

Veröffentlichung der Ergebnisse von Forschungsvorhaben im BMBF-Programm

Biologie

"Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 0312280 A gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor".

Forschungsvorhaben:
**Vergleichende Sequenzierung zweier Regionen im Gersten- und Reisgenom
(GABI-MAP)**

Förderkennzeichen:
0312280 A

Zuwendungsempfänger:
Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK), Corrensstr. 3, 06466 Gatersleben

Projektleiter:
Dr. Marion Röder, Prof. Andreas Graner

Laufzeit:
01. 03. 2000 – 30. 06. 2003

I.1. Aufgabenstellung:

Das Arbeitsprogramm umfaßte den Aufbau von BAC-Contigs für zwei Regionen des Gersten- und Reisgenoms zur Untersuchung der Mikrosyntanie zwischen dem großen Gerstengenom (ca. 5 pg) und dem etwa 10-mal kleineren Reisgenom. In einem Fall handelte es sich um den Bereich des *Rh2*-Resistenzgenes auf Gerstenchromosom 7HS, welcher Resistenz gegen das pilzliche Pathogen *Rhynchosporium secalis* vermittelt, und den syntänen Bereich auf Chromosom 6 des Reisgenomes. Die zweite Region enthält den *rym4/rym5* Locus für Resistenz gegen Barley Mild Mosaic (BaMMV) bzw. Barley Yellow Mosaic Virus (BaYMV) auf Chromosom 3HL der Gerste und die entsprechende syntäne Region auf Reischromosom 1. Für jede Region wurden BAC-Contigs aufgebaut und anschließend etwa 450 kb des Gerstengenoms sequenziert und mit der entsprechenden Region des Reisgenoms auf DNA-Sequenzebene verglichen. Mit diesem Sequenzvergleich sollte die Anzahl der Gene pro kb und die Konservierung der Genabfolge untersucht werden. Ferner wurden Informationen über Sequenzähnlichkeiten im Bereich von Resistenzgenen und über die Struktur der intergenischen Bereiche zwischen Reis und Gerste gewonnen. Ergebnisse dieser Arbeiten beantworten die Frage inwieweit die Reis/Gersten-Syntanie für die Isolation von weiteren Resistenzgenen aus der Gerste genutzt werden kann und damit die umfangreichen internationalen genomischen Ressourcen der Modellpflanze Reis für Entwicklungen und Genisolierungen in der Gerste einsetzbar sind.

I.2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde:

Am Pflanzengenom-Ressourcen Zentrum des IPK stand eine Gersten-BAC-Bibliothek der Sorte Morex zur Verfügung, welche in den USA entwickelt wurde (Yu et al. 2000). Diese Bibliothek hat einen Umfang von über 310.000 Einzelklonen mit einer durchschnittlichen Insertgröße von 106 kb. Daraus errechnet sich eine theoretisch mehr als 6-fache Repräsentation des haploiden Gerstengenoms. Für die Bibliothek standen hochdichte Koloniefilter für die Southern Hybridisierung sowie DNA-Pools für ein PCR-gestütztes Screening zur Verfügung. Für die Feinkartierung der beiden Resistenzgene standen hochauflösende Kartierungspopulationen zur Verfügung, welche im Laufe des Projektes noch erweitert wurden (vgl. I.4 und II.1).

Aus dem japanischen Reisgenomprojekt standen öffentliche Informationen über die Sequenzen der beiden syntänen Regionen aus Reis zur Verfügung (<http://rgp.dna.affrc.go.jp>).

Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Brueggeman, R.S., Muehlbauer, G.J., Wise, R.P., Wing, R.A.: A bacterial artificial chromosome library for barley (*Hordeum vulgare* L.) and the identification of clones containing putative resistance genes. *Theor Appl Genet* 101: 1093-1099 (2000).

I.3. Planung und Ablauf des Vorhabens:

Das Vorhaben gliederte sich in zwei unabhängige Teilprojekte, welche parallel in zwei verschiedenen Arbeitsgruppen durchgeführt wurden. Teilprojekt A betrifft die Sequenzierung im Bereich des *Rh2*-Resistenzgenes auf Gerstenchromosom 7H und wurde in der Gruppe Gen- und Genomkartierung unter Leitung von Herrn Dr. M. Ganai und nach dessen Ausscheiden vom IPK von Frau Dr. M. Röder durchgeführt. Teilprojekt B betrifft die Sequenzierung im Bereich des *rym4/rym5* Resistenzlocus auf Gerstenchromosom 3H und wurde in der AG Molekulare Marker unter Leitung von Herrn Prof. A. Graner und Dr. N. Stein durchgeführt.

Aufgrund der gegenüber der ursprünglichen Antragstellung um 50% gesunkenen Sequenzierkosten konnte im Laufe der Projektarbeiten die doppelte Anzahl an BACs sequenziert werden. Insgesamt wurden somit anstatt von 3 BACs pro Teilprojekt 6 BACs pro Teilprojekt sequenziert.

I.4. Wissenschaftlicher und technischer Stand zu Beginn des Projektes

Durch vergleichende Kartierung von RFLP-Sonden war bekannt, dass Reischromosomen in großen chromosomalen Bereichen eine konservierte Markieranordnung zu anderen Gramineen wie Mais oder Weizen besitzen (Ahn et al. 1993). Dies führte zu der Annahme, dass sich die Chromosomen der verschiedenen Gramineen-Arten auf ein Urgenom zurückführen lassen, welches durch Translokationen und andere genomische Rearrangements zur Herausbildung der heutigen Arten führte. Von praktischer Bedeutung ist dabei, dass Kenntnisse über die Kartierungsposition einer Markersonde in einer Spezies oft auf andere Gramineen-Spezies übertragbar sind (Moore et al. 1995; Gale & Devos 1998). Neben dieser sogenannten Makrosyntenie zeigten vergleichende Sequenzierungen von

syntänen genomischen Regionen in verschiedenen Gramineen-Arten oft auch eine Mikrokolinearität auf der Ebene einzelner Gene und deren Anordnung im Genom (Chen et al. 1997; Bennetzen 2000; Dubcovsky et al. 2001). Diese Makrokollinearität fehlt in Genregionen um sogenannte Resistenzgenanaloga, die als schnell evolvierende Loci in Zusammenhang mit Genomumstrukturierung betrachtet werden (Leister et al. 1998). Hierbei ist unklar, inwieweit die Ursache in genomischen Rearrangements im Bereich der Resistenzgene oder aber in der divergenten Evolution derselben zu suchen ist. Ziel dieses Projektes war es daher zu untersuchen, inwieweit interspezifische Mikrolinearität auch im Bereich von Resistenzgenen beobachtet werden kann.

Relevante Publikationen:

Ahn, S., Anderson, J.A., Sorrells, M.E., Tanksley, S.D.: Homoeologous relationships of rice, wheat and maize chromosomes. *Mol. Gen. Genet.* 241: 483-490 (1993)

Bennetzen, J.L.: Comparative sequence analysis of plant nuclear genomes: microlinearity and its many exceptions. *Plant Cell* 12: 1021-1029 (2000)

Chen, M., Saniguel, P., Oliviera, A.C., Woo, S.-S., Zhang, H., Wing, R.A., Bennetzen, J.L.: Microlinearity in the sh2-homologous regions of the maize, rice and Sorghum genomes. *Proc. Natl. Acad. Sci. USA* 94: 3431-3435 (1997)

Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L., Shiloff, B.A., Bennetzen, J.L.: Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiol.* 125: 1342-1353 (2001)

Gale, M.D., Devos, K.M.: Comparative genetics in the grasses. *Proc. Natl. Acad. Sci. USA* 95: 1971-1974 (1998)

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A., Schulze-Lefert, P.: Rapid organization of resistance gene homologues in cereal genomes. *Proc. Natl. Acad. Sci. USA* 95: 370-375 (1998)

Moore, G., Devos, K.M., Wang, Z., Gale, M.D.: Grasses, line up and form a circle. *Curr. Biol.* 5: 737 (1995)

Projektteil A: Sequenzierung der Region um das *Rh2*- Resistenzgen auf Chromosom 7H

Aus früheren Kartierungsarbeiten war bekannt, dass das *Rh2*-Resistenzgen gegen den Erreger der Blattfleckenkrankheit bei Gerste *Rhynchosporium secalis* auf dem kurzen Arm von Gerstenchromosom 7H (1 nach alter Nomenklatur) nahe dem Marker MWG555 kartiert (Schweizer et al. 1995). Das *Rh2*-Gen stammte dabei aus der Gerstensorte Atlas. Bei einer genaueren Kartierung von 741 F2-Pflanzen (= 1482 Gameten) der Kreuzung Atlas x Steffi konnte das *Rh2*-Gen in ein Intervall von ca. 1 cM zwischen den Markern MWG2018 und MWG555 kartiert werden (Schmidt et al. 2001). Aus publizierten Arbeiten von Kilian et al. 1995; 1997 war bekannt, dass diese Region des Gerstenchromosoms 7H Homologien zu Reischromosom 6 besitzt. Die Ausnutzung dieser Informationen aus Reis erlaubte es homologe Hybridisierungssonden für Gerste zu entwickeln, welche für den Aufbau des Gersten-BAC Contigs benutzt werden konnten.

Relevante Publikationen:

Kilian, A., Kudrna, D.A., Kleinhofs, A., Yano, M., Kurata, N., Steffenson, B., Sasaki, T.: Rice-barley synteny and its application to saturation mapping of the barley *Rpg1* region. *Nucl. Acids Res.* 23: 2729-2733 (1995)

Kilian, A., Chen, J., Han, F., Steffenson, B., Kleinhofs, A.: Towards map-based cloning of the barley stem rust resistance genes *Rpg1* and *rpg4* using rice as an intergenomic cloning vehicle. *Plant Mol. Biol.* 35: 187-195 (1997)

Schmidt, D., Röder, M.S., Dargatz, H., Wolf, N., Schweizer, G.F., Tekauz, A., Ganai, M.W.: Construction of a YAC library from barley cultivar Franka and identification of YAC-derived markers linked to the *Rh2* gene conferring resistance to scald (*Rhynchosporium secalis*). Genome 44: 1031-1040 (2001)

Schweizer, G.F., Baumer, M., Daniel, G., Rugel, H., Röder, M.S.: RFLP markers linked to scald (*Rhynchosporium secalis*) resistance gene *Rh2* in barley. Theor Appl Genet 90: 920-924 (1995)

Projektteil B: Sequenzierung der Region um den *rym4/rym5* Resistenzlocus auf Chromosom 3H

Resistenz gegenüber den Bymo-Viren BaMMV und BaYMV wird in Gerste durch verschiedene rezessive Loci kontrolliert, die über das Gerstegenom verstreut sind. Ein Locus, *rym4/rym5* wurde in einem ca. 2 cM großen Intervall zwischen den Markern MWG010 und MWG838 in zwei Population aus Igri (anfällig) x Franka (resistent, *rym4*) bzw. Igri x W122 (basiert auf Resistant Ym No.1, resistent, *rym5*) auf Gerste Chromosom 3HL kartiert (Graner und Bauer 1993, Graner et al. 1999). *Rym4* vermittelt gegenüber zwei, *rym5* gegenüber drei Virusstämmen Resistenz. Es ist ungeklärt, ob es sich bei diesem Locus um zwei unabhängige Resistenzgene oder um Allele desselben Gens handelt, da bisher, in umfangreichen Resistenztests, keine Rekombination zwischen beiden Resistenzspezifitäten gefunden werden konnte (Graner et al. 1999, Dr. Frank Ordon, persönliche Mitteilung). Vor Beginn des Projektes wurden zwei Reismarker (RZ783 und C112) flankierend zu MWG010 als kolinear zwischen Gerste und Reis identifiziert (Abb. 1, *rym4*) (Streng 2000 und eigene unveröffentlichte Daten). Studien zur Syntanie des Reis- und Gerstegenoms ergaben bereits zuvor, dass Gerste Chromosom 3H (bzw. das Consensus-Triticeae Chromosom 3) und Reis Chromosom 1 zu den am besten, über die gesamte Chromosomenlänge, konservierten Kopplungsgruppen zwischen beiden Spezies gehören (Moore et al. 1995, Smilde et al. 2001). Diese Erkenntnis konnte durch die umfangreiche Kartierung EST-basierter Marker inzwischen weiter verdichtet werden (Sorrells et al. 2003, Prasad et al., Manuskript in Vorbereitung, GABI-PLANT 0312271A). Es bestand die Einschätzung, dass der Resistenzlocus *rym4/5* sich in einem kolinearen Genomabschnitt zwischen Gerste und Reis befinden sollte. Vor Beginn des Projektes existierten unabhängige BAC bzw. BAC/YAC Contigs in relativ großem flankierenden Abstand zum Locus *rym4/5*, deren Einzelklone für die vergleichende Sequenzierung eingesetzt werden konnten (Abb. 1).

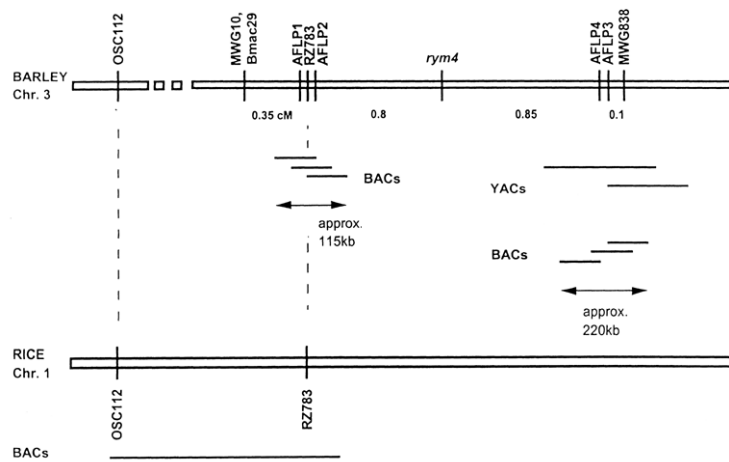


Abbildung 1:
Syntanie zwischen Gerste Chromosom 3H und Reis Chromosom 1
Ausgangssituation zur vergleichenden Kartierung des *rym4/5* Locus zu Beginn des Projektes. Distal des Gens *rym4* wurden die Reismarker RZ783 und C112 in Colinearität zwischen Gerste und Reis kartiert. In jeweils ca. 0,8 cM Entfernung vom Resistenzlocus waren bereits Contigs etabliert

Relevante Publikationen:

Graner, A., Bauer, E.: RFLP-mapping of the *ym4* virus resistance gene in barley. Theor Appl Genet 86:689-693 (1993).

Graner, A., Streng, S., Kellermann, A., Schiemann, A., Bauer, E., Waugh, R., Pellio, B., Ordon, F.: Molecular mapping and fine structure of the *rym5* locus encoding resistance to different strains of the Barley Yellow Mosaic Virus Complex. Theor Appl Genet 98: 285-290 (1999).

Graner, A., Streng, S., Kellermann, A., Proeseler, G., Schiemann, A., Peterka, H., Ordon, F. Molecular mapping of genes conferring resistance to soil-borne viruses in barley – an approach to promote understanding of host-pathogen interaction. J. Plant Diseases & Plant Protection 106: 405-410 (1999).

Moore, G., Devos, K.M., Wang, Z.Gale, M.D.: Grasses, line up and form a circle. Curr Biol 5: 737-739 (1995).

Smilde, W.D., Halukova, J., Sasaki, T.Graner, A.: New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. Genome 44: 361-367(2001).

Sorrells, M.E., La Rota, M., Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., Munkvold, J.D., Miftahudin, M.A., Ma, X., Gustafson, P.J. et al.: Comparative DNA Sequence Analysis of Wheat and Rice Genomes. Genome Res 13: 1818-1827(2003).

Streng, S.: Strategien zur Markersättigung im Bereich des *rym4* Resistenzlocus bei der Gerste. Dissertation, Shaker Verlag, Aachen, ISBN 3-8265-6876-1 (2000).

I.5. Zusammenarbeit mit anderen Stellen

Die Sequenzierung von insgesamt 12 Gersten-BACs wurde nach Ausschreibung als Unterauftrag an die Fa. AGOWA GmbH, Glienicke Weg 185, 12489 Berlin vergeben.

II.1. Erzielte Ergebnisse

Projektteil A: Sequenzierung der Region um das *Rh2*- Resistenzgen auf Chromosom 7H

(durchgeführt in der Arbeitsgruppe Gen- und Genomkartierung unter Leitung von Frau Dr. Röder)

A.1 Feinkartierung des Resistenzgenes *Rh2*

Aus vorausgehenden Kartierungen in einer Doppelhaploiden-Population war bekannt, dass *Rh2* zwischen den Markern PSR119 und MWG555 in einer distalen Region auf Chromosomenarm 7HS kartiert (Schmidt et al. 2001). Die beiden genannten Marker wurden zu einem Rekombinantscreen in einer F2-Population der Kreuzung Atlas x Steffi verwendet. Aus rund 750 F2-Pflanzen (= 1500 Gameten) wurden 15 rekombinante Pflanzen in dem Intervall PSR119-MWG555 gefunden und zu homozygot Rekombinanten weiter entwickelt. Dieser Schritt ist notwendig, da der Rhynchosporium Resistenz-Test nicht mit Einzelpflanzen durchgeführt werden kann und mehrmals wiederholt werden muss. Die Resistenz-Tests wurden in Zusammenarbeit mit Herrn Dr. Günther Schweizer von der Bayerischen Landesanstalt in Freising durchgeführt.

Im weiteren Verlauf der Arbeiten wurden weitere Marker gefunden, welche näher an *Rh2* kartieren. In einem weiteren Rekombinantscreen mit den Markern AFLP14 und B-R1 wurde die Gesamtzahl der getesteten F2-Pflanzen auf ca. 2500 (= 5000 Gameten) erhöht und weitere 7 Rekombinante in diesem Intervall gefunden, so dass die Gesamtzahl der Rekombinanten in diesem Intervall 9 beträgt. Im Verlauf des weiteren ‚Walk‘ wurden neue Marker aus den gefundenen BACs entwickelt (Abb. 2). Die Marker G-71 und G-REV kosegregieren mit den bisher vorhandenen rekombinanten Pflanzen mit *Rh2*. Daher wurde ein weiterer Rekombinantscreen mit den Markern AFLP14 und A-D23R initiiert. In insgesamt 2155 Pflanzen (= 4310 Gameten) wurden 12 Rekombinante identifiziert. Aus diesen werden derzeit homozygote Linien selektiert, um die Feinkartierung durchführen zu können. In Abb. 2 sind die verschiedenen Schritte der Rekombinantenselektion zusammengefasst.

Für das Intervall AFLP14/A-23R ergeben sich nach dem dritten Screening insgesamt 20 rekombinante Pflanzen in einer Population von 4721 F2-Pflanzen (9442 Gameten), welches einer genetischen Distanz von 0,21 cM für das Intervall bzw. 0,011 cM pro rekombinante Pflanze entspricht. Die Auflösung entspricht demnach 11-48 kb pro rekombinante Pflanze, wenn man nach Künzel (2000) zugrunde legt, dass 1 cM am telomeren Ende auf dem kurzen Arm von Chromosom 1 (7H) einer physikalischen Distanz von 1 - 4,4 Mb entspricht. Diese Größenordnung deckt sich mit den beobachteten 6 Rekombinanten /230kb = 1 Rekombinante/38 kb im Intervall E-E4 bis G-REV basierend auf insgesamt 2571 Pflanzen (=5142 Gameten) vom ersten und zweiten Screening.

A.2 Erstellung eines BAC-Contigs und Sequenzierung

Durch Ausnutzung der Gersten/Reis Syntänie wurden zwei BAC-Contigs von ca. 230 kb mit Gersten BACs der Sorte Morex erstellt. Zunächst wurde das proximale Contig aufgebaut, welches auf der rechten Seite durch die Probe B-R1 begrenzt ist (Abb. 3). Auf der linken Seite wird das Contig durch die Probe A-EST begrenzt. Während dieses Contig anfänglich mit *Rh2* kosegregierte, ist es nach der Erweiterung der Kartierungspopulation nun eine Rekombinante von *Rh2* entfernt (Abb. 3). Drei Sonden, welche aus vorhergesagten Genen aus dem syntenen Reis BAC P0514G12 (AP000616) entwickelt wurden, zeigen bei Hybridisierung Homologien zu dem proximalen Gersten BAC-Contig (Abb. 3). In Abb. 3 sind nur die drei BACs A-246J13, C-524N3 und B-288D11 dargestellt, welche vollständig sequenziert wurden.

Durch Ausnutzung der bekannten syntänen Reis-Sequenz wurden weitere Hybridisierungsproben generiert und mit der Probe E-E4 ein BAC (BAC E-621J22) auf der linken Seite von *Rh2* gefunden. Durch Subklonierung und Sequenzierung dieses BACs konnte der genomische ‚Walk‘ fortgesetzt werden und die beiden BACs F-793L23 und G-693M6 wurden nacheinander

isoliert und sequenziert. Diese bilden ein zusammenhängendes Contig von ca. 230 kb dessen Ende (Hybridisierungsproben G-71 und G-REV) mit *Rh2* kosegregiert (Abb. 3). Alle verwendeten Hybridisierungsproben und damit gefundenen BACs sind in Tabelle1 aufgelistet.

Die Sequenzierung wurde von der Fa. Agowa mit gleicher Technik wie unter Kapitel B.2 beschrieben durchgeführt. In BAC A-246J13 konnten bis jetzt zwei Lücken in repetitiven Bereichen nahe dem linken Ende nicht geschlossen werden. Eine weitere Lücke existiert in BAC C-524N3.

Tab. 1: Screening der Gersten BAC Bibliothek mit Reis-Sonden, Gersten ESTs und Gersten BAC-Subklonen aus der *Rh2* Region

DNA-Sonde	Positive BACs in Gersten BAC Bibliothek
Reis BAC-Sonde Xba6 = CDS20, 21	Gruppe1: 52C22, 60J17, 134N3, 274A20, 556G13, 750E8, Gruppe 2: 40I5, 70D24, 85H13, 172C5, 260N5, 350P22, 443M12, 626J22, 636I5, 638D10, 712I5, 721A19, 740A1, 813C22
Telomeres Contig	
Gersten-EST HW08F19 = Marker GBR0961, (Homologie zu CDS23 auf Reis BAC P0514G12, Sequenz AP000616)	Gruppe1: 61K19, 124I19, 137H16, 196B16, 202I19, 258E4, 266L19, 274E2, 344E20, 410O12, 557D13, 621J22, 641H6, 643P7, 701C18, 793L23, 346P17 Gruppe 2: 70P11, 256L8, 267C4, 337O13, 352P12, 385O24, 449G17, 610A1, 657H4, 712I5, 740A1, 814D12 Gruppe 3: 54O15, 371E7, 372F10, 532F22, 625L9, 811M24
E-E4 (621J22)	137H16, 202I19, 258E4, 266L19, 274E2, 344E20, 381N21, 557D13, 621J22
E-G6 (621J22)	73O13, 124I19, 196B16, 202I19, 258E4, 266L19, 274E2, 344E20, 346P17, 352P12, 410O12, 499P7, 557D13, 621J22, 641H6, 643P7, 701C18, 751G16, 793L23,
E-80U1 (621J22)	61K19, 72A22, 124I19, 137H16, 196B16, 258E4, 266L19, 274E2, 320O13, 344E20, 346P17, 352P12, 410O12, 499P7, 557D13, 621J22, 641H6, 643P7, 701C18, 751G16, 793L23
F-U1 (793L23)	196B16, 258E4, 274E2, 344E20, 410O12, 499P7, 641H6, 643P7, 693M6, 751G16, 793L23
G-71 (693M6)	110H17, 620M10, 650J6, 670O13, 677J6, 693M6
G-REV (693M6)	110H17, 620M10, 650J6, 670O13, 677J6, 693M6
Proximales Contig	
A-EST (246J13)	22A4, 246J13, 252H18, 259A13, 267J16, 342G18, 362C6, 380E18, 380I18, 399D20, 404G7, 417B24, 439C24, 441I24, 499J24, 515B6, 552H13, 594A19, 671F19, 681O7, 783O8,
Reis BAC-Sonde PST1 entspricht Marker A-23R, (246J13) (Homologie zu CDS27 auf Reis BAC P0514G12, Sequenz AP000616)	246J13, 252H18, 414H11, 524N3, 552H13

C-Con652-21 (524N3)	288D11, 524N3, 702E18,
B-R1 (288D11)	1D23, 46B5, 259B3, 288D11, 374L16, 401I7,

A.3 Sequenzanalyse der BAC-Contigs und Vergleich mit der syntänen Sequenz in Reis

Der erste Schritt bei der Sequenzanalyse war die Identifikation der repetitiven Regionen und dann die Annotation der Gene in den beiden Gersten BAC-Contigs. Die repetitiven Sequenzbereiche wurden mit Hilfe einer Datenbank für repetitive Sequenzen in Gräsern (TREP-Triticeae Repeat Sequence Database) bestimmt. Die konservierten Regionen wurden mit dem Dotter Programm bzw. einem automatischen Annotationsprogramm RICEGAAS (rice genome automated annotation system), welches vom japanischen RGP (Rice Genome Research Program) entwickelt wurde, und mit Blast Vergleichen am NCBI (National Center of Biotechnological Information) identifiziert.

Der Dotter Vergleich der gesamten proximalen BAC-Contig Sequenz (BACs 246J13, 524N3, 288D11) bzw. der gesamten telomeren BAC-Contig Sequenz (BACs 621J22, 793L23, 693M6) mit der gesamten Sequenz (AP000616.2) des orthologen Reis BACs, in zwei getrennten Vergleichen, deutete auf drei konservierte Regionen im proximalen BAC-Contig und zwei konservierten Regionen im Telomeren BAC-Contig. Die gleichen Regionen wurden durch RICEGAAS und Blast Vergleiche identifiziert, was darauf hindeutet, dass es sich bei den konservierten Regionen um Gene handelt. Außerhalb der Gene wurde keine Homologie festgestellt (siehe Abb. 4). Die identifizierten repetitiven Sequenzen (Transposons, Retrotransposons, Mites, Tandem Repeats) betragen im proximalen Contig etwa 86 kb von 237 kb und im Telomeren Contig etwa 82 kb von 232 kb im Contig (Abb. 5 und 6). Die konservierte Region erstreckt sich in Reis auf eine Region von 40 kb, in Gerste auf mindestens 380 kb (Abb.7), und repräsentiert den etwa 11-fachen Größenunterschied zwischen dem Gersten- und dem Reisgenom.

Proximales Contig:

Das proximale BAC-Contig umfaßt rund 237 kb. Eine detaillierte Sequenzanalyse ist in Abb. 5 dargestellt, eine Auflistung der putativen Gene in Tab. 2.

Die drei homologen Gene in Reis und dem proximalen Gersten Contig liegen in Reis in einer Region von 14 kb und in Gerste in einer Region von 140 kb, wobei in Gerste zwei der drei Gene (19 und 20) in einer Region von 15 kb geclustert liegen. Die Distanz zwischen Gen 19 und Gen 20 (in Gerste 5 kb in Reis 2kb) und zwischen Gen 20 und Gen 21 (Gerste 130 kb, Reis 2 kb) ist für Gerste größer als bei Reis, und für die beiden Genpaare sehr unterschiedlich (Abb. 7). Die unterschiedlich großen intergenen Räume von Gerste und Reis sind durch die Anwesenheit von repetitiven Elementen in Gerste bedingt. Die ersten beiden Gene 19 und 20 des proximalen Gersten Contigs besitzen die gleiche Orientierung, wie die Gene 19 und 20 (P0514G12.37 und P0514G12.39) in Reis. Das dritte orthologe Gen 21 liegt in Gerste in entgegengesetzter Transkriptionsrichtung wie in Reis Gen (P0514G12.41) vor und ist am 3' Ende verkürzt, d.h. nur das 5' Ende des Gens ist vorhanden (Abb. 5 und 6).

Telomeres Contig:

Das telomere BAC-Contig umfaßt rund 232 kb. Eine detaillierte Sequenzanalyse ist in Abb. 6 dargestellt, eine Auflistung der putativen Gene in Tab. 2.

Am 5' Ende des BAC Contigs liegt ein hypothetisches unvollständiges Resistenzgen (1). Aus dem 5' Ende des Gens wurde der Marker E-E4 entwickelt (Abb. 3 und 7). Stromabwärts im Contig in Richtung Centromer liegen zwei Gene in der Reihenfolge Betaphosphoglucomutase (Gen 2) – putatives F-Box Gen (Gen 5) – putatives F-Box Gen (Gen 7) – Betaphosphoglucomutase (Gen 8) (Abb. 7) als identische Kopien vor, die sich jeweils in entgegengesetzter Orientierung befinden. In deren Mitte liegt ein putatives Cytochrom P-450 Gen CYP₁₇₉₃ (Gen 6) aus dessen 5' Ende und 5' flankierender Region der Marker (E-G6) entwickelt wurde (Abb. 3 und 7). Stromaufwärts befinden sich zwei Polyubiquitine (Gen 9 und Gen 11), wobei Gen 9 kein Start-Methionin besitzt (Tab. 2) und zwei partielle hypothetische Gene (10) und (12) in gleicher Orientierung, die für eine Ketol-acid-

reductoisomerase kodieren. Gen 13 ist ein weiteres hypothetisches Gen. Stromabwärts liegen dann ein hypothetisches Resistenzgen (14) und ein Resistenzgenhomolog (15) mit Homologie zu einem Resistenzgen in Reis und Gerste (Tab.2). Anschließend kommen zwei vorhergesagte Gene (16 und 17) und ein hypothetisches Gen (18).

Bis auf die Betaphosphoglucomutase Gene 2 und 8, die in invertierten Orientierungen vorliegen, ist keines der übrigen Gersten Gene im Telomeren Gersten Contig in der orthologen Reisregion vorhanden, d. h. an diesem Locus im Gerstengenom auf Chromosom 1 (7H) ist die Kolinearität zur orthologen Reisregion auf Chromosom 6 unterbrochen. Ein Hinweis auf eine stärkere Rekombinationshäufigkeit genau in dieser Gerstenregion, zeigt sich bei den genetischen Untersuchungen mit den Markern 621J22-E4 und 621J22-G6. Für deren Intervall wurden für eine physikalische Distanz von etwa 52 kb im Telomeren Contig vier rekombinante Pflanzen selektiert, was 13 kb pro rekombinante Pflanze entspricht. Für den übrigen Teil beider Gersten BAC Contigs wurde ein Durchschnitt von etwa 82 kb pro rekombinante Pflanze ermittelt. Möglicherweise wurde die Kolinearität in dieser Region im Gerstengenom durch diese verstärkte Rekombinationshäufigkeit gebrochen

Tab. 2: Putative Gene in den sequenzierten Gersten BAC Contigs. Gene wurden mit den Genannotationsprogrammen RiceGAAS (rice genome automated annotation system) vom japanischen Rice Genome Research Program (RGP) oder der PEDANT Software von Biomax Informatics AG vorhergesagt.

Gen Nr.	Putative Gene in Gerste	Homologie zu <i>O. sativa</i>	Andere Homologien zu ESTs, u.a.
Telomeres Contig			
1	putatives Resistenzgen	Homologie zu hypothetischem Gen CDS 5 auf BAC B1003B09 (Sequenz AP004222.5) auf Chromosom 1	Homologie zu Resistenzgen <i>RGC2K</i> in Gartensalat (PIR:T30559), Homologie zu AB012240_8 Resistenzgen <i>RPS4 A. thaliana</i> auf Chromosom 5, TAC clone:K18C1
2	putatives Beta-phosphoglucomutase protein (BPGM)	Homologie zu Beta-phosphoglucomutase ähnlichem Gen P0514G12.30 in BAC Sequenz AP000616. Homologie zu putativem Beta-phosphoglucomutase protein AAL86484.1 bzw. NP_922801	Homologie zu Gersten EST Klon HW08F19; Homologie zu hypothetischem Protein At2g38740, G84808 <i>A. thaliana</i> ; Siehe auch Gen 8 gleiches Gen und identische Sequenz, aber entgegengesetzte Orientierung
3	vorhergesagtes Gen	kein Blast Hit	unbekannte Proteinfunktion
4	vorhergesagtes Gen	kein Blast Hit	unbekannte Proteinfunktion
5	putatives F-box Gen(<i>FB</i>)	Homologie zu hypothetischem Protein (BAA95883.1); Homologie zu hypothetischem protein AP002071_9 PAC- Klon P0675A05, Chromosom 6	N-terminal ist ähnlich zu S-locus F-box Domäne von <i>SFBd</i> (AB081648_1) von <i>Prunus dulcis</i> ; Homologie zu hypothetischem Protein (BAA95875.1) BAC F21M12; <i>A. thaliana</i> Chromosom 1; Siehe auch Gen 7 gleiches Gen und identische Sequenz, aber entgegengesetzte Orientierung

6	putatives Cytochrom P-450 Gen (<i>CYP_1793</i>)	Homologie zu putativem Cytochrom P-450 (BAC99858.1)	Homologie zu putativem Cytochrom P-450 Gen (NP_179782.1) in <i>A. thaliana</i>
7	putatives F-box Gen (<i>FB</i>)	Homologie zu hypothetischem protein (BAA95883.1); Homologie zu hypothetischem Protein AP002071_9 PAC- Klon P0675A05, Chromosom 6	N-terminal ist ähnlich zu S-locus F-box Domäne von SFBd (AB081648_1) von <i>Prunus dulcis</i> ; Homologie zu hypothetischem Protein (BAA95875.1) BAC F21M12; <i>A. thaliana</i> Chromosom 1; Siehe auch Gen 5 gleiches Gen und identische Sequenz, aber entgegengesetzte Orientierung
8	putatives Beta-phosphoglucomutase protein (BPGM)	Homologie zu Beta-phosphoglucomutase ähnlichem Gen P0514G12.30 in BAC Sequenz AP000616. Homologie zu putativem Beta-phosphoglucomutase protein AAL86484.1 bzw. NP_922801	Homologie zu Gersten EST Klon HW08F19; Homologie zu hypothetischem Protein At2g38740, G84808 <i>A. thaliana</i> ; Siehe auch Gen 8 gleiches Gen und identische Sequenz, aber entgegengesetzte Orientierung
9	putatives Polyubiquitin Gen (<i>UBQ793_1</i>)	Homologie zu Polyubiquitin 6 Gen (AAF01315.1)	Homologie zu Polyubiquitin in <i>Fragaria x ananassa</i> (AAB68045.1); Homologie zu putativem Polyubiquitin <i>UBQ10</i> in <i>A. thaliana</i> (At4g05320, BT000701_1)
10	hypothetisches Gen	partielle Homologie zu Ketol-acid-reductoisomerase aus Reis	
11	Ubiquitin Gen (<i>UBQ793_2</i>)	Homologie zu Reis Ubiquitin 6 Gen (PIR:S38669)	Homologie zu putativem Polyubiquitin <i>UBQ10</i> in <i>A. thaliana</i> (At4g05320, BT000701_1)
12	hypothetisches Gen	partielle Homologie zu Ketol-acid-reductoisomerase aus Reis	
13	hypothetisches Gen.	Homologie zu hypothetischem Reis Protein (Sequenz AAMM19038.1)	
14	hypothetisches Resistenzgen (<i>RGH793_1</i>)	Homologie zu putativem Resistenzgen (AC096691_9) BAC OSJNBa0083M08	Homologie zu Resistenzgenhomolog Pseudogen RGH3A und RGH3B in Gerste (BAC AF427791)
15	Resistenzgenhomolog (<i>RGH793_2</i>)	Homologie zu Resistenzgen <i>Pi-ta</i> (AAK00132); putativem NBS Resistenzgenhomolog (NP_911919)	Homologie zu Resistenzgenhomolog RGH2B in Gerste (AF427791_10, AAM22819.1)
16	vorhergesagtes Gen		
17	vorhergesagtes Gen		

18	hypothetisches Gen	Homologie zu Flavonol-3-sulfotransferase	
Proximalles Contig			
19	Cytochrome P-450 Gen (<i>CYP_1524</i>)	Homologie zu <i>Dwarf3</i> ähnlichem Gen (P0514G12.37, AP000616)	Homologie zu <i>CYP88A</i> Cytochrome P450 Gen (AAK11616) in <i>Hordeum vulgare</i> ; Cytochrome P450 <i>DWARF3</i> (GA ₁₂ -13-hydroxylase) in <i>Zea mays</i> (AAC49067.1)
20	<i>HVEYA</i> Gen (<i>H. vulgare</i> eyes absent Gen)	Homologie zu <i>OSEYA1</i> (eyes absent) Gen (P0514G12.39)	Homologie zu <i>Eya</i> (eyes absent CG9554-PA) in <i>Drosophila melanogaster</i> (NP_723188); <i>EYA1</i> in <i>Homo sapiens</i> (NP_742057); similar to eyes absent Gen in <i>A. thaliana</i> (AAM65149.1)
21	hypothetisches Gen Hypo41	Homologie zu hypothetischem Gen (P0514G12.41)	Homologie zu <i>PM29</i> Gen (AF117725_1) Samenreifeprotein in <i>Glycine max</i> , Homologie zu <i>LEA</i> (late embryogenesis abundant) Homolog <i>Le25</i> (PIR:S71249) in <i>A. thaliana</i>

Projektteil B: Sequenzierung der Region um den *rym4/rym5* Resistenzlocus auf Chromosom 3H
(durchgeführt in der AG Molekulare Marker unter Leitung von Herrn Prof. Graner und Dr. N. Stein)

B.1 Feinkartierung des Resistenzlocus *rym4/5* und Erstellung eines BAC-Contigs

Parallel zu dem vorliegenden Projekt wurde im Rahmen eines von der DFG geförderten Vorhabens die hochauflösende, genetische Feinkartierung, sowie eine weiterführende physische Kartierung des *rym5* Locus bearbeitet [DFG Schwerpunktprogramm 1005, 'Molekulargenetische Feinkartierung multipler Resistenzen der Gerste (*Hordeum vulgare* L.) gegen den Gelbmosaikvirus-Komplex

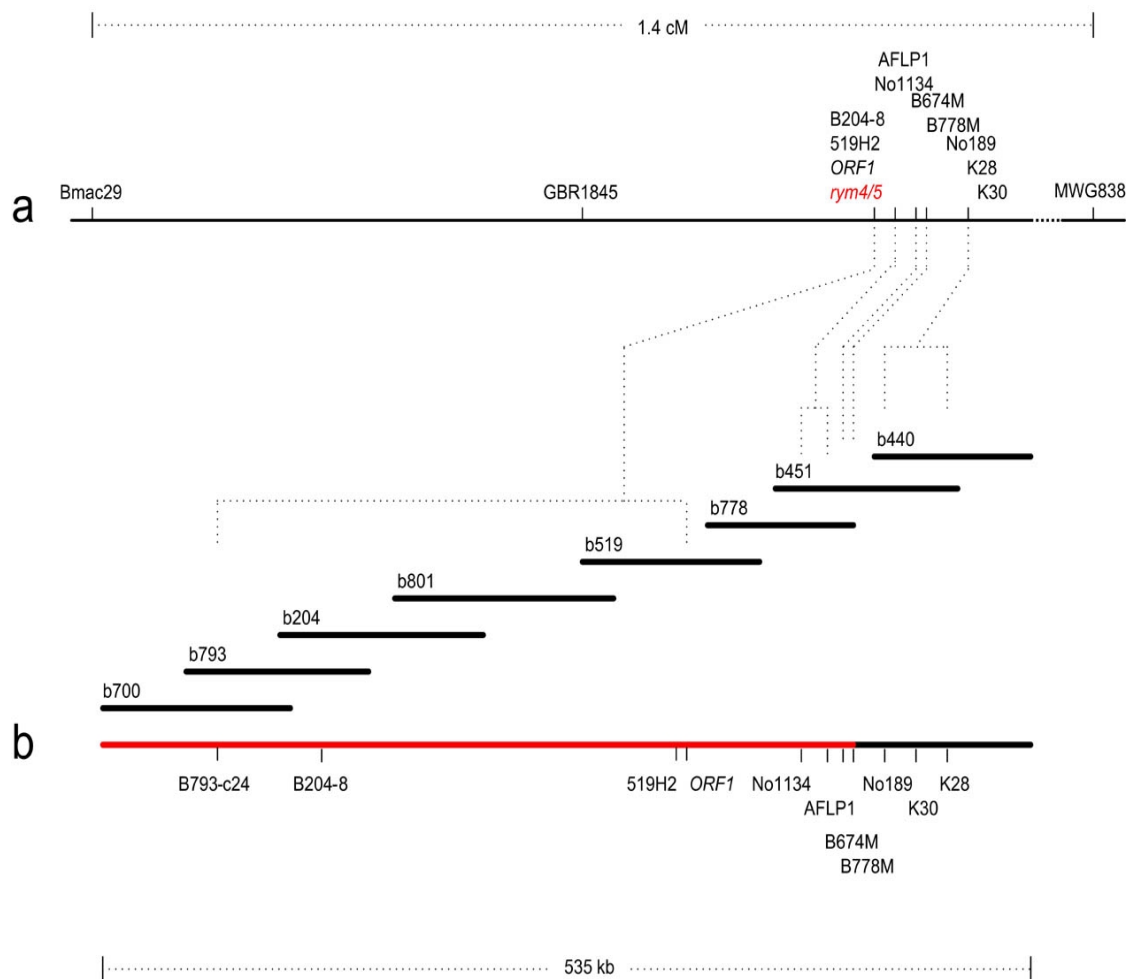


Abb. 8: Hochauflösende genetische und physische Feinkartierung des *rym4/5* Locus

a) Genetische Feinkartierung des *rym5* Gens in einer Population von 3884 F₂ Pflanzen. Die Karte besitzt eine Auflösung von ca. 0,015 cM pro Rekombinationsereignis. Ausgehend von Marker AFLP1 wurde mit der physischen Kartierung des Locus durch „Chromosome Walking“ begonnen. Alle Marker bis auf Bmac29, GBR1845, AFLP1 und MWG838 stammen aus dem Prozess der Erstellung einer physischen Karte des Resistenzlocus. b) Physische Karte bestehend aus überlappenden BACs aus der Morex BAC Bibliothek. Der Bereich, für den im Rahmen dieses Projektes eine vollständige Sequenz ermittelt wurde, ist rot markiert.

(BaYMV/BaMMV)‘, FKZ 1317/3-3, A. Graner und F. Ordon]. Im Zuge der genetischen Feinkartierung wurde eine hochauflösende Kartierungspopulation, Alraune (anfällig) x W122 (resistent, *rym5*), von insgesamt 3884 F₂-Pflanzen aufgebaut und zur Selektion von Rekombinanten im

Zielintervall um den *rym5* Locus eingesetzt. In einem umfangreichen AFLP-Screening wurden für diese Population eng gekoppelte Marker identifiziert (Abb. 8a) (0,05 cM Distanz zum Gen, Pellio 2002); einer dieser Marker diente als Ausgangspunkt für die Etablierung eines neuen BAC-Contigs. Dieser konnte durch acht „Chromosome Walking“ Schritte erweitert werden und kosegregiert mit dem Resistenzlocus über eine physische Distanz von mehr als 300 kb (Abb. 8b). Es ist bisher nicht gelungen auf diesem Contig ein distal vom Gen gelegenes Rekombinationsereignis nachzuweisen, so dass gegenwärtig nicht beurteilt werden kann, ob der Resistenzlocus bereits überschritten wurde. Sechs BACs dieses Contigs, die entweder ganz oder teilweise genetisch mit dem Resistenzlocus kosegregierten, wurden für das vorliegende Projekt zur Sequenzierung ausgewählt.

Relevante Publikationen

Pellio, B.: Hochauflösende Kartierung des *rym5* Resistenzlocus der Gerste (*Hordeum vulgare* L.). Dissertation, Shaker Verlag, Aachen, ISBN 3-8322-0752-X (2002)

B.2 Sequenzierung und Sequenzanalyse des *rym4/5* BAC-Contigs

Sequenzierung und Sequenzassemblierung wurden durch die Fa. AGOWA durchgeführt. Dabei wurden sogenannte Shotgun-Bibliotheken der zu sequenzierenden Klone erstellt und bis zu einer theoretisch zehnfachen Abdeckung der BAC Inserts (z.B. 10 x 100 kb Gesamtsequenz pro BAC) sequenziert. Verbliebene Lücken wurden durch sog. „Primer Walking“ auf Ebene der Subklone (Klone, die anhand Ihrer Teilsequenzen als Lücken-überspannend identifiziert wurden) geschlossen. Auf diese Weise wurden vier der sechs ausgewählten Klone des *rym4/5* Contigs vollständig sequenziert. In den Sequenzen der übrigen zwei Klone b519 und b778 konnten, aufgrund der hohen Dichte an stark konservierten, repetitiven Elementen oder des lokal erhöhten GC-Gehaltes, eine bzw. zwei Lücken bisher nicht geschlossen werden. Nach Abschätzung der Größe der verbliebenen Lücken (je zwischen 0,5 und 2,0 kb Länge) beträgt die Länge der Gesamtsequenz über 430.000 bp. Nach Fertigstellung der Sequenz des BACs b700 wurde eine sehr kurze Überlappung zu BAC b204 beobachtet (Abb. 8b), die vor Auswahl des Klons zur Sequenzierung nicht zweifelsfrei identifiziert werden konnte. Der einzige Marker, B204-8, der zur Verlängerung des Contigs ausgehend von b204 genutzt werden konnte, lag nicht im Überlappungsbereich mit b700 (Abb. 8b). Eine Identifikation von b700 über Fingerprint-Analysen direkt ausgehend von b204 war nicht möglich. Erst nach Vollendung der Sequenzierung von b793 gelang die Zuordnung des Klons b700 zu dem bestehenden Contig.

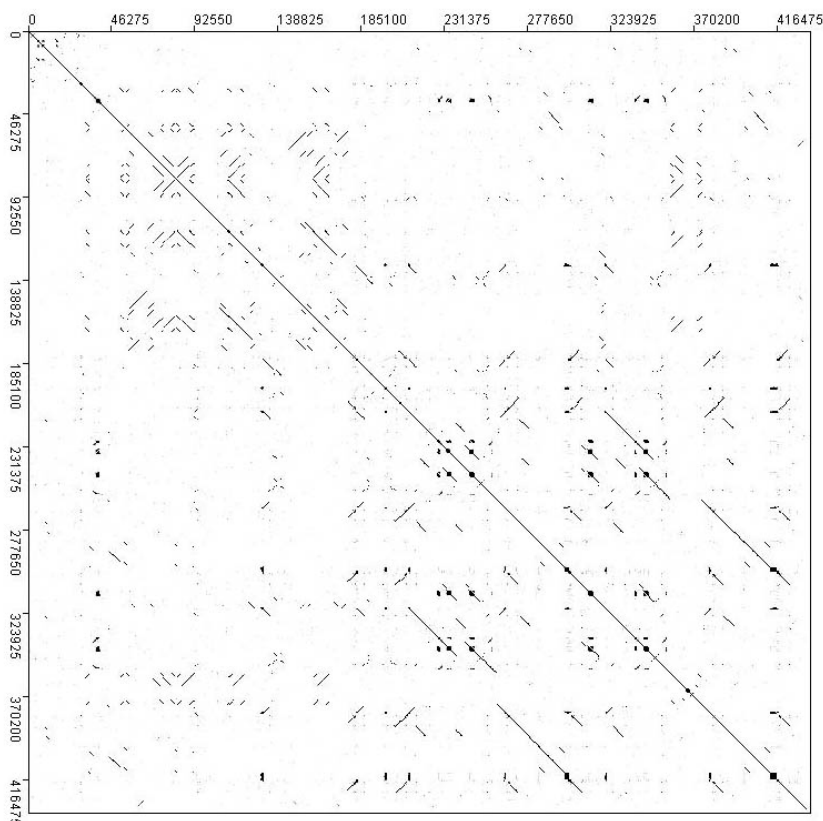


Abb. 9: Dotplot-Analyse des *rym4/5* Contigs

Die gesamte Sequenz des Contigs wurde gegen sich selbst verglichen. Die Hauptdiagonale stellt die vollständige Identität der Sequenz gegen sich selbst dar. Zusätzlich ist ein komplexes Muster an parallelen und senkrechten Linien zu der Diagonale sichtbar. Hierbei handelt es sich um Sequenzabschnitte, die in direkter oder inverser Form auf anderen Bereichen des Contigs wiederholt vorkommen. Der Contig besitzt nur in einem Bereich von 10 kb nach bp-Position 92.600 einen Abschnitt nichtrepetitiver DNA, der die identifizierten Gene *ORF1* und *ORF2* trägt.

Die Sequenzanalyse erfolgte unter Verwendung unterschiedlicher Werkzeuge. Zum einen wurden die erhaltenen Einzel-BAC Sequenzen über das *Rice Genome Automated Annotation System* (RiceGAAS, <http://ricegaas.dna.affrc.go.jp/>) analysiert. Dabei handelt es sich um ein automatisches Sequenzanalyse-Werkzeug, das eine Vielzahl verschiedener, bestehender Algorithmen (Genvorhersage, Sequenzvergleich gegen Nukleinsäure- und Proteindatenbanken, *Repeat-Masker*, etc.) beinhaltet und in die Analyse integriert. Obwohl dieses Werkzeug für die Annotation sequenzierter Reis-BACs entwickelt wurde, bietet es ein sehr effizientes System für die Schnellanalyse genomischer DNA verwandter Gräser. Darüber hinaus wurden Sequenzvergleiche gegen die umfangreichen EST-Sammlungen für Gerste und Weizen sowie gegen repetitive Elemente der Triticeen-Genome durchgeführt. Zur Visualisierung der Sequenz-Überlappung zweier BACs sowie der Verteilung repetitiver DNA-Elemente wurden die Sequenzen gegeneinander und gegen sich selbst in einem Dotplot Verfahren analysiert (Jdotter, <http://athena.bioc.uvic.ca/pbr/jdotter/>, Sonnhammer und Durbin 1995, Gene 167:GC1-10). Bei dieser Analyse werden zwei Sequenzen „Basenpaar für Basenpaar“ in einem 50 bp Fenster gegeneinander verschoben. Für dieses Fenster wird ein Sequenzähnlichkeitswert ermittelt, dem eine Graustufe von 0-256 zugeordnet wird. Bei der Darstellung des Dotplots, lässt sich dann ein Mindest- (Werte darunter werden als Weiss dargestellt) und ein Maximalgraustufenschwellenwert (Werte darüber werden als Schwarz dargestellt) wählen, um Hintergrund von signifikanten Sequenzhomologien zu unterscheiden. Längere Sequenzhomologien werden so durch Geraden sichtbar gemacht (Abb. 9).

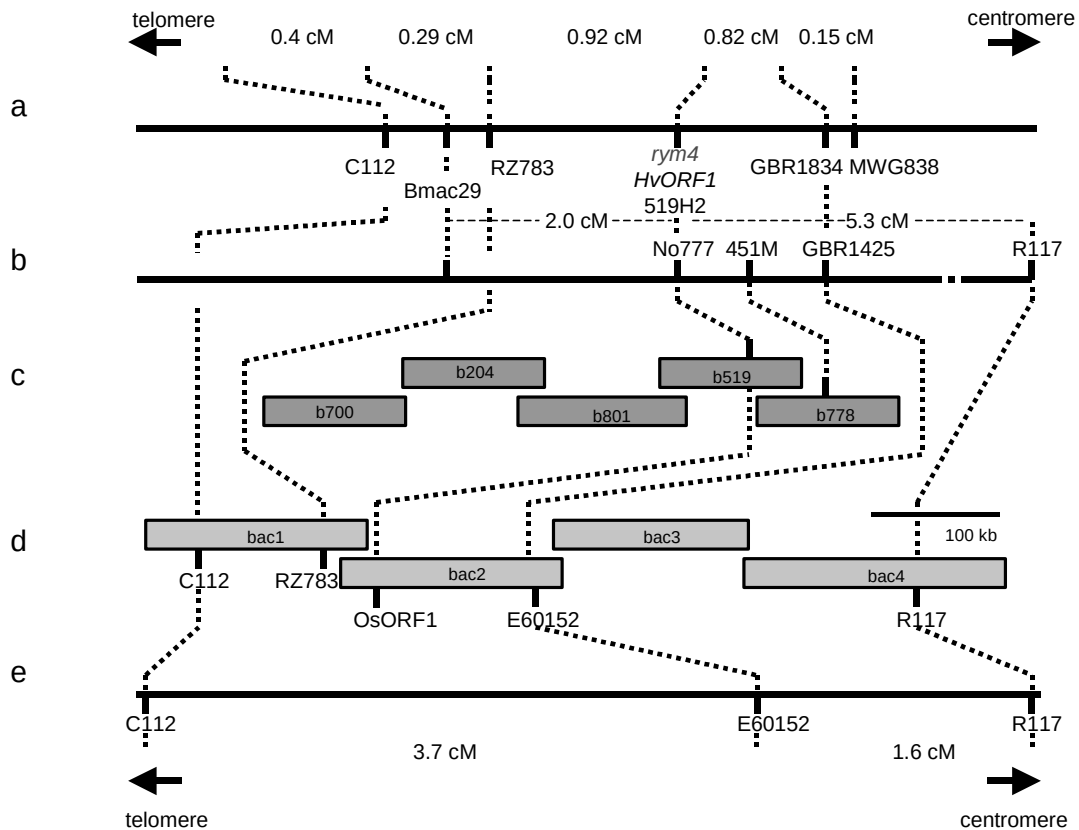


Abb. 10: Vergleichende genetische und physische Kartierung des Gerste *rym4/5* Locus mit der syntänen Region in Reis

a) Genetische Karte basierend auf 1.025 F2 Individuen der *rym4* Kartierungspopulation Igri x Franka, b) genetische Kartierung von Reis-EST Markern und BAC-Contig-abgeleiteten Gerste Markern in 94 Individuen der Gerste DH-Population OWB Dom x Rec, c) schematische Darstellung des sequenzierten Bereichs der physischen Karte des Gerste *rym 4/5* Locus, d) schematische Darstellung der physischen Karte der *rym 4/5*-orthologen Region des Reisgenoms, e) Ausschnitt der genetischen Karte des Reischromosoms 1L

Die detaillierte Sequenzanalyse zur Bestimmung der Zusammensetzung und Vollständigkeit der vorhandenen repetitiven DNA-Elemente ist noch nicht abgeschlossen. Die Dotplot-Analyse des *rym4/5* Contigs ergab jedoch einen Anteil von über 90 % Sequenzen, die in mindestens einmaliger Wiederholung auf demselben Contig zu finden sind. Hierbei handelt es sich um repetitive DNA – vorherrschende Klassen sind Vertreter aus der Familie der Gypsy-ähnlichen Retrotransposons und der TNP2-Transposons. Der einzige größere, offensichtlich nicht-repetitive Sequenzabschnitt des Contigs (ca. 10 kb Länge, bei bp-Position 92.600) beinhaltet die zwei einzigen vollständigen Gene, die auf

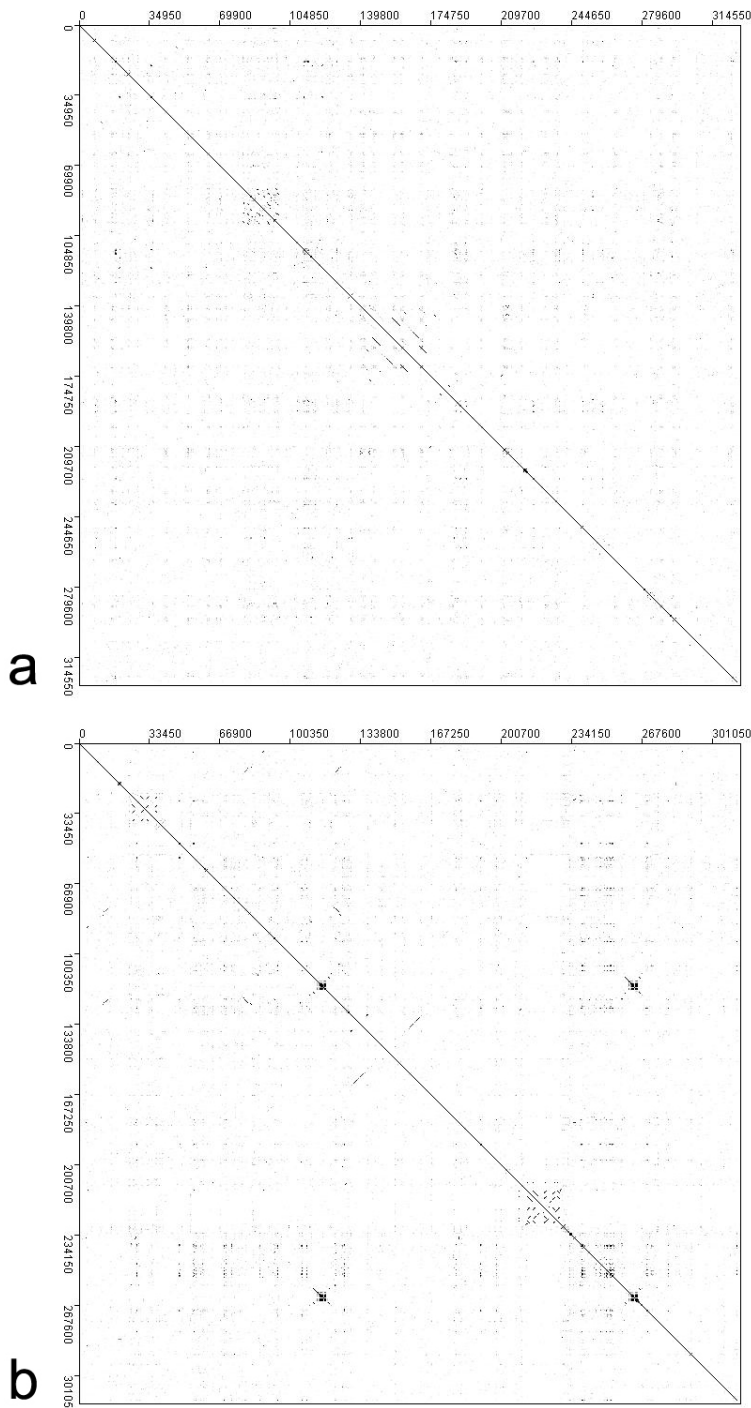


Abb. 11: Dotplot-Analyse des Reiscontigs
Der Reis-Contig wurde in zwei Portionen gegen sich selbst geplottet: a) Teilcontig bestehend aus bac3 und 4, b) Teilcontig bestehend aus bac1 und 2. Im Vergleich zu Gerste (Abb. 9) wird deutlich, dass das Reisgenom relativ wenige lokale Sequenzwiederholungen aufweist.

diesem Contig bisher identifiziert werden konnten (Abb. 9). Diese werden im folgenden als Leserahmen 1 (*ORF1*) und 2 (*ORF2*) bezeichnet. *ORF1* besitzt eine sehr hohe Sequenzhomologie zu einem Gen, dass als rezessives Resistenzgen gegen Potyviren in *Lactuca* und *Capsicum* beschrieben worden ist. Es stellt somit ein ausgezeichnetes Kandidatengen für *rym4/5* vermittelte rezessive Virusresistenz dar (BYMO-Viren gehören in die Verwandtschaft der Potyviridae). *ORF2* besitzt deutliche Homologie (> 80% Identität auf Proteinsequenzebene) zu einem Mono-Carboxylat-Transporter-ähnlichen Gen aus Reis (MCT-1 protein-like AP003535.2, *O. sativa*) und besitzt eine vergleichsweise geringe Bedeutung als Kandidatengen für Virusresistenz.

B.3 Vergleichende Sequenzanalyse der syntänen Sequenz aus Gerste und Reis

Die Region des *rym4/5* Locus auf Gerste Chromosom 3H war basierend auf der vergleichenden Kartierung zweier Reis-EST Marker als möglicherweise syntän zu einem Bereich auf Reis Chromosom 1 eingestuft worden. Im Verlaufe des Projektes wurden weitere Daten zur vergleichenden Kartierung der potentiell orthologen Regionen in Gerste und Reis ermittelt. Dies wurde ermöglicht durch Verwendung einer weiteren Kartierungspopulation und der Kartierung von sogenannten Anker-Markern in zwei oder mehr Populationen. Durch die Kartierung des Reis-EST Markers R117 in der DH-Population „Oregon-Wolfe-Barley (OWB) Dom x Rec“ konnte ein maximales kollineares Intervall zwischen Gerste und Reis definiert werden, das in Reis einem Contig aus

vier BACs mit einer Gesamtlänge von mehr als 600 kb entspricht (Abb. 10). Die Reis-BAC-Sequenzen wurden dann mittels BLASTN-Analyse gegen Gerste-EST-Sequenzen verglichen. Auf diese Weise konnten zwei Gerste ESTs, GBR1834 und GBR1425, in Gerste kartiert werden. Die beiden neuen Marker zeigen Homologie zu zwei ORFs des Reis-Klons bac3, die in enger Nachbarschaft zum Locus des Reis ESTs E60152 auf Reis BAC3 liegen (Abb. 10). Weiterhin befindet sich das eine der zwei Gerste-Gene (*ORF1*), die durch Sequenzierung des *rym4/5* Contigs identifiziert werden konnten, in Kollinearität zu Reis, so dass es sich bei diesem wahrscheinlich um ein orthologes Gen zwischen beiden Spezies handelt. Sollte es sich als Resistenzgen bestätigen lassen, wäre dieses ein Beispiel dafür, dass Reis als Modell zur Isolierung bestimmter agronomisch wichtiger Gene aus anderen Getreiden dienen kann. Interessanterweise, befindet sich das Reis-Homologe zu dem zweiten Gen *ORF2* nicht in Kollinearität zu Gerste. Es ist auf dem kurzen Arm desselben Reis Chromosoms lokalisiert. Insgesamt konnte am *rym4/5* Locus für mindestens 5 Gene/Marker Kollinearität zwischen Gerste und Reis beobachtet werden, bestätigt. Das in Bezug auf *rym4/5* relevante Intervall konnte in Reis auf ca. 200 kb zwischen den Markern E60152 und RZ783 eingegrenzt werden und enthält 39 vorhergesagte Gene oder Regionen mit signifikanter Homologie zu Reis-ESTs (Sequenzannotation durch das Reisgenomprojekt, RGP, Japan). Die Bestimmung der genetischen Position im Gerstengenom der besten Gerstenhomologe zu diesen Reisgenen wird gegenwärtig bearbeitet. Auf dem Gerste-BAC Contig einer Länge von über 500 kb konnte keiner der bisherigen flankierenden Reis Marker nachgewiesen werden und es wurden nur zwei Gene identifiziert. Dieser deutliche Unterschied in der Gendichte zwischen Gerste und Reis ist nicht überraschend. Es wird angenommen, dass Gerste und Reis eine vergleichbare Gesamtzahl an Genen besitzen bei einem ca. 12-fachen Genomgrößenunterschied zwischen beiden Spezies. Das größere Genom der Gerste wird im wesentlichen durch den hohen Anteil an repetitiven Sequenzen bedingt. Die *rym4/5* orthologe Region des Reisgenoms enthält nur sehr wenig repetitive DNA (Abb. 11). Die analysierten Sequenzen aus Reis und Gerste sind lediglich im kodierenden Bereich von *ORF1* (mehr als 80 % DNA-Sequenz Identität im Bereich der Exons) aber nicht im Bereich der repetitiven DNA konserviert. Die repetitive DNA des Gerste-Contigs ist nicht in der Reissequenz enthalten. Lediglich für ein Gypsy-Retroelement konnte Sequenzverwandtschaft zwischen Reis und Gerste beobachtet werden (Daten nicht gezeigt).

Die Ergebnisse zur vergleichenden Sequenzanalyse am *rym4/5*-Locus der Gerste und der syntänen Region des Reisgenoms lassen sich folgendermaßen zusammenfassen: Der Gerstenlocus befindet sich in einer Region stark ausgeprägter Makrokollinearität zu Reis. Das Ausmaß der Mikrokollinearität der Regionen lässt sich schwer abschätzen, weil im Gerste-Contig nur zwei Gene identifiziert wurden. Eines befindet sich in Kollinearität zur orthologen Region des Reisgenoms - das zweite, unmittelbar dem ersten benachbarte Gen befindet sich hingegen nicht in Kollinearität zu Reis. Außerhalb der Gensequenz im Bereich der repetitiven DNA lässt sich, aufgrund der geringen Sequenzhomologie, keine Kollinearität zwischen Gerste und Reis beobachten. Am *rym4/5*-Locus der Gerste befinden sich in einer vergleichbaren physischen Distanz 10-20 mal weniger Gene als in Reis, bei einem entsprechend höheren Anteil repetitiver DNA. Die Region um den *rym5* Locus bei Gerste kann somit als extrem genarm bezeichnet werden.

II. 2. Möglicher Nutzen und Verwertbarkeit der Ergebnisse

Die in Bezug auf den *rym4/5* Locus gewonnenen Ergebnisse zeigen, dass Kandidatengene für rezessive Virusresistenz an orthologen Positionen des Reisgenoms lokalisiert sein können. Sequenzkonservierung zwischen Gerste und Reis beschränkt sich auf die Genregionen. Homologe repetitive Sequenzen lassen sich in den Genomen beider Spezies identifizieren, sind aber für die Beurteilung der Genomkollinearität wertlos. Die Gendichte in einer Region des Reisgenoms ist kein Indikator der Gendichte einer orthologen Region des Gerstengenoms. Die ermittelten Daten lassen sich aufgrund ihres relativ geringen Umfangs nicht genomweit extrapolieren. Hierfür wäre eine umfangreichere Sequenzierung orthologer Bereiche des Gerste- und des Reisgenoms erforderlich.

Das identifizierte Kandidatengen für *rym4/5*-vermittelte Resistenz kann wichtige Informationen für die Entwicklung neuer molekularer Marker zur markergestützten Selektion auf Virusresistenz liefern. Sollte es sich um das Resistenzgen handeln, so wird das Potential für die Nutzung des Gens in einem

transgenen Ansatz, aufgrund des rezessiven Resistenztypus, zum gegenwärtigen Zeitpunkt als gering eingeschätzt.

In Bezug auf den Resistenzlocus *Rh2* hat das Projekt signifikante Fortschritte in der physikalischen Kartierung der Region erbracht. Ein Kandidat für ein Resistenzgen konnte jedoch im Rahmen der bisher sequenzierten BACs nicht gefunden werden, da die beiden resistenzgenhomologen Gene (Gen 14 und Gen 15 in Tabelle 2) noch durch zwei rekombinante Pflanzen von *Rh2* getrennt sind.

II.3. Fortschritte bei anderen Forschungsgruppen auf dem Gebiet

Die vergleichende Sequenzanalyse zwischen Gerste oder Triticeae-Spezies allgemein und Reis ist bisher auf wenige Loci der Triticeae-Genome beschränkt geblieben und diente meist der gezielten Genisolierung und nur in Ausnahmen der vergleichenden Genomanalyse (Büschges et al. 1997, Wei et al. 1999, Druka et al. 2000, Lahaye et al. 1998, Collins et al. 2001, Dubcovsky et al. 2001, Brunner et al. 2003). Dabei wurde beobachtet, dass sich die Sequenzhomologie zwischen Reis und Triticeae im wesentlichen auf codierende Sequenzen beschränkt und dass Regionen ausgeprägter Makrokollinearität häufig Abweichungen in der Mikrokollinearität aufweisen. An orthologen Loci sogenannter rassenspezifischer Resistenzgene ließen sich in Reis keine Kandidatengene identifizieren (z.B. *Rpg1*: Brueggemann et al. 2002) und generell scheint diese Gen-Kategorie nicht-kollinear zwischen Reis und den Triticeae organisiert zu sein (Leister et al. 1998). Für Gene mit möglicherweise allgemeinerer Bedeutung im Pflanzenreich wurden inzwischen mehrere Beispiele publiziert in denen Reis als Ressource für Kandidatengene hätte dienen können, so z.B. im Falle der Gene *Vrn1* (vernalisation response, Yan et al. 2003) und *ror2* (basal penetration resistance, Collins et al. 2003). Insgesamt wird die Gesamtzahl an Beispielen der vergleichenden Sequenzierung zwischen Reis und den Triticeae-Genomen als zu gering angesehen, um wirkliche genomweite Extrapolationen zur Genomkollinearität zu erlauben. Es gibt deshalb Initiativen zumindest die Gen-reichen Regionen der großen Gräsergenome zu sequenzieren (Mais: Chandler und Brendel, 2002; Weizen – IGROW: Gill und Appels, 2004).

Relevante Publikationen:

- Brueggemann, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., Kleinhofs, A. The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. *Proceedings of the National Academy of Sciences of the United States of America* 2002; 99: 9328-9333.
- Brunner, S., Keller, B., Feuillet, C. A large rearrangement involving genes and low-copy DNA interrupts the microcollinearity between rice and barley at the *Rph7* locus. *Genetics* 2003; 164: 673-683.
- Buschges, R., Hollricher, K., Panstruga, R., Simons, G., Wolter, M., Frijters, A., van Daelen, R., van der Lee, T., Diergaarde, P., Groenendijk, J. et al. The barley *mlo* gene: A novel control element of plant pathogen resistance. *Cell* 1997; 88: 695-705.
- Chandler, V.L., Brendel, V. The maize genome sequencing project. *Plant Physiology* 2002; 130: 1594-1597.
- Collins, N.C., Thordal-Christensen, H., Lipka, V., Bau, S., Kombrink, E., Qiu, J.-L., Huckelhoven, R., Stein, M., Freialdenhoven, A., Somerville, S.C. et al. SNARE protein mediated disease resistance at the plant cell wall. *Nature* 2003; 425: 973-977.
- Druka, A., Kudrna, D., Han, F., Kilian, A., Steffenson, B., Frisch, D., Tomkins, J., Wing, R., Kleinhofs, A. Physical mapping of the barley stem rust resistance gene *rpg4*. *Molecular and General Genetics* 2000; 264: 283-290.
- Dubcovsky, J., Ramakrishna, W., SanMiguel, P.J., Busso, C.S., Yan, L.L., Shiloff, B.A., Bennetzen, J.L. Comparative sequence analysis of colinear barley and rice bacterial artificial chromosomes. *Plant Physiology* 2001; 125: 1342-1353.
- Gill, B.S. and Appels, R. NSF/USDA/IGROW workshop on wheat genome sequencing. *Genetics* 2004, in press.
- Lahaye, T., Hartmann, S., Topsch, S., Freialdenhoven, A., Yano, M., Schulze-Lefert, P. High-resolution genetic and physical mapping of the *Rar1* locus in barley. *Theoretical and Applied Genetics* 1998; 97: 526-534.
- Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A., Schulze-Lefert, P. Rapid reorganization of resistance gene homologues in cereal genomes. *Proceedings of the National Academy of Sciences of the United States of America* 1998; 95: 370-375.
- Wei, F.S., Gobelmann-Werner, K., Morroll, S.M., Kurth, J., Mao, L., Wing, R., Leister, D., Schulze-Lefert, P., Wise, R.P. The *Mla* (powdery mildew) resistance cluster is associated with three NBS-LRR gene families and suppressed recombination within a 240-kb DNA interval on chromosome 5S (1HS) of barley (vol 153, pg 1929, 1999). *Genetics* 1999; 154: 1929-1948.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., Dubcovsky, J. Positional cloning of the wheat vernalization gene VRN1. Proceedings of the National Academy of Sciences of the United States of America 2003; 100: 6263-6268.

II.4. Erfolgte und geplante Veröffentlichung der Ergebnisse

Die Veröffentlichung der Daten zur vergleichenden Genom- und Sequenzanalyse im Bereich des *rym4/5* Locus und des *Rh2-Locus* und den orthologen Regionen des Reisgenoms ist in Form dreier Publikationen in Vorbereitung:

Perovic, D., Stein, N., Pellio, B., Streng, S., Stracke, S., Ordon, F., Graner, A. (2004): Collinearity of the barley and rice genomes and identification of a candidate resistance gene as revealed by map-based cloning at the *rym4/5* locus for BaMMV/BaYMV resistance, in preparation.

Wicker, T., Perovic, D., Zimmermann, W., Paterson, A., Graner, A., Stein, N. (2004): Genome organisation in a gene-poor region of the barley genome, in preparation.

Cossu, R., Schweizer, G., Schmidt, D., Ganai, M.W., Röder, M.S. (2004): Sequence comparison between barley and rice at a resistance gene region, in preparation.

Desweiteren sind Teilergebnisse auf dem 3. Gabi-Statusseminar präsentiert worden:

Frank, M. (2003): Plant environment interactions (biotic). 3. Gabi-Statusseminar, 11./12.02.2003, Bonn.

Cossu, R., Schmidt, D., Ganai, M.W., Röder, M.S. (2003): GABI-MAP: Analysis of a chromosomal region from rice and barley. Statusseminar GABI II, February 11th-12th, 2003, Bonn, Germany. (Poster)

Teilergebnisse wurden als Poster präsentiert:

Cossu, R., Schmidt, D., Ganai, M., Röder, M.: Analysis of a chromosomal region from rice and barley. Plant, Animal & Microbe Genomes X, San Diego, USA, Jan. 12-16, 2002.

Cossu, R., Schmidt, D., Schweizer, G., Ganai, M.W., Röder, M.S.: Towards the map-based cloning of resistance gene *Rh2* in barley. 11th Molecular Markers Symposium of the GPZ, IPK Gatersleben, 2003

Teile der Ergebnisse werden in Form von Posterbeiträgen präsentiert :

Stein, N., Perovic, D., Pellio, B., Stracke, S., Ordon, F. and Graner, A.: Chromosome walking reveals a candidate gene for Barley Mild / Barley Yellow Mosaic Virus Resistance at the Locus *rym4/5*. PAG-XII conference, 09.-14.01.2004, San Diego, USA.

Stein, N., Perovic, D., Pellio, B., Kumlehn, J., Ordon, F. and Graner, A.: Physical Mapping and Identification of Candidate Genes at the Virus Resistance Gene Locus *rym4/5*. 9th Barley Genetics Conference, 22.-26.06.2004, Brno, Czech Republic.

III. Erfolgskontrollbericht

1. Beitrag der Ergebnisse zu den förderpolitischen Zielen des Förderprogrammes

Das Projekt beinhaltete eine vergleichende strukturelle Genomanalyse von zwei Zielregionen im Gerstengenom im Vergleich zu Reis. Daneben trug es signifikant zu Fortschritten bei der markergestützten Klonierung zweier Resistenzgene bei. Insgesamt gesehen konnten in beiden Teilprojekten wichtige detaillierte Erkenntnisse zur Genomstruktur und zur Genomkonservierung zwischen zwei Modell-Gräsergenomen gewonnen werden, die einerseits bisherige Daten aus vergleichbaren Studien bestätigten, andererseits den potentiellen Nutzen der strukturellen Genomanalyse an diesen Modell-Organismen für die zukünftige Anwendung in den Genomen der Nutzpflanzen (Gerste, Weizen, Roggen) unterstreichen. Somit konnte das Projekt einen wichtigen Beitrag zu den förderpolitischen Zielen leisten.

2. Wissenschaftlich-technische Ergebnisse des Vorhabens

Das angestrebte Ziel, die vergleichende Sequenzanalyse zweier genomischer Regionen mit Resistenzgenloci in Gerste und Reis, wurde erreicht. Die gewonnenen Daten tragen zu den Erkenntnissen aus anderen Laboratorien bei, um ein genaueres Bild über die Genomevolution innerhalb der Gramineen zu gewinnen und die Nutzbarmachung von Reis als Modell-Organismus für die Gräser auszuloten.

3. Fortschreibung des Verwertungsplanes

• Erfindungen/Schutzrechtsanmeldungen und erteilte Schutzrechte

Die Anmeldung von Schutzrechten für klonierte Resistenzgene kann erst nach einer Bestätigung durch Transformation erfolgen, was nicht Bestandteil dieses Projektes ist.

• Wirtschaftliche Erfolgsaussichten nach Projektende

Die wirtschaftlichen Erfolgsaussichten des vorliegenden Projektes liegen auf verschiedenen Ebenen. Kurzfristig gesehen werden die gewonnenen Erkenntnisse eine Isolierung der betreffenden Resistenzgene erheblich beschleunigen. Darüber hinaus sind im Rahmen der Projekte neue molekulare Marker generiert worden, die eine Bedeutung für die markergestützte Selektion in Züchtungsprogrammen gewinnen könnten. Des weiteren haben die Arbeiten teilweise die Bedeutung des Modellorganismus Reis für die Isolierung von Kandidatengen in Gräsergenomen bekräftigt. Die konsequente Nutzung der Reisgenomsequenz-Ressourcen werden längerfristig eine effizientere Genisolierung aus Weizen und Gerste erlauben (mit Einschränkung bzgl. rassenspezifischer Resistenzgene).

• Wissenschaftliche Erfolgsaussichten nach Projektende

Das Projekt hat neben den Ergebnissen zur vergleichenden Sequenzierung wesentliche Fortschritte zur Klonierung zweier, wirtschaftlich bedeutender Resistenzgene in Gerste erbracht. Für den *rym4/rym5* Locus konnte ein Kandidatengen gefunden werden, welches noch durch Transformation bestätigt werden muss. Für den *Rh2*-Locus wurden substantielle Fortschritte bei der physikalischen Kartierung gemacht. Obwohl hier noch kein Kandidatengen gefunden wurde, soll in Nachfolgeprojekten die Klonierung von *Rh2* erfolgen.

• Wissenschaftliche Anschlussfähigkeit für eine notwendige nächste Projektphase

Da die eigentliche Klonierung der beiden Resistenzloci keine funktionelle Genomanalyse im engeren Sinne darstellt, wurde eine Fortführung des Projektes innerhalb von GABI II für die Förderung nicht als aussichtsreich betrachtet. Die getrennte Fortführung der beiden Teilprojekte im Rahmen von Projekten anderer Fördereinrichtungen wird jedoch angestrebt.

Generell sind die bisherigen Erkenntnisse zur Organisation der Triticeae Genome, sowie deren Gehalt an Genen nicht weitreichend untersucht. Bisherige Erkenntnisse stammen alle von wenigen, für bestimmte agronomisch relevante Gene ausgesuchten Loci im Weizen- und Gerstengenom und aus der EST-Sequenzierung. Um extrapolierbare Erkenntnisse zu erreichen, wäre die Möglichkeit zu einer umfangreicheren, zufälligen, möglicherweise chromosomenweiten vergleichenden Sequenzierung zwischen einer Triticeaespezies (Gerste/Weizen) und Reis sehr wünschenswert. Dies würde wichtige und umfangreiche Daten zur Analyse der Genstruktur in Gräsern, sowie der generellen Genomorganisation in einer kleingenomigen und einer großgenomigen Spezies einer Pflanzenfamilie liefern.

4. Arbeiten, die zu keiner Lösung geführt haben

keine

5. Präsentationsmöglichkeiten für mögliche Nutzer

Eine Präsentation für potentielle Nutzer wird wahrscheinlich erst nach erfolgreicher Klonierung und Bestätigung durch Transformation eines oder beider Resistenzloci aktuell, was nicht Bestandteil dieses Projektes ist. Ein mögliches Präsentationsforum wäre dann die GFP oder das direkte Angebot an entsprechende Industriepartner.

6. Einhaltung der Ausgaben- und Zeitplanung

Das Projekt wurde wie ursprünglich geplant in einem Zeitrahmen von drei Jahren abgewickelt. Da die Kosten für die Sequenzierung niedriger als ursprünglich geplant waren, konnte gegenüber der ursprünglichen Planung die doppelte Anzahl von BACs sequenziert werden. Aufgrund der geringen Gendichten in den beiden Zielregionen der Gerste erwies sich diese Maßnahme als essentiell für die Durchführung aussagefähiger Vergleiche zum Reisgenom und für die Charakterisierung der umfangreichen repetitiven Sequenzbereiche in Gerste.

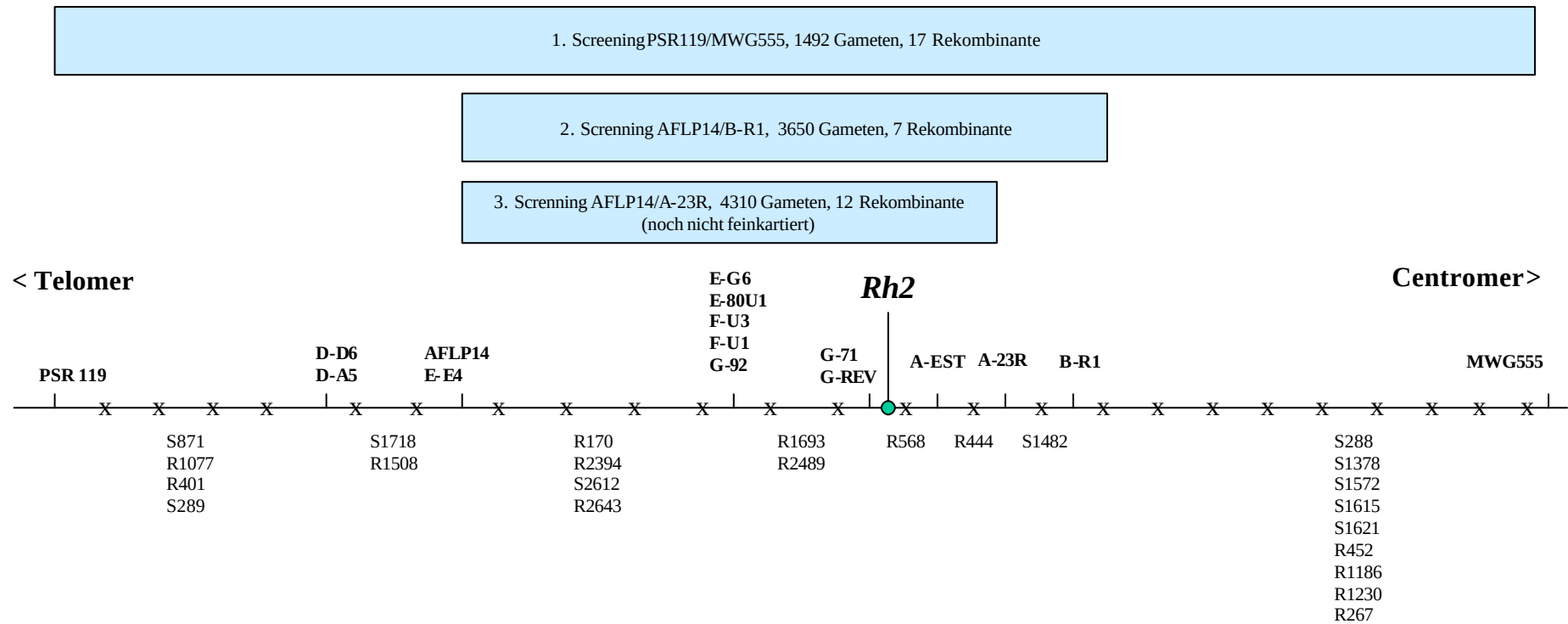


Abb. 2: Genetische Karte von Chromosom 1 (7H) in der *Rh2* Region.

Markernamen über Linie, rekombinante Pflanzen unter Linie, R = resistente Pflanzen, S = susceptible Pflanzen, x = genetische Distanz eine Rekombinante. Blaue Rahmen = Screening nach homozygot rekombinanten Pflanzen.

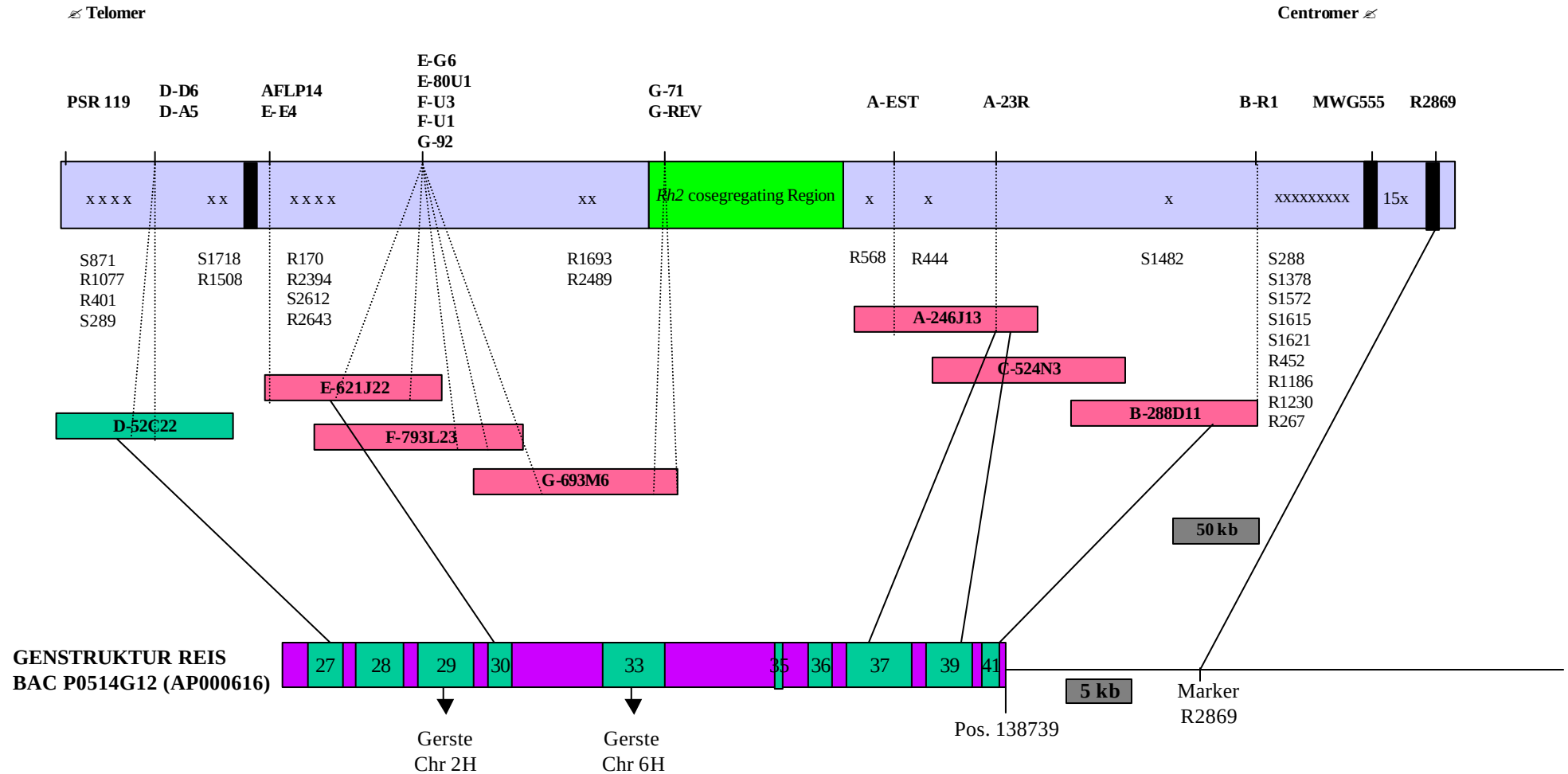


Abb. 3: Genetische und physikalische Karte der *Rh2* Region auf Chromosom 1 (7H) in Gerste. Schwarze Linien repräsentieren Proben die zwischen Reis und Gerste kreuzhybridisieren. x = Rekombinante Pflanzenlinien, S = anfällig, R = resistent, sequenzierte BACs rot, grüner BAC für Markerherstellung.

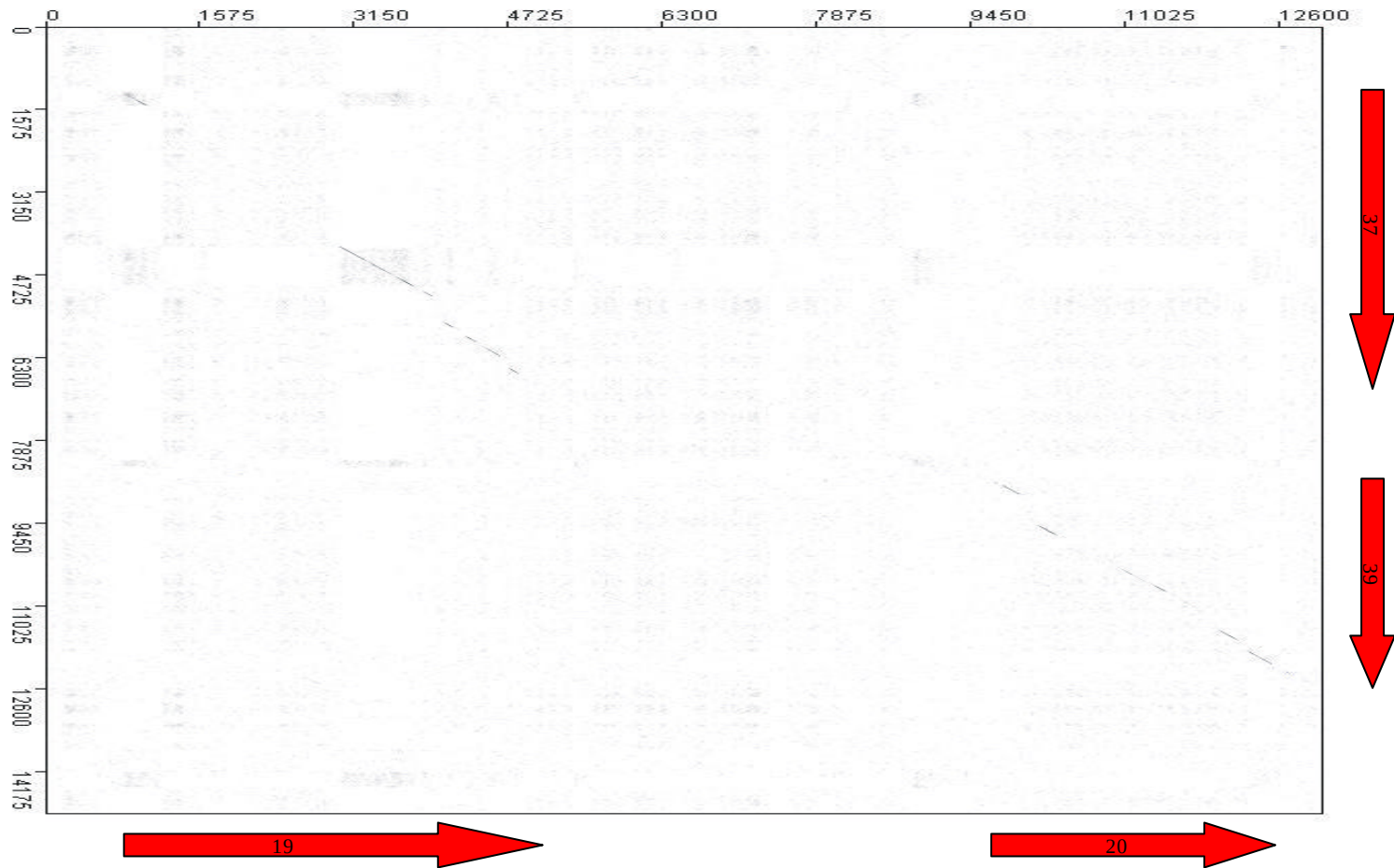


Abb. 4: Sequenzvergleich eines Ausschnitts des proximalen Gersten BAC Contigs (246J13, 524N3, 288D11) von Position 68001-81000 und des Reis PACs P0514G12 von Position 124001-138929 mit dem Programm Dotter. Die Lage der Gene in Gerste (horizontal) und Reis (vertikal) wird durch die Pfeile angezeigt. Die Exon-Intron Struktur der Gene ist erkennbar.

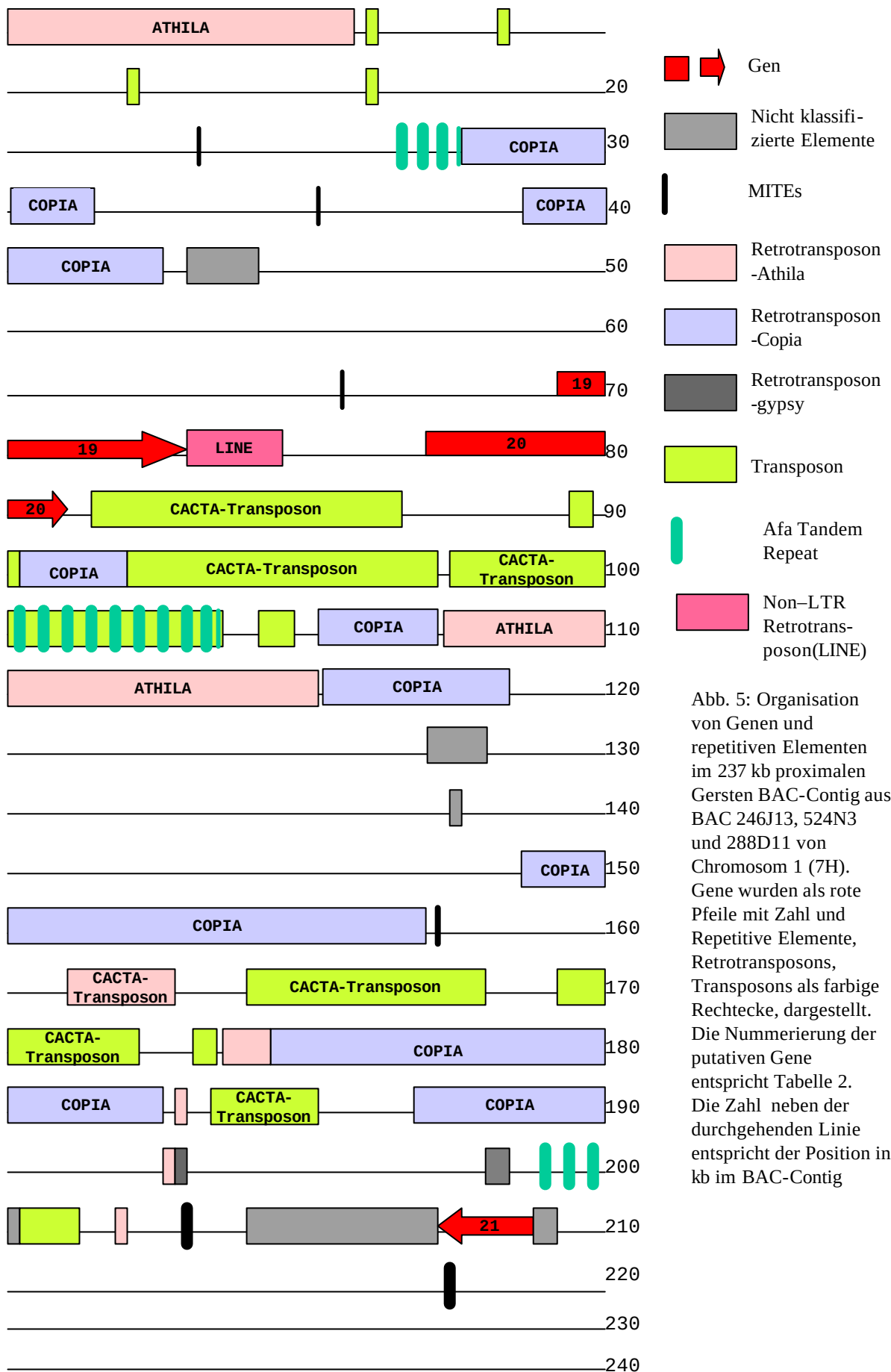
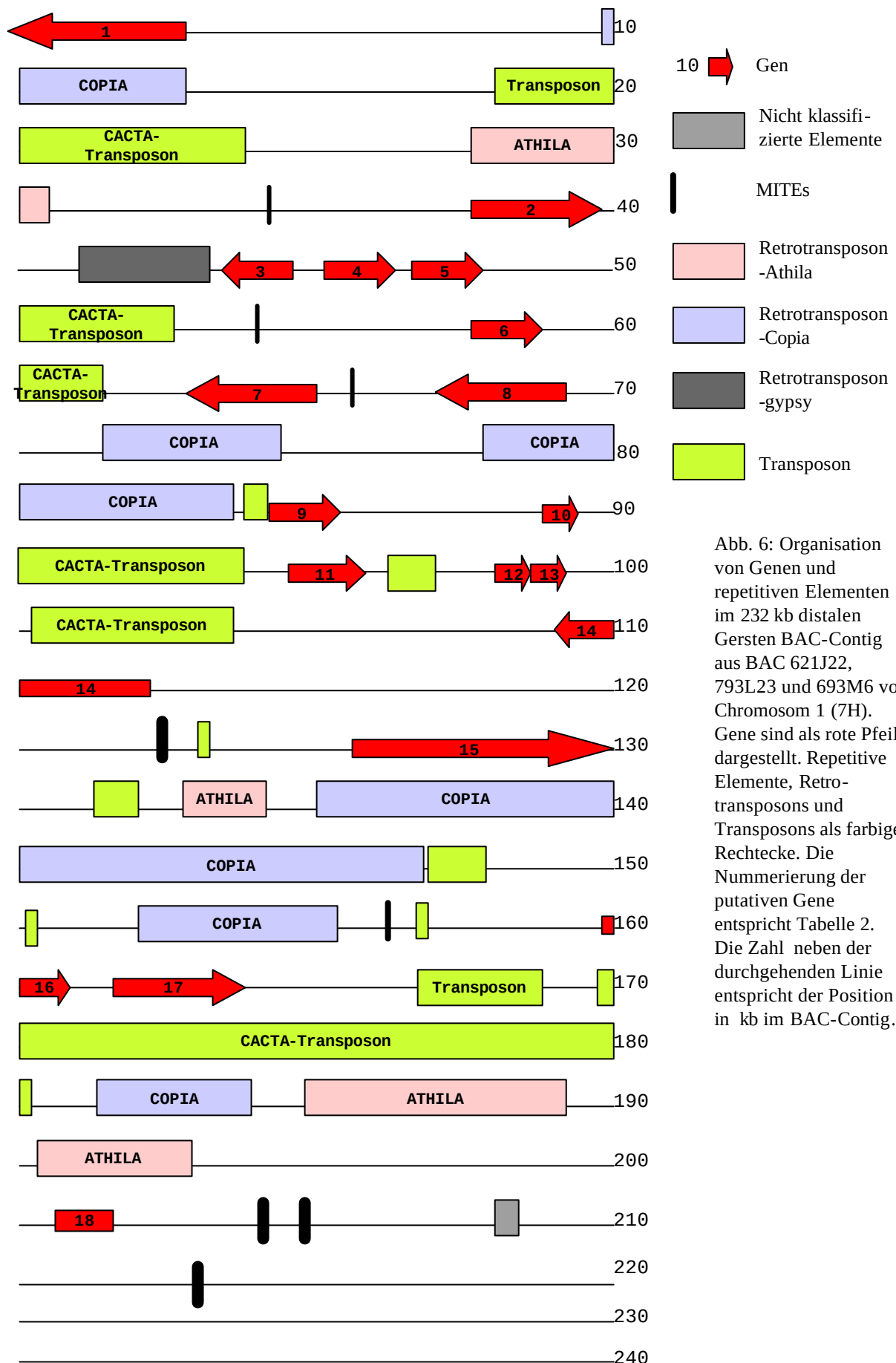
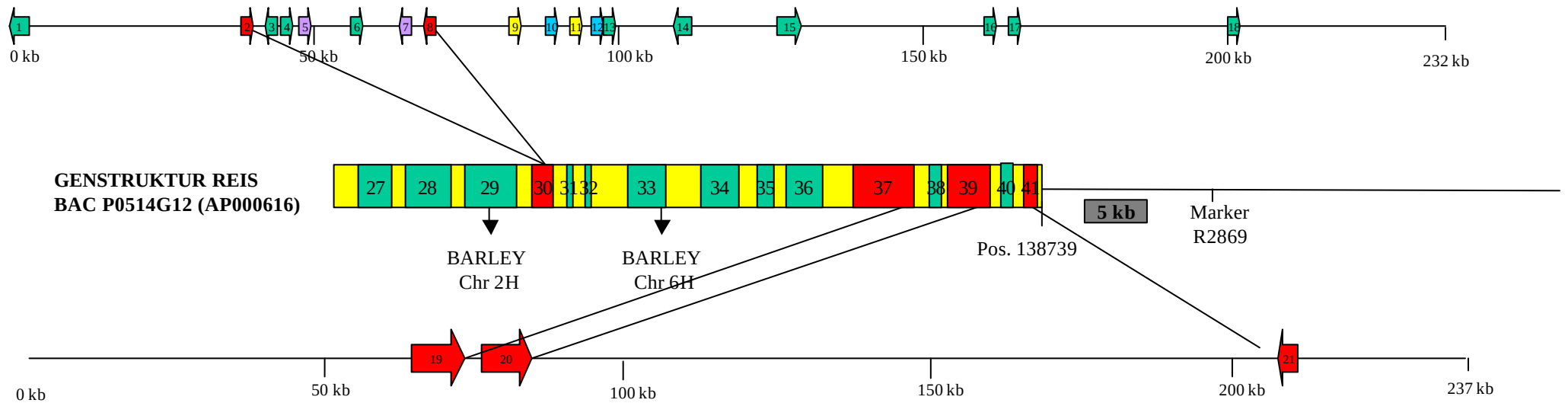


Abb. 5: Organisation von Genen und repetitiven Elementen im 237 kb proximalen Gersten BAC-Contig aus BAC 246J13, 524N3 und 288D11 von Chromosom 1 (7H). Gene wurden als rote Pfeile mit Zahl und Repetitive Elemente, Retrotransposons, Transposons als farbige Rechtecke, dargestellt. Die Nummerierung der putativen Gene entspricht Tabelle 2. Die Zahl neben der durchgehenden Linie entspricht der Position in kb im BAC-Contig



GENSTRUKTUR GERSTE TELOMERES BAC-CONTIG



GENSTRUKTUR GERSTE PROXIMALES BAC-CONTIG

Abb. 7: Organisation der Gene im telomeren und proximalen Gersten BAC Contig im Vergleich zu Reis BAC P0514G12 (AP000616). Die Rechtecke mit Zahlen entsprechen den Genen in Reis. Die Pfeile mit Zahlen entsprechen den Genen in Gerste. Rote Pfeile und Rechtecke sind homologe Gene zwischen Reis und Gerste. Im telomeren BAC Contig sind die Gene 2 und 8 bzw. 5 und 7 gleich, aber invertiert. Gene 10 und 12 sind gleich und haben die gleiche Orientierung. Gene 9 und 11 codieren beide für Polyubiquitin besitzen jedoch nicht die gleiche Sequenz.

