UvA-DARE (Digital Academic Repository)

From phenotype via QTL to virtual phenotype in *Microseris* (Asteraceae): predictions from multilocus marker genotypes

Bachmann, K.; Hombergen, J.E.M.

DOI

[10.1046/j.1469-8137.1997.00824.x](https://doi.org/10.1046/j.1469-8137.1997.00824.x)

Publication date

1997

Published in

New Phytologist

[Link to publication](#)

Citation for published version (APA):

Bachmann, K., & Hombergen, J. E. M. (1997). From phenotype via QTL to virtual phenotype in *Microseris* (Asteraceae): predictions from multilocus marker genotypes. *New Phytologist*, 137, 9-18. <https://doi.org/10.1046/j.1469-8137.1997.00824.x>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

From phenotype via QTL to virtual phenotype in *Microseris* (Asteraceae): predictions from multilocus marker genotypes

BY KONRAD BACHMANN¹* AND ERIK-JAN HOMBERGEN²

¹*Institut für Pflanzengenetik und Kulturpflanzenforschung IPK, Corrensstr. 3, D-06466 Gatersleben, Germany*

²*Hugo de Vries Laboratory, University of Amsterdam, Kruislaan 318, NL-1098 SM Amsterdam, The Netherlands*

(Received 3 April 1997; accepted 8 July 1997)

SUMMARY

Microseris douglasii (DC.) Sch.-Bip. and *M. bigelovii* (Gray) Sch.-Bip. are two small annual autogamous species of Compositae with nearly non-overlapping distribution ranges in Western North America. Specifically, *M. bigelovii* occurs directly along the Pacific coast, whilst *M. douglasii* has an inland distribution including patches of serpentine soil. Both species are variable, and artificial hybrids between them vary widely in fertility depending on the individual parents. Segregating offspring of one hybrid (strain H27) is being used to analyse the genetic basis of characters differentiating the two species by QTL mapping with RAPDs and ALFPs as molecular markers. Technical problems with mapping dominant markers in a wide cross will be briefly listed and QTL analysis will be discussed. For the genetic analysis of physiological characters, the precise definition of the characters is crucial and the methods of scoring or measuring phenotypes in different environments eventually require more time and effort than the molecular characterization. We are establishing recombinant inbred lines to provide material for more complex physiological analyses requiring several plants per genotype. An increasing number of characters is being studied in this cross, and the possibility of shared pleiotropic QTLs is high. The potential number of QTL genotypes by far exceeds the number of actual genotypes in these lines. We are characterizing the gene interactions as closely as possible and making quantitative genetic models to predict the genotypes corresponding to all possible genotypes. These predictions are converted via computer modelling into an increasingly realistic three-dimensional representation of the growing plant useful for a simulation of plant evolution.

Key words: multigene characters, phenotype prediction, AFLP, RAPD, QTL.

INTRODUCTION

The availability of single-locus molecular polymorphisms from all across the genome has revolutionized the genetics of quantitative characters. By studying the co-segregation of single-locus markers and quantitative traits, we can identify individual loci affecting quantitative traits (quantitative trait loci, QTLs; Paterson *et al.*, 1988; Lander & Botstein, 1989). When a complete genetic map of molecular markers is available, we can expect to find all QTLs having an appreciable effect on the trait. Such a powerful method has its price in the

costs of materials and the amount of work needed for a complete analysis, and QTL analysis has been primarily applied to commercially important crop species (Phillips & Vasil, 1994). Recently, PCR-based methods for the detection of markers have been introduced that work with arbitrary primers and therefore require no previous sequence knowledge. Such poorly characterized markers with mainly dominant inheritance have their drawbacks, but they permit a relatively inexpensive and fast genetical examination, and selected informative markers can easily be investigated in more detail. With these methods, primarily random amplified DNA polymorphisms (RAPDs: Welsh & McClelland, 1990; Williams *et al.*, 1990) and amplified fragment

* To whom correspondence should be addressed.

length polymorphisms (AFLPs: Vos *et al.*, 1995), QTL analysis can be performed efficiently on practically any segregating population. Increasingly, the method is being used to explore the genetics of morphological and physiological variation in natural populations (Van Houten, Van Raamsdonk & Bachmann, 1994; Bradshaw *et al.*, 1995; Bachmann & Hombergen, 1996).

Here, we want to report on an ongoing project in which the genetics of phenotypic variation in the annual species of *Microseris* is explored. The distribution and association of character states for a wide range of characters has been determined in representative populations of the three diploid annuals of the genus in California, *M. douglasii* (DC.) Sch.-Bip. (Bachmann & Battjes, 1994), *M. bigelovii* (Gray) Sch.-Bip. (Bachmann, Chambers & Price, 1984; Bachmann *et al.*, 1987; Bachmann, 1992; Van Heusden & Bachmann, 1992*b*) and *M. elegans* Greene ex Gray (Bachmann & Van Heusden, 1992; Van Heusden & Bachmann, 1992*a*), and of the closely related Chilean species, *M. pygmaea* D. Don. (Bachmann *et al.*, 1985; Van Heusden & Bachmann, 1992*c*). With minor variations for each species, a general picture of the structure and distribution of intraspecific phenotypic variation has emerged that was unexpected but might be characteristic for highly autogamous annuals distributed in isolated populations. In these populations, plants in the field are frequently completely homozygous for all morphological and molecular markers (Roelofs & Bachmann, 1995, 1997), and populations might consist of one or more inbred lines growing side by side. Some local populations are derived from single plants after achene dispersal, sometimes over hundreds of km distant from the source population (Van Heusden & Bachmann 1992*a, b*), other sites have been colonized by more than one specimen. Outbreeding is rare, but occasional crosses produce a local burst of recombination much of which remains preserved as a series of natural recombinant inbred lines (Roelofs & Bachmann, 1995, 1997). The fact that these events can be reconstructed from the existing population structure suggests that selection among genotypes must be very weak. One explanation for this is the evident phenotypic plasticity of the plants for quantitative characters with a potential adaptive value (Bachmann & Battjes, 1994). It seems that the overall phenotype of most of the annual *Microseris* in nature is depauperate as compared to the genetically determined maximum luxuriance that can be expressed under favourable glasshouse conditions (Battjes & Bachmann, 1994). The adaptive strategy seems to be aimed at assuring the production of good quality seed under suboptimal conditions rather than at maximizing seed production. Quantitative phenotypic variation in nature therefore is a very sensitive reflection of the differences in the environmental conditions among

the individual spots in which the plants have grown rather than of the genetic differences between the plants. Very extensive genetic variability can be demonstrated in the glasshouse that is swamped by plastic responses to the environment in nature and is therefore inaccessible to natural selection under the normal growing conditions. On the other hand, the patchy distribution of all species in isolated islands of suitable habitat (Chambers, 1955) suggests that there are a few, probably physiological, factors that are under strong selection and limit the number of sites where the species can become established. The same must be claimed for the factors separating the distribution area of *M. bigelovii* from that of the other species that not uncommonly occur in mixed populations. *M. bigelovii* occurs essentially in a very narrow strip of land along the immediate Pacific coast of California, Oregon and Vancouver Island.

The demonstration of heritable variation within and among the species for virtually every character that we have studied and the ease with which segregating offspring can be maintained in the form of inbreeding lines derived from single F2 individuals has recommended the plants for a more detailed genetic investigation, and *Microseris pygmaea* has probably been the first 'wild' species that has been investigated by QTL mapping (Van Houten *et al.*, 1994). Here, we report on an interspecific cross between *M. douglasii* and *M. bigelovii* in which a great many segregating phenotypic characters could be scored and analysed easily in the F2 (Bachmann & Hombergen, 1996). For technical reasons, the paper will deal mainly with characters that can be scored in individual plants. Projects on more complex characters that we are initiating will be described briefly at the end. In the framework of this symposium, we hope to show that the genetic investigation of any heritable character in a segregating population is becoming a manageable task, we want to indicate some difficulties that are likely to be encountered with the approach, and we want to show how we begin to reassemble the abundant harvest of data from the reductionist genetical approach into an integrated model of the genetic determination of the phenotype. Such a model summarizes the results efficiently, clearly indicates the explanatory value of the data, points out where essential information is missing, and will, it is hoped, someday incorporate mechanisms of developmental genetics in sufficient detail to predict the emergent properties of complex gene interactions under the influence of environmental constraints.

MATERIALS AND METHODS

The plants

Hybrid strain H27 is a cross between *Microseris douglasii* strain B14 collected near Parkfield,

California and *M. bigelovii* C94 from Uplands Park, Victoria B.C., Canada. It was selected from several crosses between the parental species since it combines a maximum of character differences with an F1 fertility of *c.* 10%, high enough to produce a sufficiently large F2. Details on this hybrid and its parents have been published elsewhere (Hombergen & Bachmann, 1995; Bachmann & Hombergen, 1996). All plants are derived by spontaneous selfing from one F1 individual. One hundred and six F2 plants were raised, thoroughly characterized phenotypically and with RAPDs.

Mapping

The preparation of a molecular marker map based on 197 RAPD markers and its use for detecting and mapping QTLs has been described by Bachmann & Hombergen (1996). The phenotypic data used for QTL determination using the RAPD map were combined from F2 data and F3 offspring averages.

About 450 plants each of the F3, F4 and F5 generations were raised in 1994, 1995 and 1996, and the same set of phenotypic characters was determined for all plants in each year. DNA was isolated from leaf fragments of 150 F5 plants for AFLP determination. From the data a new molecular marker map, this time based on AFLPs rather than RAPDs, was calculated, and QTLs were determined using the phenotypic values for the 150 F5 plants. These 150 plants are offspring of 143 F4 plants, 129 F3 plants, and 82 of the original 106 F2 plants. Of the 24 F2 lines lost to the F5, 16 are the consequence of completely sterile F2 plants.

Details on the RAPD reactions have been given by Bachmann & Hombergen (1996). For the AFLP reactions, the kit of Gibco-BRL Life Technologies was used and the protocol supplied by the manufacturer was followed. The primers used were *EcoRI* with ACA, ACG and ACC extensions and *MseI* with CAA, CAC, CAG, CAT, CTC, CTG, CTT extensions. The *EcoRI* primers were end-labelled with [γ -³²P]ATP according to the protocol.

The mapping program JOINMAP (Stam, 1993) has been used for all recent linkage maps. Several approaches have been used to test the reliability of the incomplete RAPD and AFLP maps. These include varying the statistical stringency and processing markers derived from each of the two parental strains separately.

RESULTS

The marker maps

A linkage map for the F2 including 197 markers with 3:1 segregation in 17 linkage groups with three or more markers and covering 1068 cM (Kosambi mapping function, Kosambi (1994); LOD 3.7) has

been published (Bachmann & Hombergen, 1996). Recalculations including more markers have not produced a saturated map but have demonstrated the essential stability of the linkage groups. Both parental species have $2n = 18$ chromosomes.

A total of 441 AFLP bands were scored in the F5 of which 271 segregated with the 17:15 segregation ratio expected for fourth generation inbred derivatives of an F1 individual. A linkage map was calculated with JOINMAP (Stam, 1993) using the segregations of these 271 markers at an LOD value of 6.1. The map contains 13 linkage groups with more than three linked markers and includes 258 of the 271 markers. The total map length is 815.9 cM.

Both maps have similar features: with an increasing statistical stringency (higher LOD score), originally very long linkage groups fall apart, but certain groups of markers remain stably linked even at very high LOD scores. Each map contains three linkage groups with many markers and others with increasingly fewer markers. In both cases, there is a relatively large number of unlinked markers. Some of these features are undoubtedly the result of the limited size of the mapping populations and of the dominance of the markers. It is easy to see that statistical fluctuations in the segregation will prevent the precise determination of the linear order of the large number of closely linked markers. We believe, however, that the low F1 fertility of 10% and the persistently low fertility of some inbred offspring lines are due to small chromosome rearrangements between the parental genomes that interfere with normal recombination and thereby with the construction of a saturated marker map of the hybrid.

The colinearity of the two maps will be formally studied by mapping key RAPD markers from the F2 map in the F5 population. Preliminary examination, especially the position of characteristic QTLs shows that both maps give about the same results and that the AFLP map of the F5 might be somewhat more stable than the RAPD map of the F2. This could be due to the slightly larger mapping population, but especially to the increased homozygosity and nearly 1:1 distribution of plants with or without the dominant marker.

Phenotypic characters and QTLs

The phenotypic characters that we have studied up to now are mainly those that can be determined in individual plants. The heritability of the character differences has been demonstrated by parent/offspring regression, and the genetic component of the quantitative variation has been enhanced for mapping by projecting the F3/F2 data points for each character on the F3/F2 regression line and using the distance along the regression line as a normalized quantitative value for the character (Bachmann & Hombergen, 1996).

Figure 1. Predicting phenotypes from marker genotypes: estimated time of the appearance of the first capitulum bud (d after germination) based on the linear regression against the number of alleles (0, 1 or 2) from the 'late' parent *Microseris bigelovii* C94 at four QTLs. (The regression predicts 40.3% of the phenotypic variation; $y = 87.1 + 6.4$ d per allele; $r^2 = 0.403$).

QTLs can be determined by standard mapping programs. However, these programs depend on the assumption that the marker map is exact. When the map order of the markers is not perfect, and especially when occasional unlinked markers are included in a linkage group, this can influence the results of QTL mapping. We have therefore applied a very simple procedure to check the reliability of the map for QTL determinations. By introducing the phenotypic values for each character together with all marker values for each plant in a spread sheet, the plants and their marker values can easily be sorted according to the phenotypic value of a character and simple statistical tests can be applied to detect markers showing a significant association with the quantitative character. Checking the locations of these markers on the map reveals QTLs as clusters of markers associated with a character even in the presence of some statistical noise in the map positions.

In practice, in spite of these problems, it has been easy to detect between one and six QTLs for each character that we have analysed. Most clusters of markers associated with a character contained markers from both parents, so that all three genotypes for the QTL, including heterozygotes, could be identified by combining dominant markers of opposite polarity (Bachmann & Hombergen, 1996). Predictions of QTL genotypes from marker phenotypes are always estimates based on several linked markers. Some QTLs are in stretches of markers with a single polarity, and for some characters we found closely associated single markers that could not be placed on the marker map. The influence of the QTLs on the phenotypic values was tested by a regression of the character value of each plant on the estimated number (0, 1, or 2) of positively acting alleles in that plant and determining the coefficient of determination, r^2 , i.e. the per-

centage of the statistical variance in the sample that is explained by the linear regression. This test provides crucial information: it indicates if the QTL acts additively or with dominance, and it reveals the basic structure of the genetic system affecting the character. This structure might consist of one major gene and several modifiers with considerably smaller influence, or there may be two, three or more interacting genes with about equal effects with or without detectable modifiers. We have found that, at least in practice, the difference between qualitatively acting loci and QTLs lies in the way the gene is expressed in the phenotype, and consequently in the method of detection. The assumption that quantitative characters are generally determined by many genes with individually small effects is certainly not true. The individual effects of such 'polygenes' should be below the level of detection even for quantitative characters with a high heritability coefficient. The observation that with the methods used here we have found QTLs for virtually every character that we have studied and that the cumulative effects of these QTLs explain a major proportion of the heritable component of the phenotypic variability in the segregating population shows that most of the character differences between the parental strains can be explained by the effects of a few genes.

Recognizing QTL genotypes by marker states

The conversion of marker genotypes into predicted QTL genotypes should make these independent of the type of marker used and allow a direct comparison of the results obtained with the F2 and F5 data. Since this includes the cumulative errors due to recombination between the combined markers and between markers and QTLs, the identity of QTLs obtained in the two generations with different marker sets can be determined with high statistical significance only for strong QTLs in marker-rich surroundings. The constant ('canalized') number of five paleaceous pappus parts in most annual *Microseris* is maintained by a dominant gene. In plants homozygous for a recessive allele of this gene, the number of pappus parts on an achene can be reduced to zero, it becomes variable among the achenes on one plant, and this variability is influenced both by modifier genes and by the environment (Vlot *et al.*, 1992). The major gene has been mapped close to two RAPD markers with opposite polarity in hybrid H27 (OPA-3.01 from C94 and OPA-2.06 from B14, map distance 3.2 cM; Bachmann & Hombergen, 1996). With these two markers, the genotypes for the major pappus part gene in the F2 plants have been predicted. Independent mapping of this character with AFLPs in the F5 has also revealed closely linked markers of opposite polarity for the major QTL (AGC/CTG13

from C94 and ACA/CTT56 from B14, map distance 1.2 cM), and the genotypes for this QTL have been determined for the F5 plants. Comparative data were obtained for 79 F2 plants and their F5 offspring. In the F2 plants, the QTL segregated 14 (B14/B14):45 (B14/C94):20 (C94/C94). All F5 offspring of the 14 (B14/B14) F2 homozygotes were similarly homozygous for the major QTL found with AFLPs. Of the 20 F2 plants homozygous for the genotype (C94/C94), 15 had homozygous offspring, five F5 offspring plants were scored as heterozygous. It is quite obvious that the same gene has been mapped and marked in both experiments. All discrepancies can be due to scoring mistakes resulting from recombination among the markers. As a result, homozygotes are misclassified as heterozygotes. False recognition in the F2 has increased the frequency of estimated heterozygotes to 57%, and false recognition of heterozygotes in the F5 has scored offspring of five homozygotes as heterozygous.

Predicting quantitative phenotypes, transgressive segregation

The regression of the quantitative value of a character against the number of alleles increasing the character value is a basic test for the interactions among the multiple QTLs affecting one character. Normally, the predictive value, r^2 , increases steadily when additional QTLs are used. In a few cases, one or another smaller QTL adds to the predictive value, but both together do not; occasionally an individually weak QTL does not add to the cumulative predictive value of the stronger QTLs or even decreases it. These regressions assume equal additive values for the various QTLs. Where considerable differences are expected, a multiple regression of the quantitative value against the various genotypes might improve the predictive value by assigning each QTL an individual strength. This also permits a test of the statistical significance of the contribution of each QTL to the multigene prediction. The result of this analysis is a genetic model listing the QTLs affecting a character and their interactions, primarily the direction, magnitude and possible dominance of the effect of an allele. Figure 1 shows the linear regression for four QTLs affecting the appearance of the first capitulum bud (i.e. flower induction) detected in the F5 with AFLP markers. All four loci have the same polarity: alleles from the late flowering parent, *M. bigelovii* C94, delay the appearance of the first bud. There is considerable phenotypic variation for each multigene-genotype predicted from the markers. This variation has two sources: (1) wrong genotype identification due to crossing over between the marker AFLPs and the QTL, and (2) plastic phenotypic variation. However, 40% of the total phenotypic variation is explained by the additive

model, and this corresponds roughly to the heritable component of the variation so that even the cumulative mistakes in the prediction of genotypes for four loci cannot be very high.

Transgressive segregations, i.e. hybrid offspring with character values above and below those of the parental strains, are not rare. They can be related to QTLs in which the allele from the parent with the higher quantitative value reduces the character value. An example is the leaf length difference between the parents. Leaves of *M. douglasii* B14 are longer than those of *M. bigelovii* C94. Of controls raised together with the F2, B14 had an average leaf length of 112.8 ± 8.28 mm (average of independent determinations on days 71 and 91 after germination; SD), whilst C94 had leaves of average length 105.2 ± 2.93 mm. The F2 plants raised in the same glasshouse had an average leaf length of 113.63 ± 23.78 mm with a range 33–160 mm. The average leaf length of the hybrid plants exceeded that of both parents in the F2, but was lower than that of the low C94 parent in all subsequent generations. In the F5, the following values were determined for leaf length on day 50 after germination; B14, 130.5 mm; C94, 120.8 mm (both much longer at an earlier date than in the F2); F5 hybrids, 94.0 mm (range 13–182 mm). We have observed similar effects that correspond to heterosis and inbreeding depression for several characters in these plants even though the natural parents are completely homozygous inbred lines. The wide range of F2 values is a true transgressive segregation, since the regression of the averages of five F3 offspring plants on their F2 parent values indicates a high narrow-sense heritability ($y = 0.8024 (\pm 0.0144)x$; $r^2 = 0.5092$; intercept n.s.). The two strongest QTLs detected with RAPDs in the F2 and with AFLPs in the F5 both increase leaf length with the alleles from B14, but the third QTL has the opposite polarity and explains part of the transgressive segregation. Altogether, the marker genotypes predict only 18% of the phenotypic variation in the F2 and 20% in the F5. This leaves room for more QTLs affecting leaf length with individual contributions below the statistical level of detection.

Pleiotropy

With a large number of characters, chances increase that pleiotropic QTLs are detected that participate in the genetic determination of several characters. Potential pleiotropic QTLs cannot definitively be differentiated from closely linked loci, but are very likely when functionally related characters are affected by the same major QTL. As an example, QTLs are found at essentially the same locations when the data of appearance of the first bud is mapped, when the day of anthesis of the first

capitulum is analysed or when the date at which the first mature capitulum opens is studied. The predictive values are smaller for the later dates than for the earlier one, and several tests have suggested that, once a capitulum has been initiated, the subsequent timing of its growth and differentiation is influenced by the environment without further genetic control. Our statement that we have found QTLs for practically any character that we have analysed therefore needs to be qualified by the observation that not all of these characters are genetically independent.

More interesting are characters that share pleiotropic genes but also have specific modifiers of their own. There is a general dependence for several quantitative characters of the capitulum, for instance the number of florets (and consequently achenes) per head, the number of involucre bracts and the total diameter (mm) of the spread-out involucre. The major QTL for all three of these characters maps to the same position (Bachmann & Hombergen, 1996). Less stringent mapping has revealed up to six QTLs affecting the number of achenes per head, and QTLs are always in the same position for the number of involucre bracts. When a regression of the number of involucre bracts against the number of achenes per head is performed, and the new character 'residual variation in the number of involucre bracts not explained by the regression against the number of achenes' is subjected to a QTL analysis, no independent QTL for phyllary number is detected in hybrid H27. However, QTL analysis of the character 'diameter of the involucre' is influenced by at least one QTL with significant effect (an r^2 of c. 20%) besides its determination by the pleiotropic QTL(s) for capitulum size (which alone explain c. 18% of the variance for involucre diameter). This double determination corresponds roughly to two components of the total diameter of the flattened involucre: the central disk of the receptacle that carries the florets and therefore has a diameter that is correlated with their number, and the length of the involucre bracts that is not dependent on the number of florets. Finally, another related character, the relative length of involucre and florets, i.e. the degree to which the florets are hidden in the involucre or extend far beyond it during anthesis, is determined by one to three QTLs depending on the stringency of the mapping, but none of these is identical with those affecting the absolute diameter of the involucre. This example illustrates not only the complete or partial overlap or independence of quantitative characters characterizing the capitulum at anthesis, it also underlines that the definition of characters is essentially arbitrary, but the precise way in which a character is defined strongly determines which and how many genes (QTLs) are found to affect it.

The definition of a character determines what genes are found

We consider the last point crucial for the analysis of multigene characters. Since computer-aided procedures for finding QTLs are not time-limiting, we regularly use various ways to define similar characters for QTL determination. This involves much redundant information (e.g. absolute values and differences or ratios between them) but can provide insights into the genetic structure of phenotypes that are hidden otherwise. One last example will demonstrate this point.

A very complex character of the annual species of *Microseris* is the direction in which the capitula are pointing. Usually, heads are nodding in bud, turn upright on the day of anthesis, return to a nodding position while the achenes are ripening and turn upright again when the mature fruiting head opens. This general cycle is overlain by a diurnal rhythm in which the capitula at anthesis turn upright when they open in the morning, close and return to a nodding position shortly after noon, and repeat this cycle with a diminishing amplitude on the next 2 d. The amplitude and timing also depends on the weather: heads open more fully in the sun but might stay open longer on an overcast day. There are striking differences among strains in the quantitative details of these head positions and movements, and most of these are heritable. In many strains, heads are never completely upright, and there are strains of *M. douglasii* in which 'nodding' heads do not point down but actually curve around 270° and point backwards. It is possible that this immense natural genetic variability can persist because the character has effectively no adaptive significance in these strictly autogamous plants. However, the complexity of the character and the heritability of any and all aspects should make it an interesting model case and justify the considerable amount of work needed to score all of these aspects comparatively in many individuals.

As a starting point, we have scored the head position on a scale of zero (pointing down 180°) to four (upright) repeatedly per plant in heads at anthesis and used the average value over many observations as a quantitative character to determine QTLs. This is a very crude determination of a very variable character. However, even a regression of the 450 F5 averages against the averages of their F2 ancestors has shown a significant heritability of 31.3% ($r^2 = 8.1\%$ owing to the very high individual variability). Both the evaluation of the F2 values with RAPDs and the evaluation of the F5 values with AFLPs have revealed four QTLs contributing significantly to a multiple regression of the phenotypic value against the predicted QTL genotypes. The following results were obtained from the

multiple regression of F2 values based on the four strongest QTLs determined with RAPD markers

$$\text{HEADPOS mid} = 110.8 - 13.3 \text{ nA} - 16.5 \text{ nB} - 10.5 \text{ nC} + 23.8 \text{ nD}, \quad (1)$$

and from the multiple regression of F5 values based on the four strongest QTLs determined with AFLP markers

$$\text{HEADPOS mid} = 117.1 - 13.0 \text{ na} - 10.4 \text{ nb} - 9.1 \text{ nc} + 7.9 \text{ nd}, \quad (2)$$

where the numerical values are expressed as degrees deviation from an upright position and nA–nd are the numbers of alleles (0, 1 or 2) from the *M. bigelovii* parent C94 for loci A–d.

Both equations explain c. 20% of the phenotypic variabilities in the data and agree in the fact that both analyses have found three QTLs (A, B, C and a, b, c) in which the alleles from C94 reduce the inclination, i.e. maintain a more upright position of the capitulum, and one QTL (D resp. d) that has an opposite polarity to the others and contributes to a transgressive segregation. Table 1 compares the predictions from the two models for some representative genotypes. There is no proof that the QTLs labelled A, B, C and D from the RAPD map are identical to the QTLs a, b, c and d from the AFLP map. However, the fact that one of the four strongest QTLs in the two mapping experiments (D and d) has an opposite polarity to the others makes it likely that at least this is the same locus found by two methods.

Since the genetic predictions of Table 1 describe only a small fraction of the total variation in head positions observed, we have taken the raw data on head position and determined for each plant the largest and smallest values for head position (HEADPOS) recorded among the repeated scores. Such single values are much more sensitive to chance deviations than the averages based on many observations, but the result of this approach was very surprising. The same four QTLs as before were found to be relevant, but their influence on maximum and minimum head position was entirely different. The following significant contributions remained after multiple regression:

better than the middle value alone, even though the linear extrapolation slightly overestimates the transgressive minima and maxima. The average (eqn (5)) calculated from the predicted lowest (eqn (3)) and highest (eqn (4)) head position is, of course, very similar to the prediction based on multiple regression of the individual plant average values in the F2 (eqn (1)).

DISCUSSION

In principle, a simultaneous analysis of the segregation of molecular single-locus markers and phenotypic characters is an ideal approach towards a very detailed knowledge of the genetic basis of any heritable phenotypic character difference. It identifies individual genetic loci contributing to the expression of the phenotypic character, and it provides linked molecular markers that make these loci accessible at the molecular level. Here, we have shown that QTL mapping with molecular markers can be applied to phenotypic differences between plant species in nature for which no previous genetic knowledge is available. Crosses between higher plants with major and evolutionary significant differences are often possible. As with our material, the fertility of the offspring usually is inversely related to the phenotypic differences, and we might have to accept irregularities in recombination and segregation that interfere with the establishment of a complete map. Ideally, we also would like to know the chromosomal basis for the irregular segregation. Since cytogenetic methods are not sensitive enough to detect these, this would best be done by comparative mapping of *M. douglasii* and *M. bigelovii* separately in fully fertile intraspecific crosses using markers that can be transferred between the two species. This is essentially the approach used by Rieseberg, Van Fossen & Desrochers (1995) to compare the genome of two sunflower species and their natural diploid hybrid. Here, we accept the problems associated with segregation in a wide cross in order to combine very divergent character states of many characters into one hybrid and get much information on genes underlying morphological

$$\text{HEADPOS min} = 189.2 - 29.9 \text{ nA} - 23.0 \text{ nB} + 25.2 \text{ nD} \quad (3)$$

$$\text{HEADPOS max} = 29.9 - 23.3 \text{ nC} + 20.2 \text{ nD} \quad (4)$$

$$\text{HEADPOS avg} = 109.6 - 15.0 \text{ nA} - 11.5 \text{ nB} - 11.7 \text{ nC} + 22.7 \text{ nD} \quad (5)$$

Strikingly, this analysis showed independent effects on maximal and minimal head position of the three loci with C94 alleles for high head position, whilst both characters are affected by the locus with opposite polarity. The individual regressions describe the total variability of the character much

variation in and between the two parental species. It is our aim to resolve complex multigenically determined phenotypic traits into the contributing genes and their interaction, and we want to do this in a way that allows us to predict phenotypes for all possible allelic combinations of the underlying genes.

Table 1. *Predictions for the genetically determined degree of the average inclination of the head at anthesis for various four-locus genotypes based on multiple regression models of the QTLs found in the F2 and F5 generations of hybrid H27*

Genotype (C94 alleles)				F2 prediction (RAPD map)	F5 prediction (AFLP map)
A	B	C	D		
a	b	c	d		
0	0	0	2 maximal transgressive	155.0	137.9
0	0	0	0 B14 parent	109.6	117.2
1	1	1	1 heterozygote (F1)	94.2	97.5
2	2	2	2 C94 parent	78.8	77.8
2	2	2	0 minimal transgressive	33.4	57.0

All values are degrees inclination from an upright (0°) position. The genotypes are alleles from parent C94 at four loci.

The divergence between discrete and objectively identifiable genes on the one hand and the arbitrary division of an essentially integrated, holistic phenotype into ‘characters’ becomes increasingly evident the more characters are mapped simultaneously in one cross and the more inclusive our knowledge of the genetic determination of the phenotype becomes. Interactions among gene products in and among cells in the developing plant that sense and react to environmental clues are the cause of the emergent processes and structures that integrate the effects of many genes with those of the local environment so that single-gene contributions are difficult to discern. Elsewhere (Bachmann, 1993) I have pointed out that hiding the effects of single genes by having them interact in a complex fashion, and especially by modifying their expression during plastic adaptation to the local environment, is in itself a result of natural selection. That conclusion has been clearly stated already by Sewall Wright (1931). To the geneticist trying to dissect the phenotype into single-gene effects, gene interactions and developmental plasticity have been a traditional hindrance. However, once we recognize that the complexity of interactions and the adaptation of gene expression rather than adaptation through specific alleles are essential evolutionary strategies, we see that the complex relation between genotype and phenotype becomes an important research object in its own right. Once we learn how to dissect the phenotype, we can begin to develop methods to reassemble phenotypes from genotypes.

The present study is a small and preliminary contribution towards this aim. Especially, it does not yet contribute to two of the most important aspects of genotype-phenotype interactions:

- (1) Our phenotype predictions are essentially descriptive based on empirical data. They can serve to interpolate data on phenotypes not

Table 2. *Characters for which genetic input (via genotypes of linked markers) is incorporated in a computer graphic model*

General architecture	Leaf initiation and phyllotaxis in rosette, no genetic variation, leaf initiation rate and leaf number genetically variable
Leaf	Leaf length genetically variable; leaf shape genetics not analysed, a standard shape is used
Rosette habitus	Upright or flat; genetically determined
Flower head	General architecture and phyllotaxis of florets preset; genetic input for number of outer and inner involucre bracts and florets; floret colour and relative length of florets and involucre; percentage of outer, hairy achenes, and pappus parts per achene
Scape	length and curvature genetically determined; curvature results from average angle of scape and inclination of head
Developmental timing	Via leaf initiation rate (see above), initiation of capitulum bud (and consequent rate to anthesis and seed maturity)

observed, but they are not sufficiently complex to predict emergent new properties of phenotypes beyond those observed in the experiment.

- (2) We predict essentially average phenotypes for a favourable glasshouse environment. Plastic responses are not considered. This is the major source of deviations between the phenotype predicted from marker genotypes and the true phenotype of a specific plant with that genotype.

The major contribution of our study is the possibility to predict phenotypes for many characters

Figure 2. Parameters used in the computer graphic model of capitulum curvature: the scape is represented by short concatenated elements, and their angles relative to one another are constrained by the angles α_0 – α_3 at points t_0 – t_3 . Total scape length and angles α_1 and α_3 have been analysed genetically and can be determined by predictions from marker genotypes.

simultaneously and to consider pleiotropic actions of QTLs. This takes into account at least the genetic correlations among various characters and is a first approximation to a holistic synthesis of an integrated phenotype from the detailed knowledge of single-gene genotypes.

The potential number of genotypes for the *c.* 50 QTLs detected up to now and based on the two alleles per locus present in the hybrid by far exceeds the number of plants analysed. The phenotypic variation created by recombination in the hybrid is most impressive when 150 more or less uniform offspring families are seen side by side so that heritable differences among the families are emphasized. Simultaneous transgressive segregations for several characters create new forms that could not be predicted from the differences between the parental strains. Some very evident differences, such as those in leaf shape, have not yet been analysed, but the known QTLs allow the prediction of the phenotypes for a great many multi-gene genotypes in some detail.

Traditionally these phenotypes are predicted by a table of the expected character states for the individual characters. Newer methods of computer graphics make it possible to design a relatively realistic drawing of the entire plant, or parts of it, in three dimensions, and even to simulate some of the time components of growth. A suitable descriptive technique for modelling the general architecture of herbaceous plants is that of ‘L-systems’ (Prusinkiewicz & Lindenmayer, 1990) which makes use of the iterative application of rewriting rules (‘productions’) to an initial string of letters (the ‘axiom’) and its successors. These letters might represent various plant organs such as apices or internodes so that the ‘L-systems’ describe the general growth pattern. Visually realistic models of the plants representing biologically relevant characters can be designed on this basis. Recent models of herbaceous plants based on ‘L-systems’

include one for *Hieracium umbellatum* (Prusinkiewicz, Hammel & Mjolsness, 1993) and cotton (Room, Hanan & Prusinkiewicz, 1996). Battjes (unpublished) has designed a visually realistic model of the annual *Microseris*, in which characters for which QTL data are available from this study can be varied to represent the range of phenotypes found in the segregating hybrids (Table 2). This model can be connected with the output of a quantitative genetical prediction of phenotypic values from marker genotypes so that the model uses the predicted character states in a three-dimensional graphic representation of the plant.

Figure 2 illustrates this for the curvature of the flower scape which is modeled as a series of concatenated elements so that each element in space is determined relative to the preceding element. This is implemented in three sections between characteristic points: insertion of the scape, point of inflexion, maximum before head inclination and insertion of capitulum. With four points and four angles, scape curvature can be described very accurately, and the predicted phenotypes of scape angle (α_1 in Fig. 2) and head inclination (α_3) can be accommodated by the model.

Although the model is not more than the simultaneous graphic presentation of the predicted character states for several characters, this in itself is a very convenient way to look at the predictions and to investigate the effects of allelic substitutions. It also allows the inclusion of genes affecting the timing of development by considering the genetics of the rate of production of rosette leaves and the timing of bud induction (which directly determines anthesis and ripening times). It does this by animating rosette growth and the growth and differentiation of the flower scapes. It is therefore possible to use the model for a comparison of the expected phenotypes at specific dates (days after germination).

The application of computer graphics for the realistic representation of phenotypes anticipates the generation of more, and more complex, models of genetic determination. The data for these will come from increasingly advanced mapping experiments and from the analysis of induced mutants in model systems such as *Arabidopsis*. These two approaches, from complex phenotype towards the gene and from the single mutant towards the phenotypic consequences, complement each other in revealing ever more details of plant developmental genetics. Computer models are already helpful at the level of descriptive data representation. The methods developed at this stage will become essential when emergent properties of gene interactions in development are studied and when the influence of environmental clues on these properties gets analysed.

A model for the generation of emergent properties is the generation of phyllotactic patterns. Spiral

phyllotaxis can be modelled descriptively by programming a preset angle of 137.5° between subsequent organs. There are also several theoretical models, generally based on inhibition from existing organ primordia, in which the pattern of spiral phyllotaxis arises as an emergent property (Battjes & Prusinkiewicz, 1997).

Accounting for plasticity in the prediction of phenotypes is difficult at the moment, because plastic responses have to be determined empirically for a range of environmental parameters (Sultan, 1995), and possibilities for extrapolating from one set of parameters to another are very limited. Until signal pathways in plants are known in sufficient detail to predict their interaction in a specific plant, two directions of empirical research could prepare their eventual integration: genetic differences in plastic responses of individual characters can be mapped when recombinant genotypes can be propagated vegetatively or as recombinant inbreds so that the relevant experiments can be performed, and the physiological correlations among various characters can be determined so that the integrated whole-plant response to an environmental signal can be predicted.

REFERENCES

- Bachmann K. 1992. Phenotypic similarity and genetic relationship among populations of *Microseris bigelovii* (Asteraceae: Lactuceae). *Acta Botanica* **105**: 337–342.
- Bachmann K. 1993. Effects of stress on the genome. In: Fowden L, Mansfield TA, Stoddart JL, eds. *Plant Adaptation to Environmental Stress*. London: Chapman and Hall, 275–292.
- Bachmann K, Battjes J. 1994. Variability in a predominantly self-fertilizing annual with fragmented distribution, *Microseris douglasii* (Asteraceae, Lactuceae). *Biologisches Zentralblatt* **113**: 69–95.
- Bachmann K, Chambers KL, Garu J, Price HJ. 1985. Genetic variation in *Microseris pygmaea* (Asteraceae-Lactuceae). *Beiträge zur Biologie der Pflanzen* **60**: 51–88.
- Bachmann K, Chambers KL, Price HJ. 1984. Differential geographic distribution of spatulate and pointed leaf shapes in *Microseris bigelovii* (Asteraceae-Lactuceae). *Beiträge zur Biologie der Pflanzen* **59**: 5–14.
- Bachmann K, Hombergen EJ. 1996. Mapping genes for phenotypic variation in *Microseris* (Lactuceae) with molecular markers. In: Caligari PDS, Hind DJN, eds. *Compositae: Biology and Utilization*. Kew: Royal Botanic Gardens, 23–43.
- Bachmann K, Van Heusden AW. 1992. Phenotypic similarity and genetic relationship in *Microseris elegans* (Asteraceae: Lactuceae). *Biologisches Zentralblatt* **111**: 38–53.
- Bachmann K, Van Heusden AW, Chambers KL, Price HJ. 1987. Genetic variation for the onset of flowering in *Microseris bigelovii* (Asteraceae, Lactuceae). *Beiträge zur Biologie der Pflanzen* **62**: 23–41.
- Battjes J, Bachmann K. 1994. Phenotypic plasticity of capitulum morphogenesis in *Microseris pygmaea* (Asteraceae: Lactuceae). *Annals of Botany* **73**: 299–305.
- Battjes J, Prusinkiewicz P. 1997. Modeling organ numbers in Asteracean heads. In: Barabé D, Jean RV, eds. *Symmetry in Plants*. Singapore: World Scientific Publishing. (In Press.)
- Bradshaw HD, Wilbert SM, Otto KG, Schemske DW. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* **376**: 762–765.
- Chambers KL. 1955. A biosystematic study of the annual species of *Microseris*. *Contributions of the Dudley Herbarium* **4**: 207–312.
- Hombergen EJ, Bachmann K. 1995. RAPD mapping of three QTLs determining trichome formation in *Microseris* hybrid H27 (Asteraceae: Lactuceae). *Theoretical and Applied Genetics* **90**: 853–858.
- Kosambi DD. 1994. The estimation of map distances from recombination values. *Annals of Eugenics* **12**: 172–175.
- Lander ES, Botstein D. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**: 185–199.
- Paterson AH, Lander ES, Hewitt JD, Peterson S, Lincoln SE, Tanksley SD. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction length polymorphisms. *Nature* **335**: 721–726.
- Phillips RL, Vasil IK, eds. 1994. *DNA-based markers in plants*. Dordrecht: Kluwer.
- Prusinkiewicz P, Hammel MS, Mjolsness E. 1993. Animation of plant development. *Computer Graphics* **27**: 351–360.
- Prusinkiewicz P, Lindenmayer A. 1990. *The algorithmic beauty of plants*. Berlin: Springer.
- Rieseberg LH, Van Fossen C, Desrochers AM. 1995. Hybrid speciation accompanied by genomic reorganization in wild sunflowers. *Nature* **375**: 313–316.
- Roelofs D, Bachmann K. 1995. Chloroplast and nuclear DNA variation among homozygous plants in a population of autogamous annual *Microseris douglasii* (Asteraceae, Lactuceae). *Plant Systematics and Evolution* **196**: 185–194.
- Roelofs D, Bachmann K. 1997. Genetic analysis of a *Microseris douglasii* (Asteraceae) population polymorphic for an alien chloroplast type. *Plant Systematics and Evolution* **206**: 273–284.
- Room P, Hanan J, Prusinkiewicz P. 1996. Virtual plants: new perspectives for ecologists, pathologists and agricultural scientists. *Trends in Plant Science* **1**: 33–38.
- Stam P. 1993. Construction of integrated genetic linkage maps by means of a new computer package, JOINMAP. *Plant Journal* **3**: 739–744.
- Sultan SE. 1995. Phenotypic plasticity and plant adaptation. *Acta Botanica Neerlandica* **44**: 363–383.
- Van Heusden AW, Bachmann K. 1992a. Genotype relationships in *Microseris elegans* (Asteraceae, Lactuceae) revealed by DNA amplification from arbitrary primers (RAPDs). *Plant Systematics and Evolution* **179**: 221–233.
- Van Heusden AW, Bachmann K. 1992b. Nuclear DNA polymorphisms among strains of *Microseris bigelovii* (Asteraceae: Lactuceae) amplified from arbitrary primers. *Botanica Acta* **105**: 331–336.
- Van Heusden AW, Bachmann K. 1992c. Genetic differentiation of *Microseris pygmaea* (Asteraceae, Lactuceae) studied with DNA amplification from arbitrary primers. *Acta Botanica Neerlandica* **41**: 385–395.
- Van Houten WHJ, Van Raamsdonk L, Bachmann K. 1994. Intraspecific evolution of *Microseris pygmaea* (Asteraceae: Lactuceae) analysed by cosegregation of phenotypic characters (QTLs) and molecular markers (RAPDs). *Plant Systematics and Evolution* **190**: 49–67.
- Vlot EC, Van Houten WHJ, Mauthe S, Bachmann K. 1992. Genetic and non-genetic factors influencing deviations from five pappus parts in a hybrid between *Microseris douglasii* and *M. bigelovii* (Asteraceae, Lactuceae). *International Journal of Plant Science* **153**: 89–97.
- Vos P, Hogers R, Bleeker M, Reijans M, van der Lee Th, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407–4414.
- Welsh J, McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* **18**: 7213–7218.
- Williams JKG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* **18**: 6531–6535.
- Wright S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97–159.