Defining Subregions and Estimating Benefits for a Specific-Adaptation Strategy by Breeding Programs: A Case Study

Paolo Annicchiarico,* Fadila Bellah, and Tiberio Chiari

ABSTRACT

Producing specific varieties for each subregion of a target region, instead of widely adapted varieties, may exploit positive genotype × location (GL) interactions to increase crop yields. With reference to the Algerian durum wheat (*Triticum durum* Desf.) region, as defined in a Geographic Information System (GIS), this study aimed at (i) comparing additive main effects and multiplicative interaction (AMMI) vs. joint regression modeling of GL effects, (ii) verifying the reliability of a GIS-based definition of two subregions that extended the site classification on the basis of GL effects as a function of long-term winter mean temperature in the GIS, and (iii) comparing wide vs. specific adaptation in terms of observed and predicted yield gains (the latter for scenarios of defined and undefined selection locations). Twenty-four cultivars from international centers, Europe, and North Africa were evaluated across 3 yr in a total of 47 environments by randomized complete block designs with four replications per trial. The AMMI modeling was distinctly superior to joint regression. The AMMI + cluster analysis and pattern analysis classified test locations consistently and in good agreement with the GIS-based subregion definition. Under the hypothesis of six selection environments assigned to subregions in proportion to their size (three sites in each of two years) for late stage selection, specific adaptation provided 2 to 7% greater gains than wide adaptation over the region at similar costs. The advantage of specific adaptation was much larger (39% determined on the basis of observed gains) for the smaller, stressful inland subregion, where specific adaptation may also enhance food security. Performing earlier specific selection, or adopting distinct genetic bases for each subregion, may further increase the advantage of specific breeding.

E ARLY PLANT BREEDERS advocated the exploitation of specific-adaptation effects to increase crop yields (Engledow, 1925). However, national breeding programs have mainly focused on the improvement of yield potential (Bramel-Cox et al., 1991), pursuing a wide-adaptation strategy that tended to neglect areas with lower potential for crop production (Ceccarelli, 1989). Also in less-developed countries, varieties have mostly been selected in favorable environments and then introduced with technological packages (mineral fertilizers, pesticides, irrigation) designed to significantly improve the growing environment (Simmonds, 1979, p. 356). This

high-input model of agriculture has raised the crop yields in many regions worldwide, but its further expansion may not be sustainable because of the high costs entailed and the negative impact on natural resources (Conway, 1998). A greater focus on breeding for diversified, less favorable environments has been advocated for increasing crop yields and, for staple crops, to enhance food security at the household level (Bramel-Cox et al., 1991; Ceccarelli, 1996). This implies a specific-adaptation strategy that exploits positive GL interaction effects that are repeatable over time while minimizing other genotype × environment (GE) interaction effects by selection for higher yield stability (Annicchiarico, 2002a). Breeding for wide adaptation aims to develop a variety that performs well in nearly all sites of the target region, whereas breeding for specific adaptation aims to produce different varieties, each of which performs well in a specific area (subregion) within the region. Specific adaptation may be directed not only to a favorable or unfavorable subregion, but also to different, relatively favorable areas, e.g., by exploiting different optima of time of flowering (Wallace et al., 1993). Specific breeding may enlarge the scope for marker-assisted selection, because a large portion of useful markers may be environment-specific (Xu, 2002). It also provides an ideal framework for the possible adoption of a participatory plant breeding approach (Ceccarelli, 1996). The merit of specific-adaptation breeding for globally-oriented breeding programs of large seed companies or international research centers is largely recognized. However, its potential for national breeding programs of medium-sized countries has been poorly investigated.

The scope for a specific-adaptation strategy may be explored on the basis of yield response of a germplasm pool that is representative of the available genetic base tested across a representative sample of sites within the target region (Annicchiarico, 2002a). After verifying the relative size of GL interaction effects that arise from the lack of genetic correlation among environments, candidate subregions may be identified on the basis of similarity for GL effects. Decisions on the adaptation strategy should be made on the basis of predicted yield gains offered by each strategy, and possibly verified by actual gains. Breeding for specific adaptation tends to imply (i) higher costs than wide adaptation, because of increased field testing and (ii) greater yield gains, owing to exploitation of GL effects and to increased heritability of yield as a consequence of low GL interaction

P. Annicchiarico, Istituto Sperimentale per le Colture Foraggere, 29 viale Piacenza, 26900 Lodi, Italy; F. Bellah, Institut Technique des Grandes Cultures, BP 16, El-Harrach, Alger, Algeria; and T. Chiari, Istituto Agronomico per l'Oltremare, 4 via A. Cocchi, 50131 Florence, Italy. This work was part of the project "Amélioration et renforcement du système nationale d'adaptation du blé dur en Algérie" funded by the Governments of Algeria and Italy. Received 8 Sept. 2004. *Corresponding author (bred@iscf.it).

Abbreviations: AMMI, additive main effects and multiplicative interaction; ANOVA, analysis of variance; CIMMYT, Centro Internacional de Mejoramiento de Maiz y Trigo; GE, genotype × environment; GIS, Geographic Information System; GL, genotype × location; GY, genotype × year; ICARDA, International Center for Agricultural Research in the Dry Areas; PC, principal component.

within subregion (Kang, 1998). A fair comparison of wide- vs. specific-adaptation strategies should have similar costs if the same total number of selection environments are used (Annicchiarico, 2002a). However, there are choices to be made among the several available techniques (DeLacy et al., 1996; Annicchiarico, 2002b) for modeling GL effects and/or classifying locations. In addition, classification of test sites may be insufficient for an accurate environmental definition of subregions and the appreciation of their relative size. Annicchiarico (2002b, p. 56) proposed extending the classification to nontest locations by means of environmental variables that are associated with GL effects and are capable of discriminating between site groups. However, the reliability of this procedure is unknown.

With reference to the potential durum wheat cropping region of Algeria as defined in a GIS (Delli et al., 2002), we had the following objectives: (i) to compare AMMI vs. joint regression modeling of GL effects, (ii) to verify the reliability of a GIS-based definition of two subregions that extended the site classification based on similarity for GL effects as a function of climatic variables in the GIS, and (iii) to compare wide- vs. specific-adaptation strategies in observed and predicted yield gains. Predicted gains were estimated for scenarios of defined and undefined selection locations.

MATERIALS AND METHODS

Plant Material and Environments

Twenty-four durum wheat cultivars, which were among the most widely grown or the best performing in Algeria, were evaluated for grain yield from 1998–1999 through 2000–2001. Included were four traditional cultivars derived from local landraces, two old varieties, and 18 modern varieties selected by international research centers, in Europe or North Africa, (Table 1). Test sites are reported in Table 2 and their geo-

graphical locations are in Fig. 1. Yield data for one location (site 17) were not available due to repeated crop failure. Yield data for Sites 3, 13, and 15 were not available for 1 or 2 yr (Table 2) for various reasons. Site numbers are consistent with previous reports (ITCF-IAO, 2002; Annicchiarico et al., 2002a).

Each trial was sown in autumn using a randomized complete block design with four replications and plots that were 10 m long × 1.2 m wide (6 rows, 20 cm apart). Management practices are reported in ITCF-IAO (2002) together with information on soil type and climatic data.

Definition of Subregions

Information from the first two years, which included records on six climatic variables, was used for preliminary modeling of adaptive responses and definition of subregions (Annicchiarico, 2002a; Annicchiarico et al., 2002a). Seven techniques for classifying sites into two subregions were preliminarily compared in terms of predicted yield gains over the region for specific adaptation vs. wide adaptation, assuming that the best methods for subregion definition maximize the advantage of specific adaptation. Six methods used the modeling of GL interaction effects by joint regression (Finlay and Wilkinson, 1963), AMMI (Gauch, 1992), or factorial regression (Hardwick and Wood, 1972) followed by grouping of locations either by cluster analysis or by using the main crossover point as a cut-off for defining two groups. The main crossover point is the estimated value of site mean yield for joint regression, site score on the first GL interaction principal component (PC) axis or site covariate for unidimensional factorial regression for which the crossover interactions of genotypes with sites reach the highest frequency (Annicchiarico, 2002b). These analyses were performed on log-transformed yields to remove the positive relationship between heterogeneity of genotypic variance and site mean yield of locations, which is frequent in Mediterranean regions (Yau, 1991) and may negatively affect the assessment of site similarity (Cooper et al., 1996). The regression of the within-site phenotypic variance of entry values as a function of site mean yield on a

Table 1. Name, origin, germplasm type, mean yield across environments of the region and of two subregions, and scaled score on the first GL interaction PC axis, for durum wheat genotypes.

Cultivar			Mean yield			PC 1
Name	Origin	Type†	Region	Subregion A	Subregion B	
			Mg/ha			(Mg/ha)^0.5
Hedba/Gerardo	Algeria	M	1.92	1.40	2.41	0.09
Bidi/Waha/Bidi	Algeria	M	2.30‡	1.66	2.91‡	0.26
GTA Dur	CIMMYT	M	2.45‡	1.88‡	3.02‡	0.35
Eider	CIMMYT	M	2.31‡	1.77	2.85‡	0.30
Chen's	CIMMYT	M	2.42‡	1.75	3.04‡	0.44
Sahel 77	CIMMYT	M	2.30‡	1.80‡	2.82‡	0.22
T. Polonicum/Z.B.	Algeria	O	1.85	1.49	2.17	−0.63
Hedba 3	Algeria	L	1.76	1.48	2.01	−0.76
Mexicali 75	CIMMYT	M	2.30‡	1.58	2.99‡	0.39
Kebir	ICARDA	M	2.03	1.62	2.48	−0.06
Om Rabi 9	ICARDA	M	2.20	1.66	2.71	0.02
Belikh 2	ICARDA	M	2.19	1.62	2.77	0.24
Bidi 17	Algeria	L	1.78	1.46	2.12	−0.63
Waha	ICARDA	M	2.28	1.74	2.85‡	0.21
Oued Zenati 368	Algeria	L	1.73	1.46	2.01	−0.80
M. Ben Bachir	Algeria	L	1.74	1.45	2.01	−0.65
INRAT 69	Tunisia	M	1.91	1.34	2.49	0.08
Ardente	France	M	2.07	1.63	2.52	−0.02
Vitron	Spain	M	2.15	1.52	2.79	0.35
B. Dur 1.94	CIMMYT	M	2.32‡	1.80‡	2.83‡	0.25
Ofanto	Italy	M	2.33‡	1.97‡	2.72	0.13
Simeto	Italy	M	2.31‡	1.74	2.90‡	0.47
Duilio	Italy	M	2.22	1.63	2.79	0.40
Polonicum	France	O	1.87	1.64	2.13	−0.67

† L = traditional cultivar derived from local landrace; O = old variety; M = modern variety.
‡ Not different from the top-ranking mean according to Dunnett's one-tailed test at *P* < 0.20. See Fig. 1 for definition of subregions.

Table 2. Code, name, test years, long-term winter mean temperature, mean yield, scaled score on the first GL interaction PC axis for log-transformed yield, and average phenotypic correlation between 2-yr genotype yield at the site and genotype yield across environments of the subregion including the site, for durum wheat test locations.

| Location | | Test years | Winter mean temperature† | Mean yield | PC 1 | r‡ | |
Code	Name					Subregion A	Subregion B
			°C	Mg/ha	$(\log_{10} \text{kg/ha})^{0.5}$		
1	Guelma	1,2,3	10.7	2.92	0.15	–	0.87
2	Souk Ahras	1,2,3	7.0	2.06	−0.14	0.57	–
3	El Khroub	1,3	8.3	3.22	0.02	–	0.79
4	Oum El Bouaghi	1,2,3	7.7	1.31	−0.10	0.80	–
5	Sétif	1,2,3	6.6	1.40	−0.14	0.73	–
6	EAC Dahal	1,2,3	5.1	1.87	−0.38	0.48	–
7	Beni Slimane	1,2,3	7.9	0.98	−0.06	0.53	–
8	Ain Bessam	1,2,3	8.0	1.51	−0.16	–	–
9	Oued Smar	1,2,3	11.9	3.53	0.58	–	0.90
10	Tipaza	1,2,3	11.6	1.68	0.11	–	0.79
11	Khemis Miliana	1,2,3	10.6	1.97	0.04	–	0.60
12	Djendel	1,2,3	10.5	2.91	−0.01	–	0.69
13	Tiaret Sebaine	2,3	7.5	1.68	−0.24	0.38	–
14	Rahouia	1,2,3	8.7	3.01	0.06	–	0.54
15	Tessala	1	9.8	1.60	–	–	–
16	Sidi Bel Abbes	1,2,3	10.0	2.10	0.20	–	0.78
18	Abdelkader	1,2,3	7.0	2.13	0.03	0.47	–

† Average of daily values in January and February.
‡ See Fig. 1 for definition of subregions. Site 8 (borderline) and Site 15 (one test year) were excluded from analysis.

logarithmic scale was used as a criterion for transforming data. A slope value around 2.0 suggested the need to log-transform data (Annicchiarico, 2002b, p. 53). Site classification was also performed by pattern analysis, standardizing entry yields within a location before cluster analysis (DeLacy et al., 1996). All cluster analyses adopted Ward's clustering method (De-Lacy et al., 1996) and a squared Euclidean distance as the dissimilarity measure. The AMMI + cluster analysis (the latter performed on site scores on the first PC, the only significant GL interaction PC axis) and pattern analysis provided similar classification results and were the best methods on the basis of predicted yield gains over the region (Annicchiarico, 2002a).

Site classification was extended to the total potential durum wheat cropping area through a discriminant analysis-based procedure. Long-term winter mean temperature (across January and February) for AMMI + cluster analysis and site altitude for pattern analysis were selected as significant ($P <$ 0.10) variables in the discriminant function for site groups (Annicchiarico et al., 2002a). The GIS-based definition of subregions provided by the two classification techniques was very similar (consistency >95%), showing a low-elevation, milder Subregion B (including 65–68% of the durum wheat cropping area) and a high-elevation, cold-prone Subregion A. Subregion definition by AMMI + cluster analysis was pre-ferred here because it was based on an environmental variable (the level of cold stress) that directly affected the adaptive responses of genotypes and it was somewhat more reliable in terms of variation between site groups accounted for by the discriminant function (Annicchiarico et al., 2002a). The critical winter mean temperature for discriminating a subregion was 8°C. Site 8, having this temperature (Table 2), was considered borderline and was not assigned to a subregion. Site 7, although close to the critical temperature (Table 2), was included in the cold-prone Subregion A together with Locations 2, 4, 5, 6, 13, and 18 (Fig. 1). Subregion B included Sites 1, 3, 9, 10, 11, 12, 14, 15, and 16.

Modeling of Adaptation Patterns and Verification of Site Classification

The analyses described in this section were performed on the 3-yr data set after excluding Site 15, which was represented by just one test year. The following analysis of variance (ANOVA) model was adopted, given the variable number of test years per site:

$$R_{ijkr} = m + G_i + L_j + Y_k(L_j) + B_r(Y_kL_j) + GL_{ij} + GY_{ik}(L_j) + e_{ijkr} \qquad [1]$$

Fig. 1. Geographical position of test sites within the Algerian durum wheat region and definition of two subregions that extends site classification by AMMI + cluster analysis through the application of a discriminant function to long-term winter mean temperature in a GIS (from Annicchiarico et al., 2002a; see Table 1 for site names).

where R_{ijkr} is the yield response of Genotype i in Location j, Year k, and Block r; m is the grand mean; G, L, Y, and B indicate the effects of genotype, location, years nested in locations, and blocks nested in years and locations, respectively; and e is random error. The GL interaction mean square was tested by using the average within-location genotype $\times$ year (GY) interaction mean square, under the assumption of year as random factor (Annicchiarico, 2002b, p. 23). The same model was used for estimating variance components for the region and individual subregions, assuming genotype and location were random factors.

Given the unavailability of climatic data for most sites in the third year, only AMMI and joint regression models could be compared for ability to describe GL interaction effects. Testing of GL interaction PC axes was performed by the F_R test as described by Annicchiarico (2002b, p. 38), using the within-location GY interaction as an error term. Heterogeneity of genotype regressions was tested by using deviations from regression.

The need for log-transforming data before classifying sites by AMMI + cluster analysis was verified in two steps. First, the size of the two determinants of the GE interaction variance, i.e., the lack of genetic correlation and the heterogeneity of genotypic variance among environments, was assessed together with the pooled genetic correlation among environments for entry yield. This was performed as described by Cooper et al. (1996), following ANOVAs for single experiments and a combined ANOVA including genotype, environment, and block factors. Then, the regression slope of the within-site phenotypic variance of entry values as a function of site mean yield on a logarithmic scale was estimated. Test sites were classified in two subregions by AMMI + cluster analysis and pattern analysis using the same procedures described for the 2-yr data set and results were compared with those of the GIS-based classification in Fig. 1.

The AMMI-modeled adaptive responses were based on original yields (i.e., the unit of yield gains) of top-ranking entries. They were graphically expressed as nominal yields as a function of the site score on the first GL interaction PC axis (Gauch and Zobel, 1997). Nominal yields sum the estimated entry mean yield and the product of the entry by the site scaled scores on PC 1 (excluding the site main effect, which is irrelevant for entry ranking). The consistency of current PC 1 scores for entries and locations with previous indications based on 2 yr of data (Annicchiarico et al., 2002a, 2002b) was verified by correlation analysis.

The software IRRISTAT, released by the International Rice Research Institute (IRRI) of Manila, was used for AMMI and joint regression modeling. All other analyses described here and in the next section were performed by SAS (1999) software. In particular, PROC GLM was used for different ANOVAs, while genotypic and genotype-environmental variance components were estimated by the maximum-likelihood method in PROC VARCOMP.

Definition of Selection Locations and Comparison of Adaptation Strategies

Comparing adaptation strategies based on observed or predicted yield gains requires some preliminary assumptions, such as the total number of selection environments (sites and years combinations) and the intensity of selection hypothesized for late selection stages (Annicchiarico, 2002a, 2002b). For specific breeding, sites can be assigned to subregions in proportion to their relative size, and selection is based on entry mean yield across selection environments independently for each subregion. For a wide-adaptation strategy, selection is based

on entry mean yield across all of the selection environments already defined for specific breeding. This is consistent with Lin and Butler's (1988) suggestion to choose selection sites across the region in a stratified manner and in proportion to the relative size of site groups. In this study, we hypothesized six selection environments over the region for late selection stages, assigning four environments (two sites in 2 yr) to subregion B (the larger) and two (one site in 2 yr) to subregion A (the smaller). Thus, selection for specific adaptation was based on 2-yr entry mean yields across selection environments independently calculated within subregions, whereas selection for wide adaptation was based on 2-yr entry mean yields across all six environments. Selection of three entries out of 24 (12.5%) was hypothesized in all cases.

Selection sites were chosen among those having 3 yr of data. The screening power of a given site for a target area is proportional to the phenotypic correlation between entry mean yields at the site and entry mean yields across the target environments (Cooper et al., 1996). Given the 2-yr period assumed for selection, entry mean yield at the site was always estimated from 2 yr of data, computing correlations for all possible pairs of test years at the site and holding the average correlation with target environments as an indicator of the screening power. Target environments included all environments in the relevant subregion as defined in Fig. 1, except those acting as selection environments in the assessment. To select the two selection sites for Subregion B, we also assessed the screening power of pairs of sites, avoiding pairs that were geographically very close. Site 15 was not assessed for screening power, but was included among the target environments of Subregion B here and in following analyses, whereas Site 8 (borderline between subregions) was excluded from analyses.

Observed yield gains for each adaptation strategy were computed in each subregion as the difference between the mean yield of the three entries selected in the relevant selection environments and the mean yield of all entries across the target environments belonging to the subregion. Yield gains were computed separately for each of the three pairs of test years at selection sites and then averaged. Target environments were represented by all relevant environments in the GIS-based definition of subregions, except those acting as selection environments. For each strategy, the average yield gain over the region was computed as a weighted mean of the average gains ΔG_A and ΔG_B observed within the two subregions (Annicchiarico, 2002a). Weights were the relative sizes of the Subregion A (P_A) and the Subregion B (P_B) as estimated in the GIS:

$$\Delta G = [(\Delta G_A P_A) + (\Delta G_B P_B)]/(P_A + P_B)$$
$$= (\Delta G_A P_A) + (\Delta G_B P_B). \quad [2]$$

Results were complemented by AMMI-modeled responses of selected entries and entry mean comparison across different target environments. Mean comparison, which was provided for reference (since entry selection was based on entry rank as usually done in breeding programs), used the GE interaction as the error term with $P < 0.20$. For the sake of simplicity, the comparison was relative to all possible target environments (including selection ones) and the 3-yr period.

Predicted yield gains per selection cycle, assuming adoption of the previously defined selection sites, were computed for the two strategies as correlated genetic gains in the target environments from the selection environments ($\Delta G_{T/M}$) according to Basford et al. (2004). In general

$$\Delta G_{T/M} = i r_{(S,T)} s_{p(T)} \quad [3]$$

where i is the standardized selection differential ($i \approx 1.56$ for

three selected entries out of 24: Falconer, 1989 p. 355), $r_{(S,T)}$ is the phenotypic correlation for entry mean yield between selected and target environments, and $s_{p(T)}$ is the estimated phenotypic standard deviation in the target environments. Yield gains in each subregion were predicted for each strategy separately for each of the three pairs of test years at selection sites and then averaged, providing the average gains ΔG_A and ΔG_B within individual subregions. Target environments were those belonging to the relevant subregion as defined in the GIS, excluding those acting as selection environments. For each set of target environments, $s_{p(T)}$ was estimated by the square root of the denominator of the formula for estimating the broad-sense heritability on an entry mean basis (h^2):

$$h^2 = s_g^2/(s_g^2 + s_{ge}^2/e + s_e^2/er) \qquad [4]$$

where s_g^2, s_{ge}^2 and s_e^2 are estimates of genotypic, GE interaction and pooled error variance components, respectively, across environments, and e and r are numbers of environments and replicates, respectively. The average yield gain over the region was computed for each strategy by Formula [2].

Adaptation strategies were also compared in terms of predicted yield gains per selection cycle as described by Annicchiarico (2002a). In this case, no prior choice of selection sites was assumed, generating results for sites with average screening power. The procedure implied the estimation of variance components for s_g^2, s_{ge}^2 and s_e^2 for three sets of target environments, i.e., all environments, for wide adaptation; and environments of Subregion A on the one hand and of Subregion B on the other, for specific adaptation. The average yield gain predicted over the region for a wide-adaptation strategy was computed by the formula:

$$\Delta G = ih^2 s_p \qquad [5]$$

where i is the same as in [3], and h^2 and s_p were estimated from Formula [4] and the square root of its denominator, respectively. However, e and r in Formula [4] represented the numbers of selection environments and experiment replications, respectively, hypothesized for selection (thus, $e = 6$ and $r = 4$), while variance components were relative to the whole set of environments. Predicted yield gains over Subregions A (ΔG_A) and B (ΔG_B) for a specific-adaptation strategy were computed by Formula [5], inputting in Formula [4] the estimates of variance components for the relevant subset of target environments and the appropriate e and r values ($e = 2$ for computing ΔG_A; $e = 4$ for computing ΔG_B; $r = 4$ in all cases). The average yield gain over the region predicted for specific breeding was derived from ΔG_A and ΔG_B through Formula [2].

RESULTS

AMMI vs. Joint Regression Modeling

The heterogeneity of genotypic variance components accounted for almost half of the GE interaction variance for original yields (Table 3). This finding, and the regres-

sion slope around 2.0 ($b = 1.99$, $P < 0.001$) of the within-site phenotypic variance of entry yields as a function of site mean yield on a logarithmic scale, confirmed the need to log-transform data before classifying sites by AMMI + cluster analysis. After transforming data, the heterogeneity of the genotypic variance components was much smaller relative to the lack of genetic correlation component (Table 3), and no relationship emerged between phenotypic entry variation and mean yield of locations ($b = -0.16$, $P > 0.20$). The transformation also tended to reduce the heterogeneity of experimental errors (which, expressed as variation coefficient, was reduced from 137–122%). The pooled genetic correlation between environments for entry yield was little affected by data transformation ($r_g = 0.44$ for original yields; $r_g = 0.37$ for transformed yields). Its fairly low value highlighted the large extent of GE interaction effects.

The following emerged for both original and log-transformed data from ANOVA, AMMI, and joint regression analysis results summarized in Table 4: (i) highly significant variation existed for genotype, GL interaction, and within-site GY interaction, (ii) a unidimensional AMMI model was adequate for describing the GL interaction variation, and (iii) the AMMI model was preferable to joint regression on the basis of the significant deviations from regression term ($P < 0.01$) and the higher mean square of PC 1 compared with heterogeneity of genotype regressions [the comparison based on mean squares takes account of the parsimony besides the accuracy of the model (Brancourt-Hulmel et al., 1997)]. These results justified the modeling of genotype-adaptive responses as a function of site score on the first GL interaction PC axis for original yields (Fig. 2), and the site classification by cluster analysis performed on the PC 1 score for transformed data. For both original and transformed data, PC 1 was a positive indicator of site mean yield ($r > 0.62$, $P < 0.01$) and tended to separate the sites of Subregion B, which had higher values than those of Subregion A (Table 2; Fig. 2). Principal Component 1 for transformed data confirmed the expected correlation with long-term winter mean temperature of locations ($r = 0.81$, $P < 0.001$), i.e., the discriminant variable used as a cut-off for defining the two subregions in Fig. 1. Entry PC 1 score separated the set of traditional cultivars and old varieties from improved varieties (Table 1). The former were

Table 3. Estimate of genotypic and genotype × environmental variance components for original and log-transformed yield of 24 durum wheat genotypes tested in 46 environments.

Variance component	Original	Transformed
	(Mg/ha)^2	(log$_{10}$ kg/ha)^2
Genotype	0.052	0.0023
GE interaction†	0.129	0.0056
Heterogeneity of genotypic variance	0.063	0.0018
Lack of genetic correlation	0.066	0.0038

† Always different from zero at $P < 0.001$.

Table 4. Genotypic and GE interaction ANOVA mean squares for original and log-transformed durum wheat yields, with GL interaction variation partitioned by (i) AMMI analysis and (ii) joint regression analysis.

Source of variation	DF	Original	Transformed
Genotype	23	9.88***	0.422***
GL interaction	345	0.96***	0.038***
(i) PC 1	37	5.66***	0.178***
Residual	308	0.39NS	0.021NS
(ii) Genotype regressions	23	4.65***	0.085***
Deviations from regression	322	0.70***	0.034**
GY interaction in location	690	0.48***	0.025***
Pooled error	4416	0.16	0.009

NS, Not significant at $P < 0.05$.
** Significant at $P < 0.01$.
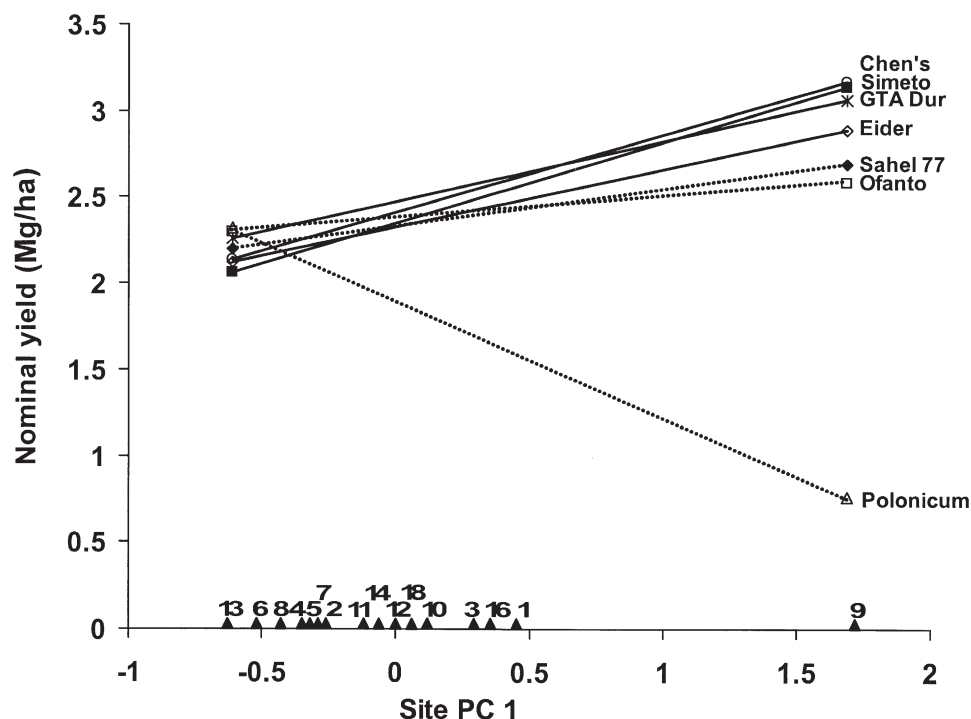*** Significant at $P < 0.001$.

Fig. 2. Nominal yield of best-performing durum wheat genotypes as a function of the score on the first GL interaction PC axis of 16 locations (site score is showed by black triangles; see Table 1 for site names).

specifically adapted to Subregion A on the basis of their highly negative score (providing a large positive AMMI-modeled GL effect when multiplied by a negative PC 1 score of sites) (Table 1). The modern varieties, however, showed relatively wide variation for adaptive response as expressed by their PC 1 score.

The site PC 1 score for original or transformed data agreed well with previous indications based on 2 yr of data ($r > 0.91$, $P < 0.001$). The consistency with 2 yr of data was even higher for the PC 1 score of genotypes ($r > 0.98$, $P < 0.001$).

Reliability of the GIS-Based Subregion Definition

The AMMI + cluster analysis and pattern analysis performed on the 3-yr data set provided the same subdivision of test sites into two groups. The groups coincided with those of the GIS-based subregion definition in Fig. 1, with the exception of Site 18, which was assigned to Subregion B. These results confirmed the reliability of the adopted definition of candidate subregions based on long-term winter mean temperature in the GIS.

Estimates of the GL interaction and within-site GY interaction variance components for original data over the region were 88 and 217%, respectively, of the genotypic variance. The zoning process allowed for a sharp reduction of the estimated GL interaction variance component, which was, on the average, only 24% of the genotypic variance within the individual subregions defined in Fig. 1.

Comparison of Adaptation Strategies

Site 4 was preferred to other candidate selection locations for Subregion A because of its top-ranking average

phenotypic correlation with target environments of the subregion (Table 2). Sites 1, 9, 10, and 16, among those of Subregion B having three test years, showed the highest phenotypic correlation with target environments of this subregion (Table 2). Considering all possible pairs formed by these sites, except that of Sites 9 and 10 [which was excluded because of the small geographical distance between these locations (Fig. 1)], the pair of sites 1 and 10 maximized the correlation with entry values across target environments ($r = 0.95$, $P < 0.001$) and was, therefore, preferred to select for Subregion B. Sites 1, 4, and 10 jointly acted as selection sites to breed for wide adaptation.

Cultivar mean yields over the region and across environments of either subregion are reported in Table 1. GTA Dur and Chen's were the top-yielding entries over the region and over Subregion B, whereas Ofanto and GTA Dur were top-yielding over Subregion A. The phenotypic correlation between the two subregions for entry mean yield ($r = 0.66$, $P < 0.001$) indicated that the consistency of genotype response between the subregions was only moderate.

In each of the three pairs of test years, the three genotypes selected for wide adaptation on the basis of mean yields across relevant selection environments were always coincident with those selected for specific adaptation to Subregion B. They were GTA Dur and Chen's in all cases, Simeto in two cases and Eider in one case. The set of entries selected specifically for Subregion A included Ofanto and GTA Dur in all cases, Sahel 77 in two cases and Simeto in one case. The wide-adaptation pattern of GTA Dur and the specific-adaptation pattern of most other selected entries are showed by AMMI-modeled nominal yields in Fig. 2 (in which

Table 5. Mean yield of selected durum wheat entries and average observed and predicted yield gains per selection cycle in two subregions and over the potential durum wheat region, for specific- and wide-adaptation strategies.

Strategy†	Mean yield‡		Observed gain‡§			Predicted gain			
	Subregion A	Subregion B	Subregion A	Subregion B	Region	Subregion A¶	Subregion B¶	Region¶	Region#
	Mg/ha								
Specific adaptation	1.899	3.031	0.233	0.372	0.327	0.199	0.509	0.409	0.316
Wide adaptation	1.833	3.031	0.167	0.372	0.305	0.181	0.505	0.400	0.304
Specific/wide ratio (%)	103.6	100.0	139.5	100.0	107.1	109.9	100.8	102.2	104.1

† Selected fraction: 3 entries out of 24. Total selection environments: 6 (3 sites by 2 years), of which 2 assigned to Subregion A (proportion of the region = 0.322) and 4 to Subregion B (proportion of the region = 0.678).
‡ Values averaged across three pairs of test years.
§ Gain computed as difference between the mean of selected entries and the mean of all entries.
¶ Relative to defined selection locations. Values averaged across three pairs of test years.
# Relative to undefined selection locations. Predicted gain from specific adaptation is 0.161 Mg/ha for Subregion A and 0.304 Mg/ha for Subregion B.

specific adaptation to Subregion A is highlighted by a dotted line). Sizeable crossover interactions between top-yielding genotypes only took place at negative PC 1 score values, where sites of Subregion A are located. The graph also shows the response of Polonicum, representative of those of traditional cultivars and old varieties, to show that this germplasm could become top-ranking only at sites with PC 1 scores that are extremely low and marginal even for the unfavorable subregion A.

Mean yields of selected entries, and observed gains relative to mean yield of all entries, are reported for the two strategies in Table 5 as average values across pairs of test years. The two strategies did not differ for Subregion B because of the coincidence of selected entries. Specific adaptation provided yield gains almost 40% higher than wide adaptation for Subregion A. This advantage implied just a 7% advantage over the region, owing to the smaller size and the smaller yield gains in absolute value of Subregion A (Table 5).

Predicted yield gains based on the defined selection locations are also reported in Table 5 as average values across pairs of test years. The advantage of specific adaptation over wide adaptation was around 10% for subregion A and was negligible for subregion B, resulting in a modest advantage (slightly over 2%) over the whole region. The comparison of strategies based on predicted yield gains over the region for undefined selection locations indicated an advantage of about 4% for specific adaptation over wide adaptation (Table 5).

DISCUSSION

The superiority of AMMI over joint regression for modeling genotype responses to locations was also noted for various cereals in Italy (Annicchiarico, 1997). In both Italy and Algeria (Annicchiarico et al., 2002b), site mean yield is mainly affected by local rainfall while adaptive responses mainly depend on winter cold stress as determined by earliness of heading. To be useful, joint regression requires that the same environmental factor be the main determinant of site mean yield and GL interaction effects.

The complete consistency of site classification between AMMI + cluster analysis and pattern analysis supports the reliability of site groups determined by using these techniques. The former technique, which is more complex than the latter, may be preferred when

modeling adaptive responses. Modeling genotype responses by pattern analysis in its ordination mode is also possible (e.g., Cooper et al., 1996), but, compared with AMMI, it suffers from inability to remove random error from the pattern (Gauch and Zobel, 1997; Annicchiarico, 2002b, p. 50).

The substantial consistency between ordinary classification of sites (by AMMI + cluster analysis or pattern analysis) and extended classification based on 2 yr of data as a function of the long-term winter mean temperature in the GIS is comforting and reinforces interest of procedures for up-scaling results from test sites to all sites within a region. The inconsistent classification of Site 18, assigned to the cold-prone Subregion A on the basis of its long-term winter mean temperature and to Subregion B on the basis of site similarity for GL effects, was already observed for 2 yr of data and attributed to the warm winters that occurred at the site during the test years (Annicchiarico et al., 2002a). Indeed, the up-scaling procedure may also compensate for the effect of unusual years (temporal up-scaling), besides improving the geographical definition and the estimation of the size of subregions (spatial up-scaling). It requires, however, that a sufficiently reliable discriminant function for site groups be defined on the basis of available environmental variables. Earlier GIS-based definitions of subregions relied on different approaches. Pollak and Corbett (1993) based it on site similarity for environmental factors that were supposed to be important (although chosen with no definite assessment of their impact on GL effects). Chapman and Barreto (1996) interfaced a GIS with simulation results for flowering date of different germplasm types. The current GIS has also been interfaced with factorial regression or AMMI-modeling of adaptive responses for cultivar recommendations, defining areas with the same top-yielding material (Annicchiarico et al., 2005).

Each of the procedures adopted for comparing adaptation strategies in terms of yield gains may have some limitations. Observed gains relate to the portion of GL effects that are of practical interest to breeders, i.e., the crossover interaction between top-ranking entries. But they may be more affected than predicted gains by small random errors, which may easily alter entry ranks. Procedures based on defined selection sites may be more realistic for breeding programs, but the estimated screening power of environments may largely depend on specific conditions during the test years that lead to particu-

larly high or low GY interaction and/or experimental error. The Basford et al. (2004) procedure, in particular, was proposed for managed selection environments, where screening power is more reliably assessed because it is less affected by specific conditions during test years. Predicted gains for undefined selection sites, on the contrary, relate to the average screening power of environments within the subregion as estimated from the sample of environments.

In conclusion, the three procedures provide complementary information for comparing strategies, while being inappropriate for estimating actual selection gains that can be expected, unless the evaluation data are based on random sample of breeding lines. The results suggest that a specific-adaptation strategy may provide a modest advantage (likely around 4–5%) over wide adaptation at similar costs. The distinctly larger efficiency showed by specific adaptation for Subregion A (10–39%) suggests for this strategy the additional merit of increasing the food security at the household level in this stressful subregion. The only moderate consistency of entry response between the two subregions ($r = 0.66$) did not result in a large advantage of specific breeding over the region because the large within-site GY interaction that is typical of Mediterranean regions limits the advantage of this strategy. Specific adaptation was largely superior to wide adaptation for barley (*Hordeum vulgare* L.) in northern Syria in a larger study (Ceccarelli et al., 1998) in which, however, the same test sites were assigned to a favorable or an unfavorable subregion depending on their mean yield in specific years, thereby reducing the impact of within-site GY interaction effects on the assessment.

Two aspects of the current assessment may have led to underestimating the relative merit of specific adaptation. Information on the actual size of each subregion, based on the proportion of the potential durum wheat area that is devoted to the crop, was lacking. If available, it would likely have increased the relative size of Subregion A (because of the paucity of alternative crops in this area) and the impact over the region of the distinctly larger efficiency of specific adaptation for this subregion. The paucity of material specifically adapted to Subregion A probably limited the advantage of specific breeding. Finally, it should be noted that all comparisons of strategies relied on the hypothesis of performing late selection on the same pool of material in the two subregions. The advantage of specific breeding may be much greater if it implied the early selection of breeding lines to either subregion on the basis of crucial adaptive traits, or the use of a distinct genetic base for each subregion. A previous study provided indications on relevant adaptive traits (Annicchiarico et al., 2002b). The AMMI-modeled yield responses in Fig. 2, possessing higher predictive ability than observed responses (Gauch, 1992), may be exploited for identifying parent material of specific interest for each subregion.

On the whole, our findings reinforce the potential of a specific-adaptation strategy for national breeding programs of medium-sized countries, in which it may contribute to increasing national yields as well as local yields in less favorable areas. Especially for poorer research systems, a relevant issue concerns the extent of experimental data and statistical effort that are required to obtain reliable indications on adaptation strategies. The less demanding procedure implying 2 yr of data, subregion definition by pattern analysis, and estimation of the relative size of subregions from the proportion of test sites that they accounted for (Annicchiarico et al., 2002a) showed an advantage of specific over wide adaptation based on predicted yield gains (11%) that was fairly close to that based on observed gains (7%). This finding encourages the basic analysis of relatively small, multilocational, multiyear data sets to obtain preliminary indications on the possible benefit of specific breeding.

REFERENCES

Annicchiarico, P. 1997. Joint regression vs. AMMI analysis of genotype-environment interactions for cereals in Italy. Euphytica 94: 53–62.

Annicchiarico, P. 2002a. Defining adaptation strategies and yield stability targets in breeding programmes. p. 165–183. *In* M.S. Kang (ed.) Quantitative genetics, genomics and plant breeding. CABI, Wallingford, UK.

Annicchiarico, P. 2002b. Genotype × environment interactions: Challenges and opportunities for plant breeding and cultivar recommendations. FAO Plant Production and Protection Paper No. 174. Food and Agriculture Organization, Rome.

Annicchiarico, P., T. Chiari, G. Delli, S. Doucene, N. Yallaoui-Yaïci, and F. Bellah. 2002a. Response of durum wheat cultivars to Algerian environments. IV. Implications on a national breeding strategy. J. Agric. Environ. Int. Develop. 96:227–259.

Annicchiarico, P., T. Chiari, F. Bazzani, F. Bellah, S. Doucene, N. Yallaoui-Yaïci, Z. Abdellaoui, B. Belloula, L. Bouazza, L. Bouremel, M. Hamou, T. Hazmoune, M. Kelkouli, H. Ould-Said, and H. Zerargui. 2002b. Response of durum wheat cultivars to Algerian environments. II. Adaptive traits. J. Agric. Environ. Int. Develop. 96:189–208.

Annicchiarico, P., F. Bellah, and T. Chiari. 2005. Repeatable genotype × location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria. Eur. J. Agron. (in press).

Basford, K.E., W.T. Federer, and I.H. DeLacy. 2004. Mixed model formulations for multi-environment trials. Agron. J. 96:143–147.

Bramel-Cox, P.J., T. Barker, F. Zavala-Garcia, and J.D. Eastin. 1991. Selection and testing environments for improved performance under reduced-input conditions. p. 29–56. *In* D. Sleper et al. (ed.) Plant breeding and sustainable agriculture: Considerations for objectives and methods. CSSA Special Publ. 18. ASA, CSSA, SSSA, Madison, WI.

Brancourt-Hulmel, M., V. Biarnès-Dumoulin, and J.B. Denis. 1997. Points de repère dans l'analyse de la stabilité et de l'interaction génotype-milieu en amélioration des plants. Agronomie 17:219–246.

Ceccarelli, S. 1989. Wide adaptation: How wide? Euphytica 40:197–205.

Ceccarelli, S. 1996. Positive interpretation of genotype by environment interactions in relation to sustainability and biodiversity. p. 467–486. *In* M. Cooper and G.L. Hammer (ed.) Plant adaptation and crop improvement. CABI, Wallingford, UK.

Ceccarelli, S., S. Grando, and A. Impiglia. 1998. Choice of selection strategy in breeding barley for stress environments. Euphytica 103: 307–318.

Chapman, S.C., and H.J. Barreto. 1996. Using simulation models and spatial databases to improve the efficiency of plant breeding pro-

grammes. p. 563–587. *In* M. Cooper and G.L. Hammer (ed.) Plant adaptation and crop improvement. CABI, Wallingford, UK.

Conway, G. 1998. The doubly green revolution. Food for all in the 21th century. Penguin, London.

Cooper, M., I.H. DeLacy, and K.E. Basford. 1996. Relationships among analytical methods used to study genotypic adaptation in multi-environment trials. p. 193–224. *In* M. Cooper and G.L. Hammer (ed.) Plant adaptation and crop improvement. CABI, Wallingford, UK.

DeLacy, I.H., K.E. Basford, M. Cooper, J.K. Bull, and C.G. McLaren. 1996. Analysis of multi-environment data– An historical perspective. p. 39–124. *In* M. Cooper and G.L. Hammer (ed.) Plant adaptation and crop improvement. CABI, Wallingford, UK.

Delli, G., P. Sarfatti, F. Bazzani, and A. Cadi. 2002. Application of GIS for agro-climatological characterization of northern Algeria to define durum wheat production areas. J. Agric. Environ. Int. Develop. 96:121–140.

Engledow, F.L. 1925. The economic possibilities of plant breeding. p. 31–40. *In* Proc. Imperial Botanical Conf. F.T. Brooks, London.

Falconer, D.S. 1989. Introduction to quantitative genetics. 3rd ed. Longman, New York.

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

Gauch, H.G. 1992. Statistical analysis of regional yield trials: AMMI analysis of factorial designs. Elsevier, Amsterdam.

Gauch, H.G., and R.W. Zobel. 1997. Identifying mega-environments and targeting genotypes. Crop Sci. 37:311–326.

Hardwick, R.C., and J.T. Wood. 1972. Regression methods for studying genotype-environment interaction. Heredity 28:209–222.

ITCF-IAO. 2002. Amélioration et renforcement du système nationale d'adaptation variétal du blé dur–Resultats des essais de comportement variétal. Institut Technique des Grandes Cultures and Istituto Agronomico per l'Oltremare, Alger and Florence.

Kang, M.S. 1998. Using genotype-by-environment interaction for crop cultivar development. Adv. Agron. 62:199–252.

Lin, C.S., and G. Butler. 1988. A data-based approach for selecting locations for regional trials. Can. J. Plant Sci. 68:651–659.

Pollak, L.M., and J.D. Corbett. 1993. Using GIS datasets to classify maize-growing regions in Mexico and Central America. Agron. J. 85:1133–1139.

SAS Institute Inc. 1999. SAS/STAT user's guide–Version 8. SAS Institute Inc., Cary, NC.

Simmonds, N.W. 1979. Principles of crop improvement. Longman, London.

Wallace, D.H., R.W. Zobel, and K.S. Yourstone. 1993. A whole-system reconsideration of paradigms about photoperiod and temperature control of crop yield. Theor. Appl. Genet. 86:17–26.

Xu, Y. 2002. Global view of QTL: Rice as a model. p. 109–134. *In* M.S. Kang (ed.) Quantitative genetics, genomics and plant breeding. CABI, Wallingford, UK.

Yau, S.K. 1991. Need of scale transformation in cluster analysis of genotypes based on multi-location yield data. J. Genet. Breed. 45:71–76.