

ALGUNAS CONSIDERACIONES SOBRE LA SELECCION DE HIBRIDOS DE MAIZ

M.A. RAPELA¹ y C. BANCHERO²

Recibido: 03/12/98

Aceptado: 09/03/99

RESUMEN

El presente trabajo tiene como fin el de revisar las causas de las aparentes discrepancias entre las recomendaciones de híbridos de maíz realizadas por instituciones públicas y privadas y el comportamiento real de los híbridos a campo.

Se discuten los problemas del diseño experimental, tamaño de parcela, número de repeticiones por ensayo experimental, número de ensayos experimentales por localidad y por año, check plots, strip tests, análisis cabeza a cabeza, parámetros de estabilidad, análisis AMMI y ensayos en campos de productores.

El incremento de la velocidad y de la efectividad de los programas de mejoramiento genético de maíz con un fin comercial puede ser obtenido de varias maneras, incluyendo tanto a los métodos clásicos como a las más recientes herramientas biotecnológicas y la ingeniería genética. Sin embargo detectar y seleccionar los mejores híbridos dentro de un conjunto de competidores de alta performance, continúa siendo una de las tareas más dificultosas de estos programas.

Palabras clave. Selección - Híbridos - Maíz - Mejoramiento

SOME CONSIDERATIONS ABOUT CORN HYBRID SELECTION

SUMMARY

The discrepancies between corn hybrid recommendations by public and private institutions and the real behavior of these hybrids in the field are reviewed.

The present paper addresses the problems associated with the experimental design, plot size, number of replications by trial, number of trials by location and per years, check plots, strip tests, head to head analysis, stability parameters, AMMI analysis, and on-farm research.

The increase of speed and effectiveness of commercial driven corn breeding programs can be obtained by several ways, including either conventional methods as well as recently developed biotechnological and genetic engineering tools. Selecting and detecting the best hybrids from a group of high performing ones, however, still continues to be one of the most critical tasks of such programs.

Key words. Selection - Hybrids - Corn - Breeding

INTRODUCCION

Algunas de las preguntas habituales que se hacen tanto agrónomos extensionistas como productores, relacionadas a cualquier material genético, sean tanto híbridos F_1 como en el caso de maíz, sorgo y girasol, así como variedades, para el caso de trigo y soja, son: i) ¿cuánta confianza se puede poner en los datos de rendimiento de estos materiales cuando aquellos fueron generados en empresas semilleras?; ii) ¿cuánta, si éstos provienen de ensayos de instituciones públicas?; iii) ¿el material ha sido efectivamente evaluado en todas las áreas para las cuales se lo recomienda?; iv) ¿por qué los datos de los ensayos son buenos y a

¹Director de Investigaciones. Agar Cross S.A. División Semillas. E-mail miguelr@agarcross.com.ar

²Cátedra de Genética, Facultad de Agronomía, Universidad Nacional de Buenos Aires. E-mail banchero@mail.agro.uba.ar

^{1,2}Co-Directores Convenio de Vinculación Tecnológica FAUBA-Agar Cross.

campo el material no demuestra una performance tan competitiva (o viceversa)?; v) ¿por qué cada compañía presenta a sus híbridos como los mejores?.

Generalmente los fitomejoradores, creadores o introductores de los híbridos, deben contestar preguntas como éstas y muchas más relacionadas todas ellas con la exactitud o credibilidad de la información que llega a público conocimiento, varias veces al año, todos los años. Las respuestas no son fáciles; por el contrario, no sólo son difíciles sino que, a menudo, son contradictorias. ¿Cuál es la razón por la que es tan difícil hacer recomendaciones predictivas relativamente seguras sobre el comportamiento de un material genético en general, o de híbridos de maíz en particular como trata el presente trabajo, con el mínimo margen de error posible?.

Los productores de maíz pueden hacer uso de una variada gama de fuentes de información para seleccionar sus híbridos como por ejemplo: a) experiencia personal, b) agentes de extensión, c) promotores de empresas semilleras, d) información pública proveniente de agencias estatales, e) comentarios de vecinos, etc. Diversos estudios en varias partes del mundo han detectado que ninguna de estas fuentes goza de absoluta confianza por parte de los productores (Bowman, 1998; Carter y Hudelson, 1992; Kalton, 1992; Miskin, 1992).

A fin de entender el origen de estas dificultades, se debe observar que para que un híbrido experimental de maíz llegue a la etapa precomercial, el fitomejorador ha seleccionado el mismo a través de Ensayos Comparativos de Rendimiento (ECR). Por ejemplo, cada ECR consta de 3 repeticiones, o sea 3 parcelas de cada híbrido lo que equivale a decir aproximadamente 21 m^2 ($7 \text{ m}^2 \times 3$, suponiendo un largo de surco de 5m y dos surcos por parcela). Supongamos que en el primer año de ensayo (etapa de ensayos preliminares) se realizaron 3 ECR, 7 en el segundo año, 10 en el tercero (etapa de híbrido experimental) y 20 en el cuarto año (etapa de híbrido pre-comercial); esto significa 40 ECRs, que es lo mismo que decir que el híbrido en cuestión, ha sido evaluado en 840 m^2 de parcelas. En otras palabras, aproximadamente 1/12 parte de hectárea.

A partir de esos datos, tomados sobre la base de diseños estadísticos probados, analizados por medio de técnicas de análisis de varianza y regresión lineal para determinar estabilidad, junto con todas las observaciones personales realizadas, el fitomejorador debe decidir qué híbrido selecciona para la fase comercial.

Llegado a este punto, el fitomejorador suele hacer Ensayos Comparativos de Rendimiento en Grandes Superficies (ECRGS), muchas veces denominados también Ensayos en Macroparcelas. Con fortuna, si las condiciones climáticas han cooperado y otra serie de imponderables no necesariamente biológicos pudieron ser salvados, se podrá contar con información de varios ECRGSs en parcelas apareadas con testigos. Suponiendo la realización (cosecha efectiva) de 20 ECRGSs, cada híbrido habrá sido ensayado en más o menos 5000 m^2 por ECRGS, lo que significa una evaluación sobre la base de 10 hectáreas.

En realidad ya no hay más datos, y lo que es más difícil, con ellos hay que contestar a todas las preguntas del principio. Es obvio que, si el material llega a la fase comercial y se afianza en ella, año tras año el caudal de información que se contará sobre el mismo irá en permanente aumento, lo que significa un correlativo incremento de la base de seguridad de las respuestas.

Stucker y Hicks (1991) de la Universidad de Minnessota, sostienen que quedan muy pocas dudas en las mentes de los investigadores agrónomos, personal de extensión y productores, que la aceptación general de los ensayos en pequeñas parcelas realizados en las estaciones experimentales ha declinado sustancialmente en los últimos 10 años. Estos autores consideran que, la mejor solución encontrada, implica la evaluación en el campo del productor utilizando enormes parcelas como unidad experimental.

La base estadística de la anterior conclusión es sencilla: a mayor tamaño de muestra, las inferencias realizadas sobre la población (el comportamiento en grandes extensiones del híbrido de maíz en este caso)

tendrán menor margen de error. Sin embargo, llevar esto a la práctica es casi imposible; partiendo de la base que un programa chico de investigación evalúa aproximadamente unos 2000 nuevos híbridos por año, la impracticabilidad del ensayo en grandes parcelas para evaluar a todo este material es obvia.

Este tipo de análisis, sólo podría quedar relegado al último eslabón de prueba, o sea en la etapa precomercial. Pero, entonces, conviene hacerse la pregunta: ¿qué fue lo que se seleccionó antes?. O mejor aún: ¿cómo estar seguro que se seleccionó lo bueno y descartó lo malo y no al revés?.

LA FILOSOFIA DETRAS DE LA PARCELA

Bradley *et al.*,(1988), han dividido los sistemas de evaluación de híbridos de maíz en tres etapas: 1) pasado (antes de 1980); 2) presente (década del 80) y, 3) futuro (1990 en adelante). Siguiendo la línea de análisis de estos autores y aplicándolo a la situación que en general ocurre en nuestro país, se pueden hacer las siguientes consideraciones:

El Pasado

Bradley *et al.*,(1988) dicen que en "el pasado" los sistemas de evaluación ponían un gran énfasis en obtener información precisa a partir de muy pocas localidades, reflejo del sobre dimensionamiento de las aproximaciones académicas del diseño experimental, a través del cual se quería obtener la máxima cantidad de información posible a partir de un grupo limitado de ambientes. El diseño experimental, tendía a maximizar la precisión para una localidad en un año determinado. Desgraciadamente, esta aproximación trataba a los ambientes como un efecto fijo (no aleatorio), y medía la performance de un híbrido en sólo uno de los infinitos ambientes posibles. En estas circunstancias, el limitado número de localidades únicamente sirve para describir el pasado de un híbrido, pero no para predecir su futura performance.

A esto se le debe sumar que los ambientes elegidos para los ensayos eran siempre los de alto rendimiento a causa de que era el rendimiento el carácter de más interés a determinar y porque los buenos ambientes dan siempre bajo nivel de coeficiente de variación. Los ambientes con stress se evitaban, y si ellos llegaban a aparecer, los datos se excluían de los análisis sumarios. El presupuesto de investigación se volcaba fundamentalmente hacia la reducción del error experimental en una localidad dada.

Interpretando lo que dicen Bradley *et al.*,(1988) y traduciéndolo en ejemplos, esto significa buscar el mejor campo posible en un área geográfica o ecológica dada, que venga con un buen manejo, adecuada rotación, sin antecesor de maíz, bajo o nulo nivel de malezas, adecuado y fuerte uso de fertilizante, aplicación de herbicidas, varias repeticiones (tres o más) en la misma localidad, muy pocas (a menudo uno o dos) localidades de ensayos, diseño del ensayo en látice, semillas seleccionadas una a una visualmente por el operador, siembra a mano (generalmente 2 semillas por golpe), raleo para dejar el stand adecuado de plantas, pasillos entre los bloques de ensayos, borduras en todo el perímetro, control de malezas si es necesario durante todo el periodo estival, y cosecha a mano.

Se puede coincidir que, para el caso de la Argentina, la descripción de Bradley *et al.*, del «pasado» de los métodos de evaluación en mejoramiento de maíz se ajusta, en general, a nuestro caso. Sin embargo, aún es posible observar en ciertos ambitos de nuestro país que «el pasado» de Bradley *et al.*, es la aproximación que se sigue utilizando como único sistema de selección o recomendación de materiales propios o de terceros. Esto no significa que el sistema sea intrínsecamente malo o técnicamente deficiente, sino que incorpora una cantidad de hechos de muy baja posibilidad de ocurrencia en la situación real, cuyas interrelaciones son además impredecibles. También, en el caso en que se realicen varias localidades, se suele observar una tendencia a no aleatorizar la primera repetición de los ECRs de cada localidad, práctica que no tiene en cuenta la influencia (positiva o negativa) entre parcelas que se incrementa cuando las diferencias entre altura de los participantes del ensayo son significativas (Clarke *et al.*, 1998).

Como resultado de estos tipos de evaluaciones, Bradley *et al.* sostienen que los híbridos que se destacan son aquellos de gran rendimiento en muy buenos ambientes. Todo tipo de problemas agronómicos (tales como la falta de espigas, caída de espigas, quebrado y vuelco, falta de tolerancia a diversos tipos de stress, baja germinación, bajo vigor, baja emergencia y hasta fallas de producción), son sistemáticamente obviados, ya que ellos generalmente aparecen asociados o son mucho más evidentes en los ambientes con estrés.

Otro elemento, ya mencionado, que se suma para confundir los resultados es el raleo. El raleo es una práctica habitual en muchos ECRs en los cuales se siembra un stand de plantas que puede llegar al doble del óptimo. Cuando la planta de maíz alcanzó una altura de alrededor de 20 cm, el operador elimina en cada parcela tantas plantas como sea necesario para llegar a la población teórica por hectárea deseada por el experimentador. El problema reside aquí en que indefectiblemente las plantas que se eliminan en la parcela son siempre aquellas de menor desarrollo relativo, cualquiera haya sido la causa que lo provocara. Esta simple acción enmascara todos los fenómenos relacionados con el vigor, poder germinativo y pureza genética de cada material, que no van a ser tenidos en cuenta en la evaluación final.

Ahora bien, suponiendo que se hayan realizado ECR de este tipo: ¿cómo se puede aseverar que un híbrido es mejor que otro?. Aquí, entra en juego el nivel de significancia para calcular las diferencias.

Segun Carmer y Walker (1988), al estimar las diferencias entre dos híbridos A y B, las posibilidades son solo tres:

DIF = (A - B), puede ser:

1. No hay diferencia
2. A es mejor que B
3. B es mejor que A

Para Carmer y Walker (1988) la estimación puede ser igual a la realidad o no, por lo que el juego queda reducido a nueve posibilidades:

Diferencia Observada	Diferencia Real		
	A es mejor que B	A es igual que B	B es mejor que A
A es mejor que B	Correcto	Error Tipo I	Error Tipo III
A es igual que B	Error Tipo II	Correcto	Error Tipo II
B es mejor que A	Error Tipo III	Error Tipo I	Correcto

El Error de Tipo I, es cuando se dice que el híbrido A es mejor que el B (o viceversa) y en realidad no lo es. Muy relacionado con éste, se encuentra el Error de Tipo II, que es cuando se dice que el híbrido A es igual al B y en realidad es diferente. El Error de Tipo III, es el denominado decisión reversa, y es cuando lo observado es directamente opuesto a la realidad. Los errores de Tipo I y III son los denominados «decisiones equivocadas».

Bajo este análisis, el mayor problema ocurre generalmente con los ECR que suman a la generalizada utilización de parcelas pequeñas, un número reducido de localidades o, a menudo, sólo una. En el mejor de los casos, el ECR viene acompañado del LSD (Least Significant Difference = Diferencia Mínima Significativa) u otro procedimiento de separación de medias. A un LSD de 0.05, se encuentran diferencias significativas basadas en la probabilidad de Error de Tipo I, pero esas diferencias se deberían aplicar

únicamente a separar los relativamente «malos» híbridos de los «buenos». Sin embargo, hay una cantidad de híbridos que se encuentran en el medio de los dos rangos y que pueden ser muy buenos. Ajustando el valor del LSD, se podría llegar hasta separar un solo híbrido del resto por su diferencia de LSD. En este caso la probabilidad de Error de Tipo I es baja, y la de Error de Tipo II desaparece. Pero en el momento en que se quiera hacer alguna estimación sobre el segundo híbrido (disminuyendo el coeficiente de LSD) se va encontrar seguramente con varios híbridos en el intervalo de LSD. En un caso así (que es el más común), existe la probabilidad real de tener un alto valor de Error de Tipo II, ya que es posible que uno de esos híbridos sea, en efecto, diferente (mejor o peor) que el resto.

Carmer y Walker (1988), sostienen que en general, en cualquier ECR donde se trata de identificar los mejores híbridos aparecen los Errores de Tipo II. A causa de las pocas (o sólo una) localidades analizadas, los resultados deben ser tomados muy cautelosamente y no se deben hacer análisis predictivos con los mismos.

El único recurso lógico según Stucker y Hicks (1991), es enfatizar la repetición de los experimentos entre localidades y años. Ya que con este procedimiento, se reduce el error standard de la diferencia entre el promedio de los híbridos, el poder analítico de este tipo de ensayos es significativamente mayor al que posee pocas (o sólo una) repeticiones. Bowman (1998) coincide con estas apreciaciones y considera que las más exactas estimaciones de la performance de los materiales se obtienen de datos provenientes de varias localidades, basado únicamente en el hecho de que el divisor para determinar el error estándar del promedio es igual al número total de observaciones.

La cuidadosa selección de los testigos que formarán parte del ECR es muy importante. La estadística, demuestra que el poder analítico de un ECR aumenta con el incremento de la magnitud de la diferencia entre los participantes. Sin embargo, para nuestro caso, esto es contraproducente. Stucker y Hicks (1991) dicen que el fitomejorador sólo debería elegir los mejores híbridos (adaptados a los ambientes a ensayar) como entradas. Como esto también es difícil (sino, ¿para qué se está haciendo el ECR?), el único recurso que le queda al fitomejorador es aumentar el número de localidades y años a ensayar.

El grado de predictibilidad de un ECR en resumen, es directamente proporcional al número de localidades ensayadas multiplicado por el número de años.

El Presente

En conocimiento de estos problemas Bradley y col (1988), dicen que la mayor parte de las empresas semilleras en EEUU ajustaron sus procedimientos de análisis a partir de la década del 80. En el presente los sistemas de evaluación reducen el énfasis de la precisión en una localidad dada para incrementar el énfasis de la precisión entre localidades. En otras palabras y según estos autores, se trata de poner el énfasis en la performance futura probable del material, más que en explicar su pasado.

El reconocimiento de que la aceptación de un material se basa en la performance estable del mismo en un rango de ambientes, llevó al desarrollo del concepto de la evaluación en áreas amplias, en donde muchos ambientes son testeados en un esfuerzo para predecir la performance en el rango de ambientes infinitos que tienen los productores.

El diseño experimental se modificó incluyendo pocas (generalmente 2) repeticiones en una localidad y muchas localidades. Los ECRGSs, son elaborados con un diseño de testigos control en los mismos, a fin de ajustar por las variaciones de fertilidad de los campos. El énfasis en más ambientes, incluye también el uso de diferentes fechas de siembra y densidades.

Bradley *et al.*, (1988) continúan diciendo que los ambientes son ahora elegidos como representantes de todas las posibilidades principales que deberán afrontar los materiales. Los ambientes con bajo rendimiento por cualquier situación de stress, son altamente deseables. Los ambientes con stress (aunque dan bajo

rendimiento, alto error experimental y altos Coeficientes de Variación), son ahora incorporados en los datos sumarios a causa de que ellos son una muestra de la población real de ambientes.

Una muy pequeña parte del presupuesto de investigación se vuelca hacia la disminución del error en una localidad individual; el presupuesto ahora se gasta en determinar la interacción genotipo-ambiente.

En un ejemplo, esto significa buscar campos promedio dentro de un área geográfica o ecológica dada, cuyos manejos, rotaciones, cultivos antecesores y nivel de malezas representen el ambiente que se quiere evaluar y no una excepción (por buena o mala) del mismo. Se aplicarán los herbicidas y la fertilización típica del productor medio de la zona y se sembrarán y cosecharán los ensayos en forma mecánica evitando en lo posible el raleo de plantas o hasta un 40% como máximo. La selección de semillas para los ensayos debe ser totalmente al azar, incorporando tanto semillas de tipo chato (centro de espiga) como redondas (punta y base de espiga). Los ensayos no deben tener más de dos repeticiones usando el diseño de bloques al azar, con borduras continuas únicamente en la periferia del ensayo y ancho de pasillos reducido al mínimo permitido por la maquinaria en uso.

Como resultado de estos cambios, Bradley *et al.*, (1988), dicen que los híbridos que ahora se seleccionan para llegar a nivel precomercial, son aquellos que muestran buena estabilidad entre ambientes junto con aceptables características agronómicas de tolerancia a vuelco y quebrado fundamentalmente. Los híbridos «nicho», tienden a desaparecer.

Paralelamente, según M.A Schmitt y S.J. Openshaw (1990), las compañías, en vistas de que los ECRs todavía no arrojaban datos confiables, desarrollaron los sistemas de ECRGSs. En general se utilizan dos sistemas de ECRGSs: el Strip Test, en el cual el número de parcelas es igual al número de participantes y el Check Plot, en el cual el número de parcelas es 1 más el número de participantes multiplicado por 2 (si el testigo check va cada un híbrido), o multiplicado por 1,5 (si el testigo check va en cada tercer lugar).

En Argentina, varias de las compañías involucradas en el negocio de semilla híbrida de maíz, comenzaron a desarrollar ECRGSs en la década del 80. El diseño más utilizado ha sido el de Check Plot. Algunas agencias de extensión del INTA y ciertos establecimientos de alto nivel tecnológico, evalúan los híbridos con el método de Check Plots.

El ECRGS es más caro y más difícil de hacer que un ECR individual y su evaluación es diferente. A pesar de todo lo que indica la lógica, los trabajos de Schmitt y Openshaw (1990) y los de Stucker y Hicks (1991), demuestran que el error experimental de los diseños de Check Plots es mayor que el de los diseños de Strip Test (sin testigos). Los resultados experimentales indican que el testigo en el Check Plot, no reduce consistentemente el error de la varianza.

Stucker y Hicks (1991), consideran el ejemplo de un Check Plot en el que cada híbrido está flanqueado por dos testigos checks. El rendimiento de cada híbrido, es corregido (ajustado) por la variación de fertilidad que se deduce del rendimiento de los testigos checks del flanco, dividido al promedio general del rendimiento de todos los testigos checks en ese lote. Sin embargo, la deducción estadística que subyace en el Check Plot indica que el valor de rendimiento ajustado del Híbrido 1 (H_1) es:

$$H_1 = H - bT_1$$

en donde H es el rendimiento verdadero del híbrido 1, T_1 es la covarianza o coeficiente de corrección del Híbrido 1 y b , es el coeficiente de regresión basado en la relación entre H_1 y H .

Lo importante aquí, dicen Stucker y Hicks (1991), es que el procedimiento de ajuste (pasar del rendimiento real del híbrido 1 al rendimiento ajustado por el testigo check) debe estar precedido por el cálculo de la eficiencia esperada del procedimiento de ajuste. Salvo que el procedimiento de ajuste demuestre que se incrementa la eficiencia del análisis, el ajuste nunca debe ser utilizado, y sólo se deben utilizar los datos reales del híbrido.

En realidad, nadie se toma el trabajo de calcular el coeficiente b , y se asume que este es igual a 1, lo

que pocas veces es así. Partiendo de la base que es necesario hacer varios Check Plots en ambientes distintos para sacar alguna conclusión útil, Stucker y Hicks (1991), dicen que el ajuste por check plots o testigos apareados, es un procedimiento inútil y, si se usan muchos ambientes, es absolutamente innecesario. Estos investigadores dicen que la industria semillera está usando un procedimiento que, en líneas generales, no es eficiente y que las organizaciones públicas o productores que no tienen acceso a información estadística competente deberían estar enteradas que están perdiendo tiempo y espacio con el uso de Check Plots en lugar de Strip Trials.

Bradley *et al.* (1988) sostienen que, en el presente, la información proveniente de ECR o ECRGSs individuales (una localidad) carece de importancia alguna tanto predictiva como descriptiva. Los convencionales análisis combinados entre localidades, clásicos de la década pasada, ahora son sinergizados con las comparaciones "cabeza a cabeza" (head to head), en las cuales dos híbridos son comparados en todos los ensayos en los que fueron sembrados juntos. Solo dos híbridos son considerados en cada análisis, por lo que el número de comparaciones aumenta considerablemente, lo que era la limitante principal del análisis combinado.

Stucker y Hicks (1991), demuestran que para los "cabeza a cabeza", el error standard de la diferencia entre el promedio de dos híbridos, es igual al cuadrado medio de la interacción genotipo ambiente dividido por e , mas la varianza genética + varianza ambiental dividido por e . Como en este caso, el termino e (número de ambientes) es muy grande, el procedimiento de análisis "cabeza a cabeza" da lugar a muy bajos LSD, lo que incrementa significativamente el poder analítico para diferenciar híbridos.

Stucker y Hicks (1991), están convencidos que muy pocos investigadores (y muchísimos menos extensionistas o productores) saben del poder de inferencia de las técnicas "cabeza a cabeza". Sin embargo agregan también, que hay mucha gente que no aprecia cuan sin sentido es la comparación "cabeza a cabeza", cuando sólo están disponibles datos de pocas localidades.

El Futuro

Bradley *et al.* (1988), dicen que los sistemas de testeo a partir de la década del 90, continuarán poniendo mas énfasis en la precisión entre localidades y años y poco énfasis en la precisión en una localidad dada. El énfasis máximo, será puesto para la predicción de la performance en el infinito número de ambientes que deberá afrontar un híbrido. El concepto de testeo en areas amplias, se extenderá para incluir el uso de todos los posibles ambientes definidos que puedan ser abarcados dentro del presupuesto de investigación. Las comparaciones entre híbridos serán hechas mediante simples diseños experimentales, tipo de bloques al azar, con una sola repetición. La única repetición dentro de una localidad, será para el caso de prueba de densidades, fechas de siembra, niveles de fertilidad, o cualquier otra variable ambiental o cultural.

Las localidades seleccionadas para los ensayos, se elegirán sobre la base de un análisis de los ambientes más probables que enfrentarán los híbridos. Los ambientes serán agrupados por ciertos elementos comunes y las localidades dentro de cada uno de los grupos de seleccionarán con el objeto de maximizar la correlación entre el presente y el futuro. El énfasis para elegir una localidad se modificará de considerar el rendimiento ambiental como indicador, a considerar la probabilidad de que esa localidad sirva como índice de futuro ambiente de desarrollo del híbrido.

Todos los datos serán incluidos en los análisis sumarios. Una muy pequeña parte del presupuesto de investigación será utilizada para controlar el error experimental en una localidad dada. La mayor parte del presupuesto, será utilizada para maximizar el valor predictivo de los datos tomados en varias localidades.

El concepto de comparación "head to head", se extenderá para incluir híbridos agrupados por estrechos márgenes de madurez relativa. Las series de comparaciones head to head serán arregladas para confeccio-

nar comparaciones head to group (series de comparaciones head to head de un híbrido contra un grupo de híbridos elite apropiados dentro de un rango de madurez).

Se desarrollarán hipótesis que ayuden a explicar biológicamente la interacción genotipo ambiente, basadas en mediciones de variables ambientales tales como la extensión del periodo de crecimiento, lluvias, temperatura, incidencia de enfermedades y fertilidad.

Stucker y Hicks (1991), tratando de hacer un análisis de lo que pueda ocurrir en el futuro, están convencidos que la mayor parte de los ensayos se harán con el modelo de bloques completamente aleatorizados, con sólo una repetición en cada nivel de ambientes y que los check plots (a fin de ajustar los rendimientos por variación de fertilidad) no se usarán mas y serán reemplazados por los strip trials.

Para ECRGSs, Stucker y Hicks (1991), y Schmitt y Openshaw (1990), coinciden en prever que el uso de grandes bloques completamente aleatorizados sera la alternativa a usar.

EL GRAN OBJETIVO: ESTABILIDAD DE RENDIMIENTOS

Por todo lo expuesto, se podría deducir que usando un buen diseño experimental, eligiendo bien los testigos a colocar en el ensayo, repitiéndolos en cuantas localidades sea posible de acuerdo a los ambientes a analizar, utilizando técnicas y procedimientos «reales» de preparación del terreno, siembra, cosecha, etc., y con un buen análisis estadístico, seríamos capaces de decir: el híbrido A es estable y rinde más que el testigo X en tales y cuales ambientes. Con todo, aún no estaríamos lo cerca que pensamos de la realidad.

Lin *et al.* (1986), consideran que los tres tipos de estabilidad paramétricos (hay otras técnicas no paramétricas), que más se usan son la estabilidad de Tipo 1, Tipo 2 y Tipo 3. La estabilidad de Tipo 1, la cual es análoga al concepto de homeóstasis, estadísticamente se basa en la determinación de la desviación a partir del efecto genotípico promedio. Aunque la estabilidad Tipo 1 es teóricamente lógica, los fitomejoradores la usan muy raramente. La razón principal, es que los fitomejoradores no solo buscan materiales con muy buena estabilidad Tipo 1, sino también que rindan bien. La estabilidad Tipo 1, está asociada con una respuesta relativamente pobre y bajos rendimientos en ambientes en donde otros materiales rinden bien.

La estabilidad de Tipo 2 (fundamentalmente el modelo de Finlay y Wilkinson, 1963), es una medida relativa que depende de los genotipos en el ensayo, de tal manera que su inferencia está confinada al grupo de híbridos analizados y no puede ser generalizada. Estadísticamente se basa en la determinación del coeficiente de regresión de cada genotipo, el cual es utilizado como parámetro de la estabilidad del mismo. A causa de que el promedio de todos los genotipos es usado para estimar la respuesta de cada ambiente, sacar inferencias a partir de este tipo de estabilidad requiere extremada cautela, salvo que los genotipos estudiados sean extremadamente representativos de los ambientes analizados. Un genotipo estable por esta definición, lo es sólo con relación a los otros genotipos del ensayo, y no se tiene ninguna seguridad de que se mantenga estable si se modifica el grupo de híbridos acompañantes. Por ejemplo, en un grupo de genotipos (A,B,C,D,E), «A» puede ser determinado estable y «B» inestable, si «A» se parece a «C,D,E» mas que lo que se parece «B». En un grupo de genotipos (A,B,F,G,H), «A» puede ser considerado inestable y «B» estable, si «A» es menos parecido a «F,G,H» que lo que se parece «B».

La estabilidad de Tipo 3, es un concepto un poco mas nuevo que los anteriores. Fue desarrollado por Eberhart y Russell (1966) y estadísticamente es un parámetro de desviación. En sí, es el residuo del cuadrado medio de la desviación a partir de la regresión definida por la ecuación de Finlay y Wilkinson (1963) de estabilidad Tipo 2. Aquí, la variabilidad de cada genotipo respecto al ambiente puede ser subdividida en una parte predecible correspondiente a la regresión y una parte impredecible correspondiente a la desviación del cuadrado medio. El modelo de Eberhart y Russell (1966), ha tenido una aceptación bastante grande. Sin embargo, como la variable independiente del modelo (el índice ambiental) no puede ser medida previa al experimento, la misma se convierte nada más que en datos que representan

el ambiente y a partir de los cuales la respuesta de los genotipos puede ser medida cuantitativamente. A causa de que el modelo es puramente empírico, la desviación del cuadrado medio no tiene propiedades determinísticas, como sería en el caso de un modelo predictivo.

Segun Lin *et al.*, (1986), el modelo de estabilidad de Tipo 3 es el menos atractivo. Los modelos Tipo 2, son útiles para comparar un grupo de genotipos pero, desde el momento en que el dato medido es relativo, el mismo no tiene una base inferencial suficientemente amplia para arribar a conclusiones generales. Los modelos Tipo 1, tienen una amplia base inferencial (no dependen de otros genotipos), pero no dan ninguna información de los patrones de respuesta de un genotipo en un rango de ambientes, que es justamente lo que se está buscando averiguar.

Gran parte de los fitomejoradores de maíz usan el modelo de estabilidad Tipo 2 de Finlay y Wilkinson (1963). El problema aquí es el siguiente: aun suponiendo que la selección de genotipos es representativa del (o de los) ambientes a analizar, la predicción del futuro comportamiento del híbrido sigue siendo muy especulativa.

Analizados todos los ensayos posibles y trazadas las rectas de regresión; ¿cómo se definen los ambientes de bajo rendimiento o los de alto rendimiento?. Tomando los de bajo rendimiento, el ambiente que causó esto puede haber tenido un suelo de baja fertilidad, o estar entoscado, o tener un horizonte A muy arcilloso, o tener un suelo con pendiente, o no llovió a siembra, o llovió mucho, o en floración no llovió, o hizo mucho calor, o hizo mucho calor en post-floración y sin lluvias, o granizo durante el llenado del grano, o etc., etc., o todo sumado. En definitiva, tanto el ambiente de bajo rendimiento como su opuesto, el de alto rendimiento, y todo el rango intermedio, constituyen un gran agujero negro sin definición precisa. Lo único que se sabe es el rendimiento promedio del grupo de híbridos del ensayo, cifra que al final caracteriza a un ambiente como malo, regular o bueno.

Sin poder discriminar, (el modelo no lo hace), cuál es o cuáles fueron las causas determinantes del potencial del ambiente, usar el modelo de Tipo 2 como predictivo no sería aconsejable. Lamentablemente no hay otra cosa a mano, pero debe quedar lo suficientemente claro que el modelo de Tipo 2 de Finlay y Wilkinson es básicamente un modelo descriptivo y no predictivo y que, utilizarlo para estos fines, conlleva un margen apreciable de inseguridad. Pese a ello, incrementando significativamente el número de localidades y genotipos totales en el análisis, el modelo de Finlay y Wilkinson podría adquirir fuerza predictiva.

El tratar de definir ambientes o localidades por atributos específicos y poder determinar cual o cuales híbridos se adapta a las mismas (es decir que tengan una interacción genotipo ambiente positiva), es el sueño de todo fitomejorador y la respuesta final que busca todo gerente de marketing. Como se vio, ninguno de los modelos vistos lleva a conclusiones definitivas. En la búsqueda de nuevos métodos de análisis, H.G. Gauch (1988) y varios colaboradores de la Universidad de Cornell, desarrollaron una nueva metodología para atacar el problema.

Zobel, *et al.*, (1988), explican que en los análisis de rendimiento de los ensayos, la suma de cuadrados total para rendimiento puede ser particionada en tres fuentes generales: el efecto principal del genotipo, el efecto principal del ambiente y la interacción genotipo ambiente (GA). Por definición, los efectos principales son aditivos y la interacción (el residuo del modelo aditivo), es no aditiva. Las tres fuentes son estadísticamente significativas y agronómicamente importantes.

Los autores dicen que el análisis de varianza, es un modelo aditivo y por lo tanto únicamente describe los efectos principales. El análisis de varianza puede determinar que existe una interacción GA, pero no arroja ninguna luz sobre las características de la misma. El análisis de componentes principales (PCA), es otro tipo muy importante de método. El PCA, es un modelo multiplicativo que tiene el problema opuesto al análisis de varianza: no describe los efectos aditivos principales. Consecuentemente la interacción, la cual es por definición el residuo del modelo aditivo, no es considerada y mucho menos analizada por este modelo.

Sobre la base de estas conclusiones, Gauch (1988) elaboró el modelo AMMI (Additive Main effects and Multiplicative Interaction), el cual incorpora tanto componentes aditivos como multiplicativos en un análisis integrado más poderoso. La presentación de este modelo en los círculos científicos produjo una gran expectativa. Aquí en la Argentina, durante el Congreso Nacional de Maíz de 1988 se invitó a uno de los colaboradores de Gauch, el Dr. Jorge Crossa (1990), a dictar conferencias sobre el tema.

Sin embargo, los resultados no han sido todo lo alentadores que era de esperar, en el sentido que el método ha demostrado su confiabilidad únicamente ante ambientes relativamente estables. Para el caso de Argentina, probablemente (o casi exclusivamente), el cíclicamente errático régimen de lluvias de su pampa húmeda, ha dificultado la clasificación de los ambientes bajo un ordenamiento someramente lógico. Por ejemplo, la localidad Chacabuco 1995/96 puede estar agrupada junto a Villa Cañas 1995/96 y tener una interacción GA positiva con el híbrido X, y en el mismo análisis la localidad Chacabuco 1997/98 puede estar junto a General Villegas 1995/96 y con interacción negativa para el mismo híbrido (resultados no publicados de uno de los autores-M.A.R.).

Para la situación de la Argentina, y toda otra región de tipo heterogéneo de crecimiento de cultivos, daría la impresión que será necesario contar con muchos más datos (localidades y años) para poder sacar alguna conclusión. Aunque el método AMMI sería relativamente más lento que los anteriores para inferir una correlación híbrido/ambiente, es muy posible que pueda ser de gran ayuda para agrupar localidades dentro de ambientes y así poder seleccionar con un criterio científico dónde realizar los ECRs y ECRGs. Esta, por ejemplo, ha sido la conclusión de Yau (1995) quien demostró que el método AMMI es en la práctica mucho más poderoso para capturar la interacción GE que los convencionales modelos de regresión lineal del tipo Finlay y Wilkinson.

En esta línea de trabajo, uno de los últimos desarrollos utilizando la técnica de AMMI, se realizó justamente con el objetivo de identificar "mega ambientes" homogéneos y subdividir una región de ambientes heterogéneos. (Gauch y Zobel, 1997). El término "mega ambiente" fue utilizado por primera vez por científicos del CIMMYT y se lo puede considerar sinónimo de región agroclimática o ecogeográfica. Estos autores entienden que la subdivisión de una región en "mega ambientes" es útil solo cuando la misma es relativamente estable, pero que los resultados preliminares demuestran la posibilidad de identificar un número razonable de "mega ambientes" de forma tal de sacar provecho de la interacción GE e identificar los genotipos útiles.

ALGUNOS OTROS PUNTOS DEL PROBLEMA

Aunque es difícil conocer el nivel de desarrollo que tiene el diseño experimental y análisis estadístico de los datos de rendimiento en las compañías semilleras, sin dudas, las dificultades señaladas en los puntos anteriores son comunes a todas las empresas e instituciones públicas dedicadas al mejoramiento de maíz.

Aquí sobresale un aspecto interesante: partiendo de la premisa de que los datos de investigación y desarrollo de todas las compañías semilleras son absolutamente veraces (no hay ninguna prueba de que no lo sean) y que el fitomejorador ha incorporado toda la información disponible sobre sus híbridos (y no solamente los datos más favorables para dichos materiales), ¿cómo se puede explicar que cada compañía presente información en la cual sus híbridos son mejores a los de la competencia?

Para contestar en parte a esta pregunta debemos tener bien en claro, que el fitomejorador al hacer un ECR o un ECRGS, lo que está tratando de hacer, en definitiva, es una imitación de la situación «real». Los ECR o el ECRGSs, son modelos de diseño estadístico que, llevados a cabo convenientemente, ayudan a inferir qué ocurrirá con un material cuando el mismo sea puesto en la situación «real». Citado por Stucker y Hicks (1991), este aspecto fue evaluado en el simposio de 1990 de la Sociedad Americana de Agronomía bajo el título «Planeamiento y conducción de la Investigación en el Mundo Real». El término «mundo real», implica que los tradicionales métodos de evaluación en pequeñas parcelas que usan las estaciones experimentales no son el «mundo real».

Cuanto más pueda el fitomejorador asemejar la situación de su ensayo al «mundo real», es posible que

su frecuencia de aciertos aumente correlativamente. Ahora bien, como cada compañía usa sus propios métodos de conducción de ensayos y obviamente selecciona los materiales que se destacan en sus evaluaciones, lo que en definitiva permitió el ensayo fue el identificar aquellos híbridos más aptos dentro del sistema de conducción empleado. Esto, no implica que esos mismos híbridos se destaquen cuando son puestos en el «mundo real». En términos más técnicos, la interacción genotipo-métodología puede ocurrir que sea más importante que la interacción genotipo-ambiente.

Una posibilidad interesante que podría ser tenida en cuenta es la de utilizar información adicional dentro de un mismo ensayo. Dicha información adicional provendría de datos correspondientes a otras variables relacionadas con el rendimiento, pero diferentes del mismo. Mediante alguna técnica de análisis multivariado (por ejemplo, análisis de variables canónicas) se podría descubrir la composición interna de variables nuevas que revelen cuáles son las verdaderas causas que hacen que dos o más híbridos se comporten diferente en una o más localidades.

Además, con esa información se podría diseñar un programa de selección -adicional al programa principal de selección por rendimiento- basado en los datos correspondientes a aquellas variables nuevas mencionadas en el párrafo anterior.

No hay compañía semillera del mundo que no haya cometido errores en el lanzamiento de algunos híbridos. La pregunta es entonces: ¿Esos materiales fallidos eran malos?. La respuesta es no. Lo que ocurrió fue que esos materiales se destacaron en la simulación del «mundo real» del ensayo, pero esta simulación por algún o varios motivos, fue mal realizada y/o interpretada.

Si no se ha parcializado el análisis dejando de lado aquellos datos “desfavorables”, es obvio asumir que toda la información que presentan las compañías semilleras sobre sus híbridos es verdadera. Lo que no se puede asegurar es que esos mismos resultados ocurran en el «mundo real», o sea cuando el híbrido es vendido y sembrado al nivel de productor en grandes superficies.

Stucker y Hicks (1991) dicen que, debido al rechazo que muchos productores tienen respecto a los resultados de ensayos de las compañías semilleras, algunos profesionales bien intencionados deciden «evaluar» los híbridos en sus propias evaluaciones. Según estos autores, el esfuerzo que estos consultores realizan es mucho más caro y en algunos casos no es más eficiente que poner los nombres de los híbridos en un sombrero y sacar al azar el 25% o el 50% de los mismos.

Si bien es posible discrepar con el argumento anterior, no menos cierto es que se debe reconocer que es muy común que cuando se dan a conocer datos de ensayos de rendimiento, no siempre se cuenta con la información completa de tareas culturales, cultivo antecesor, fecha de siembra, herbicidas y fertilizantes aplicados, diseño experimental utilizado, número de repeticiones, tamaño de la parcela, análisis de varianza, coeficiente de variación y LSD. Y aquí estamos hablando de un ensayo individual, que como ya se expuso, no tiene prácticamente ninguna utilidad de inferencia. Los ensayos combinados son muy raros. En el mejor de los casos, se realiza la combinación de ensayos realizados en la misma localidad.

Stucker y Hicks (1991), sostienen que el problema reside en que tanto los profesionales con su mejor intención, como los productores, están convencidos de que a partir de un ensayo individual puede ser inferido cuál es el mejor híbrido para una zona. El ensayo individual, siempre que haya sido conducido correctamente, sólo describe la situación pasada de una localidad, un año determinado y una fecha de siembra específica, y tiene un escaso poder de inferencia.

Para complicar más el problema, a menudo existe la tendencia de realizar este tipo de ensayos con el único objetivo -casi nunca explícito por otra parte- de demostrar hasta dónde llega el techo de rendimiento de los materiales comerciales, utilizando toda y cuanta tecnología está disponible con ese propósito. En definitiva, imitan una situación como la descrita en los métodos de selección del «Pasado» (antes de 1980).

Bowman (1998), analizando los resultados de ECR oficiales de maíz de Carolina del Norte encontró que las posibilidades de seleccionar los realmente mejores materiales de ciclo corto, intermedio y largo

sobre la base de un año de ensayos era de 55, 49 y 67% respectivamente y que esos porcentajes se incrementaban a 72, 51 y 75% para una selección de dos años de ECR combinados y a 79, 56 y 94% para una combinación de tres años de ECR combinados.

CONCLUSIONES

Los métodos de selección y estadística que utilizan los fitomejoradores en maíz, están poniendo cada vez más énfasis en el análisis de la respuesta de los materiales en áreas amplias, más que en situaciones individuales de cada localidad. El ensayo comparativo de rendimiento en microparcels (ECR), está cuestionado como única herramienta para la obtención de resultados confiables que permitan la selección de híbridos. El ensayo comparativo de rendimiento en grandes superficies o macroparcels (ECRGSs) es, por el momento, la más adecuada técnica de selección para corroborar los datos de los ensayos en microparcels. De tal forma, los programas de mejoramiento de las compañías semilleras, han incorporado el ensayo en grandes superficies o macroparcels como etapa imprescindible, previa al lanzamiento de un material y para el posterior desarrollo del mismo. La información de rendimientos basada en ensayos individuales en un año y una localidad no tiene capacidad de inferencia futura sobre la performance de los híbridos. Por el momento, sólo los ECR y ECRGSs repetidos sistemáticamente a través de varias localidades representativas de ambientes diferentes y durante varios años, y analizados por métodos estadísticos que permitan disminuir el error estándar a fin de maximizar el poder de separación de promedios, es el único procedimiento confiable para la selección y posterior recomendación de híbridos de maíz.

BIBLIOGRAFIA

- BOWMAN, D.T. 1998. Using crop performance data to select hybrids and varieties. *J. Prod.Agric.*, 11: 256-259.
- BRADLEY, J.P., K.H. KNITTLE and A.F. TROYER. 1988. Statistical methods in seed corn product selection. *J. Prod. Agric.*, 1: 34-38.
- CARMER, S.G. and W.M. WALKER. 1988. Significance from a statistician's viewpoint. *J. Prod. Agric.*, 1: 27-33.
- CARTER, P.R. and K.D. HUDELSON. 1992. University corn hybrid trials: Are results useful and reliable for growers?. *Proc. 22nd Soybean Seed Res. Conf.*, 22: 74-84.
- CLARKE, F.R., R.J. BARKER and R.M. DEPAUW. 1998. Interplot interference distorts yield estimates in spring wheat. *Crop Sci.*, 38: 62-66.
- CROSSA, J., H.G.GAUCH and R.W. ZOBEL. 1990. Additive main effects and multiplicative interaction analysis of two international maize cultivar trials. *Crop Sci.*, 30: 493-500.
- EBERHART, S.A. and W.A. RUSSELL. 1966. Stability parameters for comparing varieties. *Crop Sci.*, 6: 36-40.
- FINLAY, K.W. and G.N. WILKINSON. 1963. The analysis of adaptation in a plant breeding programme. *Aust. J. Agric. Res.*, 14: 742-754.
- GAUCH, H.G. 1988 Model selection and validation for yield trials with interaction. *Biometrics* 44: 705-715.
- GAUCH, H.G. and R.W.ZOBEL. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.*, 37: 311-326.
- KALTON, R. 1992. Public variety testing as viewed by private breeders. *Proc. 22nd Soybean Seed Res. Conf.* 22: 65-71.
- LIN, C.S., M.R. BINNS AND L.P. LEFKOVITCH. 1986. Stability Analysis: Where do we stand?. *Crop Sci.*, 26: 894-900.
- MISKIN, K.E. 1992. Value of university run state yield trials. *Proc. 22nd Soybean Seed Res. Conf.*, 22: 72-73.
- SCHMITT, M.A. and S.J. OPENSHAW. 1990. Comparison of experimental designs for variety cultivar strip trials. *Proc. 45th Annual Corn and Sorghum Res. Conf.*, 45: 54-63.
- STUCKER, R.E and D.R. HICKS. 1991. Experimental design and plot size considerations for on-farm research. *Proc. 46th Annual Corn and Sorghum Res. Conf.*, 46: 58-75.
- YAU, S.K. (1995). Regression and AMMI analyses of genotype x environment interactions: An empirical comparison. *Agron. J.*, 87: 121-126.
- ZOBEL, R.W., M.J.WRIGHT and H.G GAUCH. 1988. Statistical analysis of a yield trial. *Agron. J.*, 80: 388-393.