

***PARTE II***  
***Métodos para generar  
y analizar diversidad***



## II CAPÍTULO 1

### Polinización y fertilización *in vitro*

Susana Cardone, Gladys Pérez Camargo,  
Aurora Picca

#### 1 Introducción

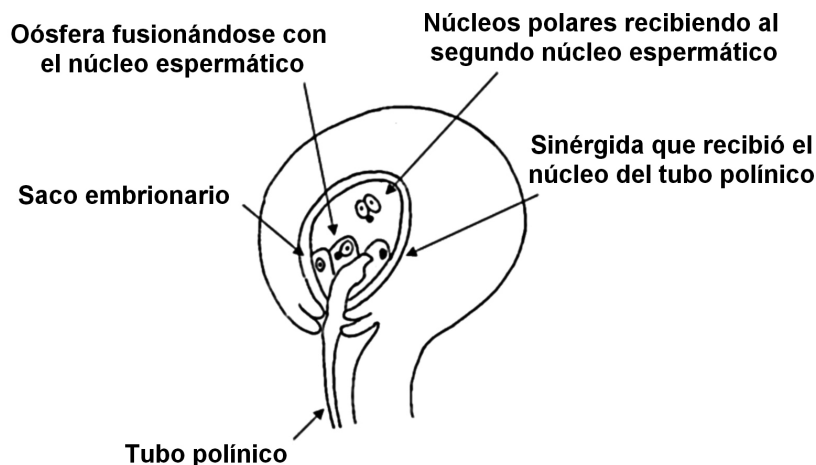
Durante la reproducción sexual de las Angiospermas los procesos de polinización y fecundación conducen a la unión de las gametas masculina y femenina para formar el embrión, que junto con el endosperma darán origen a la semilla, que generará una nueva planta. La fecundación contribuye a la diversidad genética y proporciona la base sobre la que se trabajará en fitomejoramiento.

El hecho más notable de la fertilización *in vivo* de las Angiospermas es el de la “doble fecundación”, que es un fenómeno muy complejo en comparación con lo que acontece en todos los demás seres vivos. Para que la polinización se produzca, durante el desarrollo *in vivo*, deben ocurrir una serie de interacciones y señales entre el grano de polen y los diferentes tejidos del pistilo. El grano de polen hace contacto con las células receptoras del pistilo

(células papilares del estigma) que permiten su adhesión, hidratación y germinación, extendiendo su intina y generando un tubo polínico, cuya función es transportar las células espermáticas hacia el óvulo. En ese momento se desorganiza el núcleo vegetativo mientras que la célula generativa dará lugar a las dos células espermáticas que se vuelcan en una de las sinérgidas, ésta se desorganiza y uno de los núcleos espermáticos se fusiona con la oósfera para dar el cigoto, que luego producirá el embrión. El otro núcleo masculino se reunirá con dos núcleos femeninos para dar la célula madre del endosperma que es, por lo tanto, triploide (**Fig 1**).

Durante la evolución de las plantas superiores, surgieron diferentes barreras de cruzamiento que controlan la transferencia de genes de una especie a otra o dentro de la misma especie.

A veces, el polen no puede germinar sobre el estigma de la flor aunque éste se halle receptivo. Este hecho puede deberse a limitaciones genéticas o fisiológicas. Estas barreras se denominan de **prefertilización o precigóticas** y pueden ser eludidas mediante las técnicas de **polinización *in vitro***.



**Figura 1.** Corte longitudinal de un óvulo mostrando la doble fecundación en Angiospermas

Otras veces, a pesar de la ocurrencia de la polinización, la fecundación no puede llevarse a cabo, por diversos motivos. El término “**fertilización *in vitro***” se utiliza para aquellos casos donde la fusión de las gametas aisladas se logra por distintos métodos como la electrofusión, altas concentraciones de calcio o elevado pH.

En este capítulo se utilizarán indistintamente los términos “fecundación” y “fertilización” como sinónimos, refiriéndose a la unión de las gametas masculina y femenina para originar un nuevo individuo.

En ocasiones, la fecundación ocurre normalmente pero el embrión aborta en forma precoz. Esto es lo que se conoce como **barreras posfertilización o poscigóticas**, que pueden contrarrestarse con técnicas de **rescate embrionario**.

Frente a todas estas limitaciones, las técnicas de cultivo *in vitro* y manipulación de células aisladas resultan apropiadas para lograr el proceso completo de desarrollo del embrión y del endosperma y la obtención de semillas normales. Una de esas técnicas, desarrollada por Kanta y colaboradores (1962), demostró que en algunas especies de Papaveráceas los granos de polen tenían la capacidad de germinar *in vitro*, directamente sobre los óvulos en cultivo. A partir de la misma, surgen una serie de metodologías que se conocen en general como técnicas de **polinización *in vitro***.

Las técnicas de micromanipulación desarrolladas durante los años 80 permitieron el aislamiento y la fusión *in vitro* de gametas de Angiospermas. La combinación de estas técnicas con métodos de cultivo de células únicas, posibilitaron la **fertilización *in vitro***. Se pudo entonces concretar el desarrollo de cigotos y de endospermas en forma individual, *in vitro*, demostrando que son capaces de autoorganizarse en cultivo, independientemente del tejido materno. Tanto los cigotos obtenidos *in vitro* como los que son aislados después de la polinización *in vivo* desarrollan en embriones normales y posteriormente en plantas fértiles.

Los sistemas de **polinización y fertilización *in vitro*** pretenden salvar las barreras pre y postfertilización para obtener plantas fértiles. Los recientes progresos relacionados con estas metodologías, conjuntamente con los

proyectos de genómica, posibilitarán una mejor comprensión de los procesos de reconocimiento gamético, fusión y activación de señales intracelulares que participan en los eventos tempranos de desarrollo del huevo y formación del cigoto. Estos conocimientos permitirán una aplicación más eficiente de estas técnicas en los programas de fitomejoramiento.

## 2 Polinización *in vitro*

En algunas especies, las **barreras de pre-fertilización** impiden la autofecundación o el cruzamiento, por diferentes mecanismos genéticos o bioquímicos. La incapacidad del polen de germinar en estigmas propios o extraños se denomina **incompatibilidad**. Esta puede deberse a varias razones: por ejemplo que el tubo polínico se deshaga en el estilo, o a la incapacidad del tubo polínico para alcanzar el óvulo debido a una lenta velocidad de crecimiento, o a una excesiva longitud del estilo. En estos casos, el ovario aborta antes de que el tubo polínico alcance la base del estilo.

Las barreras que obstaculizan los cruzamientos interespecíficos no son demasiado conocidas, aunque es evidente la existencia de interacciones en el proceso de reconocimiento polen-pistilo. Los mecanismos que originan la imposibilidad de los cruzamientos intraespecíficos son mucho más conocidos. Por ejemplo, el de la **autoincompatibilidad**, que involucra interacciones polen-pistilo que determinan la inhabilidad de una planta hermafrodita para producir cigotos, a través de la autopolinización. En este sistema, el intercambio de información entre el polen y el tubo polínico permite al estigma y al estilo discriminar el polen de plantas idénticas del proveniente de otros miembros de la misma especie.

Existen dos grandes sistemas de **autoincompatibilidad**, que se diferencian en sus estrategias moleculares y genéticas:

**a) autoincompatibilidad esporifítica**, en ella, el resultado de la polinización está determinado por el genotipo de la planta en la que se ha formado el grano de polen. En este tipo de incompatibilidad, presente en la familia de las Brassicáceas, proteínas de bajo peso molecular provenientes del polen, difunden desde la

exina e interactúan de manera alélica con un receptor quinasa, localizado en las células del estigma. Dicha interacción estimula la actividad fosforilante del receptor y se produce el bloqueo de la hidratación específica del grano de polen, impidiendo su germinación.

**b) autoincompatibilidad gametofítica** en la cual, la ocurrencia de la polinización está determinada por el genotipo haploide del polen. Bajo este sistema se encuentran dos mecanismos diferentes:

- Uno es típico de las Papaveráceas. En este, las señales de autoincompatibilidad se traducen por medio de proteínas, codificadas por un gen específico, que generan la movilización del calcio, fosforilando proteínas del polen que desencadenan la muerte celular del tubo polínico.
- En el otro mecanismo, presente en Solanáceas, un locus específico, el locus S, codifica para RNAsas que actúan sobre los tubos polínicos degradando al ARN del polen propio, inhibiendo su crecimiento.

Para sortear estas incompatibilidades surgieron técnicas de **polinización *in vitro***. La primera, desarrollada en *Papaver somniferum*, ha sido también empleada exitosamente en 14 familias de Angiospermas, resultando más adecuada su aplicación en aquellas especies que poseen ovarios grandes y que contienen muchos óvulos. Dentro de las especies de plantas en las que se abordó la técnica se pueden citar *Dianthus caryophyllus*, *Nicotiana tabacum*, *N. rustica*, *Argemone mexicana*, *Eschechlozia californica*, *Petunia violacea*, *Antirrhinum majus* y *Dicranostigma franchetianum*.

Bajo el nombre de “**polinización *in vitro***” se suele englobar a varias metodologías que surgieron a partir de modificaciones del procedimiento inicial, que si bien poseen puntos en común, difieren en otros. Entre estas metodologías se encuentran:

- **La polinización estigmática**, que consiste en cultivar el pistilo *in vitro* y colocar el polen sobre el estigma.

- **El injerto del estilo**, en la cual los granos de polen germinan en un estilo compatible y luego se los une al ovario de la planta madre que se desea fertilizar.
- **La polinización de óvulos aislados** que se cultivan directamente sobre el medio de cultivo. En esta técnica los óvulos son separados de su placenta; por esta causa es un procedimiento mucho más complicado, que ha sido empleado con éxito sólo en pocas especies de plantas.
- **La polinización placental**, técnica en la cual los óvulos se ponen en contacto con el polen por remoción de la pared del ovario o a través de un corte en la parte superior del mismo. En esta metodología los porcentajes de supervivencia son muy buenos, ya que los óvulos no resultan dañados durante las manipulaciones involucradas en su aislamiento. El cultivo del óvulo junto con la placenta incrementa, en general, las posibilidades de un rápido desarrollo embrionario.

La **polinización placental *in vitro*** se ha utilizado con éxito para salvar las barreras de incompatibilidad precigóticas (autoincompatibilidad e incompatibilidad interespecífica), en varias especies, inclusive cuando, experimentalmente, se intentó el cruzamiento a partir de óvulos de Angiospermas con polen proveniente de Gimnospermas. En este caso, al realizarse el análisis embriológico de los óvulos de la especie *Melandrium album*, fijados tres días después de la polinización con polen de *Pinus wallihiana*, se observó la presencia de remanentes de tubos polínicos y embriones bicelulares en los sacos embrionarios.

- **La polinización con polen mentor** Se ha utilizado, en muchas especies de plantas, una estrategia para salvar las barreras de incompatibilidad, involucrando mezclas de polen no compatible con polen compatible muerto (denominado “mentor”). El rol del polen mentor es proveer alguna función, por ejemplo, a través del aporte de sustancias proteicas,

las que facilitan la polinización del polen incompatible. Algunas de las especies en las que se ha utilizado son: *Populus*, *Malus*, *Petunia* y *Nicotiana*.

Finalmente, en todos los casos, la gameta femenina es fecundada por células espermáticas, liberadas por el tubo polínico, en forma "casi" idéntica a la vía natural, con la consiguiente formación de semilla normal en un medio de cultivo apropiado. En general, todos los medios de cultivo para germinación de los granos de polen poseen altas concentraciones de sacarosa y de ácido bórico.

No obstante, para que los procedimientos anteriormente citados sean exitosos es crucial realizar el aislamiento de los óvulos y del polen en el estado fisiológico adecuado, por lo que es indispensable llevar a cabo la determinación de la duración de los distintos estadios del desarrollo de ambos, como por ejemplo:

**a)** Determinar los tiempos de anthesis, dehiscencia de las anteras y polinización, **b)** Precisar el momento en que el estigma esté receptivo, **c)** Especificar el instante oportuno para la polinización y **d)** Establecer el tiempo adecuado para recoger el polen.

### 3 Fertilización *in vitro*

Durante años los científicos intentaron imitar los mecanismos de fecundación que acontecen *in vivo*, utilizando técnicas de **fertilización *in vitro*** de gametas aisladas; esta aplicación fue más sencilla en animales y en plantas inferiores pero no pudieron ser aplicadas a las plantas superiores hasta comienzos de la década de los noventa. Esto se debió a que la fertilización *in vivo* en las plantas superiores es un proceso muy diferente al que ocurre en los sistemas animales y de plantas inferiores, donde las gametas son de vida libre.

El primer aislamiento de células espermáticas vivas fue reportado por Cass en 1973, en el cuál se aislaron gametos masculinos de cebada por ruptura de granos de polen en una solución de sacarosa al 20%. A partir de este aislamiento se pudieron describir las características celulares de ese gameto. Esto se

repitió luego con éxito en varias especies de Angiospermas. Para el aislamiento de gametas femeninas existe un método que se emplea a menudo y que consiste en un tratamiento con enzimas diluídas para disolver las paredes celulares, acompañado de la micromanipulación para separar las otras células del saco embrionario.

En 1991, Kranz y col. reportaron la primera fusión de gametas de maíz *in vitro*, pudiendo reproducirse actualmente en esta especie el ciclo de vida completo, comenzando con la electrofusión de la célula espermática con la ovocélula, dando como resultado un cigoto, un embrión y finalmente una planta. Pese a que se ha tomado al maíz como sistema vegetal modelo, en el cual estudiar los procesos de fusión de gametas y posterior embriogénesis, también se han realizado trabajos exitosos de fusión de gametas en trigo, donde Kovacs (1995) logró la fusión de gametas *in vitro*, obteniendo como resultado estructuras multicelulares y en algunos casos del tipo de microcallos. En los últimos años se ha logrado finalmente la fusión de gametas en *Nicotiana tabacum*, la que pudiendo ser una especie modelo en dicotiledóneas para estudiar la fertilización *in vitro*, dada su capacidad para la regeneración de plantas, presenta sin embargo el inconveniente de que las gametas no se fusionan fácilmente.

Tradicionalmente, se consideró que todas las células espermáticas tenían igual posibilidad de fusionarse con la célula huevo. Sin embargo, actualmente se sabe que existen mecanismos de fusión preferencial a nivel gamético, donde un determinado morfotipo de célula espermática es reconocido. Cuando está presente este mecanismo hay una identificación de las dos células espermáticas, candidatas para la fusión celular y una elección de la que se fusionará con la ovocélula. La discriminación se basa en las diferencias de las células espermáticas respecto de su heteromorfismo citoplasmático en: tamaño, forma, volumen, superficie, contenido de organelas y asociación física con el núcleo vegetativo. Por otro lado, el contacto del polen con el gameto femenino parece estimular la maduración del mismo, incluyendo la acumulación de calcio necesaria, la organización de la célula huevo, la migración nuclear,

y la progresión del ciclo celular. Este último factor parecería ser crítico en la combinación exitosa de gametas. Además, hay señales, suministradas por el saco embrionario, que permiten estimular el crecimiento y la maduración del tubo polínico.

Para que la fertilización tenga lugar, *in vitro*, no sólo deben aislarse las gametas femeninas y masculinas sino que deben ser inducidas a fusionarse, en ausencia de las células asociadas.

Para obtener un embrión *in vitro*, los pasos a seguir son los siguientes:

### Aislamiento y selección de las gametas

Se han desarrollado métodos de aislamiento y selección de gametas para una amplia variedad de especies vegetales, entre ellas: maíz (*Zea mays*), trigo (*Triticum aestivum*), cebada (*Hordeum vulgare*), raygrass (*Lolium perenne*) y algunas brassicáceas. A fin de realizar eficientemente este paso se debe:

a) Conocer exhaustivamente la morfología floral de la especie a utilizar, con el objetivo de ubicar el saco embrionario y sus células constituyentes. b) Aislar las gametas masculinas, a partir de los granos de polen mediante choque eléctrico, osmótico y/o de pH. c) Aislar el saco embrionario y las gametas femeninas mediante microdissección individual en asociación con maceración enzimática.

Este paso es crítico y limitante ya que pueden obtenerse pocas gametas femeninas y el proceso de manipulación es muy delicado. El tratamiento de las espigas, previo a la polinización con 2,4-D, que induce a la expansión del ovario, facilita el proceso resultando en ovarios más grandes y óvulos más blandos, incrementando la eficiencia en el aislamiento de las ovocélulas.

### Fertilización *in vitro* propiamente dicha

Puede realizarse mediante la microinyección de células espermáticas o núcleos espermáticos aislados, en células individuales obtenidas de los sacos embrionarios, o por electrofusión. En este último caso las gametas se ponen en contacto en una cámara de electrofusión por alineación dielectroforética y se fusionan mediante pulsos eléctricos (Fig. 2). Los productos de fusión pueden ser entonces cultivados en

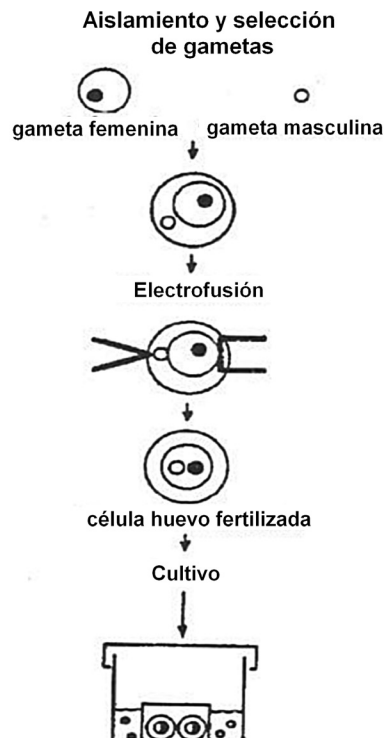


Figura 2. Fertilización *in vitro* mediante electrofusión

una solución con células nodrizas y algunos de ellos desarrollarán en proembriones.

**Fusión de las gametas.** En maíz, la secuencia de eventos para que se verifique la fusión *in vitro* es la siguiente: las gametas masculinas se obtienen a partir de granos de polen fresco sometidos a un "shock" de pH en una solución 0,5 M de manitol. Las ovocélulas y los protoplastos de las células centrales del saco embrionario se aíslan de los óvulos por tratamiento con enzimas celulolíticas y microdissección manual.

Las dos gametas aisladas se colocan de a pares, bajo condiciones de esterilidad, en un medio de fusión simple compuesto por manitol y cloruro de calcio (CaCl<sub>2</sub>). Para realizar este procedimiento se utilizan microagujas de vidrio que permiten poner en contacto las gametas femeninas y masculinas, a fin de facilitar y dirigir la fusión. La adhesión y posterior fusión de las gametas es rápida (aproximadamente 10 segundos). A los 20 minutos de realizada la fusión, deja de visualizarse el núcleo masculino y 20 horas después es posible observar mediante microscopio óptico, con contraste de

fases la presencia de dos núcleos, de manera similar a lo que se observa *in vivo*.

La formación del endosperma ocurre cuando se fusiona el núcleo de la célula espermática con los dos núcleos polares. Los estudios *in vitro*, sobre los primeros eventos después de la fusión de las células y núcleos, indican que las primeras células del endosperma resultante desarrollan un tejido característico capaz de auto-organizarse en forma separada del tejido materno. Este proceso se completa, en esta especie, aproximadamente dos horas después de la fusión *in vitro* de las células que forman el embrión.

#### **4 Cultivo de óvulos y embriones *in vitro***

El **cultivo de óvulos** es un procedimiento muy complejo debido a que la manipulación del material es difícil. El óvulo es una estructura delicada, muy pequeña, altamente hidratada, que puede ser dañada con suma facilidad. Es imprescindible realizar la microcirugía con rapidez para evitar que los tejidos sufran desecación durante el proceso. Pese a las dificultades de la técnica, en algunos cruzamientos interespecíficos como los realizados en papa, algodón y *Hevea brasiliensis*, se comprobó que el cultivo de óvulos completos es más exitoso para el rescate de embriones que el cultivo de los embriones aislados.

El **rescate de embriones** consiste en aislar los embriones de los óvulos de una planta y cultivarlos en un medio estéril que contenga los nutrientes esenciales que les permita concluir su desarrollo normal. Esta técnica es relativamente fácil de llevar a cabo en embriones maduros y sus posibilidades de éxito son muy buenas. Las dificultades se incrementan cuanto más inmaduro sea el embrión a rescatar, ya que los requerimientos nutricionales son más específicos en las etapas tempranas de desarrollo del embrión.

#### **Factores externos que condicionan el cultivo de óvulos y embriones**

Entre los factores externos más importantes que afectan el cultivo de óvulos y de embriones se citan: el medio de cultivo, la osmolaridad y los reguladores de crecimiento.

El óvulo presenta una gran cantidad de requerimientos nutricionales, que varían con el estadio de desarrollo del mismo y que es necesario identificar para cada especie en particular. Por consiguiente un aspecto fundamental a tener en cuenta en el rescate de embriones es la selección correcta del medio de cultivo adecuado para cumplir con los requerimientos durante las distintas etapas del desarrollo embrionario, siendo necesario a veces la transferencia de los embriones de un medio de cultivo a otro para garantizar un óptimo crecimiento.

Se pueden reconocer dos fases en el desarrollo embrionario con respecto a la nutrición:

- **la fase heterotrófica** en la que el embrión es dependiente del endosperma y demás tejidos maternos que lo rodean, y
- **la fase autotrófica**, en la cual el embrión es capaz de sintetizar las sustancias requeridas para su crecimiento a partir de sales minerales y azúcares. La etapa crítica de pasaje de una fase a otra varía con las especies.

La introducción de agua de coco y extractos de endospermas en el medio de cultivo han sido muy útiles ya que activan el crecimiento y desarrollo de los embriones en el cultivo de óvulos de algunas especies vegetales. Por otro lado, no existe demasiada evidencia acerca de la absoluta necesidad de los reguladores de crecimiento para el desarrollo de los embriones extraídos, sean inmaduros o maduros, excepto en el rol que cumplen en la ruptura de la dormición.

La concentración osmótica es un factor importante en el desarrollo de óvulos y embriones, dependiendo del estado de madurez de los mismos. La sacarosa presente en los medios de cultivo no sólo provee la fuente de carbono necesaria para el crecimiento sino que además mantiene la osmolaridad requerida para cada etapa del desarrollo. Está demostrado que una presión osmótica elevada previene la germinación precoz de los embriones inmaduros, evitando así la ocurrencia de malformaciones estructurales. A medida que van creciendo, los embriones necesitan ser transferidos a medios con progresivas disminuciones en los niveles de sacarosa.



Respecto de los requerimientos de temperatura, los embriones de la mayoría de las especies presentan un buen crecimiento en un rango de temperaturas que va de los 25° C a los 30° C. Aunque algunos estudios indican que la luz no es crítica para el crecimiento de los embriones, se demostró que en el caso de la cebada la luz disminuye las probabilidades de germinación precoz en los embriones inmaduros. Entre los medios de cultivo más difundidos para el cultivo de óvulos y embriones se pueden citar el de Murashige y Skoog (1962), Monnier (1976), Randolph y Cox (1943), Gamborg y col. (1976), Liu y col. (1993).

## 5 Aplicaciones de las Técnicas de Fertilización *in vitro*

### 5.1 Aplicaciones en el Mejoramiento Genético Vegetal

Las técnicas antes mencionadas han sido utilizadas para salvar algunos obstáculos que se le presentan al mejorador de plantas cuando desea realizar cruzamientos, para el logro de ciertos objetivos de mejoramiento. A continuación se mencionan ejemplos de aplicación de las metodologías de polinización y fertilización *in vitro*:

En algunas combinaciones de cruzamientos, la **polinización placental** *in vitro* permitió que se superen las barreras de incompatibilidad precigótica, obteniéndose embriones híbridos y aún plantas híbridas, imposibles de obtener mediante técnicas convencionales. Así se realizaron los cruzamientos de las siguientes especies: *Zea mays* x *Z. mexicana*, *Nicotiana tabacum* x *N. amplexicaulis*, *Melandrium album* x *Viscaria vulgaris*, *M. Album* x *Silene schafta*.

La **polinización *in vitro*** se empleó también como vía para la obtención de plantas resistentes a estreses ambientales, ya que permite seleccionar para la fecundación aquellos granos de polen que tuvieron capacidad de germinar bajo diferentes condiciones de estrés, acelerando así la obtención de plantas resistentes. Esta estrategia se desarrolló en maíz, con polen cosechado de plantas creciendo a temperaturas elevadas y condujo a la obtención de plantas tolerantes a estrés por calor. En *Brassica rapa*, la selección de polen a baja tempe-

ratura tuvo un efecto positivo en el crecimiento de la progenie en condiciones de frío, logrando acumular más peso seco que en progenies obtenidas sin la selección previa del polen.

El **cultivo de óvulos y de embriones** vegetales se emplea en mejoramiento vegetal cuando se desea: rescatar embriones abortivos derivados de la hibridación interespecífica o intergenérica, superar problemas de abscisión precoz del fruto, recuperar embriones de cruzas interespecíficas que conducen a la obtención de haploides y acortar el período de latencia que ocurre en algunas semillas por presentarse, en el endosperma o en la cubierta de la misma, inhibidores del desarrollo del embrión.

Los siguientes ejemplos ilustran las diversas aplicaciones:

Durante la producción comercial del triticale, especie obtenida artificialmente mediante el cruzamiento de trigo y centeno, el **cultivo o rescate de embriones** se utilizó para superar el alto grado de incompatibilidad del cruzamiento inicial, ya que los embriones abortaban en estadios tempranos del desarrollo debido a una incompatibilidad con el endosperma. Esta técnica y la aplicación de colchicina en plantas híbridas, fueron de fundamental importancia para salvar los problemas que se presentaron en el cruzamiento inicial y la marcada esterilidad que poseía la F1. El perfeccionamiento de estas dos metodologías fue decisiva para la producción de un gran número de triticales hexaploides y octoploides, cruzando el centeno con trigos duros y harineros, respectivamente, logrando finalmente un alto grado de fertilidad.

Las técnicas de **rescate de embrionario** son necesarias como complemento en la producción artificial de haploides. En ellas, se cultivan embriones que se forman por partenogénesis o por eliminación de cromosomas luego de la hibridación interespecífica o intergenérica. En algunos casos los embriones haploides se distinguen morfológicamente de los normales, de manera que se extraen y se los cultiva *in vitro*, en medios y condiciones adecuados (ver en este libro Parte IV-Capítulo 1). Otras veces es necesario esperar a obtener las plantas para determinar el nivel de ploidía. Esta técnica se ha utilizado con éxito en cebada, trigo y tabaco.

También se emplea el **cultivo de embriones** en casos en que el endosperma no se desarrolla normalmente, como sucede en los cruzamientos de cultivares diploides y tetraploides de *Citrus*, o cuando el mismo se desintegra después de la fecundación, como en *Oenothera biennis* x *O. muricata* o *O. biennis* x *O. lamarckiana*. Esta técnica también es de utilidad para sortear barreras de dormición o en casos donde la viabilidad de la semilla es baja, como en yuca y otras especies del género *Manihot*.

El **cultivo *in vitro* de embriones** es también una alternativa para el intercambio de semilla a través de fronteras internacionales. En un proyecto de mapeo de genes con resistencia al virus del mosaico, en yuca, desarrollado entre CIAT (Centro Internacional de Agricultura Tropical, con sede en Colombia) y IITA (Instituto Internacional de Agricultura Tropical, con sede en Nigeria) fue necesario trasladar "stocks" de semillas desde Sudamérica al continente africano, el cultivo de embriones resultó ser un excelente método para transferir germoplasma con mínimos riesgos fitosanitarios.

El **cultivo de embriones** ha ayudado también al mejoramiento genético de especies leñosas que tienen dificultad en la germinación de sus semillas, como en *Taxus mairei*, o con escasa producción de semillas como es el caso de *Pseudotsuga macrolepis*.

El **cultivo de embriones inmaduros** puede ser aplicado para disminuir el tiempo necesario en la obtención de un nuevo cultivar. Por ejemplo en soja se requieren unos diez años para este fin. Si se consiguen realizar varias generaciones por año, en contraestación, se acelera el proceso. El cultivo de embriones inmaduros ahorraría, además, el tiempo de maduración de la semilla. En este y en cada caso en particular, se debe realizar la evaluación del beneficio económico para decidir la utilización de estas técnicas.

**Cultivo de nucelas** En las plantas leñosas, los explantos de tejidos maduros pierden la capacidad de regeneración, se ha inducido entonces con éxito la **embriogénesis somática**

**a partir de nucelas** de óvulos jóvenes. Este es el caso de *Citrus sp*, *Vitis vinifera*, *Malus domestica*, *Psidium guajava* y *Pyrus serotina*.

La técnica de **rescate de óvulos** fecundados se utiliza también en aquellos cruzamientos en los que la barrera de incompatibilidad se encuentra a nivel del ovario como en el cruzamiento *Trifolium repens* x *T. hybridum*.

Las técnicas de **fertilización *in vitro*** pueden complementar, además, las metodologías de transformación de plantas. Por ejemplo, a través de la obtención de un embrión a partir de cigotos transgénicos que han sido manipulados, utilizando un protocolo que involucra la microinyección de genes en el cigoto. Esta técnica evitaría la utilización de antibióticos y genes de resistencia a herbicidas, como marcadores, durante el proceso de selección de las células transformadas.

Otra ventaja de la **fertilización *in vitro*** es la posibilidad de realizar cruzamientos entre plantas que son filogenéticamente distantes, fusionando *in vitro* sus respectivas gametas para producir cigotos híbridos. Se evitan así los inconvenientes que frecuentemente se observan en la hibridación somática, como por ejemplo la ocurrencia de inestabilidad del híbrido respecto del número de cromosomas o las dificultades núcleo e inter-citoplasmáticas (ya que en este caso es la célula huevo la que contribuye con el citoplasma del cigoto híbrido. Además, en la fusión gamética, el desarrollo temprano del cigoto está regulado por genes pertenecientes a la célula huevo, los que brindan señales celulares propias, haciendo que progrese el ciclo celular en forma exitosa con el resultado de proembriones producto de las sucesivas divisiones celulares. Pese a la potencialidad de la técnica de hibridación distante a partir de gametos, es importante la elección de especies parentales adecuadas para minimizar las dificultades que surgen producto de las diferencias filogenéticas, fisiológicas y del ciclo celular.

## **5.2 Aplicaciones para dilucidar aspectos básicos del desarrollo en plantas**

Las dificultades en el aislamiento de los ga-

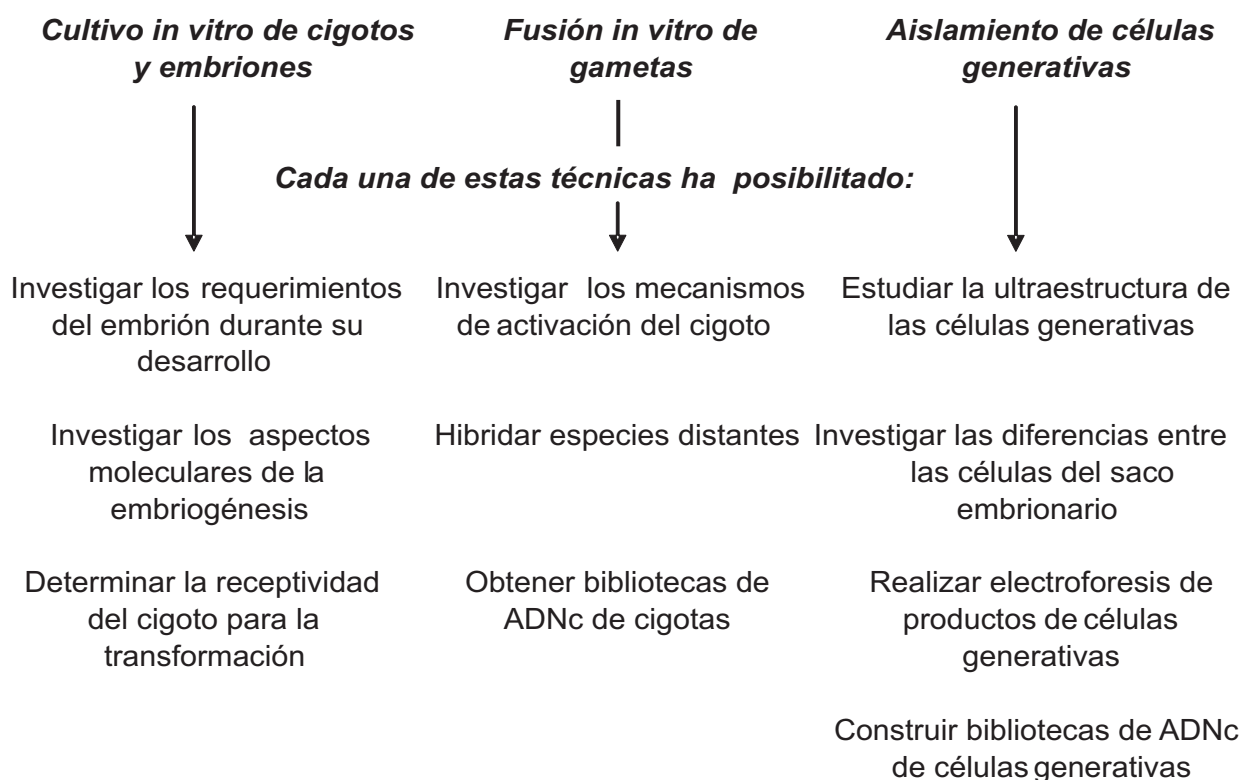
metos de las plantas superiores ha obstaculizado durante mucho tiempo la comprensión de la fisiología gamética y de los mecanismos que desencadenan la embriogénesis; los embriones cigóticos en su medio natural ofrecen dificultades debido a su tamaño pequeño y al hecho de estar inmersos en el tejido materno.

A partir del aislamiento de células espermáticas, células huevo y cigotas, se han podido realizar estudios moleculares de las células germinales, evitando las interferencias que resultan de la interacción con el tejido somático. Asimismo se ha facilitado el análisis de los efectos genéticos, epigenéticos e interacciones núcleo-citoplasmáticas inherentes a los gametos como también el estudio del desarrollo del embrión y el endosperma. De este modo, la **fertilización *in vitro*** de las plantas superiores se ha convertido en un área de investigación importante, resultando actualmente uno de los campos principales de la biología vegetal.

Los estudios sobre la **embriogénesis somática** también han permitido identificar patrones de expresión génica relacionados con este proceso, que se desencadena a través del cultivo de tejidos. El sistema mejor estudiado, en relación con la embriogénesis somática, es el de suspensiones celulares de zanahoria (*Daucus carota*). En este sistema, la remoción del 2,4-D (ácido 2, 4-diclorofenoxiacético) del medio de cultivo, genera una respuesta embriogénica por la cual los agregados celulares forman embriones que crecen, de a millares, en un medio líquido (ver en este libro Parte II-Capítulo 1). Los embriones somáticos constituyen una excelente herramienta para estudiar los procesos de desarrollo en plantas, ya que facilitan el acceso a gran cantidad de embriones libres.

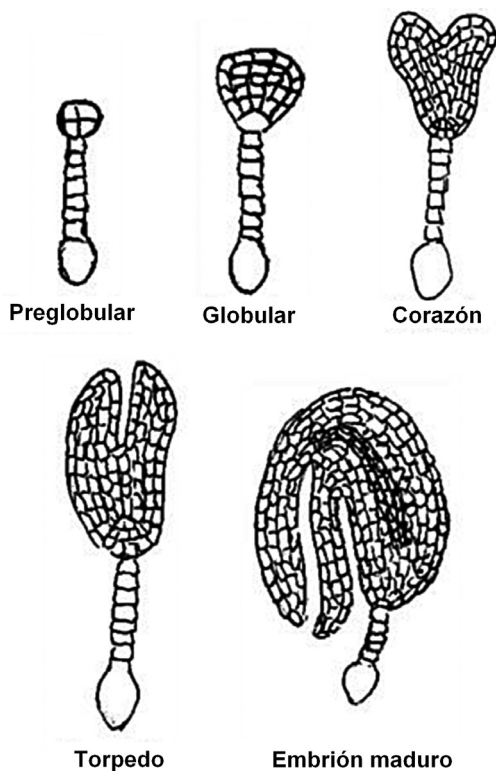
Las plantas parecen poseer un programa embriogénico preestablecido que se desencadena en respuesta a condiciones específicas. El proceso de fertilización dispara la em-

**Tabla 1.** Resumen de los avances metodológicos posibilitados por las técnicas de **fertilización *in vitro*** para el estudio del proceso de la doble fecundación de las Angiospermas (Modificado de Wang y col., 2006).



biogénesis cigótica, cuyos estadios normales de desarrollo se muestran en la **Fig. 3**. Una serie de investigaciones han demostrado que hay requerimientos para que los gametos se fusionen *in vivo*, por ejemplo la presencia de calcio, elemento que posee una actividad altamente fusogénica y que además activaría el desarrollo del cigoto. Además, al estudiar la ultraestructura celular se han encontrado diferencias en la arquitectura de los microtúbulos entre la célula huevo no fertilizada y el cigoto, demostrando que este cambio está inducido por la fertilización.

El cultivo *in vitro* y las tecnologías diseñadas para estudiar la expresión génica han permitido analizar por separado la función de genes presentes de manera específica en las células que participan en el proceso de fertilización. Con la utilización de técnicas de especiales de electroforesis bidimensional se han podido identificar proteínas de células gaméticas y cigóticas, encontrándose pocas diferencias entre los patrones proteicos de embriones somáticos y cigóticos. Ning y colaboradores (2006)



**Figura 3.** Estadios de la embriogénesis en *Arabidopsis thaliana*

estudiaron la contribución de los genes de las células espermáticas, la célula huevo y la cigota y obtuvieron evidencias de que hay clusters de genes que están presentes en la célula huevo y espermáticas, que persisten en el cigoto, mientras que hay una gran cantidad de genes que son propios del cigoto, produciendo en éste nuevos transcritos.

En el caso particular de la gameta masculina, se sabe que cada una de las funciones para las cuales está destinada (reconocimiento, fusión y adhesión con la gameta femenina), está controlada por la activación de genes. Por ejemplo, Xu y colaboradores (1999) identificaron y caracterizaron dos genes de histonas gCH2A y gCH3 que actúan durante la maduración de la célula espermática y también el gen LGC1 que se expresa exclusivamente en las células gaméticas masculinas. El producto de este último gen (una proteína) fue localizado en la superficie de las células gaméticas masculinas sugiriendo un posible rol en la interacción célula espermática-célula huevo, en el momento del reconocimiento o señalización. Además, este mismo grupo de investigación aisló, a partir de una biblioteca de ADNc, dos genes, NtS1 y NtS2, que codifican para una poligalacturonasa, enzima que tendría la función de modificar la pared celular durante la diferenciación de las células espermáticas.

Pese a las dificultades que presenta el aislamiento de las gametas femeninas (hay menor eficiencia en la obtención de células que en las masculinas), también se han logrado progresos en el esclarecimiento de sus funciones. Por ejemplo, recientes investigaciones han esclarecido detalles acerca de la expresión de genes involucrados en el gameto femenino que determinan la guía del tubo polínico y la maduración de las anteras, como por ejemplo el gen ZmEA1, en maíz, que posee una función en la atracción del tubo polínico a la micrópila. Cuando este gen está regulado negativamente, la señal direccional del tubo polínico a la micrópila es insuficiente y no se produce la fertilización.

En relación a la expresión de genes en la embriogénesis del maíz, se han identificado algunos diferencialmente expresados en la célula basal y apical del embrión bicelular. En esta

misma especie la célula huevo mostró un incremento en la abundancia de transcriptos que involucran funciones como la comunicación celular, el crecimiento y la división celular, la síntesis de ADN y factores relacionados con estreses y transducción de señales. Por otro lado, las células centrales del saco embrionario muestran una abundancia de transcriptos involucrados con los procesos relacionados con el metabolismo de la energía y la síntesis proteica. En cada caso, estos productos parecerían reflejar categorías funcionales representadas en el futuro desarrollo del embrión y el endosperma. Como en el caso ya citado de la célula espermática, se han identificado también en la célula huevo y la cigota genes de histonas que podrían representar un mecanismo regulatorio de varios genes.

El número de herramientas disponibles en la actualidad para manipular las gametas de las plantas superiores provee numerosas oportunidades para realizar progresos científicos y tecnológicos. La investigación en el área de la biología reproductiva de las Angiospermas ha entrado en una nueva fase, donde los métodos de la biología molecular permiten responder muchas preguntas acerca de los mecanismos fundamentales de la fecundación y activación del desarrollo embrionario temprano. Un resumen de **las técnicas de fertilización *in vitro*** que han permitido el acceso a nuevas investigaciones en el área molecular de este tema, se muestran en la **Tabla 1**.

El desarrollo de sistemas experimentales en varias especies, particularmente en dicotiledóneas, son necesarios para confirmar los resultados obtenidos en gramíneas, a fin de complementar la información molecular acerca del proceso de fertilización y mejorar los métodos para la producción de plantas regeneradas fértiles, con combinaciones genéticas estables y de alta eficiencia.

## 6 Lecturas recomendadas

Cass D. 1973. An ultrastructural and Nomarski-interference study of the sperms of barley. *Can J Bot* 51:601-605.  
 Chen S.; Yang, Y.; Liao, J.; Kuang, A.; Tian, H. 2008. Isolation of egg cells and zygotes of *Torenia fournieri* L. and determination of their surface

charge. *Zygote* 16:179-186 Cambridge University Press.  
 Dumas C.; Morgensen, H. 1993. Gametes and fertilization: Maize as a model system for experimental embryogenesis in flowering plants. *Plant Cell* 5: 1337-1348.  
 Faure J.; Digonnet, C.; Dumas, C. 1994. An *in vitro* System for Adhesion and Fusion of Maize Gametes. *Science* Vol 263, pp 1598-1600.  
 Faure J.; Aldon, D.; Rougier, M.; Dumas, C. 1996. Emerging data on pollen tube growth and fertilization in flowering plants. *Protoplasma*, Vol 193, Iss 1-4, pp 132-143.  
 Gamborg O.; Murashige, T.; Thorpe, T.; Vasil, I. 1976. Plant tissue culture media. *In Vitro* 12: 473- 478.  
 Ge L.; Tian, H.; Russell, S. 2007. Calcium function and distribution during fertilization in Angiosperms. *American Journal of Botany* 94 (6): 1046-1060.  
 Guevara C.; Ospina, J.; Mafla, G.; Verdier V. 1998. Zygotic embryo culture of *Manihot esculenta* Crantz: a practical approach for the safe international movement of cassava seed stocks. *Plant Genetic Resources Newsletter* 115:33-38.  
 Holm P.; Olsen O.; Schnorf M.; Brinch-Pedersen H.; Knudsen, S. 2000. Transformation of barley by microinjection into isolated zygote protoplasts. *Transgenic Res* 9:21-32.  
 Howell S. 1998. *Molecular Genetics of Plant Development*. Cambridge University Press. ISBN 0521587840, 9780521587846.  
 Kanta K.; Rangaswamy, W.; Maheshwari, P. 1962. Test tube fertilization in a flowering plant. *Nature* 194:1214-1217.  
 Kovacs M.; Barnabas, B.; Kranz, E. 1995. Electro-fused isolated wheat (*Triticum aestivum* L.) gametes develop into multicellular structures. *Plant Cell Rep.* 15:178-180.  
 Kranz E.; Bautor J.; LÖrz H. 1991. *In vitro* fertilization of single, isolated gametes of maize mediated by electrofusion. *Sex Plant Reprod* 4:12-16.  
 Kranz E.; Dresselhaus, T. 1996. *In vitro* fertilization with isolated higher plant gametes. *Trends in plant science reviews*. Vol. 1, Nro. 3, pp 82-89.  
 Kranz E.; Kumlehn, J. 1999. Angiosperm fertilization, embryo and endosperm development *in vitro*. *Plant Science*. Vol. 142, Iss 2, pp 183-197.  
 Litz R. 1991. "Cultivo de embriones y óvulos". En *Cultivo de tejidos en la Agricultura. Fundamentos y Aplicaciones*. Centro Internacional de Agricultura Tropical (CIAT). Roca y Mroginsky editores. pp 295-312.  
 Liu C.; Xu, Z.; Chua, N. 1993. Proembryo culture – *in vitro* development of early globular-stage zygotic embryos from *Brassica juncea* *Plant J.* 3:

- Cell. Dev. Biol. 2002. 18:81-105.
- Mol R.; Matthys-rochon, E.; Dumas, C. 1995. Embryogenesis and plant regeneration from maize zygotes by in vitro culture of fertilized embryo sacs. Plant Cell Reports. 14: 743-747.
- Monnier M 1976 Culture in vitro of embryo immature of *Capsella bursapastoris* Moench (L.) Rev Cytol Biol Veg 39: 1-120.
- Murashige T.; Skoog, F. 1962. A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plant. 15: 473-497.
- Ning J.; Peng X.; Qu, L.; Xin, H.; Yan, T.; Sun, M. 2006. Differential gene expression in egg cells and zygotes suggests that the transcriptome is restricted before the first zygotic division in tobacco. FEBS Lett 580:1747-752.
- Okamoto T.; Higuchi, K.; Shinkawa, T.; Isobe T.; Lörz, H; Koshiba, T.; Kranz, E. 2004. Identification of Major Proteins in Maize Egg Cells. Plant Cell Physiol. 45(10): 1406-1412.
- Ponya Z.; Barnabas, B. 2001. Microinjected fluorescent phalloidin *in vivo* reveals F-actin dynamics in isolated egg cells of wheat (*Triticum aestivum*, L.) developed *in situ* and fertilised *in vitro*. Journal of Plant. Physiology. Vol 158 (12): 1527-1539.
- Ponya, Z.; Timar, I.; Szabo, L.; Kristof, Z. 1999. Morphological characterisation of wheat (*Triticum aestivum* L.) egg cell protoplast isolated from immature and covered caryopses. Sex Plant Reprod. 11: 357-359.
- Randolph L., Cox L. 1943 Factors influencing the germination of Iris seed and the relation of inhibiting substances to embryo dormancy. Proc. Amer. Soc. hort. Sci. 43: 284-300.
- Rougier M.; Antoine, A.; Aldon, D.; Dumas, C. 1996. New lights in early steps of *in vitro* fertilization in plants. Sexual Plant Reproduction. Vol 9, Iss 6, pp 324-329.
- Russell S. 1985. Preferential fertilization in *Plumbago*: Ultrastructural evidence for gamete-level recognition in an angiosperm (cytoplasmic heritable organelles/male gamete/sexual reproduction). Proc. Natl. Acad. Sci. USA. Vol. 82, pp. 6129-6132.
- Tian H.; Yuan, T.; Russell, S. 2005 Relationship between double fertilization and the cell cycle in male and female gametes of tobacco. Sex Plant Reprod 17:243-252.
- Wang Y; Kuang A.; Russell S.; Tian, H. 2006. In vitro fertilization as a tool for investigating sexual reproduction of angiosperms. Sex Plant Reprod. 19: 103-115.
- Xu H.; Swoboda I.; Bhalla, P.; Singh M. 1999. Male gametic cell-specific gene expression in flowering plants. Proc. Natl. Acad. Sci. USA. Vol. 96, pp. 2554-2558.
- Zenkter M. 1999. In vitro pollination of excised ovaries. Acta Biologica Cracoviensia Series Botanica. Vol.41, pp 31-38.

## II CAPÍTULO 2

### Hibridación Somática

Pablo Polci y Pablo Friedrich

#### 1 Introducción

La mejora genética de las plantas cultivadas en procura de incrementar la producción viene siendo practicada por el ser humano desde hace miles de años, seguramente desde el inicio mismo de la agricultura. Para ello, el hombre ha empleado los cruzamientos con el fin de incrementar la variabilidad genética, paso inicial en el proceso de mejoramiento. De esta manera la hibridación entre especies diferentes permitió aumentar de modo significativo la productividad de diversos cultivos, como por ejemplo los cereales. Sin embargo, en otros cultivos esta estrategia no resultó exitosa a causa de esterilidad sexual o incompatibilidad genética.

La hibridación somática, es decir, la obtención de plantas híbridas a partir de la fusión de células o de protoplastos derivados de células somáticas, surgió hace unos 30 años como una herramienta muy promisoriosa para sortear problemas de incompatibilidad precigótica..

Esta técnica, si bien presenta algunas limitaciones, como una elevada esterilidad o imposibilidad de regenerar plantas en algunos casos, demostró ser útil en la mejora genética de plantas, principalmente por permitir la introgresión limitada de genes de especies silvestres en los cultivos. Se utiliza, por ejemplo, cuando se quiere transferir resistencia a enfermedades o tolerancia a estreses. También permite la obtención de citoplasmas híbridos (cíbridos). La hibridación asimétrica y cibridización sirve como puente para la transferencia de genes individuales

La técnica de fusión de protoplastos se desarrolló en la década de 1950-60, a partir de la observación de que ciertos virus pueden inducir la fusión de células animales. Sin embargo, recién en la década de 1980-90 tuvo un mayor auge debido a la aparición de métodos químicos y eléctricos más apropiados para la fusión de células vegetales. La pared presente en estas células representa una barrera que normalmente impide la fusión entre ellas. Por

ello, la obtención de plantas híbridas somáticas requiere del previo aislamiento de protoplastos mediante la digestión enzimática de las paredes celulares, para luego proceder a la fusión de protoplastos de distintos tipos parentales (distintos genotipos) y posterior regeneración de plantas a partir de los productos de fusión. La hibridación somática en plantas comenzó a desarrollarse a principios de 1970 con la aplicación de métodos químicos de fusión, y se extendió en los '80 con el empleo de métodos de electrofusión.

Es relevante destacar que, a los fines del mejoramiento genético, el sistema de protoplastos no sólo se emplea para la obtención de híbridos somáticos, sino que también constituye un material apropiado para la incorporación de ADN exógeno. Asimismo, la fusión de protoplastos tiene un uso mucho más amplio que el estrictamente de interés agronómico; en tal sentido, es de gran utilidad en estudios de biología celular así como para otras aplicaciones biotecnológicas, por ejemplo, la producción de anticuerpos monoclonales.

#### 2 Hibridación somática vs. hibridación sexual

En relación a su potencialidad para producir híbridos, existen diferencias importantes entre la fusión de protoplastos somáticos y el cruzamiento sexual:

- La hibridación sexual entre organismos de diferentes especies está limitada por barreras de incompatibilidad. En algunos casos estas barreras son precigóticas, es decir, no permiten que se produzca la fecundación. Tal limitante no existe en la fusión de protoplastos, dado que este proceso es no-específico, permitiendo la fusión de células de orígenes muy diversos, incluso de organismos muy distanciados filogenéticamente (por ejemplo, células animales y vegetales). Ello no significa que pueda regenerarse un organismo viable y fértil a partir de cualquier tipo de combinación entre células. De hecho, la búsqueda de híbridos de interés agronómico ha demostrado que el principal aporte de esta metodología es

la introgresión, en los cultivos, de genes provenientes de plantas silvestres emparentadas.

- La hibridación sexual ocurre normalmente por fusión de células haploides ( $n + n$ ), mientras que en la hibridación somática se fusionan dos protoplastos con sus genomas completos ( $2n + 2n$ ) o con uno de ellos conteniendo sólo parte de su genoma.
- Otra diferencia está dada por la herencia de los genes extranucleares, esto es, plastídicos y mitocondriales. Mientras que en la reproducción sexual el genoma citoplasmático es aportado casi exclusivamente por la gameta femenina, en la hibridación somática se combinan organelas de ambos progenitores, produciendo nuevas combinaciones de genes citoplasmáticos.

### 3 Hibridación somática vs. transformación genética

La preponderancia que han tomado las técnicas de transformación genética ha relegado a la hibridación somática a un papel de menor significación como herramienta biotecnológica aplicada al mejoramiento de los cultivos. Comparando ambas metodologías, podemos destacar las siguientes diferencias:

- En la transformación genética se incorporan uno o pocos genes exógenos en una planta, mientras que en la hibridación somática el número de genes introducidos es mayor dado que se transfieren cromosomas de una especie a otra o se combinan dos genomas completos.
- La hibridación somática permite transferir caracteres sin necesidad de contar con un conocimiento detallado de los mecanismos moleculares que determinan la expresión de dichos caracteres. Por el contrario, la transformación genética requiere la previa identificación y aislamiento de los genes que determinan la expresión del carácter de interés.
- Caracteres tales como productividad, tolerancia a ciertos tipos de estrés, resis-

tencia horizontal a enfermedades y otros son poligénicos, por lo que su transferencia es factible por hibridación somática pero no por transformación genética. Sin embargo, en la actualidad, con las nuevas herramientas aportadas por la genómica se está avanzando en el conocimiento de todos los genes involucrados en diversas vías metabólicas. Estos podrían ser transferidos en conjunto vía transgénesis.

### 4 Tipos de híbridos somáticos

Cuando se realiza la fusión de protoplastos se obtienen células con el aporte de cromosomas de dos o más núcleos. Si los protoplastos involucrados en la fusión proceden de distintos tipos parentales se originan “**heterocarios**”, y si proceden del mismo tipo parental se originan “**homocarios**”. Cuando en el heterocario ocurre la fusión de los núcleos (cariogamia), se forma una “**célula híbrida**”.

A partir de una célula híbrida, que contiene dos genomas completos, se puede regenerar una planta que, según cual haya sido el destino del material genético nuclear durante las divisiones celulares que siguen a la fusión, puede ser de tres tipos:

1. **Híbrido simétrico**, que tiene el genoma nuclear completo de cada tipo parental.
2. **Híbrido asimétrico**, que tiene el genoma nuclear completo de uno de los parentales y sólo una parte del genoma nuclear del otro parental.
3. **Cíbrido**, que contiene el genoma nuclear completo de uno de los parentales, reteniendo sólo material genético extranuclear del otro parental.

Es usual que en el proceso de regeneración de plantas híbridas ocurra una eliminación gradual de cromosomas de uno de los tipos parentales, produciéndose de este modo híbridos asimétricos o cíbridos. La práctica de la hibridación somática en plantas demostró que, en general, los híbridos asimétricos y cíbridos son más valiosos que los simétricos producidos entre plantas alejadas filogenéticamente. Por esta razón se han desarrollado métodos para



inducir la hibridación asimétrica o cibridación, alterando el genoma de uno de los protoplastos parentales. En este caso se dice que se transfiere parte del genoma de un protoplasto donante a uno receptor. Por lo tanto, teniendo en cuenta el material de partida empleado para la fusión, pueden producirse tres tipos de células híbridas:

1. **Células híbridas simétricas**, por fusión de dos protoplastos con sus genomas completos.
2. **Célula híbrida asimétrica**, por fusión de un protoplasto conteniendo su genoma completo con otro conteniendo su núcleo inviable o sólo una parte de su genoma nuclear. Los protoplastos donantes son irradiados con rayos X o Gamma, lo que provoca la fragmentación de los cromosomas. Una vez producida la fusión, dichos fragmentos tienden a eliminarse en las sucesivas mitosis, pero algunos de ellos pueden integrarse con el genoma del receptor produciendo un híbrido asimétrico. El tratamiento de irradiación parece no tener efectos deletéreos o mutagénicos sobre los genomas de organelas, probablemente debido a que cada célula posee varias copias de ellas. También puede producirse una célula híbrida asimétrica fusionando protoplastos receptores con microprotoplastos, conteniendo uno o pocos cromosomas. Los microprotoplastos se obtienen induciendo la formación de micronúcleos por tratamientos con agentes antimicrotúbulos.
3. **Cíbridos**, por fusión de un protoplasto conteniendo su genoma nuclear completo con un protoplasto enucleado o con su núcleo inviable. Los protoplastos enucleados o citoplastos pueden obtenerse centrifugando los mismos a alta velocidad en un gradiente de densidad isosmótico. Asimismo, el método de irradiación con rayos X o Gamma puede generar cíbridos en casos en que se produzca la eliminación total de los fragmentos cromosómicos. Los protoplastos que poseen su genoma nuclear completo (receptores), pueden ser tratados pre-

viamente con inhibidores metabólicos. En este caso sólo pueden sobrevivir, por complementación metabólica, los heterocariones conteniendo el núcleo del receptor y el citoplasma del donante enucleado.

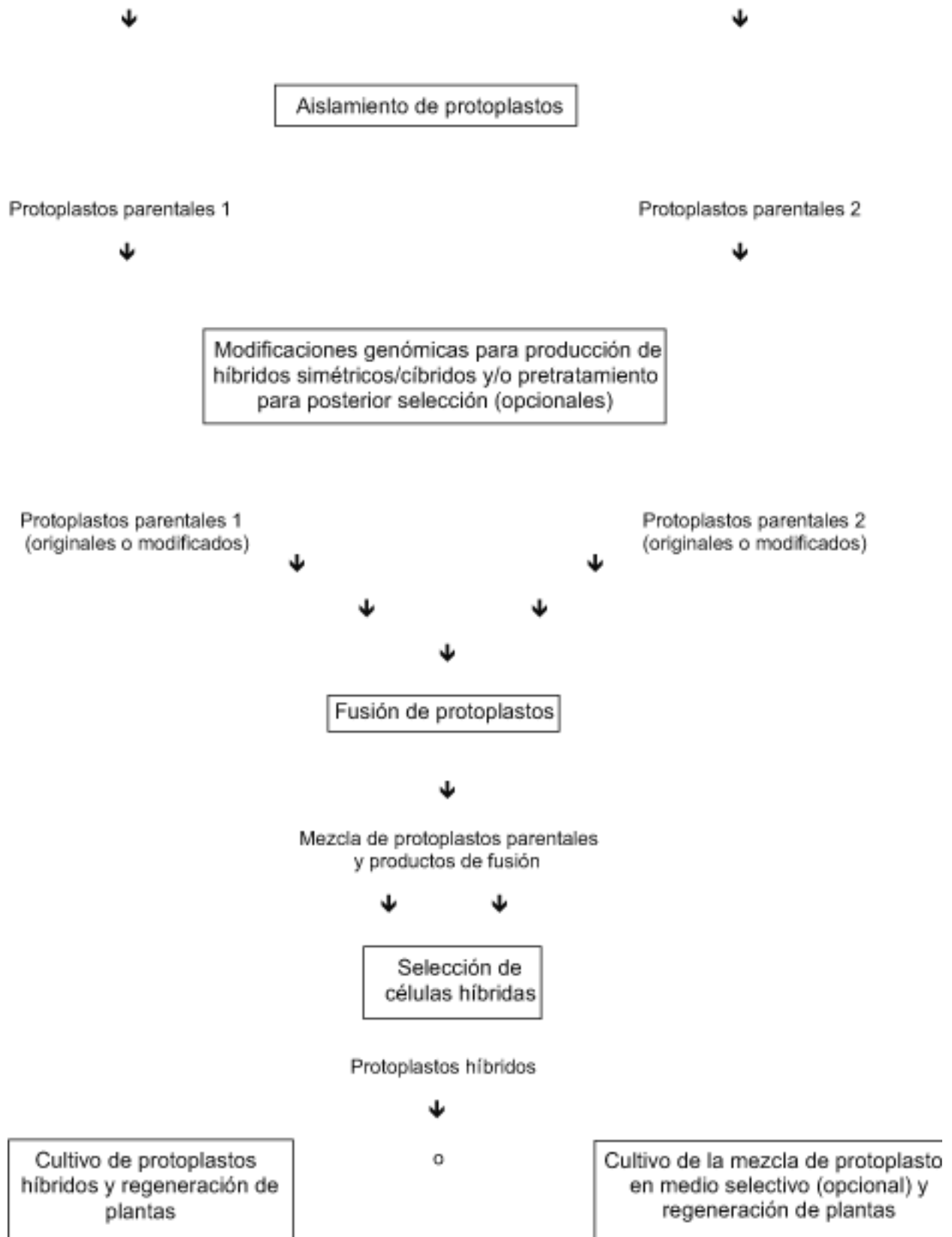
La segregación de los plástidos y mitocondrias de los dos tipos parentales también constituye una fuente importante de variabilidad en la regeneración de plantas híbridas. Es frecuente que ocurran recombinaciones de los genomas mitocondriales de ambos tipos parentales, pero ello es poco común entre plástidos. Dado que las organelas segregan independientemente unas de otras, la regla es que al cabo de pocas divisiones mitóticas las células formadas contengan plástidos provenientes de uno u otro tipo parental. Así, una célula híbrida origina una colonia de células que exhiben diferentes combinaciones de genes citoplasmáticos, pero con una mayor frecuencia de células que poseen mitocondrias recombinantes y plástidos de uno u otro tipo parental. El destino que sufre el material genético nuclear y extranuclear durante las divisiones celulares determina que, a partir de una población de células híbridas similares, se puedan regenerar plantas con diferentes combinaciones de genes nucleares, plástídicos y mitocondriales.

La Figura 1 muestra un esquema de las etapas involucradas en la hibridación somática, las cuales son tratadas a continuación.

### 5 Aislamiento de protoplastos

El desarrollo de métodos enzimáticos de aislamiento a partir de 1960 brindó la posibilidad de obtener grandes cantidades de protoplastos vegetales. Ello tuvo gran influencia en diferentes áreas de la biología y la biotecnología de plantas, dada la utilidad de los protoplastos como sistema experimental para estudios fisiológicos, bioquímicos y moleculares, así como para su aplicación al mejoramiento genético.

Los protocolos de aislamiento emplean enzimas fúngicas con actividad celulasa y pectinasa, siendo necesario en algunos casos el agregado de hemicelulasas. Las celulasas y hemicelulasas degradan la pared celular, en



**Figura 1:** Materiales y procesos involucrados en la hibridación somática.

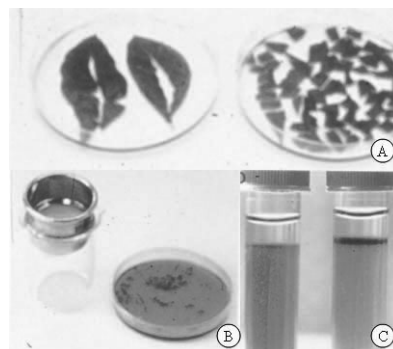
tanto que las pectinasas digieren la matriz de pectina que mantiene adheridas a las células. El uso de enzimas disponibles comercialmente posibilitó el aislamiento de protoplastos de prácticamente cualquier tejido vegetal, siempre que el mismo no esté lignificado. Se ha podido aislar protoplastos de una gran cantidad de especies vegetales y regenerar plantas a partir de los mismos, permitiendo el uso de este sistema para la propagación clonal y la modificación genética de plantas.

El número de protoplastos aislados, su viabilidad y la pureza de la suspensión obtenida dependen de varios factores, debiendo establecerse en forma empírica las condiciones óptimas para un sistema determinado. Las etapas del aislamiento y los factores a considerar en cada una de ellas se describen a continuación:

- 1. Selección del material de partida.** Influye en el resultado del aislamiento así como en el proceso de regeneración posterior. Generalmente se emplean hojas y células en cultivo líquido o sólido. Cuando se emplean hojas, la edad y las condiciones de cultivo de la planta afectan el rendimiento. Se obtienen mejores resultados con hojas jóvenes totalmente expandidas que con hojas viejas (Fig. 2A). Para facilitar el contacto de las enzimas con las células, las hojas suelen cortarse en trozos pequeños, muy finos en el caso de las monocotiledoneas. Algunas veces se extrae la epidermis inferior antes de colocar la hoja en la solución enzimática, como sucede con las dicotiledoneas. Cuando se emplean células en suspensión, se obtienen mejores resultados si se encuentran en la fase exponencial de crecimiento.
- 2. Tratamiento enzimático.** La concentración de la enzima y el tiempo de incubación requeridos para lograr la liberación de los protoplastos dependen del producto comercial utilizado. El pH generalmente se ajusta en un rango de 4,7 a 6, y la temperatura entre 25 y 30 °C. La duración del tratamiento enzimático y la relación volumen de solución/cantidad de tejido influyen en el rendimiento.

Durante el aislamiento se produce la lisis de algunas células, que liberan enzimas hidrolíticas y compuestos oxidantes que pueden dañar los protoplastos, por lo que se suelen agregar inhibidores de proteasas y antioxidantes tales como el Polivinilpirrolidona-10. El agregado de un estabilizador osmótico a la solución enzimática es esencial para evitar la ruptura de los protoplastos a medida que son liberados, siendo más estables si la solución es ligeramente hipertónica. Para ello se utilizan solutos osmóticamente activos, comúnmente manitol o sorbitol, en concentraciones que varían entre 0,3 a 0,8 M según el tipo de protoplasto. La solución enzimática es generalmente suplementada con sales, comúnmente  $\text{CaCl}_2$ , que incrementan la estabilidad de la membrana (Fig. 2B).

- 3. Limpieza y purificación de los protoplastos.** Una vez lograda la liberación de los protoplastos, se procede a la remoción de las enzimas y de los restos de tejido y de células rotas. La suspensión se pasa por un filtro de nylon o metáli-



**Figura 2:** Aislamiento de protoplastos de tabaco. **A.** Eliminación de las nervaduras centrales y seccionado de la hoja en trozos pequeños. **B.** Digestión enzimática de las paredes celulares. **C.** Centrifugación y obtención de protoplastos libres de restos de tejidos vegetales.

co con un diámetro de poro (30-100  $\mu\text{m}$ ) que permita el paso de los protoplastos y retenga los restos de tejido y grupos de células no digeridas. Posteriormente, se efectúan 3 ó 4 lavados centrifugando la suspensión por 5 minutos a baja velocidad (50-100 g). Finalmente, los protoplastos sedimentados se resuspenden en una solución salina que contenga un estabilizador osmótico. De este modo se eliminan las enzimas y restos celulares. Para lograr una mayor pureza de la suspensión es conveniente centrifugarla en un gradiente de densidad (Fig. 2C). Para el recuento de protoplastos se emplea un hemocitómetro. La viabilidad se determina generalmente empleando colorantes que son incorporados (rojo neutro, diaacetato de fluoresceína) o excluidos (azul de Evan) por las células viables; también pueden evaluarse la actividad fotosintética (emisión de fluorescencia) y la respiratoria (consumo de  $\text{O}_2$ ).

## 6 Métodos de fusión

### 6.1 Métodos químicos

Los primeros casos informados de fusión controlada y reproducible de protoplastos vegetales y de regeneración de híbridos somáticos se produjeron a principios de la década de 1970, empleando nitrato de sodio como agente fusógeno. La frecuencia de formación de heterocariones en tales casos era muy baja. Posteriormente, se lograron mejores resultados fusionando protoplastos de células del mesófilo de *Nicotiana* en un medio conteniendo alta concentración de  $\text{Ca}^{++}$  (50 mM) y pH elevado (10,5). Sin embargo, el fusógeno que alcanzó mayor aceptación es el polietilenglicol (PEG) debido a que, en comparación con los otros agentes químicos, produce mayor proporción de heterocariones y, para muchos tipos celulares, resulta menos tóxico. Además, el tratamiento con PEG genera una mayor proporción de heterocariones binucleados, con menor ocurrencia de fusiones entre más de dos protoplastos.

El procedimiento de fusión con PEG más ampliamente utilizado es esencialmente una

combinación del método original del PEG y el de  $\text{Ca}^{++}$  y pH elevados. Los protoplastos de dos tipos parentales se mezclan en una solución conteniendo 15 a 45 % de PEG (peso molecular de 1500 a 6000), durante 15 a 30 minutos. La temperatura de incubación óptima para inducir la fusión es de 35 a 37 °C pero, en la práctica, suele ajustarse a 24 °C. La densidad de protoplastos adecuada para la formación de heterocariones es de 4 a 5 % (volumen de protoplastos/volumen de suspensión). La remoción del PEG se lleva a cabo en forma gradual, siendo esto fundamental para asegurar una elevada frecuencia de heterocariones. Para ello, la suspensión se somete a sucesivos lavados con una solución alcalina (pH 9-11) de  $\text{CaCl}_2$  10-50 mM, disminuyendo progresivamente la concentración de PEG con cada lavado.

En la fusión de los protoplastos ocurren, en forma sucesiva, los siguientes eventos: (1) las membranas plasmáticas de protoplastos adyacentes entran en íntimo contacto, (2) se producen alteraciones en sitios localizados de las mismas, formándose puentes citoplasmáticos entre protoplastos vecinos, y (3) las interconexiones citoplasmáticas se expanden, provocando la fusión. La carga negativa neta de la superficie de las membranas produce la repulsión entre ellas; el tratamiento con  $\text{Ca}^{++}$  y pH elevados neutraliza las cargas superficiales, favoreciendo el contacto entre protoplastos. El PEG actuaría como un puente molecular causando alteraciones en las membranas plasmáticas que promueven la adhesión y formación de puentes citoplasmáticos entre protoplastos vecinos, produciéndose finalmente la fusión.

### 6.2 Electrofusión

En 1973 se descubrió que pulsos de elevado potencial eléctrico en un rango muy estrecho de intensidad y duración conducen a un incremento reversible, experimentalmente controlable, de la permeabilidad de la membrana celular. Este efecto se denominó ruptura eléctrica reversible para enfatizar la habilidad de la membrana para reparar tales perturbaciones. El voltaje mínimo para un protoplasto (umbral) con el cual se produce la formación de poros disminuye a mayor duración del pulso eléctrico.

Este fenómeno de ruptura reversible también es aplicado en la técnica de electroporación, con la que se pueden introducir en las células construcciones de ADN y otras moléculas de alto peso molecular.

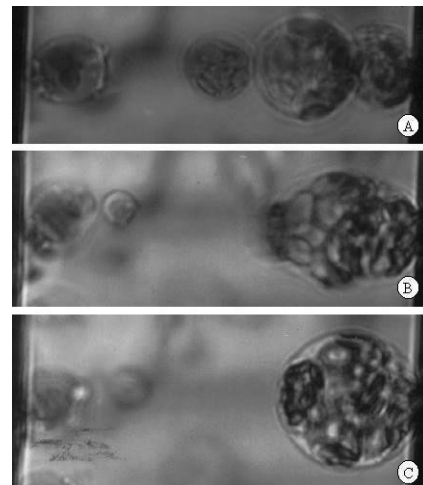
El método de electrofusión en esencia involucra dos pasos:

Los protoplastos, colocados en un medio de baja conductividad entre dos electrodos se ponen en contacto al aplicar corriente alterna de alta frecuencia (0,5 a 1,5 MHz) en un campo eléctrico no uniforme. Las cargas se separan dentro de las células formando un dipolo y generando, por lo tanto, un potencial de membrana. La fuerza del campo a ambos lados de una célula es diferente (campo no uniforme), dando origen a una fuerza neta que la empuja a una región de mayor intensidad de campo (hacia uno de los electrodos). Este proceso se denomina **dielectroforesis**. La polaridad de la célula incrementa el campo local no uniforme, atrayendo a las células vecinas. Luego, los protoplastos se aproximan unos a otros y forman cadenas paralelas a las líneas de campo aplicadas (Fig. 3A).

La fuerza ejercida sobre la célula bajo condiciones dielectroforéticas depende (1) de la intensidad de campo eléctrico, (2) del gradiente del campo, (3) del volumen del protoplasto, y (4) de la diferencia entre las constantes dieléctricas de la célula y su ambiente (así como la correspondiente diferencia entre sus conductividades).

El segundo paso consiste en la aplicación de uno o más pulsos cortos de corriente directa de 15 a 100  $\mu$ seg. con una intensidad de campo elevada (1 a 3 Kv/cm), suficiente para causar la ruptura reversible de la membrana. Para que esto ocurra es necesario que el voltaje crítico de membrana sea alcanzado en cuestión de nanosegundos a microsegundos. Ocurre entonces la **fusión** de las dos células cuyas membranas están en estrecho contacto (1 a 2 nanómetros de distancia). El proceso de resellado es principalmente atribuido a las moléculas lipídicas debido a que su coeficiente de difusión es dos órdenes de magnitud mayor que el de las proteínas. Durante este proceso pue-

den formarse puentes lipídicos entre las membranas de las dos células. En otras palabras, las membranas no se vuelven a sellar separadamente en la zona de contacto. Los puentes formados de esta manera, con continuidad citoplasmática entre las dos células, conducen a un radio de curvatura muy pequeño y a altas tensiones superficiales. Como consecuencia se forma una nueva célula esférica, ya que el proceso es favorecido energéticamente (Fig. 3B y C). La ruptura es reversible si el número y tamaño de los poros es pequeño en relación al total de la superficie de la membrana. Fuerzas de campo supracríticas, así como largos períodos de exposición, rompen áreas mayores de membrana y pueden producir la destrucción total de la célula.



**Figura 3:** Electrofusión de protoplastos de mesófilo de pasto llorón. **A.** Dielectroforesis de células. Las diferencias en intensidad de campo eléctrico hacen que las células se muevan hacia la zona cercana al electrodo. **B.** Aplicación de pulsos cortos de corriente directa, que inducen la formación de poros en la membrana, generando puentes entre las membranas de las células adyacentes. **C.** Los puentes con continuidad citoplasmática poseen un radio de curvatura muy pequeño. Las altas tensiones superficiales generadas en ese sector conducen a la formación de una nueva célula esférica.

- La frecuencia de fusión depende de varios factores:
- El genotipo.
- La procedencia de los protoplastos dentro del mismo genotipo. Diferentes poblaciones de protoplastos exhiben frecuencias de fusión diferentes, aún cuando provengan del mismo genotipo. En general, los protoplastos obtenidos a partir de hojas se fusionan más fácilmente que los provenientes de raíz o de cultivos celulares.
- El tamaño de los protoplastos. Los más grandes se fusionan más fácilmente.
- La densidad de los protoplastos.
- El número, la duración y la fuerza de los pulsos de corriente directa. La duración del pulso eléctrico es de 15 a 50  $\mu\text{seg}$ . El tiempo requerido para la fusión que sigue a la aplicación de un pulso varía desde pocos segundos a varios minutos. Esto probablemente esté relacionado a la diferencia de fluidez de la membrana en los distintos tipos celulares y a las propiedades del citoesqueleto dentro de la célula.
- El medio de fusión. Un factor limitante en la agregación dielectroforética de las células es la conductividad eléctrica del medio; si ésta es muy alta, hay mucho flujo de corriente en la solución y se producen calentamientos y turbulencias que impiden la formación de cadenas.
- El potencial osmótico del medio de fusión. La inclusión de iones en el medio puede incrementar el porcentaje de fusión. Sin embargo, existe la limitación de que la dielectroforesis debe realizarse en un medio de baja conductividad; el medio de fusión usualmente consiste en manitol (cuya concentración se ajusta según los protoplastos empleados) y  $\text{CaCl}_2$  1 mM. Valores bajos de presión osmótica en el medio de suspensión de los protoplastos puede favorecer el proceso de fusión.
- La longitud de las cadenas de protoplastos formadas durante la dielectroforesis, que además depende de factores como la densidad de protoplastos, fuerza del

campo eléctrico y tiempo durante el cual es aplicado. Cadenas más largas darán mayor frecuencia de fusión que las cadenas cortas. Pulsos de corriente más largos y de mayor voltaje producen un aumento de los eventos de fusión múltiple; obviamente, pulsos muy largos y voltajes muy elevados pueden matar a los protoplastos.

- La composición y propiedades de la membrana plasmática pueden ser afectadas por la actividad metabólica de los protoplastos y por el tipo de enzimas empleadas en el proceso de aislamiento, influyendo en consecuencia en el proceso de fusión.

Las condiciones experimentales deben ajustarse de modo de obtener la mayor frecuencia de fusión posible (número de protoplastos fusionados sobre el total de protoplastos), intentando maximizar la proporción de heterocariones (productos de la fusión de protoplastos diferentes) formados por sobre la de homocariones (productos de la fusión de protoplastos del mismo tipo parental), y minimizando la ocurrencia de eventos de fusión múltiple (fusión de más de dos protoplastos). La proporción de cada uno de los dos tipos de protoplastos en la suspensión puede fijarse a fin de incrementar la formación de heterocariones por sobre la de homocariones. El tipo de protoplasto con mayor tendencia a fusionar se mezcla con el de menor tendencia a fusionar en una relación de 1:5 a 1:10. Así, durante la formación de cadenas, los protoplastos más fusionables estarán mayoritariamente en contacto con los menos fusionables, y éstos a su vez estarán ligados entre ellos mismos. Seleccionando las condiciones -duración y voltaje de los pulsos- que promuevan sólo la fusión de los protoplastos más fusionables, los productos serán mayormente heterocariones.

Sin embargo, aún cuando las condiciones de fusión puedan fijarse de modo de incrementar la frecuencia de fusión y la proporción de heterocariones formados, la fusión de protoplastos a gran escala (macrofusión), tanto por métodos químicos como eléctricos, resultará en una

mezcla de protoplastos parentales, homocariaciones y heterocariaciones. Esto conlleva la necesidad de emplear métodos de selección de los híbridos somáticos obtenidos. Una técnica alternativa que ofrece la ventaja de no requerir una posterior selección de heterocariaciones, aunque demanda mayor manipulación, es la microfusión o electrofusión de pares de protoplastos. Esta técnica se basa en los mismos principios que la electrofusión a gran escala pero está diseñada para inducir la fusión en pares seleccionados de protoplastos. Este sistema emplea microelectrodos de platino fijados a un soporte montado debajo del condensador de un microscopio invertido. Los pares de protoplastos seleccionados se colocan en microgotas de medio de fusión sobre un soporte, recubiertas por aceite mineral. En cada evento de fusión, los microelectrodos se posicionan a cada lado de una microgota conteniendo un par de protoplastos, y se aplican los pulsos eléctricos. Después de la fusión, los heterocariaciones resultantes son transferidos a gotas con medio de cultivo para la posterior regeneración de plantas híbridas. La tasa de fusión es de aproximadamente 30 pares de protoplastos fusionados y transferidos a medio de cultivo por hora, pudiendo alcanzarse frecuencias de fusión cercanas al 50 %.

#### **Ventajas de la electrofusión:**

1. El proceso entero puede ser seguido bajo microscopio por lo que los híbridos pueden ser fácilmente identificados y transferidos a un medio nutritivo o selectivo.
2. Puede preseleccionarse el número de células a ser fusionadas. Las formaciones de dos células son favorecidas por bajas densidades en la acanaladura entre los electrodos. Para obtener productos de multifusión deben emplearse elevadas densidades.
3. El proceso de fusión es sincrónico, ya que el pulso provoca la fusión de todas las células expuestas al campo.
4. Se obtienen elevadas plasmogamia (80 - 100 %) y cariogamia.
5. La viabilidad de las células fusionadas es muy buena.
6. Este método ofrece ventajas sobre otros

como el del PEG, que: - requiere condiciones no fisiológicas, - las frecuencias de formación heterocariaciones binucleadas son bajas y variables, - resultan tóxicos para diversos tipos celulares, - el fusógeno debe ser removido antes del cultivo de los protoplastos y - bajo rendimiento de células fusionadas.

#### **7 Selección de heterocariaciones**

La fusión de protoplastos a gran escala produce una mezcla de tipos parentales y productos de fusión. Si no se efectúa una previa selección de las células híbridas, la regeneración de plantas y la posterior identificación de los híbridos demandará mucho trabajo y recursos. Dicha selección es particularmente necesaria cuando se emplean métodos químicos, que producen una baja proporción (<10%) de células híbridas. Por el contrario, los métodos eléctricos no requieren necesariamente del posterior proceso de selección dado que normalmente producen frecuencias de fusión muy elevadas o porque, en el caso de la microfusión, no se generan mezclas heterogéneas de protoplastos debido a que se fusionan dos células por vez. Cualquiera sea el caso, la condición de híbrido debe verificarse en la planta regenerada mediante diferentes análisis que se explican más adelante.

La selección de células híbridas puede llevarse a cabo empleando los siguientes métodos:

- **Selección en base a caracteres morfofisiológicos.** En ciertos casos es posible diferenciar los callos híbridos de los parentales. Por ejemplo, en algunos cruzamientos, los callos híbridos poseen un color intermedio al de los parentales. Cuando se produce un cruzamiento entre un parental con capacidad para regenerar planta con uno recalcitrante, los híbridos somáticos normalmente regeneran plantas, ya que este carácter se comporta como dominante. Si los protoplastos del genotipo respondedor son tratados con inhibidores metabólicos irreversibles (iodoacetamida, dietilpirocarbamato, etc.) no sobrevivirán, y solamente podrán regenerarse los híbridos.

- **Selección por complementación.** La selección se produce porque el medio de cultivo resulta restrictivo para el crecimiento de los parentales pero no para el de los híbridos. La complementación puede lograrse por las siguientes vías:

a) Empleando dos parentales con defectos metabólicos o genéticos no alélicos. Si dichos caracteres son recesivos, la fusión produce una célula híbrida en la que los defectos se anulan por complementación. Por ejemplo, la fusión de dos variedades albinas no alélicas (nucleares) de tabaco, permitió obtener plantas verdes.

AAbb (albina) x aaBB (albina)



AaaaBBbb (verde)

Asimismo, a partir de dos líneas mutantes no alélicas de tabaco deficientes en nitrato reductasa, se obtuvieron colonias capaces de crecer en un medio conteniendo nitrato como única fuente de nitrógeno.

b) Fusionando parentales que expresan caracteres dominantes tales como resistencia a antibióticos o herbicidas. Para ello se emplea una línea resistente a cierta droga, y otra resistente a una segunda droga, efectuándose la selección de las células híbridas en un medio conteniendo ambas drogas a concentraciones que resultan letales para las líneas parentales.

c) Una línea que presente una mutación autotrófica (por ejemplo: albinismo, deficiencia a nitrato reductasa) y un carácter dominante (por ejemplo resistencia a antibióticos o herbicidas) es de gran utilidad para la selección. Los híbridos producidos por el cruzamiento de estos parentales con cualquier genotipo silvestre pueden ser seleccionados en medio mínimo suplementado con el antibiótico o herbicida, que no permite el crecimiento de los parentales.

- **Selección por aislamiento de los heterocariones o células híbridas.** Es el sistema más confiable y de más amplio espectro de aplicación. El aislamiento puede realizarse de dos maneras:

a) Selección mecánica directa utilizando téc-

nicas de micromanipulación con micropipeta. Puede realizarse cuando los protoplastos parentales presentan rasgos que los distinguen al microscopio, permitiendo reconocer las células híbridas. Por ejemplo al fusionar protoplastos del mesófilo (verdes) con protoplastos de suspensiones celulares (incoloros). También pueden colorearse las células de ambos progenitores con diferentes colorantes vitales, como el rojo neutro y azul brillante de cresilo.

b) Citometría de flujo. Los protoplastos parentales se marcan con diferentes colorantes fluorescentes, por ejemplo, rodamina (rojo) y fluoresceína (verde amarillento). El aislamiento de los heterocariones marcados con ambos colorantes se realiza utilizando un citómetro de flujo (FACS, fluorescence-activated cell sorter), que es preciso y muy rápido. El equipo posee fotocélulas que detectan fluorescencia, pudiendo separar los protoplastos marcados con un solo fluorocromo (parentales) de aquellos doblemente marcados (híbridos).

## 8 Cultivo de protoplastos

La habilidad para regenerar plantas a partir de células y tejidos en cultivo es un componente esencial en la biotecnología para la manipulación genética y el mejoramiento vegetal. Esto resulta relativamente fácil para las dicotiledóneas herbáceas, pero ha sido bastante difícil para las monocotiledóneas, particularmente las gramíneas.

Los protoplastos se cultivan en cajas de Petri en medio líquido o semisólido, rico en compuestos orgánicos e inorgánicos, suplementado con reguladores del crecimiento, y en presencia de un estabilizador osmótico. Suelen ser cultivados en cajas de Petri con medio agarificado, donde se mezcla la solución de protoplastos con un medio de cultivo líquido a 40°C conteniendo 1,2% de agar, preferentemente agarosa (Fig. 4A). Los protoplastos quedan, así, atrapados en el medio semisólido y, luego de varios días, se ven las colonias formadas (Fig. 4B). Sin embargo, el medio líquido da mejores resultados por varias razones: (a) los protoplastos de varias especies solo pueden dividirse en medio líquido, (b) la presión osmótica del medio puede ser fácilmente reducida a los pocos días en cultivo, (c) la densidad celular puede

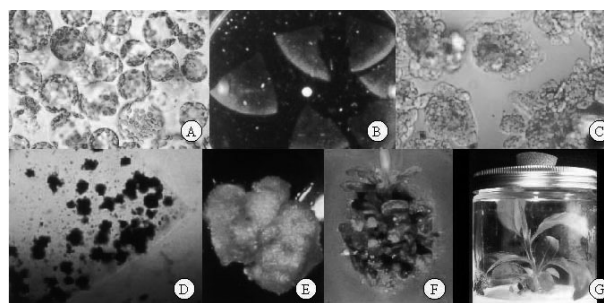


ser modificada agregando medio de cultivo o las células con especial interés pueden ser aisladas y cultivadas a alta densidad, (d) la degradación de algunos componentes del medio por la población de protoplastos puede producir algunas sustancias citotóxicas, cuya concentración alrededor de las células será menor en el medio líquido. Una técnica muy eficiente que combina medio sólido y líquido es la de cultivo en perlas de agarosa, donde los bloquitos con los protoplastos inmovilizados en agarosa son cortados en cuatro mitades y colocados en medio líquido, donde se cultivan en continua agitación.

Los protoplastos regeneran la pared celular luego de uno a cuatro días en cultivo. La presencia de pared es esencial para lograr una división regular. Luego de dos a tres semanas, producto de mitosis sucesivas, se forman microcolonias derivadas cada una de una célula, que dos semanas después se pueden ver a simple vista. Estos callos son transferidos a un medio de cultivo y a condiciones apropiadas para regenerar plantas (Fig. 4).

Los principales factores a tener en cuenta para el cultivo de protoplastos son:

- Los requerimientos nutricionales: se han utilizado en muchos casos los mismos medios de cultivo que se utilizan en el cultivo de células y tejidos, pero en general son enriquecidos con vitaminas, azúcares, aminoácidos, reguladores de crecimiento y también con aditivos inespecíficos como leche de coco, hidrolizado de caseína, extracto de levadura, etc.
- El potencial osmótico del medio de cultivo: los protoplastos requieren de protección osmótica antes de regenerar la pared celular. Normalmente se ajusta con el agregado al medio de cultivo de manitol o sorbitol.
- La densidad celular: la densidad inicial de protoplastos varía de  $10^4 - 10^5$  protoplastos/ml de medio de cultivo. Los cultivos con altas densidades producirán, a partir de cada protoplasto, colonias que deberán separarse tempranamente para evitar quimeras. Si deseamos colonias bien separadas para asegurar que provengan de un solo protoplasto, la densidad inicial debe ser baja, de 100 – 500 protoplastos/ml. Hay protoplastos de varias especies y tipos que no logran dividirse a bajas



**Figura 4:** Cultivo de protoplastos de tabaco. **A.** Protoplastos de mesófilo. **B y D.** Cultivo de los protoplastos en perlas de agarosa suspendidas en medio líquido. **C.** Detalle de la formación de callo a partir de protoplastos. **E.** Callo en medio de regeneración. **F.** Desarrollo de plantas a partir de callos de protoplastos. **G.** Planta regenerada.

densidades de cultivo. Para estos casos se desarrolló la técnica de cultivo con células nodrizas o alimentadoras. Esta técnica consiste en exponer una suspensión de  $10^6$  protoplastos/ml a rayos X a dosis que inhiben la división celular pero que quedan metabólicamente activas. Estas células se plaquean en un medio agarificado, y luego se colocan por sobre éstas los protoplastos sin irradiar, de los cuales se formarán las colonias. También suele utilizarse papel de filtro para separarlas de las células nodrizas. Otra técnica utilizada es el cultivo en microgota, donde se puede cultivar hasta un solo protoplasto. Cada microgota contiene entre 0,25 - 25  $\mu$ l de medio de cultivo, logrando densidades de  $2 - 4 \times 10^3$  protoplastos/ml. Con ayuda de un micromanipulador se coloca la célula y se la cubre con aceite mineral para evitar la deshidratación del medio.

- Los tratamientos físicos: tratamientos de choque eléctrico y térmico estimulan la división de los protoplastos..
- Las condiciones de cultivo: en general los protoplastos son sensibles a la luz, por lo cual durante el cultivo se los mantiene en oscuridad o bajo luz muy tenue.
- El origen de los protoplastos: el estado fisiológico del tejido del cual provienen y la calidad de los protoplastos son muy importantes para obtener una elevada eficiencia en la respuesta al cultivo.

### 9 Identificación de las plantas híbridas

La condición de híbrido debe verificarse en la planta regenerada aún cuando se haya efectuado algún tipo de selección para el cultivo de las células híbridas. A tal fin es conveniente efectuar diferentes evaluaciones, lo que permite una mejor caracterización del híbrido. Comúnmente se llevan a cabo los siguientes estudios:

1. **Morfológicos.** Los híbridos somáticos pueden presentar rasgos morfológicos característicos. Los mismos pueden ser intermedios a los de los parentales, la suma de ellos, u otros completamente nuevos.
2. **Citológicos.** El análisis del complemen-

to cromosómico permite revelar si el híbrido posee el total de cromosomas de cada parental, si se han fusionado más de dos protoplastos, el grado de aneuploidía y posibles translocaciones intergenómicas. Mediante hibridación *in situ* empleando sondas de ADN repetitivo específico de especie puede evaluarse con más profundidad la contribución de los genomas parentales y reestructuraciones cromosómicas en el híbrido.

3. **Isoenzimáticos.** El patrón de bandas isoenzimáticas del híbrido debe mostrar bandas de ambos parentales, pudiendo haber otras bandas que resultan de nuevas combinaciones de las subunidades enzimáticas. Habitualmente se analizan los patrones de glucosa-6-fosfato deshidrogenasas, fosfoglucoisomerasas, glutamato oxalacetato transaminasas, esterasas, peroxidasas y fosfatasa ácida. Las isoenzimas son extremadamente variables entre tejidos vegetales, por lo que para el análisis se deben emplear los mismos tejidos u órganos, y en el mismo estadio fenológico.
4. **A nivel de ADN.** La demostración de la presencia de ADN de ambos parentales es la prueba más directa de la hibridación. El análisis de ADN es independiente del tejido empleado y de la edad de la planta. Puede llevarse a cabo por RFLP (polimorfismos en la longitud de los fragmentos de restricción), RAPD (polimorfismos en el ADN amplificado al azar) o Southern blot, empleando sondas de ADN repetitivo específico de especie o de genes de ARN ribosomal.

### 10 Logros y tendencias de la hibridación somática

La hibridación somática ha permitido producir híbridos que no se habían podido obtener por cruzamientos sexuales, incrementando así el flujo de genes en los cultivos. La tabla 1 muestra algunos ejemplos de híbridos somáticos que presentan algunas ventajas agronómicas.

En sus inicios, algunos esperaban que la hibridación somática permitiera producir nuevas

**Tabla 1:** Ejemplos de híbridos somáticos que exhiben algún carácter de interés agronómico

Híbrido	Carácter
<b>Híbridos simétricos</b>	
<i>Solanum tuberosum</i> x <i>S. brevidens</i>	Resistencia al virus del enrollamiento de la hoja de papa (PLRV), al hongo <i>Phytophthora infestans</i> y a la bacteria <i>Erwinia cartovora</i> .
<i>S. tuberosum</i> x <i>S. circaeifolium</i>	Resistencia a <i>P. infestans</i> y al nemátodo <i>Globodera pallida</i> .
<i>S. tuberosum</i> x <i>S. phureja</i>	Mayor productividad de tubérculo.
<i>S. melongena</i> x <i>S. integrifolium</i> / <i>S. saintwongseis</i>	Resistencia a la bacteria <i>Pseudomonas solanacearum</i>
<b>Híbridos asimétricos</b>	
<i>Brassica oleracea</i> x <i>B. campestris</i>	Resistencia al hongo <i>Phoma lingam</i>
<i>B. napus</i> x <i>B. carinata</i> / <i>B. juncea</i>	Resistencia a <i>P. lingam</i>
<i>Nicotiana tabacum</i> x <i>N. Repanda</i> / <i>N. glutinosa</i>	Hipersensibilidad al virus del mosaico del tabaco (TMV).
<b>Cíbridos</b>	
<i>B. campestris</i> (silvestre) x <i>B. napus</i> / <i>B. campestris</i> (cultivar)	Resistencia al herbicida atrazina y esterilidad masculina citoplasmática (CMS).

especies que expresaran las características deseadas de ambos progenitores. Por ejemplo, por hibridación entre tomate y papa, se buscó producir plantas que desarrollaran tomates en la parte aérea y tubérculos en la raíz (pomato). La experiencia mostró que difícilmente puedan lograrse estos híbridos espectaculares, debiéndose más bien dirigir la técnica hacia la introgresión de pocos genes silvestres en las especies cultivadas mediante hibridación asimétrica o cibridación, manteniendo las características generales del cultivo.

Uno de los factores que dificulta la obtención de híbridos simétricos es la progresiva pérdida de cromosomas de uno de los parentales que normalmente ocurre durante la regeneración. Si bien la fusión de protoplastos permite sortear las barreras precigóticas, siguen existiendo barreras genómicas que resultan en la eliminación espontánea de cromosomas durante las divisiones celulares. El mecanismo que determina la eliminación de cromosomas no está bien comprendido, pero generalmente se retienen los cromosomas del parental que tiene el ciclo celular más corto. Uno de los factores que puede producir incompatibilidad ge-

nómica es el estado de diferenciación de los tipos celulares involucrados en la fusión. Por ejemplo, cuando se fusionan células en fase de crecimiento activo con células del mesófilo, que no se dividen, generalmente se pierden los cromosomas de estas últimas.

Los híbridos somáticos o sexuales entre especies silvestres y cultivadas contienen muchas características no deseadas de la primera, además de aquéllas deseadas. El retrocruzamiento con la especie cultivada es requerido para remover los genes no deseados del parental silvestre, pero ello no siempre es posible debido a que los híbridos suelen ser estériles. La hibridación asimétrica y la cibridación tienen la ventaja de incorporar sólo uno o pocos caracteres provenientes del parental silvestre en la especie cultivada, y producir plantas con mayor fertilidad.

Algunos caracteres deseables están codificados por el genoma extranuclear, tales como la androesterilidad citoplasmática y ciertos tipos de resistencia a enfermedades y a herbicidas. Para estos casos es de particular utilidad la cibridación dado que es un método simple para transferir estos genes.

## 11 Algunos ejemplos

Se han obtenido híbridos intergenéricos de *Panicum maximum* (+) *Pennisetum americanum* (Ozias-Atkins *et al.*, 1986), *Saccharum officinalis* (+) *Pennisetum americanum*, *Oriza sativa* (+) *Echinochloa oryzicola*, *Triticum monococcum* (+) *Pennisetum americanum*, *Festuca arundinacea* (+) *Lolium multiflorum*. El primer caso de regeneración de plantas maduras de híbridos intergenéricos (simétricos y asimétricos) en gramíneas fue el *Festulolium*.

A pesar de los esfuerzos realizados, la aplicación de esta estrategia no ha conducido a la obtención de nuevos híbridos de especies importantes, fundamentalmente debido a problemas de baja o nula fertilidad. Los métodos actuales de transferencia de genes aislados han desplazado el interés en la hibridación somática. Sin embargo son una excelente herramienta para estudiar las interacciones núcleo-citoplasmáticas entre genomas.

## 12 Lecturas recomendadas

- Bhojwani, S.S., Razdan, M.K. 1996. Protoplast isolation and culture. Somatic hybridization and cybridization. En: Plant Tissue Culture: Theory and Practice, Capítulos 12 y 13. S.S. Bhojwani y M.K. Razdan (eds.) Elsevier, Amsterdam. pp.337-406.
- Lindsey y Jones, 1992. Biología celular de la ingeniería genética. En: Biotecnología vegetal agrícola, Capítulo 6. Editorial ACRIBIA, S.A., Zaragoza. Pp.105-142.
- Matsumoto, K. 2001. Híbridos somáticos. Biotecnología Ciencia & Desarrollo, 20, mayo / junio: 26-28.
- Zimmermann, U. 1983. Electrofusion of cells: principles and industrial potential. Trends in Biotechnology, 1,5:149-156.

## II CAPÍTULO 3

### Epigenética y evolución

R.W. Masueli y C.F. Marfil

#### 1. Introducción

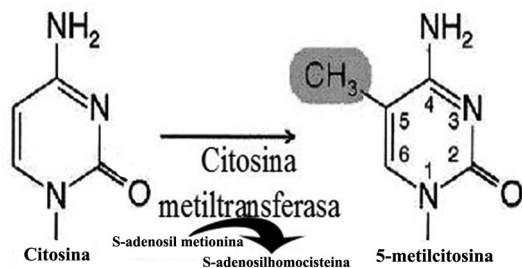
Uno de los principios básicos de la biología es que la variabilidad genética es un pilar fundamental de la evolución. Sin variabilidad la selección natural no podría actuar y la uniformidad genética conduciría a la población o especie a un callejón sin salida en el proceso evolutivo. Este concepto propuesto por Charles Darwin en 1859 fue corroborado por el redescubrimiento de las leyes de Mendel y por trabajos de genetistas un siglo después. Sin embargo, la noción de que la variabilidad genética se encuentra asentada únicamente en la secuencia de nucleótidos ha sido puesta en tela de juicio en los últimos años. Trabajos recientes especialmente en el área de la Genética Molecular muestran que la variabilidad heredable también se explica por cambios que no necesariamente implican alteraciones en las secuencias de bases del ADN. Variaciones heredables en los patrones de expresión génica se pueden producir debido a “mecanismos epigenéticos” con total ausencia de variabilidad genética. El término “epigenética” se refiere a cambios heredables en el fenotipo, y por lo tanto en la regulación génica, que no son causados por alteraciones en la secuencia del ADN. Se han descrito tres mecanismos de codificación de la información que no están basados en la secuencia de ADN. Estos son: a) Metilación del ADN, fundamentalmente de citosinas; b) Modificaciones post-traduccionales de histonas (acetilación, fosforilación, metilación, etc) y c) Mecanismos basados en el ARN. A su vez estos tres mecanismos actúan en forma conjunta para mantener el grado de compactación de la cromatina. En esta sección nos referiremos específicamente a como la hibridación interespecífica y la poliploidía inducen cambios en los patrones de metilación y dan lugar a nuevas fuentes de variabilidad.

#### 2. La metilación del ADN

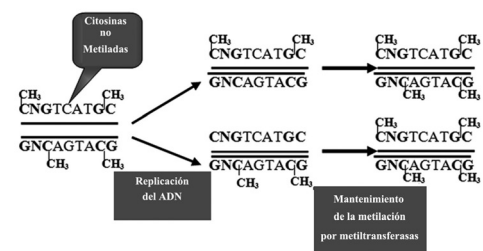
Uno de los mecanismos epigenéticos más estudiados en eucariotas es la incorporación al

ADN de un grupo metilo en la posición 5' del anillo de citosina (<sup>5m</sup>C) (Figura 1a). En plantas la <sup>5m</sup>C se distribuye en sitios donde se encuentran dinucleótidos del tipo CG o trinucleótidos CNG (donde N es cualquiera de los cuatro nucleótidos). A su vez la metilación del ADN puede dividirse en metilación de sitios simétricos y asimétricos. En el primer caso la metilación se produce en sitios CG y CNG donde las citosinas metilables se encuentran de a pares ubicadas en hebras opuestas, mientras que la metilación en sitios asimétricos se produce en citosina ubicadas en un contexto diferente en el ADN. La metilación simétrica es la más común y permite que los patrones de metilación se mantengan a través de la enzima ADN metiltransferasa que cataliza la transferencia de un grupo metilo de la S-adenosil metionina a las citosinas del ADN. En *Arabidopsis thaliana* los genes *CMT3* y *MET1* codifican para las enzimas Cromometiltransferasa 3 y Metiltransferasa 1, respectivamente, la primera metila los grupos CNG y la segunda los sitios CG. Después de la replicación del ADN sólo la hebra

1.A.



1.B.



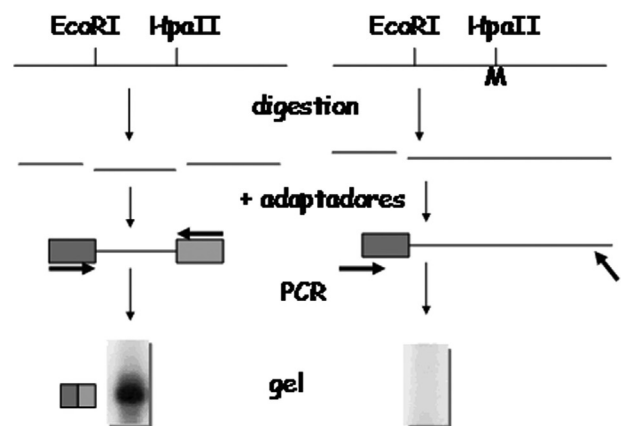
**Figura 1. A.** Agregado del grupo metilo al carbono 5' de la citosina. **B.** Mantenimiento de los patrones de metilación del ADN por las enzimas ADN metiltransferasas durante la replicación del ADN.

molde mantiene el patrón de metilación original. Las enzimas ADN metiltransferasas se encargan de metilar la hebra nueva de acuerdo al patrón de metilación de la hebra molde en los sitios CG y CNG (Figura 1b). Este mecanismo permite que se hereden los patrones de metilación tanto entre células (mitóticamente) como entre generaciones (meióticamente). La proteína DDM1, la cual es un factor remodelador de la cromatina, también participa en el mantenimiento de patrones de metilación específicos a lo largo del genoma. Los mutantes de *Arabidopsis ddm1* (*Deficient in DNA Methylation*) tienen una reducción del 70% en los niveles de <sup>5m</sup>C, lo cual provoca una amplia variedad de defectos en el desarrollo de las plantas al inducir cambios heredables en otros loci. Si bien los patrones de metilación son heredables, se ha demostrado que no son estáticos y que el estado de desarrollo de la planta y condiciones ambientales alteran estos patrones. Por otro lado, la metilación está asociada a cambios en la expresión génica. En general el incremento en la metilación (hipermetilación) de un locus no solo produce la reducción en la expresión o el total silenciamiento del mismo, sino que también reprime transcripcionalmente e inactiva transposones capaces de movilizarse a través del genoma. Se demostró que anulando (*knock out*) los genes *CMT3* y *MET1* se desmetila el genoma activándose elementos transponibles y genes que se encontraban silenciados. Estos estudios mostraron que la pérdida de metilación produce aberraciones tales como cambios morfológicos en hojas y flores, como así también en el tiempo de floración.

### 3. Métodos para analizar la variación epigenética

Los métodos comúnmente utilizados para analizar la variabilidad genética, tales como AFLP (*Amplified Fragment Length Polymorphism*), RFLP (*Restriction Fragment Length Polymorphism*), SNP (*Single Nucleotide Polymorphism*) o SSR (Microsatélites) no son adecuados para medir la variabilidad epigenética. Una modificación de la técnica de AFLP, llamada MSAP (*Methylation Sensitive Amplified Polymorphism*), en la que se reemplaza la enzima de corte frecuente *MseI* por los

isoesquizómeros *HpaII* y *MspI*, sensibles a la metilación, que permiten detectar la metilación externa o interna de las citosinas en el sitio de reconocimiento 5'-CCGG-3' de las enzimas (Figura 2). Otro método desarrollado que permite analizar la metilación es la secuenciación previo tratamiento del ADN con bisulfito, que convierte las citosinas no metiladas en uracilo quedando las citosinas metiladas intactas. Luego la secuencia tratada se amplifica por PCR reemplazando los uracilos por timinas para su posterior secuenciación. A través de estos métodos es posible analizar el estado de metilación de secuencias aleatorias (MSAP) o secuencias específicas (secuenciación por bisulfito) y determinar la variación epigenética en poblaciones naturales.



**Figura 2.** Esquema que representa el análisis de metilación del ADN a través de la técnica de MSAP (*Methylation Sensitive Amplified Polymorphism*). El ADN se corta con las enzimas de restricción *EcoRI* y *HpaII/MspI*, luego se ligan adaptadores específicos para las enzimas y se amplifican los fragmentos con primers que reconocen la secuencia de los adaptadores. Si las citosinas del sitio de reconocimiento de *HpaII* (CCGG) se encuentran metiladas, esquema de la derecha, la enzima no corta y por lo tanto no se amplifica el fragmento generando polimorfismo.

### 4. Epigenética y variación natural

El grado de metilación de genes puede variar entre plantas produciendo cambios en la expresión y nuevos fenotipos que son heredables. A las secuencias génicas que difieren en

el grado de metilación se les llama “alelos epigenéticos o epialelos”. Se han descrito varios ejemplos de epialelos que ocurren naturalmente. Un ejemplo clásico es el gen *CYCLOIDEA* en *Linaria vulgaris*, que codifica para un activador de la transcripción cuyo efecto es la presencia de un fenotipo floral de simetría radial y otro de simetría bilateral. A través del análisis molecular de este gen se demostró que en la naturaleza existían dos epialelos, uno más metilado que produce la simetría radial y otro con menor metilación asociado a la simetría bilateral. Mediciones de los patrones de metilación entre introducciones de *Arabidopsis thaliana* muestran variaciones en los patrones de metilación de más del 34% siendo menores al 1% las variaciones en metilación entre plantas dentro de una misma introducción. En una población natural de la especie silvestre de papa *Solanum ruiz-lealli* se encontró que plantas que tenían una variación genética menor al 1% (medida por AFLP) presentaban variaciones en los patrones de metilación (medida por MSAP) superiores al 18%. Estas variaciones no son producidas por polimorfismo de nucleótidos sino por variaciones en los patrones de metilación. Resultados similares se obtuvieron analizando introducciones de algodón donde la diversidad detectada por MSAP fue superior a la obtenida analizando patrones de RFLP.

### 5. Cambios epigenéticos producidos por hibridación interespecífica y poliploidía

La hibridación interespecífica y la poliploidía son dos mecanismos profundamente relacionados que han jugado roles muy importantes en la evolución vegetal. Una de sus características más sobresalientes es la capacidad de generar una amplia variación fenotípica de gran importancia en la microevolución adaptativa. La poliploidización implica la generación de un organismo con un número aumentado de complementos genómicos completos. Los juegos cromosómicos adicionales pueden provenir de la misma especie (autoploidía) o de una especie divergente (aloploidía), implicando en este último caso un proceso de hibridación interespecífica. A su vez, la hibridación interespecífica se define como la formación de un nuevo organismo por cruzamiento entre dos

especies diferentes aunque relacionadas. Tanto en el caso de la aloploidía como en la hibridación interespecífica dos genomas divergentes se unen en un núcleo híbrido, desencadenando una serie de cambios genéticos y epigenéticos que dan lugar a una gran variabilidad fenotípica (Figura 3). En híbridos aloploidios sintéticos de *Arabidopsis* se demostró que se producen inestabilidades fenotípicas asociadas con cambios en la metilación y en la expresión génica. Por otro lado en generaciones tempranas de híbridos aloploidios de trigo y *Brassica* se reportaron cambios genéticos estructurales tales como eliminacio-

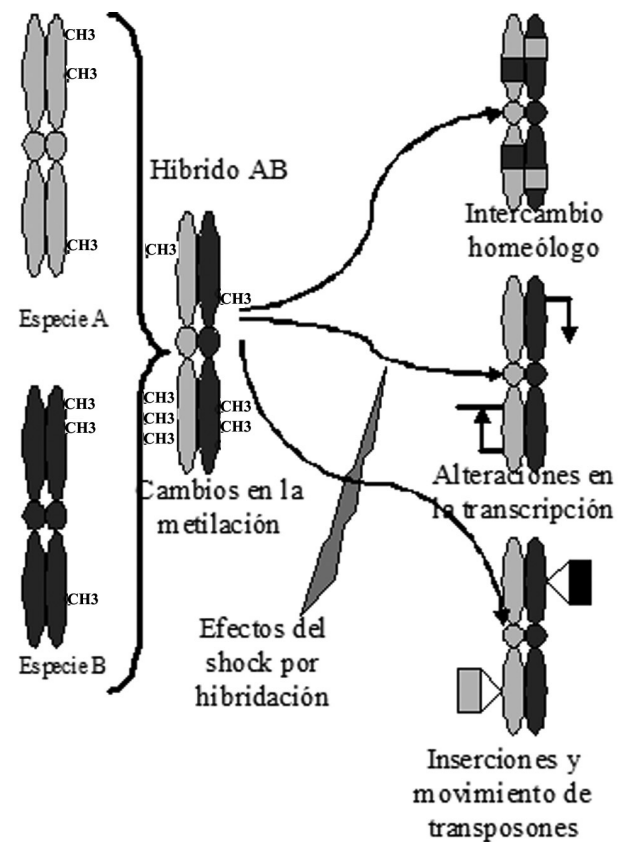
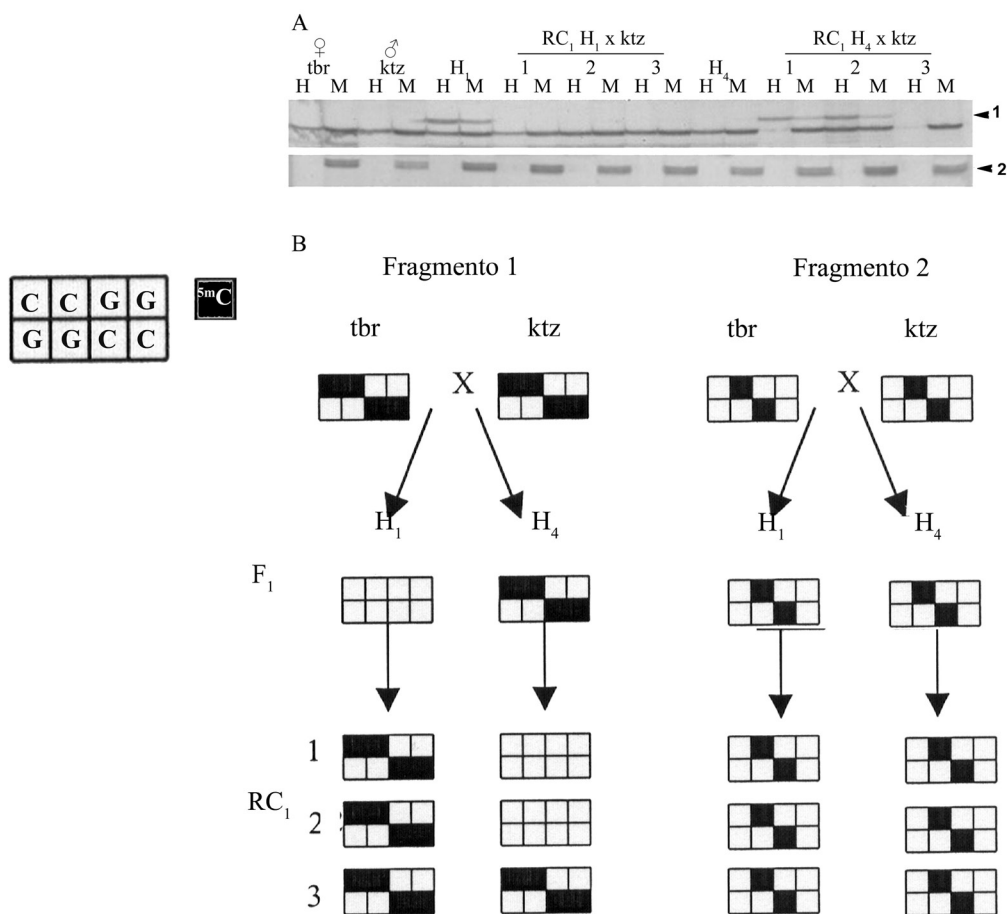


Figura 3. Esquema que representa el “estrés o shock genómico” producido por efecto de la hibridación interespecífica. El cruzamiento de dos especies que han divergido induce alteraciones en los patrones de metilación en el núcleo híbrido que producen cambios en la transcripción y reestructuraciones genéticas tales como: movimiento de transposones, rearrreglos cromosómicos y cambios en la cromatina.

nes de secuencias. En nuestro Laboratorio se obtuvo un híbrido interespecífico entre un haploide de la papa cultivada *Solanum tuberosum* ( $2n=2x=24$ ) y la especie silvestre *Solanum kurtzianum* ( $2n=2x=24$ ). Algunas plantas híbridas presentaban malformaciones florales, similares a las descritas en *Arabidopsis*, que incluían cambios en la morfología floral, pétalos disectados, anteras atrofiadas y estilos retorcidos entre otras. Por otro lado, los híbridos presentaban nuevos fragmentos de AFLP y RAPD (*Random Amplified Polymorphic DNA*) que no estaban presentes en los padres. Así mismo, el análisis de la metilación por MSAP

mostró cambios en los patrones de metilación de las plantas híbridas en relación a las especies progenitoras. Estos resultados obtenidos en distintos híbridos interespecíficos sugieren que existe una correlación entre la remodelación general de la metilación del genoma híbrido, los cambios genéticos y la alteración de la actividad génica y el fenotipo.

En resumen, tanto la poliploidía como la hibridación interespecífica inducen cambios genómicos como producto del estrés generado por la duplicación génica o por la reunión de dos genomas divergentes en un núcleo híbrido. Además de la metilación del ADN otros me-



**Figura 4.** Análisis de la segregación de los patrones de metilación en plantas de las generaciones F<sub>1</sub> y RC<sub>1</sub>. A, porción de los patrones de amplificación MSAP de *S. tuberosum* (tbr), *S. kurtzianum* (ktz), dos híbridos F<sub>1</sub> (H<sub>1</sub> y H<sub>4</sub>) y tres plantas de la retrocruza 1 (RC<sub>1</sub>) H<sub>1</sub> x *S. kurtzianum* y H<sub>4</sub> x *S. kurtzianum*. El ADN extraído fue digerido con *EcoRI/HpaII* (H) y *EcoRI/MspI* (M). El Patrón de metilación del fragmento 1 cambia en las progenies sucesivas. El fragmento 2 muestra un patrón de metilación altamente conservado. B, interpretación gráfica de los fragmentos MSAP 1 y 2 del panel A. Los cuadrados representan el sitio de reconocimiento doble cadena (CCGG) de los isoesquizómeros *HpaII/MspI*. Cuadrados negros indican citosinas metiladas.



canismos tales como, acetilación y metilación de histonas, activación de transposones y regulación por pequeños ARNs y ARN de interferencia pueden causar cambios en la expresión génica de híbridos y poliploides.

## 6. Herencia de los patrones de metilación

La obtención de mutantes que afectan la metilación del ADN ha permitido un amplio desarrollo de esta área. Las anomalías en el desarrollo observadas en mutantes *ddm1* de *Arabidopsis* son transmitidas a la progenie, inclusive en líneas que segregan respecto a la mutación *ddm1*. Estas observaciones morfológicas de plantas *ddm1* se correlacionan con datos moleculares, ya que la reducción en la metilación que afecta a la mayoría de las secuencias del genoma de estas plantas, tanto repeticiones como secuencias de bajo número de copias, pueden ser heredadas establemente a través de la mitosis y meiosis. Esto indica que la información epigenética, en forma de metilación diferencial del ADN, puede ser transmitida en plantas a través de generaciones de reproducción sexual. Otro modelo experimental que ha permitido obtener importante información acerca de la dinámica en los patrones de metilación es la obtención de alopoliploides e híbridos interespecíficos sintéticos. La obtención de progenies a través de cruzamientos sexuales controlados permite una comparación precisa entre la generación parental y la descendencia. En la Figura 4 se ejemplifica cómo a través de análisis MSAP se pueden inferir los patrones de metilación de secuencias CCGG. Estudiando generaciones sucesivas de reproducción sexual (Filial 1-F1; Retrocruza 1- RC1; etc.) se puede evidenciar si ocurren cambios en la metilación y qué estabilidad tienen los mismos. Datos obtenidos en híbridos interespecíficos sintéticos diploides y poliploides de varias especies indican que en el genoma de las plantas existen secuencias particularmente sensibles a remodelar su metilación (Fragmento 1 de la Figura 4). Por otro lado, hay secuencias con patrones de metilación altamente conservados evolutivamente, ya que especies diferentes comparten un mismo patrón, el cual a su vez se hereda invariablemente en todas las plantas obtenidas en generaciones sucesi-

vas de reproducción sexual (Fragmento 2 de la Figura 4). Desde un punto de vista evolutivo la variabilidad generada por mecanismos epigenéticos amplía la variabilidad genética ya que aparecen nuevos epiallelos que pueden ser seleccionados en poblaciones naturales. Además, las variantes epialélicas son potencialmente reversibles y generarían fenotipos más flexibles que se adaptarían a ambientes cambiantes. A partir de este tipo de evidencias, un concepto que está ampliamente documentado para fenómenos genéticos, debe ser ampliado de manera que abarque la información epigenética: en cada generación se reconstituye un reservorio epigenético enteramente nuevo, del cual surgen los individuos que son los blancos de la selección natural en esa generación.

## 6. Lecturas recomendadas

- Adams K.L. and Wendel J.F. 2005. Novel patterns of gene expression in polyploid plants. *Trends in Genetics* 21:569-543.
- Jablonka E. and Lamb M. 1995. Epigenetic inheritance and evolution. The Lamarckian dimension. Oxford University Press, New York, 346 pp.
- Kalisz S. and Purugganan M.D. 2004. Epialleles vs DNA methylation: consequences for plant evolution. *Trends in Ecology and Evolution* 19:309-314.
- Madlung A. and Comai L. 2004. The effect of stress on genome regulation and structure. *Annals of Botany* 94:481-495.
- Marfil C.F., Camadro E.L. and Masuelli R.W. 2009. Phenotypic instability and epigenetic variability in a diploid potato of hybrid origin, *Solanum ruiz-lealii*. *BMC Plant Biology* 9:21.
- Rapp R.A. and Wendel J.F. 2005. Epigenetics and plant evolution. *New Phytologist* 168:81-91.
- Riddle N.C. and Birchler J.A. 2003. Effects of reunited diverged regulatory hierarchies in allopolyploids and species hybrids. *Trends in Genetics* 19:597-600.



## II. CAPÍTULO 4

### Mutagénesis, TILLING y EcoTILLING

Alberto Prina, Alejandra Landau,  
María Gabriela Pacheco y Esteban H Hopp

#### 3.1. Mutagénesis clásica

##### 3.1.1. Introducción

El concepto de mutación en Biología fue introducido a principios del siglo pasado por de Vries, para designar a los cambios heredables de aparición súbita. Hoy en día las mutaciones se explican por alteraciones en el ADN, que van desde cambios en una o unas pocas bases (mutaciones de punto) hasta pérdidas, adiciones o translocaciones de grandes porciones de ADN (cambios estructurales) o incluso cromosomas y genomas enteros; también podrían incluirse entre ellas las alteraciones heredables de naturaleza epigenética, por ejemplo los cambios en la metilación del ADN (ver Grant-Dwonton y Dickinson, 2005). Las mutaciones son el origen primario de la variabilidad genética y, por lo tanto, cierto control sobre su frecuencia y/o espectro puede considerarse una herramienta de gran valor para el mejoramiento de las plantas cultivadas. A fines de la década del 20, Stadler desarrolló magistralmente las bases de la mutagénesis experimental en plantas cultivadas, hecho que fue contemporáneo a los trabajos pioneros de Muller sobre mutagénesis en *Drosophila*. El primer éxito comercial consistió en una mutante de tabaco de hoja clara y de gran calidad obtenida por Tollenar (Holanda) que llegó a cubrir una importante área de cultivo en Indonesia a mediados de la década del 30. Hoy en día existen en el mundo miles de cultivares inscriptos obtenidos por esta metodología. Se incluyen en esta lista, cultivos de gran importancia económica, como los principales cereales, oleaginosas, así como, numerosas especies de hortalizas y cultivos industriales. Los caracteres que han sido mejorados son muy diversos: tolerancia a herbicidas, a estrés bióticos o abióticos, calidad nutricional, industrial, etc. (para más información ver IAEA 1995, Datta 2005). Además, la ampliación de la variabilidad disponible por medio de las mutaciones inducidas

ha contribuido considerablemente a dilucidar procesos biológicos fundamentales para la vida vegetal y para la productividad de los cultivos.

Las investigaciones en busca de mejorar la efectividad y la especificidad de los tratamientos mutagénicos sobre las plantas cultivadas, y de procedimientos de manejo y selección más eficientes, tuvieron un gran auge en las décadas del 50, 60 y 70.

Hoy existe un renovado interés por esta temática al que sin duda contribuyó el enorme conocimiento generado por la biología molecular, que alejó viejos fantasmas sobre la variabilidad inducida, y el reciente desarrollo de técnicas moleculares de gran capacidad de análisis como la técnica de genética reversa denominada TILLING (del inglés, *Targeting Induced Local Lesions in Genomes*).

A continuación se discuten los principales factores a tener en cuenta para la aplicación eficiente de la metodología clásica de mutaciones inducidas en plantas cultivadas. La descripción detallada de protocolos puede consultarse por ejemplo en: Neuffer, 1993; IAEA, 1995; Koornneef, 2002.

##### 3.1.2. Tratamientos mutagénicos

###### 3.1.2.1. Agentes mutagénicos

Los agentes artificiales utilizados en mutagénesis clásica pueden dividirse según su naturaleza en dos grandes grupos: físicos o químicos. Otra manera artificial de incrementar la tasa de mutaciones es a través de las condiciones especiales del cultivo *in vitro* de células o de tejidos (variación somaclonal). También existen factores de naturaleza biológica que causan algún tipo de inestabilidad genética mayor a la habitual de la especie, como la producida por ejemplo por genes relacionados con los sistemas de reparación de ADN o por la activación de transposones.

###### Agentes físicos

Entre éstos tenemos distintos tipos de radiaciones ionizantes como los rayos X, los rayos gamma, los neutrones, los protones y las partículas alfa, cada una de ellas con diferente poder de ionización y de penetración.

Las radiaciones ionizantes al atravesar la materia producen iones radicales inestables y electrones libres, que a su vez originan una

serie de reacciones químicas directas (por ionización de un átomo en una molécula), o indirectas (por reacción de las macromoléculas con productos radiolíticos, mayormente con los radicales OH del agua). Entre otros cambios químicos que afectan el ADN ocurre la deaminación de bases y producción de 8-hidroxiadenina y de varios glicoles de timina, oxidación de alcoholes de la desoxirribosa y roturas de las uniones carbono-carbono. Los tratamientos con radiaciones ionizantes implican habitualmente altos niveles de daño cromosómico, originando roturas simples y dobles de las cadenas de ADN, y efectos fisiológicos deletéreos, que conducen a una alta letalidad celular y a una alta esterilidad de las plantas de la primera generación, todo lo cual limita la obtención de una alta tasa de mutaciones en la segunda generación.

La magnitud y el tipo de daño ocasionado por las radiaciones ionizantes son marcadamente influidos por ciertas características del material tratado, como el estado metabólico y el contenido de agua y de oxígeno. Así también, los efectos de determinada dosis pueden variar de acuerdo al mayor o menor tiempo en que se aplica (tasa de irradiación), desde unos pocos segundos o minutos (irradiación aguda) hasta meses o años (irradiación crónica). Los efectos letales de determinada dosis pueden disminuirse, por ejemplo, cuando se aplica repartida en distintos momentos, con intervalos de uno o varios días de recuperación o incluso dividida en aplicaciones agudas realizadas en años sucesivos (tratamientos recurrentes).

Las radiaciones ionizantes que más se han utilizado para inducir variabilidad con fines de fitomejoramiento son los rayos X y los gamma, con longitudes de onda del orden de fracciones de nanómetros (nm) a varios nm. Si bien los rayos gamma tienen mayor poder de penetración que los X, los dos pueden penetrar desde varios milímetros a centímetros, siendo por lo tanto adecuados para tratamientos aplicados sobre semillas o yemas vegetativas. Ambas son radiaciones de origen electromagnético de baja energía lineal liberada al atravesar la materia, por lo que son clasificadas como no densamente ionizantes, en contraposición con las densamente ionizantes, como por ejemplo

los neutrones ligeros. Las consecuencias genéticas típicas de las radiaciones ionizantes son las aberraciones cromosómicas, como las deleciones y translocaciones, que han resultado de utilidad para la ubicación de numerosos genes sobre los cromosomas. Las deleciones producidas por las radiaciones densamente ionizantes son de mayor tamaño y se agrupan, formando *clusters*, siendo más difíciles de reparar. Las de mayor tamaño difícilmente sortean el filtro de la gametogénesis, por lo que en muchos casos no son transmitidas sexualmente a las progenies. Las deleciones conducen a *knock outs* génicos y han sido utilizadas por ejemplo para inactivar genes que promueven la susceptibilidad a algunas enfermedades en trigo. En combinación con la tecnología de microarreglos pueden permitir la identificación de genes candidatos cuya expresión pueden cambiar drásticamente. En muchos casos, las deleciones pueden ser localizadas por medio de *Southern blot* o PCR (reacción en cadena de la polimerasa).

Por otro lado, las propiedades de las radiaciones ionizantes para romper los cromosomas y permitir así translocaciones han sido utilizadas desde hace más de 50 años para transferir segmentos cromosómicos entre distintas especies; técnica que se ha utilizado para mejorar la resistencia a enfermedades en trigo.

La luz ultravioleta es también una radiación de origen electromagnético, pero a las longitudes de onda habitualmente utilizadas no se considera ionizante. Su longitud de onda es varios miles de veces superior a la de los rayos X y gamma, y por su baja penetración sólo es utilizable sobre materiales muy delgados, como por ejemplo granos de polen y cultivos celulares o de tejidos. La luz ultravioleta origina un daño muy específico sobre el ADN, constituido mayormente por la formación de dímeros de timina que a su vez son eficientemente reparados por sistemas celulares muy especializados. Su efecto biológico varía considerablemente de acuerdo a la longitud de onda, siendo la de mayor absorción por el ADN aquella en el rango de 250 a 290 nm.

#### Agentes químicos

Las primeras investigaciones sobre la capacidad mutagénica de las sustancias químicas

datan de principios del siglo pasado, pero su uso en fitomejoramiento se concretó muchos años después a través de los trabajos que facilitaron el uso al disminuir los requerimientos en equipamientos e instalaciones.

Existen muchas sustancias con capacidad mutagénica, sin embargo los más utilizados con fines prácticos son los denominados supermutágenos, entre ellos varios agentes alquilantes y la azida sódica, que pueden aumentar varios cientos de veces las tasas de mutación espontánea.

A diferencia de los tratamientos con radiaciones ionizantes donde las moléculas son afectadas indiscriminadamente, los mutágenos químicos suelen ser más específicos. Los ejemplos clásicos son el ácido nitroso y los análogos de bases, productores de sustituciones del tipo transición, y las acridinas que producen mutaciones de desfase de lectura. Los mutágenos químicos que más se han utilizado en fitomejoramiento son los agentes alquilantes, denominados así por incorporar grupos alquilo a las macromoléculas. A este grupo pertenecen los alquil alcanos sulfonatos, como el etil metano sulfonato (EMS) y el metil metano sulfonato (MMS), los dialquil sulfatos, como el dimetilsulfato (DMS) y el dietilsulfato (DES); y los compuestos nitrosos, como la N-metil-N-nitroso-urea (MNU o MNH) y la N-etil-N-nitroso-urea (ENU). Se trata de compuestos electrofílicos que reaccionan con átomos que tienen un par de electrones libres como S, N y O (en ese orden, de mayor a menor fuerza nucleofílica), y de esta manera actúan sobre las bases y los grupo fosfatos del ADN y también sobre las proteínas. Los agentes de mayor reactividad, como el MMS y el DMS, tienen afinidad por los centros más nucleofílicos y en el ADN atacan mayormente el nitrógeno 7 de la guanina (N7-G) y muy poco al O6-G o al grupo fosfato. Por el contrario, los agentes de menor reactividad como ENU, tienen un mayor espectro de sitios de alquilación y afectan en mayor proporción al O6-G. Esto tiene su importancia si consideramos que la O6-alquilG sería la principal causa de apareamientos erróneos de G con T, mientras que la N7-alquilG origina depurinación y errores de reparación, que tienden a derivar en roturas cromosómicas y ocasionalmente en

deleciones de pares de bases. La MNU, también de baja reactividad, además de las alquilaciones mencionadas, forma mayor cantidad de fosfodiésteres, los que pueden derivar en roturas de cadenas de ADN y contribuir a la letalidad. El mutágeno químico que más se ha utilizado es el EMS, que habitualmente se asume como productor de mutaciones de punto del tipo transición G/C a A/T. En relación a esto es interesante advertir que hay 5 aminoácidos que en sus codones sólo tienen G o C en posiciones sinónimas, por ejemplo lisina (AAA o AAG) y tirosina (UAU o UAC), por lo que no podrían ser cambiados por otros aminoácidos mediante transiciones del tipo G/C a A/T.

Un mutágeno muy eficiente en algunas plantas cultivadas, como cebada, arveja, trigo y arroz, es la azida sódica. Esta droga no es de por sí mutagénica sino que necesita de la vía de la síntesis de L-cisteína para transformarse en el metabolito mutagénico azido-alanina. La azida no se ha visto como mutagénica en mamíferos y por lo tanto desde el punto de vista de la bioseguridad, no sería peligrosa como mutágeno en humanos, más allá de su reconocida toxicidad como veneno de la respiración oxidativa. En cebada, la azida sódica induce esencialmente sustituciones de bases, en su mayor parte transiciones, existiendo resultados divergentes en la bibliografía sobre su tipo: G/C a A/T o viceversa.

Tanto la azida como el EMS, por su propiedad de inducir mutaciones de punto, son mutágenos indicados para la obtención de series alélicas y ambos han sido utilizados para generar poblaciones para análisis de genética reversa como el TILLING (ver más adelante). Actualmente, las técnicas de TILLING están aportando una enorme cantidad de información sobre el efecto molecular de los mutágenos. Esto ha permitido detectar que más del 50 % de las mutaciones inducidas por EMS o azida sódica son de sentido cambiado, y más del 20 % son mutaciones silenciosas. En términos generales aplicando estos agentes en varias plantas cultivadas se han identificado polimorfismos en nucleótidos simples (SNPs) con una frecuencia que va de 1/50 a 1/500 kb.

Se cree importante señalar que además de las propiedades particulares de los mutágenos

hay factores biológicos que influyen sobre la tasa y el espectro de mutaciones, como la fase en que se encuentre la célula en el momento de ser tratada (G1, S o G2), la composición de bases y el tamaño del gen y, dependiendo del tipo de daño, la eficiencia de los distintos sistemas de reparación celular y las condiciones en que estos actúen. Todo esto hace que dosis similares de un mismo mutágeno puedan tener efectos biológicos muy dispares.

### 3.1.2.2. Materiales tratados

#### Órganos seleccionados para el tratamiento

En el caso de especies propagadas por semillas, éstas han sido el órgano preferido a tratar, ofreciendo la ventaja de un manejo sencillo que permite trabajar con grandes cantidades de material en condiciones relativamente fáciles de controlar. La principal dificultad del tratamiento de semillas radica en la constitución multicelular del embrión, que hace que los individuos originados a partir de la semilla tratada resulten quiméricos, es decir compuestos por células o clones celulares de distinta constitución genética. Esto es así porque en cada célula, los cambios genéticos o mutaciones que potencialmente puede producir el mutágeno pueden tener muchísimas alternativas diferentes, haciendo cercana a cero la probabilidad de que un grupo de células meristemáticas muten simultáneamente en el mismo gen y posición nucleotídica. Los clones celulares que forman la planta adulta tienen un crecimiento irregular de difícil predicción que complica el diseño de un manejo eficiente del material experimental. A pesar de esto, el tratamiento de semillas ha dado numerosísimos éxitos aplicados al mejoramiento de diversos cultivos. Como se señaló anteriormente los tratamientos físicos de semillas deben hacerse con radiaciones que tengan suficiente penetración, por ejemplo rayos X, rayos gamma, o neutrones ligeros, así como también se deberá prestar atención a los diversos factores que influyen en el efecto biológico de las radiaciones. Por otro lado, los tratamientos químicos de semillas se aplican mediante inmersión en solución acuosa y con agitación constante durante varias horas. Los tratamientos cortos suelen aplicarse después de un remojo en agua que potencia marcadamente el

efecto de los mutágenos. Luego del tratamiento químico, la semilla se lava cuidadosamente y se seca, y una vez eliminada la humedad superficial estará lista para la siembra. En la semilla tratada se inicia la generación M1, esta nomenclatura denota la primera generación de mutantes y se diferencia de una F1 en que NO es resultado de un cruzamiento, sino de un tratamiento mutagénico.

En las plantas de propagación vegetativa los equivalentes a las semillas son los órganos de multiplicación como yemas, rizomas, etc. Éstos son también órganos multicelulares y la complejidad de crecimiento de los sectores mutados es aún mayor que en las semillas. No obstante, esto no ha sido impedimento para que, en la práctica, se hayan obtenido numerosos cultivares mejorados. En general los órganos vegetativos son más sensibles que las semillas a los efectos tóxicos de los mutágenos químicos y resulta más difícil lograr aplicaciones homogéneas. Por ello, las radiaciones ionizantes con suficiente penetración han sido las más utilizadas en estos casos. Por medio de irradiación de yemas con altas dosis puede reducirse el número de células del vástago y así tender a la obtención de vástagos mutantes no quiméricos.

El problema del quimerismo de las plantas M1 puede evitarse tratando órganos unicelulares como óvulos, polen o proembriones unicelulares, y en el caso de plantas de multiplicación vegetativa, por tratamiento de yemas adventicias o de cultivos *in vitro* de células o tejidos. La utilización de mutagénesis inducida en combinación con el cultivo *in vitro* presenta además la ventaja de permitir una selección sobre un gran número de células mutagenizadas. Si bien esta selección está limitada a caracteres que puedan ser evidenciados en estas condiciones, son varios los casos que han resultado exitosos, por ejemplo con respecto a mejorar la tolerancia a temperaturas extremas y resistencia a algunas enfermedades y herbicidas.

#### Elección de los genotipos parentales

La filosofía de la mutagénesis inducida aplicada al mejoramiento es en cierta medida similar a la de la retrocruza, en el sentido que se

busca cambiar un gen (o unos pocos) conservando un fondo genético determinado. Por otro lado, se diferencia claramente de ella en que la variante alélica de interés es generada *de novo* y, por lo tanto, puede ser una variante no presente en el germoplasma actual de la especie. Es aconsejable partir de genotipos con un buen fondo genético, de manera que puedan ser mejorados por cambios en uno o unos pocos genes. Por ejemplo, puede mejorarse la resistencia genética a determinada enfermedad conservando un fondo genético de buenas características agronómicas y/o de calidad industrial o comercial. También por cambios en un solo gen puede adaptarse un genotipo a una nueva región mediante mutantes para el largo de ciclo, o por ejemplo cambiar drásticamente la calidad industrial por mutantes en las características del aceite de la semilla.

### **3.1.3. Evolución de los daños de los tratamientos mutagénicos**

#### **3.1.3.1 Desarrollo de los sectores mutantes**

Tal como se ha descrito, el tratamiento de órganos multicelulares originará plantas M1 quiméricas en las que las mutaciones inducidas sólo ocuparán un sector. El tamaño y la forma de estos sectores mutantes dependerán de varios factores:

A) *Factores inherentes al material tratado*, como el estado de desarrollo y la estructura del embrión o del meristema vegetativo y la ontogenia de la especie. Esto ha sido estudiado principalmente mediante marcadores clorofílicos y mutantes del polen, y más recientemente por técnicas basadas en la recombinación de ADN para activar marcadores específicos. En plantas con alta capacidad de macollaje, como cebada o trigo, la tendencia general es que el tejido germinativo de las espigas principales provenga de varias células iniciales (generalmente de 4 a 12) mientras que puede encontrarse que más de una espiga lateral provenga de la misma célula inicial.

B) *Factores relativos al tratamiento* como el tipo de mutágeno, dosis y condiciones de aplicación. En general los tratamientos que inducen alta letalidad celular, como por ejemplo altas dosis de radiaciones ionizantes, reducirán el número de células meristemáticas a partir de

las cuales se formará la planta M1 y por lo tanto en éstas tenderá a aumentar el tamaño de cada uno de los clones celulares.

C) *Factores relativos al mutante inducido*, especialmente en lo que se refiere a su expresión y competitividad. En especies reproducidas sexualmente los sectores mutantes tendrán que sortear dos tipos de filtros selectivos para llegar a la segunda generación (M2). En esta generación, salvo excepciones, las plantas tendrán una constitución genética homogénea en todo el soma. En el caso de plantas autógamas, la segregación de homocigotas mutantes en M2, permitirá la observación de los alelos mutantes que tengan expresión recesiva. En el caso de las alógamas para que esto ocurra será necesario realizar la autofecundación forzada de las plantas M1. En el maíz, por ser una planta diclino monoica en la que los clones celulares que forman las flores femeninas y masculinas se originan en distintas células iniciales de la semilla, la autofecundación de las plantas M1 sólo servirá para el control genealógico pero no para la segregación de homocigotas mutantes en la M2. Esto hace que la aplicación de mutágenos sobre órganos unicelulares que evitan la formación de plantas quiméricas, como el polen, resulte especialmente interesante en maíz.

Volviendo a los filtros selectivos, primero ocurrirá la selección somática o diplóntica a nivel de los tejidos quiméricos del soma de la planta M1. La competencia de un sector mutante con su entorno puede comenzar muy temprano a nivel de la célula mutante original. Luego a nivel de la gametogénesis de las plantas M1 ocurrirá la selección gamética o haplóntica. Lo arriba mencionado es válido para los genes localizados en el núcleo, pero en el caso de genes localizados en las organelas citoplasmáticas (plástidos y mitocondrias), que obedecen a reglas de la herencia más flexibles, el desarrollo de las quimeras es más complejo que para los nucleares, sumado a que la competencia del alelo mutante también puede ocurrir a nivel intracelular entre los distintos genomas de una organela. Una representación esquemática que ilustra la evolución en uno y otro caso puede encontrarse en Kirk y Tilney-Bassett (1978).

En las plantas multiplicadas vegetativamente, la selección gamética y la recombinación

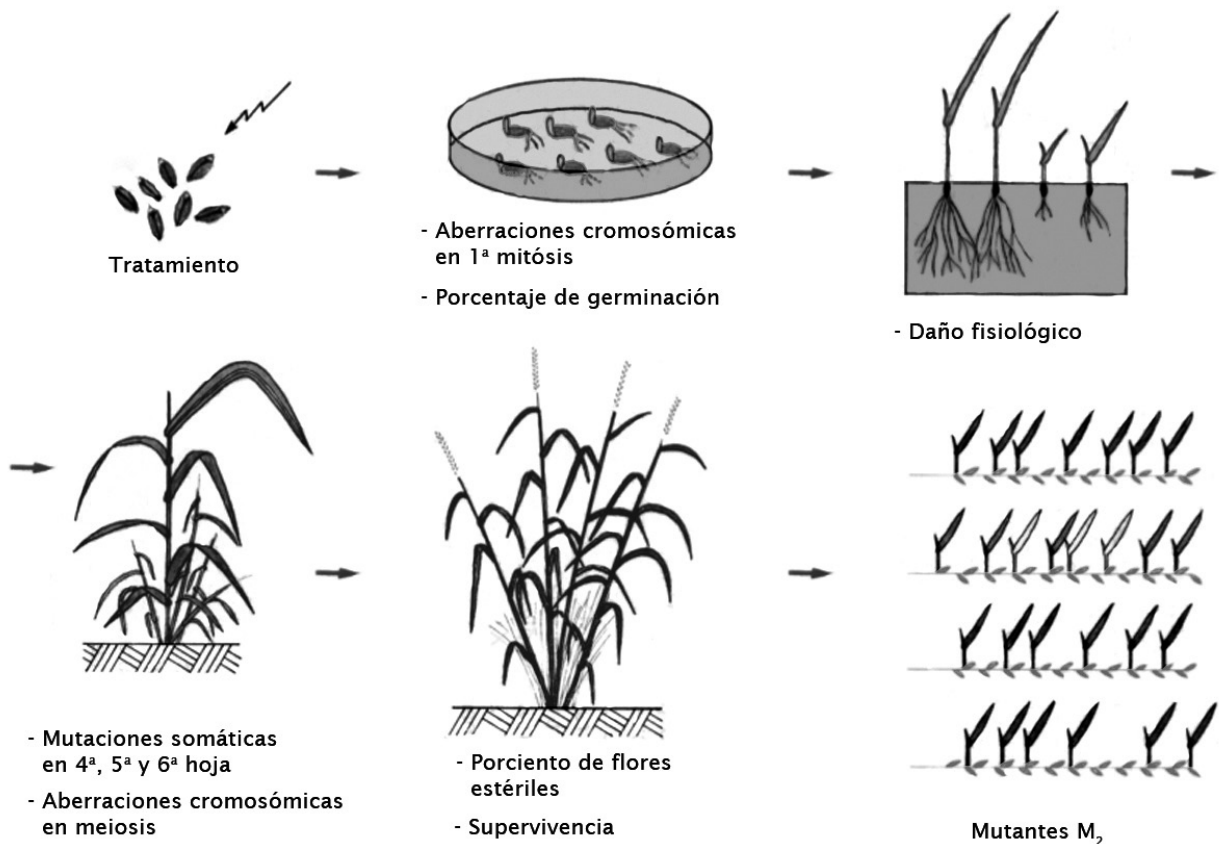
sexual estarán ausentes, a lo que se agrega que en las yemas, las células meristemáticas tienen una organización del tipo *tunica corpus*, es decir que están organizadas en capas histogénicas que cubren un cuerpo central y que al crecer se comportan, en cierta manera, como independientes. Por otro lado, estas plantas son comúnmente poliploides y con un alto nivel de *loci* heterocigotas, todo lo cual hace que un espectro muy diferente de mutantes puede estar disponible para el mejoramiento vegetal. La influencia de los genes al estado heterocigota sobre la expresión de los sectores mutantes a nivel somático fue estudiada por Prina y Favret (1988) utilizando como modelo la cebada.

Por otro lado, la poliploidía puede influir notablemente la expresión y vitalidad de las mutantes, los genotipos poliploides tienen mayor capacidad de soportar cambios drásticos como grandes deleciones de cromosomas o incluso

de cromosomas enteros. En éstos habitualmente se observa una menor proporción de cambios fenotípicos drásticos que en los diploides.

### 3.1.3.2. Expresión de los distintos daños

El esquema de la Figura 1 indica los distintos momentos en que pueden observarse los daños producidos por el tratamiento de semillas (para más detalles ver Prina, 1989). Las proporciones de los distintos daños tienen tendencias definidas de acuerdo al mutágeno y la manera en que se aplique. Si comparamos mutágenos como el EMS y la azida sódica con los rayos X, a iguales niveles de daño en las plantas M1 (sectores somáticos, aberraciones cromosómicas, letalidad), con los primeros tendremos tasas de mutaciones muy superiores en la generación M2. Tal información es útil para la elección de las dosis más adecuadas en base a observaciones en las plantas M1. En



Esquema de las principales etapas en que se registran los efectos de los tratamientos mutagénicos.



principio puede pensarse que a mayores dosis se obtendrá un mayor número de mutantes en la M2, pero la elección no es tan sencilla. En el caso de las radiaciones ionizantes suele recomendarse el uso de la dosis letal 50, buscando un compromiso entre la cantidad de plantas sobrevivientes en la M2 y la cantidad de mutaciones que ellas porten. Con los químicos los efectos letales suelen no tener una correlación directa con los efectos mutagénicos. Con variaciones importantes entre especies e incluso entre genotipos de una misma especie, las más recomendables son las concentraciones intermedias. En los tratamientos de azida se ha visto una influencia muy marcada del pH de la solución. Con estos mutágenos, el daño más limitante suele ser la esterilidad de la M1. Es importante señalar que no siempre es recomendable utilizar los tratamientos que den las más altas tasas de mutaciones en M2, ya que si bien esto aumenta la probabilidad de obtener una mutante buscada analizando una determinada cantidad de material, también aumenta la probabilidad de inducir mutaciones simultáneas en una misma célula lo que luego puede requerir un trabajo considerable para eliminar por retrocruza los alelos no deseados.

#### **3.1.4. Manejo y selección**

La semilla M1 usualmente se siembra a campo semilla por semilla para permitir la individualización de las plantas a la cosecha. En lo posible, debe realizarse en lote aislado, geográfica y/o temporalmente de otras parcelas del cultivo. Para mutantes de efecto marcado que son fácilmente distinguibles a nivel de individuos habitualmente la selección comienza en la M2. En el caso de mutantes que conducen a cambios fenotípicos leves o cuya expresión tiene una importante influencia ambiental, es más confiable realizar la selección entre familias de generaciones más avanzadas.

En forma simplificada podemos describir tres tipos básicos de manejo del material mutagenizado para pasar de M1 a M2:

**1- Manejo genealógico por progenies de espigas, panojas o ramas:** fue propuesto por Stadler en 1930, en los albores de la mutagénesis experimental. Implica un mayor trabajo inicial, pero tiene el beneficio de que permite recono-

cer las mutaciones inducidas por el tratamiento mutagénico de otras preexistentes, o de las originadas por mezcla de semillas o de polen. Por otro lado, en el caso de hallarse mutantes idénticas permite evaluar la probabilidad de que se hayan originado en el mismo o en distintos eventos mutacionales.

**2- Manejo por conjunto o pool de 1 semilla por espiga, panoja o rama:** en este caso la cosecha se limita a una semilla de cada una de las inflorescencias principales de las plantas M1 con las que se formará un conjunto (*pool*). Se basa en que cada una de las semillas del *pool* provendrá de células iniciales diferentes, y por lo tanto en un mismo *pool* no se encontrarán mutantes provenientes de un mismo evento mutacional. La formación de varios *pools* como el descrito aumentará la probabilidad de aislar una mutante determinada pero, por otro lado, no permitirá distinguir sobre el origen mutacional de mutantes idénticas que aparezcan en *pools* diferentes. Este manejo es adecuado cuando se quiere seleccionar mutantes de efectos moderados y/o de expresión altamente influida por el ambiente para lo que es recomendable comenzar la selección en familias de la generación M3.

**3- Manejo en masa:** se cosechan en masa o *bulk* todas las plantas M1 de un mismo tratamiento. También puede limitarse el tamaño de los *bulks* haciendo varios por tratamiento. Para este manejo se recomienda una siembra densa para limitar el crecimiento de las plantas M1 y así aumentar el número de los sectores M1 que serán analizados en M2 los que, de esta manera estarán más homogéneamente representados. Este manejo sólo es aplicable cuando se dispone de una metodología de selección individual barata y de aplicación masiva. Al principio es menos laborioso que los anteriores, pero no distingue la variabilidad inducida de la preexistente, ni las mezclas de semilla o polen. Además, puede incrementar notablemente el trabajo posterior si es que se obtienen muchas mutantes similares y se quiere distinguir si se originaron de manera independiente, o si provienen del mismo evento mutacional.

En cuanto a la cantidad de material a tratar es obvio que a mayor cantidad de sectores mutantes representados en el momento de la

selección, será mayor la chance de aislar una mutante determinada. Cada uno de los manejos mencionados estará limitado en forma diferente por el trabajo de conducción. Los cálculos teóricos sobre la cantidad de material a manejar en M1 y M2 son de relativa validez dada la multiplicidad de factores a tener en cuenta, muchos de los cuales son sólo parcialmente controlables o predecibles. Habitualmente estos cálculos se basan en probabilidades de *knock out* génicos, considerando a todos los genes con igual chance de ser afectados. La crianza de la M1 habitualmente requiere poco esfuerzo y espacio en comparación con las otras etapas, por lo que es recomendable criar material adicional para asegurar una población M2 suficientemente grande, lo que además puede permitir seleccionar entre tratamientos de acuerdo a observaciones sobre la M1. Algunos autores señalan que al trabajar con *Arabidopsis*, a menudo se prefiere criar gran cantidad de plantas M1 y luego someter a la selección unas 2000 a 125000 plantas M2. Mientras que en el caso del maíz, que por ser alógama requiere una autofecundación controlada para poder seleccionar mutantes recesivas, se considera que 1000 plantas M1 no quiméricas obtenidas por tratamientos de polen, pueden ser suficientes para tener chance de obtener al menos una mutante recesiva por *locus*. En autógamias, donde obviamente la autopolinización de las plantas M1 no insume un trabajo adicional, es recomendable incrementar el número de plantas M1 que pasan a la generación M2 a varios miles.

### 3.1.5. Conclusiones

Podemos señalar entonces, tres pasos fundamentales y complementarios en la mutagénesis inducida aplicada al fitomejoramiento, ellos son:

- 1) Realizar tratamientos mutagénicos apropiados para originar una importante cantidad de células mutantes, pero en lo posible con una o unas pocas mutaciones cada una (evitando la multiplicidad de cambios mutacionales en una misma célula) y, adicionalmente, con bajos niveles de otros daños que afecten la viabilidad y/o vitalidad de las células y/o de los individuos mutantes.

- 2) Hacer un manejo eficiente del material

mutagenizado en función del crecimiento de los sectores mutados en las plantas de la primera generación y, de ser posible, con un control genealógico adecuado, que permita distinguir sobre el origen de mutantes similares.

- 3) Una vez que los alelos mutantes puedan ser detectados, aplicar una metodología de selección eficiente que permita analizar grandes cantidades de material.

Si bien la mutagénesis inducida tradicional es un proceso con un amplio espectro de resultados posibles y dependiente de múltiples factores, que pueden ser relativamente controlados, lejos está de ser una técnica que depende totalmente del azar. Al éxito de esta metodología contribuirán los conocimientos de la biología de la especie y del genotipo particular a tratar, la información disponible sobre su respuesta a los tratamientos mutagénicos, el conocimiento de las bases genéticas del carácter a mejorar, etc. Sin duda que los resultados de esta metodología se verán potenciados por un trabajo multidisciplinario de genetistas, fisiólogos y mejoradores, sin descartar el aporte que puedan hacer muchas otras especialidades. El potencial de la mutagénesis clásica sin duda se ve hoy en día enormemente acrecentado por su uso en combinación con las nuevas técnicas de la biología molecular y la biotecnología.

### 3.2. TILLING y EcoTILLING

TILLING (del inglés *Targeting Induced Local Lesions in Genomes*) es una técnica de búsqueda de mutaciones focalizada en un gen o secuencia nucleotídica conocida, que permite el análisis de muestras de muchos individuos que provienen de una población previamente mutagenizada en forma artificial; o que pueden portar variación natural sin necesidad de tratamientos mutagénicos, en este último caso se la denomina EcoTILLING. Dado que se la utiliza para averiguar la función de un gen o secuencia incógnita, se la considera como una técnica de genética inversa (a partir de modificaciones genotípicas se evalúan los efectos fenotípicos de las mismas para descifrar la funcionalidad de la secuencia), con ventajas sobre el *knock out* tradicional porque los sistemas de reemplazo alélico son ineficientes en plantas superiores. Otra ventaja es que la mutagénesis con

mutágenos como el EMS o la azida sódica no necesariamente anulan completamente la función del gen, como sí lo hace el *knock out* o el *knock down* que se obtiene cuando se utiliza ARN interferente o cuando se utilizan radiaciones ionizantes que producen deleciones (como tratamiento mutagénico). Esto, potencialmente, permite evaluar los efectos de reemplazo de cada uno de los nucleótidos individuales del gen.

El TILLING permite identificar estas mutaciones puntuales aunque para poder utilizarlo se requiere conocer previamente la secuencia nucleotídica del gen o segmento genómico que se desea estudiar; sin embargo, no es necesario conocer la secuencia completa del genoma. El método fue desarrollado originalmente en *Arabidopsis thaliana* y su uso se extendió posteriormente a plantas cultivadas como el maíz, la papa, el trigo, la soja, el tomate, la lechuga y el girasol. Ha sido muy utilizada con fines de mejoramiento agronómico porque tiene la ventaja de que no conlleva las regulaciones a las que están sometidas las plantas genéticamente modificadas. Entre los genes estudiados están los relacionados a la calidad de almidón en granos (en mutantes con mayor amilopectina como los *waxy*), senescencia, enanismo, resistencia a oidios, larga vida en estante (en tomate y lechuga), reducción de isoflavonas y alérgenos en soja, modificación de la composición de aceites en girasol, etc.

El EcoTILLING en cambio, se orientó a la utilización de ecotipos como fuente de diversidad para estudiar genes relacionados a la resistencia a estrés abiótico (como por ejemplo los codificantes para dehidrininas) o para encontrar nuevos marcadores moleculares funcionales de tipo SNP (polimorfismos de un solo nucleótido, ver capítulo correspondiente en este mismo libro).

La técnica consiste esencialmente de dos pasos: 1) la aplicación de mutagénesis química inducida tradicional como se describió en este capítulo, usualmente empleando EMS (etilmetanosulfonato); y 2) la utilización de una novedosa estrategia de búsqueda y detección de las mutaciones inducidas sobre un determinado gen o secuencia nucleotídica de interés. En la Figura 2 ilustra el proceso. Usualmente se

mutagenizan semillas o granos de polen para inducir la aparición de mutaciones puntuales a lo largo del genoma. Las plantas derivadas de estas semillas tratadas con mutágenos se denominan M1 y, como ya se mencionó, suelen ser quiméricas y las mutaciones inducidas se encuentran en heterocigosis. Usualmente a las plantas M1 se las endocría dejándolas autofecundar en el caso de autógamias, o se procede a la autopolinización forzada en alógamas para poder tener control de las genealogías y asegurar su conservación. Usualmente se cosecha una semilla por planta M1 para pasar a la próxima generación, pero esto bien puede hacerse con una semilla por inflorescencia principal con muy bajo riesgo de repetir en la muestra dos mutantes provenientes del mismo evento mutacional. Por la endocría de las plantas M1 los genes mutados tenderán a hacerse homocigotas, sin embargo no es imprescindible que alcancen la homocigosis para que puedan ser detectados por TILLING. La generación M2 es segregante: cada individuo podrá ser homocigota mutante, heterocigota u homocigota salvaje en cada uno de los genes afectados, pero serán de constitución genética homogénea en todo su soma, es decir no quiméricos. De cada planta M2, que debe ser debidamente identificada, se extrae ADN de las hojas para el análisis de el/los gen/es de interés. Por otro lado, se guardan sus semillas para poder disponer en el futuro de sus descendencias (poblaciones M3). De este modo es posible conservar las mutaciones detectadas mediante el análisis de ADN, y adicionalmente disponer de poblaciones que puedan distribuirse internacionalmente entre los consorcios de investigación de cada especie (es común que se comparta el material de las poblaciones mutantes de las especies más estudiadas, como *Arabidopsis* y maíz, entre los consorcios internacionales). Para reducir costos, los ADNs procedentes de distintas plantas M2 (usualmente entre 2 y 10) se pueden mezclar para su análisis en un mismo tubo. Sobre una parte de la mezcla se realizará una amplificación por PCR, utilizando oligonucleótidos con secuencias flanqueantes al segmento genómico de interés (previamente diseñados en base a la secuencia conocida). Los productos de amplificación, que consis-

tirán en una mezcla de secuencias normales y mutadas, se desnaturalizan sometiéndolos a alta temperatura y se les permite rehibridar entre sí bajando la temperatura. Algunas de las cadenas rehibridadas podrán ser heteroduplexas, es decir, una de las cadenas provendrá de una secuencia mutada y otra de una normal produciendo una doble cadena que contiene bases mal apareadas por falta de complementariedad. Estas bases mal apareadas serán reconocidas por la enzima endonucleasa *CelI*, produciendo un corte donde encuentra la anomalía. Los productos de la digestión con *CelI* se resuelven por electroforesis en geles de poliacrilamida, y si los oligonucleótidos utilizados en la PCR fueron previamente marcados con fluoróforos (de fluorescencia rojo y verde en la figura), los productos del corte podrán ser identificados por su menor tamaño comparado con los ADNs no cortados (ver figura). La utilización de 2 fluoróforos, en lugar de uno, ayuda a reducir el número de falsos positivos.

Sobre esta técnica básica se introdujeron modificaciones como, por ejemplo, la utilización de geles no desnaturalizantes de poliacrilamida o agarosa y tinción con bromuro de etidio, que funcionó muy bien en trigo. En los casos en que se utiliza irradiación X o gamma para la inducción de mutaciones, el paso de digestión se saltea y se analiza en geles de secuenciación.

A medida que se van reduciendo los costos de secuenciación y se convierte en rutina la detección de SNPs, posiblemente reemplacen el tratamiento con *CelI* en la técnica de TILLING. Al día de hoy las publicaciones de *ultra-deep sequencing* en virus, permiten detectar genomas con variaciones en mezclas que contienen muchísimos genomas no mutados. Otra variante utilizada es la DHPLC (del inglés, *denaturing high-performance liquid chromatography*) basada en la utilización de columnas para cromatografía líquida de alta presión que permiten diferenciar ADN heterodúplex del homodúplex. Para reducir el número de muestras analizadas, varios productos de PCR y posterior digestión con *CelI* se siembran en el mismo carril de un gel (por ejemplo toda una columna). Por otro lado se siembra también toda una fila. De esta manera, la aparición de bandas de

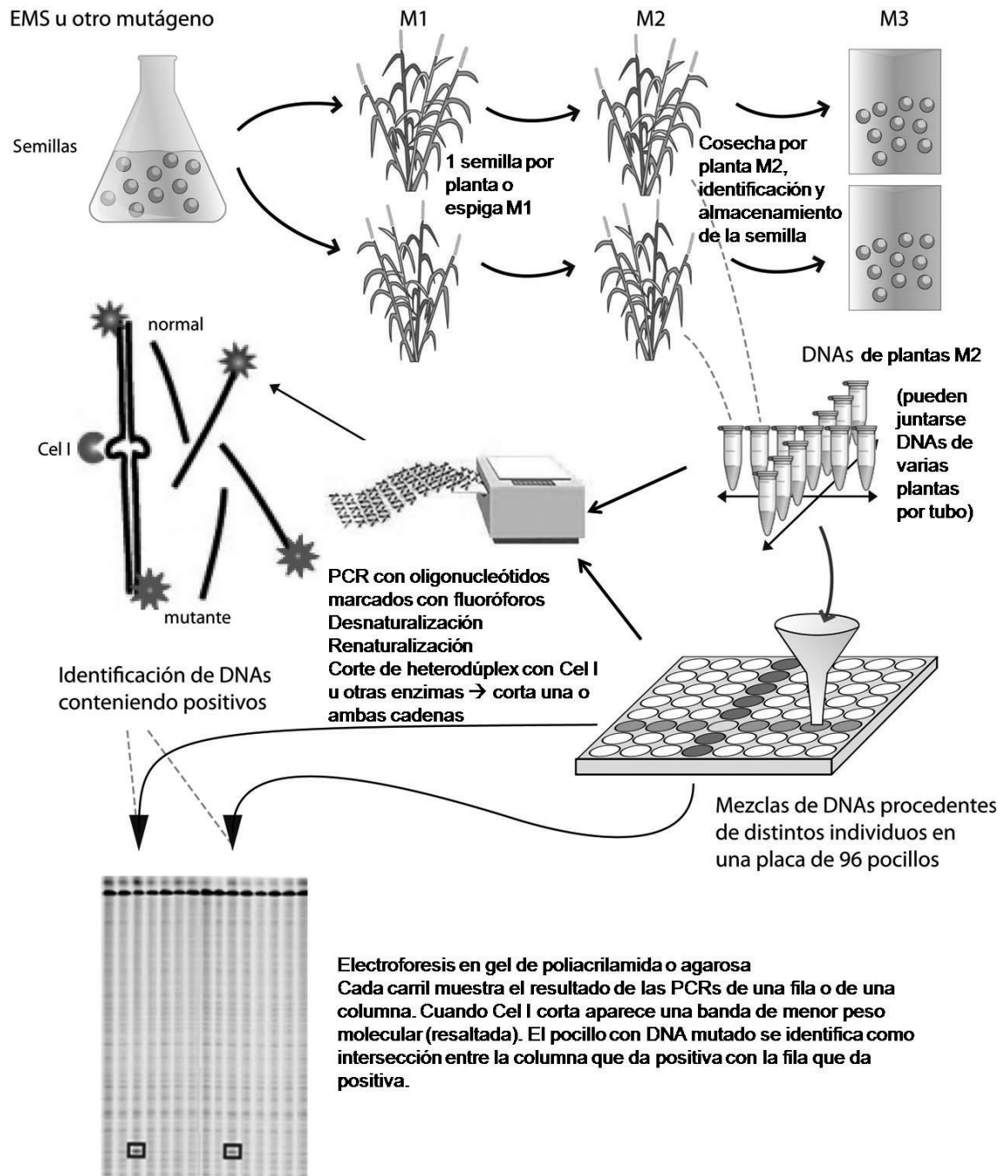
igual tamaño en carriles pertenecientes a columnas y filas permite identificar la intersección en la que se encuentra el pocillo con el ADN mutado (de un modo similar al esquema que se propone en la Figura 2).

En el EcoTILLING, menos utilizado que el TILLING, en lugar de analizar poblaciones provenientes de mutagénesis inducida, los materiales de partida que se utilizan como fuente de variabilidad suelen ser los almacenados en bancos de germoplasma, o bien ecotipos de especies no domesticadas. Los problemas de este enfoque alternativo son varios: el acceso al material puede ser limitante, los fondos genéticos son demasiado variables, las mutaciones que disminuyen la adaptación de una planta no suelen estar representadas porque son eliminadas por selección natural o artificial, etc.

### 3.3. Mutagénesis de inserción con transposones y ADN-T

Otro sistema utilizado en genética inversa para el análisis funcional de genes consiste en la activación de transposones endógenos (siendo el Ac-Ds de maíz el más emblemático), mediante cruzamientos entre genotipos que resulten en la activación de dichos transposones o mediante la introducción de transposones foráneos por ingeniería genética (por ejemplo, el sistema Ac de maíz fue introducido en papa, el *Enhancer* o *SPM* de maíz en *Arabidopsis* o el mutador Robertsomiano *Mu* en maíz).

Una de las grandes ventajas de esta técnica es que el elemento insertado proporciona un marcador molecular de secuencia conocida (*tag*) en el sitio de inserción (sitio de *knock out*), que puede facilitar el aislamiento del gen mutado por pérdida o ganancia de función debido a que la inserción es reversible. Las desventajas principales de esta aproximación son: que las mutaciones pueden ser inestables (revertir), y que la transposición tiene frecuencias bajas y no siempre la inserción es totalmente azarosa. Para potenciarla se pueden colocar promotores fuertes como el 35S del virus del mosaico del coliflor (CaMV) para controlar la expresión del gen de la transposasa. Las ventajas de hacer esto se reflejan en un aumento significativo en la tasa de mutación (hasta  $1 \times 10^{-4}$  en maíz, por ejemplo), y en que provee una secuencia



foránea conocida en el sitio de inserción, lo cual facilita el posterior clonado molecular.

Existen algunas variantes de esta técnica como la captura de genes o de promotor (*gene o promotor trapping*), en este caso se pueden incluir secuencias foráneas de genes reporteros como *uidA* (GUS), *gfp* (GFP) y sus variantes de otros colores, o genes involucrados en la síntesis de antocianinas (Lc), con o sin pro-

motores funcionales. Si el gen indicador que no contiene promotor se inserta al azar bajo el control de algún promotor endógeno, entonces se activará su expresión y se visualizará fácilmente. Otra ventaja de esta aproximación es la identificación sencilla sin necesidad de disponer de un fenotipo mutante, lo cual es útil cuando existen genes redundantes, heterocigosis o cuando se analizan genes activos

en distintos estadios de desarrollo. La función génica puede caracterizarse posteriormente en las mismas líneas si la inserción ha producido un *knock out*.

La mutagénesis de inserción con ADN-T (segmento T de transferencia del plásmido Ti) consiste en aprovechar la propiedad del sistema basado en *Agrobacterium tumefaciens*, para la transformación genética de plantas. El segmento T se inserta teóricamente al azar en cualquier sitio de la eucromatina de las plantas, permitiendo así generar colecciones de mutantes de inserción con las características ya mencionadas anteriormente para los transposones; pero con mayor estabilidad que éstos y menor preferencia por el sitio de inserción. Sin embargo, como las frecuencias de mutación son extremadamente bajas, sólo se utilizan actualmente en combinación con transposones. Es decir que se incluye un transposón activo dentro del segmento T para inducir mutagénesis por transposición en la especie transformada. Una de sus desventajas, es que su uso estará limitado a aquellas especies que son transformadas por agrobacterias.

### Lecturas Recomendadas

Birky C.W. Jr. 2001. The inheritance of genes in mitochondria and chloroplasts: Laws, mechanisms, and models. *Ann Rev Genet*, 35, 125-148.

Bouchez, D. and H. Höfte. 1998. Functional Genomics in Plants. *Plant Physiol*, 118, 725-732

Datta S.K. 2005. Role of classical mutation breeding in crop improvement. Daya Publishing House, India. 314 p.

Dong C, J Dalton-Morgan, K Vincent, P Sharp A Modified TILLING Method for Wheat Breeding *Plant Gen.* March 2009 2:39-47

Draper B.W., McCallum C.M., Stout J.L., Slade A.J., Moens C.B. 2004. A high-throughput method for identifying N-ethyl-N-nitrosourea (ENU)-induced point mutations in zebrafish. *Methods Cell Biol*, 77, 91-112.

Grant-Downton R.T. and Dickinson H.G. 2005. Epigenetics and its implications for plant biology. 1. The epigenetic network in plants. *Annals of Botany*, 96, 1143-1164.

IAEA. 1995. Manual on mutation breeding Second Edition. Technical Reports Series No 119. Joint FAO/IAEA Division of Atomic Energy in Food and Agriculture, Vienna 1995. 288 p.

Kirk J.T.O. and Tilney-Bassett R.A.E. 1978. The plastids. Elsevier, North-Holland, Amsterdam, 960 p.

Koornneef M. 2002. Classical mutagenesis in higher plants. In: PM Gilmartin and C. Bowler (eds) *Molecular Plant Biology: Practical Approach series Number 258*. Vol.1. pp 1-11.

May, BP, H Liu, E Vollbrecht, L Senior, PD Rabinowicz, D Roh, X Pan, L Stein, M Freeling, D Alexander, R Martienssen 2003 Maize-targeted mutagenesis: A knockout resource for maize *Proc Natl Acad Sci U S A*. 100: 11541-11546.

McCallum CM, Comai L, Greene EA, Henikoff S. 2000 Targeted screening for induced mutations. *Nat Biotechnol*. 18:455-7.

McCallum CM, Comai L, Greene EA, Henikoff S. 2000 Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. *Plant Physiol*. 123: 439-42.

Neuffer M.G. 1993. Mutagenesis. En: *The Maize Handbook* (Freeling M. and Walbot V. eds.). pp 212-219.

Prina A.R. 1989. Consideraciones sobre la aplicación eficiente de la mutagénesis inducida en fitomejoramiento. *Mendeliana*, 9 (1), 5-49.

Prina A.R. 1993. La mutagénesis inducida en el mejoramiento vegetal. *Bol. Genet. Inst. Fitotéc. Castelar*, 17, 9-22.

Prina A.R. and Favret E.A. 1988. Influence of marker genes on the expression of somatic mutations in barley. *J Hered*, 79, 371-376.

Ramachandran, S y S Venkatesan. 2001. Transposons as tools for functional genomics. *Plant Physiol. Biochem*. 39: 243-252

Slade AJ, Fuerstenberg SI, Loeffler D, Steine MN, Facciotti D. 2005 A reverse genetic, nontransgenic approach to wheat crop improvement by TILLING. *Nat Biotechnol*. 23:75-81.

Stadler L.J. 1930. Some genetic effects of X-rays in plants. *J. Hered.*, 21 (1), 3-19.

Uauy, C; F Paraiso, P Colasuonno, R K Tran, H Tsai, S Berardi, L Comai, J Dubcovsky 2009 A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat, *BMC Plant Biology* 9:115

van Harten A.M.. 2002. Mutation breeding of vegetatively propagated ornamentals. *Breeding for Ornamentals: Classical and Molecular Approaches* In: A. Vainstein (ed.) Kluwer Academic Publishers, The Netherlands 2002. pp 105-127.

## II. Capítulo 5

### Variación somaclonal

Susana Cardone, Sofía Olmos y  
Viviana Echenique

#### 1 Introducción

El cultivo *in vitro* representa un momento de estrés para las células y tejidos vegetales y puede desencadenar procesos mutagénicos durante el establecimiento del explanto, la inducción de callo y la formación de embriones o vástagos durante el proceso de regeneración de plantas. Algunos de los cambios genéticos ocurridos en las plantas regeneradas pueden resultar variantes atractivas, con utilidad potencial en el mejoramiento vegetal para el desarrollo de nuevas variedades.

Larkin y Scowcroft (1981) llamaron **variación somaclonal** a los cambios ocurridos en las plantas regeneradas y que son transmitidos a la progenie. Asimismo cabe citar la ocurrencia *in vitro* de cambios reversibles que pueden modificar la expresión de ciertos genes. Estos cambios que no implican alteración en la secuencia nucleotídica se denominan “epigenéticos” (Madlung y Comai, 2004). Algunos autores los consideran variantes somaclonales mientras que otros sólo incluyen en la misma aquellos cambios que no revierten en ciclos sucesivos de reproducción sexual.

Los mecanismos por los cuales ocurre la variación somaclonal no han sido completamente dilucidados. Sin embargo, se han propuesto varias causas de posible incidencia en la ocurrencia de la misma. Entre esas causas se citan: el genotipo, la fuente de explanto, el tiempo en cultivo, las condiciones y composición del medio de cultivo y la vía de regeneración. La comprensión de estas causas ayudaría a mejorar la interpretación de los procesos celulares de respuesta al estrés y permitiría definir como actúan en los procesos de evolución.

Entre los fenómenos que ocurren durante el cultivo *in vitro* se mencionan alteraciones en el cariotipo, mutaciones puntuales, recombinación somática e intercambio de cromátidas hermanas, rearrreglos génicos somáticos, activación de elementos genéticos transponibles,

amplificación y metilación del ADN y cambios en el ADN de las organelas. Por ello, aunque los caracteres morfológicos (como por ejemplo el hábito de crecimiento, la morfología floral, etc.) son fáciles de evaluar, otros cambios no se manifiestan en variantes morfológicas evidentes. Una diferencia estructural en un producto génico puede no alterar su actividad biológica lo suficiente como para producir un fenotipo modificado, por lo tanto, la variación debe analizarse a varios niveles.

Los cambios producidos son generalmente indeseables, pero la aparición ocasional de variantes no encontradas en las poblaciones naturales y que representan una ventaja desde el punto de vista agronómico, permite utilizar este fenómeno en programas de mejora vegetal. Se ha utilizado, en algunos casos, para conferir caracteres deseables a cultivares de importancia económica, entre los que se incluyen: resistencia a enfermedades, tolerancia a suelos ácidos y a salinidad. El desarrollo de nuevos cultivares por esta técnica involucra un balance entre la cantidad de variación inducida y el mantenimiento de los caracteres agronómicos del cultivar. Como ejemplo puede citarse el caso de somaclones de pasto bermuda, *Cynodon dactylon*, donde se encontró resistencia a la oruga militar en un cultivar que reunía otras buenas características, dando origen al cultivar Brazos-R3.

Si bien desde el punto de vista práctico no es una técnica muy eficiente, dado que no es posible predecir ni dirigir el tipo de variación y se hace menester trabajar con poblaciones grandes de plantas, es necesaria la comprensión del mecanismo para evitarlo en casos donde se requiere fidelidad genética, como en la micropropagación, la conservación de germoplasma y la transformación de plantas.

En algunos casos puede representar una fuente de variación rápida y de fácil acceso para ser utilizada en programas de mejoramiento, especialmente para especies con sistemas genéticos limitados o de base genética estrecha, como en el caso de la apomixis, donde la variabilidad dentro de las poblaciones naturales o cultivadas puede ser limitada.

Los primeros casos de variación somaclonal se registraron en plantas de reproducción agámica como la caña de azúcar y la papa,

pero el fenómeno ha sido observado en monocotiledóneas como trigo, maíz, avena, arroz, pasto miel (*Paspalum dilatatum*) y pasto llorón (*Eragrostis curvula*) y en dicotiledóneas como tabaco, tomate, zanahoria y col, entre muchas otras especies.

## 2 Factores relacionados con la aparición de variación somaclonal

Hemos mencionado, previamente, algunos de los posibles factores que tienen influencia en la aparición de variación somaclonal. En este punto trataremos cada uno por separado.

**Genotipo.** En general se asume que la frecuencia de cambios dependerá de variaciones preexistentes en el genotipo y de las interacciones que surgen entre el mismo y el proceso de cultivo. En arroz, ciertas variedades estables muestran niveles de variación que oscilan entre el 0 y el 1%, mientras que en otras, consideradas inestables, oscilan entre el 10 y el 27%. En *Kalanchoe*, especie ornamental, la variación somaclonal es genotipo dependiente y se la utiliza como una herramienta para la obtención de variedades. El nivel de ploidía del genotipo es otro factor a tener en cuenta. Por ejemplo en raigrás y en papa se observó que, durante el cultivo de tejidos, las variedades diploides muestran estabilidad, mientras que las tetraploides tienden a generar aneuploidías. La explicación más razonable para este hecho es que los poliploides, con más de dos juegos completos de cromosomas, están “tamponados”, y por lo tanto toleran la ganancia o pérdida de cromosomas, pudiendo así cumplir con los requisitos impuestos por la regeneración. En los diploides, la pérdida o la alteración de cromosomas que llevan genes vitales impedirá la regeneración de las plantas, llegando con éxito a esta etapa sólo aquellos explantos que tengan su complemento cromosómico completo. Existen evidencias de una mayor inestabilidad en híbridos interespecíficos cultivados *in vitro*, si se los compara con las especies parentales. Como ejemplo puede citarse el caso de los híbridos obtenidos de la cruce de *Hordeum vulgare* x *Hordeum jubatum*, donde el cultivo *in vitro* genera entre un 5% y un 10% de regenerantes haploides que suelen contener sólo el genoma de *Hordeum vulgare*.

**Explanto.** La variación observada entre plantas regeneradas puede ser una consecuencia de quimerismo en el explanto original. Las quimeras son mosaicos genéticos. Esto significa que dentro de una misma planta existen diferentes linajes celulares, esto se debe a una serie de cambios en el ADN nuclear de ciertos tipos celulares producidos durante el desarrollo. Si estos tejidos se utilizan como explantos y sus células son inducidas a dividirse y rediferenciarse, las diferentes líneas celulares pueden dar origen a plantas genéticamente diferentes. Por lo tanto, la utilización de explantos con tejidos quiméricos preexistentes puede resultar en una fuente “extra” de variación. Sin embargo, la recuperación de quimeras a partir de explantos no quiméricos es un fenómeno frecuente que puede ocurrir, a partir de procesos organogénicos, o por causas desconocidas. La utilización de explantos con tejidos organizados, como esquejes radicales o caulinares, es una buena opción si es necesario mantener estabilidad genética durante el cultivo *in vitro*. Sin embargo, diferentes genotipos reaccionan de manera distinta, aún con explantos “seguros”. El cultivo de meristemas aislados minimiza el riesgo de variación somaclonal. Esto no debe tomarse como regla, ya que utilizando técnicas moleculares fue posible detectar variación en plantas de frutilla y paraíso obtenidas a partir de meristemas. El cultivo de protoplastos parecería inducir inestabilidad genética, como fue observado en papa y tabaco; esto se ha asociado a los mayores **tiempos en cultivo** involucrados en la regeneración de plantas a partir de explantos tan pequeños.

La **vía de regeneración** también tiene un rol importante en la ocurrencia de alteraciones en plantas obtenidas *in vitro*. En especies de los géneros *Pennisetum*, *Panicum*, y *Lolium* la variación observada en cultivos embriogénicos fue relativamente menor que la que obtenida en cultivos organogénicos. Esto probablemente se debe a la gran presión de selección impuesta en la formación de los embriones, mayor que la requerida en la formación de vástagos. Se cree que el gran número de genes requeridos para la iniciación y maduración de embriones cigóticos y somáticos impediría la acumulación de mutaciones deletéreas. Sin



embargo, también se observaron variantes en plantas de café, apio y caña de azúcar regeneradas a través de embriogénesis somática. En estos casos se advirtió que algunos fenotipos anormales de embriones somáticos se asemejan a los mutantes del desarrollo embrionario obtenidos en *Arabidopsis* y maíz.

La mayor parte de la variación obtenida *in vitro* parece provenir de la **fase de callo**. Los mecanismos de iniciación de un callo parecen ser similares a la respuesta de las plantas a heridas, que se sabe que activan elementos transponibles y estimulan la inducción de enzimas y productos específicos de situaciones de estrés. La desdiferenciación que ocurre durante la inducción del callo, altera procesos celulares que resultan en cambios genómicos, uno de los más frecuentes es la inestabilidad cromosómica. La **naturaleza del callo** también puede afectar el nivel de variación obtenida. Un callo verdadero es una masa de células desdiferenciadas que proliferan desorganizadamente, lo cual probablemente genera considerable variación. En varias monocotiledóneas el callo representa, a menudo, una masa de órganos suprimidos o proembriones más que un tejido completamente desorganizado. Este tipo de callo mantiene un alto grado de estabilidad genética, como se ha observado en espárrago. En un estudio realizado en *Cymbopogon* se observó que muchas de las plantas regeneradas a partir de callos provenientes de semillas fueron anormales, mientras que las obtenidas por callos de inflorescencias fueron muy semejantes a las plantas de las cuales provenían los explantos. En ocasiones, también fue posible observar variación en plantas obtenidas por regeneración directa, es decir, sin que medie un proceso previo de desdiferenciación celular ni formación de callo.

El estado físico del **medio de cultivo** también influye en el nivel de variación obtenida. Un mismo explanto puede tener diferente comportamiento si se lo cultiva en medio sólido o en medio líquido. Otro factor importante es la **temperatura**, que puede inducir inestabilidad cariotípica y/o incrementar el número de plantas albinas y también la **deficiencia de oxígeno** que se genera durante el transcurso del cultivo. La tensión de oxígeno a la que están

expuestas las células superficiales del callo es diferente a la de las células situadas en profundidad. La anaerobiosis resultaría en la producción de etanol, el cual podría comportarse como un mutágeno. Un efecto similar podría tener la **acumulación de metabolitos** producidos por las propias células durante el cultivo.

Los **reguladores de crecimiento**, principalmente el 2,4-D (ácido 2,4 diclorofenoxiacético), han sido señalados como inductores de inestabilidad. Estos ejercen profundos efectos sobre la respiración celular, el consumo de azúcares y en el control de la división celular. Por ello se especula acerca de su rol indirecto en la inducción de cambios en el metabolismo celular y tisular de plantas creciendo *in vitro*. Aparentemente el 2,4-D induce desdiferenciación y provoca un incremento sustancial en la transcripción. Esto puede alterar la estructura de la cromatina, por lo cual se lo utiliza en gramíneas como inductor de aneuploidías. La relación auxina:citosina también suele afectar en este sentido. El 2,4-D, el AIA (ácido indolacético) y el ANA (ácido naftalenacético) han sido señalados como los responsables de los incrementos en la metilación de las citosinas, que tienen lugar durante el cultivo *in vitro*. Los sectores del ADN donde las citosinas se encuentran metiladas en posición 5 son considerados puntos calientes de mutación, ya que la desaminación de una 5-metilcitosina resulta en un cambio de citosina a timina. Un estudio realizado en papa indicaría que el tipo y la concentración de los reguladores de crecimiento en los estadios iniciales de cultivo *in vitro* no generaría variación genética. Los efectos serían más importantes durante el crecimiento del callo y la iniciación de los vástagos. Estos datos sugieren que la fase de callo sería un período sensible en el cual la manipulación hormonal afecta la estabilidad de las plantas regeneradas; de manera que las hormonas inducirían inestabilidad genética sin ser, necesariamente, el origen de la misma.

La **deficiencia o exceso de minerales** en el medio de cultivo podrían ser causa de variación. Este fenómeno se ha observado también en plantas creciendo en condiciones de campo. Un ejemplo típico lo constituyen plantas de lino sometidas a exceso o deficiencias de

azufre, nitrógeno, fósforo, calcio y magnesio. En estas plantas se observaron cambios genómicos estables y heredables. En este caso los cambios fueron inducidos por una fertilización excesiva del suelo, obteniéndose así líneas estables, denominadas "genotrofos", caracterizadas por diferente contenido de ADN nuclear y citoplasmático.

La **edad del cultivo** es otro factor que afecta el nivel de variación, aumentando la proporción de variantes en cultivos envejecidos y también en plantas obtenidas a través de varios subcultivos. Esto se ha observado en ajo, maíz, avena, tabaco, triticale y raigrás triploide. La variación puede minimizarse realizando subcultivos frecuentes de explantos jóvenes. El número óptimo de subcultivos podría estimarse empíricamente luego de realizar pruebas de fidelidad genética en las plantas regeneradas, a través de subcultivos subsecuentes. Una observación común es que en los períodos prolongados de cultivo hay una pérdida de totipotencia y que esto sucedería debido a la acumulación de mutaciones y a la alteración de los genes que son responsables de la regeneración. Sin embargo, un estudio realizado por nuestro grupo de trabajo en la UNS, en colaboración con el IBO-NE, demostró que en plantas micropropagadas de paraíso gigante las variaciones se producen al azar en los diferentes subcultivos, no habiendo una relación lineal entre el número de subcultivos y la cantidad de variación obtenida. La variación en este caso se detectó mediante la técnica de RAPDs a lo largo de de 10 subcultivos (Olmos y col., 2002).

### **3 Cambios genómicos producidos durante el cultivo de tejidos**

Las plantas poseen una amplia capacidad de acomodarse a las situaciones cambiantes de su entorno, pero en algún momento esa capacidad de amortiguación se vuelve deficiente y los organismos caen en un estado de crecimiento subóptimo que parece inducir mecanismos de cambios genómicos. Las bases metabólicas de la interacción entre el estrés y estos cambios son desconocidas. Sin embargo, en la literatura se menciona la posibilidad de ocurrencia de alteraciones a nivel del ADN a consecuencia de diferentes estreses ambientales, depen-

diendo de la intensidad del estrés y del estado de diferenciación de las células, en el momento de experimentarlo. Estos cambios ocurrirían, en parte, debido a una pérdida del control del ciclo celular. Entre las alteraciones genéticas se citan **mutaciones puntuales, cambios en la estructura o el número de cromosomas y activación de elementos genéticos transponibles**. Estos últimos causan reorganizaciones moleculares, no solo por su transposición, sino también porque generan amplificaciones y deleciones. El cultivo *in vitro* también incrementa la frecuencia de intercambio entre cromátidas hermanas o "crossing over" somático.

Varias hipótesis han intentado explicar y relacionar la serie de eventos que tienen lugar durante el cultivo *in vitro*, como son las perturbaciones del ciclo celular, las aberraciones cromosómicas estructurales y numéricas, la activación de transposones y las mutaciones génicas. Kaeppler y Phillips (1993) proponen la existencia una base molecular común para todos estos eventos mutagénicos que comprendería una alteración en los patrones de metilación del ADN. Estos cambios en la metilación podrían afectar la expresión de genes específicos, incluyendo elementos transponibles o podrían afectar la estructura de la cromatina de una manera más global, involucrando, por lo tanto, a un gran número de genes. El mecanismo por el cual suceden estos cambios no ha sido totalmente dilucidado pero se los ha vinculado con una replicación tardía de la heterocromatina que también provocaría eventos de ruptura cromosómica. Los cambios en los modelos de metilación causados por el cultivo de tejidos no serían aleatorios y en raigrás se ha propuesto la existencia de puntos calientes de inestabilidad en el ADN genómico.

Kaeppler y Phillips (1993) indican que plantas bajo estrés severo de nutrientes o de agua no experimentan el mismo tipo de cambios encontrados en los regenerantes de cultivo *in vitro*. El cultivo de tejidos podría representar, por lo tanto, un tipo muy particular de estrés, que podría inducir una respuesta única y un perfil particular de mutación. Un posible efector de esta variación podrían ser los reguladores de crecimiento.

En cuanto a los cambios en el número de cromosomas, la **poliploidía** es el más común de estos fenómenos, y estaría relacionada con

fallas en la mitosis, mientras que la aneuploidía puede surgir por segregación desigual en la mitosis o por fragmentación nuclear, seguida de mitosis. A veces estos cambios en el número cromosómico pueden afectar el fenotipo. En papa, se observaron fenotipos aberrantes por aneuploidía o por duplicación cromosómica; sin embargo, no todos los regenerantes aneuploides o poliploides presentaban cambios fenotípicos evidentes.

La variación somaclonal también puede afectar el **genoma de los plástidos y las mitocondrias**. En plantas de sorgo androestériles obtenidas *in vitro* se observó restauración parcial de la fertilidad y en plantas androestériles de maíz con citoplasma "T" (que eran, además, susceptibles a *Helminthosporium maydis*) hubo reversión de la androesterilidad y de la susceptibilidad al patógeno. Estos cambios se correspondían con mutaciones en el ADN mitocondrial. Otro caso de variación debida al cultivo *in vitro* es el albinismo, problema que se presenta frecuentemente en el cultivo de anteras de Gramíneas. En plantas albinas de trigo obtenidas por cultivo de anteras se observaron deleciones y rearreglos en el ADN de los cloroplastos. Si bien parecería que los tejidos somáticos son menos sensibles a la variación, también se ha observado albinismo en plantas regeneradas a partir de los mismos.

#### **4 Niveles de detección de la variación somaclonal**

Los mecanismos que operan para inducir variación son numerosos y diversos y, probablemente, actúan en simultáneo, conduciendo a cambios en caracteres cuali y cuantitativos. Detectar y analizar los distintos niveles de modificaciones genómicas generadas por cultivo *in vitro* reviste interés desde un punto de vista práctico, ya que las mismas podrían ser potentes fuentes de material para seleccionar variantes de interés y desde un punto de vista teórico, ya que posibilitarían realizar estudios básicos acerca de las causas de la variación y el posible control de la misma.

La ocurrencia de variación somaclonal puede detectarse con marcadores morfológicos, bioquímicos y moleculares. La correlación de dichos marcadores con caracteres agronómi-

cos es un requisito relevante para su implementación en los programas de mejoramiento.

La utilización de **marcadores morfológicos** para detectar variantes somaclonales ha sido exitosa desde el punto de vista del mejoramiento y citada en varios trabajos de investigación. El examen visual y la utilización de un analizador de imágenes posibilitaron la diferenciación entre fenotipos normales y aberrantes en *Pelargonium*. También se detectaron cambios en altura, tamaño de hojas, número y peso de semillas de plantas de maní obtenidas *in vitro*. En ananá se observaron cambios estables en el color y textura de las hojas, tamaño de los frutos y tasa de proliferación de vástagos; el cultivo de tejidos de violeta africana resultó en un 67% de variación en el color de las flores de plantas regeneradas.

A **nivel fisiológico**, tanto en cereales como en frutales se detectaron y seleccionaron variantes que presentaron resistencia a enfermedades, tolerancia a herbicidas, sales y condiciones de acidez en el suelo, entre otros. Estos estudios se basaron en la utilización de altas presiones de selección durante la regeneración, mediante la adición de agentes selectivos en el medio de establecimiento y regeneración, como inóculos, toxinas, herbicidas, condiciones de pH extremas o concentraciones salinas elevadas.

El estudio del **cariotipo** permite revelar cambios en el número de cromosomas (poliploidías, aneuploidías) y en la estructura de los mismos (translocaciones, inversiones, deleciones y duplicaciones). Las alteraciones cariotípicas constituyen una fuente importante de variación, que muchas veces es subestimada cuando se trabaja con métodos tradicionales de conteo cromosómico. Las técnicas de bandeo cromosómico brindan más información acerca de la estructura cromosómica y las alteraciones que pueden surgir en este sentido. Por ejemplo, en plantas de *Brachycome dichromosomatica* regeneradas a partir de suspensiones celulares, esta técnica permitió detectar reordenamientos cromosómicos, mientras el número cromosómico permaneció estable. La hibridación *in situ* es otra de las técnicas que permiten examinar los cariotipos de una forma más exacta y resulta particularmente útil para revelar anomalías en la estructura de los cromosomas.

Entre los **marcadores bioquímicos** utilizados para analizar plantas regeneradas y sus progenies, se encuentran las isoenzimas. Las mismas permitieron detectar variación en plantas de *Populus tremuloides* regeneradas a través de callos. El análisis se realizó mediante electroforesis en geles de almidón, utilizando los sistemas isocitrato deshidrogenada (IDH) shiquimato deshidrogenasa (SKDH), malato deshidrogenasa (MDH), fosfogluco-isomerasa (PGI) y 6-fosfogluconato - deshidrogenasa (6-PGD). Los patrones electroforéticos permitieron diferenciar las plantas control de las variantes, sin embargo, las plantas albinas obtenidas en este ensayo no presentaron polimorfismos en los loci estudiados. Tampoco se encontraron variantes cuando se estudiaron líneas de callos derivadas de embriones cigóticos y somáticos de abeto. Los 25 loci analizados, utilizando 15 sistemas isoenzimáticos, mostraron fenotipos isoenzimáticos normales. Otro ejemplo, utilizando esta metodología, se llevó a cabo en líneas isogénicas de *Trifolium pratense* L. que diferían en su capacidad de regeneración. En este caso los patrones isoenzimáticos de esas líneas resultaron idénticos para los sistemas de peroxidasas (PRX), glutamato deshidrogenadas (GDH), alcohol deshidrogenasas (ADH) y estererasas (EST).

Si bien las isoenzimas aportan información útil acerca de la variación generada *in vitro*, su análisis refleja los productos de genes expresados, cuya regulación puede estar en cierta medida afectada por factores ambientales o fisiológicos. Por otro lado, sólo se transcribe y traduce una pequeña proporción del genoma. Los **marcadores moleculares** presentan varias ventajas sobre las isoenzimas ya que permiten detectar mutaciones en diferentes tipos de secuencias y además, porque no son influenciados por el ambiente, ni por los estadios fenológicos.

Los **RAPDs** (fragmentos polimórficos de ADN amplificados al azar) permitieron detectar la existencia de polimorfismos en plantas de *Triticum tauschii* obtenidas de callo. Estos marcadores también se utilizaron para evaluar la fidelidad genética en plantas regeneradas de pino, abeto, festuca, roble y ginseng. En caña de azúcar, los marcadores RAPDs permitieron

comprobar la susceptibilidad diferencial del genotipo al pasaje por cultivo *in vitro*, además de reunir información para encarar nuevos cruzamientos en programas brasileros de mejoramiento de esta especie.

Existen numerosos ejemplos que indican que la variación molecular no siempre se expresa a nivel fenotípico y viceversa. Por ejemplo en el género *Musa*, la técnica de RAPDs reveló polimorfismos en plantas micropropagadas de fenotipo normal. Del mismo modo, plantas regeneradas de palmera datilera que presentaban fenotipos florales anormales, tuvieron patrones moleculares normales.

Otros marcadores útiles son los **AFLP** (polimorfismo en la longitud de fragmentos amplificados de ADN), que permitieron detectar la existencia de diferencias moleculares entre plantas de banana con fenotipos normales y aberrantes. Los **microsatélites o SSRs**, iniciales de su nombre en inglés "simple sequence repeats", son altamente informativos, ya que son abundantes y están dispersos a través del genoma. Estos marcadores se utilizaron para el análisis de plántulas de álamo, obtenidas por micropropagación, donde se observaron microsatélites polimórficos, en plantas fenotípicamente normales.

A veces, para detectar una variante obtenida *in vitro* es necesario realizar pruebas específicas. Por ejemplo, cuando se evaluaron somaclones de *Cynodon dactylon* en laboratorio, invernáculo y a campo, para resistencia a la oruga militar. Un somaclón, el Brazos-R3, tenía un rendimiento equivalente al del cultivar madre, pero además era resistente a dicho insecto. Esto se debe a que produce niveles más bajos de un estimulante para la alimentación del insecto. Es decir que mantenía los caracteres agronómicos positivos del cultivar parental, pero tenía un cambio en la producción de un metabolito secundario que confería resistencia a la oruga.

## **5 La variación somaclonal y su aplicación en el mejoramiento genético**

La base del mejoramiento genético es la optimización de las interacciones génicas. Para lograr combinaciones génicas óptimas o superiores se requiere de diversidad genética. Esta

diversidad, que se encuentra normalmente en las poblaciones de individuos, puede provenir de cruzamientos sexuales, mutaciones inducidas (físicas o químicas o producidas por ADN-T, por elementos transponibles o retrotransposones), hibridación somática y transformación genética. La variación somaclonal proporcionaría una fuente adicional de variabilidad genética, potencialmente útil en programas convencionales de mejoramiento.

Algunas de las **ventajas** que presenta la variación somaclonal pueden resumirse de la siguiente manera:

- Es una técnica poco costosa: requiere de un laboratorio de rutina y facilidades de campo.
- Constituye una forma rápida de generar variabilidad genética, particularmente para cultivos con base genética estrecha y que son difíciles de mejorar a través de técnicas tradicionales.
- Es exitosa para eliminar uno o pocos defectos en cultivares bien adaptados.
- Puede utilizarse para mejorar especies de propagación sexual y vegetativa.
- Las tasas de mutación son relativamente elevadas si se comparan con las tasas de mutaciones espontáneas. La variación somaclonal ocurre con una frecuencia de 1 mutación cada 100 plantas regeneradas, en contraste con la tasa esperada de mutación espontánea de  $10^{-7}$  –  $10^{-9}$  mutaciones por par de nucleótidos por generación.
- El conocimiento de las condiciones que generan inestabilidad genética durante el cultivo *in vitro*, permitiría utilizar el fenómeno como estrategia de mejoramiento y permitiría eludirlo en aquellos casos en que se requiera estabilidad genética.
- El nivel de cambios no deseados es menor que cuando se utilizan mutágenos químicos o físicos, ya que, en el caso de la variación somaclonal, la mayoría de los cambios deletéreos producidos son eliminados por el filtro que representa la regeneración de plantas.

Entre las **desventajas** cabe mencionar:

En algunos casos, las variantes somaclonales no han avanzado de la etapa de laboratorio o invernáculo, probablemente debido a que el material seleccionado tiene poca importancia práctica.

- La baja tasa de regeneración de plantas en cultivos de largo término. Generalmente en estos casos existe pérdida de la capacidad morfogénica.
- La regeneración está limitada a genotipos específicos que pueden no ser de mucho interés para los mejoradores.
- Algunos somaclones son inestables (variación epigenética).
- Algunos presentan alteraciones no deseables como aneuploidía, esterilidad, etc.

Desde el punto de vista **productivo y comercial**, una variante somaclonal debería cumplir los siguientes requisitos:

- Involucrar caracteres agronómicamente útiles.
- El nivel de expresión del carácter debe superar al de sus progenitores.
- El carácter mejorado debe estar combinado con todos los otros caracteres agronómicos de la variedad que son importantes para el cultivo.
- La variación debe ser heredable y estable (sin reversión) en la progenie. En general, los mejoradores buscan en los somaclones caracteres de importancia práctica. El recurso es utilizar cultivares altamente adaptados para modificar unos pocos caracteres, ya que resulta más sencillo mejorar selectivamente una variedad de uso corriente que crear una nueva.

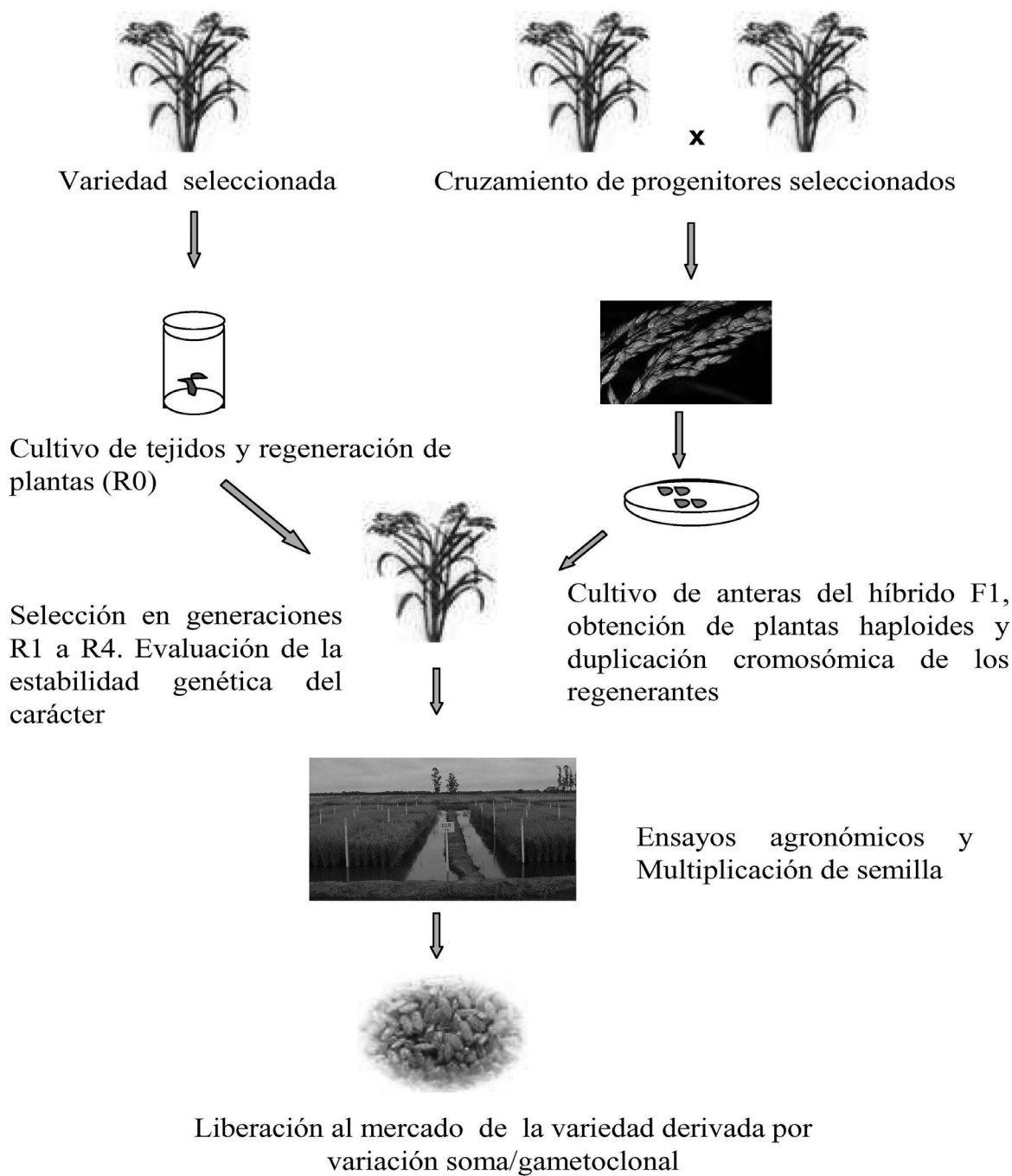
### 5.1 Estrategias para la selección de variantes útiles

Las estrategias que se han utilizado para obtener variación somaclonal comprenden:

a) **Obtención de plantas por cultivo de células o tejidos, que luego son analizadas para el carácter deseado.**

La **Figura 1** resume los pasos a seguir para la obtención de un cultivar por variación somaclonal. Para este fin se cultiva la mejor variedad

disponible y se seleccionan, entre las plantas regeneradas, aquellas que expresen el carácter de interés. Estas plantas (generación R0) se emplean para generar progenies sexuales



**Figura 1.** Estrategia para la producción comercial de variantes somaclonales (izquierda) y gametoclonales (derecha), en arroz.

(en el caso de plantas con este tipo de reproducción), sobre las cuales se vuelve a realizar la selección para el carácter agronómico buscado, tratando de mantener a la vez todas las características favorables de la variedad. En plantas autógamias, como trigo, arroz y cebada, se considera aceptable la evaluación de la estabilidad del carácter hasta la generación R4. A partir de este momento pueden realizarse cruzamientos de las plantas R4 selectas, con las líneas parentales a fin de determinar mediante el análisis de segregación, la base genética del carácter mejorado. Posteriormente, cuando se tenga semilla disponible, se realizarán pruebas agronómicas en diferentes ambientes, al menos en dos años distintos, para evaluar los efectos del componente ambiental en la expresión del carácter. El programa finaliza con la etapa de multiplicación de semillas para el lanzamiento de la variedad nueva al mercado.

Una modificación para la producción de nuevas variedades está basada en el empleo de variantes gametoclonales (**Fig. 1**). Si se utilizan células haploides de plantas F1, como fuente de explanto, las variantes obtenidas se denominan variantes gametoclonales. Los gametoclonos se refieren a genotipos haploides, derivados del cultivo de anteras. Por medio de esta metodología se podrían recuperar caracteres recombinados de ambos parentales, además de los que pudieran surgir *in vitro*. Las plantas haploides obtenidas son duplicadas mediante la exposición al alcaloide colchicina. Las evaluaciones a campo se realizan en forma conjunta a medida que los nuevos materiales van siendo multiplicados, de manera de aumentar las replicas bajo diferentes condiciones ambientales.

**b) Selección a nivel celular o selección *in vitro*.** Se realiza con el objetivo de obtener resistencia a estreses bióticos o abióticos y se utiliza generalmente para caracteres tales como resistencia a: toxinas fúngicas, herbicidas, concentraciones salinas elevadas y temperaturas extremas.

Mediante esta práctica pueden cultivarse explantos de distintos genotipos y observarse el comportamiento de los callos frente al agente selectivo, en condiciones de laboratorio. Una

vez seleccionados los genotipos resistentes se procede a la regeneración de las plantas, que son analizadas para ver si expresan la resistencia a campo y luego son introducidas en el programa de mejoramiento. La selección efectuada sobre cultivos celulares, *in vitro*, puede llevarse a cabo mediante dos modalidades:

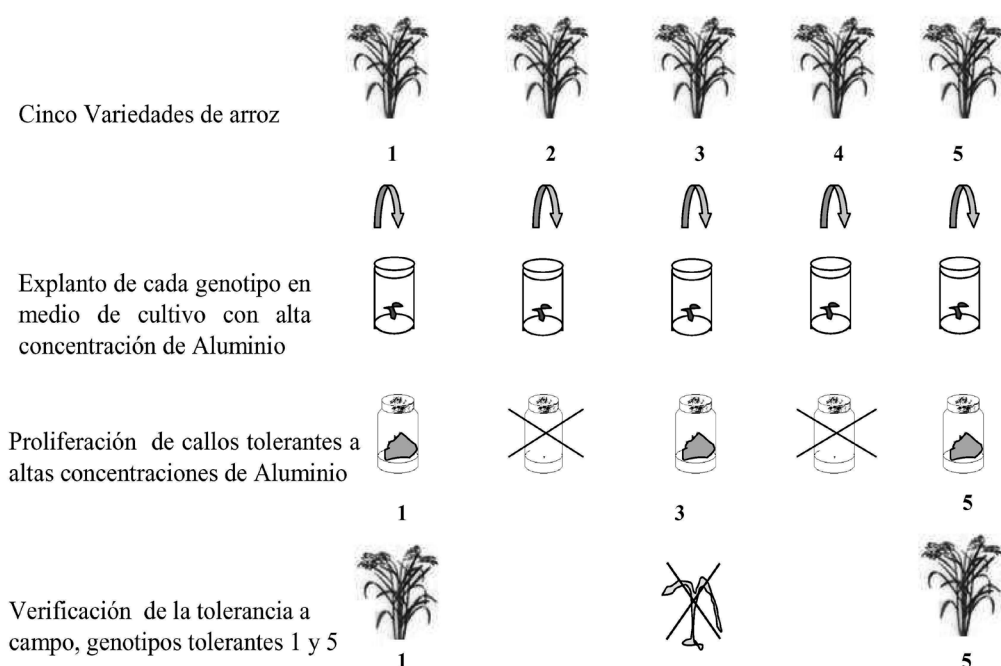
- **Selección directa** en un solo paso, donde el agente de selección es utilizado en concentraciones dobles o triples de la MIC (concentración mínima que produce un 100 % de inhibición).

- **Selección en varios pasos**, donde la concentración del agente selectivo es gradualmente incrementada en cultivos sucesivos y frecuentes.

El primero es un método simple y efectivo, ya que las células sensibles al agente morirán, permitiendo el crecimiento de las tolerantes. Sin embargo, elevados niveles de estrés serán deletéreos a nivel celular, eliminando materiales que tal vez, con un mayor nivel de diferenciación tisular hubiesen sido capaces de regenerar plantas tolerantes. La elección de un método u otro dependerá de un monitoreo preliminar que brinde información acerca de la reacción del tejido vegetal a las concentraciones letales y subletales del agente selectivo.

En la **Figura 2** se representan las etapas de la selección *in vitro*. En el ejemplo, diferentes variedades de arroz son evaluadas para tolerancia al aluminio. La inducción y la proliferación de callo en el medio de cultivo al que se ha adicionado aluminio, es indicadora de la tolerancia del genotipo al metal en cuestión. Una vez regeneradas las plantas, se requieren evaluaciones posteriores realizadas a campo, en suelos con concentraciones elevadas de aluminio, a fin de corroborar la tolerancia demostrada *in vitro*.

La búsqueda de variantes a nivel celular permite verificar y seleccionar fenotipos adecuados entre millones de células, con una inversión menor en tiempo, espacio y dinero y ha sido extensamente utilizada en diferentes especies entre las que pueden citarse: arroz, clavel, cártamo, banana, vid, frutilla y trigo. Sin embargo, no existe total garantía de que la resistencia manifestada *in vitro* se expresará a



**Figura 2.** Selección in vitro de plantas de arroz tolerantes a altas concentraciones de aluminio.

nivel de la planta entera. Hay diferentes explicaciones para este fenómeno, como por ejemplo que la resistencia o tolerancia observada se deba a cambios epigenéticos, a la naturaleza quimérica del material, al uso de toxinas no específicas (en el caso de resistencia a enfermedades), a la complejidad del carácter a ser seleccionado (como por ejemplo la resistencia a salinidad o sequía) o a la imposibilidad de las células para expresar el carácter cuando se diferencian. Para las dos estrategias que se citaron previamente, **a** y **b**, la obtención de resistencia sólo será posible si ésta existe en la población original (entonces el cultivo *in vitro* sólo serviría para recuperar los genotipos útiles) o bien si surge por variación somaclonal durante el proceso de cultivo.

## 6 Variantes somaclonales liberadas comercialmente

Dentro de los primeros registros de variación somaclonal en cultivos de importancia económica se encuentra la variación registrada en arroz, donde se informó en detalle la variación

obtenida en la progenie derivada de la autofecundación de plantas regeneradas. Se lograron avances en el mejoramiento de esta especie en caracteres cuali y cuantitativos. Ejemplos en éste y otros cultivos importantes, con variedades comerciales obtenidas por esta técnica, se detallan en la **Tabla 1**.

Además, se pueden mencionar logros en otros cultivos, por ejemplo:

En **papa**, se logró resistencia a *Phytophthora infestans*, a *Alternaria solani* y a la colonización por áfidos que transmitían el virus Y (PVY), y el virus del enrollamiento de la hoja (PLRV). En este caso la resistencia surgió a través del cultivo de protoplastos provenientes de suspensiones celulares cromosómicamente variables del cultivar "Russet Burbank".

En **banana** se obtuvo resistencia a *Fusarium* en cultivares que no habían podido ser mejorados por métodos tradicionales. No obstante, para lograr una variedad con rendimiento equiparable a las variedades no resistentes, se necesitó un trabajo posterior que permitió combi-



Especie	Carácter	Origen
<b>Geranio</b>	Arquitectura floral	Purdue Univ., USA
<b>Batata</b>	Calidad de raíces y arquitectura del canopeo	North Carolina Research Service, USA
<b>Caña de azúcar</b>	Resistencia a estrés abiótico y a enfermedades	Sugarcane Research Centre, Fiji
<b>Maíz</b>	Contenido de triptófano	Molecular Genetic, USA
<b>Tomate</b>	Contenido de materia seca	DNA Plant technology of New Jersey, USA
<b>Apio</b>	Contenido de materia seca	DNA Plant technology of New Jersey, USA
<b>Arroz</b>	Resistencia a enfermedades	Plantech Research Institute, Japan Univ. Agricultural Sciences, Hungary
<b>Mostaza</b>	Rendimiento	ICAR, India

**Tabla 1.** Especies con variedades obtenidas por la estrategia de variación somaclonal.

nar ambas características en un somaclón.

En **tabaco**, se obtuvieron cultivares tolerantes a aluminio y resistentes a herbicidas.

También, mediante la aplicación de esta estrategia se obtuvieron cambios interesantes para su explotación comercial en especies ornamentales. Variaciones en el color de las flores, la arquitectura de la planta y el número de flores por planta se citaron en **gerbera, clavel, crisantemo, tulipán, violeta africana y begonia**.

Dentro de las **especies forrajeras**, se han observado variaciones en plantas de **alfalfa** regeneradas por cultivo *in vitro*, para caracteres como altura y resistencia a *Fusarium oxysporum*. En **gramíneas forrajeras** como *Lolium* y *Festuca arundinacea* se han citado cambios, en varios caracteres como por ejemplo: morfología de planta, forma y tamaño de hoja y espiga, desarrollo floral, vigor y supervivencia, y en pasto bermuda (*Cynodon dactylon*), resistencia al moteado de la hoja

En el laboratorio de Genética del Dpto. de

Agronomía de la UNS se llevó a cabo un proyecto para obtener somaclones de pasto llorón. Se trabajó con varios cultivares de esta gramínea apomíctica y luego de la evaluación de un gran número de plantas se seleccionaron los siguientes materiales:

- Una planta diploide sexual, obtenida a partir de un cultivar tetraploide apomíctico. Este material fue registrado en el Instituto Nacional de Semillas.

- Utilizando un tratamiento de duplicación cromosómica, se trató semilla, de una planta R<sub>1</sub> derivada de la planta anterior obteniendo un genotipo tetraploide, altamente sexual, también registrado.

Estos dos materiales se utilizan actualmente para realizar estudios básicos acerca del modo de reproducción (apomixis) y en la generación de poblaciones de mapeo para el análisis del modo reproductivo y de caracteres agronómicos de importancia.

- Dos plantas heptaploides apomícticas a

partir de un cultivar tetraploide apomítico. Una de dichas plantas fue registrada como cultivar Don Luis, en honor al Ing. Luis Mroginski. Este cultivar presenta buenas características agronómicas.

La obtención de variación *in vitro*, en **trigo**, y en otros cereales como **cebada** es altamente dependiente del genotipo. En trigo, se ha obtenido resistencia a *Helminthosporium sativum*. En **cebada** se encontró tolerancia en somaclones regenerados en medios conteniendo el herbicida glifosato. En **sorgo**, se obtuvo variación estable para altura, número de macollos, mayor producción de granos y tolerancia a suelos ácidos y a sequía. En plantas regeneradas de **triticale y trigo** se informaron variaciones a nivel del ADN mitocondrial y de proteínas de la semilla, como las gliadinas. En ambas especies se informaron cambios estables en varios caracteres como por ejemplo: longitud de la espiga, número de granos por espiga, peso de granos, densidad y biomasa de la espiga, contenido de proteína en grano, altura de la planta, fertilidad, número de macollos, color del grano, tiempo de espigazón, dureza, sensibilidad al ácido abscísico y color de las glumas. Las mutaciones observadas fueron de dominantes a recesivas y viceversa.

El único antecedente en nuestro país de búsqueda de variación somaclonal en trigo, es un estudio acerca de la estabilidad *in vitro* de tres cultivares con elevados niveles de inestabilidad cariotípica espontánea. Se evaluaron caracteres tales como estructura cromosómica, patrón de gliadinas, color del grano y de las glumas, tipo de aristas, niveles de clorofila y morfología de la planta. Sólo se detectó, como consecuencia del cultivo *in vitro*, una translocación no descrita previamente y un patrón diferente de gliadinas. En este caso, se concluyó que la técnica no fue efectiva para inducir variación que pudiera ser utilizada en programas de mejoramiento (Franzone y col., 1996).

La variación somaclonal ha sido vastamente registrada y tratada en la bibliografía. Sin embargo, su utilización como estrategia para la obtención de variantes agronómicamente aprovechables, es discutida. No obstante, además de su aplicación en la generación de variabilidad genética (que resulta particular para cada

grupo de plantas, especie, genotipo y condiciones), es un fenómeno a considerar cuando se requiera utilizar cultivo de tejidos para regenerar plantas genotípicamente estables.

## 7 Lecturas recomendadas

- Bregitzer P.; Halbert, S.; Lemaux, P. 1998. Somaclonal variation in the progeny of transgenic barley. *Theor. Appl. Genet.* 96: 421-425.
- Bregitzer P.; S. Zhang; M.-J. Cho; P. G. Lemaux. 2002. Reduced Somaclonal Variation in Barley Is Associated with Culturing Highly Differentiated, Meristematic Tissues. *Crop Sci.* 42:1303-1308.
- Cassels A.C.; Croke, J.T.; Doyle, B.M. 1997. Evaluation of Image Analysis, Flow Cytometry, and RAPD Analysis for the Assessment of Somaclonal Variation and Induced Mutation in Tissue Culture Derived *Pelargonium* Plants. *Angew. Bot.* 71: 125-130.
- Cardone S.; Polci P.; Selva JP.; Mecchia M.; Pessino, S.; Hermann, P.; Cambi, V.; Voigt, P.; Spangenberg G.; V. Echenique. 2006. Novel genotypes of the subtropical grass *Eragrostis curvula* for the analysis of apomixis (diplospory). *Euphytica* 151 (2): 263-272.
- Croughan S.; S. Quisenberry; M. Eichhorn; JR., P.Colyer; T. Brown. 1994. Registration of Brazos-R3 bermudagrass germplasm. *Crop Science.* 34: 542.
- Da Silva C.; C. Mangolin; A. Mott; M. Machado. 2008. Genetic diversity associated with *in vitro* and conventional bud propagation of *Saccharum* varieties using RAPDS analysis. *Plant Breeding.* Vol 127 (2): 160-165.
- Eastman P.; Webster, F.; Pitel, J.; Roberts, D. 1991. Evaluation of somaclonal variation during somatic embryogenesis of interior spruce (*Picea glauca engelmannii* complex) using culture morphology and isozyme analysis. *Plant Cell Reports* 10: 425-430.
- Evans D.; Sharp, W. 1983. Single gene mutations in tomato plants regenerated from tissue culture. *Science.* 221: 949-951.
- Franzone P; suarez Y; Solari R; favret A; Ríos R; DÍAZ PALEO A. 1996. Somaclonal variation in three Argentinean varieties of *Triticum aestivum* with different karyotypic instability. *Plant Breed.* 115: 89-93.
- Gallego F.; Martínez, I.; Celestino, C., Toribio, M. 1997. Testing somaclonal variation using RAPDs in *Quercus suber* L. somatic embryos. *Int. J. Plant Sci.* 158(5): 563-567.
- Gauto Acosta M. 2008. Evaluación a campo de somaclones de *Lolium perenne* y sus progenies

- de medios hermanos. Trabajo de intensificación para optar al grado de Ingeniero Agrónomo. Facultad de Agronomía. (UBA). 73 pp.
- Geier T. 1991. Chromosome variability in callus produced plants. In Genetics and Breeding of Ornamental Species. 79-106. Harding & Mol. eds.
- Kaepler S. M.; Phillips, R. L. 1993. In Vitro Cell Dev. Biol. 29: 125-130.
- Jain S. Micropropagation of selected somaclones of *Begonia* and *Saintpaulia*. 1997 J. Biosci. Vol. 22, Number 5 (585-592).
- Krikorian A. 1991. Estabilidad genotípica en células, tejidos y plantas derivados del cultivo in vitro. En Cultivos de tejidos en la Agricultura. CIAT. Roca W. y Mroginsky L. Eds.
- Larkin P.; Scowcroft. W. 1981. Somaclonal variation - a novel source of variability from cells cultures from plant improvement. Theor. Appl. Genet. 60, 197-214.
- Linacero R.; Freitas Alves E.; Vasquez A. 2000. Hotspots of DNA instability revealed through the study somaclonal variation in rye. Theor Appl Genet. 100: 506-511.
- Madlung A.; Comai, I. 2004. The effect of stress on genome regulation and structure. Annals of Botany. 94:481-495.
- Olmos S.; G. Lavia, M. Di Renzo, L. Mroginski; V. Echenique. 2002. Genetic analysis of variation in micropropagated plants of *Melia azedarach* L. In Vitro Cellular & Developmental Biology Plant. 38: 617-622.
- Peschke V.M.; Phillips, R.L. 1992. Genetic implications of somaclonal variation in plants. Adv. Genet. 30: 41-75. Scandalios & Wright eds.
- Podwyszynska M. 2005. Somaclonal variation in micropropagated tulips based on phenotype observation. Journal of Fruit and Ornamental Plant Research. Vol 13: 109-122.
- Svabova L; Lebeda A. 2005. In Vitro Selection for Improved Plant Resistance to Toxin-Producing Pathogens. Journal of Phytopathology. Vol 153: 52-64.
- Umbeck P.; B. Gengenbach. 1983. Reversion of male sterile T cytoplasm maize to male fertility in tissue culture. Crop. Science. 23: 584-588.
- Venuto BC, Croughan SS, Pitman WD, Jessup RW, Renganayaki K, Burson BL. 2007. Variation among hexaploid *Paspalum dilatatum* Poir. regenerants from tissue culture. Australian Journal of Experimental Agriculture. 47:1109-1116
- Zalewska M., Lema-Ruminska, J.; Miler, N. 2007. In vitro propagation using adventitious buds technique as a source of new variability in *Chrysanthemum*. Scientia Horticultural. Vol. 113: 70-73.



## II. CAPÍTULO 6

### Aplicación de la transformación genética al mejoramiento vegetal

Marina L. Díaz, Diego C. Zappacosta,  
Pascual M. Franzone y Raúl D. Ríos

#### 1 Introducción

El mejoramiento genético vegetal se originó hace aproximadamente 10.000 años, cuando el hombre se hizo agricultor y comenzó la domesticación de las plantas. En el siglo XX, con la incorporación de los cruzamientos sexuales, el conocimiento de la biología floral de las especies, el desarrollo de la genética y de la estadística experimental, entre otros avances, se conformaron los métodos de mejoramiento actualmente utilizados. El mejoramiento genético convencional se basa en la existencia de variabilidad genética para los caracteres que se desea mejorar y la manipulación de la misma mediante la reproducción sexual. Esto hace que el aprovechamiento de la variabilidad este restringido por barreras de compatibilidad reproductiva. El reciente desarrollo de métodos de transferencia de genes que no implican cruzamientos, como la transformación genética, permite superar esta limitación, abriendo nuevas perspectivas en el mejoramiento de las plantas.

La transformación genética o transferencia de genes, técnica también conocida como ingeniería genética, permite introducir en plantas genes provenientes no sólo de otras especies vegetales muy alejadas desde el punto de vista evolutivo sino incluso de hongos, virus, bacterias y animales. Asimismo, a través de esta tecnología es posible modificar la expresión de genes presentes en el genoma de la planta. Si bien se suele denominar OGMs (organismos genéticamente modificados) a las plantas obtenidas mediante esta metodología, es oportuno considerar que todas las plantas que han sido mejoradas por el hombre a partir de su domesticación, han sido modificadas genéticamente por lo cual son OGMs y consideramos más apropiado utilizar para las plantas obtenidas mediante transformación genética la deno-

minación de plantas transgénicas. Conviene destacar que la transformación de plantas es una tecnología que aporta variabilidad genética conocida sin alterar el fondo genético por lo que puede ser considerada como una metodología conservativa de mejoramiento, similar, en ese sentido, a la mutagénesis inducida o a la retrocruza. Este último aspecto es de gran importancia ya que la creación de cultivares es un proceso acumulativo, es decir, que se desea introducir características favorables sin perder las mejoras logradas anteriormente. El germoplasma transgénico es posteriormente incorporado al proceso de mejoramiento el cual dependerá de la especie y del tipo de cultivar a obtener.

En 1983 se informaron los primeros experimentos de expresión de un transgen (gen introducido por vía asexual) en células vegetales y al año siguiente se obtuvieron las primeras plantas transgénicas (tabaco y petunia). Desde entonces se ha extendido la aplicación de esta tecnología a más de 120 especies. Las plantas transgénicas se obtienen por diversos métodos, los que han sido modificados para cada especie en particular, aumentándose de esta forma la eficiencia de los mismos.

Entre sus aplicaciones de interés agropecuario se encuentran la obtención de plantas con resistencia a estreses bióticos (virus, insectos, hongos y bacterias), a estreses abióticos (salinidad, sequía, etc.), tolerancia a herbicidas para facilitar el control de malezas y modificación de la calidad nutricional de los cultivos entre otras. La transformación genética también constituye una herramienta útil para estudios básicos que permiten conocer y/o profundizar acerca de la estructura y función de genes específicos, aspecto particularmente relevante en la era genómica.

#### 2 Requerimientos para la obtención de plantas transgénicas

Para todas las técnicas de transformación desarrolladas hasta el momento es necesario disponer del **transgen** a introducir (éste consiste en secuencias regulatorias y codificante clonadas en un vector de transformación) y de una metodología eficiente para su transferencia al genoma vegetal. Luego se induce el de-

sarrollo de plantas mediante distintas técnicas de cultivo de tejidos y se procede al análisis molecular de las plantas regeneradas *in vitro* para identificar aquellas que porten y expresen el o los transgenes en los niveles deseados. Finalmente, mediante experimentos de campo y laboratorio se estudia el comportamiento de los individuos transgénicos y su descendencia.

Por lo tanto, los elementos básicos que se requieren en trabajos de transformación genética en plantas son:

1. Un sistema eficiente de cultivo de tejidos que permita regenerar plantas completas y fértiles.
2. Vectores apropiados, que permitan el clonado del gen de interés y su transferencia al tejido blanco de transformación.
3. Un protocolo de transformación, es decir, un sistema de transferencia de genes y de selección del material transformado.
4. Herramientas de análisis molecular para detectar la presencia del transgen y los productos del mismo en la planta.

Antes de iniciar la descripción de los ítems anteriores, resulta importante destacar que para lograr una planta transgénica deben ocurrir tres procesos en una misma célula:

- a) El transgen debe ser transferido al interior de la célula,
- b) El transgen debe integrarse al ADN celular y
- c) Se debe regenerar una planta completa a partir de la célula transgénica en la que se verificaron los procesos a y b.

Dado que las células tienen diferente competencia o capacidad de respuesta para cada uno de estos procesos, la puesta a punto de un protocolo eficiente de transformación genética requiere maximizar la cantidad de células competentes para todos ellos de manera simultánea.

### 2.1 Cultivo de tejidos vegetales

Los métodos de transformación más utilizados actualmente se basan en la obtención de células transgénicas y posterior recuperación de plantas completas y fértiles a partir de las mismas a través del cultivo y selección *in vitro*. Esto es posible gracias a la propiedad de

totipotencia expresada por las células vegetales. En la mayoría de las especies se observa una importante influencia del genotipo en la respuesta al cultivo *in vitro*, razón por la cual es frecuente que se utilicen genotipos modelo (de alta respuesta) con fines experimentales los cuales no necesariamente presentan interés agronómico. Por ello, cuando se usa esta tecnología con fines de mejoramiento genético es muy importante conocer la respuesta morfo-genética de distintos genotipos con adaptación local, de modo de poder transformar directamente genotipos con valor agronómico. Por otra parte, en el cultivo *in vitro*, un prolongado período de subcultivos puede inducir la aparición de variación somaclonal no deseable ya que es conveniente introducir sólo las modificaciones que se desean y en forma controlada.

### 2.2 Construcción del vector

Para introducir un transgen es necesario que el mismo sea incorporado previamente en un vector (por ejemplo un plásmido). Para ello, mediante la tecnología del ADN recombinante, se digiere el ADN extraído de un organismo, cualquiera sea su origen, con enzimas de restricción. Estas endonucleasas tienen la capacidad de reconocer en el ADN secuencias específicas compuestas por pocos nucleótidos y posteriormente producir cortes en las mismas. Los fragmentos de restricción así obtenidos pueden ser ligados enzimáticamente a vectores de clonado, obteniéndose, de este modo, clones de ADN recombinante (ver parte I). Alternativamente, se puede utilizar una tecnología de clonado basada en el mecanismo de recombinación sitio-específica del fago lambda conocida como sistema GATEWAY (Invitrogen).

Un transgen está compuesto por una secuencia codificante (región traducible comprendida entre los codones de iniciación y terminación de la traducción) y por secuencias regulatorias que determinan tanto el momento del desarrollo de la planta como el tejido y el nivel en el que se expresará. Muchas veces los genes utilizados provienen de bacterias (procariotas) y usualmente no pueden expresarse eficientemente en células eucariotas. Por lo tanto, resulta necesario reemplazar las secuencias regulatorias bacterianas por otras aptas para su

expresión en células vegetales. En este contexto, la secuencia más importante es la correspondiente al promotor, sitio del ADN al cual se une la enzima ARN-polimerasa para iniciar el proceso de transcripción (**Fig. 1**). En la elección del promotor a utilizar en la construcción del transgen, hay que considerar la especie vegetal a transformar, ya que su nivel y patrón de expresión pueden variar al ser utilizados en especies distintas a la de origen del mismo. Dentro de los promotores se distinguen aquellos constitutivos, que permiten la expresión del gen en toda la planta en forma continua, los inducibles, que responden a factores ambientales y finalmente los específicos de tejido que permiten la transcripción en órganos y tejidos particulares (**Tabla 1**). Los promotores pueden ser modificados con el fin de aumentar la expresión de los genes que regulan. Así, pueden truncarse, adicionárseles un intrón o unírseles secuencias de otros promotores. Además, es necesario que el transgen disponga de una re-

gión no traducida en el extremo 3' del mismo, que incluya la señal de corte y poliadenilación (terminador), necesaria para el correcto procesamiento del transcripto correspondiente.

Por otra parte, el nivel y patrón de expresión del gen también puede estar afectado por la posición del mismo en el genoma de la planta, fenómeno conocido como "efecto de posición". Esto puede deberse a la presencia de otras secuencias regulatorias cercanas como intensificadores y silenciadores, a la estructura de la cromatina, al patrón de metilación, etc. Este aspecto es particularmente relevante considerando que con los métodos de transformación genética nuclear utilizados actualmente, las inserciones de los transgenes en el genoma de la planta son al azar. Así, la cantidad de proteína producida por un transgen comúnmente varía más de 100 veces entre individuos transformantes obtenidos en el mismo experimento. La solución para evitar los "efectos de posición" es dirigir la secuencia de interés a sitios específi-



**Figura 1:** Estructura molecular del transgen

**Tabla 1:** Promotores usados en la transformación genética de plantas.

<b>Promotores</b>	<b>Constitutivos</b>	<i>nos</i> (nopalina sintetasa de <i>Agrobacterium</i> ) 35 S (virus del mosaico del coliflor) <i>Ubi</i> (ubiquitina de maíz) <i>Act 1</i> (actina de arroz)
	<b>Específicos de tejido</b>	<i>TA 29</i> (tapete de la antera de tabaco) faseolina (cotiledones de poroto) <i>TobRB 7</i> (raíz de tabaco) patatina (tubérculo de papa) glutenina (endosperma de trigo)
	<b>Inducibles</b>	subunidad pequeña de Rubisco (luz) alcohol deshidrogenasa-1 (anaerobiosis) proteína de shock térmico (calor)

cos del genoma vegetal, elegidos por su patrón de expresión adecuado y estable. Esto podría lograrse usando sistemas de recombinación heteróloga sitio-específicos o mediante reemplazo génico por recombinación homóloga.

### 2.3 Métodos de transformación

La pared celular constituye un obstáculo para la entrada del ADN a la célula que todos los métodos de transformación genética tienen que superar de algún modo. Estos métodos pueden dividirse en:

a) Transformación mediada por *Agrobacterium*, vector biológico que participa del proceso de transferencia.

b) Métodos de transformación genética directa, también llamados físicos, mediante los cuales, por distintos mecanismos no biológicos, se introduce el ADN en la célula.

#### 2.3.1 Transformación mediada por *Agrobacterium*

Las bacterias del género *Agrobacterium* son Gram-negativas, aeróbicas obligadas y viven en el suelo. Ellas son capaces de desarrollar un crecimiento saprofítico o parasítico. El género comprende cuatro especies fitopatogénicas, dos de ellas ampliamente estudiadas (*A. tumefaciens* y *A. rhizogenes*). Ambas especies son capaces de infectar una amplia variedad de especies de dicotiledóneas. La patogénesis se inicia a partir de heridas, provocando la proliferación de las células individuales infectadas. Así, *A. tumefaciens* causa tumores, enfermedad que se conoce como ‘agalla de la corona’ y *A. rhizogenes* induce la proliferación de raíces dando lugar a la enfermedad denominada ‘raíz en cabellera’. La capacidad patogénica de estas bacterias está asociada a la presencia de megaplásmidos (150-200 kilo pares de bases o Kb) llamados Ti (por ‘tumor-inducing’ o inductor de tumores) o Ri (por ‘root-inducing’ o inductor de raíces), presentes en algunas de las agrobacterias. Se demostró que durante la patogénesis un fragmento de estos plásmidos llamado T-ADN (por ‘transfer-DNA’), es transferido a la célula vegetal, donde se integra al ADN cromosómico y cuya expresión causa proliferación de células de la planta a través de la síntesis y alteración de la respuesta a hormonas

vegetales. El T-ADN contiene genes, llamados oncogenes, que se expresan eficientemente en la célula vegetal infectada y producen síntesis de hormonas vegetales, responsables de la proliferación anormal del tejido. Además, contiene los genes responsables de la síntesis de opinas (fuente de carbono y nitrógeno para la bacteria).

El sistema mediado por *A. tumefaciens* es el más estudiado y utilizado en la transformación de plantas. El T-ADN está delimitado por dos repeticiones directas imperfectas de 25 pares de bases (pb) que lo flanquean, llamadas bordes derecho e izquierdo (**Fig. 2**). Estos bordes son los únicos elementos en *cis* necesarios para dirigir el procesamiento del T-ADN. Cualquier fragmento de ADN ubicado entre estos bordes puede ser transferido a la célula vegetal. Contrariamente a lo que sucede con otros tipos de secuencias móviles de ADN, como los transposones autónomos, el T-ADN no codifica los productos que median su transferencia. El T-ADN es una región relativamente grande (ca. 20 Kb) que contiene genes con secuencias regulatorias (promotores y señales de poliadenilación) típicamente eucarióticas que permiten su expresión eficiente en células vegetales.

Desde la década de 1950 se sabe que los tumores de la agalla de la corona se desarrollan si hay *A. tumefaciens* patogénicas en presencia de heridas. Un aspecto destacable de esta bacteria es que usa la respuesta de la planta a las heridas (cicatrización y defensa) como quimioatractivo y activador del proceso de patogénesis. Luego de la adhesión de la bacteria a la célula vegetal, etapa en la que están involucrados genes cromosómicos bacterianos (*chv A*, *chv B*, *chv E*, *cel*, *psc A* y *att*), se produce el procesamiento y la transferencia del T-ADN, mediados por los genes *vir* (por virulencia) que son inducidos por azúcares y compuestos fe-

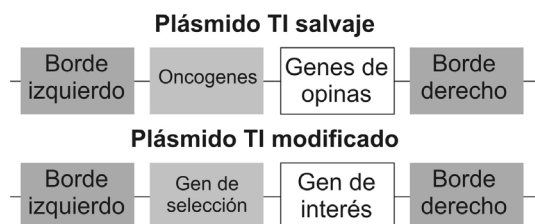


Figura 2: Estructura del T-DNA



nólicos, como la acetosiringona, producidos por células vegetales heridas. Se conoce con mucho detalle la etapa de la patogénesis que ocurre en la bacteria (ver revisión de Gelvin, 2000). La región de virulencia (de alrededor de 30 Kb) está organizada en operones esenciales para la transferencia del T-ADN (*vir A*, *vir B*, *vir D*, *vir G*) o que permiten incrementar la eficiencia de la transformación (*vir C*, *vir E*). El conocimiento de la interacción de *Agrobacterium* con las plantas se profundizará a partir de la disponibilidad (2001) de la secuencia nucleotídica completa del genoma de esta bacteria.

El desarrollo de la patogénesis causada por *Agrobacterium* representa una situación única en la naturaleza: la transferencia de un elemento genético (T-ADN) de un organismo procarionota a un organismo eucariota superior, con su subsiguiente integración y expresión en el genoma hospedante. Este mecanismo de ingeniería genética natural es aprovechado para la transferencia de genes de interés a las plantas. Para ello, los oncogenes y los genes de síntesis de opinas presentes en el T-ADN, son reemplazados por un marcador seleccionable y el gen a transferir (**Fig. 2**). El método llamado 'del explante' (Horsch *et al.*, 1984), consiste en inocular un explante (órgano vegetal) con *A. tumefaciens*, dejar la bacteria en contacto con el mismo durante un cierto tiempo en el cual se produce la transferencia del T-ADN modificado. Posteriormente se transfieren los explantes a un medio de cultivo que contiene un antibiótico que actúa como bacteriostático para detener el desarrollo de *Agrobacterium* y el agente selectivo correspondiente al gen marcador seleccionable utilizado. Una vez completado el protocolo de cultivo y selección *in vitro* se recuperan las plantas transgénicas (**Fig. 3**). El análisis por hibridación molecular de plantas transgénicas y su progenie ha demostrado que el T-ADN se incorpora al ADN cromosómico y se hereda en forma estable.

Los plásmidos Ti sin oncogenes se denominan 'desarmados', son de gran tamaño y resulta difícil introducir genes en su T-ADN usando técnicas habituales de ADN recombinante. Por esta razón se han desarrollado dos sistemas de vectores para introducir genes en *Agrobacterium*: los vectores cointegrados, y los bina-

rios. En los vectores cointegrados la inserción del ADN que contiene los transgenes dentro del T-ADN de un plásmido Ti desarmado, se logra por recombinación homóloga. En el sistema alternativo, sobre la base de la capacidad de los genes de virulencia de actuar en *trans*, movilizándolo el T-ADN presente en otro plásmido, se construyen los denominados vectores binarios que contienen un T-ADN no oncogénico en un plásmido pequeño con un origen de replicación funcional en un amplio espectro de hospedantes, características que permiten su fácil manipulación genética en *E. coli*. La introducción de un vector binario a una cepa de *Agrobacterium* que contenga un plásmido llamado 'helper' de virulencia (con la región *vir* intacta y sin T-ADN), la hace apta para la transferencia del T-ADN a las células vegetales. Si bien ambos tipos de vectores son herramientas eficientes para la transformación se aprecia una notable preferencia por el uso de vectores binarios.

La integración del T-ADN en el ADN cromosómico de la planta se produce al azar, por recombinación ilegítima, preferencialmente en regiones genómicas con actividad transcripcional y mediado por proteínas de la bacteria y de la planta. Este sistema de transformación permite transferir fragmentos de ADN de hasta 150 Kb. Para ello se utilizan vectores especiales que tienen características tanto de cromosomas artificiales de bacterias (BACs) como de vectores binarios. El uso de estos vectores permite clonar fragmentos de ADN de gran tamaño y posteriormente transferirlos al genoma vegetal vía *A. tumefaciens*. Esta tecnología es de gran utilidad para el clonado posicional de genes.

Más de 600 especies vegetales son hospedantes naturales de *A. tumefaciens*. Estas son en su mayoría dicotiledóneas y gimnospermas y más raramente monocotiledóneas. Para muchas de ellas se han puesto a punto protocolos de transformación genética basados en este vector. El interés en el desarrollo de este tipo de tecnología hizo que se expandiera el rango de hospedantes de *A. tumefaciens* a especies que no son susceptibles a este patógeno en condiciones naturales. Esto fue posible debido al amplio conocimiento que se tiene de la interacción *Agrobacterium*/célula vegetal. Así, se

ha puesto énfasis en la utilización de *A. tumefaciens* como vector de transformación en gramíneas y se han desarrollado protocolos en varias especies de este grupo (arroz, maíz, cebada, trigo y gramíneas forrajeras). Cabe notar que su aplicación es rutinaria en arroz y maíz.

La transformación con *A. rhizogenes* tiene fundamentalmente dos aplicaciones: la producción de raíces con alta tasa de crecimiento capaces de sintetizar metabolitos secundarios como productos farmacéuticos, aditivos alimentarios y cosméticos y por otra parte representa una valiosa herramienta para la estimulación de la rizogénesis en especies recalcitrantes, por ejemplo leñosas.

### **2.3.2 Transformación directa o por métodos físicos**

La primer metodología de transformación genética de plantas desarrollada fue la de *A. tumefaciens*, pero no resultó inicialmente de utilidad para abordar la transformación de especies de gran importancia económica como los cereales, debido a que éstos no son hospedantes naturales de esta bacteria. Esta situación condujo al desarrollo de los métodos físicos de transformación. En ellos el transgen es introducido en la célula vegetal mediante distintas técnicas que se detallan a continuación:

#### Transformación de protoplastos

La introducción y expresión de ADN foráneo en protoplastos fue el primer método de transferencia directa claramente demostrado en plantas. Se disponía, para algunas especies, de procedimientos eficientes para la regeneración a partir de protoplastos, células que carecen temporariamente de pared celular que es la principal barrera para la introducción de ADN en las células vegetales. Los protoplastos pueden ser aislados en forma mecánica o por un proceso enzimático que digiere la pared. Así se obtiene una suspensión que contiene millones de células individuales lo que favorece la transformación de células aisladas.

Los protoplastos pueden ser transformados utilizando polietilenglicol (PEG), electroporación, microinyección o liposomas. La transformación de protoplastos mediada por PEG es el método más comúnmente usado. Tanto el

PEG como la electroporación producen poros en la membrana plasmática por alteración de la polaridad de la misma y por ellos penetra el ADN foráneo. En el último caso se somete a los protoplastos a un campo eléctrico. Ambas técnicas permiten el tratamiento simultáneo de un gran número de protoplastos. La microinyección es un método difícil y laborioso que consiste en la introducción de ADN dentro de protoplastos individuales mediante el uso de capilares de inyección y micromanipulador y con esta metodología es posible lograr un alto grado de integración del ADN foráneo.

Cabe notar que el cultivo de protoplastos es la metodología más sofisticada de cultivo *in vitro* de plantas por lo cual representa gran complejidad experimental y no está disponible más que para algunas especies y dentro de ellas particularmente en genotipos modelo. Por ello se desarrollaron metodologías alternativas de transformación directa como la electroporación de tejidos, el uso de fibras de carburo de silicio para facilitar la entrada del ADN a las células en cultivo y el bombardeo con microproyectiles o micropartículas siendo este último el más utilizado actualmente dentro de los métodos físicos.

#### Bombardeo de micropartículas

Es un proceso por el cual micropartículas cubiertas con ADN son aceleradas por un gas comprimido e introducidas en células vegetales. Inicialmente la fuerza impulsora de las partículas estaba dada por pólvora, pero luego fue reemplazada por el sistema de helio comprimido que brinda una mejor regulación de la fuerza, distribución de microproyectiles y mayor reproducibilidad entre bombardeos. Se utilizan microproyectiles de oro o tungsteno (químicamente inertes) cuyos tamaños van desde 0,5 a 3  $\mu\text{m}$ , que al ser disparados a grandes velocidades pueden atravesar pared y membranas de la célula vegetal bombardeada sin causarle daños significativos (Klein *et al.*, 1987). Como blanco de bombardeo se pueden emplear diversos tipos de explante vegetal, desde células o protoplastos hasta plántulas completas, pasando por tejidos organizados en embriones y meristemas. El explante es generalmente sometido a un tratamiento osmótico pre y post bombardeo que produce plasmólisis celular,

evitando que el impacto y la penetración de las partículas dañen las células (**Fig. 3**). Los vectores plasmídicos que se usan en este tipo de procedimiento solo requieren un origen de replicación que permita un alto número de copias de los mismos en *E. coli*, aspecto que facilita en la práctica la preparación del ADN necesario en grandes cantidades para la transformación genética por este método.

Esta técnica, sin embargo, tiene limitaciones. Algunas especies oponen una resistencia natural a la penetración de las partículas, dada por cutículas endurecidas, paredes celulares lignificadas o superficies vellosas. Sin embargo la principal limitación del método continúa siendo la baja relación entre el total de células sometidas al bombardeo y el número de células que logran incorporar de manera permanente el transgen. A pesar de esta limitación, la versatilidad de la aceleración de partículas para introducir transgenes ha superado muchas de las barreras asociadas a otros métodos de transformación, como son el rango de huéspedes de *Agrobacterium* y las dificultades inherentes al cultivo y regeneración de protoplastos. Este método, conocido también como biolística, ha demostrado ser una buena opción para la producción de plantas transgénicas de

soja, sorgo, papaya, espárrago, caña de azúcar y trigo, entre otras.

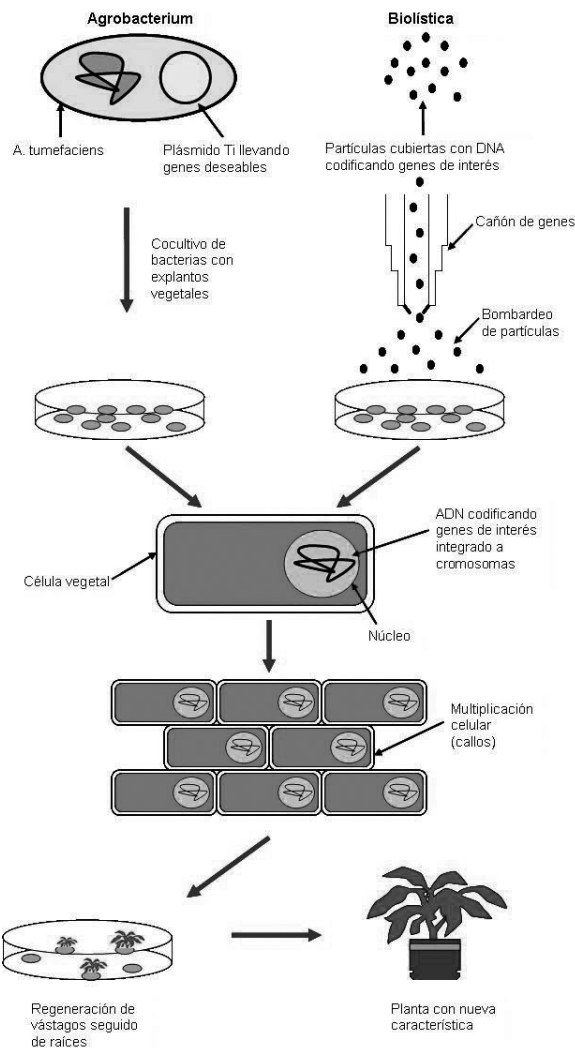
### 2.3.3 Transformación directa vs. indirecta

Con el transcurso de las investigaciones se ha logrado optimizar los protocolos de transformación para muchas especies vegetales. Ambas vías de transformación presentan ventajas y desventajas que las hacen más o menos convenientes para los distintos fines. En la **Tabla 2** se presenta una comparación entre la transformación del genoma nuclear mediada por *A. tumefaciens* y el método biolístico como representante de los métodos de transformación directa.

Uno de los procesos necesarios para obtener plantas transgénicas, es la integración del transgen al ADN cromosómico. Con relación a este proceso existen diferencias entre ambos métodos. *A. tumefaciens* posee un mecanismo natural muy eficiente para transferir, al núcleo celular, el T-ADN que contiene el transgen y producir una integración aleatoria en regiones cromosómicas con actividad transcripcional. Generalmente, los patrones de inserción de transgenes son más sencillos en comparación a los producidos por el método biolístico. Así el número de copias que se incorpora es bajo

**Tabla 2:** Comparación entre métodos de transformación del genoma nuclear

	<i>Agrobacterium</i>	Método biolístico
<b>Especies a transformar</b>	dicotiledóneas y algunas monocotiledóneas	sin limitaciones
<b>Eficiencia de transformación</b>	alta	baja
<b>Tipo de integración en el genoma vegetal</b>	aleatoria en regiones con transcripción bajo número de copias independientes precisa	aleatoria multicopia en tandem imprecisa
<b>Construcción de vectores</b>	compleja	simple
<b>Dependencia del genotipo vegetal</b>	mayor	menor



**Figura 3:** Comparación de técnicas de transformación directa e indirecta

y cuando hay más de una copia, éstas suelen distribuirse en loci independientes, lo que facilita la posterior eliminación de las copias no deseadas por segregación. Por otra parte, el mecanismo involucrado hace que se transfiera un segmento de ADN definido: el T-ADN, acompañado de proteínas que lo protejen de la acción de nucleasas. Por el contrario, en el método biolístico el ADN transferido no está asociado a proteínas, por lo que al ser parcialmente degradado solo parte del mismo llega al núcleo donde se produce, con baja eficiencia, la integración al azar en el ADN cromosómico. En este método es común que se produzcan inserciones de varias copias del transgen en

un mismo sitio del genoma con distinto grado de reordenamiento de la secuencia del transgen. Este último aspecto es considerado como una desventaja ya que tanto desde el punto de vista de la percepción pública como el de la optimización de la expresión de los transgenes, es deseable disponer de inserciones simples y bien definidas molecularmente. Esto es particularmente relevante cuando se tiene por objetivo el desarrollo de productos comerciales. Así, el fenómeno de silenciamiento de transgenes (pérdida de su expresión), observado en algunos casos, puede resultar de la presencia de varias copias del transgen. Por otra parte, cuando se utiliza el método biolístico, el seg-

mento de ADN plasmídico que se integra al ADN cromosómico no está definido como en el caso de T-ADN, sino que es de longitud variable.

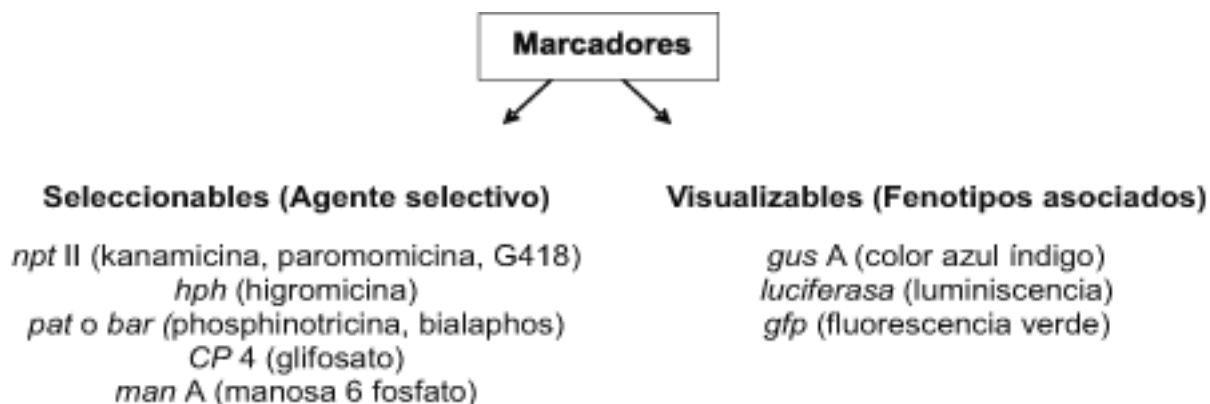
La construcción de vectores es mucho más sencilla en el caso del método biolístico. Este aspecto puede resultar importante en el contexto del desarrollo de proyectos de investigación, en los cuales muchas veces se requiere utilizar un número considerable de vectores diferentes.

Es conveniente destacar que en ambos métodos la integración del transgen en el genoma se produce al azar, como consecuencia de ello, diferentes plantas transgénicas provenientes de un mismo experimento, presentan inserciones en distintos sitios del genoma receptor. Así, la expresión fenotípica de un transgen será diferente en distintas plantas transgénicas que contienen el mismo transgen, por lo cual para clarificar la situación se considera a cada una de ellas como eventos de transformación distintos.

La eficiencia de aplicación de cualquiera de estos métodos de transformación al mejoramiento vegetal, depende en gran medida de la disponibilidad de genotipos con muy buena respuesta al cultivo *in vitro*. Finalmente, cabe mencionar que en el caso de *A. tumefaciens*, debido a que se trata de una interacción biológica planta/bacteria, la dependencia del genotipo es mucho más acentuada ya que influyen también en la eficiencia de transformación, tanto el genotipo de la planta como el de la bacteria.

### 2.3.4 Genes de selección e informadores

En el proceso de obtención de plantas transgénicas, los genes marcadores seleccionables (Tabla 3), generalmente de origen bacteriano, cumplen un rol de relevancia ya sea en forma individual para poner a punto un protocolo de transformación o como acompañantes del gen de interés que se desea introducir. El gen marcador seleccionable otorga a las células transgénicas que lo expresan una importante ventaja, con respecto a las células no transgénicas, al permitirles crecer en un medio de cultivo que contiene el agente selectivo correspondiente. Entonces, si el gen marcador codifica para una proteína que confiere resistencia a agentes fitotóxicos como antibióticos o herbicidas, las células que lo expresen podrán crecer y desarrollarse en medio de cultivo que contenga dichos agentes selectivos mientras que las células no transgénicas no lo harán (selección negativa). En otros casos, la ventaja estará dada por la posibilidad de utilizar diferencialmente un sustrato. Así, el gen *man A* de *E. coli* permite a las células vegetales que lo expresan utilizar la manosa como fuente de carbono, lo cual no es posible para las células no transgénicas (selección positiva). Es necesario tener en cuenta que algunas especies pueden poseer una resistencia natural a los agentes selectivos antes mencionados. Por ello, es importante evaluar este aspecto cuando se elige el gen de selección a utilizar, así como el agente selectivo y las concentraciones del mismo.



**Tabla 3:** Genes marcadores utilizados comúnmente en transformación de plantas

En la puesta a punto de las metodologías de transferencia de genes son asimismo de gran utilidad los genes marcadores visualizables o informadores que posibilitan la observación directa de las células que los expresan (**Tabla 3**). Estos genes codifican para proteínas que no están presentes en las células vegetales y producen un fenotipo característico de fácil y rápida observación. Así, el gen *gus A* proveniente de *E. coli* codifica para la enzima  $\beta$ -glucuronidasa. Esta proteína puede ser detectada cualitativamente por tinción histoquímica. En esta técnica, en presencia de un sustrato específico esta enzima cataliza la producción de un precipitado azul-índigo de fácil visualización. Este marcador también puede ser detectado cuantitativamente usando técnicas fluorométricas. Como ejemplo de la aplicación de estos métodos de detección, podemos mencionar el estudio de la expresión de promotores en plantas. Para ello se puede construir un vector con la secuencia codificante de *gus A* bajo el control del promotor en estudio y una secuencia de terminación de transcripción, y obtener plantas transgénicas. El estudio de las mismas permite conocer el patrón de expresión espacio-temporal del promotor (en que tipo de células y cuando se expresa) por tinción histoquímica y su nivel de expresión por fluorimetría.

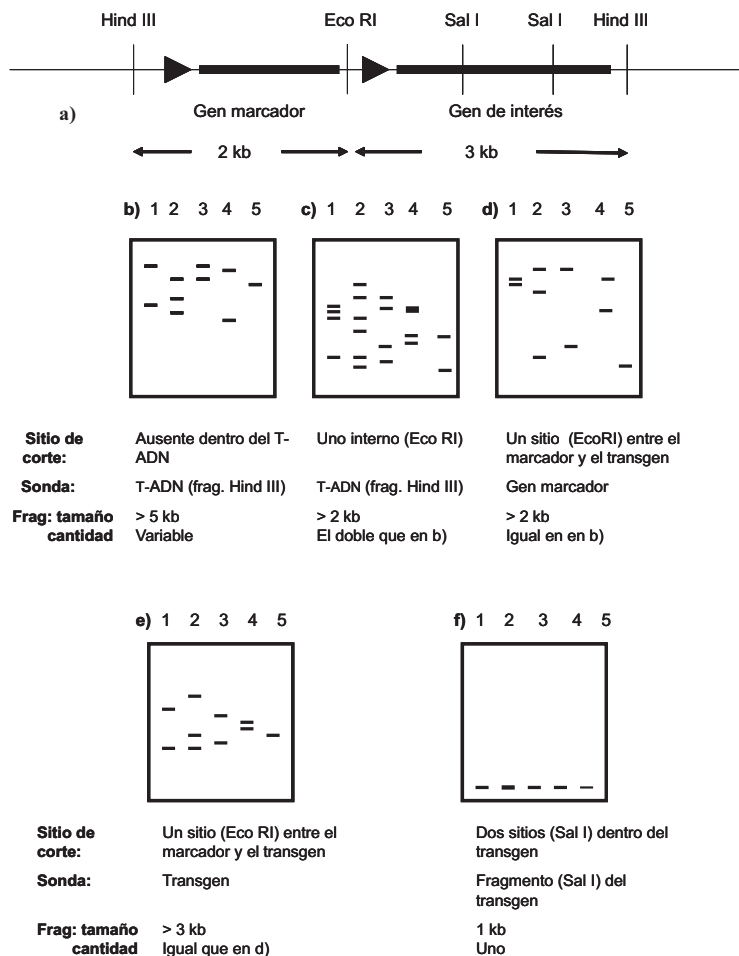
Por otra parte, el gen de la proteína fluorescente verde, *gfp* ("green fluorescent protein"), aislado de una medusa, se ha convertido en un gen marcador visualizable *in vivo* muy utilizado. Cuando se ilumina con luz ultravioleta o luz azul, tejidos que contienen células que expresan este gen, se observa un brillo verde fluorescente. Al ser este un ensayo no destructivo, es decir que permite conservar vivo el material en estudio, representa una herramienta de utilidad para visualizar diferentes etapas del proceso de transformación.

## 2.4 Técnicas de detección del transgen y sus productos

Luego de la selección *in vitro*, las plantas regeneradas deben ser analizadas por métodos moleculares para identificar aquellas que porten y expresen los transgenes en los niveles deseados. Para ello se emplean técnicas como:

- **Reacción en Cadena de la Polimerasa (PCR):** con el uso de 'primers' homólogos a la secuencia del transgen se puede determinar la presencia del mismo. Esta técnica permite hacer un 'screening' rápido de las plantas regeneradas, aunque tiene algunas limitaciones. Así, un resultado positivo indica solamente que en la muestra existe una secuencia que amplifica con los 'primers' utilizados pero no demuestra que el transgen se encuentre incorporado al genoma de la planta. Para esto último se requiere un estudio de hibridación molecular o el estudio de la transmisión sexual del transgen. La robustez de esta prueba se maximiza utilizando temperaturas de "annealing" o apareamiento de "primers" altas (más de 60 °C) y "primers" relativamente largos (más de 20 nucleótidos). Una variación de esta técnica, PCR en tiempo real, resulta de utilidad para evaluar el número de copias del transgen incorporadas al genoma de la célula vegetal (ver parte I).

- **Hibridación molecular o 'Southern blotting':** es una de las herramientas moleculares más poderosas para caracterizar las plantas transgénicas. Además de indicar la presencia o no de la secuencia de interés, da información acerca de la integración de la misma en el genoma de la célula huésped (número de copias integradas y número de loci en los cuales se produjo la integración). En la **Figura 4** se representa el análisis de un caso hipotético de transformación mediante *Agrobacterium*, aunque este es aplicable también a plantas obtenidas por métodos directos. El esquema 4-a representa parte del vector que contiene el gen de interés y el gen marcador. Si se utiliza como sonda el T-ADN completo y la digestión del ADN genómico se realiza con enzimas de restricción que no cortan dentro del T-ADN, el número de bandas del patrón representa el número de sitios de inserción. El tamaño de las bandas será mayor que el T-ADN ya que se están utilizando enzimas que cortan por fuera de este. Plantas transgénicas obtenidas en forma independiente tendrán patrones de hibridación diferentes (4-b). Por otra parte, si la enzima tiene un sitio de corte único dentro del T-ADN, utilizando la sonda antes mencionada, esta dará dos señales de hibridación



**Figura 4:** Patrones de Hidridación de Southern Blot (Tomado de Bhat y Srinivasan, 2002)

para cada sitio de inserción independiente del T-ADN (4-c). La aparición de un número impar de bandas indica que en un mismo sitio se insertaron múltiples copias, ocurrieron rearrreglos o se produjo el truncado del T-ADN. Si se digiere la muestra con una enzima de restricción que posee un sitio único de corte entre el gen marcador y el gen de interés y utilizando como sonda el gen marcador, se obtendrá una sola banda por cada inserción independiente (4-d). Si la membrana se hibrida con el gen de interés (4-e) aparecerá un patrón que tendrá el mismo número de bandas que el anterior. La ausencia de una banda indica la inserción de una copia truncada de uno de los genes y la

pérdida del sitio de corte. La aparición de una banda del tamaño del T-ADN puede indicar la presencia de repeticiones en tandem. Para estimar el número de copias del transgen en un locus transgénico, se deben utilizar enzimas de restricción con sitios flanqueantes del transgen y como sonda el transgen. Se observará una sola banda del tamaño del transgen, sin embargo la intensidad de la señal variará entre las distintas plantas analizadas de acuerdo al número de copias del mismo (4-f). Para realizar esta comparación es necesario utilizar simultáneamente estándares conteniendo un número conocido de copias del transgen. Esta determinación no es del todo precisa ya que puede

existir variabilidad en los resultados debido a diferencias en la digestión de las muestras y a la posible degradación del ADN.

En lo que respecta a la cuantificación, esta técnica ha sido reemplazada en gran medida, en los últimos años, por la PCR en tiempo real o “real time PCR”, antes mencionada, dado que esta última resulta menos costosa en reactivos y tiempo y requiere menos cantidad de material vegetal para realizar el análisis.

**- Análisis de ARN:** Tanto las técnicas de RT-PCR (transcripción reversa seguida de PCR) como la hibridación de ARN o “Northern blotting” pueden resultar valiosas herramientas para detectar los transcritos de interés presentes en la muestra. La RT-PCR presenta algunas ventajas frente al Northern blotting: se requiere poca cantidad de material, posee una alta sensibilidad y la muestra es de fácil preparación. Por otra parte, la hibridación del ARN utiliza como sonda el transgen, ofrece información sobre el tamaño del transcrito y permite su cuantificación. Cabe destacar que la RT-PCR puede también ser utilizada como técnica de cuantificación de transcritos.

**- ELISA y “Western blotting”:** aquellas plantas que poseen la secuencia de interés y la expresan como transcrito, pueden ser analizadas con estas técnicas inmunológicas a fin de determinar la presencia de la proteína codificada por el transgen.

La actividad de la proteína codificada por el transgen debe ser medida, ya sea por ensayos bioquímicos o por bioensayos, a fin de interpretar correctamente el fenotipo de la planta obtenida. Además, es conveniente confirmar la estabilidad meiótica de los transgenes estudiando su transmisión sexual a la generación siguiente.

Estas técnicas se encuentran descriptas detalladamente en la parte I y su aplicación para el análisis de plantas transgénicas en Register (1997) y en Bhat y Srinivasan (2002).

Finalmente el comportamiento de los individuos transgénicos se estudia en laboratorio, invernáculo y campo. Cabe aclarar que para realizar experimentos de invernáculo y de campo en Argentina es necesario contar con au-

torización de la Comisión Nacional Asesora de Promotores Biotecnología Agropecuaria (CONABIA) (ver parte VI).

### **3 Aporte de la transformación genética a la variabilidad genética**

Como ya se mencionó, la transformación genética permite introducir variabilidad genética novedosa en las diferentes especies vegetales. Este aporte puede ser realizado de diferentes maneras:

a) La expresión de secuencias codificantes no existentes en la especie (ejemplo: proteínas insecticidas de origen bacteriano).

b) La expresión de nuevas formas alélicas de genes que ya están presentes en el genoma (ejemplo: tolerancia al glifosato en soja).

c) La expresión de secuencias codificantes presentes en el genoma pero bajo el control de nuevas secuencias regulatorias que modifican su nivel o patrón de expresión (ejemplo: proteínas relacionadas a la patogénesis).

d) La inhibición de la expresión de genes residentes en el genoma.

Para lograr la inhibición de la expresión génica se pueden usar diferentes técnicas: cosupresión, ARN antisentido y la denominada interferencia del ARN. Todas ellas se basan en un proceso denominado silenciamiento génico que involucra la degradación del ARN mensajero producido por un gen determinado (ver capítulo siguiente).

Un inconveniente que se encuentra tanto en el mejoramiento convencional como en la ingeniería genética es la incorporación de múltiples genes. A esto se lo llama piramidización de genes y es necesario para la expresión de características complejas de naturaleza poligénica (por ej. vías biosintéticas), la obtención de resistencia a enfermedades más duradera donde se requiere más de un gen para evitar la caída de la resistencia, para otorgar resistencia conjunta a varios patógenos o la incorporación de múltiples características.

Algunas estrategias para alcanzar este objetivo son las siguientes:

a) cruzamiento de dos plantas transgénicas con distintos transgenes

b) re-transformación en eventos consecutivos



### c) co-transformación en el mismo evento

La co-transformación es la alternativa más utilizada y ventajosa, ya que los transgenes co-transformados tienden a co-integrarse en la misma posición en el genoma en la mayor proporción de los eventos transgénicos. Esto aseguraría la no segregación de los genes, lo cual es muy útil cuando se quiere incorporar una característica poligénica.

La mayoría de los procesos metabólicos que pueden ser sujetos a manipulación dependen de la interacción entre numerosos genes y el flujo dentro de las vías bioquímicas es a veces coordinado con otras vías, por lo que la manipulación efectiva sólo puede ser alcanzada coordinando el grado de expresión de los distintos transgenes. Coordinar la expresión de múltiples transgenes sobre una misma vía metabólica es muy difícil de alcanzar, ya que la expresión de los transgenes depende del lugar donde se inserten en el genoma. Existen algunas estrategias elegidas para la manipulación de estas características:

- transgenes policistronicos (un promotor con varios transgenes)
- poliproteínas

El avance técnico en esta área es uno de los mayores desafíos de la biotecnología vegetal, ya que las nuevas generaciones de transgénicos requieren de la incorporación de múltiples características.

#### 4 Obtención de plantas transplastómicas

Las células vegetales contienen tres genomas: el nuclear, el plastídico y el mitocondrial. Los métodos presentados en las secciones precedentes tienen como blanco de transformación al genoma nuclear. Sin embargo en los últimos años la obtención de plantas transplastómicas ha recibido notable atención. Así, el genoma de los plástidos (plastoma) se ha convertido en un blanco atractivo para la ingeniería genética ya que esta tecnología ofrece una serie de ventajas sobre la transformación del genoma nuclear:

1. Se pueden obtener altos niveles de expresión de los transgenes y elevados niveles de acumulación de las proteínas

codificadas por ellos (hasta el 47% de las proteínas solubles totales). Cabe notar que el nivel de acumulación observado a partir de transgenes nucleares es generalmente inferior al 1%.

2. Es posible expresar un transgen "operón" con varias secuencias codificantes bajo el control de un mismo promotor.
3. No se observa efecto de posición debido a que la integración del transgen está dirigida a un sitio particular del plastoma. Esto hace que no sea necesario obtener una gran población de transformantes, a diferencia de lo que actualmente ocurre con las plantas transgénicas nucleares.
4. No se observa silenciamiento de transgenes.
5. Para las especies que tienen transmisión materna de los plástidos (la mayoría). Se minimiza la dispersión de los transgenes por el polen debido a la baja de transmisión de los plástidos por el mismo.
6. A pesar de la naturaleza eminentemente procariota del sistema genético de los plástidos, se ha demostrado que estas organelas pueden procesar proteínas eucariotas permitiendo su correcto plegamiento y la formación de puentes disulfuro gracias a la presencia de proteínas chaperonas y sistemas enzimáticos endógenos.

Los protocolos existentes para la obtención de plantas transplastómicas se basan en la transferencia de genes por el método biolístico, un eficiente proceso de cultivo y selección *in vitro*, y la utilización de vectores con secuencias homólogas al genoma del cloroplasto. A diferencia de lo que ocurre en la transformación del genoma nuclear, la integración dirigida de los transgenes en el plastoma se realiza mediante recombinación homóloga. Para lograr esto, los vectores contienen los transgenes de interés flanqueados por secuencias con alta homología al ADN plastídico. La homología entre el ADN del vector y del ADN del cloroplasto determina la potencialidad del plásmido para ser utilizado en la transformación genética del plastoma. En este contexto, la existencia de secuencias altamente conservadas en el

plastoma de varias especies es explotada para el diseño de vectores universales, potencialmente útiles en la transformación de los cloroplastos de numerosas especies. La expresión exclusivamente plastídica de los transgenes queda asegurada a través de la utilización de promotores específicos del cloroplasto, cuya transcripción tiene características típicamente procarióticas.

## 5 Perspectivas

La continuidad en el desarrollo de tecnología de transferencia de genes en plantas es importante tanto para ampliar el rango de especies a transformar como el de genotipos dentro de cada especie. En los últimos años se nota un creciente interés en el control del nivel de expresión de transgenes así como de su regulación espacial y temporal. La expresión de los genes nucleares eucarióticos está controlada en varios niveles que involucran la transcripción propiamente dicha, el procesamiento, transporte y estabilidad de los transcriptos y su traducibilidad, así como la estabilidad, modificación y compartimentalización del producto proteico final. Por otra parte, es necesario aprovechar la experiencia acumulada en cuanto a la expresión de transgenes, de modo de diseñar protocolos de transferencia que minimicen las posibilidades de silenciamiento génico. En este contexto, está recibiendo notable atención el estudio de la recombinación homóloga como mecanismo para la transformación del genoma nuclear.

Del análisis comparativo de métodos presentado cabe destacar tres aspectos actualmente en desarrollo. Por un lado, continúa el interés en extender el uso de la transformación mediada por *A. tumefaciens* a especies que no son hospedantes naturales de la bacteria. Por otro lado, como fuera anteriormente mencionado, los métodos de transformación genética actualmente en uso requieren de genotipos con una excelente respuesta al cultivo *in vitro*, lo cual resulta en una clara limitación cuando se desea aplicar esta tecnología a los materiales usados por los mejoradores, los que generalmente no poseen esta característica. Por ello, se está tratando de reducir y si es posible evitar la fase de cultivo *in vitro* durante el proceso de

transformación, lo que permitirá ampliar el rango de genotipos a los cuales se podrá aplicar esta moderna tecnología. Así, se desarrolló un método eficiente y reproducible de transformación *in planta* para *Arabidopsis thaliana*. Este consiste en infiltrar con vacío plantas adultas de esta especie, con una suspensión de *A. tumefaciens* de modo que los tejidos y órganos reproductivos sean invadidos por la bacteria. Posteriormente se propuso una simplificación de este método que consiste en reemplazar el tratamiento de vacío y sumergir la inflorescencia en una suspensión bacteriana que incluye un surfactante. Las plantas transgénicas que se obtienen por autofecundación de las plantas así transformadas, son hem cigotas y se demostró que esta metodología produce la transformación de la gameta femenina. Finalmente, debido a la importante restricción que puede significar la utilización de métodos de transformación genética protegidos por patentes, se ha está estudiando el desarrollo de métodos que utilizan, a partir de herramientas derivadas de *Agrobacterium*, otras bacterias como vectores alternativos de transformación.

El logro de determinados objetivos como puede ser el redireccionamiento de una ruta metabólica o la modificación de un carácter complejo de interés agronómico como por ejemplo la resistencia durable a enfermedades fúngicas, requiere de la integración de múltiples transgenes y de la expresión coordinada de los mismos en la planta. Este tema está recibiendo notable atención.

Los métodos de transformación que hemos descrito se basan en la utilización de genes marcadores seleccionables. Sin embargo, hay razones por las cuales la presencia del marcador no es deseable en plantas transgénicas que se liberan al ambiente. Hay un manifiesto rechazo por parte del público con respecto a la utilización de genes de resistencia a antibióticos y en determinadas situaciones, el uso de genes de resistencia a herbicidas puede ser inconveniente. Por otra parte, desde un punto de vista experimental puede ser necesario transformar una misma planta varias veces con distintos transgenes y con la metodología convencional no hay más que un número limitado de genes marcadores disponibles. Por ello se

consideran estrategias alternativas para la obtención de plantas transgénicas libres de marcadores (Puchta, 2003). La estrategia que ha recibido más esfuerzo hasta ahora, plantea la remoción vía cruzamiento sexual del marcador seleccionable luego de haber obtenido la planta transgénica. Una posibilidad para ello es cotransformar. Esto implica que el transgen de interés y el marcador seleccionable estén presentes en vectores distintos y que posteriormente se separen los genes por segregación luego de la reproducción sexual. Sin embargo, esta puede ser una alternativa poco atractiva en circunstancias particulares como cuando no se dispone de un protocolo eficiente de cotransformación o cuando no se desea reproducir sexualmente la planta transgénica. Otra opción dentro de esta línea es la utilización del sistema de recombinación sitio-específica de origen viral Cre/lox. Esta estrategia requiere primero la obtención de plantas transgénicas portadoras del gen de la recombinasa (cre) y posteriormente la introducción del transgen flanqueado por sitios lox de reconocimiento. Más recientemente se ha propuesto como estrategia alternativa el desarrollo de métodos de transformación que no usen marcadores seleccionables. La idea subyacente es incrementar la frecuencia de transformación a través del uso de genes que promuevan la regeneración.

Finalmente, las plantas transgénicas actualmente cultivadas comercialmente, en su mayoría resistentes a herbicidas y a insectos, nos permiten ganar experiencia y acumular el conocimiento necesario para desarrollar nuevas y mejores generaciones de productos. A pesar de estos avances en el desarrollo de esta tecnología, su aplicación comercial se ve limitada por el elevado costo que tiene el pasaje por el estricto sistema de controles que regula su aplicación. Esto restringe notoriamente las especies a ser usadas en transformación y los caracteres a mejorar mediante la misma.

## 6 Lecturas recomendadas

- Bhat S., Srinivasan S. 2002. "Molecular and genetics analyses of transgenic plants: considerations and approaches". *Plant Science* 163:673-681.
- Birch R. 1997. "Plant transformation: Problems and strategies for practical applications". *Annu. Rev. Plant Physiol.Plant Mol.Biol.* 48: 297-326.
- François I, Broekaert W., Cammue B. 2002. "Different approaches for multi-transgene-stacking in plants". *Plant Science* 63: 281-295.
- Gelvin S. 2003. "Agrobacterium-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool". *Microbiology and Molecular Biology Reviews*, Mar. 2003: 16-37.
- Gepts P. 2002. "A Comparison between Crop Domestication, Classical Plant Breeding, and Genetic Engineering". *Crop Sci.* 42:1780-1790.
- Halpin C. 2005. "Gene stacking in transgenic plants – the challenge for 21<sup>st</sup> century plant biotechnology". *Plant Biotechnology Journal*, 3:141-155.
- Hanin M., Paszkowski J., 2003. "Plant genome modification by homologous recombination". *Current Opinion in Plant Biology* 6: 157-162.
- Horsch R., Fraley R., Rogers S., Sanders P., Lloyd A., Hoffman N. 1984. "Inheritance of functional foreign genes in plants". *Science* 223: 496-498.
- Klein T., Wolf E., Wu R., Sanford J. 1987. "High velocity microprojectiles for delivering nucleic acids into living cells". *Nature* 327: 70-73.
- Lee L. y Gelvin S. 2008. "T-DNA Binary Vectors and Systems". *Plant Physiology Vol.* 146: 325-332.
- Maliga P. 2004. "Plastid transformation in higher plants". *Annu. Rev. Plant Biol.* 55:289-313.
- Puchta H. 2003. "Marker-free transgenic plants". *Plant Cell, Tissue and Organ Culture* 74: 123-134.
- Register J (1997) "Approaches to evaluating the transgenic status of transformed plants". *TiBTech*, 15:141-6.
- Small I., 2007. "RNAi for revealing and engineering plant gene functions". *Current Opinion in Biotechnology* 18:148-153.



## II. CAPÍTULO 7

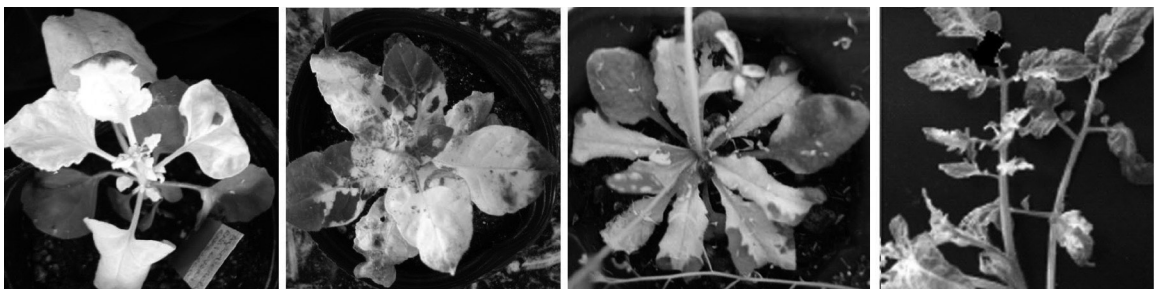
### Usos del silenciamiento génico para el análisis funcional de genes candidatos.

Cecilia Vazquez Rovere, Ariel Bazzini, Cecilia Rodríguez, Natalia Almasia y Sebastián Asurmendi

El progreso de las disciplinas genómicas ha generado una rápida acumulación de información biológica relacionada a la estructura de los genomas, su organización, su composición en número de genes, su grado de homología y sus relaciones evolutivas. Esta información puede ser aprovechada para el análisis funcional de genes promoviendo un gran avance en diferentes aspectos como el fitosanitario o el de respuestas a estreses ambientales, entre otros. Para el estudio de la funcionalidad de genes candidatos existen diversas estrategias. La más antigua (no por eso menos efectiva o utilizada) consiste básicamente en sobre-expresar un gen, es decir aumentar la expresión del gen de interés a niveles tan altos que su función o actividad puede aumentar y así ser medible u observable mediante técnicas moleculares o simplemente fenotípicamente. Por ejemplo, la sobreexpresión de genes del metabolismo puede producir un aumento importante en el crecimiento de una planta. La segunda estrategia es antagónica a la primera y consiste en disminuir los niveles de expresión del gen en estudio. Esta reducción (total o par-

cial) puede también mostrar cambios visibles o medibles, los cuales permiten evidenciar la actividad del gen. Por ejemplo si uno reduce la expresión de un gen involucrado en la síntesis de pigmentos, dicha planta puede presentar una disminución en la coloración de sus hojas. Esta disminución dependerá del grado de reducción en la expresión del gen, pudiendo presentar desde pequeñas manchas amarillentas o hasta la falta total de coloración foliar (Figura 1). Este capítulo considerará únicamente esta segunda estrategia de reducción de la expresión génica mediada por el mecanismo denominado silenciamiento génico. Los métodos basados en el silenciamiento génico presentan varias ventajas respecto a otros empleados en el campo de la genómica funcional. Una de ellas es la posibilidad de silenciar varios o todos los miembros de una familia multigénica simultáneamente. Esto es altamente relevante cuando se realizan estudios de genómica en organismos poliploides, como por ejemplo en el caso del trigo en el cual existe una gran redundancia genética (Travella y col., 2006). Otra de las ventajas es que permite obtener distintos grados de silenciamiento del gen (débil, intermedio y fuerte) siendo útil para el análisis de fenotipos asociados a genes cuya disminución total resultaría letal. Finalmente, también se puede mencionar la capacidad de generar un silenciamiento eficiente a partir de secuencias de pequeño tamaño permitiendo el estudio de genes cuya secuencia completa se desconoce (Lu y col., 2003).

Actualmente existen diversas metodologías basadas en el silenciamiento génico. Algunas de las mismas requieren de la generación de



**Figura 1.** Foto mostrando el blanqueo del tejido de plantas en las cuales se silenció el gen PDS en las especies *Nicotiana benthamiana*, *Nicotiana tabacum*, *Arabidopsis thaliana* y *Solanum lycopersicum* respectivamente de izquierda a derecha

plantas transgénicas portadoras de una construcción capaz de silenciar al gen candidato, mientras que en otras el silenciamiento se desencadena a partir de la replicación de un virus que porta en su genoma un fragmento del gen a silenciar. Dependiendo del problema específico a resolver se deberá considerar la tecnología disponible más adecuada. Es importante destacar que no todas las herramientas moleculares pueden utilizarse en todas las especies, es decir que en la elección de la metodología deberá tenerse en cuenta la especie con la que se esté trabajando. Una de las metas de la ciencia actual es masificar estas técnicas, es decir que independientemente del organismo (girasol, papa, hongo, etc.) uno pueda optar por una variedad de técnicas moleculares para solucionar el problema planteado de la forma más adecuada. Asimismo, es importante mencionar que el silenciamiento génico es un mecanismo ampliamente distribuido en los procesos naturales de los seres vivos. Todos los organismos regulan la expresión de sus genes dependiendo de su información génica, del ambiente y del estado de desarrollo, y en los organismos superiores una de las formas de regular negativamente (disminuir) la expresión de sus genes es mediante el silenciamiento. En los últimos 15 años el estudio de este mecanismo natural, ha permitido utilizarlo de forma dirigida a fin de disminuir la expresión de genes endógenos a elección.

En este capítulo se introducirán los conceptos básicos del mecanismo de silenciamiento y se desarrollarán las distintas metodologías más frecuentemente empleadas en la actualidad para determinar la función de los genes y/o analizar su diversidad funcional.

### **1. Silenciamiento génico**

El silenciamiento génico hace referencia a diversos procesos que producen la reducción secuencia-específica de la expresión génica y en los cuales la molécula involucrada es el RNA. Las moléculas de RNA que participan del silenciamiento génico presentan un rango de tamaño de 20 a 30 bases nucleotídicas por lo que se las denomina “small” RNAs (sRNAs por su siglas en inglés) o bien “silencing” RNAs debido a su participación en este fenómeno

(Baulcombe, 2004 y Ossowski y col., 2008). Distintas clases de sRNAs han sido descritas dado que se originan por diferentes vías y que poseen funciones celulares diversas (ver tabla 1). A pesar de estas diferencias, los procesos de silenciamiento comparten numerosas etapas en común. En todos los casos al inicio se produce la formación de una molécula de RNA doble cadena (dsRNA), la cual es específicamente clivada por una enzima perteneciente a la familia de las RNasas tipo III (similar a Dicer – DCL por su siglas en inglés (Fire y col., 1998) la cual genera fragmentos de 21 a 25 nucleótidos de ambas polaridades (Baulcombe y col., 2004 y Dorokhov y col., 2007). Las dos cadenas de los fragmentos resultantes son luego separadas presumiblemente por una helicasa y una de las cadenas, denominada hebra guía, es incorporada al complejo enzimático endógeno RISC (por sus siglas en inglés “RNA-induced silencing complex”) el cual es capaz de producir el silenciamiento de aquellos genes que presenten homología con la secuencia de dicha cadena incorporada. Dentro de este mecanismo general pueden existir variaciones a nivel de la formación del dsRNA, de la clase de DCL involucrada en el clivaje del RNA y de las moléculas efectoras que reclutan a los sRNAs para luego ejercer su acción. Dicha acción puede ser la modificación del nivel de metilación de las secuencias regulatorias del gen blanco, la degradación del RNA mensajero (mRNA) o bien la reducción específica de su traducción. En consecuencia, el silenciamiento génico es un mecanismo regulatorio que puede actuar en distintos niveles: transcripcional (denominado silenciamiento transcripcional, TGS), post-transcripcional (silenciamiento post-transcripcional, PTGS) y traduccional (Chapman y Carrington, 2007).

A continuación se explicarán brevemente los mecanismos de silenciamiento mencionados:

#### **1.1 Silenciamiento transcripcional**

En el silenciamiento transcripcional, su acción está dirigida hacia las secuencias regulatorias y como consecuencia el gen río abajo de dicha región deja de transcribirse. En este proceso los sRNAs dirigen modificaciones en el DNA o en las proteínas asociadas a él, por

Clase*	Descripción	Biogénesis y origen genómico	Función
miRNA	Micro RNAs	Transcriptos de genes miRNAs forman estructuras secundarias de RNA doble cadena que son procesados por Dicer o por RNAsas tipo III	Regulación post transcripcional de los transcritos de un amplio rango de genes
siRNAs primarios	sRNAs pequeños interferentes	Procesamiento de RNA doble cadena o de estructuras secundarias Por miembros de la familia Dicer	Unión a RNAs targets complementarios, guía para la iniciación de siRNAs secundarios
siRNAs secundarios	sRNAs pequeños interferentes	Se forman por la actividad de RNA polimerasas RNA dependientes, las cuales generan RNAs doble cadena	Regulación post transcripcional de los transcritos, formación y mantenimiento de la heterocromatina.
tasi RNAs (Trans acting siRNAs)	sRNAs que actúa en trans	Transcriptos de genes TAS son clivados por un mecanismo dependiente de miRNAs, luego son convertidos en RNA doble cadena por RdRP, seguido del procesamiento por Dicer	Regulación post transcripcional de los transcritos.
natsi RNAs ( natural antisense transcript-derived siRNAs)	siRNAs derivados de transcritos antisentidos naturales	Pares de transcritos sense y antisense son procesados por Dicer	Regulación post transcripcional de la expresión de genes involucrados en mecanismos de defensa y respuesta a stress
piRNAs ( Piwi - interacting siRNAs)	RNAs que interaccionan con Piwi	El mecanismo de biogénesis no se conoce completamente, sería dependiente de argonauta e independiente de Dicer	Supresión de transposones y retroelementos en líneas germinales de moscas y de mamíferos.

**Tabla 1.** Clases de sRNAs y su función.\* se han mantenidos las siglas en inglés. RdRP: sigla en Inglés para RNA polimerasa – RNA dependiente. Tabla adaptada de Carrington y col., 2007.

ejemplo produciendo la metilación de los residuos de citosina de DNA y la consecuente modificación de la cromatina. Sin embargo la metilación *per-sé* no es suficiente en todos los casos para el establecimiento del silenciamiento. A modo de ejemplo puede citarse la existencia de mutantes de los genes *sil1* y *mom1*, para los cuales se ha reportado la pérdida del TGS a pesar de que las secuencias se encuentran metiladas (Furner y col., 1998 y Amedeo

y col., 2000). Vaucheret y sus colaboradores propusieron al silenciamiento transcripcional como un sistema de control de la expresión génica de transposones y accidentalmente de transgenes con características similares a ellos (2001).

### 1.2 Silenciamiento post-transcripcional

El silenciamiento génico post-transcripcional (PTGS) es un mecanismo inducible y especí-

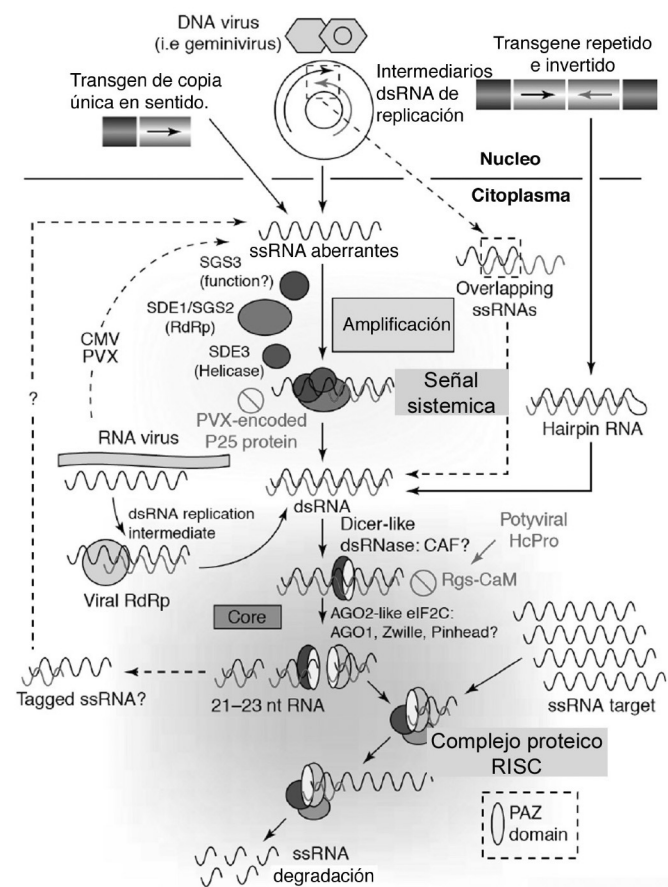
fico en el cual la molécula blanco es el RNA mensajero. Se ha propuesto que el PTGS evolucionó como un sistema de defensa antiviral en plantas dado que restringe la acumulación o la diseminación de los virus dentro de las mismas (Vance y Vaucheret 2001; Dunoyer y col., 2004). Como se mencionara previamente el PTGS se desencadena contra una molécula de RNA mediado por la homología de secuencia del sRNA, pudiendo reconocer tanto un RNA del propio individuo como el de un organismo patógeno, por ejemplo un virus. Cuando se produce el reconocimiento del RNA viral durante la infección se produce la degradación del mismo que con lleva al fracaso de la infección. Este mecanismo es también responsable del fenómeno de co-supresión de genes endógenos, del silenciamiento de los transge-

nes y recientemente también se lo ha ligado a procesos de regulación de genes endógenos. Si bien se describió por primera vez en plantas, ha sido reportado en hongos (donde se denomina "quelling"), y en sistemas animales como nematodos, insectos o mamíferos (donde se lo denomina interferencia mediada por RNA o RNAi). En plantas el PTGS es desencadenado eficientemente por intermediarios de dsRNA que pueden tener su origen en: transgenes nucleares cuyos transcriptos forman naturalmente estructuras tipo horquilla o "hairpin" (hpRNAi), genes que se expresan en altos niveles, transcritos aberrantes que pueden ser reconocidas como dsRNA o estructuras de dsRNA que se generen durante la replicación de un virus. Cualquiera sea su origen, una vez formadas las estructuras de dsRNA, las mismas son reconocidas por una enzima RNasa DCL y procesadas dando lugar a siRNAs de 21 a 25 pb. Estos siRNAs se asocian al complejo RISC, el cual es capaz de degradar específicamente todo aquel RNA que comparta homología con la secuencia que desencadenó el proceso y también pueden ser utilizados como templado por las polimerasas celulares dependientes de RNA (RdRps) para sintetizar la segunda hebra del gen a silenciar amplificando de esta forma la señal de silenciamiento (Figura 2).

Estos siRNAs se asocian al complejo RISC, el cual es capaz de degradar específicamente todo aquel RNA que comparta homología con la secuencia que desencadenó el proceso y también pueden ser utilizados como templado por las polimerasas celulares dependientes de RNA (RdRps) para sintetizar la segunda hebra del gen a silenciar amplificando de esta forma la señal de silenciamiento (Figura 2).

### 1.2.1 Amplificación y movimiento del PTGS

Una de las características del mecanismo de silenciamiento en plantas es la capacidad de ser sistémico, es decir de poder propagar a toda la planta la reducción del gen el cual resultara inicialmente disminuido en un único tejido. Por ello se ha propuesto la existencia de una señal de silenciamiento capaz de moverse a través del organismo propagando el mismo. Estas vías de movimiento de la señal se clasifican según la distancia que alcancen. Así, en la vía de movimiento de corta distancia los siRNAs generados en una célula son capaces de migrar directamente a células adyacentes a través de los plasmos-



**Figura 2.** Esquema del silenciamiento génico inducido por virus, por sobre expresión de transgén o expresión de secuencias duplicadas invertidas. Se detalla como el RNA es procesado en siRNA de 21-25 pb (Figura adaptada de Voinnet 2001)



desmos (conexiones directas entre células) y establecer en estas últimas el silenciamiento. Una vez que en la célula adyacente es degradado el mRNA homólogo al siRNA, se generan nuevos siRNAs, amplificando así la señal que se trasladará por toda la planta. Se ha reportado asimismo que esta amplificación se genera a partir de la síntesis de dsRNA mediada por una RNA polimerasa RNA dependiente RdRp (Baulcombe, 2004). En la vía de movimiento de larga distancia la señal se traslada a través del floema permitiendo el movimiento a partes muy distantes del lugar de origen. Esto fue demostrado mediante experimentos que emplean injertos en los que plantas silenciadas para un dado transgen son capaces de producir el silenciamiento de ese transgen en plantas no silenciadas. Si bien aún no se conoce cuál es la señal de silenciamiento de larga distancia, es






muy probable que sea distinta de la señal que migra de célula en célula debido a que los dos procesos pueden ser separados farmacológica o genéticamente (Dunoyer y Voinnet 2008).

La dispersión del silenciamiento tiene una gran importancia principalmente en la defensa frente a virus, ya que a partir del sitio inicial de infección, donde se indujo el silenciamiento, se produce una señal que genera la defensa antiviral en los tejidos adyacentes que de resultar previa a la llegada del virus limitará la propagación del mismo al resto de la planta.

### 1.3 Regulación génica por microRNAs

Como hemos mencionado, el silenciamiento es un mecanismo natural por el cual se regula la expresión de genes. En particular, los microRNAs (miRNAs) son un tipo específico de moléculas de RNA de bajo peso molecular (21 pb) que controlan la acumulación de mRNA endógenos al regular finamente la expresión de diferentes genes. A diferencia de los siRNAs los miRNAs

son codificados por genes endógenos, por lo tanto difieren en su biogénesis. Los miRNAs son genes endógenos completos, con la única particularidad de que no codifican para ninguna proteína, sino que su RNA mensajero es procesado generando una pequeña molécula de 21 pb aproximadamente. Los miRNAs han sido descritos sólo en organismos eucariotes, desde plantas a humanos (Bartel 2004). Resumidamente, los miRNAs se originan a partir de transcritos que se pliegan sobre sí mismos dando lugar a una estructura precursora de RNA doble cadena (denominadas pre-miRNAs), las cuales son procesadas de manera que se produce una única cadena de miRNA de pequeño tamaño (21-22 nt) mientras que la cadena complementaria, denominada miRNA\*, es degradada. En la Figura 3A se muestran ejemplos de estructuras secundarias

Construcción genética	Estructura del transcripto de RNA	% PTGS	n
Gen candidato en sentido		7	27
Gen candidato en antisentid		4	25
hpRNA (brazos de gen candidato, gen GUS como secuencia espaciadora)		58	43
hpRNA (brazos de gen candidato, intrón en antisentido como secuencia espaciadora)		65	34
hpRNA (brazos de gen candidato, intrón en sentido como secuencia espaciadora)		96	23

**Figura 3. A)** Se representan las típicas estructuras de rulo o "stem-loop" de los miRNAs. miR164 de Arabidopsis y sus homólogos de Oryza y Populus. El segmento correspondiente al miRNA maduro se muestra en color rojo (adaptado de Jones-Rhoades y col. 2006). **B)** Esquema del pre-miR164a de Arabidopsis y el pre-miR-GFP sintético. En rojo se encuentra el miR164 maduro y en verde la porción reemplazada para generar el miR artificial miR-GFP maduro. Se muestra la energía libre ( $\Delta G$ ) y la entalpía ( $\Delta U$ ) del ensamblado.

de los miRNAs. Una vez que el transcripto del miRNA es procesado y el miRNA maduro resultante es incorporado al complejo RISC, el cual al dirigirse contra el mRNA blanco reprime su expresión génica ya sea mediante degradación específica o inhibición de la traducción. En las plantas el proceso de degradación es el más común, mientras que en el reino animal la inhibición de la traducción es el proceso más frecuente. (Hutvagner y Zamore, 2002; Doench y col. 2003; Zeng y col. 2003). Estas pequeñas moléculas son extraordinariamente importantes en la regulación o modulación de los procesos celulares, un solo miRNA puede regular la expresión muchos genes por lo tanto incidir en la regulación de múltiples procesos ya que muchos de los genes blancos son factores de transcripción, porejemplo la alteración de la expresión de un único miRNA en plantas puede cambiar totalmente el fenotipo de la misma.

## 2. Métodos que emplean el mecanismo de silenciamiento para disminuir la expresión génica en plantas

Varios métodos han sido desarrollados para disminuir la expresión de un gen en estudio

basados en el mecanismo de silenciamiento génico. Entre los mismos se encuentra el uso de moléculas antisentido, de moléculas con estructura de horquilla (hpRNAi), de miRNAs artificialmente diseñados y el silenciamiento génico inducido por virus (ver Tabla 2).

A continuación se exponen algunos de los métodos más usados en la actualidad para lograr establecer el silenciamiento génico. Cada uno de estos métodos presenta ventajas y desventajas que determinan que sean más o menos aptos para el estudio particular del gen asilenciar y en la especie que se desea emplear, con lo que resulta de suma importancia evaluarlas antes de elegir la metodología. Estas metodologías pueden a su vez agruparse en dos categorías principales de acuerdo a que involucren la producción de plantas transgénicas o no. Siendo ésta una metodología un tanto lenta y laboriosa.

### 2.1 Metodologías para el análisis de genes candidatos que implican transgénesis.

La transformación estable con construcciones capaces de inducir el silenciamiento de un gen candidato es quizá la aproximación más

**Tabla 2:** Métodos basados en el mecanismo de silenciamiento génico, para disminuir la expresión de un gen.

Construcción empleada	Antisentido	HpRNAi	miRNAs artificiales	VIGS
Requerimiento de transgénesis para obtener un silenciamiento sistémico	Si	Si	Si	No
Construcción empleada				
Origen de dsRNAs que luego son procesados por DCL.	Formación de dsRNA por apareamiento del transcripto sentido y del antisentido.	Formación de dsRNA por apareamiento del transcripto sentido y del antisentido.	Explota a los precursores de miRNAs endógenos para generar sRNAs.	Los dsRNAs se forman durante la replicación viral o bien por estructuras secundarias que adquiere el virus.

simple y antigua del uso de silenciamiento génico para el análisis funcional de genes. Si bien estos métodos tienen la desventaja de requerir la producción de plantas transgénicas, la obtención de las mismas permite que el silenciamiento sea estable siendo incluso posible identificar líneas transformantes con diferentes niveles de expresión del gen en estudio. Más aún, el uso de promotores específicos y/o inducibles permiten el estudio de genes cuya disminución total sería letal en etapas tempranas del desarrollo vegetal.

### 2.1.1 Silenciamiento mediante moléculas con estructura de horquilla






Un desarrollo reciente y efectivo para desencadenar silenciamiento génico en plantas es el uso de largos precursores que formen estructuras de horquilla, es decir moléculas de doble cadena de RNA (dsRNA). Para ello se han diseñado y utilizado construcciones genéticas con secuencias repetidas e invertidas del gen a silenciar y que al ser transcritas dan lugar a

largos fragmentos de dsRNA que son eficientes precursores de siRNAs y en consecuencia del silenciamiento génico (Chuang y Meyerowitz, 2000, Wesley y col., 2001). La utilización de esta estrategia requiere un estudio minucioso de la variabilidad de secuencia del gen que se quiere silenciar, por ejemplo si éste forma parte de una familia génica se debe considerar si desea disminuir la expresión de sólo un miembro de la misma o, por el contrario, silenciar a todos los miembros de la familia. Es decir que es muy importante la elección de la secuencia que se va a utilizar en la construcción ya que ésta debe ser muy conservada, de un tamaño no menor a 100 pb y no debe presentar homología con secuencias de genes que no se desean silenciar. Una vez elegida la secuencia, ésta se coloca en sentido y en antisentido de manera que la dsRNA se forme fácilmente (Ver tabla 2). Mediante trabajos sistemáticos en los que se estudian las características de las distintas construcciones capaces de desencadenar el silenciamiento se ha demostrado que aquellas

construcciones que llevan repeticiones invertidas de un fragmento del gen candidato, separadas a su vez por un pequeño intrón, son las más eficiente (ver Tabla 3) (Wesley y col., 2001). Esta estrategia ha sido ampliamente probada en distintas especies de plantas tanto dicotiledóneas como monocotiledóneas y a tal punto que actualmente existen plásmidos genéricos disponibles para la generación de diversas plantas transgénicas (<http://www.pi.csiro.au/rnai/>).

### 2.1.2 Silenciamiento mediante miRNA artificial.

Un miRNA sintético o artificial es un pre-miRNA al cual se le permutaron las 21-22 pb del miRNA maduro y del miRNA\* por otras del gen de interés pero manteniendo las características estructurales lo más parecidas posibles al original. En estas construcciones es muy impor-

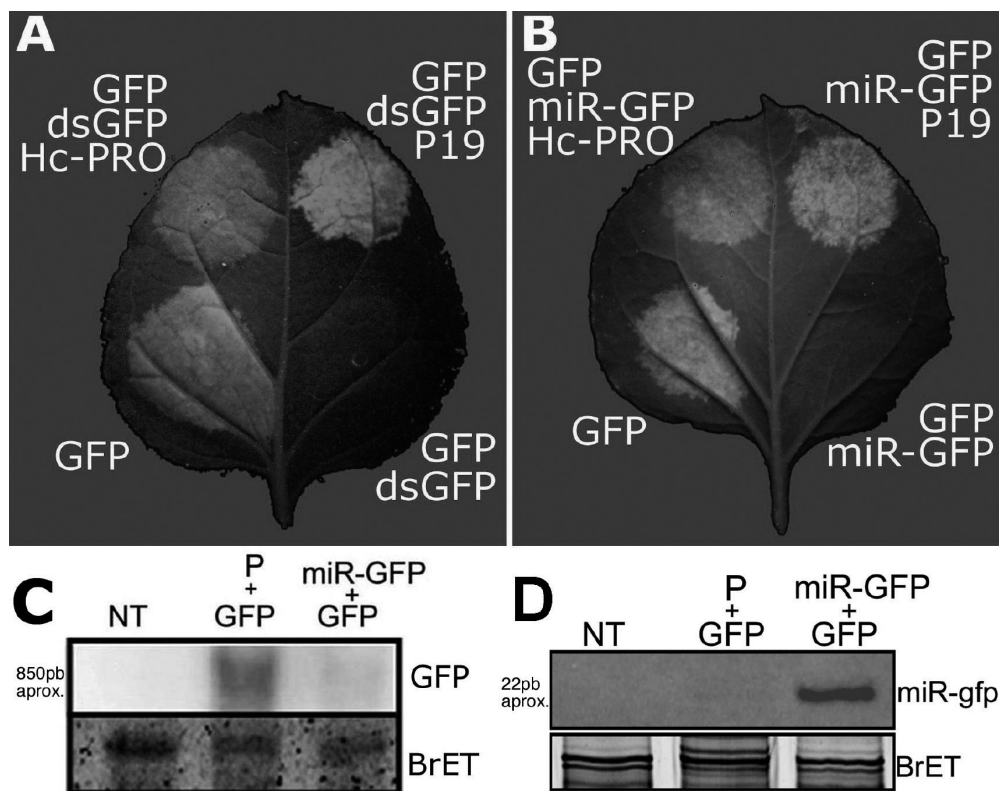
Construcción genética	Estructura del transcripto de RNA	% PTGS	n
Gen candidato en sentido		7	27
Gen candidato en antisentido		4	25
hpRNA (brazos de gen candidato, gen GUS como secuencia espaciadora)		58	43
hpRNA (brazos de gen candidato, intrón en antisentido como secuencia espaciadora)		65	34
hpRNA (brazos de gen candidato, intrón en sentido como secuencia espaciadora)		96	23

**Tabla 3.** Evaluación de la eficacia de las estructuras disparadoras del silenciamiento génico. Adaptado de Wesley y col., 2001

tante realizar un estudio de la energía libre y de la entalpía del pre-miRNA sintético ya que determinarán la estructura y en consecuencia el reconocimiento de la maquinaria del procesamiento. De esta manera el miRNA sintético es capaz de reconocer y controlar al gen blanco que se desea silenciar. Utilizando esta estrategia se han reportado el control de transgenes (Parizotto y col. 2004), de genes endógenos que originalmente no eran blancos de miRNAs (Schwab y col. 2006) e incluso se ha logrado conferir resistencia a virus (Niu y col. 2006) (Ver capítulo sobre obtención de plantas resistentes a virus). En particular nuestro grupo de trabajo ha diseñado un vector de miRNA artificial basado en la estructura del miR164a de *Arabidopsis thaliana* (Figura 3B) logrando controlar la expresión del gen reportero GFP en *Nicotiana benthamiana* (Fig 4).

Si bien esta estrategia requiere de construcciones genéticas más complejas, posee una

gran ventaja frente al silenciamiento clásico ya que minimiza los riesgos de reducción de la expresión de genes no deseados al emplear solamente 21 pb. Asimismo, si se desea silenciar un gen único dentro de una familia de varios genes muy similares la utilización de miRNAs artificiales es más adecuada que la técnica de silenciamiento por horquilla ya que es más fácil hallar 21 pb específicas que grandes porciones (mayores a 100 pb) de secuencias únicas. Para alcanzar mayores niveles de silenciamiento con esta técnica se puede recurrir a tándems de miRNAs artificiales de diferentes porciones de la secuencia blanco del gen a silenciar, aunque esto aumenta la complejidad de las construcciones genéticas necesarias. Para finalizar, la utilización de miRNA artificiales en plantas cultivadas es compatible con su uso a bajas temperatura (menor a 15C), mientras que el PTGS por horquilla podría resultar inhibido a estas temperaturas.



**Figura 4.** (A) y (B) Hoja de *N. benthamiana* agroinfiltrada con distintas combinaciones de GFP o dsGFP, Hc-Pro, P19 y miR-GFP (micro artificial) observada bajo luz UV. (C) "Northern blot" de RNA total extraído de las zonas infiltradas utilizando una sonda que reconoce al mRNA de GFP. (D) Detección del miR-GFP en gel de poliácridamida de material extraído de las zonas infiltradas. P= vector vacío.

## **2.2 Determinación de la actividad de vectores aptos para el silenciamiento de forma transitoria vía agroinfiltración**

Dado que obtener plantas transgénicas es una labor dificultosa y de larga duración, se pueden realizar aproximaciones útiles previas al desarrollo de las mismas. Así, para analizar la eficacia de una horquilla de silenciamiento o el correcto procesamiento de un miRNA artificial se puede realizar un test de silenciamiento local utilizando *Agrobacterium tumefaciens*. Esta técnica consiste en realizar co-agroinfiltraciones generalmente de hojas de *N. benthamiana* con diferentes agrobacterias transformadas; una conteniendo un plásmido expresando el gen blanco y otra conteniendo el plásmido silenciador a evaluar, es decir la misma secuencia blanco (o parte de la misma) repetida e invertida o el miRNA artificial. Una vez transcurrido el tiempo necesario para el establecimiento del silenciamiento se analizan, por ejemplo mediante la técnica de "Northern blot", los indicadores del silenciamiento en la zona agroinfiltrada. Dichos indicadores son en primer lugar una reducción en el nivel de acumulación esperado del gen blanco y en segundo lugar la aparición de sRNAs homólogos al gen blanco. En la Figura 4B se muestra un típico ensayo de inducción de PTGS mediante agroinfiltraciones de hojas de *N. benthamiana* utilizando como gen blanco el reportero GFP y como plásmidos silenciadores una construcción con estructura de horquilla que posee la secuencia de GFP en sentido y antisentido separada por un intrón (dsGFP) o bien una construcción que expresa el miRNA artificial que posee 22 pb de la secuencia de GFP (miR-GFP). Para mayor información acerca de la realización de experimentos de silenciamiento empleando la técnica de agroinfiltración se puede consultar el trabajo de Bazzini y col. (2007). [http://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0717-34582007000200002&lng=es&nrm=iso](http://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0717-34582007000200002&lng=es&nrm=iso). ISSN 0717-3458

## **2.3 Metodología para el análisis de genes candidatos que no requiere transgénesis: silenciamiento génico inducido por virus (VIGS).**

El término VIGS (silenciamiento génico indu-

cido por virus, por su sigla en inglés) se aplica a la técnica que emplea virus recombinantes a fin de disminuir la expresión génica de un gen endógeno. La construcción de vectores virales que portan en su genoma insertos derivados de secuencias génicas vegetales permite silenciar específicamente la secuencia endógena homóloga al inserto introducido en el vector viral. VIGS es una metodología que no requiere de transgénesis por lo tanto es rápida y versátil, siendo ideal para estudios a gran escala. En los últimos años ha ido adquiriendo una gran relevancia debido a la necesidad de analizar la abundante información surgida de los proyectos genómicos. Sin embargo, su mayor desventaja radica en que el nivel de silenciamiento es variable y que depende mucho de la especie a utilizar.

Brevemente, la metodología de VIGS consiste en el clonado del gen candidato en un vector viral, la incorporación de este último en *Agrobacterium* y finalmente en la inoculación de plantas con el cultivo bacteriano para lo cual existen diversos métodos. Uno de los métodos mayormente empleados es la agroinfiltración que consiste en inyectar mediante una jeringa sin aguja la solución de *Agrobacterium* en la cara abaxial de la hojas de una planta. Otro de los métodos es el denominado "agrodrench" que consiste en regar la planta a la altura del tallo con la solución deseada (Ryu y col., 2004). Tras la agroinoculación de las plantas los vectores virales son incorporados a las células vegetales donde forman nuevos genomas virales causando la infección sistémica de la planta. En cada una de las células infectadas se produce la replicación viral induciendo el silenciamiento del gen candidato (Ruiz y col., 1998; Liu y col., 2002). Existen varios vectores virales desarrollados para VIGS (que se resumen en la tabla 4), sin embargo el sistema de VIGS basado en el virus TRV es una de los más utilizados en la actualidad ya que presenta ciertas ventajas frente a otros vectores virales. Este virus ocasiona un nivel menor de síntomas en las plantas, puede infectar meristemas y tiene un amplio rango de hospedantes. TRV es un virus de RNA de polaridad positiva con un genoma bipartito. Uno de esos RNAs genómicas fue modificado de modo de eliminar

**Tabla 4:** Vectores virales para su uso en VIGS .Tabla adaptada de Godge y col., 2008.

Virus	Género	Hospedantes naturales importantes	Probado en	Genes silenciados	Referencias
Tobacco Mosaic Virus	Tobamovirus	Nicotiana tabacum	Nicotiana benthamiana, Nicotiana tabacum	PDS, PSY	Kumagai et al (1995)
Tobacco Rattle virus	Tobravirus	Amplio rango de hospedantes	N. benthamiana, Arabidopsis, Tomate, especies de Solanum, pimiento	Varios genes PDS, PSY, GFP, EDS, RAR, NPR1	Liu et al (2002b) Ratcliff et al (2001)
Potato Virus X	Potexvirus	Solanum tuberosum Brassicca Campestris	N. benthamiana, Solanum tuberosum	PDS, GFP	Ruiz et al (1998)
Barley stripe mosaic virus	Hordeivirus	Trigo, cebada, maíz	Trigo, cebada	PDS	Holzberg et al (2002) Scofield et la ( 2005)
Tomato goleen mosaic virus	Begomovirus	Lycopersicum sculetum	N. benthamiana	Su(Sulfur gene)	Peele et al (2001)

las proteínas no esenciales e incorporar un sitio múltiple de clonado en donde se puede introducir el fragmento del gen candidato que se desea silenciar, dejando las regiones no traducibles y la región codificante de la proteína de la cápside intactas. Posteriormente este vector fue mejorado por el grupo del Dr. Dinesh Kumar mediante la incorporación del promotor 35S duplicado el cual incrementa la transcripción, del terminador de la nopalina sintetasa y de la secuencia de una ribozima que permite simular el final 3' del genoma viral de forma más adecuada (Liu y col., 2002).

Así, la base del silenciamiento inducido por virus puede explicarse fundamentalmente mediante el mecanismo de PTGS. Estudios realizados en *N. benthamiana* empleando un transgén de la proteína reportera GFP muestran que el mecanismo de VIGS puede ser separado en

dos etapas, una de iniciación y otra de mantenimiento. La primera sería absolutamente dependiente de la presencia del virus, siendo los genes blancos silenciados sólo si la planta es infectada por el virus. Mientras que la segunda etapa sería independiente del mismo. (Ruiz y col., 1998, Jones y col., 1999).

### Consideraciones finales

A lo largo de este capítulo se describieron algunas de las metodologías empleadas frecuentemente para estudios de genética reversa basados en el silenciamiento génico. Como se mencionó al inicio del capítulo el empleo de estas metodologías presenta ciertas ventajas respecto a otras, sin embargo una de las consideraciones que hay que tener en cuenta al usar los métodos basados en silenciamiento génico es la especificidad del mismo con el fin de

evitar que se altere la expresión de otros genes no deseados. Estudios realizados en *Arabidopsis thaliana* indican que al emplear como inserto la región completa de un gen se alteran aproximadamente la expresión de 4 genes no deseados (Ping Xu y col., 2006). En el último tiempo han surgido varios trabajos en donde se proponen distintos modos de controlar y/o disminuir estos efectos no deseados, entre ellos se encuentran el uso de programas bioinformáticos que permiten analizar cuál es la mejor región a emplear de modo de obtener un silenciamiento lo más específico posible (Qiu y col., 2006 y Ping Xu y col., 2006). Sin embargo estas herramientas sólo son útiles cuando el genoma del organismo ha sido completamente secuenciado o bien si al menos se dispone de gran parte de su secuencia. Por otro lado también se ha desarrollado una estrategia experimental para validar la información obtenida la cual emplea genes sintéticos que codifican para el gen candidato, pero que no resultan blanco de la acción de siRNAs, de este modo si el gen silenciado es el responsable del fenotipo observado, la expresión de los mismos permitiría la recuperación el fenotipo salvaje (Kumar y col., 2006).

Brevemente, no hay una técnica mejor que otra para reducir la expresión génica, sino que cada una posee sus ventajas y desventajas y queda a merced del investigador elegir la/las más correcta/s para su sistema, considerando principalmente los tiempos, el costo, la especie a utilizar, las condiciones del ambiente y el grado de reducción del silenciamiento con el que uno desea trabajar.

Un mensaje final para involucrarse y leer más sobre este tema, es que si bien el silenciamiento es un mecanismo de regulación negativa de la expresión génicas, se ha tornado uno de los temas más importantes de las últimas décadas en biología molecular, a tal punto que los descubrimientos iniciales de este mecanismo les permitieron a los científicos Dr Andrew Z. Fire y Dr Craig C. Mello obtener el Premio Nobel de Fisiología o Medicina en el año 2006. Actualmente se están desarrollando numerosos estudios en los cuales mediante silenciamiento génico se logra obtener plantas inmunes a infecciones virales o bacterianas, disminuir la

alergenicidad de vegetales comestibles y estudiar la función de los genes.

### Bibliografía:

- Amedeo, P., Habu Y, Afsar K, Mittelsten Scheid O, Paszkowski J. 2000. Disruption Of The Plant Gene *Mom* Releases Transcriptional Silencing Of Methylated Genes. *Nature*, 405 (6783): P. 203-6.
- Bartel D.P. 2004. MicroRNAs: Genomics, biogenesis, mechanism and function. *Cell*: 116:281-297.
- Baulcombe, D.C. 2004. RNA Silencing In Plants. *Nature*, 431 (7006): P. 356-63.
- Bazzini, A.A., Mongelli, V.C., Hopp, H.E., Del Vas, M. And Asurmendi, S. 2007. A Practical Approach To The Understanding And Teaching Of Rna Silencing In Plants. *Electronic Journal Of Biotechnology*, 10 (2): 178-190. eISSN: 0717-3458.
- Chapman E. J., Carrington J. C. 2007. Specialization And Evolution Of Endogenous Small Rna Pathways *Nat Rev Genet*. Nov;8 (11):884-96
- Chuang, C.F. And Meyerowitz, E. M. 2000. Specific and heritable genetic interference by double-stranded RNA in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* 97, 4985-4990.
- Doench J.G., Petersen, C.P., And Sharp, P.A. 2003. siRNAs can function as miRNAs. *Genes & Dev*.17: 438-442.
- Dorokhov Iu, L. 2007. Gene Silencing In Plants. *Mol Biol (Mosk)*, 41(4): P. 579-92.
- Dunoyer P., Lecellier C.H., Parizotto E.A., Himber C., Voinnet O. 2004. Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. *Plant Cell* 16:1235-1250.
- Dunoyer, P. And O. Voinnet. 2008. Mixing And Matching: The Essence Of Plant Systemic Silencing? *Trends Genet*, 24(4): P. 151-4.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. And Mello, C. C. 1998. "Potent And Specific Genetic Interference By Double-Stranded Rna In *Caenorhabditis Elegans*." *Nature*, Vol. 391, No. 6669, P. 806-11.
- Furner, I.J., M.A. Sheikh, And C.E. Collett. 1998. Gene Silencing And Homology-Dependent Gene Silencing In *Arabidopsis*: Genetic Modifiers And Dna Methylation. *Genetics*, 149(2): P. 651-62.
- Godge MR, Purkayastha A, Dasgupta I, Kumar PP. 2008. Virus-induced gene silencing for functional analysis of selected genes. *Plant Cell Rep*. 27(2):209-19.
- Jones, L. Hamilton A.J, Voinet O, Thomas C.L, Maule, A.J Y Baulcombe D.C 1999. Rna-Dna Interactions And Dna Methylation In Post-Tran-

- scriptional Gene Silencing. *Plant Cell*, 11(12): P. 2291-301.
- Jones-Rhoades, M.W., Bartel, D.P., Bartel, B. 2006. MicroRNAs and their regulatory roles in plants. *Annu. Rev. Plant Biol.* 57:19–53.
- Hutvagner G. And Zamore, P.D. 2002. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297: 2056–2060.
- Kumar, D., C. Gustafsson, D. F. Klessig. 2006. Validation of RNAi Silencing Specificity Using Synthetic Genes: Salicylic Acid-binding Protein 2 is required for Innate Immunity in Plants. *The Plant Journal* 45(5): 863-868
- Liu, Y., M. Schiff, Y S.P. Dinesh-Kumar. 2002. Virus-Induced Gene Silencing In Tomato. *Plant J*, 31(6): P. 777-86.
- Lu R., Martin- Hernandez A.M, Peart J.R., Malcuit I. Y Baulcombe D.C. 2003. Virus-Induced Gene Silencing In Plants. *Methods*, 30(4): P. 296-303
- Niu, Q. W., Lin, S. S., Reyes, J. L., Chen, K. C., Wu, H. W., Yeh, S. D. And Chua, N. H. 2006. "Expression Of Artificial Micromnas In Transgenic Arabidopsis thaliana Confers Virus Resistance." *Nat Biotechnol*, Nov, Vol. 24, No. 11, P. 1420-8.
- Ossowski, S., R. Schwab, And D. Weigel. 2008. Gene Silencing In Plants Using Artificial MicroRNAs And Other Small Rnas. *Plant J*, 53(4): P. 674-90.
- Parizotto, E. A., Dunoyer, P., Rahm, N., Himber, C. And Voinnet, O. 2004. In Vivo Investigation Of The Transcription, Processing, Endonucleolytic Activity, And Functional Relevance Of The Spatial Distribution Of A Plant Mirna. *Genes Dev*, Sep 15, Vol. 18, No. 18, P. 2237-42.
- Ping Xu, Yuanji Zhang, Li Kang, Marilyn J. Roosinck, And Kirankumar S. Mysore. 2006. Computational Estimation And Experimental Verification Of Off-Target Silencing During Posttranscriptional Gene Silencing In Plants. *Plant Physiol.* Oct;142(2):429-40.
- Peng Qiu, Z. Jane Wang , K. J. Ray Liu , Zhang-Zhi Hu And Cathy H. Wu. 2006. Dependence network modeling for biomarker identification. *Bioinformatics* 23(2):198-206
- Ruiz, M.T., O. Voinnet, And D.C. Baulcombe. 1998. Initiation And Maintenance Of Virus-Induced Gene Silencing. *Plant Cell*, 10(6): P. 937-46.
- Ryu, C.M., Anad A., Kang L., Mysore K.S. 2004. Agrodrench: A Novel And Effective Agroinoculation Method For Virus-Induced Gene Silencing In Roots And Diverse Solanaceous Species. *Plant J*, 40(2): P. 322-31.
- Travella S., Klimm T. E., And Keller, B. 2006. RNA Interference-Based Gene Silencing As An Efficient Tool For Functional Genomics In Hexaploid Bread Wheat. *Plant Physiology* 142:6-20 (2006)
- Vance V., Vaucheret H. 2001. RNA silencing in plants: Defense and counterdefense. *Science* 292:2277–2280.
- Vaucheret, H. And M. Fagard. 2001. Transcriptional Gene Silencing In Plants: Targets, Inducers And Regulators. *Trends Genet.* Jan;17(1):29-35
- Voinnet, O. 2001. RNA silencing as a plant immune system against viruses. *Trends Genet* 17, 449–459.
- Wesley, S. V., Helliwell, C. A., Smith, N. A., Wang, M. B., Rouse, D. T., Liu, Q., Gooding, P. S., Singh, S. P., Abbott, D., Stoutjesdijk, P. A., Robinson, S. P., Gleave, A. P., Green, A. G. And Waterhouse, P. M. 2001. Construct Design For Efficient, Effective And High - Throughput Gene Silencing In Plants." *Plant J*, Vol. 27, No. 6, P. 581-90.
- Zeng, Y., Yi, R., And Cullen, B.R. 2003. MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms. *Proc. Natl. Acad. Sci.* 100: 9779–9784.



## II. CAPÍTULO 8

### Análisis de Experimentos Biológicos

Sofía Olmos; Miguel Di Renzo;  
Mercedes Ibañez; Nélica Winzer

#### 1 La variación biológica

La **variabilidad genética** es un hecho universal en las poblaciones reproductivas y una condición preliminar necesaria para el cambio evolutivo y para la respuesta al mejoramiento genético mediante selección o hibridación. Por otra parte son conocidos los peligros que implica la uniformidad y la falta de diversidad en las especies cultivadas y en los animales domésticos, como consecuencia de la erosión genética.

La variación que presentan los individuos de una población puede ser discontinua o continua. Los **caracteres cualitativos** de variación discontinua, como los marcadores moleculares, permiten clasificar a los individuos sin ambigüedades, ya que estos caracteres están regidos por uno o por pocos genes y el ambiente no tiene efecto en su manifestación fenotípica. En cambio, los **caracteres cuantitativos**, representados por la mayoría de los caracteres de interés agronómico, presentan variación continua y los fenotipos, que están determinados por poligenes y por efectos ambientales, son difíciles de clasificar en categorías discretas. Para su análisis deben emplearse métodos biométricos, que no miden el efecto de genes individuales o de segmentos cromosómicos específicos, sino de todo el genoma.

Durante los últimos años ha surgido un consenso entre los mejoradores y genetistas moleculares sobre la conveniencia del uso de marcadores moleculares para asistir a la selección de caracteres cuantitativos (MAS, Marker Assisted Selection). La biotecnología ha colaborado mediante la construcción de mapas genéticos saturados, esto es, con innumerables marcadores moleculares a lo largo de los cromosomas, que permiten ubicar regiones de actividad cuantitativa (QTLs, Quantitative Trait Loci). De esta manera un carácter de variación

continua puede ser manejado como si se tratara de un carácter de herencia mendeliana mediante la introgresión de un fragmento de cromosoma conteniendo un QTL por retrocruzamientos sucesivos en las variedades comerciales. Los QTLs representan un nuevo sistema de enfocar el estudio básico y el manejo práctico de los sistemas poligénicos y de la genética cuantitativa.

#### 2 Experimento científico

La ciencia, que tiene como objetivo la explicación y la predicción de los hechos, cuenta con el **método científico** como un procedimiento que contempla **un proceso en flujo desde los hechos observados hasta la proposición de hipótesis explicatorias**. Un requisito fundamental en toda ciencia fáctica es la **contrastación de las hipótesis planteadas**, poniendo a prueba las mismas mediante una confrontación con la experiencia. El **diseño experimental** crea artificialmente las condiciones para la contrastación de la hipótesis y brinda la metodología estadística correspondiente para el análisis de los datos.

El **experimento científico**, que como queda dicho es un ensayo destinado a probar la validez de una hipótesis, se basa en la reproducción de ciertos fenómenos bajo condiciones controladas con el objeto de medir el efecto que produce la modificación de uno o varios factores en estudio (tratamientos) sobre la variable dependiente.

El **ensayo**, cuidadosamente diseñado, deberá ser lo más simple posible, evitando los errores tendenciosos y sistemáticos (bias=sesgo) de manera tal que los datos recabados provean conclusiones con un amplio rango de validez. Por otra parte, para magnificar el efecto de los tratamientos y disminuir el error experimental, se procura minimizar el efecto de otros factores no controlados que ejercen influencia en las observaciones. Esto se logra utilizando materiales y condiciones experimentales lo más homogéneas posibles. De esta manera, las variaciones observadas se deberán casi exclusivamente a los tratamientos.

Para conocer el efecto de los **factores controlados** y el de los **factores no controlados** (error experimental) sobre el material experi-

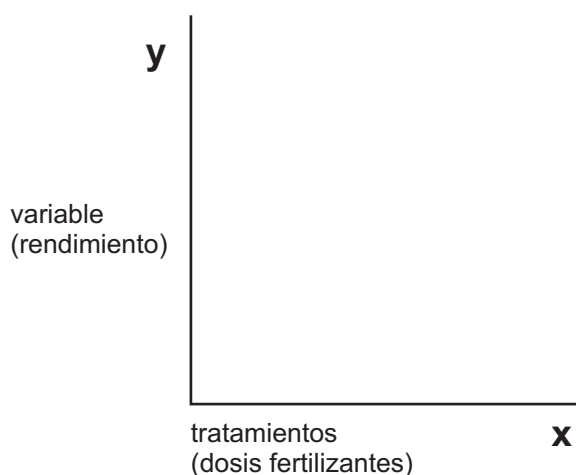
mental es necesario formular claramente una hipótesis, utilizar un diseño experimental apropiado y analizar e interpretar los resultados obtenidos.

Por ejemplo, la aplicación de distintos tratamientos (ej.: distintas dosis de fertilizante = causa), producen variaciones sobre una característica de interés (ej. rendimiento = efecto):

#### relación causal

**causa:**           ⇒       **efecto:**  
tratamientos                    cambios en la variable de interés

Esto puede graficarse de esta manera:



$$Y = f(x)$$

rendimiento = f (dosis fertilizante)

#### Terminología

- **Factor:** Es la causa cuyo efecto se quiere medir (ej. fertilizante, insecticida, medio de cultivo, genotipo, ambiente)

- **Nivel:** Es cada una de las categorías o alternativas del factor en estudio (ej. dosis de un fertilizante, dosis de insecticida, sales y pH de los medios de cultivo, variedades, diferentes años y localidades).

- **Tratamiento:** Es cada nivel del factor en estudio. Si se estudia el factor fertilizante, cada dosis (o nivel) es un tratamiento distinto. Si se prueba el efecto de concentraciones crecientes de un fertilizante los tratamientos tendrían un orden natural (*discreto ordinal*); pero si se prueban distintas combinaciones de fertilizantes

(NPK), no habría un orden entre los tratamientos (*discreto nominal*). Cuando se estudian dos (o más) factores simultáneamente, los distintos tratamientos resultan de la combinación de cada nivel de un factor con cada nivel del otro factor. Por ejemplo, si se estudia el efecto de fertilizante y competencia, cada tratamiento resulta de la combinación de una dosis de fertilizante y una densidad de siembra determinada.

- **Variable independiente:** es la causa que afecta los resultados del experimento y está representada por los distintos niveles del factor cuyo efecto se quiere medir (tratamientos) o que se quiere controlar (bloques).

- **Variable dependiente:** variable respuesta o simplemente «variable» es el carácter de interés (datos u observaciones) cuya variación permite cuantificar el efecto de la variable independiente. Ejemplos de variables dependientes: a) rendimiento (kg/ha) y diámetro de colonia (mm): cuantitativas continuas, b) tiempo a panojamiento (días), tamaño de camada, número de colonias (número): cuantitativas discretas, c) marcadores moleculares (presencia/ausencia): cualitativa binaria, d) grado de respuesta a un estrés o a un patógeno (tolerante, intermedio, susceptible): cualitativa ordinal.

- **Unidad experimental:** es un individuo o grupo de individuos, objetos, porción de material o terreno (parcela) al que se le asigna un tratamiento experimental y sobre la cual se observa una respuesta.

- **Repetición:** es el número de unidades experimentales a las que se les asigna el mismo tratamiento.

### 3 Hipótesis Estadística

Como se indicara previamente, el experimento científico es un ensayo destinado a probar la validez de una hipótesis. Las hipótesis estadísticas plantean supuestos referentes a algún parámetro poblacional. Por lo general se utiliza la **hipótesis nula** ( $H_0$ ), que se refiere a la igualdad entre los parámetros probados. Frente a la  $H_0$  se plantea una **hipótesis alternativa** ( $H_a$ ).

Para determinar si un tratamiento tuvo algún efecto, se considera a la  $H_0$  como la ausencia de efectos, lo que equivale probar la igualdad entre las medias de los tratamientos. La

Ha plantea que al menos una de las medias difiere significativamente del resto o sea que al menos uno de los niveles del factor probado tiene efecto sobre la variable de interés. Las hipótesis estadísticas también prueban otras características poblacionales, como por ejemplo la igualdad entre varianzas o entre coeficientes de regresión.

El ensayo y las pruebas de significancia permiten calcular probabilidades bajo el supuesto de que la hipótesis nula sea cierta. Si la probabilidad calculada es igual o mayor que un nivel de significancia preestablecido (de 0,05 por ejemplo) la  $H_0$  no se rechaza y se infiere que no hay diferencias significativas entre las medias de los tratamientos. Se concluye que los tratamientos no tienen efecto significativo en la variabilidad del carácter estudiado y las diferencias observadas no son mayores que las que se producirían por azar. En cambio si la probabilidad calculada es pequeña ( $P \leq 0,05$ ) se rechaza la  $H_0$  y se infiere, de acuerdo con la  $H_a$ , que hay diferencias significativas entre los tratamientos. Si  $P \leq 0,01$  se dice que las diferencias son altamente significativas.

	Ho verdadera	Ho falsa
Rechazo $H_0$	error tipo I	correcto
No rechazo $H_0$	correcto	error tipo II

El *nivel de significancia*, generalmente designado por  $\alpha$ , representa hasta qué punto, en términos de probabilidad, estamos dispuestos a aceptar un error de tipo I, es decir, rechazar una  $H_0$  que sea verdadera. Esto conlleva a que aceptemos una  $H_0$  como verdadera con una probabilidad  $1-\alpha$ . Si se fija  $\alpha=0.05$  estamos dispuestos a aceptar el error tipo I aproximadamente una de cada 20 veces.

#### 4 Análisis univariado

Los análisis univariados emplean una variable dependiente, mientras que los análisis multivariados comparan muestras sobre la base de varias variables que están relacionadas.

La significancia estadística de una diferencia entre dos medias puede verificarse mediante una prueba t o mediante una prueba F. Ambas

son equivalentes,  $t^2 = F$ . La prueba F, que es un cociente de varianzas, es una generalización de la prueba t, especialmente útil cuando se comparan más de dos tratamientos. En el análisis de la varianza (ANOVA), por ejemplo, se utiliza la prueba F para verificar la igualdad de medias.

#### Prueba F

Una prueba F consiste en un cociente entre dos varianzas:

$$F = s^2_{\text{entre}} / s^2_{\text{dentro}}$$

Para realizar la prueba F resulta necesario descomponer la varianza total ( $s^2_{\text{total}}$ ) en dos componentes:  $s^2$  entre y  $s^2$  dentro.

El análisis de la varianza es un procedimiento aritmético que permite la descomposición de la varianza total en varianzas parciales. El cociente entre las varianzas permite obtener un valor de F calculado ( $F_c$ ) el cual se compara con un F de tabla ( $F_t$ ) para un determinado nivel de significación (por ejemplo  $\alpha=0.05$ ) y los grados de libertad de los tratamientos y del error. Los programas estadísticos estiman directamente el P, la probabilidad de obtener un valor de F mayor que  $F_c$  si  $H_0$  fuera cierta.

Los supuestos básicos en que se fundamenta el ANOVA son errores independientes, normalmente distribuidos, aditividad de efectos y varianzas homogéneas para todos los tratamientos. Cuando no se cumplen se ve afectado el nivel de significación, la sensibilidad de la prueba de F y la discrepancia real de la  $H_0$ .

#### Varianza total

Se puede considerar a la variación total como la resultante de dos fuentes de variación: una fuente de variación debida a las diferencias entre grupos o tratamientos, señaladas por los promedios de éstos y otra debida a la diferencia dentro de los tratamientos (entre repeticiones), que se calcula en base a las diferencias entre las observaciones o repeticiones de cada tratamiento.

#### Varianza entre tratamientos

Para aislar la variación entre los grupos ne-

cesitamos suprimir la variación dentro de los tratamientos (entre repeticiones). Podemos obtener este resultado haciendo a todos los individuos de un mismo grupo iguales entre sí e iguales a la media del grupo. Mediante esta operación, la variación entre los grupos no se modifica mientras que toda la variación dentro del grupo queda anulada.

### **Varianza dentro de tratamientos**

Para obtener la varianza dentro de los grupos debemos tener en cuenta solamente la variación entre las observaciones de un mismo tratamiento (entre repeticiones). La variación dentro de un tratamiento se debe a las desviaciones que presentan los valores individuales en ese grupo en relación con la media de ese grupo.

### **5 Análisis multivariado**

El análisis multivariado brinda los métodos estadísticos para el análisis conjunto de varias variables que pueden estar interrelacionadas. Esta rama de la estadística comprende procedimientos y técnicas para la síntesis, la presentación y el análisis multidimensional de variables, tanto cualitativas como cuantitativas, obtenidas a partir de un número de individuos, OTU (**Operational Taxonomic Unit**), objetos o tratamientos.

Debido a que las variables se consideran en forma simultánea estas técnicas permiten realizar interpretaciones más complejas a las que surgen mediante la utilización de métodos univariados. El análisis multivariado colabora en la comprensión más adecuada y completa de las ciencias biológicas, por tal motivo en este capítulo se brindará un panorama simplificado y algunos comentarios sobre las técnicas más usuales con la finalidad de poder discernir y elegir la metodología más apropiada para analizar la información que se disponga.

Hay algunas razones que justifican el análisis conjunto de diferentes variables en lugar de analizar cada variable separadamente por métodos univariados. Por ejemplo, cuando se trata de probar hipótesis mediante procedimientos de inferencia estadística. En este caso, la aplicación del análisis univariado a cada una de las variables aumenta el error Tipo I mien-

tras que la aplicación del análisis multivariado preserva el valor de  $\alpha$  y aumenta en muchos casos la potencia del test. El análisis multivariado utiliza las relaciones (correlaciones) entre las variables o entre los objetos que el método univariado directamente no considera. El análisis multivariado aprovecha estas relaciones para buscar en los datos patrones o estructuras enriqueciendo así la descripción de los mismos. En el capítulo siguiente se desarrollarán en más detalle algunas metodologías multivariadas, seleccionándose aquellas que se utilizan con mucha frecuencia y que no requieren demasiados conocimientos previos.

### **Análisis exploratorio y confirmatorio**

Con los datos provenientes de observaciones realizadas sobre un hecho en particular es necesario, en primer lugar, realizar un **análisis exploratorio**. Posteriormente, mediante las inferencias realizadas a partir de este análisis y con la información que se pueda disponer acerca de los mecanismos que originan el proceso biológico, es posible desarrollar un **análisis confirmatorio** y así plantear un modelo descriptivo de causalidad el que, a su vez, es confirmado mediante una prueba de bondad de ajuste.

En otras palabras, toda investigación llevada a cabo especialmente en áreas nuevas, donde existen pocos antecedentes, comienza con una exploración de los datos con el objeto de reconocer cualquier estructura o patrón no aleatorio que requiera una explicación. En una segunda etapa, es posible que surja la necesidad de formular **la pregunta**, lo cual es a menudo más interesante que buscar **la respuesta** subsiguiente, ya que el objetivo primordial es, en primer lugar, generar hipótesis interesantes que promuevan estudios más avanzados. En este caso, los métodos que están diseñados para responder a preguntas específicas no serían necesarios. En cambio resultarían apropiados aquellos métodos que puedan detectar un patrón de comportamiento no previsto en los datos y que, además, abran un amplio campo de posibles explicaciones del proceso. Estas técnicas generalmente se caracterizan por el énfasis puesto en la utilización de representaciones visuales y gráficas y por la carencia de modelos

estocásticos, donde la significación estadística de los resultados pierde importancia.

La aplicación de los análisis confirmatorios se vuelve más apropiada cuando el investigador tiene en mente hipótesis bien definidas. Es aquí donde las pruebas de significación estadística son necesarias. Sin embargo, no existe una división muy estricta entre técnicas exploratorias y confirmatorias. Por ello sería importante que el investigador tenga una perspectiva flexible y pragmática para analizar sus datos y una experiencia suficiente que le permita elegir la herramienta analítica adecuada. El error más grave que se podría realizar en este sentido sería arribar a conclusiones que expliquen causas a partir de datos exploratorios y descriptivos sin haber realizado ningún tipo de diseño experimental para probarlos. Se plantea como causa probable de esta tendencia generalizada el mal uso semántico entre las terminologías estadísticas y las biológicas. El uso estadístico de términos como “**efectos**” o “**variable independiente**” no implica causalidad.

### **Síntesis de métodos: objetivos y limitaciones**

Los análisis multivariados más comúnmente adoptados, en orden decreciente, son: análisis de componentes principales (ACP), análisis de funciones discriminantes (AD), análisis de agrupamientos (conglomerados o clusters), regresión múltiple (RM), análisis multivariado de la varianza (MANOVA), análisis de correspondencia (AC), análisis de coordenadas principales (ACOO), análisis factorial (AF), correlación canónica (CC), modelos logarítmicos lineales (LOGL), escalamiento multidimensional (EMD) y la regresión logística múltiple (RL).

En la **Tabla 1** se mencionan los objetivos generales de los métodos de análisis multivariado más utilizados. Los procedimientos numerados del 1 al 7 utilizan **combinaciones lineales** de las variables, estos métodos son más eficientes con variables continuas, mientras que aquellos métodos numerados del 8 al 11, comprenden **métodos no lineales** y son más apropiados cuando se cuenta con variables cualitativas y el método 12 es adecuado para variables cuantitativas, cualitativas o la mezcla de ambas.

Los **métodos lineales** son apropiados cuando se quiere interpretar combinaciones óptimas de variables (por ejemplo, componentes principales en el análisis de componentes principales, factores en el análisis factorial, funciones discriminantes en el análisis discriminante). Quienes aplican métodos lineales usualmente asumen que los valores de las variables incrementan o decrecen regularmente y que entre ellas no hay interacciones. Estas relaciones no lineales en el análisis multivariado de la varianza aparecen en los términos de las interacciones y pueden revelar efectos no lineales importantes en el análisis de datos experimentales.

Para examinar con mayor eficiencia datos binarios (presencia/ausencia) y datos en rangos, se pueden usar otros modelos, donde las variables se relacionan mediante **funciones no lineales**. Cuando se utilizan combinaciones de variables hay que considerar que el coeficiente que afecta a una variable representa la contribución de esta variable a la combinación lineal y que este valor depende de las otras variables incluidas en el análisis. No se considera correcto el empleo de estos coeficientes individuales para la interpretación de dichas contribuciones.

Antes de aplicar cualquier técnica de análisis resulta imprescindible realizar un examen inicial de los datos. El examen inicial, básicamente, consiste en la evaluación de datos faltantes y la representación gráfica de los datos, lo que permite la identificación de «outliers» (datos fuera del rango o fuera de tipo), tendencias y/o agrupamientos preliminares e hipotetizar posibles modelos para su análisis. Además, cuando el análisis multivariado a emplear así lo requiera, se debe verificar si se cumplen los supuestos básicos de normalidad multivariada, homogeneidad en la estructura de variación y covariación, independencia de errores y linealidad.

Por otro lado, y al igual que en el análisis univariado, hay que tener cuidado con el uso del nivel de significancia ( $\alpha$ ) que se adopte, de las subsiguientes inferencias estadísticas y de las conclusiones a que se arriba luego del análisis. Las conclusiones confirmatorias solamente se justifican si los estudios estuvieron basados en un muestro apropiado. Esto significa que la inferencia está justificada en relación a la forma

**Tabla 1.** Objetivos y limitaciones de los 12 tipos de análisis multivariados más comúnmente utilizados.

Análisis	Objetivos	Limitaciones
1. Regresión múltiple (RM)	Estudiar la relación funcional entre una variable dependiente (Y) y varias variables independientes (X1, X2 ..., Xn). Predecir una variable dependiente a partir de variables independientes. Si los modelos están basados en experimentación, investigar causalidad y efectos.	Una buena predicción no habilita por sí sola a hacer inferencias sobre causalidad. La predicción debe ser realizada sólo en situaciones similares a aquellas de donde el modelo fue derivado (dentro del dominio de los valores experimentados). El procedimiento sólo considera funciones lineales de las variables analizadas. La RM es apropiada para variables Y continuas e independientes; en el caso que se busque realizar inferencias estadísticas, los errores deben ser normales, las varianzas homogéneas el muestreo debe ser aleatorio.
2. Análisis multivariado de varianza (MANOVA)	Explorar la relación entre distintas variables independientes cualitativas (tratamientos) y varias variables dependientes cuantitativas. Es un método básicamente inferencial.	El MANOVA debe cumplir con los supuestos de independencia de las observaciones, homogeneidad de matrices de varianza y covarianza para todos los grupos de tratamientos y distribución normal multivariada para todas las variables dependientes. En algunos casos con un gran número de variables dependientes el poder estadístico del ANOVA excede al obtenido con un simple MANOVA.
3. Análisis discriminante (AD)	Describir las situaciones multigrupales, encontrar las combinaciones lineales óptimas entre variables que permitan discriminar grupos de objetos. La ecuación de la función discriminante lineal puede ser usada para clasificar observaciones o para predecir a que grupo pertenece una nueva observación.	El procedimiento está indicado para variables independientes continuas, siendo ineficiente para variables que no están corregidos por varianzas o covarianzas. Con AD sólo se encontrarán las combinaciones de variables que sean lineales. Los grupos deben ser definidos a priori. El procedimiento es propuesto principalmente para variables continuas, es ineficiente para datos que no están corregidos por varianzas o covarianzas. Con ACP sólo se encontrarán las combinaciones de variables que sean lineales.
4. Análisis de componentes principales (ACP)	Describir una matriz de datos de objetos y variables en una dimensión más reducida, usualmente mediante una representación gráfica (biplot); fijar combinaciones lineales no correlacionadas de las variables originales maximizando sus varianzas. Sugerir nuevas combinaciones de variables para estudios posteriores. Describir los datos mediante la reducción de las dimensiones entre objetos y mostrarlas mediante una representación gráfica. Se puede hacer con variables cuantitativas, cualitativas o con la mezcla de ambas.	Los resultados dependen del índice de distancia elegido. El análisis produce un nuevo sistema coordenado que, a diferencia de otros métodos, no puede indicar combinaciones de variables porque sólo emplea la matriz de distancia entre objetos. Los métodos exploratorios de análisis de factores son tan estructurados que las interpretaciones son subjetivas. El procedimiento es ineficiente para datos que no están corregidos mediante correlaciones o combinaciones lineales, por lo que no resulta ideal para relaciones no lineales o datos binarios. El análisis es ineficiente para datos que no están corregidos mediante correlaciones o combinaciones lineales, por lo que no resulta ideal para relaciones no lineales o datos binarios. No se pueden hacer rotaciones como en el caso del AF lo cual dificulta la interpretación de las combinaciones lineales encontradas. Brinda una estimación de la varianza entre combinaciones lineales y no de la varianza extraída de las variables.
5. Análisis de coordenadas principales (ACOOP)		
6. Análisis Factorial (AF)	Reproducir una matriz de correlación entre variables originales suponiendo la existencia de uno o más factores no evidentes (latentes). Descubrir estructuras subyacentes en grupos de datos mediante la interpretación de estos factores.	
7. Análisis de correlaciones canónicas (CC)	Analizar la correlación existente entre dos grupos de variables del mismo grupo de objetos. Esto se hace en forma simultánea en lugar de calcular correlaciones de a pares.	

8. Regresión logística (RL)	Modelar la relación entre una variable respuesta binaria (Y) y una o más variables X cuantitativas y/o categóricas. La RL también sirve para investigar la asociación entre una variable X con otra variable en la presencia de otras variables X. Si los modelos de causalidad están basados en experimentación se podría estudiar causas y efectos. La RL brinda una alternativa para el AD entre dos grupos cuando las variables son categóricas. Investigar las relaciones conjuntas entre variables categóricas.	Una buena predicción por sí sola no permite hacer inferencias de causalidad. Las predicciones sólo deberían ser llevadas a cabo en situaciones similares a aquellas utilizadas por el modelo propuesto (dentro del dominio de los valores experimentados).
9. Modelos logarítmicos lineales LOGL.		Las variables deben ser categóricas o bien se requiere una transformación a categóricas.
10. Análisis de correspondencia AC.	Explorar gráficamente las relaciones contenidas en las tablas de contingencia. El AC permite sugerir nuevas combinaciones de variables para estudios posteriores. Describir los datos mediante la reducción de las dimensiones a través de una representación gráfica, con el objetivo de encontrar relaciones no lineales entre objetos.	Es una técnica descriptiva aconsejable para análisis exploratorios. Es muy sensible a la presencia de "outliers".
11. Escalamiento multidimensional EMD.		El análisis solamente utiliza información que está ordenada en rangos.
12. Análisis de agrupamientos o conglomerados	Clasificar objetos o variables por su semejanza o diferencias en base a mediciones de similitud o distancia. Los agrupamientos se representan en un dendrograma.	Si bien este método puede ser usado para las variables cuantitativas y/o cualitativas, los resultados dependen de las medidas de distancias y de los algoritmos elegidos para formar los agrupamientos.

en que los datos fueron tomados y en que el experimento fue realizado, más que en la técnica de análisis en sí misma.

Las inferencias sólo están justificadas cuando los datos son tomados de una muestra grande y bien definida. Cuando esto no ocurre los valores de probabilidad tal vez directamente no se deberían informar y las conclusiones a las que se arriba sólo deberían limitarse a los datos utilizados y a las condiciones en las que se desarrolló el experimento. De la misma manera resulta poco aconsejable elaborar generalizaciones demasiado amplias cuando los estudios se realizan sobre casos particulares o específicos que no son representativos.

A menudo sucede que el investigador prefiere o necesita interpretar las variables en forma separada cuando maneja un grupo de variables correlacionadas. En estos casos sería más apropiado utilizar métodos univariados, con el complemento de pruebas de Bonferroni ajustadas. Sin embargo, cuando sea posible, la consideración conjunta de las variables, mediante la

aplicación del análisis multivariado, puede brindar conclusiones más fuertes que las logradas a través de un grupo de comparaciones simples. Si se tiene cuidado durante el proceso de interpretación de resultados, las combinaciones de variables (componentes, factores, etc.) podrían ser de significativa importancia para estudiar un proceso. El uso racional de la estadística multivariada, el ajuste de un modelo y el conocimiento biológico pueden ayudar al investigador a diseñar con criterio un experimento crucial y a llegar a conclusiones consistentes.

Como dijimos anteriormente, el análisis multivariado aprovecha las relaciones entre las variables para buscar patrones o estructuras en los datos. En este sentido, podría encontrarse un origen de patrones cuando se realizan mediciones sobre grupos de objetos similares. Pueden existir casos donde no se conoce *a priori* si los grupos están ya formados, cuantos son, o cuales objetos pertenecen a cada grupo, es allí donde habría que ver que tipo de análisis es más apropiado.

En general, en esta etapa se podrían usar métodos como los de ordenamiento de objetos (donde se logra la reducción de dimensiones entre las variables o entre objetos a una o pocas dimensiones). Otra posibilidad es hacer análisis de agrupamientos donde los objetos se clasifican en categorías jerárquicas sobre la base de matrices de similitud entre los mismos. En el primer caso, los objetos son representados en un espacio gráfico en los cuales los ejes constituyen gradientes de combinaciones de variables. El método de **análisis de componentes principales**, que es un ejemplo de análisis por ordenamiento, utiliza para tal fin las estructuras de los autovectores (también llamados *eigens* ó raíces latentes) de la matriz de correlación o bien de una matriz de varianza-covarianza entre las variables originales.

El **análisis de componentes principales** y el **análisis factorial** tienen como objetivo encontrar una estructura más simple reduciendo la dimensionalidad de las variables sin perder información. Para simplificar el análisis de los datos se reduce el número de variables a un pequeño número de índices o factores. Algunas diferencias entre estas dos técnicas son que las componentes principales están definidas como una combinación lineal de las variables originales y no están basadas en un modelo estadístico particular y por lo tanto no se requiere el cumplimiento de supuestos previos. En el análisis factorial las variables están expresadas como una combinación de factores, está basado en un modelo especial y requiere el cumplimiento de distintos supuestos. Por otra parte mediante el análisis de componentes principales se busca explicar una gran parte de la varianza total, mientras que con el análisis factorial se enfatiza el estudio en las relaciones entre las variables explicadas con las covarianzas o correlaciones. El análisis factorial resulta apropiado cuando el objetivo consiste en encontrar un grupo de variables similares, altamente correlacionadas, y postular que esas similitudes provienen del hecho de que éstas son variables "latentes o factores" que actúan en forma particular sobre el proceso estudiado.

El análisis de componentes principales es muy utilizado para el análisis de ensayos multambientales comparativos de rendimiento,

donde a partir de la interacción genotipo-ambiente y QTL-ambiente se pueden identificar adaptaciones específicas, estabilidad, ideotipos y subregiones o mega-ambientes. Los gráficos biplot acompañan los resultados del análisis de componentes principales, y a través de estos se grafican de manera simultánea la variabilidad de las observaciones y de las variables.

El investigador interesado en definir grupos homogéneos de objetos o de variables, basados en la similitud o diferencias, y estudiar la estructura natural de las observaciones o de las variables puede aplicar el **análisis de agrupamientos**. Este análisis ha sido tradicionalmente utilizado para propósitos exploratorios. Dado que no es una técnica de inferencia estadística, para que los grupos exhiban la mayor homogeneidad interna y la mayor diferenciación entre grupos sólo se requiere que la muestra sea representativa y que las variables no sean multicolineales. En este análisis, se deben incluir aquellas variables que contribuyan a caracterizar los objetos y que se encuentren relacionadas con los objetivos del análisis. Se recomienda que un análisis de agrupamientos sea posteriormente complementado con un **análisis de funciones discriminantes** sobre los grupos previamente identificados.

El **escalamiento multidimensional**, involucra una serie de técnicas que ayudan al analista a identificar dimensiones subyacentes en la evaluación de objetos. Es especialmente utilizado en ciencias del comportamiento, ya que permite identificar dimensiones no conocidas que afectan el comportamiento. Se basa en evaluaciones comparativas de objetos cuando las bases de comparación son desconocidas o indefinibles. Se transforma el comportamiento de quienes perciben los objetos en distancias, las que finalmente se representan en un espacio multidimensional. Si bien mediante el **análisis de agrupamientos** los objetos se agrupan de acuerdo a sus características, con el escalamiento multidimensional no se enfoca el interés en los objetos en sí sino más bien, en como estos son percibidos. Esta técnica mide la correlación o covarianza de los estímulos derivados de las características de los objetos a ser evaluados. Así, las representaciones grá-



ficas enfatizan las relaciones entre los estímulos que se estudian.

El **análisis de correspondencia** es un procedimiento de ordenación apropiado para datos de tablas de contingencia (frecuencias). Las observaciones multivariadas se grafican en planos (biplot) para identificar las asociaciones de mayor peso entre las modalidades de varias variables cualitativas.

El **análisis de funciones discriminante** se utiliza cuando los objetos o individuos se reúnen en dos o más grupos, definidos *a priori*, frecuentemente se plantea el problema de cómo describir las diferencias entre grupos sobre la base de un conjunto de variables. El propósito básico del análisis de funciones discriminantes es estimar la relación que existe entre una variable dependiente cualitativa (machos vs hembras o genotipos resistentes vs genotipos susceptibles) y entre un grupo de variables independientes cuantitativas con distribución normal. Lo mismo que en el caso de análisis de componentes principales y análisis factorial, el análisis de funciones discriminantes se basa en la posibilidad de encontrar una combinación lineal de las variables originales. Como característica, el análisis de funciones discriminantes permite identificar aquellos grupos ya formados a los cuales pueden ser asignados nuevos objetos en estudio (individuos, genotipos). Posibilita además la identificación de cual o cuales de las variables independientes contribuye más a la diferenciación entre grupos.

El **análisis multivariado de la varianza** es un procedimiento de inferencia estadística usado para evaluar la significancia de la diferencia entre los grupos, que se hallan delimitados sobre la base de las distintas variables independientes discretas o tratamientos. Es una extensión del ANOVA, sólo que las diferencias se establecen teniendo en cuenta varias variables dependientes cuantitativas de distribución normal. Por ello este análisis es particularmente útil cuando los datos provienen de diseños experimentales. También puede ser considerado como una extensión del análisis de funciones discriminantes, aunque en éste la única variable dependiente es categórica y las variables independientes son cuantitativas, mientras que el análisis multivariado de la varianza involucra un grupo de varia-

bles dependientes cuantitativas y las variables independientes son cualitativas.

Las **correlaciones canónicas** reducen las dimensiones de dos grupos de variables provenientes de un mismo conjunto de objetos o individuos de manera que se puedan estudiar las relaciones conjuntas entre los grupos. Es similar al análisis de funciones discriminantes sólo que en las correlaciones canónicas las variables (no los objetos o individuos) son divididas en grupos por lo que el interés se centra sobre las relaciones entre los grupos de variables. Por ejemplo, las correlaciones canónicas se pueden utilizar cuando interesa conocer la relación existente entre los patrones de marcadores moleculares, que representan diferentes genotipos, y algunas características ambientales donde viven los individuos muestreados.

La **regresión múltiple** se considera el método apropiado cuando se presume que los valores de una variable continua dependen de los valores que tomen las otras variables. Esta técnica, que explora todo tipo de relaciones dependientes, tiene como objetivo predecir los cambios en una variable dependiente cuantitativa en respuesta a los cambios en las distintas variables independientes o predictoras. Los modelos de regresión múltiple constituyen la base para la estimación de parámetros en genética cuantitativa. Son ejemplos de su aplicación la descomposición del valor genotípico en el efecto medio de cada alelo y las interacciones debidas a dominancia y epistasia, el cálculo de la estabilidad de los genotipos mediante la descomposición de la interacción genotipo ambiente, el parecido entre parientes, la ubicación y la utilización de QTL mediante marcadores moleculares.

La ecuación de regresión es una combinación lineal de las variables independientes que mejor predicen la variable dependiente. La selección de las variables con mayor poder de predicción se realiza mediante aproximaciones secuenciales, llamadas estimaciones "step-wise" (pasos inteligentes). Mediante los coeficientes de regresión estandarizados es posible determinar la importancia relativa de cada variable predictora. Las variables independiente y dependiente deben ser cuantitativas si bien, mediante transformaciones previas, es posible

incluir variables independientes no métricas.

La aplicación del análisis **regresión logística** permite predecir una variable dependiente binaria (0,1) empleando tanto variables categóricas como cuantitativas. Es importante tener en cuenta que para asignar causalidad en forma justificada es necesario disponer de una cantidad adecuada de datos biológicos experimentales.

Si se cumplen los supuestos básicos, en especial la normalidad de las variables, no hay mayores diferencias entre la regresión logística y el análisis de funciones discriminantes, sólo que en el caso que sea posible formar más de dos grupos, este último resulta el más apropiado. Los **modelos logarítmicos lineales** comprenden a otros análisis que permiten mostrar las relaciones que existen entre variables categóricas.

## **6 Estudio de la interacción QTL-ambiente con ANOVA y ACP**

La selección asistida por marcadores es una inmediata aplicación de la biotecnología que puede ser usada por los mejoradores para localizar un gen específico o un segmento de cromosoma que regula el fenotipo estudiado y que está presente en un individuo o en la población de interés.

En algunas poblaciones de mapeo (poblaciones segregantes con desequilibrio de ligamiento entre los marcadores genéticos y los QTLs) tales como las RIL (líneas endocriadas recombinantes) y en especial las DH (doblehaploides) es posible multiplicar genotipos prácticamente sin variabilidad genética y probarlos en ensayos multiambientales. Los ensayos multiambientales en general brindan la posibilidad de explorar las interacciones genotipo-ambiente (GE) y QTL-ambiente (QE). Si bien el efecto de algunos QTLs explica una alta proporción de la varianza fenotípica estimada en los diferentes ambientes, el efecto de otros QTLs es detectado de manera poco consistente debido a la asociación con ambientes específicos, generando una fuerte interacción QTL-ambiente.

Se ha encontrado que en el maíz tropical las temperaturas de los diferentes sitios afectan la expresión de distintos QTLs, adaptación específica debida, al menos en parte, al efecto de

la selección artificial y a la sustitución alélica ocurrida en diferentes loci. El análisis de los patrones de interacción QTL-ambiente, permite visualizar los efectos principales de los QTLs (Q) dentro y a través de los ambientes o megaambientes, e identificar aquellos que aumentan la efectividad de la selección y la retrocruza asistida por marcadores. Frecuentemente, los modelos que se aplican para estudiar la interacción QTL-ambiente son los mismos o similares a los utilizados para estudios de interacción genotipo-ambiente, tales como el modelo AMMI (Additive Main Effects and Multiplicative Interaction) o el modelo SREG (Sites Regression). El análisis SREG combina la técnica del ANOVA y del ACP en un sólo modelo. El ANOVA permite estudiar el efecto principal de ambiente, mientras que la interacción genotipo-ambiente más el efecto principal de genotipo (G) es tratado de forma multivariada mediante el ACP.

De manera análoga con el método GGE biplot (evaluación simultánea del G + GE), se describe el QQE biplot (Q + QE) que reduce la dimensionalidad de la información obtenida sobre múltiples QTLs en múltiples ambientes. En este método se representan los patrones de interacción QTL-ambiente a partir de una tabla de dos vías (QTLs y ambientes) y por lo tanto es posible visualizar los efectos de los QTLs dentro y a través de ambientes o megaambientes. Esta tabla se descompone en componentes principales y los dos primeros componentes principales se grafican para QTLs y ambientes en la forma de un biplot. El resultado del QQE biplot contiene información sobre los efectos principales de los QTLs o efecto aditivo, resumido como el efecto medio a través de ambientes e información sobre la interacción QTL-ambiente relacionada con la estabilidad de los QTLs. El primer análisis de los QQE biplot consiste en determinar si existen diferentes megaambientes. Si es así, se podrán generar múltiples QQE biplots y los patrones de interacción se investigarán en cada megaambiente.

Para interpretar la metodología del QQE biplot se utilizó una tabla QTL-ambiente (Yan y Tinker, 2005), con ocho QTL (Q1 a Q8) y ocho ambientes (E1 a E8) (Tabla 2). El biplot fue realizado con el programa Infogen (Balzarini y Di

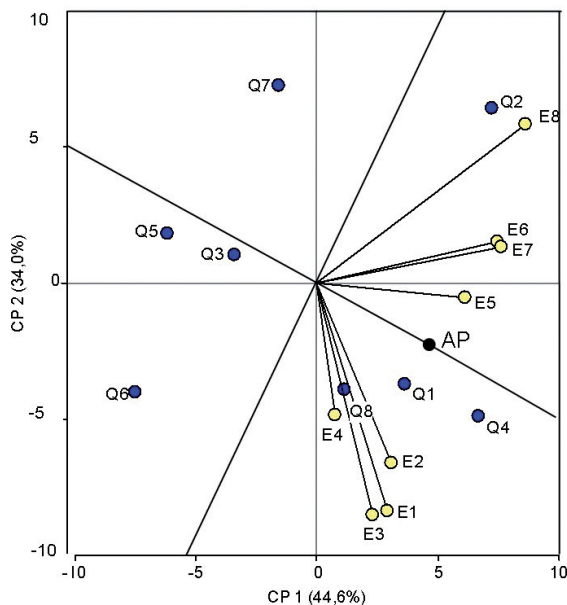
Rienzo, 2004) en base al efecto aditivo de los QTLs.

**Tabla 2.** Efectos aditivos de ocho QTLs (Q1 a Q8) en ocho ambientes (E1 a E8).

	E1	E2	E2	E3	E4	E5	E6	E7	E8
Q1	1	2	2	2	2	1	1	0	1
Q2	-2	-2	-2	-3	-4	2	3	4	6
Q3	2	-2	-2	-1	-2	-3	-2	-1	-2
Q4	2	4	4	3	1	1	4	2	1
Q5	3	-2	-2	-2	-3	-4	-3	-3	-3
Q6	0	0	0	0	0	-3	-4	-5	-6
Q7	-6	-5	-5	-4	-3	0	0	0	0
Q8	3	-3	-3	4	4	3	-3	4	-4

El efecto principal de Q, de E, y de la QE explicaron 40, 3 y 57% de la variación total, respectivamente. El biplot de las CP1 y CP2 explicaron el 79% de la variación total (Figura 1). El QQE biplot permite visualizar distintos temas relacionados con la evaluación y utilización de los QTL.

**QTLs mayores vs menores:** La distancia de los QTLs al origen permite visualizar la magnitud de sus efectos. Así, Q2, Q6, Q7, Q5 y



**Figura 1.** QQE biplot obtenido con los efectos aditivos de la Tabla 2.

Q4 pueden ser considerados QTLs con efectos mayores. Los restantes no son necesariamente menores porque sus efectos podrían aparecer en la CP3 o CP4.

**Similitud entre QTLs:** La distancia y el ángulo entre dos QTLs es un indicador de la similitud en sus respuestas a los ambientes. Así por ejemplo, Q6 es muy diferente de Q2, Q4 y Q7. Los QTLs con patrones similares pueden ser tratados de la misma forma en la selección asistida por marcadores.

**Similitud entre ambientes y clasificación de mega-ambientes:** La distancia y el ángulo entre dos ambientes miden la similitud en la expresión de los QTLs. En la Figura 1 se observan dos grupos de ambientes, E1 a E4 vs E5 a E8. Los ambientes dentro de cada grupo son similares (ángulos pequeños) pero los dos grupos son independientes (ángulo recto). Por lo tanto, se identifican dos mega-ambientes. Esto implica, que la selección de genotipos en un mega-ambiente es independiente del otro, y que las estrategias de selección requeridas en cada mega-ambiente deben ser distintas.

**Efecto principal y estabilidad de los QTL a través de ambientes:** El punto que representa el ambiente promedio (AP) se obtiene con la media de los coeficientes CP1 y CP2 a través de los ambientes. La línea que pasa a través del ambiente promedio y el origen del biplot es llamada eje x del ambiente promedio y las proyecciones sobre este eje cuantifican el efecto de los QTLs. Así, los efectos principales de los QTLs a través de ambientes son:  $Q4 > Q1 > \dots > Q7 > Q5$ . La línea perpendicular al AP que pasa por el origen se llama eje y del ambiente promedio. Las mayores proyecciones de los QTLs sobre el eje x indican poca estabilidad de los mismos. Así Q2, Q7 y Q6 son poco estables y Q1, Q3, Q4 y Q5 son relativamente estables a través de los ambientes. Para una mejor interpretación de Q8 que toma valores intermedios, se podría realizar un gráfico complementario con las CP3 vs CP4.

**Uso de los QTLs en la selección asistida por marcadores en diferentes mega-ambientes:** El biplot QQE también indica la mejor combinación de los alelos de los QTLs para la máxima expresión de los caracteres en cada mega-ambiente. Por ejemplo, la Figura 1 su-

giere que para la máxima expresión del carácter en el mega-ambiente que contiene de E1 a E4 deben ser seleccionados los alelos Q1 y Q4 del parental 1, ya que presentan los valores más altos y positivos. Los mega-ambientes pueden ser examinados en biplots separados para una mejor interpretación.

**Utilidades del biplot QQE:** Los caracteres cuantitativos y aquellos regidos por pocos genes están sujetos a interacciones QE. Las plantas portadoras de genes para resistencia a enfermedades o para enanismo son ejemplos clásicos de genotipos adaptados para responder a ambientes específicos, por ejemplo la presencia de patógenos o altas tasas de fertilizante. Por lo tanto, es necesario entender los patrones de los QTL-ambiente antes de usar los QTL en la selección asistida por marcadores.

### Lecturas recomendadas

- Balzarini, M.; Di Rienzo, J. 2004. Info-Gen: Software para análisis estadístico de datos genéticos. Universidad Nacional de Córdoba. Córdoba. Argentina.
- Everitt, B.S.; Dunn, G. 1991. Applied Multivariate Data Analysis. E. Arnold, London.
- Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. 1995. Multivariate Data Analysis. 4ta. Ed. Prentice Hall - Upper Saddle River, New Jersey.
- James, F.C.; McCulloch, C.E. 1990. Multivariate analysis in ecology and systematic: Panacea or Pandora's box?. Annual review of ecology and systematics, 21: 129-166.
- Manly, B.F. 1986. Multivariate Statistical Methods. A Primer. Chapman and Hall, London.
- Rencher, A.C. 1995. Methods of Multivariate Analysis. J. Wiley and Sons, Inc. New York.
- Yan, W.; Tinker, N.A. 2005. A biplot approach for investigating QTL-by-environment patterns. Molecular Breeding 15: 31-43.

## II. CAPÍTULO 9

### Métodos multivariados para estimar variabilidad genética

Nélida Winzer, Miguel Di Renzo, Sofía Olmos, Mercedes Ibañez.

#### 1 Introducción

La información procesada a partir de las técnicas que se desarrollan en este libro puede ser de diversos tipos y puede tratarse por otro lado, del análisis de una sola variable o de varias variables evaluadas en cada individuo o muestra.

Los análisis estadísticos a utilizar dependen de estos dos aspectos, tipo de datos y número de variables, y, además, de qué objetivo se haya planteado el investigador.

Si se trata de una sola variable se podrán aplicar las técnicas aprendidas en un curso básico de estadística (comparación de dos medias, Análisis de la Varianza, Análisis de Regresión, pruebas chi-cuadrado, etc.), o bien alguna de sus generalizaciones.

En este capítulo se desarrollarán en más detalle algunas metodologías multivariadas. Se han seleccionado aquellas que se utilizan con mucha frecuencia y que no requieren demasiados conocimientos previos.

Los datos multivariados se ordenan en una matriz. A efectos de unificar el lenguaje se considerará que cada fila de la matriz representará un individuo, una muestra, una especie; en definitiva, una OTU (*Operational Taxonomic Unit*). Las columnas representarán los caracteres o variables que se observan en cada OTU.

#### 2 Tipo de datos

**DATOS DOBLE ESTADO:** Son los también llamados datos binarios o dicotómicos. Estos datos se presentan cuando el carácter puede tomar sólo dos estados posibles. Habitualmente se codifican como "0" ó "1".

Hay que reconocer dos tipos de situaciones donde pueden aparecer datos binarios.

a) Datos de presencia o ausencia. Sólo se registra la presencia ó ausencia del carácter.

Ejemplo: Presencia o no de una banda en

una corrida electroforética de isoenzimas o marcadores moleculares.

b) Datos doble estado excluyentes. El carácter tiene sólo dos estados posibles y la asignación del "0" ó el "1" es indistinta.

Ejemplos: Carácter: Sexo. Codificación:

Hembra = 0, Macho = 1, ó

Hembra = 1, Macho = 0

Tipo de Fruto: dehiscente ó indehiscente.

Clasificación de hojas: paripinadas ó imparipinadas.

La distinción entre uno u otro tipo de dato binario es importante al momento de seleccionar la medida de asociación.

**Datos Multiestado Cualitativos:** El carácter puede tomar un conjunto finito de estados.

Ejemplos: Margen de la hoja: aserrado, lobulado, entero.

Color de la flor: blanca, roja o rosada.

Disposición de folíolos en el raquis: alternos, subopuestos, opuestos.

Identificación de los nucleótidos presentes en una determinada secuencia de ADN: A, C, T, G.

Identificación de los aminoácidos esenciales presentes en una proteína: Gly, Ala, Ser, Thr, Cys, Val, Ile, Leu, Pro, Phe, Tyr, Met, Trp, Asn, Gln, His, Asp, Glu, Lys, Arg.

Clases de embriones somáticos: globular, acorazonado, torpedo, cotiledonar.

**Datos Multiestado Cuantitativos Discretos:** Están asociados a valores de conteo.

Ejemplos: Número de hileras por espiga, número de macollos por planta, número de inflorescencias.

**Datos Multiestado Cuantitativos Continuos:** Representan propiedades que se pueden expresar en una escala continua.

Ejemplos: Biomasa de plantas, peso de semillas, longitud de partes verdes, concentración de proteína del grano, frecuencia relativa de un alelo.

#### 3 Medidas de asociación y distancia

Un objetivo frecuente suele ser la descripción del grado de parecido que hay entre las

OTUs. Para ello será necesario medir ese grado de similitud.

Se describirán a continuación algunas de las medidas que aparecen con mayor frecuencia en la literatura.

Hay dos tipos de medidas: los índices de asociación y las distancias. Los primeros miden el grado de similitud: cuanto más parecidas sean dos muestras mayor será su asociación. Las segundas miden el grado de disimilitud: cuanto más parecidas son dos muestras menor será su distancia.

Para datos binarios es usual trabajar con medidas de asociación. Entonces, para caracterizar a dos OTUs existen cuatro valores a considerar:

a = N° de caracteres donde ambas OTUs coinciden en el "1";

b = N° de caracteres donde la primera OTU tiene un "1" y la otra un "0";

c = N° de caracteres donde la primera OTU tiene un "0" y la otra un "1";

d = N° de caracteres donde ambas OTUs coinciden en el "0"

La suma de estos cuatro valores dará el total de caracteres medidos.

		OTU 1	
		1	0
OTU 2	1	a	b
	0	c	d

Si los datos son de presencia/ausencia, un criterio válido es considerar que dos OTUs son más parecidas cuantos más "unos" compartan. La coincidencia de "ceros" no aporta a la similitud porque puede estar generado por falta de información (la no amplificación de un fragmento RAPD puede ser debida a la carencia del sitio de hibridación del cebador por ausencia completa de la secuencia complementaria o como consecuencia de una mutación de punto en dicho sitio o bien, a una falla en la amplificación por ejemplo).

Dos de los índices que comparten este criterio son Jaccard y Dice.

JACCARD

$$J_{12} = \frac{a}{a + b + c}$$

Es la proporción de caracteres presentes que comparten, respecto al total de caracteres presentes en las dos OTUs.

Dice

$$D_{12} = \frac{2a}{2a + b + c} = \frac{a}{(a + b) + (a + c)} \cdot 2$$

Es la proporción de caracteres copresentes respecto al promedio de los caracteres presentes en cada OTU.

Se puede probar que Jaccard siempre dará un valor de asociación menor que Dice, salvo cuando toman el valor 0 ó 1, caso en que siempre coinciden. Más aún, ambos están relacionados mediante la ecuación  $J = D/(2-D)$  ó  $D = 2J/(1+J)$ . Esta relación tiene como consecuencia, entre otras, que en un Análisis de Agrupamientos, cuando se usa ligamiento Simple o Completo, den los mismos grupos sólo que con distintos valores de asociación.

Ejemplo: Como resultado de la aplicación de la técnica RAPD sobre seis muestras se registró la información de la **Tabla 1** sobre la presencia o ausencia de ocho bandas generadas por un cebador.

Este es un proceso típico en el análisis de la información a partir de bandas. Primero se codifica la información, se construye la matriz de datos y, en este caso, se calcularon las matrices de asociación de Jaccard y Dice a efectos de observar las características mencionadas más arriba.

Entre las muestras A y C no se observaron bandas en común, por lo tanto  $a = 0$  y tanto Jaccard como Dice dan 0.

Por otro lado, A y E comparten la presencia de las mismas bandas:  $a = 5$ ,  $b = 0$ ,  $c = 0$ . Tanto Jaccard como Dice dan 1.

Entre B y E:  $a = 3$ ,  $b = 1$ ,  $c = 2$ . Por lo tanto, de las 6 bandas presentes en una u otra muestra comparten 3. Así Jaccard vale 0.5. El total de bandas de B es 4 y el total de bandas detectadas en E es 5. El promedio de estos totales da 4.5 por lo que la asociación de Dice da  $(3/4.5) = 2/3 = 0.667$ .

Se observa también que, salvo los 0 y los 1, todos los valores de Jaccard son menores que los de Dice.

**Tabla 1:** Matriz de datos construida en base a 8 cebadores RAPDs.

	A	B	C	D	E	F
1	—			—	—	
2	—	—		—	—	—
3		—	—	—		—
4	—	—			—	—
5	—	—			—	—
6			—			—
7	—				—	—
8			—			—

→

	A	B	C	D	E	F
1	1	0	0	1	1	0
2	1	1	0	1	1	1
3	0	1	1	1	0	1
4	1	1	0	0	1	1
5	1	1	0	0	1	1
6	0	0	1	0	0	1
7	1	0	0	0	1	1
8	0	0	1	0	0	1

**Bandas**

	1	2	3	4	5	6	7	8
A	1	1	0	1	1	0	1	0
B	0	1	1	1	1	0	0	0
C	0	0	1	0	0	1	0	1
D	1	1	1	0	0	0	0	0
E	1	1	0	1	1	0	1	0
F	0	1	1	1	1	1	1	1

	A	B	C	D	E	F
A	1	0.5	0	0.333	1	0.5
B	0.5	1	0.167	0.4	0.5	0.571
C	0	0.167	1	0.2	0	0.429
D	0.333	0.4	0.2	1	0.333	0.25
E	1	0.5	0	0.333	1	0.5
F	0.5	0.571	0.429	0.25	0.5	1

	A	B	C	D	E	F
A	1	0.667	0	0.5	1	0.667
B	0.667	1	0.286	0.571	0.667	0.727
C	0	0.286	1	0.333	0	0.6
D	0.5	0.571	0.333	1	0.5	0.4
E	1	0.667	0	0.5	1	0.667
F	0.667	0.727	0.6	0.4	0.667	1

Además, ambas matrices son simétricas: la asociación entre A y E es la misma que la que se da entre E y A. Por eso, basta mostrar la mitad de la matriz de asociación.

Esta observación será válida para todas las medidas que se presenten en este capítulo.

Si los datos son binarios pero de estados equivalentes, dos muestras deberán considerarse más asociadas tanto cuando compartan

"unos" como "ceros", ya que la codificación que se les asigna es indistinta.

Un índice recomendable por su simplicidad es el de *Simple Matching* o de Concordancia Simple:

SIMPLE MATCHING  $SM_{12} = \frac{a+d}{a+b+c+d}$

Simplemente es la proporción de caracteres que comparten las dos muestras.

Si se piensa que la matriz de datos anterior corresponde a ocho caracteres binarios de estados equivalentes la matriz de asociación con el índice de Simple Matching dará:

	A	B	C	D	E	F
A	1	0.625	0	0.5	1	0.5
B		1	0.375	0.625	0.625	0.625
C			1	0.5	0	0.5
D				1	0.5	0.250
E					1	0.5
F						1

Se observa que A y E comparten todos los caracteres, por lo tanto tienen la máxima asociación. En cambio C y E no coinciden en ninguno, su asociación es 0. La asociación entre C y D y entre D y E es 0.5, lo que significa que coinciden en la mitad de los caracteres evaluados.

Además de las que se han definido hasta ahora existen muchas otras medidas para datos binarios. También se pueden definir medidas de distancia o asociación para distintos tipos de datos multiestado. Para no extender en demasía esta sección se presentarán algunas distancias definidas específicamente para frecuencias alélicas.

#### 4 Distancias Genéticas

Se pueden clasificar en dos grupos: las basadas en un modelo geométrico (Cavalli-Sforza y Rogers) y las basadas en modelos biológicos (Nei).

Si se tienen dos poblaciones de las que se cuenta con la información sobre los *loci* para  $m$  alelos y las frecuencias correspondientes, éstas pueden expresarse de la siguiente manera:

Pob. 1:  $p_1, p_2, \dots, p_m$   
 Pob. 2:  $q_1, q_2, \dots, q_m$

#### Distancia Cuerda de CAVALLI-SFORZA y EDWARDS (1967):

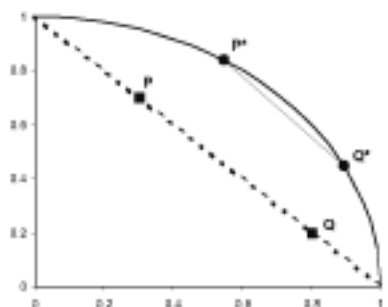
$$d_{12} = \sqrt{2 \left( 1 - \sum_{i=1}^m \sqrt{p_i q_i} \right)}$$

Los puntos  $P^* = (\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_m})$

y  $Q^* = (\sqrt{q_1}, \sqrt{q_2}, \dots, \sqrt{q_m})$  están sobre una

esfera de radio 1. La distancia Cuerda es, precisamente, la longitud de la cuerda que une esos dos puntos.

Para el caso  $m = 2$ , se ve en el gráfico la distancia cuerda entre los puntos  $P = (0.3, 0.7)$  y  $Q = (0.8, 0.2)$ . Los puntos originales son llevados a la esfera (círculo) de radio 1 y luego se calcula la distancia habitual entre esos dos puntos:  $d = 0.5324$ .



Si se consideran simultáneamente varios *loci* la distancia Cuerda se calcula así:

$$d_{12} = \sqrt{2 \left( 1 - \sum_{i=1}^r \sum_{j=1}^{m_i} \sqrt{p_{ij} q_{ij}} \right)}$$

#### Distancia de ROGERS (1972):

$$R_{12} = \frac{1}{r} \sum_{i=1}^r \sqrt{\sum_{j=1}^{m_i} (p_{ij} - q_{ij})^2}$$

En el gráfico, ésta es la distancia habitual entre los puntos P y Q.

#### Distancia Estándar de NEI (1972):

Esta distancia tiene una base biológica. Expresa la probabilidad de que dos alelos tomados al azar de dos poblaciones diferentes sean idénticos, con respecto a la probabilidad de que lo sean dos alelos tomados al azar de la misma población. Sean, como antes, las poblaciones 1 y 2 y las frecuencias alélicas de un *locus*:



$\sum p_i q_i$  se puede interpretar como la probabilidad que dos alelos sean iguales si uno ha sido elegido al azar de la población 1 y el otro de la población 2.

$\sum p_i^2$  es la probabilidad que dos alelos elegidos al azar de la población 1 sean iguales;

$\sum q_i^2$  es la probabilidad que dos alelos elegidos al azar de la población 2 sean iguales.

Los dos últimos términos equivalen a un estado de homocigosis en la población 1 y 2, respectivamente.

Nei define primero la *Identidad normalizada* (varía entre 0 y 1):

$$I = \frac{\sum_{i=1}^m p_i q_i}{\sqrt{\sum_{i=1}^m p_i^2 \sum_{i=1}^m q_i^2}}$$

y luego la *Distancia Estándar*:  $D_{12} = -\ln I$ .

Si se usan varios *loci* se trabaja con el promedio aritmético de las probabilidades definidas anteriormente resultando:

$$D_{12} = -\ln \frac{\sum_{i=1}^r \sum_{j=1}^{m_i} p_{ij} q_{ij}}{\sqrt{\sum_{i=1}^r \sum_{j=1}^{m_i} p_{ij}^2 \sum_{i=1}^r \sum_{j=1}^{m_i} q_{ij}^2}}$$

Ejemplo: Para la información de la Tabla 2, en el caso de las frecuencias de cinco alelos provenientes de dos *loci*, se tiene:  $I = 0.857986$  y, por lo tanto,  $D = 0.15317$ .

loci	1			2		
	1a	1b	1c	2a	2b	
Población 1	0.4	0.3	0.3	0.2	0.8	p
Población 2	0.6	0.05	0.35	0.5	0.5	q

**Tabla 2:** Datos de frecuencias alélicas de dos *loci* para el cálculo de la Distancia Estándar de Nei.

En las poblaciones naturales una de las formas de realizar el agrupamiento jerárquico de las mismas en grupos (variedades, líneas, genotipos), es mediante el empleo de medidas de divergencia genética. Es en este campo donde las llamadas distancias genéticas tienen su mayor aplicación. Las distancias genéticas basadas en modelos biológicos son las más empleadas en estudios de genética poblacional debido a que resumen los conocimientos obtenidos de las fuerzas evolutivas que conllevan a los cambios genéticos es decir, mediante mutaciones y deriva genética.

La medida de distancia genética estándar de Nei es más confiable cuando se utilizan datos provenientes de muchos alelos. La distancia de Nei está formulada para un modelo donde se asume que las mutaciones ocurren en forma infinita, con la misma probabilidad para todos los alelos a una tasa de mutación neutral, y que la variabilidad genética inicial en la población está en equilibrio entre la variabilidad producida por mutaciones y por deriva genética, siendo el tamaño efectivo de la población, en cada una, constante. Basándose en estos supuestos, y en que todos los *loci* sean equivalentes entre sí, se espera que la Distancia de Nei se incremente linealmente con el tiempo. Si las frecuencias alélicas en cada población han sido estimadas con pocas muestras se pueden utilizar, en cambio, algunas modificaciones a la Distancia Estándar de Nei como las propuestas por Nei (1978) y Hillis (1984). La distancia de Cavalli-Sforza y Edwards, por el contrario, asume que las diferencias entre las poblaciones no provienen de mutaciones sino que son sólo consecuencia de la deriva genética. Además, no asume que el tamaño poblacional ha permanecido constante en todas las poblaciones. Considera que el tamaño de la población cambia en forma lineal pero no con el tiempo sino con  $1/N$ , donde  $N$  es el tamaño efectivo de la población. Así, en estos casos, los conocimientos básicos de genética poblacional brindarán las herramientas necesarias en el momento de la elección del método más apropiado para el análisis de nuestros datos.

### 5 Análisis de coordenadas principales

Si se tiene una matriz de distancias entre  $N$  muestras ¿será posible representar cada

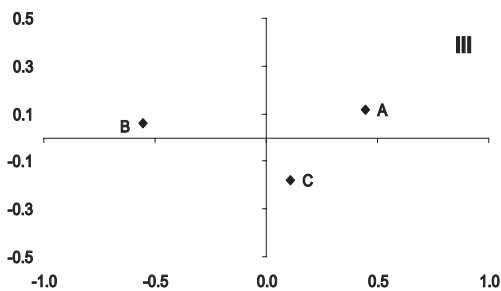
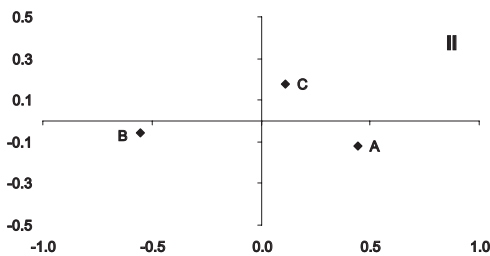
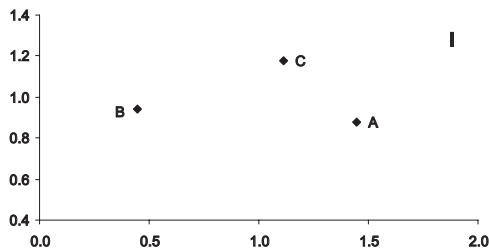
muestra mediante un punto de manera tal que las distancias resultantes reproduzcan las de la matriz?.

Veamos dos ejemplos:

(i) Si se tiene la siguiente matriz D de distancias entre las muestras A, B y C:

$$D = \begin{pmatrix} 0 & 1 & .447 \\ 1 & 0 & .707 \\ .447 & .707 & 0 \end{pmatrix}$$

se podrían hacer algunos de los siguientes gráficos:



Coordenadas de A, B y C:

- I: (1.444, 0.881); (0.445, 0.94); (1.111, 1.179)
- II: (0.444, -0.119); (-0.555, -0.06); (0.111, 0.179)
- III: (0.444, 0.119); (-0.555, 0.06); (0.111, -0.179).

Si se calcula la distancia euclídea entre los puntos en uno cualquiera de estos gráficos se ve que se han reproducido exactamente los valores de la matriz D. Por ejemplo:

$$d_E(A,B) = \sqrt{(1.444 - 0.445)^2 + (0.881 - 0.94)^2} = 1$$

en el gráfico I.

(ii) Si se tuviera la matriz de distancias entre las muestras P<sub>1</sub>, P<sub>2</sub> y P<sub>3</sub>:

$D^{\#} = \begin{pmatrix} 0 & 0.9 & 0.5 \\ 0.9 & 0 & 0.3 \\ 0.5 & 0.3 & 0 \end{pmatrix}$  Es fácil ver que es imposible representar en un plano puntos separados

por estas distancias porque si P<sub>1</sub> y P<sub>2</sub> se ubican como extremos de un segmento de longitud 0.9, no hay manera de ubicar un punto P<sub>3</sub> que esté a 0.5 de P<sub>1</sub> y a 0.3 de P<sub>2</sub>.

P<sub>1</sub>-----P<sub>2</sub>

P<sub>1</sub>-----P<sub>3</sub>? P<sub>3</sub>?-----P<sub>2</sub>

Se van a considerar, por ahora, distancias como las del ejemplo (i). Esto es, distancias para las que se puedan encontrar puntos que reproduzcan esas distancias.

Otro aspecto a considerar es el número de coordenadas necesarias para representar esos puntos. Es intuitivo que si sólo hay dos muestras a una distancia dada se pueden graficar en una recta (dimensión = 1); si hay tres muestras ya se vio que se pueden graficar en un plano (dimensión = 2); si se tuvieran cuatro muestras serían necesarias tres coordenadas (dimensión = 3); ... y si se tienen N muestras se necesitarán, en general, N-1 coordenadas. Con lo cual si se quieren graficar 40 muestras serán necesarias 39 coordenadas!!!

Es evidente que esto no es cómodo si el objetivo es obtener una representación gráfica de los puntos o muestras. Lo ideal sería obtener dos coordenadas pues se podrían graficar en un plano, o bien tres para graficar en tres dimensiones.

El problema, entonces, es encontrar las **dos mejores** coordenadas (ó las tres mejores) para

obtener esta representación. Ese es el objetivo del Análisis de Coordenadas Principales (ACOOOP): encontrar las  $k$  mejores coordenadas para las muestras de manera tal que si se calculan las distancias euclídeas entre ellas reproduzcan lo mejor posible una matriz de distancias dada. Habitualmente  $k = 2$  ó  $3$ , a éstas se las llama Coordenadas Principales. Naturalmente se perderá la información de las restantes coordenadas.

Como se mostró en el ejemplo (i) hay muchas representaciones posibles. Ahí se han dado tres pero se podrían haber dado muchas más. Una manera de reducir las opciones es obtener coordenadas de manera tal que los puntos estén centrados en el origen, como en II ó III.

Se explicará el método de obtención de las Coordenadas Principales partiendo de una matriz de Asociación  $S$ . En este caso las distancias que se van a reproducir son las definidas por:

$$d^2(i, j) = S_{ii} + S_{jj} - 2 S_{ij} \quad (1)$$

donde  $S_{ij}$  es la asociación entre la muestra  $i$  y la  $j$ . Si la asociación de una muestra consigo misma es 1, es decir  $S_{ii} = 1$ , como ocurre con la mayoría de los índices, la ecuación se reduce a:

$$d^2(i, j) = 2(1 - S_{ij}). \quad (2)$$

Se detallarán los pasos que se deberían realizar para hacer este análisis al simple efecto de poder manejar con idoneidad los programas estadísticos de que se dispongan. Los pasos 1) a 3) son responsabilidad de esos programas. El usuario deberá, fundamentalmente, indicarle qué asociación o distancia desea calcular y cuántas coordenadas quiere.

PASO 1) Centrar doblemente la matriz  $S \rightarrow S^0$ .

PASO 2) Calcular los autovalores y autovectores de  $S^0$ .

PASO 3) Calcular las coordenadas de los puntos representativos de las muestras.

PASO 4) Graficarlos.

El PASO 1 tiene como fin hacer que los puntos resultantes estén centrados. Cada elemento de la matriz de asociación  $S$  se centra por filas y por columnas:

$$S_{ij}^0 = S_{ij} - S_{i\bullet} - S_{\bullet j} + S_{\bullet\bullet} \quad (3)$$

donde es  $S_{i\bullet}$  el promedio de la fila  $i$  de  $S$ ,  $S_{\bullet j}$  es el promedio de la columna  $j$  y  $S_{\bullet\bullet}$  es el promedio de todos los elementos de  $S$ .

En el PASO 2 se encuentran:

a) una matriz  $N \times N$  donde cada columna es un vector de longitud uno. Son los *autovectores*.

b)  $N$  números ordenados en forma decreciente: los *autovalores*.

Estos elementos son característicos de cada matriz  $S^0$ . Por eso también se los conoce como vectores y valores característicos. Cada autovector está asociado a un autovalor.

Si la matriz de distancias es "representable" como la del ejemplo (i) los autovalores serán positivos o ceros (el más chico es siempre cero para estas matrices "representables").

En el PASO 3, si se desean  $k$  coordenadas, se toman los  $k$  primeros autovectores y se multiplican por la raíz del autovalor respectivo. El primer autovector da la primera coordenada de todas las muestras, el segundo la segunda coordenada y así siguiendo hasta la  $k$ -ésima. Si sólo se desean dos coordenadas para graficar los puntos en un plano se toman los dos primeros autovectores ( $k=2$ ).

En la **Tabla 3** se presenta un ejemplo sencillo.

El gráfico correspondiente es el Gráfico II que se mostró más arriba donde A es la muestra 1, B la 2 y C la muestra 3.

Las distancias (2) calculadas a partir de la matriz de asociación  $S$  dan la matriz de Distancias  $D$ . Por esa razón ambas matrices tienen la misma representación.

$$d_{12}^2 = 2(1 - S_{12}) = 2(1 - 0.5) = 1 \Rightarrow d_{12} = 1$$

$$d_{13}^2 = 2(1 - S_{13}) = 2(1 - 0.9) = 0.2 \Rightarrow d_{13} = 0.447$$

$$d_{23}^2 = 2(1 - S_{23}) = 2(1 - 0.75) = 0.5 \Rightarrow d_{23} = 0.707$$

**OBSERVACIÓN 1:** Como en el ejemplo sólo hay tres muestras los autovectores tienen tres componentes. Si se trabajara con una matriz de asociación entre 40 muestras, los autovec-

tores tendrían 40 componentes, una por cada muestra.

**OBSERVACIÓN 2:** Para la matriz S se han reconstruido totalmente las distancias porque sólo son tres muestras. Si hubiera más (N=40, por ejemplo) sólo se reconstruirían totalmente las distancias si se usan 39 coordenadas. Si sólo se usan las dos primeras se reconstruye, en general, sólo un porcentaje de las distancias. Hay dos tipos de porcentajes que se pueden calcular:

*Porcentaje de Dispersión:*

$$\alpha_1 = \frac{\lambda_1 + \lambda_2}{\lambda_1 + \dots + \lambda_N} \times 100\%$$

si sólo se usan dos coordenadas.

Este porcentaje se interpreta teniendo en cuenta que:

$$2N (\lambda_1 + \dots + \lambda_N) = \sum \sum d_{ij}^2$$

sumando sobre todos los pares de muestras. En cambio, la suma del cuadrado de las distancias reconstruidas con las dos coordenadas es igual a  $2N (\lambda_1 + \lambda_2)$ . Por lo tanto, este porcentaje mide cuánto se reconstruye del cuadrado de todas las distancias.

Si se usan más de dos coordenadas se van sumando los autovalores correspondientes en el numerador.

En nuestro ejemplo:  $\alpha_1 = 100\%$

*Porcentaje de Distorsión:*

$$\alpha_2 = \frac{\lambda_1^2 + \lambda_2^2}{\lambda_1^2 + \dots + \lambda_N^2} \times 100\%$$

Este valor mide cuánto se acercan los valores de la matriz de asociación reconstruida respecto de la matriz de asociación original  $S^0$ . Más exactamente:

$$\alpha_2 = \left( 1 - \frac{\|S^0 - S^{0*}\|^2}{\|S^0\|^2} \right) \times 100\% = \left( 1 - \frac{\sum (s_{ij}^0 - s_{ij}^{0*})^2}{\sum (s_{ij}^0)^2} \right) \times 100\%$$

En este algoritmo se ve que lo que se está comparando son las asociaciones originales,  $s_{ij}^0$ , con las reconstruidas,  $s_{ij}^{0*}$ , una a una. Para el cálculo de  $\alpha_2$  se usa la primera fórmula por lo que el lector no debe traumatizarse con la segunda.

$S = \begin{pmatrix} 1 & 0.5 & 0.9 \\ 0.5 & 1 & 0.75 \\ 0.9 & 0.75 & 1 \end{pmatrix}$	$\Rightarrow$	<p>Doble centrado</p> $S^0 = \begin{pmatrix} 0.2111 & -0.239 & 0.028 \\ -0.239 & 0.3111 & -0.072 \\ 0.028 & -0.072 & 0.0444 \end{pmatrix}$
<p>Autovalores de <math>S^0</math>: <math>\lambda_1 = 0.5167</math>   <math>\lambda_2 = 0.05</math>   <math>\lambda_3 = 0</math></p>		
<p style="text-align: center;">↓                      ↓                      ↓</p> <p style="text-align: center;"><math>v_1</math>                      <math>v_2</math>                      <math>v_3</math></p>		
<p>Autovectores de <math>S^0</math>: Columnas de V =</p> $\begin{pmatrix} 0.6173 & -0.5344 & 0.57735 \\ -0.7716 & -0.2674 & 0.57735 \\ 0.1543 & 0.8019 & 0.57735 \end{pmatrix}$		
<p>Primeras 2</p>		
Coordenadas:	$\sqrt{\lambda_1} v_1$	$\sqrt{\lambda_2} v_2$
Muestra 1⇒	0.4437	-0.1195
Muestra 2⇒	-0.5546	-0.0598
Muestra 3⇒	0.1109	0.1793

**Tabla 3.** Cálculo de autovalores y autovectores para el gráfico de coordenadas principales.

**OBSERVACIÓN 3:** Si se hubieran obtenido autovalores negativos la primera medida no es recomendable. Hasta podría darse el caso que  $\alpha_1$  fuera mayor al 100%. La segunda, en cambio, sigue teniendo sentido. Por otro lado, no se podrían calcular coordenadas asociadas a ese autovalor porque no se puede calcular  $\sqrt{\lambda_i}$

La aparición de autovalores negativos no es necesariamente un problema grave para la representación siempre que los autovalores negativos sean pocos y pequeños. Por ejemplo, si se trabaja con 50 muestras, en el PASO 2 se obtendrán 50 autovalores, de los cuales uno, seguramente, es cero. Para la representación en el plano se usan sólo los dos primeros. Si los últimos 10 son negativos no se verán afectados los cálculos de las coordenadas.

Hay asociaciones y distancias que nunca dan autovalores negativos, entre ellos se encuentran Jaccard, Ochiai, Russell y Rao, Simple Matching y la distancia euclídea.

Si en lugar de tener una matriz de asociaciones se tiene una matriz de distancias el procedimiento consiste en calcular una matriz S cuyos elementos son:

$$s_{ij} = -\frac{d_{ij}^2}{2}, \quad (4)$$

y luego seguir los PASOS 1, 2, 3 y 4 anteriores.

**OBSERVACIÓN 4:** Cada autovector de  $S^0$  es un vector de longitud 1 en una cierta dirección. Pero por cada dirección podemos tener dos vectores:  $v$  y  $-v$ . Por ejemplo:

$$v_2 = \begin{pmatrix} -0.5344 \\ -0.2674 \\ 0.8019 \end{pmatrix} \quad y \quad -v_2 = \begin{pmatrix} 0.5344 \\ 0.2674 \\ -0.8019 \end{pmatrix}$$

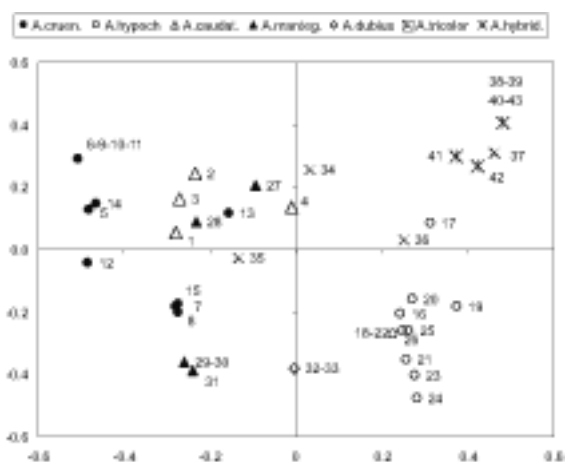
Cualquiera de estos dos vectores pueden usarse como segundo autovector y, más aún, algunos programas pueden dar como resultado el primero y otros el segundo. Como consecuencia de esta selección se obtendrá el gráfico II ó el III. Se ve que el único efecto de usar uno u otro vector es una reflexión del gráfico respecto del primer eje. Si se cambia el signo de  $v_1$  se obtendrá una reflexión respecto del segundo eje. Ninguno de estos cambios modifica la distancia entre los puntos.

No debe causar sorpresa, entonces, que con un programa se obtenga un gráfico y con otro el mismo gráfico pero con una reflexión sobre uno u otro eje, o sobre los dos.

Por otro lado, también es lícito cambiar, por alguna razón, el signo de una coordenada. Por ejemplo, esto puede ser útil cuando se quieren comparar Análisis de Coordenadas Principales realizados con distintas asociaciones o distancias sobre el mismo conjuntos de muestras.

Ejemplo:

Presentamos la caracterización de 8 cultivares comerciales y 35 cultivares en experimentación mediante análisis izoenzimático de semillas de *Amaranthus*. Estas accesiones pertenecen a siete especies: *A. caudatus*, *A. cruentus*, *A. hypochondriacus*, *A. mantegazzianus*, *A. dubius*, *A. tricolor*, y *A. hybridus* (Di Renzo et al., 2001). Del total de 43 cultivares, 31 fueron polimórficos con 7 sistemas isoenzimáticos (**Tabla 4**). Se aplicó análisis de coordenadas principales a los datos de presencia/ausencia de las 23 bandas polimórficas. En la figura puede observarse el nivel de diversidad genética dentro de este germoplasma representado por las coordenadas 1 y 2. Este análisis agrupó la mayoría de los cultivares de acuerdo a su clasificación taxonómica y mostró que existen relaciones entre diferentes accesiones. Se concluye que el análisis isoenzimático de semillas fue efectivo para caracterizar los cultivares destinados a los programas de mejoramiento de *Amaranthus*.



Al sólo efecto de que el lector pueda reparar algunos cálculos en la **Tabla 5** se dan los primeros diez elementos de los dos primeros autovectores y las dos primeras coordenadas principales.

## 6 Análisis de agrupamientos (cluster)

Esta técnica tiene como objetivo formar grupos de muestras, poblaciones, individuos u OTUs que sean similares entre sí y diferentes a los elementos de los otros grupos.

El primer problema es fijar el criterio para decidir cuándo dos elementos son similares. Si se tuviera un mazo de cartas, ¿cómo formar grupos? ¿por el valor de la carta?, ¿porque comparten el palo?. Según el criterio que se aplique los grupos que se formen serán distintos.

En la sección anterior hemos visto distintas maneras de definir la similitud o disimilitud entre OTUs. Cada una de ellas, u otras que no se han presentado aquí, plantean distintas formas de medir la similitud entre los elementos a agrupar.

Existen, además, muchos métodos para formar los grupos. Uno de los más usados son los agrupamientos *jerárquicos* que pueden, a su vez, ser *divisivos* o *aglomerativos*. Los primeros comienzan con todos los elementos en un solo grupo y en sucesivas divisiones se van formando grupos cada vez más pequeños. Los aglomerativos, por el contrario, empiezan considerando tantos grupos como elementos se tienen y se van agrupando según su grado de parecido. Los sucesivos pasos de uno y otro método se pueden representar en un gráfico denominado *dendrograma*.

En esta sección se tratarán solamente los métodos jerárquicos aglomerativos.

Aquí se plantea un segundo problema: ¿Qué tipo de ligamiento se va a utilizar? Esto es, una vez realizado el primer agrupamiento, cómo se define la distancia entre ese grupo y los restantes elementos?.

Supongamos que en un primer paso se han agrupado los elementos A y B porque son los que presentan la menor distancia entre sí. La distancia entre el nuevo grupo AB y otro elemento C se puede definir según distintos criterios:

LIGAMIENTO SIMPLE:  $d(AB,C) = \min \{d(A, C), d(B, C)\}$ ,

LIGAMIENTO COMPLETO:  $d(AB,C) = \max \{d(A, C), d(B, C)\}$ ,

LIGAMIENTO PROMEDIO:  $d(\{A,B\},C) = \frac{d(A,C) + d(B,C)}{2}$

En el caso del ligamiento simple, la nueva distancia se define es la mínima distancia entre

el elemento elemento C y cada elemento del grupo formado AB. En cambio, el ligamiento completo la nueva distancia se define tomando la máxima distancia entre C y cada elemento del nuevo grupo. Si A, B y C fueran grupos formados en pasos anteriores, se usan los mis-

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
5	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	0	1
6	0	0	0	1	0	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	0	1
9	0	0	0	1	0	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	0	1
10	0	0	0	1	0	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	0	1
11	0	0	0	1	0	1	1	1	1	0	1	0	0	0	0	1	1	1	1	0	0	0	1
12	0	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1
13	0	0	0	1	0	0	1	0	1	0	0	0	0	0	1	1	0	1	0	0	0	0	1
14	0	0	0	1	0	1	0	1	1	1	0	0	0	0	0	1	1	1	1	0	0	0	1
15	0	0	0	1	0	0	0	0	1	1	0	0	0	0	0	1	1	1	1	1	1	1	0
19	0	1	1	0	1	0	1	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	1
16	0	1	0	1	0	1	1	1	1	0	1	0	0	0	0	1	0	0	1	0	0	0	1
20	0	1	0	0	1	0	1	1	1	0	1	0	0	0	0	1	1	0	1	0	0	0	1
21	0	1	0	1	0	1	0	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	1
28	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	0	1	0	0	0	1
27	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0
1	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	0	1	0	1	0	0	1
2	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	0	0	1
3	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	0	0	0
23	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	1	0	0	0	0	1
24	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	1	0	0	1	0	0	0	1
25	0	1	0	0	1	0	0	1	0	1	0	0	0	0	0	1	1	0	1	0	0	0	1
26	0	1	0	0	1	0	0	0	1	0	1	0	0	0	0	1	1	0	1	0	0	0	1
29	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	1	0	0	1
30	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	1
31	0	1	0	0	1	0	1	0	1	0	0	0	1	0	1	1	0	0	1	1	0	0	1
32	0	1	0	1	0	1	1	1	0	0	1	1	1	1	0	1	0	0	1	0	0	0	1
33	0	1	0	1	0	1	0	1	1	0	0	1	1	1	1	1	0	1	0	0	0	0	1
18	0	1	1	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	1
22	0	1	1	1	0	0	1	1	0	1	0	0	0	0	0	0	1	1	0	0	0	0	1
7	0	1	1	0	1	0	1	0	1	0	0	0	0	0	0	1	1	1	0	1	0	1	1
17	0	1	1	0	1	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0	0	1
34	1	1	0	0	0	1	0	1	0	1	1	0	0	0	0	1	0	0	1	0	0	0	1
35	1	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	1	1
36	1	1	1	0	1	0	0	1	1	1	0	0	0	0	0	1	0	0	1	0	0	1	1
37	1	1	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
38	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
39	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
40	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
41	1	1	1	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	1	0	0	0
42	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
43	1	1	0	0	1	0	1	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	1
4	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
8	0	0	1	1	0	1	0	1	1	0	0	0	0	0	0	0	0	1	1	1	0	0	0

Tabla 4. Matriz de datos de análisis isoenzimático de 8 cultivares de *Amaranthus*.

	V1	V2	CP1	CP2	
1	-0.127	0.031	-0.279	0.055	<b>Autovalores:</b>  <b>4.81258</b> <b>3.22431</b> $\alpha_1 = 46.85$ $\alpha_2 = 78.2$
2	-0.107	0.137	-0.235	0.246	
3	-0.123	0.091	-0.270	0.163	
4	-0.005	0.074	-0.010	0.134	
5	-0.220	0.070	-0.483	0.126	
6	-0.232	0.162	-0.509	0.291	
7	-0.125	-0.112	-0.273	-0.202	
8	-0.125	-0.097	-0.274	-0.174	
9	-0.232	0.162	-0.509	0.291	
10	-0.232	0.162	-0.509	0.291	
...	...	...	...	...	

Tabla 5. Resultados de autovectores y coordenadas principales del análisis isoenzimático de 8 cultivares de *Amaranthus*.

	AE	B	C	D	F
AE	0	1	1.41	1.15	1.1
B	1	0	1.29	1.10	<b>0.93</b>
C	1.41	1.29	0	1.26	1.07
D	1.15	1.10	1.26	0	1.22
F	1.1	<b>0.93</b>	1.07	1.22	0

	AE	BF	C	D
AE	0	1	1.41	1.15
BF	1	0	1.07	1.10
C	1.41	1.07	0	1.26
D	1.15	1.10	1.26	0

	AEBF	C	D
AEBF	0	1.07	1.10
C	1.07	0	1.26
D	1.10	1.26	0

	AEBFC	D
AEBFC	0	1.10
D	1.10	0

	AE	BF	C	D
AE	0	1.1	1.41	1.15
BF	1.1	0	1.29	1.22
C	1.41	1.29	0	1.26
D	1.15	1.22	1.26	0

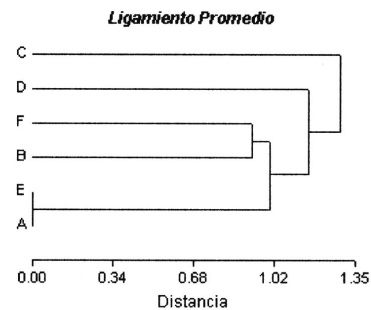
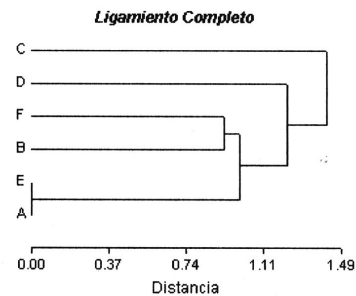
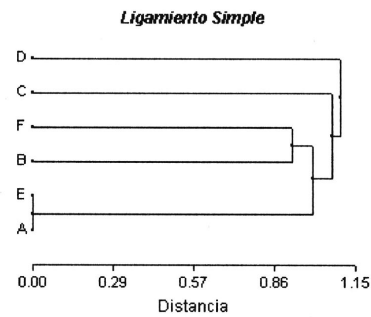
	AEBF	C	D
AEBF	0	1.41	1.22
C	1.41	0	1.26
D	1.22	1.26	0

	AEBFD	C
AEBFD	0	1.41
C	1.41	0

	AE	BF	C	D
AE	0	1.05	1.41	1.15
BF	1.05	0	1.18	1.16
C	1.41	1.18	0	1.26
D	1.15	1.16	1.26	0

	AEBF	C	D
AEBF	0	1.295	1.155
C	1.295	0	1.26
D	1.155	1.26	0

	AEBFD	C
AEBFD	0	1.277
C	1.277	0



**Tabla 6.** Diferentes clases de agrupamiento en base a la matriz de distancias.

mos algoritmos en el caso del ligamiento simple o completo. Para el ligamiento promedio hay que promediar todas las distancias entre elementos de AB y de C: donde  $d_{ij}$  es la distancia

$$d(AB, C) = \frac{\sum_{i,j} d_{ij}}{n_{AB} n_C}$$

entre el elemento  $i$  del grupo AB y el  $j$  de C;  $n_{AB}$  y  $n_C$  son los números de elementos en el grupo AB y C, respectivamente.

Este ligamiento promedio se conoce también por las siglas UPGMA (Unweighted Pair Groups Method with Arithmetic Averages).

Consideremos el ejemplo de la **Tabla 6**. Primero se realizará un agrupamiento aplicando ligamiento simple.

En primer lugar se agrupan los elementos A y E porque están a menor distancia (son iguales). Formado ese primer grupo habría que calcular las nuevas distancias del grupo AE a los restantes elementos pero como  $d(A, U) = d(E, U)$  para todos los restantes elementos U se mantienen esas distancias.

En el segundo paso se observa que la menor distancia es la de B a F; se forma ese grupo y se recalculan las distancias del nuevo grupo a los restantes. Y así hasta que se han agrupado todos. También se puede detener el proceso fijando algún otro criterio para ello. Por ejemplo, cuando se ha alcanzado una distancia máxima fijada previamente.

A continuación se aplica el ligamiento completo. Los dos primeros agrupamientos coinciden en este ejemplo, por lo que sólo se muestran los tres últimos pasos. Por último se ven los pasos cuando se aplica un ligamiento promedio. Los mismos métodos se pueden aplicar sobre matrices de asociación cambiando los conceptos de "menor distancia" por el de "mayor asociación".

En general, el ligamiento simple produce agrupamientos en cadena: a los grupos formados inicialmente se van acoplando los restantes elementos. En cambio, el ligamiento completo tiende a generar muchos grupos chicos que se reunirán en los últimos pasos.

El análisis de Agrupamientos sólo debe tomarse como un **método exploratorio**. Si originalmente los datos forman grupos bien separados cualquier método de agrupamiento dará, aproximadamente, los mismos resultados. Pero si las muestras forman un continuo, cada método y cada medida de distancia pueden conducir a resultados muy dispares. En este caso las conclusiones deben realizarse con cautela.

A veces es conveniente realizar un Análisis de Coordenadas Principales para determinar si existen grupos claramente definidos o no.

## 7 Lecturas recomendadas

- Cavalli-Sforza, L.L.; Edwards, A.W.F. 1967. Phylogenetic analysis: models and estimation procedures. *Am J Hum Gen* 19: 233-257.
- Di Renzo, M.; Bonamico, N.; Gesumaria, J. 2001. Characterisation of amaranth accessions by isozymic patterns. *Seed Sci Technol* 29 (1): 227-238.
- Everitt, B.S.; Dunn, G. 1991. *Applied Multivariate Data Analysis*. E. Arnold, London.
- Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. 1995. *Multivariate Data Analysis*. 4ta. Ed. Prentice Hall - Upper Saddle River, New Jersey.
- Hillis, D. 1984. Misuse and modification of Nei's genetic distance. *Syst Zool* 33: 238-240.
- Manly, B.F. 1986. *Multivariate Statistical Methods. A Primer*. Chapman and Hall, London.
- Nei, M. 1972. Genetic distances between populations. *Am Nat* 106: 283-292.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from small number of individuals. *Genetics* 76: 379-390.
- Rencher, A.C. 1995. *Methods of Multivariate Analysis*. J. Wiley & Sons, Inc. New York.