ORIGINAL PAPER

Genome-wide association mapping reveals epistasis and genetic interaction networks in sugar beet

Tobias Würschum · Hans Peter Maurer ·
Britta Schulz · Jens Möhring · Jochen Christoph Reif

Abstract Epistasis is defined as interactions between alleles of two or more genetic loci. Detection of epistatic interactions is the key to understand the genetic architecture and gene networks underlying complex traits. Here, we examined the extent of epistasis for seven quantitative traits with an association mapping approach in a large population of elite sugar beet lines. We found that correction for population stratification is required and that in terms of reducing the false-positive rate the mixed model approach including the kinship matrix performed best. In genome-wide scans, we detected both main effects and epistatic QTL. For physiological traits, the detected digenic and higher-order epistasis explained a considerable proportion of the genotypic variance. We illustrate that the identified epistatic interactions define comprehensive genetic networks, which may serve as starting points towards a systems-oriented approach to understand the regulation of complex traits.

T. Würschum (✉) · H. P. Maurer · J. C. Reif
State Plant Breeding Institute, University of Hohenheim,
70593 Stuttgart, Germany
e-mail: tobias.wuerschum@uni-hohenheim.de

J. Möhring
Bioinformatics Unit, Institute for Crop Production and Grassland
Research, University of Hohenheim, 70593 Stuttgart, Germany

B. Schulz
KWS SAAT AG, 37574 Einbeck, Germany

Introduction

Epistasis refers to interactions between alleles from two or more genetic loci (Carlborg and Haley 2004; Phillips 2008). The consequence of epistasis is that the phenotype of an individual cannot be predicted simply by the sum of the single-locus effects, but rather depends on the specific combinations of loci (Lynch and Walsh 1998). Recent molecular and biochemical work points to a vast number of interactions between genes and gene products (Tong et al. 2004; St Onge et al. 2007; He et al. 2010; Costanzo et al. 2010). The complexity of quantitative traits and the observed extensive interplay of factors implicated in these pathways suggest that these networks should be reflected in a similar amount of epistasis on a genetics level. Studies in animal systems have revealed major effects of epistasis on the genetic architecture and the variation of complex traits (Montooth et al. 2003; Carlborg et al. 2006). In plants, epistasis has been reported only in few studies in barley, rice and tomato (Li et al. 1997; Eshed and Zamir 1996; Xu and Jia 2007).

A major limitation of classical linkage mapping is that it is based on biparental crosses and evaluates only two alleles at a time. The detection of epistatic interactions relies on two interacting loci being polymorphic in that particular cross. In contrast, association mapping is based on diverse populations with a low frequency of fixed loci. The statistical power to detect epistasis depends on the size of the genotypic classes determined by the allele frequencies and subsequently on the population size. Detection of epistasis in association mapping therefore represents a complementary approach to linkage mapping studies.

A potential problem of association mapping is that any non-functional association between the trait and the underlying population structure will also be detected as

QTL. Population stratification can therefore lead to the discovery of many false-positive QTL (Zhao et al. 2007). Recently, different methods have been suggested which correct for population structure (*P* matrix) and familial relatedness (*K* matrix) (Price et al. 2006; Yu et al. 2006). The *P* matrix uses principal components and the *K* matrix is based on kinship estimates. Both, the *P* and *K* matrix, are commonly estimated with marker data, which can lead to an overcorrection for population stratification resulting in a low power to detect QTL. Therefore, a careful choice of an appropriate association mapping method is pivotal.

Sugar beet is of major economical importance for worldwide sugar production (Draycott 2006) and serves as an outstanding model crop to study the genetic architecture of physiological and agronomic traits with varying complexity. Linkage mapping and candidate-region association mapping approaches have identified main effect QTL underlying yield and physiological traits such as sodium, potassium and α-amino nitrogen content (Weber et al. 2000; Schneider et al. 2002; Stich et al. 2008a). Nevertheless, none of the previous studies investigated the contribution of epistasis to the genetic variation of relevant physiological and agronomic traits.

In this paper, we evaluated the applicability of association mapping in a large, experimental data set in sugar beet. We investigated different methods to correct for population stratification and found that the mixed model approach including the *K* matrix performed best. Main effect QTL were identified for seven quantitative traits. We focused on the detection of epistasis and identified two- and three-way epistatic interactions. We show that the detection of epistasis can be used to unravel the genetic architecture and gene networks of complex traits.

Materials and methods

Plant materials, field experiments and molecular markers

Our study was based on 460 elite sugar beet (*Beta vulgaris* L.) inbred lines, which were randomly derived from 112 crosses among 22 diploid sugar beet clones (Supplementary Figure S1). The number of progenies from each cross ranged from two to seven. Testcross progenies were produced by crossing the genotypes to a single diploid tester. All material used in this study was provided by the breeding company KWS SAAT AG (Einbeck, Germany).

The 460 genotypes were evaluated in routine plant breeding trials (α-lattice design) with two replicates at five locations in Germany in 2008. Data were recorded for seven traits as percentage of the mean performance of four checks based on lattice-adjusted entry means. The

evaluated traits were white sugar yield (WSY), sugar yield (SY), sugar content (SC), beet yield (BY), potassium (K), sodium (Na) and α-amino nitrogen (N).

The 460 genotypes were fingerprinted following standard protocols with 290 single nucleotide polymorphism (SNP) markers. These markers were randomly distributed across the sugar beet genome with an average marker distance of 3.5 cM and a maximum gap between adjacent markers of 18 cM. Map positions of all markers were based on the linkage map of KWS SAAT AG with a total map length of 10.23 M (unpublished data).

Phenotypic data analyses

The analyses were based on adjusted entry means calculated for each location. The following linear mixed model was used to estimate variance components of the testcrosses: $y_{ij} = \mu + l_j + g_i + e_{ij}$, where y_{ij} is the adjusted entry mean of the ith sugar beet line at the jth location, μ the intercept term, l_j the effect of the jth location, g_i the genetic effect of the ith sugar beet line and e_{ij} the error term including the genotype times location interaction effect. Locations were modeled as fixed effect and the genetic effect as a random effect. Variance components were determined by the restricted maximum likelihood (REML) method. Significance for variance component estimates was tested by model comparison with likelihood ratio tests where the halved P values were used as an approximation (Stram and Lee 1994). Heritability on an entry-mean basis was calculated as the ratio of genotypic to phenotypic variance according to Melchinger et al. (1998). Furthermore, genotypes were regarded as fixed effects and best linear unbiased estimates (BLUEs) were determined for all genotypes and traits. Phenotypic correlation coefficients (*r*) were calculated among all traits based on BLUEs of the 460 genotypes.

A Mantel test (Mantel 1967) was performed between the phenotypic and genetic distance matrices to estimate the correlation between the phenotype and the population structure. The phenotypic distance matrix was estimated as all pairwise absolute differences between BLUEs of the 460 genotypes. The genetic distances were calculated based on the modified Rogers' distances (Wright 1978).

Association mapping

A two-step association approach was applied which has been shown to posses only a slightly reduced power for the detection of marker-phenotype associations compared to a one-step approach (Stich et al. 2008a). The mixed model for the association mapping approaches was: $y_{ijp} = \mu + l_j + a_p + g_i + e_{ijp}$, where y_{ijp} is the adjusted entry mean of the ith sugar beet line at the jth location carrying allele p, μ

the intercept term, l_j the effect of the jth location, a_p the effect of allele p, g_i the genetic effect of the ith sugar beet line and e_{ijp} the residual. Locations and a_p were modeled as fixed effect whereas g_i was regarded as random effect.

Association mapping populations often show a non-functional correlation between the phenotype and the population structure, which will lead to the detection of false-positive QTL. Therefore, different methods to correct for population structure were tested. The P10 model included the first 10 principal coordinates based on the modified Rogers' distances as fixed effects (Price et al. 2006). In the K model, the variance of the random genetic effect was assumed to be $Var(g) = K\sigma_g^2$, where σ_g^2 refers to the genetic variance estimated by REML and K was a 460×460 matrix of kinship coefficients that define the degree of genetic covariance between all pairs of entries. We followed the suggestion of Bernardo (1993) and calculated the kinship coefficient K_{ij} between inbreds i and j on the basis of marker data as $K_{ij} = 1 + (S_{ij} - 1)/(1 - T_{ij})$, where S_{ij} is the proportion of marker loci with shared variants between inbreds i and j, and T_{ij} is the average probability that a variant from one parent of inbred i and a variant from one parent of inbred j are alike in state, given that they are not identical by descent. The coefficient T_{ij} was estimated using a REML method suggested by Stich et al. (2008b) setting negative kinship values between inbreds to zero. The P10K model included both, the first 10 principal coordinates and the K matrix.

For the detection of main effect QTL, a genome-wide scan for marker-trait associations was conducted. Markers with a P value < 0.01 were selected. To correct for co-linearity, the selected markers were simultaneously fitted in a final model in the order of their P values. Markers with significant ($P < 0.05$) association in the final model were declared as main effect QTL. A two-dimensional genome scan was performed and all possible marker–marker interactions were tested. The model included the detected main effect QTL as co-factors as well as the main and interaction effects of the marker pair under consideration. The Bonferroni–Holm procedure (Holm 1979) was applied to correct for multiple testing. Main effect QTL and two-way interactions, which were significant at a Bonferroni-corrected $P < 0.01$ ($P < 2.4\mathrm{e}{-7}$), were subsequently fitted in a final model to correct for co-linearity. Marker interactions with significant ($P < 0.05$) association in the final model were declared as two-way epistatic QTL. No full three-dimensional genome scan was done, because of the immense computational demand. Nevertheless, three-way interactions were identified based on the results of the two-dimensional scan. The model included the detected main effect and two-way epistatic QTL as co-factors, the three-way interaction to be tested, and the subordinated two-way interactions and single markers. Three-way interactions,

which were significant at a Bonferroni-corrected $P < 0.01$ ($P < 3.5\mathrm{e}{-5}$), were fitted in a final model together with the detected main and two-way epistatic QTL. Three-way interactions, which were still significant in the final model at $P < 0.05$ were considered three-way epistatic QTL.

The total proportion of genotypic variance (p_G) explained by the detected QTL was calculated by fitting all QTL simultaneously to obtain R_{adj}^2. The ratio $p_G = R_{adj}^2/h^2$ yielded the proportion of genotypic variance (Utz et al. 2000). The genotypic variance of the epistatic QTL was obtained as the difference in p_G between the full model and a model without the epistatic QTL. All mixed model calculations were performed using the software ASReml 2.0 (Gilmour et al. 2006).

Linkage disequilibrium (LD) was assessed by the LD measure r^2 (Weir 1996) and significance of LD was tested with Fisher's exact tests (Hill and Robertson 1968). LD computations were performed with the software package Plabsoft (Maurer et al. 2008).

Results

Composition of the data set

We analyzed an association mapping data set composed of 460 diploid sugar beet inbred lines. Plants were phenotyped at five locations resulting in high heritabilities and significant genotypic variances for the seven traits investigated in our survey (Table 1). The mapping resolution in our data set was determined by analyzing the extent of linkage disequilibrium (LD) in the population of the parental lines. Analysis of LD revealed a high amount of intrachromosomal LD and low interchromosomal LD (Fig. 1a). The plot of LD versus genetic map distance showed that LD decayed within approximately 10 cM (Fig. 1b). Consequently, the applied average marker distance of 3.5 cM provided sufficient marker coverage for genome-wide association mapping.

Models for association mapping and population structure

For markers that are unlinked to functional polymorphisms controlling the trait, it is expected that the P values are uniformly distributed and, thus, follow the diagonal line plotting observed versus expected P values (Yu et al. 2006). As markers will also be linked to QTL, we expected for traits with comparable complexity, a similar magnitude of the bulge at the beginning of the diagonal. Mantel tests between phenotypic and genetic distance matrices revealed for all traits small, yet significant associations (Table 1). The simple model, which does not include any correction

Table 1 Mean, minimum and maximum testcross performance of 460 sugar beet genotypes in percentage of the mean performance of the checks, variances of the testcrosses for genotypes (σ_G^2) and residuals (σ_E^2) and heritabilities (h^2) for white sugar yield (WSY), sugar yield (SY), sugar content (SC), beet yield (BY), potassium content (K), sodium content (Na), α-amino nitrogen content (N)

Parameter	WSY	SY	SC	BY	K	Na	N
Mean	95.66	95.91	103.00	93.18	114.10	104.60	107.30
Min	82.43	81.75	96.40	75.30	96.39	70.85	84.95
Max	106.00	106.40	110.00	107.10	139.90	158.30	143.30
σ_G^2	11.78**	13.84**	6.51**	32.44**	36.06**	145.93**	50.31**
σ_E^2	12.43	11.46	2.13	10.27	29.98	230.34	103.20
h^2	0.83	0.86	0.94	0.94	0.86	0.72	0.66
r	0.31**	0.32**	0.25**	0.34**	0.09**	0.09**	0.12**

Correlation (r) between pairwise phenotypic differences of all genotypes and the corresponding genetic distances

** Significance at $P < 0.01$
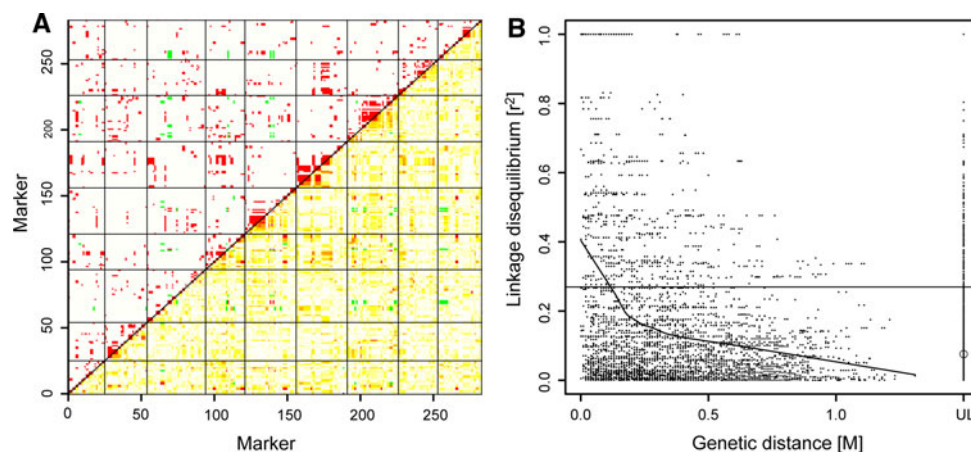
Fig. 1 Linkage disequilibrium (LD) structure measured as r^2 in the parental population. **a** Significant LD ($P < 0.05$, above diagonal) and LD (r^2) between all pairs of loci (below diagonal). The *horizontal and vertical lines* separate the chromosomes, *red coloring* indicates significant LD and higher R values ($R \geq 0.1$), respectively, *white* indicates non-significant LD or $R = 0$ and *green* shows combinations that could not be calculated. **b** LD as a function of genetic distance. *Horizontal line* refers to the 95th percentile of r^2 estimates between unlinked markers. Curve was fitted by robust locally fitted regression and shows that LD decays with genetic distance. (*UL* unlinked loci) (color figure online)

for relatedness, resulted in huge differences with more pronounced bulges for traits with a higher correlation with the population structure (Fig. 3). This clearly indicated an inflated false-positive rate. Therefore, correcting for relatedness is mandatory to avoid a high number of false-positive associations.

The genotypes included in the data set were a sample of elite lines belonging to the same heterotic group, but classified either as sugar or yield type. In accordance with this, principal coordinate analysis revealed no clearly distinct subpopulations but a family structure of sugar and yield types (Fig. 2). This suggested that correction for population structure was not necessary but family relatedness had to be considered (Fig. 3). The use of the P10K model can potentially result in an overcorrection for population stratification with a reduced power in QTL detection. Therefore, we based our results on the K model and not on the P10 or P10K models.

Genome-wide scan for main effects

The optimum estimate for the identity-by-state probability for the calculation of the K matrix was 0.475. A genome-wide scan for main effects identified QTL for all seven traits, distributed throughout the whole genome (Table 2; Fig. 4). The number of detected QTL ranged from 4 for α-amino nitrogen to 21 for sugar content. The proportion of genotypic variance explained by the detected QTL ranged from 27.9% for sodium to 82.5% for beet yield. The effect sizes of the QTL were higher for the quality-related traits (K, Na, N) compared to yield traits (WSY, SY, SC and BY) (Fig. 6).

Detection of epistasis

As for the main effects genome scan, the plot of observed versus expected P values indicated that the model
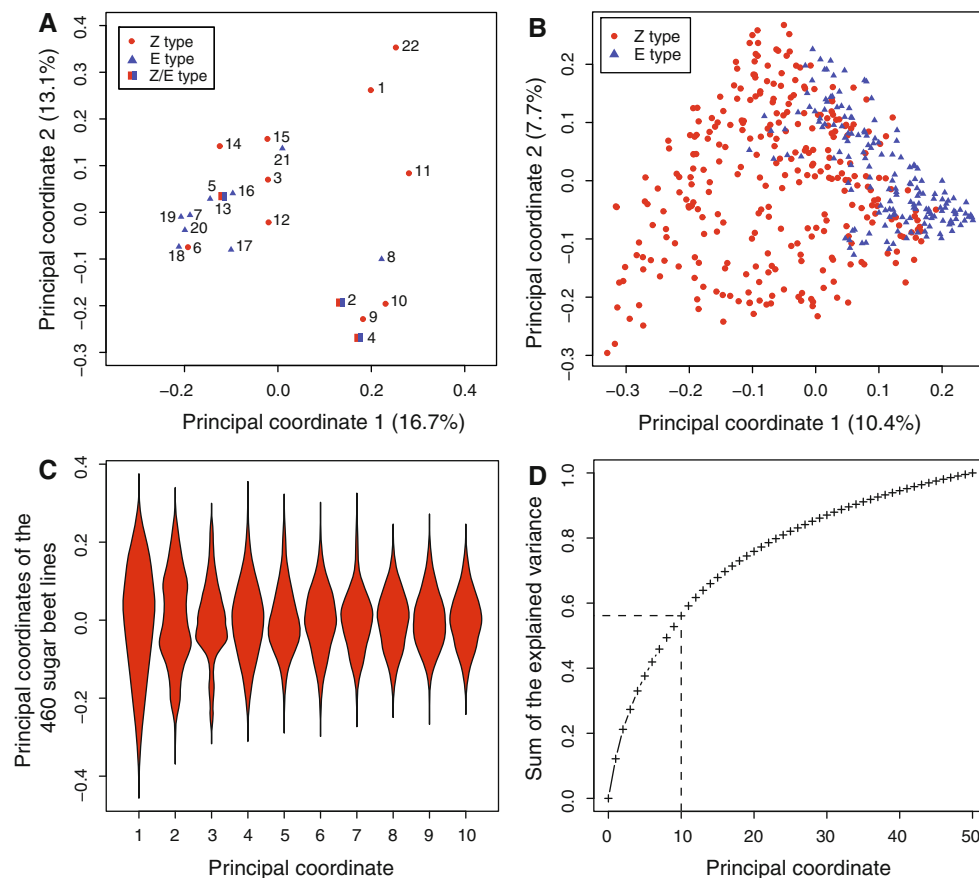
Fig. 2 a Principal coordinate analysis of the 22 parents and **b** of the 460 genotypes of the population, based on modified Rogers' distance estimates. Sugar (Z) and yield (E) types are shown. *Percentages in parentheses* refer to the proportion of variance explained by the principal coordinate. **c** Violin plot showing the density distribution of the first 10 principal coordinates for the genotypes from the population. No population structure is apparent. **d** Sum of the explained variance of the first 10 principal coordinates (56%) used to correct for population structure

including the kinship matrix should also be appropriate for epistasis scans (Fig. S2). The two-dimensional genome scan revealed significant two-way interactions for all traits. The number of identified QTL ranged from one for white sugar yield to 20 for sugar content. Contribution of epistatic effects to the genotypic variance was minimum for WSY (1.1%) and maximum for Na (37.4%) (Table 2). Most of the two-way interactions are composed of markers, which do not have a significant individual effect (Fig. 5, Table S1). For most of the detected interactions no significant LD was found suggesting that there is not a high selection pressure on specific allele combinations.

We searched for higher order epistasis by extending the detected two-way interactions towards three-way interactions. Significant three-way epistatic QTL were detected for the four traits, SY, SC, BY and K (Table 2). The number of detected QTL ranged from one to three, explaining 0.3–1.9% of the genotypic variance. The main and the epistatic QTL together explained between 40.5

(N) to 87.2% (BY) of the genotypic variance (Table 2). The effect sizes of both the two-way and the three-way epistatic QTL were found to be smaller compared to the main effects (Fig. 6).

Genetic interaction networks

To illustrate the extent and the range of epistasis, we visualized the detected two- and three-way interactions (Figure S3). The two-way interactions define a network of genetic interactions, which encompasses the whole genome. Whereas some of the epistatic QTL are intrachromosomal interactions, the majority of the two-way epistatic QTL reveal interchromosomal interactions. The extension of the two-way interactions towards higher order epistasis revealed that the detected interactions extend to loci, which have not been identified as single or as two-way interaction. Also the three-way interactions define a network covering the entire genome.
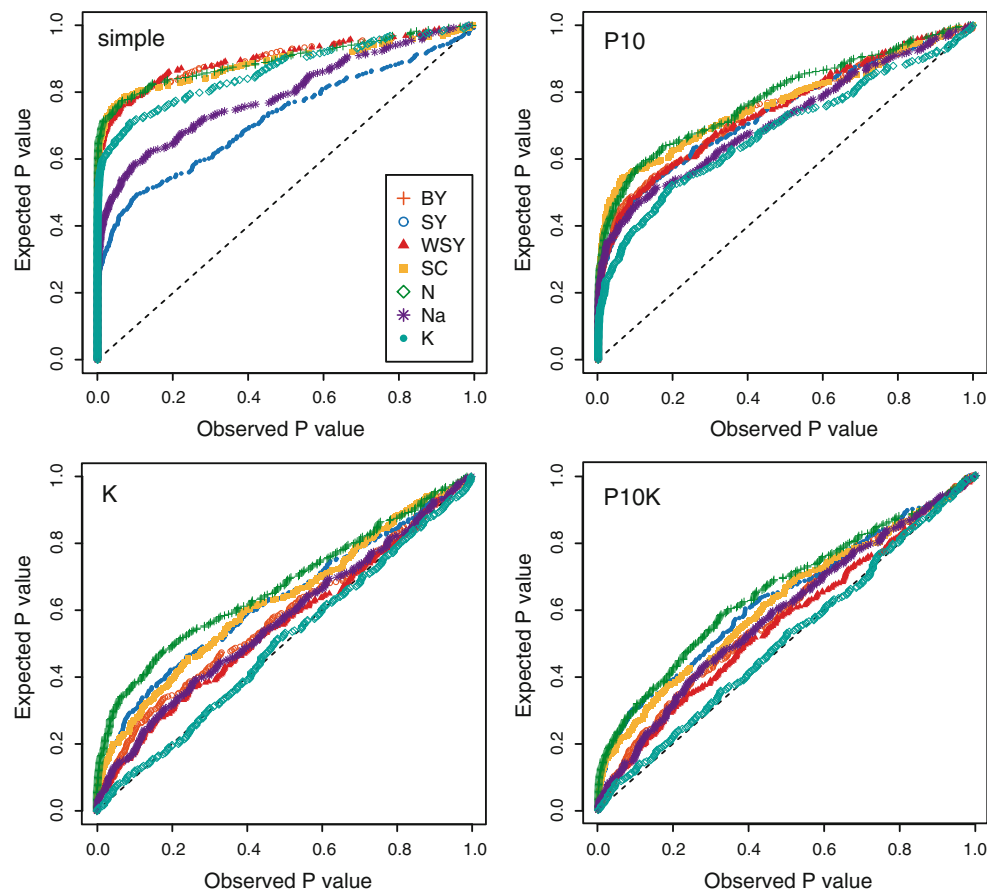
Fig. 3 Plot of observed versus expected *P* values for the four different association mapping models. No correction (simple), first 10 principal coordinates (P10), kinship matrix (*K*), and first 10 principal coordinates and kinship matrix (P10K). In each plot all seven traits are shown which have different correlations between the phenotype and the population structure (Table 1). *WSY* white sugar yield, *SY* sugar yield, *SC* sugar content, *BY* beet yield, *K* potassium, *Na* sodium, *N* α-amino nitrogen

Table 2 Number of significant main, two- and three-way epistatic QTL and the proportion of the genotypic variance (p_G) explained by these QTL
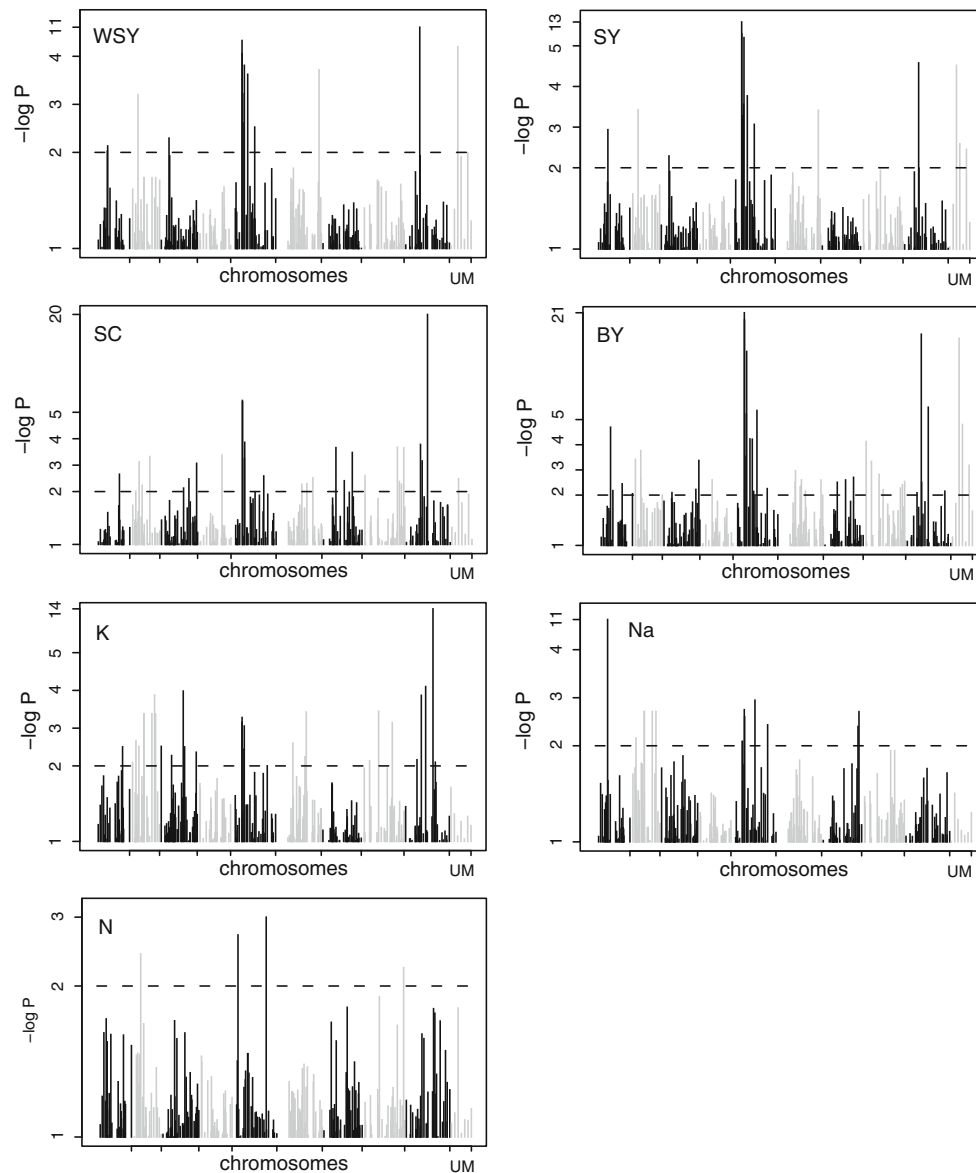
The model including the kinship matrix (*K*) was used and results are shown for *WSY* white sugar yield, *SY* sugar yield, *SC* sugar content, *BY* beet yield, *K* potassium, *Na* sodium, *N* α-amino nitrogen

	WSY	SY	SC	BY	K	Na	N
Main effect QTL							
Number	8	8	21	19	18	8	4
p_G (%)	70.7	69.2	78.4	82.5	51.6	27.9	34.0
Two-way epistatic QTL							
Number	1	4	20	8	12	10	2
p_G (%)	1.1	4.0	6.7	4.4	9.4	37.4	6.5
Three-way epistatic QTL							
Number	0	2	3	3	1	0	0
p_G (%)	0	0.9	1.9	0.3	0.5	0	0
Combined							
p_G (%)	71.8	74.1	87.0	87.2	61.5	65.3	40.5

Discussion

Epistasis is a phenomenon that has been discovered more than 100 years ago (Carlborg and Haley 2004; Phillips 2008). With the rapidly developing omics technologies, it

is likely that the concept of epistasis will become even more central to biology than it used to be (Brem and Kruglyak 2005; Schadt et al. 2003). Understanding the essential role of gene interactions and the genetic architecture of complex traits can build a foundation to examine

Fig. 4 Results of the single marker analysis with the model including the *K* matrix are shown for the seven traits. The significance threshold ($P < 0.01$)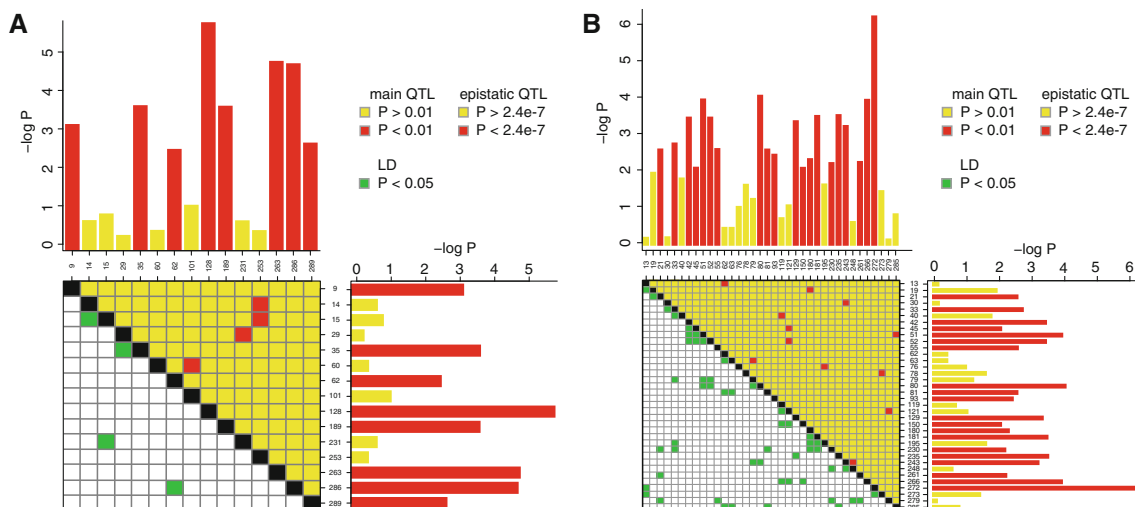 is indicated as *dashed line*. *Black and white* is used to separate chromosomes. *WSY* white sugar yield, *SY* sugar yield, *SC* sugar content, *BY* beet yield, *K* potassium, *Na* sodium, *N* α-amino nitrogen, *UM* unmapped markers

the interplay of components on various levels within biological systems. We therefore wanted to assess the potential of association approaches to identify epistasis.

Resolution of association mapping in breeding populations

The power to detect QTL, especially epistatic QTL, for quantitative traits in association mapping, greatly depends on the sample size and the heritability of the traits. Our study was based on a large experimental data set and except for the two traits, Na and N, heritability was higher than 0.8 (Table 1), facilitating the detection of epistatic interactions (Stich et al. 2008c). Association mapping is expected to enable a higher mapping resolution than traditional linkage mapping methods, as it employs LD based on historical recombinations. We found that in our data set LD decays after approximately 10 cM (Fig. 1). This is in accordance with results from breeding populations of other crops such as maize (Stich et al. 2005), barley (Kraakman et al. 2004), and wheat (Chao et al. 2007). The average r^2 between adjacent loci was 0.52. Consequently, a higher power to detect QTL of interest could be achieved by further increasing the marker density.

Fig. 5 *P* values from the single marker analysis (barplot, *top* and *right*) and from the two-way epistatic QTL (*center square*, above diagonal) are shown for **a** sugar yield and **b** potassium. Significant LD is shown below the diagonal
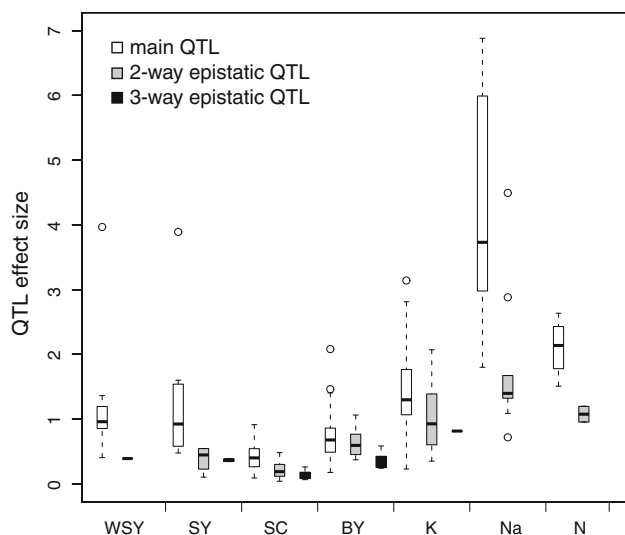
Fig. 6 Box-Whisker plots of the absolute effect sizes of the detected main and epistatic QTL for *WSY* white sugar yield, *SY* sugar yield, *SC* sugar content, *BY* beet yield, *K* potassium, *Na* sodium, *N* α-amino nitrogen

Choice of an appropriate model for association mapping

Correcting for confounding by population structure is essential for association mapping in plant populations (Zhao et al. 2007; Yu et al. 2006). An appropriate statistical model should provide an excellent compromise between correcting for population stratification to decrease the probability to detect false-positive marker–trait associations but still retaining enough information within the markers for QTL detection. The analysis of the relatedness of the 460 lines revealed that correction for population structure is not required whereas familial relatedness

should be considered (Figs. 2, 3). Therefore, trait–marker associations were investigated in detail with the *K* model, which incorporates estimates of kinship coefficients based on marker data.

Detection of main effect QTL

The genome-wide scan revealed QTL for all seven traits (Table 2). The effect sizes of the QTL were considerable and especially high for sodium (Fig. 6). Especially for traits with a high heritability, such QTL, which explain a high proportion of the genotypic variance are good candidates for a map-based cloning approach to identify the genes and the causative polymorphisms underlying these QTL.

The proportion of genotypic variance explained by the detected QTL was high, except for Na and N (Table 2). Cross-validation experiments in linkage mapping studies suggest that these values are generally over-estimated (Utz et al. 2000). Due to the inherent population structure in association mapping populations a cross-validation strategy as in linkage mapping is not feasible. We therefore suggest that validation of the observed QTL should be done in an independent sample set to obtain unbiased estimates and realistically assess the prospects of marker-assisted selection.

The role of epistasis

We performed a two-dimensional genome scan and detected two-way epistatic interactions for all seven traits (Table 2). Like previous studies in various organisms (Montooth et al. 2003; Eshed and Zamir 1996; Xu and Jia 2007), we found that the interacting loci mostly have no

significant individual main effect (Fig. 5). Consequently, any epistasis scan limited to testing only those QTL with significant main effects will fail to detect the majority of epistatic interactions.

The proportion of genotypic variance explained by the epistatic interactions was small compared to that explained by the main QTL (Table 2). The exception to this was sodium, where more of the genotypic variance was explained by epistasis than by the main effects. As for the main effect QTL, the genotypic variance explained by the epistatic interactions is likely to be over-estimated. We also identified three-way epistatic interactions for four of the traits by extending the detected two-way interactions (Table 2). We speculate that, like for the two-way epistasis, a full three-dimensional genome scan would identify three-way epistatic effects, which were missed by the applied computationally less demanding approach. Parallel computing represents a promising avenue to investigate higher-order epistasis within a reasonable timeframe.

Our results show that epistasis should not be neglected, because it contributes substantially to the genotypic variance of agronomic important traits (Table 2). If epistasis is not considered, the estimated effects of QTL may be severely biased resulting in a lower gain in MAS. In contrast, incorporating important epistatic effects in MAS through the targeted combination of specific alleles at epistatic loci will lead to an improved development of superior varieties.

Genetic architecture of agronomic important traits

We detected main effect QTL and two- and three-way epistatic interactions which define the genetic architecture and allow a first glimpse at the genetic networks shaping these quantitative traits. The large number of components involved in the expression of complex traits is reflected by a considerable amount of interactions in these networks (Figure S3). Screens for epistasis in association mapping approaches should therefore be performed as a first step to identify genetic networks. The epistatic interactions must then be verified by other experimental approaches.

In conclusion, our results show that association mapping in applied plant breeding populations provides sufficient power for the detection of both, main and epistatic QTL. Whereas linkage mapping offers a high power to detect QTL due to balanced allele frequencies, the specific design of association approaches with a low probability of fixed alleles at each locus and many combinations of these alleles in different genotypes is advantageous for epistasis scans. The combination of both as in joint linkage association mapping thus holds great potential for the detection of epistatic QTL. We show, that for physiological traits, epistasis can be an important contributor to genetic variation that should be considered in future analyses. This information on the genetic architecture of agronomic important traits can then be implemented in breeding programs to improve knowledge-based breeding of crops, urgently required in times of increasing human population size and environmental change.

References

Bernardo R (1993) Estimation of coefficient of coancestry using molecular markers in maize. Theor Appl Genet 85:1055–1062

Brem RB, Kruglyak L (2005) The landscape of genetic complexity across 5.700 gene expression traits in yeast. Proc Natl Acad Sci USA 102:1572–1577

Carlborg Ö, Haley CS (2004) Epistasis: too often neglected in complex trait studies? Nat Rev Genet 5:618–625

Carlborg Ö, Jacobsson L, Ahgren P, Siegel P, Andersson L (2006) Epistasis and the release of genetic variation during long-term selection. Nat Genet 38:418–420

Chao S, Zhang W, Dubcovsky J, Sorrells ME (2007) Evaluation of genetic diversity and genome-wide linkage disequilibrium among U.S. wheat (*Triticum aestivum* L.) germplasm representing different market classes. Crop Sci 47:1018–1030

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED et al (2010) The genetic landscape of a cell. Science 327:425–431

Draycott AP (2006) Sugar beet, 1st edn. Blackwell, Oxford

Eshed Y, Zamir D (1996) Less-than-additive epistatic interactions of quantitative trait loci in tomato. Genetics 143:1807–1817

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2006) ASReml User Guide Release 2.0 VSN International Ltd, Hermel Hempstead, UK

He X, Qian W, Wang Z, Li Y, Zhang J (2010) Prevalent positive epistasis in *Escherichia coli* and *Saccharomyces cerevisiae* metabolic networks. Nat Genet 42:272–276

Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. Theor Appl Genet 38:226–231

Holm S (1979) A simple sequentially rejective multiple test procedure. Scand J Stat 6:65–70

Kraakman ATW, Niks RE, Van den Berg PMMM, Stam P, Van Eeuwijk FA (2004) Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168:435–446

Li Z, Pinson SR, Park WD, Paterson AH, Stansel JW (1997) Epistasis for three grain yield components in rice (*Oryza sativa* L.). Genetics 145:453–465

Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Assoc., Sunderland

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Maurer HP, Melchinger AE, Frisch M (2008) Population genetic simulation and data analysis with Plabsoft. Euphytica 161:133–139

Melchinger AE, Utz HF, Schön CC (1998) Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and larger bias in estimates of QTL effects. Genetics 149:383–403

Montooth KL, Marden JH, Clark AG (2003) Mapping determinants of variation in energy metabolism, respiration and flight in Drosophila. Genetics 165:623–635

Phillips PC (2008) Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. Nat Rev Genet 9:855–867

Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA et al (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat Genet 38:904–909

Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N et al (2003) Genetics of gene expression surveyed in maize, mouse and man. Nature 422:297–302

Schneider K, Schäfer-Pregl R, Borchardt C, Salamini F (2002) Mapping QTLs for sucrose content, yield and quality in a sugar beet population fingerprinted by EST-related markers. Theor Appl Genet 104:1107–1113

St Onge RP, Mani R, Oh J, Proctor M, Fung E et al (2007) Systematic pathway analysis using high-resolution fitness profiling of combinatorial gene deletions. Nat Genet 39:199–206

Stich B, Melchinger AE, Frisch M, Maurer HP, Heckenberger M et al (2005) Linkage disequilibrium in European elite maize germplasm investigated with SSRs. Theor Appl Genet 111:723–730

Stich B, Melchinger AE, Heckenberger M, Möhring J, Schechert A et al (2008a) Association mapping in multiple segregating populations of sugar beet (Beta vulgaris L.). Theor Appl Genet 117:1167–1179

Stich B, Möhring J, Piepho HP, Heckenberger M, Buckler ES et al (2008b) Comparison of mixed-model approaches for association mapping. Genetics 178:1745–1754

Stich B, Yu J, Melchinger AE, Piepho HP, Utz HF et al (2008c) Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy. Genetics 176:563–570

Stram DO, Lee JW (1994) Variance components testing in longitudinal mixed effects model. Biometrics 50:1171–1177

Tong AH, Lesage G, Bader GD, Ding H, Xu H et al (2004) Global mapping of the yeast genetic interaction network. Science 303:808–813

Utz HF, Melchinger AE, Schön CC (2000) Bias and sampling error of the estimated proportion of genotypic variance explained by quantitative trait loci determined from experimental data in maize using cross validation and validation with independent samples. Genetics 154:1839–1849

Weber WE, Borchardt DC, Koch G (2000) Marker analysis for quantitative traits in sugar beet. Plant Breed 119:97–106

Weir BS (1996) Genetic data analysis II, 2nd edn. Sinauer Associates, Sunderland

Wright S (1978) Evolution and genetics of populations, variability within and among natural populations, vol 4. The University of Chicago Press, Chicago, p 91

Xu S, Jia Z (2007) Genomewide analysis of epistatic effects for quantitative traits in barley. Genetics 175:1955–1963

Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208

Zhao K, Aranzana MJ, Kim S, Lister C, Shindo C et al (2007) An Arabidopsis example of association mapping in structured samples. PLoS Genet 3:e4