PLANT 2030

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Highlights der Angewandten Pflanzenforschung

Zum Abschluss der Förderinitiative
PFLANZENBIOTECHNOLOGIE DER ZUKUNFT
des Bundesministeriums für Bildung und Forschung

Projektberichte

Liebe Leserinnen und Leser,

die folgenden Berichte stammen aus den Projekten der Förderprogramme PFLANZENBIOTECHNOLOGIE DER ZUKUNFT und PLANT-KBBE I-III des Bundesministeriums für Bildung und Forschung (BMBF). Die Berichte wurden für PLANT 2030 von den jeweiligen Koordinatorinnen und Koordinatoren nach Beendigung der Projekte zur Veröffentlichung angefertigt. Jedes dieser Projekte stellt eine eigene Erfolgsgeschichte dar, die hier nochmals nachvollzogen werden kann. Die Berichte liegen teilweise in englischer, teilweise in deutscher Sprache vor. Die Verantwortung für die Wahrung der Urheber- und Persönlichkeitsrechte für die Abbildungen und Texte liegt allein bei den Personen, die der PLANT 2030 Geschäftsstelle diese Berichte übergeben haben. Die Berichte sind nicht zur Verbreitung und weiteren Veröffentlichung gedacht. Zusätzlich stehen weitere Informationen zu diesen und anderen Projekten in der Projektdatenbank von PLANT 2030 auf www.Pflanzenforschung.de zur Verfügung.

Ihre PLANT 2030 Geschäftsstelle

Potsdam, im August 2018

PLANT-KBBE – PLant Alliance for Novel Technologies - towards implementing the Knowledge-Based Bio-Economy in Europe

Die Projekte der drei Ausschreibungen der PLANT-KBBE-Initiative wurden zwischen 2009 und 2014 durchgeführt. Eine vierte Ausschreibung erfolge im Jahr 2012. Die Fördermaßnahme geht auf eine gemeinsame Initiative vom Bundesministerium für Bildung und Forschung (BMBF) zusammen mit dem Ministerium für Forschung und Innovation (DGRI) in Frankreich – vertreten durch die nationale Forschungsagentur (ANR) –, dem Ministerium für Wirtschaft und Wettbewerbsfähigkeit in Spanien (MINECO), sowie dem Ministerium für Wissenschaft und Innovation in Portugal (MCTES) – vertreten durch die nationale Wissenschaftsorganisation (FCT), zurück.

Coverbild: © KWS SAAT SE

Inhalt

FRAGENOMIC Genetical genomics for improving strawberry fruit nutritional quality

Juan Muñoz Blanco[1], Amparo Monfort[2], Wilfried Schwab[3], Béatrice Denoyes-Rothan[4], Alexander Pierron-Darbonne[5], Philippe Chartier[6]

1 Departamento de Bioquímica y Biología Molecular, Edificio Severo Ochoa (C-6; Ala Norte), Universidad de Córdoba, 14071 Cordoba, Spain
2 Institut de Recerca i Tecnologia Agroalimentàries (IRTA), Laboratori de Genètica Molecular Vegetal CSIC-IRTA, 08348 Cabrils, Spain
3 Biotechnology of Natural Products, Technische Universität München, Liesel-Beckmann-Str. 1, D-85354 Freising, Germany.
4 Unité de Recherche des Espèces Fruitières et de la Vigne (UREFV), INRA CR Bordeaux, BP 81, 33883, Villenave d'Ornon Cedex, France
5 Planasa S.A, Crta de San Adrián, Km 1, Valtierra, Navarra, Spain
6 CIREF Création Variétale Fraise, Fruits Rouges, Maison Jeannette, 24140 Douville, France

Abstract
Phenolic compounds have drawn attention due to their potential nutritional benefits. The fundamental chemical reactions of the phenolics pathways in plants have been elucidated but the flux through the pathway and the control of their accumulation is not that well known. A strawberry (*Fragaria* × *ananassa*) microarray platform was used to analyze gene expression patterns associated with the accumulation of phenolics such as phenylpropanoids, flavonoids, and anthocyanins in strawberry fruit. The examination of the transcript levels, coupled with metabolite profiling data obtained by liquid chromatography hyphenated with mass spectrometry from different commercial varieties, was undertaken to identify genes whose expression correlated with altered phenolics composition. Comparative microarray analyses yielded fifteen genes that were differentially (> 200-fold) expressed in phenolics-poor versus phenolics-rich varieties. The results were validated by heterologous expression of a candidate gene, which showed the highest altered expression level (> 900-fold). The encoded protein is involved in lignin formation during strawberry fruit ripening and quantitative trait locus (QTL) analysis indicated that the genomic region of the candidate gene is associated with the fruit color trait. The result highlights the competition of the different phenolics pathways for common precursors. The list of candidate genes which was further validated by transient overexpression and down-regulation provides information about genes that are likely to impact polyphenol accumulation in strawberry fruit and could be valuable molecular markers to select phenolics-rich germplasm.

Introduction

Strawberry (Fragaria x ananassa Duch.) is an important soft fruit in the temperate climate zones of the world (Shulaev *et al.*, 2011). This attractive and tasty fruit, with its high nutritional value, is either sold in the fresh market or as a raw material for the food processing industry. Strawberry fruit flavor is highly appreciated and belongs to the three most important flavors worldwide (Raab *et al.*, 2006). However, the strawberry is a particularly perishable fruit, with the result that its shelf-life is short and its quality suffers due to rapid deterioration in unsuitable storage conditions.

Nearly all European cultivars of Fragaria x ananassa Duch derive from the original crossing of two American wild types: Fragaria chiloensis and Fragaria virginiana. Almost 1000 cultivars were mainly created by cultivar crossing

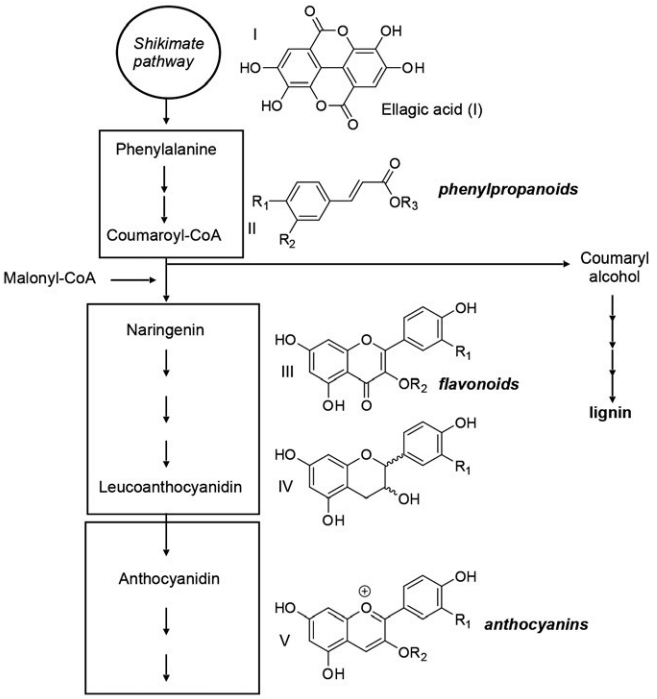


Fig. 1: *Schematic scheme of the shikimate-phenylpropanoid-flavonoid-anthocyanidin-lignin pathway.*

using this genetic basis, consequently the genetic basis of the main cultivars in Europe became very small.

Expression patterns of genes encoding different enzymes likely to be involved in the strawberry fruit ripening process have been extensively studied (Aharoni and O'Connell 2002). They include enzymes acting on the pectin fraction of the cell wall and putatively involved in fruit firmness such as β-galactosidase, pectin methylesterase, endo- and exo-polygalacturonase, and pectate lyase. Transgenic plants with reduced pectate lyase expression (< 30% of the control plant) showed higher fruit-firmness but fruits harvested from transgenic plants with reduced endo-1,4-ß-glucanase expression did not show a clear change in firmness. Moreover a gene coding a D-galacturonate reductase has been demonstrated to be related with the production of ascorbic acid in ripening strawberry fruits and two strawberry ripening related genes SAAT (coding an alcohol acyl-CoA tansferase) and FaQR (encoding a quinone reductase) involved in aroma production in ripen strawberry fruits were cloned and physiologically characterized.

The knowledge of the genetic determinants affecting both fruit quality and agronomical traits is crucial for the sustained improvement of crops (Capocasa *et al.*, 2008). Although the strawberry fruit is highly appreciated for its tasty flavor and nutritional value fruit quality attributes have been reduced or even lost because breeding of modern cultivars has mainly concentrated on agronomical traits such as fruit size and harvest yield. Therefore, enhancement of strawberry flavor, shelf life, and nutritional quality has become an important criterion in current breeding programs, even when these quality attributes are regulated by a complex genetic background and are frequently associated with negative agronomic characters (Zorrilla-Fontanesi *et al.*, 2011). The challenge for breeders, who want to produce fruit with high nutritional value while maintaining an outstanding quality, is not only the knowledge of the single trait but also what is affecting the variation and how different traits are correlated together. Breeding to increase one or more beneficial phytochemicals in fruit is likely to be achievable since many different fruit traits have already been successfully modified with breeding strategies.

Strawberry phenolics have drawn increasing attention due to their potential nutritional benefits (Aaby *et al.*, 2007). Anthocyanins, flavonoids, and phenylpropanoids are among the major phenolics that accumulate in ripe strawberry (Fragaria × ananassa) fruit. Anthocyanins give rise to the red color of the fruit, which attracts frugivores that help to disperse seeds while flavonols are thought to function as protectants against UV-B light. Proanthocyanidins contribute to defense and stress resistance and 1-O-acyl-glucose esters of cinnamic, 4-coumaric, and caffeic acids (phenylpropanoids) may serve as energy-rich substrates in plant secondary metabolism (Lunkenbein *et al.*, 2006) whereas ellagic acid may play a role in protection from predation and in plant growth regulation.
Although genetic variation of strawberry cultivars is low, the total phenolics content and composition strongly differ between genotypes. Thus, cultivated strawberry varieties are a suitable model to identify new factors regulating flux through the phenolics pathway. Besides, the genome sequence of Fragaria vesca, one of the progenitors of strawberry, can be used as a genomic reference for the genus (Shulaev *et al.*, 2011).

In this project, an examination of the transcriptome analyzed by microarray analysis, coupled with metabolite profiling analysis performed by LC-MS analysis of different strawberry genotypes, was undertaken to reveal genes whose expression levels correlated with altered phenolics composition. The suitability of the approach was confirmed by the functional characterization of candidate genes that showed the highest transcript level difference in genotypes with contrasting phenolics content. The role of the candidate genes in color and total polyphenols production was highlighted by the identification of quantitative trait loci (QTLs) linked to color in the region where the candidate genes are located and in two segregation populations. Chemical ana-

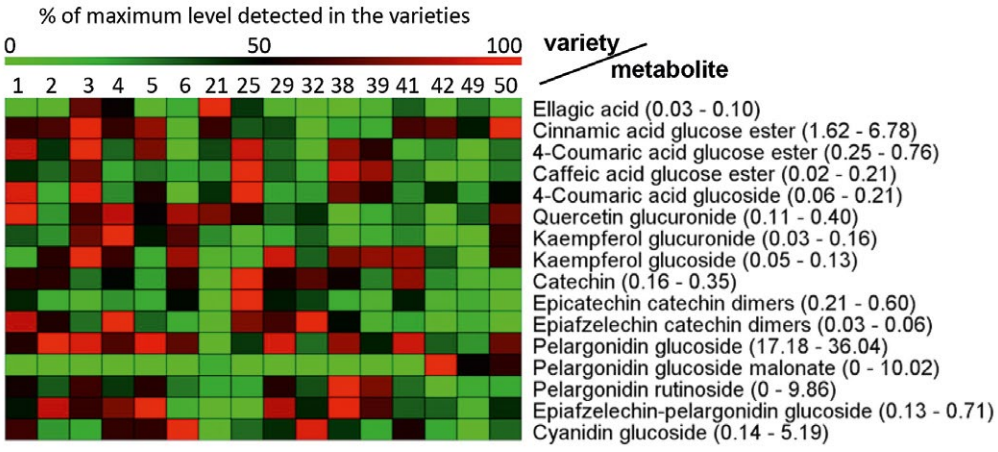


Fig. 2: *Quantification of strawberry fruit phenolics. Quantification of metabolites was performed by LC-MS analysis as ‰ equivalents of the dry weight (relative concentration). Heat map presentation of the metabolite levels in different varieties. Varieties with the highest and lowest levels of individual metabolites are shown in red and green, respectively. Minimum and maximum level of individual metabolites are shown in parentheses.*

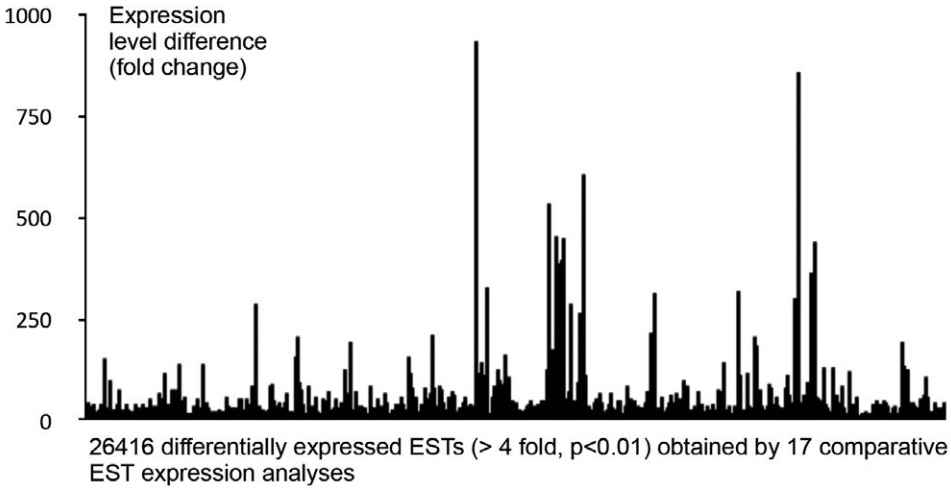


Figure 3. *Total expressed sequences tags (EST) which showed differential expression (> 4 fold, p<0.01) in 17 microarray analysis. Microarray analyses were performed on different strawberry varieties that produced significantly contrasting levels of phenolics as determined by metabolite profiling analysis.*

lysis of transgenic plants confirmed the metabolic interaction between anthocyanin and lignin biosynthesis.

Results

The major known phenolic metabolites of strawberry fruits (Figure 1) of a number of phenotypically different strawberry varieties were quantified in the positive and negative MS mode by the internal standard method by liquid chromatography (LC)-mass spectrometry (MS). Each variety accumulated a unique concentration pattern of metabolites (Figure 2).

Variety #49 accumulated relatively low levels of various soluble phenolics in the fruits, whereas #3 produced relatively high amounts of metabolites. Some metabolites, such as pelargonidin glucoside, accumulated to high levels in a number of varieties, whereas its malonylated derivative was observed in only a few genotypes. The levels of the individual phenolics were used to calculate the total amount of phenylpropanoids, including ellagic acid, flavonoids, anthocyanins, and total phenolics Variety #5 and #3 were superior genotypes with regard to the total amount of phenolics due to the high levels of anthocyanins and phenylpropanoids. In contrast, #49 and #21 produced only low levels of soluble phenolics.

In addition, unique comparative microarray analyses were conducted on 17 RNA samples isolated from receptacles of 12 different varieties that contrasted in their levels of one of 20 soluble phenolic metabolites. In total, 26,416 differentially expressed ESTs (more than 4-fold, P < 0.01; Figure 3) were detected. Most of the transcript levels differed by less than 50-fold, but some mRNAs levels exceeded 200-fold and reached up to 925-fold expression ratios. The latter were selected as top candidate genes and their protein-coding sequenes deduced from the genome sequence of F. vesca.

Due to EST redundancy, gene 19544, annotated as a putative heme-containing peroxidase 27 gene (Veitch 2004; PRX27), showed up five times among the most highly differentially expressed nucleotides. The function of this gene was analyzed in detail.

The expression pattern of PRX27 was analyzed by qPCR to understand the physiological roles. Significant transcript levels were detected in roots and red, ripe fruits, in accordance with the results of the microarray analyses. FaPRX27 transcript levels were negligible in other tissues. The full-length complementary DNA (cDNA) of FaPRX27 was cloned and introduced into Escherichia coli cells. Cell extracts after induction of the recombinant PRX27 protein production contained the 61-kD fusion protein FaPRX27-GST (for glutathione S-transferase). Soluble protein fractions were used in colorimetric peroxidase enzyme assays and enzymatic assays were additionally subjected to LC-MS analysis to detect potential products. The starting material ferulic acid and coniferyl alcohol was almost completely consumed when incubated with FaPRX27. Caffeic acid, guaiacol, vanillin, and 3,4-dihydroxybenzaldehyde were also transformed. This result clearly points to a function of FaPRX27 in lignin biosynthesis. In addition, total lignin content was analyzed in fruits of genotypes that produced extremely high and low levels of total phenolics. Genotypes #3 and #5 accumulated significantly lower concentrations of the lignin polymer than #21 and #25, consistent with the hypothesis that FaPRX27 produced the cell wall polymer at the expense of soluble phenolic compounds.

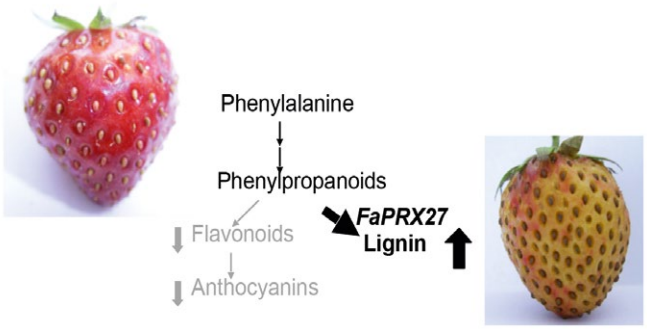


Fig. 4: Simplified sheme to show the proposed effect of FaPRX27 expression on the total amount of soluble phenolics.

Quantitative trait loci (QTLs) were detected for all quantitative traits by composite interval mapping for either Capitola or CF1116 for a French strawberry population and Camarosa or Dover for a Spanish population. In the French segregating population, in the region of FaPRX27, QTLs linked to total polyphenol and total flavonoids were detected on the male map. In the Spanish population, the PRX27 gene was co-localized with QTLs for epiafzelechin-pelargonidin glucoside and total flavonoids. In both populations, in the region of FaPRX27, a QTL linked to visual fruit color decrease was detected.

Finally, chalcone snthase (CHS) expression was down-regulated during strawberry fruit ripening by agro-infiltration of an inverted hairpin RNAi construct to confirm the competition of the anthocyanin/flavonoid biosynthesis and lignin pathways for common substrates. One-half of a strawberry fruit was agro-infiltrated, whereas the second half remained untreated. Down-regulation of CHS, as shown by the loss of pigmentation, was accompanied by a significant increase in lignin content, which correlates with enhanced firmness of the fruits. Transcript analyses by qPCR affirmed the reduced levels of CHS mRNA levels due to RNAi-mediated gene silencing and revealed the strong induction of FaPRX27 expression upon agro-infiltration.

The results clearly show that the reduced flux through the anthocyanin/flavonoid pathway (Figure 1) along with increased FaPRX27 transcript levels significantly enhanced the lignin content and associated firmness at the expense of anthocyanins in strawberry fruit (Figure 4). Manipulation of the lignin biosynthesis pathway enzymes could yield strawberry fruit with increased levels of soluble phenolic compounds such as flavonoids and anthocyanins and thus would provide fruit with increased potential health benefits.

Conclusion

Correlation of metabolite levels with the transcript patterns of different strawberry genotypes revealed candidate genes that might affect the production of flavonoid and anthocyanin in strawberry fruit. Neither transcription factors nor structural genes of the phenolics pathway appeared among the top 20 candidates. The majority of the novel genes have not been implicated in the secondary metabolism. A presumed peroxidase gene FaPRX27 was characterized in detail to verify the approach. FaPRX27 encodes a functional protein which is required for lignin formation in ripening strawberry fruit and thus competing with flavonoid/anthocyanin pathway enzymes for common substrates. Follow-up studies have already shown that other candidates on the list also affect phenolics accumulation (Song *et al.*, 2015).

Materials and methods

- *Experimental procedures have been performed as described in Ring et al., 2013; Medina-Puche et al., 2014; Fischer et al., 2014 and Song et al., 2015.*

Quellen/Referenzen

- *Aaby K, Ekeberg D, Skrede G (2007) Characterization of phenolic compounds in strawberry (Fragaria × ananassa) fruits by different HPLC detectors and contribution of individual compounds to total antioxidant capacity. Journal Agricultural and Food Chemistry 55, 4395–4406.*
- *Aharoni A, O'Connell AP (2002) Gene expression analysis of strawberry achene and receptacle maturation using DNA microarrays. Journal of Experimental Botany 53, 2073–2087.*
- *Capocasa F, Diamanti J, Tulipani S, Battino M, Mezzetti B (2008) Breeding strawberry (Fragaria × ananassa Duch) to increase fruit nutritional quality. Biofactors 34, 67–72.*
- *Fischer TC, Mirbeth B, Rentsch J, Sutter C, Ring L, Flachowsky H, Habegger R, Hoffmann T, Hanke MV, Schwab W (2014) Premature and ectopic anthocyanin formation by silencing of anthocyanidin reductase in strawberry (Fragaria x ananassa). New Phytologist 201, 440-451.*
- *Lunkenbein S, Bellido ML, Aharoni A, Salentijn EMJ, Kaldenhoff R, Coiner HA, Muñoz-Blanco J, Schwab W (2006) Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry. Plant Physiology 140, 1047–1058.*
- *Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero JL, Schwab W, Muñoz-Blanco J, Blanco-Portales R (2014) MYB10 plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of Fragaria x ananassa fruits. Journal of Experimental Botany 65, 401-417.*
- *Raab T, López-Ráez JA, Klein D, Caballero JL, Moyano E, Schwab W, Muñoz-Blanco J (2006) FaQR, required for the biosynthesis of the strawberry flavor compound 4-hydroxy-2,5-dimethyl-3(2H)-furanone, encodes an enone oxidoreductase. Plant Cell 18, 1023-1037.*
- *Ring L, Yeh SY, Hücherig S, Hoffmann T, Blanco-Portales R, Fouche M, Villatoro C, Denoyes B, Monfort A, Caballero JL, Muñoz-Blanco J, Gershenson J, Schwab W (2013) Metabolic interaction between anthocyanin and lignin biosynthesis is associated with peroxidase FaPRX27 in strawberry fruit. Plant Physiology 163, 43-60.*
- *Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP, et al. (2011) The genome of woodland strawberry (Fragaria vesca). Nature Genetics 43, 109–118.*
- *Song C, Ring L, Hoffmann T, Huang FC, Slovin J, Schwab W (2015) Acylphloroglucinol biosynthesis in strawberry fruit. Plant Physiology 169, 1656-1670.*
- *Veitch NC (2004) Structural determinants of plant peroxidase function. Phytochemical Reviews 3, 3–18.*
- *Zorrilla-Fontanesi Y, Cabeza A, Domínguez P, Medina JJ, Valpuesta V, Denoyes-Rothan B, Sánchez-Sevilla JF, Amaya I (2011) Quantitative trait loci and underlying candidate genes controlling agronomical and fruit quality traits in octoploid strawberry (Fragaria × ananassa). Theoretical and Applied Genetics 123, 755–778.*

PLANT KBBE II

SUSTAINPINE Genomic tools in Maritime pine for enhanced biomass production and sustainable forest management

Subproject: Natural variation of candidate genes involved in wood formation, growth and response to environmental stress

Barbara Vornam[1], Konstantin V. Krutovsky[1] and Reiner Finkeldey[2]

1 University of Goettingen, Buesgen-Institute, Department of Forest Genetics and Forest Tree Breeding, Buesgenweg 2, 37077 Goettingen
2 University of Kassel, Moenchebergstr. 19, 34125 Kassel

The forestry has to cope with progressive impact of the global climate change in a context of increasing economic competition among industrial forest areas. The aim of this project is the identification of key genes determining adaptive traits in conifers, which are crucial for forest productivity, conservation and management. The investigation is focused on maritime pine (*Pinus pinaster* Ait), the most advanced conifer model species for genomic research in Europe and the most widely planted species in France, Spain and Portugal. Results on this species could be easily transferred to closely related *Pinus* species and other economical and/or environmentally important conifer species.

The genetic diversity is a prerequisite of forest trees in order to keep their adaptive potential and capacity in the changing environmental conditions. The objective of this subproject was to analyse the naturally occurring nucleotide variation of candidate genes related to growth, lignin biosynthesis, wood formation and environmental stress in order to find alleles and SNPs (single nucleotide polymorphisms) under selection.

Identification of candidate genes

Based on full-length cDNA sequences (available at http://www.scbi.uma.es/sustainpine and Canales *et al.* 2014) 50 candidate genes involved in wood formation, growth and response to environmental stress were selected and identified on the genomic DNA level by Sanger sequencing in order to create reference sequences. The sequences of the amplification products represent 32 full-length genes, the promoter sequences of the *LipT*, *LEC*, *CCoAOMT* and *SPR1* genes and 18 gene fragments. Structure, function and length of the identified genes are summarized in Table 1. Among the analysed sequences (in total about 50 000 nucleotides) one alternative splicing site was found within the sequence of *Korrigan1* (endo-1,4-beta-glucanase), a gene involved in cellulose biosynthesis and therewith in the properties of wood (secondary xylem).

In order to find alleles and SNPs under positive selection seeds from 80 trees from 8 populations representing the natural distribution range of the species were sampled (Fig.1), and the DNA was isolated from the haploid megagametophytes of the seeds and amplified with the corresponding primers of the selected candidate genes. SNP detection in the population samples was performed by sequencing using the NGS platform of Illuminia. For each sample equimolar amounts of all amplicons were pooled and fragmented using sonication, then the individually barcoded sequencing libraries were prepared and sequenced in a single pool according to the Illumina protocol. After a paired-end sequencing run the data were analysed using the SNP detection function in the CLC Genomics Workbench software package (CLCbio, Aarhus, Denmark) with the following alignment options: 1) Since haploid tissue (DNA from the megagametophytes) was analysed only one allele was allowed in a sample for a SNP, 2) and only those sequences that have 90% homology with a reference sequence were analysed. Due to these options no indels were included in the analysis, but the detected SNPs should be more reliable.

Nucleotide diversity of the candidate genes in the natural populations

In total 354 SNPs were found for the candidate genes in the 8 analysed natural populations. Depending on the gene the nucleotide variation varied from completely monomorphic (*Gludump*3, *GluDeCa*, *LipT* (coding sequence), *Lac*, und *AMP1*), to weakly (e.g. *GS*1b, *DOF*5, *Korrigan*, *LEC*, *WOX*2) and highly polymorphic (e.g. *CHS*1, *Nod*, *GP*x). The nucleotide variation of the analysed polymorphic candidate genes is summarized in Table 2. The mean nucleotide diversity was relatively high – 1 SNP every 116 bp, and any randomly selected pair of haplotypes from a population differed by one site per 1000 bp on average ($\pi = 0.0011 \pm 0.0001$). Similar results were found in earlier studies of *Pinus pinaster* (DIGENFOR, BMBF 0313156 and Pot *et al.* 2005), Douglas-fir (Krutovsky and Neale 2005) and *Pinus taeda* (González-Martínez

SUSTAINPINE Genomic tools in Maritime pine for enhanced biomass production and sustainable forest management
Subproject: Natural variation of candidate genes involved in wood formation, growth and response to environmental stress

Tab. 1: *Structure, function and length of the identified candidate genes*

Candidate gene	comment	structure (5'→ 3')	involved in	Length PCR product
Glutamine synthetase (GS1b)	partial	E	C and/or N metabolism	900
Glutamine synthetase (GS1a)	partial	I / E	C and/or N metabolism	800
Dof transcription factor (DOF5)	full length	E	C and/or N metabolism	917
Glutamate decarboxylase (GluDeca)	partial	I / E / I / E	C and/or N metabolism	2500
Isocitrate dehydrogenase (ICDH)	partial	E	C and/or N metabolism	1133
Arginase (ARS20)	partial	I / E gap I / E	C and/or N metabolism	1400
Glutamate synthase Fd-GOGAT Genomic	partial	E / I / E	C and/or N metabolism	910
antimicrobial peptide 1 (AMP1)	full length	E / 3'utr	C and/or N metabolism	546
antimicrobial peptide 2 (AMP2)	full length	E / 3'utr	C and/or N metabolism	550
Sucrose-Synthase (SuSy)	partial	gap / I / E / I / gap / E / gap	C and/or N metabolism	8000
Glutamin dumper 5	full length	E	C and/or N metabolism	337
Glutamin dumper 3	full length	E	C and/or N metabolism	330
psbO	full length	E / I / E	C and/or N metabolism	1112
Phosphoenolpyruvat carboxylase (PepCaKi)	full length	E / I / E / I / E	C and/or N metabolism	911
Myb8	partial	E / I / E / I / E	Wood formation	1152
S-adenosylmethionine synthetase (SAMS)	partial	E	Wood formation	858
Xyloglucan endotransglycosidase (XyGly) 2fragments	full length	E / I / E / I / E / I / E	Wood formation	1100
Scarecrow gene regulator-like (SCL1) 4fragments	partial	E	Wood formation	4553
Korrigan	partial	E / I / E / I / E gap / I / E	Wood formation	1240
Myb1	full length	E / I / E / I / E	Wood formation	1066
Myb14	partial	E / I / E / I / E	Wood formation	840
Naringenin-Chalcone-synthase (CHS 1)	full length	E / I / E gap / I / E	Wood formation	1255
Phenylalanin amoniclyase (PAL)	full length	E / gap / E	Wood formation	2200
Clavata	identified		Wood formation	900
Cinnamyl-alcohol dehydrogenase (CAD) 2fragments	partial	E / I / E / I / E / I	Wood formation	1200
Caffeoyl-CoA O-methyltransferase (CCoAOMT) 2fragments	partial	E / I / E / I / E / I / E	Wood formation	800
Promotor(CCoAOMT)	identified	5'utr	Wood formation	720
Nitrilase ass protein (SPR1)	full length	5'utr / E / I / E / 3'utr	Wood formation	604
transcription factor lim1	full length	gap / I / E	Wood formation	3500
Cellulose synthase CES 2 coated vesicle membrane prot	full length	E / I / E / I / E / I / E	Wood formation	1522
Laccase (Lac)	partial	E / I / E / I / E	Wood formation	1132
CES A1	full length	E / I / E / I / gap / I / E / I / E / I / E	Wood formation	8000
GDP-mannose 4,6-dehydratase (Mur1)	full length	E	Wood formation	1000
Serine-threonine protein kinase (STPK)	full length	E	Sress resistance	1104
drought responsive element binding (DREB) or (Etylrespel)	full length	E	Sress resistance	721
Glutathione-S-transferase (GST)	partial	I / E / I / gap/ I / E / I/ E	Sress resistance	2500
ZIP family metal transporter (ZIPMT)	full length	E / I / E / I / E	Sress resistance	1337
Nodulin mtN3 (Nod)	full length	E/ I / E / I / E / I / E / I / E	Sress resistance	2500
Caspase	full length	E / I / E	Sress resistance	1113
Glutathione peroxidase (GPx)	full length	E/ I / E / I / E / I / E / I / E / I / E	Sress resistance	1665
Lipid transferase (LipT)	full length	E	Plant / embryo development	345
Lipid transferase promo (LipT pr)	full length	5'utr / E	Plant / embryo development	866
Lea1	full length	E / I / E	Plant / embryo development	700
Lea2	full length	E	Plant / embryo development	700
Lea3	full length	E / I / E	Plant / embryo development	500
RabGTPase	full length	gap E / I / E / gap / I / E	Plant / embryo development	6000
Gibberellin receptor like prot (GLP1)	full length	E	Plant / embryo development	974
Leafy cotyledon (LEC)	full length	5'ut r/ E / 3'utr	Plant / embryo development	620
Wuschel-related homeobox2 (WOX2)	full length	E / I / E / I / E	Plant / embryo development	1242

with: E=exon; I=intron; utr=untranslated region; gap=missing sequence

SUSTAINPINE Genomic tools in Maritime pine for enhanced biomass production and sustainable forest management
Subproject: Natural variation of candidate genes involved in wood formation, growth and response to environmental stress

Tab. 2: *Nucleotide variation in 35 genes*

Gene	N	bp	SNP	h	Hd	π	Tajima's_D
SPR9	79	605	6	8	0,591	0,0013	-0,874
Myb1	79	1066	8	12	0,807	0,0014	-0,256
ZIPMT	79	1337	16	32	0,923	0,0024	-0,026
CHS1	79	1255	18	39	0,943	0,0023	-0,656
PAL	79	2107	20	34	0,949	0,0015	-0,761
STPK	79	1104	13	23	0,789	0,0021	-0,494
Glu5	79	274	4	4	0,648	0,0031	0,104
LIM1	79	725	14	13	0,762	0,0021	-1,322
CES2	79	1522	19	25	0,813	0,0021	-0,563
GST	79	1110	11	15	0,567	0,0010	-1,417
CESA1	79	2577	16	29	0,885	0,0010	-0,715
PepCaKi	79	911	7	8	0,316	0,0007	-1,346
Nod	79	1871	27	43	0,962	0,0028	-0,442
RabGTPase	79	1133	7	8	0,507	0,0006	-1,375
Caspase	79	1098	4	5	0,252	0,0003	-1,352
GPx	79	1665	37	43	0,912	0,0038	-0,573
LipT	79	865	3	4	0,293	0,0005	-0,635
CCoAOMT	79	1846	12	15	0,798	0,0008	-1,181
GS1b	79	811	14	14	0,840	0,0024	-0,919
DOF5	79	917	4	5	0,169	0,0002	-1,501
ICDH	79	1133	12	17	0,478	0,0011	-1,500
Mur1	79	999	8	12	0,734	0,0011	-0,863
Myb8	79	1148	4	6	0,391	0,0004	-0,810
SAMS	79	857	5	5	0,616	0,0011	-0,097
DREB	79	721	3	5	0,618	0,0012	0,793
XyGly	79	1182	3	4	0,209	0,0002	-1,214
Fd-GOGAT	79	910	9	7	0,587	0,0014	-0,849
Korrigan	79	1237	1	2	0,097	0,0001	-0,625
AMP2	79	552	6	5	0,477	0,0013	-0,961
CAD	79	1824	8	14	0,760	0,0007	-0,751
LEC	79	619	1	2	0,342	0,0006	0,831
WOX2	79	1242	2	3	0,454	0,0006	1,354
Myb14	79	876	13	28	0,893	0,0034	0,336
psbO	79	1112	5	3	0,050	0,0001	-1,886
GLP1	79	974	4	4	0,367	0,0005	-0,901

Note: N= analysed samples; bp= basepairs; SNP= number of SNPs detected (without singletons);h= number of haplotypes; Hd= Haplotype diversity; π = nucleotide diversity (per site); Tajima's D= neutrality test; estimated using DnaSP5 (Librado and Rozas 2009)

et al. 2006). The nucleotide diversity of synonymous sites and non-coding regions was higher than for non-synonymous cites. Tajima's *D* (Tajima 1989) test was used to study whether the frequency spectrums of nucleotide variation in the population samples were in agreement with neutral expectations. Statistically significant Tajima's *D* was found only for the *psbO* gene. The negative value indicates negative selection or recent changes in the population size, whereas the positive values in case of the *DREB*, *CAD*, *WOX* 2 and *Myb* 14 genes indicated an excess of haplotypes relative to the expectations under mutation-drift equilibrium, assuming balancing selection.

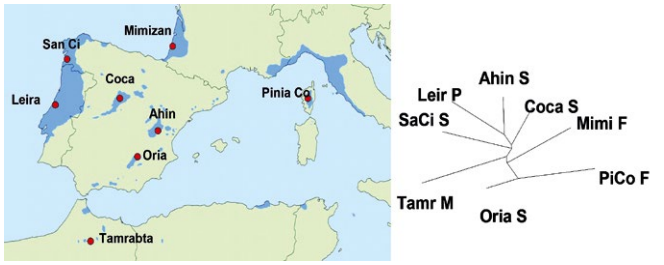


Fig. 1 Geographic position in the natural range of P.pinaster (Alia and Martin, 2003) and the unrooted Neighbor-Joining tree based on genetic distance between the analysed populations

Differentiation of the populations

Among polymorphic genes only *DOF*5, *Mur*1, *LEC* and *psb*O showed no differentiation between the populations.The most polymorphic populations were Tamrabta, from Morocco (183 SNPs; π = 0.0014 ± 0.0002) and Oria from Spain (151 SNPs; π = 0.0014 ± 0.0001). The least polymorphic populations were Pinia, Corse (Italy) and Mimizan, Landes (France) (115-117 SNPs; π = 0.0010 ± 0.0001). Most of the candidate genes showed a higher nucleotide diversity within the population than between the populations except of *SPR*1 (32%), *Gludump*5 (29%), *PepCaKi* (24%) and *GLP*1 (24%), which showed a relative high differentiation between the populations. The genetic differentiation between the populations, averaged over all candidate genes, was not great, but in consensus with their geographic distribution (Fig.1).

Conclusions

For forest trees the predicted increasing rate of global change bears the risk of reduced potential for long-term adaptation compared to annual plants with much higher reproduction rates. Therefore, the analysis ofSNPs in candidate genes and their allele frequency in natural populations should help us better understand evolutionary forces and implement conservation and breeding programs of the species in order to increase the productivity and sustainability of forests.

First, *in silico* comparison of the results described here for the tree species *Pinus pinaster* with the data already described for the German native pine species *Pinus sylvestris* showed that they can be completely transferred. So, the results could be used to describe genes related to growth, wood formation and environmental stress in *Pinus sylvestris*, which becomes more and more important for wood production in the context of climate change.

References

- *Alia, R. and Martín, S. (2003): EUFORGEN Technical Guidelines for genetic conservation and use for Maritime pine (Pinuspinaster). International Plant Genetic Resources Institute, Rome, Italy.6 pages.*
- *Canales J, Bautista R, Label P, Gómez-Maldonado J, Lesur I, Fernández-Pozo N, et al. (2014): De novo assembly of maritime pine transcriptome: implications for forest breeding and biotechnology. Plant Biotechnol. J.12:286–299.*

- *González-Martínez S. C.,Ersoz E., Brown G. R., Wheeler N. C. and Neale D.B. (2006): DNA Sequence Variation and Selection of Tag Single-NucleotidePolymorphisms at Candidate Genes for Drought-StressResponse in Pinus taeda L. Genetics 172: 1915–1926.*
- *Krutovsky K. V. and Neale D. B. (2005): Nucleotide Diversity and Linkage Disequilibrium in Cold-Hardiness- and Wood Quality-Related Candidate Genes in Douglas Fir. Genetics 171: 2029–2041.*
- *Librado, P. And Rozas, J. (2009): DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics.25:1451-1452.*
- *Pot D, McMillan L, Echt Cet al.(2005) Nucleotide variation in genes involved in wood formation in two pine species. New Phytologist, 167, 101–112.*
- *Tajima, F. (1989) Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics, 123, 585-595.*

PLANT KBBE III

CAROMAIZE Maize fortified with high-value carotenoids as raw material for industry and nutritionally enhanced animal feed

Prof. Dr. Gerhard, as coordinator

Institut für Molekulare Biowissenschaften, Goethe Universität Frankfurt

Structure and aims of CAROMAIZE

CAROMAIZE was a transnational Plant KBBE Project. The consortium of CAROMAIZE consisted of three companies, PartnerChip, Evry, France, Piensos del Segre, Balaguer, Spain, Trow Nutrition Deutschland, Burgheim, Germany, three academic partners, Goethe University, Frankfurt, Germany, Universitat de Lleida, Spain, Royal Holloway University of London, UK and two research institutes, INRA, Montpellier, France, PPM Magdeburg.

The whole project was structures into four packages. These were

- the development of two maize prototype by genetic pathway engineering using maize as a dual production platform for ß-carotene and astaxanthin in seed after engineering of astaxanthin synthesis in a high-oil line,
- the analysis of terpenoid content and integrated analysis (transcriptomics, proteomics and metabolomics), •
  the direct use of seeds of ß-carotene line in chicken feeding experiments,
- milling of astaxanthin maize and processing to an oily astaxanthin raw material and its use as an additive in fish feeding trials.

The pipeline structure of the different tasks was as illustrated: Figure 1.

Results of CAROMAIZE

Transgenic maize was used as a production platform for β-carotene and astaxanthin to generate nutrient rich animal feed. The astaxanthin accumulating maize line was engineered by introduction of a β-carotene hydroxylase and a β-carotene ketolase gene for pathway extension, overexpression of a gene for enhanced carotenoid production and a gene knock-down to direct more precursors into the β-branch of the extended ketocarotenoid pathway which ends with astaxanthin. This astaxanthin-accumulating transgenic line was crossed into a high oil-maize genotype in order to increase the storage capacity for lipophilic astaxanthin. The β-carotene maize prototype resulted from our pre-project activities.

Detailed targeted metabolic analysis was carried out with both line including the determination of carotenoids, sterols and γ-tocopherol content either of greenhouse or field-grown material. In the transgenic β-carotene maize, total carotenoid content increased 23-fold including a 1000-fold increase of β-carotene. In parallel, 5-fold higher concentrations of γ-tocopherol were obtained. In contrast, the synthesis of sitosterol, stigmasterol and campesterol was not affected. Although enhancement of carotenoid biosynthesis was targeted to the endosperm of maize kernels, a concurrent up-regulation of sterol and fatty acid biosynthesis in the embryo was measured. Integrated metabolome, transcriptome and proteome was focused on sugar metabolisms as provider of precursors for enhanced carotenoid biosynthesis and on how global seed metabolism adapted to this expanded biosynthetic processes. An integrative model was obtained to explain the metabolic regulation for the increased provision of terpenoid and fatty acid precursors, particularly glyceraldehyde 3-phosphate and pyruvate or acetyl-CoA from imported fructose and glucose.

The model was supported lower glucose and fructose pools and by higher activities of fructokinase, glucose 6-phosphate isomerase, and fructose 1,6-bisphosphate aldolase indicating a higher flux from sugars through the glycolytic pathway.

The β-carotene maize was used directly for chicken feeding trials with hens (egg production) and broilers (for meat) in laboratory trial and an industrial setting. In both cases, the use of ß-carotene maize enhanced pigmentation of meat and eggs and led to a better antioxidative status and a hig-

Fig. 1: Pipeline structure of the different tasks

CAROMAIZE Maize fortified with high-value carotenoids as raw material for industry and nutritionally enhanced animal feed

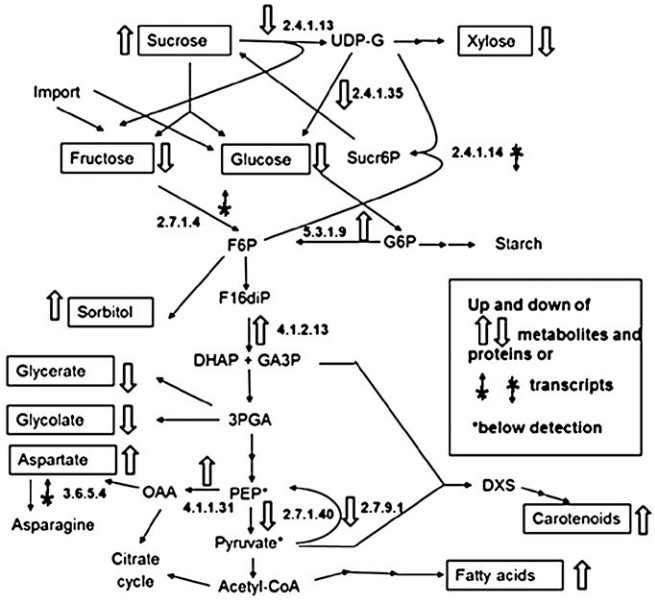


Fig. 2: Sugar metabolism based on integrated omics data

her vitamine A content (by metabolization of β-carotene) in the blood and liver compared to chicken raised on a commercial diet. Our results demonstrate that carotenoid-rich corn maintains poultry health and increases the nutritional value of poultry products without the use of feed additives.

The following figure demonstrates the difference in pigmentation of egg yolk and the intensity of broiler pigmentation:

Since the astaxanthin concentration of the astaxanthin maize was not high enough for direct addition to the fish feed, several precessing steps had to be developed. This included milling and extraction of an astaxanthin oil as raw material in an ethanol-based process. Before its use as an additive in trout feeding trials, a concentration step was necessary. By phase partitioning against several solvents, astaxanthin could be 10-fold concentrated in the oil. The trout feeding experiment with an industrial feed as positive control, a negative control devoid of astaxanthin and a feed with astaxanthin oil supplementation was carried out at the Thünen Institute, Ahrensburg, Germany. As in the previous feeding experiments, health, productivity, meat pigmentation including carotenoid profiles were determined. Although in this setting astaxanthin from maize was not formulated for optimum bioavailability and the applied amount per fish was only 60% the amount in the industrial feed, substantial red pigmentation of the trout filets was achieved with the astaxanthin from our engineered maize.

Significance and economic evaluation

The scientific impact of CAROMAIZE was mainly due to its integrated concept covering metabolic engineering of high-pigmented maize prototypes, processing of the maize material by milling, carotenoid extraction and concentration and by either direct feeding of the milled maize or by feed supplementation with the extracted carotenoids. Key outputs included fundamental knowledge at the system level to enable more precise and targeted engineering of the carotenoid pathway and the use of maize as a production platform for valuable molecules and nutrient improved animal feed. CAROMAIZE demonstrated for the β-carotene maize its bioavailability as animal feed resulting in the desired pigmentation. We could show that its use is applicable for a future economic utilization in poultry farming. For astaxanthin from our maize prototype, we could demonstrate that it can fully replace synthetic astaxanthin with respect to fish pigmentation. However, a more economical direct application avoiding processing can be reached by further optimization of the astaxanthin prototype towards higher astaxanthin concentrations.

The data provided by the determination of carotenoids and related metabolites but also by an integrated analysis at a systems level using transcriptomics, proteomics and metabolomics provided fundamental knowledge for comparison to the non-transgenic maize for regulatory purposes and revealed how the primary metabolism was adjusted to enhanced precursor supply for carotenoid biosynthesis.

For standard pigmentation of chicken and eggs, the color required for commercial purposes are met by application of marigold flower extract in combination with red paprika extract. The current cost for both pigment additives can be avoided by using the β-carotene prototype which provides

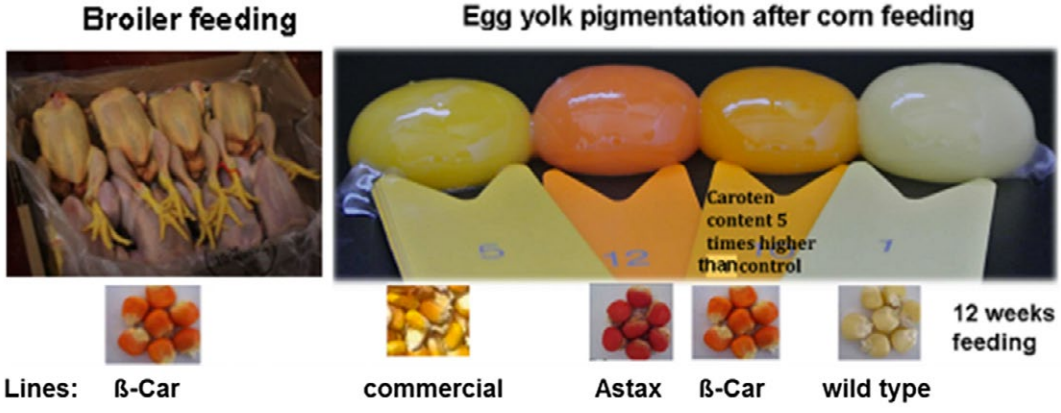


Fig. 3: The figure demonstrates the difference in pigmentation of egg yolk and the intensity of broiler pigmentation

CAROMAIZE Maize fortified with high-value carotenoids as raw material for industry and nutritionally enhanced animal feed

Fig. 4: Trout feeding experiment with an industrial feed as positive control, a negative control devoid of astaxanthin and a feed with astaxanthin oil.

high levels of yellow carotenoids by just replacing other cereals in the chicken feed. The resulting yellow pigmentation of meat and egg yolk makes the use of this maize prototype commercially appealing.

The usefulness of the astaxanthin maize as a substitute for chemically synthesized astaxanthin in fish feed was demonstrated. A final economic evaluation of astaxanthin from plant sources absolutely depends on the market price for synthetic astaxanthin which dropped in the course of this project from about 2000 to 1000 per kg. This price drop raised the economic hurdle for a biotechnologically produced astaxanthin. Our astaxanthin maize has not yet reached an economically competitive level. The astaxanthin concentration should be higher for direct feed supplementation to avoid the costs for extraction and concentration. A second generation astaxanthin maize with higher astaxanthin content can be obtained by increasing the astaxanthin concentration in additional engineering steps to enhance precursor supply and targeting a more efficient utilization of phytoene as well as a better conversion of intermediates.

Peer reviewed publications from the project

- *Huang J, Zhong Y, Sandmann G, Liu J, Chen F. (2012) Cloning and selection of carotenoid ketolase genes for the engineering of high-yield astaxanthin in plants. Planta 236, 691-699*
- *Farre G, Naqvi S, Sanahuja G, Bai C, Zorrilla-López U, Rivera SM, Canela R, Sandman G, Twyman RM, Capell T, Zhu C, Christou P (2012) Combinatorial Genetic Transformation of Cereals and the Creation of Metabolic Libraries for the Carotenoid Pathway. In Transgenic Plants Methods and Protocols. Series: Methods in Molecular Biology, Vol. 847: 419-435. Dunwell, Jim M.; Wetten, Andy C. (Eds.) 2nd ed. Springer, Berlin.*
- *Nogareda C, Moreno JA, Angulo E, Sandmann G, Portero M, Capell T, Zhu C, Christou P. Carotenoid-enriched transgenic corn delivers bioavailable carotenoids to poultry and protects them against coccidiosis.in press doi: 10.1111/pbi.12369*
- *Shu C, Berman J, Sheng Y, Wang Y, Capell T, Shi L, Ni X, Sandmann G, DellaPenna D,Christou P, Zhu C (2014) Cloning and functional characterization of maize (Zea mays L.) carotenoid epsilon hydroxylase gene. PLoS One 10, e0128758*
- *Farré G, Rivera Vélez S, Alve, R, Vilaprinyo E, Sorribas A, Canela R, Naqvi S, Sandmann G, Capell T, Zhu C, Christou P (2013) Targeted transcriptomic and metabolic profiling reveals temporal bottlenecks in the maize carotenoid pathway that can be addressed by multigene engineering. Plant J 75: 441-455*
- *Naqvi S, Zhu C, Farre G, Sandmann G, Capell T, Christou P (2011) Synergistic metabolism in hybrid corn indicates bottlenecks in the carotenoid pathway and leads to the accumulation of extraordinary levels of the nutritionally important carotenoid zeaxanthin. Plant Biotechnol J 9:384-393.*
- *Bai C, Rivera SM, Medina V, Alves R, Vilaprinyo E, Sorribas A, Canela R, Capell T, Sandmann G, Christou P, Zhu C. (2014) An in vitro system for the rapid functional characterization of genes involved in carotenoid biosynthesis and accumulation. Plant J 77:464-475.*
- *Berman J., Zorrilla-Lopez U., Gemma Farre G,. Zhu C., Sandmann G., Twyman R.M., Capell T., Christou P. Nutritionally important carotenoids as consumer products. Phytochem Rev 14:727-743*
- *Decourcelle M., Perez-Fons L., Baulande S., Steiger S., Couvelard L., Hem S., Zhu C., Capell T., Fraser PD., Christou P., Sandman G. Combined transcript, proteome and metabolite analysis of transgenic maize seeds engineered for enhanced carotenoid synthesis reveals pleotropic effects in core metabolism. J Exp Bot 66:3141-3150*
- *Zorrilla-Lopez U, Masip G, Arjo G, Bai C, Banakar R, Bassie L, Berman J, Farre G, Miralpeix B, Perez-Massot E, Sabalza M, Sanahuja G, Vamvaka E, Twyman RM, Christou P, Zhu C, Capell T (2013) Engineering metabolic pathways in plants by multigene transformation. Internat J Developm Biol 57: 565-576*

CONVIGOUR Controlling variation for germination and vigour in rapeseed

Rod Snowdon[1], Sarah Hatzig[1], Amine Abbadi[2] and Gunhild Leckband[2]

1 Department of Plant Breeding, Justus Liebig University, Giessen
2 NPZ Innovation GmbH, Holtsee

CONVIGOUR Partners: Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, KWS SAAT KG, Justus Liebig University Giessen, Leibniz Institute of Plant Genetics and Crop Plant Research (Germany), DL Seeds Inc. (Canada), RAGT Semences, GEVES (Groupe d'Etude et de contrôle des Variétés Et des Semences), INRA Joint Laboratory for Plant Genetics and Biotechnology, INRA Joint Laboratory for Molecular Physiology of Seeds (France)

In the public-private research consortium "CONVIGOUR – Controlling Variation in Germination and Seed Vigour in Oilseed Rape for optimal Yield Stability", German, French and Canadian researchers and breeders cooperated to investigate the complex genetic and environmental interactions influencing post-sowing field emergence in Germany's most important dicotyledonous crop. Funded on the German side by BMBF from 2011-2015, within the framework of the Knowledge-based Bio-Economy (KBBE) research programme, the CONVIGOUR consortium included four commercial plant breeding companies, along with five scientific institutes who contributed complementary expertise in plant breeding, genetics, genomics, biochemistry and seed physiology. The project generated comprehensive laboratory and field data for large plant populations. Multidimensional data analyses provided new insight into mechanisms underlying the variable germination performance of autumn-sown rapeseed in the face of environmental constraints.

Plant-based food, feed and biofuel production takes place under variable and changing climatic and soil conditions. Seed vigour is an important component of seed quality, with a serious impact on crop establishment and subsequent yield performance. Besides traditional seed quality criteria based on measurements of moisture content, purity, germination and phytosanitary requirements, satisfactory levels of seed vigour are essential to obtain an optimum plant stand and high productivity. The strong influence of environmental stress factors on seed vigour poses a strong challenge for seed companies to produce high vigour seeds in order to stabilise yield. A key to achieve this is the elucidation of the genetic and molecular basis of natural variation for seed vigour, in relation to external factors like temperature, water and soil. Internal factors like seed composition, concentrations of plant hormones and the micronutrient status play additional important roles in driving germination and seedling nutrition.

Of the four major crops grown in Germany, rapeseed (oilseed rape, canola: Brassica napus) has the smallest seeds. Consequently, the seeds have the lowest energy reserves to sustain seedling development after germination. This makes it essential for rapeseed seedlings to become self-sufficient as quickly as possible after germination, through fast and uniform emergence of photosynthetically active true leaves. On the other hand, sowing of European winter-type rapeseed generally takes place in August or September, when temperatures and water availability are notoriously unpredictable. The germinating seedlings therefore often have to cope with various forms of abiotic stress, ranging from water scarcity to flooding, a potentially extreme temperature range and sometimes considerable

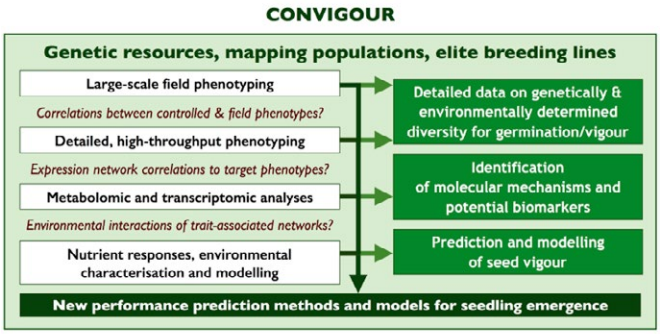


Figure 1. Seed vigour, a product of complex genetic interactions with multiple environmental influences. *Genetic control of seed germination and vigour is influenced by multiple interacting environmental factors throughout the entire growing season, which in winter-type rapeseed spans almost the whole year. Stress affecting the maternal plant, for example from drought or heat, can strongly influence the composition of the harvested seeds. Along with the temperature and humidity during seed storage, concentrations of specific hormones and metabolites in the seeds can severely influence germination capacity. Water, temperature, soil properties and microbes further modulate the emergence rate of the germinated seed. Depending on the winter climate, the speed of emergence and the number of emerged seedlings can in turn influence stand establishment, yield potential and seed properties in the following production year.*

soil encrustation. The substantial increase in arable area of rapeseed in recent decades has considerably increased the importance of such environmental constraints on seedling emergence, and breeders must spend considerable effort (including large-scale field trials to count emerging plants at multiple locations with varying environments) to accurately assess genetic variation for this important trait. Identification of gene targets or metabolic markers for selection of cultivars that are better adapted to adverse environments could potentially help to considerably improve breeding progress for seed vigour and yield.

Many valuable agronomic traits are under complex genetic control by large numbers of interacting gene loci (so-called "quantitative trait loci", or QTL), that can have variable effects depending on the specific genetic background and the environmental conditions. In some cases, trait expression is further complicated by multi-level environmental interactions. Seedling emergence and vigour traits are a prime example, because they can be strongly influenced not only by the genotype and the environment in which the seed germinates, but also by the environment in which the seeds have matured on the maternal plants (Figure 1). These complex environmental interactions are particularly pronounced in winter-type rapeseed, which has a growing season of nearly 11 months. The relevant time period influencing the seed performance, beginning with embryo fertilisation and continuing throughout the seed maturation, harvest, sowing, germination and field emergence phases, spans all four seasons from early spring through to early winter. (Fig. 1)

Emergence traits can therefore fluctuate considerably over rapeseed seedlots produced in different environments, hence their genetics and inheritance are poorly understood and breeding can be extremely difficult. To address this inherent complexity, the CONVIGOUR consortium applied multi-dimensional "omics" approaches to better understand the genetics of environmental influences on seedling emergence and vigour, particularly under abiotic stress constraints (Figure 2). Using seeds from genetically diverse populations, produced under independent production environments in different years, the CONVIGOUR partners performed deep phenotyping of germination and emergence traits in controlled-environment laboratory and growth-chamber experiments. These included high-throughput germination analyses in an automated phenotyping facility at the French national seed testing laboratory in Angers (project partner GEVES). This platform, which delivers detailed kinetic data about germination parameters, provides the possibility to screen large plant populations. By combining these data with high-resolution genetic marker information the CONVIGOUR partners were able to identify promising gene loci associated to enhanced germination performance (Hatzig *et al.* 2015). Genetic regions contributing significantly to emergence-related traits were mapped using bi-parental populations from crosses between good-emerging and poor-emerging parents, and validated genome-wide association studies (GWAS) in large diversity collections contributed by the German and French partners. (Fig. 2)

By comparing the data from the controlled-environment experiments with the emergence scores from large-scale, multi-environment phenotyping of the same plant populations under field conditions, it was possible to correlate

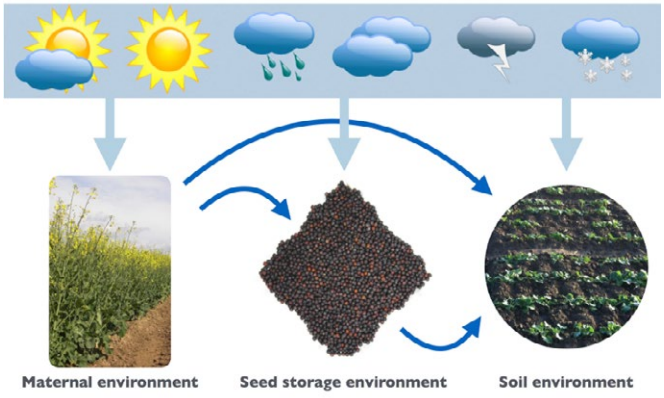


Figure 2. Overview of the CONVIGOUR project. *Multidimensional datasets related to germination and vigour under different stress constraints were collected in genetically diverse plant populations. The results led to identification of genetic and biochemical markers, performance prediction models and enhanced breeding approaches to select for high-vigour rapeseed cultivars.*

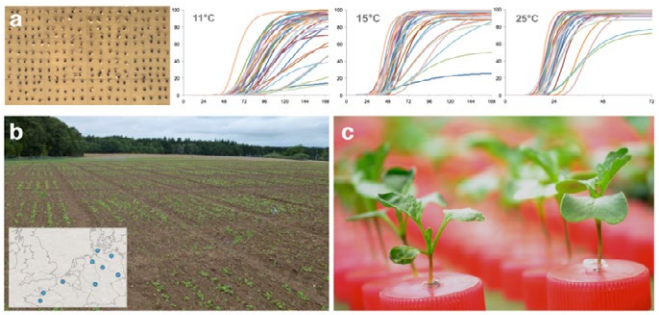


Figure 3. Translational plant breeding research from lab to field. *The CONVIGOUR project screened large plant populations in diverse field environments and controlled laboratory conditions. (a) Automated high-throughput germination screens revealed great diversity for germination responses to temperature, for example. Relating these data to emergence scores from (b) multi-location field trials enabled identification of surrogate traits that can be automatically screened in large breeding populations for detailed genetic analyses and enhanced selection. (c) Controlled growth-chamber experiments were carried out to perform sequencing-based gene expression and metabolite analysis of diverse genotypes under stress and control conditions. Kinetic models of germination were used to determine appropriate time points for investigation of gene expression networks involved in regulation of seed vigour.*

target traits showing high environmental variance to less complex, more heritable surrogate traits that can be measured in high throughput under controlled conditions. At the same time, it was possible to identify environmental factors with a decisive influence on the ability and speed of germination and emergence. This led to design of controlled-environment experiments to examine global gene expression patterns in roots and shoots of genetically diverse accessions under specific stress conditions (Figure 3). Systems-biological analyses, implementing sequencing-based transcriptome data and detailed metabolome data, allowed identification of gene expression networks regulated by genes associated with QTL for relevant vigour traits. Variants of such genes present breeders with new targets for molecular breeding. (Fig. 3)

Interestingly, key differences were also found between the regulatory responses of roots and shoots from plants with good versus poor post-germination stress tolerance, implicating a role of stress response pathways in modulation of emergence and vigour under abiotic stress constraints (Hatzig *et al.* 2014). Striking differences were also observed in the metabolite composition of good and poor germinating genotypes, respectively. Some key metabolites could be potentially useful as biochemical markers to tag seedlots with suspected sub-optimal germination capacity. By combining multiple levels of genetic, transcriptomic, metabolomic and phenotypic information, related to the germination-emergence complex and its regulation, we hope to ultimately refine genomic selection models in order to improve selection gains in breeding.

Bibliography

- *Hatzig S, Zaharia LI, Abrams S, Hohmann M, Legoahec L, Bouchereau A, Nesi N, Snowdon RJ (2014) Early osmotic adjustment responses in drought-resistant and drought-sensitive oilseed rape. Journal of Integrative Plant Biology 56:797-809*
- *Hatzig SV, Frisch M, Breuer F, Nesi N, Ducoumau S, Wagner MH, Leckband G, Abbadi A, Snowdon RJ (2015) Genome-wide association mapping unravels the genetic control of seed germination and vigor in Brassica napus. Frontiers in Plant Science 6*

FIBRAGEN Flax for omproved biomaterials through applied genomics

Katharina Haag & Jörg Müssig

Hochschule Bremen – City University of Applied Sciences , Bionik–Innovations–Centrum (B-I-C)
The Biological Materials Group , Neustadtswall 30, 28199 Bremen

I. Project summary

The long-term objective of FIBRAGEN is to expand markets for flax bio products by developing optimised feedstocks for use in advanced composite materials. Flax produces some of the longest and strongest fibres of any crop, and these fibres can be used to replace man-made materials such as glass fibres in composite materials. Establishment of a robust industry in biocomposites requires a reliable supply of high-quality flax feedstocks. The development of such feedstocks relies on the availability of molecular markers linked to traits of interest (e.g. yield, high tensile strength and matrix interface compatibility). FIBRAGEN aims to develop novel molecular markers that can be used in either linseed varieties (for dual purposed flax in Canada) or in fibre flax (for composite-dedicated crops in Europe).

A diversity panel of 66 different flax varieties and 3 RILs populations were grown and harvested in 2012 for different analyses. These plants have been phenotyped for different characteristics. Stem anatomy, fibre content and cell wall chemistry were being analysed in a subset (10 varieties) of these samples. Micromechanical tests and composite formation have been initiated on the 2012 samples. Samples have been collected and were carried out to optimise transcriptomics, proteomics and metabolomics of fibre-rich outer stem tissues from 10 selected varieties. These analyses were repeated with plants grown in 2013. Simultaneously, DNA has been extracted from the 66 diversity panel varieties and was re-sequenced in order to identify flax SNPs.

The major deliverables at the end of the project are a set of mapped SNPs and QTLs/genes, which can be used in marker-assisted selection of yield and quality, as well as a better understanding of the optimal properties of feedstocks for high performance natural fibre-reinforced composites. Together, these activities will also advance scientific understanding of cell wall development in cellulose-rich fibres.

1. The FIBRAGEN - consortium

- Within the framework of the funding measure *PLANT KBBE III* plant biotechnology for the future the transnational research project *FIBRAGEN – Flax for Improved Biomaterials through Applied Genomics* was carried out from May 2011 to March 2015.
- The consortium consisted of nine partners from France, Spain, Canada and Germany. The partners were funded each by the national funding agencies, in Germany by the Federal Ministry of Education and Research (BMBF).
- The partners within the consortium are:
  - Limagrain Europe, France (PI: Anne-Marie Bochard)
  - Terre de Lin, France (PI: Jean-Paul Trouvé)
  - Linea, France(PI: Reynald Tavernier)
  - Université de Lille, UMR SADV, France (PI: Simon Hawkins)
  - INRA UMR FARE, France (PI: Brigitte Chabbert)
  - AIMPLAS, Spain (PI: Valentin Polo)
  - Viterra, Canada (PI: Paul Dribnenki)
  - University of Alberta, Canada (PI: Mike Deyholos)
  - Hochschule Bremen HSB – City University of Applied Sciences, Bionik-Innovations-Centrum, The Biological Materials Group (PI: Jörg Müssig)

Due to funding problems, the Canadian partners could not contribute in the planned way. Partner Viterra was withdrawn from the consortium, the partner University of Alberta agreed to stay within the consortium with a lower amount of measurements to avoid the collapse of the project.

- The FIBRAGEN-consortium consists of experts from complementary research areas e.g. plant breeding, genome analysis, biology, plant cell wall chemistry, biological and composite materials. With this widespread knowledge, new insights on the relationship between the plant and the industrial application in fibre-reinforced composites were gained.

2. Planning and implementation of the project

- Within the interdisciplinary consortium a close collaboration of the partners was necessary. Therefore yearly meetings of the whole consortium accompanied by additional meetings of subgroups working on specific questions were realized. The annual meetings were organised by the partners (2011: Limagrain; 2012: AIMPLAS; 2013: Terre de Lin; 2014: HSB; 2015: Université de Lille).
- Due to the start of the project in May 2011, it was too late in the year to start with flax cultivation in 2011. This lead to delay from the originally planned schedule. To realise and analyse two subsequent years of flax cultivation, the project was elongated for 11 additional months till end of March 2015.

- As the Canadian partners could not participate in the project in the planned way, the project focused on the French breed varieties (mainly developed for fibre purpose), the oilseed-varieties mainly breed in Northern America were examined in a limited way.

3. State of the art of science and technology

- The **fineness of flax fibre bundles** is of crucial importance for the composite behaviour. At the same fibre mass fraction, finer fibre batches have a higher surface to enable load transfer from the polymer to the fibre (Thomason, 1999). The fineness of flax fibre bundles extracted from the plant stems depends on the morphology in the stem, the chemical composition and the retting and fibre extraction conditions (Akin, 2010). For the characterisation of the fibre bundle width there are several methods available; within the FIBRAGEN-project the FibreShape-system based on a high-resolution scanner and image analysis algorithms was adapted and used at HSB. The advantages of the system include relatively fast and reproducible analysis of a large number of elements and their distribution.
- **Tensile tests** are the most common method **to determine the mechanical properties of fibres and fibre bundles.** In the literature a large scatter of the properties of bast fibres is given that cannot only be explained by the natural variation of the fibre bundles, but is strongly influenced by the applied method (Nechwatal *et al.*, 2003). Within the project, a harmonisation of the methodology was achieved between the labs using different equipment. Therefore the protocols were adapted for the needs of bast fibre bundle testing. As important factors affecting the results the following reasons were identified:
  - The climatic conditions
  - The compliance of the testing machine
  - The sample selection
  - The method with which the cross-sectional area of the fibre bundles is determined.
- While the determination of tensile properties is quite common, only very little is known about the bending behaviour of bast fibres. Within a composite the fibres are not only subjected to tensile loading so the bending behaviour of a fibre is of crucial importance for a better understanding of composite properties. Within the FIBRAGEN-project an innovative measurement approach for the determination of the angle-dependent morphology and bending behaviour of single fibre bundles could be realised. Thus, for the first time a systematic analysis of the bending behaviour of flax fibre bundles was possible.
- Traditionally yarn structures are used for the **manufacturing of unidirectionally reinforced thermoset composites** for which the fibre bundles extracted from the plant have to undergo further textile processing steps (Angelov *et al.*, 2007). These steps induce damage to the fibre bundles and influence the composite behaviour (Bos *et al.*, 2002). To exclude these influences, a **lab-scale pultrusion** described by Mader *et al.* (2012) was adapted to work with 100 cm long bast fibre bundles instead of endless fibre structures. With the adapted set-up no further processing of the scutched fibre bundles was necessary. Therefore, the behaviour of the plant fibre in the composite could be analysed independently of further textile processing steps. The composite results could directly be correlated with the plant properties.

4. Collaborations

- **Within the FIBRAGEN-consortium** a close collaboration was necessary. To drive forward the composite and fibre characterisation work packages, especially the intensive cooperation with the partners INRA UMR FARE and AIMPLAS was driven forward.
- The realisation of the measurement set-up to characterise the bending behaviour of flax fibre bundles was possible with the support of **Diastron Ltd.** (New Hampshire, UK). A considerable exchange of information lead to several adaptations and optimisations in the hard- and software of the equipment.
- With the membership and active participation of Prof. Dr.-Ing. Jörg Müssig within the **Scientific Committee of the European Confederation of Flax and Hemp CELC**, current information on new developments and trends from the sector of bast fibre-reinforced composites could be implemented to the project. Furthermore, Katharina Haag, responsible for the FIBRAGEN-project within Prof. Müssig's group, could support the CELC on the **JEC Composites Show**, one of the world's largest fairs for composites in Paris, France (March 2012, 2014 and 2015). Getting in contact with representatives from the industry from all over the world gave constructive input on the demands of the industry that could be implemented into the project.
- To the initiative of **Projektträger Jülich (PTJ)** and in close collaboration with PTJ, representations of the project on the fairs „Grüne Woche" (Berlin, January 2012) and Hannover Messe (Hannover, April 2013 and April 2014) could be realised. The aims and results of the project could be shown to a broad range of visitors.

5. References

- *Akin, D.E. (2010): Flax – ASTM Standardisation and Harmonisation. In Müssig, J. (Editor, 2010): Industrial Applications of Natural Fibres. Structure, Properties and Technical Applications. John Wiley & Sons, Ltd Chichester, UK, p. 371 – 379.*
- *Angelov, I., Wiedmer, S., Evstatiev, M., Friedrich, K. & Mennig, G. (2007): Pultrusion of a flax / polypropylene yarn. Composites Part A 38 (2007), p. 1431-1438*
- *Bos, H.L., van den Oever, M.J.A. & Peters, O.C.J.J. (2002) Tensile and compressive properties of flax fibres for natural fibre reinforced composites. Journal of Materials Science 37 (2002), p. 1683-1692.*
- *Mader, A., Volkmann, E., Einsiedel, R. & Müssig, J. (2012): Impact and Flexural Properties of Unidirectional Man-Made Cellulose Reinforced Thermoset Composites. Journal of Biobased Materials and Bioenergy, 2012, 6, p. 481-492.*

- *Nechwatal, A., Mieck, K.-P. & Reußmann, T. (2003): Developments in the characterization of natural fibre properties in the use of natural fibres for composites. Composites Science and Technology 63(2003), p. 1273-1279.*
- *Thomason, J.L. (1999): The influence of fibre properties of the performance of glass-fibre-reinforced polyamide 6.6. Composites Science and Technology, 59 (16): p. 2315–2328.*

II. Detailed presentation

1. Utilization of the funding and detailed description of achieved results

Before starting with the project, the consortium defined 15 aims to be achieved within the project (see Table 1). The aims connected to the responsibilities of HSB are highlighted and further discussed. (Tab. 1)

2. Potential usability of the project outcomes

The leading breeders and suppliers of flax seeds in Europe were involved to the project as partners and are highly interested in improving fibre quality of new varieties by using the project results (e.g. molecular markers). With these future varieties the availability of flax fibres optimised for composite applications will be given.

Economic prospects

- Current developments during the project period proof that the acceptance of natural fibre-reinforced composites in the market is continuously growing. This is indicated by the growing diversity of products including high-class applications such as e.g. in the sports sector.
- An overview of current products is given by Müssig & Haag (2015). For these applications a constantly homogeneous quality of the raw material is necessary.
- With the results of the project, new flax varieties will be developed that fulfil these expectations.

Scientific and technical prospects

- The fundamental research on the relationships between the plant genome, the phenotype of the plant, the fibre properties and their behaviour in a composite will lead to optimised fibre qualities. This will strengthen the mar-

Table 1: Aims of the FIBRAGEN–project. The aims connected to the responsibilities of HSB are marked bold and further discussed.

No	Aim
1	>768 validated, mapped SNP markers with MAF >0.25 polymorphic in >10 varieties of fibre flax.
2	10 SNP markers linked to traits of fibre yield, fibre quality, and disease resistance
3	Identification of QTLs for fibre yield, disease resistance, and some aspects of quality.
4	Identification of best fibre flax and linseed varieties for composites from among existing germplasm.
5	**Definition of chemical and mechanical characteristics of an ideal flax fibre for composites** To reach this goal, flax fibres of > 10 varieties grown in two subsequent years were used for composite production with different methods and their performance was analysed. At HSB a lab-scale pultrusion line was adapted for the production of UD-reinforced composites with minimum fibre processing. As the Canadian partners could not participate in the project in the planned way, the focus was set on fibre varieties.
6	**Definition of the range of chemical, anatomical and mechanical variation in fibres of existing germplasm** The fibre bundles used for composite manufacturing were analysed regarding their mechanical properties and their fineness at HSB. The results were correlated with the performance of the composites.
7	**Identification of compositional (biochemical) predictors of fibre performance** Further flax fibre bundles grown in lines of the whole set of flax varieties were cleaned manually and analysed for their mechanical properties at HSB. Due to bad weather conditions in 2012 and the less homogeneous retting compared to the plot grown flax, larger scatter in fibre quality could be shown compared to the fibre production in quasi-industrial scale (plot grown flax).
8	Description of a physical model of heat & energy transfer during composite formation.
9	Identification of gene/protein expression patterns correlated with fibre composition and performance.
10	Identification of candidate genes for fibre development and fibre strength.
11	Development of a protocol for industrial scale application of flax in pultrusion composite manufacture Current developments within the natural fibre reinforced composites sector have shown that low twist flax yarns can be used in industrial scale pultrusion lines. As these processes could be established in the market, within the FIBRAGEN project, we focussed on lab-scale pultrusion for UD-composites without further textile processing into yarns. By doing so, another composite system (pultrusion with epoxy resin) could be analysed additional to the systems analysed by the project partners (injection moulding with PCL at INRA UMR FARE and VARTM with unsaturated polyester resin at AIMPLAS). The influence of fibre length and orientation could be analysed (compare Haag et al., 2014).

ket sector of natural fibre-reinforced composites which is dominated by small and medium-sized companies. Even though the prospects of the developments are high, the small and medium-sized companies would not be able to cover the development costs by their selves. They can benefit from the project results.
- The interdisciplinarity of the consortium depicts the unique chance to drive forward research and development from different sectors.
- The results were and are presented to interested people from different market sectors during conferences, publications in scientific journals and trade shows.

Scientific and economic integrability
- The international collaboration between the partners of the project and beyond could be broadened during the FIBRAGEN-project and will be continued in follow-up collaborations and projects. Due to the enhanced publicity (presentation on fairs) new R&D collaborations could be established. They will, based on the experience of the FIBRAGEN-project, drive forward questions along the processing chain from the fibre to the composite and will lead to new ideas and innovative products.
- As students were part of the FIBRAGEN-project they could learn about chances and specialties of flax-reinforced composites.
- With the set-up of a fibre bending equipment, new chances and possibilities to scientifically work on basic and up-to-now unsolved questions on structure-property relationships appeared.

3. Progress and contribution of others during the project
- During the FIBRAGEN-project the EU-funded project „MultiHemp-Multipurpose hemp for industrial bioproducts and biomass“ started, which follows related questions on the bast fibre plant hemp. Both consortia were in close contact; in January 2016 a common Workshop was

Table 2: Publications within the FIBRAGEN project in chronological order.

Haag, Katharina & Müssig, Jörg (2012): Comparison of tensile and flexural properties of different flax varieties and their composites. PLANT-KBBE Statusseminar, Kongresshotel Potsdam am Templiner See, March 6-8, 2012, Potsdam, DE.- Poster presentation

Müssig, Jörg & Hughes, Mark (2012): II – Reinforcements: fibres. In: Reux, Frédéric. & Verpoest, Ignaas (Ed.): Flax and hemp fibres: a natural solution for the composite industry. First edition – 2012, Paris, France, JEC composites, 2012, prepared for JEC by the European Scientific Committee of the CELC, (ISBN 978.2.9526276-1-0), p. 39-60.- Book chapter

Haag, Katharina & Müssig, Jörg (2013): FIBRAGEN – Flax for Improved Biomaterials through Applied Genomics. PLANT 2030 Statusseminar, Kongresshotel Potsdam am Templiner See, March 6-8, 2013, Potsdam, DE.- Poster presentation

Haag, Katharina & Müssig, Jörg (2013): Flax for Improved Biomaterials through Applied Genomics: New concepts to characterize the mechanical properties of flax fibre bundles. ResEff 2013 – International Conference on Resource Efficiency in Interorganizational Networks. Georg-August-Universität Göttingen, November 13-14, 2013, Göttingen, DE.- Presentation

Haag, Katharina, Beaugrand, Johnny, Fita, Sergio, Padovani, Justine & Müssig, Jörg (2014): FIBRAGEN – Flax fibre varieties for composite applications – a bottom-up approach for a high performance composites project. PLANT 2030 Statusseminar 2014, Kongresshotel Potsdam am Templiner See, March 31-April 02, 2014, Potsdam, DE.- Poster presentation

Müssig, Jörg & Haag, Katharina (2015): 2 - The use of flax fibres as reinforcements in composites. In: Faruk, Omar & Sain, Mohini (Ed.): Biofiber Reinforcements in Composite Materials, Woodhead Publishing, 2015, p. 35 – 85.- Book chapter

Haag, Katharina & Müssig, Jörg (2015): Morphology based phenotyping of flax fibre bundles - Comparison of manual and automated data analysis-. PLANT 2030 Status Seminar 2015, Kongresshotel Potsdam am Templiner See, March 4-6, 2012, Potsdam, DE.- Poster presentation

Hawkins, Simon, Bochard, Anne-Marie, Trouvé, Jean-Paul, Tavernier, Reynald, Chabbert, Brigitte, Fita. Sergio, Deyholos, Mike & Müssig, Jörg (2015): Flax for Improved Biomaterials through Applied Genomics. PLANT 2030 Status Seminar 2015, Kongresshotel Potsdam am Templiner See, , March 4-6, 2012, Potsdam, DE.- Presentation

Haag, Katharina (2015): Development of sustainable material concepts exemplarily shown for the process chain of natural fibre-reinforced composites. Ingenieurinnen Sommeruni, Universität Bremen, August 19, 2015, Bremen, DE.- Presentation

Haag, Katharina & Müssig, Jörg (2016): Scatter in tensile properties of flax fibre bundles: influence of determination and calculation of the cross-sectional area. Journal of Material Sciences, Vol. 51(17), p. 7907-7917 .- Peer-reviewed scien-

held on the topic „Fibre Quality“ in Lille, FR.
- The acceptance of flax-reinforced composites and their presence in the market was enlarged significantly since the beginning of the project. A large number of different products is available (compare overview by Müssig & Haag, 2015). In conferences like the 13th International JEC Conference (June 2-4, 2015, Houston, USA) two of six sessions were dedicated to natural fibre topics. This shows the topicality of the project and indicates that the knowledge gained within the project will be transferred into new, innovative and bio-based products.

4. Publications

- A selection of the results from the project was chosen for the presentation at the annual PLANT 2030 status seminars held in March in Potsdam, DE (2012-2015).
  In 2015, additionally an oral presentation was given presenting results from the overall consortium.
- At the ResEff-conference in Göttingen, DE, in November 2013 the background of the FIBRAGEN-project as well as results from work package 1 and 2 were presented.
- The knowledge gained within the FIBRAGEN-project could be integrated into a book chapter on flax as reinforcement fibre in composite materials.
- Further on, several publications in scientific journals were realised and are planned in detail, respectively.

An overview on the publication activities is given in Table 2.

5. References

- *Haag, Katharina, Beaugrand, Johnny, Fita, Sergio, Padovani, Justine & Müssig, Jörg (2014): FIBRAGEN – Flax fibre varieties for composite applications – a bottom-up approach for a high performance composites project. PLANT 2030 Statusseminar 2014, Kongresshotel Potsdam am Templiner See, March 31-April 02, 2014, Potsdam, DE.- Poster presentation*
- *Müssig, Jörg & Haag, Katharina (2015): 2 - The use of flax fibres as reinforcements in composites. In: Faruk, Omar & Sain, Mohini (Ed.): Biofiber Reinforcements in Composite Materials, Woodhead Publishing, 2015, p. 35 – 85.- Book chapter*

SAFQIM Sugars and fruit quality in melon

France
Vilmorin & Cie S.A., Beaucouze: Danièle Hosemans (coordinator), Frank de Langen
URGV-INRA, Evry: Abdel Bendahmane, Adnane Boualem
INRA-GAFL, Montfavet: Catherine Dogimont
INRA-Bordeaux Aquitaine, Villenave d'Ornon: Yves Gibon

Spain
CSIC-IRTA-UB, Centre de Recerca en Agrigenòmica (CRAG), Barcelona: Jordi Garcia-Mas
Universidad Politécnica de Valencia-COMAV, Valencia: Maria Belén Picó
CSIC- IBMCP-UPV, Valencia: Antonio Monforte
Semillas Fitó S.A.U, Barcelona: Torben Jahrmann

Germany
MPI für Molekulare Pflanzenphysiologie, Potsdam-Golm: John Lunn, Björn Usadel* (*present address: RWTH, Aachen)

Summary
Fruit sweetness, mainly determined by sucrose content, is one of the most important fruit quality traits in melon. Melon fruits are sink organs that depend on import of sugars from the leaves via the phloem to accumulate sucrose. In sweet melon types, sucrose levels are very low in young developing fruit, but reach very high levels in the mature fruit. Sucrose and its constituent hexose sugars also act as signal molecules that regulate the expression of genes involved in several important physiological processes. Continued metabolism of sugars is a major factor in post-harvest deterioration of fruit quality in climacteric melon types, as it leads directly to loss of sweetness and accumulation of undesirable products that decrease fruit quality, limiting the shelf life and potential geographical market for the fruit. Within the Plant-KBBE SAFQIM project we studied sucrose metabolism in melon fruit using a combination of genetic, transcriptomic and metabolomic approaches. The project used the comprehensive set of genetic and genomic tools available in melon (mapping populations, TILLING and Eco-TILLING platforms, EST collections, microarray, genome sequence) to identify a candidate genes with potential to give rise to new, potentially valuable phenotypes when mutated. There is considerable natural variation in melon and introgression of wild germplasm into cultivated melon is another approach being used to develop new varieties with higher and more stable fruit sugar content, as well as other quality and agronomic traits. The final goal of the project was to characterize newly generated melon lines containing novel alleles of genes that will improve the sugar profile and post-harvest stability of sugar levels in the fruit, which ultimately will be introduced into elite melon lines by the participating seed companies.

PROJECT AIMS
Fruit sweetness, mainly determined by sucrose, is one of the most important fruit quality traits for melon producers and consumers. The overall objective of the project is to understand the factors that regulate sugar accumulation in developing melon fruits and the stability of sugar content in mature fruits. The ultimate goal is to generate melon lines

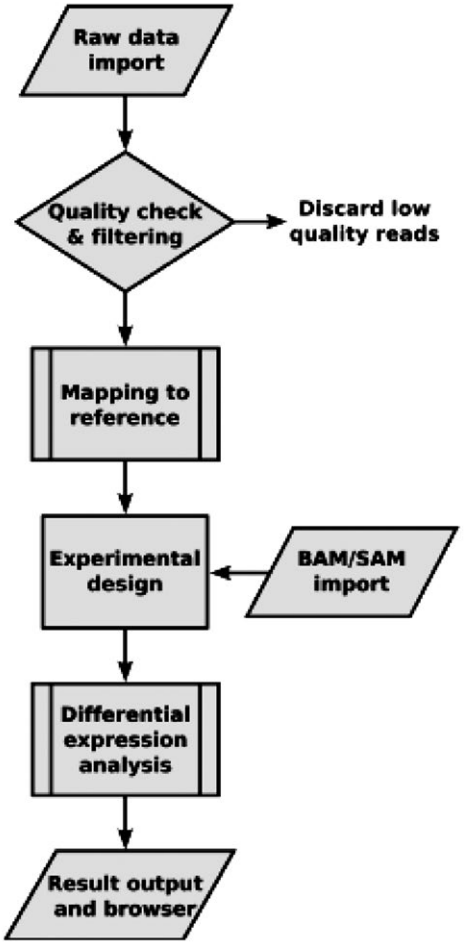


Fig.: 1: Flow chart of RNA-Seq-based differential gene-expression analysis steps provided by Robi NA.

containing new alleles of genes that will improve the sugar profile and post-harvest stability of sugar levels in the fruit.

Outcomes

One of the most significant outcomes of the project was development of the RobiNA software (Lohse *et al.*, 2012) for analysis of transcriptomic data acquired using next-generation sequencing technmologies (RNA-seq). RobiNA performs quality control checks, processes the raw reads, maps the reads to a reference transcriptome or genome, compares data between treatments, genotypes etc., performs statistical analyses of the data and gives data outputs that can be readily visualized using the MAPMAN suite of tools (Fig. 1). The RobiNA software is intuitive and easy to use by non-experts, and has been applied to a wide range of crop species and also to animal and human studies, being cited 270 times since its publication in 2012. The software is available free of charge from: http://mapman.gabipd.org/web/guest/robin.

Publications

- *Lohse M, Bolger A, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B (2012) RobiNA – a user-friendly, integrated software solution for RNA-Seq based transcriptomics. Nucleic Acids Research 40; 622-627.*

A report on the KBBE-SAFQIM project appeared in the BMBF Pflanzenforschung online journal:

- *Moosmann, A. (2013) Süßer Sommerstar - SAFQIM erforscht die Reifung der Melone. http://www.pflanzenforschung.de/de/journal/journalbeitrage/suesser-sommerstar-safqim-erforscht-die-reifung-der-mel-10068*

Publications from ERA-Net PG MELRIP project (partly supported by KBBE-SAFQIM):

- *Leida C, Moser C, Esteras C, Sulpice R, Lunn JE, de Langen F, Monforte AJ, Picó B. (2015) Variability of candidate genes, genetic structure and association with sugar accumulation and climacteric behavior in a broad germplasm collection of melon (Cucumis melo L.). BMC Genetics 16; e28.*
- *Saladié M, Cañizares J, Phillips MA, Rodriguez-Concepcion M, Larrigaudière C, Gibon Y, Stitt M, Lunn JE, Garcia-Mas J. (2015) Comparative transcriptional profiling analysis of developing melon (Cucumis melo L.) fruit from climacteric and non-climacteric varieties. BMC Genomics 16; e440.*

TREE FOR JOULES Improving eucalypt and poplar wood properties for bioenergy

Fladung, M.[1], Pakull B.[1], Meier D.[2], Schmitt U.[2]

1 Thünen Institute for Forest Genetics, Sieker Landstr. 2, D-22927 Grosshansdorf
2 Thünen Institute for Wood Research, Leuschnerstr. 91, D-21031 Hamburg

Introduction

The efforts to reduce the dependence on non-renewable fossil fuels and to mitigate climate change are leading to an increasing interest in sustainable bio-energy production. The use of lignocellulosic biomass from forest plantations as second-generation renewable bio-energy feedstock is gaining more and more attention, as trees are non-food crops, which can be grown under relatively poor soil conditions (Häggman *et al.* 2016). Fast-growing tree species such as poplar and eucalyptus grown as short-rotation coppice are easy to establish and produce high yields of lignocellulosic biomass.

Regrettably, the structure and composition of the lignified secondary cell wall, which mainly consists of cellulose (about 45%), hemicelluloses and lignin (each about 25%), render woody feedstock particularly recalcitrant to degradation for e.g. bioethanol production (cellulose is the most valuable component for biofuel-production, while lignins impair the accessibility and are considered as a barrier for the fermentation process).

For this reason, improved genetic material is needed to use trees as energy crops in an efficient manner. The project "TreeForJoules" therefore aims to uncover processes/genes regulating relevant wood cell wall properties in poplar and eucalyptus. This knowledge will be invaluable for breeding fast-growing "elite" trees for improved down-stream processing and efficient conversion into biofuels. "TreeForJoules" general aims are to (Grima-Pettenati 2013, 2014):

- identify and characterize the regulatory candidate genes (i.e. transcription factors and miRNAs) that control wood properties relevant to bioenergy;
- delineate and characterize genomic regions in eucalyptus and poplar that control wood properties of interest through comparative genetic and physical mapping, and comparative QTL mapping;
- develop highly efficient methods for wood property measurements, which can be used for the analysis of tree lines produced in the project;
- test relevant samples for their bio-ethanol and bio-oil potential.

The specific roles of the two Thünen Institutes Forest Genetics and Wood Research are to (Pakull *et al.* 2015):

- perform a transcriptome analysis of developing xylem to identify candidate genes controlling wood properties in Populus tremula;
- overexpress / downregulate these candidate genes (CG) by genetic transformation of poplar;
- analyse bioenergy potentials, saccharification, and bio-oil of CG-transgenic poplar;
- micro-phenotype transgenic wood zones through microscopic and micro-spectrophotometric investigation methods.

Results

In silico identification of CGs from compendia of expression data

The Thünen Institute of Forest Genetics performed a transcriptome sequencing analysis of developing wood in *Populus tremula* x *Populus tremuloides* by Illumina High-Seq. 2000 (GenXPro, Frankfurt; 100bp reads of a non-normalized mRNA-libary; 50 bp reads of a miRNA library). Bioinformatic analysis was done and transcriptome databases (transcription factors, miRNA) have been built.

In total, 19,139 different *P. trichocarpa* transcript identifiers could be assigned, including 707 transcription factors/transcription regulators identified to date. The miRNA was prepared from total RNA by size-fractionation using gel filtration. About 191,000 different miRNA tags were generated by Illumina sequencing with HiSeq 2000 sequencer.

Functional characterization of a limited number of selected CGs

Identification of candidate genes

The 707 transcription factors identified so far by xylem transcriptome sequencing (see before) were analyzed for their potential as candidate genes for wood properties important for bioenergy. Criteria for the selection of candidate genes were discussed in telephone conferences and personal meetings with the different partners of the project.

Candidate genes were chosen based on relative transcript levels in the transcriptome sequencing approach, *P. trichocarpa* xylem eFP expression data (http://www.bar.utoronto.ca/efppop/cgi-bin/efpWeb.cgi), and on different expression levels between tension and opposite wood which show differences in their relative composition concerning lignin and cellulose.

The NAC, LIM, KNOX and MYB transcription factors, which have been shown to regulate the monolignolpathway were considered in particular.

Characterization of candidate genes

Up to now, the Thünen Institute of Forest Genetics has chosen 13 candidate genes probably regulating wood properties for testing via stable transformation (knockdown and over-expression lines). Constructs comprising nine overexpression constructs, three RNAi-constructs and one amiRNA-construct have been produced and transformed into *Populus*.

Altogether, 23 transformation approaches have been carried out. Regenerated plants are still under selection for all constructs. Independent genetic lines with propagated individuals exist for eleven of the 13 constructs (one to 20 lines per construct up to now, more plants are still under selection). Transgenic lines of 10 (11) constructs have successfully been PCR-tested for transgene presence with gene- and T-DNA specific primers. The construct copy- number has been analyzed by Southern Blot experiments for transgenic lines representing five different constructs. RT-PCR experiments for construct expression analysis have been started.
Transgenic lines representing 10 (11) constructs have been transferred to the greenhouse. Four of these transgenic lines representing two different constructs show apparent phenotypic aberrations, including dwarfing or reduced growth and/or bended stems and stronger and irregular leaf serration or brown spotted and crumpled leaves (Fig. 1). Transgenic lines representing five different constructs were old enough for relevant wood production, and have been harvested for wood property analysis.

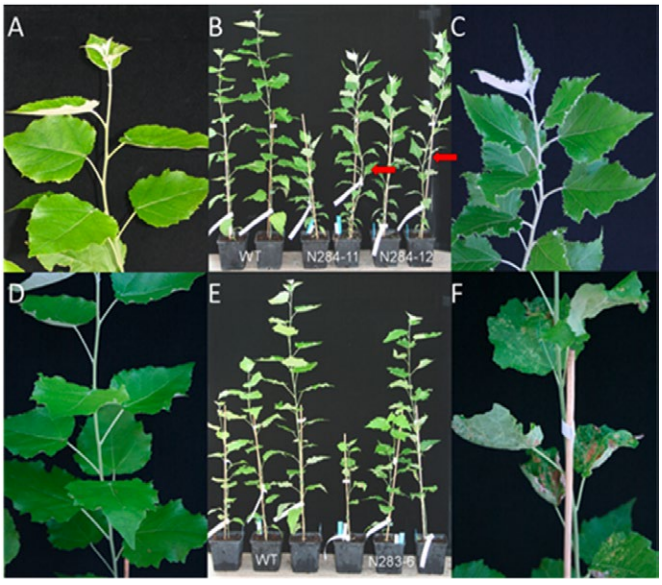


Fig. 1: Phenotypic changes within the transgenic poplar lines produced during the project. A and D: wildtype. B and C: phenotypic changes in plants overexpressing the transcription factor Potri 018G068700.1, a NAM (no apical meristem)-like protein, with high levels of xylem expression and a higher expression rate in tension versus opposite wood. E and F: phenotypic changes in plants overexpressing the transcription factor Potri. 003G195300.1, an alcohol dehydrogenase /Myb/SANT-like Transcription factor, with high levels of xylem expression and a detectable difference in expression rates in tension versus opposite wood.

Five constructs comprising candidate genes for wood properties from *Eucalyptus grandis*, produced by the working group of Jacqueline Grima-Pettenati (University of Toulouse, France), were also transformed into poplar by the Thünen Institute of Forest Genetics. The plants are still under evaluation.

Bioenergy potentials, saccharification, and bio-oil

Lignin content and sugar composition

The samples were subjected to a two-step acid hydrolysis for determination of the Klason lignin content. Klason lignin contents of the poplar hybrids varied between 19.4 and 22.5 wt%, the eucalyptus species varied between 23.7 and 33.1 wt% and the poplars varied between 15.8 and 25.5 wt%. Monomeric sugars resembling the carbohydrate composition of the cell wall were also determined by acid hydrolysis with subsequent borate anion exchange chromatography.

Micro steam-treatment

A micro steam-treatment system was constructed and a suitable method was developed and optimized (210 °C for 15 min) to give comparable results with a procedure at laboratory scale. The samples were further enzymatically hydrolyzed and yielded approximately 65wt% of fermentable sugars.

Analytical Pyrolysis-Gas Chromatography/ Mass Spectroscopy

For the estimation of bio-oil potential all samples have been subjected to analytical pyrolysis combined with GC/MS. From each sample 4 repetitive runs were performed. A typical chromatogram showing all volatile products is presented in Fig. 2.

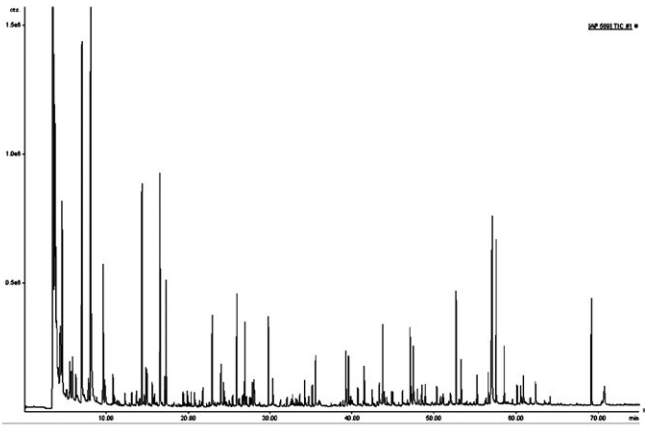


Fig. 2: Typical chromatogram of volatile components from poplar wood

The resulting data from pyrolysis have been chemometrically evaluated by multivariate data treatment. The calculations include a multistep procedure consisting of (1) peak alignment, (2) transformation (3) principal component analysis and (4) partial least square regression analysis. Preliminary results show that samples with different gene constructs could be discriminated based on some lignin and carbohydrate peaks (see Fig. 3). Further evaluation and data interpretation is still ongoing.

Preparative Mini-Pyrolysis

In order to obtain mass balances and investigate products from bio-oil production, a mini pyrolyzer was used. The whole system could be weight before and after the experiment. Approximately 100 mg were pyrolyzed, so that oil yields and char yields could be determined gravimetrically. In addition, the biooil composition could be analyzed by gas chromatography.

In comparison to standard beech wood, the oil yields are slightly lower (2-5 % absolute) and could be caused by the lower lignin content. Figure 6c demonstrates great differences in biooil yield (30-45 wt%) from INRA samples. Samples from the wild type yields typically more biooil which can be related to the lignin content.

In addition, 127 transgenic samples were received from three different partner organizations comprising 50 poplar hybrids (P. deltoids x P. trichocarpa), 24 eucalyptus hybrids (E. urophylla x E. grandis), and 38 poplars (P. tremula, P. tremuloides, P. x canescens, P. alba).

Micro-phenotyping of transgenic wood zones through microscopic and microspectrophotometric investigation methods

Structural and topochemical analyses

Beside symptomless xylem samples, slightly brownish discolored xylem portions were prepared for microphenotyping. Briefly, for transmission electron microscopy (TEM) the samples taken with razor blades from stem portions were prepared as follows: after fixation in an aldehyde solution, samples were washed, dehydrated in acetone, and finally embedded in Spurr's epoxy resin. Ultrathin sections (0.1 µm thick) were cut with a diamond knife and analysed with a Philips CM12 transmission electron microscope at accelerating voltages of 60 or 80 kV.

A parallel set of samples with same origin was prepared for cellular UV-microspectrophotometry (UMSP) in the same way to analyze lignin topochemistry. However, 1 µm semithin sections were prepared. Lignin topochemistry was analyzed in the scanning mode at a constant wavelength of 278 nm (absorbance maximum of hardwood lignin).

It was shown by cellular UV-microspectralphotometry that walls of the various cell types in hardwoods had a heterogeneous lignin distribution. At a wavelength of 278nm, i.e., maximum absorbance for hardwood lignin, most lignin was found in the cell corner regions with values between 0.25 and somewhat above 0.3. On the other hand relatively low

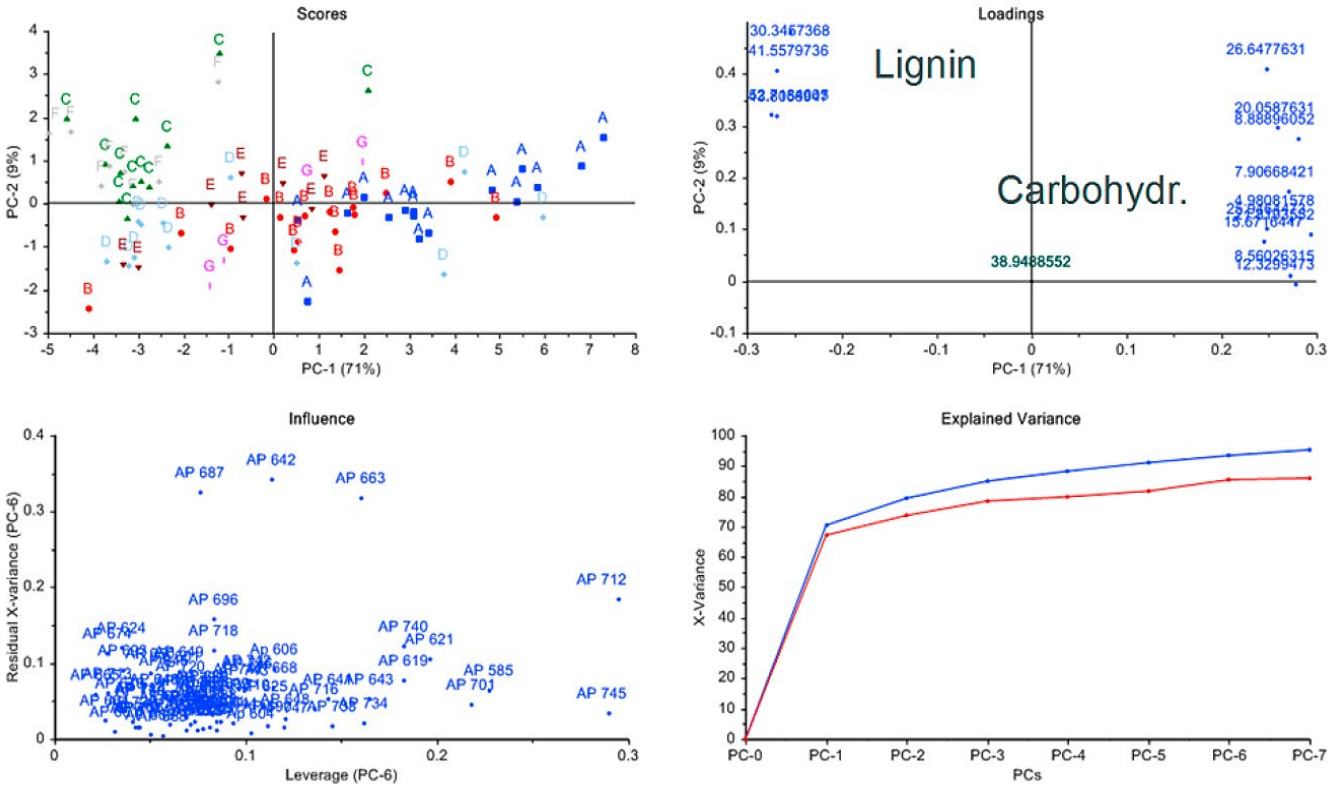


Fig. 3: Principal component analysis (PCA) of based on selected lignin and carbohydrate-derived components

absorbance was recorded for the secondary wall layers with values mostly below 0.1. When comparing the secondary walls of vessels with secondary walls of fibres, more lignin was found in vessel walls (Figs. 4a and d). In some cases, no lignin was found in inner secondary wall portions; most probably, these wall portions could be anatomically identified as so called gelatinous layers (Fig. 4b). Gelatinous layers or G-layers are composed of cellulose only, without any lignin deposited. Those fibres are formed in wood under tension stress, therefore belonging to reaction wood tissue. Varying lignin contents in cell corner regions have been regularly observed in many hardwoods. According to their fine structure, those differences became evident after staining with potassium permanganate showing either intensely stained regions or empty appearing regions (Figs. 4c and e). The different transgenic lines also often had different lignin levels. E.g. some xylem regions showed a relatively high lignin content in their fibre secondary walls as well as in their cell corner regions.
Some genetically modified poplar trees with homogeneous lignin distribution in fibres of adjacent growth rings and without tension wood showed similar absorbance values like control trees. Scans of xylem portions without tension wood fibres revealed a homogeneous lignin distribution in cell walls throughout the outermost two growth rings.

Electron microscopy added information on the cell wall architecture of topochemically varying xylem fibres. Wall thicknesses corresponded to those visualized by UV-microscopy. We observed rather thin walls with a narrow S2 layer as well as thick walls with broadened S2 layer. The formation of G-layers which were deposited onto the innermost secondary wall layer was also well demonstrated and found in several transgenic lines. Also poplar trees without genetic modifications partly showed tension wood zones. Potassium permanganate staining gave no evidence that G-layers of these samples contained any phenolic constituents which might also affect UV-absorbance behavior.

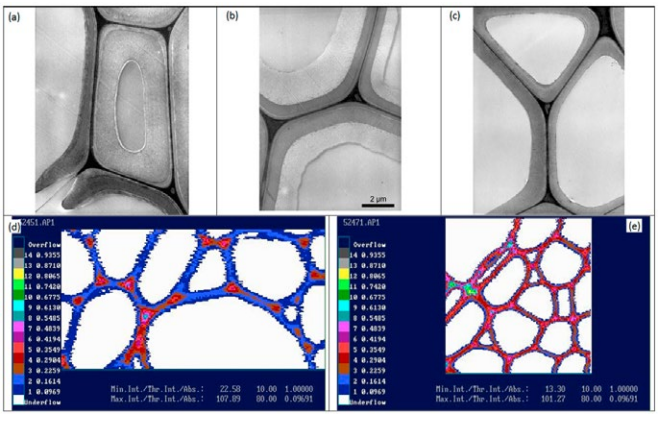


Fig. 4: Electron and UV micrographs of different poplar lines (a) Electron micrograph of poplar fibres with thick walled fibres (right) and thin walled, densely stained vessel wall (left); vessel walls contain more lignin than fibre walls; (b) Electron micrograph of poplar fibres with secondary cell walls consisting of outer S1 and S2 layer and an inner gelatinous layer; (c) Electron micrograph of poplar fibres showing cell corner with varying ultrastructure; empty appearing cell corner regions probably contain less lignin; (d) UV micrograph of poplar xylem with fibres (top) and a vessel (bottom) taken at a wavelength of 278 nm. Secondary walls lower absorbance than cell corner regions; (e) UV micrograph of poplar xylem with fibres (top) and a vessel (bottom) taken at a wavelength of 278 nm. Fiibres with a relatively high absorbance of their secondary walls.

Chemical analyses of Klason lignin contents vary more or less distinct between transgenic lines. As considered above, wall thicknesses as well as amount of tension wood fibres with reduced lignin amounts in their walls can also play a certain role for the determined variations in lignin contents.

In conclusion, when comparing structural, topochemical and chemical analyses it could not clearly be stated that reduced or increased lignin contents are alone the results of modified fibre walls with more or less lignin. The formation of tension wood with their lignin free G-layers as well as the formation of narrow or broad secondary wall layers in fibres cannot be attributed to a certain transgenic poplar line. This is due to a rather inhomogeneous xylem tissue in the investigated trees. It is recommended to exactly determine the amount of tension wood – which we did not perform in the present study – to exclude this as a major structural feature affecting chemical composition of trees.

For microphenotyping, also samples from project partners INRA Orleans, CIRAD Montpellier and the Thünen Institute of Forest Genetics were taken from various transgenic poplar trees (*Populus tremula x tremuloides*) and one series of eucalypt trees (*Eucalyptus urograndis*).

References

- *Grima-Pettenati J, Leplé J C, Gion JM, Harvengt L, Fladung M, Kamm B, Pinto Paiva J, Rodrigues JC, Costa Leal L, Cantón FR, Gallardo F, Allona I, Sixto H, Ruiz F (2013) TreeForJoules – supporting sustainable second generation biofuels. In: International innovation : dissemination science, research and technology. Auzeville: Universite Paul Sabatier, p 2*
- *Grima-Pettenati J, Leplé J C, Gion J M, Harvengt L, Fladung M, Kamm B, Pinto Paiva J, Rodrigues JC, Costa Leal L, Cantón FR, Gallardo F, Allona I, Sixto H, Ruiz F (2014) TREEFORJOULES, a Plant KBBE project to improve eucalypt and poplar wood properties for bioenergy. In: Lusser M (Ed) JRC Scientific and Policy Reports - Joint Research Centre. Proceedings of the Workshop on public-private partnerships in plant breeding, pp. 60-61.*
- *Häggman H, Sutela S, Fladung M (2016) Genetic Engineering Contribution to Forest Tree Breeding Efforts. In: Vettori C et al (Eds), Biosafety of Forest Transgenic Trees. Springer Science+Business Media Dordrecht, Forestry Sciences 82, pp. 11-29. DOI 10.1007/978-94-017-7531-1_2.*
- *Pakull B, Kersten B, Lüneburg J, Fladung M (2015) TreeForJoules Improving eucalyptus and poplar wood properties for bioenergy - genetics. In: Conference documents Plant 2040 Status Seminar, March 4 - 6, 2015 in Potsdam. pp 171-172*

PFLANZENBIOTECHNOLOGIE DER ZUKUNFT

AMAIZING Targeted molecular genetic and bioinformatic approaches to increase genetic diversity in elite MAIZe breedING populations

T. Presterl[1], M. Ouzunova[1], D. Scheuermann[1], P. Westhoff[2], K. Ernst[2], M. Frisch[3], C. Falke[3], K. Mayer[4], G. Haberer[4]

1 KWS SAAT SE (KWS), Grimsehlstraße 31, 37555 Einbeck
2 Heinrich-Heine Universität (HHU), Universitätsstr. 1, 40225 Düsseldorf
3 Justus-Liebig Universität Gießen (JLU), Heinrich-Buff-Ring 26, 35398 Gießen
4 Helmholtz Zentrum München (HGMU), Ingolstädter Landstr. 1, 85764 Neuherberg

Abstract
The aim of the AMAIZING project was the targeted transfer of favorable alleles for known QTL of the traits nitrogen-use-efficiency, cold tolerance, Northern Corn Leaf Blight (NCLB) resistance from tropic germplasm into adapted elite maize material. This should be achieved without transferring negative effects of exotic germplasm on adaptation and photoperiodic sensitivity. The transfer of the exotic alleles was controlled by cost-efficient SNP markers from the known QTL regions. Whole genome sequencing of recurrent parent lines supported the development of high density marker maps for the target region. Additionally, the allelic diversity of the produced introgression lines (IL) was investigated with new molecular technologies like targeted sequence capture. Extensive field trials with ILs revealed a series of new promising alleles, particular for resistance to NCLB.

Simultaneously an optimized strategy via biostastistical methods was developed, which allows investigating genome-wide introgression populations (IP) much cost- and time-efficiently. The BC3-DH scheme with pre-selection of complete donor chromosomes proved to be efficient in practice and was used for the development of 3 IPs. Several approaches for a powerful QTL detection in IPs were investigated and applied in the project. Field trials with IPs revealed positive effects of certain genomic regions, particular for disease resistance to NCLB.

IL with such favorable new exotic alleles will immediately be used in breeding and can be regarded as an excellent source for further fine mapping or gene cloning projects. Further work will be focused on the molecular and functional characterization of the new alleles using molecular resources generated in the project.

Results

1. Targeted allele mining in tropical germplasm and introgression into elite germplasm

The gene pool of maize is tremendously diverse (Goodman, 1985). Despite this abundance of genetic diversity, the germplasm base of current temperate elite cultivars is narrow (Troyer, 1999; Reif et al, 2005). In the tropics, maize has been adapted to very contrasting climates. In the lowland tropics, when conditions are humid, maize is cultivated under a high insect and fungal disease pressure. However, in the same regions the maize crop may also face severe droughts and high temperatures. In the highland tropics, i.e. the Andean regions, maize is grown at elevations up to 4000 m and has to cope with very low temperatures throughout its growing season. Tropical maize should therefore present a source of allelic variation that confers tolerance to diverse biotic and abiotic stresses, and that might be used to improve yield and its stability in temperate germplasm.

1.1. Selection of donors and development of target specific introgression lines

A set of 90 exotic maize accessions was available for the project. This set consists of inbred lines and populations from the tropics: 1) lines from the US GEM project (Germplasm Enhancement of Maize; Willmot & Pollak 2003); 2) CIMMYT lines; 3) additional accessions from tropical origin; 4) six landrace populations from the Andean Highlands; and 5) four flint maize landraces.
Available information on pedigree, disease resistance, stress tolerance (drought, low N) was considered for the choice of lines for the project.

In a first step, exotic accessions were crossed with five KWS elite lines (GA1-5) as a preparatory work to the project by partner KWS. In total, 289 exotic alleles were segregating in

the BC1 generation. Using the MaizeSNP50 bead Chip (Illumina) and additional information on the phenotype, the pedigree, and the origin (e.g. only tropical highland and flint maize populations for chosen for the cold tolerance QTL), the number of populations was reduced in BC2 to select approximately 10 different alleles per target QTL for an intensive backcrossing procedure. Marker-assisted backcrossing was used with three generations per year to produce homozygous Near Isogenic Lines (NILs) by at least two generations of selfing after the BC3. A set of QTL specific SNP markers was developed for all target QTL and the selection against the donor background genome was performed with an KWS internal 12k SNP array.

1.2 Genotyping of genetic resources

Molecular marker information was not available for most of the exotic materials. The MaizeSNP50 bead Chip from Illumina was used for this first genotyping. The aim was to reduce the number of donors to a set with maximized genetic diversity. Important KWS elite temperate lines were included for comparison. Donor germplasm showed high genetic diversity and the pattern of variability did not fit into known heterotic groups (e.g. flint and dent). First principle component analysis (PCoA) axes separates more pure tropical CIMMYT lines from GEM lines, which contain already 50% of temperate US germplasm (Figure 1). Hybrids and landraces showed high degree of heterozygosity and tend to cluster more in the center of the PCoA plot.

Due to the diverse degrees of heterozygosity within exotic donor accessions, individual plants were used for the backcrossing procedure to work with defined alleles. F1 plants were analysed with markers for the target regions and information was used for further selection of diverse exotic alleles. Depending on the target, the allelic diversity was very high.

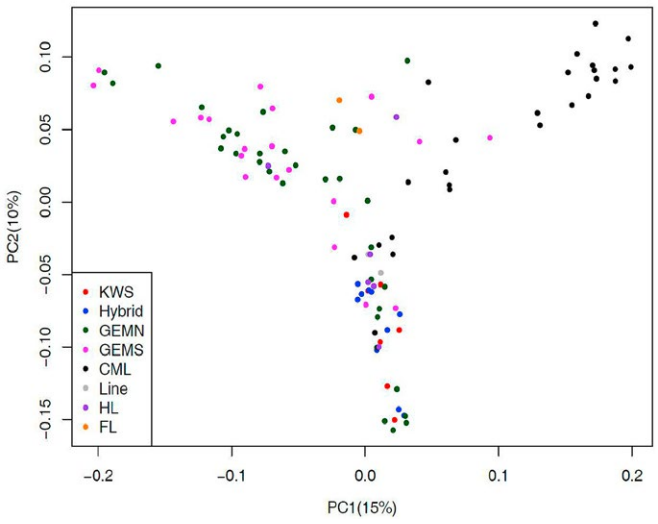


Fig.1: PCoA analysis of all exotic donors and KWS elite lines based on 50k chip marker information (Hybrid=exotic hybrids, GEMN, GEMS=GEM lines, CML=CIMMYT lines, Line=exotic line from gene bank, HL=Andean Highland landraces, FL=Flint landraces).

1.3 Molecular analysis of target regions

Five QTL for chilling tolerance (CT), nitrogen use efficiency (NUE) and Northern Corn leaf Blight (NCBL) on chromosomes 4, 5, 8, and 10 were investigated in tropical maize accessions and European elite lines. As the QTL differed considerably in size, they were analysed by different approaches. $QTL4_{CT}$ and $QTL8_{NUE}$ had already been narrowed down to a size feasible for analyses by the "Sequence capture method", while the other QTL seemed to be too large for this method. Hence, these QTL should be analysed by Infinium® technology to achieve high density genetic maps.

42 exotic donors and six recurrent parents from which the IL were developed, were used for the sequence capture hybridisation of the target QTL_{NCLB}. For the sequence capture array (Roche NimbleGen), 10 Mb on the physical map of B73 AGPv02 of the target region were selected, and the information was transferred to Eurofins Genomics (Ebersberg, Germany). In the region, 30.9% of sequences could be identified that were i) unique in the B73 genome, ii) equally distributed over the target region, and iii) useful for probe design.

The obtained sequence information ranged between 2533 Mb und 7083 Mb for the 48 samples and the 10 Mb QTL_{NCLB} region, respectively. All sequence reads were mapped to the reference genome B73 AGPv02 by using the software BWA and BLASR. In addition, all sequence reads of every sample were *de novo* assembled, and also mapped to the B73 AGPv02 reference genome. Single nucleotide polymorphisms (SNPs) as well as small insertions/deletions (InDels) were detected for every sample in comparison to the reference genome. Since the number of detected polymorphisms was very high, the annotated genes within the target region were extracted and haplotypes were determined.

A first marker development for 40 SNPs was carried out by using the KASP technique. From 40 selected SNPs, 15 of them could be transformed into functional assays. The low success rate was due to strong sequence differences between the B73 reference genome and the genome of the donor lines. Anyhow, the 15 new markers were very useful for the better molecular characterisation of the introgressed donor alleles and the identification of unique resistance alleles at this target locus.

Multiple alignments of the *de novo* contigs were performed. The analysis revealed a high number of InDels, especially outside the exons regions. The insertions were often located close to the detected SNP positions, and disabled therefore the primer binding in the respective genotype.

For $QTL4_{CT}$, fine mapping and mapped based cloning were

successful and a patent application was filed by KWS and HHU. The data of $QTL5_{CT}$, $QTL8_{NUE}$, and $QTL10_{NUE}$ were received at the end of project and are evaluated by KWS after project completion according the pipeline established for target QTL_{NCLB}

The $QTL5_{CT}$ and $QTL10_{NUE}$ were at the start of the project too large to be investigated by sequence capturing methods. Fine mapping projects for both QTL were conducted outside of the AMAIZING project in a bilateral project between KWS and HHU. For both QTL, no suitable phenotyping assays could be developed and activities for the molecular cloning were stopped. However, the availability and development of additional markers from the sequencing and genotyping resources of this project and the SYNBREED project for the target regions allowed for a reduction of the QTL intervals by re-mapping those markers in the original QTL mapping population.

Whole-genome sequencing of donor lines and recurrent parents

Using next generation sequencing (NGS), sequence resources of the three recurrent parents were generated by the Illumina HiSeq2000 sequencing platform generating 100bp paired-end reads with a ~50x coverage. Altogether, five so-called long jumping distance libraries for each parent were analyzed containing 20kb insert size and varying genomic distances (180bp, 350bp, 3kb, 8kb). The NGS sequences were used to detect SNPs as well as InDels compared to the reference genome B73 supporting the development of new markers and the identification and discrimination of introgression blocks. For that analysis, a pipeline that was developed within the SYNBREED project was used combining publically available and established software. Altogether, more than 26x106 polymorphisms and a SNP density of ~1 SNP/200bp were detected using only libraries with an insert size ≥3kb as in the other libraries many PCR duplicates occurred. Using in addition data of the SYNBREED Illumina® 50K array, false discovery rates (FDR) of ~1.5% for homozygous variants were reached. FDRs for heterozygous variants were much higher (~87%) due to the high content of repetitive DNA and duplications in the maize genome, and were excluded from the further analysis.

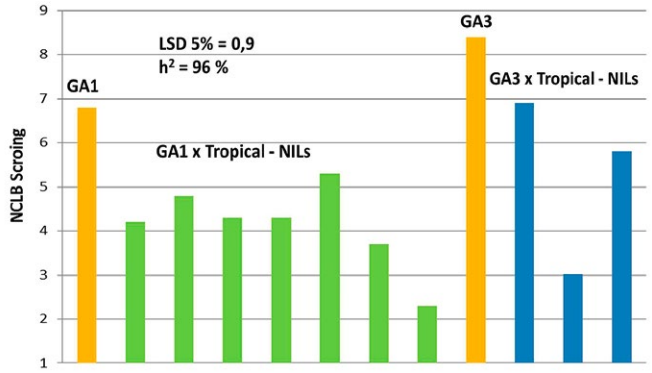


Fig. 2: Tropical alleles improve NCLB resistance. NCLB resistance was scored for recurrent parents GA1 and GA3 and ten selected (resistant in 2013 trials) NILs across three locations in 2014. 1 = resistant, 9 = susceptible.

Additionally, *de novo* assemblies of the three recurrent parents were performed using different assembly tools: Abyss (Simpson et al, 2009), SoapDenovo2 (Luo et al, 2012), Velvet (Zerbino 2010), and AllpathLG (Gnerre et al, 2011). It turned out that AllpathLG delivered the best assemblies resulting in a reconstruction of the reference B73 in the range of ~450 - 528 Mb with a 95% identity. Additionally, coding regions were defined using consensus models of spliced alignments of homologous proteins.

1.4 Phenotyping of target specific introgression lines

Experiments were conducted across several locations to assess NCLB resistance, cold tolerance, time to flowering, and biomass production (plant height). The main phenotyping of NILs was done in 2013 and 2014 using BC3S2 lines. In addition to the trait specific phenotyping, time to flowering, plant height, and other agronomic traits were assessed to characterize potential linkage drag associated with new exotic introgressions.

Phenotyping NUE and drought

The early recurrent parents were tested at 5 locations in 2013 and 4 locations in 2014 with 2 N levels each. In both years, none of the NILs with exotic introgressions showed significantly higher yield or stress tolerance compared to the recurrent parent.

The NILs from the mid-late recurrent parent were evaluated across 3 locations in 2013 and 4 locations in 2014 using 3 treatments per location (control, reduced water, low N). Results in 2013 were promising, particular under low conditions.

Phenotyping cold tolerance

Cold tolerance was evaluated across 3 to 4 locations differing in their mean temperatures during early growing period. Early plant development under cold spring conditions was assessed using early plant height and early vigor scoring as non-destructive measurements for cold tolerance. Effects of exotic introgressions were generally small and only 3 NILs for $QTL5_{CT}$ and 1 NIL for $QTL4_{CT}$ showed improved cold tolerance compared to the recurrent parent. These NILs were derived from donors from the CIMMYT Andean Highland program.

Phenotyping NCLB

Resistance against NCLB was scored under controlled conditions at 6 locations in 2013 and 2014, using 2 locations for the late, 2 locations for the early and 3 locations for the mid-late materials. Several NILs showed significantly improved resistance against NCLB compared to the recurrent parent (Figure 2). Finally, 7 NILs with recurrent parent

AMAIZING Targeted molecular genetic and bioinformatic approaches to increase genetic diversity in elite MAIZe breedING populations

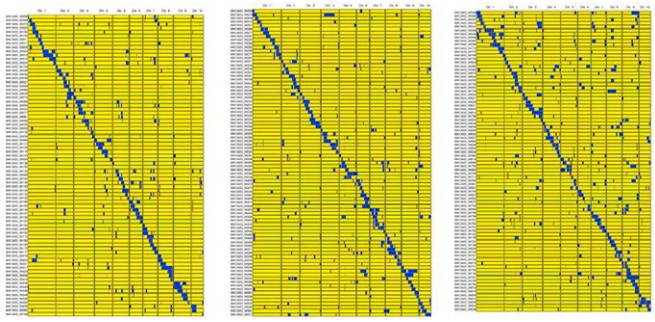


Fig. 3: Graphical genotypes of the three introgression populations (IL-Pop1, IL-Pop2, IL-Pop3).

GA1 and 3 NILs with recurrent parent GA3 were selected for further utilization in breeding. Results from will be used for a further molecular characterization of the new allelic series.

2. Optimization, development, and evaluation of introgression populations covering the complete exotic donor genome

The establishment of introgression libraries often requires several hundred backcross programs in parallel. The aim of this part of the project was to develop strategies for reducing this tremendous effort and employ probability theory to develop criteria for selection the optimal plants.

2.1 Development of biostatistics methods for construction and evaluation of optimized introgression populations

Breeding schemes with different combinations of backcross and selfing generations or DH line development were developed, and comparison of different population sizes during the backcross or selfing generations were carried out. In each generation, the best plants were selected using an in this project defined selection index.

Applying an optimized backcrossing scheme with 3 generations per year, pre-selection of complete donor chromosomes in BC1 and BC2 and a final DH step after BC3 was established. Results indicate that the applied procedure was cost effective and resulted in good donor genome coverage (Figure 3). The percentage of donor parent genome varied between 91 and 98 % in the libraries. Each IL had on average 4 to 5 % of donor genome with an average donor segment length of 20 to 22 cM.

2.2. Analysis of optimized introgression populations

Previous investigations have shown that the genetic architecture of the trait under consideration and its heritability determine the power of QTL detection with introgression populations (IP) (Falke & Frisch 2010). Applying different QTL detection models, the computer simulations showed that genome-wide prediction employing heteroscedastic marker variances had a greater power and a lower false positive rate compared with homoscedastic marker variances when the phenotypic difference between the donor and recipient lines was controlled by few genes.
Additionally, methods were developed to investigate the IP in terms of the effect of number, size and distribution of donor segments on the power and false positive rate. Positive and negative donor effects could be detected.

2.3 Genotyping and molecular selection for adaptation

Populations derived from the first selfing (F2/F3) and the first backcross (BC1S1) were developed in parallel to the IL populations and were phenotyped for adaptation (photoperiod sensitivity, flowering time) at two environments. The results were combined with the already available information on adaption QTL to detect novel genomic regions responsible for the lack of adaptation of the exotic donors. In general, the impact of adaptation QTL was large in the F2/F3 generation with an average exotic donor genome proportion of 50%. Already one generation of backcrossing reduced the significance of these QTL drastically. These results confirmed the strategy of advanced backcrossing for the utilization of unadapted exotic germplasm.

2.5 Phenotyping and data analysis

The three introgression populations could be evaluated in 2014 for line per se and testcross performance. Target environments were South and South-East Europe (Rumania, Hungary, and France). Grain yield, NUE and drought tolerance were evaluated across 4 locations with 3 treatments each (low N, low irrigation, control). Drought stress did not occur to a larger extend in 2014. Under low N conditions, several donor segments showed positive effects on grain yield without negative side effects on grain dry matter content, a measure for maturity. ILs carrying these segments will be re-tested to verify the effects.

ILs were also phenotyped for resistance to NCLB in 2014 across 2 locations in the target environments. Results indicate a good and highly significant variation for resistance. Donor segments with large effects on resistance were detected on chromosomes 1, 2, 5 in IL-Pop1, on chromosome 9 in IL-Pop2, and on chromosomes 5 and 10 in IL-Pop3. Some of the effects correlate with effects for late flowering, and will therefore not be considered for utilization. The other effects will be utilized to improve resistance against the important fungal disease NCLB.

Conclusion

Within the project AMAIZING, novel strategies for the systematic use of tropical germplasm to improve temperate maize were developed. The utilization of 3 generations per year with pre-selection of complete donor chromosomes, and state of the art marker technologies allowed for a fast and efficient development of series of exotic donor alleles in an elite background. Furthermore, several approaches for a powerful QTL detection in IL populations were inves-

tigated and applied. A series of new promising alleles was identified, particular for resistance to NCLB. ILs with positive effects will immediately be used in breeding and can be regarded as an excellent source for further fine mapping or gene cloning projects.

References

- *Falke KC, Frisch M (2010). Power and false positive rate in QTL detection with near-isogenic line libraries. Heredity, doi:10.1038/hdy.2010.87*
- *Gnerre, S. et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A 108, 1513-8 (2011)*
- *Goodman MM (1985). Exotic maize germplasm: status, prospects, and remedies. Iowa State J Res 59: 497–527*
- *Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1, 18 (2012)*
- *Reif J, Hamrit S, Heckenberger M, Schipprack W, Maurer HP, Bohn M & Melchinger AE (2005). Genetic structure and diversity of European flint maize populations determined with SSR analyses of individuals and bulks. Theor Appl Genet 111: 906-913*
- *Simpson, J.T. et al. ABySS: a parallel assembler for short read sequence data. Genome Res 19, 1117-23 (2009)*
- *Troyer AF (1999). Background of U.S. hybrid corn. Crop Sci. 39: 601–626*
- *Willmot DB & Pollak LM (2003) Molecular marker analysis of exotic maize introgression after selection. ASA-CSSA-SSSA Annual Meeting. Paper No. C08-Willmot558711-Oral*
- *Zerbino, D.R. Using the Velvet de novo assembler for short-read sequencing technologies. Curr Protoc Bioinformatics Chapter 11, Unit 11 5 (2010)*

BARLEY-FORTRESS Targeted exploitation of basal defense genes for pathogen resistance

Antony Bacic[1], Geoffry Fincher[2], Anja Hahnemann[3], Ralph Hückelhoven[4], Viktor Korzun[5], Jochen Kumlehn[6], Klaus Pillen[7], Ulrich Schaffrath[8], Günther Schweizer[9], Patrick Schweizer[6], Udo Seiffert[10] (Authors are listed in alphabetical order)

1 ARC Centre of Excellence in Plant Cell Walls, School of Botany, University of Melbourne, Parkville, VIC 3010, Australia
2 ARC Centre of Excellence in Plant Cell Walls, School of Agriculture, Food and Wine, University of Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia (UoA)
3 Saatzucht Josef Breun GmbH & Co. KG, Amselweg 1, D-91074 Herzogenaurach, Germany (BREUN)
4 Technische Universität München, Emil-Ramann-Str. 2, 85350 Freising, Germany (TUM)
5 KWS LOCHOW GMBH, Ferdinand-von-Lochow-Straße 5, 29303 Bergen, Germany (KWS-L)
6 Leibniz Institut für Pflanzengenetik und Kulturpflanzenforschung, Corrensstrasse 3, 06466 Stadt Seeland OT Gatersleben, Germany (IPK)
7 Martin-Luther-University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle/Saale, Germany (MLU)
8 RWTH Aachen University, Department of Plant Physiology (Bio III), Worringerweg 1, 52074 Aachen, Germany (RWTH)
9 Bayrische Anstalt für Landwirtschaft, Am Gereuth 8, 85354 Freising, Germany (LfL)
10 Fraunhofer Institute for Factory Operation and Automation IFF, Sandtorstrasse 22, 39106 Magdeburg, Germany (IFF)

Background

Quantitative disease resistance is the most promising approach to ensure safe and sustainable crop production, while at the same time reducing pesticide input. However, it is known that this form of resistance is inherited in a multigenic manner by quantitative trait loci, which makes its targeted exploitation by breeders difficult (Schweizer and Stein, 2011). In addition to the complexity of introgressing multiple favourable alleles into one genetic background, QTL for resistance may be linked to unfavourable alleles for other traits, such as yield, by so called "linkage drag". The above-mentioned problems can be reduced by knowledge-based approaches consisting in the targeted introgression of favourable alleles of genes with known function in quantitative resistance. One of the well-known, important barriers of plants against invading microorganisms is the cell wall. In response to attack by fungal and oomycete pathogens, which usually breach their hosts' cell walls by mechanical force or by secreted lytic enzymes, plants often form local cell wall appositions (papillae) as an important first line of defence (Huckelhoven, 2007). Although a comprehensive picture of the involvement of cell wall biosynthetic enzymes in the formation of these papillae is still lacking, especially in cereal crops, a number of strong candidate genes contributing to cell-wall related defence have been recently identified in barley and wheat (Bhuiyan *et al.*, 2009; Hu *et al.*, 2009; Chowdhury *et al.*, 2014; Chowdhury *et al.*, 2016; Douchkov *et al.*, 2016). As an alternative to resistance improvement by traditional or gene marker-assisted breeding, transgenic crops can be developed, which carry for example engineered forms of cell-wall related genes. However, usually just one gene of interest is inser-

Gene	QTL	Assoc	RNAi	OEX	Regul	Resistance donor	Used for	Chr	Pos (cM)
Unknown_1	YES	YES	YES		YES	Nigrate	1 gene BC2	2H	136,2
Secondary_metabolism_1	YES		YES		YES	Cebada Capa	1 gene BC2	6H	59,8
Cell_death_2	YES	YES	YES		YES	OWB Dom	1 gene BC2	5H	43
Secretion_cell-wall_3	YES		YES		YES	L 94 (Donor A)	4-way Stack	2H	67,4
Secretion_cell-wall_4	YES		YES	YES	YES	Scarlett S42IL-124 (Donor B)	4-way Stack	4H	135,8
Secretion_cell-wall_2	YES		YES		YES	OWB Rec (Donor C)	4-way Stack	7H	8,8
Secondary_metabolism_2	YES	YES		YES	YES	HOR2932 (Donor D)	4-way Stack	3H	66,7

Table 1: Resistance donors from Plant Genetic Resources carrying potentially favourable alleles of candidate genes co-localizing with resistance QTL. Converging evidence for a role in plant-pathogen interactions derived from mapping-, transient RNAi- and transient over-expression experiments is highlighted by "YES" in green field.

BARLEY-FORTRESS Targeted exploitation of basal defense genes for pathogen resistance in barley

LineID	Introgressed gene(s)	# QTL	Powdery Mildew field	Powdery Mildew GH	Scald field	Scald GH	M. oryzae DLA	Leaf Rust field	Average PM_scald_blast
KWSBambina			100.0	100.0	100.0	100.0	100.0	100.0	100.0
KWSBambina			100.0	100.0	100.0	100.0	100.0	100.0	100.0
50067		0	67.3	31.2	103.0	97.3	59.0	120.9	75.0
50104			131.4	68.0	103.1	96.5	77.6	99.1	95.5
50103			128.2	100.5	103.0	98.2	57.7	111.6	96.5
50002	Secretion_cell-wall_3	3	88.6	64.9	96.0	96.5	59.4	104.9	81.1
50116			96.5	18.9	115.0	98.2	99.4	112.2	83.6
50062			85.3	59.1	83.0	104.7	72.4	114.0	84.6
50109			105.9	62.2	81.3	104.7	85.5	104.7	87.9
50005			87.3	66.5	131.0	81.8	73.5	102.4	88.0
50006			112.7	8.9	130.0	103.6	93.7	104.9	89.8
50111			90.0	54.6	104.0	99.8	106.1	124.4	92.9
50112			89.8	24.7	125.0	101.4	133.8	122.0	93.2
50110			137.6	47.9	107.8	90.8	83.6	109.3	93.5
50008			100.9	28.5	104.0	104.7	134.1	124.4	94.4
50010			101.7	38.7	90.0	101.4	147.4	122.0	95.9
50016			65.5	8.4	90.0	86.7	237.2	112.2	97.6
50003			95.6	75.0	117.0	99.0	105.1	124.4	98.3
50007			98.3	2.4	123.0	103.9	164.6	109.8	98.4
50001			83.6	76.3	107.0	102.3	129.2	124.4	100.1
50115			118.8	24.6	133.0	100.3	130.4	112.2	101.4
50015			42.8	0.8	86.0	70.4	328.5	107.3	105.5
50014			96.5	19.6	108.0	95.4	209.9	112.2	105.7
50117			84.7	59.7	105.0	98.5	169.1	109.8	107.0
50113			127.5	6.1	142.0	104.7	168.2	128.8	108.9
50012			107.0	31.6	148.0	98.5	171.8	124.4	110.6
50013			111.4	44.1	160.0	101.4	142.9	112.2	112.0
50011			125.3	29.2	172.0	104.7	138.3	151.2	113.9
50004			114.8	110.6	107.0	104.7	143.2	127.3	116.1
50114			134.9	48.5	168.0	104.7	163.6	126.8	124.0
50009			103.1	77.9	140.0	101.4	213.8	102.4	127.2
50121	Unknown_1	1	86.5	47.4	103.0	94.9	36.8	100.0	72.1
50120			72.1	36.0	108.0	100.6	50.1	100.0	73.4
50118			71.2	9.1	110.0	104.7	102.1	112.2	79.4
50019			71.2	110.9	104.0	94.1	28.2	90.2	81.7
50125			81.7	32.4	124.0	101.4	91.0	109.8	88.1
50017			72.5	26.2	141.0	103.6	88.8	97.6	88.4
50124			98.3	42.3	124.0	100.3	70.4	104.9	87.1
50020			80.3	17.0	130.0	100.6	120.6	102.4	89.7
50018			72.1	70.4	140.0	98.2	71.7	97.6	90.4
50022			96.1	28.5	135.0	104.7	97.4	117.1	92.3
50021			108.3	41.7	107.0	94.1	121.9	109.8	94.6
50123			90.8	98.2	122.0	92.4	77.3	97.6	98.2
50119			85.6	85.4	132.0	103.1	90.4	107.3	99.3
50023			73.8	21.9	159.0	104.7	165.0	119.5	104.9
50126	Secondary_metabolism_2	1	52.4	28.1	108.0	82.6	27.0	100.0	59.6
50130			49.8	43.8	92.0	74.4	38.8	107.3	59.7
50102			75.6	19.5	103.0	80.5	54.9	97.7	62.1
50069			46.2	75.1	73.4	97.1	33.9	107.0	65.1
50026			65.9	48.7	101.0	90.8	37.1	104.9	68.7
50030			79.0	87.9	102.0	87.5	0.8	103.9	71.5
50027			73.4	39.5	116.0	89.2	41.8	97.6	72.0
50128			95.2	43.1	96.0	87.5	41.2	122.0	72.6
50134			80.7	51.3	132.0	90.0	39.8	109.8	74.7
50031			79.0	96.4	116.0	85.9	0.9	109.8	75.6
50135			57.2	64.9	118.0	85.1	78.0	100.0	80.6
50131			83.8	117.7	112.0	87.5	0.9	107.3	81.0
50029			69.4	94.1	108.0	91.6	59.4	95.1	84.5
50032			58.5	125.1	107.0	82.6	54.1	109.8	85.5
50028			80.3	72.7	134.0	88.4	91.5	100.0	89.4
50024			80.3	59.2	103.0	99.0	109.4	114.6	89.6
50133			100.4	105.9	127.0	100.6	46.4	109.8	98.1
50025			101.3	101.2	154.0	82.6	104.9	90.2	108.8
50126			120.1	149.3	120.0	86.7	69.0	100.0	109.0
50127			103.5	140.5	119.0	100.6	81.0	122.0	110.7
50132			73.8	102.6	138.0	90.8	165.3	102.4	113.7
50136			72.1	96.0	131.0	88.7	195.1	102.4	116.2
50095	Secretion_CW_3	1	47.2	5.1	96.9	67.9	56.1	93.5	54.7
50107			30.8	0.8	75.0	83.4	99.9	120.9	57.8
50048	Cell_death_2	2	74.7	41.4	121.0	88.4	55.0	90.2	76.1
50049			80.8	62.4	111.0	91.9	64.5	85.4	82.1
50145			104.4	87.7	124.0	83.4	31.4	102.4	88.2
50146			74.2	100.1	98.0	72.0	98.5	92.7	88.2
50038			122.3	46.2	138.0	85.9	51.8	92.7	88.4
50050			84.7	89.0	118.0	89.2	59.7	97.6	89.4
50039			92.6	47.2	124.0	90.5	93.6	97.6	89.6
50136			128.6	60.1	155.0	79.4	25.0	104.9	89.6
50051			81.2	85.2	150.0	90.0	51.0	100.0	91.5
50040			94.8	74.1	76.0	90.8	125.3	104.9	92.2
50055			125.8	75.5	111.0	77.7	75.2	102.4	93.0
50037			117.0	65.2	164.0	91.8	32.0	102.4	94.0
50140			121.0	68.6	126.0	76.9	83.6	95.1	95.2
50137			134.1	41.2	136.0	86.7	104.1	115.6	96.3
50038			129.7	38.6	176.0	91.8	46.0	126.8	96.5
50033			112.2	67.9	116.0	92.4	96.8	100.0	97.1
50139			81.2	104.0	116.0	85.1	99.9	92.2	97.2
50047			113.1	63.7	131.0	91.6	89.8	100.0	97.6
50143			102.2	86.7	134.0	100.6	104.4	114.6	101.6
50034			130.6	39.5	179.0	94.3	66.3	109.8	101.9
50042			108.3	75.2	152.0	94.9	84.7	102.4	103.0
50041			101.7	45.4	163.0	104.7	107.8	104.9	104.5
50045			115.3	89.0	142.0	90.5	84.1	124.4	105.3
50046			134.1	55.0	154.0	100.6	90.8	119.5	106.9
50141			115.7	101.0	119.0	89.2	130.7	101.0	111.1
50043			114.4	122.3	119.0	90.0	111.6	102.4	111.3
50144			121.4	92.1	142.0	81.2	150.8	114.6	111.4
50044			95.6	89.2	132.0	97.3	158.1	100.0	112.3
50142			137.6	99.4	156.0	102.3	134.5	126.0	124.1
50096	Secretion_CW_2	2	34.8	44.2	64.1	84.3	35.7	93.0	52.6
50064			34.8	29.9	66.4	72.0	69.5	107.0	54.3
50060			62.0	38.8	104.0	62.2	59.7	97.6	65.3
50059			79.8	32.7	93.0	62.2	59.5	104.9	65.4
50057			81.7	20.6	124.0	59.7	51.8	109.8	67.6
50053			82.4	8.9	88.0	90.0	84.9	102.4	69.1
50061			81.7	27.2	107.0	80.3	71.0	97.6	69.4
50150			84.3	40.9	118.0	71.2	43.6	114.6	71.6
50058			82.1	23.1	116.0	81.0	80.6	100.0	72.7
50054			119.2	6.8	92.0	94.1	70.6	100.0	75.2
50151			99.1	33.6	122.0	69.5	56.3	104.9	76.1
50154			86.0	56.6	118.0	78.5	43.3	104.9	76.5
50149			67.7	78.3	128.0	94.9	61.0	102.4	88.0
50153			103.5	88.8	96.0	79.1	80.5	100.0	87.6
50155			72.1	86.1	99.0	100.6	84.7	103.9	88.5
50152			101.3	81.9	116.0	75.3	89.4	121.5	88.6
50063			89.5	79.0	112.0	100.1	66.1	112.7	89.3
50052			122.7	8.9	144.0	95.7	87.7	117.1	90.0
50148			53.0	65.9	103.0	94.1	79.8	107.3	91.1
50155			120.1	0.8	141.0	98.2	101.1	102.4	92.1
50062			84.3	85.9	104.0	98.2	93.2	100.0	93.1
50147			107.0	108.4	124.0	94.3	80.5	124.4	98.9
50056			131.0	99.4	122.0	94.9	74.9	122.0	102.6
50105	Secretion_CW_3 & 4	2	83.3	13.9	85.9	94.9	54.3	95.3	66.4
50088			55.2	22.4	67.2	78.9	60.0	102.3	58.3
50106			32.8	5.9	75.0	88.7	109.0	100.0	60.5
50097			79.2	14.9	103.1	100.3	88.2	107.0	77.1
50063			62.9	40.3	94.5	99.8	89.8	139.5	77.5
50090			113.1	4.8	114.1	104.7	66.4	114.0	80.6
50067	Secretion_CW_2 & 4		49.8	14.1	60.2	76.1	15.9	90.7	43.2
50061			61.5	11.3	64.1	67.9	19.6	100.0	44.7
50073			43.3	13.6	64.4	90.0	27.0	74.4	51.8
50100			32.6	0.8	103.1	53.2	71.4	136.0	52.1
50068	Secretion_CW_4 & Sec.metab_2		64.3	18.2	90.6	84.3	50.0	93.0	61.5
50106			39.8	0.8	62.5	78.5	39.4	104.7	42.0
50101			64.3	34.3	73.4	51.5	15.4	100.0	47.6
50092	Secretion_CW_2 & Sec.metab_2		51.6	56.7	46.9	77.7	30.6	95.3	52.7
50094			36.2	11.1	64.1	80.2	44.4	108.3	47.2
50085			46.2	17.9	62.5	78.5	48.2	120.9	50.7
50096			55.2	7.7	100.0	75.0	24.1	114.0	52.4
50077	Secretion_CW_2 & 3 & 4	3	65.6	3.7	112.5	85.4	15.1	109.3	58.5
50074			74.2	8.6	76.6	94.9	28.2	86.0	58.5
50091			50.7	18.9	75.0	104.7	35.3	153.5	58.9
50080			78.7	29.2	70.3	86.7	36.4	90.7	60.3
50068			68.8	10.0	95.3	95.7	73.1	109.3	69.6
50093	Secretion_CW_2 & 3 & Sec.metab_2		40.3	6.6	90.6	94.9	59.4	116.3	57.8
50064			47.5	15.8	88.3	101.4	66.6	127.9	63.9
50078	Secretion_CW_2 & 4 & Sec.metab_2		49.8	13.1	50.0	77.4	32.1	95.3	44.5
50069			70.6	5.3	71.1	91.1	19.8	86.0	51.8
50099			30.8	0.8	93.8	43.1	107.8	107.0	55.1
50075			74.7	67.6	65.6	65.4	21.5	97.7	59.0
50071			74.9	50.6	76.1	80.7	29.2	93.5	62.7
50076	Secretion_CW_2 & 3 & 4 & Sec.metab_2	4	62.4	3.6	70.3	79.8	55.4	114.0	54.3
50079			71.5	13.2	67.2	79.6	40.5	102.3	54.4
50065			76.0	21.6	65.6	83.8	57.8	114.0	57.0
50070			89.1	33.8	70.3	79.9	15.8	102.3	57.8
50072			67.4	55.9	73.4	67.8	52.6	95.3	63.4
Donor D2	Resistance donors		45.2	1.1	85.6	58.9	119.6	137.2	58.1
Donor D2			40.6	1.1	104.0	58.9	165.5	136.6	74.0
Donor D3			49.3	0.8	84.0		236.7	146.3	87.5
Donor D4			27.8	4.4	84.4	71.2	9.1	137.6	39.3
Donor D4			49.8	4.4	124.0	71.2	8.2	117.1	51.5
Donor D5a			40.7	77.4	58.6	104.7	27.6	101.2	61.8
Donor D6			134.1	95.6	166.0		208.4	163.4	151.0
Donor D8a			110.4	116.5	122.2	103.9	20.7	109.3	89.2
Donor D8a			116.9	116.5	111.6	103.9	29.7	127.3	95.1

ted resulting in measurable but not sufficient trait improvement (Schweizer, 2008; Shin *et al.*, 2008; Bahrini *et al.*, 2011; Ferrari *et al.*, 2012; Chen *et al.*, 2016). This often limits efforts to further develop such transgenic events.

Barley suffers from a number of fungal diseases, the most important in Europe being powdery mildew caused by *Blumeria graminis* f.sp. *hordei*, leaf rust caused by *Puccinia graminis* f.sp. *hordei*, scald caused by *Rhynchosporium secalis*, net blotch caused by Pyrenophora teres, spot blotch caused by Bipolaris sorokiniana and Fusarium head blight caused by a complex of Fusarium *graminearum/culmorum* (Schweizer, 2014). Blast, which is caused by *Magnaporthe oryzae*, is well known as the most important disease in rice cultivation worldwide. However, this fungus is also able to colonize and cause considerable yield losses on barley and wheat, and threatens to move into temperate climatic zones as a consequence of global warming (Zellerhoff *et al.*, 2006). Overall, world-wide yield losses caused by fungal diseases averages approximately 10% in in barley, despite the use of fungicides (Oerke, 2006). This underlines the importance of developing durable, quantitative resistance for these major diseases in barley.

The project

The BARLEY-FORTRESS consortium consisting in academic and private partners from Germany and Australia set out to improve quantitative resistance in barley by a suite of approaches, which included the targeted introgression of candidate-gene alleles from resistance donors as well as the generation of transgenic plants. The main focus lied on the introgression and transformation of candidate genes for improved defence at the cell-wall, but also included the silencing of proposed susceptibility-related genes in transgenic plants. The gene-based introgression approach was performed in the elite spring barley cultivar KWS BAMBINA, which exhibits moderate to high susceptibility to different fungal pathogens. The transgenic approaches were performed in barley cultivar Golden Promise and in the elite cultivar JB DORINKA, which are easily transformable and suffer from moderate to high susceptibility to relevant fungal diseases of the crop.

Achievements

Introgression of beneficial alleles from plant genetic resources

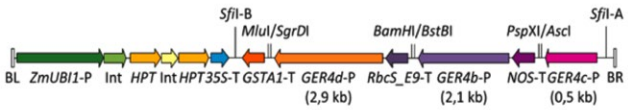
In a back-crossing approach at MLU and KWS-L candidate genes selected based on pre-existing knowledge of project

Fig. 1: Results of all disease assessments in field trials and greenhouse experiments. *Each colour-coded field represents one line/test combination. Grey, same susceptibility as KWS BAMBINA; green, more susceptible than KWS BAMBINA; red, more resistant than KWS BAMBINA. Please note the increase in intensity of red coloring in multiple introgression lines (#QTL: 2-4). Numbers inside colored fields indicate percent susceptibility relative to KWS BAMBINA.*

partners were introgressed in KWS BAMBINA singly or as gene-stacks. For this purpose, SNP markers were developed and used inside the candidate genes. In total 107 single-introgression lines and 46 lines with variable combinations of two to four candidate genes were generated. In a joint effort the generated lines, after being multiplied by KWS-L in New Zealand, were phenotyped for multiple disease resistance by MLU, KWS-L and BREUN in 2014 at 3 locations in the field, and by 4 institutions in the greenhouse. In the field, enhanced resistance of a number of lines against *B. graminis* and *R. commune*, but not against *P. graminis*, was found across locations. In the greenhouse, the partner at RWTH participated in the disease scoring of this material in response to cereal blast (*Magnaporthe oryzae*) and leaf rust (*Puccinia hordei*).

Using the primary leaf of intact plants for inoculation with *M. oryzae*, it became obvious that most of the quadruple introgression lines showed an enhanced resistance phenotype. The result for resistance to *P. hordei* was less clear and more variable among different introgression lines. Partner TUM and its subcontractor LfL tested lines for quantitative resistance (QR) against *Bipolaris sorokiniana* or *Rhynchosporium commune*. Three of four QTL (from D_4,D_5a and D_8a) were found to support QR against *R. commune*. Combinations of these QTL appeared more effective in controlling *R. commune* than each QTL alone. The same QTL were either not effective or even suppressive in terms of QR to *B. sorokiniana*. Donor 2 introgression also was contra-productive in terms of QR to *R. commune*. At IPK the resistance against *B. graminis* was tested on greenhouse-grown seedlings, and strong resistance of a number of lines was observed. Overall, a trend towards multiple disease resistance was observed in lines carrying more than one QTL introgression. Furthermore, individual lifestyles of fungi may determine as to whether individual QTL have positive or negative effects, and stacking QTL can be effective in enhancing QR to fungal pathogens.

A



B

Construct	Pathway	Promoter GER4b	Promoter GER4c	Promoter GER4d
TriOEX1	Membrane- and cell-wall crosslinking	Secretion_cell-wall_GOI10	Secretion_cell-wall_GOI5	Secretion_cell-wall_GOI8
TriOEX2	Lignin-dependent cell-wall fortification	Secondary_metab_GOI1	Secretion_cell-wall_GOI5	Secondary_metab_GOI2
TriOEX3	Lignin-independent cell-wall fortification	Secretion_cell-wall_GOI7	Secretion_cell-wall_GOI6	Secretion_cell-wall_GOI2
TriOEX4	ROS burst	Secretion_cell-wall_GOI5	Secretion_cell-wall_GOI11	Signalling_ GOI3

Fig. 2: Gene combinations of triple-overexpression (triOEX) cassettes 1 to 4. GOI1-11, gene of interest 1-11 (internal nomenclature).

Function and diversity at a QTL for powdery mildew resistance

The *Bgh* resistance QTL QPm.S42IL-4H.c on chromosome 4H (Schmalenbach *et al.* 2008) served as a model to study quantitative resistance in plant breeding. The exotic QTL allele, present in the wild barley introgression line S42IL-124, caused a significant reduction of powdery mildew symptoms in multiple field tests and, in addition, was verified in a detached leaf assay at IPK. The *Germin 4 (HvGER4)* locus containing a cluster of *germin-like* genes (Himmelbach *et al.*, 2010) turned out to be a strong candidate to explain the QTL effect because stable *HvGER4* RNAi lines tested at IPK were super-susceptible to *B. graminis*. At MLU, using a set of app. 5,000 segregating high-resolution offspring lines derived from S42IL-124 (Schmalenbach *et al.*, 2011), the presence of *Bgh* resistance in the material could be verified in a field test in 2013. However, because no recombination event within the cluster was identified we used a second population, the nested association mapping (NAM) population HEB-25, consisting of 1,420 segregating barley lines, to screen for recombinants and *Bgh* resistance in the field. Based on a genome-wide association mapping (GWAS) screen with 5,727 informative single nucleotide polymorphism (SNP) markers, a new map of the *HvGER4* locus with a widely increased genetic resolution was developed and revealed three linked *Bgh* resistance loci on chromosome 4H: The first locus corresponded to the known *Bgh* resistance gene *mlo* (position 4H: 100.6cM), and the second and third QTL mapped to positions 4H: 113.1cM and 4H: 113.7cM, respectively. However, the *HvGER4* gene cluster was placed between the latter two QTL at position 4H: 113.4cM and did not show a significant association with *Bgh* resistance. We conclude that the *Germin 4* gene cluster is presumably not causing *Bgh* resistance in HEB-25 lines. Partner IPK focused on revealing the dynamics of the *HvGER4* locus by sequencing BAC clones from eight barley accessions differing in *HvGER4* mRNA levels and powdery mildew resistance. Work still in progress suggests a considerable level of copy-number variation, which however appears not to be directly related to expression levels. Therefore, gene silencing e.g. by partial DNA methylation might play a role at this locus.

Conditional overexpression of stacks of defence-related genes

The aim of this part of the project was the generation of plants carrying triple-OEX (over-expression) cassettes under the control of the pathogen-inducible *HvGER4b*, *4c* and *4d* promoters. The rationale behind this was to reduce the costs of resistance for the plant. However, the used promoters of the *HvGER4* family are only expressed in the plant epidermis (incl. roots) and therefore, not suitable for protecting plants against pathogens that evade an epidermal penetration by entering through stomata (Himmelbach *et al.*, 2010). Candidate genes had been selected previously and were expected to render the plants more resistant against fungal pathogens (see for example (Douchkov *et al.*, 2014). A novel cloning cassette for the overexpressi-

Tab. 2: Summary of transgenic barley plants carrying triple-OEX constructs.

Construct	GOLDEN PROMISE	DORINKA
pIPKTA47	11	-
pIPKTA47_TriOEX_4bGUS	1	-
pIPKTA47_TriOEX_4cGUS	1	-
pIPKTA47_TriOEX_4dGUS	2	-
pIPKTA47_TriOEX1	31	15
pIPKTA47_TriOEX2	12	11
pIPKTA47_TriOEX3	6	19
pIPKTA47_TriOEX4	0	0

on of three different candidate genes was designed and generated at IPK. The cassette contained three different *HvGER4*-promoters to allow for different levels of pathogen-inducible transgene expression and to minimize the risk of cassette silencing by DNA methylation. This tool was used to generate constructs with different combinations of selected candidate genes (Fig. 2). At RWTH the over-expression construct 1 (triOEX1) was cloned and sent to IPK for transformation into barley. IPK generated triOEX2 and triOEX3, and TUM generated triOEX4.

The four different constructs were to be used to transform elite material of barley of the two participating breeding companies (JB FLAVOUR and KWS BAMBINA). Due to a lack of information regarding tissue culture behaviour and amenability to *Agrobacterium*-mediated gene transfer, initial experiments were performed using reporter gene constructs. Although a high number of explants were treated, both cultivars showed an unexpected recalcitrance. Therefore, constructs were used in parallel in the model cv. Golden Promise. Other elite breeding lines of partner BREUN tested subsequently were found to be amenable to transformation and consequently, line JB11787mz1 now released as new cv. JB DORINKA was used in the experimental work. Table 2 summarizes all primary transgenic barley plants which were confirmed by PCR to carry intact triple-OEX cassettes. The fact that zero primary transformant carrying cassette No 4 could be obtained probably reflects construct toxicity.

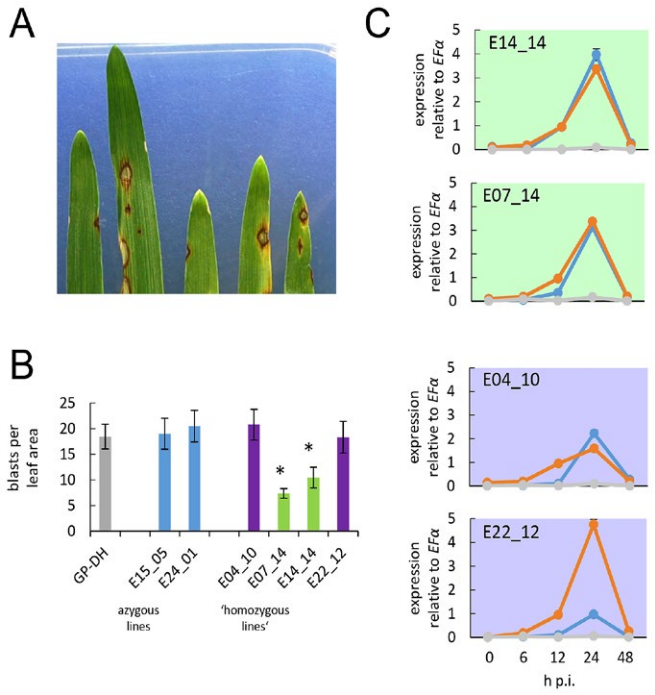


Fig. 3: Enhanced resistance against M. oryzae caused by TriOEX1. (A) Example of blast lesions on barley cv. G. Promise. (B) Enhanced resistance of some T2 lines. (C) Pathogen-induced transgene expression revealed by RT-qPCR. Transcript levels relative to the reference gene elongation factor alpha are shown. Colour code: blue, Secretion_cell-wall GOI5; red, Secretion_cell-wall GOI10; grey, Secretion_cell-wall GOI8. GP-DH, double-haploid line of G. Promise used for transformation.

As proof of principal for the functionality of the *HvGER4*c promoter, TUM analyzed GER4c::UmPEP1 transgenic barley plants, that expressed the smut virulence effector UmPEP1 specifically only after inoculation with *Bgh* and showed a super-susceptible phenotype (Hemetsberger *et al.*, 2015).

After receiving transgenic T0 plants for TriOEX1, RWTH selected promising events based on infection assays with *M. oryzae*. Enhanced resistance of derived, putative homozygous T2 lines could be confirmed and homozygosity was validated by Southern-blot analysis at IPK. Reverse transcription-qPCR revealed pathogen-induced expression of all three transgenes of the cassette (Figure 3). In sum, RWTH was able to select promising T2 lines of this const-

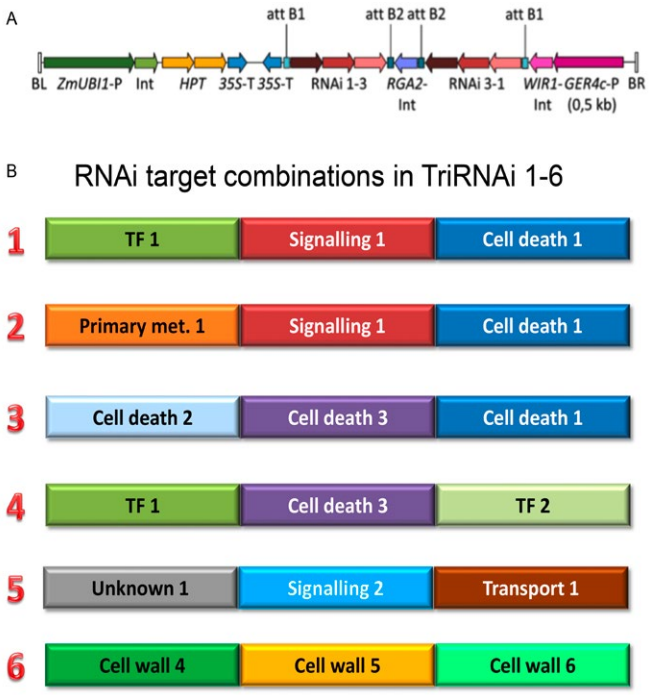


Figure 4: Gene combinations of triple-RNAi hairpin cassettes. Indicated are the processes and pathways the respective genes of interest are involved. TF, transcription factor.

Tab. 3: Summary of transgenic barley plants with triple RNAi constructs including PCR analysis.

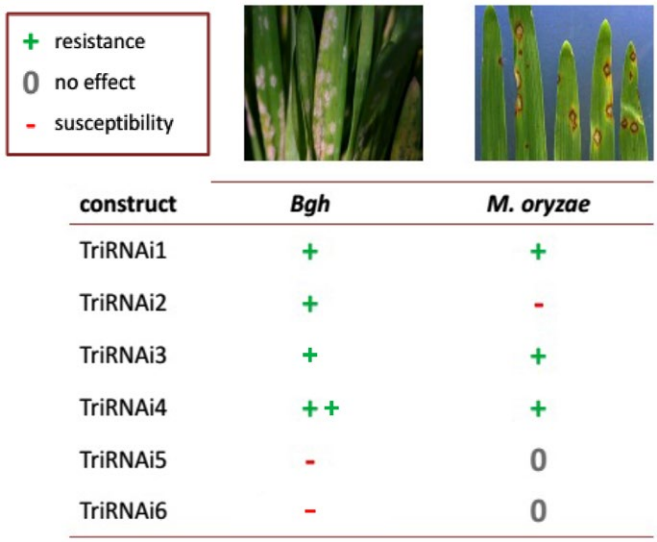
Construct	No of regenerants	only 1st repeat present	only 2nd repeat present	both repeats present
pIPKTA46_TriRNAi1	26	1	5	18
pIPKTA46_TriRNAi2	30	1	3	21
pIPKTA46_TriRNAi3	15	3	0	12
pIPKTA46_TriRNAi4	14	0	2	10
pIPKTA46_TriRNAi5	12	0	4	6
pIPKTA46_TriRNAi6	9	1	1	7

ruct with an enhanced quantitative resistance phenotype against *M. oryzae*.

The T1 progeny of TriOEX1, plus selected T2 lines were also tested at IPK and TUM for enhanced resistance to *B. graminis*. Several T1 progeny and T2 lines with significantly enhanced resistance could be identified at IPK, whereas no effect was found at TUM. These conflicting results might reflect different levels of GER4 promoter induction at the two locations, or differences in the aggressiveness of the used fungal strains. For TriOEX2 and TriOEX3 no enhanced resistance against any of the pathogens was observed.

Conditional silencing of stacks of susceptibility-related genes

As an alternative approach to enhancing resistance by triOEX constructs, we set out to reduce susceptibility by reducing the expression of potential susceptibility-related genes (selected based on their impact on powdery mildew infections). TUM and IPK defined potential host susceptibility factors as possible targets for RNAi-mediated QR (Huckelhoven *et al.*, 2013). To further validate barley alcohol dehydrogenase 1 as a suitable target for pathogen-triggered RNAi (Pathuri *et al.*, 2011), transgenic lines constitutively silenced for ADH1 expression were phenotyped. This showed significantly reduced susceptibility to *Bgh* along with strongly reduced amounts of ADH1 transcripts and reduced ADH enzyme activity in corresponding plant extracts. In order to minimize the risk of pleiotropic side effects of gene silencing, triple-RNAi constructs were generated by using the pathogen-inducible *HvGER4*c promoter. Triple RNAi constructs designed to target three different barley susceptibility genes were generated at IPK and transformed into transgenic plants (Fig. 4).



construct	*Bgh*	*M. oryzae*
TriRNAi1	+	+
TriRNAi2	+	-
TriRNAi3	+	+
TriRNAi4	++	+
TriRNAi5	-	0
TriRNAi6	-	0

Figure 5: Resistance phenotypes of TriRNAi cassettes against two fungal pathogens. T1 progeny of primary transformants was analysed in the greenhouse at the seedling stage. "++", enhanced resistance observed both at TUM and IPK. Bgh, B. graminis f.sp. hordei.

Former transformation experiments showed that not all transgenic RNAi plants contain both parts of the inverted-repeat construct. Therefore, repeat-specific primers were designed and PCR of genomic DNA was performed. Only T0 plants giving rise to both amplicons were further analyzed. A summary of all transformation efforts is given in Tab. 3.

For all constructs and genotypes, grains were harvested and distributed amongst the different partners for resistance tests. As a result, T1 families with positive responses were identified and most promising populations were used to estimate transgene copy number and to generate homozygous progeny using DH technology. Southern blot experiments with genomic DNA of T2 plants were performed and in the majority of cases homozygous progeny identified.

T1-populations derived from those events were screened at RWTH for resistance against *M. oryzae* and at IPK as well as TUM for resistance against *B. graminis*. While the disease phenotype of T1-plants belonging to construct 5 and 6 did not display a clear result, construct 1, 3 and 4 appeared to have a resistance promoting effect against *M. oryzae* (Figure 5). Promising individuals were brought to the T2-generation. Putative homozygous plants were selected based on a Hygromycin-assay and re-screened for *M. oryzae* resistance. The results confirmed the enhanced resistance phenotype for construct 4. IPK and TUM evaluated powdery mildew phenotypes of T1 progeny of transgenic plants carrying TriRNAi1 to 5. Phenotyping all events in repeated experiments revealed reproducibly reduced susceptibility of plants carrying constructs TriRNAi2 and TriRNAi4. This was seen in individual events as well as in superpools of all events carrying the same construct. The achieved QR of the transgenic lines could be reproduced in T2 and T3 generations albeit to a lower extent than in the T1 generation.

BARLEY-FORTRESS Targeted exploitation of basal defense genes for pathogen resistance in barley

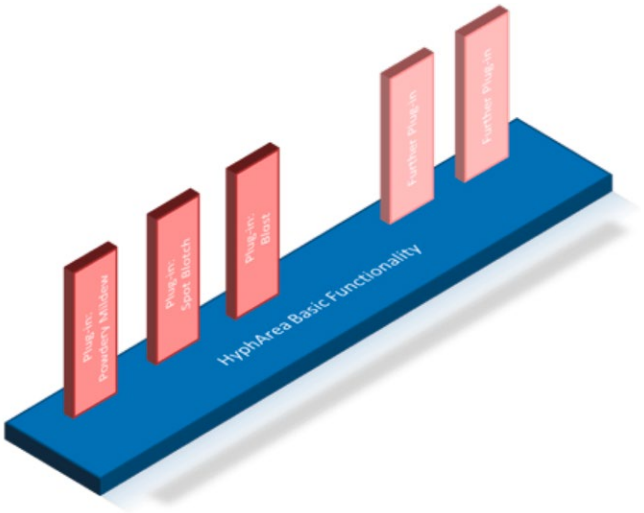


Fig. 6: HyphArea Plug-In System. Built on the HyphArea basic functionality different pathogen specific or even host-pathogen-system specific plug-in modules can be used.

Cell-wall modification and resistance

Cell wall-associated defense is key to QR of *Triticeae* to powdery mildew (Chowdhury *et al.*, 2014; Hueckelhoven, 2014; Chowdhury *et al.*, 2016; Douchkov *et al.*, 2016). TUM hence evaluated individual lines carrying TriRNAi4 for cell wall-associated defense. This suggested that TriRNAi4 supports local callose deposition in papillae because individual less susceptible lines showed higher frequencies of non-penetrated cell wall appositions that accumulated callose than corresponding azygous sister lines. Lines carrying the same construct were also evaluated at RWTH by microscopy, and this revealed a stronger penetration resistance as compared to control plants. The results obtained in both fungal interactions indicated thus a higher defense predisposition at the cell wall by silencing susceptibility-related signaling components. As part of the joint funding program with the ARC Centre of Excellence in Plant Cell Walls, Australian partners UoA as well as University of Melbourne carried out cell-wall analysis of transgenic RNAi lines silenced in *HvCslD2*. This revealed reduced cellulose accumulation in papillae of transgenic plants upon *B. graminis* attack, as well as generally reduced cellulose content in cell-walls of the epidermal layer suggesting *HvCslD2* as novel cellulose synthase (Douchkov *et al.*, 2016).

A novel tool for disease assessment

The project partner Fraunhofer IFF was responsible for developing a software tool to automatically detect hyphal colonies on leaf surfaces by means of light microscopy-based imaging and subsequent estimation of geometrical properties of detected colonies. This will allow to quantitatively assessing intermediate resistance responses occurring after initial fungal penetration but before macroscopic symptom development. For this purpose the concept and functionality of the previously developed tool HyphArea (BMBF forerunner project GABI-phenome) was extended. The major challenges in this context were to implement a modular system capable to deal with an increasing number of pathogens within the BARLEY-FORTRESS project and to fulfil the demand for high-throughput processing. The central approach to achieve these goals was implementing a plug-in system (Fig. 6) with task-specific functionality to accommodate different processing steps and/or parameters for each image operator and the individual pathogens, respectively. The availability of a successively growing library of plug-in modules enables the user to build a task-specific sequence of processing steps. Each individual module was optimised according to its processing speed to achieve a high overall performance.

In order to provide the scientific community with the current version of HyphArea it is available for download under Gnu General Public License. This also enables scientists who were not direct project partners to contribute to its further development and extension.

Summary and Outlook

In the project we were successful in providing proof of concept that candidate genes affecting the interaction of barley with the powdery mildew fungus, which had been previously identified in a number of projects including GABI-nonhost, PRO-GABI, and GABI-phenome, can provide protection by allele introgression as well as by direct gene transfer. Overall, the genes (alleles) encoding five cell-wall modifying proteins, the identity of which remains currently confidential to protect intellectual property, plus of the cellulose synthase-like gene *HvCslD2*, turned out to provide multiple fungal disease resistance. However, especially in the transgenic approach, the moderate degree of protection reached would not justify to enter directly a pipeline of translational research. Instead, the now functionally validated, pathogen-induced TriOEX and TriRNAi cassettes can be used for testing additional candidate gene (combinations) to combat pathogens that enter the plant via the epidermis (and not via stomata). The multiple resistant KWS BAMBINA lines from back-crossing provide a starting point for further exotic genome elimination and use in pre-breeding. Importantly, fine mapping will have to confirm if the positive marker effect was indeed caused by the gene itself or by one or several neighbouring, linked gene(s). We could provide proof of concept for the novel approach of targeted reduction of host susceptibility via candidate-gene silencing, at least for the cassette TriRNAi4. This also provides a starting point for site-directed mutagenesis of the validated genes by new tools such as RNA-guided Cas9. Besides the successful testing of options for trait improvement in barley, the project provided information about mechanisms of cell-wall based defense in barley, which is triggered by conserved non-self molecules and expected to be more durable than the recognition of pathogen effectors triggering programmed cell death in plants.

Literature

- *Bahrini I, Ogawa T, Kobayashi F, Kawahigashi H, Handa H (2011) Overexpression of the pathogen-inducible wheat TaWRKY45 gene confers disease resistance to multiple fungi in transgenic wheat plants. Breeding Science 61: 319-326*
- *Bhuiyan NH, Selvaraj G, Wei YD, King J (2009) Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defence in wheat against powdery mildew invasion. Journal of Experimental Botany 60: 509-521*
- *Chen T, Xiao J, Xu J, Wan W, Qin B, Cao A, Chen W, Xing L, Du C, Gao X, Zhang S, Zhang R, Shen W, Wang H, Wang X (2016) Two members of TaRLK family confer powdery mildew resistance in common wheat. Bmc Plant Biology 16*
- *Chowdhury J, Henderson M, Schweizer P, Burton RA, Fincher GB, Little A (2014) Differential accumulation of callose, arabinoxylan and cellulose in nonpenetrated versus penetrated papillae on leaves of barley infected with Blumeria graminis f. sp hordei. New Phytologist 204: 650-660*
- *Chowdhury J, Schober MS, Shirley NJ, Singh RR, Jacobs AK, Douchkov D, Schweizer P, Fincher GB, Burton RA, Little A (2016) Down-regulation of the glucan synthase-like 6 gene (HvGsl6) in barley leads to decreased callose accumulation and increased cell wall penetration by Blumeria graminis f. sp. hordei. New Phytologist: n/a-n/a*
- *Douchkov D, Lück S, Johrde A, Nowara D, Himmelbach A, Rajaraman J, Stein N, Sharma R, Kilian B, Schweizer P (2014) Discovery of genes for affecting resistance of barley to adapted and non-adapted powdery mildew fungi. Genome Biology 15: 518*
- *Douchkov D, Lueck S, Hensel G, Kumlehn J, Rajaraman J, Johrde A, Doblin MS, Beahan CT, Kopischke M, Fuchs R, Lipka V, Niks RE, Bulone V, Chowdhury J, Little A, Burton RA, Bacic A, Fincher GB, Schweizer P (2016) The barley (Hordeum vulgare) cellulose synthase-like D2 gene (HvCslD2) mediates penetration resistance to host-adapted and non-host isolates of the powdery mildew fungus. New Phytologist: n/a-n/a*
- *Ferrari S, Sella L, Janni M, De Lorenzo G, Favaron F, D'Ovidio R (2012) Transgenic expression of polygalacturonase-inhibiting proteins in Arabidopsis and wheat increases resistance to the flower pathogen Fusarium graminearum. Plant Biology 14: 31-38*
- *Hemetsberger C, Mueller AN, Matei A, Herrberger C, Hensel G, Kumlehn J, Mishra B, Sharma R, Thines M, Hückelhoven R, Doehlemann G (2015) The fungal core effector Pep1 is conserved across smuts of dicots and monocots. New Phytologist 206: 1116-1126*
- *Himmelbach A, Liu L, Zierold U, Altschmied L, Maucher H, Beier F, Muller D, Hensel G, Heise A, Schutzendubel A, Kumlehn J, Schweizer P (2010) Promoters of the Barley Germin-Like GER4 Gene Cluster Enable Strong Transgene Expression in Response to Pathogen Attack. Plant Cell 22: 937-952*
- *Hu PS, Meng Y, Wise RP (2009) Functional Contribution of Chorismate Synthase, Anthranilate Synthase, and Chorismate Mutase to Penetration Resistance in Barley-Powdery Mildew Interactions. Molecular Plant-Microbe Interactions 22: 311-320*
- *Huckelhoven R (2007) Cell wall - Associated mechanisms of disease resistance and susceptibility. Annual Review of Phytopathology 45: 101-127*
- *Huckelhoven R, Eichmann R, Weis C, Hoefle C, Proels RK (2013) Genetic loss of susceptibility: a costly route to disease resistance? Plant Pathology 62: 56-62*
- *Hueckelhoven R (2014) The effective papilla hypothesis. New Phytologist 204: 438-440*
- *Oerke EC (2006) Crop losses to pests. Journal of Agricultural Science 144: 31-43*
- *Pathuri IP, Reitberger IE, Huckelhoven R, Proels RK (2011) Alcohol dehydrogenase 1 of barley modulates susceptibility to the parasitic fungus Blumeria graminis f.sp hordei. Journal of Experimental Botany 62: 3449-3457*
- *Schmalenbach I, March T, Bringezu T, Waugh R, Pillen K (2011) High-resolution genotyping of wild barley introgression lines and fine-mapping of the threshability-locus thresh-1 using the illumina GoldenGate assay. Genes Genomes Genetics 1: 1-10*
- *Schweizer P (2008) Tissue-specific expression of a defence-related peroxidase in transgenic wheat potentiates cell death in pathogen-attacked leaf epidermis. Molecular Plant Pathology 9: 45-57*
- *Schweizer P (2014) Host and Nonhost Response to Attack by Fungal Pathogens. In J Kumlehn, S N., eds, Biotechnological Approaches to Barley Improvement. Springer, Berlin, Heidelberg, pp 197-235*
- *Schweizer P, Stein N (2011) Large-Scale Data Integration Reveals Colocalization of Gene Functional Groups with Meta-QTL for Multiple Disease Resistance in Barley. Molecular Plant-Microbe Interactions 24: 1492-1501*
- *Shin SY, Mackintosh CA, Lewis J, Heinen SJ, Radmer L, Dill-Macky R, Baldridge GD, Zeyen RJ, Muehlbauer GJ (2008) Transgenic wheat expressing a barley class II chitinase gene has enhanced resistance against Fusarium graminearum. Journal of Experimental Botany 59: 2371-2378*
- *Zellerhoff N, Jarosch B, Groenewald JZ, Crous PW, Schaffrath U (2006) Nonhost resistance of barley is successfully manifested against Magnaporthe grisea and a closely related Pennisetum-infecting lineage but is overcome by Magnaporthe oryzae. Molecular Plant-Microbe Interactions 19: 1014-1022*

CEREAL-ROOTS Nutzung von Genen zur Erhöhung der Resistenz und Toleranz der Getreidewurzel gegenüber biotischem und abiotischem Stress

Herr Prof. Dr. rer. nat. Kogel

Justus-Liebig-Universität Gießen, 35390 Gießen

Eingehende Darstellung

Aufgrund der von uns identifizierten Funktion der Gersten-MORC ATPasen in der Pflanzenimmunität und der gleichzeitigen Beteiligung an epigenetischen Prozessen, der RNA-vermittelten DNA Methylierung (RdDM), sind diese Proteine Kandidaten für die Erstellung von breiten Resistenzen (Kang *et al.* 2012; Langen *et al.* 2014; Koch and Kogel 2014; Manosalva *et al.* 2015). Dieser im Verlaufe der Arbeiten zu CEREAL ROOT fortentwickelten Hypothese folgend, haben sich die Arbeiten der AG Kogel gänzlich auf die Bearbeitung von MORC ATPasen beschränkt.

Die ursprünglich CRT1 (Compromised in Recognition of Turnip Crinkle Virus 1) genannte Proteinfamilie wurde aufgrund der Sequenzähnlichkeit mit der MORC (*microrchidia*) Untergruppe der GHKL (für **G**yrase, **H**SP90, Histidin **K**inase und Mut**L**) -ATPasen umbenannt. AtCRT1 aus *Arabidopsis thaliana* sowie seine 6 Homologen AtCRH 1-6 (CRT1 homologe 1 to 6) werden nun als AtMORC 1-7 bezeichnet. Die Gerstenhomologe wurden entsprechend ihrer Ähnlichkeit zu den Vertretern aus Arabidopsis in HvMORC1, 2, 6a, 6b und 7 umbenannt (Langen *et al.* 2014).

Wir konnten zeigen, dass durch RNAi-vermitteltes *knock-down* von *HvMORC1* oder *HvMORC2* in der Gerstensorte Golden Promise (GP) eine erhöhte Resistenz gegenüber dem Erreger des Gerstenmehltaus *Blumeria graminis* f sp. *hordei* (Bgh, Isolat A6) erreicht wird. Die Anzahl der durch *Bgh*A6 gebildeten Pusteln pro Blattfläche war in den *knock-down* Linien verglichen mit Kontroll-Pflanzen (GP Wildtyp oder mit leerem Vektor transformiertem GP (empty vector, ev) um bis zu 53% verringert. Dieses Ergebnis stimmt überein mit einem verringerten mRNA Level (Verringerung um bis zu 45% gegenüber Kontrolle).

Die Beobachtung, dass ein *knock-down* von MORC1 oder MORC2 in Gerste zu einer erhöhten Resistenz führt, steht im Gegensatz zu den Ergebnissen in *Arabidopsis*, wo ein *knock-out* der Gene zu einer erhöhten Anfälligkeit der Pflanzen gegenüber verschiedenen Pathogenen führt. Deshalb wurden die Unterschiede zwischen *Arabidopsis* und Gerste mittels transienter Komplementation (konstitutiven Überexpression von *AtMORC1* oder *HvMORC1* in *HvMORC1*-RNAi-Pflanzen) näher untersucht. HvMORC1-RNAi Pflanzen zeigten eine um 40% verminderte Anfälligkeit gegenüber *Bgh*A6. In Abkömmlingen der *HvMORC1*-RNAi Pflanzen (T2-Generation) wurden mittels particle bombardment Arabidopsis

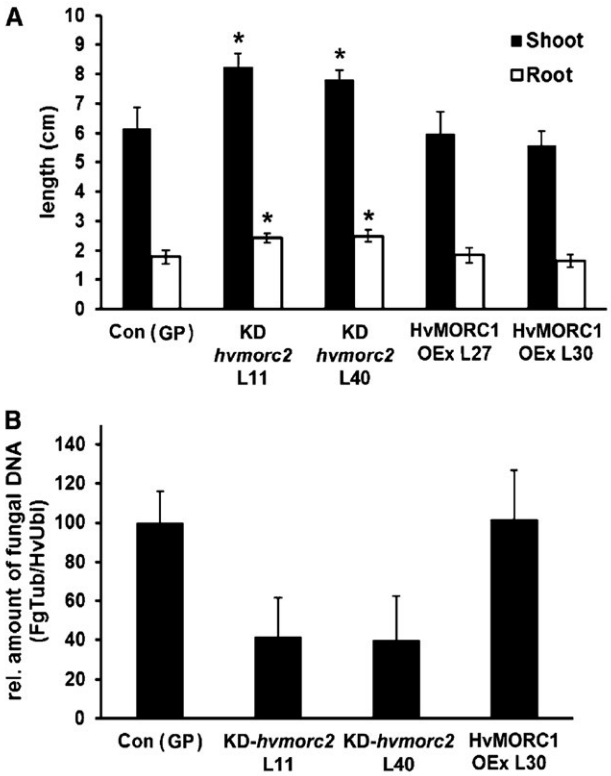


*Abb. 1: Altered expression of HvMORC1 and HvMORC2 affects basal resistance against root rot disease caused by Fusarium graminearum (Fg). A, Shoot and root lengths of infected seedlings. B, Quantification of Fg in the roots by quantitative RT-PCR based on the ratio of fungal tubulin (FgTub) to plant ubiquitin (HvUbi). Shown are phenotypic effects of fungal infections of 10-d-old cv Golden Promise (GP) seedlings (empty vector control), the knockdown lines L11 and L40 (silenced for HvMORC2), and HvMORC1 overexpressor (OEx) lines L27 and L30. For inoculation, roots of 2-d-old seedlings were dipped into a solution of 50,000 Fg macrospores. Presented are mean 6 SE of 30 seedlings from three biological repetitions. Significant changes are marked: *P > 0.05 (Student's t test).*

CEREAL-ROOTS Nutzung von Genen zur Erhöhung der Resistenz und Toleranz der Getreidewurzel gegenüber biotischem und abiotischem Stress

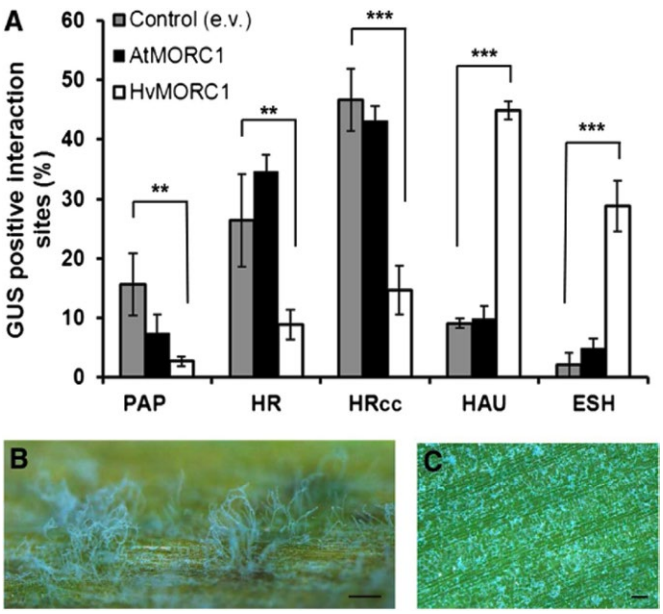


*Abb. 2: Die Überexpression von HvMORC1 in Gerste bricht die MLA12-vermittelte ETI gegen BghA6. A: Interaktion des Pilzes mit transient transformierten Epidermiszellen klassifiziert nach Ausbildung von Papillen (PAP), Hypersensitiver Reaktion (HR), HR in benachbarten Zellen (HRcc), Haustorien (HAU) und verlängerten Sekundärhyphen (ESH). B und C: Mehltaukolonien 6 Tage nach Inokulation auf Gerste cv. Sultan5 (HvMORC1 transient überexprimiert) (B) bzw. der Leervektor-Kontrolle (ev) (C). (** p < 0,001; *** p < 0,0001; Student's t-test); aus Langen et al. 2014).*

bzw. Gersten MORC (*AtMORC1* und *HvMORC1*) transient überexprimiert, danach mit *Bgh*A6 Konidien inokuliert, und hinsichtlich der Penetrationsrate mikroskopisch ausgewertet. In Übereinstimmung mit unseren früheren Daten (Kang *et al.* 2012) zeigte sich, dass die bereits erhöhte Resistenz gegenüber Gerstenmehltau durch Überexpression von *AtMORC1* nochmals um ca. 25% erhöht werden konnte, wohingegen eine Überexpression des *MORC1* Homologs aus Gerste (*HvMORC1*) eine Verminderung der Resistenz bewirkte. Dies zeigte, dass das Protein selbst und nicht die Proteinumgebung für die entgegengesetzte Funktion von MORC1 in *Arabidopsis thaliana* und *Hordeum vulgare* verantwortlich ist.

HvMORC1 Protein wurde in *E. coli* rekombinant hergestellt, aufgereinigt und biochemisch charakterisiert. Es konnte gezeigt werden, dass HvMORC1 eine Cofaktor-abhängige Endonuclease ist. Als Cofaktoren können Mn^{2+}, Co^{2+} und in geringerem Maße auch Ni^{2+} und Mg^{2+} fungieren. In Anwesenheit von ATP war eine langsamere Umsetzung von *supercoiled* in relaxierte und nachfolgend in linearisierte DNA zu beobachten. Dieser Effekt wurde im verstärkten Maß auch bei Verwendung des nicht-hydrolysierbaren ATP-Analogon AMPPnP gefunden. Durch Verwendung des spezifischen GHKL-ATPase-Inhibitors Radicicol konnte gezeigt werden, dass die inhibierende Wirkung von ATP tatsächlich auf der Bindung des Nukleotids an die Interaktionsstelle der ATPase-Domäne beruht.

Stabil transgene Gerstenlinien für Überexpression und *knock-down* von *HvMORC1* oder *HvMORC2* wurden auch hinsichtlich ihres Resistenzverhaltens gegenüber dem Getreidepathogen Fusarium graminearum untersucht. *Knock-down* von *HvMORC2* resultierte in einer verminderten Infektion der Gerste durch *F. graminearum* (Abb. 1). Durch relative Quantifizierung von *Fg*-DNA zu *Hv*-DNA mittels quantitativer RT-PCR konnte eine verminderte Kolonisierungsrate des Pilzes auf den Pflanzen bestätigt werden.

Ein verändertes Expressionsniveau von MORC Genen beeinflusst die Effektor-vermittelte Resistenz (ETI)

Es wurde untersucht, ob Gerste MORC's in der Effektor-vermittelten Resistenz (ETI) eine Rolle spielen. AtMORC1 wurde in diesen Experimenten ebenfalls mit überprüft, um die gegenteiligen Effekte der MORC Proteine aus Gerste und Arabidopsis näher zu beleuchten. Zellen der Blattepidermis von Gerste cv. Sultan5 wurden mittels particle bombardment transient transformiert. Mit Hilfe dieser Methode konnten die langen Generationszeiten stabil transformierter Gerste umgangen werden. Sultan5 besitzt das *MLA12*-Gen und ist damit resistent gegenüber *Bgh*A6, das das korrespondierende *AvrMLA12*-Gen besitzt. Die Interaktionen von *Bgh*A6 mit den transient überexprimierenden Zellen wurden auf Einzelzellebene ausgewertet (Abb. 2).

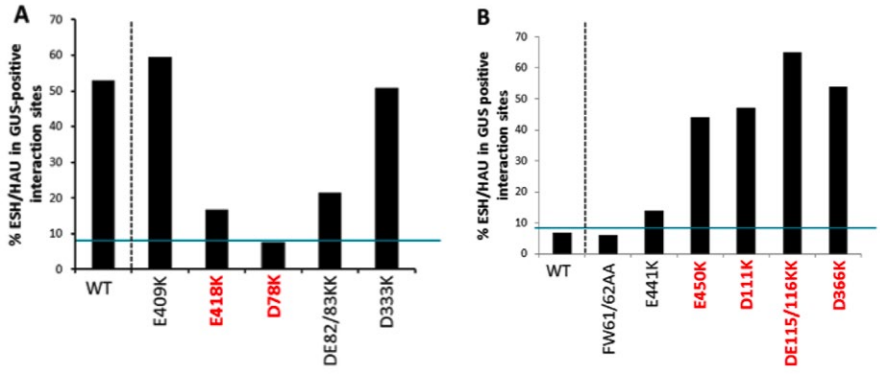


Abb. 3: Einfluss von Punktmutationen auf die MLA12-vermittelte Resistenz gegen BghA6. Gerste cv. Sultan5 wurde transient mit p35S-Überexpressionskonstrukten verschiedener Punktmutanten von HvMORC1 (A) und AtMORC1 (B) bzw. dem jeweiligen Wildtyp (WT) sowie einem GUS-Reporter-Konstrukt transformiert und mit BghA6 inokuliert. Quantifiziert wurde nach Ausbildung von ESH und/oder Haustorien in transformierten Zellen.

CEREAL-ROOTS Nutzung von Genen zur Erhöhung der Resistenz und Toleranz der Getreidewurzel gegenüber biotischem und abiotischem Stress

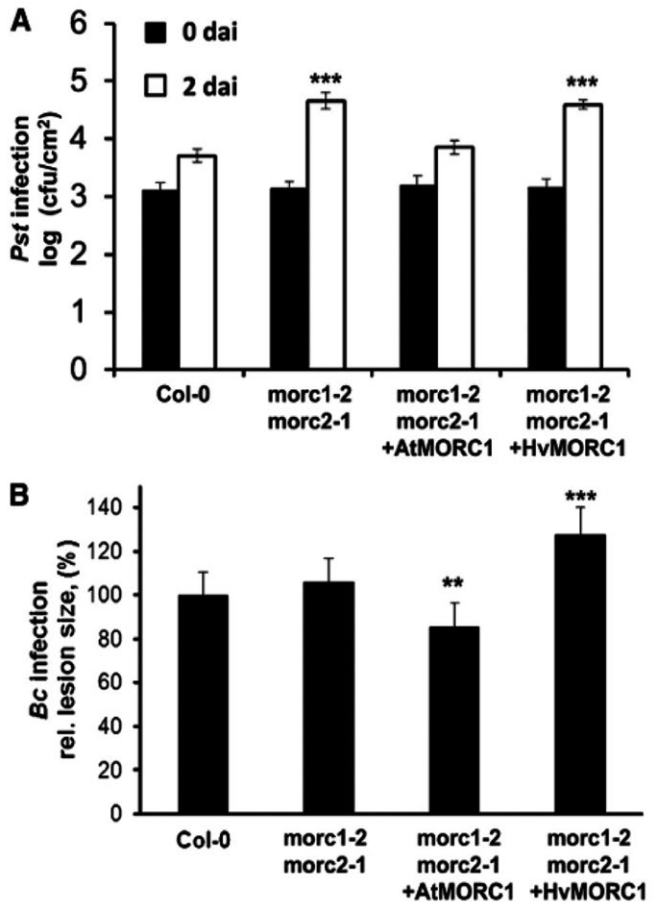


*Abb. 4: Untersuchungen zur funktionellen Komplementierung von AtMORC1 und HvMORC1 in Arabidopsis. (A) Blätter wurden mit Pst AvrRpt2 (105 cfu/ml) mittels einer Spritze infiltriert und das Bakterienwachstum 0 und 2 Tage nach Infiltration (dpi) bestimmt. (B) Blätter wurden mit 10 µl einer Sporensuspension (50.000 Konidien/ml) von Botrytis cinerea versetzt und Symptome nach 3 Tagen ausgewertet. (** p < 0,01; *** p < 0,001; Student's t-test).*

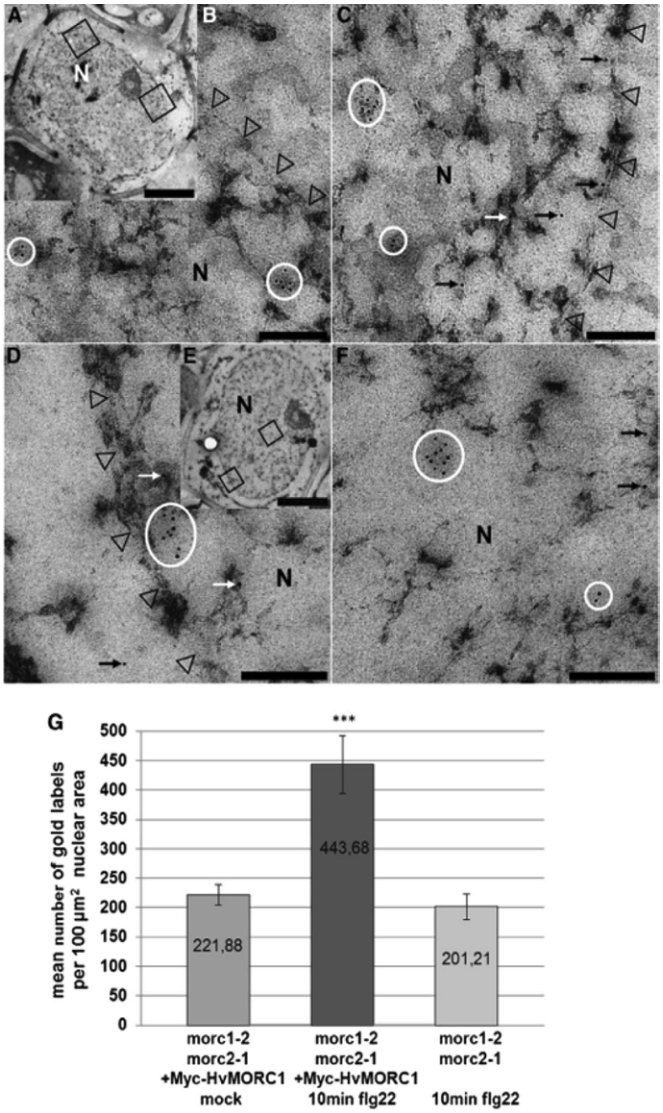


*Abb. 5: Detektion von Myc-HvMORC1 in Arabidopsis mittels TEM nach Immunogold-Markierung. A (mit B und C in höherer Vergrößerung) und E (mit D und F in höherer Vergrößerung): Nuclei von Phloemparenchym-Zellen 10 min nach PTI-Induktion mittels flg22. G: Durchschnittliche Anzahl von Gold-Markierungen je 100 µm² Kernfläche (*** p < 0,0005, Student's t-test).*

Es zeigte sich, dass die konstitutiv *HvMorc1* (p35S::HvMORC1) überexprimierenden Zellen einen Bruch der MLA12-vermittelten Resistenz gegen *BghA6* aufweisen. Die Ausbildung einer hypersensitiven Immunantwort (HR) war im Vergleich zur Leervektor-Kontrolle (GP ev) signifikant reduziert, während der Pilz häufiger Haustorien und Myzelien ausbilden konnte. Die Überexpression von *AtMORC1* in Sultan5 (homozygot für *MLA-12*) dagegen verstärkte die ETI. Die Unterschiede zur Leervektor-Kontrolle waren hier jedoch nicht signifikant.

Diese Untersuchung wurden nachfolgend auf weitere Vertreter der MORC-Familie aus Gerste ausgeweitet, um mögliche Unterschiede zwischen den Untergruppen aufzudecken: HvMORC1 und 2 (Untergruppe I), HvMORC7 (Untergruppe II) und HvMORC6a (Untergruppe III) (siehe Abb. 1). Alle untersuchten MORC Proteine von Gerste führten bei Überexpression zu einem Bruch der ETI, wobei die beiden Mitglieder der Untergruppe I den stärksten Effekt hatten.

Identifizierung und Generierung von Funktionsverlust-Mutanten von HvMORC1 und AtMORC1 durch Einführung von Punktmutationen

MORC Proteine besitzen zwei unterschiedliche Enzymaktivitäten (ATPase und Endonuklease); somit ergab sich die Frage nach dem Einfluss dieser Enzymfunktionen auf die Pathogenresistenz. Zur Identifizierung von für die Aktivität kritischen Aminosäuren wurden Sequenzvergleiche mit verwandten, sehr gut charakterisierten MutL-Proteinen aus verschiedenen Prokaryoten durchgeführt, anhand derer sich putative Funktionsverlust-Mutanten ergaben. Entsprechende Konstrukte wurden generiert und im Rahmen eines

transienten Tests im Gerstenkultivar Sultan5 hinsichtlich ihres Einflusses auf die Resistenz gegenüber *Bgh*A6 im Vergleich zum Wildtyp Protein überprüft (Abb. 3). Für MORC1 aus beiden Pflanzenspezies konnten Punktmutanten identifiziert werden, die die ATPase-Aktivität (HvMORC1 D78K, AtMORC1 D111K, AtMORC1 DE115/116KK) bzw. die Endonuklease-Aktivität (HvMORC1 E418K, AtMORC1 E450K) ausschalten, und die den jeweiligen Effekt des WT Proteins negieren. Nachweise der Proteinexpression und -stabilität *in planta* wurden bereits durchgeführt, wohingegen der Nachweis des enzymatischen Funktionsverlustes *in vitro* derzeit noch aussteht.

Untersuchung zur funktionellen Komplementierung von HvMORC1 und AtMORC1 in Arabidopsis.

Weiterführende Untersuchungen zu den funktionellen Unterschieden von AtMORC's und HvMORC's fanden in komplementären Experimenten statt. Dazu wurden Arabidopsis *morc1/morc2* Knock-out Mutanten (dKO) mit einem Überexpressionskonstrukt für *HvMORC1* bzw. *AtMORC1* transformiert und hinsichtlich ihres Resistenzverhaltens gegen *Pseudomonas syringae* pv. *tomato* (*Pst*) *AvrRpt2* (hemibiotroph) oder *Botrytis cinerea* (Bc, necrotroph) untersucht (Abb. 4). Die Überexpression von *AtMORC1* konnte die R Gen-vermittelte Resistenz gegenüber Pst bis hin zum Niveau des Wildtyps wiederherstellen, während Pflanzen, die *HvMORC1* überexprimieren, ebenso anfällig blieben wie die Arabidopsis dKO Mutante. Im Fall von *B. cinerea* konnte ein vergleichbares Verhalten nachgewiesen werden. AtMORC1 konnte die Resistenz gegenüber dem Pilz leicht verstärken, während die Überexpression von *HvMORC1* die Anfälligkeit sogar noch erhöhte.

Lokalisierung und Translokalisierung von HvMORC1 in Arabidopsis nach Induktion der Immunantwort mittels flg22

Die entgegengesetzte Wirkung von HvMORC1 und AtMORC1 könnte auf eine unterschiedliche subzelluläre Lokalisierung der Proteine erklärt werden. Für AtMORC1 konnte bereits nachgewiesen werden, dass die Induktion der Immunantwort durch flg22 zu einer Translokalisierung des Proteins vom Zytoplasma in den Nukleus führt (Kang *et al.* 2012).

Gleichermaßen konnten wir mittels Transmissionselektronenmikroskopie zeigen, dass auch HvMORC1, nach ektopischer Expression in Arabidopsis und Infiltration der Blätter mit flg22, in den Kern transloziert wird (Abb. 5).

Referenzen

- *Kang HG, Choi HW, von Einem S, Manosalva P, Ehlers K, Liu PP, Buxa SV, Moreau M, Mang HG, Kachroo P, Kogel KH, Klessig DF (2012) CRT1 is a nuclear-translocated MORC endonuclease that participates in multiple levels of plant immunity. Nature Comm 3: 1297.*
- *Langen G, von Einem S, Koch A, Imani J, Pai S, Manohar M, Ehlers K, Choi HW, Claar M, Schmidt R, Mang HG, Bordiya Y, Kang HG, Klessig DF, Kogel KH (2014) The Compromised Recognition of Turnip Crinkle Virus1 subfamily of Microrchidia ATPases Regulates Disease Resistance in Barley to Biotrophic and Necrotrophic Pathogens. Plant Physiology 164: 866-878.*
- *Koch A, Kogel KH (2014) New wind in the sails: improving the agronomic value of crop plants through RNAi-mediated gene silencing. Plant Biotechnology Journal doi: 10.1111/pbi.12226.*
- *Patricia Manosalva, Murli Manohar, Karl-Heinz Kogel, Hong-Gu Kang, and Daniel F. Klessig (2015) The GHKL ATPase MORC1 modulates species-specific plant immunity in Solanaceae. MPMI 10.1094/MPMI-12-14-0401-R.*

CLIMATE CHANGE Association genetic studies for combined heat and drought tolerance in barley

Sven Templer[1,2,*], Lars Voll[3], Anja Hanemann[4], Jutta Förster[5], Maria von Korff[2,6], Frank Ordon[1]

1 Julius Kühn-Institute (JKI), Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, D-06484 Quedlinburg, Germany
2 Max Planck Institute for Breeding Research, Carl-von-Linné-Weg 10, D-50829 Köln, German
3 Friedrich-Alexander-Universität Erlangen-Nürnberg, Division of Biochemistry, Staudtstr. 5, D-91058 Erlangen, Germany
4 Saatzucht Josef Breun GmbH & Co. KG, Amselweg 1, D-91074 Herzogenaurach, Germany
5 Saaten-Union Biotec GmbH, Hovedisser Strasse 92, D-33818 Leopoldshöhe, Germany
6 Cluster of Excellence on Plant Sciences, Heinrich-Heine-Universität Düsseldorf, Institute for Plant Genetics, Universitätsstrasse 1, D-40225 Düsseldorf, Germany
* current address: Max Planck Institute for Biology of Ageing, Joseph-Stelzmann-Str. 9b, D-50931 Köln, Germany

Introduction

Heat and drought severely limit crop productivity worldwide and often prevail in combination. Prolonged heat and/or drought periods are not limited to arid and semi-arid regions, but also increasingly affect European agriculture in the course of global climate change, that is, for Western/Central Europe, associated with a progressive redistribution of precipitation to winter months and increased temperature in spring and summer.

In breeding, barley varieties adapted to various climates have been selected, including marginal environments e.g. semi-arid regions in the Mediterranean, while barley breeding in temperate regions like Western/Central Europe has been conducted under favorable conditions focusing on grain yield and quality traits instead of tolerance to abiotic stress.

To develop breeding strategies for future climatic conditions, a deep genetic, phenotypic and metabolic understanding of adaptation strategies of barley to drought and heat conditions is required. Therefore, the aim of CLIMATE CHANGE was to generate large datasets including phenotypic, genetic and metabolic data from 451 Western/Central European spring barley breeding lines and 92 Mediterranean landraces and cultivars derived from semi-arid climates. The lines were grown for three consecutive years in field trials at three locations in Germany, two of which were naturally drought prone and one location with higher annual precipitation. During the field trials extensive phenotypic data were collected and the genetic variation of all tested lines was determined using the 9K barley iSelect SNP array. These data were later used to perform a genome wide association study in order to detect genomic regions which are involved in the adaptation to drought stress and which could ultimately be used for marker assisted selection for improving drought tolerance of new breeding lines.

Furthermore, different climate scenarios were calculated which were then used as a basis for running climate chamber experiments with defined drought as well as combined heat and drought stress conditions with a core set of 89 genetically divergent spring barley lines. The core set was compiled based on the genetic and phenotypic data of the field trials. During the climate chamber experiments morphological changes, yield and yield components as well as

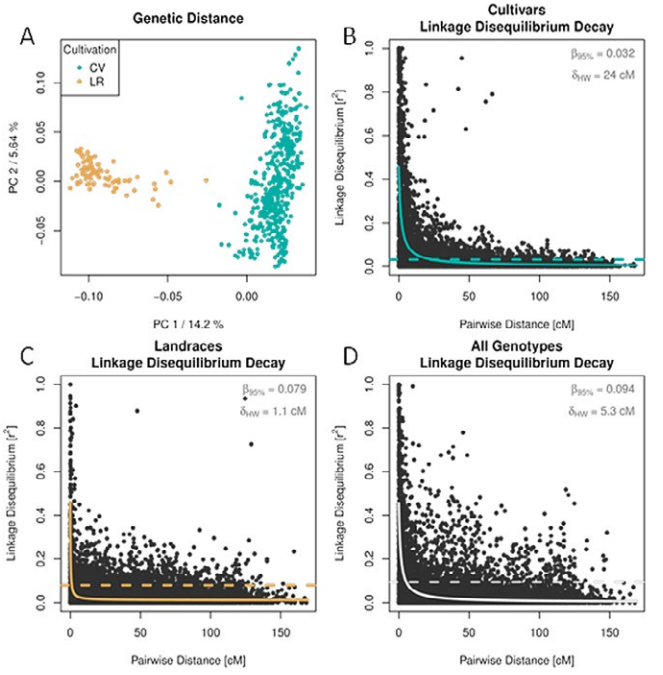


Fig. 1: Principal coordinate analysis of genetic distances in the cultivar (CV) and landrace (LR) germplasm sets (A). Average r2 linkage disequilibrium (LD) decay (δHW) is shown for the CV germplasm set (B), the LR germplasm set (C) and the combined germplasm set (D). Regression (solid) and background LD (dashed) lines in colors according to germplasm set. βHW is the amount of background LD.

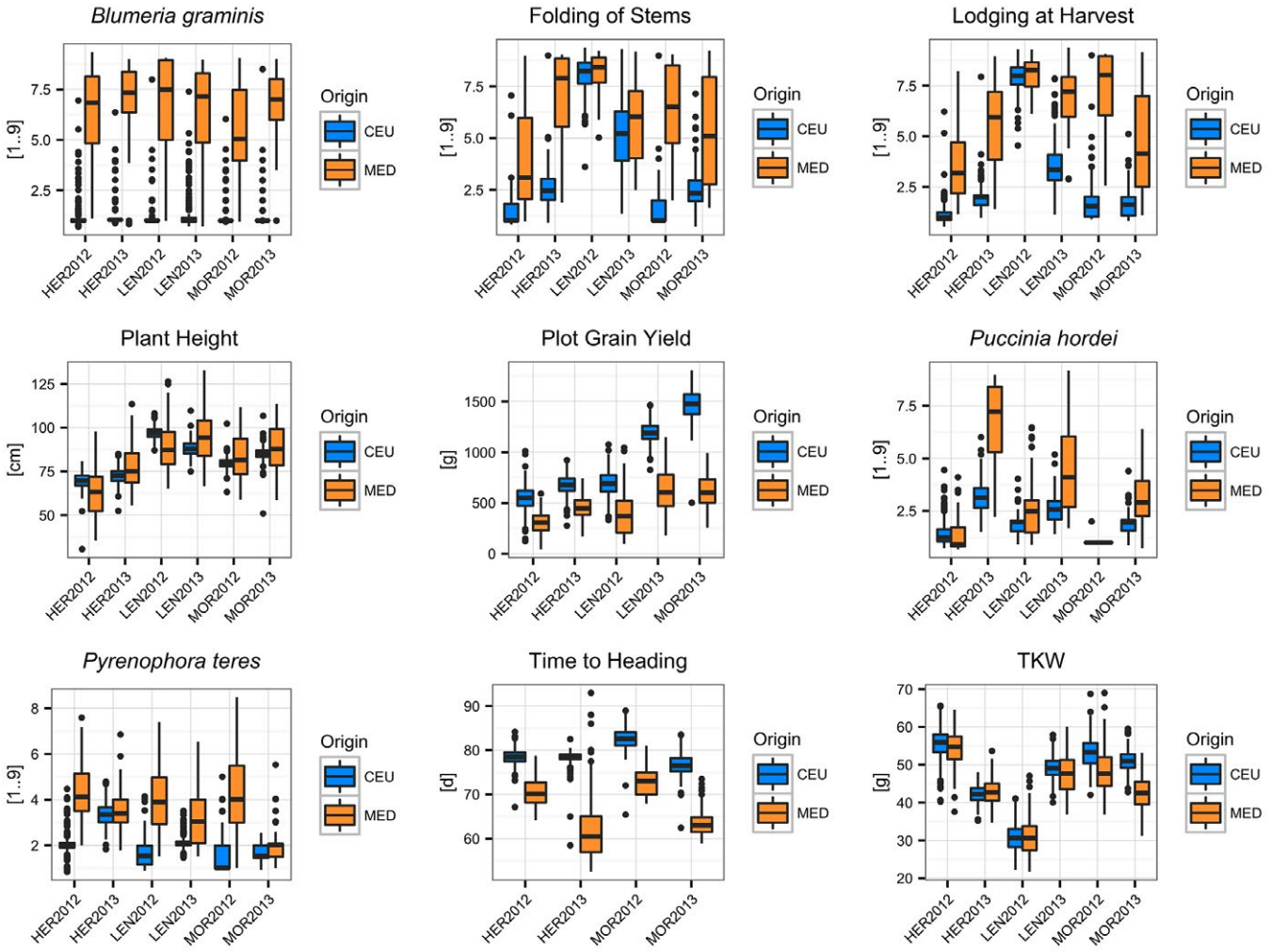


Figure 2: Boxplots showing the distribution of best linear unbiased estimates (BLUEs) in each environment (x-axis) and germplasm set separately. The unit of each trait is given on the respective abscissa label, with the degree of stem folding, lodging and disease symptoms being represented by scores reaching from 1-9. Cultivars are depicted in turquoise, landraces in yellow. TKW: Thousand kernel weight HER: location Herzogenarauch; LEN: location Lenglern; MOR: location Morgenrot.

metabolic changes of primary and secondary metabolic pathways in flag leaves were recorded. Later on, the phenotypic and metabolic data were correlated in a genome wide association study in order to detect markers and metabolites which are associated with drought and/or heat tolerance.

Results

Field Trials

The germplasm set consisting of 451 German spring barley genotypes (CV) and a diverse collection of 92 landrace genotypes (LR) collected from arid and semi-arid regions was analyzed for genetic variation with the barley Illumina iSelect 9k array comprising 7864 genome-wide single nucleotide polymorphisms (SNPs). Population structure was determined by a principle coordinate analysis on modified Roger's distances of all possible pairwise genotype comparisons. Principal coordinate analysis clustered the CV and LR genotypes into genetically distinct groups (Figure 1A). The 451 spring barley breeding lines showed allelic variation at 4455 SNP markers whereas the 92 landraces were polymorphic for 5083 SNP markers. In general, the LR collection contained a higher number of minor alleles and was significantly more diverse than the set of CV lines with 451 genotypes.

In order to investigate linkage disequilibrium (LD), we calculated the squared correlation r2 of marker pairs according to Flint-Garcia et al. (2003). To average the LD extend over the genetic distance, a linear regression was calculated according to Hill and Weir (1988). The rate of LD decay averaged for all intrachromosomal marker pairs was significantly more rapid in LR (1.1 cM) than in CV (24.0 cM) (Figure 1 B, C). Background LD showed a higher value in LR (0.079) than in CV (0.032). In the combined CV and LR sets LD decay was estimated at 5.3 cM (background r2: 0.094) (Figure 1D).

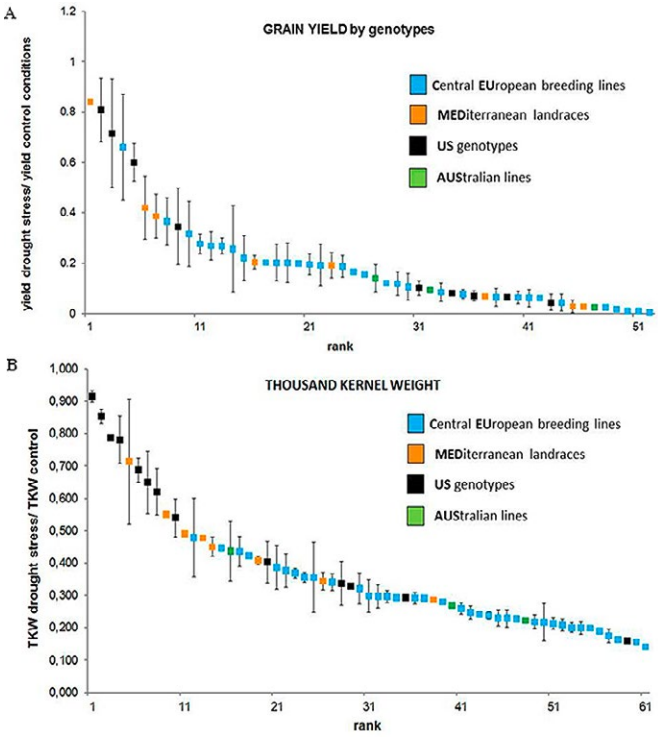


Figure 3 Grain yield (A) and TKW (B) reduction in drought stress relative to control conditions. Grain yield and TKW in drought stress relative to control conditions was calculated for each genotype and results were ordered from highest to lowest. Geographic origin of the individual genotypes is indicated in the figure. Only reliable datasets were included in the plots.

In order to calculate genotypic values for each location-year combination (called environment hereafter) the best linear unbiased estimate (BLUE) for each genotype was calculated. Phenotypic variation among the germplasm sets and environments were visualized by boxplots (Figure 2).

Phenotypic differentiation was stronger between the two germplasm sets than between locations and years. As expected phenotypic variation within and between environments was significantly higher in the landrace collection than in the breeding lines in particular for disease symptoms, stem folding and lodging. Disease symptoms showed correlations between Pyrenophora teres and Puccinia hordei. Furthermore, diesease symptoms for Puccinia hordei were negatively associated with heading time. Lodging and folding of stems was negatively correlated with thousand kernel weight (TKW). In summary, phenotypic performance significantly differed between the CV and LR germplasm sets. Higher genetic variation and higher heritabilities were observed for the LR germplasm as compared to the elite lines. The LR set was characterized by early flowering, lower yield, increased lodging and disease susceptibility compared to the CV genotypes.

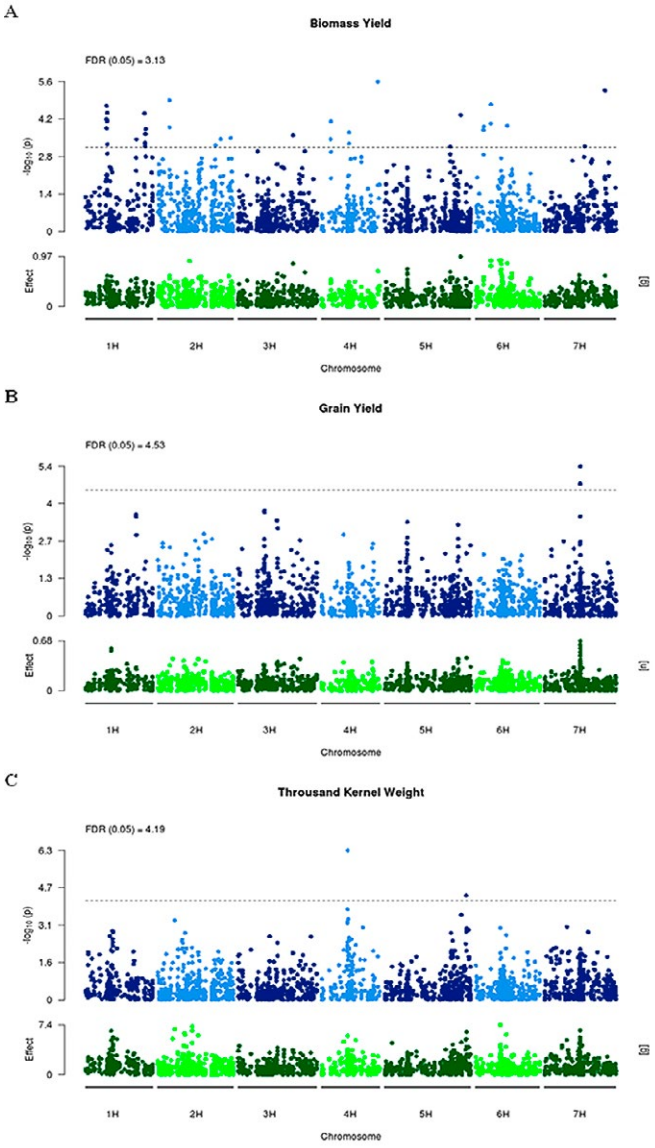


Figure 4 Manhattan plots of the association analysis for the agronomic traits Total biomass (A), Grain yield (B) and TKW (C) from a GWAS over all treatments (as fixed effect). Each dot along the x-axis represents the negative logarithmic p-value (blue) and effect size (green) of a single SNP marker.

Climate Chamber Experiments

With the climate chamber experiments we wanted to identify genetic markers and metabolic patterns which correlate with drought or combined heat and drought responses in a diverse set of barley germplasm. Therefore, a core set of genotypes was selected from the large germplasm set tested under field conditions. The 89 selected barley genotypes consisted of 27 barley landraces and 12 cultivated genotypes from Mediterranean environments and 50 spring barley elite breeding lines from Western/Central Europe.

Three independent repeats of the climate chamber experiments were performed in two years by growing four plants of each genotype per pot under three different conditions

Tab. 1: Gross changes in flag leaf metabolite contents after 5 days in the respective stress conditions relative to controls. These observations were made in all examined genotypes.

Metabolite	Response in drought stress	Response in combined stress
α-Tocopherol	↑	↑ ↑
γ-Tocopherol	↑	↑ ↑
Glutamine	↑ ↑	↑ ↑ ↑
Glycine	↑	↑ ↑
Histidine	↑	↑ ↑
Isoleucine	↑	↑ ↑
Leucine	↑	↑ ↑
Lysine	↑	↑ ↑
Phenylalanine	↑ ↑	↑ ↑ ↑
Proline	↑ ↑ ↑	↑ ↑ ↑ ↑
Tyrosine	↑	↑ ↑
Valine	↑	↑ ↑
Alanine	↓	↑
Aspartate	↓	
Serine	↓ ↓	
Glutamate	↓	
Starch	↓	

(control, drought stress and combined heat and drought stress). Drought stress was applied at the heading stage by a controlled reduction of soil water content (SWC) to 20% using a gravimetric method, while control plants were kept at 70% SWC. For the combined heat and drought treatment, the ambient temperature was changed from 20°C/16°C during day/night under control conditions to 26°C/22°C under heat seven days after the application of drought stress, when the SWC had reached 20%. The stress treatment was maintained until harvest of the plants. During the stress treatment physiological measurements were taken; at days 4 and 10 leaf temperature and quantum yield of photosynthesis were recorded and at 11 days relative leaf water content was determined. At harvest, agronomic traits such as plant height (HGT), biomass of straw (YST), grain yield (YGR), kernel number per ear (KPE) as well as thousand kernel weight (TKW) were recorded.

Furthermore, the flag leaf of all four plants per pot for all three conditions and from one experimental repeat were sampled at day 5 after the onset of stress and pooled for metabolite analysis. The leaf samples were analyzed for 57 metabolites, major antioxidants, soluble sugars, amino acids, carboxylates and phosphorylated intermediates.

The treatments had a strong effect on agronomic and physiological traits in the core set. TKW, grain yield, straw yield, photosynthetic quantum yield, and leaf water content were highest under control conditions, while stress treatment caused a reduction in these trait values (Figure 3). E.g., in drought stress, grain yield of Mediterranean and US genotypes were least affected (Figure 3A), while TKW was most stable in drought stressed US lines (Figure 3B). Our data demonstrate genetic variation in agronomic performance; in average, elite barley lines showed a stronger reduction in grain yield and TKW under drought stress than drought-adapted Mediterranean accessions and US genotypes. This indicates that the adjustment of the source-sink relationship under drought stress conditions in elite germplasm is suboptimal.

The genome wide average LD decay was observed at 3.6 cM in the core collection. Significant marker trait associations were detected for 38 traits across all three treatments. Representative results are shown for the traits biomass yield, grain yield and thousand kernel weight (Figure 4). In this genetically diverse core collection, QTLs for grain yield and TKW were detected on chromosome 7H and on chromosome 4H, respectively (Figure 4).

Treatment effects on leaf metabolites differed between drought alone and the combination of drought and heat. Most metabolites showed the same response in all examined genotypes, but some metabolites behaved differently between drought adapted accessions and elite germplasm. In general, flag leaf contents of tocopherols and most free amino acids exhibited a much stronger increase in combined heat and drought stress compared to drought stress (Table 1), while the major amino acids alanine, aspartate, serine and glutamate as well as starch contents only decreased in drought stress relative to control conditions (Table 1). While the combined drought and heat caused more severe changes in agronomic traits than drought alone, a number of metabolites were specifically altered either by drought alone or the combined stress treatments. In addition, the landrace barley genotypes adapted to dry Mediterranean environments showed improved yield and TKW and different metabolite levels under stress compared to the elite lines, and might thus harbor genetic variation for improving stress responses in elite breeding lines.

Resumen

The project CLIMATE CHANGE helped to improve our understanding of the response to drought stress and combined heat and drought stress in spring barley genotypes adapted to semi-arid growth conditions and breeding lines adapted to a temperate climate. The field experiments served to identify a genetically diverse set of genotypes differing in their drought stress response which later were subjected to extensive climate chamber experiments. Additionally, based on the data collected during the field experiments, regions in the genome which are involved in drought stress response were identified.

The climate chamber experiments provided a large number of phenotypic and metabolic data which allowed us to narrow down on metabolic pathways involved in drought and/

or combined drought and heat stress resistance. Additionally, genome wide association studies revealed genomic regions which are strongly associated with the response to drought stress. These regions are currently under further investigation in order to identify the underlying genes and to develop markers which can be used for marker assisted selection in barley breeding.

Furthermore, the knowledge gained on the drought stress response of the CLIMATE CHANGE core set genotypes led to the successful application for a research project funded by the Bavarian State that will allow further studies to investigate the physiological adaptation strategies of barley to drought stress.

Literature

- *Flint-Garcia, S.A., Thornsberry, J.M., S, E., and IV, B. (2003). Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54, 357–374.*
- *Hill, W.G., and Weir, B.S. (1988). Variances and covariances of squared linkage disequilibria in finite populations. Theoretical Population Biology 33, 54–78.*

FROWHEAT Evaluation of wheat prebreeding germplasm for frost tolerance via a genome wide and candidate gene based analysis approach

Koch Michael[2], Arana-Ceballos Fernando[1], Lohwasser Ulrike[1], Chesnokov Yuriy[3], Pshenichnikova Tatyana[4], Schondelmaier Jörg[5], Babben Steve[6], Perovic Dragan[6], Ordon Frank[6], Börner Andreas[1]

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstraße 3, 06466 Stadt Seeland OT Gatersleben, Germany
2 Deutsche Saatveredelung AG (DSV), Weissenburger Straße 5, 59557 Lippstadt, Germany
3 N.I.Vavilov Research Institute of Plant Industry (VIR), B. Morskaya Street 42-44, 190000 St. Petersburg, Russia
4 Institute of Cytology and Genetics of Siberian Branch of the Russian Academy of Sciences, Lavrentyeva 10, 630090 Novosibirsk, Russia
5 Saaten-Union Biotec GmbH, Hovedisser Str. 92, 33818 Leopoldshoehe, Germany
6 Julius Kühn-Institut (JKI), Federal Research Centre for Cultivated Plants, Institute for Resistance Research and Stress Tolerance, Erwin-Baur-Str. 27, 06484 Quedlinburg, Germany

Abstract
The most important limiting factor for the development of winter wheat is besides water the frost stress. Therefore more than 300 winter wheat varieties were characterized for their reaction to frost with the goal to get an idea about the genetic variability regarding frost tolerance in this material. The phenotyping was realized over different years and locations. Furthermore the testing of this material was done also under controlled conditions in climate chambers. In parallel the genotyping of 276 of these varieties with the 90k Infinium SNP chip was conducted. Based on the SNPs marker trait associations were calculated. Associations on chromosomes 1B and 5A turned out to be stable over several environments. The SNP approach was supplemented by a candidate gene study. For genes involved in frost response genome specific primers were developed and amplicons over the 276 varieties, which were also used for the 90k Infinium SNP chip study, produced and sequenced. In some *CBF* genes were polymorphisms observed, which showed an association to frost tolerance.

Introduction
According to the FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO) is wheat the biggest cereal worldwide with a mean harvested area of 217 million ha. Also in Eastern Europe (primarily continental climate) wheat was the most widely grown cereal with a mean harvested area of 39 million ha (2005-2008) (http://faostat.fao.org 2010). This holds also true for the EU27 and Germany where wheat was grown on 3.178 Mio ha with an average yield of 7.84 t/ha in 2009. High and stable wheat yields are of special importance for feeding the worlds growing population and to meet the demands for renewable resources. Exposure to low temperatures is one of the most important plant abiotic stress factors. Therefore, cold signalling and acclimation in plants is a topic of prime importance in breeding research and applied plant breeding since many decades. Most of this basic research was conducted in the model species *Arabidopsis thaliana*. The molecular network established in response to cold stress consists of different key genes/proteins connected via signalling pathways. The CBF (CRT/DRE Binding factors) pathway is essential in triggering response to cold exposure. In the promoters of many cold-regulated plant genes a cis-acting element with a 5 bp core sequence of CCGAC, designated as C-repeat (CRT), is present in one or more copies, for instance in the *Arabidopsis* genes COR15a (Baker *et al.* 1994) or in the *Brassica napus* gene BN115 (Jiang *et al.* 1996). This element induces the expression of cold-regulated genes and also stimulates response to dehydration stress. These C-repeat containing elements were given therefore also the name Drought responsive element (DRE). Proteins binding to this motif were for the first time identified in Arabidopsis and named CBF (CRT Binding factors) (Stockinger *et al.* 1997). Liu *et al.* (1998) identified further proteins binding to this motif and named these DREB (DRE-Binding protein). Subsequent to their discovery in *Arabidopsis*, many CBF homologues have been found in monocots and dicots. Badawi *et al.* (2007) identified 15 different TaCBF gene groups in hexaploid wheat. In *Triticum monococcum* a cluster of eleven CBF genes was mapped to the frost resistance-2 locus on chromosome 5 using a cross between the frost tolerant accession G3116 and frost sensitive DV92. Closer characterization of the mapped genes revealed that *TmCBF12*, *TmCBF14* and *TmCBF15* had the biggest impact on frost tolerance (Knox *et al.* 2008). Also in *Triticum aestivum* it was shown that deletions

in the CBF gene cluster at the Fr-B2 locus is critical for the frost tolerance (Pearce *et al.* 2013).

In classical mapping studies the loci Fr-A1 (frost resistance 1) (Sutka & Snape 1989) and Fr-A2 (Vágújfalvi *et al.* 2000, 2003) were located on the distal portion of the long arm of chromosome 5A separated by a genetic distance of about 30 cM. Dhillon *et al.* (2010) examined diploid wheat mutants lacking *VRN-1* and concluded that Fr-A1 is probably a pleiotropic effect of *VRN-1*. These loci were also mapped in barley as Fr-H1 and Fr-H2 in syntenic regions to the wheat chromosome 5A (Francia *et al.* 2004). Like in wheat the loci co-localize with *Vrn-* and CBF-genes, respectively (Stockinger *et al.* 2007). Fricano *et al.* (2009) analysed four different CBF genes in a European barley germplasm collection and showed that *HvCbf14* is associated with frost tolerance. Wheat *CBF* gene sequences were used in rye to isolate eleven rye *CBF* genes. Nine of these mapped to chromosome 5RL (Campoli *et al.* 2009). Furthermore polymorphisms in 12 candidate genes with an association to frost tolerance in rye were described by Li *et al.* 2011a and Li *et al.* 2011b. Zhu *et al.* 2014 reported that the variation in the copy number of *VRN-A1* in wheat is important for the frost tolerance.

Genetic studies of agronomic important traits in cereals have revealed that most of them are inherited quantitatively and, therefore, they are difficult to detect within the genome. Using segregation-based mapping methods a huge amount of quantitative trait loci (QTLs) have been mapped during the last decade. Examples for wheat are given by Börner *et al.* (2002), Quarrie *et al.* (2005), Huang *et al.* (2006) and Kumar *et al.* (2007).

Segregation-based approaches directly observe the linkage of marker loci and target genes in one segregating population. Another, more recent methodology for the detection of QTLs is the association-based approach, largely and effectively used in human genetics. Here a larger population of individual genotypes is analysed in order to detect associations between marker patterns and trait expressions. Association mapping is based on the linkage disequilibrium (LD) between different loci caused by the history of mutations and genetic drift in the population. Therefore, high LD is expected to be observed between loci in tight linkage, as recombination events should have eliminated LD between loci that are not in close distance (Breseghello & Sorrels 2006a, b). There are two main advantages of an association-based mapping strategy: [1] A much larger gene pool (number of alleles) is included in the analysis (not only two parents or alleles, respectively). [2] The time-consuming development of segregating populations (recombinant inbred lines, doubled haploid lines, etc.) can be avoided. Because of the huge advantages of this approach, it is now more and more established in plant genetics and breeding. One difficulty in the statistical analysis is the occurrence of 'false-positive' associations between a marker and a trait, which is caused by LD between loci that are not or only in loose linkage due to population structure. For dealing with such a hidden population structure, statistical options were developed (Falush *et al.* 2003). This situation can occur through selection, which can be artificial or natural, genetic drift and species dependent characters as the mating system. In our crop plants the domestication and breeding process may have caused such LD.

First association mapping studies performed with plants started just some years ago (Skøt *et al.* 2004, Aranzana *et al.* 2005). There are two different approaches for mapping genes or QTLs, the candidate gene approach and the genome-wide association analysis. Up to now, more studies used the candidate gene association mapping approach (e.g. Tommasini *et al.* 2007, Matthies *et al.* 2008), while for genome-wide association mapping in cereals only very few studies exist (Breseghello & Sorrels 2006a, 2006b, Kraakman *et al.* 2006, Comadran *et al.* 2008).

The development of SNP arrays (Cavanagh *et al.* 2013, Wang *et al.* 2014) also in wheat made genome wide association analysis easier applicable as SNP markers are a reliable, fast and cheap method of profiling the whole genome and useful for marker assisted selection as well. A first genome wide association mapping study was published by Crossa *et al.* (2007). The authors analysed historical multi-environment trials for the traits grain yield as well as resistance to the diseases stem rust, leaf rust, stripe rust and powdery mildew. Numerous marker trait associations were detected. Regarding frost tolerance in wheat Zhao *et al.* 2013 applied a genome wide association mapping approach to reveal a frost tolerance locus located on chromosome 5B.

Because frost is such an important abiotic factor limiting winter wheat yield, especially in the eastern European and north american growing regions with a continental climate, the project aims at improving the frost tolerance of wheat. This should be achieved by applying a whole genome and a candidate gene association genetics approach for the development of markers facilitating an efficient marker based selection procedure for frost tolerance. The objectives of the project are to:

(1) get information on the genetic variability for frost tolerance present in a set of diverse winter wheat cultivars
(2) identify genomic regions influencing winter hardiness by applying a whole genome association genetics approach
(3) resequence candidate genes for winterhardiness and frost tolerance on the set of genotypes in order to identify haplotypes related to these traits
(4) develop molecular markers based on the results obtained in the whole genome and candidate gene approach which facilitate efficient marker based selection procedure.

Materials and Methods

Plant material

A wheat panel comprising around 340 varieties/lines were selected. Most of the varieties were described as winter types. About 2/3 of the panel were registered varieties. The others breeding lines. Almost all of the material comes from regions having a continental climate with around 21% of the varieties bred in Russia, 16% of the varieties derived from North America and 6% from Scandinavia. The other varieties and breeding lines are mainly from western European countries or eastern Europe. Furthermore a few accessions from Asia and America were added.

Besides the wheat panel segregating populations were used respectively designed based on the frost tolerance testing results. The parents of the crosses showed a different level of frost stress response and segregation in the double haploid populations was expected.

Phenotyping in the field

The wheat variety/line panel was evaluated in ten field trials at five different locations in 2011/2012. Novosibirsk, Roschinskij and Pushkin are locations in different regions of Russia. In addition Gatersleben and Ranzin in Germany conducted trials. In 2012/2013 the evaluation was done on the six different locations Roschinskij, Pushkin, Gatersleben, Ranzin, Boehnshausen and Leutewitz. The testing was realized in 2011/2012 and 2012/2013 with double rows per variety/line as testing unit and each double row occurring in two replications. For describing the frost tolerance the percentage of surviving plants after winter was determined. Only in Roschinskij unreplicated microplots were used as testing unit and a 1to 9 scale was used for scoring frost tolerance. In addition to frost tolerance the traits heading, height, maturity, yield and if occurring diseases were recorded on some of the locations.

The testing of two DH-populations in 2013/14 was run in the same manner but only at the Russian locations.

Phenotyping under controlled conditions

The frost tolerance under controlled conditions was tested in climate chambers by using an established method at the Agricultural Institute, Centre for Agricultural Research, Hungarian Academy of Sciences. Seedlings of around 300 accessions from the wheat/variety panel were grown and before freezing application cold hardened. The freezing was done over six days by decreasing the temperature from 0°C to -14°C or -16°C stepwise. After the freezing treatment the temperature was gradually increased to 17°C for regeneration. The recovery was scored seven days and twenty-one days after the freezing treatment by using a 0-5 scale and looking actually at the regrowth with 0 meaning no regrowth (plants dead) and 5 normal growth (plants fully alive). Per variety/line 5 plants in 4 replications were evaluated.

DNA extraction

For each variety/line from the wheat panel plants were grown and leaf samples for extraction of genomic DNA cut at the three leaf stage. DNA was extracted according to the method of Stein *et al.* (2001).

Genotyping and genome wide association study

From the wheat panel 276 varieties/lines were genotyped with the 90k Infinium SNP array (Wang *et al.* 2014). From the 81587 SNPs 13857 completely failed and 29678 were monomorphic. In addition SNPs with more than 10% missing values were excluded and also markers with a minor allele frequency of less than 0.1. For the further analyses 16968 SNPs were finally used. The population structure was checked with STRUCTURE v 2.3.3 (Pritchard *et al.* 2000). Furthermore a dendrogram was calculated based on 249 mapped markers by using the software PAUP (Swofford 2003). For the genome wide association study a multi linear model with K-matrix as correction for population structure was in the end applied, after testing also a model with Q- and K-matrix as correction factors. Associations were calculated for each environment separately, because the broad sense heritabilities over all environments were quite low. For the association calculations the software package TASSEL (Trait Analysis by aSSociation, Evolution and Linkage; Bradbury *et al.* 2007) was used.

In addition the panel was checked with SSR markers to compare DNA lots. Furthermore height related markers were applied.

Candidate gene re-sequencing and association study

Sequences of 27 candidate genes involved in frost response were retrieved from NCBI (National Center of Biotech-

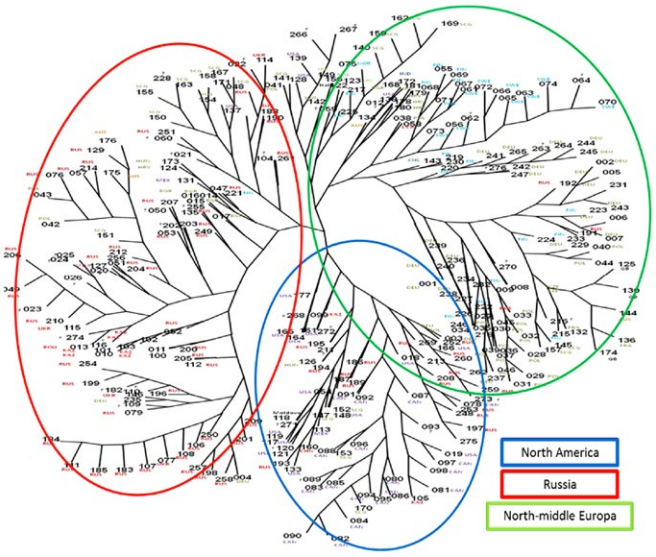


Fig. 1: PAUP tree of the wheat variety panel used for studying frost tolerance visualizing the genetic relatedness

		Years / locations	
chr.	pos. (cM)	2012	2013
1B	**56 - 65**	**Gatersleben, Pushkin, Ranzin**	**Leutewitz, Pushkin, Roschinskij**
2B	95 - 99	Gatersleben, Ranzin	Roschinskij
4B	50	Ranzin	Roschinskij
5A	**62 - 64**	**Gatersleben, Novosibirsk, Pushkin, Roschinskij**	**Leutewitz, Pushkin, Roschinskij**
6B	55	Ranzin	Pushkin
7B	139	Gatersleben, Ranzin	Pushkin
7B	160	Gatersleben, Roschinskij	Pushkin

Fig. 2: Genomic regions associated tp frost tolerance over several environments in the wheat variety panel

nology, http://www.ncbi.nlm.nih.gov/) and the full genomic sequence reconstructed, if necessary, by using the International Wheat Genome Sequencing Consortium (IWGSC, http://www.wheatgenome.org/) and/or the Bristol Wheat Genomics (http://www.cerealsdb.uk.net/) databases. The gene intron-exon-structure were figured out with internet platform 'Spidey' (http://www.ncbi.nlm.nih.gov/spidey/spideyweb.cgi) from NCBI. The intron/UTR regions sequences were used for primer development. The primers were designed by using 'Primer3' (v. 0.4.0). The amplified fragments with these primers were checked on agarose for fit to expectation. Sequencing of PCR fragments was performed by Microsynth AG (Balgach, Switzerland) using the Sanger sequencing method. Further details can be studied in Babben *et al.* (2015). For the polymorphisms in the candidate genes associations to the frost tolerance phenotypes in the field and the controlled conditions were calculated. The basis for these calculations was different than in the genome wide approach. In the candidate gene associations the number of days below -15°C was considered as covariable.

Results

Wheat panel genetic structure

Almost all of the genotyped 276 wheat varieties/lines were winter wheat, because the winter killed varieties in 2011/2012 were excluded. The population structure analysis resulted in three subgroups, which can be reasonably explained by the geographic origin of the material. The PAUP tree in Fig. 1 shows the relatedness of the material and a North American, Russian and North/Middle European cluster was observed.

Genome wide associations

Considering the three clusters in the genetic structure genome wide associations to frost tolerance, heading, maturity and height in the field were calculated by using a multi linear model with K-matrix as correction for population structure.

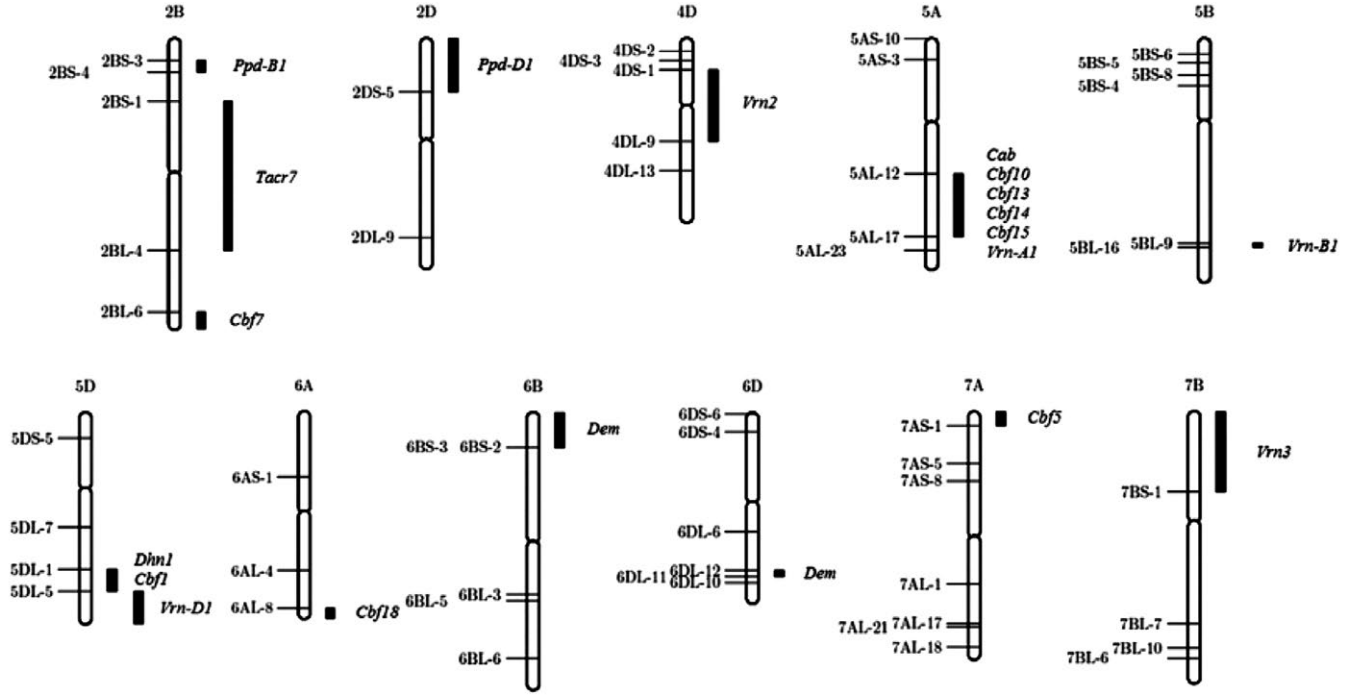


Fig. 3: Map of gene specific PCR fragments by using wheat NT- and deletion lines. In this figure only wheat chromosomes are shown harbouring mapped PCR fragments. The white bar is the chromosome, the constriction symbolised the centromere, on the left side of chromosomes deletion breaks points are listed and the black bars are the regions of mapped PCR fragments with appended candidate gene.

For heading a significant association over several environments on chromosome 5B were found. Regarding the maturity the region on chromosome 5B associated with heading was significant again, but other SNP markers. In addition chromosome 2D carries an association, which is stable over different environments as well. This region was associated to height too. In this respect several other chromosomes showed markers associated (1D, 2B, 4B, 5A, 6B, 7A).

Nevertheless the major trait of interest was frost tolerance. Six different chromosomes were associated to frost tolerance. The Fig. 2 contains the overview which genomic region on which chromosome was linked to frost tolerance on which location. Outstanding are chromosome 1B and chromosome 5A.

For frost tolerance also the haplotypes in the chromosome 1B and 5A associated regions were studied. For the locus on 1B three SNPs were used to describe two haplotypes with 234 varieties carrying an ACC pattern and 38 varieties having a CTA pattern. Whereas the ACC pattern is showing on the most environments a higher frost tolerance. The 5A locus was characterized by seven SNPs yielding three different haplotypes. One of the haplotypes was occurring only in two varieties. The other two haplotypes showed a relation similar to the 1B haplotypes. The major 5A haplotype showed up in 228 varieties, the other haplotype in 44 varieties. Like at the 1B locus the frequent haplotype had on the most environments a higher frost tolerance. Comparing the less frequent 1B haplotype group with the less frequent 5A haplotype group it was realized that 21 varieties belong to both groups. The geographic origin of these varieties is mostly France or South-East Europe. So regions where not so often frost is occurring.

The genome wide association study against the frost tolerance under controlled conditions resulted in hits on chromosome 1B, 5A and 5B. The associated region on 1B is the same like from the field related genome wide association study.

Furthermore on the chromosome 1B no candidate gene influencing frost tolerance is described.

Candidate gene polymorphisms and associations to frost tolerance

The essential prerequisite for the re-sequencing of the 27 selected candidate genes was the development of gene and genome specific primers. For all of the 27 candidate genes the re-construction of the gene structure or at least a part of it was successful. A set of 119 PCR products was obtained from 157 primers pairs designed in this. These were tested for PCR functionality and specificity. A set of 86 primer combinations from 23 candidate genes showed single band amplification (72.27%). By using Nulli-tetrasomic (NT)-lines of Chinese Spring resulted in correct assignment of amplicons originating from 19 genes and were located on 11 wheat chromosomes (Fig. 3).

Gene	Number of varieties with sequence	Number of polymorphisms	SNPs	InDels	Haplotypes
Cab	276	48	26	3	6
CBF3	276	4	4	0	2
CBF5	276	6	5	1	4
CBF10	276	2	2	0	2
CBF13	276	38	5	2	3
CBF14	276	19	7	1	3
CBF15A	276	15	6	1	2
CBF15B	276	6	0	1	2
CBF18	276	43	34	3	3
Tacr7	276	5	5	0	3
Vrn3	276	1	1	0	3
Vrn-A1	276	36	10	4	4
Vrn-B1	276	28	21	3	6
Vrn-D1	276	3	3	0	3
Ppd-B1	276	5	5	0	6
Ppd-D1	276	132	98	12	6

Fig. 4: Candidate genes with number of polymorphism and haplotypes

For the 23 genes the sequencing in the wheat panel (the same 276 like in the SNP genotyping) was realized. This was achieved by producing 45 amplicons with sequencing afterwards. All 45 obtained sequences were compared to retrieved gene models via MegaBlast. The best BLAST hit was for all against the starting sequence, from which the PCR primers were derived.

The sequences for the candidate genes were checked for polymorphisms (SNPs and Insertions/Deletions [InDels]). For 17 of the 23 candidate genes differences between the 276 wheat varieties/lines were observed. However CBF2 contained a lot of ambiguous bases. Therefore it was excluded from the further analysis. Figure 4 gives an overview about the polymorphisms and number of haplotypes in the wheat variety/line panel.

The association study of the field frost tolerance data against this polymorphisms resulted in four candidate genes (CBF3, CBF13, CBF14 and CBF15A) carrying either SNPs and/or InDels with a significant association to frost response. Depending on the associated polymorphism the effect of the different alleles on the frost tolerance was between 6.7 to 9.4%. Examining the haplotypes against the frost tolerance also CBF10 yielded a significant association to frost tolerance. The favorable haplotype showed an improvement in frost tolerance of 9.8%. All the CBFs with significant effects are located on chromosome 5A. In each of the associated genes only two haplotypes were observed. For CBF3, CBF13 and CBF14 the varieties in the wheat panel showed exactly the same behavior regarding the grouping into the haplotype groups. The smaller haplotype group consisted of 57 varieties. These varieties are coming mainly from France, South-East Europe, Scandinavia, a few from Poland and USA. Concerning CBF15A 61 varieties belong to the small haplotype group. The 57 varieties which make up the CBF3, CBF13

and CBF14 minor group and four varieties in addition. These four additional ones are coming either from South-East Europe or France. The bigger haplotype group had the higher frost tolerance level.

Comparing the candidate gene haplotype groups to the SNP haplotype groups from the genome wide association approach results in 14 varieties making up an intersection over all single small haplotype groups. Nine varieties are coming from Sweden and are related by descent, four are from France and in addition a variety from USA.

Running the analysis against the climate chamber data delivered no significant association. For polymorphisms in CBF15A and Vrn-A1 marker assays were developed.

Conclusions

A wide range of wheat germplasm was tested extensively for frost tolerance in the field and in the climate chamber as well. Although not all field trials delivered sufficient differences in frost tolerance useful for association studies the genome wide and candidate gene approach resulted in the detection of genetic effects onto the frost tolerance. In both approaches chromosome 5A was found to be associated with frost tolerance. The chromosome 5A carries besides Vrn-A1 also some CBF genes. In addition a locus on chromosome 1B was in the genome wide study described to be associated to frost tolerance. Unfortunately the genetic effects could not be validated so far, because the parents of the DH populations, which were phenotyped, carried neither a polymorphism detected by the genome wide approach nor a polymorphism discovered in the candidate gene study and further DH populations are not yet phenotyped for frost tolerance.

Acknowledgments

The authors thank the German Federal Ministry of Education and Research (BMBF) for funding the project FROWHEAT (0315953).

Literature cited

- *Anonymous (2010) Beschreibende Sortenliste 2010 Getreide, Mais, Öl- und Faserpflanzen, Leguminosen, Rueben, Zwischenfruechte: 86-94 ed. Bundessortenamt*
- *Akhunov E, Nicolet C, Dvorak J (2009): Single nucleotide polymorphism genotyping in polyploidy wheat with the Illumina GoldenGate assay. Theor Appl Genet 119: 507-517*
- *Aranzana MJ, Kim S, Zhao K, Bakker E, Horton M, Jakob K, Lister C, Molitor J, Shindo C, Tang C, Toomajian C, Traw B, Zheng H, Bergelson J, Dean C, Marjoram P, Nordborg M (2005) Genome-wide association mapping in Arabidopsis identifies previously known flowering time and pathogen resistance genes. PLoS Genet 1(5): e60. doi:10.1371/journal.pgen.0010060*
- *Arbuzova VS, Efremova TT, Laikova LI, Maystrenko OI, Popova OM, Pshenichnikova TA (1996) The development of precise genetic stocks in two wheat cultivars and their use in genetic analysis. Euphytica 89: 11-15*
- *Babben S, Perovic D, Koch M, Ordon F (2015): An efficient method for the development of locus specific primers in bread wheat (Triticum aestivum L.) and its application in re-sequencing of genes involved in frost tolerance. PLOS ONE DOI:10.1371/journal.pone.0142746*
- *Badawi M, Danyluk J, Boucho B, Houde M, Sarhan F (2007) The CBF gene family in hexaploid wheat and its relationship to the phylogenetic complexity of cereal CBFs. Molecular Genetics and Genomics 277: 533-554*
- *Baker SS, Wilhelm KS, Thomashow F (1994) The 5´-region of Arabidopsis thaliana cor15a has cis-acting elements that confer cold-, drought-, and ABA-regulated gene expression. Plant Molecular Biology 24: 701-713*
- *Börner A, Schumann E, Fürste A, Cöster H, Leithold B, Röder MS, Weber WE (2002) Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (Triticum aestivum L.). Theor Appl Genet 105: 921-936*
- *Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Bucker ES (2007): TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19): 2633-2635. doi: 10.1093/bioinformatics/btm308*
- *Breseghello F, Sorrells ME (2006a) Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. Genetics 172: 1165–1177*
- *Breseghello F, Sorrells ME (2006b) Association analysis as a strategy for improvement of quanti-tative traits in plants. Crop Sci. 46: 1323–1330*
- *Campoli C, Matus-Cádiz MA, Pozniak CJ, Cattivelli L, Fowler DB (2009) Comparative expression of Cbf genes in the Triticeae under different acclimation induction temperatures. Molecular Genetics and Genomics 282: 141-152*
- *Cavanagh CR, Chao S, Wang S, Huang BE, Stephen S, Kiani S, Forrest K, Saintenac C, Brown-Guedira GL, Akhunova A, See D, Bai G, Pumphrey M, Tomar L, Wong D, Kong S, Reynolds M, da Silva ML, Bockelman H, Talbert L, Anderson JA, Dreisigacker S, Baenziger S, Carter A, Korzun V, Morrell PL, Dubcovsky J, Morell MK, Sorrells ME, Hayden MJ, Akhunov E (2013): Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. Proc Natl Acad Sci 110(20):8057-8062*
- *Chao S, Zhang W, Akhunov E, Sherman J, Ma Y, Luo MC, Dubcovsky J (2009): Analysis of gene-derived SNP marker polymorphism in US wheat (Triticum aestivum L.) cultivars. Molecular Breeding 23: 23-33*
- *Chekurov VM & Kozlov VY (2003) Winter wheat´s main survival mechanisms in Siberia: low metabolic rate and high frost tolerance. Increasing wheat production in Central Asia through science and international cooperations - Proceedings of the First Central Asian Wheat Conference ed. Morgounov A, McNab A, Campbell KG, Paroda R: 118-121*
- *Chesnokov YV, Pochepnya NV, Börner A, Lohwasser U, Goncharova EA, Dragavtsev VA (2008) Ecological-genetic organization of plant quantitative traits and mapping of the loci determining agronomically important traits in soft wheat. Doklady Biochemistry and Biophysics 418: 36-39*

- *Comadran J; Russel JR; van Eeuwijk FA; Cecacarelli S; Grando S; Baum M; Stanca AM;*
- *Peccioni N; Mastrangelo AM; Akar T; Al Yassin A; Benbelkacem A; Choumane W; Ouabbou H; Dahan R; Bort J; Araus J L; Pswarayi A; Romagosa I; Hackett C A; Thomas*
- *WTB (2008) Mapping adaptation of barley to droughted environments. Euphytica 161,35-45*
- *Crossa J, Burgueno J, Dreisigacker S, Vargas M, Herrera-Foessel SA, Lillemo M, Singh RP, Tre-thowan R, Warburton M, Franco J, Reynolds M, Crouch JH, Ortiz R (2007) Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. Genetics 177: 1889-1913*
- *Distelfeld A, Li C, Dubcovsky J (2009) Regulation of flowering in temperate cereals. Curr Opin Plant Biol 12:178–184*
- *Dhillon T, Pearce SP, Stockinger EJ, Distelfeld A, Li C, Knox AK, Vashegyi I, Vágújfalvi A, Galiba G, Dubcovsky J (2010) Regulation of freezing tolerance and flowering in temperate cereals. the VRN-1 connection. Plant Physiology 153: 1846-1858*
- *Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecology 14: 2611–2620*
- *Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 164: 1567–1587*
- *Francia E, Rizza F, Cattivelli L, Stanca AM, Galiba G, Toth B, Hayes P, Skinner JS, Pecchioni N (2004) Two loci on chromosome 5H determine low-temperature tolerance in a ´Nure´ (winter) x ´Tremois´ (spring) barley map. Theor Appl Genet 108: 670-680*
- *Fricano A, Rizza F, Faccioli P, Pagani D, Pavan P, Stella A, Rossini L, Piffanelli P, Cattivelli L (2009) Genetic variants of HvCbf14 are statistically associated with frost tolerance in a European germplasm collection of Hordeum vulgare. Theor Appl Genet 119: 1335-1348*
- *Huang XQ, Cloutier S, Lycar L, Radovanovic N, Humphreys DG, Noll JS, Somers DJ, Brown PD (2006) Molecular detection of QTLs for agronomic and quality traits in a doubled haploid popu-lation derived from two Canadian wheats (Triticum aestivum L.). Theor Appl Genet 113: 753–766*
  http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567
- *Jiang C, Iu B, Singh J (1996) Requirement of a CCGAC cis-acting element for cold induction of the BN115 gene from winter Brassica napus. Plant Molecular Biology 30: 679-684*
- *Khlestkina EK, Röder MS, Pshenichnikova TA, Börner A (2010) Functional diversity of the Rc (red coleoptile) in bread wheat. Mol. Breeding 25: 125-132*
- *Knox AK, Li C, Vágújfalvi A, Galiba G, Stockinger EJ, Dubcovsky J (2008): Identification of candidate CBF genes for the frost tolerance locus Fr-A (m) 2 in Triticum monococcum. Plant Molecular Biology 67: 257-270*
- *Kocsy G, Athmer B, Perovic D, Himmelbach A, Szűcs A, Vashegyi I, Schweizer P, Galiba G, Stein N (2010) Regulation of gene expression by chromosome 5A during cold hardening in wheat. Mol Genet Genomics 283:351–363*
- *Knox AK, Dhillon T, Cheng H, Tondelli A, Pecchioni N, Stockinger EJ (2010) CBF gene copy number variation at Frost Resistance-2 is associated with levels of freezing tolerance in temperate-climate cereals. Theor Appl Genet 121: 21–35*
- *Kosova K, Vitámvás P, Prášil IT (2010) Expression of dehydrins in wheat and barley under different temperatures. Plant Science (article in press)*
- *Kraakman ATW, Martinez F, Mussiraliev B, van Eeuwijk FA, Niks RE (2006) Linkage disequilib-rium mapping of morphological, resistance, and other agronomically relevant traits in modern spring barley cultivars. Molecular Breeding 17: 41–58*
- *Kumar N, Kulwal PL, Balyan HS, Gupta PK (2007) QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. Molecular Breeding 19: 163-177*
- *Li, Y, Haseneyer, G, Schön, CC, Ankerst, D, Korzun, V, Wilde, P, Bauer, E (2011a): High levels of nucleotide diversity and fast decline of linkage disequilibrium in rye (Secale cereale L.) genes involved in frost response. BMC Plant Biology 11:6*
- *Li, Y, Böck, A, Haseneyer, G, Korzun, V, Wilde, P, Schön, CC, Ankerst, DP, Bauer, E (2011b): Association analysis of frost tolerance in rye using candidate genes and phenotypic data from controlled, semi-controlled, and field phenotyping platforms. BMC Plant Biology 11:146*
- *Liu Q, Kasuga M, Sakuma Y, Abe H, Miura S, Yamaguchi-Shinozaki K, Shinozaki K (1998) Two transcription factors, DREB1 and DREB2, with an EREBP/AP2 DNA binding domain separate two cellular signal transduction pathways in drought- and low-temperature-responsive gene expression, respectively, in Arabidopsis. Plant Cell 10: 1391-1406*
- *Matthies IE, Weise S, Röder MS (2009) Association of haplotype diversity in the a-amylase gene amy1 with malting quality parameters in barley. Molecular Breeding 23: 139-152*
- *Neumann K, Kobiljski B, Denčić S, Varshney RK, Börner A (2010) Genome-wide association mapping – a case study in bread wheat (Triticum aestivum L.). Molecular Breeding (2010) DOI 10.1007/s11032-010-9411-7*
- *Pearce S, Zhu J, Boldizsár Á, Vágújfalvi A, Burke A, Garland-Campbell K, Galiba G, Dubcovsky J (2013): Large deletions in the CBF gene cluster at the Fr-B2 locus are associated with reduced frost tolerance in wheat. Theor Appl Genet: DOI 10.1007/s00122-013-2165-y*
- *Pritchard JK, Stephens M, Donnelly PJ (2000) Inference of population structure using multilocus genotype data. Genetics 155: 945–959*
- *Quarrie SA, Steed A, Calestani C, Semikhodskii A, Lebreton C, Chinoy C, Steele N, Pljevljakusic D, Waterman E, Weyen J, Schondelmaier J, Habash DZ, Farmer P, Saker L, Clarkson DT, Abugalieva A, Yessimbekova M, Turuspekov Y, Abugalieva S, Tuberosa R, Sanguineti MC, Hollington PA, Aragues R, Royo A, Dodig D (2005) A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring × SQ1 and its use to compare QTLs for grain yield across a range of environments. Theor Appl Genet 110: 865–880*
- *Sanger F, Nicklen S, Coulson AR (1977): DNA SEQUENCING WITH CHAIN-TERMINATING INHIBITORS. Proc Natl Acad Sci U S A 74(12):5463–7. doi: 10.1073/pnas.74.12.5463*
- *Sears ER (1966): Nullisomictetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR, editors. Chromosome manipulations and plant genetics. p. 29–45*
- *Skøt L, Humphreys MO, Armstead I, Heywood S, Skøt KP, Sanderson R, Thomas ID, Chorlton KH, Hamilton NRS (2004) An association mapping approach to identify flowering time genes in natural populations of Lolium perenne (L.). Molecular Breeding 15: 233–245*

- *Stein N, Herren G, Keller B (2001): A new DNA extraction method for high-throughput marker analysis in a large-genome species such as Triticum aestivum. Plant Breed. 120(4):354–6. doi: 10.1046/j.1439-0523.2001.00615.x*
- *Stockinger EJ, Gilmour SJ, Thomashow MF (1997) Arabidopsis thaliana CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. PNAS 94: 1035-1040*
- *Stockinger EJ, Skinner JS, Gardner KG, Francia E, Pecchioni N (2007) Expression levels of barley Cbf genes at the Frost resistance-H2 locus are dependent upon alleles at Fr-H1 and Fr-H2. The Plant Journal 51: 308-321*
- *Sutka J (1981) Genetic studies of frost resistance in wheat. Theor Appl Genet 59: 145-152*
- *Sutka J & Snape JW (1989) Location of a gene for frost resistance on chromosome 5A of wheat. Euphytica 42: 41-44*
- *Swofford DL (2003). PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts*
- *Thomashow MF (1990) Molecular genetics of cold acclimation in higher plants. Adv Genet 28: 99-131*
- *Thomashow MF (1999) PLANT COLD ACCLIMATION: Freezing tolerance genes and regulatory mechanisms. Annu Rev Plant Phys 50: 571-599*
- *Tommasini L, Schnurbusch T, Fossati D, Mascher F, Keller B (2007) Association mapping of Stagonospora nodorum blotch resistance in modern European winter wheat varieties. Theor Appl Genet 115: 697–708*
- *Vágújfalvi A, Crosatti C, Galiba G, Dubcovsky J, Cattivelli L (2000) Two loci on wheat chromosome 5A regulate the differential cold-dependent expression of the cor14b gene in frost-tolerant and frost-sensitive genotypes. Molecular Genetics and Genomics 263: 194-200*
- *Vágújfalvi A, Galiba G, Cattivelli L, Dubcovsky J (2003) The cold-regulated transcriptional activator Cbf3 is linked to the frost-tolerance locus Fr-A2 on wheat chromosome 5A. Molecular Genetics and Genomics 274: 506-514*
- *Vitámvás P, Kosova K, Prášilova P, Prášil IT (2010) Accumulation of WCS120 protein in wheat cultivars grown at 9°C or 17°C in relation to their winter survival. Plant Breeding doi:10.1111/j.1439-0523.2010.01783.x*
- *Wang S, Wong D, Forrest K, Allen A, Chao S, Huang BE, Maccaferri M, Salvi S, Milner SG, Cattivelli L, Mastrangelo AM, Whan A, Stephen S, Barker G, Wieseke R, Plieske J, International Wheat Genome Sequencing Consortium, Lillemo M, Mather D, Appels R, Dolferus R, Brown-Guedira G, Korol A, Akhunova AR, Feuillet C, Salse J, Morgante M, Pozniak C, Luo MC, Dvorak J, Morell M, Dubcovsky J, Ganal M, Tuberosa R, Lawley C, Mikoulitch I, Cavanagh C, Edwards KJ, Hayden M, Akhunov E (2014): Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. Plant Biotechnology Journal 12(6):787-796*
- *Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. PNAS 100: 6263–6268*
- *Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303: 1640–1644*
- *Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. PNAS 103: 19581– 19586*
- *Yoshida T, Nishida H, Zhu J, Nitcher R, Distelfeld A, Akashi Y, Kato K, Dubcovsky J (2010) Vrn-D4 is a vernalization gene located on the centromeric region of chromosome 5D in hexaploid wheat. Theor Appl Genet 120:543–552*
- *Wenzl P, Carling J, Kudrna D, Jaccoud D, Huttner E, Kleinhofs A, Kilian A (2004) Diversity Arrays Technology (DArT) for whole-genome profiling of barley. PNAS 101: 9915–9920*
- *Zhao Y, Gowda M, Würschum T, Longin CFH, Korzun V, Kollers S, Schachschneider R, Zeng J, Fernando R, Dubcovsky J and Reif JC (2013) Dissecting the genetic architecture of frost tolerance in Central European winter wheat. Journal of Experimental Botany 64: 4453-4460*
- *Zhu J, Pearce S, Burke A, See DR, Skinner DZ, Dubcovsky J, Garland-Campbell K (2014): Copy number variation and haplotype variation at the VRN-A1 and central FR-A2 loci are associated with frost tolerance in hexaploid wheat. Theor Appl Genet: DOI 10.1007/s00122-014-2290-2*

HAPLOIDS Herstellung von Haploiden mittels uniparentaler Genomeliminierung

Jochen Kumlehn, Plant Reproductive Biology, IPK Gatersleben, Corrensstraße 3, 06466 Seeland/Gatersleben, Germany
Andreas Houben, Chromosome Structure and Function, IPK Gatersleben,
Corrensstraße 3, 06466 Seeland/Gatersleben, Germany
Nils Stein, Genomics of Genetic Resoureces, IPK Gatersleben, Corrensstraße 3, 06466 Seeland/Gatersleben, Germany

Aims

The aim of the HAPLOIDS project has been to establish of a novel principle of haploid production in the crop species barley and sugar beet, which relies on uniparental chromosome elimination owing to manipulated function of the centromer-specific histon CENH3. To further facilitate this approach, the well-manageable experimental model Nicotiana benthaminana was used in addition. The envisaged experiments aimed to result not only in the establishment of novel haploid technology in the above species, but also to improve our understanding of the general cell-biological processes in the context of uniparental chromosome elimination. The overall goal of the project was to provide knowledge and methodological information on uniparental genome elimination to facilitate their implementation towards substantially improved efficacy of breeding programs.

Results

Gene cloning and synthesis (Kumlehn, Houben)
To identify the sequences of the promoter region and to complete the genomic sequence of the CENH3 genes of *Nicotiana benthamiana* and of *Hordeum vulgare* (α and β type), we constructed four different genomic DNA libraries for each species using the ‚Genome walker Universal kit' (Clontech). Two successive PCRs were performed for each library using a combination of gene specific primers and universal primers (provided by Genome walker kit). We achieved about 2 kb sequences of the 5' regions of the CENH3 genes from *H. vulgare* (α, β) and *N. benthamiana*.

By BLAST search analysis using the α and β CENH3 sequences of *H. vulgare* we obtained by genome walking in combination with the previously available CDS the complete sequences of barley α and β CENH3. Finally, the promotor/intron/ exon structure of the α and β CENH3 genes were assembled by alignment of CENH3 genomic sequence with the CDS.

Knock-down/knock-out of CenH3
(Kumlehn, Houben, Stein)
NbCENH3-specific TALENs generated and validated in yeast (by CELLECTIS) were cloned into expression cassettes under control of the CaMV35S promoter with doubled enhancer and NOS terminator (Figure 1).

Primary transgenic plants expressing one or both left and right TALEN-units were produced and analysed for induced mutations in the native *NbCENH3* gene. However, no mutation could be identified. Selected primary plants were cross-combined to produce more plants expressing both TALEN-units along with one of the complementation constructs. The crossbred offspring was screened and plants expressing either both TALEN-units and one of the complementation construct or only both TALEN-units were further analysed (see WP3). Again, mutations could not be detected in the progeny within the TALEN target region.

We alternatively used RNA-guided endonuclease (RGEN) technology to target the native *NbCENH3*. We established this method in tobacco by transformation of plants carrying a single copy of the *GFP* reporter gene using a *gfp*-specific RGEN construct. The chimeric guide RNA (gRNA) was driven by a AtU6-26S promoter and contained a 20 bp protospacer providing the target sequence specificity. The endonuclease Cas9 was under control of the PcUbi4-2 promoter and the pea3A terminator (Figure 2). Primary transgenic plants expressing the Cas9 were used for further investigations involving amplification and sequencing of a ca. 560 bp fragment including the RGEN target site. Out of 11 analyzed plants, three carried various indels in the RGEN target region. Most

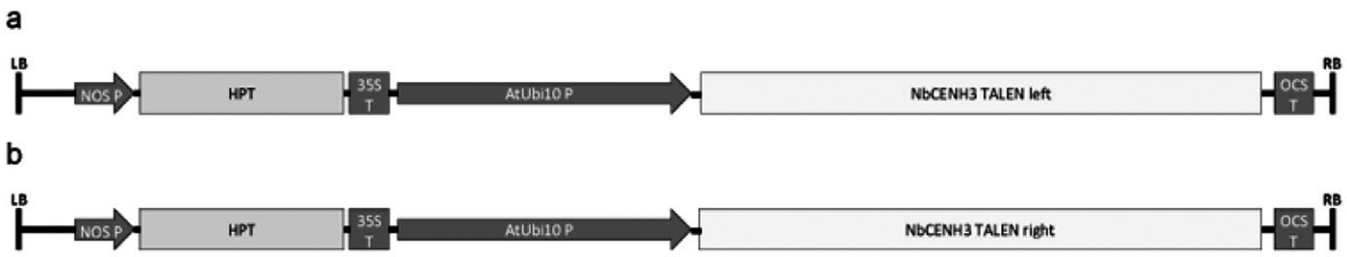


Figure 1: T-DNA constructs carrying NbCENH3-specific left (a) and right (b) TALEN-units.

Figure 2: T-DNA construct carrying gfp-specific RGEN.

of the mutations proved heritable as was shown by sequencing the *GFP* target region in T1 seedlings. To target the native *NbCENH3*, four different *NbCENH3*-specific RGENs, which recognize and bind specific sequences in the first, third or fifth exon of the native *NbCENH3*, were assembled. Re-transformation experiments using the (tailswap-) complemented plants as donor material resulted in transgenic plants used for amplification of a fragment carrying the target site of one of the *NbCENH3*-specific RGEN. However, no mutation was identified in the target region.

To elucidate whether, besides the severe conformational change using the CENH3-tailswap (Ravi and Chan, 2010; Ravi *et al.*, 2014), single point mutations in endogenous CENH3 could also affect centromere function for haploid induction, we screened an EMS-induced TILLING population of barley (Gottwald *et al.*, 2009), a diploid species having two functional variants of CENH3 (α and βCENH3) (Sanei *et al.* 2011, Ishii *et al.* 2015). Assuming that either CENH3 variant can compensate for the absence of the other, viable offspring should be observed in the presence of a loss of function allele in one of the two CENH3 variants. A total of twenty-five TILLING mutants were identified for both barley CENH3 genes.

The potential of *Hvßcenh3* L92F to act as a haploid inducer was tested. The analysis of 577 F1 plantlets obtained from crosses involving the *Hvßcenh3* L92F mutant as maternal (35 spikes) or paternal partner (for 22 wild-type barley spikes) did not reveal any haploid or otherwise hypoploid plants, whereas all 18 plantlets derived from nine wild-type barley spikes pollinated with H. bulbosum (as positive control for the procedure used to induce uniparental genome elimination) proved to be invariably haploid. This result indicates that the *Hvßcenh3* L92F mutation in the presence of native αCENH3 is not sufficient for chromosome elimination during early zygotic embryogenesis.

HvαCENH3 and HvβCENH3-specific TALENs were cloned into expression cassettes under the control of the maize UBI1 promoter and NOS terminator. The binary vectors were used for co-transformation experiments to target the native CENH3s (Figure 3).

Since a non-functional mutant of HvβCENH3 was identified via TILLING, the knock-out approaches in barley using TALENs were confined to HvαCENH3. Thus, we attempted to knock-out αCENH3 in the βcenh3 knock-out background. We performed crosses of the homozygous TILLING line 4528 lacking centromeric βCENH3-specific signals with the well-transformable barley cultivar 'Golden Promise'. All F1 hybrids carried the βcenh3 mutation in heterozygous state. After selfing, F2 immature embryos (grown on these F1 plants) were used for stable transformation with the binary vectors carrying αCENH3 specific TALENs. However, none of the transgenic plants possessed both TALEN units, and thus all still carried the native (WT) sequence of HvαCENH3.

Table 1: Overview of transgenic N. benthamiana produced by crossing primary transgenic plants expressing left and/or right TALEN-unit and/or one of the CENH3-tailswap constructs.

No. of transgenic *GFP-AtCENH3* plants with *GFP-NbH3*-tailswap	With left TALEN-unit	With right TALEN-unit	With -tailswap
14	+	-	-
11	+	+	-
34	+	-	+

Partial complementation of cenh3 mutant or RNAi-lines using recombinant CENH3-derivatives (Kumlehn)

To partially complement the native *NbCENH3*, two CENH3 derivatives were produced, namely a "*GFP*-tailswap" variant,

Figure 3: Schematic representation of the T-DNA TALEN constructs.

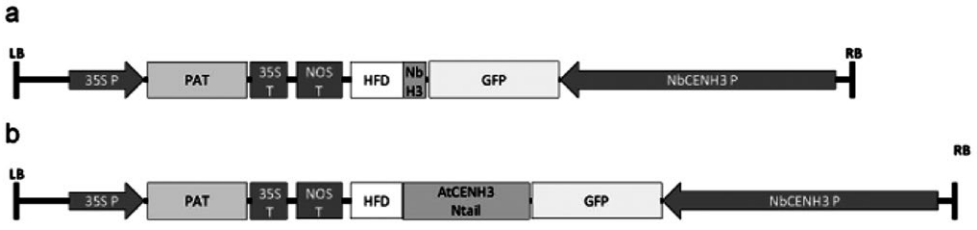


Figure 4: T-DNA constructs carrying cenh3-derivatives used for partial (pre-)complementation of cenh3 mutants of N. benthamiana. a) GFP-NbH3-tailswap, b) GFP-AtCENH3-tailswap.

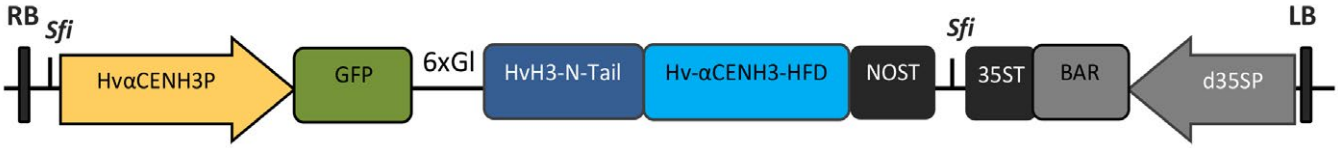


Figure 5: Schematic representation of the T-DNA of the "GFP-HvCENH3-tailswap" complementation construct for barley.

which contains the hypervariable N-tail from NbH3 fused with *GFP* and the histone fold domain from the native *Nb-CENH3*, and a so called "*GFP*-Arabidopsis-tailswap" variant containing the hypervariable N-tail from *A. thaliana* CENH3 (Figure 4, Table 1).

The transgenic plants obtained were used for re-transformation experiments using both TALEN-constructs and crossings to generate inducer lines that carry a tailswap construct in TALEN-induced cenh3 mutant background. The (tailswap) complemented plants were used as donor material for additional re-transformation experiments using different *Nb-CENH3*-specifc RGNs. The generation and analysis of these plants is still in progress.

Considering the results of the investigations using different CENH3 derivatives in Arabidopsis (Ravi & Chan 2010), we used a "*GFP*-HvCENH3-tailswap" variant, which contains the native HvCENH3 histone fold domain along with the hypervariable N-tail of canonical Histone H3 of barley translationally fused with *GFP* (Figure 5).

The resultant chimeric "*GFP*-HvCENH3-tailswap" gene does not contain the αCENH3 sequence motif targeted by the TALEN pair used to knock out the wild-type allele, so that crossing with TALEN-lines will not compromise its function. Agrobacterium mediated transformations with single TALEN units (left or right), co-transformations with both TALEN units, and co-transformation experiments with the constructs (both TALEN units and complementation construct) were performed. Transgenic barley plants were generated using either immature embryos (in case of cv. Golden Promise, GP x Barke βcenh3 mutant, and Barke βcenh3 mutant X GP) or immature pollen (cv. Igri) as explants (Table 2).

Reciprocal crossings with WT (Kumlehn)
The potential of Hvßcenh3 TILLING mutant to induce haploidy in barley was tested by reciprocal crosses with wild type barley (cv. Golden Promise). The barley TILLING line 4528 was first made homozygous for the leucine to phenylalanine substitution at amino acid 92 of ßCENH3 via generational segregation. To reduce the mutation background, it was then crossed with barley wild type cv. Golden Promise. As expected, all F1 hybrids carried the ßcenh3 mutation in heterozygous state. Upon selfing of F1 plants and sequencing ßCENH3 of F2 individuals, 14 selected homozygous ßcenh3 segregants derived from 7 independent F1 hybrids were used for reciprocal crosses with wild-type barley. In total, 58 crosses were performed, in 36 of them a mutant plant was used as a maternal parent, and in 22 crosses as a pollinator. In total, 577 plantlets were produced after embryo rescue, and for all of them genome size was determined. The analysis of these 577 F1 plantlets obtained from crosses involving the Hvßcenh3 mutant as maternal (35 spikes) or paternal partner (for 22 wild-type barley spikes) did not reveal any haploid or otherwise hypoploid plants. This result indicates that the Hvßcenh3 mutation in the presence of native αCENH3 is not sufficient for chromosome elimination during early zygotic embryogenesis.

Cell-biological analysis of chromosomal features and behavior in wild type and transgenic lines during early embryogenesis (Houben)
The organization of centromeric chromatin of diploid barley encoding two (α and β) CENH3 variants was analysed by super-resolution microscopy. Antibody staining revealed that both CENH3 variants are organized in distinct but intermingled subdomains in interphase, mitotic and meiotic centromeres. Artificially extended chromatin fibres illustrate that these subdomains are formed by polynucleosome clusters. Thus, a CENH3 variant-specific loading followed by the arrangement into specific intermingling subdomains forming the centromere region. The CENH3 composition and transcription varies among different tissues. In young embryos, most interphase centromeres are composed of both CENH3 variants, while in meristematic root cells a high number of nuclei contain βCENH3 mainly dispersed within the nucleoplasm. A similar distribution and no preferential arrangement of the

Table 2: Overview of transgenic barley produced using αCENH3 specific TALENs and 'GFP-Tailswap' complementation construct

Genetic background	No. of transgenic plants	transgenics with left TALEN	transgenics with right TALEN	transgenics with left + right TALEN	transgenics with tailswap	transgenics with tailswap + left or right TALEN
GP x Barke βcenh3 mutant	89	60	4	2	14	9
Barke βcenh3 mutant x GP	41	12	15	-	6	8
GP	122	57	9	4	33	19
Igri	53	19	10	-	19	5

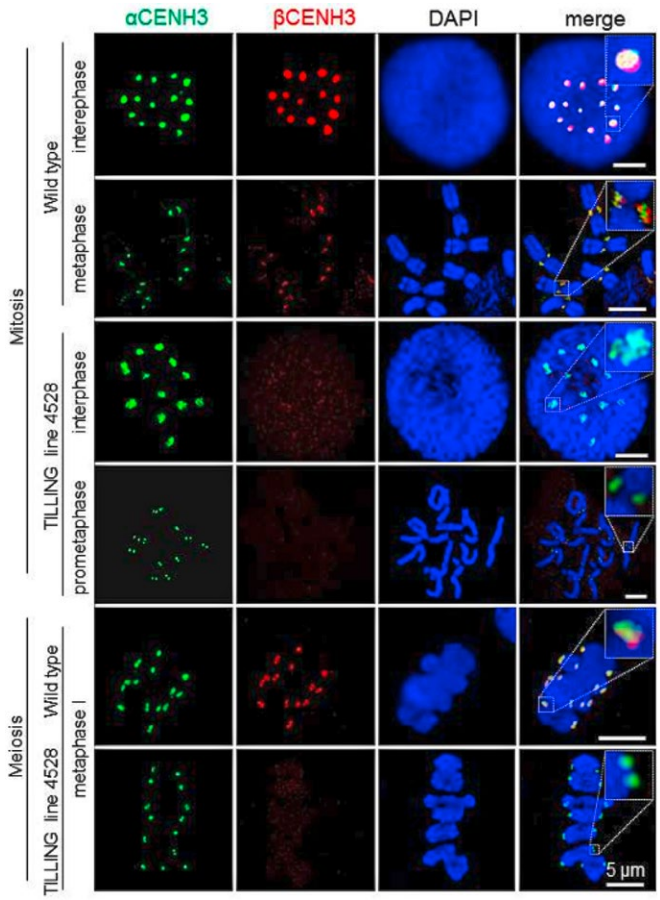


Figure 6: Centromeres of barley TILLING line 4528 lost βCENH3. Chromosomes of wild type and homozygous TILLING line 4528 after immunostaining with antibodies specific for αCENH3 (in green) and βCENH3 (in red). Note the absence of βCENH3-specific signals in line 4528.

two CENH3 variants in relationship to the spindle poles suggest that both homologs meet the same function in metaphase cells (Ishii *et al.*, 2015).

The functionality of mutated CENH3s of homozygous M2 individuals of the nine TILLING lines carrying non-synonymous mutations was determined by immunostaining of the centromeres with barley CENH3 variant-specific antibodies.

Alpha and βCENH3 signals at centromeres were revealed in all but one mutant TILLING genotype 4528 (called Hvßcenh3 L92F), carrying a homozygous leucine to phenylalanine substitution at amino acid 92. It showed no centromeric βCENH3 signals in mitotic, meiotic or interphase cells - only minor βCENH3 signals in the nucleoplasm outside of the centromeres were observed (Figure 6). Since no obvious differences in the transcription levels of both CENH3s between wild type and Hvßcenh3 L92F was found, the centromeric loading of the mutated βCENH3 seems to be impaired. Centromeres without βCENH3 are sufficient for mitotic centromere function as no obvious chromosome segregation defects, such as anaphase bridges or changes of (endopoly)ploidy could not be found (Karimi et al, under revision).

Analysis of uniparental chromosome elimination at the cellular level (Houben)

The Hvβcenh3 L92F mutation is located in the CENH3 centromere targeting domain (CATD), defined by loop 1 and α2 helix of the histone fold domain. This domain was shown to be required for centromere loading of CENH3 by Scm3/HJURP chaperons in non-plant species (Foltz *et al.*, 2009).

To prove whether the CATD mutation caused the observed impaired centromere loading, YFP was N-terminally fused to the coding sequence of *A. thaliana* CENH3 with an L/I or L/F exchange of the corresponding position (L130I or L130F) in *A. thaliana* CENH3. Double immuno labelling of transgenic *A. thaliana* with anti wild type CENH3 and anti-*GFP* revealed a significantly reduced centromere targeting, especially of the L to F mutated CENH3s compared to wild type CENH3 fused to YFP. The L/F exchange resulted in a stronger effect than L/I, likely caused by improper function or folding of the protein due to steric challenges imposed by the additional aromatic group provided by phenylalanine (Thomas *et al.*, 1995). These results indicate an impaired centromere targeting of CENH3 by a point mutation in the conserved amino acid sequence of mono and eudicot species.

To test for haploid inducer ability in A. thaliana, a genomic CENH3 construct previously used for functional complemen-

Cross (♀ x ♂)	Seed analysis							Ploidy analysis of seedlings*							
	Total seeds	Seed quality			Average number of seeds per silique (n)	Germinated seeds		Haploids		Diploids		Aneuploids		Mixoploids (haploid+aneuploid)	
		Normal (%)	Shriveled (%)	Nonviable (%)		Normal (%)	Shriveled (%)	Normal (%)	Shriveled (%)	Normal (%)	Shriveled (%)	Normal (%)	Shriveled (%)	Normal (%)	Shriveled (%)
Col x Col	100	99 (99.0)	0 (0)	1 (1.0)	32 (15)	90 (90.9)	0 (0)	0 (0)	0 (0)	90 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
L130F-1 x Col	520	341 (65.6)	39 (7.5)	140 (26.9)	23 (23)	282 (82.6)	29 (74.4)	8 (2.8)	7 (24.1)	254 (90.1)	13 (44.8)	19 (6.7)	7 (24.1)	1 (0.4)	2 (6.9)
L130F-2 x Col	442	335 (75.8)	50 (11.3)	57 (12.9)	20 (22)	307 (91.6)	25 (50.0)	0 (0)**	0 (0)	204 (97.1)**	13 (52.0)	6 (2.9)**	12 (48.0)	0 (0)**	0 (0)
L130F-3 x Col	186	165 (88.7)	13 (7.0)	8 (4.3)	32 (9)	129 (78.2)	8 (61.5)	0 (0)***	0 (0)	100 (100)***	6 (75.0)	0 (0)***	2 (25.0)	0 (0)***	0 (0)
Col x L130F-1	292	222 (76.0)	46 (15.8)	24 (8.2)	19 (15)	216 (97.3)	28 (60.9)	0 (0)	0 (0)	214 (99.1)	24 (85.7)	2 (0.9)	4 (14.3)	0 (0)	0 (0)
Col x L130F-2	100	99 (99.0)	0 (0)	1 (1.0)	44 (9)	96 (96.5)	0 (0)	0 (0)	0 (0)	96 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Col x L130F-3	50	50 (100)	0 (0)	0 (0)	25 (2)	41 (82.0)	0 (0)	0 (0)	0 (0)	41 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)

* seedling derived from seeds with a normal or shriveld phenotype ** 210 seedlings out of 307 germinated seeds were analysed *** 100 seedlings out of 129 germinated seeds were analysed

Table 3: Analysis of offsprings derived from the reciprocal cross Atcenh3 L130F x wild type A. thaliana.

tation of cenh3.1 homozygous knock-out plants (Ravi and Chan, 2010) was L130F mutated and used to transform heterozygous cenh3.1 knock-out plants (Ravi and Chan, 2010). Lines with either 1, 2 or 3 *Atcenh3* L130F transgene insertions were selected. Consistent with the result obtained with the reporter cenh3 L/F construct in A. thaliana, and sugar beet, a high proportion of *Atcenh3* L130F complemented cenh3.1 null mutants displayed limited centromeric anti-CENH3 signals in interphase nuclei. The single insertion line *AtAtcenh3* L130F-1 revealed the highest proportion of nuclei with an impaired CENH3 distribution. Unlike the tailswap-CENH3 haploid inducer (Ravi and Chan, 2010; Ravi *et al.*, 2014) the phenotype, meiosis, and seed setting of homozygous *Atcenh3* L130F lines were almost unaffected. Thus, despite a diminished centromere loading of *Atcenh3* L130F, mitosis and meiosis work sufficiently well to produce diploid offspring upon selfing.

When *Atcenh3* L130F-1 plants were pollinated with wild type A. thaliana, 25.9% and 2.8% of F1 seedlings grown from shriveled and normal seeds, respectively, were haploid, and possessed only the chromosomes of the wild type parent (Table 3). In addtion, plants derived from shriveled seeds were more often aneuploid or mixoploid. The same combination, including the other *Atcenh3* L130F lines, resulted in diploid and aneuploids only (Table 3). The reciprocal cross did not generate haploid plants. To test whether the haploid inducer efficiency correlates with the total amount of CENH3 a comparative Western blot analysis was conducted. Indeed, line L130F-1 and 2 possesed less total CENH3 than wild type and L130F-3 plants. These results suggest that depending on the amount of CENH3, a single point mutation in CENH3 is able to generate a haploid inducer in species carrying no more than one CENH3 variant.

References

- *Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-Specific Assembly of CENP-A Nucleosomes Is Mediated by HJURP. Cell 137, 472-484.*
- *Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., and Stein, N. (2009). TILLING in the two-rowed barley cultivar ‚Barke' reveals preferred sites of functional diversity in the gene HvHox1. BMC Res Notes 2, 258.*
- *Ishii, T., Karimi-Ashtiyani, R., Banaei-Moghaddam, A.M., Schubert, V., Fuchs, J., and Houben, A. (2015). The differential loading of two barley CENH3 variants into distinct centromeric substructures is cell type- and development-specific. Chromosome Research DOI: 10.1007/s10577-015-9466-8.*
- *Maeder ML, S Thibodeau-Beganny, JD Sander, DF Voytas and JK Joung (2009): Oligomerized Pool ENgineering (OPEN): An "Open-Source" Protocol for Making Customized Zinc Finger Arrays. Nat Protoc. 4, 1471–1501*
- *Puchta, H., and Fauser, F. (2014). Synthetic nucleases for genome engineering in plants: prospects for a bright future. Plant Journal 78, 727-741.*
- *Ravi, M., and Chan, S.W. (2010). Haploid plants produced by centromere-mediated genome elimination. Nature 464, 615-618.*
- *Ravi, M., Marimuthu, M.P.A., Tan, E.H., Maheshwari, S., Henry, I.M., Marin-Rodriguez, B., Urtecho, G., Tan, J., Thornhill, K., Zhu, F., Panoli, A., Sundaresan, V., Britt, A.B., Comai, L., and Chan, S.W.L. (2014). A haploid genetics toolbox for Arabidopsis thaliana. Nature Communications 5.*
- *Sanei, M., Pickering, R., Kumke, K., Nasuda, S., and Houben, A. (2011). Loss of centromeric histone H3 (CENH3) from centromeres precedes uniparental chromosome elimination in interspecific barley hybrids. Proc Natl Acad Sci U S A 108, E498-505.*
- *Thomas, P.J., Qu, B.H., and Pedersen, P.L. (1995). Defective Protein-Folding as a Basis of Human-Disease. Trends in Biochemical Sciences 20, 456-459.*

6. Publication of Results

- *Gurushidze M, G Hensel, S Hiekel, S Schedel, V Valkov and J Kumlehn (2014) True-breeding targeted gene knock-out in barley using designer TALE-nuclease in haploid cells. PLoS One 9 (3), e92046*
- *Karimi-Ashtiyani R, T Ishii, M Niessen, N Stein, S Heckmann, M Gurushidze, AM Banaei-Moghaddam, J Fuchs, V Schubert, K Koch, O Weiß, D Demidov, K Schmidt, J Kumlehn and A Houben (2015) Point mutation impairs centromeric CENH3 loading and induces haploid plants. PNAS 112, 11211-11216*
- *Budhagatapalli N, T Rutten, M Gurushidze, J Kumlehn and G Hensel (2015) Targeted modification of gene function exploiting homology-directed repair of TALEN-mediated double strand breaks in barley. G3: Genes|Genomes|Genetics 5, 1857-1863*
- *Ishii, T., Karimi-Ashtiyani, R., Banaei-Moghaddam, A.M., Schubert, V., Fuchs, J., and Houben, A. (2015). The differential loading of two barley CENH3 variants into distinct centromeric substructures is cell type- and development-specific. Chromosome Research DOI: 10.1007/s10577-015-9466-8.*

Patent

- *Houben A, R Karimi-Ashiyani, T Ishii, N Stein and J Kumlehn (2014) Generation of haploid inducer lines via mutation of a single amino acid of CENH3. EP 14182719.6, filed*

HYWHEAT Genomics- and metabolomics based prediction of hybrid performance in wheat*

C. F. Longin[1], Y. Zhao[2], H.-P. Mock[3], A. Matros[3], E. Ebmeyer[4], R. Schachschneider[5], E. Kazman[6], J. Schacht[7], M. Gowda[2], F. Mette[2], J.C. Reif[2]

1 State Plant Breeding Institute, University of Hohenheim, 70593 Stuttgart, Germany
2 Department of Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstraße 3, 06466 Stadt Seeland, Germany
3 Department of Physiology and Cell Biology, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Corrensstraße 3, 06466 Stadt Seeland, Germany
4 KWS LOCHOW GMBH, 29296 Bergen, Germany
5 Nordsaat Saatzuchtgesellschaft mbH, 38895 Langenstein, Germany
6 Lantmännen SW Seed Hadmersleben GmbH, 39398 Hadmersleben, Germany
7 Limagrain GmbH, 31226 Peine-Rosenthal, Germany

* This article contains excerpts of the habilitation thesis "Prospects and constraints for an efficient hybrid wheat breeding" of C.F. Longin. The copyright is not transferred and remains with the authors.

Introduction

Hybrid breeding is very successfully performed in many allogamous crops such as maize, sorghum, sugar beet, sunflower, and rye (Coors and Pandey 1999). The main advantages of hybrid compared to line varieties are increased performance for important agronomic traits due to heterosis (Shull 1908), higher yield stability especially in marginal environments (Hallauer *et al.* 1988), and larger return of investment for seed companies due to the built-in plant variety protection by inbreeding depression (Edwards 2001). Hybrid breeding for autogamous cereals like wheat was less successful yet. Currently, less than one per cent of the total world wheat area is planted with hybrids.

The decision to start a hybrid breeding program depends on the long-term competitiveness of hybrid compared to line breeding (Oettler *et al.* 2005). A stable yield surplus is required to justify the higher seed production costs of hybrids. The yield surplus is a function of the extent of heterosis and of the selection gain in hybrid *versus* line breeding. Nevertheless, only very few studies were available for the long-term competitiveness of hybrid wheat breeding and most of them used only small numbers of hybrids tested in few locations. Thus, more intensive research efforts were necessary to investigate the potential of hybrid breeding in wheat. This motivated us to initiate the HYWHEAT project in order to determine the extent of heterosis and the genetic variance for hybrid breeding as compared to line breeding based on large scale field trials involving high number of hybrids.

A high number of parental lines are available in hybrid breeding and, therefore, not all potential hybrid crosses can be produced and evaluated in field trials. Thus, efficient methods are required to predict the hybrid performance of untested hybrids. A further main objective of the HYWHEAT project was therefore to elaborate new predictions methods exploiting recent advances in genomics and metabolomics and compare their efficiency with traditional methods for hybrid prediction.

Heterosis and yield stability

Robust estimation of variance components requires large field trial series in order to separate genetic from environmental effects and from the genotype × environment interaction. To estimate heterosis, the experiments further need to include not only the hybrids, but also their parental lines as well as a set of important commercial varieties as benchmark. For wheat, the most important trait is grain yield, which is largely influenced by environmental effects. Thus, an adequate experimental study for grain yield has to be based on field trials at more than 3 locations with net field plot size larger than 5 m². Owing to hybrid seed production problems in wheat, nearly all of the studies reported before the start of the HYWHEAT project do not fulfill these requirements (*cf.* Longin *et al.* 2012).

The HYWHEAT consortium conducted the largest ever published field experiment for hybrid wheat. In total 135 parental lines were provided from the four participating breeding companies. The 120 female lines were crossed with 15 male lines using a factorial mating design and 1,604 out of the 1,800 potential single-cross hybrid combinations could be produced with sufficient amount of seeds. These 1,604 hybrids were evaluated in field trials together with their

Table 1. Traits measured in HYWHEAT

Traits	Unit	No. Locations
Grain yield	ton ha^{-1}	11
Heading time	Days in year	7
Plant height	cm	10
Frost susceptibility	1 (low)- 9 (high)	3
Lodging susceptibility	1 (low)- 9 (high)	3
Yellow rust susceptibility	1 (low)- 9 (high)	5
Leaf rust susceptibility	%	6
Powdery mildew susceptibility	1 (low)- 9 (high)	6
Fusarium head blight susceptibility	1 (low)- 9 (high)	7
Septoria tritici blotch susceptibility	1 (low)- 9 (high)	6
Protein content	%	6
Sedimentation volume	ml	5
Gluten content	%	2
Starch content	%	2
Hardness	%	3
Thousand kernel weight	gr	4
Test weight	kg hl-1	3

135 parental lines as well as ten commercial varieties as checks. Ten agronomic and disease resistance traits as well as seven quality traits were recorded across the different environments in the years 2012 and 2013 (Table 1). Thus, the HYWHEAT data set represent a unique basis for trustworthy conclusions regarding hybrid wheat delivering high impact research data to the scientific community.

For the estimation of the market potential of wheat hybrids, the difference between hybrids and the best line varieties, denoted as commercial heterosis, is of main interest. Especially in wheat, where hybrid seed production is complicated and, thus, expensive, a considerable yield advantage is necessary for hybrids. In our large experimental study, 620 out of the 1,604 hybrids had a larger grain yield than the best line variety "*Tobak*", which yielded across 11 locations 10.8 tons per ha (Longin *et al.* 2013; Fig. 1). Thereby, the best hybrid outyielded "*Tobak*" by 1 ton per ha (Figure 1). This advantage of hybrids was surprisingly high, taking into account that "*Tobak*" is an extraordinary high yielding line variety, which has a considerably higher yield than all other line varieties currently on the market. For instance, the second best commercial check was the widely grown variety "*Julius*", which had a grain yield of 10.2 tons per ha underlining the high yield potential of some of the wheat hybrids.

Besides high grain yield, baking quality is of crucial importance for wheat varieties with protein content representing one crucial indirect predictor. In wheat, protein content is highly negatively correlated with grain yield, which was confirmed in our study also for hybrids (Liu *et al.* 2016). In line with this observation, we determined for protein content a negative midparent and better-parent heterosis of - 2 % and - 4 %, respectively. Nevertheless, this amount of negative heterosis seems surprisingly low compared with the high positive midparent heterosis of 10% for grain yield. Consequently, it seems feasible to select for hybrids with outstanding yield having acceptable quality parameters. This was confirmed by grouping the 1,604 hybrids into quality classes based on protein content and sedimentation volume (Figure 1). Although 874 hybrids had a C quality, 683 hybrids were identified with B quality and even 47 hybrids with A quality. None of the hybrids had an E quality, which can be explained by the use of parental lines with mainly A, B and C quality. Although the groups varied in size, we determined a commercial heterosis of more than one ton per ha for all quality groups.

Besides heterosis, Duvick *et al.* (2004) and Mühleisen *et al.* (2013) reported that heterosis seems to be larger under stress environments. Furthermore, yield stability is reported to be higher for hybrids than for lines (Hallauer *et al.* 1988). In contrast to homozygous line varieties, hybrids are heterozygous and might therefore be more robust against environmental effects. Comparing the group of hybrids to

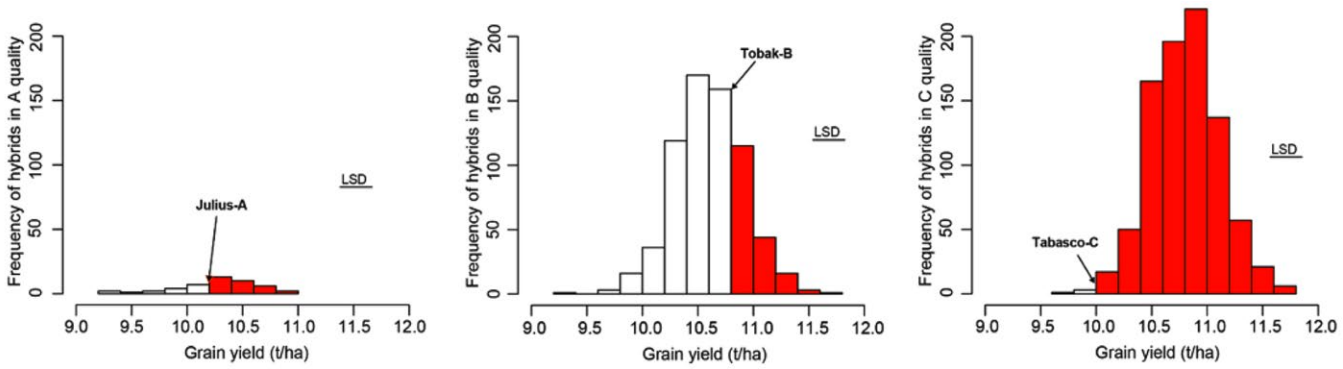


Figure 1. Phenotypic distribution for grain yield of the 1,604 hybrids grouped according to their values for protein content and sedimentation volume in different German quality classes (A, B, C). Hybrids exceeding the best check of the respective quality class are marked in red (LSD = least significant difference at the 5% probability level).

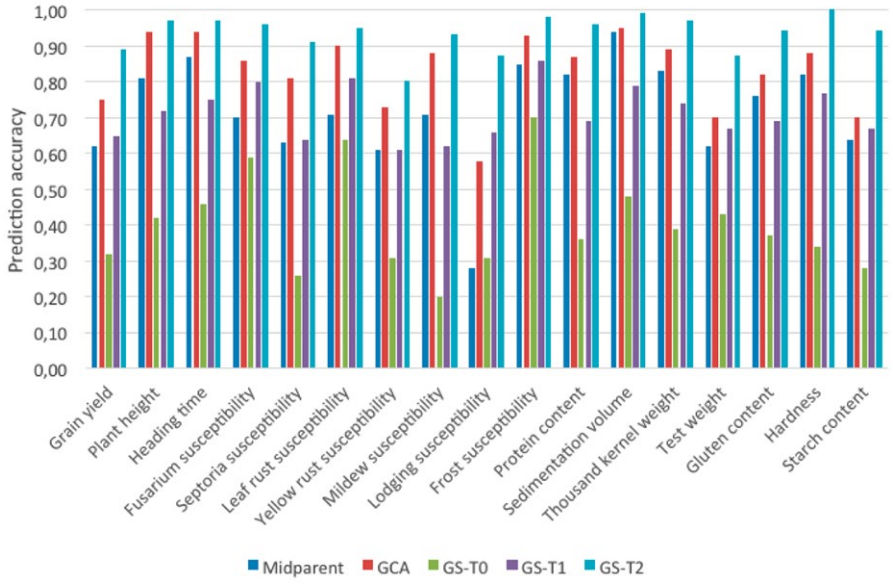


Figure 2. Accuracy to predict the hybrid performance based on the mean of the line per se performance (Midparent) or the mean GCA effects of the parents (GCA) or based on genomic selection using ridge regression best linear unbiased prediction determined with the 90,000 SNP Illumina marker chip (Würschum et al. 2013). The test set includes hybrids that had both parents (GS-T2), one parent (GS-T1), or no parent (GS-T0) in common with hybrids of the corresponding estimation set.

the group of parental lines, we determined a clearly smaller variance due to genotype × environment interaction for the wheat hybrids (Mühleisen *et al.* 2014). This could also be confirmed for hybrid barley and hybrid triticale and was in line with previous findings for wheat (*cf.* Borghi *et al.* 1994; Gowda *et al.* 2012). Consequently, hybrids seem to be more yield stable than line varieties and heterosis seems to be larger in stress environments. These findings represent important advantages of hybrid wheat to feed the growing world population even under the effects of the climate change. The consequences of the estimated quantitative genetic parameters on the design of hybrid breeding programs were then discussed in comprehensive studies (Longin *et al.* 2014a, b). In conclusion, the large phenotypic data produced in the HYWHEAT project allowed for the first time at all in wheat a robust characterization of the potential of hybrid wheat breeding showing that previous small studies has underestimated its potential.

Prediction of hybrid performance

In principle, all new lines bred in each heterotic group can serve as new parental lines and could be crossed to all existing lines from the other heterotic group(s). Thus, millions of potential single-cross combinations are feasible, whereby the challenge for the breeder is to efficiently select the most promising one (Bernardo 1994). As it is impossible to test all these combinations, prediction of hybrid performance prior to testing them is crucial and several approaches have been proposed in the literature. Briefly, they can be classified in methods based on phenotypic data and methods using molecular markers (for review, cf. Zhao *et al.* 2015a).

Prediction of hybrid performance based on phenotypic data requires field data from the parental lines, which shall be combined in the new hybrid. This can either be their line per se performance or their GCA effects. The accuracy of this prediction can be estimated by the correlation between midparent and hybrid performance *r(MP, F1)* or the mean of the GCA effects and the hybrid performance *r(GCA, F1)*. For all 17 traits investigated in our large experimental study, *r(GCA, F1)* was larger than *r(MP, F1)* (Longin *et al.* 2013; Liu *et al.* 2016; Figure 2). The traits plant height, heading time, frost susceptibility, protein content, sedimentation volume, and thousand kernel weights could be predicted with high accuracies with values above 0.8. For the remaining traits, phenotypic prediction accuracies were still substantial with values of *r(MP, F1)* = 0.62 and *r(GCA, F1)* = 0.75 observed for the most complex trait grain yield.

Alternatively, prediction of hybrid performance can be performed using marker-assisted selection. For instance, few functional markers with large effects are currently applied in wheat breeding for plant height, heading time, frost tolerance, disease resistance, and quality. Although important to improve the line per se performance, we failed in most instances to accurately predict hybrid performance based on these single markers only (Zhao *et al.* 2013; Zhao *et al.* 2014; Gowda *et al.* 2014; Miedaner *et al.* 2013; Mirdita *et al.* 2015).

Table 2. Accuracy to predict hybrids based on genomic and metabolomic selection, and the combination of both methods. The test set includes hybrids that had both parents (T2), one parent (T1), or no parent (T0) in common with hybrids of the corresponding estimation set.

Marker systems	Scenario	Accuracy
90k SNP Chip	T0	0.32
90k SNP Chip	T1	0.65
90k SNP Chip	T2	0.89
34 metabolites	T0	0.15
34 metabolites	T1	0.42
34 metabolites	T2	0.74
90k SNP Chip + 34 metabolites	T0	0.31

90k SNP Chip + 34 metabolites	T1	0.65
90k SNP Chip + 34 metabolites	T2	0.90

Genome-wide selection was proposed as promising method for hybrid prediction (Piepho 2009; Zhao *et al.* 2012; Technow *et al.* 2014). Thereby, the use of ridge regression best linear unbiased prediction (RR-BLUP) seems to be a robust method with low computing time demand (cf. Piepho *et al.* 2012; Zhao et al 2015a). The prediction accuracy of genomic selection varied widely across the 17 traits (Zhao *et al.* 2013; Zhao *et al.* 2014; Gowda *et al.* 2014; Miedaner *et al.* 2013; Mirdita *et al.* 2015; Liu *et al.* 2016; Figure 2). For all traits, the highest prediction accuracy was determined for the scenario, where both parental lines were present in hybrid combinations in the test and estimation set (GS-T2, Figure 2). This accuracy dropped rapidly when only one (GS-T1) or no parental line (GS-T0) was present in hybrid combinations in the test and estimation set. For instance, we identified for grain yield prediction accuracies based on the 90,000 SNP Illumina marker chip of 0.32, 0.65 and 0.89 for GS-T0, GS-T1 and GS-T2, respectively. This clearly underlines the central role of relatedness between estimation and test sets as driver of the accuracies of genome-wide selection.

Table 2. Accuracy to predict hybrids based on genomic and metabolomic selection, and the combination of both methods. The test set includes hybrids that had both parents (T2), one parent (T1), or no parent (T0) in common with hybrids of the corresponding estimation set.

Alternatively to genomic selection, metabolomic profiles might be used to predict hybrid performance. We used 34 metabolites measured in the flag leaves sampled from three field locations to predict the hybrid performance. The prediction accuracy based on this untargeted metabolomics was considerably lower than prediction accuracies achieved by genomic selection (Table 2; Zhao *et al.* 2015b). Furthermore, the combination of genomic and metabolomic profiles for hybrid prediction could not improve the prediction accuracy from genomic selection. Taken into account the high costs and logistical requirements for metabolomic profiling, the use of this technique for hybrid prediction in wheat cannot be recommended.

Summarizing, the HYWHEAT project resulted in robust estimates of heterosis, yield stability and variance components. Our findings underlined the potential of hybrid breeding also for wheat. Based on the comprehensive mapping population, we showed that genomic-selection is a valuable tool to accelerate selection gain in hybrid wheat breeding.

Literature cited

(publications of the HYWHEAT project marked in italics)

- *Bernardo, R.. 1994. Prediction of maize single-cross performance using RFLPs and information from related hybrids. Crop Science 34: 20-25*
- *Borghi, B., and M. Perenzin. 1994. Diallel analysis to predict heterosis and combining ability for grain yield, yield components and bread-making quality in bread wheat (T.aestivum). Theoretical and Applied Genetics 89: 975–981*
- *Coors, J.G., and S. Pandey. 1999. The genetics and exploitation of heterosis in crops. ASA, CSSA, and SSSA, Madison, WI*
- *Duvick, D.N., J.S.C. Smith and M. Cooper. 2004. Long-term selection in a commercial hybrid maize breeding program. In: Janick J. (ed) Plant breeding reviews, Vol. 24, part 2. Long term selection: crops, animals, and bacteria. Wiley, New York, pp 109–151*
- *Edwards, I.B.. 2001. Origin of cultivated wheat. In: Bonjean A.P., Angus W.J. (eds), The world wheat book-a history of wheat breeding. Vol.1, Lavoisier publishing, Paris, pp 1019–1045*
- *Gowda, M., C.F.H. Longin, V. Lein, and J.C. Reif. 2012. Relevance of specific versus general combining ability in winter wheat. Crop Science 52: 2494-2500*
- *Gowda, M., Y. Zhao, T. Würschum, C.F.H. Longin, T. Miedaner, E. Ebmeyer, R. Schachschneider, E. Kazman, J. Schacht, J.-P. Martinant, M.F. Mette, and J.C. Reif. 2014. Relatedness severely impacts accuracy of marker-assisted selection for disease resistance in hybrid wheat. Heredity 112: 552-561*
- *Hallauer, A.R., W.A. Russell and K.R. Lamkey. 1988. Corn breeding. In: Sprague, G.F., Dudley J.W. (eds) Corn and corn improvement, 3rd edn. Agron Monogr 18 ASA, CSSA, SSSA, Madison, WI, pp 469–565*
- *Liu, G., Y. Zhao, M. Gowda, C.F.H. Longin, J.C. Reif, and M.F. Mette. 2016. Predicting hybrid performances for quality traits through genomic-assisted approaches in Central European wheat. Theoretical and Applied Genetics In review.*
- *Longin, C.F.H., J. Mühleisen, H.P. Maurer, H. Zhang, M. Gowda, and J.C. Reif. 2012. Hybrid breeding in autogamous cereals. Theoretical and Applied Genetics 125: 1087-1096*
- *Longin, C.F.H., M. Gowda, J. Mühleisen, E. Ebmeyer, E. Kazman, R. Schachschneider, J. Schacht, Y. Zhao, and J.C. Reif. 2013. Hybrid wheat: quantitative genetic parameters and consequences for the design of breeding programs. Theoretical and Applied Genetics 126: 2791-2801*
- *Longin C.F.H., J.C. Reif, and T. Würschum (2014a) Long-term perspective of hybrid versus line breeding in wheat based on quantitative genetic theory. Theoretical and Applied Genetics 127: 1635-1641.*
- *Longin C.F.H., X. Mi, A.E. Melchinger, J.C. Reif, and T. Würschum (2014b) Optimum allocation of test resources and comparison of breeding strategies for hybrid wheat. Theoretical and Applied Genetics 127: 2117-2126.*
- *Miedaner, T., Y. Zhao, M. Gowda, C.F.H. Longin, V. Korzun, E. Ebmeyer, E. Kazman, and J.C. Reif. 2013. Genetic architecture of resistance to Septoria tritici blotch in European wheat. BMC Genomics 14: 858*
- *Mirdita V., G. Liu, Y. Zhao, T. Miedaner, C.F.H. Longin, M. Gowda, M.F. Mette, J.C. Reif. 2015. Genetic architecture is more complex for resistance to Septoria tritici blotch than to Fusarium head blight in Central European winter wheat. BMC Genomics 16: 430*

HYWHEAT Genomics- and metabolomics based prediction of hybrid performance in wheat

- *Mühleisen, J., H.P. Maurer, G. Stiewe, P. Bury, and J.C. Reif. 2013. Hybrid breeding in barley. Crop Science 53: 819-824*
- *Mühleisen, J., H.-P. Piepho, H.P. Maurer, C.F.H. Longin, and J.C. Reif. 2014. Yield stability of hybrids versus lines in wheat, barley and triticale, Theoretical and Applied Genetics 127: 309-316*
- *Oettler, G., S.H. Tams, H.F. Utz, E. Bauer, and A.E. Melchinger. 2005. Prospects for hybrid breeding in winter triticale: I. Heterosis and combining ability for agronomic traits in European elite germplasm. Crop Science 45: 1476–1482*
- *Piepho, H.P.. 2009. Ridge regression and extensions for genomewide selection in maize. Crop Science 49: 1165-1176*
- *Piepho, H., J. Ogutu, T. Schulz-Streeck, B. Estaghvirou, A. Gordillo, and F. Technow. 2012. Efficient computation of ridge-regression best linear unbiased prediction in genomic selection in plant breeding. Crop Science 52: 1093-1104*
- *Shull, G.H.. 1908. The composition of a field of maize. American Breeders Association Reports 4: 296–301*
- *Technow, F., T.A. Schrag, W. Schipprack, E. Bauer, H. Simianer, A.E. Melchinger. 2014. Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. Genetics 197: 1343-1355*
- *Würschum T., S.M. Langer, C.F.H. Longin, V. Korzun, E. Akhunov, E. Ebmeyer, R. Schachschneider, J. Schacht, E. Kazman, and J.C. Reif. 2013. Population structure, genetic diversity and linkage disequilibrium in elite winter wheat assessed with SNP and SSR markers. Theoretical and Applied Genetics 126: 1477-1486*
- *Zhao, Y., M. Gowda, W. Liu, T. Würschum, H.P. Maurer, C.F.H. Longin, N. Ranc, and J.C. Reif. 2012. Accuracy of genomic selection in European maize elite breeding populations. Theoretical and Applied Genetics 124: 769-776*
- *Zhao, Y., M. Gowda, T. Würschum, C.F.H. Longin, V. Korzun, S. Kollers, R. Schachschneider, J. Zeng, J.Z.R. Fernando, J. Dubcovsky, and J.C. Reif. 2013. Genetic architecture of frost tolerance in Central European winter wheat. Journal of Experimental Botany 64: 4453-60*
- *Zhao, Y., F. M. Mette, M. Gowda, C. F. H. Longin, and J.C. Reif. 2014. Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112: 638-645*
- *Zhao, Y., M.F. Mette, and J.C. Reif. 2015a. Genomic selection in hybrid breeding. Plant breeding, 134: 1-10*
- *Zhao, Y., Z. Liu, G. Liu, Y. Jiang, H.P. Maurer, T. Würschum, H.-P. Mock, A. Matros, E. Ebmeyer, R. Schachschneider, E. Kazman, J. Schacht, M. Gowda, C.F.H. Longin and J.C. Reif. 2015b. Genome-based establishment of a high-yielding heterotic pattern for hybrid wheat breeding. PNAS, doi:10.1073/pnas.1514547112*

INNO GRAIN-MALT Developing drought sustainable genotypes with improved seed yield and quality suitable for malting

Andriy Kochevenko[1], Christiane Seiler[1], Korana Surdonja[1], Malthe Schmidt[2], Burkhard Schinkel[2], Jörg Großer[2], Sonja Kollers[2], Nese Sreenivasulu[1,3], Viktor Korzun[2] and Andreas Graner[1]

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany
2 KWS LOCHOW GMBH, Bergen, Germany, Project coordinator
3 International Rice Research Institute, Los Banos, Laguna, Philippines 4031

Introduction

Barley malting quality is a complex of multiple, polygenically inherited traits, many of which are highly influenced by the environments. We have studied the molecular mechanisms which control grain yield and seed quality, in order to better understand how to maintain grain yield and improve malting quality from barley grown under stress conditions. In particular, we were interested to reveal the regulatory factors and structural genes affecting yield stability, seed quality and malting quality under both control and drought-stress conditions. Furthermore the potential of genomic selection for twelve malting quality traits in spring and winter barley breeding programs was evaluated.

Genetics of Malting Quality

For QTL analysis of yield components and malting quality, a double-haploid population was developed by crossing the two elite lines „Victoriana"(LP106) and „Sofiara"(LP104). Under terminal drought conditions „Victoriana" displays a „stay green" phenotype, whereas the mechanism observed „Sofiara" is based on remobilization of carbohydrates from stem and leaves towards the developing seed. A total of 100 DH lines were grown in replicated field plots in two distinct locations, Wohlde (North Germany) and Gatersleben (East Germany), where an additional terminal drought stress experiment was performed by means of rain-out shelter. Mature seeds were harvested and analysed for yield, thousand grain weight (TGW) and grain morphological traits such as seed length (SL), seed width (SW), and shape. Mature seed material (80 g) was micro-malted (VLB Berlin) and a series of malting quality parameters, including protein and beta glucan content (BGL), soluble nitrogen (SNI), friability (FRI), kolbach index (KOL), viscosity (VIS), extract yield (FGE), moisture (MMC), final attenuation (FLA), respiration and rootlet loss were determined.

For QTL mapping, a high density genetic map based on 1810 polymorphic SNP markers, was developed based on the Illumina Chip technology. QTL analysis resulted in 105 QTLs with additive main effects and/or additive x environment interaction effects (QE) for 25 analysed traits. These QTLs were in particular associated with yield-related traits (57), malting quality parameters (41), seed quality traits (5) and disease resistance (2) (Fig.1). Variation explained by main effects (QTLs) ranged from 0.41 to 68.77%, e.g. 10 QTLs explained more than 10% of the variance, which was substantially higher than that of QE interaction effects. Six (trait associated) QTLs VIS, SNI, KOL, FRI, BGL and powdery mildew (ME) had significant environmental interaction besides significant genetic main effects, indicating the significance of environmental factors. 18 QTLs without significant main effect revealed QE effects and include QTL for BGL, FGE, plant height (PH), seed area (SA), seed nitrogen content (SNC), SNI, thousand grain weight (TGW) and yield (YLD).

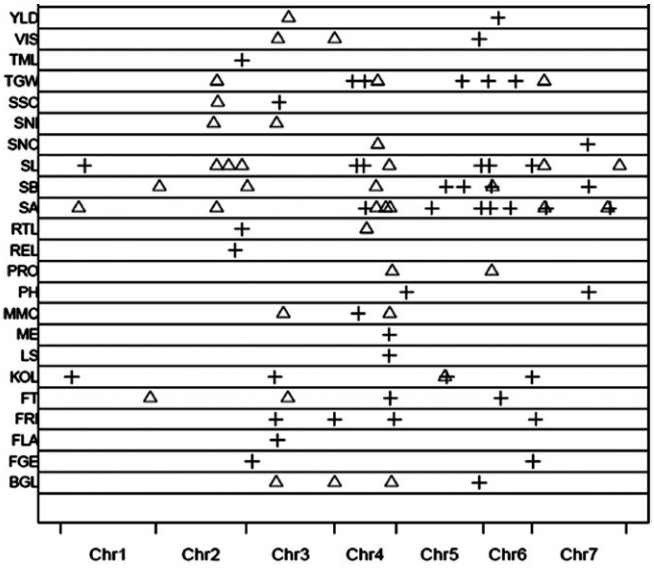


Fig. 1: Location of main-effect QTLs on chromosomes. Triangle and cross represent QTL effects contributed by Victoriana ("Δ" = LP106) and Sofiara ("+" = LP104). YLD- yield, VIS- viscosity, TML- total malting loss, TGW- thousand grain weight, SSC- seed starch content, SNI- soluble nitrogen, SNC- seed nitrogen content, SL- seed length, SW- seed width, SA- seed area, RTL- rootlet losses, REL- respiration losses, PRO- protein content, PH- plant height, MMC- malt moisture content, ME- powdery mildew, LS- leaf spots, KOL- kolbach index, FT- flowering time, FRI- friability, FLA- final attenuation, FGE- fine grind extract, BGL- ß-glucan.

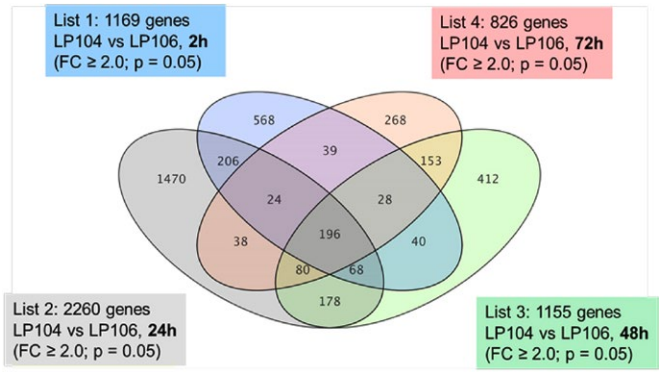


Fig. 2: Venn diagram showing differentially expressed genes between parental lines at four different stages of seed germination. A fold change of more than 2.0 at a significance level of p ≤ 0.05 was applied as selection criteria.

Seed area (14), seed width (8) and TGW (9) displayed the highest number of QTLs.

Among malting quality traits, the largest number of main effect QTLs were detected for KOL (5), which mapped to four chromosomes. Co-localization of QTLs for a number of traits represented a common feature and clusters of QTLs were observed on each chromosome except for 1H. The highest level of clustering occurred in the centromeric region of chromosome 3H and at the distal end of chromosome 4H. The cluster on chromosome 3H was mainly determined by malting quality traits. At this location, five major QTL for BGL, FRI, KOL, SNI and VIS (R2>10%) were co-localized; while the cluster on 4H contained QTLs for BGL, FRI and PRO together with several QTLs for disease resistance and morphological traits. Clusters of QTLs affecting yield-related traits such as TGW, SL, SW and SA were observed on chromosomes 2H, 4H, 5H, 6H and 7H. The cluster present on chromosome 2H was also associated with QTLs for seed starch content (SSC) and SNI, while the one on chromosome 5H involved QTLs influencing BGL and VIS.

Transcriptome profiles of micro-malted samples (72h after onset of imbibition) of double-haploid lines grown in Gatersleben and Wohlde were carried out using a 60K Agilent barley array. The normalised and filtered datasets comprised 25.055, 25.127 and 23.760 transcripts and were subsequently scrutinized for expression QTL (eQTL).

Malting quality is determined by processes that take place both during seed development and seed germination. To identify candidate genes underlying phenotypic QTL (pQTL) for malting quality in barley we carried out a germination time-course transcriptome analysis using seed material of parental lines from Gatersleben control conditions (eight stages: 0, 2, 24, 48, 72, 96, 120 and 144 hours after imbibition(hai)). The normalised data were analysed for more than 2-fold differences in transcript level between parental lines LP104 vs LP106. The Venn diagram (Fig. 2) demonstrates relations between four of these sets, representing the four most transcriptionally active stages at 2, 24, 48 and 72 hai. In total, 762 differentially expressed transcripts were analysed yielding 1826 eQTLs. Large effect eQTLs comprise genes involved in β-glucan metabolism, transcriptional regulation, protein degradation, membrane trafficking and polyamine biosynthesis. 316 eQTLs had a LOD score of more than 10, the level which is considered as a threshold for prediction of cis-eQTLs.

Transcripts stored during seed development and maturation are germination specific, and differential expression under stress could possibly affect germination and malting quality. To identify those transcripts, which are *de novo* translated during germination, polysomes were isolated from germinating seeds which were dissected into embryo and endosperm fractions. Sucrose gradients were used to separate ribosome complexes and the absorbance profile was monitored to distinguish between polysomal and non-polysomal fractions. Together with total RNA of mature seeds (0h) all samples have been subjected to transcriptome analysis. The transcriptomes of the dry seed and the micro-malted seedling (72h) share approximately 20.000 transcripts, which account for almost two thirds of all micro-malted transcripts. Nearly 15.000 genes are uniquely found in the polysomal fraction of micro-malted seedlings which are presumed to be newly synthesized and translated transcripts.

Mechanisms of terminal drought tolerance

Occurrence of drought stress during anthesis and onset of seed development is a very critical factor for yield, resulting in impaired seed set and reduced grain weight due to a decrease in photosynthetic efficiency and altered remobilization processes. Abscisic acid (ABA) is known to accumulate upon drought stress and there is a negative linear relationship between seed set and increase in ABA in reproductive tissue. ABA induces maturation leading to a shorter grain filling period. Hence, to potentially achieve optimum tolerance, it is important to achieve a stay green phenotype by reducing the elevated ABA to basal levels during this developmental period. Another strategy to achieve terminal drought tolerance is an improved remobilization of stem reserves to developing seeds. Although seed yield could be improved to a certain extent with remobilization phenotypes under drought, the resulting increase in seed nitrogen might not be beneficial for malting. Hence, improving assimilation during those critical stages is key to increasing drought tolerance in malting barley.

Transgenic stay-green lines in “Golden Promise” background with a lowered ABA content in developing seeds under terminal drought were crossed into the elite variety “Barke”. The lines express either a gene of ABA biosynthesis (AtNCED6, construct “SN”) or signaling (AtRPK1, construct “SR”) under control of a stress-inducible promoter. After crossing, DH lines were produced and used for drought stress experiments under greenhouse conditions. Measu-

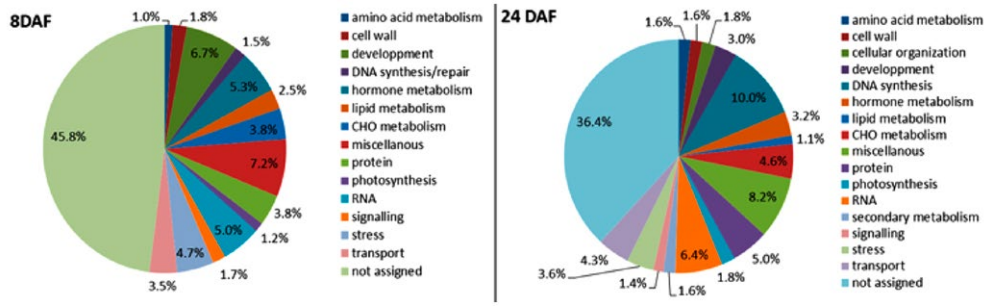


Fig. 3: Pie chart showing functional categories of differentially expressed genes and their percentage distribution. DAF - days after flowering.

red photosynthesis parameters showed that assimilation and stomatal conductance were reduced under stress in all lines but were higher when compared to Barke crosses. Yield was recorded in several drought stress experiments under greenhouse conditions. Generally, yield was reduced under stress, while TGW was less affected. Seed number per spike was not affected by drought, it even increased slightly under drought stress in the Barke x SN cross.

In one drought stress experiment samples from developing caryopsis (stages 8, 16, 20 and 24 days after flowering, DAF) have been collected from control and drought stressed plants (drought stress starting from 4 DAF until maturity) for further microarray analysis. For each developmental stage, differentially expressed genes (drought stress vs control) having a fold change of ≥ 2.0 in at least one of the analysed lines have been extracted and validated by a 2-way ANOVA. For 8, 16, 20 and 24 DAF, 600, 654, 452 and 439 differentially regulated genes were identified respectively. As an example, transcripts belonging to the category DNA synthesis/repair increased from 1.5% to 10% at 8 and 24 DAF, respectively and thus could be regarded as a potential response to long-term stress (Fig. 3). Around 5.3% of differentially regulated genes at 8 DAF are transcripts related to hormone metabolism, among those, the majority were up-regulated in Golden Promise (GP) and Barke but down-regulated in transgenic lines Barke x SN and SN x Barke.

Moreover, under drought-stress conditions, differentially regulated genes were identified in Barke, Barke x SN, SN x Barke and SN410, when compared to wild type Golden Promise (GP). For short- and long-term stress (8 and 24 DAF) 563 and 217 differentially regulated genes in at least one of the lines compared to GP were found, respectively. K-means cluster analysis revealed high expression of genes involved in transport processes and protein degradation in Barke and Barke transgenic lines but low expression in GP wild type. Putative sucrose synthase genes involved in CHO metabolism showed highest expression in Barke and Barke transgenics compared to GP or SN410. A similar cluster was identified at 24 DAF. Here, ABA-responsive genes, genes related to starch degradation, raffinose metabolism and glycolysis were found to be up-regulated in Barke and Barke transgenics as well as partially up-regulated in SN410 compared to GP. Two genes annotated as cysteine-proteinase were highly expressed in GP compared to transgenic lines and Barke probably indicating a faster senescence in GP wild type.

Tab. 1: Predictive abilities (PA) derived from a 5 by 5 cross validation with their standard deviation (sd) and sizes of the training sets (#TS) for the twelve analysed malting quality traits in spring- and winter barley. The traits are alpha and beta amylase (AMA, AMB), extract (EXT), free amino nitrogen (FAN_L), final attenuation (Fin_At), friability (FRI), beta glucan content (GLU), kolbach index (KOL), malting loss (LOSS), soluble nitrogen (NIT), protein content (PRT) and viscosity (VIS).

	Spring barley			Winter barley		
Trait	PA (mean)	PA (sd)	#TS	PA (mean)	PA (sd)	#TS
AMA	0.444	0.151	99	0.564	0.300	102
AMB	0.142	0.218	118	0.606	0.150	102
EXT	0.558	0.072	424	0.625	0.175	102
FAN_L	0.497	0.218	65	0.572	0.165	102
Fin_At	0.495	0.120	424	0.732	0.113	102
FRI	0.552	0.077	424	0.788	0.113	102
GLU	0.479	0.096	423	0.798	0.173	102
KOL	0.487	0.106	424	0.556	0.178	102
LOSS	0.419	0.101	322	0.652	0.165	102
NIT	0.583	0.074	424	0.551	0.150	102
PRT	0.521	0.064	424	0.399	0.158	102
VIS	0.449	0.111	424	0.744	0.116	102
mean	0.469	0.117		0.632	0.163	

Genomic selection of malting quality traits

Genomic selection has been applied to various plant species, mostly for yield or yield related traits including grain yield, thousand kernel weight and resistances against diseases. Until now, quality traits have not been in the scope for genomic selection, but instead were tracked by marker assisted selection. Despite the breeding material being highly related and a long history of breeding for quality traits, there is still considerable variation in those traits. This opens the opportunity to further improve those traits through the application of genomic selection.

In the frame of the INNO GRAIN-MALT project, the potential to apply genomic selection to twelve malting quality traits in breeding programs of spring and winter barley (*Hordeum vulgare* L.) was assessed. Phenotypic means were calculated over three and four years with three to five locations per year. Heritabilities for malting traits ranged from 0.50 to 0.98. Predictive abilities (PA), as derived from cross-validation, ranged from 0.14 to 0.58 for spring barley and 0.40 to 0.80 for winter barley (Tab. 1). Furthermore, it could be shown that (i) due to the genetic structure spring- and winter barley need to be analysed separately; (ii) a training set

of around 100 individuals can give sufficient PA and correlations to unobserved data (e.g. when phenotypes of one year was removed before analysis) in a breeding program with highly related material and (iii) an increased training set size leads to higher PA with a reduced standard deviation, showing the potential to improve the accuracy by increasing the training set size.

The results from this study clearly indicate the power of genomic selection to improve selection efficiency in malting barley breeding programs. Deployment of genomic selection in barley will help to reduce cost intensive phenotyping for quality traits, increase selection intensity and shorten breeding cycles.

References

- *Schmidt M., S. Kollers, A. Maasberg-Prelle, J. Großer, B. Schinkel, A. Tomerius, A. Graner, V. Korzun (2016) Prediction of malting quality traits in barley based on genome-wide marker data to assess the potential of genomic selection, Theor Appl Genet 129, pp 203-213, DOI 10.1007/s00122-015-2639-1*

PHENO VINES Development of an automated field phenotyping pipeline for grapevine breeding

Kicherer, Anna[1]; Reinhard, Töpfer[1]

1 Julius Kühn-Insitut, Institute for Grapevine Breeding Geilweilerhof, 76833 Siebeldingen

Project motivation

Pathogens being introduced in the 19th century almost destroyed viticulture in Europe at the time, making the use of plant protection unavoidable up to now. Grapevine is grown on 1% of the German agricultural area and depends on intense fungicide application. To improve sustainability a major objective in grapevine breeding is the development of fungus resistant grapevine cultivars. However, these developments of new cultivars are very time-consuming, typically taking 20 to 25 years. Marker assisted selection (MAS) has in recent years considerably increased the efficiency of resistance breeding. Markers have been developed for e.g. for downy and powdery mildew resistances and were used very successively to pyramid resistance loci. Further improvement of MAS methods in grapevine breeding will come along with the development of faster and more cost efficient genotyping techniques.

Complementary to genotyping techniques reliable methods for plant phenotyping need to be established permitting high quality, objectivity, and precision in phenotypic data recording.

High-throughput phenotyping applications need to be developed and established to further increase the efficiency of grapevine breeding through sensor assisted selection.

Nowadays a lot of different sensor technologies are available ranging from imaging to non imaging sensors. Imaging sensors range from visible light sensors (Red, Green, Blue (RGB) cameras), multispectral, hyperspectral, thermal and fluorescence cameras to three dimensional cameras and laser concepts. The level of standardisation on which phenotyping can be done, decreases from controlled environments (growth cambers, greenhouses) to the field level. Other than grapevine seedlings in the greenhouse, grapevine as a perennial plant definitively needs to be screened directly in the field. Field level phenotyping faces a variety of challenges like variable light conditions, the similarity of fore- and background and in viticulture the canopy hidden traits.

Image acquisition and analysis

Image acquisition:

The acquisition of phenotypic data in grapevine breeding is usually done directly in the field by visual estimations. In general OIV descriptors or the BBCH scale are applied to categorize the phenotypes into classes. The phenotyping is strongly limited by time, costs and the subjectivity of records, therefore only a certain number of genotypes, respectively vines can be screened during the grapevine breeding process. Due to that limitation, objectivity, automation and precision of phenotypic data evaluation is crucial in order to (a) reduce the consisting phenotyping bottleneck, (b) increase the efficiency of grapevine breeding, (c) enable further important genetic research and (d) assure improved vineyard management.

To ensure a high quality wine production balanced and stable yields are important and therefore play an important role in grapevine breeding. However, quantifying yield parameters on a single vine level is challenging, particularly when measurements need to be done on large samples ripening at the same time. Complex shapes and slight variations between genotypes make it difficult and very time-consuming. Objective manual screenings can be done on small samples but this method is rather vague especially when done by multiple persons or varying descriptive standards.

Project objectives

The general objective of the project PHENO VINES was to set up and implement a phenotyping pipeline for high-throughput field phenotyping, for the application in grapevine breeding based on yield parameters. In particularly the objectives of the project were the:

- set up of a carrier vehicle, an automated GPS based control unit with corresponding software and a sensor positioning system.
- use of cost-efficient and fast visible light sensors (RGB camera, monochrome cameras) in different environments (laboratory, greenhouse, field) and the set up of a sensor system to put on the carrier vehicle.
- definition of yield parameters detectable using RGB images and the image analysis of these parameters.

- acquisition and validation of reference data to assess the obtained sensor. Followed by the verification of the feasibility to record objective and precise phenotypic data of yield parameters using RGB images and automated image analysis.
- set up of an automated data acquisition in the field including a chain vehicle as carrier for the phenotyping platform, a sensor system, the geo-referenced data acquisition and the guarantee of an automated data handling in cooperation with an interdisciplinary team. A special aim was the establishment of the possibility to record phenotypic data on a single vine level for grapevine breeding purposes.
- interpretation an validation of the obtained phenotypic sensor data for grapevine breeding.

The aims of this project are an objective and precise determination of phenotypic parameters in the field to increase the sample size and furthermore reduce the assessment errors. These HT- phenotyping methods have the intention to provide reliable data that can also be analysed retrospective to increase the efficiency in grapevine breeding.

Fig.1: PHENObot consisting of a robotic platform, a multi-camera-system and a geo-information system.

Project successes

PHENObot

The PHENObot was developed as phenotyping platform to be used within the grapevine breeding program. It consists of an automated guided tracked vehicle system containing a control unit, a calibrated multi camera system (five cameras), a Real-Time-Kinematic GPS system and a computer for image data handling. The multi-camera-system consists of three monochrome (MC) cameras, one RGB camera, one NIR (near infrared) camera and a lighting unit (containing eight LED bars to enable adequate illumination for standardised image acquisition).

Yield parameters

Fruit shape and size are major factors determining quality and yield. Fruit shape is of bigger importance in table grape breeding (Wycislo *et al.* 2008) as the shape and moreover the uniformity of the whole cluster are crucial quality traits with great influence on consumer acceptability. Yield parameters important to grapevine breeding are the time of ripening, berry size, number of berries per cluster, cluster size and number of clusters per cane. One objective of this project was the determination of grapevine yield parameters and the evaluation of possibilities to use RGB images and image analysis to detect these parameters.

The yield parameters are shown in Table 1. The number of vines ins mainly influenced by viticultural practise and not a parameter that needs to be detected by image analysis. Shoots per vine and cluster per vine are both difficult to detect out of images during the growing season due to occlusion (see image acquisition), but can be detected better in early growing stages (number of inflorescences before flowering) or during winter (number of shoots). Berry and cluster size as well as berry number per cluster could be detected with two tools (Kicherer *et al.* 2014,Kicherer *et al.* 2013) developed for the laboratory (see image analysis). The berry size and colour (Kicherer *et al.* 2015,Roscher *et al.* 2014) could be detected out of images taken directly in the field as well as the vine balance (Kicherer et. al, unpublished).

Image acquisition: Depending on the desired application,

Table 1 Yield parameters in grapevine, influence factors and possibility to be detected by image analysis.

Yield parameters	influences genetic	viticultural	environmental	detectable by image analysis
vines per unit area		x		-
shoots per vine		x		-/+
clusters per shoot	x		x	-/+
berries per cluster	x		x	+
berry size/ berry weight	x		x	+
vine balance (generative/vegetative growth)	x	x	x	+

Abb.2: Challenges of image acquisition under field conditions. A quantity and quality of light (1=backlighting; 2, 3=shading through the canopy). B image acquisition at night using artificial light for standardizes images (1); images recorded at night showing the variation of the phenotypic yield parameters (2, 3). C similar background and occlusion through canopy or other grapes (1= January; 2= June; 3= September).

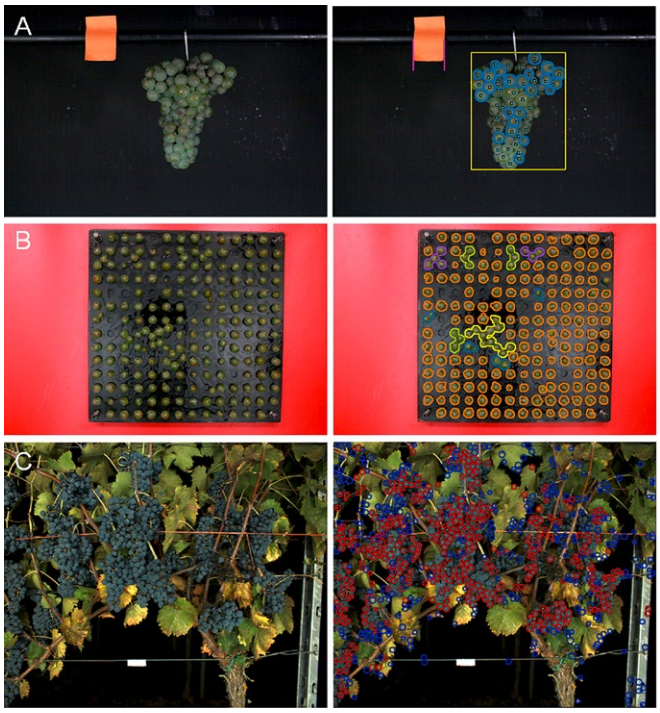


Abb.3: Image analysis tools, the original image and the result image in comparison. A Cluster Analysis Tool (CAT). B Berry Analysis Tool (BAT). C Berries In Vineyards tool (BIVcolor).

the sensors available for effective data acquisition can have different advantages and disadvantages. Using VIS sensors enables a fast and cost effective 2D (two dimensional) imaging, allowing to determine traits such as number and size based on geometric information, but also some radiometric information (monochrome, red green blue and near infrared). Visible light sensors have been used for image acquisition of yield parameters. A SLR (single lenses reflectance) camera in the laboratory (controlled environment) on the one hand, and on the other hand an industrial camera (multi-camera-system) on the PHENObot (under field conditions).

During this work we faced different challenges in assessing grapevine yield parameters especially under field conditions: (a) variation of the trait, (b) light environment, (c) similar background and (d) occlusion of traits, which will be discussed below.

(a) Depending on the genotype and the genotype-environment interaction phenotypic traits are highly variable. The variation of yield parameters can be very challenging with regard to the development of image analysis algorithms.
Yield parameters can vary in cluster size and architecture (Cubero *et al.* 2011,Kicherer *et al.* 2014), berry size (Kicherer *et al.* 2013,Tardaguila *et al.* 2012,Wycislo *et al.* 2008), shape and colour (Kicherer *et al.* 2015,Wycislo *et al.* 2008). Moreover, the colour of the surrounding area and the trait can be similar with regard to early development stages prior to varaison for coloured varieties or through the whole season for white varieties. Image analysis based only on single features like colour (Diago *et al.* 2012,Dunn and Martin 2004), shape(Rabatel and Guizard 2007) or texture (Grossetete *et al.* 2012) can be insufficient and a combination of these features is required (Nuske *et al.* 2014,Roscher *et al.* 2014).

(b) The acquisition of sensor data with image based phenotyping systems in the field can be influenced by the high variability of light quality and quantity. Permanently changing light conditions between two screenings or within the screening of one set of vines can cause problems. Challenges for the image analysis are backlighting or the even under good sun conditions the shading of the canopy. A solution can be the use of artificial light units and the image acquisition at night as equally shown by different studies (Font *et al.* 2014,Kicherer *et al.* 2015,Nuske *et al.* 2014).

(c) Fore- and background similarities of images captured in the field are particularly challenging for image analysis at an early development stage due to the missing canopy and repetition of the same objects in the background. This can either be overcome by using artificial backgrounds (Herzog *et al.* 2014,Herzog *et al.* 2014) or the application of computer vision based methods (Herzog *et al.* 2014,Klodt *et al.* 2015).

(d) Yield traits can be covered by leaves or other grapes negatively influencing the detected berries for instance. Approaches of fruit zone defoliation can enable a more accurate determination of the amount but on the other hand increase the background problematic at the same time.

Image analysis: Two high-throughput image interpretation tools for the laboratory were developed. Digital RGB images

were taken and machine vision algorithms using Matlab® were applied for image interpretation. Complementary to the RGB images detailed reference data was taken for the testing and evaluation of the image interpretation tools. The CAT (Cluster Analysing Tool) enables a fast acquisition of grape cluster size and berry size on a small scale. For more detailed information about the berry shape the BAT (Berry Analysing Tool) provides information about the berry number, size, and volume from RGB images. The image interpretation of one data set with an optional number of images runs automatically after starting the program.

Roscher *et al.* 2014 implemented the first automatic tool to detect the berry size out of images taken in the field (Berries In Vineyards = BIV). The characterization of multiple traits in a single pass is a crucial step to increase the sample throughput. Therefore this tool was extended to additionally assess the berry colour (BIVcolor) by Kicherer *et al.* (2015).

Proof of concept - application in grapevine breeding
A phenotyping pipeline has been set up for the implementation of high- throughput phenotyping in grapevine breeding. It covers the automated image acquisition directly in the field using the PHENObot, to data management, data analysis and the interpretation of obtained phenotypic data for grapevine breeding aims. The application of the phenotyping pipeline involved 2700 grapevines representing 970 accessions from the grapevine repository at the experimental vineyards of Julius Kühn-Institut Geilweilerhof.

Data acquisition: For the automated image acquisition directly in the field, the PHENObot was used. The precise GPS positions of individual vines has been surveyed and used for the path planning of the PHENObot. A software application (IggGeotagger.Ext) has been developed for the configuration and monitoring of the image acquisition process fulfilling two main tasks: the image acquisition task controls the cameras and the image transport and storage; the communication task handles the communication between the control unit of the PHENObot and the image acquisition computer. An image acquisition process performs different steps: The communication task waits for a message from the PHENObot control unit. As soon as the predefined position is reached it sends a specific message covering the position, orientation and the corresponding plant ID. After that the communication task starts the image acquisition task which triggers all cameras and the lightning unit, saves the images and generates the file names (containing the plant ID). Once the image acquisition task has finished, the communication task sends an acknowledgment message to the PHENObot control, signalling that it can move to the next vine.

Data management: All Images obtained by the IggGeotagger.Ext are stored in a database (IMAGEdata) which is linked to the *PLA* (Plant Location Administration) - A common management tool for experimental areas in the Julius Kühn-Institut. The geo-information of each grapevine and the associated plant ID is stored in this central database. Based on the image names, which contain the plant ID, every image is uniquely assigned to a single grapevine.

Data analysis: As phenotypic trait for the phenotyping pipeline the detection of berries and the determination of the berry size and colour was done using the Berries In Vineyards (BIVcolor) tool (Kicherer *et al.* 2015,Roscher *et al.* 2014).

Data acquisition took about 20 seconds per vine, which was followed by the automatic image analysis to extract objective and precise phenotypic traits. It was possible to capture images of 2700 vines within 12 hours using the PHENObot and subsequently automatic analysis of the images and extracting berry size and berry colour. This phenotypic evaluation in the grapevine repository has not been done to the same extend any time sooner. With this analysis proof of principle was demonstrated (Kicherer *et al.* 2015). The pilot pipeline provides the basis for further development of additional evaluation modules as well as the integration of other sensors.

References

- *Cubero, S., N. Aleixos, E. Moltó, J. Gómez-Sanchis, and J. Blasco (2011) Advances in Machine Vision Applications for Automatic Inspection and Quality Evaluation of Fruits and Vegetables. Food and Bioprocess Technology 4, 487-504. doi: 10.1007/s11947-010-0411-8.*
- *Diago, M.-P., C. Correa, B. Millán, P. Barreiro, C. Valero, and J. Tardaguila (2012) Grapevine Yield and Leaf Area Estimation Using Supervised Classification Methodology on RGB Images Taken under Field Conditions. Sensors 12, 16988-17006.*
- *Dunn, G.M. and S.R. Martin (2004) Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. Australian Journal of Grape and Wine Research 10 (3), 196-198.*
- *Font, D., T. Pallejà, M. Tresanchez, M. Teixidó, D. Martinez, J. Moreno, and J. Palacín (2014) Counting red grapes in vineyards by detecting specular spherical reflection peaks in RGB images obtained at night with artificial illumination. Computers and Electronics in Agriculture 108, 105-111. doi: http://dx.doi.org/10.1016/j.compag.2014.07.006.*
- *Grossetete, M., Y. Berthoumieu, J.P. Da Costa, C. Germain, O. Lavialle, and G. Grenier (2012) Early estimation of vineyard yield: site specific counting of berries by using a smartphone. Proceedings of the Infomation Technology, Automation and Precision Farming. International Conference of Agricultural Engineering - CIGR-AgEng, Valencia, Spain pp. 6 pp.*

- *Herzog, K., A. Kicherer, and R. Töpfer, „Objective Phenotyping of Bud Burst by Analyzing Grapevine Field Images", in 11th International Conference on Grapevine Breeding and Genetics. (International Society for Horticultural Science (ISHS), Beijing, China, 2014), pp. 379-388.*
- *Herzog, K., R. Roscher, M. Wieland, A. Kicherer, T. Läbe, W. Förstner, H. Kuhlmann, and R. Töpfer (2014) Initial steps for high-throughput phenotyping in vineyards. Vitis 53, 1-8.*
- *Kicherer, A., K. Herzog, M. Pflanz, M. Wieland, P. Rüger, S. Kecke, H. Kuhlmann, and R. Töpfer (2015) An automated phenotyping pipeline for application in grapevine research. Sensors 15 (3), 4823-4836.*
- *Kicherer, A., R. Roscher, K. Herzog, W. Förstner, and R. Töpfer, „ Image based Evaluation for the Detection of Cluster Parameters in Grapevine", in 11th International Conference on Grapevine Breeding and Genetics (International Society for Horticultural Science (ISHS), Beijing, China, 2014), pp. 335-340.*
- *Kicherer, A., R. Roscher, K. Herzog, S. Šimon, W. Förstner, and R. Töpfer (2013) BAT (Berry Analysis Tool): A high-throughput image interpretation tool to acquire the number, diameter, and volume of grapevine berries. Vitis 52 (3), 129-135.*
- *Klodt, M., K. Herzog, R. Töpfer, and D. Cremers (2015) Field Phenotyping of Grapevine Growth Using Dense Stereo Reconstruction. BMC Bioinformatics, 16:143 DOI:110.1186/s12859-12015-10560-x. doi: DOI:10.1186/s12859-015-0560-x.*
- *Nuske, S., K. Wilshusen, S. Achar, L. Yoder, S. Narasimhan, and S. Singh (2014) Automated Visual Yield Estimation in Vineyards. Journal of Field Robotics 31, 837-860. doi: 10.1002/rob.21541.*
- *Rabatel, G. and C. Guizard (2007) Grape berry calibration by computer vision using elliptical model fitting. Proceedings of the 6th European Conference on Precision Agriculture ECPA, Skiathos , Greece pp. 581-587.*
- *Roscher, R., K. Herzog, A. Kunkel, A. Kicherer, R. Töpfer, and W. Förstner (2014) Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields. Computers and Electronics in Agriculture 100, 148-158. doi: http://dx.doi.org/10.1016/j.compag.2013.11.008.*
- *Tardaguila, J., M.P. Diago, J. Blasco, B. Millán, S. Cubero, O.L. García-Navarrete, and N. Aleixos, „Automatic estimation of the size and weight of grapevine berries by image analysis", in International Conference of Agricultural Engineering (Valencia, Spain, 2012).*
- *Wycislo, A.P., J.R. Clark, and D.E. Karcher (2008) Fruit shape analysis of *Vitis* using digital photography. HortScience 43 (3), 677-680.*

POP MASS Development and use of novel gene technologies to increase biomass yield in the woody perennial Populus spec.

Fladung, M.[1], Brügmann, T.[1], Becker, D.[2], Nietsch J.[1]

1 Thünen Institute for Forest Genetics, Sieker Landstr. 2, D-22927 Grosshansdorf
2 Universität Hamburg, Ohnhorststr. 18, D-22609 Hamburg

Introduction

The objective of the project is the generation of qualitatively optimized poplar trees with increased wood yield, modified lignin content and optimized plant architecture above the progress that can be achieved by conventional methods. Poplars with increased biomass as well with modified lignin content are of high economic interest serving as feed material for the direct retrieval of energy or of biopolymers in the pulping industry, respectively (Hägman *et al.* 2016).

Candidate genes for biomass production play a fundamental role in the different biomass-controlling pathways e.g. in vegetative growth capacity, vascular cambium activity, and secondary cell wall formation (Demura and Ye 2010) (Fig. 1). In the sub-project of the Thünen Institute of Forest Genetics, seven candidate genes that are involved in vegetative growth and/or suppression of flowering are being tested. A classical biotechnological approach is used to modify expression levels of these genes by over-expression or down-regulation (RNAi or amiRNA) via *Agrobacterium*-mediated transformation.

At first, the strategy is to use a transgenic approach to modify expression levels of selected candidate genes by using antibiotic resistance genes as selection marker. The transgenic poplar plants will be analysed in respect to biomass yield, wood property characteristics and tree architecture. Selected lines will be tested under greenhouse conditions for optimized characteristics.

Secondly, the development of cisgenic engineered poplars is aimed which proved to be much higher accepted in the society than transgenic ones. Therefore cisgenic poplars with increased biomass yield have a very high potential for field studies and plantation forestry.

Results

Selection and cloning of genes relevant for poplar transformation

Seven candidate genes involved in vegetative growth and/or flowering were selected for poplar transformation as candidate genes to increase biomass formation in the woody perennial *Populus* spec. With Arabidopsis mutants, it was shown by Melzer *et al.* (2008) that the known flowering genes *SOC1 (SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1)* and *FUL (FRUITFULL)* have an impact onto wood formation. Both genes are involved in regulation of meristem's transition from vegetative to reproductive stage (Demura and Ye 2010). Beside *SOC1* and *FUL*, five additional genes with so far widely unknown function were chosen. In *Populus* sp., they are expressed in both developing xylem and catkins or developing xylem and roots as expressed in the poplar eFP-browser (http://bar.utoronto.ca/efp_poplar/cgi-bin/efpWeb.cgi). All candidate genes have been cloned into conventional binary vectors for over-expression/down-regulation resulting in more than 20 different combinations.

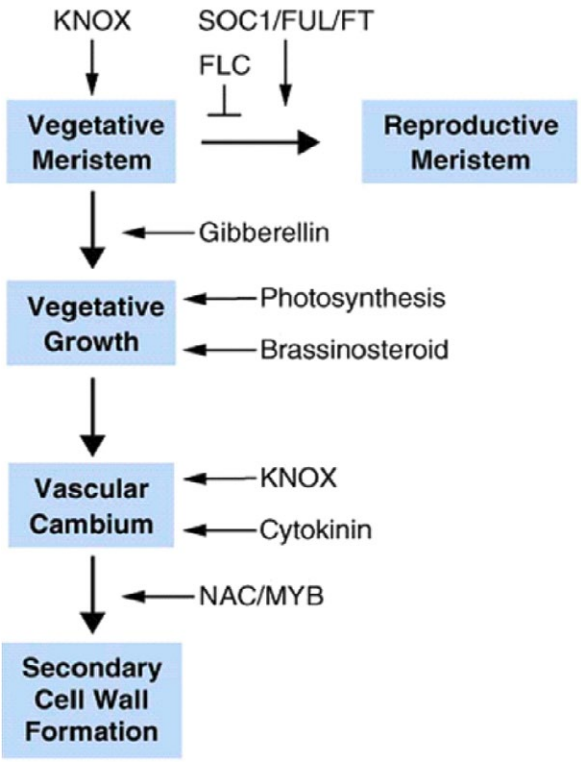


Fig. 1: Influence of SOC1 and FUL onto biomass formation. Both SOC1 and FUL act as regulator for the transition from vegetative to reproductive meristem identity (Demura and Ye 2010).

POP MASS Development and use of novel gene technologies to increase biomass yield in the woody perennial Populus spec.

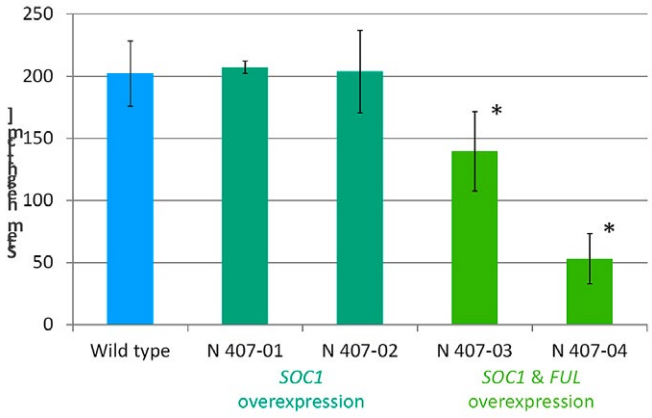


*Fig. 2: Plant height of independent overexpression poplar lines. Shown are mean values ± SD, significance marked with * ($p \leq 0,05$).*

P1 could adapt the efficient *P. tremula* transformation system also to *P.* × *canescens* successfully. As model organisms, the fast growing, well-regenerating and well-studied clones "INRA 717 1-B4" (*P.* × *canescens*, Leple *et al.* 1992) and Esch5 (*P. tremula* × *P. tremuloides*) were used for transformation. The selection of transgenic lines was assured by antibiotic resistance (either kanamycin or hygromycin). Over 40 transformations were accomplished via *Agrobacterium* yielding in numerous independent transgenic lines for both overexpression and down regulation. In addition to single transformations, double transgenic lines with various combinations were produced. Inclusive already existing plant material, more than 35 different transgenic lines are available for molecular analyses. The *in vitro* plants were transferred to soil and cultivated under greenhouse conditions for phenotype analyses. To date, nearly every produced transgenic line is available in greenhouse. There, over 2,500 plants are under observation.

Several noticeable phenotypes with altered biomass formation were identified. Double transgenic trees for the overexpression of both SOC1 and *FUL* are significant smaller than non-transgenic control plants (Fig. 2, 3). This is evidence, that both MADS box genes have an corporate influence on biomass formation. Due to genome duplication events (Tuskan *et al.* 2006), more than one copy of *SOC1* and *FUL* are identifiable in poplars. In the next step, a biomass increase should be achieved with the knockdown of all poplar orthologues of *SOC1* and FUL.

Genetic modifications of the other candidate genes resulted at times in significant smaller or higher plants. To date, uniform growth alterations in lines overexpressing/down regulating the same gene were not discovered. Expression studies have unraveled the degree of transgene transcription. **Lines with superior growth** are currently under detailed investigation and promising candidates for future applications.

Fig. 3: Double transgenic SOC1 / FUL phenotypes in the greenhouse. Double transgene poplars (left) with overexpression of FUL and SOC1 after eight month growth are significant smaller.

Transformations have been and also are being performed for partner labs (P2, P4), e.g. nine for P2 (FU Berlin, T. Schmülling, M. Riefler). Transgenic plants were produced and provided to P2 for further analyses.

In cooperation with the Fraunhofer-Institut für Angewandte Polymerforschung, to whom plant material from selected transgenic lines with an altered growth was provided, were analysed in respect to wood chemical constitution. While little variations in minor wood components were found, the "big three" lignin, cellulose and hemicelluloses remained unmodified.

POP MASS Development and use of novel gene technologies to increase biomass yield in the woody perennial Populus spec.

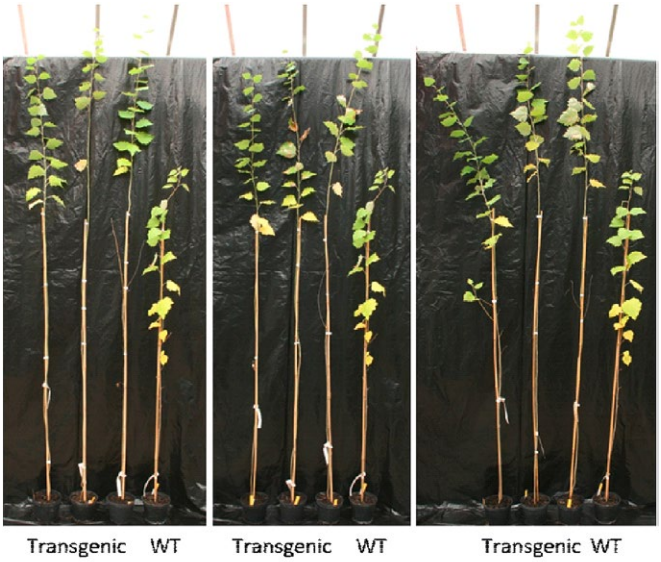


Fig. 4: Noticeable phenotypes of transgenic poplars. Knockdown of a xylem-flower-gene led to higher plants than the wildtype.

Taken together, we have shown that the two genes *SOC1* and *FUL* have a strong corporate impact on wood formation in poplars. Double knockdown of both genes is expected to increase biomass formation. Due to poplar genome duplications, two copies of *FUL* and three copies of *SOC1* are targets for the knockout. In the project prolongation, knockout will be achieved with the newest genome editing technique CRISPR/Cas9. Worldwide, we are under the first workgroups to establish this genome editing technique for poplars.

Generation of cisgenic poplar transformants

The main focus of the POP MASS-project is the generation of new poplar lines with increased wood yield and optimized plant architecture. Since in Europe transgenic trees will not be permitted for cropping, an important goal of the project is the development of cisgenic poplar lines. Cisgenesis describes the genetic modification of a recipient plant with one or more genes from the same or a genetically crossable plant species. The genes are transferred unchanged with their introns and control elements in the new genome. After crossing out, no foreign genes or DNA-sequences are passed to the progeny. It is hoped that politicians and consumers, aware of the limitation of the genetic modification within the natural species boundaries, tone down their reservations towards green genetic engineering.

A possible "cisgene approach" for the genus *Populus* was tested. In the first phase of POP MASS, two different methods of direct DNA uptake namely the biolistic and the PEG-based protoplast transformation method were exercised to provide an opportunity to transfer gene fragments containing only cisgenic elements directly into poplar. As a preliminary test for a subsequent cisgenic approach, *gfp* and *dsRed* were applied as visual reporter genes with the intention to optimize not only the transformation parameter but also the conditions of the in vitro culture to receive a maximum quantity of transformed cells. The results of the investigation confirmed, that the biolistic and the protoplast transformation method can be considered suitable to transfer gene fragments directly into poplar tissue. It has succeeded in regenerating firmly *gfp* and *dsRed*-expressing transgenic poplars.

Regenerating callus is used to establish biolistic transformation in poplar. Optimization of the *P.* × *canescens in vitro* regeneration system was needed. Cloning of *gfp* and *dsRed* in appropriate biolistic vectors have successfully been completed. Testing of bombardment parameters was performed. Transformed cells were selected via fluorescence tests, and *gfp* and *dsRed* transgenic calli have been obtained. Illuminating poplar plants have been regenerated.

In the project prolongation, these optimized parameters are now applied to transfer the poplar *acetolactate synthase* gene (*ALS*) as an endogenous candidate selection marker gene for the cisgenic approach into the poplar genome of INRA 717 1B4. The *ALS*-enzyme catalyzes the first step in the synthesis of the branched-chain amino acids valine, leucine and isoleucine. All three are essential. Inhibitors of *ALS* are used as herbicides. However, a naturally occurring mutation in the *ALS*-gene of poplar confers resistance to the herbicide chlorsulfuron and thus is well suited to distinguish transformed poplar plants from non-transformed. The ultimate objective is to transfer possible target genes for

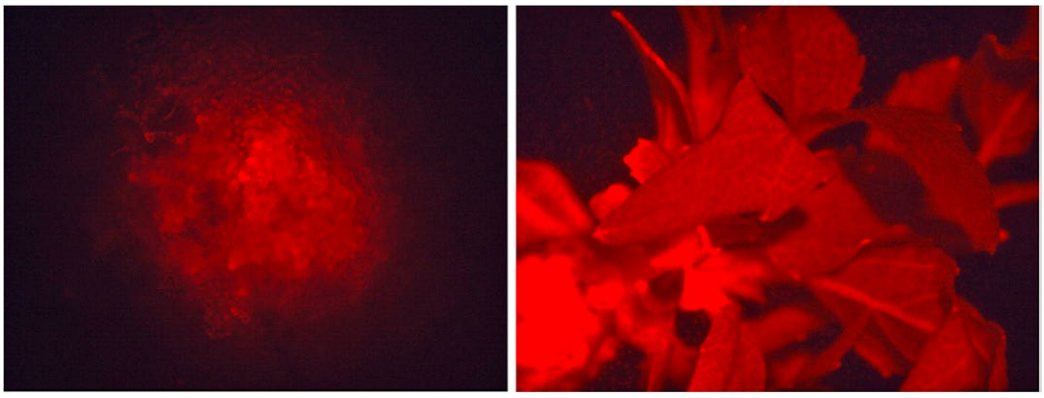


Fig. 5: Left: Dsred-expressing callus three months after transformation. Right: Development of a firmly „dsRed" transgenic poplar six months after transformation growing on ½ MS-medium.

the improvement of poplar together with the *ALS*-gene and select the cisgenic plants by means of their resistance to the herbicide chlorsulfuron.

The state of the experiments is that mesophyll protoplasts of INRA 717 1B4 have been transformed with plasmid-DNA, containing the *ALS*-gene. The transformation was performed with 40 % PEG 1500. In each experiment, simultaneous to the transformation with ALS, also transformation with the *gfp*-gene was performed. Since the efficiency of these transformations was high, it is assumed that shortly ALS-positive plantlets will be regenerated. Currently, one putative chlorsulfuron-resistent line is under regeneration.

Molecular characterization of transgenic lines

Regenerated plants are under selection for all constructs. Independent transgenic lines with propagated individuals exist for all constructs (1 to 11 lines per construct up to now). For all other constructs, transformations have been completed. Transgenic lines of 9 constructs have successfully been PCR-tested for transgene presence using transgene- and T-DNA specific primers, and transgenic lines of 3 constructs have already been tested in Southern blot analyses to determine transgene copy number. Selection of single-copy containing lines has already been started. Analysis of the stability of gene expression (targeted genes) is currently been done on basis of antibiotic resistance during *in vitro* culture. A series of qPCR experiments will provide insights in expression of the transferred genes.

Evaluation of growth behavior, morphology, and wood anatomical characteristics

Transgenic lines of 6 constructs have been transferred to soil and are being cultivated in the greenhouse. Phenotyping (growth behavior, branching) is being performed or will be evaluated next winter season (height of plants, stem diameter, etc.). Best growing transgenic lines were selected and harvested last winter season for wood anatomical and chemical characteristics. The results of the analyses will be available soon.

A pH-D thesis by Tobias Brügmann will be submitted to University Hamburg in 2016.

References

- *Demura T, Ye Z-H (2010) Regulation of plant biomass production. Curr Opin Plant Biol 13: 299-304.*
- *Häggman H, Sutela S, Fladung M (2016) Genetic Engineering Contribution to Forest Tree Breeding Efforts. In: Vettori C et al (Eds), Biosafety of Forest Transgenic Trees. Springer Science+Business Media Dordrecht, Forestry Sciences 82, pp. 11-29. DOI 10.1007/978-94-017-7531-1_2.*
- *Leple JC, Brasileiro ACM, Michel MF, Delmotte F, Jouanin L (1992) Transgenic poplars: expression of chimeric genes using four different constructs. Plant Cell Reports 11: 137-141.*
- *Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T (2008) Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nat Genet 40: 1489-1492.*

PRE-BREED YIELD Capturing genomic diversity in rapeseed for precision breeding

Rod Snowdon[1], Amine Abbadi[2] and Gunhild Leckband[3]

1 Department of Plant Breeding, Justus Liebig University, Giessen
2 NPZ Innovation GmbH, Holtsee
3 German Seed Alliance GmbH, Cologne

The public-private research consortium "*PRE-BREED YIELD*– Precision Breeding for Yield Gain in Oilseed Rape" represents an outstanding example for pre-competitive research cooperation in the plant breeding sector. The interdisciplinary *PRE-BREED YIELD* consortium[1], funded by BMBF from 2011-2015 within the framework of the Plant Biotechnology for the Future programme, included seven commercial plant breeding companies, along with seven scientific institutes. The consortium partners contributed diverse, complementary expertise spanning the fields of plant breeding, genetics, genomics, statistics, plant physiology, abiotic stress resistance and plant phenotyping. Through development of extensive, highly novel plant populations, *PRE-BREED YIELD* provided a unique pre-breeding resource to enhance sustainable agricultural production. By associating these populations with large-scale datasets from high-resolution genome analyses and novel, precise phenotyping strategies, the project established a valuable legacy for knowledge-based breeding of high-performing commercial cultivars in one of Germany's most important crops.

Modern crop breeding is frequently dependent on germplasm that survived domestication processes with low genetic diversity, as a result of strong founder effects. Normally, only a small proportion of the diversity present in the primary gene pools of crop plants is readily accessible to breeders. A prime example for this phenomenon is rapeseed (oilseed rape, canola: *Brassica napus*; Figure 1), a very young agricultural species with an extremely narrow primary gene pool due to strong eco-geographical selection and intensive artificial selection in breeding.

Overcoming genetic bottlenecks is an ongoing, difficult challenge for crop breeders, as diversity needed for sustainable breeding is depleted in primary gene pools. Recovery of lost diversity, by crossing with exotic forms, related species or wild relatives, requires costly and time-consuming "pre-breeding" efforts. These represent a considerable investment that may not lead to significant commercial returns for many decades. Breeders therefore depend strongly on public research support for pre-breeding activities.

On the other hand, huge recent advances in high-resolution genomic technologies open completely new possibilities

Fig. 1: Brassica napus (rapeseed, oilseed rape, canola). *Through intensive breeding to remove undesirable seed quality traits and improve agronomic performance, today's rapeseed varieties yield large quantities of an extremely healthy vegetable oil, containing an optimal composition of highly desirable unsaturated fatty acids, along with a nutritious seed meal low in anti-nutritive glucosinolates. Rapeseed oil is also ideal for use as a biofuel, and the meal delivers one of Europe's most valuable sources of non-imported, non-GM protein for livestock feeding. Within a few decades in the late 20th century, breeding breakthroughs took rapeseed from agricultural obscurity to become the second most important oilseed crop in the world. Paradoxically, however, this resounding success simultaneously eroded diversity for important secondary traits, particularly resistance to diseases and abiotic stress. Recovering useful diversity for these traits is of paramount importance to sustain future productivity under reduced agricultural inputs and meet additional challenges posed by climatic change. Image: R. Snowdon.*

1 PRE-BREED YIELD Partners: German Seed Alliance GmbH, Norddeutsche Pflanzenzucht Hans-Georg Lembke KG, Deutsche Saatveredelung AG, KWS SAAT KG, Limagrain GmbH, Syngenta Seeds GmbH, Bayer Crop Science Raps GbR, Justus Liebig University Giessen, Georg August University Göttingen, Forschungszentrum Jülich GmbH, Max Planck Institute for Plant Breeding Research, University of Bonn, Julius Kühn Institute for Resistance Research and Stress Tolerance, Leibniz Institute of Plant Genetics and Crop Plant Research

for directed discovery and exploitation of useful genetic variation in breeding. As a case-study for genomics-based development of pre-breeding resources, *PRE-BREED YIELD* explored new avenues to apply genomic data in knowledge-based discovery and implementation of rich variation from the primary and secondary gene pools of rapeseed. A combination of "omics" technologies with precise plant phenotyping strategies was implemented to optimise the utilisation of novel diversity for rapeseed breeding. The extensive plant resources and large-scale genomic data developed during the project provide a long-term basis for sustainable agricultural production through knowledge-based development of improved elite varieties, in line with the National Strategy "Bio-economy 2030".

The "breeders paradox" in rapeseed

It is well known that reduced genetic diversity in breeding pools causes significant erosion of heterotic potential. This "breeders' paradox" is particularly prominent in rapeseed, where the recent domestication history and stringent selection for specific seed qualities extremely created genetic bottlenecks within eco-geographically divergent international breeding pools (Snowdon *et al.* 2015).

Brassica napus is one of the youngest known crop species. It arose under anthropogenic influence, probably just a few hundred years ago, from spontaneous interspecific hybridisations between Mediterranean and Asian cabbage species. These rare hybridisation events gave rise to morphologically divergent allopolyploid[2] forms that were selected by humans for use as oilseed, fodder rape, kale and rutabaga crops. Due to intense breeding, present elite rapeseed varieties adapted to European growing conditions have very restricted diversity. Asian oilseed forms and other exotic *B. napus* contain potentially interesting diversity for development of diversified breeding pools for hybrid breeding, but their use for rapeseed breeding is difficult due to their inherent lack of essential adaptation and seed quality characters. Furthermore, because the species derived from just a few interspecific hybridisation events, the diversity in the primary gene pool is limited compared to most other crops. Fortunately, however, with the help of tissue culture techniques it is possible to readily recreate *de novo* forms of "synthetic" *B. napus* by crossing the two parental species. This opens the way to generate novel diversity that goes far beyond that present in the primary gene pool.

Genome-assisted recovery of crop genetic diversity

Large collections of highly diverse but poorly adapted *B. napus* accessions were available from previous work of the *PRE-BREED YIELD* partners. These included genetically fixed inbred lines derived from European genebank collections, along with extremely diverse synthetic *B. napus* derived from crosses between the progenitor species of rapeseed. Both sets of materials represent completely novel diversity in comparison to modern oilseed rape cultivars, however such variants can be extremely difficult to deal with in breeding programmes because they often lack essential adaptation traits and generally have extremely poor agronomic performance. Nevertheless, such sources are rich in hidden diversity that can be extremely valuable to adapt crops to new environmental challenges, for example as a consequence of climate change.

In *PRE-BREED YIELD*, an unprecedented effort was made to capture diversity from these novel *B. napus* gene pools, using a so-called "nested" crossing strategy that was originally developed in the USA for maize (McMullen *et al.* 2009; Yu *et al.* 2008). Nested crossing designs fix genetically diverse chromosome segments from highly diverse founder lines in the genetic background of an adapted, elite cultivar. By crossing 50 highly diverse *B. napus* founders with a single, elite rapeseed cultivar, and backcrossing particularly exotic materials to improve their adaptation to central European growing environments, the *PRE-BREED YIELD* breeders fixed the genetic diversity from the exotic founders in a vast interrelated population of over 2500 *B. napus* offspring. This huge nested association mapping (NAM) population carries sufficient adaptation for direct implementation in breeding, enabling detailed analysis of performance in field trials and controlled-environment experiments.

Resources for gene discovery and genome-based breeding

In combination with detailed phenotype information, high-resolution genome data from NAM populations represents a powerful resource for analysis and discovery of

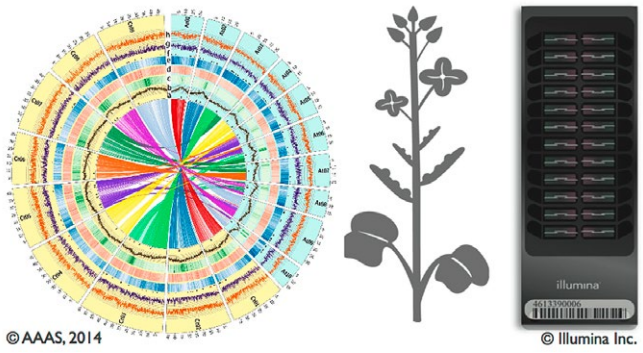


Figure 2: Enabling resources for rapeseed research and breeding. *The Pre-BreedYield consortium made important contributions to (a) the groundbreaking publication of the Brassica napus reference genome sequence (reprinted from Chalhoub et al. 2014 with permission from the American Association for the Advancement of Science) and (b) the high-density Illumina Brassica 54k SNP genotyping array (Clarke et al. 2016). These essential enabling resources for high-resolution genetic analysis provide the basis for genome-based breeding and detailed genome research in Europe's most important oilseed crop.*

2 Allopolyploid: Containing two or more sub-genomes derived from a hybridisation between different progenitor species.

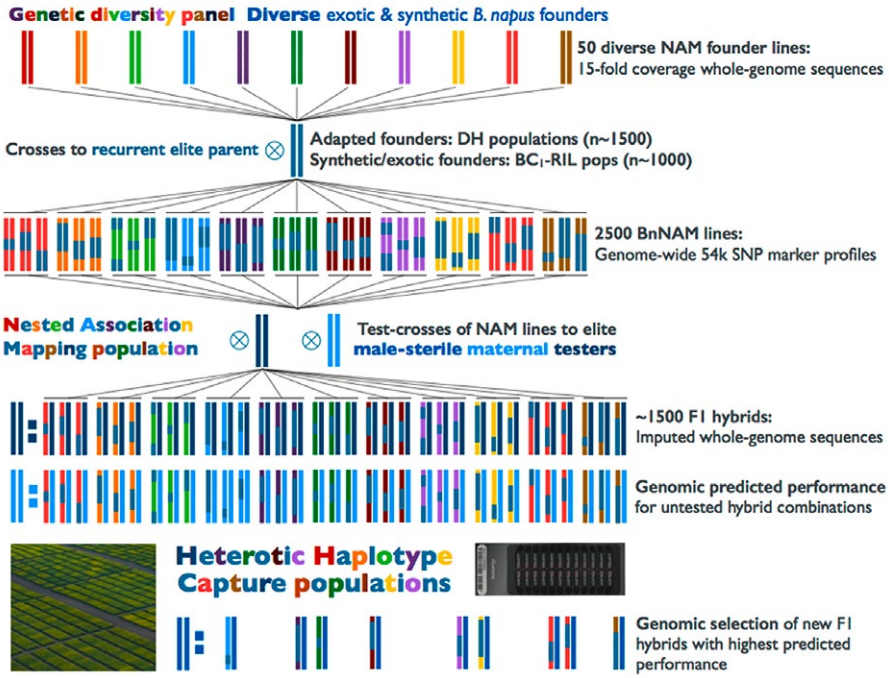


Figure 3: From heterotic patterns to improved yield performance. *The PRE-BREED YIELD NAM and NAM-hybrid populations represent a hugely valuable reservoir of diversity for future breeding. By implementing genome-based selection strategies based on genome wide SNP profiles of the populations, along with the whole-genome sequences of the founders, it is now feasible to estimate the performance of "virtual" hybrids based solely on heterotic patterns measured in the genome profiles of the parents. This allows breeders to focus resources on generating and field-testing only those hybrid combinations with best predicted performance. leading to a potentially strong increase in genetic gain per breeding cycle.*

chromosome regions and genes responsible for inheritance of important traits. In *PRE-BREED YIELD*, the latest ultra-high-throughput genome sequencing technologies were used to decode the whole-genome sequences of all 51 founder accessions of the new NAM populations. Bioinformatic analysis revealed more than 4 million SNP single-nucleotide polymorphism (SNP) sequence variants throughout the rapeseed genome (Schmutzer *et al.* 2015), so far the most comprehensive catalogue of genome sequence diversity to be generated for *B. napus*. The sequence data also uncovered a completely unexpected degree of genome structural variation, with chromosome rearrangements revealed as a major evolutionary driver of trait variation both in synthetic and natural *B. napus* forms. This finding became one of the key features in the publication of the rapeseed genome sequence, in the leading scientific journal Science, by an international consortium involving *PRE-BREED YIELD* participants (Chalhoub *et al.* 2014) (Figure 2a).

Another extremely important recent advance in rapeseed genomics and breeding, enabled with the help of data contributions from *PRE-BREED YIELD* partners, was the highly successful, high-density Illumina SNP genotyping array for rapeseed (Clarke *et al.* 2016) (Figure 2b). By genotyping the 2500 lines of the rapeseed NAM populations with the 54,000 SNP markers on the array, every chromosome segment in every NAM line could be traced back to the original parental sequence. This enables the imputation of whole-genome sequences for the entire interrelated population. Besides ultra-high resolution genetic mapping of important traits, this vast dataset also provides a previously unimaginable basis for genomics-assisted breeding to improve resistance, sustainability and hybrid vigour.

Heterotic haplotype capture: Increasing yield potential through exploitation of heterosis

Directed exploitation of heterosis, or hybrid vigour, achieved through breeding of hybrid varieties from crosses between genetically divergent gene pools, has been one of the most important contributors to yield increases in major crops over recent decades. The ability to link genome-wide SNP data to genome structural variation also provides a unique opportunity to fix chromosome segments associated with hybrid performance in breeders' gene pools. Extending the NAM concept beyond fixed inbred lines, the *PRE-BREED YIELD* breeders crossed a subset the 1500 NAM lines to a common male-sterile mother line, generating more than 1000 interrelated hybrids with varying degrees of hybrid vigour (heterosis) contributed by the different exotic genome donors. Extensive phenotype data was generated from multi-year, multi-location field trials to investigate the yield performance of these hybrids under different environmental conditions, for example under reduced nitrogen fertilisation input. Specific subpopulations were targeted based on detailed characterisation of the founder lines for drought stress or nutrient deficiency, for example. Interestingly, despite the inherently poor agronomic performance of all founder accessions compared with modern elite cultivars, some of the experimental hybrids generated from the NAM lines nevertheless exhibited a surprisingly strong yield performance comparable to hybrids generated with the same male-sterile tester. This demonstrates the high yield potential that is hidden in previously unexplored genetic resources.

To our knowledge, the vast population of interrelated NAM hybrids, associated with high-resolution genome informa-

tion and extensive field phenotype data, is unique for any crop species to date. Using the genome-wide genetic profiles of each NAM line, it is possible to trace chromosome segments associated with hybrid vigour back to the original genome sequence of their founder accessions. This provides an unprecedented basis to associate complementary chromosome patterns with hybrid performance. In future, analysis of such "heterotic haplotypes" (Snowdon *et al.* 2015) (Figure 3) can help breeders to generate and identify the most promising crossing partners for new hybrid combinations with optimised performance.

Challenges for the future: Maintaining and implementing valuable crop research resources

Vast experimental plant populations and genomic data resources, like those generated in *PRE-BREED YIELD*, provide an enormously valuable basis for future crop research and breeding. Besides their implementation in long-term breeding programmes of the participating breeders, the *PRE-BREED YIELD* NAM populations have already formed the basis for numerous ongoing research activities on a national and international level. Examples include continuing public-private research projects to increase resistance to important diseases (German-French Plant KBBE consortium *GeWiDis*), to develop systems-biological approaches for predicting hybrid performance under environmental constraints (BMBF eBio consortium *PROGReSs*), or to develop new machine-learning approaches for hybrid performance prediction (BMBF consortium *BreedPatH*). Ongoing spinoffs in research and breeding and among the most important long-term outputs of *PRE-BREED YIELD*. The close ties between industrial and academic partners established during the project ensure ongoing transfer of research results into commercial innovation in the plant breeding sector.

On the other hand, managing and maintaining such huge plant and data resources represents a huge ongoing challenge for the *PRE-BREED YIELD* partners. The enormous size and pre-breeding nature of the rapeseed NAM populations and their founders makes generation and distribution of seeds an extremely labour-intensive and costly enterprise that cannot be sustained by breeders in parallel to their commercial breeding operations. Long-term public investment in appropriate infrastructure for seed storage, multiplication and distribution is therefore essential to perpetuate these highly valuable plant resources and make them available for future research. Furthermore, the full exploitation of the vast *PRE-BREED YIELD* data resources can only be realised in the context of a publicly available data storage, sharing, visualisation and analysis platform. Ongoing research funding, for implementation of the *PRE-BREED YIELD* genome data into a public crop genome informatics portal, is therefore an essential next step to ensure the long-term added value of the unique resources generated by this highly successful consortium.

Bibliography

- *Chalhoub B, Denoeud F, Liu SY, Parkin IAP, Tang HB, Wang XY, Chiquet J, Belcram H, Tong CB, Samans B, Correa M, Da Silva C, Just J, Falentin C, Koh CS, Le Clainche I, Bernard M, Bento P, Noel B, Labadie K, Alberti A, Charles M, Arnaud D, Guo H, Daviaud C, Alamery S, Jabbari K, Zhao MX, Edger PP, Chelaifa H, Tack D, Lassalle G, Mestiri I, Schnel N, Le Paslier MC, Fan GY, Renault V, Bayer PE, Golicz AA, Manoli S, Lee TH, Thi VHD, Chalabi S, Hu Q, Fan CC, Tollenaere R, Lu YH, Battail C, Shen JX, Sidebottom CHD, Wang XF, Canaguier A, Chauveau A, Berard A, Deniot G, Guan M, Liu ZS, Sun FM, Lim YP, Lyons E, Town CD, Bancroft I, Wang XW, Meng JL, Ma JX, Pires JC, King GJ, Brunel D, Delourme R, Renard M, Aury JM, Adams KL, Batley J, Snowdon RJ, Tost J, Edwards D, Zhou YM, Hua W, Sharpe AG, Paterson AH, Guan CY, Wincker P (2014) Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345:950-953*
- *Clarke WE, Higgins EE, Plieske J, Wieseke R, Sidebottom C, Khedikar Y, Batley J, Edwards D, Meng J, Li R, Lawley CT, Pauquet J, Laga B, Cheung W, Iniguez-Luy F, Dyrszka E, Rae S, Stich B, Snowdon RJ, Sharpe AG, Ganal MW, Parkin IA (2016) A high-density SNP genotyping array for Brassica napus and its ancestral diploid species based on optimised selection of single-locus markers in the allotetraploid genome. Theor Appl Genet*
- *McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li HH, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell SE, Peterson B, Pressoir G, Romero S, Rosas MO, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz JC, Goodman M, Ware D, Holland JB, Buckler ES (2009) Genetic Properties of the Maize Nested Association Mapping Population. Science 325:737-740*
- *Schmutzer T, Samans B, Dyrszka E, Ulpinnis C, Weise S, Stengel D, Colmsee C, Lespinasse D, Micic Z, Abel S, Duchscherer P, Breuer F, Abbadi A, Leckband G, Snowdon R, Scholz U (2015) Species-wide genome sequence and nucleotide polymorphisms from the model allopolyploid plant Brassica napus. Sci Data 2:150072*
- *Snowdon RJ, Abbadi A, Kox T, Schmutzer T, Leckband G (2015) Heterotic Haplotype Capture: precision breeding for hybrid performance. Trends in Plant Science 20:410-413*
- *Yu JM, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178:539-551*

RYE SELECT Genome-based precision breeding strategies for rye

Eva Bauer1, Wiltrud Erath[1], Chris-Carolin Schön[1], Manfred Schönleben[1], Thomas Schmutzer[2], Uwe Scholz[2], Angela-Maria Bernal-Vasquez[3], Jens Möhring[3], Hans-Peter Piepho[3], Thomas Miedaner[4], Ivan Barilar[5], Karl Schmid[5], Bernd Hackauf[6], Andres Gordillo[7], Viktor Korzun[7], Malthe Schmidt[7], Brigitta Schmiedchen[7] and Peer Wilde[7]

1 Technische Universität München, Plant Breeding, Freising;
2 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Bioinformatics and Information Technology, Stadt Seeland OT Gatersleben;
3 University Hohenheim, Biostatistics, Stuttgart;
4 University Hohenheim, State Plant Breeding Institute, Stuttgart;
5 University Hohenheim, Crop Biodiversity and Breeding Informatics, Stuttgart;
6 Julius Kühn-Institute (JKI), Institute for Breeding Research on Agricultural Crops, Groß Lüsewitz;
7 KWS LOCHOW GMBH, Bergen (Coordinator of the RYE SELECT project)

Introduction

In the light of global environmental changes, rye (*Secale cereale* L.) promises an excellent adaptation capacity to growth-limiting factors such as drought, frost, or nutrient deficiencies. Rye (haploid genome size of 7.91 Gb) is the only cross-pollinated small-grain cereal and in Germany mainly bred as a hybrid crop. Hybrid breeding requires extensive testing and selection programs in two divergent genepools. For harnessing the full yield potential of rye, genome-based breeding strategies play an important role in increasing selection efficiency and accelerating breeding processes. Our aims in the RYE SELECT project were to i) improve the genomic toolbox for rye by establishing a draft genome sequence which was instrumental for developing a high-density SNP genotyping array and target-specific gene-based markers, ii) adapt and develop state-of-the-art genome-based breeding methods for improving yield and yield stability, iii) identify and validate genomic regions involved in important traits.

Genomic resources for rye

Using the Illumina HiSeq2000 platform we sequenced the rye inbred line Lo7. In total, we produced ~78x genome coverage of paired end (46x) and mate pair (32x) reads from different libraries. Using the paired end sequences a rye WGS assembly with ~1.58 million contigs and a total sequence length of 1.69 Gbp was constructed. In collaboration with the group of Klaus F.X. Mayer (Helmholtz Centre Munich) we assigned 50% of these WGS contigs (representing 67% of all nucleotides) and ~14% of the mate pair sequences to one of the seven rye chromosomes. For hierarchical scaffolding we used the WGS contigs and mate pair sequences as well as the chromosome assignments of both data sets. As result ~2.8 Gbp scaffolds were produced where more than 66% of the represented base pairs were assigned to the seven rye chromosomes (Fig.1).

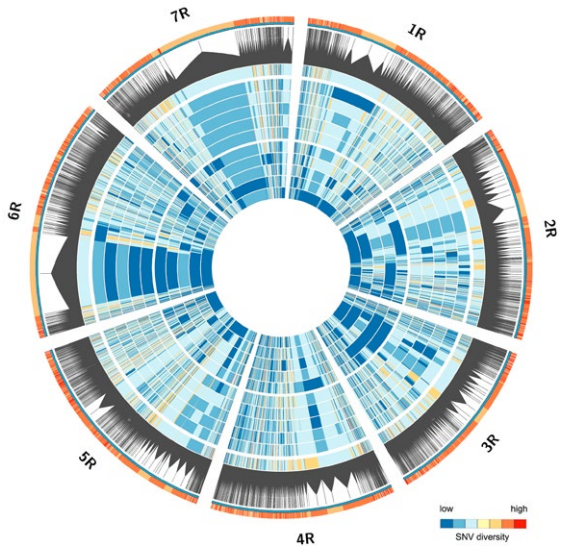


Fig. 1: SNV density and distribution along the rye nuclear genome
The figure depicts WGS contigs that were anchored to the rye genetic framework (combined representation of 90k genetic map and rye genome zipper). The outer circle shows the position of genes along the genetic map of the seven chromosomes (1R to 7R). The gene density was plotted as heatmap ranging from low (yellow) to high (red). The following circle illustrates the anchoring of the genetic map to the physical sequence contigs. In total, 44,371 WGS contigs were anchored to 2.063 unique genetic map positions of the rye genetic framework. The heatmaps of the subsequent circles present the SNV density along the anchored WGS contigs for each of the re-sequenced rye genotypes. A low number of SNVs is illustrated in blue segments and high number of SNVs in red segments. From inside to outside, the tracks represent five inbred lines from the seed parent pool (Lo90, Lo115, Lo117, Lo176, Lo191), five inbred lines from the pollen parent pool (Lo282, Lo298, Lo310, Lo348, Lo351) and the S. vavilovii accession. (Source: Uwe Scholz, Thomas Schmutzer, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Bioinformatics and Information Technology).

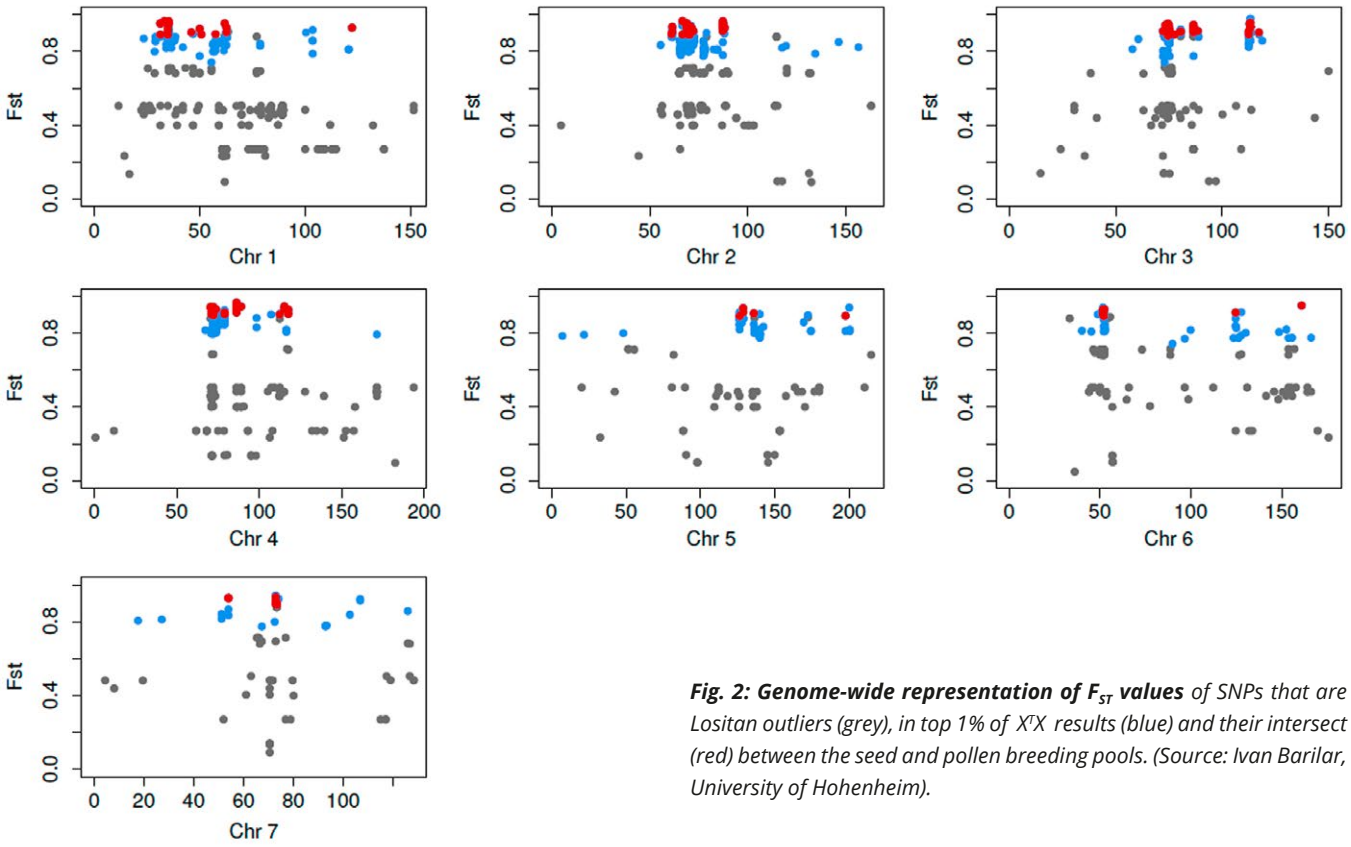


Fig. 2: Genome-wide representation of F_{ST} values *of SNPs that are Lositan outliers (grey), in top 1% of X^TX results (blue) and their intersect (red) between the seed and pollen breeding pools. (Source: Ivan Barilar, University of Hohenheim).*

Additionally, we sequenced ten diverse inbred lines and one *Secale vavilovii* accession with mid-coverage (average 16x) to develop a high-density SNP array. SNP discovery was performed by mapping of paired end reads of these 11 genotypes to the Lo7 WGS assembly followed by variant calling. Comprehensive filtering of variants led to a final number of 8.3 millions unique polymorphic sites. These were further filtered for design of a 600k Affymetrix genotyping array. The performance of SNPs on the 600k array was evaluated by analysing a large panel of diverse rye samples from elite germplasm and genetic resources as well as a mapping population. In total, 316,546 (52.7%) polymorphic SNPs resulted in scorable genotypes and thus could be experimentally validated. These SNPs are a highly valuable resource for genome-based breeding, fine-mapping of specific target regions and diversity studies in rye.

In a recombinant inbred line population from the cross Lo7 x Lo225 a high-density genetic map was constructed which aligns well with previous rye genetic maps. The map comprises 87,820 markers, representing 44,371 Lo7 WGS contigs and 3,022 contigs from sequence resources generated in previous projects (e.g. GABI RYE-EXPRESS). The high-density genetic map thus allowed anchoring of these sequence contigs along the rye genome and enables cross-species comparisons with other grass genomes at an unprecedented coverage and density.

Population genetics

The eleven sequenced genomes together with the *Secale vavilovii* outgroup were used in an initial population genetic analysis of rye genetic diversity, in particular with respect to differences between the pollen parent and seed parent pool. We calculated descriptive population genetic parameters such as Tajima's *D*, nucleotide diversity and F_{ST} both across the genome and for individual contigs. Methods for selection detection identified contigs and genes with footprints of artificial selection during plant breeding. We further analysed a larger set of seed and pollen parent lines as well as genetic resources that were genotyped with the 600k SNP array. The application of population differentiation (X^TX) and of selective sweep (Lositan) tests on a subset of 78,731 SNPs with assigned genomic positions revealed that different genomic regions changed patterns of genetic diversity due to selection independently in the seed and pollen parent pool (Fig. 2). This pattern change likely results in selection on fertility traits or heterosis-related genes. Further genomic regions were identified that likely were selected in response to adaptation to Central European climate. The results of this analysis provide a detailed insight into the genetic diversity of the available rye material and provide information that can be used for future management of breeding material diversity.

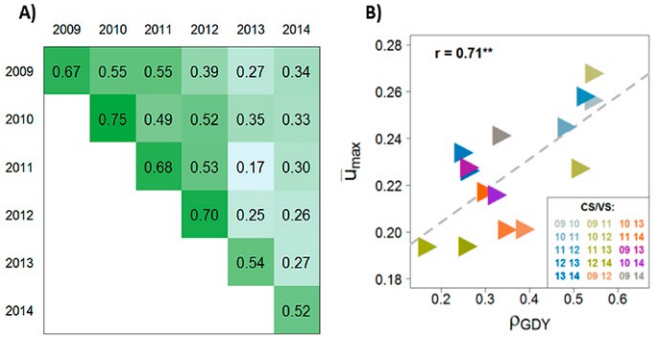


Fig. 3: Genomic prediction in rye within and across six breeding cycles *A) Within (diagonal elements) and across cycle (off-diagonal elements) prediction accuracies for grain dry matter yield from GBLUP with 10 × 5 fold cross-validation using constant calibration (N=208) and validation set (N=52) sizes. B) Across cycle prediction accuracies for grain dry matter yield from GBLUP (x-axis) plotted against the average maximum kinship (y-axis). Color coded triangles indicate cycles in calibration (CS) or validation set (VS). Results are shown for all possible pairwise cycle combinations, with one cycle forming the calibration (N=208) and one cycle the validation set (N=52), respectively. (Source: Manfred Schönleben, Eva Bauer, Technische Universität München,*

Genome-wide breeding approaches

The concept of genomic selection (GS) is currently widely adapted in plant and animal breeding. GS integrates genotypic and phenotypic information from a training population for the prediction of genomic breeding values of unphenotyped individuals with maximum accuracy. At the phenotypic level, we evaluated spatial and non-spatial models for multi-location field trials using weakly connected trial series and scrutinized the outlier detection method used by KWS LOCHOW GMBH (PlabStat-based). We then linked the phenotypic with the genotypic analysis exploring the consequences of the different models on the prediction performance (Bernal-Vasquez *et al.* 2014) and proposed and compared outlier methods under a genomic selection cross validation scenario (Bernal-Vasquez *et al.* 2016).

We investigated the prediction performance of genomic best linear unbiased prediction (GBLUP) within and across six interconnected cycles of a hybrid rye breeding program using cross validation, and evaluated systematically the importance and effect of the kinship structure on the prediction performance of economically important traits in rye. Within breeding cycles, GBLUP yielded intermediate to high prediction accuracies for the traits grain dry matter yield, plant height and thousand kernel weight. In across cycle scenarios, we observed for all traits promising prediction accuracies on intermediate levels. Across cycle genomic prediction for cycles 2009 to 2012 yielded significantly ($p<0.05$) higher accuracies, compared to cycles 2013 and 2014, which can be attributed to the higher level of relatedness between the earlier cycles (Fig. 3A). For grain dry matter yield we observed a significant correlation ($r=0.71$, $p<0.003$) between prediction accuracies and the across cycle maximum kinship (Fig. 3B). By augmenting the calibration sets with phenotypes of multiple cycles and by increasing the calibration set size, predictions for all traits could be further improved. Our study paved the way for the implementation of genomic selection in hybrid rye breeding by determining important parameters and insights in the optimal design of calibration sets for model training to maximize prediction accuracies.

Targeted approaches

We further targeted specific genomic regions for agronomically important "must-have" traits and identified genomic regions involved in frost tolerance by QTL mapping in a population of recombinant inbred lines (RILs). The mapping population was developed from a cross of a plant from a highly frost tolerant selected fraction of the Canadian 'Puma' population with the elite line Lo157. The RILs were genotyped in the F_4 and the F_5 generations with genome-wide SNPs. In each of three years, line *per se* and testcross performance was assessed in controlled freezer tests and in multilocation field trials in Canada and Russia. Both, in freezer and in field trials, the largest proportion of phenotypic variance for frost tolerance was explained by a major QTL on chromosome 5R at the *Fr-2* locus which harbors a cluster of *Cbf* transcription factors, known to be involved in the frost responsive network. A comparison of predictive ability obtained from the QTL-based model with different genomic prediction models showed that GBLUP models including markers flanking main effect QTL and the variable selection method LASSO were best suited for selection of the most frost tolerant candidates. Medium to high genetic correlations among traits assessed in the two phenotyping platforms revealed the suitability of the controlled environment as a pre-screening test for frost tolerance. In addition, a high genetic correlation between line per se and testcross performance was observed in the freezer test which is favourable for early selection on the line per se level in this controlled environment.

Building on a rye introgression library established in previous GABI projects, genomic segments carrying novel alleles for traits such as pollen-fertility restoration and quality were reduced in size by recombination and marker-based selection fine-mapped and further characterized for line per se and testcross performance. The use of effective restorer-of-fertility (*Rf*) genes provides an enabling technology in hybrid rye breeding programs. The superior pollen-fertility restoration ability of the backcross lines with the gene Rfp^{Avg} on chromosome 4RL from the Iranian donor Altevogt 14160 was confirmed across locations and years in greenhouse and field. Plant height, grain yield, protein content, soluble pentosane content and falling number, however, were significantly affected compared to sub-introgression lines without this gene. In particular, BC lines with Rfp^{Avg} were, on average, 10 cm taller, and had 3.1 dt/ha lower grain yield. The barley draft genome sequence enabled the

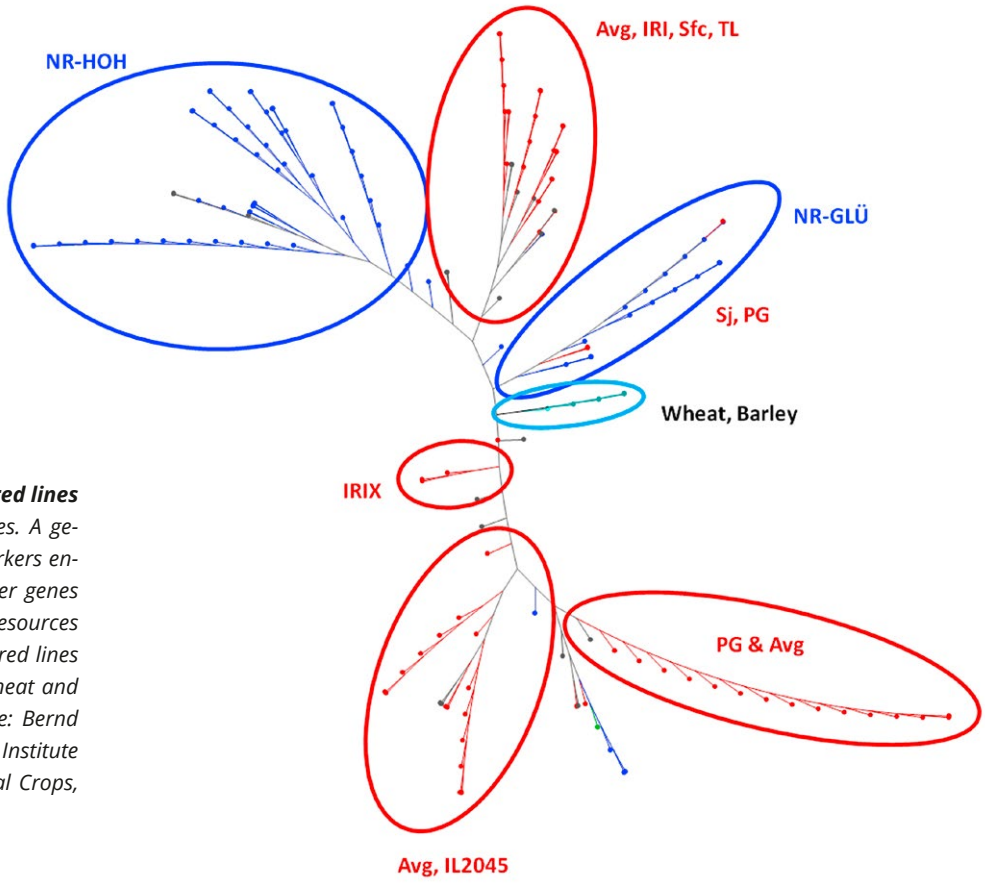


Fig. 4: Phylogenetic tree of rye inbred lines *carrying restorer of fertility (Rf) genes. A genetic fingerprint based on linked markers enables to discriminate efficient restorer genes originating from different genetic resources of rye (red). A set of non-restorer inbred lines (blue) were included as a control, wheat and barley served as outgroups. (Source: Bernd Hackauf, Julius Kühn-Institute (JKI), Institute for Breeding Research on Agricultural Crops, Groß Lüsewitz).*

targeted enrichment of novel selection markers for the 4R segment and allowed the fine mapping of *Rfp*^{Avg}. Based on these markers, a haplotype signature has been established for the restorer gene locus. This molecular approach enabled to differentiate phenotypically indistinguishable restorer genes, which have been trapped from rye genetic resources of different geographic origins (Fig. 4). The established markers render the systematic introgression of restorer genes representing individual haplotypes into elite germplasm with a precision not feasible before. The obtained results open the perspective to examine, whether a linkage drag associated with the restorer gene currently used in practical hybrid rye breeding can also be observed for restorer genes with alternative haplotypes.

In the sub-introgression lines that were developed for all agronomic traits, effects from the previous GABI project could be verified in many cases, although most lines were taller and had lower grain yield than the recurrent parent. One sub-introgression line with a donor segment on chromosome 1R, however, had 6.7 dt/ha higher grain yield in testcrosses, another line with a donor segment on chromosome 5R was significantly shorter. Concerning quality traits, individual sub-introgression lines with higher starch, lower protein and higher pentosane content, needed for superior baking quality, could be detected. For improving feeding quality, two lines with lower soluble pentosane content would be useful. In conclusion, despite the agronomic inferiority of the original Iranian donor population, individual sub-introgression lines displayed traits and trait combinations, including pollen-fertility restoration, that are highly valuable for practical rye breeding.

References

- *Bernal-Vasquez A-M, Möhring J, Schmidt M, Schönleben M, Schön C-C, Piepho H-P (2014) The importance of phenotypic data analysis for genomic prediction - a case study comparing different spatial models in rye. BMC Genomics 15:646*
- *Bernal-Vasquez A-M, Utz H-F, Piepho H-P. (2016) Outlier detection methods for lattices: a case study on the transition from ANOVA to REML Theor and Appl Genet DOI 10.1007/s00122-016-2666-6*

SELECT Selection and identification of molecular markers in specific genomic regions of agronomic importance and for chromosome-specific mapping in allopolyploid wheat for accelerated breeding

Christine Zanke[1], Marion S. Röder[1], Andreas Polley[2], Andrea Eichhorn[2], Eva Graner[2], Jörg Plieske[2], and Martin W. Ganal[2]

1 1Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany
2 TraitGenetics GmbH, Gatersleben, Germany (Project ccordinator)

Figure 1: Examples for wheat lines with different seed size and TGW

Introduction

Molecular markers are nowadays an indispensable tool for plant breeding in all major crop plants. They are mainly used for the analysis of genetic diversity in the entire genome, for marker-assisted backcrossing and for the tagging of individual genes or QTL (quantitative trait loci) along with the development of diagnostic markers for breeding. In order to achieve a higher level of precision during marker-assisted plant breeding increasing numbers of molecular markers are necessary over the entire genome or for the high-density saturation of specific chromosomal regions. Such marker development is especially difficult in allopolyploid crop species such as wheat since this species is hexaploid (consisting of the three genomes ABD with seven chromosomes each), it has a very large genome (approx. 16 Gbp), and nearly 90% of the wheat genome consist of highly repeated sequences which are less suitable for marker development. Because of that, it would be desirable to have a validated genome analysis and marker development system that is specifically focussed onto the single copy fraction (which are mostly the genes) especially for saturating specific genomic regions. The primary purpose of this project is to develop such technologies and test them for the precise mapping of seed size and thousand-grain weight QTL.

Seed quality trait loci and genetic material

Seed size and thousand-grain weight (TGW) are two important traits that constitute yield components in wheat. Some

Table 1: Selection of homozygous BC2F2-Introgression lines for TGW QTL

Population	Trait	No. of selected loci	Chromosome(s)	No. of BC2-F2 lines tested	No. of selected homozygous Introgression lines
Ephoros x Mulan	TGW	3	2D, 4A, 6D	408	180
Mikon x Inspiration	TGW	1	3A	624	43
Interet x Sogood	TGW	1	2D	144	41
Hereford x Expert	TGW	1	1B	132	40

SELECT Selection and identification of molecular markers in specific genomic regions of agronomic importance and for chromosome-specific mapping in allopolyploid wheat for accelerated breeding

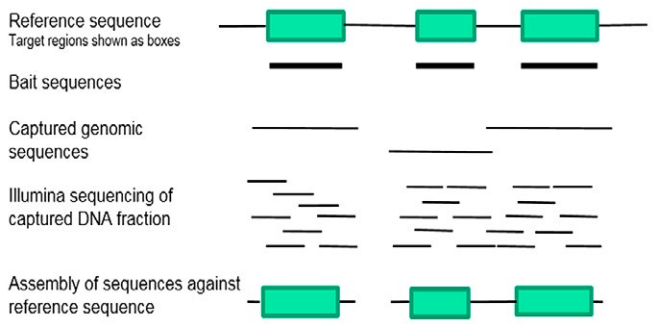


Figure 2: Principles of Sequence Capture

individual loci affecting these traits have been identified earlier and characterized through the development of Nearly Isogenic Lines (NILs) as, for example, a seed size QTL (Figure 1) on chromosome 7D (Röder *et al.* 2008). Within this project, we have also performed an association genetic analysis of phenotypic data for TGW in order to identify additional loci (QTL) for TGW (Zanke *et al.* 2015). A total of 16 loci on 13 wheat chromosomes were selected from these results that contain interesting TGW QTL.

With the long term goal of specifically characterizing, delimiting and fine-mapping some of these 16 TGW QTL on the individual wheat chromosomes, we have worked on the development of NILs for six loci (Table 1) affecting TGW. Specifically, we have performed foreground and background marker-assisted selection during two backcross generations in order to isolate these QTL. At the end of the project, BC_2S_2 material has been generated that contains an individual QTL each in a 95-97% recurrent background which is now ready for field evaluation.

For general and genome-wide mapping purposes, the standard ITMI-DH (International Triticeae Mapping Initiative) population (Sorrells *et al.* 2011) was grown and cross-mapped with selected SSR and SNP markers being located in interesting QTL regions. For precise mapping of selected TGW regions, crosses between varieties with specific QTL were generated. F_2-populations of the respective crosses were grown and will be used for the mapping of SNP markers in the selected QTL regions created with the marker identification procedures described below. These F_2-populations constitute furthermore an interesting resource for genome-wide mapping of SNP markers and comparing their order between different populations and with the wheat genome sequence.

Sequence capture

In the technical part of project, the goal was to establish and employ a system to identify and select specific markers in defined chromosomal regions in order to determine the genetic variability in the genes in such regions in an effective way without sequencing the entire wheat genome. For this, sequence capture techniques should be established and employed for the efficient identification of useful SNPs in wheat. Sequence capture is a technology that is based on the selective capturing of single or low copy genomic sequences in a genome followed by subsequent sequencing of this enriched fraction using high throughput NGS (Next Generation Sequencing). With Sequence Capture, it is possible to limit the analysed fraction of a genome effectively to, for example, only the sequences of the genes or single-copy sequences of specific chromosomal regions and thus reducing the amount of DNA sequencing necessary for the effective identification of molecular markers (Figure 1). For wheat, this means that only a few percent (2-3%) of the genome will actually be analysed but in this fraction almost the entire genetic variability in genes or single copy sequences will be represented.

Within the project, the use of the Sequence Capture approach towards the identification and development of high quality markers in alloployploid wheat should be demonstrated for (i) the specific saturation of chromosomal regions that contain specific QTL associated with the TGW trait for which NILs have been developed and (ii) for the entire wheat genome to saturate specific genomic regions with other QTL.

Sequence Capture for specific genomic regions

In order to test Sequence Capture technologies for specific genomic regions with a previously known DNA sequence or genome sequences from the wheat reference genome, we have applied the Agilent SureSelectXT Target Enrichment System. This system permits the specific selection of single copy genome regions for Sequence Capture. Here, we selected one genomic region on chromosome 7D for which a sequenced BAC contig was available. Through bioinformatic analysis, all single copy sequences in this sequence context have been identified and subsequently used as baits for sequence capture. The other selected regions analyzed with this system contained mainly entire gene sequences present in specific regions with TGW QTL that have been identified by the use of available genome sequences and synteny data with other Gramineae. A total of 24 lines with defined genetic backgrounds were used for Sequence Capture (total captured region of approx. 3 MBp) and subsequent NGS sequencing on the Illumina platform. Bioinformatic analysis identified more than 21000 SNPs in the captured regions and genetic mapping of some markers in the form of individual KASP assays revealed a validation/conversion rate of more than 70% of the identified SNPs. As far as the functional markers could be mapped in populations, most of these markers map to the selected regions. This demonstrates that Sequence Capture can be used for the targeted saturation of genomic regions for which a sequence is available with extremely high numbers of molecular markers and that these markers can be employed in

SELECT Selection and identification of molecular markers in specific genomic regions of agronomic importance and for chromosome-specific mapping in allopolyploid wheat for accelerated breeding

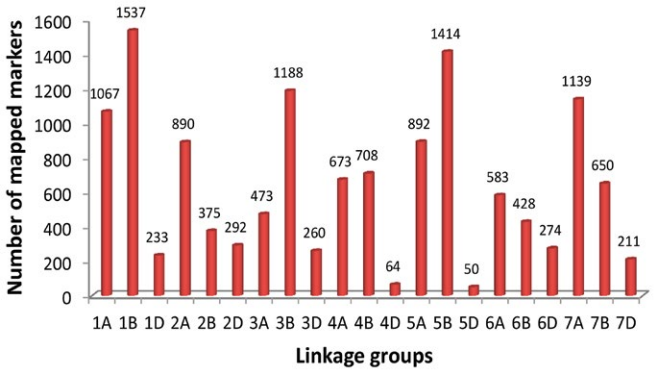


Figure 3: Chromosomal distribution of markers from sequence capture that were mapped in one of the mapping populations via Affymetrix array analysis. Mapped markers on chromosomes with specific QTL are highly enriched.

fine-mapping of regions containing specific seed size and TGW QTL.

Sequence capture with a genome-wide exome capture system

During the course of the project, a Nimblegen whole genome exome capture system became available that contained the exon sequences of most wheat genes in more than 97000 contigs with a total size of 126.9 Mbp. This capture system was employed by TraitGenetics to perform sequence capture on a set of 25 wheat lines that were representative for a wide range of European wheat material on the one side and contained lines that harbored specific seed size and TGW QTL on the other side. After the Sequence Capture, NGS sequencing and bioinformatic analyses, a total of more than 3.1 million high confidence SNPs were identified in the wheat exome of these lines.

Development of a 135k Affymetrix SNP array for high density mapping of SNPs in specific regions of the wheat genome

From the SNPs that were identified in the two different sequence capture projects, we have selected approximately 135000 SNP markers that were mainly located in the respective genomic regions with interesting QTL supplemented by additional markers that were spread over all 21 wheat chromosomes. These SNPs were employed for the design of an Affymetrix SNP genotyping array. The array was subsequently used for the genotyping of DNA from four developed mapping populations and a representative set of European wheat varieties. Analysis of the genotype calls from the Affymetrix array revealed that a very high percentage (>85%) of the selected markers are of high quality in terms of allele calling, >40,000 markers were polymorphic in wheat varieties and, for example, a total of 13,401 of those markers could be mapped onto one of the QTL mapping populations (Figure 3). Current mapping data and chromosomal assignments demonstrate that it has been possible to saturate selected QTL regions with large numbers of SNP markers located in and adjacent to wheat genes. A haplotype analysis revealed that for each QTL, diagnostic marker sets can be established which can be used for following the respective QTL alleles in breeding material.

Summary

We have successfully developed procedures for the targeted saturation of specific regions of the wheat genome that contain interesting seed size and TGW QTL with molecular markers (SNPs). This opens the door to the cost efficient development of markers for specific alleles of interesting genes and QTL. The now established genetic material (introgression lines and segregating populations) can be used in the future for the fine mapping of such QTL in large populations and for the comparison of the genetic maps in these regions to the wheat genome reference sequence. In the long term, this will allow to elucidate the functional basis of the genes that underlie these interesting QTL and for the accelerated use of interesting QTL alleles in breeding with molecular markers.

References

- *Röder, M.S., Huang, X.Q., Börner, A.: Fine mapping of the region on wheat chromosome 7D controlling grain weight. Funct. Integr. Gen. 8: 79-86 (2008)*
- *Sorrells, M.E., Gustafson, J.P., Somers, D., Chao, S., Benscher, D., Guedira-Brown, G., Huttner, E., Kilian, A., McGuire, P.E., Ross, K., Tanaka, J., Wenzl, P., Williams, K., Qualset, C.O.: Reconstruction of the synthetic W7984×Opata M85 wheat reference population. Genome 54: 875-882 (2011)*
- *Zanke, C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., Neumann, F., Eichhorn, A., Polley, A., Jaenecke, C., Ganal, M.W., Röder, M.S.: Analysis of main effect QTL for thousand grain weight in European winter wheat (Triticum aestivum L.) by genome-wide association mapping. Frontiers in Plant Science 6: 644 (2015)*

SUNRISE Genomics assisted breeding in sunflower for better yield potential, stability and efficiency

Silke Wieckhorst[1], Milena Ouzunova[1], Volker Hahn[2], Maren Livaja[3], Chris-Carolin Schön[3], Hartmut Luerßen[4], Martin Ganal[4]

1 KWS SAAT SE, Grimsehlstraße 31, 37555 Einbeck
2 Universität Hohenheim, Fruwirthstraße 21, 70593 Stuttgart
3 Technische Universität München (TUM), Liesel-Beckmann-Straße 2, 85354 Freising
4 TraitGenetics GmbH, Am Schwabeplan 1b, 06466 Stadt Seeland OT Gatersleben

Abstract
The project SUNRISE aims at broadening the spectrum of crops for oil production in European markets to increase biodiversity in agriculture which secures productivity and economic returns. Sunflower cultivation faces challenges from abiotic and biotic factors as production is shifting from areas of high productivity to more marginal areas with lower yield potential. The challenge of sunflower production is thus to breed sunflowers adapted to such environments while still maintaining or increasing seed yield. One of the major limiting factors for growing sunflower in parts of Europe are highly destructive fungal diseases. Because the introgression of new resistance alleles is time-consuming and interferes with genepool management, sunflower breeding is a challenging task.
SUNRISE lays the foundation for genome-based breeding in sunflower which is an important prerequisite for long-term competitiveness of sunflower as a crop. The main target is to improve yield and yield potential in sunflower by adapting innovative breeding strategies aided by high-throughput genotyping technologies, mining for "must have" alleles, and developing a platform for *in vivo* haploid induction. These approaches will lead to an increased selection efficiency and accelerate breeding processes. SUNRISE contributes to a sustainable agriculture by widening the agrobiodiversity in European markets with respect to oil crops for human consumption, industrial uses, and bioenergy.

The SUNRISE consortium of four partners from universities and industry focused on the following topics: i) Genome-wide SNP detection and development of a multiplex array for high-throughput genotyping, ii) Development of F_2 populations for fine mapping of the resistance genes Pl_2, Pl_6 and Pl_8 against *Plasmopara halstedii* and the identification of closely flanking molecular SNP markers, iii) Cloning of the resistance locus Pl_{ARG}, iv) Development of optimized breeding populations with important traits for the restorer and maintainer pool, v) Development of a TILLING resource.

1. Marker discovery and development of an Infinium 25K genotyping array

Before the start of the SUNRISE project in 2010, mainly cost intensive simple sequence repeat (SSR) markers were available for sunflower genotyping (Tang *et al.* 2002). The prerequisite for cost efficient high throughput genotyping are high numbers of markers for genome-wide screening and for the identification of diagnostic markers for marker assisted selection (MAS). This can only be achieved with SNP (single nucleotide polymorphism) markers that can be analyzed in large numbers and in a highly parallel fashion.

Amplicon and whole genome sequencing: For amplicon sequencing, 48 maintainer and restorer lines representative of current elite breeding material were selected. Primers were designed based on 5,955 EST sequences of the *Helianthus annuus* UniGene EST Set (http://www.ncbi.nlm.nih.gov/genbank/dbest). 3,356 well working primer pairs were used for amplification of all genotypes. The amplicons were barcoded for each pool and sequencing was performed using Illumina NGS analysis. Four sunflower inbred lines representing two main groups of the sunflower gene pool (two restorer; SUN48-0003, SUN48-0006 and two maintainer lines; SUN48-0025, SUN48-0026) were selected for whole genome sequencing (WGS).

SNP detection: A multi-step selection procedure was employed to obtain high-confidence bi-allelic SNP markers with stable cluster performance.

1) Discovery of SNPs based on *de novo* assembly including UniGene and read mapping (616,781 SNPs)
2) Haplotype specific SNP selection mainly from the sequenced amplicons (36,382 SNPs)
3) Inclusion of 10,640 pre-validated SNPs (Bachlava *et al.* 2012)
4) Removal of duplicated SNPs with *in silico* Illumina Infinium assay design (22,299 SNPs)
5) Assembly of a final set of variants for a sunflower SNP array validated by genotyping, high quality variants with stable cluster performance (20,502 SNPs)

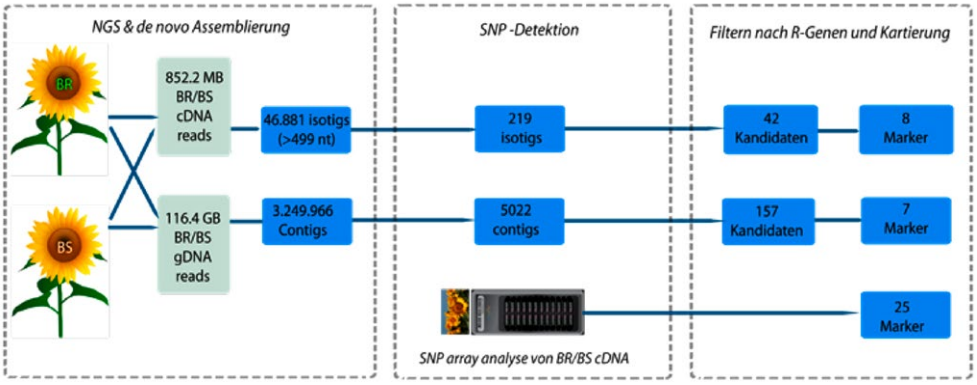


Figure 1: Identification of R gene candidates and molecular markers for the target PI_{ARG}

Genotyping and cluster file development: With the developed SNP Illumina array, a diversity set of lines, hybrids and mapping populations of 1,152 samples was genotyped. A cluster file was developed whereby each SNP marker was evaluated by visual inspection using the GenomeStudio software.

Genetic mapping: Genetic maps were established with the polymorphic SNP markers based on one RIL population and four F2 mapping populations, and an integrated BIN map was calculated which covers 13,503 SNP markers.

Use of SNP array: The genotypic data were used to carry out a diversity study as well as an ADMIXTURE analysis based on a sunflower CoreCollection of 287 samples (Mandel *et al.* 2013). The four groups of nonoil restorer, oil restorer, nonoil maintainer and oil maintainer could be clearly distinguished. QTL analysis and genome-based prediction were performed for 113 RILs of the population NDBLOSselxCM625. Phenotypic data of one morphological and three resistance traits against the pathogen *Sclerotinia sclerotiorum* were included from a previous study (Micic *et al.* 2005). Two major QTL for *Sclerotinia* tolerance were identified (Livaja *et al.* 2015).

2. Development of optimized breeding populations for joint linkage association mapping

Optimized breeding populations are an important prerequisite for the mapping of important traits. Such populations were established for the maintainer and restorer pool using a factorial design since this design has a higher power for QTL detection (Stich 2009). Four elite sunflower maintainer lines were crossed with eight lines carrying diverse traits (resistance against *Orobanche*, resistance against *Plasmopara*, drought tolerance, high oleic content) which are derived from genetic resources. The same was done for the restorer pool. For all populations, RILs were developed via selfing until F5 generation (maintainer pool 3,445 RILs and for the restorer pool 3,000 RILs). This material represents an outstanding genetic resource for further research into the detection of QTL for important breeding traits.

3. Fine-mapping "must have" alleles and cloning of a gene conferring resistance against Plasmopara halstedii

3.1. Fine-mapping PI_2, PI_6 and PI_8

Three F_2 populations were developed for the resistance genes PI_2 (728 F_2 individuals), PI_6 (728 F2 individuals), PI_8 (385 F_2 individuals) against Plasmopara halstedii. For each population, 100 F2:3 families were phenotyped with the race 304 (PI_2) or 710 (PI_6, PI_8). Developed SNP markers were used for genotyping of the three F_2 populations. In depth phenotyping was done for recombinant F_2 individuals to fine map the position of the resistance loci PI_2, PI_6 and PI_8. For all three resistance genes, closely flanking SNP markers were identified which allow an efficient and fast selection of resistance genes against Plasmopara halstedii as well as the pyramiding of several resistance genes in one sunflower genotype.

3.2. Fine mapping and cloning of the PI_{ARG} locus

Fine mapping of PI_{ARG}: PI_{ARG} is a resistance locus which protects sunflower against all known races of Plasmopara halstedii. PI_{ARG} was mapped on linkage group 1 (Dußle *et al.* 2004, Wieckhorst *et al.* 2010). For fine mapping and cloning of the resistance locus PI_{ARG}, the two F_2 populations (cms) HA342xARG1575-2 (2,141 F_2 individuals) and ARG1575-2xRHA265 (2,344 F_2 individuals) were used. The F_2 population ARG1575-2xRHA265 developed within the SUNRISE project has a higher recombination frequency than the F_2 population (cms)HA342xARG1575-2. The F_2 population ARG1575-2xRHA265 was genotyped with markers from linkage group 1. In this process, a total of 170 recombinant inbred lines were identified. The F_2 recombinant plants were analyzed with further markers from a bulked-segregant-transcriptome-analysis (BSTA) (Livaja *et al.* 2013, Figure 1) and SNP markers from Bachlava *et al.* 2012 which were detected via analysis of a susceptible and resistant bulk of F2 individuals generated from the population (cms) HA342xARG1575-2 (Figure 1). The $F_{2:3}$ families (or $F_{3:4}$ families) of 111 recombinant F_2 individuals were phenotyped with race 730 and 304 to fine map the target region of PI_{ARG}. Ten molecular markers were identified which cosegregate with PI_{ARG} and which can be used for MAS.

Cloning of PI_{ARG}: To identify PI_{ARG} gene candidates and to develop further molecular markers, a further bulk segregant approach was performed. F_2 individuals of population ARG1575-2xRHA265 were selected to develop a susceptible bulk of 14 F_2 individuals and a resistant bulk of 17 F_2 individuals which showed recombination events above or below the PI_{ARG} target region.

Genomic DNA of the two bulks of ARG1575-2xRHA265, the resistant bulk of (cms)HA342xARG1575-2, RHA265, HA342, ARG1575-2 (resistance donor), and ARG1575-4 (susceptible) was used for whole genome sequencing on Illumina HiSeq 2000. For that, two types of sequencing libraries were generated: shotgun libraries (SG) with 250-500 bp insert size and long jumping distance libraries (LJD) of 3 kbp fragment size.

For each sample, a *de novo* assembly and a SNP calling was performed comparing the sequences of the resistant and susceptible bulk (Figure 1). In total, 5,022 contigs were identified. Blast analysis against the Plant Resistance Gene database (PRGdb) resulted in 157 contigs with similarity to known R genes. A Blast2GO analysis was performed to determine the putative function of the R gene candidates. Fifteen contigs matched to the class of receptor (-like) and protein kinases and 11 contigs showed similarity to the candidate gene RGC151 (Wieckhorst *et al.* 2010) of the TIR-NBS-LRR class. SNPs between the parental lines ARG1575-2 and RHA265 were confirmed for seven contigs and used for mapping on linkage group 1.

The molecular marker RGC151 cosegregates with PI_{ARG} (Wieckhorst *et al.* 2010). Via Sanger sequencing, five copies of the candidate gene were identified for the susceptible parent RHA265 as well as five candidate genes for the resistance donor ARG1575-2. Molecular markers were developed based on the identified sequences to integrate the candidate genes into the genetic map. Two copies of the resistant parent ARG1575-2 and one copy of the susceptible parent have an open reading frame (ORF), and code for putative proteins of the TIR-NBS-LRR class. Constructs for sunflower transformation were developed for the two candidate genes of ARG1575-2 which are currently under evaluation.

4. Development of a TILLING resource

In order to identify induced mutations for specific sunflower genes, a TILLING (targeting induced local lesions in genomes) population was developed based on three sunflower genotypes RHA857, NDRLOS and HA89. Seeds (0.75%, 6h and 18h; 1%, 6h) and pollen (0.75%, 6h) were treated with different concentrations of Ethyl-methanesulfonate (EMS). As a result of that approach, a long-term storage TILLING population of 2,875 M1 plants has been constructed for screening to identify functional variation in genes of interest.

5. Summary of the SUNRISE project

Within the SUNRISE project, major genomic and genetic resources were developed for sunflower that is an important prerequisite for accelerated sunflower breeding. The developed 25K SNP array covers SNPs which are almost evenly distributed across 17 linkage maps. We showed that the array is very suitable for genetic diversity analysis, QTL and genome-based prediction approaches. With the development of closely linked molecular markers for the "must have" resistance genes against *Plasmopara halstedii* a fast and easy selection of these genes as well as pyramiding of resistance gene is possible. Thus, the SNP array and SNP markers for "must have" traits enable the development of more efficient breeding strategies. The optimized breeding populations and the TILLING resource will support the identification of further important traits and closely linked molecular markers for selection in the future. All these new tools and resources pave the way for effective genomics-assisted breeding in sunflower.

References

- *Bachlava E, Taylor CA, Tang S, Bowers JE, Mandel JR, Burke JM, Knapp SJ (2012) SNP discovery and development of a high-density genotyping array for sunflower. PLoS One 7: e29814*
- *Dußle CM, Hahn V, Knapp SJ, Bauer E (2004) PlARG from Helianthus argophyllus is unlinked to other known downy mildew resistance genes in sunflower. Theor Appl Genet 109 (5):1083-1086*
- *Livaja M, Wang Y, Wieckhorst S, Haseneyer G, Seidel M, Hahn V, Knapp S, Taudien S, Schön C-C, Bauer E (2013) BSTA: a targeted approach combines bulked segregant analysis with next- generation sequencing and de novo transcriptome assembly for SNP discovery in sunflower. BMC Genomics 14 (1):628*
- *Livaja M, Unterseer S, Erath W, Lehermeier C, Wieseke R, Plieske J, Polley A, Luerßen H, Wieckhorst S, Mascher M, Hahn V, Ouzunova M., Schön C-C, Ganal MW (2015) Diversity analysis and genomic prediction of Sclerotinia resistance in sunflower using a new 25K SNP genotyping array. Theor Appl Genet DOI: 0.1007/s00122-015-2629-3*
- *Mandel JR, Nambeesan S, Bowers JE, Marek LF, Ebert D, Rieseberg LH (2013) Association mapping and the genomic consequences of selection in sunflower. PLoS Genet 9(3):e1003378*
- *Micic Z, Hahn V, Bauer E, Schön CC, Melchinger AE (2005) QTL mapping of resistance to Sclerotinia midstalk rot in RIL of sunflower population NDBLOSsel × CM625. Theor Appl Genet 110:1490–1498.*
- *Stich B (2009) Comparising of mating designs for establishing nested association mapping populations in maize and Arabidopsis thaliana. Genetics 183:1525-1534*
- *Tang S, Yu JK, Slabaugh MB, Shintani DK, Knapp SJ (2002) Simple sequence repeat map of the sunflower genome. Theor Appl Genet 105:1124-1136*
- *Wieckhorst S, Bachlava E, Dussle CM, Tang S, Gao W, Saski C, Knapp SJ, Schön CC, Hahn V, Bauer E (2010) Fine mapping of the sunflower resistance locus PlARG introduced from the wild species Helianthus argophyllus. Theor Appl Genet 121 (8):1633-1644*

VALID Validation and Identification of Important Marker-Trait Associations for the Development of Improved Wheat Varieties

Marion S. Röder[1], Jie Ling[1], Christine Zanke[1], Bernd Rodemann[2], Christiane Balko[3], Frank Ordon[4], Sonja Kollers[5], Viktor Korzun[5], Erhard Ebmeyer[5], Odile Argillier[6], Gunther Stiewe[7], Thomas Zschäckel[7], Jörg Plieske[8], Martin W. Ganal[8]

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Stadt Seeland OT Gatersleben, Germany
2 Julius Kühn Institute (JKI), Braunschweig, Germany
3 JKI, Groß-Lüsewitz, Germany
4 JKI, Quedlinburg, Germany
5 KWS LOCHOW GMBH, Bergen, Germany
6 Syngenta France S.A.S., Orgerus, France
7 Syngenta Seeds GmbH, Bad Salzuflen, Germany
8 TraitGenetics GmbH, Stadt Seeland OT Gatersleben, Germany (Project coordinator)

Introduction

Wheat is the most important crop in Germany based on acreage. In 2013, 3.1 million hectares of wheat were grown and 25 million tons were harvested (http://faostat.fao.org). The average yield was approximately 8.0 t/ha. While there was a steady increase of yield in wheat in the 1980's and 1990's, a plateau has been reached by the beginning of the new century and high fluctuations were observed in different years since then (Figure 1). These have been mainly attributed to environmental effects, such as drought, heat and frost. Therefore, the availability of high yielding varieties which provide stable yields under varying climatic conditions is of utmost importance for plant breeders and farmers.

The goal of the VALID project was to apply genome wide association studies (GWAS) to identify major genetic loci for yield and other important agronomic parameters, as well as for tolerance to biotic and abiotic stresses in recent European winter wheat varieties. During the association analysis, phenotypic data from field evaluations will be combined with genotypic data based on molecular markers (Figure 2). By applying appropriate statistical models, it will be possible to calculate so-called marker-trait associations (MTAs), which link specific molecular markers to specific regions in the genome. QTL (quantitative trait loci) regions contributing to complex traits, such as grain yield, can then be identified. However, the whole analysis is highly dependent on environmental factors, genetic material and population structure of the material. Therefore validation strategies are required, before the detected significant markers can be applied in plant breeding and variety development.

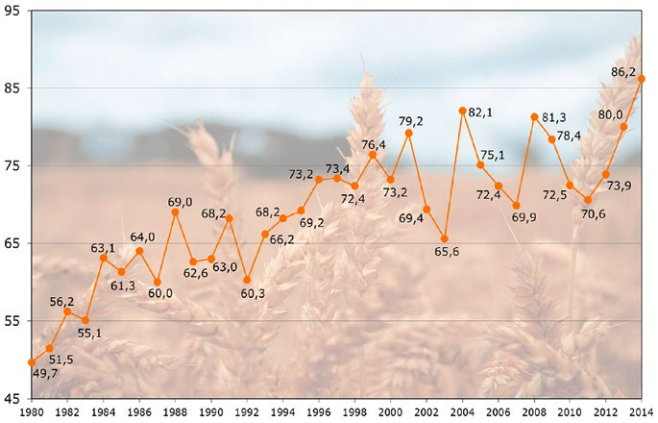


Figure 1: Historical development of yield in German winter wheat.

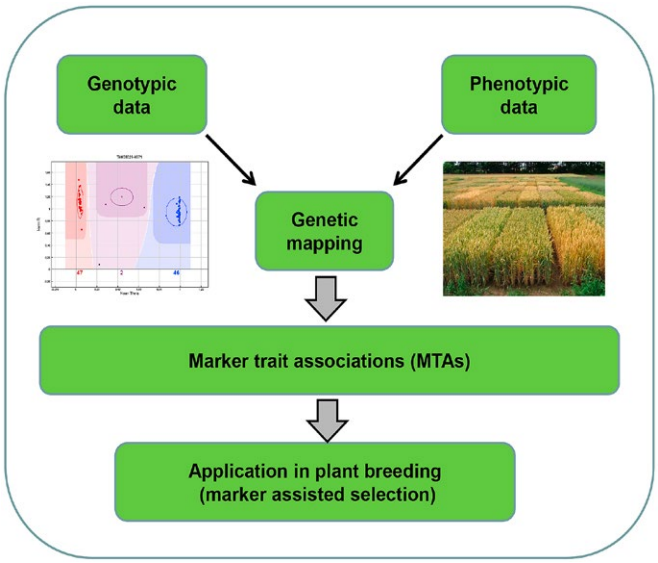


Figure 2: Association analysis allows the combination of phenotypic and genotypic data and results in marker-trait associations which can be applied in plant breeding.

VALID Validation and Identification of Important Marker-Trait Associations for the Development of Improved Wheat Varieties

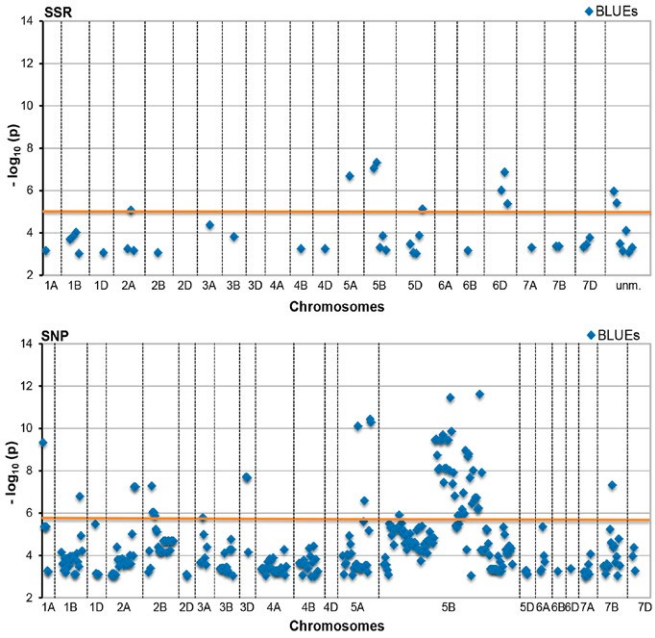


Figure 3: Manhattan plots for heading date depict significant marker-trait associations for SSR (top) and SNP markers (bottom). Blue diamonds represent best linear unbiased estimations (BLUEs) across all eight environments. The x-axes show the 21 chromosomes of wheat.

Detection of marker-trait associations for yield and other traits

The initial analysis was based on a set of 358 winter wheat varieties plus 14 spring wheat varieties (GABI-WHEAT panel) representing the European wheat breeding gene pool. Phenotyping in the field was conducted by the industrial partners KWS LOCHOW GMBH and Syngenta at a total of eight locations in Germany and France over two years. For genotyping a set of 732 genome-wide SSR (simple sequence repeats) was initially used which was now complemented by newly available SNP-markers (single nucleotide polymorphism). For this, the company TraitGenetics GmbH generated data with several genotyping arrays including a 90k iSELECT ILLUMINA SNP array resulting in 21200 polymorphic markers of which 7761 were mapped on a reference map. Additionally, a 35k Affymetrix SNP array has been used resulting in 20052 polymorphic markers of which 4231 could be mapped. Together, both SNP arrays produced 16688 unique patterns (=haplotype blocks) across all varieties and provided a comprehensive genome coverage. A mixed linear model was applied to combine the marker data with the field data for a large set of individual traits. Figure 3 shows significant MTAs for heading date based on field data in eight environments and SSR and SNP markers (Zanke *et al.* 2014a). Similar analyses were conducted for plant height (Zanke *et al.* 2014b) and thousand grain weight (Zanke *et al.* 2015). In all cases up to several hundred markers/loci were identified to be associated with a given agronomic trait.

Figure 4: Rain-out shelter experiment in Groß-Lüsewitz.

In addition to the agronomic field data, the Julius Kühn Institute (JKI) in Braunschweig conducted resistance tests for various important fungal pathogens. These data were analyzed in the same way as the agronomic data. Significant MTAs were discovered for resistance to *Fusarium* head blight (Kollers *et al.* 2013a; Jiang *et al.* 2015), resistance to *Septoria* blotch caused by *Mycosphaerella graminicola* (Kollers *et al.* 2013b), resistance to tan spot caused by *Pyrenophora tritici-repentis* (Kollers *et al.* 2015) and resistance to eyespot caused by *Oculimacula yallundae* (syn. *Pseudocercosporella herpotrichoides*). Based on such data, the investigated material could, for example, be clearly classified with respect to different types of introgressions.

Rain-out shelter tests for tolerance to drought

In the era of climate change, yield stability under varying environmental conditions and the tolerance to drought are of increasing importance for variety development. Therefore, a subset of the varieties was tested in rain-out shelters with controlled irrigation at the JKI in Groß- Lüsewitz in 2013 and 2014 (Figure 4). Drought conditions were applied before anthesis when the flag leaf was visible. The same varieties were tested in parallel in the field at the same location. Grain yield, yield components and phenological traits, as well as physiological traits, such as chlorophyll content, accumulation of free proline and total content of protein, starch and soluble sugars were assembled. Several yield-related traits were strongly affected by the drought treatment including yield, biomass above ground, ears per plot and grains per plot at harvest. A preliminary association analysis yielded significant markers related to the drought stress conditions. However, further phenotypic testing in future years will be necessary to confirm the data.

Validation of the detected MTAs

For validation of the initially detected MTAs, an additional panel of varieties called VALID with 137 varieties was established and tested for yield, heading date, plant height, thousand-kernel weight and resistance to *Fusarium* head blight. The results were compared to the initial GABI-WHEAT panel of 372 varieties. The validation rates, i.e. the con-

firmation rate for the identified loci was only moderate to low. Low significances in the VALID panel were most likely caused by the smaller number of accessions compared to the initial panel and by fluctuating phenotypic field testing due to environmental differences. Nevertheless, a number of loci could be identified that can be used for marker-assisted selection in wheat breeding in the future.

As an additional strategy, MTAs identified in the WHEAT project were selected for validation in specifically developed bi-parental and introgression populations. For this purpose, the two breeding companies developed three doubled-haploid populations segregating for specific genomic regions, in which MTAs for yield had been identified. Furthermore, a number of backcross populations were developed by all partners, in which specific genomic regions associated with yield, resistance to *Fusarium* and *Septoria* were introgressed into another elite variety. The long-term goal is to test these newly developed lines for the effect caused by the specifically introgressed regions in order to verify the MTAs obtained in the association panel of varieties. This could lead directly to an improvement of already existing varieties for the respective traits. All backcross generations were monitored with molecular markers for the presence of the introgressed fragments and with genotyping arrays for the genetic background. Currently, the populations are in the BC2 stage with more than 90% recurrent parent genome and are now propagated and tested in the field.

General conclusions

The data and analyses performed in the project VALID provide a solid fundament for the genetic analysis of several key traits in European winter wheat, including yield and yield-related parameters, quality parameters, tolerance to several important fungal pathogens and preliminary data about drought stress tolerance. Up-to-date genome-wide genotyping technologies including the application of a 90k iSELECT ILLUMINA and a 35k Affymetrix SNP array were combined with solid and repeated phenotyping in the field. The concepts for validation of the obtained MTAs are novel, but require a long term genetic approach or the further characterization and isolation of individual loci/genes that affect these traits. The analysis provides also an extremely valuable resource for the improvement of wheat yield through various genetic improvement strategies such as targeted introgressions and genomic selection. Overall the outcome of the project VALID can be regarded as pioneering in the analysis of the genetic architecture of European winter wheat and place the VALID consortium at the forefront of genetic analysis of winter wheat in Europe. This has been demonstrated by numerous publications in international peer-reviewed journals and several oral presentations at international conferences.

References

- *Jiang, Y., Zhao, Y., Rodemann, B., Plieske, J., Kollers, S., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Ling, J., Röder, M. S., Ganal, M.W., Mette, M.F., Reif, J.C.: Potential and limits to unravel genetic architecture and predict the variation of Fusarium head blight resistance in European winter wheat (Triticum aestivum L.). Heredity 114:318-326 (2015)*
- *Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M.W., Röder, M.S.: Whole genome association mapping of Fusarium head blight resistance in European winter wheat (Triticum aestivum L.). Plos One 8: e57500 (2013a)*
- *Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M.W., Röder, M.S.: Genetic Architecture of resistance to Septoria tritici blotch (Mycosphaerella graminicola) in European winter wheat. Mol Breed 32:411-423 (2013b)*
- *Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M.W., Röder, M.S.: Genome wide association mapping of tan spot resistance (Pyrenophora tritici-repentis) in European winter wheat. Mol Breed 34:363-371 (2014)*
- *Zanke, C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Argillier, O., Stiewe, G., Hinze, M., Beier, S., Ganal, M.W., Röder, M.S.: Genetic architecture of main effect QTL for heading date in European winter wheat. Frontiers in Plant Science 5:217 (2014a)*
- *Zanke, C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., Neumann, K., Ganal, M.W., Röder, M.S.: Whole genome association mapping of plant height in winter wheat (Triticum aestivum L.). PLOS ONE 9:e113287 (2014b)*
- *Zanke, C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., Korzun, V., Argillier, O., Stiewe, G., Hinze, M., Neumann, F., Eichhorn, A., Polley, A., Jaenecke, C., Ganal, M.W., Röder, M.S.: Analysis of main effect QTL for thousand grain weight in European winter wheat (Triticum aestivum L.) by genome-wide association mapping. Frontiers in Plant Science 6:644 (2015)*

Impressum

Herausgeber
PLANT 2030 Geschäftsstelle
c/o Max-Planck-Institut für Molekulare Pflanzenphysiologie
Am Mühlenberg 1 · 14476 Potsdam
E-Mail: PLANT2030@mpimp-golm.mpg.de

PLANT 2030 Geschäftsstelle
Dr. Matthias Arlt, Geschäftsstellenleiter
Dr. Hanna Berger
Dr. Christiane Hilgardt
Dr. Henrike Perner
Joram Schwartzmann
Juliane Vosswinkel

Redaktion
Dr. Christiane Hilgardt

Layout
Dirk Biermann Grafk Design, Potsdam

ISBN 978-3-947237-97-5

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung