

Aplicación de la genómica y la proteómica al mejoramiento genético de los cultivos

Iván Galindo Castro*

Laboratorio de Genómica y Proteómica. Centro de Biotecnología. Fundación Instituto Internacional de Estudios Avanzados IDEA. Carretera Nacional Hoyo de la Puerta, Miranda

RESUMEN

Las herramientas biotecnológicas y el desarrollo de las llamadas disciplinas “ómicas” suponen una reducción importante en los tiempos para la generación y preselección de materiales, para predecir los cruces más exitosos, acelerar el desarrollo de nuevos cultivares, potenciar su explotación genética y ofrecerlos al consumidor. La gran mayoría de los caracteres agronómicos se comportan de una manera compleja, como consecuencia de la intervención desigual de múltiples genes y el concurso de muy variados componentes ambientales. Sobre este comportamiento, que dio origen a lo que hoy conocemos como genética cuantitativa, se sientan las bases del mejoramiento genético moderno a través del estudio del aporte cuantitativo de dichos caracteres (denominados QTL). La genética cuantitativa clásica ha tenido cierto éxito en cubrir las expectativas de los mejoradores genéticos, sin embargo existen muy pocos descriptores estadísticos para explicar la complejidad biológica subyacente al proceso de mejoramiento, lo cual siembra muchas dudas sobre su aplicabilidad real, en función de las exigencias cada vez más fuertes del desarrollo agrícola moderno. Las tecnologías “ómicas” están generando una enorme cantidad de información a partir de múltiples y muy variados organismos, creando la necesidad de una “biología sistémica” que permita integrar multidimensionalmente esa información biológica a través de redes dinámicas de interacciones moleculares. La biología sistémica apunta entonces hacia la síntesis del conjunto de datos complejos surgidos a partir de varias jerarquías genómicas (genoma, transcriptoma, proteoma, metaboloma) en modelos matemáticos útiles para explicar de manera cuantitativa, la biología funcional en términos de cómo interactúan las unidades funcionales codificadas en el genoma. Esta información preliminar puede ser ampliamente informativa para incluir o excluir genes candidatos potenciales. En el futuro inmediato podrían surgir novedosos algoritmos capaces de leer las secuencias reguladoras de los genes y deducir, a partir de ellas, los perfiles de expresión. Aunque la genómica funcional

*Autor de correspondencia: Iván Galindo

E-mail: igoalindocastro@gmail.com

puede ofrecer vías más expeditas en ese sentido, es importante tomar en cuenta las particularidades del sistema genotipo-fenotipo-ambiente para poder alcanzar programas de mejoramiento exitosos en el desarrollo de los cultivos del futuro. En Venezuela todavía es muy incipiente el uso de la genómica y la proteómica en el mejoramiento genético, sin embargo resultados preliminares han permitido identificar genes candidatos asociados a resistencia a *Magnaporthe oryzae*, agente causal de la piricularia en el arroz.

Palabras clave: Biología Sistémica de los cultivos, modelaje de cultivos, genómica funcional, mapeo de QTL, micromatrices de ADN, electroforesis en geles bidimensionales.

Application of genomics and proteomics to crops genetic breeding

ABSTRACT

The advent of the biotechnological tools and the emergence of the “omics” disciplines have promised a substantial time reduction in breeding and selection of higher productive materials, predicting the most successful crosses, accelerating the development of new cultivars, and enhancing their genetic potential, to finally reach the consumer tables. Most agronomic characters express in a complex manner, as a consequence of an unequal intervention of multiple genes influenced by diverse environmental effects. On the basis of this behavior, which originated what we know as quantitative genetics, modern genetic breeding is funded through studying the quantitative outcome of these characters, known as QTL.

Classic quantitative genetics have had a relative success covering the expectations of genetic breeders; nevertheless there are just a few statistical descriptors that try to explain the subjacent biological complexity underlying the genetic improvement process, shading doubts about the potential applicability of this approach to solve the more exigent needs of the modern agriculture development. The “omics” techniques are producing a huge amount of biological information from multiple and very different organisms, requiring a systemic approach that allows the integration of that information in a multidimensional way through the construction of dynamic networks of molecular interactions. Hence, systems biology aims to synthesize the biological data sets coming from different genomics scale jerarquization (genome, transcriptome, proteome, metabolome) into reliable mathematical models capable of explaining, in quantitative terms, how the functional units coded in the genome interact with each other. This preliminary information can be very informative in terms of including or excluding potential candidate genes. In the near future, novel algorithms capable of explaining how gene regulatory sequences work will be available. Even though functional genomics may offer even more expedite ways on this aspects, it is important taking into account the singularities of each genotype-phenotype-ambient system to develop the crops of the future by more successful breeding programs. In Venezuela, the application of genomics

and proteomics on genetic breeding is still incipient, although preliminary results have allowed the identification of potential candidate rice genes associated with resistance to *Magnapothae orizae*, the etiological agent of the rice blast disease.

Key words: Crop systems biology, crop modeling, functional genomics, QTL mapping, DNA microarrays, 2-D gel electrophoresis.

El mejoramiento genético en la agricultura moderna

El mejoramiento genético de los cultivos se ha sustentado tradicionalmente en la habilidad del mejorador para observar con mucha precisión el comportamiento de los diferentes caracteres agronómicos de los cultivos, para así seleccionar y cruzar los materiales más promisorios a partir de las poblaciones segregantes de acuerdo con una vasta experiencia acumulada por décadas.

El desarrollo reciente de la biotecnología agrícola ha causado una gran conmoción en las expectativas de producción de nuevos materiales comerciales, así las herramientas biotecnológicas y el desarrollo de las llamadas disciplinas “ómicas” suponen una reducción importante en los tiempos para la generación y preselección de materiales, para predecir los cruces más exitosos, acelerar el desarrollo de nuevos cultivares, potenciar su explotación genética y ofrecerlos al consumidor. En este sentido, los retos más importantes que desafían al mejoramiento genético actual se centran fundamentalmente en i) poder reducir en forma importante los tiempos para la generación y preselección de materiales, ii) predecir con cierta precisión los cruces más exitosos, iii) generar más rápidamente los materiales élite, y iv) desarrollar al máximo el potencial genético de los materiales a ser colocados en el mercado.

La expresión de la gran mayoría de los caracteres fenotípicos se comporta de una manera muy compleja, como consecuencia de la intervención desigual de múltiples genes y el concurso de muy variados componentes ambientales, incluyendo el propio manejo agronómico del cultivo. La modulación de la expresión genética por distintos efectos ambientes representa un largo camino para lograr explicar las conexiones entre los múltiples genes y la expresión de fenotipos como el rendimiento de grano, por ejemplo, que para la mayoría de los cultivos es crucial para la agricultura.

De tal forma que la manifestación de un fenotipo particular no es más que la suma de los efectos genéticos y ambientales, más la propia interacción genotipo-ambiente. Sobre este comportamiento, que dio origen a lo que hoy conocemos como genética cuantitativa, se sientan las bases del mejoramiento genético moderno a través del estudio del aporte cuantitativo de dichos caracteres (denominados QTL).

Los marcadores moleculares y la identificación de QTL

La genética cuantitativa clásica ha tenido mucho éxito en cubrir buena parte de las expectativas de los mejoradores genéticos; sin embargo, existen muy pocos descriptores estadísticos para explicar la complejidad biológica subyacente

al proceso de mejoramiento, lo cual siembra muchas dudas sobre su aplicabilidad real en función de las exigencias cada vez más fuertes del desarrollo agrícola moderno. Adicionalmente, existe la creencia que el conocimiento de las secuencias de los genomas completos de las plantas reducirá la importancia de la genética cuantitativa en el futuro cercano, ya que permitirá movernos desde la incertidumbre estadística hacia el marco del conocimiento de la funcionalidad de los propios genes. Sin embargo, el escenario más probable es que los modelos genéticos existentes, creados a partir de la genética cuantitativa clásica, puedan ser adaptados para crear nuevos modelos que conformen la plataforma postgenómica con la cual acceder al manejo de la información masiva que se genera actualmente.

Típicamente, el análisis de los QTL depende de la realización de los cruces entre dos genotipos individuales que difieren en el carácter de estudio, a partir de la cual se deriva una población segregante consistente de una gran población de líneas individuales. Estas líneas individuales son genotipificadas usando marcadores moleculares y fenotipificadas a través de experimentos para estudiar los caracteres de interés y entonces establecer cualquier asociación cercana entre los marcadores y el carácter fenotípico estudiado. Este análisis de caracteres es un atajo para conectar una región particular del genoma con el fenotipo sin la necesidad de información adicional sobre los pasos intermedios a nivel de los transcritos, proteínas o metabolitos.

Estos marcadores moleculares están basados en pequeños cambios en las secuencias del ADN (deleciones o adiciones de bases o sustituciones) o en la variación del número de repeticiones de pequeñas secuencias, la mayoría de ellas situadas en las regiones no codificantes del genoma. Los marcadores ocurren de manera natural, y son abundantes en la mayoría de las especies y heredadas de manera mendeliana. Los marcadores moleculares son por tanto muy útiles para construir mapas saturados a lo largo de los cromosomas, característica esencial para localizar posteriormente las regiones del genoma que confieren la variación del carácter cuantitativo (QTL).

Uno de las desventajas del análisis de QTL mediante cruces biparentales es el experimento en si mismo, donde se utilizan sólo dos parentales por vez y por tanto es muy difícil representar toda la variación genética del germoplasma bajo estudio.

A pesar de que los mejoradores pueden usar directamente los resultados del análisis de QTL mediante el mejoramiento asistido por marcadores moleculares, todavía esto sigue siendo una suerte de caja negra. Diversos estudios han demostrado que la combinación de QTL y los análisis bioquímicos o fisiológicos han generado nuevas consideraciones en las relaciones causales entre los caracteres, que podrían ser difíciles de obtener mediante aproximaciones convencionales. Estos análisis metabólicos resultaron muy útiles para caracterizar las bases reguladoras de la variación general de los caracteres cuantitativos, proveyendo herramientas comparativas fundamentales para caracterizar los QTL (Walsh, 2001). El origen de la variación de la expresión genética se puede derivar de un polimorfismo localizado físicamente en las cercanías de un gen o a partir de un sitio distante en el genoma. Afortunadamente, el primero parece tener, con mayor frecuencia, los

efectos fenotípicos más importantes (Hughes *et al.*, 2006; Keurentjes *et al.*, 2007)

El mejoramiento genético en la era postgenómica

La búsqueda de genes candidatos, es decir, aquellos genes que se encuentra en el entorno de una región del genoma asociada con la expresión de un fenotipo, se ha estado estableciendo con la utilización de marcadores moleculares para acotar esa región. Sin embargo con el uso de las micromatrices de ADN y el estudio comparativo de los perfiles de expresión de los ARN mensajeros (transcriptómica) con la ayuda del análisis de imágenes y la bioinformática, se puede lograr también la identificación de los genes potencialmente asociados a la expresión de un fenotipo particular.

Las tecnologías “ómicas” están generando una enorme cantidad de información a partir de múltiples y muy variados organismos, creando la necesidad de una “biología sistémica” que permita integrar multidimensionalmente esa información biológica a través de redes dinámicas de interacciones moleculares. La biología sistémica, propuesta por Kitano (2002), apunta entonces hacia la síntesis del complejo conjunto de datos surgidos a partir de varias plataformas jerárquicas de información (genoma, transcriptoma, proteoma, metaboloma) en modelos matemáticos útiles para explicar de manera cuantitativa, la biología funcional en términos de cómo interactúan las unidades de información codificadas en el genoma. De manera similar, la biología sistémica aplicada a las plantas, tal como la definió Minorsky (2003), plantea el uso de algoritmos computacionales para predecir el comportamiento de una célula vegetal a partir de los datos genómicos y proteómicos (Hammer *et al.*, 2004).

En general, la secuencia del ADN en si misma puede proveer pistas muy importantes para la identificación de potenciales genes candidatos, aunque también el creciente auge de la proteómica ha generado un catálogo extensivo de motivos de proteínas, ofreciendo la posibilidad de hacer algunas deducciones, todavía muy tímidas, acerca de las funciones de los marcos abiertos de lectura, tales como asociaciones a membrana, mecanismos de transporte, tráfico hacia organelos particulares, etc. Esta información preliminar puede ser ampliamente informativa para incluir o excluir potenciales genes candidatos. En el futuro inmediato, seguramente surgirán novedosos algoritmos capaces de leer las secuencias reguladoras de los genes y deducir, a partir de ellas, sus perfiles de expresión bajo diversas condiciones y su conexión funcional con otros genes.

De las disciplinas “ómicas” a la biología sistémica de los cultivos

Lo que podríamos llamar la zona caliente de investigación en estos momentos es la formulación de algoritmos matemáticos para modelar la entramada red del “interactoma” en los principales cultivos. En estos modelos se debe reflejar, no sólo la interconexión entre los actores y los procesos celulares, sino también su grado de interacción espacio-temporal. Se debe modelar en forma dinámica la reacción de todo el sistema ante las perturbaciones desde el entorno subcelular

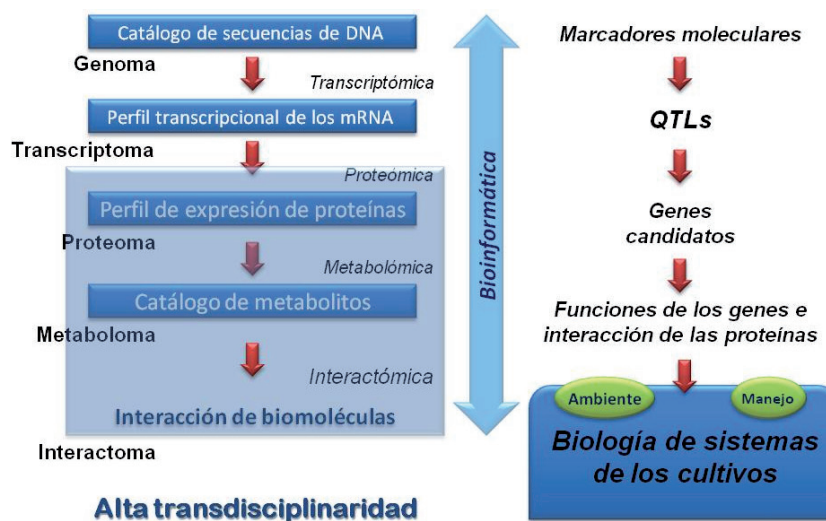


Figura 1. Aplicación de las disciplinas “ómicas” al mejoramiento de los cultivos. La estructura jerárquica de las distintas plataformas de información biológica (transcriptómica, proteómica, metabolómica) puede ser analizada con el uso de complejos algoritmos matemáticos a fin de generar modelos predictivos dinámicos del “interactoma” y de esta forma explotar más eficientemente el potencial genético, tomando en cuenta tanto su interacción con el ambiente como el propio manejo agronómico del cultivo. Esta visión holística de la interacción funcional entre los genes (y proteínas) es que lo que se ha dado en llamar la biología sistémica de los cultivos.

hasta el ambiental, reacción que puede ser totalmente insospechada e incluso contraria a la propia intuición del investigador (Figura 1).

En las últimas décadas, el impacto más importante en el mejoramiento de la mayoría de los cereales sobre el potencial genético de los cultivos se ha alcanzado mediante la explotación de unos pocos genes (como los del enanismo e insensibilidad al fotoperíodo) capaces de redistribuir la partición de masa a favor de los órganos cosechados del cultivo. Esto permitió alcanzar incrementos importantes en el rendimiento de los cultivos, especialmente durante la llamada “revolución verde”. Sin embargo, el mejoramiento empírico, aún utilizando mejoramiento asistido por marcadores (MAS) para manipular y mejorar aún más dicha partición todavía no parece arrojar valores cercanos a sus límites teóricos (Wollenweber *et al.*, 2005).

Para poder alcanzar aportes significativos semejantes a los de la revolución verde tendremos que trabajar de manera simultánea con muchos genes, cuya acción concertada permita manifestar el fenotipo deseado de una manera estable. Para ello, indudablemente requeriremos enfoques totalmente transdisciplinarios y sistémicos con los cuales abordar los grandes problemas de disponibilidad de alimentos, recursos hídricos y energéticos de manera sustentable y más armónica con el ambiente, considerando los cambios globales que se suscitan actualmente en nuestro planeta. Aunque la genómica funcional puede ofrecer vías más expedi-

tas en ese sentido, es importante tomar en cuenta las particularidades del sistema genotipo-fenotipo-ambiente para poder alcanzar programas de mejoramiento exitosos en el desarrollo de los cultivos del futuro.

La genómica funcional apunta al descubrimiento de las funciones de los genes, básicamente a través del estudio masivo del transcriptoma, el proteoma y el metaboloma, combinado con la bioinformática, para el manejo y análisis de la data y también mediante el análisis funcional de genotipos mutantes. Es decir, este nuevo paradigma metodológico se fundamenta en que estas disciplinas ómicas en lugar de estudiar un gen a la vez y a partir de experimentos individuales, permiten predecir la función de un enorme conjunto de genes a partir de las redes de interacción que regulan los complejos procesos biológicos (interactoma). De esta forma, la biología de sistemas surge como una necesidad para sintetizar ese enorme conjunto de datos provenientes de los estudios ómicos en modelos matemáticos y complejos algoritmos que permitan dilucidar cómo responden las redes de interacción biológica en las células ante los cambios o perturbaciones del micro y macro ambientes (Thomas, 2007)

Por tanto, la biología sistémica de los cultivos emerge como una nueva concepción metodológica multi y transdisciplinaria con la cual abordar el modelaje de caracteres fenotípicos complejos y con relevancia en la producción de alimentos y de generación de energía, vinculando la información entre las ómicas derivadas de la comprensión de la red de interconexiones de los procesos bioquímicos y fisiológicos (Peleman Rouppe, 2003; Yin *et al.*, 2004; Yuan *et al.*, 2008). Para ello es necesario mapear tanto los niveles de organización como los sistemas de interacción que participan en los diferentes procesos que modulan la expresión fenotípica (Seo *et al.*, 2011).

Claramente, se requiere precisar las diferentes escalas temporales, espaciales y estructurales para los distintos componentes, rutas y procesos del modelo, para desarrollar una disciplina intensiva en el uso de la bioinformática. Tales modelos podrán permitir el estudio *in silico* de la respuesta del cultivo a través de la modulación de la expresión genética frente a los distintos escenarios ambientales, conformando una herramienta muy poderosa para simular el comportamiento de los caracteres complejos de los cultivos durante del mejoramiento genético.

En Venezuela todavía es muy incipiente el uso de la genómica y la proteómica en el mejoramiento genético; sin embargo, resultados preliminares han permitido identificar genes candidatos para resistencia a *Magnaporthe oryzae*, agente causal de la piricularia en el arroz, a través del uso de estas disciplinas ómicas (Franco, 2011). Efectivamente, al comparar los perfiles de ARNm y de expresión de proteínas entre dos isolíneas de *Oryza sativa* (una resistente y otra susceptible) al ser infectadas con un aislado de *M. oryzae*, se logró identificar 187 genes candidatos, de los cuales aproximadamente el 50% presentó una función totalmente desconocida. La importancia de estos genes sobre la definición del fenotipo resistente apenas comienza a ser evaluado, pero tal como se observa en la Figura 2, al vincular las regiones cromosomales donde se localizan estos genes

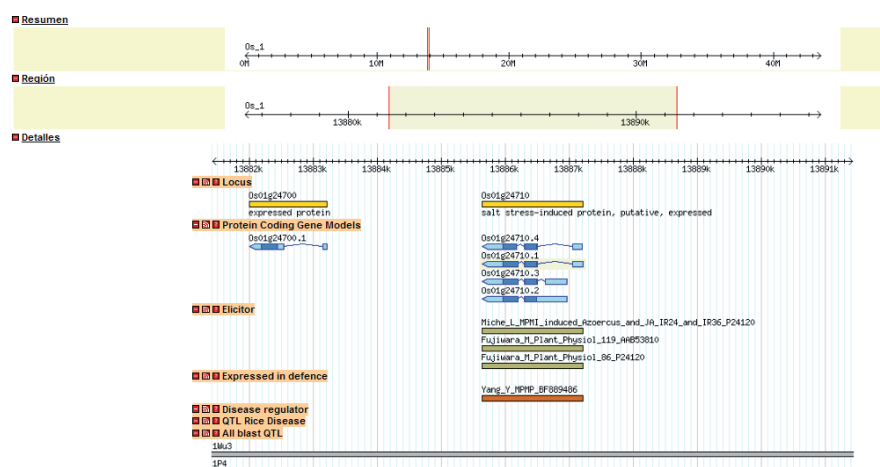


Figura 2. Asociación de genes candidatos con QTL descritos para estrés biótico. Una vez vez identificados los genes candidatos, mediante perfiles de transcripción de mRNA (transcriptómica) o de expresión de proteínas (proteómica), se puede validar su contribución a la expresión del fenotipo mediante su colocación con QTL ya descritos en la literatura o en bases de datos como OriGeneDB (http://orygenesdb.cirad.fr/cgi-bin/gbrowse/odb_japonica/#search). En el ejemplo, se visualiza el caso del gen LOC_Os01g24710.1 correspondiente con la descripción de Salt stress-induced protein, putative, expressed.

candidatos con la información correspondiente a los QTL descritos para estrés biótico, las coincidencias permitieron ratificar la pertinencia de la identificación realizada mediante la comparación de los transcriptomas y proteomas de los materiales estudiados.

CONCLUSIONES

Con el avance de las nuevas disciplinas ómicas y su aplicación al estudio de las características fenotípicas de los cultivos se establece la posibilidad cierta de mejorar sus rendimientos potenciales con base en el incremento del uso eficiente de los recursos genéticos y la resistencia a las plagas y enfermedades. La biología sistémica de los cultivos surge como un nuevo enfoque que permite conectar la genómica funcional y el modelaje de los cultivos para asistir a los programas de mejoramiento genético de los principales rubros agrícolas, particularmente de los cereales. Con esta nueva visión hacia una disciplina con un alto grado de intervención de la bioinformática se podrá modelar *in silico* la respuesta fina de los cultivos bajo diversos escenarios micro y macroambientales, convirtiéndose en una poderosa herramienta para afianzar el mejoramiento genético de los cultivos. La revolución de las ómicas está creando una nueva revolución en el campo del mejoramiento de las plantas que no puede ser ignorado. En Venezuela es neces-

rio potenciar el manejo de estas novedosas disciplinas y ponerlas al alcance de los mejoradores. La creación de capacidades en infraestructura y talento humano, así como, la promoción de este nuevo enfoque es fundamentalmente importante para contribuir a lograr nuevos estándares de productividad y sostenibilidad en materia agrícola, para colocarlos a la disposición de nuestra sociedad.

REFERENCIAS

- Franco, M. 2011. Evaluación de los mecanismos de resistencia a *Pyricularia grisea* en arroz (*Oryza sativa*) a través de perfiles masivos de expresión de ARNm y proteínas. Tesis Doctoral en Biología Celular. Escuela de Biología. Facultad de Ciencias, UCV. Caracas. 147 p.
- Hammer, G.; T. Sinclair; S. Chapman; E. Van Oosterom. 2004. On systems thinking, systems biology, and the in silico plants. *Plant Physiology* 134: 909-911.
- Hughes, K.; J. Ayroles; M. Reedy; J. Drnevich; K. Rowe; E. Ruedi; C. Caceres; K. Paige. 2006. Segregating variation in the transcriptome: cis regulation and additivity of effects. *Genetics* 173: 1347-1355
- Keurentjes, J.; J. Fu; I. Terpstra; J. García; J. Ackerveken; L. Snoek; A. Peeters; D. Vreugdenhil; M. Koornneef; R. Jansen. 2007. Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci USA* 104: 1708-1713.
- Kitano, H. 2002. Systems Biology: A brief overview. *Science* 295:1662-1664.
- Minorsky, P. 2003. Achieving the in silico plant. Systems biology and the future of plant biological research. *Plant Physiology* 132: 404-409.
- Peleman, J.; J. Rouppe. 2003 Breeding by Design. *Trends in Plant Science* 8: 330-334.
- Seo Y.S.; M. Chern; L.E. Bartley; M. Han; K.H. Jung. 2011 Towards establishment of a rice stress response interactome. *PLoS Genet.* 7: 1-20.
- Thomas, H. 2007. Systems biology and the biology of systems: how, if at all, are they related? *New Phytologist* 177: 11-15.
- Walsh, B. 2001 Quantitative genetics in the age of genomics. *Theoretical and Population Biology* 59: 175-178.
- Wollenweber, B.; J. Porter; T. Lubberstedt. 2005 Need for multidisciplinary research towards a second Green Revolution. *Current Opinion in Plant Biology* 8: 337-341.
- Yin, X.; P. Struik; M. Kropff 2004. Roles of crop physiology in predicting genotype-phenotype relationships. *Trends in Plant Sciences* 9: 426-432.
- Yuan, J.; D. Galbraith; S. Dai; P. Griffin; N. Stewart. 2008 Plant Systems Biology comes of age. *Trends in Plant Sciences* 13: 165-170.