

## 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 GATEWAY<sup>TM</sup>-Klonierungssystems (Invitrogen) konstruiert. Ausgehend vom Entryvektor können die klonierten cDNAs mittels sequenzspezifischer DNA-Rekombination, die durch spezifische Rekombinasen katalysiert wird, in verschiedene GATEWAY<sup>TM</sup>-Expressionsvektoren überführt werden.

Nach der Klonierung der Pistill-cDNAs in den Entryvektor, der Transformation in *E. coli* DH5 $\alpha$ -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 His<sub>6</sub>-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 mediumbefüllten 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<sup>+</sup>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  $\mu$ 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 ( $\mu$ l)	Gesamtvolumen ( $\mu$ l)	Nr.	Fraktionsvolumen ( $\mu$ l)	Gesamtvolumen ( $\mu$ 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 $\alpha$  für Pistill bzw. SCS1/pSE111 für Infloreszenz, wobei die Transformationseffizienz der DH5 $\alpha$ -Zellen bei  $\geq 1 \times 10^8$  koloniebildende Einheiten (KBE)/ $\mu\text{g}$  pUC 18 und der SCS1/pSE111-Zellen bei  $\geq 3,7 \times 10^7$  KBE/ $\mu\text{g}$  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

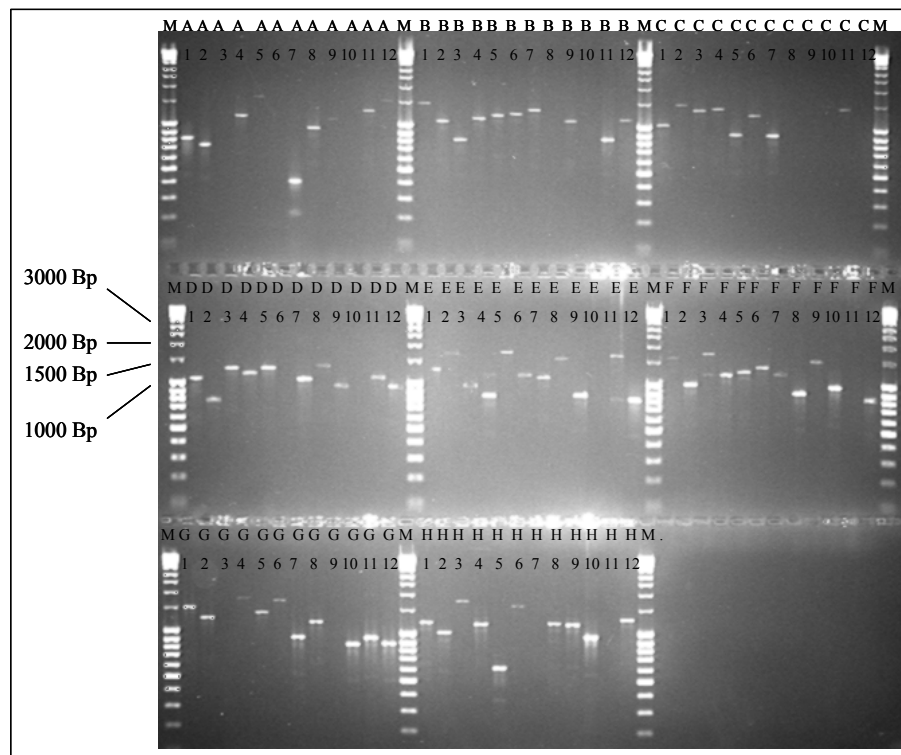
cDNA-Bibliothek hergestellt (Plattenummer 1-103, 106-108), die dann als Grundlage für die Erstellung einer geordneten Expressionsbibliothek (Kap. 3.3), sowie eines Uniklonsets (Kap. 3.4) diene.

## 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.

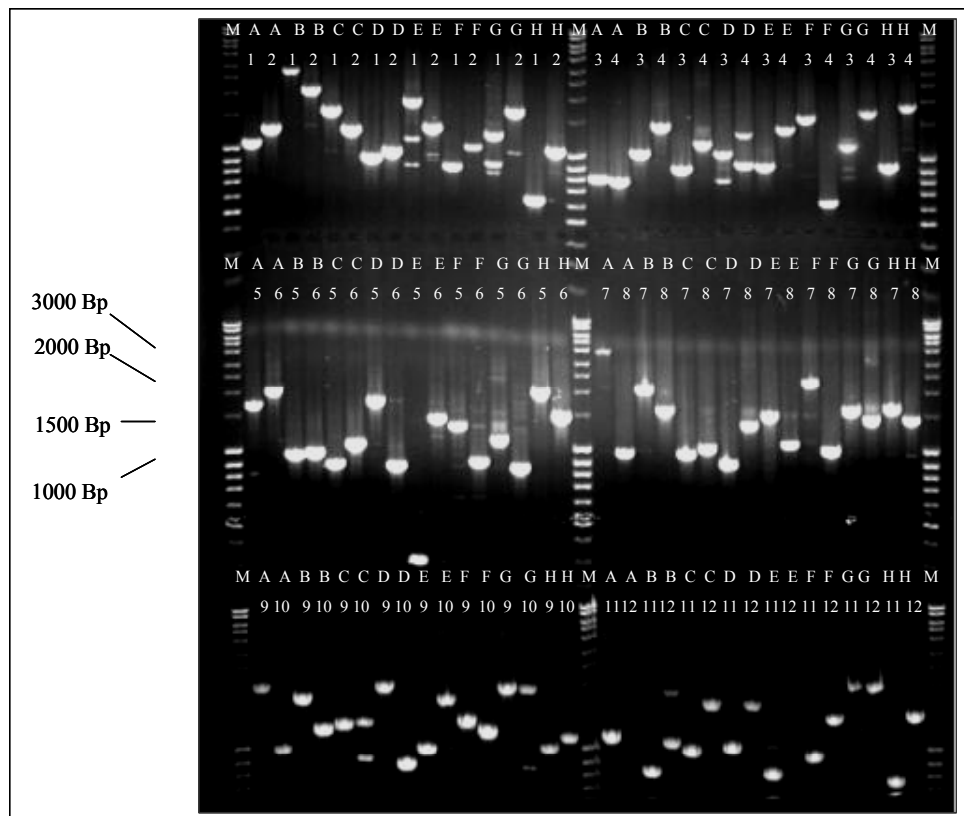
Koordinaten	[Bp]
A3	300
A6	1.400
A10	1.700
B8	1.300
B10	kein Insert
C8	1.300
C9	1.300
C10	1,000
C12	1.700
D6	1.500
D10	1.500
E2	1.200
E10	kein Insert
F11	1.300
G3	1.600
G9	kein Insert
H7	850
H11	1.650

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 His<sub>6</sub>-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

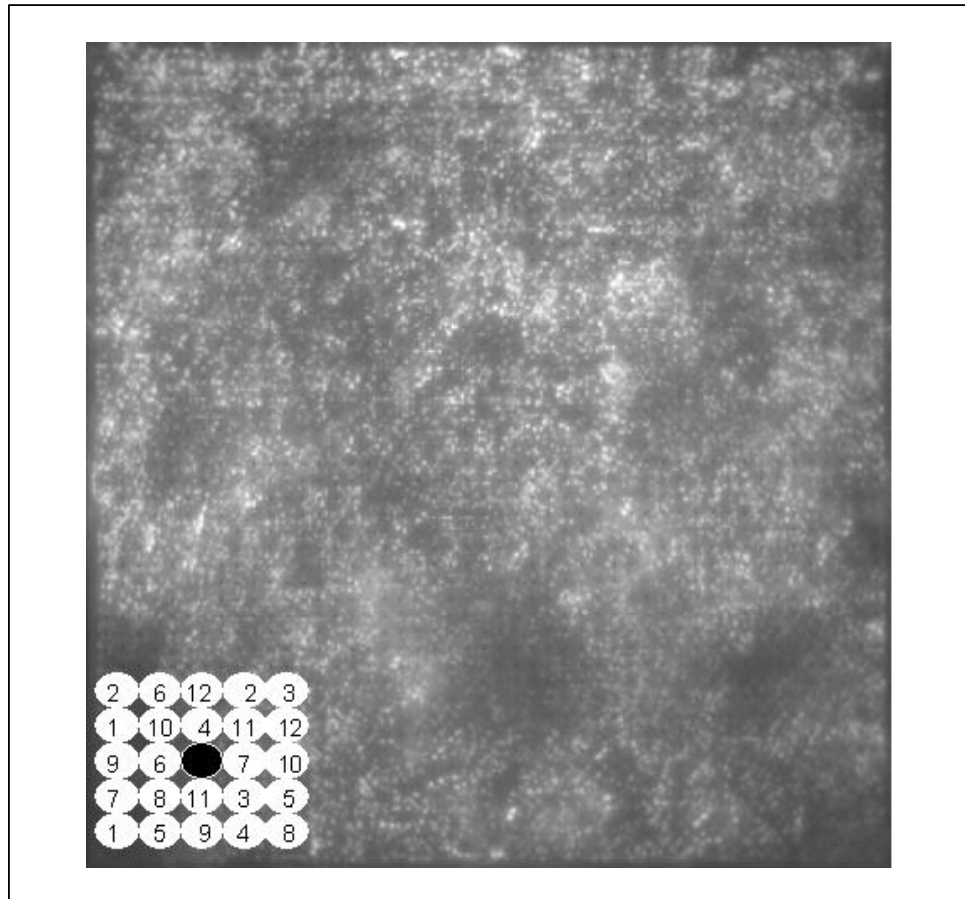
#### 3.3.1 Identifizierung potentieller Expressionsklone der ATM1 Bibliothek mit Hilfe von Hochdichte-Proteinfiltern

Als potentielle Expressionsklone wurden hier Klone bezeichnet, die mit einem anti-RGS His<sub>6</sub>-Antikörper detektiert werden konnten. Klone, deren Inserts im Leserahmen zum RGS His<sub>6</sub>-Tag lagen, lieferten ein getagtes Protein und damit ein Signal mit dem tagspezifischen Antikörper. Klone, deren Inserts nicht im Leserahmen zum RGS His<sub>6</sub>-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 His<sub>6</sub>-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 His<sub>6</sub>) 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 His<sub>6</sub> positiven Klone auf dem *Image* erfolgte anschließend unter Verwendung des Programms „Visual Grid 2.01“ (vgl. 2.2.2). Anti-RGS His<sub>6</sub> 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 His<sub>6</sub>-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

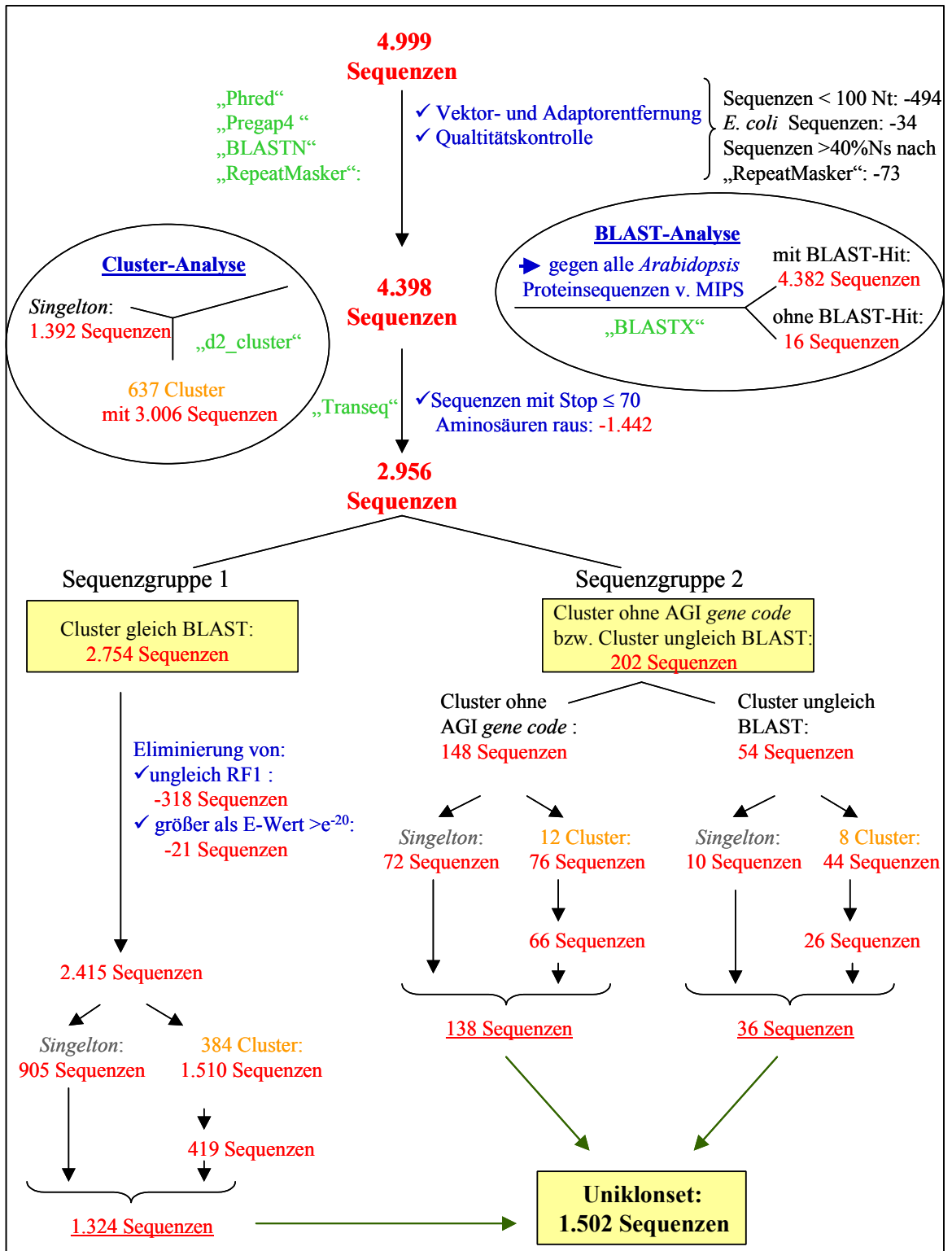
#### 3.4.1 Sequenzierung

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 His<sub>6</sub>-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](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](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
<b>Gesamt</b>	<b>637</b>	<b>3.006</b>

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  $\leq 70$  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 His<sub>6</sub>-Tag befanden. Dabei entspricht der Leserahmen 1 dem Leserahmen des RGS His<sub>6</sub>-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.

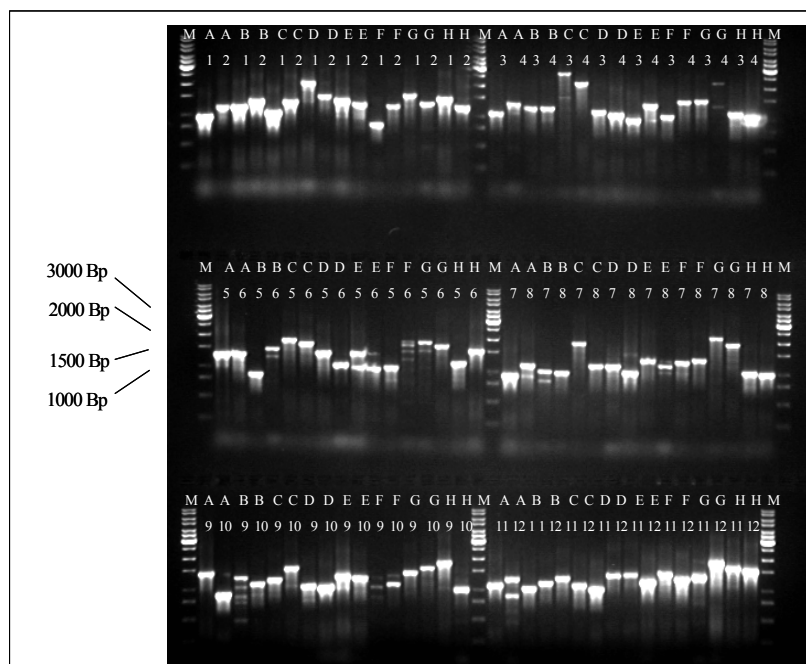
Die 12 Cluster der 76 Sequenzen ohne Cluster-AGI *gene code* setzten sich aus 10 Clustern mit weniger als 6 Sequenzen und 2 Clustern mit mehr als 6 Sequenzen zusammen. Von den ersten Clustern ( $< 6$  Sequenzen) gingen alle 23 Klone und von den zweiten Clustern ( $> 6$  Sequenzen) gingen von 53 Sequenzen 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 (Plattenummer: 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; Plattenummer: 315).



**Abb. 3.5:** Ermittlung der Insertlängen der 96 vollständige cDNA-Transkriptionsfaktoren. 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.

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.

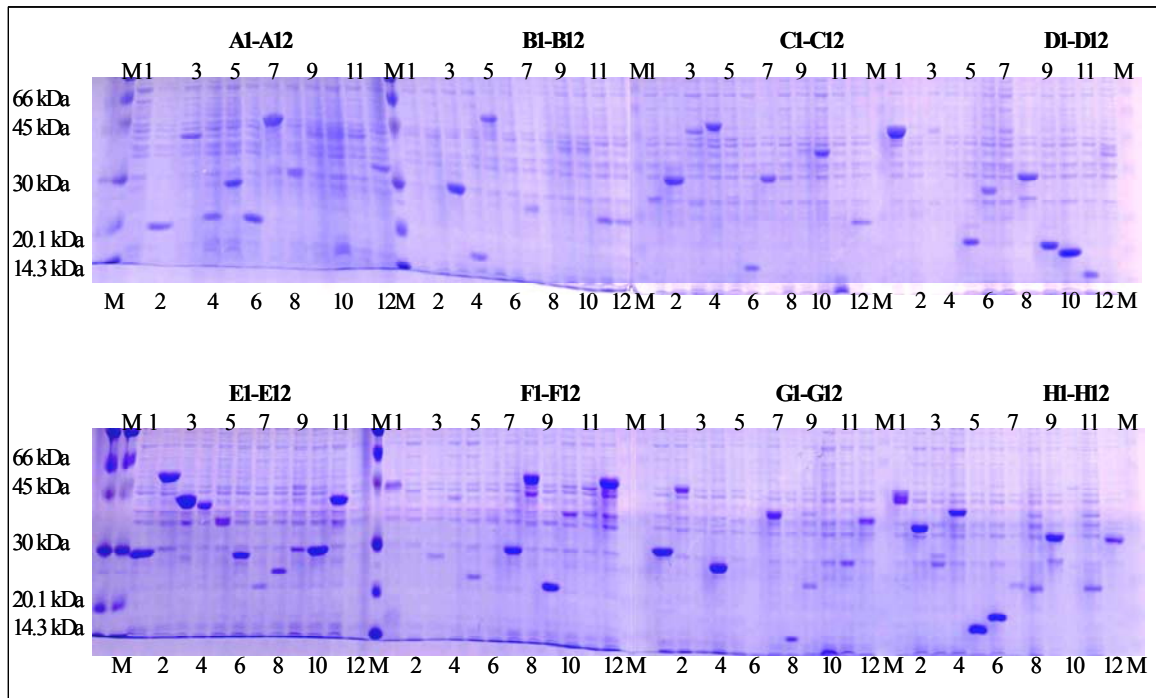
Somit lagen zusätzlich zu den 1.502 Uniklonset-Klonen 192 vollständige cDNA-Klone vor. Diese nunmehr insgesamt 1.694 Klone bildeten zusammen das Erweiterte Uniklonset, das im Weiteren verwendet wurde.

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**

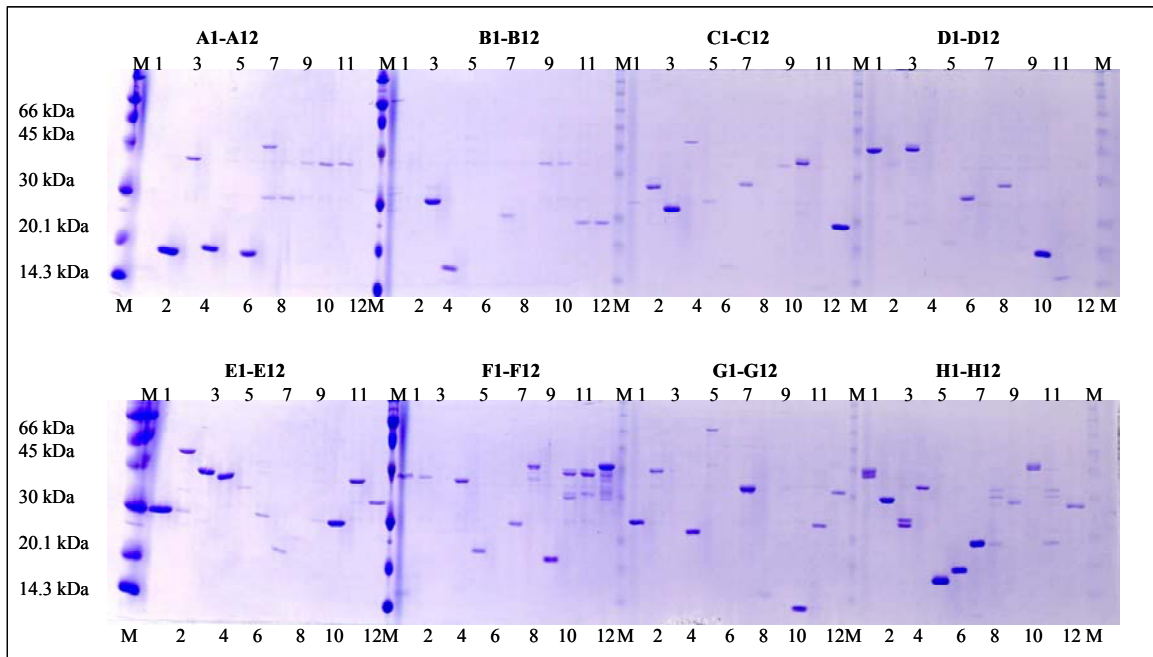
Um eine erste Analyse des neu erstellten Uniklonsets durchzuführen, wurden 96 Klone des Uniklonsets der Mikrotiterplatte 311 Quadrant A1 exprimiert und im 96 *well* Format aufgereinigt (vgl. 2.2.5.4). Dazu wurde zunächst ein Teil der Zellpellets nach der Expression mittels SDS-Polyacrylamid-Elektrophorese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

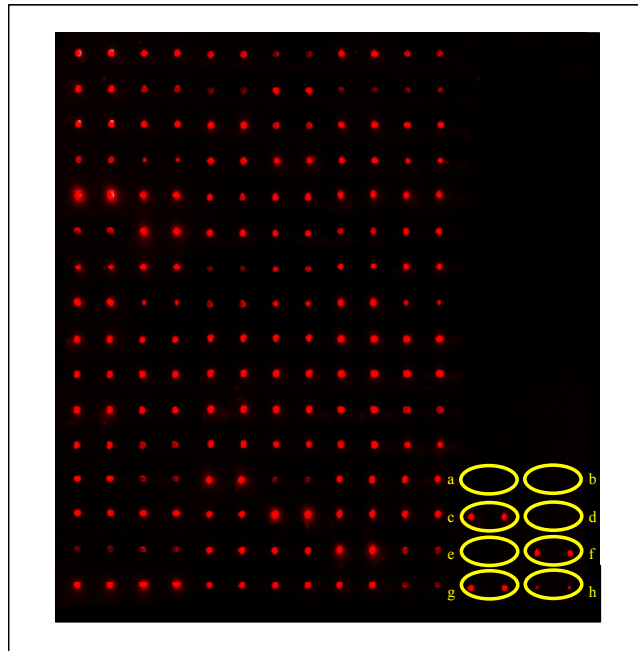
Anschließend wurden die Zellpellets im 96 *well* Format mittels NiNTA-Agaroselösung über Filterplatten unter denaturierenden Bedingungen aufgereinigt. Ein Teil der Eluate, welche die gereinigten Proteine enthalten, wurde auf SDS-Polyacrylamidgele aufgetragen. Ein Gel wurde direkt nach der Elektrophorese 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 His<sub>6</sub>-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  $\mu\text{m}$ ) gespottet (Abb.3.8). Der Microarray wurde zunächst mit einem anti-RGS His<sub>6</sub>-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 His<sub>6</sub>-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 His<sub>6</sub>-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 His<sub>6</sub>-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 FAST<sup>TM</sup>-Slides gespottet und für Phosphorylierungsversuche eingesetzt.

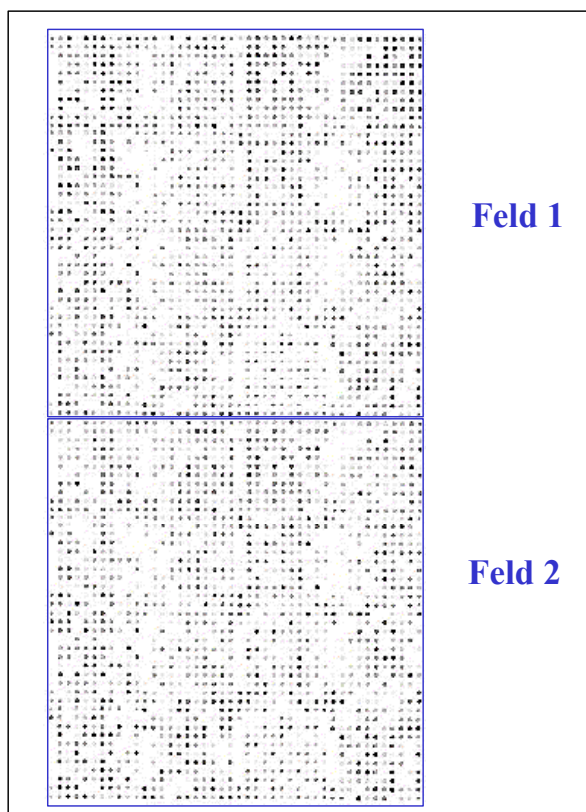
### 3.4.6 Bradford-Test

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 His<sub>6</sub>-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 His<sub>6</sub>-Antikörper.** Die Proteine sind jeweils in zwei identischen Feldern gespottet worden. Der Nachweis rekombinanter Proteine erfolgte mit Hilfe eines anti-RGS His<sub>6</sub>-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 His<sub>6</sub>-Antikörper und Kaninchen anti-Maus IgG3-Cy3) keine Signale detektiert. Damit wurde gezeigt, dass keine unspezifische Bindung des Zweitantikörpers an die rekombinanten Proteine erfolgt.

Im Fluoreszenz-*Image*, das nach der Inkubation mit dem anti-RGS His<sub>6</sub>-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 gewonnenen 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 rekombinanten Gerste-Proteinen (Kramer *et al.*, 2004). Diese *proof-of-principle* Studie wurde mit 768 gespotteten Proteinen durchgeführt, wobei potentielle Targets der Casein Kinase2 $\alpha$  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

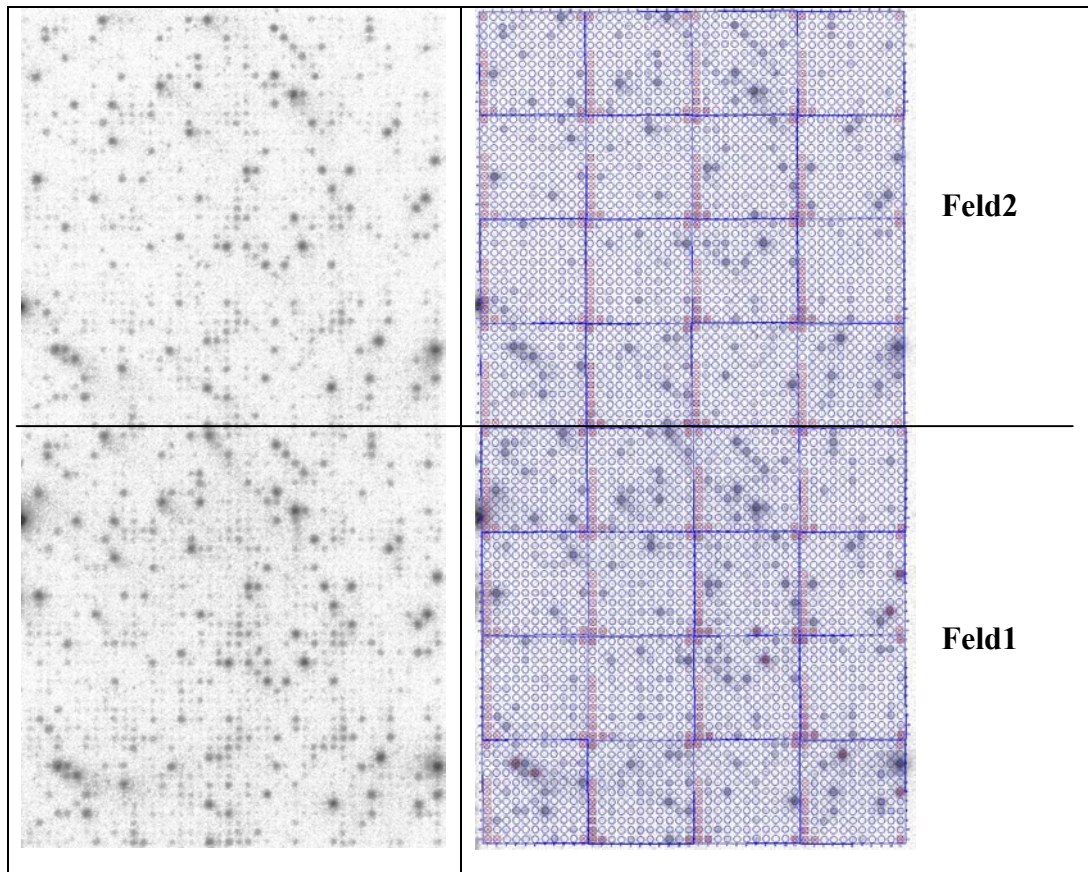
Zunächst wurde anhand von Phosphorylierungsstudien mit der Proteinkinase A die maximale Spotdichte ermittelt, bei der bei einem Phosphorylierungs-Assay einzelne Signale noch unterscheidbar sind.

Das Erweiterte Uniklonset wurde zu diesen Zwecke in einem 14 x 14-Pattern (Spotabstand: 321  $\mu\text{m}$ ) und in einem 15 x 15-Pattern (Spotabstand: 300  $\mu\text{m}$ ) gespottet. Bei beiden Spotdichten war eine Detektion der Proteine unter Verwendung des anti-RGS His<sub>6</sub>-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  $\mu\text{m}$ ) 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

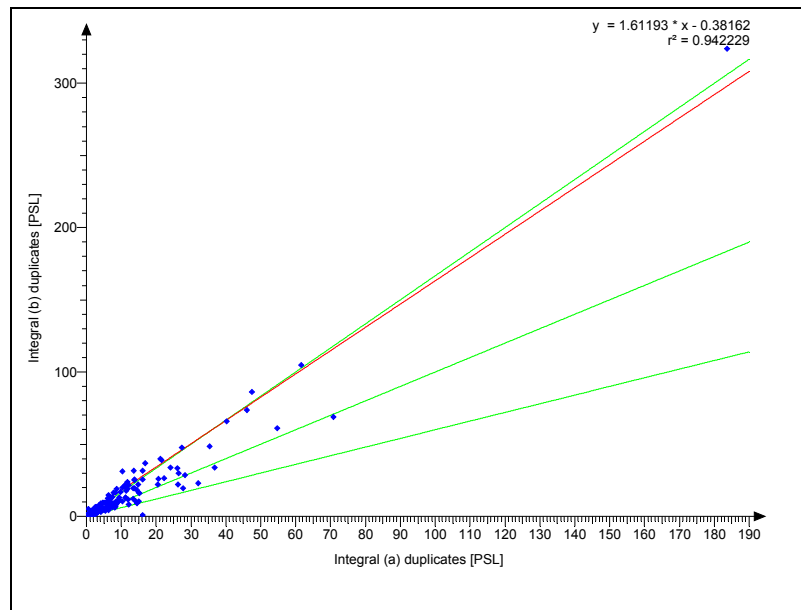
Diese Versuche zur Etablierung der Methode erfolgten ausgehend von je zwei Microarrays aus unabhängigen Versuchsreihen. Zu jedem Phosphorylierungs-Versuch (siehe auch Kap. 0) wurde generell ein Kontroll-Microarray mitgeführt. Bei den Kontrollen enthielt der Phosphorylierungsansatz  $\gamma$ -ATP (<sup>33</sup>P), 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/ $\mu\text{l}$ , versetzt mit Maus anti-RGS His<sub>6</sub>-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.

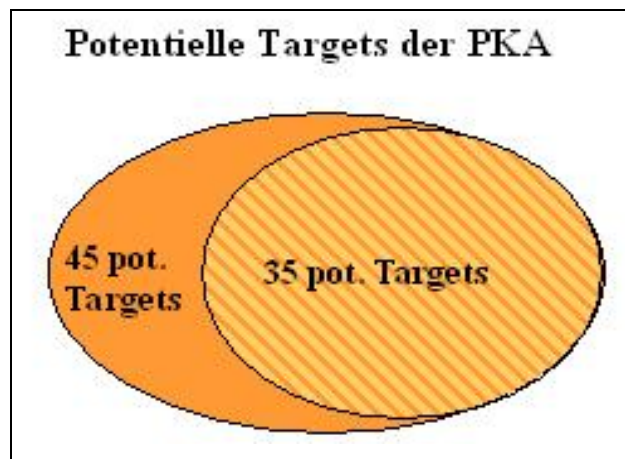
Die Datenpunkte, die mehr als 25% von ihrem Mittelwert abwichen, wurden manuell markiert und von der weiteren Auswertung ausgeschlossen. Des weiteren wurden Datenpunkte, die von „Aida Array Matrix“ automatisch als nicht auswertbar markiert wurden, manuell kontrolliert und gegebenenfalls ihr Ausschluss 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.

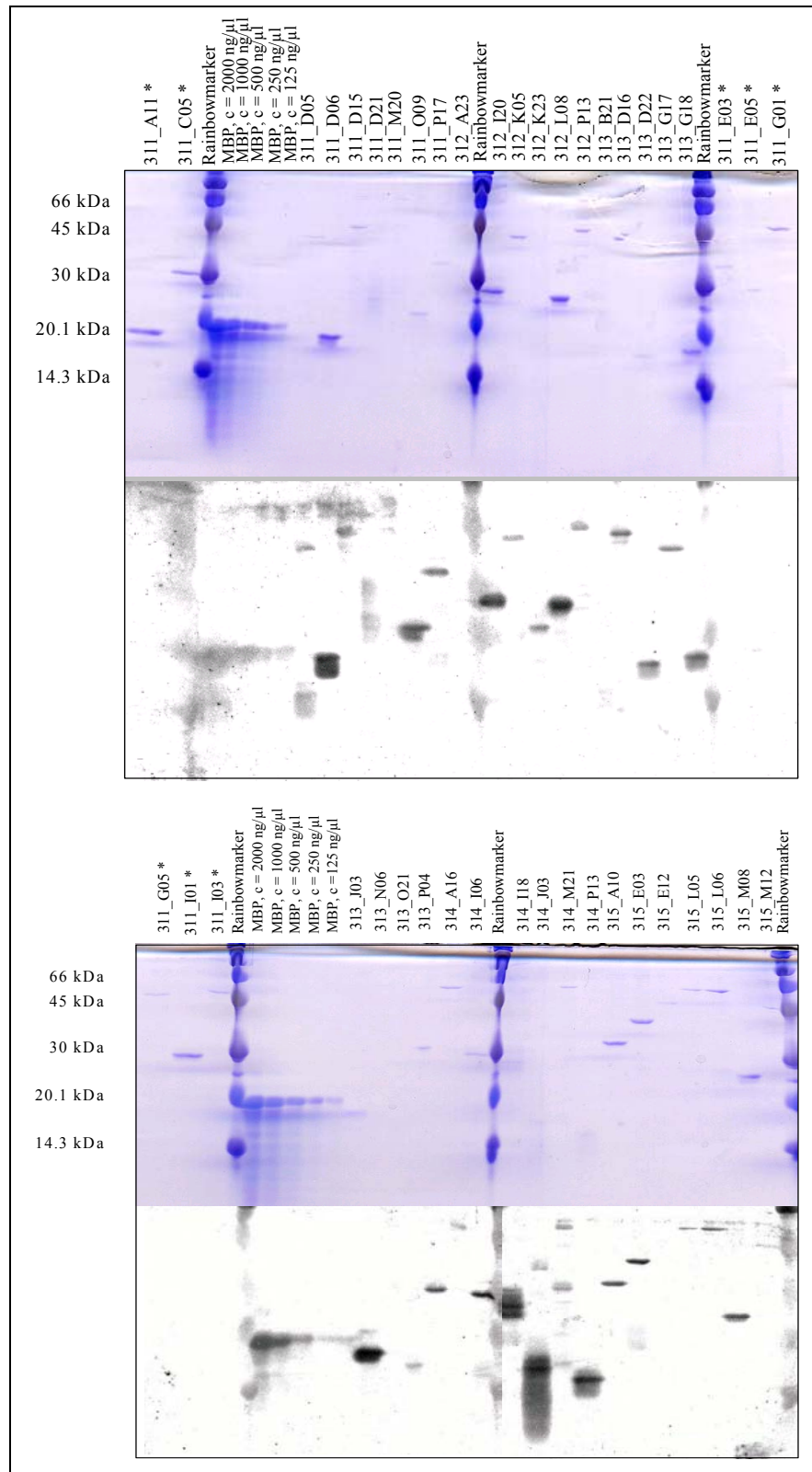
Wie aus Abb. 3.12 ersichtlich wird, konnten ausgehend von den gespotteten Proteinen des Erweiterten Unikonsets 35 potentielle Phosphorylierungs-Targets (PKA-Set) detektiert werden. Diese sind in Tab. 3.5 zusammengefasst.

**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 <sup>+</sup> -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%

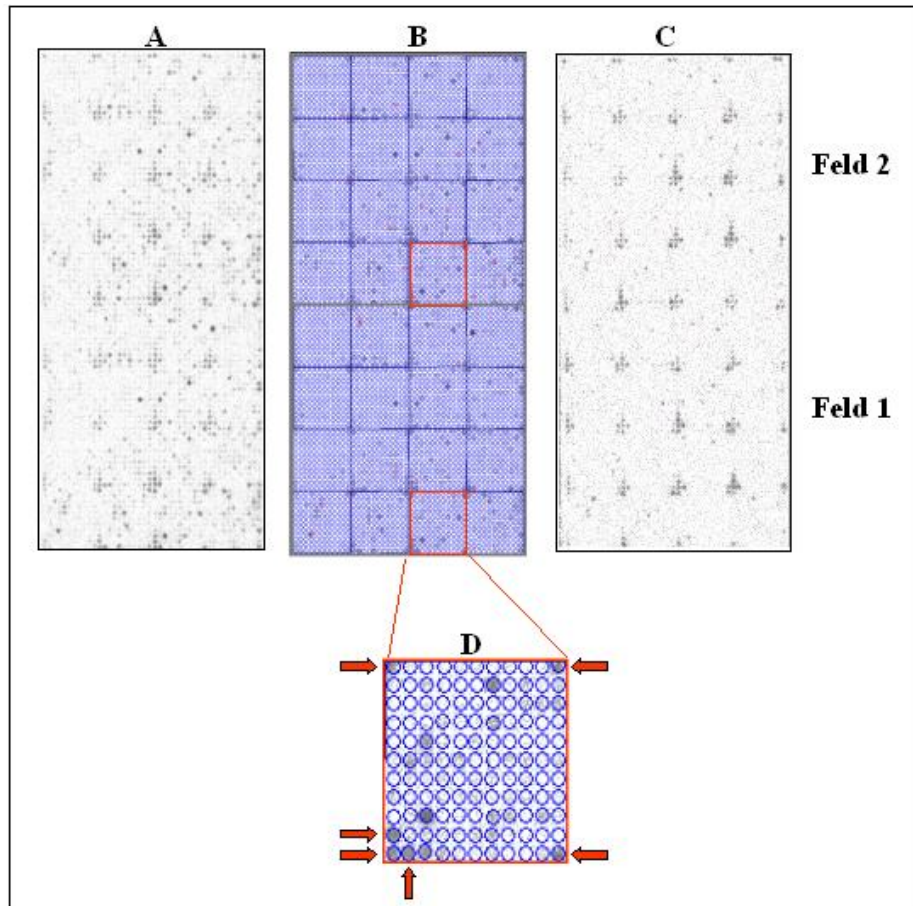
In der Tabelle sind die Ergebnisse des Vergleichs der Serin- und Threoningehalte des PKA-Sets und des Uniklonsets dargestellt. PKA = Proteinkinase A.

### 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.

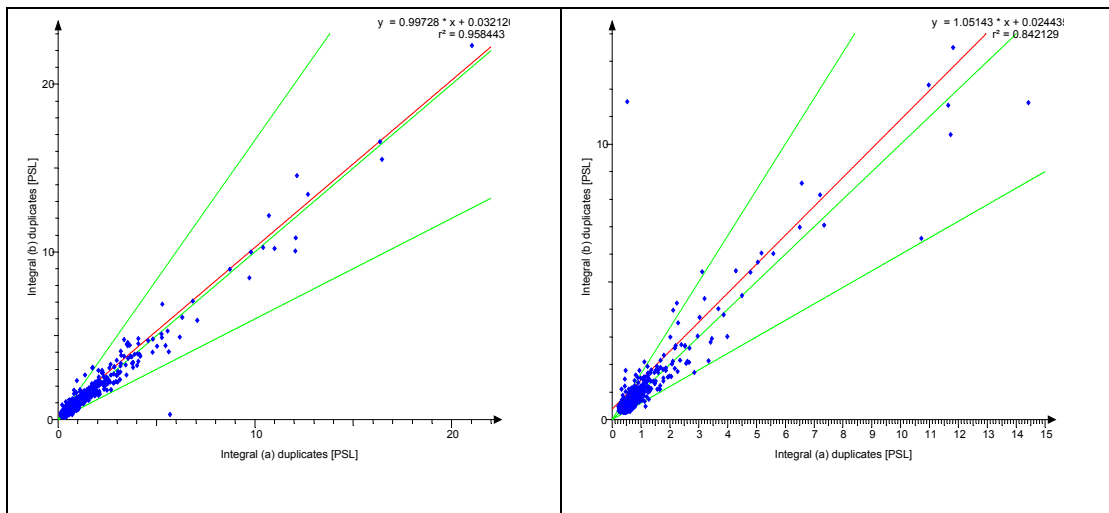
Jeder Kinase-Versuch wurde auf jeweils zwei Microarrays in zwei unabhängigen Versuchsreihen durchgeführt. Anhand von Vorversuchen wurde eine geeignete Konzentration der einzusetzenden MPKs ermittelt. Dabei zeigte sich, dass die MPK3 (Konzentration: 100 ng/μl) in geringeren Konzentrationen eingesetzt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





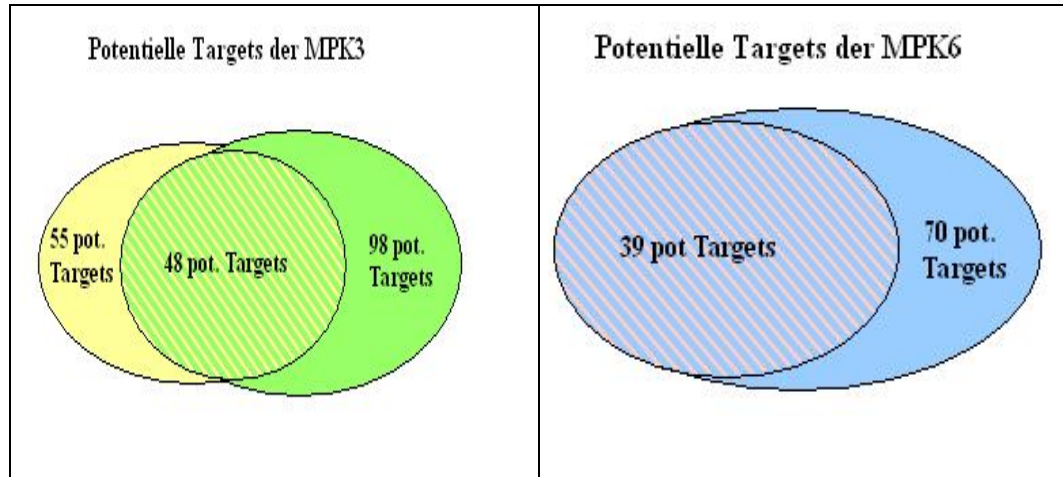
**Abb. 3.14: Image der eingescannten Microarrays mit den gespotteten Proteinen des Erweiterten Unikonsets nach der Phosphorylierung durch die MPK3 bzw. MPK6.** Die Proteine wurden in zwei identischen Feldern auf den Microarrays angeordnet. Die Detektion erfolgte mit einer Imaging Plate, die nach der Exposition (MPK3: 6 h und MPK6: 48 h) ausgelesen wurde (vgl. 2.2.6.5). Darstellung mit einem dynamischen 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/ $\mu$ l, versetzt mit anti-RGS His<sub>6</sub>-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

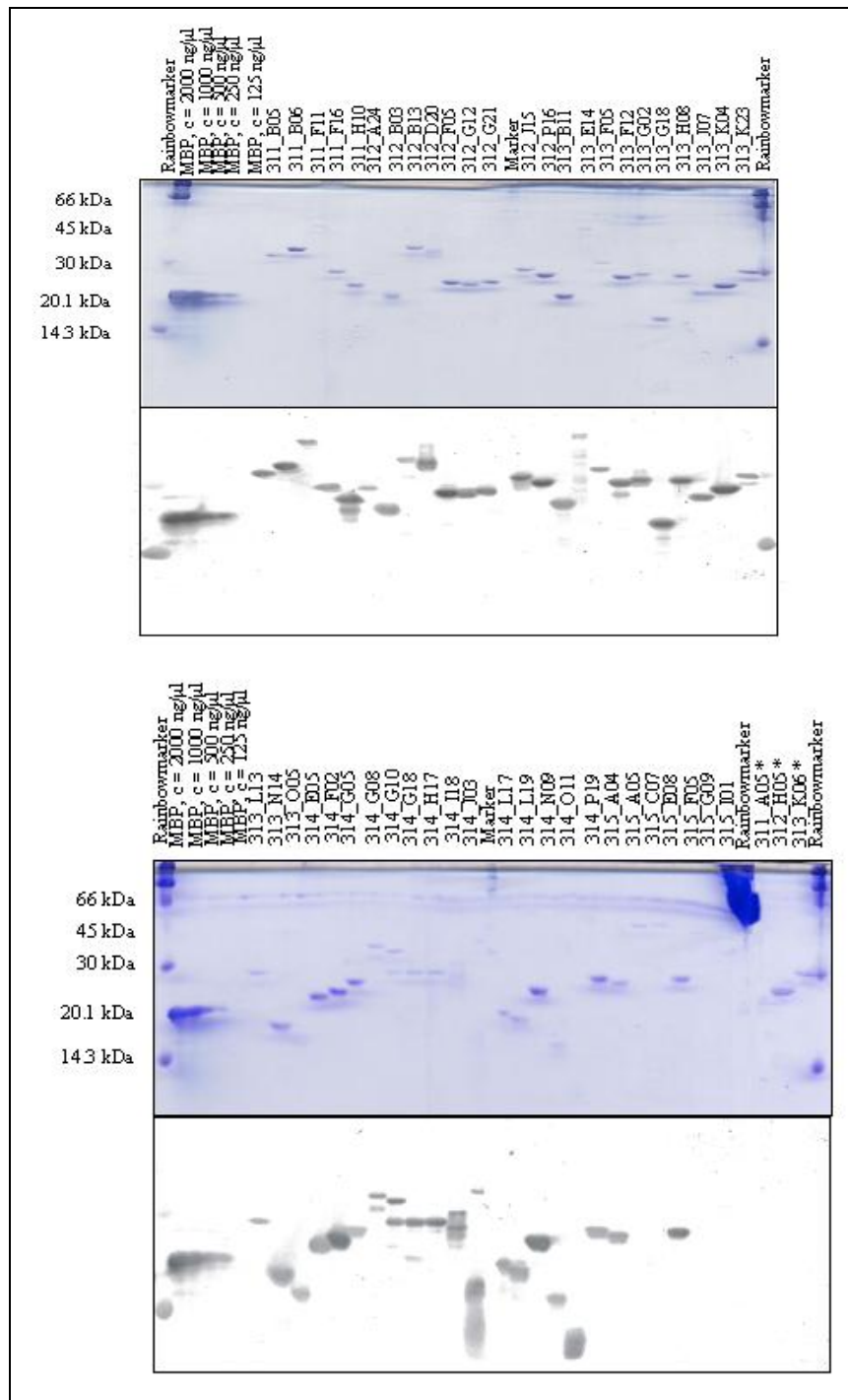
In der Tabelle sind die potentielle Phosphorylierungs-Targets des MPK3-Sets dargestellt. MPK3 = Mitogen-aktivierten Proteinkinase 3.

**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	cucumisin-like serine protease	cucumisin-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/ $\mu$ l, 1.000 ng/ $\mu$ l, 500 ng/ $\mu$ l, 250 ng/ $\mu$ l, 125 ng/ $\mu$ 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).

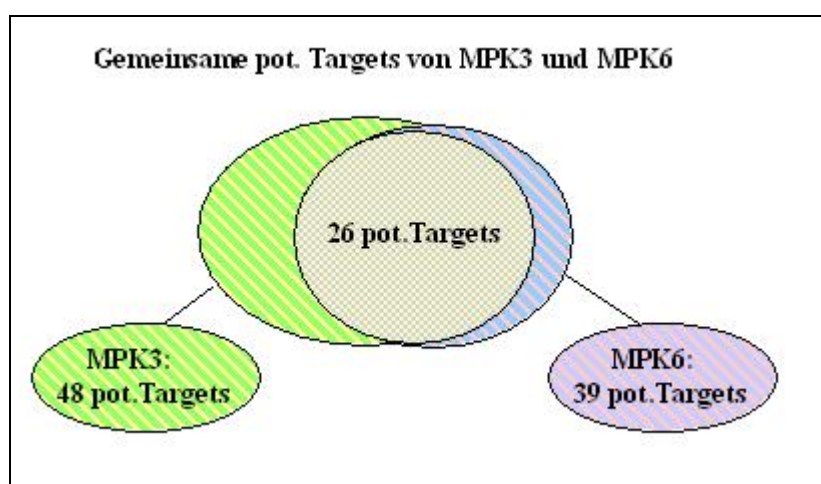


Abb. 3.18: Schematische Darstellung zur Ermittlung der gemeinsamen potentiellen Targets der MPK3 und der MPK6. Pot. = potentiell.

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.

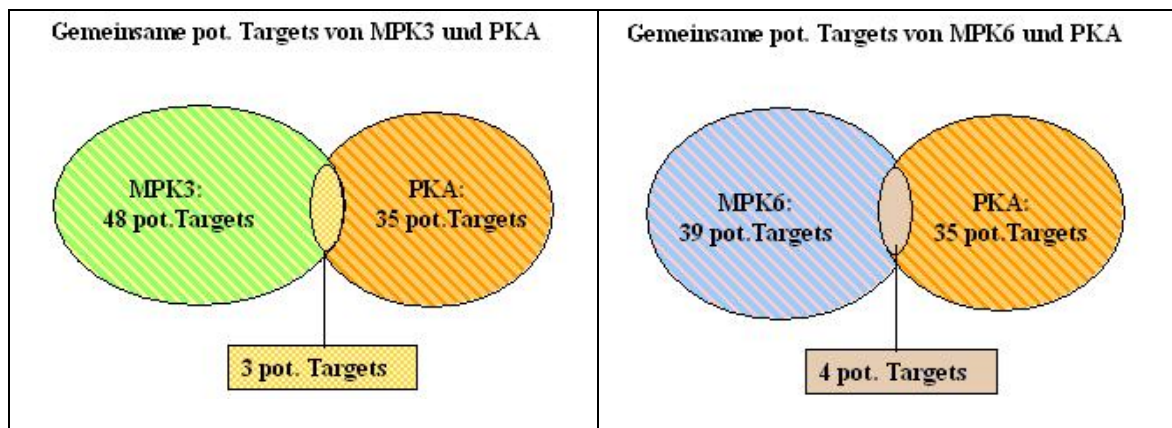


**Tab. 3.12: Potentielle Targets, die spezifisch für MPK6 sind.**

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.