Indian Journal of Agricultural Sciences **83** (12): 1372–76, December 2013/Article

The new method of identification of extremes groups based on epistatic interaction effect using molecular markers

JAN BOCIANOWSKI[1]

Poznañ University of Life Sciences, Wojska Polskiego 28, 60-637 Poznañ, Poland

Received: 28 March 2012; Revised accepted: 18 October 2013

ABSTRACT

Additive × additive epistatis interaction is an important and integral feature of the genetic architecture of quantitative traits. Epistasis is difficult to detect in classical quantitative analyses based on resemblance between relatives in outbred populations. In traditional quantitative genetics estimation of the additive × additive epistasis effect on the basis of phenotypic observations requires identification of the groups of extreme lines, i.e. group of minimal lines and group of maximal lines. Minimal and maximal lines contain, respectively, only alleles decreasing and only alleles increasing the value of the trait. Because the genotype has not been known, to extreme groups were usually put lines with minimal and maximal expression of the observed traits. Estimates obtained by this way are biased. Therefore, in this paper estimation of the parameter connected with the additive × additive epistasis genes interaction effect on the basis of extreme groups of doubled haploid lines identified on the basis of molecular marker data is presented. Numerical analysis was made on the basis of three populations of doubled haploid lines of barley (*Hordeum vulgare* L.), viz. 120 DH lines from the Clipper × Sahara 3771 cross, 145 DH lines from the Harrington × TR306 cross and 150 DH lines from the Steptoe ´ Morex cross. In total, 157 sets of observations were analyzed. The estimates of additive × additive epistasis effects from proposed (extreme lines) method are much less biased than the estimates obtained by quantile method.

Key words: Barley, Double haploid lines, Epistasis, Extreme lines

Epistatic additive × additive (non-allelic) interaction is an important genetic basis of quantitative traits in different plants. It is important to identify and estimate additive × additive epistasis interaction because this parameter can influence decisions about usefulness of the breeding material for obtaining new transgressive genotypes (Bocianowski 2012). Moreover, epistasis must be included in a model for the unbiased estimation of genetic components (Sharma *et al.* 2012).

QTLs with epistatic effects have been detected for different quantitative traits in many plants (Lee *et al.* 2006, Yang *et al.* 2007, Zhang *et al.* 2008, 2009; Krajewski *et al.* 2012). Previous studies have demonstrated that the epistatic interaction was prevalent in quantitative trait (i.e. the number of internodes in the primary lateral branch, the average length of these internodes, the proportion of male or staminate spikelets in the primary lateral inflorescence, and the number of branches in the primary lateral inflorescence, grain weight, peduncle number, pod number per peduncle, seed number per pod, thousand seed weight, grain yield and protein content) inheritance (Doebley *et al.* 1995, Yu *et al.* 1997).

In traditional (classical) quantitative genetics estimation of the additive × additive (*aa*) epistasis effect on the basis of phenotypic observations requires identification of the groups of extreme lines, i e group of minimal lines and group of maximal lines. Minimal and maximal lines contain, respectively, only alleles responsible for decreasing and only alleles responsible for increasing the value of the trait (Choo and Reinbergs 1982). Because the genotype has not been known, to extreme groups were put lines with minimal and maximal expression of the observed traits. Bocianowski *et al.* (1999) showed the methods of identification of extreme groups of double haploid lines in estimation of parameters connected with the additive and additive × additive interaction effects. As preferred (giving the smallest bias) the authors propose the quantile method, in which as minimal (maximal) lines are taken the ones with the mean values smaller (bigger) than 0.03 (0.97) quantile of the empirical distribution of means. Estimates obtained by this way are biased. Can it be possible to eliminate or to reduce this bias? Because the quantitative trait is determined by large number of genes therefore, the probability of obtaining extreme genotypes is very small. If ten genes influence on value of quantitative

Department of Mathematical and Statistical Methods (e mail: jboc@up.poznan.pl)

trait, then the number of different genotypes equals 1024. In experiment the number of lines has equaled from 100 to 200. Hence, the probability of obtainment at least one extreme line has equal from about 0.1 to 0.2. However, the probability of obtainment both the minimal and maximal lines equals from about 0.01 to 0.04. Moreover, the value of quantitative trait is in large degree determined by environment is also taken as it is an additional bias. The development of molecular genetics and other methods for studying genotypes at the DNA level offered new tools in analysis of quantitative traits independently from environmental conditions (Bocianowski *et al.* 2011). Hence, in this paper estimation of the parameter connected with the additive × additive epistasis genes interaction effect on the basis of extreme groups of doubled haploid lines identified on the basis of molecular marker data is presented.

MATERIALS AND METHODS

A population of 120 doubled haploid lines derived from a cross between the boron-sensitive Australian malt barley variety Clipper and the boron-tolerance Algerian landrace Sahara 3771 at Waite Agricultural Research Institute, University of Adelaide, Australia (Karakousis *et al.* 2003) was used as the first dataset in this study. The lines were investigated with respect to four phenotypic traits: beta-amylase activity, BA; alpha-amylase activity, AA; beta-glucanase activity, BG; cyst nematode resistance, CCN. We use observations of 183 molecular markers (SSR and RFLP).

The second example dataset is the barley Harrington × TR 306 doubled haploid (DH) population (Tinker *et al.* 1996), a well known population from the North American Barley Genome Mapping Project (http://wheat.pw.usda.gov/ggpages/maps/Hordeum). The data matrix consisted of 145 doubled haploid lines. The DH lines were analysed for seven phenotypic traits (weight of grain harvested per unit area, WG; number of days from planting until emergence of 50% of heads on main tillers, NH; number of days from planting until physiological maturity, NM; plant height, H; lodging transformed by arcsin $\sqrt{x}/100$, L; 1000 kernel weight, KW; test weight, TW). For this purpose, the map was composed of 127 molecular markers (mostly RFLP) with the mean distance between markers equal to 10.62 cM. Results shown below concern observations from five locations (in four locations observations were made over two years): ON92a – Ailsa Craig, Ontario, 1992; ON93a – Ailsa Craig, Ontario, 1993; ON92b – Elora, Ontario, 1992; ON93b – Elora, Ontario, 1993; MB92 – Brandon, Manitoba, 1992; MB93 – Brandon, Manitoba, 1993; QC93 – Ste-Anne-de-Bellevue, Quebec, 1993; SK92a – Outlook, Saskatchewan, 1992; SK93a – Outlook, Saskatchewan, 1992.

The third dataset mapping population, consisting of 150 doubled haploid lines of barley obtained from the Steptoe × Morex cross, used in the North American Barley Genome Mapping project and tested at sixteen environments (Kleinhofs *et al.* 1993, Romagosa *et al.* 1996, http://wheat.pw.usda.gov/ggpages/SxM). The linkage map used consisted of 223 molecular markers, mostly RFLP, with mean distance between markers equal to 5.66 cM. The DH lines were analysed for eight phenotypic traits (alpha amylase, AA; diastatic power, DP; grain protein, GP; grain yield, GY; height, H; heading date, HD; lodging, L; malt extract, ME; Hayes *et al.* 1993). Grain protein, lodging and malt extract were transformed by arcsin $\sqrt{x}/100$.

If in the experiment we observed n significantly different biparental homozygous (doubled haploid, DH or recombinant inbred, RI) plant lines, we obtain an n-vector of phenotypic mean observations $y = [y_1\ y_2\ ¼\ y_n]'$ and q n-vectors of marker genotype observations m_l, l=1, 2, …, q. The i-th element (i = 1, 2, …, n) of vector m_l is equal to –1 or 1, depending on the parent's genotype exhibited by the i-th line. The total additive × additive epistasis interaction effect aa can be estimated by the formula (Surma *et al.* 1984)

$$\breve{a}a_{ex} = \frac{1}{2}(\bar{L}_M + \bar{L}_m) - \bar{L}, \quad (1)$$

where $\bar{L}_m$ and $\bar{L}_M$ denote the means for the groups of minimal and maximal lines, respectively, $\bar{L}$ denotes the mean for all lines.

The group of minimal lines (L_m) consists the lines which have the largest number of markers observations equal to -1. Analogously, the group of maximal lines (L_M) contains the lines which have the largest number of markers observations equal to 1. The groups of extreme lines defined by this way are the best estimators of groups of lines contain alleles only decreasing or increasing the value of the quantitative trait.

RESULTS AND DISCUSSION

In total, 157 sets of observations (four traits of DH lines from Clipper × Sahara 3771 cross, seven traits in nine environments of DH lines from Harrington × TR 306 cross and eight traits in 16 environments (only three traits – GY, HD and H – were observed in all 16 environments) of DH lines from Steptoe × Morex cross) were analyzed.

DH lines from Clipper × Sahara 3771 cross

For Clipper × Sahara 3771 doubled haploid lines, line DH 165 have the largest number of 1, however DH 284 line characterized the largest number of –1. Table 1 contains the results of the estimates of the additive × additive epistasis

Table 1 Estimates of additive × additive epistasis interaction effects for 120 doubled haploid lines of barley obtained from the Clipper × Sahara 3771 cross

Trait	Beta-amylase activity	Alpha-amylase activity	Beta-glucanase activity	Cyst nematode resistance
Epistasis effect	×205.3	4.061	10.9	8.6

interaction effects for four quantitative traits of Clipper × Sahara 3771 doubled haploid lines. The estimates were equal to -205.3, 4.061, 10.9 and 8.6 for BA (Beta-amylase activity), AA (Alpha-amylase activity), BG (Beta-glucanase activity) and CNN (Cyst nematode resistance), respectively. In all four cases the signs of estimates were opposite to signs of estimates obtained on the basis of quantile method (Bocianowski 2008).

DH lines from Harrington × TR306 cross

The additive × additive epistasis effects of quantitative traits for doubled haploid lines from Harrington × TR 306 cross were presented in Table 2. Only in one case (for NH in ON93b) the epistasis effect was equal to 0. Fig 1 shows the summary of the comparison between additive × additive epistasis interaction effects estimated on the basis of method proposed in this paper (extreme lines method, $\hat{a}a_{ex}$) and on the basis of quantile method ($\hat{a}a_q$ - results obtained by Bocianowski 2012) in the form a box-and-whisker diagram of the observed values $\hat{a}a_{ex}/\hat{a}a_q$ for Harrington × TR 306 doubled haploid lines (variability over nine environments). For the NH and TW the additive-by-additive epistasis effects calculated on the basis of both methods were similar. However, the range of the calculated coefficients is quite large, from -52.57 for KW in one of the environments, to 29.94 for NM.

Table 2 Estimates of additive × additive epistasis interaction effects for doubled haploid lines from Harrington × TR 306 cross

Environment	Trait						
	WG[#]	NH	NM	H	L	KW	TW
ON92a[&]	–3.45	0.610	0.659	–1.631	–0.047	–0.724	–0.168
ON93a	25.22	0.076	–0.331	3.993	0.242	–1.472	–1.367
ON92b	–23.14	0.772	0.079	2.262	0.242	–0.193	0.005
ON93b	–9.68	0	0.224	–1.359	0.145	–0.554	–0.389
MB92	11.98	0.293	0.728	1.479	0.065	0.003	–0.239
MB93	16.83	0.386	1.903	0.938	–0.018	1.112	0.147
QC93	27.94	0.272	0.403	–5.533	–0.323	1.016	0.818
SK92a	12.36	0.648	–1.882	–4.217	–0.302	1.189	1.222
SK93a	–65.13	–0.290	–0.617	1.109	0.028	–0.040	0.551

[&]ON92a – Ailsa Craig, Ontario, 1992; ON93a – Ailsa Craig, Ontario, 1993; ON92b – Elora, Ontario, 1992; ON93b – Elora, Ontario, 1993; MB92 – Brandon, Manitoba, 1992; MB93 – Brandon, Manitoba, 1993; QC93 – Ste-Anne-de-Bellevue, Quebec, 1993; SK92a – Outlook, Saskatchewan, 1992; SK93a – Outlook, Saskatchewan, 1992.

[#]WG – weight of grain harvested per unit area; NH – number of days from planting until emergence of 50% of heads on main tillers; NM – number of days from planting until physiological maturity; H – plant height; L – lodging; KW – 1000 kernel weight; TW – test weight.

Table 3 Estimates of additive × additive epistasis interaction effects for Steptoe × Morex doubled haploid lines barley population

Environment	Trait							
	AA[#]	DP	GY	GP	HD	H	L	ME
ID91[&]	6.508	20.36	0.089	0.0019	–1.550	–3.554		0.0088
ID92	3.866	24.67	0.610	–0.0062	–1.253	–1.183		0.0205
MA92			0.349		0.210	–3.730	0.008	
MN92	10.280	11.16	–0.404	–0.0399	–2.323	–9.420		0.0454
MTd91			–0.504		–0.180	–1.168		
MTd92	–0.227	20.11	–0.106	0.0009	–1.070	2.145	0.197	0.0103
MTi91	2.597	11.29	–0.419	–0.0107	–1.193	–6.820		0.0251
MTi92	3.083	22.32	–0.066	0.0108	–1.483	–6.953	0.197	0.0161
NY92			0.453		–1.193	–3.833	0.113	
ON92			0.355		0.980	–1.467	0.222	
OR91	6.571	26.09	0.784	–0.0033	–0.660	–5.393		0.0243
SKg92			–0.227		0.427	–0.594		
SKk92			–0.228		0.047	–2.953		
SKo92			0.242		–0.947	–3.940	–0.004	
WA91	7.471	13.55	–0.079	–0.0137	–0.363	–6.277		0.0183
WA92	4.376	11.91	0.518	–0.0020	–1.387	–0.667		0.0198

[&]ID91 – Aberdeen, Idaho, 1991; ID92 – Tetonia, Idaho, 1992; MA92 – Brandon, Manitoba, 1992; MN92 – Crookston, Minnesota, 1992; MTd91 – Bozeman, Montana, dry, 1991; MTd92 – Bonzeman, Montana, dry, 1992; MTi91 – Bozeman, Montana, irrigation, 1991; MTi92 – Bozeman, Montana, irrigation, 1992; NY92 – Ithaca, New York, 1992; ON92 – Guelph, Ontario, 1992; OR91 – Klamath Falls, Oregon, 1991; SKg92 – Goodlae, Saskatchewan, 1992; SKk92 – Kcfr, Saskatchewan, 1992; SKo92 – Outlook, Saskatchewan, 1992; WA91 – Pullman, Washington, 1991; WA92 – Pullman, Washington, 1992.

[#]WG – weight of grain harvested per unit area; NH – number of days from planting until emergence of 50% of heads on main tillers; NM – number of days from planting until physiological maturity; H – plant height; L – lodging; KW – 1000 kernel weight; TW – test weight.

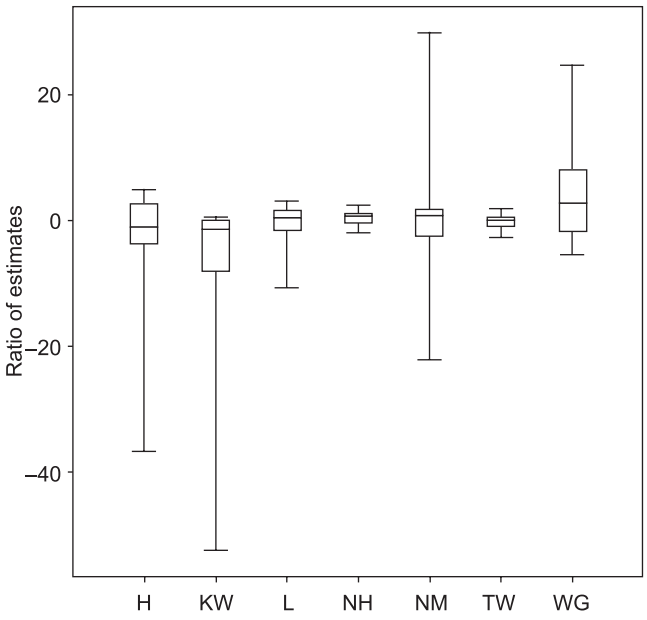

Fig 1 Relative comparison of additive × additive epistasis effects estimated on the basis of method proposed in this paper ($\hat{a}a_{ex}$) and on the basis of quantile method ($\hat{a}a_q$) - results obtained by Bocianowski (2012): box-and-whisker diagram of the values $\hat{a}a_{ex}/\hat{a}a_q$, classified by the observed phenotypic traits for Harrington × TR306 doubled haploid lines

In 31.75% cases coefficient was larger than 1. It means that new method gives absolute values of epistasis effects larger than absolute values obtained on the basis of quantile method. In 52.46% cases $\hat{a}a_{ex}/\hat{a}a_q$ coefficient was larger than 0. It means that epistasis effects for both methods had the same signs. The negative value denotes that estimates $\hat{a}a_{ex}$ and $\hat{a}a_q$ have opposite signs.

DH lines from Steptoe × Morex cross

Table 3 contains additive × additive epistasis effects for Steptoe Morex doubled haploid lines of barley population. For DP and ME epistasis effects were positive in all environments. Differences in epistasis effects estimation results among various environment/trait combinations have resulted in unreliable indications of the significance of additive-by-additive × environment interaction. In 48.89% cases coefficient was larger than 1, while in 61.11% larger than 0.

REFERENCES

Bocianowski J. 2012. Analytical and numerical comparisons of two methods of estimation of additive × additive interaction of QTL effects. *Scientia Agricola* **69**(4): 240–6.

Bocianowski J, Kozak M, Liersch A and Bartkowiak-Broda I. 2011. A heuristic method of searching for interesting markers in terms of quantitative traits. *Euphytica* **181**: 89–100. (http://www.springerlink.com/content/u6637288648521j1/)

Bocianowski J, Krajewski P and Kaczmarek Z. 1999. Comparison of methods of choosing extreme doubled haploid lines for genetic parameter estimation. *Colloquium Biometryczne* **29**: 193–202.

Borràs-Gelonch G, Denti M, Thomas W T B and Romagosa I. 2012. Genetic control of pre-heading phases in the Steptoe × Morex barley population under different conditions of photoperiod and temperature. *Euphytica* **183**: 303–21. (http://www.springerlink.com/content/k4830173515556v4/)

Choo T M and Reinbergs E. 1982. Estimation of the number of genes in doubled haploid populations of barley (*Hordeum vulgare*). *Canadian Journal of Genetics and Cytology* **24**: 337–41. (http://www.nrcresearchpress.com/doi/abs/10.1139/g82-035)

Doebley J, Stec A and Gustus C. 1995. *Teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* **141**(1): 333–46. (http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1206731/)

Hayes P M, Liu B H, Knapp S J, Chen F, Jones B, Blake T, Franckowiak J, Rasmusson D, Sorrells M, Ullrich S E, Wesenberg D, and Kleinhofs A. 1993. Quantitative trait locus effects and environmental interaction in a sample of North American barley germ plasm. *Theoretical and Applied Genetics* **87**: 392–401. (http://www.springerlink.com/content/g156pk6582740125/)

Henshall J M and Goddard M E. 1999. Multiple-trait mapping of quantitative trait loci after selective genotyping using logistic regression. *Genetics* **151**: 885–94. (http://www.genetics.org/content/151/2/885.short)

Karakousis A, Barr A R, Kretschmer J M, Manning S, Jefferies S P, Chalmers K J, Islam A K M and Langridge P. 2003. Mapping and QTL analysis of the barley population Clipper × Sahara. *Australian Journal of Agricultural Research* **54**(12): 1137–40. (http://www.publish.csiro.au/paper/AR02180.htm)

Kleinhofs A, Kilian A, Saghai Maroof M A, Biyashev R M, Hayes P, Chen F Q, Lapitan N, Fenwick A, Blake T K, Kanazin V, Ananiev E, Dahleen L, Kudrna D, Bollonger J, Knapp S J, Liu B, Sorrells M, Heun M, Franckowiak J D, Hoffman D, Skadsen R, Steffenson and B J. 1993. A molecular, isozyme and morphological map of the barley (*Hordeum vulgare*) genome. *Theoretical and Applied Genetics* **86**: 705–12. (http://www.springerlink.com/content/u2187gh8lu5407u1/)

Krajewski P, Bocianowski J, Gaw³owska M, Kaczmarek Z, Pniewski T, Œwiêcicki W and Wolko B. 2012. QTL for yield componenets and protein content: a multienvironment study of two pea (*Pisum sativum* L.) populations. *Euphytica* **183**: 323–36. (http://www.springerlink.com/content/9q1v050501p3q563/)

Lee S Y, Ahn J H, Cha Y S, Yun D W, Lee M C, Ko J C, Lee K S, Eun M Y. 2006. Mapping of quantitative trait loci for salt tolerance at the seedling stage in rice. *Molecules and Cells* **21**(2): 192–6. (http://www.ncbi.nlm.nih.gov/pubmed/16682812)

Liu T, Thalamuthu A, Liu J J, Chen C, Wang Z and Wu R. 2011. Asymptotic distribution for epistatic tests in case–control studies. *Genomics* **98**(2): 145–51. (http://www.sciencedirect.com/science/article/pii/S088875431100125X)

Ohno Y, Tanase H, Nabika T, Otsuka K, Sasaki T, Suzawa T, Morii T, Yamori Y and Saruta T. 2000. Selective genotyping with epistasis can be utilized for a major quantitative trait locus mapping in hypertension in rats. *Genetics* **155**: 785–92. (http://www.genetics.org/content/155/2/785.short)

Rahman L, Khanam S, Roh J-H and Koh H-J. 2011. Mapping of QTLs involved in resistance to rice blast (*Magnaporthe grisea*) using *Oryza minuta* introgression lines. *Czech Journal of Genetics*

and Plant Breeding **47**(3): 85–94. (http://www.agriculture journals.cz/publicFiles/47675.pdf)

Romagosa I, Ullrich S E, Han F and Hayes P M. 1996. Use of the additive main effects and multiplicative interaction model in QTL mapping for adaptation in barley. *Theoretical and Applied Genetics* **93**: 30–7. (http://www.springerlink.com/content/n31383725v718x35/)

Sharma A, Kapur P and Katoch V. 2012. Generation mean analysis to estimate genetic parameters for desirable horticulture traits in garden pea (*Pisum sativum*). *Indian Journal of Agricultural Sciences* **82**(3): 201–6. (http://epubs.icar.org.in/ejournal/index.php/IJAgS/article/view/15938)

Surma M, Adamski T, Kaczmarek Z. 1984. The use of doubled haploid lines for estimation of genetic parameters. *Genetica Polonica* **25**: 27–32.

Tinker N A, Mather D E, Rossnagel B G, Kasha K J, Kleinhofs A, Hayes P M, Falk D E, Ferguson T, Shugar L P, Legge W G, Irvine R B, Choo T M, Briggs K G, Ullrich S E, Franckowiak J D, Blake T K, Graf R J, Dofing S M, Saghai Maroof M A, Scoles G J, Hoffman D, Dahleen L S, Kilian A, Chen F, Biyashev R M, Kudrna D A and Steffenson B J. 1996. Regions of the genome that affect agronomic performance in two-row barley. *Crop Science* **36**: 1053–62. (https://www.crops.org/publications/cs/abstracts/36/4/CS0360041053)

Weil M W, Brown B W and Serachitopol D M. 1997. Genotype selection to rapidly breed congenic strains. *Genetics* **146**: 1061–9. (http://www.genetics.org/content/146/3/1061.short)

Yang D L, Jing R L, Chang X P and Li W. 2007. Identification of quantitative trait loci and environmental interactions for accumulation and remobilization of water-soluble carbohydrates in wheat (*Triticum aestivum* L.) stems. *Genetics* **176**(1): 571–84. (http://www.genetics.org/content/176/1/571.short)

Yu S B, Li J X, Tan Y F, Gao Y J, Li X H, Zhang Q F and Maroof M A S. 1997. Importance of epistasis as the genetic basis of heterosis in an elite rice hybrid. *Proceedings of the National Academy of Sciences of the United States of America* **94**(17): 9 226–31. (http://www.pnas.org/content/94/17/9226.short)

Zhang K P, Tian J C, Zhao L, Liu B and Chen G F. 2009. Detection of quantitative trait loci for heading date based on the doubled haploid progeny of two elite Chinese wheat cultivars. *Genetica* **135**(3): 257–65. (http://www.springerlink.com/content/m844885g7415k727/)

Zhang K P, Tian J C, Zhao L and Wang S S. 2008. Mapping QTLs with epistatic effects and QTL × environment interactions for plant height using a doubled haploid population in cultivated wheat. *Journal of Genetics & Genomics* **35**(2): 119–27. (http://www.sciencedirect.com/science/article/pii/S1673852708600 17X)