CHAPTER 17

Breeding for quantitative variables Part 4: Breeding for nutritional quality traits

Yusuf Genc, Julia M. Humphries, Graham H. Lyons and Robin D. Graham

17.1 INTRODUCTION

Malnutrition is the most important cause of mortality in the global human population. More than a quarter of children less than five years old suffer from protein-energy malnutrition, as determined by rates of stunting and underweight. Of these, 70 percent are in Asia, 26 percent in Africa and 4 percent in Latin America. Stunting in resource-poor populations is usually associated with reduced mental development (Stephenson, Latham and Otteson, 2000).

Dietary deficiencies of the micronutrients iron (Fe), zinc (Zn), vitamin A (in the form of pro-vitamin A carotenoids), selenium (Se) and iodine (I) are widespread globally, affecting well over half of the world's population, and often occur concurrently (WHO, 2003). These deficiencies increase the risk of severe disease in approximately 40 percent of the world's population (Graham, Welch and Bouis, 2001). Se, Fe, Zn and vitamins A, B and C have immunomodulating functions and thus influence the susceptibility of a host to infectious diseases and their courses and outcomes (Bhaskaram, 2002; Failla, 2003).

Most of the Fe in the body occurs in combination with proteins as the oxygen-carrying pigments haemoglobin in red blood cells and myoglobin in muscle cells. Deficiency results in reductions in haemoglobin (anaemia) and tissue Fe (myoglobin and Fe-containing enzymes), and in lethargy (Jones, 1997).

Zn is a component of over three hundred enzymes, involved in carbohydrate metabolism, DNA synthesis, protein synthesis and digestion, and bone metabolism. Deficiency can result in reduced growth rate, skin lesions and increased susceptibility to infection (Jones, 1997).

Vitamin A deficiency is a major cause of blindness, growth retardation and increased susceptibility to infection. It commonly occurs in association with protein and Zn deficiency (Wahlqvist, 1997). This chapter will deal with the plant precursors to vitamin A: carotenoids such as β-carotene, and the non-provitamin A carotenoids, including lutein and zeaxanthin. Carotenoids are responsible for many of the orange, red and yellow colours seen in plants and animals. A small proportion of the over 600 named carotenoids are precursors to vitamin A, and are essential for the prevention of vitamin A deficiency. β-carotene has the highest provitamin A activity. Carotenoids that are not precursors to vitamin A also have an important role in health and nutrition as antioxidants and in the maintenance of sight, and those most commonly found in staple foods are lutein and zeaxanthin. Lutein and zeaxanthin are abundant in maize, and lutein is the dominant carotenoid in both bread and pasta wheat.

Se is an integral component of at least three systems required for normal cell metabolism, and has antioxidant, anti-cancer and anti-viral effects (Arthur, 1999). I is involved in growth, development and metabolic regulation, through its role as a component of thyroid hormones (Hetzel and Pandav, 1996). Moreover, interactions between Se and I are important in the body. Both micronutrients are required for thyroid hormone synthesis, activation and metabolism, and the thyroid gland has the highest Se and I concentrations of all organs (Kohrle, 1999).

HarvestPlus is a Biofortification Global Challenge Program of the Consultative Group on International Agricultural Research (CGIAR). It is coordinated by the International Centre for Tropical Agriculture (CIAT) and the International Food Policy Research Institute (IFPRI). Genetic bio-fortification is a strategy of

breeding staple crops such as rice, wheat, barley, maize, cassava, potatoes and beans with the ability to fortify themselves with micronutrients. It offers a sustainable, cost-effective alternative to other strategies such as individual supplementation and fertilization, which is more likely to reach those most in need and has the added advantage of requiring no change in current consumer behaviour to be effective (Graham, Welch and Bouis, 2001). Once a one-off investment is made to breed bio-fortified seed, recurrent costs are low and germplasm can be shared globally. Bio-fortification, commercial fortification and supplementation are complementary strategies for reaching malnourished populations. Furthermore, bio-fortification can increase farm productivity as certain micronutrients, such as Zn and Se, that improve human nutrition can help plants resist diseases and other environmental stresses (HarvestPlus, 2007).

Breeding criteria for micronutrient-enriched staple food crops have been reviewed recently (Welch and Graham, 2004). These criteria include (i) maintaining crop productivity, (ii) evidence for stability of micronutrient enrichment traits across various edaphic and climatic zones, (iii) demonstration of significant effects of enriched micronutrients on human health, (iv) demonstration of bio-availability of enriched micronutrients for human nutrition, and (v) consumer acceptance.

In this chapter, we will focus on genetic potential, genotype × environment interactions, screening protocols, breeding strategies for enhancing grain micronutrient accumulation. Physiological and molecular mechanisms of uptake, translocation and deposition of micronutrients in the grains or other edible parts of major staple food crops such as wheat, rice, maize, beans and cassava, which are consumed by billions of people in resource-poor nations will also be discussed. Sufficient genotypic variation in the trait to be selected is necessary for conventional breeding to be feasible, so we will discuss this as a first step, with reference to the five key micronutrients. Breeding principles discussed in this chapter are applicable to both traditional and participatory plant breeding.

17.2 GENOTYPIC VARIATION OF MICRONUTRIENT CONCENTRATION IN STAPLE FOOD CROPS

17.2.1 Fe & Zn

Over the last decade, there have been considerable efforts in several international research centres such as the International Maize and Wheat Improvement Centre (CIMMYT) in Mexico and the International Rice Research Institute (IRRI) in the Philippines, to identify wheat and rice germplasm with high Fe and Zn concentration, which has also been the subject of several reviews (Graham *et al.*, 1999; Rengel, Batten and Crowley, 1999; Cakmak *et al.*, 2000).

Wheat

At CIMMYT, in one study, 170 wheat lines selected out of 550 initially screened lines were grown in a replicated trial (Ortiz-Monasterio and Graham, 2000). This study identified three promising sources of high grain Fe and Zn concentration: wild species, landraces and breeding lines. Fe concentration was in the range of 25 to 56 mg/kg dry weight (DW), while Zn concentration varied from 25 to 65 mg/kg DW. In a second study, a group of 154 lines from the breeding programme were grown together. Lines were identified with up to 73 mg Fe/kg DW and 92 mg Zn/kg DW. Our group at the University of Adelaide also observed significant variation in grain Zn and Fe concentrations in commercial cultivars and advanced breeding lines

TABLE 17.1
Zn and Fe concentrations (mg/kg DW) in grains of durum and bread wheat genotypes grown in standard potting mix with all nutrients supplied adequately in a glasshouse

	No. of entries	Fe		Zn	
		Mean conc. (SD)	Range	Mean conc. (SD)	Range
Modern bread wheat (*T. aestivum*)	25	36 (6)	27–53	39 (8)	25–53
Synthetic hexaploid wheat (*T. aestivum*)	36	41 (8)	32–67	41 (9)	28–66
Durum wheat (*T. dicoccon*)	24	42 (7)	29–56	51 (6)	39–62
(*T turgidum*)	191	33 (8)	17–62	30 (12)	12–81

Note: Standard deviation, SD, is given in parentheses.
Source: Genc *et al*., unpublished.

TABLE 17.2
Zn and Fe concentrations (mg/kg DW) in grains of bread wheat genotypes in field trials in Australia

Location and year	No of entries	Fe		Zn	
		Mean conc. (SD)	Range	Mean conc. (SD)	Range
Birchip-2000	28	37 (3)	31–41	16 (2)	12–19
Birchip-1999	42	38 (2)	32–42	25 (3)	20–31
Birchip-1998	39	42 (4)	36–55	23 (2)	19–31
Horsham-1998	30	33 (3)	27–40	16 (1)	13–19
Bute-1997	42	32 (3)	27–38	18 (2)	15–21
Lameroo-1996	35	33 (3)	27–39	20 (2)	15–24

Note: Standard deviation, SD, is given in parentheses
Source: Graham *et al.,* unpublished.

of wheat in glasshouse and field trials conducted over a number of years in Australia (Tables 17.1 and 17.2). In general, grain Zn and Fe concentrations were higher in the glasshouse than in field studies, which can be attributed to better growing conditions (well-watered and fertilized) in the glasshouse than in the field. In the field studies, grain Zn concentration was in the range of 12 to 31 mg/kg DW, a narrower range than that found at CIMMYT. This narrower range and also lower values (<15 mg/kg DW) in grain Zn concentration are indicative of Zn deficiency in the field. Grain Fe concentration varied from 27 to 55 mg/kg DW. Recently, a large-scale screening by Cakmak *et al.* (2004) has identified wild wheat accessions with even higher Fe and Zn concentrations than those reported previously. In this comprehensive study of 825 accessions, including wild emmer wheat (*Triticum turgidum* subsp. *dicoccoides*), grain concentrations were 14 to 190 mg/kg DW and 15 to 109 mg/kg DW for Zn and Fe, respectively. In this study, no yield data were reported, thus we do not know whether high concentrations are associated with low yield. In the meantime, despite lower Fe and Zn concentrations than in wild species, more screening is needed of elite germplasm (modern wheat genotypes and advanced breeding lines) for high Fe and Zn concentration in the grain, as they have already improved agronomic performance (Graham *et al.*, 1999).

Rice

There also exists considerable genotypic variation for grain Zn and Fe concentration in rice. Researchers at IRRI and the

University of Adelaide (Australia) have evaluated a large set of brown rice varieties (1138), including breeding lines and wild rice and its derivates, and observed a wide range in Fe (6–24 mg/kg DW) and Zn (14–58 mg/kg DW) concentrations in the grain (Gregorio *et al.,* 2000). Some of these high-Fe and-Zn varieties were further tested alongside the two most popular cultivars in Asia, IR36 and IR64, in the same soil and year. The traditional variety, Jalmagna, had much higher grain Fe and Zn concentration than high-yielding IR36 (22 vs 12 mg Fe/kg DW; 32 vs 21 mg Zn/kg DW in Jalmagna and IR36, respectively). Moreover, aromatic rices are often reported to have higher grain Fe and Zn concentration than non-aromatic rices (Graham, Senadhira and Ortiz-Monasterio, 1997; Gregorio *et al.,* 2000).

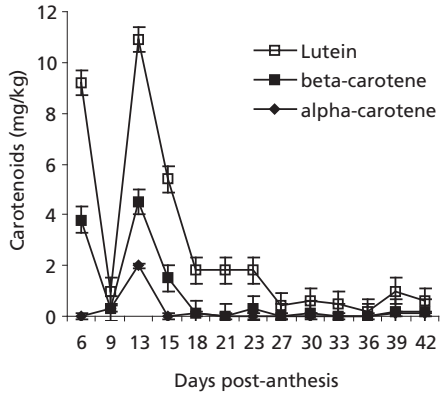
There is some evidence that breeding for either high Fe or Zn may also result in higher concentrations of other nutrients, as there is occasionally a significant positive correlation between Fe and Zn concentrations in the wheat and rice grain (Ortiz-Monasterio and Graham, 2000; Graham, Senadhira and Ortiz-Monesterio, 1997; Genc *et al.,* unpublished). This is also evident in a recent study by Vasconcelos *et al.* (2003), who introduced a soybean *ferritin* gene into *indica* rice, and found much higher Fe and Zn concentrations in the grains of transgenic plants compared with non-transgenic plants. However, as no seed weight data were provided, we do not know whether these high concentrations were associated with small seed size (concentration effect). Nevertheless, our calculations of their data established a significant positive correlation between grain Fe and Zn ($r^2 = 0.54$).

17.2.2 Vitamin A

The potential for finding genetic variation that can form the basis for breeding crops with increased carotenoid concentrations is great, given that all photosynthetic organisms have substantial concentrations of these compounds. However, in many staple crops it is necessary for the plant to store carotenoids in non-photosynthetic tissues, such as the tuber of the sweet potato, or in tissues that no longer have a photosynthetic capacity when harvested, as in wheat and maize. It is possible that the consumed portion of the crop that once had photosynthetic capacity may retain the carotenoids accumulated during the photosynthetic period post-degradation of the chlorophyll. However, in root crops, carotenoids must accumulate in non-photosynthetic tissues, and therefore need to be transported there from other photosynthetic tissues, or synthesized *de novo*.

A report by Graham and Rosser (2000) compared the synthesis patterns during maturation of both lutein and β-carotene in bread and durum wheat varieties (Figure 17.1). These results concur with an earlier report of Lacroix and Lier (1975), and indicate that a potential benefit may be gained by harvesting wheat at the immature (green) stage. However, appropriate storage and preservation methods are necessary in order for immature wheat to be stored for any period of time without spoilage. Such a method has been used for centuries in Middle Eastern countries, where wheat is harvested green and dry roasted to produce a product called *freekeh*. Substantial amounts of both β-carotene and lutein can be conserved from the photosynthetic stage in the roasted product by this method (Humphries and Khachik, 2003). This method of storage may be a valuable starting point for adoption into local cultures for

FIGURE 17.1
Time trend of synthesis and metabolism of various carotenoids in the caryopsis of wheat (var. Krichauff) from anthesis to maturity



Source: Graham and Rosser, 2000. Reprinted with permission from the UNU Food and Nutrition Bulletin (2000) 21: 404–409.

preservation of carotenoids present in immature wheat.

Selection against highly pigmented varieties in several staple foods has led to very little variation for this trait in modern cultivars. However, germplasm banks where landraces and old varieties are stored hold the key to retrieving genetic variation. These valuable resources can be used as a source of variation for the introduction of desirable traits back into commercially profitable and locally grown varieties, for nutritional benefit. Reports from screening of genetic resources obtained from germplasm banks indicate that many of the older varieties have substantial variation for carotenoid concentration.

Maize

White maize was previously highly desired for reasons of cleanliness and apparent purity, while maize varieties with high concentrations of pigmentation were used as stock feed. While people suffered from vitamin A deficiency, their stock remained healthy due to consumption of yellow maize (Brunsen and Quackenbush, 1962). Collaboration between nutritionists and agriculturalists resulted in the production of high-β-carotene maize adapted to local conditions, resulting in a reduction in vitamin A deficiency-associated diseases.

There have been several reports of genetic variation for carotenoids in maize. One of the first was that of Brunsen and Quackenbush (1962), who showed that the total concentration of carotenoids varied significantly between high- and low-carotene inbred lines, with an even greater variation between low and high for provitamin A carotenoids. Blessin *et al*. (1963) reported a range of 0.9 to 4.1 mg/kg for carotenes, and 18.6 to 48 mg/kg for xanthophylls in 39 maize inbreds. In the same year, Quackenbush *et al*. (1963) observed concentrations of up to 7.3 mg/kg carotenes, and a range of 2 to 33 mg/kg lutein in 125 inbred lines. The value of extensive screening for high-accumulation varieties was indicated in the report of Egesel (1997), who found a range of just 0.13 to 2.9 mg/kg β-carotene in 200 maize families. More recently, 16 yellow seeded maize lines were reported to have 143 to 278 mg/kg carotenoids (Maziya-Dixon *et al*., 2000). This study measured total carotenoid concentrations rather than defining provitamin and non-provitamin carotenoids, and although valuable for calculation of total carotenoid intake, gives no idea of the provitamin A potential of the cultivars.

Wheat

Wheat is another staple food that has been subjected to selection for colour, though with different outcomes depending on the end use, either bread or pasta. Bread wheat varieties (*Triticum aestivum*) have been

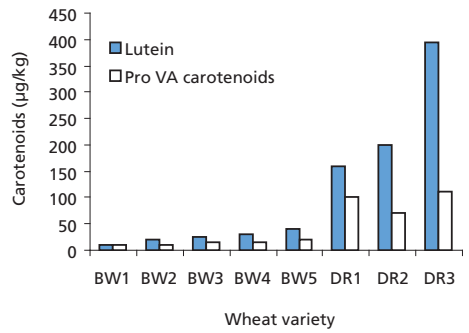
the subject of selection against pigmentation, though ironically, despite breeding efforts, it was still deemed necessary to use chemical bleaching rather than plant breeding alone to achieve the bright white colour demanded by bread wheat consumers. The process of bleaching has now been discontinued. However, the generations of selection against pigmentation has resulted in current cultivars containing very low concentrations of carotenoids. In Australia, there is an exception, a South Australian wheat by the name of Krichauff, which has relatively high concentrations of carotenoids in comparison with other current cultivars. It is also possible that the wheat used to make tortillas in South America may have substantial carotenoid concentrations that could be exploited for nutritional gain.

Matus-Cadiz *et al.* (2003) reported that the lutein concentration of 79 diverse spring wheat varieties from Australia and Canada ranged between 1.8 and 3.7 mg/kg, and also reported genotype × environment (G×E) effects that will be discussed in the G×E section of this chapter. An extensive survey of bread wheat varieties (Humphries *et al.*, 2004) revealed considerable genetic variation for both provitamin and non-provitamin A carotenoid concentrations in germplasm from the CIMMYT germplasm bank. Results from the range of concentrations obtained from bread wheat combined with those for durum wheat varieties are given in Figure 17.2.

Alternatively, the importance of colour in durum wheat (*Triticum durum*) used for pasta, has led to extensive studies into carotenoid concentrations. Several of these studies are presented below, and together with other reports not included here, indicate that despite the apparent potential for a source of high concentrations of carotenoids in different cultivars, concentrations of provitamin A carotenoids are usually low. It appears that selection for the colour provided by lutein has led to varieties with an abundance of the enzymes responsible for hydroxylation of the provitamin A carotenoids, and consequently the durum wheats are not thought to be a useful source of genetic variation for increased provitamin A carotenoids.

Some of the first reports of the dominance of non-provitamin A carotenoids in durum wheat were published in 1935 (Markley and Bailey, 1935a, 1935b), and later it was confirmed that only a small proportion of carotenes were present in comparison to lutein (Munsey, 1938). This was followed up by Zechmeister and Cholnoky (1940) and Lepage and Sims (1968), who reported lutein ester concentration and no provitamin A carotenoids. A more recent evaluation of the carotenoid composition of durum wheat (Hentschel *et al.*, 2002) revealed a range of lutein concentrations from 1.5 to 4 mg/kg, and no carotenes.

FIGURE 17.2
Carotenoid concentrations of five bread wheat varieties (BW) and three durum wheat varieties (DR), representing the range of values found in collections from the germplasm banks in CIMMYT and ICARDA



Source: Graham and Rosser, 2000. Reprinted with permission from the UNU Food and Nutrition Bulletin (2000) 21: 404–409.

Crosses with barley

Tritordeum lines, which are a cross between the wild barley species *Hordeum chilense* and diploid, tetraploid and hexaploid wheat have consistently shown higher concentrations of carotenoids than wheat (Alvarez, Urbano and Martin, 1994), and are considered a useful source for increasing the carotenoid concentration of durum wheat. However, using this cross to increase concentrations is dependent on interactions between the genetics of the parents, and the final concentrations cannot be reliably predicted.

Ancient wheat

Einkorn (*T. monococcum*), an ancient diploid wheat, has been reported to have yellow coloration (D'Egidio, Nardi and Vallega, 1993; Abdel-Aal, Hucl and Sosulski, 1995; Borghi *et al*., 1996). When compared with other ancient wheat varieties, spelt (*T. aestivum* subsp. *spelta*), emmer (*T. turgidum* subsp. *dicoccum*), Kamut (*T. turgidum* subsp. *turanicum*) and Khorasan (*T. turgidum* subsp. *turanicum*) einkorn lines generally had higher concentrations of lutein (mean 8.1 ± 0.26 µg/g) (Abdel-Aal *et al*., 2002).

Cassava

The carotenoid concentration of cassava roots has been closely correlated to the intensity of the root colour. However, within groups of the same tuber colour, genotypic variation has also been reported, from 6 to 24 mg/kg fresh weight (FW) (Chavez *et al*., 2000). This variation within colour types necessitates individual analyses to determine individual concentrations, and colour alone cannot be relied upon to give accurate estimations of concentrations.

One of the largest reported analyses of cassava for carotenoid concentration was conducted by Iglesias *et al.* (1997), who screened a total of 632 accessions from the CIAT germplasm bank collection of 5500. The distribution of concentrations ranged from 1 to 24 mg β-carotene/kg FW. Those varieties with the deepest coloration towards orange had the greatest concentration of carotenoids, which is consistent with the report of Chavez *et al*., (2000).

The variability for carotene concentrations in cassava has been reported for accessions obtained from germplasm banks in India (Moorthy *et al*., 1990) and Brazil (Ortega-Florez, 1991). The highest concentrations were below 8 mg of β-carotene equivalents/kg FW, which is one-third the highest concentration reported by Iglesias *et al*. 1997. However, the potential for rapidly increasing carotene concentrations using recurrent selection was reported by Jos *et al*. (1990). Using this method, it is possible to increase the concentration by three times, from 4.2 to 13.8 mg/kg FW after 2 cycles of selection and recombination (Jos *et al*., 1990). It is therefore possible, in theory, to obtain concentrations of β-carotene up to 72 mg β-carotene/kg FW.

17.2.3 Se & I

To address dietary Se deficiency, agronomists and plant breeders have adopted complementary strategies to develop crops with higher Se content. The first is an agronomic (fertilizer) approach, discussed elsewhere (Lyons *et al.*, 2003, 2004; Broadley *et al.*, 2006). The second strategy is to develop varieties with improved Se accumulation and tolerance traits by either conventional breeding or genetic modification. To implement this approach, a comprehensive characterization of the interactions between Se and sulphur nutrition was conducted in *Arabidopsis* (White *et al.*, 2004). If sufficient genotypic variation exists in Se accumulation within a crop species, and if this variation is heritable, conventional plant breeding could

provide an alternative to agronomic biofortification and thus minimize the need for Se fertilizers (Broadley *et al.*, 2006).

Few data have been published on varietal differences for Se accumulation for most crop species. However, in *Lycopersicon* (tomatoes and related plants), four-fold differences in shoot Se accumulation have been found (Pezzarossa *et al.*, 1999), and in *Brassica* (broccoli) a significant genotype effect for Se concentration in heads was found in hybrids, but not inbreds. However, the effect of environment was around ten times stronger than that for genotype (Farnham *et al.*, 2007).

Wheat

In surveys and trials conducted by our group, involving diverse wheat germplasm and a total of eleven datasets in South Australia and Mexico, grain Se concentrations were in the range of 5 to 720 µg/kg DW, but much of this variation was associated with spatial variation in soil-available Se. South Australian soils are renowned for their microspatial variability, which makes detection of genotypic differences in grain Se density difficult. No significant genotypic variation in grain Se density among modern commercial bread or durum wheat varieties was detected in this study (Lyons *et al.*, 2005), which agrees with earlier research (Noble and Barry, 1982; Grela, 1996; Tveitnes, Singh and Ruud, 1996). However, the ancient diploid wheat, *Aegilops tauschii* L. and rye (*Secale cereale*) were found in our studies to be 42 percent and 35 percent higher ($p < 0.001$; $p = 0.03$), respectively, in grain Se concentration than other cereals in separate field trials, and in a hydroponic trial rye was 40 percent higher *(*$P < 0.001$) in foliar Se content than two wheat landraces. *Ae. tauschii* was also higher in Zn, Fe and Mn than other wheats in the trial. Other wild wheat relatives may also be efficient accumulators of Se and other minerals (Piergiovanni *et al.*, 1997).

Studies of genotypic variation of I in diverse plant species have yielded variable findings. A Japanese survey found no significant difference in leaf I concentration for plants grown on similar soils (Yuita *et al.*, 1982), while an earlier survey of different pasture species grown together in the field found a thirty-four-fold variation in leaf I concentration, with perennial ryegrass (*Lolium perenne* L.) the most efficient I accumulator (Johnson and Butler, 1957).

Evidence is scarce for significant genotypic variation in grain density of I in wheat (Shinonaga *et al.*, 2001), and no variation was detected in our South Australian trial, where varieties were grown at three locations, with three replications. I concentrations were low, typically less than 20 µg/kg (range 10–60 µg/kg DW) in whole grain (Lyons *et al.,* unpublished).

Maize

Little research has been conducted on genotypic variation in Se or I concentration in maize kernels. Our group has conducted a limited survey of diverse maize genotypes grown in the United States of America (Illinois) and Nigeria. No significant variation was detected for either micronutrient; however, the soils at both sites were low in available Se, resulting in kernel Se levels at the Nigerian site, for example, of just 5 µg/kg DW (range 3–10 µg/kg DW). Such low levels may have limited expression of possible varietal Se differences (Lyons *et al.*, unpublished).

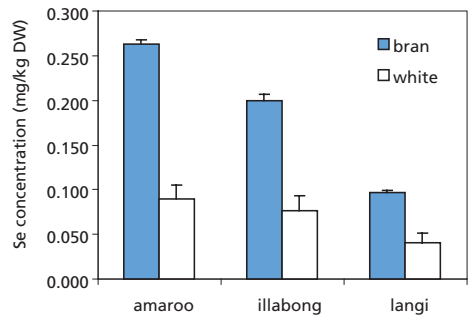
Rice

Rice appears to be a more promising cereal for genotypic variation in Se, with differences detected in other studies (Nan and

Han, 1993; Zhang *et al.*, 2004) as well as our own, which involved several varieties grown together in New South Wales, Australia. Three varieties differed in Se concentration in bran (means [SE] of 97 [12], 200 [17], 263 [14] µg/kg DW), while one was lower than the others in endosperm Se concentration (40 [2], 83 [5] µg/kg DW) ($P < 0.001$) (Lyons *et al.*, 2005b; Figure 17.3).

Genotypic variation in I concentration in rice bran was apparent in our Australian study, with all three cultivars different ($p = 0.02$). The mean I concentrations in the bran [SE] in µg/kg DW were 910 [50], 770 [10], and 500 [50]. There was no difference in I concentration in the endosperm for the three cultivars (Lyons *et al.*, unpublished).

FIGURE 17.3
Effect of genotype on Se concentration in rice bran and white (polished) rice. Bars represent means of three replications ± SE



Source: Lyons *et al.*, 2005b.

17.3 GENOTYPIC VARIATION IN STAPLE CROPS FOR DISTRIBUTION OF MICRONUTRIENTS WITHIN EDIBLE COMPONENTS

17.3.1 Fe & Zn

Wheat

Almost all studies to date have dealt with nutrient concentration in the whole grain, and there are few data available on the distribution of micronutrients in different grain fractions. Lyons *et al.* (2005c) studied distribution of grain Fe, Zn and other nutrients across the seed tissues of four wheat genotypes, and reported that the proportion of Zn in the grain (percent of total grain Zn content) was in the following order: endosperm (including aleurone layer) > embryo > bran (72–78, 11–27 and 2 percent, respectively). The proportion of Fe in the grain fractions followed the same order as Zn (79–91, 3–19 and 2–9 percent for endosperm, embryo and bran, respectively).

Rice

The highest proportion of Fe was located in pericarp, including aleurone layer (43 percent), followed by endosperm (42 percent) and embryo (10 percent) (Boyd *et al.*, 1972). The proportion of Zn was 60, 42 and 14 percent in pericarp (including aleurone layer), endosperm and embryo, respectively. It is interesting to see that when all proportions for Zn are added up, we get a value of over 100 percent, which may be due to contamination during dissection, processing or analytical errors associated with a very small sample size (Boyd *et al.*, 1972).

The distribution of nutrients in cereal grains is important for human nutrition. For obvious reasons, high concentration in the endosperm is desirable as endosperm makes up the majority of the grain. It is well known that a significant proportion of nutrients is lost in milling residue (Burk and Solomons, 1985). A further reduction in nutrient content occurs in the polishing process in rice (Graham *et al.*, 1999). Therefore, in breeding programmes, selection for higher Fe and Zn concentration in the endosperm would not only result in lower losses of Zn and Fe, but enhance bio-availability due to lower levels of phytate and fibre

components in endosperm than in bran (Lyons *et al.,* 2005c). Further research with a large number of genotypes is required to determine the extent of genotypic variation in distribution of Zn in the grain, and also to assess the potential for breeding for this trait.

17.3.2 Se & I

As noted above, genotypic variation was found by our group in rice bran and, to a lesser extent, endosperm for Se, but only in bran for I. Se concentration in rice bran was around three times that in endosperm, while I concentration in bran was around nine times that in endosperm (mean 730 vs 80 µg/kg DW) (Lyons *et al.,* 2005b). As most rice is eaten in the polished form, with the bran removed, selection for a higher proportion of grain Se and I stored in endosperm may be worthwhile, as noted for Zn and Fe above.

17.5 GENOTYPE × ENVIRONMENT INTERACTIONS AND THEIR EFFECTS ON MICRONUTRIENT TRAITS

G×E interaction is an important issue for plant breeders. A significant G×E interaction implies that rankings of genotypes differ with environment, indicating the need for testing at various sites or seasons. For example, a genotype may exhibit high-micronutrient-density traits in one environment, but not in others. A significant G×E interaction can be classified as either (i) non-crossover, where the ranking of genotypes remains consistent in different environments and it is the degree of accumulation that is affected; or (ii) crossover, where the rank of individual cultivars is affected by the environment (Baker and Kosmolak, 1977). G×E interactions and statistical methods for analysis and interpretation of G×E interactions have been reviewed elsewhere (Kang, 1990; Hill, Becker and Tigerstedt, 1998; Annicchiarico, Chapter 20 this volume). From a breeding point of view, it is important to understand the nature and extent of G×E interactions for designing breeding strategies and selection procedures (Eiseman, Cooper and Woodruff, 1990). Much effort has been directed to understanding G×E interactions in relation to yield, but little attention has been given to grain nutrient density. Graham *et al.* (1999) recognized that environment (soil type, fertilizer management and climate) can have a strong influence on nutrient density of grains; thus it is important to grow out the seed to be compared for at least one generation in the same environment to minimize the variation in nutrient density associated with previous growing conditions. Only then can valid comparisons of genetically controlled variation be made.

Fertilizer management can also influence micronutrient density in the grain. In studies with maize varieties grown at different nitrogen (N) and water levels, Feil *et al.* (2005) found that N fertilization reduced grain concentrations of Zn and Mn, which was attributed to higher grain yield as a result of N application. In contrast, continuous irrigation did not affect grain nutrient density. However, the rankings of varieties remained unchanged by water regime and N levels, pointing to stability of varietal differences in grain nutrient density over a range of N and water levels. At present, there is little information available on the effects of fertilization on grain micronutrient density in rice or wheat.

17.5.1 Fe & Zn

Wheat

From field trials conducted by our group in different locations and years

TABLE 17.3
Fe and Zn concentrations (mg/kg DW) in grains of wheat cultivars grown at different locations and years in Australia

Cultivar	Zn concentration						Fe concentration					
	Lameroo 1996	Bute 1997	Birchip 1998	Horsham 1998	Birchip 1999	Birchip 2000	Lameroo 1996	Bute 1997	Birchip 1998	Horsham 1999	Birchip 1999	Birchip 2000
Excalibur	22	18	21	17	25	18	37	32	44	35	40	37
Krichauff	20	19	25	18	24	18	36	34	46	38	39	38
Songlen	24	20	25	19	31	16	37	35	45	38	42	35
Trident	18	17	25	19	23	17	35	32	45	35	38	38
Yallaroi	20	17	22	16	24	14	32	32	45	36	40	37
Mean	21	18	24	18	26	17	35	33	45	36	40	37
Standard deviation	2	1	2	1	3	2	2	1	1	2	2	1

Source: Graham *et al.,* unpublished.

TABLE 17.4
Fe and Zn concentrations (mg/kg DW) in grains of wheat cultivars grown at two levels of Zn fertilization at Lameroo (1996) and Bute (1997) in South Australia

Cultivar	Zn concentration						Fe concentration					
	Lameroo			Bute			Lameroo			Bute		
	Nil	+Zn	Mean*	Nil	+Zn	*Mean	Nil	+Zn	*Mean	Nil	+Zn	*Mean
Barunga	14	22	18	15	20	17	38	36	37	33	31	32
Cascades	14	21	17	11	20	15	38	35	36	35	31	33
Excalibur	14	22	18	11	18	14	36	37	36	34	32	33
Frame	12	19	15	12	19	15	32	31	32	35	34	35
Halberd	13	23	18	12	17	14	35	33	34	36	35	36
Janz	12	18	15	11	16	13	29	29	29	29	27	28
Krichauff	12	20	16	12	19	15	38	36	37	37	34	35
RAC750	12	22	17	13	21	17	33	31	32	31	31	31
RAC809	11	19	15	10	16	13	36	35	35	34	30	32
RAC812	11	17	14	12	20	16	33	31	32	35	34	34
RAC820	15	22	18	10	15	13	33	33	33	34	32	33
RAC826	13	21	17	12	18	15	33	34	34	36	33	34
RAC832	13	21	17	14	24	19	30	31	31	38	35	37
RH911996	15	21	17	10	16	13	32	33	32	31	31	31
RH912025	14	20	17	11	18	14	36	31	34	33	35	34
Songlen	14	24	19	11	20	15	36	37	37	36	35	35
Tammin	14	20	17	10	17	14	37	35	36	34	31	32
Trident	12	18	15	11	17	14	35	35	35	35	32	33
WI334	11	17	14	11	16	14	33	33	33	35	33	34
WI94063	11	18	15	13	19	16	31	30	30	29	30	30
WI94091	10	14	12	12	18	15	29	27	28	32	29	31
Yallaroi	17	21	19	12	17	15	34	32	33	32	32	32
Yanac	11	17	14	11	16	13	31	30	31	35	31	33
Mean	13	20		12	18		34	33		34	32	
Range	10–17	14–24		10–15	18–24		29–38	27–37		29–38	27–35	

NOTES: Genotype × Zn fertilization interaction for grain Zn and Fe concentrations was non-significant, while genotype × location interaction was significant ($LSD_{0.05}$=3 for Zn and Fe). +Zn treatment received granular zinc (zinc oxysulphate, 32 percent Zn) at a rate of 7 kg/ha at seeding and a foliar spray (Zincsol, 16.7 percent Zn) at a rate of 2 L/ha at the early growth stage.
Source: Graham *et al.,* unpublished.

in Australia, Zn application resulted in an increase in grain Zn concentration in all varieties and sites. However, few genotypes in these trials were retained year after year, thus G×E interactions could not be analysed for all environments. However, when we analysed grain Zn and Fe data for the five genotypes tested in 6 environments (Table 17.3) or 23 genotypes tested in two environments (Table 17.4), responses of genotypes differed with environments (locations and years), indicating the presence of G×E interactions. When we subjected the data for grain Zn concentration (adequate Zn only) in Table 17.5 to Spearman's Rank Correlation Test (r_s), we found a non-significant correlation between rankings of genotypes in two different environments (r_s = 0.223), suggesting that rankings of genotypes differ with environment.

Most recently, significant G×E interactions were also reported for grain Zn and Fe concentrations, which would make direct selection for these traits difficult (Oury *et al.*, 2006). The ranges in grain Zn and Fe concentrations of adapted material in this study (15–35 and 20–60 mg/kg DW for Zn and Fe, respectively) were wider than those found in our field study (Table 17.5), which might be attributed to differences in genotypes and environments between the two studies. These limited studies suggest that there is a need for further field trials at different locations and years to determine or confirm the extent and nature of G×E interactions and their effects on grain Zn and Fe concentrations in wheat.

Rice

It was reported that high Fe and Zn traits were expressed in all rice environments (Graham *et al.*, 1999) and G×E interactions were sufficiently moderate (Gregorio *et al.*, 2000), suggesting that breeding for high Fe and Zn traits is a worthwhile effort. However, there was some evidence of G×E interaction in extreme environments. Although these limited studies are encouraging, there is clearly a need for further studies in this area.

17.5.2 Vitamin A

No G×E effect on lutein concentration was reported by Matus-Cadiz *et al.* (2003) when they investigated the effect of genotype, year and location in Australian and Canadian wheat varieties. However, further statistical analysis of the data revealed that 12 of the 79 cultivars showed significant crossover genotype-by-year interactions, indicating that in different years those cultivars reported changed in lutein concentrations that affected their rank.

17.5.3 Se

While genotypic differences may exist in modern wheat varieties, they are likely to be small in comparison with background soil variation. Soil Se is uneven in distribution and availability, with total Se concentrations ranging from less than 0.1 to more than 100 mg/kg DW (Berrow and Ure, 1989). Areas that are notably low in Se include parts of China, Siberia, central Africa, eastern Europe and New Zealand (Combs, 2001). In studies of grain Se concentration in wheat grown in South Australia, our group has found substantial microspatial (that is, metre-to-metre) variation in levels of available Se in soils. For example, at one trial site near Bordertown, south-east of Adelaide, we found a six-fold variation in grain Se concentration in four replicates of one wheat cultivar, grown together in the same field (Lyons *et al.*, 2004). Hence

the detection of what may be relatively small (for example, 10 percent) genotypic variations in Se uptake efficiency between wheat cultivars under these field conditions is virtually impossible. Background soil variation in available I has also been found to be substantial at the South Australian sites we have used, although less so than for Se. This large microspatial variation in soils makes it difficult, if not impossible, to accurately assess genotypic differences across environments for Se and I. This and narrow genotypic variation reported so far may be the reasons why to date there have been no studies reported on G×E interactions for Se and I.

17.6.SCREENING AND ANALYTICAL METHODS FOR MICRONUTRIENTS IN FOOD CROPS

Where should screening be carried out: field or greenhouse? The principles of both controlled environment and field screening are reviewed elsewhere (Graham, 1984), and therefore will not be dealt with in detail here.

17.6.1 Fe & Zn

The data presented in Table 17.4 suggest that screening for grain Zn concentration should be carried out in optimal growing conditions, as variation in grain Zn concentration under Zn deficient conditions is rather narrow. Unlike traits such as agronomic Zn efficiency, screening at the early growth stage does not appear to be suitable for detecting or identifying genotypes with the ability to load more Zn into the grain, due probably to the overriding importance of re-mobilization of Zn from leaves into grain occurring towards maturity. The evidence for this comes from a study in barley (Lonergan, 2001) in which two of the four chromosomal regions (also known as Quantitative Trait Loci, QTL) identified were found to co-segregate for grain Zn accumulation and vegetative Zn accumulation at anthesis, indicating little prospect for screening for grain Zn accumulation even as early as anthesis. Moreover, the only QTL detected for shoot Zn concentration or content (chromosome 4) did not co-segregate with Zn concentration or content at either anthesis or maturity, suggesting that screening for grain Zn accumulation at the early stage will not be reliable or relevant to grain Zn accumulation.

Inductively Coupled Plasma Optical Emission Spectrometry (ICPOES) is commonly used to determine mineral nutrient concentrations in plant tissues, and allows the determination of interactions among the various essential nutrients, effects that can be large and important. This method is fast and reliable, but can be costly for breeding programmes in both developed and developing countries, where tens of thousands of samples are handled each year. Are there alternative and cheaper methods to ICPOES? The researchers at the University of Adelaide (Australia) have developed a rapid, cheap and user-friendly assay for determination of Fe in the grain of rice or wheat (Choi, Graham and Stangoulis, 2007). This new cost-effective assay consists of two phases. In phase one, the assay is used to identify high grain-Fe lines from thousands of samples, while in phase two, the high-Fe lines identified in phase one are confirmed by ICPOES.

17.6.2 Vitamin A

Several standard procedures for extraction and identification of carotenoids from plant material have been used since the discovery and naming of carotene in the early 19th century. Separation of carotenoids initially involved a two-step chromatographic

method involving open-column and thin-layer chromatography (TLC). These two methods have been combined in high performance liquid chromatography (HPLC), which is now the preferred method for carotenoid analysis.

Spectrophotometric analysis, TLC and HPLC all require extensive extraction procedures using organic solvents that are both costly and toxic. While there is no doubt that these methods are necessary for elucidation of specific isomers and absolute quantitative analysis, this reduces the scope for identification of high-carotenoid parent lines. Given the participatory focus of this book, a fast and accurate method of identifying high carotenoid concentrations would vastly increase the number of lines that could be screened to identify suitable parents by persons with little organic chemistry background.

Spectrophotometric determination of wheat grain xanthophyll concentration following extraction of flour or meal with water-saturated butanol is well established (AACC, 1983). Similarly, reflectance spectrophotometric measurement of flour colour is commonly used (Oliver, Blakeney and Allen, 1992), as is the relationship between Commission Internationale l'Eclairage (CIE) b* and extractable yellow pigments (Mares and Campbell, 2001). Colour determined by CIE classifies colour in three dimensions: L*, brightness; a* red to green colour; and b* yellow to blue colour. CIE colour is influenced by inherent genotypic characteristics, environmental conditions and stresses during grain production, the milling procedure and by the size of flour particles and bran flakes, which is caused by differences in grain hardness and moisture content of the grain at milling. Variation in L* affects the measurement of b* and potentially could result in errors in estimating carotenoid content (Mares and Campbell, 2001).

Current methods for the identification of wheat genotypes high in specific carotenoids involve HPLC and are slow, costly and highly labour intensive. The chemical structure of carotenoids indicates that a correlation with colour is likely and it is therefore possible that divergent selection for colour in bread and pasta wheat has influenced the carotenoid content of these species. Determination of a correlation between a fast and accurate colour measurement, such as that obtained from the Minolta Chroma Meter, and carotenoid concentration determined by HPLC, could vastly increase the number of samples that could be screened in a given period.

In a recent report (Humphries *et al.*, 2004), whole-meal wheat, including both bread and durum varieties, and triticale samples were analysed for their carotenoid content by HPLC, and also for colour using reflectance spectrophotometry (CIE L*a*b*). A positive correlation between CIE b* (yellowness) and lutein concentration was shown in all wheat groups, but was strongest in the durums. There was little correlation between CIE L* (lightness) or CIE a* (redness) and lutein, α- or β-carotene. By contrast, the b* value correlated well with the concentration of α- and β-carotene, and therefore the vitamin A activity, though those wheat groups that did not have a strong correlation were those with the lowest CIE b* values. The durum wheat had the highest CIE b* value and the highest lutein concentration, but a relatively low concentration of β-carotene.

17.6.3 Se & I

Because of the high soil variation in available Se and I, screening for higher

Se and I traits in cereals needs to include hydroponic trials and pot trials using a standardized growth medium, backed up by field studies conducted on soils that are relatively uniform in Se and I. Selenate is the most mobile Se form and the dominant available Se form in well-aerated, neutral to alkaline soils (Cary and Allaway, 1969), while selenite is the major form taken up by rice in flooded paddy soils of lower pH (Wang and Gao, 2001). Thus the composition of hydroponic culture media needs to be tailored to the relevant field situation. Using solution culture containing selenite as the dominant available Se form, Zhang *et al.* (2004) have found genotypic variation in Se concentration in the leaves of rice seedlings, and the levels are well correlated with those in grain.

Genotypic variation in I uptake in rice may be explained by differences in the oxidising power of the roots, which can oxidise the iodide ion to form molecular I, which is then absorbed more readily. A significant correlation was found between the oxidising power of rice roots and the uptake of I (Yamada *et al.,* 2005), hence this may prove to be a suitable screening method.

Commonly used methods of Se analysis include hydride ICPOES, ICP mass spectrometry, and fluorimetry. Sample preparation for hydride ICPOES involves digestion with a nitric+perchloric acid mixture, followed by hydrochloric acid digestion, then treatment with sodium borohydride (Tracy and Moller, 1990). ICP mass spectrometry, in which plasma is used as the ionization source, is a highly sensitive method for Se and I analyses (Hieftje and Vickers, 1989). Another commonly used (and time-honoured) method for both Se and I analysis is the fluorimetric method. This is based on the reaction of 2,3-diaminonaphthalene (DAN) with Se (IV) to form a fluorescent Se-DAN complex, piazselenol (Koh and Benson, 1983). Samples for I analysis are typically prepared using tetramethyl ammonium hydroxide (TMAH) extraction.

17.7 INHERITANCE OF MICRONUTRIENT ACCUMULATION IN FOOD CROPS

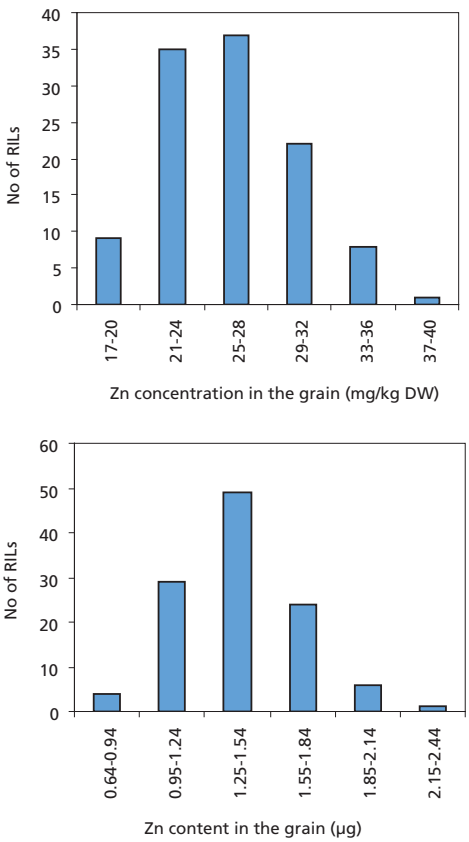
Apart from the existence of genetic variation, breeding for enhanced grain nutrient content also requires knowledge of genetic control mechanisms.

17.7.1 Fe & Zn

Wheat

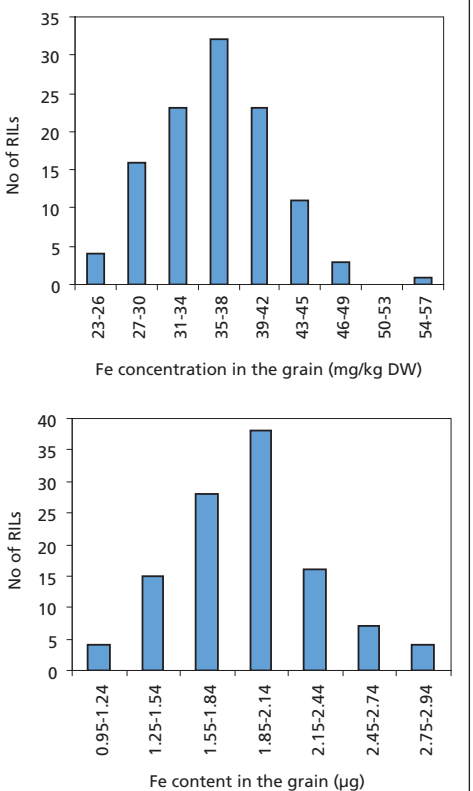
At present, there is little or no information available on genetics of Zn and Fe accumulation in wheat grain. The continuous variation in grain Zn and Fe concentration and content within Recombinant Inbred Lines (RIL) (n=113) derived from Opata × Synthetic cross (Figures 17.4 and 17.5) indicates that the two traits are quantitative and controlled by several genes (Genc *et al.,* unpublished). It is interesting to note that some RILs had higher Zn and Fe concentration and content than either of the parents, suggesting a transgressive segregation, which is probably due to these lines carrying favourable allele combinations from both parents. This multi-gene control hypothesis is supported by recent studies that identified several chromosomal regions associated with grain Zn concentration (Shi *et al*., 2007; Genc *et al*., 2009) and content (Shi *et al*., 2007). However, these field studies were conducted at a single location and QTLs identified were mapped to either different chromosomes or to different regions of the same chromosome. Therefore, there is a need to test mapping populations at multiple sites and years to validate the QTLs and also to determine the extent of GE interactions on grain Zn and Fe

FIGURE 17.4
Frequency distribution of grain Zn concentration and content in Recombinant Inbred Line population of Opata × Synthetic grown in standard potting mix in a glasshouse (Zn concentration, 24 and 20 mg/kg DW; Zn content, 1.29 and 1.08 µg/grain for Opata and Synthetic, respectively)



Source: Genc *et al.*, unpublished.

FIGURE 17.5
Frequency distribution of grain Fe concentration and content in Recombinant Inbred Line population of Opata × Synthetic grown in standard potting mix in a glasshouse (Fe concentration, 33 and 30 mg/kg DW; Fe content, 1.74 and 1.63 µg/grain for Opata and Synthetic, respectively)



Source: Genc *et al.*, unpublished.

traits. Identification and validation of QTLs associated with high grain Fe and Zn traits will accelerate breeding for these complex traits. Marker-assisted selection is discussed in Chapter 19 of this volume.

Rice

Genetic analysis of grain Fe concentration using four traditional varieties, three advanced lines and three IRRI-released varieties revealed the presence of a large genetic effect (additive and non-additive gene action) and small environmental effects (Gregorio *et al.*, 2000). Narrow-sense and broad-sense heritabilities were 44 percent and 88 percent, respectively. This study suggested that selection for a high grain-Fe trait should be delayed as late

as F_5 generation where dominance effect is not evident. This study also identified three chromosomal regions associated with a high grain-Fe trait (chromosomes 7, 8 and 9), providing evidence for multi-gene control for this trait, in which case, selection as late as F_5 may mean that lines bearing the fullest expression of high Zn concentration may be few in number and so lost in earlier generations unless populations are large.

17.7.2 Vitamin A

Maize

It is important to be aware of reciprocal differences in their contribution to kernel content of carotenoids, as this will affect the inheritance of the traits. Several studies have shown that the pollen parent affects carotenoid concentrations of the F_1 seed of reciprocal crosses (Mangelsdorf and Fraps, 1931; Johnson and Miller, 1938; Randolf and Hand, 1940; Grogan *et al.*, 1963)

However, a study by Egesel (2001) found that the female parent had the greatest influence on carotenoid concentrations in open pollinated kernels. Another study reported broad sense heritability of 33 percent for β-carotene and 47 percent for β-cryptoxanthin (another carotenoid with provitamin A activity) (Wong, 1999).

To produce colour in the maize kernel, numerous genes are necessary for structural and regulatory mechanisms. For carotenoid production, three genes have been reported to be relevant to carotenoid concentration. The *Y1* (yellow 1) gene on chromosome 6 encodes for phytoene synthase (Buckner *et al.*, 1996), an essential enzyme in the carotenoid pathway; the *VP9* (viviparous 9) gene on chromosome 7 is associated with ζ-carotene desaturase; and the *VP5* gene encodes for phytoene desaturase (Wong *et al.*, 2004). Carotenoids are produced in the starchy endosperm, and because of this, the yellow or white colours are only seen when the aleurone layer is colourless. The endosperm is triploid in nature, two from the maternal parent and one from the paternal parent. Thus the maternal parent can give a good indication of the expected carotenoid concentration (Egesel *et al.*, 2003). This results in a heterozygote containing two dominant or two recessive alleles, resulting in phenotypic differences (Table 17.5). For example, in the monohybrid cross for *Y1*, the endosperm can be one of four genotypes, which produce variation in colour (Symcox, Shadley and Weber, 1987). The genotypic differences correlate to the colour of the endosperm, and to the carotenoid concentration.

TABLE 17.5
Genotype and phenotype of maize heterozygote, and contributions from paternal and maternal parents

Endosperm genotype	Maternal contribution	Paternal contribution	Phenotype
Y1Y1Y1	Y1Y1	Y1	Yellow
Y1Y1*y*1	Y1Y1	*y*1	Light yellow
*y*1*y*1Y1	*y*1*y*1	Y1	Pale yellow
*y*1*y*1*y*1	*y*1*y*1	*y*1	White

Cassava

It was initially reported that the inheritance of root colour was simple and that a single dominant gene was responsible for yellow colour (Hershey and Ocampo, 1989). However, since that report, inheritance for carotenoid concentration has been found to be under the control of two genes, one responsible for transport to the non-photosynthetic roots, the other for accumulation within these storage organs (Chavez *et al.*, 2000). These two genes do not function independently of each other, rather each affects the expression of the other, though the mechanisms behind this are as yet unreported. The two major genes

are also combined with other genes with smaller effects that affect accumulation.

Wheat

A study into the genetic origin of an increase in carotenoid pigments in the cross between a wild barley (*Hordeum chilense*) and durum wheat located this trait to the α-arm of chromosome 7H^{ch} (Alvarez, Martin and Martin, 1998). Screening of 35 lines with various Tritordeum lines revealed that although the presence of the H^{ch} genome is responsible for increased carotenoid concentrations in these lines it is difficult to predict the effect of the interaction between the barley and wheat genetics. *H. chilense* is therefore a useful but not entirely reliable source of increased carotenoid concentrations for *T. durum* (Alvarez, Martin and Martin, 1999).

17.8 PHYSIOLOGICAL AND MOLECULAR MECHANISMS OF MICRONUTRIENT UPTAKE, TRANSLOCATION, RE-MOBILIZATION AND ACCUMULATION

17.8.1 Fe & Zn

For a breeding programme to be successful, it is important to understand the processes leading to accumulation of nutrients in the grain. Obviously an increase in accumulation of Fe and Zn in the grain of any plant species will require higher uptake, translocation or re-mobilization from source (leaves) to sink (grain). The role of these processes in relation to accumulation of Fe, Zn and other micronutrients has been reviewed recently (Grusak, Pearson and Marentes, 1999); thus it will not be discussed in detail here. It is interesting to note that Fe has been the most studied micronutrient in rice, while Zn has been researched to a larger extent in wheat. There is some suggestion that increasing the levels of Fe-chelating agents (phytosiderophores, nicotianamine, organic acids), reducing agents (ferric reductase), enzymes and transport proteins in the root cells could enhance Fe uptake and transport (Grusak, Pearson and Marentes, 1999). However, higher uptake and transport does not necessarily imply higher accumulation in the grain (phloem loading). For example, a pea mutant of cultivar Sparkle (*brz*) accumulated 36-fold higher Fe in the leaves compared to Sparkle, but did not have higher Fe in the seeds (Grusak, 1994). The author concluded that the rate limitation to phloem Fe loading was due to an unidentified ligand species that would complex with Fe prior to phloem loading rather than the availability of Fe as substrate. This study was followed by the study of Marentes and Grusak (1998), who demonstrated that a second mutant of cultivar Sparkle (*dgl*) had 2.5-fold higher Fe concentration in the embryo compared to Sparkle (163 and 65 mg/kg DW, respectively). This mutant also had higher Fe concentration in the seed coat. The authors used radiotracer ^{59}Fe to determine the movement of Fe in the seed coat, and found that Fe was symplastically phloem loaded. They further suggested that Fe resided within the non-vascular seed coat cells, and that the cells at the inner surface of the seed coat may facilitate the release of Fe to the embryo apoplast. The form of Fe in the seed coat or embryo is still not known at present.

A recent study in rice suggests that nicotianamine (NA) and nicotianamine synthase (NAS) genes (OsNAS1, OsNAS2 and OsNAS3) are also involved in long-distance transport of Fe (Inoue *et al.*, 2003), apart from their roles in the release of low-molecular weight compounds, phytosiderophores, from the roots of graminaceous plants. This release of phytosiderophores solubilizes rhizospheric

Fe(III) thus increasing plant uptake of Fe, for example in rice (Takagi, 1976; Takahashi *et al.,* 2001). Most recently, a rice metal-NA transporter, OsYLS2, has been linked to phloem transport and translocation to the grain of Fe (Koike *et al.,* 2004). Increasing the expression of transporters such as OsYSL2 and OsNASs in rice and other species can enhance Fe, and to some extent Zn, accumulation in the grains. However, further studies are required to determine the extent of genotypic variation in NA concentrations and its relative contribution to Fe and Zn accumulation in several varieties with low and high Fe and Zn in the grain.

In contrast to Fe, there is little information available on Zn translocation in the plant and transport to the grain. An earlier study (Longnecker and Robson, 1993) suggested that organic complexes with citrate and malate may be important in re-mobilization of Zn in the phloem. A recent study indicates that this may not be the case, as no relationship could be found between the presence of complexes or ligands and loading of Zn into the wheat grain (Pearson *et al.,* 1996). However, this result does not rule out the possibility of other endogenous chelates that may play a role in the long-distance transport of micronutrients (Welch, 1995). At the same time, transport to the grain is thought to occur predominantly via the phloem (Pearson *et al.,* 1995), and is well regulated (Herren and Feller, 1997). This regulation of Zn transport has been reported by Pearson, Rengel and Graham (1999), who suggested that low-Zn grain was not a strong sink for Zn, while in the case of high-Zn grain, there may be a barrier preventing excessive accumulation in the grain. It has also been suggested that phytosiderophores and nicotianamine may facilitate Zn uptake and transport in the plant (Scholz, Seifert and Grun, 1987; Treeby, Marschner and Romheld, 1989; Zhang, 1991; Cakmak *et al.,* 1996), but this needs to be established in future studies.

The positive correlations between Fe and Zn concentrations in cereal grains provide evidence that both nutrients may be taken up and translocated to the grain through the same process. However, as the correlation coefficient is never 1, there must be other mechanisms specific to Fe or Zn uptake or translocation to the grain, which warrants further investigation.

Having briefly discussed uptake and translocation of Fe and Zn to the grain, two questions arise: "How much of the Fe and Zn in the plant ends up in the grain?" and "Where are these nutrients deposited in the grain?" Contrary to earlier suggestions of poor re-mobilization of Fe from vegetative tissues to the grain (4 percent in rice, Marr *et al.,* 1995; 20 percent in wheat, Miller *et al.,* 1994), a recent growth-room study with wheat reported that 77 percent of total shoot Fe was re-mobilized to the grain at maturity (Garnett and Graham, 2005). The lower re-mobilization of Fe in the earlier studies involving field-grown plants were attributed to (i) precipitation of Fe in the apoplasm (inactive Fe) at high concentrations, which may result in non-re-mobilization, (ii) saturation of either the grain loading or phloem loading, and (iii) contamination of plant tissues by soil (references in Garnett and Graham, 2005). Miller *et al.* (1994) reported that, in wheat, 70 percent of Zn in the leaves was re-mobilized to the grains, while only 42 percent of shoot Zn re-mobilized to the grain in the study of Garnett and Graham (2005). The differences in the amounts of Zn re-mobilized in these two studies may be due to differences in genotypes and experimental conditions. It has been suggested that,

in wheat, relatively large amounts of Zn are transported into crease or inner pericarp tissues via the crease phloem, and translocation to the embryo and endosperm continues throughout grain development (Pearson *et al*., 1998). As Zn status of the grain improves, more Zn is distributed to the inner pericarp and less Zn to the endosperm, outer pericarp and embryo (Pearson, Rengel and Graham, 1999).

17.8.2 Se & I

In most plants, uptake, transport and assimilation of selenate is the same as for sulphate, and leads to synthesis of selenocysteine and selenomethionine; selenocysteine is then incorporated into proteins (Lauchli, 1993). Hence, the transfer of sulphate/selenate transporter genes from a Se accumulator like *Astragalus bisulcatus* may be useful for phytoremediation of high-Se areas (Goodson *et al.*, 2003). However, this strategy may not assist with Se uptake on soils with low Se availability, where most Se is present as selenite, selenide and elemental Se forms (Cary and Allaway, 1969). Selenite absorbed by the roots undergoes a series of reduction reactions, including conversion to selenide, and finally a reaction with O-acetylserine to form selenocysteine (Tsang and Schiff, 1978). Because of shared transporters, sulphate in growth media inhibits uptake of selenate (Ferrari and Renosto, 1972), and sulphite may inhibit uptake of selenite, but further studies are required to confirm this.

Iodine species of lower oxidative state and molecular weight (iodide, -1 and 116, respectively) are absorbed more readily than the heavier, higher valency forms (iodate, +5 and 214, respectively) (Umaly and Poel, 1971). I is transported mostly in the xylem, hence little is re-translocated from the leaves into the grain, where most is stored in the bran layers and lost during milling or polishing (Muramatsu *et al.*, 1989). To date, little has been reported on the physiology of I in the plant system and further studies are needed, especially on the forms in which I is transported and stored.

17.9 CONCLUSIONS

17.9.1 Fe & Zn

There is substantial evidence for genotypic variation to justify breeding efforts towards developing high grain-Fe and -Zn varieties. However, our knowledge of genetics, physiological mechanisms responsible for high grain Fe and Zn trait and G×E interactions is very limited, and now it is time to focus on these areas. One important point we should mention is that these proposed studies should be supplemented by bio-availability studies in animals and humans. There have been some concerns with respect to poor bio-availability of these nutrients due to naturally occurring high phytate concentrations in the grains. However, a study in rats reported that bio-availability of Fe and Zn remained constant in low- and high-density genotypes of cereals and beans (Welch *et al.,* 2000). So it is a reasonable argument that there will be an increase in absorption of these nutrients as their concentrations increase in the grain, despite their low bio-availability compared to animal food sources, as observed in the Philippine rice study (Haas *et al.*, 2005), though in that study, the varieties differed simultaneously in Fe and Zn, giving rise to potential interactions in the gut. We believe that breeding for these traits is a worthwhile approach given the impact the small increment in absorption of these nutrients will have on the lives of billions of people who are reliant on staple food crops such as rice and wheat for their dietary requirements of Fe and Zn. Finally and importantly, if

breeding for high grain-Fe and -Zn traits is to be successful and the varieties adopted by farmers, the high grain-Fe and -Zn traits must be linked to high yield. This has been achieved in rice (Gregorio, 2002), and results from wheat trials are also encouraging (R.M. Trethowan, pers. comm.).

17.9.2 Vitamin A

Despite extensive selection against pigmentation in several staple foods, genetic variation for carotenoid concentration can still be revealed by screening varieties available from germplasm banks, as illustrated in this chapter. Even within those staple crops that have substantial concentrations of carotenoids, the value of screening ancient varieties for sources of higher accumulation is obvious. The time consuming and expensive nature of carotenoid analysis still remains a significant restriction, though recent progress in the development of fast screening methods for wheat will expedite mass screenings for this staple crop. Although much work has been done in elucidating sources of increased carotenoid concentrations in staple foods there is still much to do before we obtain concentrations that can alleviate vitamin A deficiencies. In addition, it is not merely enough to develop lines with high carotenoid concentrations; they must be adapted to local conditions, and also be culturally acceptable and the carotenoids bio-available.

17.9.3 Se & I

The limited investigations carried out to date suggest that rice may be the most promising of the major cereals for breeding to improve grain Se and I density, although further screening of all the major cereals may reveal more germplasm that can enhance these traits. For rice, in particular, further pot trials and field trials conducted at sites with different soil types and including a wide range of germplasm grown together are needed to confirm whether sufficient genetic variability exists to enable selection for uptake and grain loading efficiency of Se and I. Previous studies suggest that Se and I delivered through bio-fortified cereals are highly bio-available (Jiang, Cao and Jiang, 1997; Lyons *et al*., 2003).

ACKNOWLEDGEMENT

Figures 17.1 and 17.2 were reprinted with permission from the UNU Food and Nutrition Bulletin (2000) 21: 404–409.

REFERENCES

AACC [American Association of Cereal Chemists]. 1983. Approved methods of the AACC. Method 8-01, 8th edition.

Abdel-Aal, E.-S.M., Hucl, P. & Sosulski, F.W. 1995. Compositional and nutritional characteristics of spring einkorn and spelt wheats. *Cereal Chemistry*, 72: 621–624.

Abdel-Aal, E.-S.M., Young, J.C., Wood, P.J., Rabalski, I., Hucl, P., Falk, D. & Fregeau-Reid, J. 2002. Einkorn: a potential candidate for developing high lutein wheat. *Cereal Chemistry*, 79(3): 455–457.

Alvarez, J.B., Martin, L.M. & Martin, A. 1998. Chromosomal localisation of genes for carotenoid pigments using addition lines of *Hordeum chilense* in wheat. *Plant Breeding*, 117: 287–289.

Alvarez, J.B., Martin, L.M. & Martin, A. 1999. Genetic variation for carotenoid pigment content in the amphiploid *Hordeum chilense* × *Triticum turgidum* conv. *durum*. *Plant Breeding*, 118: 187–189.

Alvarez, J.B., Urbano, J.M. & Martin, L.M. 1994. Effect on flour quality from inclusion of the *Hordeum chilense* genome into the genetic background of wheat. *Cereal Chemistry*, 71: 517–519.

Arthur, J.R. 1999. Functional indicators of iodine and selenium status. *Proceedings of the Nutrition Society,* 58: 507–512.

Baker, R.J. & Kosmolak, F.G. 1977. Effects of genotype-environment interactions. *Canadian Journal of Plant Science*, 68: 405–410.

Berrow, M.L. & Ure, A.M. 1989. Geological materials and soils. *In* M. Ihnat, ed. *Occurrence and distribution of selenium*, pp. 226–228. Boca Raton, USA, CRC Press.

Bhaskaram, P. 2002. Micronutrient malnutrition, infection, and immunity: an overview. *Nutrition Reviews,* 60: S40–S45.

Blessin, C.W., Dimler, C.O., Grogan, C.O. & Campbell, C.W. 1963. Carotenoids of corn and sorghum. *Cereal Chemistry*, 40: 436–441.

Borghi, B., Castagna, R., Corbellini, M., Heun, M. & Salamini, F. 1996. Breadmaking quality of Einkorn wheat (*Triticum monococcum* subsp. *monococcum*). *Cereal Chemistry*, 73(2): 208.

Boyd, L.O., De Boland, A.R. & Koirtyohann, S.R. 1972. Distribution of phytate and nutritionally important elements among the morphological components of cereal grains. *Journal of Agricultural and Food Chemistry,* 20: 718–721.

Brunsen, A.M. & Quackenbush, F.W. 1962. Breeding corn with high provitamin A in the grain. *Crop Science*, 2: 344–347.

Broadley, M.R., White, P.J., Bryson, R.J., Meacham, M.C., Bowen, H.C., Johnson, S.E., Hawkesford, M.J., McGrath, S.P., Zhao, F-J., Breward, N., Harriman, M. & Tucker, M. 2006. Biofortification of UK food crops with selenium. *Proceedings of the Nutrition Society,* 65: 169-181.

Buckner, B., Miguel, P. S., Janik-Buckner, D., & Bennetzen, J. L. 1996. The *y1* gene of maize codes for phytoene synthesis. *Genetics*, 143: 479–488.

Burk, R.F. & Solomons, N.W. 1985. Trace elements and vitamins and bioavailability as related to wheat and wheat foods. *American Journal of Clinical Nutrition,* 41: 1091–1102.

Cakmak, I., Yilmaz, A., Kalayci, M., Ekiz, H., Erenoglu, B. & Braun, H. 1996. Zn deficiency as a critical problem in wheat production in Central Anatolia. *Plant and Soil,* 180: 165–172.

Cakmak, I., Ozkan, H., Braun H.J., Welch R.M. & Romheld V. 2000. Zinc and iron concentrations in seeds of wild, primitive and modern wheats. *Food and Nutrition Bulletin*, 21(4): 401–403.

Cakmak, I., Torun, A., Millet, E., Feldman, M., Fahima, T., Korol, A., Nevo, E., Braun, H.J. & Ozkan, H. 2004. *Triticum dicoccoides*: An important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. *Soil Science and Plant Nutrition,* 50: 1047–1054.

Cary, E.E. & Allaway, W.H. 1969. The stability of different forms of selenium applied to low-selenium soils. *Soil Science Society of America Proceedings*, 33: 571.

Chavez, A.L., Bedoya, J. M., Sanchez, T., Iglesias, C., Ceballos, H. & Roca, W. 2000. Iron, carotene, and ascorbic acid in cassava roots and leaves. *Food and Nutrition Bulletin*, 21(4): 410–413.

Choi, E.Y., Graham, R. & Stangoulis, J. 2007. Semi-quantitative analysis for selecting Fe- and Zn-dense genotypes of staple food crops. *Journal of Food Composition and Analysis*, 20: 496–505.

Combs, G.F. 2001. Selenium in global food systems. *British Journal of Nutrition*, 85: 517–547.

D'Egidio, M.G., Nardi, S. & Vallega, V. 1993. Grain, flour and dough characteristics of selected stains of diploid wheat. *Cereal Chemistry*, 70(3): 298–303.

Eiseman, R.L., Cooper, M. & Woodruff, D.R. 1990. Beyond the analytical methodology: better interpretation and exploitation of genotype by environment interaction in breeding. *In* M.S. Kang, ed. *Genotype by environment interaction and plant breeding*, pp. 108–117. Baton Rouge, USA, Louisiana State University Agriculture Center.

Egesel, C.O. 1997. Inheritance and genetic variability in carotenoids, tocols and related traits in maize. MSc thesis, University of Illinois, Urbana-Champaign, USA.

Egesel, C.O. 2001. Genetic variation among maize genotypes for carotenoids and tocopherol compounds. PhD thesis, University of Illinois, Urbana-Champaign, USA.

Egesel, C.O., Wong, J.C., Lambert, R.J. & Rocheford, T.R. 2003. Gene dosage effects on carotenoid concentration in maize grain. *Maydica*, 48(3): 183–190.

Failla, M.L. 2003. Trace elements and host defense: recent advances and continuing challenges. *Journal of Nutrition,* 133: 1448S–1451S.

Farnham, M.W., Hale, A.J., Grusak, M.A. & Finley, J.W. 2007. Genotypic and environmental effects on selenium concentration of broccoli heads grown without supplemental selenium fertilizer. *Plant Breeding,* 126(2): 195–200.

Feil, B., Moser, S.B., Jampatong, S. & Stamp, P. 2005. Mineral composition of the grains of tropical maize varieties as affected by pre-anthesis drought and rate of nitrogen fertilisation. *Crop Science,* 45: 516–523.

Ferrari, G. & Renosto, F. 1972. Regulation of sulfate uptake by excised barley roots in the presence of selenate. *Plant Physiology,* 49: 114–116.

Garnett, T.P. & Graham, R.D. 2005. Distribution and remobilisation of iron and copper in wheat. *Annals of Botany,* 95: 817–826.

Genc, Y., Verbyla, A.P., Torun, A.A., Cakmak, I., Willsmore, K., Wallwork, H. & McDonald, G.K. 2008. Quantitative trait loci analysis of zinc efficiency and grain zinc concentration in wheat using whole genome average interval mapping. *Plant and Soil*, 314: 49-66.

Goodson, C.C., Parker, D.R., Amrhein, C. & Zhang Y. 2003. Soil selenium uptake and root system development in plant taxa differing in Se-accumulating capability. *New Phytologist,* 159(2): 391–401.

Graham, R.D. 1984. Breeding for nutritional characteristics in cereals. *Advances in Plant Nutrition,* 1: 57–102.

Graham, R.D. & Rosser, J.M. 2000. Carotenoids in staple foods. *Food and Nutrition Bulletin*, 21(4): 404–409.

Graham, R.D., Senadhira, D. & Ortiz-Monasterio, I.A. 1997. A strategy for breeding staple-food crops with high micronutrient density. *Soil Science and Plant Nutrition,* 43: 1153–1157.

Graham, R.D., Welch, R.M. & Bouis, H.E. 2001. Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods: principles, perspectives and knowledge gaps. *Advances in Agronomy,* 70: 77–142.

Graham, R.D., Senadhira, D., Beebe, S., Iglesias, C. & Ortiz-Monasterio, I. 1999. Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Research,* 60: 57–80.

Gregorio, G.B. 2002. Progress in breeding for trace minerals in staple crops. *In* H.E. Bouis, ed. *Plant breeding: a new tool for fighting micronutrient malnutrition*, pp. 500S–502S. Orlando, USA.

Gregorio, G.B., Senadhira, D., Htut, H. & Graham, R.D. 2000. Breeding for trace mineral density in rice. *Food and Nutrition Bulletin,* 21: 382–386.

Grela, E.R. 1996. Nutrient composition and content of antinutritional factors in spelt

(*T. spelta* L) cultivars. *Journal of the Science of Food and Agriculture,* 71: 399–404.

Grogan, C.O., Blessin, C.W., Dimler, R.J. & Campbell, C.M. 1963. Parental influence on xanthophyll and carotenes in corn. *Crop Science*, 3: 213–214.

Grusak, M.A. 1994. Iron transport to developing ovules of *Pisum sativum* L. I. Seed import characteristics and phloem iron-loading capacity of source regions. *Plant Physiology,* 104: 649–655.

Grusak, M.A., Pearson, J.N. & Marentes, E. 1999. The physiology of micronutrient homeostasis in field crops. *Field Crops Research,* 60: 41–56.

Haas, J.D., Beard, J.L., Murray-Kolb, L.E., del Mundo, A., Felix, A. & Gregorio, G.B. 2005. Iron-bio-fortified rice improves the iron stores of nonanemic Filipino women. *Journal of Nutrition,* 135: 2823–2830.

HarvestPlus. Breeding Crops for Better Nutrition 2007. URL: www.harvestplus.org

Hentschel, V., Kranl, K., Hollmann, J., Lindhauer, M.G., Bohm, V. & Bitsch, R. 2002. Spectrophotometric determination of yellow pigment content and evaluation of carotenoids by high-performance liquid chromatography in durum wheat grain. *Journal of Agricultural and Food Chemistry*, 50: 6663–6668.

Herren, T. & Feller, U. 1997. Influence of increased zinc levels on phloem transport in wheat shoots. *Journal of Plant Physiology,* 150: 228–231.

Hershey, C.H. & Ocampo, C. 1989. Description of new genetic markers in cassava. *Cassava Newsletter*, 13: 1–5.

Hetzel, B.S. & Pandav, C.S. 1996. *SOS for a billion. The conquest of Iodine Deficiency Disorders*. Delhi, Oxford University Press.

Hieftje, G.M. & Vickers, G.H. 1989. Developments in plasma source mass spectrometry. *Analytica Chimica Acta,* 216: 1–24.

Hill, J., Becker, H.C. & Tigerstedt, P.M.A. 1998. *Quantitative and ecological aspects of plant breeding*. London, Chapman and Hall. 275 p.

Humphries, J.M. & Khachik, F. 2003. Distribution of lutein, zeaxanthin, and related geometrical isomers in fruit, vegetables, wheat, and pasta products. *Journal of Agricultural and Food Chemistry*, 15(5): 1322-1327.

Humphries, J.M., Mares, D.J. & Graham, R.D. 2004. Application of reflectance colour measurement to the estimation of carotene and lutein content in wheat and triticale. *Journal of Cereal Science*, 40(2): 151–159.

Iglesias, C., Mayer, J., Chavez, L.S. & Calle, F. 1997. Genetic potential and stability of carotene content in cassava roots. *Euphytica*, 94: 367–373.

Inoue, H., Higuchi, K., Takahashi, M., Nakanishi, H., Mori, S. & Nishizawa, N.K. 2003. Three rice nicotianamine synthase genes, OsNAS1, OsNAS2 and OsNAS3, are expressed in cells involved in long-distance transport of iron and differentially regulated by iron. *Plant Journal,* 36: 366–381.

Jiang, X.M., Cao, X.Y. & Jiang, J.Y. 1997. Dynamics of environmental supplementation of iodine: four years' experience of iodination of irrigation water in Hotien, Xinjiang, China. *Archives of Environmental Health*, 52(6): 399–408.

Johnson, J.M. & Butler, G.W. 1957. Iodine content of pasture plants. *Physiologia Plantarum*, 10: 100–111.

Johnson, L.J. & Miller, E.S. 1938. Variation in carotenoid pigment concentration among inbred and crossbred strains of corn. *Cereal Chemistry*, 15: 345–350.

Jones, G.P. 1997. Minerals. *In* M.L. Wahlqvist (editor). *Food and nutrition*, pp. 249–254. Sydney, Australia, Allen & Unwin.

Jos, J.S., Nair, S.G., Moorthy, S.N. & Nair, R.B. 1990. Carotene enhancement in cassava. *Journal of Root Crops*, 16: 5–11.

Kang, M.S. 1990. Understanding and utilization of genotype by environment interaction in plant breeding. *In* M.S. Kang, ed. *Genotype by environment interaction and plant breeding*, pp. 52–68. Baton Rouge, USA, Louisiana State University Agriculture Center.

Koh, T.S. & Benson, T.H. 1983. Critical reappraisal of fluorimetric method for determination of selenium in biological materials. *Journal AOAC,* 66: 918–926.

Kohrle, J. 1999. The trace element selenium and the thyroid gland. *Biochimie,* 81(5): 527–533.

Koike, S., Inoue, H., Mizuno, D., Takahashi, M., Nakanishi, H., Mori, S. & Nishizawa, N.K. 2004. OsYSL2 is a rice metal-nicotianamine transporter that is regulated by iron and expressed in the phloem. *Plant Journal,* 39: 415–424.

Lacroix, L.J. & Lier, J.B. 1975. Carotenoids in durum wheat: developmental patterns during two growing seasons. *Canadian Journal of Plant Science*, 55: 679–684.

Lauchli, A. 1993. Selenium in plants: uptake, functions, and environmental toxicity. *Botanica Acta,* 106: 455–468.

Lepage, M. & Sims, R.P.A. 1968. Carotenoids of wheat flour: their identification and composition. *Cereal Chemistry*, 45: 600–611.

Lonergan, P.F. 2001. Genetic characterization of QTL mapping of zinc nutrition in barley (*Hordeum vulgare* L.). PhD thesis, University of Adelaide, Australia.

Longnecker, N.E. & Robson, A.D. 1993. Distribution and transport of zinc in plants. *In*R.D. Graham, R.J. Hannam & N.C. Uren, eds. *Zinc in soils and plants*, pp. 113–124. Dordrecht, Netherlands: Kluwer Academic Publishers.

Lyons, G.H., Stangoulis, J.C.R. & Graham, R.D. 2003. High-selenium wheat: biofortification for better health. *Nutrition Research Reviews,* 16: 45–60.

Lyons, G.H., Lewis, J., Lorimer, M.F., Holloway, R.E., Brace, D.M., Stangoulis, J.C.R. & Graham, R.D. 2004. High-selenium wheat: agronomic bio-fortification strategies to improve human nutrition. *Food, Agriculture and Environment*, 2(1): 171–178.

Lyons, G.H., Ortiz-Monasterio, I., Stangoulis, J. & Graham, R. 2005a. Selenium concentration in wheat grain: Is there sufficient genotypic variation to use in breeding? *Plant and Soil,* 269: 369–380.

Lyons, G.H., Ortiz-Monasterio, I., Genc, Y., Stangoulis, J.C.R. & Graham, R.D. 2005b. Can cereals be bred for increased selenium and iodine concentration in the grain? *In* C.J. Li *et al*., eds. *Plant nutrition for food security, human health and environmental protection*, pp. 374–375. Beijing, Tsinghua University Press.

Lyons, G.H., Genc, Y., Stangoulis, J.C.R., Palmer, L.T. & Graham, R.D. 2005c. Distribution of selenium and other nutrients in wheat grain, and the effect of processing on wheat selenium content. *Biological Trace Element Research,* 103(2): 155–168.

Mangelsdorf, P.C. & Fraps, G.S. 1931. A direct quantitative relationship between vitamin A in corn and the number of genes for yellow pigmentation. *Science*, 73: 241–242.

Marr, K.M., Batten, G.D. & Blakeney, A.B. 1995. Relationships between minerals in Australian brown rice. *Journal of the Science of Food and Agriculture,* 68: 285–291.

Marentes, E. & Grusak, M.A. 1998. Iron transport and storage within the seed coat and embryo of developing seeds of pea (*Pisum sativum* L.). *Seed Science Research,* 8: 367–375.

Mares, D.J. & Campbell, A.W. 2001. Mapping components of flour and noodle colour

in Australian wheat. *Australian Journal of Agricultural Research*, 52: 1297–1309.

Markley, M.C. & Bailey, C.H. 1935a. Determination of the carotenoid pigment concentration of small samples of whole wheat. *Cereal Chemistry*, 12: 49–53.

Markley, M.C. & Bailey, C.H. 1935b. The nature of the pigments of the gasoline extract of wheat. *Cereal Chemistry*, 12: 33–39.

Matus-Cadiz, M.A., Hucl, P., Perron, C.E. & Tyler, R.T. 2003. Genotype × environment interaction for grain colour in hard white spring wheat. *Crop Science*, 43: 219–226.

Maziya-Dixon, B., Kling, J.G., Menkir, A. & Dixon, A. 2000. Genetic variation in total carotene, iron and zinc contents of maize and cassava genotypes. *Food and Nutrition Bulletin,* 21(4): 419–422.

Miller, R.O., Jacobsen, J.S. & Skogley, E.O. 1994. Aerial accumulation and partitioning of nutrients by hard red spring wheat. *Communications in Soil Science and Plant Analysis,* 25: 1891–1911.

Moorthy, S.N., Jos, J.S., Nair, R.B. & Sreekumari, M.T. 1990. Variability of beta-carotene content in cassava germplasm. *Food Chemistry*, 36: 233–236.

Munsey, V.E. 1938. Application of the neutral wedge photometer to the measurement of carotenoid pigments in flour and macaroni products. *Journal of the Association of Official Agricultural Chemists,* 21(2): 331-352.

Muramatsu, Y., Uchida, S., Sumiya, M., Ohmomo, Y. & Obata, H. 1989. Tracer experiments on transfer of radio-iodine in the soil-rice plant system. *Water Air Soil Pollution*, 45: 157–171.

Nan, Z.H. & Han, L.Z. 1993. Elite black rice Long Jin 1. *Crop Genetic Research*, 3: 41.

Noble, R.M. & Barry, G.A. 1982. Survey of selenium concentrations in wheat, sorghum and soybean grains, prepared poultry feeds and feed ingredients from Queensland. *Queensland Journal of Agriculture and Animal Science,* 39: 1–8.

Oliver, J.R., Blakeney, A.B., & Allen, H.M. 1992. Measurement of flour color in color space parameters. *Cereal Chemistry*, 69(5): 546–551.

Ortega-Florez, C.L. 1991. Carotenes with pro-vitamin A activity and levels of cyanide in different cultivars from the state of Sao Paulo. MSc thesis, Universidade de Sao Paulo, Sao Paulo, Brazil.

Ortiz-Monasterio, I. & Graham, R.D. 2000. Breeding for trace minerals in wheat. *Food and Nutrition Bulletin,* 21: 392–396.

Oury, F.X., Leenhardt, F., Remesy, C., Chanliaud, E., Duperrier, B., Balfourier, F. & Charmet, G. 2006. Genetic variability and stability of grain magnesium, zinc and iron concentrations in bread wheat. *European Journal of Agronomy*, 25: 177–185.

Pearson, J.N., Rengel, Z., Jenner, C.F. & Graham, R.D. 1995. Transport of zinc and manganese to developing wheat grains. *Physiologia Plantarum,* 95: 449–455.

Pearson, J.N., Rengel, Z., Jenner, C.F. & Graham, R.D. 1996. Manipulation of xylem transport affects Zn and Mn transport into developing wheat grains of cultures ears. *Physiologia Plantarum,* 98: 229–234.

Pearson, J.N., Rengel, Z., Jenner, C.F. & Graham, R.D. 1998. Dynamics of zinc and manganese movement in developing wheat grains. *Australian Journal of Plant Physiology,* 25: 139–144.

Pearson, J.N., Rengel, Z. & Graham, R.D. 1999 Regulation of zinc and manganese transport into developing wheat grains having different zinc and manganese concentrations. *Journal of Plant Nutrition,* 22: 1141–1152.

Pezzarossa, B., Piccotino, D., Shennan, C. & Malorgio, F. 1999. Uptake and distribution of selenium in tomato plants as affected by genotype and sulphate supply. *Journal of Plant Nutrition*, 22: 1613–1635.

Piergiovanni, A.R., Rizzi, R., Pannacciulli, E. & Della Gatta, C. 1997. Mineral composition in hulled wheat grains: a comparison between emmer (*Triticum dicoccon* Schrank) and spelt (*T. Spelta* L) accessions. *International Journal of Food Science and Nutrition*, 48(6): 381–386.

Quackenbush, F.W., Firch, J.G., Brunson, A.M. & House, L.R. 1963. Carotenoid, oil, and tocopherol content of corn inbreds. *Cereal Chemistry*, 40: 250–259.

Rengel, Z., Batten, G.D. & Crowley, D.E. 1999 Agronomic approaches for improving the micronutrient density in edible portions of field crops. *Field Crops Research,* 60: 27–40.

Randolf, L.F. & Hand, D.B. 1940. Relation between carotenoid content and number of genes per cell in diploid and tetraploid corn. *Journal of Agriculture*, 60: 51–64.

Scholz, G., Seifert, K. & Grun, M. 1987. The normalizing factor for the tomato mutant chloronerva. 26. The effect of nicotianamine on the uptake of Mn^{2+}, Zn^{2+}, Cu^{2+}, Rb^{+} and PO_4^{3-} by the tomato mutant *chloronerva*. *Biochemie und Physiologie der Pflanzen,* 182: 189–194.

Shi, R., Li, H., Tong, Y., Jing, R., Zhang, F. & Zou, C. 2007. Identification of quantitative trait locus of zinc and phosphorus density in wheat (*Triticum aestivum* L.) grain. *Plant and Soil*, 306: 95–104.

Shinonaga, T., Gerzabek, M.H., Strebl, F. & Muramatsu, Y. 2001. Transfer of iodine from soil to cereal grains in agricultural areas of Austria. *Science of the Total Environment*, 267(1-3): 33–40.

Stephenson, L.S., Latham, M.C. & Otteson, E.A. 2000. Global malnutrition. *Parasitology,* 121: S5-S22.

Symcox, K. D., Shadley, J. D. & Weber, D. F. 1987. Detection of the time of occurrence of non-disjunction induced by the r-x1 deficiency in *Zea maize* L. *Genome*, 29: 782–785.

Takagi, S. 1976. Naturally occurring iron-chelating compounds in oat- and rice-root washings. *Soil Science and Plant Nutrition,* 22: 423–433.

Takahashi, M., Nakanishi, H., Kawasaki, S., Nishizawa, N.K. & Mori, S. 2001. Enhanced tolerance of rice to low iron availability in alkaline soils using barley nicotianamine aminotransferase genes. *Nature Biotechnology,* 19: 466–469.

Tracy, M.I. & Moller, G. 1990. Continuous flow vapour generation for inductively coupled argon plasma spectrometric analysis. Part 1: Selenium. *Journal AOAC,* 73: 404–410.

Treeby, M., Marschner, H. & Romheld, V. 1989. Mobilization of iron and other micronutrient cations from a calcareous soil by plant-borne, microbial and synthetic metal chelators. *Plant and Soil,* 114: 217–226.

Tsang, M.L.S. & Schiff, J.A. 1978. Studies of sulfate utilization by algae. *Plant Science Letters,* 11: 177–183.

Tveitnes, S., Singh, B.R. & Ruud, L. 1996. Selenium concentration in spring wheat as influenced by basal application and top dressing of selenium-enriched fertilizers. *Fertilizer Research*, 45: 163–167.

Umaly, R.C. & Poel, L.W. 1971. Effects of iodine in various formulations on the growth of barley and pea plants in nutrient solution culture. *Annals of Botany,* 35: 127–131.

Vasconcelos, M., Datta, K., Oliva, N., Khalekuzzaman, M., Torrizo, L., Krishnan, S., Oliveira, M., Goto, F. & Datta, S.K. 2003. Enhanced iron and zinc accumulation in transgenic rice with the *ferritin* gene. *Plant Science,* 164: 371–378.

Wahlqvist, M.L. 1997. Vitamins and vitamin-like compounds. *In* M.L. Wahlqvist, ed. *Food and nutrition,* pp. 222–248. Sydney, Australia, Allen & Unwin.

Wang, Z.J. & Gao, Y.X. 2001. Biogeochemical cycling of selenium in Chinese environments. *Applied Geochemistry,* 16: 1345–1351.

Welch, R.M. 1995. Micronutrient nutrition of plants. *Critical Reviews in Plant Science,* 14: 49–82.

Welch, R.M., House, W.A., Beebe, S., Senadhira, D., Gregorio, G.B. & Cheng, Z. 2000. Testing iron and zinc bioavailability in genetically enriched beans (*Phaseolus vulgaris* L.) and rice (*Oryza sativa* L.) in a rat model. *Food Nutrition Bulletin,* 21: 428–433.

Welch, R.M. & Graham, R.D. 2004. Breeding for micronutrients in staple food crops from a human nutrition perspective. *Journal of Experimental Botany,* 55: 353–364.

White, P.J., Bowen, H.C., Parmaguru, P., Fritz, M., Spracklen, W.P., Spiby, R.E., Meacham, M.C., Mead, A., Harriman, M., Trueman, L.J., Smith, B.M., Thomas, B. & Broadley, M.R. 2004. Interactions between selenium and sulphur nutrition in *Arabidopsis thaliana. Journal of Experimental Botany,* 55(404): 1927–1937.

WHO [World Health Organization]. 2003. *Iodine Deficiency Disorders*. WHO, Geneva, Switzerland. Available at www.who.int/en/

Wong, J.C. 1999. Molecular marker mapping of chromosomal regions associated with levels of carotenoids and tocopherols in maize kernels. MSc thesis, University of Illinois, Urbana-Champaign, USA.

Wong, J.C., Lambert, R.J., Wurtzel, E.T. & Rocheford, T.R. 2004. QTL and candidate genes phytoene synthase and beta-carotene desaturase associated with the accumulation of carotenoids in maize. *Theoretical and Applied Genetics*, 108: 349–359.

Yamada, H., Takeda, C., Mizushima, A., Yoshino, K., & Yonebayashi, K. 2005. Effect of oxidizing power of roots on iodine uptake by rice plants. *Soil Science and Plant Nutrition,* 51(1): 141–145.

Yuita, K., Nobusawa, Y., Shibuya, M. & Aso, S. 1982. Iodine, bromine and chlorine contents in soils and plants of Japan. 1. Iodine, bromine and chlorine contents in soils and plants of the basin of the Miomote River. *Soil Science and Plant Nutrition*, 28(3): 315–336.

Zechmeister, L. & Cholnoky, L. 1940. Carotenoids of Hungarian wheat flour. *Journal of Biological Chemistry*, 135: 31–36.

Zhang, F., Romheld, V. & Marschner, H. 1991. Release of zinc-mobilizing root exudates in different plant species as affected by zinc nutritional status. *Journal of Plant Nutrition,* 14: 675–686.

Zhang, L., Wang, X., Zhou, X. & Shi, W. 2004. Genotypic difference in selenium accumulation in rice seedlings at early growth stage and the relation with selenium content in rice seeds. *In* Abstracts of the International Symposium on Trace Elements and Health, Hangzhou, China, 10–13 October 2004. pp. 28–29.

CHAPTER 18

Breeding for quantitative variables Part 5: Breeding for yield potential

José L. Araus, Gustavo A. Slafer, Matthew P. Reynolds and Conxita Royo

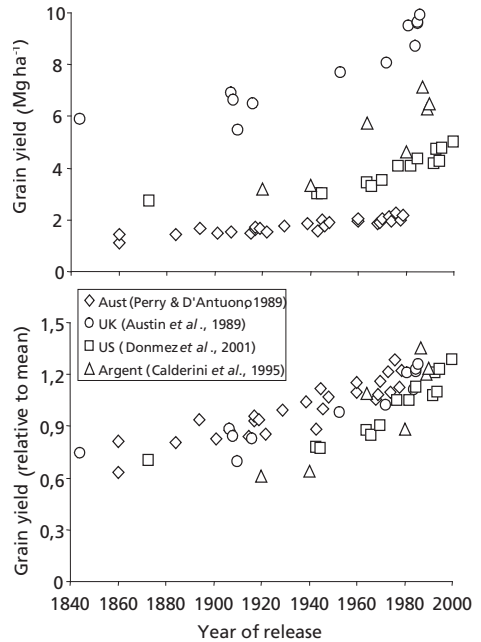
18.1 YIELD POTENTIAL

The combination of the continuing increase in the world's population and the lack of expansion, or even reduction, of arable lands (so as to maintain agricultural sustainability; Cassman *et al*., 2003), has lead to a situation where the relative importance of breeding to further raise yield potential and adaptiveness will be even greater than in the past (Slafer, Araus and Richards, 1999; Araus *et al*., 2002). This is no minor expectation: plant breeding in general, and cereal breeding in particular, have been remarkably successful during the second half of the twentieth century. They have substantially contributed to an increase in production at a rate faster than population growth, despite the Earth's population having increased faster than ever in the history of mankind (more than doubling in half a century). Despite increased demands for irrigation and chemical fertilizers, the technological progress made in cereal cultivation has actually led to a decline in the cost of cereal production per unit of output. Without this growth in productivity, people in many developing countries would have been forced to further extend cultivation onto marginal lands, thus aggravating the problem of how to sustain the natural resource base (IRRI, 1996). In the case of wheat, yield has been genetically improved virtually everywhere (Calderini, Reynolds and Slafer, 1999). The magnitude of the improvement has depended upon the environmental conditions of the region (Figure 18.1, top). However, when wheat yields are expressed as a percentage of the mean yield of the trial in which they were assessed, the data seem to converge on a single trend (Figure 18.1, bottom).

Nevertheless, the scenario is slightly more complex. Lately, yields have not been increasing at the pace seen from the 1950s to the 1990s (e.g. Calderini and Slafer, 1998; Conway and Toenniessen, 1999). As the population continues rising, and with it demand for human consumption (Reynolds, Sayre and Rajaram, 1999), food shortages will be unavoidable if food production increases do not return to previous rates, at least so as to match population growth (Khush, 1999). Therefore, increasing cereal yield is, even more so than in the past, an important challenge. If genetic yield potential is not increased in a sustainable manner, untouched natural ecosystems will go under the plough to meet greater demands, especially in the developing world.

Improvements in cereal productivity to meet the above requirement will not be easy without further technological

FIGURE 18.1
Yield of cultivars released in different years in Australia, United Kingdom, United States of America and Argentina in absolute (t/ha; top panel) and relative terms (bottom panel)



Source: Austin, Ford and Morgan, 1989

breakthroughs in shifting yield ceilings. The increase in cereal production has so far been obtained mostly from irrigated land, through the diffusion of improved varieties and agronomic practices suitable for specific ecosystems. Moreover, investment in irrigation has allowed the conversion of rainfed ecosystems when suitable. However, the rising costs of irrigation and the problems of management, cost recovery and the maintenance of existing systems constrain the further expansion of irrigation. Increasing siltation of reservoirs and canals, lowering of underground water levels, and the accumulation of salt in already irrigated soils are causing environmental concerns (IRRI, 1996). Furthermore, new growth sectors, such as industry and tourism, as well as increasing population and urbanization, are all competing for water resources. Moreover, sustainability concerns constrain the adoption of intensive agronomic practices (MacIlwain, 2004), bringing into question the desirability of the further expansion of irrigation (Araus, 2004).

Abiotic stresses frequently constrain the growth and productivity of major crop species such as cereals. They have been specifically covered in Chapter 16 and therefore we will only marginally refer to them here.

While genetic increases in yield potential are best expressed in optimal environments, they are also associated with better yields under drought (Trethowan, van Ginkel and Rajaram, 2002; Araus *et al*., 2002), nitrogen deficiency (Ortiz-Monasterio *et al*., 1997) and heat-stressed environments (Reynolds *et al*., 1998). However, this is disputed by several authors, as discussed in Chapters 3, 13, 14 and 16.

This chapter will discuss techniques for improving yield potential (and eventually adaptiveness to unfavourable environmental conditions) for small-grain cereals (such as wheat, barley and rice) and how a physiological understanding can contribute to reach such a goal. The study of crop physiology can assist cereal breeding in different ways: (i) improving the understanding of the factors that determine crop yield and adaptation through the pedagogical (for syncretic) concept of ideotype and, as a consequence, improving crop simulation models; (ii) defining particular 'secondary' traits to select for (analytical breeding) when choosing parents for crossing or screening in segregating populations; (iii) indicating the kind of genetically modified organisms (GMOs) that are worth developing and how to test them; and (iv) phenotyping associated with marker-assisted selection (MAS).

Special emphasis will be devoted to those aspects that may be particularly useful to the National Agricultural Research Systems (NARS) of developing countries, where research budgets are limited and prioritization is necessary. As such, we will focus on alternatives for evaluating secondary traits in an economical way, and prospects for the new array of molecular techniques available.

18.2 ANALYTICAL BREEDING

Genetic improvement may be achieved through selection either:

- directly, for a primary trait (such as grain yield) in a target environment (Ceccarelli and Grando, 1996). This has been referred to as empirical or pragmatic breeding; or
- indirectly, for a secondary trait, that must be putatively related to a higher yield potential or to an improved behaviour of the crop when it is grown in a stressful environment. This is known as analytical or physiological breeding.

Traditionally, breeders have achieved yield increases by intercrossing elite lines and selecting the highest- and most stable-

yielding offspring that express disease resistance and appropriate end-use quality. Thus, during the past 50 years, most of the progress in major cereals came from yield increases made possible through the gradual replacement of traditional tall cultivars by dwarf, and fertilizer-responsive varieties with superior harvest indices. These varieties were deployed as part of a package that included irrigation, fertilizers, pesticides and mechanization in a development strategy termed the Green Revolution, a term coined in March 1968 by William S. Gaud, the Director of the United States Agency for International Development (USAID).

The genetic and physiological bases governing yield are still quite poorly understood (Reynolds, Sayre and Rajaram, 1999), as yield is a quantitative trait under multigenic control, characterized by low heritability and a high genotype-by-environment (G×E) interaction (Jackson *et al.*, 1996). For these reasons, new and more strategic approaches must be explored if wheat yields are to keep pace with demand. Moreover, as empirical breeding seems to be reaching a plateau, different approaches, complementing empirical with analytical selection methodologies, may be needed to further improve grain yields. In such a context, analytical breeding, drawing on a physiological understanding of G×E interactions, may be an option. The multi-site testing of elite lines is unavoidable, and so the contribution of physiology to interpret the nature of G×E interactions, one of the critical drawbacks that the breeders have to face, may be crucial for future yield gains.

18.2.1 Identifying physiological traits

One approach to identifying potential secondary traits relies on selecting genotypes released as a result of previous breeding programmes. These genotypes are cultivated simultaneously under controlled conditions, thereby eliminating the effects on yield of varying management practices (Slafer *et al.*, 1994) and allowing the comparison of any physiological bases underlying the differences in yield capacity. Most of the traits identified in retrospective analyses have been shown to be constitutive in nature; that is, to be expressed in the absence of stress.

Retrospective studies: physiological changes associated with genetic improvement in grain yield

Understanding the contributions made in the past by successful wheat breeding may provide clues to help identify alternatives for breeders to further increase yield. Knowing the changes in physiological traits associated with genetic gains in yield potential is essential to improve the understanding of yield-limiting factors and to inform future breeding strategies. These studies may afford some clues regarding the physiological changes underlying the genetic gains in yield achieved in the past.

As well as the regular publication of this type of study, where cultivars released at different eras have been compared for yield and morpho-physiological determinants, there have also been several reviews, synthesizing the main findings from such studies (e.g. Calderini, Reynolds and Slafer, 1999). In the following section, we review the main attributes responsible for genetic gains in wheat yield in the past, including more recently published studies. The attributes are divided into four categories: time to flowering, and plant height; biomass production and partitioning; main yield components; and, lastly, cross-category interactions.

Time to flowering, and plant height

The timing of flowering is one major trait related to the adaptation of cultivars to particular growing areas, thus determining crop performance under the prevalent field conditions (e.g. Perry and D'Antuono, 1989; Passioura, 1996, 2002; Richards, 1996a; Slafer and Araus, 1998). This is why time to flowering (phenological adjustment) is one of the first attributes optimized by breeding programmes (Slafer, 2003). Consistent changes in this trait are therefore only to be expected in regions in which lengthening or shortening the growing season may have represented advantages for adaptation compared with cultivars released earlier. A scenario frequently reported in the literature, where the manipulation of time to heading may have a strong impact on adaptation, is that of regions characterized by a Mediterranean climate, i.e. a dry hot summer and humid, temperate winter (Perry and D'Antuono, 1989; Loss and Siddique, 1994; Acevedo *et al*., 1999). In these environments, the crop's vegetative and early reproductive phases occur under reasonably good water availabilities. However, as the season progresses, drought becomes more intense and frequent water stresses occur during the late reproductive phases. After anthesis, grains fill under rather severe water and heat stresses. The analysis of long-term trends in time to flowering for cultivars released in different eras reveals that, in most cases, there seemed to be little or no change in regions with climates different from the Mediterranean, while reducing time to flowering has been a successful strategy when breeding for environments characterized by terminal stresses. This is because earliness is probably the most effective solution for increasing yield where drought during grain filling is a common event (Passioura, 1996; Slafer and Whitechurch, 2001). Aside from its adaptive advantage, the reduction in crop duration has permitted an increase in cropping intensity for cereals such as rice, and also allowed land to be used for growing non-cereal crops in cereal-based farming systems (IRRI, 1996).

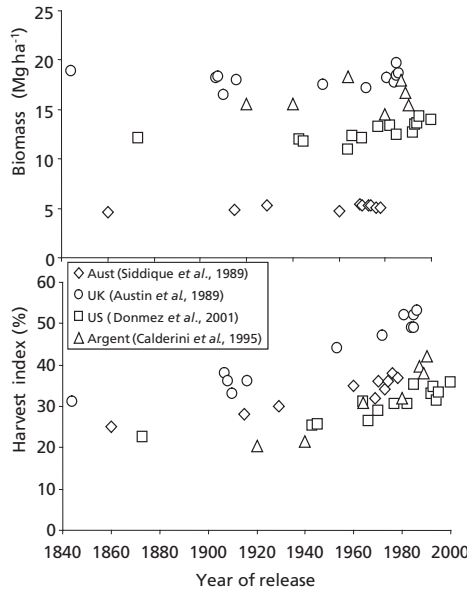
Breeders have always selected for reduced height when the initial elite material is taller than a threshold. For plant statures above this threshold (70–100 cm in wheat; Richards, 1992; Miralles and Slafer, 1995), there is no gain in biomass while there is a proportional reduction in harvest index. Below this threshold, the further gain in harvest index does not compensate for the loss in biomass, due to extremely poor radiation distribution within the canopy and consequent reductions in radiation-use efficiency (e.g. Miralles and Slafer, 1997). Thus, reducing height down to an optimum range increases yield potential (a similar biomass more efficiently partitioned to grains) and simultaneously reduces the risk of lodging causing yield penalties. Consequently, plant height has been reduced universally through wheat breeding during the 20th century (see a number of cases reviewed by Calderini, Reynolds and Slafer, 1999, and more recently reported studies: Donmez *et al.*, 2001; Brancourt-Hulmel *et al*., 2003; Jiang *et al*., 2003; Ramdani *et al*., 2003).

Reduced plant height has been one of the major causes of lack of adoption of modern varieties in a number of crops grown by farmers in developing countries as animal feed and where the way in which total biomass is portioned is irrelevant (Ceccarelli *et al*., 2000, Annicchiarico *et al.*, 2005).

Dry matter production and its partitioning

As yield has been greatly increased by genetic improvement during the past century (Figure 18.1) and grain yield is simply a

FIGURE 18.2
Biomass (t/ha; top panel) and harvest index (bottom panel) of cultivars released at different times in Argentina, Australia, United Kingdom and United States of America



Source: Austin, Ford and Morgan, 1989

fraction of the dry matter accumulated by the crop during the growing season, one might have expected a trend for increased biomass as newer higher-yielding cultivars have been released. However, in almost all cases in which the physiological bases of yield improvements have been analysed, final biomass was not associated with the year of release of the cultivars (Figure 18.2, top), with only a few exceptions to this general trend. Thus, for example, recent genetic gains in grain yield in the United Kingdom have resulted from a combination of improved growth rate in the pre-anthesis period, which has driven increases in the number of grains per unit area, and a larger source for grain filling through increases in stem soluble carbohydrate reserves (Shearman *et al.*, 2005).

While the lack of a consistent increase of biomass through breeding implies that neither leaf photosynthesis (Austin, 1989) nor radiation-use efficiency (Calderini, Dreccer and Slafer, 1997) are related to past yield increases, in a few cases positive relationships between maximum leaf photosynthesis and the year of release of cultivars have been reported (Fischer *et al.*, 1998; Jiang *et al.*, 2003). In the same way, Shearman *et al.* (2005) reported genetic increase across time for pre-anthesis radiation-use efficiency, which was correlated with wheat grain yield progress in the United Kingdom. However, except for that last study (Shearman *et al.*, 2005), these differences have not been associated with increases in biomass production. Looking in closer detail at post-anthesis, it seems, however, that modern cultivars may be characterized by higher radiation-use efficiency levels than their predecessors. This was the pattern observed in the only retrospective analysis of this characteristic that we are aware of (Calderini, Dreccer and Slafer, 1997). This increased radiation-use efficiency may just be a consequence of the increased demand of a larger sink in modern cultivars (see below) compared with their older counterparts. These kinds of results are in line with evidence from near-isogenic lines for semi-dwarfism (Miralles and Slafer, 1997) and with the fact that the number of grains growing after anthesis positively influences photosynthetic efficiency (Richards, 1996b). Also, it was recently reported that increased post-anthesis biomass may be achieved by increased sink strength through positive feedback to photosynthesis (Reynolds *et al.*, 2004). A higher stomatal conductance seems to be responsible, at least in part, for the higher photosynthetic rates (Fischer *et al.*, 1998; Araus *et al.*, 2002).

Therefore, the genetic gains in grain yield in most countries were virtually entirely due to modifications in harvest index (Figure 18.2, bottom), probably related to reductions in plant height (see above). In the few studies in which biomass partitioning was analysed before maturity in cultivars released in different eras, it seemed clear that the partitioning effects observed at maturity, as differences in harvest index, were already established by anthesis as variations in the spike-to-stem ratio (Siddique, Kirby and Perry, 1989; Slafer and Andrade 1993).

A critical issue in this respect is that harvest index in most modern cultivars seems to be close to its biological maximum (ca. 60 percent; Austin, 1980). Consequently, even though breeding in the past has been very successful in increasing grain yield through reducing height and increasing harvest index, it appears imperative to find alternatives for improving biomass—while maintaining harvest index—if further genetic gains in yield are to be expected.

Main yield components

The two main yield components are the number of grains per unit area and the averaged individual grain weight. Most experiments analysing genetic improvement effects on yield have considered these components. The vast majority of the reported studies found that, while selecting for higher-yielding cultivars, wheat breeders have consistently increased the number of grains per unit land area and produced either no trend or even a trend to slightly reduced individual grain weight (Calderini, Reynolds and Slafer, 1999; Donmez *et al*., 2001; García del Moral *et al*., 2002; Brancourt-Hulmel *et al*., 2003), though earlier studies have found a slight increase in this component during the 20th century (Cox *et al.,* 1988) or part of it (Calderini, Dreccer and Slafer, 1995).

The increased number of grains per unit and per area of the modern cultivars compared with their predecessors seems to be the consequence of a higher survival of floret primordia, as the number of potential florets per spike is apparently quite similar (e.g. Slafer and Andrade, 1993; Miralles *et al*., 2002). Thus, the higher survival of floret primordial—a process taking place during the last half of the stem elongation period (Kirby, 1988)—appears to be the most important factor leading to higher yield potential in modern cultivars (Slafer *et al*., 1994). This is in agreement with the finding that semi-dwarfing genes, which contribute significantly to increased yields in many breeding programmes throughout the world, increase the number of grains per unit land area by increasing the survival of floret primordia (Miralles *et al*., 1998). Actually, the giberellic acid-insensitive dwarfing genes *Rht-B1b* and *Rht-D1b*, the two most important commercially, have been reported to reduce plant height by around 18 percent, simultaneously exerting large pleiotropic effects improving spike fertility (Flintham *et al*., 1997).

Associations between attributes changed while selecting for higher yield

The genetic improvement of wheat yield may be understood more mechanistically by inspecting and interpreting the relationships between the main attributes of growth and partitioning of the yield components described above. Understanding the mechanistic bases by which breeding has successfully increased yield may shed light on possible future alternatives.

Genetic gains in grain yield were almost unequivocally due to gains in the number of grains per unit land area, with no gains and even slight losses in the average weight of the grains, probably as a consequence

of the lower size of the cellules induced by insensitivity to giberellic acid. This has given rise to a frequently found negative relationship between grain number per unit land area and the average weight of those grains (Slafer, Calderini and Miralles, 1996). In fact this negative relationship is also frequent when yield is increased by management practices. Although the most common interpretation in the literature has been an increased competition among grains as the number of grains per unit are is increased; we argue that the negative relationship is not competitive in nature (Slafer, 2003). This is because, even though wheat breeding has been reducing the degree of post-anthesis sink limitation to yield (e.g. Kruk, Calderini and Slafer, 1997), the photosynthetic capacity during grain filling together with the pre-anthesis assimilate reserves seem to be in excess of the demands of the growing wheat grains during post-anthesis (e.g. Richards, 1996b; Slafer, Calderini and Miralles, 1996; Borras, Slafer and Otegui, 2004; Reynolds *et al*., 2004). The main conclusion from this overall analysis is that, with only a few exceptions, wheat yield is limited by sink during grain filling and that further increases in yield depend upon increases in sink-strength after anthesis (to either further increase grain number per unit land area or to increase potential size of the individual grains).

The number of grains per unit land area is determined by various factors, including plants per unit land area, spikes per plant, spikelets per spike and grains per spikelet. These factors are sensitive to growing conditions throughout the entire period from sowing to anthesis (Slafer and Rawson, 1994). However, the number of grains per unit area seems to be far more sensitive to changes in growth partitioning over a rather short window of time immediately before anthesis (coinciding with stem elongation) than to any changes in growth occurring before this time (e.g. Kirby, 1988; Savin and Slafer, 1991; Fischer, 1993; Demotes-Mainard and Jeuffroy, 2001). In all these cases, there was a strong relationship between grain number per unit area at maturity and spike dry matter per unit area around anthesis (Slafer, 2003). Breeding seems to have increased grain number per unit area precisely through the same mechanism: crop growth has not clearly and systematically been affected by breeding for higher yield, but there has been a consistent trend to reduce plant height (a trait determined during the stem elongation phase), leading to increased partitioning towards the growing spikes resulting in an increased spike dry matter associated with more grains per unit land area (e.g. Siddique, Kirby and Perry, 1989; Slafer and Andrade, 1993). Studies demonstrating a consistent increase in yield caused by *Rht* genes in a wide variety of conditions also point to the importance of this mechanism (e.g. Brooking and Kirby, 1981; Fischer and Stockman, 1986; Miralles *et al*., 1998).

In this context, one can understand the positive relationship between the number of grains per unit area and harvest index that can be found almost without exception when comparing modern and old wheat cultivars. The number of grains is simply a reflection of changes in partitioning operating before anthesis, resulting in lower vegetative biomass.

18.2.2 Yield potential versus stress adaptation: G×E interaction

The limitations of empirical breeding are more evident when selecting for stress (e.g. drought) adaptation due to the existence of important G×E interactions, and higher within-site variability that also diminishes

heritability (h^2) (Richards, 1996a; Araus *et al*., 2002) even though this can not be generalized (Al Yassin *et al*., 2005; Comadran *et al*., 2008). Thus, although selecting for yield *per se* in the targeted environment may be sensible if the stress is uniformly severe (e.g. Ceccarelli and Grando, 1996), this is not the case in many realistic situations. For example, cultivars tested in a particular set of stressful conditions may not behave well in another set (Cooper *et al*., 1997). Moreover, a crossover effect in the yield of genotypes of high and low yield potential when regressed against the environmental index over a wide range of conditions is not often found unless the severe conditions are too extreme. This may indicate that in general, genotypes selected under high yielding environments will perform better than those with lower yield potential when grown in rather wide range of yielding environments (Calderini and Slafer, 1999; Slafer and Araus, 2007).

Constitutive whole-plant traits have a major role in affecting plant water use and plant dehydration avoidance under stress. These largely determine some of the negative relations between yield potential and the ability to sustain yield under severe water shortage (Blum, 2005). Under most dryland situations where crops depend on unpredictable seasonal rainfall, the maximization of soil moisture use is a crucial component of drought resistance (avoidance), which is generally expressed in lower water-use efficiency (WUE) (Blum, 2005) and may explain the positive correlations frequently found under Mediterranean conditions between carbon isotope discrimination ($\Delta^{13}C$, see below) and grain yield (Araus *et al.* 1998a, 2002, 2003c). However, selection for yield under drought stress resulted in a dehydration-avoidant phenotype that is rarely compatible with a high yield potential phenotype. If selection can address factors of stress adaptation in addition to yield under stress, perhaps higher yield potential and drought resistance can be recombined (Blum, 2005).

As mentioned earlier, this is one of the most controversial topics in plant breeding and the views presented in this paragraph represent one of the philosophies concerning G×E interaction. Different views and interpretations are discussed elsewhere in this volume (for example, Chapters 4, 14, 16 and 20).

Physiological avenues for increasing yield potential

As stressed above, retrospective studies with wheat indicate that improvement in yield has more often been associated with increased partitioning of biomass to the grain than it has with increased overall biomass (Austin *et al*., 1980; Waddington *et al*., 1986; Sayre, Rajaram and Fischer, 1997; Calderini, Reynolds and Slafer, 1999). Since harvest index is estimated to have an upper limit of just over 60 percent (Austin, 1980), and since this limit is already being approached (Shearman *et al*., 2005), it is becoming more important than ever to understand the physiological and genetic bases of radiation-use efficiency and biomass determination if yield is to go on increasing (Araus *et al*., 2003b).

Increases in biomass have started to be reported in spring wheat (Reynolds, Sayre and Rajaram, 1999) and winter bread wheat (Shearman *et al*., 2005). One study has revealed increases in biomass of about 10 percent in spring wheat specifically associated with the introduction of the long arm of chromosome 7D from a distant relative of wheat, *Lophopyrum elongatum,* into a number of wheat backgrounds (Reynolds *et al*., 2001). Detailed physiological investigation revealed that the basis of this increase in

biomass was associated with a small increase in assimilation rate during the spike growth stage, and a much larger increase in photosynthetic rate during grain filling, leading to an increased number of grains per spike (Reynolds, Pellegrineschi and Skovmand, 2005). Further experiments, in which grain number was increased artificially in elite lines with a brief light treatment during the rapid spike growth stage, showed that these lines posses a photosynthetic capacity in excess of that needed to fill the grains they would normally set (Reynolds, Pellegrineschi and Skovmand, 2005).

One way to exploit this excess photosynthetic capacity would be to increase grain number. CIMMYT is experimenting with a number of approaches, one of which being to exploit the large-spike trait. Large spikes themselves do not necessarily result in a higher yield should the trait not be in balance with other plant characteristics. For example, genotypes with large spikes often have small and shrivelled grain. Large-spike genotypes frequently tiller less, presumably because they carry the tiller inhibitor gene (Richards, 1988), which results in what is known as yield compensation. The challenge is therefore to bring traits together in a balanced way such that increased grain number is matched by an adequate vascular system with the ability to fill all of the additional grains, and a good tillering capacity is combined with large spikes.

Another trait being explored is the so called multi-ovary characteristic, which causes a single floret to set up to four kernels instead of just the usual one (Reynolds, Pellegrineschi and Skovmand, 2005). Currently the trait suffers from the problem of low kernel weight, but pre-breeding is underway with different spike architectures to try to better accommodate the large number of grains in terms of space and vascular connections. Traits that have shown association with improved yield in populations of random sister lines include above-ground biomass at flowering, spike mass at flowering, and duration of rapid spike growth phase (Reynolds, Pellegrineschi and Skovmand, 2005).

An alternative approach to further raise the number of grains per unit land area might be to lengthen the stem elongation phase (hypothesized by Slafer, Calderini and Miralles, 1996; Slafer *et al.,* 2001). The hypothesis would be that a longer stem elongation phase may result in more crop growth during this phase, higher spike dry matter at anthesis and subsequently more grains being filled. The hypothesis only makes for a viable solution if the length of the stem elongation phase can be manipulated and if the expectedly higher biomass accumulated during the phase is not counterbalanced by reduced partitioning to the spikes. These manipulations might involve genes responsible for sensitivity to photoperiod or for earliness *per se*, as the duration of the stem elongation phase seems to be governed by photoperiod response (e.g. Slafer and Rawson, 1997; Miralles and Richards, 2000; González, Slafer and Miralles, 2002) and to present genetic variation in its minimum duration, the intrinsic earliness (grown under long photoperiods after removing vernalization requirements; Slafer, 1996). Details of this hypothetical alternative can be found in Slafer *et al*. (2001). Briefly, there is clear genetic variation in the duration of stem elongation, even when holding constant the duration of the entire period to anthesis (Slafer and Rawson, 1994; Kernich, Halloran and Flood, 1997). At least part of this variation may be due to sensitivity to photoperiod. Such sensitivity varies between phenological phases (Slafer and Rawson,

1996; González, Slafer and Miralles, 2002). Exposing plots to photoperiod extensions during the stem elongation phase—natural day length was maintained before this phase—produced a change in the duration of the phase, associated with changes in the number of fertile florets and grains due to modifications in spike dry matter at anthesis (González, Slafer and Miralles, 2003a, b). This suggests that the mechanism by which photoperiod alters the final number of grains per unit area is the same as that determined by radiation interception during stem elongation (González, Slafer and Miralles, 2005). Therefore, isolating and subsequently manipulating (traditionally or through marker assisted selection) the genetic bases controlling sensitivity to photoperiod (as in this example, or genetic bases of differences in earliness *per se*) during stem elongation, might be an effective avenue for increasing yield.

18.3 THE PRACTICAL USE OF SECONDARY TRAITS

The putative secondary traits for a breeding programme assisted by analytical selection can be used:

- for the selection of parents to be included in the crossing block; or
- sas direct selection criteria for screening among a large number of genotypes (i.e. segregating populations) and when the amount of seed available is too small to carry out field trials with replications (i.e. the evaluation of double-haploid lines).

Whereas intensive work is continuously being carried out by physiologists to increase yield potential, few breeders routinely use the latest developed physiological criteria in their mainstream breeding programmes. One reason may be the difficulty in evaluating the response to the selection of secondary traits, this being an essential requirement for their incorporation into breeding programmes. The real value of a given trait may only be assessed by determining the genetic gain in segregating populations following selection. However, many traits are not available in well adapted genotypes, and their validation frequently requires the development of appropriate breeding material, which is costly and time consuming (Royo *et al.*, 2005). Moreover, the evaluation of some of the traits proposed by plant physiologists can be time-consuming, and sometimes even expensive, which is not practical for application to the thousands of entries that comprise the segregating generations of breeding programmes. In addition, selection in segregating populations requires screening at the plant level or between very small plots, thus hindering the use of traits that require large field plots to be assessed.

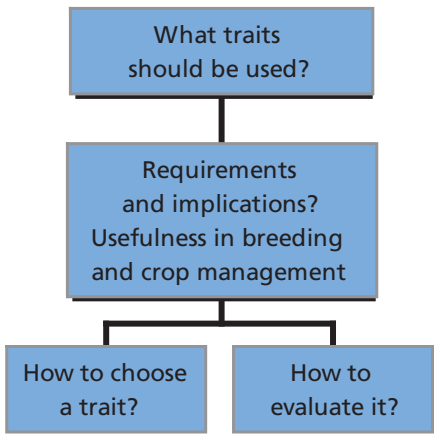
Nevertheless, some analytical or indirect selection criteria have been used for decades in breeding programmes. Plant height, days to heading or to maturity, photoperiod or vernalization responses, spike length, disease reaction, tillering capacity or grain weight are examples of traits usually evaluated in wheat and barley breeding programmes, both conventional and participatory (Ceccarelli *et al.*, 2000, 2003) so as to provide relevant information about the performance of genotypes.

Any trait to be chosen must fulfil a set of requirements related to relevance in terms of crop performance, as well as how it can be measured. These aspects are discussed below and summarized in several diagrams (Figures 18.3, 18.4 and 18.5).

18.3.1 How to choose a trait?

Two comprehensive manuals have been developed recently by CGIAR Centers for cereals such as wheat and rice. Both

FIGURE 18.3
The two main aspects when considering a candidate trait in a breeding programme



are free access. One (Reynolds, Ortiz-Monasterio and McNab, 2001) deals with how physiology may help in crop breeding, while the other (Fischer *et al.* 2003) refers to the overall context of a given breeding programme, where physiological traits could be used.

For a secondary trait to be useful in a breeding programme, it has to comply with various requirements (Araus *et al.*, 2002; Lafitte, Blum and Atlin, 2003) (Figures 18.3 and 18.4), including:

1. It should be genetically correlated with grain yield in the environmental conditions of the target environment, i.e. the relationship with yield has to be causal not casual.
2. It should be less affected by environment than is grain yield, i.e. it should have higher heritability than the yield itself, and so less G×E interaction.
3. Genetic variability for the trait must exist within the species.
4. In the case of traits addressed to breeding for stress-prone environments, the trait should not be associated with poor yields in unstressed environments. Unfortunately, this is the situation for many traits selected because they confer tolerance instead avoidance to a given stress (Araus *et al.*, 2002, 2003a).

FIGURE 18.4
Different requirements when choosing a candidate trait

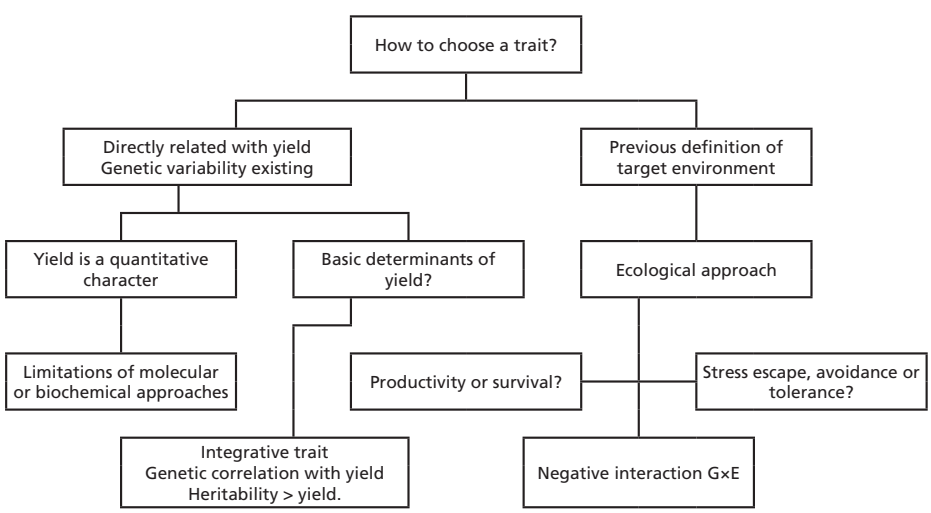
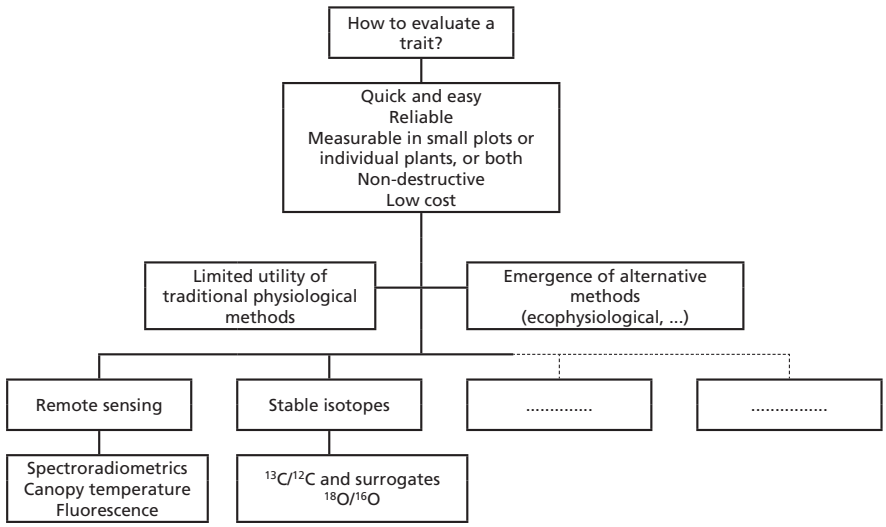


FIGURE 18.5
How a candidate trait should be measured



5. It should be possible to measure the trait rapidly and more economically than the yield itself, and in a reliable manner.
6. It can be assessed in individual plants or in very small plots.

We can predict whether the use of a secondary trait can enhance expected progress in selection by calculating its genetic correlation with yield (point 1) and heritability (point 2). Any trait that fulfils the first three requirements will provide the breeders with a useful prediction tool. While this may be enough for a breeding programme, for a direct confirmation of the value of a trait (validation), several approaches can be taken, including the development of lines expressing well the secondary trait and the assessment of their performance in the target environment, as well as the identification of the co-segregation of quantitative trait loci (QTLs) for the trait as well as for yield (Lafitte, Blum and Atlin, 2003). Once a given trait is chosen as a candidate, a practical way to incorporate it into a breeding programme is necessary (i.e. to fulfil point 5). This may be particularly pertinent when evaluating germplasm in segregating populations.

Secondary traits may be particularly suited to improving selection response for stress conditions, if they: (i) improve precision (in those case where the heritability of yield is reduced by stress); (ii) avoid any confounding effects of stress timing in yield (e.g. drought and flowering dates); (iii) focus the selection on a specific type of stress; and (iv) are cheaper, easier and faster to measure than grain yield.

Moreover, when choosing traits, it is necessary to keep in mind an eco-physiological perspective. For example, in the case of drought, most traits proposed by stress physiologists appear to be associated with stress tolerance (i.e. tolerate cell dehydration). If the target environment is under severe stress-prone conditions it may be helpful to select for this kind of trait. In most circumstances, however, the

main effect of drought is to reduce grain yield without killing the plant. In such cases, breeding for higher yield potential plus traits conferring stress avoidance (i.e. to avoid cell dehydration) may be in general a better choice (see examples in Araus *et al*., 2002, 2003a, b) but there is ample evidence that this is the exception rather than the rule (Chapter 16). At the same time, and indeed in most cases, trait evaluation should be carried out under field conditions, avoiding those experimental situations (growth chambers, greenhouses, pots) that are far from the agricultural growing environment.

18.3.2 Which traits to use in practice?

While many traits have been studied for their use in breeding for drought resistance, there is a general consensus among breeders that only a few of them can be recommended for use in practical breeding programmes at this time. For example, CIMMYT (Reynolds, Ortiz-Monasterio and McNab, 2001) and IRRI (Lafitte, Blum and Atlin, 2003) recommend the use of flowering and maturity dates, spike fertility, changes in green biomass (e.g. leaf death score) and canopy temperature. The manuals developed by both institutions include a comprehensive explanation of how to measure these traits. In practical terms, these traits seem valuable when breeding for higher yield potential and adaptation to moderate degrees of stress. Development of modern equipment and new analytical tools should facilitate future measurements of additional physiological traits in the field.

Phenology

This is the most widely used of secondary traits, due to its ease of measurement (see section 18.2.1. in this chapter). If the pattern of water deficit is predictable in a given region, selection for a flowering date that does not coincide with the period of water deficit (i.e. it exhibits an escape strategy) is a very effective way to improve drought adaptation (Araus *et al.*, 2002). The limitations to this approach are that very early varieties may suffer yield penalty in good seasons, while late-in-season freezing episodes may affect spike fertility.

Spike fertility

When stress occurs near flowering, the most sensitive growth stage, the main yield component affected is the percentage of fertile spikelets. This trait, important for any cereal, is critical for rice under water stress (Lafitte, Blum and Atlin, 2003).

Plant growth and senescence

Stress may accelerate the senescence of leaves. To measure leaf desiccation, it is possible to make a visual integration of the symptoms, translated to a ranking. Also, to check for early senescence of leaves, particularly the flag leaves, portable chlorophyll meters are extensively used. At the same time, stresses such as drought strongly affect leaf expansion (Royo *et al.,* 2004) and thus plant growth (Villegas *et al.*, 2001) and further yield (Araus *et al.,* 2002). Therefore the total green biomass evaluated at a critical plant stage (i.e. anthesis) or its change over time is a potentially powerful traits. A feasible evaluation of total biomass is only possible in practice through indirect methods, such as using spectroradiometers to measure the spectra of light reflected by the canopy (Aparicio *et al.*, 2000; Araus, Casadesús and Bort, 2001; Royo *et al.,* 2003). In many cases, however, the wide range of different spectroradiometrical indices have not fulfilled expectations when evaluating field plots for yield and their adaptation

to environmental conditions. The scarce use of spectral reflectance measurements as tools for screening in breeding programmes may be attributed to several reasons: (i) a wide range of variability for the measured trait must exist within the set of genotypes in order to be detected by the apparatus (Royo *et al*., 2003); (ii) the devices commercially available nowadays only allow measurements at canopy level, i.e. on medium to large plots, while, as noted above, selection in early segregating generations is based on individual plants or spikes cultivated in small plots; and (iii) the misleading use of spectral reflectance indices. Aside from the 'classical' vegetation indices related to green biomass (i.e. Normalized Difference Vegetation Index – NDVI; Simple Ratio Vegetation Index – SR), other indices are strongly affected by differences in green biomass (Araus, Casadesús and Bort, 2001). Therefore, the information provided by indices such as water index (WI – a measure of plant tissue content) or photochemical reflectance index (PRI – a measure of photosynthetic efficiency) is confounded by differences in biomass. Indices other than vegetation index only allow one to track physiological changes (i.e. in photosynthetic efficiency, pigment content and so forth) when differences in biomass do not exist across accessions or when used to track changes over time as a response to stress (see Tambussi *et al.,* 2002 for PRI). As an alternative to the use of indices, models constructed using the complete VIS/NIR reflectance spectra have proven to be accurate in ranking durum wheat genotypes by their grain production, although they did not provide a proper quantification of yield (Ferrio *et al*., 2005). In this regard, alternative techniques, such as the use of a conventional, affordable, digital camera (see below), may provide complementary information, such as the portion of the soil occupied by green biomass, that may help de-confound biomass from the information derived from spectral indices (Casadesús, Biel and Savé, 2005).

Canopy temperature

Because a major role of transpiration is leaf cooling, canopy temperature and its reduction relative to ambient air temperature are an indication of how much transpiration cools the leaves under a demanding environmental load. Higher transpiration represents colder leaves and higher stomatal conductance, both aspects favouring net photosynthesis and crop duration. Relatively lower canopy temperature in drought-stressed crops indicates a relatively better capacity for taking up soil moisture or for maintaining a better plant water status. Thus, higher transpiration is a positive trait when selecting for higher yield potential or better adaptation to mild to moderate stresses. The same may be inferred for higher carbon isotope discrimination when this trait is positively correlated with grain yield. Although canopy temperature may seem very easy to measure, in practice there are methodological problems, particularly in Mediterranean drought environments (Royo *et al.,* 2002; Araus *et al.,* 2002) or where there is not a homogeneous canopy. In fact, screening by canopy temperature measurements under drought stress can be done only during the vegetative growth stage, after full ground cover has been attained, before inflorescence emergence (Lafitte, Blum and Atlin, 2003), at high vapour-pressure deficits in recently irrigated crops and without the presence of wind or clouds (Royo *et al.*, 2005).

Carbon isotope discrimination

Other putative traits, while potentially useful, are less widely accepted, despite being very promising. Such is the case for discrimination against the stable isotope $\Delta^{13}C$, which is limited by the cost of its determination. CSIRO Plant Industry has released recently the two first commercial wheat varieties (cv. Drysdale in 2002, and cv. Rees in 2003) selected for high transpiration efficiency using $\Delta^{13}C$. These varieties are cultivated under rainfed conditions and rely solely upon the precipitation accumulated prior to planting. They have been selected based on their low $\Delta^{13}C$ (and thus high transpiration efficiency), fitting with what has been postulated with regards to this trait. However for Mediterranean environments, $\Delta^{13}C$ (particularly when measured in mature grains) is frequently positively correlated with grain yield (Araus *et al*., 1998a, 2003c; Villegas *et al*., 2000; Condon *et al*., 2004). One of the reasons for this positive relationship is that a genotype exhibiting higher $\Delta^{13}C$ is probably able to maintain a better water status (see Araus *et al*., 2002; Condon *et al*., 2004).

Many other traits, however, cannot yet be recommended as part of an ongoing breeding programme, particularly those that are expensive or difficult to measure. However, some can be used for the selection of parents. Also, QTLs may be mapped for traits such as root characteristics. Nevertheless, the relationship between these loci and drought resistance is not well established (Mackill, Fukai and Blum, 2003). For example, other traits, such as chlorophyll fluorescence for trait evaluation, have long been proposed (Baker and Rosenqvist, 2004), but their application in the field may be strongly affected by crop phenology (Araus *et al.,* 1998b).

18.3.3 How to measure those traits inexpensively?

Carbon isotope discrimination

Given the relatively high costs associated with carbon isotopic analysis (about € 10 per sample), several surrogate approaches have been proposed that are much cheaper, faster and easier to handle. The option most studied has been to use the mineral, or simply the total, ash content of leaves (Masle, Farquhar and Wong, 1992; Mayland *et al.*, 1993; Araus *et al.,* 1998a) or grains (Febrero *et al.*, 1994; Araus *et al.*, 1998a; Voltas *et al.*, 1998). Another promising alternative relies on the estimation of $\Delta^{13}C$ through the Near Infrared Spectroscopy (NIRS) technique (Clark *et al.,* 1995; Ferrio *et al.*, 2001), which carries with it the further advantage of being non-destructive.

Leaf colour

Leaf colour is extensively used due to the speed and ease of use of portable chlorophyll meters (such as the Minolta SPAD™), as well as for the physiological significance of the trait itself. Total chlorophyll content has been extensively used for managing nitrogen fertilization. It provides a high quality, standardized tool for nitrogen management, measured at anthesis in drought-stressed Mediterranean environments. It may be also useful for the screening of the protein content of wheat grains (Rharrabti *et al*., 2001). Moreover, since it is an indicator of early senescence, it has been reported to be positively correlated with wheat yield (Araus *et al*., 1997; Rharrabti *et al*., 2001), and indeed SPAD measurements are routinely taken in breeding programmes. However the cost of a portable chlorophyll meter (at best, at least € 1000) makes this device unaffordable for many breeding programmes in developing countries. This is why IRRI, in collaboration

with NARS, has a multi-panel leaf colour chart developed and calibrated for use with rice throughout Asia (Shukla *et al.*, 2004; IRRI, 2005).

Biomass assessment

Field spectroradiometers able to measure the spectrum of light reflected by the canopy have been expensive devices (exceeding € 12 000). However, the situation is changing. Designed initially for nitrogen management, a simple and easy-to-handle spectroradiometer, such as GreenSeeker™, has become a potentially very useful device for breeding (Figure 18.6). It gives only the basic spectroradiometric indices of green biomass, such as NDVI (as well as the inverse of the SR). Nevertheless, as noted above, these single indices are in fact the most useful for routine breeding purposes. Moreover, as the GreenSeeker includes its own radiation source, it may be used independently of the atmospheric conditions (sunny or cloudy day) and, more importantly, its cost is comparable (or even less) that of a SPAD (below € 3 000).

Conventional digital cameras are low cost devices that could be adopted for generalized use in a number of agricultural applications, including plant breeding. Through adequate processing of the information contained in digital pictures, it is possible to evaluate total green biomass much more cost effectively than with land-based portable spectroradiometers. In addition, digital pictures may also provide information that is not currently acquired through spectral reflectance measurements,

FIGURE 18.6
An example of cheap, easy-to handle spectroradiometer: the GreenSeeker. Evaluation of dry weight of pot plants

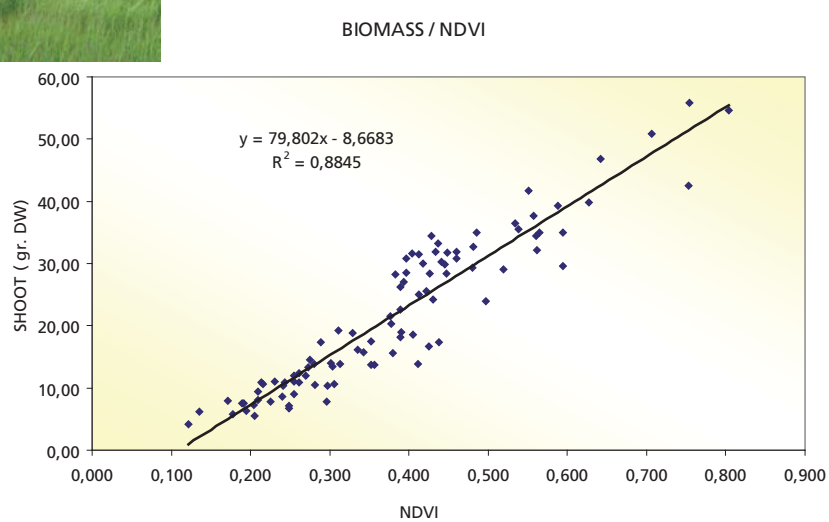
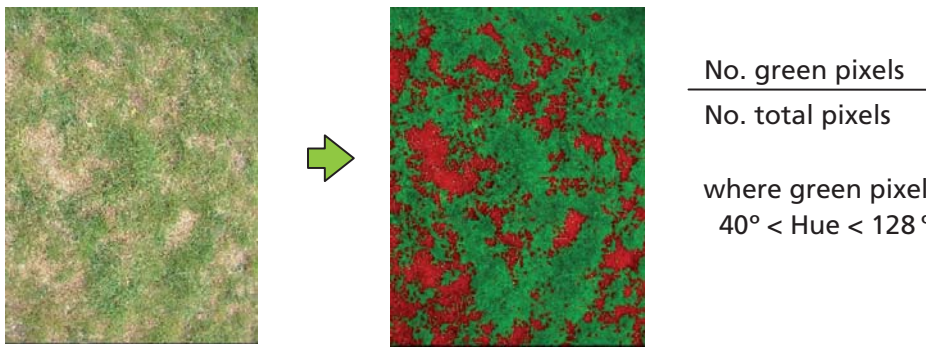


FIGURE 18.7
Some applications of digital photography: Ratio of green area to total area. Easy-to-calculate estimator of green cover



Courtesy of Dr J. Casadesús.

such as the degree of soil covered by the crop (Figure 18.7), the proportion of yellow leaves, or even yield components such as the number of spikes per unit land area (Casadesús, Biel and Savé, 2005).

18.3.4 When to use the traits?

Grain yield is the primary trait for selection in breeding programmes aimed at both increasing yield potential and adaptation to stress-prone environments. When breeding for drought adaptation, the conceptual model used considers yield under drought to be a function of: (i) yield potential; (ii) the flowering date (which indicates whether the crop will avoid drought stress); and (iii) traits that provide drought resistance.

Most breeders select strongly for traits other than yield in the early segregating generations, and do yield testing only at later stages, when a certain level of homozygosity has been achieved and sufficiently large seed quantities are available. While acknowledging the importance of secondary traits, most breeding programmes are not able to integrate them into their selection schemes for the reasons mentioned previously. The decision to advance or reject a genotype is often complex, and in practical terms breeders most often use a system of multiple cut offs. The usual approach is to carry out the selection of early generations in stress-free environments, in order to optimize the expression of desired traits in the plant and simultaneously maximize heritability and response to selection. In early generations, breeders select genotypes that presumably achieve the levels required for the primary traits evaluated in segregating populations (resistance to diseases, plant type, plant height, growth cycle, spike fertility, etc.), choosing only those with potential good score expression in the next generation, when they will be assessed again. Quality is frequently evaluated early, at the family level, in order to detect those crosses with desirable characteristics. It may also be worth evaluating in early generations for some additional physiological traits that may give an indication of yield potential or crop adaptation to abiotic stresses.

At subsequent stages, in more advanced generations, multi-site field experiments are

conducted in order to study the adaptation of lines to the target environments. The combination of yield data with data regarding secondary traits in environments ranging from well watered to high stress allows one to ascertain the adaptability of genotypes to a wide range of conditions, thus making possible more reliable decisions. At this level, selection is mostly based on the main goals of the programme, usually focusing on important commercial traits. However, data on secondary traits may be decisive at this stage in interpreting and explaining G×E interactions, mostly when the heritability of the secondary traits is higher than that of yield, and the genetic correlation of these traits with yield in the target environment is high.

Examples of how to implement indirect selection for physiological traits for drought resistance in different cereals are illustrated in Fischer *et al*. (2003)

18.4 BEYOND BREEDING

18.4.1 Social context

Hall's (2001) definition of ideotype has a wider sense than that first proposed by Donald (1968), being a plan of the phenotype of a cultivar that will perform optimally in a specific set of climatic, soil, biotic and socio-cultural conditions. Emphasis on socio-cultural aspects is now accepted as an important part of the concept of ideotype and embraces the concept of the participatory breeding approach.

Broad-based research, including research on socio-economic aspects of cereal production, is needed to characterize cereal ecosystems in terms of people and their environment; to improve understanding of farming systems, indigenous knowledge, and farmers' practices; and to refine the definitions of the kinds of technologies that should be developed. This will require close interaction with target beneficiaries (IRRI, 1996).

As Bänziger *et al* (2000) state in a comprehensive breeding manual for abiotic stress breeding published by CIMMYT, if the variety being developed for improved tolerance to any stress is unacceptable to farmers for other reasons and is not adopted, all the research work invested in that variety will be wasted. It is critically important, therefore, that farmers be involved in the selection and testing process, and that researchers pay careful attention to farmers' views on what constitutes an appropriate and attractive variety under their circumstances. In such a context, farmer participatory plant breeding represents:

- A dialogue between farmers and scientists to solve agricultural problems.
- A way to increase the impact of agricultural research by developing technologies that are more widely adopted.
- A path to more productive, stable, equitable and sustainable agricultural systems.

Abiotic stress and genetic response for adaptation to stress will depend upon the choice of target environment. To that end, farmer participatory breeding emphasizes three aspects: farmers' knowledge, farmers' ability to experiment, and farmer exchange of information and technologies. Thus, breeding programmes should include participatory on-farm trials, managed by farmers, as part of the testing of a new cultivar. This may ensure that selection has been effective, and that progress made at the station will be transferable to the farm level. Participatory trials should be run concurrently with advanced multiple-environment trials. Moreover, testing for grain quality (and other quality attributes of the crop), in consultation with farmers from the target population environments, is cheaper than replicated yield testing. Hence,

quality screening should be carried out before multiple-environment trials, so as to discard those varieties whose quality would be unacceptable to farmers (Atlin, 2003). A comprehensive methodological approach for farmer participatory breeding can be also found in Ceccarelli *et al.* (2000, 2003).

18.4.2 Crop management and sustainability of cropping systems

Breeding is just half of the equation for more productive and sustainable crops, the other half being agronomic management. The progress that has been achieved for grain yield has been the result of combining improved varieties with appropriate crop management strategies (Cooper *et al.,* 2004). Increased demand for staple crops has resulted in the intensification of agriculture all over the developing world, and one of the most serious consequences of this is soil degradation. When soil is no longer part of a natural ecosystem, if not properly managed, its physical and biological properties become degraded and productivity declines. This has been documented, for example, in the rice–wheat systems of the Punjab in South Asia (Timsina and Connor, 2001). Left unchecked, this process eventually leads to soil loss through erosion and problems of chemical imbalance, such as salinity. Water scarcity adds to the problem, and is intensified in poor soils as they lose the capacity to absorb and retain moisture. The key to this downward spiral is the loss of soil organic matter. While breeding can improve the tolerance of cultivars to salinity and reduced moisture, it is not a sustainable solution if soils continue to degrade. Since soil tillage is the principal cause behind the declining levels of organic matter, crop management practices that minimize tillage whilst keeping crop residues on the soil surface result in healthier soils. These practices are known as conservation (or zero) tillage, and are a potent tool for stabilizing and improving soils (Bradford and Peterson, 2000)

Although conservation tillage is not new (Cline and Hendershot, 2002), it is an alternative strategy that agronomists have been promoting recently in Asia and other parts of the world as a means of combating declining crop productivity (Hobbs, Giri and Grace, 1997). Conservation tillage has significant agricultural and environmental benefits (see Bradford and Peterson, 2000). Improved soil health means less erosion and higher productivity, as well as reducing the probability of encroachment of crops into natural ecosystems. Less carbon emissions are produced thanks to the reduced use of fossil fuels, there is less oxidation of soil organic matter, and less burning of crop residues. Other environmental impacts over time may include less pesticide use as a result of suppression of weeds by residues, and more integrated control of insects. Dust levels in the atmosphere are also reduced because crop residues protect the soil from wind and there is less soil disturbance during field operations. The adoption of zero-tillage on 1.3 million hectares of wheat in South Asia has been achieved through the initiatives of the Rice–Wheat Consortium for the Indo-Gangetic Plains (www.rwc.cgiar.org/rwc).

A management approach for increasing input use efficiency in irrigated environments is the cultivation of crops on raised beds, such that soil movement is reduced, resulting in greater water and nitrogen use efficiency as well reduced herbicide use (Sayre and Moreno-Ramos, 1997; Sayre, 2004). Another technology that is likely to have water-saving applications and is very compatible with raised bed cultivation is alternate-furrow irrigation.

It is well established that when plant roots detect a drying of the soil profile, they send chemical signals to their leaves resulting in a reduction in transpiration rate mediated by reduced stomatal conductance (Davies and Zhang, 1991). The result of this response is a decreased growth rate, but an increase in water-use efficiency. This innate drought response mechanism can be exploited by a modification in irrigation strategy, whereby the side from which plants, growing on top of ridges, are irrigated alternately (Davies, Wilkinson and Loveys, 2002). This was tested in furrow-irrigated maize in China and proved to be highly effective in permitting substantially reduced water application while increasing harvest index and thus not significantly reducing yield (Kang *et al.*, 2000). This simple but effective modification in irrigation strategy has clear potential benefits that can probably be adapted to many irrigated cropping systems.

Precision agriculture is also being applied to increase resource use efficiency. For example, agronomists at Oklahoma State University have adapted spectral reflectance sensor technology for use by farmers to precisely calculate the crop requirement for in-season nitrogen application. By measuring an index (NDVI) early in the crop cycle, levels of N in crop can be predicted (www.nue.okstate.edu). The value of this technology has already been demonstrated in Mexican and Ecuadorian farmers' fields in collaboration with CIMMYT scientists, enabling them to apply precise amounts of fertilizer for yield optimization while minimizing N runoff.

This Tillage Revolution, as it is sometimes referred to, is gaining momentum worldwide (MacIlwain, 2004; Nelson, Giles and Gewin, 2004). The adoption of resource-conserving crop management practices opens up new possibilities for plant breeding, specifically targeted to specific features of crop management. For example, it has been shown that tillage practices influence the composition and intensity of mycorrhizal fungi over a range of soil depths. New ideotypes may be better adapted to emergence and growth among crop residues, to soils with higher levels of organic nutrient sources, and to an increased level and diversity of soil fauna and micro-organisms. This potential synergy between breeding and innovative input-use-efficient crop management practices has been referred to as the Doubly Green Revolution (Conway, 1997).

ACKNOWLEDGMENTS

This study was supported in part by the European research project OPTIWHEAT (INCO-STRIP 015460) and by the Spanish Ministry of Science and Technology projects AGL-2006-13541-C02-1 (for J.L. Araus), AGL 2006-07814/AGR (for G.A. Slafer) and AGL-2006-09226-C02-01 and RTA 2005-00014-C04-01 (for C. Royo).

REFERENCES

Acevedo, E., Silva, P., Silva, H. & Solar, B. 1999. Wheat production in Mediterranean environments. *In* E.H. Satorre & G.A. Slafer, eds. *Wheat: Ecology and Physiology of Yield Determination,* pp. 295–331. New York, USA, Food Product Press.

Al Yassin, A., Grando, S., Kafawin, O., Tell, A. & Ceccarelli, S. 2005. Heritability estimates in contrasting environments as influenced by the adaptation level of barley germplasm. *Annals of Applied Biology,* 147: 235–244.

Annicchiarico, P., Abdellaoui, M., Kerkouli, M. & Zerargui, H. 2005. Grain yield, straw yield and economic value of tall and semi-dwarf durum wheat cultivars in Algeria. *Journal of Agricultural Science,* 143: 57–64.

Aparicio, N., Villegas, D., Casadesús, J., Araus, J.L. & Royo, C. 2000. Spectral reflectance indices for assessing durum wheat biomass, green area, and yield under Mediterranean conditions. *Agronomy Journal*, 92: 83–91.

Araus, J.L. 2004. The problem of sustainable water use in the Mediterranean and research requirements for agriculture. *Annals of Applied Biology*, 144: 229–272.

Araus, J.L., Casadesús, J. & Bort, J. 2001. Recent tools for the screening of physiological traits determining yield. *In* M.P. Reynolds, J.I. Ortiz-Monasterio & A. McNab, eds. *Application of Physiology in Wheat Breeding*, pp. 59–77. Mexico, D.F., CIMMYT.

Araus, J.L., Amaro, T., Zuhair, Y. & Nachit, M.M. 1997. Effect of leaf structure and water status on carbon isotope discrimination in field-grown durum wheat. *Plant Cell & Environment*, 20: 1484–1494.

Araus, J.L., Amaro, T., Casadesús, J., Asbati, A. & Nachit, M.M. 1998a. Relationships between ash content, carbon isotope discrimination and yield in durum wheat. *Australian Journal of Plant Physiology*, 25: 835–842.

Araus, J.L., Amaro, T., Voltas, J., Nakkoul, H. & Nachit, M.M. 1998b. Chlorophyll fluorescence as a selection criterion for grain yield in durum wheat under Mediterranean conditions. *Field Crops Research*, 55: 209–224.

Araus, J.L., Slafer, G.A., Reynolds, M.P. & Royo, C. 2002. Plant breeding and water relations in C_3 cereals: what should we breed for? *Annals of Botany*, 89: 925–940.

Araus, J.L., Bort, J., Steduto, P., Villegas, D. & Royo, C. 2003a. Breeding cereals for Mediterranean conditions: ecophysiological clues for biotechnology application. *Annals of Applied Biology*, 142: 129–141.

Araus, J.L., Slafer, G.A., Reynolds, M.P. & Royo, C. 2003b. Physiology of yield and adaptation in wheat and barley breeding. *In* H.T. Nguyen & A. Blum, eds. *Physiology and Biotechnology Integration for Plant Breeding.* Part I. Physiological basis of yield and environmental adaptation, pp. 1–49. New York, USA, Marcel Dekker.

Araus, J.L., Villegas, D., Aparicio, N., García Del Moral, L.F., El Hani, S., Rharrabti, Y., Ferrio, J.P. & Royo, C. 2003c. Environmental factors determining carbon isotope discrimination and yield in durum wheat under Mediterranean conditions. *Crop Science*, 43: 170–180.

Atlin, G. 2003. Improving drought tolerance by selecting for yield. *In* K.S. Fischer, R. Lafitte, S. Fukai, G. Atlin & B. Hardy, eds. *Breeding Rice for Drought-Prone Environments,* pp. 14–22. Los Baños, Philippines, IRRI (Available at: http://www.knowledgebank.irri.org/drought/drought.pdf; accessed 15 Jan. 2009).

Austin, R.B. 1980. Physiological limitations to cereals yields and ways of reducing them by breeding. *In* R.G. Hurd *et al.*, eds. *Opportunities for increasing crop yields*, pp. 3–19. Boston, USA, Pitman for Association of Applied Biology.

Austin, R.B. 1989. Genetic variation in photosynthesis. *Journal of Agricultural Science (Cambridge)*, 112: 287–294.

Austin, R.B., Ford, M.A. & Morgan, C.L. 1989. Genetic improvement in the yield of winter wheat: a further evaluation. *Journal of Agricultural Science (Cambridge)*, 112: 295–301.

Austin, R.B., Bingham, J., Blackwell, R.D., Evans, L.T., Ford, M.A., Morgan, C.L. & Taylor, M. 1980. Genetic improvement in winter wheat yield since 1900 and associated physiological changes. *Journal of Agricultural Science (Cambridge)*, 94: 675–689.

Baker, N.R. & Rosenqvist, E. 2004. Review article. Applications of chlorophyll fluorescence can improve crop production

strategies: an examination of future possibilities. *Journal of Experimental Botany*, 55: 1607–1621.

Bänziger, M., Edmeades, G.O., Beck, D. & Bellon, M. 2000. *Breeding for Drought and Nitrogen Stress Tolerance in Maize: From Theory to Practice.* Mexico, D.F., CIMMYT (available at: http://www.cimmyt.org/Research/Maize/map/guides_tools/BredDrougMaize/BredDroug.htm; accessed 15 Jan. 2009).

Blum, A. 2005. Drought resistance, water-use efficiency, and yield potential—are they compatible, dissonant, or mutually exclusive? *Australian Journal of Agricultural Research*, 56: 1159–1168.

Borras, L., Slafer, G.A. & Otegui, M.E. 2004. Seed dry weight response to source–sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Research*, 86: 131–146.

Bort, J., Brown, R.H. & Araus, J.L. 1996. Refixation of respiratory CO_2 in the ears of C_3 cereals. *Journal of Experimental Botany*, 47: 567–1575.

Bradford, J.M. & Peterson, G.A. 2000. Conservation tillage. *In* M.E. Sumner, ed. *Handbook of soil science*. pp. G247–G269. Boca Raton, USA, CRC Press.

Brancourt-Hulmel, M., Doussinault, G., Lecomte, C., Bérard, P., Le Buanec, B. & Trottet, M. 2003. Genetic improvement of agronomic traits of winter wheat cultivars released in France from 1946 to 1992. *Crop Science,* 43: 37–45.

Brooking, I.R. & Kirby, E.J.M. 1981. Interrelationship between stem and ear development in winter wheat: The effects of a Norin 10 dwarfing gene, *Gai/Rht2*. *Journal of Agricultural Science (Cambridge),* 97: 373–381.

Calderini, D.F., Dreccer, M.F. & Slafer, G.A. 1995. Genetic improvement in wheat yield and associated traits. A re-examination of previous results and the latest trends. *Plant Breeding*, 114: 108–112.

Calderini, D.F., Dreccer, M.F. & Slafer, G.A. 1997. Consequences of breeding on biomass, radiation interception and radiation use efficiency in wheat. *Field Crops Research*, 52: 271–281.

Calderini, D.F., Reynolds, M.P. & Slafer, G.A. 1999. Genetic gains in wheat yield and main physiological changes associated with them during the 20th century. *In* E.H. Satorre & G.A. Slafer, eds. *Wheat: Ecology and Physiology of Yield Determination,* pp. 351–377. New York, USA, Food Product Press.

Calderini, D.F. & Slafer, G.A. 1998. Changes in yield and yield stability in wheat during the 20th century. *Field Crops Research*, 57: 335–347.

Calderini, D.F. & Slafer, G.A. 1999. Has yield stability changed with genetic improvement of wheat yield? *Euphytica,* 107: 51–59.

Casadesús, J., Biel, C. & Savé, R. 2005. Turf colour measurement with conventional digital cameras. *In* J. Boaventura Cunha & R. Morais, eds. 2005 EFITA/WCCA Joint Congress on IT in Agriculture, Universidade de Trás-os-Montes e Alto Douro, Vila Real, Portugal, 25–28 July 2005, pp. 804–811.

Cassman, K.G., Dobermann, A., Walters, D.T. & Yang, H. 2003. Meeting cereal demand while protecting natural resources and improving environmental quality. *Annual Review of Environment and Resources,* 28: 315–358.

Ceccarelli, S. & Grando, S. 1996. Drought as a challenge for the plant breeder. *Plant Growth Regulation*, 20: 149–155.

Ceccarelli, S., Grando, S., Tutwiler, R., Baha, J., Martini, A.M., Salahieh, H., Goodchild, A. & Michael, M. 2000. A Methodological Study on Participatory Barley Breeding. I. Selection Phase. *Euphytica,* 111: 91–104.

Ceccarelli, S., Grando, S., Singh, M., Michael, M., Shikho, A., Al Issa, M., Al Saleh, A.,

Kaleonjy, G., Al Ghanem, S. M., Al Hasan, A.L., Dalla, H., Basha, S. & Basha T. 2003. A Methodological Study on Participatory Barley Breeding. II. Response to Selection. *Euphytica*, 133: 185–200.

Clark, D.H., Johnson, D.A., Kephart, K.D. & Jackson, N.A. 1995. Near infrared reflectance spectroscopy estimation of ^{13}C discrimination in forages. *Journal of Range Management*, 48: 132–136.

Cline, J. & Hendershot, R. 2002. Conservation tillage. In *Encyclopaedia of Soil Science*, pp. 223–226.

Comadran, J., Russell, J.R., van Eeuwijk, F., Ceccarelli, S., Grando, S., Stanca, A.M., Francia, E., Pecchioni, N., Akar, T., Al-Yassin, A., Benbelkacem, A., Choumane, W., Ouabbou, H., Bort, J., Araus, J.-L., Pswarayi, A., Romagosa, I., Hackett, C.A. & Thomas, W.T.B. 2008. Mapping adaptation of barley to droughted environments. *Euphytica*, 161: 35–45.

Condon, A.G., Richards, R.A., Rebetzke, G.J. & Farquhar, G.D. 2004. Breeding for high water-use efficiency. *Journal of Experimental Botany*, 55: 2447–2460.

Conway, G. 1997. *The doubly green revolution: food for all in the twenty-first century.* Ithaca, USA, Cornell University Press.

Conway, G. & Toenniessen, G. 1999. Feeding the world in the twenty-first century. *Nature*, 402: C55–C58.

Cooper, M., Stucker, R.E., DeLacy, I.H. & Harch, B.D. 1997. Wheat breeding nurseries, target environments, and indirect selection for grain yield. *Crop Science*, 37: 1168–1176.

Cox, T.S., Shroyer, R.J., Ben-Hui, L., Sears, R.G. & Martin, T.J. 1988. Genetic improvement in agronomic traits of hard red winter wheat cultivars from 1919 to 1987. *Crop Science*, 28: 756–760.

Davies, W.J., Wilkinson, S. & Loveys, B. 2002. Stomatal control by chemical signalling and the exploitation of this mechanism to increase water use efficiency in agriculture. *New Phytologist*, 153: 449–460.

Davies, W.J. & Zhang, J. 1991. Root signals and the regulation of growth and development of plants in drying soil. *Annual Review of Plant Physiology and Plant Molecular Biology*, 42: 55–76.

Demotes-Mainard, S. & Jeuffroy, M.H. 2001. Partitioning of dry matter and nitrogen to spike throughout the spike growth period in wheat crops subjected to nitrogen deficiency. *Field Crops Research*, 70: 153–165.

Donald, C.M. 1968. The breeding of crop ideotypes. *Euphytica,* 17: 385–403.

Donmez, E., Sears, R.G., Shroyer, J.P. & Paulsen, G.M. 2001. Genetic gain in yield attributes of winter wheat in the Great Plains. *Crop Science*, 41: 1412–1419.

Febrero, A., Bort, J., Voltas, J. & Araus, J.L. 1994. Grain yield, carbon isotope discrimination and mineral content in mature kernels of barley, under irrigated and rainfed conditions. *Agronomie*, 2: 127–132.

Ferrio, J.P., Bertran, E., Nachit, M.M., Royo, C. & Araus, J.L. 2001. Near infrared reflectance spectroscopy as a new surrogate analysis for $\Delta^{13}C$ in mature kernels of durum wheat. *Australian Journal of Agricultural Research*, 52: 809–816.

Ferrio, J.P., Villegas, D., Zarco, J., Aparicio, N., Araus, J.L. & Royo, C. 2005. Assessment of durum wheat yield using visible and near-infrared reflectance spectra of canopies. *Field Crops Research*, 94: 126–148.

Fischer, R.A. 1993. Irrigated spring wheat and timing and amount of nitrogen fertilizer. II. Physiology of grain yield response. *Field Crops Research*, 33: 57–80.

Fischer, R.A. & Stockman, Y.M. 1986. Increased kernel number in Norin 10 derived dwarf wheat: Evaluation of a cause. *Australian Journal of Plant Physiology*, 13: 767–784.

Fischer, R.A., Rees, D., Sayre, K.D., Lu, Z.-M., Condon, A.G. & Larqué-Saavedra, A. 1998. Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Science*, 38: 1467–1475.

Fischer, K.S., Lafitte, R., Fukai, S., Atlin, G. & Hardy, B., eds. 2003. *Breeding Rice for Drought-Prone Environments.* Los Baños, Philippines, IRRI. 98 p. (Available at: http://www.knowledgebank.irri.org/drought/drought.pdf; accessed 15 Jan. 2009).

Flintham, J.E., Borner, A., Worland, A.J. & Gale, M.D. 1997. Optimising wheat grain yield: effects of *Rht* (gibberellin-insensitive) dwarfing genes. *Journal of Agricultural Science (Cambridge),* 128: 11–25.

García del Moral, L.F., Roca, L.F., Ramdani, A., Rharrabti, Y., Villegas, D. & Royo, C. 2002. Genetic improvement effects on durum wheat yield and yield components in the Mediterranean region. *In* F.J. Villalobos & L. Testi, eds. *Proceedings VII Congress of the European Society for Agronomy*, pp. 91–92. Córdoba, Spain.

González, F.G., Slafer, G.A. & Miralles, D.F. 2002. Vernalization and photoperiod responses in wheat reproductive phases. *Field Crops Research,* 74: 183–195.

González, F.G., Slafer, G.A. & Miralles, D.F. 2003a. Floret development and spike growth as affected by photoperiod during stem elongation in wheat. *Field Crops Research*, 81: 29–38.

González, F.G., Slafer, G.A. & Miralles, D.F. 2003b. Grain and floret number in response to photoperiod during stem elongation in fully and slightly vernalised wheats. *Field Crops Research*, 81: 17–27.

González, F.G., Slafer, G.A. & Miralles, D.F. 2005. Photoperiod during stem elongation in wheat: is its impact on fertile floret and grain number determination similar to that of radiation? *Functional Plant Biology*, 32: 181–188.

Hall, A.E. 2001. *Crop Responses to Environment.* Boca Raton, USA, CRC Press.

Han, F., Romagosa, I., Ullrich, S.E., Jones B.L., Hayes, P.M. & Wesenberg, D.M. 1997. Molecular marker-assisted selection for malting quality traits in barley. *Molecular Breeding*, 3: 427–437.

Hobbs, P.R., Giri, G.S. & Grace, P. 1997. Reduced and zero tillage options for the establishment of wheat after rice in South Asia. CIMMYT/Rice-Wheat Consortium for the Indo-Gangetic Plains. Rice-Wheat Consortium Paper Series, No. 2. Mexico DF, CIMMYT. 24 p.

IRRI [International Rice Research Institute]. 1996. IRRI toward 2020. Los Baños, Philippines, IRRI. 43 p.

IRRI. 2005. Annual Report of the Director General, 2004-05. Vol. 15. Los Baños, Philippines, IRRI. 135 p.

Jackson, P., Robertson, M., Cooper, M. & Hammer, G. 1996. The role of physiological understanding in plant breeding; from a breeding perspective. *Field Crops Research,* 49: 11–37.

Jiang, G.M., Sun, J.Z., Liu, H.Q., Qu, C.M., Wang, K.J., Guo, R.J., Bai, K.Z., Gao, L.M. & Kuang, T.Y. 2003. Changes in the rate of photosynthesis accompanying the yield increase in wheat cultivars released in the past 50 years. *Journal Plant Research*, 116: 347–354.

Kang, S., Liang, Z., Pan, Y., Shi, P. & Zhang, J. 2000. Alternate furrow irrigation for maize production in an arid area. *Agricultural Water Management*, 45: 267–274.

Kernich, G.C., Halloran, G.M. & Flood, R.G. 1997. Variation in duration of pre-anthesis phases of development in barley (*Hordeum vulgare*). *Australian Journal of Agricultural Research* 48: 59–66.

Khush, G.S. 1999. Green revolution: preparing for the 21st century. *Genome*, 42: 646–655.

Kirby, E.J.M. 1988. Analysis of leaf, stem and ear growth in wheat from terminal spikelet stage to anthesis. *Field Crops Research*, 18: 127–140.

Kruk, B., Calderini, D.F. & Slafer, G.A. 1997. Source-sink ratios in modern and old wheat cultivars. *Journal of Agricultural Science (Cambridge),* 128: 273–281.

Lafitte, R., Blum, A. & Atlin, G. 2003. Using secondary traits to help indentify drought-tolerant genotypes. *In* K.S. Fischer, R. Lafitte, S. Fukai, G. Atlin & B. Hardy eds. *Breeding Rice for Drought-Prone Environments,* pp. 37–48. Los Baños, The Philippines, IRRI (Available at: http://www.knowledgebank.irri.org/drought/drought.pdf; accessed 15 Jan. 2009).

Loss, S.P. & Siddique, K.H.M. 1994. Morphological and physiological traits associated with wheat yield increases in Mediterranean environments. *Advances in Agronomy*, 52: 229–276.

Mackill, D.J., Fukai, S. & Blum, A. 2003. Indirect selection for physiological traits for drought tolerance. *In* K.S. Fischer, R. Lafitte, S. Fukai, G. Atlin & B. Hardy, eds. *Breeding Rice for Drought-Prone Environments,* pp. 49-54. Los Baños, The Philippines, IRRI (Available at: http://www.knowledgebank.irri.org/drought/drought.pdf; accessed 15 Jan. 2009).

MacIlwain, C. 2004. Organic: Is it the future of farming? *Nature,* 428: 792–793.

Masle, J., Farquhar, G.D. & Wong, S.C. 1992. Transpiration ratio and plant mineral content are related among genotypes of a range of species. *Australian Journal of Plant Physiology*, 19: 709–721.

Mayland, H.F., Johnson, D.A., Asay, K.H. & Read, J.J. 1993. Ash, carbon isotope discrimination and silicon as estimators of transpiration efficiency in crested wheatgrass. *Australian Journal of Plant Physiology*, 20: 361–369.

Miralles, D.J. & Richards, R.A. 2000. Response of leaf and tiller emergence and primordium initiation in wheat and barley to interchanged photoperiod. *Annals of Botany*, 85: 655–663.

Miralles, D.J. & Slafer, G.A. 1995. Yield, biomass and yield components in dwarf, semi-dwarf and tall isogenic lines of spring wheat under recommended and late sowing dates. *Plant Breeding*, 114: 392–396

Miralles, D.J. & Slafer, G.A. 1997. Radiation interception and radiation use efficiency of near isogenic wheat lines with different height. *Euphytica,* 97: 201–208.

Miralles, D.F., Katz, S.D., Colloca, A. & Slafer, G.A. 1998. Floret development in near isogenic wheat lines differing in plant height. *Field Crops Research*, 59: 21–30.

Miralles, D.J., Rharrabti, Y., Royo, C., Villegas, D. & García del Moral, L.F. 2002. Grain setting strategies of Mediterranean durum wheat cultivars released in different periods (1900–2000). AAB (Association of Applied Biologists) Conference *Genotype to phenotype: narrowing the gaps*. Royal Agricultural College, Cirencester, UK.

Nelson, L., Giles, J. & Gewin, V. 2004. Organic FAQs. *Nature,* 428: 796–798.

Ortiz-Monasterio, R.J.I., Sayre, K.D., Rajaram, S. & McMahon, M. 1997. Genetic progress in wheat yield and nitrogen use efficiency under four N rates. *Crop Science*, 37: 898–904.

Passioura, J.B. 1996. Drought and drought tolerance. *Plant Growth Regulation*, 20: 79–83.

Passioura, J.B. 2002. Environmental biology and crop improvement. *Functional Plant Biology*, 29: 537–546.

Perry, M.W. & D'Antuono, M.F. 1989. Yield improvement and associated characteristics of some Australian spring wheat culti-

vars introduced between 1860 and 1982. *Australian Journal of Agricultural Research*, 40: 457–472.

Ramdani, A., Martos, V., Villegas, D., García del Moral, L.F., Rharrabti, Y. & Royo, C. 2003. Changes on some agronomic and physiological traits of durum wheat cultivars released in Italy and Spain throughout the 20th century. Vol. 1. *In* N.E. Pogna, M. Romanò, E.A. Pogna & G. Galterio, eds. Proceedings of the 10th International Wheat Genetics Symposium, pp. 163–166. Paestum, Italy.

Reynolds, M.P., Ortiz-Monasterio, J.L. & McNab, A., eds. 2001. *Application of Physiology in Wheat Breeding.* Mexico D.F., CIMMYT. 240 p. (available at: http://www.cimmyt.org/research/wheat/map/research_results/wphysio/WPhysiology.pdf).

Reynolds, M.P., Pellegrineschi, A. & Skovmand, B. 2005. Sink-limitation to yield and biomass: a summary of some investigations in spring wheat. *Annals of Applied Biology*, 146: 39–49.

Reynolds, M.P., Sayre, K.D. & Rajaram, S. 1999. Physiological and genetic changes in irrigated wheat in the post-Green Revolution period and approaches for meeting projected global demand. *Crop Science*, 39: 1611–1621.

Reynolds, M.P., Singh, R.P., Ibrahim, A., Ageeb, O.A., Larqué-Saavedra, A. & Quick, J.S. 1998. Evaluating physiological traits to complement empirical selection for wheat in warm environments. *Euphytica*, 100: 85–94.

Reynolds, M.P., Calderini, D.F., Condon, A.G. & Rajaram, S. 2001. Physiological basis of yield gains in wheat associated with the LR19 translocation from *A. elongatum. Euphytica,* 119: 137–141.

Reynolds, M, Condon, A.G., Rebetzke, G.J. & Richards, R.A. 2004. Evidence for excess photosynthetic capacity and sink-limitation to yield and biomass in elite spring wheat. *In* T. Fischer *et al*., eds. *New directions for a diverse planet.* Proceedings of the 4th International Crop Science Congress, Brisbane, Australia.

Rharrabti, Y., Villegas, D., García del Moral, L.F., Aparicio, N., ElHani, S. & Royo, C. 2001. Environmental and genetic determination of protein content and grain yield in durum wheat under Mediterranean conditions. *Plant Breeding*, 120: 381–388.

Richards, R.A. 1988. A tiller inhibitor gene in wheat and its effect on plant growth. *Australian Journal of Agricultural Research*, 39: 749–757.

Richards, R.A. 1992. The effect of dwarfing genes in spring wheat in dry environments. I. Agronomic characteristics. *Australian Journal of Agricultural Research*, 43: 517–522.

Richards, R.A. 1996a. Defining selection criteria to improve yield under drought. *Plant Growth Regulation*, 20: 157–166.

Richards, RA. 1996b. Increasing yield potential in wheat - source and sink limitations. *In* M.P. Reynolds, S. Rajaram & A. McNab, eds. *Increasing Yield Potential in Wheat: Breaking the Barriers,* pp. 134–149. [Proceedings of a] CIMMYT International Symposium, CIANO, Ciudad, Obregon, Mexico. Mexico DF, CIMMYT.

Royo, C., Villegas, D., García del Moral, L.F., El Hani, S., Aparicio, N., Rharrabti, Y. & Araus, J.L. 2002. Comparative performance of carbon isotope discrimination and canopy temperature depression as predictors of genotype differences in durum wheat yield in Spain. *Australian Journal of Agricultural Research*, 53: 561–569.

Royo, C., Aparicio, N., Villegas, D., Casadesús, J., Monneveux, P. & Araus, J.L. 2003. Usefulness of spectral reflectance indices as durum wheat yield predictors under contrasting Mediterranean conditions.

International Journal of Remote Sensing, 24: 4403–4419.

Royo, C., Aparicio, N., Blanco, R. & Villegas, D. 2004. Leaf and green area development of durum wheat genotypes grown under Mediterranean conditions. *European Journal of Agronomy*, 20: 419–430.

Royo, C., García del Moral, L.F., Slafer, G., Nachit, M.N. & Araus, J.L. 2005. Selection tools for improving yield-associated physiological traits. *In* C. Royo, M.N. Nachit, N. Di Fonzo, J.L. Araus, W.H. Pfeiffer & G.A. Slafer, eds. *Durum Wheat Breeding: Current Approaches and Future Strategies,* pp. 563–598. New York, USA, Haworth Press.

Savin, R. & Slafer, G.A. 1991. Shading effects on the yield of an Argentinean wheat cultivar. *Journal of Agricultural Science (Cambridge)*, 116: 1–7.

Sayre, K.D. 2004. Raised-bed cultivation. *In* R. Lal, ed. *Encyclopaedia of Soil Science*. Marcel Dekker, Book Site Online publication.

Sayre, K.D. & Moreno-Ramos, O.H. 1997. Applications of raised-bed planting systems to wheat. Wheat Special Report no. 31. Mexico D.F., CIMMYT. 36 p.

Sayre, K.D., Rajaram, S. & Fischer, R.A. 1997. Yield potential progress in short bread wheats in northwest Mexico. *Crop Science*, 37: 36–42.

Shearman, V.J., Sylvester-Bradley, R., Scott, R.K. & Foulkes, M.J. 2005. Physiological processes associated with wheat yield progress in the UK. *Crop Science*, 45: 175–185.

Shukla, A.K, Ladha, J.K., Singh, V.K., Dwivedi, B.S., Balasubramanian, V., Gupta, R.K., Sharma, S.K., Singh, Y., Pathak, H., Pandey, P.S., Padre, A.T. & Yadav R.L. 2004. Calibrating the leaf color chart for nitrogen management in different genotypes of rice and wheat in a systems perspective. *Agronomy Journal*, 96: 1606–1621.

Siddique, K.H.M., Kirby, E.J. M. & Perry, M.W. 1989. Ear to stem ratio in old and modern wheats; relationship with improvement in number of grains per ear and yield. *Field Crops Research*, 21: 59–78.

Slafer, G.A. 1996. Differences in phasic development rate amongst wheat cultivars independent of responses to photoperiod and vernalization. A viewpoint of the intrinsic earliness hypothesis. *Journal of Agricultural Science (Cambridge)*, 126: 403–419.

Slafer, G.A. 2003. Genetic basis of yield as viewed from a crop physiologist's perspective. *Annals of Applied Biology*, 142: 117–128.

Slafer, G.A. & Andrade, F.H. 1993. Physiological attributes related to the generation of grain yield in bread wheat cultivars released at different eras. *Field Crops Research*, 31: 351–367.

Slafer G.A. & Araus J.L. 1998. Improving wheat responses to abiotic stresses. *In* A.E. Slinkard, ed. Proceedings of the 9th International Wheat Genetics Symposium, Vol. 1, pp. 201–213. Saskatoon, Canada, University of Saskatchewan Extension Press.

Slafer, G.A. & Araus, J.L. 2007. Physiological traits for improving wheat yield under a wide range of conditions. *In* J.H.J. Spiertz, P.C. Struik & H.H. van Laar, eds. *Scale and Complexity in Plant Systems Research: Gene-Plant-Crop Relations,* pp. 145–154. Dordrecht, Netherlands, Springer.

Slafer, G.A., Araus, J.L. & Richards, R.A. 1999. Physiological traits to increase the yield potential of wheat. *In* E.H. Satorre & G.A. Slafer, eds. *Wheat: Ecology and Physiology of Yield Determination*, pp. 379–415. New York, USA, Food Product Press.

Slafer, G.A., Calderini, D.F. & Miralles, D.J. 1996. Generation of yield components and compensation in wheat: Opportunities for further increasing yield potential. *In* M.P. Reynolds, S. Rajaram & A. McNab,

eds. *Increasing Yield Potential in Wheat: Breaking the Barriers*, pp. 101–133, [Proceedings of a] CIMMYT International Symposium, CIANO, Ciudad, Obregon, Mexico. CIMMYT, Mexico DF, Mexico.

Slafer, G.A. & Rawson, H.M. 1994. Sensitivity of wheat phasic development to major environmental factors: a re-examination of some assumptions made by physiologists and modellers. *Australian Journal of Plant Physiology*, 21: 393–426.

Slafer, G.A. & Rawson, H.M. 1996. Responses to photoperiod change with phenophase and temperature during wheat development. *Field Crops Research*, 46: 1–13.

Slafer, G.A. & Rawson, H.M. 1997. Phyllochron in wheat as affected by photoperiod under two temperature regimes. *Australian Journal of Plant Physiology*, 24: 151–158.

Slafer, G.A. & Whitechurch, E.M. 2001. Manipulating wheat development to improve adaptation and to search for alternative opportunities to increase yield potential. *In* M.P. Reynolds, J.L. Ortiz-Monasterio & A. McNab, eds. *Application of Physiology in Wheat Breeding*, pp. 160–170. Mexico D.F., CIMMYT. 240 p. (available at: http://www.cimmyt.org/research/wheat/map/research_results/wphysio/WPhysiology.pdf).

Slafer, G.A., Calderini, D.F., Miralles, D.J. & Dreccer, M.F. 1994. Pre-anthesis shading effects on the number of grains of three bread wheat cultivars of different potential number of grains. *Field Crops Research*, 36: 31–39.

Slafer, G.A., Abeledo, L.G., Miralles, D.J., Gonzalez, F.G. & Whitechurch, E.M. 2001. Photoperiod sensitivity during stem elongation as an avenue to raise potential yield in wheat. *Euphytica,* 119: 191–197.

Slafer, G.A., Araus, J.L., Royo, C. & García Del Moral L.F. 2005. Promising eco-physiological traits for genetic improvement of cereal yields in Mediterranean environments. *Annals of Applied Biology*, 146: 61–70.

Tambussi, E.A., Casadesus, J., Munné-Bosch, S. & Araus, J.L. 2002. Photoprotection in water stressed plants of durum wheat (*Triticum turgidum* L. var. *durum*): changes in chlorophyll fluorescence, spectral signature and photosynthetic pigments. *Functional Plant Biology*, 29: 35–44.

Timsina, J. & Connor, D.J. 2001. Productivity and management of rice-wheat cropping systems: Issues and challenges. *Field Crops Research*, 69: 93–132.

Trethowan, R.M., van Ginkel, M. & Rajaram, S. 2002. Progress in breeding for yield and adaptation in global drought affected environments. *Crop Science*, 42: 1441–1446.

Voltas, J., Romagosa, I., Muñoz, P. & Araus, J.L. 1998. Mineral accumulation, carbon isotope discrimination and indirect selection for grain yield in two-rowed barley grown under semiarid conditions. *European Journal of Agronomy*, 9: 147–155.

Villegas, D., Aparicio, N., Nachit, M.M., Araus, J.L. & Royo, C. 2000. Photosynthetic and developmental traits associated with genotypic differences in durum wheat yield across the Mediterranean basin. *Australian Journal of Agricultural Research*, 51: 891–901.

Villegas, D., Aparicio, N., Blanco, R. & Royo, C. 2001: Biomass accumulation and main stem elongation of durum wheat grown under Mediterranean conditions. *Annals of Botany*, 88: 617–627.

Waddington, S.R., Ranson, J.K., Osmanzai, M. & Saunders, D.A. 1986. Improvement in the yield potential of bread wheat adapted to northwest Mexico. *Crop Science*, 26: 698–703.

CHAPTER 19

Marker-assisted selection in theory and practice

Andrew R. Barr

19.1 INTRODUCTION

19.1.2 Definitions and background

Marker-assisted selection (MAS) involves the use of genetic markers to follow regions of the genome that encode specific characteristics of a plant. For example, a marker genetically linked to a disease-resistance locus can be used to predict the presence of the resistant or the susceptible allele. The reliability of the prediction will depend upon the closeness of the genetic linkage. Markers that co-segregate with the target trait are absolutely reliable and can be regarded as diagnostic.

To be effective, the markers must detect a polymorphism between the plants being analysed. The nature of the polymorphism will vary depending on the marker system being used. Initially markers based on protein differences were widely used. These were based on variation in protein size detected by size fractionation electrophoresis or related techniques, or differences in protein charge or iso-electric point. Specific proteins or enzymes could be detected by staining for total proteins or using the enzyme activity to produce or remove a coloured substrate. Iso-electric variants are referred to as isozymes and were, for many years, extremely important markers for specific chromosomes and chromosome regions. They suffered from two major weaknesses:

- They are limited in number and are often difficult to detect or assay. The total number of isozyme loci that can be assayed is generally fewer than 100 in a well characterized species.
- They detect only low levels of polymorphism. The assays are dependent on revealing a shift in protein mobility through either an altered size or altered iso-electric point, involving a change in at least one amino acid in the protein. Therefore only changes in the coding sequence of a limited number of genes can be detected and changes that lead to an amino acid substitution or deletion. In addition, many such changes are not selection neutral but may involve a change in enzyme function that can be detrimental to other characteristics of the plant.

Detection of sequence variation at the DNA level offers several important advantages over protein-based markers. There are essentially an unlimited number of such DNA markers since sequence variation, in the form of single-base changes, insertions and deletions, or large sequence differences, are abundant. Indeed, in many systems where detailed analysis has been carried out, a sequence difference occurs between two sexually compatible individuals at a frequency of greater than one in every three hundred bases. Most of these changes will be have no detectable effect on plant performance and are selection neutral. In addition, a large range of techniques are now available for detecting sequence variation. The available techniques differ in their ability to detect variation, the ease of assay and the number of loci that can be assayed simultaneously. While early work focused around the use of restriction fragment length polymorphisms (RFLPs) detected via Southern Hybridization, polymerase chain reaction (PCR)-based marker systems are now more widely used, in particular microsatellite or simple sequence repeat (SSR) and amplified fragment length polymorphism (AFLP). New techniques not requiring gel electrophoresis are gaining in importance. These include mass spectrometric and Taq-Man type assays. As sequence information becomes more widely available for major crop species, we can expect to see alternatives to gel-based

detection systems applied to identify single nucleotide polymorphisms (SNPs).

Marker development and implementation can be divided into the following steps:

- identify parents differing in the trait of interest;
- develop a population of plants segregating for the trait of interest (using doubled haploids, single-seed descent or F_2 populations);
- screen the population for the traits of interest;
- construct linkage maps of the cross, using bulked segregant analysis or genomic fingerprinting;
- identify molecular markers that co-segregate with the trait of interest;
- test the applicability and reliability of the markers in predicting the trait in related families (also referred to as marker validation);
- produce clear and simple protocols for assaying the markers;
- modify breeding strategy to optimize use of MAS relative to alternative selection techniques; and
- implement into the breeding programmes.

19.2 DEVELOPING MOLECULAR MARKERS

The technique used to identify a molecular marker linked to a trait of interest will depend upon the type of trait and the resources and marker systems available. There are several features of the trait that will be important in devising the most efficient marker development strategy. These include:

- the mode of inheritance—whether simple or polygenic;
- the heritability of the trait and the reliability of the alternative (bio-assay, biochemical analysis, etc.); and
- the ease of bio-assay (this may restrict the size of the population that can be assayed).

The process used to decide on the best approach to marker development is summarized in Figure 19.1. There are five fundamental questions to be asked.

1. Is a marker needed? Marker development is slow and expensive, so it must be clear that there will be a real benefit to the breeding programme through access to a molecular marker. This will involve comparing the cost of marker development and use with more traditional screening methods.
2. Is the trait simply inherited or is it controlled by many genes? If the inheritance is simple, then bulked segregant analysis would usually be the preferred method for marker identification. However, if inheritance is controlled by several genes or is strongly influenced by the environment (low heritability), full map construction or linkage disequilibrium mapping would be the best options.
3. Cost of map construction? The shift to automation and high-throughput marker analysis is dramatically reducing the cost and time involved in marker development. For many species, map construction will be sufficiently cheap to replace Bulked Segregant Analysis (BSA).
4. Is one parent critical or a common source of trait? The question relates to the diversity of germplasm available for screening. If there are multiple sources of the desirable trait, linkage disequilibrium mapping offers the options of localizing loci from multiple sources, while full map construction will only allow the genes from one source to be localized.

TABLE 19.1
Comparison of major marker systems

Marker system	Usual loci per assay	DNA amount	Approx. time per assay	Comments
RFLP	1	5.0 g	5 days	Co-dominant, reliable, often low-level polymorphism
SSR	1	0.2 g	5 hours	Co-dominant, reliable, large number of alleles
AFLP	50	0.2 g	1 day	Mostly dominant, reliable, low level of polymorphism but high multiplexing capacity
RAPD	10	0.2 g	5 hours	Dominant, unreliable in some situations
SNP	1	0.05 g	1 to 5 hours	Co-dominant or dominant, very rapid result depending on technology platform. High development cost
DArT	>100, limited only by chip	5 ng	1.5 to 2 days	Dominant, ideal for fingerprinting as many loci generated from single sample

5. Is there access to large-scale phenotyping systems? Linkage disequilibrium mapping is dependent upon the analysis of a large pool of germplasm and will require substantially more marker assays than full map construction. This is most effectively achieved through automation and high-throughput genotyping.

The marker systems that are currently available, and their relative advantages and disadvantages are summarized in Table 19.1.

The data in Table 19.1 represent averages for these marker systems. Although Table 19.1 would suggest that random amplified polymorphic DNA (RAPD) markers are the most efficient, the unreliability of these markers and the inability to transfer them between crosses has greatly limited their use. The full potential of newer systems, such as Diversity Arrays Technology (DarT) (Jaccoud *et al.*, 2001), is as yet untapped.

Essentially five basic systems have been used to develop markers linked to traits of interest. The relative merits of each and their applicability are outlined below.

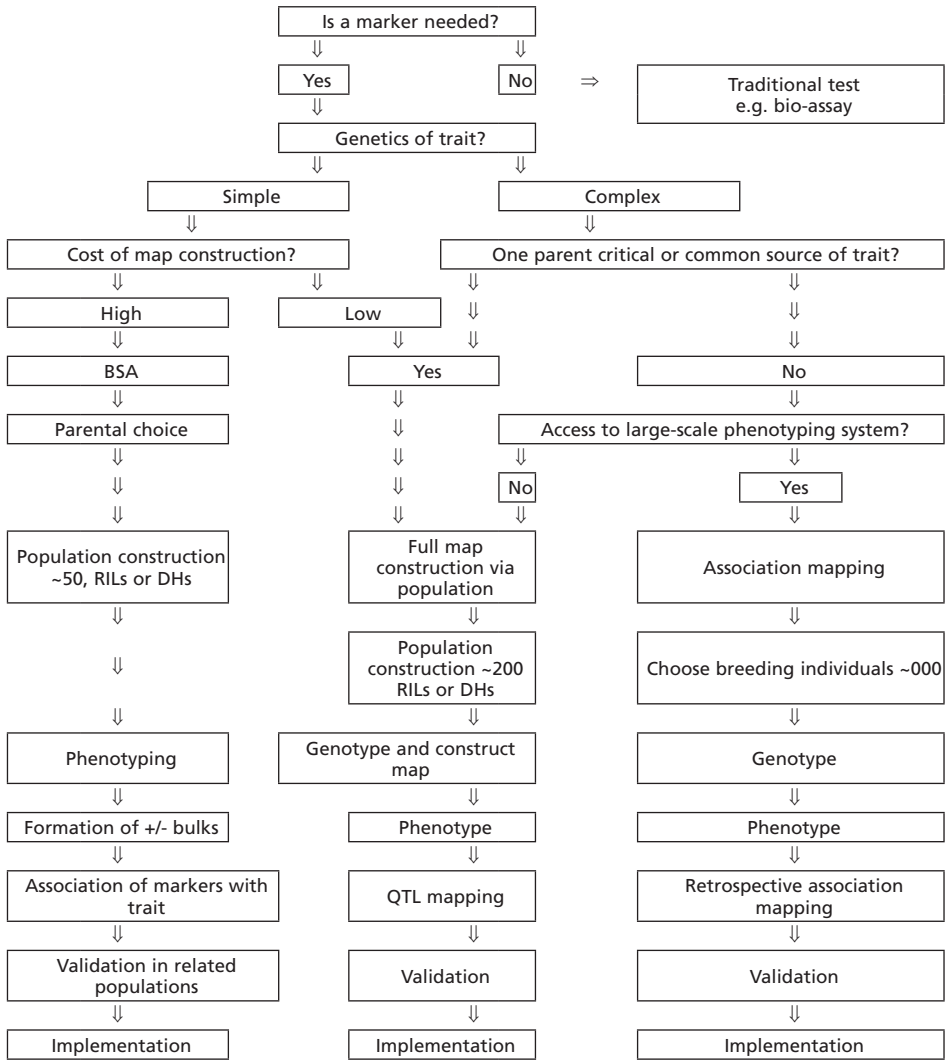
19.2.1 Fully mapped approaches

Complete linkage maps generated from screening the progeny of a cross have provided the basis for most early marker development work. In this process, a population is produced from parents differing in the trait of interest. Molecular markers, able to detect polymorphisms between the two parents, are screened against the population so that each line will have been assayed for the target trait or traits and a large number of markers. Linkage between the markers is assessed based on the segregation pattern. Several public and commercial software packages are available to assist in the map construction. In the first phase of map construction, the markers are divided into linkage groups, usually based on paired comparisons. Markers that fall into a linkage group are then ordered through two- and three-point analysis.

It is often difficult to assign marker-based linkage groups to known chromosomes. For most major crop species, good quality linkage maps have been published. Markers from these maps can be used as reference points to determine chromosome designations. For some species, such as wheat, chromosome substitution, translocation or deletion lines can be used to determine the chromosome designation for specific probes, and hence linkage groups.

With a good quality linkage map, the target trait can be mapped relative to the molecular markers by measuring

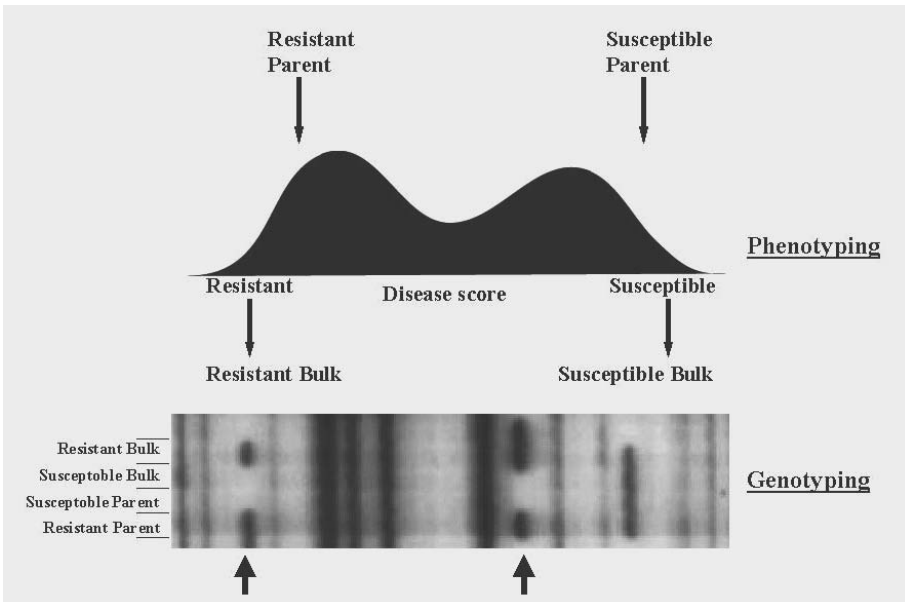
FIGURE 19.1
Marker development strategies



the frequency of trait co-segregation with markers.

Full map construction is difficult and labour intensive, particularly in species with a large number of linkage groups, where linkage maps must be constructed for each chromosome. Usually ten to twenty markers are desirable for each chromosome to give reasonable genome coverage (about 1 every 10 to 20 cM). The work involved in screening this number of markers is considerable if markers such as RFLPs are used. For some cultivated species, the germplasm base may be small and the level of polymorphism between varieties or cultivars low. In order to reveal sufficient

FIGURE 19. 2
Bulked segregant analysis in the identification of AFLP markers linked to a disease resistance locus



Parents, differing in their resistance phenotype, are used to produce a population segregating for resistance. This population is phenotyped for resistance and two bulks are generated: resistant and susceptible. The parental DNA and the DNA pooled from the bulks are analysed with a range of AFLP markers to identify bands present in the resistant parent and the resistant bulk (arrows).

polymorphisms for full genome coverage it may be necessary to screen several thousand probes. Combining RFLP with various other marker systems offers an alternative approach. AFLP markers are particularly useful due to the high multiplexing ratio. Even though the level of polymorphism for AFLP bands is usually lower than for RFLPs, the large number of bands that can be scored in a single assay makes this a valuable marker system for enhancing linkage maps.

It is also often important to consider how many traits of significance may be segregating in the population used for marker development. If several important traits are segregating, the benefits derived from constructing a full linkage map may outweigh the time and costs associated with full map construction relative to Bulked Segregant Analysis (BSA).

Linkage maps of most major crop species have now been constructed, usually based around RFLP markers. These established maps are a valuable resource and a good starting point in identifying the most appropriate markers to use to obtain good genome coverage.

19.2.2 Bulked Segregant Analysis

BSA is a technique that allows rapid, cheap development of markers for simply inherited traits (Michelmore, Paran and Kesseli, 1991). The first step involves pooling individuals from the two phenotypic extremes of a segregating F_2, doubled haploid or similar

population. DNA isolated from the two pools is then screened with DNA markers, usually RAPD or AFLP, and polymorphic bands identified. Clear polymorphisms seen between the two pools will be derived from regions of the genome that are common between the individuals that made up the pools. The remainder of the genome will be randomly contributed by the parents and should show no polymorphisms between the pools. The principle of this technique is illustrated in Figure 19. 2.

This technique has clear advantages for marker identification relative to construction of complete linkage maps. Only a few weeks are required to screen the pools with the AFLP or RAPD primers and the cost will be between 20 percent and 30 percent of that for full map construction. For traits controlled by a single gene, such as many disease resistances, only five to ten plants are required for each bulk, meaning that only small populations are required. However, there are also disadvantages to this technique.

- The method is not effective (although not impossible) for complex traits controlled by several unlinked loci. The technique can be used for traits controlled by fewer than three major loci.
- The actual location of the gene of interest is not revealed, as only linked markers are detected.
- The method can only be efficiently used with PCR-based marker systems such as SSR, RAPD and AFLP markers. For AFLP and RAPD markers, it is usually necessary to convert the marker to a sequenced tagged site (STS) marker, and this is not a simple task. This problem is largely overcome through the use of SSR markers or where very good quality maps are available of AFLP markers. The effort involved in the initial screening of the pools may be greater for SSR markers and a large number will be required, but the benefits in having markers that are immediately applicable in selection programmes may overcome this drawback. Good collections of SSRs are available for most major crop species, but are limited for minor species.

19.2.3 Partial maps

The 'partial map' strategy represents a compromise between the limited resolution of a BSA and the expensive and, in some cases unnecessarily detailed, fully mapped approach. The main advantage of the partial map approach is obviously the cost savings in comparison with full map construction.

As the name suggests, partial maps are constructed of selected regions of the genome. A partial mapping approach is normally applied when some prior knowledge of the genetic control of the trait or gene locations is available, which suggests that a full map is not required for development of the marker–trait association but where the genetic control is too complex for BSA to be successfully used.

Other scenarios where partial maps are useful are where BSA has been used to develop a marker–trait association and consensus maps used to obtain a tentative genomic location for the gene, but where:

- the trait appears closely linked to other traits of interest, and hence linkage relationships are crucial; or
- the marker density on existing maps is too poor for useful markers to be developed.

19.2.4 Pedigree-based analysis

Pedigree-based whole-genome marker development (sometimes called analysis by descent, or similar) is a technique for developing and using marker-assisted

selection in pragmatic breeding programmes by using pedigree information to identify markers linked to traits based on identity by descent in actual breeding populations. This approach therefore bridges the development, validation and implementation stages discussed earlier by making use of the pedigree, phenotypic and genotypic information collected during the day-to-day activities of a breeding programme.

Paull *et al.* (1998) used this approach to develop a marker in wheat linked to stem rust resistance, Jordan *et al.* (1999) identified a marker for sorghum midge resistance, and Eisemann *et al.* (2004), while working with wheat and barley germplasm, proposed an extension to this technique termed 'pedigree-based whole genome marker development', and were developing new software and systems to maximize the information gleaned from such analysis. 'Graphical genotypes' (Young and Tanklsey, 1989) are often integral to pedigree analysis; they are discussed in Section 19.5.2, below.

19.2.5 Linkage disequilibrium mapping

The principle techniques used to identify marker trait associations in crop species have been based around the construction of linkage maps or the use of BSA. Both techniques are based around the production, genotyping and phenotyping of special populations. The populations are generally developed from two varieties that show a major difference in the traits targeted for mapping. This leads to several problems.

- There is usually a high cost associated with phenotyping, particularly for traits requiring extensive field trials or complex analysis, such as many aspects of processing quality and yield. Consequently the number of replicates and of sites are often limited, reducing the sensitivity of some of the analyses.
- The lines (varieties) used to construct the populations are often out-of-date by the time the marker and trait information is available. Many marker development projects for annual crops are using populations that were established five or more years before the marker development work. This reduces the value of the information gathered and the scope of its implementation.
- The structure of the populations limits the types of traits that can be mapped and many of the subtleties of adaptation cannot be analysed.
- Mapping is frequently restricted to known traits for which a well-defined bio-assay is available.
- Validation of the marker is required in populations other than the original mapping population to ensure that the selected allele has a similar effect in different genetic backgrounds.

These limitations in existing mapping strategies can be addressed through association or linkage disequilibrium (LD) mapping. LD mapping is based on seeking associations between phenotype and allele frequencies.

Linkage disequilibrium (see also Chapter 2) is based around the association of marker loci with traits at the population level. Equilibrium is seen in large populations over many generations where there is no selective advantage, or disadvantage, associated with a particular allelic combination in a region of the genome. Disequilibrium appears where selective pressure increases, or decreases, the frequency of particular alleles or allelic combinations. It can be measured through an estimation of changes in allele frequencies as a result of the selective pressure. In this way particular regions of the genome can be associated with particular traits.

There are four potential advantages of this approach in mapping genes in crop species.

- It provides a new perspective for trait mapping, because it uses population structures (based largely around pedigree) and phenotypic data that differ from those used for full map construction of BSA. Consequently, we can expect to see new marker trait associations and targets for more detailed analysis.
- LD mapping also provides detailed fingerprinting information on a large number of lines and varieties and this information will be valuable in several of the breeding strategies outlined below.
- The LD method uses real breeding populations, the material is diverse and relevant and the most important genes (for adaptation, etc.) should be segregating in such populations. The breeder is also integrally involved in the process and this may lead to improved rate and efficiency of validation and adoption. Plant breeders are often reluctant to grow and assess a huge number of lines with little or no potential for direct commercial outcome, as required for complete map construction. The advantage of LD mapping to the breeder is that mapping and commercial variety development is conducted simultaneously.
- Pattern analysis of marker data might detect complex combinations (even epistatic interactions) between alleles at several loci that underlie the superior individuals in a breeding population. This might prove difficult to isolate and validate using the full mapping approach.

Molecular markers offer an easily quantifiable measure of genetic variation within crop species. Many crop species are based on a narrow germplasm pool and display a low level of polymorphism between cultivars. This has hampered the identification of molecular markers linked to agronomically important traits, complicating the differentiation of varieties and the analysis of genetic variability.

19.3 IMPLEMENTING MAS IN BREEDING PROGRAMMES

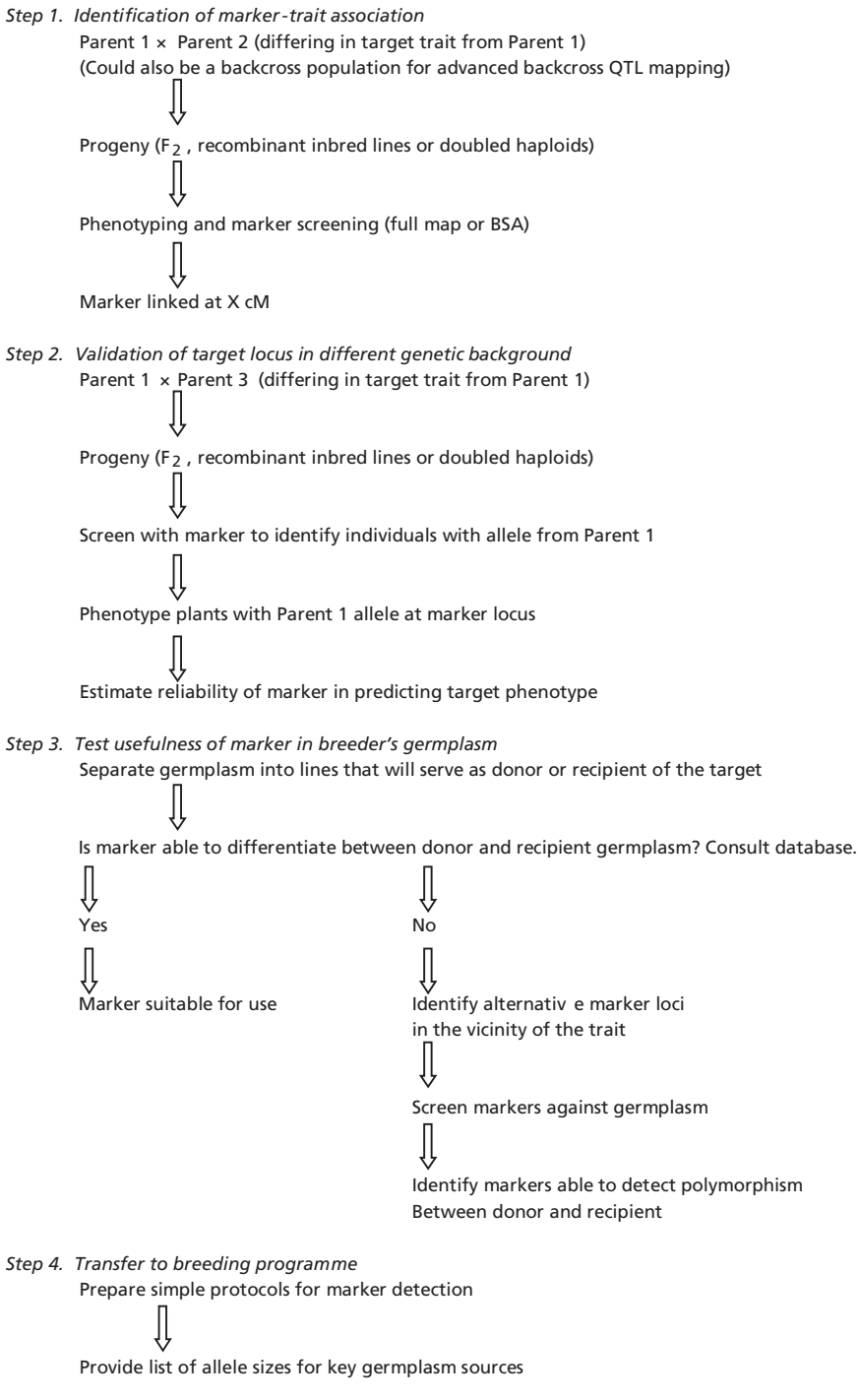
The identification of markers associated with a trait of importance represents a first step in marker application. Several further steps are needed before a marker can be used in a practical breeding programme (outlined in Figure 19.3). The key element is marker validation.

19.3.1 Marker validation

Irrespective of the technique used to identify a marker linked to a target trait, the association has been found in a particular set of circumstances, usually in the progeny of a specific cross (in fully mapped, BSA, pedigree analysis and partial mapped methods). The reliability of the marker can be estimated from the closeness of linkage. As the next step, the ability of the marker to predict the target phenotype must be tested in further populations. Usually one would aim to keep one parent in common in the first test. Therefore, the predictive value of the marker in identifying the phenotype of the plant can be estimated and compared with the original mapping result.

The usefulness of the marker in screening and assaying germplasm in use in the breeding programme is assessed as the next step. The parents likely to be used in the marker screening programme are screened for polymorphisms between those that have and do not have the target trait. Given the narrow germplasm base of many crop plants, a high proportion of markers will fail

FIGURE 19.3
Development, validation and implementation of markers for MAS schemes



this test of applicability. It is then necessary to see if other markers can be found in the vicinity of the target locus that could be used in screening. Ideally one would aim to have about ten marker loci within 10 cM of the target locus to have a reasonable chance of being able to track the locus in diverse germplasm. At this point, markers such as SSRs are particularly valuable, since they tend to detect a greater level of diversity than RFLP-based marker systems.

Finally, a clear protocol for use of the markers must be provided. This will include the following information:

- which molecular marker to use for a given germplasm combination;
- the protocol for the marker assay and the expected size of the alleles;
- the reliability of the marker in terms of the actual success rate of the assay. For example, SSR markers can vary in their ease of assay, with some working almost 100 percent of the time and others showing less than 80 percent success rate. This is often due to the sequence of the primers; and
- the closeness of linkage. This provides a measure of the success of the marker in predicting the target locus or allele.

19.3.2 Structure and function of marker implementation laboratories

Many marker implementation laboratories have evolved from molecular biology research laboratories, and were therefore not well prepared for the demands of achieving regular, high-throughput marker results for demanding plant breeders. Alternatively, some of the newer marker labs have been added *de novo* to plant breeding facilities and may lack the technical backup of adjacent molecular biologists. In either scenario, many marker labs have had difficult gestation periods before they satisfactorily service their plant-breeder clients.

Several crucial factors are required to achieve success in marker implementation labs:

- the plant breeders should be involved in the management and day-to-day activities of the marker lab;
- staff should be selected with technical skill to develop and implement markers, but also with a 'service ethic' that recognizes the importance of timeliness and accuracy;
- the target throughput over time needs to be identified so that the optimum technology platforms can be applied;
- a capital budget is needed that allows for upgrades in the marker platforms;
- the logistics of the entire process needs to be considered, from growing the plants in easy-to-sample, and subsequently cull, containers (often with bar code identification), rapid and accurate label generation for plant and DNA samples, and data management;
- a bio-informatics capability is needed that manages the plants, plant samples, DNA, marker data and selection decisions, and is compatible with the plant breeder's software; and
- access to research molecular biologists should be available for problem solving and rapid recognition of worthwhile new technology.

19.3.3 Marker types

Markers need to be easy and cheap to use. The first widely used marker system was based on RFLPs. They were expensive and technically demanding to apply to many species, particularly for polyploid species or those with a large genome size, such as wheat or onion. Emphasis then shifted to microsatellite markers and AFLPs—

both PCR-based assay systems—that are cheaper and more easily used than RFLPs. The protocols for various marker systems and the pros and cons of the options are widely discussed in the literature and so are not dealt with here (see e.g. Higgins *et al*., 2003; Russell *et al.*, 1997).

Linked markers

Most of the markers deployed in plant breeding at present are linked to the trait of interest, and alleles are detected via polymorphism in non-functional DNA, not variation within the functional genes. This places inherent limitations on the use of the marker screening data. The need to test all parents in the current crossing block to determine the allele associated with the desired genotype has already been discussed. However, it is also necessary to modify the population size of the selected group to allow for recombination between the marker locus and the functional gene.

Diagnostic markers (sometimes called 'perfect markers')

The attraction of diagnostic markers (where the marker is derived from the actual functional gene sequence or shows no recombination with the target gene) is almost overwhelming, particularly when coupled with a simple analysis system, such as SNP. It may be feasible to delete (or at least greatly reduce) the time consuming validation and parental polymorphism stages, and increase the precision of marker assays. However, the costs of developing such markers, while obviously not a limitation for the human genome projects, may be more prohibitive for many crops.

Detailed studies of the structure of key candidate genes, and the characterization of alternative alleles, are generating a new class of diagnostic molecular markers. Increasingly, significant polymorphisms are being identified within the introns of structural genes. These allelic markers may consist of insertion or deletion events, or polymorphic SSR, which are both amenable to gel-based detection systems currently utilized for molecular plant breeding. These markers offer the obvious advantage of not recombining with the gene of interest, and also save implementation resources in establishing parental polymorphisms.

As the research focus shifts from traditional QTL mapping to trait dissection and functional genomics, further candidate genes for MAS will be identified. Coupled with the broad interest in allele mining to identify functional genetic variation, identification of further diagnostic markers is expected.

19.3.4 Polymorphic markers

Mapping a QTL or gene is often only a start for the implementation of a successful MAS scheme. In many situations, several or many markers around the locus of interest are required because of lack of polymorphism among the parents used in the breeding programme. That is, the desired allele of the target is not always linked with the same allele at the marker locus.

Looking at Table 19.2, for a hypothetical cross of Variety 1 × Variety 2, MAS will be effective using either Marker1 or Marker2 as the desired "R" allele is associated with either the 'A' allele at Marker1 or the 'b' allele at Marker2. For the hypothetical cross Variety 2 × Variety 3, there is no polymorphism available between the parents at the marker2 locus, so MAS would have be based on selection at the marker1 locus where the A allele is linked to the desired R allele. For the cross Variety 1 × Variety 4, no polymorphism is available at either of the known marker loci, and so a new marker must be identified. As the number of parental

TABLE 19.2
Hypothetical parent and marker genotypes demonstrating the importance of 'marker polymorphism' in MAS schemes (see text for explanation)

Line	Genotype at Marker locus 1	Genotype at Marker locus 2	Genotype at target gene locus
Variety 1	A	b	R
Variety 2	a	B	r
Variety 3	A	B	R
Variety 4	A	b	r

genotypes increases, so too does the potential problem of lack of polymorphism, and for many breeding programmes this problem has severely curtailed the use of MAS by:

- reducing the number of crosses where MAS can be applied; and
- limiting the complexity of crossing schemes where MAS can be applied. Thus, 4-way crossing plans or convergent schemes are often difficult for MAS application due to lack of polymorphism among parents.

As mapping, QTL analysis, genomics and marker development progress, the number of markers has increased, making this problem potentially less serious in the future for the major crops. Similarly, diagnostic markers render this problem redundant.

19.3.5 Modification to breeding strategies

The power of marker-assisted breeding technologies makes it a very attractive technology for breeders. However, breeders who have been using markers for a long period make the following observations:

- Some parents and some traits do not suit MAS, due to lack of polymorphism between common parents.
- No markers may be available for one QTL contributing to the expression of a particular trait, but perhaps there are for other QTLs affecting the trait.

One extreme consequence is a risk that only parents and populations that suit MAS are constructed and selected. Hence, one must ask "Is the risk of skewing the breeding population greater than the potential gains in efficiency by using MAS?"

Computer simulations have shown that MAS is more effective than phenotypic selection when population sizes are large and heritability is low (Whittaker, Haley and Thompson, 1997; Lande and Thompson, 1990). Modelling also indicates that the combination of genotypic selection (i.e. MAS) and phenotypic selection can be especially powerful. Modification to backcrossing, pedigree and progeny schemes for self-pollinated crops and hybrid breeding to better exploit the power of MAS are discussed in Sections 19.4, 19.5 and 19.9, below.

19.3.6 Intellectual property issues

The use of some marker systems may be constrained by intellectual property protection. Further, research licences may not enable users to apply MAS for selection of commercial varieties without the consummation of a commercial licence. The simplest advice is to ensure you have freedom to operate before embarking on any MAS project.

19.4 USE OF MAS IN BACKCROSS PROGRAMMES

Backcrossing has been widely applied in breeding to transfer a simply inherited trait into an elite cultivar by repeated backcrossing to the elite variety (usually 3 to 6 times) after the initial cross between the donor parent and the elite (or more correctly: recurrent) parent. The gene of interest from the donor parent is usually tracked through the backcrossing stages by using a bio-assay for, say, disease resistance. Selfing is pursued after the backcrossing to fix the genes of interest. Field testing

follows to select the best line from amongst the selfed progeny, usually for similarity to the recurrent parent, expression of the transferred trait and sometimes for improved performance. This process is sometimes called defect elimination, especially when an otherwise elite cultivar has an obvious flaw, which can be rectified by backcrossing.

Breeders like this technique because of its relatively predictable outcomes, yet it is widely viewed as a very conservative breeding strategy and hence many breeding programmes do not allocate a large percentage of their resources to backcrossing. Further, breeders criticize backcrossing for other reasons:

- more rapid genetic improvement is possible by conventional pedigree or progeny methods of breeding;
- there may be no trait(s) of sufficient importance to warrant a dedicated backcross programme;
- an absence of a phenotypic selection system suited to rapid identification of single plants during the backcrossing phase;
- an absence of elite varieties that warrant 'defect elimination' or trait enhancement;
- a long development time; or
- there is little prospect of simultaneous improvement in more than a single trait.

However, molecular markers can benefit a backcross breeding strategy in many ways and turn it from a conservative strategy improving only one trait at a time, to a more aggressive strategy where many traits are simultaneously improved while retaining favoured linkage blocks. The benefits of MAS in backcrossing will be examined, first in general (the following paragraph) and in detail in the seven sections that follow.

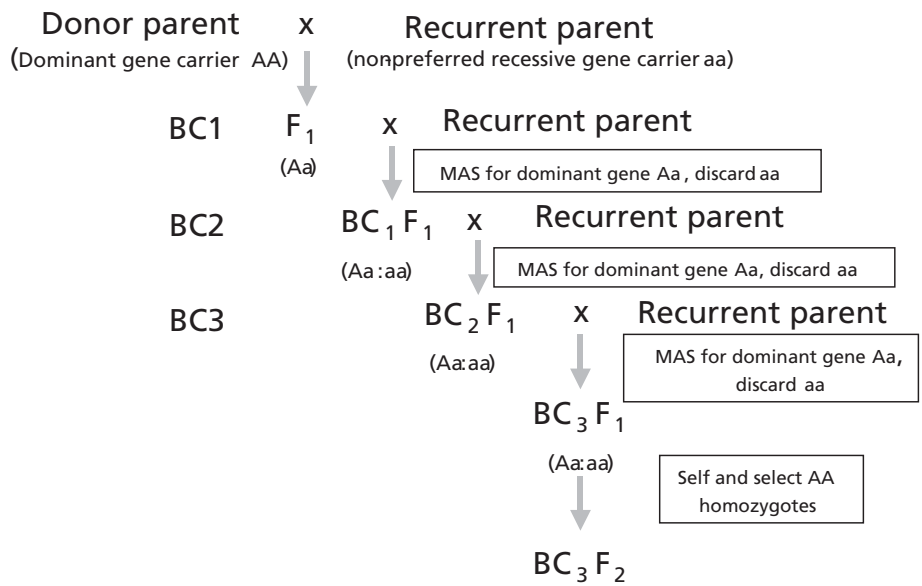
The marker genotype of an individual plant can be determined at very early development stages and therefore plants carrying the gene or genes of interest can be identified prior to flowering and backcrossed once they begin flowering. Up to three backcrosses can be made in this way in one season and plants can be grown in optimum conditions for cross-pollination. Some molecular markers, such as RFLPs, show a co-dominant mode of inheritance and therefore heterozygous individuals can be identified. This is of particular benefit when introgressing a recessive gene. Molecular marker-assisted selection is based on genotype (not phenotype) and therefore is not subject to environmental variation and lower assay error. Molecular markers can be used for selecting regions of the recurrent parent genome unlinked to the introgressed region. This reduces the number of backcross generations required to recover the recurrent parent genotype and increases the probability of obtaining a suitable introgression product.

19.4.1 Transfer of a single dominant gene

This is perhaps the simplest MAS application attempted by most breeders. Here, MAS is used to follow the gene of interest through the backcrossing phase (Figure 19.4) and may directly substitute for a bio-assay (such as a disease resistance test) or a chemical assay (such as an isozyme test). A co-dominant marker is used to distinguish the heterozygous carriers (*Aa* genotype) from the non-carriers (*aa* genotype) during backcrossing generations, to ensure that the gene of interest is not lost before the fixation (inbreeding) stage.

The beauty of the marker test here is that the plants can be sampled at the 1- to 2-leaf stage and a result is available

FIGURE 19.4
Application of MAS to transfer a single dominant gene (Genetic structure of each generation shown in brackets; MAS action shown in text box)



within 5 days, ensuring that the carriers are quickly identified, perhaps re-potted, fertilized and transferred to ideal conditions for crossing when flowering occurs. In contrast, some bio-assays (such as nematode resistance) may be so slow that the plants will have already flowered and crossing must already have taken place before the result is available, meaning that crosses made to non-carrier plants are discarded and represent a waste of effort.

Beckmann and Soller (1986) showed the frequency of the favourable allele was substantially increased where MAS was applied during backcrossing. Care should be taken to ensure that the number of individuals selected for each backcross generation allows for recombination between the marker and the gene of interest, so that false positives do not represent a risk to success.

19.4.2 Transfer of a single recessive gene

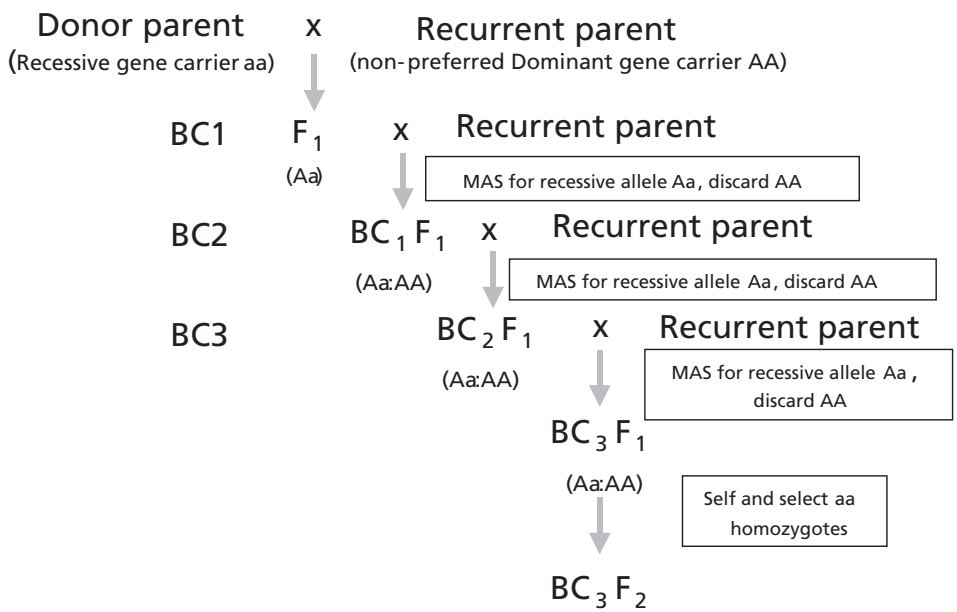
When backcrossing a recessive trait using a phenotypic screening system, a generation of selfing is required in between each backcross to expose the carriers of the recessive gene of interest; in contrast, where a co-dominant marker is available to distinguish the Aa genotype from the AA genotype (Figure 19.5), backcrossing can proceed at the same speed as for a dominant gene.

Hence, it is possible to save the selfing step required between cycles of phenotypic selection in a conventional backcrossing programme.

19.4.3 Selection of several genes simultaneously

One of the most exciting applications of MAS in backcrossing is the potential to transfer multiple traits. It is difficult to

FIGURE 19.5
Application of MAS to transfer a single recessive gene (Genetic structure of each generation shown in brackets; MAS action shown in text box)



conceive and manage practical backcrossing strategies using phenotypic selection to simultaneously transfer two traits, let alone more. In contrast, the application of MAS is only limited by handling an appropriate $BC_n \times F_1$ population to ensure recovery of the required number of individuals heterozygous for all target loci. Hence, quite complex and ambitious defect elimination strategies can be developed and pursued. This may involve introgression of several traits from a single donor parent or simultaneous defect elimination streams with different donor parents (but the same recurrent parent), which are subsequently merged to achieve the transfer of all the targeted traits.

Conceptually, this approach is quite powerful, but in practice the difficulties encountered with this approach have been:

- availability of polymorphic markers as the number of donor parents increases, although this difficulty could be reduced by fixing genes in each backcross stream, by selfing or doubled-haploid production, prior to intercrossing lines from each stream; and
- managing populations of an appropriate size.

19.4.4 Transfer of QTL

Backcrossing has traditionally not been the method of choice for the introgression of quantitative characters into breeding populations. It is often assumed that many genes with small additive effects condition quantitative traits and therefore the probability of recovering a satisfactory number of these genes through backcrossing is very small. While there have been a number of reports of success, most attempts failed to achieve full expression of the donor parent trait. Success appears to be

largely dependent on the number of genes controlling the trait, the heritability of the trait, the ease and timing (generation) of selection, and the desired level of recovery of the recurrent parent phenotype.

QTL marker-assisted selection has the potential to greatly enhance the efficiency of quantitative trait introgression, particularly those derived from exotic germplasm. In conventional selection for a trait of low heritability (most quantitative traits), an individual's phenotype is more greatly influenced by non-genetic factors, such as environment. For a mapped QTL, the phenotypic effect is estimated from the data on many individuals; thus, the influence of non-genetic factors should be reduced substantially (Paterson *et al*., 1991). Therefore genotypic selection for quantitative traits should be more efficient than phenotypic selection so long as a large proportion of the additive genetic variance is associated with the marker loci. In a backcrossing programme, the gains in efficiency are not solely restricted to reduction in selection error, but relate also to logistics and total turnover time. Backcrossing is often the most appropriate breeding strategy for quantitative traits controlled by a relatively small number of loci. As the number of loci increases, the number of backcross individuals that must be grown to have a high probability of recovering each marker allele at all loci increases to a point where backcrossing becomes inefficient.

In its simplest form, there is little difference between the backcross strategy required to transfer a QTL and that for a major gene. Nevertheless, there are important differences. First, there is ambiguity about the location of the QTL, often making the use of flanking markers necessary. This usually increases the size of the introgressed segment, unless the region of interest is finely mapped. Second, unlike a single major gene such as resistance to a nematode, where phenotypic selection during backcrossing is an alternative to MAS, the only way of selecting, for example, a complex grain quality trait is with genotypic selection using markers. It is not possible to test many aspects of grain quality on single plants grown in controlled conditions. Another issue important to QTL transfer is accuracy of selection, and recent studies show MAS significantly improved the accuracy of selection for traits of low heritability. Finally, the most important issue becomes validation of the QTL. This process involves testing the phenotypic effect of the particular chromosomal region derived from a mapping population initially in a range of related germplasm, and then in progressively less-related germplasm.

19.4.5 Selection of donor parent

Many authors have now published phylograms for varieties and genotypes in germplasm pools for major crops in many different agro-ecological zones. A similarity index based on marker data can then be used to select from among a number of possible donors for a desired trait, where the ideal is the minimum genetic distance from the proposed recurrent parent. This will theoretically reduce the number of backcrosses required to recover the recurrent parent phenotype.

19.4.6 Recovery of recurrent parent genotype

In a backcrossing programme, the donor parent genome proportion decreases with an increasing number of backcrosses. As the proportion of donor parent genome decreases, so also should the contribution of the donor parent to phenotypic expression. Therefore, if the proportion of donor

parent genome varies amongst individuals in the same backcross generation, then selection for those individuals will greatly improve the efficiency of the backcrossing programme.

In some species, the range in the percentage of donor parent marker alleles in BC_1 progenies was 8–60 percent, compared with the expectation of 25 percent. It would therefore be possible to save 1 or 2 backcrosses by using MAS to select for a high frequency of recurrent parent alleles, speeding the recovery of the background genotype.

Hospital and Charcosse (1992) and Openshaw, Jarboe and Bevis (1994) used simulations to compare the recovery of recurrent parent genotype following various backcrossing schemes in maize. These studies have further refined the strategies required to optimize backcrossing. Stam and Zeven (1981) estimated that the typical segment length of donor parent DNA retained after three backcrosses was 51 cM in a 100 cM chromosome. Frisch, Bohn and Melchinger (1999) showed that heavy selection in the BC_1 in a simulated maize system was essential to reduce the size of the donor segment, along with tightly linked flanking markers. Frisch, Bohn and Melchinger and Hospital and Charcosse both recommend two to three markers per 100 cM, distributed across the remainder of the genome, for rapid recovery of the donor parent. Frisch, Bohn and Melchinger (1999) are confident that reductions of 2 to 4 backcrosses in a 6-backcross strategy are possible.

AFLPs are well suited to this purpose, subject to an appropriate technology-use licence. SSR-based protocols are also ideal and, depending on source, may be less constrained by intellectual property restrictions. DarT technology may also have application in this area.

The minimization of the length of the intact chromosomal segment of donor type carried (or 'dragged') along with the target gene ('linkage drag') can be achieved by selecting for individuals that are heterozygous at the target locus and homozygous for recurrent parent alleles at two markers flanking the target locus. The estimation of the population size required to achieve this for various flanking marker scenarios has been made in a computer model called Popmin (Decoux and Hospital, 2002), further enhancing the sophistication and precision of the backcross strategy.

19.4.7 Accuracy of MAS during backcrossing

The next general consideration is the accuracy of MAS during backcrossing, which will be determined by:

- the genetic distance between the marker locus and the functional gene; and
- the number of markers available near the gene of interest.

For each backcross progeny selected, the probability of losing the target allele by recombination when selection is performed on a linked marker locus is simply equal to r, the recombination frequency between the marker and target loci. If this is continued for t generations of backcrossing, the probability of losing the target allele by recombination is $1 - (1 - r)^t$. Holland (2004) provides the following examples:

> *If the marker locus exhibits 10 percent recombination with the target gene, there is a 10 percent chance of losing the target allele each generation, and a 27 percent chance of losing the target allele after three generations of backcrossing. Tightly linked markers, of course, do much better: three generations of backcrossing and selection on a marker locus with 1 percent recombination*

with the target gene has only a 3 percent chance of losing the target allele.

If tightly linked markers are not available or you wish to make certain that the gene of interest in included within the introgressed segment, flanking markers can be used. Again, Holland (2004) provides a worked example:

If marker loci A and B flank the target locus, one would select backcross progeny that have both A and B alleles from the donor parent. The probability of losing the target allele with flanking marker selection is equal to the probability of selecting a double recombinant progeny from among the doubly heterozygous backcross progeny. If the flanking loci have recombination frequencies r_A and r_B, respectively, with the target locus, the probability of losing the target allele due to double crossovers within the selected region (ignoring crossover interference) is: $r_A . r_{B/(1-} r_A - r_{B+2} r_A r_{B)}$. This probability can be much lower than the probability of losing the target allele based on selection for a single marker locus. For example, if the flanking markers each have 10 percent recombination frequency with the target locus, there is only a 1.2 percent chance of losing the target allele after a single generation. In any case, with tighter linkage, the chance of losing the target allele and the amount of linkage drag are reduced.

19.4.8 Backcrossing in cross-pollinated species

Many of the opportunities and roles of MAS in backcrossing in cross-pollinated and hybrid crops are similar to those covered earlier in Section 19.4. One subtle difference is that for crops such as maize, where a great deal of information about the genetic control of characters is known, it may be more common to transfer 1 to 5 chromosomal segments associated with the expression of a semi-quantitative trait than may be the case in other crops (Stuber, Polacco and Senior, 1999).

19.4.9 Summary of backcross breeding using MAS

The advent of molecular marker technology has dramatically increased the potential efficiency, precision and flexibility of backcross breeding. However, these gains are yet to be realized by most practical breeding programmes. The first commercial varieties selected from backcrossing enhanced by MAS have now been commercialized, and more will follow shortly. The challenge is now for new crops and breeders to realize the potential of the increased power of backcrossing.

19.5 USE OF MAS IN PEDIGREE OR PROGENY BREEDING

In many breeding programmes, backcrossing accounts for only a small percentage of the total germplasm. While improvements in backcross efficiency are important, the biggest potential impact of MAS in most programmes for self pollinated plants lies in pedigree or progeny breeding systems, based on so-called 'forward crossing' strategies (Holland, 2004).

For perennials, and especially long-lived perennials, which take years to reach their reproductive stage, markers are especially valuable. For instance, if marker+trait associations for fruit quality in stone fruits or pulp yield in forest trees were available, the impact of markers would be especially high, as only those seedlings with some chance of success need to be progressed to field trials, greatly increasing the cost-effectiveness of the breeding work.

In the following sections, the important issues and opportunities in the application of MAS in pedigree and progeny breeding schemes are discussed.

19.5.1 Characterization of the germplasm pool

The first attempt of many breeding programmes to utilize molecular markers in their wider breeding programme is often the characterization of the genetic diversity of their germplasm pool (e.g. Melchinger *et al.*, 1994). Markers are chosen that have wide genomic coverage (even better if hotspots of specific interest within the genome are known), which are then assayed over key representatives of the germplasm base of the programme, including significant ancestral material, successful parents, representative genotypes from other major improvement programmes in the world, and current elite varieties.

This information can be used to:

- choose donor parents closely related to recurrent parents for backcrossing, as described above;
- choose parents with wide genetic diversity to maximize opportunities for transgressive segregation for traits of interest or to define possible 'heterotic groups'; or
- map traits based on pedigree and marker similarity, as demonstrated by Paull *et al.* (1998) for Sr 22 in wheat. Forster *et al.* (1997) demonstrated AFLP associations with salt tolerance within *Hordeum spontaneum* germplasm.

19.5.2 Linkage block analysis and selection

Breeders have long suspected that certain chromosomal regions carry critical clusters of genes (linkage blocks) that have been highly conserved during selection. The term 'national parks' has been coined to describe these linkage blocks. Few breeders have attempted to characterize linkage blocks in the breeding material and most do not have the tools to exploit this information. However, such linkage blocks have now been characterized using molecular markers in many crops, including barley and maize. The usual methodology is to establish a genotypic database of genotype × marker loci. This information can be used to develop graphical genotypes, a technique where genotypes of lines are visualized in a range of software applications, which make it easier to interpret complex datasets. These software applications can be used to select individuals that carry alleles closest to an ideal combination, as defined by the linkage block analysis. Currently, the most advanced software for graphic genotypes resides in the private sector, although several programmes are currently available in the public domain, e.g. GGT, Hypergene® and Geneflow®.

19.5.3 Key recombination events

One of the greatest positives to come from marker development programmes has been the increased knowledge of the genome and the physical and genetic control of important traits. After the first decade of research in this area, breeders can feel much more confident of designing an appropriate strategy to transfer a trait from one parent to another, or indeed improve a quantitative trait. Hence, breeding strategies can be modified to:

- recognize coupling and repulsion linkage between key traits;
- ensure that a breeding population is of sufficient size to capture the key QTLs contributing to a quantitative trait; and
- ensure key major genes form the base of each important segregating population

19.5.4 Validation of F_1s

Although most breeders prefer not to acknowledge the problem, up to 5 percent of all crosses made in pragmatic breeding programmes may be incorrect, i.e. they are selfs or encompass incorrect parents. Molecular markers are useful in identifying and eliminating incorrect crosses, and absolutely essential in the production of mapping populations, genetic studies and for crosses destined for doubled-haploid production.

19.5.5 Enrichment of complex cross F_1s

Three- and four-way crosses are attractive since they potentially allow a greater range of desirable traits to be simultaneously incorporated into elite progeny. However, in practice, many breeders find that the frequency of elite progeny is very low. Hence, they prefer to take the longer route of making simpler crosses, fixing desirable alleles and then intercrossing selected lines from each of the simpler crosses. One alternative is to use MAS to increase the desirable allele frequency for each locus contributed from a quarter parent from 25 percent of progeny, to 50 percent by screening the top cross F_1 or four-way cross F_1s. This has proven to be single biggest use of MAS in some programmes, as it has huge leverage on the frequency of desirable alleles across the whole germplasm pool.

Many traits can be screened in complex crosses, but the conundrum posed by this strategy is that although marker technology has made complex crosses more useful, the added complexity may ultimately constrain its application by lack of polymorphism among parents.

19.5.6 Pyramiding genes

A special case worth mentioning is that of 'pyramiding genes', i.e. accumulation of a number of genes affecting the same trait, e.g. to build a resistance mechanism with greater chance of showing durable resistance or gain greater expression of a quantitative trait. The use of molecular markers, and MAS in particular, offers great advantages over classical genetics, as shown in Table 19.3. Not only is the breeder more certain of the number of alleles contributing to the overall pyramid, but the time to release of the new variety is reduced by avoiding the need to dissect the pyramid using testcrosses and progeny testing.

TABLE 19.3
A comparison of building and dissecting resistance gene pyramids using classical versus molecular breeding techniques

Step	Classical breeding	Molecular breeding
Building pyramids	Make cross between different resistant parents	Make cross between different resistant parents
	Select for most highly resistant phenotype	Marker-assisted selection for different contributing loci
	Use 'ultra' virulent races, when available	Applicable to major and minor genes
	Very difficult for genes of small effect	Genotype of final lines firmly established, except for possible recombination between marker use and gene (use flanking markers)
	Genotype of final line may not be clear	
Dissecting pyramids	Retrospective analysis of developed lines required to prove all important alleles are included	No need, as genotype known from marker assay
	Resistant × Susceptible cross made, F_1 to test dominance relationships, F_2 analysis for number of genes	
	Progeny test and/or Test-cross to confirm	

19.5.7 Early generation selection

The selection theory required to implement MAS in early generations is similar to other forms of selection, although MAS is closer to 'simultaneous' rather than 'tandem' (or sequential) selection, which is often a feature of early-generation phenotypic selection. In general, breeders visually select traits of high heritability in early generations because it is not possible to effectively select for yield (and some other complex traits) in rows or small plots. Computer simulations have shown that MAS is more effective than phenotypic selection when population sizes are large and heritability is low (Whittaker, Haley and Thompson, 1997; Lande and Thompson, 1990). This challenges breeders and geneticists to design and implement practical MAS strategies to effectively select complex traits in early generations.

MAS of early-generation fixed lines is therefore the ultimate goal of the many programmes. However, at present, this is fantasy for public-sector breeders, as a number of limitations dictate that MAS can only be used in early-generation screening for very important material, comprising a small fraction of the total programme. The key limitations are the cost of DNA extraction, availability of suitable markers, staff resources for sample and data handling, and the costs (fixed and recurrent) of high-throughput systems.

19.5.8 Trait dissection

The use of QTL analysis to dissect traits is proving a powerful tool. Trait dissection relies on the co-incidence of QTLs derived for different characters, which is then used to infer associations. For instance, in barley, the QTL for a thin husk is co-incident with QTL for malt extract, inferring that malt extract (the economically important trait) is associated with a thin husk. Intuitively this was reasonable, and subsequent genetic and biochemical analysis showed this to be true. Other QTLs for malt extract were co-incident with QTLs controlling levels of starch-degrading enzymes—again, this association was subsequently proven to be causal. While trait dissection may not be as powerful an approach to the resolution of complex characters as expression profiling or functional genomics, it is certainly proving a useful, practical tool in many breeding programmes.

19.5.9 Summary of use of markers in backcrossing and pedigree and progeny breeding systems

In the previous two sections, the applications, opportunities and limitations of MAS in backcrossing and pedigree and progeny breeding systems were discussed. In Table 19.4, these roles are summarized for a typical breeding programme for a self-pollinated crop.

19.6 USE OF MAS IN THE EXPLOITATION OF PRIMITIVE GERMPLASM

Genetic diversity in modern cultivars of some crop species is diminishing due to the high selection pressure for specific quality and performance characteristics. This may ultimately lead to a reduced capacity for breeders to respond to new disease pressures and may restrict breeders from making major improvements in yield and quality parameters. Introgression of germplasm from outside the current gene pool is essential. Such introgressions have typically sought to introduce major genes, such as those for disease resistance, via backcrossing—this type of project has been a long-standing and successful part of many breeding programmes, and it is easy to envisage a role for MAS along the lines discussed earlier.

TABLE 19.4
Summary of roles for MAS in breeding programmes in self-pollinated crops

Phase	Pedigree and Progeny	Backcrossing	Marker roles
Parental choice	A × B		Understand the genetic relationships between current and future members of the germplasm pool. In hybrids, estimate likely heterosis through diversity analysis.
		A × B	Choose parents with small genetic distance to reduce number of backcrosses. Develop strategy for introgressing and selection of "quarantine traits". Transfer transgenes into elite lines
Crossing		BC_1 F_1 × A BC_2 F_1 × A BC_n F_1 × A	Identify progeny Carrier of trait Low percentage donor genome Small introgression segment Carrying domestication traits (for wild × cultivated crosses) Save 1 generation per backcross for recessive traits
	A × B		Check veracity of cross
	(A × B) × C		Enrichment of F_1s – markers used to characterize germplasm
	(A × B) × (C × D)		Enables more complex crossing strategies
Segregating generations	F_2 to F_n		High throughput markers required to select desired alleles Markers used to identify desired progeny in parent building schemes
		Fixation	Use markers to choose lines close to recurrent parent to 'fast-track' line to release by reducing evaluation requirements as before
Evaluation of fixed lines	Year 1 Limited sites, limited replications Year 2 More sites, more replications Year 3 Regional trials Year 4 National list Year 1 Year 5 National list Year 2		Limited role until QTL for adaptation/quality validated
		Year 1 Limited sites & seasons Year 2 Regional trials Year 3 National list	Whole genome marker analysis can identify individuals close to recurrent parent, thereby saving expensive yield and quality testing
Pure seed			Random genome survey plus key economic traits to ensure purity Markers used to compare re-selections for bulking
Commercialization			Use markers to provide evidence in "essentially derived" discussions Markers used to identify or compare new varieties against other varieties of "common knowledge"

However, MAS may be able to do much more. The QTLs for agronomic and quality traits in many programmes are largely from elite breeding lines and cultivars, with relatively little effort currently being invested in the identification of QTLs from outside this elite (and narrow) gene pool. Experience from some crops (e.g.

tomato, barley, wheat), indicate major improvements in quantitative characters can be made through the exploitation of exotic germplasm. However, this is a difficult and time consuming task for most plant species, and presents two major challenges, both of which can be better met with MAS than with conventional tools. These are considered in the following sections.

19.6.1 Backcross strategies to manage domestication traits

The domestication process for annual plants often requires modification of:

- seed dormancy (usually a reduction);
- seed dispersal (improve seed retention for mechanical harvesting);
- uniformity (greater synchrony of flowering to ensure more even ripening); and
- removal of leaf, seed or fruit toxins and anti-nutritional factors.

An example of the domestication process in the latter twentieth century was the conversion of the wild blue lupin (*Lupinus angustifolius*) to a sweet, domestic white lupin used widely for stock feed. While some of the domestication traits were dominant alleles expressed so that visual (phenotypic) selection was possible, seed toxins and dormancy selection was more difficult, with the result available after the opportunity for crossing had passed. More recent attempts to domesticate the related species *L. atlanticus* and *L. pilosus* have been greatly enhanced by use of MAS for the domestication traits, applied to seedling plants so that selected individuals could be crossed in that generation. Where further traits are required from wild germplasm, gene introgression would be greatly facilitated by MAS against all weedy traits during the backcrossing phase.

19.6.2 'AB-QTL' techniques to identify and introgress QTLs

Unadapted or wild-species germplasm has been used mainly as a source of major genes for disease and insect resistance, which can be readily introgressed into adapted types through backcrossing. While it is relatively simple to identify disease resistance in unadapted germplasm, it is difficult to identify accessions that are likely to carry genes for quantitative traits, as the unadapted germplasm is almost always inferior to adapted germplasm for these traits (Tanksley and Nelson, 1996). Several studies in tomatoes have demonstrated that mapped QTLs isolated from wild accessions can substantially improve commercial tomato varieties. Traditional QTL mapping techniques, however, have several deficiencies when applied to unadapted germplasm. These include the high frequency of undesirable genes (e.g. shattering and sterility), making the collection of meaningful data difficult. Also, epistatic interactions are statistically difficult to detect, yet are likely to occur in high frequency in conventional populations (the most desirable QTLs are those not requiring epistatic interactions).

Tanksley and Nelson (1996) proposed that these problems could be surmounted with QTL analysis by delaying analysis to an advanced backcross generation (BC_2 or BC_3). This would overcome agronomic problems associated with individuals within a mapping population carrying a high proportion of wild germplasm. The detection of QTLs with epistatic effects would be much reduced. Deleterious effects due to linkage drag would also be less likely.

The potential of this method to produce commercial cultivars has been successfully demonstrated in tomatoes by Tanksley *et al.* (1996). The potential in other crop species is currently being tested, with the first

papers on wheat, barley and rice just being published. It is likely to be more difficult in those crops with more complex genomes. However, this methodology may facilitate the exploitation of relatively underutilized genetic variation in wild ancestors and related landraces of many crop species.

19.7 USE OF MAS IN PARTICIPATORY PLANT BREEDING STRATEGIES

Participatory plant breeding has been proposed as a useful method of simultaneously achieving genetic progress through plant breeding, as well as facilitating seed production and variety adoption. The technique has special application in resource-poor farming communities. The basic principles of a participatory scheme are that:

- the breeder generates a relevant pool of genetic variability for the target environment;
- the breeder multiplies seed of these segregating populations for distribution to local participators (usually farmers or locally-based agronomists);
- the breeder facilitates trials of these populations;
- the participators identify the single plants or lines they feel are relevant to their agro-ecological zone and end uses;
- seed is multiplied by the breeder, who then facilitates merit trials at local level;
- the participating farmers and breeder jointly select elite lines, which are further multiplied for regional testing and seed production; and
- the participating farmers spread seed of new elite material for commercial production.

Similar schemes are now widely used in the CGIAR network and their full potential has still to be unlocked. MAS may have some useful roles to play in participatory schemes.

19.7.1 Enrichment of germplasm pools

The success of the participatory approach will depend on the frequency of desirable alleles in the segregating populations supplied to the participating farmers. Hence, there is a role for MAS applied to pre-select TC_1F_1 or BC_1F_1 individuals carrying desired alleles, or, if greater marker throughput is possible, individuals could be selected from segregating populations for distribution to the participating farmers.

19.7.2 Pre-breeding for the participating farmers

If the participating farmers in the breeding programme have a set of minimum criteria (such as disease resistance genes, plant development or stature genes, quality traits, etc.), which they would like in a range of adapted backgrounds from an earlier cycle of the participatory breeding process, these could be provided to the breeder for enrichment before return to the participating farmers.

19.7.3 Seed production

A successful participatory programme will produce many lines for commercialization for different, small, regional areas. The task of seed production and maintenance can therefore be complex. Marker fingerprinting of lines could be used to ensure the integrity of seed stocks, both prior to release to the participating farmers and to maintain pure seed in the longer term.

19.8 USE OF MAS IN TRANSGENIC BREEDING PROGRAMMES

MAS has been extremely important in the selection and monitoring of transgenics following the development of transgenic plants. The reluctance of the food, feed and

processing industries to accept transgenic varieties has restricted the application of the technology. However, it is hard to predict shifts that might occur in public attitudes over the next ten to twenty years—the timeframe of a breeding programme. It is therefore important that genetic engineering be included in the planning of breeding methodologies for many crops, in addition to those where commercial transgenic varieties are currently available. The expectation is that many genes will be amenable to manipulation via genetic engineering. The transgenics will need to be closely monitored through the breeding programmes to satisfy regulatory requirements and to ensure their rapid deployment. This will have the effect of increasing the number of loci that need to be tracked and will require breeding strategies that effectively integrate the transgenic screening with other aspects of MAS and breeding targets.

Some companies have combined marker-assisted identification of transgenics with Geographical Positioning Systems (GPS)-based bioinformatics technologies to track and monitor transgenics in field trials. This is often a requirement of various gene technology regulation authorities operating in many countries during the pre-commercial evaluation phase.

Markers also play a crucial role in determining the position, size and number of inserts of the functional transgenics. In addition, markers can be used to track the selectable marker used during transformation and assist in the selection of transgenics free of the selectable marker, post-transformation.

19.9 PRACTICAL APPLICATION OF MARKER-ASSISTED SELECTION IN BREEDING PROGRAMMES

In this section, six examples of practical application of MAS in the South Australian barley breeding programme are described to illustrate some of the principles described in this chapter. The progression of the type and scale of MAS can be followed by looking at papers presented at the International Barley Genetics Symposiums of 1996, 2000 and 2004 (Langridge *et al.*, 1996; Barr *et al.*, 2000, 2004).

19.9.1 Transfer of a single dominant gene

This was the first MAS application attempted by the barley programme and has passed the ultimate test: it resulted in the release of a new variety that has been a commercial success.

Example 1: Transfer of resistance to Cereal cyst nematode into the malting variety Sloop

The transfer of the gene *Ha*2 for resistance to Cereal cyst nematode (CCN) from the feed variety Chebec to the malting quality variety Sloop was initiated in 1994. MAS was applied using the RFLP marker *Xawbma*21 for 3 cycles of backcrossing. Doubled haploid plants were produced from the BC_3F_1 and 66 percent of the regenerants were classed as CCN resistant. This exceeds the expected proportion (48 percent, for a marker 3 cM distal from *Ha*2) as it seems that a region associated with improved regeneration from tissue culture is linked in coupling to the *Ha*2 gene from cv. Chebec. It is estimated that the use of MAS saved at least two years compared with phenotypic selection. The best of the CCN-resistant BC_3-derived Sloop types was released in 2002 and registered under Plant Breeders Rights in Australia as cv. Sloop SA. In 2002, all 200 tonne of breeders' seed was sold to farmers eagerly awaiting the new malting variety with the resistance required to manage CCN infestations in

their cropping rotations. By 2005, it is estimated that nearly 100 000 ha would have been sown to Sloop SA in South Australia.

Example 2: Transfer of resistance to BYDV into the malting variety Sloop

A second example is the successful transfer of Barley Yellow Dwarf Virus (BYDV) tolerance conditioned by the gene *Yd2* from cv. Franklin into cv. Sloop, using MAS with the PCR marker YLM (Paltridge *et al.,* 1998). MAS was employed through two backcross cycles and the BC_2F_2 derived lines were field tested in plots inoculated with viruliferous aphids. *Yd2* was successfully transferred to the BC_2 lines, and losses due to BYDV were dramatically reduced.

19.9.2 Transfer of a single recessive gene

Provided a co-dominant marker (such as a RFLP or SSR) is available, it is possible to avoid the selfing step required between cycles of phenotypic selection in a conventional backcrossing programme. This then makes the transfer of a recessive gene, both conceptually and practically, little different to the above example of a major gene. In barley, an example of a recessive trait with linked markers is the *mlo* gene for resistance to Powdery mildew, which has been transferred to many new lines via MAS during backcrossing.

19.9.3 Selection of several genes simultaneously

One of the most exciting applications of MAS in backcrossing is the potential to transfer multiple traits. It is difficult to conceive and manage practical backcrossing strategies using phenotypic selection to simultaneously transfer two traits, let alone more. In contrast, the application of MAS is only limited by handling an appropriate $BC \times F_1$ population to ensure recovery of the required number of individuals heterozygous for all target loci. In a defect-elimination strategy with the cultivar Sloop, the South Australian Barley Improvement Programme (SABIP) introduced genes for resistance to CCN (either *Ha2* or *Ha4*), tolerance to BYDV (*Yd2*), resistance to Spot form of Net blotch (SFNB) (*Rpt4*) and manganese efficiency (*Mel1*) into cv. Sloop in parallel backcross streams. Beginning in 1996, these independent streams have been progressively merged. Sloop types with combinations of these traits entered field trials in 2000.

The difficulties encountered with this approach have been:

- availability of polymorphic markers as the number of donor parents increases, although this difficulty could be reduced by fixing genes in each backcross stream, by selfing or doubled-haploid production, prior to intercrossing lines from each stream; and
- managing populations of an appropriate size.

Example 3: Backcross conversions — Transfer of resistance to SFNB and CCN to cv. Gairdner malting barley from cv. Keel feed barley

Cv. Gairdner is a widely adapted malting variety bred by Agriculture Western Australia. It has a good disease resistance profile, with resistance to Leaf scald, BYDV tolerance (*Yd2 locus*), Powdery mildew (presumably cv. Franklin's two genes) and net form of Net blotch (unknown), but it is susceptible to Leaf rust, SFNB and CCN. The high yielding, feed variety cv. Keel has resistance to SFNB (*Rpt4* plus another locus (or loci) conditioning adult plant resistance), Leaf scald (unknown) and CCN (*Ha4).*

We designed a fast-track strategy using single-seed descent, which was initiated in 1998, to rapidly introduce resistance to SFNB and CCN from cv. Keel into cv. Gairdner. We used MAS and bio-assays in large populations of the BC_1F_1 and BC_1F_2 to select for *Rpt4* – the gene conferring seedling resistance to SFNB. BC_1F_4 individuals were multiplied over summer and placed into Stage 1 field trials in 2000. Following phenotypic selection for SFNB and CCN resistance and Gairdner plant type, the remaining 43 individuals were tested for *Yd2*, *Rpt4*, *Bmy1* (Gairdner carries the SD1 allele for β-amylase) and three malt extract QTLs (1H, 2HL, 5H). The malt extract QTLs are important to maintain the good malting quality of Gairdner in the face of introgression of potentially undesirable alleles from the feed variety Keel. Four lines were promoted to Stage 4 trials in 2001, based on the marker profile and their field performance in 2000. In this cross, the Gairdner allele for malt extract on 1H was the most significant, providing 2 percent higher extract than the Keel allele. This strategy would have been greatly assisted by markers for CCN resistance (we did not know Keel carried *Ha 4* until 2000) and the 'second' adult plant gene in SFNB resistance. Further, it would have been useful to have access to a wider range of SSR markers and the ability to deploy them cost effectively. While this project started with a BC_1F_1 population of over 150 individuals per generation, there were still too few selected lines by the BC_1F_7. Nevertheless, promising lines completed Stage 4 trials in 2001, with excellent yield, malt quality and disease resistance. Seed (500 kg) was produced of the best 3 lines in 2002 and release was to proceed in 2006, subject to satisfactory performance in commercial malting and brewing trials in 2005.

The next phase in the Gairdner defect elimination is to introduce genes for boron tolerance, leaf rust resistance, leaf scald resistance and the SD3 allele for thermostable β-amylase. The crossing strategy involves merging five streams (Table 19.5).

By the time the seed of this complex cross is available, it is hoped that markers for the 'missing' genes will be available and that it will be in practise possible to select for the Gairdner background genotype using SSR markers. We have avoided using AFLPs for selecting the recurrent parent genotype because of the cost of implementing a commercial licence.

From Table 19.6, we can see the genotypes of the final 22 lines, which were derived from a combination of genotypic (MAS for the loci shown) and phenotypic selection (yield, malt extract, malt diastatic power, and resistance to CCN and SFNB). The graphical genotypes show the effectiveness of MAS for changing the allelic frequency at the target loci, but also show that unselected regions, such as EBMAC501 on chromosome 1H, BMAC310 and large regions of chromosome 7, have more Keel alleles than expected. These regions may carry Keel alleles that affect the traits targeted by phenotypic selection, but are as yet uncharacterized in QTL studies.

TABLE 19.5
Trait donors for five streams of the Gairdner defect elimination programme

CCN + SFNB stream	Boron stream	SD3 stream	Leaf scald stream	Leaf rust stream
Gairdner/Keel// Gairdner	Gairdner/DH115/ / Gairdner	Gairdner/SD3// Gairdner	Novel alleles from *Hordeum spontaneum*	Complex cross of Gairdner and Fanfare

TABLE 19.6

Marker genotype of 22 lines from the cross Gairdner/Keel//Gairdner assessed at 47 marker loci

Chromosome	Primer	cM	WI3584	WI3585	WI3586	WI3587	WI3588	WI3599	WI3600	WI3601	WI3602	WI3603	WI3604	WI3673	WI3674	BX98A061D/8	BX98A061D/12	BX98A061D/13	BX98A061/129	BX98A061/1130671	BX98A061/1130672	BX98A061/1241739	BX98A061/1241747	BX98A061/4515	Keel alleles	Gairdner Alleles
1H	EBMAC501	37	B	A	A	B	B	B	n/r	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	BMAC213	52	A	A/B	A	A	A	A	B	A/B	A	A	A	A	A	A	A	B	A	A	A	A	A	A		
	BMAC032	97.3	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
2H	BMAC134	4	B	B	A	B	B	A/B	A	A	B	A	B	A	A	B	A	A	A/B	B	B	A	A	A	Possible sf	
	BMAC881	71	A	A	A	A	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		Malt extract
	BMAC878	72	A	A/B	A	n/r	B	A	A	A	n/r	A	A	A	A	A	A	A	A	A	A	A	A	A		
	BMAC140	78	A	A/B	A	A/B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
	AWBMS56	85	A	A	A	A/B	B	B	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A		
	BMAC125	89.9	A	A	A	B	A/B	B	A	A	n/r	A	B	A	A	n/r	B	A	A	A	A	A	A	A		
	BMAC114	90	A	A	n/r	A	A	A	A	A	B	A	B	B	A	B	B	A	A	n/r	A	A	A	A		
3H	BMAC209	31	A	A/B	A	A	A	A	A/B	A/B	A	A/B	A	B	B	A	A	A	B	A	A	A	A	A		
	BMAC906	45	A	A	A	A	A	A	A/B	A	A	A	A	B	B	A	A	A	B	A	A	A	A	A		
	BMAC603	45	A	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	A	A	A+	A+		
	BMAC112	56	A	A	A	A	A	A	A	A	A	*	A	B	B	A	A	A	A	A	A	A	A	A		Yd2
	HVM060	76	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A/B	A	A	A	A	A		
	BMAC122	90	A	A	A	A	A	A	A/B	A/B	A	A	A	B	B	A	A	A	B	A	A	A	A	A		
	EBMAC871	90	A	A/B	A	A	A	A	A/B	A/B	A	A/B	A	B	B	A	A	A	B	A	A	A	A	A		
	EBMAC708	115	A	A	A	A	A	A	A	A	A	A	A	B+	B	A	A	A	A	A	A	A	A	A		
	HVM062	115	A	A/B	A	A	A	A	A	A	A	A	A	B	B	A	A	A	A/B	A	A	A	A	A		
4H	GMS089	23	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	n/r	B	A	A	A	A	A		
	BMAC310	29	A	A/B	B	B	B	B	A	A	B	B	A	A	A/B	A	A	A	B	A	A	B	B	A		
	BMAC53	30	B	A/B	B	B	B	B	A	A	B	B	A	A	A/B	B	B	B	B	A	B	B	B	A		
	BMAC884	39	A	A	A	A	A	A	A	A	A	A	A	A	A	B	B	B	B	A	A	A	A	A		
	HMLOH1A	70	A	A/B	B	A	A	A/B	A	A	A	B	A	A	A	A	A	B	A	A	A	B	A	A		
	BMAC419	86	A	B	*	A	*	A	A	A	A	*	A	A	*	A	A	B	A	A	A	B	B	A		
	BMY1	102	A	A/B	A	A	A	A	A/B	B	A	A	A	A	A	A	A	A	A/B	B	A	A	A	A/B		Bamy1
5H	BMAC113	29	A	B	B	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	A	B	B	A		
	BMAC887	29.5	A	B	B	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	A	B	B	A		
	BMAC096	33	A	B	B	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	A	B	B	A		
	GMS061	33	A	B	A	A	A	A	B	A	B	A	A	B	B	A	A	A	B	B	n/r	A	A	A		
	BMAC822	58	A	B	n/r	B	A	A	B	B	A	A	A	B	B	A	A	A	B	B	B	B	B	B	CCN Ha4	
	HVLEU	79	A	B	B	A	A	A	A	A	B	B	A	A	A	B	A	A	A/B	A	A	B	B	A		
	GMS01	105	A	A	n/r	n/r	A	A	n/r	A/B	n/r	n/r	A	B	n/r	A	A	n/r	B	A	n/r	n/r	n/r	n/r		Malt extract
6H	BMAC316	8	A	B	n/r	B	A	B	A	A	n/r	A/B	A	A	A	A	A	B	A	A	A	A	B	B		
	BMAC173	73	A	A	A	A	A	A	B	A	A	A	A	A	A	A	B	B	B	A	A	A	A	A		
	BMAC040	145	n/r	A	A	A	A	A	A	A	A	A	A	B	B	A	A	A	B	A	A	n/r	A	A		
7H	BMAC806	16	A	B	B	A	B	B	B	A	A	B	B	A	A	A	B	A	A	A	A	B	B	A		
	BMAC167	65	B	*	A	B	B	B	B	A	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	EBMAC827	65	B	A	A	B	B	B	n/r	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	BMAC21	70	B	A/B	A	B	B	B	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	BMAC607	74	B	A	A	B	B	A	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	BMAC817	79	B	A	A	B	B	B	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	HVSS1	79	B	B	A	B	A/B	A/B	A	A	B	B	A/B	A/B	n/r	B	B	B	A	B	B	B	A	n/r		
	AWBMS37	80	B	A	A	B	B	B	B	B	B	A	B	A	A	B	B	A	A	B	B	A	A	B		
	HVCMA	85	B	B	A/B	B	B	B	A	A	B	B	B	B	A/B	B	B	B	A	B	B	B	A	A		
	BMAG 120	100	A	B	B	*	B	B	n/r	B	A	B	*	*	*	A	A	n/r	*	A	*	B	B	n/r	seedling sf	
	BMAC135	170	A/B	B	B	n/r	A	n/r	B	B	n/r	B	n/r	B	B+	n/r	n/r	n/r	n/r	B	B	B	A/B	A/B		
C CNR			seg	seg	seg	r	s		r	seg	r	s	s	r	r	s	s	s	r	r	r	r	r	r		
	%Gairdner		77	62	78	63	60	65	73	76	64	80	73	63	64	67	65	79	56	73	72	70	75	78	70	

Keel alleles are shown in blue (with designation of "B"), Gairdner alleles are shown in yellow ("A"), heterozygotes in green ("A/B") and 'no result' is shown in white ("n/r"). The primers are shown in column 2 together with their position on a consensus map in column 3. In the column headed 'Keel alleles' are genomic regions where MAS was applied for Keel traits and in the column headed 'Gairdner alleles' are genomic regions where MAS applied for Gairdner alleles. The resistance to CCN is shown on the bottom row (as determined by a bio-assay) and can be compared to marker alleles for locus BMAC222 on chromosome 5H.

19.9.4 Transfer of QTLs

There are now several examples of backcrossing to transfer QTLs in barley. These include quantitatively inherited Stripe rust resistance (Toojinda *et al.*, 1999), distilling quality (Powell, pers. comm.), malting quality (Thomas *et al.*, 1995; Han *et al.*, 1997; Hayes *et al.*, 1993) and boron toxicity tolerance (Jefferies *et al.*, 1999). The Australian National Barley Molecular Marker Programme has progressed with the analysis of the genetic control of malt quality traits to the point where MAS is possible for malt extract, diastatic power and wort viscosity.

In its simplest form, there is little difference between the backcross strategy required to transfer a QTL and that for a major gene. Nevertheless, there are

important differences. Firstly, there is ambiguity about the location of the QTL, often making the use of flanking markers necessary. This usually increases the size of the introgressed segment, unless the region of interest is finely mapped. Finally, the most important issue becomes validation of the QTL. This process involves testing the phenotypic effect of the particular chromosomal region derived from a mapping population, initially in a range of closely related germplasm, then in progressively less related germplasm. Collins *et al.* (1999) and Coventry *et al.* (1999) have undertaken this task for malt extract and for diastatic power, respectively, and have recommended QTLs that are amenable to marker-assisted introgression. The opportunity to manipulate complex quality traits via MAS and backcrossing is now available. This was not previously possible without time-consuming progeny testing cycles in a conventional backcrossing system using phenotypic selection. The other issue important to QTL transfer is accuracy of selection. Two studies (Lande and Thompson, 1990; Zhang and Smith, 1992) have shown that MAS significantly improved the accuracy of selection for traits of low heritability.

Example 4: Feed barley conversion — Transfer of malt quality alleles from cvs. Alexis, AC Metcalfe and Haruna nijo to the disease-resistant feed barley cv. Keel.

Typically, when breeding malt quality barley varieties, we attempt to limit the percentage of feed-type germplasm in a pedigree to a maximum of 25 percent, as the genetic control of malting quality is complex and top quality is difficult to retain if we introgress too much genetic material from inferior lines. However, resistance to five leaf diseases, plus CCN, plus adaptation to our tough Australian growing conditions, is also a complex genetic assignment. Recent advances in the genetics of malt quality and in MAS for malt quality have provided the possibility of converting feed-quality lines into malting varieties – a strategy opposite to conventional wisdom.

In our programme, progress in breeding for yield and disease resistance has been more rapid in feed barley than in malting barley, for several reasons:

- the number of breeding priorities set for malting barley (>12) compared with feed (6–7) diminishes the rate of progress;
- physical quality specifications have been much easier to achieve in feed barley; and
- a wider range of germplasm is of direct use in the improvement of feed barley than in malting barley.

Consequently, there is at least 10 percent difference in yield between malt and feed varieties, as well as significantly poorer disease resistance in the malt varieties.

A current project aims to introgress the key genetic loci influencing malt extract, diastatic power and fermentability from Canadian, European and Japanese material into elite Australian feed lines. The initial focus is on the feed variety Keel (released in 1999), the second phase will focus on WI3385 (expected for commercial release in 2003), and new recurrent parents will be added as elite lines are identified in the feed barley breeding programme.

A breeding strategy to convert elite feed varieties into malting quality lines has been developed based on transferring the important genes for malting quality from cvs. Alexis, Haruna nijo and AC Metcalfe into feed varieties. The strategy is based on backcrossing with extensive

use of molecular markers for the most rapid and efficient development of new lines. In the first year of the project, this approach has successfully developed cv. Keel backcross lines (BC_1 and BC_2) that carry the key quality genes from the three malting barley parents, including malt extract, on chromosomes 1H, 2HS, 2HL and 5H, and the ß-amylase gene on 4H, which influences diastatic power and fermentability. These lines were to enter field trials in 2002 for selection based on agronomic performance, prior to yield trials and malt quality evaluation in 2003.

The key malting quality genes from cvs. Alexis, Haruna nijo and AC Metcalfe are also being combined by merging the three cv. Keel backcrossing programmes. BC_3 generations will also be developed for each of the Keel backcrossing programmes. Molecular markers will be used to identify elite individuals from these intercrosses and the BC_3F_1 generations. The elite lines will also be used to develop doubled haploid populations by the end of 2003.

19.9.5 Simple graphical genotypes to demonstrate ancestry

Example 5: Use of graphical genotypes presented in a spreadsheet to show the contribution of the landrace barley CI 3576 to modern Australian varieties.

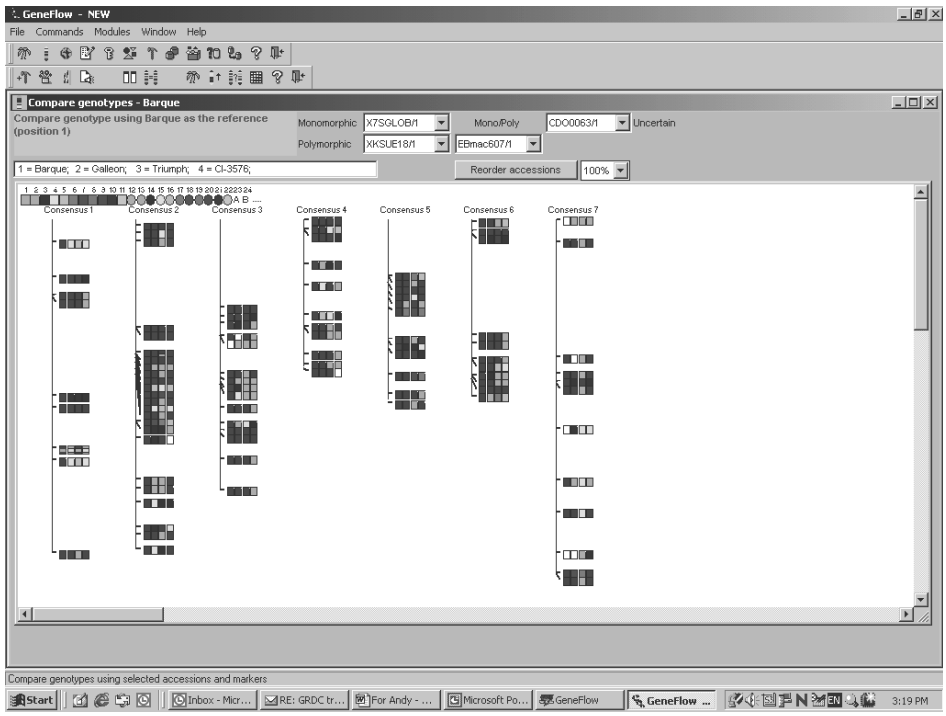
Atmojdo (2002), in an unpublished Master's thesis of the University of Adelaide, developed simple graphical genotypes of a number of Australian barley varieties of descent from an important land race known as CI 3576. The genetic composition (Figure 19.6) of the cultivars implied that something important from CI 3576 in the region around Bmac093 is retained during breeding and selection in Australia. Subsequently, mapping studies revealed that a key development locus with an allele conferring early flowering was contributed from CI 3576.

FIGURE 19.6

Graphical genotype of chromosome 2H of barley for 9 Australian barley cultivars derived from the landrace CI 3576, showing the retention of key linkage blocks associated with adaptation and disease resistance

Cultivars are shown in columns and marker loci in rows, with alleles for various SSR and RFLP loci shown both by colour (CI 3576 alleles in red) and fragment size and allelic designation. (Atmojdo, unpublished M.Ag.Sc. Thesis).

FIGURE 19.7
Graphical genotypes prepared in Geneflow® of Barque, Galleon, Triumph and CI 3576 comparing alleles at all available SSR loci



19.9.6 More sophisticated graphical genotypes to analyse flow of alleles in breeding programmes

Example 6: The use of Geneflow® software to analyse and present the contribution of parental genotypes to the breeding of the modern feed barley cv. Barque.

For whole-genome strategies to be easily interpreted and ultimately implemented in practical breeding, analysis and visualization of the data becomes crucial. In Example 5, the use of an Excel™ spreadsheet was demonstrated. A more powerful tool is Geneflow® (http://www.geneflowinc.com/). This package can manage pedigree, genotypic and phenotypic databases, and perform analyses combining this information to:

- give graphical genotypes;
- follow allele flow through complex pedigrees;
- compare genotypes from pedigree and backcross programmes against the preferred genotype; and
- much more!

One example is shown in Figure 19.7, where the contributions of parental genotypes (cvs. Triumph and Galleon – and one Galleon's key parents is CI 3576) to a new feed barley, cv. Barque, are presented across the 7 chromosomes of barley.

19.10 THE FUTURE OF MAS

There are several developments that we can expect in marker technologies over the next few years. The techniques for assaying

DNA variation between lines will become cheaper and allow higher throughput than is at present possible. While the cost per assay may come down, the cost of running an advanced molecular marker lab will go up, reflecting the need for sophisticated DNA analysis and detection systems and robotics. The cost of DNA isolation will continue to be a major limitation, although this should also come down to some extent. We can also confidently expect the number of useful traits tagged with molecular markers to increase still further. The application to crop improvement of genetic engineering, Targeting Induced Local Lesions in Genomes (TILLING), and chimeric DNA gene engineering will further increase the number of genes (loci) that will be monitored through a breeding and selection programme. These trends will encourage breeders to conduct detailed molecular analyses of lines and will greatly expand the amount of information that is available on each line. With these points in mind, we believe that the following issues will dominate the development and application of MAS.

19.10.1 Polymorphic markers

In many implementation laboratories, lack of polymorphism amongst parents is now the most important factor limiting MAS applications, especially for screening top-cross and four-way cross F_1s. Depending on the purpose of the MAS application, many more markers than are currently available will be required. For instance, for whole-genome scanning, assuming Polymorphic Information Content (PIC) = 0.25, coverage every 10 cM and a genome size of 1200 cM, 480 SSRs would be required. If MAS at targeted loci is required, the number of SSRs may in practice be around 10 to account for most parental combinations.

19.10.2 Validation of markers

Most groups undertaking MAS underestimate the cost of the validation process required after the initial identification of a potential region of interest in the mapping population. In many cases, finding a linked marker in a mapping population has proved relatively straightforward, whereas implementing it in a pragmatic breeding programme has not. The validation process is defined as the testing of an allele for its effect in genetic backgrounds other than the original mapping population, the characterization of the polymorphisms for a range of candidate markers in all combinations of parents, and the bench testing of the selected enzyme-marker/marker-protocol. Approximately 15–25 percent of the total MAS resources may be devoted to this process in some marker laboratories. The attractiveness of MAS would therefore be greatly enhanced by reducing the overheads associated with validation. It may be that diagnostic markers, such as SNPs, are one possible way to reduce these costs.

19.10.3 Marker throughput

The limitations on marker throughput were discussed briefly in the section on early-generation screening.

The primary limitations are:

- the cost and time for DNA extraction. Rapid DNA isolation systems are available that may reduce the cost but, as mentioned above, this will continue to be a major cost in MAS;
- the availability of closely linked, polymorphic markers. Public and private efforts in barley genomics may provide sufficient SSR markers to supplement those currently available and thereby overcome this limitation. Some breeders still have concerns about the use of MAS leading to narrowing of the germplasm pool to only

those parents and traits for which validated, polymorphic markers are available;
- staff resources for handling the plants, DNA and marker allele data. Many breeding programmes have funded their initial work in MAS from special funds. The time is fast approaching for public-sector breeders, and did several years ago for private breeders, where resource re-allocation within the breeding team is required to further implement MAS;
- new, robotic systems may be capable of analysing 10^4 to 10^5 loci per day, and will challenge the information processing capability of all but the largest breeding programmes; and
- the cost of robotics, Taq polymerase, non-gel-based discrimination systems, etc., for true, high-throughput systems. While such systems show great potential for handling targeted numbers, they appear at present to be beyond the budget of most barley breeders. However, SNP markers may be analysed on chips or use mass spectrophotometric methods, which could greatly reduce costs.

19.10.4 Risks and limitations of using MAS

The risk most frequently noted is the temptation to use only parents for which either markers or polymorphic markers exist, thus narrowing genetic diversity. In particular, this may concentrate use of a few, well-characterized disease-resistance genes to the exclusion of less well documented sources. This risk can be minimized by breeder discretion allocating a proportion of the programme to new or uncharacterized sources. It might be more useful to think of this problem as a challenge: How can marker technology be used to expand our useful gene pool?

There is also concern that a strong emphasis on markers in a breeding programme may lead to a focus on breeding strategies based on the technology rather than on the most appropriate strategy for a particular environment. For example, backcrossing is highly amenable to MAS, but is a conservative breeding strategy and should not become the prime focus of a breeding programme. Again, this issue can be addressed if the breeders are aware of the potential problem and include a diversity of strategies in their programme.

Another limitation is the power of QTL discrimination (Melchinger *et al.*, 1998). Many QTLs indicated in mapping population studies 'disappear' in validation populations. These QTLs may have been cross specific, subject to genotype environment interaction effects or illusory. Illusory QTLs may have been artefacts of small mapping populations, error in the phenotyping experiments or reflect fundamental limitations of QTL analysis methods.

19.10.5 Application of MAS to yield improvement

New approaches to the improvement of grain yield remain the Holy Grail of plant breeding. Yield, according to conventional wisdom, is conditioned by many genes of small effect, so how can MAS play an important role in manipulating such a complex trait? Many mapping studies have measured grain yield and applied QTL analysis (Teulat *et al.*, 2001; Marquez-Cedillo *et al.*, 2001; Baum *et al.*, 2001). Australian researchers have also measured grain yield in 10 barley mapping populations (Higgins *et al.*, 2003). The number of QTLs for grain yield varied widely, from two up to eight. This range will be due to:
- the population size of the mapping population and quality of the maps, which affects the resolution possible;

- the nature of the parents and the number of genes segregating;
- difficulties with detecting QTLs affecting yield when major genes controlling development (e.g. photoperiod response) or stature (e.g. *sdw* gene for semi-dwarfism) are segregating in the population;
- genotype × environment interaction. A number of studies report QTL × environment interactions (Piepho, 2000; Teulat *et al*., 2001; Marquez-Cedillo *et al*, 2001; von Korff *et al.,* 2008); and
- complex genetic interactions that cloud the QTL analysis.

While barley breeding programmes are successfully using MAS to transfer major genes and QTL for quality (Iguarta *et al*., 2000) and Stripe rust (Hayes, Toojinda and Vivar, 1999), progress with transferring QTLs affecting yield is less impressive. Kandemir *et al*. (2000) attempted to transfer three QTLs identified in the Steptoe × Morex mapping population into an elite malting line (cv. Morex) by marker-assisted backcrossing. The near-isogenic lines (NILs) developed did not have improved yield, although traits often related to yield, such as plant height, maturity, etc., were changed. They concluded that these yield QTLs must interact with other genes in the donor parent to give full expression, or, alternatively, they may affect harvestable yield through reduced lodging and head shattering, which may have been observed during the evaluation of the Steptoe × Morex mapping population, but not during the evaluation of the NILs.

To improve the likelihood of success, a number of steps must be taken. First, the mapping population sizes used must be increased: in the Australian barley context, population sizes range from 80 to 250, with a median of 180. The design and analysis of mapping population experiments is also important, and with appropriate techniques, such as spatial analysis and simultaneous yield and QTL analysis, it will be possible to more accurately define yield QTLs. Alternatively, different approaches, such as association mapping, may be more suitable. Clearly, it is crucial to know what other traits are coincident with the yield QTLs and it is important to seek the underlying cause of the effect of a QTL on yield. Only then will it be possible to confidently expect progress in this challenging area.

Mapping techniques may never have the precision to confidently define more than a handful of QTLs for any one trait, and this role may lie with genomics approaches.

In maize simulations, Bernardo (2001) shows that application of MAS in selection for traits controlled by very large numbers of genes might even be counterproductive. He argues that MAS has its biggest advantage over phenotypic selection when the number of genes is small (<10).

ACKNOWLEDGEMENT

This chapter has been prepared by the author from discussions with colleagues from the University of Adelaide and SARDI based at the Waite Campus, including Peter Langridge, Ken Chalmers, Steve Jefferies and Jason Eglinton. Peter Langridge and Ken Chalmers have graciously agreed to allow material jointly prepared for other publications for use in this chapter. Recent discussions with Haydn Kuchel, Australian Grain Technologies, have also contributed to the manuscript.

REFERENCES

Atmojdo, M. 2001. Major linkage blocks important to Australian barley breeding. M.Ag. Sci. Thesis submitted 2001, University of Adelaide, Australia.

Barr, A.R., Jefferies, S.P., Warner, P., Moody, D.B., Chalmers, K.J. & Langridge, P. 2000. Marker-assisted selection in theory and practice. *In* Proceedings of the 8th International Barley Genetics Conference, Adelaide, South Australia, October 2004, pp. 167–178.

Barr, A.R., Eglinton, J.K., Collins, H.M., Vassos, E. & Roumeliotis, S. 2004. Malting quality in barley – a 25-piece puzzle. pp. 58–71, *in* J. Spunar and J. Janikopva (eds). Oral Presentations. Proceedings of the 9th International Barley Genetics Symposium, Brno, Czech Republic, 20–26 June 2004. Available at: http://www.ibgs.cz/photos/Text/Oral_Presentations.pdf.

Baum, M., Grando, S., Backes, G., Jahoor, A., Sabbagh, A. & Ceccarelli, S. 2003. QTLs for agronomic traits in the Mediterranean environment identified in recombinant inbred lines of the cross 'Arta' × *H. spontaneum* 41-1. *Theoretical and Applied Genetics*, 107: 1215–1225.

Beckmann, J.S. & Soller, M. 1986. Restriction of fragment length polymorphism in plant genetic improvement. *Plant Molecular Cell Biology,* 3: 197–250.

Bernardo, R. 1999. A model for marker-assisted selection among single crosses with multiple genetic markers. *Theoretical and Applied Genetics*, 97: 473–478.

Bernardo, R. 2001. What if we knew all the genes for a quantitative trait in hybrid crops? *Crop Science,* 41: 1–4.

Collins, H.M, Logue, S.J., Jefferies, S.P., Stuart, I.M. & Barr, A.R. 1999. A Study of the physical, biochemical and genetic factors influencing malt extract. pp. 2.44.1–2.44.6, *in* Proceedings of the 9th Australian Barley Technical Symposium.

Coventry, S., Barr, A.R., Jefferies, S.P., Chalmers, K., Logue, S.J. & Langridge, P. 1999. Validating Quantitative Trait Loci for diastatic power in populations of malting barley (*Hordeum vulgare L.).* pp. 3.16.1–3.16.5, *in* Proceedings of the 9th Australian Barley Technical Symposium.

Decoux, G. & Hospital, F. 2002. Popmin: A programme for the numerical optimization of population sizes in marker-assisted backcross programmes. *The Journal of Heredity*, 93(5): 383–384.

Eisemann, B., Banks, P., Butler, D., Christopher, M., DeLAcy, I., Jordan, D., Mace, E., MacGowan, P., McIntyre, L., Poulsen, D., Rodgers, D. & Shepherd, J. 2004. Pedigree-based genome mapping for marker-assisted selection and recurrent parent recovery in wheat and barley. Poster 3.4.1 In *New directions for a diverse planet.* Proceedings of the 4th International Crop Science Congress, Queensland, Australia, 26 September–1 October 2004 (available at http://www.cropscience.org.au/icsc2004/poster/3/4/1/1422_eisemannb.htm; accessed 16 Jan. 2009).

Frisch, M., Bohn, M. & Melchinger, A.E. 1999. Comparison of selection strategies for marker-assisted backcrossing of a gene. *Crop Science,* 39: 1295–1301.

Forster, B.P., Russell, J.R., Ellis, R.P., Handley, L.L., Robinson, D., Hackett, C.A., Nevo, E., Waugh, R., Gordon, D.C., Keith, R. & Powell, W. 1997. Locating genotypes and genes for abiotic stress tolerance in barley – a strategy using maps, markers and the wild species. *New Phytologist,* 137(1): 141–147.

Han, F., Romagosa, I., Ullrich, S.E., Jones, B.L., Hayes, P.M. & Wesenberg, D.M. 1997. Molecular marker-assisted selection for malting quality and traits in barley. *Molecular Breeding,* 3: 427–437.

Hayes, P.M., Liu, B.H., Knapp, S.J., Chen, F., Jones, B., Blake, T., Frankowiak, J., Rasmusson, D., Sorells, M., Ullrich, S.E., Wesenberg, D. & Kleinhofs, A. 1993. Quantitative trait locus effects and environmental interaction in a sample of North American barley germplasm. *Theoretical and Applied Genetics,* 87: 392–401.

Hayes, P.M., Toojinda, T. & Vivar, H. 1999. Mapping and using Stripe Rust resistance genes in a malting quality background. pp. 64–66, *in* Proceedings of the Australian Plant Breeding Conference, Adelaide, Australia.

Higgins, T.J., G.R. Steed, J.C. Fegent & S. Banerjee. 2003. Better barley faster: the role of marker-assisted selection. *Australian Journal of Agricultural Research,* 54: 1065–1408.

Holland, J.B. 2004. Implementation of molecular markers for quantitative traits in breeding programmes – challenges and opportunities. Paper 3.4, in *New directions for a diverse planet.* Proceedings of the 4th International Crop Science Congress, Queensland, Australia, 26 September–1 October 2004 (available at http://www.cropscience.org.au/icsc2004/symposia/3/4/203_hollandjb.htm; accessed 16 Jan. 2009).

Hospital, F. & Charcosse, A. 1992. Marker-assisted introgression of quantitative trait loci. *Genetics,* 147(3): 1469–1485.

Iguarta, E., Edney, M., Rossnagel, B.G., Spaner, D., Legge, W.G., Scoles, G.J., Eckstein, P.E., Penner, G.A., Tinker, N.A., Briggs, K.G., Falk, D.E. & Mather, D.E. 2000. Marker-assisted selection of QTL affecting grain and malt quality in two-row barley. *Crop Science,* 40(5): 1426–1433.

Jaccoud, D., Peng, K., Feinstein, D. & Kilian, A. 2001. Diversity arrays: A solid state for sequence information independent genotyping. *Nucleic Acids Research*, 29(4): Article e25, 7 p.

Jefferies, S.P., Barr, A.R., Karakousis, A., Kretschmer, J.M., Manning, S., Chalmers, K.J., Nelson, J.C., Islam, A.K.M.R. & Langridge, P. 1999. Mapping of chromosome regions conferring boron toxicity tolerance in barley (*Hordeum vulgare*). *Theoretical and Applied Genetics,* 98: 1293–1303.

Jordan, D.R., Henzell, R.G., Tao, Y.Z., Godwin, I.D., Cooper, M. &McIntyre, C.L. 1999. Pedigree analysis provides a useful adjunct to QTL analysis in sorghum. *In* Proceedings of the 11th Australian Plant Breeding Conference, Adelaide, Australia, 19–23 April, pp. 222–223.

Kandemir, N., Jones, B.L., Wesenberg, D.M., Ullrich, S.E. & Kleinhofs, A. 2000. Marker-assisted analysis of three grain yield QTL in barley (*Hordeum vulgare* L.) using near-isogenic lines. *Molecular Breeding*, 6(2): 157–167.

Lande, R. & Thompson, R. 1990. Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics,* 124: 743–756.

Langridge, P., Lance, R.C.M. & Barr, A.R. 1996. Practical application of marker-assisted selection. *In* Proceedings of the 7th International Barley Genetics Symposium, Saskatoon, Canada, pp. 141–149.

Marquez-Cedillo, L.A., Hayes, P.M., Kleinhofs, A., Legge, W.G., Rossnagel, B.G., Sato, K., Ullrich, S.E. & Wesenberg, D.M. 2001. QTL analysis of agronomic traits in barley based on the doubled haploid progeny of two elite North American varieties representing different germplasm groups. *Theoretical and Applied Genetics,* 103(4): 625–637.

Melchinger, A.E., Graner, A., Singh, M. & Messmer, M. 1994. Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. *Crop Science,* 34: 1191–1199.

Melchinger, A.E., Utz, H.F. & Schon, C.C. 1998. Quantitative trait locus (QTL) mapping using different testers and independent population samples in maize reveals low power of QTL detection and large bias in estimates of QTL effects. *Genetics,* 149: 383–403.

Michelmore, R.W, Paran, I. & Kesseli, R.V. 1991. Identification of markers linked to disease resistance genes by bulked seg-

regant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of the United States of America,* 88: 9828–9832.

Openshaw, S.J., Jarboe, S.G. & Bevis, W.D. 1994. Marker assisted selection in backcross breeding. pp. 41–43. In *Proceedings of the Symposium Analysis of molecular data.* Corvallis, Oregon, USA, 5-6 August 1994. American Society for Horticulture Science and Crop Science Society.

Paltridge, N.G., Collins, N.C., Bendahmane, A. & Symons, R.H. 1998. Development of YLM, a co-dominant PCR marker closely linked to the *Yd2* gene for resistance to barley yellow dwarf disease. *Theoretical and Applied Genetics*, 96: 1170–1177.

Paterson, A.H., Damon, S., Hewitt, J.D., Zamir, D., Rabinowitch, H.D., Lincoln, S.E., Lander, E.S. & Tanksley, S.D. 1991. Mendelian factors underlying quantitative traits in tomato: comparison across species, generations and environments. *Genetics,* 127: 181–197.

Paull, J.G., Chalmers, K.J., Karakousis, A., Kretdschmer, J.M., Manning, S. & Langridge, P. 1998. Genetic diversity in Australian wheat varieties and breeding material based on RFLP data. *Theoretical and Applied Genetics*, 96: 435–446.

Piepho, H.P. 2000. A mixed-model approach to mapping quantitative trait loci in barley on the basis of multiple-environment data. *Genetics,* 156(4): 2043–2050.

Russell, J.R., Fuller, J.D., Macaulay, M., Hatz, B.G., Jahoor, A., Powell, W. & Waugh, R. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRS and RAPDS. *Theoretical and Applied Genetics*, 95(4): 714–722.

Stam, P. & Zeven, A.C. 1981. The theoretical proportion of the donor genome in near isogenic lines of self-fertilisers bred by backcrossing. *Euphytica*, 30: 277–238.

Stuber, C.W., Polacco M. & Senior, M.L. 1999. Synergy of empirical breeding, marker-assisted selection and genomics to increase crop yield potential. *Crop Science*, 39: 1571–1583.

Tanksley, S.D. & Nelson, J.C. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theoretical and Applied Genetics,* 92: 191–203.

Tanksley, S.D., Grandillo, S., Fulton, T.M., Zamir, D., Eshed, Y., Petiard, V., Lopez, J. & Beck-Bunn, T. 1996. Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative *L. pimpinellifolium*. *Theoretical and Applied Genetics*, 92: 213–224.

Teulat, B., Merah, O., Souyris, I. & This, D. 2001. QTLs for agronomic traits from a Mediterranean barley progeny grown in several environments. *Theoretical and Applied Genetics*, 103(5): 774–787.

Thomas, W.T.B., Powell, W., Waugh, R., Chalmers, K.J., Barua U.M., Jack, P., Lea, V., Forster, B.P., Swanston, J.S., Ellis, R.P., Hanson, P.R. & Lance, R.C.M. 1995. Detection of quantitative trait loci for agronomic, yield, grain and disease characters in spring barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics*, 91: 1037–1047.

Toojinda, T., Baird. E., Booth, A., Broers, L., Hayes, P., Powell, W., Thomas, W., Vivar, H. & Young, G. 1998. Introgression of quantitative trait loci (QTLs) determining Stripe rust resistance in barley; an example of marker-assisted line development. *Theoretical and Applied Genetics*, 96: 123–131.

von Korff, M., Grando, S., Del Greco, A., This, D., Baum, M. & Ceccarelli, S. 2008. Quantitative trait loci (QTL) associated

with adaptation to Mediterranean dryland conditions in barley. *Theoretical and Applied Genetics*, 117: 653–669.

Whittaker, J.C., Haley, C.S. & Thompson, R. 1997. Optimal weighting of information on marker-assisted selection *Genetical Research,* 69: 137–144.

Young, D. 1999. A cautiously optimistic vision for marker-assisted breeding. *Molecular Breeding,* 5: 505–510.

Young, N.D. & Tanksley, S.D. 1989. RFLP analysis of the size of chromosomal segments retained around the *tm-2* locus of tomato during backcross breeding. *Theoretical and Applied Genetics,* 77: 353–359.

Zhang, W. & Smith, C. 1992. Computer simulation of marker-assisted selection utilizing linkage disequilibrium. *Theoretical and Applied Genetics*, 83 813–820.

CHAPTER 20

Coping with and exploiting genotype-by-environment interactions

Paolo Annicchiarico

20.1 INTRODUCTION

Future increase in agricultural production is limited by both the difficulty of expanding the cultivated land area (Evans, 1998), and the demands of a high-input model of agriculture (Conway, 1998). Ongoing climatic changes are expected to increase the frequency and severity of drought and other abiotic stresses in many tropical, subtropical and Mediterranean areas (Fischer, Shah and Velthuizen, 2002). To improve the availability and stability of crop production, national or regional breeding programmes need to define a strategy to produce (when economically convenient) improved germplasm capable of maximizing the agricultural potential of specific farming systems while minimizing the occurrence of crop failures or very low yields. Some major elements of this strategy are discussed in this chapter, which focuses on the difficulties and opportunities related to breeding for environments that may differ in space (across locations), in time (across cropping years) and in crop management. This section will introduce the main relevant concepts and definitions. Two following sections will be devoted to targeting of germplasm to attain optimal adoption, and breeding strategy. Two final sections will briefly discuss the control of micro-environmental variation and suitable support software.

Following on, the term 'genotype' usually designates a cultivar (genetically homogeneous, e.g. a pure line, or heterogeneous, e.g. an open-pollinated population) rather than an individual's genetic make-up. Environment pools the set of climatic, soil, biotic and management conditions for the crop in a given location-year (annuals) or location-crop cycle (perennials) combination. The target region of a breeding programme comprises the population of potential environments for the crop. It could be defined geographically (e.g. the cropping area for a national programme), or possibly by the relative frequency of major types of environment.

Purely environmental yield effects, reflecting the different ecological potentials of sites, years and management conditions, are not of direct concern for breeding or targeting varieties. Genotype main effects (i.e. differences in genotype mean yield) provide the only relevant information, when genotype × environment (G×E) interaction effects are absent or ignored. G×E effects that lead to inconsistent genotype ranking across environments are frequently too large to be ignored. They tend to be large when there is wide variation among genotypes for traits conferring tolerance to one or more stresses (e.g. drought; low or high temperatures; soil salinity; nutrient deficiency; pests; diseases; grazing) and, concurrently, wide variation among target environments for incidence of the same stress(es) (Kang, 1998). Landraces and old cultivars tend to respond relatively better in less favourable environments (Ceccarelli, 1994; van Oosterom, Bidinger and Weltzien, 2003), although the opposite response has also been observed for material that evolved in favourable environments (Annicchiarico and Piano, 2005). The level of matching of genetic determinants of phenological development (e.g. photoperiod and vernalization requirements) with site factors that affect the length of the growing season may also imply large G×E interaction across fairly vast regions (Wallace, Zobel and Yourstone, 1993). Finally, differences in genetic structure among genotypes may contribute to G×E interaction, because variety types characterized by low levels of heterogeneity (e.g. pure lines, clones, single-cross hybrids) or heterozygosity

(e.g. pure lines) are less buffered against environmental variation than other types, such as open-pollinated populations or mixtures of pure lines, owing to their lower richness in adaptive genes (Becker and Léon, 1988).

A few widely used terms, such as adaptation and yield stability, need be unequivocally defined at this stage. In an evolutionary biology context, adaptation is a process, adaptedness is the level of adaptation of plant material to a given environment, and adaptability is the ability to show good adaptedness in a wide range of environments. In a plant breeding context, the first two terms relate to a condition rather than a process, indicating the ability of the material to be high-yielding in a given environment or given conditions (to which it is adapted) (Cooper and Byth, 1996). Genotype adaptation is usually assessed on the basis of yield responses (although other variables, e.g. gross benefit, may be considered), and undergoes modification when better performing material becomes available. Breeding for wide adaptation and for high yield stability and reliability have sometimes been considered one and the same, insofar as the latter two terms indicate a consistently good yield response across environments. However, only the adaptive responses to locations, geographical areas, farming practices or other factors that can be controlled or predicted prior to sowing can be exploited by selecting and growing specifically-adapted genotypes. For example, the knowledge of specific adaptation to past years, as shown by positive genotype × year (G×Y) interaction effects, cannot be exploited in future years, as the climatic conditions that generate year-to-year environmental variation are not known in advance. Therefore, some authors have proposed applying the yield stability concept only in relation to genotype responses over time, using the adaptation concept in relation to responses in space (Barah *et al.*, 1981; Lin and Binns, 1988). This view, accepted here, agrees with the farmer's view that location is a constant—not variable—factor, and yield consistency over time is the only relevant component of a genotype's yield stability.

Breeding for wide adaptation aims to develop a variety that performs well in nearly all the target region, whereas breeding for specific adaptation aims to produce different varieties, each of which performs well in a definite area (subregion) within the region. Early plant breeders advocated the usefulness of selection for specific adaptation (Engledow, 1925), and Falconer (1952) suggested that specific breeding may be preferable for environmentally-contrasting subregions on the basis of selection theory. However, breeding programmes have mostly considered G×E interactions as simply a hindrance to crop improvement, while pursuing, even in less developed countries, a wide-adaptation strategy that tended to promote varieties with high yield potential alongside technical packages designed to significantly improve the environment (Simmonds, 1979: 356). This trend has been favoured by the perspectives of rapid yield gain offered by high input levels, the greater profitability of targeting seed markets in more productive areas, and the belief that selection in favourable areas produces a substantial yield gain also in less favourable areas. The difficulty in sustaining and expanding high-input agricultural systems, and the mounting evidence of the dimension of G×E effects demonstrated between favourable and stress-prone environments (e.g. Ceccarelli, 1989; Bänzinger, Bertrán and Lafitte, 1997), have led to reconsideration

of some opportunities offered by G×E interactions (Simmonds, 1991; Ceccarelli, 1996). In particular, G×E effects can be exploited to select and grow varieties that show positive interaction with the location and its prevailing environmental conditions (exploitation of genotype × location interaction), or low frequency of poor yield or crop failure (exploitation of within-site genotype × year interaction). Coping with G×E interactions, rather than ignoring them, is also required when breeding for wide adaptation (e.g. Cooper *et al*., 1995).

In fact, only the genotype × location (G×L) interaction that is repeatable in time can be exploited by selecting and growing specifically-adapted material. The non-repeatable G×L interaction is the genotype × location × year (G×L×Y) (or the genotype × location × crop cycle) interaction in the analysis of variance (ANOVA) of multi-site, multi-year data sets (also called Multi-environment Trials or MET), in which the time factor is crossed with location. This term and the G×Y interaction are pooled in the within-site G×Y interaction in ANOVA models holding the time factor nested into location. The term including the non-repeatable G×L interaction acts in all cases as the error term for (repeatable) G×L interaction (see Section 20.2.2). Analysing G×L effects instead of G×E effects makes the genotype adaptive responses consistent with the proposed concept of adaptation, and simpler to model (see Section 20.2.1). Several reports summarized by Annicchiarico (2002a) indicate the need for repetition in time of genotype testing in annual crops, because the estimation of G×L effects based on a single year's data tends to be largely inflated by non-repeatable effects, due mainly to within-site, year-to-year variation in climatic factors. Repetition in time may not be needed for perennials on the ground of results for alfalfa (Annicchiarico, 1992, 2002b), suggesting that the variation in environmental factors encountered by genotypes across a three-year crop cycle is wide enough to act as a buffer against the occurrence of non-repeatable G×L effects.

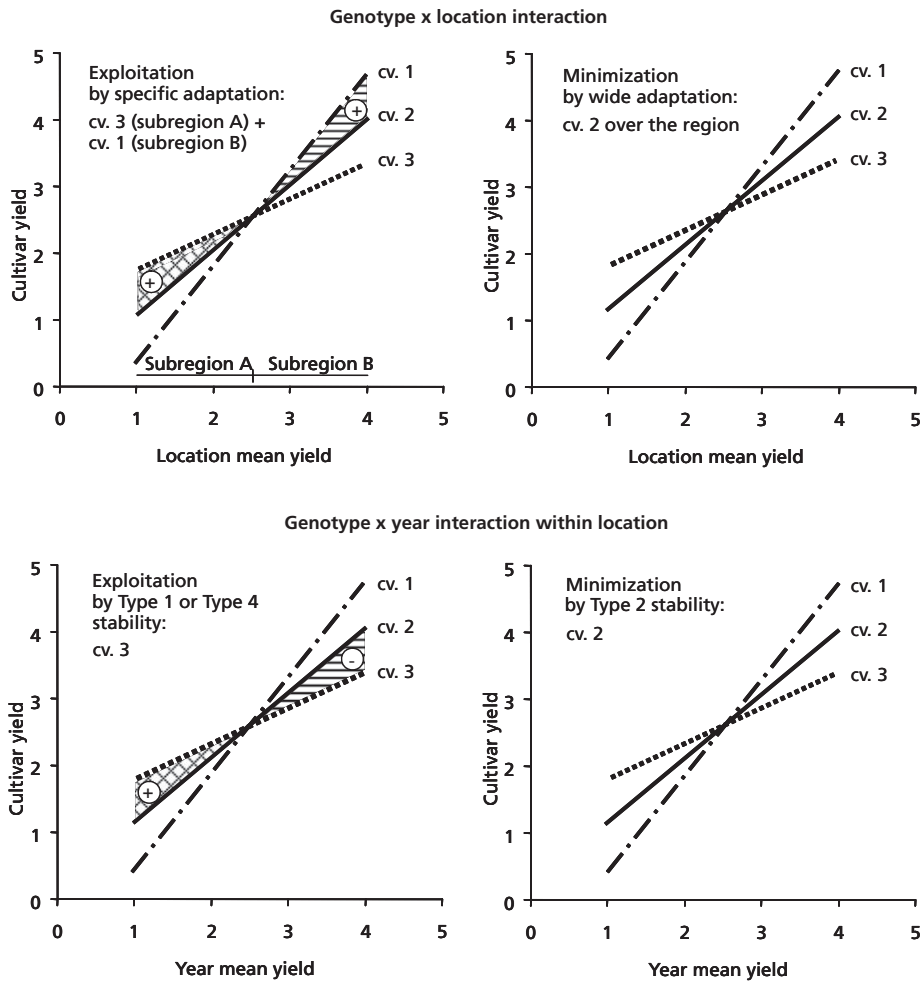
Repeatable G×L interaction effects can be exploited by breeding and growing specifically-adapted varieties, or minimized by selecting and growing widely-adapted material. The implications of either choice are shown graphically in Figure 20.1 (upper part) with respect to three hypothetical high-yielding cultivars that differ for adaptive response, as revealed by a simple model, i.e. the response to site mean yield. Minimizing G×L effects by growing the cultivar with lowest G×L interaction over all the region, i.e. cultivar 2, implies a yield penalty relative to growing specifically-adapted germplasm, i.e. cultivar 3 in the low-yielding subregion A along with cultivar 1 in the high-yielding subregion B. Likewise, aiming at selecting a variety like cultivar 2 in the context of a wide-adaptation strategy may imply lower yield gains over the region than breeding distinct germplasm for each subregion. However, the choice between exploiting or minimizing the G×L effects requires assessment of the yield gains at parity of costs (see Section 20.3.4).

The G×E interactions with the time factor can also be either exploited or minimized, by growing or selecting germplasm according to either of two concepts of stability. The first, termed 'static' by Becker and Léon (1988), is analogous to the biological concept of homeostasis: a stable genotype tends to maintain a constant yield across environments. A stable genotype according to the 'dynamic' concept of stability implies a yield response in each environment that is always parallel to the

mean response of the tested genotypes, i.e. zero G×E interaction. Lin and Binns (1988) proposed a second stability measure that refers to consistency of yield across years within location (rather than across generic environments). This measure relates to the static stability concept, but is nearer to a farmer's view of stability. The effect of the stable or dynamic concept of stability is displayed in Figure 20.1 (lower part) with respect to within-location response to year mean yield of three hypothetical cultivars. Growing cultivar 2 in all years, the most stable cultivar according to the dynamic concept, minimizes the within-site G×Y interaction effects. Growing the most stable material based on the static concept, i.e. cultivar 3, minimizes the year-to-year yield variation and implies a yield penalty in favourable, high-yielding years and a

FIGURE 20.1

Exploitation by specific adaptation, and minimization by wide adaptation, of spatial G×E interaction (above); and exploitation by the static yield stability concept, and minimization by the dynamic yield stability concept, of temporal G×E interaction (below)



yield gain in unfavourable, low-yielding years. Exploiting positive G×Y effects in unfavourable years increases the security of food production or agricultural income at national and household levels (Simmonds, 1991), making the static stability concept far more attractive than the dynamic one in a wide range of cases, and particularly for public institutions and less favoured agricultural regions.

High yield stability may be associated with low mean yield (or low stability with high mean yield), complicating variety targeting or selection. As an extreme example of high stability associated with low yield, consider a genotype that yields just above zero (high static stability) or is consistently the bottom-yielding (high dynamic stability) across years or environments. Obviously, a less stable, higher-yielding genotype would be preferable. The practical interest of combining high levels of mean yield with yield stability has led to development of the yield reliability concept. A reliable genotype has consistently high yield across years (or environments) (Section 20.2.7).

Farmers' practices frequently imply an awareness of the effects of G×E interactions. Landraces are preferred to improved varieties in stress-prone areas, reflecting their specific adaptation and higher yield stability (Almekinders, Lowaars and Bruijn, 1994). Harvest security for farmers producing near subsistence level is associated with wider crop heterogeneity, obtained by mixture and multiple cropping of different landraces (Clawson, 1985). In relatively favourable areas, mixing landrace and improved variety seed, and introgressing varietal germplasm into landraces, improves crop reliability via better response to favourable years (vom Brocke *et al*., 2003). The integration of farmers into national or regional breeding programmes fits well and enlarges the opportunities to breed for specific adaptation and yield stability (Section 20.3.1).

The adoption of biotechnologies (marker-assisted selection; genetic engineering) will hardly eliminate the need to cope with G×E interactions, because genetically based trade-offs between yield potential and tolerance to major stresses (e.g. drought; Ludlow and Muchow, 1990) and the need to choose a definite level of earliness for grain crops (Wallace, Zobel and Yourstone, 1993) will limit the possibility of assembling all useful genes in a single variety. Breeding for specific adaptation may even broaden the scope for marker-assisted selection, because a large portion of QTLs (Chapter 2 in this volume) and useful markers can be environment specific (e.g. Romagosa *et al*., 1999; Ribaut *et al*., 2007).

Targeting varieties, and defining a breeding strategy in relation to G×E interactions, are distinct objectives and will be treated separately, as they require in part different assumptions and analytical techniques, even when they can be investigated using the same set of multi-environment experiment data. Information will be given on a subset of techniques considered of primary interest on the grounds of the information generated, the ease of application (also in relation to software availability) and the limited amount of input data required. Most of them are discussed in greater detail in the book by Annicchiarico (2002a). Valuable information on general or specific aspects can also be found in the books of Gauch (1992), Basford and Tukey (2000) and Yan and Kang (2003), and in those edited by Cooper and Hammer (1996), Kang and Gauch (1996) and Kang (2002), as well as in papers cited hereafter for specific issues.

The modelling techniques described in Section 20.2 can also be used by breeding programmes for purposes other than variety targeting, e.g. for studying the adaptive responses of genotypes that represent definite ideotypes or plant types, or germplasm that differs for selection method, geographical origin or the presence of specific genes (thereby verifying different working hypotheses). The same techniques could also be applied to other contexts, e.g. for assessing the adaptation and yield stability of management practices or cropping systems (Piepho, 1998) or, more generally, for studying the interaction between two factors (one or both qualitative) (Gauch, 1992).

20.2 TARGETING CULTIVARS TO LOCATION

20.2.1 Overview

Targeting cultivars is a concern of public or private seed companies, who wish to verify the area of adaptation and the agronomic value of novel germplasm. This information, needed for proper planning of marketing and advisory schemes, is particularly useful if breeding contemplated no definite adaptation target. One genotype that is top-yielding across the target region (wide targeting), or two or more genotypes each of which is top-yielding in a distinct subregion (specific targeting), may be promoted to commercial cultivar status among several candidate entries.

Targeting cultivars is also a concern of public institutions committed to MET for defining cultivar recommendations. Different recommendation domains, i.e. subregions that are the object of specific recommendation, can be identified if locations are characterized by different top-yielding genotypes. Contemplating more than one recommended cultivar can be a sensible choice, particularly for wide subregions or genetically homogeneous variety types (pure lines, clones), as it may limit the risk of disasters arising due to the unforeseen susceptibility to a biotic or abiotic stress of the only cultivar recommended for a vast area.

Genotype targeting information can be very useful for breeding programmes also for: (i) locating elite parents for crossing, or promising populations for recurrent selection; and (ii) highlighting the success and the shortcomings of their breeding work for specific areas.

G×L interaction effects are termed as 'crossover' (alias 'qualitative') when they imply a change of genotype rank across environments, and 'quantitative' when they imply a simple variation in the extent of differences between genotypes with no change in ranking across environments. The definition of subregions as sets of locations with the same top-yielding material implies that the relevant G×L effects for targeting cultivars are the crossover type which modify the top ranks of genotypes in each environment. Subregions may be defined based on geography alone, or also according to farming practices (e.g. irrigated vs. rainfed cropping). For seed companies, subregions with limited extension or negligible advantage from specifically-adapted cultivars may be merged with larger, relatively similar subregions when the additional cost of multiplying and marketing specifically-adapted varieties is likely to outweigh the expected benefit. The recommendation by public institutions of specifically-adapted cultivars is generally convenient for maximizing regional yields and increasing the biodiversity of cultivated material.

Genotype targeting exploits the information from MET to predict yield responses in future years and, as far as possible,

TABLE 20.1
Number of subregions, and predicted and observed yield gain over the pair of most-grown cultivars, for the pair of top-yielding cultivars as predicted by different methods (observed yield data; modelled data alone or interfaced with a Geographical Information System), and comparison of models for predictive ability according to two criteria, for durum wheat in Algeria

Method [1]	No. of subregions [2]	Average gain (%)[3]		Model comparison [4]	
		Prediction [2]	Validation [5]	Sum of s_C^2 (t/ha)^2	% G×L SS / % G×L DF
Observed data	15	24.4	6.9	—	—
JR	3	12.5	10.3	0.030	5.80
AMMI	3	10.7	9.7	0.031	6.03
FR	4	10.8	11.8	0.038	3.98
AMMI + GIS	2	9.4	8.6	—	—
FR + GIS	3	11.4	9.4	—	—

NOTES: (1) Key to methods: JR = joint regression, AMMI = Additive Main effects and Multiplicative Interaction with one PC axis, and FR = two-covariate factorial regression modelling; GIS = Geographical Information System. (2) Based on modelling of 24 genotypes across 17 test sites in 1998/1999 and 1999/2000. (3) Across values of 16 individual test sites; source: Annicchiarico, Bellah and Chiari, 2006. (4) See Section 20.2.6 for explanation of criteria; source: Annicchiarico *et al*., 2002. (5) In 2000/2001.

at new sites. Modelling entry yields can clarify the adaptive responses, facilitate the variety targeting, and improve the prediction of future responses. This is showed in Table 20.1 with reference to the site-specific recommendation of two top-yielding cultivars based on cultivar yields on the site as observed (entry means across two years and four replications per year) and as modelled by each of three techniques described later (neglecting here modelling interfaced with a Geographical Information System – GIS).

In comparison with observed data, modelled data implied much fewer subregions (3 or 4 instead of 15, for 17 test sites), thereby facilitating the variety targeting, and provided 3–5 percent higher yields on average in a validation data set. Modelling was a highly cost-efficient activity here: the best model allowed for nearly doubling (+71 percent) the gain from adopting better cultivars in comparison with observed data, while requiring just a modest additional cost in comparison with the evaluation costs. In addition, observed data largely overestimated the predicted gain from improved choice of cultivars, while modelled data provided realistic predictions of these gains in the validation data set (Table 20.1). Modelled data can predict cultivar differences on the site better than observed data because all model parameters (hence, all plot values used for estimating these parameters) concur to estimate each genotype by location cell mean instead of only the plot values for the specific cell mean, thereby reducing the amount of uncontrolled error variation (so-called 'noise') in the estimated G×L effect (Gauch, 1992). For trials repeated over time, the noise relates mainly to the error term for G×L interaction (i.e. G×L×Y interaction, or within-site G×Y interaction). Modelling can decrease the number of subregions because noise effects allowing lower-yielding material to occasionally appear top-yielding have been reduced (Gauch and Zobel, 1997). Modelling can also facilitate the extension of results to new sites (see Section 20.2.6).

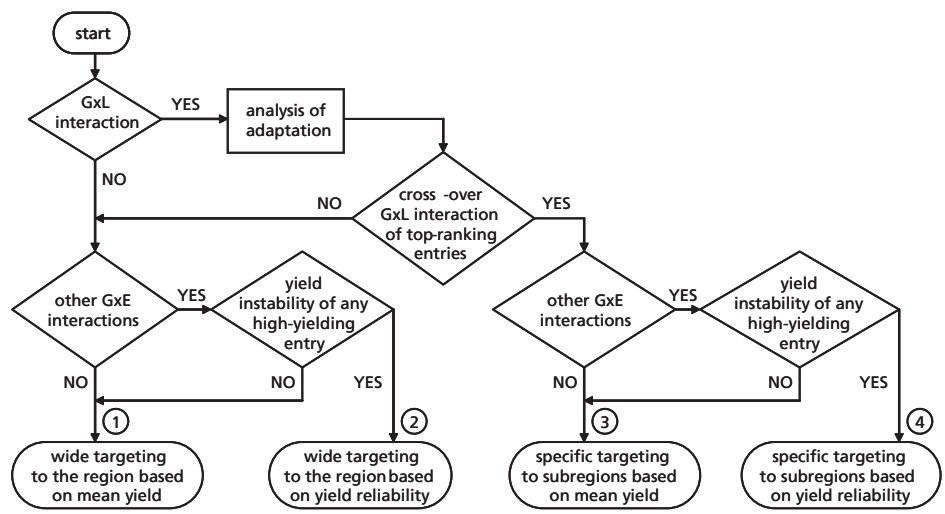
As anticipated, analysing G×L effects instead of G×E effects simplifies the modelling of adaptive responses, besides being conceptually sound. This is shown in Table 20.2 for a model described in Section

TABLE 20.2
Complexity of adaptation patterns as depicted by the number of significant principal component (PC) axes, in the analysis of genotype × environment (G×E) and genotype × location (G×L) interaction effects in six data sets

Data set	No. of sites	No. of years	No. of genotypes	Significant PC axes (1)	
				G×E	G×L
Bread wheat	31	3	18	5	2
Durum wheat 1	6	3	9	4	2
Durum wheat 2	5	2	15	4	1
Durum wheat 3	6	2	12	4	1
Maize 1	11	3	13	4	0
Maize 2	11	3	11	3	1

NOTE: (1) $P < 0.01$ according to F_{GH2} test, in an Additive Main effects and Multiplicative Interaction (AMMI) analysis.
Source: Annicchiarico (1997).

FIGURE 20.2
Flow chart of steps for targeting genotypes from analysis of multi-location yield trials repeated over time



NOTES: G×L = genotype × location; G×E = genotype × environment (environment as location–year or location–crop cycle combination).

20.2.4. Responses that are remarkably complex when evaluated on a G×E basis (requiring three or more dimensions for a convenient multivariate representation) become relatively simple on a G×L basis (requiring two dimensions at most). A simple model also facilitates the extension of results to new sites (Section 20.2.6).

The main analytical steps for targeting cultivars on the basis of multi-location yield trials repeated over time are summarized in Figure 20.2. There are four possible conclusions, which imply the targeting of same material over the region or distinct material to subregions and, in both cases, the inclusion or exclusion of yield stability

in the assessment of genotype merit. Specific targeting can be envisaged when, in the presence of significant G×L interaction in the ANOVA, the G×L effects as modelled in the analysis of adaptation imply the rank inversion of top-yielding genotypes between locations. Highly significant G×E interaction (e.g. $P < 0.001$) for ANOVA effects other than G×L interaction supports the interest of yield stability and the variety targeting based on yield reliability, but only in the presence of differences in stability among well-performing entries. Obviously, yield stability is not an issue for trials not repeated over time.

20.2.2 Types of data and ANOVA models

Genotype evaluation trials may be carried out either at research stations and experimental farms, or on farm or village land. In the former case, it is imperative that the crop management is as close as possible to that of the target population of farmers' environments (Ceccarelli, 1994). In the latter, the trials tend to have few or no replications, to reduce the number of plots per farmer site. Unreplicated trials performed contemporaneously by various farmers at nearly the same site may act as complete blocks (or large incomplete blocks) for the site, targeting two complete replications per site. Should only one unreplicated trial be adopted on each site, experimental errors may be estimated from the variation of some replicated entries randomly assigned to plots. The sample of test sites should encompass the major cropping areas and farming practices in the target region, to reflect the variation in climatic, soil, biotic and crop management factors. The number of sites may vary depending on the size of the region and the variation in environmental factors, but it should probably not fall below 5 or 6. Repetition in time is recommended for annual crops.

An inventory of useful ANOVA models with indication of relevant error terms for F tests (which depend on assumptions on genotype and location as fixed or random factors) can be found in Annicchiarico (2002a: 23). Year is always a random factor, while genotype is always fixed in the context of this section. The following class of models can be applied to trials not repeated over time (as it is frequently the case for perennials) laid out in a randomized complete block (RCB) design:

$$R_{ijr} = m + G_i + L_j + B_r(L_j) + GL_{ij} + e_{ijr}$$

where (here and in following formulae): R_{ijr} = response of the genotype i in the location j and block r; m = grand mean; G = genotype, L = location, and B = block main effects; and e_{ijr} = random experimental error. With respect to mean values of genotype (m_i) and location (m_j) and the observed genotype-location cell mean (m_{ij}), ANOVA main effects and G×L effects are estimated as:

$G_i = m_i - m$;

$L_j = m_j - m$;

$GL_{ij} = m_{ij} - m - G_i - L_j = m_{ij} - m_i - m_j + m$.

The error term for the mean square (MS) of G×L interaction is the pooled experimental error.

The following class of ANOVA models can be used for multi-location trials repeated at each site during same cropping years using a RCB design:

$$R_{ijkr} = m + G_i + L_j + Y_k + B_r(L_j\,Y_k) + GL_{ij} + GY_{ik} + LY_{jk} + GLY_{ijk} + e_{ijkr}$$

where R_{ijkr} = yield response of the genotype i in the location j, year k and block r, and Y = year (or crop cycle, for perennials)

effect. The error terms for G×E effects are: (i) G×L×Y interaction, for G×L interaction; (ii) pooled experimental error, for G×L×Y interaction; and (iii) pooled error (if location is fixed factor), for G×Y interaction.

This last class of models is particularly useful when test sites differ for number and/or timing of test years. It contemplates the year factor nested into (instead of crossed with) location:

$$R_{ijkr} = m + G_i + L_j + Y_k\,(L_j) + B_r\,(L_j\,Y_k) + GL_{ij} + GY_{ik}\,(L_j) + e_{ijkr}\,.$$

Non-repeatable G×L effects are here included in the G×Y interaction within site, which is tested on the pooled experimental error and acts as the error term for G×L interaction.

The possible heterogeneity of experimental errors is a major concern only for trials not repeated over time (where the pooled error acts as error term for G×L interaction). In this case, a transformation of variable may be envisaged when experimental errors vary as a function of the environment mean yield (Dagnelie, 1975: 367; Annicchiarico, 2002a: 28). Analysing balanced data sets is preferable (estimating missing plot values according to the design). The unbalance due to adoption of different experiment designs or to varying number of experiment replications can be overcome by performing the ANOVA on genotype–location–year cell means, estimating the pooled experimental error and converting the ANOVA sum of squares (SS) into values relative to plot data analysis, as described in Annicchiarico (2002a: 21). The absence of some genotype–location–year cell mean, or the variable number of test years per site, can be taken into account by using corrected SS (usually Type III) in the combined ANOVA, and estimating the genotype by location cell means subjected to analysis of adaptation as least squares means (i.e. adjusting for the lack of orthogonality in the data; Patterson, 1997). No observation for some genotype–location cell mean is a serious problem that may be dealt with by specific techniques and support software (e.g. Gauch, 1992: 157; 2007). A complete matrix of genotype by location cell means may also be obtained by eliminating from the data set some genotypes or locations with missing values.

Climatic, soil, biotic and crop management data of locations and morphophysiological traits of genotypes (even limited to a subset of test sites) can help identify environmental factors and adaptive traits that contribute to G×L interaction. Environmental data may also help extend the results to new sites, and are needed (from all environments) for factorial regression modelling.

20.2.3 Joint regression model

This model was developed by Yates and Cochran (1938) and proposed again, in slightly different forms, by Finlay and Wilkinson (1963), Eberhart and Russell (1966) and Perkins and Jinks (1968). In Perkins and Jinks' model (here applied to G×L rather than G×E interaction analysis), the GL_{ij} effects are modelled as a function of the location mean value (m_j) or the location main effect (L_j), which represents an indicator of the ecological potential of the site for the crop:

$$GL_{ij} = \beta_i\,L_j + d_{ij} = \alpha_i + \beta_i\,m_j + d_{ij}$$

where β_i is the regression coefficient of the genotype i and d_{ij} is the deviation from the model; in the second expression an intercept value α_i (equal to $-m\,\beta_i$) is also present. The β coefficients, with a mean value equal to zero, and the genotype mean

yields are the relevant estimated parameters of genotype adaptation. The expected (or modelled) yield response of the genotype *i* at the site *j* is:

$$R_{ij} = m + G_i + L_j + \alpha_i + \beta_i m_j = m_i + L_j + \alpha_i + \beta_i m_j .$$

Finlay and Wilkinson (1963) proposed a simpler description of genotype response to site mean yield, based on the coefficient b_i, which is equal to $(\beta_i + 1)$:

$$R_{ij} = a_i + b_i m_j$$

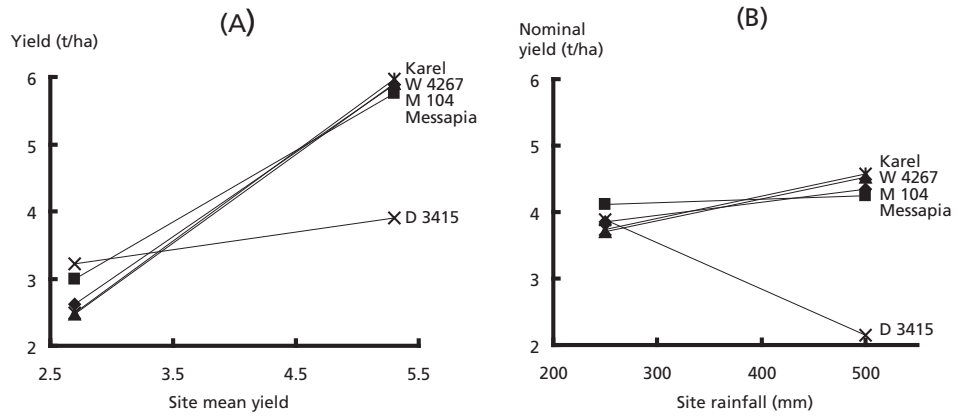
where the intercept values a_i (different from previous α_i) are equal to $(m_i - m\, b_i)$. Specific adaptation to high-yielding and low-yielding sites descends from b_i distinctly higher and lower than unity, respectively, in the presence of relatively high m_i. Conversely, b_i around unity indicates a lack of specific adaptation (and wide adaptation, if combined with high m_i). No definite indications on genotype adaptation can be inferred solely from *b* values.

The ANOVA G×L interaction SS and degrees of freedom (DF) are partitioned into two components: (i) the heterogeneity of genotype regressions, with DF = $(g - 1)$ for *g* genotypes; and (ii) the deviations from regressions, with the residual G×L interaction DF. The SS proportion accounted for by the former term, i.e. the model R^2, may be obtained: (i) by summing up the model SS across separate regression analyses of G×L effects as a function of site mean yield performed for each genotype, and calculating its proportion on the total SS of the regressions; or (ii) as the ratio of model to total SS in the analysis of covariance of all G×L effects as a function of genotype main effect and genotype × site mean yield interaction. For integration with ANOVA results (relative to plot data), the SS for the two terms can be calculated by multiplying the model R^2 and its complement to one, respectively, by the ANOVA G×L interaction SS. The genotype regressions MS should preferably be tested using the deviations from regressions MS as the error, whereas this latter MS is tested on the appropriate error MS for the ANOVA G×L interaction. It is preferable to test each *b* coefficient for difference to unity (or each *ß* coefficient for difference to zero) using the deviation from regression of the individual genotype as the error term (as provided by a separate regression analysis for each genotype). Also non-significant regression parameters should be used for modelling. The inference on expected yields is valid in the range of observed site mean yields.

Targeted genotypes depend on the site mean yield. For example, the adaptive responses of top-yielding material reported in Figure 20.3(A) suggest the presence of three subregions: (i) high-yielding (site mean yield >4.5 t/ha), where 'Karel' is top-ranking but 'W 4267' and 'M 104' yield almost as well; (ii) medium-yielding (mean yield 3–4.5 t/ha), where 'Messapia' is the top-ranking cultivar; and (iii) low-yielding (mean yield <3 t/ha), where 'D 3415' is preferable.

For this and following models, a simple outline of material statistically inferior to the top-yielding entry at variable levels of site mean yield may be provided by the mean value of Dunnett's one-tailed critical difference (Annicchiarico, 2002a: 34). The assessment is improved by computing Dunnett's critical difference for each test site on the ground of its specific error term, i.e. the G×Y interaction (for trials repeated over time) or the pooled experimental error (Annicchiarico, Bellah and Chiari, 2006).

FIGURE 20.3
Yields modelled by joint regression (A), and nominal yields modelled by factorial regression on rainfall from January to June (B), for durum wheat varieties across Italian locations



Source: Annicchiarico, 2002a.

More precise and complex methods for cultivar comparison at specific values of site mean yield (or other environmental covariate) are available (e.g. Piepho, Denis and van Eeuwijk, 1998). It is recommended to adopt less critical Type 1 error rates, e.g. $P < 0.20$, to achieve a better balance with Type 2 error rates. Even so, the site-specific recommendation of fairly large sets of statistically ($P < 0.20$) not different cultivars provided markedly lower yields than that of two top-yielding cultivars (regardless of statistical comparisons) for durum wheat in Algeria (Annicchiarico, Bellah and Chiari, 2006), confirming the high Type 2 errors (mainly due to large within-site G×Y interaction) implied by statistical differences. Type 2 errors may be very high also when the pooled experimental error acts as the error term (Kang, 1998).

20.2.4 AMMI models

The use in agricultural research of Additive Main effects and Multiplicative Interaction (AMMI) models was proposed by Kempton (1984), but became popular after Gauch's (1992) comprehensive monograph. Genotype and location main effects are estimated by ANOVA. The G×L interaction matrix is subjected to a double-centred principal components analysis (i.e. two simultaneous analyses: in the one, the genotypes are individuals and the sites original variables; in the other, vice versa), which models the G×L effects according to a multiplicative term whose estimated parameters relate to statistically significant principal components (PC) axes:

$$GL_{ij} = \sum u_{in}\, v_{jn}\, l_n + d_{ij} = \sum (u_{in} \sqrt{l_n})$$
$$(v_{jn} \sqrt{l_n}) + d_{ij} = \sum u_{in}' v_{jn}' + d_{ij}$$

where u_{in} and v_{jn} are eigenvectors (scaled as unit vectors, i.e. $\sum u_i^2 = \sum v_j^2 = 1$) of the genotype i and the location j, respectively, and l_n is the singular value (i.e. the square root of the latent root or eigenvalue), for the PC axis n; and d_{ij} is the deviation from the model. The further scaling of eigenvectors through multiplication by $\sqrt{l_n}$

allows for a straightforward estimation of the GL_{ij} effects expected on the PC axis n by multiplying the scaled genotype (u_{in}') and location (v_{jn}') scores on that axis (Gauch, 1992: 85). A genotype-location pair has a largely positive G×L effect expected on a given PC axis if the genotype and location PC scores are high and with same sign (while different signs implies a largely negative G×L effect). The simultaneous sign change of all genotype and location scores on a given PC axis leaves the estimated G×L effects unchanged.

There are several possible AMMI models characterized by a number of PC axes ranging, for g genotypes and l locations, from zero (AMMI-0, i.e. additive model) to a minimum between (g – 1) and (l – 1). The full model (AMMI-F), with the highest number of PC axes, provides a perfect fit between expected and observed data. Models including one (AMMI-1) or two (AMMI-2) PC axes are frequently appropriate in the presence of significant G×L interaction (Table 20.2). For AMMI-2 models, the scaled scores of genotypes and locations in the space of PC 1 and PC 2 may be reported in a single graph (biplot) to appreciate site or genotype similarity for G×L effects, and graphically estimate these effects. The AMMI-1 biplot displays mean values on the abscissa and PC 1 scores on the ordinate axis of genotypes and locations, showing all determinants of genotype performance.

For integration with ANOVA results, the ANOVA G×L interaction SS is divided into portions relating to each significant PC axis and to a residual term. The SS for each PC can be obtained as the proportion of G×L interaction SS accounted for by the PC multiplied by the ANOVA G×L interaction SS. The DF for the PC axis n is (Gauch, 1992: 85): ($g + l - 1 - 2n$). The G×L interaction SS and DF not accounted for by significant PC axes are pooled in the residual G×L interaction term. The F_R test is commendable for statistical testing of PC axes in a wide range of situations (Piepho, 1995; Annicchiarico, 2002a: 38). This simple test (usually not provided by statistical software) verifies the significance of the residual G×L interaction in each AMMI model, beginning with AMMI-0. By an ordinary F ratio, the MS of the residual is tested on the error MS for the ANOVA G×L interaction (for AMMI-0, the test coincides with the ANOVA F test). A significant result implies the addition of one more PC to the model (i.e. the significance of the newly-added PC).

In the selected AMMI model, the expected response of the genotype i in the location j is:

$$R_{ij} = m + G_i + L_j + \sum (u_{in}' v_{jn}') = m_i + L_j + \sum (u_{in}' v_{jn}').$$

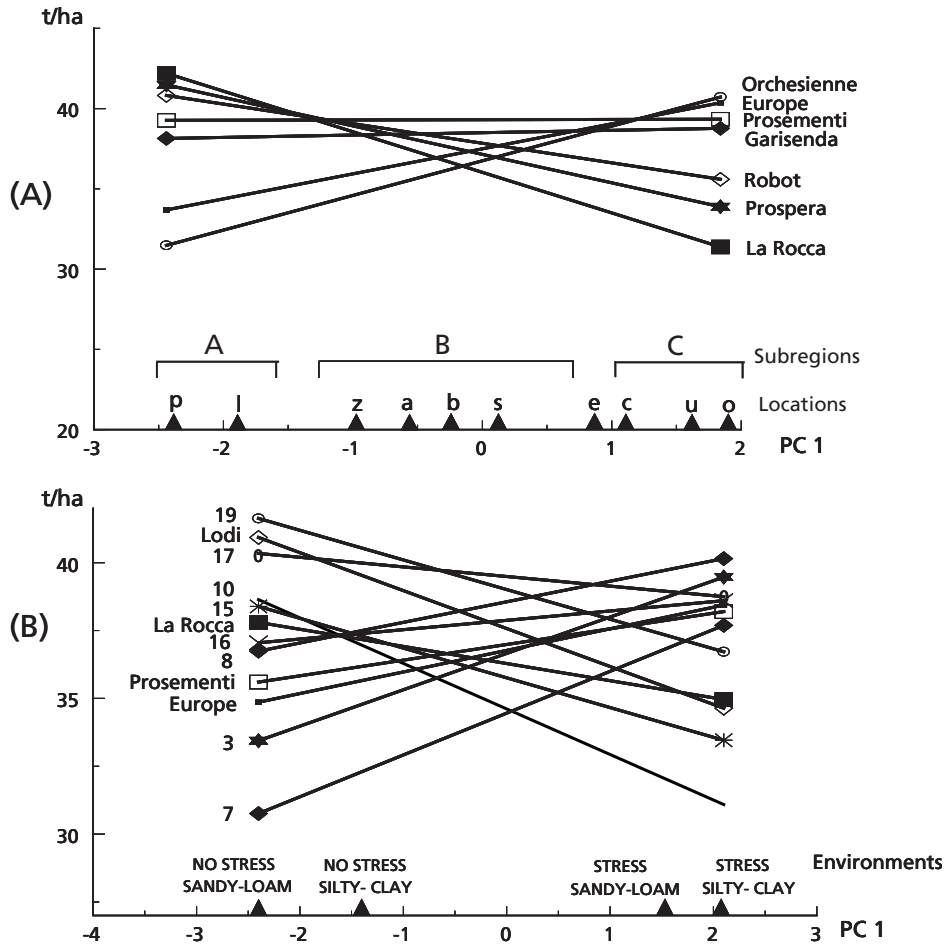
As the changes in genotype rank across sites only depend on the multiplicative term, the adaptive responses can conveniently be represented as a function of the scaled scores of locations on the statistically significant PC axes. The location main effect, which has no influence on genotype ranks and complicates the graphic representation of adaptation patterns, may be eliminated, thereby modelling the yield responses as nominal yields (Gauch and Zobel, 1997). For AMMI-1 models, nominal yields (N_{ij}) can be estimated as:

$$N_{ij} = m_i + (u_{i1}' v_{j1}').$$

They can be represented by straight lines as a function of the scaled PC 1 score of sites reported in abscissa, as shown in Figure 20.4(A) for alfalfa varieties grown

FIGURE 20.4

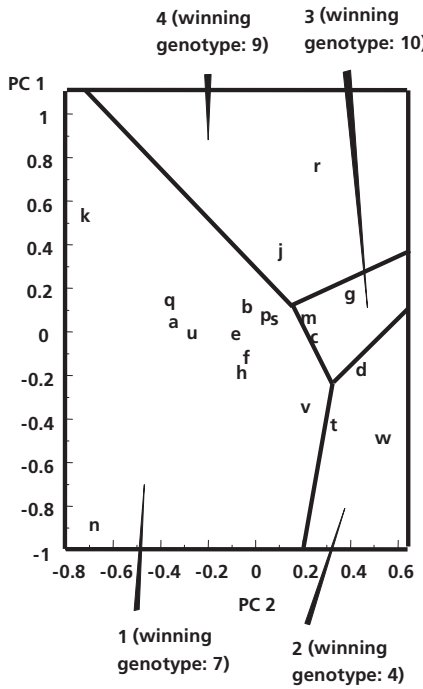
Nominal yields of top-yielding alfalfa cultivars as a function of the site score on the first G×L interaction principal component (PC) axis in a reference data set (A) or the score on the first G×E interaction PC axis of four artificial environments created by the factorial combination of two drought stress levels (almost nil or high) by two soil types (B)



Source: Annicchiarico and Piano, 2005.

in locations of northern Italy. For each genotype, N_{ij} values may be calculated just for the two sites with extreme PC 1 scores (sites 'p' and 'o'; if the software outputs R_{ij} values, $N_{ij} = R_{ij} - L_j$) and the two values connected by a straight line. The results suggest the presence of three recommendation domains: (i) 'La Rocca', 'Prospera' and (to a lesser extent) 'Robot' for the sites 'p' and 'l'; (ii) 'Prosementi' and (to a lesser extent) 'Garisenda' and 'Robot' for the sites 'z', 'a', 'b', 's' and 'e'; and (iii) 'Orchesienne', 'Europe' and (to a lesser extent) 'Prosementi' for the sites 'u', 'o' and 'c'. These subregions are almost coincident with those indicated in Figure 20.4(A),

FIGURE 20.5
Scores on the first two G×L interaction PC axes of 20 Italian locations (coded by letters), and definition of four subregions for bread wheat variety recommendation by grouping sites that have same expected top-yielding genotype (coded by number)



Source: Annicchiarico and Perenzin (unpublished results).

which actually are candidate subregions for breeding as identified by a cluster analysis of site PC scores (see Section 20.3.3).

For AMMI-2 models, the graphical expression of nominal yield responses would require three dimensions (one for each PC axis, and one for yield). However, subregions may graphically be represented as in Figure 20.5, in which the sites (displayed in the space of PC 1 and PC 2) that have same top-yielding material according to AMMI-2 modelled yields are grouped together (Gauch, 1992). This graph can also be outputted by freely-available software (see Section 20.5). In the example, genotype 7 may be targeted to 12 sites, whereas genotypes 2, 3 and 4 are of specific interest for small subsets of sites. Targeting has to be based on listed expected yields of genotypes in each test site when displaying more than one top-yielding cultivar per site in AMMI-2 models or when adopting AMMI-3 or more complex models.

Environmental variables can be related to PC scores of locations by correlation or regression analysis to reveal factors that are likely to affect G×L interaction, characterize subregions, and to scale up results (see Section 20.2.6). Likewise, correlations of morphophysiological traits (possibly recorded in a subset of sites) with adaptation parameters of genotypes (mean yield and PC scores) may reveal traits associated with specific or wide adaptation.

Other AMMI models including PC axes with different quantitative weights (instead of truncated models including or excluding a given PC axis) have been proposed to increase the predictive accuracy (Cornelius and Crossa, 1999). The genotype main effect (G) plus genotype × environment (GE) interaction (GGE) biplot analysis (Yan and Kang, 2003) is another model usable for GL interaction analysis. It applies singular value decomposition to a matrix of genotype–location cell means with the environmental effects removed (rather than a matrix of G×L effects). The first PC axis tends to summarize the genotype main effects and the other PC axes the G×L effects, but additional procedures may be needed to clearly separate these effects, while having some disadvantage relative to AMMI for graphical representations and in other respects (Gauch, Piepho and Annicchiarico, 2008).

20.2.5 FACTORIAL REGRESSION MODELS

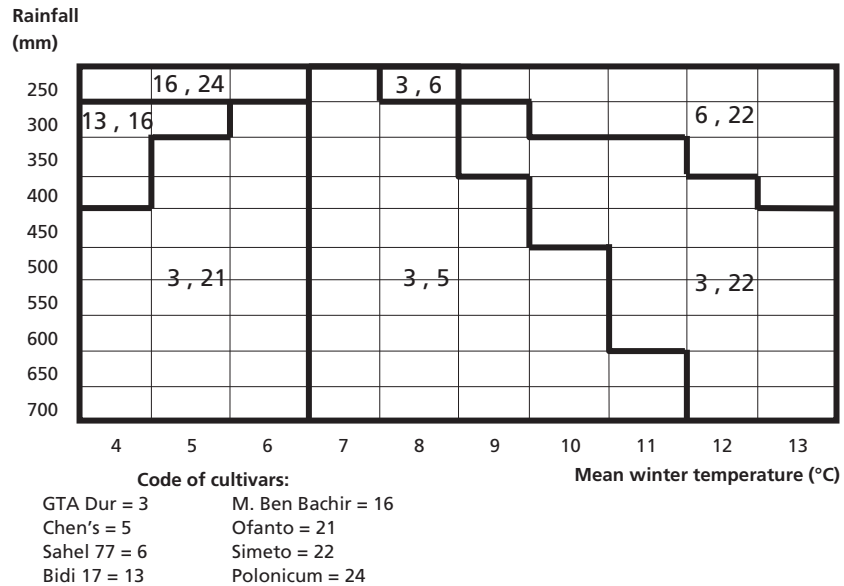
Modelling genotype adaptive responses to environmental covariates was proposed by Hardwick and Wood (1972). Covariates are usually quantitative, but qualitative ones can be incorporated through a set of dummy variables (Piepho, Denis and van Eeuwijk, 1998). Regressions are usually linear, but quadratic terms may be included as additional covariates. Denis (1988) described also the use of genotypic covariates together with environmental ones (not considered herein). The GL_{ij} effects are modelled as a function of the mean value on the site j of the environmental variable n (V_{jn}):

$$GL_{ij} = \alpha_i + \sum \beta_{in} V_{jn} + d_{ij}$$

where: α_i = intercept value, and β_{in} = regression coefficient on the covariate n, for the genotype i; and d_{ij} = deviation from the model. For one covariate equal to site mean yield, the model coincides with Perkins and Jinks' joint linear regression. Positive and negative β values indicate a positive G×L effect in sites with high and low level, respectively, of the environmental variable concerned (while specific adaptation also requires relatively high entry mean yield, besides the positive G×L effect).

The model is constructed with the progressive addition of the most important covariates (Denis, 1988). The best one-covariate model is identified on the basis of the proportion of G×L interaction SS accounted for (i.e. the model R^2), which may be computed for each covariate either: (i) by summing up the model SS across separate regression analyses of G×L effects as a function of site mean value of the covariate performed for each genotype, and calculating its proportion on the total SS of the regressions; or (ii) as the ratio of model to total SS in the analysis of covariance of all G×L effects as a function of the genotype factor and its interaction with site value of the covariate. For integration with ANOVA results, the SS for the covariate and the residual G×L interaction can be calculated by multiplying the model R^2 and its complement to one by the ANOVA G×L interaction SS. For g genotypes, each covariate accounts for $(g-1)$ GL interaction DF while the residual G×L interaction pools the remaining DF. MS values for the covariate and the residual G×L can be tested on the error MS for the ANOVA G×L interaction. After identifying the best one-covariate model, the best two-covariate model is found by comparing (on the basis of their R^2) the possible multiple linear regression models including the best single covariate (previously identified). Again, the R^2 for each two-covariate model may obtained either by summing the results of multiple regression analyses of G×L effects executed for individual genotypes, or through an analysis of covariance of all G×L effects including, beside the genotype factor, a genotype × covariate site mean value interaction term for each of the two covariates. The MS of any added covariate is the ratio between the portion of G×L interaction SS accounted for by the additional covariate, i.e. the partial regression SS (calculated as the difference in R^2 between the two-covariate and the one-covariate model, multiplied by the ANOVA G×L interaction SS), and its DF. If the added covariate of the two-covariate model with highest R^2 is significant, the best three-covariate model can be searched for, and so forth until no significant covariate can be added. The estimation and testing for difference to zero of the β_{in} parameters is limited to the selected model (preferably

FIGURE 20.6
Expected pair of top-yielding durum wheat cultivars for each of 10 annual rainfall by 10 mean winter temperature values of Algerian locations in a factorial regression model



Source: Annicchiarico *et al*., 2002.

using as error term the deviations from the model of the genotype *i*). Also non-significant estimated parameters should be used for modelling. The inference is valid in the range of observed covariate values.

Careful selection of environmental variables prior to analysis, on the basis of common sense and their putative importance in G×L interaction, is recommended, to limit the calculation process and avoid the risk of multicollinearity. This risk tends to be small (Vargas *et al.*, 1999) but may be eliminated through a partial least squares regression model (Aastveit and Martens, 1986).

The expected yield of the genotype *i* in the location *j* according to the selected model is:

$$R_{ij} = m + G_i + L_j + \alpha_i + \sum \beta_{in} V_{jn} = m_i + L_j + \alpha_i + \sum \beta_{in} V_{jn} .$$

Its expression as nominal yield:

$$N_{ij} = m_i + \alpha_i + \sum \beta_{in} V_{jn}$$

may simplify its calculation and, for one-covariate models, allows modelling of the genotype responses to the environmental variable as straight lines (connecting the N_{ij} values calculated only for the extreme covariate levels in the graph), as shown in Figure 20.3(B) for response to site mean rainfall of the same cultivars already modelled by joint regression (modelling actual yields as straight lines would require a perfectly linear response to rainfall for site mean yield). Specific targeting may be envisaged here for two subregions: (i) relatively low rainfall (<400 mm), where 'Messapia' is preferable; and (ii) relatively high rainfall (≥400 mm), where 'Karel' and 'W 4267' are preferable.

For two-covariate models, nominal yield responses would require a three-dimensional graphical representation. However, it is possible to display the expected top-yielding material and the subregions as a function of sets of site values for the covariates, as shown in Figure 20.6 for two top-yielding cultivars in each combination of 10 rainfall by 10 winter temperature levels (a denser grid of points would provide more fine-tuned indications). Targeting has to rely on lists of expected genotype yields for more complex models.

20.2.6 MODEL COMPARISON AND SCALING UP OF RESULTS

Joint regression is a simple and popular model, but cannot describe G×L effects that are ecologically complex (e.g. because of variably occurring environmental stresses) or mainly affected by a different environmental factor relative to site mean yield (Annicchiarico, 1997). Factorial regression allows for explicitly assessing the relationships of environmental variables with G×L effects, thereby improving our understanding of G×L interaction (both in general and for single genotypes). It also simplifies the definition of subregions, because the genotype responses to new sites can easily be predicted as a function of the site mean value for the significant covariates. Its use may be limited, however, by the unavailability of environmental data in the complete set of test environments, or by the modest explicative value of the available covariates. With AMMI analysis, sites with missing environmental data can be used for modelling but excluded from analyses that assess the relationships of these data with G×L effects. AMMI modelling has a broader range of application, but makes the definition of subregions less straightforward than for regression models.

The model predictive ability depends on the accuracy (high SS) and the parsimony (low DF) of its G×L interaction parameters (Gauch, 1992: 134). Uni-dimensional models (joint linear regression, AMMI-1 and one-covariate factorial regression models) can be compared for predictive ability based on the MS value of their G×L interaction parameter (which takes account of both characteristics). The sum of the estimated variances of the G×L interaction terms of the model could provide a general criterion for model comparison (Annicchiarico, 2002a: 49). The variances of G×L interaction PC axes (which are uncorrelated), or those of environmental covariates (which relate to partial regression SS for each added covariate), can be summed up because they add independent pieces of information. Extending Becker's (1984) procedure for estimating the variance of genotype regressions to PC axes or environmental covariates, the variance of any component of the G×L interaction (s_C^2) can be estimated from its MS (M_C) and the error MS for the ANOVA G×L interaction as it follows (with respect to notations in Section 20.2.2, and M_e = pooled error MS): $s_C^2 = (M_C - M_e) / r\,l$, for trials not repeated over time; $s_C^2 = (M_C - M_{GLY}) / r\,y\,l$, for location and year crossed factors; and $s_C^2 = (M_C - M_{GY(L)} / r\,y\,l$, for the year factor nested into location. Brancourt-Hulmel, Biarnès-Dumoulin and Denis (1997) proposed a second criterion equal to the following ratio of SS to DF globally accounted for by the G×L interaction parameters of the model: % G×L interaction SS / % G×L interaction DF. An empirical assessment suggested the superiority of the former criterion, whose indication of the greater predictive ability of the two-covariate factorial regression over joint regression or AMMI-1 was confirmed by the greater

ability of this model to predict the top-yielding material (i.e. the information of practical interest) in a validation data set (Table 20.1). In contrast, the Brancourt-Hulmel, Biarnès-Dumoulin and Denis' (1997) criterion ranked factorial regression as the least predictive.

The definition of subregions, initially limited to test sites, should ideally be scaled up to the entire target region. The possible procedures vary depending on the modelling technique and the availability and the type of environmental data. Whenever possible, the model established in relation to genotype responses to test year values of relevant environmental variables (site mean yield; climatic covariates; etc.) is exploited to predict future responses on the basis of long-term or mean values of these variables on new sites or test sites. Test sites may be reassigned to other subregions, if top-ranking material happens to differ for long-term conditions. One site may belong to more subregions depending on the crop management (e.g. irrigation level), if relevant. For joint regression, sites can be assigned to subregions depending on their long-term mean yield (as known from statistical records or, possibly, as predicted by mean values of relevant environmental variables). For factorial regression, nominal yields of genotypes can be estimated as a function of site mean values of the significant covariates. For models with one or two covariates, graphical expressions such as Figure 20.3(B) or Figure 20.6 can greatly simplify the scaling up (reading the best-yielding material as a function of the covariate value(s) on the site).

Extending results is more complex for AMMI models. If no environmental data is available or no relationship is found between environmental variation and ordination on significant G×L interaction PC axes of sites, subregion definition relies on the supposedly close relationship between geographical proximity and similarity for top-yielding material of the sites (attributing new sites to the subregion to which the nearest test site belongs). If a test site is representative of a given area, this area can be attributed to the subregion including the test site. Correlations of environmental variables with PC axes of sites (possibly assessed for a subset of test sites) can help characterize subregions and assign new sites to the most similar subregion according to the mean level of these variables on the site. Taking a step further, an equation for estimating the scaled PC scores of sites as a function of environmental variables recorded in test years may be searched for by a stepwise multiple regression analysis for each significant PC axis. The selected equation(s), if able to explain a fairly large portion of variation (e.g. $R^2 \geq 60$ percent), can be exploited to predict the site PC score and, thereby, the expected nominal yields of genotypes on the site (Annicchiarico, 2002a: 57).

The opportunity to interface GIS data allows for a very fine-tuned geographical definition of subregions for scaling up factorial regression or AMMI results (Annicchiarico, Bellah and Chiari, 2006). Alternatively, a simple 'Decision-aid System' could be developed that outputs the expected nominal yields of the cultivars (possibly together with an average value of Dunnett's one-tailed critical difference) as a function of the inputted site mean value of relevant environmental variables.

Scaling up may introduce a bias due to neglecting some important environmental variable or inaccurately estimating its effects. This bias was assessed for durum wheat variety targeting in Algeria, where genotype responses to test sites or new sites were predicted as a function of long-

term values in a GIS of winter mean temperature and rainfall over the season, which were selected as covariates for both factorial regression and in a multiple regression for predicting the site PC 1 score (Annicchiarico, Bellah and Chiari, 2006). Data for test sites in a validation data set revealed that GIS-based recommendations implied just a slight yield decrease relative to those based on conventional modelling (Table 20.1), while greatly enlarging the scope for site-specific targeting.

20.2.7 Taking account of yield stability

Static and dynamic stability concepts were introduced in Section 20.1. A few measures of static stability will be considered herein, reflecting the following advantages of static measures over dynamic ones: (i) somewhat higher repeatability or heritability; (ii) estimation independent from the set of tested genotypes (which allows for broader generalization); (iii) less ambiguous agronomic interpretation; and (iv) relevance for increasing food security or agricultural income (Lin, Binns and Lefkovitch, 1986; Annicchiarico, 2002a: 81). The repeatability across different sets of environments may vary, depending on the crop and the region. It increases with the temporal scale of the assessment, but remains distinctly lower than that of genotype mean yield across environments. The assessment of yield stability requires numerous test environments (eight or more) to be reliable, given its high sampling error (Kang, 1998).

The entry regression as a function of environment or year mean yield (Figure 20.1) may have various drawbacks as a stability measure (poor ability to describe G×E effects; too few years to assess stability over time). Lin and Binns' (1988) measure of average within-site temporal stability across years (or crop cycles) is the MS for year within location ($M_{y(l)}$) in an ANOVA limited to location–year cell means of the relevant genotype that also includes the location factor. High stability is indicated by low $M_{y(l)}$. However, the estimate provided by $M_{y(l)}$ is inflated by the experimental error variance. An unbiased estimate can be provided by the temporal stability variance $S_{y(l)}^2$ (Annicchiarico, 2002a: 81):

$$S_{y(l)}^2 = M_{y(l)} - (M_e / r)$$

where M_e = pooled error in the general ANOVA (performed earlier according to Figure 20.2), and r = number of experiment replications. $S_{y(l)}^2$ and $M_{y(l)}$ are equivalent for ranking genotypes, but the former is recommended for testing genotype differences and for adoption in yield reliability indexes. Well-performing genotypes can be compared for $S_{y(l)}^2$ value by ordinary tests for variance comparison such as Fisher's bilateral test (for two entries) or Hartley's test (for more entries), preferably using less critical Type 1 error rates (e.g. $P < 0.05$ or $P < 0.10$). For l test locations and y test years, DF = $l\,(y - 1)$ for each estimated $S_{y(l)}^2$ parameter.

The environmental variance (S^2) measures the static stability across environments. With reference to notations in Section 20.2.2 and for e = number of environments, its value for the genotype i is: $S_i^2 = \sum (R_{ij} - m_i)^2 / (e - 1)$. Greatest stability is $S^2 = 0$. This measure is simpler to compute than $S_{y(l)}^2$ but requires a more complex procedure for genotype comparison based on Ekbohm's test (described in Annicchiarico, 2002a: 82) or other tests.

Temporal stability, which meets farmer's perception of stability, is relevant for site-specific targeting of cultivars (path 4 in Figure 20.2) and can be used also for wide targeting (path 2 in Figure 20.2). The

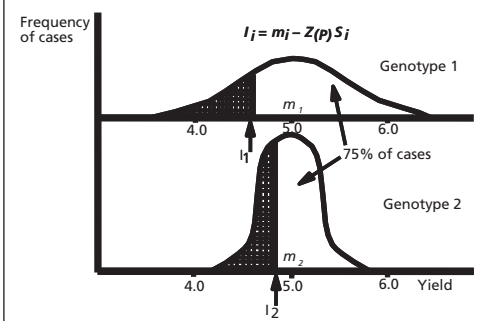
environmental variance is preferable for wide targeting when considering that G×L effects, although not directly relevant to farmers, do influence the consistency of response and the value for farmers of the cultivars across the region. If high-yielding genotypes differ for the relevant stability measure, yield stability can be accounted for by targeting genotypes according to their yield reliability instead of their yield response (Figure 20.2).

Kataoka (1963) proposed a simple index of reliability of yield (or other economic variable) that accounts for the environmental variance. This index estimates the lowest genotype yield that is expected for a probability P fixed according to the level of farmers' risk aversion. For example, P = 0.95 (lowest yield expected in 95 percent of cases) indicates high concern for disastrous events, i.e. marked risk aversion, with little consideration for mean yield response. P may vary from 0.95 (for subsistence agriculture in unfavourable regions) to 0.70 (for modern agriculture in very favourable regions) (Eskridge, 1990). For the genotype i, the index is:

$$I_i = m_i - Z_{(P)}\, S_i$$

where m_i = mean yield, and S_i = square root of the environmental variance, are parameters of the distribution of genotype yields as estimated from the sample of entry yield values; and $Z_{(P)}$ = percentile from the standard normal distribution for which the cumulative distribution function reaches the value P ($Z_{(P)}$ is 0.675 for P = 0.75, 0.840 for P = 0.80, 1.040 for P = 0.85, 1.280 for P = 0.90, and 1.645 for P = 0.95). For the two hypothetical cultivars with same mean yield and different stability reported in Figure 20.7, the advantage in yield reliability of the more stable-yielding cultivar 2 over cultivar 1 is nil when ignoring the stability characteristics ($Z_{(P)}$ = 0), sizeable for the considered level of risk aversion ($Z_{(P)}$ = 0.675), and widening for increasing levels of risk aversion. The same approach may be used for taking account of temporal stability only in the wide targeting of cultivars over the region, substituting the square root of the temporal stability variance ($S_{y(l)i}$) in place of S_i in the index formula for genotype i. Kataoka's approach has been extended to derive indexes also for measures of dynamic stability (Eskridge, 1990). For only two compared genotypes, an alternative measure of reliability of one cultivar is its estimated probability to outperform the other entry (Piepho, 1998).

FIGURE 20.7
Frequency distribution of yield values across environments of two genotypes having same mean yield (m_i) and contrasting yield stability as measured by the square root of the environmental variance (S_i), and genotype yield reliability (I_i) as lowest yield that is expected in 75 percent of cases



The site-specific yield responses of genotypes as modelled by analysis of adaptation can take account of the temporal stability of genotypes (Annicchiarico, 2002a: 85). The adaptive responses are estimates of the mean value of yield or nominal yield that is expected for each genotype on each site. This value is affected by year-to-year varia-

tion in proportion to the level of instability over time of the entry. Imposing a yield reliability assessment basically requires the estimation of a lower confidence bound for each response that depends on its variation in time and the specified level of risk aversion. The following procedure provides, for a modest level of calculation, an approximate solution that may apply to any previous formula for calculating yields or nominal yields from estimates of genotype mean yield (m_i) and interaction parameters (Sections 20.2.3 through 20.2.5). For the genotype i, it suffices to substitute ($m_i - Z_{(P)}$ $S_{y(l)i}$), where $S_{y(l)i}$ and $Z_{(P)}$ corresponds to previous notations, for m_i. For example, for the AMMI-1 model:

$$N_{ij}' = m_i - (1.28\ S_{y(l)i}) + (u_{i1}'\ v_{j1}')$$

where N_{ij}' = nominal yield reliability of the genotype i at the site j, when estimating the lowest response expected in 90 percent of cases ($Z_{(P)} = 1.28$). Compared with nominal yields (relative to mean responses, i.e. $Z_{(P)} = 0$), the AMMI-1 responses for nominal yield reliability are parallel but lowered to an extent that is modest for stable material and severe for unstable one (possibly modifying the crossover points between top-ranking entries that determine the limits of subregions). This simple approach for modelling yield reliabilities assumes that the year-to-year variation for entry yield on each site is substantially constant across the range of site mean yields (joint regression), PC scores (AMMI) or covariate values (factorial regression). This assumption may hold even when site mean yields vary widely (Annicchiarico, 2002a: 86), because unfavourable sites frequently display large temporal variation of genotype yields due to wide year-to-year extent of climatic stresses. Otherwise, a data transformation could be envisaged (Annicchiarico, 2002a p. 86).

Information on the most reliable genotypes on each site may be extended to new sites nearly as described in Section 20.2.6, the only difference concerning the introduction of the term ($- Z_{(P)}\ S_{y(l)i}$) in formulae for estimating the entry yields or nominal yields on new sites.

20.3 DEFINING A BREEDING STRATEGY

20.3.1 Overview

Global-oriented or large international breeding programmes are forced to identify some transcontinental or transnational subregions (within which to breed for wide adaptation). For breeding programmes targeting a medium-sized region, a specific-adaptation strategy may (e.g. Ceccarelli, Grando and Impigli, 1998; Annicchiarico, Bellah and Chiari, 2005; Annicchiarico, 2007) or may not (e.g. Singh *et al.*, 1992) increase the overall selection gains. When convenient, this strategy helps national breeding programmes withstand the increasing competition exerted in national markets by international seed companies (by exploiting G×L effects at a scale inaccessible for these companies). In addition, it enhances the security of food production or agricultural income in a sustainable manner, by: (i) fitting cultivars to less favoured environments, instead of altering these environments (possibly with costly or environment-unfriendly inputs) to fit widely-adapted cultivars; and (ii) safeguarding the diversity of cultivated material. Finally, it may facilitate the technological adaptation of varieties, by fixing characteristics of specific interest to subregions.

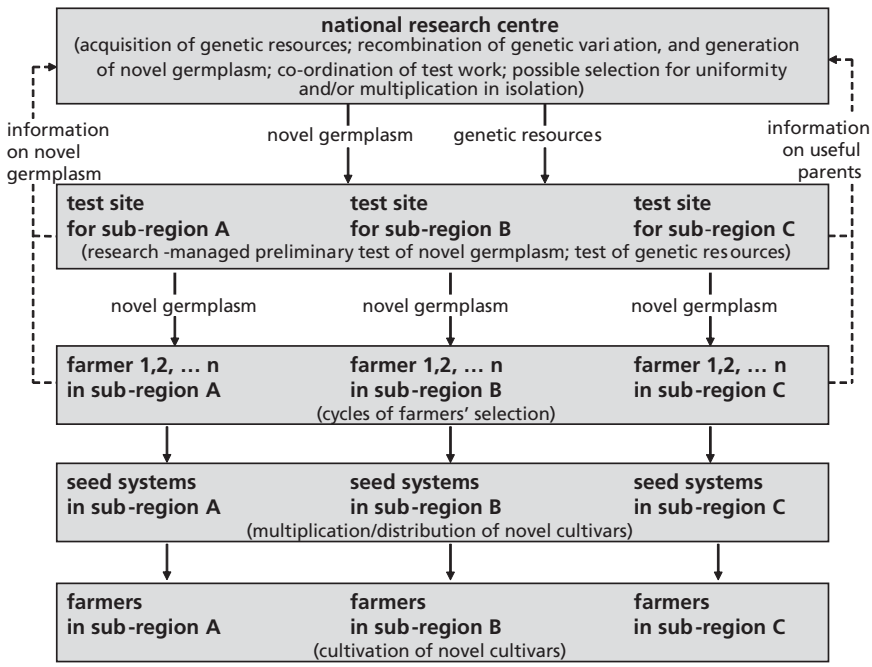
Breeding programmes of neighbouring countries that share a similar diversity of target environments could cooperate to share resources and become, on the whole,

more cost-efficient. Their cooperation might widen the scope for a specific-adaptation strategy, by enlarging the size of possible subregions (e.g. a transnational drought-prone area). International research centres could contribute to this strategy by selecting segregating material specifically adapted to distinct transnational agro-ecological zones (Ceccarelli, 1996). The size of non-repeatable G×L effects and practical considerations usually limit to two or three the number of subregions that a national programme or a few cooperating ones might possibly target.

Despite its potential interest, increased yield stability has usually been a minor breeding target. The adoption of variety types with high levels of heterozygosity and/or heterogeneity has been limited by the fewer opportunities for maximizing the yield potential that such types may offer and, sometimes, by the difficulty of meeting the uniformity level set for variety registration by many national legislations. Within a given variety type, the selection is hindered by the modest heritability and repeatability of the stability trait and the fairly high number of environments required for a reliable estimation.

When the target region is subject to large G×E interactions, effective breeding requires MET. Breeding for specific adaptation and yield stability provides an ideal framework for linking formal breeding with participatory plant breeding and overcoming the limits of each individual approach. In particular, the availability of farmers' selection environments can overcome the difficulty of formal selection programmes to adequately sample the different target environments and variety users. A national (or regional) research centre that serves the different subregions (Figure 20.8) could centralize the following tasks that participatory plant breeding could hardly assume: (i) performing sufficient crossing and hybridization work to produce largely diversified, possibly subregion-specific, novel germplasm; (ii) producing in isolation large and broadly-based composites (which is very important for outbred crops; Witcombe, 2001); and (iii) coordinating the testing of novel germplasm and large collections of genetic resources through advanced procedures that minimize the micro-environmental variation and, for novel germplasm, allow for the combined analysis of entry yields across unreplicated and possibly largely unbalanced farmers' trials (which is particularly important for inbred crops; Atlin, Cooper and Bjørnstad, 2001). This centre may also perform a preliminary screening of novel germplasm and genetic resources, while leaving the main early selection stage(s) and the main screening of genetic resources as a researcher-managed activity performed in one representative site for each subregion (the national centre might, or might not, also act as one such site). The overall selection scheme varies depending on the crop breeding system, e.g. including the farmer-managed selection among advanced lines produced by single-seed descent or among and within bulks (derived from different crosses or F_2 plants) for inbreds, and the phenotypic selection between and within composite populations for outbreds. The genotypic selection of outbreds for each subregion (e.g. based on half-sib progenies) could mostly be only research-managed. The selected varieties might enter the local seed systems (e.g. multiplied and traded by farmers-entrepreneurs), the formal seed systems, or both (e.g. after selection for uniformity prior to official registration) (Figure 20.8). In the absence of participatory breeding, specific selection is mainly performed at the site representative of a given subregion, but might be complemented

FIGURE 20.8
A specific-adaptation strategy implying the generation of novel germplasm in a national (or regional) research centre, the research-managed early selection of novel germplasm and evaluation of genetic resources in distinct subregions, and subsequent farmers' selection stages in each subregion



by within-subregion on-farm evaluation of candidate varieties.

Identifying optimal selection environments is a basic element of the breeding strategy. It is of paramount importance when selecting for wide adaptation in the presence of large G×E effects, where selection should be devised across environments that contrast for these effects and are jointly able to reproduce the genotype mean responses over the region. Optimal selection environments are of crucial importance in all cases for recurrent selection or genotypic selection of outbreds, which are severely constrained by the small number of possibly usable environments.

Sometimes, selection may partly be performed in managed or artificial environments instead of agricultural sites. These environments differ for one or more environmental factors strictly related to G×L (or G×E) interaction occurrence, and reproduce the factor levels that feature different subregions for specific adaptation or contrasting sites (or environments) for wide adaptation.

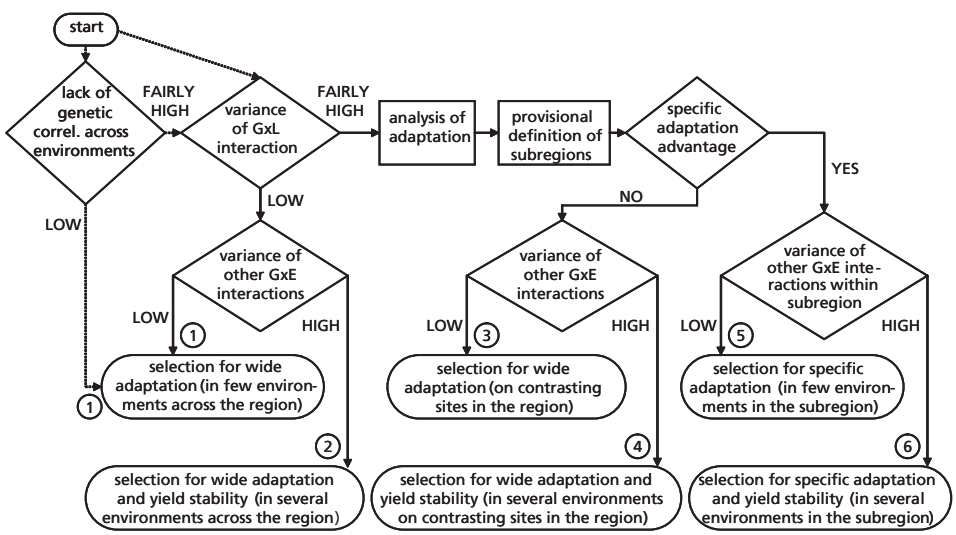
MET data that sample adequately the possible genetic base and the target environments may help breeding programmes to define a strategy in relation to G×E interactions. Its elements include the adaptation strategy, the yield stability targets, the selection environments and, possibly, the

parent material, the genetic structure, the ideotype and the single adaptive traits, and the role of participatory plant breeding. Decisions on most of these elements may conveniently be verified by further experimental work. In particular, wide- vs. specific-adaptation strategies may be compared on the basis of actual yield gains. A fair comparison of adaptation strategies implies similar costs by assuming the same total number of selection environments. Specific adaptation may then allow for higher or lower yield gains than wide adaptation, because the advantage of exploiting the portion of G×L effects that relates to subdivision of the target region (i.e. the genotype × subregion interaction) may be offset by the greater error of the estimated entry means that arises from the lower number of selection environments within each subregion (the error depends mainly on the size of G×E interaction over the region for wide adaptation and within each subregion for specific adaptation: see Section 20.3.4).

For targeting genotypes, only the G×L effects that imply a change of top-ranking genotype(s) across locations are relevant (as these effects define the subregions). For defining the adaptation strategy, all G×L effects that arise from lack of genetic correlation among sites for entry response are relevant, because the results for a given data set are extrapolated to produce information on the G×L effects that are likely to be met in future breeding for the region. A candidate subregion includes the locations that are similar in terms of genetic correlation.

The main analytical steps to provisionally define the adaptation strategy and yield stability targets from multi-location yield trials repeated over time are summarized in Figure 20.9. G×L or G×E effects that arise from heterogeneity of genotypic variance

FIGURE 20.9
Flow chart of steps for defining the adaptation strategy and yield stability targets from analysis of multi-location yield trials repeated over time



NOTES: G×L = genotype × location; G×E = genotype × environment (environment as location–year or location–crop cycle combination).

rather than lack of genetic correlation among environments are irrelevant for breeding (as they merely modify the size of the entry differences) and should be removed if they are too large. The relative size of the lack of genetic correlation and the heterogeneity of genotypic variance among environments may be estimated or, more simply, the need for transforming data could be verified (see Section 20.3.2), before estimating the variance components relative to spatial and temporal G×E interaction. An analysis of adaptation aimed to define candidate subregions (by classifying test sites on the basis of their similarity for G×L effects) may be justified if the G×L interaction variance is significant and moderately large relative to the genotypic variance, e.g. ≥30–35 percent (Atlin *et al*., 2000). Two (or possibly three) candidate subregions may be identified that are large enough to be of practical interest and lend themselves to a definition based on geography, environmental factors or farming practices. Wide- and specific-adaptation scenarios can be compared in terms of yield gains predicted from original yield data of the same data set. Wide adaptation may be preferred owing to low G×L interaction variance or to high G×L interaction variance with no clear advantage of specific breeding, with different implications for the choice of selection environments (the analysis can also help locate these environments). Adaptation strategies may also be compared according to predicted gains in other data sets (e.g. including a few sites representative of the candidate subregions), or actual gains from following selection work.

Yield stability may be justified as a target when the overall variance accounted for by the relevant G×E interaction components (either the G×Y plus G×L×Y interaction or the within-site G×Y interaction, depending on the ANOVA model) is large relative to the genotypic variance component (e.g. ≥200 percent). Decisions on the stability target are subregion-specific (depending on results for the relevant subset of sites) and can be affected by other elements (e.g. costs of additional selection environments; emphasis on food security policies).

20.3.2 Types of data and estimation of variance components

The current requirements for use of data sets are more stringent than those for targeting genotypes, given the larger inference space of the analysis. Ideally, the sample of sites should represent the different areas and cropping systems in proportion to their importance, and the germplasm sample should represent the genetic base of local interest (by including the main cultivar types and origins, or breeding lines derived from recombination from the major germplasm groups). The lack of random sampling of entries is not a limitation, because a set of elite varieties or breeding lines may represent the genetic base better than a random sample.

The G×E interaction variance components relative to the lack of genetic correlation and the heterogeneity of genotypic variance among environments can be estimated as described by Cooper, DeLacy and Basford (1996). If the latter term has larger variance than the former, it should be reduced by a suitable data transformation. This may occur when the environment mean yields vary widely (Annicchiarico, 2002a: 51). In this situation, the regression of the within-site phenotypic variance of genotype yields (averaged over test years and replications) as a function of site mean yield, with both terms expressed on a logarithmic scale, provides a quick option for verifying the need for transforming data and indicates the proper transformation (Annicchiarico,

2002a: 53). The regression slope $b \approx 2$ suggests a logarithmic transformation; $b \approx 1$ suggests a square root transformation; b not statistically significant from 0 discourages any transformation.

The reference ANOVA models are those reported in Section 20.2.2, holding genotype, location and year as random factors. The model without year factor is relevant also for estimating the variance components for genotype and G×E interaction (as required in Section 20.3.4), upon substitution of environment for the location factor. Adopting a Restricted Maximum Likelihood (REML) method allows to estimate variance components also for unbalanced data sets (Patterson, 1997).

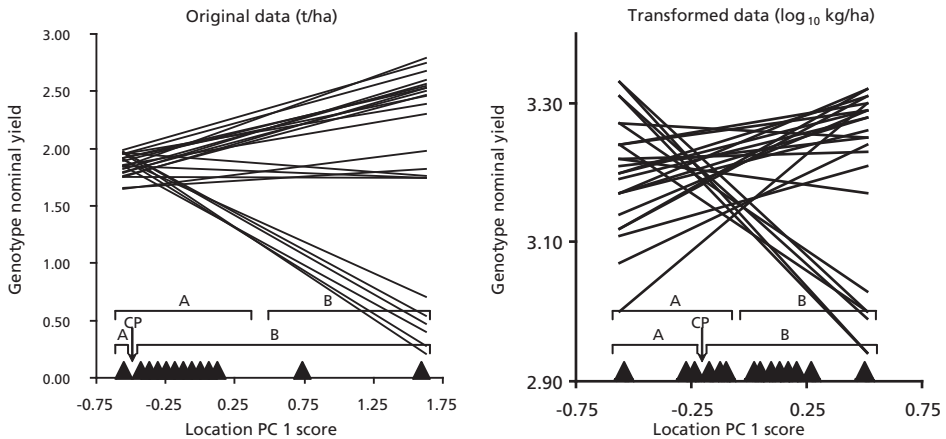
20.3.3 Classifying test sites and defining candidate subregions

Classifying test locations according to their similarity for G×L effects is needed to empirically identify candidate subregions for specific breeding, and is useful, especially in a wide-adaptation prospect, to locate selection environments. There are many possible classification methods (DeLacy *et al.*, 1996a; Annicchiarico, 2002a). Pattern analysis in its classification mode implies the hierarchical cluster analysis of sites performed on the matrix of genotype by location original yields (averaged across test years and replications) standardized within location to zero mean and unit standard deviation (thereby eliminating also the possible effect of heterogeneity of genotypic variance among sites), using a squared Euclidean distance as the dissimilarity measure and Ward's clustering method (DeLacy *et al.*, 1996a). This analysis is rapid (through freely-available software) and has the theoretical advantage of providing a grouping of sites that reflects the opportunities for exploiting indirect selection responses to locations (as these responses are maximized within potential subregions and minimized across subregions; Cooper *et al.*, 1996). It also allows for jointly analysing different data sets (e.g. relative to different test years) that have several common sites but few or no common entries, using a procedure proposed by DeLacy *et al.* (1996b).

Site classification may also follow the site ordination by modelling techniques (performed on transformed data, when appropriate). The cluster analysis of locations—using Ward's method—may be carried out on the site characteristics affecting the modelled adaptive responses, namely, the mean yield for joint regression, the significant PC scores for AMMI, and the significant covariates for factorial regression. No standardization of PC score variables is required, because the variation in site score on each PC axis is proportional to the importance of each variable. This approach has the theoretical advantage of assessing the site similarity after reducing the noise portion in the G×L interaction matrix [with the same aim, Crossa *et al.* (1991) suggested to submit the G×L matrix of AMMI-modelled yields to pattern analysis]. AMMI + cluster analysis also offers the opportunity to quickly pool the results of two or more data sets that have same test locations and different genotypes, by performing the cluster analysis on site scores of all significant PC axis issued by AMMI analyses of the different data sets (Annicchiarico *et al.*, 2006).

A truncation criterion for cluster analysis by any of the above techniques may be the lack of significant G×L interaction (at a given P level) within a group of sites (going backward from the last fusion stage to the formation of groups, and performing an ANOVA for each newly-formed group on

FIGURE 20.10
Nominal yields of durum wheat genotypes as a function of the scaled G×L interaction PC 1 score of 14 Algerian locations (black triangles), and definition of two subregions (A and B) based on cluster analysis of site PC 1 scores (above) and the estimation of the main crossover point *CP* (below), for original and log-transformed data



NOTE: Based on data by Annicchiarico, Bellah and Chiari, 2005.

data from the member sites). Other criteria are also available (DeLacy *et al.*, 1996a). In most cases, the maximum number of subregions (two or three) is fixed *a priori*.

Singh, Ceccarelli and Grando, (1999) proposed the estimation of the main crossover point in the joint regression model, i.e. the site mean yield where the interaction of crossover type between genotypes reaches the highest frequency, as a cut-off for dividing the test sites into two subregions (one high- and one low-yielding). This approach was extended to AMMI-1 and to one-covariate factorial regression models by Annicchiarico (2002a).

The cited techniques often provide different classification results for the same data. Their empirical comparison based on the assumption that a useful technique tends to maximize the selection gains for a specific-adaptation strategy suggested the superiority of pattern analysis and AMMI + cluster analysis over the other methods (based on cluster analysis or the main crossover point), while confirming the need for modelling transformed data when most G×E effects are due to heterogeneity of genotypic variance among sites (Annicchiarico, 2002b). Figure 20.10 provides some insight into reasons for these results. Responses for original yields implied larger G×L effects for higher-yielding sites (here tending to higher PC 1 score), owing to their higher genotypic variance. As a consequence, many low-yielding sites were grouped together by cluster analysis, because they tended to appear less distinct than they really were. The data transformation (selected as outlined in Section 20.3.2) removed the heterogeneity of genotypic variance between low- and high-yielding sites (while maintaining the lower genotype variance for sites with intermediate PC 1 score that is intrinsic to the adaptive responses). Cluster

analysis, unlike the main crossover criterion, subdivided the sites on the basis of their discontinuities for G×L effects as expressed by the relevant characteristic, i.e. the site PC 1 score, thereby justifying its better classification ability.

A two-subregion scenario is the first to compare with wide adaptation, as its lower interest relative to wide adaptation rules out more complex specific-adaptation scenarios. A small subregion that hardly justifies any specific breeding may be merged with the most similar large subregion (unless farmers have very specific and different preferences). An indication of the approximate proportion of the target region occupied by each subregion is useful for comparing the adaptation strategies and possibly for other reasons (e.g. estimation of seed markets). A very rough indication is provided by the proportion of test sites assigned to each subregion in the analysis of adaptation. The scaling-up of subregion definition over the region may be somewhat arbitrary, by assigning an area represented by a given test site to the subregion in which the test site was classified, or characterizing the subregions according to their mean values for environmental variables correlated with G×L interaction PC scores of sites (e.g. Annicchiarico, 1992). These variables may also be exploited for a discriminant analysis of subregions which may provide a thorough up-scaling procedure and, for the frequent case of two subregions, can easily be performed by multiple regression (Annicchiarico, 2002a: 59). The possible interfacing of this analysis with GIS allows for a detailed geographical definition of subregions (Annicchiarico, Bellah and Chiari, 2005). As an alternative, a few well-performing reference genotypes characterized by contrasting adaptation and a specific ranking order in each subregion could be used to indirectly characterize new sites on the basis of their response in small trials (Fox and Rosielle, 1982). For example, the cultivars 'La Rocca', 'Prosementi' and 'Orchesienne' in Figure 20.4(A) could be used for assigning new sites to one of three subregions identified by AMMI + cluster analysis. Additional information may allow estimating the relative size of subregions in terms of crop growing area or, especially for private companies, seed market importance.

The assessment of site similarity for G×L effects is recommended to explore the opportunities for specific breeding. It revealed, for example, that winter cold stress is more important than drought stress in determining subregions for winter cereals in Italy (Annicchiarico, 1997) and Algeria (Annicchiarico, Bellah and Chiari, 2005). Sometimes, however, candidate subregions for comparing adaptation strategies are defined *a priori* on a geographical, climatic or crop management basis (e.g. Atlin *et al*., 2000). The investigation of environment similarity for G×E effects in fairly small data sets including a few test sites representative of distinct geoclimatic areas and several environments per site (issued by different test years and, possibly, different crop managements per year) may provide useful indications for the adaptation strategy (e.g. Annicchiarico and Iannucci, 2008), as specific breeding requires that location is the main determinant of environment classification and crossover G×E interactions.

20.3.4 Comparison of wide- vs. specific-adaptation strategies

The comparison may concern predicted or actual yield gains, and vary depending on the crop breeding system and the selection procedures. It hypothesizes one selection cycle (over a definite time span) on research-managed selection sites or, possibly, managed environments, but more

selection cycles may be envisaged, especially for actual yield gains. The adaptation strategies imply similar costs by assuming the same total number of selection environments (as number of sites × number of years), assigning sites to subregions roughly in proportion to their relative size. Specific selection is based on entry mean yields over selection environments of the target subregion. Selection for wide adaptation is based on entry mean yields over all test environments hypothesized for specific breeding, in agreement 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.

The following method (Annicchiarico, 2002b) aims to compare adaptation strategies for inbred lines or clones in terms of predicted yield gains over the target region from undefined selection sites, using original yields of a data set possibly used also for defining candidate subregions. In general, the average predicted yield gain over E environments can be estimated as (DeLacy *et al.*, 1996a): $\Delta G = i\, h^2\, s_p$ (where i = standardized selection differential, h^2 = estimated broad sense heritability on a genotype mean basis, and s_p = square root of the estimated phenotypic variance across environments). In particular:

$$h^2 = s_g^2 / [s_g^2 + (s_{ge}^2 / E) + (s_e^2 / E\, R)] \quad [20.1]$$

where s_g^2, s_{ge}^2 and s_e^2 are estimates of the variance components for genotype, G×E interaction and pooled experimental error, respectively, and E and R = numbers of selection environments and experiment replications. The s_p term is equal to the square root of the denominator in equation [20.1]. In the formulae for prediction of yield gains, h^2 values are calculated from variance components estimated by a REML method for genotype and environment as random factors using all test environments (for wide adaptation) or their subsets (for specific adaptation), fixing E and R as hypothesized for selection. R and i are set to constant values in the assessment. Hereafter, E_A and E_B represent the number of selection environments for two subregions (A and B, respectively) in a specific-adaptation scenario, whereas $E_{AB} = E_A + E_B$ is the number of selection environments that are used, in a wide-adaptation scenario, for parallel selection across the subregions. P_A and P_B represent the proportion of the target region occupied by subregions A and B ($P_A + P_B = 1$), as estimated by the proportion of test sites classified in each subregion or, more precisely, after scaling up the subregion definition (for private companies, they might rather estimate the relative commercial importance of the subregions). E_A and E_B should be roughly proportional to P_A and P_B, respectively (e.g. for $E_{AB} = 6 = 3$ sites by 2 years, and $P_A = 0.64$: $E_A = 4 = 2$ sites by 2 years, and $E_B = 2 = 1$ site by 2 years). The average predicted gain (per unit area) by a wide-adaptation strategy is:

$$\Delta G_W = i\, h_{AB}^2\, s_{p(AB)} \quad [20.2]$$

where h_{AB}^2 and $s_{p(AB)}$ are obtained from equation [20.1] after estimating the components of variance for the whole set of environments in the data set, and inserting E_{AB} and R values as appropriate. The average predicted yield gain over the region provided by breeding for specific adaptation (ΔG_S) arises from a weighted mean of the gains ΔG_A and ΔG_B predicted for the subregions A and B, respectively:

$$\Delta G_A = i\, h_A^2\, s_{p(A)}\ ;\ \Delta G_B = i\, h_B^2\, s_{p(B)}\ ;$$
$$\Delta G_S = [(\Delta G_A\, P_A) + (\Delta G_B\, P_B)] / (P_A + P_B) =$$
$$(\Delta G_A\, P_A) + (\Delta G_B\, P_B) \quad [20.3]$$

where heritability and phenotypic variance values are obtained from equation [20.1] after estimating the components of variance for the subset of test environments belonging to subregion A (values h_A^2 and $s_{p(A)}$) or B (h_B^2 and $s_{p(B)}$), and inserting E_A or E_B and R in the equations as appropriate. This procedure can easily be extended to three (or more) subregions, computing the specific-adaptation gains across more than two subregions. Another procedure for the same context (i.e. predicted yield gains for inbred lines or clones from undefined selection sites) was proposed by Atlin *et al*. (2000). Piepho and Möhring (2005) expanded this approach by considering a more complex scenario that maximizes the selection gains by using for specific selection also the data from other subregions. These data are given a weight proportional to their relevance for the target subregion.

Another adaptation strategy may contemplate the selection in one subregion also for another subregion, so that the yield gains in the latter subregion derive from correlated responses in an indirect selection context. Formulae for comparing this strategy against the previous ones in terms of predicted gains (taking conveniently into account the effect of G×E interactions within subregions, unlike more simplified approaches) are reported elsewhere (Annicchiarico, 2002a: 68; 2002b).

For open-pollinated species, the comparison of strategies based on predicted gains from selection of populations (as represented by cultivars in a MET data set) is frequently out of context, as selection mainly concerns individual plants. Preliminary indications may be obtained by comparing the top-ranking cultivars according to each adaptation strategy, as if each cultivar was used as the unique genetic base and could provide, when chosen, a constant gain from intra-population selection (e.g. Annicchiarico and Piano, 2005). Predicted gains could be computed as the difference between the top-ranking cultivar and the mean of all tested entries.

Reliable predicted gains for comparing adaptation strategies in open-pollinated species may be obtained from multi-environment progeny testing of individuals as half-sib or full-sib families or as selfed progenies. It suffices to substitute the appropriate narrow sense heritability term (Wricke and Weber, 1986) for the broad sense heritability in formula [20.1] when using this equation for estimating ΔG_W and ΔG_S according to formulae [20.2] and [20.3]. For example, the following equation and the square root of its denominator allow for estimating h^2 and s_p, respectively, for half-sib progeny testing targeted to selection of parent material as clone or selfed progeny:

$$h^2 = 0.5\, s_a^2 / [0.25\, s_a^2 + (0.25\, s_{ae}^2 / E) + (s_e^2 / E\, R)] \qquad [20.4]$$

where s_a^2, s_{ae}^2 and s_e^2 are estimated variance components relative to the additive genetic variance, the interaction of additive genetic effects with environment and the pooled error, respectively, and E and R = numbers of hypothesized selection environments and experiment replications. The REML analysis performed on family plot values for the relevant sets of test environments provides the estimate of s_e^2, and allow estimation of the other variance components from the variance among families (s_g^2) and the family × environment interaction variance (s_{ge}^2) (assuming no inbreeding in the tested material) as:

$$s_a^2 = 4\, s_g^2;\ s_{ae}^2 = 4\, s_{ge}^2.$$

The previous procedures assume undefined selection sites. Thus, they relate to the average screening ability of sites within each subregion (as estimated from the sample of test sites). It is also possible to compare the adaptation strategies for predicted yield gains from selection in managed environments or in previously-defined, nearly-optimal selection sites (Annicchiarico, Bellah and Chiari, 2005). Predicted gains are correlated gains from the defined selection environments to the target environments. For inbred lines or clones, they are (DeLacy *et al*., 1996a):

$$\Delta G_{T/S} = i\, r_{(S,T)}\, s_{p(T)} \qquad [20.5]$$

where $r_{(S,T)}$ = phenotypic correlation for entry mean yield between selection and target environments, and $s_{p(T)}$ = phenotypic standard deviation in the target environments. The basic difference with previous formulae for predicted gains of inbreds or clones is the substitution of the appropriate $r_{(S,T)}$ term in place of h^2. For agricultural sites, however, the gains relate to estimates of site screening ability that may largely be affected by specific conditions during the test years.

A promising specific-adaptation strategy may be compared with wide adaptation in terms of actual yield gains. Selection is performed both independently within subregion (specific adaptation) and jointly across subregions (wide adaptation) at previously-defined selection sites (or managed environments), contemplating the same total number of selection environments (roughly assigned in proportion to the relative size of subregions). The selected groups are compared across a sample of environments, assessing the gains provided by each strategy in each subregion (e.g. in terms of difference between the group mean and the mean of a set of high-yielding cultivars) and weighting them on the relative importance of the subregions. For example, Ceccarelli, Grando and Impiglia (1998) selected barley genotypes within a large germplasm pool for wide and for specific adaptation to an unfavourable (A) and a favourable (B) subregion, reporting the mean values of the selected groups tested in the two subregions. For one set of material (Cohort 89), the yield gains estimated with respect to a set of high-yielding control cultivars were: ΔG_A = 0.03 t/ha, and ΔG_B = 0.08 t/ha, for specific adaptation; ΔG_A = –0.03 t/ha, and ΔG_B = 0.08 t/ha, for wide adaptation. Based on these values, the advantage of the former strategy is manifest. If the subregions were of equal size ($P_A = P_B$ = 0.50), the gain over the region from specific (ΔG_S) and from wide (ΔG_W) breeding could be estimated as:

$\Delta G_S = (\Delta G_A\, P_A) + (\Delta G_B\, P_B) = (0.03 \times 0.50) + (0.08 \times 0.50) = 0.055$ t/ha per cycle
$\Delta G_W = (\Delta G_A\, P_A) + (\Delta G_B\, P_B) = (-0.03 \times 0.50) + (0.08 \times 0.50) = 0.025$ t/ha per cycle

implying a greater efficiency of specific breeding equal to $\Delta G_S/\Delta G_W = 0.055/0.025 = 220$ percent.

Large data sets for inbred lines or clones that include many entries and several test years may also be used for assessing actual yield gains of adaptation strategies, after defining candidate subregions and selection sites. For example, Annicchiarico, Bellah and Chiari (2005) used two years and a total of three sites of a three-year durum wheat data set for wide or subregion-specific entry selection, and the remaining environments to assess actual gains in each subregion (estimated as yield difference between the mean of selected entries and the mean of all entries), averaging results across the three possible pairs of selection years (Table 20.3).

TABLE 20.3
Mean yield of selected durum wheat entries, and average observed and predicted yield gains per selection cycle in two subregions and over the Algerian durum wheat cropping region, for specific- and wide-adaptation strategies

	Specific adaptation	Wide adaptation	Specific/wide ratio (%)
Mean yield (t/ha) [a b]			
Subregion A	1.899	1.833	103.6
Subregion B	3.031	3.031	100.0
Observed gain (t/ha) [a b c]			
Subregion A	0.233	0.167	139.5
Subregion B	0.372	0.372	100.0
Region	0.327	0.305	107.1
Predicted gain (t/ha) [a]			
Subregion A [d]	0.199	0.181	109.9
Subregion B [d]	0.509	0.505	100.8
Region [d]	0.409	0.400	102.2
Region [e]	0.316	0.304	104.1

[a] 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).
[b] Values averaged across three pairs of test years.
[c] Gain computed as the difference between the mean of selected entries and the mean of all entries.
[d] For defined selection locations (using equation [20.5]). Values averaged across three pairs of test years.
[e] For undefined selection locations (using equation [20.2] for wide adaptation, and equation [20.3] for specific adaptation; in the latter, ΔGA = 0.161 t/ha and ΔGB = 0.304 t/ha).
Source: Annicchiarico, Bellah and Chiari, 2005.

Gains over the region for each strategy were weighted means of those in each subregion (as shown for the barley data) and indicated the greater efficiency of specific over wide breeding ($\Delta G_S/\Delta G_W$ = 107.1 percent) as a consequence of greater gain (+39.5 percent) in the stressful subregion A. This procedure is less reliable than comparisons for actual gains on a larger and independent genotype sample. However, its results agreed closely with those of predicted gains for undefined selection sites ($\Delta G_S/\Delta G_W$ = 104.1 percent), while being less consistent with those of predicted gains for the same set of selection sites ($\Delta G_S/\Delta G_W$ = 102.2 percent) (Table 20.3).

Especially for cross-pollinated crops, the lack of sufficiently large data sets may limit the comparisons of adaptation strategies based on predicted gains, giving impulse to those based on actual gains. One example was given by Annicchiarico (2007) for phenotypic selection of alfalfa in artificial environments capable of reproducing the adaptive responses occurring across the three subregions shown in Figure 20.4(A). Direct selection for specific adaptation targeted each of the contrasting subregions A and C, exploiting correlated selection gains for the intermediate subregion B. To reduce the evaluation costs, the selection environments also acted as test environments for the selections (possibly introducing some bias relative to agricultural sites).

All cited procedures hypothesize growing the novel germplasm in all selection environments. Their indication of some advantage for specific breeding probably implies larger gains after optimizing other elements of the breeding strategy by considering, at least to some extent: (i) the allocation of novel germplasm to only one

subregion on the basis of crucial adaptive traits (or molecular markers) assessed preliminarily at the main research centre; or (ii) the use of a distinct genetic base for each subregion. Indications for these elements may be provided by the MET data set, with or without further research. Anyway, the comparison of adaptation strategies could not take account of some positive effects of breeding for specific adaptation (Section 20.3.1) that are difficult to quantify.

20.3.5 Definition and use of selection environments

In the presence of sizable G×L interaction, the main research centre may host a preliminary selection stage if its screening ability for the target region is high. According to formula [20.5] in Section 20.3.4, the screening ability of a site (or a managed environment) is proportional to the phenotypic correlation between entry yields on the site and entry mean yields over the target environments. The phenotypic correlation takes account of the genetic correlation between selection and target environments and the broad sense heritability on the site (Cooper, DeLacy and Basford, 1996). When breeding for specific adaptation, this preliminary selection stage may also allow for the allocation of material to a specific subregion on the basis of adaptive traits.

Multi-environment data can also help locate optimal selection sites for research-managed selection (also usable for detecting parent germplasm of specific interest for subregions; see Figure 20.8). Preliminary indications may be obtained from site ordination in the analysis of adaptation or site classification for G×L effects. The optimal selection site for a given subregion has the highest screening ability for the relevant target environments. When adopting more sites within a subregion, it is the joint screening ability of the sites (as indicated by phenotypic correlations between selection and target environments for entry yields) that should be maximized (e.g. Annicchiarico, Bellah and Chiari, 2005).

Selection for wide adaptation in the presence of sizable G×L interaction should be performed across sites that contrast for G×L effects (as hypothesized for comparing adaptation strategies) and are jointly capable of maximizing the screening ability (as indicated by phenotypic correlations), rather than across sites that maximize individually the screening ability (which are implicitly similar for G×L effects). Contrasting sites offer the opportunity for disclosing and selecting material capable of assembling different adaptive traits of interest for the region (Calhoun *et al*., 1994). Thus, optimal selection sites may be identified for wide or specific adaptation by the same procedures. Phenotypic correlations between selection and target environments also allow the assessing of the lower yield gain expected from adopting suboptimal sites.

Managed or artificial selection environments that reproduce the genotype adaptive responses and do not imply very high implementation costs can partly replace agricultural selection sites to reduce costs (especially when optimal sites belong to remote areas or have little infrastructure) or to increase the selection gains (especially when agricultural sites are subject to wide G×Y interaction due to unpredictable climatic conditions). For example, the artificial environments in Figure 20.4(B) were established on the ground of the positive correlations of soil clay content and drought stress level with PC 1 score of alfalfa test sites in Figure 20.4(A) (Annicchiarico, 1992). They could reproduce the adaptive responses occurring in three subregions (as shown by three reference varieties: Figure 20.4),

and may be used to select for wide or specific adaptation (Annicchiarico, 2007). Selection under the natural conditions of the breeding centre (located in subregion A), implying sandy-loam soil and negligible stress, would produce varieties specifically adapted to these conditions, such as cultivar 'Lodi' in Figure 20.4(B).

Managed environments may also be used to breed for wide adaptation to regions featured by large within-site G×Y interaction and small repeatable G×L effects due to wide year-to-year climatic variation. In such regions, agricultural sites in individual years frequently misrepresent the target environments, leading to low selection gains (Cooper, DeLacy and Basford, 1996). An optimal set of managed environments can be identified by assessing the joint screening ability of these environments (Cooper *et al*., 1995, 1997). Federer and Scully (1993) proposed statistical designs to select material for wide adaptation across a factorial combination of two or three crop management or physical factors that reproduce the variation for environmental variables associated with G×E effects.

Selecting on agricultural sites for wide adaptation to climatically unpredictable regions may increase its efficiency by a procedure proposed by Podlich, Cooper and Basford (1999). A large sample of target environments is classified on the basis of G×E interaction effects, identifying a few major groups whose relative frequency is estimated and that are characterized either by a specific response of some probe genotypes or a definite value of some crucial climatic variable(s). Each new selection environment is classified according to the response of the probe genotypes (grown along with the tested material) or the relevant climatic variable(s), and is given a weight on the future MET-based entry selection that is proportional to the frequency of its group.

The optimal number of selection sites, years and experiment replications for inbred lines of clones in a region or subregion may be investigated after estimating the genotypic and genotype-environmental variance components for a representative sample of elite material and target environments. The aim is maximizing, for about same costs, the yield gain: $\Delta G = i\, h^2\, s_p$, where h^2 and s_p are computed by the following equation and the square root of its denominator, respectively, for L locations, Y years and R replication numbers hypothesized for selection (Cooper *et al*., 1999):

$$h^2 = s_g^2 / (s_g^2 + s_{gl}^2 / L + s_{gy}^2 / Y + s_{gly}^2 / LY + s_e^2 / RLY)$$

Estimates may also relate to more complex scenarios, e.g. two-stage selection (Grüneberg *et al*., 2004) or among-cross (bulk) plus within-cross pure line selection (Cooper *et al*., 1999). Research-managed selection trials would usually include at least two replications, while the number of selection years is kept low (often no more than two) so as not to delay the release of varieties. Thus, decisions mainly regard the number of selection sites. Selecting also for yield stability may lead to increases in this number (Figure 20.9), if socio-economically convenient.

Predicted gains for different scenarios relative to managed environments depend on the h^2 value over selection environments (as affected by hypothesized L, Y and R values) and the genetic correlation between selection and target environments (DeLacy *et al*., 1996a; Qiao *et al*., 2004).

In various contexts, only an unbalanced data set (with possibly many missing genotype-environment cell means) may be

available for entry selection. Best Linear Unbiased Prediction (BLUP) entry means as estimated by a REML method should be used in this case. BLUP entry main effects (P_i) differ from Best Linear Unbiased Estimate (BLUE) entry main effects (G_i, as provided by least squares means) because they are shrunk to a greater extent for entries with less observations, to take account of the greater uncertainty introduced by less MET. In particular, P_i effects are shrunk in proportion to the difference to unity of the broad sense heritability on an entry mean basis (DeLacy *et al.*, 1996a), i.e. $P_i = (h^2\ G_i)$, where $h^2 = s_g^2 / [s_g^2 + (s_{ge}^2 / e_i) + (s_e^2 / e_i\ r_i)]$, and e_i and r_i are numbers of environments and experiment replications for the entry i (the latter as harmonic mean). The variance components are constant values estimated by a REML method for the entire data set. This procedure is simpler and more reliable than other methods that also consider the broad sense heritability in single trials (DeLacy *et al.*, 1996a). The BLUP means can be obtained by a REML analysis performed on the genotype-environment cell means, holding genotype and G×E interaction as random effects (while environment is fixed) (Hill and Rosenberger, 1985). Specific models may be applied to unreplicated trials or more complex data structures, e.g. sites and farms (or years) within site (Smith, Cullis and Thomson, 2001; Coe, 2002). BLUP entry means also provide more realistic predictions of yield gains from actually selected entries than do BLUE means.

Breeding for wide or specific adaptation can account for yield stability by selecting entries according to Kataoka's index of reliability (see Section 20.2.7) instead of mean yield over selection environments. This measure is justified by the fact that all G×E effects (including G×L ones) influence the consistency of response and the value for farmers of a variety across the target region or subregion.

20.3.6 Identification of genetic resources, adaptive traits and useful markers

The analysis of adaptation can also produce information on germplasm which, within a given adaptation strategy, is of special interest as parent for crosses or as population for recurrent selection in view of its adaptive response. In general, evidence points to a moderate heritability of adaptation parameters (Becker and Léon, 1988). For example, crosses for wide adaptation could be envisaged between genotypes: (i) possessing high mean yield and the desired adaptive response (as indicated by $b \approx 1$ in joint regression, β value near zero in factorial regression, and genotype PC score near zero in AMMI analysis); or (ii) between pairs of genotypes possessing high mean yield and specific adaptation of contrasting type, such as 'Orchesienne' and 'La Rocca' in Figure 20.4(A). Genotypes may also be classified for adaptive response by pattern analysis (Cooper, DeLacy and Basford, 1996). The analysis of adaptation of genetic resources with contrasting origin may highlight the relationship of the environment of origin to the adaptive response as determined by evolutionary adaptation. Finally, the comparison of different variety types may contribute to decisions of breeding programmes on the genetic structure of novel germplasm (Brancourt-Hulmel, Biarnès-Dumoulin and Denis, 1997).

The analysis of adaptation may also provide preliminary indications on traits contributing to wide or specific adaptation, by correlations of the estimated genotype adaptation parameters with morphophysiological traits (possibly recorded in a subset of test

sites) (Annicchiarico, 2002a). Additional correlations for distinct subregions may contribute to highlight the adaptive traits of local interest. For qualitative traits, or quantitative traits largely definable by just a few contrasting levels (e.g. tall vs. semi-dwarf), the relationship of the trait with adaptive responses can be: (i) inferred visually, by indicating also the plant type in genotype ordination diagrams; or (ii) estimated, by averaging the adaptive responses across genotypes belonging to the same plant type.

The identification of adaptive traits by the above procedures usually needs be confirmed by further results relative to a large genotype sample tested in a few sites (or managed environments) representing different subregions or contrasting environments. The assessment may include a wider set of traits; assess also curvilinear relationships of yield with trait levels; and thoroughly assess the value of single traits or sets of traits as indirect selection criteria. Recent crop simulation models that incorporate gene action may contribute to define adaptive traits by predicting the impact of single traits or trait combinations on genotype adaptive responses to different subregions or contrasting environments (Chapman *et al*., 2002).

Genotype adaptive responses and QTLs for yield may be studied concurrently, to locate molecular markers that could assist the selection for wide or specific adaptation and help locate parent combinations with complementary useful characteristics. QTLs can be accommodated in a factorial regression model by including, in place of the genotype factor, a set of genotypic covariates accounting for the different QTLs (Vargas *et al*., 2006). Also AMMI analysis can be used to map QTLs associated with wide or specific adaptation (Romagosa *et al*., 1996). The information on QTLs and useful markers may derive from experiments performed on just a few sites (or in managed environments) that represent contrasting subregions or environments (e.g. Ribaut *et al*., 2007). Such environments may also be used to compare selection strategies that exploit markers of wide or subregion-specific interest in terms of actual yield gains.

20.4 COPING WITH MICRO-ENVIRONMENTAL VARIATION

Genotypes vary across experiment replications, owing to micro-environmental variation for soil fertility, soil depth or other factors. The portion of this variation that is not controlled by the experimental design produces a special type of G×E interaction (e.g. the genotype × block interaction in a RCB design) that represents the experimental error in the data analysis. Adopting efficient experimental designs and convenient blocking of treatments, and exploring and correcting for the within-block spatial variability, are important tools to minimize this error and increase the accuracy of entry comparisons and, in selection trials, the selection gain. This is particularly important in a participatory programme where most of the trials are conducted in farmers' fields.

The availability of suitable software allows for the ordinary use of efficient experimental designs. For early selection stages, the possibly large number and the limited seed amount of the tested entries may lead to adopting an augmented row-column design with unreplicated entries and replicated controls, such as those proposed by Federer, Nair and Raghavarao (1975) or by Lin and Poushinsky (1983). The latter design, in which most control entries are allocated systematically (at the centre of

subplots, surrounded by eight test entries) while the others are randomly placed to estimate the experimental error, allows for an accurate and flexible adjustment of test entry values as a function of: (i) the row and column effects; or (ii) the covariate represented by the value of systematic controls (which estimates the local yield potential). As a development of Federer, Nair and Raghavarao's (1975) design, a partially replicated design implies two replications for a subset of test entries to improve the estimation of the experimental error (see also Section 20.3.1).

Resolvable incomplete block designs, in which the incomplete blocks can be grouped to form a complete replication of the entries, are preferable to non-resolvable ones because this double-blocking structure allows for better error control and for the possible analysis as a RCB for missing data (Basford *et al*., 1996). To partly overcome the constraints of lattice designs for numbers of tested entries and of plots within incomplete block, Patterson, Williams and Hunter (1978) devised a new class of resolvable designs termed alpha lattices in which, given k plots per incomplete block, the number of entries g may be whatever multiple of k. The efficiency of these designs increases when k approaches the square root of g.

The layout of farmer-managed selection trials depends on the total number of test entries and the plots available per farm. Within these constraints, it should preferably allow for estimating an experimental error on the farm, e.g. by unreplicated trials including a few replicated entries, or on the site and its surroundings, by assigning a complete block to each of few farms or an incomplete block to each of several farms. In less favourable instances or when farm size is small (see also Chapter 9 in this volume), each farm within a subregion may host an incomplete block.

Experimental designs cannot control the effects of environmental variation within complete or incomplete blocks. These effects may be large in stress-prone sites and in trials with many entries. For plots arranged in a rectangular row-column array, the ANOVA residual E_{mn} of the plot located in row m and column n may be modelled spatially as a function of: (i) uni- or bi-directional fertility gradients estimated by polynomial functions of row (R_m) and/or column (C_n) numbers, e.g. a second-degree polynomial response surface; (ii) the covariate X_{mn} represented by the mean value of the ANOVA residuals for the neighbouring plots (which estimates the local yield potential), through a nearest-neighbour (or Papadakis') method: $E_{mn} = b\ X_{mn} + e_{mn}$ (where e_{mn} is the non-modelled residual) (Brownie, Bowman and Burton, 1993). As a third, more complex approach, Cullis and Gleeson (1991) proposed modelling the residuals as a function of the spatial autocorrelation between pairs of neighbouring plots assessed separately along rows and columns. If there is spatial pattern, the variance between residuals of neighbouring plots will be lower than that between residuals of plots far apart, leading to autocorrelation (estimated by a moving average). This two-dimensional separable autoregressive spatial model of first order (AR1 × AR1) can describe many spatial patterns and has frequently proved adequate. Therefore, Gilmour, Cullis and Verbyla (1997) proposed generally fitting this model and then displaying by various diagnostic tools (e.g. the variogram of non-modelled residuals along rows and columns) the possible presence of remaining pattern, to be modelled by additional covariates (e.g. a cyclic row or column effect due

to sowing or harvesting operations) in a REML analysis. The best model is usually selected so as to minimize the Akaike Information Criterion. Trial-specific spatial parameters and other parameters relative to trial design, genotype, environment and G×E effects may all be included into a comprehensive model for estimating BLUP entry means (Smith, Cullis and Thomson, 2001; Smith *et al*., 2002).

Spatial analysis is a trial-specific, iterative exercise that may be relatively time-consuming, especially for complex models or large sets of selection trials. It has tended to reduce the experimental error to a large degree in RCB and to a sizeable degree in lattice designs in extensive studies (Smith *et al*., 2002; Singh *et al*., 2003). However, an assessment of this tool and the two designs in terms of predicted yield gain for wheat selection trials in Australia indicated the sizeable advantage of spatial analysis over the ordinary lattice design analysis in the range of one to five selection trials (test environments), as well as the sizeable advantage of lattice over RCB up to about seven trials (Qiao *et al*., 2004). On the whole, this study highlighted the basic importance of using a lattice design, suggesting modelling spatial variability only in small sets of trials or in trials that exhibit unusually high experimental error.

20.5 COMPUTER SOFTWARE

Nearly all the reported analytical techniques are difficult to apply by breeders without suitable and relatively user-friendly software. IRRISTAT, renamed as CROPSTAT from its sixth version onwards, has been developed by the International Rice Research Institute (IRRI, 2007). It is freely available and has specific modules for plant breeding designs (alpha lattices, augmented row-column design, etc.) and ANOVA, joint regression, AMMI and pattern analysis (outputting most of the relevant graphics). This and other software usually focus on the analysis of G×E effects, and may require some modification to the recommended procedures when analysing G×L effects in trials repeated over time. The use of CROPSTAT in this context has been described by Annicchiarico (2002a) for the main analyses considered in this chapter, including factorial regression analysis (for which no specific module is available). Some recently-implemented tools, such as a REML analysis module usable for estimation of variance components and spatial analysis and a module for generic cluster analysis, have widened or simplified the assistance to data analysis offered by this software.

MATMODEL is software for AMMI and joint regression modelling, which is also available in a free version (Gauch, 2007). It is particularly useful for handling missing data, and its output includes the AMMI-based definition of subregions for genotype targeting. GEBEI is free software for pattern analysis, whose utilities have largely been included into CROPSTAT. It has the unique feature of allowing classification of sites in largely unbalanced data sets according to DeLacy *et al.* (1996b) (contact: Professor K.E. Basford, University of Queensland).

Useful commercial software includes: (i) AGROBASE, allowing one to perform ANOVA, joint regression and AMMI modelling, analysis of unreplicated and replicated designs, and spatial analysis; (ii) INTERA (Decoux and Denis, 1991), with modules for factorial and joint regression and AMMI modelling, ANOVA, and classification of sites and genotypes; (iii) GGE BIPLOT (Yan and Kang, 2003), useful for joint regression and

AMMI modelling, ANOVA, GGE biplot and pattern analysis; and (iv) ASREML (Gilmour *et al*., 2006), powerful for REML and spatial analyses.

GENSTAT, compared with other powerful generic software (e.g. SAS; S-PLUS), includes specific procedures for many analyses (e.g. joint and factorial regression; AMMI; REML; spatial trend) and has a policy of free licensing to institutions of less developed countries. Sets of SAS instructions for several aspects of G×E interaction, REML and spatial analyses are reported in Kang (2003). SAS instructions for specific analyses can also be found in many individual papers.

REFERENCES

Aastveit, A.H. & Martens, H. 1986. ANOVA interactions interpreted by partial least squares regression. *Biometrics*, 42: 829–844.

Almekinders, C.J.M., Lowaars, N.P. & de Bruijn, G.H. 1994. Local seed systems and their importance for an improved seed supply in developing countries. *Euphytica*, 78: 207–216.

Annicchiarico, P. 1992. Cultivar adaptation and recommendation from alfalfa trials in northern Italy. *Journal of Genetics and Breeding*, 46: 269–278.

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

Annicchiarico, P. 2002a. Genotype × environment interactions: challenges and opportunities for plant breeding and cultivar recommendations. *FAO Plant Production and Protection Paper*, No. 174 (available at: http://www.fao.org/docrep/005/Y4391E/y4391e00.HTM; accessed 18 Jan. 2009).

Annicchiarico, P. 2002b. Defining adaptation strategies and yield stability targets in breeding programmes. *In* M.S. Kang (ed). *Quantitative genetics, genomics and plant breeding*, pp. 365–383. Wallingford, UK, CABI.

Annicchiarico, P. 2007. Wide- versus specific-adaptation strategy for lucerne breeding in northern Italy. *Theoretical and Applied Genetics*, 114: 647–657.

Annicchiarico, P. & Iannucci, A. 2008. Breeding strategy for faba bean in southern Europe based on cultivar responses across climatically-contrasting environments. *Crop Science*, 48: 983–991.

Annicchiarico, P. & Piano, E. 2005. Use of artificial environments to reproduce and exploit genotype × location interaction for lucerne in northern Italy. *Theoretical and Applied Genetics*, 110: 219–227.

Annicchiarico, P., Bellah, F. & Chiari, T. 2005. Defining subregions and estimating benefits for a specific-adaptation strategy by breeding programs: a case study. *Crop Science*, 45: 1741–1749.

Annicchiarico, P., Bellah, F. & Chiari, T. 2006. Repeatable genotype × location interaction and its exploitation by conventional and GIS-based cultivar recommendation for durum wheat in Algeria. *European Journal of Agronomy*, 24: 70–81.

Annicchiarico, P., Chiari, T., Bellah, F., Doucene, S., Yallaoui-Yaïci, N., Bazzani, F., Abdellaoui, Z., Belloula, B., Bouazza, L., Bouremel, L., Hamou, M., Hazmoune, T., Kelkouli, M., Ould-Said, H. & Zerargui, H. 2002. Response of durum wheat cultivars to Algerian environments. I. Yield. *Journal of Agricultural Environment for International Development*, 96: 157–187.

Annicchiarico, P., Russi, L., Piano, E. & Veronesi, F. 2006. Cultivar adaptation across Italian locations in four turfgrass species. *Crop Science*, 46: 264–272.

Atlin, G.N., Cooper, M. & Bjørnstad, 2001. A comparison of formal and participatory breeding approaches using selection theory. *Euphytica*, 122: 463–475.

Atlin, G.N., Baker, R.J., McRae, K.B. & Lu X. 2000. Selection response in subdivided target regions. *Crop Science*, 40: 7–13.

Bänzinger, M., Bertrán, F.J. & Lafitte, H.R. 1997. Efficiency of high-nitrogen selection environments for improving maize for low-nitrogen target environments. *Crop Science*, 37: 1103–1109.

Barah, B.C., Binswanger, H.P., Rana, B.S. & Rao, G.P. 1981. The use of risk aversion in plant breeding; concept and application. *Euphytica*, 30: 451–458.

Basford, K.E. & Tukey J.W. 2000. *Graphical Analysis of Multiresponse Data: Illustrated with a Plant Breeding Trial*. Boca Raton, USA, Chapman & Hall/CRC Press.

Basford, K.E., Williams, E.R., Cullis, B.R. & Gilmour, A. 1996. Experimental design and analysis of variety trials. pp. 125–138, *in* Cooper & Hammer, 1996, q.v.

Becker, H.C. 1984. Analysis of genotype × environment interaction with partitioning of environmental effects into effects of locations and years. *Vortrage fur Pflanzenzüchtung*, 7: 209–218.

Becker, H.C. & Léon, J. 1988. Stability analysis in plant breeding. *Plant Breeding*, 101: 1–23.

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

Brownie, C., Bowman, D.T. & Burton, J.W. 1993. Estimating spatial variation in analysis of data from yield trials: a comparison of methods. *Agronomy Journal*, 85: 1244–1253.

Calhoun, D.S., Gebeyehu, G., Miranda, A., Rajaram, S. & van Ginkel, M. 1994. Choosing evaluation environments to increase wheat grain yield under drought conditions. *Crop Science*, 34: 673–678.

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

Ceccarelli, S. 1994. Specific adaptation and breeding for marginal conditions. *Euphytica*, 77: 205–219.

Ceccarelli, S. 1996. Positive interpretation of genotype by environment interaction in relation to sustainability and biodiversity. pp. 467–486, *in* Cooper & Hammer, 1996, q.v.

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

Chapman, S.C., Hammer, G.L., Podlich, D.W. & Cooper, M. 2002. Linking bio-physical and genetic models to integrate physiology, molecular biology and plant breeding. *In* M.S. Kang, ed. *Quantitative genetics, genomics and plant breeding*, pp. 167–187. Wallingford, UK, CABI.

Clawson, D.L. 1985. Harvest security and intraspecific diversity in traditional tropical agriculture. *Economic Botany*, 39: 56–67.

Coe, R. 2002. Analyzing data from participatory on-farm trials. pp. 23–34, *in* M.R. Bellon & J. Reeves, eds. *Quantitative analysis of data from participatory methods in plant breeding.* Mexico, D.F., Mexico: CIMMYT.

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

Cooper, M. & Byth, D.E. 1996. Understanding plant adaptation to achieve systematic applied crop improvement – A fundamental challenge. pp. 5–23, *in* Cooper & Hammer, 1996, q.v.

Cooper, M. & Hammer, G.L. eds. 1996. *Plant adaptation and crop improvement*. Wallingford, UK, CABI.

Cooper, M., DeLacy, I.H. & Basford, K.E. 1996. Relationships among analytical methods used to study genotypic adaptation in multi-environment trials. pp. 193–224, *in* Cooper & Hammer, 1996, q.v.

Cooper, M., Woodruff, D.R., Eisemann, R.L., Brennan, P.S. & DeLacy, I.H. 1995. A selection strategy to accommodate genotype-by-environment interaction for grain yield of wheat: managed-environments for selection among genotypes. *Theoretical and Applied Genetics*, 90: 492–502.

Cooper, M., Stucker, R.E., DeLacy, I.H. & Harch, B.D. 1997. Wheat breeding nurseries, target environments, and indirect selection for grain yield. *Crop Science*, 37: 1168–1176.

Cooper, M., Rajatasereekul, S., Somrith, B., Sriwisut S., Immark, S., Boonwite, C., Suwanwongse, A., Ruangsook, S., Hanviriyapant, P., Romyen, P., Pornuraisanit, P., Skulkhu, E., Fukai, S., Basnayake, J. & Podlich, D.W. 1999. Rainfed lowland rice breeding strategies for northeast Thailand. II. Comparison of intrastation and interstation selection. *Field Crops Research*, 64: 153–176.

Cornelius, P.L. & Crossa, J. 1999. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment trials. *Crop Science*, 39: 998–1009.

Crossa, J., Fox, P.N., Pfeiffer, W.H., Rajaram, S. & Gauch, H.G. 1991. AMMI adjustment for statistical analysis of an international wheat yield trial. *Theoretical and Applied Genetics*, 81: 27–37.

Cullis, B.R. & Gleeson, A.C. 1991. Spatial analysis of field experiments – an extension to two directions. *Biometrics*, 47: 1449–1460.

Dagnelie, P. 1975. *Théorie et méthodes statistiques. Vol. 2.* 2nd ed. Gembloux, Belgium, Les Presses Agronomiques.

Decoux, G. & Denis, J.B. 1991. INTERA. Logiciels pour l'interpretation statistique de l'interaction entre deux facteurs. Versailles, France: Biométrie INRA.

DeLacy, I.H., Basford, K.E., Cooper, M., Bull, J.K. & McLaren, C.G. 1996a. Analysis of multi-environment data – An historical perspective. pp. 39–124, *in* Cooper & Hammer, 1996, q.v.

DeLacy, I.H., Basford, K.E., Cooper, M. & Fox, P.N. 1996b. Retrospective analysis of historical data sets from multi-environment trials – Theoretical development. pp. 243–267, *in* Cooper & Hammer, 1996, q.v.

Denis, J.B. 1988. Two-way analysis using covariates. *Statistics*, 19: 123–132.

Eberhart S.A. & Russell W.A. 1966. Stability parameters for comparing varieties. *Crop Science 6*: 36–40.

Engledow, F.L. 1925. The economic possibilities of plant breeding. pp. 31–40, *in Proc. Imperial Botanical Conference.* London, F.T. Brooks.

Eskridge, K.M. 1990. Selection of stable cultivars using a safety-first rule. *Crop Science*, 30: 369–374.

Evans, L.T. 1998. *Feeding the ten billion: plants and population growth.* New York, USA, Cambridge University Press.

Falconer, D.S. 1952. The problem of environment and selection. *American Naturalist*, 86: 293–298.

Federer, W.T. & Scully, B.T. 1993. A parsimonious statistical design and breeding procedure for evaluating and selecting desirable characteristics over environments. *Theoretical and Applied Genetics*, 86: 612–620.

Federer, W.T., Nair, R.C. & Raghavarao, D. 1975. Some augmented row-column designs. *Biometrics* 31: 361–374.

Finlay, K.W. & Wilkinson, G.N. 1963. The analysis of adaptation in a plant-breeding programme. *Australian Journal of Agricultural Research*, 14: 742–754.

Fischer, G., Shah, M. & van Velthuizen, H. 2002. *Climate change and agricultural vulnerability*. Vienna, IIASA.

Fox, P.N. & Rosielle, A.A. 1982. Reference sets of genotypes and selection for yield in unpredictable environments. *Crop Science*, 22: 1171–1175.

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

Gauch, H.G. 2007. *MATMODEL Version 3.0: Open source software for AMMI and related analyses*. Ithaca, USA, Cornell University. Downloadable from: http://www.css.cornell.edu/staff/gauch.

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

Gauch, H.G., Piepho, H.P. & Annicchiarico, P. 2008. Statistical analysis of yield trials by AMMI and GGE: further considerations. *Crop Science*, 48: 866–889.

Gilmour, A.R., Cullis, B.R. & Verbyla, A.P. 1997. Accounting for natural and extraneous variation in the analysis of field experiments. *Journal of Agricultural, Biological and Environmental Statistics*, 2: 269–273.

Gilmour, A.R., Gogel, B.J., Cullis, B.R., Welham, S.J. & Thomson, R. 2006. *ASReml User Guide Release 2.0*. Hemel Hempstead, UK, VSN International.

Grüneberg, W.J., Abidin, E., Ndolo, P., Pereira, C.A. & Hermann, M. 2004. Variance component estimation and allocation of resources for breeding sweet potato under East African conditions. *Plant Breeding*, 123: 311–315.

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

Hill, R.R. & Rosenberger, J.L. 1985. Methods for combining data from germplasm evaluation trials. *Crop Science*, 25: 467–470.

IRRI [International Rice Research Institute]. 2007. *CROPSTAT Version 6.1*. Los Baños, Philippines, IRRI. Downloadable from: http://www.irri.org/science/software/cropstat.asp.

Kang, M.S. 1998. Using genotype-by-environment interaction for crop cultivar development. *Advances in Agronomy*, 62: 199–252.

Kang, M.S. ed. 2002. *Quantitative genetics, genomics and plant breeding.* Wallingford, UK, CABI.

Kang, M.S. ed. 2003. *Handbook of formulas and software for plant geneticists and breeders.* New York, USA, Haworth Press.

Kang, M.S. & Gauch, H.G. eds. 1996. *Genotype-by-environment interaction*. Boca Raton, USA, CRC Press.

Kataoka, S. 1963. A stochastic programming model. *Econometrika*, 31: 181–196.

Kempton, R.A. 1984. The use of biplots in interpreting variety by environment interactions. *Journal of Agricultural Science,* 103: 123–135.

Lin, C.S. & Binns, M.R. 1988. A method for analyzing cultivar × location × year experiments: a new stability parameter. *Theoretical and Applied Genetics*, 76: 425–430.

Lin, C.S. & Butler, G. 1988. A data-based approach for selecting locations for regional trials. *Canadian Journal of Plant Science*, 68: 651–659.

Lin, C.S. & Poushinsky, G. 1983. A modified augmented design for an early stage of plant selection involving a large number of test lines without replication. *Biometrics*, 39: 553–561.

Lin, C.S., Binns, M.R. & Lefkovitch, L.P. 1986. Stability analysis: where do we stand? *Crop Science*, 26: 894–900.

Ludlow, M.M. & Muchow, R.C. 1990. A critical evaluation of traits for improving crop yields in water-limited environments. *Advances in Agronomy*, 43: 107–153.

Patterson, H.D. 1997. Analysis of series of variety trials. *In* R.A. Kempton & P.N. Fox, eds. *Statistical methods for plant variety evaluation,* pp. 139–161. London, Chapman & Hall.

Patterson, H.D., Williams, E.R. & Hunter, E.A. 1978. Block designs for variety trials. *Journal of Agricultural Science,* 90: 395–400.

Perkins, J.M. & Jinks, J.L. 1968. Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity*, 23: 339–356.

Piepho, H.P. 1995. Robustness of statistical tests for multiplicative terms in the additive main effects and multiplicative interaction model for cultivar trials. *Theoretical and Applied Genetics*, 90: 438–443.

Piepho, H.P. 1998. Methods for comparing the yield stability of cropping systems - A review. *Journal of Agronomy and Crop Science*, 180: 193–213.

Piepho, H.P. & Möhring, J. 2005. Best linear unbiased prediction of cultivar effects for subdivided target regions. *Crop Science*, 45: 1151–1159.

Piepho, H.P., Denis, J.B. & van Eeuwijk, F.A. 1998. Predicting cultivar differences using covariates. *Journal of Agricultural, Biological and Environmental Statistics*, 3: 151–162.

Podlich, D.W., Cooper, M. & Basford, K.E. 1999. Computer simulation of a selection strategy to accommodate genotype-environment interactions in a wheat recurrent selection programme. *Plant Breeding*, 118: 17–28.

Qiao, C.G., Basford, K.E., DeLacy, I.H. & Cooper, M. 2004. Advantage of single-trial models for response to selection in wheat breeding multi-environment trials. *Theoretical and Applied Genetics*, 108: 1256–1264.

Ribaut, J.M., Fracheboud, Y., Monneveux, P., Banzinger, M., Vargas, M. & Jiang C. 2007. Quantitative trait loci for yield and correlated traits under high and low nitrogen conditions in tropical maize. *Molecular Breeding*, 20: 15–29.

Romagosa, I., Ullrich, S.E., Feng, H. & Hayes, P.M. 1996. Use of additive main effects and multiplicative interaction in QTL mapping for adaptation in barley. *Theoretical and Applied Genetics*, 93: 30–37.

Romagosa, I., Feng, H., Ullrich, S.E., Hayes, P.M. & Wesenberg, D.M. 1999. Verification of yield QTL through realized molecular marker-assisted selection responses in a barley cross. *Molecular Breeding*, 5: 143–152.

Simmonds, N.W. 1979. *Principles of crop improvement*. Harlow, UK, Longman.

Simmonds, N.W. 1991. Selection for local adaptation in a plant breeding programme. *Theoretical and Applied Genetics*, 82: 363–367.

Singh, M., Ceccarelli, S. & Grando, S. 1999. Genotype × environment interaction of cross-over type: detecting its presence and estimating the crossover point. *Theoretical and Applied Genetics*, 99: 988–995.

Singh, M., Malhotra, R.S., Ceccarelli, S., Sarker, A., Grando, S. & Erskine, W. 2003. Spatial variability models to improve dry-land field trials. *Experimental Agriculture*, 39: 1–10.

Singh, S.P., Gutierrez, J.A., Urrea, C.A., Molina, A. & Cajiao, C. 1992. Location-specific and across-location selections for seed yield in populations of common bean, *Phaseolus vulgaris* L. *Plant Breeding*, 109: 320–328.

Smith, A., Cullis, B. & Thomson, R. 2001. Analyzing variety by environment data using multiplicative mixed models and adjustments for spatial field trend. *Biometrics*, 57: 1138–1147.

Smith, A., Cullis, B., Luckett, D., Hollamby, G. & Thomson, R. 2002. Exploring variety-environment data using random effects AMMI models with adjustments for spatial trends. Part 2: Applications. *In* M.S. Kang, ed. *Quantitative genetics, genomics and plant breeding*, pp. 337–351. Wallingford, UK, CABI.

van Oosterom, E.J., Bidinger, F.R. & Weltzien, E. 2003. A yield architecture framework to explain adaptation of pearl millet to environmental stress. *Field Crops Research*, 80: 33–56.

vom Brocke, K., Weltzien, E., Christinck, A., Presterl, T. & Geiger, H.H. 2003. Effects of farmers' seed management on performance and adaptation of pearl millet in Rajasthan, India. *Euphytica* 130: 267–280.

Vargas, M., Crossa, J., van Eeuwijk, F. A., Ramìrez, M.E. & Sayre, K. 1999. Using partial least squares regression, factorial regression, and AMMI models for interpreting genotype × environment interaction. *Crop Science*, 39: 955–967.

Vargas, M., van Eeuwijk, F.A., Crossa, J. & Ribaut, J.M. 2006. Mapping QTL and QTL × environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. *Theoretical and Applied Genetics*, 112: 1009–1023.

Wallace, D.H., Zobel, R.W. & Yourstone, K.S. 1993. A whole-system reconsideration of paradigms about photoperiod and temperature control of crop yield. *Theoretical and Applied Genetics*, 86: 17–26.

Witcombe, J.T. 2001. The impact of decentralized and participatory plant breeding on the genetic base of crops. *In* H.D. Cooper, C. Spillane & T. Hodgkin, eds. *Broadening the genetic base of crop production*, pp. 407–401. Wallingford, UK, CABI.

Wrike, G. & Weber, W.E. 1986. *Quantitative genetics and selection in plant breeding*. Berlin: W. de Gruyter.

Yan, W. & Kang, M.S. 2003. GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists. Boca Raton, USA, CRC Press.

Yates, F. & Cochran, W.G. 1938. The analysis of groust DTPps of experiments. *Journal of Agricultural Science*, 28: 556–580.