

Claves moleculares de la biodiversidad vegetal

Javier Fuertes Aguilar, Irene Villa-Machío

Grupo Biología Evolutiva de Plantas
Real Jardín Botánico, RJB-CSIC

BIOLOGÍA MOLECULAR EN EL ESTUDIO DE LA BIODIVERSIDAD VEGETAL

La incorporación de las técnicas de biología molecular en las disciplinas de estudio de la biodiversidad vegetal está indisolublemente ligada a la incorporación de la idea de que el estudio de la diversidad vegetal debe reflejar los procesos evolutivos que la han generado.

Como consecuencia del desarrollo de los enfoques evolutivos en el estudio de la biodiversidad, a inicios de los años 80 del pasado siglo se produce la convergencia de cuatro corrientes hasta entonces aisladas en distintas áreas de conocimiento. En primer lugar, la difusión en los estudios vegetales de las técnicas basadas en la caracterización electroforéticas de isoenzimas producto de la llamada moderna síntesis evolutiva y habi-

tualmente circunscrita a estudios de genética de poblaciones, mejora vegetal, especiación y microevolución. En segundo lugar, la progresiva traslación a los estudios filogenéticos, hasta entonces basados en la cladística morfológica, de los datos cada vez más accesibles, de secuencias de genes de regiones del ADN cloroplástico y nuclear. Una tercera corriente recogía el inmenso conocimiento acumulado acerca de la variación citogenética observada en las plantas vasculares (números de cromosomas, niveles de ploidía, aneuploidía, alopoliploidía) que confirmó la importancia de fenómenos como la evolución reticulada en la mayoría los linajes vegetales. Y, finalmente, la identificación mediante estudios funcionales y experimentales en especies modelo de la base genética del desarrollo y su extensión a otras especies (enfoques evo-devo), que permitió comenzar a entender las bases de la impresionante diversificación morfológica de las

angiospermas. Hasta entonces, solo los datos micro- y macromorfológicos, y en mucha menor medida los fitoquímicos (metabolitos secundarios) y cromosómicos (números de cromosomas), sustentaban la base de las evidencias científicas utilizadas en las disciplinas centradas en el estudio de la biodiversidad vegetal. Igualmente, la Sistemática, es decir la clasificación de las especies de

plantas en grupos, carecía de evidencia científica sólida que trascendiera lo puramente descriptivo. Con la síntesis de las cuatro corrientes descritas llegó la consolidación de la Sistemática filogenética y su plasmación en una clasificación y una nomenclatura que reflejara la evolución de los linajes vegetales, así como el establecimiento de un marco filogenético donde encajan todas las especies conocidas. Este hito es particularmente significativo en

el caso de las 369.000 especies de plantas con flores que se agrupan en 416 familias.

PECULIARIDADES DE LAS PLANTAS EN RELACIÓN CON OTROS SERES VIVOS

Los linajes que constituyen el mundo vegetal (algas, briófitos, pteridófitos, gimnospermas y angiospermas) presentan una serie de diferencias en relación con otros grupos de seres vivos como los animales, hongos y procariontes que requieren que tanto los diseños experimentales de los estudios, como las técnicas de análisis estén dotados de rasgos propios. Mencionaremos tres de los más importantes.

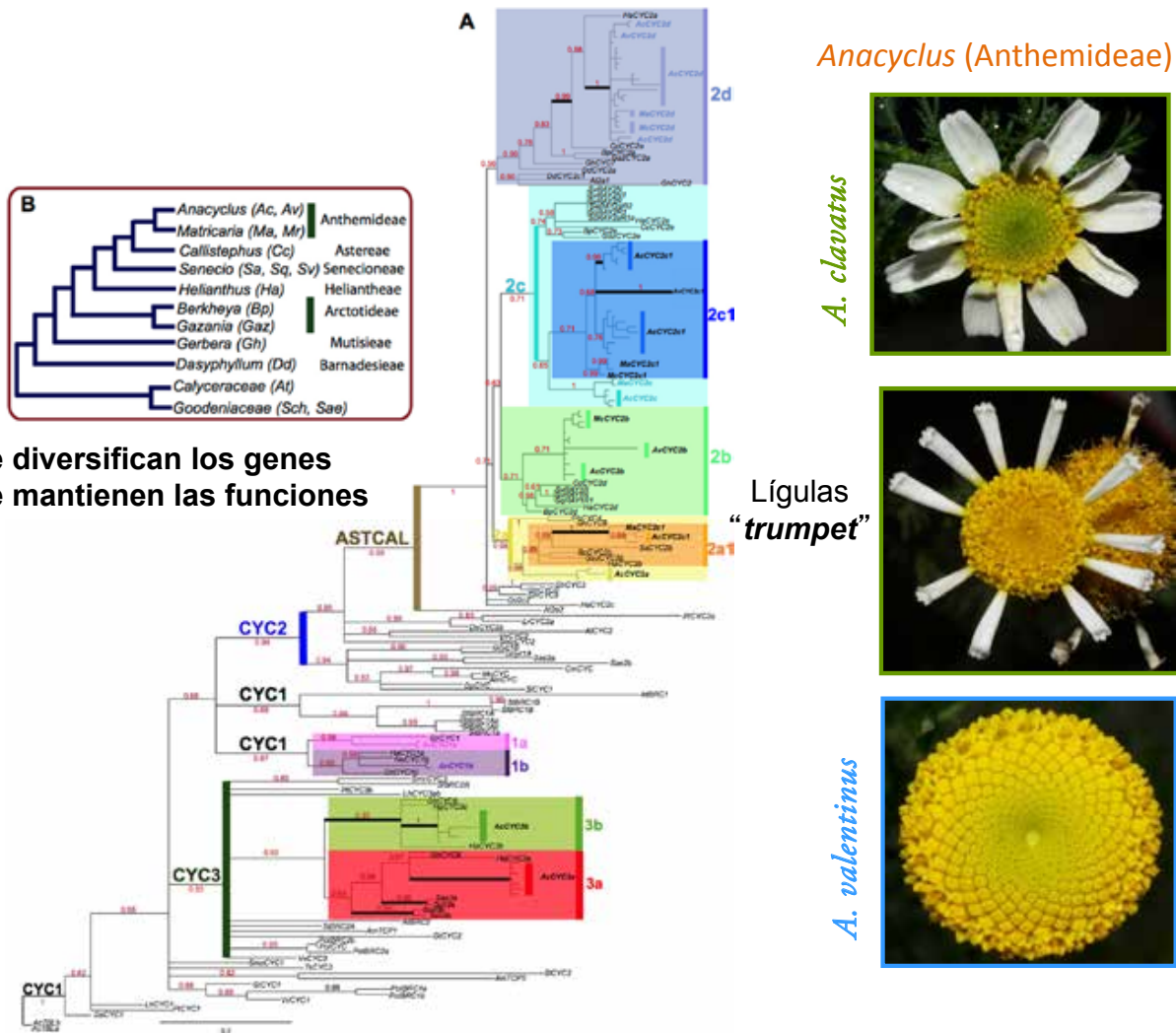
1. La existencia de los cloroplastos, con su correspondiente genoma dotado de un tipo de herencia (mayoritariamente materna) distinta de la biparental del genoma



Figura 1

Paris japonica (*Melanthiaceae*), una monocotiledónea con un tamaño de genoma de $1C=127,4$ pg, constituye el ser vivo con el mayor tamaño relativo de genoma conocido, aproximadamente 60.000 veces el tamaño del genoma haploide humano.

Créditos: *By alpsdake - Own work, Public Domain.*



- Se diversifican los genes
- Se mantienen las funciones

Figura 2

Evolución del grupo de genes *Cycloidea* en las Compuestas, mostrando que los parálogos *CYC2* conservan su función en la determinación de la simetría floral en flores en tubo y en lígula dentro del género *Anacyclus* (Bello et al. 2017 *Front. Plant Sci.* 8: 589.) <https://doi.org/10.3389/fpls.2017.00589>.

Créditos: Javier Fuertes Aguilar & M. Angélica Bello.

del núcleo, influye de manera drástica en la incongruencia filogenética frecuentemente observada al contrastar estudios basados en genomas cloroplásticos y genomas nucleares. Esta singularidad de los genomas plastidiales se exagera cuando se estudia la evolución en las más de 4.500 especies parásitas de plantas, cuyas funciones fotosintéticas perdidas distorsionan las comparaciones con los cloroplastos de plantas fotosintéticas.

2. La evidencia acumulada acerca de la prevalencia de la hibridación entre especies de plantas vasculares como generadora de diversidad genética, funcional y morfológica. Se conocen innumerables casos de especies de plantas bien documentados molecularmente cuyo ori-

gen se encuentra en un suceso de hibridación entre dos especies preexistentes.

3. La enorme capacidad de aumentar el tamaño del genoma, a través de dos mecanismos: por un lado por su facultad de incorporar y mantener en su genoma elementos transponibles de diversos tipos (hasta un 85% del tamaño del genoma), manteniendo este su funcionalidad y dotando a este de una plasticidad infrecuente en otros grupos de seres vivos. Y por otro el extendido fenómeno en las plantas vasculares de la poliploidía; es decir, la duplicación del número de copias del genoma presentes en el núcleo de la progenie (WGD, *whole genome duplication*) sin que medie reducción meiótica

del número de cromosomas. Este extendido fenómeno constituye la base de la evolución del tamaño genoma en la mayoría de las plantas vasculares (por ejemplo, se estima que más del 50% de las angiospermas proviene de un ancestro próximo con un cambio en el nivel de ploidía). La combinación de los dos mecanismos descritos ha llevado a que las plantas sean los organismos con los mayores tamaños de genomas conocidos en cualquier ser vivo.

Una consecuencia de las citadas peculiaridades se refleja en que los estudios moleculares y técnicas de plantas incorporan ciertos rasgos característicos. Por ejemplo, a menudo es difícil distinguir genes que se han diversificado mediante la divergencia de los organismos (genes ortólogos) de aquellos que han resultado de los fenómenos de duplicación total (poliploidía) o parcial del genoma (genes parálogos). Otro aspecto característico de los estudios moleculares en plantas es que la representación de los patrones evolutivos está mejor representada por gráficos de tipo “red” o por árboles filogenéticos cuya topología incorpora reticulación (“neighbor-net”) que por árboles con ramas. De esta forma, es posible reconciliar los procesos de diversificación e hibridación.

A pesar de todas las dificultades a día de hoy, y gracias a la incorporación de herramientas analíticas basadas en Máxima Verosimilitud e Inferencia Bayesiana, ha sido posible el asentamiento de un conjunto de conocimientos, en donde el origen de la extraordinaria variabilidad morfológica, ecológica y funcional observada en los grupos de plantas terrestres, que en palabras de Darwin constituía el “abominable misterio de la diversificación de las plantas con flores”, no solo puede ser explicada sino que es posible incluso datarla congruentemente con el registro fósil conocido.

REGIONES DEL GENOMA USADAS EN LA INFERENCIA DE LAS RELACIONES FILOGENÉTICAS ENTRE PLANTAS

La mayoría de los estudios realizados hasta la generalización de las técnicas de secuenciación masiva (NGS) están basados en secuencias moleculares tanto de genes como de regiones no codificantes, tales como espaciadores del genoma plastidial. Al contrario que en los animales, donde el genoma mitocondrial ha constituido la base de los estudios, en las plantas el genoma mitocondrial apenas presenta variabilidad. La naturaleza no recombinante y la haploidía del ADN cloroplástico hacen que sea adecuada para estimar la historia evolutiva tanto entre linajes alejados (familias, órdenes) como en poblaciones de plantas, aun cuando su tasa de mutación sea menor que la del núcleo. Sin embargo, su herencia uniparental hace que las genealogías de regiones plastidiales tengan limitaciones para poder reflejar la

filogenia de las especies ya que solo reflejan la historia genealógica de uno de sus progenitores. Para compensar esta desventaja ha sido frecuente el uso de secuencias de marcadores nucleares, sobre todo del ADN ribosómico nuclear que en plantas tiene dos espaciadores transcritos internos (*ITS*). Algunos de los motivos que justifican el uso generalizado de marcadores nucleares son, además de su herencia biparental, la disponibilidad de varios sets de cebadores universales para PCR útiles en un gran número de grupos taxonómicos, su estructura multicopia que facilita la amplificación y el hecho de que, debido a su tasa de mutación, son marcadores moleculares adecuados para estudios evolutivos a nivel de especie. Al contrario que en gran parte de los linajes animales, la búsqueda de una “bala de plata” que pudiera servir de código de barras genético (*DNA barcoding*) como la secuencia del gen del COI mitocondrial, que pudiera generalizarse como herramienta diagnóstica en la identificación de especies vegetales, ha constituido un relativo fracaso. No existe un consenso y actualmente se acepta que al menos dos genes son necesarios para conseguir una identificación aproximada de una especie vegetal (*psbA-trnH*, *ITS*, *ndhF*). Los genes varían según los grupos y en todo caso deben hacer frente a los efectos anteriormente descritos provocados por la evolución reticulada, la evolución concertada sesgada tras la frecuente hibridación y la incongruencia entre los genomas nucleares y cloroplásticos de las especies.

TÉCNICAS DE SECUENCIACIÓN MASIVA (NGS) Y BIODIVERSIDAD VEGETAL

En 2010 había apenas 25 genomas secuenciados de plantas; siete años después son 250 y para 2025 existe la fiable predicción y el compromiso de que se habrán secuenciado 10.000 genomas.

Desde 2010 se han impuesto las técnicas de secuenciación masiva de nueva generación (*Next Generation Sequencing*) cuya plétora de variaciones se aplica desde el ensamblaje completo de genomas hasta la identificación de variabilidad individual de SNP (*single nucleotide polymorphisms*). El ideal de poder comparar entre sí una gran cantidad de genomas completos está próximo a cumplirse, aunque las limitaciones bioinformáticas ha hecho que una parte considerable de los estudios de NGS en plantas están centrados en el estudio de la variabilidad molecular dentro de una especie en particular o grupos de especies relacionadas. Existen diferentes técnicas para la generación de datos NGS aplicables a estudios filogenéticos y filogeográficos. La más sencilla, y muy utilizada en la detección de parálogos, es la secuenciación de amplicones, que consiste en secuenciar mediante plataformas de NGS productos de PCR que han sido previamente amplificados. Otro método es el denominado “*target enrichment*”, también llamado

“sequence capture”, que implica una captura selectiva de regiones del genoma previas a la secuenciación NGS. Esta técnica requiere un conocimiento genómico previo para diseñar las sondas encargadas de marcar el ADN, pero tiene la ventaja de poder trabajar con ADN de baja calidad (por ejemplo procedente de ejemplares de herbario). Y por último, los métodos de secuenciación de ADN “asociados a sitios de restricción” o RADseq, en los que se secuencian fragmentos de ADN que previamente han sido digeridos por enzimas de restricción que cortan en regiones específicas del genoma. Son técnicas que generan una representación reducida del genoma. RADseq comprende técnicas como RAD (*restriction site associated DNA*), ddRAD (*double digestion RAD*) o GBS (*Genotyping-by-sequencing*) entre otros, que se diferencian principalmente en el número de enzimas empleadas para fragmentar el genoma y por la existencia de un paso adicional para la selección del tamaño de los fragmentos.

HITOS LOGRADOS E IMPACTO DE LAS TÉCNICAS MOLECULARES EN VEGETALES

A pesar de todas las dificultades, la aplicación de las técnicas moleculares ha permitido la consecución de una serie de hitos insospechados apenas hace unos años. En primer lugar, la construcción de un árbol fiable de la vida de prácticamente todas las plantas conocidas sobre la Tierra, en cuyo marco la identificación de especies nuevas descubiertas o de difícil adscripción es prácticamente inmediata. El marco evolutivo que ha permitido reconstruir la evolución de los rasgos morfológicos incluso con una reconstrucción de la flor ancestral y la datación molecular donde ha sido posible incorporar el registro paleobotánico de fósiles vegetales.

La combinación de técnicas de estudios de la expresión génica con la de los árboles filogenéticos ha permitido la identificación de genes responsables de rasgos de importancia evolutiva, como por ejemplo los genes responsables de la simetría floral (genes *TCP*), que determinan selectivamente el tipo de polinizadores que visita cada especie, asegurando su aislamiento reproductivo y la ocupación de distintos nichos mediante los procesos de especiación.

La caracterización mediante técnicas moleculares ha generalizado la identificación inequívoca de mate-

rial vegetal aplicada a múltiples áreas de la ciencia (ecología, identificación forense, farmacognosia), la industria (control de calidad de alimentos) y la agricultura (identificación de semillas, malas hierbas). El acceso de los métodos moleculares ha permitido incluso la puesta en valor de la herencia histórica albergada en las colecciones de historia natural (herbarios). Los métodos tradicionales de la botánica han dado paso a su uso como fuente de datos moleculares.

RETOS DEL FUTURO

El futuro del uso de las técnicas moleculares aplicadas al estudio de la biodiversidad vegetal coincide con dos retos planteados en la comunidad científica. Por un lado, la creciente facilidad en las adquisición de datos genómicos que frecuentemente no van acompañados de herramientas bioinformáticas accesibles que sostengan el rigor requerido en los estudios comparados. Y, por otro lado, la trágica evidencia de que justo cuando en la historia de las ciencias de la biodiversidad se accede a los datos necesarios y suficientes para su estudio, interpretación y utilización por el ser humano, se asiste a una de las mayores crisis de extinción en la historia de seres vivos causada por la destrucción del medio natural. En las condiciones actuales, e incluso en ausencia de consideración del cambio climático global, en menos de un siglo se habrán extinguido en su hábitat más de un tercio de las especies de plantas con flores conocidas en la actualidad.

PARA LEER MÁS

- Angiosperm Phylogeny Group; 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV, *Bot J Linn Soc*, 181 : 1–20.
- Cronk Q C B, Bateman RM, Hawkins JA. (eds); 2002. Developmental Genetics and Plant Evolution. CRC Press. Boca Raton, 453 p.
- Grant V; 1981. Plant speciation, 2nd ed. Columbia University Press. Nueva York, 432 p.
- Kew Royal Botanic Gardens; 2017. The State of the World's Plants Report 2017. www.stateoftheworldsplants.com.
- Soltis DE, PS Soltis, JJ Doyle; 1998. Molecular Systematics of Plants II: DNA Sequencing. Springer Science & Business Media, Berlin, 574 p.
- Spencer V, Kim M; 2017. ‘Re’CYC’ling molecular regulators in the evolution and development of flower symmetry. *Semin Cell Dev Biol*. <http://dx.doi.org/10.1016/j.semcdb.2017.08.052>.