

Selection Limits - How Close Are We?

[S.K. St. Martin](#)

**Department of Horticulture and Crop Science
2021 Coffey Rd.
Columbus, OH 43210-1086 USA**

S. K. St. Martin, Dep. of Horticulture and Crop Science, Ohio Agric. Res. and Dev. Ctr., The Ohio State Univ., 2021 Coffey Rd. Columbus, OH 43210-1086. Salaries and research support provided by state and federal funds appropriated to the Ohio Agric. Res. and Dev. Ctr., The Ohio State Univ. This report is Journal Article no. HCS 01-06.

Abstract

Theoretical and empirical studies of selection limits have provided a way to use estimates of genetic gain and effective population size to calculate the plateau (selection limit) beyond which further gains are impossible. Literature pertinent to genetic gain for soybean yield in North America was surveyed for this purpose. Most results indicate that the highest-yielding existing cultivars are far from the calculated selection limit. Selection limits are not an imminent threat to North American soybean breeders, but this does not imply that efforts to broaden the germplasm base are wasted.

Introduction

A population undergoing genetic change encounters a selection limit when the genetic variation is exhausted and no further change is possible, at least in the desired direction. For a breeding program aimed at improving productivity, this may take the form of a “yield plateau” that renders further efforts futile.

Robertson (1960) discovered that the theoretical selection limit was a function of Ns or Ni , where N = population size, s = fitness, and i = standardized selection differential. For additive genes and low values of Ni , total genetic advance (i.e., the cumulative genetic change at the time the selection limit was reached) was $2N$ times the gain in the first generation. Total advance could be much greater, however, for recessive genes, especially if initial frequency were low.

Hill (1982) was the first to look closely at the effect of mutation on long-term selection response. Mutation delays the onset of limits by providing a mechanism to introduce new genetic variation. Hill concluded that it would take approximately 20 cycles of selection before mutation became a significant source of genetic variation.

Linkage disequilibrium, in the form of increased repulsion linkages, increases as a consequence of selection [“Bulmer’s effect” (Hospital and Chevalet, 1996)]. These linkages may reduce the short-term rate of genetic gain but in effect store genetic variation that can later be released by intermating.

Weber and Diggins (1990) analyzed data from 27 published, long-term (50 cycles or more) selection experiments, mostly with model systems, but including the Illinois high-protein and high-oil maize populations (Dudley and Lambert, 1992). In Figure 4 of their paper, they presented the relationship between effective population size and the ratio of cumulative response after 50 cycles (R_{50}) to first-generation response (R_1). Three separate sigmoidal curves in the Figure represent theoretical response, with and without mutation, and realized response, taken from the published long-term experiments. The ratio R_{50}/R_1 increases as effective population size increases. The three curves are roughly parallel, with the realized responses giving somewhat smaller values of R_{50}/R_1 than the theoretical curves. Weber and Diggins (1990) noted that experiments with effective population size less than 20 reached a selection limit prior to completing 50 cycles. In these cases, the value of R_{50}/R_1 represents the total amount of gain achieved at the limit, in relation to the gain made in the first cycle.

St. Martin (1982) used coefficients of parentage to estimate the effective population size for the North American soybean breeding program in maturity groups 00 to IV. He concluded that the effective population size was 11 to 15 inbred lines, equivalent to 5.5 to 7.5 non-inbred individuals selected per cycle.

Concern has frequently been expressed that the relatively narrow germplasm base in North American soybean breeding may limit progress (e.g., Carter et al., 1998; Sneller, 1999). A selection limit is the most drastic limitation to progress. The objective of this paper is to use the response curves published by Weber and Diggins (1990) and estimates of effective population size to calculate selection limits for grain yield for North American soybean breeding, using previously published estimates of genetic gain.

Materials and Methods

All published papers estimating rates of genetic gain for yield in North American soybean breeding were included in these calculations. These papers were reviewed by Specht et al. (1999), with the exception of the most recent studies (Wilcox, 2000; Uston et al., 2001). In most of these experiments, researchers tested old and recent cultivars in a common experiment and used regression of cultivar yield on year of release to estimate genetic gain per year. Wilcox (2000) used 60 years of regional test data, regressing the yield of the best entries on the year of the test.

With an effective population size between 5.5 and 7.5 (St. Martin, 1982), the empirical curve of Weber and Diggins (1990) yields a value of $R_{50}/R_1 \gg 10$. (Slightly larger values are obtained if the theoretical curves are used instead.) Therefore, a selection limit was calculated as $Y_0 + 10R_1$, where Y_0 = initial yield and R_1 = gain achieved in the first cycle. A cycle of selection was assumed to require 10 years; hence the value of R_1 was taken as 10 times the rate of gain/year estimated in the paper surveyed. With sporadic exceptions, the first U.S. and Canadian varieties that derived from crosses were released beginning in 1939 (Bernard et al., 1988). Therefore, the initial date for the breeding program was taken as 1930. Genetic gain achieved prior to this first cycle of crossing was due to selection among introduced cultivars. This gain was not relevant to the

calculations, and therefore plant introductions were excluded from the regression analysis unless they represented the most recent cycle, just before the modern releases began in 1939. This necessitated recalculating the author's regression lines in a few cases. The regression lines were used to obtain estimated yields for 1930, by extrapolation if necessary, and these values were used for Y_0 .

A second selection limit was calculated using an updated estimate of effective population size. The estimate of effective population size (St. Martin, 1982) is obsolete and limited in scope, derived from Northern public cultivars released up to 1980, i.e., two or more breeding cycles ago. More recently, Sneller (1994, 1999) found no evidence for an increase in mean coefficient of parentage in the past 20 years. Indeed, his (1999) mean value of 0.17, which includes public and private cultivars from both the North and the South, is less than the value of 0.25 reported by St. Martin (1982). So too is his estimate for Northern cultivars (0.21). For a given effective population size, mean coefficient of parentage is expected to increase with each breeding cycle. The smaller mean coefficient of parentage in recent reports suggests a larger effective population size than that of the earlier paper. Revisiting the calculations of St. Martin (1982), and assuming two or three additional cycles since his study, we conclude that a mean coefficient of 0.17 would result from an effective population size of 27 to 38 inbred lines, equivalent to 13.5 to 19 non-inbred individuals. Applying these values to the empirical curve in Figure 4 of Weber and Diggins (1990), we obtain a value of R_{50}/R_1 in the range from 15 to 20. The second selection limit was therefore calculated as $Y_0 + 15R_1$.

Results and Discussion

Calculated selection limits are shown in [Table 1](#). In most cases, even the limits derived from the assumption of low effective population size were well above the yield level of the best cultivar. The limits were also generally greater than any regional test means that are commonly achieved. In two cases, involving maturity groups 0 and VI, negative estimates for rate of gain implied that the limit has already been reached. Both negative estimates, however, derived from cultivars tested outside their areas of adaptation, i.e., group 0 cultivars in Nebraska and group VI cultivars in Florida. Because there is significant germplasm exchange between maturity groups, it is difficult to argue for a selection limit in a single maturity group in isolation. Considering Table 1 as a whole, it seems clear that a selection limit for yield in North America is not imminent.

There are at least two plausible arguments that the true selection limits are higher than those in Table 1. First, although spontaneous mutation is not likely to create significant useful variation in the early cycles, new sources of genetic variability are now being added to the soybean germplasm in the form of transgenes. The contribution of genes from outside the species has not been taken into account in the literature on selection limits, but, because they expand the genetic possibilities, they cannot decrease the selection limit and may increase it. Even if a particular gene has no direct effect on seed yield, its incorporation into the working germplasm base offers the opportunity to select for genes that feature complementary interactions with the transgene.

Second, changes in the production environment may unlock genetic variation that was not previously expressed. These changes may involve agronomic practices or environmental phenomena. Martin and Hallauer (1980) attributed increased genetic variances in a maize breeding population to the change from hand-harvesting to mechanical harvesting of test plots. Specht et al. (1999) discussed the implications for soybean productivity of the increase in CO₂ concentration of the earth's atmosphere, a change which is expected to continue in the future. Interactions between new environments and existing genes may offer greater scope for selection advance than is possible in a static environment.

Use of molecular markers in marker-assisted selection should have no effect on the selection limit. If effective, however, marker-assisted selection may permit the limit to be reached more quickly. In the near future, the cost of mapping quantitative traits precludes extending this tool beyond a narrow base of elite parents, suggesting a scenario of rapid progress in the short term as known markers are exploited, followed by a period of slow progress or a temporary plateau while a broader set of germplasm is explored for rare, favorable genes. Bernardo (2001) has raised doubts as to the effectiveness of marker-based selection for traits controlled by a large number of genes.

There is also an argument that selection limits could become important to some soybean breeders in the near future. Barriers to intermating create isolated sub-populations, each with its own effective population size. Such a barrier has been introduced to soybean breeding with the practice of patenting new varieties and restricting their use as germplasm outside the organization that developed them. The largest companies active in North American soybean breeding are reluctant to license their patented cultivars for use as germplasm, a practice that could result in closed populations for some breeders or groups of breeders. Perhaps it would be useful for each company to calculate its own selection limit, as imposed by these restrictions on the flow of germplasm. Because patents expire after 20 years, however, their long-term effect on the industry's collective limit may be small.

Although selection limits do not pose an immediate threat, a pertinent but much more difficult question is whether the current effective population size is optimal for the amount of genetic gain needed in the short and long term. The initial narrowing of the North American soybean germplasm base (i.e., up to about 1980) came about because of successful cooperative regional testing, with cooperative release of a relatively small number of broadly-adapted cultivars, and free exchange of germplasm between breeders. These factors led to each breeder having roughly the same elite genetic material to work with. High selection intensity was accompanied by both low effective population size and high short-term gains. It is difficult to argue against cooperative testing, cooperative release, and germplasm exchange. It is also difficult to fault a soybean breeder of, say, 1950, for a focus on short-term progress in the face of a world population then increasing at roughly 2% annually and an expected increase of like magnitude in the demand for crops. With current annual population growth rate reduced to 1.2% and expected to fall further (United Nations Population Division, 2000), perhaps it is easier now to justify an interest in long-term goals. An encounter with a selection limit might be just as

destructive 100 years from now as it would be today. Even with no immediate threat of a selection limit, it is possible that current efforts to broaden the germplasm base will prove extremely helpful to future generations.

References

Bernard, R. L., G. A. Juvik, E. E. Hartwig, and C. J. Edwards, Jr. 1988. Origins and pedigrees of public soybean varieties in the United States and Canada. U. S. Dept. of Agric. Tech. Bull. No. 1746.

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

Boerma, H. R. 1979. Comparison of past and recently developed soybean cultivars in maturity groups VI, VII, and VIII. *Crop Sci.* 19:611-613.

Carter, T. E., Jr., R. L. Nelson, P. B. Cregan, H. R. Boerma, P. Manjarrez-Sandoval, X. Zhou, W. J. Kenworthy, and G. N. Ude. 1998. Project SAVE (Soybean Asian Variety Evaluation) - Potential new sources of yield genes with no strings from the USB and public/private research. Proc. 28th Soybean Seed Research Conference, pp. 68-83.

Dudley, J. W., and R. J. Lambert. 1992. Ninety cycles of selection for oil and protein in maize. *Maydica* 37:1-7.

Hill, W. G. 1982. Predictions of response to artificial selection from new mutations. *Genet. Res. Camb.* 40:255-278.

Hospital, F., and C. Chevalet. 1996. Effects of population size and linkage on optimal selection intensity. *Theor. Appl. Genet.* 86:775-780.

Luedders, V. D. 1977. Genetic improvement in yield of soybean. *Crop Sci.* 17:971-972.

Martin, J. M., and A. R. Hallauer. 1980. Seven cycles of reciprocal recurrent selection in BSSS and BSCB1 maize populations. *Crop Sci.* 20:599-603.

Robertson, A. 1960. A theory of limits in artificial selection. *Proc. Roy. Soc. London Series B.* 153:234-249.

St. Martin, S. K. 1982. Effective population size for the soybean improvement program in maturity groups 00 to IV. *Crop Sci.* 22:151-152.

Salado-Navarro, L. R., T. R. Sinclair, and K. Hinson. 1993. Change in yield and seed growth traits in soybean cultivars released in the southern USA from 1945 to 1983. *Crop Sci.* 33:1204-1209.

Sneller, C. H. 1994. Pedigree analysis of elite soybean lines. *Crop Sci.* 34:1515-1522.

Sneller, C. H. 1999. Diversity within commercially used gene pools. Proc. World Soybean Res. Conf. VI:176-184.

Specht, J. E., D. J. Hume, and S. V. Kumudini. 1999. Soybean yield potential - a genetic and physiological perspective. Crop Sci. 39:1560-1570.

Specht, J. E., and J. H. Williams. 1984. Contribution of genetic technology to soybean productivity - retrospect and prospect. Chap. 3 in W. R. Fehr (ed.). Genetic contribution to yield gains of five major crop plants. CSSA Special Publ. 7. Crop Sci. Soc. of Amer. and Amer. Soc. of Agron., Madison, WI.

United Nations Population Division. 2000. World population prospects: the 2000 revision. <http://www.undp.org/popin/wdtrends/wdtrends.htm>.

Ustun, Ali, Fred L. Allen, and Burton C. English. 2001. Genetic progress in soybean of the U.S. Midsouth. Crop Sci. 41:993-998.

Voldeng, H. D., E. R. Cober, D. J. Hume, C. Gillard, and M. J. Morrison. 1997. Fifty-eight years of genetic improvement of short-season soybean cultivars in Canada. Crop Sci. 37:428-431.

Weber, K. E., and L. T. Diggins. 1990. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. Genetics 125:585-597.

Wilcox, J. R., W. T. Schapaugh, Jr., R. L. Bernard, W. R. Fehr, and M. H. Niehaus. 1979. Genetic improvement of soybeans in the Midwest. Crop. Sci. 19:803-805.

Wilcox, J. R. 2000. Sixty years of progress in improving soybean cultivars. Agron. Abstr. p. 95.

Table 1. Selection limits for soybean seed yield calculated from published estimates of genetic gain for North America.

Reference	Maturity group	Rate of gain*	Highest yield**	Calculated selection limit***	
				low Ne	high Ne
		(kg/ha)/yr	kg/ha		
				low Ne	high Ne
Luedders, 1977	II to IV	17.6	2428 ('Kent')	3416	4297

Boerma, 1979	VI	10.4	2434 ('Tracy')	3013	3531
Boerma, 1979	VII	17.9	2565 ('Bossier')	3585	4480
Boerma, 1979	VIII	14.8	2543 ('Hutton')	3225	3965
Wilcox et al., 1979	II	17.3	3383 ('Amsoy 71')	4291	5158
Wilcox et al., 1979	III	13.2	3160 (('Woodworth'))	3902	4561
Specht & Williams, 1984	00	33.5	----	4243	5918
Specht & Williams, 1984	0	-3.9	----	----	----
Specht & Williams, 1984	I	8.6	----	3826	4256
Specht & Williams, 1984	II	4.3	----	4025	4240
Specht & Williams, 1984	III	13.6	----	4330	5010
Specht & Williams, 1984	IV	9.7	----	3776	4261
Salado-Navarro et al., 1993	VI	-9.9	2366 ('Davis')	----	----
Salado-Navarro et al., 1993	VII	13.4	2552 ('Braxton')	3054	3723
Salado-Navarro et al., 1993	VIII	17.1	2685 ('Coker-368')	3376	4232
Voldeng et al., 1997	00	13.4	3652 ('PS42')	3738	4410
Voldeng et al., 1997	0	12.1	3846 ('9071')	4067	4674
Wilcox, 2000	00	21.7	----	3390	4474
Wilcox, 2000	0	25.8	----	3947	5237

Wilcox, 2000	I	30.4	----	4455	5976
Wilcox, 2000	II	29.3	----	4615	6078
Wilcox, 2000	III	30.6	----	4657	6188
Wilcox, 2000	IV	29.5	----	4514	5990
Ustun et al. 2001	V,VI,VII	14.4	2720 ('Hutcheson')	3240	3960

*omitting all cultivars derived from plant introductions except for the most recent such cultivars.

**for best-yielding entry in study.

***Assumes cumulative gain is 10 or 15 times gain/cycle, depending on effective population size (Ne). Low Ne (5.5 to 7.5) estimated by St. Martin (1982); high Ne (13.5 to 19) based on more recent estimates. Also assumes 1 cycle = 10 years. Initial yield calculated for 1930 by extrapolation using linear regression equation from reference. Dash (----) indicates negative gain, limit already reached with existing cultivars.