Regulators determining seed maturation: a genetical genomics approach

Sreenivasulu N., Pietsch C., Radchuk V., Strickert M., Röder M., Weschke W., Wobus U.

in

Molina-Cano J.L. (ed.), Christou P. (ed.), Graner A. (ed.), Hammer K. (ed.), Jouve N. (ed.), Keller B. (ed.), Lasa J.M. (ed.), Powell W. (ed.), Royo C. (ed.), Shewry P. (ed.), Stanca A.M. (ed.).
Cereal science and technology for feeding ten billion people: genomics era and beyond

Zaragoza : CIHEAM / IRTA
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 81

2008
pages 431-434

Regulators determining seed maturation: A genetical genomics approach[1]

N. Sreenivasulu†, C. Pietsch, V. Radchuk, M. Strickert, M. Röder, W. Weschke, U. Wobus
Leibniz-Institut für Pflanzengenetik und Kulturpflanzenforschung (IPK),
Corrensstraße 3, D-06466 Gatersleben, Germany
†E-mail for correspondence: srinivas@ipk-gatersleben.de

SUMMARY – We analysed the transcriptome of the two major parts of barley seeds, embryo and endosperm, throughout development using a 12 k cDNA array of seed-expressed unigenes. We focussed on transcriptional regulators/transcription factors (TFs), kinases and signalling networks. Correlation analysis of gene expression patterns demonstrated interactions of transcriptional and hormonal regulators during grain maturation. We propose that ABA in the starchy endosperm is controlling starch synthesis *via* SNF1 kinase, whereas ABA-responsive genes active during embryo maturation are involved in the control of lipid storage. Via ABA responsive element binding factors, ABA is furthermore believed to control acquisition of desiccation tolerance but the data also suggest an ABA independent but interactive pathway via DREB transcription factors. In an attempt to gain insight into the underlying genetic factors that govern differences in storage product accumulation we compared changes in gene expression during seed development among 42 homozygous lines carrying defined segments of a wild barley (*Hordeum spontaneum*) genome introgressed into the elite *H. vulgare* cv. 'Brenda' background and dissected the candidate regulatory genes involved in altering storage products.

Introduction

During the 1960's despite the lack of molecular genetic knowledge the "Green Revolution" contributed significantly to the improvement of cereal crops by the development of new dwarf-varieties, which resulted in tripling of yield (Khush, 1999). In order to cope with the further increasing world population it is necessary to double further the current crop yields mainly by making full use of the genetic potential of crop plants. The genomics revolution of the past decade resulted in the generation of a draft genome sequence of the cereal model crop rice (International Rice Genome Sequencing Project, 2005) and vast amounts of Expressed Sequence Tags (EST)-resources of large genome size cereal crop species (http://www.tigr.org/tdb/tgi/plant.shtml). Such sequence information in combination with transcriptomics, metabolomics and proteomics data generated for developing seeds continues to contribute significantly to the better understanding of the genetic make up of cereal seeds. For developing "knowledge based breeding" programs one needs to understand the orchestrated genetic architecture of gene expression controlling seed filling and yield-related parameters obtained from diverse germplasm and specific genetic material. Identified favourable alleles can be targeted by two approaches: by developing transgenic plants and through the development of near-isogenic lines (NILs) and recombinant-inbred lines (RILs), which eventually act as pre-breeding material for crop improvement. In this review we provide an overview of the strategies covering: (i) correlation-based network identification of key regulators controlling tissue-specific storage events in barley seeds; and (ii) implementation of a genetical genomics approach (Jansen and Nap, 2001) to track favourable alleles related to seed traits based on introgression lines (segments of a wild barley *Hordeum spontaneum* genome within the elite cultivar 'Brenda') in an doubled haploid back cross 3 (BC3-DH) population of barley.

Regulators determining tissue-specific seed maturation programs in developing barley seeds

Mature barley seeds are predominantly composed of the triploid endosperm and a diploid embryo. Both accumulate specific storage products such as starch/storage proteins and lipids, respectively.

[1] This work is supported within the German Plant Genome Research Program (GABI SEED-II grant by BMBF).

The storage pattern is mainly influenced by the orchestrated genetic architecture of gene expression. In particular, transcriptome analysis using profiling techniques in developing seeds (Sreenivasulu *et al.*, 2004) and EST-based *in silico* expression analyses during barley plant ontogeny including developing and germinating seeds (Zhang *et al.*, 2004) provides a wealth of mainly correlative data which help to dissect on the transcriptional level primary pathway networks and direct specific in depth studies to especially interesting network nodes (see Review Wobus *et al.*, 2005). Developmental changes at the transcriptome level also mark a dramatic transcriptional reprogramming in endosperm during the transition from pre-storage to storage phase with a peak of gene expression related to ATP production, photosynthesis and starch accumulation (Sreenivasulu *et al.*, 2004). Subsequently, expression of storage protein transcripts encoding various classes of hordeins and protease inhibitors are found to be tightly co-expressed in endosperm tissue. Activation of the above mentioned specific storage processes is not seen in the embryo; instead, activation of the triacyl glycerol (TAG) lipid biosynthetic pathway and the synthesis of globulins are noted.

Key regulators expressed along with starch, storage protein and oleosin transcripts during barley seed development

Cluster analysis of expression data revealed robust and dynamic patterns of tissue and/or stage-specific gene activity patterns in filial endosperm and embryo tissues representing key transcriptional programs active during seed maturation. Further, co-expressed gene clusters characteristic for the particular physiological status of a tissue during storage processes in developing seeds have been identified. These clusters contain groups of functionally related genes together with possible regulators thus allowing the deduction of new putative regulatory networks, which probably contribute to tissue-specific physiological events. We focused on regulators such as transcription factors, kinases and signaling networks involved in starch/storage protein accumulation in endosperm tissue and lipid accumulation in the embryo conferring desiccation tolerance. The transcriptome data of endosperm and embryo development obtained in two-day intervals from custom made high-density 12,000 cDNA macro array of barley seed (Sreenivasulu *et al.*, 2006) were analyzed with the J-express software. This analysis tool ranks other functionally related genes according to a strong expression correlation match with query genes during endosperm and embryo development. We used this approach to find regulatory genes tightly co-expressed with ADP-glucose pyrophosphorylase (AGP) small subunit transcript, B1 hordein storage protein transcript and oleosin transcripts representing tri-acyl glycerol biosynthesis. The results shed clues about the possible involvement of key regulators involved in triggering starch, storage protein and lipid accumulating transcripts during seed development. Predicting *cis* elements in promoter regions (the orthologous rice genes) of co-expressed functionally related genes provided further evidence for the deduced regulatory networks.

Prominent transcription factors expressed together with key genes of starch biosynthetic pathway during the storage phase of endosperm belong to the bZIP (5), C3H/C3HC4 (3), MYB (2), chromatin remodelling factors (2), WRKY (1), SBP (1), ABI3/VP1 (1), B3 (1) and unclassified transcription factors (14). Among well characterized bZIP class transcripts we observed expression of ABA response element binding factors (ABF3 and ABI3) co-expressed with SNF1 and starch biosynthetic genes during first storage peak of endosperm development. We found ABRE elements in the SNF1 kinase promoter region, which indicates a possibly positive role of ABA in triggering these regulators. ABA positively interacts with sugar signalling pathways in controlling key starch biosynthesis genes via SNF1 kinase (Halford and Paul, 2003). In potato tubers, suppression of SNF1 by antisense strategies resulted in altering sucrose synthase activity (Purcell *et al.*, 1998) and pointed in addition to a key role in redox modulation of AGP (Tiessen *et al.*, 2003). Based on our current studies, we also propose that SNF1 expression in endosperm is mediated by ABA via ABF3/ABI3, which in turn might be responsible in regulating key starch biosynthesis genes. It is also noteworthy that among less abundant transcription family members, we observed co-expression of the ABI3/VP1 homologue (a B3 family member), a leafy cotyledon (LEC) CCAAT binding factor from the HAP family and a DET1 family member along with starch biosynthetic genes during endosperm development. Based on genetic studies, participation of ABI3/VP1, LEC and DET transcription factors in promoting reserve accumulation is well documented in dicots (Finkelstein *et al.*, 2002).

Transcription factors preferentially co-expressed with B1 hordein storage protein transcripts during main storage phase of endosperm development includes NAC (5), chromatin remodeling (4), DOF (2), bHLH (2), JUMONJI (1) and unknown transcription factors (16). It was shown recently that two DOF

transcription factors (SAD and BPBF) act as activators of B1 storage protein genes (Díaz *et al.*, 2005) during the maturation phase. During our current study (Sreenivasulu *et al.*, 2006) we found: (i) co-expression of two DOF family members, SAD and BPBF transcription factors, along with B1 hordein storage protein transcripts; and (ii) enrichment of prolamin box *cis* elements in upstream sequences of (rice) prolamin class storage protein genes (D, B1 and B3 hordeins).

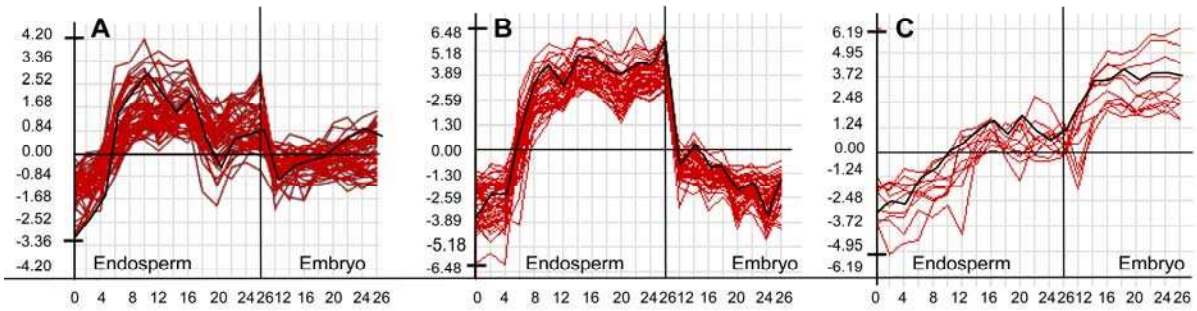

Fig. 1. Regulatory genes co-expressed together with (A) ADP-glucose pyrophosphorylase small subunit transcript involved in starch biosynthesis, (B) hordein B1 storage protein transcript and (C) oleosin 1 transcript from triacyl glycerol biosynthesis. The expression profiles of AGPase, hordein B1 and oleosin 1 are shown in black and co-regulated transcription factor profiles are shown in grey. The developmental scale (days after flowering) is on the X-axis and gene expression levels (log transformed values) on the Y-axis.

Transcription factors co-expressed along with oleosin transcripts encode ABF transcription factor from bZIP class, a dehydration responsive element binding factor from the AP2/EREBP family and two unclassified transcription factors. ABA-responsive genes active during embryo maturation appear to be involved in the control of lipid storage and desiccation tolerance. Our data also suggest that triggering of desiccation related genes is controlled by an ABA independent but interactive pathway via DREB transcription factors. Participation of the above mentioned transcription factor family members in triggering expression of TAG biosynthesis and desiccation related genes via an ABA-dependent and ABA-independent manner has been discussed in detail recently (Sreenivasulu *et al.*, 2006). In summary, described correlative evidence suggests the existence of a broad network of regulators triggering tissue specific physiological events in the endosperm and embryo during ongoing storage product accumulation.

A genetical genomics approach to track favourable alleles related to seed traits in barley

In an attempt to gain insight into the genetic factors underlying important seed traits a set of barley introgression lines [*Hordeum spontaneum* (HS) segments in a *H. vulgare* cv. 'Brenda' background] was analyzed. Comparison of two populations 'Brenda' × HS213 (Li *et al.*, 2005) and 'Brenda' × HS584 (Li *et al.*, 2006) for grain yield and relevant parameters such as spike number per plant, spikelet number per spike, grain number per spike and thousand grain weight resulted in the identification of 15 QTLS and 80 QTLs, respectively. The identified QTLs were verified to be relatively independent of environmental factors (Li *et al.*, 2005, 2006). These characterised introgression lines shown to affect seed traits in barley are a valuable asset as an alternative resource to genome-wide insertion mutants and open up new ways for the discovery of genetic networks contributing to phenotypes. Gene expression levels are regarded as phenotypes and eQTLs calculated and mapped on the chromosomes [eQTL's are peaks of statistical significance in a genome-wide scan for linkage between genetic markers and transcript abundances (Gibson and Weir, 2005)]. In addition to correlative genetic network identification methods used to analyse introgression lines with uniform genetic backgrounds we additionally used a genetical approach to explore the contribution of genetic variation on gene transcription and yield related traits during barley seed development. The combined data from the different approaches will finally permit to nominate candidate genes involved in trait expression and to define regulatory networks.

Detection of putative genetic networks underlying seed traits in barley introgression lines. To assess the genetic architecture of transcript regulation for grain yield, a custom made 12 k barley seed array was used to systematically explore variance in mRNA abundance in developing seeds (4, 8, 16 and 25 DAF) of 92 doubled haploid introgression lines (BC3-DH) generated based on two populations 'Brenda' × HS213 (Li *et al.*, 2005) and 'Brenda' × HS584 (Li *et al.*, 2006). Initially we focused on 42 introgression lines generated based on the first population 'Brenda' × HS213 and addressed the key question whether there is variation of transcript abundance between the 42 introgression lines. Measuring the transcript variation between 'Brenda' *vs* all 42 introgression lines across 4 developmental time points covering pre-storage phase (4 DAF), intermediate phase (8 DAF), early storage phase with the peak of starch synthesis (16 DAF) and later storage phase with ongoing storage protein and lipid accumulation (25 DAF) allowed us to discern that putative genetic networks differ between lines and to identify important regulators controlling main storage events during barley seed development. To place the observed variance of transcript abundance into a broader physiological context of storage product accumulation we implemented correlation analysis of 1,440,000 expression points and defined pathway networks influencing seed traits. The results allowed us to identify groups of introgression lines showing altered changes in storage pathways. These correlation networks will be extended to additionally measured data of metabolite, enzymatic and protein levels.

Detection of expression(e)QTLs for seed traits in barley introgression lines. Based on the array analysis data, we performed an eQTL analysis of the introgression lines (BC3-DH) described above. In total 179 eQTLs were detected by applying a 1% genome wide significance threshold derived by 1000 permutations (for details see Doerge and Churchill, 1996) and 1347 on a 5% genome wide significance threshold including seven eQTL hotspots. These hotspots identified on chromosomes 2H, 4H, 5H and 6H correspond in partially with yield-QTLs in the 'Brenda' × HS213 (Li *et al.*, 2005) and 'Brenda' × HS584 (Li *et al.*, 2006) populations. The coincidence of yield- and expression QTL points towards the participation of pleiotropic trans-regulatory factors in barley seed development. We are currently developing additional molecular markers in order to map the position of these eQTLs in the barley genome at higher resolution to differentiate, for instance, between trans- or cis-regulatory effects on gene expression. As a future outlook the candidate regulatory genes involved in seed filling could be used to proof gene-to-trait relationships to confirm their role in grain quality and yield.

References

Díaz, I., Martínez, M., Isabel-LaMoneda, I., Rubio-Somoza, I. and Carbonero, P. (2005). *Plant Journal*, 42: 652-662.

Doerge, R.W. and Churchill, G.A. (1996). Permutation tests for multiple loci affecting a quantitative character. *Genetics*, 142: 285-294.

Finkelstein, R.R., Gampala, S.S.L. and Rock, C.D. (2002). *Plant Cell*, 14: S15-S45.

Gibson and Weir (2005). *Trends in Genetics*, 21: 616-623.

Halford, N.G. and Paul, M.J. (2003). *Plant Biotechnology Journal*, 1: 381-398.

International Rice Genome Sequencing Project (2005). *Nature*, 436: 793-800.

Jansen, R.C. and Nap, J.P. (2001). *Trends in Genetics*, 17: 388-391.

Khush, G.S. (1999). *Genome*, 42: 646-655.

Li, J., Huang, X.Q., Heinrichs, F., Ganal, M.W. and Roder, M.S. (2005). *Theoretical & Applied Genetics*, 110: 356-363.

Li, J., Huang, X.Q., Heinrichs, F., Ganal, M.W. and Roder, M.S. (2006). *Genome*, 49: 454-466.

Purcell, P.C., Smith, A.M. and Halford, N.G. (1998). *Plant Journal*, 14: 195-202.

Sreenivasulu, N., Altschmied, L., Radchuk, V., Gubatz, S., Wobus, U. and Weschke, W. (2004). *Plant Journal*, 37: 539-553.

Sreenivasulu, N., Radchuk, V., Strickert, M., Miersch, O., Weschke, W. and Wobus, U. (2006). *Plant Journal*, 47: 310-327.

Tiessen, A., Prescha, K., Branscheid, A., Palacios, N., McKibbin, R., Halford, N.G. and Geigenberger, P. (2003). *Plant Journal*, 35: 490-500.

Wobus, U., Sreenivasulu, N., Borisjuk, L., Rolletschek, H., Panitz, R., Gubatz, S. and Weschke, W. (2005). *Recent Research Development in Plant Molecular Biology*, 2: 1-29.

Zhang, H., Sreenivasulu, N., Weschke, W., Stein, N., Rudd, S., Radchuk, V., Potokina, E., Scholz, U., Schweizer, P., Zierold, U., Langridge, P., Varshney, R.K., Wobus, U. and Graner, A. (2004). *Plant Journal*, 40: 276-290.