Mapping of yield, yield stability, yield adaptability and other traits in barley using linkage disequilibrium mapping and linkage analysis.

Arnold T.W. Kraakman

Promotor:

Prof. dr. R. G. F. Visser
Hoogleraar in de Plantenveredeling

Co-promotor:

Dr. F. A. van Eeuwijk
Universitair hoofddocent, Laboratorium voor Plantenveredeling

Promotiecommissie:

Prof. dr. ir. P. C. Struik (Wageningen Universiteit)
Prof. dr. R. F. Hoekstra (Wageningen Universiteit)
Prof. dr. M. A. M. Groenen (Wageningen Universiteit)
Prof. dr. I. Romagosa (University of Lleida, Spain)

Mapping of yield, yield stability, yield adaptability and other traits in barley using linkage disequilibrium mapping and linkage analysis.

Arnold T.W. Kraakman

Proefschrift
ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. Dr. Ir. L. Speelman,
in het openbaar te verdedigen
op dinsdag 7 juni 2005
des namiddags te vier uur in de Aula.

CIP-DATA Koninklijke Bibliotheek, Den Haag

Kraakman, A.T.W.

Mapping of yield, yield stability, yield adaptability and other traits in barley using linkage disequilibrium mapping and linkage analysis.

Thesis Wageningen University, The Netherlands, 2005
With references – with summaries in English and Dutch

ISBN 90-8504-205-4

Keywords: barley, linkage disequilibrium mapping, linkage analysis, molecular markers, AFLP, SSR, yield, yield stability, yield adaptability, genotype × environment interaction, leaf rust resistance, barley yellow dwarf virus, plant height, heading date, rachilla hair length, lodicules size.

Voor Wim en Afra,

twee bijzondere mensen

Voor Rikje, Rianne, Wouter en Mirjam,

mijn liefde, mijn leven

Bibliographic abstract

Identification and mappping of Quantitative Trait Loci (QTLs) in plants is mostly done through linkage analysis. A segregating mapping population is created from a bi-parental cross and linkages between trait values and mapped markers reveal the positions of QTLs. In this study we explored linkage disequilibrium (LD) mapping of traits in a set of modern barley cultivars. LD between molecular markers was found up to a distance of 10 centimorgan, which is large compared to other species. The large distance might be induced by LD increasing factors such as inbreeding and the fact that the population is most likely based on a rather small set of founding genotypes. Associations between markers and traits were found for yield, yield stability, leaf rust resistance (LR), barley yellow dwarf virus resistance (BYD), plant height, and days to heading. Trait-associated markers from LD analysis were located in regions where already QTLs for the trait considered had been reported from studies based on bi-parental crosses. In addition, new QTLs were found for yield, yield stability, LR and BYD. We expect that LD mapping will become a valuable extension to conventional QTL analysis in plant breeding.

Special attention was given to traits describing genotype × environment interactions. Statistical models were used to define measures for yield adaptability and yield stability without including environmental factors directly in the models. Adaptability was defined as the responsiveness of the genotype to the environment, and stability was defined as the unexplained deviation from the statistical model. LD mapping in barley cultivars resulted in marker-trait associations for yield stability. In addition, linkage analyses in four doubled-haploid populations resulted in detection of many QTLs for adaptability, but only a single QTL for stability. We concluded that adaptability measures were genetically better defined than stability measures and that selection for adaptability should be possible.

Contents

Chapter 1

General introduction and outline

Chapter 1

General Introduction and outline

Cultivated barley (*Hordeum vulgare* L.) is the world's fourth most important cereal crop, after wheat, maize, and rice. It is grown over a broader environmental range than any other cereal. The popularity of barley is due to its broad ecological adaptation, utility as a feed and food grain, and superiority of barley malt for use in brewing (Poehlman, 1985).

Barley provides an excellent system for genome mapping and genetic studies, due to (1) its diploid nature, (2) low chromosome number (2n=14), (3) relatively large chromosomes (6-8 μm), (4) high degree of self fertility, and (5) ease of hybridization. Its only drawback is the size of the genome, which is relatively large with 5.3×10^9 bp for a haploid (Bennett and Smith, 1976).

Genetic variation is large in the *Hordeum* genus, due to mutants and a large number (32) of *Hordeum* species. The species in the genus range from diploid to polyploid, and they are perennial as well as annual. Cultivated barley, *Hordeum vulgare* spp.*vulgare*, can be enriched with genes from its ancestor *Hordeum spontaneum*, and also less related species can be used as gene donor, *viz. Hordeum bulbosum*.

In genebanks approximately 250.000 *Hordeum* accessions are stored. With the growth of this germplasm collection, the need for a more efficient use has increased. The Core Collection concept (Brown, 1989), an initiative that also includes barley (Hintum 1992), improved the accessibility of a large source of genetic variation.

BARLEY BREEDING

Barley is grown for fodder, human consumption, and the brewing of beer and whisky. The main breeding objectives are high yield, and resistance to biotic and abiotic stresses. Furthermore, malting cultivars need to have high malting quality, which includes plump

kernels, rapid and uniform germination, and optimal values for protein content and enzymatic activity.

As barley is a relatively strict inbreeder, barley lines are almost completely homozygous. F1-hybrids are produced by emasculation of the female parent and adding the pollen of the male parent one to three days later to the bagged female spike. The F1 can be developed into inbred lines by self-fertilization, but also by the production of doubled haploids (DH). The most frequently applied techniques to obtain DHs are the *bulbosum* method (Kasha and Kao, 1970), and the anther culture (Friedt and Foroughti-Wehr, 1981). DHs are a fast road to homozygosity, but selection is only possible after the DHs have been created. Selfing is time consuming, as at least 7 or 8 cycles of selfing are necessary to reach homozygosity, but in the later stages of this process many inadequate lines can be discarded already.

Qualitative and quantitative traits

A distinction can be made between qualitative and quantitative traits. Qualitative traits are mostly determined by only one or a few loci, and they are relatively easy to select for as the phenotype expresses the genotype. Quantitative traits are mostly determined by many genes (quantitative trait loci or QTLs), and more complex to select for. Some examples of qualitative traits in barley are the number of rows (*v* gene on chromosome 2), and black lemma and pericarp (*B* gene on chromosome 5). Examples of quantitative traits are yield, many disease resistances, plant height, and days to heading.

Genotype × Environment interaction

Selection for many traits is not only being complicated by their quantitative nature, but also by the interaction between genotype and environment (GE). As a result of this interaction, the ranking order of varieties may change as the growing conditions (environments) change. Yield is a complex, polygenic trait that is strongly influenced by environmental factors. The changes of yield in relation to environmental changes are studied in the context of the concepts yield adaptability and yield stability. Adaptability can be described as the reaction of the genotype to environmental factors, often defined in terms of linear or quadratic functions (Lin *et al*., 1986). Stability is often described as the variability of the genotype's performance around its mean performance. A well known measure for adaptability is the slope of the regression of yield for an individual cultivar on the mean yield (over all cultivars) across environments (Finlay and

Wilkinson, 1963). A well known measure for stability is the mean square of deviations from the Finlay-Wilkinson regression line (Eberhart and Russell, 1966).

The polygenic basis of adaptability and stability, combined with the probable interaction between loci, makes selection for these traits very complicated. Large scale testing across a range of environments is required, which is labor and time intensive. Identification of the loci involved would help a great deal, especially if markers are linked to those loci and indirect selection on the basis of markers is possible.

Several researchers have conducted multi-environment trials for various traits in different plant species, *e.g.* drought resistance in cotton (Saranga *et al*., 2001), photoperiod plasticity in *Arabidopsis* (Ungerer *et al*., 2003), growth and yield in rice (Hittalmani *et al*., 2003), and yield in barley (Teulat *et al*., 2001; Romagosa *et al*., 1996; Voltas *et al*., 2001; Malosetti *et al*., 2004). They all succeeded in identifying loci that interacted with the environment, so loci underlying GE. Some loci for GE co-localized with loci for the trait mean expression, while others appeared at positions where no QTLs for the mean expression were found.

GENETIC MAPPING OF TRAITS

Genetic linkage maps

The breeding process can be enhanced by using the linkage between markers and traits, which enables indirect selection on markers avoiding the phenotypic assessment of traits. An important step towards the establishment of such linkages is the development of genetic maps. One of the first well developed classical genetic maps for barley included isozymes and morphological markers (Sogaard and von-Wettstein-Knowles, 1987). Later on, molecular markers were added, beginning with RFLP and PCR markers (Shin *et al*., 1990), and these maps became more dense (Graner *et al*., 1991; Heun *et al*., 1991; Kleinhofs *et al*., 1993) enabling the mapping of many important agronomic qualitative and quantitative traits. New molecular markers were developed, improving the barley genetic map with AFLP markers (Waugh *et al*., 1997; Qi *et al*., 1998; Yin *et al*., 1999), and with microsatellite markers (Ramsay *et al*., 2000; Pillen *et al*., 2000; Holton *et al*., 2002). The main advantage of AFLP markers is that in relatively little time many genotypes can be evaluated and identified with many markers. The main advantage of SSR markers is that they are multi-allelic and therefore

very powerful in breeding studies where a broad germplasm is evaluated (Russell *et al*., 1997; Struss and Plieske, 1998).

Mapping traits

Genetic mapping of traits comes down to finding linkage between mapped markers and phenotypic trait observations. Finding such linkage can be done in several ways. Two commonly used approaches are: (1) linkage analysis using a bi-parental mapping population segregating for the trait(s) of interest, or (2) linkage disequilibrium mapping using a well chosen (natural) population of lines, accessions, or cultivars.

1. Linkage analysis

The method of linkage analysis is well developed for bi-parental crosses between inbred lines. Estimation of recombination rates between loci allows the construction of a genetic linkage map. Besides, associations between a traits and marker alleles identify the genomic regions in which the loci controlling the trait are located. In this way, QTL locations and effects are determined.

The precision of QTL mapping depends on both the size of the mapping population and the genetic variation that is attributable to a particular QTL. More precise mapping (fine mapping) of QTLs requires different approaches, such as the construction of backcross inbred lines (BILs) or recombinant inbred lines (RILs). BILs are homogeneous lines containing small introgressed donor genome fragments. RILs are inbred lines that have undergone several rounds of random mating, increasing the potential number of recombination events.

Disadvantages of linkage analysis are that the results can not be a priori extrapolated to other crosses (although several studies have shown that certain QTL regions may be common to several crosses). Furthermore, only two alleles at a given locus can be studied simultaneously, and special mapping populations have to be developed and tested for the trait of interest.

Linkage analysis depends on the linkage disequilibrium that is introduced in the mapping population. Random mating produces a population in which all loci are in linkage equilibrium, so that no association between markers and trait loci can be established. In a mapping population, however, the degree of linkage disequilibrium depends on the recombination events that have taken place during a single or a few meioses following the cross between the parents.

2. Linkage disequilibrium mapping

The method of linkage disequilibrium (LD) mapping (or association mapping) is relatively new. Until recently, it was only used in human and animal genetics. LD, also known as gametic phase disequilibrium, gametic disequilibrium, and allelic association, is the nonrandom association of alleles at different loci. It is the correlation between polymorphisms that is caused by their shared history of mutation and recombination.

LD mapping is done with a natural population in which association between traits and markers exists due to linkage disequilibrium. The degree of LD in a germplasm depends on the recombination events that have taken place in history (Nordborg and Tavaré, 2002). It is a result of the interaction between many factors, *e.g.* the mating system, recombination rate, selection, and population subdivision (Ardlie *et al*., 2002; Flint-Garcia *et al*., 2003; Rafalski and Morgante, 2004; Zondervan and Cardon, 2004; see Table 1). Not all LD occurring in a germplasm is due to linkage between loci. LD between unlinked loci can occur, attributable to population structure, admixture, outcrossing events, selection, etc. Therefore, observed associations between markers and traits should be interpreted with care.

Table 1: Factors affecting linkage disequilibrium (LD) in a population.

Factor	Effect on LD
Genomic rearrangements	Rearrangements suppress local recombination leading to LD increase in the vicinity
Mutation rate	High mutation rate decreases overall LD, but LD around a newly created mutated allele remains high until dissipated by recombination
Recombination rate	High recombination lowers LD
Balancing selection	Increases LD
Genetic isolation between lineages	Increases LD
Population admixture	Increases LD
Population size	Small populations have more LD
Population subdivision	Increases LD
Selection	Locally increases LD
Selfing / outcrossing	High / low LD

Although both linkage analysis with a segregating population and LD mapping with a germplasm have the same underlying principle of using linkage disequilibrium for identifying

trait associated chromosome regions, there are a number of differences (see Table 2). Which approach is the most useful depends both on the genetic material and the purpose of the mapping analysis.

Table 2: Comparison of linkage analysis and LD mapping

Factor	Linkage analysis	LD mapping
Analysis methodology	Well developed	Under development
Fine mapping	Possible with many markers and a large population	Depending on history
Genetic variety	Only two parents	Broad genetic background
Number of alleles per locus	Two	Many
Phenotypic testing	Extra work	Often already done
Population	Segregating population from a biparental cross (F1, F2, BILs, RILs, DHs)	Germplasm
Population history	Known	Often unknown

Measuring LD

A variety of statistics have been used to measure LD (see reviews of Jorde (2000) and Flint-Garcia *et al.,* 2003). The two most common statistics are r^2 and *D'*. Consider a pair of loci with alleles *A* and *a* at locus one, and *B* and *b* at locus two, with allele frequencies π_A, π_a, π_B, and π_b, respectively. The resulting haplotype frequencies are π_{AB}, π_{Ab}, π_{aB}, and π_{ab}. The basic component of all LD statistics is the difference between the observed and the expected haplotype frequencies, $D_{ab} = (\pi_{AB} - \pi_A \pi_B)$. The distinction between the two statistics lies in the scaling of this difference. The first of the two measures, r^2, is calculated as:

$$r^2 = \frac{(D_{ab})^2}{\pi_A \pi_a \pi_B \pi_b}$$

Actually, r^2 can also be considered as the squared correlation coefficient between the two loci. The second measure, *D'*, is calculated as:

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_b , \pi_a \pi_B)} \quad \text{for } D_{ab} < 0$$

$$|D'| = \frac{(D_{ab})^2}{\min(\pi_A \pi_B , \pi_a \pi_b)} \quad \text{for } D_{ab} > 0$$

D' and r^2 reflect different aspects of LD and perform differently under various conditions. Whereas r^2 summarizes both recombinational and mutational history, *D'* measures only recombinational history and is therefore the more accurate statistic for estimating recombination differences. However, *D'* is strongly affected by small sample sizes, resulting in highly erratic behavior when comparing loci with low allele frequencies. For the purpose of examining the resolution of association studies, the r^2 statistic might be more appropriate, as it is indicative of how markers might correlate with the QTL of interest. In Figure 1 examples of LD are given with the corresponding r^2 and *D'*.

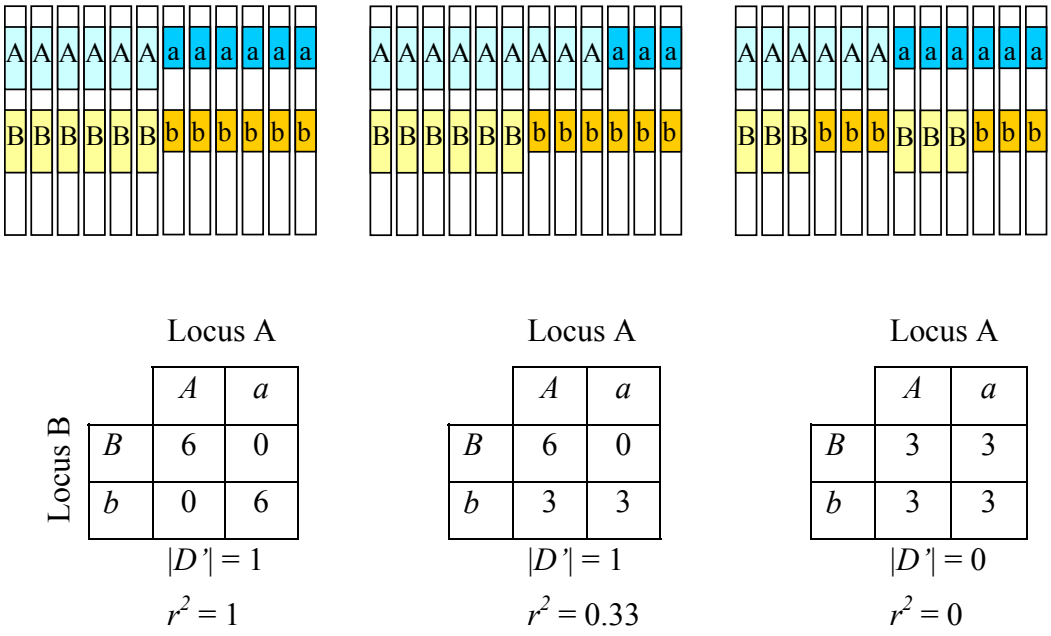


Figure 1: Examples of linkage disequilibrium and the corresponding r^2 and *D'*.

LD in humans and animals

LD in humans has been studied extensively (see review of Pritchard and Przeworski, 2001). LD estimates vary largely, from 5 kb (Reich *et al*., 2001) to 500 kb (Taillon-Miller *et al.,* 2000) (see Table 3). This heterogeneity is caused by differences in loci, sample populations, and chromosome type.

In cattle (*Bos Taurus*) LD extended to 10 cM (Farnir *et al*., 2000). This large LD distance is likely induced by the fact that the Dutch black and white dairy population originated for 40% from the top ten ranked sires, resulting in a narrow germplasm.

LD in plants

In maize (*Zea mays* spp.*mays*), several studies with a wide range of populations showed that LD varies substantially with the population chosen. In a diverse group of maize germplasm, Tenaillon *et al*. (2001) found LD up to 200 bp, while in inbred lines and elite maize LD diminished after 1500 bp and 100.000 bp, respectively (Remmington *et al*., 2001, and Rafalski, 2002, respectively). In the latter germplasm it is likely that recombinational inactive repetitive regions were included in the genome area under study, explaining (partly) the extended LD.

Table 3: LD estimates in animal, human and plant species**.**

Species	Population	LD range	Reference
Cattle	Black and white dairy	10 cM	Farnir *et al*., 2000
Human	Nigerian	5 kb	Reich *et al*., 2001
Human	Finnish, Scandinavian and CEPH samples	500 kb	Taillon-Miller *et al*., 2000
Arabidopsis	Global sample of 76 accessions	250 kb (1 cM)	Nordborg *et al*., 2002
Arabidopsis	Local population	> 50 cM	Nordborg *et al*., 2002
Maize	Diverse group	< 1kb	Tenaillon *et al*., 2001
Maize	Elite lines	> 100 kb	Rafalski, 2002
Maize	Inbred lines	1.5 kb	Remmington *et al*., 2001
Soybean	25 genotypes	50 kb	Zhu *et al*., 2003
Sugar beet	9 inbreds	3 cM	Kraft *et al*., 2000
Sugarcane	59 modern cultivars	10 cM	Jannoo *et al*., 1999

In Arabidopsis (*Arabidopsis thaliana*) LD occurred over much larger distances, because Arabidopsis is a highly selfing crop. Nordborg *et al*. (2002) observed LD over 250 kb (equivalent to 1 cM), and in isolated populations LD did not even decay until 50 cM.

In sugarcane (*Saccharum* spp.) cultivars, LD was found up to 10 cM (Jannoo *et al*., 1999). This long stretch could be attributed to several characteristics of this germplasm. First, a majority of modern cultivars was derived from an interspecific cross between *S.officinarum* ×

S.spontaneum, followed by backcrossing with *S.officinarum* creating a bottleneck in breeding history. Second, sugarcane is vegetatively propagated, as a result of which most cultivars have undergone less then ten meioses since the original cross. Both factors increase the LD present in this germplasm.
Soybean showed little decay in LD over 50 kb (Zhu *et al*., 2003). LD in sugar beet (*Beta vulgaris*) diminished after 3 cM (Kraft *et al*., 2000).

LD mapping studies in plant crops

Only few examples exist of LD mapping of genes in plant crops. Earlier attempts for establishing associations between traits and markers across germplasm collections concerned oat, rice, maize, sea beet, barley and wild barley. In oat, Beer *et al.* (1997) studied 64 landraces and cultivars, and they found associations between thirteen quantitative traits and markers. In rice, Virk *et al*. (1996) used multiple linear regressions to predict six traits with marker alleles. In maize, Thornsberry *et al*. (2001) found associations between Dwarf8 polymorphisms and flowering time. In sea beet, Hansen *et al.* (2001) found LD between AFLP markers and the bolting gene in four populations. In barley, Igartua *et al*. (1999) concluded that marker-trait associations for heading date, found in mapping populations, were, to some extent, maintained in 32 cultivars. In wild barley, Ivandic *et al.* (2003) found association between markers and the traits water-stress tolerance and powdery mildew in 52 lines.

The methodology of LD mapping in plant crops is still under development. The focus is on issues such as detection of and dealing with population structure (Pritchard, 2000a+b), the use of pedigree information (Jannink *et al*., 2001), and choosing thresholds when performing multiple tests in a genome wide association study (Weller *et al.*, 1998).

RESEARCH GOALS AND OUTLINE OF THIS THESIS

This thesis aimed at investigating the properties of a novel method of genetic analysis and improving knowledge about the genetics of yield stability and yield adaptability. Phenotypic data from past multi-environment trials were analyzed for association between molecular marker genotypes and yield stability and yield adaptability.

More specifically, the following research questions were addressed:

1. What is the degree of LD present in modern barley cultivars?
2. Can LD mapping be used to map QTLs and major genes for a diverse set of traits?
3. Can we map measures for yield stability and yield adaptability in mapping populations based on bi-parental crosses, and do the results agree with the results of the LD mapping?

In **Chapter 2**, LD mapping in barley is explored. A germplasm of 148 modern spring barley cultivars was fingerprinted with AFLP markers. We determined the extent of LD in our germplasm by calculating the association between markers, and by graphically displaying these associations against the map distances. Yield data from the official Danish barley variety trials were used and yield adaptability and yield stability was established for all genotypes. Finally, the associations between markers and the traits yield, yield adaptability and yield stability were determined. We addressed special issues in LD mapping, such as population structure, and the determination of significance thresholds in multiple testing.

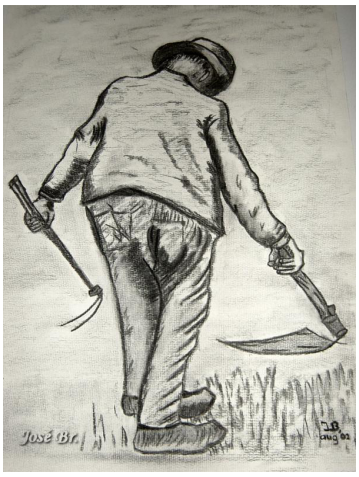
In **Chapter 3**, the search for associations between markers and traits was extended to a variety of traits other than yield. A phenological trait (days to heading), plant architecture trait (plant height), resistance traits (leaf rust, BYD), and morphological traits (Rachilla hair length, Lodicule size) were studied. The set of AFLP markers was enlarged with multi-allelic SSR markers. The presence of known major genes and QTLs in the cultivars was confirmed, and new genes were searched for.

In **Chapter 4**, yield, yield stability, and yield adaptability in a mapping population based on the cross between two cultivars is described. The cultivars were selected on the basis of their difference in stability and adaptability, and their difference for loci which were associated with stability and adaptability according to chapter 2. We tried to map QTLs for yield, yield stability and yield adaptability in this population, and also in three other DH-populations available from the North American Genome Mapping Project. The main objectives of this study were the validation of the association study in chapter 2, and further investigation of yield stability and yield adaptability in order to elucidate the genetics behind those traits.

Finally, in **Chapter 5** the results obtained in the different chapters are described in the light of past, ongoing, and future research in the field of LD mapping and in the field of yield stability and yield adaptability.

Chapter 2

Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars



Chapter 2

Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars

Arnold T. W. Kraakman, Rients E. Niks, Petra M. M. M. Van den Berg, Piet Stam, Fred A. van Eeuwijk

Genetics 168: 435-446 (September 2004)

Keywords: modern spring barley cultivars, AFLP, association mapping, linkage disequilibrium mapping, yield stability and yield adaptability

ABSTRACT

Associations between markers and complex quantitative traits were investigated in a collection of 146 modern two-row spring barley cultivars, representing the current commercial germplasm in Europe. Using 236 AFLP-markers, associations between markers were found for markers as far apart as 10 cM. Subsequently, for the 146 cultivars the complex traits mean yield, adaptability (Finlay-Wilkinson slope), and stability (deviations from regression) were estimated from the analysis of variety trial data. Regression of those traits on individual marker data disclosed marker-trait associations for mean yield and yield stability. Support for identified associations was obtained from association profiles, i.e., from plots of *p*-values against chromosome positions. In addition, many of the associated markers were located in regions where earlier QTLs were found for yield and yield components. To study the oligogenic genetic base of the traits in more detail, multiple linear regression of the traits on markers was carried out, using stepwise selection. By this procedure, 18 to 20 markers were selected that accounted for 40 to 58% of the variation. Our results indicate that association mapping approaches can be a viable alternative to classical QTL approaches based on crosses between inbred lines, especially for complex traits with costly measurements.

INTRODUCTION

The genetic dissection of complex traits still presents a challenge. The oligo/polygenic character of complex traits, combined with interactions between loci, makes the task a priori difficult and intricate. In addition, environmental factors will trigger and modify gene actions, and thereby further complicate the analysis. Yield is the classical example of a complex trait. Yield fluctuations in relation to environmental factors are often described in terms of adaptability and stability. The latter can be considered to constitute complex traits on their own. Parameters quantifying adaptability and stability require observations across a range of environments for their estimation. The parameters are typically defined in terms of linear and quadratic functions of the genotype by environment (GE) interaction (LIN *et al*., 1986).

Adaptability has been studied from several perspectives, manifested by special conferences of breeders and geneticists (TIGERSTEDT, 1997) and physiologists (THOMAS and FARRAR, 1997). Geneticists incline to explanations in terms of favorable epistatic combinations of alleles (ALLARD, 1997). Physiologists focus on the stress response and developmental genes involved. FORSTER *et al*. (2000) stated that developmental genes have strong pleiotropic effects on a number of performance traits in barley, but CATTIVELLI *et al.* (2002) concluded that little is known about the regulatory mechanisms controlling stress responses, mainly because all stress responses involve many genes.

The polygenic basis of complex traits has consequences for the application of quantitative trait locus (QTL) mapping methodology, as many markers that are associated with the trait need to be identified. Typically, for QTL mapping, a cross between two inbred lines is made and the co-segregation of alleles of mapped marker loci and phenotypic traits allows the identification of linked markers. For complex traits with GE-interaction, this approach implies large scale testing of special mapping populations across a range of environments. Several researchers have conducted such multi-environment trials for various traits in different plant species, *e.g.* drought resistance in cotton (SARANGA *et al.*, 2001), photoperiod plasticity in Arabidopsis (UNGERER *et al.*, 2003), growth and yield in rice (HITTALMANI *et al.*, 2003), and yield in barley (TEULAT *et al.*, 2001; ROMAGOSA *et al.*, 1996; VOLTAS *et al.*, 2001). They all succeeded in identifying loci that interacted with the environment, so-called stability loci. Some loci for stability co-localized with loci for mean expression of the trait, while others appeared at positions where no QTLs for the mean expression were found. This finding leaves

inconclusive the debate about the genetic base of stability raised in the evolutionary biology literature. Two types of genetic control for stability were described by VIA *et al.* (1995). In the allelic sensitivity model, the constitutive gene is itself regulated in direct response to the environment, whereas in the gene regulation model one or more regulatory loci are under the direct influence of the environment and the constitutive genes are switched on and off by the regulatory gene(s). Co-localization of QTLs exhibiting GE-interaction and QTLs for stability parameters would point in the direction of allelic sensitivity models. QTLs for stability parameters appearing elsewhere than QTLs for the trait itself would indicate a regulatory gene model.

In this article we explore the possibilities of mapping traits in a collection of modern cultivars, instead of in a segregating population derived from a bi-parental cross. We looked at methodology that has become popular in human genetics under names such as association mapping and linkage disequilibrium (LD) mapping. The success of LD-mapping is obvious from the series of disease genes that have been fine-mapped. For a review, see CARDON and BELL, 2001. Therefore, quantitative geneticists working in crop plants have started to adapt the methodology to their situation (e.g. JANNINK and WALSH, 2002; see GAUT and LONG, 2003 for a review of LD in crop plants).

In the plant breeding context, LD-mapping has several advantages over classical linkage analysis using segregating populations. First, broader genetic variation in a more representative genetic background can be included in the analyses. Second, LD-mapping may attain a higher resolution. Third, multi-trial phenotypic data stored in databases can be linked to marker characterizations of the involved cultivars. Especially the latter advantage is important when evaluation of the trait is time and money consuming, as is the case with mean yield, adaptability and stability.

A genome-wide LD-scan requires many markers, the number depending on the level of LD. In sugar beet, LD extended up to 3 cM (KRAFT *et al*., 2000), while in some Arabidopsis populations LD exceeded even 50 cM (NORDBORG *et al*., 2002). In contrast, in maize LD diminished already after 2,000 bp (REMINGTON *et al*., 2001). As no data are known for barley, a first objective of our research was to obtain an estimate of the level of LD in barley. Our germplasm consisted of 146 modern two-row European spring barley cultivars. They were, homozygous, diploid lines, created by inbreeding or by doubling haploids. As the cultivars were grown all over northwest Europe during the last decade, including the United Kingdom,

France, Germany, Sweden, Denmark, and The Netherlands, they were therefore representative for a large part of the European germplasm.

The main objective of this article was the detection of associations between marker alleles and the quantitative traits mean yield, yield adaptability and yield stability in a set of modern spring barley cultivars. Yield adaptability was defined as the slope of the regression of yield for an individual cultivar on the mean yield (over all cultivars) across environments (FINLAY and WILKINSON, 1963). Yield stability was defined as the mean square of deviations from the Finlay-Wilkinson line (EBERHART and RUSSELL, 1966). We used data from the official Danish barley variety trials for the national and recommended lists from 1993 to 2000. Although many QTLs have been found for yield (see: http://barleyworld.org/NABGMP/qtlsum.htm), only few have been reported for yield adaptability and yield stability (VOLTAS *et al*., 2001; MALOSETTI *et al.*, 2004). Yield stability is considered an important attribute of good cultivars, but selection for yield stability is too time and money consuming to be carried out routinely.

Earlier attempts for establishing association between traits and markers across germplasm collections concerned oat, rice, maize, sea beet, and barley. In oat, BEER *et al*. (1997) found associations between markers and 13 quantitative traits in a set of 64 landraces and cultivars. In rice, VIRK *et al*. (1996) predicted the value for six traits using multiple linear regression. In maize, THORNSBERRY *et al*. (2001) found associations between Dwarf8 polymorphisms and flowering time. In sea beet, HANSEN *et al*. (2001) mapped the bolting gene, using AFLP markers in four populations. In barley, IGARTUA *et al.* (1999) concluded that marker-trait associations for heading date, found in mapping populations, were, to some extent, maintained in 32 cultivars. IVANDIC *et al.* (2003) found association between markers and the traits water-stress tolerance (chromosome 4H) and powdery mildew resistance in 52 wild barley lines. Chromosome 4H is, according to FORSTER *et al.* (2000), known for many loci involving abiotic stress tolerance, including salt tolerance, water use efficiency, and adaptation to drought environments.

This article is, to the best of our knowledge, the first publication on the extent of LD in a large collection of commercial barley cultivars, and on the usage of LD to explore the genome for markers linked to complex traits as mean yield and yield stability.

MATERIALS AND METHODS

Plant material and quantitative traits: Yield data of 146 modern European two-row spring barley cultivars were obtained from the official Danish variety trials over the period 1993 to 2000. Each year new cultivars were added to the trials, while others were discarded. The number of cultivars tested per year varied between 49 and 66. The number of locations at which a cultivar was tested varied between the years: 15 for 1993, 13 for 1994, and 5 for 1995 to 2000. Cultivars were tested in varying numbers of environments (year by location combinations), with a minimum of 5, a maximum of 50, and an average of 15 environments per cultivar. Each trial consisted of two replicates. More details can be found at http://www.planteinfo.dk.

The yield trials were either treated or not treated with chemicals to control leaf diseases. For treated and untreated trials, Finlay-Wilkinson regression coefficients were estimated as a measure for yield adaptability (b_i; FINLAY and WILKINSON, 1963). As a measure for yield stability, mean squared deviations from regressions were estimated (s_i^2; EBERHART and RUSSELL, 1966). Both statistics were based on the regressions of yields for individual genotypes in a trial on an environmental index, the latter supposed to express the general growing conditions in the trial. We estimated the environmental index by the environmental effects obtained from the fit of an additive model (phenotype = genotype + environment). Values of s_i^2 were log transformed for subsequent analyses.

Yield, stability and adaptability will be called YLD, STAB and ADAP, respectively, with subscript *tr* or *untr* referring to treated and untreated trials.

AFLP markers: The testing authorities supplied us with seed of all the cultivars tested in 1999. For cultivars not tested in 1999 seed was provided by the original breeders. Collection of DNA from leaf tissue and AFLP analysis were done as described by QI and LINDHOUT (1997). Fourteen primer combinations were employed: E33M54, E35M48, E35M54, E35M55, E35M61, E37M33, E38M50, E38M54, E38M55, E39M61, E42M32, E42M48, E45M49, and E45M55. Individual markers were identified following the profiles of QI and LINDHOUT (1997; also see http://wheat.pw.usda.gov/ggpages/Qi/). Markers were scored for presence (1) or absence (0) of a band. When two markers were very closely linked, or when they were allelic, the marker with most missing values was discarded. In total 286 polymorphic markers were

scored within this germplasm. For analyses, 236 markers with band frequencies between 5 and 95% were used.

Map position based on an integrated map: Map positions of markers were derived from an integrated map using three segregating populations: 1) L94 × Vada, 568 markers (QI and LINDHOUT, 1997); 2) Apex × Prisma, 252 markers (YIN *et al.*, 1999); 3) GEI119 × Gunhild, 137 markers (KOOREVAAR, 1997). The integrated map was constructed with the software package JoinMap (VAN OOIJEN and VOORRIPS, 2001). The assumption was made that AFLP markers with equal gel mobility were identical (ROUPPE VAN DER VOORT *et al*., 1997; WAUGH *et al*., 1997). The role of the integrated map is critical in our study. Every genetic map created with real life data, so probably including scoring and other errors, will give rise to some mistakes in the order of the marker loci. The integration of three different maps into one is another source of errors. For that reason, the AFLP-data were checked with great care, and any suspicious marker was removed. Furthermore, we carried out an extra control measure in the form of reference gels including all markers and all parental lines, to double check gel mobility and to minimize erroneous equal labeling of markers.

The number of markers common to two or three populations was 89, varying from eight on chromosome 1 to eighteen on chromosome 7. To constrain the number of possible map orders, five loci per chromosome provided a 'skeleton map' (fixed order) to which other markers were added. The fixed order loci were chosen to cover well the chromosomes from the map of QI *et al*. (1998). The latter map was aligned to the RFLP-map of the Proctor × Nudinka population (BECKER *et al.*, 1995).

Goodness of fit of proposed marker orders and positions on chromosomes were tested by a statistic that measured the overall discrepancy between map distances based on 'direct' estimates of recombination frequencies between individual markers on the one hand and the fitted map distances based on all available pairwise recombination frequencies on the other hand (STAM, 1993). This statistic approximately follows a chi-square distribution under the null hypothesis of a correct order of the markers on the map, with degrees of freedom equal to the total number of pairwise distances minus the number of adjacent pairs of markers on the chromosomes**.**

Population structure: To investigate possible structure in the set of cultivars various analyses were performed. First, an agglomerative hierarchical cluster analysis was performed on band incidence. As the measure for proximity, the Jaccard coefficient was chosen, while for the cluster algorithm average linkage, also known as UPGMA, was used (GORDON, 1981).

Second, a correspondence analysis was applied to the cultivar by marker matrix of band incidences (GREENACRE, 1984) and the plot of cultivar scores on the first two axes was used to investigate population structure. Finally, a Bayesian model based clustering was performed as described by PRITCHARD *et al*. (2000). The basis of this clustering method is the allocation of individual genotypes to groups in such a way that Hardy Weinberg equilibrium and linkage equilibrium are valid within clusters, whereas these forms of equilibrium are absent between clusters. As we worked with homozygous lines, we adapted the method to our situation by using the method to detect exclusively association between marker loci, while ignoring the within marker locus situation. The analysis was applied once to the complete set of all markers and once to a set of moderately independent markers.

Linkage Disequilibrium: A commonly used measure for quantifying and comparing LD in the context of LD-mapping is the squared correlation coefficient r^2 between pairs of biallelic markers (PRITCHARD and PRZEWORSKI, 2001). We have calculated r^2 between all pairs of loci and plotted it against the genetic distance in centimorgans to determine the map distance across which LD can occur within our set of cultivars.

Marker – trait associations: Pearson correlation coefficients were calculated among the traits YLD, ADAP, and STAB (treated and untreated) on the one hand, and band incidences for markers on the other hand. This is effectively equivalent to t-tests using marker incidence as grouping variable. The test statistic for Pearson correlations, $t^* = r \cdot (n-2)^{½} / (1-r^2)^{½}$, with r the correlation and n the number of observations, follows a $t_{(n-2)}$ distribution under the null hypothesis. To control for multiple testing, we tested at a false discovery rate (FDR) of 0.20 (BENJAMINI and HOCHBERG, 1995). The false discovery rate, q*, is defined as the expected proportion of true null hypotheses within the class of rejected null hypotheses. In practice, the procedure works as follows. Let $H_{(1)}, H_{(2)}, \ldots, H_{(m)}$ represent a series of hypotheses sorted by increasing p-value, $P_{(1)}, P_{(2)}, \ldots, P_{(m)}$, so that $P_{(1)} \leq P_{(2)} \leq \ldots \leq P_{(m)}$. Then the hypotheses $H_{(1)}, H_{(2)}, \ldots, H_{(k)}$ are rejected, where k is the largest i for which $P_{(i)} \leq (q^* i)/m$. In analogy to LOD profiles in QTL testing, association profiles were created by plotting p-values for marker-trait correlations against chromosome position. Association profiles graphically display the LD-region around an associated marker and can help in the assessment of the 'credibility' of a marker-trait association. To verify the relevance of our marker-trait associations, we checked the literature for QTLs in the regions near markers with significant trait association.

In addition to studying marginal marker-trait associations, i.e., correlations between markers and traits without correction for associations with other markers (cf. simple interval mapping),

YLD, ADAP and STAB were regressed on markers using multiple linear regression (*cf*. composite interval mapping) in an attempt to investigate conditional marker-trait associations. The final objective of this exercise was to obtain an estimate of the minimum and maximum theoretical trait values achievable by selective choice of marker alleles. Two methods for model construction were used. First, a stepwise regression procedure (MONTGOMERY and PECK, 1982) with an F-value for entering the regression model, F_{in}, of 4 and an F-value for dropping out of the model, F_{out}, of 1 was used. The marker set for model building was the full set of markers. In this way a model with a good combination of markers out of the complete set of markers was selected. Second, a regression model was constructed on the basis of the subset of markers that had significant correlation on an individual basis with the trait. In this second model, we used a combination of the individually best markers to predict the response, no selection was applied any more. The differences in predictions from both models will illustrate the necessity to account for correlations between markers. We chose as goodness-of-fit statistics the amount of explained variation adjusted for the number of regressors (R^2_{adj}; MONTGOMERY and PECK, 1982).

RESULTS

Yield, stability and adaptability: Table 1 presents several statistics concerning YLD, ADAP and STAB are given. Mean YLD_{tr} was higher than YLD_{untr}, as expected. The correlation between the treated and untreated version of YLD, ADAP and STAB was highly significant. YLD was weakly negatively correlated with STAB, treated and untreated.

Integrated map and map position: The final integrated map, based on three crossing populations, consisted of 811 AFLP-markers on a genome of 1052 cM (Kosambi mapping function) with eight gaps >10 cM and one gap >20 cM (data not shown). The quality of the integrated map was good, considering the low values for the goodness-of-fit statistics for map order across the chromosomes (see MATERIAL AND METHODS). Of the 236 markers that were found to be polymorphic across the cultivars, 123 appeared also on the integrated map of the crossing populations. The other 113 markers were not mapped, because they were apparently present or absent in both parents of the populations. Coverage figures for the 123 mapped markers showed twelve gaps between 10 and 20 cM, six gaps between 20 and 30 cM, and seven gaps of >30 cM. However, some of the 113 unmapped loci may be located inside those gaps.

Table 1: Descriptive statistics for yield (YLD), adaptability (ADAP), and stability (STAB). The yield trials were either treated (tr) or not treated (untr) with chemicals to control leaf diseases.
*, **, *** $P< 0.01, 0.001, 0.0001$

	YLD (kg/ha)		ADAP (b_i)		STAB [$\ln(s^2_i)$]	
	Treated	Untreated	Treated	Untreated	Treated	Untreated
Average	5779.3	5367.9	1.001	0.9978	1.8639	1.6041
Minimum	4841.0	4123.6	0.713	0.595	-1.8363	-3.5166
Maximum	6377.1	6037.9	1.49	1.254	4.0054	3.5825
Variance	2.7046	3.835	0.1061	0.0978	0.8192	0.8426
Correlations:						
YLD_{tr}						
YLD_{untr}	0.90 ***					
$ADAP_{tr}$	-0.06	-0.08				
$ADAP_{untr}$	-0.11	-0.19	0.76 ** *			
$STAB_{tr}$	-0.25 *	-0.34 ***	0.15	0.15		
$STAB_{untr}$	-0.29 **	-0.45 ***	0.02	0.06	0.60 ***	

Population structure: The 236 AFLP markers allowed unique identification of each cultivar. To investigate population structure, which could cause associations in the absence of linkage, we performed three types of analysis. A hierarchical cluster analysis with proximity defined by Jaccard coefficients and average linkage as clustering algorithm produced a dendrogram that hinted at the existence of two sub-groups. Correspondence analysis confirmed this split in the germplasm (Figure 1). The split could not be explained by geographic arguments, or by a separation of fodder and malting barleys. Various analyses using the Bayesian clustering methodology described in PRITCHARD *et al*. (2000) did not provide information on possible population structure. The posterior probabilities for the numbers of clusters remained about constant or kept steadily increasing with the number of clusters without individual varieties being allocated clearly to specific clusters. In both cases we concluded for absence of population structure.

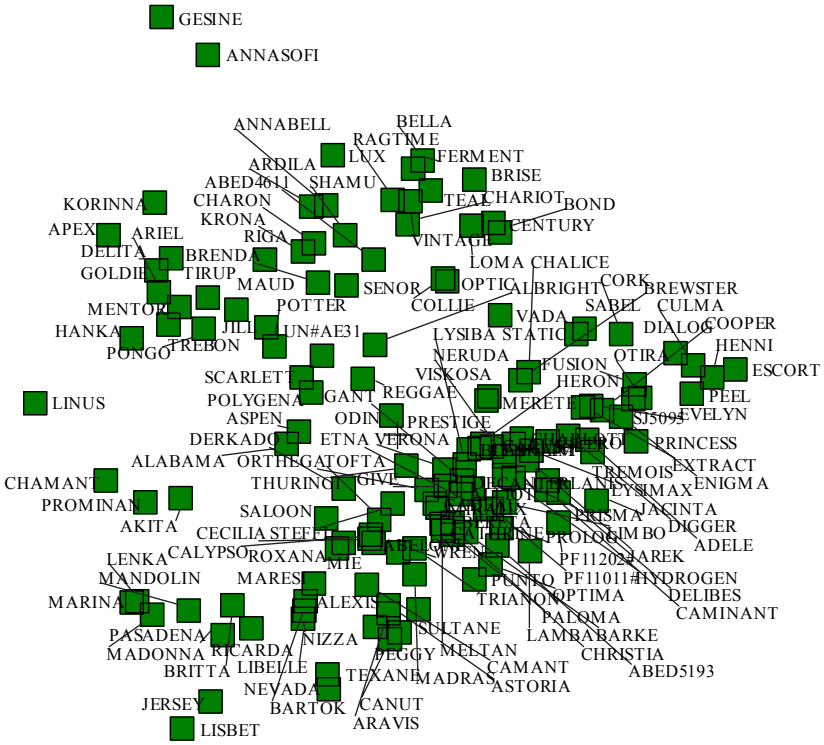


Figure 1: Correspondence analysis plot for 146 modern barley cultivars based on 236 AFLP-markers. The germplasm roughly falls apart in the subgroups at the top left and the bottom right in the plot.

Linkage Disequilibrium: Figure 2 gives LD as a function of genetic distance. LD was very common for distances <10 cM. Occasionally, LD occurred between loci farther apart. The r^2 between unlinked loci on different chromosomes was always <0.28, except for two markers on chromosomes 3 and 5, which had an r^2 of 0.40. These two markers also exhibited markedly

different band frequencies between the two subgroups found by the cluster and correspondence analysis. In contrast to *a priori* expectation, some marker pairs that were close together on the integrated map, were not correlated across the cultivars and so were in linkage equilibrium (LE). To check whether this unexpected apparent LE could be explained by misplaced markers due to the integration of maps from different mapping populations, we investigated the closely linked marker pairs in more detail. There were in total 53 marker pairs with distance < 1 cM, of which 32 had a significant correlation ($p<0.01$), while 19 pairs were not significantly correlated ($p>0.01$), and thus in LE. Thirteen of the nineteen pairs in LE contained two markers that were mapped using different populations, while six pairs consisted of two markers that were mapped in the same population. The three loci pairs in LE with the shortest distance between them (<0.06 cM) were all mapped in the L94×Vada population. This shows that the map integration in itself could not be the only explanation for apparent LE on short distances.

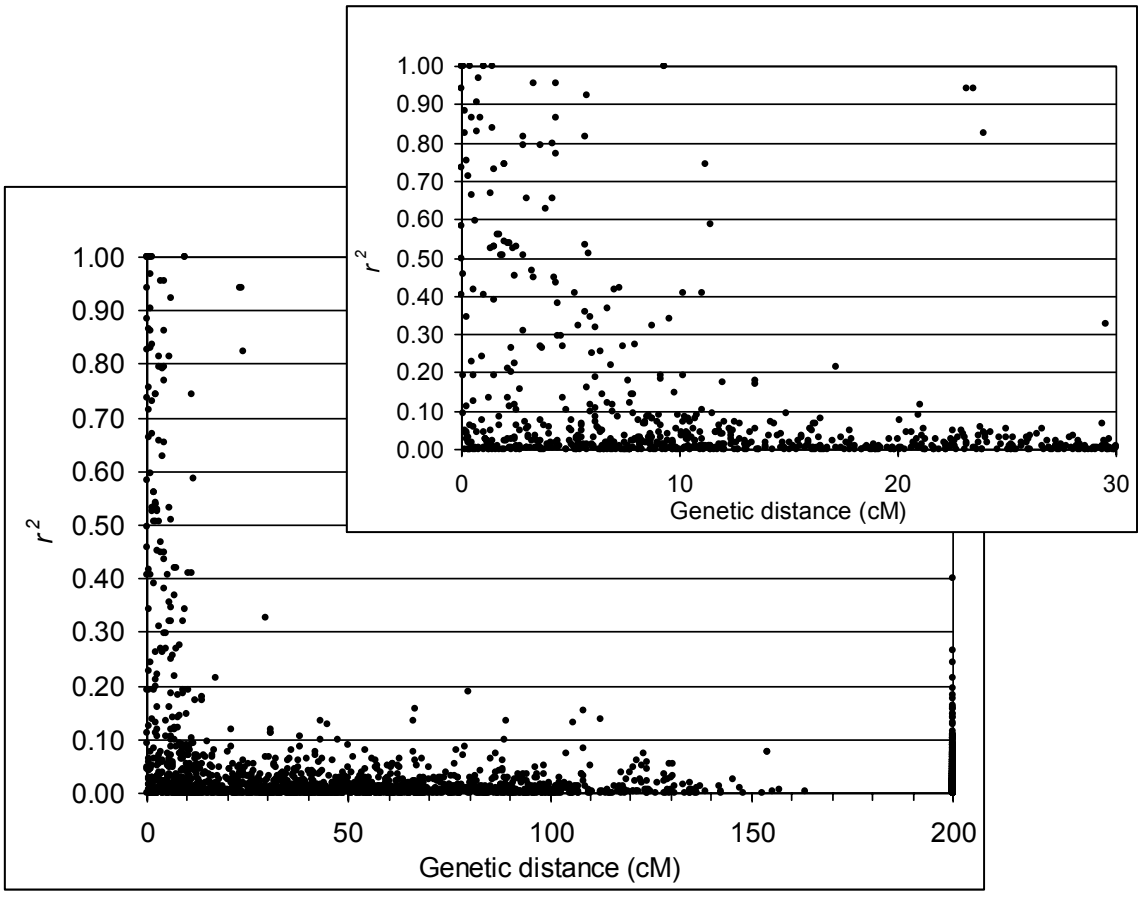


Figure 2: Linkage disequilibrium (r^2) as a function of genetic distance for 123 AFLP loci on the barley genome. LD has been determined with 146 modern barley cultivars; the genetic distance has been determined with an integrated map from three segregating populations. A genetic distance of 200 centimorgan was chosen to represent unlinked loci on different chromosomes. The inset provides an enhanced view of LD decay.

Table 2: Correlation of AFLP markers with yield, adaptability and stability. Only AFLP markers with a significant marker-trait correlation are given (P<0.01). A q indicates that the False Discovery Rate control value calculated according to BENJAMINI and HOCHBERG (1995) is <0.20; see MATERIAL & METHODS. The position on the chromosome is given in centimorgans (cM) from the top of the short arm.

Significance *, **, ***, ****, $P < 0.01, 0.005, 0.001, 0.0001$, respectively.

			Yield						Adaptability (b_i)						Stability [ln(s^2_i)]					
		Position	Treated			Untreated			Treated			Untreated			Treated			Untreated		
AFLP	Chrom	(cM)	*r*	*P*	*q**	*r*	*P*	*q**	*r*	*P*	*q**	*r*	*P*	*q**	*r*	*P*	*q**	*r*	*P*	*q**
E38M55-205	1	95.9	0.02			0.04			-0.24	*		-0.18			0.06			0.07		
E38M50-119	1	105.8	-0.10			-0.08			0.24	**		0.19			0.06			-0.12		
E42M48-270	2	6.7	0.13			0.15			-0.13			-0.16			-0.23	*	*q*	-0.06		
E35M54-412	2	33.7	-0.21			-0.28	***	*q*	-0.18			-0.08			0.15			0.21		
E45M55-086	2	36.2	0.08			0.12			-0.06			-0.13			-0.26	**	*q*	-0.15		
E42M32-333	2	94.5	-0.21			-0.22	*	*q*	-0.05			-0.03			-0.10			-0.05		
E35M48-250	3	19.5	-0.29	***	*q*	-0.29	***	*q*	-0.04			-0.06			0.10			0.14		
E45M55-142	4	45.8	0.10			0.16			0.05			0.05			-0.26	**	*q*	-0.22		
E33M54-282	4	47.8	-0.18			-0.25	*	*q*	-0.05			-0.02			0.15			0.21		
E42M48-139	4	48.3	0.10			0.12			-0.06			0.07			-0.22	*		-0.12		
E45M55-212	4	86.1	-0.22	*		-0.21		*q*	0.15			0.08			0.03			0.05		
E35M55-262	4	105.0	-0.14			-0.16			-0.13			-0.17			0.30	***	*q*	0.16		
E37M33-083	5	129.5	-0.17			-0.22	*	*q*	0.14			0.04			0.12			0.04		
E42M48-103	6	35.1	0.20			0.20			-0.04			-0.05			-0.32	***	*q*	-0.15		
E38M55-114	7	0.0	0.27	**	*q*	0.28	**	*q*	-0.06			-0.09			-0.06			-0.09		
E38M54-247	7	7.4	0.30	***	*q*	0.34	****	*q*	-0.06			-0.05			-0.11			-0.04		
E38M50-385	7	28.5	-0.15			-0.23	*	*q*	-0.17			-0.04			0.12			0.15		
E38M50-355	7	32.4	0.24	*		0.31	***	*q*	-0.06			-0.12			-0.21			-0.24		
E35M48-380	Unmapped		-0.27	**	*q*	-0.25	**	*q*	-0.20			-0.12			0.15			0.22	*	
E35M54-069	Unmapped		0.19			0.25	**	*q*	0.17			0.03			-0.17			-0.22	*	
E35M61-106	Unmapped		-0.22			-0.23		*q*	-0.22			-0.19			0.14			0.25	*	
E38M50-382	Unmapped		0.24	*		0.19			0.00			0.02			-0.05			-0.04		
E38M50-388	Unmapped		-0.07			-0.10			0.16			0.26	*		0.03			0.09		
E38M50-390	Unmapped		0.21			0.28	***	*q*	0.07			-0.04			-0.25	**	*q*	-0.26	**	
E38M55-110	Unmapped		-0.08			-0.08			0.23	*		0.15			0.00			0.00		
E42M32-187	Unmapped		0.20			0.26	**	*q*	0.15			0.02			-0.24	**	*q*	-0.22		
E42M32-271	Unmapped		-0.07			-0.05			0.02			0.12			-0.33	****	*q*	-0.25	**	

Association: Table 2 gives an overview of markers with their genome positions and correlations with traits. For the correlations, *p*-values and *q** values of the FDR analysis are presented. All markers with $q^* \leq 0.20$ belong to a group for which the proportion of false positives is no greater than 0.20. Only markers with a $p<0.01$ for at least one of the traits are shown.

Taking $q^* \leq 0.20$ as the threshold, 4 markers could be identified for YLD_{tr}, 15 markers for YLD_{untr}, and 8 markers for $STAB_{tr}$. No markers with significant association for $STAB_{untr}$ and $ADAP_{tr/untr}$ were found at $q^* \leq 0.20$. The most significantly correlated markers for $YLD_{tr/untr}$ were located at the top of chromosome 7 (7.4 cM) and chromosome 3 (19.5 cM). The most significant correlations for $STAB_{tr}$ were for a marker with unknown position and for markers on chromosomes 4 and 6. In general, markers were correlated with only one of the traits. Two unmapped markers formed an exception as they were correlated with both YLD and STAB. As none of the markers found associated with a trait differed in allele frequency between the two subgroups of cultivars identified by the cluster analysis and correspondence analysis, we concluded that the associations were not caused by substructure in the germplasm.

In Figure 3 the *p*-value of the correlation is given as a function of map position for a selection of trait-chromosome combinations. For YLD_{untr} a peak appeared on chromosome 2 at 34 cM with a rapid decline at 5 cM before the peak and 1 cM after the peak. The same peak showed up in the YLD_{tr} graph, but then with a lower magnitude. For both YLD_{tr} and YLD_{untr}, peaks appeared on chromosome 3 at 20 cM. No mapped markers were located before this peak, and the markers shortly beyond this peak showed a fast decrease in correlation, suggesting LD across a short distance. On chromosome 7 (5H), there were peaks at 7 cM and at 32 cM. The first peak at 7 cM was preceded by a significant correlation at 0 cM, suggesting LD over a distance of at least 7 cM. The second peak at 32 cM decayed already 1 cM before and 2 cM after the peak.

For $STAB_{tr}$, peaks were found at chromosomes 2, 4, and 6. All peaks faded rapidly. On chromosome 4 at 46 cM, the graph jumped up and down in the 46-48 cM area. After the first peak at 46 cM, a drop followed and then a second (smaller) peak followed at 48 cM.

In Table 3 an overview is given of the trait-associated markers, their map position, and related QTLs found in the same region by other authors. All of our YLD-associated markers and three of the STAB-associated markers were found in a region where at least once before a yield QTL has been reported. In addition, two of the three STAB-associated markers also coincided with a region known to exhibit QTLxE interaction (VOLTAS *et al*., 2001; MALOSETTI *et al.*, 2004).

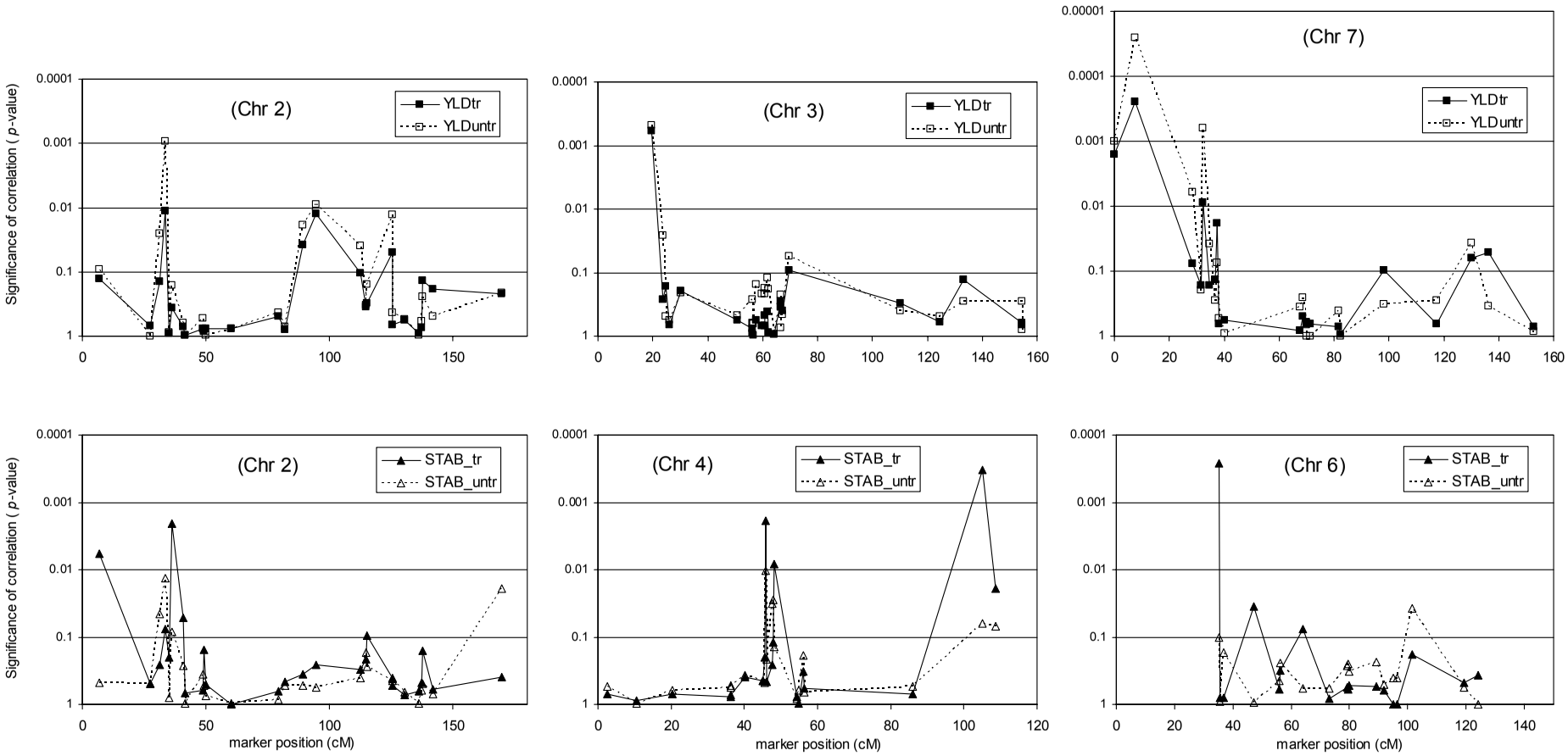


Figure 3: Association profiles showing the *p*-values of correlation between marker and trait against the position of the marker on the chromosome. (Top): yield on chromosomes 2, 3 and 7. (Bottom): yield stability on chromosomes 2, 4 and 6.

Table 3: Trait-associated markers and QTLs reported in literature in the same region. Significant marker-trait associations, marked with *q* in Table 2, linked to QTLs reported in literature. The traits were yield (YLD) and yield stability (STAB). The yield trials were either treated (tr) or not treated (untr) with chemicals to control leaf diseases. The Bin positions were determined using the Bin maps available at http://www.barleyworld.org, where an overview of known QTLs also can be found. The position on the chromosome is given in centimorgans (cM) from the top of the short arm.

Associated in this research			Reported in literature			
Chrom.	Position (cM)	Associated with	Bin	QTL for	Population	Author
2	33.7	$YLD_{tr/untr}$	3-4	Yield	Steptoe/Morex	Hayes *et al*., 1993
2	94.5	YLD_{untr}	10	Yield	Steptoe/Morex	Hayes *et al*., 1993
3	19.5	$YLD_{tr/untr}$	3	Yield	Blenheim/Kym	Bezant *et al*., 1997
4	47.8	YLD_{untr}	5	Yield	Harrington/Morex	Marquez-Cedillo *et al*., 2001
4	86.1	YLD_{untr}	8-10	Yield	Harrington/TR306	Tinker *et al*., 1996
5	129.5	YLD_{untr}	12-13	Yield	Harrington/Morex	Marquez-Cedillo *et al*., 2001
7	0 / 7.4	$YLD_{tr/untr}$	1-2	Yield	Blenheim/Kym	Bezant *et al*., 1997
				Yield	Apex/Prisma	M. Jarso (unpublished results)
7	28.5 / 32.4	$YLD_{tr/untr}$	3	Yield	Harrington/TR306	Tinker *et al*., 1996
2	6.7	$STAB_{tr}$	1	–	–	–
2	36.2	$STAB_{tr}$	3-4	Yield	Steptoe/Morex	Hayes *et al*., 1993
				G×*E*	Steptoe/Morex	Voltas *et al*., 2001
						Malosetti *et al*., 2004
4	45.8	$STAB_{tr}$	5	Yield	Harrington/Morex	Marquez-Cedillo *et al*., 2001
4	105.0	$STAB_{tr}$	11	Yield	Steptoe/Morex	Hayes *et al*., 1993
				G×*E*	Steptoe/Morex	Voltas *et al*., 2001
				Stress-response	several	Forster *et al*., 2000
6	35.1	$STAB_{tr}$	3-4	–	–	–

Multiple linear regression: Using all 236 markers, mapped and unmapped, we tried to describe variation in YLD, ADAP and STAB by a linear regression model including marker predictors. Stepwise regression resulted in regression models containing 18 to 20 markers (Table 4). The R^2_{adj}, adjusted for the number of predictors in the model, was 55/56% for $YLD_{tr/untr,}$ 45/40% for $ADAP_{tr/untr}$, and 56/58% for $STAB_{tr/untr}$. Therefore a large amount of the variation of these traits could be described by regression on markers (band incidence). By choosing the adequate marker profile, i.e. by creating a hypothetical marker genotype, the regression models could be used to predict minimum and maximum *theoretical* trait values. For $YLD_{tr,}$ the minimum and maximum value were 3631 and 7804 kg/ha, respectively. This is

much less and more, respectively, than the realized minimum and maximum of 5779 and 6377 kg/ha. So, if a genotype with all the favorable alleles for the selected set of markers could be created, this genotype would theoretically yield 7804 kg/ha. A similar transgression can be observed for the other traits.

Table 4: Predicting YLD, ADAP and STAB with multiple linear regression analysis using a subset of markers. The subset was chosen either using stepwise regression starting with the full set or based on the significance of correlation of the markers with the trait of interest. R^2_{adj} is the adjusted R^2. Realised mean, min(imum), and max(imum) values are given for comparison. Predicted min and max values were calculated using the regression model with the least/most favorable allele configuration. Yield is in kg·ha^{-1}. NA, not available. [a] ± standard error.

Trait	Selection of markers	No. of selected markers	R^2_{adj} (%)	Realized			Predicted with selection	
				Mean	Min	Max	Min	Max
YLD_{tr}	Stepwise	19	54.5	5779	4841	6377	3631 ±330[a]	7804 ±305
	$P < 0.05$	29	32.9				4782 ±418	6603 ±401
	$P < 0.01$	7	25.8				5414 ±65	6155 ±66
	$q^* < 0.20$	4	21.7				5588 ±37	6143 ±67
YLD_{untr}	Stepwise	18	56.4	5368	4124	6038	2494 ±343	7117 ±159
	$P < 0.05$	35	40.2				3944 ±299	6967 ±318
	$P < 0.01$	13	29.9				4400 ±209	6309 ±207
	$q^* < 0.20$	19	34.0				4395 ±221	6131 ±219
$ADAP_{tr}$	Stepwise	18	44.9	1.00	0.60	1.25	0.41 ±0.08	1.70 ±0.08
	$P < 0.05$	25	17.4				0.76 ±0.07	1.27 ±0.08
	$P < 0.01$	3	10.9				0.91 ±0.02	1.05 ±0.01
	$q^* < 0.20$	0	NA				NA	NA
$ADAP_{untr}$	Stepwise	18	40.3	1.00	0.71	1.49	0.55 ±0.06	1.68 ±0.08
	$P < 0.05$	14	21.3				0.84 ±0.03	1.30 ±0.05
	$P < 0.01$	1	1.8				0.99 ±0.01	1.05 ±0.03
	$q^* < 0.20$	0	NA				NA	NA
$STAB_{tr}$	Stepwise	18	55.7	1.60	-3.52	3.58	-6.29 ±0.72	9.10 ±0.87
	$P < 0.05$	21	25.7				-2.19 ±0.74	3.81 ±0.73
	$P < 0.01$	9	22.7				-0.76 ±0.46	2.40 ±0.37
	$q^* < 0.20$	8	23.2				-0.69 ±0.42	2.29 ±0.23
$STAB_{untr}$	Stepwise	20	57.5	1.86	-1.84	4.00	-7.27 ±0.80	9.55 ±0.80
	$P < 0.05$	18	18.2				-0.68 ±0.60	3.27 ±0.57
	$P < 0.01$	5	13.2				0.71 ±0.40	2.01 ±0.27
	$q^* < 0.20$	0	NA				NA	NA

Performing the regression with the subset of only those markers that showed significant marker-trait correlations on an individual basis, and so without further selection by a regression subset procedure, led in all cases to lower R^2_{adj}. In addition, predicted minimum and maximum values were less extreme, and in most cases did not exceed realized minima and maxima.

The final sets of selected markers by the two different strategies contained only a very modest overlap. Across the six traits under study, the maximum observed overlap amounted to five markers, roughly a quarter of the selected set by stepwise regression.

DISCUSSION

The main findings for the collection of barley cultivars we studied are: (1) LD extended to as far as 10 cM distance, (2) markers were associated with the traits yield and yield stability, and (3) the markers could be useful for selection.

LD: LD stretched over a distance of at least 10 cM. It is difficult to give the number of markers needed for a genome-wide scan, because LD will vary over the genome in relation to, among other factors, varying recombination rate and selection. Contrary to expectation, we also found LE between some closely linked markers. The same observation on LD at larger distances and LE at short distances was found in *Arabidopsis* (NORDBORG *et al*., 2002).

In comparison to other species, an LD interval up to 10 cM is large. Only in *Arabidopsis*-populations were larger distances found (>50 cM), but this was in populations founded by only a few genotypes and after extreme inbreeding (NORDBORG *et al*., 2002). In sugar beet lines, LD was <3 cM (KRAFT *et al*., 2000) and in maize LD diminished over a distance of 2,000 bp (REMINGTON *et al*., 2001). Many factors influence LD (see ARDLIE *et al*., 2002), but the most probable cause for the high level of LD in barley is the fact that it is an inbreeder. In addition, the current population of cultivars descended from a limited number of founding types (RUSSELL *et al.*, 2000) in which some haplotypes were lost and others were preserved, which will have increased LD. Finally, selection can increase LD, for instance by a hitchhiking effect, in which the alleles at flanking loci of a locus under selection can be rapidly swept to high frequency or fixation.

A major complication in LD studies like the one undertaken in this article is the appearance of false-positive marker-trait associations due to population structure. Bayesian cluster analysis

following PRITCHARD *et al*. (2000) gave no clue to the existence of such structure. A hierarchical cluster analysis and correspondence analysis did point to the existence of two subpopulations. However, fortunately, no trait-associated markers were in the set of markers discriminating between the two subpopulations, so we concluded that identified marker-trait associations were not a consequence of population structure, but very probably were indeed caused by linkage.

Association: Association between markers and traits (YLD, ADAP and STAB) was examined in three ways: (1) significance of marker-trait correlations, (2) LD-profiles over chromosomes (*p*-values against chromosome position), and (3) marker-trait associations found in other (QTL) studies.

Establishing a significance threshold for marker-trait associations is critical. In genome-wide LD-mapping, many markers are tested simultaneously, and some correction for multiplicity of testing is required. Well known approaches include Bonferroni-like procedures (*e.g*. HOLM, 1979), and permutation tests (CHURCHILL and DOERGE, 1994). Both kinds of approaches aim at controlling the type I error; that is, the probability of obtaining any false positive should be below a specified level, usually 0.05. As a result, the power (or the proportion of correctly identified positives) of these approaches can become very low. HOLLAND and COPENHAVER (1987) improved the Holm method with respect to power, but it remained conservative with impaired power. Instead of controlling the type I error, BENJAMINI and HOCHBERG (1995) advocated the control of the so-called false discovery rate (FDR). FDR was defined as the expected proportion of true null hypotheses within the class of rejected null hypotheses. The multiplicity control in FDR is directed at not surpassing a particular percentage of false positives (wrongly rejected null hypotheses, marker-trait associations that 'in reality' do not exist) within the set of identified positives. We argue that for our purposes - an exploratory genome-wide LD scan - an FDR control for multiplicity is more appropriate than type I control. Identification of associated markers in LD-mapping could be followed by the creation of a segregating population, polymorphic for the involved loci, in which the association is confirmed or refuted. In a similar vein, WELLER *et al.* (1998) demonstrated the utility of an FDR approach in the genetic dissection of complex traits.

In any LD mapping, it will be informative to examine the flanking markers of trait-associated markers. A chromosome-wide association profile containing a trait-associated marker will show whether the associated marker stands out or whether a smooth rise and fall appears before and after the marker. The latter pattern might point to real association, although

it still remains possible that LD extends over such a short distance that a ragged profile appears. Therefore, a smooth association profile confers confidence with respect to the identified marker-trait association, but a ragged profile not necessarily invalidates a found association.

Another kind of confirmation for identified associations came from reported QTLs from linkage analysis studies. All of the YLD-associated markers coincided with earlier reported yield QTLs. Most of the earlier reported QTLs were found in crosses within North American germplasm, while we only used European material. This suggests that, at least for yield, the North American germplasm genotypically resembles the European germplasm. An explanation might be that North American cultivars and European cultivars have common ancestors. Support for this hypothesis is given by FISCHBECK (2003), where it is stated that barley seeds were introduced to North America from many countries, especially from Central, Northern and Eastern Europe.

Furthermore, three of the STAB-associated markers were located in a region of known yield QTLs, and two of those three (on chromosomes 2 and 4) also coincided with a region earlier found to exhibit QTL×E interaction (Table 3). In addition, the STAB-associated marker on chromosome 4 is located in the region where several stress-responsive genes have been found (FORSTER *et al.*, 2000).

The question on the feasibility of selection on stability is an old one. Heritability for stability measures is generally low (BECKER and LEON, 1988; LEON and BECKER, 1988; LIN and BINNS, 1991; SNELLER *et al*., 1997). We have found markers associated with stability, but we do not know the nature of the genes linked to these markers. Three of five of the STAB-associated markers were in a region where yield QTLs also have been found, suggesting the presence of environmentally affected yield QTLs. The other two STAB-associated markers were in a region where so far no yield or yield-related QTLs were reported, suggesting environmentally affected regulatory genes. However, if yield QTLs were present at those locations, their irregular expression might be the reason for their non-identification so far.

Multiple linear regression: The question whether markers could be useful for predicting phenotypic responses, was answered with multiple linear regression, explaining traits by band incidence of markers. When subsets of 18 - 20 markers were selected from the total set of markers using stepwise regression, between 40 and 58% of the variation could be explained. We predicted the theoretical minimum and maximum for all traits according to the final regression model by choosing the favorable alleles (1 or 0 depending on the sign of the effect)

for the selected markers. The predicted minimum and maximum values were far beyond the observed minimum and maximum values. This could be explained by the absence of genotypes with exclusively (un)favorable alleles, but also by the fact that accumulating alleles almost always result in a lower effect than one might expect on the basis of adding up the effects of all the alleles. Nevertheless, selection on the basis of these markers might result in genotypes with superior yield and/or stability potential.

The marker-trait assocation models were fitted by regression under the assumption that individual varieties represented independent units. Of course, this assumption will have been violated by pedigree relations between the varieties. At first sight it may seem attractive to take account of these pedigree relations by inclusion of a relationship matrix in a mixed model analysis of the same data. However, several considerations have prevented us from changing from a standard regression model to a mixed model analysis. First, the pedigree information for collections of varieties as included in the present study is typically strongly incomplete. Second, the use of a relationship matrix is a logical consequence of the use of polygenic models for quantitative traits, but its use in oligogenic QTL models is far less natural. The estimator of the genetic correlation between genotypes in a polygenic model is a function of the expected identity by descent across the whole of the genome. However, in an oligogenic QTL model, the use of the expected identity by descent across the whole genome in the estimation of genetic correlations becomes questionable. In the latter case, the use of local identity-by-descent relations on the positions of the QTLs would seem more appropriate. These local identity-by-descent measures may be estimated from the allele composition of trait-associated markers as described by MILLIGAN (2003). The reliability of such estimates is still a matter of discussion and for that reason we preferred to use equally weighted independent varieties above disputably weighted and correlated varieties.

It may be contested that linkage will preclude the attainment of optimal allele configurations. However, closely linked markers were very seldom included in the stepwise regression models, because of the nature of this subset selection procedure. The predictions from the final stepwise regression models were thus supposed to represent a reasonably optimal combination of alleles on different loci. In contrast, the regression model based on the set of markers that were individually highly correlated with the trait did not take into account linkage relations between loci. Therefore, with this model, far less extreme minimum and maximum responses were obtained.

In conclusion, LD mapping seems to have clear potential for improving barley, especially for complex traits, like yield and yield stability, for which measurements are costly and time-consuming. Combining existing phenotypic variety trial data and genotypic marker characterizations within an LD approach may prove to be highly profitable.

Acknowledgements

Lars Kjær of the Danish Agricultural Advisory Center is gratefully acknowledged for the barley field data and seed of recently tested cultivars. More than thirty breeding companies from all over Europe have sent us seed of their cultivar(s), which we greatly appreciate. We thank Pim Lindhout and Fien Meyer-Dekens for performing AFLP analyses. This research is supported by the Technology Foundation STW, applied science division of Netherlands Organisation for Scientific Research NOW, and the technology programme of the Ministry of Economic Affairs of The Netherlands.

Chapter 3

Linkage disequilibrium mapping of loci for disease resistance, plant height, heading date and morphological characters in modern spring barley cultivars

Chapter 3

Linkage disequilibrium mapping of loci for disease resistance, plant height, heading date and morphological characters in modern spring barley cultivars

A. T. W. Kraakman, F. Martinez, B. Mussiraliev, F. A. van Eeuwijk, R. E. Niks

In preparation for submission.

Abstract

A set of 148 modern spring barley cultivars was explored for the extent of linkage disequilibrium (LD) between genes governing traits and nearby marker alleles. Associations of agronomically relevant traits (days to heading, plant height), resistance traits (leaf rust, Barley Yellow Dwarf (BYD)), and morphological traits (rachilla hair length, lodicules size) with AFLP markers and SSR markers were found. Known major genes and quantitative trait loci (QTLs) were confirmed, but also new putative QTLs were found. The LD mapping clearly indicated the common occurrence of *Rph3*, a gene for hypersensitivity resistance against *Puccinia hordei*, and also confirmed the QTL *Rphq2* for prolonging latency period of *Puccinia hordei* in seedlings. We also found strong indication for a hitherto not reported gene for resistance or tolerance to BYD on chromosome 2, linked to SSR marker HVM054. Our conclusion is that LD mapping is a valuable additional tool in the search for major genes and QTLs.

Introduction

Unraveling the genetics of traits is a quest for many. An important step is finding the location on the genome of genes that govern the trait phenotype. Molecular markers provide a helpful tool for this. When markers have been found in the neighborhood of a gene, they can be used for marker assisted selection.

In plants, many genes have been mapped on the genome. Especially multigenic traits have profited from the development of QTL mapping technology. Typically, for QTL mapping, two contrasting genotypes are crossed, e.g. two inbred lines, and the association between mapped marker loci and phenotypic traits allows the identification of QTLs on the genome. Recently, instead of bi-parental crosses, collections of cultivars, lines, or landraces are used to identify marker–trait associations in rice (Virk *et al.* 1996), oat (Beer *et al.*, 1997), and barley (Igartua *et al.*, 1999; Kraakman *et al.*, 2004). This methodology has become popular in human genetics under names as association mapping or linkage disequilibrium (LD) mapping, and led to a number of successes (see for a review Cardon and Bell (2001)). In plant genetics, using a collection of cultivars has a number of advantages over the use of a bi-parental cross. Firstly, broader genetic variation in a more representative genetic background will be available in the population. This implies that one is not limited to the marker and trait loci that happen to differ between two parents. Secondly, LD mapping may attain a higher resolution, because of the use of all meioses accumulated in the breeding history. Thirdly, historic phenotypic data on cultivars can be used to link markers to traits, without the need for new trials with special mapping populations. The methodology for associating markers and traits in a collection of cultivars are still under development. Some basic issues were addressed by Jannink and Walsh (2002).

In order to be able to identify marker - trait associations, LD has to occur in the germplasm. LD will tend to decay with genetic distance between the loci under consideration, because genetically distant loci are more likely to have recombined in the past than tightly linked loci. In populations, for any pair of linked polymorphic loci LD decreases over generations, because of accumulation of recombination. Finally the loci will be in linkage equilibrium (LE), *i.e.* alleles are not preferentially paired anymore. LD may occur in a population due to selection, for instance when an important trait is regulated by multiple loci, or due to recent introductions

of genotypes. The process of decrease of LD to reach LE depends on the opportunities of genetic recombination between the allele pairs of the loci under consideration. For effective recombination double heterozygotes are required, and these are much more common in allogamous than in autogamous plant species. Therefore, LD will tend to be more obvious after repeated inbreeding, as in autogamous species, than in outcrossing species. In several crops LD has been studied and quantified. In sugar beet, LD extended up to 3 cM (Kraft *et al.*, 2000), while in some *Arabidopsis* populations LD exceeded even 50 cM (Nordborg *et al.*, 2002). In barley, Kraakman *et al.* (2004) found LD on distances up to 10 cM.

The focus of this article is on LD mapping in modern spring barley cultivars for a variety of traits. We studied both major genes and QTLs in different categories: phenological traits, resistance traits, and morphological traits. Positions of markers that show high association with trait values were compared with those of loci known to govern the trait. We used pedigree information to understand whether certain cases of LD between a marker and a trait might be explained by particular ancestors of modern cultivars.

Our barley germplasm consisted of 148 two-row European spring barley cultivars, *i.e.* homozygous, diploid lines. All cultivars were released between 1985 and 1997, except the older cultivar Vada, which was released in 1957. Vada was incorporated as reference in our leaf rust experiments. The cultivars were grown all over North-West Europe during the last fifteen years, including the UK, France, Germany, Sweden, Denmark, and The Netherlands, and were therefore representative for modern West-European spring barley germplasm.

The traits under study were leaf rust (LR) resistance, resistance to barley yellow dwarf virus (BYD), plant height (PH), days to heading (DTH), and length of rachilla hairs and size of lodicules. Below some information on the genetics of the trait and known genes or QTLs is given.

Leaf rust

Leaf rust, caused by *Puccinia hordei* G.Otth, is becoming an increasingly important pathogen of barley in temperate areas of the world, causing severe yield losses in susceptible varieties. Two types of resistance have been described for barley: (1) qualitative resistance, expressed as a hypersensitive host response (chlorotic or necrotic spots), controlled by major genes, designated *Rph* or *Pa*, and (2) quantitative or partial resistance, non-hypersensitive, expressed as a reduced number of uredia, controlled by few to many genes, designated *Rphq* (Parlevliet and Van Ommeren, 1975; Qi *et al*., 1998b). The first type of resistance is race-specific (Jin and

Steffenson, 1994). The second type is mainly race-non-specific, although small isolate specific effects have been reported as well (Parlevliet, 1978).

In Europe, *Rph3*, *9*, and *12* are the most frequently used resistance genes in cultivars, although virulence against those genes occurs commonly (Niks *et al.*, 2000; Dreiseitl and Steffenson, 2000). The paucity of widely effective qualitative resistance genes has increased the importance of quantitative resistance genes in breeding programmes. The main components of partial resistance against leaf rust are longer latency period, reduced infection frequency, and low sporulation rate, but infection sites are not associated with plant tissue chlorosis or necrosis (reduced infection type, IT) as in resistance caused by *Rph* genes. Contrary to qualitative resistance, quantitative resistance remains effective even after widespread agricultural use over an extended period. In European spring barley cultivars, the level of partial resistance is generally high and still increasing (Niks *et al*., 2000).

At least 13 QTLs for resistance to leaf rust have been reported in barley mapping populations, on nearly all chromosomes at least one (Table 1).

Table 1: Known QTLs for partial resistance to leaf rust, their position on the genome, and the reporting authors. Qi[a]=Qi *et al.*, 1998b; Qi[b]=Qi *et al.*, 2000; Spa=Spaner *et al.*, 1998; Tho=Thomas *et al.*, 1995; Kic=Kicherer *et al.*, 2000; Bac=Backes *et al*., 2003.

Chromosome	Short arm	Around centromere	Long arm
1 (7H)	Qi[b]	Qi[a,b]	Qi[b]
2 (2H)	Spa	Qi[a,b], Kic	Qi[a,b], Kic, Bac
3 (3H)			Tho
4 (4H)	Qi[b], Kic	Qi[a]	
5 (1H)			
6 (6H)		Qi[b], Spa	Bac
7 (5H)	Qi[a,b]		Qi[b], Spa, Tho

Qi *et al*. (1998b) reported QTLs detected in a mapping population from a cross between partially resistant cultivar Vada and the susceptible line L94. The main QTLs for latency period were *Rphq2*, *3* and *4* on chromosome 2, 6, and 7, respectively, where the latter was only effective in adult plants. Vada is a cultivar often used as parent in European pedigrees, while L94 is an Ethiopian landrace. Vada was incorporated in our set of cultivars as a reference. The parents of the biparental mapping populations of Spaner *et al.* (1998), Thomas *et al.* (1995), Kicherer *et al.* (2000), and Backes *et al*. (2003) were not in our germplasm.

BYD

BYD is caused by a virus that belongs to a group of related viruses infecting barley, wheat, oat, rye, and other grasses. Major symptoms are stunted growth and leaf yellowing. Furthermore, the number of ears per plant and kernel weight is reduced, heading date is delayed, and the plants are more susceptible to abiotic stress and fungal diseases compared to healthy plants (D'Arcy, 1995). The symptoms are dependent on the plant developmental stage and the virulence of the virus isolate. Development of resistant cultivars is laborious, because testing and selection are difficult due to low heritability and the requirement for aphid mediated infection of plants. No monogenic complete resistance to BYD is known in barley. Genes conferring tolerance or resistance are *ryd1*, derived from the cultivar 'Rojo' (Suneson 1955), and *Ryd2*, identified in Ethiopian landraces (Rasmusson and Schaller, 1959). Due to its low efficiency, *ryd1* has probably rarely been used in barley breeding, and its chromosomal position remains unknown. *Ryd2*, however, has been incorporated into many barley cultivars from at least five donor lines (Burnett *et al.,* 1995). *Ryd2* has been mapped to the centromeric region of the long arm of chromosome 3 (Collins *et al.*, 1996). QTLs for BYD resistance in barley have been reported by Toojinda *et al.* (2000) and Scheurer *et al.* (2001). In addition to *Ryd2* on chromosome 3, they found QTLs on the top of chromosome 1 and on the long arm of chromosome 4. Furthermore, Scheurer *et al.* found two QTLs on both ends of chromosome 2, and Toojinda *et al.* found a QTL on top of chromosome 5.

A recently reported novel major gene, *Ryd3*, on chromosome 6 (Niks *et al.*, 2004) is unlikely to occur already in modern barley cultivars.

Plant Height and Days To Heading

Assessment of the traits PH and DTH is very straightforward, and is therefore among the standard observations in barley cultivar trials and QTL mapping studies. QTLs for PH in barley have been reported by Yin *et al.* (1999), Qi *et al.* (1998b), Kicherer *et al.* (2000), Hayes *et al.* (1993), and Bezant *et al.* (1996). QTLs for PH occurred on all chromosomes. Especially chromosomes 1, 2 and 7 were mentioned frequently, and many QTLs on these chromosomes mapped in different studies to similar positions, so they presumably refer to the same loci.

QTLs for DTH have been reported by many authors on all chromosomes (*e.g.* Qi *et al.*, 1998b, 2000; Kicherer *et al.*, 2000; Bezant *et al.*, 1996). Especially the short arm of chromosomes 1 and 2 appeared to carry QTLs for DTH.

Rachilla hair length and lodicula size

The rachilla and lodicules are organs that can be observed in harvested grains. A rachilla is a branch of the main rachis lying in the ventral side of a barley grain. A rachilla may be covered with short hairs or long hairs. A known locus for rachilla hair length, *srh*, is positioned on chromosome 7 around position 100 cM in BIN 9 (Wolfe *et al.*, 1996; Franckowiak, 1997).
Lodicules are a pair of thin fleshy structures which lie underneath the base of the lemma on the dorsal side of the grain. Lodicules can vary in shape and size. By swelling they play a role in opening of the floret during flowering. Lodicule size is therefore associated with the degree of fusarium headblight resistance and cleistogamy (Zhu *et al*., 1999). A known locus for cleistogamy (*Cly2*) is found around position 150 cM on chromosome 2 (Turuspekov *et al.*, 2004).

An overview of many QTLs in barley for many traits can be found at www.barleyworld.org/NABGMP/qtlsum.htm.

In this article we explore a set of modern spring barley cultivars for the extent of LD between genes governing the traits introduced above and nearby marker alleles using AFLP markers and SSRs. The results are compared with positions of genes that have been reported for these traits, and with pedigree information of the cultivars.

Materials and methods

Materials

A total of 148 modern European two-row spring barley cultivars were used, representing commercial germplasm used all over North-West Europe in the past fifteen years.

Leaf rust *(P. hordei)* isolate 'IVP2000' was used, which is virulent against the hypersensitivity genes *Rph9* and *12*, and avirulent against the commonly present *Rph3* gene (Niks *et al*., 2000). Only cultivars without effective hypersensitivity to 'IVP2000' provided the data on the basis of which the level of partial resistance was assessed.

BYD occurred by natural infection. The experiment was adjacent to the 2001 verification trial as reported by Niks *et al.* (2004), and the infection was therefore probably due to the same viral strain(s) as in that trial.

Seedling tests: Infection Type and Latency Period

Seedlings were raised in 37 × 39 cm plant boxes. Each box contained 6 or 7 cultivars, and L94 and Vada as references. Per cultivar two to five seedlings were available. About 10 days after sowing the first leaves were fixed in horizontal position and inoculated in a settling tower. Each box received 3.5 mg inoculum, which amounts to about 200 spores per cm^2. About ten times as much volume of *Lycopodium* powder was added to the inoculum to obtain a homogeneous distribution of the inoculum over the plant boxes. After incubation at a relative humidity of 100% overnight, the seedlings were transferred to a greenhouse.

IT: Infection Type was determined on the seedling leaves using the scale of 0-9 described by McNeal *et al*. (1971). This scale takes into account degrees of necrosis, chlorosis and sporulation associated with the infection sites. IT0: no symptoms; IT1 and IT2: minute or small necrotic/chlorotic flecks, respectively, no sporulation; IT3, 4, 5, 6: trace to much sporulation, respectively, and pustules surrounded by clear and serious necrosis and many flecks without sporulation (IT3) to clear chlorosis around the pustules and few flecks without sporulation (IT6); IT 7, 8 and 9: abundant sporulation, with some (IT7) to no chlorosis at all (IT9).

RLP: The latency period (LP) of seedlings was evaluated by estimating the period (hours) at which 50% of the ultimate number of pustules became visible. The relative latency period

(RLP) was calculated relative to the LP measured on L94 seedlings, where the LP on L94 = 100, as described by Parlevliet (1975). RLP is a component of partial resistance that is conveniently estimated in a monocyclic experiment in a greenhouse compartment. RLP was measured in two consecutive experiments in 2001 and 2003. We estimated the genotypic means for RLP from the fit of an additive model (phenotype = genotype + trial).

Field tests: AUDPC, BYD, PH, and DTH

In 2001, the cultivars were planted in a randomized block design with two replications, where the second replication contained fewer cultivars because of seed limitations. The plot size was 0.75×1.25 m^2. In each strip of 14 plots the reference cultivar L94 was represented as one of the plots. The plots of barley lines alternated with plots of oats to limit the inter-plot interference. One month after sowing, more than 350 young L94 plants, raised in pots, were inoculated with leaf rust isolate IVP2000 in the greenhouse. One week later, the sporulating L94 plants were placed in the alleyways between the field plots in front of each barley plot. Approximately 10 days later, when the plants in the L94 reference plots started to sporulate, the spreader plants were removed.

The amount of leaf rust infection in the field was determined three times on the upper three leaves of three tillers per plot. The leaf rust pustule counts were transformed according to the scale proposed by Parlevliet and Van Ommeren (1984). The Area Under the Disease Progress Curve (AUDPC) was calculated and used as trait for linkage disequilibrium mapping. Furthermore, days to heading (DTH) was evaluated as the number of days from sowing until 50% of the plants in the plot had headed. Plant height (PH) was measured as the number of centimeters between ground and the basis of the ear, averaged for eight plants in the final stage of plant development. Barley Yellow Dwarf (BYD) occurred by natural infection. BYD tolerance was assessed on adult plants on June 28, July 1st, and July 4th. The severity of symptom development was scored on a 0–5 scale. The score took into account both the proportion of the plot that was affected and the severity of the symptom development per plant. A detailed description of the different levels of scoring can be found in Niks *et al*. (2004)

For all traits, we estimated their genotypic means from the fit of an additive model (phenotype = genotype + trial). These genotypic means were used for the assessment of associations between traits and markers.

Seeds: Rachilla hair length and lodicule size

The observations of both rachilla hair length and lodicule size were done on at least two grains per accession. The rachilla was evaluated directly under binocular preparation microscope and the size of the lodicules after removing the basal part of the lemma. Rachilla was rated as short (0) or long (1), and lodiculae as small (0) intermediate (1) or large (2).

Genotyping and map construction

The cultivars were genotyped with fourteen AFLP- primer combinations and twelve microsatelites (SSRs). The AFLPs were run as described by Qi *et al*. (1997, 1998a), and resulted in 286 polymorphic markers. For analyses, 236 markers with band frequencies in between 5% and 95% were used. For 123 markers the map position could be established, using an integrated map of three segregating populations (see for details Kraakman *et al.*, 2004). The integrated map is available through http://www.dpw.wau.nl/pv/. Some extra AFLPs and SSRs were added to the integrated map on the basis of their significant association with AFLP markers on our integrated map. The position of AFLPs could be confirmed with information from other mapping populations like L94 × 116-5 (Qi *et al*., 2000) and L94 × C123 (unpublished). The SSRs we used were Bmac018, Bmag009, HVM14, HVM22, HVM65, HVM74, Bmag323, Bmac134, HVM054, HVM054, Bmac163, and Bmac316. The selection of SSRs was based on their map position according to Ramsay *et al.* (2000). We selected SSRs in regions where the integrated AFLP-map showed gaps, especially when in those regions interesting genes might be located. For instance, the first six SSRs mentioned above were selected in the neighborhood of the leaf rust QTL *Rphq3* on chromosome 6. The protocols used were according to Macaulay *et al.* (2001) and Ramsay *et al.* (2000). The electrophoreses were run on a DNA sequencer 4200 (LI-COR) under 1500 Volt, 40 mA, 45°C buffer temperature, and 25 Watt power. The data were collected with a single scanning laser on two infra-red frequencies, and automatically saved to a computer image.

Linkage disequilibrium mapping and statistical analysis

Marker-trait associations for AFLPs were calculated as Pearson correlation coefficients (r^2) between the trait values and band incidences for markers. This is effectively equivalent to t-tests using marker incidence as grouping variable. The test statistic for Pearson correlations, $t^* = r \cdot (n-2)^{½} / (1-r^2)^{½}$, with r the correlation and n the number of observations, follows a $t_{(n-2)}$ distribution under the null hypothesis. For the trait IT, we used Spearman rank correlation

instead of Pearson correlation. For SSRs, we treated the individual SSR band sizes as separate binary indicator variables. To assess the association of each SSR with traits, we used multiple regressions on the allele indicator variables. The corresponding R^2 and *p*-value were used for further analyses.

To control for multiple testing, we used a procedure that estimates the false discovery rate (FDR) (Benjamini and Hochberg, 1995). The FDR is the proportion of false positive tests among the significant tests. FDR for individual tests can be expressed as a *q*-value that represents the expected proportion of false positives incurred when calling the observed value for the test-statistic significant. FDR *q*-values can be used alongside the commonly used *p*-values. FDR can be useful in genome wide studies where often many markers are tested for association with many traits (Weller *et al.*, 1998; Storey and Tibshirani, 2003; Kraakman *et al.,* 2004). Storey and Tibshirani (2003) adapted a version of FDR that takes into account the density distribution of realized *p*-values to estimate the FDR for a batch of tests. We used their software called QVALUE, with default settings (π_0-method=smoother, FDR level=0.05). This software can be downloaded at www.genomine.org/qvalue.

Population structure

Population structure can result in false associations, and should therefore be prevented or taken into consideration if it occurs. To investigate possible structure in the set of cultivars we performed three analyses. First, a cluster analysis, second a correspondence analysis, and finally an analysis based on a Bayesian model (Pritchard *et al.*, 2000). Details about these analyses can be found in Kraakman *et al.* (2004).

Results

Assessment of resistance and plant development traits

The set of cultivars showed variation for all traits observed (Table 2). The IT of leaf rust infection was for 33 cultivars low (≤6; hypersensitive reaction) and for 113 cultivars high (>6; compatible reaction). The rachilla hair length was short for 17 cultivars and long for 131 cultivars. The lodicule size was small/intermediate/large for 33/13/101 cultivars, respectively.

Table 2: Descriptive statistics for Infection Type (IT), Relative Latency Period (RLP), and AUDPC for leaf rust resistance, and Barley Yellow Dwarf Virus tolerance (BYD), Days to Heading (DTH), and Plant Height (PH) of 148 cultivars. DTH is given in days, PH in cm.

			Percentiles					
Trait	N	Minimum	5%	25%	50%	75%	95%	Maximum
IT	146	1	2	9	9	9	9	9
AUDPC	140	69.9	101.0	189.4	223.2	234.6	245.8	258.7
" " IT>6	107	166.2	194.7	217.5	229.8	236.5	247.2	258.7
RLP	111	101.6	104.3	107.4	109.5	113.0	118.0	124.0
BYD	140	0.33	0.52	1.99	2.67	3.00	3.50	4.00
DTH	140	51.0	52.0	53.0	54.5	56.5	60.0	64.0
PH	140	63.8	66.4	72.2	76.3	79.5	84.2	93.1

IT and AUDPC were significantly and positively correlated: effective hyper-sensitivity, expressed as low IT, resulted in low AUDPC scores in the field (Table 3, Figure 2). Also significantly correlated were AUDPC with RLP, DTH with PH, and Rachilla hair length with BYD.

Table 3: Correlations between trait values in 148 spring barley cultivars. For IT Spearman rank correlation is given, for all other traits Pearson correlation coefficient.
Significance *, **, *****, $P < 0.01, 0.001, 0.000001$, respectively.

Trait	IT	AUDPC	AUDPC IT>6	RLP	BYD	Heading	Height	Rachilla length
IT								
AUDPC	0.70 *****							
AUDPC IT>6	-0.07	1.00						
RLP	-0.13	-0.29 *	-0.32 *					
BYD	0.04	0.07	0.12	-0.25				
DTH	0.13	0.14	-0.05	0.08	-0.02			
PH	-0.04	-0.01	0.09	-0.06	-0.12	-0.22 *		
Rachilla length	0.00	0.02	0.00	-0.23	0.31 **	0.06	0.00	
Lodicule size	-0.08	-0.01	0.25	-0.12	-0.09	-0.10	0.14	-0.05

In Figure 2 it is shown that cultivars with an IT>6 had an AUDPC higher than 166, and cultivars with an IT<= 6 had an AUDPC lower than 160. In the latter group was one exception, Hanka, which had a high AUDPC (215), despite a hypersensitive reaction (IT=3) in the seedling stage.

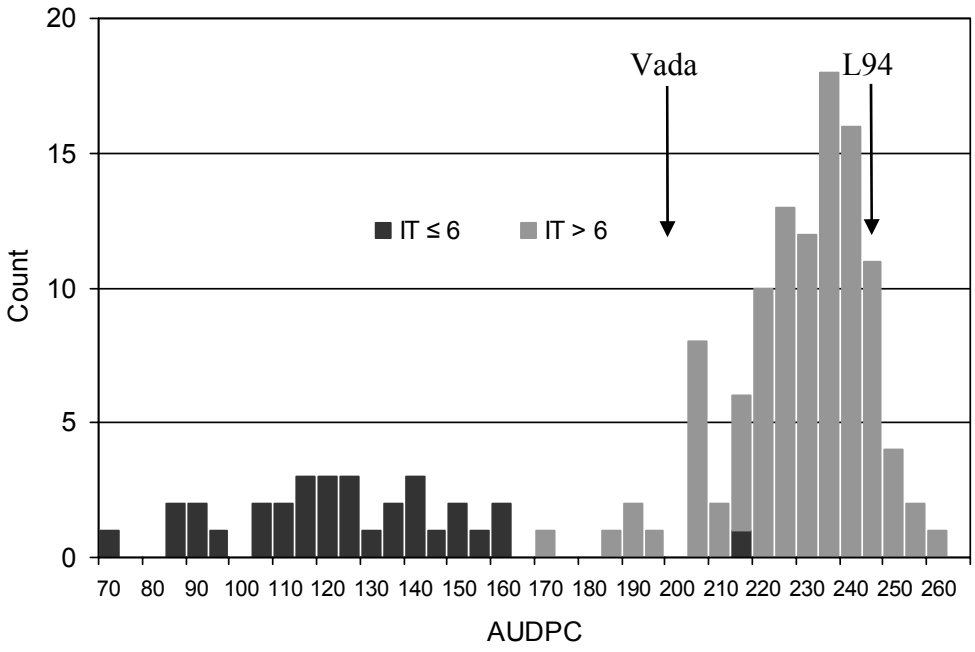


Figure 2: Histogram of AUDPC: Frequency distribution of phenotypes for AUDPC of all cultivars. Values of Vada (202.2) and L94 (247.2) are shown by arrows.

AFLPs and SSRs

Information on the AFLPs used can be found in Kraakman *et al.* (2004). In this study the set of molecular markers was extended with twelve SSRs (Table 4). The number of alleles found per SSR ranged from two to seven. For those cultivars in our set that had also been genotyped for SSR markers by other authors (Ramsay *et al.*, 2000; Russell *et al.*, 2000; Macauly *et al.*, 2001), the size of amplification products was in agreement with those reports.

Table 4: Overview of SSRs used, their position on the genome, and the number of alleles found in the cultivars. The position on the chromosome is given in cM from the top of the short arm.

SSR	Chromosome	Position (cM)	Number of alleles
BMAC018	6	79.4	3
BMAC134	2	49.2	7
BMAC163	7	31.6	2
BMAC316	6	36.8	5
BMAG009	6	79.4	2
BMAG323	7	42.5	6
HVM014	6	79.4	2
HVM022	6	79.4	3
HVM036	NA	NA	3
HVM054	2	95.5	5
HVM065	6	79.4	4
HVM074	6	79.4	3

Population structure

The 236 AFLP markers allowed unique identification of each cultivar. To investigate population structure, which could cause associations that are not due to linkage, we performed three types of analysis. Various analyses using the Bayesian clustering methodology described in Pritchard *et al*. (2000) did not suggest a distinct subpopulation structure. The posterior probabilities for the numbers of clusters remained either about constant, or kept steadily increasing with the number of clusters without individual varieties being allocated clearly to specific clusters. In both cases we concluded for absence of subdivision of the population into distinct germplasm groups. However, hierarchical cluster analysis as well as correspondence analysis hinted at the existence of two sub-groups (see Kraakman *et al*., 2004). This split in the germplasm could not be explained by geographic arguments, or by a separation of fodder and malting barleys.

Overview of all associations:

All significant marker-trait associations are shown in Table 5, for both mapped and unmapped markers. Furthermore, the mapped associated markers are graphically shown in Figure 3 with relevant QTLs reported in literature. For all traits the most significant marker-trait associations will be discussed.

Table 5: All marker-trait associations found in 148 cultivars for markers where at least one association had a false discovery rate (FDR) of less than 5% (q<0.05). AFLP-trait associations are expressed as the Pearson correlation coefficient between AFLP value and trait value, except for IT where Spearman rank correlation is used. For SSRs, the trait values were explained with multiple regressions with the alleles as indicator variables. The square root of R^2 is given in the same column as where r is given for AFLPs. Marker map positions are based on an integrated map (Kraakman *et al.*, 2004), except for the markers shown in italics, which were positioned based on the map of L94 × 116-5 (Qi *et al.*, 2000) or L94 × C123 (unpublished).

The significance of correlation/regression is shown as p-value (simple testing), and as q-value (multiple testing). Significance of p *, **, ***, ****, p < 0.01, 0.001, 0.0001, 0.00001, respectively. Correlations printed in bold type had an FDR less than 5%.

The RLP could only be measured on those cultivars that showed IT>6.

Known QTLs: Reference to authors who reported a QTL in the same region. The authors were: (a) Kicherer *et al*., 2000, (b) Spaner *et al*., 1998, (c) Toojinda *et al*., 2000, (d) Scheurer *et al.*, 2001, (e) Qi *et al*., 1998b, (f) Bezant *et al.*, 1996, (g) Yin *et al.*, 1999, (h) Chelkowski *et al*., 2003, (i) Jin and Steffenson, 1994, (j) Hayes *et al.*, 1993, (k) Backes *et al*., 2003, (l) Tinker *et al*., 1996, and (m) Marquez-Cedillo *et al*., 2001.

Table 5: Associations between mapped markers and traits.

				AUDPC high IT			RLP			IT			BYD			DTH			PH		
No.	marker	chr	pos	*r*	*p*	known		*r*	*p*	known	*r*	*p*	known		*r*	*p*	known	*r*	*p*	known	*r*	*p*	known
1	E38M54-472	1	9	-0.17				0.27	*		-0.13			1	**-0.48**	****	d	0.07			0.01		
2	*E42M32-231*	*1*	*24*	0.01				0.03			-0.02				0.05			0.00			1	**-0.27**	* j,g
3	E38M50-119	1	106	0.04				-0.24			0.05				-0.18			0.08			2	**0.30**	** f,g,e
4	E39M61-255	1	151	-0.15				-0.07			**0.74**	****	h,i		0.04			0.10			-0.04		
5	E35M54-183	1	163	-0.04				-0.03			-0.06			2	**-0.28**	**	c	-0.16			0.12		
6	E35M61-228	2	49	0.25			1	**-0.34**	*	a,e	0.05			3	**0.47**	****	d	-0.02			0.05		
7	E35M54-078	2	89	0.04			2	**0.34**	**	a	-0.18				-0.12			0.12			-0.22	*	
8	HVM054	2	95	0.32				0.37			0.22			4	**0.67**	****		0.25			0.20		
9	E42M48-405	2	136	-0.22			3	**0.33**	**	a,e,k	0.18				0.15			0.00			-0.11		
10	E38M54-294	2	137	0.04				-0.10		a,e,k	**-0.29**	**			0.10			0.02			0.15		
11	E38M55-251	2	138	-0.22				**0.42**	****	a,e,k	0.22				-0.06			-0.03			-0.10		
12	E42M48-376	2	142	-0.21				**0.37**	***	a,e,k	**0.28**	**			-0.03			-0.08			-0.15		
13	E45M55-172	3	56	-0.06			4	**0.38**	***		-0.03				-0.07			-0.04			0.03		
14	E35M54-310	3	61	0.07				**-0.37**	***		0.06				0.08			0.05			-0.03		
15	*E45M55-154*	*3*	*76*	-0.03				**-0.35**	**		0.18				0.17			0.01			0.14		
16	*E45M49-164*	*4*	*0*	0.07				0.10			0.21			5	**0.34**	**		0.06			-0.21		
17	E35M55-302	4	45	0.19				-0.15			-0.03			6	**0.25**	*		-0.08			-0.07		
18	E38M50-355	7	32	-0.18				0.18			-0.14				-0.12			1	**-0.29**	*	l,m	0.16		
19	BMAG323	7	43	0.26				0.10			0.20				0.22			0.32			3	**0.37**	* m
20	E42M48-087	7	69	**-0.29**	*			0.06			-0.11				0.02			0.22	*		-0.19		
21	E42M32-200	7	70	**-0.38**	***			0.25			0.14				0.14			0.04			0.00		
22	E38M55-128	7	71	**-0.38**	***			0.02			0.06				0.18			-0.04			-0.12		
23	E42M48-279	7	82	**-0.30**	*			0.10			0.02				-0.02			-0.06			-0.04		

Table 5 - ***continued:*** Associations between unmapped markers and traits.

			AUDPC high IT			RLP			IT			BYD			DTH			PH		
No.	marker	chr	pos	*r*	*p*	known	*r*	*p*	known	*r*	*p*	known	*r*	*p*	known	*r*	*p*	known	*r*	*p*	known
24	E33M54-063	unm.		-0.07			**0.36**	**		-0.03			-0.08			-0.03			-0.01		
25	E33M54-095	unm.		-0.04			-0.03			-0.03			0.19			**-0.28**	**		0.00		
26	E33M54-421	unm.		0.24			-0.19			-0.01			**0.29**	*		0.10			-0.06		
27	E35M48-095	unm.		0.03			-0.13			-0.08			-0.01			-0.14			**0.28**	**	
28	E35M48-111	unm.		-0.02			-0.08			-0.09			0.08			-0.11			**0.33**	***	
29	E35M48-233	unm.		0.12			-0.01			0.12			0.01			-0.11			**0.29**	*	
30	E35M54-147	unm.		0.18			-0.07			-0.08			0.02			0.09			**-0.32**	*	
31	E35M54-265	unm.		0.10			-0.10			-0.03			**0.28**	*		0.01			0.04		
32	E35M55-306	unm.		-0.10			0.18			-0.03			**-0.28**	**		-0.10			0.08		
33	E37M33-256	unm.		0.11			-0.03			0.07			**-0.31**	**		-0.10			0.13		
34	E38M50-135	unm.		-0.13			-0.03			0.00			0.17			0.03			**-0.28**	*	
35	E38M50-149	unm.		-0.09			**0.30**	*		0.18			-0.11			0.03			0.14		
36	E38M50-263	unm.		0.29	*		**-0.35**	**		-0.21			-0.09			-0.06			0.16		
37	E38M50-456	unm.		-0.02			**0.31**	*		**-0.35**	***		-0.04			-0.08			0.07		
38	E42M32-069	unm.		0.03			**-0.37**	**		0.15			0.10			-0.02			-0.17		
39	E42M32-156	unm.		-0.08			**0.37**	***		-0.06			-0.09			-0.04			0.02		
40	E42M32-160	unm.		-0.05			**0.35**	**		0.03			0.02			0.01			-0.03		
41	E42M32-211	unm.		0.02			-0.16			0.07			**0.29**	**		0.00			-0.06		
42	E42M32-228	unm.		-0.07			0.02			-0.11			0.04			-0.20			**0.30**	**	
43	E42M32-273	unm.		0.08			**-0.32**	*		0.07			0.06			-0.08			0.07		
44	E45M49-176	unm.		0.03			0.00			-0.12			**0.39**	***		-0.09			-0.10		
45	E45M49-226	unm.		0.12			**-0.32**	**		0.05			0.17			0.13			0.04		
46	E45M49-388	unm.		-0.08			0.09			-0.15			0.05			-0.13			**0.28**	*	
47	E45M55-164	unm.		0.09			-0.12			0.10			0.12			-0.06			**0.27**	*	

Linkage Disequilibrium mapping of leaf rust resistance

IT: *Rph*3 confirmed

Marker E39M61-255 on chromosome 1 at 151 cM was highly associated with IT (Spearman's r=0.74; p<1E-6) (Table 6). In this genome region a major gene for leaf rust resistance, *Rph3*, has been identified (Chelkowski *et al.*, 2003). As the marker-trait association of this marker was very high in this set of cultivars, we can conclude that *Rph3* was responsible for most of the variation in IT.

Table 6: Occurrence of cultivars with a high or low Infection Type of *Puccinia hordei* isolate IVP2000 and allele for marker E39M61-255.

	Infection Type			
E39M61-255	Low (≤6)	High (>6)	Unknown	Total
0	27	9	-	36
1	4	101	2	107
Unknown	2	3		5
Total	33	113	2	148

Resistance to leaf rust (=low IT) corresponded to absence of the band (AFLP value zero). Table 6 shows that 27 of 36 cultivars with E39M61-255 = 0 had a low IT. Note that nine cultivars with E39M61-255 = 0 still had a high IT. All nine had the corresponding high AUDPC. Out of 107 cultivars with the susceptible phenotype for marker E39M61-255 101 had a high IT, and four cultivars had low IT.

Another significant peak in association with IT was found on chromosome 2 around 138 cM. This peak coincides with a peak in RLP, and will be discussed later. Finally, one of the unmapped markers (E38M50-456) was highly associated with IT. In our set of cultivars this marker was moderately correlated with E39M61-255 (r=-0.28, p=0.001), suggesting linkage to *Rph3*.

RLP, AUDPC: search for partial resistance for leaf rust

RLP and AUDPC are quantitative measures for leaf rust resistance. RLP is positively correlated with resistance level, since it is based on the time required for maturation of rust pustules in a monocyclic experiment. AUDPC is based on the level of infection in a field trial, and hence, is negatively correlated with resistance level. RLP was measured on cultivars with

high IT, so without effective genes for hypersensitivity resistance. This way we concentrated on variation in level of partial resistance in the cultivar set, without confounding with hypersensitive resistance. For the same reason the LD analysis on AUDPC was carried out on those cultivars that had high IT in seedling stage.

AUDPC was significantly associated with four AFLP-markers on chromosome 7 between 69 and 82 cM. Presence or absence of E42M32-200 resulted in an average AUDPC of 216 or 229, respectively ($LSD_{0.05}$=7.2). For E38M55-128 the AUDPC was 219 or 231, respectively ($LSD_{0.05}$=6.2). This locus had not been implicated before in mapping studies on field experiments in which partial resistance level was determined. Conversely, QTLs that had been reported to reduce AUDPC of *P. hordei*, like those of Vada reported by Qi *et al.* (1998b), did not appear in the present study.

RLP, measured in seedlings, was significantly associated with many markers (Table 5). If we consider markers within 20 cM of each other as indicative for one and the same putative QTL, we can distinguish three QTLs on chromosome 2 (Figure 3 and 4) and one QTL on chromosome 3. In addition, nine unmapped markers were associated with RLP. Out of the four QTLs indicated by LD mapping, three were at positions where QTLs for quantitative resistance to *P. hordei* had been reported before.

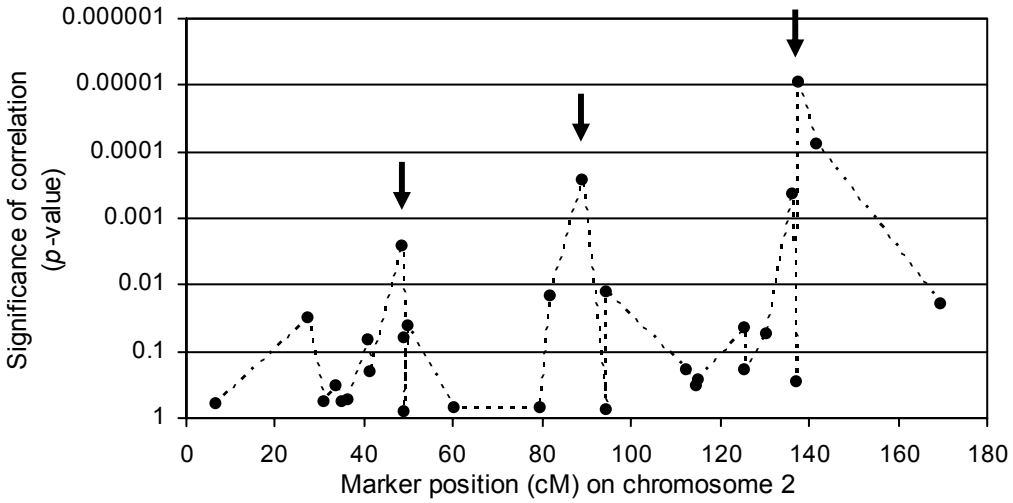


Figure 3: Association profile of RLP on chromosome 2. The significance of correlation between marker and trait is shown. The position of three putative QTLs is indicated with an arrow.

In Figure 3 an association profile is given for chromosome 2. Note that around 138 cM all markers except one had a very low *p*-value, so were associated with RLP. The markers with a

significant *p*-value were all 'Vada-markers', i.e. the band for such a marker also occurred in Vada. The one marker that was not significantly associated with RLP (E38M54-294) was an 'L94-marker'. The marker that was most significantly associated with RLP was E38M55-251 (r=0.42 p=9E-6). The average RLP on the 40 cultivars with a band for that marker was 112.6, while the 62 cultivars without the band gave an average RLP of 108.9 ($LSD_{0.05}$=1.63). Furthermore, when all possible marker configurations for the three Vada-markers and the one L94-marker were investigated, only four haplotypes occurred in the set of cultivars (Table 7). The most common configuration was the haplotype of L94. The 33 cultivars with this haplotype gave a low RLP. A comparably low RLP was found on the haplotype with only one marker difference (E38M54-294 changed from 1 to 0). The Vada haplotype occurred 19 times, and the cultivars with this haplotype gave a significantly higher RLP than the ones with the L94 haplotype. A similarly high RLP was observed on the 15 cultivars with a haplotype differing only for marker E42M48-205. Apparently, the allele change of this marker did not affect the level of partial resistance.

Table 7: Frequency and average RLP and AUDPC values of the most common marker haplotypes in the *Rphq2* region. All other configurations occurred four or less times in the set of cultivars. For RLP the difference in subgroups is given based on Bonferoni corrected LSD(0.05).

E42M48-205 (Vada) 136.2 cM	E38M54-294 (L94) 137.4 cM	E38M55-251 (Vada) 137.7 cM	E42M48-376 (Vada) 141.8 cM	Count	Average RLP	Average AUDPC
1	0	1	1	19	112.5 [a]	221.4
0	0	1	1	15	112.2 [a,b]	223.7
0	1	0	0	33	109.6 [b,c]	228.9
0	0	0	0	20	107.9 [c]	229.4

Marker E38M55-251 has been reported to be the peakmarker for an important QTL, *Rphq2*, for partial resistance against *P. hordei* (Qi *et al*., 1998b). This locus is at very short distance of powdery mildew gene *MlLa*, which has been reported to be contributed by *Hordeum laevigatum*, one of the parents of Vada (Arru *et al.* 2002; Giese *et al.* 1993; Jensen and Jørgensen, 1991). Therefore, also *Rphq2* is likely to have been derived from *H. laevigatum*. In our set of cultivars, 113 cultivars had *H. laevigatum* in their ancestry, of which 70 also had Vada as ancestor. However, analyzing only cultivars for which we both know their ancestry and their state for marker E38M55-251, 42 of the 76 cultivars derived from *H. laevigatum* did

not carry E38M55-251 (Table 8). There was no association between the presence of *H. laevigatum* in the ancestry and the presence of E38M55-251 (χ^2 not significant). The average RLP for cultivars derived from *H. laevigatum* was not different from cultivars missing *H. laevigatum* in their ancestry (see HSD in Table 8). However, when both the cultivar was derived from *H. laevigatum* and it carried marker E38M55-251 there was a significantly higher RLP. Apparently, the presence of *H. laevigatum* is only indicative for higher RLP when combined with E38M55-251.

Table 8: Frequency of cultivars carrying *Rphq2* peakmarker E38M55-251 with or without *Hordeum laevigatum* or Vada in their ancestry. Per subgroup the average relative latency period (RLP) is shown. Tukey HSD (0.05) is given for the RLP averages per subset.

Ancestry contains *H. laevigatum*	E38M55-251	Count	RLP	HSD
0	0	11	107.7	a
1	0	42	109.3	a,b
0	1	3	111.5	a,b
1	1	34	112.8	b
Vada				
0	0	39	107.3	a
1	0	36	111.0	b
0	1	12	112.3	b
1	1	30	112.9	b

Combined association for IT and RLP: chromosome 2 at 138 cM (*Rphq2)*

Three Vada-markers on chromosome 2 around 138 cM were positively correlated with RLP, and also positively correlated with IT. This implies that they are associated to resistance (prolonged LP) and to susceptibility (high, compatible IT) at the same time. In our set of cultivars, only five out of 30 cultivars (17%) with low IT had marker E38M55-251, the peak marker for *Rphq2*, while 40 out of 104 cultivars (38%) with high IT had this marker. The difference is highly significant (χ^2; p=1.5E-5). This suggests that if in breeding programs advanced germplasm carried an effective gene for hypersensitivity (low IT) breeders relatively frequently did not select *Rphq2*, since they could not judge the level of partial resistance.

Linkage Disequilibrium mapping of other traits

BYD: many associated markers

Many markers were found to be associated with the level of BYD. Assuming that markers within 20 cM of each other indicate the same putative QTL, we found two QTLs each on chromosome 1, 2, and 4 (**Error! Reference source not found.**). Furthermore, six unmapped markers were significantly associated with BYD symptoms (Table 5).

Table 9: Means for BYD symptom level for HVM054 alleles are shown. The different subsets are based on Tukey HSD(0.05).

HVM054 (allele migration on gel in bp)	N	Subset 1	Subset 2
150	45		2.682
154	11		2.287
158	15	0.993	
162	5		2.399
167	57		2.679

Special notice should be taken of SSR marker HVM054 on chromosome 2. With an *r* of 0.67 (*p*=9E-15) the association between this marker and BYD was extremely significant. Table 9 shows that HVM054 had five different alleles in the cultivars, of which one resulted in a significantly lower score for BYD. The mean BYD value for SSR allele size 158 bp was 0.993, indicating that cultivars carrying this SSR allele were much more resistant or tolerant to BYD than the cultivars carrying the other four alleles (mean symptom level at least 2.287). An analysis of the ancestries of all fifteen cultivars did not reveal any common ancestor which could have donated the resistance gene.

Three of the six QTLs for BYD resistance indicated by LD mapping have been reported earlier (Table 5, Figure 3). QTLs on chromosome 1 and 2 coincided with QTLs reported by Toojinda *et al*. (2000) and Scheurer *et al*. (2001). The other three QTLs have not been reported before, including the QTL that was highly associated with one allele of HVM054 (Table 9). The other QTLs of Scheurer *et al*. and Toojinda *et al*. were not confirmed in LD mapping, although high correlation was found for markers on chromosome 3 at 154 cM (*r*=0.24, *p*=0.0057*)*, on chromosome 4 at 86 cM (*r*=-0.25, *p*=0.0036), and on chromosome 5 at 98 cM (*r*=-0.24, *p*=0.0039). On all these locations Toojinda *et al*. found QTLs for BYD.

DTH and PH

For the trait DTH one QTL was found on chromosome 7 at 32 cM, and one unmapped marker was significantly associated. For PH, associated markers were found on chromosome 1 and 7. Besides, eight unmapped markers were significantly associated with PH. In total three QTLs for PH were indicated with LD mapping. On the positions of all these three indicated QTLs genes for PH had been reported before (Table 5, Figure 3).

Rachilla and lodicula

The known loci for rachilla hair length (*srh*) and lodicule size were not confirmed in our LD mapping (data not shown). No associations were found between rachilla hair length and markers in the neighborhood of *srh* (short rachilla hair) on chromosome 7 at 90 cM. One marker (BMAG323) on chromosome 7 at 42.5 cM did show significant correlation with rachilla hair length. A known gene for lodicule size is on chromosome 2 around 150 cM (Turuspekov *et al*., 2004), but no associated markers were found in that region. However, lodicule size associated markers were found on chromosome 4 around 50 cM. This area, as well as the area around 150 cM on chromosome 2, is known for several loci for Fusarium Head Blight resistance, a trait that tends to be associated with closed flowering (Zhu *et al*., 1999; De la Pena *et al*., 1999).

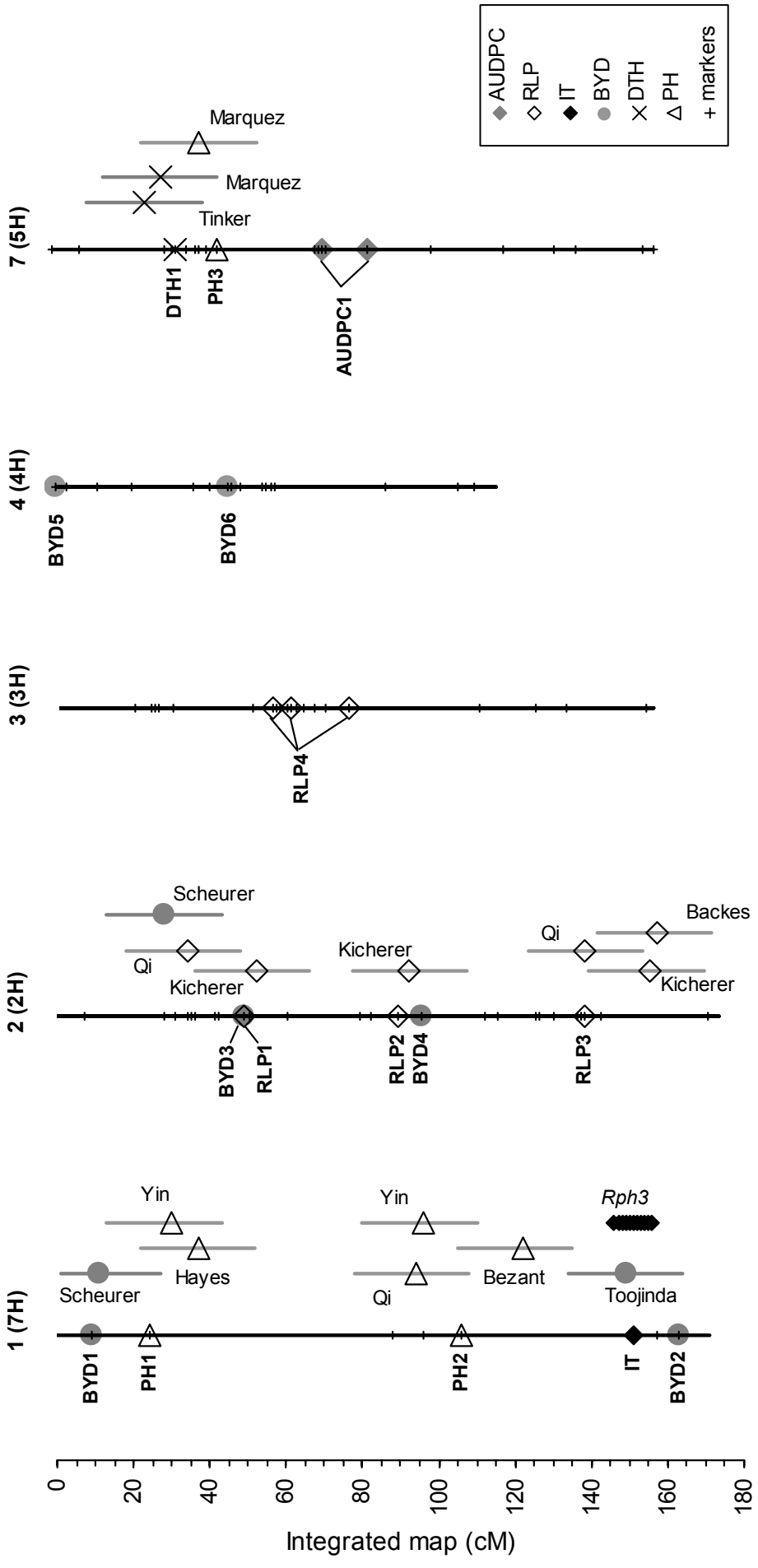


Figure 4: Overview of associated markers (left of chromosome) and QTLs reported in literature (right of chromosome) for five barley chromosomes. The traits were leaf rust (*P. hordei*) resistance (IT=Infection Type, AUDPC=area under the disease progress curve for adult plants in the field, RLP=relative latency period of seedlings), BYD (Barley Yellow Dwarf), DTH (Days To Heading), and PH (Plant Height). The map distances are according to the integrated map (IM) as described by Kraakman *et al.* (2004). The positions of QTLs from literature have been recalculated to positions on the integrated map following: position IM = position literature x (chrom. length literature/chrom. length IM). All QTLs from literature have been drawn with 30 cM intervals. For the QTLs found with LD mapping only the associated markers were indicated, without intervals. Markers within 20 cM of each other were considered to be linking to the same QTL. The authors to which is referred can be found in Table 5.

Conclusions & Discussion

Linkage disequilibrium existed in this set of modern spring barley cultivars. This LD could be used to map a variety of traits with a set of 236 AFLP marker loci and 12 SSR marker loci. We found major genes as well as minor genes, for morphological, resistance, and for other agronomically relevant traits. Many of the trait - associated markers were located in a region where other authors had reported QTLs for the same trait after linkage studies with segregating populations. For leaf rust resistance and BYD the LD mapping suggested novel QTLs.

Rust resistance

A major gene for leaf rust resistance, *Rph3* on chromosome 1, was very significantly associated with an AFLP-marker in this set of cultivars. It was remarkable that the absence of a band for AFLP marker E39M61-255 was linked to resistance. After all, for AFLP markers in a set of cultivars the common presence of a band is highly informative, whereas the common absence of a band is less informative. Intrinsic to the AFLP technique is that the probability of a change from 0 to 1 for a specific band (primer combination plus specific mobility) is much smaller than a reversion from 1 to 0, so that many reversions are easier to envisage than even one extra forward change. According to Dollo's Law, it is harder to gain a complex feature than to lose it (Farris, 1977). In this case, 27 out of 36 cultivars received a chromosome segment without the AFLP-band from a genetic source with the *Rph3* gene. Apparently, the marker band occurred in European spring barley germplasm at a frequency of nearly 1 before the introduction of *Rph3*, but the donor possessed, linked to *Rph3*, the rare 0 allele of the marker. The occurrence of nine cultivars (Table 6), which lacked the band but did have high IT, could be due to a number of reasons. For instance, these accessions may have obtained a different 0 allele of the marker from another, non *Rph3* carrying source. Also, linkage could have been broken between this marker and *Rph3*.

Four cultivars, Hanka, Abed5193, Enigma, and Tofta, had the marker of E39M61-255 that is associated with the susceptibility allele *rph3*, but they showed hypersensitivity (IT 3 or 4 on the 0–9 scale). Their resistance is probably due to effective *Rph* genes elsewhere on the genome. Remarkable in this group was Hanka, probably carrier of *Rph7* (Niks *et al*., 2000). In seedlings the resistance was effective (IT=3), but in the field Hanka showed a surprisingly high AUDPC (215). In previous trials Hanka had high resistance also in the field, and virulence to *Rph7* still

has not been reported in the European *P. hordei* population (Niks *et al*., 2000). We have no explanation for this unexpected result on Hanka.

QTLs for rust resistance have been reported by many authors, *e.g.* Qi *et al*. (1998b; 2000), Kicherer *et al.* (2000), and Spaner *et al.* (1998). Qi *et al.* (1998b; 2000) consistently found QTLs for RLP: *Rphq2*, *3*, and *4*. The LP prolonging alleles were contributed by Vada, a cultivar that features in the ancestry of 70 of the 148 cultivars. LD mapping in the cultivars showed that *Rphq2* on the distal part of chromosome 2 was also present and effective in this germplasm. However, in the regions of *Rphq3* and *Rphq4*, on chromosomes 6 and 7, no markers were found to be associated with RLP or AUDPC. First, *Rphq2* will be discussed, and then *Rphq3* and *4* will be elaborated on.

Rphq2:

Rphq2 was found in our set of cultivars with significant association between RLP and three AFLP markers on chromosome 2. The same markers were also found by Qi *et al.* (1998b) to be peak markers for *Rphq2*. In contrast to Qi *et al.*, these markers were not associated with AUDPC. One marker within the *Rphq2* region (Qi *et al.*, 1998b) was not associated with RLP: E38M54-294 (Figure 3). The allele giving an amplification product was from L94 (L94-marker), and would be expected to give a strong negative association with RLP. In the L94 × Vada RIL population of Qi *et al.* (1998b), E38M54-294 was associated with RLP. In the RIL population there is no doubt that the band on gel is from L94, but in the cultivar set the band could also originate from another part of the genome and co-migrate by identical mobility (Waugh *et al.*, 1997; Koopman and Gort, 2004). However, the association between E38M54-294 and the Vada markers was just as high as the mutual correlation between the Vada markers in the *Rphq2* region, indicating that no other band has been interfering. The lack of association in our germplasm might be due to the fact that breeding germplasm lines that carry the L94-allele, without carrying an effective *Rph* gene, would tend to have an unacceptably low level of partial resistance. The breeder would reject them, unless the absence of *Rphq2* is compensated by the presence of (an)other QTL(s) elsewhere on the genome. This would lead to the interesting hypothesis that, for characters for which many loci can contribute, association between marker alleles and positive, agronomically desirable alleles, is more likely to be found than association between marker alleles and agronomically undesirable alleles.

Rphq3 and 4:

In this set of cultivars we did not find association between markers and RLP or AUDPC that could indicate variation for *Rphq3* on chromosome 6 and *Rphq4* on chromosome 7. Among the

possible explanations for this lack of association between *Rphq3*- and *Rphq4*-linked markers and RLP and AUDPC are the following: (1) Maybe the indicative markers are not sufficiently close to the loci for partial resistance to have resulted in LD, (2) maybe loci for other agronomically relevant characters are linked to the locus for partial resistance, but with the favorable alleles in repulsion phase, stimulating recombination and/or selection against the allele for partial resistance, or (3) the partial resistance-enhancing effect of *Rphq3* and *Rphq4* may depend on interaction with other genes in the genotype, making their effect genotype dependent. Support for the last explanation can be found in NILs that have been developed from the L94 × Vada population (Berloo *et al.*, 2001). It appeared that *Rphq4* was only moderately effective in a L94 background (pers.comm. Niks and Marcel).

AUDPC was only associated with markers on chromosome 7 around 70 cM. Those markers were also present and segregating in the L94 × Vada population, but did not show association with AUDPC there. This is unexpected, and further research is required to explain this apparent contradiction. It was the only association with AUDPC found in the present set of cultivars.

BYD

BYD was associated with many mapped and unmapped markers, indicating putative QTLs on chromosome 1, 2, and 4. The indicated chromosome regions seem to coincide with one QTL that had been reported by Toojinda *et al.* (2000) and two QTLs reported by Scheurer *et al.* (2001). Remarkably, the commonly used major gene *Ryd2* for BYD resistance was not found with LD mapping, although it cannot be ruled out that one or more of the unmapped markers is associated with *Ryd2*. Another possibility is that in the set of cultivars *Ryd2* is not present. Testing the material on the *Ryd2* linked markers YLM (Jefferies *et al.,* 2003) and YLP (Ford *et al*., 1998) would indicate which and how many cultivars are likely to possess *Ryd2*.

New putative QTLs for BYD have been found on chromosomes 1, 2, and 4. Of special interest is SSR marker HVM054 on chromosome 2. The association between this marker and BYD was extremely significant. One SSR allele is associated with a very low symptom score compared to the other four SSR alleles. To the best of our knowledge, no gene for BYD resistance or tolerance has been reported in this region. Fifteen cultivars carried the marker allele associated with the resistance. They had no obvious common ancestry.

DTH and PH

DTH was significantly associated with only two markers. This may seem surprising, as many QTLs for heading date have been reported. In a mapping population of Blenheim × Kym, Bezant *et al.* (1996) found QTLs for DTH on each chromosome and on some chromosomes even more than one. This showed that QTLs for DTH are abundantly present even in one biparental mapping population, so we may assume that in a set of cultivars with many different ancestors an overwhelming number of QTLs affecting DTH may be present and effective. It will then be hard to detect LD between a marker and one of the many QTLs. Furthermore, the range of DTH values for all cultivars was rather small (Table 2), suggesting high adaptedness for earliness of all genotypes. Lack of genetic variation further may have hampered the detection of association between markers and DTH.

PH was associated with three mapped markers and with eight unmapped markers. All mapped associated markers were in a region were QTLs for PH have been reported before. Similar to DTH, many QTLs for PH have been reported. In contrast to DTH, for PH we did find four QTLs with LD mapping. This suggests that for PH fewer QTLs are present in the cultivar germplasm than for DTH, so LD between a marker and a QTL was not obscured by the possible compensation at other QTLs.

Combined association for IT and RLP: chromosome 2 at 138 cM (*Rphq2*)

Cultivars carrying *Rph3* were less likely to carry *Rphq2* for partial resistance to *P. hordei*. We presume that if in breeding programmes advanced germplasm carried an effective gene for hypersensitivity (low IT), breeders relatively frequently did not select *Rphq2*, since they could not judge for level of partial resistance. This may be regarded as evidence supporting the theory of the "Vertifolia effect". This effect has been defined by VanderPlank (1963) as the loss of horizontal resistance in the process of breeding for vertical resistance. Parlevliet (1981) argued that, however logical the "Vertifolia effect" appears, it is not a general phenomenon. As example of evidence against this effect he mentioned the barley cultivar 'Cebada Capa' which combines a widely effective gene for hypersensitive complete resistance to *P. hordei* (*Rph7*) but has a high level of partial resistance in its genetic background (Parlevliet and Kuiper, 1977). This evidence is based on an example of a possibly casual cultivar, whereas the evidence appearing from our study is based on frequencies in a wide set of cultivars.

The value of LD studies

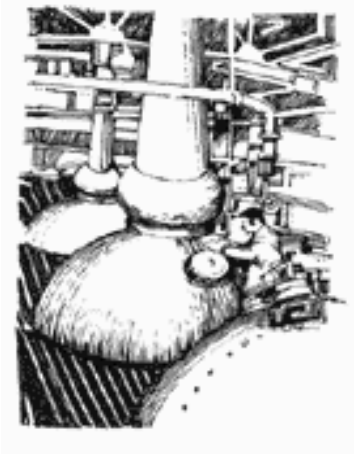
A number of the associations between markers and trait values were for markers located in a region where already QTLs for the trait considered had been reported. This indicates that QTLs detected in mapping populations from biparental mapping populations were widely represented in this set of cultivars, and that they could be detected with LD mapping. Associations between traits and markers in regions that had not been implicated before to affect the trait suggest new QTLs. Such new QTLs were found for AUDPC, RLP, and BYD. Probably, those traits have never been studied in a mapping population where both parents differed for the QTLs involved, or there were no markers in that region. The newly suggested QTLs in the present study should be validated in a study with an appropriate mapping population. Such a population should segregate for the contrasting alleles of the associated marker. Our article indicates that LD studies are efficient in indicating novel genes for important agronomic characters that subsequently can be validated in specific biparental crossing populations, and in confirming QTLs that have been detected in biparental mapping populations.

Acknowledgements

Lars Kjær of the Danish Agricultural Advisory Center is gratefully acknowledged for providing seed of recently tested cultivars. More than thirty breeding companies from all over Europe have sent us seed of their cultivar(s), which we highly appreciate. We thank Pim Lindhout, Fien Meyer-Dekens, and Petra van den Berg for performing AFLP analyses, Thierry Marcel for SSR analyses, and Fabien Debellis for SSR analyses and help with the observations in the field. This research was supported by the Technology Foundation STW (project WBI 4785), applied science division of NWO and the technology programme of the Ministry of Economic Affairs of The Netherlands.

Chapter 

QTLs for yield, yield stability and yield adaptability in doubled haploid populations of barley



Chapter 4

QTLs for yield, yield stability and yield adaptability in doubled haploid populations of barley

Kraakman, A.T.W., M. Malosetti, F. Meijer-Dekens, R.G. F. Visser and F. A. van Eeuwijk

In preparation for submission

Abstract

Yield adaptability and yield stability are important traits, but difficult to select for by conventional phenotypic selection. To investigate the possibilities for marker assisted selection on these complex traits, we studied the genetic basis for commonly used adaptability and stability measures by QTL analysis. The adaptability measures were the genotypic slopes of the regressions of individual performances on mean yield and the genotypic scores on AMMI principal components. The stability measures were Shukla's stability variance and the Eberhart-Russell stability variance. These parameters were estimated for four populations of doubled haploids stemming from crosses between inbred lines. Each population was separately evaluated in multi-environment trials. The first population was produced from a cross between Henni and Meltan, two modern cultivars contrasting for adaptability and stability. The other three populations consisted of publicly available data from the North American Barley Genome Project. No QTLs were found for Shukla's stability variance and only one for Eberhart-Russell stability variance, making selection for them difficult to impossible. Many QTLs were identified for the adaptability measures based on regression and AMMI, suggesting that marker assisted selection for these traits should be possible. The results of the QTL analyses for the doubled haploid populations were compared to the results of an association analysis with a collection of European cultivars. No clear cut correspondences were found between both approaches.

Introduction

The primary objective of plant breeders is to produce genotypes with high and consistent performance across environments. The selection process is hindered by the interaction between genotype (G) and environment (E), GE-interaction, especially when crossover interaction occurs, as the rank order of genotypes changes over environments.

Assessing the phenotypic stability and adaptability of genotypes requires trials in many environments. Indirect selection for these traits, for instance using molecular markers, would greatly improve the breeding process. Needed for that are a meaningful measures for adaptability and stability and markers linked to loci determining those traits. Both issues are addressed in this article. We describe our use of a segregating population of spring barley, created with two genotypes contrasting for stability and adaptability in an earlier study. We estimated commonly used stability and adaptability measures, and we tried to identify and map QTLs for these traits.

In this introduction, we first define stability and adaptability, then we give a brief overview of modeling approaches for GE-interaction analysis, and finally a number of stability/adaptability measures is described. We conclude the introduction with a summary of the research objectives.

Definitions

The terms adaptation, adaptability, and stability are often used in quite different senses (Becker and Leon, 1988; Lin *et al.* 1986). *Adaptability* is often defined as the ability to show good adaptedness to a range of environments. *Stability* is the consistency of phenotypic expression around some reference level of expression. So, a genotype with wide range adaptability performs stable over a wide range of environments, while a genotype with specific adaptability performs stable in a small range of environments.

Two different concepts of stability exist, referred to as the static concept and the dynamic concept (Leon, 1985). With regard to the static concept, a stable genotype performs consistently regardless of the variation of the environmental conditions – its variance across environments is small. Compared with other genotypes, a stable genotype in this conception

performs good in bad environments and bad in good environments. Therefore, this concept is less interesting to breeders, especially because high stability is associated with low average performance (Annicchiarico, 2002). In the dynamic concept, stability is defined as deviation from the predicted performance for a genotype in a series of environments. A stable genotype shows little deviation from this predicted response. It is not required that the predicted genotypic response to environmental conditions should be equal for all genotypes. This way, stability is not only relative to the environments in which it was assessed, but also relative to the other genotypes that were tested. Becker (1981) referred to the static and dynamic concepts of stability as the biological and agronomic concepts, respectively. In this article we will limit the scope to the dynamic concept, as this is the most interesting concept for breeders. Central in the dynamic concept are the models used to predict the genotypic response to environments.

Models for GE analysis and measures for stability and adaptability

Analysis of GE-interaction, and thus the assessment of stability and adaptability, is based on biometrical models. Parametric and nonparametric approaches have been proposed. The nonparametric approaches are mostly based on rank orders of genotypes. They do not make assumptions on the distribution of observed values or of variance homogeneity. We limited our research to parametric approaches, as they receive more attention in literature and appear to be more popular by breeders. For more information on the non-parametric approaches the reader is referred to Annicchiarico (2002).

We present three relatively simple models for analyzing GE-interaction. Each model forms the basis for one or more stab/adap measures. The difference between the models is the way they model GE-interaction. All equations for the models are given in Table 1 in a simplified notation.

The first model, the stability variance model (Shukla, 1972), is based on a two-way analysis of variance (ANOVA) model including genotypic and environmental main effect, and an interaction effect that is assumed to be random with variance depending on the genotype.

Table 1: Overview of stability and adaptability measures used in this research, and the models on which they are based. The part of the equation that differs between models is printed in bold type. The error terms have been left out for simplicity.

Measure			Based on model	
	Symbol	Description	No.	
Stab	σ^2_{SH}	Stability variance	1	$y_{ij} = \mu + g_i + \boldsymbol{e_j} + \boldsymbol{s_{ij}}$
Adap	β_i	Slope of regression	2	$y_{ij} = \mu + g_i + \boldsymbol{\beta_i e_j + d_{ij}}$
Stab	σ^2_{ER}	Deviation from regression	2	$y_{ij} = \mu + g_i + \boldsymbol{\beta_i e_j + d_{ij}}$
Adap	*pc1, pc2, pc3*	AMMI principal components	3	$y_{ij} = \mu + g_i + \boldsymbol{e_j + \sum_{l=1}^{L} u_{li} v_{lj} + \delta_{ij}}$

where y_{ij} is the mean phenotype of genotype or cultivar i in environment j,

μ is the grand mean across the whole experiment,

g_i is the main effect for genotype i,

e_j is the j_{th} environmental main effect, or the environmental index,

s_{ij} is the random genotype-environment interaction effect,

β_i is the linear regression of y_{ij} on e_j,

d_{ij} is the deviation from the linear regression,

u_{li} is the score for genotype i, corresponding to multiplicative term l,

v_{lj} is the environmental score in term l of environment j,

δ_{ij} is the deviation from the multiplicative model for interaction.

The second model is based on the regression approach, suggested already by Yates and Cochran (1938) and later adapted by many others (e.g. Finlay and Wilkinson, 1963; Eberhart and Russell, 1966). The GE-interaction is now expressed as $\beta_i e_j + d_{ij}$, where β_i is the linear regression coefficient for genotype i in a regression on the environmental mean e, and d_{ij} is the deviation from this regression. The regression is mostly done on the environmental mean of all genotypes, or on the mean of check cultivars. In this model it is assumed that there is a linear relationship between the genotypic response to the environment and the environmental mean. Knight (1970) suggested that the usefulness of this approach decreases when reduced performance is due to many different environmental factors as those observations are placed side by side in the regression. Furthermore, the relationship between genotypic response and

the environment might not be linear, but for instance threshold regulated. In spite of these drawbacks, the regression approach has been the most popular approach so far, mainly due to its simplicity. After all, no environmental conditions have to be monitored, and the calculations are very straightforward.

The third and last model belongs to the so-called AMMI-models (Gollob, 1968; Mandel, 1971). In this model, the main effects (G and E) are treated as additive, while the GE-interaction variance is modeled with multiplicative terms (Additive Main, Multiplicative Interaction). The GE-interaction is now expressed as:

$$\sum_{l=1}^{L} u_{li} v_{lj} + \delta_{ij} ,$$

where L is the number of multiplicative terms, and u_{li} and v_{lj} are scores for genotype i and environment j, respectively. The environmental scores could be understood as hypothetical environmental variables to which the genotypes respond. The scores for genotypes and environments can be obtained by a singular value decomposition of the matrix of residuals from additivity. If a clear linear response exists of genotypes to the mean performance of all genotypes in the environments, as assumed in the regression approach, one of the $u_{.i}$ will represent the regression coefficient β_i, the genotypical sensitivity to the environment, and the corresponding $v_{.j}$ expresses e_j.

In Table 1, the parts of the model that describe genotypic responses to environmental descriptors represent *adaptability*, while the remaining unexplained part of GE-interaction represents *stability*. Note that under this definition, in theory we could model the performance of genotypes in environments perfectly and only adaptability parameters would remain, while the stability of all genotypes would become zero (fully stable).

Stability and adaptability measures

Expressing a genotype's stability and adaptability in one statistic is attractive both for breeders and farmers. Many stability measures have been proposed (see reviews of Becker and Leon, 1988; Lin *et al*., 1986; Lin and Binns, 1994; Annicchiarico, 2002; Kang, 2002). Measures based on the regression approach have been adopted most widely, although actual breeding for stability based on those measures is not very common.

In this research we selected four stability/adaptability (stab/adap) measures based on the models described in the previous paragraph: σ^2_{SH}, the stability variance (=$var(s_{i.})$), β_i, the slope

of regression on the mean, σ^2_{ER}, deviation from regression *(=var(d_i)*), and *pc1, pc2,* and *pc3*, the AMMI principal component scores (= u_{1i}, u_{2i}, and u_{3i})(see Table 1). Note that σ^2_{ER} and σ^2_{SH} could be regarded a parameter for the goodness of fit of the model, and they are therefore sometimes disqualified as stability measure. Nevertheless, we decided to include σ^2_{ER} and σ^2_{SH}, especially because they are often considered in literature.

Stability and adaptability as defined in these measures have a different meaning from the definitions in the first part of this introduction. The difference is due to the difference in view between a breeder (and farmer) and a statistician. Where the breeder wants adaptability in the sense that the genotype should perform as good as feasible in target environments, the statistician defines adaptability as the sensitivity to the environment. For the breeder, stability is consistency in performance, while for the statistician stability is the part of GE-interaction variance that is not explained by the model.

Objectives of this research

Marker assisted selection for adaptability and stability using molecular markers would greatly improve the selection process.

The modern spring barley cultivars Henni and Meltan differed in mean yield, yield stability and yield adaptability, according to the results from the Danish variety testing in the years 1993-1999 (Kraakman *et al*., 2004). A DH-population from Henni and Meltan was used to assess yield, yield stability and yield adaptability in six contrasting environments. The same traits were evaluated in publicly available datasets of three DH-populations. For all datasets, QTLs were mapped for yield and stab/adap measures.

The objectives of this research were:

1. Estimate four stab/adap measures for the Henni × Meltan DHs,
2. Map QTLs for yield and stab/adap measures in the Henni × Meltan DH-population,
3. Map QTLs for yield and stab/adap measures in three other publicly available segregating populations, and compare the results with the Henni-Meltan population,
4. Compare the found QTLs by linkage analysis with the QTLs for yield and yield stab/adap measures found with linkage disequilibrium mapping (Kraakman *et al.*, 2004).

Material and methods

Plant material

One hundred and sixty microspore-derived doubled haploid (DH) lines were produced from the F1 of Henni × Meltan by anther culture (Davies and Morton, 1998). All DH-lines were used for construction of a genetic linkage map, and a subset of 118 DH-lines was used for yield trials. Both Henni and Meltan are modern European two-rowed spring barley varieties, mainly used for fodder. Henni (1995) is a cross between Baronesse and 84160.1.3.3 made by Nordsaat. Meltan (1991) is a cross between D80-20 and Tellus MMMDDN made by Svalof Weibull. Based on evaluation data from the Danish variety trials between 1993 and 2000, Kraakman *et al*. (2004) characterized Henni as a stable genotype with high adaptability, and Meltan as an unstable genotype with low adaptability. Adaptability and stability were expressed as β_i (slope of regression) and σ^2_{ER} (deviation from regression), respectively.

Genotyping and construction of genetic linkage map

The following 27 AFLP primer combinations were employed (written as *Eco*RI(*Mse*I)): 32(55,61), 33(55,58,61), 35(*48*,*54*,*55*,*61*), 37(32,*33*,38), 38(*50*,*54*,*55*), 39(55,*61*), 40(32,38), 41(32), 42(*32*,40,*48*,51), and 45(*49*,*55*,58). An explanation of the primer combinations can be found in Qi and Lindhout (1997). Twenty two of them (underlined) have been used by Qi *et al*. (1998) to construct a high density map of L94 × Vada, and thirteen (italics) have been used in the linkage disequilibrium study in barley of Kraakman *et al*. (2004). Markers were identified following the AFLP profiles of Qi and Lindhout (1997; also available on http://wheat.pw.usda.gov/ggpages/Qi).

JoinMap 3.0 (Stam and Van Ooijen, 1996) was used for linkage grouping and map construction. Linkage groups were assigned to the corresponding chromosomes by using common markers on the maps of L94 × Vada (Qi and Lindhout, 1997), Apex × Prisma (Yin *et al*., 1999), GEI119 × Gunhild (Koorevaar, 1997), and L94 × 116-5 (Qi *et al*., 2000).

Phenotypic data collection

Yield performance was scored in six contrasting environments (Table 2). Yield was measured as the weight of grain harvested per plot and converted to kilograms per hectare. The individual trials differed in the number of replicates (1, 1.25 and 2) and row and column balance in

relation to the set of genotypes. For the purpose of this article, trials were individually analyzed imposing an appropriate model for the specific design features.

Table 2: Environments in which Henni, Meltan, and their doubled haploids were tested.

Code	Country	City	Latitude/longitude		Year	Factor
NL*sa*	Netherlands	Achterberg	51º97'	5º61'	2002	Sandy soil
NL*cl*	Netherlands	Wageningen	51º95'	5º64'	2002	Clay soil
SP*ni*	Spain	Forada	41º52'	1º1'	2004	Not irrigated
SP*ir*	Spain	Gimenells	41º35'	0º32'	2004	Irrigated
UK*ni*	UK Scotland	Dundee	56º47'	-3º5'	2004	Not irrigated
UK*ir*	UK Scotland	Dundee	56º47'	-3º5'	2004	Irrigated

Other multi-environment datasets of DH populations

Multi-environment yield data of three other DH-populations were used to compute the stab/adap measures. These datasets were produced by the North American Barley Genome Project (NABGP). The crossing parents of the populations were Steptoe × Morex (SM), Harrington × Morex (HM), and Harrington × TR306 (HT) (Table 3). Steptoe and Morex are six-rowed varieties, Harrington and TR306 are two-rowed varieties. More information on SM and HT can be found at http://wheat.pw.usda.gov/GG2, and for HM at http://wheat.pw.usda.gov/ggpages/HxM/, or in the references mentioned in Table 3. We used the basemaps for QTL mapping.

Table 3: Overview of four DH-populations of which the data were used in this research.

Code	Cross	#DHs	#Loci	#Environments	First published
HeMe	Henni × Meltan	118	274	6	This thesis
SM	Steptoe × Morex	150	223	16	Hayes *et al*., 1993
HM	Harrington × Morex	140	107	9	Hayes *et al*., 1997
HT	Harrington × TR306	145	127	28	Tinker *et al*., 1996

Statistical analyses

For each trial genotypic predictions for the 118 DH lines from the HeMe population across the six environments were calculated in a two-way genotype by environment table of means. Various two-way models, differing in the model for the GE-interaction, were fitted to this table of means. We first fitted a two-way ANOVA model of the form: $y_{ij} = \mu + g_i + e_j + s_{ij}$ with $s_{ij} \sim N(0,\sigma^2_{SH})$. The genotypic-specific residual variance can be interpreted as a stability parameter (Shukla, 1972). Secondly we fitted the model: $y_{ij} = \mu + g_i + \beta_i e_j + d_{ij}$ with $d_{ij} \sim N(0,\sigma^2_{ER})$, where the phenotypic response (y_{ij}) is regressed on an environmental index (e_j) describing the quality of the environment (calculated as the average of all the genotypes in this particular environment). The slope, β_i, is genotype-specific and interpreted as an adaptability parameter (Finlay and Wilkinson, 1963). Analogous to the situation in Shukla model, the assumption of homogeneity of variance for the residual is relaxed to allow for genotype-specific deviations from the model (σ^2_{ER}). This genotype-specific parameter was interpreted as a stability parameter by Eberhart and Russell (1966). Finally, we fitted an AMMI model of the form: $y_{ij} = \mu + g_i + e_j + \Sigma u_{Li} v_{Lj} + d_{ij}$ with $d_{ij} \sim N(0,\sigma^2)$ and L the number of multiplicative terms (Gollob, 1968; Mandel, 1971; Gauch, 1992).

Due to difficulties in computing Shukla's σ^2_{SH} in HT, we used only 23 out of 28 environments for estimating this measure in this population. Prior to QTL analyses the natural logarithm (ln) was taken of σ^2_{SH} and σ^2_{ER}.

QTL analysis

QTL analyses were performed on the estimates for the adaptability and stability measures using restricted MQM mapping (MapQTL 5.0; Van Ooijen, 2004). After initial interval mapping, the markers with the highest LOD values ('peak markers') were taken as co-factors for multiple-QTL mapping using the restricted MQM-method (Jansen, 1993). When new LOD peaks appeared, new peak markers were added to the co-factor set until a stable LOD profile was reached. The genomewide $p<0.01$ significance threshold was determined according to Van Ooijen (1999).

Results

Linkage map:

In total 294 polymorphic markers were scored, giving an average of 10.8 markers per primer combination, ranging from 5 (E35M55, E38M54, E39M55) to 19 (E42M40, E42M51). After quality checks and removal of allelic loci 274 markers remained for map making. The final map was 1056 cM long (Kosambi function) and had chromosomes between 104 cM (chr 4) and 178 cM (chr 7). There was sufficient coverage of the genome, although four chromosomes had two gaps larger than 25 cM (chr 1, 3, 4, and 5). The largest gap was 50 cM on chromosome 5. On chromosome 4 significant distorted segregation of 22 loci occurred between 40 and 103 cM. This was shown before in another DH-population by Sayed *et al*. (2002). The locus order and linkage map were in agreement with earlier published maps of L94 × Vada, Apex × Prisma, GEI119 × Gunhild, and L94 × 116-5.

A basemap was constructed for QTL mapping (Figure 1). For this basemap, each marker was removed that had a pairwise recombination percentage with another marker of less than 5%. The marker with most missing values was removed. The 115 loci on the final basemap were in the same order as on the full map. The total length of the basemap was 1099 cM, approximately the same as the full map.

Yield data

The mean yield per hectare for the HeMe DH-population varied from 3269 kg (SP*ni*) to 6644 kg (SP*ir*) (Table 4). In four of the six environments both parents yielded more than the mean of the DHs. This could be genetically regulated, but also the difference in seed quality between the parents and the DHs might have played a role. In two of the six environments Henni performed significantly better than Meltan, while in the other four environments the differences were not significant.

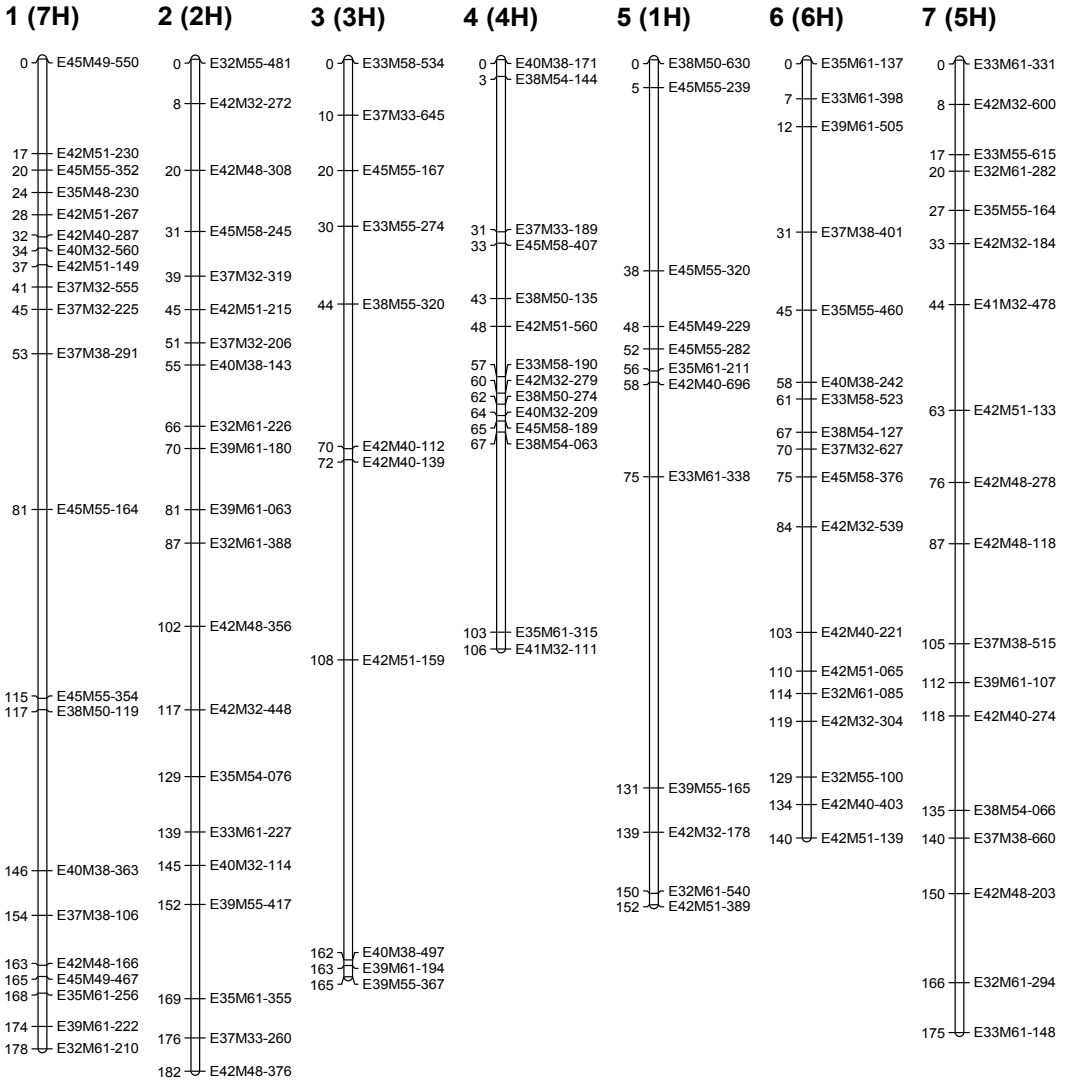


Figure 1: Linkage map (basemap) with 115 AFLP markers from the Henni × Meltan barley cross showing approximate cM positions (Kosambi function).

Table 4: Yield (kg ha-1) of Henni, Meltan, and their 118 doubled haploid (DH) progeny in six environments. Parental values in bold were significantly higher than the other parent ($p<0.01$).

	Parents		DH lines	
Environment	Henni	Meltan	Mean	St.dev
Nl*sa*	**5341**	4413	4283	415
NL*cl*	**4841**	3504	3749	319
Sp*ni*	3141	3325	3269	419
Sp*ir*	7001	6823	6644	288
Uk*ni*	5225	5403	5179	433
Uk*ir*	5631	5318	5241	332
Average	5197	4798	4728	368

Stability/adaptability measures in all four DH-populations.

The total amount of GE-interaction in the HeMe population was rather low: the mean Shukla's stability variance, σ^2_{SH}, was much lower for HeMe (Table 5) as compared to the other three DH-populations (Table 6 to Table 8). The slope of regression, $\beta_{i,}$ ranged from 0.843 to 1.147 in HeMe, with a standard deviation of 0.062. This was a narrow range as compared to the other populations. The mean variance of regression σ^2_{ER} was for all populations around 1.1, but the minimum and maximum varied across populations. In the AMMI analyses the first *PC* in HeMe summarized already 47% of the total variance. In the other populations this percentage was 31%, 34%, and 15%.

The correlations between mean yield on the one hand and stab/adap measures on the other hand were not significant for σ^2_{SH} and σ^2_{ER}. For β_i significant correlations with mean yield of 0.60 (SM) and 0.38 (HT) were found. For all populations at least one of the AMMI genotypic scores was significantly correlated with mean yield.

The highest mutual correlations between stab/adap measures were found between σ^2_{SH} and σ^2_{ER}; they ranged from 0.66 (HT) to 0.98 (HeMe). The correlations between σ^2_{SH} and β_i and between σ^2_{ER} and β_i were low and not significant for all populations. At least one of the AMMI genotypic scores was highly correlated with β_i for each of the populations, up to -0.81 in HM (PC3), and 0.70 in HeMe (PC2).

Table 5: Descriptive statistics for mean yield and four stability/adaptability measures for the **Henni** × **Meltan** DH-population in six environments. Significant correlations in **bold** (p<0.01).

	Mean yield (kg/ha)	Shukla	Regression approach	
		σ^2_{SH}	β_i	σ^2_{ER}
Mean	4,728	0.051	1.000	1.197
Minimum	4,050	0.002	0.843	0.001
Maximum	5,447	0.314	1.147	7.626
St.dev	0,310	0.051	0.062	1.218

Correlations

	Mean yield	σ^2_{SH}	β_i	σ^2_{ER}
Mean yield				
σ^2_{SH}	-0.20			
β_i	-0.18	-0.06		
σ^2_{ER}	-0.20	**0.98**	-0.03	
PC1	0.09	0.04	**0.52**	0.03
PC2	-0.03	-0.20	**0.70**	-0.15
PC3	**-0.41**	0.08	**0.25**	0.08

AMMI Principal Components
Proportion of GE
0.47
0.29
0.13

Table 6: Descriptive statistics for yield and four stability/adaptability measures for the **Steptoe** × **Morex** DH-population in 16 environments. Significant correlations in **bold** (p<0.01).

	Mean yield (kg/ha)	Shukla	Regression approach	
		σ^2_{SH}	β_i	σ^2_{ER}
Mean	5,289	0.620	1.000	1.071
Minimum	4,124	0.102	0.587	0.186
Maximum	6,505	1.634	1.446	2.866
St.dev	0,414	0.300	0.164	0.516

Correlations

	Mean yield	σ^2_{SH}	β_i	σ^2_{ER}
Mean yield				
σ^2_{SH}	-0.15			
β_i	**0.60**	-0.04		
σ^2_{ER}	-0.18	**0.97**	-0.08	
PC1	**0.32**	-0.11	**0.33**	-0.09
PC2	**-0.43**	0.09	**-0.33**	0.05
PC3	0.18	0.12	**0.68**	0.02

AMMI Principal Components
Proportion of GE
0.31
0.20
0.18

Table 7: Descriptive statistics for yield and four stability/adaptability measures for the **Harrington** × **Morex** DH-population in 9 environments. Significant correlations in **bold** (p<0.01).

	Mean yield (kg/ha)	Shukla	Regression approach	
		σ^2_{SH}	β_i	σ^2_{ER}
Mean	4,313	0.357	1.000	1.078
Minimum	2,952	0.055	0.169	0.130
Maximum	5,350	1.712	1.868	3.390
St.dev	0,456	0.232	0.295	0.624
Correlations				
Mean yield				
σ^2_{SH}	0.05			
β_i	0.10	-0.19		
σ^2_{ER}	0.05	**0.82**	-0.13	
PC1	-0.13	-0.08	**-0.29**	-0.08
PC2	**-0.31**	**-0.32**	**0.28**	**-0.29**
PC3	-0.07	0.07	**-0.81**	0.03

AMMI Principal Components
Proportion of GE
0.34
0.17
0.12

Table 8: Descriptive statistics for yield and four stability/adaptability measures for the **Harrington** × **TR306** DH-population in 28 environments (σ^2_{SH} in 23 environments; see M&M). Significant correlations in **bold** (p<0.01).

	Mean yield (kg/ha)	Shukla	Regression approach	
		σ^2_{SH}	β_i	σ^2_{ER}
Mean	4,801	0.245	1.001	1.051
Minimum	4,125	0.088	0.764	0.381
Maximum	5,389	0.556	1.283	2.588
St.dev	0,214	0.095	0.108	0.433
Correlations				
Mean yield				
σ^2_{SH}	-0.03			
β_i	**0.38**	0.11		
σ^2_{ER}	0.04	**0.66**	0.13	
PC1	0.10	0.10	-0.01	0.17
PC2	**0.28**	-0.14	0.13	0.05
PC3	0.16	-0.07	**-0.35**	-0.18

AMMI Principal Components
Proportion of GE
0.15
0.13
0.11

QTLs for Yield in HeMe

We considered a peak in the LOD profile a QTL if the LOD threshold was exceeded and if the LOD dropped in the surrounding region. The second criterion was relaxed when the profiles of other traits made it clear that high LOD values over a large range were probably due to different QTLs.

Seven QTLs for mean yield were found, spread over chromosomes 1, 2, 4, and 7 (Figure 2, Table 9). Every mean yield QTL occurred at least in one individual environment, and at most in five environments. For none of the mean yield QTLs it was true that they corresponded to yield QTLs in all of the individual environments. In total eleven yield QTLs were found in individual trials. The two most pronounced mean yield QTLs were located on the short arm of chromosome 1. Both QTLs had a positive effect for the Henni alleles, while all other yield QTLs had a positive effect for the Meltan alleles (Table 9). The amount of explained variance was 30% and 22% for mean yield, the highest percentage of all mean yield QTLs.

QTLs for Stab/Adap measures in HeMe

No QTLs were detected for Shukla's σ^2_{SH} or for σ^2_{ER} (Figure 2, Table 9). Four QTLs for β_i were found. The first one on chromosome 1 at 61 cM was right in the region where a yield QTL was found. The difference in magnitude of this yield QTL in the individual environments could explain the presence of a QTL for β_i. The other three QTL for β_i were not exactly on the same location as yield QTLs. The QTL on chromosome 2 at 100 cM was close to a yield QTL, but the QTLs on chromosome 2 (at 10 cM) and 6 (at 65 cM) were not even close to a yield QTL. This suggested the presence of genes regulating yield adaptability without regulating yield.

QTLs for AMMI genotypic scores were identified on chromosome 1, 2, and 4. Both QTLs for *PC2* coincided with QTLs for β_i, confirming that β_i and *PC2* partly express the same aspect of GE (see also Table 5). In the same manner both QTLs for *PC3* coincided with QTLs for mean yield. The QTLs for *PC1* on chromosomes 2 (164 cM) and 4 (106 cM) did not coincide with QTLs for yield or β_i.

QTLs for Yield and Stab/Adap measures in other DH-populations

In the DH-populations SM, HM, and HT 12 QTLs for mean yield were found spread over all chromosomes (Table 9, Figure 2). No QTLs for σ^2_{SH} were found, and only one for σ^2_{ER} in HT. Four QTLs for β_i were identified, three in SM, and one in HM. Mapping AMMI genotypic scores, we identified 18 QTLs in total, spread over all chromosomes. Most QTLs for β_i and genotypic scores were found in SM, while in this population fewest yield QTLs were found.

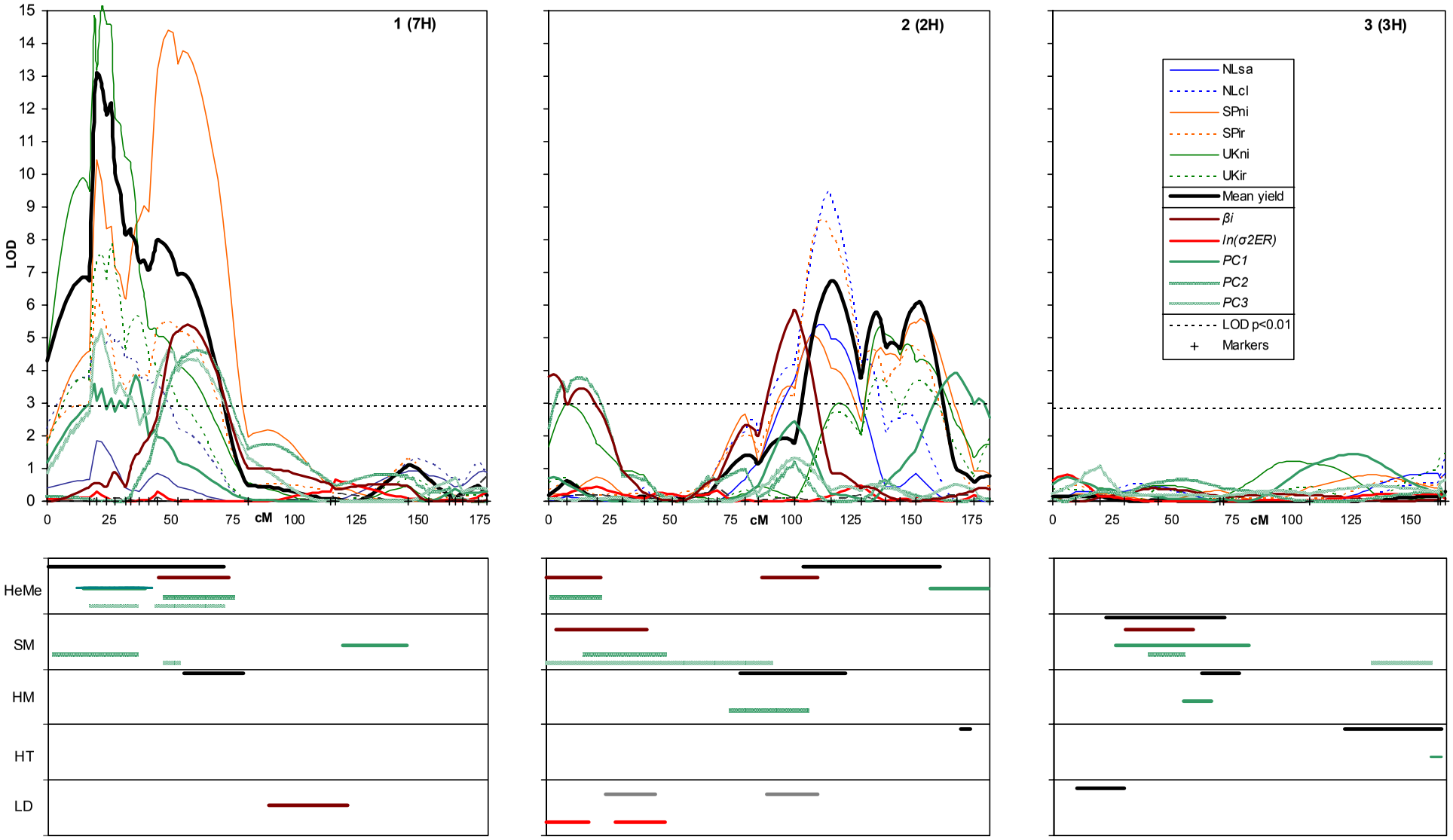


Figure 2a: see explanation on next page.

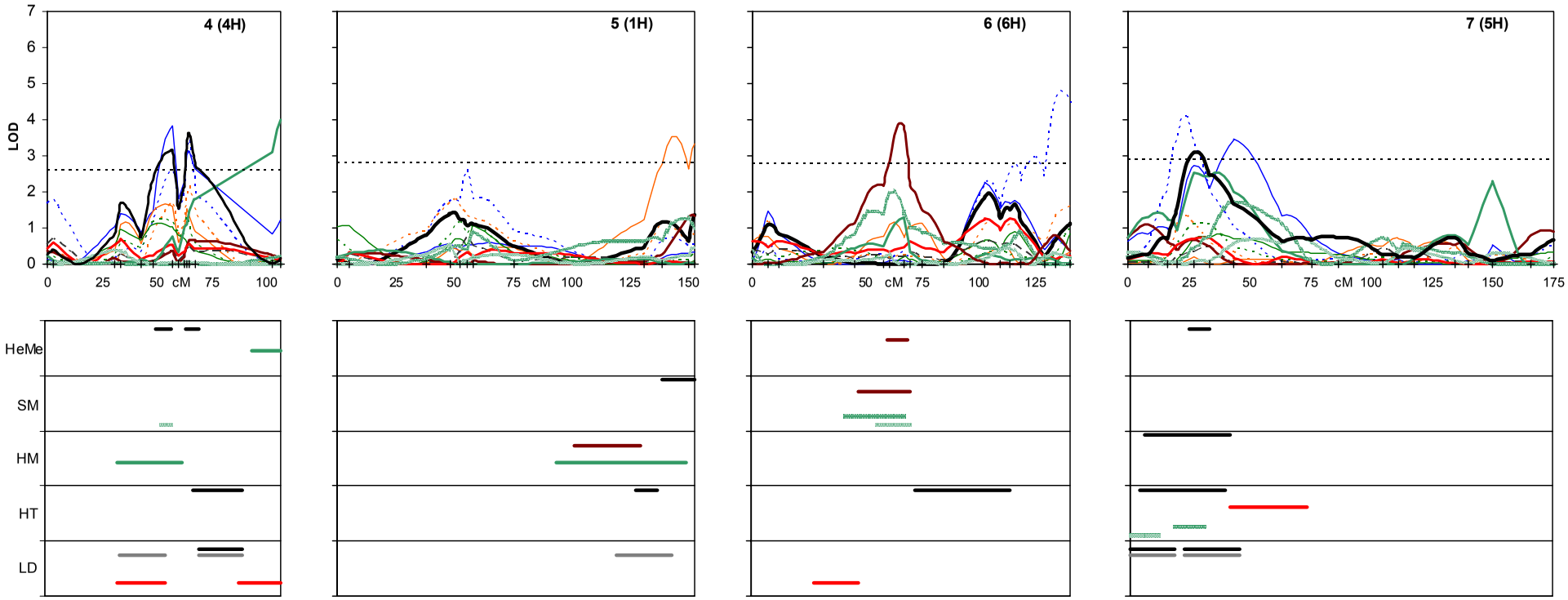


Figure 2b: Top boxes: QTL profiles for all traits of the Henni × Meltan population. The LOD scores of restricted MQM mapping are given. The traits are: yield in six individual environments (thin lines), mean yield, and stab/adap measures β_i, σ^2_{ER}, and Ammi *PCs*. The stability measure σ^2_{SH} has been left out as no QTLs were found. The p<0.01 genome wide significance threshold is given by a dashed line. The position of the marker loci is shown with dashes on the x-axis.

Bottom boxes: An overview of QTLs for mean yield and all stability and adaptability traits is given for all populations: HeMe=Henni × Meltan, SM=Steptoe × Morex, HM=Harrington × Morex, HT=Harrington × TR306, and LD=cultivars. QTL analysis results from the LD study are given in black (trials treated for leaf diseases) or grey (trials untreated for leaf diseases). The map positions have all been expressed in distances from the top of the short arm of the integrated map of Kraakman *et al*. (2004).

Table 9: Overview of QTLs for Yield and Stability/Adaptability measures in the Henni × Meltan population. The estimated effect of substitution of the Henni alleles with Meltan alleles is shown, with between brackets the percentage of explained variance. A + or – denotes the sign of the allele substitution effect for yield QTLs that were not significant. Numbers in **bold** type identify QTLs that have also been found in SM, HM, HT, or the cultivars. Traits σ^2_{ER} and σ^2_{SH} have been left out as no QTL was identified.

		Stability/adaptability measures				Yield (kg/ha)						
Chr	Pos	β_i	PC1	PC2	PC3	Mean	NLsa	NLcl	SPni	SPir	UKni	UKir
1	20		**-0.076 (9)**		0.063 (20)	-348 (30)	-	-198 (9)	-453 (27)	-228 (14)	-503 (36)	-322 (24)
	61	0.055 (19)		**0.102 (20)**	**0.068 (21)**	**-296 (22)**	-	-	**-541 (40)**	**-230 (15)**	-	-
2	10	**0.046 (14)**		**0.078 (11)**		+	+	+	-	+	+	+
	100	-0.056 (19)				+	+	+	+	+	+	+
	112					**242 (15)**	**336 (16)**	**295 (21)**	**299 (13)**	**276 (22)**	+	+
	135					254 (15)	+	+	+	+	354 (16)	262 (15)
	156					+	+	+	304 (13)	201 (11)	+	+
	164		0.092 (11)			+	+	+	+	+	+	+
4	56					154 (6)	274 (10)	152 (5)	+	+	+	+
	65					**163 (6)**	**249 (8)**	**177 (7)**	+	+	+	+
	106		-0.089 (9)			+	+	+	-	-	-	-
5	140					+	+	+	**230 (8)**	+	+	+
6	65	**0.042 (11)**				+	-	-	-	-	+	+
	136					+	+	199 (9)	+	+	+	+
7	25					**144 (5)**	+	191 (9)	+	+	+	+
	48					+	251 (9)	+	-	+	+	+
Total		4	3	2	2	7	4	5	5	4	2	2

Comparison of QTLs in all populations

In none of the DH-populations a QTL for Shukla's σ^2_{SH} was found, and only one QTL for σ^2_{ER} was identified (Table 10). Apparently, the stability measures were not regulated in such a way that QTLs could be identified. QTLs for β_i occurred in HeMe, SM and HM, and QTLs for AMMI *PC* scores occurred in all populations.

Table 10: Overview of the number of QTLs per chromosome for Yield and Stability/Adaptability measures in all populations. The traits were explained in Table 1, and the populations were introduced in Table 3.

		Regression approach		AMMI approach
	Mean yield	β_i	σ^2_{ER}	PC1/2/3
	Chromosome	Chromosome	Chromosome	Chromosome
Population	1 2 3 4 5 6 7 Tot	1 2 3 4 5 6 7 Tot	1 2 3 4 5 6 7 Tot	1 2 3 4 5 6 7 Tot
Cultivars	. 2 1 2 1 . 2 8	1 1	. 2 . 2 . 1 . 5	- - n/a - - - -
HeMe	2 2 . 2 . . 1 7	1 2 . . . 1 . 4	 0	4 2 . 1 . . . 7
SM	. . 1 . 1 . . 2	. 1 1 . . 1 . 3	 0	3 2 3 1 . 2 . 11
HM	1 1 1 . . . 1 4	. . . . 1 . . 1	 0	. 1 1 1 1 . . 4
HT	. 1 1 1 1 1 1 6	 0	. . . . . . 1 1	. . 1 . . . 2 3
Total	3 4 3 3 2 1 3 19	. 3 1 . 1 2 . 7	. . . . . . 1 1	7 5 5 3 1 2 2 25

Many QTLs found in the HeMe population were in regions where QTLs for the same trait were identified in other populations (see bold type in Table 9). The yield QTLs on chromosome 5 (~140 cM) and 7 (~25 cM) were identified in two different DH-populations and in the LD mapping. The β_i QTLs on chromosome 2 (~10 cM) and 6 (~65 cM) were identified in both HeMe and SM. The same applies to QTLs for AMMI scores on chromosome 1 (~20 and ~61 cM), and 2 (~10 cM).

Conclusions and Discussion

Yield adaptability and yield stability are important traits, but complicated for plant breeders to select for. The objective of this study was to assess the usefulness of some well known stability and adaptability measures. We used linkage analysis with four barley DH-populations, and mapped QTLs for yield and stability/adaptability measures. Our focus was on measures belonging to the dynamic concept (Leon, 1985).

Correlation between stab/adap measures

We found a positive correlation between mean yield and β_i in two populations (SM and HT), while the correlation in the other two populations was not significant. A positive correlation was also reported by (Weber and Wricke, 1990). Highly significant correlations were also found between the adaptability measures β_i and the AMMI scores, confirming the findings of Nurminiemi *et al.* (2002) in a large set of spring barley cultivars. Also the stability measures σ^2_{ER} and σ^2_{SH} were highly positively correlated. Comparably, Nurminiemi found significant correlations between Tai's λ and σ^2_{SH}. Tai's λ is a measure equivalent to σ^2_{ER} (Tai, 1971). Both σ^2_{SH} and σ^2_{ER} are stability measures, expressing unexplained GE-interaction. The high correlation between these measures for all populations suggests that both underlying models, Shukla's Stability Variance model and the Regression Approach, did not differ much with respect to the predicted responses across environments. The regression on the mean thus did not explain much of the GE.

QTLs in Henni × Meltan population

In total sixteen regions with QTLs were detected in the Henni × Meltan population, of which eleven for yield, seven for stab/adap measures, and two QTL regions coincided for both yield and stab/adap measures.

For the eleven regions with yield QTLs seven regions contained mean yield QTLs. For those seven regions, yield QTL effects in individual environments varied in magnitude, although they were of the same sign.

For stab/adap measures we found seven QTLs. All of them coincided with changing magnitudes for yield, and three coincided with yield effects with different signs. No QTLs for stability were detected, so σ^2_{SH} and σ^2_{ER} were not measures regulated by genes in such a way

that linkage with markers could be established. Adaptability QTLs were found for β_i (4) and AMMI *PC*s (7).

Especially interesting was a QTL for both yield and stab/adap on chromosome 1 at 61 cM. The explained variance for both traits was high, and the effect of substitution of Henni alleles with Meltan alleles was estimated as -296 kg ha^{-1} for yield and +0.055 for β_i. So, lower mean yield was combined with higher responsiveness to better environments. The largest yield effects for individual trials for this QTL location appeared in the Spanish environments.

The combined effect of QTLs for yield and stab/adap measures could make it difficult to select for only one of the traits. Therefore, the β_i QTLs on chromosome 2 at 10 cM and chromosome 6 at 65 cM are very interesting. Both QTLs were isolated on the chromosome, thus no yield QTL was in the neighborhood. Selection for only β_i should then be possible.

QTLs in SM, HM, HT

In the DH-populations SM, HM, and HT the number of mean yield QTLs was 2, 4, and 6, respectively. All the QTLs were reported before, for SM by Hayes *et al.* (1993) and Romagosa *et al.* (1996), for HM by Hayes *et al.* (1997) and Marquez-Cedillo *et al.* (2001), and for HT by Tinker *et al.* (1996). The number of QTLs for stab/adap measures was 14, 5, and 3, respectively. These numbers suggest a negative relationship between the number of detectable QTLs for mean yield and for stab/adap: the more stab/adap QTLs, the less mean yield QTLs. It seems coherent that more GE interaction has led to more QTLs for stab/adap measures, obscuring QTLs for mean yield and reducing the heritability of yield.

The stab/adap QTLs were mainly for β_i and AMMI *PC*s. Only one QTL for σ^2_{ER} was detected (HT), and no QTLs for σ^2_{SH}. This pattern of the presence of adaptability QTLs and the absence (or oddness) of stability QTLs was consistent for all DH-populations. Apparently, the response to the environment in the form of β_i or *PCs* can be modeled and mapped to the genome, while the remaining variability in the form of σ^2_{SH} or σ^2_{ER} cannot. The position of all the QTLs for stab/adap measures were in agreement with QTL×E effects reported earlier by the authors mentioned above. The remarkable abundance of QTL×E in the SM population, combined with only a few mean yield QTLs (Hayes *et al.*, 1993), was confirmed by our analysis.

Comparison of HeMe with SM, HM, HT, and Cultivars

Nine of the 17 QTLs from HeMe coincided with QTLs in SM, HM, HT, and the cultivars. The criterion for coincidence was the presence of an overlap between the intervals for QTL location. The matching QTLs were for yield (6), β_i (2), and *PCs* (3). The β_i QTLs were of special interest, because the two β_i QTLs in HeMe, highlighted in the previous paragraph, were also present in SM.

We looked at stab/adap measures in the context of the dynamic stability concept. All measures were therefore depending on (i) the other genotypes tested, and (ii) on the environments in which testing was carried out.

(i) The genotypes in the DH-populations differed in many aspects: European material (HeMe) versus American material (SM, HM, HT), feed varieties (Henni, Meltan, Steptoe, TR306) versus malting varieties (Morex, Harrington), two-rowed (Henni, Meltan, Harrington, TR306) versus six-rowed (Steptoe, Morex). The fact that we found overlap in QTLs between those populations supports the idea that the original genotypic source of modern barley cultivars is rather small and that only 20 founding genotypes account for 80% of the genotypic variation (Russell *et al.*, 2000). The difference in genotypes does not only create differences in the active loci, but under the dynamic stability concept it also influences the calculations of the stab/adap measures. Especially the regression approach and the AMMI approach, which are based on a description of the environment by using the performance of all genotypes, will be affected by the selection of included genotypes.

(ii) The environments in which the DH-populations were tested differed a lot. We did not investigate the environmental variables and we therefore do not know what positive or negative factors were present. We included only models without direct input of environmental variables. The environment can be represented as the environmental mean (regression approach) or as environmental score (AMMI). If GE-interaction is due to a certain environmental condition, a relationship might be established between the environmental mean or score and this environmental condition.

Selected genotypes versus unselected genotypes

The DH-populations were unselected populations representing the whole range in yield and stability that was possible by combining the parents. The set of cultivars from the LD-mapping study, however, consisted of commercial lines resulting from selection in many years and

locations. Therefore, only genotypes were present that had high yield and proved to be stable over many environments. Differences in detected QTLs between the LD mapping and the linkage analysis in the DH-populations could be due to this difference in selected versus unselected genotypes. The variation in yield and stab/adap might have been narrowed down. We tested this hypothesis with the SM population. We created subsets of 75 lines (50%) and compared QTL mapping in the subsets with the results from the full sets. The subsets were based on (i) the highest mean yield, (ii) the highest slopes, and (iii) random 'selection'. The random sets were used a reference for the other subsets, as the reduction in number of lines already reduces the power of QTL analyses. The random set selection was repeated ten times. The results (not shown) showed that (a) a reduction in sample size indeed reduced the power of QTL analyses, (b) the selection of only genotypes with the highest mean yield significantly reduced the power of QTL analysis for mean yield, as compared to the random sets, and (c) the selection of only genotypes with the highest slopes reduced the power of QTL analysis for slope, but the difference with the random set was not significant. Especially for slope, the random sets showed high variance with sometimes LOD scores for QTL presence being comparable to the full set of DHs, and sometimes LOD scores being close to zero. The reduction in variation for the trait of interest due to selection clearly could obscure the detection of QTLs in sets of commercial cultivars.

Final conclusion

Breeding for stability or adaptability relies on the heritability of stab/adap measures. High repeatability, estimated as the correlation of genotypic assessments across independent datasets, indicates a high degree of genetic determination. Several studies have investigated this aspect for different measures giving non-conclusive results (see Annicchiarico, 2002 for references). In general, σ^2_{SH}, β_i and σ^2_{ER} can vary largely in repeatability, but on average their repeatability is low and always lower than the repeatability of the measures for the mean expression of the trait. AMMI derived measures allow for a slight increase of repeatability compared to σ^2_{SH}, β_i and σ^2_{ER}.

Our research showed that the stability measures σ^2_{SH} and σ^2_{ER} did not lead to the detection of QTLs, while the adaptability measures β_i and AMMI *PC*s led to QTLs in three of the four DH-populations. Therefore, the adaptability measures seem to be meaningful measures, while stability measures are probably not. Selection for adaptability measures using molecular markers should be possible.

Acknowledgements

We thank Anton Vels for assistance with crossing Henni and Meltan and preparing field trials, Advanta Seeds BV for the creation of the DHs, Sergio Ceretta for multiplication of the DHs in Uruguay. Ignacio Romagosa and Bill Thomas supervised the field trials in Spain and Scotland, respectively, where the trials were included in the trials of the EU-Incomed project Mapping Adaptation of Barley to Drought Environments (ICA3-2002-10073). Furthermore, Pim Lindhout and Petra van den Berg are acknowledged for assistance in the AFLP analyses, and Stefanos Kalogiannis for improving the genetic linkage map. Part of this research was supported by the Technology Foundation STW, applied science division of Netherlands Organization for Scientific Research, and the technology program of the Ministry of Economic Affairs of The Netherlands.

Chapter 

Summarizing discussion

Chapter 5

Summarizing Discussion

We used a germplasm of modern two-row spring barley cultivars and several barley doubled haploid (DH) populations to investigate the potential of linkage disequilibrium (LD) mapping of many traits, and to study yield stability and yield adaptability in both a germplasm and in DH populations.

The main findings of our research were:

(1) LD between molecular markers was found up to 10 cM in our germplasm of barley cultivars (*chapter 2*),

(2) LD mapping of traits is feasible as marker-trait associations were found for yield and yield stability, resistance to leaf rust and barley yellow dwarf virus, plant height and days to heading (*chapter 2 and 3*),

(3) QTLs for yield, yield stability and yield adaptability were mapped in four DH-populations, partly overlapping between the DH-populations and partly overlapping with QTLs mapped in the cultivars (*chapter 4*).

We will now expand on these results, discuss them in the light of past and current research, and give possibilities for future research which are most needed and promising in our opinion.

1. LINKAGE DISEQUILIBRIUM IN BARLEY

a) LD measured with molecular markers

This research: We measured LD in 148 modern two-row spring barley cultivars using 236 AFLP markers. Associations between markers, expressed as the correlation coefficient r^2 between them, were found up to a distance of 10 cM. Genome wide scans for marker- trait association seem therefore feasible in a germplasm like ours.

Discussion: In comparison to other species, an LD interval up to 10 cM is large (see Chapter 1, Table 3). Only in *Arabidopsis* larger distances were found (> 50 cM), but this was

in populations founded by only a few genotypes and after extreme inbreeding (Nordborg *et al.,* 2002). Comparable or smaller distances were found for sugarcane (10 cM), sugar beet (3 cM), maize (200 bp up to 100 kbp), and soybean (50 kbp) (see chapter 1 for references).
In chapter 1 (Table 1) a number of factors were reported affecting LD in a population. Many factors that can increase the level of LD in a population probably played a role in our germplasm, *e.g.* selection, inbreeding, population admixture, and a small effective population size increasing genetic drift. These factors will be discussed under section (b) of this chapter.

Future research on LD in species should include the investigations of more crops, more populations within crops, and the integration of models for processes at the population level with processes on the DNA level.
Investigations of more crops and more populations within crops could give more insight into the factors affecting LD: the reproduction system (selfing/outcrossing), genetic isolation, population structure, population admixture, selection, and population size. Our finding of LD up to 10 cM could then be compared with results in barley germplasms of winter barley, or six-rowed barley, or barley cultivars from other parts of the world, populations of landraces, or with many other crops that are inbreeders like barley or outbreeders like maize.
Investigations at the DNA level could enhance insight on the effect on LD of factors such as recombination rate, mutation rate and genomic rearrangements. Recombination rates are not homogeneous across the human genome (Stumpf, 2002). Crossovers appear to be localized in short hot-spots that separate longer stretches of DNA. Markers within these low-recombination blocks show increased levels of LD. The same effect was found in maize (Weil, 2002; Yao *et al.,* 2002; Fu *et al.,* 2002). Also in barley heterogeneity of recombination rates was found along individual chromosomes (Künzel *et al.,* 2000). To date, the direct relation between hot-spots and structure of LD has not been demonstrated in plants, but it is likely that this relation exists. Predicting LD levels solely on the basis of physical distance will than be problematic. An interesting observation in this respect was done by Eisenbarth *et al.* (2000). They found in the human genome that chromosome segments with a high level of G+C showed low LD, up to a 100-fold difference with the surrounding segments. This might imply that the level of G+C in a chromosome region could be predictive for the level of LD. The research to recombination rates, mutation rates and genomic rearrangements is still at an embryonic stage. Progress in this area could help explain and predict levels of LD in different crops and in different chromosomes and chromosome regions with crops.

In conclusion, LD on distances up to 10 cM seems promising for LD mapping of traits. However, one should keep in mind that this distance of 10 cM does not imply that a map with a marker every 20 cM is dense enough to find any existing trait locus. In some regions it will be enough, but in other regions more markers could be required. The big question is: which region requires which density?

b) LD mapping of traits

This research showed that LD mapping in our germplasm of cultivars resulted in the detection of marker-trait associations for yield and yield stability, leaf rust (LR) resistance, barley yellow dwarf (BYD) virus resistance, plant height (PH), and days to heading (DTH) (*chapter 2 and 3*). A major gene for LR (*Rph3*), known from literature, was confirmed, as well as QTLs for yield, LR, BYD, PH and DTH. Besides, novel QTLs for yield, LR and BYD were detected.

Discussion: This is not the first time that marker-trait associations were reported. Earlier reported associations between markers and traits in germplasm collections include oat (Beer *et al.,* 1997), rice (Virk *et al.,* 1996), maize (Thornsberry *et al.,* 2001), sea beet (Hansen *et al.,* 2001), barley (Igartua *et al.,* 1999), wild barley (Ivandic *et al.,* 2003), and potato (Simko *et al.,* 2004).

Until recently, all QTL analyses in plants were based on a segregating population from a bi-parental cross (linkage analysis). The advantages of linkage analyses are that the population history is known, that LD is only based on linkage, and that the methodology has evolved to a certain maturity (Mackay, 2001; Hackett, 2002). Disadvantages of linkage analysis are that only one or a few meioses have occurred, so the resolution of mapping is low. Furthermore, the genetic background is typically narrow as usually only two parents are used, so that only two alleles per locus can be studied. Besides, phenotypic data has to be collected from scratch requiring extra time and money, especially when complex traits are concerned.

In human genetics QTL analysis is mostly based on LD mapping in germplasms. Nowadays, LD mapping is also becoming more popular for QTL analysis in plant crops. The advantages of LD mapping over linkage analysis are that existing germplasms can be used, for which often phenotypic data are already available. As many meioses have accumulated in history, especially in outbreeders, the resolution of mapping is much higher than with linkage analysis. Furthermore, a broad genetic background can be scanned at once, including multiple alleles per

locus. Disadvantages of the LD approach are that the methodology is still under development, although progress is being made. For example, Jannink and Walsh (2002) give an overview of issues in association mapping in plant populations, and Bink and Meuwissen (2004) propose a theoretical approach for fine mapping of QTLs with LD in inbred plant populations.
Combined employment of linkage analysis and LD mapping is also being considered (Xiong and Jin, 2000; Wu *et al.,* 2002), although they only give theoretical examples without real world data.

LD in germplasms is not only due to linkage, but also to other factors. Most of these involve demographic aspects of a population, and they tend to disturb the relationship between LD and the physical distance between loci. Furthermore, if LD is due to something else than linkage, the association found will be observable only in the population under study and can not be extrapolated to other populations. Literature reviews describing many of these factors have been written by Jorde (2000), Cardon and Bell (2001), Ardlie *et al.* (2002), Flint-Garcia *et al.* (2003) and Rafalski and Morgante (2004). We will now elaborate on population structure, population admixture, population age, population size, and selection.
Population structure: The presence of population structure, i.e. the presence of groups within the population differing in allele frequencies, can result in spurious associations not based on physical linkage (Pritchard and Rosenberg, 1999). Statistical methodologies can help in detecting population structure, and correct for it. Pritchard *et al.* (2000a,b) have developed an approach that incorporates estimates of population structure directly into the association test statistic. This approach has been used in an association mapping study with flowering time in maize (Thornsberry *et al.*, 2001). The number of false positives was reduced with 80%, and the resolution of mapping was greatly increased.
In our germplasm we did find some population structure using the marker data, but the markers that were indicative for the grouping were not involved in marker-trait associations. We therefore concluded that the marker-trait associations we found were not due to population structure (*chapter 2*).
Population admixture: Admixture or migration (gene flow) between populations can create LD. Initially, LD will be proportional to the allele frequency differences between the populations and unrelated to the distance between markers. In subsequent generations this spurious LD will vanish faster than LD based on linkage. In theory, this allows admixture mapping (Darvasi and Shifman, 2005). Zhu *et al.* (2005) give an example of admixture

mapping with hypertension loci in humans mapped in African Americans. Individuals from Nigeria and Europe provide estimates of allele frequencies for the African and European ancestral populations, respectively. Admixture mapping falls between linkage analysis and LD mapping in many aspects (see Darvasi and Shifman, 2005). The success of this approach will depend heavily on the time since admixture occurred and the frequency differences of the loci of interest in the parental populations. Application of admixture mapping in cultivar populations is conceivable, for example when new resistance genes have been introduced from exotic sources causing allele frequencies to shift in the regions containing the resistance loci.

Population age: LD reflects the history of recombination. In general, many studies in humans demonstrated higher levels of LD in recently founded populations than in older populations (see Jorde, 2000). Younger populations, therefore, may be most useful for initial detection of marker-trait associations at larger distances. Subsequently, older populations, in which recombinations have accumulated, may be more useful for fine-mapping. Out population of barley cultivars could be considered very new as modern breeding started around early 1900. According to Russell *et al.* (2000) a small set of 20 founding genotypes from that period contain 80% of all genetic variation in modern European cultivars. If that was the beginning of our population it may be considered extremely young and the level of LD will be high. On the other hand, barley is a crop that was domesticated many centuries ago, so age is a relative concept in that respect.

Population size: Genetic drift in isolated small populations can increase LD. On the other hand, rapid population growth will reduce LD due to reduced genetic drift. A bottleneck in the history of a population will increase the level of LD. If Russell *et al.* (2000) are correct and the modern cultivar population was founded mainly by a small set of genotypes, this may be considered a bottleneck in history that increased the level of LD.

Selection: Selection increases LD by a hitchhiking effect in which an entire haplotype that flanks a favored variant can be swept to high frequency or even fixation. Besides, epistatic selection for combinations of alleles at two or more loci will cause LD between the concerning regions. Our population existed of cultivars from several breeders. We can be sure that selection has occurred and that the LD increasing effects described above have taken place.

Future research on LD mapping of traits must address all factors mentioned above in order to be able to predict in which crops and in which populations LD mapping might be useful. LD studies with more crops and with more populations within crops are required. LD mapping

provides opportunities for QTL analysis, especially in existing germplasms with phenotypic databases already available. For example, the exploration of seed bank core collections can help unlocking the genetic potential from the wild. A number of issues need to be addressed promptly including the applied marker systems, the use of integrated maps, pedigree data, phenotypic data, and the determination of significance in association mapping.

Applied markers systems: The best marker system for LD mapping in germplasms should be highly automated to precipitate the work, yield informative and multi-allelic markers, and enable comparisons between germplasms. The most applied marker systems in barley are RAPDs, RFLPs, AFLPs and SSRs. Comparisons between these marker systems have been given by several authors, *e.g.* Russell *et al.* (1997a) and Powell (1996). At the moment, SSRs seem to be the choice for genetic diversity studies in barley (Becker and Heun, 1995; Russell *et al.,* 1997b; Struss and Plieske, 1998; Pillen *et al.,* 2000; Matus and Hayes, 2002; Karakousis *et al.,* 2003a). Many SSRs have been developed and mapped (Liu *et al.,* 1996; Ramsay *et al.,* 2000; Künzel and Waugh, 2002; Li *et al.,* 2003, Ablett *et al.,* 2003). SSRs can be applied easily, once they are developed, and because they are multi-allelic they are highly informative in germplasms. Furthermore, they can be used to compare populations. A disadvantage of SSRs might be that they appear to be sensitive to mutation (Ellegren, 2002). The best marker system for LD mapping is probably a combination of several systems, combining the ease of data generation of AFLPs with the multi-allelism of SSRs and the transferability between populations of RFLPs. Thomas *et al.* (1998) give an example where they combined AFLPs for map making with SSRs for comparing with other studies and tracing the origin of a segment in breeding history.

Integrated maps: As no genetic linkage maps based on germplasms can be made, LD mapping relies on linkage maps from mapping populations. Several integrated maps are known in barley. The BIN map from Kleinhofs and Graner (2001) is probably the best known integrated map, as it is referred to by many authors and used by the North American Barley Genome Project. The BIN-map integrates three DH-populations and contains mainly RFLPs, but also AFLPs and SSRs are incorporated. Other integrated maps are an RFLP map from Qi *et al.* (1996), integrating four barley populations, an SSR-map from Ablett *et al.* (2003) integrating eight Australian DH-populations, an AFLP-map from Kraakman *et al.* (2004) integrating three barley populations, and an integrated map with SSRs, RFLPs and AFLPs (Karakousis *et al.,* 2003b). Map integration methodology still needs attention, and with more synergy between all

efforts world wide a valuable dense multi-marker barley map could be feasible within a few years.

Pedigree information and Phenotypic data: Marker data will soon no longer be a bottleneck, but the phenotypic data and pedigree information may hamper the unlocking of many existing germplasms by LD mapping. The amount and quality of phenotypic data will be limiting, and the collection of those data should be coordinated and harmonized. Besides, as we have seen that population dynamics have a large effect on the level of LD, pedigree information should be of high quality and complete to fully employ the possibilities of LD mapping.

Significance of marker-trait associations: The rapidly growing databases with marker data and phenotypic data raised the question of determining significance thresholds for association mapping. When many traits are tested with many markers genome-wide, a Bonferroni correction for the multiple testing will cause a very low power to detect associations. We applied a novel approach in which a maximum false discovery rate is set for the found associations (*chapter 2 and 3*; Benjamini and Hochberg, 1995; Storey and Tibshirani, 2003). More discussion is required to come to a standard methodology and to be able to compare the numerous results on association mapping that will be reported in the upcoming years.

Final remarks on LD mapping

LD mapping will be a valuable extension to the geneticist's and breeder's toolbox. LD mapping of traits will be especially interesting for complex traits, where measurements are costly and time consuming. As cultivars are already tested for many years in many environments, exploiting the databases with phenotypic data of cultivars in LD studies for traits as yield, yield stability and yield adaptability seems a very obvious and beneficial path for future research.

2. GENOTYPE × ENVIRONMENT ANALYSIS FOR YIELD

In this research we studied yield, yield stability and yield adaptability using a set of cultivars and four DH-populations of barley. Yield adaptability was statistically defined as the responsiveness to the environment, and stability was defined as the deviation from the applied model. LD mapping in the cultivars resulted in marker-trait associations for yield and yield stability, of which many in a region of the genome where QTLs for yield or QTL×E interaction

were reported earlier in literature (*chapter 2*). Linkage analysis with the DH-populations showed many QTLs for yield and yield adaptability and one QTL for yield stability (*chapter 4*). There was a high degree of coincidence between the DH-populations for the location of yield QTLs and there was moderate coincidence for yield adaptability QTLs. Furthermore, QTLs for the adaptability measure based on Finlay-Wilkinson regression (β, the slope of regression) often coincided with the adaptability measures based on an AMMI approach (*PCs*, principal components). Interestingly, not all QTLs for yield adaptability coincided with QTLs for the main trait yield, suggesting an independent regulation of adaptability. In conclusion, adaptability measures appeared to be more useful than stability measures and marker assisted selection for adaptability seems feasible.

Discussion: We have selected four aspects of our research on GE to discuss, namely the definition of stability and adaptability, a comparison with results in literature, the usefulness of measures for stability and adaptability, and the existence of genes for aspects of GE-interaction.

Definitions of stability and adaptability:

Many definitions of stability and adaptability are possible and many have been given in literature (see reviews of Lin *et al*., 1986; Becker and Leon, 1988; Annicchiarico, 2002). The definition will always be based on the perspective of the researcher. An important contrast can be found between breeders/farmers on the one hand and statisticians on the other. Adaptability is for breeders/farmers the ability to adapt to all environmental conditions present in the set of goal environments. For a farmer this will be the one location where the farm is and all conditions occurring in time, while for the breeder the scope is broader and many locations in many years are considered. A highly adapted genotype will then give high yield in all occurring environments. A statistician, however, often defines adaptability as the response of a genotype to the changing conditions. A genotype with high adaptability will then have low yield in bad environments and high yield in good environments. The quality of the environments is often measured by the average result of a group of genotypes. An example of a measure for adaptability is the slope of the regression on the mean yield of all genotypes over many environments. This regression model can predict a genotype's performance in an environment, and the deviation from the regression can then be considered as a measure for the genotype's stability. More general one could say that adaptability is a parameter in a model describing the genotype's response to the environment, and stability is the deviation from the

model's prediction of the genotype's performance. Under this definition the quality of the model determines the genotype's stability. Furthermore, in theory a model could predict the genotype's performance perfectly and its stability will be complete.

Compare our results with literature:

The DH-populations Steptoe × Morex (SM), Harrington × Morex (HM), and Harrington × TR306 (HT) have been used by others to map QTL. All authors reported main effects, but also QTL×E interaction effects.

In the SM population many effects for QTL×E were found (Hayes *et al.,* 1993, 1996; Romagosa *et al.,* 1996; Voltas *et al.,* 2001). The most significant effects for QTL×E were on chromosome 2, 3 and 6. In the same regions we found QTLs for AMMI *PCs* in this population. Romagosa *et al.* (1996) found QTLs at the same location for exactly the same *PCs*. Furthermore, we found QTLs for β, the slope of regression, at the same location on chromosome 3 and 6. Only the QTL on chromosome 3 coincided with a QTL for the main effect. Hayes *et al.* (1996) reported many more QTL×E effects on all chromosomes, but it seems that they used a low significance threshold. Romagosa *et al.* (1996) reported another six QTL×E effects on chromosome 1, 2, 4, 5, and 7. Three of them coincided with a QTL for a *PC* in our analyses.

The HT population has been studied by Tinker *et al.* (1996). They reported three primary QTLs for QTL×E effect on chromosome 1, 3, and 7. The latter two were confirmed in our analysis by QTLs for *PCs*. In general, GE in HT was low compared to SM.

The HM population has been studied by Marquez-Cedillo *et al.* (2001). They found two QTLs for QTL×E effect on chromosome 2 and 5. Both were confirmed in our analysis by QTLs for *PC*, and on chromosome 5 also a QTL for β was found. In our analyses we found two more QTLs for *PCs* on chromosome 3 and 4. No QTLs for the main effect were found on chromosome 4 and 5, neither by Marquez-Cedillo *et al.* (2001) nor in our analyses.

In conclusion, the QTLs for β and *PCs* we found in SM, HT and HM were in agreement with QTL×E effects and QTLs for *PCs* reported earlier. The coincidence of QTLs for QTL×E effects between populations is often considered to be low (Voltas *et al.,* 2002). In our data, however, a number of regions on the genome showed QTLs for β and *PCs* in two or more populations (Figure 2, Chapter 4). Considering the fact that the measures we have chosen depend on both the environments and on the genotypes, which both differ among the datasets, these results could be qualified as promising for the potential of selection for adaptability.

Usefulness of measures for stability and adaptability:

The usefulness of stability and adaptability measures depends largely on their genetic determination or heritability. Heritability can be indicated by the repeatability of a measure or by detection of QTLs. As we have found many QTLs for β and *PCs,* these measures seem heritable and selectionable. In contrast, only one QTL for a stability measure was found, so these measures seem not heritable. Becker and Leon (1988) reported moderate heritability for adaptation parameters. Sneller *et al.* (1997) showed in soybean that the repeatability of β and *PCs* was moderate, while for both Shukla's and Eberhart-Russell's stability variance it was low. Nurminiemi *et al.* (2002) demonstrated in spring barley that β and *PCs* were correlated, and *PCs* were usefull for determining the magnitude and nature of QTL×E. In a review by Anicchiarico (2002) he concluded that the repeatability of stability and adaptability measures varies largely between crops and datasets, but it is always lower than the repeatability of the main trait. AMMI derived measures allow for a slight increase of repeatability compared to other dynamic measures.

Genes for GE-interaction:

In the Henni × Meltan population (HeMe) we found QTLs for adaptability measures coincident with QTLs for the main trait, but also solitaire on the genome. Similar effects for the other DH-populations were found. Ungerer *et al.* (2003) described two models for the genetic regulation of GE: (1) the allelic sensitivity model, in which gene effects change in different environments, and (2) the gene regulation model, in which regulatory genes change their effect in different environments. These models are not mutually exclusive. Under the first model QTLs for the main effect should coincide with QTLs for the GE effect, while under the second model QTLs for the GE effect can be located separately. Our results support both models as both coinciding and separate QTLs were found.

Little is known about the regulatory system leading to stress response and adaptation (Cattivelli *et al.,* 2002), although some examples of the interaction between stress responsive genes and environmental factors have been demonstrated (Maestri *et al.,* 2002).

Future research should address the interaction between genotypes and the environment at several levels, *e.g.* gene regulation, plant physiology, and crop physiology. Are QTLs for the slope of the regression useful for breeders? Will breeders select for optimal slope? As slope is depending on environments and genotypes included in the analysis it is quite remarkable that we have found QTLs for this trait coinciding between populations. However, there is no

guarantee that the same QTLs will be found when these populations will be tested in other environments, as nothing is known about the environmental factors that caused the effects or about the regulation within the plant. The measure slope is a "black box" measure, not giving information on the underlying interaction between genes and environmental conditions. We explored relatively simple models describing GE without direct inclusion of environmental factors. AMMI can give most information, in this respect, as the hypothetical environments that are fitted could be related to environmental factors in order to disclose the most important factors. However, those factors could also be included in the analysis directly, provided that the measurements are available. Selecting the right factors to include in the model is the next challenge. Examples of this approach are given by Voltas *et al.* (1999a,b) and Malosetti *et al.* (2004). Further enhancements of GE models are the integration with crop physiology and the decomposition of complex traits (*e.g.* yield) into components interacting with the environment. This approach is followed in the MABDE-project (Mapping Adaptation of Barley to Drought Environments). Cultivars, landraces and DH-populations of barley are used to study drought resistance integrating genetics, physiology and environmental factors.

Final remarks on GE analysis for yield

Linkage analysis in four DH-populations demonstrated that QTLs for adaptability measures can be identified. Selection for these measures seems feasible, especially because a number of QTLs coincided between populations, and because some QTLs for adaptability did not coincide with a QTL for the main trait. As adaptability expressed the response of a genotype to the environment, further research should focus on the identification of relevant environmental factors and integrate those factors in GE-analysis.

References

Ablett, G. A., A. Karakousis, L. Banbury, M. Cakir, T. A. Holton *et al.*, 2003 Application of SSR markers in the construction of Australian barley genetic maps. Australian J Agric Res 54: 1187-1195.

Allard, R. W., 1997 Genetic basis of the evolution of adaptedness in plants, pp. 1-12 in *Adaptation in Plant Breeding*, edited by P. M. A. TIGERSTEDT, Selected papers from the XIV EUCARPIA Congress on Adaptation in Plant Breeding, Jyväskylä, Sweden.

Annicchiarico, P., 2002 Genotype × environment interactions. Challenges and opportunities for plant breeding and cultivar recommendations, pp. 115 in *Plant production and protection paper*. FAO, Rome.

Ardlie, K. G., L. Kruglyak, and M. Seielstad, 2002 Patterns of linkage disequilibrium in the human genome. Nature Reviews Genetics 3: 299-309.

Arru, L., R. E. Niks, P. Lindhout, G. Valé, E. Francia *et al.*, 2002 Genomic regions determining resistance to leaf stripe (*Pyrenophora graminea*) in barley. Genome 45: 460–466.

Backes, G., L. Madsen, H. Jaiser, J. Stougaard, M. Herz *et al.*, 2003 Localisation of genes for resistance against Blumeria graminis f.sp. hordei and Puccinia graminis in a cross between a barley cultivar and a wild barley (*Hordeum vulgare* ssp. *spontaneum*) line. Theor Appl Genet 106: 353-362.

Becker, H. C., 1981 Correlations among some statistical measures of phenotypeic stability. Euphytica 30: 451-458.

Becker, H. C., and J. Leon, 1988 Stability analysis in plant breeding. Plant Breeding 101: 1-23.

Becker, J., and M. Heun, 1995 Barley microsatellites: allele variation and mapping. Plant Molecular Biology 27: 835-845.

Becker, J., P. Vos, M. Kuiper, F. Salamini, and M. Heun, 1995 Combined mapping of AFLP and RFLP markers in barley. Mol. Gen. Genet. 249: 65-73.

Beer, S. C., W. Siripoonwiwat, L. S. O'donoughue, E. Souza, D. Matthews *et al.*, 1997 Associations between molecular markers and quantitative traits in an oat germplasm pool: Can we infer linkages? Journal of Quantitative Triat Loci 3.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc 57: 289-300.

Benneth, M. D., and J. B. Smith, 1976 Nuclear DNA Amounts in Angiosperms. Philos Trans R Soc Lon Biol Sci 274: 227-274.

Berloo, R. van, H. Aalbers, A. Werkman and R. E. Niks, 2001 Resistance QTL confirmed through development of QTL-NILs for barley leaf rust resistance. Molecular Breeding 8: 187-195.

Bezant, J. H., D. A. Laurie, N. Pratchett, J. Chojecki, and **M. J. Kearsey, 1996** Marker regression mapping of QTL controlling flowering time and plant height in a spring barley (*Hordeum vulgare* L.) cross. Heredity 77: 64-73.

Bezant, J., D. Laurie, N. Pratchett, J. Chojecki and M. Kearsey, 1997 Mapping QTL controlling yield and yield components in a spring barley cross using marker regression. Molecular Breeding 3: 29-38.

Bink, M. C. A. M., and T. H. E. Meuwissen, 2004 Fine mapping of quantitative trait loci using linkage disequilibrium in inbred plant populations. Euphytica 137: 95-99.

Brown, A. H. D., 1989 Core collections: a practical approach to genetic resources management. Genome 31: 818-824.

Burnett, P. A., A. Comeau and C. O. Qualsett, 1995 Host plant tolerance or resistance for control of barley yellow dwarf, PP. 321-343 in Barley yellow dwarf. Forty years of progress, edited by C. D'Arcy and P. A. Burnett. APS, St.Paul.

Cardon, L. R., and J. I. Bell, 2001 Association study designs for complex diseases. Nature Reviews Genetics 2: 91-99.

Cattivelli, L., P. Baldi, C. Crossati, M. Grossi, G. Valé *et al.*, 2002 Genetic bases of barley physiological response to stressful conditions, pp. 307-360 in *Barley Science*, edited G. A. SLAFER, J. L. MOLINA-CANO, R. SAVIN, J. L. ARAUS, and I. ROMAGOSA, The Haworth Press, Binghamton, NY.

Chelkowski, **J., M**. **Tyrka**, **and A. Sobkiewicz, 2003** Resistance genes in barley (*Hordeum vulgare* L.) and their identification with molecular markers. J Appl Genet 44: 291-309.

Churchill, G. A., and R. W. Doerge, 1994 Empirical threshold values for quantitative trait mapping. Genetics 138: 963-971.

Collins, N.C., N.G. Paltridge, C.M. Ford, and R.H. Symons, 1996 The *Yd2* gene for barley yellow dwarf virus resistance maps close to the centromere on the long arm of barley chromosome 3. Theor Appl Genet 92: 858-864.

D'Arcy, **C.J., 1995** Symptomatology and host range of barley yellow dwarf. In D'Arcy CJ and Burnett PA (eds.) Barley yellow dwarf: 40 years of progress. The American Phytopathological Society, St.Paul, Minnesota, pp 9-28.

Darvasi, A., and S. Shifman, 2005 The beauty of admixture. Nature Genetics 37: 118-119.

Davies, P. A., and S. Morton, 1998 A comparison of barley isolated microspore and anther culture and the influence of cell culture density. Plant Cell Reports 17: 206-210.

De la Peña, R., K.P Smith, F. Capettini, G.J. Muehlbauer, M. Gallo-Meagher, R. Dill-Macky, D. Somers, and **D.C. Rasmusson, 1999** Quantitative trait loci associated with resistance to Fusarium head blight and kernel discoloration in barley. Theor Appl Genet 99: 561–569.

Dreiseitl, A, and B. J. Steffenson, 2000 Postulation of leaf-rust resistance genes in Czech and Slovak barley cultivars and breeding lines. Plant Breeding 119: 211-214.

Eberhart S.A., and W. A. Russell, 1966 Stability parameters for comparing varieties. Crop Sci. 6: 36-40.

Eisenbarth, I., G. Vogel, W. Krone, W. Vogel and G. Assum, 2000 An Isochore Transition in the NF1 Gene Region Coincides with a Switch in the Extent of Linkage Disequilibrium. Am J Hum Genet 67: 873-880.

Ellegren, H., 2002 Microsatellite evolution: a battle between replication slippage and point mutation. TRENDS in Genetics 18: 70.

Farnir, F., W. Coppieters, J.-J. Arranz, P. Berzi, N. Cambisano *et al.*, 2000 Extensive Genome-wide Linkage Disequilibrium in Cattle. Genome Res. 10: 220-227.

Farris, J. S., 1977 Phylogenetic analysis under Dollo's Law. Systematic Zoology 26: 77-88.

Finlay, K. W., and G. N. Wilkinson, 1963 The analysis of adaptation in a plant-breeding programme. Aust. J. Agric. Res. 14: 742–754.

Fischbeck, G., 2003 Diversification through breeding, pp. 29-50 in *Diversity in Barley*, edited by R. von Bothmer, T. van Hintum, H. Knüpffer, and K. Sato. Elsevier Science, Amsterdam, The Netherlands.

Flint-Garcia, S. A., J. M. Thornsberry and E. S. Buckler IV, 2003 Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54: 357-374.

Ford, C. M., N. G. Paltridge, J. P. Rathjen, R. I. Moritz, R. J. Simpson *et al.*, 1998 Rapid and informative assays for Yd2, the barley yellow dwarf virus resistance gene, based on the nucleotide sequence of a closely linked gene. Mol Breed 4: 23-31.

Forster, B. P. , R. P. Ellis, W. T. B. Thomas, A. C. Newton, R. Tuberosa *et al.*, 2000 The development and application of molecular markers for abiotic stress. J Exp Botany 51: 19-27.

Franckowiak, J., 1997 Revised linkage maps for morphological markers in barley, *Hordeum vulgare*. Barley Genetics Newletter 26: 9-21.

Friedt, W., and B. Foroughi-Wehr, 1981 Anther culture of barley: plant regeneration

and agronomic performance of diploid progenies, pp. 690-698 in M. ASHER (ed.): *Proc 4th Int Barley Genet Symp,* Edingburgh Univ. Press, Edingburgh, UK.

Fu, H., Z. Zheng and H. K. Dooner, 2002 Recombination rates between adjacent genic and retrotransposon regions in maize vary by 2 orders of magnitude. PNAS 99: 1082-1087.

Gauch, H. G. J., 1992 *Statistical analysis of regional trials: AMMI analysis of factorial design*. Elsevier, Amsterdam.

Gaut, B. S., and A. D. Long, 2003 The lowdown on linkage disequilibrium. The Plant Cell 15: 1502-1506.

Giese, H., A. G. Holm-Jensen, H. P. Jensen and J. Jensen, 1993 Localization of the Laevigatum powdery mildew resistance gene to barley chromosome 2 by the use of RFLP markers. Theor Appl Genet 85: 897-900.

Gollob, H. F., 1968 A statistical model which combines features of factor analysis and analysis of variance techniques. Psychometrika 33: 73-115.

Gordon, A.D., 1981 *Classification.* Chapman and Hall, London.

Graner, A., A. Jahoor, J. Schondelmaier, H. Siedler, K. Pillen *et al.*, 1991 Construction of an RFLP map of barley. Theor Appl Genet 83: 250-256.

Greenacre, M. J., 1984 *Theory and application of correspondence analysis*. Academic Press, London.

Hackett, C. A., 2002 Statistical methods for QTL mapping in cereals. Plant Molecular Biology 48: 585-599.

Hansen, M., T. Kraft, S. Ganestam, T. Säll, and N-O. Nilsson, 2001 Linkage disequilibrium mapping of the bolting gene in sea beet using AFLP markers. Genet Res Camb 77: 61-66.

Hayes, P. M., B. H. Liu, S. J. Knapp, F. Chen, B. Jones *et al.*, 1993 Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. Theor. Appl. Genet. 87: 392-401.

Hayes, P. M., G. Briceño and J. Cerono, 1996 Using QTL in barley, pp. 182-188 in *Proceedings of the VII Barley Genetics Symposium*, edited by G. SCOLES and B. ROSSNAGEL. Univ.of Saskatchewan, Canada.

Hayes, P. M., J. Cereno, H. Witsenboer, M. Kuiper, M. Zabeau *et al.*, 1997 Characterizing and exploiting genetic diversity and quantitative traits in barley (*Hordeum vulgare*) using AFLP markers. Journal of Quantitative Trait Loci 3.

Heun, M., A. E. Kennedy, J. A. Anderson, N. L. V. Lapitan, M. E. Sorrells *et al.*, 1991 Construction of a RFLP map for barley (*Hordeum vulgare*). Genome 34: 437-447.

Hintum, T. J. L., 1992 The barley core collection, pp. 703-707 in *Proc 6th Int Barley Genet Symp*, edited by L. MUNCK. Munksgaard International Publishers, Helsingborg, Sweden.

Hittalmani, S., N. Huang, B. Courtois, R. Venuprasad, H. E. Shashidhar *et al.*, 2003 Identification of QTL for growth- and grain yield-related traits in rice across nine locations of Asia. Theor Appl Genet 107: 679-690.

Holland, B. S., and **M. D. Copenhaver, 1987** An improved sequentially rejective Bonferroni test procedure. Biometrics 43: 417-423.

Holm, S., 1979 A simple sequentially rejective multiple test procedure. Scand J Statist 6: 65-70.

Holton, T. A., J. T. Christopher, L. Mcclure, N. Harker and R. J. Henry, 2002 Identification and mapping of polymorphic SSR markers from expressed sequences of barley and wheat. Mol Breed 9: 63-71.

Igartua, E., A. M. Casas, F. Ciudad, J. L. Montoya and I. Romagosa, 1999 RFLP markers associated with major genes controlling heading date evaluated in a barley germ plasm pool. Heredity 83: 551-559.

Ivandic, V., W. T. B. Thomas, E. Nevo, Z. Zhang and B. P. Forster, 2003 Associations of simple sequence repeats with quantitative trait variation including biotic and abiotic stress tolerance in *Hordeum spontaneum*. Plant Breeding 122: 300-304.

Jannink, **J. L., and J. B. Walsh, 2002** Association mapping in plant populations, pp. 59-68 in *Quantitative Genetics, genomics and plant breeding*, edited by M. S. Kang. CABI, Wallingford, UK.

Jannink, J. L., M. C. A. M. Bink and R. C. Jansen, 2001 Using complex plant pedigrees to map valuable genes. TRENDS in Plant Science 6: 337-342.

Jannoo, N., L. Grivet, A. Dookun, A. D'Hont and J. C. Glaszmann, 1999 Linkage disequilibrium among modern sugarcane cultivars. Theor Appl Genet 99: 1053-1061.

Jansen, R. C., 1993 Interval mapping of multiple quantitative trait loci. Genetics 135: 205–211.

Jefferies, S. P., B. J. King, A. R. Barr, P. Warner, S. J. Logue *et al.*, 2003 Marker-assisted backcross introgression of the *Yd2* gene conferring resistance to barley yellow dwarf virus in barley. Plant Breeding 122: 52-56.

Jensen, H. P., and J. H. Jørgensen, 1991 Genetics of '*Laevigatum*' resistance and virulence in barley and powdery mildew. Barley Genetics Newletter 21: 34-37.

Jin, Y., and B. J. Steffenson, **1994** Inheritance of resistance to *Puccinia hordei* in cultivated and wild barley. J Hered 85: 451-454.

Jorde, L. B., 2000 Linkage Disequilibrium and the Search for Complex Disease Genes. Genome Res. 10: 1435-1444.

Kang, M. S., 2002 Genotype-Environment interaction: progress and prospects, pp. 221-243 in Quantitative genetics, Genomics and Plant Breeding, edited by M. S. KANG. CABI Publishing, Oxon, UK.

Karakousis, A., K. J. Chalmers, A. R. Barr, G. A. Ablett, T. A. Holton *et al.*, 2003a Potential of SSR markers for plant breeding and variety identification in Australian barley germplasm. Australian Journal of Agricultural Research 54: 1197-1210.

Karakousis, A., J. P. Gustafson, K. J. Chalmers, A. R. Barr and P. Langridge, 2003b A consensus map of barley integrating SSR, RFLP, and AFLP markers. Australian Journal of Agricultural Research 54: 1173-1185.

Kasha, K. J., and K. N. Kao, 1970 High frequency haploid production in barley. Nature 225: 874-876.

Kicherer, S., G. Backes, U. Walther, and A. Jahoor, 2000. Localising QTLs for leaf rust resistance and agronomic traits in barley (*Hordeum vulgare* L.). Theor Appl Genet 100: 881-888.

Kleinhofs, A., A. Kilian, M. Saghai Maroof, R. Biyashev, P. M. Hayes *et al.*, 1993 A molecular isozyme and morphological map of the barley (*Hordeum vulgare*) genome. Theor Appl Genet 86: 705-712.

Kleinhofs, A., and A. Graner, 2001 An integrated map of the barley genome, pp. 187-199 in *DNA-Based markers in plants*, edited by R. L. PHILIPS and I. K. VASIL. Kluwer Academic Publishers, The Netherlands.

Knight, R., 1970 The measurement and interpretation of genotype-environment interactions. Euphytica 19: 225-235.

Koopman, W. J. M., **and G. Gort, 2004** Significance Tests and Weighted Values for AFLP Similarities, Based on Arabidopsis in Silico AFLP Fragment Length Distributions. Genetics 167: 1915-1928.

Koorevaar, G., 1997 QTL for date of ear emergence and lodging in barley. M.Sc. Thesis, Wageningen University and Research Center, Laboratory of Plant Breeding, Wageningen, The Netherlands.

Kraakman, A. T. W., R. E. Niks, P. M. M. M. van den Berg, P. Stam and F. A. van Eeuwijk, 2004 Linkage Disequilibrium Mapping of Yield and Yield Stability in Modern Spring Barley Cultivars. Genetics 168: 435-446.

Kraft, T., M. Hansen, and N-O. Nilsson, 2000 Linkage disequilibrium and fingerprinting in sugar beet. Theor Appl Genet 101: 323-326.

Künzel, G., and R. Waugh, 2002 Integration of microsatellite markers into the translocation-based physical RFLP map of barley chromosome 3H. Theor Appl Genet 105: 660-665.

Künzel, G., L. Korzun and A. Meister, 2000 Cytological integrated physical restriction fragment length polymorphism maps for the

barley genome based on translocation breakpoints. Genetics 154: 397-412.

Leon, J., 1985 Beiträge zur Erfassung der phänotypischen Stabilität unter besonderer Berücksichtigung unterschiedlicher Heterogenitäts- und Heterozygotiegrade sowie einer zusammenfassende Beurteilung von Ertragshöhe und Ertragssicherheit, pp. Christian Albrechts Universität, Kiel, Germany.

Leon, J., and H. C. Becker, 1988 Repeatability of some statistical measures of phenotypic stability – Correlations between single year results and multi-year results. Plant Breeding 100: 137-142.

Li, J. Z., T. G. Sjakste, M. S. Röder and M. W. Ganal, 2003 Development and genetic mapping of 127 new microsatellite markers in barley. Theor Appl Genet 107: 1021-1027.

Lin, C. S., and M. R. Binns, 1991 Genetic properties of four types of stability parameter. Theor Appl Genet 82: 505-509.

Lin, C. S., and M. R. Binns, 1994 Concepts and methods for analyzing regional trial data for cultivar and location selection, pp. 271-297 in *Plant Breeding Reviews*, edited by J. JANICK. John Wiley & Sons, New York.

Lin, C. S., M. R. Binns and L. P. Lefkovitch, 1986 Stability analysis: Where do we stand? Crop Sci. 26: 894-900.

Liu, Z. W., R. M. Biyashev and M. Saghai Maroof, 1996 Development of simple sequence repeat DNA makers and their integration into a barley linkage map. Theor Appl Genet 93: 869-876.

Macaulay, M., L. Ramsay, W. Powell, **and R. Waugh, 2001** A representative highly informative genotyping set of barley SSRs. Theor Appl Genet 102: 801-805.

Mackay, T. F. C., 2001 The genetic architecture of quantitative traits. Annu.Rev.Genet. 35: 303-339.

Maestri, E., A. Malcevschi, A. Massari and N. Marmiroli, 2002 Genomic analysis of cultivated barley (*Hordeum vulgare*) using sequence-tagged molecular markers. Estimates of divergence based on RFLP and PCR markers derived from stress-responsive genes, ans simple-sequence repeats (SSRs). Mol Genet Genomics 267: 186-201.

Malosetti, M., J. Voltas, I. Romagosa, S. E. Ullrich and F. A. van Eeuwijk, 2004 Mixed models including environmental covariables for studying QTL by environment interaction. Euphtyica 137: 139-145.

Mandel, J., 1971 A new analysis of variance model for non-additive data. Technometrics 13:1-18.

Marquez-Cedillo, L. A., P. M. Hayes, A. Kleinhofs, W. G. Legge, B. Rossnagel *et al.*, 2001 QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. Theor Appl Genet 103: 625-637.

Matus, I. A., and P. M. Hayes, 2002 Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats. Genome 45: 1095-1106.

McNeal, F. H., C. F. Konzak, E. P. Smith, W. S. Tate, and T. S. Russell, 1971 A uniform system for recording and processing cereal research data. USDA Agricultural Research Service, Washington, D.C., ARS 34-121.

Milligan, B.G., 2003 Maximum-Likelihood estimation of relatedness. Genetics 163: 1153-1167.

Montgomery, D. C., and E. A. Peck, 1982 *Introduction to linear regression analysis*, 2nd edn., John Wiley & Sons, New York.

Niks, R. E., A. Habekuß, B. Bekele and F. Ordon, 2004 A novel major gene on chromosome 6H for resistance of barley against the barley yellow dwarf virus. Theor Appl Genet 109: 1536-1543.

Niks, R. E., U. Walther, H. Jaiser, F. Martinez, D. Rubiales, O. Andersen, K. Flath, P. Gymer, F. Heinrichs, R. Jonsson, L. Kuntze, M. Rasmussen, **and E. Richter, 2000.** Resistance against barley leaf rust (*Puccinia hordei*) in West-European spring barley germplasm. Agronomie 20: 769-782.

Nordborg, M., and S. Tavaré, 2002 Linkage disequilibrium: what history has to tell us. TRENDS in Genetics 18: 83-90.

Nordborg, M., J. O. Borevitz, J. Bergelson, C. C. Berry, J. Chory *et al.*, 2002 The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nature Genetics 30: 190-193.

Nurminiemi, M., S. Madsen, O. A. Rognli, Å. Bjornstad and R. Ortiz, 2002 Analysis of the genotype-by-environment interaction of spring barley tested in the Nordic Region of Europe: Relationships among stanility statistics for grain yield. Euphytica 127: 123-132.

Ooijen, J. W. van, 1999 LOD significance thresholds for QTL analysis in experimental populations of diploid species. Heredity 83: 613–624.

Ooijen, J. W. van, 2004 MapQTL 5. Software for the mapping of quantitative trait loci in experimental populations, pp. 63. Kyazma BV, Wageningen, Netherlands.

Parlevliet, J.E. and A. van Ommeren, 1975 Partial resistance of barley to leaf rust, *Puccinia hordei*. II. Relationship between field trials, micro plots test and latent period. Euphytica 24: 293-303.

Parlevliet, J.E. and H. J. Kuiper, 1977 Resistance of some barley cultivars to leaf rust, *Puccinia hordei*: polygenic, partial resistance hidden by monogenic hypersensitivity. Neth.J.Pl.Path. 83: 85-89

Parlevliet, J.E., 1975 Partial resistance of barley to leaf rust, *Puccinia hordei*. I. Effect of cultivar and development stage on latent period. Euphytica 24: 21-27**.**

Parlevliet, J.E., 1978 Further evidence of polygenic inheritance of partial resistance in barley to leaf rust, *Puccinia hordei*. Euphytica 27: 369-379.

Parlevliet, J.E., 1981 Disease resistance in plants and its consequences for breeding. Proc.Pl. Breeding Symp. II Ed. K.J. Frey, Iowa State Univ., Ames: 309-364

Parlevliet, J.E., and A. van Ommeren, 1984 Interplot interference and the assessment of barley cultivars for partial resistance to leaf rust, Puccinia hordei. Euphytica 33: 685–697.

Pillen, K., A. Binder , B. Kreuzkam, L. Ramsay, R. Waugh *et al.*, 2000 Mapping new EMBL-derived barley microsatellites and their use in differentiating German barley cultivars. Theor Appl Genet 101: 652-660.

Poehlman, J. M., 1985 Adaptation and distribution, pp. 2-17 in *Barley*, edited by D. C. RASMUSSON. ASA/CSSA/SSA, Madison, Wisconsin.

Powell, W., E. Baird, A. Booth, P. Lawrence, M. Macaulay *et al.*, 1996 Single locus and multi-locus molecular assays for barley breeding and research, pp. 174-181 in *Proceedings of the VII Barley Genetics Symposium*, edited by G. SCOLES and B. ROSSNAGEL. Univ.Extension Press, Univ.of Saskatchewan, Canada.

Pritchard, J. K., and N. A. Rosenberg, 1999 Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet 65: 220-228.

Pritchard, J. K., M. Stephens and P. Donnely, 2000a Inference of population structure using multilocus genotype data. Genetics 155: 945-959.

Pritchard, J. K., M. Stephens, N. A. Rosenberg and P. Donnelly, 2000b Association mapping in structured populations. Am J HumGenet 67: 170-181.

Pritchard, J. K., and M. Przeworski, 2001 Linkage disequilibrium in humans: models and data. Am J Hum Genet 69: 1-14.

Qi, X., P. Stam and P. Lindhout, 1996 Comparison and integration of four barley genetic maps. Genome 39: 379-394.

Qi, X., and P. Lindhout, 1997 Development of AFLP markers in barley. Mol Gen Genet 254: 330-336.

Qi, X., F. Fufa, D. Sijtsma, R. E. Niks, P. Lindhout *et al.*, 2000 The evidence for abundance of QTLs for partial resistance to Puccinia hordei on the barley genome. Molecular Breeding 6: 1-9.

Qi, X., P. Stam, and P. Lindhout, 1998a Use of locus-specific AFLP markers to construct a high-density molecular map in barley. Theor Appl Genet 96: 376-384.

Qi, X., R. E. Niks, P. Stam, and P. Lindhout, 1998b Identification of QTLs for partial

resistance to leaf rust (*Puccinia hordei*) in barley. Theor Appl Genet 96: 1205-1215.

Rafalski, A., 2002 Applications of single nucleotide polymorphisms in crop genetics. Current Opinion in Plant Biology 5: 94-100.

Rafalski, A., and M. Morgante, 2004 Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. TRENDS in Genetics 20: 103-111.

Ramsay, L., M. Macaulay, S. Degli-Ivanissevich, K. Maclean, L. Cardle *et al.*, 2000 A simple sequence repeat-based linkage map of barley. Genetics 156: 1997-2005.

Rasmusson, D.W., and C.W.Schaller, 1959 The inheritance of resistance in barley to the yellow dwarf virus. Agron J 51: 661-664.

Reich, D. E., M. Cargill, S. Bolk, J. Ireland, P. C. Sabeti *et al.*, 2001 Linkage disequilibrium in the human genome. Nature 411: 199-204.

Remington, D. L., J. M. Thornsberry, Y. Matsuola, L. M. Wilson, S. R. Whitt *et al.*, 2001 Structure of linkage disequilibrium and phenotypic associations in the maize genome. PNAS 98: 11479-11484.

Romagosa, I., S. E. Ullrich and F. Han, 1996 Use of the AMMI model in QTL mapping for adaptation in barley. Theor Appl Genet 93: 30-37.

Rouppe van der Voort, J. N. A. M., P. M. Zandvoort, H. J. van Eck, R. T. Folkertsma, R. C. B. Hutten *et al.*, 1997 Use of allele specificity of comigrating AFLP markers to allign genetic maps from different potato genotypes. Mol Gen Genet 255: 438-447.

Russell, J. R., J. D. Fuller, M. Macaulay, B. G. Hatz, A. Jahoor *et al.*, 1997a Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. Theor Appl Genet 95: 714-722.

Russell, J. R., J. Fuller, G. Young, W. T. B. Thomas, G. Taramino *et al.*, 1997b Discriminating between barley genotypes using microsatellite markers. Genome 40: 442-450.

Russell, J. R., R. P. Ellis, W. T. B. Thomas, R. Waugh, J. Provan *et al.*, 2000 A retrospective analysis of spring barley germplasm development from 'foundation genotypes' to currently successful cultivars. Molecular Breeding 6: 553-568.

Saranga, Y., M. Menz, C.-X. Jiang, R. J. Wright, D. Yakir *et al.*, 2001 Genomic dissection of genotype x environment interactions conferring adaptation of cotton to arid conditions. Genome Research 11: 1988-1995.

Sayed, H., H. Kayyal, L. Ramsey, S. Ceccarelli and M. Baum, 2002 Segregation distortion in doubled haploid lines of barley (*Hordeum vulgare* L.) detected by simple sequence repeat (SSR) markers. Euphytica 125: 265-272.

Scheurer, K.S., W. Friedt, W.Huth, R.Waugh, and F.Ordon, 2001. QTL analysis of tolerance to a German strain of BYDV-PAV in barley (*Hordeum vulgare* L.). Theor Appl Genet 103: 1074-1083.

Shin, J. S., S. Chao, L. Corpuz and T. Blake, 1990 A partial map of the barley genome incorporating restriction fragment length polymorphism, polymerase chain reaction, isozyme, and morphological marker loci. Genome 33: 803-810.

Shukla, G. K., 1972 Some statistical aspects of partitioning genotype-environmental components of variability. Heredity 29: 237-245.

Simko, I., S. Costanzo, K. G. Haynes, B. J. Christ and R. W. Jones, 2004 Linkage disequilibrium mapping of a *Verticilium dahliae* resistance quantitative trait locus in tetraploid potato (*Solanum tuberosum*) through a candidate gene approach. Theor Appl Genet 108: 217-224.

Sneller, C. H., L. Kilgore-Norquest and D. Dombek, 1997 Repeatability of yield stability statistics in soybean. Crop Science 37: 383-390.

Sogaard, B., and P. Von-Wettstein-Knowles, 1987 Barley: genes and chromosomes. Carlsberg Res Commun 52: 123-196.

Spaner, D., L. P. Shugar, T. M. Choo, I. Falak, K. G. Briggs, W. G. Legge, D. E. Falk, S. E. Ullrich, N. A. Tinker, B. J. Steffenson, and D. E. Mather, 1998

Mapping of disease resistance loci in barley on the basis of visual assessment of naturally occurring symptoms. Crop Science 38: 843-850.

Stam, P., 1993 Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant Journal 3: 739-744.

Storey, J. D., and R. Tibshirani, 2003 Statistical significance for genomewide studies. PNAS 100: 9440-9445.

Struss, D., and J. Plieske, 1998 The use of microsatellites marekrs for detection of genetic diversity in barley populations. Theor Appl Genet 97: 308-315.

Stumpf, M. P. H., 2002 Haplotype diversity and the block structure of linkage disequilibrium. TRENDS in Genetics 18: 226-228.

Suneson, C.A., 1955 Breeding for resistance to barley yellow dwarf virus in barley. Agron J 47: 283.

Tai, G. C. C., 1971 Genotype stability analysis and its applications to potato regional trials. Crop Sci 11: 184-190.

Taillon-Miller, P., I. Bauer-Sardiña, N. L. Saccone, J. Putzel, T. Laitinen *et al.*, 2000 Juxtaposed regions of extensive and minimal linkage disequilibrium in human Xq25 and Xq28. Nature Genetics 25: 324-328.

Tenaillon, M. I., M. C. Sawkins, A. D. Long, R. L. Gaut, J. F. Doebley *et al.*, 2001 Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. *mays* L.). PNAS 98: 9161-9166.

Teulat, B., O. Merah, I. Souyris, and D. This, 2001 QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. Theor Appl Genet 103: 774-787.

Thomas, H., and J. F. Farrar, 1997 *Putting plant physiology on the map: genetic analysis of developmental and adaptive traits.* Cambridge University Press, Cambridge.

Thomas, W. T. B., E. Baird, J. D. Fuller, P. Lawrence, G. R. Young *et al.*, 1998 Identification of a QTL decreasing yield in barley linked to Mlo powdery mildew resistance. Molecular Breeding 4: 381-393.

Thomas, W. T. B., W. Powell, R. Waugh, K. J. Chalmers, U. M. Barua *et al.*, 1995 Detection of QTL for agronomic, yield, grain and disease characters in spring barley (*Hordeum vulgare* L.). Theor Appl Genet 91: 1037-1047.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, and D. Nieksen, 2001 Dwarf8 polymorphisms associate with variation in flowering time. Nature Genetics 28: 286-289.

Tigerstedt, P. M. A., 1997 *Adaptation in plant breeding.* Kluwer Academic Publishers, Dordrecht, The Netherlands.

Tinker, N. A., D. E. Mather, B. G. Rossnagel, K. J. Kasha, A. Kleinhofs *et al.*, 1996 Regions of the genome that affect agronomic performance in two-row barley. Crop Sci. 36: 1053-1062.

Toojinda, T., L.H.Broers, X.M.Chen, P.M.Hayes, A.Kleinhofs *et al.*, 2000 Mapping quantitative and qualitative disease resistance genes in a doubled haploid population of barley (*Hordeum vulgare* L.). Theor Appl Genet 101: 580-589.

Turuspekov, Y., Y. Mano, I. Honda, N. Kawada, Y. Watanabe and T. Komatsuda, 2004 Identification and mapping of cleistogamy genes in barley. Theor Appl Genet 109: 480-487.

Ungerer, M. C, S. S. Halldorsdottir, M. D. Purugganan and T. F. C. Mackay, 2003 Genotype-environment interactions at quantitative trait loci affecting inflorescence development in *Arabidopsis thaliana*. Genetics 165: 353-365.

Ooijen, J. W. van, and R. E. Voorrips, 2001 JoinMap 3.0, software for the calculation of genetic linkage maps. Plant Research International, Wageningen, the Netherlands.

VanderPlank, J. E., 1963 Plant Diseases: Epidemics and Control. Academic Press, New York.

VanderPlank, J. E., 1968 Disease Resistance in Plants. Academic Press, New York.

Via, S., R. Gomulkiewicz, G. De Jong, S.M. Scheiner, C.D. Schlichting and P.H. van Tienderen, 1995 Adaptive phenotypic

plasticity: consensus and controversy. Trends Ecol. Evol. 10: 212-217.

Virk, P. S., B. V. Ford-Lloyd, M. T. Jackson, H. S. Pooni, T. P. Clemeno *et al.*, 1996 Predicting quantitative variation within rice germplasm using molecular markers. Heredity 76: 296-304.

Voltas, J., F. A. van Eeuwijk, A. Sombrero, A. Lafarga, E. Igartua *et al.*, 1999a Integrating statistical and ecophysiological analyses of genotype by environment interaction fro grain filling of barley I. Individual grain weight. Field Crops Research 62: 63-74.

Voltas, J., F. A. van Eeuwijk, J. L. Araus and I. Romagosa, 1999b Integrating statistical and ecophysiological analyses of genotype by environment interaction fro grain filling of barley II. Grain growth. Field Crops Research 62: 75-84.

Voltas, J., I. Romagosa, S. E. Ullrich and F. A. van Eeuwijk, 2001 Identification of adaptive patterns in the 'Steptoe × Morex' barley mapping population integrating genetic, phenotypic and environmental information, pp. in *Proc.7th Quantitative Trait Locus Mapping and Marker-Assisted Selection Workshop*, Valencia, Spain.

Waugh, R., N. Bonar, E. Baird, B. Thomas, A. Graner *et al.*, 1997 Homology of AFLP products in three mapping populations of barley. Mol. Gen. Genet. 255: 311-321.

Weber, W. E., and G. Wricke, 1990 Genotype × environment interaction and its implication in plant breeding, pp. 1-19 in Genotype-by-environment interactions and plant breeding, edited by M. S. KANG. LSU, Baton Rouge, Louisiana.

Weil, C. F., 2002 Finding the crosswalks on DNA. PNAS 99: 5763-5765.

Weller, J. I., J. Z. Song, D. W. Heyen, H. A. Lewin and M. Ron, 1998 A new approach to the problem of multiple comparisonsin the genetic dissection of complex traits. Genetics 150: 1699–1706.

Wolfe, R. I., P. M. Hayes and L. Shugar, 1996 Multiple Dominant and Recessive Genetic Marker Stock Development. Barley Genetics Newsletter 25: 57-59.

Wu, R., C.-X. Ma and G. Casella, 2002 Joint linkage and linkage disequilibrium mapping of quantitative trait loci in natural populations. Genetics 160: 779-792.

Xiong, M., and L. Jin, 2000 Combined linkage and linkage disequilibrium mapping for genome screens. Genetic Epidemiology 19: 211-234.

Yao, H., Q. Zhou, J. Li, H. Smith, M. Yandeau *et al.*, 2002 Molecular characterization of meiotic recombination across the 140-kb multigenic a1-sh2 interval of maize. PNAS 99: 6157-6162.

Yates, F., and W. G. Cochran, 1938 The analysis of groups of experiments. J. Agric. Sci. 28: 556-580.

Yin, X., P. Stam, C. J. Dourleijn and M. J. Kropff, 1999 AFLP mapping of quantitative trait loci for yield-determining physiological characters in spring barley. Theor Appl Genet 99: 244-253.

Zhu, H., L. Gilchrist, P. Hayes, A. Kleinhofs, D. Kudrna, Z. Liu, L. Prom, B. Steffenson, T. Toojinda, **and H. Vivar**, **1999** Does function follow form? Principal QTLs for Fusarium head blight (FHB) resistance are coincident with QTLs for inflorescence traits and plant height in a doubled-haploid population of barley. Theor Appl Genet 99: 1221–1232.

Zhu, X., A. Luke, R. S. Cooper, T. Quertermous, C. Hanis *et al.*, 2005 Admixture mapping for hypertension loci with genome-scan markers. Nature Genetics 37: 177-181.

Zhu, Y. L., Q. J. Song, D. L. Hyten, C. P. van Tassell, L. K. Matukumalli *et al.*, 2003 Single-Nucleotide Polymorphisms in Soybean. Genetics 163: 1123-1134.

Zondervan, K. T., and L. R. Cardon, 2004 The complex interplay among factors that influence allelic association. Nature Reviews Genetics 5: 89-101.

Summary

Mapping of yield, yield stability, yield adaptability and other traits in barley using linkage disequilibrium mapping and linkage analysis

Objectives

Identification and mapping of Quantitative Trait Loci (QTLs) in plants is mostly done through linkage analysis. A segregating mapping population is created from a bi-parental cross and linkages between trait values and mapped markers reveal the positions of QTLs. In human genetics the most common approach is linkage disequilibrium (LD) mapping in a natural population. Observed LD can be based on linkage, but it is also affected by population dynamics. Application of LD mapping in plants can be valuable for many reasons, *e.g.* existing databases with phenotypic data and marker data could be utilized. We explored LD mapping in barley using a set of modern cultivars.

Yield as such is a complex trait, and the interaction of Genotype x Environment (GE) makes it even more complex to analyse. GE is assessed in multi-environment trials and analyzed by specific statistical models. Measures for yield adaptability (changes in mean response in relation to environmental factors) or yield stability (variation around the mean response) describe aspects of GE for which breeders may want to select.

The three main questions in this thesis were:

(1) What is the degree of linkage disequilibrium in modern barley cultivars?

(2) Can LD mapping be used to identify QTLs for a diverse set of traits?

(3) Can stability or adaptability measures be mapped by either LD or linkage analysis?

Results

(1) What is the degree of linkage disequilibrium in modern barley cultivars?

In chapter 2, a set of 148 modern spring barley cultivars from North-Western Europe was fingerprinted with AFLP markers. Associations between markers, expressed as r^2, demonstrated LD up to a distance of 10 cM distance which is large, compared to other species. The large distance might be induced by LD increasing factors such as inbreeding and the fact that the population is most likely based on a rather small set of founding genotypes.

(2) Can LD mapping be used to identify QTLs for a diverse set of traits?

In chapter 2 and 3, marker-trait associations were reported in the cultivar set for yield, yield stability, leaf rust resistance (LR), barley yellow dwarf virus resistance (BYD), plant height, and days to heading. A number of trait-associated markers were located in a region where already QTLs for the trait considered had been reported. This indicates that QTLs detected in a mapping population based on a bi-parental cross can be detected with LD mapping. Besides, trait-associated markers located in regions that had not been implicated before suggested new QTLs. Such new QTLs were found for yield, yield stability, LR and BYD. LD in germplasms is not only due to physical linkage in the genome, but also to other factors such as demographic aspects of a population. A number of these factors are discussed, *e.g.* population structure, population size, population admixture, and selection. Another issue in LD mapping is the determination of significance thresholds. The large numbers of traits and markers decrease the power of conventional multiple testing methods such as Bonferroni to an impractical level. We applied False Discovery Rates as an alternative method.

(3) Can stability or adaptability measures be mapped by either LD or linkage analysis?

Selection for yield is complicated by GE. Multi-environment trials are required to evaluate a genotype's performance. Statistical modeling can be applied to define measures for adaptability and stability. Adaptability is the responsiveness to the environment and stability is the unexplained deviation from the statistical model. LD mapping in cultivars resulted in marker-trait associations for yield stability (chapter 2). Linkage analysis with four doubled haploid populations resulted in many QTLs for adaptability, but only one for stability (chapter 4). Furthermore, coincidence of QTL locations between the DH- populations was high for yield and moderate for yield adaptability. Finally, QTLs for adaptability did not always coincide with QTLs for the mean yield.

Conclusions

(1) LD between molecular markers was found up to 10 cM in our germplasm of barley cultivars (chapter 2),

(2) LD mapping of traits is feasible as marker-trait associations were found for yield and yield stability, resistance to leaf rust and barley yellow dwarf virus, plant height and days to heading (chapters 2 and 3),

(3) QTLs for yield, yield stability and yield adaptability were mapped in four DH-populations, partly overlapping between the DH-populations and partly overlapping with QTLs mapped in cultivars (chapter 4).

To our belief LD mapping will become a valuable extension to conventional QTL analysis. This approach will be especially interesting for complex traits, where measurements are costly and time consuming, because existing databases can be utilized.

Samenvatting

Lokalisatie van loci voor opbrengst, opbrengststabiliteit, –adaptibiliteit en andere eigenschappen in gerst met linkage disequilibrium studies en genetische koppelingsanalyses.

Vraagstelling

Identificatie en kartering van Quantitative Trait Loci (QTLs) in planten gebeurt meestal met behulp van genetische koppelingsstudies. De koppeling tussen moleculaire merkers en eigenschapswaarden wordt vastgesteld in een segregerende populatie uit een kruising tussen twee contrasterende genotypen. In de humane genetica wordt meestal gebruik gemaakt van linkage disequilibrium (LD) studies in een natuurlijke populatie. LD kan dan gebaseerd zijn op koppeling, maar het wordt ook beïnvloed door populatiedynamica. De toepassing van LD-studies in planten zou waardevol kunnen zijn, bijvoorbeeld door gebruik te maken van bestaande databases met fenotypische informatie en merkerdata. Wij hebben LD bestudeerd in een set moderne gerstrassen.

Opbrengst is een complexe eigenschap, en door Genotype x Milieu interactie (GMI) wordt het nog complexer. GMI wordt geanalyseerd met proeven in diverse milieus en speciale statistische modellen. Uit deze modellen kunnen maten worden afgeleid voor opbrengst-adaptabiliteit (veranderingen in gemiddelde opbrengst door milieufactoren) of opbrengst-stabiliteit (variatie rond deze gemiddelde opbrengst). Adaptibiliteit en stabiliteit beschrijven verschillende aspecten van GMI waarvoor veredelaars mogelijk willen selecteren.

De drie hoofdvragen van dit onderzoek waren:

(1) Hoeveel LD is er aanwezig in moderne gerstrassen?
(2) Kunnen met LD-studies QTLs geïdentificeerd worden?
(3) Kunnen QTLs voor maten voor opbrengstadaptibiliteit of –stabiliteit worden gekarteerd met behulp van LD-studies of genetische koppelingsstudies?

Resultaten

(1) Hoeveel LD is er aanwezig in moderne gerstrassen?

In hoofdstuk 2 werd een set van 148 moderne gerstrassen uit noord-west Europa genetisch gekarakteriseerd met AFLP-merkers. Associaties tussen merkers onderling, uitgedrukt in r^2,

toonden aan dat er nog LD was tot 10 cM afstand. Die afstand is groot in vergelijking met andere soorten. Deze grote afstand zou veroorzaakt kunnen worden door LD stimulerende factoren zoals zelfbevruchting en het feit dat deze gerstrassen waarschijnlijk afkomstig zijn van een kleine set oorspronkelijke genotypen.

(2) Kunnen met LD-studies QTLs geïdentificeerd worden?

In hoofdstuk 2 en 3 werden associaties tussen merker– en eigenschapswaarden beschreven voor de gerstrassen voor opbrengst, opbrengststabiliteit, dwergroest-resistentie, gerstevergelingsvirus-resistentie (BYD), planthoogte en vroegheid. Een aantal van deze associaties lagen in een regio waarvoor reeds eerder QTLs waren gerapporteerd. Dit duidt erop dat QTLs die in een genetische koppelingsstudie zijn gevonden ook met behulp van LD-studies kunnen worden geïdentificeerd. Ook werden merker-eigenschap associaties gevonden voor opbrengst, opbrengststabiliteit, dwergroestresistentie en BYD-resistentie in regio's waarvoor nog niet eerder QTLs waren gevonden. Dit duidt op de ontdekking van nieuwe QTLs.
LD in een natuurlijke populatie is niet alleen een gevolg van fysieke koppeling, maar het wordt ook beïnvloed door andere factoren. De meeste van deze factoren betreffen demografische aspecten van een populatie. Voorbeelden hiervan werden besproken, waaronder populatiestructuur, populatiegrootte, populatiemenging en selectie. Een ander vraagstuk in LD-studies is de vaststelling van significantie van associaties. Standaardmethoden voor multipele tests, zoals Bonferroni, blijken onwerkbaar doordat het statistische onderscheidingsvermogen teveel daalt door de grote hoeveelheid eigenschappen en merkers. Wij hebben daarom een alternatieve methode toegepast die gebruik maakt van *False Discovery Rates*.

(3) Kunnen QTLs voor maten voor opbrengstadaptibiliteit of –stabiliteit worden gekarteerd met behulp van LD-studies of genetische koppelingsstudies?

Selectie voor opbrengst wordt bemoeilijkt door GMI. Om de prestatie van een genotype te kunnen beoordelen zijn proeven in diverse milieus nodig. Met statistische modellering werden maten voor adaptibiliteit en stabiliteit gedefinieerd. Adaptibiliteit is de respons op milieufactoren en stabiliteit is de resterende onverklaarde afwijking van het statistische mdel. LD-studies met gerstrassen toonde associatie aan tussen opbrengststabiliteit en merkers (hoofdstuk 2). Genetische koppelingsstudies met vier populaties van verdubbelde haploiden (DH), afkomstig uit een F1, resulteerden in de identificatie van vele QTLs voor adaptibiliteit en één voor stabiliteit (hoofdstuk 4). De overeenkomst in kaartpositie tussen verschillende DH-

populaties was hoog voor opbrengst-QTLs en redelijk hoog voor adaptibiliteit-QTLs. Opmerkelijk was dat adaptibiliteits-QTLs niet altijd samenvielen met opbrengst-QTLs.

Conclusies

(1) LD tussen merkers werd gevonden tot op 10 cM in deze gerstrassen (hoofdstuk 2),

(2) Identificatie van QTLs met LD-studies is mogelijk, want er werden geassocieerde merkers gevonden voor opbrengst, opbrengststabiliteit, dwergroestresistentie, BYD-resistentie, planthoogte en vroegheid (hoofdstuk 2 en 3),

(3) QTLs voor opbrengst, opbrengstadaptibiliteit en –stabiliteit werden geïdentificeerd in vier DH-populaties op genoomposities die vaak overeenkwamen tussen DH-populaties en soms overeenkwamen tussen DH-populaties en de gerstrassen (hoofdstuk 2 en 4).

LD-studies zullen een waardevolle aanvulling vormen op de bestaande methodieken voor QTL-analyse. Dit geldt met name voor complexe eigenschappen, waar metingen kostbaar en tijdrovend zijn, omdat bestaande data van bijvoorbeeld rassenproeven kunnen worden benut.

Nawoord

Na een lange periode verbonden geweest te zijn aan Vakgroep Plantenveredeling waarbinnen ik verschillende onderwerpen heb bestudeerd is het aantal mensen aan wie ik dank verschuldigd ben groot. De kans dat ik iemand vergeet is dan ook aanzienlijk, dus ik begin maar vast met een welgemeend dank-je-wel voor iedereen die ik vergeten ben.

Mijn eerste aanstelling bij Plantenveredeling was op het gebied van de toepassing van lerende computertechnieken in de veredeling van tuinbouwgewassen. In eerste instantie werkte ik samen met Professor Jan Parlevliet en Ies Bos, maar de plaats van Jan Parlevliet werd spoedig overgenomen door Piet Stam. Samen met Piet en Ies probeerde ik het wetenschappelijke deel van mijn project vorm te geven. Piet, ik kon altijd bouwen op jouw rust en puur onafhankelijke denkwijze. Indien ik verstrikt raakte in een kluwen van gedachten luisterde jij geduldig om na enige tijd op te merken dat ik er misschien ook op geheel andere wijze tegen aan kon kijken. Ies, onze verbintenis gaat al verder terug, namelijk toen ik nog student was en jij mijn begeleider bij een afstudeervak. Bijzonder prettig waren voor mij jouw enthousiasme voor mijn ideeen en plannen. Je eerlijkheid en persoonlijke interesse maken jou voor mij uniek. Gedurende dit onderzoek heb ik samen met Leo van Eijk van Zaadunie (tegenwoordig Syngenta) software gemaakt voor veredelaars. Leo, jouw praktische kijk, snelle manier van denken, gastvrijheid en openheid zal ik niet licht vergeten.
Mijn tweede aanstelling bij Plantenveredeling was op het gebied van kartering in gerst en heeft geleid tot dit proefschrift. Ook deze keer was Piet Stam erbij betrokken. De grote kracht en schrijver van het project was Johan Dourleijn. Johan, jij bent een prettig mens om mee samen te werken, niet alleen vanwege je inhoudelijke kennis, maar ook vanwege de rust en het respect dat je uitstraalt. Ik was teleurgesteld toen je wegging naar Advanta BV om daar Vestigingsmanager te worden. Gelukkig bleef je betrokken als lid van de STW-gebruikerscommissie. Johan's plek als projectleider werd ingenomen door Fred van Eeuwijk. Fred, jouw kennis op het gebied van statistiek is zeer groot. Ik moest soms wennen aan jouw manier van communiceren, maar eenmaal gewend kreeg ik steeds meer oog en waardering voor jouw bijzondere kwaliteiten als statisticus en als netwerker in de wereld van de rekenmeesters uit de plantenveredeling. In de laatste fase is de plaats van Piet Stam als promotor overgenomen door Richard Visser. Richard, onze betrokkenheid is van korte duur geweest en het onderwerp ligt een redelijk eind buiten jouw kernexpertise, maar ik ben je

dankbaar voor de ideeen die je hebt aangedragen, je inzet om teksten snel te lezen, en de concrete wijze waarop je samen met mij hebt toegewerkt naar de voltooing van mijn promotie.

Naast begeleiders en (co-)promotoren zijn er nog vele anderen betrokken geweest bij mijn onderzoek. Rients Niks, mijn onofficiële co-promotor, heeft een grote bijdrage gehad. Rients, het was prettig hoe ik kon bouwen op jouw geestdrift, nieuwsgierigheid, kennis van gerst en vele graanziekten. Gesprekken met jou zijn een bevestiging van de stelling dat ieder antwoord leidt tot een vermenigvuldiging van het aantal vragen dat je hebt. I also want to thank Marcos Malosetti en Fernando Martinez. Both of you have contributed to this research in your own specific way, and I have experienced both of you as very pleasant persons to be with. Muchas gracias! Van alle collegae en kamergenoten die ik heb gehad wil ik speciaal Ralph van Berloo en Johan Schut noemen. Jullie boden mij je vermogen tot meedenken en discussieren, maar ook vriendschap en een luisterend oor. Oliver Kiplagat was my officemate in the last year. Thanks for your friendship. You motivated me daily by being ahead one step in the proces of finishing your thesis. I look forward to visiting you in Kenya in April! Herman van Eck is een collega, buurman en vriend ineen. Je bent een bijzonder persoon met een geheel eigen denkwijze en een voor mij prikkelend pakket aan meningen die vaak leidt tot lange en inspirerende discussies over werk en andere zaken.

Praktische zaken rond mijn onderzoek waren nooit tot een goed einde gebracht zonder de hulp van Unifarm op het veld en in de kas, de hulp van TUPEA op computergebied, en de hulp van het AFLP-lab. Van Unifarm wil ik speciaal Anton Vels bedanken voor zijn inzet. Van TUPEA wil ik Hans Romberg niet ongenoemd laten. We hebben lang met elkaar te maken gehad en ik heb het contact altijd kunnen waarderen. Na het noodlottige ‘misverstand’ in 2004, dat mij al mijn data kostte, denk ik dat we beiden geleerd hebben op zowel technisch als relationeel gebied. Het AFLP-werk was nooit gelukt zonder Petra, Fien en Pim. Uiteraard kwam ik altijd ongelegen en moest het gisteren klaar, maar desondanks redden jullie het telkens weer. Dank!

Een aantal studenten hebben me geholpen: Hans van Os, Joel Schwartz, Bakhyt Mussiraliev, Fabien Debellis en Stefanos Kalogiannis. Bedankt, bedankt, rahmet, merci en efharisto! En Hans, het was spannend tot het laatst, maar uiteindelijk promoveer ik toch nog 6 dagen eerder!

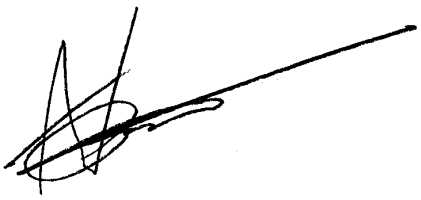
Sergio Ceretta supervised the very important multiplication of the Henni × Meltan population in Uruguay. Thanks even more for your three months stay in The Netherlands in 2004. You inspired me to arrange a period for myself to work full time on my thesis. And it worked.

De mensen van STW en de STW-gebruikerscommissie wil bedanken voor hun meedenken en Advanta Seeds BV voor hun belangrijke bijdrage in het maken van de Henni × Meltan verdubbelde haploiden.

Jos Stuyt heeft op bijzondere wijze het afronden van dit boekje mogelijk gemaakt. Daarnaast kwam ik je regelmatig tegen in het bos bij hardloopwedstrijden. Dit jaar liet je me je hielen zien, maar ik hoop dat je me volgend jaar gelegenheid geeft tot revanche.

Tot slot bedank ik uit de grond van mijn hart mijn ouders, Rikje en mijn kinderen. Pa en ma, bedankt voor jullie onvoorwaardelijke liefde en aandacht. Jullie hebben mij veel gegeven. Rikje, jij hebt het proces het meest van dichtbij meegemaakt. Ik dank je voor je geduld en betrokkenheid en de ruimte die je me gegeven hebt om me volledig te kunnen bezig houden met het afronden van het proefschrift. Laten we dit vieren in Afrika! Rianne, Wouter en Mirjam, mama's boekje was al af en papa was er nog steeds mee bezig. Nu is het klaar en ligt het voor jullie. Vinden jullie het mooi? Ik wil jullie bedanken voor de afleiding die jullie me gaven en de wijsheid die jullie uitdragen. Kunnen jullie nog helpen de uitnodigingen erin doen en ze in enveloppen stoppen?

Wageningen, 5 april 2005



Curriculum vitae

Arnoldus Theodorus Willebrordus Kraakman werd op 16 maart 1967 geboren in Noord-Scharwoude (gemeente Langedijk, NH). Na met goed resultaat de R.K. Mariaschool te hebben doorlopen mocht hij naar het Han Fortmancollege te Heerhugowaard. In 1985 haalde hij zijn VWO-diploma en vertrok hij naar de Landbouwuniversiteit te Wageningen. In 1991 ontving hij zijn ingenieurstitel in Plantenveredeling met als specialisatie het kwekersprofiel. Speciale aandacht gaf hij aan selectiemethodieken, informatica en plantenfysiologie. Hij deed een stage bij Fides (chrysant- en Kalanchoeveredeling) in De Lier en gedurende een groot, gecombineerd Plantenveredeling/Informatica afstudeervak bij Novartis Seeds in Enkhuizen (destijds Zaadunie).

Zijn eerste onderzoeksproject richtte zich op de toepassing van lerende computerprogramma's in de veredeling van tuinbouwgewassen. Met behulp van de eigenschappen van ouderlijnen werden de waarde van F1-hybriden voorspeld. Hij werkte samen met Novartis Seeds en inventariseerde de wensen van veredelaars, ontwierp een lerend expertsysteem, begeleidde de programmeur en testte het resultaat. Voor zijn tweede onderzoeksproject stapte hij over op gerst. Moleculaire merkers werden gebruikt voor kartering van loci voor diverse eigenschappen met zowel mappingpopulaties op basis van twee ouders, als ook met een set moderne gerstrassen. Dit werk wordt beschreven in dit proefschrift.
Naast zijn werk als wetenschappelijk onderzoeker is Arnold sinds een aantal jaren actief als loopbaancoach. Hij begeleidt mensen bij het verbeteren van hun motivatie en werkplezier in hun huidige baan, of helpt ze door het inventariseren van hun talenten en persoonskenmerken om te kiezen voor heel ander werk.

Arnold is gehuwd met Rikje van de Weerd en vader van Rianne, Wouter en Mirjam.

This PhD thesis contains the results of research conducted at the Laboratory of Plant Breeding of Wageningen University and Research Centre in The Netherlands.

The doubled haploids of the Henni × Meltan cross were made by Advanta Seeds BV.

The field trials of chapter 4 in Spain and Scotland were included in the trials of the EU-Incomed project Mapping Adaptation of Barley to Drought Environments (ICA3-2002-10073).

This research was supported by the Technology Foundation STW (project WBI 4785), applied science division of NWO and the technology programme of the Ministry of Economic Affairs of The Netherlands.

The thesis was printed by Ponsen & Looijen, Wageningen.