Evaluating F_3 Lines of Single Crosses for Yield Improvement in Cotton

J. Lançon
CIRAD-CA-Programme Coton, PARAB, Recherche Coton et Fibres, 01 BP 715, Cotonou, Bénin.

ABSTRACT

In Benin, average cotton yields have stagnated for several years. Although the main factors are known to be decreasing levels of inputs by the farmers, yield improvement has become a research priority for cotton research as a whole, including plant breeding. Selecting variables with low heritability like yield requires specific experimental field design, more demanding than those used for less complex characters. The design in which single crosses are evaluated through the performance of an F_3, 30 lines sample as described in this paper intends to meet this requirement. The lines are grown in cropping conditions that are similar to those of commercial fields. After examining the general parameters status under which the design is most relevant, examples show how to interpret the data: (1) from the statistical analysis, the crosses are compared and their lines ranked in each cross; then (2) with an appropriate quantitative genetics model, the genetic variability of the traits under selection can be described more precisely. Finally, the cost-effectiveness of this design is discussed and compared to more common designs.

Introduction

In 50 years, research in general and plant breeding in particular has increased cotton production appreciably in French speaking Africa. Productivity and fiber quality have improved and the varieties bred are among the best in the Upland family (CIRAD-CA, 1997). These results were obtained with pure line cotton cultivars that were bred with proven methods for average or highly heritable traits (ginning out-turn, earliness, fiber quality). However these methods have limitations for improving productivity when expression is hidden by environment and complexity of the genetic effects in action, including dominance and epistatic interactions (Lançon, 1994). Breeders have to reconsider this strategy because:

1) genetic variability and heritability of most traits tends to decrease so we can expect non additive, especially epistatic effects, to gain importance;
2) cotton production is moving to the private sector and economic and political power is shifting in favour of cotton producers;
3) weakening extension services and increases in input prices are leading to greater agronomic diversity and a trend towards extensive farming.

Cotton breeding must adapt to these changes that deal with producer priorities, in particular hardiness and yield, and to take account of the real or potential decrease in additive genetic variability.

This paper describes a design that was studied on cotton by Lançon (Lançon, 1995; Lançon 1995a), based on early evaluation of crosses (Gallais, 1988) that should fill both requirements

Preliminary Considerations

Selection in the F_2 generation

Response to selection may be written as:

$$R = i\, h^2_N\, s_P = i\, s^2_A / s_P$$

where s^2_P is the phenotypic variance, s^2_A the additive genetic variance and i is the selection intensity, which varies according to the fraction of selected plants.

Selection efficiency in the F_2 is then directly linked to the importance of additive genetic variance and to narrow heritability h^2_N.

For many agronomic traits, heritability estimated on individual F_2 plants is low (Table 1). Most of the variation is not additive, rendering selection inefficient and biased by micro-environmental or non-additive genetic variation.

Selection within or between crosses?

At constant means, the option may be either to concentrate breeding efforts on crosses selected *a priori* or to do more crosses and select the most promising. Simulations by Fouilloux (1981) show that the choice depends on the trait's heredity. As long as the additive genetic variance remains significant, it is preferable to cross carefully selected parents but when the additive genetic variance is low, selecting between crosses becomes more efficient. Several decades of active cotton breeding in Francophone Africa have shown that the additive part of the genetic variability utilized for improving productivity has decreased and that the best strategy involves multiplying and evaluating numerous crosses.

How to evaluate crosses?

To select few crosses among numerous, one has to consider their mean value and their additive genetic variance (Lançon *et al.*, submitted) :

1) the parents are good indicators only in the case of favourable gene repulsion and in the absence of any epistatic effect;
2) the F_1 generation does not segregate and its mean is biased by dominance;
3) F_2 variance includes genetic and environmental variation and its mean includes some dominance;
4) the F_3 generation segregates: it can be organized in lines and grown at normal plant densities that integrates competition effects. Genetic components of the mean and variance can be estimated more easily under more realistic cropping conditions.

The EE design

F3 lines are tested at crop spacing *e.g.* in a Fisher block layout with 3 replications. Each cross is represented by a sample of at least 30 F_3 lines, in order to estimate between and within line variability. When added to the trial, the parents may be used as checks and provide an estimate of the environmental variance.

Statistical analysis

Are there differences between cross means?

In the analysis of variance and appropriate tests, if the line data are obtained from several plants, a within-line variance may be estimated directly and compared to that provided by the analysis of variance (Table 2).

Which are the best crosses?

For each cross, a genetic progress expectation may be evaluated from the F_3 line mean and variance (Table 3) at a given selection intensity (Lançon 1995a). Between F_3 lines variance is assumed to be worth half the additive genetic variance (A) of the cross (*cf* § 2.2.2).

Which are the best lines in each cross?

The analysis shown in § 2.1.2 can be applied to the F_3 lines. Generally, only the mean is used but if several data per line are available a variance may be estimated.

Genetic analysis

Mean genetic analysis

The EE trial can also be utilized to estimate parameters necessary for the genetic description of the trait. The model uses Kearsey and Pooni's notation (1996) where effects are written a (additive), d (dominance), a.a (*cis* epistasy) and d.d (*trans* epistasy).

If one considers only the dominance effects, d, and dominance x dominance interaction, d.d, the F3 generation mean can be written as:

$$\overline{F3} = m + [a] + \frac{1}{4}[d] + \frac{1}{16}[d.d]$$

Also, in the presence of additive x additive epistasy, a.a, each parent mean is written:

$$\overline{P1} = m + [a] + [a.a]$$

$$\overline{P2} = m - [a] + [a.a]$$

from which can be derived:

$$\overline{P1} = m + [a.a] = \frac{\overline{P1} + \overline{P2}}{2}$$

Genetic analysis of the variance

The F_3 genetic variance may be divided into between lines variance, GS^2_B, and within lines variance, GS^2_W.

According to Kearsey and Pooni's model (1996) and considering only the main 3 components of the genetic variance *ie* A, the additive variance, D, the dominance variance and AA, the *cis* epistatic variance, becomes:

Total genetic variance:

$$GS^2 = 3/4.A + 3/16.D + 9/16.AA$$

Between lines genetic variance:

$$GS^2_B = 1/2.A + 1/16.D + 1/4.AA$$

Within lines genetic variance:

$$GS^2_W = 1/4.A + 1/8.D + 5/16.AA$$

If the parents are not part of the EE trial, only two variances are computed and two parameters estimated *ie* A and E.

From Table 2 ("mean squares expectations") and the preceding model, we have:

$$s^2_B = GS^2_B = 1/2.A$$

$$s^2_W = GS^2_W + V_E = 1/4.A + V_E$$

and

$$A = 2.s^2_B$$

$$V_E = s^2_W - 1/2.s^2_B$$

When the parents are included in the trial, V_E may also be estimated through:

$$(V_{P1}+V_{P2})/2$$

If the result gets close enough to that obtained in the preceding case, it can be accepted as an acceptable estimator of environmental variance.

An extra equation is available for estimating a third parameter. In general, D will be preferred to AA, even though the coefficient of AA is inferior to that of D. The system of equations becomes:

$$V_E = (V_{P1}+V_{P2})/2$$

$$s^2_B = 1/2.A + 1/16.D$$

$$s^2_W = 1/4.A + 1/8.D + V_E$$

and

$$A = 4/3\ (2\ s^2_B - s^2_W + V_E)$$

$$D = 16/3\ (2\ s^2_W - s^2_B - 2\ V_E)$$

The Whole Breeding Design

Description

A program that includes an EE trial targets low heritability traits, especially yield improvement, but selection pressure must be maintained on other traits.

The F_2 generation is sown at nursery spacing and economically important traits with good heritability and genetically not linked with yield may be screened *e.g.* ginning out-turn, seed index or seed coat fragments rate in the fiber. At least 30 F_2 plants are selected to produce 30 F_3 lines that are evaluated in the EE trial. On their performance in the EE trial, superior lines of the 50% best crosses are followed by conventional breeding. Note that evaluating the crosses productivity in the F_3 should reduce the number of genotypes in station yield tests in later breeding steps.

Cost Estimates

The respective costs of conventional pedigree-selection or EE programmes may be compared with two easy to compute parameters:

1) the total area covered by the trials;
2) the number of fiber analyses required.

With EE, the area covered by the F_3 population is obviously larger than it is with a conventional programme (Table 4). The number of observed lines must represent the whole cross and must include most of its variability, both within and between lines. An F_2 generation could be represented by 200 plants but the F_3 requires at least 30 lines with 30 plants per line.

The EE trial is kept to a reasonable size by:

1) restricting each elementary plot length (6 m); and
2) observing the lines at the crop stand (0.3 x 0.8 m).

With these restrictions, the area covered by an F_3 population is only about 450 m².

The number of fiber analyses is important because they require external, expensive expertise, their number being an indicator of laboratory work require (Table 5). Designs that include an EE trial do not require fiber analyses before the F_3 generation and then only for lines. The following generations reinforces the results.

Pedigree selection starting in the F_2 needs about the same number of analyses per cross (around 200) as programmes with F_3 EE. The number of analyses is similar for similar numbers of hybrids annually.

Conclusion

The ability to screen many crosses early in the breeding process under normal growing conditions, facilitates orienting a breeding program towards improvement of traits with relatively low heritability. This flexibility allows breeders to elaborate better answers to foreseeable changes in African cotton production.

References

CIRAD-CA. (1997): Manuel pour l'amélioration du cotonnier. D. Dessauw and B. Hau (Éd)., Montpellier (France).

Fouilloux, G. (1981): Sélection intra ou inter croisements chez les espèces autogames. Sélectionneur Français, 29:53-59.

Gallais, A. (1988): Predicting of the mean and variance of lines and hybrids which can be derived from a random mating population. Theoret. Appl. Genet., 75:625-627.

Kearsey, M.J. and H.S. Pooni. (1996): The Genetical Analysis of Quantitative Traits. Chapman and Hall, London

Lançon, J. (1995): Plant spacing and cotton breeding for improving productivity: predicting the line value of a single cross hybrid. In: Proc. Meeting working group "Cotton breeding". U. Kechagia and F. Xanthopoulos. (Ed). FAO, Adana, Turkey. Pp. 12-19.

Lançon, J., 1994.- L'amélioration du cotonnier au CIRAD-CA. Séminaire Traitements

statistiques des essais de sélection, série Colloques du CIRAD. Pp. 275-293.

Lançon, J. (1995).- Effet de la densité de semis en sélection sur l'amélioration génétique du cotonnier : interactions, structures de corrélations, hétérosis et valeur en lignées. Thèse de doctorat. Université Paris-Sud (Orsay). 119p.

Table 1. Broad and narrow sense heritability estimates at the plant level (as %) at crop (42,000 plants/ha) and nursery (10,000 plants/ha) spacings.

Trait	GOT	SI	NCF	NBV	HT	LBV
Broad heritability						
Crop spacing	75	66	55	16	57	26
Nursery spacing	69	69	30	14	52	39
Narrow heritability						
Crop spacing	75	29	7	16	33	14

Nursery spacing	69	17	15	14	30	11

(after Lançon *et al*, submitted for publication).
GOT: ginning out-turn (%); SI : seed index (g); NCF : number of bolls borne by the fruiting branches; NBV: number of vegetative branches; HT : plant height (cm); LBV: length of the longest vegetative branches (cm)

Table 2. Analysis of variance of an EE trial including c crosses, f lines per cross and r replicates.

Source of variation	df	MS	E (MS)	Test
Between crosses	c-1	m1	$s^2_{Wl} + r\, s^2_{Bl} + rf\, Q(C)$	m1/m2
Between lines within crosses	c x (f-1)	m2	$s^2_{Wl} + r\, s^2_{Bl}$	m2/m3
Within lines	(c x f) x (r-1)	m3	s^2_{Wl}	
Total	(c x f x r)-1			

Table 3. Potential genetic progress estimates in each cross.

Cross	Lines mean	Between lines variance	Expected genetic progress
A x B	X_{AB}	VarAB	X_{AB} + 0.798 x _(0.20 x Var_{AB})
A x C	X_{AC}	VarAC	X_{AC} + 0.798 x _(0.20 x Var_{AC})
B x D	X_{BD}	VarBD	X_{BD} + 0.798 x _(0.20 x Var_{BD})
etc			

h_N^2 = 0,10; Var_{AB} represents half of the additive genetic variance of the AxB cross; and the selection pressure is 0.5 corresponding to k = 0.798

Table 4. Areas occupied by two breeding designs based on conventional pedigree selection or including an early evaluation trial.

On station breeding activities	Pedigree selection (1)	Early Evaluation
Crosses	1.2 per 1000 m² of selection	1.2
F2 generation	25 to 30% of the segregating material	10%
EE trial in F3	25 to 30% of the segregating material	55%
F4 and following generations	40 to 45% of the segregating material	35%
On station evaluation trials	35 to 40% of the total breeding program	35 to 45 %
TOTAL for 10 crosses (m²)	*13,000 m²*	*13,000 m²*

(1) after Lançon (1994) and rectified for a 0.8 m instead of 1 m inter-row

Table 5. Number of analysis per cross.

On station breeding activities	Pedigree selection	Early Evaluation
F2 generation	70 plants	
F3 EE trial	20 lines and 50 plants	2 x 30 lines
F4 generation	10 lines and 35 plants	1/2 x (30 lines and 125 plants)
F5 generation	5 lines and 20 plants	1/2 x (30 lines and 125 plants)
F6 and following generations	5 lines and 10 plants	1/2 x (5 lines and 30 plants)
Total	*225*	*230*