	33	1	227	3
atm3_A9		At1g65290		1,00E-64	14	1	126	2
atm2_F7	1	At1g67050	hypothetical protein	3,00E-41	73	1	264	1
atm1_H9	1	At1g67060	hypothetical protein	1,00E-74	97	1	182	1
atm2_E2		At1g70600	60S ribosomal protein L27A	1,00E-80	39	1	146	3
atm2_A8	1	At1g74670	GASTT-like protein	1,00E-33	219	1	80	3
atm3_G9	1	At1g75880	anter-specific proline-rich -like protein (APG-like)	2,00E-48	94	1	374	1
atm3_F8	1	At1g77270	hypothetical protein	3,00E-06	241	1	558	1
atm1_A8	1	At1g78040	similar to phosphoglycerate mutase 1 spP31217; similar	5,00E-95	44	1	171	2
atm1_D2	2	At1g79040	hypothetical protein	3,00E-57	34	1	140	1
atm2_B3		At1g79040	hypothetical protein	1,00E-48	38	1	140	2
atm6_D3	1	At1g79550	hypothetical protein	9,00E-27	109	1	401	1
atm3_C8	1	At1g80920	J8-like protein	6,00E-38	118	1	163	1
atm3_H10	1	At2g01350	putative nicotinate-nucleotide pyrophosphorylase	8,00E-68	89	1	323	2
atm3_D3	1	At2g02400	putative cinnamoyl-CoA reductase	e-103	15	1	318	3
atm4_E3	1	At2g06520	hypothetical protein	4,00E-37	45	1	116	3
atm4_D8		At2g15300	putative receptor-like protein kinase	4,00E-44	137	1	744	2
atm3_E4	1	At2g20820	unknown protein	1,00E-35	51	1	112	3
atm3_G8	1	At2g20870	unknown protein	2,00E-70	18	1	140	3
atm1_H3	1	At2g20920	unknown protein	e-103	32	1	287	2
atm6_B4	1	At2g21190	putative ER lumen protein retaining receptor	e-123	30	1	263	3
atm2_B4	1	At2g21660	glycine-rich RNA binding protein	4,00E-44	36	1	176	3
atm1_A5	1	At2g23780	putative RING zinc finger protein	2,00E-78	134	1	227	2
atm2_A12	1	At2g23930	putative small nuclear ribonucleoprotein E	1,00E-31	63	1	80	3
atm1_C8	1	At2g27750	unknown protein	5,00E-35	82	1	113	1

atm3 C10	1	A12g28950	putative expansin				2,00E-64	42	1	257	3
atm3 G1		A12g29550	tubulin beta-7 chain				4,00E-76	45	1	449	3
atm2 E11	1	A12g29960	cyclophilin				1,00E-67	23	1	201	2
atm4 B6	1	A12g30570	photosystem II reaction center 6.1KD protein				2,00E-40	54	1	133	3
atm1 F7	3	A12g30860	glutathione S-transferase				e-108	30	1	215	3
atm1 G4		A12g30860	glutathione S-transferase				e-109	30	1	215	3
atm3 E7		A12g30860	glutathione S-transferase				2,00E-83	31	1	215	1
atm5 H3	1	A12g31360	delta 9 desaturase				4,00E-74	40	1	311	1
atm4 F5	1	A12g31610	40S ribosomal protein; contains C-terminal domain				6,00E-97	55	1	250	1
atm4 G9	1	A12g32060	40S ribosomal protein S12				4,00E-64	51	1	144	3
atm3 A1	1	A12g33150	3-ketoacyl-CoA thiolase				2,00E-73	101	1	462	2
atm6 A1	1	A12g33450	putative chloroplast 50S ribosomal protein L28				5,00E-64	15	1	140	3
atm3 A10	2	A12g34420	photosystem II type I chlorophyll a/b binding protein				3,00E-94	42	1	265	3
atm6 E3		A12g34420	photosystem II type I chlorophyll a/b binding protein				5,00E-65	34	1	265	1
atm1 F9	1	A12g35370	glycine decarboxylase complex H-protein				2,00E-57	24	1	165	3
atm5 D4	1	A12g38270	unknown protein				6,00E-82	9	1	293	3
atm6 C1	1	A12g38530	putative nonspecific lipid-transfer protein				1,00E-48	41	1	118	2
atm3 G4	4	A12g38540	putative nonspecific lipid-transfer protein				3,00E-64	24	1	118	3
atm4 B4		A12g38540	putative nonspecific lipid-transfer protein				2,00E-64	33	1	118	3
atm5 B7		A12g38540	putative nonspecific lipid-transfer protein				3,00E-64	49	1	118	1
atm5 H9		A12g38540	putative nonspecific lipid-transfer protein				2,00E-64	49	1	118	1
atm6 B8	1	A12g39760	hypothetical protein				e-107	23	1	408	2
atm2 D6	1	A12g41430	dehydration-induced protein (ERD15)				5,00E-50	77	1	163	2
atm1 D1	1	A12g41470	unknown protein				2,00E-89	67	1	365	1
atm1 F12	1	A12g42540	cold-regulated protein cor15a precursor				3,00E-56	69	1	139	3
atm4 A7		A13g02470	S-adenosylmethionine decarboxylase				2,00E-06	667	1	366	1
atm5 F7	1	A13g02600	putative phosphatidate phosphohydrolase				2,00E-60	101	1	314	2
atm1 D9	1	A13g04710	ankyrin-like protein				7,00E-59	127	1	456	1
atm3 H4	1	A13g08030	unknown protein				6,00E-79	74	1	365	2
atm4 C9	1	A13g11400	putative eukaryotic translation initiation factor 3				1,00E-82	98	1	294	2
atm6 B9	1	A13g11940	putative 40S ribosomal protein S5				e-104	24	1	207	3
atm3 B6	1	A13g12145	leucine-rich repeat protein FLR1				6,00E-89	1	1	203	1
atm6 D9	1	A13g12370	hypothetical protein				1,00E-58	132	1	171	3
atm2 B9	2	A13g17970	phosphoprotein phosphatase, putative				1,00E-52	112	1	143	1
atm2 D1		A13g17970	phosphoprotein phosphatase, putative				9,00E-57	112	1	143	1
atm2 A6	1	A13g22210	hypothetical protein				1,00E-25	16	1	69	1

atm2_D10	1	At3g26740	light regulated protein, putative	1,00E-56	35	1	141	2
atm3_E10	1	At3g27160	unknown protein	2,00E-52	41	1	183	2
atm5_D7	1	At3g27850	50S ribosomal protein L12-C	1,00E-69	40	1	187	1
atm6_G5	1	At3g44310	nitrilase 1	1,00E-36	14	1	346	2
atm5_F12	1	At5g46010	actin depolymerizing factor 1 (ADF1)	1,00E-45	87	1	139	3
atm5_B1	1	At3g47470	CHLOROPHYLL A-B BINDING PROTEIN 4 PRECURSOR homolog	7,00E-62	81	1	251	3
atm4_C5	1	At3g48890	putative progesterone-binding protein homolog.Atmp2	2,00E-96	27	1	233	3
atm4_G1	1	At3g49010	60S ribosomal protein L13, BBC1 protein	2,00E-98	36	1	206	3
atm1_B5	4	At3g49270	hypothetical protein	2,00E-23	189	1	121	3
atm1_F8		At3g49270	hypothetical protein	2,00E-64	168	1	121	3
atm2_H1		At3g49270	hypothetical protein	1,00E-29	189	1	121	3
atm2_H2		At3g49270	hypothetical protein	3,00E-51	189	1	121	3
atm6_B5	1	At3g52590	ubiquitin / ribosomal protein CEP52	1,00E-69	14	1	128	2
atm2_H4	1	At3g56680	putative protein	1,00E-33	219	1	353	3
atm3_G7	1	At3g56940	leucine zipper-containing protein AT103	6,00E-65	31	1	409	1
atm2_C3	2	At3g58680	transcriptional coactivator - like protein	4,00E-69	57	1	142	3
atm6_F1		At3g58680	transcriptional coactivator - like protein	4,00E-74	48	1	142	3
atm3_A8	1	At3g60500	nucleolar autoantigen - like protein	1,00E-80	120	1	438	3
atm6_B3	1	At4g00430	probable plasma membrane intrinsic protein 1c	2,00E-78	8	1	295	2
atm4_C3	1	At4g03210	putative xyloglucan endotransglycosylase	e-101	36	1	290	3
atm1_G6		At4g03280	putative component of cytochrome B6-F complex	2,00E-74	14	1	229	2
atm1_F2	1	At4g03520	putative M-type thioredoxin	2,00E-83	61	1	186	1
atm2_A1		At4g05320	polyubiquitin (ubq10)	7,00E-40	53	1	464	2
atm2_B1	2	At4g10340	chlorophyll a/b-binding protein - like	9,00E-38	23	1	280	2
atm6_E6		At4g10340	chlorophyll a/b-binding protein - like	2,00E-65	40	1	280	1
atm3_A12	1	At4g12800	probable photosystem I chain XI precursor	6,00E-62	9	1	219	3
atm5_F3	1	At4g15140	putative protein	2,00E-63	6	1	175	3
atm5_G10	1	At4g16500	cysteine proteinase inhibitor like protein	1,00E-28	13	1	117	1
atm6_D4	1	At4g17390	60S ribosomal protein L15 homolog	1,00E-77	17	1	204	2
atm3_C3	1	At4g18140	putative protein	9,00E-55	74	1	307	2
atm2_H5	1	At4g21620	putative protein	2,00E-15	21	1	131	3
atm1_E7	1	At4g22880	putative leucoanthocyanidin dioxygenase (LDOX)	3,00E-69	134	1	356	2
atm5_H6	1	At4g23630	putative protein	6,00E-23	77	1	275	2
atm5_C1	2	At4g23670	putative major latex protein	9,00E-64	40	1	151	1
atm5_D10		At4g23670	putative major latex protein	2,00E-75	40	1	151	1
atm4_D12	1	At4g25030	putative protein	3,00E-60	48	1	344	3

atm5_H11	1	At4g28340	hypothetical protein					3,00E-23	124	1	159	1
atm2_H7	1	At4g29040	26S proteasome subunit 4-like protein					2,00E-50	12	1	443	3
atm3_C9	1	At4g32150	synaptobrevin-like protein					4,00E-58	198	1	219	3
atm1_D4	1	At4g32470	ubiquitin-cytochrome c reductase - like protein					2,00E-49	52	1	122	1
atm5_E6	1	At4g33865	ribosomal S29 subunit					1,00E-30	18	1	56	3
atm2_D12	1	At4g34870	peptidyl prolyl isomerase (cyclophilin)					1,00E-74	39	1	172	3
atm3_F1	1	At4g36800	ubiquitin--protein ligase-like protein					9,00E-74	57	1	183	3
atm3_C5	1	At4g39980	2-dehydro-3-deoxyphosphoheptanate aldolase					1,00E-63	178	1	525	1
atm1_H5	1	At5g02790	putative protein					e-116	115	1	235	1
atm3_G6	1	At5g02840	putative protein					6,00E-56	68	1	307	2
atm4_B10	2	At5g04800	40S ribosomal protein S17					1,00E-74	86	1	141	2
atm4_B11		At5g04800	40S ribosomal protein S17					2,00E-72	86	1	141	2
atm3_G11	1	At5g07550	glycine-rich protein PUTG1					9,00E-46	37	1	106	1
atm1_H6	2	At5g10980	histon H3 protein					9,00E-72	65	1	136	2
atm1_H7		At5g10980	histon H3 protein					3,00E-52	65	1	136	2
atm3_F2	1	At5g14150	putative protein					9,00E-36	143	1	383	2
atm4_C10	1	At5g15090	voltage-dependent anion-selective channel protein hsr2					7,00E-98	58	1	274	1
atm1_A12	1	At5g20290	putative protein					3,00E-41	30	1	222	3
atm6_A4	1	At5g20740	ripening-related protein - like					6,00E-70	12	1	205	3
atm5_C11		At5g22430	unknown protein					7,00E-73	31	1	172	1
atm6_E4		At5g22430	unknown protein					1,00E-67	37	1	172	1
atm2_E8	1	At5g23120	photosystem II stability/assembly factor HCF136					5,00E-26	30	1	403	3
atm3_B8	2	At5g24780	vegetative storage protein Vsp1					8,00E-83	8	1	270	2
atm5_B8	1	At5g27850	60S ribosomal protein - like					e-100	4	1	187	1
atm1_A7		At5g38410	ribulose biphosphate carboxylase small chain 3b					1,00E-81	40	1	181	1
atm3_G10	1	At5g40580	20S proteasome beta subunit PBB2					3,00E-66	41	1	274	2
atm6_B2	1	At5g40950	50S ribosomal protein L27					8,00E-96	23	1	198	2
atm4_A9	1	At5g42780	putative protein					2,00E-89	8	1	242	2
atm5_H4	1	At5g45775	ribosomal protein L11-like					2,00E-83	45	1	182	3
atm3_D6	1	At5g47570	unknown protein					4,00E-56	22	1	125	1
atm3_F12	1	At5g47700	60S acidic ribosomal protein P1-like protein					1,00E-35	73	1	113	1
atm3_A5	1	At5g50040	putative protein					2,00E-72	2	1	175	2
atm5_G4	1	At5g50680	ubiquitin activating enzyme					3,00E-68	17	1	320	2
atm6_H5	1	At5g51410	arginine-aspartate-rich RNA binding protein-like					5,00E-18	151	1	334	1
atm2_B5	1	At5g54290	cytochrome c biogenesis protein precursor (gb AAF35369.1)					e-101	26	1	354	2
atm3_D10	1	At5g54580	putative protein					2,00E-56	113	1	156	2

atm6 C4	2	At5g54770	thiazole biosynthetic enzyme precursor (ARA6) (sp Q38814)	4,00E-47	12	1	349	3
atm6 C8		At5g54770	thiazole biosynthetic enzyme precursor (ARA6) (sp Q38814)	9,00E-53	27	1	349	3
atm2 C8	1	At5g56020	putative protein	1,00E-44	52	1	230	1
atm4 D1	1	At5g58500	putative protein	3,00E-89	78	1	182	3
atm4 A2	1	At5g59310	nonspecific lipid-transfer protein precursor - like	1,00E-56	22	1	110	1
atm3 G5	1	At5g60360	AALP protein	8,00E-44	184	1	358	1
atm2 E6	1	At5g60750	putative protein	4,00E-69	157	1	319	1
atm6 A10	1	At5g61240	CF-5 disease resistance protein - like	9,00E-66	10	1	380	1
atm5 G8	1	At5g61660	structural protein - like	1,00E-04	127	1	105	1
atm4 A5	1	At5g62300	ribosomal protein S20 - like	2,00E-66	58	1	124	1
atm2 B7	1	At5g64080	putative protein	3,00E-73	2	1	173	2
atm4 H8	1	At5g66290	unknown protein	4,00E-77	77	1	202	2
atm6 A3	1	At5g66320	GATA-binding transcription factor-like protein	1,00E-19	139	1	339	1
atm3 B12	1	At5g66880	protein kinase, 41K (EC 2.7.1.-) (pir S71172)	2,00E-57	128	1	361	2
atm1 H1	1	At5g67330	natural resistance-associated macrophage protein	5,00E-39	60	1	512	3
atm4 B5	1	At2g21250	putative NADPH dependent mannose 6-phosphate reductase	e-115	2	2	309	2
atm6 D1	1	At2g47400	putative chloroplast protein CPI2	6,00E-54	2	2	124	2
atm1 C4	1	At3g01470	homeobox-leucine zipper protein HAT5 (HD-ZIP protein 5)	0,004	513	2	272	3
atm6 E5	1	At4g04020	putative fibrillin	2,00E-86	2	2	318	2
atm3 H3	1	At4g04610	5'-adenylylsulfate reductase	3,00E-53	29	2	465	2
atm3 D4	1	At4g21580	putative NADPH quinone oxidoreductase	8,00E-85	2	2	325	2
atm3 C2	1	At1g67290	hypothetical protein	1,00E-44	1	3	615	1
atm1 E12	1	At5g52930	fructose biphosphate aldolase - like protein	1,00E-78	1	3	358	1
atm5 G5	1	At5g06270	putative protein	3,00E-39	2	4	122	2
atm5 F5	1	At1g11850	unknown protein	5,00E-07	2	5	108	2
atm1 H11	1	At1g42970	hypothetical protein	e-107	2	5	447	2
atm4 B3	1	At1g44575	photosystem II 22kDa protein, putative	7,00E-35	1	6	265	1
atm6 A5	2	At3g20670	histone H2A, putative	7,00E-55	2	6	132	2
atm4 D4	1	At4g29350	profilin 2	5,00E-70	77	6	131	2
atm1 D3	1	At1g22300	hypothetical protein	2,00E-70	71	7	254	2
atm2 C10	2	At1g70600	60S ribosomal protein L27A	7,00E-81	1	7	146	1
atm2 D7	1	At2g45960	aquaporin (plasma membrane intrinsic protein 1B)	2,00E-82	151	7	286	1
atm2 E10	2	At4g33220	pectinesterase - like protein	9,00E-14	215	8	477	2
atm2 H9		At4g33220	pectinesterase - like protein	6,00E-29	215	8	477	2
atm3 F11	1	At2g47170	ADP-ribosylation factor 1	8,00E-78	2	9	181	2
atm3 E1	1	At5g11420	putative protein	1,00E-48	87	9	366	3

				e-112	2	11	235
atm4 B9	1	At1g16470	multicatalytic endopeptidase				
atm3 H6	1	At5g14470	putative protein	8,00E-70	22	11	366
atm4 E7	1	At5g47200	ras-related small GTP-binding protein-like	2,00E-80	1	11	159
atm6 G3	1	At1g12760	hypothetical protein	0,098	18	13	368
atm5 A11	2	At1g29930	hypothetical protein	1,00E-81	82	13	267
atm3 E6	1	At5g65260	poly(A)-binding protein II-like	3,00E-69	66	13	220
atm4 F3	1	At2g22090	putative RNA-binding protein	5,00E-67	1	14	347
atm3 D1	1	At5g61410	ribulose-5-phosphate-3-epimerase	5,00E-48	115	14	280
atm2 F2	9	At1g67090	hypothetical protein	4,00E-94	51	15	180
atm2 F9		At1g67090	hypothetical protein	2,00E-65	51	15	180
atm3 B11		At1g67090	hypothetical protein	3,00E-64	60	15	180
atm5 B10		At1g67090	hypothetical protein	5,00E-91	48	15	180
atm5 D8		At1g67090	hypothetical protein	4,00E-80	48	15	180
atm6 C5		At1g67090	hypothetical protein	3,00E-89	49	15	180
atm6 E12		At1g67090	hypothetical protein	3,00E-33	48	15	180
atm6 G10		At1g67090	hypothetical protein	3,00E-88	48	15	180
atm1 B6	2	At5g38430	ribulose biphosphate carboxylase small chain 1b	1,00E-91	53	15	181
atm1 B8		At5g38430	ribulose biphosphate carboxylase small chain 1b	3,00E-88	53	15	181
atm4 F12	1	At1g62500	putative proline-rich cell wall protein (pir S2985);	0,014	2	16	297
atm6 A12	1	At5g26000	myrosinase precursor	8,00E-81	66	16	541
atm1 D6	1	At3g11630	putative 2-cys peroxidoxin	3,00E-59	76	17	266
atm3 E9	1	At1g16650	Hypothetical protein	6,00E-44	87	18	489
atm3 H2	1	At1g23390	unknown protein	1,00E-94	39	18	394
atm6 H4		At1g29930	hypothetical protein	8,00E-68	1	18	267
atm3 D12	1	At3g62870	60S RIBOSOMAL PROTEIN L7A protein	2,00E-46	47	19	256
atm3 A7	1	At2g23150	putative metal ion transporter (NRAMP)	8,00E-61	65	20	509
atm4 G7	1	At4g34670	Putative S-phase-specific ribosomal protein	1,00E-42	121	20	262
atm4 B12	1	At5g49440	unknown protein	5,00E-77	421	22	181
atm1 D5	1	At1g62940	4-coumarate coenzyme A ligase, putative	5,00E-94	1	24	542
atm2 H11	1	At5g02380	metallothionein 2b	9,00E-28	136	24	77
atm1 C2	1	At3g09390	metallothionein-like protein	4,00E-25	71	25	81
atm1 C2	1	At3g09390	metallothionein-like protein	5,00E-26	71	25	81
atm5 A9	1	At1g60170		2,00E-67	119	27	511
atm2 D5	1	At3g56800	calmodulin-3	7,00E-65	99	27	149
atm3 E2	1	At4g18270	putative protein	3,00E-24	1	28	313
atm3 C4	1	At3g21020	unknown protein	0,94	239	29	310

atm3_H12	1	At5g44080	bZIP protein AtbZIP13			2,1	155	31	315	-2
atm1_E8	1	At3g15630	unknown protein			8,00E-39	2	32	107	2
atm3_H11	1	At3g61470	Lhca2 protein			2,00E-78	124	32	257	1
atm3_B7	1	At5g63040	putative protein			2,00E-30	246	34	366	3
atm5_E4	1	At1g78630	Hypothetical protein			3,00E-93	3	38	241	3
atm1_C11	1	At3g16950	dihydroipoamide dehydrogenase [pd]			9,00E-16	3	38	570	3
atm4_B2	1	At1g04850	unknown protein			4,00E-06	411	40	400	3
atm2_A10	2	At3g57490	40S ribosomal protein S2 homolog			8,00E-56	168	42	276	3
atm2_A9		At3g57490	40S ribosomal protein S2 homolog			2,00E-70	168	42	276	3
atm1_E5	2	At4g34620	putative ribosomal protein S16			9,00E-36	3	42	113	3
atm1_E6		At4g34620	putative ribosomal protein S16			8,00E-36	3	42	113	3
atm4_B1	1	At3g12930	hypothetical protein			1,00E-84	3	46	238	3
atm5_F1	1	At2g41840	40S ribosomal protein S2			3,00E-37	184	48	285	1
atm2_C12	2	At1g25170				1,00E-34	585	51	469	-1
atm1_G1	3	At2g45180	putative proline-rich protein			1,00E-44	164	52	134	2
atm1_G2		At2g45180	putative proline-rich protein			1,00E-44	164	52	134	2
atm4_A1		At2g45180	putative proline-rich protein			1,00E-44	167	52	134	2
atm2_D11		At1g25170				4,00E-27	573	53	469	-2
atm4_C7	1	At5g14740	CARBONIC ANHYDRASE 2			2,00E-79	142	53	331	1
atm6_E11	1	At5g23750	putative protein			1,00E-04	220	55	202	1
atm6_A8		At3g20670	histone H2A, putative			5,00E-10	166	60	132	1
atm5_A2	1	At4g21960	peroxidase prx1			e-109	3	63	323	3
atm5_C9	1	At5g60910	MAD box containing protein NAP1-1 - like			2,00E-37	60	63	154	3
atm6_A9	1	At1g24070	glucosyltransferase, putative			1,00E-46	5	69	552	2
atm4_H9	1	At1g18210	hypothetical protein			1,00E-42	2	76	170	2
atm2_C7	3	At5g22430	unknown protein			8,00E-35	19	79	172	1
atm2_F4	1	At1g05680	putative indole-3-acetate beta-glucosyltransferase			e-100	1	80	453	1
atm2_G10	2	At5g21940	putative protein			1,00E-59	3	80	264	3
atm2_G9		At5g21940	putative protein			9,00E-54	3	80	264	3
atm1_A11	1	At3g24830	60S ribosomal protein, putative			3,00E-66	1	83	206	1
atm6_B7	1	At5g09810	ACTIN 2/7 (sp P53492)			e-130	2	85	377	2
atm4_A3	1	At1g70490	putative ADP-ribosylation factor 1			6,00E-49	1	86	181	1
atm3_F3	1	At3g55770	transcription factor L2			1,00E-53	2	86	199	2
atm1_A6	2	At5g38410	ribulose biphosphate carboxylase small chain 3b			3,00E-81	323	96	181	2
atm3_E3	1	At5g16710	valine-tRNA ligase-like protein			2,00E-81	3	97	1093	3
atm1_B3	2	At4g03280	putative component of cytochrome B6-F complex			9,00E-89	238	100	229	1

atm4 C1	1	At3g16240	delta tonoplast integral protein (delta-TIP)	3,00E-67	1	101	250	1
atm5 C12	1	At1g30630	coatamer-like protein, epsilon subunit	2,00E-85	1	107	292	1
atm4 F4	1	At1g76970		4,00E-92	1	113	387	1
atm4 D10	1	At3g60390	homeobox-leucine zipper protein HAI3	0,27	296	117	315	-3
atm6 D12	1	At1g10890	unknown protein	1,00E-06	3	118	592	3
atm6 G6	1	At3g50500	protein kinase SPK-2	3,00E-99	1	119	362	1
atm4 G8		At1g67090	hypothetical protein	1,00E-29	3	123	180	3
atm1 G12	1	At3g01980	unknown protein	3,00E-90	5	123	296	2
atm3 E5	1	At3g13800	hydrolase, putative	3,00E-38	14	125	282	2
atm5 H1	1	At4g32260	H+-transporting ATP synthase chain 9 - like protein	6,00E-26	1	125	219	1
atm1 D7	3	At4g05320	polyubiquitin (ubq10)	e-107	3	126	464	3
atm3 A4	1	At4g26970	putative aconitase	7,00E-43	108	133	907	3
atm2 C2	1	At2g35020	putative UDP-N-acetylglucosamine pyrophosphorylase	9,00E-66	107	134	502	2
atm1 C10	1	At4g08590	putative zinc finger protein	3,6	307	135	432	1
atm5 C7	1	At3g08580	adenylate translocator	6,00E-96	1	138	381	1
atm5 A3		At5g62700	tubulin beta-2/beta-3 chain (sp P29512)	6,00E-91	3	144	450	3
atm6 H10		At5g62700	tubulin beta-2/beta-3 chain (sp P29512)	e-102	3	144	450	3
atm1 F3	1	At3g51410	putative protein	2,00E-24	36	145	255	3
atm3 B5	1	At3g62530	putative protein	3,00E-63	103	150	342	1
atm1 F10	2	At5g16840	putative protein	1,00E-46	2	150	259	2
atm1 F11		At5g16840	putative protein	1,00E-46	2	150	259	2
atm1 D12	1	At5g24770	vegetative storage protein Vsp2	1,00E-64	2	150	265	2
atm3 F5	1	At1g69680	hypothetical protein	6,00E-14	1	152	186	1
atm1 G9	2	At2g29550	tubulin beta-7 chain	e-104	41	152	449	2
atm5 D11	1	At2g44720	unknown protein	2,00E-11	4	152	239	1
atm2 D4	1	At3g01500	carbonic anhydrase, chloroplast precursor	e-103	2	154	336	2
atm3 B2	1	At5g38420	ribulose biphosphate carboxylase small chain 2b	2,00E-08	1	157	181	1
atm6 C9	1	At1g28290	proline-rich protein, putative	2,2	511	162	359	1
atm3 A6	1	At1g67870	unknown protein	0,005	346	163	191	1
atm4 B8	1	At5g08680	H+-transporting ATP synthase beta chain (mitochondrial)	7,00E-91	1	164	559	1
atm2 E12	1	At1g62050	hypothetical protein	3,00E-69	1	167	605	1
atm2 F8	1	At5g23940	acyltransferase	3,00E-93	2	169	484	2
atm1 H2	1	At5g20720	chloroplast Cpm21 protein	3,00E-39	1	174	253	1
atm1 B2	1	At1g27580	hypothetical protein	5,00E-10	52	175	380	1
atm5 F10	1	At3g49850	MYB-like protein	6,1	503	175	295	2
atm5 G3	1	At1g12830	unknown protein	2,5	334	176	213	1

atm1_F6	1	At1g76010	unknown protein	6,00E-36	1	178	350	1
atm4_C4	1	At4g11320	cysteine proteinase-like protein	5,00E-74	2	178	371	2
atm3_F6	1	At5g37490	putative protein	1,00E-55	1	178	435	1
atm6_G1	1	At3g12290	hypothetical protein	8,00E-62	2	183	299	2
atm2_B2	1	At5g02240	putative protein	2,00E-29	1	184	376	1
atm5_A8	1	At5g05610	nucleic acid binding protein-like	2,00E-32	2	186	241	2
atm1_C1	2	At5g01530	chlorophyll a/b-binding protein CP29	2,00E-54	2	187	290	2
atm1_C3		At5g01530	chlorophyll a/b-binding protein CP29	1,00E-54	2	187	290	2
atm5_E10		At5g24780	vegetative storage protein Vsp1	1,00E-42	1	189	270	1
atm5_G7	1	At1g56340	hypothetical protein	1,00E-80	2	194	425	2
atm2_A11	1	At4g39280	phenylalanyl-trna synthetase - like protein	2,00E-52	31	197	428	1
atm3_B1	3	At5g62700	tubulin beta-2/beta-3 chain (sp P29512)	1,00E-92	1	197	450	1
atm1_D8		At4g05320	polyubiquitin (ubq10)	2,00E-99	3	202	464	3
atm1_E11	1	At3g47080	putative protein	9,00E-72	2	203	515	2
atm1_A2	1	At3g10410	putative serine carboxypeptidase precursor	3,00E-78	1	209	516	1
atm6_B10	1	At1g76930	extensin	0,37	510	213	373	-1
atm3_H8	1	At3g07340	unknown protein	3,00E-89	2	213	456	2
atm4_F9	1	At3g58660	putative protein	4,00E-40	3	213	446	3
atm5_D5	1	At3g12610	hypothetical protein	5,00E-80	2	231	372	2
atm3_H7	1	At2g30200	putative malonyl-CoA:Acyl carrier protein transacylase	2,00E-16	477	232	367	-1
atm4_C11	1	At3g50660	steroid 22-alpha-hydroxylase (DWF4)	1,00E-89	2	239	513	2
atm3_E11	1	At1g69870	putative peptide transporter	5,00E-80	1	243	644	1
atm5_E1	2	At3g20920	hypothetical protein	1,00E-65	3	245	365	3
atm5_E2		At3g20920	hypothetical protein	5,00E-65	3	245	365	3
atm2_B8	1	At1g11860	aminomethyltransferase-like precursor protein	1,00E-89	3	248	408	3
atm5_G1		At3g02470	S-adenosylmethionine decarboxylase	3,00E-53	3	248	366	3
atm3_D7	1	At5g47210	putative protein	4,00E-40	2	248	357	2
atm3_C7	1	At3g21670	nitrate transporter	7,00E-82	3	274	590	3
atm1_E10	1	At1g03900	unknown protein	0,23	2	276	585	2
atm4_D2	2	At3g21600	unknown protein	2,00E-32	2	289	374	2
atm4_D3		At3g21600	unknown protein	2,00E-32	2	289	374	2
atm5_F2	1	At1g09840	shaggy-like protein kinase, kappa	7,00E-77	1	290	421	1
atm6_A2	1	At3g54190	putative protein	1,00E-92	3	292	454	3
atm1_G3	1	At4g34110	poly(A)-binding protein	5,00E-86	2	292	629	2
atm6_E10	1	At5g08410	lipic acid synthase - like protein	2,00E-26	27	294	477	3
atm1_F4	2	At5g44190	putative protein	4	319	299	386	-2

atm1_F5		At1g44190	putative protein		4,1	319	299	386	-2
atm5_C10	1	At1g08510	acyl-(acyl carrier protein) thioesterase, putative		2,00E-60	1	304	412	1
atm5_B12	1	At3g02350	hypothetical protein		e-124	2	304	561	2
atm5_G9	1	At5g08570	pyruvate kinase		1,00E-90	3	304	510	3
atm2_G3	3	At3g02470	S-adenosylmethionine decarboxylase		2,00E-15	1	311	366	1
atm3_A3	1	At1g55530			2,2	2	312	351	2
atm4_A8	1	At1g26250	hypothetical protein		1,1	245	316	443	2
atm1_A1	1	At1g21750	putative protein disulfide isomerase precursor		4,00E-33	1	319	501	1
atm2_A3	1	At2g16570	amidophosphoribosyltransferase		8,00E-71	2	320	566	2
atm6_F7	1	At4g32940	gamma-VPE (vacuolar processing enzyme)		6,00E-95	2	322	494	2
atm4_A11	1	At1g52400	beta-glucosidase, putative		e-108	3	323	528	3
atm3_C1	1	At1g60200			8,00E-55	2	332	781	2
atm3_G3	1	At3g54220	SCARECROW1		7,00E-79	2	339	653	2
atm1_E9	1	At4g23040	putative protein		1,00E-67	2	339	577	2
atm2_G5	1	At2g26460	unknown protein		e-109	1	349	585	1
atm2_E4	2	At2g29650	putative Na ⁺ -dependent inorganic phosphate cotransporter		1,00E-81	3	350	512	3
atm2_E5		At2g29650	putative Na ⁺ -dependent inorganic phosphate cotransporter		4,00E-81	3	350	512	3
atm6_B12	1	At3g55940	phosphoinositide-specific phospholipase C - like protein		4,00E-50	761	353	584	-1
atm6_C7	1	At5g56030	HEAT SHOCK PROTEIN 81-2 (HSP81-2)(spP55737)		1,00E-64	2	355	699	2
atm4_A12	1	At5g57655	xylose isomerase		2,00E-58	3	355	472	3
atm5_H10	1	At3g09440	heat-shock protein (At-hsc70-3)		1,00E-68	2	364	649	2
atm4_E2	1	At3g19420	putative tyrosine phosphatase		e-112	3	369	611	3
atm5_A10	1	At3g13870	root hair defective 3		2,00E-96	1	380	802	1
atm5_E7	1	At5g35630	glutamate-ammonia ligase (EC 6.3.1.2) precursor,		4,00E-22	129	385	430	3
atm5_E8	1	At5g35630	glutamate-ammonia ligase (EC 6.3.1.2) precursor,		4,00E-22	129	385	430	3
atm4_H5	1	At2g27350	unknown protein		2,00E-53	60	392	522	3
atm5_D1	1	At1g05560	hypothetical protein		1,00E-61	2	404	519	2
atm6_G9	1	At4g38430	putative protein		1,00E-72	1	406	548	1
atm2_B6	1	At4g38770	extensin - like protein		3,00E-05	407	433	448	2
atm6_D2	1	At3g55980	putative protein		1,00E-66	3	436	586	3
atm2_C5	1	At3g53230	CDC48 - like protein		9,00E-86	5	441	815	2
atm3_G2	1	At2g45810	putative ATP-dependent RNA helicase		2,00E-32	1	457	528	1
atm3_F4	1	At5g51150	putative protein		1,00E-18	2	470	531	2
atm4_E12	1	At5g50310	putative protein		1,00E-40	250	477	596	1
atm5_E11	1	At1g08540	plastid RNA polymerase sigma-subunit (SIG1)		1,00E-22	2	481	572	2
atm3_B3	1	At2g01570	putative RGA1, gibberellin response modulation protein		5,00E-48	1	488	587	1

atm3 B10	1	At4g04310	putative transposon protein	8,5	574	529	1011	-2
atm1 C12	1	At3g48670	putative protein	6,00E-49	1	541	644	1
atm6 F4	1	At5g02500	dnaK-type molecular chaperone hsc70.1	1,00E-46	1	548	651	1
atm2 A7	1	At5g63410	putative protein	8,00E-32	2	557	637	2
atm6 C10	1	At5g17920	5-methyltetrahydropteroyltriglutamate--homocysteine	1,00E-66	1	559	765	1
atm6 B11	1	At1g51570	anthranilate phosphoribosyltransferase, putative	e-104	3	574	776	3
atm3 F7	1	At3g45140	lipoxigenase AtLOX2	e-108	3	580	870	3
atm2 D8	2	At3g16000	myosin heavy chain-like protein	9,00E-51	137	596	724	2
atm2 E1		At3g16000	myosin heavy chain-like protein	9,00E-46	103	596	724	1
atm1 A3	1	At4g24190	HSP90-like protein	3,00E-52	140	600	823	2
atm3 D5	1	At1g29350	unknown protein	3,00E-77	3	614	858	3
atm1 E3	7	At2g18790	phytochrome B	e-110	128	619	1172	2
atm2 B11		At2g18790	phytochrome B	e-100	81	619	1172	3
atm2 D9		At2g18790	phytochrome B	1,00E-70	85	619	1172	1
atm3 A2		At2g18790	phytochrome B	8,00E-53	146	619	1172	2
atm3 H1		At2g18790	phytochrome B	3,00E-72	140	619	1172	2
atm4 D9		At2g18790	phytochrome B	e-110	110	619	1172	2
atm5 H8		At2g18790	phytochrome B	e-100	128	619	1172	2
atm5 B2	1	At4g16250	phytochrome D	5,00E-66	123	623	1164	3
atm3 C6	1	At1g07110	hypothetical protein	7,00E-45	112	656	744	1
atm5 A4	1	At1g15690	hypothetical protein	1,00E-47	2	679	770	2
atm1 A4	1	At5g05170	cellulose synthase catalytic subunit (gb/AAC39336.1)	e-115	2	772	1065	2
atm1 G5	1	At3g26240	hypothetical protein	1,00E-88	3	775	922	3
atm6 G12	1	At3g28770	hypothetical protein	4,00E-09	176	919	2081	2
atm4 G4	1	At3g14570	hypothetical protein	3,00E-48	2	1171	1973	2

Anhang 8.3

Schematische Darstellung der ersten Zeilen der HTML-Tabelle 1 (Proteinexpressions-Unterbibliothek):
http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionCloneSet.shtml

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
sequence name	GenBank accNo	cluster number	cluster size	cluster and Blast comparison	AGIs in cluster	first stopcodon position [aa]	BLAST Hit	putative genefunction	e-value	start query [nt]	start subject [aa]	length subject [aa]	direction	frame
201a01.p1	CK121741	5504	1	+	At3g20530	56	At3g20530	protein kinase, putative	3e-84	245	2	386	forward	2
201a04.p1	CK121740	5503	1	+	At1g67400 At3g43400	-	At1g67400	unknown protein	e-120	118	61	328	forward	1
201a05.p1	CK121742	5505	1	+	At2g33540	145	At2g33540	CTD phosphatase-like 3 (CPL3)	6e-77	19	1104	1241	forward	1
201a07.p1	CK121743	5506	1	+	At1g68530	-	At1g68530	very-long-chain fatty acid condensing enzyme (CUT1)	e-152	4	158	497	forward	1
201a08.p1	CK121744	118	1	+	At1g74920	66	At1g74920	betaine aldehyde dehydrogenase like protein	9e-34	16	442	501	forward	1
201a09.p1	CK121745	5508	1	+	At5g66170	145	At5g66170	senescence-associated protein sen1-like protein	2e-73	25	1	136	forward	1
201a10.p1	CK121746	5509	1	+	At1g02205	218	At1g02205	CER1 protein	e-117	16	414	625	forward	1
201a13.p1	CK121747	2471	3	+	At2g46020	-	At2g46020	putative SNF2 subfamily transcriptional activator	e-147	4	941	1245	forward	1
201a16.p1	CK121748	5511	1	+	At1g71220	101	At1g71220	putative UDP-glucose: glycoprotein glucosyltransferase	2e-50	22	1582	1674	forward	1
201a19.p1	CK121749	5513	1	+	At2g37470	99	At2g37470	putative histone H2B	4e-49	140	31	138	forward	2

Anhang 8.4:

Schematische Darstellung der ersten Zeilen der HTML-Tabelle 2 (Uniklonset):

http://gabi.rz-berlin.mpg.de/projects/Arabidopsis_Proteomics/ArabidopsisExpressionUniCloneSet.shtml

Das Format wurde an DIN A 4 angepaßt, die Inhalte sind identisch.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
unigene name	sequence name	GenBank accNo	cluster number	Blast cluster and comparison	AGIs in cluster	first stopcodon position	Results from BLAST against MIPS annotated protein sequences		e-value	start query	start subject	length subject	Results from BLAST against TIGR5 annotated protein sequences		e-value
							BLAST-Hit	putative genefunction					BLAST-Hit	putative genefunction	
311_A01	A01201a04.p1	CK121740	5503	+	At1g67400	-	At1g67400	unknown protein	e-120	118	61	328	At1g67400	phagocytosis and cell motility protein ELMO1-related (Mus musculus)	0.0
311_A03	A03201a05.p1	CK121742	5505	+	At2g33540	145	At2g33540	CTD phosphatase-like 3 (CPL3)	6e-77	19	1104	1241	At2g33540	CTD phosphatase-like protein 3 (CPL3) identical to thaliana GI:22212705; contains Pfam profile PF03031; NLI interacting factor	0.0
311_A05	A05201a07.p1	CK121743	5506	+	At1g68530	-	At1g68530	very-long-chain fatty acid condensing enzyme (CUT1)	e-152	4	158	497	At1g68530	very-long-chain fatty acid condensing enzyme (CUT1) identical to very-long-chain fatty acid for cuticular wax biosynthesis and pollen fertility; Millar, A. A., et al., Plant Cell (1999)	0.0
311_A07	A07201a09.p1	CK121745	5508	+	At5g66170	145	At5g66170	senescence-associated protein sen1-like protein	2e-73	25	1	136	At5g66170	senescence-associated family protein. contains similarity to ketoconazole-resistant protein GI:928938 and senescence-associated protein GI:1046268 from (Arabidopsis thaliana)	0.0
311_A09	A09201a10.p1	CK121746	5509	+	At1g02205	218	At1g02205	CER1 protein	e-117	16	414	625	At1g02205	CER1 protein identical to maize gl1 homolog (glossy1 locus) GI:1209703 and CER1 GI:1199467 from (Arabidopsis thaliana)	0.0
311_A11	A11201a16.p1	CK121748	5511	+	At1g71220	101	At1g71220	putative UDP-glucose: glycoprotein glucosyltransferase	2e-50	22	1582	1674	At1g71220	UDP-glucose:glycoprotein glucosyltransferase. putative similar to UDP-glucose:glycoprotein glucosyltransferase precursor GB:Q09332 (SPQ09332) from Drosophila melanogaster. (gi:7670746) and (gi:11346464) from Homo sapiens citrate synthase, mitochondrial, putative strong similarity to SP120115 Citrate synthase, mitochondrial precursor (Arabidopsis thaliana); contains Pfam profile PF02085: Citrate synthase	e-160
311_A13	A13201a23.p1	CK121751	5515	+	At2g44350	-	At2g44350	citrate synthase	e-125	19	42	473	At2g44350	SECI14 cytosolic factor family protein / phosphoglycerate transfer family protein similar to polyphosphoinositide binding protein Ssh2p (GI:2739046) (Glycine max); contains Pfam PF00650: CRAL/TRIO domain; contains Pfam PF03765: CRAL/TRIO, N-terminus	0.0
311_A15	A15201a01.p1	CK121752	155	+	At3g51670	235	At3g51670	putative protein	e-128	22	183	409	At3g51670		0.0

Anhang

Anhang 8.5:

Ermittlung der durchschnittlichen Proteinkonzentration von drei 96er Platten mittels Bradford-Test. Ausgehend von den Platten 311-313 wurde jeweils von der Quadranten A2 die Konzentration [$\mu\text{g/ml}$] bestimmt. n.b.= nicht bestimmbar

311_A2 [$\mu\text{g/ml}$]												
	1	2	3	4	5	6	7	8	9	10	11	12
A	98	103	185	50,2	14,3	35,8	22,7	81,2	72,9	68,1	21,5	40,6
B	112	13,1	135	40,6	7,17	172	58,5	226	43	40,6	121	69,3
C	216	227	154	102	141	47,8	141	20,3	88,4	115	57,3	33,5
D	39,4	44,2	14,3	9,56	140	103	160	117	2,39	16,7	200	41,8
E	51,4	117	39,4	98	105	64,5	9,56	198	25,1	50,2	14,3	84,8
F	75,3	28,7	141	63,3	115	47,8	161	178	40,6	93,2	8,36	92
G	20,3	n.b.	1,19	47,8	41,8	51,4	84,8	n.b.	n.b.	n.b.	26,3	n.b.
H	59,7	28,7	1,19	119	23,9	72,9	3,58	34,6	71,7	98	21,5	60,9

312_A2 [$\mu\text{g/ml}$]												
	1	2	3	4	5	6	7	8	9	10	11	12
A	278	51,4	71,7	170	104	254	9,56	118	15,5	172	165	64,5
B	130	129	111	254	71,7	39,4	76,5	225	46,6	168	58,5	172
C	55	208	231	106	226	184	145	148	220	74,1	111	198
D	111	155	170	190	26,3	274	191	71,7	10,8	32,3	50,2	37
E	151	195	82,4	157	161	214	34,6	180	162	137	135	45,4
F	57,3	111	219	166	65,7	62,1	270	127	87,2	n.b.	159	31,1
G	225	174	32,3	35,8	8,36	121	57,3	124	194	162	164	173
H	81,2	128	148	179	133	122	108	124	186	155	137	51,4

313_A2 [$\mu\text{g/ml}$]												
	1	2	3	4	5	6	7	8	9	10	11	12
A	1,19	153	32,3	28,7	168	46,6	87,2	241	179	71,7	293	116
B	200	286	123	213	151	95,6	185	56,2	166	188	200	17,9
C	145	104	33,5	82,4	167	207	104	114	75,3	8,36	68,1	197
D	146	213	n.b.	89,6	177	136	102	56,2	121	19,1	75,3	52,6
E	253	148	43	137	280	17,9	47,8	133	147	8,36	110	103
F	200	254	63,3	88,4	140	165	n.b.	134	164	110	146	96,8
G	49	205	208	135	238	106	57,3	157	60,9	102	96,8	108
H	n.b.	72,9	56,2	145	151	289	33,5	81,2	88,4	76,5	125	59,7

Danksagung

Herrn Prof. Dr. Hans Lehrach danke ich für die Vergabe des Themas, die Möglichkeit der Durchführung der Experimente am Max-Planck-Institut für molekulare Genetik und der Begutachtung der Arbeit.

Herrn Prof. Dr. Thomas Schmülling möchte ich für sein Einverständnis und seiner Mühe danken, diese Arbeit zu begutachten.

Des weiteren möchte ich mich bei Dr. Birgit Kersten, Dr. Jürgen Kreuzberger und Dr. Claus Hultschig für ihr stetes Interesse am Fortgang der Arbeit, sowie für ihre andauernden Hilfsbereitschaft als auch für ihre Diskussionsbereitschaft bedanken. Insbesondere Dr. Birgit Kersten danke ich für ihre Geduld, ihrem Optimismus und ihrem Engagement, welche sehr dazu beigetragen haben, schwierige Phasen zu überwinden. Auch Dr. Harald Seitz möchte ich hier für seine stete Hilfsbereitschaft danken.

Weiterhin danke ich „meiner Bioinformatikerin“ Dr. Svenja Meyer für die Hilfe bei der umfangreichen Zusammenfassung der Ergebnisse.

Ebenso bedanke ich mich bei allen Mitgliedern und ehemaligen Mitgliedern der Protein-Arbeitsgruppen für die gute Zusammenarbeit. Ein herzlicher Dank geht an dieser Stelle an Andrea König, Alexandra Poßling, Silke Wehrmeyer und Armin Kramer.

Weiterhin danke ich Thomas Przewieslik und Thomas Nietzsche für ihren steten Einsatz bei der Arbeit an den Robotern.

Ein besonderer Dank geht auch an das Institut für Pflanzenbiochemie nach Halle an Dr. Justin Lee und Prof. Dr. Dierk Scheel, ohne deren aktiven MAP-Kinasen die Phosphorylierungsstudien nicht möglich gewesen wären.

Für die Bereitstellung sowohl der mRNAs als auch der Entryklone möchte ich mich bei Dr. Richard Immink und Prof. Dr. Gerco Angenent bzw. Dr. Ralf Stracke und Prof. Dr. Bernd Weisshaar bedanken.

Nicht zuletzt möchte ich meinen Freunden und meiner Familie danken, deren Unterstützung, Verständnis und Hilfsbereitschaft weit über diese Arbeit hinausgehen. Annette, Claudia, Jan, Jasmin und Kerstin danke ich für ihre Beiträge zur Fertigstellung der Arbeit und für vieles mehr.